



UNCHARTED

Understanding, Capturing and Fostering the Societal Value of Culture



The UNCHARTED project received funding under the Horizon 2020 Programme of the European Union

Grant Agreement number: 870793

Deliverable number	D3.9
Title	Report on the representations of cultural value in cultural information systems

Due date	Month 33
Actual date of delivery to EC	15 th November 2022

Included (indicate as appropriate)	Executive Summary	<input type="checkbox"/>	Abstract	<input type="checkbox"/>	Table of Contents	<input checked="" type="checkbox"/>
---	-------------------	--------------------------	----------	--------------------------	-------------------	-------------------------------------

Project Coordinator:

Prof. Arturo Rodriguez Morató

Universitat de Barcelona

Email: rodriguez.morato@ub.edu

Technical Coordinator:

Antonella Fresa

Promoter S.r.l.

Email: fresa@promoter.it

Project WEB site address:

<http://www.Uncharted-culture.eu>

Context:

Partner responsible for deliverable	UNIBO; TRI
Deliverable author(s)	Paolo Ferri, Simone Napolitano, Andrea Carlo Lo Verso, Chiara Giulia De Leo, Rebecca Levy Orelli, UNIBO; Ola K. Berge, Åsne Dahl Haugsevje, Ole Marius Hylland, TRI.
Deliverable version number	1
Dissemination Level	Public

Statement of originality:

This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.

TABLE OF CONTENTS

Introduction.....	4
First pair – Content aspects.....	5
1. Cultural statistics in Norway (case 4.1)	5
1.1. Case background	5
1.2 Methodology	6
1.3 Analysis and discussion.....	6
2. Regional cultural observatories in Italy (case 4.2).....	10
2.1 Case background	10
2.2 Methodology	10
2.3 Analysis and discussion.....	11
3. Pairwise comparison – cases 4.1 and 4.2	17
Second pair – Processual aspects	20
4. The Cultural Index Norway. Benchmarking local cultural activity (Case 4.3).....	20
4.1 Case background	20
4.2 Methodology	21
4.3 Findings.....	21
4.4 The system of evaluation.....	23
4.5 Tensions and their dynamics	24
5. LUQ – The process of accreditation of regional museums (case 4.4)	26
5.1 Case background	26
5.2 Methodology	26
5.3 Findings.....	27
5.4 The system of evaluation.....	30
5.5 Tensions and their dynamics	31
6. Pairwise comparison - cases 4.3 and 4.4	32
7. Topic-level discussion	34
References	36

Introduction

In the overall rationale of the UNCHARTED project, cultural information systems serve the crucial role of providing input to policy makers, who on that basis try to steer the actions of organizations in the field of cultural production and heritage management. Thus, the construction of cultural information systems and their implementation marks a key phase to understand how change is triggered and what values inform transformation in the field of cultural production and heritage management. In this report, cases are grouped in pairs to have a better understanding of these dynamics.

The first group of case studies (4.1 and 4.2) looks at the statistics published annually by regional and national observatories to detect the kind of information collected, especially their quality and availability. Taking a perspective focusing on the content, they analyse how cultural observatories structure and categorize their datasets, in order to clarify to what degree a plurality of values is covered. The main methodological approach for this pair of cases is documentary analysis, supplemented by interviews to understand the rationale behind thematic and methodological choices.

The second pair of cases (4.3 and 4.4) looks at the ranking of municipal (Norway) and museal (Italy) cultural production and consumption. This pair of cases takes a processual perspective to focus on the activities and value-related issues taking place once the data collection strategy is designed and rolled out. By taking such a process perspective, this second pair of cases allows us to understand the initial configurations of values underlying the implementation of information systems and how these are re-conciliated, along with underlying tensions and conflicts between the different actors. The methodology for this pair of cases is observation, both participant (4.3) and non-participant (4.4), in-depth interviews and documentary analysis.

Both pairs of cases illustrate how the value dynamics and the plurality of values detected in WP1 and WP2 are here combined into uniform criteria of evaluation. As stated, Topic 4 analysis will be enriched by insights coming from cases performed under Topic 1- 3. On that note, this Topic can provide valuable insights for WP4 analysis, since all the cases described above fit well into a preliminary assessment of how cultural values can inform cultural policies at both regional and national level.

In the following, for each case, we introduce its background and the methodology used. Then we move on to detail the valuation grammars identified and types of tensions in valuation and their dynamics. Due to the structure of the first two cases, the valuation grammar here is detailed through the concepts of data availability and quality, categorization, and scope of coverage. In the second pair of cases the valuation grammar is detailed after the findings, through the sections focusing on the systems of evaluation, the tensions and their dynamics. After each pair of cases, a pairwise comparison analyzes the key issues emerging from the cases. In the final section we compare the underpinnings of these pairwise analyses by discussing critically their topic-level implications.

First pair – Content aspects

1. Cultural statistics in Norway (case 4.1)

1.1. Case background

The fundamental cultural information system in Norway is managed by the national agency of statistics, Statistics Norway (SN). SN was established in 1876, and the agency is formally organized as an entity under the Ministry of Finance. Their work is regulated by the Statistics Act (full name: Act relating to official statistics and Statistics Norway), in which the first paragraph describes the general aim: “to promote the development, preparation and dissemination of public statistics, that might contribute to public education and support analysis, research, decision making and general debate”. The Act also describes the main tasks and the general independence of the agency: It is the central authority for public statistics, it shall also conduct research and analysis, and it should be professionally independent in performing their work as described by law.

SN has collected, aggregated, analysed, and presented numbers to describe the cultural sector for more than five decades. This kind of cultural statistics is and have been based on a combination of survey data and different kinds of registry data, often aggregated from the reporting of cultural institutions. Broadly speaking, this kind of statistics have been covering cultural participation, cultural interest, cultural production and different aspects of cultural economy.

In this case, we are primarily interested in the statistics aiming to capture different kinds of cultural participation. The primary sources for this kind of statistics are the regular studies *Norwegian Cultural Barometer* and *Norwegian Media Barometer*. Some information is also to be found in the annual publication *Norwegian Cultural Statistics*. These two barometers have been produced since 1991, and a continuous challenge have been to identify categories and questions that reflect the actual participative practice among the Norwegian population. The most recent version of these publications is *Norwegian Cultural Barometer 2021*, published in late May 2022.¹ The main point of these publications is to systematically measure the cultural and media consumption of the Norwegian population. A large part of these studies consists of tables and overviews of the number of visits and patterns of use of different forms of cultural and media provisions. The publications from SSB on culture and media consumption are key sources for policymakers at the national level, including the Ministry of Culture.

There are two different consumption studies involved here. A pertinent and continuous challenge is to separate constructively between what should be measured in the two different barometers, which is becoming increasingly challenging as cultural consumption is turning digital.

As these different statistical reports set out to measure actual use of different cultural offerings, as well as interest in and active participation in different activities, there is also a diversity of statistical information in the published studies. The point of interest for this specific case study is what sort of cultural values that are attempted to be quantified through this kind of information systems. Building upon (some of) the cultural values identified in the UNCHARTED work packages WP1 and WP2, to what degree is national cultural statistics as represented by the efforts of Statistics Norway a source for information about such values? And to what degree can statistics potentially capture the broad scope of values attributed to cultural participation and production? These questions will be answered through an analysis of the development of categories, methodologies and questions in the Norwegian culture and media barometers from 1991 until 2022.

The research questions asked in this case study are the following:

- What sort of values are quantified and measured in cultural statistics and how does the questions, categories and classifications affect the perspectives of measuring values?

¹ https://www.ssb.no/kultur-og-fritid/kultur/artikler/norsk-kulturbarometer-2021/_attachment/inline/10aa6cb1-a8d4-4996-979e-752b7247f461:483529cfa9aeb1e67260503bbadcdadfb77ef431/SA170-Kulturbarometer2021_web.pdf

- To what degree can statistics capture the broad scope of values attributed to cultural consumption, participation, and production?

1.2 Methodology

This case is primarily based on document analysis. The primary empirical data consists of the seven editions of The Norwegian Cultural Barometer between 1994 and 2021 and selected editions of The Norwegian Media Barometer between 1995 and 2020. Some information from the recent publication Norwegian Cultural Statistics 2020 has also been used in the analysis. All reports, as well as selected datasets, are published on the website of Statistics Norway (SN). The analysis of the statistic reports, categories and attached values is complemented by an interview with representatives from SN. We conducted one interview with two employees responsible for the cultural and media barometers and general cultural statistics in Statistics Norway. The interview was conducted the day after the publication of *The Norwegian Cultural Barometer 2021*. The main topics of this interview were the organization of their work as part of the agency, the development of surveys and statistics, including the choice of categories and questions, the measuring of quality and value, and the relationship between cultural policy and cultural statistics.

The questions and the perspectives of this case are inspired by previous analyses of the politics of statistics. In the academic debate on culture, the number of studies of cultural categorization and statistics seems to be rather limited, with a notable exception in Paul DiMaggio's paper *Classification in Art* (DiMaggio 1987). This case and the analysis assume that categories are *productive*, in the sense that they have direct consequences for the knowledge produced. An analysis of the development and history of Statistics Norway describes this facet in this way: «There is power in the hands that count. This is visible in categories used to sort and classify, but maybe first of all in the decisions on what to count and investigate” (Lie and Roll-Hansen 2001). The point is that statistical categories might affect how the cultural sector is perceived, administered and governed by cultural policy.

In our case, our analytical interest lies in the relation between statistical measurement methodology, statistical categories and cultural values. What values are being measured (or not) in what way, and how do the questions, categories and classifications of statistics affect the perspectives of measuring values? A sub-question asked, is: When and in what fields do categories change? In order to answer these questions, we have systematized and summarized the different categories, questions and classifications in the culture and media barometers. The guiding indicators for this systematization have been 1) *Category* of cultural or media expression (e.g. Radio, Cinema, Theatre), 2) *Sub-category* (e.g. Paper/Web, different audio media), 3) *Type of cultural provider* (including specific providers, and only applicable for selected categories, e.g. different kinds of newspapers), 4) *Genre* (content-based distinction between different forms of expressions, e.g. book genre, TV-show category, concert genre), 5) *Measurement/unit* (e.g., minutes spent, number of activities, use/non-use). In addition, we have also looked at what participant variables were included in the surveys. These were gender, age group, education level, geographical region, level of urbanization, income level, household type and household size.

1.3 Analysis and discussion

The main focus of this section is to look at what is being measured, how it is measured, and, consequently, what is *not* being measured in these information systems. As an integrated part of this, we will also comment upon the possible tensions inherent in the measurement practices, as well as upon the attempts to resolve these tensions.

Data availability and quality

Statistical Norway is a national, governmentally sanctioned and supported agent, and the main actor of the cultural statistics in question. As referred to above, the majority of governmental papers and strategies on cultural policy relies on data from SN to outline the current status in different sectors. Their work and their autonomy being regulated by law. Another important actor is the Ministry of Culture, being the commissioner and financier of the culture and media barometers. As mentioned by the responsible employees in SSB, the development of their surveys and questions is also the result of a close cooperation with representatives

from the field of cultural production, a third category of actors involved in the work on statistics on culture. Given the role and ambition of SN, their statistics cover the entire country, and they aim to present numbers that are representative for the entire population. For the survey-based part of their publications, this means that their statistics are based on background information on *age, gender, education level, geographical region, level of urbanization, income level, household type and household size*. Statistically speaking, the available information is seemingly of a high quality. While there are no explicit consistencies with supranational information systems (e.g. Eurostat) in the two barometers, a few selected categories in the general Norwegian Cultural Statistics publication comply with European standards. This pertains especially to definitions and divisions for cultural industries, including NACE codes.

