

Build Your Own Timeline Algorithm

Davide Eynard @ SFSCON, 7 Nov 2025

 @mala@fosstodon.org

 [mozilla-ai/byota](https://github.com/mozilla-ai/byota)

BYOTA is... A (pacifist) call for arms



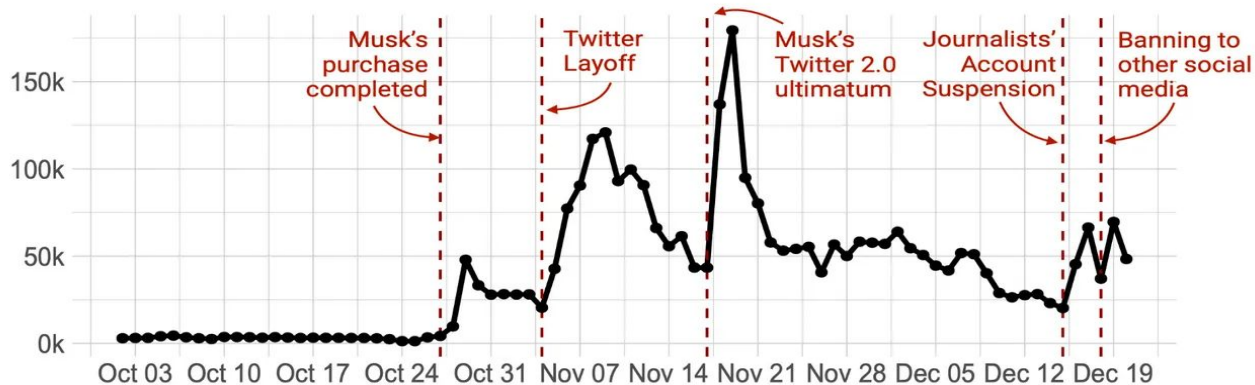
I need you, you, you

How it started

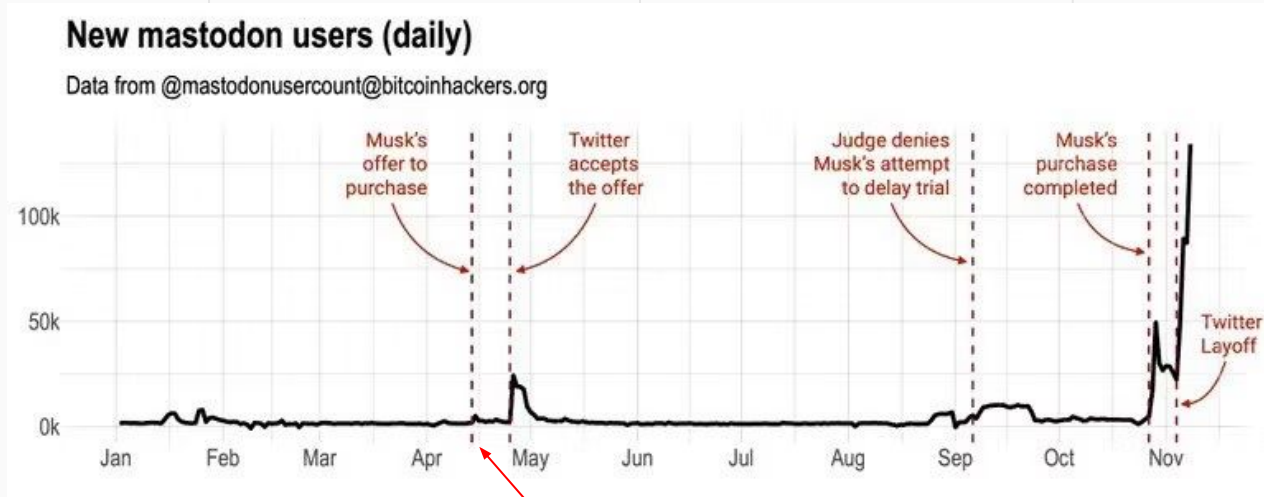
#TwitterMigration

New mastodon accounts (daily)

Data from @mastodonusercount@bitcoinhackers.org



#MyMigration



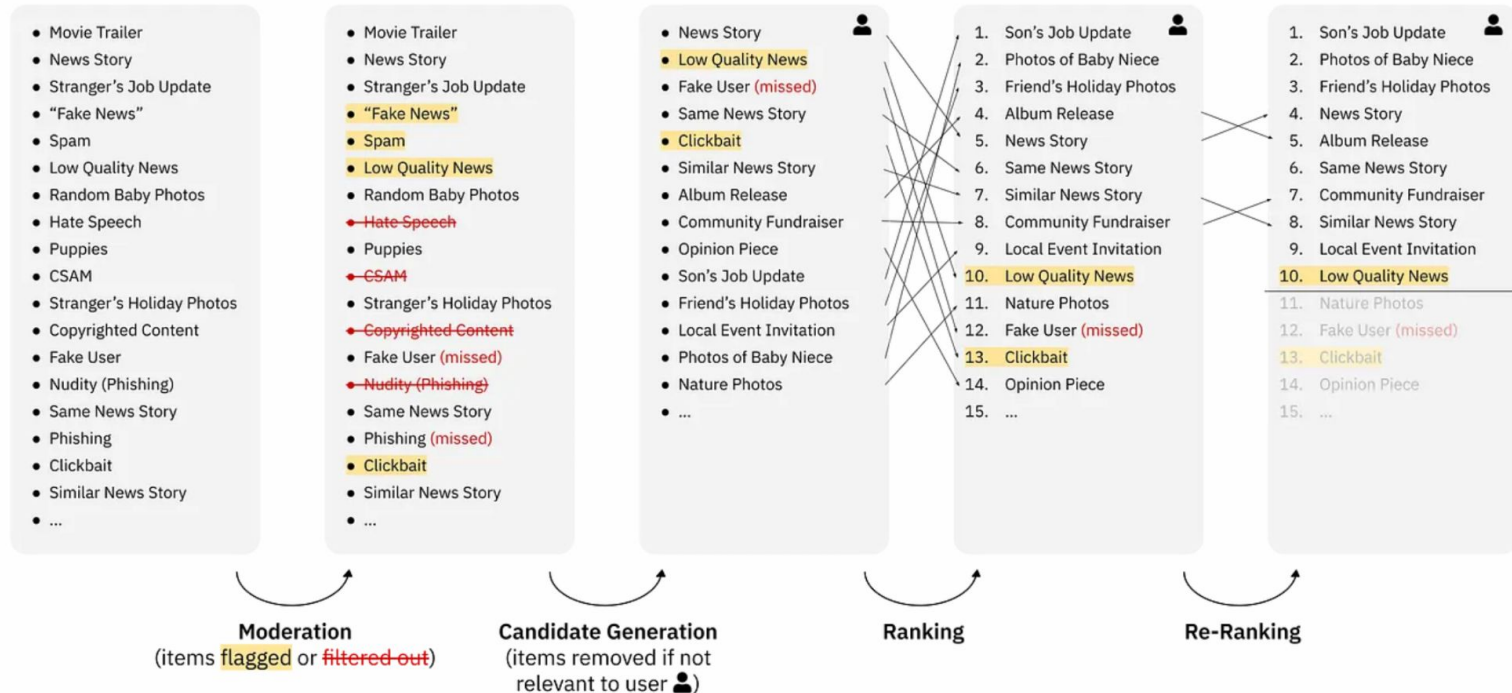
Timeline Algorithms

Timeline Algorithms

Define *whether* and *how* posts appear on your timeline

Algorithm	Remove?	How
Revchron	No	Sort by datetime in ascending order
Spam / Bots	Yes	Remove post if spam / bot
Moderation	Yes/No	Remove / Downrank if post is abusive
Recommender	No	Uprank if post is more likely to receive engagement

Timeline Algorithms



Luke Thorburn et Al.: How Platform Recommenders Work (2022)

One the shoulders of giants



Ex-Searchtodon

Formerly: Private Timeline Search for Mastodon

Jan Lehnardt



Fedi-Feed

Philip Kreißel

Tootfinder

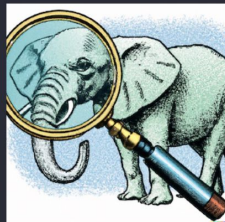
Opt-in global Mastodon full text search. [Join the index!](#)

Search

Full text search on Mastodon

Imagine searching any post on Mastodon. This is now possible - at least for the posts of users who opt in.

