

# The controversy of AI in healthcare

Due to technological and medical development we have become better at curing and treating diseases and physical ailments. Furthermore, we are facing demographic changes which are causing a rise in the elder population of Denmark, resulting in an increasing number of patients requiring treatment. Additionally, healthcare professionals are at a global shortage, and proponents of AI in healthcare argue that AI will be able to streamline the healthcare system by e.g., performing faster and more accurate diagnosis, thereby creating more time for healthcare professionals.

On the contrary, critics of AI in healthcare argue that there are ethical issues with using algorithm-driven technology to treat patients, and that the risk of failure must be addressed. IBM created the AI Watson which was supposed to give oncologists treatment recommendations based upon a large amount of data; however, it failed due to lack of representative data which caused wrong recommendations for treatment and was therefore taken off the market.

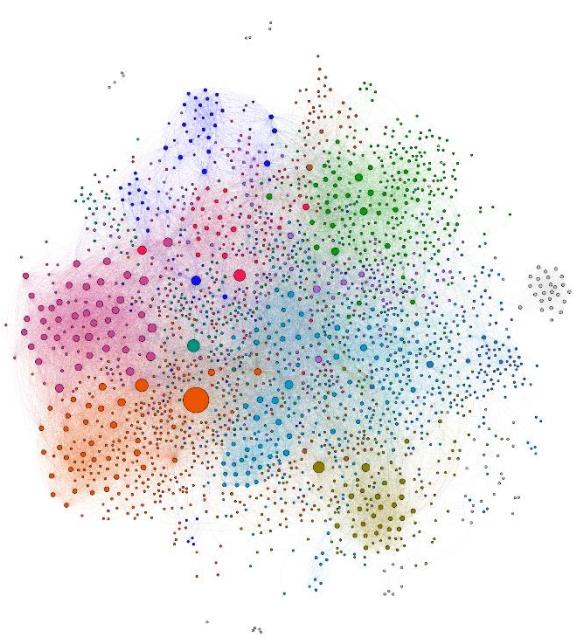
Thus, the main controversy is whether we can rely on artificially intelligent treatment and diagnostic tools?

AI in healthcare is a hot topic and a major controversy due to the diverse debate. With our controversy atlas, we thus ask: *What is AI actually about and who has a voice in the debate on Wikipedia?*

## 1. What is AI according to Wikipedia?

Initially, our intention was to explore what AI is according to Wikipedia. When building a network of articles, the network organized into clusters that reflected the subcategories within the AI category on Wikipedia. The corpus of articles on AI and its subcategories are structured neatly and divided

into clearly defined areas of science and knowledge, e.g., Robotics and Applied AI. At the same time, the density of the network and overlap of clusters (i.e. topics) indicates how AI as a field does not exist in isolation, because it draws and depends on several different fields that come together to form this cutting-edge technology. As such, at least in the Wikipedia realm, it is not possible to talk meaningfully about AI without talking about all its constituting parts.



*Figure 1. Network of articles in the AI category on Wikipedia. Articles as nodes, edges as hyperlinks. Node size indicates how often the article is linked to by other articles.*

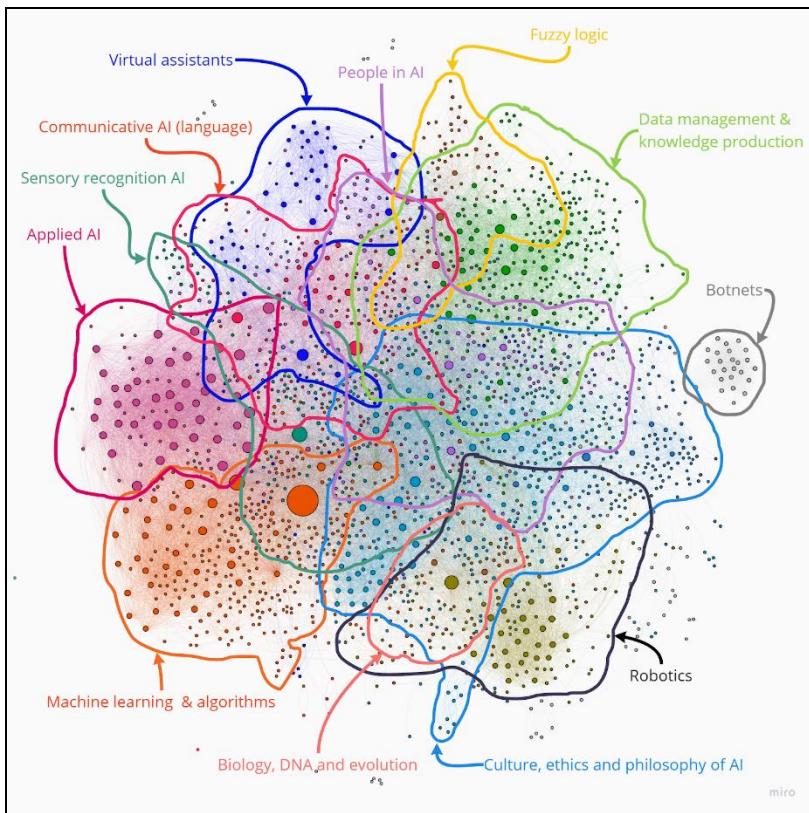


Figure 2. Clusters identified as topics in the article network.

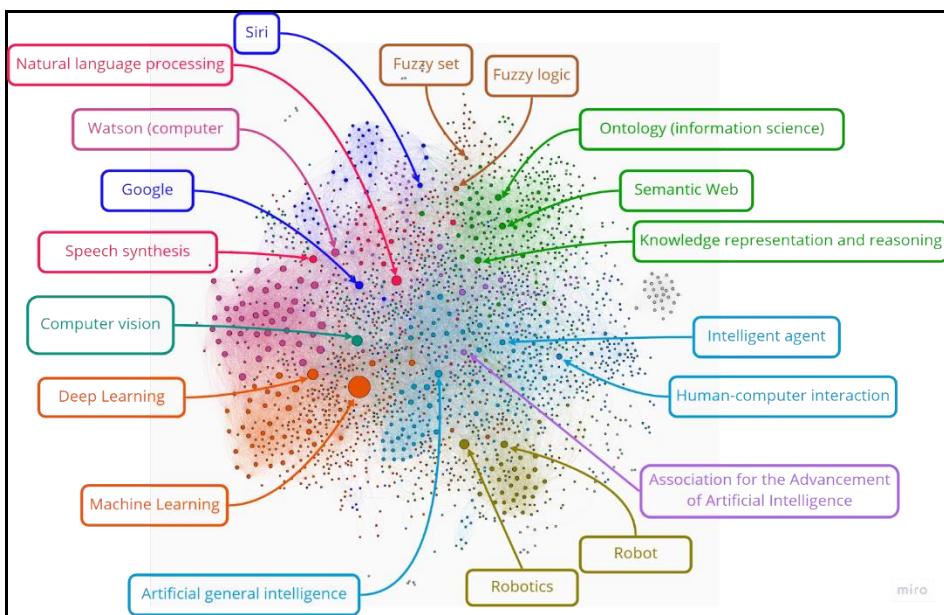
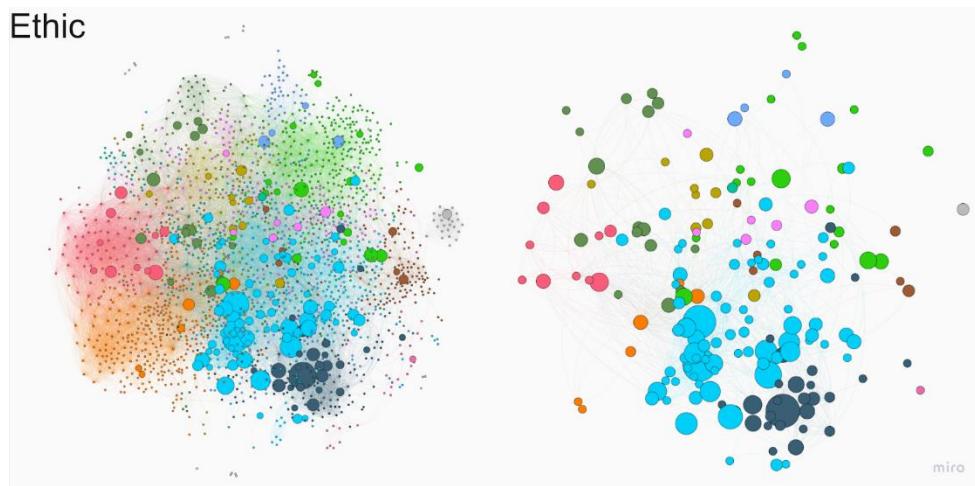


Figure 3. Central articles by in-degree in the article network.

## 2. Where do terms related to healthcare and controversial issues occur in the AI article network?

We hypothesize that healthcare is a debated topic within AI, within all the different areas, since Protocol 1 revealed there is not one cluster specifically related to healthcare itself and the articles with a bigger occurrence of words related to healthcare most likely also contain debates on ethics and privacy.

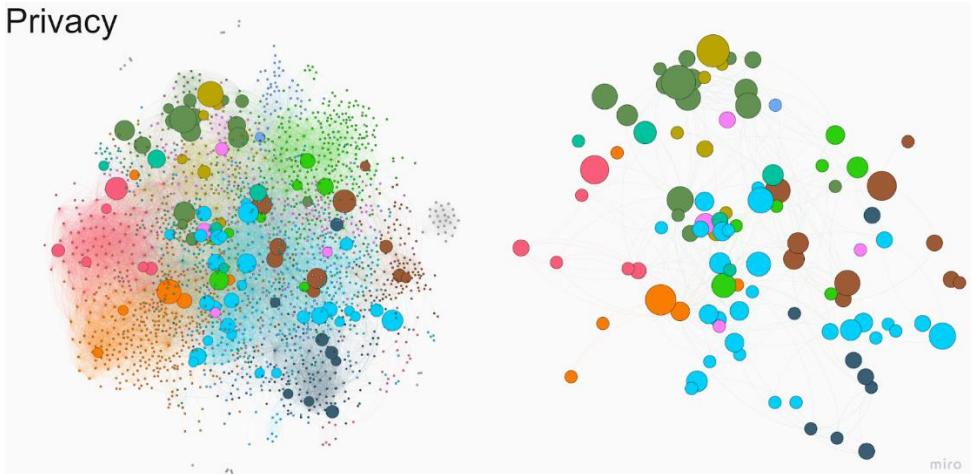
Protocol 2 explores whether the words *health*, *medic*, *ethic* and *privacy* occur in the same areas of our network from Protocol 1.



*Ethic* is most dominant in the subcategory Culture, Ethics and Philosophy of AI (CEP). *Ethic*'s dominance in CEP makes sense, as this subcategory mainly contains several articles specifically on ethics. CEP is located centrally in the network, which shows how ethics is referred to and debated within other subcategories of AI.

