

*Congress «Humanities vs Sciences & the Knowledge Accelerating in
Modern World : Parallels and Interactions».
Moscow, Russian Federation, 12-16 October 2022*

Countries of the World in the Global Online Information Space: Quantitative Representation, Dynamics, Tonality



*Alexander SHARIKOV,
Professor of the National
Research University Higher
School of Economics, Moscow,
Russian Federation*

Global Online Information Space (GOIS) Research Project

General Research Questions :

- 1. How are different countries represented in GOIS?***
- 2. What are the dynamic patterns of GOIS?***
- 3. What is the mood (positive, negative or neutral) in the information publications ?***

Data base **FACTIVA**

Owner: Dow Jones

Access to: 23000 resources

Number of languages: 26

Number of texts in 2021: over 106 mln



Among resources – web sites of key media:

- **Over 5000 newspapers and journals** – *The New York Times, Wall Street Journal, Financial Times, Le Figaro, Toronto Star, Die Welt, Corriere della Sera, 人民日报 (People's Daily), Times of India, Correio Braziliense, et al.*
- **Over 500 information agencies** - *TASS, 新华社 (Xinhua), Asian News International, Globo, The Associated Press, Reuters, France Presse, Deutsche Presse-Agentur, L'Agenzia Nazionale Stampa Associata, The Canadian Press, 共同通信社 (Kyodo News) et al.*
- **Sites of key television channels and radio stations**
- **Key internet-resources**
- **Social media, blogs, etc.**

Some Characteristics of the Global Internet

The Internet has a **complex structure**. Semantic content is located **mainly in the Web zone**, one of the Internet subsystems.

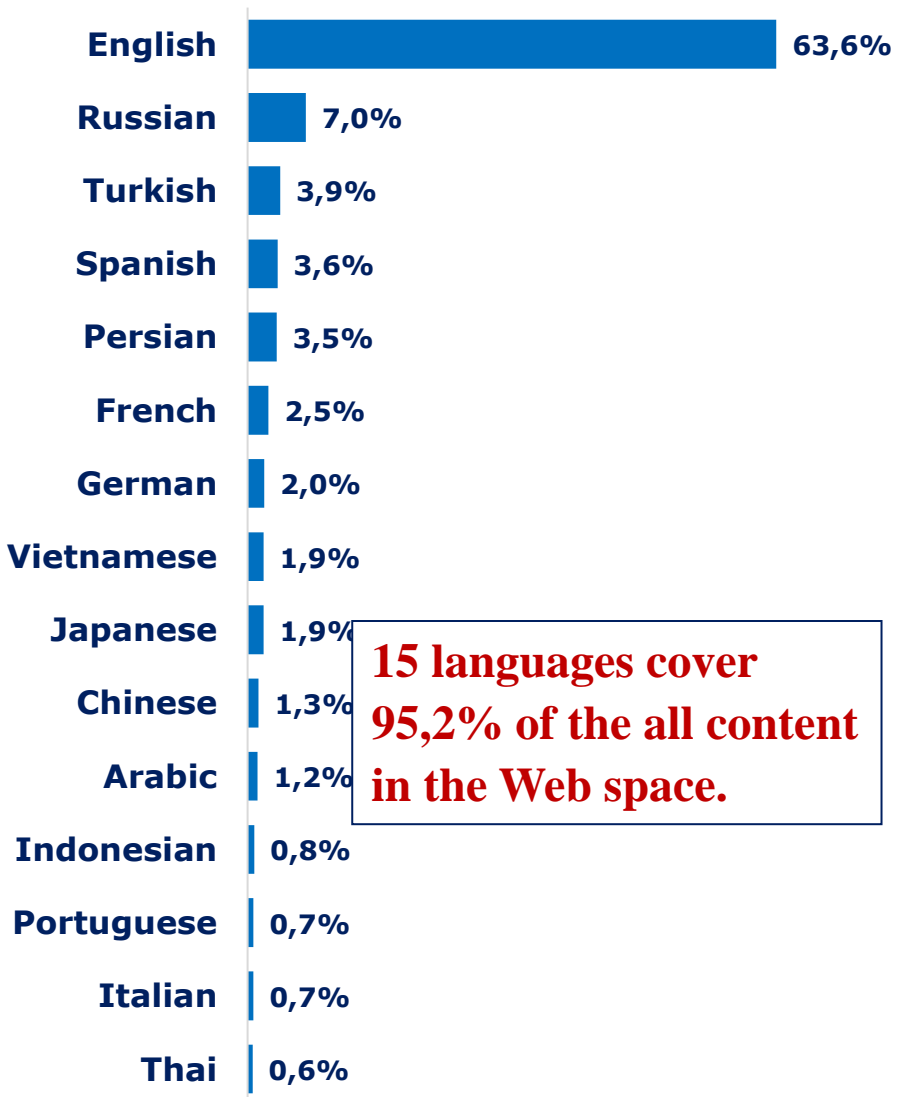
From the user's point of view, structuring goes not so much along geographic and administrative boundaries as in **the languages of use**.

Number of Internet users (penetration) at the 30 June 2022 was estimated at **5,4 billion** (<https://www.internetworldstats.com/stats.htm>).

The total number of websites in 2022 will reach **1.99 billion** (www.internetlivestats.com).

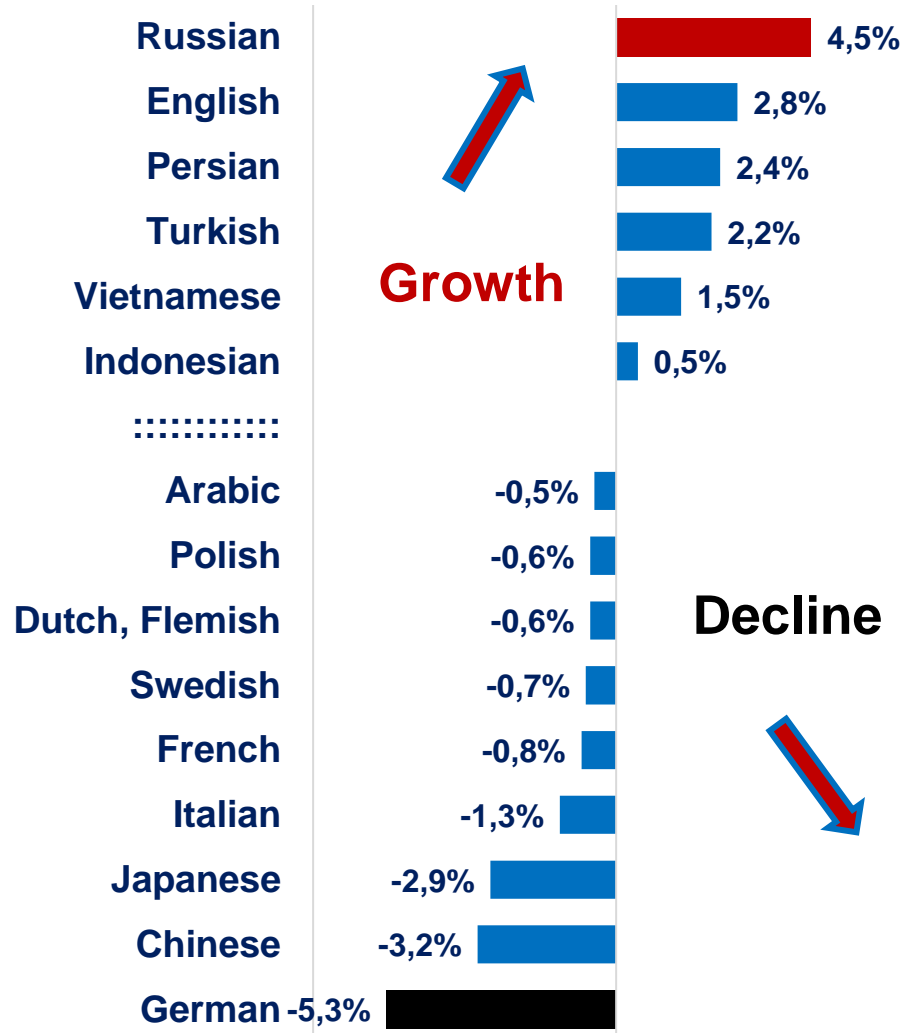
The total number of used content languages in June 2022 was about **180** (<https://w3techs.com>).

