



Article

# Social Media Analytics as a Tool for Cultural Spaces—The Case of Twitter Trending Topics

Vassilis Pouloupoulos <sup>1,\*</sup>  and Manolis Wallace <sup>1,†</sup> 

Knowledge and Uncertainty Research Laboratory, University of the Peloponnese, 221 31 Tripolis, Greece; wallace@uop.gr

\* Correspondence: vacilos@uop.gr

† These authors contributed equally to this work.

**Abstract:** We are entering an era in which online personalities and personas will grow faster and faster. People are tending to use the Internet, and social media especially, more frequently and for a wider variety of purposes. In parallel, a number of cultural spaces have already decided to invest in marketing and message spreading through the web and the media. Growing their audience, or locating the appropriate group of people to share their information, remains a tedious task within the chaotic environment of the Internet. The investment is mainly financial—usually large—and directed to advertisements. Still, there is much space for research and investment in analytics that can provide evidence considering the spreading of the word and finding groups of people interested in specific information or trending topics and influencers. In this paper, we present a part of a national project that aims to perform an analysis of Twitter’s trending topics. The main scope of the analysis is to provide a basic ordering on the topics based on their “importance”. Based on this, we clarify how cultural institutions can benefit from such an analysis in order to empower their online presence.

**Keywords:** social media analysis; cultural spaces; cultural informatics; trending topics



**Citation:** Pouloupoulos, V.; Wallace, M. Social Media Analytics as a Tool for Cultural Spaces—The Case of Twitter Trending Topics. *Big Data Cogn. Comput.* **2022**, *6*, 63. <https://doi.org/10.3390/bdcc6020063>

Academic Editors: Konstantinos Kotis and Dimitris Spiliotopoulos

Received: 7 April 2022

Accepted: 26 May 2022

Published: 2 June 2022

**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Nowadays, many people use online services for business or for leisure. This phenomenon has grown to the point of being a huge part of everyday life for many people around the globe, especially as a result of the COVID-19 pandemic’s lockdowns. On these grounds, people have started to be concerned about their online profiles, having even created personas. The Internet and its services became a main stream of socialization at some points, as people were unable to physically meet in person. Social media platforms, which have already gained a lot of attention, became the main communication channels, and teleconferencing services replaced daily meetings.

Under these strange circumstances, many businesses changed (and keep changing) their business models, starting to empower their online presence; some of them pivoted to online only. A large number of people around the globe are working remotely, and more and more companies are announcing to their employees that they will continue with the “remote work” model for the coming months at least. However, a number of companies, including culture-related spaces, have a business model that is based on the presence of their audience (visitors). Cultural organizations are in the strange situation of their “customers” desiring the physical presence of culturally significant objects (say, in order to enjoy a piece of art), despite the fact that it is prohibited to reach or even approach the cultural objects.

A large number of culture-related organizations have already turned to the Internet as a space for message spreading and audience acquisition. They have also invested in cutting-edge technology in order to offer unique experiences to visitors. Technology has also helped to perform interdisciplinary research in the field of culture, giving birth to

museum and cultural informatics. The EU has already made a very large investment in culture and cultural heritage, by providing funds for research projects, having as the apogee the announcement of 2018 as the year of cultural heritage. It is interesting that an important number of the funded projects are related to information technology, either related to cultural heritage preservation or museum informatics. The main axes of technology in culture are related to digitization, virtual reality, preservation, semantic analysis, crowdsourcing, networking and IoT (Internet of Things). This is proof that cutting edge technologies are widely used in modern museums, enabling maximization of the visitors' stimulation. Modern cultural spaces are trying to take advantage of innovation and new techniques and technologies. Still, a large number of them remain attached to classical research and methodologies; without calling this the wrong side, concerns about the world changing must be taken into consideration seriously.

The pandemic has slightly changed the point of view for a large number of businesses, including cultural spaces in general. They understood that there is a strong need to support alternative ways of spreading their messages and attracting people. Moreover, physical presence is now considered "risky" in several places, leading to solutions that include technology and "remote access". In parallel, during the lockdowns of the COVID-19 pandemic, people realized that culture is part of their everyday life. A number of cultural spaces in Greece started providing open access to culture-related content. The Greek National Opera announced a number of online performances during April 2020 (<https://www.nationalopera.gr/en/news-features/item/3117-ministry-of-culture-and-sports-the-gno-free-broadcasts-continue-with-great-success> (accessed on 30 May 2021)). The content was presented live, using platforms such as YouTube and Facebook, and was available for free access some hours after it finished its live broadcast. On other occasions, theatrical shows were broadcast online at a voluntary price. The aforementioned were organized within a small time frame from the moment the quarantine was announced, meaning that the technological infrastructure and readiness are at a very high level. The majority of online performances from famous cultural organizations reached sufficient levels of spectators (online users), but still, the information reached the people that were, and are, searching for plays regardless of the lockdown due to the pandemic. The possibility of reaching a larger audience, especially at a time at which there seems to be a paradigm shift regarding the engagement of people towards arts and culture, remains at low levels. The readiness level in alternative presentations is very high, but still, there are very low expectations regarding the audience that is reached and is convinced by the modern communication channel.

Efforts towards discovering the audience of a business in general have been the subject of research ever since the Internet became popular. Once people were using the Internet for socializing, a wide variety of businesses started turning to deep marketing analysis on the medium. The cultural organizations were late to adopt such technologies and techniques. However, even today, with the extensive usage of technology in our everyday lives, so much so that in a sense we are considered as online or offline in our lives, cultural spaces have not yet adopted the kinds of technology that will help them reach broader audiences and provide information about their content and messages.

In this paper, we present how social media analytics, as part of the Greek National Project "PaloAnalytics", can be helpful for cultural spaces. We focus on the analysis of Twitter's trending topics and relate its outcomes to cultural spaces in order to prove how they can benefit from the simple correlation of their content with the analysis that can be performed to social media. "PaloAnalytics" is a nationally funded project that intends to focus on the basic challenges that organizations face and is related to the implementation of a universal monitoring tool with links and interconnections between collected data. It is analyzed by a number of different researchers in various languages. The project as a whole can be a useful tool for large companies, but select parts of it, implemented within the scope of the project, can be beneficial for culture-related organizations. On these grounds, we selected a module that can be used as a stand-alone tool and presented how it can be

used so as help organizations put the focus of their online efforts into specific actions and posts, in order to achieve higher penetration to the medium and better public awareness.

The rest of the paper is organized as follows. Section 2 presents research performed in the field of cultural informatics and social media analysis. The next section gives detailed evidence on the algorithmic approach of the proposed solution. Section 4 presents the modules that were implemented from a technical perspective and the flow of information within the system. Experimental results are presented in Section 5, and the paper concludes with Section 6 that presents the discussion and future work.