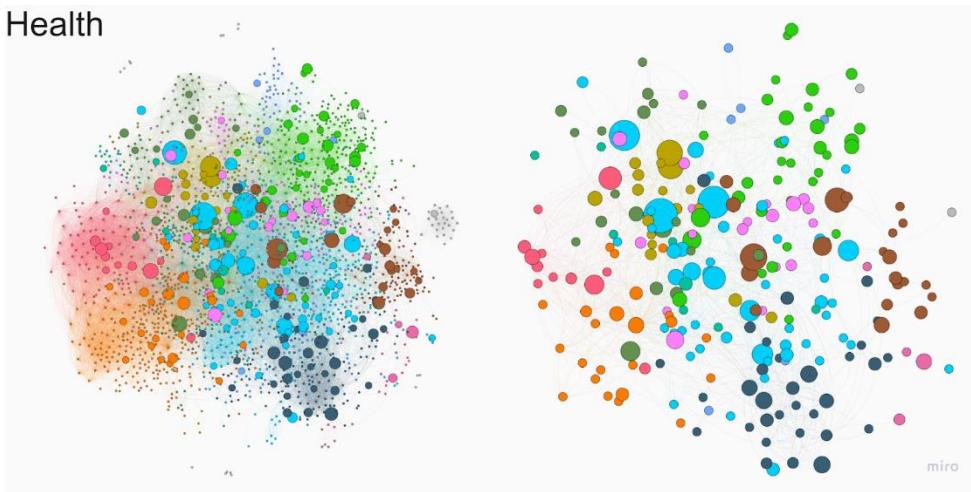
Analyzing *privacy*, the visualization differs; one may argue ethics and privacy are closely related value-driven words, however, this is not the case within AI on Wikipedia.

## Privacy



A subcategory which differs between the two visualizations is CEP; *ethic* occurs more often than *privacy*. The sparse occurrence of *privacy* is not located in the same area as *ethic* is, indicating that *privacy*, when discussing *ethic*, is not a typical concern within the CEP area.

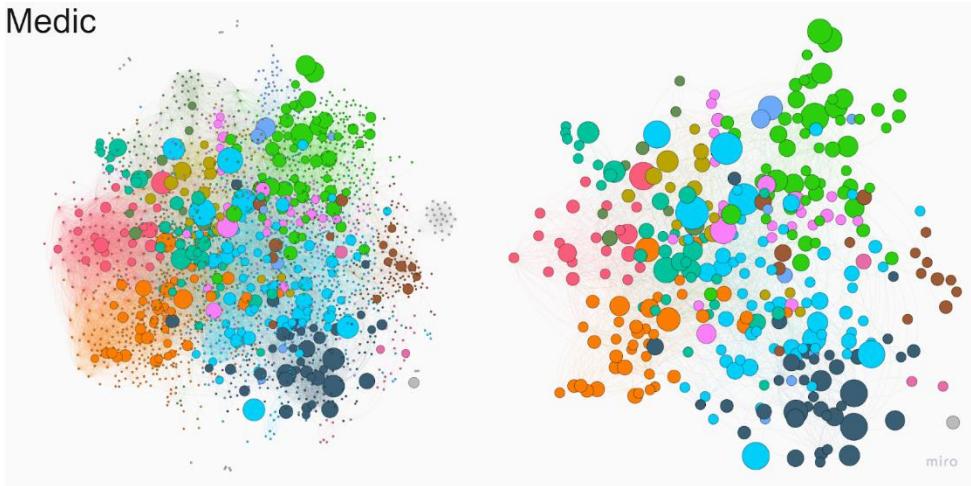
## Health



*Health* is debated within all areas of AI. The subcategory where *health* is debated the most is within CEP.

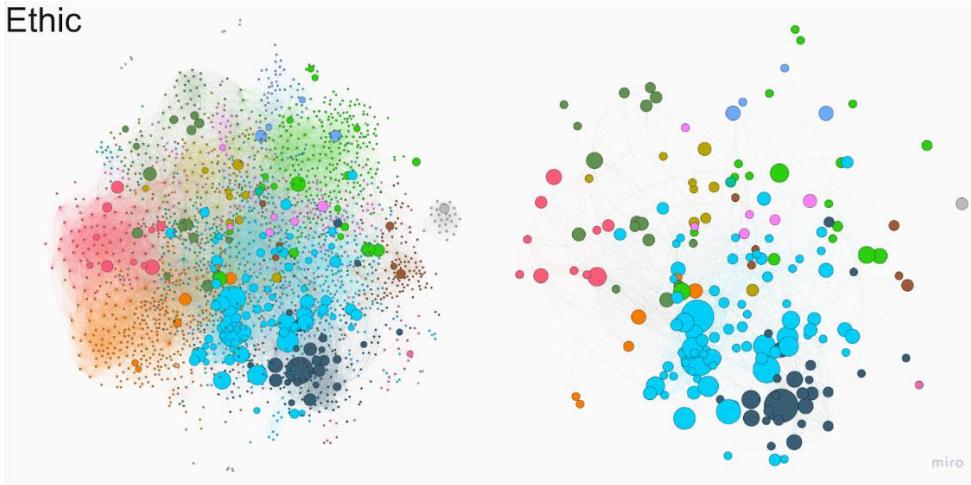
In comparison to *health*, *medic* occurs in more articles in all the different clusters.

Medic

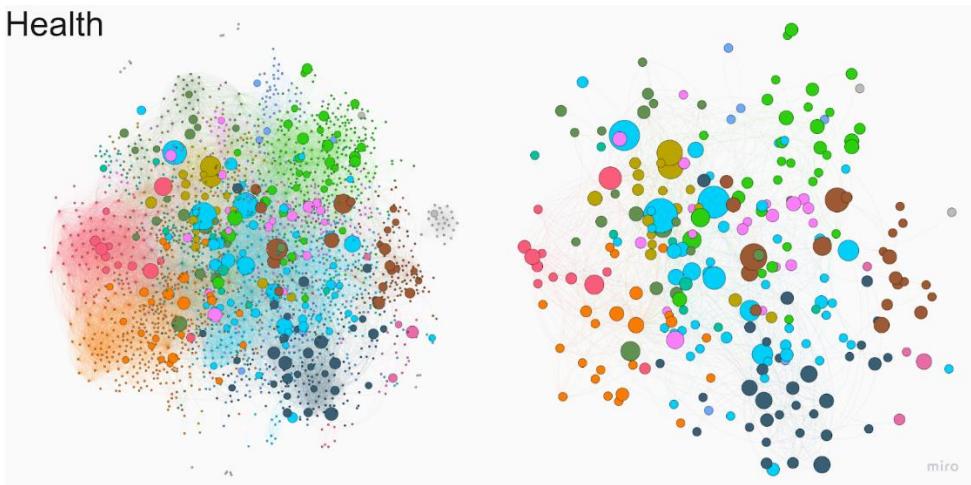


In CEP, the articles with the biggest occurrences of *health* and *medic* are located in the same area. *Medic* is more evenly distributed throughout the entire network: on Wikipedia, *medic* (and versions hereof) might be a more typical choice of word in the AI in healthcare debates.

Ethic



Health



Comparing *health* and *ethic*, they do not occur in the same areas. Generally, *ethic* and *privacy* occur rarely, and there's hereby not a lot of coherence with the articles containing healthcare related words. There is only a small overlap of occurrences between the two sets of words *ethic* and *privacy*, and *health* and *medic*; this showcases how the two sets of words are rarely debated within the same articles.

### 3. Do spikes in edit activity reflect external concurrent events pertaining to AI?

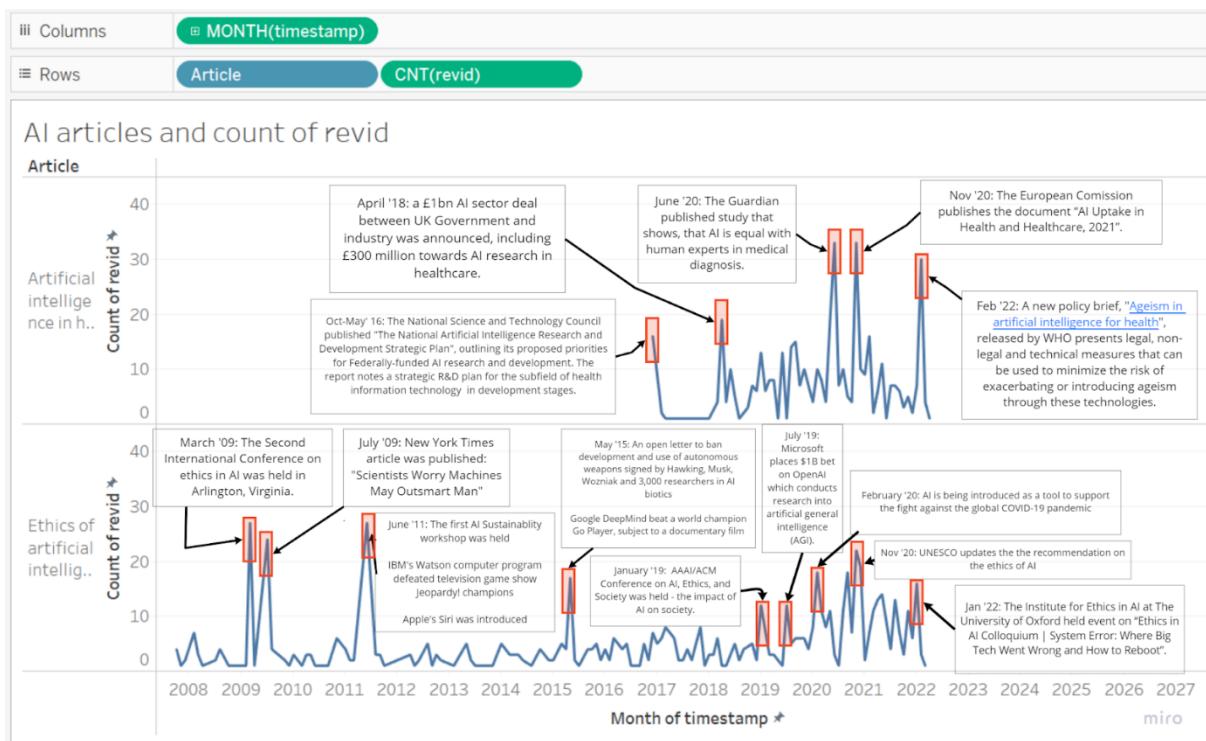


Figure 4. Timeline of edit activity on two articles in the AI category.