Weight of languages (%) in the Internet Web space (as of January 1, 2022)



15 languages cover 95,2% of the all content in the Web space.

The difference in the weights of languages in the Internet Web space between 2021 and 2011



Growth

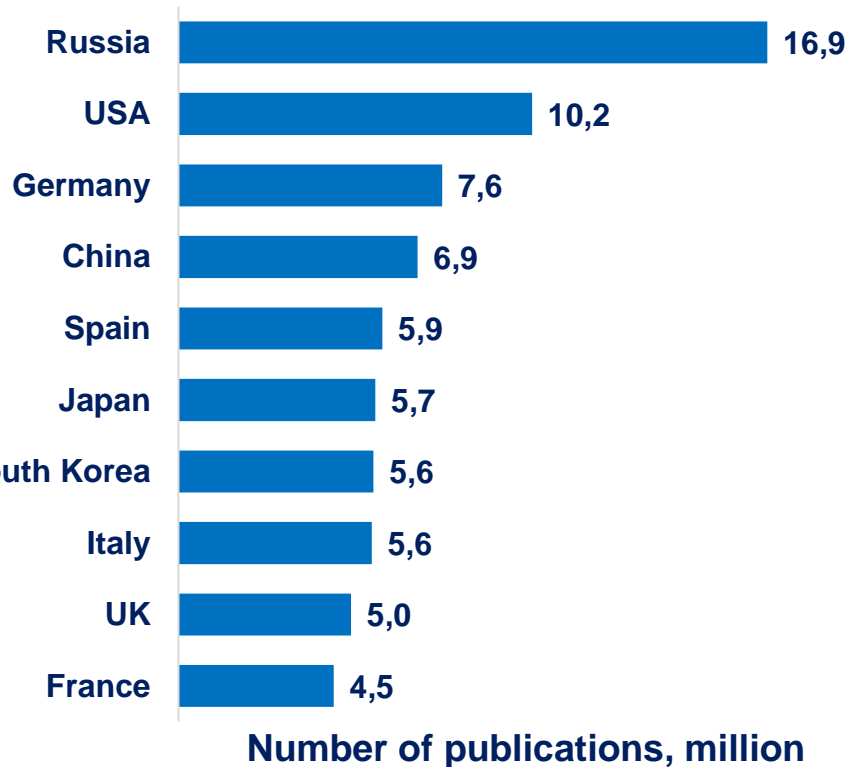
Decline



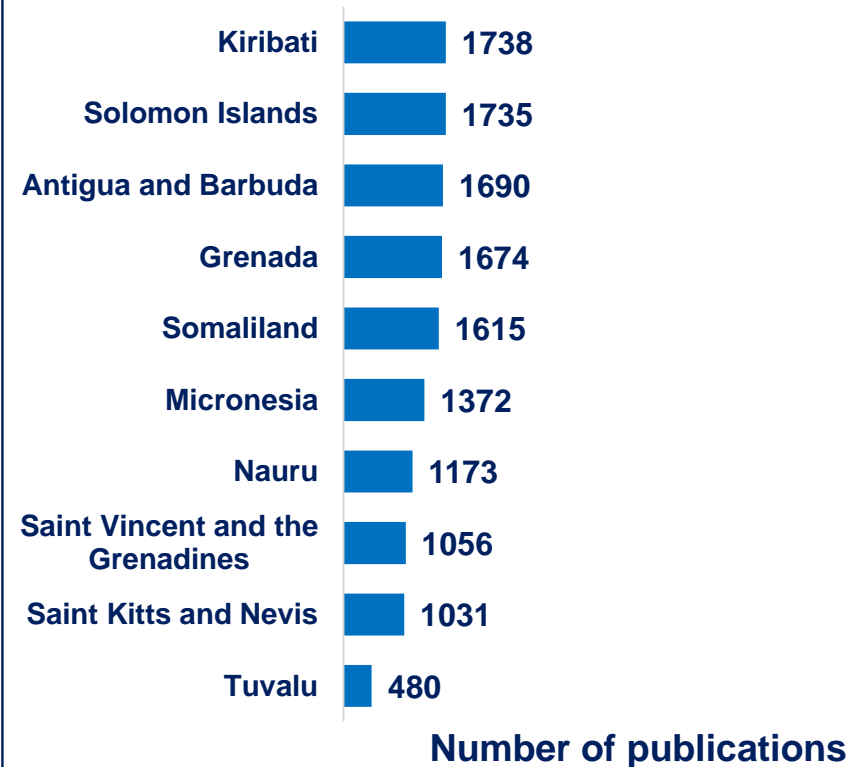
How different countries are represented in GOIS ?

There is a thesis about “visible” and “invisible” countries. With the help of *FACTIVA*, it was possible to establish that there are no “invisible” countries among **204** recognized and partially recognized ones. All of them are “visible”.

The 10 most mentioned countries in 2020 according *FACTIVA* (25 languages in total)



The 10 least mentioned countries in 2020 according *FACTIVA* (25 languages in total)

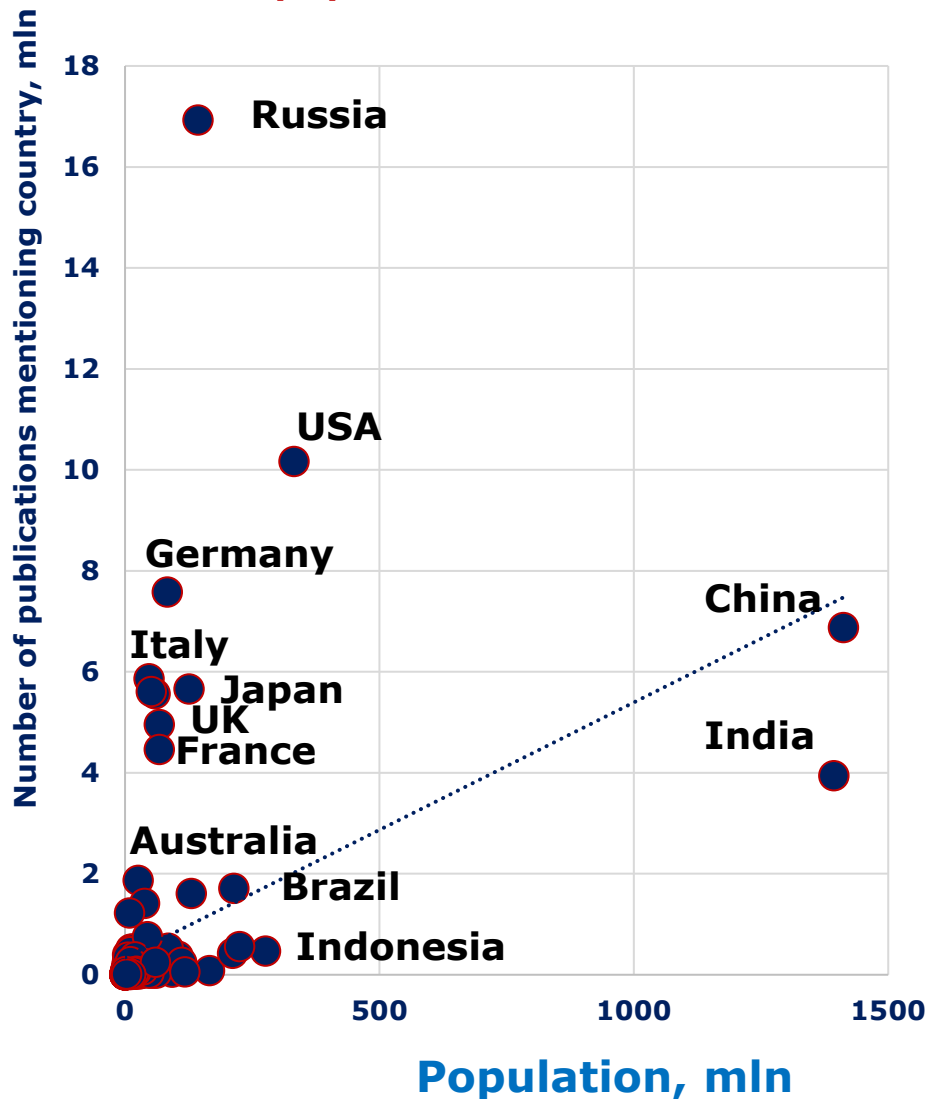


The most frequently mentioned countries within the FACTIVA data base in 2021 and in February-May 2022 (in total in 26 languages)

