



inDICEs

Measuring the Impact of Digital Culture

Deliverable 1.1

inDICEs Methodological Toolbox



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

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

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



D1.1 – inDICES Methodological Toolbox

Final version

Grant Agreement number:	870792
Project acronym:	inDICES
Project title:	Measuring the impact of Digital Culture
Funding Scheme:	H2020-DT-GOVERNANCE-13-2019
Project co-ordinator name, Title and Organisation:	Simonetta Buttò, Director of the Central Institute for the Union Catalogue of the Italian Libraries (ICCU)
Tel:	+39 06 49210425
E-mail:	simonetta.butto@beniculturali.it
Project website address:	http://inDICES-culture.eu/

Authors:

Pier Luigi Sacco, FBK

Manlio De Domenico, FBK

Oriol Artime, FBK

Maria Tartari, IULM University

Federico Pilati, IULM University

Contributing partners:

Sabrina Pedrini, Alma Mater Studiorum University, Bologna

Alba Irollo, Europeana Foundation

Katinka Böhm, WLT

Sara Uboldi, University of Modena and Reggio Emilia

Elisa Pellegrini, EFHA

Rasa Bocyte, NISV

Document History

- 01.10. 2020– Draft Version
- 15. 10. 2020– Draft Version
- 21. 10. 2020– Draft Version
- 02. 12. 2020– Draft Version
- 21.12. 2020- FIRST SUBMISSION TO PARTNERS
- 30.12. 2020 - first internal review
- 08. 01. 2021 - first internal review
- 09. 01. 2021 - first internal review
- 11. 01. 2021 - first internal review
- 19. 01. 2021 - first internal review
- 04. 02. 2021 - first internal review
- 24. 02. 2021 - SECOND SUBMISSION TO PARTNERS
- 05. 03. 2021 - second internal review
- 18. 03. 2021 - second internal review
- 29. 03. 2021 - FINAL VERSION

Table of Contents

1. Executive Summary	5
2. Introduction: inDICES Methodological Toolbox	7
3. Objectives	10
4. Theoretical Framework: Digital, Culture and Participation	14
5. Targeted strategies of data acquisition	17
5.1 Data gathering on online cultural productions and reproductions	17
5.2 Data gathering on CHI digitization	21
6. Targeted Indicators	22
6.1 Indicators on CHI digitization	23
6.2 Indicators on users behavior	25
6.3 Indicators on Economic impact	26
6.4 Indicators on IPR status of CHI's	27
7. Conclusions	28
Annex 1. Explanation of the indicators on digital platforms' users behavior	29
Annex 2. Indicators on Economic Impact	35

1. Executive Summary

This deliverable aims to describe inDICES' targeted methodologies of research and analysis, briefly describing the Methodological Toolbox as the set of targeted methodologies defined through an in-depth review of the current state of the art approaches, both in relation to the research questions and to the kind of available and acquired data and their structure. These are the main components of the Methodological Toolbox:

- inDICES Theoretical Framework
- Targeted strategies of data gathering and data analysis
- Available datasets
- Set of useful indicators

This set of methodologies is aimed at supplying inDICES with new data in order to develop reports on CHI digitization status, to support the Self-assessment tool development, to populate the Observatory Platform and to offer insight in the understanding of the users behavior (in their use of different personas), as well as in the understanding of the social and economic impact of digitized culture as a basis for the design of evidence-based policies in the future context of the Digital Single Market (DSM).

As discussed in this deliverable and in Deliverable D1.3, users will have access to various types of data and content resources on CHI digitization via the inDICES Repository and the Observatory Platform's dashboard. These data include:

1. datasets/data dumps that are analyzed as part of inDICES, and the results made available through the Open Observatory,
2. relevant data provided by third-parties,
3. online content gathered on an ongoing basis, made accessible through WLT's Visual Analytics Dashboard in close to real-time,
4. on-line datasets on cultural production collected by FBK from the following sources/ web 2.0 platforms:
 - Wikipedia
 - Tiktok
 - Deviantart
 - Alltheater
 - IMDB

The whole of the material provided by inDICES partners, in its variety of sources and formats, will be organized and presented in the Repository that is part of the inDICES Observatory.

In this first phase of the project (M1-M12), the WP1 proceeded with data gathering by pursuing the following activities:

1. conducting and monitoring the state-of-the-art on data gathered from statistical institutes and other data gathering institutions;
2. managing the development of a new data gathering tool for the already available data, starting from, and aligned with, the ENUMERATE observatory reference case;
3. developing the possibility of generating or acquiring new data both through conventional and innovative channels, including data scraping on social media where possible, through natural experiments, and direct engagement of cultural content communities.

Data gathering processes and tools are detailed in Deliverable D1.3.

2. Introduction: the inDICEs Methodological Toolbox

The Methodological Toolbox is the set of techniques tailored to inDICEs data gathering, analysis and presentation, as an output of the WP1, and designed by means of a collective endeavor and research, according to the original project design and its evolutions and developments that resulted from the actual research work and from the partners' suggestions and comments collected in the meetings conducted during the first year of work.

The Methodological Toolbox is composed of a set of tools that are amenable to subsequent refinement, by choosing which methods and techniques to deploy, depending on the specific needs emerging from the research and from the inDICEs project development.

As planned in the submitted inDICEs proposal, here we provide a literature meta-analysis on existing cultural data, statistical reports, indicators, and tools for data collection and analysis, in order to understand the different modes of cultural production and participation in digital environments according to the inDICEs theoretical framework - based on the analysis and integration of the following frameworks, which are of special relevance for the purpose of our research: the Europeana Impact Framework, the UNESCO Culture for Development Indicators (CDIS), and the Culture 3.0 framework - and to define the most useful strategy for further data collection and analysis. Although in principle there are a plurality of frameworks to consider, we have chosen the ones listed above for the following reasons. The Europeana Impact Framework has been specifically designed to assess the various dimensions of the impact of cultural heritage participation. The Culture 3.0 framework, which has been originally developed on request of the EU Open Method of Coordination table on Cultural and Creative Industries, has been one of the main references of the Europeana Impact Framework itself. The latter's dimensions of impact have been selected by considering the 8-tier impact scheme of the former. Moreover, the Culture 3.0 framework has provided a methodological foundation for the New European Agenda for Culture, which has adopted 4 tiers out of its 8-tier scheme as the strategic priorities for the cultural crossovers. Taken together, these two tools can be considered today as a conceptual backbone of the EU conceptual approach to the impact of digital participation. Finally, the UNESCO CDIS represents the most widely adopted culture-related indicator framework in a global perspective, also due to the overarching role of UNESCO as the main umbrella institution in the cultural domain. These choices are therefore justified in the light of policy relevance considerations.

In parallel, WP1 parsed existing datasets in order to extract, catalogue, re-organize and optimize useful indicators for further research developments, as a benchmark for the

framework behind the data analyses that will be carried out by both the observatory and the research group.

The Methodological Toolbox is composed of:

- **Theoretical Framework [chapter 4]**

- The Theoretical Framework provides a convenient conceptual background to assess, and to align to, the current state of CHI digitalization, thanks to the integration of the following reference frameworks: the Europeana Impact Framework and Enumerate project, the UNESCO Culture for Development Indicators (CDIS), and the Culture 3.0 paradigm; its composition is underpinned by the reframing of data and information, currently fragmented across various reports and sources such as the reports on CH digitization and socio-economic impact: ENUMERATE, NEMO, EGMUS, CDSI, EUROSTAT, DESI, Etis, UNCTAD, Eu Open Data Portal, SotCommons [see Deliverable D1.3]. It will be presented and made accessible to the users in the Repository (see figure 1).

- **Tailored strategies of data gathering and analysis [chapter 5]**

- Built on the basis of the whole set of methodologies applied by FBK and WLT in the project's first year of data gathering and analysis, as well as those planned for data acquisition in the next period plan of inDICES.

Such a set of tools provides inDICES with techniques that allow to extract numerous types of quantitative information also from texts, and provide the basis for sophisticated qualitative assessments to address specific research questions.

- **Available datasets**

- Existing datasets on CHI digitization and socio-economic impact: ENUMERATE, NEMO, EGMUS, EUROSTAT, DESI, Etis, UNCTAD, Eu Open Data Portal;
- New datasets: web sources, social networks, 3.0 platforms of co-creation of cultural contents [detailed in Deliverable D1.3];
- New datasets on CHI digitization and socio-economic impact.

This resource provides inDICES with a constantly growing base of open datasets in order to feed the Open Observatory and extract useful indicators and metrics for monitoring the performances of CH and Cultural and Creative Sector (CCS) digitization in the Digital Single Market (DSM). These datasets will be organized and presented in the Repository made available on the inDICES platform.

- **Set of useful indicators [chapter 6]**

- Indicators on CHI digitization;
- Indicators on social and economic impact of CCS digitization;
- Indicators on 3.0 platforms users behavior.

The aim of these tools is to provide inDICEs with new evaluation metrics to be used in the preparation of reports on CHI digitization status and to support the development of the self-assessment tool, to feed the Observatory Platform and its analysis users' behavior (in their use of different personas), to understand the social and economic impact of digitized culture and to facilitate the design of evidence-based policies in the future context of the DSM. These indicators will also be organized and presented in the Repository to the benefit of users and researchers.