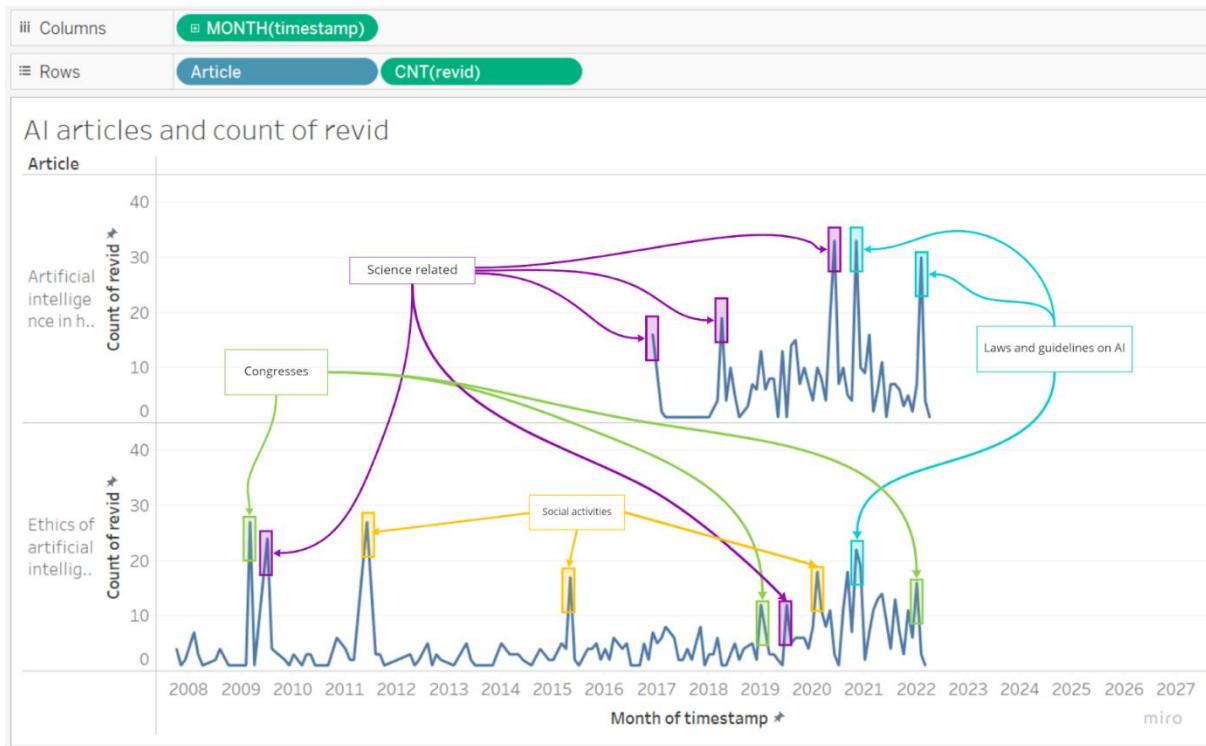


Figure 5. Types of external events happening concurrently with edit activity spikes.

AI in healthcare is experiencing great development; multiple congresses have been held in relation to AI, and many academic papers have been published on the matter. Furthermore, many legal and financial measures have been implemented since the two wikipedia articles have been published. Based on these findings, we hypothesize the spikes reflect the external events pertaining to AI happening in the same period.

However, exploring the most active editors on spikes and the content of their edits revealed that the edits generally don't reflect what is happening in the 'real world' - with the single exception of when *AI in healthcare* was created. At this point, the creator of the article did address the strategy report published in the same month.

The spikes on both articles do not seem to correlate with each other which was a surprise, since we had a preconceived idea that ethics and health as topics would go hand in hand.

An interesting aspect when comparing the two articles was that *AI in healthcare* was not created until 2016, whereas *Ethics of AI* was created in 2008. A reason could be that, unlike a more universal topic such as ethics which we saw had its own ‘hub’ that other articles link to, healthcare in AI may have been addressed more sporadically across the entire network up until December 2016 - at which point a user decided to create a dedicated page for the topic.

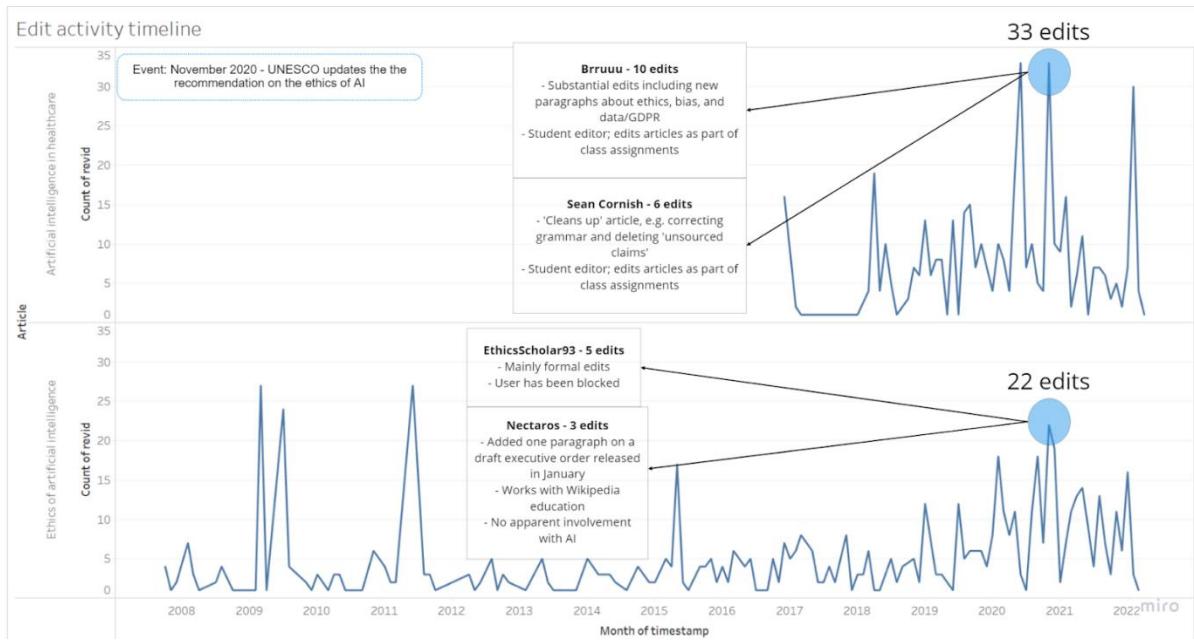
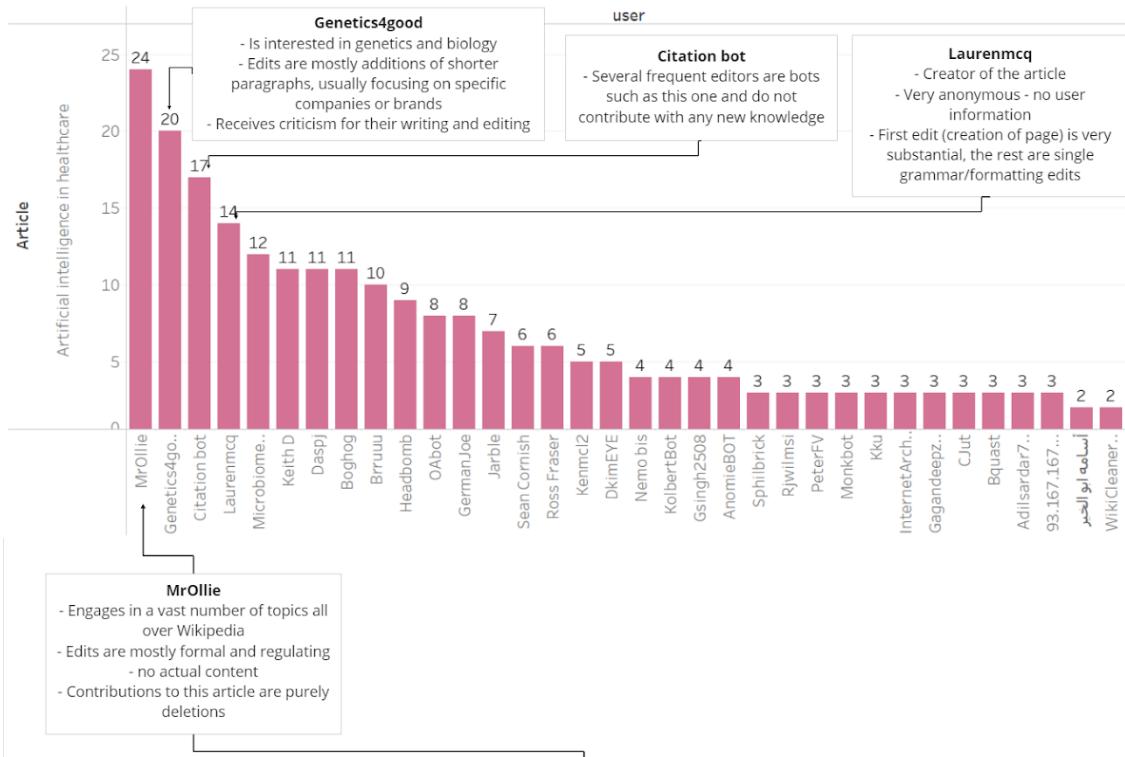


Figure 6. Example of what happened during an edit activity spike, here in November 2020.

## 4. Who are the most dominant editors, and what characterizes their editing behavior?

Editor activity - all time



Editor activity - all time

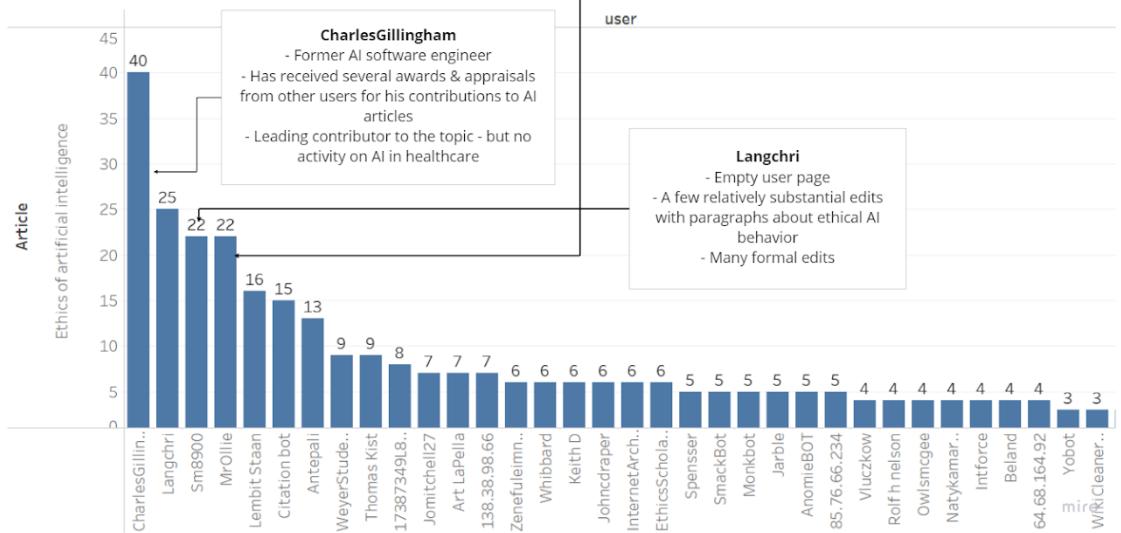


Figure 7. Bar chart of the most active editors of all time on two articles in the AI category.

AI as a topic is closely interconnected with a vast number of different areas of science and knowledge, and various issues central to the controversies around AI are not equally represented across these areas. These qualities are not merely inherent to the AI debate; they're a result of how Wikipedia as a crowdsourcing encyclopedia selects and structures the knowledge visible to its readers. Since anyone can contribute to articles, the corpus of knowledge is first and foremost the result of the work of a large number of "strangers", i.e., editors. Investigating who the dominant editors are can provide valuable insights about who sets the tone in the AI sphere on Wikipedia.

