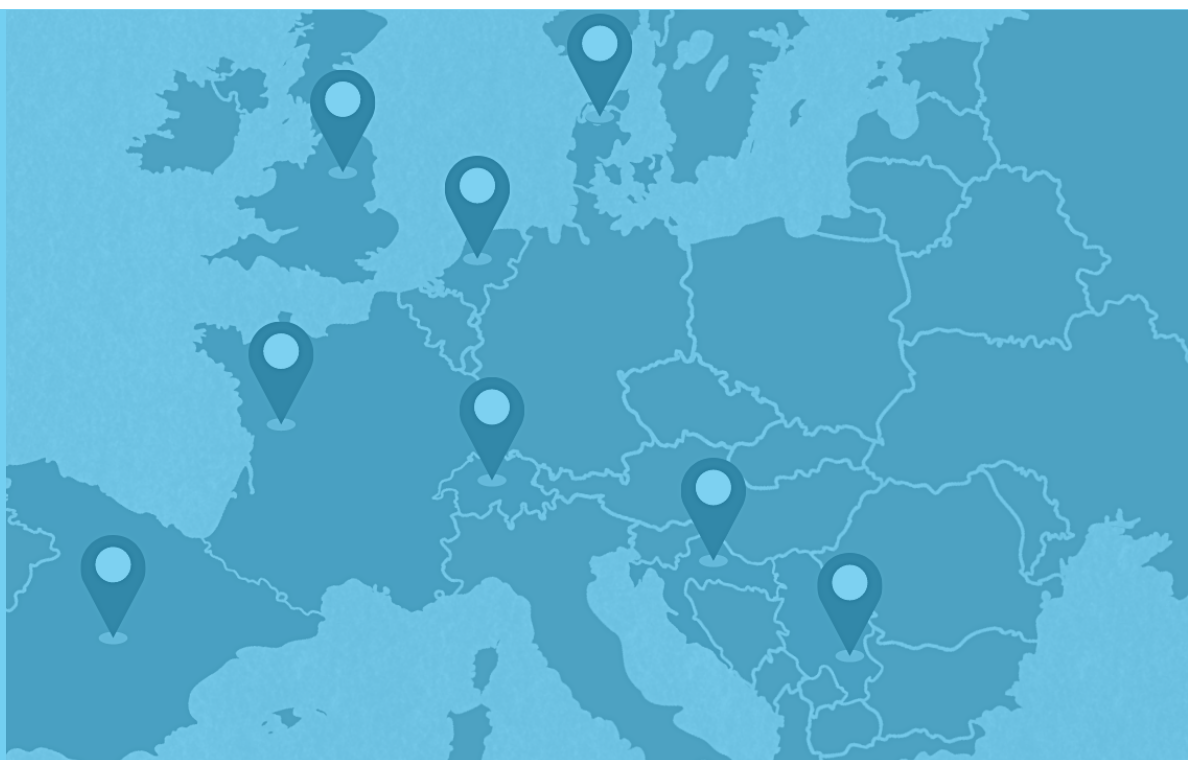


# invent

EUROPEAN INVENTORY  
OF CULTURAL VALUES



## Topics

- Literature-Arts
- Economics-Business-Work
- Identities and Boundaries
- Music-Concerts-Festivals
- Spaces and Places
- Policies and Politics
- Society and Inequalities
- Film-TV-Theatre
- Other Topics

## Data scraping of online content

### REPORT I

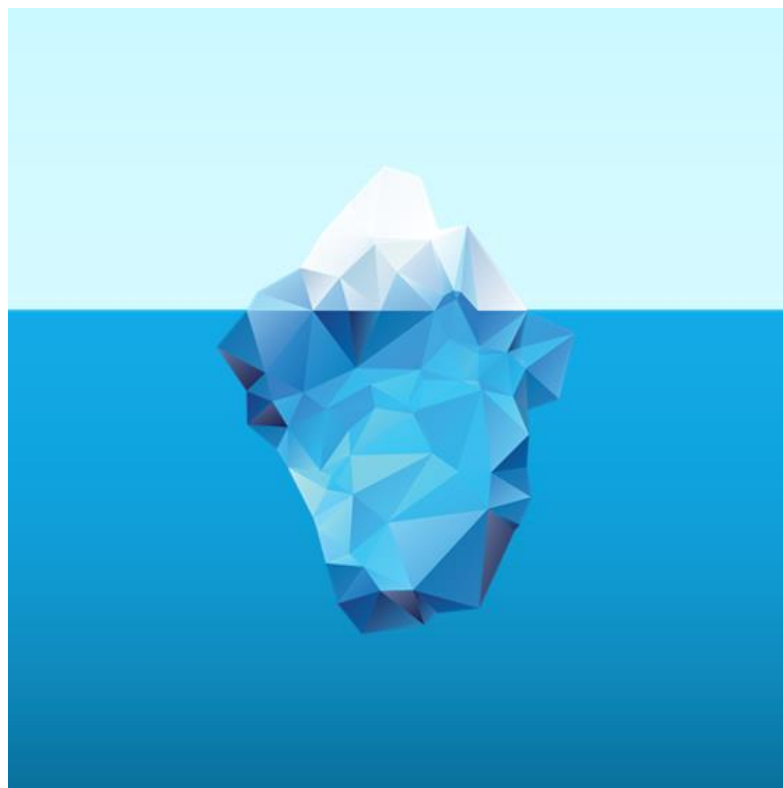
JULY 2021



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 870691

# invent

## EUROPEAN INVENTORY OF SOCIETAL VALUES OF CULTURE AS BASIS FOR INCLUSIVE CULTURAL POLICIES



### Deliverable 5.1

#### **Data scraping of online content - Report 1**

Grant agreement ID: 870691

Submission date: 30-07-2021

Beneficiary leading deliverable: UCPH



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 870691

## Deliverable description

<b>Grant Agreement ID</b>	<b>870691</b>
<b>Project name</b>	European inventory of societal values of culture as a basis for inclusive cultural policies in the globalizing world
<b>Project acronym</b>	INVENT
<b>Project start date</b>	1-2-2020
<b>Project duration</b>	42 months
<b>Project website</b>	<a href="https://inventculture.eu/">https://inventculture.eu/</a>
<b>Deliverable number</b>	D5.1
<b>Deliverable name</b>	The first yearly report on the data scraping of online content
<b>Work package</b>	WP5 – Cultural is digital/boundless
<b>Work package leader</b>	University of Copenhagen (UCPH)
<b>Work package participants</b>	EUR, UAB, UoH, UCPH, TAU, UZH, ISSIP, CECS, ENS
<b>Authors</b>	Geffen Ben David, Susanne Janssen, Jinju Kim, Jordi López-Sintas, Franziska Marquart, Ratko Nikolić, Lucas Page Pereira, Ossi Sirkka, Sebastian Weingartner, Neta Yodovich, Željka Zdravković
<b>Type</b>	Report
<b>Version</b>	1
<b>Draft/Final</b>	Final
<b>Number of Pages</b>	143
<b>Due date</b>	31-07-2021
<b>Submission date</b>	30-07-2021
<b>Keywords</b>	culture, social media, topic modelling, comparative analysis, Covid-19, Europe

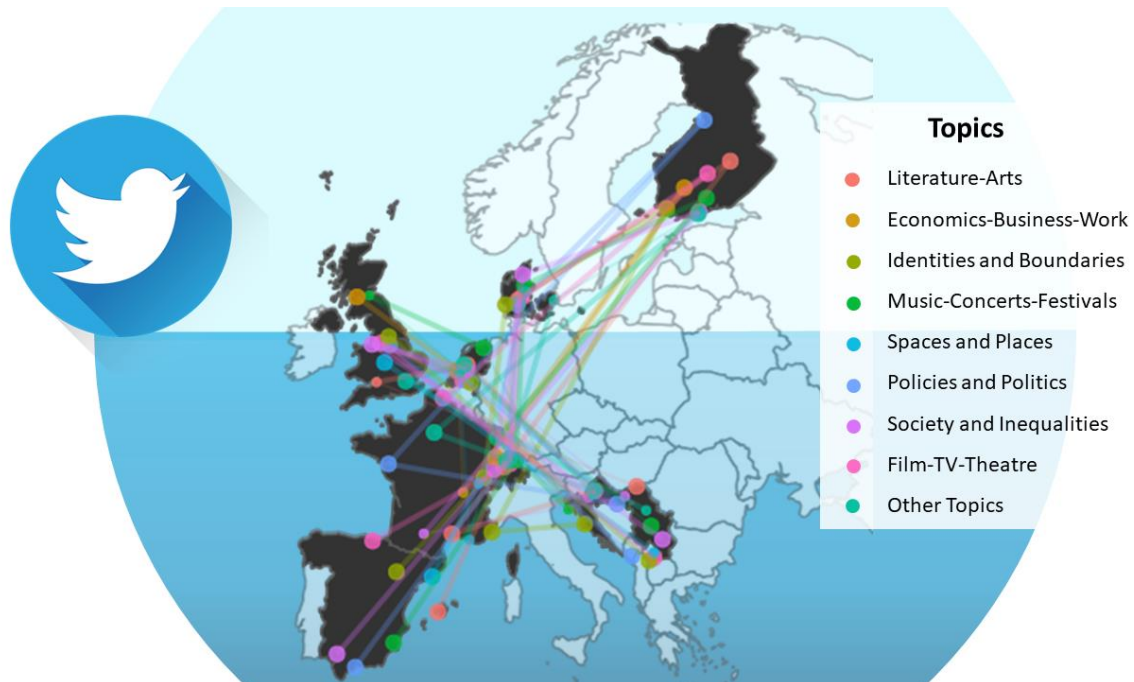


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 870691

# Talking About Culture on Twitter

A comparative analysis of culture-related topics  
in nine European countries in 2019 and 2020

<i>Geffen Ben David</i>	<i>(University of Haifa)</i>
<i>Susanne Janssen</i>	<i>(Erasmus University Rotterdam)</i>
<i>Jinju Kim</i>	<i>(Universitat Autònoma de Barcelona)</i>
<i>Jordi López-Sintas</i>	<i>(Universitat Autònoma de Barcelona)</i>
<i>Franziska Marquart</i>	<i>(University of Copenhagen)</i>
<i>Ratko Nikolić</i>	<i>(University of Belgrade and CECS)</i>
<i>Lucas Page Pereira</i>	<i>(University of Versailles-Saint-Quentin)</i>
<i>Ossi Sirkka</i>	<i>(Tampere University)</i>
<i>Sebastian Weingartner</i>	<i>(University of Zurich)</i>
<i>Neta Yodovich</i>	<i>(University of Haifa)</i>
<i>Željka Zdravković</i>	<i>(University of Zadar and ISSIP)</i>





## Executive Summary

In the INVENT project, we aim to develop new perspectives and methodologies for capturing the wider societal value of culture. To understand how Europeans view culture and its societal values, we follow a bottom-up approach to researching and mapping cultural diversity in Europe. In doing so, we employ – among other things – data scraping techniques to collect digital content and gain a better understanding of how Europeans express themselves online about culture. Two main research questions guided our investigation in the first data scraping phase: First, in mapping the cultural ‘Twittersphere’ in all INVENT partner countries, we asked which topics are associated with the keyword ‘culture’ in conversations on Twitter during the years 2019 (pre-Covid19) and 2020 (during Covid19). Second, we were interested in the types of actors that engage in conversations about culture on Twitter and the extent to which these interact with each other.

The data scraping process began with the social media platform Twitter in January 2021 using R programming language. We decided to focus on Twitter for several reasons. First, Twitter is prominently used for sharing news stories, but also enables the emergence of stories from below, based on individuals’ experiences, opinions, and struggles, making it a highly relevant source for studying understandings of culture from the bottom-up. Second, Twitter offers one of the most accessible websites for research in the social sciences. In 2021, the platform introduced an advanced Academic Research product track that facilitates access to the platform’s historical archive for qualified researchers. To scrape relevant tweets, we used the keyword “culture” in the INVENT project’s respective country languages; in some cases, the sample was further extended with English-language “culture” tweets as well. We focused on tweets posted from users’ home locations in each country and analyzed the tweets through topic modelling, a powerful and effective method for detecting relevant hidden structures in the data.

We identified eight thematic topic groups across the nine INVENT countries: Tweets and conversations about “Film/TV/Theatre”, “Literature/Arts”, “Music/Concerts/Festivals”, “Policies and Politics”, “Identities and Boundaries”, “Society and Inequalities”, “Economics/Business/Work”, and “Spaces and Places”. As these topics are identified in almost all countries, they are at the heart of discussions related to culture in Europe. Additionally, we detected several country-specific topics that are unique to the particular context in each country, such as conversations on “MeToo” or “Health”. Such context-specific cultural topics could be of interest to cultural policymakers, but they also highlight the need to acknowledge cultural discussions on the regional and national level. Our analysis shows that there is a widely shared basis of culture that refers not only to the arts, leisure, and creativity (“culture as arts”), but also to more elementary aspects of human life such as politics, inequality, and economics (“culture as way of life”). Another similarity between many countries, albeit to varying degrees, is the relationship between culture and the emerging Covid19 pandemic, where Twitter users discussed the shutting down of cultural venues and the policies implemented to save at-risk events and venues. The most active users in our data are media and cultural organizations and individuals such as journalists and artists, that account for a significant share of tweets in each country. This means that the perceptions and viewpoints of specific organizations and professions are to some extent overrepresented in our data. Still, we found a large variety of users in each country who contribute to cultural discussions on Twitter, indicating that analyzing Twitter data is indeed useful in studying culture in a bottom-up fashion. Despite choosing a general keyword, our approach restricted the data collection to tweets that explicitly mention the word “culture”, and posts about other cultural offerings were, therefore, not included in our data. This report is the first phase out of three in our web scraping research. Based on our findings, for the next phase, we are considering looking into active users or particular cultural offerings that were not reflected in our current findings.

## Contents

<b>Deliverable description.....</b>	<b>0</b>
<b>Executive Summary .....</b>	<b>2</b>
<b>Contents.....</b>	<b>3</b>
<b>Introduction .....</b>	<b>5</b>
1. The context of the INVENT project.....	5
2. Web scraping .....	5
3. Social media platforms and Twitter .....	6
<b>Methodology.....</b>	<b>8</b>
1. Data Scraping.....	8
2. Pre-Processing .....	10
3. Topic Modelling .....	10
<b>Findings for Croatia .....</b>	<b>12</b>
1. Twitter usage in Croatia .....	12
2. Data description .....	12
3. User analysis.....	13
3. Hashtag analysis .....	15
4. Keywords analysis .....	17
5. Topic modelling .....	18
6. Discussion and conclusion .....	22
<b>Findings for Denmark .....</b>	<b>23</b>
1. Twitter usage in Denmark .....	23
2. Data description .....	23
3. User analysis.....	25
3. Hashtag analysis .....	28
4. Keyword analysis .....	30
5. Topic modelling .....	33
6. Discussion and conclusion .....	37
<b>Findings for Finland .....</b>	<b>39</b>
1. Twitter usage in Finland .....	39
2. Data description .....	39
3. User analysis.....	39
4. Hashtag and keyword analysis .....	41
5. Topic modelling .....	43
6. Discussion and conclusion .....	46
<b>Findings for France .....</b>	<b>48</b>
1. Twitter usage in France .....	48
2. Dataset description .....	48
3. Hashtag analysis .....	49
4. Topic modelling .....	53
5. Discussion and conclusion .....	58
<b>Findings for the Netherlands .....</b>	<b>60</b>
1. Twitter usage in the Netherlands.....	60
2. Data description .....	60
3. User analysis.....	61
4. Hashtags and keyword analysis.....	65
5. Topic modelling .....	68
6. Discussion and conclusion .....	76
<b>Findings for Serbia.....</b>	<b>78</b>
1. Data description .....	78
2. User analysis.....	79

3. Hashtag and keyword analysis .....	81
4. Topic modelling .....	83
5. Discussion and conclusion .....	88
<b>Findings for Spain .....</b>	<b>89</b>
1. Twitter usage in Spain .....	89
2. Data description .....	89
3. Hashtags and keywords analysis .....	90
4. Topic modelling .....	92
5. Discussion and conclusion .....	98
<b>Findings for Switzerland .....</b>	<b>100</b>
1. Twitter usage in Switzerland .....	100
2. Data description .....	100
3. User Analysis .....	102
4. Hashtag and keyword analysis .....	103
5. Topic modelling .....	108
6. Discussion and conclusion .....	111
<b>Findings for the United Kingdom .....</b>	<b>112</b>
1. Twitter usage in the UK .....	112
2. Data description .....	112
3. Hashtag and keyword analysis 2019 .....	114
4. Hashtag and keyword analysis 2020 .....	116
5. Topic Modelling .....	117
6. Discussion and conclusion .....	122
<b>General Discussion and Conclusion .....</b>	<b>124</b>
1. Country comparison .....	124
2. Time comparison .....	130
3. Limitations and future research .....	131
4. Policy implications .....	132
<b>References .....</b>	<b>134</b>
<b>Appendix .....</b>	<b>137</b>
Tables A1 and A2 .....	137
R Script used to download tweets utilizing Twitter Academic research product track .....	138
List of Tables .....	139
List of Figures .....	140

## Introduction

### 1. The context of the INVENT project

In the INVENT project, we aim at developing new perspectives and methodologies for capturing the wider societal value of culture. In acknowledgment of the importance of culture in our changing modern world, we investigate the impact of ongoing processes such as migration, globalization, and digitalization on citizens' perceptions of and participation in culture across Europe. We pay special attention to the use of digital technologies and to the ability of the Internet to provide a broader, more democratic access to culture and heritage. In order to create inclusive and participatory cultural policies suited to the age of globalization, we require scientific insights into how European citizens use online and digital media as a) forms of cultural consumption and b) forms of digital infrastructure that enable access to cultural content and deliberation (see Mihelj, Leguina, et Downey 2019; Weingartner 2020).

To understand how Europeans view culture and its societal values, we follow a bottom-up approach to researching and mapping cultural diversity in Europe. We consider digital cultural participation an important aspect of citizens' general perceptions of and participation in (European) culture. Notably, digital participation is increasingly conceptualized as a complementary and integral component of public participation as a whole (Mihelj, Leguina, and Downey 2019; Smith 2005). It should therefore be examined and evaluated both separately and in conjunction with more 'traditional' forms of cultural engagement offline. We thus ask what citizens define as 'cultural' topics, and under which circumstances they discuss culture in the digital world. Importantly, culture can be something that is shaped and adjusted by individuals in a collective online space (e.g., on social network platforms) and through user-generated content. Yet investigations into digital cultural expression and participation remain rare (Highfield, Harrington, and Bruns 2013; Agostino 2018; but see, e.g., MacDowall and de Souza 2018). Not only do we know little about the platforms used by European citizens to discuss cultural topics, we also have only few insights into the type of cultural engagements and discourses that take place online. In the INVENT project, we therefore identify key local actors and networks by assessing digital cultural participation online, and we investigate to which extent public participation brings together citizens and policy makers in public decision-making (Smith 2005).

### 2. Web scraping

To address these challenges, we employ web scraping techniques to collect digital content and gain a better understanding of how Europeans express themselves online about culture through evaluative, conversational, and creative-productive modes. With the help of web scraping, researchers can collect large amounts of digital data via automated data extraction from the World Wide Web, often making use of available code packages based on common programming languages (e.g., C/C++, R, Python). One of the method's main advantages is its simplicity: When scraping content from the Web, researchers may efficiently and unobtrusively 'observe' the creation of and engagement with content online without interfering with the object of their study. In addition, and depending on the research interest, scholars may collect thousands of relevant cases in a very short amount of time (Landers et al 2016), which significantly facilitates the research process and can help in the study of large-scale social phenomena.

While web scraping is increasingly employed across scientific disciplines, it is seldom used in research on digital cultural participation. In the INVENT project, we employ the method to

complement other approaches of data collection. Web scraping allows us to study forms of expression that are difficult to examine via (standardized) questionnaires, and enables us to investigate citizens' 'natural behavior' in the increasingly relevant digital sphere. In the first of three phases of web scraping, we assess what citizens talk about when they talk about 'culture' online – specifically, on the social media platform Twitter.

### 3. Social media platforms and Twitter

Internet access in European countries has increased significantly over the last decade. Indeed, if in 2011 71% of European households (EU-27) had access to the Internet, in 2020 more than 91% had an Internet connection (

**Table 1).** The daily use of the Internet has also become a reality for most Europeans, who also use it to access social networks. These platforms have become a space for entertainment, for strengthening ties with friends and family, for keeping up to date with news, and for free self-expression for more than half of Europe's residents. This use, however, varies from country to country: whereas in Denmark 80% of people are connected to a social network in 2019, this holds for only 40% of French residents.

**Table 1.** Internet access, frequency of use and participation in social networks in Europe (%), 2011 and 2019

	2011			2019		
	<i>Households with access to the internet*</i>	<i>Individuals accessing the internet daily**</i>	<i>Individuals participating in social networks***</i>	<i>Households with access to the internet*</i>	<i>Individuals accessing the internet daily**</i>	<i>Individuals participating in social networks***</i>
Denmark	90	78	55	95	92	81
Spain	63	47	35	91	78	59
France	76	62	36	90	77	42
Croatia	61	44	32	81	71	58
Netherlands	94	79	46	98	92	67
Finland	84	76	45	94	90	67
Switzerland	91 <sup>1</sup>	76 <sup>1</sup>	43 <sup>1</sup>	96	89	53
UK	83	70	50	96	91	73
Serbia	64 <sup>1</sup>	55 <sup>1</sup>	49 <sup>1</sup>	80	72	55
EU - 27	72	54	36	90	77	54

**Source:** Eurostat \*ISOC\_CI\_IN\_H, \*\*ISOC\_CI\_IFP\_FU, \*\*\*ISOC\_CI\_AC\_I

\* Households with at least one person aged 16-74.

\*\*/\*\* Persons aged 16-74.

<sup>1</sup> Data for 2014

Since their inception, social media have permeated every aspect of our everyday lives: from the way in which we engage with our friends or find a new job to how we follow our favorite celebrity. They have also become major platforms for self-presentation and self-expression, where people can share their opinions, experiences, and achievements and showcase their creative products, cultural activities, and lifestyle purchases. In our exploratory survey on cultural consumption during the

Covid-19 pandemic, we found that during the initial lockdown in 2020, many citizens across Europe turned to social media to engage with their families and friends, and to access culture (Katz-Gerro et al. 2020).

Among the commonly used social media platforms, Twitter offers one of the most accessible websites for research in the social sciences. Twitter is an online microblogging site that was originally developed in 2006. With a maximum of 280 characters since 2017, individuals are invited to share posts about their private lives, news, popular culture, and more. These tweets reach posters' personal network of followers and beyond through the option of "retweets". With its fast-paced posting regime, Twitter has been recognized as an effective technological tool that can transform how people share knowledge and interact with one another (Newgarden 2009). In contrast to other platforms, which put a stronger emphasis on users' social engagement, Twitter, with its massive sources of information, also enables its users to keep themselves informed about current events (Duncombe 2019; Vergeer 2015). Not only does the platform serve as a news-sharing arena, it also dictates its own current affairs: on Twitter, stories emerge from below, based on individuals' experiences, opinions, and struggles (Duncombe 2019; Hutchins 2011). This makes Twitter a perfect source for studying understandings of culture from the bottom up. Notable examples include the MeToo movement, which started on Twitter (Bogen, Bleiweiss, and Orchowski 2019), or the series of protests and rebellions forming the Arab Spring (Meraz and Papacharissi 2013), which also gained momentum through the platform. In recent years, social scientists have embraced Twitter as an accessible research terrain providing a large quantity of real-time data. With around 500 million tweets posted each day and an open application programming interface (Duncombe 2019; Komorowski, Huu, and Deligiannis 2018), Twitter offers a plethora of accessible data that can be used for research. Ovadia (2009), for instance, explains that with its easy keyword or hashtag search, Twitter allows researchers to track online discussions on any topic.

Because Twitter is prominently used for sharing news stories, scholars have used the platform to explore political discourse online on topics such as elections, climate change, or social movements (Duncombe 2019; Fownes, Yu, and Margolin 2018; Meraz and Papacharissi 2013; Vergeer 2015). Researchers have also turned to Twitter to examine discussions regarding cultural consumption (Gutiérrez-Martín and Torrego-González 2018; Highfield et al. 2013; Hutchins 2011). As one example, Stever and Lawson (2013) focused on audiences' unmediated engagement with celebrities or influencers. Moreover, researchers study real-time tweets posted during live events, such as football matches, broadcasts of a new episode of a beloved TV series, or live singing competitions, such as the Eurovision Song Contest (Highfield et al. 2013; Hutchins 2011; Wood and Baughman 2012). Other scholars used Twitter for investigations into cultural discussions in a broader sense, such as Florini's (2014) research on "Black Twitter", where individuals who identify as Black use the platform to share their stories and develop their own jargon and jokes.

In the spring of 2021, the platform further eased the access for qualified researchers by introducing an advanced Academic Research product track that also provides access to the platform's full archive – all content published since its establishment in 2006 (Tornes and Trujillo 2021). For the first time, scientists have the possibility to not only track developments or events unfolding on Twitter in real time, but may also investigate network dynamics and user engagement retrospectively and without paying for the content. Twitter therefore provides an ideal research environment for the study of digital cultural participation in the context of the INVENT project.

Two main research questions guided our investigation: First, in mapping the cultural 'Twittersphere' in all INVENT partner countries, we asked which topics are associated with the keyword 'culture' in conversations on Twitter during the years 2019 and 2020. Second, we were interested in the types of actors that engage in conversations about culture on Twitter (e.g., cultural institutions, citizens,



individual artists/groups, cultural influencers), and the extent to which these interact with each other (e.g., through following/retweeting behavior). The unique composition of the INVENT team allowed us to not only assess these questions within each of the partner countries, but also to compare findings for different European countries with regard to the use of Twitter in digital cultural discussion and participation. Furthermore, we studied the extent to which conversations about culture on social media were affected by the Covid-19 pandemic, comparing data from 2019 with 2020 (Katz-Gerro et al. 2020).

## Methodology<sup>1</sup>

### 1. Data Scraping

All nine institutions of the INVENT project participated to analyze online data for every country in the project. The data scraping process began with Twitter in January 2021 using the R programming language. After a few preliminary processes of data scraping, we decided to apply two ways of scraping corresponding to each country due to the vast differences in the number of available tweets between the countries. Two distinct means for downloading Twitter data were: 1) utilizing an academic track with the new V2 API (Application Program Interface) and 2) utilizing a Premium account with R package *rtweet* (Kearney 2019). In the next paragraphs, these approaches are introduced.

#### Academic research product track and V2 API

Twitter offers a new *Academic research* product track and with this academic account we could theoretically access as many as *10 million* tweets per month for free (Twitter 2021). The Academic research product track provides historical data scraping with search filtering only by tweets location (place country). Therefore, in an additional step, we filtered by the location of the Twitter account (profile country) in order to detect discussions from people who actually live in the country, not tourists etc. In countries with a smaller pool of available tweets, the tweets location filtering reduced sizes of available tweets significantly. Consequently, only the Spanish, French, and UK teams were able to obtain a sufficient amount of data scraped based on tweets location, and, hence, could use the academic research track, whereas other teams employed the premium account (see below).

To utilize the academic account to obtain Twitter data, we had to develop custom R code to download the tweets. A script available in GitHub was used to connect to V2 API (Schoch 2021). The maximum number of tweets obtained with a single call to V2 API is 500, so we created a data frame containing the desired start and end dates and looped the function over every row of the data frame. The code used to download tweets is presented in the Appendix.

#### Rtweet and Twitter Premium API account

After preliminary inspection of the data quality achieved with a chargeable Twitter Premium API account, we bought premium access for the Croatian, Danish, Dutch, Finnish, Serbian, and Swiss teams. The R package *rtweet* provides relatively easy access to the full archive endpoint in Twitter API, which was required to get access to historical data (for example, posts tweeted one year ago). The price of these accounts is determined by the maximum number of tweets available per month. Utilizing premium accounts together with *rtweet* yielded the possibility of filtering tweets directly by the country of the user (profile-country), which produced good and reasonably priced data sets for these countries. The essence of the R code used to download tweets were in the form of the

---

<sup>1</sup> The authors would like to acknowledge the important methodological contribution of Jinju Kim, postdoctoral researcher in the Spanish INVENT team, to the process of data collection and analysis. Also the methodological support of Ossi Sirkka and Lucas Page Pereira from the Finnish and French team is highly appreciated.

following:

```
library(rtweet)
full_archive_tweets <- search_fullarchive("kulttuuri,
  profile_country:fi -is : retweet",
  n = 100000,
  fromDate = 201901010000,
  toDate = 202012312359,
  env_name = "Environment_name")
```

### Data sets

To obtain meaningful online data from Twitter, we used the key word “culture” in the respective languages used in the member countries of our project. For example, to obtain tweets addressing culture in Spain, the keyword “cultura” was utilized. The only exception in this strategy was Finland, in which a combination of the word “kulttuuri” and its 14 most relevant conjugations in singular was decided to be used (see Table A1 in the appendix). We determined that this step ensured more robust comparability between Finnish and other data sets. Generally, this procedure allowed us to download all tweets containing the word “culture” (or the corresponding language equivalents) in the tweet text or metadata in the chosen time frame.

To get more insight into the data availability and quality, we proceeded to download a preliminary sample of tweets in each country from two weeks in February 2021. With these samples and their analysis, we obtained a preliminary understanding of, among other things, relationships between hashtags connected to keyword “culture” in each country. During this phase, we also decided to change our data downloading strategy from using the hashtag “#culture” in respective languages to using the simple keyword “culture” without the hashtag. This yielded a significant increase in the number of available tweets in each of the countries. The results of these preliminary analyses confirmed that Twitter was a valuable source for investigation about culture. Based on this positive result, it was decided to proceed with more advanced analyses on historical Twitter data accessed by means of a combination of academic and premium accounts.

**Table 2.** *Scraping methods and number of tweets analyzed for each country*

Country	Scraping method	Number of Tweets <sup>1</sup>
Croatia	Twitter Premium account	12,334
Denmark	Twitter Premium account	54,286
Finland	Twitter Premium account	68,366
France	Academic research product track	51,935
Netherlands	Twitter Premium account	185,843
Serbia	Twitter Premium account	60,716
Spain	Academic research product track	78,908
Switzerland	Twitter Premium account	82,429
United Kingdom	Academic research product track	102,835

<sup>1</sup> Number of tweets analyzed from 2019-2020 with keyword “culture” in respective languages.

After sensitive planning and considering the optimal number of tweets in each of the countries, we proceeded to download full historical Twitter data sets for each country from January 2019 to December 2020. This time frame yielded the possibility of taking the impact of Covid-19 into account in the analysis, as it was approximately halfway in the time frame when the pandemic started in Europe. Table 1 presents the summary of data scraping methods and the number of tweets obtained in each of the countries. More detailed information on each country’s dataset is included in Table A2 in the appendix.

## 2. Pre-Processing

*Accessing and cleaning data.* After the data scraping, tweets have been pre-processed and analysed with the *ComTxt* package (Kim 2021) developed on the R language and Environment for Data Analysis (R-Core Team 2021). Twitter data scraped by the academic API account were rearranged to the data frame with 38 columns. For comparative purposes, all the column names were matched with tweets data scraped by Premium API with the R package *rtweet* (Kearney 2019). This assured that Twitter data scraped by academic API and premium API had the same format.

The first preprocessing step involved geographic information. Twitter provides two types of location. One is the location where the tweets are posted (i.e. place geographic information), and the other is the users' home location (i.e. profile geographic information). The geographic preprocessing captured the home location (city and country) provided by the users in their profile. Then we selected the tweets posted from users' home locations in each country. Tweets posted in countries that did not match the user's home location were removed. The R package *Map* (Deckmyn 2018) was used to obtain the exact geographic coordinates (i.e. latitude and longitude) according to profile city information. Additionally, we deleted tweets posted by bots (non-human Twitter users). This information is given by Twitter as a source such as 'Twitter bot for Mac' or 'Twitter bot for Android'.

The tweets in each country contained various languages, including the country's official language and English. We selected the tweets with their official languages and English for some specific countries. The other languages which had less than one-digit percentages were not included for the analysis.

Ultimately, the text data of the tweets were pre-processed in order to enable later text analyses. First, we removed Emoticons, URLs, hashtags, and mentions. Second, we removed stopwords and words that do not have any meaning such as prepositions. Since these words were different in each language, stopwords were written separately for each language with the R package *stopwords* (Benoit, Muhr, and Watanabe 2021). An important third step in text preprocessing is lemmatizing, which is a powerful and advanced computational text analysis technique. Lemmatizing reduces all words to their "lemma" using a lexicon in combination with regular conjugation rules. For instance, lemmatization returns "is" and "were" to their lemma "be". Since we analyzed at least nine different languages - including German, Dutch, French, and Spanish - lemmatization provided better results for textual preprocessing than stemming (Jacobi, van Atteveldt, and Welbers 2016; Haselmayer and Jenny 2017). Thereby, we processed the lemmatization for our data in all different languages. Lemmatization provided also part of speech (POS) tagging, which indicates the type of words (e.g. verb, noun, or preposition). Using this automatic filtering we selected noun, proper noun, and adjective. This selection process is known as the best strategy for topic modelling (Jacobi et al. 2016). After all these steps finally tweets were saved and ready for topic modeling analysis.

## 3. Topic Modelling

Topic modelling is a powerful and effective procedure for finding useful hidden structures in a collection of documents (Zhu, Feng, and Li 2018). Topic modelling is a computational linguistics model that uses a particular algorithm to identify recurring patterns of words in textual documents. In this way, topics can summarise and classify a large body of documents (Alghamdi and Alfalqi 2015; Jacobi et al. 2016). The output of topic modeling is a set of "bags" of words that are likely to appear often together in the same context according to an estimated probability (Barde and Bainwad 2017). These topics summarise the themes discussed in a collection of documents (Kim and López Sintas 2021). As a daily life example, documents are like shopping bags and their content are indicators of the shops at the market. The goal of topic modelling is to identify the shops from the content of

shopping bags (da Silva, Hruschka, and Hruschka 2014). This method is considered the most appropriate model for analysing a huge amount of text documents without manual coding, thus reducing time and costs significantly (Jacobi et al. 2016).

R package ComTxt (Kim 2021) applied *Mallet* for topic modelling analysis. *Mallet* (Machine Learning for Language Toolkit) is a Java-based package for Natural Language Processing or text mining. Mallet is used for text classification, text clustering, information retrieval, and topic modelling. The Latent Dirichlet allocation (LDA) model (Blei, Ng, and Jordan 2003; Mimno 2013) is employed to identify patterns of word co-occurrence in documents, which do not necessarily match theoretical concepts to infer topics (Jacobi et al. 2016).

To conduct a topic model it is required to decide how many topics,  $K$ , the LDA model should classify the words in the documents. A commonly used procedure to select the optimal  $K$  is the *perplexity* procedure, where a lower perplexity indicates a better prediction (Blei et al. 2003). To calculate the perplexity, the LDA model is trained on a portion of data and computes the perplexity index with several topics (Jacobi et al. 2016). Another way of calculating the optimal  $K$  number is using the topic coherence criterion. The topic coherence score measures the degree of semantic similarity between high scoring words in the topic. However, it is important to stress that these computational predictions should only be used as a guide to making an initial selection of the number of topics. The interpretability of topics is a more important criterion in establishing a meaningful division of topics (Jacobi 2015).

## Findings for Croatia

### 1. Twitter usage in Croatia

In this chapter we analyze how Twitter users frame cultural issues in Croatia. Twitter is the most popular platform for discussing and framing hot topics in contemporary societies and allows us to focus on specific actors that can be defined as opinion makers in Twitter space.

Croatia is a country in South East Europe, that has been a European Union member since 2013. In January 2021, it had a population of 4.09 million inhabitants – of whom 51.8% are female and 48.2% are male. The population is decreasing because of aging and increased post-EU accession because of migration of younger generations. The majority of the population live in cities (57.7%), while 42.3% live in rural areas (Kemp 2021).

The number of internet users in Croatia increased by 145 thousand (+4.6%) between 2020 and 2021 and in January 2021 amounted to 3.27 million, i.e., 80% of the population. The number of social media users increased with 4.8% between 2020 and 2021 to 2.80 million (January 2021), equalling 68.4% of the total population (Kemp 2021).

Twitter in Croatia is not among the most popular social media in general (Facebook prevails and Instagram among younger users), and especially not in cultural institutions, or among civil society organizations in culture (web pages, Facebook, and Instagram prevail). However, potential audience reach with an advertisement on Twitter is approximately 280,000 people (Kemp 2021). This allows us to focus on specific groups and individuals that increasingly taking over the role of opinion makers in the Croatian Twitter space and make comparable cross-national analyses to those for other countries included in the INVENT project.

### 2. Data description

Using Twitter's premium account API, we scraped tweets about culture in Croatia. The sample was defined by the keywords such as 'kultura', 'kulture', 'kulturi', 'kulturu', 'kultutama', 'kulture', 'kulturom' (declension forms of the word "kultura" in the singular). We collected data for two years (2019 and 2020). To answer the research questions, we compare results for two different periods: (1) before the Covid-19 pandemic in 2019 and (2) during the Covid-19 pandemic in 2020<sup>2</sup>.

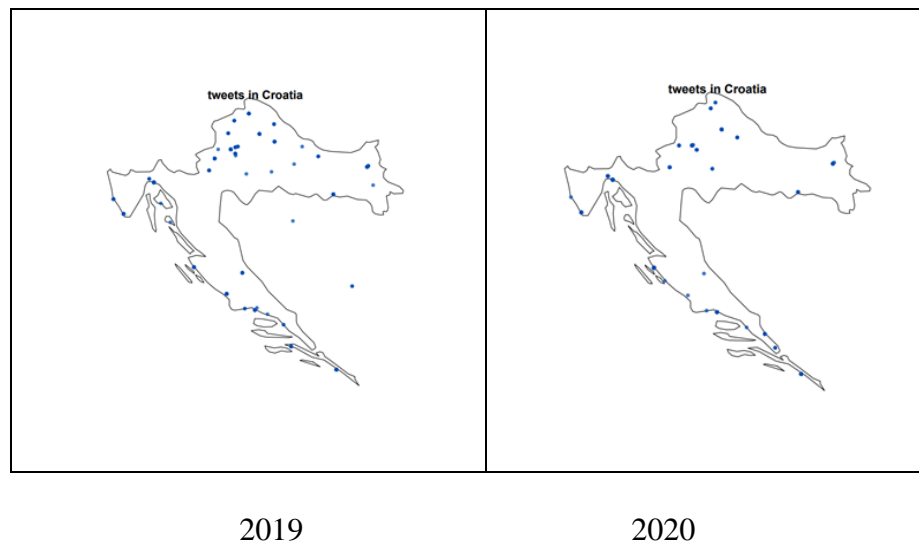
A total of 18,626 tweets were scraped initially. After filtering out retweets and geolocation information 12,687 remained. The tweets contained 32 different languages, but about 94% of the tweets was in Croatian. The other 6% included tweets in Serbian, Czech, English, German, Italian, etcetera. Only tweets written in the Croatian<sup>3</sup> language were further analyzed. The distribution of the profile location is presented in **Figure - HR - 1** and **Table - HR - 2**.

---

<sup>2</sup> Before Covid-19 period was defined from January 1th till December 31th 2019, and Covid-19 period was defined from January 1th till December 31th 2020.

<sup>3</sup> Tweets in Croatian language were set to 'undefined' when Twitter data were downloaded. We inspected the data to establish out which tweets were in the Croatian language.

**Figure - HR - 1.** Geographical location of the tweets about culture in 2019 and 2020 in Croatia ( $n = 7,800$ )



Most tweets about culture were sent from the capital city of Zagreb. A smaller number came from other urban centres in Croatia. Empty space on the map shows that almost no tweets with geographic location were found in more rural area and less developed regions (Lika-Senj County) in Croatia. For 2020, the profile location map is less diversified than for 2019.

**Table - HR - 1.** Top 10 geolocations in 2019 and 2020

2019		2020	
City	Freq.	City	Freq.
Zagreb	2,447	Zagreb	2,512
Rijeka	424	Zadar	555
Dubrovnik	278	Rijeka	395
Zadar	130	Dubrovnik	284
Osijek	92	Split	82
Split	63	Osijek	64
Križevci	39	Križevci	45
Pula	33	Šibenik	41
Vrgorac	12	Pula	36
Kastav	9	Bjelovar	24

### 3. User analysis

The analysis for 2019 included 6,421 tweets in total. For 2020, 5,913 tweets were included. In this section, we identify the key actors in defining the topics about culture discussed on Twitter. To answer the research questions about the most active users we did a comparative analysis for 2019 and 2020. An overview of the most popular users can be obtained from **Table – HR - 2**.



**Table - HR - 2.** Users with most followers and users with most tweets about culture in 2019 and 2020

2019				2020			
User	No of followers	User	No of tweets	User	No of followers	User	No of tweets
DNEVNIKhr	231,886	eZadar	479 (7,5%)	DNEVNIKhr	231,886	NinoslavSafaric	477 (8,1%)
24sata_HR	184,735	NinoslavSafaric	420 (6,5%)	24sata_HR	184,735	duvjesnik	209 (3,5%)
VladaRH	181,904	duvjesnik	230 (3,6%)	VladaRH	181,904	HanzoRed	140 (2,4%)
KolindaGK	114,448	studentskiHR	215 (3,3%)	radio_101	94,303	Rijeka2020	133 (2,2%)
radio_101	94,303	scena_hr	184 (2,3%)	novilisthr	90,169	GlasHrvatske	105 (1,8%)
novilisthr	90,169	gradrijeka	145 (2,1%)	novahr	40,385	lokalnahravska	92 (1,6%)
LaFamiliaFilm	87,794	HanzoRed	133 (1,9%)	ivospigel	26,756	BrankaStajner	75 (1,3%)
novahr	40,385	lokalnahravska	123 (1,9%)	N1infoZG	25,372	HRTkultura	68 (1,2%)
ivospigel	26,756	HDSZAMP	101 (1,6%)	jutarnjihr	25,308	PrigorskiH	68 (1,2%)
N1infoZG	25,372	Ludvig4	101 (1,6%)	mojHRT	23,466	Zady_direct	65 (1,1%)

Tweets from 1,897 unique users were collected. In 2019 and 2020, there were respectively 946 and 951 unique users.

#### *Users with the most followers*

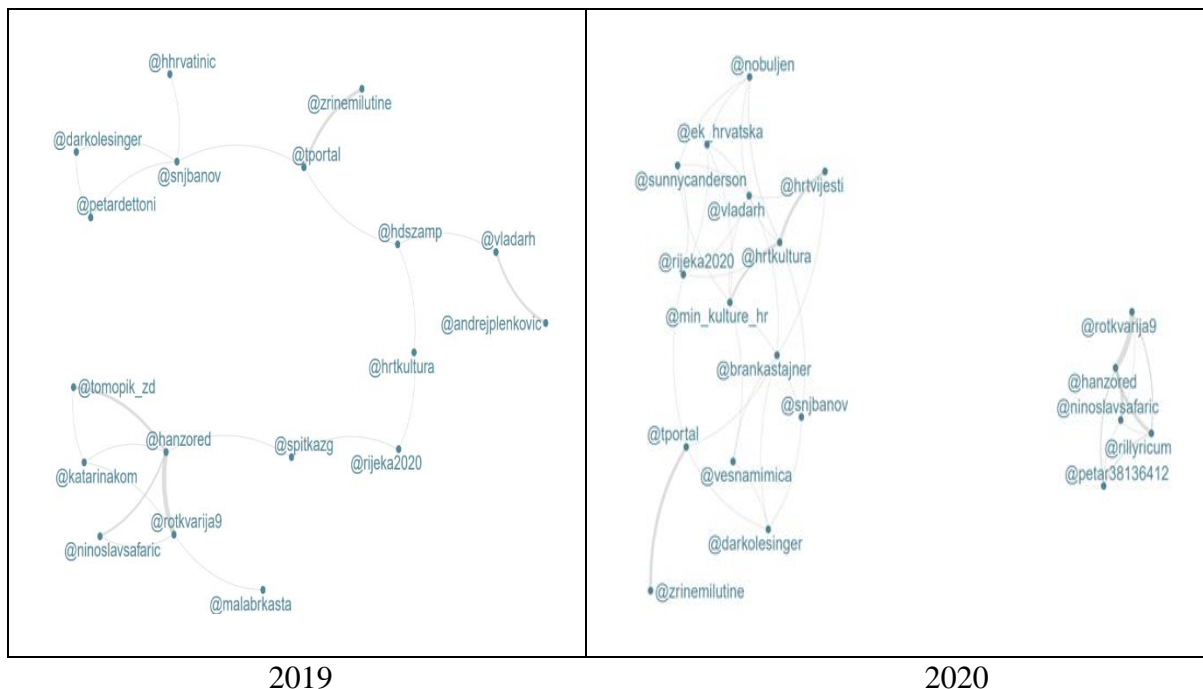
Comparing users with the most followers in the pre-Covid-19 period in 2019 and during the Covid-19 pandemic in 2020, we find similar results.

In both years, the most popular users (i.e., Twitter accounts) were those belonging to Croatian media companies, such as public television portals (mojHRT), commercial television portals (DNEVNIKhr, novahr), online newspapers with national (24sata\_HR, jutarnjihr), and regional (novilist, N1infoZG) topics, as well as a popular capital city radio station (radio\_101). Among the most popular users we also find the Croatian government's account and the former Croatia's president account (Kolinda Grabar Kitarević was still president in 2019). The private account of the businessman and columnist Ivo Špigel (ivospigel) was also popular.

**Figure - HR - 3** visualizes the strongest connections between accounts for the two periods in the forms of mentions.

As has already been established in the literature, Twitter is a social medium with a significant potential for news and journalism usage. Although it is one of the many platforms on which the news media operate, Twitter's architecture makes it more akin to a broadcast medium and it is therefore closely associated with the ongoing transformations in the news media industry (Doğu 2020). This is evident from the presented data on the most popular Croatian users.

**Figure - HR - 2. Network of mentioned user accounts in Croatia 2019 and 2020**



### Number of culture-related tweets

The most active user 'eZadar' (local news portal) is responsible for 7,5% of tweets about culture in 2019, mostly tweeting about local cultural events. In 2020 the most active user was NinoslavSafaric (book author, field astrology) who posted 8,1% of tweets in 2020 and around 6,5% tweets in 2019 (**Table – HR - 2**).