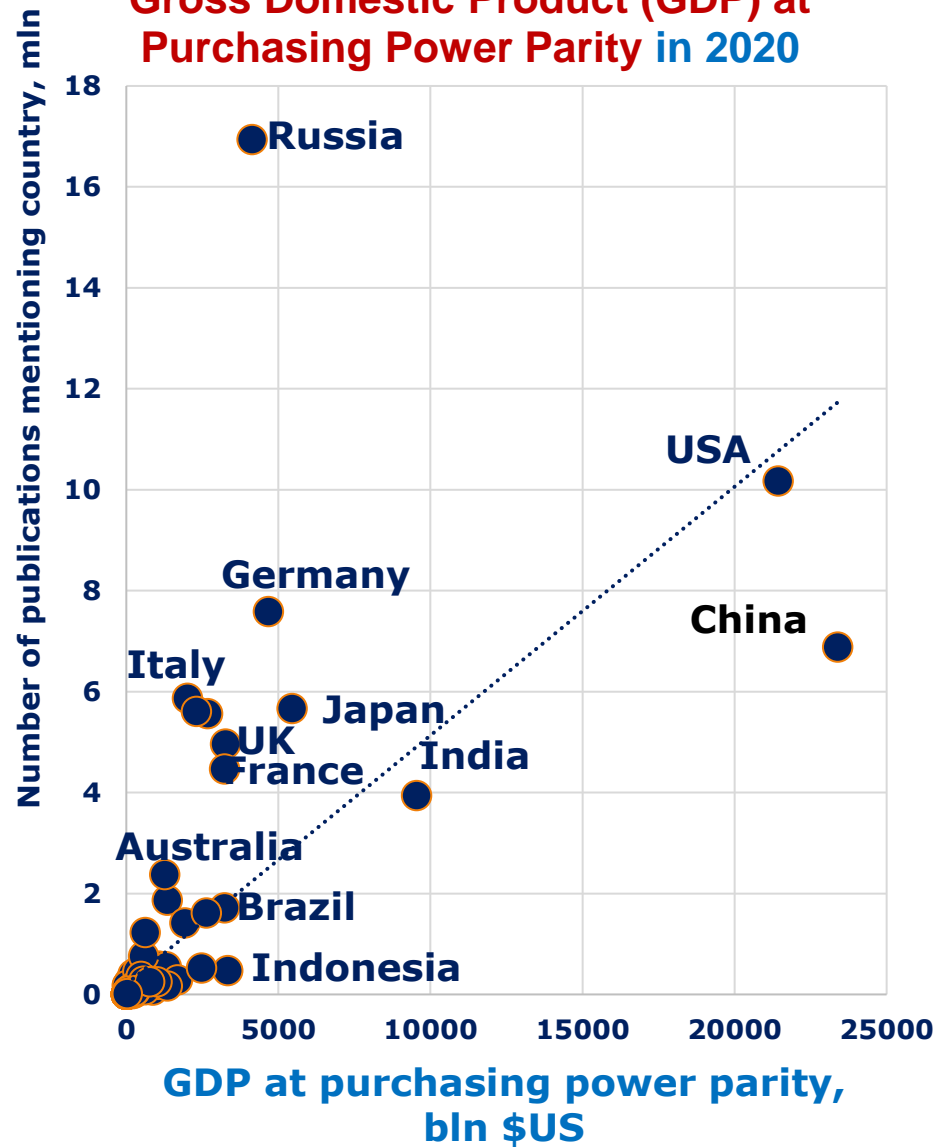
2021			February-May 2022		
	Country	Number of materials mentioning the country, %		Country	Number of materials mentioning the country, %
1	Russia	14,63	1	Russia	19,24
2	USA	8,79	2	USA	8,61
3	Germany	6,31	3	Ukraine	6,55
4	Spain	5,12	4	Germany	6,39
5	China	5,09	5	China	5,74
6	Italy	5,01	6	Spain	5,25
7	South Korea	4,40	7	Italy	4,66
8	Japan	3,88	8	South Korea	3,93
9	India	3,86	9	India	3,75
10	UK	3,81	10	UK	3,41

Sources: *FACTIVA, UN*

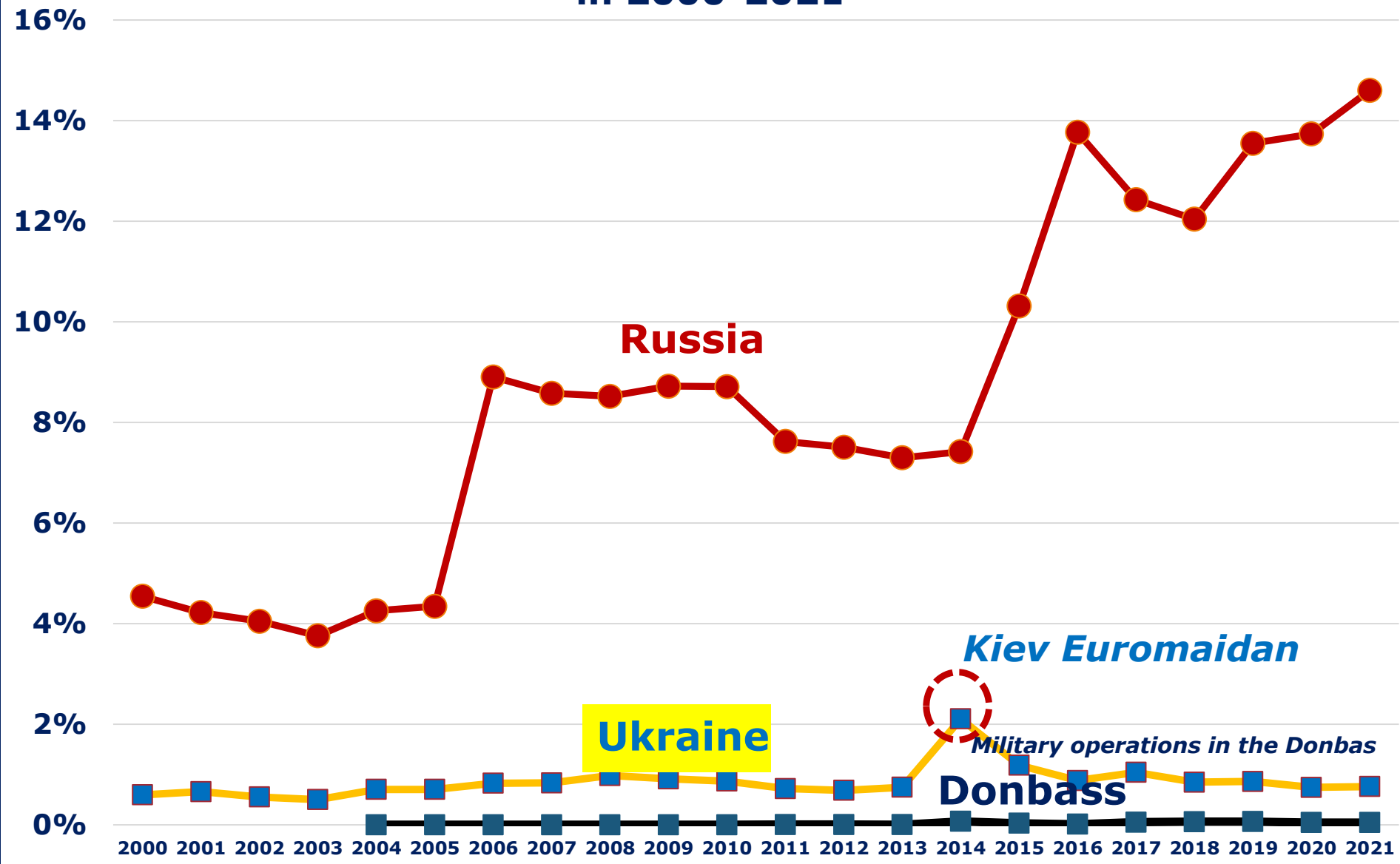
Dependence of the number of publications mentioning countries in the FACTIVA database on their **population in 2020**



Dependence of the number of publications mentioning countries in the FACTIVA database on their **Gross Domestic Product (GDP) at Purchasing Power Parity in 2020**