Qualitative inquiry into the user- and talk pages of the most active users revealed that the behavior and background of editors are very diverse. For instance, the user MrOllie engages with many different topics on Wikipedia and is one of the most frequent editors on the AI articles. However, he is only contributing with formal and regulating edits, not adding new knowledge to the articles - in fact he mostly removes content that, according to him, violates Wikipedia's rules. On the other hand, CharlesGillingham, who is also one of the most dominant editors on *Ethics of AI* could, contrary to MrOllie, be considered an expert due to his background as a former AI software engineer and his great know-how within the field of AI.

## Protocol 1

What is AI according to Wikipedia? Visualizing a network of articles related by hyperlinks in their text

### Artificial intelligence on Wikipedia

40 subcategories and 400 pages in the AI category

Call API to pages from category at level 1

Colab

List of 2714 Artificial intelligence page titles

CSV

Call API to build network of articles connected by hyperlinks in their text

Colab

Network of 2714 artificial intelligence pages

Articles as nodes, edges as hyperlinks

GEXF

**HARVEST**

**VISUALIZE**

MAP

Slides

Network viz

PNG

**ANNOTATE**

Clusters from force layout and Modularity  
*Themes central to AI*

Central nodes by in-degree  
*Identify key articles by articles linking to it*

Apply force-directed layout with ForceAtlas2

Label size: ranking, in-degree (0,5-6)

Colour nodes by modularity in partitioning

Delete node 'Artificial intelligence'

Size nodes by in-degree (1-150)

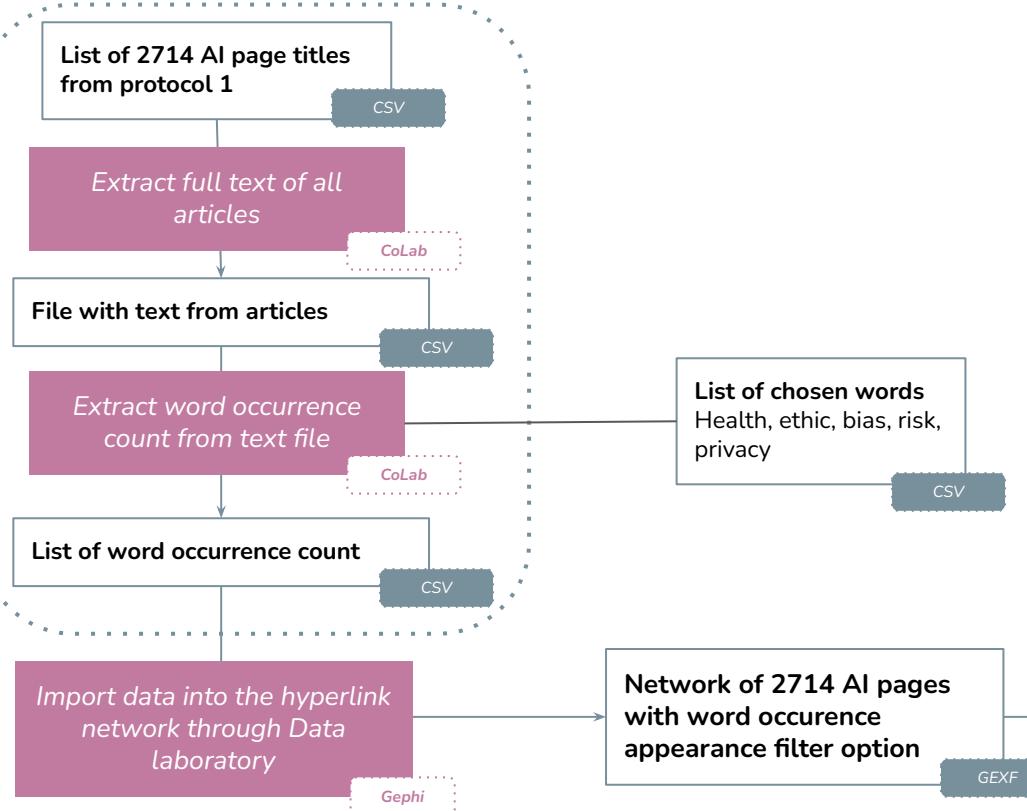
Apply in-degree range filter with min 5

Gephi

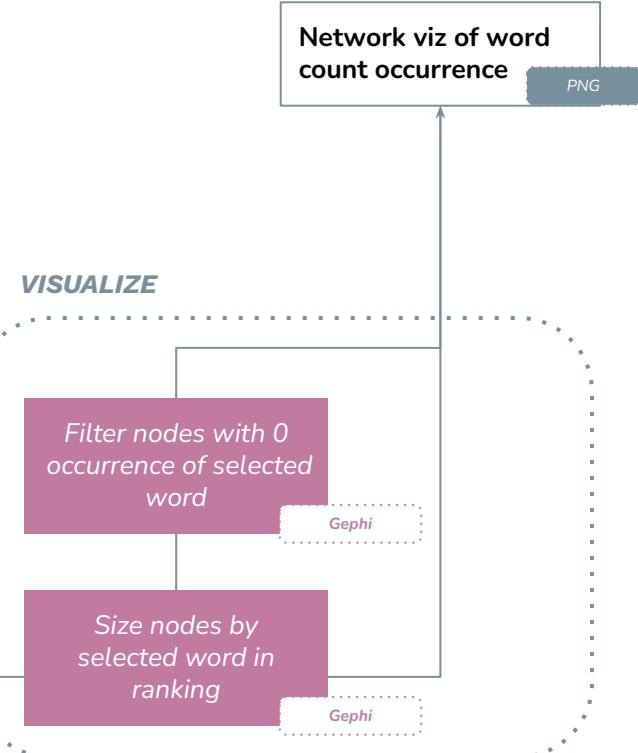
## Protocol 2

Where do terms related to healthcare and controversial issues occur in the AI article network? Visualizing how keywords occur in the network of AI articles

### HARVEST

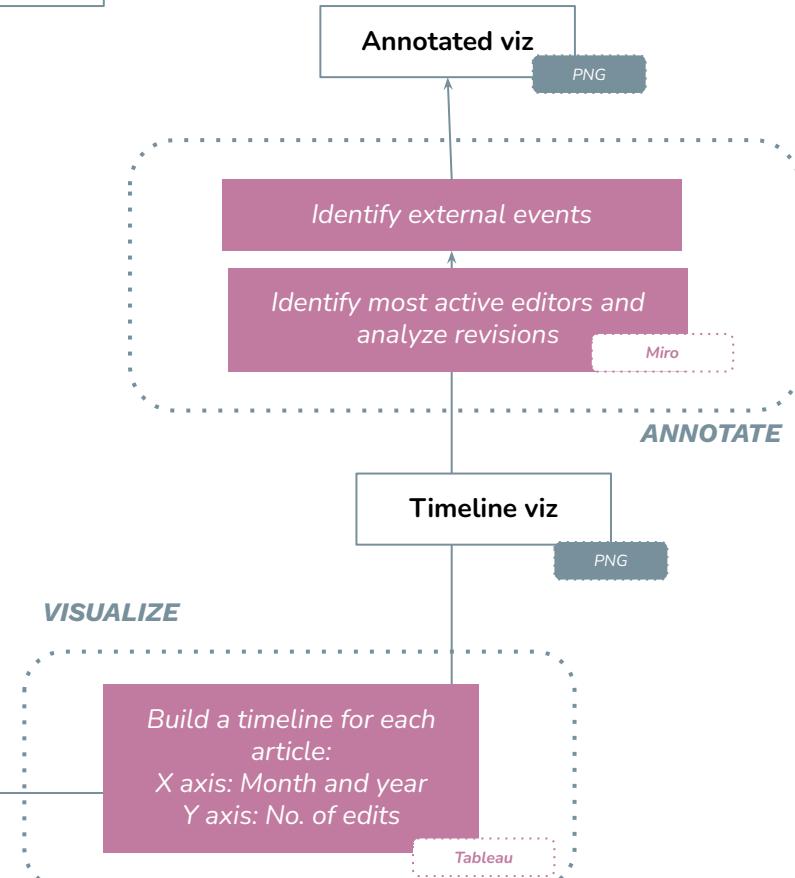
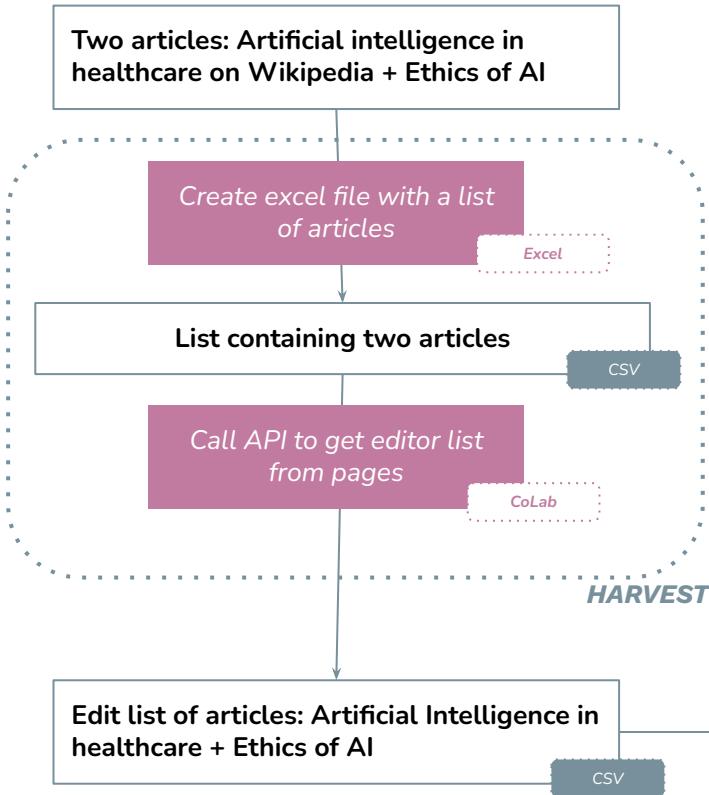


