

POLITECNICO DI MILANO

**School of Industrial and Information Engineering
Master of science in Management Engineering**

**INTERACTIVE DASHBOARDS AND
DATA INTEGRATION: DESIGN AND
TOOL DEVELOPMENT FOR MUSEUMS**



**POLITECNICO
MILANO 1863**

Supervisor: Michela Arnaboldi

Co-supervisor: Deborah Agostino

Authors

Stefania Castelnovo matr. 921005

Riccardo Pinna matr. 915069

Academic Year 2020/2021

“...there are known knowns; there are things we know we know. We also know there are known unknowns; that is to say we know there are some things we do not know. But there are also unknown unknowns—the ones we don't know we don't know. And if one looks throughout the history of our country and other free countries, it is the latter category that tends to be the difficult ones”

Donald Rumsfeld

TABLE OF CONTENTS

1	<i>Abstract</i>	8
1.1	Italian Version	8
1.2	English Version	11
2	<i>Introduction</i>	14
2.1	Research question	14
2.2	Objective	16
3	<i>Literature Review</i>	21
3.1	Framework Analysis: Balance Score Card and KPIs	21
3.2	User-Generated-Content platforms and Social Media	27
3.2.1	UCG Platforms Evolution	32
3.2.2	UCG Platforms Classification	36
3.2.3	UCG Platforms Business Applications	41
3.2.4	Measuring the degree of corporate social media use	43
3.2.5	Social Media Analytics	55
3.2.6	Sentiment analysis	57
3.3	Interactive Dashboard	60
3.3.1	Dashboard development	61
3.3.2	Dashboard's qualities, shortcomings and impact over the organization	63
3.3.3	Data Analytics and Big Data	68
3.3.4	Database and the Entity-Relationship Model	71
3.4	The research gap	74
4	<i>Conceptual Framework</i>	76

4.1	An overview of Museums organization in Italy	76
4.2	The CCO Matrix - a new strategic model for performance monitoring	78
4.2.1	Communication	86
4.2.2	Collection	98
4.2.3	Organization	108
4.2.4	CCO Matrix overview	118
5	Methodology	120
5.1	Dashboards Observation	121
5.1.1	Dashboard Best Practices identification	121
5.2	From a model to a tool: CCO Dashboard construction	131
5.3	Literature review	141
5.4	Tool development and application	144
5.4.1	Proof of concept I	144
5.4.2	Proof of concept II	147
6	Results and discussion	149
6.1	Proof of concept I	149
6.2	Proof of concept II	154
6.2.1	March 2021 Prototype	154
6.2.2	The challenge of Data gathering	160
6.2.3	Response from the stakeholders	163
7	Conclusions	167
7.1	Limitations and further research	171
8	Bibliography	174

9	<i>Annex</i>	181
9.1	Guidelines on how to proper use Social Media in business	181
9.1.1	Four Points about using social media	181
9.2	Deep Dive into main user-generated content platforms used by cultural institutions: Facebook, Twitter, Instagram, Google Maps and Trip Advisor	183
9.3	Social media Reputation Dashboard	195
9.4	Microsoft Power Bi	201

TABLE OF FIGURES

Figure 1 - Harvard Business Review: The high-performance organization, July-August 2005	23
Figure 2 - The cascading process example (Arnaboldi, Azzone and Giorgino, 2015).....	25
Figure 3 - An example of a strategy map (Arnaboldi, Azzone and Giorgino, 2015)	26
Figure 4 - Illustrative representation of marketing evolution	42
Figure 5 - Influencer Marketing, Social Media Marketing, TV Advertisement, Radio Advertisement, Printed Advertisement trend in terms of interest overtime compared.....	43
Figure 6 - Extracting insights from big data (Labrinidis and Jagadish, 2012)	70
Figure 7 - Example of entity set	72
Figure 8 - Structure of relationship between entities.....	72
Figure 9 - The research pannel	81
Figure 10 - CCO Matrix Strategic Map overview	85
Figure 11 - CCO Matrix: Communication Section indicators.....	86
Figure 12 - CCO Matrix: Collection section indicators.....	98
Figure 13 - CCO Matrix: Organization section indicators	109
Figure 14 - CCO Matrix Overview.....	119
Figure 15 - "Dark mode" search query popularity on Google over time. Source: Google Trends	125
Figure 16 - Channel analytics panel from the Youtube Studio Dashboard	129
Figure 17 - Google's Stock Performance Dashboard	129
Figure 18 - Visualization of the difference between proper and improper hierarchy	138

Figure 19 - MIBACT Dashboard view template.....	150
Figure 20 - Visualizing data through different level in hierarchy.....	152
Figure 21 - Screenshot of the Comunicazione - Reputazione view	153
Figure 22 - screenshot of the Main view for MANN.....	158
Figure 23 - screenshot of the Engagement view for MANN	159
Figure 24 - screenshot of Engagement Benchmark view for MANN	160
Figure 25 - Mock up of percentage change implementation.....	198

1 **Abstract**

1.1 *Italian Version*

La centralità e l'importanza anche a livello internazionale dei beni culturali italiani è indiscussa, tuttavia sotto il profilo manageriale solo dopo alcune recenti riforme si è posta maggiore attenzione, in particolare ai Musei. Tema particolarmente rilevante, enfatizzato anche dalla riforma cosiddetta Franceschini (REF) è la capacità di misurare le prestazioni, attraverso indicatori quantificabili, coprendo sia le attività più tradizionali, come ad esempio il numero di visitatori, ma anche le performance nel mondo online.

In Italia, la nomina di *museo* può indicare svariati tipi di diverse istituzioni culturali la cui principale missione è quella di valorizzare e proteggere il vasto patrimonio culturale della nazione. Pertanto, si tratta di enti il cui successo non è misurabile in termini puramente finanziari. Essi operano con diversi gradi di autonomia, relazionandosi con il territorio, offrendo servizi al cittadino locale e al visitatore internazionale, in un contesto globale interessato da profondi cambiamenti e in transizione verso la digitalizzazione.

I cruscotti digitali di indicatori sono uno strumento di business intelligence di rilevanza crescente. I progressi e la diffusione tecnologica rendono questi strumenti sempre più potenti, flessibili ed accessibili. Queste caratteristiche li rendono un candidato ideale per l'implementazione di una soluzione multidimensionale di monitoraggio delle prestazioni e supporto decisionale.

In questo contesto, la tesi si propone tre obiettivi:

- Il primo è quello di delineare un modello per la misurazione delle prestazioni di istituzioni culturali in grado di catturare la dinamicità del segmento, e di applicarlo mettendo a punto, costruendo e lanciando uno strumento di business intelligence potente e flessibile che offra una panoramica sullo stato generale dei singoli così come dell'insieme dei musei in tempo reale, utile sia alle istituzioni sia al regolatore che le supervisiona;
- Il secondo è quello di integrare ai dati tipicamente impiegati nella misurazione delle performance anche fonti non convenzionali, principalmente social network e altre piattaforme basate su contenuti generati da utenti, per costruire indicatori che superino i modelli più tradizionali e dipingano una visione più ampia, rilevante e attuale, estraendo e presentando informazioni declinate verso obiettivi di qualità e servizio richiesti dall'utilizzatore finale;
- Il terzo obiettivo è quello di inquadrare e comprendere le difficoltà che derivano dall'implementazione di questa soluzione e quali siano i potenziali impatti positivi e negativi all'interno dei processi e strutture dell'organizzazione esistente.

L'elaborato di tesi presenta un'analisi della letteratura accademica che descrive i modelli più tradizionali e i concetti di base applicabili nell'arte della misurazione delle performances, seguita da un approfondimento sull'evoluzione e sullo stato attuale delle piattaforme social di interesse per il nostro studio. Infine, ampio spazio è dedicato ai cruscotti di indicatori, cosa sono, come si sono sviluppati e quali sono le sfide e conseguenze che emergono nel design e implementazione di tali strumenti. Per il raggiungimento degli obiettivi, il fine dell'analisi è quello di determinare un modello solido e le linee guida da seguire per formulare ed impostare un cruscotto di indicatori

interattivo utile ad utenti decisori in campo di istituzioni culturali. Le indicazioni di lavoro sono poi validate dalla loro messa in pratica con la formulazione e presentazione di un cruscotto per il monitoraggio e l'assistenza alla verifica e raggiungimento degli obiettivi di qualità del *Servizio museale nazionale*, in collaborazione con la *Direzione generale Musei* del Ministero per i beni e le attività culturali e per il turismo (MiBACT) e il *Museo Archeologico Nazionale di Napoli*.

1.2 English Version

The focus and the importance of the Italian cultural heritage, also on an international level, is not in question, however from a managerial perspective, especially on museums, more attention to the topic has been drawn by some recent reforms. A topic of particular relevance, emphasized by the so called Franceschini reform (REF) is the ability to measure performances, through quantifiable indicators, covering both traditional activities, such as visitors' number, as well as performances in the online world.

In Italy, the word *museum* may indicate numerous kinds of different cultural institutions, as long as they share the common role of keepers of the vast cultural heritage the Nation possesses. Hence, they are entities of which performances cannot be measured in purely financial terms. Museums in Italy operate with different degrees of autonomy and intertwining with the territory, offering services to local citizens and international visitors, while the global context around them goes through deep transformations, transitioning into digitalization.

Digital dashboards of indicators are business intelligence tools of rising relevance. Technological progresses and diffusion result in these tools getting more and more powerful, flexible and accessible. These characteristics make them an ideal fit as a multidimensional solution for performance monitoring and decision-making support.

This thesis proposes three research objectives:

- the first one is to outline a cultural institutions' performance measurement model capable of capturing the dynamicity of the segment, and to apply it by designing,

building and launching a powerful and flexible business intelligence tool, that offers a snapshot of the general status of both a single as well as of a group of museums in real time, useful for both the institutions and the regulator supervising them;

- the second one is integrating non-conventional sources of data, mostly social network and other user generated content platforms, to data that are typically put in use when measuring performances, to build key performance indicators that go beyond traditional models and paint a wider, relevant and actual picture, extracting and presenting information leaning towards quality and service objectives required by the final user;
- the third objective is to frame and understand the challenges that come from the implementation of such a solution and what is the potential positive and negative impact on the organization's existing processes and structure.

This thesis presents an academic literature review that describes traditional models and basic concepts applicable in the art of measuring performances, followed by an exploration of the evolution and current state of the social platforms that are relevant for this study. Finally, ample space is dedicated to dashboards, what they are, how were they developed and what are the challenges and consequences that emerge from the design and implementation of such tools. To achieve the objectives, the target of the analysis is to determine a solid model and guidelines to formulate and set up an interactive dashboard of indicators useful to decision-makers in the cultural institutions field. The work instructions are then validated by their application to formulate and present a dashboard for monitoring and assisting the achievement of the *Sistema museale nazionale's* objectives, in collaboration

with the Ministry of Cultural Heritage and Activities' *Museums general Direction* and the *National Archeological Museum of Naples*.

2 Introduction

2.1 *Research question*

The International Council of Museums (2007) defines a museum as *a non-profit-making, permanent institution in the service of society and of its development, and open to the public, which acquires, conserves, researches, communicates and exhibits, for purposes of study, education and enjoyment, material evidence of people and their environment.*

The production of museum services is facing multiple challenges, some external and other internal to the organization. The main external challenges for the museum's management arise from the economic crisis (Basso, Casarin and Funari, 2018). Specifically, during economic crisis (as the latest crises due to the COVID-19 pandemic) we observe both a reduction in the public funding for arts and a change in the consumption of culture, due to a decrease in the personal income of many agents.

Moreover, museums sponsors have changed their attitude towards funding and ask for a more precise and proactive assessment of the actual returns that can be achieved; even donors request a more detailed planning and information on supported activities (Basso, Casarin and Funari, 2018). As a result, museums are increasingly coping with competitive situations, trying to attract both users and donors, competing with many organizations operating in the field of the free time market (Basso, Casarin and Funari, 2018).

This thesis aims at determining an effective tool to monitor the performances of museums and cultural institutions, and subsequently developing a tool able to integrate the different points of view of the internal and the external stakeholders involved, a tool that can capture

the dynamics of the segment and provide real-time outputs and powerful insights: this goal can be reached through the construction of dynamic dashboard.

On this note the thesis answers to one main questions:

How do you build a museums performances measurement system, integrating unconventional data with conventional ones, presented through an interactive and flexible dashboard, that captures the point of view of different stakeholders involved and the dynamism of the sector, and what are the challenges and effects from building and implementing such a solution?

In order to properly answer this question, the thesis starts by presenting an analysis of the available academic literature regarding existing traditional frameworks. The analysis then continues presenting the potential of unconventional data extracted from user generated content platforms, as well as of interactive dashboards. For unconventional data, the research starts by defining what User Generated Platforms are and how data coming from these sources can be used in a value-added way, through for example data elaboration and sentiment analysis. A considerable effort has been spent in describing how datasets extracted from online user-generated content platforms such as social networks can be effectively implemented and integrated with more conventional data to produce and deliver valuable information. The results of this analysis are presented in chapter 3.2 “User-Generated Content platforms and Social Media” and in the Annex. As for interactive dashboards, the research followed a two-step process that consists on observing first and creating last and it revolves around the key word “dashboard” and the key tool “Power Bi”. Chapter 3.3 “Interactive Dashboard” of this thesis contains an important digression regarding what dashboards are and how can they be implemented to offer value to both

single decision makers and organizational processes as a whole. Chapter 5.1 “Dashboard observation” outlines some common best practices that can be observed in a pool of publicly available dashboards. Then, the thesis proposes a new framework in chapter 4. “Conceptual Framework” that captures the dynamism of the industry, the point of view of internal and external stakeholders and lays the foundations to a business tool construction. This framework is then adapted in two working interactive dashboard prototypes designed and developed following the best practices identified before hand and tailored for both a single institution (the National Archeological Museum of Naples) as well as for the biggest Italian regulator (the Museum General Service Direction office of the Italian Ministry of Cultural Heritage and Activities), both of which offered meaningful insights in answering this question in a satisfying manner. The results are presented in chapter 6 “Results and discussion”.

The desire for investigating this matter stems from researcher curiosity: tools such as interactive dashboards have been implemented successfully in many business areas, applying the same concepts to museums in which performance evaluation is a challenging and not standardized task which can yield both practical benefits for the parties involved as well as an interesting and worthwhile contribution to the academic literature.

2.2 Objective

The objective of this thesis is to answer to the question presented on the previous chapter in three different main stages:

- outlining a model to measure museum performances;
- extracting meaningful and valuable information from user generated content platforms and integrating these information in a measurement system;
- understanding the potential, the shortcomings, the challenges and the effects of an interactive dashboard application and developing it.

This expected result is to support museums to face new challenges by providing this sector with a new innovative model and tool to proper monitor its performances and, in consequence, increase its efficiency and stakeholder's satisfaction. The research focuses on designing a new reporting tool through Microsoft's Power Bi services, to explore the possibilities the software offers in building a dashboard that is flexible, powerful and a user-friendly reporting and business intelligence tool. The research put effort into innovating not only the tool design and functionalities, but also the data sources used for data analysis and monitoring, by including and leveraging unconventional data harvested from user-generated-content platforms online.

Regarding the subject of this research, we chose to focus on museum and cultural institutions because the literature alone cannot provide to them an up-to-date performance measurement system in a satisfying extent. Moreover, what's available heavily leans towards internal data, completely missing out on a more holistic view of museum performances deriving from the analysis of different points of view in terms of internal and external stakeholders. Consulting available studies about the analysis of data coming from social media and user-generated-content platform reveals the lack of horizontal and

comprehensive indicators that track an institution reputation across different social media, focusing instead vertically on a single platform at the time.

Helping museums and cultural institutions to innovate and move forward to more advanced and comprehensive tools for business is very beneficial as museums are one of the most illustrative examples of the concept of cultural heritage, and providing culture is a matter of public good allocation on the same line of other services such as health, education, as well as various administration services (del Barrio and Herrero, 2013), a service that should always be available and deserves proper attention.

The challenge of the analysis lies in the fact that assessing performance for cultural institutions, and particularly for museums, is a more subjective and value-driven exercise than for companies, given the uniqueness of cultural goods, resulting from their symbolic and intangible significance. Museum defines a wide group of different institutions that pursue many different goals, there is no single quantitative or qualitative metric against which performance can be evaluated and ranked, and further complications come from the difficulty involved in collecting reliable and representative data from the institutions. Finally, many stakeholders are involved and need to be considered, and according to Speckbacher, “there is no clearly defined stakeholder group with homogeneous expectations and objectives that can be placed at the top of the hierarchy” in the museums and cultural institutions industry.

In order to undertake this remarkable task, the work is structured in 4 main parts.

1. **Literature Review:** “*Chapter 3. Literature review*” presents the results of the review and observation of the three main subjects of the desertion, being framework and

indicators definition, user-generated-content platforms data integration and interactive dashboards.

The literature review provides an understanding and overview of these elements and their value in pursuit of the objective of the research.

2. **Conceptual Framework:** "*Chapter 4. Conceptual Framework*" presents the construction of a new framework tailored of museums and describes the theoretical and empirical evidence that support its soundness.
3. **Methodology:** "*Chapter 5. Methodology*" it describes how all the process has been conducted and describes how the model can be translated into a practical tool and how it has been done.
4. **Results and Discussion:** "*Chapter 6. Results and Discussion*" it presents the final output of the research, its practical application, the main findings, alongside with the description of the process of validation of the tool by the Ministry of Cultural Heritage and the National Archeological Museum of Naples.

The work concludes with the conclusions which include the benefits and shortcomings of the model.

3 Literature Review

In this chapter we present the results of the scoping literature review. The result of the review of the literature is divided into three main chapters:

- Framework Analysis: Balance Score Card and KPIs
- User-Generated-Content platforms and Social Media
- Interactive Dashboard

corresponding to the three subjects of study that we focused on in order to answer to the three main questions of the thesis:

- outlining a model to measure museum performances;
- extracting meaningful and valuable information from user generated content platforms and integrating these information in a measurement system;
- understanding the potential, the shortcomings, the challenges and the effects of an interactive dashboard application and developing it.

with a last additional small chapter presenting the gap found in the literature analysis that our research aims to fill.

3.1 Framework Analysis: Balance Score Card and KPIs

The following framework analysis is composed by two main parts: the first related to the Balance Scorecard tool analysis and, since the quality of the BSC depends on the indicators selected, the second part is focuses on indicators.

The framework proposed in chapter 4 takes as input the above-mentioned topics and overcomes the problems linked to the application of a traditional BSC to non-for-profit organizations, in particular to museums.

A *balanced scorecard* is a strategic management performance metric used to measure businesses performances. It was first introduced in 1992 by David Norton and Robert Kaplan from the Harvard Business School, who took previous metric performance measures and adapted them to include nonfinancial information. A Balance Scorecard identifies how various internal business functions influence external results and are used to measure and provide feedback to organizations.

The balanced scorecard allows managers to look at the business from four important perspectives and it provides answers to four basic questions (Kaplan and Norton, 1992):

- How do customers see us? (*customer perspective*). The balanced scorecard demands that managers translate their general mission statement on customer service into specific measures that reflect the factors that really matter to customers.
- What must we excel at? (*internal business perspective*). The balance scorecard focuses on those critical internal operations that enable them to satisfy customer needs and meet their expectations.
- Can we continue to improve and create value? (*innovation and learning perspective*). The balanced scorecard identifies the parameters that the company considers most important for competitive success, adapting to their ever-changing nature.

- How do we look to shareholders? (*financial perspective*). The balance scorecard measures whether the company's strategy, implementation, and execution are contributing to bottom-line improvement.

The Balanced Scorecard Links Performance Measures

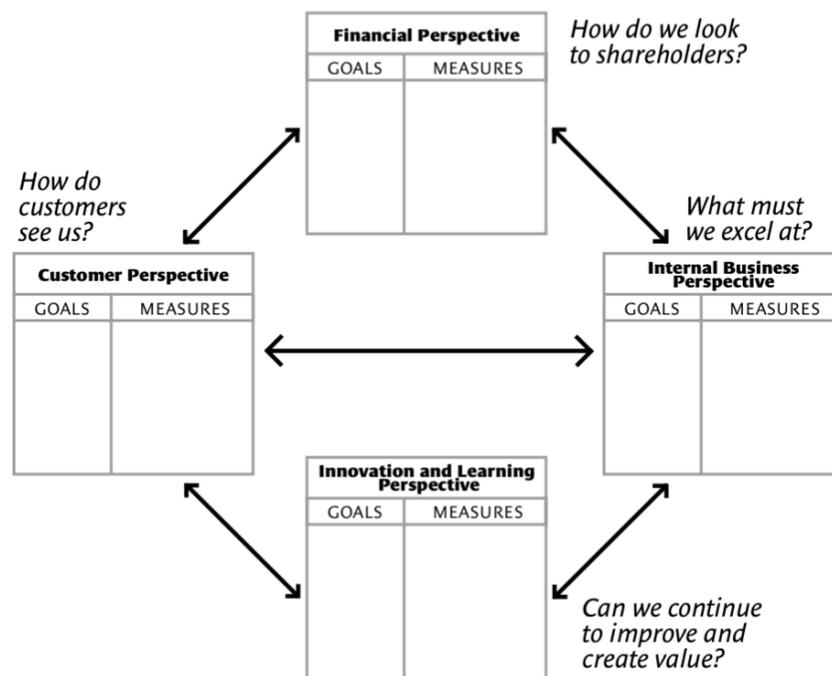


Figure 1 - Harvard Business Review: The high-performance organization, July-August 2005

The balanced scorecard can be used as a strategic management system and a tool to translate strategy into action since it provides information about the company as a whole. An

organization may use the balanced scorecard model to implement strategy mapping to see where value is added within an organization.

The quality of the BSC depends on the indicators selected and included in each of the four quadrants, which should be linked by causal relationships and updated according to changes in the company's strategy. To achieve this result, a tailored design is required for each company. In the second-generation BSC, Kaplan and Norton illustrate the process to follow in order to achieve that (Arnaboldi, Azzone and Giorgino, 2015).

The first step would be the construction of a strategic map of enterprise goals. It starts by identifying general goals, further divided into subobjectives through a cascade process. Usually, relations are from the top to the bottom or within the same level. The cascading process can, however, skip one or more areas. (Arnaboldi, Azzone and Giorgino, 2015)

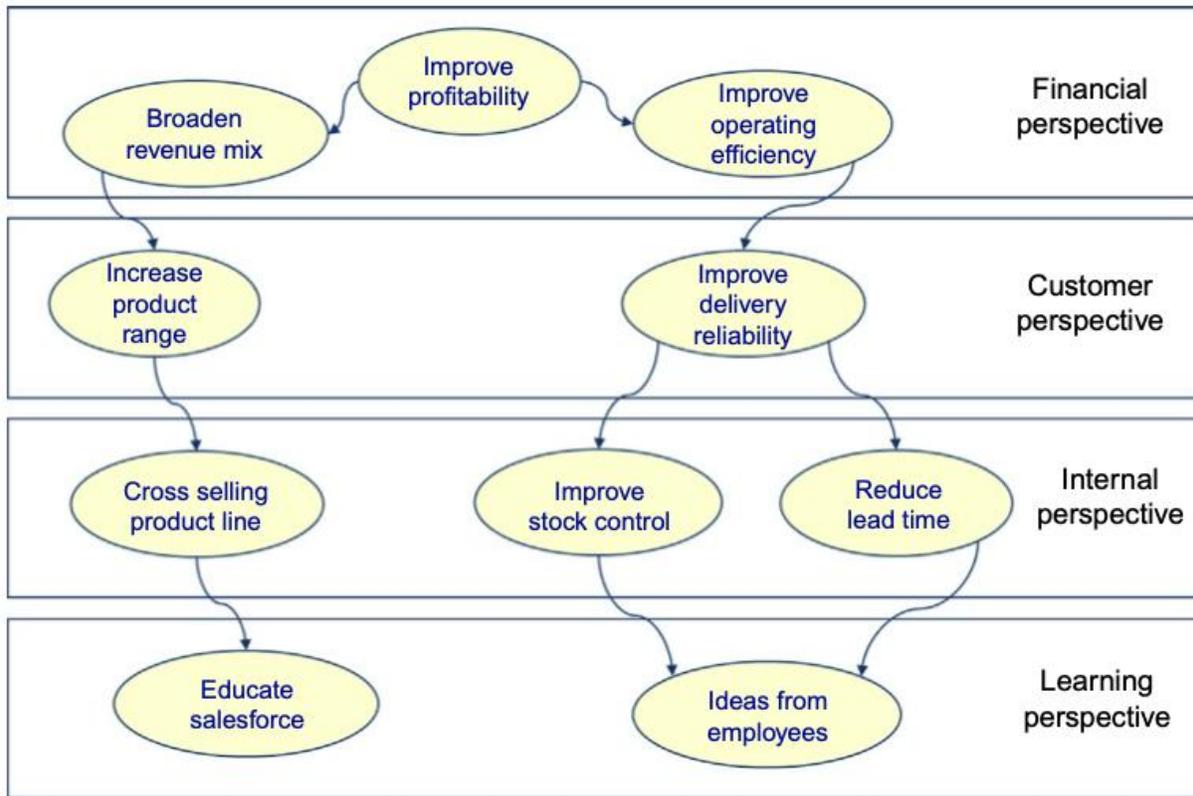


Figure 2 - The cascading process example (Arnaboldi, Azzone and Giorgino, 2015)

Once the map is ready and the objectives are laid out, each of them is assigned with an indicator; the four sets of indicators defined become the indicators to be included in the BSC (Arnaboldi, Azzone and Giorgino, 2015).

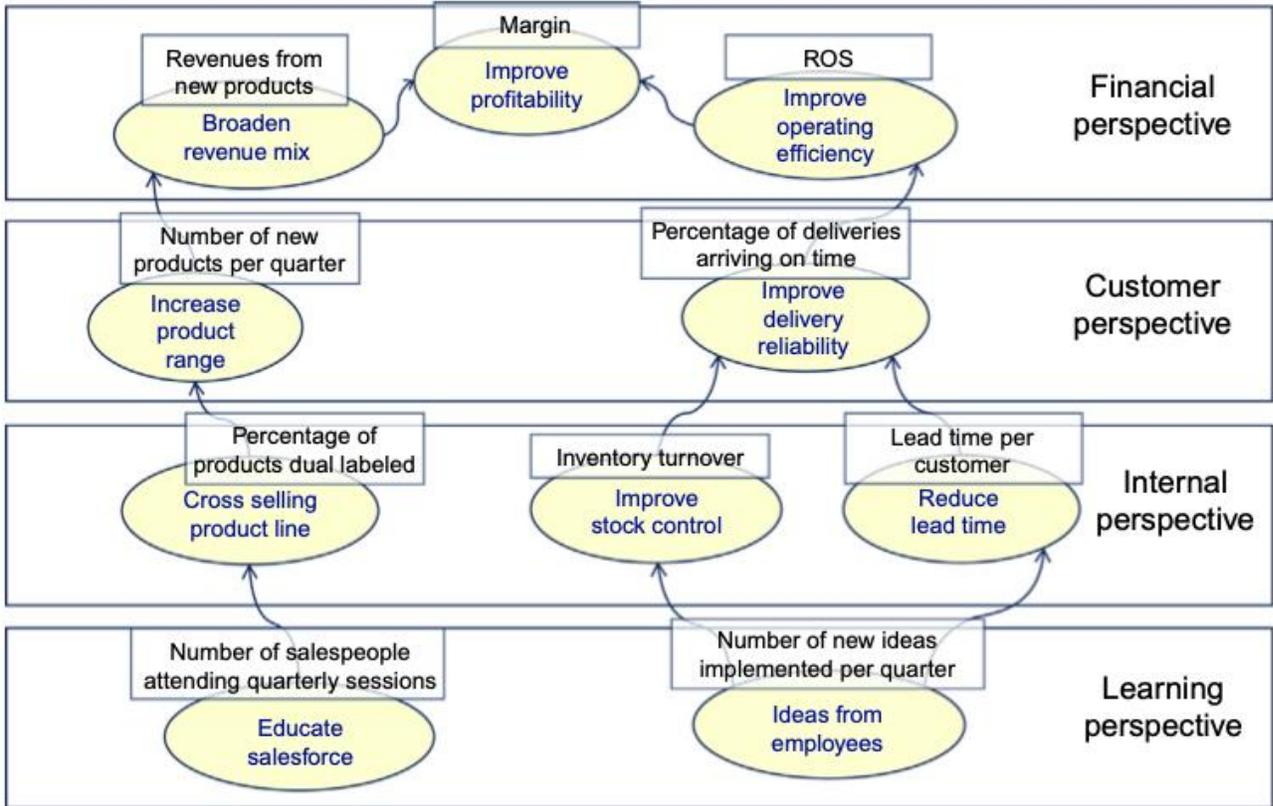


Figure 3 - An example of a strategy map (Arnaboldi, Azzone and Giorgino, 2015)

Different types of indicators can be adopted for measuring performances (Arnaboldi, Azzone and Giorgino, 2015):

- *value-based indicators*, aim to control enterprises, measuring the creation of value and its risks in a more or less direct way.
- *accounting indicators*, indicators constructed starting from financial statements: the profit & loss (P&L) account, balance sheet, and cash flow statement.
- *value drivers*, indicators that provide earlier signals (drivers) of the value creation

In the business world, market forces serve as feedback mechanisms. Performance is relatively easy to quantify through accounting indicators, since such metrics can be compared thereby ensuring that the companies producing the best results will attract new capital.

