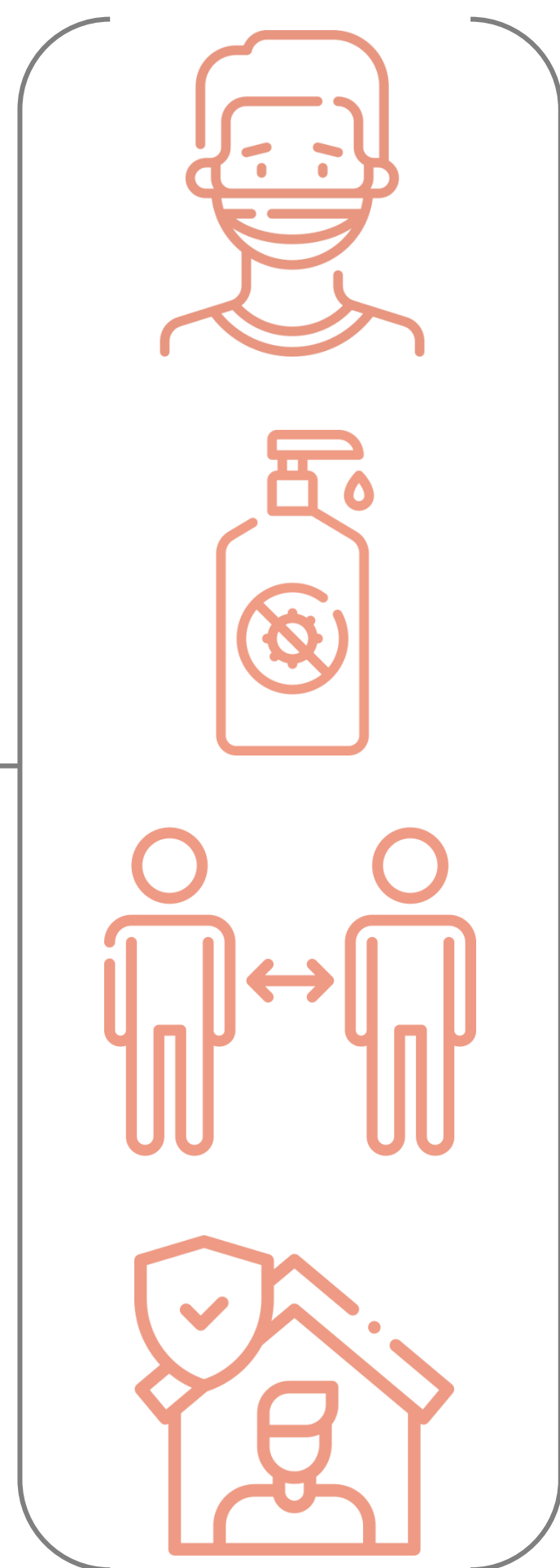# Social media insights about COVID-19 in Portugal: a text mining approach

Presented by Joao C Ferreira (ISCTE and INOV at Portugal) in CONASENSE 2021 SYMPOSIUM

Work of a Master thesis of Carolina Ferraz Marreiros supervised by Joao C Ferreira and Ricardo Ribeiro

Motivation

COVID-19
spread

Measures
taken

Social and
economic impact

Sharing on
social media

2

istar _iscte

Main Objectives

The purpose of this dissertation is to understand a perception of reality shared on social media, that may also be useful to support decision making within COVID-19 pandemic.

- **What are the main topics emerging in the social media about the Pandemic context in Portugal?**
- **What themes were identified with the most positive to most negative sentiment?**
- **It is possible to identify parallels between the events that took place in Portugal, in the context of COVID-19, and the sentiments expressed in the social networks?**

