



FACULTY OF
APPLIED SCIENCES

Geotag Coverage & Mutual Attention

Oleksandr Zaytsev
Pavlo Myronov
Kateryna Liubonko

Supervisor: Bruno Gonçalves

Datasets

- ❖ Wikipedia dumps of pages' geographical coordinates

Questions

- ❖ How does the spatial distribution of geotags change from language edition to language edition?
- ❖ Does this reflect the relative importance that each culture attributes to each other?



Languages in question

1. Ukrainian
2. English
3. Russian
4. Polish
5. Japanese
6. Chinese

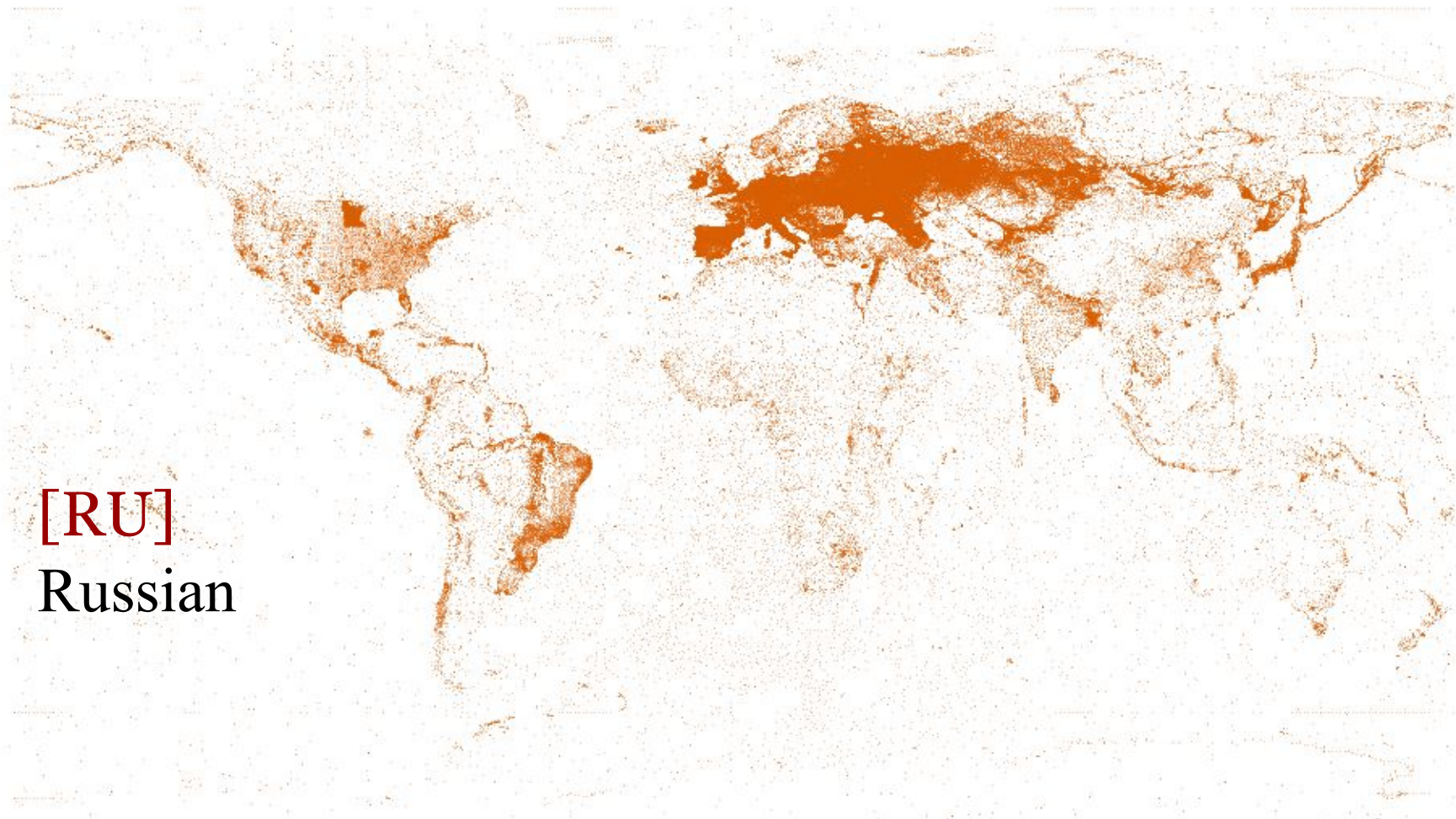
7. Spanish
8. German
9. Hebrew
10. Belarusian
11. Hungarian

[Step 1] Earth



[EN]

English



[RU]

Russian

A world map with a light gray background. The United Kingdom is highlighted in a solid red color. The rest of the world's landmasses are shown in a light gray tone, and the oceans are white.

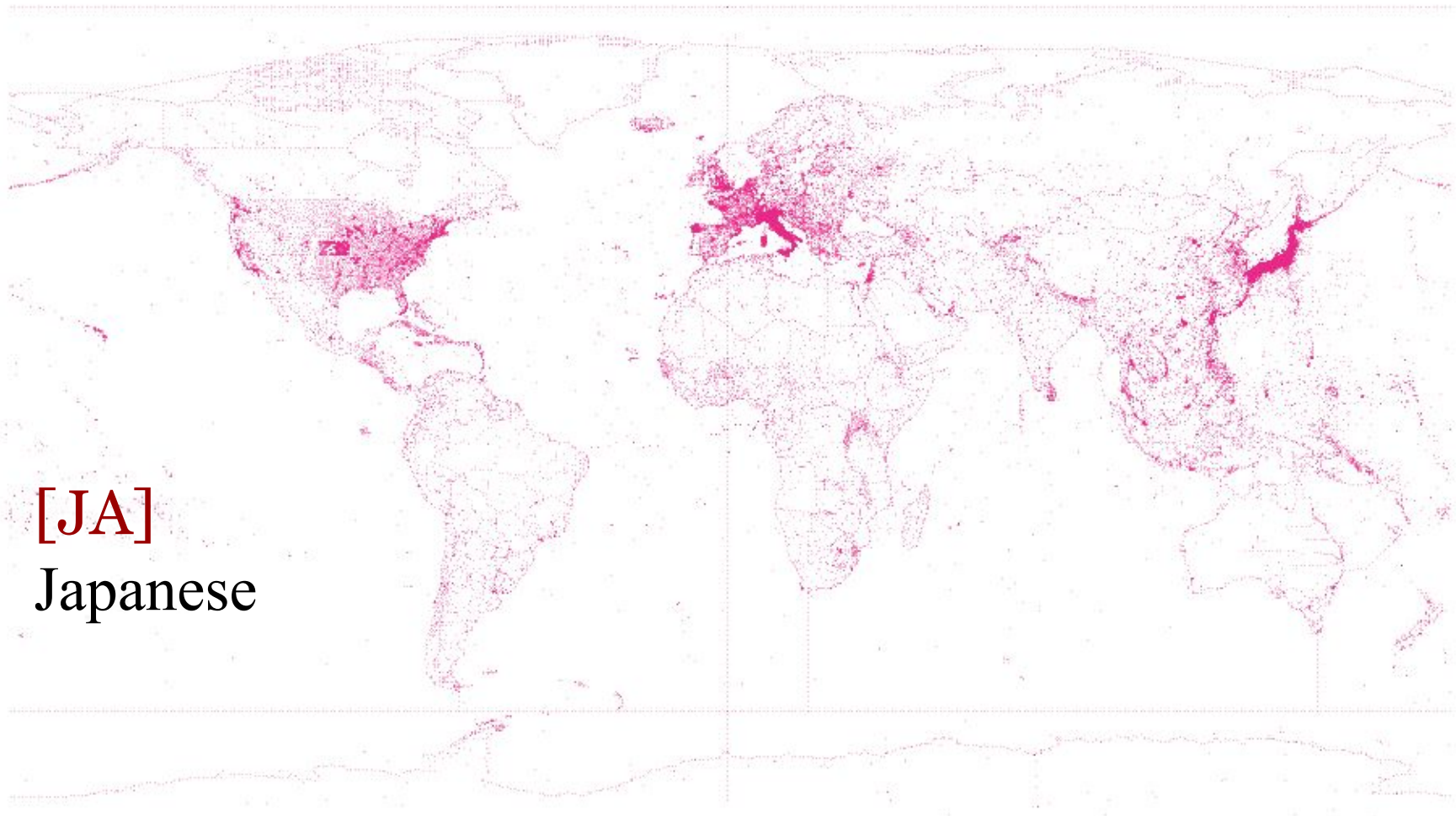
[UK]