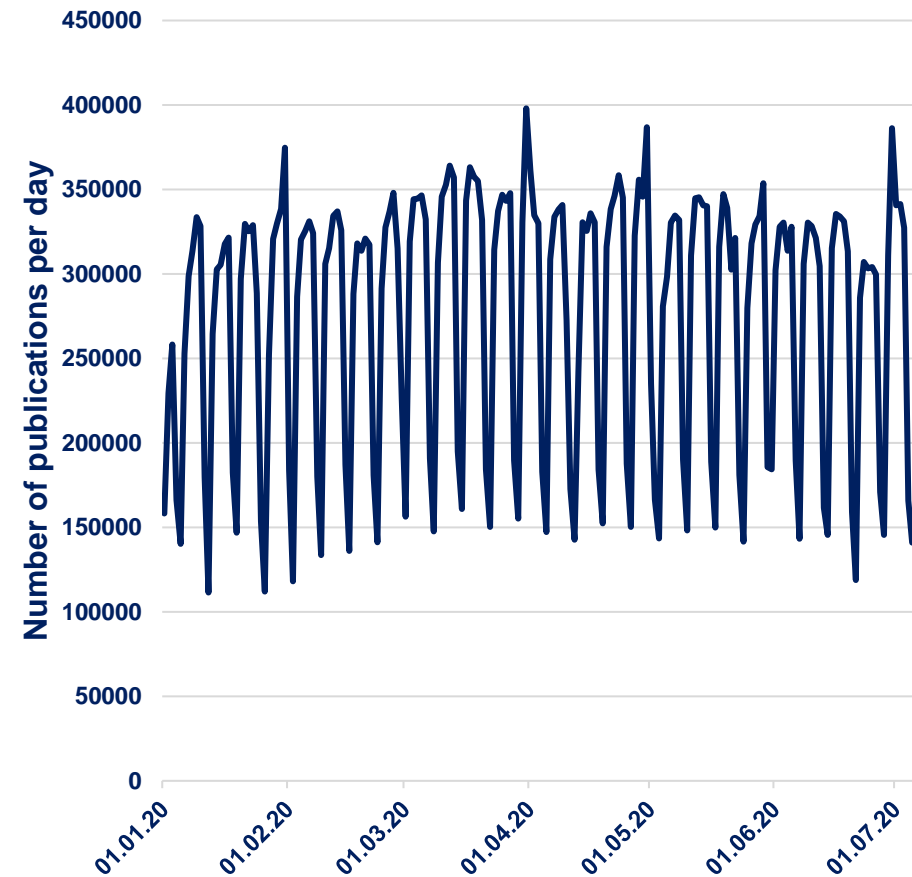
Dynamics of the percentage of materials mentioning Russia, Ukraine & Donbass within the Factiva data base in 2000-2021



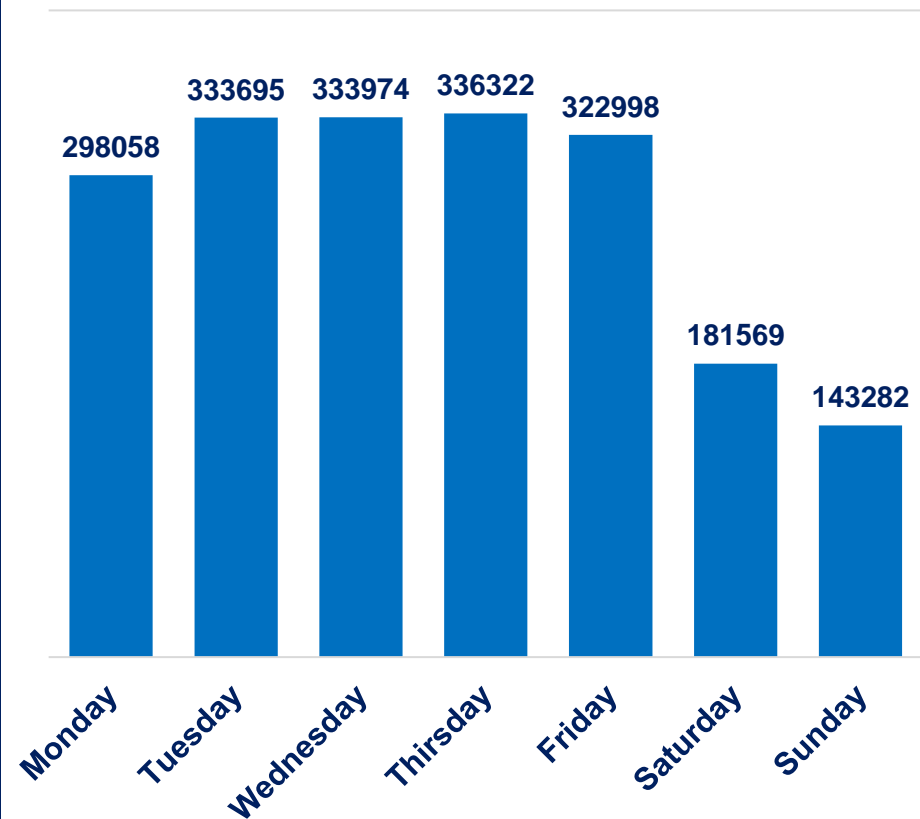
What are the dynamic patterns of GOIS?

Another objective process is related to the regularity of the appearance of publications. Weekly cyclicality detected.

Dynamics of the daily number of new publications on the resource base of the FACTIVE system in the first half of 2020



Average number of publications that formed the resource basis of the FACTIVE system in January-July 2020 by day of the week

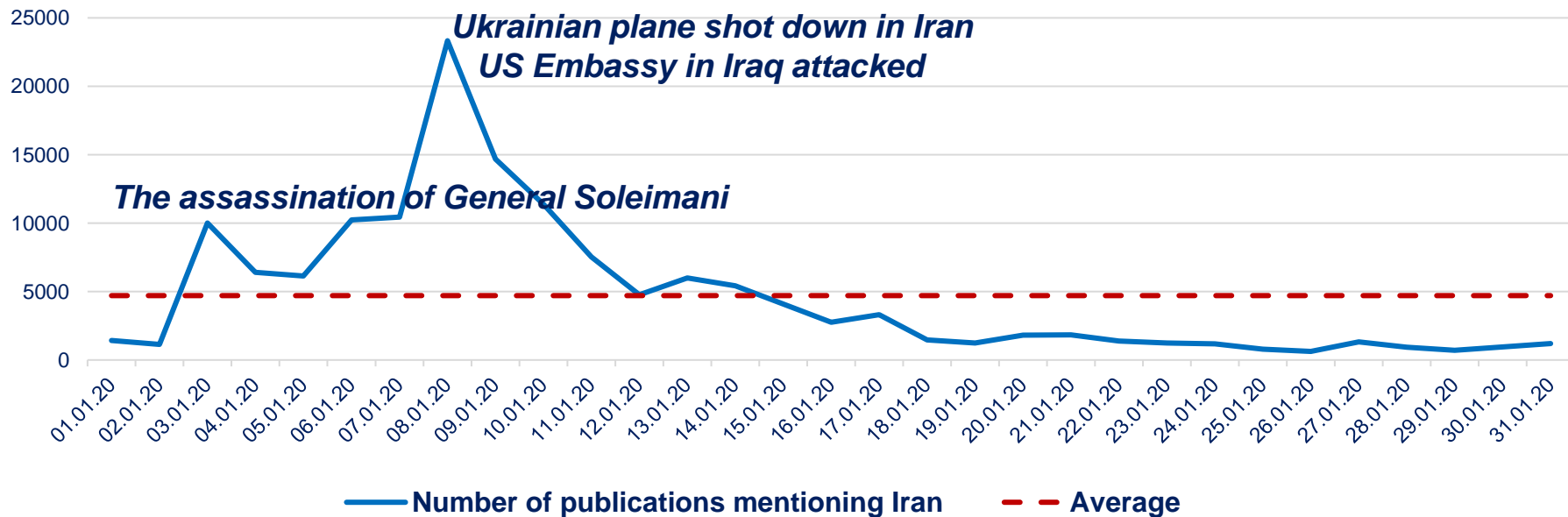


What are the dynamic patterns of GOIS?