However, considering the diversity of goals museums pursue, assessing performance is a more subjective and value-driven exercise for museums than it is for companies.

The application of the Balanced Scorecard to non-profit organizations, and therefore also to museums faces a basic challenge, which can be overcome by adjusting the BSC with respect to the original setting. The basic challenge posed by non-profit organizations to BSC concerns the fulfilment of the social purpose of their activity: the social purpose is usually directed toward multiple primary stakeholders, not only toward shareholders (Basso, Casarin and Funari, 2018).

The classification of value driver indicators presented beforehand is solid and proven. It is however necessary to establish the characteristics of the specific indicators geared towards social media. This necessity stems from the fact that the way social-media related indicators are sourced and calculated is fundamentally different compared to standardized reporting methods.

3.2 User-Generated-Content platforms and Social Media

Using the standard classification, indicators generated from social media data are clearly a value driver indicator type. They fit the characteristics of value drivers, and the assumption that social media generate value is proven by the literature as well as by the evolution that many industries, notably the marketing and entertainment industry, underwent in the past decade. However, it is not easy to fit this class of indicator into the *performance* or *resource* indicators. Under the performance indicator label, social media indicators can be seen as reporting on marketing activities, but they can only marginally be tied to revenues as the correlation between the performances on social media and revenues and costs of the company is very complex and very hard to isolate. A different approach can be used, and social media indicators can be put under the resource indicator label, included in the intellectual capital, as the relations and reputation built up by a social media account are a resource that needs to be nurtured and from which value needs to be harvested. What sets social media indicators apart from indicators relating marketing effort and intellectual capital is mainly the approach used.

Data used to calculate social media indicators are sourced through the different UGC platforms. The main advantage of this is that both the data sourcing and the data processing can be automated and present common standard practices that are applicable cross-platforms. Techniques such as the sentiment analysis can be applied to generate data that are easier to interpret.

Because of that, social media indicators are fundamentally different from indicators such as a *brand value* that needs extensive research and cannot be calculated as easily. The latter relies on data from many different sources and sometimes not readily available, complicating

the data sourcing process even further with sources such as customer questionnaire that adds a layer of complexity.

It is also important to understand the implication that each platform bring with its data, as different platforms offer data that could be comparable, but have different meaning, as platforms have different demographic, different functions, different use scenario and are subject to different trends and users.

User-generated-content (UGC)

User-generated-content (**UGC**), alternatively known as user-created content (**UCC**), is any form of content, such as images, videos, text, and audio, that has been posted by users on online platforms such as social media and wikis. It comes from regular people who voluntarily contribute data, information, or media that then appears before others in a useful or entertaining way, usually on web. The use of such contents has seen a rapid growth in recent years, in part because it's fairly inexpensive to obtain as users normally supply it for no charge (Krumm, Davis and Narayanaswami, 2005).

There are many types of user-generated content platforms:

- *Internet forums*, where people talk about different topics;
- *Blogs*, services where users can post about many topics;
- *Wikis*, such as Wikipedia and Wikia allow users, sometimes including anonymous users, to edit the content;

- *Social networking sites* alternatively known as *Social Media* like Facebook, Instagram, Twitter, where users interact writing messages, posting images or links, and sharing content;
- *Media hosting sites* such as YouTube allow users to post content;
- *Other types of user-generated-content platforms* as Trip Advisor or Google Local Guides a functionality of Google Maps;

Later-on we will focus on the most used Social Media as of January 2020 (Instagram, Facebook and Twitter) and on Trip Advisor and Google Local Guides a functionality of Google Maps, which jointly represent the set of user generated platforms most used by cultural institutions when implementing communication strategies.

Social media

Social Media is a group of Internet-based applications that build on the ideological and technological foundations of *Web 2.0*, that allow the creation and exchange of *User-generated-content* (Kaplan and Haenlein, 2010).

Actually, social media can be thought of in a few different ways. In a practical sense, it is a collection of software-based digital technologies, usually presented as apps and websites, that provide users with digital environments in which they can send and receive digital content or information over some type of online social network. In this sense, we can think of social media as the major platforms and their features, such as Facebook, Instagram, and Twitter. We can also in practical terms of social media as another type of digital marketing channel that marketers can use to communicate with consumers through advertising. But

we can also think of social media more broadly, seeing it less as digital media and specific technology services, and more as digital places where people conduct significant parts of their lives. From this perspective, it means that social media becomes less about the specific technologies or platforms, and more about what people do in these environments. To date, this has tended to be largely about information sharing, and, in marketing, often thought of as a form of online word of mouth (**WOM**) (Appel, Grewal, Hadi and T. Stephen, 2019).

In synthesis social media are digital platforms, services and apps built around the convergence of content sharing, public communication and interpersonal connection (Burgess, Marwick and Poell, 2019)

Definition of social media elements relevant for the thesis:

- *Post*, something such as a message or picture that users can publish on a website or using social media (Cambridge Dictionary). Posts can take different forms depending on the social media, but in general they are either text, a video, a picture or a gif;
- *Follower*, every social media features a feed composed by UGC selected based on different factor. Users can become follower of a content creator to make sure that his new contents are always featured in the feed;
- *Tag*, a link to someone's profile from a photo or comment on a website such as Instagram or Facebook (Cambridge Dictionary). Generally, when a tag is used, the person tagged is being notified;

- *Hashtag*, a hash (#) followed by a word, used on social media for describing the general subject of a post (Cambridge Dictionary). This can be used for post or contents relating to trends or marketing campaign;
- *Comment*, a mean through which users can interact with posts, by commenting it. The comments can usually be seen underneath the post;
- *Sharing*, a mean through which users can share a post with their followers, by displaying it on their profile;
- *Stories*, is a format that allows people to share photos or short videos that will disappear within 24 hours. Stories can be enhanced by the use of stickers, emoji, GIFs, real-time face-tracking augmented reality effect-, post-production filters or camera effects.

According to statistics as of January 2020, 4.54 billion people are active internet users, encompassing 59% of the global population (Clement, 2020). Social media usage has become an integral element to the lives of many people across the world. In 2019 2.95 billion people were active social media users worldwide.

3.2.1 UCG Platforms Evolution

Web 1.0: "Read Only Web"

The first websites of the World Wide Web, labelled today as Web 1.0, allowed a narrow and chosen set of people to create the website, resulting in a one-way flow of information

towards the user. The information framework of Web 1.0 sites was simple and static, user involvement in terms of interaction was highly restricted since they had no control over the content of the website unless they were the authors.

Some of the important site features that mark out a Web 1.0 identified by O'Reilly founder and CEO of O'Reilly Media are presented:

- *Web 1.0 sites are static*, they contain information that might be useful, but there's no reason for a visitor to return to the site later;
- *Web 1.0 sites aren't interactive*, visitors can only visit these sites; they can't impact or contribute to the sites. Most organizations have profile pages that visitors can look at but not impact or alter;
- *Web 1.0 applications are proprietary*, under the Web 1.0 philosophy, companies develop software applications that users can download, but they can't see how the application works or change it.

The leap from Web 1.0 to Web 2.0 is not only characterized by redefining the users role on the World Wide Web, but also involved a technological improvement (Viswanathan, Mathur, Yammiyavar, 2010).

Web 2.0: "Read-Write Web"

Web 2.0 is term that was introduced in 2004 and refers to the second generation of the World Wide Web (Pal, 2012). The shift to this more interactive web from Web 1.0 generally occurred as a result of technological changes that made the internet, and the ability to develop content, more accessible.

The structure of the Web 2.0 emphasizes user-generated content, participatory culture and interoperability for end users, it has extended the user role which has been empowered with the right to create content and publish it on the web all by themselves while interacting with other users. Additionally, the dynamic process that arises between users is heterogeneous in space and time.

There was a huge change in user behavior, user demands and needs. Instead of merely reading a Web 2.0 site, a user is invited to contribute to the site's content by commenting on published articles or creating a user account or profile on the site, which may enable increased participation (Kabari and Osuo-Genseleke, 2008).

Some of the important site features that mark out a Web 2.0 identified by G.Cormode, B. Krishnamurthy are presented:

- *Web 2.0 users as first-class entities in the system*, with prominent profile pages, including such features as: age, sex, location, testimonials, or comments about the user by other users;
- *Web 2.0 users can form connections between them*, via links to other users who are “friends”, membership in “groups” of various kinds;
- *Web 2.0 users have the ability to post content in many forms*: photos, videos, blogs, comments and ratings on other users’ content, tagging of own or others’ content, and some ability to control privacy and sharing.

These features fulfilled user needs such as collective intelligence, collaborated work, learning and community behavior as well as social interactions (Viswanathan, Mathur, Yammiyavar, 2010).

Some examples of features considered part of Web 2.0 are listed below:

- *Blogs*, also known as Web logs, these allow users to post thoughts and updates about their life on the Web;
- *Wikis*, sites like Wikipedia and others enable users from around the world to add and update online content;
- *Social networking*, sites like Facebook and Instagram allow users to build and customize their own profile and communicate with friends;
- *Web applications*, a broad range of new applications make it possible for users to run programs directly in a Web browser (Pal, 2012).

Web 2.0 technologies provide a level user interaction that was not available before. Websites have become much more dynamic and interconnected, producing "online communities" and making it even easier to share information on the Web. Because most Web 2.0 features are offered as free services, sites like Wikipedia and Facebook have grown at amazingly fast rates (Pal, 2012).

Web 3.0: "Read-Write-Execute Web"

Decentralized Web or Web 3.0 describes services build on the internet which do not depend on any single "central" organization to function.

Search Engines (e.g. Google), Social Networks (e.g. Facebook), Chat Apps (e.g. WhatsApp) have grown huge by providing centralized services on the internet. The Decentralized Web envisions a future world where services such as communication, currency, publishing, social networking, search, archiving etc. are provided not by centralized services owned by single

organizations, but by technologies which are powered by the people: their own community (Hodgson, 2016).

For the sake of convenience, we conceptualize Web 3.0 as a new platform which would include the following features identified by Ganesh Viswanathan, Punit Dutt Mathur, and Pradeep Yammiyavar:

- *Portability*, infotainment anywhere, anytime;
- *Personal*, focused on the individual, as opposed to community;
- *Dynamic and Contextual content*, handling the aspect of Semantic Web

When addressing the meaning of Web 2.0 one must present the meaning of the so-called Semantic Web. Semantic Web has the goal to make Internet data machine-readable and it is a feature of the World Wide Web which allows to match the requests of people and machines while they are dynamically interactive. For example, a search engine that understand who you are, what you've been doing and what you are going to do next all by itself in an intelligent way.

One of the biggest challenges of presenting information on the web is that applications cannot provide context to data, and, therefore, can't understand what is relevant. Through the use of some sort of semantic markup (or data interchange formats), data could be put in a form not only accessible to humans via natural language, but able to be understood and interpreted by software applications as well, applications that can speak to each other directly and interpret information for humans.

3.2.2 UCG Platforms Classification

As User Generated Platforms are a complex subject which includes numerous approaches, tools and techniques, it can be useful to try to identify the main categories. Based on the definitions of (Kaplan and Haenlein, 2010) the following types of UGC platforms can be identified:

- *Collaborative blogs or collaborative projects*, Collaborative projects enable the joint and simultaneous creation of content by many end-users and are, in this sense, probably the most democratic manifestation of **UGC**. Within collaborative projects, one differentiates:
 - Wikis, websites which allow users to add, remove and change text-based content
 - Social bookmarking applications, which enable the group-based collection and rating of Internet links or media content.

The main idea underlying collaborative projects is that the joint effort of many actors leads to a better outcome than any actor could achieve individually

- *Blogs, Microblogs, and Internet Forums*, are special types of websites that usually display date-stamped entries in reverse chronological order [OECD, 2007]. They are considered personal web pages where individuals can communicate with one another through texts and other multimedia like videos, audios, and pictures (Dao, 2015).
- *Content Communities*, the main function is sharing media contents (e.g., videos, pictures, audio clips, and PowerPoint presentations) between users. Internet users can create an account for each of the Content Communities, upload media content to

the sites, and share the media content with one another. The simplest way to share media content is to send a link to others or to post a link to the personal blog. Internet users can access media content by clicking on the link that leads to the content page (Dao, 2015).

- *Social Networking*, Facebook, LinkedIn, Instagram are considered the most common social networking sites. The site applications allow the Internet users to create personal profiles, invite others to join the site, access the profiles of other users, share information (text, pictures, videos, and other links from the other sites), and send emails as well as instant messages to each other (Dao, 2015).
- *Virtual Game Worlds*, virtual worlds are platforms that replicate a three-dimensional environment in which users can appear in the form of personalized avatars and interact with each other as they would in real life (Kaplan and Haenlein, 2010).
- *Virtual Social Worlds*, this type allows Internet users to choose personalized avatars, their behaviors, their lives, and their acts in their virtual lives that are similar to their real lives. The users can meet each other at a certain place in the virtual world to do certain things together as they do in real life (Dao, 2015).

Kaplan & Haenlein (2010) classified social media sites into these six types based on:

- Media reach, including social presence and media richness
- Social process, including self-presentation and self-disclosure.

Regarding the *media-related component* of Social Media:

- *Social presence* theory (Short, Christie and Williams, 1976) states that media differ in the degree of “social presence” defined as the acoustic, visual, and physical contact

that can be achieved. Social presence is influenced by the intimacy (interpersonal vs. mediated) and immediacy (asynchronous vs. synchronous) of the medium, and can be expected to be lower for mediated (e.g., telephone conversation) than interpersonal (e.g., face-to-face discussion) and for asynchronous (e.g., e-mail) than synchronous (e.g., live chat) communications. The higher the social presence, the larger the social influence that the communication partners have on each other's behavior.

- *Media richness theory* (Daft and Lengel, 1986) is based on the assumption that the goal of any communication is the resolution of ambiguity and the reduction of uncertainty. It states that media differ in the degree of richness they possess, that is, the amount of information they allow to be transmitted in a given time interval and that therefore some media are more effective than others in resolving ambiguity and uncertainty.

Regarding the *social-process component* of Social Media:

- *Self-presentation theory* (Goffman, 1959) states that in any type of social interaction people have the desire to control the impressions other people form of them. On the one hand, this is done with the objective of influencing others to gain rewards, on the other hand, it is driven by a wish to create an image that is consistent with one's personal identity.
- *Self-disclosure* is the conscious or unconscious revelation of personal information (e.g., thoughts, feelings, likes, dislikes), consistent with the image one would like to

give. Applied to the context of Social Media, we assume that a second classification can be made based on the degree of self-disclosure it requires

Combining both dimensions leads to a classification of Social Media which we have visualized in Table 1 below.

Table 1 - Classification of Social Media by social presence/media richness and self-presentation/self-disclosure

Self-presentation/Self-disclosure	high		Social networks (i.e. Facebook) Virtual communities (blogs, micro-blogs)	Virtual social world (i.e. Second Life)
	medium	Virtual communities (commercial review websites) (i.e. TripAdvisor)	Content communities (i.e. Youtube, Pinterest)	
	low	Collaborative projects (i.e. Wikipedia)		Virtual games (i.e. Xbox, Playstation)
		low	medium	high
		Social presence/Media richness		

For the purpose of this thesis, we considered five major platforms: Facebook; Instagram; Twitter; TripAdvisor; Google Maps. Despite being similar in concept, each platform has unique features and user demographics, and can gather and provide different data. A more detailed deep dive over these five platforms is available at Chapter 9.2 (Annex).

3.2.3 UCG Platforms Business Applications

As we can conclude from the previous analysis, the jungle of Social Media changed the way consumers used Internet, users behavior radically changed from a passive approach that included reading and watching content to a more proactive approach: sharing, modifying and discussing about Internet content. *This Social Media phenomenon can have a sound effect on a firm business or institution success in terms reputations, revenues and survival, in particular traditional marketing methods can no longer sustain a business, social connectivity is becoming the key to marketing.*

Communication about brands happens, with or without permission of the firms in question and the power has been taken from those in marketing and public relations by the individuals and communities that create, share, and consume blogs, tweets, Facebook entries, movies, pictures, and so forth (Kietzmann, Hermkens, McCarthy and Silvestre, 2011).

In the last years there has been a shift from traditional marketing, companies using the “4Ps” to manage marketing-mix variables to Relationship Marketing, that focus on the identification a long-term retention of customers. This kind of new way of performing marketing activities points out that is more profitable to keep satisfy existing customers rather than constantly renew the customer base. In the case of cultural institutions loyal customers not only may be keener to take part of new events or temporarily expositions but a customer loyal base would unblock the e-WOM mechanism, then information for already loyal customers can generate leads and launch a positive loop.

- *Electronic word-of-mouth communication (eWOM)*: is any positive or negative statement made by potential, actual or former customers about a product or company which is made available to multiples of people and/or institutions and is pread over the internet.
- *The term "4Ps"*: also known as "Marketing Mix" is a foundation model for businesses developed by E. Jerome McCarthy in th 1960s, historically centered around product, price, place, and promotion (also known as the "4 Ps").

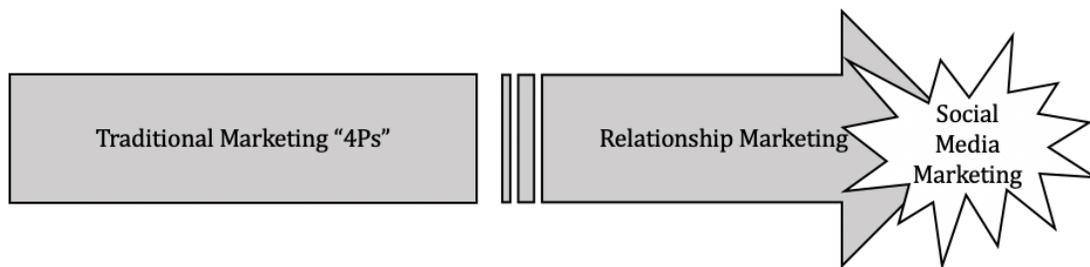


Figure 4 - Illustrative representation of marketing evolution

We conduct an analysis using Google Trends tool to compare the evolution of customers interest in Social Media Marketing, Influencer Marketing, Radio Advertisement, Tv Advertisement and Printed Advertisement in the last 6 years (from 2004 to 2020).

An eye-ball analysis is enough to acknowledge the growth of influencer marketing and social media marketing, compared to more traditional marketing tools as Tv, Radio Marketing and Printed Advertisiement.

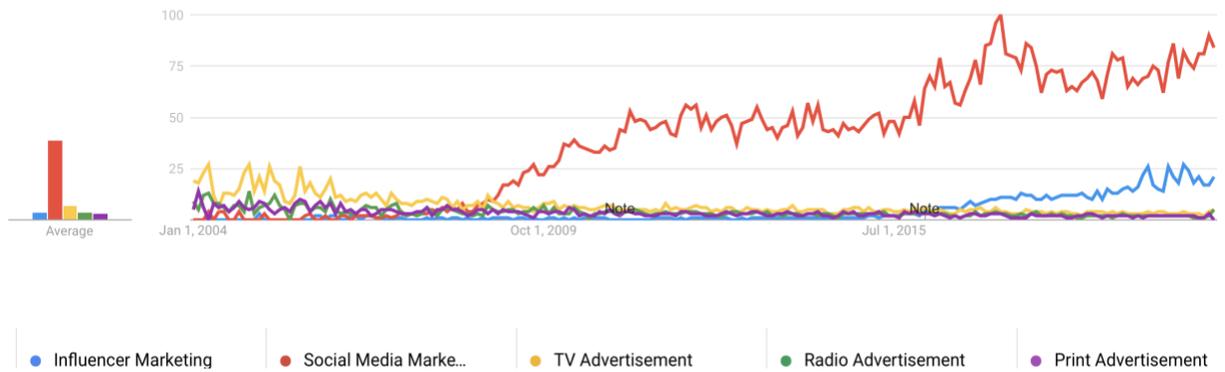


Figure 5 - Influencer Marketing, Social Media Marketing, TV Advertisement, Radio Advertisement, Printed Advertisement trend in terms of interest overtime compared

Although it is clear that for better or for worse social media is very powerful, many executives are reluctant or unable to develop strategies and allocate resources to engage effectively with social media (Kietzmann, Hermkens, McCarthy and Silvestre, 2011). Consequently, firms regularly ignore or mis-manage the opportunities and threats presented by creative consumers (Berthon, Pitt, McCarthy and Kates, 2007).

3.2.4 Measuring the degree of corporate social media use

Making good use of user-generated platforms is not enough if a model to track its success is not well defined. Finding an effective way to measure the degree of corporate social media use is, and has been a growing field of research since the arise of Web 2.0.

In this paragraph we will present an approach developed by Thomas Aichner from University of Padova and Frank Jacob from ESCP Europe, developed in 2015 which aims to provide a model with which to measure the degree of corporate social media use or, in other

words, the extent to which companies are exploiting the potentialities of single or multiple social media platforms. This is, however, explicitly different from using metrics to assess the success of social media activities, as it is purely measuring how intensively a pre-defined group of social media is utilised, taking into account the frequency of social media activity by the brand as well as the related user reactions (Aichner and Jacob, 2015).

The main assumption of the analysis relays on the idea that companies, firms or cultural institutions have to adopt innovative ways to reach potential customers through user-generated-content platforms social media. It is well known that, generally speaking, it can be very helpful for companies or cultural institutions to set themselves objectives that should be properly tracked and monetarized in order to assess the “as is” situation and to understand where to put their effort in order to improve their results.

In this model the effort and results will be expressed in terms of number of posts, number of interactions and average response time to enquiries.

When taking into account the Social Media world, it is not enough for companies to simply ‘be there’ and to have a social media account. They need to update their profiles and websites regularly and be highly reactive to customer requests in social media in order for their presence there to be effective. Posts, competitions, content uploads and other activities have to be planned well in advance and require companies to coordinate their use of different social media (Aichner and Jacob, 2015).

In the table below a study that shows the importance and impact of social media for different corporate functions will be presented.

Table 2 - Importance of social media for different corporate functions.

Type of social media	Corporate function					
	R&D	Marketing	Customer service	Sales	HR	Organisation
Blogs						
Business networks						
Collaborative projects						
Enterprise social networks						
Forums						
Microblogs						
Photo sharing						
Products/services review						
Social bookmarking						
Social gaming						
Social networks						
Video sharing						
Virtual worlds						

Importance: (empty) none or almost none; low; medium; high; very high

This model we are going to describe proposes a method of how to measure the extent to which companies are using a pre-defined group of social media, called the ‘degree of corporate social media use’ (CSMU). This indicator is useful and should be combined with social media metrics in order to draw better conclusions about where to increase or intensify social media activities. In addition, it can be used to easily compare the degree of social media use by companies and brands (Aichner and Jacob, 2015).

The Model

Before entering into the four-steps approach that describes the model some definitions will be presented, in order to guarantee a straightforward reading.

Corporate social media use (CSMU): defined as the degree to which companies or brands are utilising a pre-defined group of social media. It is essential to have a standardized ratio scale in order to simplify the measurement, and to allow a reliable, clear and transparent comparison of companies' and brands' social media use (Aichner and Jacob, 2015).

Active user: is defined as a registered user who logs in to his or her account, irrespective of whether or not he or she actually reads, posts or shares something, and regardless of how long he or she remains on the social media website. For those platforms that do not require a registration to use the most important functions, e.g. YouTube, the number of unique page visitors is taken instead (Aichner and Jacob, 2015).

Social media impact factor (SMIF): it is based on the monthly active users of the social media included in the analysis and determines the relative importance of each social media platform in the given framework. It may change according to the number of active users over time and if more platforms are added to the model (Aichner and Jacob, 2015).

Social media use (SMU): In order to find the optimal degree of utilisation for each user-generated-content platform, the activities three well-known and popular global brands have been investigated by Thomas Aichner and Frank Jacob¹:

- Intel Corporation (Intel), an American semiconductor chip maker;
- Samsung Mobile, a business unit of Samsung Electronics Corporation, a South Korean electronics company;

The Walt Disney Company (Disney), an American entertainment and media enterprise.

Based on the activities of these three companies, an index of social media use (SMU), ranging from zero (0 = no use at all) to one (1 = full use) is established. The average values serve as a benchmark that equals one (1 = full use), representing the optimal degree of utilisation of the specific platform (Aichner and Jacob, 2015).

Step 1: monthly active users

First, the five leading global social media in terms of monthly active users are identified.

The ones used in this model are the following, describing the ranking as of January 2014:

¹ As global brands usually have several accounts on most social media, e.g. a Facebook account in different countries and/or in different languages, the corporate account with the highest number of fans of Intel, Samsung Mobile and Disney was selected for the study. The data were collected from 1 November 2013 to 31 December 2013, and include text, pictures, video and job postings on each of the respective platforms, which have all been assessed individually. It does not include, however, announcements of events, replies to customer requests and special activities such as competitions, unless they are promoted with a traditional text, picture or video posting.

- facebook.com (Facebook)
- youtube.com (YouTube)
- plus.google.com (Google+)
- linkedin.com (LinkedIn)
- twitter.com (Twitter)

Step 2: social media impact factor

Second, the social media impact factor (SMIF) is calculated. It is based on the monthly active users of the five social media included in the analysis, and may change according to the number of active users over time and if more platforms are added to the model. The SMIF is calculated by dividing the active users of each platform (AUplatform) by the sum of the active users of all platforms ($\sum AU_{platform}$) included in the model (Aichner and Jacob, 2015).

$$SMIF\ Platform = \frac{AU\ platform}{\sum AU\ platforms}$$

The number of social media can easily be reduced or expanded by taking into account fewer or more websites. When more platforms are included in the model, the relative importance of each is reduced.

Step 3: social media use

For each user-generated-content platform a criteria to calculate the corporate use of the respective platform is assessed is established.

This index may not only change depending on the type of social media, but differences also surface between platforms of the same type based on the tools and range of functions offered by the platform. This model provides specific criteria about how to calculate the use of five individual social media platforms, namely Facebook (social network), YouTube (video sharing), Google+ (social network), LinkedIn (business network) and Twitter (microblog). The use of other social media can be calculated in the same or in a similar way, depending on the range of applications, e.g. no option for users to comment on posts (Aichner and Jacob, 2015).

The index is calculated by taking into account publicly available information about the frequency of social media activity by the brand (text, picture, video and job postings) as well as the related user reactions (comments, shares, 'likes' and job applications) by its followers, subscribers or fans. As all social media platforms offer different features, the SMU of a platform (SMUplatform) is a customised function of the number of the social media activities by the brand and the user reactions to each individual activity (Aichner and Jacob, 2015).

$$SMU Platform = f(\text{social media activities, user reactions})$$

Deep dive into user reactions

It is common knowledge that while a 'like' (or a '+1', 'thumbs up', etc., depending on the social media platform) expresses a certain degree of positive interest, comments may also be negative and need to be classified accordingly in the analysis of the five social media, in this model a conversion rate is defined as follows:

Positive comments in every Social Media show a higher degree of identification so are weighted 5 times the weight of 'likes'.

Shares (or retweets', etc., depending on the social media platform), are weighted 10 times the weight of 'likes' because they have an important multiplier effect.

In the case of YouTube, 'likes' are weighted 100 times higher than views.

Both the analysis of the five social media platforms and the SMU formulas are adjusted to the type of social media and its individual applications. In order to calculate the average values, the user reactions are weighted by the number of corporate social media activities, e.g. comments by video uploads or job applications by job postings. Companies can use the following formula to calculate the SMU of Facebook and will receive a value between zero (0 = no use at all) and one (1 = full use). The result of the equation can exceed the optimal value of one, in which case the SMU shall be equal to one (Aichner and Jacob, 2015).

Facebook

Companies can use the following formula to calculate the SMU of Facebook and will receive a value between zero (0 = no use at all) and one (1 = full use). The result of the equation can exceed the optimal value of one, in which case the SMU shall be equal to one.

$$SMU\ Facebook = posts * \frac{\emptyset 'likes' + \emptyset comments * 5 + \emptyset shares * 10}{fans} * FbConstant$$

$$FbConstant = \left(monthly\ posts * \frac{'likes'' + comments * 5 + shares * 10}{fans} \right)^{-1}$$

The FBConstant is calculated using the average values of likes, comments, shares fans and posts of the selected companies used for the benchmark analysis (in the case of the Thomas Aichner and Frank Jacob model data from Intel, Samsung Mobile and Disney was used to conduct the analysis).

YouTube

On YouTube, companies can create corporate video channels and upload videos, e.g. product tests, commercials and interviews. The videos can usually be watched by anyone who visits the website, while 'liking' and commenting on them is restricted to registered users, who can also subscribe to a channel. Even though videos can be shared in other social media, and despite the fact that this is an important factor, shares are not considered in this research, as it is impossible to reveal the number of shares. It is important to mention that, in contrast to postings in social networks or in microblogs, where posts are usually popular only for a few days or for as long as they are visible on the user's news page, visitors to video-sharing platforms generally search for videos using the platform's search function, and can find and watch videos that were uploaded years earlier. Therefore, in order to guarantee consistency, the analysis of the SMU of YouTube and other video-sharing platforms should be done within a time range of two to ten weeks after the videos have been posted. In addition, and in contrast to Facebook, users can 'dislike' a video. In light of this, Table 6 shows the number of subscribers, the number of video uploads, the average numbers of views, the adjusted average number of 'likes' minus 'dislikes', and the average number of comments for each of the three corporate YouTube channels analyzed (Aichner and Jacob, 2015).