Methodology
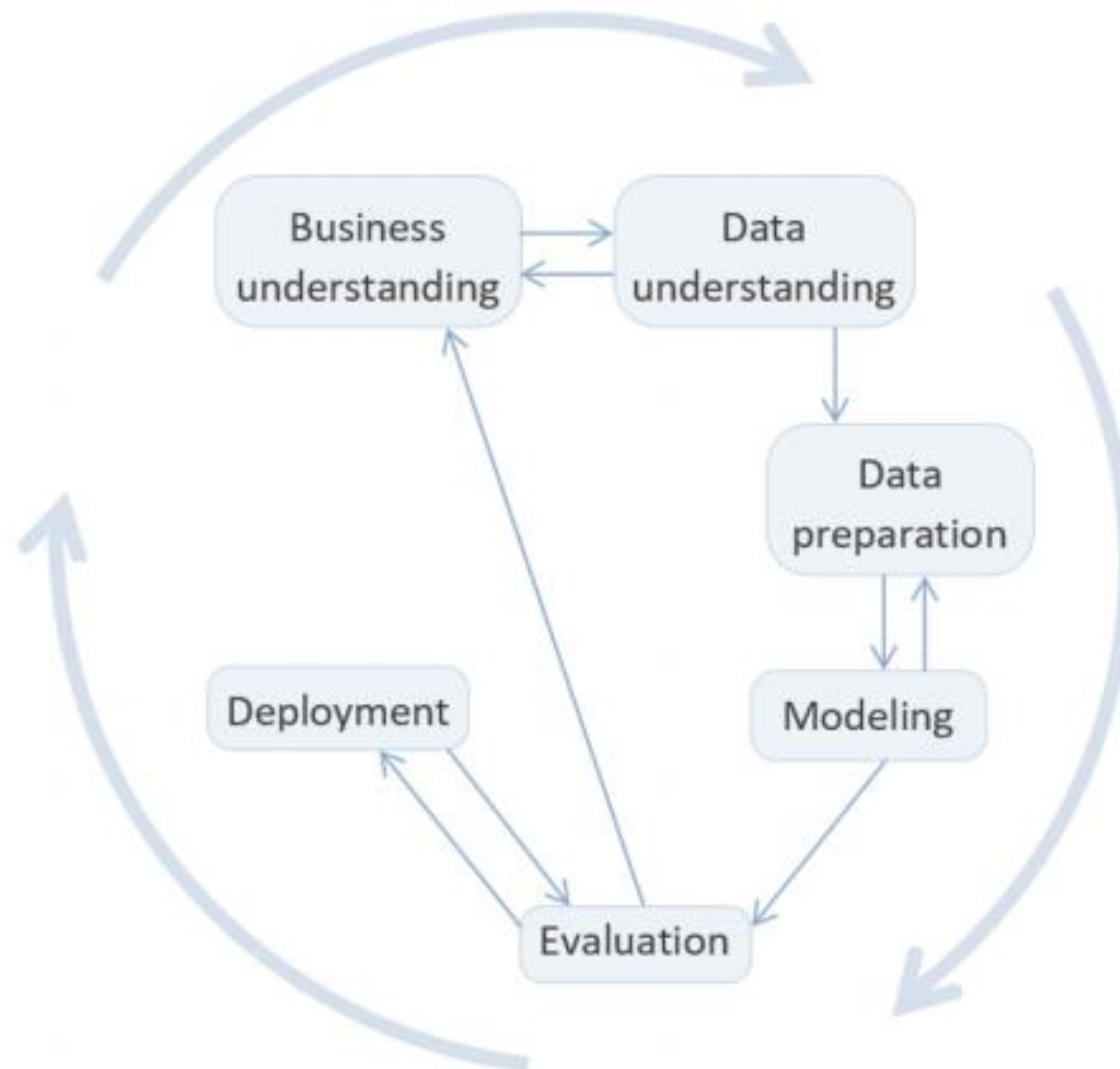
**CRISP-DM methodology**
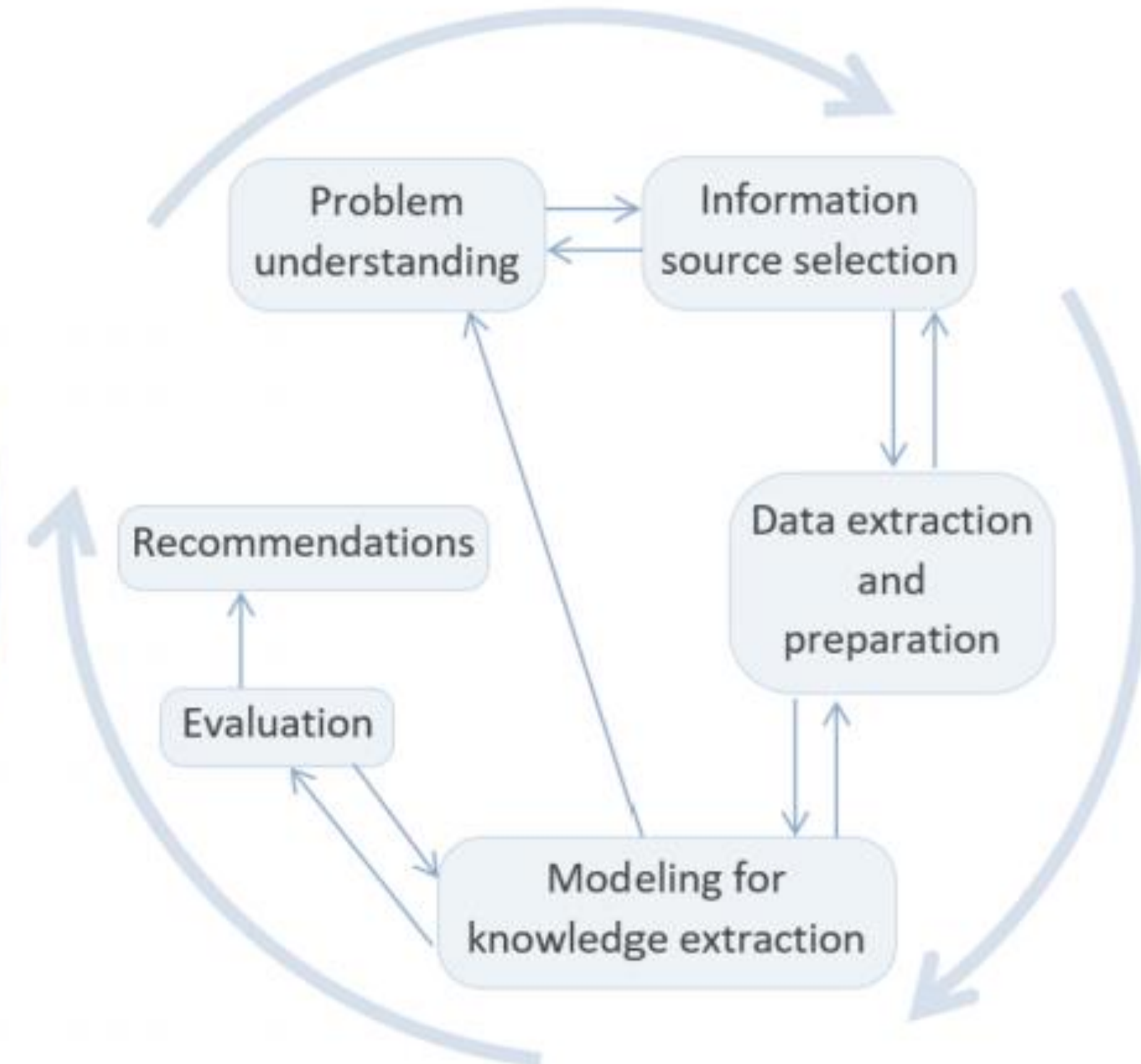
Methodology adopted

istar _iscte

Related Work

Queries:

- Articles
- written in English or Portuguese
- After 2019
- In Scopus and B-On
- with key words on the title or abstract

Key words	Scopus	B-On
(covid OR corona)	74 697	258 699
(covid OR corona) AND (twitter OR tweet OR reddit)	490	22 022
(covid OR corona) AND Portugal	149	316
(covid OR corona) AND "social media"	4	7
(covid OR corona) AND Portugal AND (twitter OR tweet OR reddit)	3	0
(covid OR corona) AND (twitter OR tweet OR reddit) AND sentiment	114	21 995
(covid OR corona) AND (twitter OR tweet OR reddit) AND (traking OR symptom OR trend)	81	33 576
(covid OR corona) AND (twitter OR tweet OR reddit) AND "public opinion"	14	33 602
(covid OR corona) AND "social media" AND (topic OR sentiment)	51	33 725
(covid OR corona) AND "social media" AND "public sentiment"	14	33 597
(covid OR corona) AND (twitter OR tweet OR reddit) AND sentiment AND topic	44	33 594