On May 24, 2022, the first cultural barometer in five years was published by SN. This publication is of particular interest, as it has been developed during an extraordinary situation for the cultural sector, with the pandemic's evident and fundamental effect on cultural consumption.

In this publication, we see that there are a couple of fundamental changes in this publication relevant to this analysis. Firstly, there are a number of *digital* cultural provisions included in the statistics for the first time: Digital concerts, digital theater, digital library services etc. Secondly, the survey has also asked respondents questions of whether they would like to use certain cultural provisions more, and of the reasons that they don't.

Categorization

The issue of categorization is a core issue in this case, and this issue relates especially to what extent the categories and the development of these have the capacity to capture and incorporate different and possibly emerging values. This includes the potential tensions between different values.

The systematization of categories and classifications in the two statistical publications clearly reveals a rather complex and multidimensional system of cultural consumption measurement. On a general level, the effort by the agency to capture the width of cultural and media consumption in statistically useful categories, is a good illustration of the many axes of information that needs to be included to capture cultural value. The most fundamental distinction is on the *Category* level, where the statistics and the publications distinguish between broad categories of consumption and/or sectors. In the Cultural Barometer, these categories are *Cinema, Sports events, Public library, Art exhibition, Theatre/musical, Museum, Popular music concert, Classical music concert, Ballet/dance and Opera/opera*. These categories change and evolve rather slowly. Since the first version of the Cultural Barometer in 1991, there have only been two major changes to this categorization: 1. In the 2008 version of the barometer, the categories of Popular music concert and Classical music concert were merged into a single Concert category. (While this change might be introduced to make data collection easier, as it is challenging to separate between genres, it also coincides with a general cultural policy turn in including popular music in its realm.) 2. From the 2004 version, the Cultural Barometer included the category of Cultural Festival. In the most recent version of the Media Barometer, the equivalent categories are *Newspapers, Television, Radio, Audio media, Weeklies, Books, Magazines, Cartoons, PC, Video media, Internet, Digital games and Cinema*. Not surprisingly, these categories have been subject to more frequent revisions and changes. The following is an illustrative list on some of the added categories in the Media Barometer publications:

- 1995: Cartoons, Home computer use
- 1996: Teletext (later to be removed)
- 2000: Internet
- 2002: Cinema
- 2007: Computer games, Telephone use

In addition, a number of changes have been made to the types and genres of media consumption included. This includes lists of specific media providers, e.g. radio channels, TV channels and newspapers. Furthermore, there are also genre classifications within the different sectors, e.g. for cinema, Comedy, Crime/Action, Social Drama, Historical, Love/Romance, War films, Science Fiction, Disaster/Horror, Western.

In the category of Concerts, the last versions of the Cultural barometer have used this list of musical genres to differentiate between different concerts: Classical/opera, Church music, Contemporary Music, Pop/rock, Marching bands, Jazz, Singer-Songwriter, Blues, Folk/Ethnic music, Country, as well as the broader category of “Containing choir singing” (!). The use of this kind of genre classifications has changed more rapidly (Danielsen 2006).

On another level, the use and frequency of use of different cultural provisions is a topic for the statistical investigations. Across the different main categories of cultural provision and the sub-categories of genre, the statistical surveys cover the use of cultural provision across different time spans: in an average day (like in the Media Barometer), during the last 3, 6 or 12 months, as well as the number of visits. The Media Barometer also request respondents to estimate the time spent in using different kinds of media. In other words, the value of cultural participation and/or consumption is attempted captured through sheer numbers of visits, views and minutes. Although not explicitly stated in these information systems, we think it is fair to say that cultural consumption is considered to be inherently valuable in itself. Furthermore, there are also different questions on the *access to*, the *interest in* (very/quite/a little/not interested), and the active *participation* in different forms of culture. (E.g.: Do you play an instrument?).

The questions regarding interest represents an interesting attempt to measure the relationship between citizens and specific cultural expressions, or, put otherwise, to measure the cultural values/valuation of citizens. These questions were, for unknown reasons, not included in the 2021 version of the Cultural Barometer (SN 2022). To illustrate the results: In the 2016 edition (SN 2017), the barometer showed that concerts was the cultural provision being valued the most, with 22 percent stating to be *very interested*. On the other end of the scale, only 2 percent said they were very interested in opera, and as many as 67 percent described themselves as *not interested*.

Part of the question here is whether “current information systems are able to grasp tensions between changing values”, to quote the UNCHARTED project description. An important part of the relevant changing values here, or rather, potential modes of valuation, relates to technological developments in general, and digitalization in particular.

Previous studies, including studies done by Telemark Research Institute, show that the challenges of measuring cultural consumption among children and youth are especially pertinent. Young cultural consumption and participation is increasingly digital and screen-based. This means that cultural consumption is dominated by the cultural consumption that takes place on digital platforms. According to the Norwegian Media Authority, the vast majority of 9–18-year-olds play some form of computer game daily, and among boys, the proportion is 96 percent (Norwegian Media Authority 2020: 92). Furthermore, nine out of ten 9–18-year-olds use YouTube daily (ibid.). This platform is both a social medium, a channel for creativity, a place for sales and commercial activities, a place for training and instruction, a reference for information and more. The same can be said about other social media platforms such as TikTok, Snapchat and Instagram. This type of cultural consumption/participation (the distinction itself is questionable in a digital, platform-based context) and these platforms present challenges for both cultural statistics, cultural research and cultural policy: Cultural policy tools and cultural statistical categories must be developed to reflect actual digital cultural use and cultural production. In order for cultural statistics to be able to provide a detailed and comprehensive picture of the population's cultural consumption, these statistics needs to evolve. For example, a statistical division of different forms of cultural expression or genres should make sense for the cultural consumers themselves.

Today's cultural use is characterized by the fact that one and the same cultural expression can be experienced on different units and different digital platforms, that new dissemination platforms are added, and others disappear, and that new forms of expression are also developed. How should cultural and media statistics relate to such a digital cultural complexity? In line with the statistics' balance between precision and continuity, a key question is whether measurements of cultural practice should primarily be developed on the basis of cultural policy categories and cultural management traditions, or whether it should to a greater extent reflect cultural practice? (Danielsen 2006).

A more specific tension is between the statistics on media consumption and the statistics on cultural consumption. Like the policies that regulate these sectors, these surveys are based on a division of "culture" and "media" as the subject of separate investigations. This applies to the Norwegian Culture Barometer and the Norwegian Media Barometer, which are central and widely used sources of knowledge about Norwegian culture and media consumption. The cultural barometer measures selected forms of cultural participation and consumption, while the media barometer measures the use of various media, including both traditional and new, digital, media.

There are considerable overlaps between the two publications, and the most recent publication strengthened this tendency. Reading books, going to the cinema, listening to music, watching theatre etc. are covered in both surveys. The inclusion of digital streaming of culture in the Cultural Barometer makes it increasingly challenging to see what should be included in the categories of "culture" and "media".

Scope of coverage

A fundamental question is to what extent the information system(s) described in this case aims and succeeds in incorporating a totality of (social) values attributed to cultural consumption and participation. These values would include, e.g., values like well-being, collectivity, aesthetical experience, entertainment, fulfilment etc. Unsurprisingly, there are few or no attempts in these varieties of cultural statistics to capture these cultural values. There are, arguably, certain traces of these values in the questions regarding interest and active participation (playing an instrument, taking part in performances etc.). The former questions explicitly ask for the personal valuation of different forms of culture, while the latter more implicitly maps the role of active cultural participation in citizens' lives. At the same time, there are no questions in the distributed consumption surveys that ask for information on the emotional aspects of consuming culture and media, which arguably is the core of consumption values.

Related to this, we would argue that the fundamental tension for this case is the one between quantitative measurement and qualitative cultural values. One of the interesting questions to ask in the closer analysis of the development of the cultural statistics is whether there are attempts to resolve this tension. That was also a topic for the interviews with the representatives of Statistics Norway. Another relevant tension might be apparent in the inclusion of different categories of cultural expressions in governmentally sanctioned statistics. Being included in (or excluded from) public statistics is also a certain sign of political recognition.

In addition to this, there is also a general and fundamental statistical tension also relevant in this case: between relevance on the one hand and the importance of comparable time series. This constitutes a tension in statistical work that might prove difficult to resolve. In effect, this represents two different and potentially incommensurable quality parameters for statistics. In a field that develops quickly, like in the field of digital cultural provision and participation, this is especially challenging.

One of the interviewees from Statistics Norway commented upon the issue of measuring quality in the following manner:

It is of course very demanding to find and ask good questions and the most difficult thing is that there is so much of the value of cultural life that is difficult to measure with numbers. There is something about quality assessment.... And what is good quality within culture? What is the value of culture to an individual? It is a huge challenge that we never stop discussing - how we value culture. It is not necessarily the case that because something is well attended, it has great value. Yes, it is very demanding, measuring value. I'm from the humanities, so for me it's a bit like... I'm very, very tempted to cross over to the qualitative all the time, but here we work in a quantitative industry, so we have to try to figure it out. There is something about finding good numbers then, but I become like, "can we not just write, or somehow use other parameters?".

The publication of the cultural barometer in May 2022 included some attempts to resolve the tensions between analogue and digital cultural consumption, made especially relevant by the increase in digital cultural consumption during the pandemic.

2. Regional cultural observatories in Italy (case 4.2)

2.1 Case background

Regional cultural observatories monitor cultural phenomena in a specific territory by collecting and analyzing data of different kinds. However, the organizations that define themselves as "observatories" do not share the same characteristics: there is, in fact, a great variety in terms of organizational forms, governance structures and activities performed. Correspondingly, other organizations do not define themselves as "observatories" but carry out monitoring activities of cultural phenomena.

In Italy, the origin of regional observatories is closely linked to two events:

- The establishment of the regions in the 1970s has led public institutions to become gradually more interested in the cultural phenomena that occurred in their territory.
- The enactment of a law for the financing of performing arts has led to the foundation of the National Performing Arts Observatory (1985), which subsequently became a prototype for other similar organizations.

Between 1985 and 1989, the first Italian local cultural observatories were established. Although they were partly modelled upon the National Performing Arts Observatory, they were highly heterogeneous in terms of the activities performed and the territorial focus of their analysis – i.e., municipality, province, or region. For example, cultural observatories were founded in the Municipality of Lecce (1986), the Province of Trento (1987) and the Lombardy Region (1988).

Between 1990 and 2000, there is a moment of settlement that allowed the stabilization of some monitoring practices, the legal recognition of some organizations through the issuance of regional laws (e.g., in Lombardy) and the foundation of two new observatories that will stand out for the abundance and continuity of their work: The Performing Arts Observatory of the Emilia-Romagna Region and the Cultural Observatory of Piedmont.

The year 2001 represents a landmark for the Italian political context when a constitutional reform assigned greater responsibilities to local authorities, including cultural sector-related tasks. This event entailed a relaunch of regional cultural observatories, including the foundation of new observatories in the regions of Marche, Friuli-Venezia-Giulia, and Apulia.

The case study aims at exploring the current situation of regional observatories of culture in Italy, to grasp the variety of approaches and internal structures that these organizations adopt for measuring and monitoring cultural phenomena. This analysis allows for shedding light on the quality of data used, the categorization efforts exerted, and the scope of values captured by these local cultural information systems.

