French Gaia-X Hub

POSITION PAPER

WG Cultural and Creative Industries

v.1.2 – 06.01.2022

gaia-x

POSITION PAPER

WG Cultural and Creative Industries

Prefigurator of the working group :

Louis-Cyrille Trébuchet, CIO, France Télévisions

Rapporteur :

Véronique Demilly, Project Manager, France Télévisions
Veronique.Demilly@francetv.fr

Table of contents

1 LIST OF PARTICIPANTS

As of 17 May 2022, the participants in the working group are from the following organisations:

- Arenamétrix
- ARPP
- BnF
- BnF Partenariats
- Cap Digital
- Comité Colbert
- CST
- Dag Conseil
- Deezer
- EBU
- Epsilon – groupe Publicis
- France Télévisions
- France Télévisions Publicité
- Frédéric Goldsmith
- Hachette
- INA
- Institut Français
- Médiamétrie
- Minalogic
- Ministère de la Culture
- Monark
- Le Monde
- Nomalab
- One Point
- Panodyssey
- Plurimédia
- Procirep
- Radio France
- RMN – Grand Palais
- Sacem
- Salto
- Singulart
- SNPTV

2 THE CULTURAL AND CREATIVE INDUSTRIES SECTOR

OVERALL DÉFINITION

The cultural and creative industries (CCI) are those sectors of activity whose main purpose is the creation, development, production, reproduction, promotion, dissemination or marketing of goods, services and activities that have a cultural, artistic and/or heritage content.

They can be recognised by several characteristics:

- The intersection between economy and culture;
- Creativity at the heart of the activity;
- Artistic, cultural or creative content inspired by the creation from the past;
- The production of goods and services frequently protected by intellectual property - copyright and related rights;
- The dual nature: economic (generation of wealth and employment) and cultural (generation of values, meaning and identity);
- Innovation and creative renewal;
- Public demand and behaviour that are difficult to anticipate;
- A sector marked by the diversity of methods of remuneration of work, in particular wage-earning and the payment of literary and artistic property rights, as well as by the predominance of micro-enterprises.

These industries have been defined as « any industry that originates from individual creativity, skill and talent and has the potential to generate wealth and employment through the creation and exploitation of intellectual property ».

THE CCI SECTORS

The overall definition of the Cultural and Creative Industries clearly states that it is a grouping of various sectors. Depending on the study or source, the groupings include more or less business sectors.

Thus, the EY[1] study for France Créative, whose figures are frequently quoted, defines the ICC sector as comprising the following 10 business sectors :

- Visual arts ;
- Music ;
- Performing arts ;
- Cinema;
- Television ;
- Radio ;
- Video games;
- Book
- Press

[1] « *L'économie mosaïque - 3e Panorama des Industries Culturelles et Créatives en France* » - study EY and France Créative – November 2019

- Advertising and communication.

The European Commission, in its proposal for a regulation establishing the Creative Europe programme for 2021 to 2027, considers that the CCI sector includes all sectors whose activities are based on cultural values or on artistic and other creative expressions, whether individual or collective.

These activities may include the development, creation, production, dissemination and preservation of goods and services embodying cultural, artistic or other creative expression, as well as related tasks such as education or management. They will have the potential to create innovation and employment, in particular through intellectual property. These sectors thus include architecture, archives, libraries and museums, arts and crafts, audiovisual (including film, television, video games, multimedia) and advertising, tangible and intangible cultural heritage, design (including fashion), festivals, music, theatre, performing arts, literature, publishing, radio and visual arts.

With a view to development on a European scale, the Cultural and Creative Industries working group of the French Hub Gaia-X is likely to welcome any company meeting the European definition of the CCI sector.

HETEROGENEITY IN THE SECTOR

Unlike other sectors, the CCI sector is made up of structures of very different sizes and operating methods. Alongside large private companies or large public institutions, there are many very small companies or structures based on associative models, with very few permanent staff, self-employed people and casual entertainment workers. In some parts of the CCI sector, the majority of the structures are micro-enterprises.

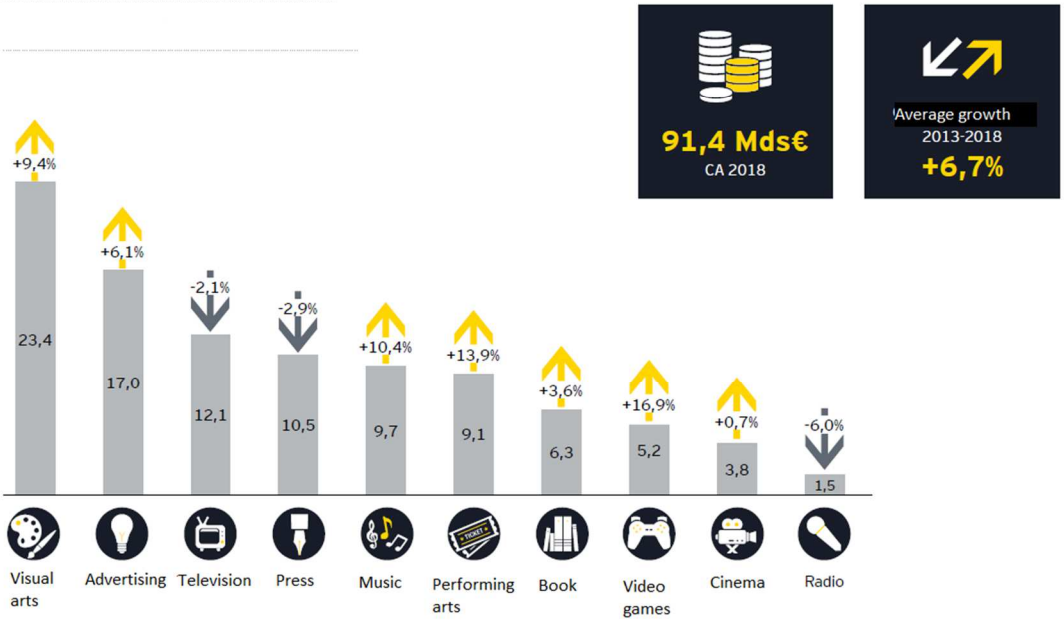
A group of 600 professions, 300,000 companies and 1.3 million people contribute to the vitality and growth of the CCI sector in France (1.2 million firms and over 8 million people contribute in Europe for an economic weight of 3.95% of EU value added – EUR 477 billion).

Revenues in the CCI sector come from a variety of sources. It can be income from commercial activity as well as from public subsidies. The CCI sector is subject to regulations specific to its different components, which respond to the objective of preserving cultural diversity as well as to other public interest objectives such as the strengthening of social links, the development of educational activities, the leverage effects on tourism activity or territorial attractiveness.

Given the different structures within the cultural and creative industries, many of them have few resources to develop their skills in the digital domain, particularly when this is not their core business. There are significant maturity gaps on data and IT issues within the sector.

CCI Income in 2018 in France

Source : Analyses EY, « Le poids économique direct de la culture en 2017 », DEPS du ministère de la Culture / INSEE



3 IMPORTANCE OF DATA FOR THE SECTOR

BACKGROUND

1. Digitisation of the CCI economy

The widespread of the digitalisation of industry is one of the key elements in the conquest of the European Union's strategic autonomy. The Cultural and Creative Industries sector is no exception to this ambition.

All CCI sectors are now affected by this transformation.