Related Work
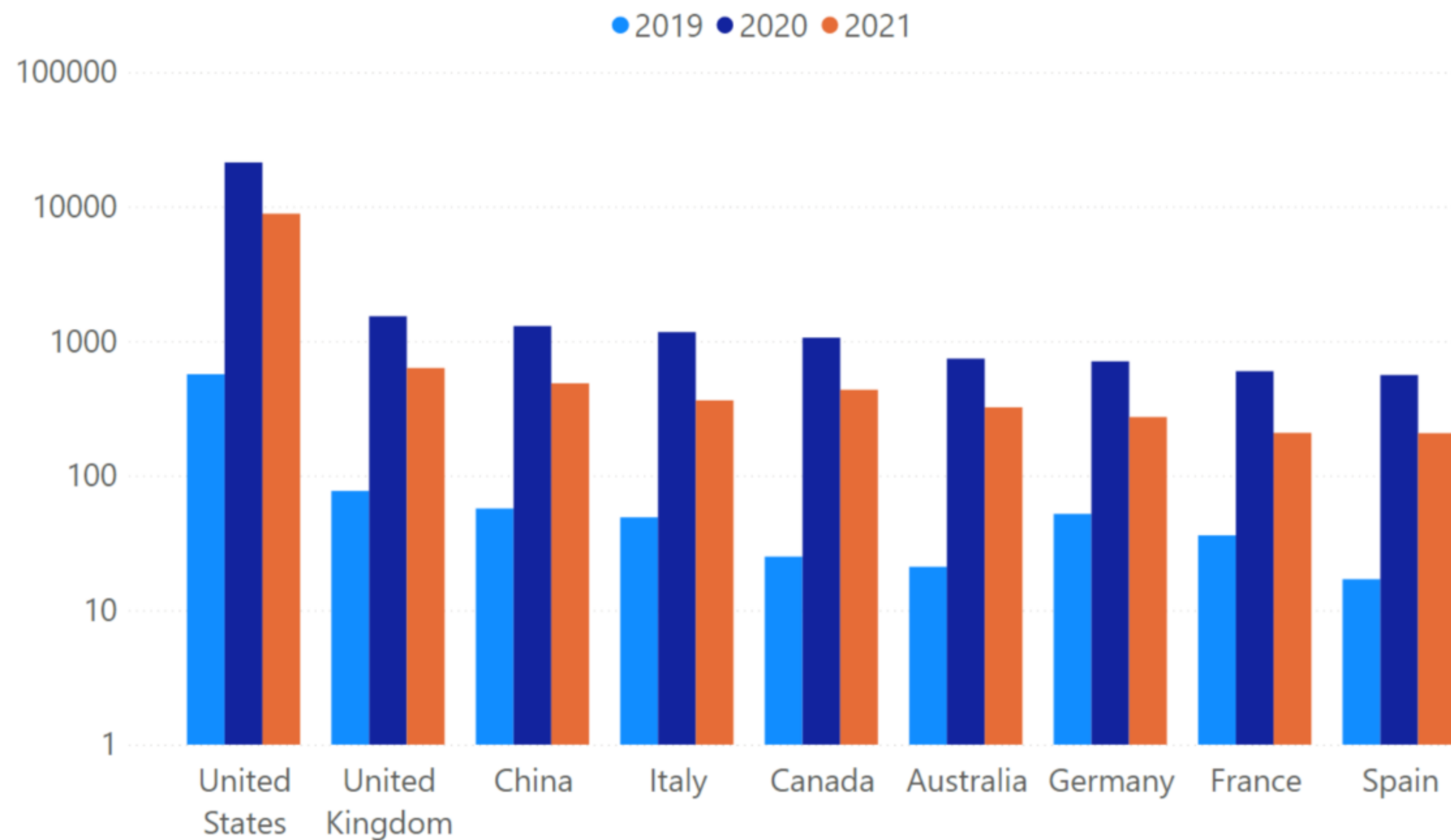
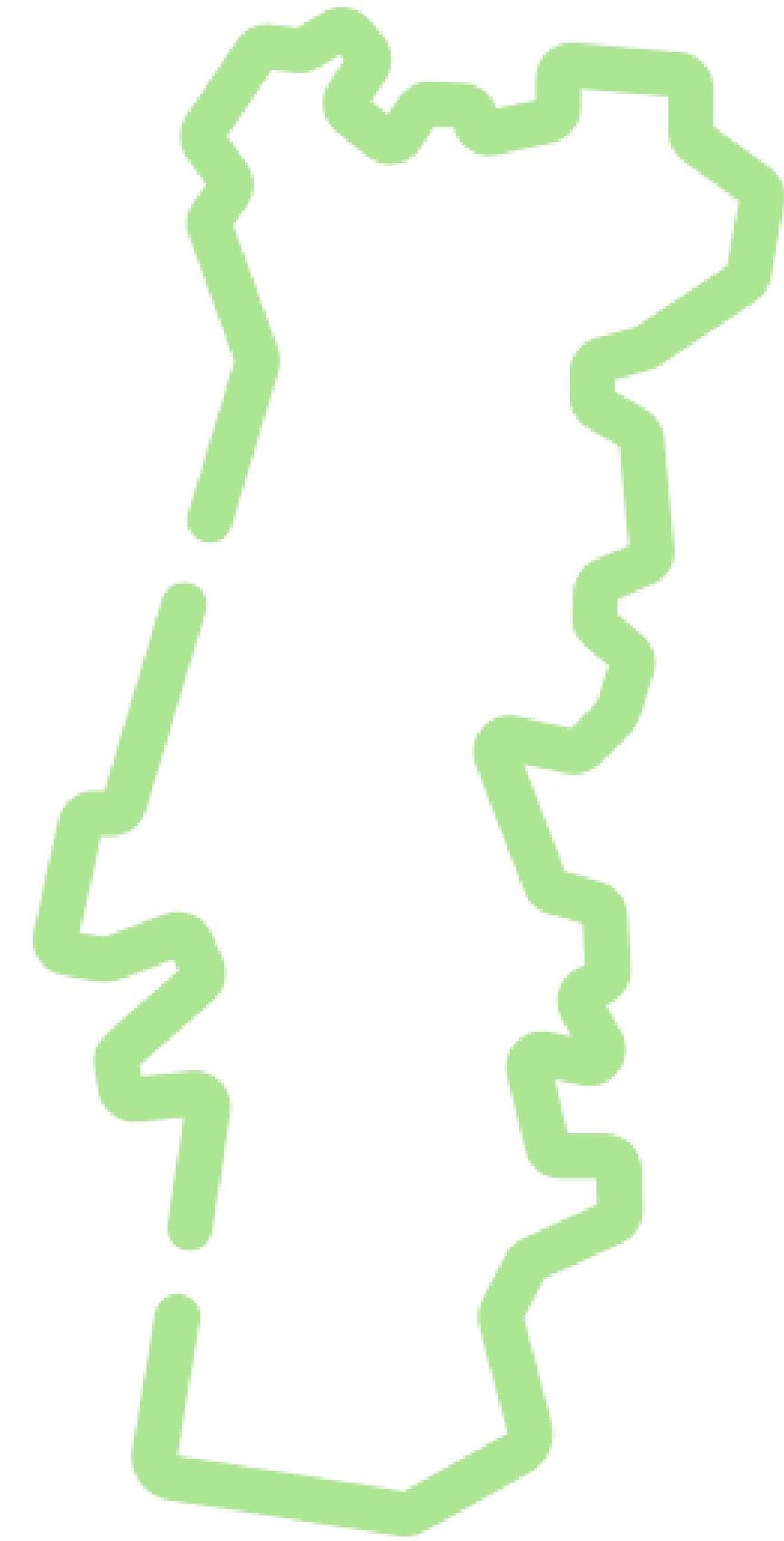Number of publications related to COVID-19 for year and country
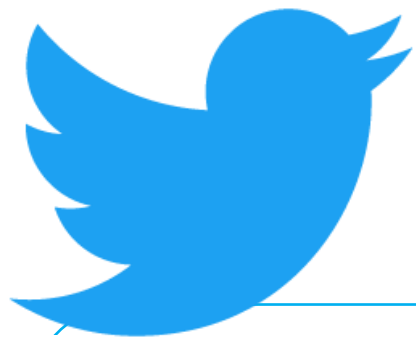
Prototype Build