## 2. Related Work

From a research perspective, the combination of technology and culture was initiated from what is called “museum informatics”. The research in this field started more than 40 years ago, with Bearman being a pioneer trying to empower the use of technology, at least in relation to archives and internal organizations ([1–3]). Initial efforts dealt with databases and cultural object registration and internal organization of cultural spaces and definition of prototypes for the classification of objects, mainly in museums ([4,5]). However, as technology evolved, more and more “technological advances” were found in cultural venues. Digitization and digital collections ([6–10]), visitor participation ([11–13]), ByoD ([11,14,15]), or AR and VR ([16–21]) as part of a museum’s procedure are only some of the examples of early or later adoptions. Nowadays, the interdisciplinary research that is performed in culture and technology intends to bring together cultural spaces with the latest advances in technology. P.F. Marty has performed extensive research in the area, covering all the bases that lead to the aforementioned and will formulate the future of culture in relation to technology ([22–24]).

Talking of which—the future of culture in relation to technology—we should note that it seems that people tend to have a mixed type of life including a lot of digital alongside with the “analog” part. In this sense, extensive research is performed considering the role of museums and cultural spaces in the online world. More specifically, we examine how the problem of a museum’s presence in the world of social media can be tackled.

Research on behaviors and interactions is common, usually in relation to social media. Research on Instagram accounts related to art was performed and it seems that “likes and comments are greatly influenced by interactions with confusion and curiosity being a big reason to engage” [25]. This reveals that “active participation” plays an important role in social media presence. It is, on the other hand, interesting to examine all the issues from several different angles when dealing with “data-driven” arts, as there is always the possibility of “opportunities or chimeras” [26]. In this work, the authors conclude that in an era that is filled with a plethora of data, following a data analytics approach would benefit ACOs (Arts and Culture Organizations). Research analysing the situation in relation to Greek museums proved that, despite the fact that a main stream of the museum message is based on images of the exhibits, only a few museums use Instagram. At the same time, a large number of users tag museums and cultural objects, without any kind of interaction from the cultural spaces [27]. Many concerns regarding the stance of the museums and cultural spaces were raised during the COVID-19 pandemic. Democratization of the procedures of the museums is also a novel approach in the world of the web [28]. In this research, the authors explore the unfulfilled potential for democratizing museums by exploring aspects of Instagram (Instagram—<https://www.instagram.com/> (accessed on 30 May 2021)). Their findings show that museums are using the medium in a way that is not attractive, as they keep a more “traditional promotional” stance or “an authoritative knowledge-telling attitude”. The aforementioned mean that the museum has to become participatory and inclusive in alternative ways.

Inspiration among visitors is a factor that museums are really interested in, and research was performed on this issue with regards to data analysis on Twitter [29]. Within the aforementioned research, we implemented prototype tools to collect information from the medium, in order to find “expressions of inspiration in Tweets”. From the findings,

there is an indication that “social media may have some potential as a source of valuable information for museums, though this depends heavily upon how annotation exercises are conducted”.

A literature review on the usage of social media by museums is presented in [30]. The review leads to the assumptions that the museum communication can benefit with the use of social media and that the educational role of the museum can be enhanced. Furthermore, it is stated that “beyond social media effectiveness, museums managers lag into dialogical communication”. This proves that there is a strong need for tools and methods to be provided to museums and cultural spaces, in order to have a differentiated view of the data in social media. Digging more for specific use cases in museums, it seems to be complex for small organizations to have a decent presence in social media, but another interesting aspect of the usage of social media in museums can be revealed [31]. It seems that “evidence that the museum staff acts as ‘social media champions’ represents a qualitative indicator of an increase in the employees’ commitment and loyalty to the organization”. This reveals another part of the online presence of the museums in regards to the engagement of the employees. The interactive and participatory presence of museums in social media gives the opportunity to the employees to be a part of the communication strategy, which is not possible otherwise.

Social media analysis is not something new. A study back in 2012 tried to present the uses and evaluations of social media in American museums [32]. At that time (10 years ago), the results “indicate that American museums believe becoming involved with social media is important, but they are not using the sites at high levels of dialogic engagement”. The latter reasoning remains a problem for cultural spaces till today. It takes huge effort to decide on which part of social media to “invest time” into so as to maximize the performance on the respective platforms. In this case, we are trying to help museums focus on specific topics that are already trending alongside the medium. Another case, the one of the C  a Valley Museum and Archaeological Park, is presented in [33]. From the analysis, it seems that as long as museums stand for “guardians of human memories”, the author believes that “museums have an important role to play in providing widespread access to their collections, facilitating scientific research and fostering its use for educational, leisure or recreation”. The case of ephemeral storytelling with social media was researched in [34]. This research indicates that the usage of the “stories” feature of specific social media are a means to engage a larger audience, and they conclude that “museums should adapt their policies and programs to current social media communication behaviors to remain relevant and be a part of what and how people share their lives”. It is obvious that the culture-related spaces need to be aligned with contemporary trends. More recent research on the part that is related to nowadays tech-savvy audience reveals that “Tech-savvy tourists will enjoy and appreciate the overall digital experience, thus identifying themselves with and feeling part of the museum, becoming loyal, and willingly providing economic support” [35]. In fact, it is obvious from a number of research outcomes that at the end of the day, museums have sufficient online communication, but there seems to be zero interaction [36]. From a number of researches regarding what is expected from the museum in its online presence today, the results remain the same: the museum has to interact and participate.

The aforementioned research proves that the community that researches cultural informatics is aware of the power of social media and the stance of people towards technology. We live in an era where people tend to be more and more digital (or online). In this light, the use of social media seems to be a one-way road for the cultural spaces.

As a matter of fact, using social media in general and having a decent profile means three things:

- Providing detailed information;
- Performing information dissemination;
- Participating.

While the first two things necessary in order to maintain a decent profile are easy to understand, the third one remains a difficult issue. In the case of a cultural space

using social media, providing detailed information means having a complete profile, with correct and precise information considering the communication, information about the organization, working hours, and booking tickets and online shopping. These features are required by the end-users when visiting a social media profile or page of a culture-related space. Considering information dissemination, it is expected that social media presence is frequently enriched with information about the actions of the organization. Usually, one can find information about exhibitions, objects, educational programs, visits or special dates. The very fast pace of the data in social media implies that this dissemination of information has to be frequent. Due to the fact that it is impossible to have very high frequency of information dissemination, another factor is proposed for the cultural spaces. Disseminating information in a cultural space means a long time of preparation, and as such, it is difficult to achieve disseminating information very frequently. In this work, we propose the factor of participation as a key factor in order to keep the cultural space up to date (as a term in social media) and consider its appearance as modern and synchronous, while in parallel keep “appearing” more in front of people on social media in order to spread the message they carry as information carriers. Our proposal is the utilization of systems that are able to perform alternative types of analytics on social media big data in order to help cultural spaces take advantage of the information created.