Ukrainian



[ZH]

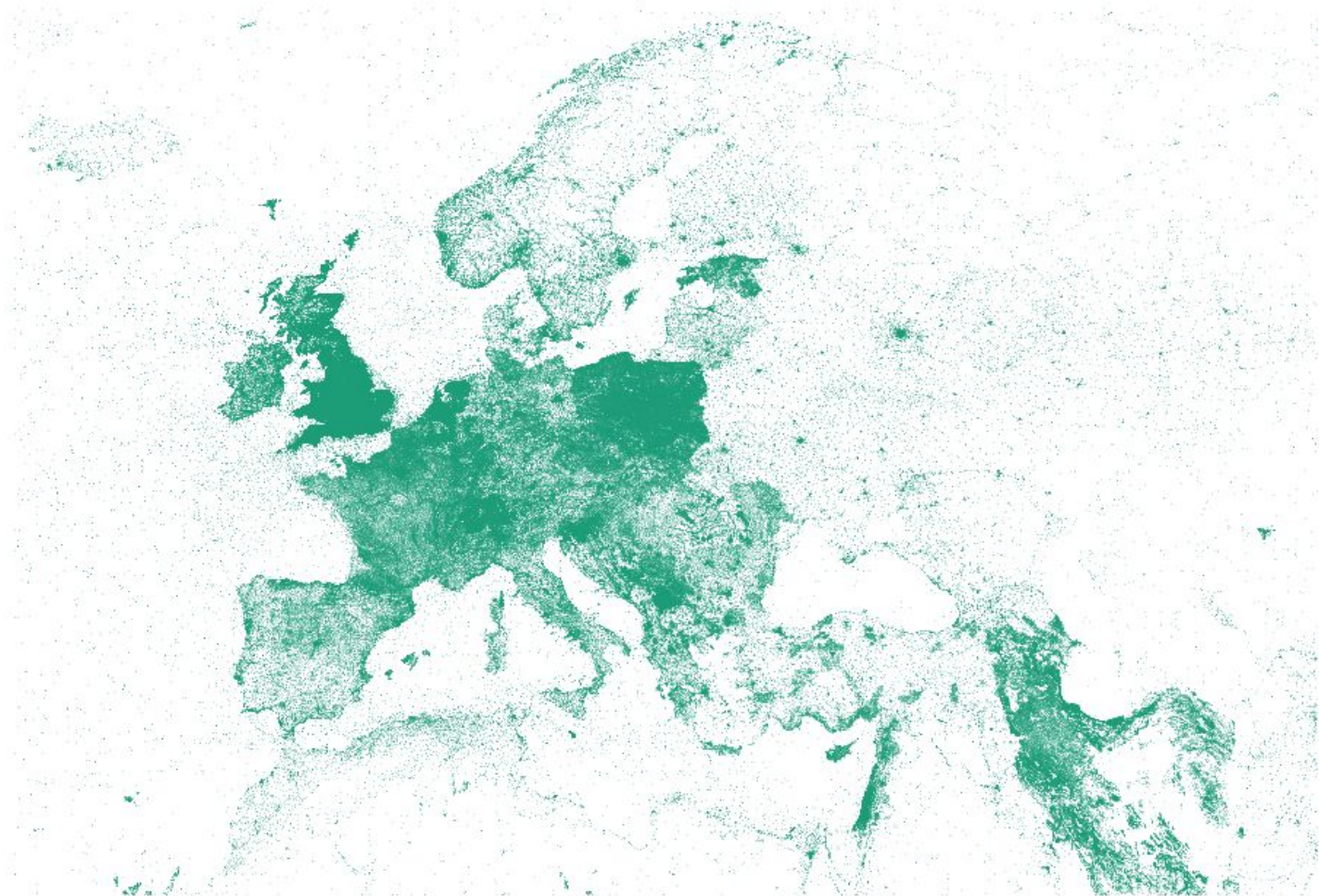
Chinese



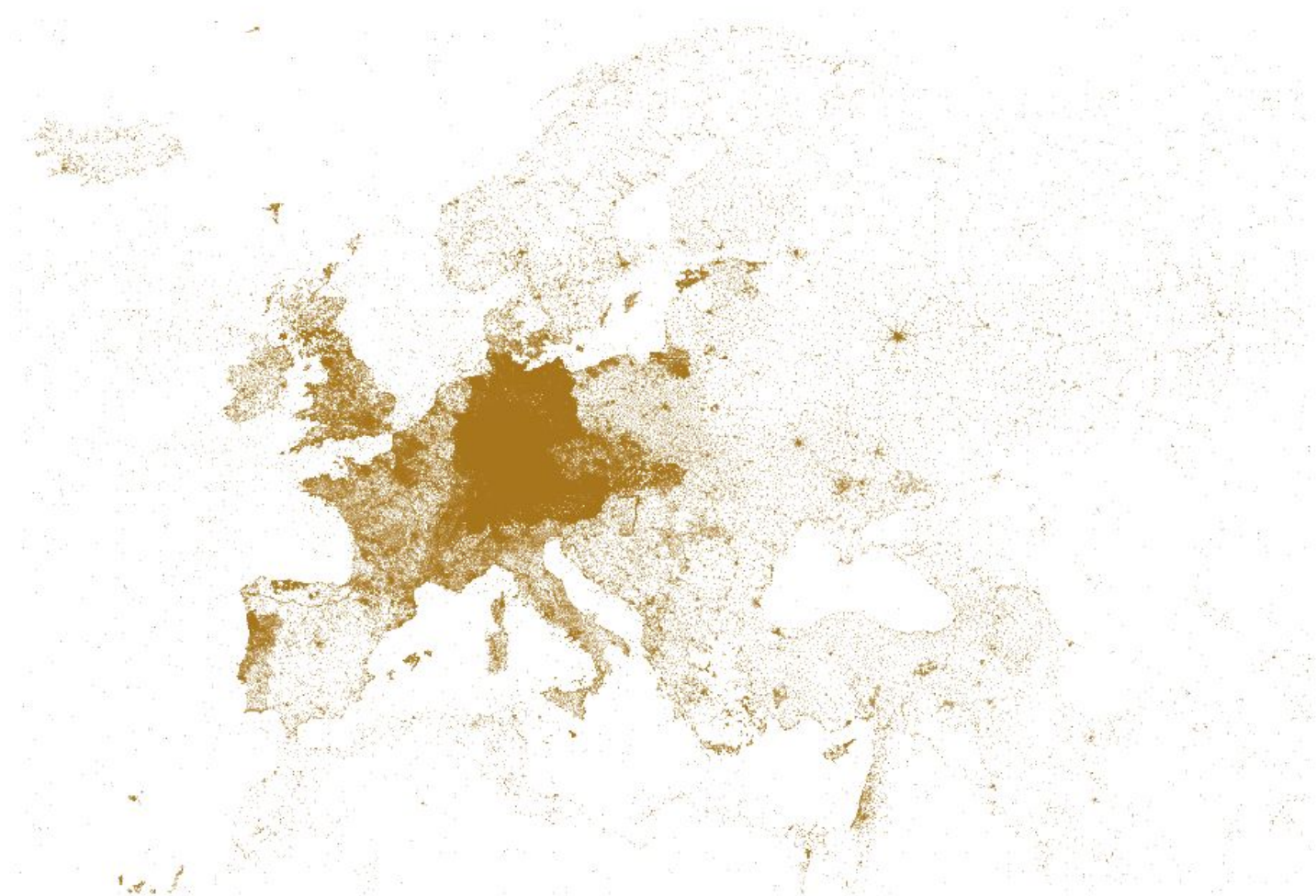
[JA]
Japanese

[Step 2] Europe

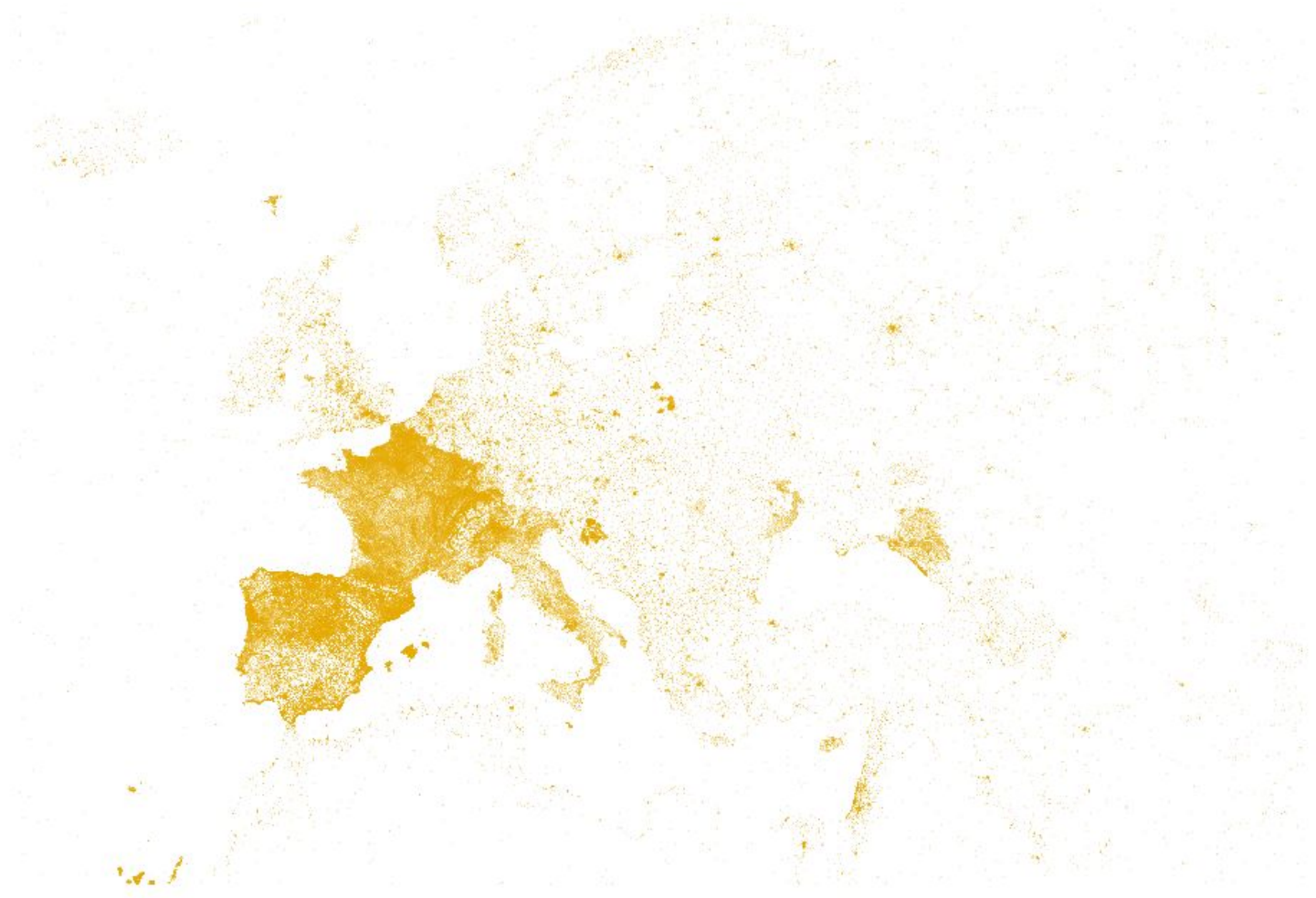
