



inDICES

Measuring the Impact of Digital Culture

Deliverable 1.4

Report on data gathering v2



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 870792.

The sole responsibility for the content of this publication lies with the authors. It does not necessarily represent the opinion of the European Union. Neither the EASME nor the European Commission is responsible for any use that may be made of the information contained therein.

The research leading to these results has received funding from the European Commission's Horizon 2020 Programme (H2020-DT-GOVERNANCE-13-2019) under grant agreement n° 870792.



inDICES

Measuring the Impact
of Digital Culture

D1.4 – Report on data gathering v2

Version 1.0

Grant Agreement Number: 870792
Project Acronym: inDICES
Project Title: Measuring the impact of Digital CultureE
Funding Scheme: H2020-DT-GOVERNANCE-13-2019
Project Coordinator Name, Title and Organisation: Simonetta Buttò, Director of the Central Institute for the Union Catalogue of the Italian Libraries (ICCU)
Tel: +39 06 49210425
E-mail: simonetta.butto@beniculturali.it
Project Website: www.indices-culture.eu

Authors:
Maria Tartari, FBK-IULM
Federico Pilati, FBK-IULM
Katinka Böhm, WLT
Pier Luigi Sacco, FBK
Manlio De Domenico, FBK

Contributing Partners:
Oriol Artime, FBK
Riccardo Gallotti, FBK
Flavia Massara, ICCU

Reviewers:
Rasa Bočytė, NISV
Marzia Piccininno

Document History

- 18.06.2021 – Document Created from Template
- 19.08.2021 - Index definition
- 15.09.2021 - internal review and proofreading
- 28.09.2021 - response to internal reviewers' comments and redefinition

Table of Contents

- 1. Executive summary**
- 2. Objectives**
- 3. Data gathered to Month 21**
 - 3.1. Methodology
 - 3.2. Museums of Art with Virtual Tour (World)
 - 3.2.1. Instagram
 - 3.2.2. Facebook
 - 3.3. Most Visited Museums in Europe
 - 3.3.1. Instagram
 - 3.3.2. Facebook
 - 3.4. Libraries
 - 3.4.1. Instagram
 - 3.4.2. Facebook
 - 3.5. Archives
 - 3.5.1. Instagram
 - 3.5.2. Facebook
 - 3.6 Update on Crawled Web Content Datasets
 - 3.6.1 CHI Knowledge Graph - GLAM Entities
- 4. Preliminary observations**
- 5. Usefulness and Limits**
- 6. Future Plans**
- 7. Conclusions**

1. Executive summary

This deliverable describes the data collection processes and status, and the datasets gathered in the months M15-M21 by WP1. In this period, the data analysis focused on CHIs' social networks users' behavior, especially within the Covid-19 wave of forced digitization¹.

The inDICES data collection processed and/or stored consists of:

- building and analysing a set of case studies, composed by lists of Facebook and Instagram accounts, per macro-sectoral areas of cultural and creative institutions. The sectors' case studies that have been identified and analysed are:
 - European National Libraries sector;
 - European Archival Institutes sector;
 - Most Visited European Museums;
 - Museums that employ the Virtual Tour tool.
- outlining repositories of metadata enriched with documents gathered from the Web and from social media sources and extending the Visual Analytics Dashboard knowledge graph with entities specific to the inDICES use cases such as the GLAM institutes.

This data was gathered with the purpose to:

- detect trends regarding the levels of digital cultural participation of the most used social network sites, with a focus on the impact of Covid-19 wave of forced digitization on users' behavior and CHIs behavior;
- enrich inDICES repositories with sources gathered from the Web and from social media.

2. Objectives

The objective of this deliverable, which is the second inDICES data gathering periodical report, is to describe the status of data gathering activities, with specific information on the methodology, quality, reliability and accessibility of the information gathered during the months 15-21 by WP1.

The report is aimed at describing the results achieved by this work package in this specific time window (2019-2021), as well as to give an overview of the preliminary data analysis conducted on the behavioural patterns of the case studies' users, namely the users of groups of CHIs, selected by expert partners of the inDICES project, differentiated per Cultural and Creative Sector (CCS).

The aim of the present data collection is to proceed with an analysis of the European sectors of the National Libraries, the National Archives and two selected groups of museums, namely the Most Visited European Museums and the Museums that provide the tool of the Virtual Tour at the international level. The main goal is to detect the trends regarding the levels of digital cultural participation of the users of the different sectors' institutions. This analysis has been conducted with a specific focus on a temporal window that can help make a comparison of the relationship between users and CHIs' digital platforms before, during and after the pandemic.

This is the second of four data gathering periodic reports that outlines the status of the first phase of the inDICES project, namely the data gathering activity. The data gathering strategy of the M15-21 of work has been devised, on one side, to collect new data on CHIs digital platforms' users behaviour; on the other side, to enrich inDICES repositories with sources gathered from the Web and from social media. In particular, strategies of acquisition of relevant data through Facebook and Instagram have been defined and implemented.

These resources are likely to be useful for a wide range of decision makers, researchers and practitioners in cultural and creative sectors. The reports that

contain aggregated data per CCS will be openly available in the Repository of the inDICES Observatory Platform.

Our data analyses address the following target groups:

- Policy Makers, who can be informed about the state-of-the-art of the virtual relation between Cultural Heritage Institutions and their users, and supported in the definition or improvement of policies of active participation via social platforms and in making consequential budget allocation choices;
- Cultural Heritage/Creative Sector Practitioners, who aim to attract and understand user experiences for their work, to make use of tools and resources for their professional development, can be supported and guided in better appreciating new ways to spark active participation, to develop digital strategies, tools or practices favoring bottom-up and collective co-creation;
- Researchers and Special Interest Groups searching for relevant data and information on case-studies about the state-of-the-art of the digitization of Cultural Heritage Institutions per sector.

In order to express how this data collection and analysis support inDICES objectives and activities, it is important to underline that the analyses conducted through the study of Facebook and Instagram platforms relative to cultural sectors such as museums and libraries are extremely relevant for three main reasons, among others. Firstly, thanks to the study of the temporal evolution of different metrics regarding the sectors taken into consideration, it will be possible to describe how the cultural sectors studied for inDICES have used digital tools as social media; secondly, through the analyses carried out, it will be possible to outline the impact that communication on social media has had on the fruition of cultural content shared through these platforms; on the basis of these two results, in the further period of analysis it will be possible to compare which types of content are most engaging for a real experience of active participation of online users of the cultural sector.

Regarding the originality of the work, we can say that this first set of analysis has advanced inDICES knowledge of the digitization processes that are occurring in the field of CHI; in fact, until now, it was not possible to find any equally detailed analysis in depth over a period of time as long as that was analysed. We can therefore say that, despite being only the beginning of a series of analyses regarding the use of social networks and the study of the digitization of the cultural sector, the preliminary results brought in this deliverable are central to building a more informed research and development strategy both within inDICES project and the European cultural sector. Finally, the open question regarding the measurability of trends in digital cultural production, consumption and behaviors can then be more critically analysed thanks to the following report.

3. Data gathered to Month 21

In these last months of work [M15-21], WP1 kept on collecting data to accomplish the inDICES objectives of carrying out a close observation of the behaviour and competitiveness of CHIs wanting to be integrated into the Digital Single Market (DSM) and to develop a constant dialogue with their target audiences via digital platforms, including professionals and entrepreneurs from the digital cultural and creative content sectors, in order to implement good practices of digital cultural active participation.

According to the early results that emerged from the first set of data analysis on CHIs case studies, we are going to offer some first suggestions for policy recommendation guidelines.

3.1 Methodology

As reported in D1.3, in order to map the current situation about digital cultural participation of CHIs users, during the first 12 months of the inDICES project, WP1 gathered a large amount of data from online sources, with special attention to social networks. During the months 12-21 of activity, WP1 proceeded with the building of a first set of case studies (lists of CHIs) per macro-sectoral areas of cultural and creative institutions (namely per CCS). In the next few months, European Theatres, European Archeological sites, Fashion GLAMs, a case study list regarding the macro-sector of the Audiovisual, in order to observe the institution of the GLAM sector and other institutions pertaining to cultural and creative domains as classified by the literature (NESTA, 2006; Thorsby, 2008; Santagata, 2009; KEA, 2019)¹

The sectors' case studies that have been identified and analysed to month 21 are:

- European National Libraries sector
- European Archives Institutes sector
- Most Visited Museums of Art in Europe
- Art Museums with Virtual Tour (World)

The lists have been drafted by WP1 internal partners and experts; each list has been built according to criteria that are explained in the introduction of each related section, and contain a number between 30 and 50 institutes per sector that are active in Instagram and/or Facebook. Instagram and Facebook have been chosen as the two most widely used and demographically heterogenous social platforms; moreover, the latest is the World's most used social platform. Indeed, according to the “We are social” report (2021)², at the European level, the number of active social media users compared to the total population is between 79 and 72 %. Moreover, both the comparison and the juxtaposition of the two social platforms is meaningful because the percentage of users overlapping (ages 16 to 64) is between 85.5 and

¹ NESTA. 2006. *Creating Growth How the UK Can Develop World Class Creative Business*. NESTA Research Report. NESTA, London; Thorsby, D. (2008). *Modelling the cultural industries*. *International journal of cultural policy*, 14(3), 217-232.; Santagata, W. (2009). *Libro bianco sulla creatività: per un modello italiano di sviluppo*. EGEA spa.; KEA new model (2019), see: <https://keanet.eu/opinions/culture-nowhere-or-everywhere/> .

² <https://wearesocial.com/digital-2021>

74.8%. The average age of Facebook users is slightly higher. Data has been collected from July 2019 to July 2021, aggregated per month.

As a premise, it is important to underline that, unlike museums, most archives and libraries do not have physical exhibitions or at least that's not their main goal (e.g. they might have smaller spaces to hold temporary exhibitions/events). Hence their usage of social media differs from that of museums.

WP1 analyses unpacked data in terms of the following metadata:

- Cultural and Creative Sector;
- Country;
- Month of publication;
- Type of relation between users and CHI as to cultural production impact;
- Form of content sharing (Photos, Links, Statuses, Facebook Videos, Facebook Live, YouTube Videos, Other Videos, Albums, IG Videos, IGTV).

To go into detail of the social platforms taken into account, the data analysis contains the following metrics, which are a selected list of social media analytics. In the present deliverable, as reported in the Preliminary Observations section, we considered only the most meaningful metrics that could provide consistent and useful information for the project's objectives.

Instagram:

Account, Codename, URL, Total Interactions, Likes, Comments, Views, All Interaction Rate, Albums, Photos, Video, IGTV, Total Posts, Album Posts, Photo Posts, Video Posts, IGTV Posts, Posts Per Day, Album Likes, Photo Likes, Video Likes, IGTV Likes, Album Comments, Photo Comments, Video Comments, IGTV Comments, Album Video Views, Video Video Views, IGTV Video Views, Followers, Follower Growth, Follower Growth %.

Facebook:

Page, Codename, URL, Total Interactions, Likes, Comments, Shares, Owned Post Views, Owned Views from Shares, Owned Total Views, Percentage Views from

Owned Posts, Views on Shared Posts, Views While Live, Video Time, 3-Min+ Videos, Loves, Wows, Hahas, Sads, Angrys, Cares, All Reactions, All Interaction Rate, Photos, Links, Statuses, Facebook Videos, Other Videos, Total Posts, Photo Posts, Link Posts, Status Posts, Owned Video Posts, Shared Video Posts, Other Video Posts, Posts Per Day, Page Likes, Page Growth, Page Growth %, Page Followers, Page Follower Growth, Page Follower Growth %, Interaction Rate Calculated.

The data analysis presented in this deliverable has been conducted by the *Fondazione Bruno Kessler* research team (WP1) using the analytical tool CrowdTangle (from 3.2 to 3.5.2 section), a content discovery and analytics platform designed to provide content creators with the data and insights they need. CrowdTangle Intelligence gives researchers a way to monitor the performance of a social channel over time, as well as to directly benchmark it against other accounts. Long-term performance figures can help publishers detect overall trends and more easily analyze what content is working and what is not. Intelligence allows us to monitor up to 100 social accounts (on any of our platforms including Facebook, Instagram, Twitter and Reddit) and see overall account-level statistics over time with graphs and charts. This tool then allows an easy comparison of the accounts next to each other and to export the whole analytics for further use.

Next to the data gathered and analyzed via CrowdTangle, WP1 partner *WebLyzard* extended their content ingestion pipeline to process content gathered from Web and Social Media sources in the CCS. It is then stored as metadata-enriched documents in the WLT repositories, indexed, and made available through the Visual Analytics Dashboard. Furthermore, the WLT knowledge graph that stores both factual and semantic information and is part of the data enrichment process, is being extended with entities specific to the inDICES use cases such as European GLAM institutes (section 3.6).

3.2 Art Museums with Virtual Tour (World)

As part of the inDICES data analysis activity, we decided to focus upon two comparable case studies in the Museums sector: the case of the Most Visited Museums of Europe in 2019 and the case of the Art Museums with Virtual Tour.

The Virtual Tour of museums is the digital tool par excellence that allows digital users to interact virtually with the spaces and collections of a museum or cultural institution. Virtual Tours are an essential resource to promote and strengthen the engagement of global visitors, including those who are not able to access the museum premises physically. On one hand, there is the necessity to address regular visitors with loyalty marketing initiatives, on the other, museums wish to enlarge the number of their visitors on a global scale. Online exhibition platforms with virtual tours have been instrumental to this purpose, accelerating a process that has been ongoing for the last two decades (Resta et al., 2021)³. The objectives of the data collection and analysis of this case study are the observation of the initiatives of the selected museums and of the choices of their users during the pandemic period compared to the previous year. The natural comparison in terms of virtual presence, engagement and user participation is with the case-study we presented at the 3.3 section, namely the “Most Visited Museums of Art in Europe”. On the basis of the insights derived from the analysis, we can thus offer strategy recommendations for CHIs that may influence future investments in tools for increasing online presence and communication, the willingness to explore innovative digital techniques, and more generally the preparedness to develop full-fledged digital strategies as a key pillar of museum strategy.

³ Resta, G., Dicuonzo, F., Karacan, E., & Pastore, D. (2021). The impact of virtual tours on museum exhibitions after the onset of covid-19 restrictions: visitor engagement and long-term perspectives. *SCIRES-IT-SCientific RESearch and Information Technology*, 11(1), 151-166.

In this case study, we also include non-European cultural institutions that offer virtual tours. This is due to the need to span a large enough number of institutions to form a meaningful sample of analytical interest. We tried to maintain the samples used for the Facebook and Instagram platforms as close as possible, provided that not all the museums selected had both comparably active Instagram and Facebook pages that could be amenable to data analysis.

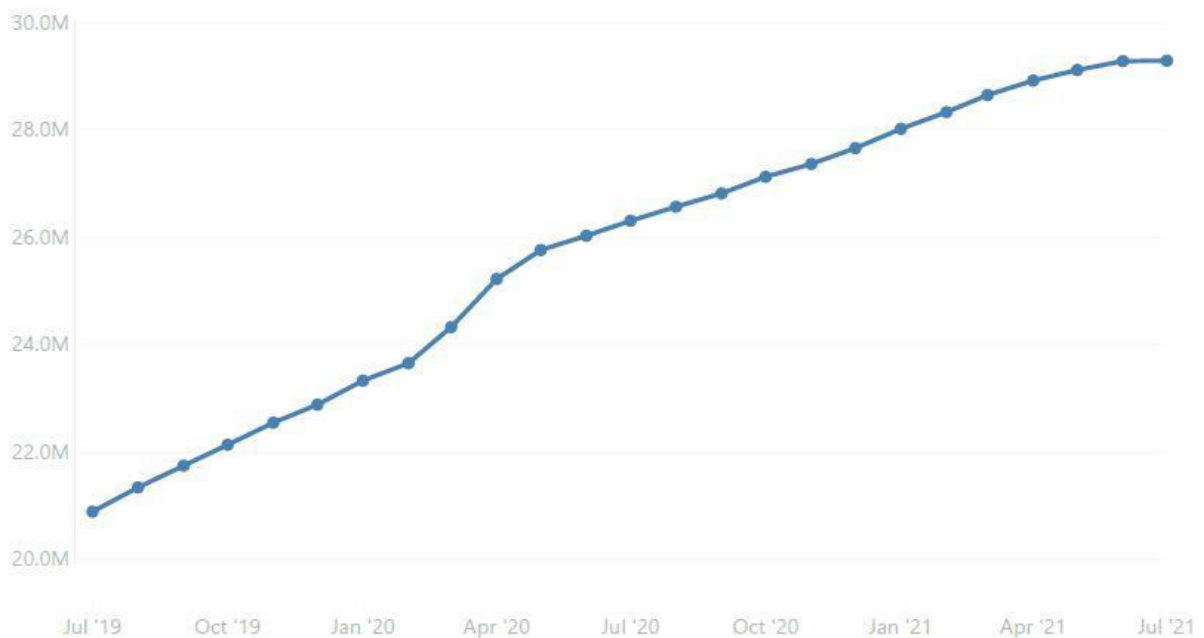
The list for this case study is as follows:

Van Gogh Museum, Museo Nacional del Prado, Château de Versailles, British Museum, Musée du Louvre, The Metropolitan Museum of Art New York, National Museum of the U.S. Air Force, Anne Frank House, Musée d'Orsay, The State Hermitage Museum, National Gallery of Art Washington, Rijksmuseum Amsterdam, Guggenheim Bilbao Museum, The Getty Museum, The Museum of Modern Art (MoMA), Museo Nacional de Antropología, Smithsonian National Museum of Natural History, Museo Egizio Torino, National Women's History Museum, Solomon R. Guggenheim Museum, Museu Picasso Barcelona, The Dalí Museum, MASP - Museu de Arte de São Paulo Assis Chateaubriand, Pinacoteca di Brera, Musei Capitolini, Musei Vaticani, Galleria degli Uffizi, Studio Ghibli Museum, The Pergamon Museum, National Museum of Modern and Contemporary Art (South Korea).