3. Objectives

The objective of this deliverable is to describe the work processes behind the development of the Methodology, and the definition of the inDICEs Methodological Toolbox. This deliverable draws from the activities carried out during the first 12 months of the inDICEs project within Work Package 1 (WP1) by all consortium partners (FBK, KU Leuven, Stichting Europeana, Platoniq, Centrul Cultural Clujean, Beeld en Geluid, EFHA, Michael Culture, Deutscher Museumsbund, WLT, CHT, PIN SCRL), and describes the results achieved by this work package. The deliverable includes a set of activities carried out to define the objectives and tasks aimed at the definition and calibration of the basic research methods of the project, and to the consequent activities of data gathering and organization.

In particular, we consider strategies of acquisition of relevant data through social media and more generally from use of relevant digital platforms in the context of the European culturescape. In addition, forms of participatory data acquisition from an open collaborative effort of specific communities of prosumers, interest groups, cultural and creative professionals and companies have been evaluated and carried out. These resources are likely to be useful to different profiles of researchers and practitioners in the cultural and creative sectors. Once properly organized, the data will be fed into the inDICEs Observatory Platform. Once the available techniques and data have been established, it becomes possible to operate a more careful choice of the actual techniques to be employed and to test them on the data, also by means of specific pilot experiments. This task then defines in more detail the actual toolbox that will be used throughout the project, to arrive at a more refined and specific formulation of the research questions to be addressed.

In what follows, we provide an extremely summarize account of the main features of the three main relevant frameworks for the inDICEs project, and of their defining features that are of special interest for our analysis. We also review some of the basic analytical approaches available to carry out cultural analysis, briefly discussing their main features and potential.

Starting from a clearer assessment of the available data as presented in Deliverable D1.3a, targeted strategies of data acquisition and gathering are then devised, both to fill at least in part the most serious gaps, and to conduct complementary kinds of analysis to the ones made possible by already available sources.



Figure 1: Repository integration flowchart

As more specifically described in Deliverable D1.3, in the “Plan for the next period” section it is described how the data we are collecting and the methodological tools we are developing will be used: who are the users we address, what kind of information we want to provide them with, what tools they will have at their disposal and how they will be integrated into the inDICES platform.

It proved necessary to introduce within the Open Observatory an additional tool, identified during WP1 internal and collective meetings as a **Repository**. Complementing the content resources of the Visual Analytics Dashboard, a Data Repository is going to be realized as an appropriate, subject-specific location where inDICES users can directly access raw data and reports, which will be synthesized and visualized via boxplots. The development of the Repository will fully comply with the D1.2 Data register [M18] objectives, which regard the final design of the structure and characteristics of data gathered, their sources, and their usefulness and limitations, and an organizational scheme that allows their effective accessibility within the Open Observatory.

For what concerns **inDICES users**, as shown in the Table 1, the inDICES Open Observatory is addressing various groups of potential users that were hypothesized, and classified into five different categories: artists, policy makers, researchers, cultural heritage practitioners, and special interest groups.

Table 1 (for a detailed description on inDICES target, see Deliverable D4.1)

Profile	Priorities	Interest in inDICES
Creative Communities and Artists	Impact assessments, contacts, tools to increase engagement, business models and references	Increase their social, cultural and economic impact, getting expert input, and brainstorming solutions within a larger community
Policy makers	Clearly communicated case studies, policies, and data sets	Studies, narratives, data from CHIs, indicators, and access to new contacts
Researchers	Searchable data sets and specific profiles to network with and study	Access to different data with different filtered criteria, the ability to play with data and be part of a working group
Cultural Heritage Practitioners	Sharing experiences, understanding user experience for their work, tools and resources, impact assessments, and performance benchmarks	Recommendations, provide data, modernise the cultural heritage sector, participate with collaborators to solicit advice and also make decisions collectively
Special Interest Groups	Access contacts, learn CHI practices, explore data from CHIs and learn more about user experiences	Meeting interesting contacts to find new project partners, learning new technologies, and sharing experiences

The Participatory Platform of InDICES in conjunction with the embedded widgets of the Visual Analytics Dashboard addresses the following target groups:

- The *Creative Sector*, which has an interest to enhance its social, cultural and economic impact, obtaining expert feedback and co-creating solutions within a larger community;
- *Cultural Heritage Practitioners* who aim to share and understand user experiences for their work, to make use of tools and resources, to carry out impact assessments, as well as performance benchmarks to provide the heritage sector with useful recommendations, and to engage associate and partners to solicit advice and facilitate collective decision-making;

- *Special Interest Groups* searching for contacts to find new project partners, learning new technologies, and sharing experiences;
- *Policy Makers and Researchers* in search of new contacts and field-specific working groups and expertise.

The *Repository* is a tool addressing mainly:

- Policy makers in search of studies, narratives, data from CHIs, and indicators;
- Researchers, Practitioners and users from the Creative Sector who want direct access to raw datasets, various data sources searchable through a number of different filtering criteria, the ability to play with data, and data sets and specific profiles to network with and to study;
- Special Interest Groups who want to explore data from CHIs.

4. Theoretical Framework: Digital, Culture and Participation

In this section, we introduce the state of the art as to the relations among culture, participation and digitization. The different contributions and resources mentioned are consistent with InDICEs strategies, components, aims, and resources.

As already discussed above, the InDICEs model design is based on the analysis and integration of the following reference frameworks: the Europeana Impact Framework, the Enumerate project, the UNESCO Culture for Development Indicators (CDIS), and the Culture 3.0 paradigm. The rationale for this choice has been illustrated in Section 2. Starting from the co-design phase, the aforementioned frameworks helped us to establish: InDICEs methods of analysis, logical structure, internal relationships between WPs, and limitations. In particular, they have been useful to define models for accessing and analyzing data in: WP1, “Participatory research methods and analysis”; WP3, “Change management and policy recommendation”; WP4, “Open Observatory”; and T5.3, “Designing a model for digital community participation as a driver of impact”.

The Europeana Impact Framework and Enumerate project: The Europeana Impact [Framework](#) provides a first important basis for designing and assessing impact in digital cultural heritage. It focuses upon various dimensions of economic impact from cultural heritage participation and access. On the basis of this framework, specific models can be developed to provide more detailed insight into every possible area of impact. The specific modelling extension of specific areas of the impact framework will be accompanied by the definition of suitable indicators that provide a synthetic measurement of the main outcomes. Europeana contributed to Enumerate, a project funded under the EC’s ICT Policy Support Programme, that investigates the state of digitisation of cultural heritage institutions in Europe, particularly museums, libraries, archives. It aims to provide a baseline of data that can inform decisions at the national and EU policy level, and it is based on gathering statistical information through a network of national coordinators. Since 2011, it has run four surveys. The last survey was conducted in 2017 and was structured around six topics: 1) digital collections; 2) digitisation activity; 3) digital access; 4) participation; 5) digital preservation; 6) digital expenditures. Nearly 1,000 institutions took part in it, and 82% of them claimed to have a digital collection or to be in the process of launching a digitisation project. The Europeana PRO website currently hosts the [Enumerate Observatory](#), where one can also find all the [Documentation](#) related to the project and the datasets of the surveys, anonymised and made available as [raw data](#).

The **UNESCO Culture for Development Indicators (CDIS)**¹: The role of culture in development is today recognized not only by the culture community but also increasingly acknowledged by the development community. References to the importance of culture both as a driver and enabler for sustainable development have been included in recent major documents that chart the path for a renewed development agenda, including the 2005 UNESCO Convention on the Protection and Promotion of the Diversity of Cultural Expressions, and the United Nations 2030 Sustainable Development Agenda.

In spite of this promising context, the difficulties encountered to date in quantifying the contribution of culture have led to its marginalization in national and international development strategies. UNESCO has developed the CDIS methodology in response to this challenge. Adopting a holistic approach, the CDIS covers seven interrelated policy dimensions that address the multi-faceted contribution of culture to sustainable development, as formulated and elaborated upon in UNESCO CDIS Methodology Manual²:

1. Economy
2. Education
3. Governance
4. Social Participation
5. Gender
6. Communication
7. Heritage

CDIS is taken into account also because it encourages cross-references among policy dimensions, so that the results of measurement exercises illustrate the contribution of culture to the creation of economic, social and cultural value, as well as contribute to a better policy design to enhance the impact and effectiveness of development interventions.

Culture 3.0: To understand the effects of the digital revolution on modes of cultural and creative production and on their economic and social impact, the framework of Culture 3.0 (Sacco 2018) was developed by distinguishing between three regimes of production: Patronage (Culture 1.0), today mostly applying to non prevalently market-oriented sectors such as visual arts, performing arts, museums and heritage; Cultural and Creative Industries (Culture 2.0), applied to industrialised forms of cultural and creative production based on the structured distinction between producers and audiences; and open communities of practice (Culture 3.0), where production and distribution of content are not necessarily market

¹ <https://en.unesco.org/creativity/activities/cdis>

² https://en.unesco.org/creativity/sites/creativity/files/cdis_methodology_manual_0_0.pdf

mediated and where the distinction between producers and users becomes blurred to an increasing extent [see Figure 2: regimes of cultural production].