In general, we can identify several actors that stand out in defining Twitter discussions on culture in Croatia. The users who actively posted tweets related to keyword 'culture' are mainly local news portals (ezadar, duvjesnik, lokalnahravska, PrigorskiH), public television portals (HRTkultura), commercial television portals (novahr), online newspapers with national and regional focus (jutarnjihr, novilist), a central students' news portal (studentskiHR), independent news and lifestyle portal (scena.hr), the city of Rijeka (European Capital of Culture in 2020), as well as the institution HDSZAMP that represents authors and holders of music copyrights and royalties. Several private users (Ninoslavsafaric, Ludvig4, BrankaStajner, HanzoRed) were also represented.

### 3. Hashtag analysis

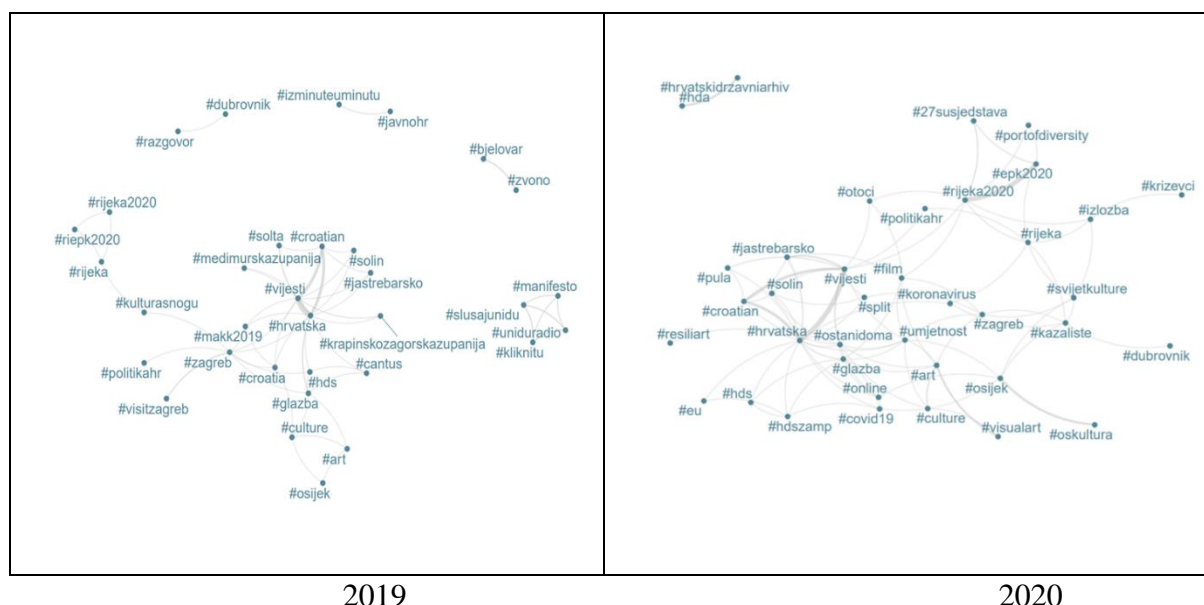
When users write about culture in Croatian Twitter space, they rarely use hashtags. More precisely, only around 12% (803 tweets) of total tweets in 2019 and around 13% (686 tweets) of all tweets in 2020 used hashtags. Top 15 hashtags for each year are presented in Table Cro\_3. Most hashtags are shared between 2019 and 2020. The three most popular hashtags in 2019 are #vijesti (news), #hrvatska (Croatia) and #zagreb (i.e., the name of the capital city). In 2020, the most popular hashtags remained the same, however with #rijeka2020 (the name of the city that was European Capital of Culture in 2020) taking the first place, and #epk2020 (European Capital of Culture 2020) taking the fourth place.

**Table - HR - 3.** The most popular hashtags except #kultura (#culture)' in 2019 and 2020

Hashtag 2019 (N=803)	Freq.	Hashtag 2020 (N=686)	Freq.
#vijesti	57	rijeka2020	51
#hrvatska	53	#vijesti	48
#zagreb	38	#hrvatska	44
#n1info	25	#epk2020	34
#krizevci	24	#zagreb	28
#rijeka2020	22	#osijek	21
#rijeka	20	#art	19
#croatian	19	#rijeka	17
#hds	18	#croatian	16
#medimurskazupanija	16	#jastrebarsko	16
#dubrovnik	15	#krizevci	14
#glazba	14	#najava	14
#zvono	13	#oskultura	14
#krapinskozagorskazupanija	12	#n1info	13
#express	10	#hrvatskidrzavniarhiv	12

**Table - HR - 3** shows that many of the 15 most popular hashtags refer to the cities and counties in Croatia which broadcast information and news about culture in their local areas.

**Figure - HR - 3.** Top hashtags network of tweets in Croatia in 2019 and 2020



To understand the hashtag network, we look at links between hashtags: the stronger the link between hashtags, the larger is the number of connections that a certain hashtag has to other hashtags. The network in **Figure - HR - 4** indicates that hashtags shown are largely not related to each other. The strongest connection in both periods is between the hashtags #vijesti, #croatia and #hrvatska. In 2020, the city of Rijeka was awarded the title of European Capital of Culture 2020 for its program 'The Port of Diversity', and that resulted in the connection between hashtags #rijeka2020 and #epk2020. In 2020, we can also see hashtags related to the pandemic like #Covid19,

#ostanidoma and #koronavirus.

#### 4. Keywords analysis

When looking at the most common keywords related to culture in Croatian Twitter space, we also differentiate between the years 2019 and 2020 (**Table - HR - 4**). The strength of the connections between the different keywords is visualized in two networks (see **Figure - HR - 4**).

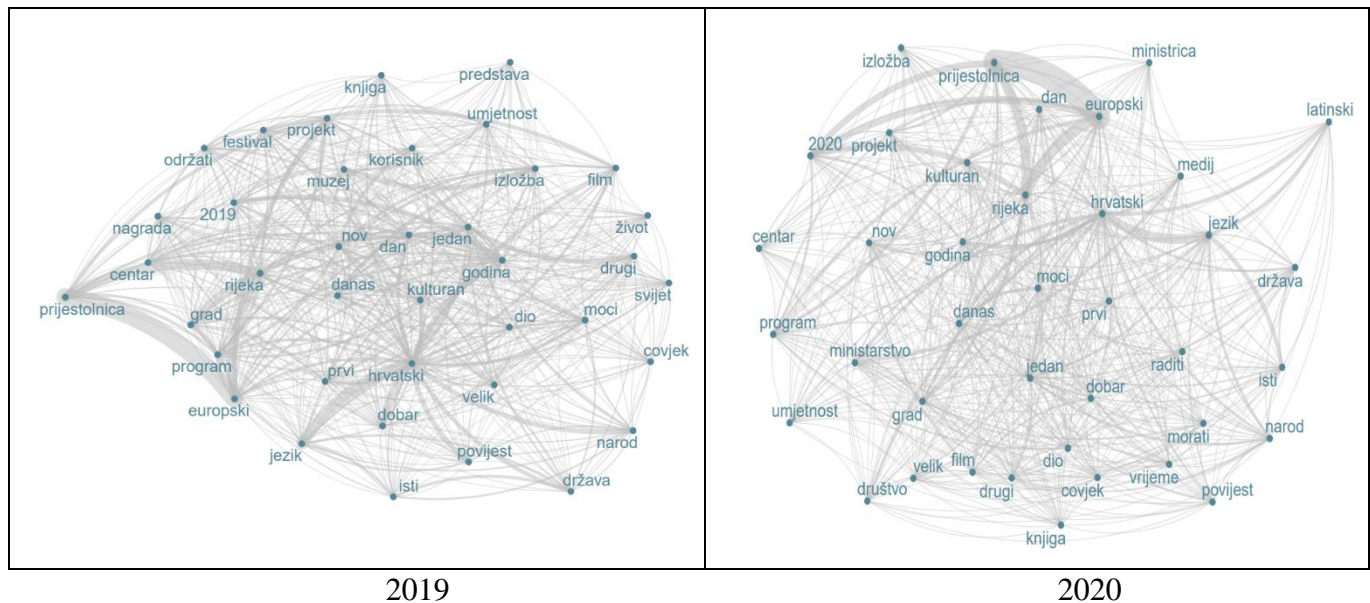
Keyword 'hrvatski' (Croatian) was the most frequent word connected with culture in both 2019 and 2020. As evident from **Table - HR - 4**, in 2020, when the city of Rijeka was the European Capital of Culture, the use of the word 'European' increased. It is interesting to note that, during the year 2020, keywords connected to the pandemic did not often appear in the tweets. For example, the word 'coronavirus' appeared only 35 times and the word 'pandemic' only 42 times. However, the word "Izložba" (Exhibition), which was the fifth most frequently used word in 2019, did not appear at all among the 15 most frequent words connected to culture in 2020 (for obvious, pandemic-related reasons).

**Table - HR - 4. Keywords for Twitter**

Keywords 2019 (N=31910)	Freq.	Keywords 2020 (N=30912)	Freq.
Hrvatski (Croatian)	859	Hrvatski (Croatian)	771
Godina (year)	441	Novi (new)	372
Novi (new)	425	Moći (to be able)	360
Dan (day)	369	Europski (European)	344
Izložba (exhibition)	346	Jedan (one)	327
Jedan (one)	306	Godina (year)	319
Korisnik (user)	302	Jezik (language)	311
Danas (today)	289	Prijestolnica (capital)	284
Grad (city)	287	Danas (today)	282
Moći (to be able)	275	Grad (city)	278
Jezik (language)	251	Rijeka (Rijeka – city)	276
Program (program)	249	Kulturan (cultural)	241
Europski (European)	243	Drugi (other)	217
Film (film)	233	Čovjek (human)	210
Projekt (project)	228	Program (program)	207

The keyword networks in **Figure - HR - 4** visualize a strong connection of the words 'prijestolnica' (Capital), 'program' (program), 'europski' (European) and 'Rijeka'. These words refer to the European Capital of Culture Rijeka. Other more visible connections are around the keyword 'hrvatski' (Croatian). This keyword connected to a variety of other terms in 2019, including 'jezik' (language), 'film' (movie), 'dobar' (good), 'narod' (people) and 'država' (country). In 2020, the same word ('hrvatski' - Croatian) connected to the words 'jezik' (language) and 'godina' (year).

**Figure - HR - 4.** *Top hashtags and keywords network of tweets in Croatia in 2019*



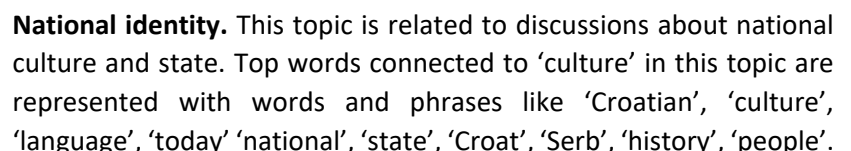
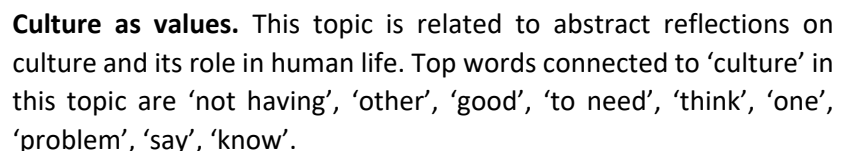
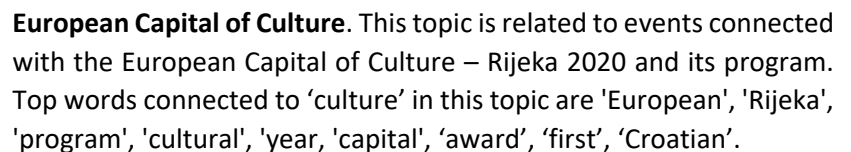
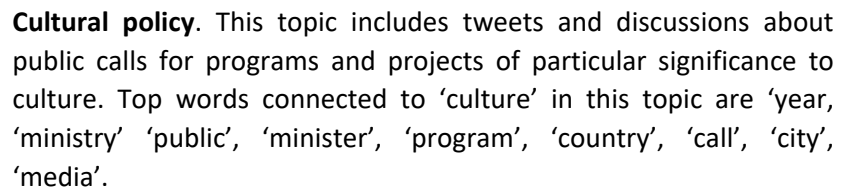
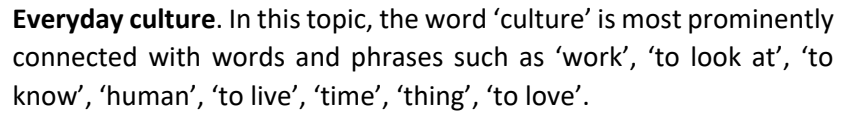
It should be noted that the tweets about the European Capital of Culture in 2020 remained well represented in spite of most of the program being cancelled due to the pandemic. Actually, the connection between the words “European”, “capital” and “Rijeka” formed the strongest connection in the visualization. This shows that the Twitter space served a purpose in disseminating the content related to the European Capital of Culture in extraordinary circumstances. It proved to be important for promotion of culture and cultural policy.

In general, among Twitter accounts discussing culture in Croatia three types of actors stand out: public and commercial media outlets, local and regional internet portals, and private users. They are all mostly located in urban centres. Media outlets and portals mostly disseminate news about culture, while private users comment on cultural topics more broadly and communicate their views regarding those topics.

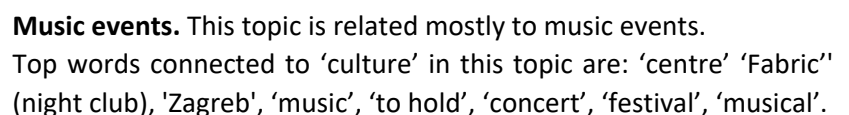
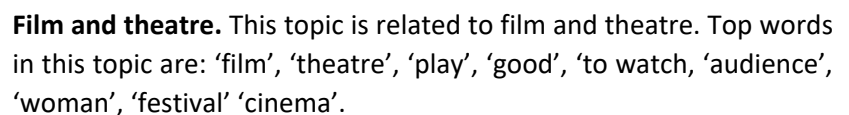
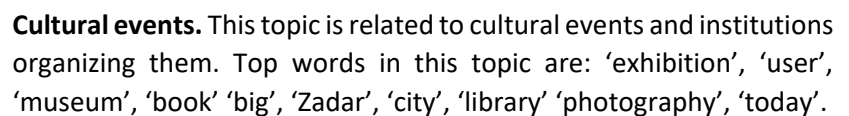
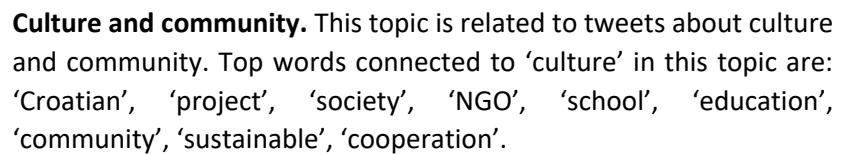
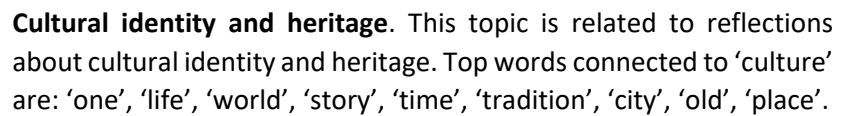
## 5. Topic modelling

This section presents the results of the topic modelling analysis, which was used to extract the topics within the Twitter data separately for the years 2019 and 2020. The results are presented in the following sub-sections. The content of each topic is presented as a word cloud, representing a collection of words that are semantically relevant. The optimal number of topics was extracted based on perplexity and coherence analysis. We extracted 12 topics for the year 2019 and 5 topics for the year 2020.

Each topic was labeled based on the 10 words with the highest probabilities that best reflect the subject matter of that topic.







## Topics 2020



**Art.** This topic is related to different fields of the arts. It is connected with words such as 'film', 'exhibition', 'art', 'museum', 'Zagreb', 'festival', 'program', 'project', 'year', 'workshop'.



**National identity.** The topic related to national identity connects the word 'culture' with words like 'Croatian', 'language', 'nation', 'state', 'Latin', 'today', 'history'.



**European Capital of Culture.** This topic is related to events connected with the European Capital of Culture – Rijeka 2020 and its program. Top words connected to ‘culture’ in this topic are ‘European’, ‘capital’, ‘Rijeka’, ‘cultural’, ‘city’, ‘program’, ‘exhibition’, ‘public’, ‘media’.



**Books.** This topic includes tweets and discussions about general topics connected with books and literature. The relevant words are ‘one’, ‘story’, ‘time’, ‘book’, ‘life’, ‘big’, ‘small’, ‘write’, ‘author’.



**Culture as values.** This topic is related to abstract reflections on culture and its role in human life. Top words connected to 'culture' in this topic are 'be able to', 'not having', 'know', 'human', 'other', 'to need', 'work'.

## 6. Discussion and conclusion

In this chapter we analyzed what Twitter users in Croatia communicated on culture in 2019 and 2020. Our exploratory analysis concentrated on three research questions: 1) mapping the key actors in defining the topics discussed; (2) defining the main topics and discourses in Twitter discussions on culture; and (3) identifying similarities and differences between the period before Covid-19 pandemic, and the period in which the Covid-19 became present.

As regards mapping the key actors, we identified three types of actors that stand out in defining Twitter discussions on culture in Croatia. These are: public and commercial media outlets, local and regional internet portals, and private users. We have also found out that these actors are mostly located in urban centres, as well as that media outlets and portals mainly disseminate news about culture, while private users comment on cultural topics more broadly.

As regards the main topics and discourses in Twitter discussions of culture, we have noticed that their scope was broader in 2019, that is before Covid-19 pandemic. In that year the topics included: everyday culture, cultural policy, European Capital of Culture, culture as values, national identity, cultural identity and heritage, culture and community, cultural events, film and theatre, and music events. In 2020, the topics were: art, national identity, European Capital of Culture, books, and culture as values.

During the Covid-19 pandemic, in the year 2020, topics related to various cultural events, as well as to culture and community, did not appear in Twitter space. In contrast to that, art and books appeared as topics. Abstract topics such as culture and values, and national identity were retained. European Capital of Culture tweets continued, although most of the program of this manifestation was cancelled, because the organizers kept their media presence also using Twitter. It is interesting to note that not only topics connected to events happening in public spaces disappeared, but also those connected to everyday topics, cultural policy, as well as cultural identity and heritage.

In general, what happens in the Croatian Twitter space mostly relates to transmission of news about culture and cultural events. Comments and discussions are less represented.

Compared to other countries, the topics specific for the Croatian Twitter space in 2019 and 2020 included the European Capital of Culture, and reflections on culture as value. In 2019, topics on cultural identity and heritage were represented, as well as community and culture (strongly connected to education and youth, but also including NGOs, as well as social, urban, and sustainable development).

Topics not represented in culture-related discussions in the Croatian Twitter space include: sports, business and organization, and cancel culture. Covid-19 did not appear as a separate topic in Croatian Twitter discussions of culture.

## Findings for Denmark

### 1. Twitter usage in Denmark

In Denmark, 7% of citizens above the age of 12 access Twitter every day (Schwartz and Scheutz 2021). 17% of Danes have a Twitter profile and access the platform to communicate with others who share their interests, to create awareness for topics or events they are interested in, and to voice their opinion (Slots- og Kulturstyrelsen, 2020). In 2021, 12% accessed the platform for any purpose, while 5% of Danes used Twitter specifically for news. Data from a 2015 survey (Andersen et al. 2021) show that of those Danes using Twitter, 53.8% are between the ages of 18 and 29, and approximately 20% are 30 to 40 and 50 to 69 years old, respectively. Only a few Danish Twitter users are older than 70 (5.6%), and 44.16% of them are female.

### 2. Data description

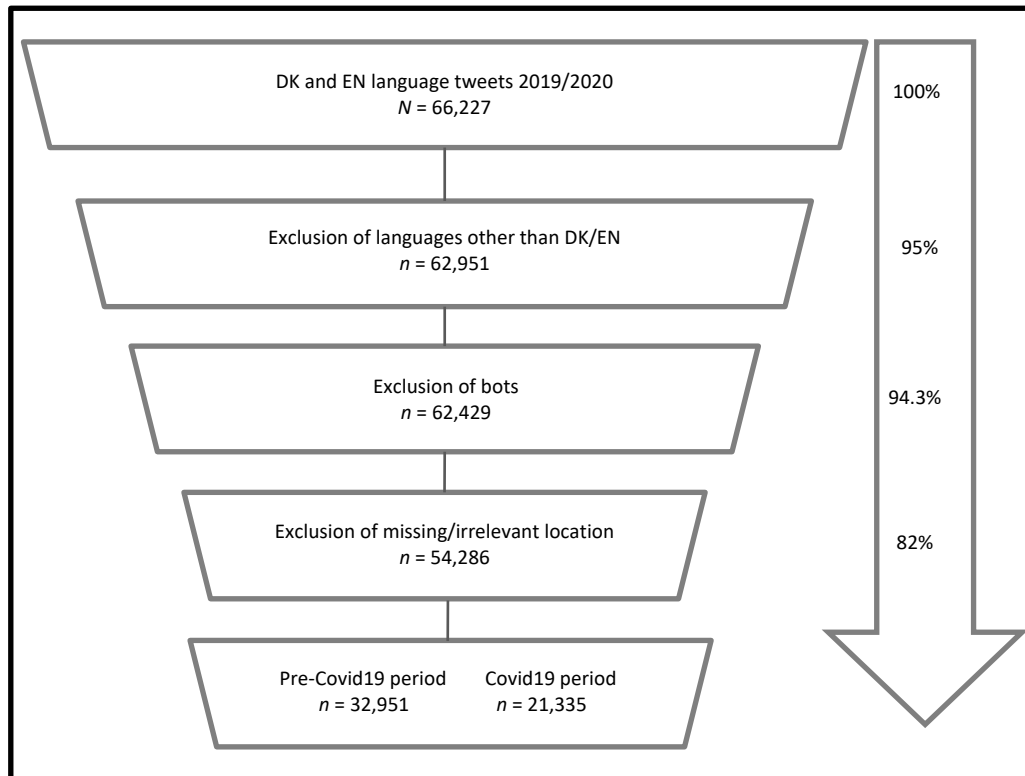
Using Twitter's Premium API, we scraped 48,238 tweets using the keyword "kultur" (Danish: culture) that were posted between January 2019 and December 2020. Given the importance of English as a posting language in the Danish Twittersphere (Laitinen et al. 2018), we added a query on "culture" for the same period, resulting in an additional 17,989 tweets for the two years. From this initial dataset, we excluded all cases ( $n = 3,276$ ) that Twitter's automated classifier identified as neither being written in Danish ( $n = 46,456$ ) nor English ( $n = 16,495$ ). That concerned mostly undefined tweets ( $n = 1,106$ ; e.g., only URLs) or texts in Norwegian ( $n = 864$ ), French ( $n = 311$ ), and Swedish ( $n = 271$ ). The remaining 62,961 cases were assessed for potential bots using Tweetbot, and another 522 tweets were removed from the dataset. The sample was filtered for users' profile location on the basis of the the location\_country operator provided by Twitter's Premium API; approximately 13% of tweets ( $n = 7,143$ ) were removed in this step.

In order to understand whether the Danish culture-related communication on Twitter differed before and during the Covid19 pandemic, we split the data according to a tweet's posting date. The first Danish citizen tested positive for the virus on February 27, 2020 (Statens Serum Institut, 2020), and we divided the data into tweets posted between January 1, 2019 and February 26, 2020 (pre-Covid19 phase;  $n = 32,951$ ), and February 27, 2020 and December 31, 2020 ( $n = 21,335$ ). **Figure - DK - 1** provides an overview of the data cleaning and preparation process.

### Geo-preprocessing

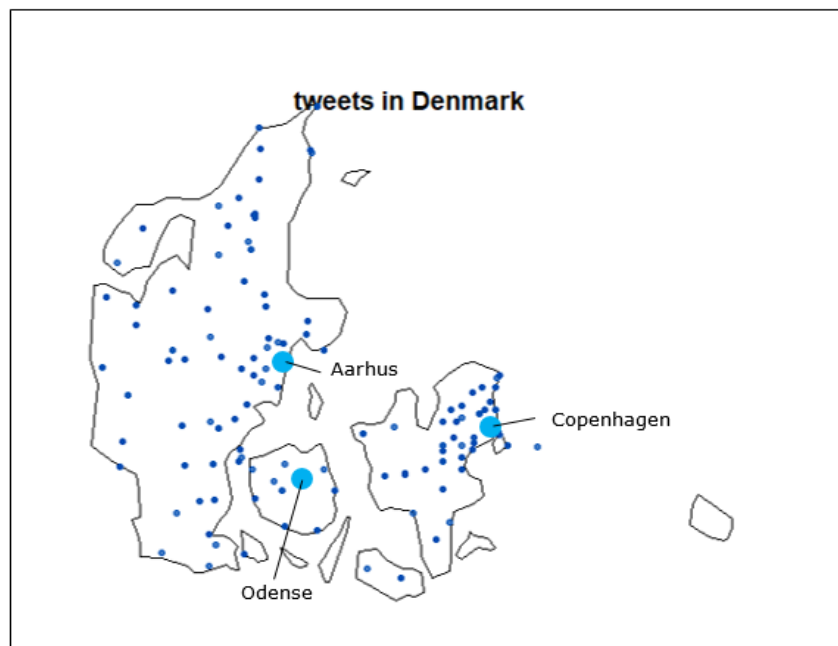
The majority of Danish 'culture'-tweets were sent from the capital Copenhagen (33%; see **Figure – DK -2**). When including the (individually listed) Copenhagen municipality Frederiksberg and the larger capital region ('Hovedstaden'), the share increases slightly further (35%). Another 3.4% and 1.7% of tweets were sent from the two second- and third largest cities, Aarhus and Odense, respectively.

**Figure - DK - 1. Steps taken in data cleaning and preparation**



**Note:** Percentages represent share of original data basis (100%).

**Figure - DK - 2. Geolocation of Danish- and English-language tweets in Denmark (n = 54,286)**



### 3. User analysis

In the following, we present separate user analyses for the two phases before and during the Covid19 pandemic (simplified to 2019 and 2020 to ease reader understanding). An overview of the most important account rankings can be obtained from **Table - DK - 1**, and

**Figure - DK - 3** and **Figure - DK - 4** visualize the strongest connections between accounts for the two periods in the forms of mentions. A total of 8,784 unique users can be identified in both periods.

#### Number of (culture-related) tweets

For 2019, the dataset includes posts from 5,783 unique Twitter accounts, with an average of 5.7 tweets per user between January 2019 and February 26 2020. During this period, the most culture-related tweets ( $n = 7,785$ ;  $>18$  tweets a day) were sent by @KulturNyt, an account that links to news on cultural topics posted on the websites of the main Danish news outlets (e.g., Politiken, DR). The related account @FilmPosten ( $n = 1,742$  culture-related tweets) focuses on information from the movie industry. Although both accounts post several times a day, they have few followers or friends and are thus only loosely connected within the Danish Twittersphere. In contrast, the largest Danish daily newspaper Politiken sent approximately 1.8 tweets per day related to culture, but reached a much larger audience (240,000 followers as per June 2021). Politiken has a clear cultural profile and can be considered a powerful cultural agenda-setter in Denmark (Kristensen and From 2018).

In 2020, 5,344 users in the Danish dataset posted an average of four culture-related tweets per account between the end of February and December. Newspaper Politiken increased its share of culture-related tweets during the Covid19 period and leads this list ( $n = 1,427$ ), while @Filmposten still holds the second place. The bot account @wallnot\_dk, which forwards newspaper articles that are normally hidden behind a paywall, moved up in this list ( $n = 372$  culture-related tweets, from previous 9th to 3rd place). Wallnot, similar to Filmposten, has only few followers but is closely connected to Politiken from which's website the bot regularly posts content. This connection is also visible in the network visualization of mentions (see

**Figure - DK - 3** and **Figure - DK - 4**). With the exception of Politiken, we can document an overall decrease in the number of culture-related tweets by the Top10 accounts during 2020 compared to the pre-Covid19 period (see

**Figure - DK - 3**, right). Presumably, the largest Danish news media focused on disseminating different types of information during these months, and/or had less culture-related content to share due to a lack of cultural events. The TOP10 user accounts in terms of the number of culture-related tweets (i.e.,

**Figure - DK - 3**, right-hand columns) accounted for 37.5% (2019) and 18.6% (2020) of all posts including the words 'culture' or 'kultur'. The heavily skewed distribution in the pre-Covid19 period is in large part attributable to @KulturNyt; the user account functions as a news aggregator across different news media organizations (see above).

The most statuses overall (i.e., not limited to content about culture) were posted by users of unverified (private/personal) accounts. In 2019, these included, for example, "Happy Life Expert" @JanAnkerstjerne (419,838 posts), human rights activist @hamrahshowaziz ( $n = 314,426$  posts), pensioner and tennis-fan @markalannixon (263,774 posts), food blogger @ldahNielsen ( $n = 260,227$  posts), and bold.dk, a Danish platform for news on national and international football. In 2020, @hamrahshowaziz and @markalannixon still feature in the list, but it is led by wrestling enthusiast @Allan\_Straarup\_ ( $n = 470,313$  posts) and also includes Manga/Anime fan @artemisthemp ( $n = 210,805$  posts). In general, we see that users who post many tweets often identify with a sports-



related fandom.

**Table - DK - 1.** Comparison of Danish culture-related user accounts before and during the Covid19 pandemic with regard to the number of followers and tweets about culture

Pre-Covid 19 period n = 32,951			Covid19 period n = 21,335		Pre-Covid 19 period n = 32,951		Covid19 period n = 21,335	
	User	No. of followers	User	No. of followers	User	N° of tweets about culture	User	N° of tweets about culture
1	MrLachris	506,137	sybogames	989,412	KulturNyt	7,785	politiken	1,427
2	tv2newsdk	288,008	LEGO_Group	797,372	FilmPosten	1,742	FilmPosten	1,017
3	dupreeh	272,829	DRBreaking	384,086	politiken	749	wallnot_dk	372
4	AndersFoghR	268,721	tv2newsdk	288,008	Hillerod_Posten	459	KulturNyt	290
5	Fedeabe	263,736	Fedeabe	263,736	ugeposten	384	videnskabdk	187
6	politiken	240,618	politiken	240,618	videnskabdk	298	Heeger	172
7	lasserimmer	222,293	lasserimmer	222,293	jyllandsposten	253	forskerzonen	152
8	larsloekke	200,109	LEGOIdeas	190,527	wallnot_dk	251	Loafertolboll	138
9	kamper	188,921	kamper	188,921	Pernillebendix	219	jyllandsposten	125
10	WHO_Europe	161,325	WHO_Europe	161,325	PRINTSVONJANE	212	BoergeOf61	94

**Note:** Highlighted accounts appear in both periods.

## Number of followers and friends

The account with the largest number of followers (n = 506,137) tweeting about culture in the pre-Covid19 phase belongs to Michael Christensen @MrLachris, followed by the Danish broadcaster TV2 (@tv2newsdk, n = 288,008 followers) and the professional Danish CS:GO player/gamer Peter Rasmussen (who is part of the internationally famous Astralis-team; @dupreeh, n = 272,829 followers). The list of top accounts with respect to followers also features, for example, former Prime Minister of Denmark Anders Fogh Rasmussen, the newspaper Politiken, and the regional office of the World Health Organization @WHO\_Europe. During the Covid19 period in 2020, the distribution looks slightly different: The Danish gaming company @sybogames now leads the list of accounts with the most followers (n = 989,412), with the Danish Lego group coming second (n = 797,372 followers). Several media outlets (DR, TV2, Politiken) and the WHO remain among the TOP10 accounts with regard to follower numbers, as does the Danish journalist and comedian Anders Lund Madsen (n = 263,736 followers). Notably, the public broadcasters DR and TV2 have a public service obligation to fulfill with regard to cultural news and information.

The list of Twitter users that follow the most other accounts (constant in both periods) includes, for example, @kamper (n = 57,091 friends) who posts about gaming and the movie industry, Fantasy author @SchmidtJesper (n = 45,758 friends), software developer @carstenbreum (n = 18,442 friends), and 'social media and online marketing expert' @LedaGRasmussen (n = 14,080 friends).

Overall, the analysis of Danish users' Twitter activity related to culture shows a mix of very active private/individual accounts posting about sports and fandom, established media institutions, and prominent industry representatives (LEGO, gaming companies). We also see a mix of accounts specializing in culture as art and industry (e.g., movies and books), popular/commercial culture, or broadly categorized in culture as life. @Videnskabdk, an account offering 'the best and most important' research news, and @forskerzonen where Danish-based researchers disseminate their knowledge and expertise for the benefit of society, frequently mention each other, but are rather

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----



Downloaded from <http://ajphaphysocpharm.sagepub.com/> at 10:06 11 November 2014

However, when it comes to user interaction and engagement, it is worth noting that these examples constitute the exception rather than the rule: The majority of tweets in our data received no likes (54.7%), retweets (84.8%), or replies (72.5%). Similarly, the majority of 69.4% tweets did not constitute a response or reaction to a previous post (e.g., by a different user), but were independent from any discourse. This confirms the notion that users in the Danish cultural Twittersphere interact very little with each other, but mostly use the platform to communicate their own opinion, or to share something they consider (news-) worthy (see also Slots-og Kulturstyrelse 2020).

At the same time, user engagement seems to increase between 2019 and 2020, including with regard to the respective average number of likes (3.02 versus 6.88), retweets (0.41 versus 0.73), quotes (0.09 versus 0.17), and replies per tweet (0.49 versus 0.78). This may be attributable to the Covid19 situation and the fact that citizens in Denmark spent more time on social media platforms in general and more frequently accessed Twitter to interact with others as well. Unfortunately, our data does not allow us to draw any inferences about the extent to which this increase in user engagement is specific to content on 'culture', a sign of a broader shift in communication patterns in the country during the pandemic, or a reflection of increased popularity of the platform in general.

**Figure - DK - 5. LEGO\_Group status post 'Grand Piano'**



Source: <https://bit.ly/3xsLDsp>

### 3. Hashtag analysis

In preparing the dataset for the language preprocessing, we first removed duplicate entries (n = 381) and then separately excluded stopwords (source: stopwords-iso) from the Danish (170 stopwords)) and English language tweets (1,298 stopwords). For most of the analyses described below, we continue to differentiate between the pre-Covid19 period of 2019 and early 2020 and the Covid19 period during the remaining months of 2020. However, we do not examine English- and Danish tweets separately, unless explicitly stated otherwise.

**Table - DK - 2. Hashtag statistics 2019/2020 Denmark**

	Pre-Covid19 period	Covid19 period
Total number of hashtags	4,671	2,370
Number of unique hashtags ( <i>n</i> )	2,002	1,114
Percentage of tweets incl. hashtags	8.54%	7.05%
Average number of hashtags	1.7	1.6
Max. number of hashtags	9	9

The use of hashtags is not very common in the Danish cultural Twittersphere; on average, only around 8% of all tweets include (at least one) hashtag(s). **Table - DK - 2** provides an overview of the two periods under investigation, with few discernible differences in the use of hashtags between 2019 and 2020.

**Table - DK - 3** lists the TOP10 hashtags in Denmark for 2019 and 2020. Given the overall infrequent use of hashtags, we kept English- and Danish-language tweets in the same dataset, but excluded both ‘#kultur’ and ‘#culture’. The country-specific #dkpol, #dkmedier, and #dkkultur hashtags refer to Danish political, media-related, and cultural policy/topics, respectively. A weekly newspaper in the Danish region Zealand (Sjælland) with the same name (@ugeposten) uses the #ugeposten hashtag. #metoo ranks highly in both years, and during 2020, different terms (e.g., #Covid19dk) refer to the pandemic. However, it should be noted that the frequencies and associated percentages are marginal; although we talk about the TOP10 hashtags related to culture in Denmark, very few are employed in more than 5% of tweets that include (at least one) hashtag(s).

**Table - DK - 3. TOP10 hashtag frequency and percentage Denmark, excluding “culture”, “kultur”**

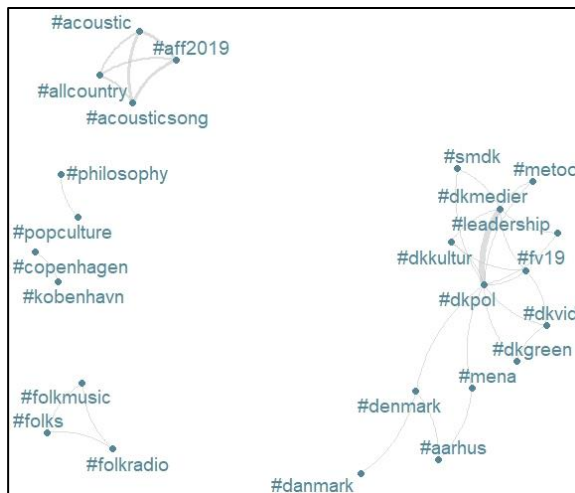
Pre-Covid19 period				Covid19 period		
Hashtag	Freq.	Percentage within tweets incl. hashtags ( <i>n</i> = 2,768)		Hashtag	Freq.	Percentage within tweets incl. hashtags ( <i>n</i> = 1,483)
1 #ugeposten	379	13.7%		#dkpol	241	16.3%
2 #dkpol	276	10%		#dkmedier	114	7.7%
3 #dkmedier	165	6%		#metoo	99	6.7%
4 #acoustic / #acousticsong	62	2.2%		#philosophy	39	2.6%
5 #metoo	59	2.1%		#Covid19dk	38	2.6%
6 #denmark	57	2.1%		#denmark	38	2.6%
7 #dkkultur	50	1.8%		#dkkultur	36	2.4%
8 #aff2019	31	1.1%		#popculture	33	2.2%
9 #fv19	29	1%		#Covid19	25	1.7%
10 #aarhus	23	0.8%		#coronavirus	13	0.9%

**Note:** Both Danish- and English language tweets are included; highlighted hashtags appear in both periods.

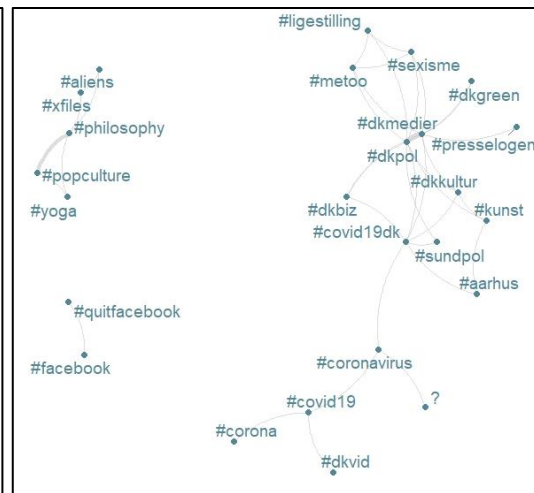
The scarcity with which hashtags are used in the Danish cultural Twittersphere also reflects in the hashtag networks (see **Figure - DK - 7** and **Figure - DK - 6**). Few, if any, connections can be observed between the different conversational areas. One exception is the link between the hashtags #dkpol and #dkmedier, both of which are also widely used in other contexts by Danish Twitter users. In

2019, tweets related to acoustic- and folk music formed a connected issue each, whereas the corona virus and Facebook were discussed with common hashtags in 2020.

**Figure - DK - 7.** Hashtag network Denmark pre-Covid19 period



**Figure - DK - 6.** Hashtag network Denmark Covid19 period



#### 4. Keyword analysis

When looking at the most common keywords, we also differentiate between Danish and English language tweets and visualize the strength of the connections between the different keywords in four networks (see **Figure - DK - 8**). As can be seen in **Table - DK - 4** and **Table - DK - 5**, many keywords in the two languages overlap in each period, indicating that the English- and Danish language cultural Twittersphere in Denmark discuss similar topics. Tweets in Danish refer to movies more frequently both in 2019 and 2020, whereas English posts address philosophy more often than their Danish counterparts. In both languages, references to 'Danish' and 'Denmark' are frequent, as are discussions relating to women. Possibly, the latter connects back to the aforementioned importance of the “#metoo” debate in the country. With the exception of 'Danish' and 'people', none of the keywords appears in more than seven percent of the respective tweets.

**Table - DK - 4.** TOP15 Danish- and English language keyword frequency and percentages, **pre-Covid19**

	Danish keyword	Translation	Freq. (n)	Percentage within Danish tweets (n = 25,483)	English keyword	Freq. (n)	Percentage within English tweets (n = 7,247)
1	dansk	danish	3,135	12.3	people	851	11.7
2	film	movie	1,736	6.8	danish	371	5.1
3	danmark	Denmark	1,004	3.9	time	351	4.8
4	kvinde	woman	953	3.7	country	290	4
5	bog	book	898	3.5	article	240	3.3
6	barn	child	882	3.5	company	223	3.1
7	verden	world	866	3.4	denmark	217	3
8	historie	history/story	846	3.3	pop	214	3
9	ung	young	826	3.2	history	198	2.7
10	tid	time	783	3.1	woman	196	2.7
11	politi	police	782	3.1	language	177	2.4
12	liv	life	752	3	life	148	2
13	land	country	745	2.9	american	144	2
14	hjerne	heart	732	2.9	philosophy	144	2
15	folk	people	721	2.8	change	142	2

**Note:** Highlighted keywords appear in both the Danish and English-language dataset in the same period; “kultur” and “culture” were excluded. Percentages assume that every word occurs a maximum of one time per tweet, but multiple occurrences are possible.

**Table - DK - 5. TOP15 Danish- and English language keyword frequency and percentages, Covid19 period**

	Danish keyword	Translation	Freq. (n)	Percentage within Danish tweets (n = 15,200)	English keyword	Freq. (n)	Percentage within English tweets (n = 5,975)
1	dansk	danish	1,591	10.5	people	778	13
2	kvinde	woman	703	4.6	time	356	6
3	artikel	article	633	4.2	cancel	316	5.3
4	film	movie	629	4.1	danish	249	4.2
5	folk	people	518	3.4	pop	249	4.2
6	tid	time	496	3.3	country	248	4.2
7	danmark	Denmark	492	3.2	article	195	3.3
8	land	country	488	3.2	philosophy	187	3.1
9	menneke	human being	470	3.1	life	179	3
10	historie	history/story	467	3.1	denmark	178	3
11	problem	problem	464	3.1	history	168	2.8
12	vigtig	important	412	2.7	white	162	2.7
13	verden	world	389	2.6	company	154	2.6
14	sammen	together	385	2.5	american	150	2.5
15	barn	child	379	2.5	cultural	144	2.4

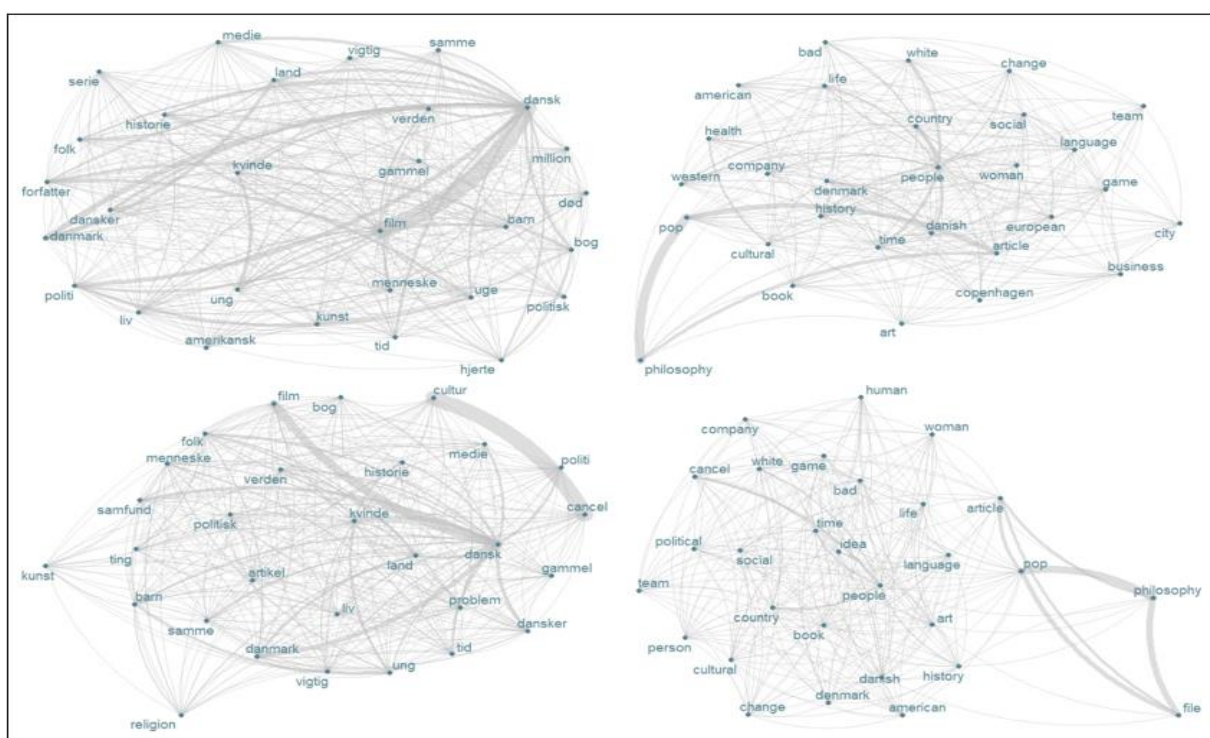
**Note:** Highlighted keywords appear in both the Danish and English-language dataset in the same period; “kultur” and “culture” were excluded. Percentages assume that every word occurs a maximum of one time per tweet, but multiple occurrences are possible.



The keyword networks in **Figure - DK - 6** visualize the centrality of the attribute ‘Danish’ in the Danish language tweets (left-hand side); the keyword connects to a variety of other terms including police, author (‘forfatter’), TV show (‘serie’) and movie (‘film’). In 2020, the term ‘cancel’ is strongly connected to what can be assumed to be ‘culture’ (but without the last ‘e’) in Danish conversations. Note that keywords were analyzed for individual words only and do not include word combinations here, and both ‘kultur’ and ‘culture’ were excluded from the keyword analysis to facilitate interpretation of the networks. When including ‘culture’ in the 2020 English keywords network, we can indeed observe a strong connection between ‘cancel’ and ‘culture’ here as well. Therefore, we can presume that “cancel culture” played an increasingly important role in the Danish cultural Twittersphere in 2020 – irrespective of the language used in the tweets.