3.2.1 Instagram

Follower Growth [figure 1 below]: This metric measures the number of new followers that the entire case study's list of museums gained on Instagram over that set period of time. This provides an indication of the "share of conversation" captured and consequently of the success on a certain platform, but not of the level of active participation of the users. The percentage of Instagram followers of the list of Museums with Virtual Tours increased by 40,2% in two years. From January 2020 to around June 2020, a time window that corresponds to the first lockdown period, we can observe the highest growth. However, it is important not only to

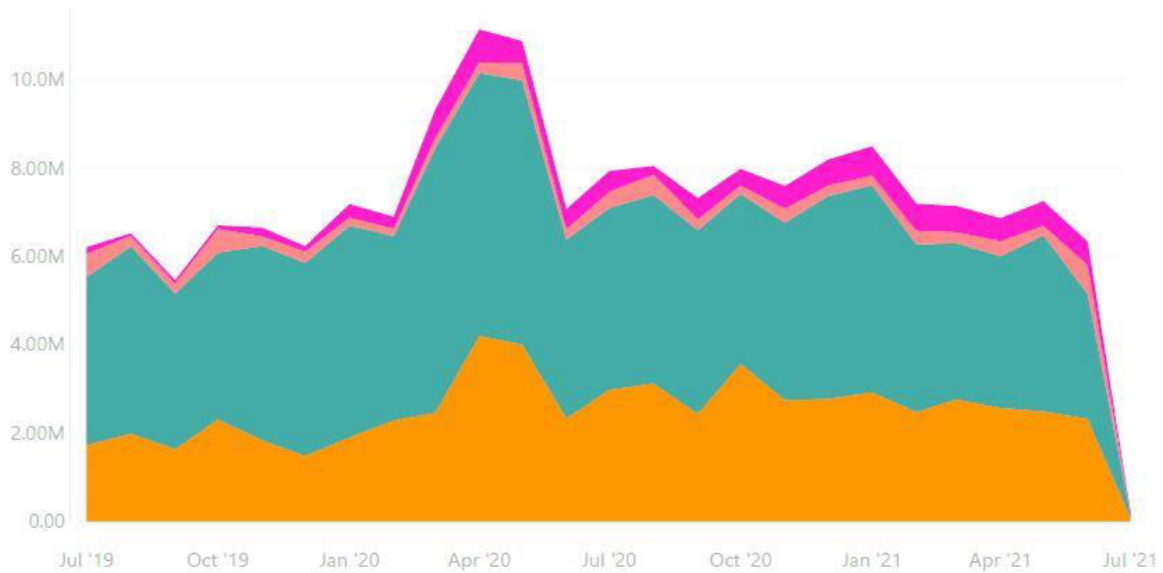
measure how many new followers the Institutions got, but also to compare that to the other metrics.



FOLLOWER GROWTH > 07/01/19 - 07/01/21

>	FOLLOWERS	GROWTH	% GROWTH
	29.29M	+8.40M	+40.20%

Total interactions [figure 2 below]: Total interactions represent the sum of different social media actions, namely reactions, comments and shares. Interactions are also known as engagement, and represent a metric that can reveal an active response of the users if positively compared to a growing line graph regarding interaction rate. Also from this point of view, the trend reveals a main peak, namely the spring of 2020, which corresponds to the first lockdown period. Photos is way up the most engaging type of content, followed by albums and IGTV.



ALL POSTS > TOTAL INTERACTIONS > 07/01/19 - 07/01/21



TOTAL INTERACTIONS

180.93M

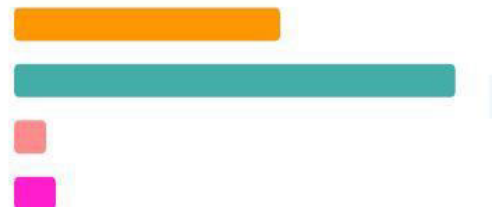
AVG. MONTHLY INTERACTIONS

7.24M

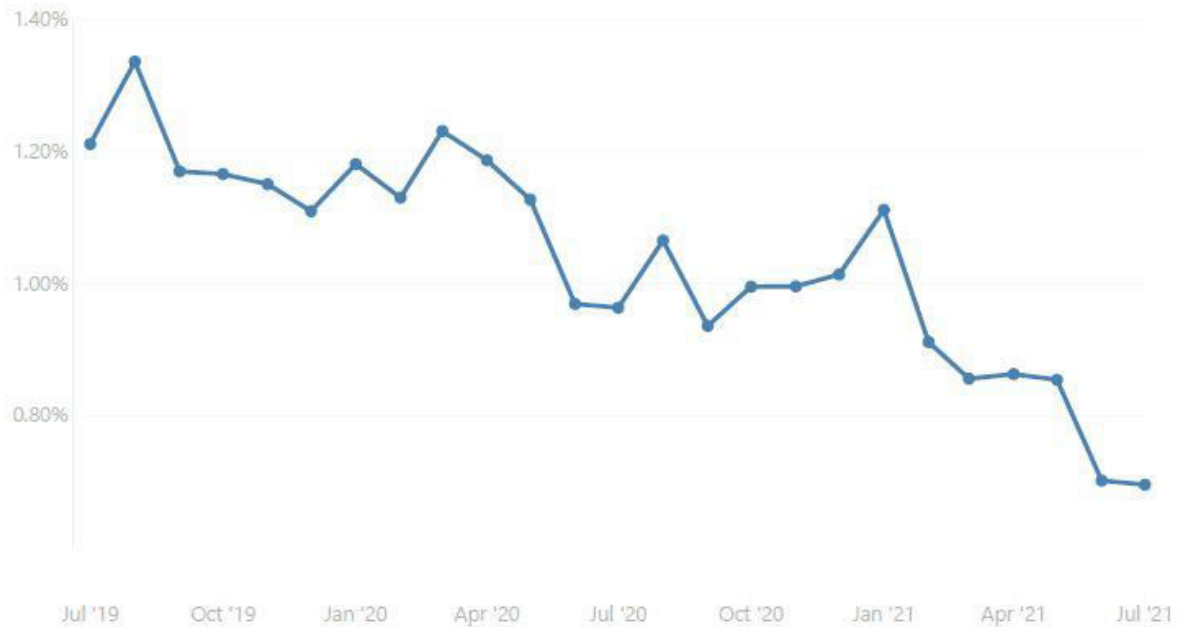
MONTHLY INTERACTION RATE

1.04%

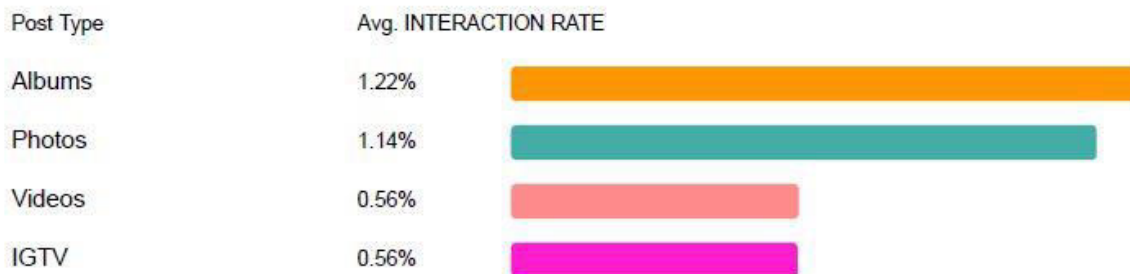
Post Type	Total	%
Albums	61.68M	34.09%
Photos	102.33M	56.56%
Videos	7.34M	4.06%
IGTV	9.57M	5.29%



Interaction rate [figure 3 below]: The Interaction Rate is calculated by adding up all the interactions on every post from every account in the list (suitably weighted), and then dividing it by the total number of posts and by the average size (follower count/page likes) of the respective account. Here too we can find various peaks, the most important of which is observed in the fall/winter 2020, in correspondence with the second lockdown. In general, we can observe how the interaction rate strongly decreased in the last two years.

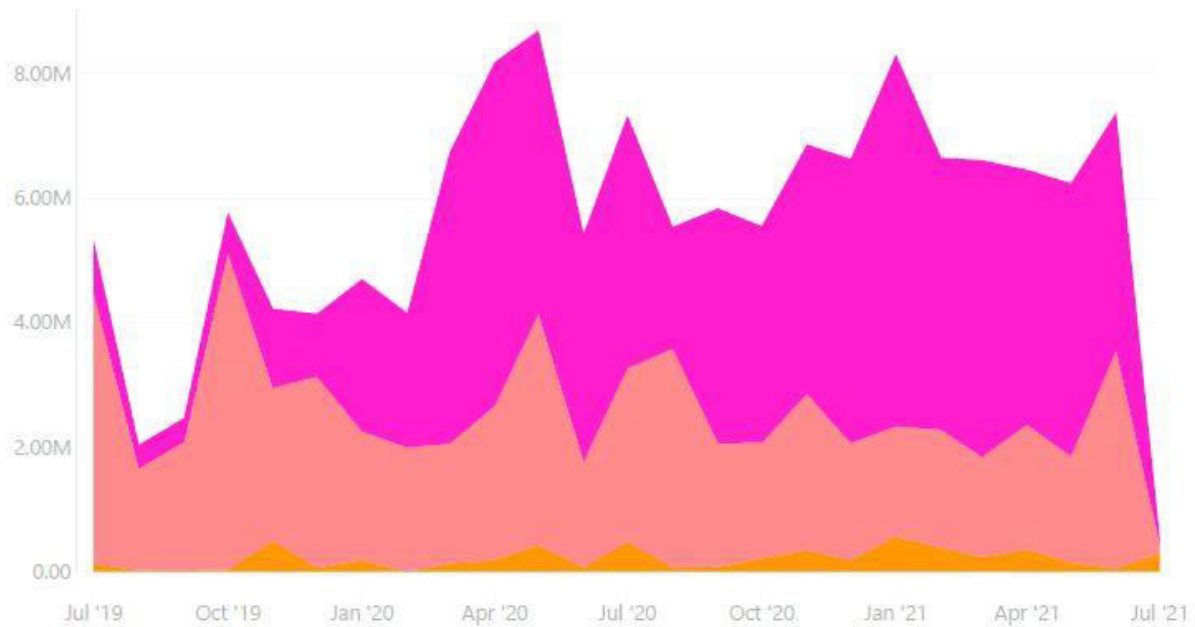


ALL POSTS > AVG. INTERACTION RATE > 07/01/19 - 07/01/21



Total views [figure 4 below]: IGTVs content is the most viewed with 55% of views (77 millions of views for 2200 IGTV videos), which is different from the “Reels” (Video). Reels and IGTV are two different functions of the Instagram social network. The Reels are small videos of up to 15 seconds that are intended to amaze, entertain and engage immediately as an implicitly competing content to those that are found on TikTok. IGTV is instead a function that allows users to follow and share longer, more articulate videos than those normally available on Instagram. IGTV

offers live content and videos in strictly vertical format, to optimize them for mobile access. They can last up to 1 hour, depending on the popularity of the creator. Creators can use this channel to communicate, inform, and comment, enabling in the meantime users to interact through reactions and comments. Also here we find different peaks of activity. The highest increase of views is again found in Spring 2020.



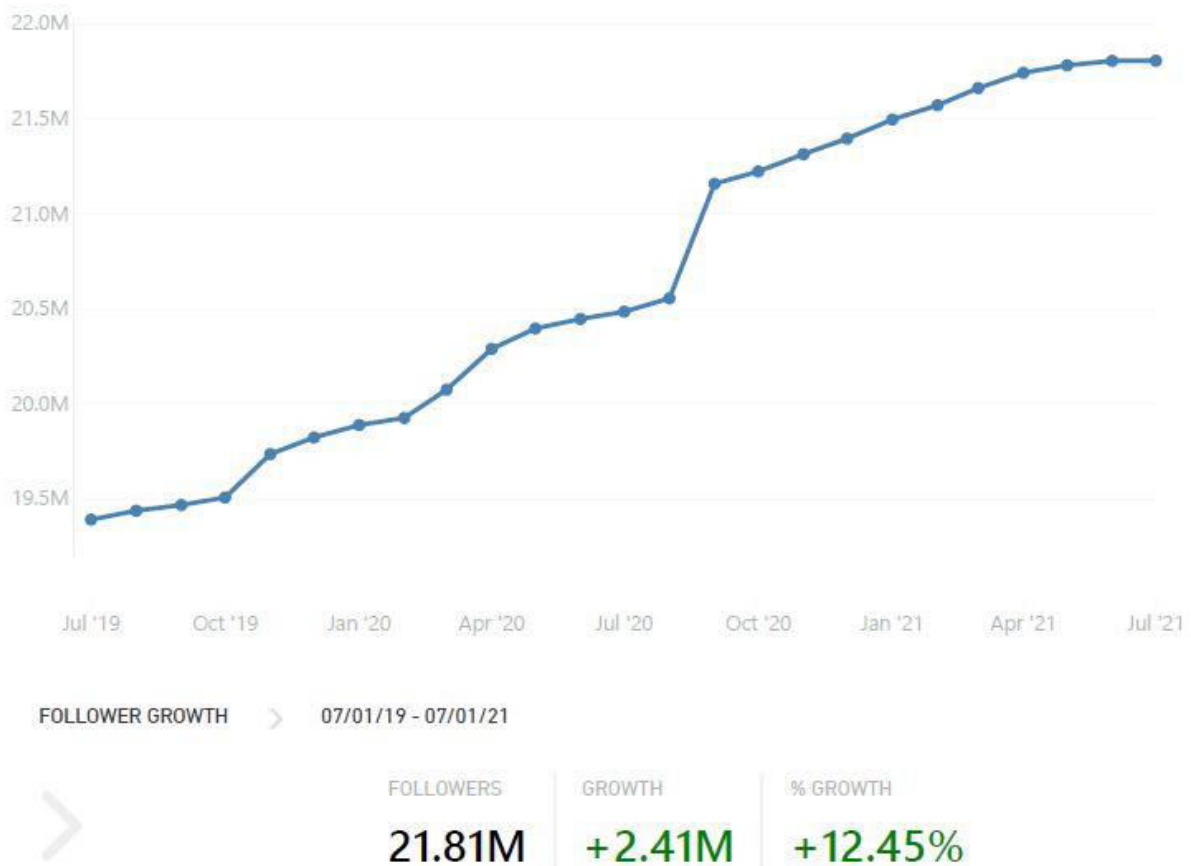
ALL POSTS > TOTAL VIEWS > 07/01/19 - 07/01/21

TOTAL VIEWS: **142.00M** | TOTAL POSTS: **9,833**

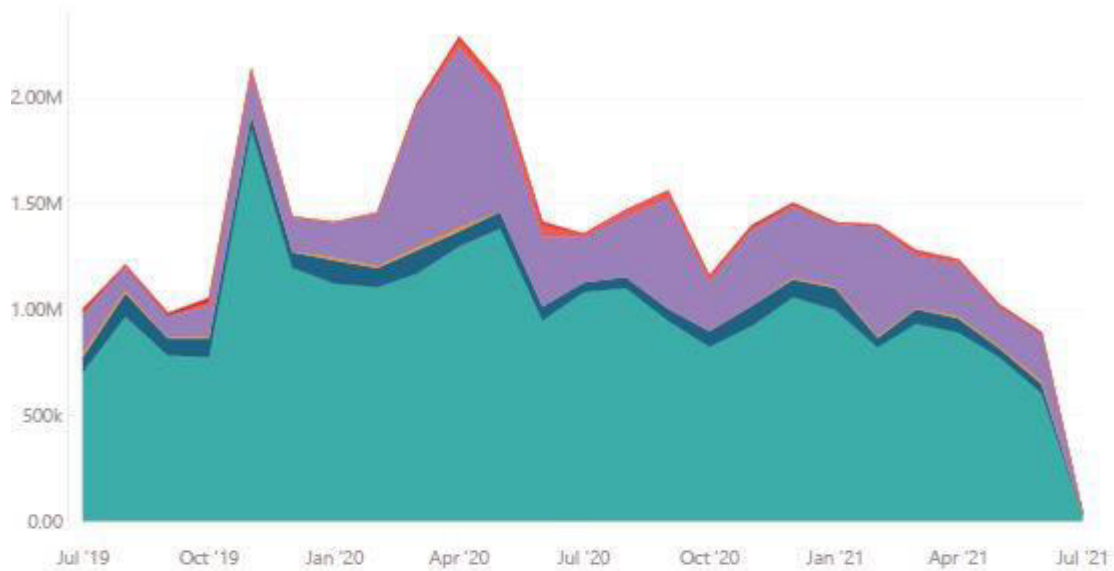
Post Type	# Views	%	# Posts
Albums	5.24M	3.69%	6.1K
Videos	59.73M	42.07%	1.5K
IGTV	77.02M	54.24%	2.2K

3.2.2 Facebook

Follower Growth [figure 5 below]: This metric measures the number of new followers that the entire case study’s list of museums gained on Facebook over that set period of time. The percentage of Facebook followers of the list of Museums with Virtual Tours increased by 12,5% in two years. In the summer of 2020 (September 2020 +603.1K), we can observe the highest growth. From winter 2020, we can observe a constant but slow growth.



Total interactions [figure 6 below]: As to Facebook users’ engagement with Museums with Virtual Tours. The trend reveals a main peak, namely the spring of 2020, which corresponds to the first lockdown period but these data are not necessarily correlated: we can observe a high peak also during fall 2019, before the pandemic. On Facebook, the type of content that triggers the most of the reactions and engagement so far is “Photos”.

