Project Acronym: **Me-Mind**

Grant Agreement no: **101008704**

Project Title: **Museum Events - Measuring Impact on local eNvironment with Data analytics**

D2.2 - Guidelines report

Data and Impact. Guidelines on how data helps to understand the impact of the CCIs

Revision: v1.0

Date: 20/12/2021

Authors: Pille Runnel, Pille Pruulmann-Vengerfeldt, Agnes Aljas, Kaja Tampere (ERM), Adriana De Cesare (FST), Marzia Cerrai (FST), Gianluigi Ferrari (UniPi), Cédric Gagnaire (DDS)

Revision History and contributions

Version	Status	Name, organisation	Date	Changes
0.1	ToC	Agnes Aljas (ERM)	11/11/2021	Structure of the text
0.2	First version draft	Agnes Aljas, Pille Pruulmann-Vengerfeldt, Pille Runnel, Kaja Tampere (ERM), Gianluigi Ferrari (UniPi), Marzia Cerrai, (FST), Gianluigi Ferrari (UniPi), Cédric Gagnaire (DDS)	13/12/2021	Full draft of all text
0.3	Second version draft	Agnes Aljas, Pille Pruulmann-Vengerfeldt, Pille Runnel, Kaja Tampere (ERM), Marzia Cerrai, (FST)	15/12/2021	Added adjustments; overall control of the text
0.4	Review of the second version	Agnes Aljas, Pille Pruulmann-Vengerfeldt, Pille Runnel (ERM), Marzia Cerrai (FST), Gianluigi Ferrari (UniPi), Cédric Gagnaire (DDS)	16/12/2021	Addressed peer review comments
1.0	Final Version	Pille Runnel, Agnes Aljas, Pille Pruulmann-Vengerfeldt, (ERM)	23/12/2021	Minor corrections

Review and approval

Action	Name, organisation	Date
Reviewed by	Marzia Cerrai (FST)	23/12/2021
Approved by	Coordinator	24/12/2021

Distribution

No.	Date	Comment	Partner
1.0	30/12/2021	Submitted to the EACEA	Fondazione Sistema Toscana

Application area

This document is a formal output for the EACEA, applicable to all members of the Me-Mind project and beneficiaries. This document reflects only the author's views and the European Union is not liable for any use that might be made of information contained therein.

Copyright notice

Table of abbreviations

Abbreviations	Description
D2.2	Deliverable 2.2
IF	Internet Festival
FST	Fondazione Sistema Toscana
DDS	Domestic Data Streamers
UNIPI	Università di Pisa
ERM	Eesti Rahva Muuseum
ToC	Table of Contents
WPx	Work package x
KPI	Key performance indicator
Tx.x	Task x.x
GSOs	Generic Social Outcomes
M/E	Museum/Event

CCIs	Cultural and Creative Industries
CDIS	Culture for Development Indicators

Table of Contents

List of Figures

1.Introduction of the project

Me-Mind is a project co-funded by Creative Europe programme and carried out in collaboration of four European partners: two cultural organisations (the Estonian National Museum and the Pisa Internet Festival), a creative agency (Domestic Data Streamers) and a research institution (University of Pisa). The project stems from the concrete needs of the two cultural organisations, but aims to provide relevant insights to Cultural organisations across Europe.

1.1 Scope of the deliverable

This document provides guidelines for small and medium-sized cultural organisations at the crossroads of culture, creativity and entrepreneurship, which are interested in learning how to better use data for impact measurement and communication. We also propose possible forms of collaboration between cultural organisations and creative agencies and between cultural organisations and research institutions for identifying relevant data sources, data collection and analysis The guidelines draws on three key elements of the project - 1) scoping the needs and requirements for understanding the impact of CCIs, 2) preliminary data collection and analysis experiments, and 3) audience-oriented interactive data installations. In addition to the project partner's collaboration, the deliverable draws on 13 interviews with European museums and festivals, establishing their needs for understanding the impact of CCIs.

1.2 Structure of the deliverable

This document is divided into the following main sections:

- Section 1 provides the executive summary of this deliverable;
- Section 2 summarises the main opportunities and challenges related to data collection, analysis and presentation in the cultural organisations;
- Section 3 gives the overview about the approaches to impact measurement, including the case studies at the two participating institutions;
- Section 4 provides the overview of the audience-related data collection possibilities;
- Section 5 gives an introduction for suggested data collection methods;
- Section 6 proposes approaches to communicating data for impact via data visualisation;
- Section 7 summarises the approach via the data driven model generation and gives recommendations based on the analysis for data-driven approach in organisational decision-making.

2. What we have learned so far

2.1 What is data

- Each organisation has internal data that is produced within the organisation, and additional data can be sourced from external stakeholders.

- Each organisation also has primary and secondary data. Primary data is purposefully collected, concerning a specific challenge of understanding the impact of the organisation or activities. Secondary data has been produced to understand other aspects of CCIs, but that can still be meaningfully (re)interpreted for understanding impact.

- Data can take different forms. There can be little of it in rich detail (qualitative data), or a lot of it in basic numbers of texts (quantitative data). Additionally, data that can inform our decisions can be texts, numbers, visual and audio-visual material and other forms. All these require different kinds of data collection and analysis skills.

2.2 Internal data challenges

- It is important to increase cultural organisations' awareness of the data they have inside. Most organisations make annual reports, project reports, collect financial data, carry out social media and media coverage analyses.

- Give internal data a better structure. Regular review of what data are collected internally and systematisation and reflection based on the data collected gives useful insights based on already collected data. There is no universal best structure, but the structure needs to speak to the current challenges of the cultural organisations, and data needs to be translatable from one part of the organisation (e.g. accounting) to another (e.g. outreach).

- It is important to increase data literacy, as it would allow for 1) understanding internal data sources and being able to relate them to the questions the organisation is facing, 2) understanding how to structure and organise internal data sets so that they can return meaning when the cultural activity we want to measure is over.

2.3 External data challenges

- CCIs should collect data that belong to external organisations. A lot of data can be collected from structured external databases of other organisations, based on an agreement between the parties.
- More data literacy relating to external sources will help with making sense of what is available, as not all public/external data sources are directly relevant, meaningful or easily accessible. External data is often unstructured and needs to be collected from different sources.