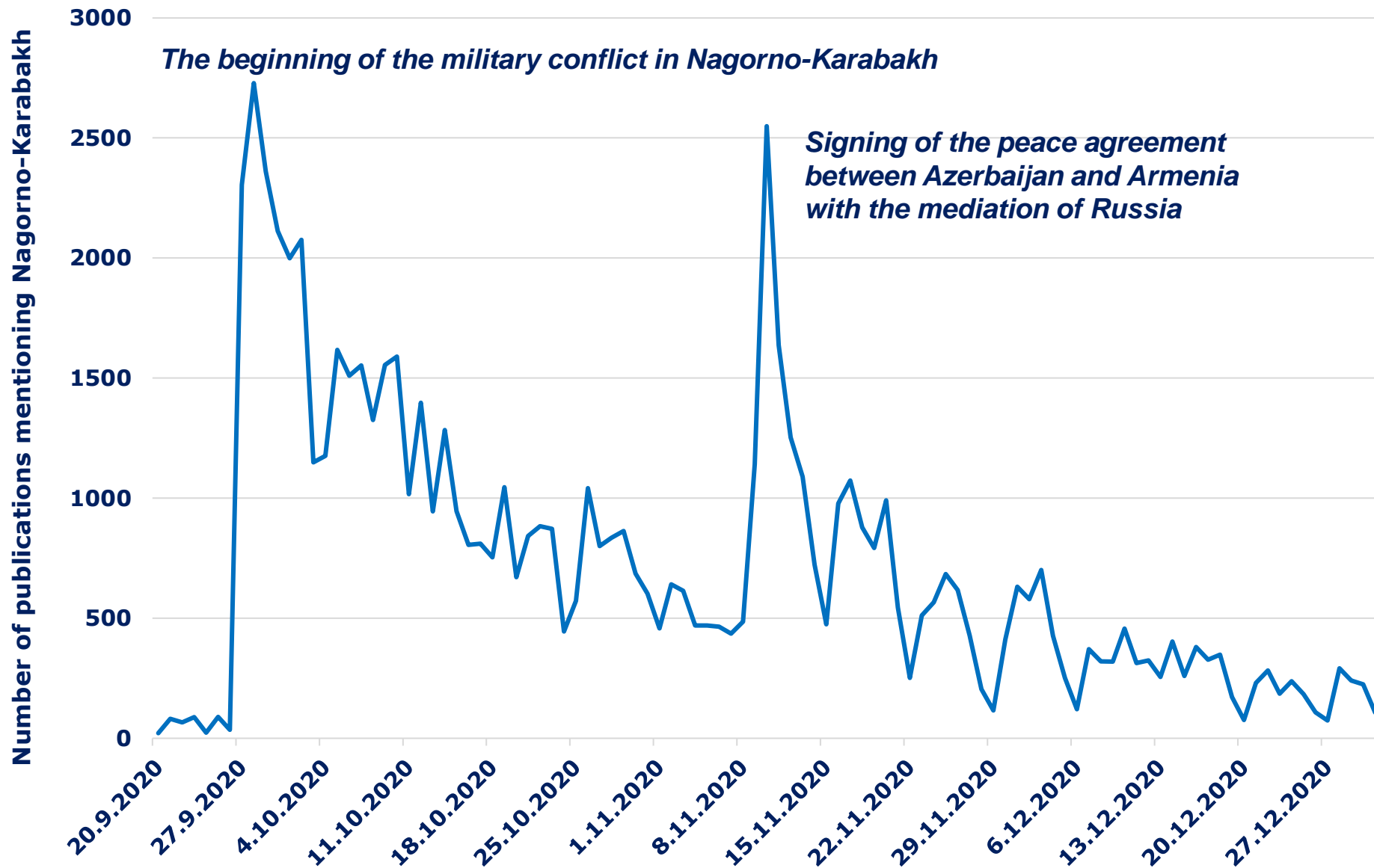
Subjective processes are related to editorial policy of each mass-media.

From time to time, there are **media splashes** when a topic is picked up by other media and there is a rapid increase in the number of publications on this topic. Sometimes a media splash turns into a **media storm**, which can last quite a long time.

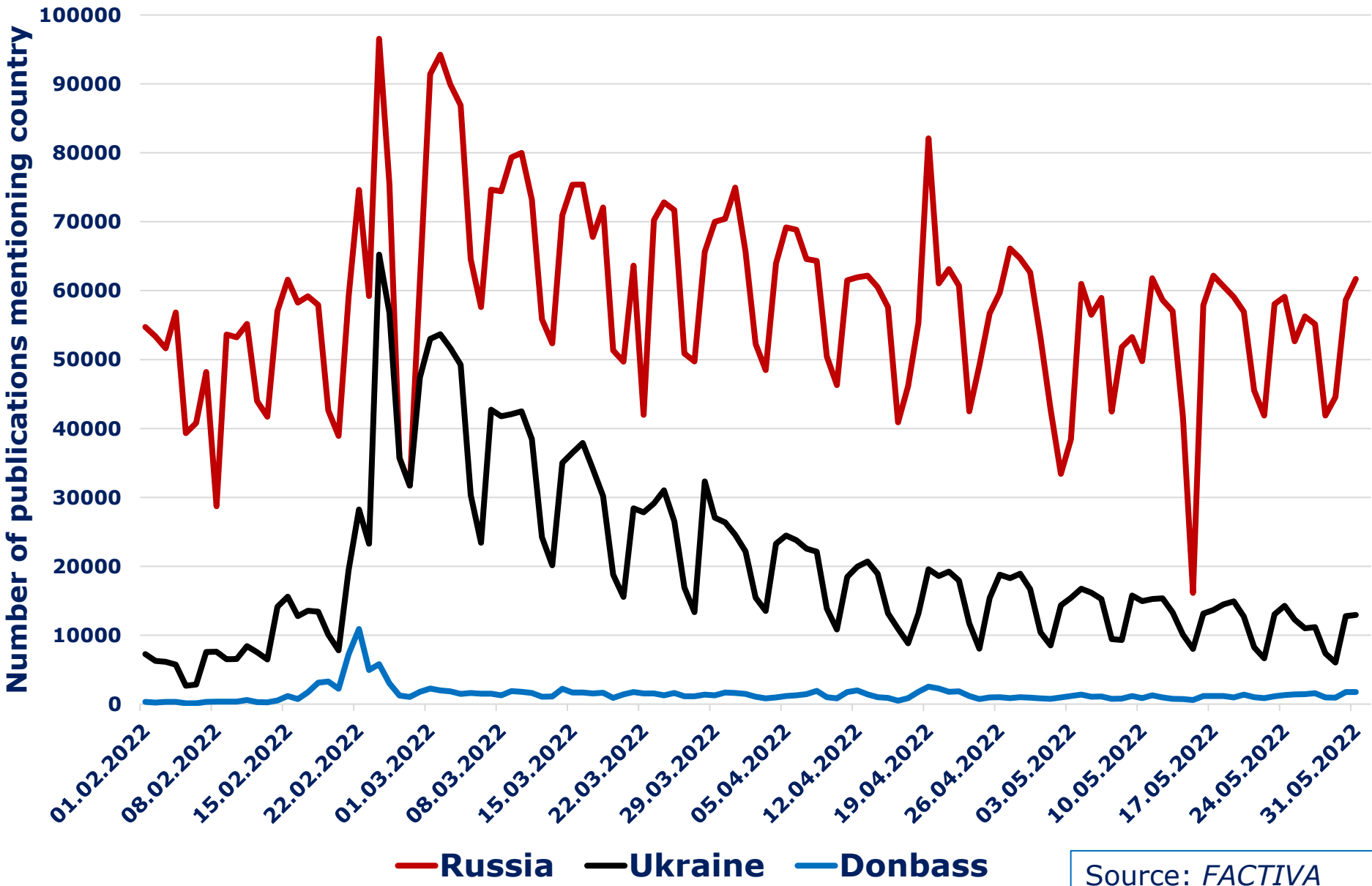
The number of publications with the mention of Iran by day in January 2020 in total in 25 languages of the FACTIVA system



The number of materials mentioning Nagorno-Karabakh daily, from September 20 to December 31, 2020, in a total of 25 languages



The number of publications mentioning Russia, Ukraine and Donbass from February 1 to May 31, 2022 by day in a total of 26 languages in the "FACTIVA" system



— Russia — Ukraine — Donbass

Source: *FACTIVA*

The strongest media storms in the post-Soviet countries in 2020 (on a monthly basis)