### 3. Algorithmic Approach

In this work, we present a novel method of social media analytics and, more specifically, the case of Twitter analysis within the scope of a Greek national project entitled PaloAnalytics. The idea that lies behind the analysis is that Twitter provides detailed information about its trending topics per area. In parallel, people tend to put their focus on trending topics, usually described as viral. Virality, as a keyword, is related to the Internet today and seems to be the life boat for both people and social media in the chaotic world of data! In order to take advantage of the fact that Twitter provides direct access (through an API) to information related to posts and users, we put the focus on the following procedure and facts. Cultural spaces, as already mentioned, have the problem of participation. In the chaotic world of data on the Internet, it is very difficult to decide where to put the efforts and time for participation. The idea is to find a means of trending topic classification in such a way so as to denote the trending topics that are mostly related to people that use them to participate in serious conversations. This seems to be a good start for a museum to initiate a conversation in order to achieve message spreading and audience acquisition and engagement. Firstly, trending topics related to a place are provided by a direct API call. The API is said to be caching the results for at least 5 min, which means that searching for trending topics very frequently is not possible. Nevertheless, a frequency of 3 times per hour is considered to be enough for our experiment in order to discover and analyze trending topics in time. Secondly, the users’ behaviour towards trending topics is more or less as follows: most people tend to react to trending topics, using either the “like” or “retweet” features of the medium. As a matter of fact, a retweet is somewhat more powerful than a simple like. Furthermore, there are different user profiles on Twitter: people whose profile includes trending topics with a comment, which are usually liked and retweeted, and people whose profile includes mainly retweets of trending posts, produced by the first category of people. Others are just interactive with the aforementioned categories (like or list). Inevitably, a number of other behaviours in the medium exist, but as long as our source of information is trending topics, in this work we will not analyze them. Moreover, by checking the profiles of the aforementioned people based on the followers/following metric, we assume the following four alternatives:

- People with a large number of followers and small number of following;
- People with a small number of followers and large number of following;
- People with small number of followers and following;
- People with large numbers of followers and following.

As a matter of fact, people in the fourth category seem to be rare and did not come up as users of our experimental procedure. By focusing on people of the other three categories, we assume the following generic comments:

- People with a large number of followers seem to be generally more affecting and try to have an account (profile) that has more “original” posts, or at least they write a personal comment even if they are reproducing information.
- People with small numbers of followers are generally reproducing (retweeting in our case).
- Among people with similar post rates, people posting more comments (mentioning) usually have larger number of followers than people that usually do not post comments.
- People with large numbers of followers follow people with large numbers of followers as well, but it is very rare to follow people with a small number of followers.

Research is performed on the exploration of the different user types and user profiles on social media and how they may possibly affect other people ([37–42]). The authors have already presented an analysis on the personalities of social media users in [43,44], where detailed information about the profiles of influencing personalities is presented. In particular, the research work concludes with the assumption that a mixture of an Influencing and Dominant user profile (according to the DiSC personality test [45]), with a writing style that includes viral topics and excess sentiment, seems to be the most influential user type.

At the end of the day, despite the fact that one is presented with the information related to people that she or he follows, it is more than clear that a “Twitter wall” is filled up with information deriving from trending topics and users that seem to be “influencers”. This is not to be a source of blame, as trending topics are topics that are flooding the network, so they will inevitably be present in several “walls”. In parallel, people that are interacting more, either by posting more things or by being followed by more people, or by mentioning people in their posts, are more likely to come up in a wall.

Taking the aforementioned into consideration, we examine the trending topics of the medium so as to extract information on how they could affect the behaviour of cultural spaces in their online presence. As trending topics are topics that in the end will come up on one’s dashboard, it is something to follow in order to recognize and uncover online places for action and message spreading. In fact, in cases where a cultural space needs a place to start its interaction, finding trending topics that can initiate serious conversations is the right place.

As the role of the museum nowadays is changing, the cultural spaces have to face the reality of the digital world accompanied with its rules and culture. As a matter of fact, it seems like the cultural spaces have to act immediately and decisively so that they will change the “flat” culture of the web. Social media is a place to interact and present the alternative view of culture. Our research is focused on recognizing trending topics deriving from people with high “influence” value in order to discover ways to interact.

The idea of our approach is the alteration of the analysis of the trending topics of Twitter, so as not to rank them only according to their volume (which seems to be how they are presented), but add some qualitative features related to a “score”, which is assigned to users who post data about trending topics. The idea behind the scoring of trending topics is to help museums realize which are the topics that are more probable to be parts of serious conversations.

On this occasion, trending topics inside a “retweet” by a user with low number of followers should score less than trending topics inside an original “tweet” by a person with large numbers of followers. According to the aforementioned idea, we conclude with the procedure depicted in Algorithm 1.

**Algorithm 1** Twitter trending topics analysis.

---

```

1: procedure TRENDSFETCH(woeid)
2:   trends ← fetchTrendsAPI
3:   while trends do
4:     tweetssearchForTweets(trend)
5:     for hasMoreTweets(tweets) do
6:       tweet ← TweetInformation
7:       save(tweet)
8:       user ← getUserInformation(tweet)
9:       save(user)
10:      userPower ← calculateUserPower(user, tweet)
11:      save(userPower)
12:    end for
13:    trendPower ← calculateTrendPower(userPower, trend)
14:    save(trendPower)
15:  end while
16: end procedure

```

---

The idea is to assign a “power” to each trend, according to the power deriving from the posts that include the trending topic and the “power” of each user of the medium.

In order to measure the power of the users, we utilize metrics such as:

- Number of posts;
- Number of followers;
- Number of following;
- Time registered to the medium;
- Is the user verified or not;
- Retweets of posts that include trending topics;
- Original posts that include trending topics.

The following algorithm provides a power for the user.

$$userPower = verified * (a * (followers/statuses) + b * statusFrequency + c * (followers/following)) \quad (1)$$

$$userPostInfluence = trendingTopicPosts/numberOfPosts \quad (2)$$

$$userPostPower = retweetTrendingPosts/numberOfPosts \quad (3)$$

The first equation utilizes four different metrics. The first factor is related to the user verification. If the user is verified, then the corresponding parameter is set to 1.2, else it is set to 1. On this occasion, people with verified profiles have 20% more power than people without verified profiles. The second metric is the rate of followers based on the number of statuses (posts). This metric provides evidence on the rate of people following according to each tweet posted. The third metric is the status frequency, which is a metric providing information on how many tweets are posted by the user within the time the user is using the social media account. Finally, another factor is the ratio of followers to following. In general, this ratio provides evidence on the number of followers related to the number of users being followed by the user. The factor can have a large variety of controversial values; as such, it is calculated according to the following:

$$followers/following = \begin{cases} 0 & < 1 - \text{the user is following more than their followers} \\ 1 & 1 \leq x \leq 10 - \text{the user has at most 10 times more followers than following} \\ 2 & > 10 - \text{the user has more than 10 times more followers than following} \end{cases}$$

The parameters a and b are weights on how much each of the two aforementioned parameters should count in the final result. According to experimentation, it seems that a ratio of a/b of around 4 provides a balance in the results. Parameter c depicts the weight that should be used for the followers/following rate. In cases where we need to focus on this ratio, the value of c should be close to 30. A generic value should be at around 15.

Equations related to postInfluence and postPower are used in order to recognize each users' post penetration within the medium. These metrics are continuously updated leading to differentiations of the total user power.

As a second step, the qualitative parameters of userPower together with postInfluence and postPower are used when reevaluating the power of trending topics. Each topic has a number of posts in which it appears, and each of these posts has a number of "likes" and retweets. Both are indicators of the trending topic power. In parallel, each of the retweets is performed by a user, whose power is already evaluated and continuously recalculated. As such, the following equation is used in order to measure a postPower.