- Be aware of the opportunities and challenges of social media data. Social media operations can provide cultural organisations with large amounts of data. Critical awareness is needed to understand what is behind each data point provided in the social media reports. Awareness of the proprietary nature of the data that is being collected (and made available) by the social media industry, is also necessary.

2.4 Expectations to data

- Stakeholders interviews showed awareness of the need for data collection but acknowledged that it is not often done. Firstly, data is often not collected due to the lack of time and different priorities. Secondly, there is a perceived lack of data collection skills.
- There was an acknowledgement that data-based decision making could be useful for achieving a better performance of the organisations. However, oftentimes also performance is not strategically planned in the organisations. The organisations often lack good tools or methodologies for strategic planning with the help of data analysis.
- There is a pressure from both in- and outside of the institutions to produce a lot of different data, but what to do with the collected data afterwards? Organisations need support in giving value to the data in the decision-making process, including better governance.
- As data is getting increasingly complex, there is less confidence in the reliability and comparability of the metrics used by the organisations. In the ideal world, there would be a manual on how to measure impact, helping to understand, how to find metrics or what should be the metrics.
- There is a lot of data enthusiasm in cultural organisations. It is perceived that data gives a better understanding of organisational performance and enables better, data-based decision making on the basis of big data, open data etc.
- The public is at the centre of CCIs work. Therefore the need to understand the needs, values and expectations of their current and potential audiences and how their work affects it, is crucial.

2.5 Collecting more data

- Each organisation collects data differently and for different reasons. Often the data that is collected for some particular purpose, can be exploited in other contexts. However, there are a lot of unexploited uses, due to the fragmented nature of the data culture in the organisation, or simply because the need to exploit information from data that arises over time or with some particular conditions. The potential of most internal data in any context, therefore, remains latent.
- If an organisation wants to collect data from its audience, it has to keep in mind that no topic is too complex to address if you make sure to make it understandable and relatable for everyone in your audience. For example, you won't ask them about the "big theme of culture", but about what they felt when they read a book, listened to a song or went to a museum. It makes it easier for the users to explain, remember and relate and in this way you can easier collect information from your audience.

- Use different data collection methods. A bold physical installation can serve as a hook for a digital experience through which to engage audiences and collect data from them.
- In the actual data collection process with the physical installation, it is important to have somebody who supports the participants if they are filling in the survey. It helps to overcome the barriers. The assistants can help people or attract new participants, and can engage in the data conversations, adding an additional rich layer of engagement.
- When collecting and analysing data about the engagement with cultural experiences, it is important to keep in mind that for some people the most meaningful cultural experiences are not about culture, but about sharing. That does not mean that the cultural experience has not had an impact, but rather the opportunity to share it with other people amplifies their experience and makes it more memorable.
- Different data sources support different knowledge - for example, quantitative data enables to give a more shallow overview of larger processes, whereas qualitative data helps better to explain the contexts, but it will be more difficult to collect and process in larger numbers.
- Impact measurement should be related to long-term organisational goals and ambitions, which are the basis for gathering existing data or collecting new ones.

2.6 Analysing data that is available

- CCIs have intuitive capabilities of data analysis and they apply them in everyday work. Data work starts with simple tools like Excel, and CCIs are able to interpret the contexts and relations behind the numbers. Competencies in managing the data (data literacies) are still limited, especially when we look at the huge metadata sets. Also, data analysis does not happen by magic usage of technologies. Data analysis also consists of the processes of cleaning data and assessing its reliability. Often data comes from multiple sources, therefore it is important to avoid duplications, unnecessary fields, to titling and formatting the incomparable fields.
- Therefore, for more in-depth insights, data analysis would benefit from the integration of data science into CCIs. Data scientists can help to derive meaningful results from unstructured and unclear data sets and put it into the context of organisational challenges and opportunities. They build mathematical models to understand the underlying relations and correlations between data. When collaborating with CCIs, data scientists must be flexible and explain all the steps required for the task at hand.
- The design of the mathematical models underlying the analysis (and related algorithms and software tools) is a fundamental step that should not be underestimated since each model is related to a particular type of analysis. Consequently, the type of analysis is tied to the limitations of the mathematical model itself.
- Data are as valuable as the use of it. Integration and cleaning of internal data are the main challenge as data comes from multiple sources, often with limited explanation. Data integration is best learned through experience. It is one of those things that will take a lot of effort and time to learn.
- The data analysis is a continuous process where each phase provides feedback to others. A clear understanding of the objectives of the analysis supports the choice of

sources in the collection phase and the identification of the types of operations carried out in the phase of data processing.

- The results of data analysis often require logical reasoning to explain why a certain correlation exists (especially with people who are unfamiliar with the nuances of data science). A proper explanation will help in understanding the results of the analysis. One of the ways to do it is data visualisation.
- To conduct a sentiment analysis on the responses to questionnaires that an organisation had administered to its stakeholders, it is required to implement an iterative and structured process to widen the number of answers and to ensure their meaningfulness.

2.7 Visualising data

- Data is often perceived as very technical and hard to understand, because it is often based on numbers. Data should be understood quickly by everyone in your target group. Data visualisation is the perfect way to do it. Data visualisation helps to better understand the trends and patterns in data, especially with big datasets and spreadsheets. When data scientists have created machine learning algorithms to understand the trends, relationships and correlations, data visualisations make it easier to understand and interpret.
- An innovative way to make your data presentation easier to understand to your audiences, presentation could involve interaction. By interacting with the data visualisation, the visitor will be immersed in it, becoming more involved. This will make it easier for them to understand it.
- Data visualisation does not need to be digital. Sometimes interesting interaction can be achieved when things are used to represent data in creative ways making the data analysis results meaningful and relatable.

3. What is impact and how to measure impact?

The impact can be defined as the outcomes or benefits of organisational actions (cultural, business-related) that demonstrate a change. Me-Mind is focused on finding innovative ways to a) understand the type and quantity of impact of the CCIs, b) measure and c) communicate social and economic impact and value created by CCIs in their operational environment and in their relationships with various stakeholders. Compared to other sectors, CCIs have additional dimensions and roles, which create both direct and indirect impact addressing such wide-scale impact areas as identified in the UNESCO framework of Culture for Development Indicators (CDIS), including economy, education, governance, social, gender, communication and education. While we suggest that the impact of the CCIs could be assessed in the wider context of impact areas, this does not mean that there is a direct, causal relation between service-level efforts of the CCIs and the wider societal level results.