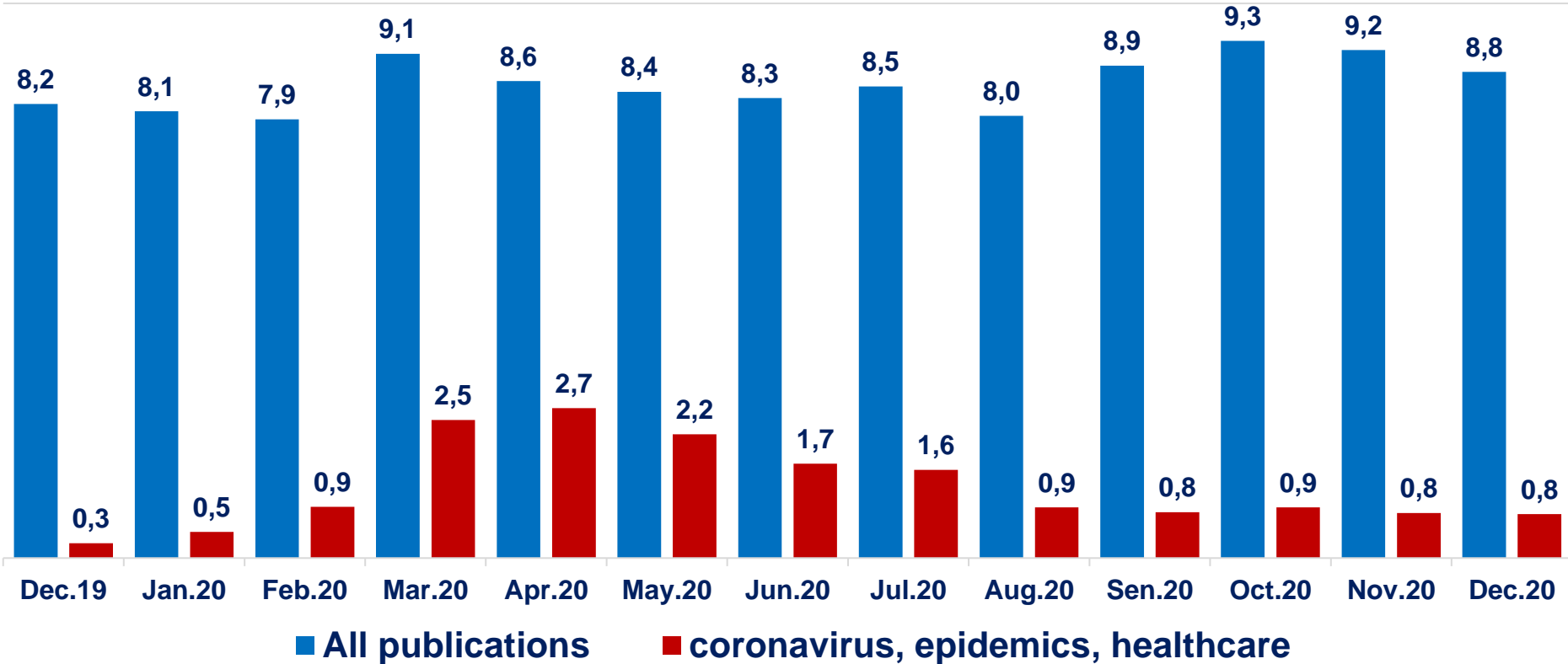
	Maximum monthly growth, %	Event
Nagorno-Karabakh	939%	The military conflict in Nagorno-Karabakh
Belarus	364%	Presidential elections and opposition protests
Abkhazia	302%	Resignation of the President of Abkhazia
Moldova	185%	Presidential elections and opposition protests
Armenia	153%	The military conflict in Nagorno-Karabakh
Kyrgyzstan	137%	Parliamentary elections, opposition protests, resignation of the President of Kyrgyzstan
Azerbaijan	120%	The military conflict in Nagorno-Karabakh
South Ossetia	96%	Assassination attempt on the Interior Minister of South Ossetia and his investigation
Donbass	46%	A truce was reached in the military confrontation with Ukraine
Russia	45%	Coronavirus epidemic

Source: *FACTIVA*

What are the dynamic patterns of GIOS?

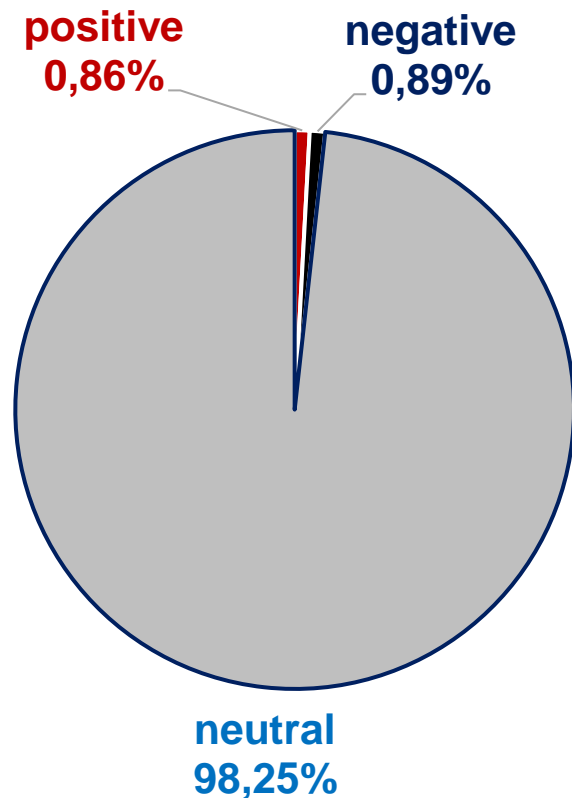
But there are situations when a **media storm** unfolds over a long period of time, but is not associated with media splashes.

Dynamics of the number of materials (in millions) that made up the resource base of the FACTIVE system for the months from December 2019 to December 2020, and the number of materials about the coronavirus epidemic

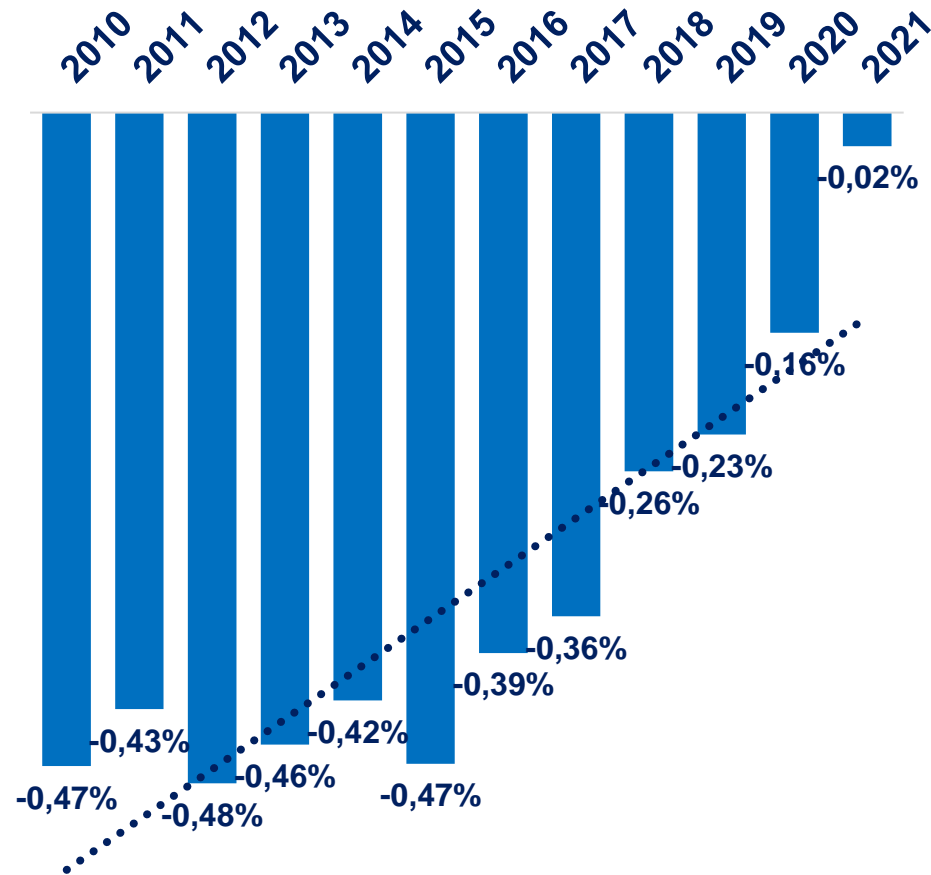


What is the mood (positive, negative or neutral) in information publications ?