CovidSocialSen
sing Platform

Information Source Selection

Data Extraction and Preparation

Key words used:
- **pandemia (pandemic);**
- **epidemia (epidemic);**
- **sars-cov2 ;**
- **covid;**
- **teletrabalho (telework);**
- **stayhome;**
- **FiqueEmCasa (stay at home);**
- **covax ;**
- **confinamento (lockdown);**
- **quarentena (quarantine).**

istar _iscte
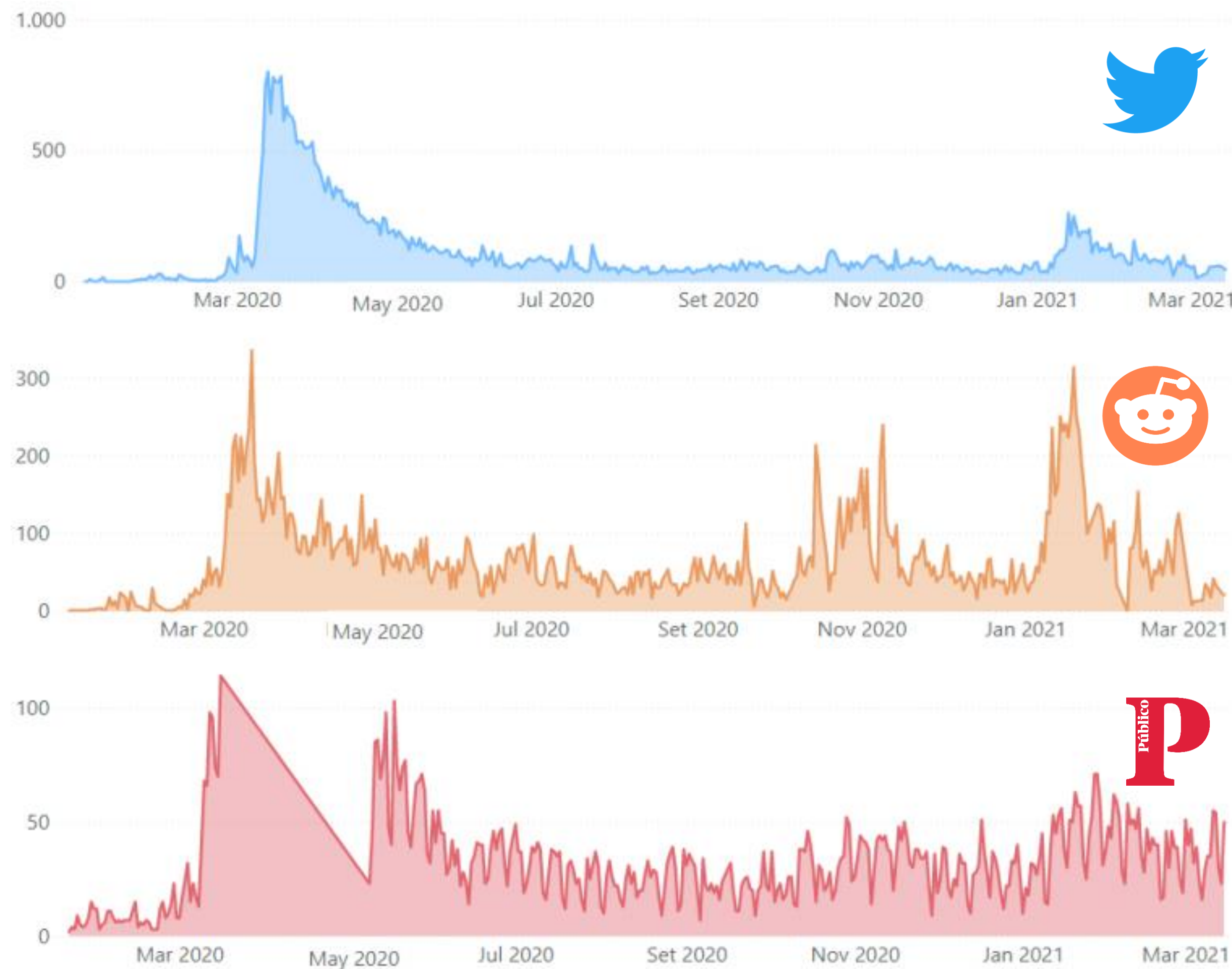
Data Extraction and Preparation

- Format standardisation
- Lower case unification
- Elimination of duplicate records
- Elimination of records with insufficient information
- Stopwords removal
- Remove double spacing
- Delete numbers
- Punctuation removal
- Link removal (Twitter)
- Elimination of retweet and user information (Twitter)
- Removal of audio and video information (Twitter)

istar _iscte

Modeling for Knowledge Extraction

- **Short text topic modelling (STTM) - based on LDA**

Portugue se cases

After lockdown

Modeling for Knowledge Extraction

Topic name	Top 6 words
Daily life in lockdown	casa, vou, dia, tudo, passar, saudades
Portuguese cases	portugal, pessoas, casos, ser, todos e virus
COVID-19 in Portugal	saúde, vamos, isolamento, casa, obrigado, mais
After lockdown	fazer, acabar, vou, casa, vai, dia
Possibility of having COVID-19	vou, espero, dia, provavelmente, sei, teste

Modeling for Knowledge Extraction

- **LeIA (Lexicon for Adapted Inference) – based on VADER (Valence Aware Dictionary and sEntiment Reasoner)**

Tweet	Neutral	Negative	Positive	Compound	Final result
A quarentena está a fazer mal a algumas pessoas, pelo que vejo...	0.645	0.355	0.0	-0.659	Negative
A força que nos faz mover, hoje foi dedicada aos que estão na linha da frente!	0.755	0.0	0.245	0.757	Positive
Nesta quarentena fiquei a saber todos os detalhes do meu quarto	1.0	0.0	0.0	0.0	Neutral