Understanding the impact has several challenges and it is important to distinguish the different elements of the impact discussions. When studying impact, it is easy to confuse data coming from different parts of the process. The following diagram will give an overview of the different

stages of the impact process, including the possible points of data collection at each stage. It also demonstrates the feasibility challenge when it comes to actual impact.

Figure 1. Stages of impact processes as possible points of data collection.

1. Inputs
   ○ Money, people, time, resources, intellectual and material contributions.
2. Activities
   ○ What cultural organisations do? Overview of activities and actions done to achieve impact?
3. Outputs
   ○ How many people have visited the museum/festival? How many publications, things in collections, nuggets of collaborative efforts etc?
4. Outcomes
   ○ What is learned, what is remembered from the experience?
5. Impacts
   ○ How learning and experiencing has changed lives?

While it is relatively easy to collect data about inputs, activities and outputs, it is more difficult to collect information about outcomes, and even more challenging to understand the impact of the cultural experiences on the people or the impact of the CCIs on the town or region. The impact analysis, hence, is almost always secondary, drawing on assumed relationships, as it is hard to isolate and attribute different influences.

An example from the Estonian National Museum is related to real-estate prices. In 2016 ENM opened a new building at the former Soviet airfields. Initially, the controversial location was cleaned from toxic waste, and a new museum infrastructure was created, completely changing the area. The new museum building became very popular among the visitors. Parallel to this process, both city and county engaged in urban development projects in the region both to support the museum, but also to expand housing in their suburbs. Several years later, the prices in the area have increased (but so have prices increased everywhere), but the community that originally saw the value of the proximity of the museum (the museum is mentioned in all the real-estate advertisements) is complaining about the noise levels of the events the museum is organising. In this anecdotal evidence, it is difficult to attribute positive, negative or scale of the impact to urban development, as several parallel processes have been at play. Yet it is also clear that without the collaborative efforts of the state, county, city, museum and the people living in the area, the urban development would be far from what it is now.

When analysing the impact, it is thus important to understand that collecting evidence about impact needs careful reflection and analysis of 1) what processes we assume to have an impact and 2) what kind of impact we hope to have, 3) what kind of information is available, 4) what is possible to collect, 5) and what is the timeline of the impact.

The better our understanding of our own inputs and actions, the more systematic the data collection and analysis processes are, the better possibilities we have to understand the impact. The challenge is to define a clear set of actions and desirable impacts in order to have a chance for systematic and comprehensive analysis of it.

3.1 Context of the EU projects

Impact measurement is often oriented towards external advocacy and validation, such as the support of specific decision-makers. In this case, cultural institutions are under the pressure of external stakeholders to demonstrate the impact of their work. Accordingly, the impact indicators, through which to evaluate success, might be defined from outside. On the other hand, when a drive to understand impact comes from within the organisation for internal management or self-evaluation purposes, including the needs for feedback, validation and input for future planning and programming, the organisation is freer to understand the value and relevance of the particular indicators for decision-making and make its own choices. The language of impact assessment varies: we can talk about impact, returns, benefits, value, relevance etc. This makes the Me-Mind project have many relevant parallel processes with other EU funded projects.

"MOI! Museums of Impact" (2020-2021)[1,]

MOI! Museums of Impact is a European cooperation project co-funded by the Creative Europe Programme. The project aims to develop a self-evaluation framework with a people-and impact-oriented perspective to be used by the European museum sector. The self-assessment framework helps museums in capacity-building to better meet the demands of a changing society. The project aims to assist museums in increasing their impact by supporting them in evaluating and refocusing their work according to their strengths and strategies.

"PAT – Performing Arts in (a world in) Transition" (2014-2021)[2]

Perform Europe is a two-part project which aims to rethink cross-border performing arts presentations in a more inclusive, sustainable and balanced way. Its goal was to develop innovative business/management and self-organisation models. Secondly, it aimed to identify a set of indicators for the measurable impact of the artistic projects to better defend arts funding. At stage 2, the goal was to empower contemporary performing arts professionals with skills, competencies and know-how that help them to successfully adapt/react to changes in their social, economic and political contexts (caused by globalisation, demographic transformation and the digital shift).

---

[1] https://www.museovirasto.fi/en/museum-sector-development/moi-museums-of-impact
[2] https://www.ietm.org/en/performing-arts-in-a-world-in-transition-i/activities

"Create to Connect -> Create to Impact: Changing the world with theatre?"
(2018-2022)
A joint effort of ten European cultural organizations to create long-lasting connections of artists, cultural operators, researchers and audiences. The project seeks new production models for engaging the audience in new, innovative ways and creates new public arenas through dialogue and participation. In order to measure and act upon impact, the CtC -> CtI's partner The Research Centre of the Slovene Academy of Sciences and Arts uses an ethnographic research approach to assess the impact of partner organizations' activities in society.

"SoPHIA – Social Platform for Holistic Impact Heritage Assessment" (2020-2021)[3]
SoPHIA aims to promote collective reflection within the cultural and political sector in Europe on the impact assessment and quality of interventions in European urban environmental and cultural heritage. It creates a Social Platform for different stakeholders to work jointly towards an impact assessment model, quality standards and guidelines for future policies and programmes. The work is organized around four dimensions – social, cultural, economic and environmental impact – which helps to identify the challenges and opportunities linked to cultural heritage interventions in Europe.

"inDICEs – measure the impact of digital culture" (2020-2022)[4]
The goal of the project is to empower policy-makers in the cultural heritage sector to understand the social and economic impact of digitisation and engagement with digital culture. This will help to address the need for innovative (re)use of cultural assets and reform the sector with new business models and legal frameworks. As a result, the cultural heritage institutions would be able to increase their positive contribution to CCIs and society at large.