### VISUALIZE



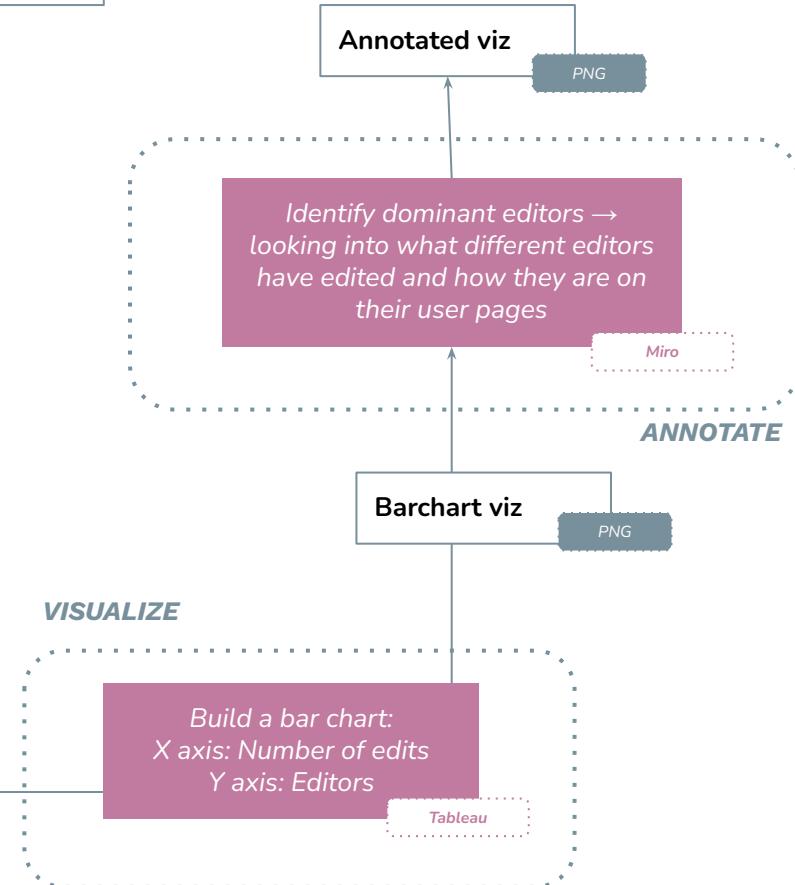
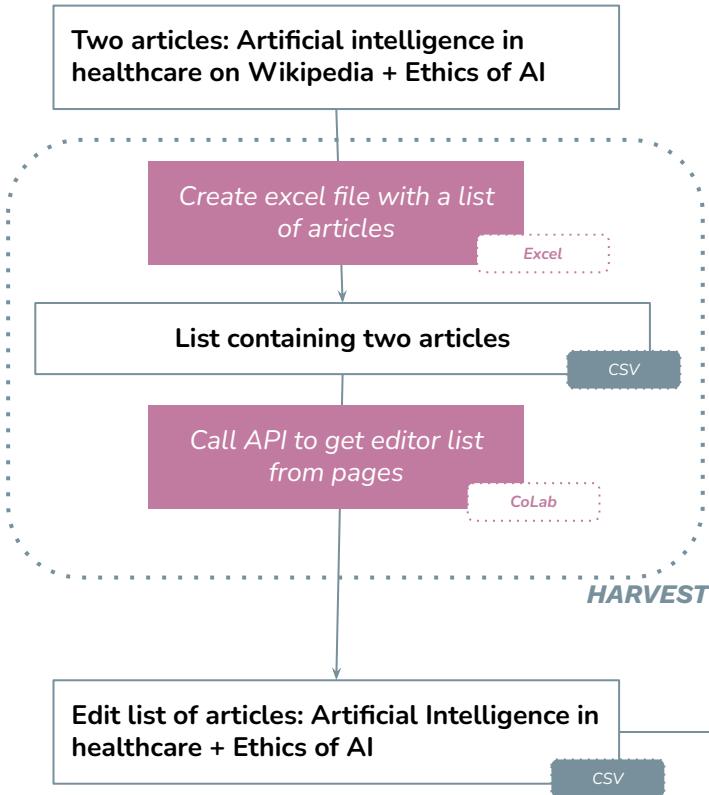
### Protocol 3

Do spikes in edit activity reflect external concurrent events pertaining to AI? Visualizing a timeline of edit activity on two Wikipedia articles