Tootfinder indexes all public posts of consenting users and makes them searchable for 3 months. If you want to be part of it, [join the index](#).



Matthias Bürcher

Mastodon Digest

Sourced from your timeline over the past 12 hours

Here are some popular posts you may have missed:

Matt Hodges



fediview

Plus 

Adam Hill

Problems of Timeline Algorithms

Choose one or more of:

- Biased
- Not transparent
- Complete lack of control
- Not privacy preserving
- Huge computational cost
- Require deep ML knowledge
- Cannot be run locally

(Also true for search & recommendations!)

Breaking the Assumptions

Choose one or more of:

- Biased
- Not transparent
- Complete lack of control
- Not privacy preserving
- Huge computational cost
- Require deep ML knowledge
- Cannot be run locally

(Also true for search & recommendations!)

- Open, substitutable models
- Clear, simple methods
- Make code available / easy to run
- Don't share personal data
- Use lighter models
- Out-of-the-box embeddings
- Build it so it runs locally :-)

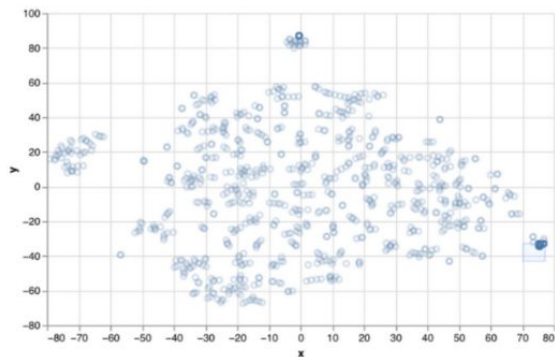
... All the above work for search & recos too!

BYOTA

BYOTA Architecture



Getting data from my Mastodon account...



id	text
int64	object
	unique: 4
	nulls: 0
1.1e17	1.1e17
182	113420501365488432
	Assigning your copyright to the FSF helps defend the GPL and keep software free. Thanks to Kristofer Hjelmtorp, Christopher Floess, and Seyed Sajad Kahani for assigning their copyright to the FSF! #GCC #Emacs More: https://u.fsf.org/3ht #CopyrightAssignments

3.57 iter/s - Total time 11.2s 40 / 40

Enter a post id or some free-form text to find the most similar posts:

Free software, digital rights, privacy

local_id	text
int32	object
	unique: 3
	nulls: 0
102	787
422	Software development is political. Software will shape t...
745	Did you know that an FSF associate membership funds t...
413	Did you know that an FSF associate membership funds t...
787	#ThankGNU! to Adarsh Melethil, David Bruce, Deborah Ja...
102	#ThankGNU! to Adarsh Melethil, David Bruce, Deborah Ja...

5 rows, 2 columns

Page 1 of 1 Download





llamafire

- Multi-platform, single-file LLMs-language models
- sentence-transformers/all-MiniLM-L6-v2
(code/papers, training data)
- 100% local
- Optimized to run on slower hardware
- OSS, will run (almost) forever

We use it in BYOTA to calculate status embeddings



- Reactive notebook for Python
- Easily shareable, runs as an application
- Runs in your browser as a WASM-powered standalone HTML file! 🧠 😄

Configuration

Mastodon Credentials

Login: Password:

Timelines

☒ Home ☐ Local ☐ Public
☐ Hashtag ☐ List

Embeddings

Embedding server URL

Caching

☐ Run in offline mode

Clear Submit

1. auth_form

```
1 # check for anything invalid in the form
2 no_stop(invalid_form(auth_form),
3         no_md("**Submit the form to continue.**"))
4
5 # login (and break if that does not work)
6 mastodon_client = byota_mastodon.login(client_cred_filename,
7                                       user_cred_filename,
8                                       auth_form.value.get("login"),
9                                       auth_form.value.get("pw"))
10
11 no_stop(mastodon_client is None,
12         no_md("**Authentication error.**"))
13
14 # instantiate an embedding service (and break if it does not work)
15 embedding_service = llamafile_embedding_service(
16     auth_form.value["emb_llamafile_url"])
17 }
18
19 no_stop(
20     not embedding_service.is_working(),
21     no_md("**Cannot access llamafile embedding server.**"),
22 )
23
24 # collect the names of the timelines we want to download from
25 timelines = []
26 for k in timelines_dict.keys():
27     if auth_form.value["k"] == "home":
28         tl_string = timelines_dict[k]
29         if tl_string in ["tag", "list"]:
30             tl_string = f"/{auth_form.value["k"].replace(" ", "_")}"
31         timelines.append(tl_string)
32
33 # set offline mode
34 offline_mode = auth_form.value["offline_mode"]
```

Getting data from my Mastodon account...

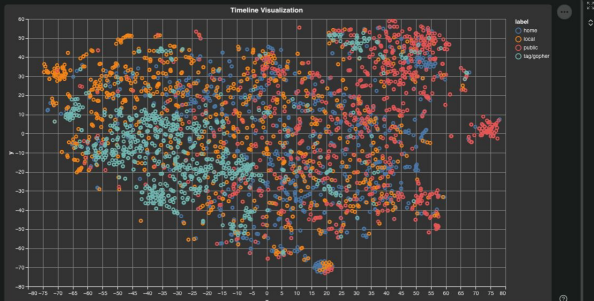
1. # Getting data from my Mastodon account...

```
1 paginated_data = build_cache_paginated_data(mastodon_client,
2                                              timelines,
3                                              cached_timelines,
4                                              paginated_data_file)
5 no_stop(paginated_data is None, no_md("**Issues connecting to Mastodon**"))
6
7 dataframes = build_cache_dataframes(paginated_data,
8                                     cached_dataframes,
9                                     dataframes_data_file)
10
11 no_stop(paginated_data is None, no_md("**Issues connecting to Mastodon**"))
12
13 loading cached paginated data from dump_paginated_data.pkl
14 loading cached dataframes from dump_dataframes.pkl
```