3.2 Case study: An example of defining the indicators for impact measurement

Me-Mind started with the collecting of needs and expectations of two vertical sectors (events and museums), using its two case studies. This work was based on the UNESCO framework of Culture for Development Indicators (CDIS) for understanding the impact of the cultural sector[5]. Thus the work was also testing the applicability of the multidimensional framework and whether it could serve as an alternative to a large number of approaches and models to impact assessment often applied in the organisations, limited to a particular sector. The seven areas from the UNESCO CDIS framework include inspired identification of indicators and data points for the Me-Mind set of cross-sectoral indicators. But this assessment exercise also showed that there are other relevant indicators, essential for the involved organisations, but not covered by the CDIS. The exercise with the framework helped us to identify more than one

---

[3] https://sophiaplatform.eu/en
[4] https://indices-culture.eu/
[5] https://en.unesco.org/creativity/sites/creativity/files/cdis_methodology_manual_0_0.pdf

hundred indicators, to which we could also identify data sources. However, the preliminary analysis of the data shows that lack of context can give little new insights for the participating organisations, and the challenge is to be able to understand the diversity of causes behind the fluctuations in numbers.

3.3 Lessons learned

- Each organisation already collects vast amounts of different data, but they are rarely brought together on the same overview.
- It is difficult to distinguish between desired impact, input and action data, as we often tend to assume action or output to be stand-in for impact.
- The different data sets are challenging computational analysis, as there is a lot of noise, formatting issues and other problems.
- Even with the small case of museum and festival, we saw a number of indicators that were of varying relevance. Data that was easily available and not particularly interesting for one organisation was seen as highly relevant and desirable for the other. The different organisational practices of data collection and use can vary a lot.
- Each cultural organisation needs to be able to identify/customise their own indicators, linking them to particular use cases identified with concrete activities/strategic decisions. List of all the possible indicators becomes overwhelmingly large, unstructured, and unfeasible to manage for any cultural organisation alone.
- Indicators need to be sector-specific. For instance, indicators relating to heritage sustainability and collections are very important to museums, but not at all relevant to festivals.
- Defining indicators, it's a good exercise that imposes organisations to actualise and give shape to the complex concept of impact. Conducting a self-assessment of the relevance of each identified indicator and the level of ease of access to the data needed to answer that indicator helps plan a clearer strategy for conducting impact measurement.

3.4 Needs for understanding impact

Additional stakeholder interviews show that impact measurement is very important for their organisational performance, yet challenging to practice in ways that give input for decision-making. Organisations are familiar with the concept of impact. They have an overall good understanding of what they need to demonstrate to stakeholders, such as government, local administration, donors or private sponsors. Often the impact is assumed to be the same as the numeric input, activities or output indicators, including finances, visitor numbers, media coverage, and social media reviews. The interviewed expressed interest to have a better understanding of the impact of their organisations and to reach beyond the numbers, which they usually have to report to external stakeholders. The CCIs also question if the metrics they use are reliable, accurate and comparable, as these often provide contradictory results.

Some of the questions that organisations face relate to the trends in CCIs development. For instance, there is interest to see the value of the cultural sector for the wellness sector, impact to health, serving better the ageing societies etc. The attention to the questions of climate

breakdown and sustainability also challenge CCIs to identify new indicators to better understand the impact in that area. COVID-19 has forced CCIs to move many of their COVID-19 has forced CCIs to move many of their activities online. That in turn has created the need for understanding digital engagement. As museums are more often holders of large collections, then compared to cultural events, they are also more interested in their collections and research. They need to understand their relevance and impact on society. This cannot be understood on the basis of assessing the size of one's collections only.

Our interviews show that open data presents a new challenge for the CCIs. As society is pushing towards open data movements and FAIR data practices in collaboration with cultural and academic institutions, there seems to be a lot of data around. But reaching clever insights from this amount of data is still a challenge for most of the interviewed.

Some of the CCIs in our interviews also mentioned that impact measurement models have been developed and tested at the sectoral level, but the complexity of the approach means that this is not applicable at the single cultural institution. We will continue by comparing our originally developed impact indicators and identified data points with the interview results to ensure a robust business model canvas.

4.Case studies about audiences

Below we describe two experiments we conducted on project use cases - the Museum and the event. The experiments represent two different methodologies that a cultural organisation can apply in order to understand what kind of impact its activities have upon its audiences. The first case demonstrates utilizing an existing internal data set of the organisation. The second case study demonstrates creating an info-experience that engages the audience in a dialogue.

4.1 Using digital data for understanding audiences

Context:
Museum has a lot of digital data, which is not often the object of analysis. Since October 2016, the Estonian National Museum (ERM) has operated in the new building. The museum has two large permanent exhibitions. Data in this example concerns the Estonian exhibition Encounters (3500 m2) which is a complex and information-dense environment, consisting of hundreds of artefacts, multimedia objects and texts, which does not have a clear path, but where visitors can choose their way in the exhibition.

Data:
The exhibition texts are currently available in seven languages. The ticket holder has to choose their preferred language when buying the ticket, and receives a language card/smart ticket that can be used to change language at the digital exhibits. The system tracks contacts made between a smart ticket and the screens. System also records which local network addresses are used, giving information about the language choice, time of the use, based on the local network addresses of the text screens. The dataset analysed is composed of 12 log

files. The museum holds data from the five past years and hundreds of thousands of visitors, which has not been analysed yet.

Possible questions:
- Is it possible to study the visitor trajectories in the exhibition space?
- Can we compare different language users, different seasons visitors etc. behaviour?
- How can this information be visualised?
- In which ways this data can be combined with other data sets?
- How to follow the interests of the visitors? How to interpret the data, which does not explicitly provide any meaningful information about it?

Knowledge obtained:
- The dataset allows tracing the path of visitors inside the exhibition and the most used digital and analogue objects in the exhibition.
- Most visitor trajectories include frequent interactions with devices, with an average of 17 devices queried.
- The devices most used are those at the entrance of the exhibition. Common patterns of visitor trajectories did not emerge: the routes analysed are all different from each other.
- Visualisation can support additional knowledge gains.

Challenges:
- The e-tickets dataset gives an overview of the most used items and trajectory people are taking in the exhibition, but the context of why these screens are in use remains unknown.
- With complementary qualitative visitor interviews or other methods, the museum could have a better understanding of the choices people make on the content in the exhibition space.

Potential uses of data:
- The museum would use this information to make better decisions when exhibition renewals are in progress.
- The museum's e-ink based system of digital labels has offered the global museum field the solution of using a potentially unlimited number of languages in an exhibition environment. Applying the visitor analysis linked to these digital labels would potentially add a new layer to that service.