positive:
compound >= 0.05

negative:
compound <= 0.05

neutral:
-0.05 < compound > 0.05

Application to the Portuguese Context

P1 — Since January 1st until the first infected in Portugal

2/26/2020   3/16/2021

July, 1 — Calamity state ends and begins the state of alert

Nov, 8 — Calamity state ends and begins the state of emergency

Mar, 16 — End of the second lockdown

P2   P3   P4   P5   P6   P7   P8

Mar, 18 — First emergency state in Portugal

May, 2 — Emergency state ends and begins the state of calamity

Sep, 14 — Alert state ends and begins the state of contingency

Jan, 15 — Beginning of the second lockdown

Application to the Portuguese Context
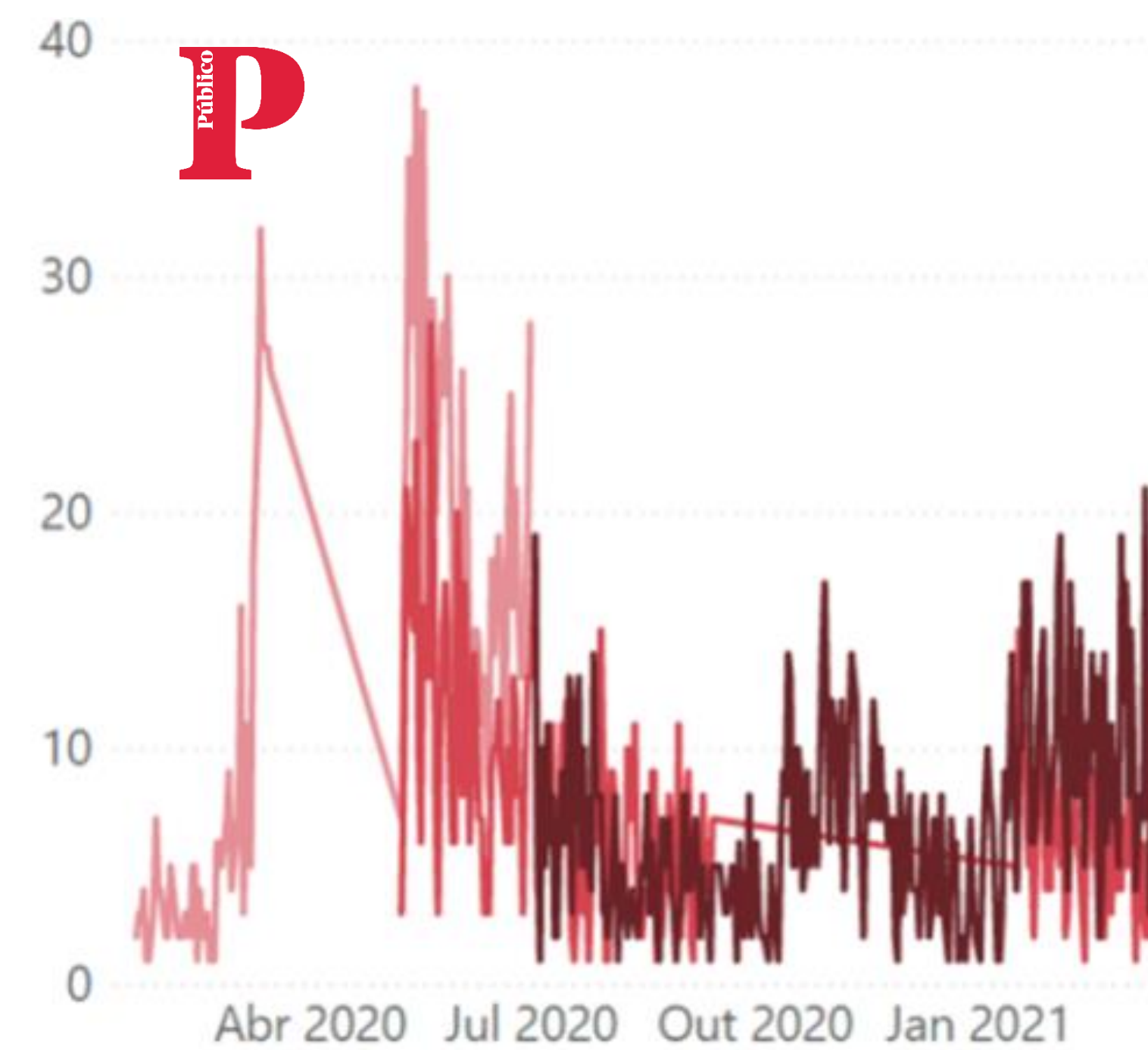
Application to the Portuguese Context