The Culture 3.0 framework allows us to understand how different production regimes each follow their own logic, so that, when reflecting about the impact of digitization on cultural participation practices, it is important to distinguish how institutions operating within different regimes face different sets of opportunities and constraints. This framework has been important in the economy of the INDICES project for the selection of specific case studies covering digital platforms that exemplify different regimes, and to check to what extent such difference reflects into differences in modes of production and access.

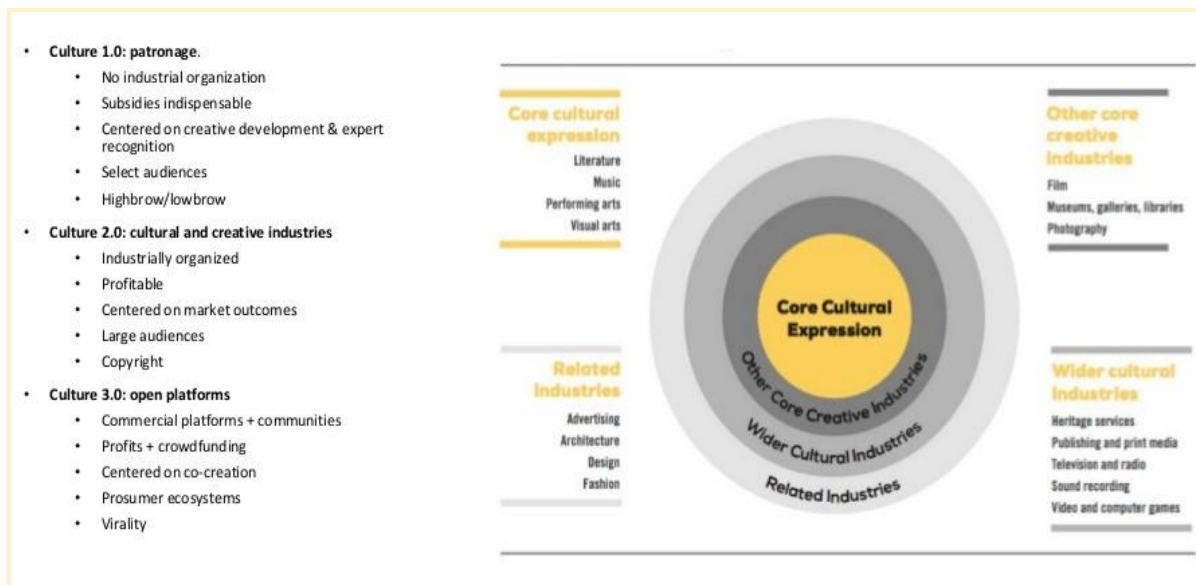


Figure 2: regimes of cultural production

5. Tailored strategies of data gathering

5.1 Data gathering on online cultural productions and reproductions

For this first year of work, inDICEs selected two combined strategies for data gathering, Web crawling and data ingestion through the APIs of content platforms. These strategies allowed inDICEs to ingest large amounts of test data. For the Open Observatory, inDICEs collected datasets on the current, available online cultural productions and reproductions, as well as digital content from Web sources (news, social networks and stakeholder Website) that is accessible via the Visual Analytics Dashboard.

This first phase of data gathering, as anticipated in section 2, is followed by a phase of model making, and one of analytics and interpretation. In the model making phase, once data has been gathered for a certain platform, we build convenient models that allow us to investigate some of the main structural features of the platform in terms of modes, levels and quality of participation, according to cases. In doing so, we apply established techniques from computational social science, in order to derive quantities such as the distribution of user-generated content production across users in the case of Culture 3.0 kinds of platforms, or the levels of access and characteristics of users in the case of Culture 1.0 and 2.0 kinds of platforms where user generated content is either not contemplated or relatively peripherals with respect to the platform's value proposition and logic of functioning. Once the modeling phase has been carried out and key quantities and values have been computed, we proceed to the analysis and interpretation of the data, which are the core of the third phase. The representativity of the data will be ensured by our big data approach that works on very high volumes of data according to the best standards of the computational social science literature. In terms of potential data, our approach is able to extract the relevant information content from any kind of production/participation data that might be available, and therefore, also in view of the difficulty of finding databases of sufficient size and quality in the cultural participation sphere, we do not pose strong restriction in terms of 'desirable' features of datasets but rather reverse-engineer our approach on the basis of what is objectively and realistically available. At the end of this cycle, we expect to obtain methodologically innovative results that shed new light on several features of the processes of content production and access on digital platforms that may be very useful for the design of cultural digital participation strategies.

To gather web content, the WLT Web crawler collects and updates Web pages to be added to the InDICEs content archive, based on the Java open source Apache Storm-Crawler (released under the terms of the ASF 2.0 License). Depending on the dynamics of the

sample to be captured, the pages are collected in daily or weekly intervals. RSS feeds are used to ingest more topical content every 10 minutes. The crawler introduces voluntary bandwidth limits per site to minimize the resulting load on third-party servers. The InDICES data collection process respects the Web site owner's robots.txt settings (a text file placed in the top directory, which is used by site administrators to restrict access to files and directories on a Web server). URLs to be crawled are collected using a Google Spreadsheet configuration template. We currently plan to further automate this source management process by programmatically creating spreadsheets and populating them with existing URLs as well as the volume of ingested content for monitoring and validation purposes.

To gather user-generated and social media content, InDICES uses the official APIs provided by the various networking platforms – strictly adhering to these platform's usage restrictions and only accessing the public portion of the content. To customise the content ingestion process, a combination of account names (gather all postings from the specified accounts) and Regular Expression-based search terms (for accessing content via streaming APIs) is being used. The terms to be gathered are collected using the configuration template mentioned above.

To define and guide the data gathering process, a configuration template has been provided by WLT to gather content and engagement data. It reflects what citizens and professional stakeholders produce and share. This resonates with the main project objectives, which is the development of tools for iterative measurement and prediction of audience engagement with digital content published online, in order to provide recommendations for its improvement. The template helps to ensure high-quality content from the specified source and to avoid common and too generic terms to reliably identify quality content. The aim of the resulting filter, which is still being refined, is to categorize the most meaningful and reliable keywords for web content extraction.

In order to understand the effects of the digital revolution on modes of cultural and creative production, we are contributing with a range of different actions to develop a methodology to measure the economic and social impact of digitization of cultural (heritage) assets as to the access to cultural goods and services and their modes of production.

The effort made for this action has been to hypothesize a system of groups and subgroups of keywords according to the framework of Culture 3.0 (Sacco 2018) which distinguishes, as already remarked, between three regimes of production "Patronage (Culture 1.0), today mostly applying to non prevalently market-oriented sectors such as visual arts, performing arts, museums and heritage; Cultural and Creative Industries (Culture 2.0), applied to

industrialised forms of cultural and creative production based on the structured distinction between producers and audiences; and open platforms/communities of practice (Culture 3.0), where production and distribution of content is not necessarily market mediated and where the distinction between producers and users becomes blurred to an increasing extent” as it appears in the inDICES objectives.

WP1 initially carried out a literature analysis (see SOURCES below) which enabled inDICES to define the keyword list and to check for its appropriateness. With a first attempt, WP1 enlisted three groups of terms to provide a hierarchical structure to define specific concepts that identify the most relevant themes of each regime. According to users stories and personas developed from WP4 (see Deliverable D4.1), WP1 re-focused on the process and reworked the initial structure of the configuration template proposed by WLT, leading to the following tables:

CULT PROD 1.0: Institutions/venues, Main examples, Production, Expressions, Actors

Group	Subgroup 1.0-1	Subgroup 1.0-2	Subgroup 1.0-3	Subgroup 1.0-4	Subgroup 1.0-5
CULT PROD 1.0	Institutions/venues	Main examples	Production	Expressions	Actors
art patronage	museum	cultural foundation	lecture	performing art	
artistic patronage	library	Aishin Foundation	frontal lesson	puppetry	connoisseur
renaissance patronage	theatre	Alexander Calder Foundation	frontal teaching	circus art	aesthetic
cultural elite	archive	Andy Warhol Foundation	passive audience	opera	art master
cultural institution	cultural institution	Antonio Saura Foundation	encyclopedia	comedy theatre	renaissance workshop
art institution	art institution	Aperture Foundation	Chambers' Cyclopaedia	pantomime theatre	patron
cultural hierarchy	cultural circle	ARTER Space for Art	Encyclopédie	drama theatre	maestro
cultural circle	cultural academy	Arthema Foundation KAI10	Encyclopædia Britannica	kabuki	art client
cultural academy	art foundation	Astrup Fearnley Museet	Lexicon Technicum	improvisational theatre	art buyer
expertise	art gallery	Audain Art Museum	Harmsworth's Universal Encyclopy	commedia dell'arte	art commissioner
public funding for culture	art system	Ballroom Marfa	Children's Encyclopaedia	ballet	amanuense monk
cultural philanthropy	art fair	Barjeel Art Foundation	World Book	unique piece of art	art-star
fine arts		Barnes Foundation	Encarta	aura Benjamin	
culture gate-keeper		Basil & Elise Goulandris Foundat	Stanford Encyclopedia of Philoso	mural	
high-brow art		Benesse Art Site Naoshima		frescoe	
art funder		Berardo Collection Museum		monument	
art foundation		Bernhard Heiliger Foundation		architecture	
cultural policy		BINZ39 Foundation			
public funding for culture		Blenheim Art Foundation			
cultural heritage		Bonniers Konsthall			
cultural legacy		C/O Berlin Foundation			
elitarian culture		Calouste Gulbenkian Foundation			
culture conservation		Calvert 22			
art conservation		Camargo Foundation			
economy of patronage		Cantor Foundation			
		Carmignac Foundation			
		Cc Foundation			
		Centro Botin			
		Cisneros Fontanals Art Foundation			
		Clyfford Still Museum			
		Colección Fortabat			
		Cragg Foundation			
		Crystal Bridges Museum			
		Cue Art Foundation			
		Dali Foundation			
		Darat Al Funun			
		De 11 Lijnen			