While, for the recorded music, video game and audiovisual sectors, new creations have been digital from start to finish for several years now, from their creation to their exploitation in all its forms, for certain other industries, such as cinema, live performance or books, essential uses are still dependent on the movement of members of the public to physical locations to access works, such as cinema screenings, live performances in public places or the sale of books in bookshops. However, digitisation is also affecting these modes of access through the digitisation of theatrical projection tools, the development of live streaming or the electronic book or the sale of physical books on online selling and/or referencing platforms.

Digitisation is not necessarily synonymous with total dematerialisation or the disappearance of physical modes of access to works. On the other hand, it is conceived in a systemic way as a transformation of industrial processes in order to take advantage of data, networks and computing power available for their creation, exploitation and discovery. It disrupts the relationship with the public.

The business models of the cultural and creative industries are diversifying thanks to the possibilities offered by the combination of digital and electronic communications. The development of artificial intelligence, metavers, NFTs, as well as immersive reality, offer new perspectives whose speed of realisation requires constant innovation, challenging the cultural industries to constantly adapt and, better still, to anticipate the upheavals to come.

Among the major business models available, all based on intellectual property rights and services offered to the public, are :

- the sale of products such as physical media (books, CDs or vinyl, DVDs or Blu-Rays, etc.) or virtual media (NFT, etc.);
- the sale of access rights to a show, a work, a heritage object, etc., in a physical location or online, generally based on a « ticketing » system that is increasingly digitised and rich in usage data;
- revenues linked to subscriptions to content aggregation services, linear or non-linear, centralised or not;
- revenues from public exposure to advertising, the targeting of which is a growing challenge, whether for linear services (radio, television, etc.) or not, sometimes combined with the subscription model (freemium).

2. Platforms for accessing works

The « platformisation » of access to works consists of the aggregation, by a third party, of creations in the form of digital files which are then made available to the public via its indexing and recommendation tools.

The development of communication networks, their democratisation and the emergence of new technologies and materials have led to the emergence of many new uses, driven to a large extent by global players with considerable resources and representing real « gatekeepers » to works. These upheavals must not allow the value of works to escape their creators and economic partners. Appropriate economic balances must therefore be constantly sought in order to distribute the value fairly between the stakeholders and the new intermediaries on the electronic communications networks.

An essential feature of this « platformisation » is the automation of part of the modalities of access to works and therefore of the public's capacity to discover them, through algorithms, for which digital data represent the « fuel ». As experience goods, works must be accessible in the form of extracts or, in any case, referenced information, so that the public can become interested in them. Any bottleneck in access to works has considerable distorting effects if alternatives are not available with equal ease.

This is why the value of uses tends to remain with the intermediaries who, through their power, can create an imbalance in their favour, which requires regulation.

The purpose of legislation is to re-establish the balances necessary for a dynamic circulation of works compatible with the diversity of their creation and distribution.

One of the challenges is the availability of data on works for operators who make them available to the public or reference them in order to promote public access to them, as well as data on the uses of these works for the benefit of the creative industries to guide their activity and develop their dynamism.

3. A European strategy

The European data strategy should enable the EU to become a leading player in a data-driven society. The creation of a single market for data will make the flow of data within the EU and between sectors more fluid, to the benefit of businesses, creators, public administrations and the public.

Data is at the heart of the digital transformation. It shapes the way we produce, consume and live. Access to and the ability to use an ever-increasing amount of data is essential for innovation and growth. Data-driven innovation can bring significant tangible benefits to citizens - for example in the form of improved mobility - and to the European economy, such as more efficient policy-making and better public services.

Several texts in preparation or recently adopted thus aim to establish an open, secure, trusted and transparent ecosystem for a better flow of data across Europe and the development of new business models. The ambition is to create a single market for data in line with the values of the European Union.

- Data Governance Act
- Digital Market Act (DMA)
- Digital Services Act (DSA)
- Data Act (DA)

European regulation, through the draft DSA and DMA regulations, has a considerable role to play in inserting new players into the European ecosystem and ensuring fairer competition with traditional operators, notably the media.

The Data Governance Act will put in place robust mechanisms to facilitate the re-use of certain categories of protected public sector data, build trust in data intermediation services and foster data altruism across the EU.

It is an important component of the European Data Strategy, which aims to strengthen the data-driven economy, increase wealth and well-being and give Europe a competitive edge for the benefit of its citizens and businesses.

The Data Governance Act creates a framework to promote a new business model - data intermediation services - that will provide a secure environment in which businesses or individuals can share data.

For businesses, these services can take the form of digital platforms that will allow voluntary sharing of data between businesses or facilitate compliance with data sharing obligations set by law. These services will allow businesses to share their data without fear of misuse or loss of competitive advantage.

4. French support

In line with the State's digital strategy, the Ministry of Culture published, on 15 September 2021, a "Cultural data and content" roadmap which aims, in particular, to strengthen the accessibility and discoverability of cultural content by enriching metadata and facilitating the sharing of this data.

In addition, several financing initiatives earmarked by the public authorities concern the cultural and creative industries, in particular from the Banque des territoires (CDC) and the BPI.

In addition, the 2030 recovery plan contains a substantial section on CCIs.

As regards digital data more specifically, the French government has supported the Gaia-X initiative since its inception and encourages the development of a European Cloud infrastructure and ecosystem capable of ensuring the free circulation of data and creating value in Europe.

In any case, innovation is a central axis for CCIs to benefit from public support related to digital matters.

THE DIFFÉRENT TYPES OF DATA

1. Contents

Digitised content, digital twins or natively digital content constitute a first type of data whose circulation must be ensured while protecting access. This data has an intrinsic value in that it constitutes the works themselves.

The issue of digitisation of works is also becoming crucial. These digitised works must be stored, identified and protected. The standardisation of formats, the description of content, the reversibility and maintenance of stored data are issues to be addressed by this working group.

2. Metadata

Metadata is data associated with another piece of data. It is essential in order to ensure the discoverability of the data to which it is associated and the remuneration of the rights holders. Its quality is therefore essential. Metadata covers many fields [2] :

Category	Description	Examples
Unique Identifiers	Strings of alphanumeric characters that serve to stably identify a document, resource or entity, regardless of its nature. In principle, an identifier should be unique for each resource.	ISBN (book) ISAN (film) ISRC (musical recording) ISWC (musical work) IPN (*International Performer Number*) ISNI (*International Standard Name Identifier* for authors, composers, performers, producers, etc.), Etc.
Descriptive metadata	Describe the content accurately and objectively	Author's name Title of the work Materials Date of creation Author's date of birth, etc.
Administrative metadata	Provide information on how the files were created, saved, modified.	Date of creation of the file Name of the company creating the file Format used Identification of the source document that was scanned, etc.
Legal (or intellectual property) metadata	Indicate whether the work is copyrighted, identify the rights holders and the collecting society, indicate whether the work is licensed for use, and determine the source statements that must appear with the work.	Name of the copyright owner or manager Owner's identifier Wording of the source statement to appear with the title of the work
Enrichment metadata	Provides additional and sometimes subjective information about content. The scope of enrichment metadata is vast. It can be used as an added value by digital cultural content providers to differentiate themselves from the competition. This enrichment data can be produced by artificial intelligence algorithms.	Biography of the artist Photographs Notes Song lyrics Artist's social network addresses Cover / poster image Text of a review, etc.

[2] Inspired from « *Etat des lieux sur les métadonnées relatives aux contenus culturels* » de l'Observatoire de la culture et des communications du Québec (2017).