The English data (right-hand side) confirm the relevance of the term ‘philosophy’, although the networks also show that the keyword is not central, but strongly connected to only few other terms (e.g., ‘file’, ‘pop’). The most important keyword in terms of its centrality in the network is ‘people’. When comparing the pre-Covid19 (top row) and Covid19 period (bottom row), we see few differences in the general organization of the networks, with the exception of the ‘cancel culture’ connection and a sudden increase of discussion connections related to ‘religion’ in the Danish language tweets in 2020.

**Figure - DK - 8.** TOP30 Danish- and English language keyword network per period



**Note:** From top to bottom, left to right: Danish language tweets pre-Covid19 period, English language tweets pre-Covid19 period, Danish language tweets Covid19 period, English language tweets Covid19 period.



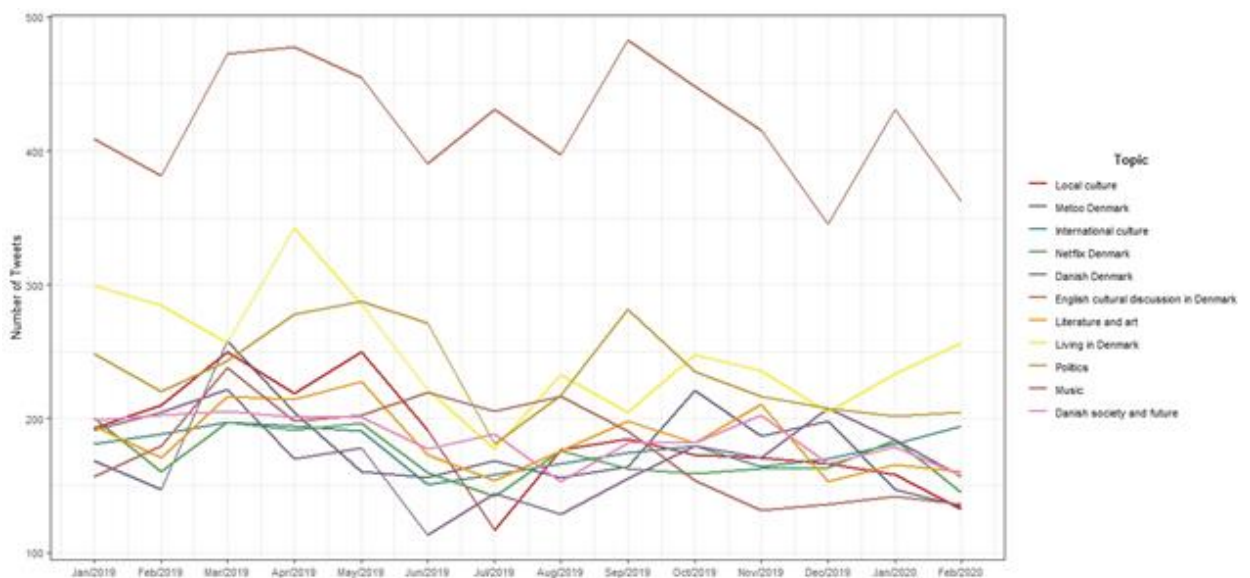
## 5. Topic modelling

In the keyword analysis described above, we saw numerous overlaps among the TOP15 most frequently used English- and Danish language terms for 2019 and 2020. On the one hand, this may indicate that users tweeting in the two languages do not form entirely distinct issue publics, but discuss similar topics. On the other hand, some keywords are unique to content in English, suggesting that there are at least some issues that are debated in one language but not the other. Therefore, we continue our analysis of the Danish data by including both the Danish- and English language tweets simultaneously in the topic modelling, but do not analyze them individually. As before, we look at the pre-Covid19 and Covid19 period separately in order to understand whether either the amount or type of topics that are discussed with relation to culture have changed during the pandemic.

For both periods, perplexity scores decrease with increasing number of topics  $k$ , which indicates a better model fit the more topics are differentiated (Jacobi et al. 2016). Topic coherence varies considerably across the tested models, but is highest for 11 or 28 topics for the pre-Covid19 phase. For the sake of parsimony and clarity, we report 11 topics here. During the remainder of 2020, coherence scores are highest for six topics, and we continue with this solution.

**Figure - DK - 9** and **Figure - DK - 10** provide an overview of the prevalence and development of each topic over the course of the two periods. In the following, we visualize the most important keywords within the topics in a wordcloud and briefly discuss their content.

**Figure - DK - 9.** Development of topic prevalence over time, pre-Covid19 period 2019



[illegible]

6. Topic 6 only includes English language words and summarizes a variety of cultural terms in reference to Denmark and its country and people. It is also by far the most prominent topic in 2019 (see **Figure - DK - 9**). Due to its language focus, we name it '**English cultural discussion in Denmark**'.

[illegible]

A word cloud of terms related to digital media and politics. The most prominent words are 'vigtig', 'medie', 'social', 'artikel', 'kritik', 'valg', 'politisk', and 'politiker'. Other visible words include 'digital', 'problem', 'ansvar', 'kulturminister', 'debat', 'arbejde', 'støtte', 'parti', 'penge', 'kunst', 'brug', 'ledelse', 'regering', 'abonnement', 'facebook', 'forskning', 'offentlig', 'spørgsmål', 'nyhed', and 'valgkamp'. The words are arranged in a dense, overlapping cluster with varying font sizes and orientations.







4. Similar to 'Cinema and TV', topic four merges discussions on previously separate cultural conversations. It includes references to story/history, art (kunst), music, sport/football, and the minister of culture. Following the previous categorization in topic 1, we call it **'Danish cultural discussion in Denmark'**.

5. The Metoo debate is summarized in a unique topic again, in which we find frequent mentions of woman (kvinde), problem, critique, important (vigtig), political (politisk), bad (dårlig), sexist, and sexism. Like in 2019, this topic is **'Metoo Denmark'**.



6. Finally, topic six during the pandemic on 'Living in Denmark' is analogue to its predecessor in 2019, but it also references elements associated with the previous 'Danish society and future' topic (which is not too surprising giving their similarities). For 2020, we keep the label **'Living in Denmark'**.



## 6. Discussion and conclusion

Overall, we can identify a number of similar discussions taking place in the Danish Twittersphere before and during the Covid-19 pandemic. Notably, in each period, one English-language topic dominated the cultural debate on Twitter. These have a rather broad focus on culture in general and are hence 'catch-all' issues specific mostly in their language use. We also see that the debate on #metoo and sexism in Denmark took place on Twitter, not least during a time in late 2020 when allegations of sexual misconduct in the largest Danish media outlets (DR, Politiken) surfaced and received wide societal attention (see the spike in **Figure - DK - 10**). Religion is also a recurring although not distinct topic, specifically in reference to Islam and Muslim citizens. On the one hand, it should be noted that it is difficult to make any inferences about the negative/positive sentiment of this (or really any) discussion in our data due to the lack of sentiment or emotion in the language. On the other hand, the co-occurrence of the word 'racism' in this context may aid understanding.

Some of the topics that were discernible in 2019 did not reappear in 2020, but were merged to larger issues. In general, these are more difficult to interpret, and we believe it necessary to complement the computational approach followed here with qualitative assessments, for example by analyzing individual tweets within the clustered topics to gain a better understanding of their meaning and the initial events or occasions. Additional analysis should take a closer look at the developments of the different topics over the two-year period (i.e., as initiated in the visualizations of Figure DK\_12 and DK\_13), which could further aid interpretation of which events 'make' or 'break' a cultural discussion on Twitter in Denmark.

Many topics that could be expected to be discussed in relation to culture (e.g., sports, gaming, travelling, food) do not appear at all, or are part of broader themes and mentioned only in passing. Aside from frequent discussions of Netflix shows (which have an important standing in Denmark), the majority of culture-related tweets seem to refer to traditional 'high-art' forms and/or policy content, with only few popular culture or lifestyle aspects. At the same time, we saw that the most popular Danish Twitter users post about gaming. These seemingly inconsistent findings should be

further investigated as well.

One aspect that so far seems distinct to the Danish Twittersphere is the use of very broad country-specific hashtags that are applied in a variety of different context (e.g., #dkpol, #dkkultur). Rather than creating individual hashtags that fit a newly emerging topic, Twitter users in Denmark seem to rely on these established combinations – if they use hashtags at all. As mentioned, this is only the case in eight percent of all tweets collected during the period under investigation. It is difficult to assess the extent to which this is a phenomenon unique to discussions about culture, but given the comparably low dissemination of the platform in Denmark, it is possible that there is simply no established Danish hashtag-culture on Twitter.

Lastly, we noted that only few interactions take place within the Danish Twittersphere. Posts by private users often appear isolated, and the majority of tweets are neither liked nor retweeted or answered. Hardly any data on Danish Twitter users are available, and it would be worthwhile to widen the methodological approach applied here to include (both qualitative and quantitative) interviews with users in order to understand their motivations for accessing the platform, following other accounts, and posting themselves.



## Findings for Finland

### 1. Twitter usage in Finland

In 2017, 11% of Finnish people older than 10 years used Twitter. Twitter usage was more common among men (14 %) than women (8%). Twitter usage was most common among young people of ages 15–24, of whom about 20% were using Twitter. Twitter use diminishes with age, and in the age group of 55–64-year-old people, only 6% were using Twitter. Among occupational groups, Twitter was most popular among higher officials and students, of whom about 20% used Twitter. It therefore seems that the Finnish Twittersphere is occupied by predominantly by males, younger people and people in good occupational positions (Kohvakka and Saarenmaa 2019).

### 2. Data description

Tweets were retrieved from Twitter Full Archive endpoint, which required subscription of premium API access. To obtain the data, we used two separate queries. For Finnish tweets, the keyword “kulttuuri” and the 9 most relevant conjugations (see Appendix FI-1) in singular were used. For Tweets in English, the keyword “culture” was used. Both searches were made utilizing the “place\_country:fi” operator, available with the Premium Full Archive account. Retweets were excluded from the sample by restricting the search with “-is:retweet” operator.

This strategy yielded 93,691 tweets with the Finnish keyword query and 24,506 with the English keyword query, composing a data set with the total size of 118,197 tweets. While the total number of languages present in the data set is 41, a total of 97 % of the tweets were written in Finnish (79%) and English (18%). Therefore, we will exclude tweets that are not classified as Finnish or English. This yields a total of 115311 tweets.

Despite acquiring datasets with both Finnish and English keywords for culture, we found that 1) merging both Finnish and English datasets did not provide good results with topic modelling and 2) conducting the analysis separately for Finnish and English datasets did not yield enough meaningful insight to justify the inclusion of the English analysis to the final report, as it would virtually double the length of the report. Therefore, we decided to present only the analysis conducted with the data acquired by Finnish keywords for culture.

To begin our analysis of tweets mentioning “kulttuuri” and its most important conjugations, we split our dataset to compare content of the tweets before and during the Covid-19 crisis. To achieve this, we split our data sets to consist of tweets posted before and after the date 29.1.2020, which was the day the first Covid-19 infection was found from Finland (Yle Uutiset 2020). After data cleaning and deletion of unfit tweets for the analysis, the total number of Finnish language tweets reduces from 93,691 to 68,366. Splitting the data to correspond pre- and during Covid-19 yields two datasets: 36,814 tweets for before the first infection (hereafter referred also as 2019) and 31,552 tweets for after the first infection (2020).

### 3. User analysis

There are 14,805 unique users in the data set. We did not find notable differences between pre- and during Covid-19 period among active users in terms of, for example, activity, and we will present user analysis without considering the distinction of pre- and during Covid-19 period. The average number of tweets posted by each user is 6.31, but there are some significantly more active users. The most active users were news media accounts such as the culture sections of Karjalainen (@KarjalainenFiK), Kouvola Sanomat (@KouvolanSanomat) and Kymen Sanomat

(@KymenSanomat). Some other active institutional accounts, besides news media, that are tweeting about culture include accounts of the capital of Finland, @helsinki, the Ministry of Education and Culture @okmfi, and the University of Helsinki, @helsinkiuni. The undoubtedly most active user of the dataset was the user @kulttuurisi with almost 9,000 tweets in the span of two years. This user is a bot that shares content related to culture from other users and media sites and is excluded from the forthcoming analysis. With this user excluded, 10 most active users in the data created 18 % of the tweets in the whole dataset.

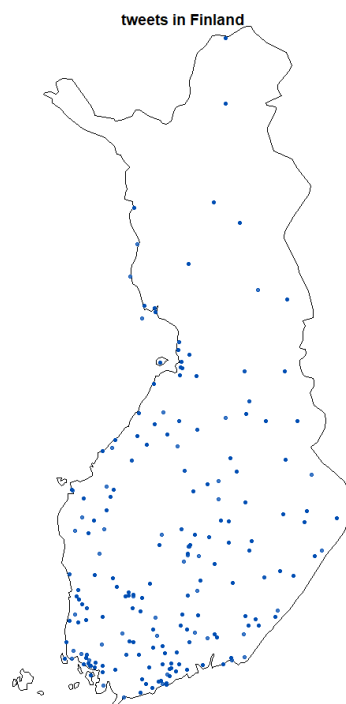
Verified users compose about 4% of unique users in the data. Some active and verified users that do not represent any institutions include journalists such as @JussiPullinen, @ElinaLappalainen and @JuttaSarhimaa, and politicians with more and less long careers such as @VilleNiinisto and @MiikkaKeranen. Active non-verified users represent a more diverse set of people.

Users with the largest number of followers include @alexstubb, former Finnish prime minister, @TuomasEnbuske, well-known Finnish media person, and media institutions such as @yleuutiset, @iltasanomat and @Kauppalehti. These users with the highest count of followers, however, only were not that influential in terms of creating culture related content, as their tweet counts in the data set ranged from only 3 to 43.

In the dataset, 76% of the tweets contain geographical information about the location of the user. The majority of the tweets in the dataset are from the major cities in Finland, such as Helsinki, Tampere and Turku. The ten cities that had the highest number of tweets contain a total of 69% of all the tweets with geographical information.

**Figure - FI - 1** shows the distribution of tweets in Finland. The density of points indicates higher number of tweets posted from the region. The small number of points on the map compared to the total size of the dataset is explained by the fact that only a small percentage of users provide necessary geographical data.

**Figure - FI - 1.** *Geographic distribution of the tweets in the dataset in Finland*



#### 4. Hashtag and keyword analysis

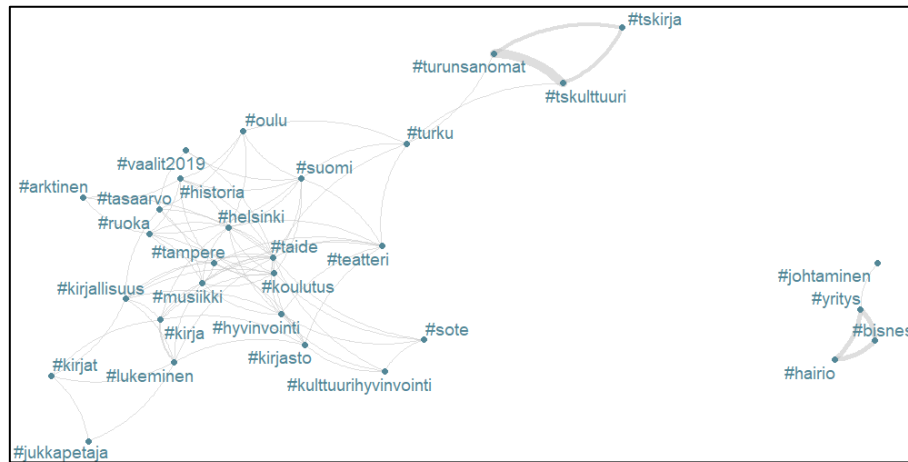
Out of the total tweets, in both 2019 and 2020, 15% included a hashtag in the tweet. In total, 32,447 were used, composing 3,782 unique hashtags. The mean number of hashtags per tweets where they were used is 1.9. The most common hashtags are presented in Table FI-1, grouped by whether they were used before or after the first confirmed Covid-19 case in Finland in 29.1.2020. Hashtag “#kulttuuri” is excluded since it was used as the keyword in acquiring the tweets. **Table - FI - 1** includes also most common keywords in the cleaned data. Some of the most common hashtags pre-Covid-19 include references to art (#taide), music (#musiikki), books (#kirjat, #kirja), and theater (#teatteri). Some other common hashtags include hashtags for news media and big cities, for example, #turunsanomat and #Helsinki. In the same time period, some of the most common keywords include terms such as Finnish, human, Finland, part, art, work and world. After the first confirmed Covid-19 case, hashtags related to the Covid-19 crisis (#korona, #koronavirus) emerged among the most frequent hashtags, and so did #well-being (#hyvinvointi). The most common keywords remained mostly same pre- and during Covid-19 in Finland.

**Table - FI - 1.** Fifteen most common hashtags and keywords before and during Covid-19

Pre-Covid-19				During Covid-19			
Hashtag	Freq.	Keyword	Freq.	Hashtag	Freq.	Keyword	Freq.
#taide	215	suomalainen	2,448	#taide	179	ihminen	1952
#tskulttuuri	132	ihminen	2,374	#korona	81	suomalainen	1830
#turunsanomat	100	suomi	2207	#musiikki	54	suomi	1812
#Kulttuuri	82	osa	1337	#Kulttuuri	50	osa	1207
#Helsinki	68	taide	1272	#tskulttuuri	46	taide	1189
#kirja	68	työ	1153	#Helsinki	44	työ	908
#Tampere	68	maailma	1150	#kirjat	38	kirja	907
#tskirja	68	nainen	1127	#turunsanomat	33	maailma	901
#musiikki	57	kirja	1097	#koronavirus	32	lapsi	814
#teatteri	55	lapsi	1068	#teatteri	32	maa	813
#kirjat	54	maa	1015	#hyvinvointi	31	nainen	761
#johtaminen	53	kieli	1003	#kirja	29	kieli	706
#lukeminen	51	n	927	#ruoka	29	n	700
#kirjallisuus	49	nuori	809	#Tampere	28	mies	668
#turku	46	mies	744	#liikunta	27	nuori	648

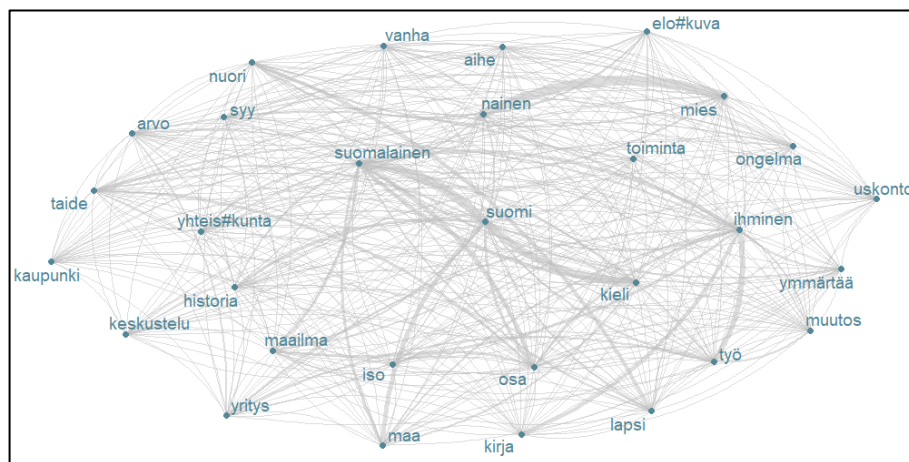
The network structure of the 30 most combined hashtags before first confirmed Covid-19 case is presented in **Figure – FI - 2**. We can detect three clusters of hashtags. On the left, some more traditional domains of culture, such as #art, #book, #music and #literature are found, but also #food, #well-being, and #equality. This part of the network describes multiple notions of culture. A small cluster of hashtags in the top of **Figure - FI - 2** is connected to Turku, the European capital of culture in 2011. This cluster might emerge from local media promoting their culture-related content linked to books. The third cluster, which relates to business or management jargon, is clearly separated from the rest of the network, suggesting that it is more distant from rest of the culture-related hashtags.

**Figure - FI - 2.** Network of hashtags in 2019. Thirty most commonly combined hashtags and their relationship to other common hashtags



In **Figure - FI - 3**, the most common keywords and their relations in the cleaned data before the first Covid-19 case are presented. From the Figure we can see that words such as Finnish and Finland, woman and man, and young and child are co-occurring frequently. It is also notable that unlike in the network of hashtags, the keywords are less related to domains linked to narrow understandings of culture, book, art and movie being the only keywords among the top 30 presented in the network.

**Figure - FI - 3.** Network of keywords in 2019. Thirty most commonly used keywords in the cleaned data and their relationship to other common keyword.



In **Figure - FI - 4**, the network structure of the 30 most combined hashtags after the first confirmed Covid-19 case is presented. Compared to a similar network presented in **Figure - FI - 2**, some notable differences can be seen. Hashtags related to Covid-19, #korona, #koronavirus and #koronafi, seem to have emerged as important co-occurring hashtags when discussing about culture in Finland. The distinct cluster related to Turku is still present in the network, but co-occurring hashtags related to business or management jargon have fallen from the top 30 network.



\_\_\_\_\_



1. *Journal of the American Medical Association*, 1997; 277: 1001-1005.

**Figure - FI - 6. 10 Wordclouds corresponding to each topic found in the 2019 tweets dataset.**  
Topics are arranged from left to right



**Culture in the narrow sense.** This topic contains words related to a narrow understanding of culture. Words such as music, theater, artwork, concert and show are important for this topic. This topic seems to also contain promotion of cultural items and shows, since words like critique, ticket, perform and agenda are included in this topic.

**Congratulations and promotion.** This topic contains words related to promotion and congratulating winners of different events in multiple fields of society. Words such as young, win, game, award and great are important for this topic.

**Public funding.** This topic is related to discussion about public funding and how and where tax money should be distributed. Words such as municipality, society, art, funding, research, sport, government, political and money are important for this topic.

**Finnish identity.** This topic is focused on Finnish historical and contemporary social and cultural discussions and national culture. Important words include Finland, Finnish, language, citizens and history.

**Journalism.** This topic is related to journalism and discussion about news articles. Words such as people, media, problem, understanding, article, question and discussion are important for this topic.

**Social and cultural debates.** This topic is clearly related to conversations about equality and immigration. Words such as man, woman, power, religion, racism, problem and values are important for this topic.

**Cultural reviews.** This topic is related to reviews and promotion of cultural domains related to



narrow culture. Words such as picture, magazine, review, music, shine, document and funny are important for this topic.

**Business or management jargon.** This topic contains quite clearly words related to white-collar working life, such as consulting and leadership. Important words for the topic are, for example, change, company, leader, leadership, action, vision, organization and learning.

**Well-being promotion.** This topic is related to promotion of local well-being, especially for younger people. Important words include art, space, youth, child, city, knowledge, future, school, family, event and discussion.

**Books and literature.** This topic is clearly related to discussion about books and literature. Words such as book, story, writer, text and recommendation are important for this topic.

Next, we will move on to interpret the analysis of tweets created after 29.1.2020, which is the date of the first reported Covid-19 infection in Finland. The resulting topics of this analysis are presented in **Figure - FI - 7**.

**Figure - FI - 7.** Wordclouds corresponding to each topic found in the 2020 tweets dataset. Topics are arranged from left to right.



**Social and cultural debates.** This topic is clearly related to conversations about equality and immigration. Words such as man, woman, power, religion, racism, problem and value are important for this topic. This topic is similar as the one found from one year earlier.

**Global culture.** This topic relates to discussion about global themes and how they relate to Finland. Important words include Finnish, world, border, international and appropriation.

**Literature and journalism.** This topic related to both literature and journalism. Words such as book, literature, (art)work and writer are important for this topic. On the other hand, also interview, magazine and article are important for this topic, suggesting that reading related to both books and new media are part of this topic.

**Public funding.** This topic is related to discussions about public funding and how and where tax money should be distributed. This topic was present in 2019 as well. In 2020, however, the word *korona* (Covid-19) appeared among the important words. Other words that are important for this topic are society, municipality, art, funding, research, sport, government, political, science and money.

**Business or management jargon.** This topic contains quite clearly words related to white-collar working life, such as consulting and leadership. Most important words for the topic are, for example, society, meaning, value, change, company, leader and future. This topic is similar as the one found 2019.

**Congratulations and promotion.** This topic is a combination of domains related to the narrow sense of culture, such as music, exhibition, artists, document and tv, and words related to promotion and congratulating, such as wins, deserve, great and reward. This topic was found in 2019 as well.

**Politics and cancel culture.** This topic is related to discussion about politics and the contemporary phenomena, cancel culture. Important words for this topic are for instance people, politics, political, cancel culture, question and problem.

**General discussion.** This topic is related to more general discussion in Twitter, possibly about news articles. Words such as old, topic, idea, interesting, promote, and media are important words for this topic.

**Finnish identity.** This topic is similar to the one found in 2019, and it relates to discussions about historical and contemporary Finnish debates and national culture. Important words for this topic are Finnish, Finland, history, language and people.

**Promotion of cultural events.** This topic relates to the promotion of cultural events. Important words for this topic are theater, space, home, show, buy, concert and gig.

## 6. Discussion and conclusion

Overall, there seems to be plenty of topics that are discussed when Finnish people mention “*kulttuuri*” or some of its conjugation in a tweet. These topics seem to remain constant despite the Covid-19 pandemic taking place in 2020, as the only topic that included the word “*korona*” among the important words, public funding, were also found in 2019. Despite the word “*korona*” being important for only one of the topics, the network analysis of hashtags revealed that hashtags related to the Covid-19 crisis were co-occurring frequently with other hashtags used when the word “*kulttuuri*” was present in the tweets. Some new topics emerged from 2019 to 2020, however, but it is hard to say whether this is because of the Covid pandemic or not.

Most of the topics were found both before and during the Covid pandemic. For example, the discussion about public funding was found as its own topic before Covid. After the pandemic started, the word *korona* appeared as an important word for this topic, reflecting the cultural sector that lost sources of income; a discussion about whether this sector should be compensated took place on Twitter.

Discussions about social and cultural debates and issues were among the most important topics. Tweets about immigration, gender equality, Finnish identity, politics and cancel culture were distributed among several topics in both before and during the Covid crisis.

Discussions related to the more narrow understanding of culture were found from several topics as well. The perspectives on the discussion did vary; some topics were more closely connected to promoting events, others to reviews. Literature was the only domain of culture as arts that composed its own topic, although in 2020 it was shared with journalism. Literature being its own topic seem reasonable when considering the lively public discussion and cultural initiatives around the topic lately in Finland.

Discussions about work culture were found both before and during the Covid pandemic. The appearance of this topic is interesting because all the other topics are related to either social and cultural debates or promoting and discussing artistic cultural items. These topics of white-collar working life, related to, for example, consulting and leadership, suggest that Twitter might be a platform for promoting discourses of successful business management.

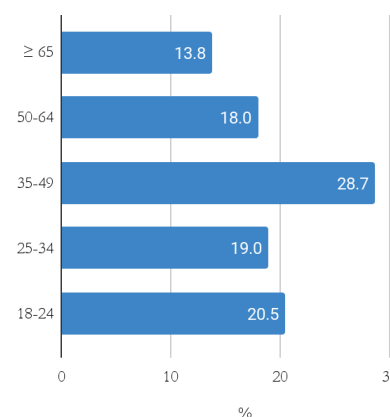
## Findings for France

### 1. Twitter usage in France

In France in 2019, the number of Twitter users aged 18 and over is estimated at 10.7 million people (21%)<sup>4</sup>, of which 40% access the platform every day or so. Most users are men (60%) in their 40s, living in big cities (especially in the Paris metropolitan area) and those with a university degree are over-represented.

As we can see (**Figure - FR - 1**) all age groups are well represented on Twitter. However, according to IFOP, the platform's diffusion is more widespread among the young public: the share of 18-24 year olds on Twitter (41%) is much higher than that of other age groups, which varies between 11% for people aged 65 or more and 26% for people aged 25-34 (IFOP 2019).

**Figure - FR - 1. Twitter users by age (2019)**



Source: Prepared by the authors based on data from IFOP (2019)

### 2. Dataset description

In 2019, 26,248 tweets containing the word “culture”, posted from France and written in french were retrieved from Twitter’s archive. In 2020, the number of tweets drops to 25,687. In both years most of tweets were sent from big cities (**Table - FR - 1, Figure - FR - 2**), Paris alone contributing with more than a quarter of all tweets. But there is a main difference in terms of the concentration of Tweets: firstly, the number of cities in the dataset increased by 4.8%, from 3,484 in 2019 to 3,653 in 2020. Furthermore, the number of Tweets posted from Paris has been reduced by 3.5 percentage points (i.e. a decrease of 12.6%). This is a direct effect of the Covid-19 pandemic: on the one hand, due to the confinements and the new framework for telework adopted in France, some of the inhabitants of the large urban centres (especially the inhabitants of the Paris metropolitan area) have moved more or less temporarily to smaller towns, mostly to the villages where their holiday homes are located. On the other hand, the crisis of Covid-19 had a strong impact on tourism in several cities: hotel reservations fell by 50% in France, but this drop reached -65.1% for the Paris metropolitan area and -46.1% for region *Provence-Alpes-Côte d'Azur* (Dangerfield et Mainguené 2021). Thus, the seasonal population of these cities has also been reduced.

<sup>4</sup> Considering Twitter does not publish official statistics about users in France, the estimation of the amount of Twitter’s users varies according to methods used by different survey and polling agencies. According to IFOP (Dabi and Dubrulle 2019), 21% (±1,5%) of persons declare to have a Twitter account, that means between 10 million and 11,5 million of persons (1); Harris Interactive (2019) estimates it in 20% (±1%), that means between 10,1 and 11,2 million of monthly active users (2); and according to Médiamétrie (2020), there are in France 15,2 millions of unique monthly access to Twitter in France, that means 23% (3).

(1) Persons aged 18 and over in 2019 residing in European France;

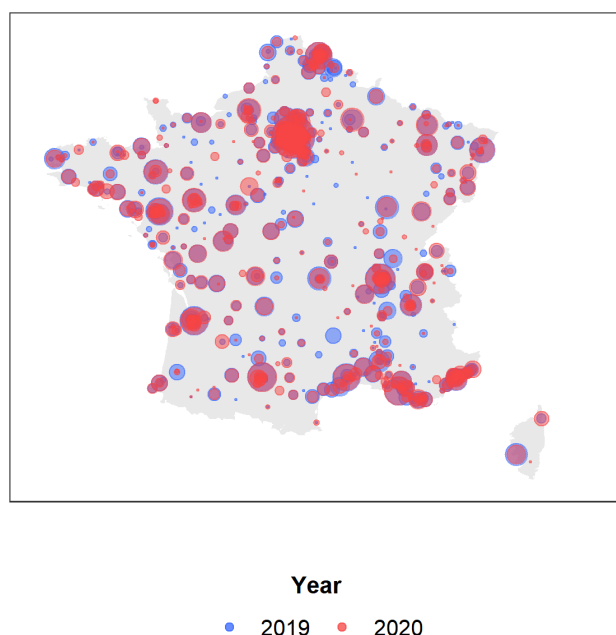
(2) Persons aged 15 and over in 2019 residing in European France;

(3) Persons aged 02 and over in 2019 residing in European France.

**Table - FR - 1.** Most active cities in terms of posted tweets (%)

City	2019	2020
Paris	27,89	24,38
Lyon	2,18	2,25
Toulouse	1,58	1,78
Bordeaux	1,4	1,62
Marseille	1,32	1,18
Montpellier	1,2	1,1
Nantes	1,08	1,02
Lille	0,98	0,98
Nice	0,98	0,86
Strasbourg	0,89	0,77
Rennes	0,66	0,75

**Figure - FR - 2.** Most active cities in terms of posted tweets (2019-2020)



### 3. Hashtag analysis

Regarding the use of hashtags, this practice does not concern the majority of tweets related to “culture” in France: only 43% of tweets posted in 2019 contained a hashtag and almost 54% of users never tagged their posts this year. These values decreased in 2020, when 36% of tweets were hashtagged and 60,8% of users did not tag any of their tweets. **Table – FR- 2** shows the number of hashtags and the number of tweets posted by the ten most active profiles in 2019 and 2020. As we can see, the practice of tagging tweets is more akin to professional use of Twitter: among the very active users, institutional profiles post on average more than ten hashtags per tweet, while journalists and politicians post just over three and citizens almost do not hashtag their tweets. Although the number of hashtags is correlated with the number of tweets posted, the most active users in terms of tweets are mostly individual profiles and not institutional profiles. The difference in terms of the number of hashtags observed in 2019 and in 2020 can be interpreted, thus, as the result of the reduction in the level of commercial cultural activities during the Covid-19 pandemic and, through this, its dissemination on Twitter.

**Table - FR - 2. Most active profiles on Twitter: number of hashtags and tweets (2019-2020)**

User	Description	2019		2020	
		Hashtags	Tweets	Hashtags	Tweets
@Bourg_en_Bresse	City of <i>Bourg en Bresse</i> official profile	250	48	.	.
@Breizhdrauhl	Citizen	.	.	0	73
@CamilleBataille	Cultural affairs manager of the city of <i>Bussy-Saint-Georges</i>	271	61	126	29
@CeFoLiAc	Language Training Centre	501	29	256	22
@celinecalvez	Politician - deputy of the region <i>Haut-de-Seine</i>	149	43	192	60
@Delcroixx	Editor-in-chief on culture - <i>Le Figaro</i>	1185	766	.	.
@DelphesDelvaux	Citizen	4	269	.	.
@emarolle	Editor-in-chief on culture and television - <i>Le Parisien</i>	601	670	703	881
@Florence31000	Bookseller	0	1	312	31
@gabiobfrance	Journalist	1	30	5	90
@Ghemmour5	Citizen	0	194	0	129
@I_M_NOBODY27	Politician - Ex-national secretary of UDI	.	.	210	44
@IsaacAmar5	Citizen - teacher of philosophy	0	1	0	76
@IsabelleLayer	Cultural journalist ( <i>France Info</i> ) and actress	209	76	57	43
@JFGuyot	Journalist - <i>Agence France Press</i>	14	4	192	60
@LaurentBibas	Citizen	1	10	6	67
@LaurentQueige	Manager of a startup incubator on tourism	73	7	205	21
@Leboulonnais1	Citizen	2	34	7	75
@LiveTweetPsg	Official profile of PSG football club	0	161	0	141
@lolomakossa	Director of Communications at <i>France Télévisions</i>	152	87	90	61
@MaeceneArts	Foundation promoting arts and cultural events	321	27	109	7
@mangapop_lemans	Shop specialized in japanese pop culture	510	47	286	32
@PEPITOBLACK	Artist official page	358	26	9	1
@ROUSSELTHOMASC1	Cultural mediator	2	129	4	50
@TheDarkPixel	Digital Media Strategist	69	77	29	50
@tvmountain	Web TV of region of <i>Chamonix Mont-Blanc</i>	243	24	.	.
@VDuponchelle	Journalist - Art & culture - <i>Le Figaro</i>	0	1	239	500
@vizouine	Digital Marketing Expert	4	3	571	100
@yannickhoppe	Politician - former mayor of city of <i>Le Bourget</i>	393	90	132	48

. Data not available

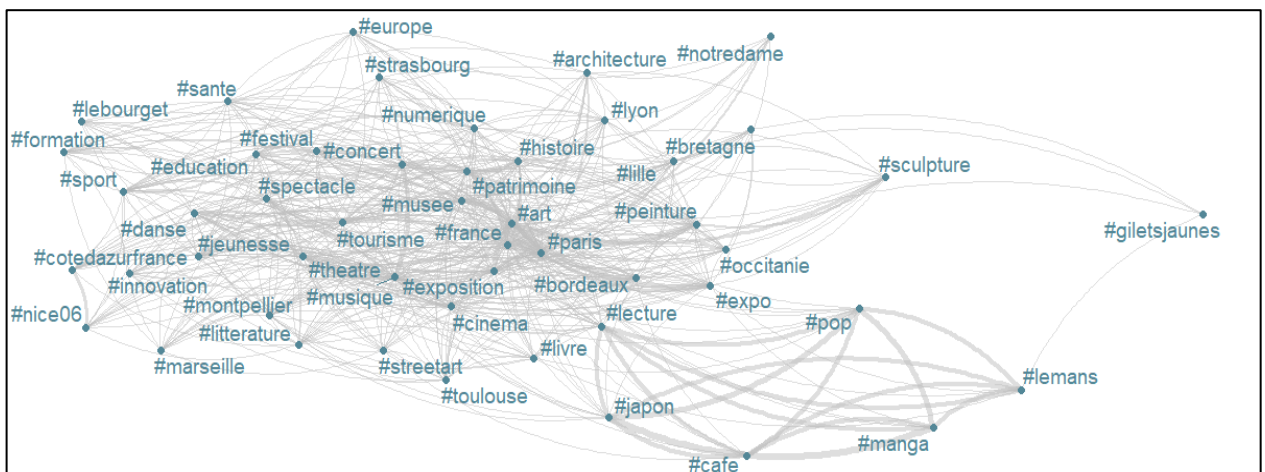


Among the hashtags, we find four main themes related to “culture” in 2019 (**Table - FR - 3, Figure FR - 3**): places related to tourism (yellow); fields of culture (green), cultural activities (blue), education (purple). In addition, we observe a very active user that composes a “hub” in the bottom left of the network (it is a shop specialized in japanese pop culture and manga). The first thing to note is the centrality of the hashtags #art, #france and #paris. These hashtags are, in turn, linked to places of culture (cinema, theatre, museum etc.) and major areas of culture (music, painting, history, heritage etc.). The Mediterranean cities (Nice, Montpellier, Marseille etc.), on the other hand, are more linked to issues of education, innovation and youth. This is because we find in this region several training and language centres that use Twitter as a means of dissemination. We can also see the link between Strasbourg and Europe, that is due to the fact that the European Parliament is located in this city. Brintany and southern France (Occitany, Bordeaux and Toulouse ) are more related to exhibitions of paintings, street art and sculptures. Finally, we can identify two major events in France in 2019 : the yellow jackets movement [gilets jaunes], isolated in the right of the graph ; and the Notre-Dame de Paris fire.

**Table - FR - 3. Twenty most posted hashtags in 2019**

Hashtag	N	Hashtag	N	Hashtag	N	Hashtag	N	Hashtag	N
#art	579	#musique	236	#musee	142	#cinema	129	#expo	110
#paris	456	#theatre	201	#numerique	142	#tourisme	129	#concert	103
#patrimoine	314	#exposition	175	#education	141	#festival	116	#montpellier	93
#france	244	#sport	143	#histoire	137	#bordeaux	115	#innovation	91

**Figure - FR - 3. Network of fifty hashtags -2019**

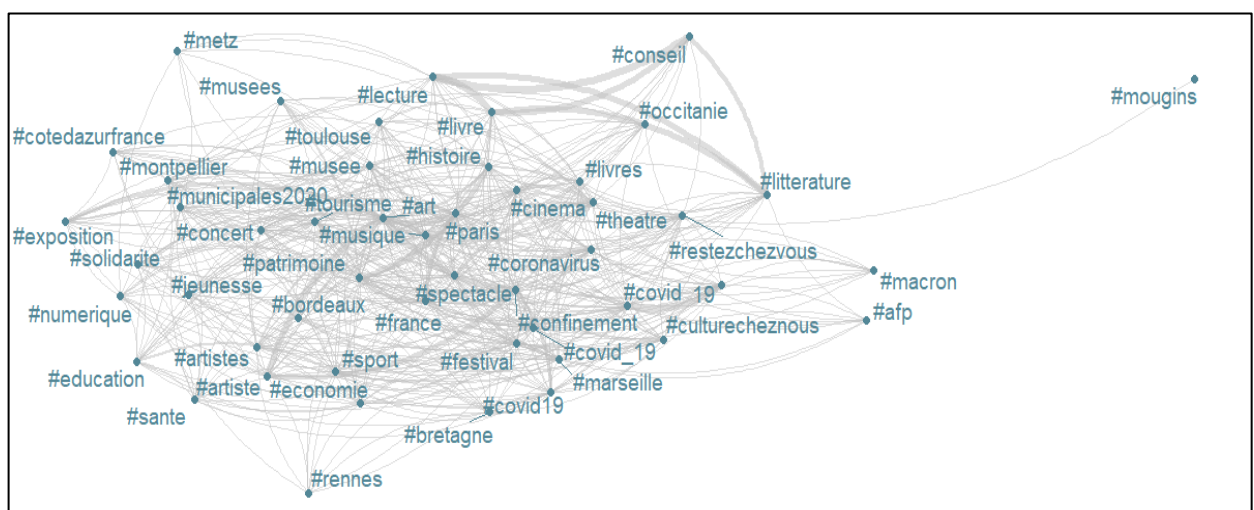


In 2020, we observed a reduction in the number of hashtags (**Table - FR - 4**). Compared to 2019, the hashtag #france had its use reduced by 28%, the hashtag # by 40% and the hashtag #paris, a reduction of 50%. We still have the four main themes observed in 2019 - places related to tourism (yellow); fields of culture (green), cultural activities (blue), education (purple), but three new topics have emerged: hashtags related to the Covid-19 pandemic (magenta); and categories of cultural practices “indoor” (red) and the mayor and city Council elections of 2020 (orange). But the network (**Figure - FR- 4**) of hashtags from 2020 is less structured than the 2019 network. As we can see in the network (the “indoor activities” are strongly connected to the hashtag recommendation [#conseil] and to hashtags related to the Covid-19 pandemic, that suggests that Twitter was used as a tool to find ways of using free time in a different way than is usually done. The proximity of the city of Marseille to the hashtags linked to Covid-19 results from the fact that this city hosts the research centre coordinated by Dr. Didier Raoult, who has been at the centre of controversy throughout 2020.

**Table - FR - 4. Twenty most posted hashtags in 2020**

Hashtag	N	Hashtag	N	Hashtag	N	Hashtag	N	Hashtag	N
#art	345	#Covid19	193	#cinema	162	#sport	109	#tourisme	102
#patrimoine	236	#municipales2020	192	#theatre	156	#Covid_19	105	#lecture	97
#paris	232	#musique	176	#coronavirus	133	#education	103	#litterature	80
#confinement	221	#france	175	#histoire	116	#livre	102	#spectacle	80

**Figure - FR - 4. Hashtags network -2020 Twitter's topics**

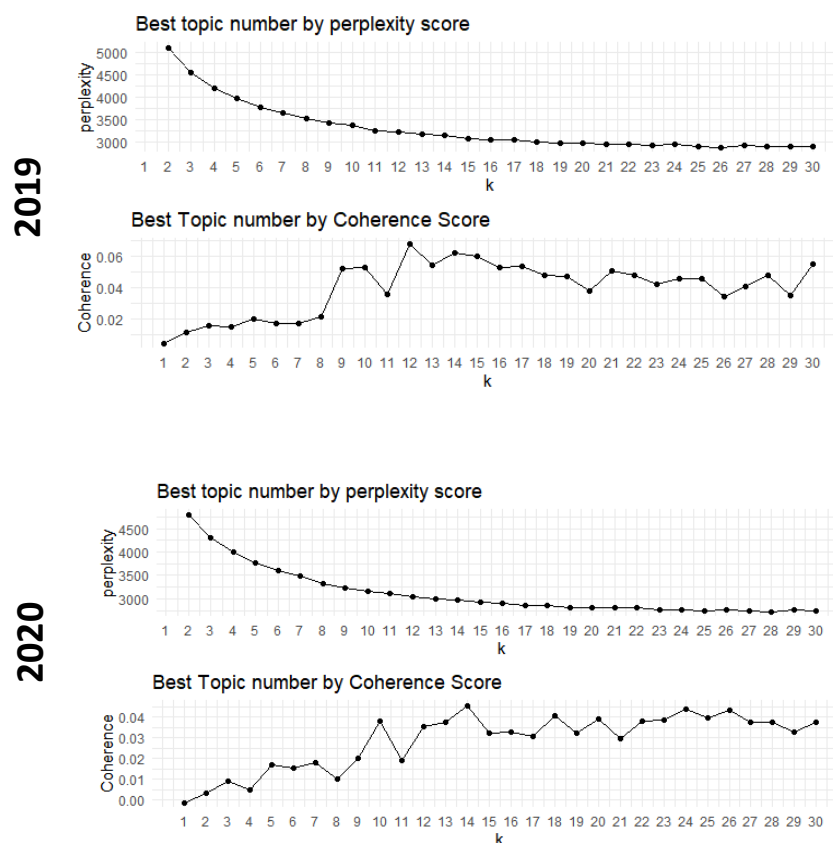


#### 4. Topic modelling

This section compares the topic of the tweets posted in 2019 and 2020. To make this comparison, we used a Latent Dirichlet Allocation (LDA) of the tweets. However, we only took into account nouns, adjectives and pronouns to make more meaningful clusters, totalling 123,916 terms after lemmatization in 2019 and 125,947 terms in 2020.

The main challenge of any classification procedure is to produce clusters that are both statistically pertinent and that we are able to interpret. In the french case (**Figure - FR - 5**), in both years after reaching 10 topics, we observe a reduction in the level of coherence of the clusters. This fact suggests that a higher number of classes will produce groups that are fairly close in terms of composition. This is why we have decided to work with 10 topics on culture that were discussed on Twitter along these years.

**Figure - FR - 5. Best topic number**



2019



**Public commotion.** This cluster is marked by the burning of *Notre-Dame de Paris*, by the tributes of two years after the death of Johnny Halliday (one of the most famous french singers) and by the death of a young man found in the *Seine* river. All these events had a great public impact.



**Business and economic trends.** This topic includes tweets about the digital economy, globalization, telework technologies and the development of a green economy.



**The new season of cultural events.** This cluster deals with the new season of theatre, music festivals and football upcoming, publicizing these events.



**Literature.** This cluster contains tweets about novel launches, literary awards and cultural programmes (radio and TV).



**World languages and cultures.** This cluster combines exhibitions of Japanese pop culture, spanish recipes ("black rice with white squid"), archaeological excavations in Israel and the death of the french philosopher Michel Serres.

2019



**General knowledge.** In this topic were classified Tweets about "general knowledge quizzes" in TV shows, in school tests and about how one can improve their general culture.



**Art exhibitions.** The cluster contains tweets that publicize art exhibitions and vernissages.



**Far right political clash.** Tweets (and replies) about the position of the far right in France: "Christian culture" vs "Muslim culture", discussions about the "rape culture" based on the case of Neymar (who was charged of rape in 2019), discussions about gender issues and arabic world issues.



**Culture and education in the third sector.** Publicity of cultural, educational and sport activities organized by non-profit organizations and local governments. Promotion of call for tenders for new and of already established partnerships between cities and non-profit organizations for local development



**Series, films and music videos.** This cluster gathers tweets about the release of films and series, as well as suggestions and reviews. It also includes discussions about music videos, music tv shows and awards (*Victoires de la musique*, Cannes festival, Cesar award).