CULT PROD 2.0: Cinema/tv, Publishing, Music/Labels, Design/Fashion, Platforms

Group	Subgroup 2.0-1	Subgroup 2.0-2	Subgroup 2.0-3	Subgroup 2.0-4	Subgroup 2.0-5
CULT PROD 2.0	Cinema/tv	Publishing	Music/Labels	Design/Fashion	Platforms
cultural and creative industry	cultural infotainment	NN	Stone Music Entertainment		cultural entertainment
reproducible content	edutainment	e/o	Genie Music		cultural relations
cultural producer	Hollywood production company	Mondadori	European Music Council		culture action
cultural consumer	Warner Bros.	Penguin	Warner Music		HBO
mainstream	Sony Pictures Motion Picture Group	Coconino	Emi		Hulu
sub-cultures	Walt Disney Studios	London Review of Books	Sony Music		Amazon Prime Video
indie	Universal Pictures	The A.V. Club's	BMG		Netflix
copies	20th Century Fox	IndieWire	Universal Music Group		Spotify
copyright	Paramount Pictures	The New York Times	PolyGram		Sky
blockbuster	Lionsgate Films	Variety's	Island Record		Disney Plus
podcast	The Weinstein Company	The A.V. Club's	Elektra		Youtube Premium
experience economy	Metro-Goldwyn-Mayer Studios	Grand Bell Awards	BlueNote		
entertainment machine	DreamWorks Pictures	The Wall Street Journal	Roc-A-Fella		
	A24		DJM Record		
	Annapurna				
	PlanB				
	Smokey House				
	Cannes Film Festival				
	Venezia Film Festival				
	Berlinale				
	Sundance Film Festival				
	Tribeca Film Festival				
	Cinema Ritrovato				
	Biografilm				
	Future Film Festival				
	Giffoni				
	Paté				
	Gaumont				
	Hong Kong				
	Deauville				
	LA Film Festival				
	Rotten Tomatoes				
	Edimbourg Film Festival				
	CJ Entertainment				
	Metacritic				
	Critics' Choice Movie Award				
	Academy Award				

CULT PROD 3.0: Actors, Sources, Practices, Platforms, Specific phenomena

Group	Subgroup 3.0-1	Subgroup 3.0-2	Subgroup 3.0-3	Subgroup 3.0-4	Subgroup 3.0-5
CULT PROD 3.0	Actors	sources	practices	Platforms	Specific Phen
cultural commons	influencer	open source platform	content co-creation	SoundCloud	Directioner
prosumer	cultural prosumer	cultural crowdsourcing	art-based community development	FilmHub	K-Pop
online community	user	copyleft	tactical urbanism	Instagram	Beliber
coworking	hater	copyfree	cultural public practices	TikTok	Ferragnez
cultural hub	troll	sharing community	bottom-up practice	Youtube	
net-art	follower	sharing content platform	empowerment	Twitch	
cultural heritage sector	blogger	content ecosystem	right to the city	GitHub	
Faro Convention		post-copyright business model	culture-led social innovation	Reddit	
heritage community			urban art	Gab	
open communities of practice			community-led art practice	4chan	
cultural digitalization			Freebooting	8chan	
crowdsourcing			Trolling	Quora	
wiki			culture-driven behavioral change		
intersectional					
viral					
social-media					
participatory platform					
community sense-making					

To ensure the long-term sustainability of data gathering and adjust the filter settings on an ongoing basis, inDICES plans to not only engage consortium partners but also third-party communities in participative research activities, and develop strategies of democratization of cultural production and access. Specifically, inDICES is about to launch an Open Call for Open Sources, as described in Deliverable D1.3. The aim is to provide a critical mass of relevant and fully annotated content that reflects open sources of cultural production and

reproduction. The Open Call will be addressing not only inDICES partners and their communities, but also a huge cross-national targeted group of already profiled users (personas), who can contribute to inform Observatory administrators on the most interesting and valuable open sources about on-line cultural contents and off-line re-utilizable sources and case-studies.

5.2 Data Gathering on CHI Digitization

As detailed in the last section of the Deliverable D1.3, in the next six months [M15-M21] further real-time data on cultural heritage online coverage and digitisation of CHIs will be collected, annotated and made searchable by the Visual Analytics Dashboard.

In addition, WP1 and WP3 will work together for integrating the Self-Assessment Tool with additional indicators (on users' behavior, IPR, and other information on CHI digitization) that will produce the specific data needed to align inDICES work with the already existing reports on CHI digitization and to carry the research on. The set of data thus gathered will contribute to both customize the Visual Analytics Dashboard and provide content for the Repository.

6. Targeted Indicators

The aim of this section is to present the first set of target indicators for dataset analysis, according to the main inDICES goals of measuring and analysing:

- Digitized CHI cultural productions and reproductions levels in terms of access, participation, preservation, expenditures and socio-economic impacts;
- Models of cultural participation online and of users' behavior.

According to the consultation conducted by leading experts and inDICES partners organizations, and according to the new developed strategies of data analysis implemented on the new sets of data gathered in this first year, we extracted, selected and targeted a first set of useful indicators which will improve the effectiveness of the inDICES platform and which will reasonably aspire to become a permanent reference for researchers, professionals and policy-makers operating not only in the CH sector but more generally in the whole sphere of cultural and creative production.

The sets of indicators presented in the following subsections have a twofold aim, and are addressed to two different inDICES tools, according to inDICES users' target. As stated in Deliverable D4.1, inDICES targeted personas may be characterized as follows, in relation with two different needs:

- a 'end-user' public interested in pre-organized and already processed data, measured by pre-selected indicators, visualizable via Visual Analytics Dashboard or via the Repository's reports (box plots and graphics);
- a more 'data fluent' public interested in raw datasets and in the indicators' list, available in the inDICES Repository.

For these reasons, as described in the Plans for the next period section from Deliverable D1.3, an additional tool will be added to inDICES Open Observatory, namely a Repository.

The following criteria were used to select indicators included in this database:

- **Relevance:** there is a clear relationship between the targeted indicators and the main goals of the inDICES project, oriented to the evaluation of digitized culture, its production and reproduction, its socio-economic impacts and its users' behavior;
- **Accuracy:** the indicators actually measure what they are expected to measure on the basis of the inDICES main goals;
- **Importance:** the measurement captures existing and new information that is objectively relevant and that can be of high interest for program effectiveness;
- **Usefulness:** the results point to areas for improvement: the indicators capture information both to support the progress of inDICES strategies, priorities, and programming, and of future targeted users by enabling them to develop their own

research or pursue their curiosity. With this aim in mind, the inDICES WP1 future plans will introduce an additional tool for data and indicators accessibility;

- Feasibility: data can be obtained with reasonable and viable effort. With this aim, inDICES WP1 future plans will introduce an additional tool for data gathering;
- Credibility: the indicators have been recommended - and are being used - by leading experts and inDICES partners organizations in their institutional reports;
- Validity: to the extent possible, the indicators have all been field-tested;
- Distinctiveness: the indicators lack redundancy and do not measure dimensions that are already captured by other indicators.

In the next six months [M15-M21] the most updated data on digitisation of cultural heritage advanced by CHIs will be gathered through the Visual Analytics Dashboard, according to the first list of indicators that have been identified and to a new list of indicators that will be defined with the help of a group of inDICES targeted users, such as researchers, CHI practitioners and policy makers, who will be consulted during the inDICES 2nd Consultation Workshop.

6.1 Indicators on CHI digitization

The indicators related to the levels of CHI digitization have been extracted, skimmed and reported by careful parsing of the most important thematic reports available: Enumerate, NEMO and EGMUS (last publication).

The following selected first indicator package will be available in Repository, the new tool that will be developed in the inDICES Open Observatory, in order to be a useful instrument for inDICES targeted users' research.