[EN]
English



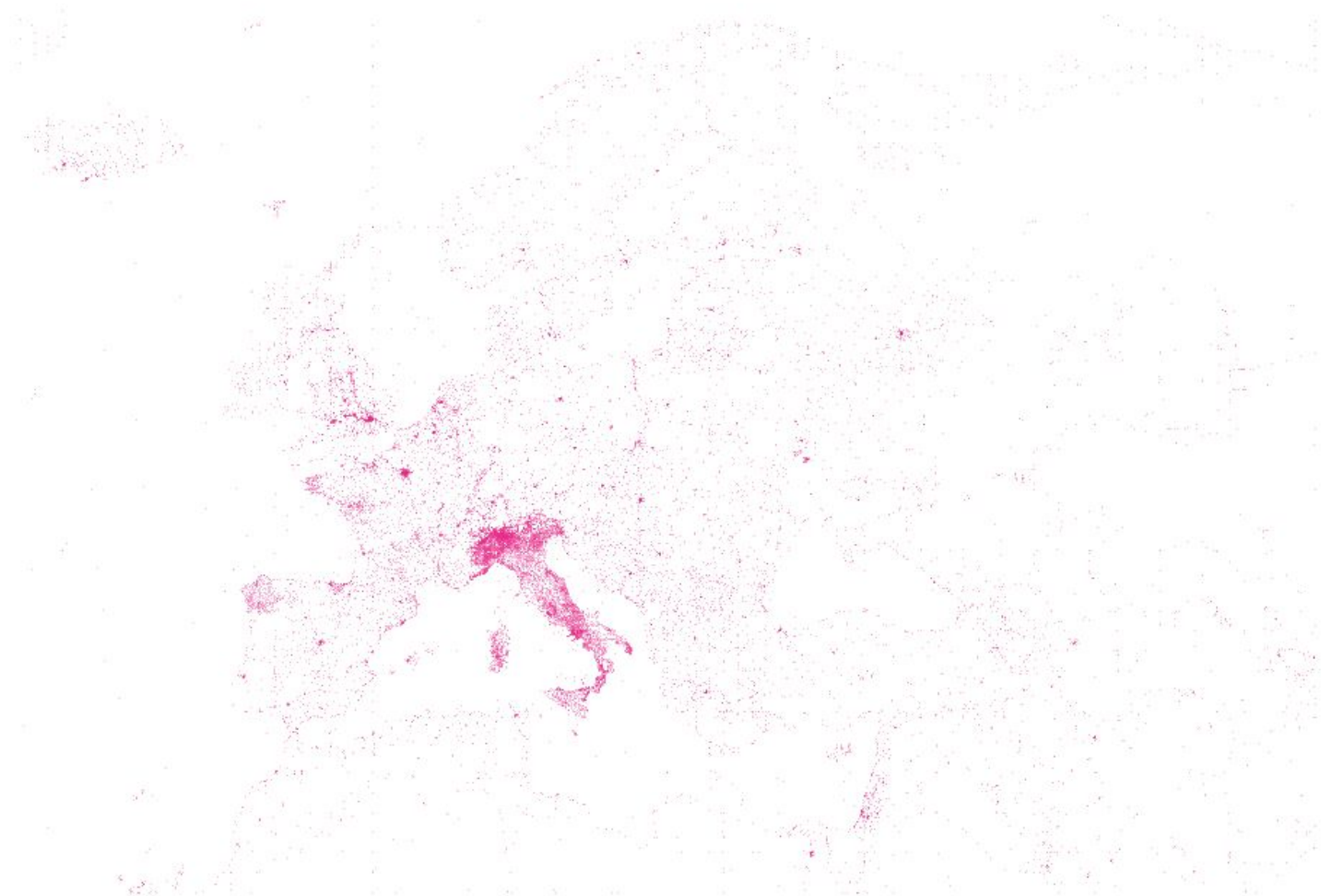
[DE]
German



[ES]
Spanish



[JA]
Japanese



[PL]
Polish

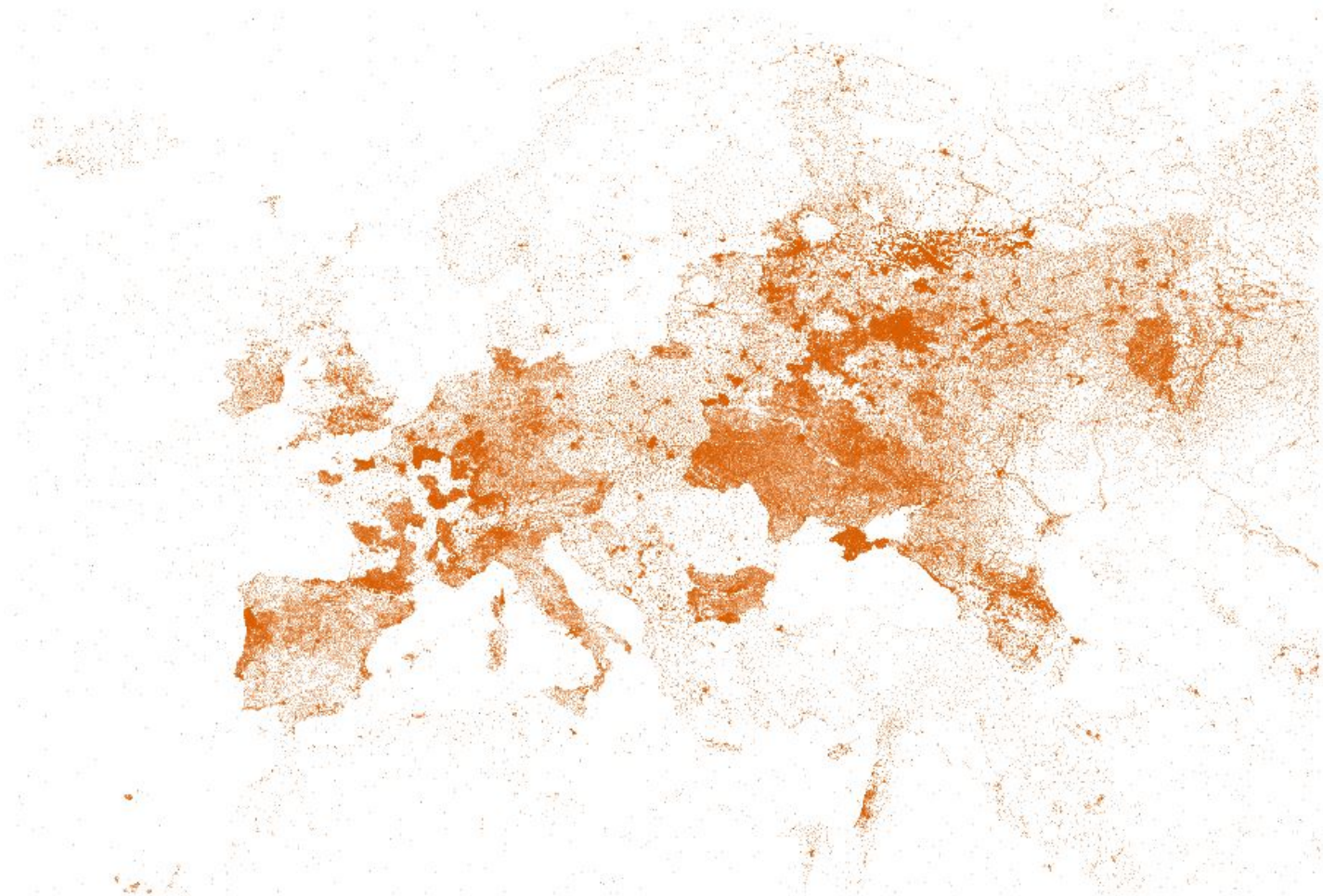


[UK]