In order to facilitate the development of applications and the exchange, sharing or enhancement of this data, it is becoming necessary for the sector to use common ontologies. To this end, the European Broadcasting Union (EBU) has developed and maintains the EBUCore and EBU CCDM standards which aim to facilitate the circulation of content between producers and broadcasters. This recommendation can be used without financial compensation. The widespread use of shared standards would simplify processes and increase efficiency.

Similarly, in an initial move to harmonise its various collections, the INA[3] has developed and deployed, within its management perimeter (legal deposit and regulatory devolution of public broadcasting), a data model freely inspired by FRBR / CIDOC CRM. This model is characterised by its scalability and flexibility, by the importance given to controlled vocabularies, articulated with very generic entities. It thus covers any type of object, audiovisual, written, photographic or web, broadcast or not, for which it describes both the intellectual nature (content) and the materialization, physical or digital (container). This principle of agnosticity and adaptability to any management context makes this model deployable in any content management environment, of any kind. This is why the INA proposes to open this model widely, to support its dissemination and facilitate its appropriation and use by all interested parties.

3. Consumption data

Consumer data includes both usage data generated automatically when a particular cultural content is consumed online and audience or sales data (or measurements) collected in a more traditional way.

This data is essential for understanding uses and markets and for adapting offers to demand. They are also necessary for the valuation of advertising space, which is an essential source of revenue for certain CCI players.

The DMA provides, in particular, for imposing access for content providers to certain consumption data collected by the major dominant platforms (« gatekeepers »). The Data Act (legislative proposal published in February 2022) aims to stimulate the sharing of consumer data generated through the use of certain connected objects, as well as to encourage the development of standards for the establishment of « data spaces ».

4. Personal data

Personal data are data relating to an identifiable natural person. It is not forbidden to use this type of data, but special precautions must be taken. The processing of personal data is highly regulated. The European Union has developed a specific regulation dealing with the particularity of this data: the GDPR (General Data Protection Regulation) of 27 April 2016.

As far as the CCI sector is concerned, this data comes into play in particular in the management of subscriptions to a platform or in the targeting of advertising and the management of consents.

5. Other data

Each industrial process generates its own data flow: logistic data, machine logs, energy consumption, execution reports, etc. This data can also be better exploited.

[3] INA (Institut National de l'Audiovisuel) : National Audiovisual Institute -

ENSURING DISCOVERABILITY THROUGH METADATA

Faced with the explosion in the amount of cultural content available and accessible, particularly via the Internet, visibility and the ability to connect a work with an ad hoc audience are becoming major issues.

The World Wide Web is itself an almost infinite space of exploration in which the public evolves thanks to tools such as search engines that use algorithms leading to natural referencing (SEO) or paid referencing (SEM) of content.

Making a work visible therefore means associating it with metadata that will enable algorithms to detect its compatibility with users' tastes or to propose it as a discovery. In a number of CCI sectors, this metadata often remains fragmented and incompatible with the metadata of other sectors due to the lack of common standards. This results in under-use or use limited to compatible platforms. There are, however, useful mappings for moving from one system to another.

The first challenge is therefore to generalise the use of one or other of the existing systems and to start the movement towards a form of standardisation. The second challenge will be to ensure the quality of the metadata associated with the works. A form of governance should probably be put in place. Such governance would make it possible, in particular, to establish good practice guides for metadata relating to cultural content.

REFINING PROPOSALS WITH USAGE DATA

Whatever the business model, today, access to consumer usage data, data related to their Internet browsing habits or their tastes, and audience data, is essential. This data makes it possible, in particular, to adapt the supply of content, its promotion and/or exploitation methods, to develop new offers or to optimise advertising communication and measure its effects.

When cultural players obtain access to data on the exploitation of their content (either from platforms and service distributors or from their own distribution channels such as their website), they do not always have the means (technical, human or financial) to harmonise, analyse and extract all the information they could benefit from, particularly with a view to establishing an optimal strategy for the discoverability of their content. This is particularly the case for smaller cultural structures.

Moreover, each player in the value chain currently collects this data on its own behalf and processes it separately, which only gives it a very fragmented view of users' tastes and needs.

The entire cultural sector would benefit from the most exhaustive pooling of all this data possible. The resulting analyses would be more refined and would allow better targeting of cultural offers. The public would also benefit, thanks to a better match between the proposals and their tastes, interests and prospects for discovery

GARANTEEING THE REMUNÉRATION OF RIGHT HOLDERS

In view of the mass and diversity of the ways in which works are exploited, the size of the flows, the variety of right holders, contracts and economic models, the global nature of the exploitations that must comply with different legislation and the speed of reaction required to keep up with technical developments and uses, the capacity to process data associated with works is crucial to ensure the management of rights and remuneration for right holders and their economic partners. It is obviously

essential that all the data relating to a work can be processed by the same processes in order to ensure fair remuneration for rights holders. In particular, content consumption data on large platforms should be made accessible under reasonable conditions.
The work on data formats and the means of collecting and analysing them is at the heart of the capacity of the cultural and creative industries to be fully involved in the changes underway.
It should also be noted that many players collect and process the same data with slightly different objectives, but that there are many opportunities to pool and simplify data flows.

RATIONALISE RESOURCES

In order for the different cultural domains to benefit from their data, companies and organisations would benefit from working together and channelling their efforts, taking into account the scale of the task for the smallest organisations in the different sectors, which do not have the same means as the large ones to invest in technological infrastructures or in the standardisation of their metadata.

These challenges cannot be met without the adoption of a common language (or at least harmonised languages) and a global vision by the different stakeholders in the cultural fields, where the needs of their own organisation are no longer the only determinants of the importance of a metadata field. For this to happen, it is important that everyone is aware of the potential of data for the full development of the cultural sector and that everyone understands the impact and interest for the development of their activity.

It is possible to establish links between the various cultural sectors or domains, bridges that would allow for a minimal convergence. For example, the widespread adoption of the same naming system in the various cultural sectors would facilitate the bringing together of sources of varied origins. Whether the cultural content is a show, a film, a television programme, a book, a museum object, a musical recording or a heritage item, a single naming system would make it possible to identify the entities relating to this content in a unique way (name of the author, the place, the production company, etc.). It would distinguish names that might otherwise be confused and link data about a particular name.

INTEREST OF THE GAIA-X INITIATIVE

The Gaia-X initiative aims to develop an open, secure and transparent ecosystem allowing each actor in the economy to access resources of various kinds under transparent and non-discriminatory conditions. These resources can be computing capacities, data processing services, access to data, authentication services, etc.

Thanks to a labelling of services based on « Policy Rules » in line with European values, Gaia-X responds to Europe's ambition to establish a single market for data (and associated processing) in a secure, trusted and interoperable Cloud infrastructure.

Gaia-X also intends to support the creation of sectoral data spaces. The concept of data space refers to an ecosystem of exchange, sharing and provision of data between trusted partners, whether or not for a fee. It is not about copying or repatriating data centrally, but about ensuring that each data holder has full control over the conditions of access to their data: who, when, under what conditions. By giving visibility to existing data and facilitating access to different data sources, sectoral data spaces are intended to bring about new services and create value for Europeans, in Europe and beyond. The Gaia-X Association intends to develop architectural and/or software building blocks to ensure secure gateways between data spaces.

4 USE CASES

The use cases described in this section are intended to lead to the launch of development projects.

The realisation of these projects will constitute the basis for the creation of a sectoral data space for the Cultural and Creative Industries, regrouping the Media part and the Cultural Heritage part.

Parties interested in any of the initiatives described in the following section are invited to contact the French Gaia-X Hub.