Companies can use the following formula to calculate the SMU of YouTube and will receive a value between zero (0 = no use at all) and one (1 = full use). The result of the equation can exceed the optimal value of one, in which case the SMU shall be equal to one.

$$SMU YouTube = video\ uploads * \frac{\emptyset views + \emptyset'likes' * 5 + \emptyset comments * 500}{subscribers} * YBConstant$$

$$YBConstant = \left(monthly\ video\ uploads * \frac{\emptyset views + \emptyset'likes' * 5 + \emptyset comments * 500}{subscribers} \right)^{-1}$$

The YBConstant is calculated using the average values of views, likes, comments and subscribers of the selected companies used for the benchmark analysis (in the case of the Thomas Aichner and Frank Jacob model data from Intel, Samsung Mobile and Disney was used to conduct the analysis).

Google+

Similar to Facebook, as of 2015 Google+ offered companies the possibility to create a dedicated page, write text postings, and upload pictures or videos. Registered users could add the page to their circle, which means that were are subscribing for all news posted by the respective Google+ page. If users like a post, picture or video, they can '+1' ('like') or share it with their individual network (Aichner and Jacob, 2015).

Companies can use the following formula to calculate the SMU of Google+ and will receive a value between zero (0 = no use at all) and one (1 = full use). The result of the equation can exceed the optimal value of one, in which case the SMU shall be equal to one.

$$SMU GooglePlus = posts * \frac{\emptyset'+1s' + \emptyset shares * 10 + \emptyset comments * 5}{in\ circles} * GooglePlus\ Constant$$

$$GooglePlus\ Constant = \left(monthly\ posts * \frac{\emptyset '+1s' + \emptyset\ shares * 10 + \emptyset\ comments * 5}{in\ circles} \right)^{-1}$$

The Google Plus Constant is calculated using the average values of '+1s', shares, comments, and in circles of the selected companies used for the benchmark analysis (in the case of the Thomas Aichner and Frank Jacob model data from Intel, Samsung Mobile and Disney was used to conduct the analysis).

Twitter

On Twitter, as of 2015 companies could create an account and post short tweets (text or picture postings) with a maximum length of 140 characters. Registered users could add the tweets to their favourites (similar to 'likes'), post an answer (comment) or retweet (share) the message. However, the answers to posts are not taken into account when calculating the SMU of Twitter, as it is difficult to identify which and how many answers are actually related to a specific post (Aichner and Jacob, 2015).

Companies can use the following formula to calculate the SMU of Twitter and will receive a value between zero (0 = no use at all) and one (1 = full use). The result of the equation can exceed the optimal value of one, in which case the SMU shall be equal to one.

$$SMU\ Twitter = \frac{\emptyset\ favourites + \emptyset\ retweets * 10}{followers} * Twitter\ Constant$$

$$Twitter\ Constant = \left(monthly\ posts * \frac{\emptyset\ favourites + \emptyset\ retweets * 10}{followers} \right)^{-1}$$

The Twitter Constant is calculated using the average values of posts, favourites, retweets and followers of the selected companies used for the benchmark analysis (in the case of the

Thomas Aichner and Frank Jacob model data from Intel, Samsung Mobile and Disney was used to conduct the analysis).

Step 4: calculating the degree of corporate social media use

As both the social media impact factor (SMIF) and the equations to calculate the social media use (SMU) are identified for the most important social media in terms of monthly active users, it is possible to calculate the degree of corporate social media use (CSMU) of any company. Companies can choose whether to include only one or up to five important social medias to determine the degree of social media use. If they choose to assess only one platform, the SMIF does not need to be used, as its function is to weight the SMU of the individual platforms by their relative importance. Whenever one, two or three websites are removed from the model, the SMIF needs to be recalculated. However, while a reduction of the number of social media in the model is unproblematic, adding new platforms requires a prior determination of the SMU equation (Aichner and Jacob, 2015).

$$CSMU_{company} = \sum_{\text{Social media } i}^N SMIF_i * SMU_i$$

N= Number of Social Media included in the model

Conclusions

The model presented in this article can help market researchers and marketing managers to measure and compare the degree of corporate social media use (CSMU), both for individual platforms and for up to five different social media;

The tool presented provides a simple, clear and consistent framework that allows comparison of the use of social media among brands, companies, markets and industries;

Practitioners can apply the CSMU in order to identify where to increase or reduce social media activities since it enhances the general comparability of quantitative studies about the use of social media;

The CSMU helps companies understand whether or not to allocate or shift human and financial resources, which increases the probability of success in the online world. The CSMU is a powerful yet easy to understand indicator that allows companies to monitor internal social media activities, build industry averages and keep track of their competitors (Aichner and Jacob, 2015).

3.2.5 Social Media Analytics

Social media analytics refer to the analysis of structured and unstructured data from social media channels, a field emerged alongside Web 2.0 (Gandomi and Haider, 2015).

Social media offer two kinds of data:

- user generated contents;
- relationships and interactions between network entities.

Relevant **UGC** platforms and the interactions and relationships between the entities on those platforms have been discussed in previous sections of this thesis. This division between type of data determines two different groups of social media analytics: *content* based and

structure based analytics. The latter can be represented through a nodes and edges model, in which nodes are the entities, but edges change their meaning depending on existing relationships (like friendship on Facebook) or active interaction.

Various techniques have recently emerged to extract information from the structure of social networks.

Community detection also referred to as community discovery, extracts implicit communities within a network. For online social networks, a community refers to a sub-network of users who interact more extensively with each other than with the rest of the network. Often containing millions of nodes and edges, online social networks tend to be colossal in size. Community detection helps to summarize huge networks, which then facilitates uncovering existing behavioural patterns and predicting emergent properties of the network (Gandomi and Haider, 2015). Community detection presents similarities to clustering (Aggarwal, 2011), which is a mining technique used to partition data sets. Community detection has applications in several areas, including product recommendation systems.

Social influence analysis it refers to techniques that are concerned with modelling and evaluating the influence of actors and connections in a social network, as the behavior of an actor in a social network is affected by others (Gandomi and Haider, 2015). The elements evaluated are the participants' influence, the strength of connections, and the patterns of influence diffusion in a network. Social influence analysis is useful to evaluate brand awareness and in viral marketing.

Link prediction is a technique to specifically address the problem of predicting future linkages between the existing nodes in the underlying network (Gandomi and Haider, 2015),

considering that the structure of social networks is not static and continuously grows through the creation of new nodes and edges in time, to understand and predicts its dynamics and the future occurrence of interaction, collaboration or influence among entities. Link prediction techniques outperform pure chance by factors of 40–50, suggesting that the current structure of the network surely contains latent information about future links (Liben-Nowell and Kleinberg, 2003).

In the context of online social media, the primary application of link prediction is in the development of recommendation systems, such as Facebook's "People You May Know", YouTube's "Recommended for You", and Netflix's and Amazon's recommender engines (Gandomi and Haider, 2015).

3.2.6 Sentiment analysis

Sentiment analysis corresponds to the process of identifying the sentiment associated with a piece of text. It usually relies on applying machine learning techniques to classify documents based on a collection of features extracted from the text using other NLP techniques, such as the presence of certain words or the coverage of some topics. (Antoniou, Dimitriou and Pereira, 2019)

With the explosion of Internet-based social media, society has witnessed not only a new tool for information sharing and spread but also a whole new economy. Since the new generation of smartphones that arrived in 2007 with the iPhone and thereafter the Android the economy of social media has impacted the classic economy on a very frequent basis. Thus,

followers of social media can easily share their opinion on their experiences and preferences as well as share their choices on new trends generated by the media. This crossbreeding effect changed many traditional businesses as marketing research, advertising, and communication significantly boosted e-commerce (Frunza, 2016).

Approaches to sentiment analysis

Researchers can use different approaches to evaluate pieces of text, but they can be boiled down to:

- Rule based
- Machine learning

Rule based means that the system follows a set of manually crafter rules to evaluate. For example, researcher can compile a list of relevant terms and assign them a score, which can be negative or positive. The sum of all the scores in a text will provide a final evaluation. This method can be naïve, as it doesn't take into consideration the context of the words used. It is possible, however, to elevate the approach by adding expression and using more advanced processing techniques.

The goal of **machine learning** is to program computers to use example data or past experience to solve a given problem (Alpaydin, 2010). The approach uses algorithm that allow to train an automatic classifying program that, once fed with a text, will produce a tag (that can be positive, neutral or negative).

Types of sentiment analysis

The three most common types of sentiment analysis are:

- polarity (positive, negative or neutral)
- emotion (anger, happiness, sadness, etc.)
- intention (interested, not interested)

Polarity can be expanded, using a scale of 5 or even more.

Beyond text

People online do not use text alone to express opinion. A key feature of many social network and other **UGC** websites is the *like*. The main obstacle in using the like counts as a reliable information for engaging content is the fact that it expresses only a *positive* sentiment. Any analysis of engagement based on likes will result in the exclusion of polarizing or controversial topics. On this regard, Facebook, the biggest social network, implemented a feature to diversify the sentiment expressed by its like. This allows to run a simpler sentiment analysis based exclusively on the number of interactions. Evaluating the engagement rate of a content based on comments is also lacklustre: comments can be analyzed through sentiment analysis techniques, but this kind of analysis will exclude potentially embarrassing or delicate topics such as medical conditions or sex education. Reviews made by users online generally feature a polar evaluation system that can range between a binary evaluation (like or dislike, used by Youtube, Netflix and Steam, typically implemented to evaluate digital content on platforms that provide the content itself), a one to five evaluation (used by Google, Amazon and Tripadvisor, typically implemented to evaluate physical products or places), a one to ten evaluation (used by Metacritic or IMDb, typically used to evaluate movies and TV shows). These evaluation metrics provide an

instant polarity sentiment analysis that allows for an easy understanding and comparison by anyone.

3.3 Interactive Dashboard

Interactive dashboards are tools that enable viewing and interacting with complex underlying data using visualization methods such as charts, tables, maps or even text, all on a single display. They are used to monitor performances and support decision making by providing an easy and quick mean to access key performance indicator. They can be visualized and interacted with through different systems.

The earliest concepts of digital dashboard derived from the study of Decision Support Systems (DSS), which are a flexible, interactive, computerized approach intended to support administrators in their decision-making activities and which are capable of providing direct, personal support for complex managerial decisions (Turban, Fisher and Altman, 1988). In the late 1980s, making a dashboard was already technologically feasible, but the idea got mainly popularized through the introduction and success of Kaplan and Norton's Balance scorecard in 1996 (Eckerson, 2010). Since then, dashboards evolved for different use, scope, industries and technologies. As the sheer amount of available data increased, dashboard proliferated in many industries to better visualize them. One of the most agreed on definition of dashboard is a visual display of the most important information needed to achieve one or more objectives, consolidated and arranged on a single screen so the information can be monitored at a glance.

Dashboards are useful for summarizing data and alleviating information overload by utilizing robust visualization principles. They are used to improve performance, enhance

business operations, improve strategic decision making, ease the integration of institutional data with decision making practices, improve routine monitoring, track processes effectiveness, minimize data complexity, communicate organisations' values to diverse stakeholders, support data-driven decision making, and enable real-time monitoring of dynamically updating data.

Dashboards can be used to monitor single projects or department performances, can present more or less degrees of interactivity and holisticness, can use different kinds of data and being represented for different kinds of users and different devices. Many experts and authors proposed different classifications for dashboard types depending on these factors.

BIDashboard.org proposes a classification based on area of use:

- strategic – related to company's strategic foals
- tactical – measuring progression of ongoing projects
- operational – monitoring activities

However, it is difficult to find a classification on which different researchers agree on, as dashboards are relatively under-studied and the concept itself is very broad and flexible.

3.3.1 Dashboard development

When designing a dashboard on a technical level, the main concern should be the flexibility of the dashboard itself. The underneath system has to support the eventual modification of components, that can be gradually added or removed without affecting the correct working conditions of the other ones.

Dashboards are composed by three layers:

- data
- analysis
- presentation

Data will be collected from different sources and then organized. They come in different formats, the most popular ones being csv or json, and put through SQL systems in order to be properly collected and transformed, analyzed and researched, visualized and managed. Once data is collected in the SQL system, it is analyzed using instruments such Python or R algorithms. These analyses can be iterative as the data transformation need to be performed in service of the presentation.

As for the presentation, there are typically two dashboard GUI templates, including the one-page style and drill-down style (Farmanbar and Rong, 2020). Data can be presented through different means. The instrument must be able to display two things: if the indicators are in an acceptable “healthy” range and if there are progresses (Ragouzeos, Gandrup, Berrean, Li, Murphy, Trupin, Yazdany and Schmajuk, 2019), which is achieved by establishing setting standards and/or benchmarks and measuring data through time. Once the datasets are collected, some data manipulation techniques can be implemented for the user to aggregate and filter data, make trend analysis through time, enhance seasonality or other pattern, or different visualization tools. There are many adaptation techniques that can improve the experience of the user. Ideally, the user is able to go through the information provided by the dashboard in a matter of seconds or minutes. The most popular recent example of dashboard is the one kind developed to monitor the spread of the coronavirus (Covid-19) by universities or world health agencies. The emergence of many public different dashboard

presenting the same data shows how the approach on making a dashboard can widely change even if the data are the same. Researchers in the field agree that dashboard are relatively under-explored and under-researched. The lack of proper design guidelines has resulted in the development of many dash boards that involve many challenges. Most dashboards are poorly designed displays and most of them fail due to designers focusing on making them aesthetically appealing more than functionally effective (Alhamadi, 2020).

3.3.2 Dashboard's qualities, shortcomings and impact over the organization

This sub-chapter present a study on what has to be taken into consideration when building a management tool that fits and evolve existing processes and hierarchies in a given organization. The literature highlights the qualities as well as the shortcomings manifested by new and modern tools, regardless of the Power Bi implementation that this thesis later presents in its results.

Dashboards can be developed targeting not only decision makers and managers to support their work, but also stakeholders that do not ultimately decide, in order to improve the communication between them and the awareness of the latter (Ragouzeos, Gandrup, Berrean, Li, Murphy, Trupin, Yazdany and Schmajuk, 2019). A shared or even public dashboard may also add a new dimension that is the transparency of the decision-making process (Kitchin, Lauriault and McArdle, 2015).

One factor that is driving a renewed interest towards interactive dashboard creation is the increasingly accessibility to real time data collection. Dashboards represent the best solution to extract value from data sourced iteratively with high frequency (Farmanbar, 2020).

Setting up indicators allows for benchmarking. This means that dashboard have the ability to motivate policy changes deemed necessary to alter their relative rating, facilitating the process of learning by simply comparing to others. The benefits of benchmarking can also be extended to stakeholders, as they can hold decision makers accountable (Kitchin, Lauriault and McArdle, 2015). Indicator data can provide a rational, neutral, comprehensive, and commonsensical evidential basis for monitoring and evaluating the effectiveness of services and policy, to develop new interventions, and to learn and manage through measurement. Businesses and institutions can then be *steered and controlled* with strong leadership, solid coordination, powerful planning instruments and high-quality information, as they are made by a set of knowable and manageable systems that act in a rational and hierarchical way. However, interpreting institutions like this is somewhat limiting, as there are cases in which complex and interdependent systems exists, that influence each other in unpredictable ways. Governance may be multilevel and complex in nature, requiring consensus building and cooperation, and its performance system may not be reduced to targets and metrics. As such, indicators are but one element, albeit an important one given their factual, standardized, time-series nature, in a tangled web of processes that help guide strategic decision-making, acting as a common source of information and learning tool, rather than being an evaluation tool of operational programs or departments or the primary driver of policy and action (Kitchin, Lauriault and McArdle, 2015).

There's a critical misunderstanding relating dashboard initiatives, which is that it is implicitly assumed that benchmarks and indicators have no inherent politics or agenda and can be taken at face value (Kitchin, Lauriault and McArdle, 2015). Where people play a central role in data collection, it is assumed that a form of mechanical objectivity is deployed

that adheres rigorously to defined rules and uses a systematic method to produce detached, impartial and transparent data that are free from researcher bias and preferences, from local customs, culture, knowledge and context (Porter, 1995). This is of course not correct: data are generated as a product of many minds working within different situations, framed and shaped within contexts and structures, and this is true for their conception, measurement and use. They are generated withing protocols, measurement scales, standards and organizational processes will be negotiated and debated. *Data are never raw, but always already cooked* (Kitchin, Lauriault and McArdle, 2015).

There's a political aspect about how indicators are measured, communicated, visualized and used, the process tends to be led by stakeholders, and this also relates to how these data should be accessed (open vs centralized and technocratic approach). In all cases, a selection of indicators is framed by whether the desired data are readily available and the contingencies and costs of filling any gaps. Indicators can forward or defend a particular agenda, enact a compromise, or being data driven (selected because the data exists or omitted because the data is not available) (Kitchin, Lauriault and McArdle, 2015). Moreover, indicator, benchmarking and dashboard initiatives have a deep normative effect, shape governance, modify institutional behavior, condition workers, influence decision-making and shape spending patterns (Franceschini, Galetto and Maisano, 2007). Indicators can be used instrumentally (e.g., for problem solving and decision-making), conceptually (e.g., understand and interpret a situation), tactically (e.g., for delaying a strategy, substitute for action, deflect criticism), symbolically (e.g., to provide reassurance or place promotion), and politically (e.g., ammunition to support a particular position). It is also a way to manage at a distance and to alleviate decision makers from taking full responsibility for the actions they

may take or supports them to make hard decisions and avoid backlash. By becoming embedded in the thoughts, practices and institutions, a dashboard can influence the decision of users on many levels. The effects can range from more productivity as well as alienation, and as the *game is set*, decision maker can display unpredictable behavior to influence the measurements in their own self-interest (Kitchin, Lauriault and McArdle, 2015). This is especially true when dashboard expose managers to the scrutiny of others or the public.

Interactive dashboards present a possibly negative effect as their instrumental rationality manifests. When a dashboard is observed, from the perspective of the user, the scientific and practical instrumental knowledge, which is the knowledge that can be translated as hard facts, trumps over knowledge derived from practice, deliberation and experience. It undermines less systematic knowledge such as policy analysis, interviews, focus groups, surveys, etc. This can be considered a positive effect, but it also presents some shortcomings. It doesn't reflect the fact that different policies, activities and initiatives may go at different speeds and may have long-term goals that are not manifested by the benchmarking and indicators. It also assumes that there's a universal valid measure and method across places, resulting in benchmarking efforts between entities that have nothing in common regardless their history, trends, ambitions, access to resources, etc. Dashboards are an instrument that give the illusion that picturing the totality of a domain is possible, while in reality it's more of a black box in which data goes in, and that ends up occluding local forms of knowledge (Kitchin, Lauriault and McArdle, 2015).

There are some technical and methodological issues at play as well. One problem is that data are often taken at face value, however the underlying situation is more complex. Truth is:

data is generalized and approximated through its production, there's always some concern regarding their qualities (veracity, precision, fidelity, bias, consistency, reliability). When used as proxy, there should also be a concern regarding their ability to represent what they are meant to. This also depends on the technology used to gather data and the measurement regime. In order to perform benchmarking, entities belonging to the same entity class are typically divided into statistical groups (e.g., countries divided by regions), which end up being relatively arbitrary. The problem stands in the fact that changing the pattern changes the observations. As an example, the performances of the city of Dublin in benchmarks change drastically whether the City Council, the County or the Greater Dublin Area is taken into consideration. What these scaling and aggregation effects produce are a range of ecological fallacies, wherein false conclusions can be drawn due to how the data are collated, categorized and presented.

As for the data gathering itself, public data has proven to be really difficult to source. There are security and privacy related issues even for existing data, sensitive data that needs to be anonymized, and may not be open and re-usable. Implementing a system that employs public data presents challenges such as: different data resolution; large heterogeneity of data type; data quality heterogeneity; legal restriction; great variety of data structures. Most data are recurringly generated for this kind of projects, which in turn requires a systematic resolution of all of these issues.

Dashboards should be able to communicate if the monitored object registers numbers in acceptable parameters and if an improvement through time is being achieved, which is useful because in turn they can be used to improve quality through a data driven approach,

by establishing institutional quality standards that are shared between stakeholders and professionals while actively contrast the well-established cognitive bias of “illusory superiority”, whereby we tend to overestimate our abilities and achievements, something that regulation alone cannot damper (Ragouzeos, Gandrup, Berrean, Li, Murphy, Trupin, Yazdany and Schmajuk, 2019).

When used to benchmark different entities, dashboards allow to overcome the challenge of mismatch between the priorities faced on local level, as each entity (and its controller) has a personal agenda on top of working in a different context. This issue can be partially overcome with transparency and data sharing to pertinent stakeholders. In summary, dashboard initiatives can help set standards and improve relationships in a network of professional with different needs as they feel more involved and share data together.

3.3.3 Data Analytics and Big Data

Business Intelligence engineering requires data management. In this section of the literature review, it is presented the result of our research over data analytics and big data management, as they are the cornerstone of any reporting system.

Challenges in data management span three dimensions (Laney, 2001) called *three Vs*:

- *Volume*, refers to the magnitude of data, and it can be reported in terabytes or petabytes
- *Variety*, Variety refers to the structural heterogeneity in a dataset. Technological advances allow firms to use various types of structured, semi-structured, and

unstructured data. Structured data, which constitutes only 5% of all existing data (Cukier, 2010), refers to the tabular data found in spreadsheets or relational databases. Text, images, audio, and video are examples of unstructured data, which sometimes lack the structural organization required by machines for analysis. Spanning a continuum between fully structured and unstructured data, the format of semi-structured data does not conform to strict standards. Extensible Markup Language (XML), a textual language for exchanging data on the Web, is a typical example of semi-structured data. XML documents contain user-defined data tags which make them machine-readable (Gandomi and Haider, 2015).

- *Velocity*, refers to the rate at which data are generated and the speed at which it should be analysed and acted upon (Gandomi and Haider, 2015). The growth of digital devices such as smartphones and sensors led to a tremendous increase in data creation rate. User data are collected and elaborated in real time to create personalized user experiences.

The concept of Big Data has unclear origins and it evolved through time. It is naturally related to *size*. However, definitions of big data usually use all the *three Vs* as a common framework (Gandomi and Haider, 2015). For example, Gartner, Inc. defines big data as:

“Big data is high-volume, high-velocity and high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making” (Gartner IT Glossary, n.d.).

Additional *Vs* have been proposed through time, the most notable are:

- *Value*, proposed by Oracle, it expresses the value density of data, which is proportional to the analysis data went through;
- *Veracity*, proposed by IBM, represents the unreliability inherent of some sources of data
- *Variability* (and complexity), proposed by SAS, these two dimensions refer to the fact that data velocity isn't consistent (variability) and data come from many different sources (complexity). The resulting challenge is to connect, match, clean and transform all of these data received from different sources at different and inconsistent speed

Big data are useless if they are not analyzed. The overall process of extracting insights from big data can be broken down into five stages (Labrinidis and Jagadish, 2012) shown in Figure 6.

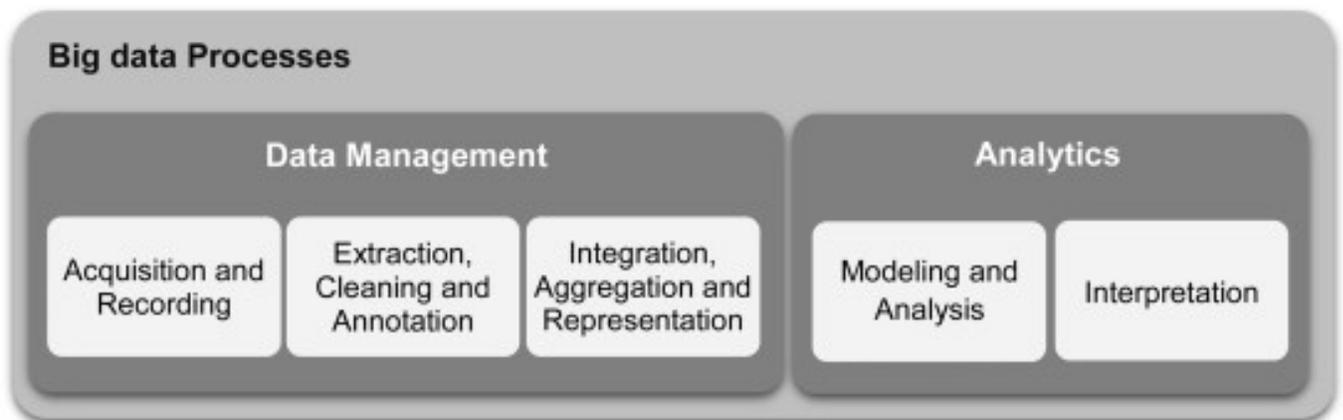


Figure 6 - Extracting insights from big data (Labrinidis and Jagadish, 2012)

3.3.4 Database and the Entity-Relationship Model

A database is a collection of information organized so that data can be easily stored, managed, updated and retrieved. To view data logically, the most widely used model is the Entity-Relationship model, published in 1976 from computer scientist Peter Pin-Shan Chen. The idea is that entities and relationships can describe the real world in the more natural way, while providing a high degree of data independence. What follow is a brief explanation of how the model work and how it is used practically to create a database, and it is completely based on the paper *The Entity-Relationship Model-Toward a Unified View of Data*.

An entity is a *thing* which can be distinctly identified. A specific person, company, or event is an example of an entity. A relationship is an association among entities. For instance, *father-son* is a relationship between two *person* entities.

Entities are classified into different entity sets such as EMPLOYEE, PROJECT, and DEPARTMENT. A relationship set is the mathematical relation among n entities from an entity set. A relationship MARRIGE would be between two entities from the same entity set PERSON. Relationships have roles, so in the case of MARRIGE, there will the the entity *husband* and the entity *wife*. The information about an entity or a relationship is obtained by observation or measurement, and is expressed by a set of attribute-value pairs. An attribute can be formally defined as a function which maps from an entity set or a relationship set into a value set. For instance, the entity set PERSON has an attribute AGE that maps into the value set NUMBER-OF-YEARS. Sometimes, the attribute and the value overlap, such as in the case of an attribute called ID-NUMBER that maps into the value set ID-NUMBER. Relations also have attributes. In practice, an entity set can be represented as:

	Attribute 1	Attribute 2	...
Entity 1	Value 11	Value 12	...
Entity 2	Value 21	Value2 2	...
...

Figure 7 - Example of entity set

A relationship can be represented as:

Entity Set 1	Entity Set 2	Relationship
Entity 1 (set 1)	Entity 1 (set 2)	Value 11 (set 12)
...

Figure 8 - Structure of relationship between entities

Primary keys are attributes that allows to identify entities inside an entity set. Basically, an entity key is a group of attributes such that the mapping from the entity set to the corresponding group of value sets is one-to-one. If it's not possible to find such one-to-one mapping on available data, or if simplicity in identifying entities is desired, an artificial attribute and a value set is defined so that such mapping is possible. This translates in semantically meaningful keys, such as *codes* that identify a single entity. Relationships also have primary keys. Since a relationship is identified by the involved entities, the primary key of a relationship can be represented by the primary keys of the involved entities. There are different kinds of relationships described as *one to many* or *one to one* or *many to many*. For instance, an employer will have many employees, but an employee will have a single employer, so their relationship is *one to many*.

The steps to set up a database are the followings:

1. identify the entity and relationship sets of interest;;
2. identify the semantic information in the relationship sets such as whether a specific relationship set is an 1 to n mapping;
3. define the value set and attributes;
4. organize data and decide primary keys.

Data integrity

Data integrity refers to the accuracy and consistency (validity) of data over its lifecycle (Brook, 2020). The integrity of data can be compromised by different means. The end goal is to prevent it, so that data remain intact and unaltered when transferred or replicated. Data integrity ensures recoverability and searchability, traceability to origin and connectivity, and for such reason is very important. On top of that, it improves performances, reusability and reliability, which are all qualities required for systems on which decision makers may rely on to evaluate and take decisions. Some sources of data alteration can be:

- human error
- transfer error
- bugs, viruses and malware, hacking or other cyber threats
- compromised hardware
- physical alteration to the devices

The entity-relationship model tackles the problem of data integrity by setting ups constraints:

1. Constraints on *allowable values* for a value set
2. Constraints for *permitted values* to an attribute
3. Constraint between existing value sets or specific values – (example, a column representing an attribute can be the sum of two other attributes)

3.4 The research gap

From the literature analysis performed we observe two main relevant trends:

- the need of a framework that matches the needs of museums regarding performance monitoring, so the need of tailored framework which fits the purpose of this different kind of industry;
- the growing role of social media in modern society, and in particular, on every kind of business;
- the evident and, in the case of cultural institutions, unexploited potential of data visualization tools.