Indicator	Area	Information conveyed
Number of artworks digitized per institute	Collection and Digitalization	Collection Care
Number of collections digitized per institute	Collection and Digitalization	Digital Collection / Activities
Number of CHI with want to develop strategies of digitization	Collection and Digitalization	Digital Strategies
Number of CHI with written strategies of digitization	Collection and Digitalization	Written Strategy
Number of BDC per institute	Collection and Digitalization	Born Digital Collection
Number of heritage Collection Catalogue in collection database per	Collection and Digitalization	Heritage Collection Catalogue in collection database

institute		
Number of Digital Access per year per institute	Digital Access/Participation	Digital Access (available online heritage collection)
Number of available online digitally reproduced and born digital per institute	Digital Access	Available online digitally reproduced and born digital
Copyright conditions (how owner) content and metadata	Digital Access	Copyright conditions (how owner) content and metadata
Percentage income improvement after digitization	Digital Access	Reasons for providing digital access
Number of channels used for access digital collections	Digital Access	Channel used for access digital collections
Delta btw Offline and/or online access	Digital Access	Offline and/or online accessibility
CHI Main online access channels	Digital Access	Main online access channels
Number of institutions measuring	Participation	Percentage of institutions measuring
Number of visit to digital collection and website	Participation	How institutions are measuring
Number of CHI with Digital preservation strategy	Digital Preservation	Digital preservation strategy
Percentage of International standard for digital preservation	Digital Preservation	International standard for digital preservation
Internal and external budgets	Digitisation Expenditure	Internal and external budgets
Incidental and structural costs	Digitisation Expenditure	Incidental and structural costs
Inhouse costs and outsourced costs	Digitisation Expenditure	Inhouse costs and outsourced costs
Costs of each activity	Digitisation Expenditure	Costs of each activity
Number of Paid staff and volunteers	Digitisation Expenditure	Paid staff and volunteers
Percentage of external funding	Digitisation Expenditure	Funding
Number of museum using computers (and for which purpose)	Museum and Digitization	Museum and Digitization
Number of museums with website	Museum and Digitization	Museum and Digitization
Number of museum possessing own website	Museum and Digitization	Museum and Digitization
Number of museum updating themselves website	Museum and Digitization	Museum and Digitization
Number of museums connected to museum portal (1 or more)	Museum and Digitization	Museum and Digitization

6.2 Indicators on users behavior

The indicators related to the user behavior [as detailed in Annex 1] in cultural platforms are derived from two types of data: network of interactions and time series. These tools refer to the modeling phase of our work cycle as detailed above, so that we can have network models or time series models, according to cases and necessities. To be precise, networks give information about the underlying organizational architecture of complex systems, mapping complicated interactions onto the simple language of nodes and edges. This turns out to be convenient, due to the algorithmic and visualization power that the realm of network theory offers. Regarding the time series analysis, it turns out beneficial in scenarios in which we have the time evolution of a variable (e.g., number of visits in a digital collection, shares of an artwork in online social platforms, frequency of the user interactions) and no other metadata is available. This allows us to unveil hidden patterns in the temporal dimension, complementing the topological point of view of network theory.

In the following we outline the proposed indicators for these dimensions, together with their explanation and how to compute them.

Table with all the indicators:

Indicator	Type of data needed	What information conveys
Gini index for the degree distribution	Network of interactions	Inequality in the social capital
Assortativity index	Network of interactions	Level of homophily between the elements of the system
Clustering coefficient	Network of interactions	Compactness of the relations
Average connectivity length	Network of interactions	Facility of global communication
Network efficiency	Network of interactions	How efficiently a network exchanges information
Community modularity	Network of interactions + community structure	Strength of the community structure
Burstiness coefficient	Network of interactions + community structure	Measure of the intensity of highly active events
Fano factor	Activity time series	Measure of the intensity of highly active events
Memory coefficient	Activity time series	Type of correlations in the temporal behavior

6.3 Indicators on Economic Impact

The following first set of selected indicators on the economic impact of the digitization of culture, in line with inDICEs objectives, have been extracted, skimmed and presented by a selective parsing of the most important thematic existing reports such as those by EUROSTAT and DESI, which have been analysed and referenced, with special attention to socio-economic impacts of culture, in Deliverable D1.3. The Digital Economy and Society Index is a European Commission’s yearly index that aims at measuring the digital competitiveness of Member States and its evolution; the international DESI index extends the analysis to other 18 non-EU-countries. Finally, it is also of interest to refer to the Women in Digital Scoreboard, that provides an analysis of the women’s inclusion in digital entrepreneurship, careers and jobs.

These indicators, which will be partially integrated in the additional tool developed, as explained in inDICEs future plans, namely a survey addressed to the CHI sector, made of indicators drawn from data analysis of the users’ behavior on web 2.0 platforms, and of selected indicators from the most relevant existing reports on CHI digitization. The following selected indicators will be available in the Repository, the new tool that will be developed in the inDICEs Open Observatory, in order to provide a useful instrument for inDICEs targeted users’ research, and to contribute to the inDICEs project as to the understanding of the social and economic impact of digitized culture, in order to help design effective evidence-based policies in the future context of the DSM.

The complete table of indicators is available in Annex 2. Below we report a reduced version.

inDICEs Objectives	Dimensions	Subdimensions/ Indicators
<ul style="list-style-type: none"> • understanding the need of all publics • supporting highly qualified professionalism • promoting the cultural use of new technologies • stimulating learning through culture • positioning access to cultural upstream and transversally in all cultural policy making 	Connectivity	4 sub
<ul style="list-style-type: none"> • understanding the need of all publics • supporting highly qualified professionalism • promoting the cultural use of new technologies • positioning access to cultural upstream and transversally in all cultural policy making • overcoming linguistic barriers 	Human capital	2 sub

<ul style="list-style-type: none"> • understanding the need of ALL publics • supporting highly qualified professionalism • improving funding & procedures • overcoming linguistic barriers • Advancing mobility and exchange • promoting the cultural use on new technologies • stimulating learning through culture • positioning access to cultural upstream and transversally in all cultural policy making 	Use of Internet Services	3 sub
<ul style="list-style-type: none"> • understanding the need of ALL publics • supporting highly qualified professionalism • improving funding & procedures • overcoming linguistic barriers • Advancing mobility and exchange • promoting the cultural use on new technologies • stimulating learning through culture 	Integration of digital technology	2 sub
<ul style="list-style-type: none"> • supporting highly qualified professionalism • improving funding & procedures • positioning access to cultural upstream and transversally in all cultural policy making • raising awareness of the legal frameworks on access to culture 	Digital public services	1 sub
supporting highly qualified professionalism	Cultural Employment	7 sub
supporting highly qualified professionalism	Enterprises in cultural sector	6 sub
supporting highly qualified professionalism	International trade in cultural goods/services	4 sub
<ul style="list-style-type: none"> • supporting highly qualified professionalism • understanding the need of ALL publics • stimulating learning through culture • promoting the cultural use on new technologies 	Participation (cultural sphere)	21 sub

6.4 Indicators on IPR Status of CHI's

Indicators drawn from data analysis carried out on the WP2 survey on intellectual property rights and CHIs, e.g. questions on IPRs, IPR status of CHI's collections, licenses used, etc. will be integrated both in the Visual Analytics Dashboard and in the Repository.

7. Conclusions

The Methodological Toolbox is a targeted, viable and flexible methodology, consisting of a set of tools that can be subsequently refined by choosing which methods and techniques to deploy, depending on the specific needs emerging from the research and from inDICES' project development path.

In order to understand the different modes of cultural production and participation in digital environments according to the inDICES theoretical framework, based on the analysis and integration of the following relevant frameworks: Europeana Impact Framework, the UNESCO Culture for Development Indicators (CDIS), and the culture 3.0 framework, inDICES carried out a literature meta-analysis on existing cultural data, statistical reports, indicators, and tools of data collections and analysis, in order to single out the most useful strategy for inDICES further data collection and analysis.

In parallel, WP1 inquired existing datasets in order to extract, catalogue, re-organize and optimize useful indicators for further research developments, so as to flesh out the structure supporting further analysis of data that will be gathered by both the Observatory Platform and the inDICES partners (FBK research group).

On the basis of the remarks presented in this deliverable, the WP1 working group believes that it is essential, in the next six months of project implementation and data gathering and analysis, to maintain a high level of flexibility in the design of strategies and in the development of the tools necessary to meet the goals of the project, in accordance with the complexity and the high quality of the desired final output.

Annex 1. Explanation of the indicators on digital platforms' users behavior

- **Gini coefficient for the degree distribution.** The Gini coefficient (also known as the Gini index or Gini ratio) measures how far a distribution of a set of values is from an egalitarian scenario, where all measured values are the same. It is frequently used in economics to assess the wealth inequality within countries, but it has found applications in other areas, such as in education, ecology, chemistry or engineering, among others. Here we propose to compute a Gini index for the degree distribution of a complex network, where the degree corresponds to the number of links a node (most of times a user, but not only) has. The degree distribution is the simplest measure of centrality in a network, offering a way to rank the nodes according to a topological descriptor. Therefore, the Gini index of the degree distribution will give an idea about the level of heterogeneity in the relationships of the elements of our system.