2.2 Methodology

Our work started with online exploratory research aiming at mapping regional observatories throughout the Country. This overview resulted in a heterogeneous list of 27 organizations. We classified each observatory by specifying the region where it is located, whether there is a website or not, and the sectors or subsectors analyzed – i.e., performing arts, museums, culture in general. At this stage, we were particularly interested in understanding data availability, in terms of the reports available on their website. Following this preliminary analysis, a sample of 8 observatories was selected, according to the following criteria:

- Availability of reports in their websites: thanks to this process, we filtered the observatories sharing the results of their activities on their official websites. Surprisingly, many observatories did not publish anything online.
- Geographical distribution: we choose one observatory for each region to preserve the representativeness for most of the Italian territory.
- Focal sector: we preserved the diversity of the area of interests, including in our sample both "general-purpose" observatories and "subject-oriented" observatories.

This first set of criteria led us to focus upon the cultural observatories of 8 regions: Apulia, Emilia-Romagna, Lombardy, Marche, Piedmont, Sicily, Tuscany, Veneto. However, through a second analytical step, we decided to further reduce the sample for the sake of focusing our research only on currently active observatories who produce or re-elaborate data which is then made public. We therefore excluded from our analysis: i) the observatory of Sicily which does not report on cultural data but is more akin to a census service for artists; ii) the observatory of Veneto that produced only thin documents containing scarce information; iii) the observatory of Marche which does not produce its own reports, but only refers to data produced by a different statistical division of the regional administration; iv) the observatory of Lombardy that is inactive since 2015.

We therefore decided to focus our analysis on four observatories that disclose a good amount of data and that give some degree of representativeness of the sample, namely those of the regions Apulia, Emilia-Romagna, Piedmont and Tuscany (see table 1 below). Among these, Apulia's and Emilia-Romagna's ones are observatories on the performing arts, while Piedmont and Tuscany are general-interest ones. This subsample gives us therefore a sufficient degree of variance on the kinds of regional observatories.

Region	Apulia	Emilia-Romagna	Piedmont	Tuscany
Documents	3	56	67	11
Pages (tot)	83	~2500	~4000	507
Time span	2019-2021	1999-2021	1999-2021	2015-2021

Table 1. Documentary data sources of observatories in the final sample

We analyzed the latest periodic reports published by the observatories of our subsample and mapped all the data items (tables and figures) therein displayed. Through a coding procedure, we aggregated the data items according to a number of categories. For example, under the category "cultural participation" we included all those data items that refer to the extent to which citizenship is variously involved in cultural endeavors (e.g., number of visitors in museums, percentage of citizens who have read a book in the past year). After having outlined all these categories, we noticed they refer to three main areas of investigation (topics) of WP3: Cultural Consumption; Cultural Production; Cultural administration. We used this classification to interpret our fine-grained analysis of categories and reconnect them to broader frames of meaning.

To further inform and imbue with sense the analysis of the data published in the reports, we also collected interview data thanks to six informants who have a responsibility role in the observatories of our final sample (average interview time: 75 minutes)

2.3 Analysis and discussion

We first briefly present the features of the four selected observatories focusing on institutional aspects, activities, and data collected. Then we compare them based on our three focal issues: data quality and availability, categorization, scope of coverage.

Overview of the selected observatories

Performing Arts Observatory of the Apulia Region

The observatory was established in 2008 through a regional law on the public financing of cultural activities in the performing arts. The observatory is part of the regional department of Culture and Tourism, and its members are representatives of the Unions, Municipalities, and Provinces of the region.

The report considered was published in 2020 titled: "Financing of the performing arts in Apulia" (40pp).

Our analysis of the individual data processing revealed that the entire information published by the observatory revolves around the public financing to the performing arts sector. The data presented, for

example, show the distribution of funding on the regional territory or the quantity of resources assigned by type of beneficiary.

Performing Arts Observatory of the Emilia-Romagna Region

Like the Apulia's observatory, the Emilia-Romagna observatory was established through a regional law for the financing of cultural organizations in the performing art sector (1999). However, its current governance mechanism involves the joint participation of the regional administration and a foundation on the performing arts participated by public bodies (regional administration and municipalities).

The observatory produces an annual report that integrates different periodic monitoring activities into a single document: the main topics they address revolves around employment, demand and supply, venues, and financing of the performing arts sector.

This periodic monitoring activity performed by the observatory is complemented by a conspicuous number of ad-hoc studies (28 studies produced since 1999): research works aiming at capturing specific phenomena through a multitude of methodological approaches that combine quantitative and qualitative data. For example, in the years they have produced reports on multiculturalism, theater and mental health, the role of bank foundations in financing the performing arts, etc.

Cultural Observatory of the Piedmont Region

Founded in 1998, it is managed through a public-private partnership involving the regional administration, private foundations, a regional research center, and the chambers of commerce of the region.

Every year, the observatory publishes two documents, containing internally produced data on museums and cinema, and a broader report, which integrates these data with external information to give a general overview of the cultural sector.

In parallel with this ordinary periodic monitoring, the observatory carried out ad-hoc studies. In total, there are 60 studies published from 2000 to 2021 on the observatory's official website, focusing on various cultural sectors. The nature of these works is mixed, containing both quantitative and qualitative analysis. For example, there are in-depth analyses or longitudinal analyses of specific sectors (i.e., on cinema or the cultural heritage of Piedmont); studies on specific phenomena or territories (e.g., young theater companies, cultural projects on Alpine valleys); impact studies on specific cultural events.

Cultural Observatory of the Tuscany Region

This Observatory was established in 2015 as a division of the regional institute for economic programming (IRPET), a public research center focused on economic matters. It doesn't publish periodic monitoring reports as the others: the main periodic output of the activities is published on its website, where quantitative data can be consulted through interactive tables. The site is divided in two parts: one concerns data about all the Italian regions, another focuses on Tuscany territory: our analysis takes into consideration only the latter.

Data items about culture in Tuscany are nineteen and refers to the 2015-2017 timespan, sharing information about every municipality of the territory.

Besides this collection of quantitative data, the observatory also produces notes and reports on ad-hoc studies. It publishes two kinds of documents: notes and monographic studies. The former are short reports of small-scale studies or re-elaborations and interpretations of existing data; the latter are instead more detailed analyses sometimes performed in collaboration with private research centers.

So far it has produced 11 documents and 8 of these have been published since 2020. Understandably, most of these documents revolve around the impact on the cultural sector of Covid-19 pandemic.

Comparative analysis

Data availability and quality

The regional observatories in Italy are linked to the regional administrations with varying degrees of closeness. In general, they work as research centers focused on the cultural sector.

In some cases, they collect and produce their data, which are nonetheless always complemented with other, external, data sources. For instance, the main source of data about cultural consumption is SIAE, the Italian copyrights collecting agency, while the main source of employment data in the cultural sector is INPS, the Italian public retirement institute. Moreover, a large amount of data reported in the periodic monitoring activity come from ISTAT, the national statistical system, and directly from the local administrations.

A large part of the work of regional observatories is thus related to the harmonization of all these data coming from different sources. Our informants confirmed that this is a necessary but critical exercise because it is hard to match different data and compare them. This difficulty may potentially undermine the quality of the observatories' work:

Our effort is to recollect the data coming from the local administrations and make them compatible with the more standardized statistical systems of ISTAT or INPS [...] We need to objectivize all these data and make explicit their interrelations, because one of the potential problems of these regional observatories is that they may become self-referential when using data produced by the local administration, which would be useless to make comparisons across regions and grasp the phenomena in a precise manner. [*representative of the Tuscany regional observatory*]

Sometimes these measurement and comparison struggles are explicitly stated in the textual part of the periodic reports. For instance, Emilia-Romagna's section on the analysis of economic performances of performing art ventures include a methodological note where these efforts are clearly spelled out:

In the case of revenues, it is a matter of recomposing a very articulated picture, given the heterogeneity of the financial resources available to the different performing arts ventures to finance their activities, depending on whether they come from internal, public, or private sources. [*Emilia-Romagna report on economic performances of performing art ventures, 2005*]

Oftentimes all this work is not fully disclosed, partly because observatories perform measurement and monitoring that are directly handed to the administrations for their internal use, and partly because part of their elaborations is too complex and unpolished to be published in a compact format.

Another, more general, concern regarding the functioning of the observatories is the issue of their neutrality in elaborating and publishing data. Ideally, the regional observatories of culture would be institutions collecting, elaborating, and analyzing data to support the formulation of cultural policies, independently from the political interests of the local governments. More concretely, the strict linkage between observatories and regional administrations oftentimes hampers the achievement of this independence. Our informants all expressed their concerns on this matter and reported examples of some cases in which data reporting have been instrumentalized to depict a situation which could create a favorable perception of the cultural policies implemented by the regional governments.

In short, a diffused concern about the functioning of the regional observatories is that their activities might be affected by political influences. This concern potentially raises issues on the reliability of the data reported and published.

Categorization

One of the central problems our informants frequently experience during data collection and analysis derives from the ambiguities inherent in defining the boundaries of cultural sectors. For instance, our informant of the Emilia-Romagna observatory stressed that multiple criteria can be used to define the artists working in the performing arts. In some cases, artists are included in the analysis if they have worked even just one day during the whole year; in other cases, a more composite analysis is performed to provide evidence that artists are involved in a permanent work activity.

A precise delimitation of the different cultural sectors is often needed to understand which organizations or ventures can be included in a specific cultural production activity. In some cases, using the ATECO codes – the Italian equivalent of Standard Industrial Classification (SIC) system – does not provide a sensible categorization. For instance, in the performing arts, it might be necessary to distinguish companies that work

in the circus industry (which is a cultural activity) from companies running amusement parks (which is not a cultural activity): these two activities are included in the same ATECO code.

These categorization issues create methodological challenges that can be overcome through additional efforts to deepen the analyses and inform the data reporting activities with insights deriving from more grounded research works. This is mostly why the observatories with a larger endowment of resources (most notably the Emilia-Romagna and the Piedmont ones), along with the periodic monitoring, produce a voluminous number of ad-hoc studies that make use of multiple methods and oftentimes deal with qualitative data analysis.

Scope of coverage

Our analysis of the data items reported in the periodic monitoring activity of the different regional observatories was based on a coding procedure through which we classify each item according to the following categories:

1. **Cultural expenditure** concerns monetary data on the expenditure of visitors, spectators, or users in general of cultural venues or products
2. **Cultural participation** concerns quantifications of visitors, spectators, or users in general of cultural venues or products
3. **Cultural supply** concerns quantifications of openings of cultural venues and number of cultural events
4. **Employment** concerns quantifications of workers in the cultural sector
5. **Public expenditure** concerns quantifications of direct expenses of local administrations (e.g., wages and maintenance costs)
6. **Public funding** concerns quantifications of funds granted by local or state administrations to cultural ventures
7. **Economic performances** concerns quantifications of revenues or value added generated by cultural ventures
8. **Customer-satisfaction** concerns data on user satisfaction on cultural ventures or products
9. **Tourism** concerns quantifications of arrivals on a specific territory
10. **Mixed** concerns data that try to combine different area of interests in the same table or figure

Subsequently, we aggregated these categories according to the main spheres involved in cultural activities (areas of interest): Consumption, Production, Administration. Table 2 below displays the results of this analysis.

What emerges is that observatories vary in the proportion to which they refer to the different areas of interests: Emilia-Romagna is primarily interested in reporting data on production, Piedmont's first area of interest is instead consumption, Apulia exclusively reports data related to the administration level, Tuscany is mostly interested in production but also significantly interested in consumption.