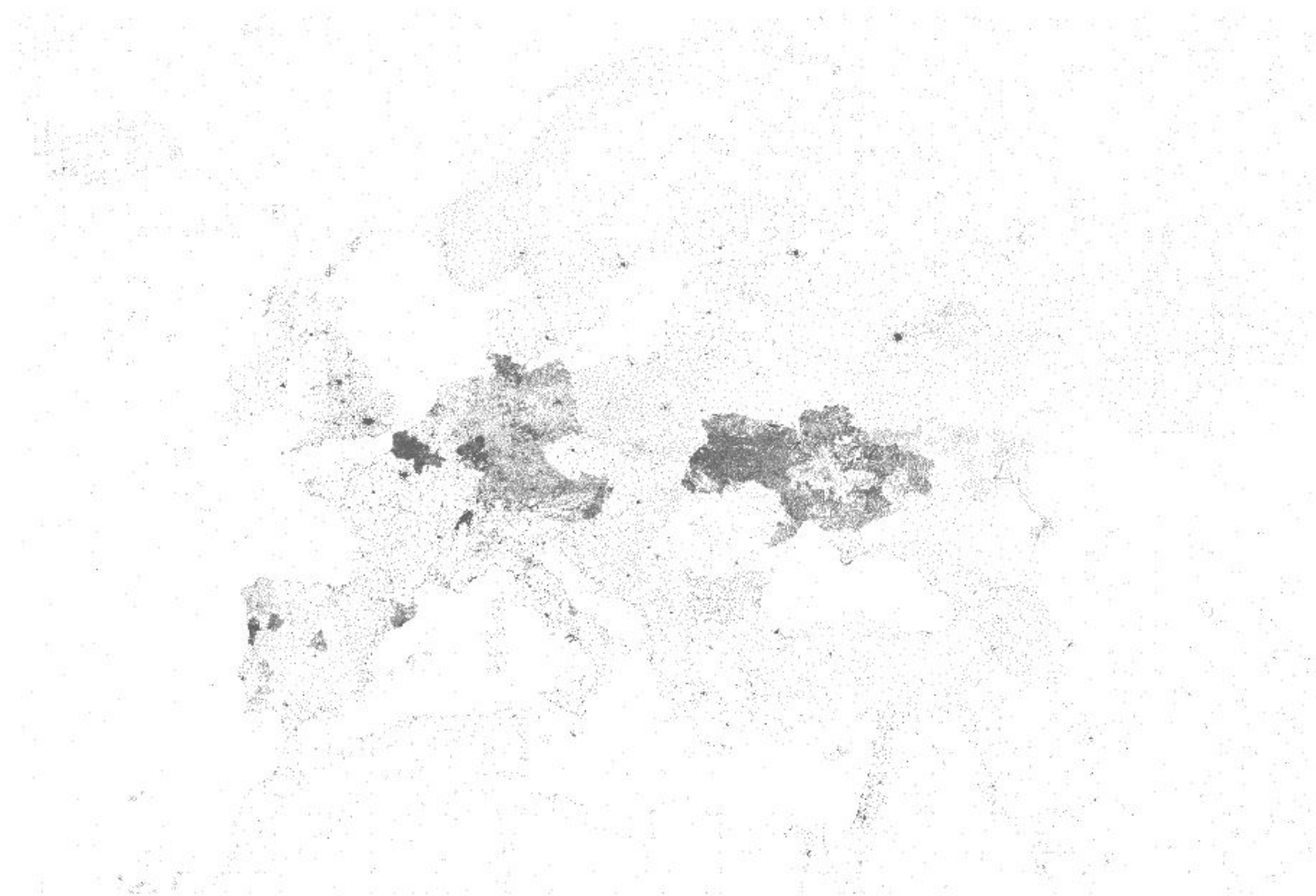
Ukrainian



[RU]
Russian

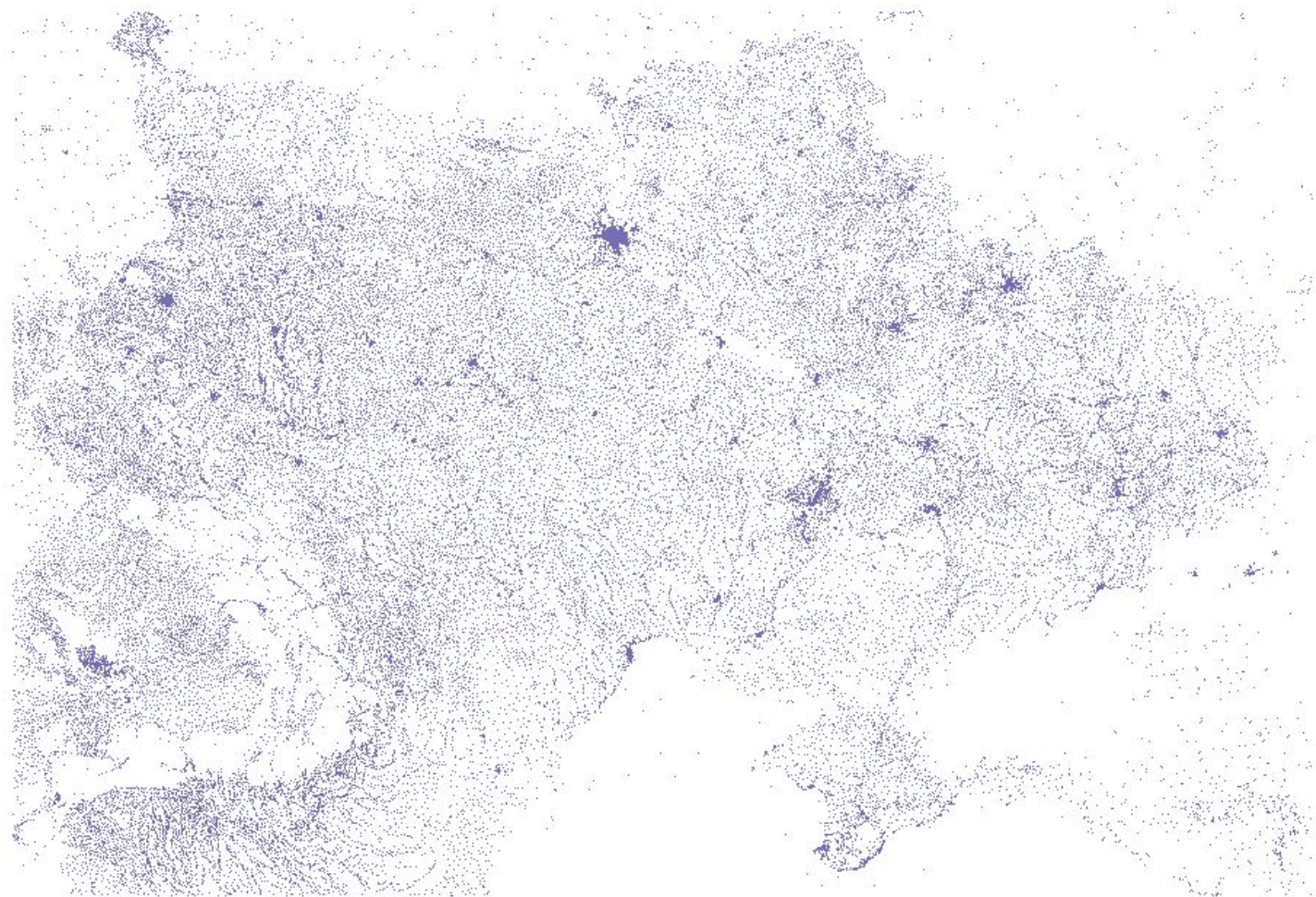


[ZH]
Chinese



[Step 3] Ukraine

[UK]
Ukrainian



[RU]
Russian