The study of the literature revealed some gaps:

- most of the work relating the monitoring of performances in cultural institutions worldwide do not take into consideration the relevance that social media data have in estimating and expressing the reputation and in turn the value that museums are able to extract from the cultural heritage they manage. These works tend to be outdated and produced before social media platform and the general online presence of an institution of any kind became such an important part of every business.

- many studies dedicated to dashboards have been conducted, some using a comprehensive approach, some focusing on a specific scenario, however it has never been done satisfyingly enough for a cultural institution.

To date only few studies focused their attention on the introduction of social media and interactive dashboards for cultural institutions performances monitoring. Moreover, the introduction of the new and different perspective given by the incumbency of social media is not well exploited and it is, often, undervalued. When using social media for business, it is not only important to set objectives and define a strategic plan, but in order to reach the company or cultural institution goals, an adequate monitoring is key. The world of social media is fast-paced and in continuous evolution so each measurement system must be tailored based on the specific requirements or goal of the firm, company or cultural institution.

Additionally, the studies that revolve around cultural institutions typically focus on how to set up quality standards for the institution as a single independent entity rather than facing the problem with an holistic approach, completely disregarding any benefit that may come from shared data and purpose between institutions.

In conclusion, the literature alone cannot provide an up-to-date performance measurement system for museums and regulators that is set up to take advantage of the possibilities offered by dashboards in a satisfying extent, so further research is required.

4 Conceptual Framework

While the idea of finding a way to measure a museum's performance is somewhat controversial - how does one assign monetary value to artistic and cultural offerings? - it is also long overdue, largely because comprehensive and reliable methodologies are in such short supply. Financial ratios and shareholder value may be good yardsticks for business, but evaluating a cultural institution is more complex. Any measurement of performance must consider artistic and intellectual contributions, not just financial strength (Egloff and Zorloni, 2012).

4.1 An overview of Museums organization in Italy

In order to give the reader a background of the landscape and scope of our work, the following sub-chapter summarizes the museums governance in Italy. The highest level of governance is the General Direction of Museums, which is part of the Ministry. The General Direction is split between two parts:

- General Direction Service I – administration
- General Direction Service II – Sistema museale nazionale (National Museum System)

Together, they work to promote research and diffusion of knowledge and culture, to enhance the national cultural heritage and make it accessible and competitive internationally. The main drivers are accessibility, retention, local networks of institutions and preservation.

The General Direction Service I is dedicated to administration and control. It's dedicated to activities connected to collections, accounting and control, legal function and human resource. It is also responsible for connection and partnerships with other prestigious

institutions in Italy and abroad, both for collection leasing and expositions, as well as purchasing and any cultural or scientific interest that may stem from these connections.

The General Direction Service II coordinates the activities of more autonomous institutions. Its work is more focused on the local territories, in which they promote collaboration between public and private institutions. Service II is in charge of quality control, establishing the criteria and elaborating the evaluations through qualitative and quantitative parameters. It also dedicates to ensure accessibility and fruition through initiatives and programs, to increment the value offered to tourists.

The General Direction Service II accomplish its goals by managing the “Sistema museale nazionale” (National museum system), which is a network composed by over 5000 museums and places of cultural relevance, connected together to create synergies and improve fruition, accessibility and sustainability of the national cultural heritage. The network may include public as well as private or independent institutions.

Public museums in Italy are divided in two groups:

- autonomous
- public

Autonomous museums are 32 institutions or cultural locations that earned special autonomy thanks to their particular relevance. The General Direction directs, guides, coordinate and control these museums, while they have autonomy on financing, accounting, organization and research matter.

There are 400 public museums that are managed from the General Direction. Since 2014, public museums refer to their regional directions called *Polo museale*, and organ in between the local institutions and the General Direction. There are a total of 17 *poli*, one per region, excluding Sicilia, Trentino Alto Adige and Valle d'Aosta because of their special autonomy.

4.2 The CCO Matrix - a new strategic model for performance monitoring

The following sub-chapter features the strategic model proposed for the monitoring of museums performances, which takes as inputs:

- Literature Analysis
- Conceptual Framework analysis
- Analysis of Italian Museums strategy

The output is the CCO Matrix - Communication, Collection and Organization Matrix - that supports museums performances monitoring. The insights gleaned through this approach can help funders and administrators make more rational, productive decisions about how to allocate scarce resources. The matrix is articulated in three main sections:

- Communication
- Collection
- Organization

These three defined sections reflect the structure of the Italian Ministerial Decree N°113 (21/02/2018 «Adozione dei livelli minimi uniformi di qualità per i musei e i luoghi della cultura di appartenenza pubblica e attivazione del Sistema museale nazionale»). The

sections are the output of an analysis made by the Italian Ministry of Culture and inspired in international best practices.

Each of these sections has been analyzed across four perspectives, inspired in the BSC model introduced in the previous portion of this document:

- Visitors and reputation
- Internal processes
- Innovation and learning
- Economic resources

The most innovative part of the model resides in the introduction of the reputation perspective.

According to an analysis made on a sample comprises 38 museum executives and three museum association executives from Washington, DC, New York, and London a key success measure in the eyes of six executives is the museum's ability to build a good reputation (Zorloni, 2010). Moreover, a theme that emerged from the interviews was the idea that reputation was considered by museum executives as highly important as a means to achieve the museum's mission.

But how reputation was measured?

- two museums stated that they periodically engage in formal surveying in order to assess the visitor's impression of the museum's programs (Zorloni, 2010);
- three museums stated that they are engaging in formal brand audits to assess the museum's reputation (Zorloni, 2010);

- other respondents, working in marketing departments, cited the ability to get respect among their peers and media coverage for the museum and its programs as indirect indicators of their influence within the field (Zorloni, 2010);
- for many others, measurement of reputation is more concerned with informal information, largely driven by word of mouth within the museum community or press coverage of the museum and its programs (Zorloni, 2010).

These responses seem to suggest the hypothesis that, in the absence of widely accepted measures of performance, reputation can easily become a signal of quality and therefore a surrogate for museum effectiveness (Zorloni, 2010).

The data were collected over 7 months (from April to July 2008 in London; November 2008 in Washington, DC, and from February to April 2009 in New York) through face-to-face interviews with executives of 13 visual art museums and three museum associations. The research panel was represented by the following organizations listed in the table below (Zorloni, 2010).

	Name of the organization	City
1	Solomon R. Guggenheim Foundation	New York
2	Museum of Arts and Design	New York
3	International Center of Photography	New York
4	Whitney Museum	New York
5	Jewish Museum	New York
6	The Metropolitan Museum of Art	New York
7	El Museo del Barrio	New York
8	Smithsonian Institution	Washington, DC
9	National Museum of the American Indian	Washington, DC
10	National Museum of African American History & Culture	Washington, DC
11	Phillips Collection	Washington, DC
12	Freer Sackler Galleries	Washington, DC
13	Tate	London
14	Association of Art Museum Directors	New York
15	American Association for State and Local History	Nashville, TN
16	American Association of Museums	Washington, DC

Figure 9 - The research pannel

This model introduces a new way to measure museums reputation: trough indicators build with data collected from social media data.

To get to the strategic matrix the first step has to be the construction of a strategic map. The International Council of Museums (2007) defines a museum as *a non-profit- making, permanent institution in the service of society and of its development, and open to the public, which acquires, conserves, researches, communicates and exhibits, for purposes of study, education and enjoyment, material evidence of people and their environment.*

The museum strategy must be aligned with its mission, so the following strategic map could be tailored and slightly modified in order to meet specific needs. Nevertheless, as the fulfilment of a strategy is not measurable by itself, the general strategy of a museum has been broken down into sub-goals, which are meant to be more easily measurable keeping in mind

that understanding performance means having a detailed knowledge of all stakeholders and their needs.

In order to adapt the balance scorecard framework described in chapter 3.1 for this specific industry where the product is a service and the performances to be measured are different, we assume that a museum generally tries to use the resources at hand in order to optimize the fulfillment of the museum mission. Under this assumption, an output orientation seems a natural choice.

The following paragraph describe how the sections of the balance scorecard have been adapted to match museums needs and goals:

- How do customers see us? (*ex customer perspective*), in the new model this section will be named *visitors and reputation* since in the case for museums visitors are the main customers and a key success measure for a museum is its ability to build a good reputation. Reputation is considered by museum executives as highly important as a means to achieve the museum's mission.
- What must we excel at? (*ex internal business perspective*), in the new model this section will be named *internal processes*, the model must ensure that the museum processes and efforts are aligned with stakeholders' expectations, it is important to guarantee that the internal processes support the expected external perspective, rather than focusing on the internal perspective per se
- Can we continue to improve and create value? (*innovation and learning*), this section will remain the same, since as for a for-profit business, continuous innovation and learning perspective fits cultural institutions mission and goals

- How do we look to shareholders? (ex *financial perspective*), in the new model this section will be named *economic resources*, a cultural institution must use its resources efficiently and effectively to achieve the trust of the public, but the word “finance” reminds of traditional financial indicators from financial statements (profit & loss account, balance sheet, and cash flow statement) which are not always available for every museum and, moreover, are not the only way to assess efficient financial management.

The analysis resulted in the definition of precise strategy direction for every section and across every perspective:

- Communication
 - Improve potential visitors' awareness (*Visitors and Reputation perspective*)
 - Broaden marketing and communication efforts (*Economic Resources perspective*)
 - Promote educational initiatives (*Innovation and Learning perspective*)
 - Enhance internal processes that support communication efforts (*Internal Processes perspective*)
- Collection
 - Increase visitors' interest through collection exploitation (*Visitors and Reputation perspective*)
 - Increase investments in collection improvement and enlargement (*Economic Resources perspective*)
 - Improve collection accessibility (*Innovation and Learning perspective*)

- Favor collection enjoyment (*Internal Processes perspective*)
- Organization
 - Better understand museum visitors' characteristics to march their needs (*Visitors and Reputation perspective*)
 - Track the impact of different revenues streams (*Economic Resources perspective*)
 - Increase the presence of qualified personnel (*Innovation and Learning perspective*)
 - Improve customer experience (*Internal Processes perspective*)
 - Promote environmental sustainability (*Internal Processes perspective*)

Social media indicators are used to track the success of the social media platform. This is crucial, especially for an institution such as a museum, that deals with the public and thrive in creating local and international communities around it. Social media indicators can measure the success of single social media activities with great precision, as well as giving an overall measure of performance, thanks to data that can be easily tied up to specific actions while also being perfectly capable of being aggregated. Each advancement in the metric can be traced down to a very specific time, sometimes down to the seconds, allowing to control trends overtime really easily without requiring extra studies or manual computation.

An overview of the CCO Matrix is presented on the table below.

	COMMUNICATION	COLLECTION	ORGANIZATION
VISITORS AND REPUTATION	<ol style="list-style-type: none"> 1. Visitors Proactivi 2. Net Promoter Score 3. External Exposure Indicator 4. Followers Growth 5. Engagement rate 	<ol style="list-style-type: none"> 1. Temporal exhibitions impact on visitors 2. Special events occurrence growth 	<ol style="list-style-type: none"> 1. Visitors classification - a collection of indicators
ECONOMIC RESOURCES	<ol style="list-style-type: none"> 6. Marketing and Communication to total budget 7. Social Media management budget to marketing and communication 8. Donors contribution rate 9. Communication impact on tickets selling 	<ol style="list-style-type: none"> 3. Investment on collection maintenance 	<ol style="list-style-type: none"> 2. Income sources - a collection of indicators
INNOVATION AND LEARNING	<ol style="list-style-type: none"> 10. Learning activities growth 	<ol style="list-style-type: none"> 4. Accessibilities to minorities index 5. Collection digitalization 	<ol style="list-style-type: none"> 3. Personnel qualification 4. Personnel training hours
INTERNAL PROCESSES	<ol style="list-style-type: none"> 11. Content creation growth rate 	<ol style="list-style-type: none"> 6. Support services index 	<ol style="list-style-type: none"> 5. Hospitality quality 6. Ticket digitalization 7. Environmental sustainability

Figure 10 - CCO Matrix Strategic Map overview

For every section of the strategic map, precise indicators linked with the previously mentioned strategic goals have been identified. In the following paragraph a description of each indicator will follow encompassing:

- The name of the indicator
- How the indicator is calculated
- Indicator position in the matrix
- Indicator meaning and linkage to the strategic goals

- Additional Notes (e.g., possible adjustments)

4.2.1 Communication

	COMMUNICATION	COLLECTION	ORGANIZATION
VISITORS AND REPUTATION	1. Visitors Proactivity 2. Net Promoter Score 3. External Exposure Indicator 4. Followers Growth 5. Engagement rate	1. Temporal exhibitions impact on visitors 2. Special events occurrence growth	1. Visitors classification - a collection of indicators
ECONOMIC RESOURCES	6. Marketing and Communication to total budget 7. Social Media management budget to marketing and communication 8. Donors contribution rate 9. Communication impact on tickets selling	3. Investment on collection maintenance	2. Income sources - a collection of indicators
INNOVATION AND LEARNING	10. Learning activities growth	4. Accessibilities to minorities index 5. Collection digitalization	3. Personnel qualification 4. Personnel training hours
INTERNAL PROCESSES	11. Content creation growth rate	6. Support services index	5. Hospitality quality 6. Ticket digitalization 7. Environmental sustainability

Figure 11 - CCO Matrix: Communication Section indicators

1. Visitors Proactivity

Name of the indicator: Visitors Proactivity Rate

Calculation

$$\frac{\text{Tripadvisor reviews} + \text{Facebook reviews} + \text{Google Maps Reviews}}{\text{Number of visitors}}$$

Matrix positioning:

Communication - Visitors and reputation: The indicator shows the visitors proactivity in social media communication channels that allow users to write reviews. It links social media data to museums daily operations data. It is a proxy of the communication effort since it shows how many visitors are prone to leave a review, how many visitors are aware of the presence of the museum in those communication channels and take time to generate a review.

Link to strategy:

Visitors proactivity is linked to the strategic goal *“Improve potential visitors’ awareness”*, the more the reviews the more the soundness of the museum social media impact. Moreover, if the museum takes measures to encourage visitors to leave reviews, or undertakes remarkable actions of any kind, the web and the visitors will react.

Additional notes:

Regarding the numerator, TripAdvisor, Facebook and Google Maps reviews are taken in into account in this analysis, since (as of December 2020) these two platforms are the most used, but the platforms selected can be changed and updated to best fit the museum strategy in UCG platforms.

2. Net promoter score – a collection of three indicators

Name of the indicator: Net Promoter Score (promoters’ rate, passives rate, detractors’ rate)

Calculation

Promoters rate – calculated as

$$\frac{\# \text{ positive reviews on Facebook + Google Maps}}{\text{Total \# of reviews Facebook + Google Maps}}$$

Passives rate – calculated as

$$\frac{\# \text{ neutral reviews on Facebook + Google Maps}}{\text{Total \# of reviews Facebook + Google Maps}}$$

Detractors rate – calculated as

$$\frac{\# \text{ negative reviews on Facebook + Google Maps}}{\text{Total \# of reviews Facebook + Google Maps}}$$

Matrix positioning:

Communication - Visitors and reputation: The indicator is a proxy the museum reputation on social media, it must be analyzed alongside with the indicator “Visitors Proactivity Rate”. It answers to the following question: “Of the visitors writing a review how many are promoters, how many are passives and how many are detractors?”

Link to strategy:

The Net Promoter Score is linked to the strategic goal “*Improve potential visitors’ awareness*”, positive reviews (reviews by promoters) boost the museum reputation, increasing potential visitors’ curiosity and desire to visit the museum. On the other hand, passives and detractors’ reviews, represent valuable feedback, a red flag that something should be improved. A deep dive into those reviews is highly recommended.

Additional Notes:

Regarding the numerator, Facebook and Google Maps reviews are taken into account in this analysis, since (as of December 2020) these two platforms are the most used, but the platforms selected can be changed and updated to best fit the museum strategy in UCG platforms.

3. External exposure indicator

Name of the indicator: External exposure indicator

Calculation

$$\frac{\text{Number of hashtags} + \text{Number of GeoTags} + \text{Number of tags across platforms}}{\text{Number of visitors}}$$

Matrix positioning:

Communication - Visitors and reputation, external posts (those that are not generated by the museum account) are also of high importance when measuring the museum reputation on social media. Tags, geotags (location), and hashtags are a way through which visitors interact with the museum.

Link to strategy:

The external exposure indicator is linked to the strategic goal “*Improve potential visitors’ awareness*”, a powerful side effect of social media resides in the indirect exposure of the accounts: external users (not necessarily active followers of the museums accounts) can

indirectly contribute to build the reputation of museums in Social Media platforms. For instance, visitors can upload a picture taken in the museum, in which they include a GeoTag (location), or relevant hashtags exposing the museum to all of their followers and contributing in building the museum reputation. Tracking these external posts allows to measure the silent and uncontrolled echo of the museums on social media platforms.

Additional Notes:

Regarding the numerator, tags, geotags and hashtags must be gathered from all the social media platforms in which the museum is present. It is heavily reliant on a complete dataset that tracks tags across different platforms, considerably increasing the effort in data gathering.

4. Followers growth

Name of the indicator: Followers and growth rate

Calculation

$$\frac{\text{Actual Number of followers in a selected social media platform}}{\text{Number of followers in a selected social media platform in the previous period}}$$

Matrix positioning:

Communication - Visitors and reputation, the indicator measures the growth of the followers in each social media platform, so it is a proxy of the museum's reputation. New followers today can be future visitors tomorrow, through this metric one can clearly observe if the effort in social media management is translated into tangible results.

Link to strategy:

The followers growth rate is linked to the strategic goal *“Improve potential visitors’ awareness”*, the more followers, the more potential visitors aware of the museum initiatives, facilities and collection, each follower can then unlock the Word of Mouth mechanism and increase the awareness of new potential visitors. To track this measure is key to observe progress.

Additional Notes:

The indicator must be calculated for each social media in which the museum is active.

5. Engagement rate

Name of the indicator: Engagement rate

Calculation

$$\frac{\text{Interactions (wieghted likes, comments, tweets) acorss Instagram, Facebook, Twitter}}{\text{Number of followers}}$$

Matrix positioning:

Communication-Visitors and reputation, the engagement rate measures the followers interactions against the number of followers, if greater than one it means that on average every follower interacts with the museum account on social media. This indicator is a proxy of museums reputation on social media.

Link to strategy:

The engagement rate indicator is linked to the strategic goal “*Improve potential visitors’ awareness*”, the more the followers are engaged, the more they interact with the museum. More interactions, shares, comments and likes mean a grater virality of the museums social media accounts, which in the end improve potential visitors awareness about the museums.

Additional Notes:

When performing the addition of the so-called *interactions* (likes, comments, etc.) at the numerator of the indicator, it is possible to improve the meaningfulness of the information by establishing an informed method to weight the different type of interactions, in order to make them summable. For instance, a spontaneous like in Facebook is less an interaction respect to a comment written by a user and again it is less an interaction respect to the action of sharing a post, as the level of commitment and involvement of the user changes and this aspect cannot be ignored. As a reference, in the model presented in *chapter “Measuring the degree of corporate user-generated content”: a model*, Thomas Aichner and Frank Jacob define a conversion rate is defined as follows: *positive comments* in every social media show a higher degree of identification so are weighted 5 times the weight of *likes*; *shares* or *retweets* are weighted 10 times the weight of *likes* because they have an important multiplier effect.

6. Marketing and Communication to total budget

Name of the indicator: Marketing Communication to total budget

Calculation

$$\frac{\text{Budget allocated to marketing and communication}}{\text{Total Budget}}$$

Matrix positioning:

Communication – Economic Resources, this indicator measures the communication efforts in terms of economic resources invested.

Link to strategy:

The marketing communication to total budget indicator is linked to the strategic goal “*Broaden marketing and communication efforts*” and it must be tracked in order to understand if the “visitors and reputation” perspective reflects the efforts made from the “economic resources” perspective.

7. Social Media management budget to marketing and communication budget

Name of the indicator: Social Media management budget to marketing and communication budget

Calculation

$$\frac{\text{Budget allocated to social media management}}{\text{Budget allocated to marketing and communication}}$$

Matrix positioning:

Communication – Economic Resources, this indicator measures which percentage of the investment in marketing and communication is allocated to social media management.

Link to strategy:

The Social Media management budget to marketing and communication indicator is linked to the strategic goal “*Broaden marketing and communication efforts*” keeping an eye of social media, since they are becoming the most important communication channel with customers (visitors). It must be tracked in order to understand if the “visitors and reputation” perspective reflects the efforts made from the “economic resources” perspective in terms of social media management efforts.

Additional Notes:

If a Social Media Manager is not well-defined but another figure carries out the tasks related to social media management, the % of the timespan dedicated to those activities multiplied by the wage of the employee must be included in the numerator.

8. Donors contribution rate

Name of the indicator: Donors contribution rate

Calculation

$$\frac{\text{Income from donations or sponsorships}}{\text{Total Income}}$$

Matrix positioning:

Communication – Economic Resources, this indicator belongs to the communication area and economic resources perspective since a better communication strategy increases awareness between potential donors and their propensity to interact with the museum.

Link to strategy:

The donors contribution rate indicator is linked to the strategic goal “*Increase museum sponsors awareness*”. We assume that a solid communication campaign influences positively donations flow.

9. Communication impact on tickets selling

Name of the indicator: Communication impact on tickets selling

Calculation

$$\frac{\text{Revenues from tickets sold through social media}}{\text{Revenues from tickets sold online}}$$

Matrix positioning:

Communication – Economic Resources, the indicator measures the role played by social media on tickets selling, it links the communication area to the economic source perspective.

Link to strategy:

The marketing communication impact on tickets selling is linked to the strategic goal “*Increase museum sponsors awareness*”. This indicator allows to track if visitors are using the museum social media accounts, as a matter of fact, if a visitor reaches the ticket selling website through social media, it means that the communication efforts are working. The indicator it is a great proxy to establish effectiveness of the overall social media page and the communication efforts.

Additional Notes:

The revenues from tickets sold through social media can be measured by using the so called “referral link”, also known as an affiliate link. The referral link contains the affiliate ID that tracks how visitors land on a webpage. It is part of the basic website analytics tools, more than a social media analytics tool. It contains, however, an important information about how many users visiting the museum’s social media page are prompted to visit the museum’s website and, how many of them makes a purchase.

10. Learning activities growth

Name of the indicator: learning activities growth

Calculation

$$\frac{\text{Projects, labs, seminars events or guided tours organized for schools high schools or universities}}{\text{Projects, labs, seminars events or guided tours organized for schools high schools or universities in the previous period } t}$$

Matrix positioning:

Communication – Innovation and learning, the indicator allows to measure the effect of the communication effort from an innovation and learning perspective, in particular tracks the growth of the museum relationship with the local community and its cultural institutions.

Link to strategy:

The learning activities growth index is linked to the strategic goal “*promote educational initiatives*” and it must be tracked in order to monitor its trend and take action when it decreases.

Additional Notes:

Regarding the denominator, the “period t” used as benchmark must be chosen by each museum and depends on data availability.

11. Content creation growth rate

Name of the indicator: Content creation growth rate

Calculation

$$\frac{\text{Content created for Social Media (number of posts + number of stories + number of tweets)}}{\text{Content created for Social Media (number of posts + number of stories + number of tweets) in the previous period } t}$$

Matrix positioning:

Communication – Internal processes, the indicator shows the internal processes must be organized in order to support social media activity by generating inputs for posts, stories or tweets: a “sharing” culture is key.

Link to strategy:

The content creation growth index is linked to the strategic goal “*Enhance internal processes that support communication effort*” and it must be tracked in order to monitor its trend, foster its growth and take action when it decreases.

Additional Notes:

For this analysis we consider Instagram and Facebook posts and Twitter tweets, but the indicator can be tailored in order to match the social media strategy of the museum and the

platforms used. Regarding the denominator, the “period t” used as benchmark must be chosen by each museum and depends on data availability.

4.2.2 Collection

	COMMUNICATION	COLLECTION	ORGANIZATION
VISITORS AND REPUTATION	<ul style="list-style-type: none"> 1. Visitors Proactivity 2. Net Promoter Score 3. External Exposure Indicator 4. Followers Growth 5. Engagement rate 	<ul style="list-style-type: none"> 1. Temporal exhibitions impact on visitors 2. Special events occurrence growth 	<ul style="list-style-type: none"> 1. Visitors classification - a collection of indicators
ECONOMIC RESOURCES	<ul style="list-style-type: none"> 6. Marketing and Communication to total budget 7. Social Media management budget to marketing and communication 8. Donors contribution rate 9. Communication impact on tickets selling 	<ul style="list-style-type: none"> 3. Investment on collection maintenance 	<ul style="list-style-type: none"> 2. Income sources - a collection of indicators
INNOVATION AND LEARNING	<ul style="list-style-type: none"> 10. Learning activities growth 	<ul style="list-style-type: none"> 4. Accessibilities to minorities index 5. Collection digitalization 	<ul style="list-style-type: none"> 3. Personnel qualification 4. Personnel training hours
INTERNAL PROCESSES	<ul style="list-style-type: none"> 11. Content creation growth rate 	<ul style="list-style-type: none"> 6. Support services index 	<ul style="list-style-type: none"> 5. Hospitality quality 6. Ticket digitalization 7. Environmental sustainability

Figure 12 - CCO Matrix: Collection section indicators

1. Temporal exhibitions impact on visitors

Name of the indicator: temporal exhibitions impact on visitors

Calculation

$$\frac{\text{Average number of visitors during temporal exhibitions}}{\text{Average Number of visitors}}$$

Matrix positioning:

Collection - Visitors and reputation the indicator highlights the impact of temporal exhibitions on the number of visitors, even if the focus is on temporal exhibitions, so on the exhibition of pieces of collection which are not owned by the museum but temporarily borrowed from other cultural institutions, the increase of visitors has an impact on the museum reputation. The facilities in which the collection is mounted belong to the museum and so it is the service visitors receive. The echo that a temporal exhibition has on the number of visitors must be monitored in order to better meet visitors needs and to stimulate their curiosity towards the museum as a whole, not only towards the temporal exhibitions that the museum holds.

Link to strategy:

The temporal exhibitions impact on visitors is linked to the strategic goal "*Increase visitors' interest through collection exploitation*" increasing the number of temporal exhibitions, increases the awareness about the museum between potential visitors. Different temporal exhibitions could help the museum to meet the different needs customers have, and to attract them to the museum boosting their interest toward the cultural institution.

Additional Notes:

This indicator could be calculated both taking into account a certain period of time (including, eventually several temporal exhibitions) or it could be calculated for every temporal exhibition that the museum holds.

2. Special events occurrence growth

Name of the indicator: Special events occurrence growth

Calculation

$$\frac{\text{Number of special events held in the museum in the time horizon selected}}{\text{Average number of special events held during the year}}$$

Matrix positioning:

Collection - Visitors and reputation the indicator measures the percentage increase or decrease of special events occurrence in a certain period of time. It is important to monitor this indicator because its trend could directly influence other indicators related to visitors flow into the museum. Moreover, the number and quality of events held in a museum influence its reputation, visitors attending special events could be future museum visitors. Participants personal experience could lead to a positive or negative review of the museum facilities and services both in Social Media and in the most traditional Word or Mouth.

Link to strategy:

The special event occurrence growth is linked to the strategic goal “*Increase visitors’ interest through collection exploitation*”, museums are no longer simply repositories for artifacts, but vibrant cultural institutions with unique programming that changes all the time. Holding special events such as conferences, laboratories, parties, congresses, film screenings etc. is an indirect way to exploit the museum collection, it boosts potential customers curiosity and

interest towards the museum in a different and new way having an impact on the museum perception and reputation.

Additional Notes:

Special events include every event a museum hold that it is not considered a recurring activity.

3. Investment in collection maintenance

Name of the indicator: Investment in collection maintenance

Calculation

$$\frac{\text{Investment in restoration}}{\text{Total annual budget}}$$

Matrix positioning:

Collection- Economic Resources, the indicator shows the budget percentage allocated in restoration, it is important to keep track of the indicator in order to better understand the impact of restoration on the total annual budget and better allocate economic resources.

Link to strategy:

The investment in collection maintenance s linked to the strategic goal *“Increase investments in collection improvement”* the museum collection has to be well preserved, not only the museum facilities and services need special investments, but the most important asset of a

museum remains its collection: considering a long-term horizon it's the collection well preservation is the key value driver.

4. Accessibility to minorities index

Name of the indicator: Accessibility to minorities index

Calculation

$$\frac{\text{Number of best practices on accessibility to minorities services achieved}}{\text{Total number of best practices on accessibility to minorities services}}$$

Best practices:

- Well-defined paths designed for wheelchairs
- Braille captions
- Tactile maps
- Contrasts of luminance
- Exposition at wheelchair height
- Presence of trained staff
- Presence of ramps

Measure these aspects with a binary variable that takes either value 0 (denoting the absence of the service) or 1 (denoting the presence of the service).

Matrix positioning:

Collection- Innovation and learning, the indicator belongs to the innovation and learning perspective since it measures the level of innovation in terms of best practices put in action by the museum in order to support collection accessibility to minorities. Best practices could be replaced by new ones with the development of new technologies and so, to keep track of the best practices implemented by the museum against all the best practices that could be implemented is a very important activity.