Area of interest Category	Emilia-Romagna	Piedmont	Apulia	Tuscany	Tot
Consumption	16.95%	51.56%	0.00%	31.58%	27.78%
Cultural expenditure	8.47%	0.00%	0.00%	0.00%	3.86%
Cultural participation	8.47%	46.09%	0.00%	31.58%	22.22%
Customer-satisfaction	0.00%	3.91%	0.00%	0.00%	1.21%
Tourism	0.00%	1.56%	0.00%	0.00%	0.48%
Production	55.37%	39.06%	0.00%	52.63%	42.75%
Cultural supply	13.56%	33.33%	0.00%	11.72%	36.84%
Economic performances	28.25%	11.90%	0.00%	21.09%	0.00%
Employment	13.56%	0.00%	0.00%	6.25%	15.79%
Administration	21.47%	9.38%	100.00%	15.79%	26.09%
Public expenditure	0.00%	5.47%	0.00%	15.79%	3.62%
Public funding	21.47%	3.91%	100.00%	0.00%	22.46%
Other	6.21%	0.00%	0.00%	0.00%	3.38%
Mixed	6.21%	0.00%	0.00%	0.00%	3.38%
Tot	100.00%	100.00%	100.00%	100.00%	100.00%

Table 2: Results of coding data items. The percentages in the table's entries indicate the proportion of data items included in the last available report relating to each specific category

This analysis, triangulated with the insights we derived through our interview data allows for outlining a broader consideration on the main value dimensions that are variously captured by the Italian regional observatories. On the one hand, a prevalent attention toward consumption denotes an underlying conception of culture as a marketable activity which can generate economic benefits: although some data relates to the involvement of citizens in cultural activities without explicit economic denotation, in most cases data on consumption are reported as a proxy of economic well-being of the region. On the other hand, attention on the administration level might imply the conception that culture is a public good that can generate broader benefits for the population. Finally, attention on the production level can surely include a relevant economic dimension, but employment and supply of cultural activities are also connected to broader social benefits generated by culture. In short, we identify two main value dimensions around which the attention of the observatories gravitates: *Wealth* creation and *Welfare* creation.

The two dimensions are not antithetic perspectives of how culture can be valuable in society but rather represent the poles of a range of possibilities: each of the observatories we have analyzed falls somewhere in between the two poles. Importantly, both perspectives share the underlying assumption that culture is primarily a productive activity, which may contribute to the overall economic development of the regions.

The Welfare pole is best represented by the Apulia observatory for performing arts: the observatory was established through a regional law aimed at financing subjects and organizations working in the performing arts sector. The Observatory collects all data about the subjects asking for these financing, and about all those instances in which funds were not granted and for what purpose. This entails a viewpoint according to which cultural endeavors are fundamentally resource-demanding activities that must be supported by the policymakers.

The Emilia Romagna Observatory for performing arts also apparently leans toward the Welfare pole. Similar to the Apulia observatory, it was established through a regional law for the financing of cultural activities. They collect and report vast amounts of data on the financial viabilities of the organizations and enterprises that apply for public funding, in fact they are the second in our sample on the quantity of data provided about the Administration level. Some of the *ad-hoc* studies conducted by this observatory also suggest that cultural activities, even when they are unattractive for the general public and may not have a great economic impact (e.g., Opera), must be preserved by the public administrations. Nonetheless, the Emilia-Romagna observatory also provides data about entrepreneurial activities ongoing in the region and the

audience/visitors expenditure signaling that culture may give rise to economic wealth. Thus, they provide a rather mixed perspective.

The most markedly positioned in the Wealth dimension, is the observatory of the Piedmont region. This observatory is the one most decidedly collecting and analyzing data at the Consumption level. Furthermore, in their ad-hoc studies, they often perform impact evaluations of the wealth generated by important cultural events taking place in the region. The underlying perspective is therefore one of culture as a resource-generating endeavor. This distance from the Welfare pole is also evident by the fact that the Administration level is the one receiving the lowest level of attention, among all the observatories in our sample.

The Tuscany observatory is also mostly positioned toward the Wealth dimension. The data they collect and report are however less focused on Consumption than Piedmont but much more than Emilia-Romagna. They are mostly focused on Production activities, and our informant working in this observatory confirmed they conceive culture as an endeavor fostering the economic development of the region. However, they significantly consider culture as a matter to be administered and financed by policymakers. In fact, the attention they put on the Administration level is higher than Piedmont.

This kind of analysis also spotlights an inherent complexity in reporting about the values of cultural activities: while the economic impact of culture can be measured in precise monetary terms, the broader social benefits deriving from cultural activity do not have any easily quantifiable measurement criterion. This is why, for example, the Emilia-Romagna performing arts observatory has invested many resources in performing research works based on interview data or focus groups to make sense of the positive impact that cultural activities may have on processes of social inclusion of migrants or disabled people. Additionally, complex societal phenomena affecting cultural sectors can hardly be appraised through quantitative measurement only. For example, the exceptional condition experienced by the cultural sector during the COVID-19 pandemic seems to have led the Cultural Observatory of Piedmont to a necessary contextualization of quantitative data, otherwise difficult to interpret. In their latest report we can read:

What we have called the end of quantitative, the end of the obsession only with the number of admissions to museums and shows, was promptly confirmed by the prolongation of the lockdown in 2021, which simply cancelled the possibility of indefinitely extending the public of cultural venues. [*Piedmont's yearly report, 2021*]

This quote reveals a critical approach to measuring the value of culture through audience quantification. What the authors of the report suggest is that quantitative data remain useful for comparisons between regions, but they are not enough to fully grasp “the complex of positive impacts of culture”.

3. Pairwise comparison – cases 4.1 and 4.2

The two case studies here compared are representative of different systems for the collection and elaboration of data on cultural activities. The focus of these analyses is the content of the information reported by these different systems.

More precisely, we develop our comparison along the three main issues characterizing topic 4 (see D3.1):

1. *Data availability and quality*: How do cultural information systems address potential inconsistencies between multiple data sources and how are data collection efforts coordinated across different administrative levels?
2. *Categorization*: How can cultural information systems grasp tensions between changing values, capturing recent trends in cultural production and technological advancements?
3. *Scope of the coverage*: To what extent do cultural information systems encompass in their records the plurality of values that have emerged from WP2 (Aesthetic, Economic, Democratic/participation, Sustainability, Emotion-based valuations) in relation to the distinct areas (cultural consumption, cultural production and heritage management, public administration)?

Data availability and quality

The context and the organization of data collection is differing between the Statistics Norway case and the Italian regional observatories case(s). For SN, the geographical scope is the entire country, and the ambition is to use collected survey data to present a representative overview on (i.a.) cultural consumption for the entire population. For the Italian observatories, the geographical scope is the regional one, and the focus of the different information systems varies between different regions, partly due to political priorities and partly due to differences in data availability. What seems to be a certain political pragmatism in the use and collection of data in the Italian observatories is clearly different from the Norwegian case where the works produced by SN do not translate into actionable cultural policies.

On the topic of data availability and quality, there are also differences between the two cases. In Norway, Statistics Norway represents a reputable and long-lived institution that operates at the national level and is endowed with enough resources to collect and elaborate all the data internally. Data availability is therefore not an issue, and the strong institutional ethos revolving around autonomy and independence from political interests guarantees the quality of data. This autonomy is furthermore backed up by law.

The situation among the regional observatories in Italy is more complicated. They are all variously connected to regional political bodies, and, in some cases, they directly depend on administrations. They are usually not endowed with enough resources to collect and elaborate all the data they need, and therefore they typically need to aggregate and harmonize data coming from different sources. This potentially raises issues on the quality of the data reported.

To sum up, we may conclude that the governance mechanisms, legislation, and the organizational infrastructure that support the existence and functioning of cultural information systems play a key role in ensuring the quality and the availability of data.

Categorization

The Culture and Media Barometers mainly report data about cultural consumption and participation. On the one hand the main difficulty is about ensuring the comparability of data collected in different times by granting some degree of consistency and continuity to the categories used to perform the measurement. On the other hand, it is also important that statistics capture broad changes in the cultural sectors. In other words, there is a potentially permanent tension at stake here, between consistent and relevant information categories. Furthermore, partly as a consequence of adjusting to new consumption patterns, the boundaries between sectors and between Media and Culture in general may potentially blur over time.

In the Italian regional observatories, the focus is not only centered around consumption, but around the entire supply-chain of culture, from financing, to production, to consumption. Here the main difficulty

encountered in the categorization process regards the problematic definition of who is to be included in the data as worker in the cultural sectors, and what firms and organizations can be defined as “cultural”.

To sum up, the main categorization issues encountered by cultural information systems regard the highly fluid and changeable nature of the boundaries of cultural sectors/activities: How to distinguish a consumption act of culture from a consumption act of media? Through which criteria can we reliably identify workers and organizations in cultural sectors? How to account for technological and social evolutions that make the boundaries between sectors more permeable? All these currently are open questions that can be addressed through a bricolage approach to produce categories that best capture different definitions of cultural sectors, or through acknowledging that some overlaps between apparently different definitions are inevitable. These categorization efforts are, implicitly, also (e)valuation work, in the sense that to be recognized as a cultural worker or an act of cultural consumption means to fall within the realms of a value-based cultural policy.

Scope of coverage

In Statistics Norway, the focus on participation signals an acknowledgement of consumption of culture as intrinsically valuable. However, it is difficult to capture more nuanced value aspects connected to consumption: well-being, aesthetical experience, entertainment, etc. While acknowledged as a challenge by SN officers, how to measure these facets and whether it should be the task of a governmental agency is an open question. In any case it will be necessary to go beyond purely quantitative approaches to capture the variety of values related to cultural consumption.

In Italy, regional observatories all share the underlying conception that cultural activities are productive activities and may therefore generate both wealth – i.e., improvement in economic performances of the regions – and welfare – i.e., the support of non-breaking even cultural activities by regional administrations. Some observatories are more prone to provide data and measurements reflecting the welfare dimension because of resource constraints rather than a specific political direction. On the other hand, more staffed observatories mix welfare and wealth dimensions. In both cases, observatories are mostly centered around economic-oriented values of culture. To fully account for the many value considerations underlying the data collected and reported, some observatories deem indispensable to perform qualitative ad-hoc research works that, while not allowing for exact measurements, are more suitable to capture complex dynamics in the cultural sector and thus complement the quantitative monitoring.

Capturing nuances of value aspects is a critical challenge that can be addressed by either adding or changing measurement categories or by complementing quantitative statistics with qualitative research works.

Concluding remarks

The pair of cases on Statistics Norway and the Italian regional observatories provides insights on the key challenges pertaining to collecting, elaborating, and reporting data on culture. Our analyses highlight that a transversal problem for cultural information systems is coping with the shifting boundaries of cultural sectors. The fluidity in the definitions of what culture is, and the wide spectrum of values connected to culture create tensions along three different core dimensions:

1. *Capturing cultural participation*: The overlaps and contrasts between the Media and Culture barometers published by SN demonstrate the difficulty in distinguishing the semantic boundaries between what is participation in cultural activities from the use of and/or participation in media services.
2. *Capturing culture as a productive activity*: The focus of regional observatories onto the wealth and welfare dimensions suggests that culture is in all effects a generator of economic activities. However, this facet implies challenges related to establishing a dialogue between cultural statistics and labor or industrial production statistics.
3. *Capturing ongoing trends*: Although traditional information systems rely on statistical methods to collect, measure, and report valuable data, our case studies both highlight the importance of complementing quantitative analyses with qualitative ones. Qualitative data would be more appropriate

to capture all the value nuances embedded in cultural activities, but they do not lend themselves to a compact format compatible with the reporting standards of a functional information system.