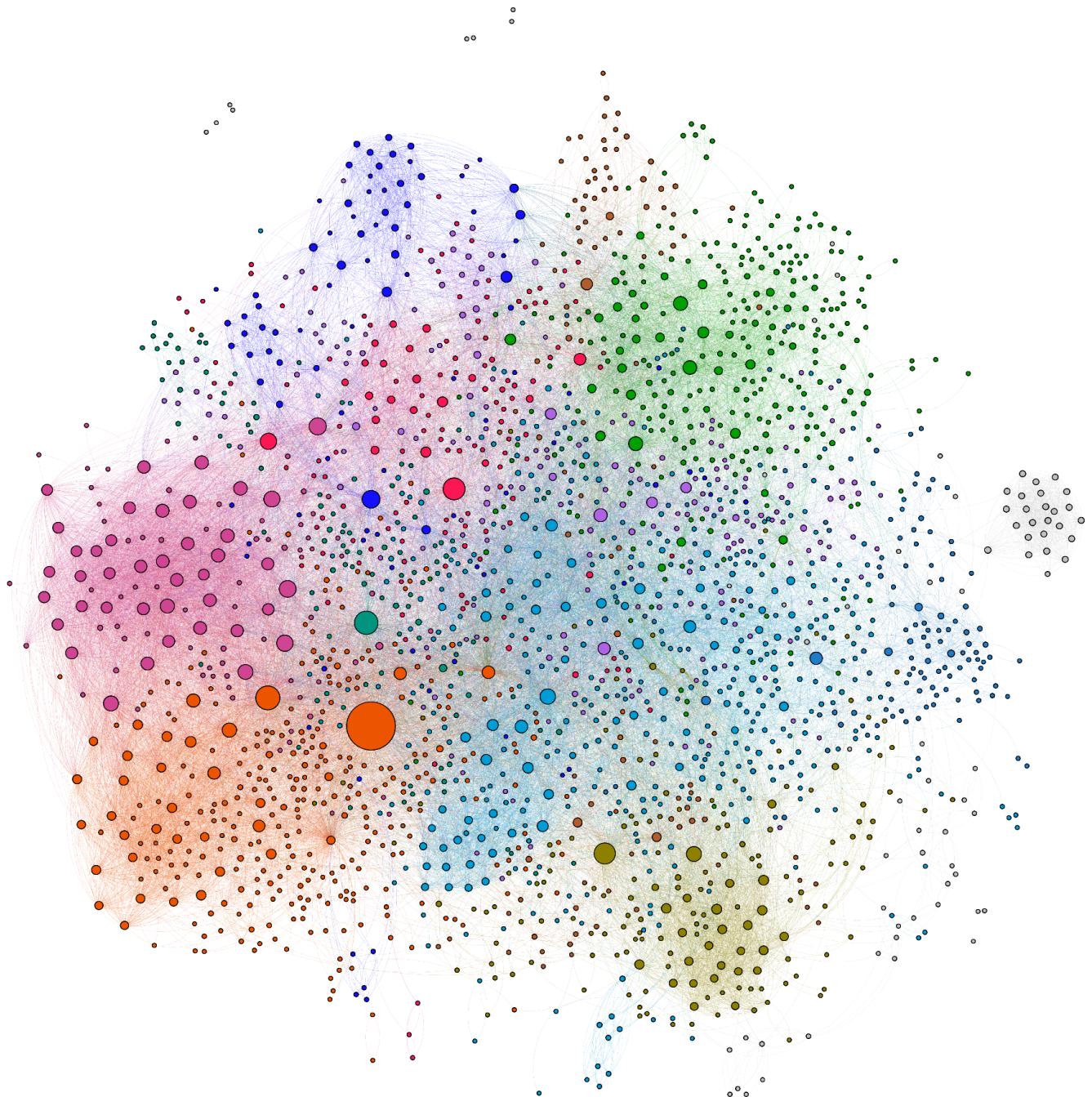


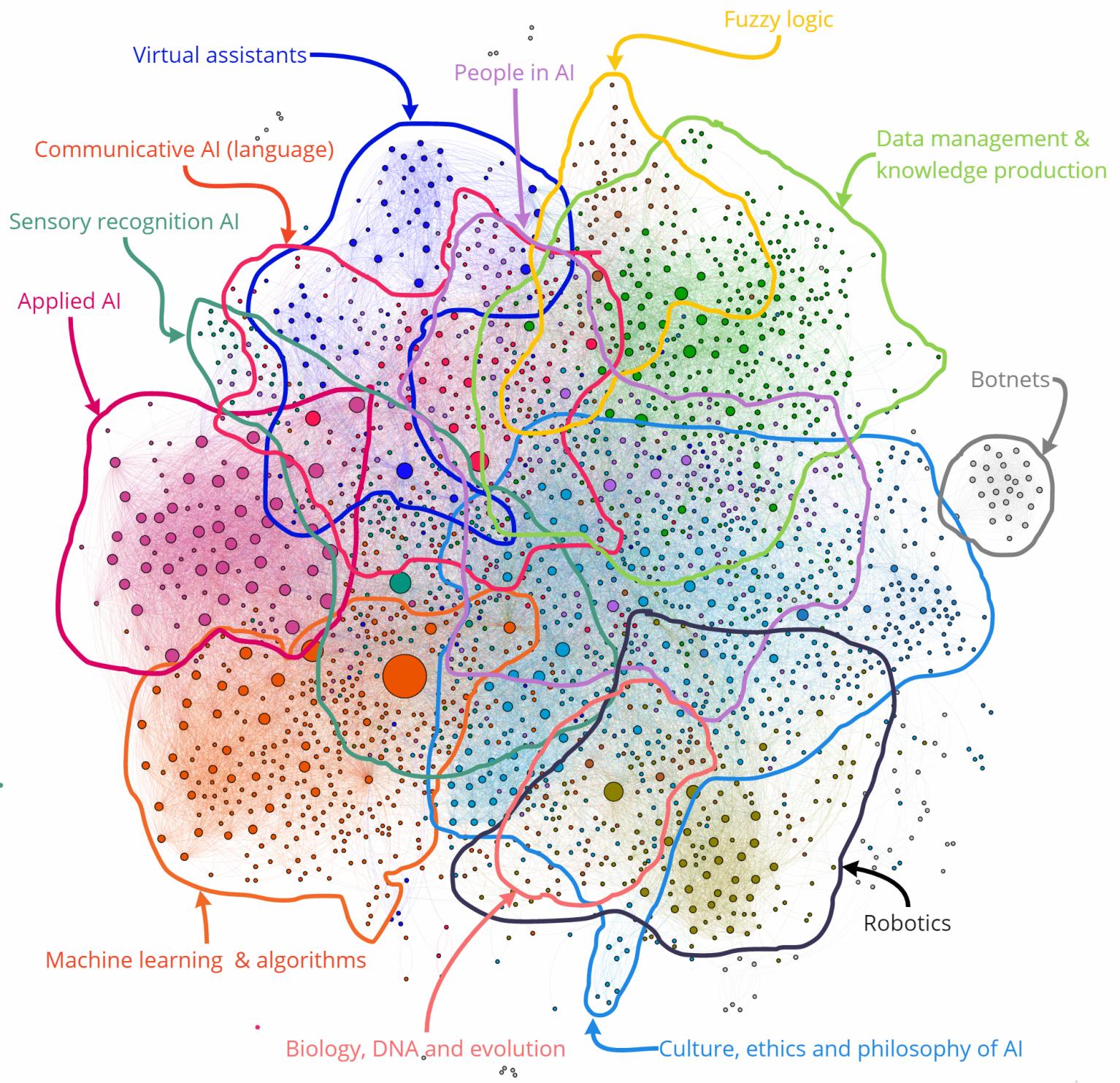
## Protocol 4

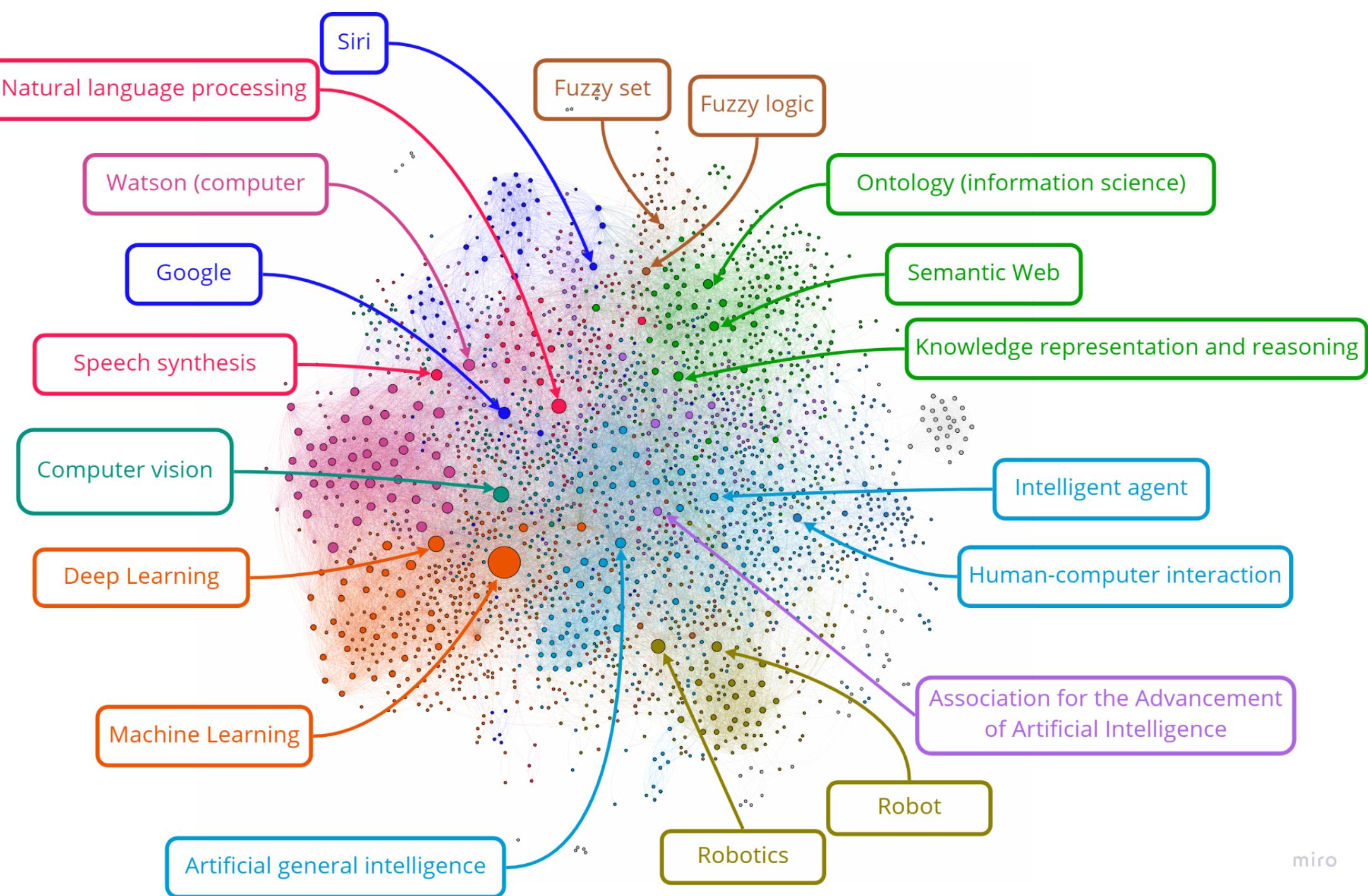
Who are the most dominant editors, and what characterizes their editing behavior?  
Visualizing editor activity on two Wikipedia articles