Link to strategy:

The accessibility to minorities index is linked to the strategic goal *“Improve collection accessibility”*. To guarantee an appropriate accessibility to the collection to all kind of customers is one of the main goals of cultural institutions, that is why an indicator which tracks the best practices implemented vs best practices available and implemented by other institutions is important in order to understand where the cultural institution stands and what can be improved.

Additional Notes:

The best practices previously listed derive from the literature analysis and are inspired in particular by the best practices defined by the Italian museums. Nevertheless, as previously stated, best practices can change, and so can the visitors’ needs, so, the list of best practices must be updated in order to compute a sound indicator aligned with technology evolution.

5. Collection digitalization

Name of the indicator: collection digitalization

Calculation:

$$\frac{\text{Number of collection pieces inventoried and digitalized}}{\text{Collection size}}$$

or

$$\frac{\text{Number of digitalized exhibition rooms}}{\text{Total number of exhibition rooms in the museum}}$$

Matrix positioning:

Collection- Innovation and learning, the indicator measures the collection digitalization process, this represents a proxy both of the innovation efforts and of the educational opportunities given by the museum. In fact, a digital collection allows people from all over the world to have access to it and so to have the chance to explore it and learn from it, this could also increase the curiosity towards the museum artworks and the propensity of potential customers to become actual visitors.

Link to strategy:

The collection digitalization index is linked to the strategic goal “*Improve collection accessibility*”, the collection digitalization is, in fact a way to broaden the collection accessibility for people from all over the world that can't have physical access to the museum but also for situations in which there is an exogenous factor that obstacles the museum accessibility to customers (lockdown due to a pandemic, floods, special laws, etc.)

Additional Notes:

Two indicators are presented, the former that takes into account the number of pieces inventoried and digitalized, the latter that takes into account the number of exhibition rooms digitalized.

A quick overview of some of best-known museums digital offering is proposed in order to show the growing relevance of the collection digitalization phenomenon indicator is proposed:

- The Louvre: has free online tours of three famous exhibits, including Egyptian Antiquities.
- Solomon R. Guggenheim Museum: The works of Pablo Picasso, Piet Mondrian, Jeff Koons, and Franz Marc are just some of the 625 artists whose work are a part of the Guggenheim's collection online.
- Smithsonian National Museum of Natural History: a 360-degree room-by-room tour of every exhibit in the museum is available.
- Van Gogh Museum: You can get up close and personal with the impressionist painter's most famous work thanks to Google Arts & Culture.
- Getty Museum: Los Angeles's premiere gallery has two virtual tours, including "Eat, Drink, and Be Merry," which is a closer look at food in the Middle Ages and Renaissance.
- The Vatican Museum: The Sistine Chapel, St. Peter's Basilica, and Raphael's Room, are just some of the sites you can see on the Vatican's virtual tour.

- Thyssen-Bornemisza Museum: Madrid's must-see art museum has the works of some of the continent's most celebrated artists like Rembrandt and Dali available online.
- Georgia O'Keeffe Museum: Six virtual exhibits are available online from this museum named for the "Mother of American modernism."
- National Museum of Anthropology, Mexico City: a dive into the pre-Hispanic history of Mexico with 23 exhibit rooms full of Mayan artifacts is available.
- British Museum, London: The Rosetta Stone and Egyptian mummies are just a couple of things that you're able to see on a virtual tour of the museum.
- NASA: Both Virginia's Langley Research Center and Ohio's Glenn Research Center offer online tours for free. Also, you can try some "augmented reality experiences" via The Space Center Houston's app.
- National Women's History Museum: Have a late International Women's Day celebration with online exhibits and oral histories from the Virginia museum.
- Metropolitan Museum of Art: Though the Met Gala was cancelled on 2020 because of Covid-19 pandemic, you can still have a peak at the The Costume Institute Conversation Lab, which is one of the institution's 26 online exhibits.
- High Museum of Art, Atlanta: This museum's popular online exhibits include "Civil Rights Photography" — photos that capture moments of social protest like the Freedom Rides and Rosa Park's arrest.

- Detroit Institute of Arts: Mexican art icon Frida Kahlo is the focal point of two of the four available online exhibits.
- Rijksmuseum, Amsterdam: The Golden Age of Dutch art is highlighted in this museum which includes the work of Vermeer and Rembrandt.
- National Museum of the United States Air Force: You can't take a ride in Franklin D. Roosevelt's presidential airplane, but you can check it out, in addition to other military weapons and aircraft, online in the Air Force's official museum.
- MoMA (The Museum of Modern Art): New York's extensive collection is available for view online.
- Museum of Fine Arts, Boston: The 16 virtual exhibits include a special section on 21st Century Designer Fashion.

6. Support services index

Name of the indicator: Support services index

Calculation:
$$\frac{\text{Number of best practices on support services achieved}}{\text{Total number of best practices on support services}}$$

Best practices:

- Presence of audio guides
- Presence of video guides
- Free flyers availability

- Free online information availability
- Museum app availability
- Museum free wifi for users
- Pannels with description of each art piece
- Bookshop online
- Multilingual guided visits

Measure these aspects with a binary variable that takes either value 0 (denoting the absence of the service) or 1 (denoting the presence of the service).

Matrix positioning:

Collection- Internal process, the indicator belongs to the internal process perspective since the achievement and implementation of best practices in terms of collection access support services derives from well-organized internal processes which allow the implementation, maintenance and update of those services.

Link to strategy: The support services index is linked to the strategic goal “*Favor collection enjoyment*” all the previously listed services that represent best practices are implemented in order to favor the collection enjoyment. Support services allow customers to have a better understanding of the cultural institution offering and encourage a sort of interaction between visitors and the museum collection.

4.2.3 Organization

	COMMUNICATION	COLLECTION	ORGANIZATION
VISITORS AND REPUTATION	1. Visitors Proactivity 2. Net Promoter Score 3. External Exposure Indicator 4. Followers Growth 5. Engagement rate	1. Temporal exhibitions impact on visitors 2. Special events occurrence growth	1. Visitors classification - a collection of indicators
ECONOMIC RESOURCES	6. Marketing and Communication to total budget 7. Social Media management budget to marketing and communication 8. Donors contribution rate 9. Communication impact on tickets selling	3. Investment on collection maintenance	2. Income sources - a collection of indicators
INNOVATION AND LEARNING	10. Learning activities growth	4. Accessibilities to minorities index 5. Collection digitalization	3. Personnel qualification 4. Personnel training hours
INTERNAL PROCESSES	11. Content creation growth rate	6. Support services index	5. Hospitality quality 6. Ticket digitalization 7. Environmental sustainability

Figure 13 - CCO Matrix: Organization section indicators

1. Visitors classification - a collection of indicators

Name of the indicator: Visitors Classification

Calculation

Visitors category – calculated as

$$\frac{\#visitors\ in\ a\ category}{\#total\ visitors}$$

Local visitors – calculated as

$$\frac{\#local\ visitors}{\#total\ visitors}$$

Foreign visitors – calculated as

$$\frac{\#foreign\ visitors}{\#total\ visitors}$$

Matrix positioning:

Organization - Visitors & Reputation: this set of indicators is computed to better understand and visualize *who* visits the museum, dividing the visitors by different categories. It is set in the organization column because it can be directly impacted by a great number of functions and operations across the institution. It falls into the *Visitors & Reputation* perspective as visitors are the central theme of this indicator.

Link to strategy:

The public is an important proxy of the appeal that the cultural institution is able to exert over different categories of people through its activities across all functional areas, including but not limited to marketing and social media effort. Understanding who are the visitors and most importantly who do not visit the museum, as well as how these numbers change overtime, reveal what the museum is lacking, which are the initiatives that work for the museum and which are not. Additionally, visitors are a source of income on top of representing an important target and goal towards the mission of the museum itself. The visitors classification index is linked to the strategic goal “*Better understand museum visitors’ characteristics to match their needs*”.

Additional Notes:

This set of indicators can be boiled down to a ratio. It's a flexible indicator that can adapt to different data visualization approach. The main difference between the first indicator (visitors' category) and the other two, is the nature of the classification. In the first indicator, visitors can be easily classified based on the type of tickets they use to enter the museum (regular, student, elderly, etc.), hence the indicator can be tracked by simply counting the different tickets that are scanned at the entrance of the museum. The second and third ones require an extra study to determine the origin of the visitor, but it can lead to more meaningful information as museums' effort towards building an offering in a network of local entities are increasing, and the appeal on locals and tourists should be evaluated separately.

2 . Income sources – a collection of indicators

Name of the indicator: Income sources

Calculation:

$$\frac{\#income\ from\ a\ source}{\#total\ income}$$

Matrix positioning:

Organization - Economic Resources: this indicator fits perfectly under the organization column as it is interested by potentially all activities of the museum that are not for support, and it is a financial indicator, falling right into the economic resources perspective.

Link to strategy:

The income sources indicator serves as a simple information about what operations of the institution generate money, and its evolution in different periods of time can be easily observed. This gives both a measure of the success of an operation on a financial term, as well as an idea about the activities that are self-financing themselves and the ones that are not. The income sources indicators are linked to the strategic goal “*Track the impact of different revenue streams*”.

Additional Notes:

As a financial indicator, it doesn't react to changes immediately and requires for accounting operations to be performed before it can be computed, so it serves as a view over the past periods instead of a real time indicator.

3. Personnel qualification

Name of the indicator: Personnel qualification

Calculation:

$$\frac{\#qualifications}{\#total\ personnel}$$

Matrix positioning:

Organization - Innovation and learning: the indicator tries to estimate the value of the human resources across the whole organization, which is not an easy and straightforward task, but it is fundamental to draw an outlook on the future in terms of innovation, expansion and research.

Link to strategy:

This is a typical indicator made to measure the unquantifiable asset that is human resources. Human capital is part of the *intellectual capital* (Arnaboldi, Azzone and Giorgino, 2015) and it is specifically important for an institution such as a museum that is not dedicated to profit, but to research and culture. The personnel qualification indicators are linked to the strategic goal “*Increase the presence of qualified staff*”.

Additional Notes:

This indicator can be geared towards more specific measurements if necessary, depending on how the data is collected. Qualifications can be filtered out or weighted differently.

4. Personnel training hours

Name of the indicator: Personnel training hours

Calculation:

$$\sum_{i=1}^n x_i$$

where x is the number of hours of training that the employee i received.

Matrix positioning:

Organization - Innovation and learning: the indicator tries to estimate the value of the human resources across the whole organization by flipping its perspective compared to the *personnel qualification* and looking at how much the institution itself invest to grow its human capital, an action that enters into a wider and long-term program for innovation and learning.

Link to strategy:

Just like the personnel qualification ratio, the number of training hours is a proxy of the value of the human capital. However, it also brings another piece of information, as it is a measurement of how the museum allocates resources to invest in its human capital, which in turn is a measure of how attractive can be for potential employees and how much the museum is committing and investing towards its betterment. The personnel training hours indicators are linked to the strategic goal “*Increase the presence of qualified staff*”.

Additional Notes:

The summation can be filtered out to focus on employees of a specific type or from a specific area. Depending on how much information is available, it can also be filtered by time, it can be displayed as an average, etc. These and more possibilities all emerge during the transition from framework to dashboard.

5. Hospitality quality

Name of the indicator: Hospitality quality

Calculation

Hospitality quality – calculated as

$$\frac{\#best\ practices\ matched}{\#total\ best\ practices}$$

Matrix positioning:

Organization – Internal processes: hospitality is a value of the user experience that depends on a collection of different small elements with different owners inside the organization, which need to work together to design and provide an up to standard quality interaction path to users.

Link to strategy:

The Hospitality Quality is key into providing a good experience to visitors and in return positively affect the pursuit of many of the museums goal either directly (better access to the collections) and indirectly (reputation between locals through word of mouth and between tourists through feedback and review online, as well as between other local institutions to collaborate with). For these reasons, it is crucial to monitor the success of a museum into providing quality services both on the spot and overtime, as consistency is very important for this metric. The Hospitality Quality is linked to the strategic goal “*Improve customer experience*”.

Additional Notes:

The quality parameters are most likely self-evaluated through a guideline that can be either be established by a different institution or self-established. The indicator as presented uses a binary system, with each practice being either up to standard or not up to standard.

6. Ticket digitalization

Name of the indicator: Ticket digitalization

Calculation:

$$\frac{\#tickets\ sold\ via\ digital\ channel}{\#total\ ticket\ sales}$$

Matrix positioning:

Organization – Internal processes: the ticket digitalization indicator aims at assessing the solidity of the purchase process for customers and the back-end operations that enable different channels for sales.

Link to strategy:

Digitalize the purchase experience is specifically important for any business in 2020, as e-commerce grows and consumption of bought online services increase, an already existing trend that got boosted even more during the coronavirus crisis (Casaleggio Associati, 2020). The indicator also serves as a way to highlight which channels are effective and are preferred from the customers. The Ticket digitalization indicator is linked to the strategic goal “*Improve customer experience*”.

Additional Notes:

This indicator can be computer through a numerator’s specification increase. As an example, the numerator could be the number of tickets sold through the official website to people that started their navigation from the social media page of the institution.

7. Environmental sustainability

Name of the indicator: Environmental sustainability

Calculation:

$$\sum_{x=1}^8 x_i$$

By counting the number of items present (assigning $x = 0$ when an item is absent and $x = 1$ when it is present and then summing up), we obtain an index for environmental sustainability whose values are comprised between 0 and 8 (Basso, Casarin and Funari, 2017).

The items are:

- use of energy-saving lamps;
- presence of high efficiency systems;
- use of water saving devices;
- differentiated waste collection;
- use of ecofriendly products;
- thermal building insulation;
- monitoring of the museum's energy consumption;
- procurement of power supply.

Matrix positioning:

Organization – Internal processes: environmental sustainability is achieved through attention and efforts that span every process throughout the whole organization.

Link to strategy:

In recent years sustainability issues have aroused more and more attention in the museum sector (Basso, Casarin and Funari, 2017) as well as in every other industry. Environmental performances are relevant for both internal and external accountability (Arnaboldi, Azzone and Giorgino, 2015) and as such, they need to be reported. The Environmental sustainability index is linked to the strategic goal “*Support environmental sustainability*”.

Additional Notes:

The calculation method presented is an empirical recognized method that simplify auto-evaluation and comparison. Depending on the context, especially for single institutions, the calculation method can be updated to better reflect the sustainability level of the institution. The indicator only takes into account *environmental* elements and no other sustainability items such as presence of ramps for accessibility.

4.2.4 CCO Matrix overview

The CCO Matrix proposed is not the final output of our research, but it is an instrument which makes the construction of an efficient dashboard straightforward, the benefits and the potentiality of the CCO Matrix transformed into a dynamic and navigable dashboard will be presented in the next chapter.

The matrix stands at the base of the dashboard initiative as it guides both its data sourcing as well as its visual implementation. It is important, at this stage, to underline and recognize the limits of the matrix. The approach used to build it is an hybrid as it presents a data driven approach while conforming to the specific public agenda of the Ministry of Cultural Heritage and Activities.

	COMMUNICATION	COLLECTION	ORGANIZATION
VISITORS AND REPUTATION	<ul style="list-style-type: none"> 1. Visitors Proactivi 2. Net Promoter Score 3. External Exposure Indicator 4. Followers Growth 5. Engagement rate 	<ul style="list-style-type: none"> 1. Temporal exhibitions impact on visitors 2. Special events occurrence growth 	<ul style="list-style-type: none"> 1. Visitors classification - a collection of indicators
ECONOMIC RESOURCES	<ul style="list-style-type: none"> 6. Marketing and Communication to total budget 7. Social Media management budget to marketing and communication 8. Donors contribution rate 9. Communication impact on tickets selling 	<ul style="list-style-type: none"> 3. Investment on collection maintenance 	<ul style="list-style-type: none"> 2. Income sources - a collection of indicators
INNOVATION AND LEARNING	<ul style="list-style-type: none"> 10. Learning activities growth 	<ul style="list-style-type: none"> 4. Accessibilities to minorities index 5. Collection digitalization 	<ul style="list-style-type: none"> 3. Personnel qualification 4. Personnel training hours
INTERNAL PROCESSES	<ul style="list-style-type: none"> 11. Content creation growth rate 	<ul style="list-style-type: none"> 6. Support services index 	<ul style="list-style-type: none"> 5. Hospitality quality 6. Ticket digitalization 7. Environmental sustainability

Figure 14 - CCO Matrix Overview

5 Methodology

This chapter features the methodologies and tools applied during the investigation of the objective of the thesis. The research effort can be described as double sided, split between an academical approach and a practical approach, ultimately providing a contribution to the academic literature while offering practical advantages, insight and a useful business tool to museums and the Ministry of Cultural Heritage and Activities.

This chapter has the purpose to:

- provide an overview of how the literature review has been performed for every part of it;
- elucidate how the new identified framework can be applied into a real case and how the proofs of concept performed have been built.

In order to build the result of this thesis, four main inputs have been taken into consideration:

1. *Social Media Reputation Dashboard*, a tool built to monitor the Social Media Reputation of 100 selected museums in the Italian territory, made possible by the collaboration between the Mibact and Politecnico di Milano, of which an extended analysis has been conducted and described in Chapter 9.3;
2. *Decreto del Ministero dei Beni e delle Attività Culturali e del Turismo N°113 del 21/02/2018* «Adozione dei livelli minimi uniformi di qualità per i musei e i luoghi della cultura di appartenenza pubblica e attivazione del Sistema museale nazionale», an important decree setting up the National Museum System and its quality standards

and conditions, as well as public documentation regarding the Ministry of Cultural Heritage and Activities, available on the official website, which paints the context in which the ministry operates. These information are further detailed in Chapter 4.1;

3. The *literature review* presented in Chapter 3;
4. The Dashboards Observation presented in Chapter 5.1.

5.1 Dashboards Observation

The academic literature offers many high-level concepts on how to design and technically build a dashboard. However, it focuses more on the expected results rather than offering immediately applicable design guidelines to build one. To overcome this issue, this research focused on studying already existing, observable public dashboards. The aim is to identify and implement a design language and features that users are already familiar with, to make the dashboard understandable and intuitive to any kind of user.

5.1.1 Dashboard Best Practices identification

The dashboards taken into consideration for this observation effort include:

- *Youtube Studio Dashboard*: its purpose is to help content creator on Youtube monitor their videos and account performances;
- *Google Admin Dashboard*: this dashboard is part of the administrator console for Google G Suite accounts;
- *Covid 19 – Situazione Italia*: dashboard made to monitor the evolution of the Covid19 spread in Italy;

- *Facebook and Instagram performance panel*: these dashboards are made to track the performances of an account on the relative platform, the latter specifically on mobile only;
- *Tableau Public*: a collection of open dashboards made on the Tableau platform;
- *Squarespace performance dashboard*: it allows to monitor the performances of a website and track user interactions;
- *Twitch's streaming management and summary*: a collection of two dashboards to use during or past a streaming session on the Twitch platform, to measure performances;
- *Dublin Dashboard*: made to monitor in real time the city of Dublin;
- *Smithsonian Dashboard*: the Smithsonian Institution is the world's largest museum, education, and research complex, and has developed an open dashboard for its stakeholders to assess its performances.

What follow is the result of this observation.

Color coding

One of the most used tool in data visualization is simply *colors*. Colors are an effective medium for communicating meaning. In data visualization, color sets the tone and enforces a message for the underlying visual display. It creates a certain atmosphere and can turn an unassuming visualization into an emotion-filled data story. Of all the design elements in a given data visualization – the headings, the analysis, the comparisons, and so on – color is arguably one of the most important and can speak to your audience in many ways (Murray, 2019).

Color coding is a flexible tool that can be bent in many ways. It is for example possible to track a specific event through data visualization and coloring it with a strong dark color, while leaving everything else pale. This creates a contrast between the empathic strong dark color which *literally* makes everything else pale in comparison. An example can be observed in *Flu Pandemic* from Justin Davis, available on Tableau Public.

Colors can also be used to connect different metrics. By highlighting two related metric with the same color, it is possible to associate the two metrics without having the user dig deeper. This concept is used with great effect in *No more Malaria* by Daniel Cairoli.

The most common use case for color, however, is the simple green/red dynamic that indicates a bad or a good performance. A particular data is going to be colored green if it indicates a good performance, or in red to indicate an underwhelming one. This concept can be extended to different shades of green and red to indicate how big the impact is on either ends. This concept is extensively used in dashboards that update themselves with new periodic data. It is used in the *Youtube Creator Dashboard* to indicate the increase or decrease of a specific metric compared to the previous period T.

Well-chosen colors reduce the time to insight for your viewers and helps them understand your message sooner and more easily. (Murray, 2019).

Dark themed design

During the last few years it has been possible to observe a growing trend. Users seem to have taken interest into *dark themed design*, a color scheme design that uses light-colored icons and text on top of dark-colored backgrounds. Dark mode has been first introduced as default

on Windows Phone 7 and Samsung TouchWiz as a battery saving feature on devices mounting an OLED display. From 2015 to 2018, the popular services Discord, Youtube and Google's Android all introduced an optional *dark mode*. However, Google Trends reveal that the interest towards dark mode really started around June 2018 and spiked in October 2019 (figure 15), when Apple released iOS 13 featuring native dark mode support. Since then, many other popular applications introduced dark mode, including:

- popular messaging-service Whatsapp;
- social media platforms Facebook, Reddit and Instagram;
- Chromium-based (Google Chrome, Opera, Microsoft Edge) and Firefox web-browser;

Most notably, the *prefers-color-scheme* property was added to CCS (Cascading Style Sheets), a cornerstone technology of the world wide web and the language for describing the presentation of Web pages, including colors, layout, font, size, spacing of content, interaction and much more (World Wide Web Consortium). This allows front-end web developer to set up a responsive design that enable dark mode on websites for people that may prefer it.

The popularity of dark mode comes not only from the battery saving feature and general aesthetic, but also from the benefits brought to user health. Blue light emitted from devices at night suppress melatonin secretion, negatively affecting sleeping, making it a prime reason why people don't get enough sleep. It also *may* contribute to the causation of cancer, diabetes, heart disease, and obesity (Harvard Health Publishing; 2020).

For these reasons, dark mode has worked its way into dashboard design as well, and it is observable in many instances such as in the *Dublin Dashboard*. Dark-themed dashboard do not get in the way of correctly color coding the visual elements of the dashboard, and Power

124

Bi features a dark-themed preset to quickly turn a regular dark-on-light dashboard color-scheme into a light-on-dark one.

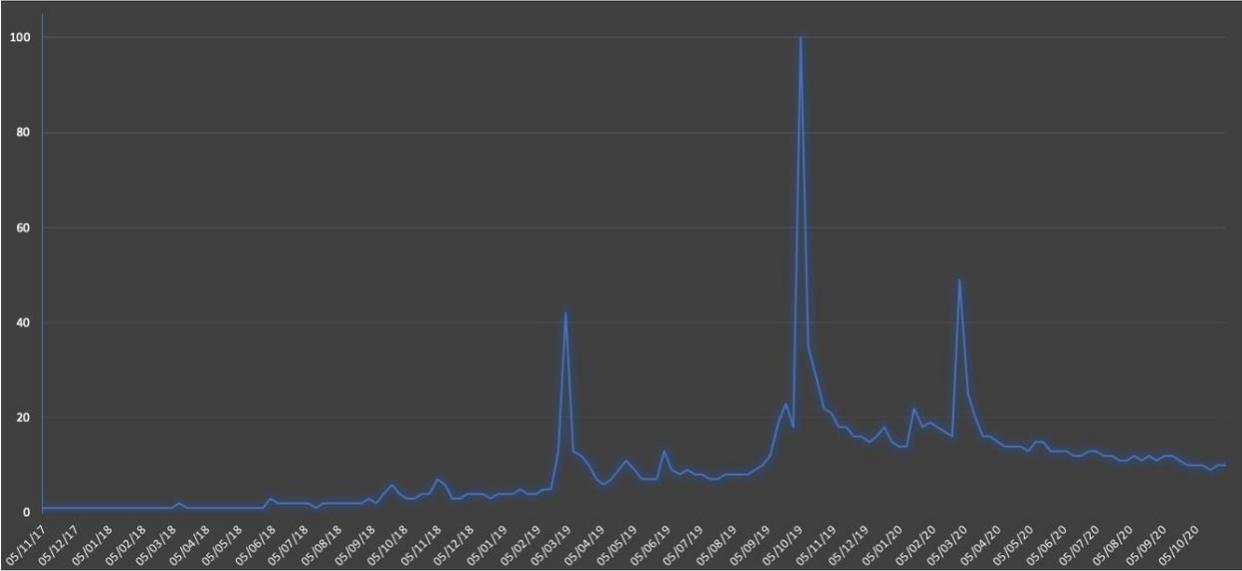


Figure 15 - "Dark mode" search query popularity on Google over time. Source: Google Trends

Filtering

Dashboard filters make it easy to provide different combinations of data from a single dashboard, without the need separate dashboards for different sets of users (Salesforce definition). Each group can have a filter that makes sense for it, personalizing the dashboard with easy-to-understand user inputs. Filters are observable in many different dashboards, but also on services such as e-commerce websites or careers portals. They generally appear on the left of the screen or on the top, as check boxes that can be toggled on and off, but they can vary depending on the nature of the filtered object. A popular example of filter that do

125

not usually appear as a checkbox is the time-period considered. A time period filter usually appears as a slider that allow to select starting and ending dates on a timeline, or as a drop-down style menu that allow to pick between popular time periods relative to the current day, such as: *current month; current semester; current year; past 365 days; past 30 days; etc.* In some instances, such as *Covid 19 – Situazione Italia*, a time period can be selected from a virtual calendar. Filters are usually applied to all visuals across the dashboard. There are examples of filters specific to a visual that are usually selectable on the visual itself, making the single object in the dashboard customizable. This is also observable in the *Covid 19 – Situazione Italia* dashboard.

Benchmarking

Benchmarking is a critical part of performance management system and it is a practice that can and should be integrated in any dashboard, as it is a feature that have the better chance to impact and direct the behavior of the decision maker and the organization. It is observable accompanied with real-time dashboards for smart cities, as benchmarking sets an aspirational and competitive agenda for areas/cities in terms of their relative performance to other locales and thus can be used to motivate policy changes deemed necessary to alter their relative Regional Studies (Kitchin, Lauriault and McArdle, 2015). The *Dublin Dashboard* visually compares some of its performances with the Mid-Est area and the entire state of Ireland average. Some municipalities use indicator and benchmarking initiatives to underpin forms of new managerialism, wherein they are used to guide operational practices with respect to specified targets and to provide evidence of the success or failure of schemes, policies, units and personnel, with the performance being used to discipline

underperformance, reward those meeting and exceeding targets, and to guide new strategies, policy, and budgeting (Craglia, Leontidou, Nuvolati and Schweikart, 2004). Observation from city-dashboards is important as cities are an example of a complex multilevel governance. Because of that, indicators are but one element, albeit an important one given their factual, standardized, time-series nature, in a tangled web of processes that help guide strategic decision-making, acting as a common source of information and learning tools about elements of a city, rather than being an evaluation tool of operational programs or departments or the primary driver of policy and action (Kitchin, Lauriault and McArdle, 2015). In these cases, benchmarking is a somewhat limited tool that helps decision making by painting a clearer picture and providing a better context to understand data, without being a decisive assessment in determining the success or failure of operations.

Time to time comparison

Time to time comparison is a direct comparison between periods, and it is a fundamental feature that is observable in any dashboard that measures performances over time. It can be presented with an existing indicator or by itself, and in many forms. The main strengths of the feature are:

- it is immediately easy to understand
- depending on how it is visualized, it takes up very little space on a dashboard
- it is very easy to color-code consistently
- it gives an immediate direct performance feedback

Using a time-to-time comparison, the user can have a feedback on his most recent work with a glance. When paired with proper filtering, it allows to quickly evaluate the real impact of a

specific object (task, set of tasks, campaign, etc.) or of an external event. Many different applications are observable. The *Youtube Studio dashboard* employs many different visualization methods. In a single panel, the time to time comparison is made using both the percentage difference between the two selected time periods and the absolute value.

The percentage difference is used to indicate the evolution of indicators that are relevant on the time-frame, such as views generated and collective hours of view time generated. It is not important for the user to know the total sum of these values from the day the account is created, as these indicators are used to calculate the performances of the account in present day and past performances have no impact on the future.

The absolute difference is instead used alongside the subscriber count. Notice how (in figure 16) the subscriber count refers to the total number of subscriber and not the subscriptions gained during the period. This is because all the subscription generated in all periods of time are equally relevant for the user.

To conclude the observation, the absolute value difference is used alongside an everlasting counter, while the percentage difference is used alongside a counter that is filtered and calculated on a specific interval of time.

Youtube Studio Dashboard also uses clever visual indicators alongside the time-to-time comparison (figure 16). The color indicates a negative (red) or positive (green) performance compared to the other time taken into consideration, and also adds a second visual aid as an arrow that goes up or down to indicate, again, the performance's quality.

The same concepts can be also observed on many stock performance dashboards (figure 17), such as the one integrated within the Google Search engine, which uses both the absolute value and the percentage difference, the color coding and the up and down arrow.