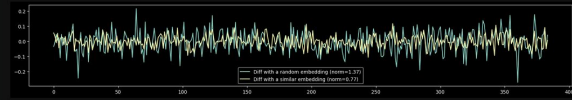
My timeline(s)

```
1 # My timeline(s)
2
3 calculate_embeddings
4 embeddings = build_cache_embeddings(embedding_service,
5                                     dataframes,
6                                     cached_embeddings,
7                                     embeddings_data_file)
8
9 loading cached embeddings from dump_embeddings.pkl
```

```
1 import numpy as np
2
3 def time(dataframes, embeddings, perplexity, random_state=42):
4     """Runs dimensionality reduction using T-SNE on the input embeddings.
5     Returns dataframes containing status id, text, and 2D coordinates
6     for plotting.
7     """
8     time = T-SNE(n_components=2, random_state=random_state, perplexity=perplexity)
9
10    all_embeddings = np.concatenate([v for v in embeddings.values()])
11    all_projections = time.fit_transform(all_embeddings)
12
13    dfs = []
14    start_idx = 0
15    end_idx = 0
16    for k in embeddings:
17        end_idx = len(embeddings[k])
18        df = dataframes[k]
19        df["*"] = all_projections[start_idx:end_idx, 0]
20        df["y"] = all_projections[start_idx:end_idx, 1]
21        df["label"] = k
22        dfs.append(df)
23        start_idx = end_idx
24
25    return pd.concat(dfs, ignore_index=True), all_embeddings
26
27 df, all_embeddings = time(dataframes, embeddings, perplexity=16)
28
29 chart = m.viz.altair.Chart(
30     alt.Chart(df, title="Timeline Visualization", height=500)
31     .mark_point()
32     .encode(
33         x="*",
34         y="y",
35         color="label"
36     )
37 )
```



```
1 no.vstack(
2     [
3         chart,
4         chart.value["id", "label", "text"] if len(chart.value) > 0 else chart.value,
5     ]
6 )
```



Timeline re-ranking

User statuses

Number of pages to load ☐ Exclude reblogs

Timelines

☐ home
☐ local
☒ public

Clear Submit

1. rerank_form

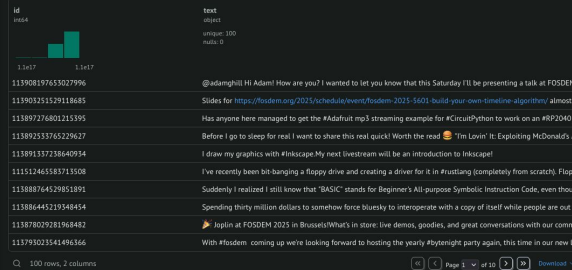
Account Statuses (replies=True, reblogs=True)

3.88 reblogs / Total time 1.28s

```
1 # check for anything invalid in the form
2 no_stop(rerank_form.value is None,
3         no_md("**Submit the form to continue.**"))
4
5 user_statuses = byota_mastodon.get_paginated_statuses(mastodon_client,
6                                                       max_pages_rerank_form.value["num_user_status_pages"],
7                                                       exclude_reblogs_rerank_form.value["exclude_reblogs"])
8
9 loading page 1: max_id = None
10 loading page 2: max_id = 13386284484266339
11 loading page 3: max_id = 133771326236064274
12 loading page 4: max_id = 133616597353386528
13 loading page 5: max_id = 133395097776558585
```

Your statuses:

1. # Your statuses:



Configuration

Mastodon Credentials

Login: Password:

Timelines

☒ Home ☐ Local ☐ Public

☐ Hashtag ☐ List

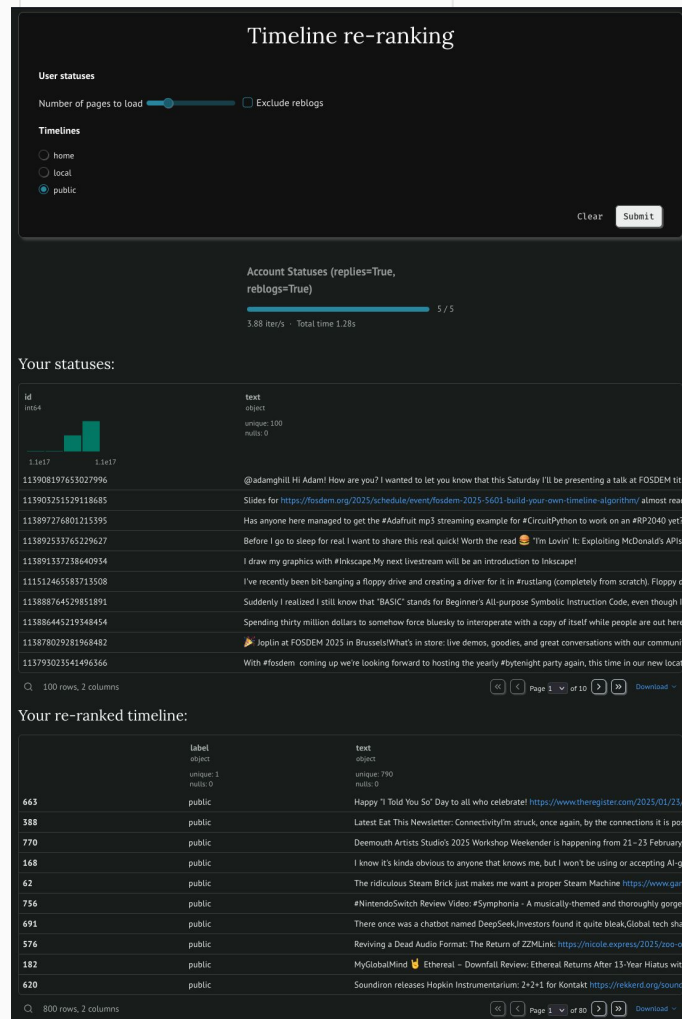
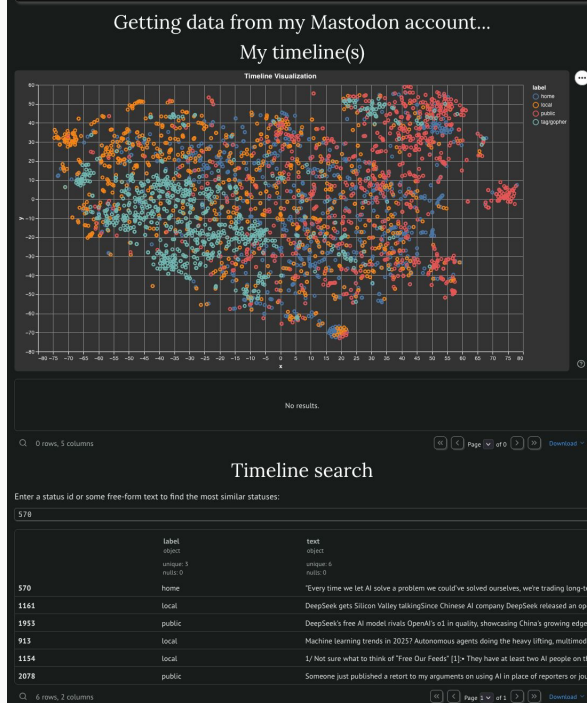
Embeddings

Embedding server URL

Caching

☐ Run in offline mode

Clear Submit



Configuration

Mastodon Credentials

Login: Password:

Timelines

☒ Home ☐ Local ☐ Public

☐ Hashtag ☐ List

Embeddings

Embedding server URL

Caching

☐ Run in offline mode

Clear



Timeline search

Enter a status id or some free-form text to find the most similar statuses:

	label	text
	object	object
	unique: 1	unique: 6
	ratio: 0	ratio: 0
2116	public	Software Libre, alternativa para la independencia de Meta y garantía de la
891	local	Visit #openSUSE this coming weekend for all things #opensource at #FOSD
856	local	Best Cross-Platform Apps for #Linux, #Windows, and #Mac in 2023https://w
1294	local	The Value of Open Source Software https://papers.ssrn.com/sol3/papers.cfm
2969	tag:python	Por qué importa el software libre y las plataformas descentralizadas https://
2968	tag:python	Por qué importa el software libre y las plataformas descentralizadas https://