1st period – until first Portuguese infected – 1st January to 26th February

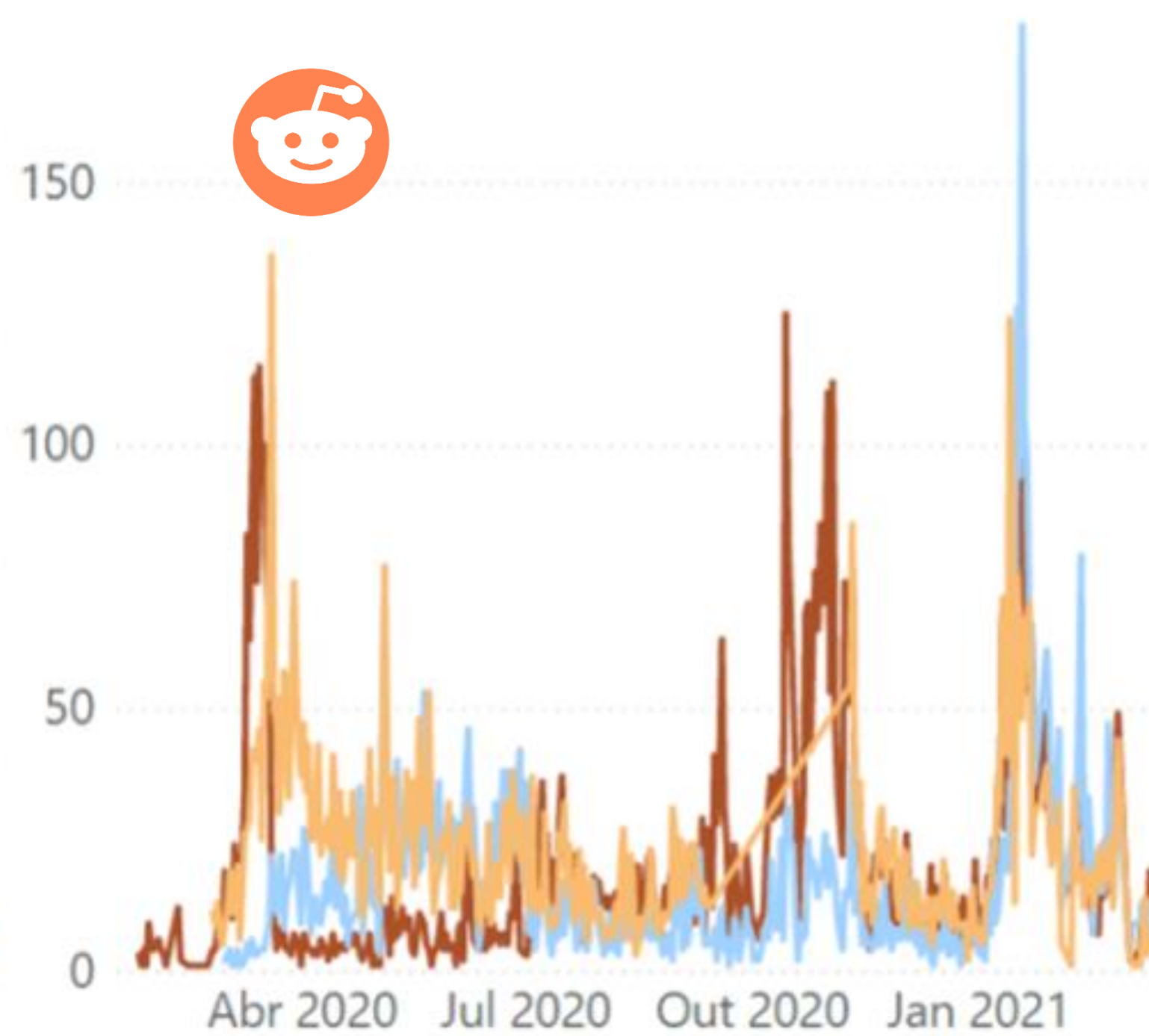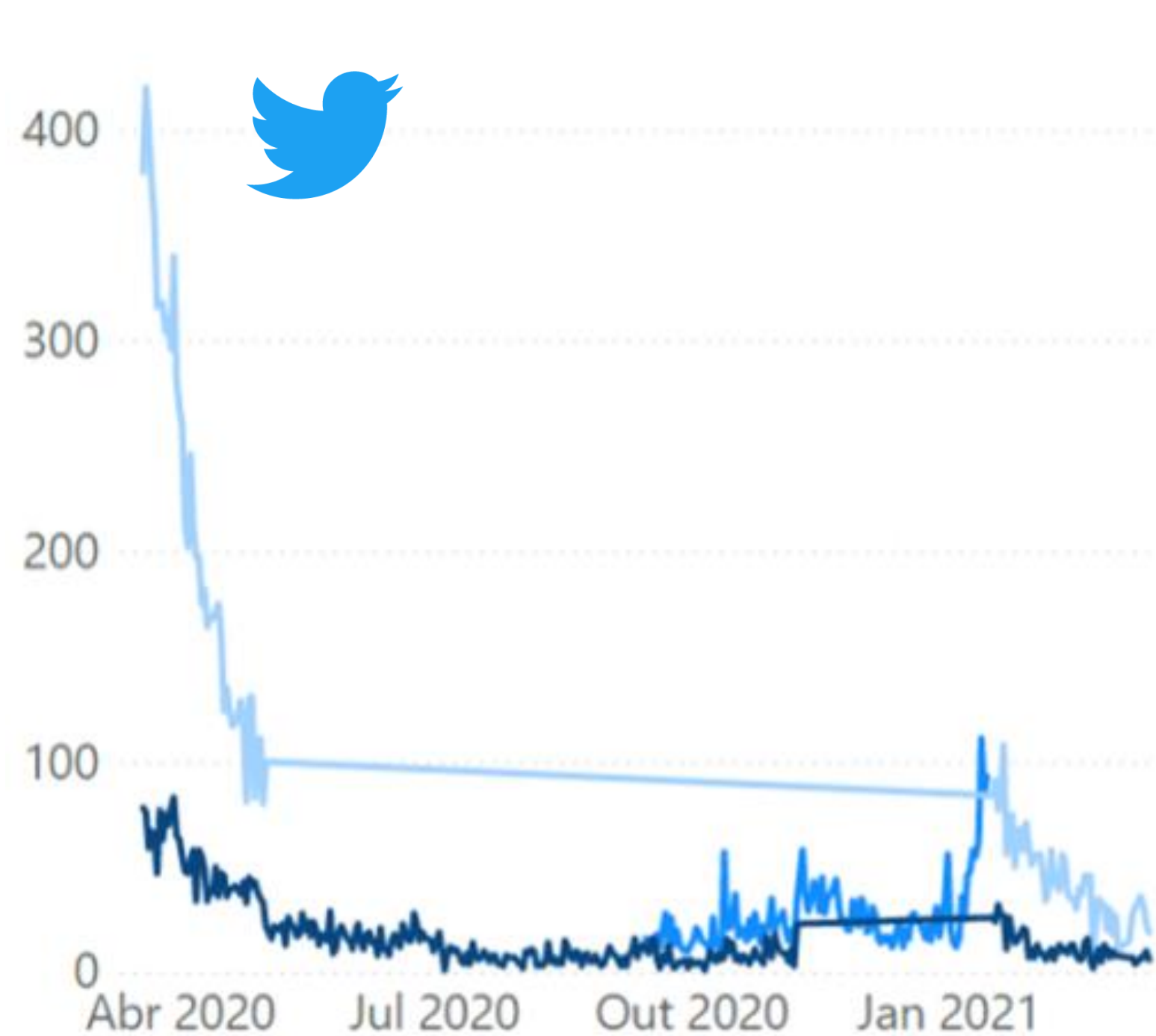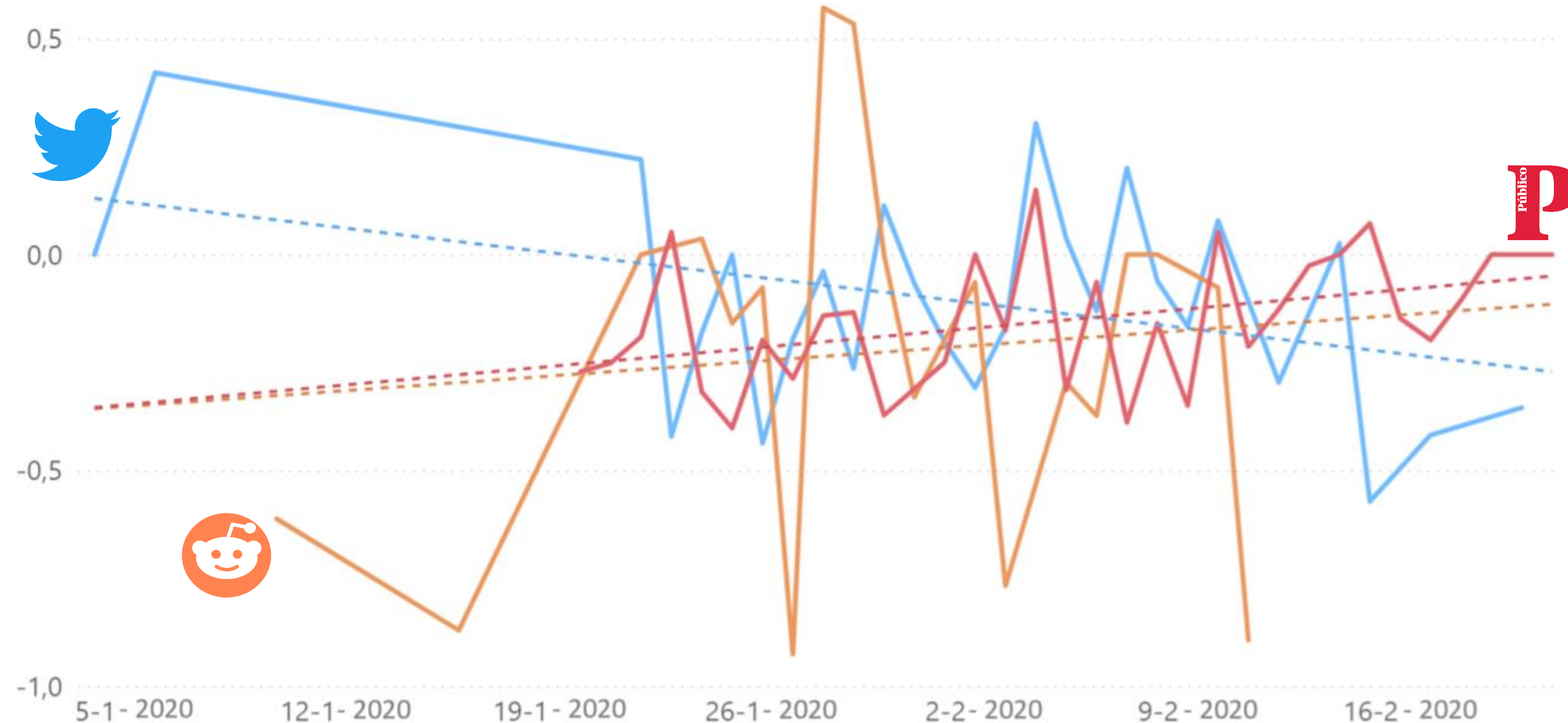**Sentiment trend line for the topic "COVID-19 in the world"**