The Gini index can be computed as follows, see Fig.1 for a sketch. For a set of observations, in our case, the degrees of the nodes in the network, one computes their cumulative distribution sorting nodes from lowest to highest degree. This will give us the convex cumulative curve, from which we can easily compute the area of the regions A and B. In practice, it suffices to compute only one of them, since the relation $1/2 = A + B$ is always held. The Gini index G is given by the ratio $A/(A+B)$, i.e., the farther the cumulative curve is from the diagonal, the larger the inequality. In practical terms, one can compute it using the relative mean absolute difference,

$$G = \frac{\sum_{i=1}^N \sum_{j=1}^N |k_i - k_j|}{2N \sum_{i=1}^N k_i}$$

where N is the size of the network and k_i is the degree of node i .

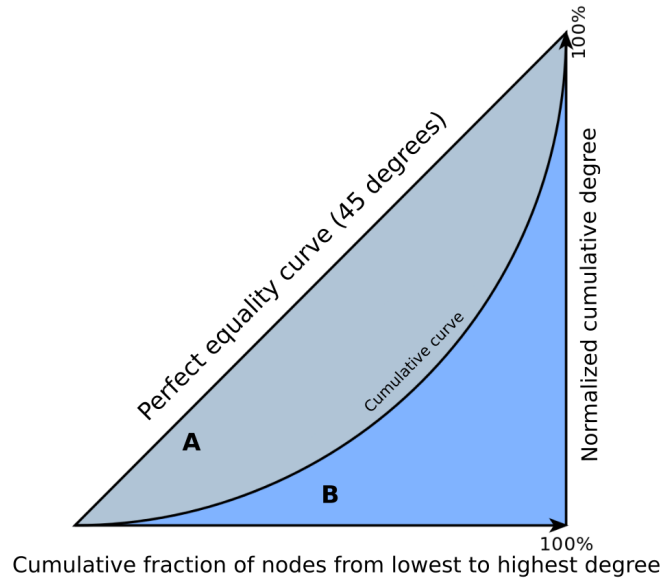


Fig 1. Sketch of the geometric idea behind the Gini index.

- Assortativity index.** In sociology it is well-known the phenomenon that people tend to interact more with those peers that think alike. This organizational tendency of relationships --- friendships, business relations, acquaintances, etc. --- extends to race, age, nationality, language, income, educational level, among many other features. This is called homophily or assortative mixing, and can be measured and visualized with different methods. Our proposal is to apply these methods to cultural systems in which there is a networked structure sustaining the interactions, be them people or activities/products created by people. This way, we can ascertain the level of homophily in different cultural socio-technical systems. To do so, we only need the topology of the interactions and a feature defined at a node level.

Qualitatively, a system is assortative when a significant fraction of the edges run between nodes of the same type, and it is disassortative when the contrary occurs.

Quantitatively, the assortativity index can be defined as

$$AI = \frac{1}{2m} \sum_{i=1}^N \sum_{j=1}^N (A_{ij} - \frac{k_i k_j}{2m}) \delta(g_i, g_j),$$

which is nothing more than the total number of edges running between nodes of the same type ---the number of edges within groups--- minus the number of edges between nodes if edges would be placed uniformly at random. We need to subtract this quantity to obtain large assortativity indexes in non-trivial cases but small in trivial ones. In the above equation m stands for the total number of edges, A_{ij} is a matrix whose elements are 1 if nodes i and j are connected and 0 otherwise, k_i is the

number of edges of node i and $\delta(g_i, g_j)$ is a Kronecker delta, where g_i is the type of group class of node i .

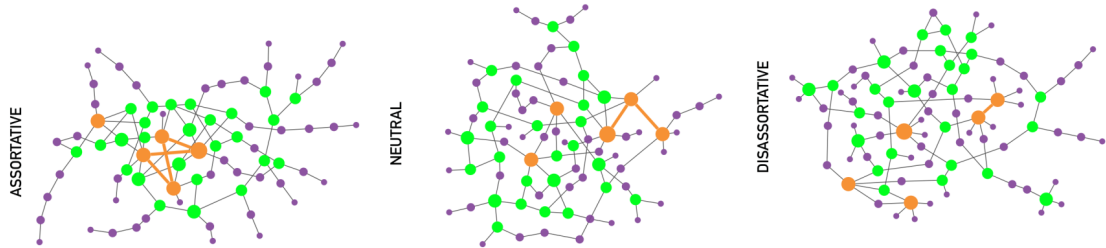


Fig 2: The three typical assortativity patterns in networks, where the color code is related to the number of connections. In positively assortative graphs, hubs are connected among themselves, while low-degree nodes do similarly among them. In neutrally assortative networks there are preferences and links can be thought to be drawn uniformly at random. Finally, in disassortative or negatively assortative networks, hubs tend to connect to low-degree nodes.

Fig. from *Network Science*, A.L. Barabasi, CUP (2015).

- Clustering coefficient.** This indicator is intimately related to the notion of transitivity and social balance theory. In common parlance, it is said that “the friend of my friend is also my friend”, and other combinations follow changing *friend* by *enemy*. This indicator deals with triadic relations, therefore going beyond the dyadic relationships encoded in the edges of a network. If a node i is connected to nodes j and k , then the clustering coefficient gives an estimation of the likelihood that j and k are connected as well (see Fig.2), i.e., we are measuring the compactness of social relations and it is a first step towards understanding social balance relationships. First we define the local clustering C_i of a node i as the ratio between the number of pairs of neighbors of i that are actually connected and the total number of pairs of neighbors of i . The latter is always equal to $k_i(k_i - 1)/2$. We can compute the clustering coefficient by just averaging over all nodes, hence

$$C = \frac{1}{N} \sum_{i=1}^N C_i.$$

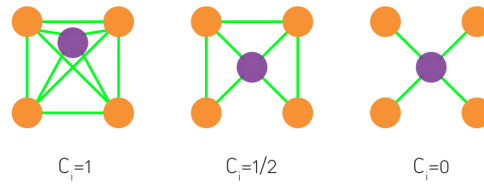


Fig 3: Sketch showing the idea behind the local clustering coefficient. On the left, the clique is complete, thus the clustering is 1. On the right, none of the neighbors of the central node are connected among them, thus its clustering is 0. In the middle, an intermediate situation is shown.

- Average shortest connectivity length.** A path in a network is the distance, counted as number of links, between a source node and a target node. The shortest path between those nodes is the most probable path through which information can flow, becoming less probable to take alternative paths if they are longer. This indicator computes the average value of this shortest path length, among all pairs of nodes, so it sheds light on how quick information can flow across the network. High values of this average correspond to networks displaying the so-called “small-world effect”, where two random nodes are much more closer than in well-organized, structured networks, while low values are related with networks that are not well-integrated and global communication is difficult. To compute it, we just need to average over the distance of all pairs of nodes, i.e.,

$$l = \frac{1}{N^2} \sum_{ij} d_{ij}.$$

There are efficient algorithms to compute the shortest path distance, e.g., like those of Prim’s or Dijkstra’s.

- Network efficiency.** This indicator determines how efficiently a network exchanges information and integrates in a single measure the two previous indicators, identifying $1/l$ and C as first approximations of the Efficiency evaluated on a global and on a local scale, respectively. It is a handy indicator for complicated networks, displaying link weights, disconnected clusters and sparseness. It is defined as

$$E = \frac{1}{N(N-1)} \sum_{i=1}^N \sum_{j=1, j \neq i}^N \frac{1}{d_{ij}}.$$

- Community modularity.** All the indicators so far give information about either the local or the global properties of the network of interactions. However, in the middle of these two extremes, there is usually a non-trivial mesoscale structure, many times in the form of community structure. Communities, roughly speaking, are groups of nodes that are tightly connected among them and sparsely connected to other groups. Community modularity gives an idea of the amount of isolation between the

nodes of a community from the rest of the network, being large for densely connected communities, 0 for random networks, null models, or single-community networks, and it is negative when there is a number of communities of the order of the network size. A good community partition is identified with a maximum value of the modularity. However, it is a difficult task to obtain such a maximum value. Luckily, there are very efficient algorithms, readily available, that implement the modularity maximization such as the Louvain algorithm.

- **Burstiness coefficient:** Typical temporal activation patterns found in socio-technical systems display considerable levels of irregularities. Bursty time series are characterized by small periods of high activity, followed by long periods of inactivity, therefore inducing very heterogeneous, often power-law, distributions of inter-event times. An inter-event τ is the time elapsed between two consecutive events, e.g., sharing the same artwork on a social platform or entering in the same online exhibition.

To quantify this phenomenon of burstiness, the burstiness coefficient B can be employed as indicator:

$$B = \frac{\sigma_{\tau} - m_{\tau}}{\sigma_{\tau} + m_{\tau}},$$

where m_{τ} and σ_{τ} are the mean and the standard deviation of the inter-event time probability distribution. The coefficient is bounded between 1, corresponding to very heterogeneous activity patterns, and -1 , corresponding to a completely regular, periodic signal.