This procedure implies that from the trending topics retrieved from the medium, we search for original posts (no retweets or mentions fetched) and we re-evaluate the "power" of the posts according to several factors. Each time a trending topic is mentioned in a tweet, its "power" is re-evaluated following the algorithm that is presented.

$$postPower = (a * totalShares + (1 - a) * totalLikes) * userPower \quad (4)$$

In this way, each trending topic fetched is provided with a score, which is the sum of the power of the post. It is updated every 20 min, which is the time interval of the system fetching trending topics. The following section presents the complete system architecture in order to support the recorded data and experimental results on the system execution, and how the results of the system can lead to beneficial results for culture-related spaces.

#### 4. Architecture

The system is based on modular architecture, so as to be able to create each system independently. Figure 1 presents the subsystems of the implemented solution.

##### 4.1. Flow of Information

The designed system utilizes data fetched from Twitter. The fetched data, which include information about trending topics, users and tweets, are directly stored in a relational database, while in parallel, new data are calculated. New data include an assigned user power, the user Post Influence factor, the user Post Power variable and a calculated trending topic power. These metadata are stored in parallel with all the original information fetched from the medium. This information is stored both in the relational database and in a time-series database and a document-based database. After data are stored, they are analyzed in order to be combined and filtered, and they are then provided as grouped information through a Visualization tool (Grafana (Grafana: The open observability platform, <https://grafana.com/> (accessed on 30 May 2021))) and a RESTful API (built with Laravel (Laravel: The PHP framework for web artisans, <https://laravel.com/> (accessed on 30 May 2021))).

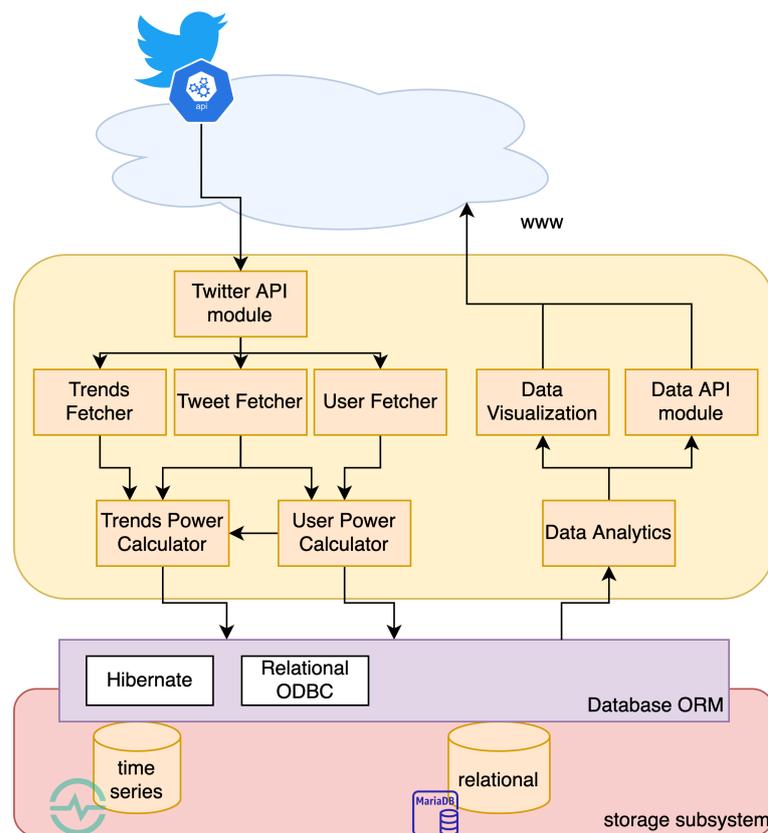


Figure 1. System Architecture.

#### 4.2. Storage Subsystem

The database subsystem consists of three different database management systems and is supported by a database ORM. The storage consists of a time-series database for the storage of information to be visualized in real time, a relational database for permanent storage of all fetched and analyzed data, and a document-based database for storing information to be easily analyzed by taking into account semantic information. Detailed information on the standards of data storage is out of the scope of the present work.

#### 4.3. Data Fetcher

The data fetcher is a system that is responsible for fetching the required information. Information is fetched by utilizing three Twitter API calls (Twitter API—<https://developer.twitter.com/en/docs/twitter-api> (accessed on 30 May 2021)). The first system is responsible for fetching the trending topics. It utilizes the Trending Topics API, which returns a list of “keywords” as the trending topics in a specific area. After the trending topics are received, each of the topics is used as the input to a Twitter Search API endpoint. This returns a number of posts that are relevant to the trending topic. From this procedure, any retweet or mentioned post is omitted, so that only original posts including the trending topics are saved. Finally, for each of the tweets, information about the user that posted it is extracted, followed by an amount of metadata including number of followers, number of posts, if she or he is verified, etc. All the data deriving from the data fetcher are saved to the relational database.

#### 4.4. Power Calculators

The data that are fetched from Twitter are analyzed in order to evaluate a so-called “power”. The purpose of this re-evaluation is to measure the strength of a trending topic among the users of the medium. The strength can be considered as the ability to spread in the community. This power is already visible by the volume of tweets containing the

trend, but is mainly dependent on countable factors such as the number of likes or retweets. We put our focus on adding another parameter in order to create an algorithm for power evaluation, which is the power of each user. Information about the algorithms is discussed in Section 3. User and trending topic power evaluation leads to storing metadata in the database that accompany the medium's information.

#### 4.5. Data Analytics

Information that is stored in the three different types of databases is analyzed in order to achieve the next step, that is, visualization and access to the structured data through a RESTful API. The analytics include procedures that combine information in trending lines. For example, as long as the fetching algorithm fetches trending topics every 20 min, the trending topics reoccur over time and have to be combined in order to form a "trending line" of information that evolves in time. Another type of analysis that is performed in order to feed mainly the RESTful API is searching for different behaviours in the trending topics. The analytics lead to three different types of trending topics. The first type is topics that appear, have a quite quick time with a high level of power and fade out within a finite period of time (usually within some hours, which is computed as around 12–16 for Greek Twitter). The second type are topics that appear in a very short period of time (usually an hour or less), achieve a low score, and disappear. Finally, the third type is trending topics that either remain present for a long period of time (more than 36 h), or appear in a periodical manner (e.g., every day, the same hour, with the same behaviour). The aforementioned findings can be found both in the visualization and in the RESTful API. For each type of topics, there is evidence on how a cultural space should act. Starting from what is easy to understand is that the second type of posts act like comets. Their appearance is short, fast and superficial, which means that one should not spend much time on these topics. The third topics, the periodical ones, are related to a recurring event, usually a TV show. These topics are not to be given much attention as well. On the other hand, given the fast pace at which the Internet moves, and more specifically the data on the Internet, attention must be paid to the trending topics that remain active for a medium period of time. These are topics that remain active for a period of a day and receive a lot of attention by several different types of users. The analysis performed provides information about such topics and can guide a cultural space to engage with them.

#### 4.6. Data Presentation

The system concludes two types of data presentation in order to support the part of our experimentation that deals with the connection to cultural spaces. Both the data visualization module and the API are used in order to obtain a presentation of the analyzed information and support our findings, in order to enhance the presence of cultural spaces on the web, and specifically social media. Visualization is achieved with the help of Grafana, in which a dashboard including different kinds of real-time information is set up, and the API, which is provided with the help of a Laravel installation that holds the RESTful API endpoints. Both of them are presented in Section 5.