Channel analytics

Current subscribers

576

-1 in last 28 days

Summary

Last 28 days

Views 820 ↓ 13%

Watch time (hours) 19.8 ↓ 19%

Figure 16 - Channel analytics panel from the Youtube Studio Dashboard



Figure 17 - Google's Stock Performance Dashboard

Screen space allocation

Dashboards need to fit the user screen in a tidy manner. This concept is expressed both in practical observable example as well as throughout the literature. The layout needs to accommodate resizing based on the user's screen size while maintaining readability. The location of different panels inside the dashboard layout, as well as their color, give an implicit hint to the user about their priority and value. According to the Tableau official website, most viewers scan web content starting at the top left of a web page. The recommendation is that once the main purpose of the dashboard is clear, the most important view element needs to be placed so that it occupies or spans the upper-left corner of the dashboard. The layout itself needs to help the user understand where to look for what. Dashboard-wide filtering is usually located on the top or on the left column, while filters for specific visualizations are located on the visualization itself. It is possible, however, to observe dashboards constructed on vertical webpages. They typically feature a bar on top of the screen containing main indicators categories, which are then divided into sub-groups. Each group is made in a block, and the pile of block forms a vertical scrollable page. This approach is used when each block is accompanied by a text explaining how to interpret the data, how the data was collected or how to interact with the data visualization. *Dublin Dashboard* and the *Smithsonian Dashboard* both use this approach, which appears to be ideal for monitoring public institutions through websites that can be viewed by anyone and on any device. The "pile of block" design guides the user while making him understand how to read the data, and is simple to adapt to any device.

Interactivity

From the observation of many different dashboards, it emerges how mostly all of them allow the user to directly interact with them and personalize their experience. The types of interaction made available to the users are generally of three types:

- filtering;
- navigation;
- drill-down.

Filtering in dashboards is mostly done through dedicated panels such as slider or checkboxes. It is sometimes possible to apply filter by interacting directly with the objects in the dashboard, as seen in the Google engine integrated Stock Dashboard, in which users can select a timeframe by selecting and dragging directly on the graph, like they would do for a string of text.

Navigation in a dashboard needs to be fast and clear. Many dashboard elements, usually smaller, are dedicated to navigating the dashboard and move between different views.

Drill-down is observable in dashboards such as *Youtube Studio*, in which clicking on some objects of the dashboard allow to dig deeper into and obtain more information specific to an object or entity.

5.2 From a model to a tool: CCO Dashboard construction

This paragraph depicts the approach and technique developed and deployed to build a business intelligence tool from the framework developed and presented in Chapter 4. To leap forward and moving from a static reporting instrument to a meaningful and powerful tool,

the aim is to build an interactive dashboard using the framework as a guide, having in mind the final goal of letting the user explore data visualization quickly, in an intuitive way, while enabling him to understand it and to extract valuable information from it.

The approach to build a dashboard described in this section has been drafted starting from the literature review as well as from some reasonable assumptions about what data is either available or can potentially be collected. The framework aims at being adaptable to different cultural institutions and has been used to create a prototype using Power Bi for our stakeholder (the General Direction of the Ministry of Cultural Heritage and Activities and the National Archeological Museum of Naples). For these reasons, the approach starts from the data and indicators and moving forward, rather than starting from a dashboard model and working backwards.

Power Bi Desktop allows to easily add dataset and gradually add, change or remove features of the dashboard down the line without writing a single line of code nor disrupting the existing service, a key feature of a solid dashboard (Farmanbar and Rong, 2020), enabling an iterative incremental approach.

Thanks to Power Bi, the process of building the dashboard doesn't need to be linear and can be reiterated multiple time to achieve the desired goal. Hence, the process presented here is not to be interpreted as rigid. The output of any step of the process can be revised and put into discussion, and at some stages this re-iteration is required to meet all the targets, which are making a dashboard that is useful and to make it convenient enough to enter the daily routine of its users. The overall process can be divided into the following activities:

- 1. Study of the framework**

2. Establishing overall design philosophy and view navigation
3. Preparing the dataset
4. Feeding and transforming datasets into Power Bi
5. Establishing relations between datasets
6. Translating indicators into visual representation
7. Adjustment in style and appearance
8. Technical testing
9. Functional testing

1. Study of the framework

The first step between the framework and the dashboard is understanding the framework itself. The framework provides valuable information in regard on how the dashboard needs to be constructed on a logical level, establishing relationships between each element of the dashboard to gear it in a way that allows different part to work independently while creating synergies and leveraging the same resources in terms of space, datasets, calculation, appearance, relations and filtering. At the same time, the dashboard design needs to translate these relations to the user experience. Once the framework has been applied to produce a draft of indicators that are both meaningful and extensive, the next step is to establish the dashboard's navigation and overall design.

2. Establishing overall design philosophy and view navigation

The framework covers a vast amount of data and different areas and information. To make it work, the dashboard needs to be designed using different views that the user can navigate freely. This allows for a creation of different compartments, and in the case of the first prototype, each compartments correspond to different cells of the matrix. Each one gets his own space and view (page), so it is important to design a consistent user interface (UI) across all of them to move from one view to the other. One way to do it is to create an additional view serving as a main menu and containing interactive objects such as clickable buttons that allow the user to navigate the views. In each view, there should be at least one way to go back to this main menu view. On top of that, other options are adding additional visual interactive elements that allow the user to navigate faster, such as arrows that allow to move through the different views sequentially, or buttons that directly lead to different views. To completely establish the overall design philosophy, we found that the more efficient way is to build a template to guide the population of each view. In the template, the space of each view needs to be allocated to different elements of the view itself, which in our case were four:

- data visualization area;
- title area;
- filter area;
- navigation area.

This approach improves usability as the user immediately recognize the different space allocated, without needing to resort to any kind of explicit instruction such as a textbox describing each area. In the final product presented to our stakeholder, each section of the

matrix has been assigned a different color, which is used in some elements of the views such as the background or the text. This differentiates distinct views with the same layout, keeping the benefits of the latter while not confusing the user regarding what section it's visualizing at any given moment.

3. Preparing the dataset

It is fundamental that the database is properly constructed in making the dashboard work. Power Bi Desktop supports excel files, but that doesn't mean that any excel file can be fed to it. Data tables needs to follow the logic described by Chen's entity-relationship model, which means that each line needs to represent a single entity and to contain one or multiple keys. Power Bi Desktop has some meaningful data transformation features, but it still requires dataset to be competent and properly structured. There should always be an entity set dedicated to the institutions and one dedicated to different days (literally a list of dates), as all other datasheets describe a relationship between these two entities. All datasheets should describe entities in the same exact way (example: *Colosseo* has to appear the same in all datasheets, it can't be *Colosseo* in one datasheet and *Parco del Colosseo* in another one), a problem that is commonly addressed with the use of sequential codes for entities. This characteristic allows to set up filtering based on the two entities or their hierarchies and categories, such as filter by semester or filter by type of institution.

4. Feeding and transforming datasets in Power Bi

Power Bi Desktop supports a shockingly high number of different sources of data. Once data tables are loaded, it is possible to transform them. Transforming datasets is required to extract meaningful information, as well as to ensure that the data injection worked. Each

single step of the transformation process is logged from Power Bi and can be independently eliminated or modified even after the dataset has undergone different transformation. The first requirement should be simply making sure that the header of each column is correct, and each column has been defined with the right data type (string, integer, decimal number, date, etc.). Through Power Bi, it is possible to clean the dataset by removing rows based on value (example: removing empty rows) or substitute values (example: substituting “third quarter” with a date “30th September 2020”). Other common transformations are adding column using Power Bi’s Data Analysis Expressions (DAX) and quick calculations, as well as separating, joining or removing column (example: joining the *day*, *months* and *year* columns into a single *date* column).

5. *Establishing relationships between datasets*

Before starting to work on the visual representation of the dashboard, it is required one last step, which is establishing relationships between datasets. This is generally done by Power Bi itself, but the process can be checked manually to ensure that filtering and drilling down works properly in the dashboard: establishing relationships between data table. While using this framework, all datasets contain the column *Museum* which is used to easily establish direct relationships between datasets, in three different forms:

- one-to-many
- many-to-many
- one-to-one

6. *Translating indicators into visual representation*

Indicators, in the form described by the framework, are really just numbers, while the dashboard allows to visualize information in graphical way, which brings many benefits in terms of representation and understanding. Graphics can leverage on visual aids such as color coding to deliver an information in a compact space, and it is sometimes possible to obtain a meaningful information from a dashboard with no more than a glance. During the dashboard creation, each indicator is evaluated and adjusted to be better visualized. A single indicator does not equate to a single visual. Indicators can be repeated, and more than one indicator can be represented in a single visual. Example: visualization models such as an histogram can be used to compare different values through time, or just the trend of a single value, or both. The disposition and space allocation are very important, as any given view cannot be crammed with objects. As seen in the literature review, not all visuals hold necessarily the same importance, and the visuals on the top should generally be the most meaningful and the one that takes up the most space. In the prototype prepared for our stakeholder, an histogram would occupy three fifth of the space allocated to indicators in almost all the views, expanding from one side of the screen to the other and holding many different information. With a complete dataset, data can be filtered and compared by time-period. Establishing the hierarchy for time-data (example: daily → monthly → quarterly → annual) is a step that needs to be done and entirely dependent on data gathering frequency. This necessity comes from the fact that different data could be gathered at different times, and the dashboard needs to aggregate different data sources to produce coherent indicators. As a practical example:

$$Indicator X = \frac{sum A_{daily}}{sum B_{monthly}}$$

In the example, A is a data collected each day, while B is a data collected each month. Setting up an indicator like this may lead to some mistake. If the user filters the entire dashboard by selecting a time-period between the 15th of January and the 27th of March, the ratio is going to be calculated as the sum of A collected during half of January, every day of February and most of March, while the sum of B is going to be January (which includes the first fourteenth days of January) and February, while excluding March entirely. This indicator is immediately invalidated by such a practice. It is important then to establish a hierarchy and completely exclude the information regarding the day of the months to which A refers to, making the indicator reliable and fixing the issue (figure 18).

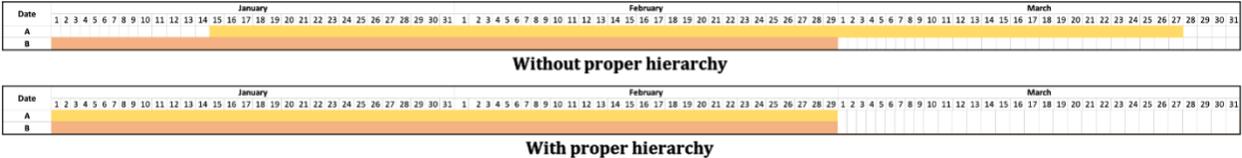


Figure 18 - Visualization of the difference between proper and improper hierarchy

This problem is especially present for social media. To gather non-social media data, for a dashboard that aims at describing the situation of many different cultural institutions at once, the reasonable assumption is that questionnaire will need to be sent out and collected. This operation can take months, and questionnaire can change overtime or be only partially filled. On the contrary, user-generated-platforms' APIs allow to gather data with a terribly high frequency, and metadata collected by these platforms contain information about the exact second in which an entity has been generated (example: when a comment was posted). This creates a terrible disbalance between the two kind of dataset that need to be addressed when dealing with social media indicator and how they behave in the dashboard.

The process of turning indicators into visual objects is the moment in which most of the good practices observed in the literature review should be applied. However, every dashboard is different, and depending on what exactly is the focus of the designer and of the user, there's not a rigid approach to follow. The coronavirus crisis of 2020 spawned many different attempts at public dashboards geared towards monitoring the evolution of the pandemic, and many different approaches has been shown with different degrees of success, there's not a single specific way to do it.

7. Adjustment in style and appearance

Once the visual objects are set, the style and appearance of the dashboard needs to be checked in order to verify the coherence between all elements. This mainly comes down to three elements:

- formatting of text
- colors
- spacing and sizing

All text labels, so every title, axis label, numbers, etc, should be coherent. Labels can differ between types, but not in the same category: the color, the size, the font, the spacing between letters, their position in respect to other elements should all be the same (example: text alignment for each visual object title should be set in the same way or have a thought out reason to be different).

Colors include the color of the background, the color of different text labels, the color of the background of each single visual object, the color of the borders, the color of lines, points, bars, selectors and of all kind of visual elements inside the dashboard.

Spacing and sizing is also important. Buttons' size should reflect their priority and they have to be all correctly aligned, the borders thickness has to be consistent, as well as the spacing between different blocks.

Power Bi can help with some of these tasks as it can automatically revamp the style through some presets, but spacing and sizing is a time consuming task that needs to be done manually. Thankfully, even if Power Bi allows for dragging and dropping elements on the dashboard, it is possible to control both size and position of each element down to each single pixel by manually inputting coordinates and length and height.

8. Technical testing

Technical testing is a required step before letting the user try the dashboard out, to ensure that everything works properly. This means testing the navigation to make sure it works as intended, verify the correct interaction with different visuals and especially trying out different filtering options, to make sure that all visuals respond correctly to the input and all dataset relations are properly set up. During this step, conditional formatting should also be tested (in our prototype, this feature has been used to shift color of an indicator from red to green depending on the value).

9. Functional testing

The last step is the feedback. What works and what doesn't work in the dashboard, what is being used more and what is being used less. What is distracting and what isn't. This feedback has to come mainly from the final user itself and is most likely, but not only, related to the transition between indicators to visuals.

As discussed at the beginning of the chapter, these steps are not to be taken as isolated compartments, but they lay the foundation for a more effective and organized approach that should somewhat limit logical mistakes while going from the framework to the dashboard.

5.3 Literature review

The following is the methodology employed to execute the literature review presented in this thesis. The effort spent in studying the academic literature channeled towards two main broad topics:

- data reporting and business intelligence tools;
- user generated content platforms.

While researching interactive dashboards and data visualization as a whole, the goal was to analyze how to create a more comprehensive business intelligence tool that do not stop at communicating data, but also becomes an instrument used to take operational decisions, and that can be adapted easily to different scenario thanks to its flexibility.

Dashboard is a flexible and broad concept that can adapt to many different industries and use. For these reasons, the study of the literature focused on three main lines:

- how dashboards affect the status quo, the organization and the processes already present in the environment in which they are introduced;
- what are data, how to manage them, what are the underlying technologies and logical models behind a business intelligence tool;
- what are the approaches, the challenges and the objectives in dashboard design.

Attention has been put to both understanding how to design and build a dashboard, as well as how the dashboard interacts with the existing organization, processes, and people. To achieve this, both recent and legacy papers have been taken into consideration. Additionally, better understanding of these concepts has later been acquired thanks to the practical application of these concepts in building the prototypes presented in Chapter 6.

In regards of social media and user-generated-content platforms, the nature of the data generated by these platforms, and the best way to enhance these data to turn them and visualize them in a meaningful and clear way to the final user, the goal was to clearly establish the main elements of our analysis (social media and UGC-based services) and investigate the gap between them and their successful implementation into a business intelligence tool. We relied on five main kinds of sources, which are:

- *recent* academic papers – paper sourced mainly from scopus.com and Google Scholar, published between 2018 and 2020, free to use, generally focused on very specific scenarios both related to the extrapolation and the presentation of data;
- *legacy* academic papers – papers that contain important concept definitions that are considered standards today, they have been used in the analysis and definition of generally understood concepts such as Web2.0, social media platform and big data;

- official platform's webpages for businesses – these are the support website built for businesses that are interested in extracting the best possible value from platforms such as *Instagram* or *Tripadvisor*, used as a main source of information regarding the functionalities offered by the different platforms that can potentially become unvaluable data sources;
- marketing reports – as internet and social media are an everchanging aspect of society that gain importance over any kind of business, many reports keep track of how the services and the response to these service is evolving, and these information has proven to be important to better contextualize the data generated and captured online;
- observation – insight is provided by many user generated content platforms themselves, as over the year they developed easy, accessible dashboard for the performance analysis of a single account.

The overall analysis includes 29 papers among the other data sources. For recent academic papers, the research was performed by looking into specific keywords such as:

- social media
- *social media name* (Facebook; Instagram; Twitter; LinkedIn; etc)
- *UCG platforms* (Google; Maps; Tripadvisor; AirBnb; etc)
- data visualization
- data analytics
- dashboards
- cultural institutions

- ...

5.4 Tool development and application

What follows is the methodology adopted in developing the prototypal dashboards discussed in Chapter 6 “Results and discussion”.

5.4.1 Proof of concept I

On November 26th 2020, we presented a working demo of a dashboard made with Power Bi Desktop to the Museum General Direction office. The demo was composed by five working views based on four cells of the COO Matrix presented in Chapter 4.2, with the addition of a landing page. These views are:

- Landing page
- Organization – Visits and revenues
- Organization – Welcoming and ticketing
- Collection – Catalogue quality
- Communication – Reputation

The presentation has been held in Italian. Despite being the first demo presented, it wasn't the first demo prepared during the development of the thesis. As such, some experience and knowledge on how to build one were already gained and partially affected the approach used to build this one. To prepare the demo, we first decided what views we wanted to develop, focusing on creating a well-balanced and diversified dashboard that would allow us to show

off as many visually intuitive features as possible in order to spark the interest of our stakeholder and convey the flexibility and strengths of Power Bi without resorting to any kind of technical speech. Once the choice of the views was done, some ground rules about the data have been set:

- data would be referring to 18 nameless museums;
- museums would be split and filtered between administration area and type;
- all data would be referring to the second and third quarter of 2020.

All data sources have been made as excel files and generated using the online realistic data generator Mockaroo (<https://www.mockaroo.com>) and/or through manual generation by employing Excel's own random functions. Mockaroo offers a free online service to generate tables up to 1000 rows that can be filled with 143 different types of data. For this project, four types of data have been mainly implemented:

- custom lists (string of text)
- integer number
- decimal number
- dates

Other data types have been initially used but later discarded. Information about dates in which data are collected have been all added manually. Rows have been filled following a sequential or random approach, and in some case following a custom distribution of values. This was done to ensure that every possible combination of *value x museum x quarter* was obtained, as datasheets are relationships between the date and museums entity sets. For this reason, the first table to be made has been the *Museums list*, containing museum's name, type

and administration type. All other table were made just before or during the making of the first dashboard view in which they were required.

Despite the fact that extensive care was adopted into generating data, part of the data transformation has been performed directly on Power Bi Desktop. This was done mainly for three reasons:

- data have not been tailored to work on the go on Power Bi Desktop to ensure that the demo would represent a feasible final product working on real data;
- to further test the transformation capability of Power Bi Desktop;
- to adapt data to the imminent requirements that were not immediately clear during the preparation process.

Once the data were ready, they were loaded on Power Bi Desktop. Before they could be used, two more steps were required:

- minor data transformation;
- establishing relations between existing tables through the key value *Museum*.

The design of the first view (Organization – Visitors and revenues) has been used as a blueprint for the other views, in order to maintain continuity through a recognizable pattern in the design. Every view feature filter options located on the left side of the screen. To ensure the same filters worked in every view, the filters featured are all used to filter museums, which are the common entity in all the data tables.

Buttons for the landing page and the navigation system were added after the four main views were completed.

5.4.2 Proof of concept II

On March 19th 2021, we presented a working demo of a dashboard made with Power Bi Desktop to the Museo Archeologico Nazionale of Napoli. A refitted version of this demo has been presented again to the Museum General Direction on March 29th 2021.

The demo was composed by five working views:

- Home
- Followers
- Activity Online
- Engagement
- Content

Additionally, the Follower, Activity Online and Engagement views all featured a secondary view dedicated to benchmarking functions.

Both presentations have been held in Italian. The prototype was built from the ground up leveraging the experience collected with the previous presentation. The demo had been developed to show off the possibilities offered by Power Bi leveraging real dataset collected from 100 different museums. Datasets included:

- Facebook, Instagram and Twitter accounts of 100 Italian museums
- Facebook, Instagram and Twitter posts made by aforementioned accounts
- news online relating the 100 Italian museums
- TripAdvisor reviews relating the 100 Italian museums
- Hashtags used in Social Media posts

The data was collected between January and December 2020 and relates to the whole year.

Datasets required extensive transformation. Mainly, tables relating different months of the year have been combined into a single one, date and time information have been split, many useless columns have been removed to improve performances of the Power Bi Desktop application. Additional tables have been manually crafted in a couple of instances in which data was not complete.

All tables are connected to the two main entity sets: museum (through ID code) and date.

To allow for a dynamic comparison between museums using filters selected by a user, many tables have been duplicated and connected to a second copy of entity sets.

The design of the first view (Followers) has been used as a blueprint for the other views, in order to maintain continuity through a recognizable pattern in the design. Every view feature filter options located on the top of the screen. On the left, a navigation panel has been added later.

The home page was designed for last, following a mock-up with precise space allocation for each element. All the colors used have been chosen with the <https://coolors.co> services, ensuring complementarity between them.

6 Results and discussion

The Ministry of Cultural Heritage and Activity monitors the performances of a network of museums by gathering data from every single one of them. To further advance the research and validate the framework, we created two concrete examples on how the model can be applied to generate a business intelligence tool capable of presenting information extracted by datasources referring to different entities. The objective is to offer a dashboard that helps monitor and navigate this data not only for the administrator's duties and operations, but also for the single institution that may require a personal evaluation.

6.1 Proof of concept I

To proceed in the model application and create a prototypal holistic dashboard, the effort relied on a data sample that was not sourced directly by the ministry, but rather generated randomly under the basic constraint that data included are supposed to be available already, based on their presence in the reports for other years, or easily gatherable if necessary, based on common sense. In the latter category are included all the data sourced from social media that we know for a fact to be easily accessible as lengthy documentation about social media API are available online and the same data sources were used in the *Social media reputation dashboard*.

The tool is composed by thirteen single-views, twelve of them corresponding to the different cells of the COO matrix and one being the main navigation menu. Users may navigate the different views through buttons. The design allows to reach any view in only two clicks or less.

The dataset is composed by two entity sets:

- museums
- time

All the information presented in the dashboard are relationship sets between the two entity classes. For this reason, all the data can be filtered by museum and by time.

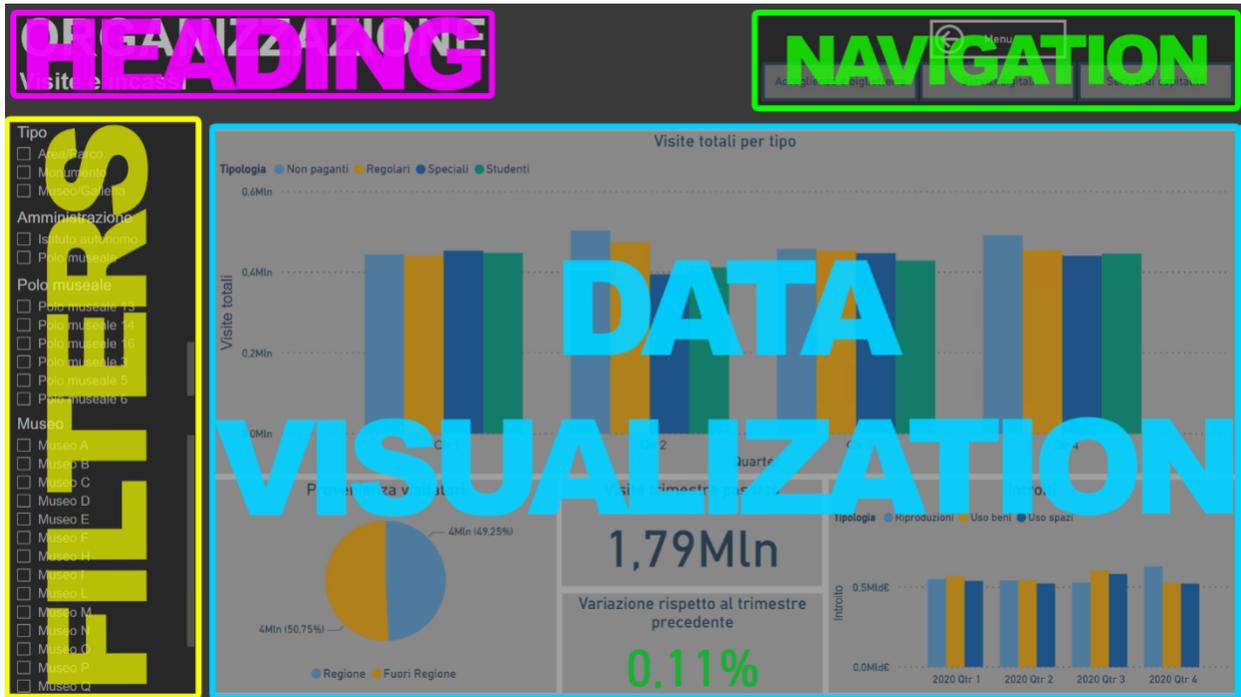


Figure 19 - MIBACT Dashboard view template

The views are split into different areas

- heading: it includes a bigger text indicating the category and a smaller one underneath indicating the focus of the single view;

- filtering: on the left side the user can access to a series of filters regarding the museum entities to be shown based on type, administration or even by selecting specific museums;
- navigation: on the top right users are offered a one-click navigation path towards a different view of the same macro-topic or two-click navigation path towards any other view of the tool, to ensure easy and quick navigation between views;
- data visualization: the larger portion of the view is dedicated to data visualization, with the most-regarded object put on top and taking up the most space.

Filtering can be performed through the panels on the left. It is however possible to move between different hierarchies of time and museums for specific customized visualization objects. Most of the data can be visualized split in months or quarters or years, depending on the user preferences. Some graphs are also re-arrangeable based on museum type if the data is a sum or percentage of museums that satisfy a specific condition (such as a quality standard).

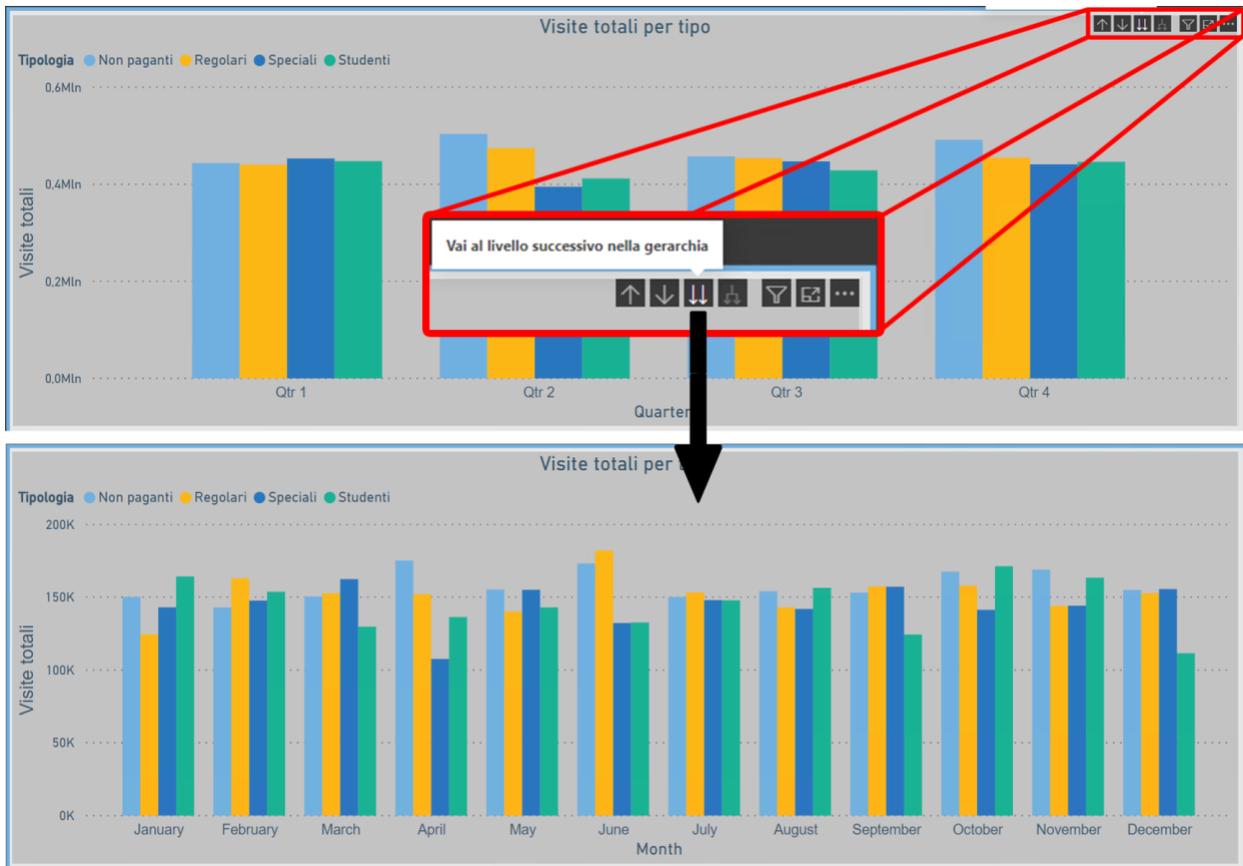


Figure 20 - Visualizing data through different level in hierarchy

The output design implements many of the best practices outlined in Chapter 5.1 “Methodology – Dashboard Best Practices observation” such as color coding, dark theme and screen space allocation. As the output was made for a presentation using mock-up data, the effort was focused on the dashboard appearance and presentation first.