Second pair – Processual aspects

4. The Cultural Index Norway. Benchmarking local cultural activity (Case 4.3)

4.1 Case background

In 2011, *The Cultural Index Norway (CIN)*, was initiated by Telemark Research Institute (TRI), to benchmark local and regional cultural activity. The index built on similar concepts developed by TRI, benchmarking regions in terms of labor migration, business climate, etc. Organizing and analyzing available register data from approximately 50 different public as well as private sources, summed up in ten indicators, this new index suggested an “objective” picture of the state of cultural activity within regions and municipalities, in a form that allowed comparison between the different entities. From the start, CIN soon gained popularity, not least because of the ranking of municipalities and regions, a function that triggered heavy media coverage in local press.

According to the researcher initiating the CIN project (hereafter *the CINnovator*), there were two motivating factors behind the innovation that were considered particularly important. In most European countries, regions and municipalities hold an important position within cultural policy. According to the website culturalpolicies.net, municipalities in most European countries account for 30-50% of the total public spending on culture. In Norway, the number is approximately 50%. This important position calls for instruments to systematically measure and compare how the municipalities put their money to work, in terms of cultural output within different sub sectors and genres. Still, within Norwegian cultural policy research, most of the research has been focused on national cultural administration, and the local and regional level has attracted less interest (Berge 2022). According to the CINnovator, the CIN was intended to bridge this research gap. Ranking the municipalities according to a set of relevant indicators, would both enable the municipalities to improve over time in categories where they underperformed *and* produce interest for regional cultural policy. The second motivation was to give a description of the *actual* level of local/regional cultural activity, as opposed to other, competing benchmarking tools or competitions that were based on less systematic, qualitative assessments of local cultural activity or engagement. Pre CIN, indexes were unscientific, often produced by organizations representing regions and municipalities, and often aimed at promoting the success of one municipality rather than critically engaging in how culture is funded, offered, and used.

From the start, the CIN was developed in collaboration with Norwegian municipalities and county municipalities. That means that the index has undergone several phases, and consequently several points of revision, to stay relevant and feasible to the regional cultural sector, hence developing from a crude ranking of the municipalities to a presumably more sophisticated system of cultural information and measurement. Consequently, the index has been widely used by municipalities as a basis for cultural planning and implementation. The index is freely accessible², but municipalities and regions are offered to purchase a report that is more comprehensive and analytically developed.

Since the CIN is both quantitatively based and serves as a basis for cultural planning and follow-up, it constitutes a relevant case for a study of cultural information systems, and how such a system may represent, adapt to, or cater for cultural values. In this case study, we consequently ask:

- 1) *What kind of representations of cultural value may be identified in CIN?*
- 2) *What negotiations have taken place to make the indexes feasible, by whom, and what are the values informing such negotiation?*
- 3) *What tensions arise from such representations and negotiations?*

² From 2019, the CIN became web based (see <https://kulturindeks.no/>).

4.2 Methodology

The study of the CIN case is based on empirical material consisting primarily of documents and qualitative interviews. The document study has examined a sample of 11 CIN reports published in the ten-year period 2011–2021. In addition, this study included a memo from the team that initiated the index in 2011, describing the process of starting, developing and maintaining/improving the index.

Interview data comes from individual interviews with a) the CINnovator, and b) cultural administrators in six municipalities. The six were sampled based on a set of criteria including a history of purchasing or subscribing to the CIN, coming from all parts of Norway, having different degrees of urbanization (from large cities to rural municipalities), holding different demographic profile, and performing differently in CIN (high, mid, low). Each of the interviews with the municipality representatives was conducted digitally, recorded and partially transcribed. The qualitative interview with the CINnovator was done live, recorded, and later transcribed. All interviews were in accordance with GDPR and approved by the *NSD - Norwegian centre for research data*.

Doing research on your own organization and “products” requires extra caution. Firstly, it is important to stress that none of the researchers involved in the study have been directly involved in the development and marketing of the CIN. However, taking the potential risk of biases or reduced validity into account, it is equally important to stress the need for a high degree of self-reflexiveness and transparency in the research process. Here, we lean on methodology from the rich literature of practice-based and auto-ethnographic research (Schön 2017, Nelson 2013).

4.3 Findings

A general project idea is that cultural information systems produce and represent diverse forms of cultural value. Another is that this value is an object of negotiation. To analyze this in the case of CIN, it is relevant to look at how the index and its design has changed, and who initiated such changes. A general finding is that an important objective with the index has been to design and operate a tool that is perceived as useful and flexible among its users. In the following, we outline and analyze how this process emerged in *four* more or less distinct phases:

Phase 1 –innovation: According to the CINnovator, the starting point of the index was pragmatically selecting interesting indicators. In 2011, firstly, the TRI team, which consisted of the CINnovator and two-three more members of the TRI staff, asked themselves how to operationalize cultural supply and consumption/participation at a municipality level based on quantitative data. Secondly, they started to look at what data that was possibly accessible to do so. The approach that was selected included a combination of the two tasks, where every possible local cultural policy measure was followed by a search for possible data. The process in this initial phase can be summed up in four questions or points of reflection to make a stable ground for the index: The first question was how to define culture. Soon, it was decided to limit the scope to an expressive, humanistic cultural definition rather than an anthropological “whole way of life” definition (Geertz 1973, Williams 1983 [1958]). Since the index was developed to be relevant in a cultural *policy* discourse, it was further decided to choose a definition of culture used in studies of cultural policy in general, and in Norwegian cultural policy in particular (Duelund 2003, Mangset and Hylland 2017). This included e.g., performing arts, visual arts, literature, cinemas, museums, and cultural heritage. In Norway and several other countries, cultural policy has since the 70s covered what is often referred to as an expanded definition, i.e., one that includes both leisure activities, voluntarism, and sports. Despite the methodological and definitory consistency it would have given, the team decided to exclude sports. Even though sports are included in local cultural policy to some extent, there are still separate measures and a separate policy for this. In terms of leisure and voluntarism, TRI chose to include only activities connected to an expressive cultural definition, such as those of choirs, school bands and historical societies, and consequently to exclude activities such as those exercised by political, environmental, or humanitarian organizations.

The second question raised prior to monitoring local culture life and cultural supply, concerned the scope of data measuring the specific topics. To be applicable for the purpose, the data had the following requirements: They needed to include a *geographic element*, e.g., that culture *took place* somewhere. This included visits,

events or that something was resident in a certain location. Thus, the data had to have a national scope and cover all municipalities in Norway, or, alternatively, multiple data-sources covering all municipalities had to be available. Also, the data had to be available *annually*, providing new datasets each year, this to secure consistent retrospective time series, something TRI based on previous index experience considered very important.

Thirdly, in order to promote demographic “neutrality”, it was necessary to reflect on how to avoid data that favored either small or large municipalities when displaying them per capita.

Finally, because of pragmatic reasons, the selection was limited to existing data, *accessible* in one or several databases. Even though the selected methodology implied to firstly select cultural topics and then to search for available data, nevertheless, the process came to include producing new, relevant topics during the data search. In some cases, the data was not easy to attach to the existing category and thus singled out as an independent category.

According to the CINnovator, based on these choices, ten general indicators were chosen (see Table 1).

Indicator	Source
Museums	Data from museums on <i>total number of museum visits, number of paid museum visits, number of exhibitions and events.</i>
Concerts	Data from 13 ticket-providers, covering most concerts and ticket sales of a certain size. However, there are differences regarding quality between the different ticket providers.
Performing arts (theater and dance)	Data from a) 30 national theater and dance companies, and b) Arts Council Norway data on the national theater companies. Private theaters, fringe theaters or amateur theaters are not included.
Cinema	Data from the organization Film & Kino, which collects detailed information on the <i>number of movie-displays and the number of visitors.</i>
Library	All Norwegian municipalities have to offer library services. Data on <i>Number of visits, number of lending's, participants on events in libraries and the number of persons lending one or several books during a year.</i>
Art for Children	Data from a) public schools for children's art education: <i>number of teaching hours, diversity of the supply and number of pupils attending one or several courses,</i> and b) The Cultural Rucksack: <i>supply and attendance every year.</i>
Artists	Data from 20 artist organizations in Norway (26 810 members) on total number of <i>musicians and composers, visual artists, performing artists and of authors</i> per municipality, and <i>total amount of grants for artists</i> per municipality.
Cultural workers	Data on three categories of culture employment within the European NACE-standard on level 4, <i>cultural dissemination, artistic production, and media,</i> and includes all enterprises, both public and private.
Voluntarism	Data on choirs, school/marching bands, local history, crafts, and a visual art association.
Grants	Data on various grants from a) <i>Arts Council Norway: art production and art dissemination</i> and b) <i>Frifond</i> (funding voluntary work within youth culture), and c) cultural heritage protection.

Table 1. CIN indicators

Since the first edition of CIN in 2011, the different indicators have remained the same. However, the sub indicators, the choice of data sources and, perhaps most importantly, the way the data has been analyzed and reported, have been adjusted and refined several times.

The general design of the first reports followed the intentions of the project, highlighting how the municipalities scored according to the indicators, individually and in combination. The municipality to score the highest in total, was announced the cultural municipality of the year. Also, in this first edition, a feature

was introduced where the municipalities that ordered the report got to choose five reference municipalities to which they were compared along all the ten indicators of local cultural offer and consumption.

Phase 2 – adjustments: Whilst the first phase from 2011 was constituted of finding indicators and making the ranging/benchmarking tool, already in 2012, several minor innovations followed. As the ranking produced quite a bit of buzz in local and regional newspapers, and the CIN therefore was well established and known, in the second phase, focus shifted towards producing simple reports available to the municipalities that were interested. Hence, dissemination and communication marked this phase. In addition, the second phase also included fixing bugs and attending to comments from both municipalities' representatives and other researchers in TRI. Several adjustments were made to increase the degree of accuracy and capture the diversities of local cultural contexts. Mostly, in this phase the changes were small (bugs mostly being related to data processing errors), and according to the CINnovator, most feedback from users did not lead to changes, but a growing sense of the need for pedagogical communication on how the index was assembled. This increased pedagogic effort, explaining the “whys” of the benchmarking, initiated in this second phase, has, according to the CINnovator, continued up until today, indicating that quantitative approaches to measuring culture are unfamiliar and perhaps even a little disturbing to many actors within the cultural field.

Phase 3 – the academic turn: In 2014, the largest revision of CIN since the start took place. This happened as the CIN was included in a national research project funded by the Norwegian research council, studying why some municipalities and regions score better than others on important growth and attractiveness indicators (businesses, population etc.). Now the level of analysis was heightened, and focus shifted from ranging to what characterizes a successful municipality regarding culture. Also, a new feature was introduced, where the municipalities were ranged according to an expectation value, i.e., a theoretical level given by various demographic and geographic factors. To some extent, this latter innovation answered critical voices that argued that the early version of CIN did not pay enough attention to the quantitative side of local cultural services. Mostly, these were cities that questioned why they did not get paid benchmark-wise for their large offer of culture.

Phase 4 – on the web: The most recent phase started in 2021, as the CIN was made available online. All data are free, but analyses and digital reports are available for subscribers only. Now, the ranking is toned even more down, in favor of the actual data, especially in how the CIN results are communicated to media. An important part of this has been to make clearer how little difference there sometimes is between many of the municipalities, despite the fact that they end up with quite different rankings (the difference between the municipality ranked 186 and the one ranked 286 could be marginal). This was possible by going from a static taxonomy scale to a more dynamic point-scale. According to the CINnovator, this change was initiated by the researchers, based on a sense that the PR-effect from ranking was less needed as the CIN was well established among users. Now focus shifted towards more complex analyses, something that to some degree tends to be overshadowed by the ranking element. Also, this was done to make the CIN more interesting to academic peers and cultural policy authorities, communities where the sentiment traditionally has been that ranking is crude and that it promotes instrumentalism.