4.2 Audience engagement as data collecting method

Figure 2. Il Nodo della Cultura installation at IF 2021

Context:

Internet Festival (IF) is a large event that takes place annually in the city of Pisa (Italy): onsite for 4 days across 10 different venues of the city and online (since 2020, because of the pandemic). Of its nature as widespread (there is no proper entrance where the public checks in to the festival) and fully free event (there is no entrance ticket), the festival has always had difficulty gathering feedback, data, and structured information from its audience. Moreover, it was always a challenge to get the IF audience's time and attention to answer the questions of the organisers.

Me-Mind experimented with a space to gather visitors' data and opinions at the festival edition in October 2021. To do this, an artistic installation was created that served as a hook for people to answer a digital survey. The installation was conceived as a space to start conversations with the visitors of the festival about culture and how culture has an impact on their lives. Gamification techniques were applied for triggering the dialogue. Answering the questionnaire was split between onsite and online: four questions could be answered at the onsite panels through playful interactions, rest of the questionnaire was filled in online. Once the respondent had completed the online survey, her/his answer to one of the digital survey questions was printed in real-time by thermal printers inside the installation, joining the collection of anonymous stories which continuously grew during the festival.

Data:

The data set includes 550 answers collected to the questionnaires during the 2021 edition of IF through the installation. The survey was available in two languages, Italian and English, and is made of questions linked to indicators developed in the first phase of the study. Accordingly, the data contains different information about the impact, feedback, and perception of the IF audience. The detailed level of the survey allows us to observe information at several layers and from different points of view. In particular, the questions referred to 8 macro-areas of the UNESCO framework, including:

- economic area;
- education area;
- cultural rights, cultural diversity, and cultural development;
- the cultural fabric of the city and beyond;
- gender equality and sensitivity to minorities' issues;
- social participation area;
- communication area;
- important factors (that cannot be classified in one of the above categories).

Knowledge obtained:

The received insight involves information regarding the social-demographic background of the IF visitors (origin, age, education, profession), their feedback about the festival itself, and their experiences and explanations on different cultural events. High level of detail of the questions in the questionnaire allows us to study the dimensions connected to our festival, e.g., demographic-economic and socio-cultural, and the dynamics between them.

Challenges:

- The responses to open-ended questions can provide free text on which to apply Entity Extraction, Sentiment Analysis and Emotion Detection. As per choice of questionnaire design, the team decided to keep the number of open questions small, resulting also in few open-ended answers with limited explanatory power.
- When designing questionnaires, the CCIs have a balancing act to do - more open ended questions give larger corpus of free text to analyse, but runs the risk of respondents being tired and the questionnaire to be hard to complete. Fewer open ended questions give less free text, but often more completed surveys.
- CCIs can also opt to use closed questions (e.g. multiple choice or Likert scales) to understand audience emotions. In that case, there is no need to conduct more complex data analysis methods, but it also limits the free expression of emotion, and constrains it to pre-determined adjectives.

Potential uses of data:

- The festival would use the socio-demographic information acquired in order to better plan its own next calendar on a base of segmentation of the target/time.
- The knowledge acquired with the Machine Learning and Natural Language Processing techniques such as keywords and entity extraction would help the festival to identify topics for its next edition.

- Sentiment Analysis and Emotion Detection results would help the festival to identify negative, positive, or neutral attitudes toward IF and then to balance or implement actions towards specific targets.

5.Overview on data collection and analysis methods

5.1 Overview on the data collection methods

Data collection is the systematic process of gathering data from different sources. Whether one is developing an analytical model for marketing or business purposes, data collection allows choosing what kind of information would provide the best understanding of the problem at hand. CCIs should always have an active role in decisions relating to data collection, especially if they want to get insights, make forecasts, and evaluate the impact of cultural actions in a way that creates significant value. While methods and aims may differ between application fields, the overall process of data collection remains largely the same.

Before starting data collection, you have to understand the goal of the analysis since having clear goals allows identification of the needs. This will inform you, how much you need to rely on **primary data collection approaches**, collecting data directly about your questions. Primary data collection approaches are mostly linked to obtaining more knowledge and insights from the audiences which are at the core of any organisational operations and decisions for the CCIs.

We list below the primary data collection methods and provide a few examples.

Surveys and Questionnaires

Goal: To understand the general characteristics or opinions of a group of people.

Surveys and questionnaires are a means of obtaining data from targeted respondents in order to generalise the results to a broader public. Almost everyone involved in data collection relies on surveys and questionnaires to obtain credible data and insights from their target audience. One of the key challenges with surveys and questionnaires is recognising the relationship between the overall population (broader public, target audiences), and the sample (the people who actually respond to the surveys).

This method is especially relevant for CCIs, and there are plenty of accessible guides available. But several key aspects have to be considered when this data collection method is adopted:
- *Surveys can be done online*. Online surveys are becoming more and more prevalent every day, however, the sample is often very hard to control in the online surveys unless a professional company is recruited or targeted sampling strategy is used.

- *Online surveys can be accessed anytime and anywhere*. Online accessibility provides CCIs with a significant advantage: collect data from your target audience with ease, skipping the step of manual data entry and additional human error possibilities.
- *Low price*. Compared to the other data collection methods, creating surveys and questionnaires does not require big budgets.
- *Wide range of data*. When utilising surveys and questionnaires, CCIs will acquire the ability to collect different data types such as opinions, preferences, and so on.

Interviews

Goal: To gain an in-depth understanding of perceptions or opinions on a topic.

An interview is a meeting between two individuals in which the interviewer asks the interviewee questions in order to gather information. An interview not only collects personal information from the interviewees, but it is also a way to acquire insights from people. Here is the summary of the advantages of this method in data collection:

- *Conducting interviews can help reveal more data about the subject*. Interviews can assist CCIs in explaining, understanding, and exploring perspectives, behavior, and experiences of people.
- *An interview can be open-ended*. Unlike other methods, interviews can be both structured supporting a more comparative approach or open-ended, enabling interviewers to ask follow-up questions in order to better understand the context.
- *Interviews can be done online.* With the advances of video conferencing, more and more interviews are conducted online, allowing to reach a broader set of respondents, reducing the costs of travel to informants and interviewers.