CHARACTÉRISTICS OF GAIA-X PROJECTS

Projects supported by Gaia-X will be eligible for national and European funding.

In order to obtain support from Gaia-X, they must meet the association's objectives:

- To be able to move to a European scale;
- Participate in the creation of an open, non-proprietary ecosystem without a monopolistic approach;
- Adopt the rules and characteristics of the policy rules, architecture and labels and be able to be associated, in the long term, with a data space.

ORGANISATION OF USE CASES ACCORDING TO THE VALUE CHAIN

As mentioned in the previous chapters, the Cultural and Creative Industries sector includes a diversity of actors who all play a role in the creation of value in the sector. Their activities are related to:

- Creation ;
- Production;
- Manufacturing;
- Publishing;
- Distribution;
- Selling.

The use cases described below involve one or more levels of this value chain.

USE CASES FOR CREATION, PRODUCTION OR MANUFACTURING

1. Deposit and repository of works - Promotion of creation

1.1 Creation of a repository of Digital Twins (IWA project - Interoperability Within the Ark)

On the initiative of:
Monark, agency for the creation of digital twins of architectural heritage, sensitive natural sites and remarkable objects.

Potential participants:

Garou, multi-user XR platform for designers and companies, metavers.

The issue:

The digitisation of real estate, furniture and remarkable objects for the purposes of conservation, transmission, enhancement and communication makes sense in our contemporary digital society. The creation of digital twins of our architectural or pictorial heritage, of certain typical places, held by the owners of the works, could constitute a new source of income for the owners of the goods and of inspiration for the Cultural and Creative Industries sector.

To date, there is no professional platform to enable the CCI sector to fully exploit the resources linked to digital twins. Legal uncertainties, lack of sharing, interoperability, dialogue or lack of visibility often lead to the re-digitisation of works.

The major hyperscalers are investing massively in new distribution media such as virtual reality (e.g. Facebook with Meta), augmented reality or video mapping (architectural projection). In order to avoid the capture of value linked to digital twins by these actors, the IWA project aims at the emergence of a French or European actor who would organise the collection, storage, description, visibility and marketing of digital twins data according to precise and transparent specifications.

The project also aims to raise the awareness of CCI players on the importance of transposing their material assets into the virtual world and the benefits they could gain from it.

Solution provided:

IWA is a database of remarkable digital twins that can be used as virtual sets by film, gaming and contemporary art professionals...

The IWA project proposes to:

- Set up a platform to make digital twins available;
- Create digital twins;
- Secure the use of these twins through fingerprinting and NFTs in a blockchain;
- Commercialise (license) the use of digital twins by paying the owner of the physical asset.

Through a platform using WebGL (GPU) technology, all actors in the cultural and creative industries will be able to consult the available digital twins.

Expected benefits:

Consolidate the intangible heritage of the cultural and creative industries and enhance it internationally.

Monark allows institutional and private investors to invest in a digital twin and to monetise it through scenarios open to numerous opportunities. IWA is an alternative to the GAFAs to keep control of the content and retain its value in Europe.

Difficulties/barriers identified:

There are several factors that make the project more complex to implement:

- The lack of standardisation of digital twin formats;
- Owners and institutions are unfamiliar with this new market;
- The absence of a common platform for the CCI (Cultural and Creative Industries) sector;
- The hegemony of the GAFA (Google, Apple, Facebook, Amazon...).

Project partners:

- Software publisher, specialising in 3D audio capture and playback
- Provider of scalable, interoperable, sovereign and integrated Blockchain solutions
- 3D animation production company

- Research institutes (possibly))

Main Gaia-X technologies at stake:

Data platform: sharing, exchange of data between several players in the ICC sector

Security and confidentiality: authentication, identity federation management, data management, data security, consent management and usage control.

Compliance: data portability, rights management and certification**.**

1.2 Creation/exhibition of a platform for the description of works

On the initiative of:

Cap Digital – Competitiveness cluster Paris/Région (1.000 members)

Potential participants:

Consortium being set up

The issue:

This project for a works description platform consists of a marketplace allowing the identification and description of all works produced and distributed in Europe, to facilitate their exhibition and distribution. This marketplace for the circulation of descriptive data on works would meet three needs:

1. Provide access to open data describing works in order to circulate the essential descriptive information conducive to their exploitation/dissemination. This data would mainly be provided by producers and their partners, in order to constitute an alternative to the American IMdB platform in the audiovisual and film sector;
2. Allow the secure circulation of more strategic data, this time involving the interconnection of private servers of industrial members (producers, post-producers, laboratories, rights holders, etc.) aimed mainly at listing, exhibiting and selling digital media (masters, PADs, subtitles, dubbing, etc.) and/or rights. This need could be technically based on the national collaborative project TAMIS which is already underway (see below) and brings together all the major French players (Ficam, CST, laboratories, etc.);
3. Offer software connectors (APIs) associated with a business model adapted to the creation of value by startups.

1.3 Creating a catalogue of digital assets

On the initiative of:

CST

Potential participants:

The issue:

More and more productions today use digital assets, monuments, landscapes, buildings, etc. These assets are usually created on demand for a given production and are archived at the end of the production. However, some of the assets created could be reused in other productions at a much lower cost than creating a new digital asset. It would then be necessary to know that these assets exist and how to access them.

1.4 Creation and exhibition/enhancement of a catalogue of sets

On the initiative of:

CST

Potential participants:

The issue:

Audiovisual productions for television or cinema use many set elements, often manufactured on demand. The reuse of these sets is encouraged within the framework of the eco-responsible Ecoprod approach. However, reuse remains low because no database, to date, lists all the sets manufactured and available with the descriptive elements needed to make an informed choice (material, colour, storage location, condition, availability, etc.).

Expected benefits:

Such a referencing, with standardised metadata, would make it possible to save raw materials and to make the construction of the sets profitable thanks to a reinforced valorisation. A trans-European database would allow local know-how to be enhanced and would increase the base of available sets.

1.5 Deposit of works and secure ownership certificate in the blockchain

On the initiative of:

Potential participants:

France Télévisions

The issue:

Any digitised work becomes easily reproducible. Blockchain and NFT (Non Fungible Token) technologies should make it possible to guarantee the intellectual property of these works to the right holders and to ensure their authenticity. They thus facilitate the fight against piracy, in particular thanks to the traceability of the blockchain.

1.6 Digitisation and sustainability of heritage content

On the initiative of:

Potential participants:

The issue:

Widespread access to telecommunication networks facilitates access to digital works. On the other hand, the audiovisual heritage (mainly cinema) remains largely stored on physical media. Access to these works remains difficult.

A major project to digitise these works would enable them to be promoted and distributed to a wide audience. This digitisation obviously goes hand in hand with the creation of all the metadata necessary for referencing and discoverability of the works.

The aim would be to achieve:

- Digitisation of works
- Cataloguing
- Enrichment of metadata
- Format maintenance
- Reliable storage space
- Associated data to ensure readability of content
- Accessibility of archives

2. Facilitate and document manufacturing

2.1 Portal for sharing data with content producers

On the initiative of:

France Télévisions with the support of Cap Digital

Potential participants:

Producers of published and broadcast works

The issue:

The exploitation of digital works requires a large volume of metadata in order to reference the work as accurately as possible. This referencing then allows for better discoverability, both directly through manual exploration of the platform and, above all, through direct or indirect recommendation and personalisation.

Today, this metadata is complex for broadcasters to collect and obtain. Producers and rights holders do not have systematic processes to create, store and distribute this metadata.