### 5. Experimental Results

The experimentation procedure is separated into two different parts. During the first part, we examine the use of the system and the results from the algorithms' application to the data. We present the results from Twitter trending topics as they are fetched from the corresponding API (without information) and how the system adds metadata for extra analysis. During this procedure, we examine how we utilize the extra information in order to offer data visualization and endpoints from the RESTful API. The second procedure utilizes results from the visualization procedure in order to show how the information presented can be used by a culture-related space in order to provide a benefit during their online presence.

### 5.1. Trending Topics on Steroids, Ordered

The first part of the system collects information about trending topics, tweets and users in order to enhance the information related to the trending topics. When data are fetched from the medium concerning the trending topics, they include a lot of accompanying information according to Table 1.

**Table 1.** Example of original Data for Trending Topics (place ID = Greece).

Name	Time	First Appearance	Volume
#covid19gr	16 October 2020 18:10:05	12 April 2020 09:10:05	null
elon musk	5 April 2022 01:30:04	22 October 2020 19:30:03	233,504
ukraine	3 April 2022 20:10:04	25 February 2022 06:10:02	1,411,402

In general, information from the original endpoint includes the name of the trending topic, a URL to search for it, and the impact it appears to have within the medium, if the latter is available. In fact, impact is available for a small number of trending topics, especially in the place of Greece where the system was mainly tested. As such, this information is not taken under consideration in our procedures.

Respectively, after retrieving the tweets related to each trending topic, information about the users are retrieved as presented in Table 2. The information presented are the ones that concern our algorithmic procedure, while original data include more information.

**Table 2.** Sample of original Data for Users.

Name	Verified	Followers	Following (Friends)	Favorites	Statuses	Created At
Black H_	false	977	2136	7	5	7 July 2019 12:50:34
Matthildi M_	false	28,048	8549	1536	495	25 May 2011 19:55:00
Nick G_	false	13,582	9	100	261	30 September 2013 21:08:23

The table does not present users with small numbers of followers, as these users are by default given a very low userPower according to the described algorithm.

The data that we retrieve are enriched with more information, that is, the userPower, which is calculated according to the algorithm presented in Section 3 and the number of trending topics that the user has posted, and the user post power and the user influence power. The parameters are recalculated every time a user’s post is returned as a result in the search-for-posts procedure.

According to our algorithm, the aforementioned sample profiles receive a userPower score. The first profile presents a user with a very low number of posts (statuses), a large number of followers (in comparison to the statuses), and a very high number of following. It seems like a typical profile that has a large number of followers both from the friends (following) and from another factor that is “hidden”. This is the type of user that comments a lot, which is the reason that the user has quite a large number of followers. In our procedure, we omit any statuses that are related to conversations and are not original posts.

The second profile has a significantly large power due to the large number of followers. The large number of followers is achieved both by the large number of friends (following) and by the favorites, which reveal very high interaction with others.

The third user is the one with the highest score. It seems that this user has achieved a very large number of followers without having to follow a large number of users, or having to favorite, but it seems that they post on a not so regular basis.

According to our algorithm, Nick G is given a power score of 100.34, Mathildi M a power score of 67.23 and Black H a power score of 4.33. The means that for every trending topic that includes a post from one of the aforementioned users, the trending topic power score will be increased by the score of the user. We should note that this score is counted

only once for the specific moment in time that the system will fetch the specific post, related to the trending topic. For each post that is re-fetched as a search result at a different moment in time in the future, the post power will be recalculated. This means that we can possibly measure the instant and accumulated power score of a trending topic in time.

Table 3 presents lists of trending topics fetched at a specific time and how they are ordered after they were given a power score.

**Table 3.** Trending topics list received on October 2020.

List 1	List 1 Ordered
amka	Prothipourgos (91.81)
KETHEA	KETHEA (78.43)
Rodo	#Tourism (68.88)
Prothipourgos	Rodo (35.43)
#XFactorGR	amka (34.28)
#EuroLeague	#EuroLeague (22.90)
#Tourism	#XFactorGR (20.02)

The words are fetched from Greek Twitter and they are originally in Greek language. It is important to note that tourism as a keyword is given a higher score, while xFactor is given a lower score. This simplistic example provides us with information on how a museum can benefit from the trending topics sorting procedure. When it seems that tourism becomes a trending topic, it is given a high score so that a museum can find the right time and space to start interactions in the medium. After all, culture is interconnected with tourism.

The system evaluates the input and decides on the trending topic score. It is able to perform a cold start as it does not have to calculate any value based on its own historical data, but only on historical data provided by the medium. According to qualitative analysis on the trending topics from several different system cycles, it seems that the system is able to calculate a score for the trending topics, so that the topics that can be of interest to a culture-related space seem to be enhanced and score higher. This means that the system is able to create an ordered list of the trending topics in such a way to promote these keywords that are usually present in more formal environments and serious conversations.

### 5.2. Information Visualization

For the visualization of information, the Grafana tool is utilized. Information visualization is an easy way to present large amounts of information to people, so that they can possibly obtain desired information without having to “dig” through big data. The environment is connected to a time-series database that holds all the information produced by the procedures that are already mentioned. The data that a time-series database needs are the values of the parameters to be displayed, followed by a timestamp. In our study, we store information about the trending topics and their power, as calculated by the internal procedures, together with the total number of posts (volume), favorites (likes) and retweets. The analytics mechanism performs a grouping of the trending topics to be easily depicted in graphs. Figure 2 presents a sample dashboard of a timeline presenting the evolution of trending topics in time.

By selecting a specific term from the legend, it is possible to obtain visualized information about the specific term. Figure 3 presents the presentation of the evolution in time of a specific term.

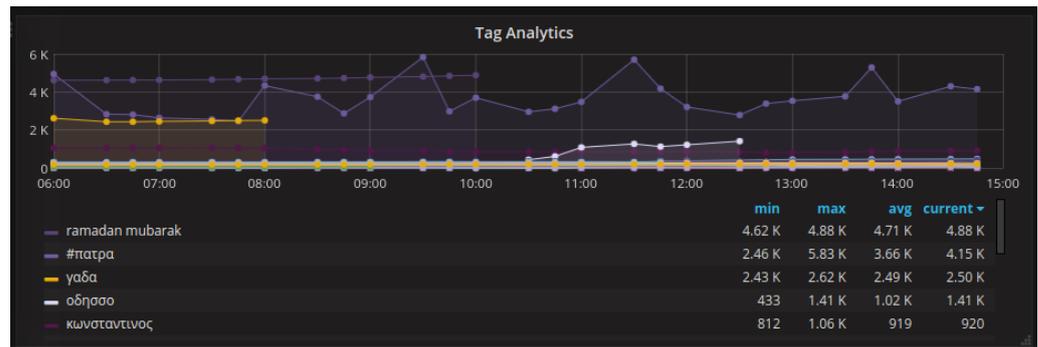


Figure 2. Sample Visualization of trending topics.