Observations

Goal: To observe and understand the behaviour of individuals inside of a certain context or a place at a specific period (time, day, year).

The observation as a data collection method involves seeing and evaluating the behaviour of groups of individuals or environments in which they are detected. Observations can take place at the events, in the museums, or in the digital spaces. In our digital world, the capacity to collect and analyze the behaviour of individuals in online social networks at higher granularity provides massive datasets supporting the observations of interesting patterns of interactions among individuals.

Here are some advantages of Observation as a data collection method:
- *Ease of data identification*. This data collection method does not require technical skills to identify sources of data.
- *Offers detailed data collection*. Observations on individuals give CCIs the ability and freedom to analyse the specific behaviours and actions they are interested in.

Internal Data and Documents

Goal: Identify and analyse internal data, or data collected for other purposes within an organisation that would be useful for understanding organisational impact.

This data collection method involves systematising and analyzing CCIs existing internal data to track or understand operational and business activities over a specific time period. The data may include staff reports, data from internal databases, ticket information, exhibition brochures, to cite a few.

There are the significant advantages of using internal data and documents as a data collection method.
- *The data is already available*. There is no need for CCIs to conduct any active collection because the information is already available.
- *Easy tracking of collected data*. Internal data will allow CCIs to evaluate and understand the history of a specific cultural event.

In addition to the primary data collection methods **secondary data collection approaches** can provide additional, important insights.

Open data

Goal: Collect data on specific topics available on public repositories

Open data sets are usually secondary targets in data collection. The sets contain information that anyone can access, use and share. This data might come from real-life observations and it has likely been collected for a purpose which is not directly emerging from your particular point of interest or research question. Open data usually are made available in a common, machine-readable format under suitable licenses. Typically, licenses permit people to use the data set in any way they want, including transforming, combining and sharing it with others, even commercially. The main advantage of using the open data collection method from public institutions is given by the accuracy and the integrity of the collected data. Secondly, public and open source datasets save time (do not have to be built from scratch) and enable collaboration with other parties, if necessary.

Examples of open datasets are:
a. Kaggle is a house-hold name by now amongst data professionals. Kaggle hosts massive open source public data across various domains. (https://www.kaggle.com/datasets)
b. Data Driven hosts competitions for Data Scientists to tackle the real world challenges, solving of which leads to relevant social impact. They have datasets from their competitions. (https://www.drivendata.org/)
c. KDD is a competition held by ACM on Knowledge Discovery and Data Mining, and it hosts datasets with detailed data dictionaries and instructions. (https://www.kdd.org/kdd-cup).
d. Google has hosted tons of datasets on Google Public Datasets. One can browse through their dataset collection using BigQuery. (https://cloud.google.com/bigquery/public-data/).
e. The AWS dataset has hundreds of datasets. AWS hosts the largest open datasets available (https://registry.opendata.aws/)

To work with Open Data, the strategy we propose for CCIs is to explore data sets made available by public institutions. For instance, to evaluate the impact that cultural events can have on the tourism of the territory, the primary candidate for information is the flow of tourist arrivals and presences, possibly with monthly granularity, municipal detail and focus on the area. Typically, these data are collected and analyzed by public institutions such as tourist offices or the statistics or research offices of government agencies. Different countries host their own Open Data repositories, seeking to make public governance data open and reusable, alternatively, requests can be made to the public offices to give out specific data collected for the governance purposes.

5.2 Online data scraping

The internet is one of the most used data sources, including both primary and secondary data sources. Among other insights, internet data sources enable analysing interactions between cultural institutions and their audiences. But there are also other web resources which the CCIs can identify as relevant to their decision-making areas.

A variety of approaches and software tools supporting data collection from online interactions are available. For instance, Data scraping: the process of collecting data from a website and saving it as a local file on a storage system. It is the most effective data collection tool that CCIs can use to gather information from social networks. CCIs may customize their scraping criteria or parameters to selectively target a specific attribute. Scraping supports the collection of both qualitative and quantitative data.

6. Data visualisation for impact

Knowing where to look at for finding data and developing practices that support data-based decision making and increase impact can be advanced by applying smart data visualisation techniques. While basic data visualisation approaches, such as presenting most important information to stakeholders in an visually engaging way, is the standard of contemporary communication and outreach practices for CCIs, Me-Mind project explores the possibilities of applying more advanced, artistic data visualisation approaches particularly in the domain of engaging the audiences by exploring the social and economic impact of the participating cultural organisations IF and ERM. We aim to connect data and design through practical action: designing meaningful narrative and emotional storytelling approaches for facilitating the info-experiences of the audiences. The key advancement here is replacing one-way reporting with experimental two-way, real time communication.

Current tools which can be fully digital, combine analogue and digital elements, or be entirely analogue, enable to collect and visualise data in real time. It is possible to display data derived from smaller (emerging) data sets or huge quantities of data and show data in a dynamic way, in motion. The strength of designing these info-experiences is the interaction of the audience: their participation turns it into a whole experience, as well as creates the data set, which is

communicated back in a meaningful way in real time. Audiences can be placed at the very centre of data collection.