6 rows, 2 columns

68

79

Page 8 of 1

88

Download

Timeline re-ranking

User statuses

Number of pages to load ☐ Exclude reblogs

Timelines

☐ home

☐ local

☒ public

Clear

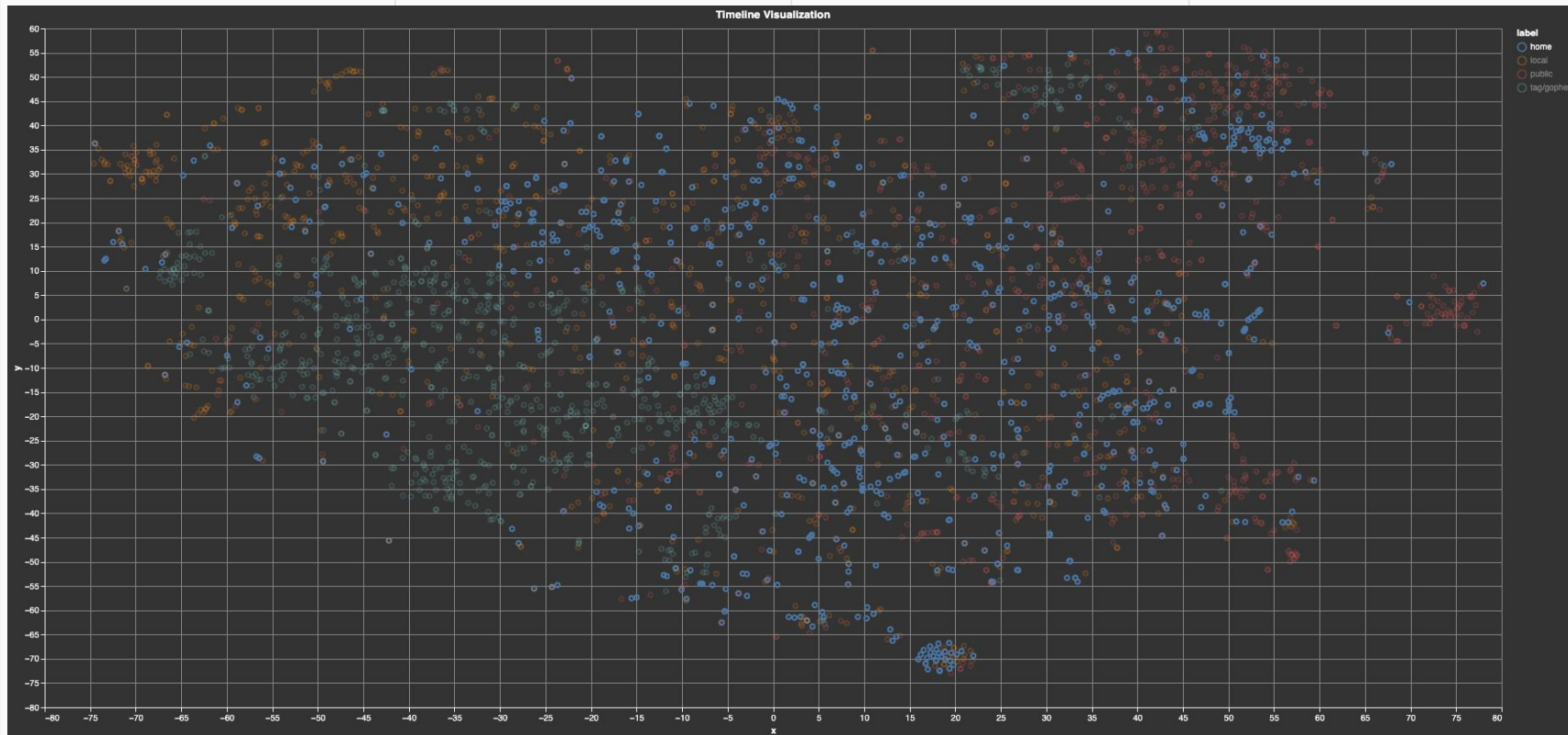
My own posts, re-ranked according to their similarity to posts in tag/bookstodon

Getting data
from my
Mastodon
account...