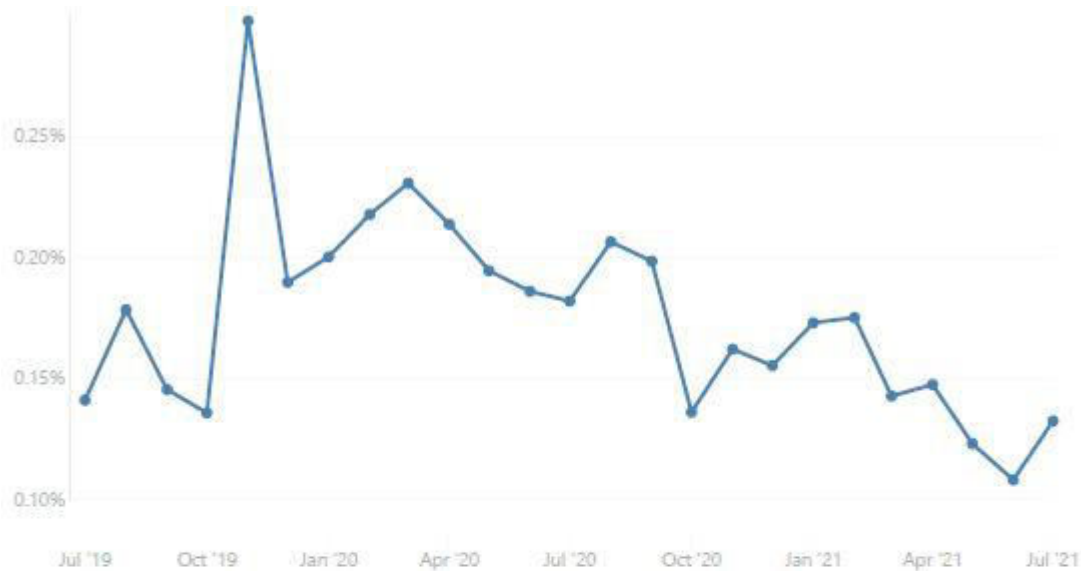


ALL POSTS > TOTAL INTERACTIONS > 07/01/19 - 07/01/21

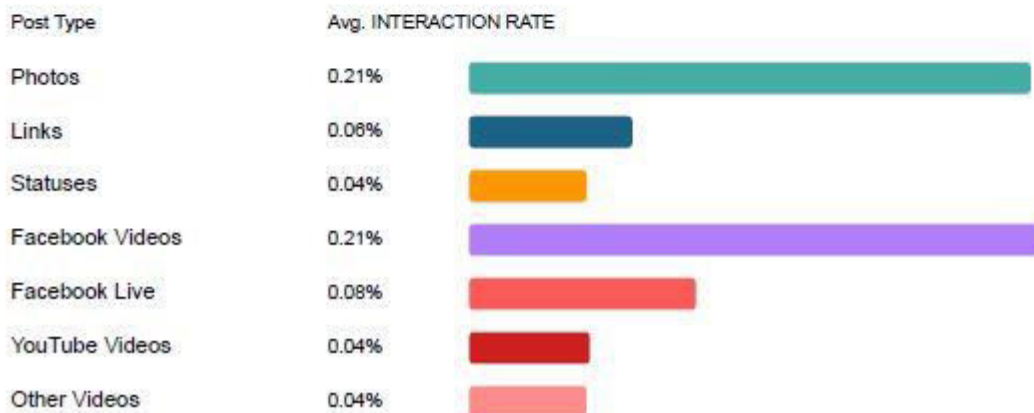
TOTAL INTERACTIONS: **34.23M**
 AVG. MONTHLY INTERACTIONS: **1.37M**
 MONTHLY INTERACTION RATE: **0.18%**

Post Type	Total	%
Photos	24.38M	71.17%
Links	1.77M	5.16%
Statuses	191.9K	0.56%
Facebook Videos	7.32M	21.40%
Facebook Live	407.1K	1.19%
YouTube Videos	161.8K	0.47%
Other Videos	17.2K	0.05%

Interaction rate [figure 7 below]: We can observe how the interaction rate undergoes strong fluctuations but, in general, it seems that Museums with Virtual tours cyclically retain the attention of their users by stimulating their active participation. Both photos and videos seem to be engaging channels.



ALL POSTS > AVG. INTERACTION RATE > 07/01/19 - 07/01/21



Total views [figure 8 below]: In the last two years, the posted videos have been around 5000 and obtained up to 135 millions of views, with 3 peaks that correspond to the 3 lockdown periods between 2020 and 2021. Videos created and shared on Facebook pages from the selected museums of this case study seem to be a very effective channel for reaching out to digital users, but the trend of the graph line is discordant with that of the interaction (engagement) rate.



ALL POSTS > OWNED VIDEOS > TOTAL VIEWS > 07/01/19 - 07/01/21

> TOTAL VIEWS FROM PAGE POSTS FROM SHARES TOTAL VIDEOS

140.51M **100.28M** **40.24M** **4,949**

Post Type	# Views	%	# Videos
Facebook Videos	134.39M	95.64%	4.4K
Facebook Live	6.12M	4.36%	563

3.3 Most Visited Museums of Art in Europe

The case study list of the Most Visited Art Museums in Europe is built on a 2019 report, “Leading museums in Europe in 2019, based on attendance”⁴. The main objectives of the data collection and analysis of this case study are the observation of the initiatives of the selected museums and of the responses of their users in the pandemic period compared to the previous year, in order to detect current trends in terms of digital participation in the Museums sector. To select the sample, we considered the 20 most visited museums in Europe, six of which are located in the UK. These six museums accounted for over 30 million visitors combined in 2019, and have a strong digitization program, in particular the British Museum, which is the most visited public national museum in the world and attracts roughly six million visitors per year. In order to provide a homogeneous coverage of European countries, as the museums located in the UK gather the majority of the most receptive and best performing museums and thus inevitably bias the sample in that they present fundamental differences with respect to all the other museums, we decided to exclude them from the case study’s list. We tried to maintain the samples used for Facebook and Instagram as similar as possible, given that not all the museums selected had both comparably active Instagram and Facebook pages that could be amenable to data analysis.

The list considered in the case study is as follows:

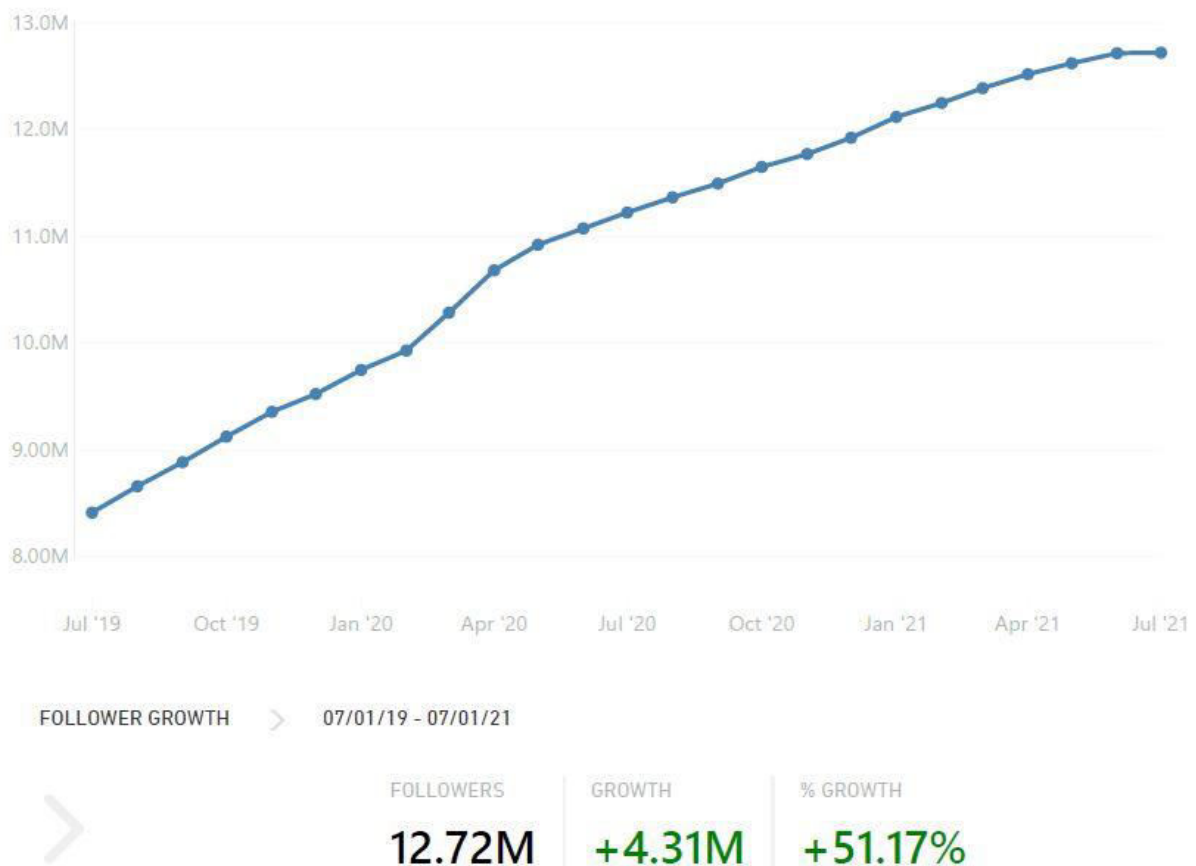
Museo Nacional del Prado, Musée du Louvre, Gallerie degli Uffizi, Centre Pompidou, Rijksmuseum Amsterdam, Museo Guggenheim Bilbao, Musée d'Orsay, Zamek Królewski w Warszawie, Museo Nacional Thyssen-Bornemisza, Museo Reina Sofía, Museu Picasso Barcelona, Belvedere Museum, Grand Palais - RMN, Kunsthistorisches Museum Vienna, Musée du quai Branly - Jacques Chirac, Fondation Beyeler, SMK - Statens Museum for Kunst, Musée de l'Orangerie, Mucem, Städel Museum, Serralves, Fine Arts Belgium, Galleria dell'Accademia di

⁴ see: <https://www.statista.com/statistics/747942/attendance-at-leading-museums-in-europe/> (last accessed September 2021)

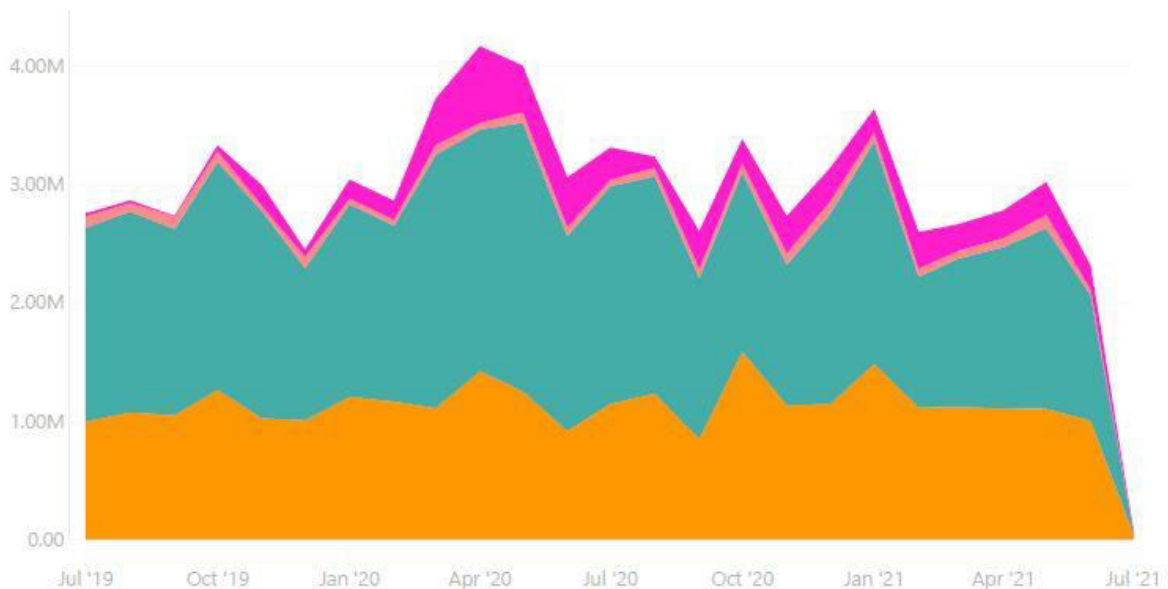
Firenze, Deutsches Historisches Museum, Museu Coleção Berardo, Musei Vaticani, Cecilie Hollberg, Deutsches Historisches Museum, Fondation Louis Vuitton, Kunsthistorisches Museum Wien, Le Grand Palais, Mucem, Musée de l'Orangerie, Musée d'Orsay, National Gallery Prague, SMK - Statens Museum for Kunst, Städel Museum, Stedelijk Museum Amsterdam

3.3.1 Instagram

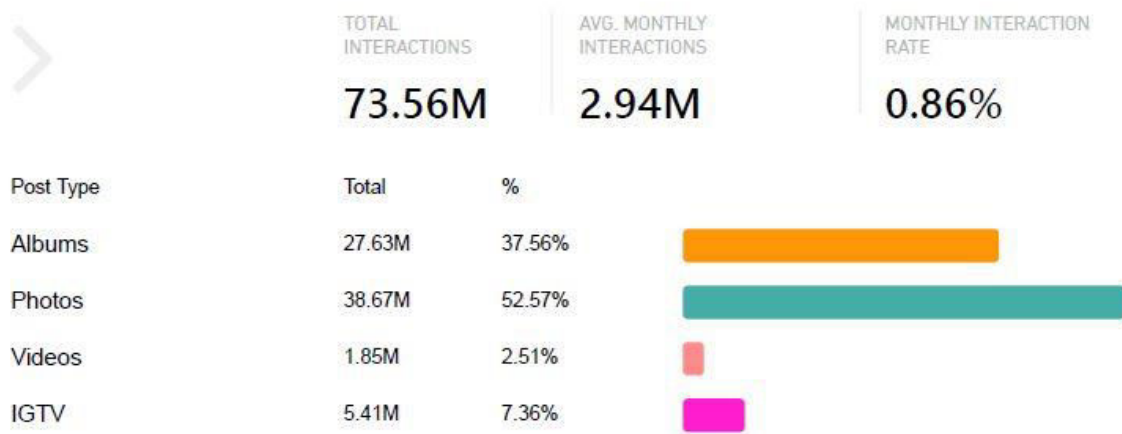
Follower Growth [figure 9 below]: The number of new followers that the entire case study's list of museums gained on Instagram over the last two years increased more than 51%. From January 2020 to around June 2020, a time window that corresponds to the first lockdown period, we can observe the highest growth. However, we can observe positive growth throughout.



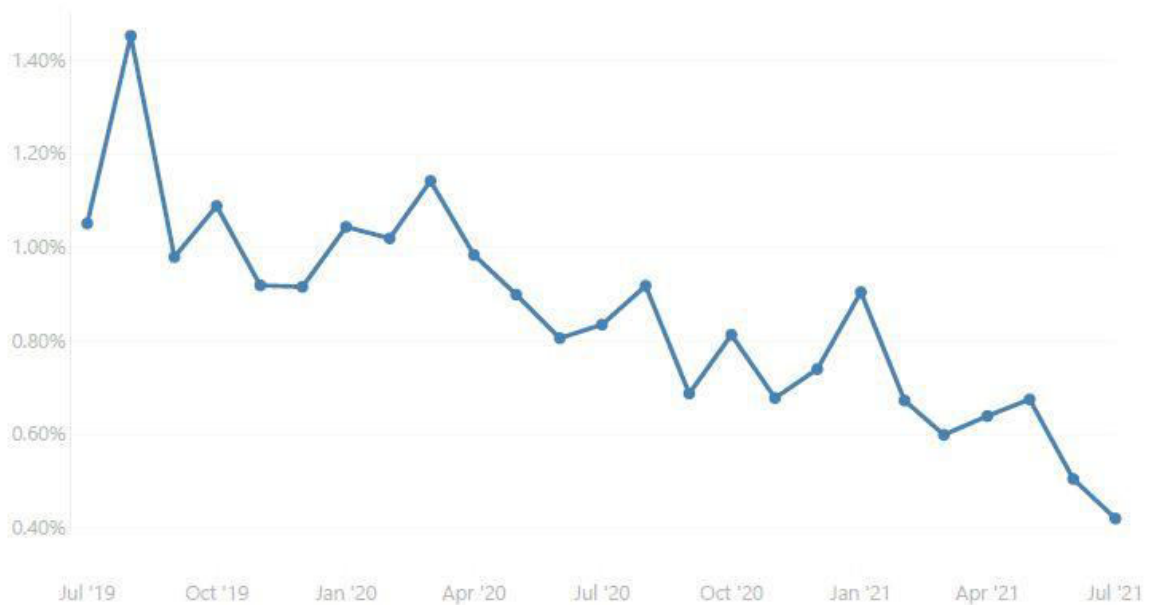
Total interactions [figure 10 below]: The trend reveals various peaks that grow and decrease cyclically, highlighting waves of renewal of interest toward, and interaction with, the selected museums by their digital visitors.