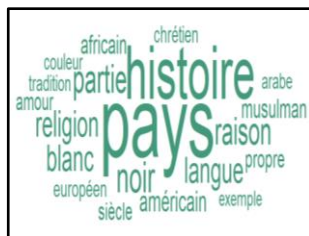
2020



**Football and journalism.** This cluster brings together Tweets about football and, especially PSG, that has a very active web page called *Le bouillon de culture du jour* [The culture broth of the day]. We also find tweets on journalism and popular science.



**Local projects and 2020 mayoral elections.** Dissemination of calls for tenders of local projects on heritage and on activities for young people. Promotion of public debates on mayoral and city Council elections.



**Far right on multiculturalism.** Tweets (and replies) about the position of the far right in France, especially on the impact of “foreign cultures” on “french culture”.



**The upcoming season of cultural events.** This cluster deals with the new season of painting and photo exhibitions, music festivals and upcoming cultural activities.



**Cultural policy under Covid-19.** This topic includes Tweets on the policy-making during the Covid-19 pandemic, especially regarding the cultural sector policies.

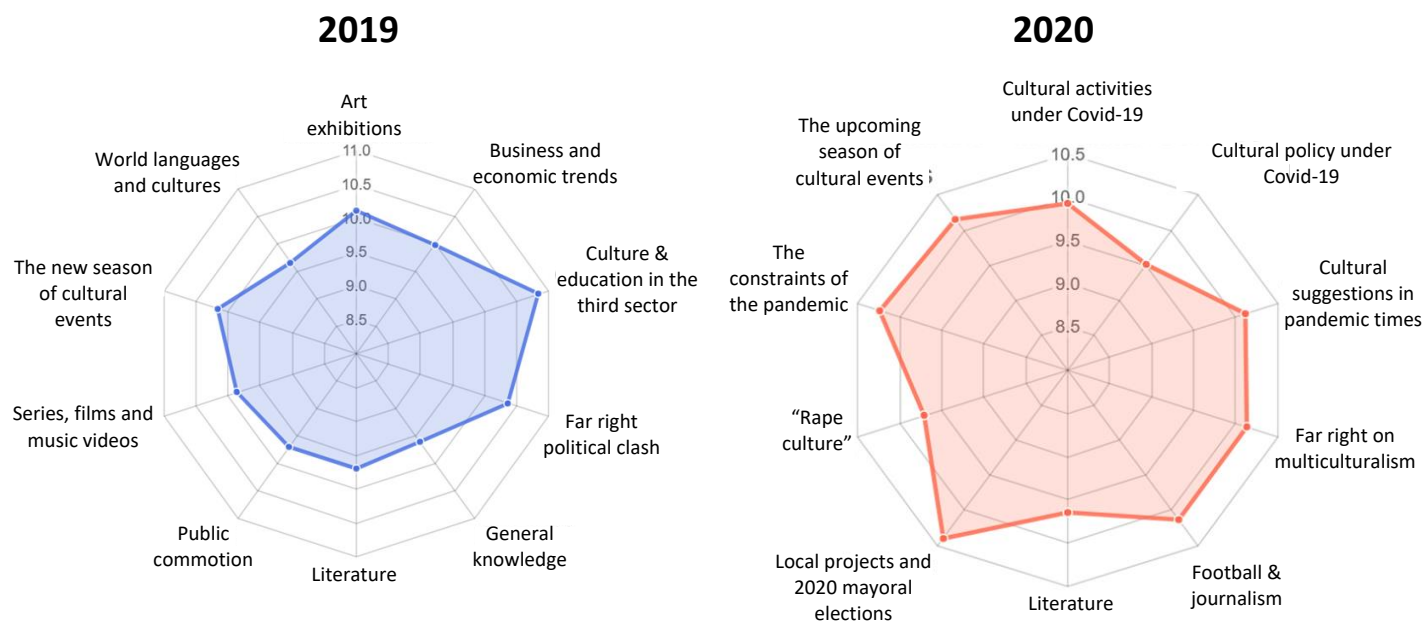


**“Rape culture”.** This topic includes tweets on violence against women, discrimination and gender issues. Furthermore, this topic contains tweets on Harvey Weinstein’s trial, the case of Gerard Darmanin (Minister of Public Action and Accounts charges to have sexual favors from a woman in exchange for his support on her judicial case).

## 5. Discussion and conclusion

In terms of inequality between thematic sizes, we observe a relative balance (**Figure - FR - 6**). In 2019, *Culture and education in the third sector* was slightly more discussed than other topics; in 2020 the *Local projects and 2020 mayoral elections* topic attracts the most attention from the public.

**Figure - FR - 6. Topics distribution 2019-2020**

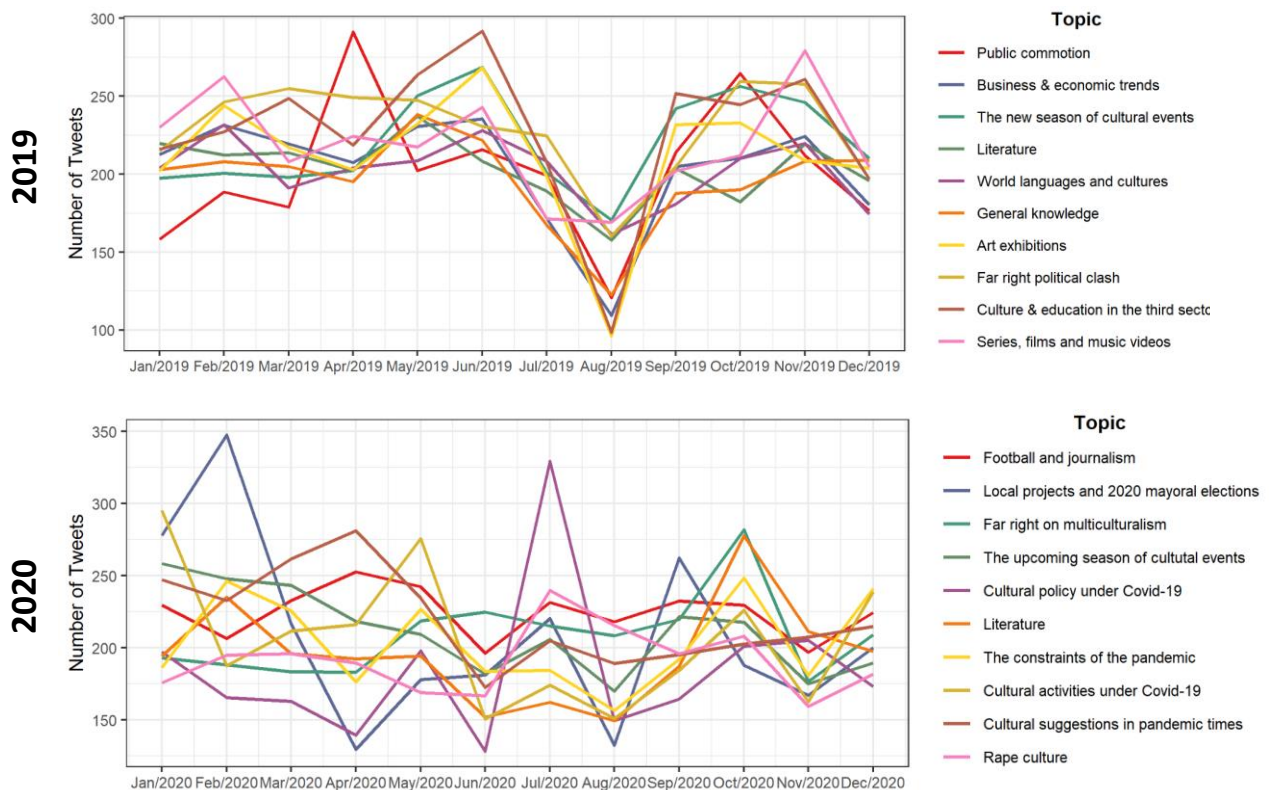


It should be noted that these different themes are not evenly distributed throughout the year (**Figure - FR - 7**). First, we can see how the distribution of tweets throughout 2020 has smaller variations than in 2019: the reduction in online activity due to the August vacancies observed in 2019 is weaker in 2020. Secondly, while in 2019 the flow of tweets follows a general trend, in 2020 there are strong differences between the time of year when certain topics were discussed. This is, for example, the case for *Cultural suggestions in pandemic times* and *Football and journalism*, which were discussed extensively during the two confinements (March-May and October-December), and in between has fallen sharply.

Due to the pandemic, tweets discussing the new season of cultural events have a very different behaviour between the two years: in 2019, tweets on *the upcoming season of cultural events* increased consecutively until June, then we see a rapid drop due to the summer vacancies, then an increase from the end of August which will lead to a new peak in October. In 2020, the number of tweets will decrease consecutively until August, after a small increase in July, then they will increase to a peak in September, but at a lower level than in the previous year.

In both years, some events boost some discussion. So, in 2019, the Notre-Dame de Paris fire in April and the Johnny Halliday tributes in December are the peaks of the *Public commotion*. Likewise, in July 2020 we see a peak in the *Cultural policy under Covid* because it is the end of confinement restrictions due to the Covid-19 pandemic in France and, in particular, because it will be followed by a major change in ministers of state.

**Figure - FR - 7. Occurance of the topics throughout the year**



As we have seen, there are big differences between the years 2019 and 2020, especially due to the impact of the Covid-19 pandemic: not only the number of hashtags and tweets have been reduced, but the content of the topics have strongly changed. Of course, we still find discussions around the main artistic fields (cinema, literature, theatre, exhibitions etc.), but the way in which these fields have been approached has been completely turned upside down: from publicizing events or commenting on launches, we have moved on to a discussion of public policies for each specific area. That is, the pandemic has deposited a political layer - already present on Twitter in general and pointed many times by analysts - on "conversations" about the use of free time and leisure practices.

## Findings for the Netherlands

### 1. Twitter usage in the Netherlands

In 2019 and 2020, Twitter respectively counted 2,5 million and 2.8 million users in the Netherlands. The use of Twitter has been quite stable for years. In 2019, Twitter usage declined, but in 2020 the number of users increased again with 13 percent to the level of 2018. In 2019 and 2020, respectively 18% and 20% of the Dutch population older than 14 years used Twitter (Oosterveer 2020). Among millennial respondents (aged 20-39 years) 26% said they used Twitter (Tankovska 2021).

According to publicly available demographics taken from the annual Newscom social media research for 2016, males were overrepresented among Twitter users. Moreover, a relatively high proportion of Twitter users was higher educated (50%); only LinkedIn users were more often higher educated (65%), but Facebook (39%), WhatsApp (42%) and YouTube (46%) included fewer higher educated users. The share of lower educated users amounted for Twitter to 13% and for LinkedIn to 6% in 2016, as opposed to 20% for Facebook and 16% for both WhatsApp and YouTube (Boerema 2016). The share of intermediate educated users amounted to 37% for Twitter, which was more or less comparable to other social media with the exception of LinkedIn (29%). In 2016, 7% of Dutch Twitter users were 15 to 19 years old, 39% were in their twenties or thirties, 45% were 40 to 64 years old and 9% were aged 65 or older (Statistica 2016).

In 2020, the number of daily users amounted to 1,13 million, which represented a significant growth of 25 percent (224,000 users) compared to 2019. In 2020, more than one in three users used Twitter every day. In the age groups of 15-19, 20-39, and 40-64 year old people, respectively 9%, 10%, and 9% used Twitter on a daily basis (2Bfound 2021); among respondents aged 65 years and older only 4% used Twitter every day.

### 2. Data description

Using the Twitter premium account API, we scraped 242,441 tweets using the keyword “cultuur” (Dutch term for “culture”) that were posted in 2019 and 2020 from an account with profile location nl. After removing both tweets created by tweetbot and tweets for which the location was missing or irrelevant, a total of 186,068 tweets with profile location in the Netherlands were retained. From this initial dataset we excluded all cases, that were posted in a language other than Dutch. This involved the removal of English language tweets ( $n = 2,591$ ), tweets in 29 other languages ( $n = 1,793$ ) and tweets for which the language was undefined. The resulting dataset included 178,722 Dutch language tweets with keyword “cultuur”.

Because of the salience of the English language in the Netherlands and the Dutch Twittersphere, we also did a query for tweets that were posted with the English language word “culture” in 2019 and 2020. This yielded an additional 81,910 tweets for the years 2019 and 2020. After removing both tweets created by tweetbot and tweets for which the location was missing or irrelevant, a total of 50,163 tweets with profile location in the Netherlands were retained. Most of these tweets (79%) were in English ( $n = 39,663$ ), while Dutch language tweets accounted for 14% ( $n = 7,121$ ). The remaining tweets were undefined ( $n = 1,145$ ) or in another language ( $n = 2,234$  covering 42 different languages). The English language tweets will be analysed and compared with the Dutch language tweets in a follow-up study.

For our current analyses we combined all Dutch language tweets with keyword “cultuur” or keyword “culture” into one datafile ( $n = 185,843$ ). This chapter only pertains to these Dutch language tweets. In order to establish whether the culture-related communication on Twitter differed before and

during the Covid19 pandemic, we split the data according to a tweet's posting date. On January 31, 2020, the first news items about possible corona infections in the Netherlands appeared, and we divided the data into tweets posted between January 1, 2019 and January 30, 2020 (pre-Covid19 phase;  $n = 92,322$ ), and January 31, 2020 and December 31, 2020 ( $n = 93,521$ ).

## Locations

Based on the geographical coordinates' information, we have information about 555 unique locations from which tweets about culture in 2019 and 2020 were posted. Figure 1 shows the distribution of the tweets in The Netherlands. The density of dots shows the number of tweets posted from different parts of the country. **Table - NL - 1** shows the top 10 locations in 2019-2020: one fifth of all tweets are posted from Amsterdam, while Rotterdam and Utrecht respectively account for 6 and 5.4 percent. Taken together, the top-10 locations accounted for 45.8 percent of all tweets. The ranking of locations as well as the geographical distribution of tweets was highly similar before and during the Covid19 pandemic.

**Table - NL - 1. Top 10 locations**

Locations	N	%
Amsterdam	38,174	20.5
Rotterdam	11,113	6.0
Utrecht	10,931	5.4
Groningen	7,164	5.4
Nijmegen	3,363	1.8
Den Haag	3,231	1.7
Eindhoven	3,179	1.7
Amersfoort	3,002	1.6
Haarlem	2,576	1.4
Leeuwarden	2,483	1.3
<b>Total Top-10</b>	<b>85,216</b>	<b>45.8</b>

**Figure - NL - 1. Geolocations of tweets in the Netherlands ( $n = 185,843$ )**



## 3. User analysis

The aggregated dataset for 2019 and 2020 contains 185,843 culture-related posts from 29,859 unique Twitter accounts, with an average of 6.2 tweets over this two-year period. The 25 most active users accounted for 20,888 tweets, representing 11.2% of the total number of tweets, whereas more than a quarter (27.8%) of all tweets was published by the 200 most active user profiles that posted at least one culture tweet per week on average. Overall, the national daily newspaper *de Volkskrant* was the most frequent poster, followed by news portal *nieuwsblik.nl*, the regional newspaper *Dagblad van het Noorden*, the Christian newspaper *Nederlands Dagblad*, and *Culturele Avonturiers*, an online platform about cultural entrepreneurship.

The five most active user profiles on average posted between 1.5 and 3 tweets per day with keyword culture. About half of the top 25 users concern news media, including national and regional newspapers (7), a news portal (1), local news sites (3), and a local RTV station (1). Other highly active users are specialized culture and leisure websites (4) and individual users (7), including two newspaper journalists.



Below, we present separate user analyses for the two phases before and during the Covid19 pandemic (referred to as 2019 and 2020 to ease reader understanding). For 2019, the dataset contains posts from 19,499 unique Twitter accounts, with an average of 4.7 tweets per user between 1 January 2019 and 30 January 2020. For 2020, the dataset includes posts from 19,504 unique Twitter accounts, that on average posted 4.8 tweets, the period from 31 January till 31 December 2020. In 2019 and 2020, the 10 most active user profiles respectively accounted for 7.4% and 6.8% of all tweets.

### Number of culture-related tweets pre-Covid and during Covid

The top-25 of most active users is by and large the same in 2019 and 2020, with some shifts in the rankings (see **Table -NL - 2**). A few users became less active or inactive in 2020, whereas some other users posted far more tweets in 2020 than in 2019. Among the latter, we find “Kunsten92”, an established interest organisation for the cultural and creative sector in the Netherlands. Its increased activity on Twitter in 2020 appears connected with the Covid-19 pandemic. National newspaper *de Volkskrant* was in 2020 still among the most active users accounts but posted considerably less tweets than in 2019. The activity of other newspapers remained more or less the same. Overall, we find more individual user accounts in the top-25 for 2020 (7 vs. 3 in 2019).

**Table -NL - 2.** Most active users in 2019 and 2020 according to number of tweets about culture<sup>a,b</sup>

2019 Users	Description	N Tweets	2020 Users	Description	N Tweets
1 <b>volkskrant</b>	<i>national newspaper</i>	1461	nieuwsbliknl	<i>news portal</i>	903
2 nieuwsbliknl	<i>news portal</i>	865	Ndnl	<i>Christian newspaper</i>	878
3 dvhn_nl	<i>regional newspaper</i>	756	dvhn_nl	<i>regional newspaper</i>	769
4 Vlaardingen24	<i>local news site</i>	624	lc_nl	<i>regional newspaper</i>	642
5 CultAvonturiers	<i>Platform cultural entrepreneurship</i>	600	Watskeburtnmi1	<i>neighbourhood app</i>	567
6 Grootsneek [28]	<i>local news site</i>	581	Boeken_Lezen	<i>website about books</i>	520
7 MoormanMark	<i>journalist (VK)</i>	538	<b>volkskrant</b>	<i>national newspaper</i>	514
8 Boeken_Lezen	<i>website about books</i>	498	recreatiegids	<i>leisure website</i>	476
9 ED_Cultuur [100]	<i>regional newspaper</i>	488	CultAvonturiers	<i>Platform cultural entrepreneurship</i>	424
10 recreatiegids	<i>leisure website</i>	430	MoormanMark	<i>journalist (VK)</i>	387
11 ndnl	<i>Christian newspaper</i>	414	onlineCW	<i>cultural website</i>	365
12 Watskeburtnmi1	<i>neighbourhood app</i>	413	Vlaardingen24	<i>local news site</i>	348
13 DeDrontenaar	<i>local newspaper</i>	396	coristus	<i>citizen</i>	286
14 SpeelmanLambert[57]	<i>manager non-profit</i>	363	SimonKadijk	<i>citizen, theologian</i>	282
15 bredavandaag	<i>local news site</i>	362	Kunsten92 [135]	<i>Interest org. cultural sector</i>	268
16 Dordtcentraal [75]	<i>local newspaper</i>	344	bredavandaag	<i>local news site</i>	261
17 lc_nl	<i>regional newspaper</i>	339	Dennispkroeze [30]	<i>Christian activist</i>	257
18 coristus	<i>citizen</i>	329	SchiedamVandaag	<i>local news site</i>	250
19 SchiedamVandaag	<i>local news site</i>	322	DeNater [-]	<i>citizen</i>	244
20 OU_Nederland [-]	<i>Open University NL</i>	317	EdwinvMeerkerk [32]	<i>cultural scholar</i>	241
21 onlineCW	<i>cultural journalism site</i>	313	mariabode_nl [96]	<i>catholic newspaper</i>	236
22 Politiek [-]	<i>political news site</i>	300	FJRoebersen [-]	<i>citizen</i>	230
23 MaassluisNu [83]	<i>local news site</i>	294	DeDrontenaar	<i>local newspaper</i>	222
24 ED_Regio [-]	<i>regional newspaper</i>	258	OmroepTilburg	<i>RTV station local</i>	215
25 UdensWeekblad [-]	<i>local newspaper</i>	251	Cultuurpers [117]	<i>cultural journalism site</i>	211

<sup>a</sup> Marked in lightgreen are accounts that appear in both years in the top 25.

<sup>b</sup> For accounts that only appear in one year, in brackets, it is indicated which position the account occupied in the other period.

A dash [-] means that the account was not among the 200 most popular accounts that tweeted about culture in the other period.



**Table -NL-3. Most popular user accounts tweeting about culture in 2019 and 2020<sup>a,b</sup>**

	2019 Users	Description	N followers	2020 Users	Description	N followers
1	KLM [0]	Royal Dutch Airways	2,320,801	vangoghmuseum	Museum	1,655,624
2	Vangoghmuseum	Museum	1,655,621	MinPres	prime minister	1,216,360
3	MinPres	prime minister	1,216,355	Devooravondtv [-]	TV talk show	1,071,151
4	volkskrant	national newspaper	785,144	HumbertoTan [-]	TV talk show host	920,933
5	Youpvan thek [-]	artist/comedian	765,600	volkskrant	national newspaper	785,156
6	telegraaf	national newspaper	655,507	Giels [-]	radio DJ	717,365
7	AT5	Amsterdam tv station	539,609	telegraaf	national newspaper	655,509
8	RTLnieuws	TV news program	518,722	LoesjeNL [-]	Ngo	567,065
9	AlexanderNL	Entrepreneur	489,530	AT5	Amsterdam tv station	539,598
10	KNVB	Dutch Football Association	475,777	RTLnieuws	TV news program	518,735
11	VI_nl	football news	472,629	AlexanderNL	Entrepreneur	489,526
12	claudiadebreij	artist/comedian	461,772	KNVB	Dutch Football Ass.	475,779
13	AmsterdamNL	City of Amsterdam	411,276	VI_nl	football news	472,636
14	ADnl	national newspaper	399,398	claudiadebreij	artist/comedian	461,772
15	nrcslimleven	national newspaper	384,758	astro_andre [-]	Dutch astronaut	459,025
16	parool	Amsterdam newspaper	365,838	AmsterdamNL	City of Amsterdam	411,277
17	Nieuwsuur	TV news program	356,515	ADnl	national newspaper	399,411
18	NOSsport [-]	TV sports program	335,536	nrcslimleven	national newspaper	384,755
19	op1npo	TV talk show	320,910	parool	national newspaper	365,841
20	3FM	music radio station	309,489	Nieuwsuur	TV news programme	356,523
21	Gerardjoling [-]	artist/singer	280,017	op1npo	Tv talk show	320,913
22	DeSpeld	online news (satirical)	277,654	3FM	music radio station	309,489
23	Jorisluyendijk [-]	Journalist	268,260	Politie [-]	Police	295,617
24	thierrybaudet	politician	261,627	DeSpeld	online news satirical)	277,665
25	Rijksmuseum [27]	Museum	255,019	thierrybaudet	Politician	261,627

<sup>a</sup> Marked in blue are accounts that appear in both years in the top 25.

<sup>b</sup> For account that only appear in one year, in brackets, it is indicated which position the account occupied in the other period. A dash [-] zero means that the account was not among the 200 most popular accounts that tweeted about culture in the other period.

### Number of followers pre-Covid and during Covid

shows the 25 most popular user accounts – according to the number of followers) - that tweeted about culture 2019 and 2020. In the pre-Covid period, the list features the Dutch Royal Airways (KLM) at the top, followed by the Van Gogh Museum, and the Dutch prime minister. The list of most popular user accounts also includes several famous Dutch artists - two comedians and a singer, the Dutch football association (KNVB), and a right-wing politician, next to various news media and TV shows.

During the Covid-19 period in 2020, the ranking looks slightly different. Most notable is the disappearance of the KLM from the 2020 ranking, and the emergence of a popular TV talk show, a talk show host, a radio DJ, and the creative collective Loesje among the highest-ranking user accounts. Loesje is a Dutch based, international collective which has become famous by its public posters which feature humoristic, and often mildly critical statements.

Comparing the rankings in **Table -NL - 2** and **Table – NL -**, the national newspaper *de Volkskrant* is the only user account that occupies a top position in both rankings.

### User engagement and interaction

Almost half of the culture-related tweets in 2019 (45.7%) and 2020 (48.7%) was liked by one or more other users. For retweets and replies, these percentages were much lower. In 2019 and 2020, about a quarter of the tweets was retweeted (24.3% versus 23.8%), whereas respectively 25% and 29.5 percent received a reply. In 2019, 40 percent of the tweets featured mentions of other user

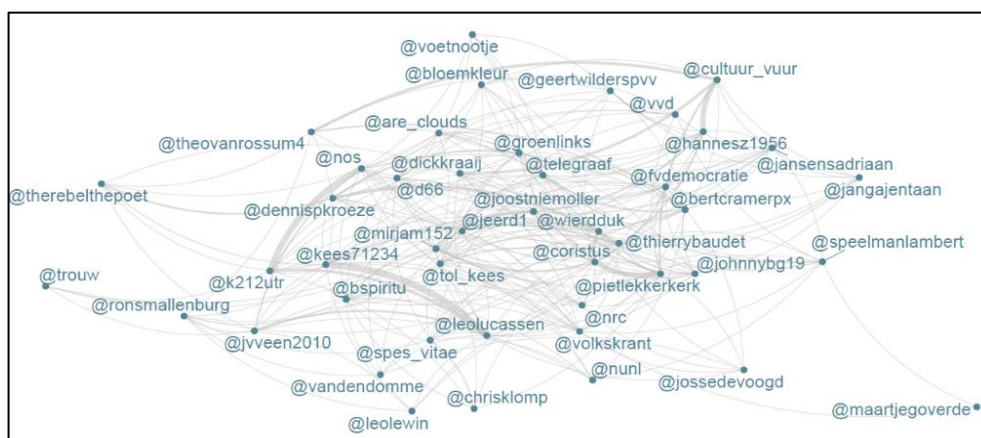
accounts, as opposed to 47 percent 2020. On average, tweets contained 1.7 and 1.8 mentions respectively.

User engagement and interaction slightly increased during the Covid-19 period, considering the somewhat higher proportions of tweets that received likes and replies, the mean number of likes (2.06 versus 3.84), retweets (0.89 versus 1.07), and replies per tweet (0.51 versus 0.72), and the percentage of tweets that contained mentions of other users (40% vs. 47%). This may be connected with the Covid-19 situation and the fact that people in the Netherlands were more active on social media platforms in general and more frequently accessed Twitter to interact with others as well.

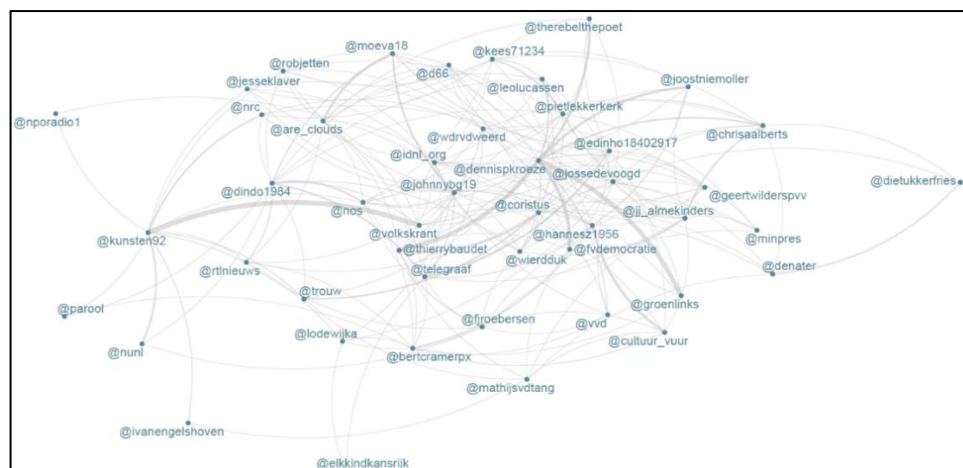
Still, in both years, more than half of all tweets had no visible connection to posts of other users (e.g. retweet, mention, or reply). Many users who tweeted about culture primarily used Twitter to voice their own opinions or to share other information they deemed newsworthy.

**Figure - NL - 2** and **Figure - NL - 3** show the networks of mentioned user-accounts that tweeted about culture before and during the Covid-19 pandemic. Cultural institutions like museums or artists do not appear in the networks. Instead, the networks include many of the “usual suspects” on Twitter, like political parties (@fvdemocratie, @groenlinks, @d66, @vvd), politicians (@thierrybaudet, @geertwilderspvv, and, in 2020 also @robjetten, @jesseklaver, @lodewijka), news media (@nos, @telegraaf, @nrc, @volkskrant, @nu.nl, @trouw), journalists (@joostniemoller, @chrisaalberts), and some very active scholars, such as Leo Lucassen, a social historian specializing in migration.

**Figure - NL - 2.** Network of mentioned user accounts in 2019 / pre- Covid19



**Figure - NL - 3.** Network of mentioned user accounts in 2020 / during Covid19

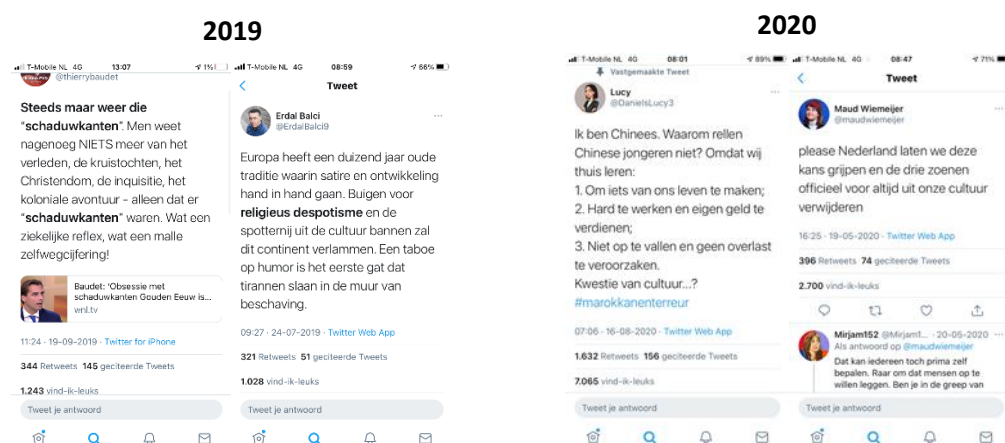


In both years, we also find activist groups like @cultuur\_vuur, which claims to fight for the victory of Dutch culture and Christian traditions and against socialism, Islam (“islamisering”), and “climate force” (“klimaatdwang”). Another example is @TheRebelThePoet, the account of poet and anti-racism activist, Jerry King Luther Afriyie, who founded Nederland wordt Beter / Kick-Out Zwarte Piet (*The Netherlands will get better / Kick Out Black Pete*).

In 2020, several new activist accounts popped up in the mentions network, for example @idnl.org, an anti-migration movement which strives to preserve the Dutch identity against foreign influences. The 2020 network also features the accounts of the Dutch prime-minister (@minpres) and minister of Culture, Education and Science (@ivanengelshoven), as well as @kunsten92, the interest organisation for the cultural sector which became more active during the Covid-crisis. As note before, this increased activity is probably due to Covid-19 related discussions about crisis support for the cultural sector.

**Figure - NL - 4** shows some of the most liked tweets in 2019 and 2020. In the most liked post of 2019 Dutch right wing politician Thierry Baudet laments the increasing attention for the dark sides of the Dutch colonial past and the Dutch Golden Age. Both in 2019 and 2020, many favourite tweets were either posts by Baudet, or critical responses to his viewpoints. Another top 3 post in 2019 came from the Dutch-Turkish writer and journalist Erdal Balci who warns for religious despotism that seeks to ban mockery from culture. The most liked post in 2020 was a critical statement on bad behaviour of some youth of Moroccan descent by a user who identified as Chinese. The third most liked post in 2020 was more light-hearted, pleading to use the opportunity provided by Covid-19 to forever abandon the Dutch greeting habit of kissing someone three times.

**Figure - NL - 4.** Most liked tweets in 2019 and 2020



#### 4. Hashtags and keyword analysis

When excluding references to #cultuur ( $n = 2,638$ ) and #culture ( $n = 66$ ), only 45,907 tweets that were posted in 2019 and 2020 (14% of total tweets) featured hashtags. On average, each tweet contained 1.8 hashtags and the number of unique hashtags was 13,772. The large majority of tweets with keyword culture did not use hashtags, and the tweets that did carry hashtags used a wide variety of terms. Table 5 gives some figures about the use of hashtags in 2019 and 2020. In 2020, the percentage of tweets with hashtags (12%) was lower than in 2019 (16%).

Many of the most frequently used hashtags that occurred with keyword culture in 2019 and 2020 were the names of cities and municipalities: #sneek, #vlaardingen, #schiedam, #amsterdam, #tilburg, #maassluis, #utrecht, #denhaag, #wassenaar, #tilburg, #castricum, #enschede, # groningen, #heiloo, #zoetermeer. Remarkable is the high frequency of hashtags carrying the names of smaller cities. In 2019, #Sneek (n = 605) was used far more often in connection with keyword culture than #Amsterdam (n = 185).

**Table - NL - 4** presents the twenty most popular hashtags before and during the pandemic, with the exception of the afore mentioned names of municipalities, #cultuur, and #culture. In both years, #kunst (art) #vacature (vacancy), and #nieuws (news) belonged to the top 5. Other hashtags that were used relatively often in both periods are #erfgoed (heritage) and #diversiteit (#diversity). The most used hashtags in 2020 include three pandemic related hashtags: #corona, #coronacrisis and #coronavirus, while the emergence of #netflix might be connected with Covid-19 as well. The top-20 for 2020, also features several new #hashtags referring to photography and film: #fotografie, #photography and #portretten (portraits).

**Table - NL - 4.** Use of hashtags and twenty most common hashtags in 2019 and 2020<sup>a</sup>

	2019	2020	Top-20 2019	N	Top-20 2020	N
Total hashtags (n)	24,819	21,088	#vacature	288	#kunst	257
Unique hashtags (n)	8,735	6,962	#kunst	285	#corona	232
Mean hashtags per tweet	1.7	1.8	#nieuws	234	#vacature	194
% tweets with hashtags	16%	12%	#atnavacature	154	#fotografie	141
Max hashtags per tweet	14	12	#lokaal	147	#nieuws	118
#cultuur	1,246	1,392	#omroep tilburg	146	#photography	111
			#boekenlezen	138	#coronacrisis	100
			#erfgoed	122	#nieuws	118
			#fondsenwerving	105	#erfgoed	96
			#music	104	#coronavirus	94
			#diversiteit	88	#film	94
			#museum	86	#portretten	92
			#taal	73	#expositie	85
			#persbericht	71	#vrouwen	82
			#vrouwen	69	#netflix	81
			#duic	67	#zwartepiet	80
			#regio	66	#diversiteit	78
			#fvd	63	#tkcultuur	76
			#theater	62	#lokaal	75
			#sport	57	#zzp	72

<sup>a</sup> Hashtags with names of cities/municipalities are not included in the top-20. Marked in green are #hashtags that appear in the top-20 of both years in the top 20.

**Figure - NL - 5** presents the connections between the most commonly used hashtags in tweets about culture. In both periods, hashtags related to cultural offerings - #kunst (art), #muziek (music); #theater; #museum and major cities (#amsterdam, #utrecht) are at the centre of the network. In both periods we find a small, separate cluster of hashtags related to Tilburg and its local broadcasting station (#omroep tilburg). The 2020 network shows some clear differences compared to the 2019 network. Corona-related hashtags, especially #corona became very central in culture-related tweets, while also #film and #Netflix emerged as important co-occurring hashtags. #Diversity moved centre stage in the 2020 network, showing more linkages with multiple other hashtags. The 2020 network also features a novel hashtag cluster concerning photography.

[illegible]

## 2020

The most common keywords are very similar in both years. Also, keywords that are no longer in the top 20 for 2020 are still among the forty most used keywords in 2020. The high prevalence of words like “sector”, “instelling” and “geld” in 2020, is likely due to the Covid-19 crisis and related discussions on Twitter about crisis support for the cultural sector.

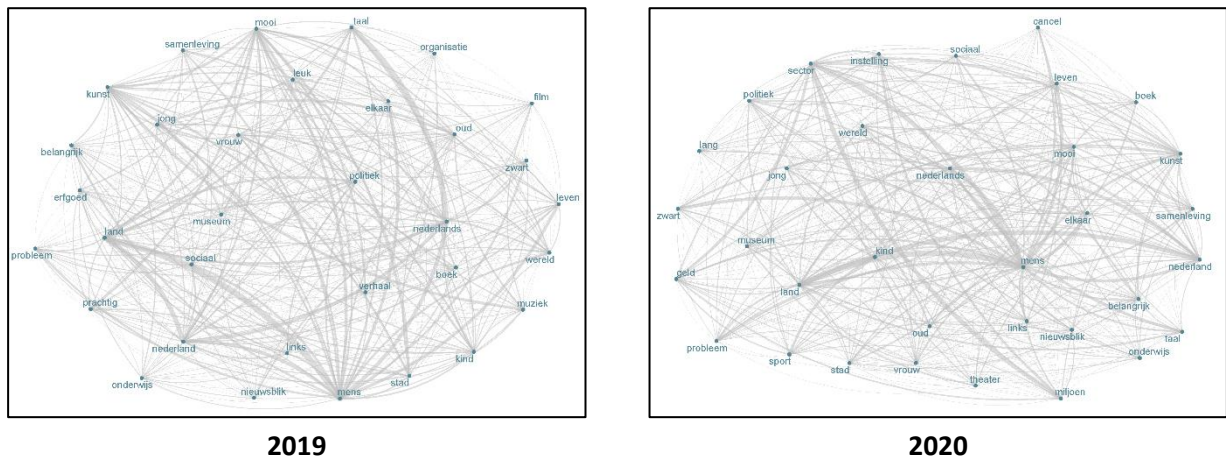
Keywords 2019			Keywords 2020		
Dutch	English	N	Dutch	English	N
goed	<i>good</i>	6,107	goed	<i>good</i>	6,961
mens	<i>human/people</i>	5,026	mens	<i>human/people</i>	6,720
nieuw	<i>new</i>	5,024	jaar	<i>year</i>	4,709
jaar	<i>year</i>	4,865	nieuw	<i>new</i>	4,650
Nederlands	<i>Dutch</i>	3,648	sector	<i>sector</i>	4,551
mooi	<i>beautiful</i>	3,544	country	<i>land</i>	3,614
groot	<i>big</i>	3,303	groot	<i>big</i>	3,562
land	<i>country</i>	3,080	Nederlands	<i>Dutch</i>	3,555
kunst	<i>art</i>	3,075	kunst	<i>art</i>	3,258
Nederland	<i>Netherlands</i>	2,666	mooi	<i>beautiful</i>	3,219
echt	<i>real / really</i>	2,421	echt	<i>real / really</i>	2,986
vrouw	<i>woman</i>	2,029	Nederland	<i>Netherlands</i>	2,970
kind	<i>child</i>	1,905	tijd	<i>time</i>	2,957
erfgoed	<i>heritage</i>	1,879	instelling	<i>institution</i>	1,983
tijd	<i>time</i>	1,791	politiek	<i>politics</i>	1,826
dag	<i>day</i>	1,721	leven	<i>life</i>	1,791
boek	<i>book</i>	1,693	belangrijk	<i>important</i>	1,783
stad	<i>city</i>	1,647	geld	<i>money</i>	1,776
politiek	<i>politics</i>	1,579	zwart	<i>black</i>	1,776



**Figure NL - 6** shows the most common keywords and their relations in both years, revealing many co-occurrences of the words country (*land*), human (*mens*), Dutch (*Nederlands*) and Netherlands in both years. During Covid, the words sector, institution (*instelling*), and million (*miljoen*) were very often used in combination with one another. During the pandemic child (*kind*) appears to have become a more central keyword, while also the word cancel emerged as a prominent term.

**Figure - NL - 6.** Top 40 Keyword network of tweets, 2019 and 2020

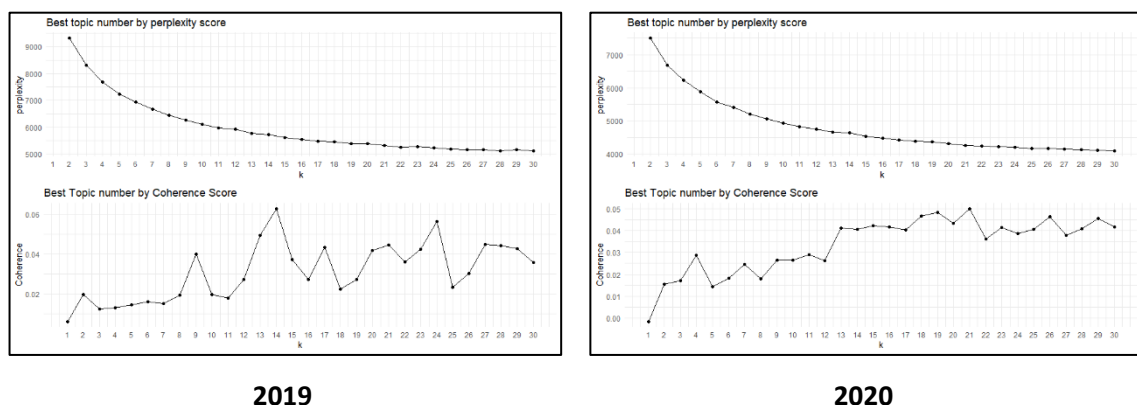
(less informative words like new, year, time, and week have been excluded)



## 5. Topic modelling

This section explores the most prevalent culture-related topics that were discussed on Twitter. Separate analyses for 2019 and 2020 were conducted to trace any temporal developments. To detect the optimal number of topics  $k$  for Mallet topic modelling in R, we inspected the results for  $k=5$  to  $k=30$ . Based on measures for perplexity and coherence (see **Figure - NL - 7**) as well as semantic interpretation of the results for 2019 and 2020, a model with respectively 17 and 19 topics appeared to best capture the relevant topics in terms of perplexity, coherence, and topic variety.

**Figure - NL - 7.** Best topic number  $k$  in 2019 and 2020



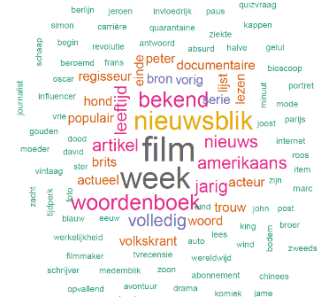
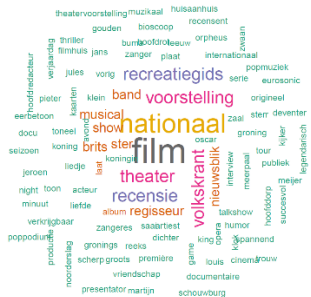
Below we give a short description and comparison of the topics identified in the period before and during the pandemic. Many topics (14 out of 22) were prevalent in both periods, but eight topics were only manifest in either 2019 or 2020. Overall, the topics can be divided into four overarching clusters of topics: (1) *Cultural Offerings and News*, (2) *Places and Destinations*, (3) *Policies and Politics*, and (4) *Culture and Society*. Word clouds for topics that were found for 2019 and 2020 are respectively presented in the left hand and right column of the following pages.



## CULTURAL OFFERINGS AND NEWS

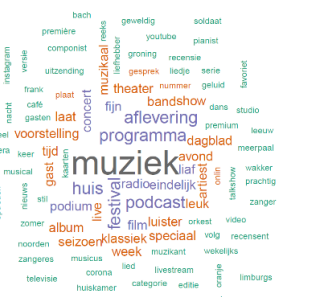
### 1. Film, TV and theatre (2019/2020).

This topic concerns film, tv, and performing arts offerings, with film as a focal point in both years. In 2019, the topic has a wider scope, also incorporating words like theatre, performance, musical, show, and band, whereas in 2020, the focus seems to be more exclusively film and other audio-visual offerings such as documentaries, and tv series. The tweets also refer to the origins of offerings (e.g. British, American) as well as to “makers” using words like director, actor, star, and presenter.



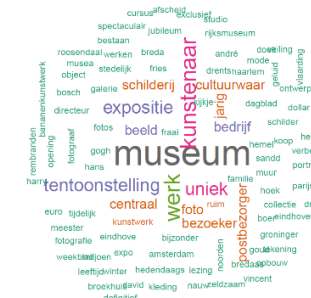
### 2. Music offerings (2019/2020).

This topic is all about musical offerings in both years. However, in 2019, the focus is mainly on live musical events/locations, considering the prevalence of words like festival, place, cafe, theatre, church, venue, and event. In 2020, these words are also present, but in addition we find many words that refer to other modalities for enjoying music, such as (at) home, podcast, radio, listening, album, living room, livestream, online, and YouTube.



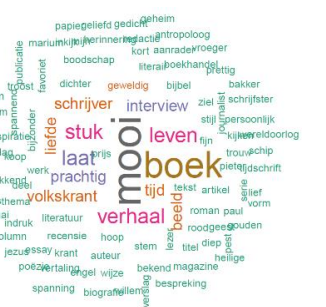
### 3. Museums and visual arts (2019/2020).

This cluster focuses in both years on museums, exhibitions (expositie; tentoonstelling), visual artists (kunstenaar, fotograaf), art works (werk, beeld, schilderij, kunstwerk, foto), attendance (bezoeker, bezoek) and activities (lezing, tour). The tweets also refer to museum directors, collections, and awards.



### 4. Books and literature (2019/2020).

This topic deals with books and literature, with words like book (boek), story (verhaal), writer (schrijver), poet (dichter), novel (roman), literature, poetry (poëzie). It includes many (positive) adjectives - e.g. beautiful (mooi) and other terms connected with literary reviewing and

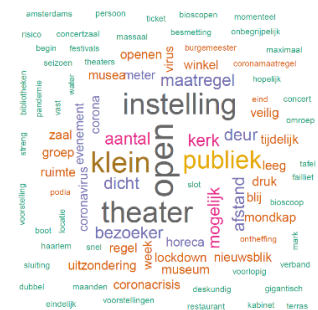


[illegible]

This topic was only found in 2019. It contains tweets about fame, cultural awards and competitions, using terms like famous (*bekend*), prize (*prijs*), success, name, winner (*winnaar*), jury, songfestival, award ceremony (*uitreiking*), favourite (*favoriet*), and prestigious (*prestigieus*).

This topic, for obvious reasons, only emerged in the 2020 data. It centers on Covid-related restrictions and measures for visiting cultural venues - e.g. theaters, museums, zaal (*hall*), stages (*podia*), festivals, cinemas (*bioscopen*), church (*kerk*), the catering industry (*horeca, restaurants*), and shops (*winkel*). Tweets about restrictions and measures include terms like open, closed (*dicht*), number (*aantal*), possible (*mogelijk*), distance (*afstand*), mouthmask (*mondkap*) measurement (*maatregel*), rule (*regel*), lockdown, and safe (*veilig*).