- **Fano factor:** In line with the previous indicator, an alternative way to measure the heterogeneity of the temporal activity patterns is the Fano factor, which is broadly employed in neuroscience research and statistics,

$$F = \sigma_W^2 / m_W.$$

W indicates a finite time window. The Fano factor measures the variability in the number of events within W in relation to the mean number of events.

- **Memory coefficient:** Both temporal indicators presented above assume that consecutive actions are independent of each other. However, this is not usually true, since there might be correlations that reinforce behaviors, for example, the more a user is familiarized with, say, a tool or a technique, the more he/she is eager to use it. Moreover, a temporal series can be bursty even if the burstiness coefficient is not positive, just by virtue of these correlations (see Fig 4). An indicator that is able to measure these correlations is the memory coefficient,

$$M = \frac{1}{n_\tau - 1} \sum_{i=1}^{n_\tau - 1} \frac{(\tau_i - \langle \tau \rangle_1)(\tau_{i+1} - \langle \tau \rangle_2)}{\sigma_1 \sigma_2},$$

where $n_\tau + 1$ is the total number of events, and $\langle \tau \rangle_1$ (respectively $\langle \tau \rangle_2$) and σ_1 (respectively σ_2) are the sample mean and standard deviation of the inter-event times $\{\tau_i | i = 1, \dots, n_{\tau-1}\}$ (respectively $\{\tau_{i+1} | i = 1, \dots, n_{\tau-1}\}$).

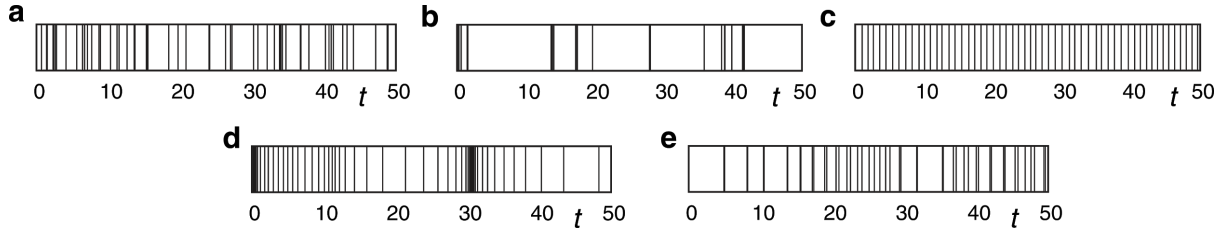


Fig 4: Time series with different burstiness and memory coefficients. Every vertical tick corresponds to an event, and the inter-event time is the horizontal distance between two consecutive ticks. In **a**, $(B,M) = (-0.05, 0.02)$. In **b**, $(B,M) = (0.44, -0.04)$. In **c**, $(B,M) = (-0.81, -0.02)$. In **d**, $(B,M) = (-0.05, 0.90)$. In **e**, $(B,M) = (-0.05, -0.74)$.

Fig. readapted from K.I. Goh & A.L. Barabasi. EP, 81(4), 48002 (2008).

Annex 2. Indicators on Economic Impact

INDICES OBJECTIVES	Selected resources			
	DESI DIMENSIONS	SUBDIMENSIONS	SUBD number	SUBD definition
	1. CONNECTIVITY			
Understanding the need of all publics				
supporting highly qualified professionalism		1a. Fixed Broadband take up	1a1	Overall fixed broadband take up
promoting the cultural use of new technologies			1a2	At least 100 Mbps fixed BB take-up
stimulating learning through culture		1b Fixed broadband coverage	1b1	Fast BB (NGA) coverage
positioning access to cultural upstream and transversally in all cultural policy making				Fixed very high capacity network coverage
		1c Mobile Broadband	1c1	4G coverage
			1c2	Mobile broadband take up
			1c3	5G readiness
		1d Broadband price index	1d1	Broadband price index
Understanding the need of all publics	2. HUMAN CAPITAL			
supporting highly qualified professionalism		2a Internet user skills	2a1	At least Basic Digital Skills
promoting the cultural use of new technologies			2a2	Above basic digital skills
positioning access to cultural upstream and transversally in all cultural policy making			2a3	At least basic software skills
overcoming linguistic barriers		2b Advaced skills and development	2b1	ICT specialists
			2b2	Female ICT specialists
			2b3	ICT graduates
	3. USE OF INTERNET SERVICES			
understanding the need of ALL publics		3a Internet use	3a1	People who never used the internet
-supporting highly qualified professionalism			3a2	Internet users
- improving funding & procedures		3b Activities on line	3b1	News
-overcoming linguistic barriers			3b2	Music, videos and games
-Advancing mobility and exchange			3b3	Video on demand

-promoting the cultural use on new technologies			3b4	Video calls
- stimulating learning through culture			3b5	Social networks
positioning access to cultural upstream and transversally in all cultural policy making			3b6	Doing an online course
		3c Transactions	3c1	Banking
			3c2	Shopping
			3c3	Selling online
	4. INTEGRATION OF DIGITAL TECHNOLOGY			
understanding the need of ALL publics		4a Business digitisation	4a1	Electronic Information sharing
-supporting highly qualified professionalism			4a2	Social media
- improving funding & procedures			4a3	Big Data
-overcoming linguistic barriers			4a4	Cloud
-Advancing mobility and exchange		4b E-commerce	4b1	SMEs selling online
-promoting the cultural use on new technologies			4b2	E-commerce turnover
- stimulating learning through culture			4b3	Selling on line cross-border
	5. DIGITAL PUBLIC SERVICES			
supporting highly qualified professionalism		5a e-government	5a1	E-government users
improving funding & procedures			5a2	Pre-filled forms
positioning access to cultural upstream and transversally in all cultural policy making			5a3	Online service completion
raising awareness of the legal frameworks on access to culture			5a4	Digital public services for businesses
			5a5	Open data
	EUROSTAT DIMENSIONS	SUBDIMENSIONS	SUBD number	Subdimention definition
supporting highly qualified professionalism	1. CULTURAL EMPLOYMENT			
		males		
		females		
		from 15 to 29 years		

		from 30 to 39 years		
		from 40 to 49 years		
		from 50 to 59 years		
		from 60 to 64 years		
		65 years or over		
		less than primary, primary and lower secondary education (levels 0-2)		
		upper secondary and post-secondary non tertiary education (levels 3-4)		
		tertiary education (levels 5-8)		
		no response		
		printing and reproduction of record and media		
		other manufacturing		
		publishing activities		
		motion picture, video and television programme production, sound recording and music publishing activities		
		programming and broadcasting activities		
		other professional, scientific and technological activities		
		creative, arts and entertainment activities		
		libraries, archives, museums and other e-activities		
		other NACE activities		
		no response		
		Employees with a permanent job		
		Self employed persons		
		Employed persons working full-time		
		Employed persons with one job only		

		Individuals, 15 to 29 years old		
		Individuals with high formal education		
		Employees with a permanent job		
		Self-employed persons		
		Employed persons working full time		
		Employed persons with one job only		
supporting highly qualified professionalism	2. ENTERPRISES IN CULTURAL SECTORS			
		Book publishing		
		Publishing of journals and periodicals		
		Publishing of computer games		
		Motion picture, video, and TV programme production, sound recording and music publishing activities		
		Programmes and broadcasting activities		
		News agency activities		
		Architectural activities		
		Specialised design activities		
		Economical indicator for structural business statistics		
supporting highly qualified professionalism	3. INTERNATIONAL TRADE IN CULTURAL GOODS			
supporting highly qualified professionalism	4. INTERNATIONAL TRADE IN CULTURAL SERVICES			
understanding the need of ALL publics	5. CULTURAL PARTECIPATION			
stimulating learning through culture				
promoting the cultural use on new technologies				
		Frequency of participation in cultural or sport activities in the last 12 months by sex, age, educational attainment level and activity type (ilc_scp03)	Not in the last 12 months	
			From 1 to 3 times	

			At least once	
			Cultural activities (cinema, live, performance or cultural sites)	
			Cinema	
			Live performances (theatre, concerts, ballets)	
			Cultural sites (historical monuments, museums, art galleries or archaeological sites)	
			Less than primary, primary and lower secondary education (levels 0-2)	
			Upper secondary and post-secondary non-tertiary education (levels 3-4)	
			tertiary education	(levels 5-8)
			males	
			females	
		Frequency of participation in cultural or sport activities in the last 12 months by income quintile, household type, degree of urbanisation and activity type (ilc_scp04)		
			Not in the last 12 months	
			From 1 to 3 times	
			At least once	
			Cultural activities (cinema, live, performance or cultural sites)	
			Cinema	
			Live performances (theatre, concerts, ballets)	
			Cultural sites (historical monuments, museums, art galleries or archaeological sites)	
			Cities	
			Towns and suburbs	
			Rural areas	

			single person	
			two adults	
			two adults with dependent children	
			three or more adults	
			first quintile	
			second quintile	
			third quintile	
			fourth quintile	
		Reasons of non-participation in cultural or sport activities in the last 12 months by sex, age, educational attainment level and activity type (ilc_scp05)		
			Financial reasons	
			No interest	
			None in the neighbourhood	
			other	
			Cultural activities (cinema, live, performance or cultural sites)	
			Cinema	
			Live performances (theatre, concerts, ballets)	
			Cultural sites (historical monuments, museums, art galleries or archaeological sites)	
			Less than primary, primary and lower secondary education (levels 0-2)	
			Upper secondary and post-secondary non-tertiary education (levels 3-4)	
			tertiary education	
		Reasons of non-participation in cultural or sport activities in the last 12 months by income quintile, household type, degree of urbanisation and activity type (ilc_scp06)		