Example of these implementations can be observed in Fig. 21



Figure 21 - Screenshot of the Comunicazione - Reputazione view

Fig. 21 is a screenshot of the *Reputation* view. The dominant color for all Communication views is green and it can be observed in various elements such as the borders of elements and the color font. This view still follows the same main template described earlier and the layout is mindful of the screen-real estate principles, with the most important indicator situated on the high left corner and taking up the most space. The data visualization includes clear benchmarks, comparison between different times as well as variations, both through visual aids as well as direct data-value visualization and ratio calculation. The screenshot also shows how an additional filter for dates, in the form of a slider, could be added, while the *Livello medio interazioni* histogram display data divided between museums type, allowing the user to change hierarchy as he pleases.

The presentation consisted in a live demo exploring four different views plus the menu. It focused on showing off the design language as well as features enabled by Power Bi while highlighting all the best practices implemented. The representative of the Ministry of cultural heritage and activities expressed particular satisfaction and interest for the visual representation, including the use of color coding and big clear interactive visualization objects for data, claiming that *appearances matter* (“l’occhio vuole la sua parte”).

Despite the stakeholder showing genuine interest, further development for this dashboard has been halted at this stage and no data is available for actual active use of the tool from the final user.

6.2 Proof of concept II

The first dashboard prototype served its purpose in testing out the macro steps laid out in Chapter 5.2 and presenting the features of Power BI to a stakeholder that has both the role of regulator nation-wide and decision maker. Once the approach in presenting and visualizing information has been approved, the next step of this researched focused on making these lesson work in a real scenario with real dataset.

6.2.1 March 2021 Prototype

The second dashboard prototype has been built and personalized for two different stakeholders:

- Direzione Generale Musei (Museum General Direction) which is part of the competent ministry;
- Museo Archeologico Nazionale di Napoli.

The first stakeholder is no different from the one that approved the first prototype, it's part of the Ministry of Cultural Heritage and Activities and has the role of regulator, decision-maker and supervisor for the different cultural institutions present in Italy. The second stakeholder is one of the largest and most active museum in Italy, selected because of its relevance (almost 700 thousand visitors in 2019 according to ISTAT), its presence online and its proven positive attitude towards innovation.

The major roadblock to the development of a working dashboard has been the data collection. Data regarding museums are disperse, not complete and require going through bureaucracy to get. For these reasons, in order to complete the research in a timely manner, the second prototype is strongly data driven, as it relies on data that was available at the time. It mainly focuses on a single cell of the COO Matrix, being the *Visitors and Reputation - Communication*. The dashboard relies exclusively on data gathered from Social Media and UGC-platforms during 2020 and uses the Museums entity dataset compiled by Politecnico di Milano beforehand. As a result, the whole initiative steered towards a stronger data driven approach. It is however worth noting that the data implemented in the dashboard were collected starting from the active input and request from the Ministry of Cultural Heritage

and Activities, making the final product inherently leaning towards the agenda of the stakeholder.

As per the presentation itself, the second dashboard has been built from the ground up and uses a slightly different approach in how it navigates compared to the first one. However, just like the first prototype, it is divided in different views, completely disregarding a one-page approach. Being exclusively reliant on social media data, the dashboard is strongly based of the first cell of the COO Matrix (Visitors and Reputation – Communication), which has been expanded in four main views:

- Followers - view dedicated to monitoring trends in followers growth across Instagram, Twitter and Facebook;
- Activity online – view dedicated to monitoring trends in museums’ activity online across Instagram, Twitter and Facebook, specifically how much content they post online (posts);
- Engagement – view dedicated to monitoring the effectiveness of online activity, mainly using quantitative measures;
- Content – view dedicated to monitoring the content not crafted and published by museums, which in the dataset is limited to users’ reviews on TripAdvisor and news published online that mention the museum.

Just like the first prototype, the dataset is composed by two entity sets:

- museums
- time

All the information presented in the dashboard are relationship sets between the two entity classes. For this reason, all the data can be filtered by museum (single museum or different categories of museums) and by time.

All of the categories are in line with already existing differentiation, which can be geographical areas such as regions or province, as well as a differentiation based on type of institution considered, which is already being used by the Italian National Institute of Statistics (ISTAT). An additional category called “size” have been made, based on the number of visitors each museum received during 2019.

The approach to the main navigation hub view completely changed compared to the first prototype, mainly because there are only four views. Instead of just having buttons, the view presents four quadrants representing the different views, featuring some of the visual elements available in the respective full view. Clicking on a quadrant allows to navigate to the full view.

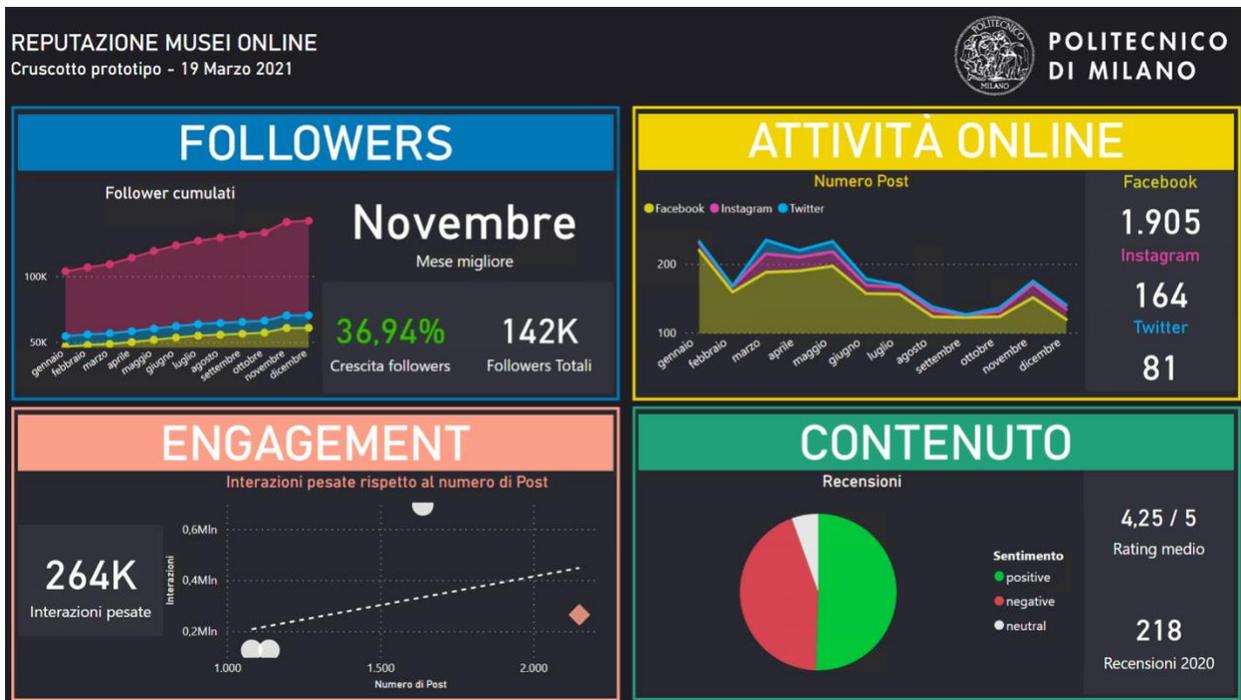


Figure 22 - screenshot of the Main view for MANN

The views are split into the same areas of the first prototype, but laid out differently

- heading: there's no heading anymore, instead the name of the view and the navigation toolbar are integrated;
- filtering: filters are on the top of the view and mostly consist in drop down menu, different views may have more or less filtering options, but the order in which they are laid from left to right is consistent between views;
- navigation: navigation functions are on the left side of the views, operated through a column in which the current view is highlighted;
- data visualization: the larger portion of the view is dedicated to data visualization, with the most-regarded object put on top and taking up the most space.



Figure 23 - screenshot of the Engagement view for MANN

Each view uses one of four selected complementary colors for its titles and relevant elements, to emphasize the difference between them. The dashboard uses the same dark-mode practice employed in the first prototype, using a more blue-ish tone rather than a grey one as it's complementary to the colors of the different views.

Compared to the first dashboard, the second prototype also features a special additional view for Followers, Activity Online and Engagement called Benchmark, that allow user to directly compare its performance/the average performance (for MANN and Mibac respectively) with the performances of museums filtered based on region, type, size or even down to a specific museum.

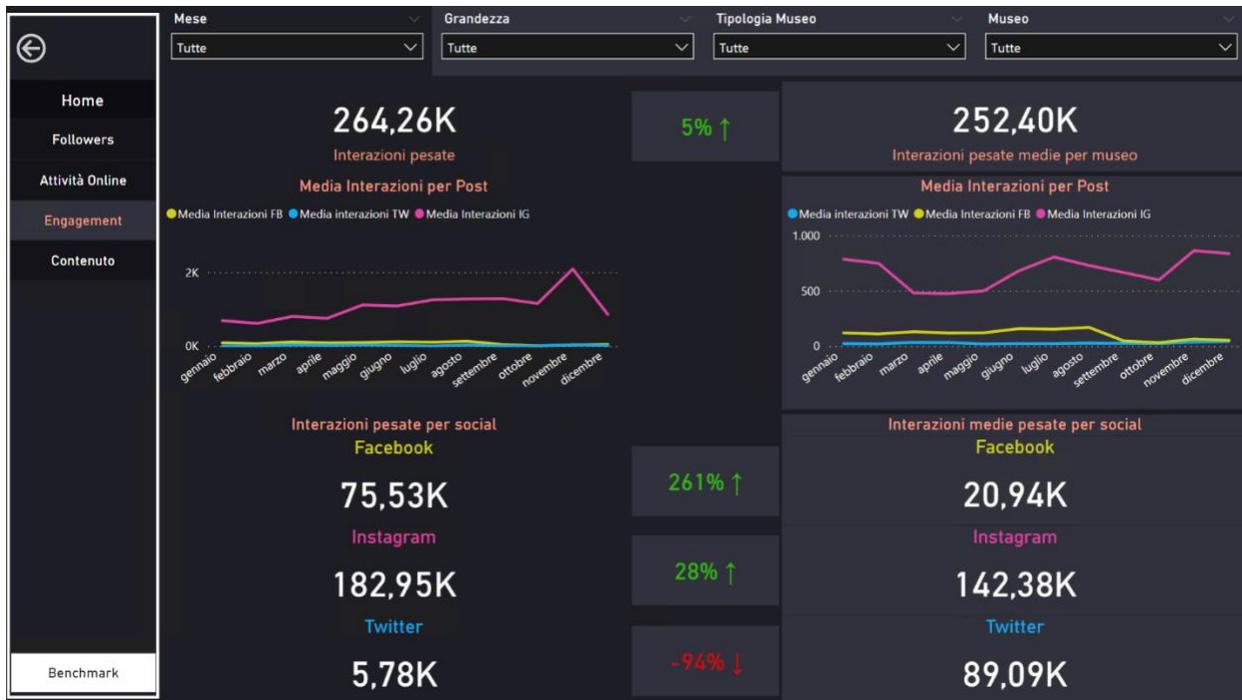


Figure 24 - screenshot of Engagement Benchmark view for MANN

6.2.2 The challenge of Data gathering

Data gathering is a fundamental activity that requires great effort from the interested parties. To make a leap forward and make a dashboard that doesn't stop at being a simple reporting tool but goes behind and truly becomes an interactive business intelligence tool to be consulted every week or every day, the data gathering needs to have strong foundation. Datasets need to be truly complete at all time, as the software relies on complete dataset to make calculations and these calculations can easily break and fall apart the moment data are missing or are inconsistent. Data needs to be collected periodically and sets of data acquired at different times need to be complete and directly comparable. For the first prototype, the

dataset has been crafted by us using random number generation, and as specified in Chapter 6.3.1 “Methodology - Proof of concept I”, it has been done ensuring that “every possible combination of *value x museum x period* was obtained”. Focusing on the second dashboard, overcoming this challenge by using of datasets already available and collected automatically via software has proven to be only partially true. Software is imperfect and hiccups happened at all times during 2020, resulting in data missing or being completely unreliable. The data collected is still plenty in quantity and can be elaborated and interpreted manually to craft insightful reports, but it is not sufficient for a dashboard that would ideally require less manual intervention and present recurringly fetched datasets. Data analyst could potentially cherry-pick indicators and visualization that give the most valuable information starting from the data that are available, but Power Bi cannot make an educated guess on what data is reliable and what is not. Given that data gathering requires effort from many different parts, the approach to building a dashboard needs to be incremental, testing out solutions that would require increasingly bigger and richer dataset, involving different parts of the organization at different times, but making sure that once the data is being collected, it is collected properly.

What follows are two instances of this data gathering challenge manifesting while crafting and presenting the second prototype dashboard. While studying the trends in follower’s growth during 2020, a strong dive in followers count emerged in February 2020. The way followers are tracked is very simple: the algorithm collects data directly from the Social Media account information, so that each account has its total followers collected at different times. However, if the algorithm stops collecting data for some reason, Power Bi visualization falls apart. This happened with the dataset employed for this research, as not all social media

account data were collected at all time. Specifically, in February two of the biggest accounts were not monitored, resulting in a reported loss of half a million followers and a consequent extremely negative growth rate followed by an extremely positive one the next month. To deal with this kind of issues we tried two different approaches, both of which came with some drawbacks:

- manually adding the missing data, in this example by estimating the original follower count in February based on the data from January and March;
- calculating the average total follower counts each month considering only the numbers of account that were actually tracked, and using that to display growth and trends.

The first solution poses the issue of actually manually generating datasets to fill the incomplete ones, which is a time-consuming task that doesn't address the issue but only fixes the instances in which the issue manifests clearly as it completely distort the expected results in the dashboard.

The second solution is more *elegant* as it is invisible to the user, but poses the issue of an incomplete information which gets less valuable the more incomplete the dataset is. It wouldn't work, for example, if data regarding follower is missing for all accounts in one month, which happened in January 2021.

The second instance of data gathering challenge manifesting that is worth reporting is in the calculation of the *Visitors Proactivity Rate* ($\frac{\text{number of reviews}}{\text{number of visitors}}$), as the dataset being used refers to 2020 while the most recent annual number of visitors data refers to 2019. This is a

problem of effort in collecting data and making it available, and we approached the issue by presenting the KPI using the 2019 visitors data instead of the unavailable 2020 one. Considering how Covid-19 negatively affected the visits in 2020, the KPI came out extremely low, as reviews were scarce in 2020 but visitors were not in 2019. The reason why we included it anyway is simply to show the KPI to the stakeholder and let him evaluate if it is worth it to collect the visitors data more often and how often.

6.2.3 Response from the stakeholders

The second prototype has been presented to three representatives of the Museo Archeologico Nazionale di Napoli in the 19th of March 2021. The feedback has been generally positive, the dashboard has been able to convince all the people to whom it's been presented, including the ones that showed some open skepticism towards the tool at first.

The visual elements that draw the most attention were definitely the ones that allowed for a direct comparison between different museums, being the *benchmark dedicated* views and the two funneling graph featured in views *Followers* and *Activity Online* that showed the total number of followers and posts published in 2020 per each museum in the same region (Campania). The attitude towards these elements by the stakeholders revealed a strong sense of competition and willingness to proof their work. The stakeholder focused more on their position compared to the museums that bested them without questioning extensively how the numbers were plotted and which were the metrics, proving that their interest lies in having the best results possible and confirming that the attitude towards this dashboard is similar to the City Dashboard scenario presented in the literature, in which users tend to

accept and take the information in the dashboard at complete face value, while preferring and leaning towards the ones that defend their results. From this feedback it is possible to formulate the hypothesis that comparing institutions that know each other and are close together on a specific metric is a strong push for the decision maker to act in order to improve the specific performance and see his institution climbing the ranks in the dashboard visualization to validate its work.

Most of the questions received by the professionals from the National Archeological Museum of Naples revolved around the possibility to add or change functionalities in the dashboard, which most of the time resulted in an affirmative answer as Power Bi allows a great degree of customization. However, one question in particular revealed an attitude towards the dashboard that we didn't take into consideration during our research: the stakeholder asked if it was possible to automatically generate a PDF containing pre-set comparisons between institutions. This question completely disregarded the flexibility of dynamic filtering, asking instead for fixed pre-set filtering as setting up the comparisons themselves across different views would just result in extra work. On top of that, they specifically requested a PDF file, which again completely disregard Power Bi's ability to visualize its content on different devices, online, at any time, and with updated data. The intention is to build a backlog of PDF files at different dates, completely disregarding the filtering by date functions of Power Bi. This indicates either a resistance in using unconventional tools or/and a failure from our part in presenting these tools properly and making them understandable.

A second presentation has been held on the 29th of March 2021 to the Direzione Generale Musei. The presentation had a totally different focus compared to the first one. It has been

more centered around the tool rather than the results and the single views. The greatest part of the presentation was dedicated to how the tool can evolve and fit into the strategy of the stakeholder while integrating more data. During the presentation, an example of data interpretation has been offered, comparing two different communication strategies from two different museums as well as analyzing and interpreting data, to understand why two Italian regions (Tuscany and Sardinia) presented opposite performances in their engagement to post ratio. From the presentation's feedback emerged some very important aspects regarding how the dashboard, as a tool, can be viewed and employed by a regulator, and what benefits it offers towards its role and the organization itself. Being the regulator, the General Direction provided a substantially different point of view compared to the National Archeological Museum of Naples. The presentation was an overall success, and was able to spark enthusiasm and interest in the professional we contacted as the dashboard can help her in her daily job and make her more relevant in the overall strategy of her office and ministry, as it enhances the visibility of what she does. It's a system that allows itself to be the object of interest from a higher level because of its features, as well as from a lower level, where the characteristics of the tool regarding flexibility and easiness to use are appreciated by who effectively manage the data.

For the regulator, the tool is also extremely valuable on an organizational level, as it can potentially overcome difficulties regarding data ownership. In our specific scenario, many important museums in Italy fly completely off the radar of the General Direction as, for a politically driven decision, they are not supervised by the ministry, but rather from local regulators. A tool such as the dashboard can be used to prompt a commitment from all

parties to gather and standardize data, to be integrated in a common tool that can be shared and adds value to everyone.

As per the dashboard itself, the regulator appreciated the ability to compare two different museums at the same time. The most criticized aspect of the dashboard has been the museum's division between size, as the size of a museum is a matter of discussion that can generate antipathy from the institutions involved, who in the past have proven to be sensible to the matter.

Both stakeholders have shown a limited understanding of the underlying shortcomings and issues of the dashboard during the brief presentation, as anticipated by the literature review. The professionals from both the Ministry and the National Archeological Museum of Naples showed no interest whatsoever in the metrics and methodologies used to collect the data and display it. They took the data completely at face value, not questioning their veracity, their quality, their accuracy and fidelity or how consistent and reliable they are. This is the effect of an instrument that tends to hide the burdens of its implementation and gives the impression that everything can be done with it.

Finally, for the purpose of pitching a dashboard, it is important to underline how despite the interest towards building one was generated from the first proof of concept as well, using real data proved to be more effective in prompting an active response to harvest more data and proceed with extending both the institutions subject of the investigation as well as the relationship that the data describe, such as the number of visitors and the quality indicators that are at the base of the framework presented in Chapter 4.

7 **Conclusions**

This thesis had been developed with the intention of building and proposing a new comprehensive business intelligence tool targeted towards cultural institutions, specifically Italian museums, that can capture the dynamics of the segment and provide real-time outputs and powerful insights by integrating conventional and unconventional data. The study of the literature revealed some gaps:

- the KPIs defined in the classic frameworks as seen in chapter 3.1 do not fit the purpose of measuring museums performances and aren't able to capture the dynamics of this particular industry
- recent dashboard studies have been conducted focusing more over traditional businesses, and not satisfyingly enough over cultural institution, which subsequently means that the value added potentially provided to museums by this tool has not been investigated properly
- no studies have investigated the possibility to integrate unconventional sources of data (eg. coming from social media) with conventional type of data (eg. coming from operations, finance, etc.) for museums performances monitoring
- the already established KPIs that rely on social media data, built to monitor businesses reputation, focus on performances relating a single social media platform, keeping the user from getting insights from a holistic view of the overall and transversal reputation in the web and the sentiments that users get from it

The present thesis focused on working to overcome these gaps and has the merit of having delivered a working final product to a stakeholder, with low capabilities in developing monitoring tools and low analytical skills, but high influence in the system, and spark interest in further developing the tool.

The thesis contributes to the academic literature by filling the gap of the inexistence of horizontal indicators that capture information and insights from more than one social media. It also bridges the gap created by the lack of studies defining an holistic monitoring model that integrates unconventional and conventional data for museums. The previous gaps were filled by outlining a model featuring 26 KPIs, each one described clearly, with precise metric, formulas, and produced with data that can be harvested from a clear identifiable conventional and unconventional source.

Moreover, the thesis contributes to the literature also by highlighting the process in terms of analysis, implementation challenges, necessities and advantages when transforming a monitoring framework into an interactive dashboard, shifting from a “static” to “dynamic” monitoring tool tailored to the needs of cultural institutions, in particular museums.

The solidity of the instrument is further strengthen by the ability of showing both an acceptable “healthy” range as well as if there are progresses (Ragouzeos, Gandrup, Berrean, Li, Murphy, Trupin, Yazdany and Schmajuk, 2019), thanks to its benchmarking capabilities and iterative periodical data collection.

The main contribution of this thesis is to carry out this vision towards a tangible product. Delivering a working dashboard to a stakeholder accomplished many tasks at once:

- it validated the model presented in Chapter 4 developed to effectively measure museums performances, as well as the processes put in action to generate a working tool starting from the theoretical model.
- it introduced them to the concept of interactive dashboard, demonstrating the latter to be a competent solution to measure museums performances, a critical tool for regulators as dashboard initiatives have a deep normative effect, shape governance, modify institutional behavior, condition workers, influence decision-making and shape spending patterns (Franceschini, Galetto and Maisano, 2007) and are a crucial part of processes that help guide strategic decision-making, acting as a common source of information and learning tool (Kitchin, Lauriault and McArdle, 2015).
- it highlights the possibilities that modern tools offer in producing a working, useful dashboard with limited resources regarding time and knowledge and overcomes the challenges that arise from the implementation of such as solution. As expected, it is possible to set up a dashboard initiative as a flexible process, allowing to make changes and adapt the dashboard to the needs of the users very quickly. In our experience, as explained in Chapter 6.2, refitting the same design for two different stakeholders has proven to be rather easy.

The feedback received from the professionals that witnessed our presentation confirms that the research successfully achieved all these objectives. Delivering the dashboard has proven to be a success, convincing both skeptical professional as well as the ones that demonstrated interest in the idea.

The contribution that this work has given to the research comes from different sides:

- the first achievement, as previously mentioned, was to provide the dashboard, which was considered by the stakeholders involved valuable and appropriate for their knowledge and processes.
- the second achievement is the one in the interest of further application for other cultural institutions, as the model presented in Chapter 4 can be potentially employed by any cultural institution. The reach of the model goes beyond single institution, as it was prepared bearing in mind the needs of a big regulator such as the Ministry itself. For this reason, the model can be employed to set up and monitor the achievement of long-term goals regarding a considerable number of cultural institutions. Cultural institutions are a particular activity of which the success can't be easily evaluated in terms of financial achievement and of which missions are vast in scope and far reaching in target.
- the third achievement is of interest of further application for any business or individual or regulator alike, that wished to build a dashboard in Power Bi. Most dashboards fail due to designers focusing on making them aesthetically appealing more than functionally effective (Alhamadi, 2020). The study of the literature and the process of creating the tool has included the observation of many real life application of a dashboard, which this thesis meticulously analyzed, laying down the best practices in dashboard design in Chapter 5.1, as well as logical steps that need to be taken into consideration to create the tool itself. The thesis essentially includes, on top of a model for cultural institution, a method on how to turn any model into a working dashboard. All of these steps have already been validated by the

stakeholders that approved on the dashboard resulted from the application of these indications.

To sum up, this thesis has proven that business intelligence tools can meet the interest of regulators and businesses alike, that non-conventional data such as social media data can be mixed together with conventional data to harvest meaningful information and offers all the tool needed to measure the performances of cultural institutions, including a model and a method.

7.1 Limitations and further research

This research has faced some limitations that open up possibilities for further studies. These limitations presented themselves in every stage of the research and are mostly related to our stakeholders.

The chapter is dedicated to highlighting these limitations and explaining what kind of further effort is needed to move the research forward towards a more refined and up to date result, as well as how this thesis navigated around them and possible suggestions to how to remove them.

The purpose of this research was to study and effective way to measure museums performances through the implementation of business intelligence tool that relied on both conventional and unconventional data, and the stakeholders that have been engaged to verify the solution have been chosen from the beginning. Despite the fact that the model has been built to be of great value for regulators and single institutions alike all around the world,

we developed a bias towards the stakeholder, and the model has been inevitably influenced by the needs of the Italian Ministry of Cultural Heritage and Activities, which resulted in giving a privileged focus to studies relating Italian cultural institutions and public documentation released by the Italian regulator itself. The model would ideally need further rework to be successfully tested and validated by international institutions. However, we believe that it's still represents a strong foundation as the research relied also on international papers, resulting into the inclusion of values shared among the international community.

The stakeholder we worked with is not easily approachable, and this was only worsened by the fact that they are very large institutions dealing with the worst pandemic of recent history. For this reason, it's been impossible to harvest meaningful datasets nor insight on its internal processes in a timely manner. As laid out in Chapter 6.2.2, creating a complete dataset requires a continuous effort from the stakeholder that we simply couldn't ask for. As a result, our first prototype was based on dummy data, and the second one relied on datasets collected in a very limited environment (Social Media) and for a different purpose. It's been impossible to investigate the *daily* or *weekly* needs of the regulator or single institution in order to craft a dashboard that would be taken and examined very frequently. The final goal of the validation ended up leaning more towards selling the concept of an interactive dashboard to the stakeholder, opening up the possibilities to future, more engaged collaboration. As for the model and its impressive array of 26 KPIs, it as well suffered from the same limitation, as these KPIs really lean towards monitoring long term goals despite presenting themselves in a form that incentivize continuous small improvements that can be detected on a timely basis.

Social Media are an everchanging industry and different trends needs to be taken into account when deciding upon what indicators and factors matter and what not. As this research relied on previously collected datasets regarding social media activity, the practical focus shifted towards working with that dataset rather than expanding it to harvest information that would be more relevant for the final user. For the future, however, it is necessary to take into consideration changes in trend like the growing insignificance of the Twitter platform (Stelzner, 2020), the rise of new Social Media platforms as well as the growing importance of new content types such as livestreams or *stories*, of which data have not been collected. Ignoring these aspects somewhat benefited our research as the second prototype contained data and information that the stakeholder was familiar with, easing the introduction of the new business tool concept, but it is where the study over this topic should pick up next in order to further increase the contribution that non-conventional data such as social media can give to the decision maker.

8 Bibliography

Aichner, T. and Jacob, F., 2015. Measuring the Degree of Corporate Social Media Use. *International Journal of Market Research*, 57(2), pp.257-276.

Alhamadi, M., 2020. Challenges, Strategies and Adaptations on Interactive Dashboards. *Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization*.

Alpaydin, E., 2014. *Introduction To Machine Learning*. Cambridge: MIT Press.

Antoniou, C., Dimitriou, L. and Pereira, F., 2018. *Mobility Patterns, Big Data And Transport Analytics*. 1st ed.

Appel, G., Grewal, L., Hadi, R. and Stephen, A., 2019. *The future of social media in marketing*. *Journal of the Academy of Marketing Science*, 48(1), pp.79-95.

Arnaboldi, M., Azzone, G. and Giorgino, M., 2015. *Performance Measurement And Management For Engineers*.

Basso, A., Casarin, F. and Funari, S., 2018. How well is the museum performing? A joint use of DEA and BSC to measure the performance of museums. *Omega*, 81, pp.67-84.

Berthon, P., Pitt, L., McCarthy, I. and Kates, S., 2007. When customers get clever: Managerial approaches to dealing with creative consumers. *Business Horizons*, 50(1), pp.39-47.

Birdas, T. J., Rozycki, G. F., Dunnington, G. L., Stevens, L., Liali, V., & Schmidt, C. M., 2019. "Show Me the Data": A Recipe for Quality Improvement Success in an Academic Surgical Department. *Journal of the American College of Surgeons*, 228(4), 368–373.

Brook, C., 2020. *What Is Data Integrity? Definition, Best Practices & More*. [online] Digital Guardian. Available at: <https://digitalguardian.com/blog/what-data-integrity-data-protection-101>.

Burgess, J., Marwick, A. and Poell, T., 2019. *The Sage Handbook Of Social Media*. London: Sage Publications.

Casaleggio Associati, 2020. *E-commerce in Italia nel 2020*.

Chen, P., 1976. The entity-relationship model — toward a unified view of data. *ACM Transactions on Database Systems*, 1(1), pp.9-36.

Cormode, G. and Krishnamurthy, B., 2008. *Key differences between Web 1.0 and Web 2.0*. *First Monday*.

Craglia, M., Leontidou, L., Nuvolati, G. and Schweikart, J., 2004. Towards the Development of Quality of Life Indicators in the 'Digital' City. *Environment and Planning B: Planning and Design*, 31(1), pp.51-64.