By asking audiences/stakeholders to take part in the data visualisation exercise enables becoming a real-time learning organisation and understanding the participants better. By using emotion-related data visualisation, CCIs can learn how the participants feel about certain topics or perceive them. Those results can help to define, modify or approve previous knowledge or the interpretation of some topic by the audiences or other audience-related questions in real time. Integrating this emotion-filled data with data from other sources CCIs can tell meaningful stories and create impactful strategies.
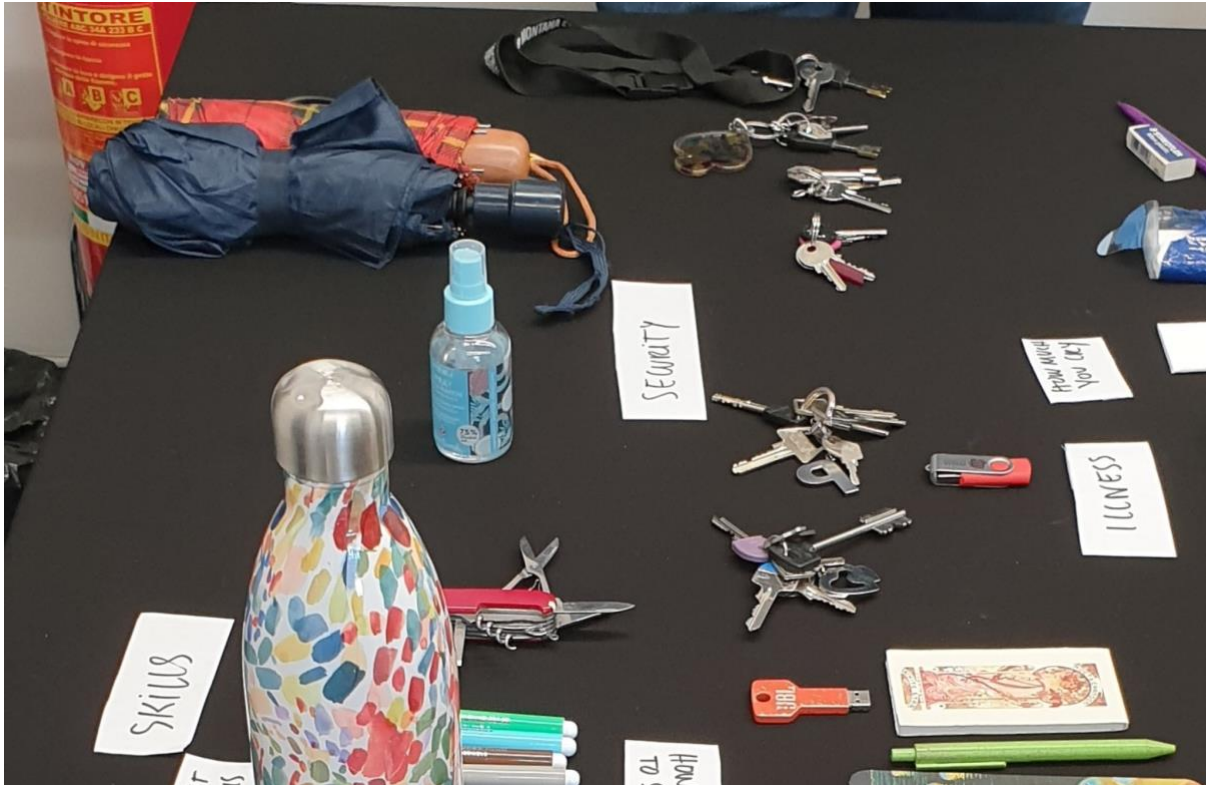
Figure 3. Different sets of house keys of the workshop participants serve as emotion-filled data.

As an example of such practices, we introduce a short set of exercises carried out in the workshop organised for the participants of the IF by Domestic Data Streamers. The goal of the workshop was to learn how to turn data into impactful stories we care about. After splitting the participants into two groups, the group members laid all the items they had in their pockets or handbags to the tables. The participants had to observe and discuss these objects and try to group them in a way that would enable them to visualise specific topics, which had emerged from the sets of objects. Just by using those personal objects, the groups were able to represent a scale of how anxious people were regarding COVID-19 (playfully organising the face masks with different protection levels, or how much they had invested into the security of their homes (comparing different sets of house keys, see Figure 3).

A second exercise consisted of asking participants a simple question using templates that are accompanied by two different "voting everyday objects". In this way, they could answer the main question while adding another level of filtering in the answers. The workshop participants were asked, what are their deepest scars - either physical or mental. By choosing a classic band-aid or a more colourful one for giving an answer, the participants could also choose to tell whether they got these scars as an adult or a child (see Figure 4).
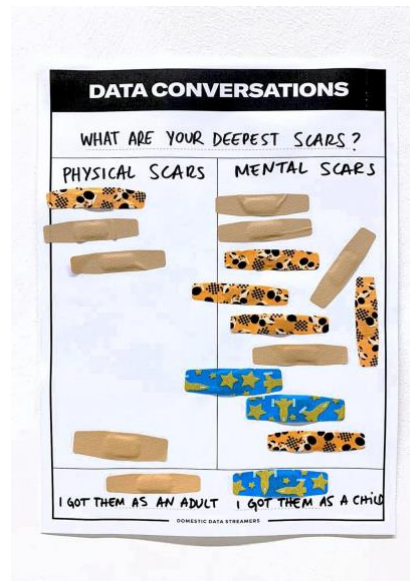

Figure 4. Data collecting worksheet for engaging workshop participants in sensitive conversations.

Although these examples can be perceived as too simple or superficial for any deeper insights, the design of these interactions enables us to talk about sensitive and private issues that would otherwise remain hidden, if asked directly in the interview or group discussion in an unfamiliar context. Thus it also adds another level of engagement between the participants instead of just delivering the answer and opinion to the workshop lead or facilitator, collecting the answers. Last, but not least, making data collection more fun attracts people. This can serve as an introduction to a more in-depth conversation with them to gather even more impactful and insightful data.

The workshop at the IF also conveyed the message that data is something we can generate ourselves at all levels of the institution, around the topic we are interested in, using tools as simple as a piece of paper and a set of stickers. This example of sets of playful exercises also suggests that no matter what the question is, focus on smart, but simple data collection methods also serves as an enabler of basic interpretation of data, which can be carried out without advanced skills in data analytics. The most powerful interpretation and communication can happen, when data shows the story itself.