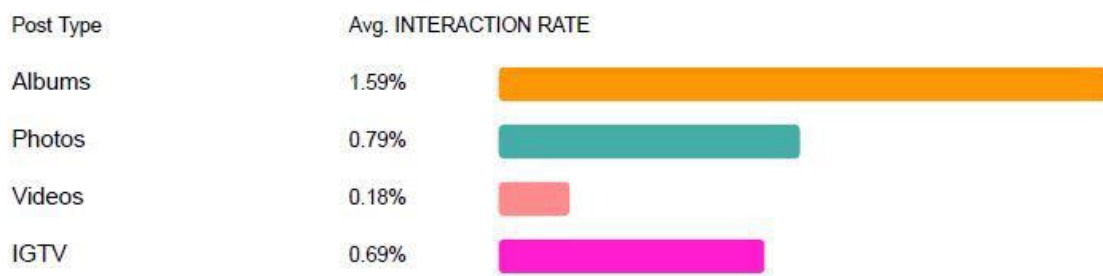
ALL POSTS > TOTAL INTERACTIONS > 07/01/19 - 07/01/21



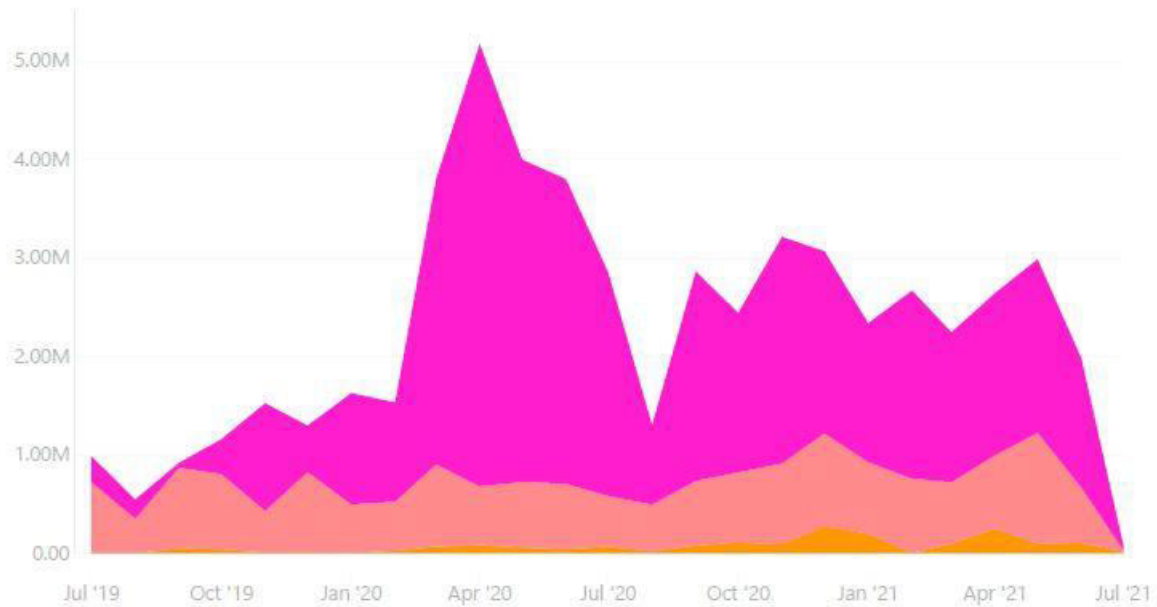
Interaction rate [figure 11 below]: The Interaction Rate is here characterized by various peaks of renewed waves of active participation by visitors but, in general, we can observe how the interaction rate decreased in the last two years overall.



ALL POSTS > AVG. INTERACTION RATE > 07/01/19 - 07/01/21



Total views [figure 12 below]: IGTV is again the most viewed channel of content sharing, with around 38 millions of views on 2200 posts. It is worth mentioning that this figure is way smaller than the corresponding one for the global panel of museums with Virtual Tours, further highlighting the importance of developing targeted digital tools for online visitors. Also here we can find different peaks, but in the time window related to the first semester of 2020 we once more find the highest increase of views.



ALL POSTS > TOTAL VIEWS > 07/01/19 - 07/01/21

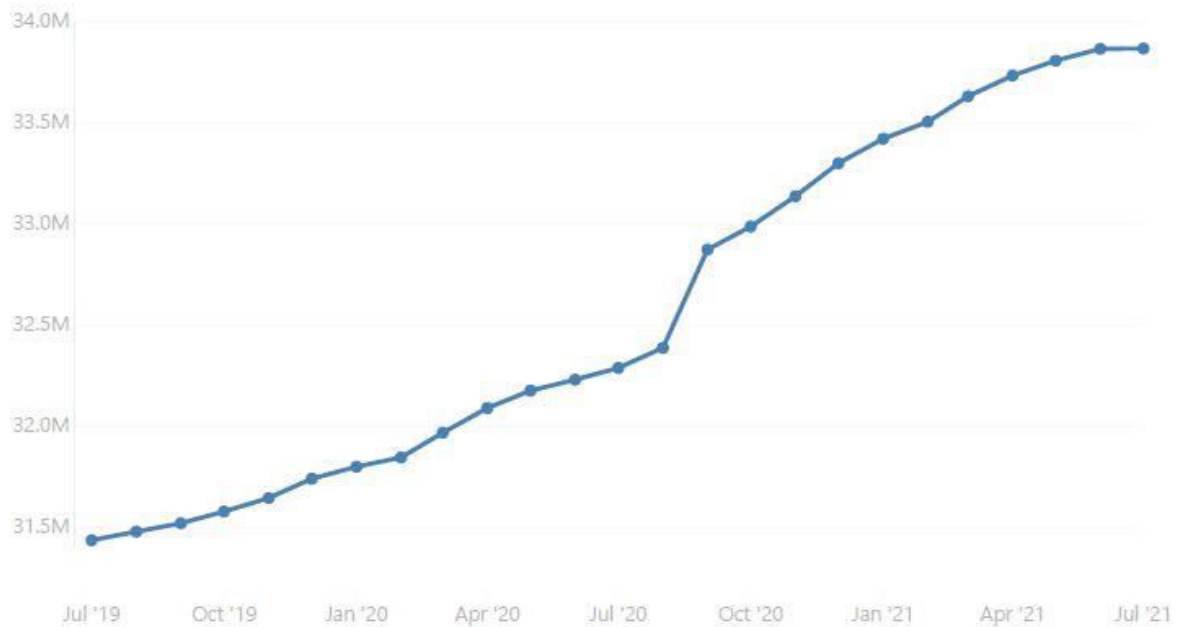
TOTAL VIEWS: **57.00M** | TOTAL POSTS: **10,197**

Post Type	# Views	%	# Posts
Albums	1.82M	3.19%	5.2K
Videos	16.33M	28.65%	2.8K
IGTV	38.86M	68.16%	2.2K

3.3.2 Facebook

Follower Growth [figure 13 below]: The percentage of Facebook followers of the list of the most visited Museums grew only 7.7% in the last two years, with a faster

growth during summer of 2020.



FOLLOWER GROWTH > 07/01/19 - 07/01/21



FOLLOWERS

33.87M

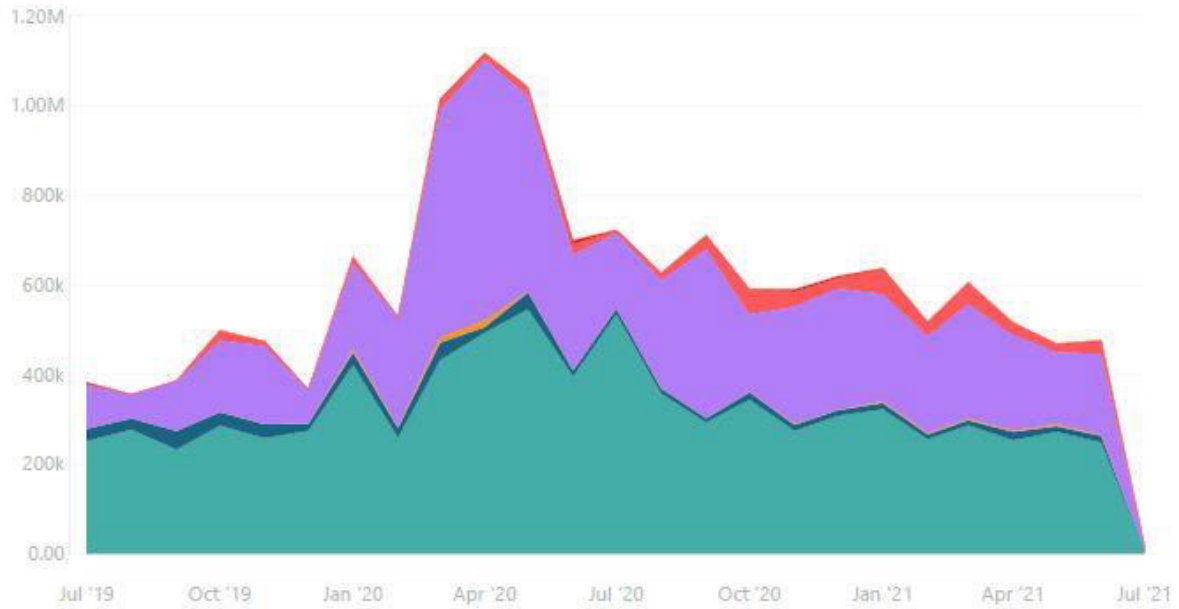
GROWTH

+2.43M

% GROWTH

+7.74%

Total interactions [figure 14 below]: The trend reveals a main peak, namely the spring of 2020, which once more corresponds to the first lockdown period. On Facebook, the types of content that trigger the most reactions and engagement are “Photos” and “Videos”.



ALL POSTS > TOTAL INTERACTIONS > 07/01/19 - 07/01/21



TOTAL INTERACTIONS

14.65M

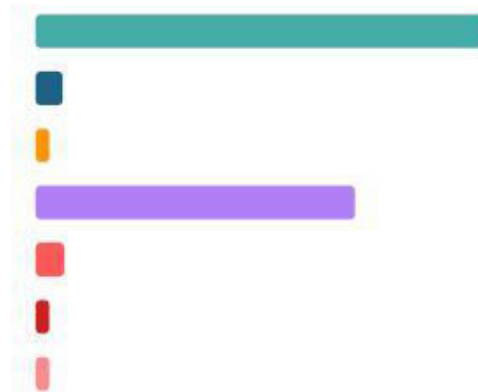
AVG. MONTHLY INTERACTIONS

586.1K

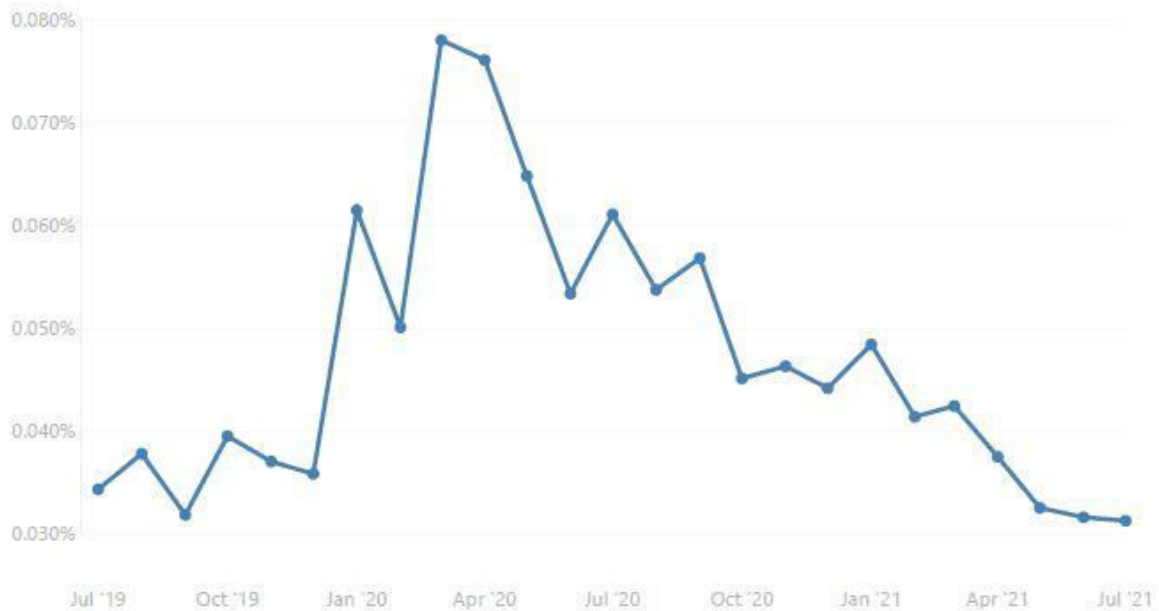
MONTHLY INTERACTION RATE

0.05%

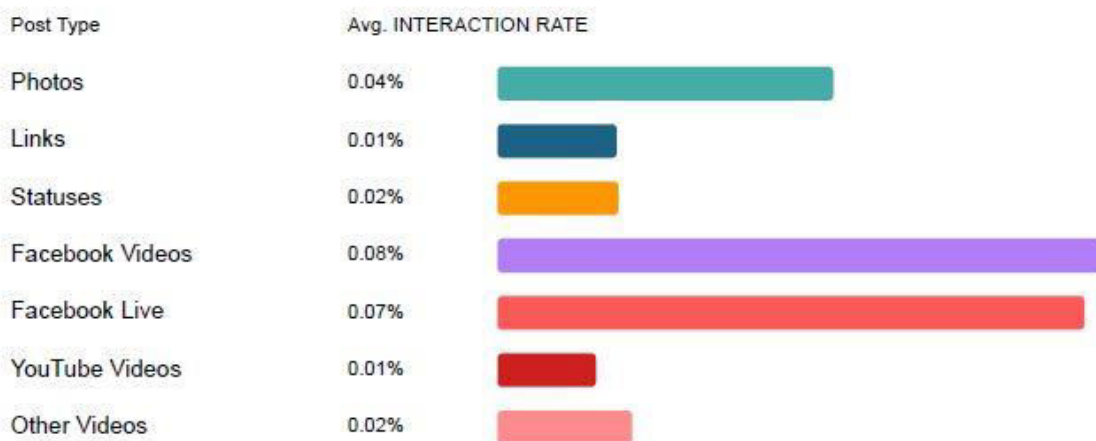
Post Type	Total	%
Photos	7.92M	54.03%
Links	472.6K	3.23%
Statuses	61.7K	0.42%
Facebook Videos	5.65M	38.56%
Facebook Live	501.9K	3.43%
YouTube Videos	44.8K	0.31%
Other Videos	5.0K	0.03%



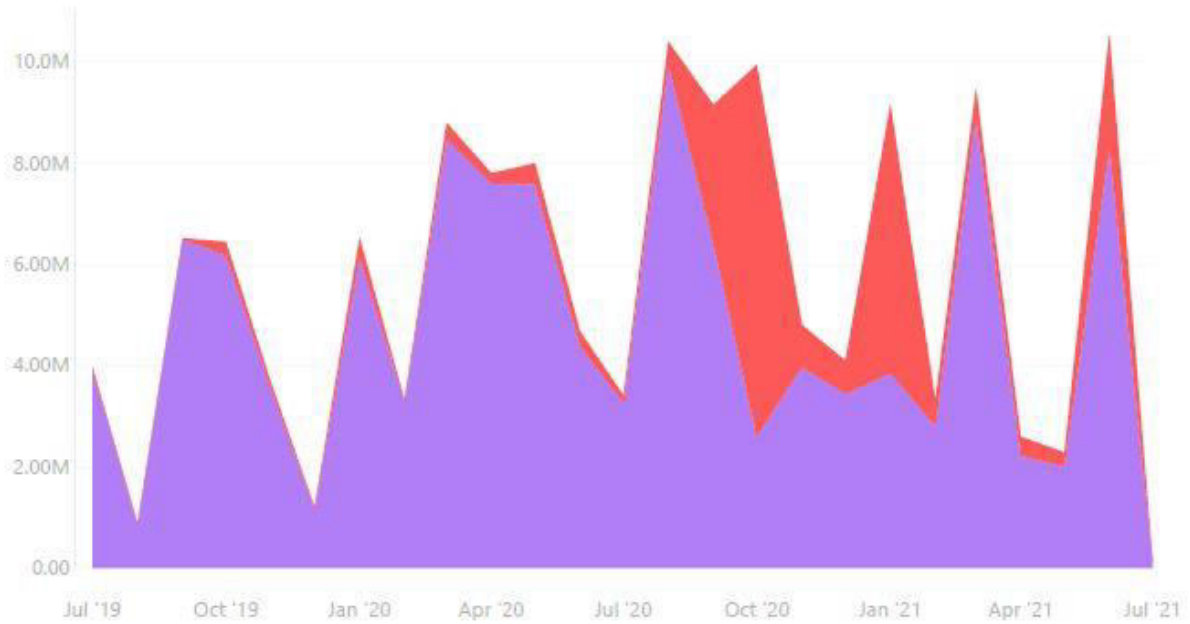
Interaction rate [figure 15 below]: The interaction rate percentage is higher here regarding Facebook Videos and Live streaming, with various peaks all along 2020.



ALL POSTS > AVG. INTERACTION RATE > 07/01/19 - 07/01/21



Total views [figure 16 below]: In the last two years the posted videos have been around 6700 and obtained up to 141 millions of views, and various very high peaks: the museums' Facebook pages users are quite assiduous in following the contents posted by their preferred museums.



ALL POSTS > OWNED VIDEOS > TOTAL VIEWS > 07/01/19 - 07/01/21



Post Type	# Views	%	# Videos
Facebook Videos	117.47M	83.16%	6.2K
Facebook Live	23.78M	16.84%	625

3.4 Libraries

This case study is built on the list of the Facebook and Instagram accounts of the European National Libraries, a list that homogeneously covers the whole territory of Europe. We selected institutes that are defined as a “national library”, namely libraries specifically set up by the Government with the aim of archiving and preserving national publishing production and legal deposits. National libraries collect rare, valuable publications or works that are particularly significant not only because they preserve their own specific culture, but also because they are valuable heritage assets for the international community. The main goals of this research is to analyse the Library institutes’ sector in terms of the virtual relationship with their community of users in a temporal window that goes from mid 2019 to mid 2021, thus covering the Covid-19-related forced digitization revolution.