Cukier, K., 2010. *Data, Data Everywhere*. [online] The Economist. Available at: <https://www.economist.com/special-report/2010/02/27/data-data-everywhere>

Daft, R. and Lengel, R., 1986. Organizational Information Requirements, Media Richness and Structural Design. *Management Science*, 32(5), pp.554-571.

Dao, D., 2015. Social Media Classification Scheme in Online Teaching and Learning Activities: A Consideration for Educators. *International Journal of Education and Social Science*, 2(4), pp.85-94.

del Barrio, M. and Herrero, L., 2013. Evaluating the efficiency of museums using multiple outputs: evidence from a regional system of museums in Spain. *International Journal of Cultural Policy*, 20(2), pp.221-238.

Eckerson, W., 2013. *Performance Dashboards: Measuring, Monitoring and Managing Your Business*. Hoboken, N.J.: Wiley.

Egloff, C. and Zorloni, A., 2012. *Art & Business: Measuring A Museum'S Performance*. [online] BCG. Available at: https://www.bcg.com/publications/2012/public_sector_arts_culture_art_and_business_measuring_museums_performance.

Farmanbar, M. and Rong, C., 2020. Triangulum City Dashboard: An Interactive Data Analytic Platform for Visualizing Smart City Performance. *Processes*, 8(2), p.250.

Franceschini, F., Galetto, M., & Maisano, D., 2007. *Management by measurement: Designing key indicators and performance measurement systems*. Berlin: Springer.

Frunza, M., 2016. *Solving Modern Crime In Financial Markets*. Waltham, MA: Academic Press, an imprint of Elsevier.

Gandomi, A. and Haider, M., 2015. *Beyond the hype: Big data concepts, methods, and analytics*. *International Journal of Information Management*, 35(2), pp.137-144.

Harvard Health Publishing, 2020. *Blue light has a dark side*. [online] Available at: <https://www.health.harvard.edu/staying-healthy/blue-light-has-a-dark-side>.

Hodgson, M., 2016. *A Decentralized Web Would Give Power Back To The People Online*. [online] Techcrunch. Available at: <https://techcrunch.com/2016/10/09/a-decentralized-web-would-give-power-back-to-the-people-online>.

Kabari, L. and Osuo-Genseleke, M., 2018. A Comparative Study on Internet Generations. *International Research Journal of Engineering and Technology (IRJET)*, 05(05), pp.3760-3766.

Kaplan, A. and Haenlein, M., 2011. *The early bird catches the news: Nine things you should know about micro-blogging*. *Business Horizons*, 54(2), pp.105-113.

Kaplan, A. and Haenlein, M., 2010. *Users of the world, unite! The challenges and opportunities of Social Media*. *Business Horizons*, 53(1), pp.59-68.

Kaplan, R. and Norton, D., 1992. The Balanced Scorecard - Measures That Drive Performance. *Harvard Business Review*, pp.71-79.

Kietzmann, J., Hermkens, K., McCarthy, I. and Silvestre, B., 2011. *Social media? Get serious! Understanding the functional building blocks of social media*. *Business Horizons*, 54(3), pp.241-251.

Kitchin, R., Lauriault, T. and McArdle, G., 2015. Knowing and governing cities through urban indicators, city benchmarking and real-time dashboards. *Regional Studies, Regional Science*, 2(1), pp.6-28.

Krumm, J., Davies, N. and Narayanaswami, C., 2008. *User-Generated Content*. IEEE Pervasive Computing, 7(4), pp.10-11.

Labrinidis, A. and Jagadish, H., 2012. *Challenges and opportunities with big data*. *Proceedings of the VLDB Endowment*, 5(12), pp.2032-2033.

Laney, D., 2001. *3D Data Management: Controlling Data Volume, Velocity, And Variety*, META Group.

Liben-Nowell, D. and Kleinberg, J., 2003. The link prediction problem for social networks. *Proceedings of the twelfth international conference on Information and knowledge management - CIKM '03*.

Murray, E., 2019. *The importance of color in data visualization*. [online] Forbes. Available at: <https://www.forbes.com/sites/evamurray/2019/03/22/the-importance-of-color-in-data-visualizations/?sh=7f9520f157ec>.

Pal, S., 2012. Web 2.0: A New Tool For Teaching and Learning in Electronic Environment. *Journal of the young librarians association*, 05, pp.107-113.

Peters, B., 2017. *Square Vs. Landscape Video—\$1.5K Worth Of Experiments: Here's How They Compare*. [online] Buffer (blog). Available at: <https://blog.bufferapp.com/square-video-vs-landscape-video>.

Porter, T. M., 1995. *Trust in numbers: The pursuit of objectivity in science and public life*. Princeton, NJ: Princeton University Press.

Ragouzeos, D., Gandrup, J., Berrean, B., Li, J., Murphy, M., Trupin, L., Yazdany, J., & Schmajuk, G., 2019. "Am I OK?" using human centered design to empower rheumatoid arthritis patients through patient reported outcomes. *Patient Education and Counseling*, 102(3), 503–510.

Salesforce. *Filtering a Dashboard*. [online] Available at: https://help.salesforce.com/articleView?id=sf.dashboard_filters_overview.htm&type=5
[Accessed 4 December 2020].

Short, J., Christie, B. and Williams, E., 1976. *The Social Psychology Of Telecommunications*. Ann Arbor, Mich.: UMI.

Speckbacher, G., 2003. The economics of performance management in nonprofit organizations. *Nonprofit Management and Leadership*, 13(3), pp.267-281.

Stelzner, M., 2020. *2020 Social Media Marketing Industry Report*.

Clement, J., 2020. *Global digital population as of October 2020*. [online] Statista. Available at: <https://www.statista.com/statistics/617136/digital-population-worldwide/>.

Tableau. 2020. *Best Practices For Effective Dashboards*. [online] Available at: https://help.tableau.com/current/pro/desktop/en-us/dashboards_best_practices.htm
[Accessed 4 December 2020].

Turban, E., Cameron Fisher, J. and Altman, S., 1988. Decision Support System in Academic Administration. *Journal of Educational Administration*, 26(1), pp.97-113.

Viswanathan, G., Mathur, P. and Yammiyavar, P., 2010. *From Web 1.0 To Web 2.0 And Beyond: Reviewing Usability Heuristic Criteria Taking Music Sites As Case Studies*. Indian Institute of Technology Guwahati.

Weijers, M., Bastiaenen, C., Feron, F., & Schröder, K., 2021. Designing a Personalized Health Dashboard: Interdisciplinary and Participatory Approach. *JMIR Formative Research*, 5(2), e24061.

World Wide Web Consortium. *HTML & CSS*. [online] Available at: <https://www.w3.org/standards/webdesign/htmlcss#whatcss>.

Zorloni, A., 2010. Managing performance indicators in visual art museums. *Museum Management and Curatorship*, 25(2), pp.167-180.

9 Annex

9.1 Guidelines on how to proper use Social Media in business

Next, we will provide a set of guidelines inspired on the Andreas M. Kaplan and Michael Haenlein analysis splitting the advice into two sections, four points about using social media and four points about being social.

9.1.1 Four Points about using social media

Choose carefully, there plenty of Social Media applications, and new ones are appearing on the horizon every day. If you still need time to run your core business, you simply cannot participate in them all, especially since “being active” is one key re- quirement of success. Choosing the right medium for any given purpose depends on the target group to be reached and the message to be communicated. On the one hand, each Social Media application usually attracts a certain group of people and firms should be active wherever their customers are present (Kaplan and Haenlein, 2011)

1. **Pick the application, or make your own**, once you know which game you’re playing, the next decision involves understanding of the basic idea behind Social Media. It’s all about participation, sharing, and collaboration, rather than straightforward advertising (Kaplan and Haenlein, 2011).
2. **Ensure activity alignment**, if you decide to rely on various Social Media, or a set of different applications within the same group, in order to have the largest possible

reach, it is crucial to ensure that your Social Media activities are all aligned with each other.

3. **Media plan integration**, What is true for different types of Social Media also holds for the relationship between Social Media and traditional media: Integration is key! While you may consider these two arenas to be completely different, in customers eyes they are both part of the same: your corporate or cultural institution image (Kaplan and Haenlein, 2011).

Four points about being social

1. **Be active**, If you want to develop a relationship with someone, it is always advisable to take the lead and to be active. Social Media are all about sharing and interaction, so ensure that your content is always fresh and that you engage in discussions with your customers (Kaplan and Haenlein, 2011).
2. **Be interesting**, if you would like your customers to engage with you, you need to give them a reason for doing so. The first step is to listen to your customers. Find out what they would like to hear; what they would like to talk about; what they might find interesting, enjoyable, and valuable. Then, develop and post content that fits those expectations (Kaplan and Haenlein, 2011).
3. **Be humble**, before you enter any application, first take some time to discover it and to learn about its history and basic rules. Only once you have gained the necessary understanding, start to participate (Kaplan and Haenlein, 2011). Otherwise, many readers can perceived account as a fake.

4. **Be unprofessional** firms, companies or cultural institutions would be wise to avoid overly-professional content offerings but instead, try to blend in with other users

9.2 Deep Dive into main user-generated content platforms used by cultural institutions: Facebook, Twitter, Instagram, Google Maps and Trip Advisor

Facebook

Basic information

Facebook is a social network founded in 2004. Users can add friends to have their activities featured in their feed. Users can interact with a user post by liking it, sharing it or commenting on it. There are two entities: pages and profiles. Pages are generally made for businesses, public personalities or institutions, and allow the owner to feature personalized call to action. Facebook also features third party apps and tools for instant private messaging, fundraising and more.

Facebook's demographic and use available relevant data

- Facebook is the biggest social network on the planet;
- Facebook host over 1.59 billion active daily users;
- 93% of marketers use the platform daily;

- 75% of all internet users with an annual income higher than USD 75,000 possess an account;
- Facebook presents a vastly diverse userbase in terms of age: out of all the seniors aged over 65 that use internet, 62% have an account.

Table 3 - Facebook's demographic

Age of people online	% possessing a Facebook Account
18 - 29	88%
30 - 49	84%
50 - 64	72%
+65	62%

How does it work

Posts on Facebook can include many media including a personal video conference call, tags, locations, live, fundraising, gif, videos or pictures, activities or *moods*. These are written from the homepage or the profile page of the user and shared to all the user's friend.

Comments on Facebook are displayed under a post. People can make comments of other comments but cannot go beyond the first level. They are sorted based on the number of interactions, but also based on variables specific of the user, such as if they are written by a friend.

Followers on Facebook are of two types. People can follow *pages*, but the relationship between two private user profile is bilateral and called *friendship*. Personal profiles cannot have more than 5000 friends.

Chat on Facebook is available to users and it is called Messenger. Users can decide between private instant messaging or videocall. On mobile, the service requires a separate app. Users can send text, photos and videos, stickers, GIFs, emoji, money, and they can *like* any message. The app supports group chats. Users can send messages to anyone regardless of the friend status or privacy settings. The receiving user won't receive any notification of messages from unknown people, and these messages are filtered and displayed in a secondary tab.

Facebook supports **hashtags** and **tags**. To tag someone, the user needs to type "@" followed by the person's or business's name. Tags can be used in both comments and posts.

Pages

Facebook has an extremely variegated offering of services built on top of the core social media functions. The main value for businesses comes from the fact that Facebook has a dedicated profile type, called *page*, for institutions, businesses, causes, brands or high public profile, which allows to deliver information about the business such as description, location and contacts directly to anyone. Unlike personal profiles, pages do not relay on the *friendship* system. Instead, people that *like* the page become fan and the page's contents are displayed in their feed. A page can have an unlimited number of fans, while a personal profile cannot exceed 5,000 friends.

A Facebook page is free of charge and takes only few minutes to set up. It offers many different functions, making it extremely versatile.

Connecting with people

To connect with people, pages features **posts** and **stories**, as well as an inbox that gathers together comments, private messages sent through Messenger or even through Instagram's dedicated messaging tool, if the accounts are connected. This is key to interact with potential customers and provide them assistance. Through pages it is also possible to create and manage groups, which are a dedicated space to communicate about a specific product or service or topic.

Business tools

Pages feature business tools such as appointments, to let customer book an appointment directly on the page. Pages can also organize events and promote them through the page, followers can make everyone aware of the fact that they will participate, and can buy tickets directly from Facebook. Jobs is a tool that allows to reach, engage and hire candidates through the management of job posting. Shops is a tool that allows to share inventory or services and integrates a buying process to sell directly to customers.

On top of these features, pages feature personalized call to actions and, for hotels or places, user generated reviews and scoring system. These features can be turned on and off from the Facebook page directly.

Twitter

Basic information

Twitter is a social network founded in 2006. Users can communicate through the exchange of tweets, which may contain photos, videos, links, and text. These messages are posted on user's profile, sent to followers, and are searchable on Twitter search. Users can interact with tweets by posting a response tweet, liking the tweet or retweeting.

Twitter's demographic and use data

Twitter is a declining social media: according to the *2020 Social Media Marketing Industry Report*, its relevance has been decreasing in the past years in favor of Instagram.

- according to Twitter itself, round 80% of its users are affluent millennials;
- Twitter is most popular in the United States, Japan, Russia and United Kingdom;
- Twitter is mainly used as a source of news and trends and is able to generate around a ~31% stronger emotional response and a ~28% higher memorability compared to the average social media post;
- the average visitor spends 12 minutes and 54 seconds every day on the website;
- Twitter is the 54th most engaging website worldwide according to Alexa ranking and is mainly used in the U.S;
- 42% of Twitter users access on the platform every day.

How does it work

Users interaction revolves around **posts**, that are called *tweet*. They can contain a picture or a video, and a short text (280 character maximum).

Tweets can receive **comments**, which are just other *tweets*. When visiting a user profile, the tweet used as comments are displayed in a separate tab, along with the tweet they are responding to. To display a tweet alongside its comments, the tweet needs to be clicked on.

Twitter supports instant messaging through the platform itself. This is a non-public conversation that can be initiated only with someone that follows the account starting the conversation. This rule can be ignored if a previous conversation has been already started in the past between the two parts, or if the receiving user opter to receive direct messages from anyone. Twitter also supports group chats. Businesses can receive messages from anyone. Users can block accounts to prevent having any kind of private conversation with them, both one-to-one and public.

On Twitter it is possible to **share** other user's tweets through a function called *retweet*. The retweeted tweet is going to be displayed on the user's personal profile and to its followers' feed.

Stories are not present on Twitter as of September 2020.

Instagram

Basic Information

Instagram is an American social networking service originally launched on October 2010 on iOS. It is now owned by Facebook, and supports iOS, Android and Fire OS. Its original and core purpose is to share pictures or short videos online, with the possibility to enhance them through filters.

Instagram's demographic and use available relevant data

- Instagram is the only social media in which female users are the majority;
- than one billion monthly active users;
- 65% of global Instagram audience is composed by user aged between 18 and 34;
- more than 200+ million "instagrammers" (Instagram users) visit a business profile every day;
- around 1/3 of the most viewed **stories** on Instagram are made by businesses;
- more than 60% of users claim to use Instagram to discover new products.

How does it work

Instagram allows its users to create **posts** that feature either a picture or a video. The original premise of Instagram was to post square (1x1 aspect ratio) pictures, but the system now allows for more vertical and horizontal layouts. Videos cannot be longer than one minute. The posts can also feature more than one picture or video displayed through a *carousel*. Users can navigate the different pictures by swiping on the touchscreen. Every post can feature a description, hashtags and the location in which the picture was taken. The

Users can **comment** a post and answer to comments creating a sub-discussion. Both posts and comments support the use of tags, and in comments they are used as a mean to address someone with a response.

Instagram features a **messaging** and videocall system called Instagram direct. On direct people can also send pictures with limited view time. Instagram direct is also a mean through

which people can react and respond to someone else Instagram story. It also support audio messages and GIFs.

User can only **share** someone else's story if they are tagged in it. Users can share someone else's post by featuring it on their story or sending it to someone else through direct. People can also share Instagram stories through their own stories if they are tagged on the original one. It is however not uncommon to see people acquiring a screen-capture of the screen to share a story in which they are not tagged.

User can leave a **like** on posts, comments and messages on direct. Throughout 2019, Instagram began testing and then maintained a change to its service to hide the like counter on posts from everyone with the exception of the post's owner. This change contrasts practices such as *like-begging*, making *likes* a more reliable proxy of user engagement.

Stories

Instagram has been the first social network to reach over 500 million stories made in a single day, at the beginning of 2020. The feature, pioneered by the competitor Snapchat, is today one of the main business tool for companies. On top of paid advertising, which allows to feature an ad among the feed of the users, stories offer tools to give traction to a product or a cause. Facebook, Instagram's owner company, provides a free software to develop and create easily augmented reality face-tracking effects, that are applied in real time on people's faces. Anyone can create one, and companies can sponsor their products or services in creative ways by literally branding the users. Other ways to use stories to grow connection to the users is to use tool such as questionnaire directly inside the story, tagging and reposting.

Google

Basic information

Google is a provider of many different services, most notably its search engine. Its inclusion in this deep dive analysis stems from the fact that, like any other social network company, Google's core business is advertisement based on collected user data, and any effective digital marketing strategy must consider it. Google is a company founded in 1998 and it is considered one of the *Big Four* leading US tech companies.

Relevant services

- *Google Maps*, is a web-mapping service. It features satellite imagery, aerial photography, street maps, 360° interactive panoramic views of streets, real-time traffic conditions and route planning for traveling by foot, car, bicycle and air or public transportation
- *Google Search Engine*, is the core service of the company. As many other search engines, it includes a repository in which company representative can add, for free, their key information, such as logo, name, contacts, business hours, address, pictures and website. These information are displayed as search results of any related search query, in a dedicated box

These two services are combined together so that Google Maps is able to display on the maps all of the businesses that register on Google's repository. This enables services such as search based on location or proximity. An example of this are search queries such as:

Cupcake store near me or Cupcake store near Bovisa

Users can directly or indirectly provide additional information such as the affluency to a determined place at any time and day, provide user-made pictures of the business or its products and services, and provide reviews.

Reviews on Google are made up of three components:

- text;
- pictures;
- rating (one to five stars; mandatory).

Other metadata include the time of the review, the user and the location from which the review is written. Google also implement a feature to allow user to filter reviews based on specific keywords.

Google's demographic and use available relevant data

Google is consistently ranked #1 in Alexa Ranking, making it the most popular website of the planet. This popularity forged the verb *google*, which means *to search on the internet for information about (the object)*. Google's popularity is also to be attributed to Chrome, a browser application that holds about 69% of global market share as of June 2020, and that features Google as its default search engine. Most notably, according to Forbes, Google pays Apple USD 7 billion annually to be the default search engine on iOS, on top of the 2.5 billion active Android devices using Google services, which makes up most of the Android-based devices in existence. Google resulting search engine market share is of 90.46%, it is virtually known and used by every person that have access to the internet. People make 63,000

searches every second. Google's services are many and all accessible from a single account. To access the mailing service Gmail alone, on 2018 there were 1.5 billion active accounts.

How does it work

Google's search engine or Google Maps will display on screen businesses relevant to the area and interest after users try to search for something. When more than one business is relevant to the search, the search engine will display a map indicating where the three most relevant businesses are and some information about them. It is also possible to visualize more. One of the information provided is the average rating of the reviews and the number of reviews. Any user add a review. Reviews are public and not anonymous.

Trip Advisor

Basic information

Tripadvisor is the world largest travel website with user generated content. It was founded in February 2000. Users can look for tourism-related businesses listing such as hotels, transportation, lodging, travel experiences and restaurants, and leave reviews for other people to see.

Tripadvisor's demographic and use available relevant data

- 463 million monthly unique visitors (as of September 2020);
- 867 million reviews (as of September 2020);

- 8.7 million accommodation, airlines, cruises, experiences and restaurant (as of September 2020);

How it works

Tripadvisor is a service accessible through website or mobile app. It is not a social network, and it doesn't require any subscription for accessing its basic functionalities. The website hosts a number of accommodation solutions, airlines, cruises, experiences and restaurants. Users can look for them specifically or discover them based on location and interests. Listing on Tripadvisor can happen in two ways: either a representative of the business requires a listing, or a user writes a review, initiating a listing. Businesses listing initiated by user reviews can be later claimed by the legitimate owner. Once an owner claim a business, he has to provide specific information depending on the category (accommodation, restaurant or attraction). Example: a restaurant will have to specify the category of restaurant (fast food, café, etc.), the pricing, the cuisine type and eventual special features. A user visiting a listing can see prices, availability, details and description of the listing, and book directly from Tripadvisor. On top of that, there are the user generated content: the **reviews**. Reviews on trip advisor includes six elements:

- title
- score (from one to five)
- text
- pictures (optional)
- metadata (such as when the review was written or where is the reviewer from)
- information about the travel (type and period of the year)

Reviews are displayed in the user language automatically thanks to Google Translate. Reviews can be filtered based on the season in which they were written, the type of travel (solo, couple, with family, with friends or business). The owner of the listing can respond to any review, while users can give a **like** to the review, an action called *useful*, to indicate a review worth reading. To write reviews, people need to create a user a **profile** that will display his nickname, the reviews and recent activities, along with other information such as the home country, personal picture, website and small biography. To stimulate user to share reviews and pictures, Tripadvisor added badges to unlock by reaching specific achievement such as *posting 3 pictures* or *getting useful votes* to the written reviews, while increasing *level* to become a better contributor.

9.3 Social media Reputation Dashboard

As of November 2020, the Museums General Direction office of the Ministry of Cultural Heritage and Activities tracks the online presence and reputation of 100 selected museums in the Italian territory through the aforementioned Social Media Reputation dashboard. This represent a strong starting point for analysis as it is focused on data extracted from user generated content platforms, which are an important element of this study. By observing how the social media data source is being used in this dashboard today it is possible to evaluate the elements that work best, assess what is still missing and ultimately deliver a degree of innovation and usefulness that is grounded in reality and that reflects the immediate needs and interests of the regulator. What follow is an analysis of the current

state of art of the Social media Reputation dashboard, including its strength, weaknesses and gaps.

The dashboard has been developed on two layers: one layer is accessible exclusively to the MIBACT and shows aggregate data of all of the museums in Italy. A second layer is provided to each museum, and it contains data for that museum alone. Each view includes information displayed through histograms, pie charts and absolute numbers covering of the following outputs:

- *Total number of contents* for each social media platform considered (Twitter, Instagram, Facebook), average per account, maximum and minimum in a single account;
- *Total number of public museums accounts* for each social media platform considered (Twitter, Instagram, Facebook) in Italy;
- *Total number of news* and its evolution over time;
- *Average number of interactions and its evolution over time*. Interactions are taken into account separately and include Instagram, Facebook and Twitter likes, Instagram and Facebook comments, Facebook shares and Twitter retweets;
- *Average number of followers* per account on Instagram, Twitter and Facebook;
- Average of total engagement for Instagram, Facebook and Twitter =
$$\frac{\text{Interactions}}{\text{Number of followers}};$$
- Popularity on each platform (Instagram, Facebook and Twitter) calculated as
$$\frac{\text{Average number of likes}}{\text{Number of followers}};$$

- Virality on Facebook and Twitter calculated as $\frac{\text{Average number of shares}}{\text{Number of followers}}$;
- Commitment on Facebook and Instagram calculated as $\frac{\text{Average number of comments}}{\text{Number of followers}}$;
- *Total number of reviews* on Facebook, Google Maps and Tripadvisor and its increment over time, filtered by sentiment;
- Number of reviews per language;
- *Most common hashtag and tags in posted content* visualized as a cloud of words (exclusive to single museum view);

Literature analysis established the utility of a dashboard as dependent from the user ability to:

- obtain a quick snapshot of aggregate data, which would take longer to look at separately;
- obtain insightful and meaningful information which can enable well defended actions to improve tangible aspects that are then reflected in the KPI's presented in the dashboard.

Regarding the first point, the dashboard currently available properly shows data collected across different user generated content platforms through intuitive infographics, making possible to have a snapshot of the situation in different platforms in a glance. Nevertheless, some points of the dashboard could be improved as important data and information that reflecting museum reputation in social media is missing.

Regarding the second point, many elements of the dashboard are presented in a way that do not provide any meaningful information regarding the relationship of the raw number to any

action or initiative taken by the social media manager, the museum or the ministry as a whole.

By applying some of the practices outlined in the literature review regarding dashboard design and utility, we theorized some improvements that could be implemented by leveraging the dataset currently collected.

1. *Percentage change compared to previous period T* - regarding the absolute values presented in the dashboard (such as number of contents published, number of news, number of reviews, average number of post, likes, etc.) it would be useful to show the *percentual increase* compared to a previous period of time T chosen by the user for each indicator. This can be applied to both dashboard layers (at a narrower level, so for each museum, and to the MIBAC exclusive view) and it would provide to the dashboard user the ability to establish if the actions taken had a measurable effect on the metrics. It is a very intuitive way to visualize this data and it doesn't require any more space as it can be integrated in the area already allocated to the absolute data (Figure 25).



Figure 25 - Mock up of percentage change implementation

2. *Benchmarking* - in the current dashboard, single museums have no access to the metrics of their peers. Including the average performance of all museums in specific metric, both in terms of *absolute numbers* and *increase compared to the previous time T*, enables to user to measure both his current status compared to other museums

and the effectiveness of his actions netting exogenous variables. This addition can be pushed even further with a categorization of museums based on kind, location and/or popularity. Just like the previous point, benchmarking is also a great addition to the dashboard as it doesn't add any additional space requirement.

3. *Interactions counting soundness* - the relative KPI *average total engagement* calculated

as $\frac{\text{Interactions}}{\text{Number of followers}}$ doesn't take into account the different relevance that those

interactions have. When performing the sum of the so-called *interactions* (likes, comments, etc.) at the numerator of the indicator, it is possible to improve the meaningfulness of the information by establishing an informed method to weight the different type of interactions. A spontaneous like in Facebook is less an interaction respect to a comment written by a user and again it is less an interaction respect to the action of sharing a post, as the level of commitment and involvement of the user changes and this aspect cannot be ignored. As a reference, Thomas Aichner and Frank Jacob model, presented in Chapter 3.2.4 of the literature review, defines a conversion rate is defined as follows:

- a. *positive comments* in every Social Media show a higher degree of identification so are weighted 5 times the weight of *likes*;
- b. *shares* or *retweets* are weighted 10 times the weight of *likes* because they have an important multiplier effect.

4. *The geographical distribution of reviewers* – it's a data that is not linked to any relevant metrics. In particular, it would be insightful to understand from where the negative reviews are coming from. The current dashboard filters the reviews by language. However, for single Museums, it would be more insightful to understand if their

supporters and/or haters are tourists or locals, providing information about the attractiveness of the museum to different kind of visitors. Aggregating these data in the Mibact exclusive view would also provide a quick measure of the sentiment that locals and foreigners have towards Italian museums.

5. *Visual improvements* - improvements to the dashboard can also be merely in terms of visualization:
 - a. *collapsing indicators* - with the objective to reduce as much as possible the space requirements, one improvement to the dashboard would be collapsing different indicators. This is important to improve readability and adaptability of the dashboard on different screen sizes, especially considering smartphone's and tablet's screen size.
 - b. *color coding* - the current dashboard presents data without providing any visual queue about how good the data is. One simple implementation of this concept is to color-code different values depending on their goodness.
6. *Integrated KPIs that provide horizontal insights across different social media platforms*- the current available dashboard presents information of the status of each single social media (Twitter, Facebook and Instagram) separately, and makes no attempt at giving a holistic view of the social media reputation of museums as a whole. As mentioned in the Chapter 9.1, what is true for different types of social media platforms individually, also holds for the relationship between there social media platforms: integration is key. While the social media manager may consider different platforms to be completely different, in customers eyes they are part of the overall corporate or cultural institution image (Kaplan and Haenlein, 2011).

9.4 Microsoft Power Bi

Microsoft Power Bi is a service platform developed by Microsoft for self-service or enterprise business intelligence, that allows to connect and visualize any data. It is mainly composed by a desktop-based interface called *Power Bi Desktop* and cloud-based business intelligence services.

Power BI aims at bridging the gap between data and decision making through three main functions:

- *Data visualization*: through the *Power BI Desktop* component, user can create reports and dashboard and interconnect data to enable drill-down functions. Dashboards are flexible, easy to personalize and modify, adaptable to any device (including mobile phones) and can even feature AI-powered information.
- *Scalability*: Power BI can integrate with most known sources of data, including any Microsoft service (Excel, Access, Azure) and also Salesforce, SharePoint and more.
- *Empowering*: Power BI can be shared throughout the whole organization and anyone can benefit from it to make business decisions.

Power Bi Desktop is a free application installable on a local computer that let the user connect to, transform, and visualize data. The application is used to connect directly to different data sources and combine them into a data model. The data model allows to build sharable visuals collections. The most common use cases for Power Bi Desktop are as follow:

- Connect to data
- Transform and clean data to create a data model

- Create visuals to provide representations
- Create reports on one or more pages
- Share reports through the Power BI service

The strength of Power Bi desktop resides on the fact that it allows to create compelling reports and doing all the aforementioned tasks through an accessible interface. New features are added monthly to the application. Power Bi Desktop supports a vast number of visual representations, including additional downloadable ones. Power Bi also removes completely the need to test the underlying software itself, it just works as intended, and user can rely on an accessible learning curve and support from the community and Microsoft to sort problems out.