Although by no means a new phenomenon, this topic became only manifest in the culture-related tweets for 2020. This attention for digital projects and initiatives could well be connected with the Covid-19 crisis to which we find some references like corona time (*coronatijd*), alternative (*alternatief*), map (*kaart*), and plan. The word art (*kunst*) is very central to this topic, which furthermore includes words like online, digital, website, project, initiative, activity, information, collaboration (*samenwerking*), and creativity.



## PLACES AND DESTINATIONS

### 8. Urban environment (2019).

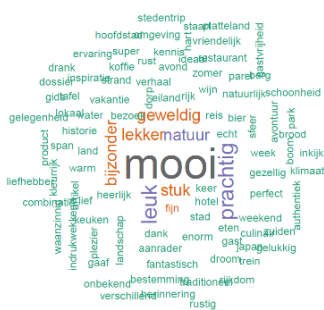
This topic was only manifest in the pre-Covid tweets, incorporating references to art, major Dutch cities (Amsterdam, Utrecht, Rotterdam, The Hague, Tilburg, Eindhoven), and many terms relating to the urban environment and city planning, such as city (*stad*), centre (*centrum*), house (*huis*), space (*ruimte*), area (*gebied*), plan, green (*groen*), neighbourhood (*buurt*), quarter (*kwartier*, *wijk*), inner city (*binnenstad*), mayor (*burgemeester*), alderman (*wethouder*), citizen (*bewoner*), breeding ground (*broedplaats*), architecture, spatial (*ruimtelijk*), and map (*kaart*).

Although we did not find this topic for 2020, some of the words in this cluster are incorporated in the Cultural Policy topic (10) for 2020.



### 9. Tourism and Leisure (2019).

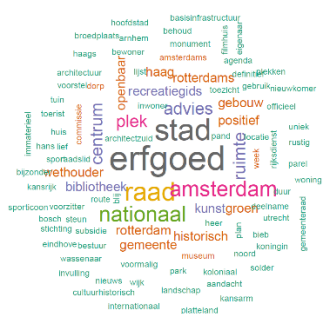
This topic includes many positive qualifications - e.g. beautiful (*mooi*), magnificent (*prachtig*), fun (*leuk*), special (bijzonder) and delicious (*lekker*) – together with words like nature (*natuur*), journey (*reis*), vacation (*vakantie*), hotel, weekend, summer (*zomer*), city trip (*stedentrip*), beach (*strand*), country side (*platteland*), restaurant, cuisine (*keuken*), climate (*klimaat*), visit (*bezoek*) and destination (*bestemming*). This topic was not prevalent in the 2020 tweets, which is probably due to leisure and travel restrictions during the Covid-pandemic.



## POLICIES AND POLITICS

### 10. Cultural policy (2019/2020).

This topic concerns cultural policy at the level of Dutch cities, provinces, and the national government, regarding cultural heritage (*erfgoed*; *monuments*), archives, museums, and the arts (*kunst*). It features words like institution (*instelling*), city (*stad*), council (*raad*), advice (*advies*), foundation (*stichting*), province (provincie), subsidy (*subsidie*), money (*geld*), million (*miljoen*), budget-cut (*bezuiniging*), and audience (*publiek*). The topics looks quite similar in both years, but in 2020, this topic also incorporates items that appeared in the 2019 'Urban environment' cluster.







## 15. Societal impacts Covid-19 (2020).

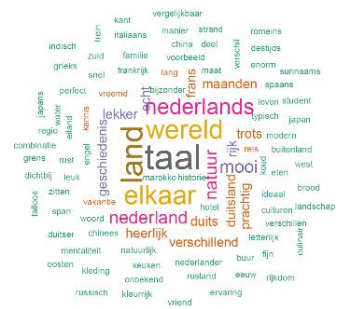
Although art (*kunst*) is a central term here, this topic is much broader than topic 11, which was all about financial support for the cultural sector. The present topic contains many references to the (financial) problems and the future (*toekomst*) of other sectors - notably healthcare (*zorg; gezondheidszorg, ziekenhuis*), but also education (*onderwijs*), sport, tourism, aviation (*luchtvaart*), agriculture (*landbouw*), and the economy (*economie*) more generally, jobs (*baan*) and employment (*werkgelegenheid*). Tweets appear to point to the importance (*belangrijk*) of and need for government support and policy measures, considering words like *overheid* (government) and corona policy (*coronabeleid*), but compared to topic 11 this is not that clear-cut.



## CULTURE AND SOCIETY

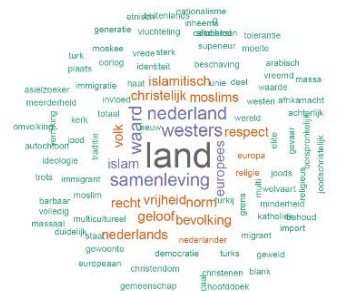
**16. Languages and cultures (2019/2020).**

This topic prominently features words like language (*taal*), Dutch (*Nederlands*), world (*wereld*), history (*geschiedenis*), word (*woord*), century (*eeuw*) century), modern, French (*Frans*), German (*Duits*), Germany (*Duitsland*), and American (*Amerikaans*), but also includes mentions of other languages, countries, and cultures. The 2019 tweets relatively often refer to academic education and research (e.g. university student, knowledge, education, docent, scientist, college). The 2020 tweets more often highlight the attractiveness of countries and cultures from a tourist perspective, pointing to their distinct cultural characteristics, but also talking about specifics of the natural environment.



### 17. Religion in multicultural society (2019/2020).

This topic incorporates words like country (*land*), society (*samenleving*), value (*waard*), Western, Islam, Muslims, Christian, Jewish (Joods), religion, Netherlands, norm, respect, right (*recht*), freedom (*vrijheid*), hate (*haat*), backward (*achterlijk*), enrichment (*verrijking*), and civilisation (*beschaving*). The tweets reflect the ongoing societal debates and tensions regarding the position of

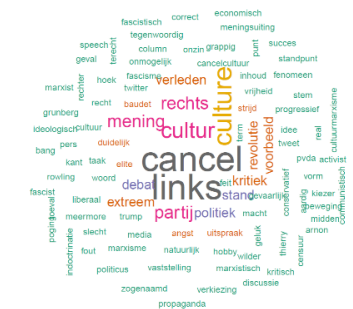






## 21. Reflections on present, past and future (2019/2020).

This topic focuses on the recent phenomenon of cancel culture, including the words cancel, culture, opinion (*mening*), criticism (*kritiek*), mistake (*fout*), wrong (*verkeerd*), statement (*uitspraak*), extreme (*extreem*), censorship (*censuur*), as well as terms referring to media, tweet, column and twitter and some specific examples such as Thierry, Baudet, Trump, and Rowling. The topic contains many references to the political arena, for example party (*partij*), politics (*politiek*), left (*links*), right (*right*), elections (*verkiezingen*), voter (*kiezer*), marxism, communism, and fascism.

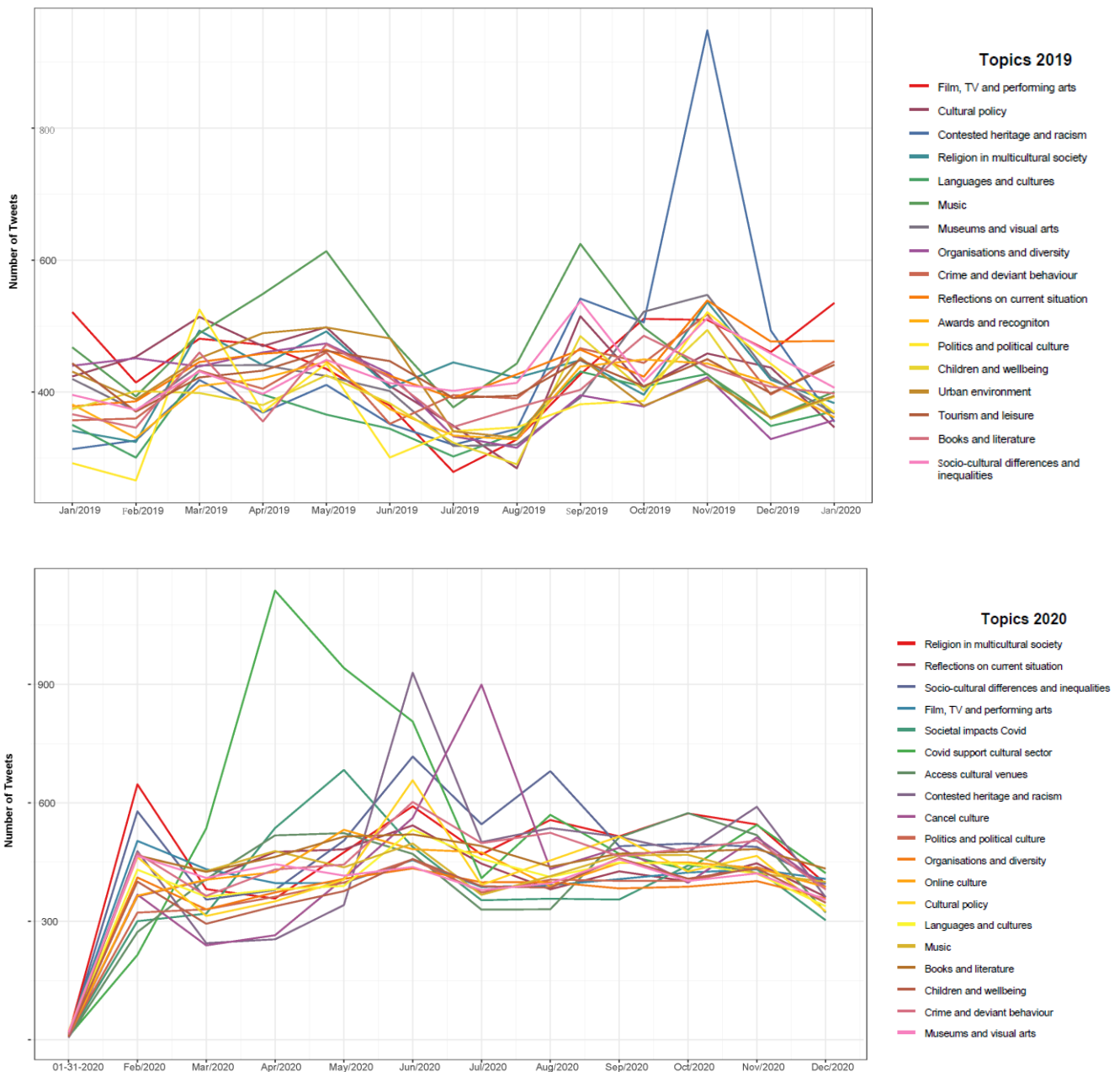


The attention on Twitter for the various topics is not evenly distributed across time but shows considerable fluctuations. In 2019, the summer period (June-July-August) shows a clear decline in the number of tweets about most topics, in particular for *Film*, *TV and performing arts*, *Politics and political culture*, *Cultural policy*, and *Children and wellbeing*. In December, towards the Christmas break, attention to most topics is also at a low. The picture for the summer of 2020 is different; although attention to most topics is also relatively low, the decline in attention is limited and mainly confined to the month July. Moreover, whereas during the summer of 2019 none of the topics witnessed a peak in attention, attention to *Contested heritage and racism*, and *Cancel culture* strongly increased in June and July 2020, respectively.

75

Covid-crisis, the graph shows more and bigger variations for many of the topics. In April and May 2020, we find clear peaks in the discussion about *Covid support for the cultural sector* and *Access to cultural venues*, while in June 2020, attention to *Social inequalities and differences* and *Cultural policy* was at a high. In November 2020 we observe again a rise in attention for *Contested heritage and racism*, although far less pronounced than in 2019.

**Figure - NL - 8. Topic prevalence and development in 2019 and 2020**



## 6. Discussion and conclusion

### Continuity in topics discussed before and during Covid-19

Our comparison of topics before and during the Covid-19 pandemic showed that there is a lot of continuity in the topics that Twitter users talk about.

Firstly, an important part of the cultural-related tweets in both periods concerned cultural offerings

and news, divided into clear sections - *Film, television and theatre, Books and literature, Music, and Museums and visual arts* - like the ones that are commonly used in Dutch cultural journalism and news media. The prevalence of this structure is likely due to the continuous stream of tweets from a variety of Dutch news media that report on the new offerings, events, and development in these longstanding fields of cultural production. However, discussions in the Dutch Twitter sphere on these specialized cultural fields are by no means limited to news and opinions about new cultural products and events. In both periods, *Cultural policy* is a recurring subject of debate, often triggered by policy measures, decisions, or advices, by local or national governmental, funding, and advisory bodies. In both years we also find significant attention to *Languages and cultures*, addressing features of both Dutch and foreign language(s) and culture(s).

Secondly, both before and during the pandemic, a very large portion of the culture-related tweets concerns wider political and socio-cultural issues that have also otherwise been widely debated in Dutch society in recent years. For both periods, we find a number of highly similar topics, that we labelled *Politics and political culture, Organisations and diversity, Contested heritage and racism, Religion in a multicultural society, Socio-cultural differences and inequalities, Crime and deviant behaviour, and Children and wellbeing*. Some of these topics seem more clearly linked to the activity of political actors, or (cultural) organisations than others. It would be interesting to take a closer look into the share of different types of users in these discussions, as well as the nature and contents of their tweets.

### **Differences in topics discussed before and during Covid-19**

Next to continuity, we also found clear differences in the topics discussed on Twitter before and during Covid. Many of these differences appear to a large extent caused by the pandemic.

First, a number of new topics emerged in 2020 that capture various pandemic-related developments and discussions in the cultural sphere: *Access to venues during Covid, Online culture, Covid support cultural sector, and Societal impacts Covid*.

Second, the disappearance of *Tourism and leisure* as a distinct topic in 2020 is probably related to Covid-19 as well, although we still found some attention to features of tourist destinations (e.g. cuisine and restaurants, attractiveness of natural environment) as part of the *Languages and cultures* cluster. The connection with the Covid-crisis is less clear cut for the disappearance of *Awards and recognition* and the *Urban environment* as distinct topics. Still there might be a connection, e.g. the absence of physical award ceremonies and shifting priorities of city governments during Covid could have played a role here.

Thirdly, we should note, that although we found highly similar topics in both years, the focal points of some topics changed during the Covid-19 period. This was clearly the case for *Politics and political culture, Crime and deviant behaviour* and *Socio-cultural differences and inequalities*, as well as for *Film, TV and theatre*, which concentrated more on audio-visual offerings, and *Music*, which in 2020 included more attention to alternative, 'corona proof' modalities for enjoying music.

Last but not least, *Cancel culture* appeared as a novel topic during the Covid-19 crisis, which clearly boomed in the summer of 2020. Once more, it would be worthwhile to inspect what kind of users discussed this topic and have a closer look at the contents of their tweets.

### **'Missing' topics**

Some topics one would expect to find in connection with key word culture did not emerge – at least not as distinct topics - from our data, neither before nor during Covid-19. These "missing" topics include, among other things, gaming and lifestyle elements such as sports, food, and fashion. Sports got some attention as part of the *Children and wellbeing* cluster, but not in the sense of the cultural dimensions of sports.

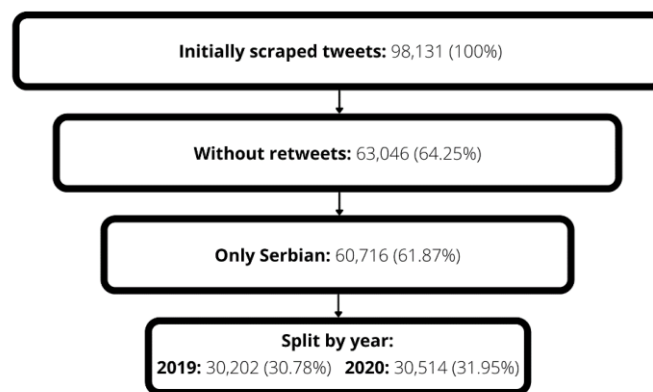
## Findings for Serbia

### 1. Data description

Tweets were scraped using premium Twitter API subscription and the criteria for scraping were that a) they were published in Serbia between the 1<sup>st</sup> of January 2019 and the 31<sup>st</sup> of December 2020 (including both dates); and that b) they contained any form of word "culture" in the Serbian language, either in Latin or Cyrillic script ("kultura", "kulture", "kulturi", "kulturu", "kulturom", "kulturo", "култура", "културе", "култури", "културу", "културом", "културо").

As can be seen in **Figure - RS - 1** below, a total of 98,131 tweets were scraped initially, while after filtering out retweets, 63,046 remained. The tweets contained 32 different languages out of which Serbian<sup>5</sup> was dominant (60,716 tweets, or 96.3% of the sample). Other languages with more than 200 tweets in the sample included Slovene (644, or 1.02%), Hungarian (491, or 0.78%), English (227, or 0.36%) and Czech (218, or 0.35%). Tweets in other languages constituted around 2.5% of the sample. Only tweets written in Serbian (60,716) were further analyzed. Since it was expected that the Twitter communication about culture would be significantly impacted by the Covid-19 pandemic, the sample was further split by year to parts of roughly the same size (30,202 tweets from 2019 and 30,514 tweets from 2020) to allow for a comparison between 2019 and 2020.

**Figure - RS - 1.** Steps taken in data cleaning and preparation. Percentages represent share of the originally scraped tweets (100%)

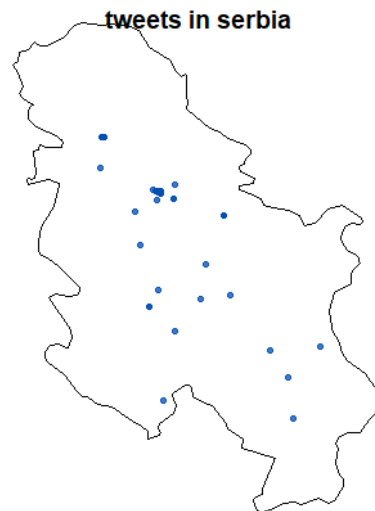


As far as the geographic location of the users is concerned, 99.12% (60,178) of analyzed tweets contained profile location, while less than 1% contained geographic coordinates. According to the profile location, most of the analyzed tweets came from Belgrade (36.35%), Novi Sad (6.42%), Niš (2.19%), and Kragujevac (1.05%), while 17.42% of tweets had their profile location listed as Serbia.

**Figure - RS - 2** shows the location of the tweets that contained geographic coordinates.

<sup>5</sup> Twitter recognized only 7,607 tweets as explicitly written in Serbian while the language of the rest was "undefined". Inspection of the tweets with an "undefined" language showed that these were either written in an informal variant of Serbian, or one of the languages that were previously considered Serbo-Croatian. As it is, due to an overwhelming linguistic similarity, impossible to precisely determine which language they were written in, all of these languages were considered to be Serbian throughout the rest of the analysis.

**Figure - RS - 2.** Geographic distribution of analyzed tweets (those with available coordinates)



## 2. User analysis

As previously mentioned, the sample was split by year and a separate analysis of users and their interaction was conducted for both years. Results of the analyses can be found below. The overview of the most important user rankings for both years can be found in **Table - RS - 1**. **Figure - RS - 3** and **Figure - RS - 4** show the networks of mentions between the users for both years with the strength of the connection (number of mutual mentions in the analyzed tweets) represented by the thickness of the line.

The 2019 dataset contained tweets published by 4,925 unique users out of which 91 were verified. The average number of tweets per user was 6.37. The most prolific users of the period were mainly media portals (@AktuelnoNet, @Krstarica, @juznevesti, @OnlineDanas, @N1infoBG, @VestiDobre, @NovostiOnline, @ekultura011) with the famous Serbian film director Srđan Dragojević (@srdjandrango) and an unverified private profile @Mackapersijska being exceptions. Their daily tweeting rate ranged from 1 (@srdjandrango) to 2.44 (@AktuelnoNet). As far as ranking by the number of followers is concerned, the situation was different with only 4 media related profiles entering the top 10 (@Blic\_online, satirical news outlet @njuznet, @b92vesti, and @OnlineDanas). The user with the most followers (687,016) in 2019 was the domestic pop singer Vlado Georgiev (@VladoGeorgiev). The Top 10 ranking by followers also contained Serbian president Aleksandar Vučić (@avucic), writer Isidora Bjelica (@isidorabjelica), opposition leader Vuk Jeremić (@jeremic\_vuk), actress Bojana Maljević (@BojanaMaljevic), and pop band Ničim Izazvan (@NichimIzazvan).

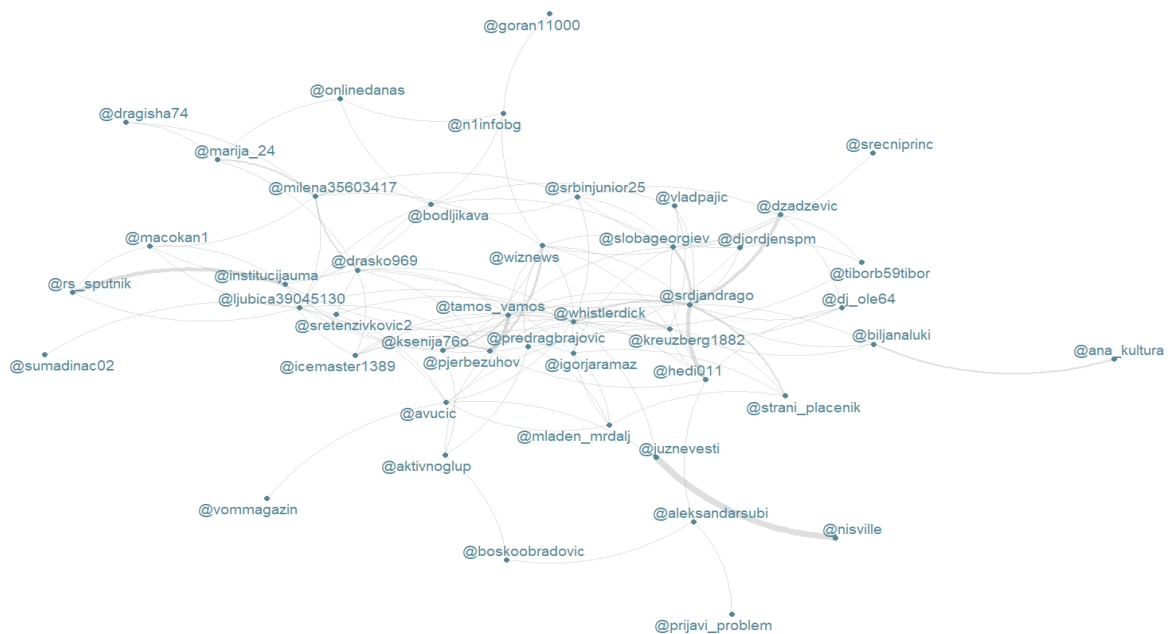
The 2020 dataset contained tweets published by 5,158 unique users with the same number of verified profiles (91) as in 2019. The average number of tweets was 6.14 with half of the top 10 ranking by the number of tweets remaining the same as in 2019 (@AktuelnoNet, @juznevesti, @OnlineDanas, @N1infoBG, @srdjandrango). New additions were newly founded media outlet @novarsonline, pro-government tabloid @InformerNovine, a local media outlet from Novi Sad @nshronikars, and two unverified profiles @AleksandarRak15 and @Vucko\_Kurijacki. Their daily tweeting rate ranged from 1.05 (@nshronikars) to 3.15 (@novarsonline). Six of the 2020 top 10 users by the number of followers remained the same as in 2019 (@VladoGeorgiev, @njuznet, @jeremic\_vuk, @OnlineDanas, @BojanaMaljevic, @NichimIzazvan); they were joined by the TV host Ivan Ivanović (@veceivan), who was also the most popular user tweeting about culture in 2020.

(742,416 followers), famous turbo folk singer Svetlana Ražnatović (@CecaRaznatovic), politician Aleksandar Šapić (@AcaSapic), and the N1 media outlet (@N1infoBG).

**Table - RS - 1.** Comparison of users with most published tweets about culture and most followers between the years. Highlighted accounts appear in both periods

2019		2020		2019		2020	
User	N° of tweets	User	N° of tweets	User	N° of followers	User	N° of followers
AktuelNoNet	891	novarsonline	1149	VladoGeorgiev	687016	veceivan	742416
Krstarica	881	N1infoBG	875	Blic_online	459019	VladoGeorgiev	687016
juznevesti	677	AktuelNoNet	741	njuznet	390471	njuznet	390471
OnlineDanas	590	AleksandarRak15	652	avucic	361431	CecaRaznatovic	299002
N1infoBG	581	OnlineDanas	651	isidorabjelica	256506	jeremic_vuk	190922
VestiDobre	526	srdjandrango	583	b92vesti	206705	AcaSapic	174469
Novostionline	522	Vucko_Kurjacki	490	jeremic_vuk	190925	OnlineDanas	166841
ekultura011	439	InformerNovine	439	OnlineDanas	166840	BojanaMaljevic	158365
Mackapersijska	426	juznevesti	437	BojanaMaljevic	158404	NichimIzazvan	154489
srdjandrango	365	nshronikars	382	NichimIzazvan	154489	N1infoBG	154160

**Figure - RS - 3.** Network of mentions from the 2019 dataset





The graph displays a network of Twitter accounts. The nodes are labeled with their usernames, and the edges represent connections between them. The network is highly interconnected, with many nodes having multiple connections. The layout is circular, with nodes arranged in a roughly circular pattern around a central cluster. Some nodes are more isolated than others.

Key nodes and their connections include:

- @luznevesti** (top) connected to **@avucic** and **@pjerbezuhov**.
- @avucic** (top) connected to **@luznevesti**, **@pjerbezuhov**, and **@nikolatanasic**.
- @pjerbezuhov** (top) connected to **@avucic**, **@luznevesti**, **@nikolatanasic**, **@ljubica39045130**, **@mladen\_mrdalj**, **@borovic3**, **@tonistankovic**, **@dazivimsan**, **@dzadzevic**, **@vojvotkinja\_m**, **@wiznews**, **@tonistankovic**, **@rihaantonela**, **@bodljikava**, **@bijanaluki**, **@antonmilan**, **@srdjandrango**, **@slobageorgiev**, **@clandes969**, **@ksenija76o**, **@branjislav1978**, **@dragisha74**, **@macokan1**, **@opet8ja**, **@hedi011**, **@onlinedanas**, **@megacomfilm**, **@blic\_online**, **@arhipelagb**, **@seecult**, **@bookvarmagazin**, **@rs\_sputnik**, **@novarsonline**, **@veceivan**, **@n1infobg**, **@gajiczeljkoyah1**, **@branjislav\_dj**, **@srdjandrango**, **@antonmilan**, **@slobageorgiev**, **@clandes969**, **@ksenija76o**, **@branjislav1978**, **@dragisha74**, **@macokan1**, **@opet8ja**, **@hedi011**, **@onlinedanas**, **@megacomfilm**, **@blic\_online**, **@arhipelagb**, **@seecult**, **@bookvarmagazin**, **@rs\_sputnik**, **@novarsonline**, **@veceivan**, **@n1infobg**, **@gajiczeljkoyah1**, **@branjislav\_dj**, **@srdjandrango**, **@antonmilan**, **@slobageorgiev**, **@clandes969**, **@ksenija76o**, **@branjislav1978**, **@dragisha74**, **@macokan1**, **@opet8ja**, **@hedi011**, **@onlinedanas**, **@megacomfilm**, **@blic\_online**, **@arhipelagb**, **@seecult**, **@bookvarmagazin**, **@rs\_sputnik**, **@novarsonline**, **@veceivan**, **@n1infobg**, **@gajiczeljkoyah1**, **@branjislav\_dj**, **@srdjandrango**, **@antonmilan**, **@slobageorgiev**, **@clandes969**, **@ksenija76o**, **@branjislav1978**, **@dragisha74**, **@macokan1**, **@opet8ja**, **@hedi011**, **@onlinedanas**, **@megacomfilm**, **@blic\_online**, **@arhipelagb**, **@seecult**, **@bookvarmagazin**, **@rs\_sputnik**, **@novarsonline**, **@veceivan**, **@n1infobg**, **@gajiczeljkoyah1**, **@branjislav\_dj**, **@srdjandrango**, **@antonmilan**, **@slobageorgiev**, **@clandes969**, **@ksenija76o**, **@branjislav1978**, **@dragisha74**, **@macokan1**, **@opet8ja**, **@hedi011**, **@onlinedanas**, **@megacomfilm**, **@blic\_online**, **@arhipelagb**, **@seecult**, **@bookvarmagazin**, **@rs\_sputnik**, **@novarsonline**, **@veceivan**, **@n1infobg**, **@gajiczeljkoyah1**, **@branjislav\_dj**, **@srdjandrango**, **@antonmilan**, **@slobageorgiev**, **@clandes969**, **@ksenija76o**, **@branjislav1978**, **@dragisha74**, **@macokan1**, **@opet8ja**, **@hedi011**, **@onlinedanas**, **@megacomfilm**, **@blic\_online**, **@arhipelagb**, **@seecult**, **@bookvarmagazin**, **@rs\_sputnik**, **@novarsonline**, **@veceivan**, **@n1infobg**, **@gajiczeljkoyah1**, **@branjislav\_dj**, **@srdjandrango**, **@antonmilan**, **@slobageorgiev**, **@clandes969**, **@ksenija76o**, **@branjislav1978**, **@dragisha74**, **@macokan1**, **@opet8ja**, **@hedi011**, **@onlinedanas**, **@megacomfilm**, **@blic\_online**, **@arhipelagb**, **@seecult**, **@bookvarmagazin**, **@rs\_sputnik**, **@novarsonline**, **@veceivan**, **@n1infobg**, **@gajiczeljkoyah1**, **@branjislav\_dj**, **@srdjandrango**, **@antonmilan**, **@slobageorgiev**, **@clandes969**, **@ksenija76o**, **@branjislav1978**, **@dragisha74**, **@macokan1**, **@opet8ja**, **@hedi011**, **@onlinedanas**, **@megacomfilm**, **@blic\_online**, **@arhipelagb**, **@seecult**, **@bookvarmagazin**, **@rs\_sputnik**, **@novarsonline**, **@veceivan**, **@n1infobg**, **@gajiczeljkoyah1**, **@branjislav\_dj**, **@srdjandrango**, **@antonmilan**, **@slobageorgiev**, **@clandes969**, **@ksenija76o**, **@branjislav1978**, **@dragisha74**, **@macokan1**, **@opet8ja**, **@hedi011**, **@onlinedanas**, **@megacomfilm**, **@blic\_online**, **@arhipelagb**, **@seecult**, **@bookvarmagazin**, **@rs\_sputnik**, **@novarsonline**, **@veceivan**, **@n1infobg**, **@gajiczeljkoyah1**, **@branjislav\_dj**, **@srdjandrango**, **@antonmilan**, **@slobageorgiev**, **@clandes969**, **@ksenija76o**, **@branjislav1978**, **@dragisha74**, **@macokan1**, **@opet8ja**, **@hedi011**, **@onlinedanas**, **@megacomfilm**, **@blic\_online**, **@arhipelagb**, **@seecult**, **@bookvarmagazin**, **@rs\_sputnik**, **@novarsonline**, **@veceivan**, **@n1infobg**, **@gajiczeljkoyah1**, **@branjislav\_dj**, **@srdjandrango**, **@antonmilan**, **@slobageorgiev**, **@clandes969**, **@ksenija76o**, **@branjislav1978**, **@dragisha74**, **@macokan1**, **@opet8ja**, **@hedi011**, **@onlinedanas**, **@megacomfilm**, **@blic\_online**, **@arhipelagb**, **@seecult**, **@bookvarmagazin**, **@rs\_sputnik**, **@novarsonline**, **@veceivan**, **@n1infobg**, **@gajiczeljkoyah1**, **@branjislav\_dj**, **@srdjandrango**, **@antonmilan**, **@slobageorgiev**, **@clandes969**, **@ksenija76o**, **@branjislav1978**, **@dragisha74**, **@macokan1**, **@opet8ja**, **@hedi011**, **@onlinedanas**, **@megacomfilm**, **@blic\_online**, **@arhipelagb**, **@seecult**, **@bookvarmagazin**, **@rs\_sputnik**, **@novarsonline**, **@veceivan**, **@n1infobg**, **@gajiczeljkoyah1**, **@branjislav\_dj**, **@srdjandrango**, **@antonmilan**, **@slobageorgiev**, **@clandes969**, **@ksenija76o**, **@branjislav1978**, **@dragisha74**, **@macokan1**, **@opet8ja**, **@hedi011**, **@onlinedanas**, **@megacomfilm**, **@blic\_online**, **@arhipelagb**, **@seecult**, **@bookvarmagazin**, **@rs\_sputnik**, **@novarsonline**, **@veceivan**, **@n1infobg**, **@gajiczeljkoyah1**, **@branjislav\_dj**, **@srdjandrango**, **@antonmilan**, **@slobageorgiev**, **@clandes969**, **@ksenija76o**, **@branjislav1978**, **@dragisha74**, **@macokan1**, **@opet8ja**, **@hedi011**, **@onlinedanas**, **@megacomfilm**, **@blic\_online**, **@arhipelagb**, **@seecult**, **@bookvarmagazin**, **@rs\_sputnik**, **@novarsonline**, **@veceivan**, **@n1infobg**, **@gajiczeljkoyah1**, **@branjislav\_dj**, **@srdjandrango**, **@antonmilan**, **@slobageorgiev**, **@clandes969**, **@ksenija76o**, **@branjislav1978**, **@dragisha74**, **@macokan1**, **@opet8ja**, **@hedi011**, **@onlinedanas**, **@megacomfilm**, **@blic\_online**, **@arhipelagb**, **@seecult**, **@bookvarmagazin**, **@rs\_sputnik**, **@novarsonline**, **@veceivan**, **@n1infobg**, **@gajiczeljkoyah1**, **@branjislav\_dj**, **@srdjandrango**, **@antonmilan**, **@slobageorgiev**, **@clandes969**, **@ksenija76o**, **@branjislav1978**, **@dragisha74**, **@macokan1**, **@opet8ja**, **@hedi011**, **@onlinedanas**, **@megacomfilm**, **@blic\_online**, **@arhipelagb**, **@seecult**, **@bookvarmagazin**, **@rs\_sputnik**, **@novarsonline**, **@veceivan**, **@n1infobg**,

Out of 60,716 analyzed tweets, 6256 (10.3%) contained a hashtag while there were 3396 unique hashtags. Most popular hashtags (apart from #culture which was filtered out) were most likely to reference a region or a city (#novisad, #beograd, #vojvodina, #nis), one of the domestic public media services (#rtsvesti, #rtvojvodine) or literature (#knjiga - #book, #knjizevnost - #literature).

2019		2020		2019		2020	
Keyword	Freq.	Keyword	Freq.	Hashtag	Freq.	Hashtag	Freq.
Srbija	2005	Srbija	1852	#dobrevesti	486	#dobrevesti	326
godina	1764	godina	1737	#beograd	322	#novisad	253
srpski	1517	nov	1588	#knjiga	270	#vojevodina	228
nov	1502	nemati	1500	#novisad	224	#srbija	144
Beograd	1452	srpski	1488	#beograduzivo	220	#rtvojevodine	123
knjiga	1420	čovek	1383	#rtsvesti	208	#Covid19	208
film	1276	drugi	1232	#knjizevnost	183	#knjiga	115
nemati	1269	velik	1201	#rtvojevodine	176	#beograd	114
dom	1249	film	1184	#vojevodina	176	#nis	91
dan	1228	dobar	1053	#srbija	156	#beograduzivo	86

81



#### 4. Topic modelling

In order to compare how the Twitter topics related to culture differed between 2019 and 2020 (when it was expected that the COVID-19 pandemic would be a significant factor in shaping them), the scraped tweets were split by year and a separate topic model was created for both years. The optimal number of topics ( $K$ ) was determined based on perplexity and coherence analysis and was  $K = 9$  for 2019 and  $K = 11$  for 2020. Figure RS\_8 shows a graphic summary of topics from 2019. The topics are shown as word clouds starting from the first in the top left corner, to the last (9<sup>th</sup>) in the bottom right corner (Figure – RS – 7).

*Figure - RS - 7. Topics that dominated Twitter discussions about culture in Serbia (2019)*



**Ministry of Culture:** Topic 1 from 2019 is related to the Serbian Ministry of Culture and Information and the debates surrounding it. It is represented by words and phrases like "culture" ("kultura", 0.16), "broad" ("opšti", 0.013), "minister" ("ministar", 0.013), "not having" ("nemati", 0.012), "Vučić" (last name of the Serbian president, 0.004), "job" ("posao", 0.04), "upbringing" ("vaspitanje", 0.004) and "connection" ("veza", 0.002), reflecting the debates about Minister's competency, nepotism in regards to the allocation of the public funds for culture, as well as employment policy.

**Film and TV:** Topic 2 is related to film, TV and actors and revolves around words and phrases like "film" ("film", 0.044), "photo" ("foto", 0.041), "actor" ("glumac", 0.015), "story" ("priča", 0.012), "show" ("serija", 0.008), "died" ("preminuo", 0.006), "director" ("reditelj", 0.005) and "famous" ("poznat", 0.005). Apart from dealing with the most prominent TV and cinema releases during the year, it also reflects mourning and discussions surrounding the deaths of several prominent domestic actors and directors.

**Urban Cultural Policy:** Dealing primarily with urban cultural policy, Topic 3 from 2019 is represented by words like "city" ("grad", 0.021), "cultural" ("kulturan", 0.021), "ministry" ("ministarstvo", 0.015), "European" ("evropski", 0.012), "monument" ("spomenik", 0.011), "capital" ("prestonica", 0.008), "protection" ("zaštita") and "budget" ("budžet", 0.005) reflecting debates about the values and aesthetics of newly erected monuments, transparency regarding their funding, as well as protection and preservation of the existing sights of cultural significance.

**Cultural Distinction:** Reflecting the frequent use of "lack of culture" in the context of delegitimizing one's opponent in a heated Twitter debate, Topic 4 is related to words and phrases like "culture" ("kultura", 0.081), "not having" ("nemati", 0.024), "home", ("kuća", 0.006), "normal" ("normalan", 0.005), "school" ("škola", 0.004), "fucking" ("jebati", 0.004) and "upbringing" ("vaspitanje", 0.004).

**Cultural Events:** Topic 5 is related to various cultural events like festivals, theatre shows, concerts and exhibitions. It's represented by words and phrases like "Belgrade" ("Beograd", 0.041), "festival" ("festival", 0.028), "theatre" ("pozorište", 0.025), "concert" ("koncert", 0.016), "music" ("muzika", 0.01), "museum" ("muzej", 0.008), audience ("publika", 0.007) and international ("međunarodni", 0.006).

**National Identity:** Topic 6 deals with concepts related to national identity and elements that comprise it like "people" ("narod", 0.023), "history" ("istorija", 0.02), "language" ("jezik", 0.02), "country" ("zemlja", 0.017), "Serb" ("Srb", 0.017), "tradition" ("tradicija", 0.012), "national" ("nacionalan", 0.006) and "faith" ("vera", 0.006), but also includes words like "other" ("drugi", 0.02) and "Croatian" ("hrvatski", 0.007).

**Education:** Topic 7 is centered on education; its importance and the role politics play in it. It is represented by words and phrases like "education" ("obrazovanje", 0.014), "right" ("prav", 0.013), "society" ("društvo", 0.012), "government" ("vlast", 0.01), "law" ("zakon", 0.009), "politics" ("politika", 0.009), "important" ("važan", 0.007) and "system" ("sistem", 0.005).

**Territory:** Centered around words and phrases like "Serbia" ("Srbija", 0.045), "great" ("velik", 0.041), "world" ("svet", 0.019), "first" ("prvi", 0.015), "Kosovo" ("Kosovo", 0.009), "Russian" ("ruski", 0.008), "church" ("crkva", 0.006), "remembrance" ("sećanje", 0.005) and "NATO" ("NATO", 0.003), topic 8 is similar to topic 6 (national identity) but focuses more on questions like Serbian territory, status of Kosovo and the international relations surrounding it.

**Literature:** Topic 9 deals with literature, newest literary works and their authors, as well as awards and recognition they received. It is represented by words and phrases like "book" ("knjiga", 0.049), "art" ("umetnost", 0.032), "award" ("nagrada", 0.024), "literature" ("književnost", 0.015), "library" ("biblioteka", 0.01), "novel" ("roman", 0.01), "author" ("autor", 0.007) and "publish" ("objaviti", 0.004).

Fout! Ongeldige bladwijzerverwijzing. shows the distribution of tweets related to the aforementioned topics throughout 2019. In general, interest in most of the topics peaks twice during the year, first in March and then less intensively in October, while it is lowest during the summer months, particularly in July. Seasonal variation in interest for a particular topic is intuitively strongest with the topic of **Cultural Events** and weakest with the topic of **National Identity** which remains relatively equally distributed throughout the year.

**Topic**

- Ministry of Culture
- Film & TV
- Urban Cultural Policy
- Cultural Distinction
- Cultural Events
- National Identity
- Education
- Territory
- Literature



("novac", 0.009) and "million" ("milion", 0.007) accompanied as well. Furthermore, words like "healthcare" ("zdravstvo", 0.007), "means" ("sredstvo", 0.006) and "needed" ("potreban", 0.005) also appeared.

**Societal Dialog:** Topic 2 from 2020 shared some of its content with the Topic 4 (**Cultural Distinction**) from 2019 including words like "culture" ("kultura", 0.116), "not having" ("nemati", 0.026) and "upbringing" ("vaspitanje", 0.008), but focused on societal dialog and the culture of civilized exchange of opinions with words like "dialog" ("dijalog", 0.007), "respect" ("poštovanje", 0.006), "opinion" ("mišljenje") and "behavior" ("ponašanje", 0.004).

**Urban Cultural Policy:** Topic 3 from 2019 remained relevant in 2020 as well with debates surrounding the aesthetics and funding of the newly erected monument of Stefan Nemanja (a medieval Serbian ruler) intensifying: "city" ("grad", 0.027), "monument" ("spomenik", 0.025), "Belgrade" ("Beograd", 0.015), "Stefan" (0.006), "Nemanja" (0.004), "heritage" ("nasleđe", 0.004).

**Cultural Events (COVID-19):** Topic **Cultural Events** was dominant in 2020 as well with the addition of the COVID-19 pandemic into the debates within it, so apart from aforementioned words from 2019, words like "corona" ("korona", 0.012), "virus" ("0.008"), "help" ("pomoć") were also present.

**National Identity:** Topic 6 from 2010 (**National Identity**) appeared as relevant in 2020 with slightly different word distributions and with the addition of the words that constituted the topic **Territory** a year before.

**Film and Theatre:** In 2020, film was discussed alongside theatre and mainly from the lens of related events and locations like "festival" ("festival", 0.029), "show" ("predstava", 0.022), "cinema" ("bioskop", 0.011), "premiere" ("premijera", 0.006) and "Fest" (a domestic film festival, 0.006).

**Literature:** With slightly different word distributions, the topic of Literature remained as relevant as it was in 2019.

**Illegitimate culture:** Represented by words and phrases like "non-culture" ("nekultura", 0.02), "actor" ("glumac", 0.02), "trash" ("smeće", 0.019), "poison" ("trovati", 0.018), "generation" ("naraštaj", 0.017), "famous" ("poznat", 0.011), "YouTube" (0.011) and "Kurir" (a pro-government tabloid, 0.01), Topic 8 from 2019 revolved around forms of culture perceived as illegitimate and the consequences of their consumption by the public with the focus on the youth.

**Education:** With a slightly different word distribution, just like in 2019, topic **Education** dealt with the importance of education and educational policy, as well as with the need for reforming the domestic educational system ("development", "razvoj", 0.006; "change", "promeniti", 0.005; "system", "sistem", 0.005).

**Emigration:** Sharing a lot of content with the **Cultural Distinction** topic from 2019 ("culture", 0.077; "not having", 0.013; "fucking", 0.004; "normal", 0.004), Topic 10 from 2020 dealt with the perceived "lack of culture" in the country in the context of reasons for needing to emigrate and start one's life elsewhere represented by words like "love" ("voleti", 0.009), "leave" ("otići", 0.008), "life" (život, 0.008), "stay" ("ostati", 0.008), "family" ("porodica", 0.006), "found" ("osnovati", 0.003) and "happiness" ("sreća", 0.003).

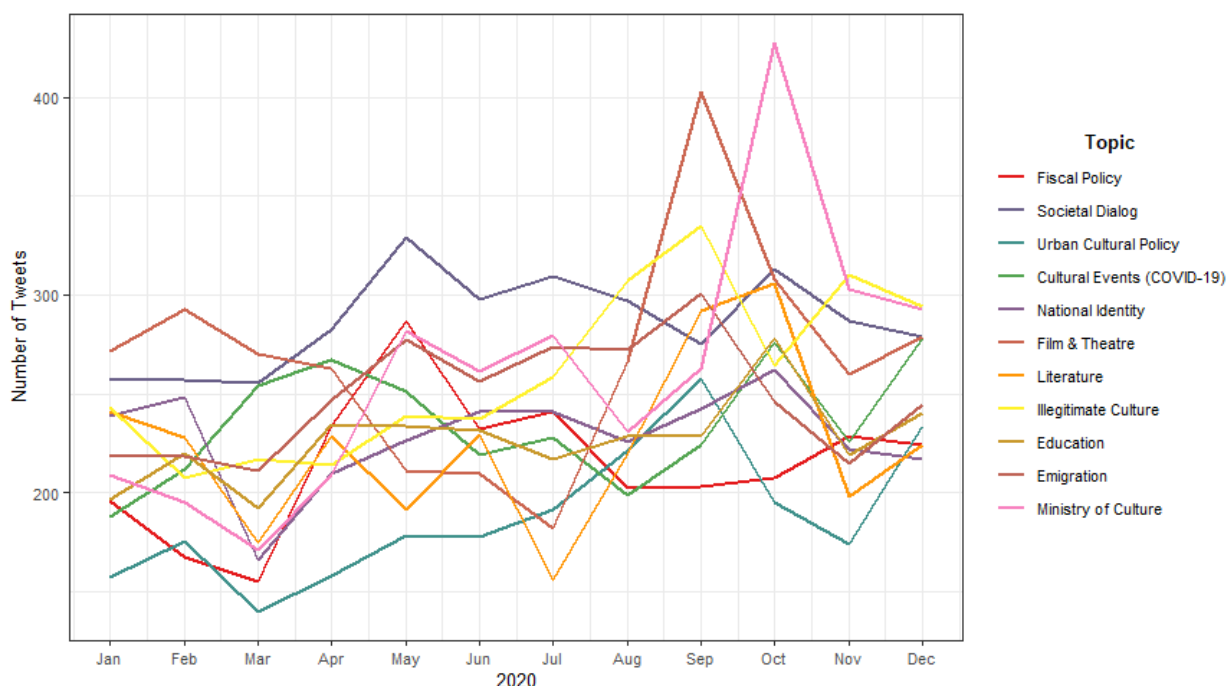
**Ministry of culture:** Change of the person acting as the Minister of Culture and Information didn't make this topic irrelevant in 2020. However instead of the debates about competency and employment policy, the focus shifted to the fight for the freedom of the press ("law", "zakon", 0.01; "fight", "borba", 0.007; "medium", "medij", 0.006; "freedom", "sloboda", 0.005).