4.4 The system of evaluation

Configuration of actors and organizations

The main actors are TRI and the municipalities (and to some extent the county municipalities), with TRI on the “provider” side, and the municipality/ies on the “client” side. Looking at the configuration between actors, we observe that the evaluative system has a direct, dialogue-based nature, where the development of the CIN has acted out in phases. However, external actors and factors such as a Norwegian Research Council funded research project has provided a framework for more advanced and self-reflexive innovation, something that we consider crucial to refine the CIN and give it necessary legitimacy.

Formality, rational elaboration, publicity and salience

Taking place within an academic/scientific framework, and besides making use of qualitative data, the evaluation (benchmarking) procedure holds a high level of formality. However, several semi-qualitative

indicators, e.g., counting well reputed artists residing in a municipality, contribute to open a space for negotiation in terms of the CIN in addition producing a qualitative profiling of the municipalities.

Context and time

The benchmarking of Norwegian municipalities should be seen in light of a tradition where the municipalities have been expected to qualitatively (not quantitatively) deliver quite similar public services regardless of size and population. This rests on a cultural policy that for many decades has focused on cultural democratization (Mangset and Hylland 2017), but also a Norwegian political tradition for decentralization and subsidizing rural areas heavily. Consequently, the benchmarking is often perceived as a tool to make small but targeted steps towards improved local cultural services. In terms of timing, improvements in both the municipalities' cultural services and the CIN's accuracy have emerged slowly and carefully, supported by the academic context framing the benchmarking.

4.5 Tensions and their dynamics

The CIN case contains several tensions. On an overarching level, the CIN is seen as a tool to measure participation and/or use. Since consumption/use recently has grown in importance as a central source of legitimacy of culture (directly) and cultural policy/funding (indirectly), tensions rise on principal grounds. E.g., several of the bureaucrats we interviewed talked about politicians that in budget work found leverage in participation numbers (high or low) displayed in the CIN. When asked if quantitative indicators and data potentially can be awarded too much weight in policy processes, one such bureaucrat answered, "well, yes, indirectly". Indirectly, in the way that such data were rarely used to sustain concrete budget or policy proposals, but rather, according to this informant, made headlines in the local newspaper, and consequently influenced politicians in one direction or the other. We have also detected internal tensions within the municipality administrations/ bureaucracies caused by the bureaucrats using the CIN as a leverage in their interventions with the policy level. Here, they have taken advantage of how quantitative data tend to evoke a sense of evidence, more than perhaps qualitative analyses do.

More specifically, we have detected tensions related to how some of the bureaucrats feel that the index misses certain indicators, that the existing ones misrepresent their specific context, or that the index does not fully reflect the intangible values of culture. One defining quality or feature with the index is that the indicators are mainly quantitative. Hence, they only indirectly measure or represent qualitative aspects of culture. This is something that the municipalities that apply the index are aware of. When we asked them in the interviews what the index measures, a representative of one such municipality e.g., answered:

It measures.... Well, what is measurable? So, numbers. Probably, it also tries to explain the numbers a bit in text form, which goes a little beyond the numbers. Which tend to become somewhat analytical and normative sometimes. But it measures the different fields within culture, sets them up against another and not least sets up a municipality's culture against comparable municipalities and national figures.

And,

Quality, where that comes in, it is difficult... What is measurable is mostly quantity and money and audience numbers and allocations. How good you are at obtaining grants of course may say something about quality, and based on that, we try to interpret the non-measurable. But not everything that counts can be counted...

This view seems to be quite representative of the municipalities we interviewed. Here it might be useful to know that most of these bureaucrats come from, and/or see themselves as some sort of representative of the cultural field. The index directly measures quantitative aspects of the local cultural life, when it comes to qualitative indicators, this is both an area of interpretation and negotiations. Interpretation in the sense that the municipalities translate or mediate the quantitative into qualitative assessments, negotiations in the sense that such translations leave a large room for forming how qualitative interpretations are to be understood. One example is, as mentioned in the citation above, when quantitative numbers on how much national grant allocations artists in one municipality earn is interpreted into a qualitative assessment of the

same artists and thus the general quality in that place. Directly, the indicator does not say anything about quality, but indirectly it does. How specifically and to what degree it represents such quality, is thus an object of negotiation.

Many Norwegian municipalities and regional governments use the CIN tool as a benchmarking and knowledge base for cultural administration and policy making, and the annual CIN ranking of Norwegian municipalities gets significant media attention every year. Hence, many municipalities see the index as a regional competition, in which they aim at strategic goals like e.g., to climb from past year(s) ranking, to improve within indicators where they underperform, or to find the strategically “right” position, where the municipality spends a decent – not too large, not too little – amount of money on culture. This has not changed although the index has become increasingly less focused on ranking. To some extent, the TRI therefore has “lost control” over the CIN and how it is used locally and regionally. As the index has been revised to become increasingly feasible to its users, perhaps most prominently by toning down its ranking feature, that very “gimmick” has shown a remarkably flexible and dynamic tool for local policy making and bureaucratic work. The question remains however whether its quantitative profile has blurred how cultural value is assessed locally.

5. LUQ – The process of accreditation of regional museums (case 4.4)

5.1 Case background

In 2018, the Italian Ministry of Culture announced the creation of a National Museum System (*Sistema Museale Nazionale*, hereafter SMN). The SMN aims at connecting museums and improving their management and overall accessibility (DM 113/2018, art 2.1). The system would include museums selected according to quality standards defined by the Ministry, and covering three areas: Organization, Collections, and Outreach. In addition to rewarding “quality”, the initiative can be seen as an attempt at stimulating not-yet-compliant museums to increase their overall performance and be eventually included in the system.

The aim of this case is to observe how the process is enacted in the Emilia Romagna region. The potential target of the accreditation procedure consisted of about 500 highly diverse museums in terms of size, disciplinary field, and governance. Our study is based on documentary analysis, in order to track the genesis of the quality standards, and non-participant observation during the data verification period, aiming at understanding the process of design and deployment of an information system.

The case study is organized as follows. After the methodological section, we analyze the questionnaire. Then we describe how the accreditation process was designed and implemented at a regional level in Emilia Romagna. The analysis is based on data collected during non-participant observation of the meetings where museum accreditation was decided. Findings are then discussed according to five analytical dimensions: context and time, actors, level of formality, and tension dynamic.

5.2 Methodology

The case study employs a mixed-method approach, combining documentary analysis, interviews, and participant observation.

The documentary analysis focuses on the decrees and regulations issued by the Ministry and the ER region. In addition, reports and academic articles dealing with the LUQ were collected and analysed. We devoted particular attention to the questionnaire originally designed by the Ministry and then adopted by the ER region. We analyzed questions and scores carefully to understand how quality was quantified.

Participant observation was performed in two ways. First, we joined the online meetings on the SPC team in charge of the accreditation procedure. During these collective meetings – nine in total over nine weeks – team members discussed how to implement the accreditation process, interpreting the criteria and dealing with challenging cases.

We complemented observation with interviews with the team members aimed at understanding better what was discussed during the collective meetings. Furthermore, since the team members decided to distribute among them the workload of the analysis of the museums’ answers, we set up online calls with individual team members and asked them to comment on the evaluations they were performing as they analyzed museums’ questionnaires one by one (3 sessions, 15 hours of fieldwork, 69 pgs. of transcription). We also had access to the questionnaires filled out by the regional museums (157) who participated in the survey.

Using the extensive evidence on the process, we adopted a temporal bracketing procedure to grasp the processual nature of the phenomenon (Langley, 1999). Each week was analyzed, focusing on:

- What it is about: the agenda of the SPC team for the given week;
- Issue of the week: the main issue(s) emerging from the previous week and informing the SPC teamwork of the given week;
- Mood of the SPC team: the atmosphere within the team as observed during group meetings;
- Organizational arrangements: the main team-level operational choices taken to keep in motion the teamwork and go ahead;
- Judgment devices: means used to gather information from multiple parties such as questionnaires, personal networks, and expertise (Karpik, 2010);

- Output: the major decision about the methodological steps to take while evaluating.

In the findings section, we will present the outcome of our observation following this week-by-week structure.

To further complement and validate our findings, we interviewed the directors of six regional museums (total duration of 4.5 hrs., 40 pgs. of transcription) after the publication of the list of the museums accredited to the system. The museums identified were different, especially in terms of institutional status, collection type, and questionnaire scores.

5.3 Findings

Designing the evaluation process at the Ministerial level

The idea of creating standards of quality for museums (*Livelli Uniformi di Qualità*, Uniform Levels of Quality - LUQ hereafter) dates back to the 90s and proceeded in fits and starts for almost 30 years, also due to political instability. Just to give an idea, from 2018 to 2022, there were two Ministers of Culture under four different governments, namely Franceschini (Gentiloni government), Bonisoli (first Conte government), and then Franceschini again (Conte and Draghi governments), and two directors of the Directorate General for Museums (Antonio Lampis from 2017 to 2020 and Massimo Osanna from September 2020).

The accreditation process unfolds differently depending on the region and on the ownership of the collections. State-owned museums, including University museums, are credited directly by the Ministry. All other museums – i.e., municipal, regional or private – are accredited by specific regional-level agencies, although through different channels, depending on whether a system of accreditation is already in place at a regional level (DM 113/2018). Museums located in regions with an accreditation system comparable to the national one (for example, Lombardy, Calabria, Emilia Romagna, and Tuscany), can be automatically enlisted to the SMN. On the other hand, regions without such systems should set up an accreditation body at the regional level to implement the procedure.

According to the Ministry, the accreditation process aims at fostering transparency and comparability among museums. As proudly stated on the DG museums websites, the accreditation process will produce a score that anyone will be able to read. Besides, the score can assess improvements over time, following best practices in the field of public services evaluation, like Tripadvisor. A crucial role in translating this ambitious program into practice is played by a Ministerial questionnaire, which quantifies quality by transforming a set of guidelines into a score.

The questionnaire allocates 80 points to the “minimum standards” and 20 points to additional, not crucial standards, which are called “improvement objectives”. The questionnaire also includes “info” questions and checklists, which intend to collect information on the museum's characteristics without assigning any score³.

Limiting our analysis to the questions on the “minimum standards”, we uncovered that:

- The most covered topic is Organization (12 pages out of 22; 91 questions out of 162): this topic concerns information about the museum's legal status, the activities, the management structure and personnel, the building adequacy/requirements, and the accounting procedures. For instance, the questionnaire attributes different scores to each positive answer to questions like: "Does the museum have a statute or guidelines?" (4.00 pts) or "Does the museum produce an annual plan of activities?"(1.20 pts).
- The Outreach section covers 7 pages and 40 questions; here, topics include external and internal signaling devices such as labels, room numbers and planimetry, educational activities, and relationships with the public, territory, and stakeholders. Points are awarded if, for instance, a museum

³ We received the Ministerial questionnaire from SPC. Of note, the Directorate General for Museums has never published the questionnaire in the website where all decrees and regulations relating to the accreditation process are published.

has “essential tools of information and guidance” (2.4 pts) or the “activities towards the surrounding territory mentioned in the museum programmatic documents”(0.80 pts).

- The Collections section covers 3 pages and 31 questions and considers basic principles in managing a collection: conservation, handling, acquisitions, entry and exit, research activities, and deposit management. Questions in this area ask, for instance, “Is there a person in charge of the handling activities?” (0.80 pts), or “Does the museum produce an inventory document with progressive and unambiguous numeration?” (1.60 pts).