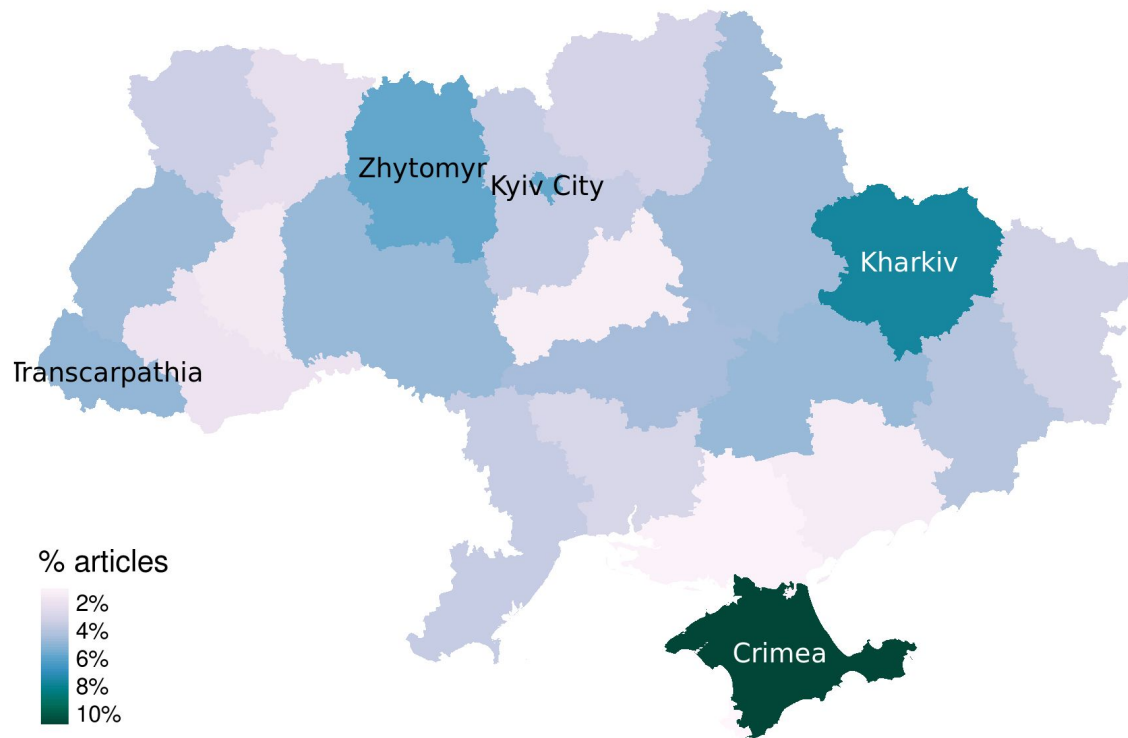


[PL]
Polish



[Step 4] Ukrainian Regions

[BE] Belorussian

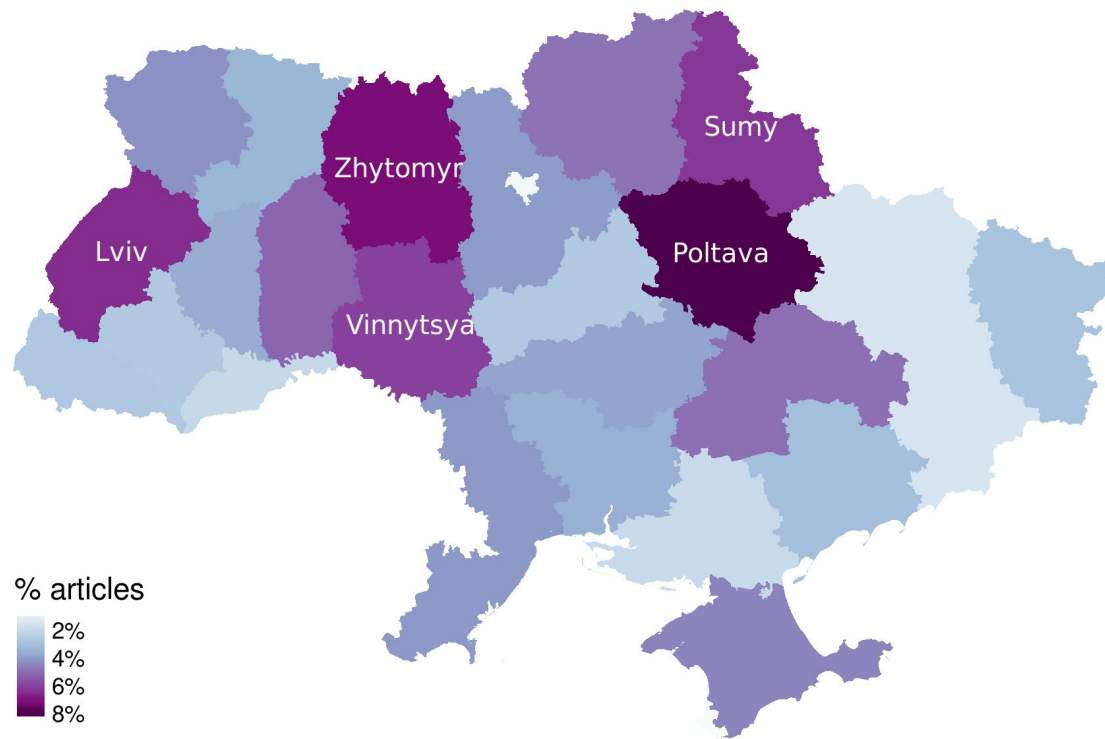


[HU] Hungarian

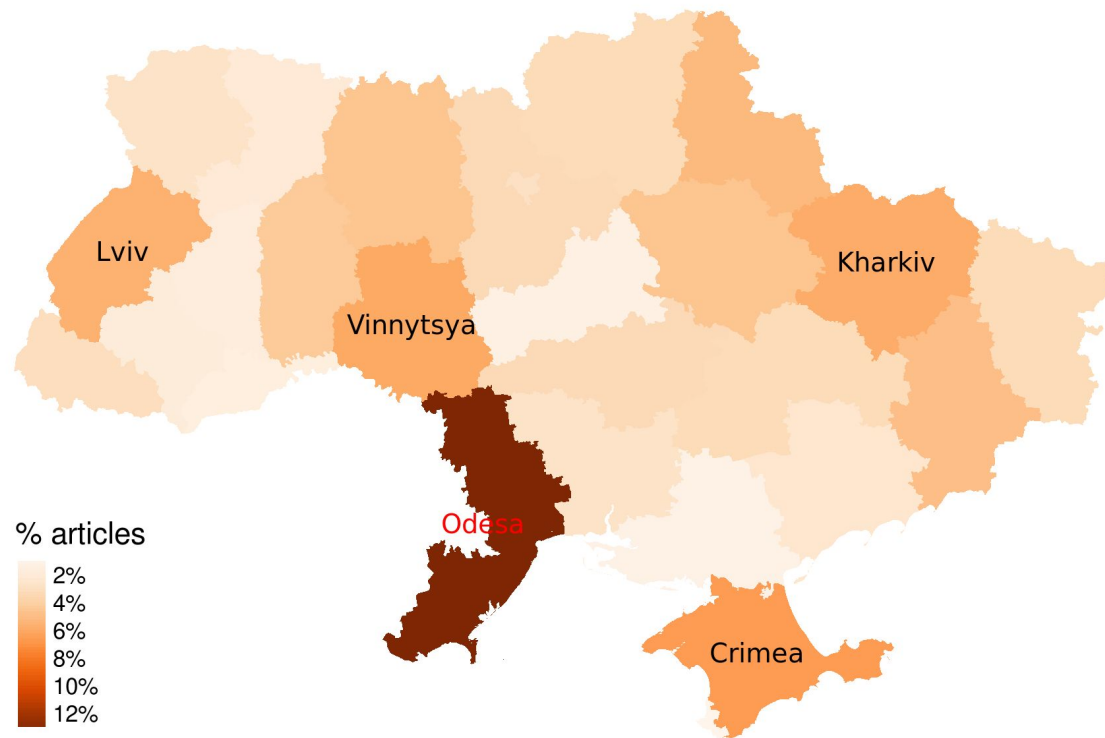


[RO]