The ratio between positive, negative and neutral content in 2021 in 9 languages in total



Dynamics of the tonal balance in 2010-2021 in 9 languages in total

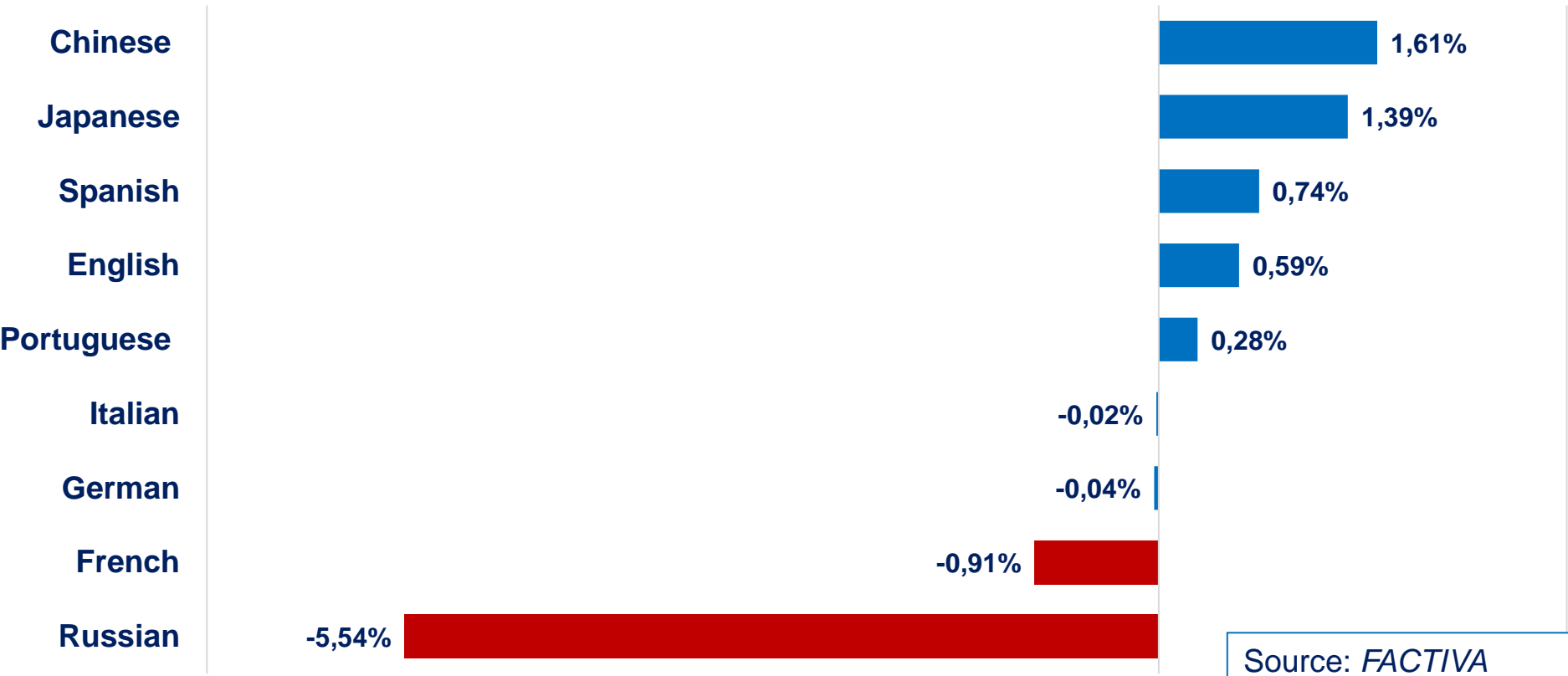


Source: FACTIVA

Does the ratio between positive and negative tonality in information materials differ in different language zones ?

Let's call the **tonality balance** expression $Bt = ((P - N)/A) * 100\%$, where *N* is the number of negative, *P* is the number of positive, and *A* is the number of all materials.

Tonality balance of the publications in 9 language zones in 2020



Source: *FACTIVA*

Tonality balance distribution by country in 2020

The smallest tonal balance

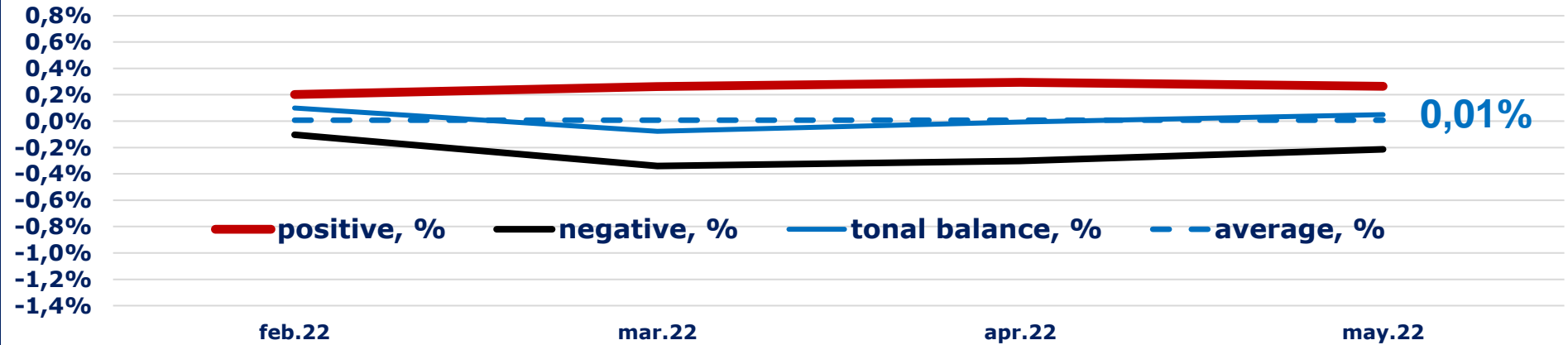
<i>Ukraine</i>	<i>-9,8%</i>
<i>Armenia</i>	<i>-11,2%</i>
<i>Nagorno-Karabakh</i>	<i>-13,8%</i>
<i>South Ossetia</i>	<i>-16,1%</i>
<i>Donbass</i>	<i>-20,5%</i>

The greatest tonal balance

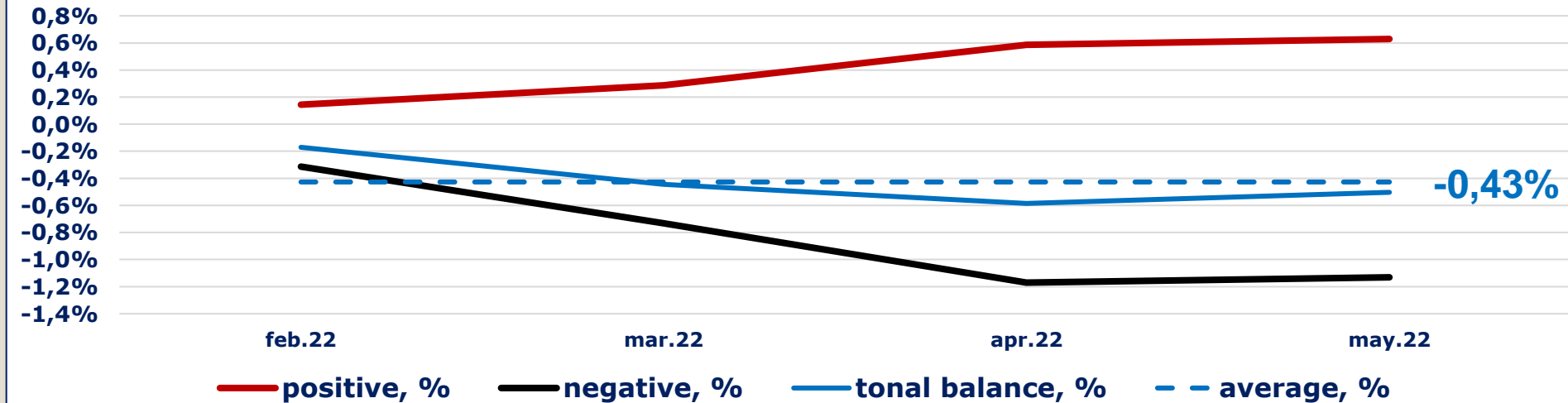
<i>Vietnam</i>	<i>1,5%</i>
<i>Nauru</i>	<i>1,3%</i>
<i>Japan</i>	<i>1,2%</i>
<i>Thailand</i>	<i>1,2%</i>
<i>China</i>	<i>1,0%</i>