It is crucial to highlight two of the questionnaire’s characteristics. First, most of the questions are procedural and are limited to assessing the presence of processes and rules, regardless of their effectiveness. In other words, points are awarded if a museum has a plan of activities, irrespective of their numbers and, quite paradoxically, quality. Second, the scoring system is somewhat ambiguous. In its official communications, the Directorate General for Museums states that a museum is accredited if it receives at least six points out of 10. However, the questionnaire is organized on a 100-point scale, where “minimum standards” add up to 80. It remains, therefore, difficult to understand the threshold that allows a museum to enter the national museum system if any.

Unfolding the evaluation in Emilia Romagna

The case of Emilia Romagna, which is the focus of this paper, is somewhat peculiar. Despite having implemented an accreditation process earlier (112 museums received a quality stamp in 2012) and being, in theory, exempted by the Ministerial accreditation procedure, the regional administration decided to re-run the accreditation, as the results of the 2012 survey were deemed to be outdated. Furthermore, the regional administration saw the new accreditation process as an opportunity to update the available data on regional museums and create, to use the words of regional officers, a “snapshot” of museums in the region. To perform the evaluation, the Cultural heritage service (*Servizio Patrimonio Culturale*, SPC hereafter) – the Emilia Romagna regional department in charge of the accreditation process – decided to adopt a questionnaire designed by the Ministry, which translates standards into a set of questions, complemented with scores.

In Emilia Romagna, the agency in charge of the accreditation process was initially the Cultural heritage institute (IBACN). IBACN, a regional, semi-autonomous entity, was absorbed in 2020 by the Regional Cultural Heritage Office (*Servizio Patrimonio Culturale*, SPC hereafter), a department of the regional administration in charge of allocating regional grants to museums, organizing training initiatives, and monitoring museum offer in the region. The region staffed SPC with almost all former IBACN employees. However, as far as the accreditation process is concerned, the manager and mostly all the members of the accreditation team changed. Only one team member in a non-apical position passed on to the new team.

SPC inherited from IBACN the decision to adopt the ministry-designed questionnaire to finalize the accreditation and introduced only minor changes to the structure of some questions and the scoring⁴. It then issued a call asking regional museums to fill out the questionnaire if they wanted to join the regional - and then the national - museum system. Notably, the call did not specify the scoring system or the accreditation threshold. Instead, it simply stated that SPC had 75 days, about nine weeks, to review the applications and decide on the museums’ accreditation.

In the following pages, we focus on what happened during the assessment period, providing evidence collected during the non-participant observation of the SPC evaluation team’s work, complemented by the individual interviews. In addition, we will highlight the differences between each week, according to the

⁴ For instance, the question concerning core museums’ activities (“Does the museum or place of culture have suitable and adequate spaces for the functions of conservation, permanent exhibition, reception, information, ticketing, services also for people with disabilities?”) had a score of 5.6 points. Other questions concerning cataloging or inventory reported scores between 0.8 and 1.6.

agenda, the emerging issues, the team's mood and organizational arrangements, the judgment devices, and the outcome of the week.

During weeks 1-2, the SPC agenda was informed by the start of the appraisal process. The aim was to examine each museum's answer and decide whether to accredit the museum or not. A total of 214 museums participated, of which 157 completed the whole questionnaire. The mood of the team was quite tense because the questionnaires had yet to be delivered by the data provider although the 75 days of the appraisal process had already started. Although still unavailable, the questionnaires and the associated scores were chosen as the first essential judgment device to evaluate the regional museums.

In these first meetings, SPC decided that the analysis would be limited to the scores relating to the minimum standards (80 points over the total score of 100), thus ignoring data relating to the "improvement objectives." Without much discussion, the team agreed on setting the threshold to be accredited at 80 points. Several reasons were at the basis of this decision. First, as implicitly suggested by the questionnaire structure and scores, this threshold is the sum of the scores attached to minimum standards, so that a museum obtaining 80 points was expected to have full possession of the standards. Second, this threshold was set by IBACN before and then approved by SPC. Third, SPC adopted this ambitious target because it was aimed at mapping "leaders and laggards" in the region. According to them, the data collected could have then been used to design a roadmap for improvement, focusing on the standards that were lacking the most. Furthermore, the team decided to divide the workload relating to the analysis of questionnaires, with each of the four analyzing 40 forms.

In week 3, SPC eventually received the questionnaires and started the analysis. The team's mood was desperate: apparently, only 8 out of the 157 museums had obtained 80 points. The results were poor and unexpected. The team members started to question the validity of the score. Was the questionnaire not comprehensible enough? Or did it not mirror the regional reality?

Making sense of the apparent misfit, SPC elaborated on the idea that the evaluation technology designed by the Ministry rewarded elements that usually characterize world-class museums, which, in Italy, are mostly managed directly by the State - i.e. The Uffizi. As commented by SPC, state museums tick all the questionnaire boxes: they show a high degree of formalization of procedures, they have qualified personnel, they produce their financial statements timely, and they have a website. In such a system, smaller yet "good" museums like the ones scattered around the region are not adequately valued. After acknowledging the flaws within the Ministry evaluation technology, SPC members agreed to perform the appraisal process in a more active way, with several interventions and data iterations that took place from weeks 4 to 7.

During week 4, SPC members tested different thresholds to see how the total number of accredited museums would vary. The team discovered that the 54 museums that scored more than 75 points were "fairly good" and would deserve to be accredited. Although the number of accredited museums increased, the mood was still tense, with a hint of despair, as the office started confronting ER accreditation with other benchmark regions, such as Lombardy, which accredited 200 museums.

During week 5, the SPC team eventually lowered the eligibility threshold to 70. Despite the new threshold, the disorientation among SPC members was still palpable. While evaluating museums, team members expressed doubts over the answers provided. Thus, SPC decided to verify the information contained in each questionnaire by calling museums by phone. The aim of the phone checks was twofold. On the one hand, the team wanted to ensure there were no mistakes in the answers (tick on the wrong response,). On the other hand, the team wanted to qualify the responses, contextualizing them (i.e., museums being too severe in evaluating the quality of their artwork conservation systems as they might not be up to date but still present). Thus, for the upcoming weeks, the SPC team planned to activate its professional network as an additional judgment device, pairing it with the questionnaire evaluating the museums.

SPC direct interventions were possible because, as previously mentioned, the ministerial regulations allowed substantial leeway in managing the "new" accreditation procedure in regions with an accreditation system already in place, like in Emilia Romagna (art.4). In addition, museums in the region were not aware of the threshold, which was discussed only "internally" during SPC meetings.

During weeks 6 and 7, the team proceeded with the evaluation of those museums that scored between 70 and 80 points. The mood was quite positive and the process was unfolding more smoothly than in the previous weeks.

At the end of week 7, the team worked in two different, although interrelated directions. First, the team tried to isolate the questions considered unavoidable and necessary for museums to obtain accreditation. These questions, after intense discussions, were reduced to three: absence of architectural barriers to guarantee access to disabled visitors; presence of safety regulations; opening hours also guaranteed upon request. Second, the team tried to create the final lists according to the newly established criteria (museums scoring over 70), devoting specific attention to museums scoring more than 70 but that did not satisfy the three minimum standards listed above, and those scoring under 70 but that should be included as deemed by SPC good regional museums.

Between weeks 8 and 9, the appraisal process was finalized. During week 8, SPC was not yet satisfied with the new evaluation method as some “eligible” museums had not achieved satisfactory scores and vice versa. Regarding the former situation, the team discussed the case of a museum that had already received the quality stamp by the Ministry due to its distinctive governance and tried to make sense of the mismatch, eventually accrediting the institution. To legitimize the approach and solve the most “difficult” cases, SPC decided to enlarge the team by adding two experts in the regional museums as additional sources of professional expertise. The main aim in this phase was to finalize the list of eligible museums while avoiding blunders, such as excluding important and high-quality museums or including museums of doubtful quality that just performed well in the questionnaire. At this point, the credited museums were 91 over 157. The mood had shifted dramatically. Despite not being wholly convinced by the final list, the team members were proud of their work.

Finally, during the last meeting in week 9, all team members reported their overall evaluation, validated using the referral of the two additional regional experts. They were satisfied with the list, featuring now 102 museums.

5.4 The system of evaluation

Configurations of actors and organizations

The main actors are the Ministry of Culture, SPC, and the museums based in Emilia Romagna. Looking at the configuration between actors, we observe that the evaluative system has an indirect, two-step nature, whereby SPC acts as a broker between the objects of evaluation – the museums – and the final evaluator – the Ministry. As we have seen, this intermediate step is crucial to weaken the quantitative logic put forward by the Ministry.

Formality, rational elaboration, publicity and salience

On the surface, the accreditation procedure seems formal and rigid: museums fill out a survey, collect points, and eventually get accredited if the score is satisfactory. In addition, deadlines are set for museums and regional agencies. Respondents are forced into binary, yes/no questions, with little room for justifying their selection.

However, decrees and regulations both at a ministerial and regional level leave a crucial issue open: the threshold to be credited. By not formalizing a threshold and avoiding making the score public, the Ministry first and SPC second made room for adjusting quantities with qualities and vice versa.

Context and time

Although in the last 20 years there had been several accreditations of regional museums performed by local agencies, these attempts have been sporadic and conducted on an emergent fashion according to local contingencies. The evaluation process presented above is part of the process of creation of a permanent system of museums: by design, the system has been conceived as an open one, with systematic calls for museums willing to be credited. Every year “windows” are opened, so that applications can be submitted and then evaluated through an appraisal process. In this case we describe the very first implementation of

the process, a first step endowed with several mistakes and un-forecasted issues. As testified by our evidence, there was widespread acceptance among the evaluators that, in the long run, the process had to be refined and adjusted, thus leading, year after year, to a better one.

5.5 Tensions and their dynamics

The main tension relates to the encounter of two different logics of evaluation: “comparative ranking”, and “qualifying expertise”. Comparative ranking can be explained as the possibility to use the scores to classify elements as orders of worth and is mainly related to the questionnaire; qualifying expertise is the rare and highly specialized knowledge and can be explained as the possibility to use expertise to qualify synthetic information. The former was brought forth by the nation-wide breadth of the museum system, in which was pivotal the idea of attaching a score to each object to be evaluated, translating the possession of standards into a number; the latter was displayed several times during the appraisal process, and was enacted to invert such translation, i.e. convert the score of each object in a consistent set of information about the actual possession of standards. The possibility granted to SPC to intervene in the process helped to stabilize these values and perform a selective, though representative evaluation within the expected timelines. The whole process must be contextualized by taking into account the features of the decentralized structure of the Italian public administration. What the central State rules is adopted and implemented by regional governments with varying degrees of freedom. Behind the rhetoric of a standardized evaluation process, a fragmented scenario emerges, whereby museums are assessed differently depending on the region where they are based.

Overall, the practice of evaluation is caught between three fires:

- The logic of the questionnaire: although contested, numbers provide, to use words of SPC people, an “objective and rational” justification for the output of the evaluative practice, in our case the list of accredited museums.
- The fact that an analogous evaluative practice was taking place in other regions: what emerged during the fieldwork was the need to produce a list of enough museums compared to the country's “top” regions, namely Tuscany and Lombardy. Failing to do so would discredit the practice, negatively affect the regional administration's reputation and provide a non-reliable comparison of the state of the art of museums in the region.
- The need to rank and be selective because, at the end of the day, not all museums deserve to be accredited.

As shown, the evaluative practices experienced a critical turning point in week 3: the “objective and rational” methodology provided a picture that no one would have expected. In the following meetings, the tension was managed first by making sense of the limit of the questionnaire, deemed biased towards state-level museums and/or incorrectly filled out. Then, additional judgment devices are introduced to counterbalance the quantitative rationale. Namely, professional judgment and knowledge about museums in the territory are added to the discussion. In parallel, SPC works on the questionnaire by narrowing down the number of standards needed to “pass” the selection. The team converged on a shorter set of quantitative criteria, complemented by a richer understanding of the regional museums' qualities, to produce the finalized list. Eventually, numbers are adjusted to make SPC's experience-based selection legitimate also from a quantitative point of view.