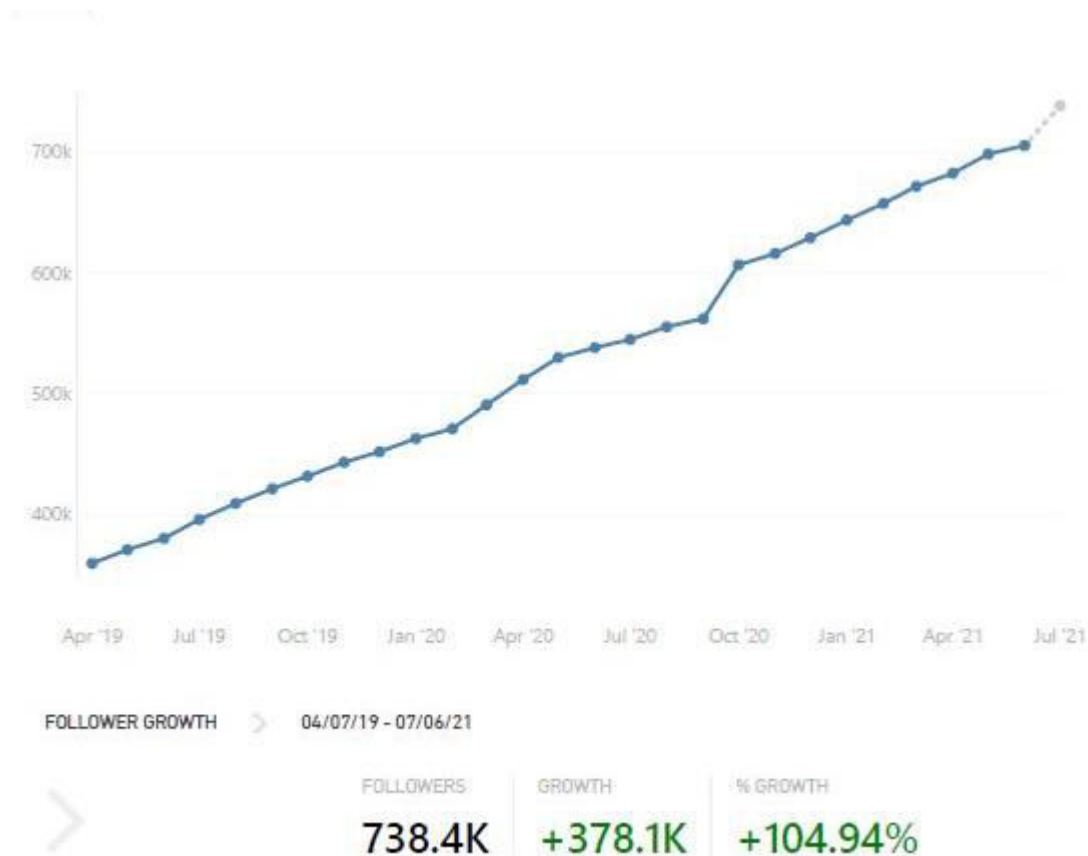
The list covered in the case study is as follows:

The Vatican Library, National Library of Spain, National Library of Portugal, National Library of the Republic of Moldova, National Library of Romania, The Central National Library of Florence, The Central National Library of Rome, National Library of Albania, National Library of Poland, National Library of France, National Library of Luxembourg, The National Library of Malta, British Library, National Library of Denmark, German National Library, National Library of Estonia, The National Library of Finland, Royal Library of the Netherlands, Royal Library of Belgium, National Library of Sweden, National University Library of Iceland, National Library of Latvia, National Library of Ireland, National Library of Liechtenstein, Martynas Mazvydas National Library of Lithuania, National Library of Azerbaijan, National Library of Turkey, National Library of Montenegro, National and University Library in Zagreb, National and University Library of Bosnia and Herzegovina, National and University Library St. Kliment Ohridski, National and University Library of Slovenia, National Library of the Czech Republic, National Library of Norway, National Library of Hungary, National Library of Austria, Swiss National Library, Slovak National Library,

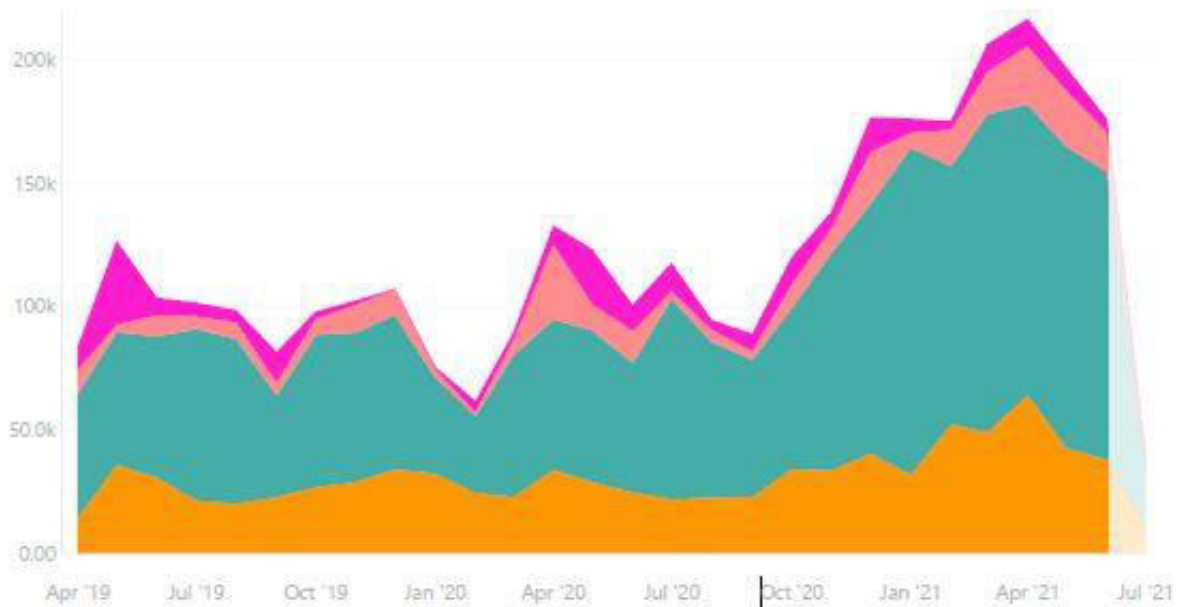
National Library of Greece, Cyprus Library, National Library of Serbia, St. St. Cyril and Methodius National Library, V. Vernadsky National Library of Ukraine, Boris Yeltsin Presidential Library, Russian State Library, National Library of Russia, National Library of Armenia, National Parliamentary Library of Georgia

3.4.1 Instagram

Follower Growth [figure 17 below] The trend reveals a 105% growth and two different peaks: the first one during spring 2020 and the second one during fall 2020.



Total Interactions [figure 18 below] “Photo” is the most important and engaging type of content. The trend shows a big increase in the interactions during 2021, with a peak that begins to mount in the fall 2020, namely during the second lockdown period.



ALL POSTS > TOTAL INTERACTIONS > 04/07/19 - 07/06/21

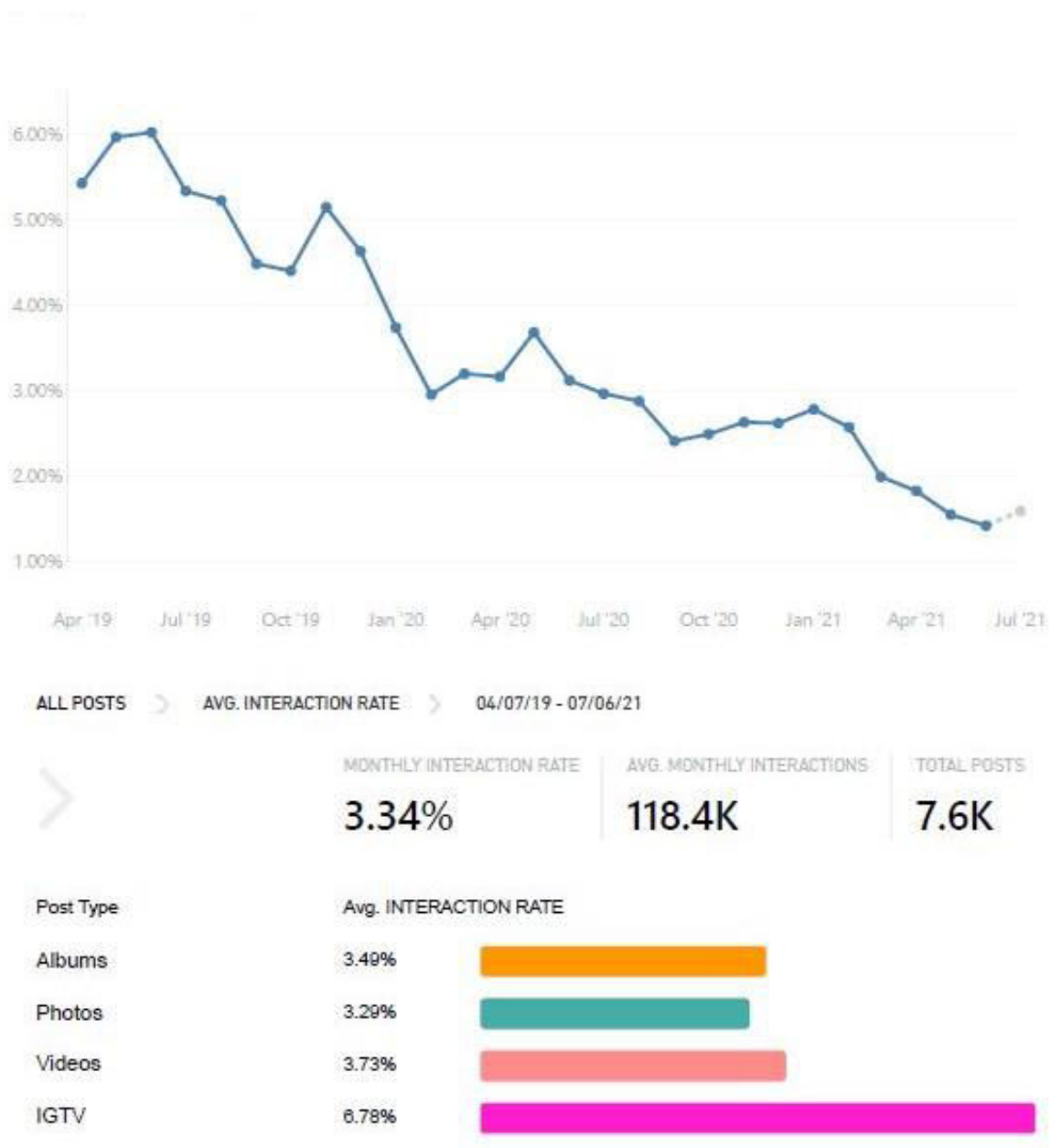
TOTAL INTERACTIONS
3.41M

AVG. MONTHLY INTERACTIONS
118.4K

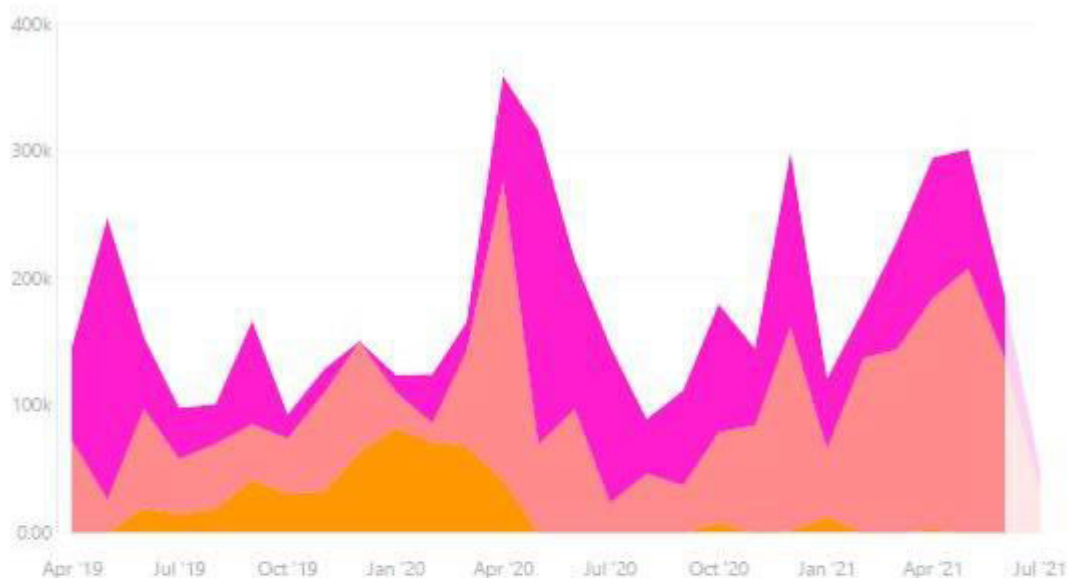
MONTHLY INTERACTION RATE
3.34%

Post Type	Total	%
Albums	870.3K	25.49%
Photos	2.03M	59.36%
Videos	288.6K	8.45%
IGTV	228.6K	6.69%

Interaction rate [figure 19 below]: The Interaction Rate is calculated by adding up all the interactions on every post from every account in the list (suitably weighted), and then dividing it by the number of posts and by the average size. As the previous graph shows, while content posted increased, the interaction rate diminished over time, revealing a decline in user interest.



Total views [figure 20 below]: Videos are the most viewed content, with more than 2 millions views (both videos and IGTV) per only around 500 posts published in two years per category. Also here we can find three different peaks: spring 2020, fall 2020 and winter 2021.



ALL POSTS > TOTAL VIEWS > 04/07/19 - 07/06/21

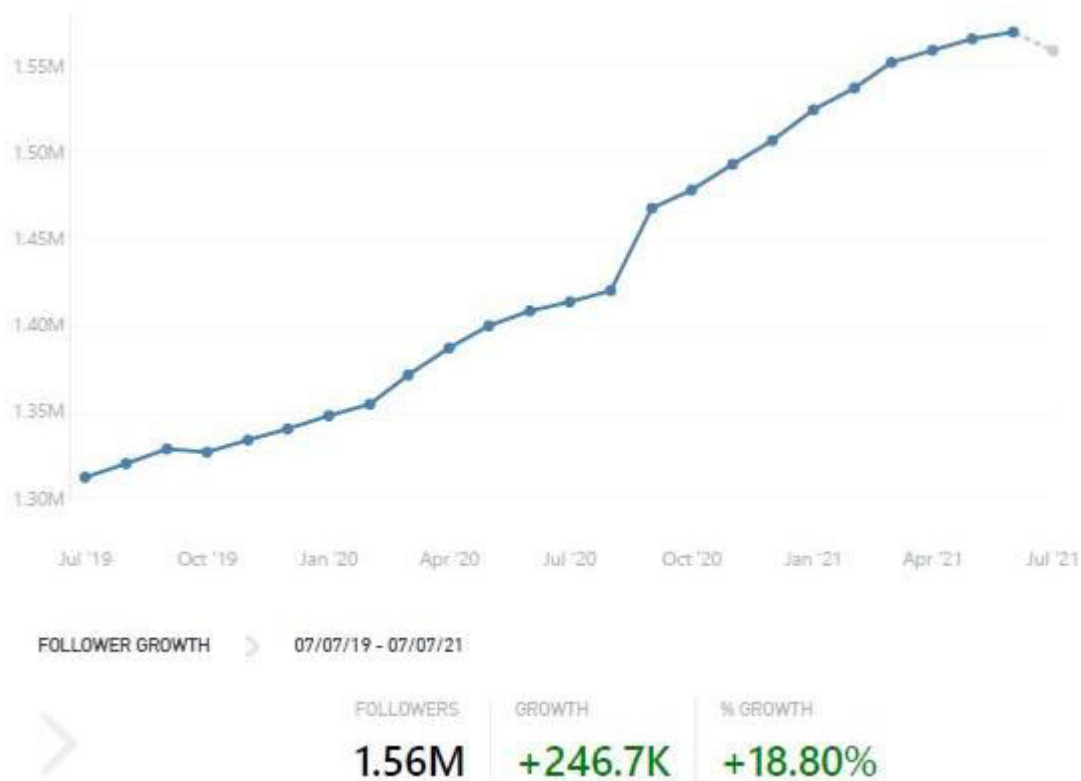
TOTAL VIEWS: **4.91M** TOTAL POSTS: **2,834**

Post Type	# Views	%	# Posts
Albums	509.3K	10.37%	1.8K
Videos	2.37M	48.17%	570
IGTV	2.04M	41.47%	484

3.4.2 Facebook

Follower Growth [figure 21 below]: Here the time window is slightly reduced if compared with the data available for Instagram. For Libraries, the trend reveals two different peaks: the first one during spring 2020, and the second one during fall

2020. This is probably due to the pandemic-related public health measures (such as lockdowns and “stay at home” campaigning). Nevertheless, the growth of followers is still ongoing and this could point out a large, permanent effect on social media users which is of particular interest.