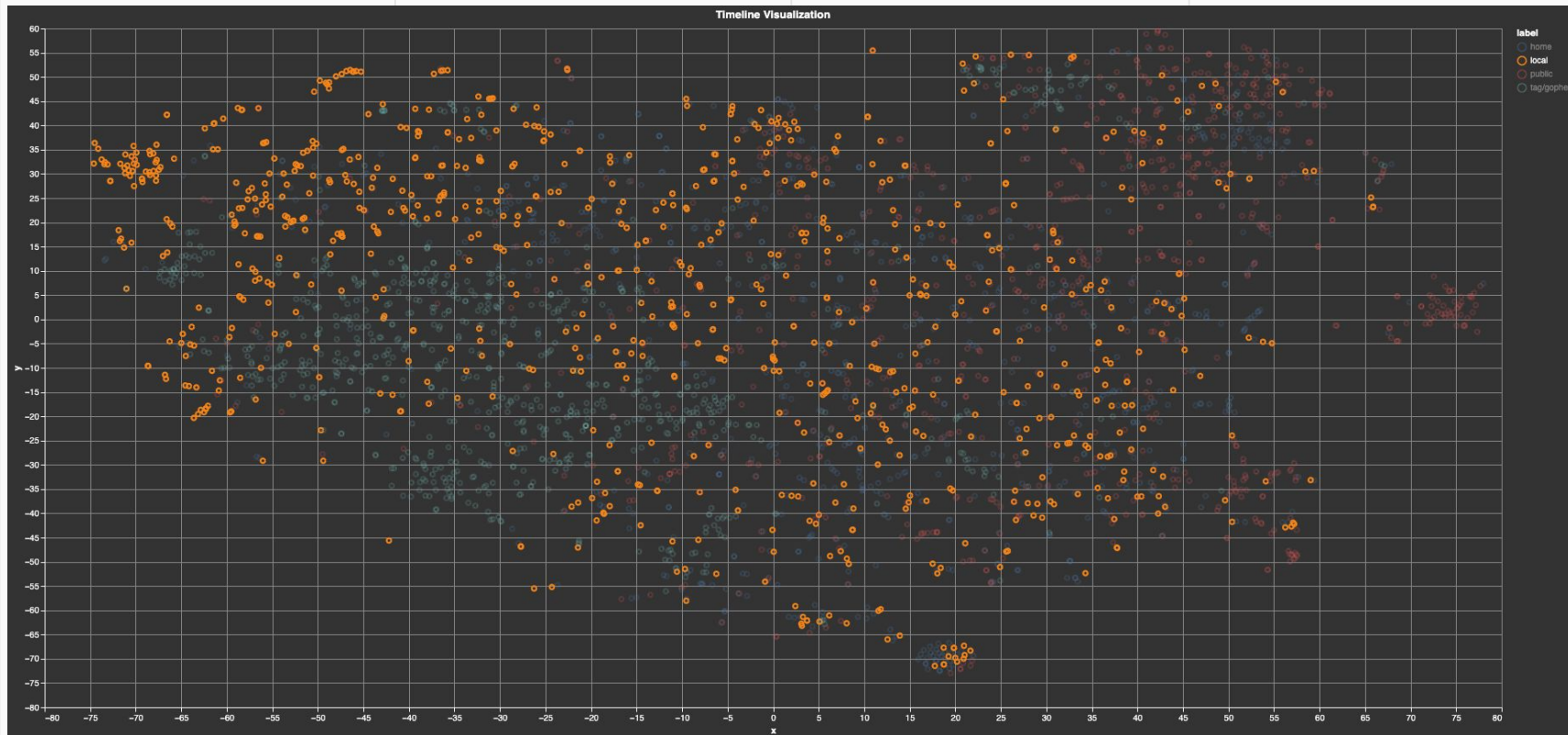
My timeline(

BYOTA in practice

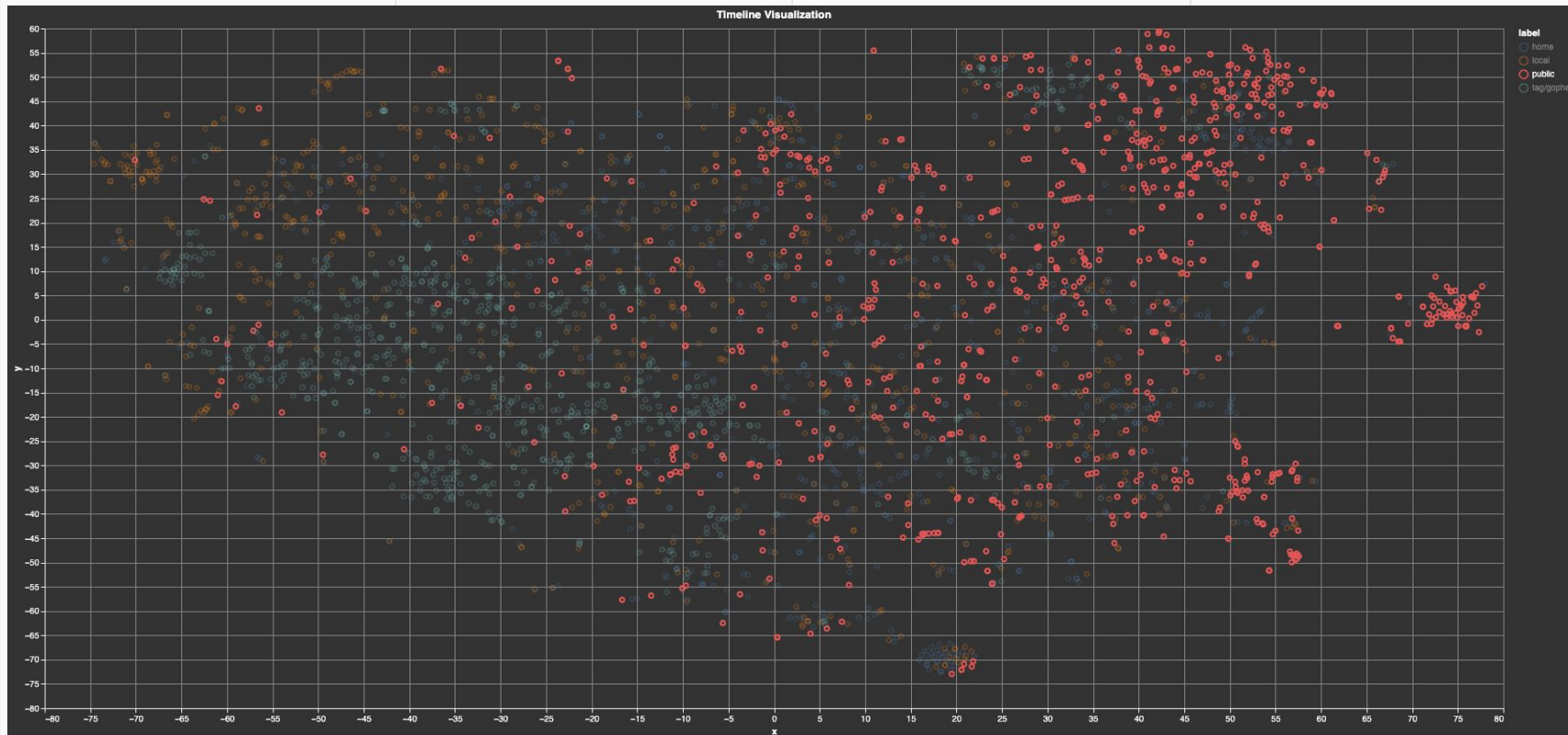
Embeddings visualization



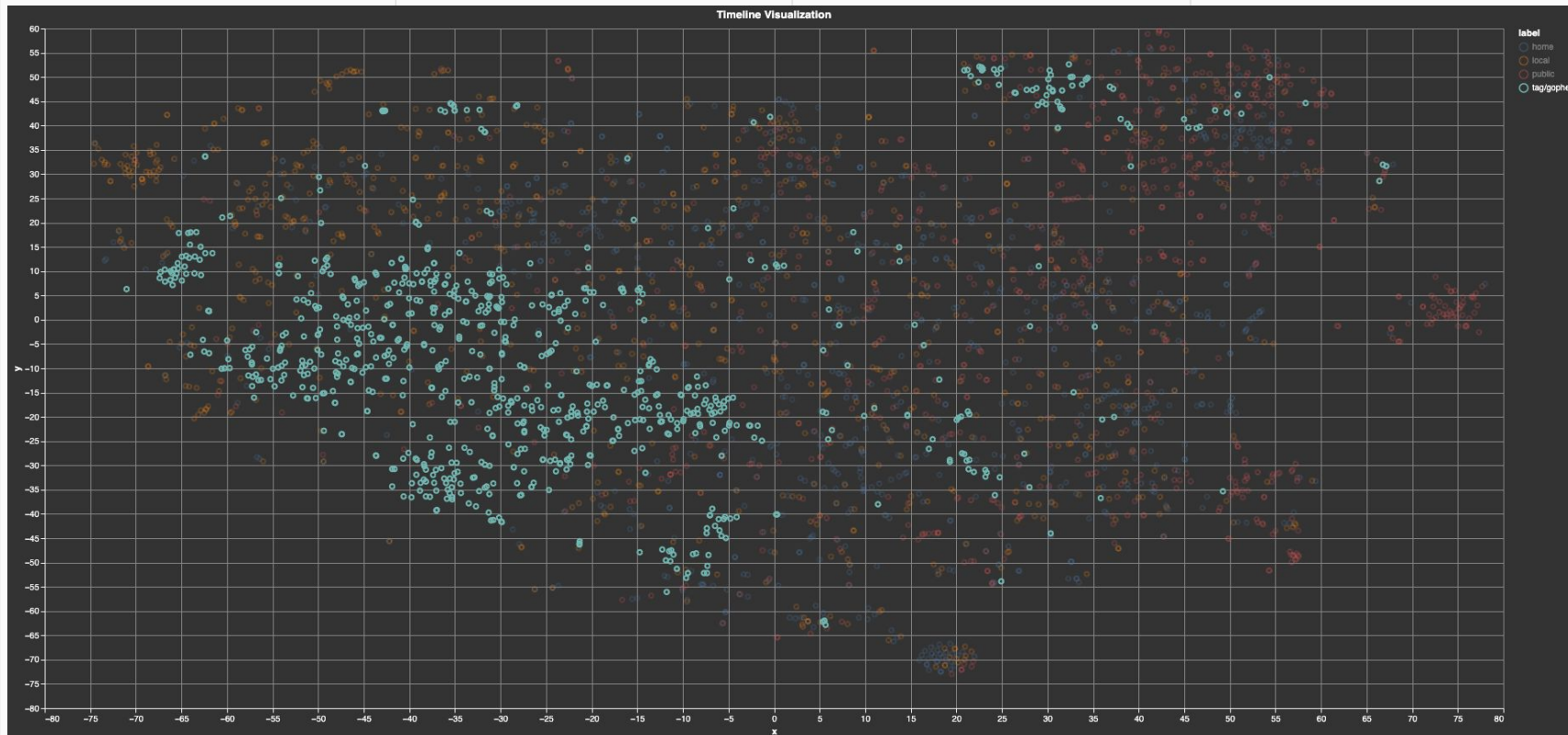
Embeddings visualization



Embeddings visualization



Embeddings visualization



People united by... World Series



People united by... Flights



People united by... Reading



People united by... Halloween

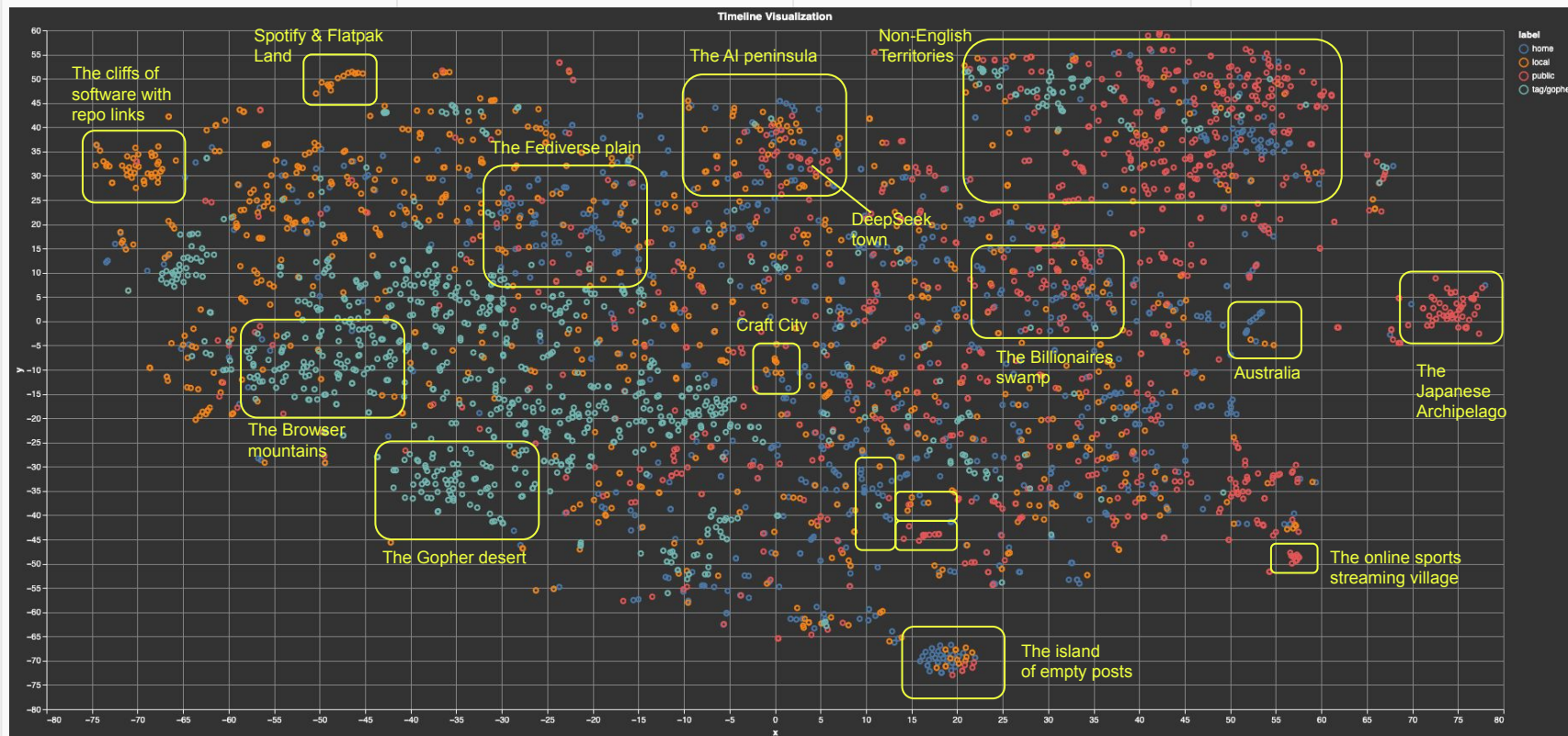


People united by... Language (or not?)

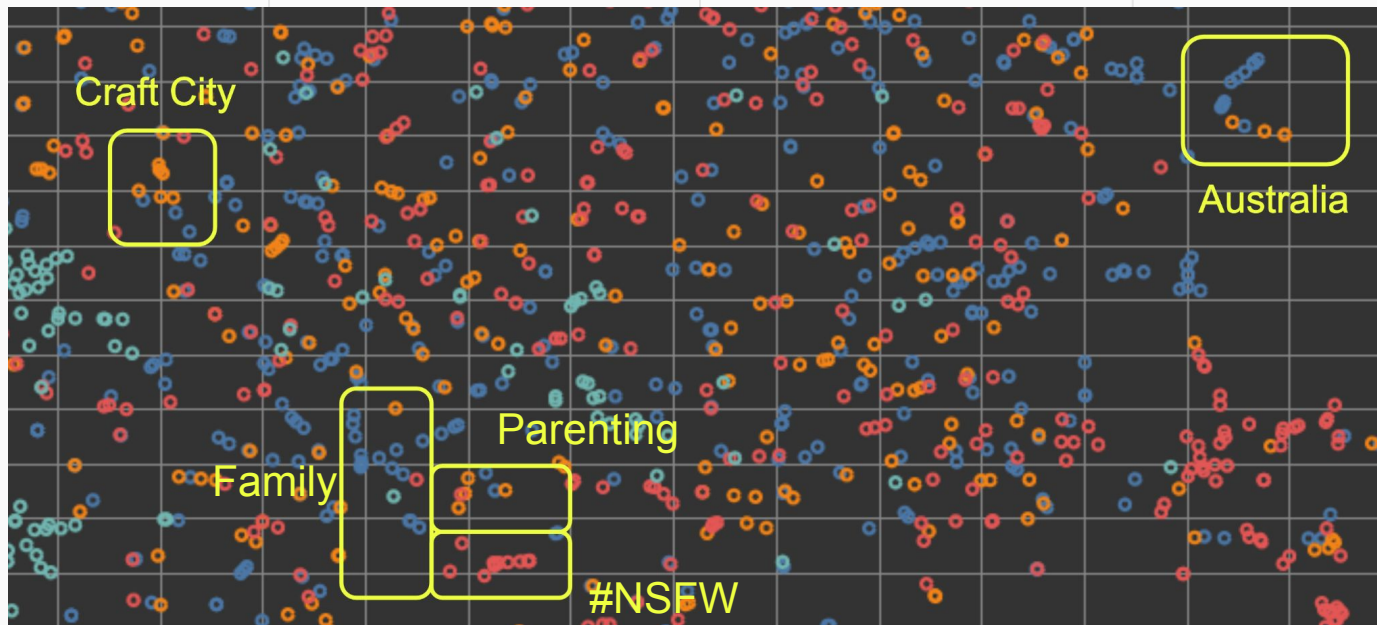


	id object	label object	text object	🔍
24	115479533899245296	fosstodon.home	Il giornalismo nell'era dell'impunità https://www.valigiablu.it/minacce-giornaliste-...	
1464	115479967360543380	fosstodon.public	🍷 J'ai nommé ce plat "Soleil d'Automne" 🍷 #food #nourriture #cuisine #champignons	
1769	115479938592298522	fosstodon.public	📰 Digital research 2025: як змінилися аудиторія та споживання? Аудиторія в Укр...	
1539	115479959006800920	fosstodon.public	🚩 EINHEIT DES VERHINDERUNGSFALLS: Kein Neustart der Lohnfortzahlung Das Ge...	
1744	115479941089212782	fosstodon.public	🚩 Betriebsrats-Widerspruch allein reicht nicht! Trotz fristgerechten Widerspruchs m...	
3553	115479733228653709	macaw.public	🇪🇸 🇮🇹 🇩🇪 "Gondoliere" ❤️ (@ RJ 132 → Wien Meidling) #NowTravelling https://trae...	
1485	115479965499343514	fosstodon.public	Florence Welch wil niet beroemer worden dan ze nu is Florence Welch is tevreden ...	
3410	115479822985206402	macaw.public	❄️ La neige s'est invitée à la fin du front à #Chamrousse où il est tombé à peu près...	
1592	115479954352429714	fosstodon.public	У китайському готелі почали використовувати гуманоїдних роботів для обслугову...	
1403	115479974114616070	fosstodon.public	Традиційно сьогодні розіграємо ножа!Практичний і зручний для виконання повся...	