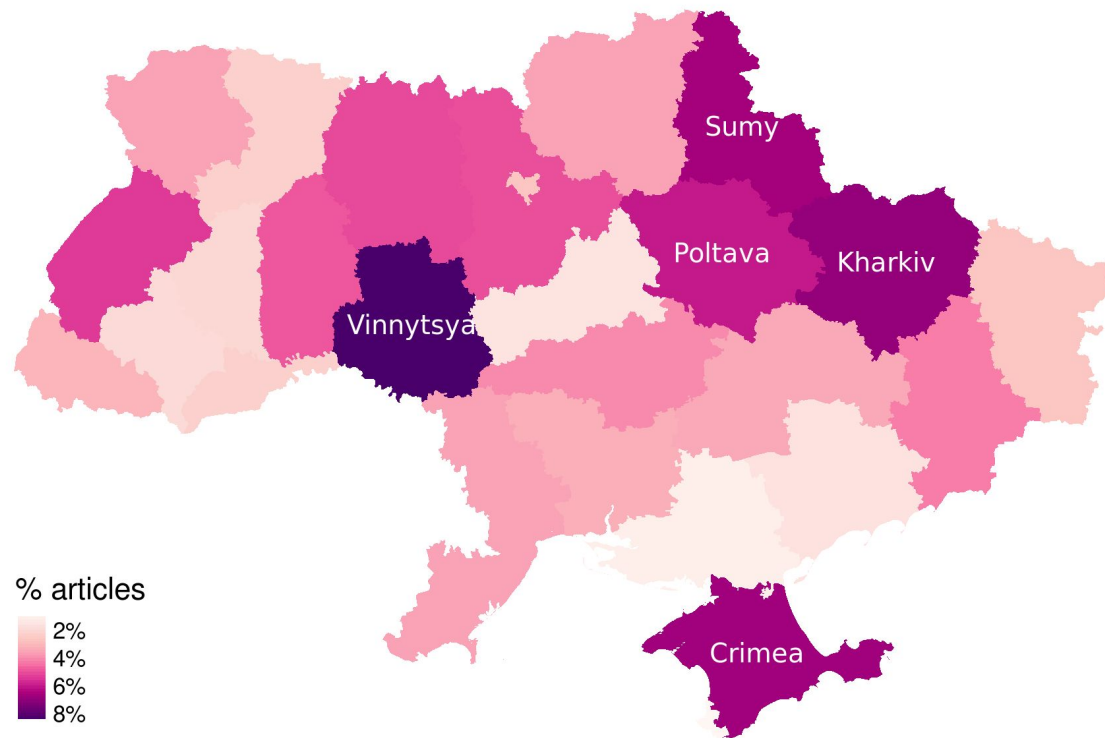
Romanian



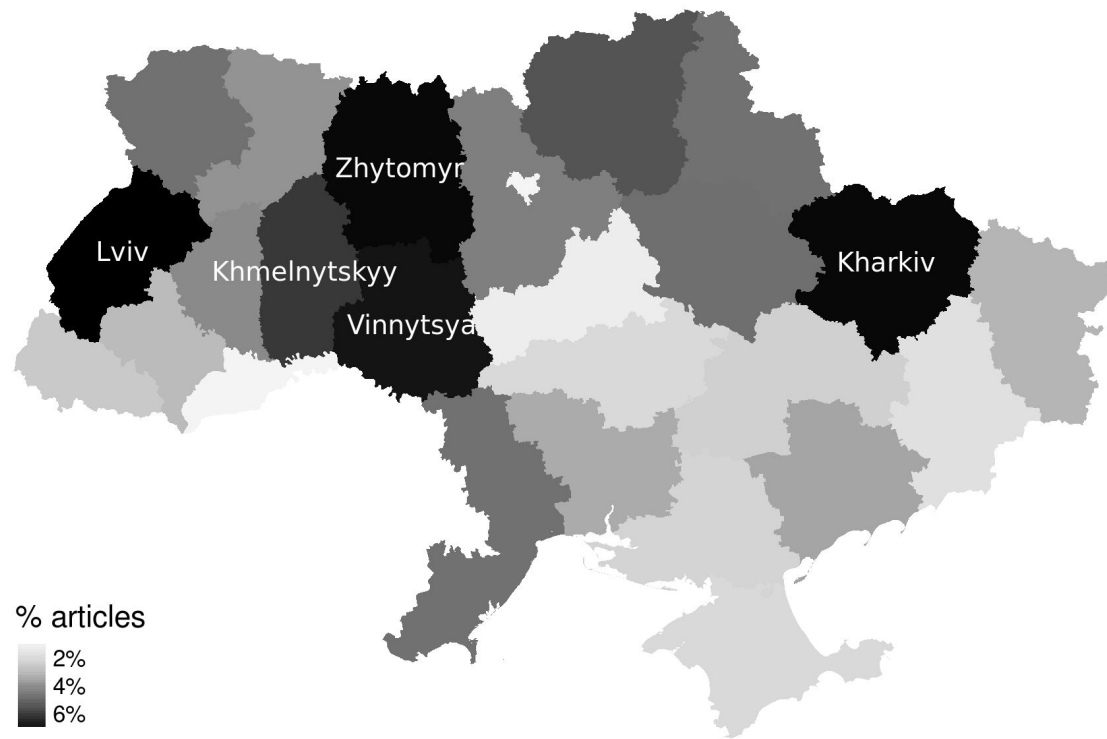
[ES] Spanish



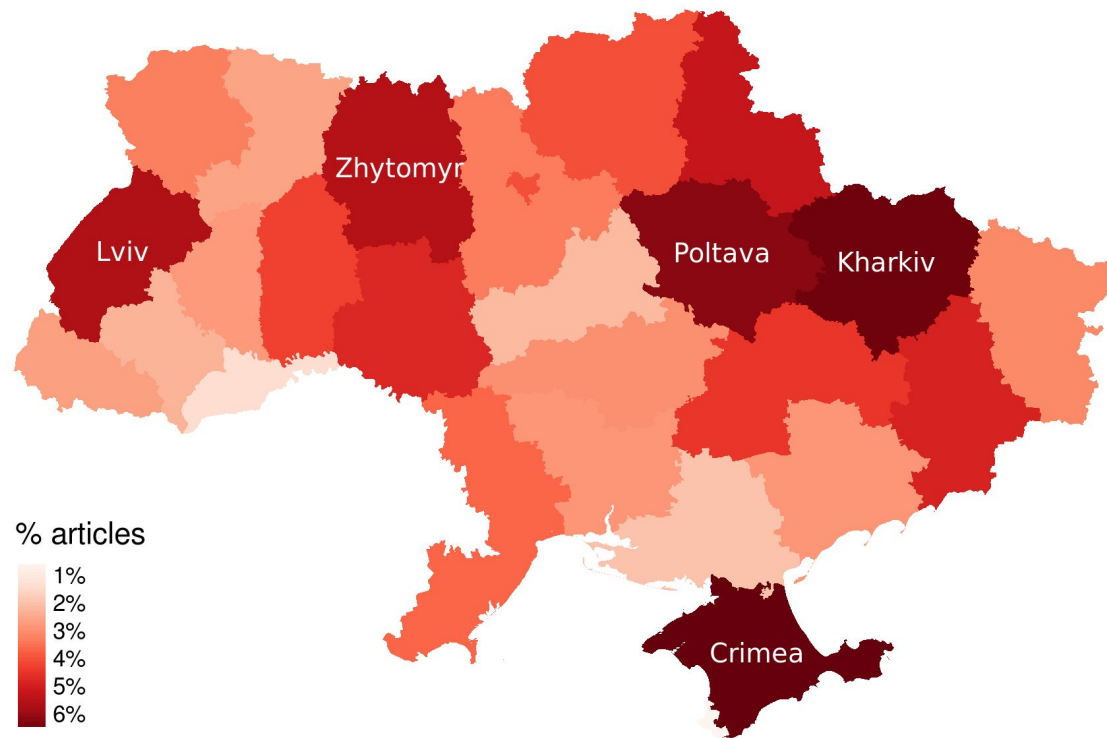
[JA] Japanese



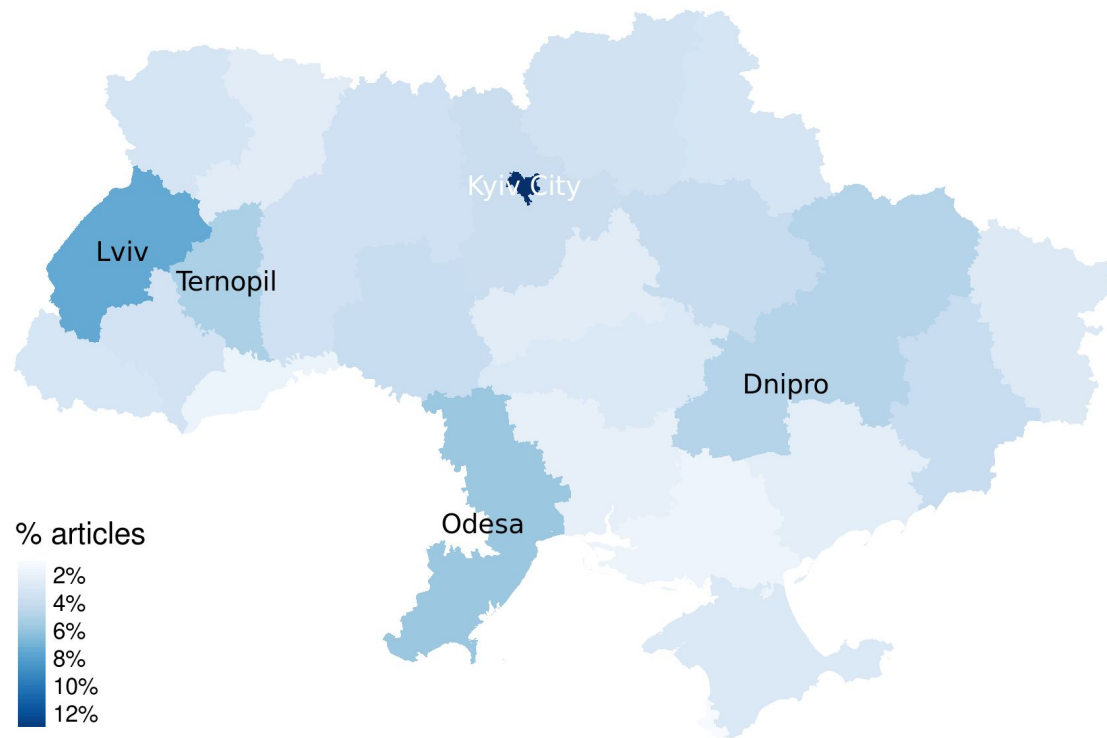
[ZH]
Chinese



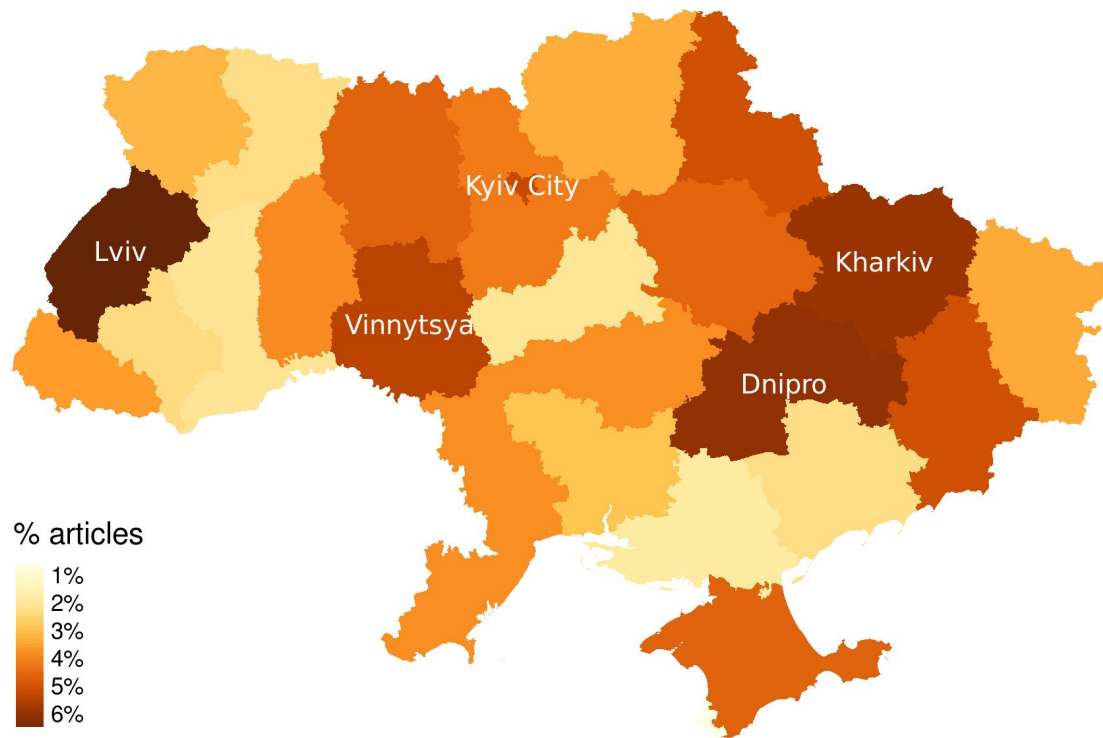