**Figure RS – 10** shows an overall much more balanced distribution of interest in the recognized topics in 2020 in comparison to 2019. For most of the topics, seasonal variation seems to be almost gone



which may be explained by the absence of the usual summer holiday season due to COVID-19 related lockdowns and travel restrictions. Only topics with significant peaks are **Film and Theatre** and **Illegitimate Culture** which peak together in September, **Societal Dialog** which peaks in May (just before the parliamentary elections in June) and **Ministry of Culture** which peaks in October.

**Figure - RS - 10.** Distribution of tweets related to each of the topics (2020)



### Comparison of 2019 and 2020 topics

Most of the topics found in the 2019 model remained relevant in 2020 as well (**Ministry of Culture**, **Urban Cultural Policy**, **Cultural Events**, **National Identity**, **Education** and **Literature**) with some of them incorporating new elements due to the ongoing Covid-19 pandemic. This is particularly visible with the topic **Cultural Events** which gained words like "corona", "virus" and "help", as well as with **Fiscal Policy** which shared most of its content with **Education** (2019) but also incorporated words like "healthcare", "means" and "needed" reflecting the debates about the allocation of public funds in the context of the economic crisis created by the pandemic.

On the other hand, some of the topics from 2019 split into subtopics with different focuses. The topic **Cultural Distinction** split into subtopics that focused on a) the consequences of "lack of culture" in the country on the possibility of fostering a functional **Societal Dialog**, and b) on the dilemma whether to stay in the country or emigrate (**Emigration**) viewing other countries as "more cultured" and therefore more suitable for life and happiness.

The topic **Film and Actors** from 2019 dropped its focus on actors and incorporated theatre instead, while the **Territory** topic from 2019 was incorporated into the overarching **National Identity** topic in 2020. On the other hand, actors were incorporated into the new topic **Illegitimate Culture** that dealt with forms of culture considered illegitimate while focusing on the influence of these forms on the youth ("poison", "trash", "generation"). Actors and YouTube celebrities were particularly singled out in this context.

## 5. Discussion and conclusion

The analysis has shown that per year on average, Twitter users in Serbia publish around 30,000 original tweets mentioning culture, with most of them being written in Serbian and coming from one of the largest cities like Belgrade, Niš, Novi Sad or Kragujevac. Not many of those users have verified profiles and those that do are most likely to belong to groups like politicians, political parties, media outlets, prominent actors and performers, as well as international organizations and foreign embassies. Hashtags are rarely used, with only 1 in 10 tweets containing them, while the most popular hashtags usually refer to cities, regions, public media organizations and literature. Users that dominate Twitter discussions about culture in Serbia are usually media outlets and politicians.

As far as the content found in the discussions within the analyzed tweets is concerned, the conducted LDA analysis singled out topics like **Ministry of Culture**, **Urban Cultural Policy**, **Cultural Events**, **National Identity**, **Education** and **Literature** as consistently relevant throughout the observed period (2019 and 2020) while their focus varied slightly between the years. One of the most prominent causes of this variation was the Covid-19 pandemic which, intuitively, most seriously influenced the topic of **Cultural Events** while overall shifting the attention of other topics towards questions about public funds allocation in the context of the related economic crisis.

The question of the unresolved status of Kosovo appeared consistently throughout the observed period having its own topic in 2019 (**Territory**), while being incorporated into the broader topic of **National Identity** in 2020.

Another interesting finding is the widespread use of culture, or more precisely, the lack of it as a strategy for delegitimizing one's opponent in a discussion (**Cultural Distinction**, 2019), explaining the economic and political hardships in the country (**Societal Dialog**, 2020) and providing a reason for one's decision to emigrate (**Emigration**, 2020).

The question of legitimacy of newer forms of media content (**Illegitimate Culture**, 2020) was also present with the focus on content created by YouTube influencers and its perceived negative influence on country's youth.

## Findings for Spain

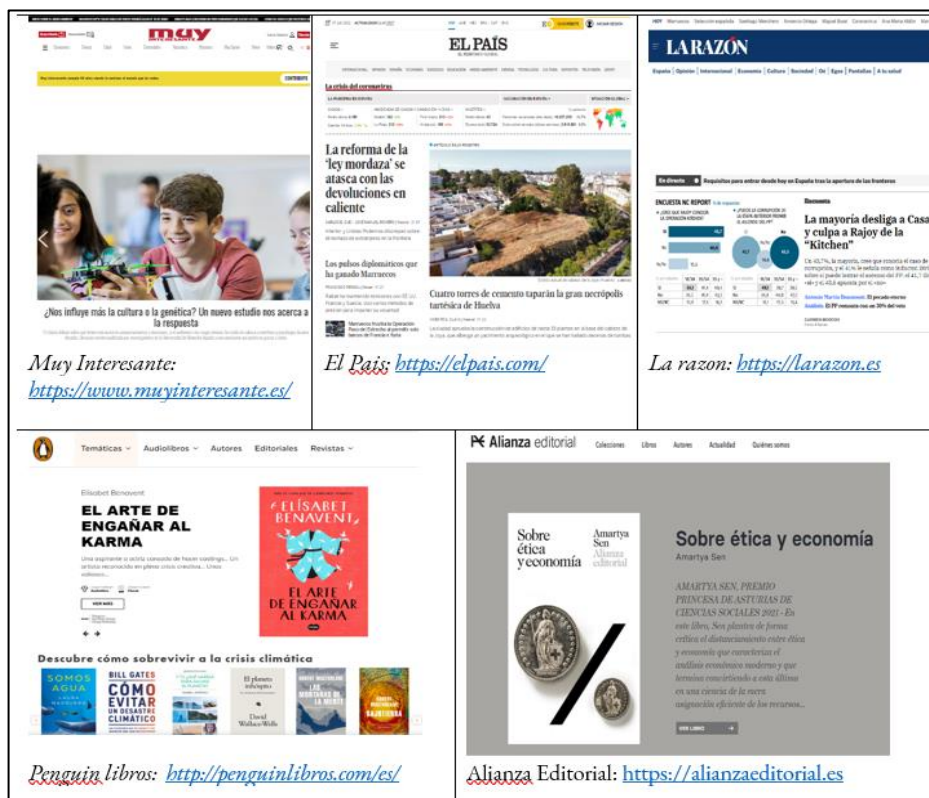
### 1. Twitter usage in Spain

According to the 'Digital 2021 report' (Kemp 2021), 8.1 million Spaniards have become active users on social networks. Users of social networks in Spain are between 16 to 64 years old. Twitter is the fourth most used social network (52.6 %). After a year of pandemic, all platforms have increased their penetration and their time of use, a 27.6% increase. Today, 80% of the Spanish population is on social networks. However, there is a significant contrast in usage distribution by the location, 80% of users are in urban areas. In 2021, Twitter reports that the potential audience reached with an advert on Twitter is approximately 7.5 million people in Spain. Over 62.8% of Twitter users in Spain were men.

### 2. Data description

Using Twitter academic account API, we scraped keywords as 'cultura' and 'culture' with place country (ES) from 2019 to 2020. After removing tweets created by tweet bots, 83,645 tweets were collected with profile locations in Spain. A total of 35 various languages were collected. The main language was Spanish (83% , 65,357) and Catalan (14%, 11,301) and English (3%, 2,250). Other languages show less than 1%. Therefore, we selected the final data for analysis with three languages (Spanish, Catalan, and English), which total 78,908 tweets. Especially we decided to include English tweets since it added meaningful insight later in topic modelling analysis especially relates to the metropolitan cities in Spain. To compare tweets before and during the Covid-19 pandemic, we have divided these two periods according to the date of the first Covid news in Spain, 22nd January 2020 (Reuters Institute 2020; Trilla 2020). The number of tweets before Covid (hereafter 2019) is 44,124 and after Covid arrival (2020) the number of tweets is 34,784.

**Figure - ES - 1. The user with the highest number of followers**



**Muy Interesante:** <https://www.muyinteresante.es/>

**El País:** <https://elpais.com/>

**La razon:** <https://larazon.es>

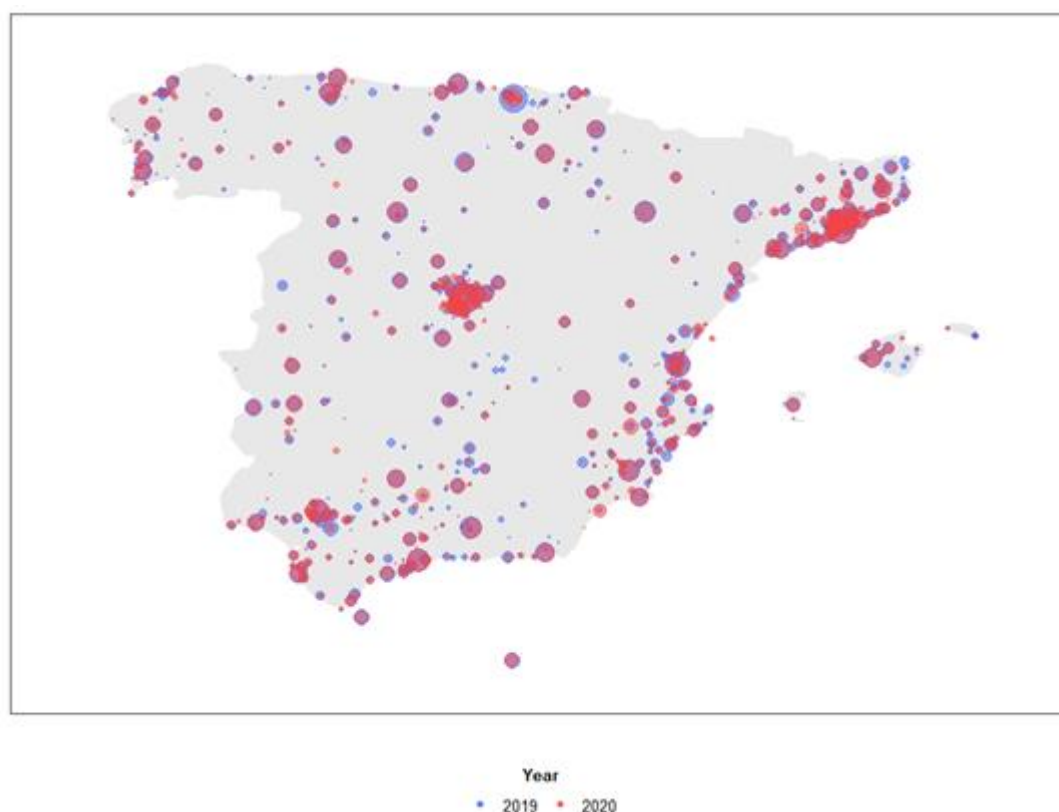
**Penguin Libros:** <http://penguinlibros.com/es/>

**Alianza editorial:** <https://alianzaeditorial.es>

There are 23,836 unique users, and the average number of tweets per user is 3.6 tweets. The unique user in 2019 is 14,783 and 13,103 in 2020 which means about 4000 users are repeated in two years. We cannot find notable differences between active users between 2019 and 2020. In 2019, the users who had the highest number of followers were all verified public users, such as magazines ([Muy Interesante](#)), journals ([El Pais](#), [La Razon](#)), celebrities ([Sir Cheeto](#)), museums ([Museo Thyssen](#)), book publisher ([Penguin Libros](#)), or the government departments ([Ministerio de Sanidad](#)). These users also appear as the most followed users in 2020, including the magazines and broadcasting companies (Muy Interesante, Europa FM europafm.com, La Razon larazon.es), celebrities (Cristina Pedroche, Goyo Jiménez), museums ([Museo Thyssen](#)), or book publishers ([Alianza Editorial](#)). Although the most active user in 2019 is the journal ([Deia](#)) that posted 2,924 tweets alone (6.7 % of total tweets in 2019), the other most active users are mainly academics, authors, librarians; 11% of tweets in 2019 were posted by these top 10 users. Similarly, in 2020, the most active users are also primarily academics and other individual users, who together account for 3% of tweets in 2020. Around 51 % of total users mention other users. On average, per tweet, users name 2 or 3 users.

After preprocessing data, all tweets contain location information based on their profile location. In total, 80% of total tweets have geographical coordinates information. There is information about 1,794 unique locations: 20% of tweets are from Madrid (13,101 tweets) and Barcelona (8,927 tweets), and between 3 and 4% of tweets are posted from other big cities such as Bilbao, Valencia, Sevilla. **Figure - ES - 2** shows the distribution of the tweets in Spain. The density of points indicates the higher number of tweets posted from the region.

**Figure - ES - 2.** Tweets distribution map in Spain before and during COVID



### 3. Hashtags and keywords analysis

Around 37% (29,404 tweets) of total tweets use hashtags. In total, 106,034 hashtags are tagged, and the number of unique hashtags is 29,731. On average, each tweet contains 3 or 4 hashtags (3.6 on

average). **Table - ES - 1** presents the most popular hashtags and keywords except ‘cultura (#cultura)’ in 2019 and 2020. Most hashtags and keywords are shared between 2019 and 2020. Especially the keywords are mostly shared except **book (libro)** in 2019 or **life (vida)** in 2020 (see

**Table - ES - 1**). The most popular keywords in Spain are ‘país (country)’, ‘casa (house)’, and ‘gracias (thanks)’. The most popular hashtags in 2019 are **#gente (people)** and **#ocio (leisure)**. These two hashtags are used 3.8% of total tweets, and 3% of total tweets use both hashtags side by side. While in 2020, there are new hashtags regarding Covid, such as **#safeculture (#culturasegura)** and **#Covid19**.

**Table - ES - 1.** Hashtags for Twitter (N= 106,034) and Keywords for Twitter (N = 58,286)

Hashtags				Keywords			
2019	Freq	2020	Freq	2019	Freq	2020	Freq
#gente (people)	2748	#culturasegura (safe culture)	660	país (country)	2041	país (country)	1247
#ocio (leisure)	2453	#musica (music)	412	casa (home)	1707	gracias (thanks)	1178
#arte (art)	598	#arte (art)	318	gracias (thanks)	1351	cultural	1070
#teatro (theatre)	474	#teatro (theatre)	235	mundo (world)	1341	casa (home)	1036
#musica (music)	445	#Covid19	212	historia (history)	1317	historia (history)	964
#curiosidades (curiosity)	423	#madrid	197	español (spanish)	1248	españa (Spain)	964
#madrid	346	#alertaroja	188	españa (Spain)	1222	español (spanish)	965
#turismo (tourism)	346	#espana (Spain)	164	arte (art)	1178	arte (art)	946
#art	291	#historia (history)	156	cultural (cultural)	1068	mundo (worlds)	918
#historia (history)	287	#gastronomia (gastronomy)	152	libro (book)	1017	vida (life)	895

**Table - ES - 1** presents the most frequently used words, while the semantic network in **Figure - ES - 3** describes how these keywords are related to each other. **Figure - ES - 3** compares hashtags between 2019 and 2020. As you can see, beforeCovid, the hashtag ‘#leisure (#ocio)’ appears as an essential hashtag. While in 2020, a few regional names appear, such as Coruña in Galicia and Andalucia, together with hashtags related to Covid, such as #corona, #Covid19, #safeculture, #we can do event (#hacemoseventos). Apart from this difference, we find also similar hashtags such as book (#libro), lecture (#lectura), truism (#turismo), etc. Keyword networks in 2019 and 2020 are



**Figure - ES - 3.** Top hashtags and keywords network of tweets in Spain before and during Covid



Based on perplexity and coherence analysis, the optimal topic number  $K$  for each year is 10 in 2019 and 12 in 2020. To define the best topic number  $K$ , we estimate several models, from  $K = 10$  to  $K = 14$ . It appears that  $K = 10$  represents the best results in terms of topic variety. Next, we present the result of topic modelling with topic number  $K$  10 in each period.







**Legitimate culture.** Topic 6 seems to illustrate **legitimate culture** such as ‘art (arte, 0.046)’, ‘museum (musio, 0.023)’, ‘picture (pintura, 0.006)’, photo (fotografía, 0.005), image (imagen, 0.009)’ and ‘exposition (exposición, 0.021)’. This legitimate culture also includes ‘architecture (arquitectura, 0.045)’ and ‘historical (historico)’ monuments.



**Nationality and identity.** Topic 7 displays **nationality and identity** with words like ‘country (país, 0.073)’, ‘Spain (españa, 0.043)’, ‘spanish (español, 0.044)’ and ‘history (historia, 0.036)’. This topic also shows words like ‘European (europeo, 0.008)’ and ‘Europe (europa, 0.006)’ together with ‘nation (nación, 0.003)’ and ‘civilization (civilización, 0.002)’.



**Social divisions.** Topic 8 concerns **social divisions**. The top words in this topic are ‘person (persona, 0.024)’ and ‘people (gente, 0.022)’. This topic depicts divisions in gender: women (mujer, 0.019/ feminismo, 0.003) versus men (hombre, 0.012/ machista 0.003), in politics (política, 0.011): right (derecho 0.012) versus left (izquierda, 0.007), in classes: rich (rico, 0.002) versus poor (pobre, 0.003). This topic also notes about bullfight (tauromaquia, 0.002), a traditional cultural event in Spain causing strong debates about animal welfare. We can also find words like liberty (libertad, 0.009) and respect (respeto, 0.009) with ignorance (ignorancia, 0.003).



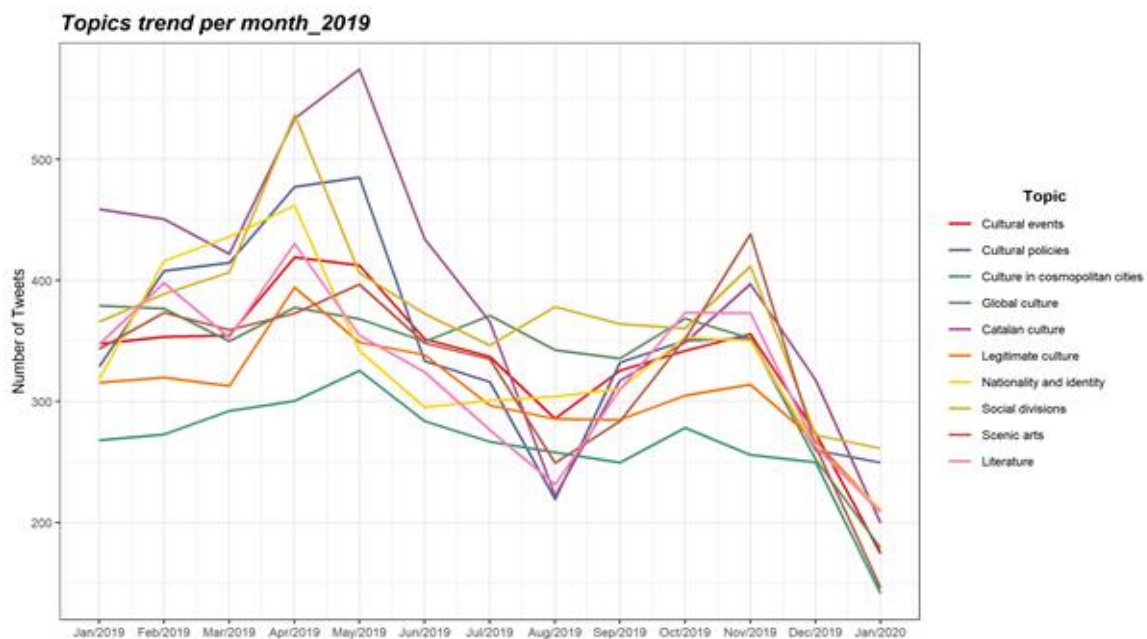
**Scenic arts.** Topic 9 is interpreted as **scenic arts**, recorded and live concerts, and is correlated with music (música, 0.03), with the word ‘casa (home, 0.06)’ which has the highest probability in this topic. This topic indicates the ‘musical (0.014)’, ‘festival (0.022)’ and ‘concert (concierto, 0.020)’ together with the words ‘night (noche, 0.014)’ and ‘theater (teatro, 0.026)’. It illustrates various genres of music such as ‘classic (clásico, 0.008)’, ‘disco (0.006)’, ‘jazz (0.002)’, ‘flamenco (0.005)’ and ‘opera (0.002)’.



**Literature.** Topic 10 is about **literature**, with the word ‘book (libro, 0.04)’ and ‘prize (premio, 0.026)’. This topic seems to include the pride of the national (nacional, 0.012) author (autor, 0.009) and his books.

**Figure - ES - 4** shows the monthly trend of 10 topics. Overall, the hottest topics are **Catalan culture** and **Social divisions** in 2019. Especially there are many comments on these two topics between January and May 2019. However, these topics are gradually decreased by August. From September to November, the topic regarding **Scenic arts** evolves rapidly. The tweets in culture in cosmopolitan cities constantly remain as the not-so-popular topic.

**Figure - ES - 4. Topics trend per month before COVID 19**



## Topic modelling results 2020



**Catalan culture.** Topic 1 is related to **Catalan culture**, with high probability words such as ‘Catalan (català, 0.014)’, ‘Catalunya (0.004)’, ‘Tradicional (0.001)’, and ‘history (història, 0.003)’. In 2019, Topic 5 was very similar to this topic.



**Cultural policies related to the pandemic.** Topic 2 is correlated with **cultural policies**, with words like ‘sector (0.003)’, ‘public (público, 0.020)’, and ‘government (gobierno, 0.014)’ like Topic 2 in 2019. However, this time we can see the words like ‘pandemic (pandemia, 0.010)’ and ‘coronavirus (0.007)’ with high probabilities. Looking at the topic it seems like people are asking for help (apoyo, 0.010), subsidy (subvención, 0.007), and mediation (medida, 0.013) from the government for cultural industry (industria, 0.005) and business (empresa, 0.010)’. The cultural sectors mentioned in this topic are ‘turism (turismo, 0.010)’ and ‘cinema (cine, 0.008)’.



**Legitimate culture.** Topic 7 is understood as **legitimate culture** and appears to be similar to Topic 6 in 2019. This topic includes ‘art (arte, 0.045)’, ‘exposition (exposición, 0.014)’ , ‘artista (artist, 0.023)’ but also ‘history (historia, 0.030)’, ‘city (ciudad, 0.026)’ and ‘heritage (patrimonio, 0.016)’ with city like ‘Madrid (0.013)’ , ‘Barcelona(0.003)’ , ‘León (0.005)’ , ‘Vaencia(0.003)’ and ‘ Galicia(0.003)’.





**Education.** Topic 8 is examined as **education**, is a new topic discussed during 2020. Due to the virus (0.005), the education (educación, 0.032) and educational system (sistema, 0.005) have been forced to change (cambio, 0.006), using digital (0.010) devices (equipo, 0.005). This new way (formación, 0.005) of teaching in class (clase, 0.008) and learning (estudio, 0.006) is described as an opportunity (oportunidad, 0.006).



**Scenic arts.** Topic 9 is analysed as **Scenic arts**, similar to Topic 9 in 2019, correlated with words like 'music (música, 0.030)', 'theatre (teatro, 0.027)', and 'concert (concierto, 0.022)'. However, in 2020 due to Covid, we can find the words like 'security (seguridad, 0.009)' and 'cancelation (cancelación, 0.006)' in this topic. Since this pandemic requires keeping a social distance (distancia, 0.004), it seems that YouTube (YouTuber, 0.004) overtakes the place of those life events.



**Audio-visual culture.** Topic 10 is interpreted as **audio-visual culture** and is correlated with words like 'musical (0.023)', 'image (imagen, 0.016)', 'video (vídeo, 0.010)', and 'audio-visual (0.006)'. It is interesting to note that we can find English words mingled in this topic like 'popular (0.008)', 'idea (0.010)', and 'Spain (0.005)'. Also, this topic includes the words like 'internet (0.004)', 'Twitter (0.002)', 'tweet (0.003)', and 'link (0.004)'.

Figure - ES - 5. Topics trend per month during COVID 19

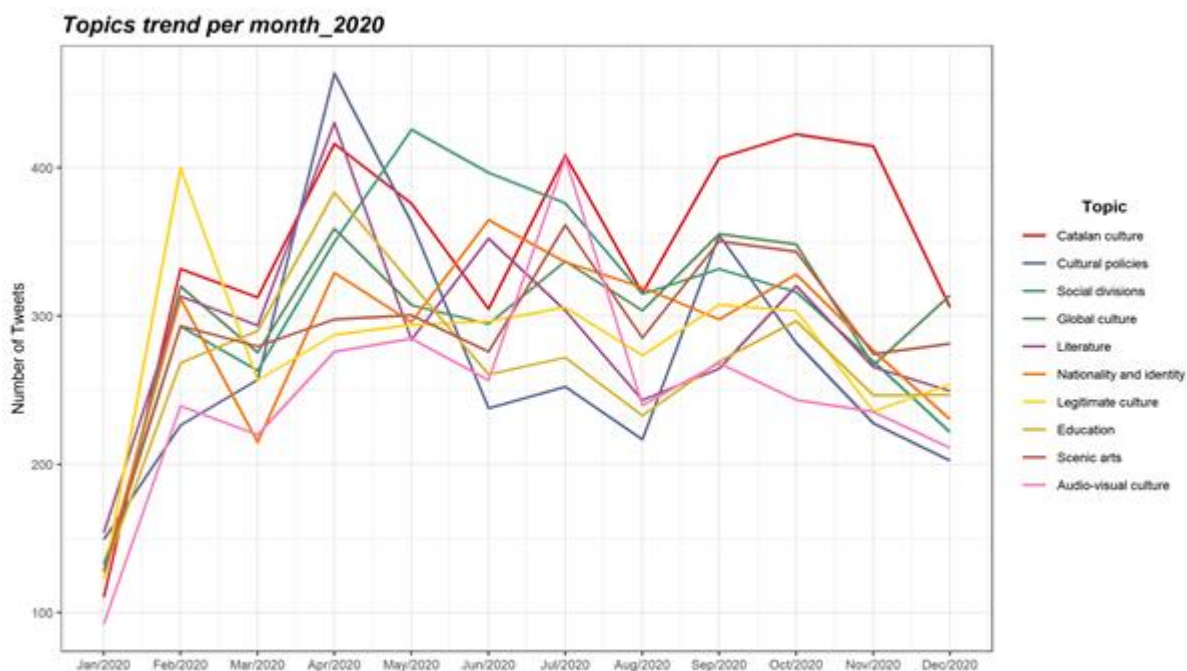


Figure - ES - 5 shows that **Legitimate culture** is the hottest topic in February. However, it shows the sudden fall and remains constant in the following months. Cultural policies topic is the most crucial in April, but its proportion declines during the summer period and jumps up again in September. The topic of **Social division** arises as to the hottest topic in May and remains one of the important

topics after all. The topic of **Catalan culture** constantly remains one of the hottest topics in all 2020. Especially from August to December it continues to be the leading topic.

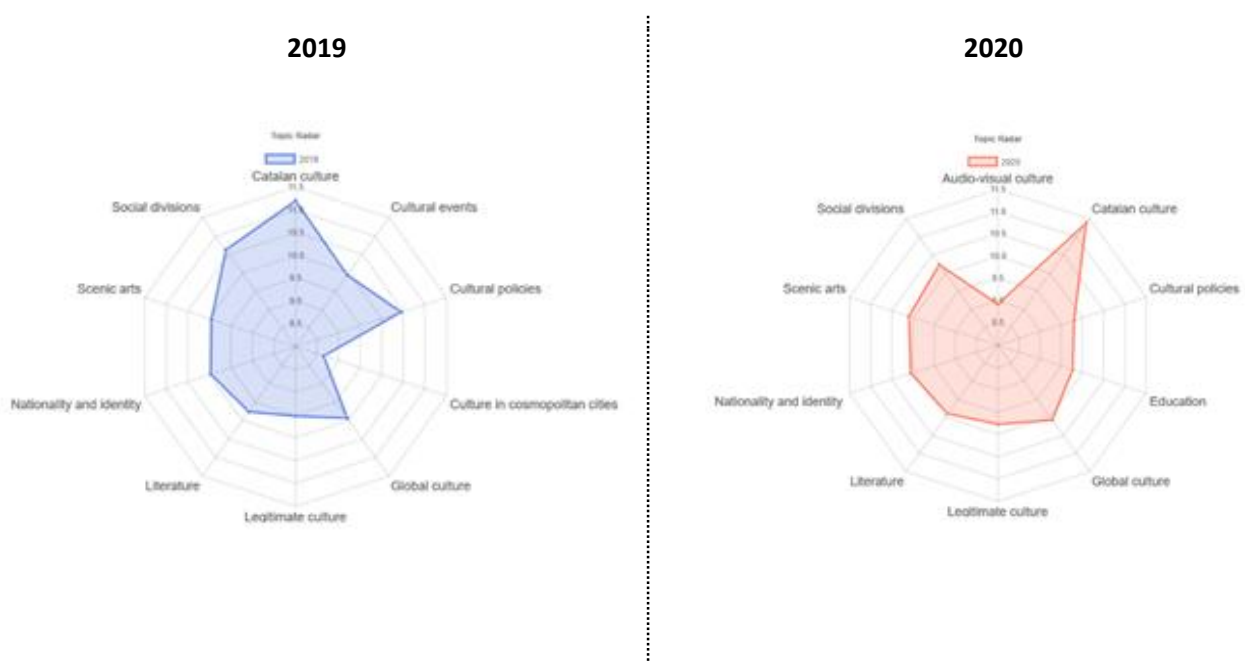
## 5. Discussion and conclusion

The topics discussed before, and during, the pandemic lockdown are similar, but adding some topics related to the Covid, and even some topics have changed their contents during the pandemic period.

**Figure - ES - 6** depicts the distribution of topics for each year.

The topics **Catalan culture**, **Nationality and identity**, **Legitimate culture**, **Cultural policies**, **Social division**, **Global culture**, **Scenic arts**, and **Literature** are found in both years. Shared topics on **Catalan culture**, **Nationality and identity**, **Scenic arts** have similar words with similar probabilities. **Catalan culture** appears as the most important topic of culture in Spain in both 2019 to 2020. In both years the words included in this topic are similar: Catalunya, Catalan, history, identity, language, etc. This result suggests that the local culture, especially Catalan culture, noticeably stands out from Spanish culture. The topic on **Nationality and identity** is correlated with words like Spain, Spanish, country, European and so. This topic suggests that people talking about culture on Twitter express their nationalities and identities but also share a European identity. The topic **Legitimate culture** includes various types of arts expressions, from museums and exhibitions to architecture and historical monuments.

**Figure - ES - 6.** 10 topic Distribution in Spain before (2019) and during COVID19 (2020)



Before Covid, when people talked about **Cultural events** it seems that people experience local culture on a short trip on weekends. However, this topic does not appear in 2020. The topic of **Culture in Cosmopolitan cities** appears only in 2019 with mainly English words. However, this topic also disappears in 2020. Instead, the topic on **Audiovisual culture** in 2020 contains English words with Internet forms like Twitter and Links. In 2020, the topic of **Education** appears as a new topic. It



is interesting to note that the topic of education is included as an important discussion of culture on Twitter during the pandemic. The system of education has experienced a critical change due to the pandemic situation. People are talking more about the new ways of teaching and learning with digital tools.

Among shared topics between 2019 and 2020, the topics about **Cultural policies, Social division, Global culture, Scenic arts, and Literature** changed the contents during the pandemic. Particularly the topic of **Cultural policies** has changed to discuss governmental support for the cultural sector and business. On Twitter, people show concern about the tourism and cinema business. In both years there is the topic of **Social Divisions** between gender and social status. In 2020, the topic of social divisions expanded to include generational differences and education levels. However, in both years, the **Social Divisions** topic consists of the words respect and freedom with high probabilities. We interpret it that on Twitter, people talk about growing social differences in Spanish society. The topic of **Global culture** has also changed in 2020. In 2019 **Global culture** is correlated to audio-visual products like TV series and films. While in 2020, the topic of **Global culture** connects with the desire to travel and have in-person experiences. The lockdown situation may cause people to look forward to travelling again and see other cultures in the world. **Scenic arts** in 2019 are mainly about various live music events across many genres, from classic to disco. However, in 2020 this topic includes words related to the music industry crisis during the lockdown situation and shows using YouTube instead of going to live events.

In summary, people on Twitter talk about topics about culture regardless of the pandemic situation, such as **Social divisions, Cultural policies**, and link their cultural discussions with their **Nationality and identity**, particularly the pride of **Catalan culture**. The essential cultural products in Spain are **Legitimate culture, Literature, and Scenic Arts**. However, the contents of these topics have changed according to the pandemic crisis and reflect what matters to the Spaniards. On Twitter, there is a growing concern for supporting the cultural sector, but people also see the pandemic crisis as an opportunity to transform culture into digital and increase the size of their market.

## Findings for Switzerland

### 1. Twitter usage in Switzerland

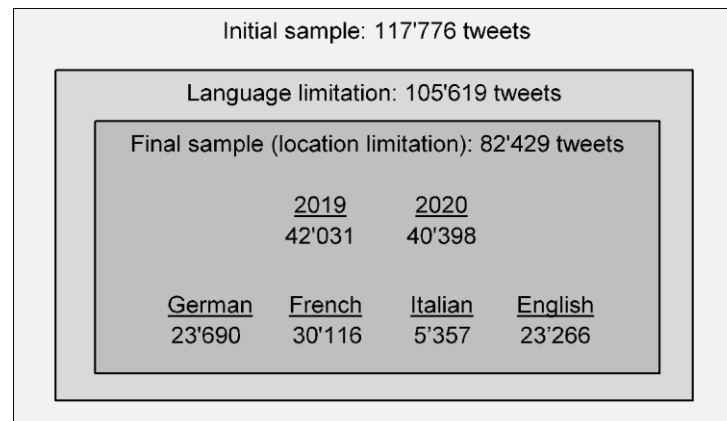
Generally, valid information on the characteristics of users of specific social media platforms is hard to find. Nevertheless, there is some empirical data that allows us to draw at least a rough picture of Twitter users in Switzerland. First, the 2019 survey of the Swiss Federal Statistical Office on Internet usage in Switzerland finds that being part of social online networks and posting texts, pictures, or videos online is particularly prevalent among younger and highly educated individuals (Bfs 2021b). Second, regarding the usage of Twitter more specifically, the 2020 Swiss Yearbook on the Quality of the Media (FÖG 2020) reports that Twitter is only the sixth most popular social media platform for information seeking (6% of the general population). For the latter purpose, Facebook (30%), WhatsApp (26%), YouTube (24%), Instagram (11%), and LinkedIn (6%) are more important. Hence, Twitter appears to be more of a minority medium that is not representative for the general Swiss population.

### 2. Data description

#### Search strategy and number of tweets

The Swiss team used Twitter's full-archive keyword search with a premium account to search for culture-related tweets in 2019 and 2020 (complete years). The "profile" operator was used in the search query to restrict the results to tweets from Swiss Twitter profiles only. Since Switzerland is a multilingual country with German, French, and Italian as the three main languages, the search string included the corresponding keywords for culture simultaneously ("kultur", "culture", "cultura"). Since the French and English word for culture is identical and since there was no additional language restriction in the search query, this search query also covered all possible English tweets with reference to culture. The initial search returned 117'776 tweets related to culture in Switzerland for both years. After deletion of tweets written in other languages apart from German, French, Italian, and English, 105'619 tweets remained in the sample. We decided to keep English language tweets in the sample because Switzerland is a country with a large share of immigrants and expats among the resident population, which makes international discourses typically written in English very relevant. The large share of English language tweets in the final sample (28%, see **Figure - CH - 1**) confirms this. Finally, all tweets were deleted that did not contain sufficient information on their geo-location. The final sample used for all subsequent analyses consists of 82'429 tweets, which is a quite substantial number for a country as small as Switzerland. The distribution of tweets across the two year studied is fairly equal, with 2019 yielding slightly more tweets related to culture than 2020 (also see **Figure - CH - 1**). Although the separation between two time periods aims at tracking changes caused by the occurrence of the Covid19 pandemic, we did not identify the exact date of appearance of the virus or media coverage about it in Switzerland. Rather, we used the calendar years 2019 and 2020 to separate a "pre-Covid" period from a "during-Covid" period.

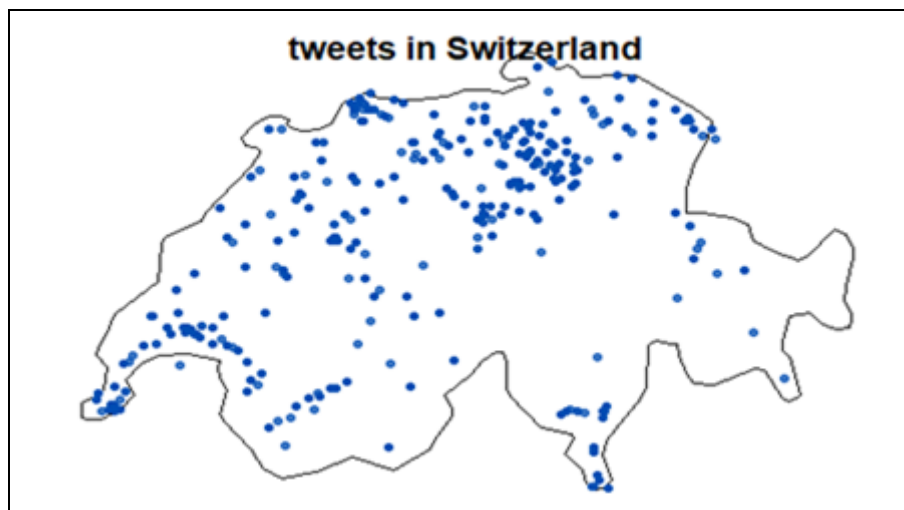
**Figure - CH - 1.** Number of tweets related to culture in Switzerland in 2019 and 2020



### Location of tweets

As can be read from **Figure - CH - 1**, tweets in French language comprise the largest part of the data (30'116), whereas tweets in Italian language are least frequent (5'357). Hence, the French-speaking part of Switzerland is overrepresented in the data, as it accounts for only 23 percent of Switzerland's population (BfS 2021a). On the other hand, tweets in German language (68% of the population, including Swiss German) are clearly underrepresented. With the present data it is not possible to say whether this is because French speakers are generally more active on Twitter or because they tweet more about culture, i.e. use the word "culture" more often.

**Figure - CH - 2.** Location of tweets related to culture in Switzerland in 2019 and 2020



**Figure - CH - 2** reveals that the tweets in the data originate from almost all parts of Switzerland. However, mountainous regions in central and eastern Switzerland are hardly represented. Overall, tweets are clearly concentrated in urban areas such as Zürich, Geneva or Basel (see **Figure - CH - 1** for more details). This is not very surprising given the topic of culture. Culture is often closely related to several types of art, which in turn is most frequently practiced in urban regions. Nevertheless, our data are widespread enough to cover also other possible manifestations of culture.

**Table - CH - 1.** Top 10 locations of tweets related to culture in Switzerland in 2019 and 2020

City	Frequencies
Zürich	12233
Geneva <sup>(a)</sup>	10023
Lausanne	6239
Basel	4940
Bern <sup>(b)</sup>	3981
Ticino	3153
Geneve & Lausanne	1922
Sion	900

(a) Sum of tweets for “geneva” and “geneve”.

(b) Sum of tweets for “bern” and “berne”.

### 3. User Analysis

Overall, the 82'429 tweets in the final sample come from 13'070 unique users. Hence, the average number of tweets per user is 6.3. However, the distribution of tweets across users appears to be heavily skewed. As the left panel of **Table - CH - 2** shows, only a small number of users seems to produce a large part of the tweets about culture. The ten most active users in the data posted 21'827 tweets, representing 26.5 percent of all tweets in the sample. These ten most active users are all of rather institutional type, there are no private persons among them. Most prominently, eight media organizations are tweeting about culture: five of them are from the French speaking part of Switzerland (news\_suisse, LeTemps, RTSinfo, Lematinch, tdgch), two from the German speaking part (srfkultur, Medien\_News), and one from the Italian speaking part (laregione). Somewhat surprisingly, the third most active tweeter about culture is a business organization (swissbusiness) that primarily seeks to connect business people. This already indicates that not only culture in terms of art is relevant in Switzerland, but also something like “business culture” (see analyses in the following sections). The major public funding organization for arts and culture in Switzerland, *ProHelvetia*, is only the tenth most active tweeter about culture, represented by its current director Philippe Bischof (PhBischof; who is also part of the Swiss stakeholder network in the INVENT project).

Comparing the most active users in 2019 and 2020 reveals that the structure of users is fairly stable over time (see the middle and right panel of **Table - CH - 2**). Almost all users that are among the top ten in 2019 are so in 2020, too. However, there is some fluctuation in the concrete rank position and thus in the level of activity of some users. The relative position of swissbusiness, Lematinch, and tdgch is lower in 2020 (also in absolute terms), meaning that these two users are less active in the second year studied. In contrast, laregione, RTSinfo, and PhBischof are more active in 2020 compared to 2019. Furthermore, separating the user analysis by year shows that there is one user among the top ten in each year who is not in that group when inspecting the overall sample. In 2019, EventkalenderCH is the tenth most active user. The fact that this user is not among the top ten in 2020 might be due to the Covid19-related shutdown, as this user’s main business is to publish the dates of cultural and other public events. In 2020, EsatogluR appears as the only private person among the top ten. He is rather new to Twitter (joined in 2018), which may be a reason why he was

less active in 2019. He describes his activity on Twitter as “the personal perspective of a historian working in the media”. Hence, he seems to be close to the media sphere as well. Finally, even though the composition of the top 10 users is rather stable over time, the total number of users has changed. While there are 7'830 users tweeting about culture in 2019, this number increased to 8'578 in 2020 (plus 9.6 percent). This rise could be a reason for why the skewness of the distribution of tweets across users has decreased over time. In 2019, the top ten users account for 30.3 percent of tweets (12'727 out of 42'031), whereas in 2020 they account for only 23.6 percent of tweets (9'514 out of 40'398). Hence, one could conclude that discourses about culture on Twitter are getting more democratized.

**Table - CH - 2.** Top 10 users by number of tweets related to culture in Switzerland in 2019 and 2020

2019+2020		2019		2020	
User	Freq.	User	Freq.	User	Freq.
news_suisse	4187	news_suisse	2377	news_suisse	1810
LeTemps	4063	LeTemps	2261	LeTemps	1802
swissbusiness	2486	swissbusiness	1907	Laregione	1484
laregione	2407	Lematinch	1295	RTSinfo	1279
srfkultur	2016	srfkultur	1091	srfkultur	925
RTSinfo	1904	tdgch	1003	swissbusiness	579
Lematinch	1846	laregione	923	Lematinch	551
Medien_News	1370	Medien_News	898	Medien_News	472
tdgch	1003	RTSinfo	625	EsatogluR	338
PhBischof	545	EventkalenderCH	347	PhBischof	274
<b>Total</b>	<b>21827</b>	<b>Total</b>	<b>12727</b>	<b>Total</b>	<b>9514</b>
% of all tweets	26.5	% of all tweets	30.3	% of all tweets	23.6

#### 4. Hashtag and keyword analysis

In a second step, the text of the tweets is analyzed according to its content. Beforehand, text data had to be cleaned, which involved the following processes: (1) deletion of stopwords, URLs, and emoticons; (2) lemmatization of words; (3) keeping only nouns, pronouns, and adjectives. After the cleaning process, 80'927 tweets are left for analysis. In order to track temporal developments, all subsequent analyses are separate for the year 2019 (41'046 tweets) and the year 2020 (39'841 tweets). Moreover, the words “kultur”, “culture”, and “cultura” are excluded from the analysis.

##### Hashtag and keyword analysis 2019

**Table - CH - 3** shows the most frequently used keywords and hashtags in all tweets related to culture in 2019. In both cases, there are no Italian words among the top 20, which is most likely due to the relatively small share of Italian tweets in the sample. However, there is quite a broad range of topics represented in the data. Focusing on keywords first, reveals that the word “people” is the most

frequently uses word. This is probably because it is a very general word and it belongs to a language that many people can speak (about 28 percent of tweets in the data are in English). Other general and therefore less interesting words among the top 20 keywords are “neu” (new), “suisse” (Swiss), “grand” (great), “schweizer” (Swiss), “time”, “monde” (world), “day”, or “change”. From the remaining more concrete words, most are related to culture in terms of arts (“film”, “festival”, “art”). But there are also references to the business sphere (“company”, “business”), technical developments (“digital”, “video”), France as an important neighboring country of Switzerland (“français”, “france”), and gender issues (“woman”).