The producer portal is intended to be a place where this metadata can be delivered at once or incrementally to support the production process. It also allows the broadcaster's teams to assess the quality and relevance of the metadata provided in order to correct or complete what needs to be corrected. Once validated, this data is transferred to the France Télévisions data product Programme for use on each of the broadcasting vectors.

2.2 Exchange of production metadata (TAMIS project)

On the initiative of:

CST

Participants:

The current operational participants of the project:

- The CST, which supports the project and its service providers,
- MediaADN
- Lum::Invent (formerly Media-IO)
- Startin'blox

With the support of:

- The ISAN-IA agency
- The Procirep
- Cap Digital - Paris Region competitiveness cluster
- Titra

The issue:

In a context of increasing dematerialised deliveries of audiovisual works, the number and variety of associated media files are multiplying, requiring increasing efforts in referencing and monitoring by the various players in the sector.

ISAN registration has been requested by the CNC in France since 2017, but its actual use remains timid in audiovisual deliverables. This is also the case for other identifiers, such as the EIDR requested by OTT platforms.

The validation of versions is done in an artisanal manner (email, SMS, etc.), generating misunderstandings and confusion in the ordering and delivery process. The registration of variants and events is centralised in the hands of the right holder, making it more difficult for the other actors in the chain to declare new deliverables.

Solution provided:

The TAMIS project aims to promote the exchange of manufacturing metadata, beyond identifiers, in order to optimise exchanges between actors in the manufacturing chain.

The envisaged solution is based on the Solid architecture, which is based on a stack of existing Web technologies. This technology is promoted in particular by Tim Berners Lee to return to a truly decentralised web of data. The Solid technology stack is based on established web technologies, semantic modelling, such as RDF, the use of restful APIs, WebId authentication based on OpenID, to name a few.

For semantic modelling, the project envisages the use of EBU-CCDM, a generalisation of EBUCore proposed by the EBU.

The TAMIS-1 project provides for an evaluation of the options considered and a fine-tuned modelling of the use cases.

The objective is to continue with a TAMIS-2 project, which would be a proof of concept (PoC) involving typical representatives of the sector's value chain, with a laboratory, a licensing entity, a distributor and a rights holder.

The deliverables would be, in addition to detailed documentation, an SDK to facilitate the development of connectors, and an easy-to-implement "plug and play" brick with a graphical interface allowing low-tech actors to do manual input to enter the TAMIS loop.

Expected benefits:

The effective circulation of identifiers is a systemic change. The massive adoption of this distributed solution by the audiovisual industry will optimise the exchange of these identifiers and the metadata that must be associated with them within the audiovisual production chain. Tamis will thus make it possible to:

- Streamline the production process of a work:
  - facilitate collaboration between actors in the chain by reducing the effort required to adapt information systems;
  - generate productivity gains through reliable automation of the various processes and by reducing input errors;
  - optimise the organisation of the schedules of the various stages of production of a work
  - reduce the "Time to Market".
- contribute to the monitoring of the marketing and exploitation of audiovisual works produced;
- promote the circulation of information related to the remuneration of rights holders;
- highlight the skills of the market by improving the visibility of the players;
- contribute to digital sobriety.

Difficulties/barriers identified:

- To attract film producers to a technical project;
- Mobilise investments from companies in the sector to finance a connector;
- Aggregate the different controlled vocabularies used in the industry (EBUCore or EBU-CCDM, or the MDDF ontology proposed by MovieLabs, the Hollywood studios' laboratory).

Main Gaia-X technologies at stake:

The architecture of the Gaia-X federation of services proposes identity and trust services, which are a priori resolved by the Solid architecture chosen, but a more in-depth study may reveal technological and architectural convergences.

On the other hand, the cataloguing service must be able to index and expose entry points to the service.

At this stage, this approach is rather envisaged technically as a Solid connector to the Gaia-X federated services.

2.3 Service to help calculate the carbon footprint of CCIs

On the initiative of:

Cap Digital - Paris/Region competitiveness cluster (1,000 members)

Potential participants:

Consortium being set up.

The issue:

The accurate measurement of the carbon footprint of the publishing, production and distribution workflows will progressively become necessary to get funding. The project aims to bring together the

actors in the value chain who have already calculated the carbon footprint of their activities. Several objectives are defined:

- to create a repository of measurements by type of activity
- to standardise measurement methodologies;
- to implement a European service to help in computing these footprints, in the form of a software platform.

2.4 Fighting fake news

On the initiative of:

France Télévisions

Potential participants:

France Télévisions, INA

The issue:

More and more amateur images, published on social networks in particular, show part of a news story.

For the public, but also for the companies that produce information, the challenge is to distinguish real images from false or out-of-context images in order to separate the real from the false.

Solution provided:

Create a platform that would allow all volunteer factcheckers to tag verified images (still or video) and store their fingerprints.

This platform would be a tool for citizens and/or journalists to query the fingerprint bank to find out whether the image has already been used or not, whether it is falsified or out of context.

A POC has already been carried out at France Télévisions.

Expected benefits:

To enable citizens to have an easy-to-use tool that would prevent them from spreading false information.

Difficulties/barriers identified:

Engage as many fact-checkers as possible to make the database efficient enough. Cost of fingerprints. Maintenance of the platform which would be public and free.

Main Gaia-X technologies at stake:

Data platform: sharing, exchange of data between several players in the CCI sector.

Security and confidentiality: authentication, identity federation management, data management, data security, consent management and usage control.

Compliance: data portability, rights management and certification.

2.5 Artificial intelligence-based services platform for enriching programme metadata

On the initiative of:

France Télévisions, INA

Potential participants:

lum::invent (ex Media-IO)

Audiovisual Public Services

The issue :

Based on an open-source project initiated in 2018: https://media-cloud.ai, the objective is to offer multimedia data processing services (images, moving images, sound, subtitles, interactive data, etc.) that can be used on an existing open-source platform instance made available in the form of a Cloud service, for users who do not have the software engineering skills required to implement it.

Beyond the classical processing of audiovisual files and streams, this platform allows the orchestration of processing based on Artificial Intelligence from explicable models.

This platform could be developed jointly by various players in the sector, particularly public players in France and Europe.

3. Artificial intelligence and self-regulation of advertising content – Invenio project

On the initiative of:

ARPP

The Autorité de Régulation Professionnelle de la Publicité -ARPP- is the French advertising self-regulatory body. Founded in 1935, the organisation now has nearly 700 paying members, brands, agencies, media, platforms, advertising technology solution providers (AdTech), etc., who together adopt ethical rules, undertake to respect them and entrust ARPP with the task of ensuring that these rules are properly applied.

Project ecosystem:

- Digital advertising agencies;
- Website publishers;
- Inventory monetisation platforms;
- Platforms/Social Networks.

The issue:

One of the main challenges facing the ARPP today is the multiplication of advertising channels and creations, particularly since the advent of programmatic advertising, and their format (display, video, audio, text, etc.), making the control of advertising content complex.

Faced with the expectations of civil society and public authorities, professionals must demonstrate their ability to implement an efficient self-regulatory system, modern in the technologies it uses, agile in its ability to adapt to changes in the market, representative in its inclusion of all players in the advertising industry and in its ability to control advertising, in its inclusion of all the players in the value chain, capable of maintaining high ethical standards in a market that is undergoing rapid digital change and has become complex, and of continuing to work towards legal, ethical, fair and respectful advertising for all audiences, in closer consultation with civil society.