			Financial reasons	
			No interest	
			None in the neighbourhood	
			Cultural activities (cinema, live, performance or cultural sites)	
			Cinema	
			Live performances (theatre, concerts, ballets)	
			Cultural sites (historical monuments, museums, art galleries or archaeological sites)	
			First quintile	
			Fifth quintile	
			Households without depend children	
			Households with depend children	
		Frequency of practicing of artistic activities by sex, age and educational attainment level (ilc scp07)		
			every day	
			every week	
			once a month	
			several times a month	
			not in the last 12 months	
			at least one a years	
			Less than primary, primary and lower secondary education (levels 0-2)	
			Upper secondary and post-secondary non-tertiary education (levels 3-4)	
			tertiary education	
			From 16 to 29 years	
			16 years or over	
			from 65 to 74 years	
			males	
			females	

		Frequency of practicing of artistic activities by income quintile, household type and degree of urbanisation (ilc_scp08)		
			every day	
			every week	
			once a month	
			several times a month	
			not in the last 12 months	
			at least one a years	
			Households without depend children	
			Households with depend children	
			first quintile	
			second quintile	
			third quintile	
			fourth quintile	
			Cities	
			Town and suburbs	
			Rural areas	
		Persons reading books in the last 12 months by sex and age (cult_pcs_bka)		
			years....	
			1 book or more	
			less than 5 books	
			from 5 to 9 books	
			10 books or more	
			males	
			females	
			from 25 to 64	
		Persons reading books in the last 12 months by sex and educational attainment level (cult_pcs_bke)		

			years....	
			1 book or more	
			less than 5 books	
			from 5 to 9 books	
			10 books or more	
			males	
			females	
			Less than primary, primary and lower secondary education (levels 0-2)	
			Upper secondary and post-secondary non-tertiary education (levels 3-4)	
			tertiary education	
		Persons reading newspapers in the last 12 months by sex (cult_pcs_nws)		
			years...	
			every day	
			at least once a week	
			less than once a month	
			at leas once a month	
			never	
			males	
			females	
		Persons reading newspapers in the last 12 months by age (cult_pcs_nwa)		
			years...	
			every day	
			at least once a week	
			less than once a month	
			at leas once a month	
			never	
			from 25 to 64 years	
		Persons reading newspapers in the last 12 months by educational attainment level (cult_pcs_nwe)		

			years...	
			every day	
			at least once a week	
			less than once a month	
			at least once a month	
			never	
			Less than primary, primary and lower secondary education (levels 0-2)	
			Upper secondary and post-secondary non-tertiary education (levels 3-4)	
			tertiary education	
		Persons participating in cultural activities in the last 12 months by sex and age (cult_pcs_caa)		
			years...	
			Cultural activities (cinema, live, performance or cultural sites)	
			Cinema	
			Live performances (theatre, concerts, ballets)	
			Cultural sites (historical monuments, museums, art galleries or archaeological sites)	
			never	
			from 1 to 6 times	
			at least once	
			more than 6 times	
			no response	
		Persons participating in cultural activities in the last 12 months by sex and educational attainment level (cult_pcs_cae)		
			years...	
			Cultural activities (cinema, live, performance or cultural sites)	
			Cinema	

			Live performances (theatre, concerts, ballets)	
			Cultural sites (historical monuments, museums, art galleries or archaeological sites)	
			never	
			from 1 to 6 times	
			at least once	
			more than 6 times	
			no response	
			males	
			females	
			Less than primary, primary and lower secondary education (levels 0-2)	
			Upper secondary and post-secondary non-tertiary education (levels 3-4)	
			tertiary education	
		Individuals - internet activities (isoc_ci_ac_i)	years...	
			Internet use: creating websites or blogs	
			Internet use: consulting wikis (to obtain knowledge on any subject)	
			Internet use: reading online news sites, newspapers/ news magazines	
			Internet use: playing or downloading games	
			Internet use: listening to music (eg. Web radio, music streaming)	
			Internet use: watching internet streamed TV (live or catch-up) from TV broadcasters	
			Internet use: watching video content from sharing services	
			Internet use: watching video content from commercial or sharing services	

			Internet use: playing or downloading games, listening to music, or watching internet streamed TV or videos	
			Percentage of individuals	
			Percentage of individuals who used Internet in the last 3 months	
			Individuals, 16 to 24 years old	
			Individuals, 25 to 64 years old	
			Individuals, 65to 74 years old	
			Individuals, 75 years old or more	
		Internet purchases by individuals (until 2019) (isoc_ec_ibuy)	years...	
			online purchases: films/music	
			online purchases: books/magazine newspapers	
			online purchases: ticket for events	
			online purchases: films/music, delivered or upgraded online	
			online purchases, downloaded or accessed from websites or apps: e-books	
			online purchases, downloaded or accessed from websites or apps: e-books, e-magazine, e-newspapers	
			online purchases, downloaded or accessed from websites or apps:films/music, e-books, e-magazines, e-newspapers	
			percentage of individuals who used Internet with the last years	
		Purpose of mobile internet use (isoc_cimobi_purp)	percentage of individuals	
			Percentage of individuals who used	

			Internet in the last 3 months	
			Percentage of individuals who used a handheld device to access internet	
			Mobile internet use via handheld device: for reading or downloading online news/newspapers/new magazines	
			Mobile internet use via handheld device: for reading or downloading online books or e-books	
			Mobile internet use via handheld device: for downloading games, images, video or music	
			Mobile internet use via handheld device: for using podcast service to automatically receive audio or video files of interest	
			Individuals, 16 to 24 years old	
			Individuals, 25 to 34 years old	
			Individuals, 35 to 54 years old	
			Individuals, 55 to 64 years old	
		Individuals - use of cloud services (isoc_cicci_use)		
			years...	
			Percentage of individuals who used Internet in the last 3 months	
			used internet storage space to save documents, pictures, music, videos or other files	
			internet storage space use: to save or share documents, pictures, music, videos or other files	
			internet storage space use: to save or share photos	
			internet storage space use: to save or share e-books or e-magazines	

			internet storage space use: to save or share music	
			internet storage space use: to save or share videos including films, tv programmes	
			Reason for using internet storage space use: access to large libraries of music, tv programmes or films	
			Used software run over the internet for editing pictures or videos	
			used internet but did not use software run over the internet for editing pictures or videos	
			used internet to save or share files or edit pictures or videos	
			used services over the internet for plying music or video files uploaded or saved in internet storage space	
			used the internet but did not use services over the internet for plying music or video files uploaded or saved in internet storage space	
			used internet storage space to save or share files or play music or videos	
		Time spent, participation time and participation rate in the main activity by sex and age group (tus 00age)		
			Cinemas in the city: very satisfied	
			Cinema in the city: rather satisfied	
			Cinema in the city: rather unsatisfied	
			Cinema in the city: not at all satisfied	
			Cinema in the city: don't know/not answer	
			Cultural facilities such as concert halls, theatres, museums and libraries in the city: very satisfied	

			Cultural facilities such as concert halls, theatres, museums and libraries in the city: rather satisfied	
			Cultural facilities such as concert halls, theatres, museums and libraries in the city: rather unsatisfied	
			Cultural facilities such as concert halls, theatres, museums and libraries in the city: not at all satisfied	
			Cultural facilities such as concert halls, theatres, museums and libraries in the city: don't know?no answer	
		Time spent, participation time and participation rate in the main activity by sex and household composition (tus_00hhstatus)		
			total	
			time spent (hh:mm)	
			participation time (hh:mm)	
			participation rate (%)	
			males	
			females	
			free time study	
			handcraft and producing textiles and other care for textiles	
			teaching, reading, and talking with child	
			entertainment and culture	
		Time spent, participation time and participation rate in the main activity by sex and educational attainment level (tus_00educ)		
			Less than primary, primary and lower secondary education (levels 0-2)	
			Upper secondary and post-secondary non-tertiary education (levels 3-4)	

			tertiary education	
			time spent (hh:mm)	
			participation time (hh:mm)	
			participation rate (%)	
			free time study	
			handcraft and producing textiles and other care for textiles	
			teaching, reading, and talking with child	
			entertainment and culture	
		Time spent, participation time and participation rate in the main activity by sex and self-declared labour status (tus_00selfstat)		
			employed persons working full-time	
			employed persons working part-time	
			unemployed person	
			students	
			homemakers	
			retired persons	
			time spent (hh:mm)	
			participation time (hh:mm)	
			participation rate (%)	
			free time study	
			handcraft and producing textiles and other care for textiles	
			teaching, reading, and talking with child	
			entertainment and culture	
			computer games	
			computing	
			hobbies and games except computing and computer games	
			reading	
			reading, except books	
			TV and videos	
			radio and music	