# Protocol 1

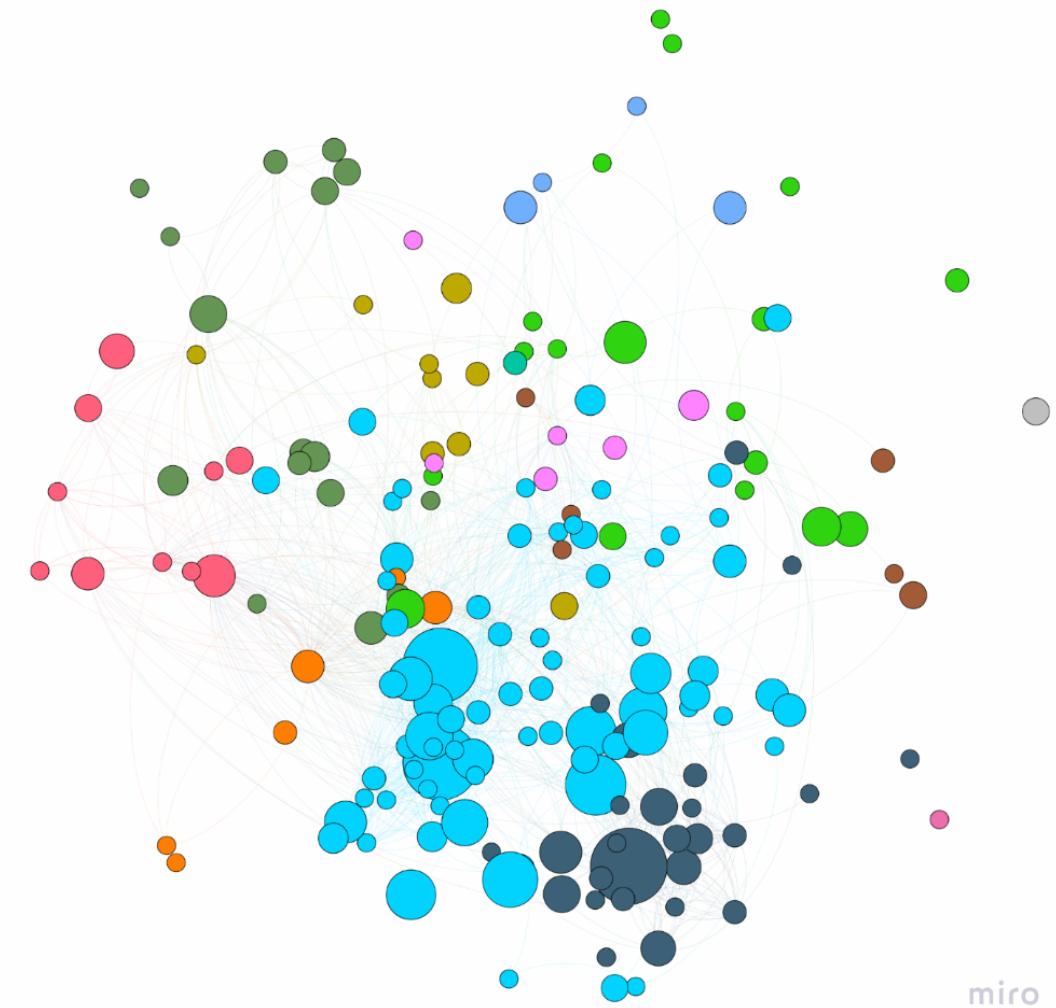






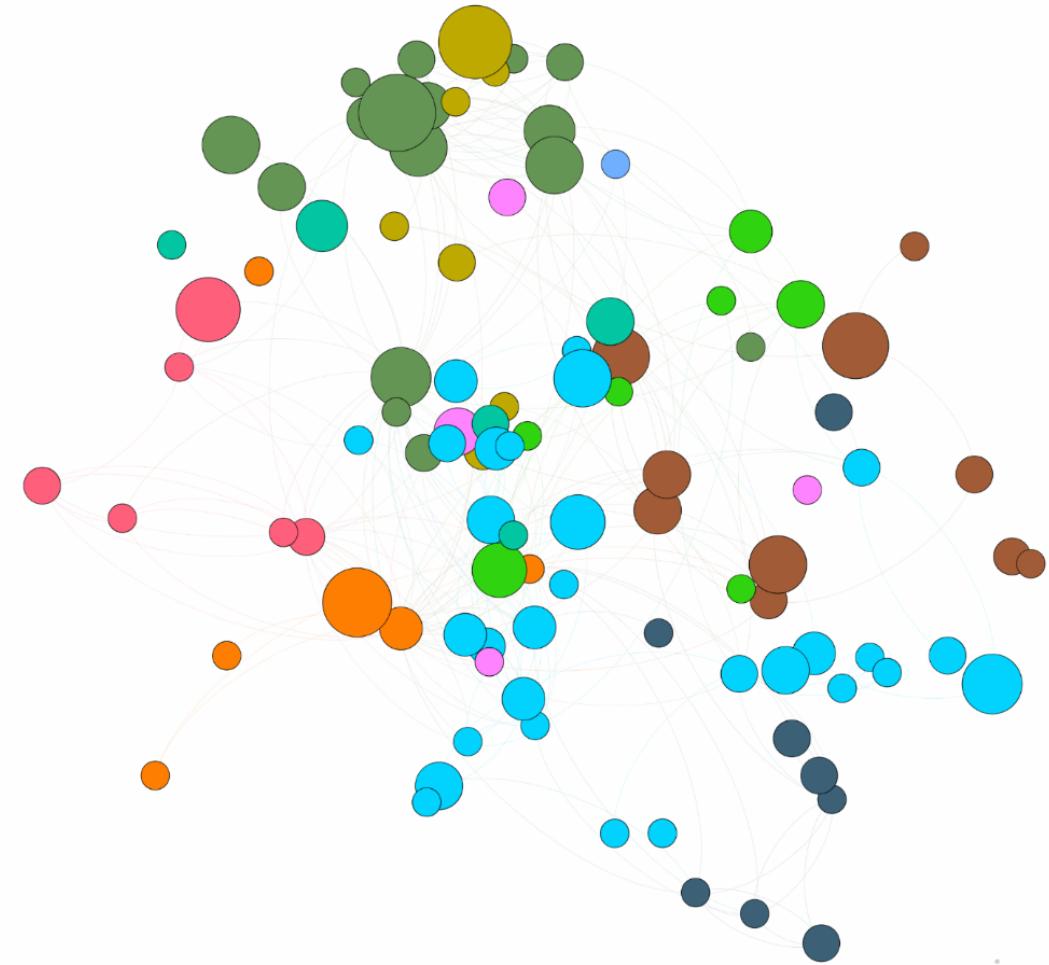
## Protocol 2

Ethic



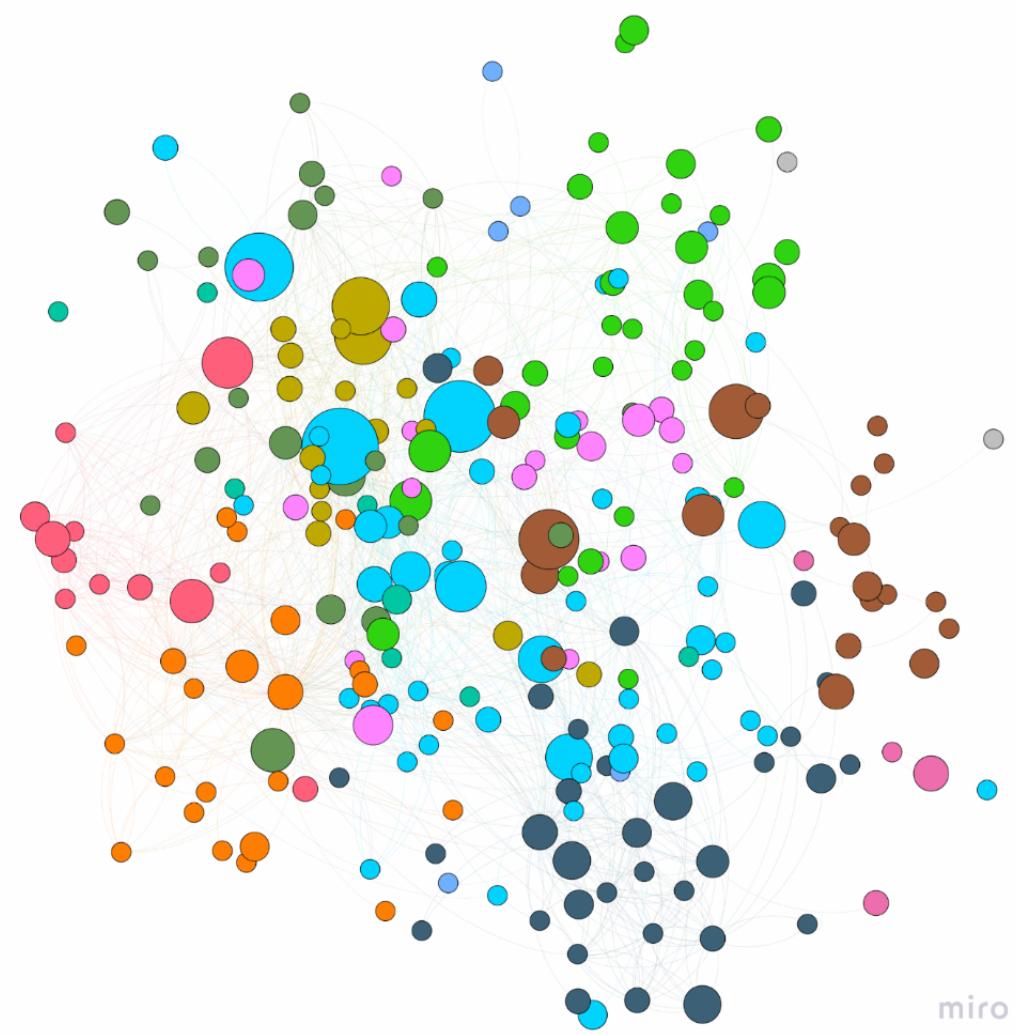
miro

# Privacy

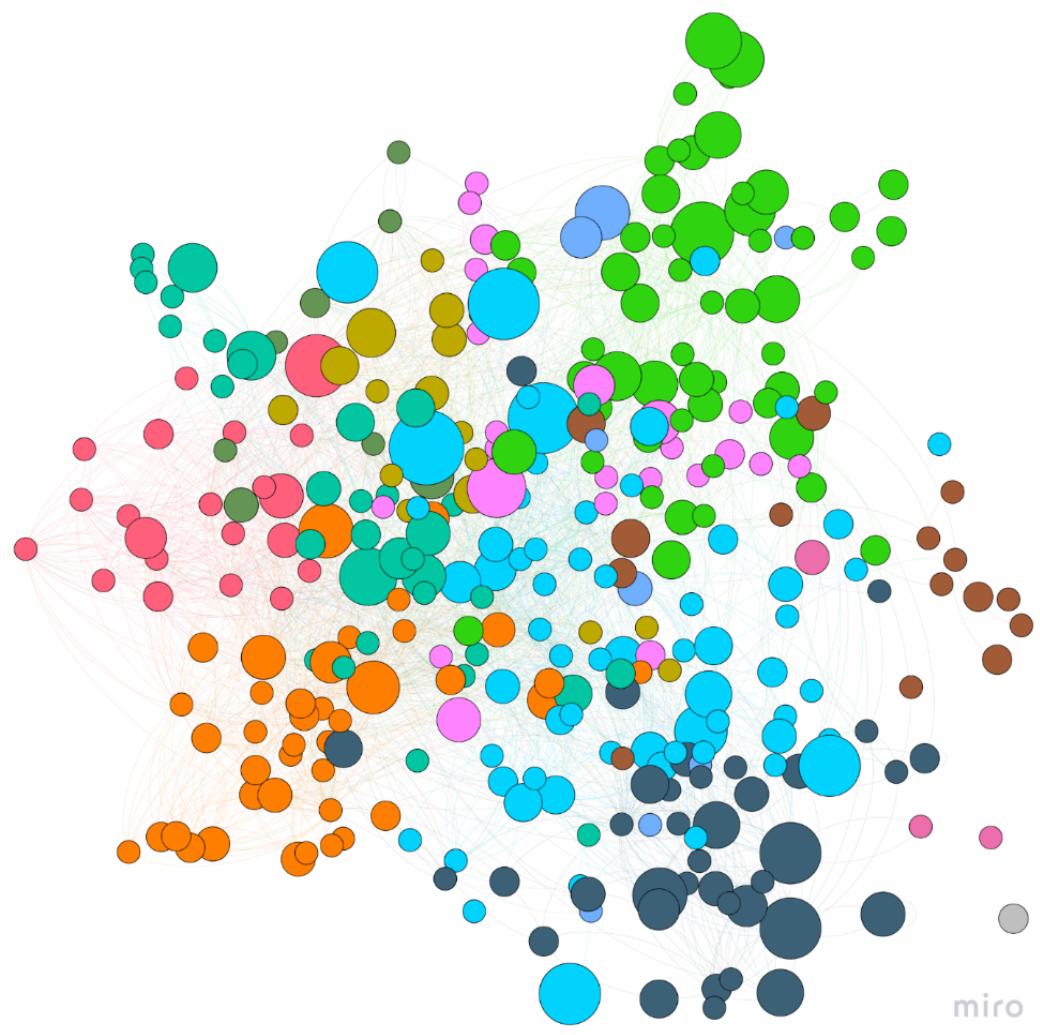


miro

# Health



# Medic



miro

# Protocol 3

iii Columns

MONTH(timestamp)

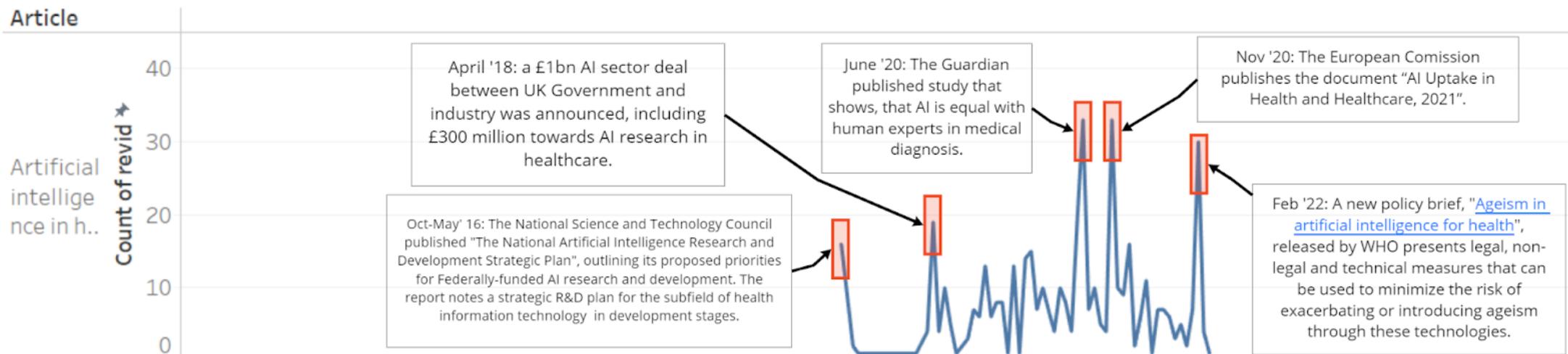
Rows

Article

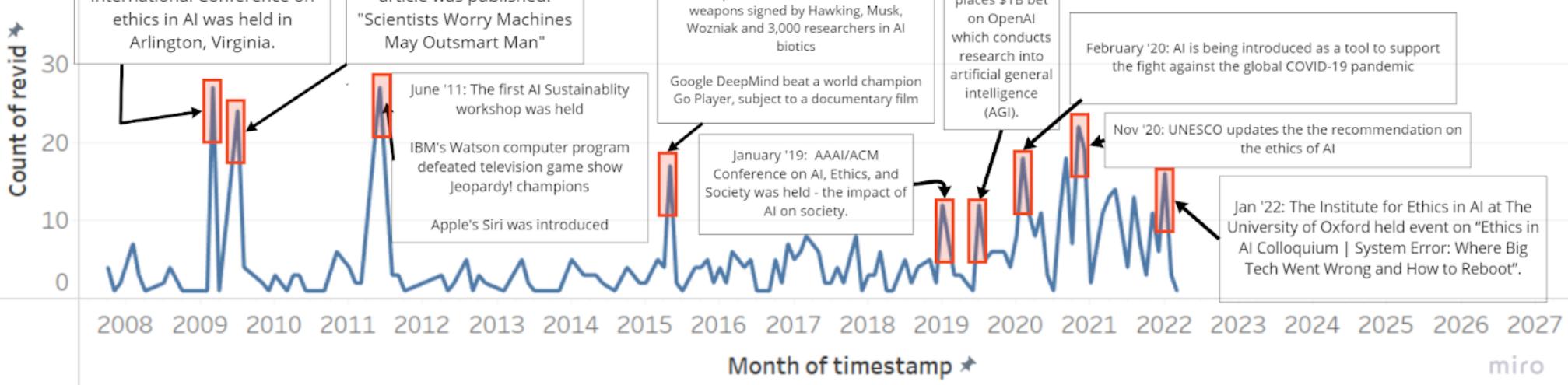
CNT(revid)

## AI articles and count of revид

Article



Ethics of  
artificial  
intellig...