Solution:

As advertising is a combination of images, text and/or sound, the technologies of automatic natural language processing, computer vision, convolutional neural networks, etc., based on machine learning, can help detect basic elements in advertising content, such as alcohol (cf. the "Evin" law), the size of characters (legibility of legal notices), men and women (contemporary representation of genders, ages, diversity in advertising), specific textual or audio references (health claims, environmental claims, etc.), etc.

Compliance As a Service solution: To provide digital advertising market players with a catalogue of detectables to help ensure the legal and ethical compliance of advertising content. Indeed, the advent of programmatic advertising and the automation of campaigns, particularly in RTB (real-time bidding system) have made it difficult to identify and manually filter advertising practices that are contrary to legal or ethical rules.

Expected benefits:

The provision of automatic detection models for suspected non-compliance will help :

- Secure the responsibility of the players in the value chain and enable them to demonstrate their ethical approach to all stakeholders (consumers, public authorities, brands, etc.).
- Implement a healthy and trusted digital environment, allowing the market to develop over the long term.

European scale

ARPP and its European counterparts federated by EASA (European Advertising Standards Alliance) exchange information within the framework of a "Data driven self-regulation Committee" working group. The tech partners of the European organisations are invited to attend and source codes are shared. The aim is to make these tools available on a European scale for common infringements.

Difficulties/barriers identified:

One challenge identified lies in the constitution of correctly annotated datasets, of sufficient volume, and corresponding to "non-compliance" use cases (e.g. misleading environmental claims, vehicle on a natural area, excessive food consumption, etc.) in order to have models whose accuracy is deemed sufficient.

USE CASES FOR PUBLISHING, DISTRIBUTION OR BROADCASTING

1. Unified user consent management

On the initiative of:

DAG Conseil

Potential participants:

France Télévisions

The issue:

As mentioned above in this document, revenue from public exposure to advertising is one of the three sources of CCI-related revenue.

These advertising revenues are based on:

- Audience size ;

- The qualification of audiences;
- The ability of the content publisher to address an audience and expose them to a message.

The current paradigm is for each publisher (or each advertising agency associated with a pool of publishers) to build its own audience in order to monetise it. For several years now, each advertising agency has therefore been working to increase the size of its audience. However, faced with the growing demand for targeting expressed by advertisers, these networks are then obliged to segment their audiences, often with the risk of finding themselves individually (agency by agency) with audience pools that are much too small.

Furthermore, the acquisition of advertising consent is currently given by the user for each data controller (often an advertising agency or a publisher). This fact leads to the major difficulty for the consumer to trace/port his consent as the GDPR regulation allows. This difficulty leads to mistrust and attrition from data sharing by consumers.

The project consists in sharing within an ecosystem grouping together all the publishers/advertising agencies, the consent, the uses of content and the centres of interest collected from the user of the content. This ecosystem would allow the user to manage his consent for several publishers/ advertising agencies within the same interface.

This management of consents through this HuB ecosystem (which does not replace the data warehouses held by the various advertising agencies) would enable the players in the sector to develop and propose new service offers to advertisers:

- Cross-agency statistics;
- Cross-agency segmented audiences;
- Anonymised or pseudo-anonymised information on the entire sector.

Solution provided:

The project consists in setting up a governance, storage and analysis layer for a precise list of data on which all the players in the sector agree, while guaranteeing the user's sovereignty over the management of his data and compliance with the GDPR.

This data shared by the players must not only be listed, but must be formulated in a standardised way in order to promote interoperability between the main players, i.e. publishers, advertisers and media agencies.

This list of data and the governance arrangements must be defined during the project's scoping phase and could fall within the scope of the "Data Space" that the WG calls for in this document.

Expected benefits:

- Benefits for publishers and agencies :
  - Better respect of the evolution of user consent over time;
  - Minimisation of legal risks linked to non-compliance with GDPR texts;
  - Better knowledge of user behaviour and interests (granularity to be defined during project scoping);
  - Possibility of monetising the data and/or knowledge "as a service" through the shared ecosystem.
- Benefits for users:
  - Better visibility of the status of their consents;
  - Better experience in updating their consents;

- Better visibility of their predictive interests and ability to update them.
- Benefits for advertisers :
  - Possibility of having visibility of audience volumes across agencies;
  - Better knowledge of user affinities.
- Profit on the regulator side :
  - Better applicability of texts.

Difficulties/barriers identified:

As with many projects involving the sharing of data, it is likely that the main obstacle is the willingness of publishers and/or agencies to share certain data.

Project partners/ecosystem:

- The main publishers and the main digital and TV agencies;
- Media agencies;
- Fixed and mobile operators;
- Sovereign AD exchange and digital programmatic platforms.

Main Gaia-X technologies at stake:

The functionalities proposed by a Cloud technical base + data governance through WEB3 technologies seem to make it possible to build a decentralised ecosystem such as the one required for the proposed project.

The scope of the technical requirements will have to be defined at the end of the project scoping phase.

2. Accelerate remuneration of rights holders

On the initiative of:

SACEM

Potential participants:

France Télévisions

Solution provided:

Media Hub Project.

Sacem protects, represents and defends the interests of over 182,520 members in France and abroad, authors, composers and publishers. In the service of creation, it has the largest repertoire in the world and is committed to musical diversity by supporting numerous projects. Sacem is a private, not-for-profit company managed by its members, whom it supports through 3 essential missions:

- Collecting and distributing royalties
- To promote and support creators
- To defend and protect its members

Sacem collects and distributes royalties for public broadcasting (audiovisual media, concert halls, festivals, internet services, cinemas, shops, etc.) and for the reproduction of the works it represents on media (discs, videos, legal digital files, DVDs, CD-Roms, video games, etc.). Copyright is distributed as closely as possible to the actual use of the works: 80% work by work, 9% by survey and 11% taking into account consumption habits.

The issue and stakeholders

In order to ensure the collection and distribution of royalties for public broadcasting in the audiovisual media, SACEM currently uses statements provided by each media. Each media invests dedicated resources in monitoring its channels and compiling these statements. Despite this, key information about the programmes and associated musical works is sometimes missing to allow for the fair distribution of the rights associated with the authors-composers and publishers of these works. This in turn creates a significant workload for the collecting societies to analyse the statements provided by broadcasters and to achieve the best possible distribution.

The improvement of this process thus concerns all stakeholders in the audiovisual creation ecosystem:

- Programme producers. They have original data on the broadcast works, but communicate them little or poorly to the other players;
- Media. They fulfil their legal obligation individually, potentially calling on service providers to produce and enrich the broadcast data, which results in significant human, technical and financial costs for them;
- Collective Management Organisations (CMOs). They receive data that is often incomplete or inconsistent, which requires a huge amount of manual verification and enrichment work that results in a long processing time, whereas the accuracy and speed of payment of rights is a key issue for the creators' activity;
- Rightsholders, at the end of the chain. They depend for their remuneration, both in terms of accuracy and timeliness, on the quality of the data and the implementation of efficient workflows between the stakeholders mentioned above.

While this is not a new issue, it is becoming more and more pressing for traditional players as the audiovisual sector becomes more digital. Indeed, digital technology has generated a massive increase in the consumption of audiovisual content online, and therefore a colossal increase in the amount of data to be processed by CMOs. At the same time, digital technology provides a tremendous opportunity to facilitate the sharing and analysis of data.