[RU] Russian



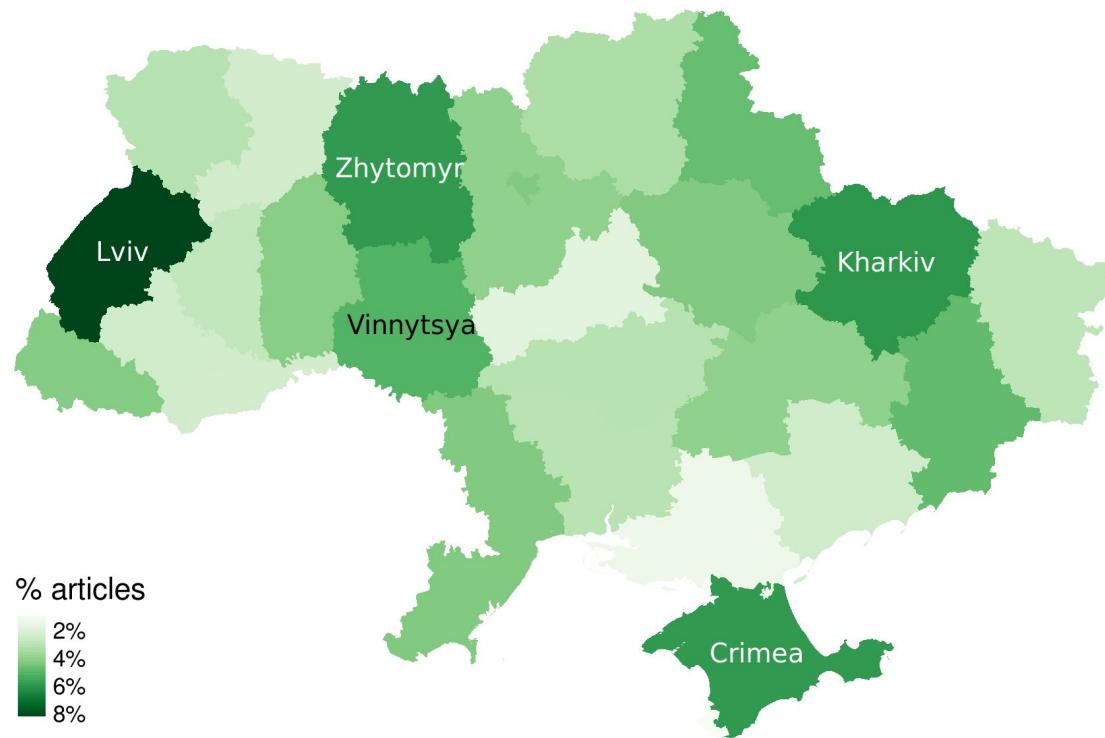
[UK] Ukrainian



[DE] German



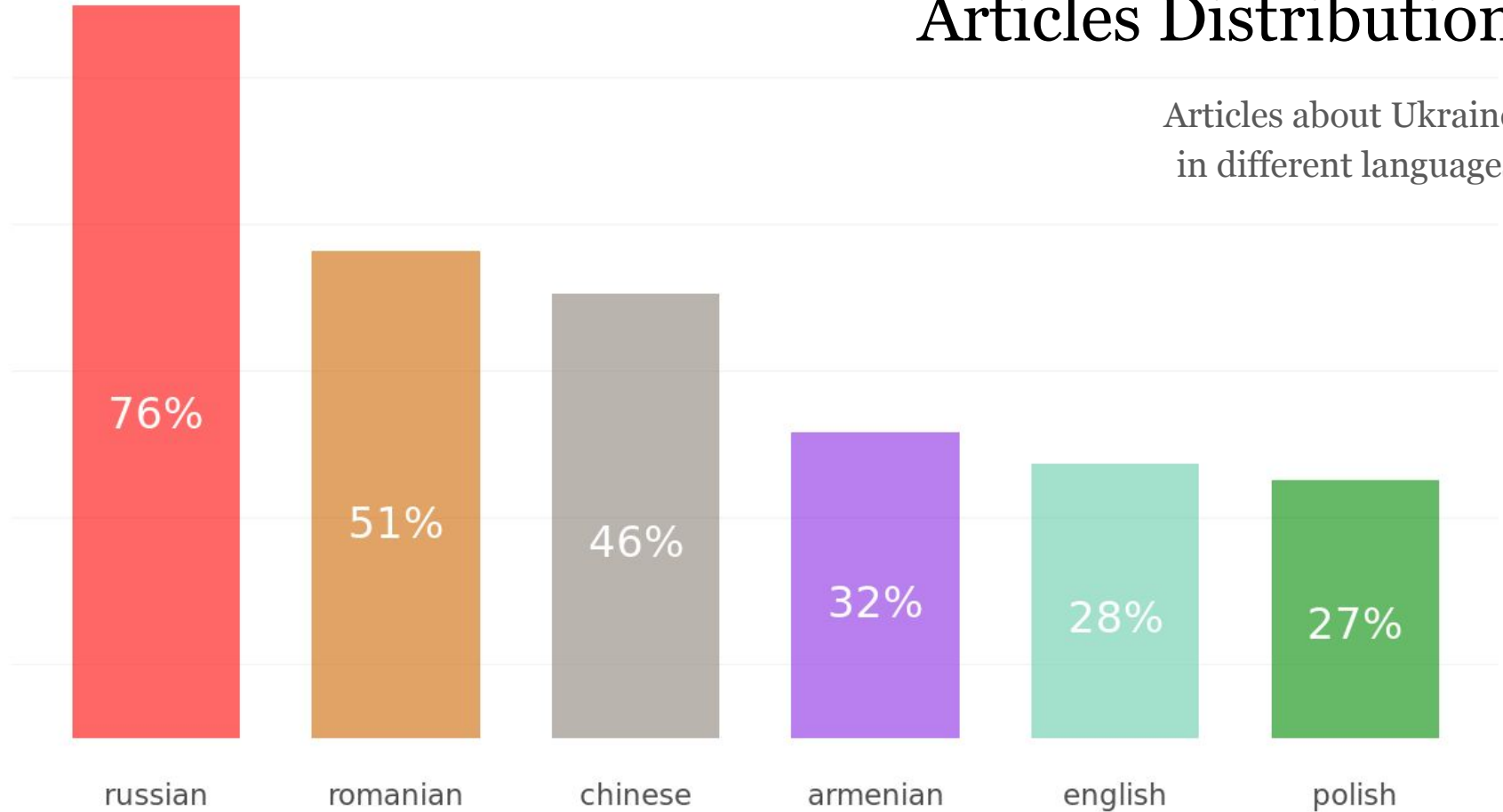
[EN]
English



[Step 5] Language Links

Articles Distribution

Articles about Ukraine
in different languages



Next steps

1. Try to understand what caused the formation of clusters in the space of geotags
2. Fill the missing data with geotags from the *langlinks* dataset
3. Analyse differences in coverage. Which pages are missing from which editions?

Thank **you!**