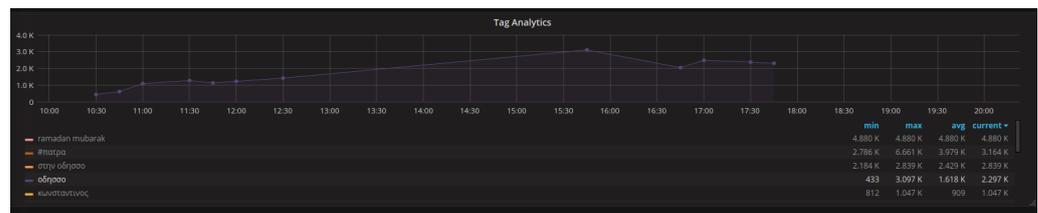


Figure 3. Single term evolution in time.

Grafana gives the possibility to build dashboards, where the time frame and interval of the analysis can be modified, in order to have different views of the information. Apart from depicting information about all the trending topics in real time, it is possible to create dashboards for each of the different trending topics that appear on the screen. The dashboard includes other parameters that are recorded in the time-series database to have a spherical approach of the collected data. Figure 4 presents a comparison of the trend power metric with the retweets and favorites over time. The representation is proof of the usage of the proposed implemented mechanism. More specifically, despite the fact that the trend keeps having higher and higher numbers of retweets and favorites, the score reaches a peak, and then it drops significantly. Furthermore, while the trend is spread across the medium, the people that write original posts related to it do not have high user current power, meaning that the trend is not spread anymore by people that tend to be influencers, losing in this way its power in the medium.

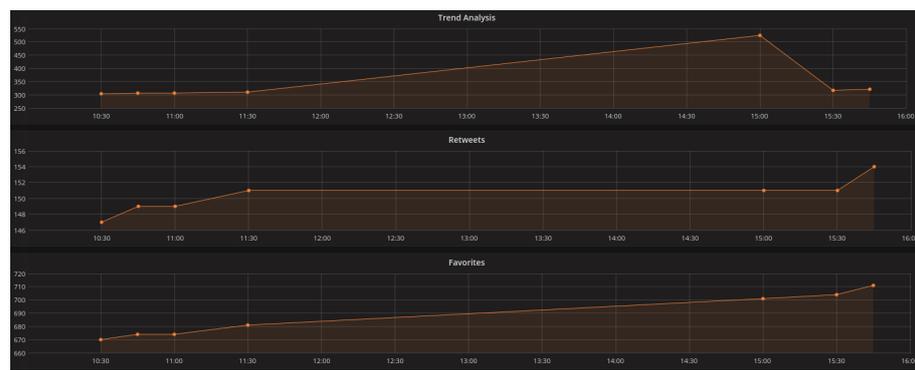


Figure 4. Trend power compared to retweets and favorites over time.

According to the collected information, the data that occur within a time of a day, concerning Greek Twitter, include 50–90 trending topics and 9000–15,000 collected posts. Inevitably, this information is impossible to be parsed manually, and the system is able to provide a simple means of data visualization in order to locate trending topics that may be of interest to a cultural space. In parallel, a testing procedure that was conducted including data from UK Twitter proves that the aforementioned numbers are increased to reach 180–200 trending topics daily with more than 50,000 tweets collected. The numbers

prove that the proposed approach can provide a solution to the problem of the vast amount of data and their daily analysis.

In parallel, a second module is able to provide specific information about the behaviour of the trending topics. This is the RESTful API module that can provide information that can be of extreme usefulness. These three are the basic parts of the API:

- Present during the last three hours;
- Persisting trends (for long time);
- Fastest growing trends.

The information from the API is provided in JSON (JSON—JavaScript Object Notation, <https://www.json.org/> (accessed on 30 May 2021)) format as presented in Figure 5.

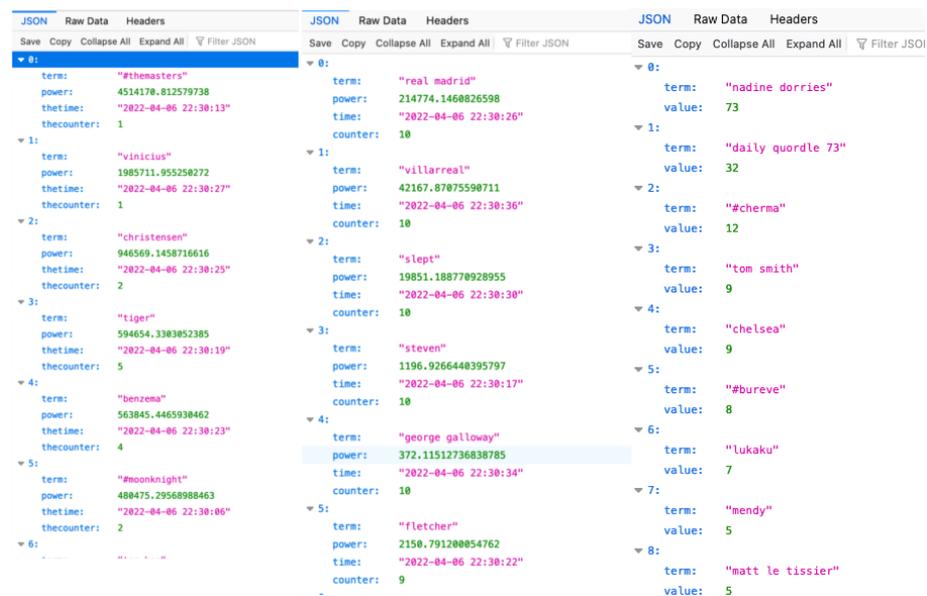


Figure 5. Sample results from JSON data execution.

The information provided by using this type of data export can be easily become an input to any system that supports creating a dashboard with JSON input. Despite the fact that information is not quite clear for a person that is not accustomed to raw data, a number of dashboards are based on JSON data. Figure 6 presents the visualization of data through a simple service (jsontoChart (JSONtoChart)—<https://jsontochart.com/> (accessed on 30 May 2021)).

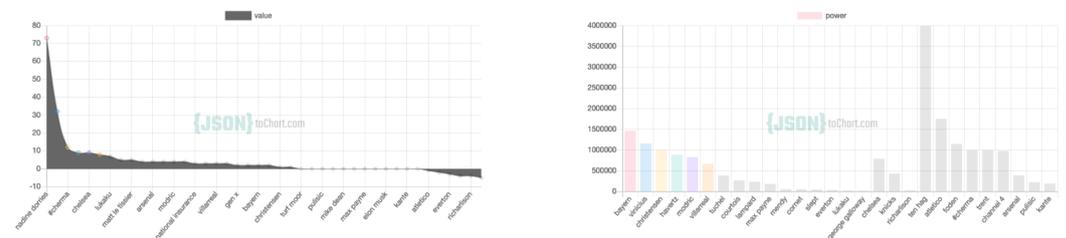


Figure 6. JSON Data to Charts.

### 5.3. Benefits of Usage

Despite the fact that a number of steps and cutting-edge technologies are used to support cultural spaces under the umbrella of “cultural informatics”, there still seems to be a number of steps that need to be taken considering the web presence. The part that seems to be the weakest considering the cultural spaces has to do with the online participation. Participation, in contrast to a generic online strategy, cannot be pre-defined as it is a dynamic procedure that changes from medium to medium and from user to user.