Producers, media and collective management organisations thus face a triple challenge:

- Technological: speed up data processing times, make them more reliable and simplify their reconciliation and enrichment; limit the production and manual processing of data;
- Economic: to reduce the costs of these activities by pooling some of the processing and investment required to build the IT tools. Possibly, in a second phase, create new revenues from the new data thus generated;
- Political: develop a partnership strategy between the players in the sector to defend a common vision of cultural creation and technological sovereignty.

Project goals:

In order to respond to these issues, the "Media Hub" is positioned transversally within the ecosystem. Its mission will be to provide producers, media and CMOs with a common technological platform, allowing for the acceleration and reliability of the process of declaring data concerning works and their broadcasting, by:

- Facilitating the production of broadcasting reports by the media;
- Ensuring the quality and completeness of the data, in particular the data relating to the works, included in these statements;
- Sharing the necessary data between the stakeholders in the rights distribution process;
- Pooling and industrialising technical resources between the partners in the form of a platform

- Ensuring the security and sustainability of the platform by allowing partners to choose the level of confidentiality of the data shared within the Hub;
- Enabling the reproduction of processes that can eventually be used to offer derived commercial services

Expected benefits

The deployment of the "media hub" will bring several benefits:

- Fairer and faster remuneration of rights holders and thus overall support for creation in a responsible and sustainable model;
- Overall cost reduction and productivity gains for all stakeholders;
- Simplification of the processes of collecting and processing the data necessary for the distribution of rights for the CMOs and broadcasters
- Taking into account new uses and innovation in the field of collective management, in order to perpetuate the French leadership in this sector.

Main Gaia-X technologies at stake:

Data platform: sharing, exchange and potentially intelligence of data between several players in the CCI sector, with the potential involvement of public authorities, to defend a French-style model for defending copyright and creation, while taking into account the evolution of uses in the face of the risk of disintermediation.

Security and confidentiality: authentication, management of identity federation, management, securing of data, management of consent and control of use.

Compliance: data portability, rights management, onboarding and certification.

3. Collaborative platform for the creation and distribution of editorial content - CREA project

On the initiative of:

Agencia EFE, BSC SME, Cap Digital, Ceipes, Panodyssey, Pro Progressione, Venetian Cluster, Voxeurop, Worldcrunch.

Panodyssey is a creative and collaborative platform to meet professionals and non-professionals creators to share their fiction or nonfiction content. It gathers writers and readers around common values ethics and the digital world.

The platform is designed to let authors to develop their creativity by exploring new narrative formats. Panodyssey also integrates native features that make it possible for authors to foster their community and manage their earnings based on their objectives - all the while keeping copyrights in mind and ensuring transactions are secured thanks to blockchain technology.

Consortium members:

The consortium (CREA) gathers 9 international partners and 5 countries (Bulgaria, Spain, France, Hungary, Italy) in the media, culture, education and technology sectors. It is created and supported by the European Commission as part of the Creative Europe Lab Innovation programme.

Members of the Creative Room European Alliance consortium:

- AGENCIA EFE: Spanish press agency, European leader in the audiovisual sector (1939)
- BSC SME: Bulgarian non-governmental organization in the cultural sector (1996)

- CAP DIGITAL: European cluster for digital & ecological transformation (2006)
- CEIPES: Italian institution for education and development (2007)
- PANODYSSEY: French start-up in the digital content sector (2018)
- PRO PROGRESSIONE: Multidisciplinary Hungarian organization for the arts and culture (2008)
- VENETIAN CLUSTER: Italian cluster for culture and innovation, based in Venice (2014)
- VOXEUROP: Multilingual online media with an European reach (2014)
- WORLDCRUNCH: International English-language media (2011)

This consortium is open to new associated partners.

The main objective is to create a new European digital space thanks to an innovative approach built around digital content creation and diffusion. The main focus is on the social topics such as fake news and media training, which are priorities for the EU. By 2030, Panodyssey aims to reach 10 million digital creators and 100 million online users, and become a local alternative to the GAFAM.

Constantly evolving, the consortium builds new European economic partnerships to become a reference ecosystem in the audiovisual, cultural and creative sectors.

Project funding:

This project is fully funded for the duration of its activities for 2022 and 2023. The European Commission contributes 60% and the founding partners of CREA finance the remaining 40%.

Issue and ambition:

The objective of the CREA, backed by the entire sector, is to create a European alternative to no longer depend exclusively on American platforms. It is essential to bring out local operators to promote and economically enhance the content of creators in Europe.

New forms of hybrid creations and new creative formats are a major challenge for innovation in the cultural sectors.

Vision and values of the european collective is to offer professional solutions, in line with European standards, and facilitate cross-border cooperation in the sectors of cultural and creative industries.

To combat disinformation, Panodyssey has set up a secure system to guarantee the certification of the identity (ID Check) of content creators. Therefore, fake profiles, fake content and bots do not have access to the platform.

The lack of advertising is a complementary and effective way to prevent the economic mechanisms that finance disinformation. The action of CREA will continue this work of education, in particular with its activity "Media Education Laboratory", by addressing societal themes, such as disinformation, inclusion and digital accessibility, through various initiatives, involving all cultural, audiovisual and creative sectors.

These sectors are particularly sensitive to the protection and enhancement of intellectual property. To control the use of content, without the authorization of the rights holders, the solution is equipped with processes of awareness and management of intellectual property. Indeed, the certification of identity makes it possible to guarantee the authorship of the works and to fulfill the legal obligations in terms of financial transactions in complete transparency.

To industrialize the valuation and protection of intellectual property, the content platform includes blockchain technology, IP expertise and an IP Smart Contract.

Expected benefits:

- **Expected benefits for content creators**: Enhance the creation of European content through the protection of intellectual property and the enhancement of works of the mind. Foster new online experiences around creative and collaborative writing.
- **Benefits for users**: Allow citizens to have a tool that is easy to use and guarantees the information disseminated. Promote accessibility to quality content around literature and knowledge. Deploy a new vision around an algorithm chosen by the user.
- **Benefits on the regulatory side**: Anticipation of the implementation of new European standards such as the Digital Services Act.

Strengths:

- An ambitious project, bringing together quality European players and specialists in their sector.
- A team with technological and/or sectoral expertise
- A project supported and co-financed by the European Commission
- An ability to bring together stakeholders in the cultural and creative sectors across Europe.

USE CASES FOR THE SALE OF CULTURAL GOODS

1. Aggregation of ticketing data

On the initiative of:

Ministry of Culture

Participants:

Arenametrix

The issue :

Ticketing is a strategic element in the cultural field, essential for the economic models of operators, but also for public knowledge, communication and loyalty.

The dematerialisation of ticketing systems raises multiple issues and the sector is currently lacking a player capable of bringing together and making visible the offers made by cultural structures with multiple statuses, operating in very diverse fields of activity and generating a very high number of transactions (more than 50 million transactions for the main French museums in 2019 alone, to which should be added theatres, monuments and other cultural venues with a dedicated ticketing system).

In addition, the emergence of new operators specialising in the electronic sale and distribution of tickets and the development of online purchasing and service platforms have significantly modified the competitive landscape (verticalization of certain groups, concentration and emergence of global players, new intermediaries from online communication or payment, etc.). Finally, the sectors and players show a wide variety of maturity in terms of dematerialisation; the smallest players are currently the least well equipped. The evolution of uses in favour of the large ticket distribution or accommodation and tourism platforms poses a significant risk to the visibility of the cultural diversity of the territories.