Month of timestamp

miro

iii Columns

MONTH(timestamp)

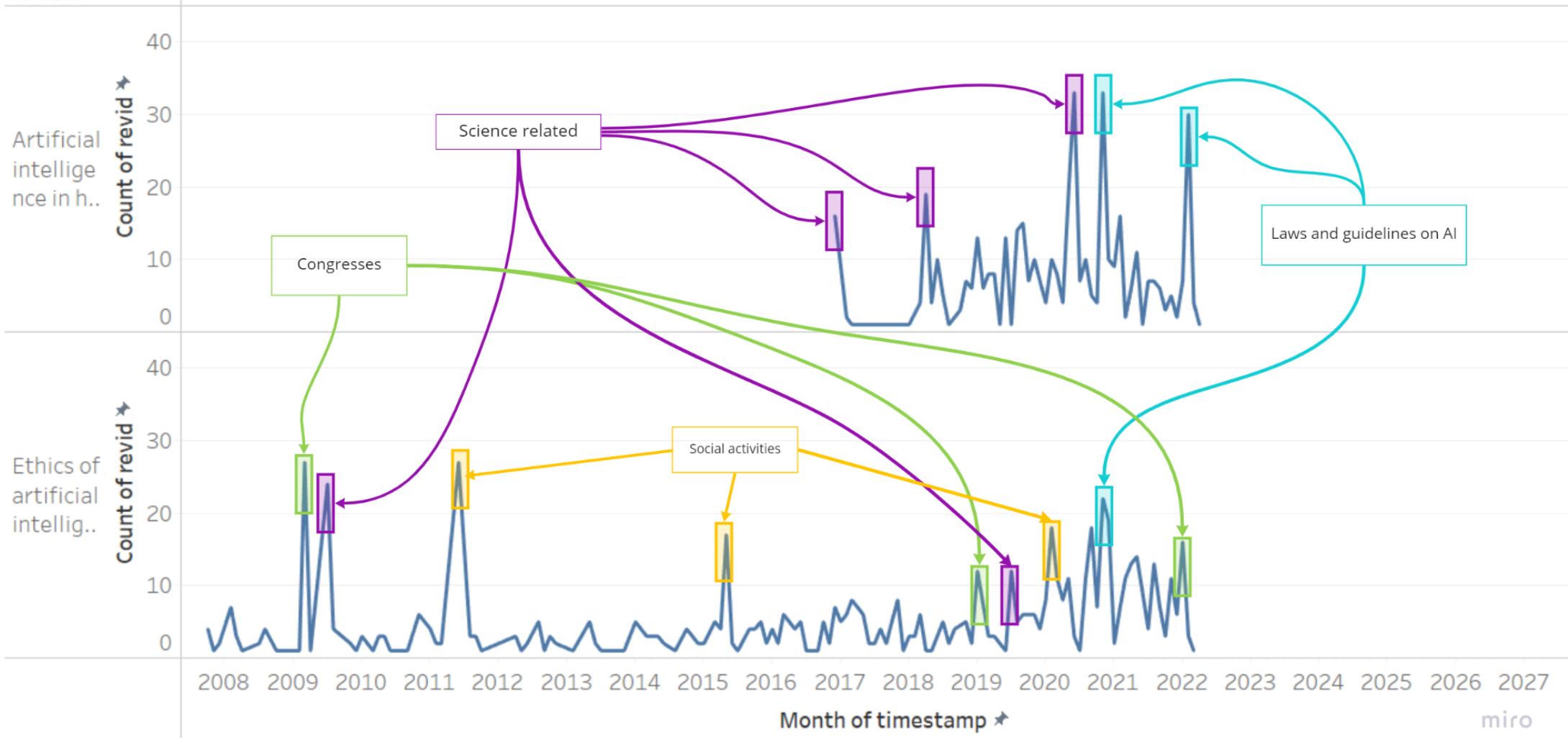
Rows

Article

CNT(revid)

## AI articles and count of revid

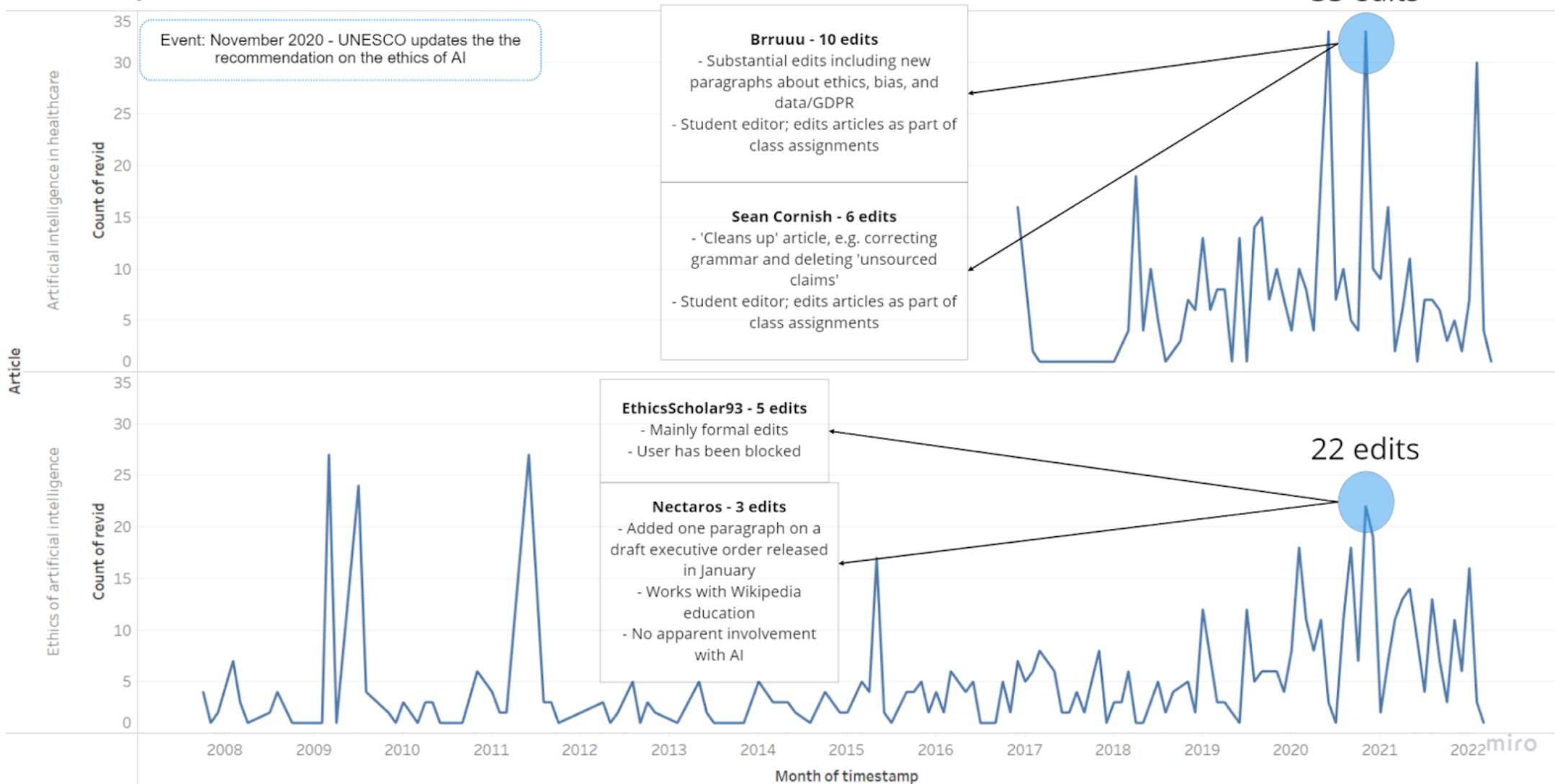
Article



miro

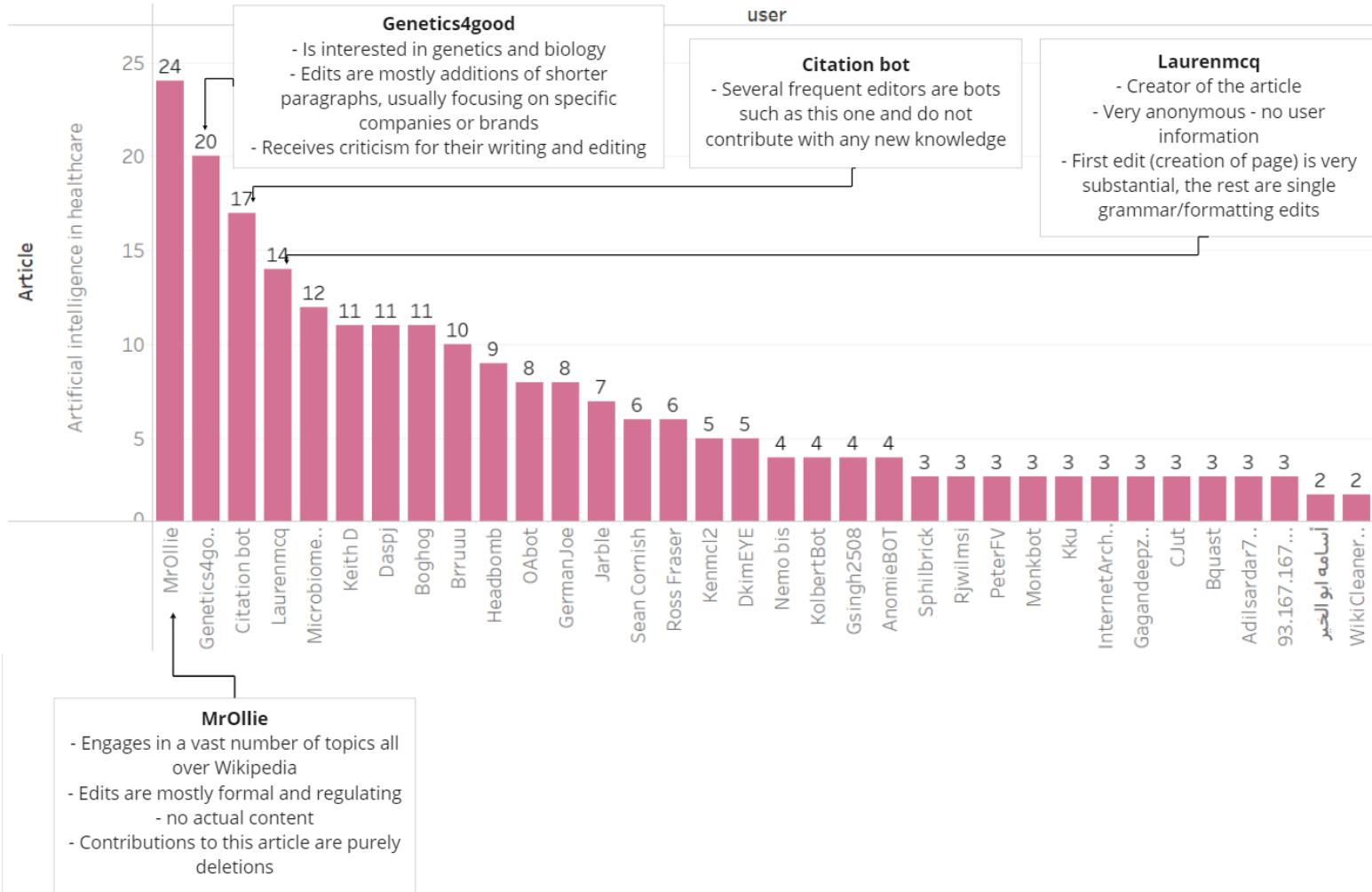
## Edit activity timeline

33 edits



# Protocol 4

## Editor activity - all time



## Editor activity - all time