**Table - CH - 3.** Top 20 keywords and hashtags in tweets related to culture in Switzerland in 2019

Keywords	Freq.	Hashtags	Freq.
people	913	#schweiz	187
film	912	#rtsculture	165
festival	814	#srf	150
neu	755	#arts	130
suisse	694	#srfkultur	108
company	617	#event	107
grand	556	#innovation	82
schweizer	502	#konzert	80
art	461	#digitaltransformation	70
time	408	#art	68
monde	391	#suisse	65
digital	373	#digital	61
franais	369	#ai	60
france	332	#switzerland	59
business	329	#metoo	57
day	312	#hobbies	56
video	309	#hotelmarketing	56
change	307	#lausanne	53
team	305	#rtsinfo	53
woman	296	#film	50
<b>Total</b>	<b>52478</b>	<b>Total</b>	<b>18729</b>

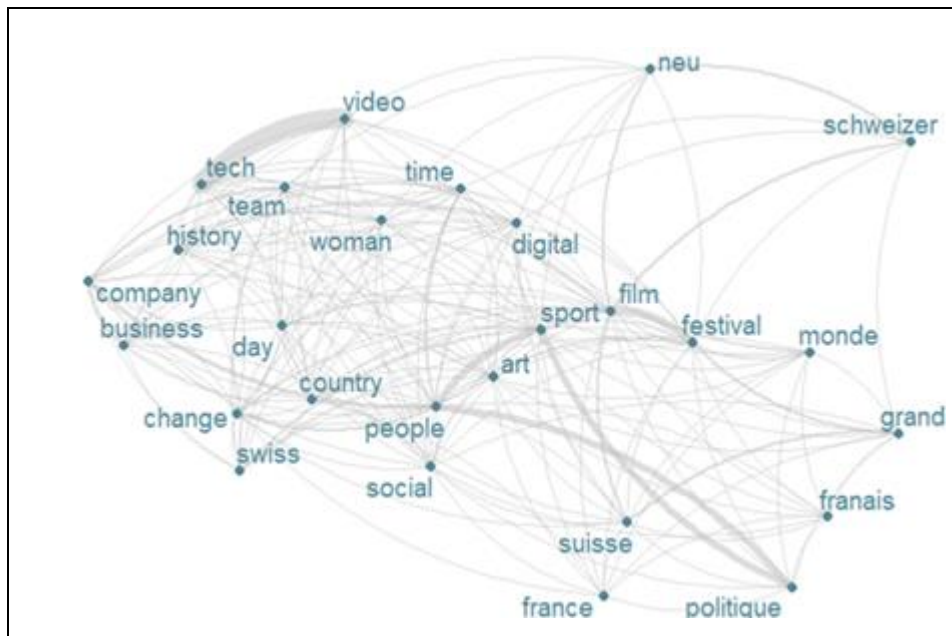
Almost all of these topics can also be identified when looking at the most frequently used hashtags (right panel of **Table - CH - 3**). Most prominently, there are representations for arts (#arts, #event,



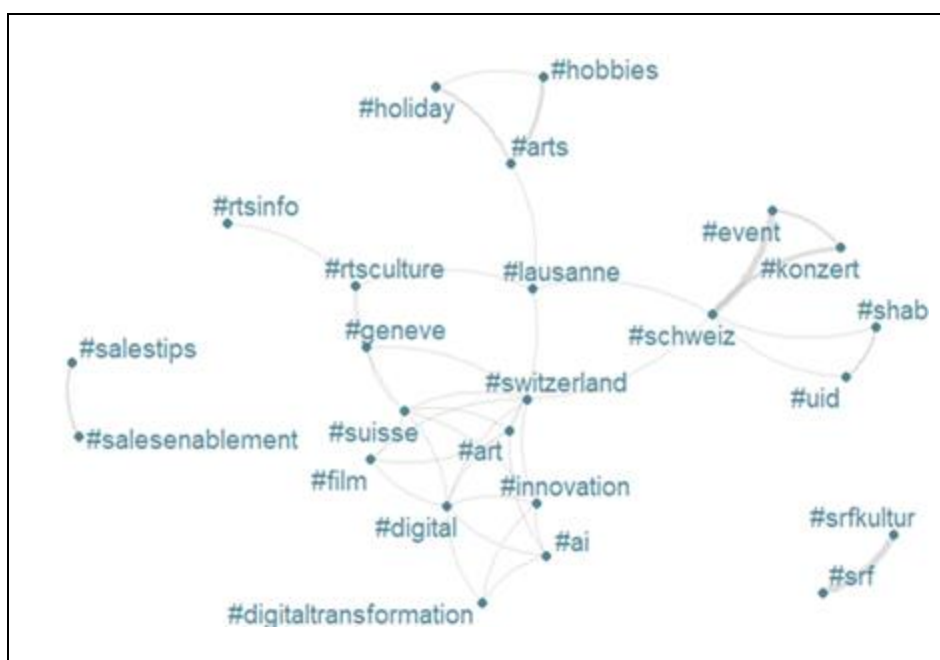
#konzert, #art, #film), but also for technical developments (#digitaltransformation, #digital, #ai), business (#hotelmarketing), and gender issues (#metoo). Moreover, hashtags are frequently used to refer to specific media institutions popular in Switzerland (#rtsculture, #srf, #srfkultur, #rtsinfo). All four hashtags mentioned refer to the Swiss public service broadcasting corporation and their culture-related sub-sections (for French-speaking and German-speaking Switzerland).

In order to better understand the interrelation and embeddedness of keywords and hashtags, **Figure - CH - 3** and **Figure - CH - 4** present the network structure of the top 30 most frequently used keywords and hashtags in 2019.

**Figure - CH - 3.** Network of top 30 keywords of tweets related to culture in Switzerland in 2019



**Figure - CH - 4.** Network of top 30 hashtags of tweets related to culture in Switzerland in 2019



Generally, keywords are much more interrelated than hashtags. This can be explained by the fact that keywords result from natural language production, whereas hashtags only highlight singular expressions and are used much less frequently. The most striking co-occurrences of keywords are “tech” and “video”; “people”, “sport” and “politique”; “film” and “festival”; and “company”, “business” and “people”. As regards hashtags, the most prominent co- occurrences are #event, #schweiz, and #konzert; #srf and #srfkultur; and #arts and #hobbies. This partly overlaps with the topics identified in the previous paragraph and thus indicates a rather broad conception of culture in our data.

### Hashtag and keyword analysis 2020

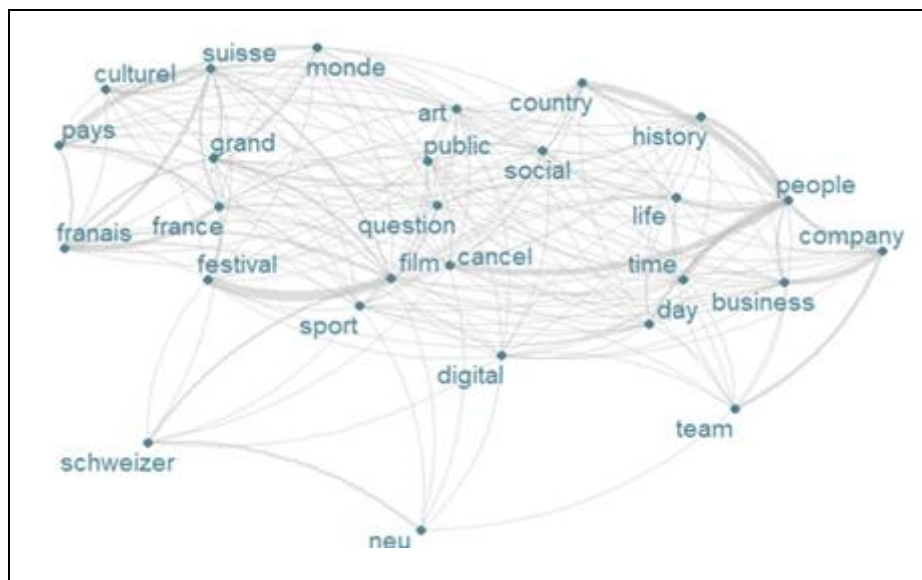
In general, there are great similarities between the most frequently used keywords and hashtags in 2019 and those in 2020 (see **Table - CH - 4**). Among the top 20, we still find keywords and hashtags

**Table - CH - 4.** Top 20 keywords and hashtags in tweets related to culture in Switzerland in 2020.

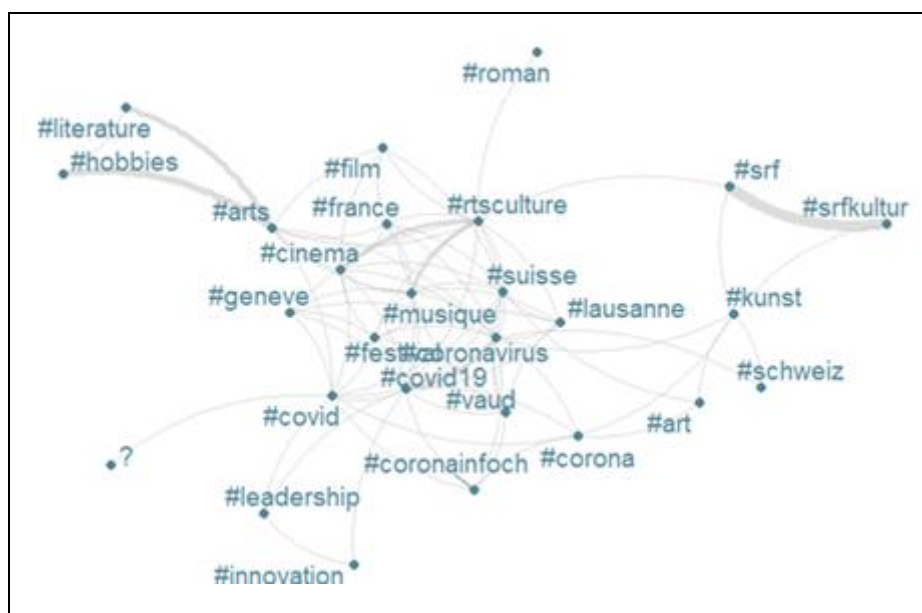
Keywords	Freq.	Hashtags	Freq.
people	1091	#rtsculture	260
suisse	850	#Covid19	215
film	751	#arts	187
cancel	740	#coronavirus	163
neu	686	#srf	145
company	631	#cinema	92
time	586	#corona	91
franais	562	#musique	83
grand	535	#srfkultur	83
monde	494	#art	74
art	489	#suisse	65
festival	466	#lausanne	54
schweizer	423	#hobbies	52
country	411	#schweiz	51
social	405	#vaud	49
sport	391	#rtsinfo	47
france	387	#leadership	44
coronavirus	379	#literature	44
pays	373	#Covid	41
question	370	#innovation	40
<b>Total</b>	<b>54114</b>	<b>Total</b>	<b>15660</b>

referring to topics such as arts, business, and France. However, keywords and hashtags referring to technical developments (esp. digitization) or gender issues (#metoo) are not present in the top 20 any longer. Instead, two new topics appear: First, the occurrence of the Covid19 pandemic in early 2020 is clearly visible in the data. As a keyword (“coronavirus”), but most prominently as a hashtag. When summing up the frequencies of all hashtags relating to the pandemic in the top 20 (# Covid19, #coronavirus, #corona, #Covid), it is the by far most frequently topic referred to in a hashtag in 2020 (510 mentions). This mirrors the huge impact Covid19 has had on the cultural sphere, among others. Second, the word “cancel” appears as the fourth most frequently used keyword in 2020. This is most likely due to public discourses on “cancel culture” that emerged in many European countries in 2020. The central role of the Covid19 pandemic and “cancel culture” in Switzerland in 2020 is also visible in the network structures of most frequent keywords and hashtags in **Figure - CH - 5** and **Figure - CH - 6**.

**Figure - CH - 5.** Network of top 30 keywords of tweets related to culture in Switzerland in 2020



**Figure - CH - 6.** Network of top 30 hashtags of tweets related to culture in Switzerland in 2020



## 5. Topic modelling

The final step of the analysis is the identification of the most prevalent topics related to culture that are discussed on Twitter in Switzerland. Again, analyses are separate for the years 2019 and 2020 in order to track temporal developments. To detect the optimal number of topics  $k$  for *Mallet* topic modelling in R, we inspected the results for  $k=7$  to  $k=12$  in each year. Using measures for perplexity and coherence as well as semantic interpretation of results, it appeared that in each year nine topics best summarize the relevant topics discussed on Twitter. As can be seen in more detail below, every topic clearly belongs to a specific language spoken in Switzerland. In both 2019 and 2020, three topics are in French, three in English, two in German, and one in Italian. This represents the structure of the data as described above quite well. Keep in mind, however, that German language is clearly underrepresented in our data relative to the population size (29% of tweets, but 65% of population). The two following sub-sections give short descriptions of the topics identified in 2019 and 2020, corresponding word-clouds for each topic are depicted in **Figure - CH - 7** and **Figure - CH - 8**.

### Topics 2019

- 1) **Popular and tech culture (English):** The most influential words here are “video”, “swiss”, “life” and “tech”, referring to recent technologies used in everyday life. But also other elements of popular culture are present, such as “magazine”, “music”, “food” or “design”. The topic seems to locate the Swiss account of these themes in an international context, not only because it is in English but also because the word “international” explicitly appears.
- 2) **Literature and art in Ticino (Italian):** The only Italian language topic identified. It accounts for culture mostly in a narrow sense, discussing various art forms present in Ticino. Most prominent are literature (“storia”, “libro”, “lingua”), visual art (“arte”, “museo”), and film (“cinema”, “serie”).
- 3) **Business culture and digital change (English):** This topic clearly relates to the business world (“company”, “business”, “corporate”, “organization”) with a special focus on developments of digital technologies (“digital”, “change”, “innovation”). But also labor relations are discussed (“team”, “leader”, “employee”).
- 4) **Film and performing arts (French):** This topic deals with culture as arts, with a specific focus on all types of art that are staged in one way or another. Film is the most discussed art form here (“film”, “festival”, “cinéma”, “star”, “scène”, “acteur”), followed by music (“musique”, “chanteur”, “album”) and theater (“théâtre”).
- 5) **Swiss-German politics and society (German):** Political and societal issues with a focus on German-speaking Switzerland (“schweizer”, “schweiz”, “deutsch”). Politics in general stand in the forefront (“politisch”, “politik”) and there is a certain trend towards youth-related (“jung”) and gender-related (“frau”) debates. This is embedded in other societal issues like education (“bildung”) or urbanity (“stadt”).
- 6) **Social inequality (English):** Here, social inequalities of several social groups (“social”) are discussed. Most prominent are ascribed inequalities between genders (“woman”), religions (“religion”), and races (“black”). Discussions seem to include in particular the role of the media (“media”, “internet”) and of violent conflicts (“violence”).

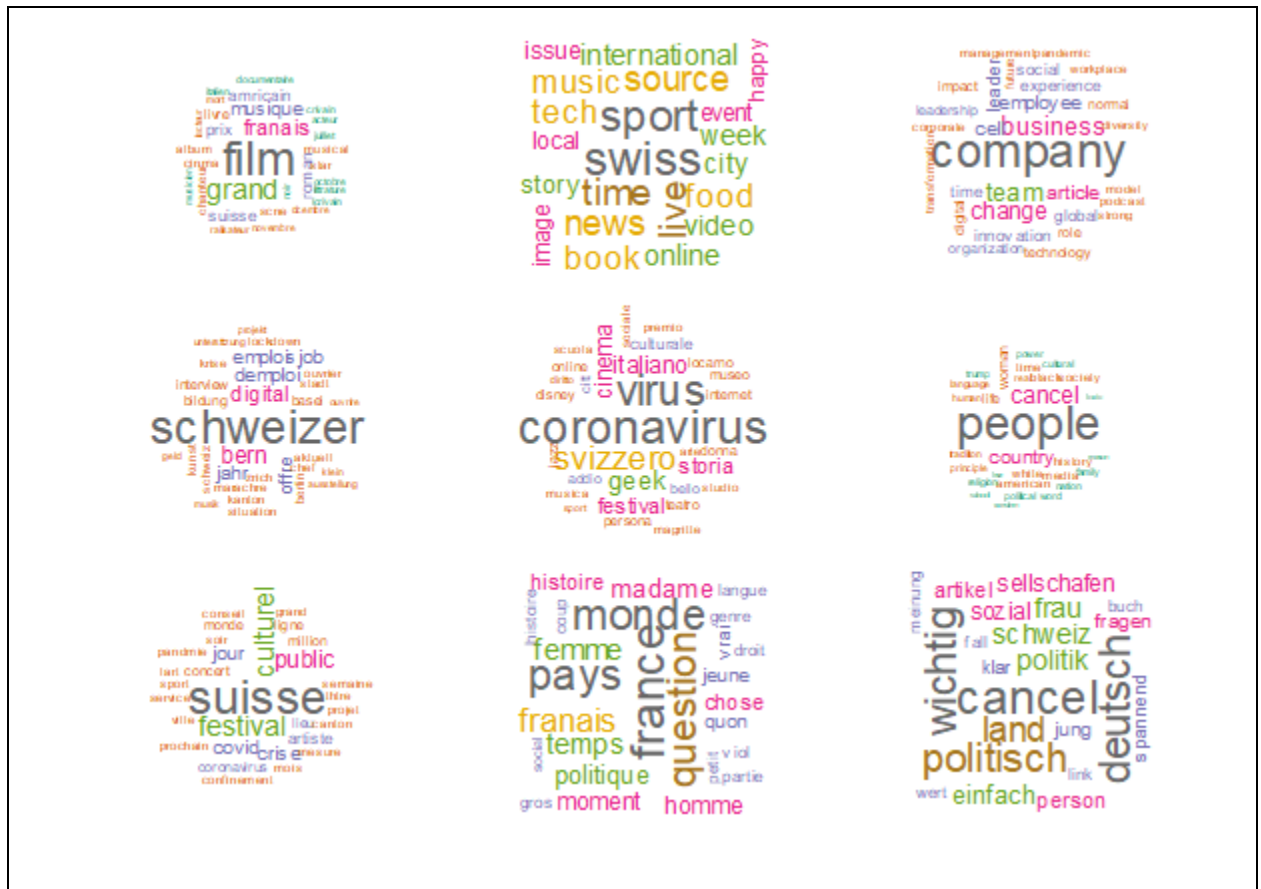
[illegible]

- ## Topics 2020

- 109



**Figure - CH - 8. Topics identified in tweets related to culture in Switzerland in 2020**



- 2) **Sport, popular, and tech culture (English):** Here, there are clear parallels with topic 1 above. However, additional to the “swiss” perspective on “international” “tech” and popular (“music”, “food”) culture, there is a strong focus on “sport”, especially “live” and “online”.
- 3) **Business culture and leadership (English):** This is a reiteration of topic 3 in 2019 on business culture (“company”, “business”, and “organization”) with a new twist. Digitization does not play a prominent role here. Instead, the focus is more on labor relations, advancement, and leadership (“team”, “employee”, “leader”, “innovation”, “experience”, “global”).
- 4) **Job market and education (mainly German):** This is an additional topic related to the economic sphere. In contrast to topic 3 it focuses on the job market (“offre”, “emploi”, “interview”, “chef”), mainly in German speaking Switzerland (“schweiz”, “bern”, “basel”, “kanton”). Additionally, the role of education (“bildung”) is discussed here.
- 5) **Covid19 and the cultural sphere in Ticino (Italian):** Again, the only Italian language topic identified in 2020. As in 2019, it is concerned primarily with culture in a narrow sense (“cinema”, “festival”, “storia”, “teatro”, “bello”). However, it is all overshadowed by the Covid19 pandemic and its impact. The words “coronavirus” and “virus” are most influential here. The pandemic could also be an explanation for the fact that also the “internet” is now discussed in this topic.
- 6) **Social inequality and cancel culture (English):** This topic is an augmented version of topic 6 in 2019. Again, ascribed inequalities between genders (“woman”) and races (“black”, “white”) are discussed, just as the role of the “media”. But now there is strong reference to debates about the so-called cancel culture. This is evident not only in the strong influence of the word “cancel”, but also in expressions referring to the use and development of language (“language”, “history”,



“tradition”. Additionally, the word “life” is now influential here, which probably points to the “Black Lives Matter” movement.

- 7) **Covid19 and the French-speaking cultural sphere (French):** Similar to topic 5, this topic is about the impact of the Covid19 pandemic on the cultural sphere (“Covid”, “crise”, “coronavirus”). Also in French-speaking Switzerland, culture is primarily perceived in a narrow sense in this context (“festival”, “artiste”, “concert”). Different from the Italian-speaking discourse, public financial support for the people affected by the pandemic seems to be additionally addressed here (“million”, “canton”).
- 8) **The French neighbor (French):** This topic is almost identical with topic 8 in 2019. In the 2020 discussion on France (“france”, “français”, “pays”), however, the gender theme is somewhat more prominent (“femme”, “madame”, “homme”), while the history and sport are less prominent.
- 9) **Swiss-German politics and cancel culture (German):** This is another topic that was present in similar form already in 2019 and has been augmented in 2020 by debates about cancel culture. Swiss-German political issues are still at the core (“politisch”, “politik”) with special focus on gender issues (“frau”). But “cancel” is the most influential word now.

## 6. Discussion and conclusion

Overall, topic modelling, keyword analysis, and hashtag analysis showed that culture is discussed in a variety of ways in the Swiss Twittersphere. Hence, we can see a fairly broad understanding of culture in our data. While there of course is a clear reference to several forms of art (especially film, but also visual arts and music), topics discussed when searching for the keyword “culture” are as diverse as technology, business, sports, politics, social inequality, language, or food. In other words, we can empirically demonstrate that many aspects of the Swiss lives are affected in one way or another by culture and this by cultural policies.

Moreover, we have seen that there is a lot of stability in the topics related to culture in Switzerland. But that doesn't mean there's no dynamic. On the one hand, the specific focus of a discussion may change even though the larger topic is stable. For example, there is a discussion about business culture in both years, but with a focus on digitization in 2019 and a focus on leadership in 2020. On the other hand, there are new themes emerging in 2020 that have not been present in 2019, most prominently Covid19 and “cancel culture”. However, such new themes do not appear as separate topics next to the existing ones. Rather, new themes get enmeshed with existing discourses or can even dominate them (e.g. in the cases of “Covid19 and the cultural sphere in Ticino” “Swiss-German politics and cancel culture”). The temporal development of discourses about culture, at least in the Swiss Twittersphere, thus takes the form of evolution rather than of disappearance and appearance. Besides, this also indicates that a medium like Twitter is prone to downplay the importance or relevance of certain themes as soon as new themes become more popular or acute.

Finally, it needs to be noted that there are some themes that appear in several topics. For instance, gender issues are prevalent in topics 5, 6, and 8 in 2019 and in topics 6, 8, and 9 in 2020. Similarly, digitization, Covid19, or cancel culture can be detected in multiple topics. Such themes could be denoted as “meta-topics”, as they are of importance in various contexts.

## Findings for the United Kingdom

### 1. Twitter usage in the UK

Based on statistical information provided by Umpf , Twitter is the third most used social media platform in the UK, with 15.5 million adult users (to compare, 37.4 million British adults use Facebook, 32.1 million use YouTube, 7.9 million use LinkedIn). Twitter is mostly used among 18-24 year olds in the UK, with no gender differences.

### 2. Data description

Using Twitter academic account API, we scraped tweets by searching for the keyword “culture”, as it was posted on Twitter in the years 2019 and 2020. For the purpose of our analysis, we compared tweets posted before and after the Covid-19 pandemic. Based on reports in the UK, the 29th of January marks the day of the arrival of the virus into the country . Thus, the 2019 data includes all tweets posted before the spread of the pandemic in the UK, whereas the 2020 tweets include those posted after the 29th of January. A total of 102,835 tweets was scraped in both years. In 2019, 55,108 tweets were scraped, and 47,727 tweets were scraped from the year 2020.

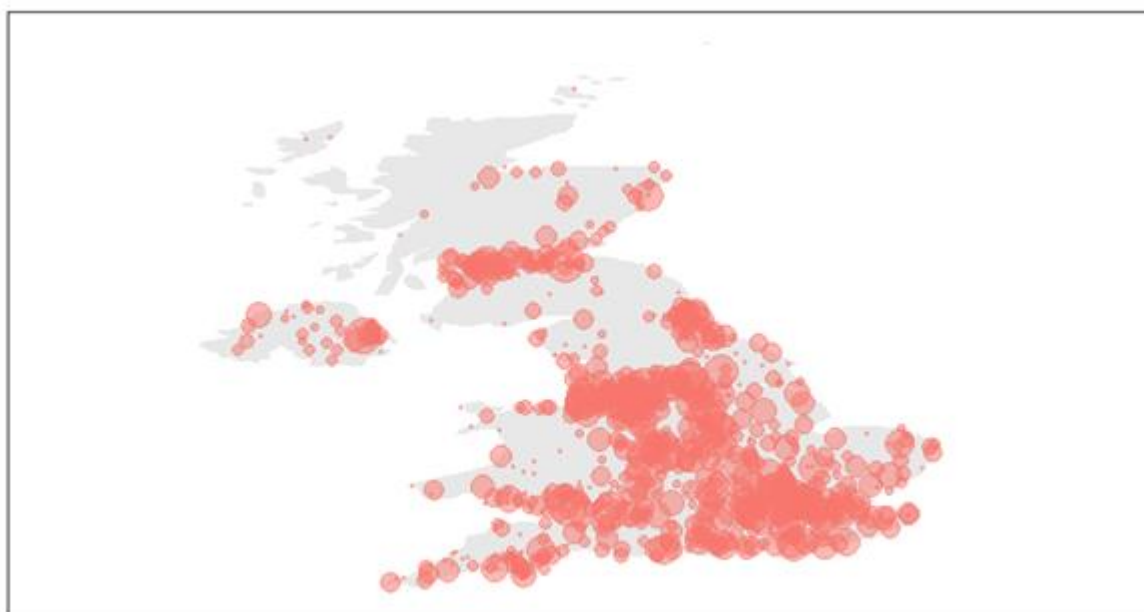
The tweets we have scraped were posted by 21,258 unique users (10,289 in 2019, 10,969 in 2020). Average number of tweets per user was 3.2 tweets. Looking into the top users who posted tweets using the keyword “culture”, we found many verified accounts of famous figures such as: Jimmy Carr (comedian), politicians (Natalie Bennett), and the royal family. Our top users also include social activists (Heather Peterson and Dr Amalina Bakri). However, no cultural organizations were part of our top users (except one museum), as we might have presumed. We were surprised to find many individuals who work for the NHS (the national health system in the UK). The significant participation of such individuals in the discourse about culture on Twitter explain our findings that will be introduced later. 62% of total users mention other users. On average, per tweet, users mention more than 3 users (3.8). In 2019, the top users represent 12% of all scraped tweets, in 2020, they made 14% of all tweets.

The top ten locations from which users tweeted about culture were: London (7,129 tweets), Manchester (1,160 tweets), Glasgow (809 tweets), Edinburgh (733 tweets), Salford (676 tweets), Birmingham (649 tweets), Liverpool (555 tweets), Bristol (547 tweets), Leeds (476 tweets), and Cardiff (438 tweets). This demonstrates that most tweets were posted by users who live in major, metropolitan cities in England (London and Manchester), Scotland (Edinburgh and Glasgow), and Wales (Cardiff). The only city that is exceptional in this list, is Salford, a working-class city that is considered as part of Greater Manchester.

**Table - UK - 1. Most active profiles using the key-word "Culture"**

Year	User	Description	Number of followers	Number of Tweets Using "Culture"
2019	@jimmycarr (verified)	Comedian	6,731,169	1,759
	@xstex	Marketing	201,537	1,267
	@Timothy_Hughes (verified)	Religious leader/singer	78,490	924
	@natalieben (verified)	Politician	170,432	786
	@webmasterdave	Researcher	112,823	572
	@heatherpaterson	Social activist	115,256	335
	@Moonbootica	NHS employee	3,880	264
	@WhatsOnOLIO	Food sharing app	1,088	207
	@DrUmeshPrabhu	Medical director	32,689	195
	@ErnestBilko	Museum	4,539	98
	@RoyalFamily (verified)	Royals	4,587,844	1,857
	@ChrisSmalling (verified)	Football player	1,791,762	1,325
	@DrAmalinaBakri (verified)	NHS employee	739,636	1,002
2020	@xstex	Marketing	201,537	610
	@natalieben (verified)	Politician	170,432	537
	@thomaspower	Entrepreneur	300,893	463
	@Moonbootica	NHS employee	3,880	325
	@WhatsOnOLIO	Food sharing app	1,088	318
	@Orgetorix	Researcher	3,282	174
	@DrUmeshPrabhu	Medical director	32,689	62

**Figure - UK - 1.** Tweets distribution map in UK before and during Covid-19



### 3. Hashtag and keyword analysis 2019

In total, 83,679 hashtags, and 54,003 keywords were found. The most popular keywords were “people”, “staff”, and “time”, while the most used hashtags were #bullying, #leadership, and #NHS. This means that most of these keywords and hashtags referred to a broader understanding of the term. Most tweets reflected the heated discussion regarding the NHS, UK’s National Health System, and the quality of service they provide. Looking into the keywords and hashtags, only one word/hashtag from the top 10 relates to culture in the narrow – “art”. Table 1 illustrates the most popular hashtags and keywords (excluding the keyword “culture”):

**Table - UK - 2.** Top 10 Hashtags and Keywords

Hashtags	Freq.	Keywords	Freq.
#bullying	1920	people	5062
#leadership	1845	staff	4962
#nhs	1401	time	2637
#patients	1152	day	2591
#speakup	904	patient	2441
#london	624	harm	2286
#art	542	care	2143
#patientsafety	488	happy	1976
#ptsafety	362	art	1917
#leaders	341	leader	1878

Similarly, in the hashtags hub (see **Figure - UK - 2**), we see NHS-related hashtags. We also find that London is considered a linchpin of British society, with connected hashtags like #UK, #Community, #British. Lastly, we find a cluster of hashtags that reflects to a narrow sense of culture, such as #Dance, #Photography. These hashtags are particularly related to hashtags describing African heritage (#Africa, #Afrobeats, #Afroculture).

---

1000

#### 4. Hashtag and keyword analysis 2020

In 2020, a total of 79,474 hashtags, and 51,590 keywords were tagged. The most popular keywords were “people”, “time”, and “art”, while the most popular hashtags were #speakup, #art, and #NHS (see **Table - UK - 3**). This means that most keywords and that appeared the most in tweets that included the keyword “culture”, referred to a broader understanding of the term. Most tweets reflected the heated discussion on the NHS, UK’s National Health System, and the quality of service they are able to provide, especially in light of the pandemic. Looking into the most used hashtags we see an emphasis on the arts more than in 2019. This is possibly due to lockdown restrictions that have put a pause on the sector as a whole. Table 2 illustrates the most popular hashtags and keywords (excluding the keyword “culture”):

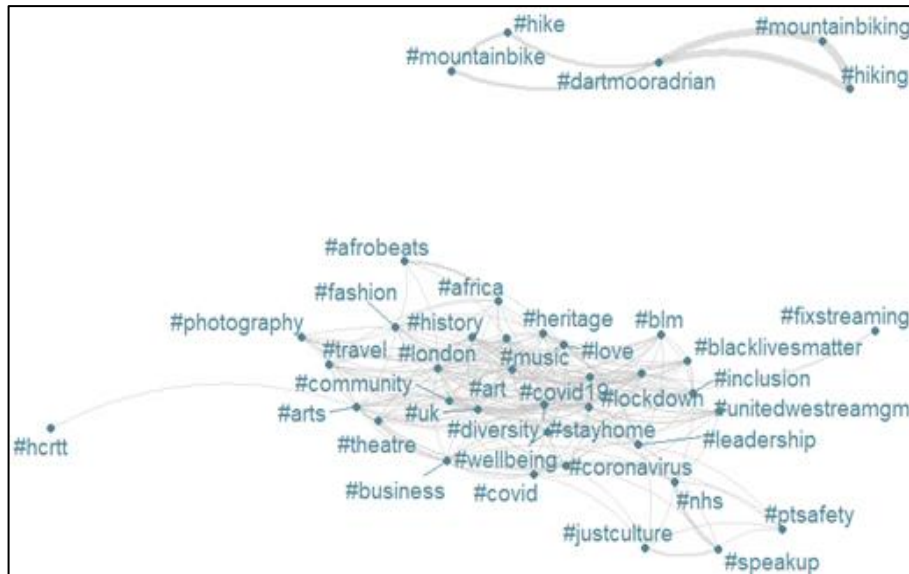
**Table - UK - 3. Top 10 Hashtags and Keywords**

Hashtags	Freq.	Keywords	Freq.
#speakup	344	people	5442
#art	297	time	3320
#nhs	278	art	1823
#forthearts	232	staff	1823
#arts	200	day	1757
#leadership	193	week	1643
#london	163	calendar	1573
#dartmooradrian	160	liturgical	1568
#Covid19	149	history	1454
#music	145	country	1453

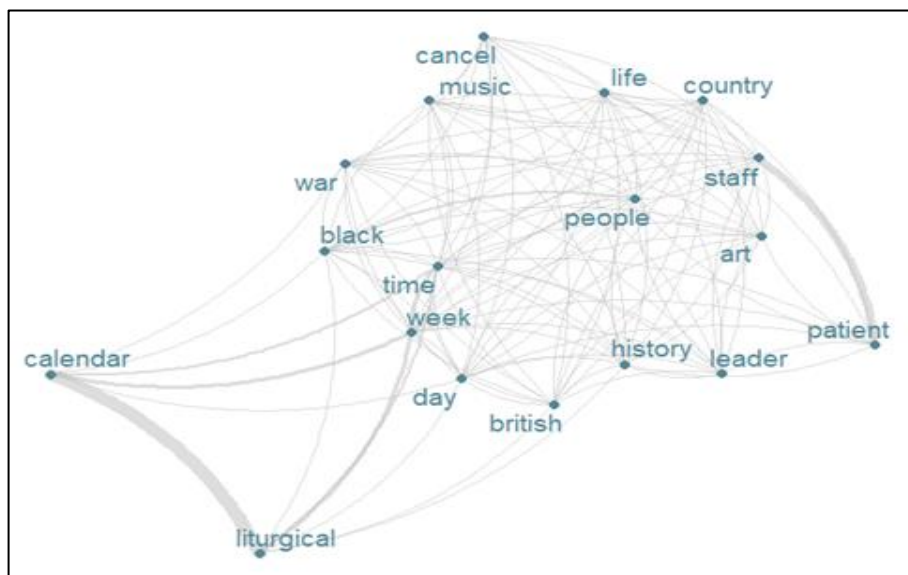
Looking into the hubs of keywords and hashtags we see a reflection of the Covid-19 pandemic. In the keywords hub (see Graph 4), there is a clear connection between “staff” and “patient”, reflecting on the Covid-19 pandemic. Similarly, we find many pandemic-related hashtags like #Covid19, #Lockdown, and #Stayhome. We also detect keywords and hashtag relating to the BLM movement which was prominent during 2020: “Black” and “War” are tied in the keywords hub, while the hashtags hub includes #BLM, #Blacklivesmatter, #Africa, #Afrobeats. In the keywords hub there is a strong connection between calendar and Liturgical, which possibly reflects on the promotion of religious discourse on British Twitter. In the hashtags hub (see **Figure - UK - 4**) we see that #Theatre is not only related to the #Arts, but also to #Wellbeing. This possibly reflects on the ways in which the lack of culture and arts can impair one’s wellbeing, especially during the pandemic.



**Figure - UK - 4. Top 20 Hashtags Network - 2020**



**Figure - UK - 5. Top 20 Keywords Network - 2020**



## 5. Topic Modelling

The following section compares the topic of the tweets posted in 2019 and 2020, as analyzed through Topic Modelling on R. To define the best topic number K, we estimate several models, from K = 10 to K = 14. It appears that K = 10 represents the best results in terms of topic variety. We now introduce the ten topics that were discussed in 2019 and 2020 when users used the word “culture”.



(2) Organizations: This cluster includes words such as “team” (0.041), “business” (0.017), “company” (0.013), and “strategy” (0.009). These words demonstrate the collective action that occurs in the workplace.



(4) Urban Leisure: Here we have words such as “city” (0.036), “event” (0.026), “food” (0.022) and “street” (0.009). These demonstrate the strong connection between cultural practices and leisure time.

**time**



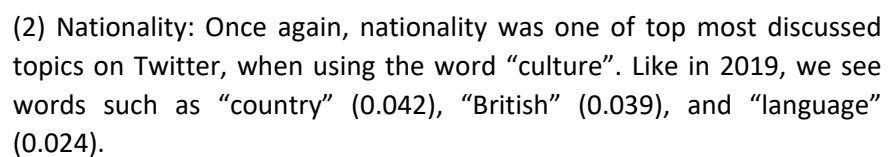
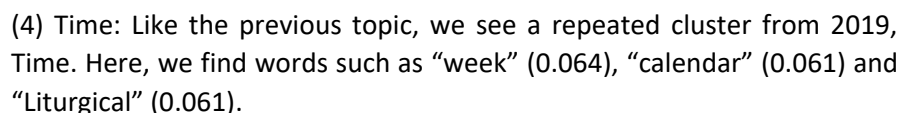
(6) Political Identities: In this cluster of topics we have words such as, “people” (0.13), “women” (0.025), “black” (0.016), and “white” (0.013).

(8) **Nationality:** Our last topic incorporates words related to nationality, such as “country” (0.04), “British” (0.03), and “language” (0.03). In post-Brexit Britain, this topic is not surprising, as discussions on who is British or European continue to dominate the everyday discourse.

(10) Education: The words in this topic are related to the education system, like “school” (0.026), “child” (0.018), and “education” (0.012).

The chart displays the monthly volume of tweets for various topics in 2019. The Y-axis represents the 'Number of Tweets' from 0 to 800, and the X-axis represents the months from January to December. The topics tracked are Positive culture, Organizations, Health care, Urban leisure, Time/Holiday, Political identities, Social divisions, Nationality, Sport, and Education. Nationality consistently has the highest number of tweets, peaking at over 800 in July. Organizations and Positive culture also show high volumes, with Organizations peaking in November and Positive culture peaking in June. Other topics like Health care, Urban leisure, Time/Holiday, Political identities, Social divisions, and Education show more fluctuation, with peaks occurring at different times throughout the year. Sport and Education generally have the lowest tweet counts.

Topic	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Positive culture	600	580	550	580	600	650	620	560	560	580	580	540
Organizations	580	600	600	520	580	680	580	420	580	660	700	420
Health care	450	530	300	480	580	450	500	320	330	390	350	210
Urban leisure	420	450	480	480	500	500	500	400	400	420	400	350
Time/Holiday	400	380	450	400	400	450	450	400	480	450	420	280
Political identities	420	420	400	380	450	480	450	420	400	420	450	350
Social divisions	220	250	150	180	200	180	180	120	190	220	230	170
Nationality	700	750	700	580	620	750	820	660	640	620	640	660
Sport	420	450	400	400	450	450	450	400	400	420	450	350
Education	450	550	500	420	520	580	550	350	550	600	580	350

[illegible]



(6) Political Identity: This cluster is also a repetition from 2019, including words such as “people” (0.141), “Black” (0.032), “woman” (0.021) and “cancel” (0.037).

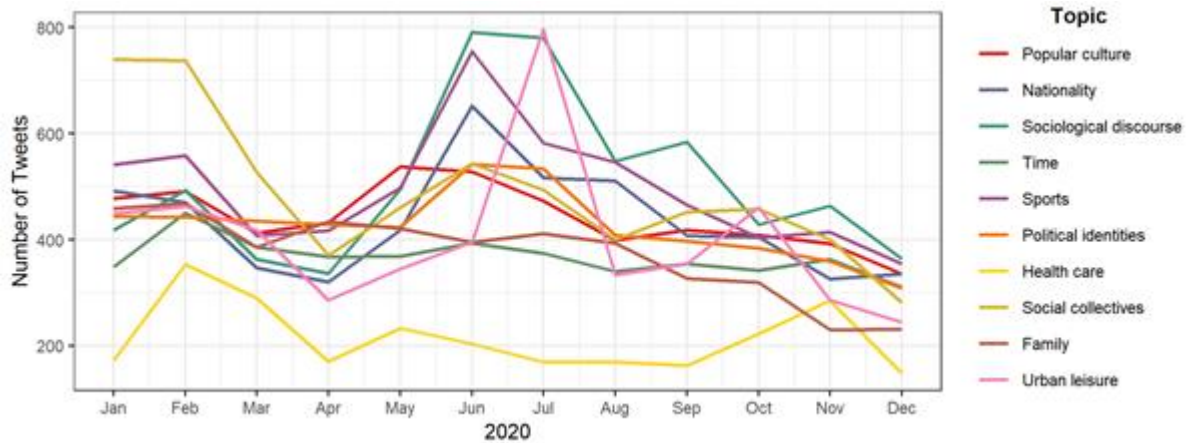
A word cloud of terms related to the nursing profession. The most prominent words are 'staff' and 'patient'. Other visible words include 'nurse', 'care', 'work', 'life', 'stress', 'burnout', 'compassion', 'dedication', 'teamwork', 'communication', 'empathy', 'resilience', 'growth', 'learning', 'collaboration', 'support', 'motivation', 'passion', 'dedication', 'teamwork', 'communication', 'empathy', 'resilience', 'growth', 'learning', 'collaboration', 'support', 'motivation', 'passion'. The words are arranged in a circular pattern, with some words appearing multiple times.

(8) Social collectives: This cluster includes different structures of collectives or communities, such as “team” (0.027), “business” (0.015), and “school” (0.021).



(10) Urban Leisure: In this cluster we have words such as “city” (0.027), “museum” (0.01), and “local” (0.015). We also have the word “support” (0.017), which might suggest that during the 2020 lockdown, there was more encouragement online to support local culture.

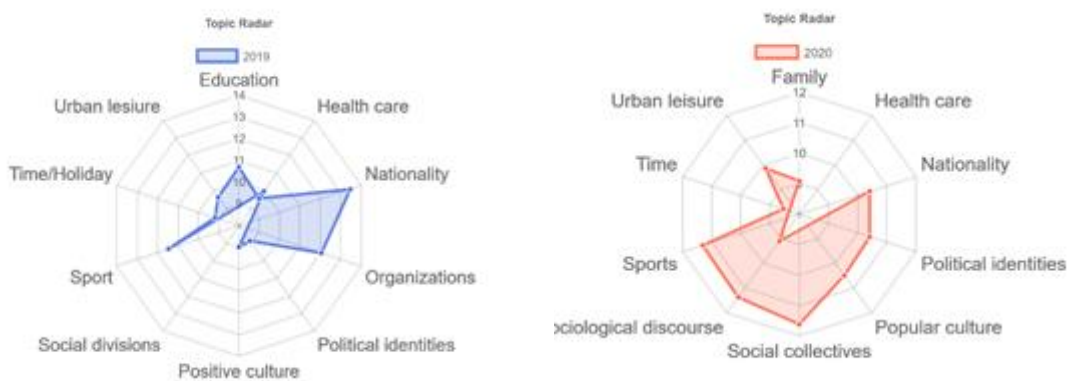
**Figure - UK - 7. Occurance of the topics throughout 2020**



In first glance, we notice that 2020 was a tumultuous time in the UK, where different topics drew focused attention, but quickly were replaced by other discourses. For instance, discussions about the social collective were at their peak during the beginning of the pandemic, but then were significantly neglected, probably due to UK's full lockdown. Other topics also correspond with lockdown measurements. For instance, we can see that discussions on sports peaked in May, once the government declared its roadmap to opening up sports matches. Similarly, sociological discussions were prominent during June and July, during the peak of the BLM movement. Health care was the least discussed topic throughout the year.

## 6. Discussion and conclusion

**Figure - UK - 8. Topics distribution 2019-2020**



Comparing topic radar from 2019, pre-Covid times and 2020, we can see that while “nationality” was a frequently discussed topic in both years, it was discussed less in 2020. “Sports” was also a recurring topic that was frequently discussed in both years. In 2020, the UK Twittersphere focused more on social matters than in 2019, which could be understandable due to the state of emergency that raised heated social discussions. A new topic that emerged in 2020 was “family”, which is also possibly due to lockdown, and people spending either more or less time with their families and relatives.



In summary, we learn from the findings of the data scraping that when people in the UK talk about “culture”, they talk about almost anything - whether it is a more political discourse about identities, health care system, education and nationalities, or culture in a more narrow sense, such as popular culture, arts and leisure. As expected, a change was spotted between 2019 and 2020 due to the Covid-19. The findings presented here call for a more specific definition of culture, both for its expressions, and for the way it is mentioned in social media.

## General Discussion and Conclusion

This section puts the results of each country's Twitter analysis into comparative perspective, focusing on the topics identified with *Mallet* topic modelling. There will be two types of comparison: First, a spatial comparison between countries and, second, a temporal comparison between the pre-Covid and during-Covid periods. In order to enable better comparative interpretations, we merged all the topics identified in each country and period to broader groups of topics. This structure is depicted in Table 3. We identified eight thematic topic groups: "Film/TV/Theatre", "Literature/Arts", "Music/Concerts/Festivals", "Policies and Politics", "Identities and Boundaries", "Society and Inequalities", "Economics/Business/Work", and "Spaces and Places". All of them can be found in almost all countries, so these topic groups are at the heart of discussions related to culture in Europe. Additionally, there is a group called "Other Topics", which contains topics that could not be allocated to any of the other groups. In most cases, these "other topics" are found only in one or two countries. So, they represent country-specific topics. We should note that also the concrete manifestations of the eight thematic groups show variations by country. Hence, country-specific characteristics can also be found within the eight thematic groups. We will discuss the similarities and differences between countries in greater detail in the following subsections.

### 1. Country comparison

Overall, Table 3 shows that there are great similarities between the culture-related discussions of each country's Twittersphere. On the one hand, when people talk about culture, they often talk about culture in a narrow sense, i.e. culture as some form of art. Three of the eight thematic topic groups refer to the sphere of arts and creativity and all countries contribute at least one topic to each group. Hence, there are Twitter users in all countries who relate culture to (a) filmmaking, cinemas, television, online streaming, theatres, and similar kinds of performing arts ("Film/TV/Theatre"); (b) books, literature, or visual arts presented in museums or galleries ("Literature/Arts"); and (c) music, concerts, festivals, and similar events ("Music/Concerts/Festivals").

On the other hand, the meaning of culture is clearly not restricted to the arts. The five remaining thematic topic groups cover a wide range of political, societal, and economical discussions. More precisely, culture in a broader sense can refer to such diverse topics as (d) political debates about the distribution of public funds (especially for the arts sector) or the educational system ("Policies and Politics"); (e) the demarcation of boundaries between social groups, often on the level of nation states and therefore aiming at reinforcing national identities ("Identities and Boundaries"); (f) the cohesion of societies and the unequal opportunities of privileged and underprivileged groups, most prominently inequalities between genders and races ("Society and Inequalities"); (g) economic developments related to organizational structures, management, employment relations, or digitization ("Economics/Business/Work"); and (h) differences between particular spatial entities such as regions (including a rural-urban divide), countries, or cities ("Spaces and Places"). The fact that all topic groups (a) to (h) can be detected in almost all countries studied indicates that there is something like a common core of what culture means in Europe. Of course, this does not mean that the eight thematic topic groups encompass everything that culture can possibly be understood to be. The simple fact that there are "Other Topics" related to culture thwarts this interpretation.