By recognizing the online trends, the cultural spaces are able to locate a place to start their online participation and interaction.

Our approach is to establish trends that will lead to serious conversations, and as such, this is the reason we research user and trend scores. In this way, we try to empower the part of a cultural space strategy that has to do with participation in social media as we narrow down the entry paths for serious conversations. Of course, trending topics that concern a museum or a cultural space may not occur on an hourly or even daily basis. However, our research proved that analysis on trending topics from a medium and a classification to promote trends that lead to serious conversations can be a “place to start” in the chaotic world of the Internet.

## 6. Conclusions and Future Work

This paper presented the research procedures that were performed as part of a research task of the Greek National Project entitled PaloAnalytics. The research work focuses on the trending topics of Twitter and intends to enhance them with qualitative data. More precisely the scope of the research is to measure the impact of each user and project this information onto the trending topics provided by the medium itself. We proposed an algorithm that takes into account users’ information to create an impact score, while in parallel we tried to use this score in order to create an ordering of the trending topics. While the original scope of the research project was limited to trending topic classification, we proceeded to a further step. We tried to interconnect the solution with a problem that exists in museums and cultural spaces. The problem is called participation and is related to the absence of an online presence of the museums when it comes to online conversations. By trying to alter parts of the implemented algorithm in order to classify higher trending topics that derive from specific types of users (with high influence), we claimed that we can possibly help museums and cultural spaces locate a place to initiate a serious conversation.

Technology is nowadays a part of museums and cultural spaces. Modern advances in technology have attracted a lot of spaces and people are expecting museums to have web presence, use high-end technology, digitally communicate their message or their educational activities, and generally have a larger radius of reach. On the other hand, some technological “necessities”, such as social media analysis for usage in culture-related spaces, are a concern that affects both people in a museum and researchers related to cultural informatics. This is because it is a matter of high importance, as more and more people tend toward a mixed type of life, while in parallel it seems that the culture of the web tends to be very flat. From the literature review, it seems that several problems have been solved, but participation remains a major issue. Participation cannot be foreseen or predicted. It has to be a dynamic procedure that should adapt both to people and the reality of digital life, which is constantly changing.

In this work, we presented a mechanism that is able to analyze a medium (Twitter) in order to enrich it with information that could be useful for organizations such as museums. This is because the analysis performed can lead to revealing spaces on the Internet where cultural spaces can interact and participate. The results from the system execution show that the qualitative analysis of the medium related to trending topics, in real time, without having to extract very large amounts of data from it, can lead to significantly high-quality results for organizations seeking conversations to intervene and fulfill their “participation” obligation.

The next research step is the actual application of the system within an organization (preferably cultural) in order to apply the results of the system execution in a real environment. However, still, while it is quite straightforward to measure the impact of the information dissemination (reach and reactions), on the contrary, it still remains an issue to measure the impact of the participation procedure.

**Author Contributions:** Conceptualization, V.P. and M.W.; Methodology, V.P.; Software, V.P.; Validation, V.P. and M.W.; Formal Analysis, V.P. and M.W.; Investigation, M.W.; Resources, V.P. and M.W.; Data Curation, V.P.; Writing—Original Draft Preparation, V.P. and M.W.; Writing—Review and

Editing, V.P.; Visualization, V.P.; Supervision, M.W. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research has been co-financed by the European Union and Greek national funds through the Operational Program Competitiveness, Entrepreneurship and Innovation, under the call RESEARCH—CREATE—INNOVATE (project code: T1EDK-03470).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Bearman, D.A.; Lytle, R.H. *The Power of the Principle of Provenance*; Archives and Museum Informatics: Toronto, ON, Canada, 1985; Volume 21, pp. 14–27.
- Bearman, D. *Collecting Software: A New Challenge for Archives & Museums*; Archives and Museum Informatics: Toronto, ON, Canada, 1987. Available online: [https://www.archimuse.com/publishing/col\\_soft/col\\_soft.Ch7.pdf](https://www.archimuse.com/publishing/col_soft/col_soft.Ch7.pdf) (accessed on 6 April 2022).
- Bearman, D. *Functional Requirements for Collections Management Systems*; Archives and Museum Informatics: Toronto, ON, Canada, 1987.
- Bierbaum, E.G. Records and access: Museum registration and library cataloging. *Cat. Classif. Q.* **1988**, *9*, 97–111. [\[CrossRef\]](#)
- Bitner, R. *Nomenclature for Museum Cataloguing; A System for Classifying Man-Made Objects*; American Association for State and Local History: Nashville, TN, USA, 1980.
- Appel, S. Copyright, digitization of images, and art museums: Cyberspace and other new frontiers. *UCLA Ent. L. Rev.* **1998**, *6*, 149. [\[CrossRef\]](#)
- Bertacchini, E.; Morando, F. The future of museums in the digital age: New models for access to and use of digital collections. *Int. J. Arts Manag.* **2013**, *15*, 60–72.
- Hirtle, P.B.; Hudson, E.; Kenyon, A.T. *Copyright and Cultural Institutions: Guidelines for Digitization for US Libraries, Archives, and Museums*; Forthcoming, U of Melbourne Legal Studies Research Paper No. 434; Cornell University Library Press: Ithaca, NY, USA, 2009.
- Hylland, O.M. Even better than the real thing? Digital copies and digital museums in a digital cultural policy. *Cult. Unbound* **2017**, *9*, 62–84. [\[CrossRef\]](#)
- Terras, M.M. The Rise of Digitization. In *Digitisation Perspectives. Educational Futures Rethinking Theory and Practice*; Rikowski, R., Ed.; SensePublishers: Rotterdam, The Netherlands, 2011; Volume 46.
- Ciolfi, L.; Bannon, L.J.; Fernström, M. Including visitor contributions in cultural heritage installations: Designing for participation. *Mus. Manag. Curatorship* **2008**, *23*, 353–365. [\[CrossRef\]](#)
- Rey, F.B.; Casado-Neira, D. Participation and technology: Perception and public expectations about the use of ICTs in museums. *Procedia Technol.* **2013**, *9*, 697–704. [\[CrossRef\]](#)
- Winter, M. Visitor perspectives on commenting in museums. *Mus. Manag. Curatorship* **2018**, *33*, 484–505. [\[CrossRef\]](#)
- Petrelli, D.; O'Brien, S. Phone vs. Tangible in Museums: A Comparative Study. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, Montreal, QC, Canada, 21–26 April 2018.
- Sayre, S. Bring It On: Insuring the Success of BYOD Programming in the Museum Environment. In Proceedings of the Museums and the Web 2015, Chicago, IL, USA, 8–11 April 2015; Proctor, N., Cherry, R., Eds.; Museums and the Web LLC: Silver Spring, MD, USA, 2015.
- Antoniou, A.; Dejonai, M.I.; Lepouras, G. 'Museum escape': A game to increase museum visibility. In Proceedings of the International Conference on Games and Learning Alliance, Laval, France, 9–10 December 2019; Springer: Cham, Switzerland, 2019; pp. 342–350.
- Clini, P.; Quattrini, R.; Bonvini, P.; Nespeca, R.; Angeloni, R.; Mammoli, R.; Dragoni, A.F.; Morbidoni, C.; Sernani, P.; Mengoni, M.; et al. Digit (al) isation in Museums: Civitas Project—AR, VR, Multisensorial and Multiuser Experiences at the Urbino's Ducal Palace. In *Virtual and Augmented Reality in Education, Art, and Museums*; IGI Global: Hershey, PA, USA, 2020; pp. 194–228.
- Lee, H.; Jung, T.H.; tom Dieck, M.C.; Chung, N. Experiencing immersive virtual reality in museums. *Inf. Manag.* **2020**, *57*, 103229. [\[CrossRef\]](#)
- Shehade, M.; Stylianou-Lambert, T. Virtual reality in museums: Exploring the experiences of museum professionals. *Appl. Sci.* **2020**, *10*, 4031. [\[CrossRef\]](#)
- Vassilakis, C.; Kotis, K.; Spiliotopoulos, D.; Margaris, D.; Kasapakis, V.; Anagnostopoulos, C.N.; Santipantakis, G.; Vouros, G.A.; Kotsilieris, T.; Petukhova, V.; et al. A semantic mixed reality framework for shared cultural experiences ecosystems. *Big Data Cogn. Comput.* **2020**, *4*, 6. [\[CrossRef\]](#)
- Vital, R.; Sylaiou, S. Digital survey: How it can change the way we perceive and understand heritage sites. *Digit. Appl. Archaeol. Cult. Herit.* **2022**, *24*, e00212. [\[CrossRef\]](#)