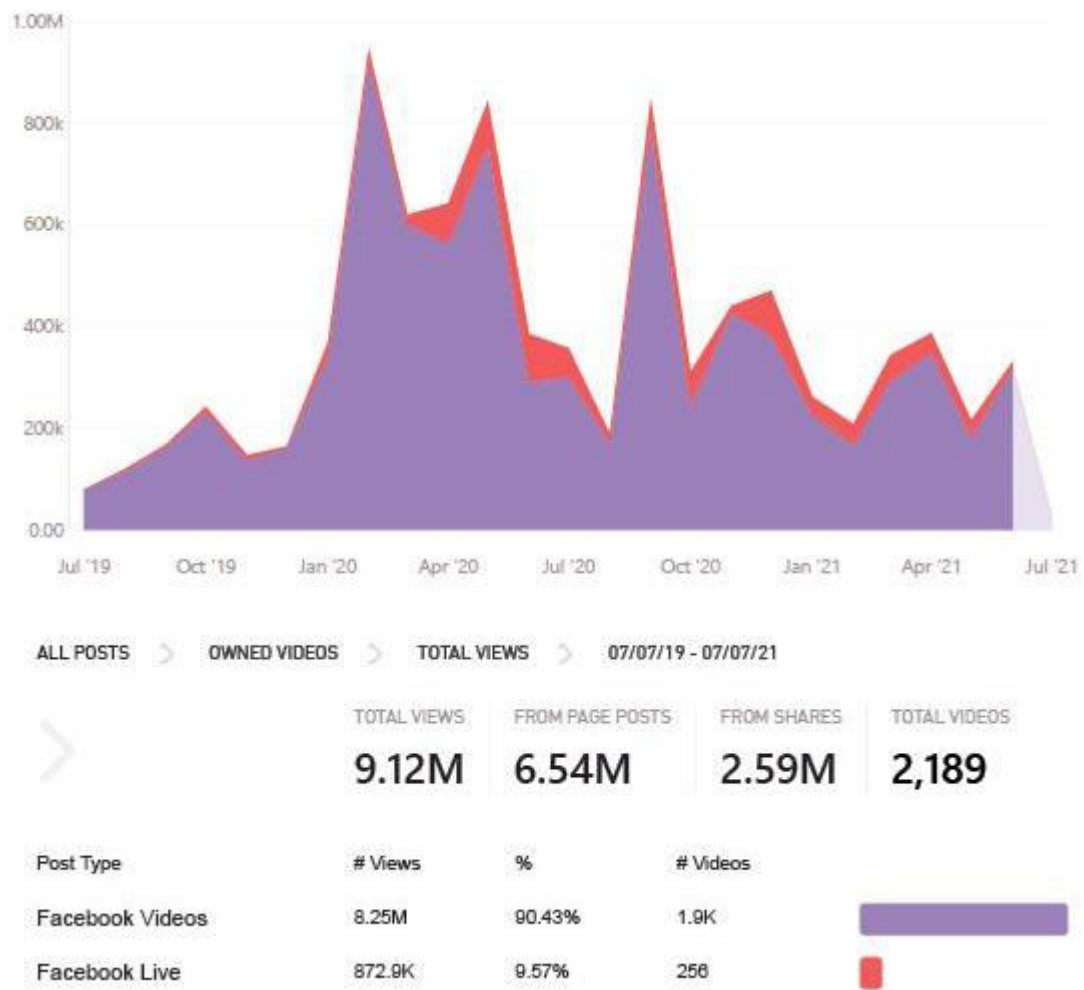
Total interactions [figure 22 below] Total interactions represent the sum of different social media actions, such as likes, comments or reactions to a post. Also from this point of view, the trend reveals three different peaks, namely the spring and fall of 2020 and winter 2021. Photo is by far the most engaging type of content, followed by videos and external links.



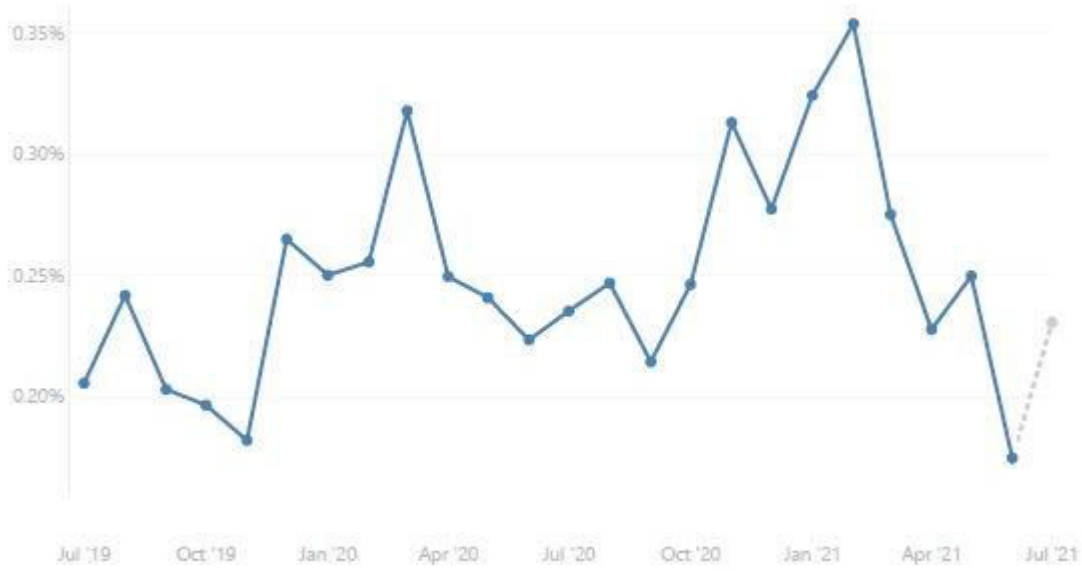
ALL POSTS > TOTAL INTERACTIONS > 07/07/19 - 07/07/21



Total Views [figure 23 below]: The chart represents a simple count of how many times a video is viewed. As we can see, the time series shows three distinct peaks in video views, namely January 2020, May 2020 and September 2020. Interestingly, the first of them, which is also the highest, precedes the pandemic crisis. This pattern cannot therefore merely be attributed to changes in users' choices related to the new situation created by the pandemic.

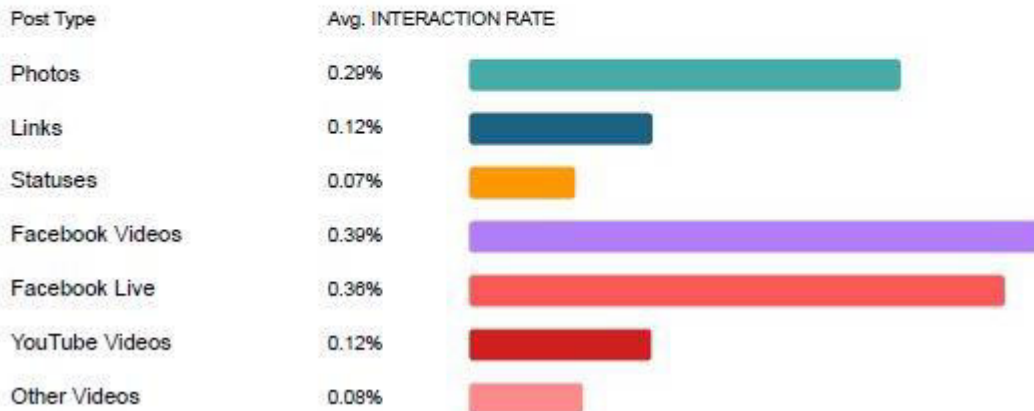


Interaction rate [figure 24 below] The graph line is pretty erratic for the whole two years. Here also we can find three peaks, namely spring 2020, fall 2020 and winter 2021, mainly referred to Facebook videos and Live streaming.



ALL POSTS > AVG. INTERACTION RATE > 07/07/19 - 07/07/21

	MONTHLY INTERACTION RATE	AVG. MONTHLY INTERACTIONS	TOTAL POSTS
>	0.24%	88.8K	28.4K



3.5 Archives

This case study is built on the list of the Facebook and Instagram accounts of the European Archives that are active in those social platforms. The list homogeneously covers the whole territory of Europe and includes both National Archives and Archive Institutes that collect a wide spectrum of cultural and creative products and documents that are historically valuable heritage assets not only for the country, but also for the international community.

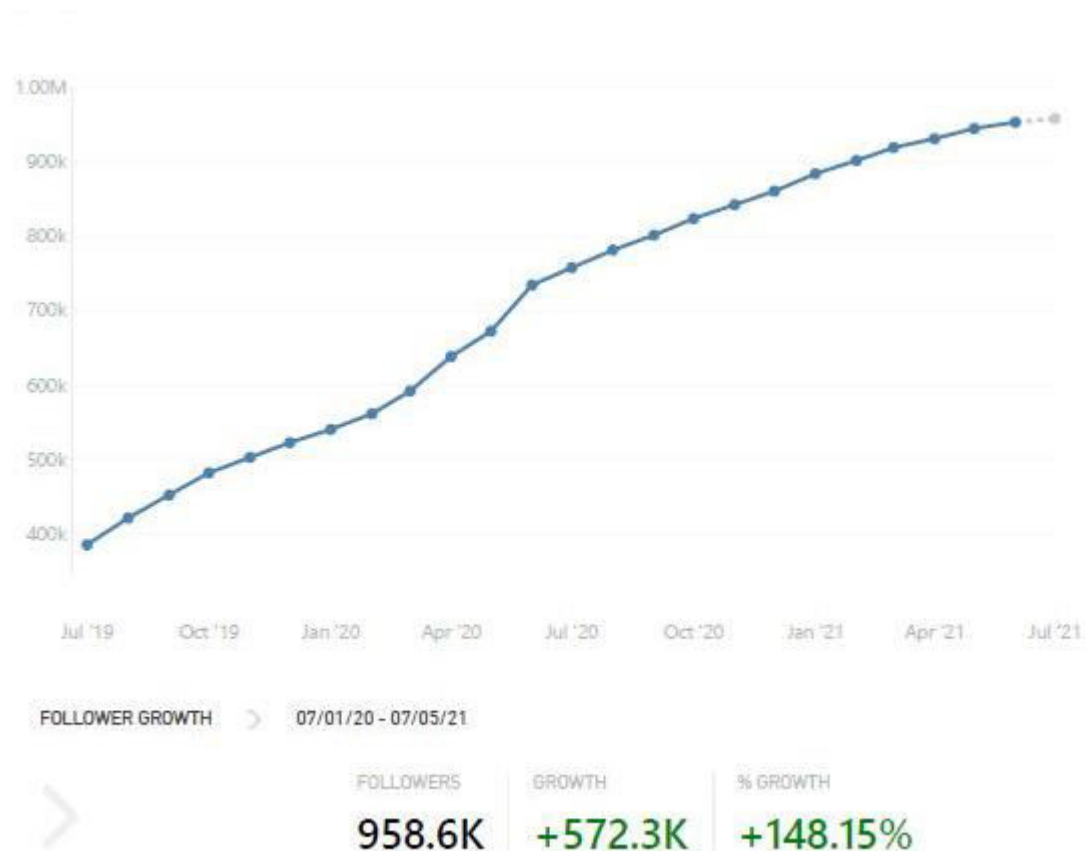
The main goal of this analysis is to analyze how Archive institutes connect to their community of digital users in a temporal window that goes from mid 2019 to mid 2021, observing how the Covid-19-related forced digitization impacted upon such relation.

The list of institutions covered by the case study is as follows:

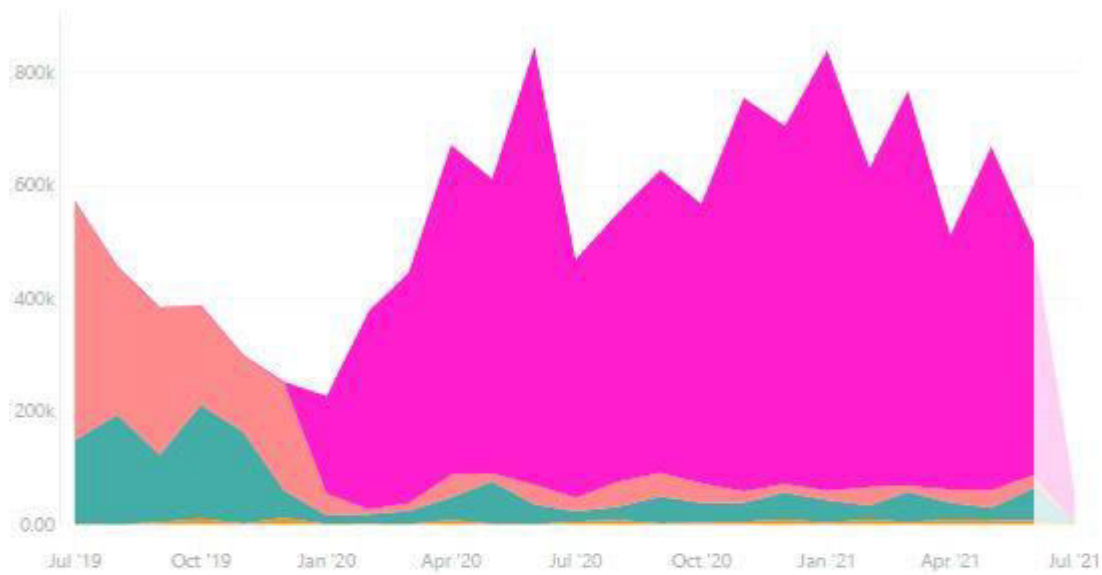
BBC Archives, Austrian State Archives, Rijksarchief, Cinematek (Royal Film Archive of Belgium), The Archives of Bulgaria, Croatian State Archives, Národní filmový archiv, the Danish National Archives, National Archives of Estonia, National Archives of Finland, INA (France), Archives nationales of France, Das Bundesarchiv, Bauhaus Archive, Greek Film Archive, National Archives of Hungary, RTE archives, National Archives of the Netherlands, Central Archives of the State, Luce (Italy), National Archives of Latvia, Lithuanian State Archiv, Lithuanian State Modern Archives, National Archives Luxemburg, National Archives of Malta, Netherlands Institute for Sound and Vision, National Archives of Ireland, National Archives of Norway, Filмотека Народова - Інститут Аудіовізуалнь, Narodowe Archiwum Cyfrowe, Arhivele Nationale ale Romaniei, Historical archives Ljubljana, Archivos Estatales, Archivo Lafuente, Riksarkivet Sverige, Swiss Film Archive, Schweizerisches Bundesarchiv.

3.5.1 Instagram

Follower Growth [figure 25 below]: While the trend reveals a stable growth process, we also see that, after spring 2020, the slope becomes momentarily steeper, suggesting a temporary acceleration in the attraction of new users.



Total Interactions [figure 26 below]: By looking at the graph, it seems that archives started using the IGTV tool precisely in correspondence to the early signals of the pandemic, in January 2020, and thus likely not as a consequence of it, as the lockdown leasures were not only in places, but not even contemplated at that moment. From then on, IGTV became the most utilized tool by Archives to share contents with their digital users on Instagram, reaching around 9 millions of interactions.



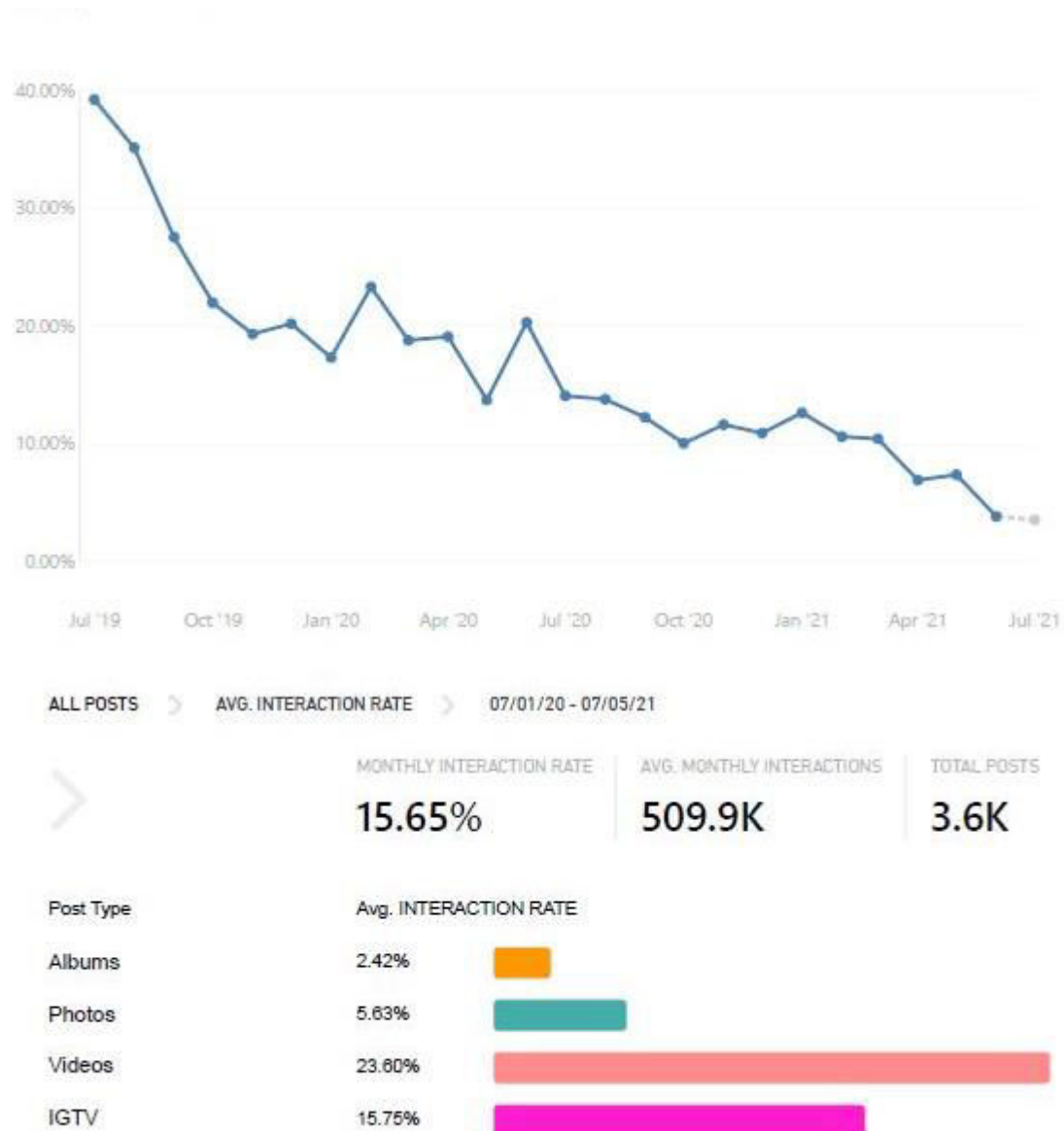
ALL POSTS > TOTAL INTERACTIONS > 07/01/20 - 07/05/21