Embeddings visualization



Embeddings visualization



Semantic Search

Semantically similar posts are returned from multiple timelines

Timeline search

Enter a status id or some free-form text to find the most similar statuses:

I am a fan of free software and everything opensource

	label object	text object
593	fosstodon:home	You already have the skills needed to support the #FreeSoftware movement, even if you don't h...
920	fosstodon:local	Want cool #LibreOffice merch? And to expand your skillset? Join us this month: https://blog.docu...
65	fosstodon:home	People with #FreeSoftware experience make excellent workers! Learn how you can advertise yo...
1209	fosstodon:local	Tools That Just WorkApps that do one thing, do it well, and quietly make my days smoother. http...
993	fosstodon:local	Wishing everyone a great time at OpenAlt starting today in Faculty of Information Technologies,...
1152	fosstodon:local	https://www.pgedge.com/blog/pgedge-goes-open-source #postgresql

Re-Ranking

Timeline re-ranking

User statuses

Download user statuses from:

- ☒ fosstodon
☐ macaw

Number of pages to load ☒ Exclude reblogs

Timeline to rerank

Timeline to rerank:

- ☒ fosstodon:home
☐ fosstodon:local
☐ fosstodon:public
☐ fosstodon:tag/caturday
☐ macaw:local
☐ macaw:public
☐ macaw:tag/caturday

Clear

Submit

Re-Ranking

Your statuses:

id object	text object
115083357438964205	Fortunately, the ecosystem is still open enough that one can provide compatible open alternatives:- tools can be easily swapped ...
115083357430312648	The thing is, models did improve but the huge boost in performance you have experienced is brought by the use of tools. Even jus...
115083357419301018	Agents bring both the chance for open, personal, private AI and the risk of opening new pathways to enshittification.Example of t...
115045887950633227	Oh yes, the code: it's at https://github.com/aittalam/ch35tlf you want to build your own client using the library, you can install it ...
115045887944717559	http://3564020356.org/ch35t/wasm/ is the endpoint. The riddle you'll find is the one that gives you access to my old website in c...
115045887939283403	"All this effort and you still connect to a website? How's that decentralized?"I have now become just one endpoint: anyone can (st...
115045887932024010	#ch35t is also a python library that parses its own format. Until yesterday, it came with a CLI client to read a json, validate it, sho...
115045887926471701	What does a #ch35t look like? Check out http://3564020356.org/ch35t/examples/ for a few examples.What riddles can you create...
115045887920193221	#Ch35t is called this way because "it's a chest with a little bit of 3564020356.org in it"👉It is first and foremost a format, defined ...
115045887912947303	If I had to change one thing from my old 356* hacking challenge website, that would be its centralized nature. Having a curated s...

200 rows, 2 columns

Page 2 of 20 Download

Your re-ranked timeline:

	label object	text object
398	fosstodon:home	YunoHost 13 (Trixie) beta, spooookiiie edition!Following the alpha-testing opened a few wee...
26	fosstodon:home	Wikidata:Con begins today!Well we did it again - 9 sessions from lightning talks, presentation...
58	fosstodon:home	YunoHost 13 (Trixie) beta, spooookiiie edition!Following the alpha-testing opened a few weeks ago, we are happy to announce that we are moving to the beta-testing stage for #YunoHost 13!The announcement is here : https://forum.yunohost.org/t/yunohost-13-0-trixie-spooky-beta/40656 Time has come to test it extensively (if you feel confident enough in your technical skills) and share your feedbacks!Happy selfhosting!
155	fosstodon:home	@Fediverse We've
193	fosstodon:home	Do I trust such a local LLM beyond such very specific and rather simple, technical questions? No...
283	fosstodon:home	Come to the @PrototypeFund Demo Day in Berlin where we will be presenting "Reflection", our ...
388	fosstodon:home	Almost finished coding the last major piece of the Loops app, the Camera.It would be done by n...
324	fosstodon:home	Fediverse Report 141 - this week's #fediverse newsMastodon is working towards making one of ...
591	fosstodon:home	Fediverse Report 141 - this week's #fediverse newsMastodon is working towards making one of ...
552	fosstodon:home	Fediverse Report 141 - this week's #fediverse newsMastodon is working towards making one of ...

672 rows, 2 columns

Page 1 of 68 Download

Re-Ranking by tag

Your own posts, re-ranked according to their similarity to posts in fosstodon:tag/bouldering

	text object	scores float64
155	Back from a mildly satisfying gym #bouldering session and I realize I do not talk about this enou...	7.8666839552
154	Also I do not find a lot of recent posts when I look for #bouldering #climbing #indoorclimbing	7.2781694099
106	Just back from the climbing gym where I saw two people, separately, checking out slack betwee...	7.0418709622
49	#LoRa + #bouldering = #climbtastic? 🤔#meshtastic #climbing #Weekend	6.8754765148
113	Walk-Tube-Train-Plane-Bus-Walk and I am in San Donà di Piave. Now sleep, then https://www.en...	6.34246246
112	Back on a plane for my flight back home from 3 great days at #EndSummerCamp. I am grateful f...	5.6126990633
145	Weekend plan:- add Gemini support to #picogopher- build web sandbox for #ch35t- build a #bra...	5.1980869574
62	Back from an intense work week, my first thought is "I need some rest" but that's immediately fo...	5.0549179675
40	Mountain hike, middle school party, lunch with the whole family, prepared dinner and bolognes...	4.8805070409
165	I took many flights in my life, and I have been traveling more and more in the last few years, but...	4.8180404027

🔍 200 rows, 2 columns

⏪ ⏩ Page 1 ▾ of 20 ⏪ ⏩ Download ▾

Re-Ranking by tag

Your own posts, re-ranked according to their similarity to posts in fosstodon:tag/books

	text object	scores float64
137	I finished reading Attack Surface! https://craphound.com/category/attacksurface/ It's been quite a ...	8.1781680911
117	I just finished reading Morphotropic by @gregeganSF and I loved it!The way I learned about it ...	7.8890159816
196	I am reading Jenny Odell's "How to Do Nothing: Resisting the Attention Economy", and I am enjo...	7.5804552178
128	#AmReading #bookstodon #books Massimo Temporelli: "Noi siamo tecnologia" (We are are tech...	7.5375652196
125	I am reading Dan Siegel's "Brainstorm: the power and purpose of the teenage brain"#AmReading...	7.1451445645
17	#weekend update:- finished reading @pluralistic 's The Bezzle, started Picks & Shovels- prepare...	6.8922864652
149	Related material: http://davide.eynard.it/zelif/20210323%20-%20CELI.pdf 🙄	6.8515546448
145	Weekend plan:- add Gemini support to #picogopher- build web sandbox for #ch35t- build a #bra...	6.8213979953
166	"Imagine caring deeply about something, so deeply that you feel compelled to master it. Imagin...	6.7136874617
105	Just saw the Ada and Zangemann movie 💙 It's so good! https://mastodon.social/@torproject/11...	6.3896691575

☐
200 rows, 2 columns

<<
<
Page 1 of 20
>
>>
Download

Too good to be true?