22. Marty, P.F.; Jones, K.B. (Eds.) *Museum Informatics: People, Information, and Technology in Museums*; Taylor & Francis: Tokyo, Japan, 2008; Volume 2.
23. Marty, P.F. Museum informatics and collaborative technologies: The emerging socio-technological dimension of information science in museum environments. *J. Am. Soc. Inf. Sci.* **1999**, *50*, 1083–1091. [[CrossRef](#)]
24. Marty, P.F. Unintended consequences: Unlimited access, invisible work and the future of the information profession in cultural heritage organizations. *Bull. Am. Soc. Inf. Sci. Technol.* **2012**, *38*, 27–31. [[CrossRef](#)]
25. Kang, X.; Chen, W.; Kang, J. Art in the age of social media: Interaction behavior analysis of Instagram art accounts. *Informatics* **2019**, *6*, 52. [[CrossRef](#)]
26. Nuccio, M.; Bertacchini, E. Data-driven arts and cultural organizations: Opportunity or chimera? *Eur. Plan. Stud.* **2021**, 1–18. [[CrossRef](#)]
27. Amanatidis, D.; Mylona, I.; Mamalis, S.; Kamenidou, I.E. Social media for cultural communication: A critical investigation of museums' Instagram practices. *J. Tour. Herit. Serv. Mark. (JTHSM)* **2020**, *6*, 38–44.
28. Bosello, G.; Haak, M.V. #Arttothepeople? An exploration of Instagram's unfulfilled potential for democratising museums. *Mus. Manag. Curatorship* **2022**, 1–18. [[CrossRef](#)]
29. Gerrard, D.; Sykora, M.; Jackson, T. Social media analytics in museums: Extracting expressions of inspiration. *Mus. Manag. Curatorship* **2017**, *32*, 232–250. [[CrossRef](#)]
30. Vassiliadis, C.; Belenioti, Z.C. Museums & cultural heritage via social media: An integrated literature review. *Tourismos* **2017**, *12*, 97–132.
31. Lazzeretti, L.; Sartori, A.; Innocenti, N. Museums and social media: The case of the Museum of Natural History of Florence. *Int. Rev. Public Nonprofit Mark.* **2015**, *12*, 267–283. [[CrossRef](#)]
32. Fletcher, A.; Lee, M.J. Current social media uses and evaluations in American museums. *Mus. Manag. Curatorship* **2012**, *27*, 505–521. [[CrossRef](#)]
33. Fernandes, A.B. "But will there be visitors?" Public outreach efforts using social media and online presence at the Côa Valley Museum and Archaeological Park (Portugal). *Internet Archaeol.* **2018**, *47*. [[CrossRef](#)]
34. Villaespesa, E.; Wowkowych, S. Ephemeral storytelling with social media: Snapchat and Instagram stories at the Brooklyn Museum. *Soc. Media+ Soc.* **2020**, *6*, 2056305119898776. [[CrossRef](#)]
35. Zollo, L.; Rialti, R.; Marrucci, A.; Ciappei, C. How do museums foster loyalty in tech-savvy visitors? The role of social media and digital experience. *Curr. Issues Tour.* **2021**, 1–18. [[CrossRef](#)]
36. Contri, M. Museums and their audience: Towards dialogic communication through social media? *Int. J. Digit. Cult. Electron. Tour.* **2020**, *3*, 22–35. [[CrossRef](#)]
37. Utami, E.; Hartanto, A.D.; Raharjo, S. Systematic Literature Review of Profiling Analysis Personality from Social Media. *J. Phys. Conf. Ser.* **2021**, *1823*, 012115.
38. Apriyanto, S.; Nurhayaty, A. Born In Social Media Culture: Personality Features Impact In Communication Context. In Proceedings of the 2nd ICoLLiT (International Conference on Language, Literature and Teaching), Padang, Indonesia, 22–23 August 2019.
39. Campbell, W.; Baseman, E.; Greenfield, K. Content+ context networks for user classification in Twitter. In Proceedings of the Neural Information Processing Systems (NIPS), Stateline, NV, USA, 5–10 December 2013.
40. De Choudhury, M.; Diakopoulos, N.; Naaman, M. Unfolding the event landscape on Twitter: Classification and exploration of user categories. In Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work, Seattle, WA, USA, 11–15 February 2012; pp. 241–244.
41. Pennacchiotti, M.; Popescu, A.M. A machine learning approach to Twitter user classification. In Proceedings of the International AAAI Conference on Web and Social Media, Barcelona, Spain, 17–21 July 2011; Volume 5.
42. Uddin, M.M.; Imran, M.; Sajjad, H. Understanding types of users on Twitter. *arXiv* **2014**, arXiv:1406.1335.
43. Pouloupoulos, V.; Vassilakis, C.; Antoniou, A.; Lepouras, G.; Theodoropoulos, A.; Wallace, M. The Personality of the Influencers, the Characteristics of Qualitative Discussions and Their Analysis for Recommendations to Cultural Institutions. *Heritage* **2018**, *1*, 239–253. [[CrossRef](#)]
44. Pouloupoulos, V.; Vassilakis, C.; Antoniou, A.; Lepouras, G.; Wallace, M. Personality Analysis of Social Media Influencers as a Tool for Cultural Institutions. In Proceedings of the Euro-Mediterranean Conference, Nicosia, Cyprus, 29 October–3 November 2018; Springer: Cham, Switzerland, 2018; pp. 236–247.
45. Scullard, M.; Baum, D. *Everything DiSC Manual*; Wiley: Hoboken, NJ, USA, 2015.