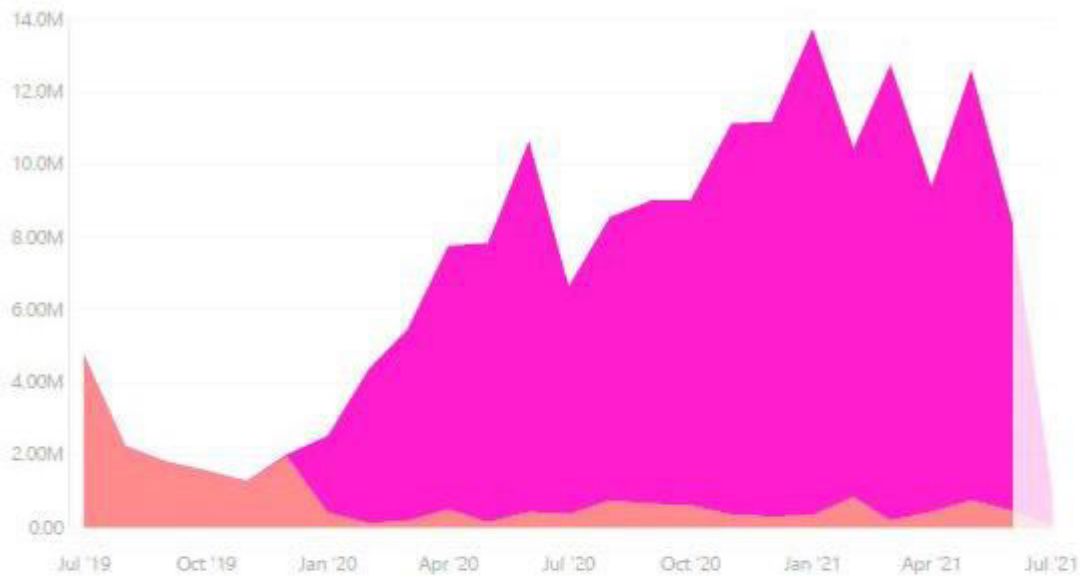
Post Type	Total	%
Albums	123.8K	0.94%
Photos	1.51M	11.48%
Videos	1.92M	14.55%
IGTV	9.63M	73.03%

A horizontal bar chart showing the percentage distribution of post types. The bars are colored: IGTV (magenta, 73.03%), Videos (red, 14.55%), Photos (teal, 11.48%), and Albums (orange, 0.94%).

Interaction rate [figure 27 below]: Interaction rate, however, does not match the trend of the previous graph, showing a decreasing user interaction.



Total views [figure 28 below] Also here it seems that archives started using the instrument of IGTV precisely in correspondence with the early signals of the pandemic, in January 2020. From that moment on, IGTV reached around 154 millions of views.



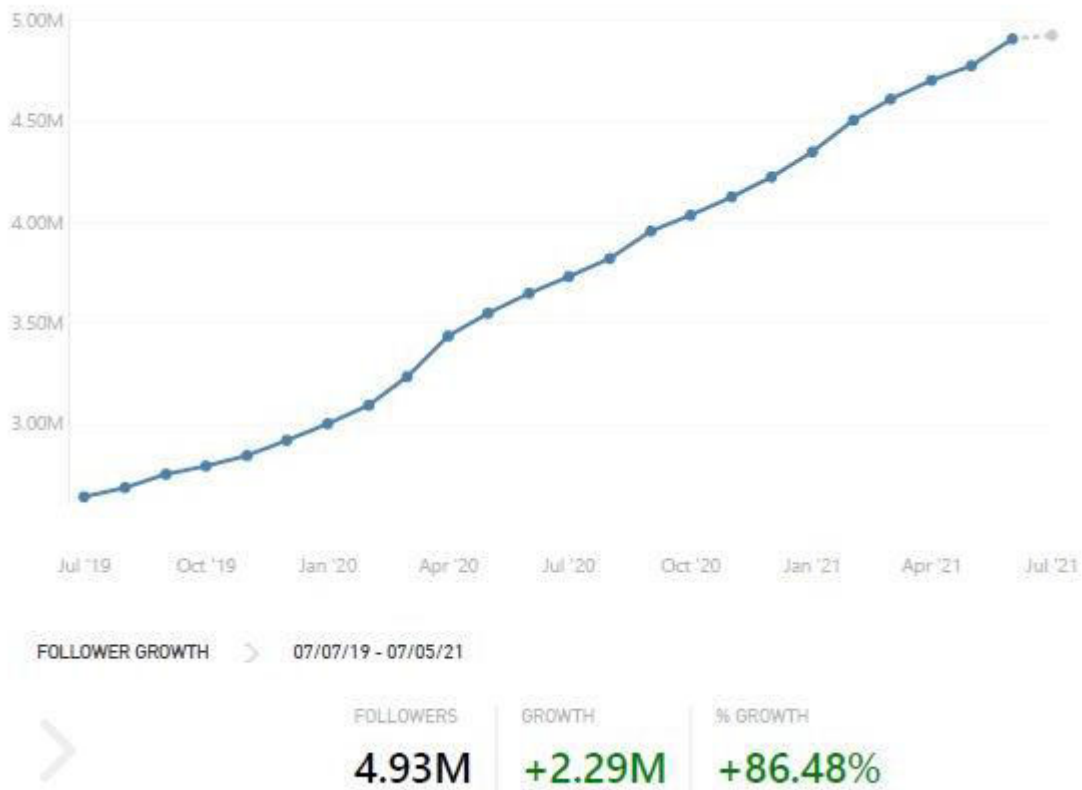
ALL POSTS > TOTAL VIEWS > 07/01/20 - 07/05/21

> TOTAL VIEWS **176.24M** TOTAL POSTS **2,376**

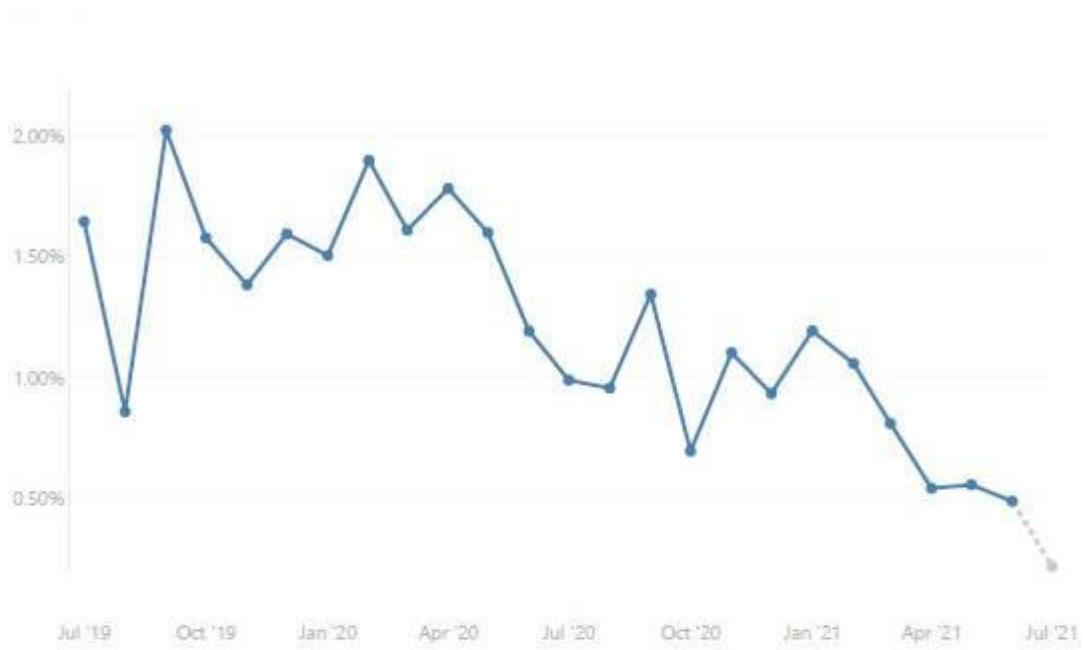
Post Type	# Views	%	# Posts
Albums	18.0K	0.01%	347
Videos	21.80M	12.28%	279
IGTV	154.62M	87.73%	1.8K

3.5.2 Facebook

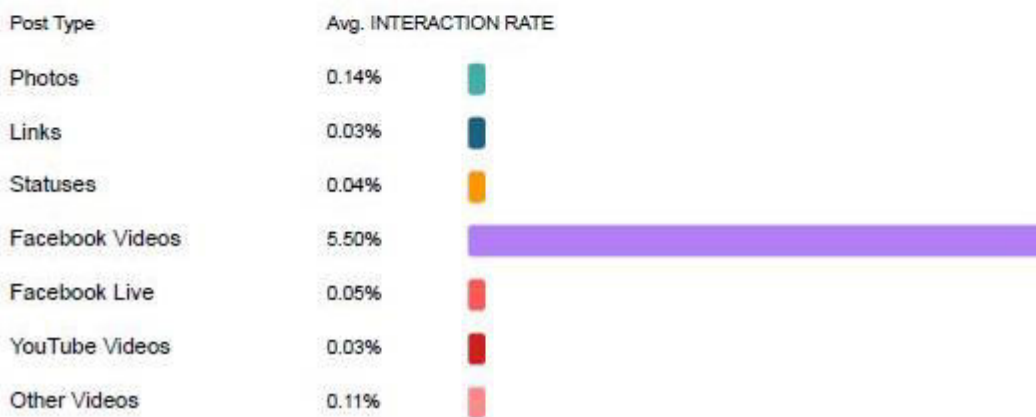
Follower Growth [figure 29 below]: As for the Instagram graph, while the trend reveals a stable growth process, it is possible to see that, after spring 2020, the slope becomes steeper, suggesting an acceleration in the attraction of new users. Indeed, in two years, the archives considered in this case study gained 86.5% of Facebook followers.



Total Interactions [figure 30 below] The Facebook video content is by far the one that engaged users the most, with the highest peak corresponding to the first lockdown period, namely spring 2020. The Facebook videos shared by the Archives triggered more than 41 millions of interactions in two years.

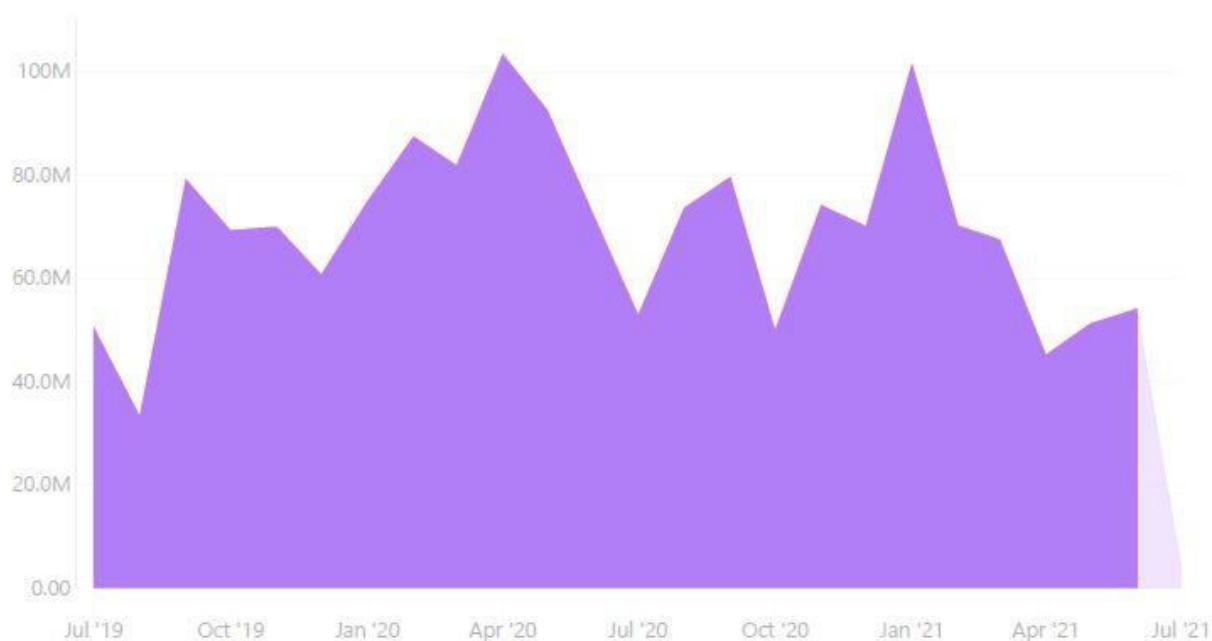


ALL POSTS > AVG. INTERACTION RATE > 07/07/19 - 07/05/21



Total views [figures 32 and 32a below]: The staggeringly high total number of views (1.67 billions) of Facebook videos brought us to unpack the data [figure 31a below] to figure out the reason why Archives managed to reach such a big users pool. In fact, the result was far from homogeneous across the sample; the BBC Archives

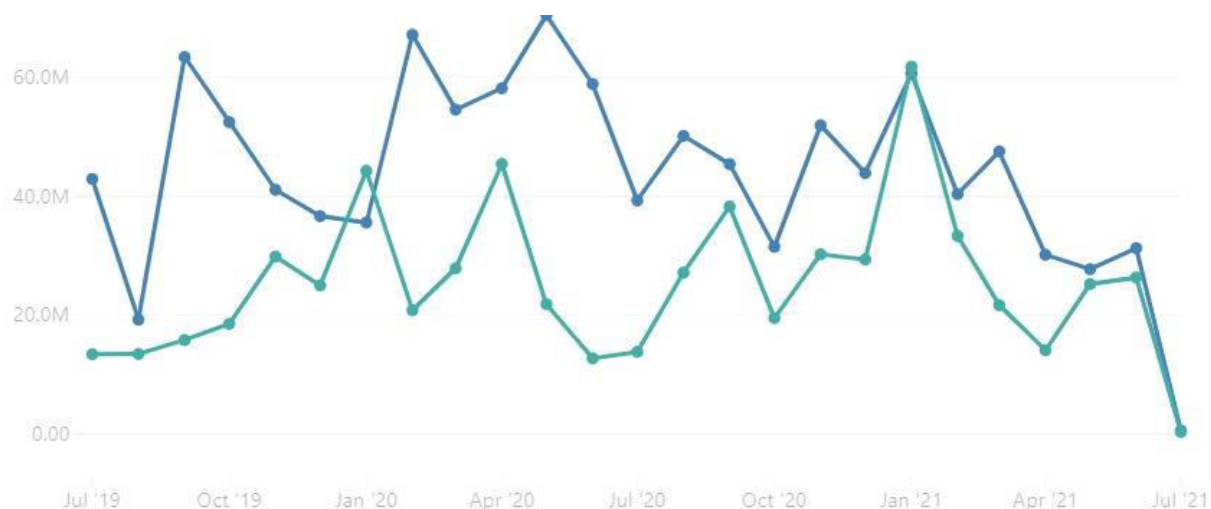
and the INA - Institut national de l'audiovisuel together cover almost the totality of the views.



ALL POSTS > OWNED VIDEOS > TOTAL VIEWS > 07/07/19 - 07/05/21

	TOTAL VIEWS	FROM PAGE POSTS	FROM SHARES	TOTAL VIDEOS
>	1.67B	1.19B	478.70M	7,017

Post Type	# Views	%	# Videos	
Facebook Videos	1.67B	99.99%	7.0K	<div style="width: 99.99%; height: 10px; background-color: #6666ff;"></div>
Facebook Live	137.9K	0.008%	64	<div style="width: 0.008%; height: 10px; background-color: #ff0000;"></div>



ALL POSTS > OWNED VIDEOS > COMBINED TOTAL VIEWS > 07/01/19 - 07/01/21

	TOTAL VIEWS	FROM PAGE POSTS	FROM SHARES	TOTAL VIDEOS
>	1.73B	1.24B	495.40M	5,801

NAME	# VIEWS	PAGE POSTS	SHARES	# VIDEOS
INA	1.10B	821.35M	279.99M	2.7K
BBC Archive	629.82M	414.41M	215.41M	3.1K

3.6 Update on Crawled Web Content Datasets

This section provides an update on the datasets of crawled web content gathered by WLT as previously reported in D1.3.

Major changes include the activation of Italian News content and the integration of further sources from the cultural and creative sector. For this the underlying sources that are addressed during data mirroring have been extended with web sites gathered by the other project partners, including web sites of the institutions that were taken into consideration for the analysis reported in Section 3. All datasets are continuously updated and reflect any newly published content of a source.

News Sources

The news sources have been extended with a set of 46 major Italian news outlets, such as *La Repubblica*, *L'Unione Sarda*, and *Rai News*, to substantially enhance the coverage of the Italian content pipeline. Since May 2021, a total of 236000 Italian News documents have been ingested, led by *Libero Quotidiano* and *La Repubblica* from which 37 750 and 18 770 documents were processed, respectively.