Rather than abandoning the questionnaire, SPC bridges the (e)valuation of quality as envisaged by the Ministry and the information provided by regional museums at the local level to eventually achieve a satisfactory outcome. SPC reasoning results in a list of accredited museums in line with the Ministerial logic, similar in numbers to the lists of other benchmark regions, and consistent with SPC knowledge of the territory.

6. Pairwise comparison - cases 4.3 and 4.4

By taking a process perspective, the second pair of cases explored in this deliverable focuses on the use and impact of information in decision-making within the cultural sphere. It explores the design of information systems and whether value tensions during their implementation. The primary methodological approach in these cases is participant (4.3) and non-participant (4.4) observation.

The two cases show a strikingly similar feature in that both information systems are designed, more or less explicitly, as self-regulatory/disciplinary devices for those responsible for managing cultural activities – i.e. a municipal bureaucrat or a museum director. The systems are, in fact, based on voluntary adherence: municipalities in case 4.3 and museums in case 4.4 decide freely to be part of these systems. Besides, they involve a “continuous improvement” rhetoric as they roll out yearly, stimulating comparison on the diachronic scale. The assumption/belief at the heart of both information systems is the same: the quantitative commensuration of the evaluated objects (municipalities or museums) along selected dimensions and their ranking will stimulate comparison and lead to improvements. Although rooted in different countries (CIN in Norway, LUQ in Italy), these informative systems are meant to ensure coordination from a center in decentralized administrative structures. A relevant difference relates to the Italian context’s more explicit, top-down nature (the ministry designs LUQ). On the other hand, in Norway, the operation is launched and managed by a private actor. In this sense, the self-regulation or soft power element is even more robust in the Norwegian scenario.

Both cases show that relevant tensions can emerge when cultural information systems are excessively geared on numbers and quantification. In 4.3 the *excessive* focus on quantitative indicators and ranking of the original system entailed debates and tensions between TRI and municipalities about the actual effectiveness of the system in adequately representing the municipalities’ cultural activities. On an overarching level, such tensions related for example to how the index was seen to overfocus on participation, with a concurrent fear that this could jeopardize local cultural policy prioritizations of production. Moreover, tensions and issues arose also in the relationship between cultural policy research and local cultural administration, as the level of communicating how hard quantitative data may represent soft cultural values was regarded insufficient. In case 4.4 the excessive reliance of the Ministry’s questionnaire on quantifying “quality” led to several tensions at the national-local node, marked by the perception of the questionnaires’ inadequacy. The questionnaire’s quality representation was geared on national museums, such as the Uffizi, rather than regional, smaller museums. In other words, the concept of quality as measured by the questionnaire could not apply to the museums to be evaluated, leading to tensions. Namely, these tensions dwelled on how the same value (quality) had to be interpreted during the evaluation practice or which evaluative procedure had to be followed to remedy the questionnaire’s shortcomings.

Despite tensions, quantification is not given up in any of the contexts. Somewhat, tensions are mitigated in both cases by weakening and toning down the ranking usage. In case 4.3 tensions were mitigated through adjustments such as refining the index and communicating more effectively the underlying methodology, accounting for suggestions on possible improvements and including in the index a ponderation of the demographics of the different Norwegian regions. In case 4.4 the ranking was coupled with additional means for qualifying the information pooled by the SPC team when this was evaluating museums, i.e. additional professional judgments and referrals coming from experts and the professional connections of the SPC team. The ranking usage was also weakened when SPC decided to lower the threshold for eligibility. Indeed, SPC was not obliged by the law to make public the threshold of eligibility, nor the score obtained by each museum. Thus, the threshold could be modified and lowered so that SPC used it simply as an informal and internal “rule of thumb” for the evaluation. Overall, the large room for intervention granted to SPC by law reveals that the procedure was flexible and only apparently formal.

The mitigations are possible because of the specific configurations of the actors involved in the two cases. In other words, the quantification could be weakened because the evaluator and the evaluated entities (municipalities or museums) could communicate. In case 4.3, TRI and the municipalities could communicate and mitigate tensions because of the direct, “provider-user” type of relationship; the dialogue established with the municipalities, who were invited to respond and give feedbacks on the CIN, allowed TRI to take in

relevant suggestions for improving the index, to establish a channel for transferring knowledge on quantitative analyses to municipalities. As a result, these mitigations revealed the affordances of CIN as a tool to measure culture and provide effective benchmarks to municipalities. In case 4.4, the indirect, two-step nature of the system allowed SPC to act as a broker between the Ministry and the regional museums, facilitate and stabilize the evaluative practice, and finally perform a selective though representative evaluation within the expected timelines.

Overall, comparing cases 4.3 and 4.4 allows for a deeper understanding of cultural information systems and the use of the information produced. Relying on ranking constitutes a quick and time-saving mean for collecting information and systematizing data. However, when it comes to the effectiveness of these mechanisms in representing the complex and multidimensional nature of culture, ranking has to be toned-down or “qualified” with additional, specific knowledge, to be pooled with procedures and practices which are time-consuming and highly informal.

7. Topic-level discussion

In this deliverable, we analyzed highly diverse information systems with the aim of understanding tensions relating to how values are represented in cultural information systems. The topic aims to investigate data availability and quality, their categorization approach and their capability to cover the plurality of cultural values. Our sample includes two pairs of cases: the first pair (4.1 and 4.2) focus on the contents of the information produced by national (cultural statistics from Statistics Norway) or regional observatories (Regional cultural observatories in Italy). The second pair of cases (4.3 and 4.4) focuses on processual aspects of cultural information systems and includes CIN- Cultural Index Norway and LUQ-National museum levels of quality.

An information system is an infrastructure designed to collect and aggregate data, with the purpose of making the information produced available and meaningful. Besides this activity, an information system gains traction once its activities have a certain degree of continuity. The collecting and aggregating of data, and the subsequent production of meaningful information, will vary depending on the field of application. However, these activities are usually enacted with some common, though contrasting, features that can be grouped in two ideal-types of information management.

On the one hand information can be used or conveyed by measuring the object of evaluation through scores and indices. Quantification allows the display of information synthetically, through means of commensuration, which are highly efficient. One of the vantage points of using a rough number to convey information is its potential for comparison, which can lead to efficient considerations on similar disciplines and fields of application, usually through ranking and classifications. Its efficiency implies also a high level of replicability, which can lead to regular updates of the information produced and the possibility of coupling it with an improving rhetoric. Finally, it enables a “bird’s eye view”, providing snapshots of reality which are broad, wide but fundamentally flat and horizontal. All these features resemble an additive process of information, which we call “lumping”. With this term, we refer to a mechanism of data management in which information is produced by addition and conveyed through scores or indexes.

On the other hand, information can be used or conveyed by accounting for more detailed and specific kinds of information. The multiple ways in which information is used and conveyed usually encompass dialogical, verbal communication. These characteristics make this approach unapt to efficient/quick copy and replication, but extremely insightful to have vertical, deep views on contextual details. All these features resemble a process of dissection of information which we call “chunking”. With this term, we refer to a mechanism of data management in which information is produced by segmentation and conveyed through qualification and contextualization.

The cases of Topic 4 are similar in the fact that the methodology used at the onset is very close to lumping, thus suggesting that the actors have a widespread trust in it. However, in all cases actors refer a sense of inadequacy when commenting on the quality of the information produced, the categories employed and the scope of coverage. For instance, in case 4.1 the dissatisfaction originates from the fact that what is measured - i.e. consumption - only partially captures what actors would like to measure - i.e. participation. In case 4.2, the tension translates into a substantive difficulty in capturing the economic value of culture due to the peculiar nature of cultural activities. In case 4.3, the different interpretations lead to a tension related to how some of the bureaucrats feel that CIN misses certain indicators and that the existing ones misrepresent their context, or that the index does not fully reflect the intangible values of culture. In case 4.4, the sense of inadequacy stems from acknowledging that the ministerial questionnaire misrepresents the quality of regional museums.

These tensions are mitigated by taking advantage of some features of quantifying, or “lumping”, and of qualifying, or “chunking”. As per lumping, its beauty rests on lumping’s potential for synthesis and replicability over time. The cases in which the tension is mitigated show that lumping can be coupled with a rhetoric of improvement, which is a valid device for justification, even for the data collector/analyst himself/herself. This is evident in case 4.2, where actors do work on the databases to make them interoperable. Also, in case 4.3 actors fix bugs and attend to comments to improve the adequacy of the index

over time. In case 4.4, evaluators and evaluated do find reciprocal trust, and a valid justification to mitigate the tension, once museums are told that the process has a yearly rollout, aimed at improving and fine-tuning the information produced over time.

However, in all cases these actions were also complemented by working on chunking. Chunking has its main advantage in allowing deep and detailed observations. In case 4.1, an attempt to solve the tension was done by subdividing the categories. In case 4.2, by coupling the ordinary production of data with spot reports on specific and timely issues. In case 4.3, several adjustments were made to increase the degree of accuracy and capture the diversities of local cultural contexts. Case 4.4 features the SPC team changing the methodology by qualifying the questionnaire's scores through phone calls, meetings with experts and on-site visits, which provides a fine-grained contextualization and the possibility to adapt the evaluation to the criteria.

Overall, the cases of Topic 4 on information systems point to the complex and multidimensional character of the cultural field. Producing meaningful information in such an environment is not an easy task, given that lumping doesn't seem enough either by itself and in mitigating the sense of inadequacy. Although the cases show that lumping has to be complemented by chunking to produce these results, it is notable to assess that there is no actor giving up lumping as a first, initial step. This perhaps may be due to the sense of evidence that quantifying evokes, making it a powerful device for justification of actions to those ruling and managing.

References

- Berge, O. (2022). A land of milk and honey? Public funding for culture in Norway. In Sokka, S., Berge, O. K., Bille, T., Johannisson, J., Kaaber, E., & Lindqvist, K. (2022). *Cultural policy in the Nordic welfare states: aims and functions of public funding for culture*. Nordic Council of Ministers.
- Bjørnsen, E. (2012). Norwegian cultural policy—A civilising mission? The Cultural Rucksack and abstract faith in the transforming powers of the arts. *Poetics*, 40(4), 382-404.
- Danielsen, A. (2006). *Behaget i kulturen. En studie av kunst- og kulturpublikum*. Oslo: Arts Council Norway.
- Duelund, P. (Ed.). (2003). *The Nordic cultural model*. Copenhagen: Nordic Cultural Institute.
- Geertz, C. (1973). *The interpretation of cultures* (Vol. 5019). Basic books.
- Heinich, N. (2020). A pragmatic redefinition of value (s): Toward a general model of valuation. *Theory, Culture & Society*, 37(5), 75-94.
- Karpik, L. (2021). Valuing the unique. In *Valuing the Unique*. Princeton University Press.
- Langley, A. (1999). Strategies for theorizing from process data. *Academy of Management review*, 24(4), 691-710.
- Mangset, P., & Hylland, O. M. (2017). Kulturpolitikk. Organisering, legitimering og praksis. *Nordisk Kulturpolitisk Tidsskrift*, 196.
- Nelson, R. (2013). *Practice as research in the arts: Principles, protocols, pedagogies, resistances*. Springer.
- Schön, D. A. (2017). *The reflective practitioner: How professionals think in action*. Routledge.
- Williams, R. (1983). *Culture and society, 1780-1950*. Columbia University Press.