What about performance?

From bert to llama, time to embed 800 statuses

				2016 MB Pro		2024 M3 Max	
Model Name	Architecture	Size	Params	llamafiler 0.9.0	ollama 0.5.7	llamafiler 0.9.0	ollama 0.5.7
all-MiniLM-L6-v2.F16	bert	46 MB	22.6M	52	42	21	11
nomic-embed-text-v1.5	nomic-bert	274MB	137M	280	302	26	15
mxbai-embed-large-v1-f16	bert	670MB	334M	586	715	35	24
e5-mistral-7b-instruct	llama	5GB	7.11B	-	-	193	126

What about performance?

From bert to llama, time to embed 800 statuses

				2016 MB Pro		2024 M3 Max	
Model Name	Architecture	Size	Params	llamafiler 0.9.0	ollama 0.5.7	llamafiler 0.9.0	ollama 0.5.7
all-MiniLM-L6-v2.F16	bert	46 MB	22.6M	52	42	21 4.3	11
nomic-embed-text-v1.5	nomic-bert	274MB	137M	280	302	26	15
mxbai-embed-large-v1-f16	bert	670MB	334M	586	715	35	24
e5-mistral-7b-instruct	llama	5GB	7.11B	-	-	193	126

Embedding servers such as llamafile run locally

39

Is it truly local?

marimo runs wasm code in your browser. This is what you download:

Status	Method	Domain	File	Initiator	Type	Transferred	Size
200	GET	test-files.pythonhosted.org	marimo_base-0.10.17-py3-none-any.whl	:10 (fetch)	octet-stream	cached	666.91 kB
200	GET	files.pythonhosted.org	Markdown-3.6-py3-none-any.whl	:10 (fetch)	octet-stream	cached	105.38 kB
200	GET	cdn.jsdelivr.net	narwhals-1.10.0-py3-none-any.whl	:10 (fetch)	wasm	cached	1.12 MB
200	GET	files.pythonhosted.org	pymdown_extensions-10.8.1-py3-none-any.whl	:10 (fetch)	octet-stream	cached	250.83 kB
200	GET	cdn.jsdelivr.net	PyYAML-6.0.2-cp312-cp312-pyodide_2024_0_wasm32.whl	:10 (fetch)	wasm	cached	502.84 kB
200	GET	cdn.jsdelivr.net	altair-5.4.1-py3-none-any.whl	fetch	wasm	cached	5.45 MB
	GET	cdn.jsdelivr.net	typing_extensions-4.11.0-py3-none-any.whl	fetch	wasm	31.83 kB (raced)	140.47 kB
	GET	cdn.jsdelivr.net	Jinja2-3.1.3-py3-none-any.whl	fetch	wasm	111.43 kB (raced)	494.18 kB
200	GET	cdn.jsdelivr.net	MarkupSafe-2.1.5-cp312-cp312-pyodide_2024_0_wasm32.whl	fetch	wasm	cached	31.67 kB
200	GET	cdn.jsdelivr.net	jsonschema-4.21.1-py3-none-any.whl	fetch	wasm	cached	279.58 kB

(1) Python deps

(2) Mastodon statuses

...

(3) Embeddings

Status	Method	Domain	File	Initiator	Type	Transferred	Size	0 ms
200	GET	fosstodon.org	home?max_id=113880828320214162	:10 (xhr)	json	20.25 kB	110.35 kB	
200	OPTIONS	fosstodon.org	home?max_id=113880828320214162	xhr	plain	672 B	0 B	
200	GET	fosstodon.org	home?max_id=113880364682334614	:10 (xhr)	json	21.34 kB	112.50 kB	
200	OPTIONS	fosstodon.org	home?max_id=113880364682334614	xhr	plain	672 B	0 B	
200	GET	fosstodon.org	home?max_id=113880249393565733	:10 (xhr)	json	1.17 kB	2 B	
200	OPTIONS	fosstodon.org	home?max_id=113880249393565733	xhr	plain	672 B	0 B	
200	POST	localhost:8080	embedding	:10 (xhr)	json	5.47 kB	5.18 kB	
200	POST	localhost:8080	embedding	:10 (xhr)	json	5.47 kB	5.18 kB	
200	POST	localhost:8080	embedding	:10 (xhr)	json	5.49 kB	5.20 kB	
	POST	localhost:8080	embedding	:10 (xhr)				

What is next?

Conclusions

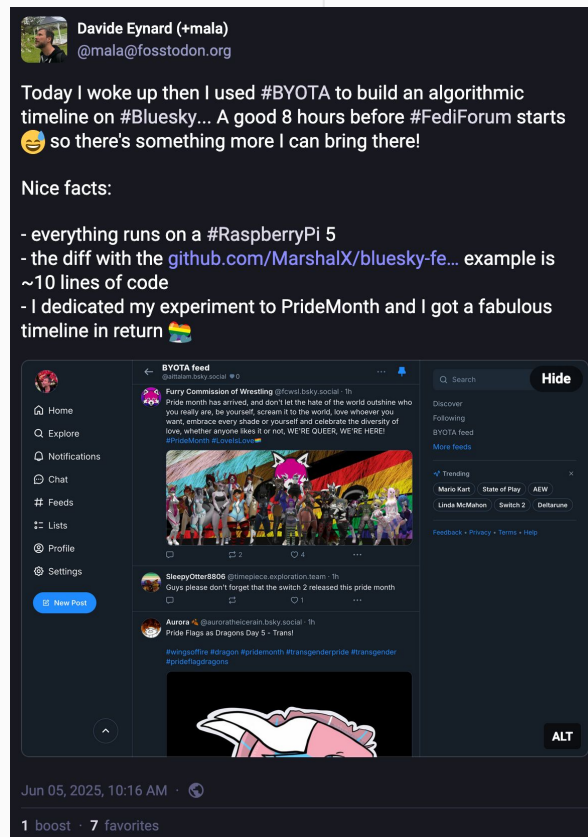
“Yet another application” has its limitations as an approach... But is helpful to convey the message:

Search and recommendation *can* be useful,
when they are tools *you* own

What does *own* mean in this case?

How can this help people build *their own solutions*?

Play with BYOTA and LMK! :-)



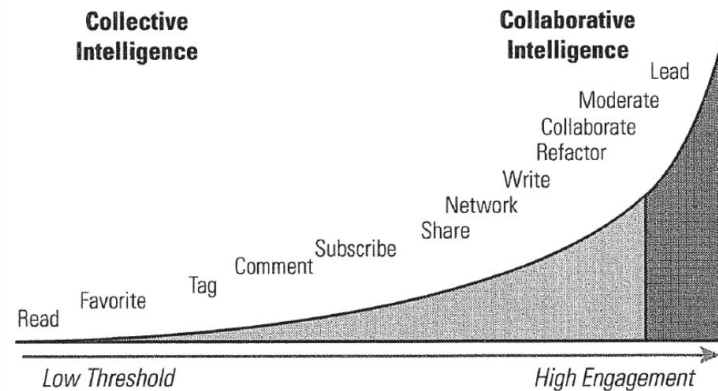
Can I run it NOW?

Of course!

- Demo on [HF spaces](#)
- Demo in a container (docker / podman)
- Run on real data in a container (with Mastodon credentials)
- Manually deploy everything locally



Figure 2.2 The power law of participation



Source: Mayfield 2006

Thank you!

 @mala@fosstodon.org

 [mozilla-ai/byota](https://github.com/mozilla-ai/byota)