Dynamics of tonal characteristics of publications mentioning Russia and Ukraine in February-May 2022 in 9 languages in the FACTIVA database

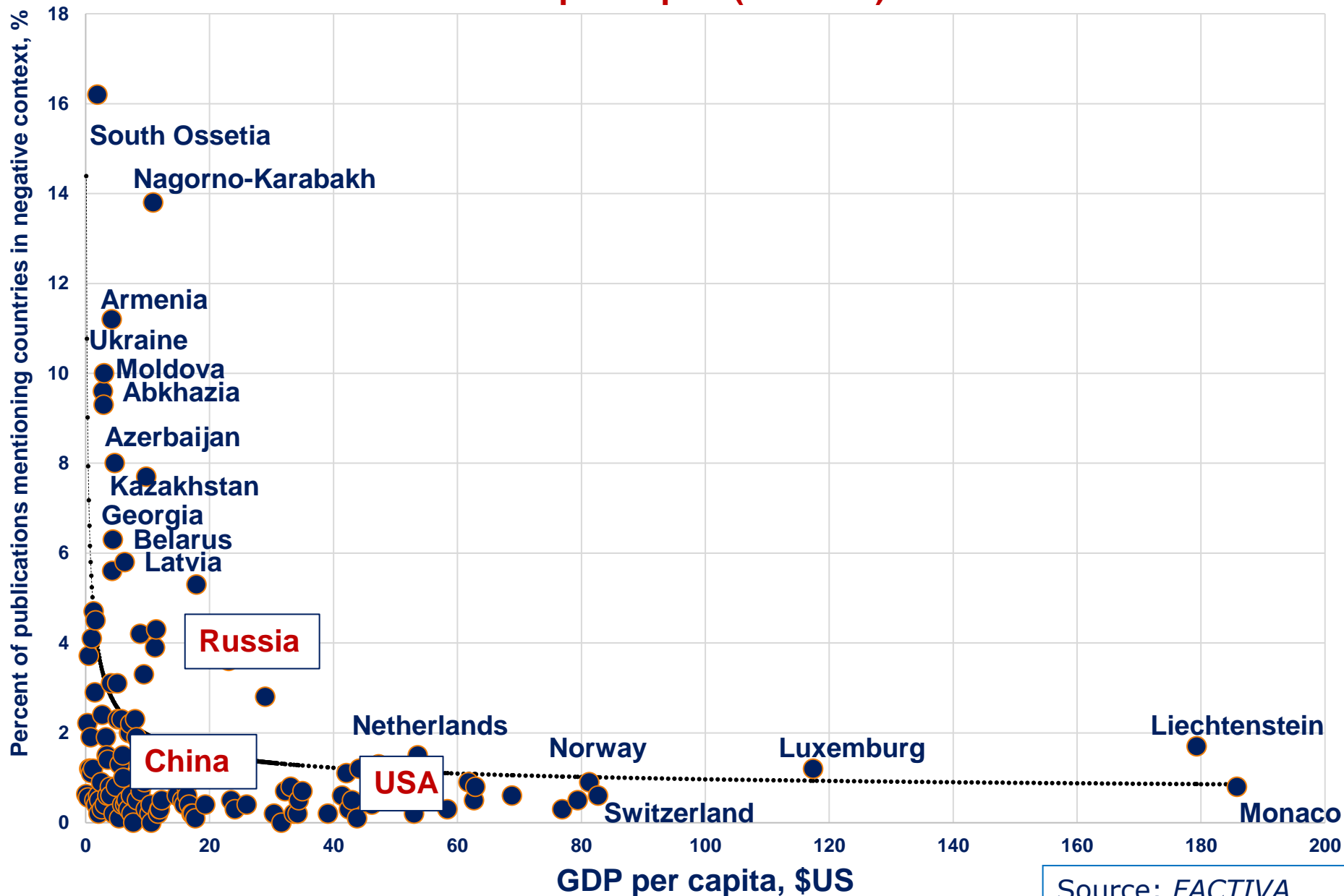
Russia



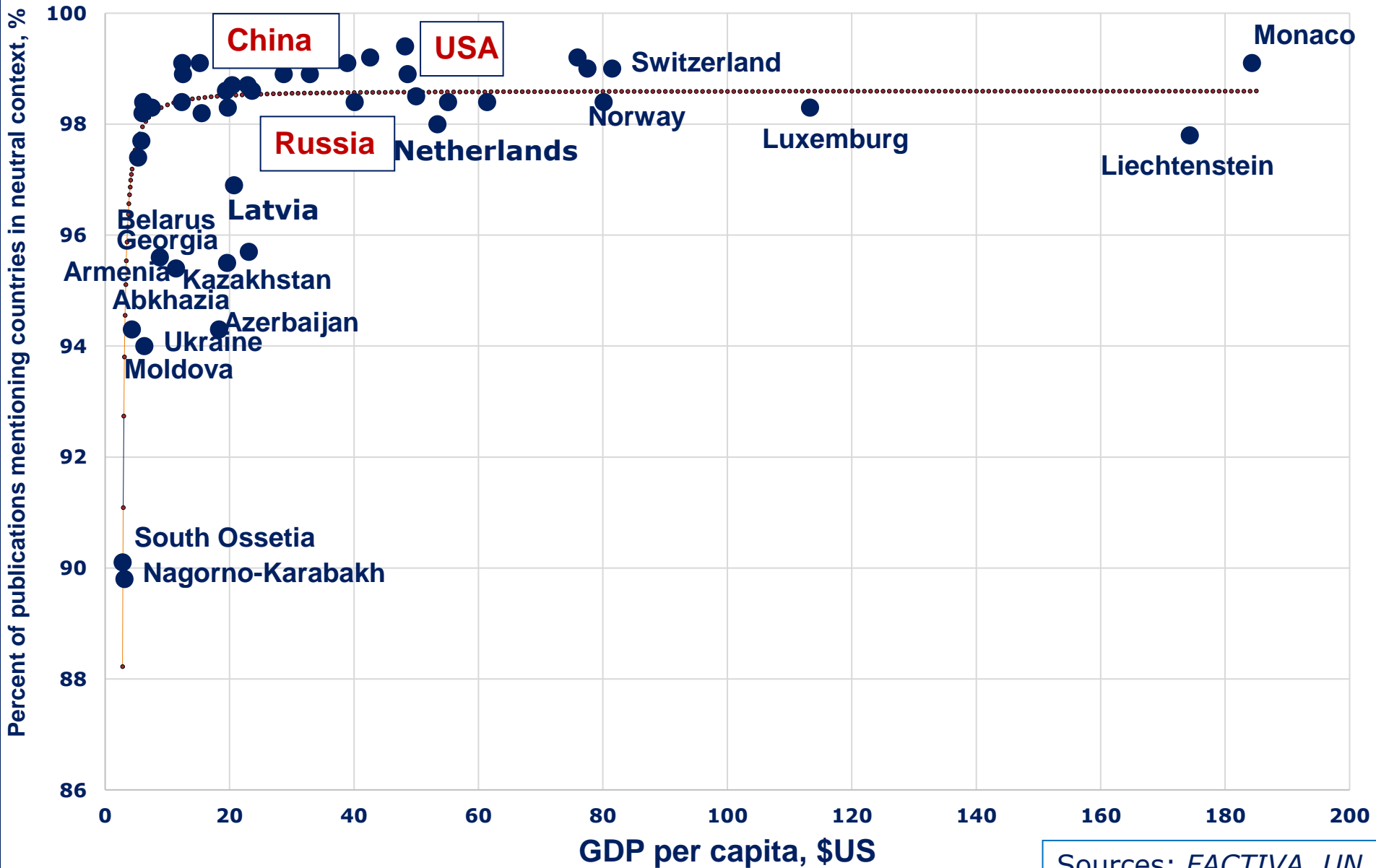
Ukraine



The dependence of the percent of negative materials mentioning countries in 9 languages (FACTIVA system data) and their GDP per capita (UN data) in 2020



The dependence of the percent of neutral materials mentioning countries in 9 languages (FACTIVA system data) and their GDP per capita in 2020



Sources: *FACTIVA*, *UN*

Thank you!