Application to the Portuguese Context
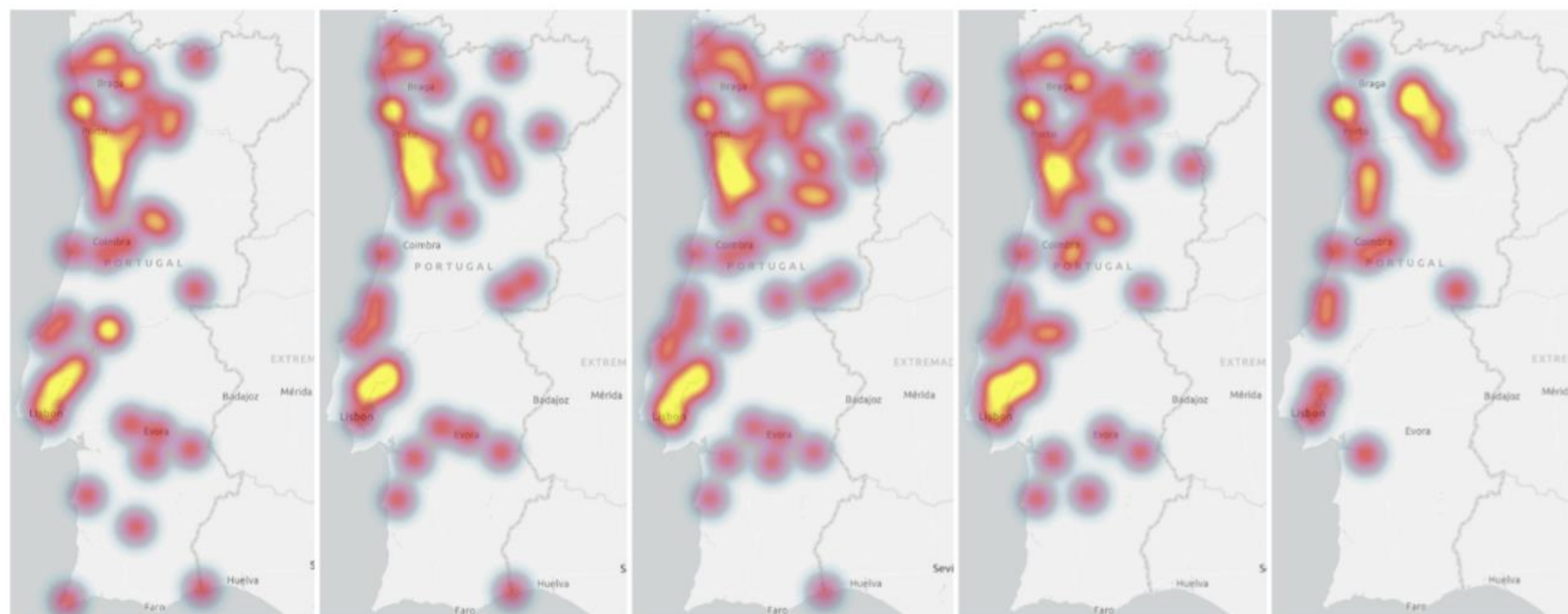
After lockdown

COVID-19 in Portugal

Daily life in lockdown

Portuguese cases

Possibility of having COVID-19

3rd period – first state of emergency – 18th March to 2nd May

8th period – second lockdown – 15th January to 16th March 2021

Daily life in lockdown

Football

Lockdown

Portuguese cases

Work from home

Application to the Portuguese Context

278 (6,9%)
1083 (27,1%)
342 (8,5%)
377 (9,4%)
554 (13,8%)
617 (15,4%)
576 (14,4%)

- COVID-19 teste
- Social events
- Avante party
- Portuguese cases
- Vaccine
- Government measures
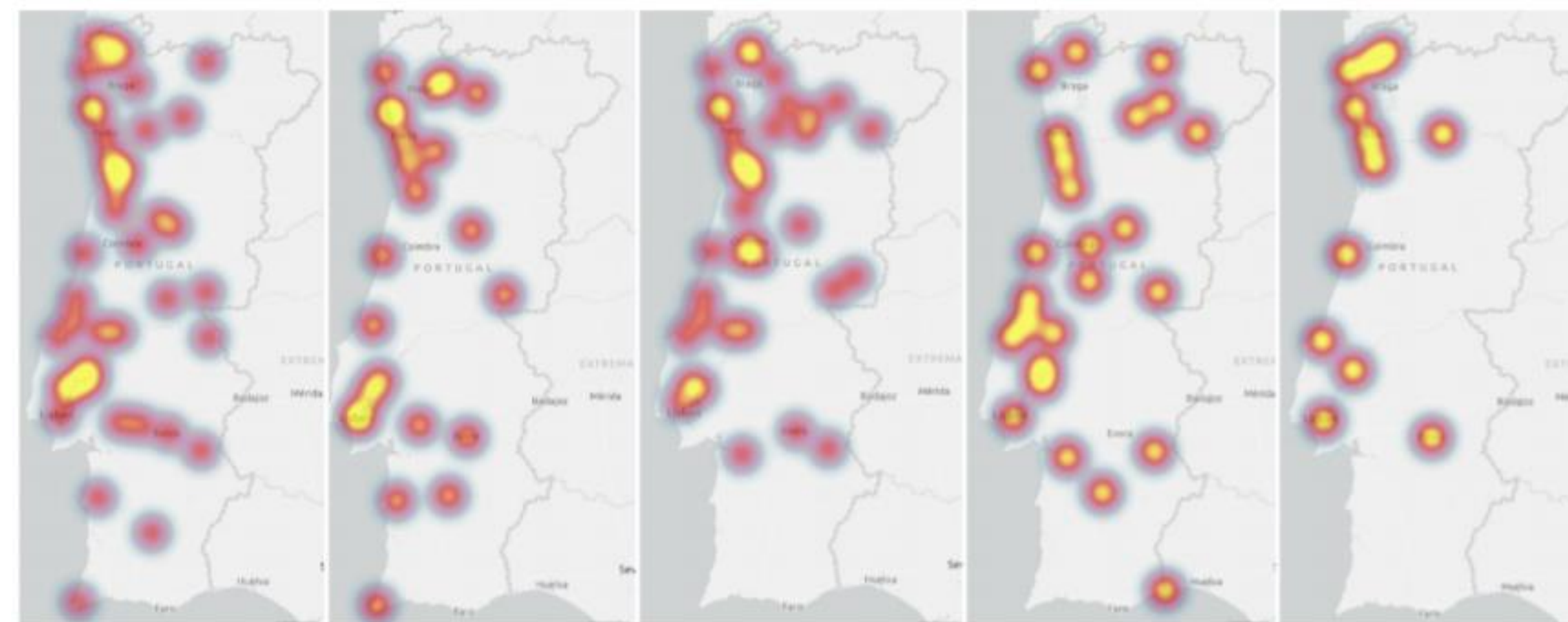- App StayAwayCovid
- Football