Since the last reporting in March 2021, a total of 11 million news documents were ingested across all languages. Of those, ~18 000 explicitly mention “cultural heritage” and a further 5 000 contain very generic terms related to digital culture (“digital art”, “digital museum”, “creative industry”). In the coming months, more tailored queries will be launched with the help of other project partners to better map the digital cultural ecosystem, and better identify contents as to topics of interest.

Web Sources

Since the last reporting the repository of cultural heritage-related content has been extended to over 1500 inDICES topic-specific web sources from the cultural and creative sector, and further by an additional ~2k web sites of Dutch GLAM organizations as part of a pilot case, where URLs were extracted from Wikidata and manually curated with the help of project partners. The majority of new source URLs were provided by project partners and are associated with different cultural regimes, such as “contemporary art”, “performance/theatre”, “music” or “GLAM”, identified in the work of WP1.

The identified cultural sources were added to the miscellaneous web sample and any newly published content is automatically picked up by the WLT crawling architecture and processed by the content extraction and enrichment pipeline. Resulting contents are added to the Web Corpus (ref. D1.2) and accessible through REST APIs as annotated WLT Documents, or through the inDICES dashboard, where they can be queried and statistics and document metrics can be visualized.

Social Media Sources

For social media sources, a handful of Twitter accounts were identified on top of the key terms used for social media queries, from which content is aggregated. The entity extraction performed on Wikidata for Dutch GLAMs also included information on associated social media accounts that was first reviewed, and respective Twitter and Facebook accounts were included into the set of collected social media. Based on this use case we plan to identify further social media accounts to monitor for GLAMs from different European countries.

Currently, over 100 Twitter accounts of interest have been identified. Since January 2021, a total of 73 million Tweets have been gathered and their text analyzed by keywords, named entities and sentiment. Due to the extension to Dutch sources, 2.7 million of those Tweets are in Dutch. We are hoping to further extend this set to include Facebook and YouTube accounts that publish cultural and creative content of interest across Europe.

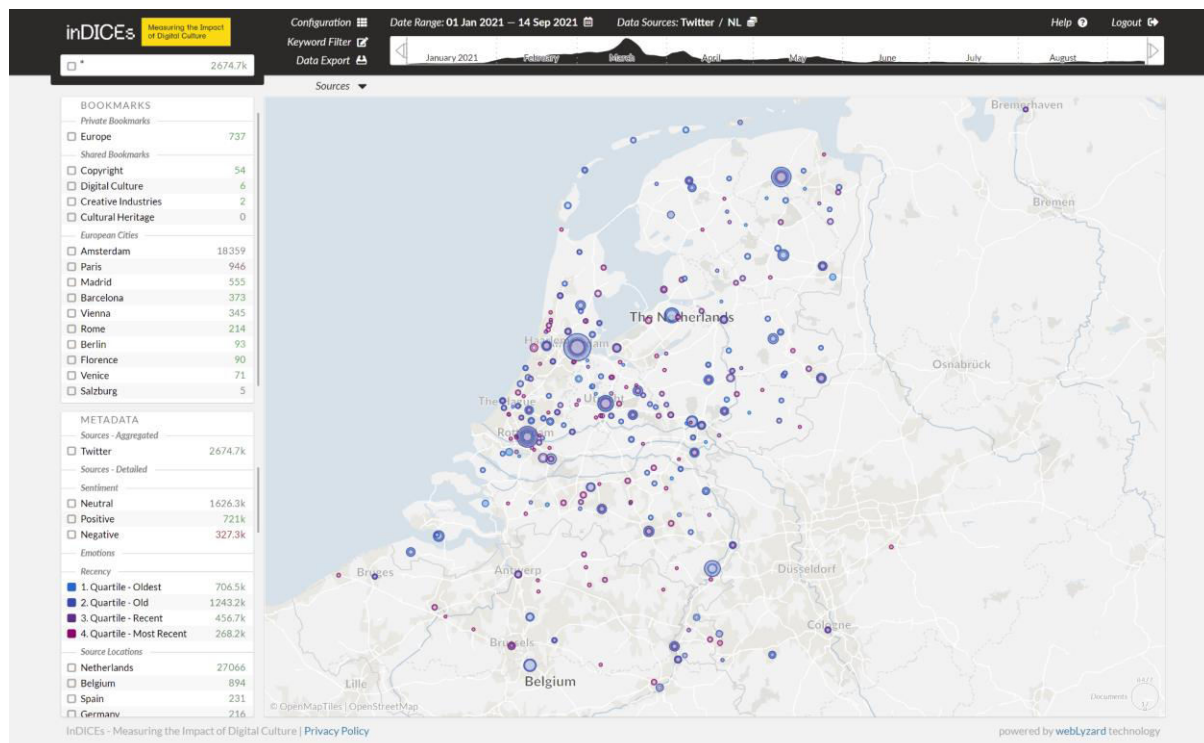


Figure 33: Geographic Map of gathered Dutch Twitter Content between January 2021 and mid September 2021. The circles mark the (long, lat) coordinates of cities, countries, landmarks, etc.

discovered in the Tweets, while circle size indicates the number of Twitter documents pertaining to a particular location.

3.6.1 CHI Knowledge Graph - GLAM Entities

Next to the repositories of metadata enriched documents that WP1 has been gathering from the Web and from social media sources, we have started extending the Visual Analytics Dashboard knowledge graph that is used to store both factual and semantic information about named entities, with entities specific to the inDICES use cases. The knowledge graph is a RDF triplestore, where stored factual knowledge is derived from external and internal sources, and can be accessed throughout the information extraction pipeline that processes the gathered documents. Entities and metadata are initially processed from *Linked Open Data* (LOD) sources such as *Wikidata* for organizations or persons and *Geonames* for locations. Furthermore, events and anniversaries are extracted from both Wikidata and external iCalendar (ical) datasets and integrated into the knowledge graph. As a result, the graph database currently stores over 2.5 million entities from different domains. To appropriately incorporate the focal project topics and improve the metadata annotations in that area, a pilot was conducted to extract entities of Dutch GLAMs from Wikidata. For that, a SPARQL query was defined to extract museums, archives and libraries with their associated metadata attributes (labels, alternate labels, inception date, description, websites etc.) (ref: D4.2). This resulted in a csv sheet of 2194 entities, of which a few were discarded as decommissioned or out-of-date. After a manual review by project partners, additional metadata information, such as further social media accounts, were added to a selected subset of ~50 of the most well-known institutions. As mentioned in the previous sections, the extracted sources (websites and Twitter channels) were added to the data gathering configuration, so that any newly posted content of those institutions is automatically mirrored into the respective repository.

The entities were then transformed into RDF triples using standard LOD predicates, and stored in the WLT Fuseki Apache Jena triple store that retains custom sets of entities. Since the entities were originally sourced from Wikidata as Open Data, we retain their Wikidata identifiers and customize them in our own triple store. Through the use of a dedicated subgraph with an associated graph name, the complete set of triples can be easily retrieved, modified, and extended with additional metadata. Retaining a modifiable copy of the original LOD entry and the triple store architecture it is possible to further add entity-specific metadata to the entities, for example data gathered through the self-assessment tool, that is not currently available in an aggregated form, such as size of an organization or number of employees. As with all our data that is stored as RDF triples, an Elasticsearch layer is employed as a distributed caching mechanism that mitigates the performance bottlenecks of RDF triplestores. The named entity recognition and annotation process was adjusted to include the dedicated GLAM graph into the entity set used for recognition and disambiguation, in order to detect mentions in text and annotate surface forms with the associated GLAM organization. This enables the user to track the online coverage about specific organizations over time.

4. Preliminary observations

The aim of this paragraph is to give a partial and preliminary overview, according to the available data gathered and analysed from Month 15 to Month 21, aimed at detecting the trends regarding the levels of digital cultural participation of the users of the different sectors' institutions, regarding the sectors that we have seen in the “Data gathered to Month 21” section, from 3.2 to 3.5.2.

We decided to focus upon two comparable case studies in the Museums sector: the case of the Most Visited Museums of Europe in 2019 and the case of the Museums of Art with Virtual Tour at a global level. The research question came from the interest in checking whether the Virtual Tour of museums, the digital tool par excellence that allows digital users to interact virtually with the spaces and

collections of a museum or cultural institution, can still be considered an effective tool in the panorama of digital cultural participation, and to what extent it can foster active participation and engagement.

As to the comparison between the two museums' case studies, the most meaningful remarks are as follows:

- Both Instagram and Facebook pages show a lively and constant active engagement with their users before and after the pandemic: the museums cyclically spark new waves of attention by their users, stimulating their participation. It seems that the Covid-19 forced digitization did not damage the Museum's online presence and the capacity to reach their users.
- The Facebook average interaction rate is much higher in the museums with Virtual Tours than in the museums without it, underlining how tailored tools for digital participation have a potential that could be further exploited.
- The percentage of Facebook followers of our sample of Museums with Virtual Tours increased by 12,5% in two years. On Instagram they grew by 40%. The EU most visited Museums increased the number of their followers by only 7% on Facebook and by 51% on Instagram: the latter social network is definitely the preferred channel of the Museum's digital users to keep track with their favorite cultural institutions.
- Museums with Virtual Tour, on Instagram, reveal low active participation: the tool that triggers the most of the reactions and engagement so far is "Photos" with around 100 millions of interactions, but the views (adding videos and IGTV) are around 137 millions. These data, in addition to revealing a prevalently passive participation, show that the Virtual Tour has not yet managed to fully cater for users' engagement with videos or with the video channel in general.
- In general, by comparing the two cases of the Most Visited Museums and the Museums with Virtual Tour, there are no clear trends that reveal that the Virtual Tour instrument is effectively engaging and triggering active participation by social platforms users, failing to take full advantage of the

opportunities offered by digital platforms and resources for the sake of digital community development and empowerment, and of digital co-creation processes.

As to Libraries and Archives, the pandemic has accelerated digital content production and interaction, compared to Museums institutions which maintain similar levels of digital activity throughout the analyzed time window.

As the relationship between follower growth and interaction rate, we can observe that if the number of followers grows and the interaction rate remains stable, this means that the new followers interact in ways that do not differ substantially from those of previous users, and therefore engagement campaigns targeting new visitors have worked. In the opposite case, if the number of followers grows and the interaction rate drops, it is plausible to conclude that the new followers are inactive and only the hard core of already registered and active users continues to really interact. In our samples, despite a constant follower growth, the interaction rate decreases: this may imply a lack of interest from users or a large share of inactive/non-engaged public, which can be related not only to the type of content produced but also to the production practices themselves, that need to become more horizontal and to exploit in a more inclusive and substantial way the potential of the digital platforms. Either way, it reflects the general overproduction of digital content in the pandemic months. Facebook users were probably already engaged and active, which may mean that Libraries, which represent a point of reference for their local communities, maintain a strong relationship with their community also digitally.

The case studies of the Archives and the Libraries present some similarities, such as:

- The pandemic has led to a substantial extra effort in Instagram digital interaction by Archives/Libraries to reach out to their users (compared to e.g. Museums).

- Compared with Instagram, there is a higher interaction with Facebook users, which could be related to the fact that such users are probably older compared to those of other social media.
- We can observe how the peaks of interaction correspond to the lockdown periods, and this correlation can be meaningful if we think about the impact of the Covid-19's forced digitization of both content production and access.
- Follower growth vs. interaction rate: the two lines present an indirectly proportional growth that can signal relative lack of interest or inactive/non-engaged users; Archives/Libraries social network users do not interact actively with the online cultural content shared by such institutions.

As to Archives specifically, the most meaningful observations are as follows:

- Looking at the trend of the graphs, the effort in producing content grew along the time window, and grew when Archives probably became aware of the need to be more propositional (corresponding to the second lockdown, Instagram, IGTV).
- In particular, Archives started to utilize IGTV in correspondence with the early signals of the pandemic, choosing it as the main channel of communication and engagement with their digital users. Although this move slightly anticipated actual lockdowns, its consolidation can be considered a reaction to the Covid-19's forced digitization.
- The Facebook video views reveal very high numbers but the sample is influenced by two leading Archives which by themselves cover the vast majority of the views, namely the BBC Archives and the INA Institut national de l'audiovisuel.

As to Libraries specifically, the most meaningful observations are as follows:

- With much less effort (7.6k Instagram vs 28.4k Facebook total posts) the total growth of Instagram reached 104% compared to 18% of Facebook. We can imagine that the Facebook users were already engaged, or that Instagram is the preferred channel for users to reach out to their National Libraries.

- Even if the “Photo” type of content remains the favorite, Libraries also share a lot of video content.

We have so far observed, in museums and libraries, a drop in engagement associated with the COVID-19 pandemic. The drop in engagement was in particular remarkable for Instagram, whereas on Facebook the effect was less striking. Both cases are focussed on GLAM institutions that are traditionally meant to be physically visited and offer cultural content that is meant to be timeless. We could possibly expect, with the next data gathering activity, that other case studies considering cultural industries less linked to the physical experience and offering more ephemeral experiences would be different. Indeed, Instagram is a kind of social network where one shares "special moments" that are perhaps harder to create in the context of the GLAM institutions than in other industries like, for example, fashion, where content consumption is not traditionally limited by physical access and, most importantly, where the ephemeral nature of the content conjures up in rendering a particular moment "special".

5. Usefulness and Limits

The information of the data analysis that is presented in the present document is likely to be useful for a wide range of potential users and stakeholders in cultural and creative sectors, thanks to the data accessibility as described in deliverable [D1.2](#), through the inDICES Open Observatory, in conjunction with the embedded widgets of the Visual Analytics Dashboard.

For what concerns the limits of these analysis, it is important to underline that our data are aggregated, and consequently they need to be eventually unpacked and parsed more in detail according to the specific interests of the inDICES partners and stakeholders.

6. Future Plans

In the following section, the plans for the next 3 months [M22-M24], in the perspective of the future deliverable D1.5, are summarized. The plans regard firstly the WP1 work, which will be constantly integrated and eventually re-shaped with the collaboration of the WP3 in the preparation of a set of policy recommendations; the WP4, regarding the Open Observatory tool integration (the Visual Analytics Dashboard and the Repository); the WP5 for the academic dissemination of the scientific analysis and results emerged in the frame of the inDICES project; and all the members of the consortium more generally.

The plans are aimed at the full compliance with the D1.2 Data register [M21] objectives, which regard the definitive design of the structure and characteristics of data gathered, and of their sources, and a discussion of their usefulness and limitations, within an organizational scheme that allows their effective accessibility through the Observatory Platform.

The goals which inherently pertain the WP1's next 3 months work plans are:

Goal 1: identify a small group of experts within the inDICES partner consortium who will define the missing case-studies lists about Cultural Heritage Institutions from further sectors, namely (among others): European Theatres, European Archeological sites, Fashion GLAMS, a case study list regarding the macro-sector of the Audiovisual (whose criteria have to be collectively defined via the inDICES internal Hypothesis Assembly) and integrate them within the Visual Analytics Dashboard.

Goal 2: as per the case-studies lists present in this deliverable, these lists will be analysed from the perspective of the users' behavior, users' active or passive participation, the relation between them and the CHIs digital performances on Facebook and Instagram (with a specific focus on the Covid-19's forced digitization impact) and will be eventually unpacked and inquired according to the specific interests that will emerge from the inDICES partners and stakeholders.

Goal 3: inquire the existing lists and the collected web data via the Visual Analytics Dashboard with [queries](#) that have been already partially discussed, proposed and defined by inDICES partners and experts, in order to implement:

- trend analysis for various CCS through web posts;
- impact analysis of specific social campaigns;
- impact analysis of CCS with respect to specific socio-cultural trends;
- investigating socio-economic impacts on some different CCS;
- network analysis and mapping of cultural digital platform ecosystems;
- psychosocial analysis of web content in various CCS.

A prospect of content categorization will be necessary, where cultural content sources will be assigned more specific categories based on their sector, in order to identify them by area of interest and filter respective content.

Goal 4: definition, drafting and academic dissemination in the form of scholarly papers on the scientific analysis and results obtained in the frame of the inDICES project.

7. Conclusions

The analyses conducted during the two-year period between 2019 and 2021 generally show a significant increase in the online activities of the social media pages of the cultural institutions surveyed. In particular, the pandemic and the forced closure of physical activities seem to have been particularly beneficial to the growth of the followers of several institutions that worked on their digital outreach. From this point of view, however, it is not yet clear if in the long term the effect will stay positive or, instead, it will eventually die out. Some signs of a likely continued growth can be seen, but further monitoring is necessary to understand more. Another related point concerns the heterogeneity of the strategies adopted by the various cultural institutions: it is not clear whether the entire sector has benefited from this push or whether some top players have reaped more significant benefits than others. Finally, a last point is related to the real uptake of the Culture 3.0 production regime, which still seems to be in its early phase. It is necessary for the CHIs to better enable people to actively participate in meaningful sense-making processes, to exploit the possibilities that the digital platforms can offer in terms of

co-creation processes, digital community empowerment, development of new soft skills and shared knowledge resources. The low rate of active interaction of users with digital cultural content and the still too low level of engagement with cultural institutions can reflect forms of cognitive poverty and deprivation, which often strongly correlate with other socio-economic deficiencies. The real challenge for CHIs is to fully tap into the extraordinary opportunities offered by digital platforms and resources in order to impact and to contribute to the progress of civil society for the sake of human development, quality of life, and social empowerment.

In this sense, inDICES will undertake to constantly monitor trends and understand where the digital users migrate, detecting the constantly evolving trends in terms of cultural digital participation.