**Table 3.** Allocation of all topics related to culture identified in all countries to topic groups

Country	N	Film/ TV / Theatre	Literature/ Arts	Music/ Concerts/ Festivals	Policies and Politics	Identities and Boundaries	Society and Inequalities	Economics/ Business/ Work	Spaces and Places	Other Topics
HR 2019	10	Film and theatre	Cultural events	Music events	Cultural policy European Capital of Culture	National identity Cultural identity & heritage	Culture and community			Everyday culture Culture as values
HR 2020	5		Art Books		European Capital of Culture	National identity				Culture as values
DK 2019	11	Netflix Denmark Danish cinema	Literature and art	Music	Politics	English cultural discussion in DK Living in Denmark	Metoo Denmark Danish society and future		Local culture International Culture	
DK 2020	6	Cinema and TV	Danish cultural discussion in DK <sup>(1)</sup>	Danish cultural discussion in DK <sup>(1)</sup>	Danish cultural discussion in DK <sup>(1)</sup>	English cultural discussion in DK Living in Denmark	Metoo Denmark			Cancel culture
FI 2019	10	Culture in the narrow sense <sup>(1)</sup> Congratulations and promotion <sup>(1)</sup>	Culture in the narrow sense <sup>(1)</sup> Congratulations and promotion <sup>(1)</sup> Cultural reviews <sup>(1)</sup> Books and literature	Culture in the narrow sense <sup>(1)</sup> Congratulations and promotion <sup>(1)</sup> Cultural reviews <sup>(1)</sup>	Public funding Well-being promotion	Finnish identity Social and cultural debates <sup>(1)</sup>	Social and cultural debates <sup>(1)</sup>	Business or management jargon		Journalism
FI 2020	10	Congratulations and promotion <sup>(1)</sup> Promotion of cultural events <sup>(1)</sup>	Congratulations and promotion <sup>(1)</sup> Literature and journalism	Congratulations and promotion <sup>(1)</sup> Promotion of cultural events <sup>(1)</sup>	Public funding (Covid19)	Finnish identity Social and cultural debates <sup>(1)</sup> Global culture	Social and cultural debates <sup>(1)</sup>	Business or management jargon		Politics and cancel culture General discussion

Country	N	Film/ TV / Theatre	Literature/ Arts	Music/ Concerts/ Festivals	Policies and Politics	Identities and Boundaries	Society and Inequalities	Economics/ Business/ Work	Spaces and Places	Other Topics
FR 2019	10	The new season of cultural events <sup>(1)</sup>	Literature	The new season of cultural events <sup>(1)</sup>	Culture and education in the third sector	Far right political clash		Business and economic trends	World languages and cultures	Public commotion
		Series, films, and music videos	Art exhibitions							General knowledge
FR 2020	10	Cultural activities under Covid19 <sup>(1)</sup>	The upcoming season of cultural events <sup>(1)</sup>	The upcoming season of cultural events <sup>(1)</sup>	Local projects and 2020 mayoral elections	Far right and multiculturalism	Rape culture			Football and journalism
		Cultural suggestions in pandemic time	Literature	Cultural activities under Covid19 <sup>(1)</sup>	Cultural policies under Covid19					The constraints of the pandemic
NL 2019	17	Film, TV, and theatre	Museum and visual arts	Music offerings	Cultural policy Politics and political culture	Languages and cultures Religion and multicultural society	Organizations and diversity <sup>(1)</sup> Socio-cultural differences and inequalities	Organizations and diversity <sup>(1)</sup>	Urban environment Tourism and leisure	Crime and deviant behavior Reflections and present, past, and future
		Awards and recognition <sup>(1)</sup>	Books and literature		Children and wellbeing	Contested heritage and racism				
NL 2020	19	Film, TV, and theatre	Museum and visual arts	Music offerings	Cultural policy Politics and political culture	Languages and cultures Religion and multicultural society	Organizations and diversity <sup>(1)</sup> Socio-cultural differences and inequalities	Organizations and diversity <sup>(1)</sup>		Online culture Crime and deviant behavior Reflections and present, past, and future
		Access to venues during Covid19 <sup>(1)</sup>	Books and literature Access to venues during Covid19 <sup>(1)</sup>	Access to venues during Covid19 <sup>(1)</sup>	Children and wellbeing Covid19 support for cultural sector	Contested heritage and racism	Societal impacts of Covid19 <sup>(1)</sup>	Societal impacts of Covid19 <sup>(1)</sup>		Cancel culture

Country	N	Film/ TV / Theatre	Literature/ Arts	Music/ Concerts/ Festivals	Policies and Politics	Identities and Boundaries	Society and Inequalities	Economics/ Business/ Work	Spaces and Places	Other Topics
RS 2019	9	Film and TV	Cultural events <sup>(1)</sup> Literature	Cultural events <sup>(1)</sup>	Ministry of Culture  Urban cultural policy Education	National identity	Cultural distinction		Territory	
RS 2020	11	Film and theatre	Cultural events (Covid19) <sup>(1)</sup> Literature	Cultural events (Covid19) <sup>(1)</sup>	Ministry of Culture  Urban cultural policy Education  Fiscal policy	National identity	Societal dialogue  Emigration			Illegitimate culture
ES 2019	10	Global culture	Legitimate culture Literature	Cultural events Scenic arts	Cultural policies	Nationality and identity  Catalan culture	Social divisions		Culture in cosmopolitan cities	
ES 2020	10	Audio-visual culture	Legitimate culture Literature	Scenic arts (Covid19)	Cultural policies related to the pandemic  Education (Covid19)	Nationality and political identity  Catalan culture	Social divisions		Global culture	

Country	N	Film/ TV / Theatre	Literature/ Arts	Music/ Concerts/ Festivals	Policies and Politics	Identities and Boundaries	Society and Inequalities	Economics/ Business/ Work	Spaces and Places	Other Topics
CH 2019	9	Literature and art in Ticino <sup>(1)</sup>	Literature and art in Ticino <sup>(1)</sup>	Urban art and entertainment <sup>(1)</sup>	Swiss-German politics and society	The French neighbor	Social inequality	Business culture and digital change	German-speaking city culture	Popular and tech culture
		Film and performing arts	Urban art and entertainment <sup>(1)</sup>							
CH 2020	9	Film and performing arts	Covid19 and the cultural sphere in Ticino <sup>(1)</sup>	Covid19 and the French speaking cultural sphere <sup>(1)</sup>	Covid19 and the French speaking cultural sphere <sup>(1)</sup>	The French neighbor	Social inequality and cancel culture	Business culture and leadership		Sport, popular, and tech culture
		Covid19 and the cultural sphere in Ticino <sup>(1)</sup>			Swiss-German politics and cancel culture			Job market and education		
UK 2019	10	Positive expressions of culture <sup>(1)</sup>		Positive expressions of culture <sup>(1)</sup>	Health care	Political identities <sup>(1)</sup>	Social divisions			Time/religious holidays
					Education	Nationality	Political identities <sup>(1)</sup>	Organizations	Urban leisure	Sports
UK 2020	10	Popular culture <sup>(1)</sup>	Popular culture <sup>(1)</sup>		Health care	Political identities <sup>(1)</sup>	Sociological discourse			Time
						Nationality	Political identities <sup>(1)</sup>	Social collectives	Urban leisure	Sports
							Family			

<sup>(1)</sup> These topics contain elements that are attributable to multiple topic groups per country and year.  
Note: The exact meaning of each topic is described in more detail in the corresponding country chapter.



Nevertheless, our analysis clearly shows that there is a widely shared understanding of culture that refers not only to the arts, leisure, and creativity, but also to identities, politics, and social inequalities, as well as to more mundane aspects of human life such as business, work, and economics.

One has to take into account, however, that these results are strongly influenced by the composition of users posting and discussing on Twitter. In most countries established news media organizations (e.g. newspapers, broadcasting stations) are the most active Tweepers related to culture. Hence, our data could be biased towards topics that are particularly relevant for such news media. On the other hand, one could argue that news media are also keen on reaching as many people as possible in order to sell more copies or generate more traffic on their websites. Consequently, they often seek to cover a very broad range of topics, which in turn would be an advantage for our analyses.

Another similarity between many countries, albeit not all and to varying degrees, is the relationship between culture and the emerging Covid19 pandemic in 2020 (notably in Finland, France, the Netherlands, Serbia, Spain, and Switzerland). Interestingly, however, Covid19 does not appear as a separate topic or topic group. Rather, it is discussed within existing topics from several topic groups. For instance, in Finland, France, the Netherlands, Spain, and Switzerland Covid19 impacts discourses about public subsidies for the arts and related cultural policies. In France, the Netherlands, Serbia, and Switzerland, Twitter users additionally discuss the lockdown of cultural venues and events (e.g. cinemas, museums, festivals), and, at least in France and the Netherlands, possible (digital) alternatives. In the Netherlands, France, and Spain, Covid19 appears also related to discussions about broader cultural issues such as healthcare, education, telework, sports, and tourism.

Despite these similarities across countries, there are also some telling differences worth mentioning. Every country exhibits some specific thematic features and/or user characteristics related to culture that are not found in other countries. Hence, despite a common core of meanings of culture, the concrete manifestations of culture may vary according to particular, often national contexts. Therefore, in the following, we will describe specific characteristics of culture country by country:

*Croatia (HR)* stands out as the only country that features a separate topic on “Everyday culture”, discussing issues like work, knowledge, or love connected to the realm of culture. Moreover, the discourse on “Policies and politics” in 2019 and 2020 is strongly shaped by Rijeka, Croatia’s third largest city, being a European Capital of Culture in 2020. This is explainable by the fact that the city of Rijeka was one of the most active Twitter users in the present sample. At the same time, Croatia is the country where some core topics of culture seem to be less developed, since, at least in our data, we could not find any references to “Economics/Business/Work” and “Spaces and Places”.

In *Denmark (DK)*, discussions about social inequalities are clearly dominated by the hashtag “metoo”, focusing on gender inequalities and violence against women. Also, Denmark is one of four countries where emerging public debates about “cancel culture” are prominently present. Together, this indicates a rather strong representation of global cultural phenomena in this relatively small European country. Apart from that, Danish cultural discourses in a narrow sense are very much related to the film industry, which reflects the fact that news media specialising in film issues are very active on Twitter in Denmark, at least during the period under study.

The outstanding feature of *Finland (FI)* in the present analysis is its explicit emphasis on the role of the media and journalism when talking about culture. This refers primarily to more established news media and the quality of the articles they publish (topics “Journalism” and “General discussion”). In addition, the fact that “cancel culture” is also a highly debated topic in Finland fits in well with this. On the other side, Finnish Twitter users do not explicitly relate culture to any specific “Spaces and Places”

*France (FR)* shows two interconnected specific features. On the one hand, discussions about the French national identity are directly tied to the political far right movement and its relation to foreign cultures and religions. On the other hand, there is a topic on “Rape culture” that focuses on violence against women and gender relations (similar to Denmark). At least in 2019, these two topics are mixed up in discussions about gender issues in the Arabic world, with the case of soccer player Neymar serving as a prominent example. Furthermore, France is a country where, besides news media and administrative bodies, ordinary citizens are among the most active Twitter users in the field of culture.

*The Netherlands (NL)* are a special case in several respects. Generally, the country yields the highest numbers of topics identified, both in 2019 (17) and 2020 (19). This is most likely due to the fact that the largest number of tweets on culture was found in the Netherlands (in total 185,843). Therefore, some cultural issues were identified as separate topics only there: digitization and digital cultural offerings (“Online culture”); crime, violence, and social norms (“Crime and deviant behavior”), and “Tourism and leisure”. Additionally, discussions about belonging to the national society (often in terms of languages and religion) are particularly pronounced, albeit not limited to the Dutch data. The same is true for the case of “cancel culture”, which represents a separate topic in the Netherlands in 2020.

*Serbia (RS)* could be described as featuring the most politicized cultural discussions on Twitter. Seven out of 20 topics were allocated to the “Policies and politics” group, with the “Ministry of Culture” being a debated political topic in itself. Also the spatial dimension of cultural discussions is heavily politically biased, as the topic “Territory” has a strong focus on the relationship between Serbia and Kosovo. The political gradient of Serbian discourses about culture is underscored by the fact that many politicians are among the very active Twitter users in the Serbian data.

The specific characteristic of *Spain (ES)* is, somewhat similar to the case of Kosovo in Serbia, a debate about boundaries and identities related to the autonomous region of Catalonia. The topic “Catalan culture” deals with differences in the languages, traditions, and history between Catalonia and the Spanish heartland. Interestingly, this goes along with Spain being the only country where intellectuals and academics are the most active Twitter users related to culture.

Debates about culture in *Switzerland (CH)* are strongly influenced by the multilingualism of the country. This is manifested in the fact that many discussions - be it about culture in the narrow sense or in a broader sense - are restricted to one of the three language regions (German, French, Italian). Moreover, one third of the topics identified in Switzerland are in English language. This reflects the pronounced presence of migrants and expats in the country. In line with this, these English-language topics mainly revolve around Swiss popular culture, the business world, and social inequality.

The *United Kingdom (UK)* stands out especially with cultural discourses that deal with the national health care system NHS, with a particular focus on the staff working there. Not surprisingly, then, individuals working for the NHS are among the most active Twitter users in the UK sample. At the same time, Britain is a country where sports (especially soccer) are an integral part of discussion about culture. The latter is found in some other countries as well (e.g. France or Switzerland), but not to the same extent as in the United Kingdom.

## 2. Time comparison

In order to establish whether the culture-related communication on Twitter differed before and during Covid19, we collected and compared data for two periods: before and during the pandemic.

Overall, for most countries, we find a lot of stability in the topics that are discussed before and during the pandemic in relation to culture (see Table 3 and the individual country chapters). This stability is likely connected with the prominence of various categories of institutional and professional users that routinely use Twitter and other social media for dissemination and discussion of content related to their areas of specialization and interest. These highly active users include, for example, news media, (cultural) organizations, policy bodies, NGOs, journalists, scholars, activists, and artists who mostly maintain and sometimes even increase their presence on Twitter during the pandemic.

However, this continuity of topics does not imply that nothing changes in the Twitter discourses about culture.

In some countries (Croatia, Denmark), the range of topics and discourses that we can identify in the 2020 data decreases compared to the pre-Covid period, whereas in some other countries (the Netherlands and Serbia) a wider range of topics appears to be discussed. For most countries the number of distinct topics remains the same before and during the pandemic.

Furthermore, in most countries, the specific focus or content of some topics clearly changes over time, even though the larger topic remains present. For instance, the focus of the Swiss Twitter discussion on “Business culture” shifts from “digitization” to “leadership”. Similarly, the scope of the Spanish discussion about “Social divisions” broadens to include generational differences and education levels next to gender and social status.

In most countries, also new topics emerge in the 2020 Twittersphere that are not (very) manifest in 2019, a prominent example being “Cancel culture”. In some countries (Denmark, the Netherlands, Finland), it appears as a separate topic, whereas in other countries it is enmeshed with existing discourses, for instance, with the ongoing social inequalities debate in the Swiss case.

Last but not least, as already noted in the previous section, the impact of Covid-19 is clearly visible in the 2020 data for most countries, especially in discussions about cultural offerings and venues (e.g. limited access, online alternatives), and policies and public funding (e.g. crisis support for culture and other sectors), but also in culture-related discussions about tourism, sports, leisure, healthcare, and education. In some countries, new, Covid-related topics emerge, such as “Education” in the Spanish case, related to people discussing alternative, digital tools for teaching and learning, and “Family” in the U.K. case, discussing Covid-related changes in family life (e.g., spending much more time with one’s family than before, or being unable to visit one’s relatives).

### 3. Limitations and future research

We conducted exploratory research on comments posted on Twitter. Before collecting and analyzing the tweets, we needed to specify a strategy for the keywords or hashtags to be targeted in our search. After ample deliberation and pilot searches in each country, we decided to search for keywords and focus on the most general keyword, “culture”. This strategy allowed us to obtain a complete sense of culture without an initial boundary cut. Starting research with the scraping of sizeable textual data on Twitter, we espoused a bottom-up approach to discover the structures and themes emerging from the data for comparative analysis. Despite choosing a general keyword, this approach restricted the Twitter data to tweets that explicitly mention the keyword “culture” in the text or in the metadata. Tweets about cultural offerings that did not feature the word “culture” were, therefore, not included in our data. We assume that due to this sampling strategy limitation and/or because of the characteristics of the most avid users of Twitter — certain popular cultural topics such as video games, celebrities, lifestyles, or food are not captured very well by our data.

We have found many different types of users in the data. However, a small proportion of users is responsible for significant to large shares of the culture related tweets in each country, namely journalists and media organizations, cultural institutions, NGOs, private citizens, professional athletes, artists, cultural industries, large international companies, or entrepreneurs. In 2020, for instance, the share of the 10 most active user accounts in the total number of tweets ranges from 6.4 % in the Netherlands to 24% in Croatia and Switzerland. This skewed distribution of tweets thus results in an overrepresentation of the contributions and opinions of a limited number of media and cultural organizations and mainly public figures. Still, there is a very long tail of users in each country who contribute to cultural discussions on Twitter. On the other hand, some of the most active users have many followers, meaning that their influence may be decisive.

We found significant differences in the number of topics identified in each country. Although we applied the same strategy for topic modelling, the decisions were not taken based on the coherence and perplexity scores alone, but also – as is common in this type of analysis – on closer inspection of the various topic models and the interpretability of topics (Jacobi 2015). Team members in each individual INVENT country eventually determined how many topics were relevant in each period, which leaves room for interpretation and, potentially, human bias. Calculations of reliability help ensure that these potential biases are kept at a minimum. We also believe that the diversity in topics and topic numbers across our countries accurately reflects the empirical reality of cultural discussions on Twitter. It would in fact be highly unlikely for the same number and debates to occur across Europe.

Consequently, we see several opportunities for further research that could overcome some of these limitations. First, further research could scrape data from other social media that young people more commonly use to express their views of culture. For example, the topic of video games might appear in the analysis, as it did in some countries in our pilot study about the meanings of culture. Second, it would be interesting to compare the topics found in the analysis reported here with qualitative insights from the INVENT survey. For example, we asked citizens in all nine countries what first comes to their mind when they think about the word “culture”. Additionally, based on the survey, we could match the topics and social groups to delve into the specific social group’s cultural topic on their favourite online platforms.

Third, further analysis of the most active users could be carried out, taking into account their influence over other Twitter users. Future research could also compare characteristics of influencers in all nine countries and explore the significance of their cultural impact

#### 4. Policy implications

The analysis has identified topics shared by all nine countries as well country-specific issues. Among the expected topics are film, literature, arts, cultural events, policies, identity, social inequalities, economics, and spatial references. Even though these topics are familiar, their content may be specific to each country context. That means that a pan-European cultural policy may focus on spreading certain content related to these topics across all European countries. In addition, these policies may pinpoint the structural similarities among Europeans.

According to the country-specific topics, country uniqueness may be promoted as particular cultural topics that need to be known. For example, Catalan culture is a country-specific topic in Spain. This topic shows the distinctive culture which also differs from other autonomous regions of Spain. This result offers insight to the Spanish and European Commission to communicate better to reduce the political conflict. In other examples, we see that the *Me Too* movement is still a trending topic in Denmark, while the health care system is a discussion that is central to the UK. The Dutch

Twittersphere gives much attention to racism, crime, and deviant behaviour, while rape culture is prevalent in France, and politics is a salient subject in Serbia. Finnish Twitter users appear highly interested in journalism, while Switzerland is strongly influenced by language issues. Such context-based cultural topics could be of interest to cultural policymakers as well. In any case, they highlight the need to acknowledge and differentiate between cultural discussions on the regional and national level.

Last but not least, identifying cultural influencers' activities and exploring how they define culture and engage with their followers in online media would be highly relevant. This strategy could provide novel and useful insights for policymakers about relevant spokespersons, the spread of influencers' opinions on cultural policies, and potential ways in which policymakers and other actors in the field, can more effectively use social media for communication with the public.

## References

- 2Bfound. 2021. 'Social media trends 2021: de belangrijkste ontwikkelingen'. *2Bfound*. Retrieved 5 July 2021 (<https://2bfound.nl/blog/social-media-trends-ontwikkelingen/>).
- Agostino, Deborah. 2018. 'Can Twitter Add to Performance Evaluation in the Area of Performing Arts? Reflections from La Scala Opera House'. *The Journal of Arts Management, Law, and Society* 48(5):321–38. doi: 10.1080/10632921.2018.1431985.
- Alghamdi, Rubyyi, and Khalid Alfalqi. 2015. 'A Survey of Topic Modeling in Text Mining'. *International Journal of Advanced Computer Science and Applications* 6(1):147–53.
- Andersen, Kim, Jakob Ohme, Camila Bjarnøe, Mats Joe Bordacconi, Erik Albæk, and Claes Holger de Vreese. 2021. *Communication and Public Engagement Project*. University of South Denmark.
- Barde, Bhagyashree Vyankatrao, and Anant Madhavrao Bainwad. 2017. 'An Overview of Topic Modeling Methods and Tools'. Pp. 745–50 in *2017 International Conference on Intelligent Computing and Control Systems (ICICCS)*. Madurai: IEEE.
- Benoit, Kenneth, David Muhr, and Kohei Watanabe. 2021. *Stopwords: Multilingual Stopword Lists*.
- BfS, Bundesamt für Statistik. 2021a. 'Languages'. Retrieved 31 May 2021 (<https://www.bfs.admin.ch/bfs/en/home/statistics/population/languages-religions/languages.html>).
- BfS, Bundesamt für Statistik. 2021b. 'Omnibus 2019 (OMN2019): Erhebung Zur Internetnutzung'. Retrieved 31 May 2021 (<https://www.bfs.admin.ch/bfs/de/home/statistiken/kultur-medien-informationsgesellschaft-sport/erhebungen/omn2019.html>).
- Blei, David M., Andrew Y. Ng, and Michael I. Jordan. 2003. 'Latent Dirichlet Allocation'. *The Journal of Machine Learning Research* 3(null):993–1022.
- Boerema, Desiree. 2016. 'Opleidingsniveau Social Media gebruikers Nederland ★ iDees online'. *iDees online*. Retrieved 9 July 2021 (<https://www.ideesonline.nl/opleidingsniveau-social-media-gebruikers-nederland/>).
- Bogen, Katherine W., Kaitlyn Bleiweiss, and Lindsay M. Orchowski. 2019. 'Sexual Violence Is #NotOkay: Social Reactions to Disclosures of Sexual Victimization on Twitter.' *Psychology of Violence* 9(1):127–37. doi: 10.1037/vio0000192.
- Dabi, Frédéric, and Jean-Philippe Dubrulle. 2019. *Les Français et Twitter: Ifop Pour FashTweet*. FD/JPD. 116318. Paris: IFOP.
- Deckmyn, Original S. code by Richard A. Becker and Allan R. Wilks R. version by Ray Brownrigg Enhancements by Thomas P. Minka and Alex. 2018. *Maps: Draw Geographical Maps*.
- Doğu, Burak. 2020. 'Turkey's News Media Landscape in Twitter: Mapping Interconnections among Diversity'. *Journalism* 21(5):688–706. doi: 10.1177/1464884917713791.
- Duncombe, Constance. 2019. 'The Politics of Twitter: Emotions and the Power of Social Media'. *International Political Sociology* 13(4):409–29. doi: 10.1093/ips/olz013.
- Florini, Sarah. 2014. 'Tweets, Tweeps, and Signifyin': Communication and Cultural Performance on "Black Twitter"'. *Television & New Media* 15(3):223–37. doi: 10.1177/1527476413480247.
- FÖG, Forschungszentrum Öffentlichkeit und Gesellschaft. 2020. *Qualität Der Medien. Jahrbuch 2020*. Basel: Schwabe Verlag.
- Fownes, Jennifer R., Chao Yu, and Drew B. Margolin. 2018. 'Twitter and Climate Change'. *Sociology Compass* 12(6). doi: 10.1111/soc4.12587.
- Gutiérrez-Martín, Alfonso, and Alba Torrego-González. 2018. 'The Twitter Games: Media Education, Popular Culture and Multiscreen Viewing in Virtual Concourses'. *Information, Communication & Society* 21(3):434–47. doi: 10.1080/1369118X.2017.1284881.
- Harris Interactive. 2019. *Social Life 2019*. Paris: Harris Interactive.
- Haselmayer, Martin, and Marcelo Jenny. 2017. 'Sentiment Analysis of Political Communication: Combining a



- Dictionary Approach with Crowdcoding'. *Quality & Quantity* 51(6):2623–46. doi: 10.1007/s11135-016-0412-4.
- Highfield, Tim, Stephen Harrington, and Axel Bruns. 2013. 'TWITTER AS A TECHNOLOGY FOR AUDIENCING AND FANDOM: The #Eurovision Phenomenon'. *Information, Communication & Society* 16(3):315–39. doi: 10.1080/1369118X.2012.756053.
- Hutchins, Brett. 2011. 'THE ACCELERATION OF MEDIA SPORT CULTURE: Twitter, Telepresence and Online Messaging'. *Information, Communication & Society* 14(2):237–57. doi: 10.1080/1369118X.2010.508534.
- Jacobi, Carina, Wouter van Atteveldt, and Kasper Welbers. 2016. 'Quantitative Analysis of Large Amounts of Journalistic Texts Using Topic Modelling'. *Digital Journalism* 4(1):89–106. doi: 10.1080/21670811.2015.1093271.
- Katz-Gerro, Tally, Itay Greenspan, Femida Handy, and Yaara Vered. 2020. 'Environmental Behavior in Three Countries: The Role of Intergenerational Transmission and Domains of Socialization'. *Journal of Environmental Psychology* 71:101343. doi: 10.1016/j.jenvp.2019.101343.
- Kearney, Michael W. 2019. 'Rtweet: Collecting and Analyzing Twitter Data'. *Journal of Open Source Software* 4(42):1829. doi: 10.21105/joss.01829.
- Kemp, Simon. 2021. *Digital 2021: Croatia*.
- Kim, Jinju. 2021. ComTxt: Twitter text analysis. *GitHub repository*: <https://github.com/BeaJJ/ComTxt>.
- Kim, Jinju, and Jordi López Sintas. 2021. 'Social TV Viewers' Symbolic Parasocial Interactions with Media Characters: A Topic Modelling Analysis of Viewers' Comments'. *Social Sciences & Humanities Open* 3(1):100129. doi: 10.1016/j.ssaho.2021.100129.
- Kohvakka, Rauli, and Kaisa Saarenmaa. 2019. 'WhatsApp suosituin – some on suomalaisten arkea iän mukaan vaihdellen | Tieto&trendit'. Retrieved 5 July 2021 (<https://www2.tilastokeskus.fi:443/tietotrendit/artikkelit/2019/whatsapp-suosituin-some-on-suomalaisten-arkea-ian-mukaan-vaihdellen/>).
- Komorowski, Marlen, Tien Do Huu, and Nikos Deligiannis. 2018. 'Twitter Data Analysis for Studying Communities of Practice in the Media Industry'. *Telematics and Informatics* 35(1):195–212. doi: 10.1016/j.tele.2017.11.001.
- Kristensen, Nete Nørgaard, and Unni From. 2018. 'Cultural Journalists on Social Media'. *MedieKultur: Journal of Media and Communication Research* 34(65):76–97. doi: 10.7146/mediekultur.v34i65.104488.
- Laitinen, Mikko, Jonas Lundberg, Magnus Levin, and Rafael Martins. 2018. 'The Nordic Tweet Stream: A Dynamic Real-Time Monitor Corpus of Big and Rich Language Data'.
- MacDowall, Lachlan John, and Poppy de Souza. 2018. '"I'd Double Tap That!!": Street Art, Graffiti, and Instagram Research'. *Media, Culture & Society* 40(1):3–22. doi: 10.1177/0163443717703793.
- Meraz, Sharon, and Zizi Papacharissi. 2013. 'Networked Gatekeeping and Networked Framing on #Egypt'. *The International Journal of Press/Politics* 18(2):138–66. doi: 10.1177/1940161212474472.
- Mihelj, Sabina, Adrian Leguina, and John Downey. 2019. 'Culture Is Digital: Cultural Participation, Diversity and the Digital Divide'. *New Media & Society* 21(7):1465–85. doi: 10.1177/1461444818822816.
- Mimno, David. 2013. *Mallet: A Wrapper around the Java Machine Learning Tool MALLET*.
- Newgarden, Kristi. 2009. 'Annotated Bibliography – Twitter, Social Networking and Communities of Practice'. *The Electronic Journal for English as a Second Language* 13(2).
- Oosterveer, Danny. 2020. 'Social Media in Nederland 2020: TikTok Is Domein van Kinderen, Uittocht Jongeren Op Facebook'. *Marketingfacts*. Retrieved 5 July 2021 (<https://www.marketingfacts.nl/berichten/social-media-in-nederland-2020>).
- Ovadia, Steven. 2009. 'Exploring the Potential of Twitter as a Research Tool'. *Behavioral & Social Sciences Librarian* 28(4):202–5. doi: 10.1080/01639260903280888.
- Reuters Institute. 2020. '¿Podría llegar el coronavirus chino a España?' *Ultima Hora*. Retrieved 5 July 2021 (<https://www.ultimahora.es/noticias/internacional/2020/01/22/1136307/coronavirus-china-puede-llegar-espana.html>).

- Schoch, David. 2021. *Get Tweets from the Academic Research Product Track*.
- Schwartz, Julie, and Sofie Scheutz. 2021. 'TikTok er blevet voksen'. *DR*. Retrieved 5 July 2021 (<https://www.dr.dk/om-dr/fakta-om-dr/medieforskning/medieudviklingen/2020/tiktok-er-blevet-voksen>).
- da Silva, Nádia F. F., Eduardo R. Hruschka, and Estevam R. Hruschka. 2014. 'Tweet Sentiment Analysis with Classifier Ensembles'. *Decision Support Systems* 66:170–79. doi: 10.1016/j.dss.2014.07.003.
- Slots-og Kulturstyrelse. 2020. *Mediernes Udvikling i Danmark. Sociale Medier 2020 –Brug, Indhold og Relationer [Media Development in Denmark. Social Media 2020 -Use, Content and Relationships]*. Mediernesudvikling.slks.dk.
- Smith, Robin S. 2005. 'Theories of Digital Participation'. in *GIS for Sustainable Development*, edited by M. Campagna. CRC Press.
- Statista. 2016. 'Netherlands: Age Distribution of Twitter Users 2016'. *Statista*. Retrieved 5 July 2021 (<https://www.statista.com/statistics/579598/age-distribution-of-twitter-users-in-the-netherlands/>).
- Stever, Gayle S., and Kevin E. Lawson. 2013. 'Twitter as a Way for Celebrities to Communicate with Fans: Implications for the Study of Parasocial Interaction.' *North American Journal of Psychology* 15(2):339–54.
- Tankovska, H. 2021. 'Netherlands: Social Media Millennials 2020'. *Statista*. Retrieved 5 July 2021 (<https://www.statista.com/statistics/697047/share-millennials-actively-using-social-media-the-netherlands-by-social-network/>).
- Tornes, Adam, and Leanne Trujillo. 2021. 'Enabling the Future of Academic Research with the Twitter API'. *Twitter*. Retrieved 5 July 2021 ([https://blog.twitter.com/developer/en\\_us/topics/tools/2021/enabling-the-future-of-academic-research-with-the-twitter-api](https://blog.twitter.com/developer/en_us/topics/tools/2021/enabling-the-future-of-academic-research-with-the-twitter-api)).
- Trilla, Antoni. 2020. '¿Está preparada España si llegan casos del coronavirus?' *La Vanguardia*. Retrieved 5 July 2021 (<https://www.lavanguardia.com/ciencia/20200122/473046473766/coronavirus-espana-riesgos-epidemia-infeccion-wuhan.html>).
- Twitter. 2021. 'Twitter API for Academic Research | Products'. Retrieved 8 July 2021 (<https://developer.twitter.com/en/products/twitter-api/academic-research>).
- Vergeer, Maurice. 2015. 'Twitter and Political Campaigning: Twitter and Political Campaigning'. *Sociology Compass* 9(9):745–60. doi: 10.1111/soc4.12294.
- Weingartner, Sebastian. 2020. 'Digital omnivores? How digital media reinforce social inequalities in cultural consumption'. *New Media & Society*, online first. doi:10.1177/1461444820957635.
- Wood, Megan M., and Linda Baughman. 2012. 'Glee Fandom and Twitter: Something New, or More of the Same Old Thing?' *Communication Studies* 63(3):328–44. doi: 10.1080/10510974.2012.674618.
- Yle Uutiset. 2020. 'Suomen ensimmäinen koronavirustartunta varmistui, THL: Tapaus valitettava, mutta ei odottamaton – Yle seurasi hetki hetkeltä'. *Yle Uutiset*. Retrieved 5 July 2021 (<https://yle.fi/uutiset/3-11181717>).
- Zhu, Qile, Zheng Feng, and Xiaolin Li. 2018. 'GraphBTM: Graph Enhanced Autoencoded Variational Inference for Biterm Topic Model'. Pp. 4663–72 in *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*. Brussels, Belgium: Association for Computational Linguistics.

## Appendix

### Tables A1 and A2

**Table - A1.** Conjugations of the word “kulttuuri” used to retrieve data from Finnish Twitter and the corresponding grammatical case in singular

Word	Grammatical case in singular
Kulttuuri	Nominative
Kulttuurin	Genitive
Kulttuurია	Partitive
Kulttuurina	Essive
Kulttuuriksi	Translative
Kulttuurissa	Inessive
Kulttuurista	Elative
Kulttuuriin	Illative
Kulttuurilla	Adessive
Kulttuurille	Allative

**Table - A2.** Descriptive information on each country’s dataset

Country	Tweet N 2019/2020	User N 2019/2020	Start Covid19 period	Most active Users 2019	Most active Users 2020	Share of tweets of top 10 users 2019/2020
<b>Croatia</b>	6421 / 5913	946 / 951 (total 1,897)	1/1/2020	News media Private users City of Rijeka European Capital of Culture 2020	News media Private users City of Rijeka European Capital of Culture 2020	33% / 24%
<b>Denmark</b>	32,951 / 21,335	5,783 / 5,344	27/2/2020	Mostly related to news media Movie industry	Mostly related to news media	38% / 19%
<b>Finland</b>	36,814 / 31,552	- / - (total 14,805)	29/1/2020	Mostly related to news media Admin. bodies	Mostly related to news media Admin. bodies	18% (both periods)
<b>France</b>	26,248 / 25,687	-	1/1/2020	News media Admin. bodies Ordinary citizens	News media Admin. bodies Ordinary citizens	9.5% / 8%
<b>Netherlands</b>	92,322 / 93,521	19,499 / 19,504 (total 29,859)	31/1/2020	Mostly news media	Mostly news media Individual users	7% / 7%
<b>Serbia</b>	30,202 / 30,514	4,925 / 5,185 (total 7,690)	1/1/2020	Mostly news media One Film director	Mostly news media	10% / 11%
<b>Spain</b>	44,124 / 34,784	14,783 / 12,103 (total 23,836)	22/1/2020	Mainly academics, authors, librarians	Mainly academics, authors, librarians	11% / 3%
<b>Switzerland</b>	42,031 / 40,398	7,830 / 8,578 (total 13,070)	1/1/2020	Mostly news media	Mostly news media	30% / 24%
<b>UK</b>	55,108 / 47,727	10,289 / 10,969 (total 21,258)	31/1/2020	Famous figures (comedians, politician, royals) Individuals working for NHS	Famous figures (comedians, politician, royals) Individuals working for NHS	12% / 14%

## R Script used to download tweets utilizing Twitter Academic research product track

```
library(jsonlite)
library(httr)
library(tidyverse)

# Get the function to connect Twitter V2 API from GitHub
source("https://gist.githubusercontent.com/schochastics/1ff42c0211916d73fc98ba8ad0dcb261/ra
aw/040422b5e1378ef4c30150d4927a3991f53bc922/get_tweets.R")

# Use bearer token credential for academic app
bearer_token_academic_app <- "Our private bearer token"

# Create data frame containing the start and the end day of the request
dates_raw <- seq.POSIXt(from = ISOdate(2019,1,1), to = ISOdate(2020,12,31), by = "day")
dates_fromDate <- paste0(gsub(" ", "T", dates_raw), "Z")
dates <- data.frame(dates_fromDate = dates_fromDate,
                    dates_toDate = c(dates_fromDate[-1], NA)[-length(dates_fromDate),])

# Initiate empty list
tweets_list <- list()

# loop over dates, use get_tweets() to obtain max 500 tweets that includes #kulttuuri from every
interval
for(i in seq_len(nrow(dates))){
  tweets_list[[i]] <- get_tweets(
    q = "culture, place_country:GB",
    n = 500,
    start_time = dates[i,1],
    end_time = dates[i,2],
    token = bearer_token_academic_app)
  Sys.sleep(15) # Wait between requests
}
```

## List of Tables

<b>Table 1.</b> Internet access, frequency of use and participation in social networks in Europe (%), 2011 and 2019....	6
<b>Table 2.</b> Scraping methods and number of tweets analyzed for each country.....	9
<b>Table 3.</b> Allocation of all topics related to culture identified in all countries to topic groups.....	127
<b>Table A1.</b> Conjugations of the word “kulttuuri” used to retrieve data from Finnish Twitter.....	139
<b>Table A2.</b> Descriptive information on each country’s dataset.....	139
<b>Table - HR - 1.</b> Top 10 locations in 2019 and 2020 .....	13
<b>Table - HR - 2.</b> Users with most followers and users with most tweets about culture in 2019 and 2020.....	14
<b>Table - HR - 3.</b> The most popular hashtags except #kultura (#cultural)’ in 2019 and 2020.....	16
<b>Table - HR - 4.</b> Keywords for Twitter.....	17
<b>Table - DK - 1.</b> Comparison of Danish culture-related user accounts before and during the Covid19 pandemic with regard to the number of followers and tweets about culture.....	26
<b>Table - DK - 2.</b> Hashtag statistics 2019/2020 Denmark. ....	29
<b>Table - DK - 3.</b> Top 10 hashtag frequency and percentage Denmark. ....	29
<b>Table - DK - 4.</b> Top 15 Danish- and English language keyword frequency and percentages, pre-Covid19 period	30
<b>Table - DK - 5.</b> Top 15 Danish- and English language keyword frequency and percentages, Covid19 period.....	31
<b>Table - FI - 1.</b> Fifteen most common hashtags and keywords before and during Covid-19 . ....	41
<b>Table - FR - 1.</b> Most active city in terms of posted tweets (%) .....	49
<b>Table - FR - 2.</b> Most active profiles on Twitter: number of hashtags and tweets (2019-2020) .....	50
<b>Table - FR - 3.</b> Twenty most posted hashtags in 2019 .....	51
<b>Table - FR - 4.</b> Twenty most posted hashtags in 2020 .....	52
<b>Table - NL - 1.</b> Top 10 locations of tweets related to culture in the Netherlands in 2019 and 2020.....	61
<b>Table - NL - 2.</b> Most active users in 2019 and 2020 according to number of tweets about culture.....	62
<b>Table - NL - 3.</b> Most popular user accounts tweeting about culture in 2019 and 2020 .....	63
<b>Table - NL - 4.</b> Use of hashtags and twenty most common hashtags in 2019 and 2020 .....	66
<b>Table - NL - 5.</b> Twenty most common keywords in 2019 and 2020.....	67
<b>Table - RS - 1.</b> Users with most published tweets about culture and most followers in 2019 and 2020 .....	80
<b>Table - RS - 2.</b> Comparison of 10 most popular keywords and hashtags between the years.....	81
<b>Table - ES - 1.</b> Hashtags for Twitter and Keywords for twitter .....	91
<b>Table - CH - 1.</b> Top 10 locations of tweets related to culture in Switzerland in 2019 and 2020.....	102
<b>Table - CH - 2.</b> Top 10 users by number of tweets related to culture in Switzerland in 2019 and 2020 .....	103
<b>Table - CH - 3.</b> Top 20 keywords and hashtags in tweets related to culture in Switzerland in 2019. ....	104
<b>Table - CH - 4.</b> Top 20 keywords and hashtags in tweets related to culture in Switzerland in 2020. ....	106
<b>Table - UK - 1.</b> Most active profiles using the key-word “Culture” .....	113
<b>Table - UK - 2.</b> Top 10 Hashtags and Keywords .....	114
<b>Table - UK - 3.</b> Top 10 Hashtags and Keywords .....	116

## List of Figures

<b>Figure - HR - 1.</b> Geographical location of tweets in Croatia in 2019-2020 .....	13
<b>Figure - HR - 2.</b> Top hashtags network of tweets in Croatia in 2019 and 2020 .....	16
<b>Figure - HR - 3.</b> Top hashtags and keywords network of tweets in Croatia in 2019 .....	18
<b>Figure - DK - 1.</b> Steps taken in data cleaning and preparation. ....	24
<b>Figure - DK - 2.</b> Geographical location of Danish- and English-language tweets in Denmark.....	24
<b>Figure - DK - 3.</b> Network of mentioned user accounts in Denmark, pre-Covid19 period .....	27
<b>Figure - DK - 4.</b> Network of mentioned user accounts in Denmark, Covid19 period.....	27
<b>Figure - DK - 5.</b> LEGO_Group status post 'Grand Piano'. ....	28
<b>Figure - DK - 6.</b> Hashtag network Denmark, Covid19 period. ....	30
<b>Figure - DK - 7.</b> Hashtag network Denmark, pre-Covid19 period. ....	30
<b>Figure - DK - 8.</b> TOP30 Danish- and English language keyword network per period. ....	32
<b>Figure - DK - 9.</b> Development of topic prevalence over time, pre-Covid19 period .....	33
<b>Figure - DK - 10.</b> Development of topic prevalence over time, Covid19 period . ....	36
<b>Figure - FI - 1.</b> Geographical location of tweets in Finland .....	40
<b>Figure - FI - 2.</b> Network of hashtags in 2019. ....	42
<b>Figure - FI - 3.</b> Network of keywords in 2019 .....	42
<b>Figure - FI - 4.</b> Network of hashtags in 2020 .....	43
<b>Figure - FI - 5.</b> Network of keywords in 2020 .....	43
<b>Figure - FI - 6.</b> Wordclouds corresponding to each topic found in the 2019 tweets dataset. . ....	44
<b>Figure - FI - 7.</b> Wordclouds corresponding to each topic found in the 2020 tweets dataset . ....	45
<b>Figure - FR - 1.</b> Twitter users by age (2019) .....	48
<b>Figure - FR - 2.</b> Most active city in terms of posted tweets (2019-2020).....	49
<b>Figure - FR - 3.</b> Network of fifty hashtags -2019 .....	51
<b>Figure - FR - 4.</b> Hashtags network -2020 Twitter's topics .....	52
<b>Figure - FR - 5.</b> Best topic number .....	53
<b>Figure - FR - 6.</b> Topics distribution 2019-2020 .....	58
<b>Figure - FR - 7.</b> Occurance of the topics throughout the year .....	59
<b>Figure - NL - 1.</b> Geographical location of tweets in the Netherlands .....	61
<b>Figure - NL - 2.</b> Network of mentioned user accounts in the Netherlands in 2019 / pre-Covid-19.....	64
<b>Figure - NL - 3.</b> Network of mentioned user accounts in the Netherlands in 2020 / during Covid19 .....	64
<b>Figure - NL - 4.</b> Most liked tweets in the Netherlands, 2019 and 2020 .....	65
<b>Figure - NL - 5.</b> Top 50 Hashtags network in the Netherlands, 2019 and 2020 .....	67
<b>Figure - NL - 6.</b> Top 40 Keyword network of tweets in the Netherlands, 2019 and 2020 .....	68
<b>Figure - NL - 7.</b> Best topic number k in 2019 and 2020.....	68
<b>Figure - NL - 8.</b> Topic prevalence and development in the Netherlands, 2019 and 2020 .....	76
<b>Figure - RS - 1.</b> Steps taken in data cleaning and preparation.. ....	78
<b>Figure - RS - 2.</b> Geographical location of tweets in Serbia.....	79
<b>Figure - RS - 3.</b> Network of mentions in 2019.....	80
<b>Figure - RS - 4.</b> Network of mentions in 2020.....	81
<b>Figure - RS - 5.</b> Networks of top 20 keywords in 2019 and 2020.....	82
<b>Figure - RS - 6.</b> Networks of top 20 hashtags in 2019 and 2020.....	82
<b>Figure - RS - 7.</b> Topics that dominated Twitter discussions about culture in Serbia (2019) .....	83
<b>Figure - RS - 8.</b> Distribution of tweets related to each of the topics (2019) .....	84
<b>Figure - RS - 9.</b> Topics that dominated Twitter discussions about culture in Serbia (2020) .....	85
<b>Figure - RS - 10.</b> Distribution of tweets related to each of the topics (2020) .....	87



<b>Figure - ES - 1.</b> The user with the highest number of followers.....	89
<b>Figure - ES - 2.</b> Tweets distribution map in Spain before and during COVID .....	90
<b>Figure - ES - 3.</b> Top hashtags and keywords network of tweets in Spain before and during COVID .....	92
<b>Figure - ES - 4.</b> Topics trend per month before COVID 19 .....	95
<b>Figure - ES - 5.</b> Topics trend per month during COVID 19.....	97
<b>Figure - ES - 6.</b> 10 topic Distribution in Spain before (2019) and during COVID19 (2020).....	98
<b>Figure - CH - 1.</b> Number of tweets related to culture in Switzerland in 2019 and 2020.....	101
<b>Figure - CH - 2.</b> Location of tweets related to culture in Switzerland in 2019 and 2020 .....	101
<b>Figure - CH - 3.</b> Network of top 30 keywords of tweets related to culture in Switzerland in 2019.....	105
<b>Figure - CH - 4.</b> Network of top 30 hashtags of tweets related to culture in Switzerland in 2019. ....	105
<b>Figure - CH - 5.</b> Network of top 30 keywords of tweets related to culture in Switzerland in 2020. ....	107
<b>Figure - CH - 6.</b> Network of top 30 hashtags of tweets related to culture in Switzerland in 2020. ....	107
<b>Figure - CH - 7.</b> Topics identified in tweets related to culture in Switzerland in 2019.....	109
<b>Figure - CH - 8.</b> Topics identified in tweets related to culture in Switzerland in 2020.....	110
<b>Figure - UK - 1.</b> Tweets distribution map in UK before and during Covid-19.....	114
<b>Figure - UK - 2.</b> Top 20 Hashtags Network - 2019.....	115
<b>Figure - UK - 3.</b> Top 20 Keywords Network - 2019.....	115
<b>Figure - UK - 4.</b> Top 20 Hashtags Network - 2020.....	117
<b>Figure - UK - 5.</b> Top 20 Keywords Network - 2020.....	117
<b>Figure - UK - 6.</b> Occurance of the topics throughout 2019 .....	119
<b>Figure - UK - 7.</b> Occurance of the topics throughout 2020 .....	122
<b>Figure - UK - 8.</b> Topics distribution 2019-2020 .....	122