135 (4,1%)
1186 (36,0%)
699 (21,2%)
1112 (33,8%)

- Government measures
- Work from home
- Portuguese cases
- English posts
- Social events
- COVID-19 in Portugal
- Economy

321 (16,6%)
431 (22,3%)
386 (19,9%)
412 (21,3%)
387 (20,0%)

- Economy
- Social events
- COVID-19 in Portugal
- Government measures
- Vaccine

Application to the Portuguese Context

6th period – Since the end of alert state until the beginning of the second state of emergency – 14th September to 8th November

	Topic Name	Nº of posts	Average sentiment
Twitter	Alcanena	186	-0.06
	Portuguese cases	403	-0.11
	Government measures	240	-0.09
	Lockdown	841	-0.15
	COVID-19 tests	1025	-0.07
	App StayAwayCovid	573	-0.13
Reddit	Portuguese cases	609	-0.37
	Government measures	2 085	-0.26
	COVID-19 tests	683	-0.32
	App StayAwayCovid	439	-0.13
Público	Portuguese cases	360	-0.01
	COVID-19 in Portugal	230	-0.06
	National Health System	270	-0.07
	Government measures	314	-0.12
	Donald Trump	225	-0.04

istar _iscte

Evaluation

The evaluations were performed by three experts:

- a researcher and specialist in communication and public health,
- a resident public health doctor and consultant in health communication and data visualization,
- the vice president of the National Association of Public Health and coordinator of the combat against COVID-19 in the Azores.
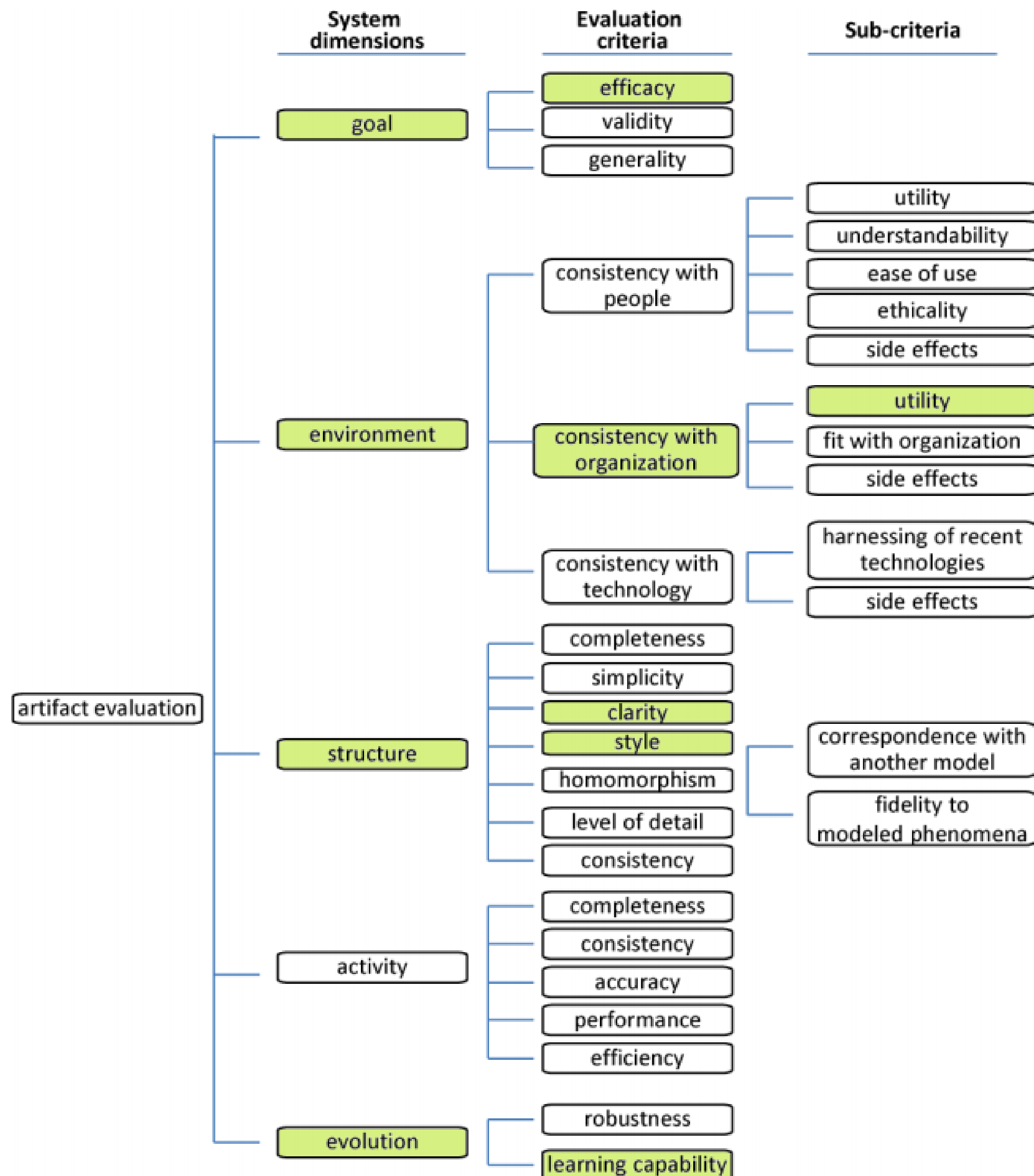
Evaluation

Criteria	Objective	Eval #1	Eval #2	Eval #3
Efficacy	Effectively inform about the topics and respective sentiment discussed on social media over time, regarding the pandemic situation in Portugal	LA	FA	TA
Consistency with organization /Utility	Obtain insights about a precision of the reality that the Portuguese share in the social media, which may contribute to help decision making of the medical community in Portugal	PA	LA	TA
Clarity	Providing clear and easily understandable information from the dashboards created	TA	LA	LA
Style	Providing appealing and understandable dashboards with straightforward interpretations	LA	TA	TA
Learning capability	Automatically learning about COVID-19 Portuguese insights regarding the discussion of topics and their associated sentiment on social media, during the first year of pandemic	PA	TA	LA

**Not Achieved (NA)
[0-15%]**

**Partially Achieved (PA)
]15-50%]**

**Largely Achieved (LA)
]50-85%]**

**Totally Achieved (TA)
]85-100%]**

istar _iscte

Conclusions

- **We conclude that by using social media, such as Twitter and Reddit, we can reflect a sense of reality with regard to the pandemic situation in Portugal**
- **As for the number of data collected in each of the eight periods, it is possible to conclude that it is influenced by the pandemic events in Portugal**
- **Also, the period where the average sentiment score was registered as more negative corresponds to the first alert state implemented in Portugal, since July 1 until September 14**
- **And the period where the amount of data analysis is higher corresponds to the period of the first compulsory confinement**

istar _iscte

Thank You