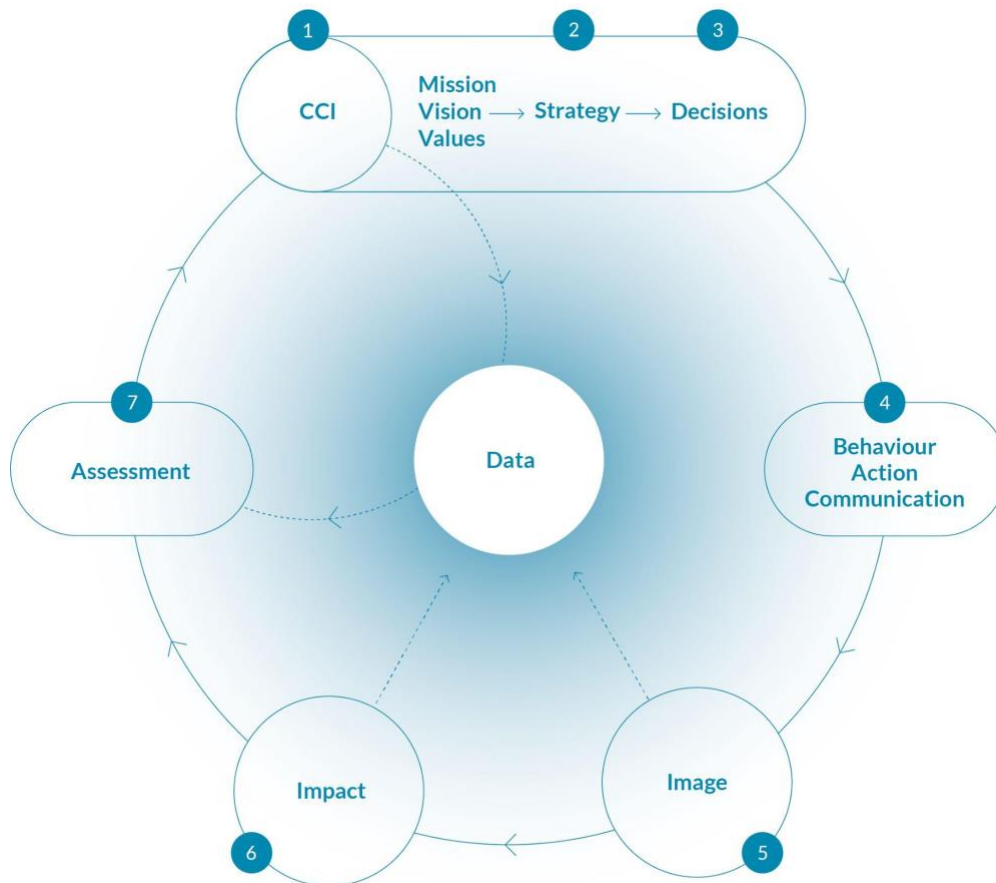# 7. Conclusions and recommendations. Data driven model generalization

Figure 5. Data driven model for CCIs

A) Recommendations regarding organisational development

Based on the activities in the Me-Mind project, we suggest that organisations benefit from developing a culture oriented to taking data-based actions and decision-making (Figure 5). Data based problem solving and decision making require a clear understanding of the mission and vision (2) of your CCI (1). Decisions (3) are a fundament for impact. In actual work, not all decisions in cultural organisations can be strategically grounded.

Assessing data collected through the lens of strategic planning facilitates to measure and understand the impact of decisions before taking action. It is advisable to assess the impact of specific actions.

Hence, impactful organisation, and the ability to understand impact starts from the clarity of the processes of what happens inside an organisation.

It is important for organisations to have:

- A clear understanding of the mission and vision and values inside the CCI, as understanding impact starts from the questions - **what needs to be impacted**.
- **Readiness to reflect and analyse the organisation from inside** and action-oriented organisational culture (data-based actions and decision making).
- **Competence of data collection, analysis, and contextualisation**.

CCIs can have an impact through behaviour, actions or communication activities (4). Planning the assessment and identifying possible data sources, data collection activities and analysis and linking them to strategy questions is more complex when impact measurement has the ambition to cover all the activities of a larger organisation under one umbrella. **Starting with specific actions and seeking to understand the impact of those actions makes the task easier.**

According to the UNESCO framework, CCIs have an impact (6) in the following seven major areas: Economy (both direct and indirect), Education, Governance, Social, Gender, Communication and Heritage. However, as discussed above, not all of these impacts have been easy to measure or capture, and oftentimes it cannot be expected that a particular process, service or activity in the organisation has a clear causal link with the impact on these generalised impact areas. In the interviews we carried out with cultural organisations, it became apparent that the needs for impact assessment are more often identified from the inside out, including more specific and articulated impact areas, such as local businesses, tourism industry, wellness, environment, urban development, COVID or growing digital presence.

The assessment (7) of the impact can be generic (across all the impact dimensions) related to a concrete action or behaviour, or specific, seeking to assess impact in a specific area. The impact assessment needs data to be collected and analysed, and the results of the assessment need to give feedback to the organisation, supporting the next actions to be taken. **For successful impact measurement, and data-driven organisation, the above-described circular loop needs to be constant within CCIs, as data is used to assess, and where needed, adapt organisations strategy and decisions**.

B) Recommendations regarding forms of collaboration

The Me-Mind project explores the opportunities of the cross-sectoral approach for identifying relevant data sources, data collection and analysis, bridging cultural institutions with creative sector and research institutions. Our case studies suggest that in the case of the CCIs, it is not always necessary to enlist the help of a market research company for data collection, analysis and presentation. In certain cases, a collaboration between a cultural organisation and a creative agency should be considered. The decision, what kind of data collection actions are needed should stem from the identification of needs and expectations of the CCI.

Traditionally, market research companies deliver their results in the form of reports and presentations. Such delivery remains a formal information experience both to the organisation and its audiences/stakeholders. In the Me-Mind consortium, the collaboration between the cultural organisation and creative agency is about experimenting with real-time information gathering, which is simultaneously data collection, presentation and interpretation experience. It is well justified to experiment and develop approaches, such as artistic dissemination, keeping human elements at the centre of impact discussions especially in the cultural sector, as the work of the CCIs is about cultural and artistic experiences and services. In the case of the approaches, where data collecting and data visualisation stages merge into another, they become part of the core services or products offered by the CCIs, being thus a good fit for analysing and communicating impact in CCIs.

A three-way collaboration involving CCIs, data science and creative agencies as partners can be challenging, but also a rewarding experience. In these collaborations, CCIs need to take a leading role in deciding on data collection, analysis, and presentation. As final beneficiaries, they need to be able to incorporate the results in their strategic decision making. Data scientists participation is important as they will allow broadening the understanding of what data can be collected and analysed, but CCIs need to be contextualising and interpreting the results, as data rarely works in isolation. Creative agencies contribute with their expertise in visualising, and audience engagement, as data is just one aspect of the impactful life in CCIs. Participating in data-based conversations, engaging data in decisions, and planning strategically are important elements in data-based organisation work that can be supported by data science, and creative industries, but need to be initiated and carried through by CCIs.