For a large majority of producers (CCI, tourism, leisure, etc.), ticket distribution, accommodation and tourism platforms have become essential for customer acquisition and economic development. This model naturally leads to a fragmentation of the value of the offers and has harmful consequences for producers and users:

- The battle for direct access to consumers and their data;

- Large mark-ups on ticket prices;
- Lack of visibility for small players without solutions.

Solution required:

In order not to create a dissymmetry that is particularly damaging to the smallest players, the pooling and aggregation of data from the various ticketing systems would enable the entire sector to better understand the behaviour and aspirations of the public and thus be able to adapt the programming of cultural events and promote encounters between artists and the public. This data could also be shared with other sectors (tourism, transport, etc.) to adapt their own offers.

The data to be reconciled are:

- Sales data;
- Sales locations;
- Available customer information.

Expected benefits:

- Better knowledge of audiences;
- Adaptation of cultural programming;
- Visibility of the offers of small structures;
- Possibility to create combined offers with tourist offers.

Stakeholders:

- Producers ;
- Broadcasters;
- Ticketing intermediaries.

2. Securing the ticketing system

Ensure traceability of ticketing and fight against fraud and the black market.

TRANSVERSAL USE CASES

1. Alliance Culture Data : project for a data exchange platform for the Cultural and Creative Industries sector

On the initiative of:

BnF-Partenariats

BnF-Partenariats, a private law subsidiary of the BnF (Bibliothèque Nationale de France), was created in 2012 to increase the digitisation and visibility of the BnF's collections by building partnerships with numerous digital and CCI players (Publishing, Press, Online Music...).

BnF-Partenariats has undertaken to federate the CCIs around a sectoral data exchange platform project, following the example of other industries.

Initially initiated and formalised by BnF-Partenariats in Alliance Culture Data, this project now needs one or more players capable of ensuring its support and development.

Project ecosystem:

The CCI sector.

The issue:

Data: a pillar of tomorrow's economy

Data is at the heart of the digital transition taking place in many industries. According to the European Commission, the amount of data generated by public bodies, businesses and citizens is expected to increase fivefold between 2018 and 2025. Properly orchestrated and shared, this data is a source of organisational optimisation, innovation and monetisation. Numerous initiatives aimed at facilitating the circulation of data throughout a sector have already been launched, particularly in agriculture, agri-food, health, tourism, etc., and have thus contributed to making data a key driver of their development. At the same time, the emergence of a single market for data in accordance with the European strategy - embodied by the Data Governance Regulation - should provide a more favourable legal and economic environment for the circulation of data.

The sharing and valorisation of data in the Cultural and Creative Industries is a source to be exploited.

Sharing data within the Cultural and Creative Industries (CCI) involves several key issues for the economic development and structuring of the sector:

- Generating value and strengthening the competitiveness of CCI companies;
- Unite the sectors in the face of the giants of the data economy;
- Innovate, stimulate the start-up ecosystem and develop AI.

However, the large amount of data produced over the years by CCIs as their economy becomes more digital is currently under-exploited and under-valued.

Obstacles identified to the development of data exchange within the Cultural and Creative Industries

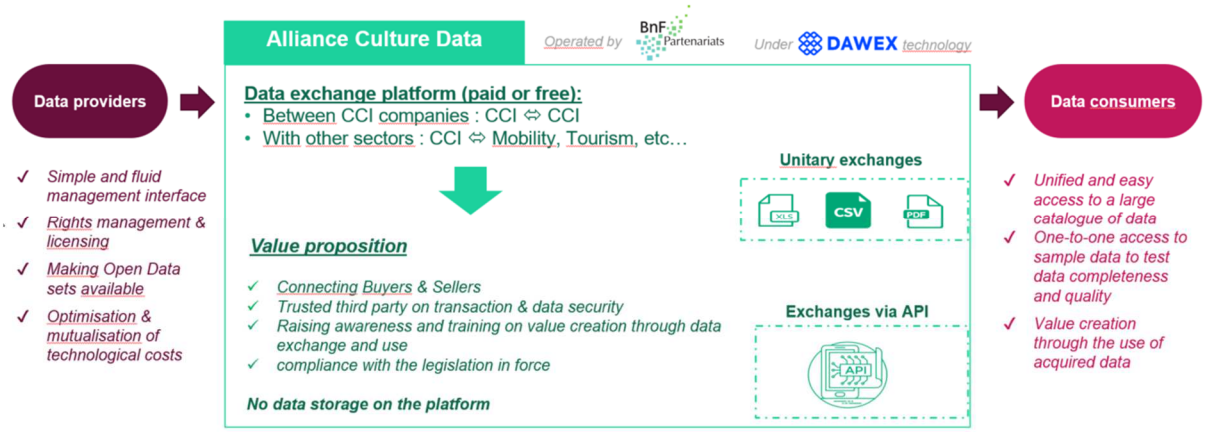
A number of financial, technical, legal and commercial constraints hinder data exchange within the CCI. The inability of each actor to remove these obstacles alone currently favours the siloing of data and does not allow the sector to take full advantage of its value.

Despite a few isolated initiatives, CCIs do not currently have a reliable and shared technical infrastructure that allows them to exchange data in a secure manner while guaranteeing traceability and the proper use of their data.

Alliance Culture Data project:

1- The platform

ACD is defined as a platform allowing its actors to provide or acquire, free of charge or against payment, any type of data. The platform constitutes itself as a "data intermediary" within the meaning of the Data Governance Regulation. The technical solution of the platform would be provided by the French company Dawex whose technology is already used by several industries.



2- Specific services around data

- Upstream (services independent of data exchanges) :
  - Structuring, quality control and cleaning of data
  - Training in data governance
  - Estimation of the value of the data
  - Packaging of the data to meet specific use cases
- Downstream (services following the exchange of data with buyers and offered by specialised companies via the platform):
  - Analysis of quantitative and qualitative data
  - Provision of data skills, data marketing and data strategy services
  - Packaging of data to meet specific needs

The ACD value proposition will thus provide a concrete response to the constraints currently hindering access to and exploitation of cultural data and aims to develop a massive and value-creating use of data, in the image of other industrial sectors.

Strong positive externalities for the sector

The positive externalities linked to the exchange and sharing of data made possible by the platform are numerous, both for the sector and the players in it, but also for sectors outside the CCI, thus making it possible to multiply the links and exchanges between the CCIs and other industries. More than thirty possible use cases and applications facilitated by ACD were identified in interviews conducted by BnF-Partenariats with CCIs.

ACD also appears to be a lobbying tool that can unite CCIs and influence strategy and regulation on data exchange with the web giants.

An economic model built to allow Alliance Culture Data to be self-financing beyond its initial financing

Because of its strategic interest for the sector and the positive externalities it generates, Alliance Culture Data is a project that makes it possible to prepare the conditions for creating value for the future. It is also a large-scale project both in terms of its scope and the degree of innovation it entails.

The economic model of the Alliance Culture Data project is based on two sources of revenue:

- Sales of subscriptions to the platform plus onboarding services
- Sales of ancillary services offered, upstream and downstream of data exchange on the platform

The project requires an estimated €6M in seed funding.

The project is intended to be financially autonomous from public support at the end of the start-up phase. The deployment trajectory planned to date allows ACD to be self-financing after 5 years following its launch and to generate a positive return on investment after 8 years.

3 CROSS-CUTTING NEEDS

DEFINITION OF AN ONTOLOGY OF METADATA RELATED TO CULTURAL AND CREATIVE CONTENT

On the initiative of:

Potential participants:

The issue:

This project aims to map the metadata used to label cultural and creative data, to define the standards (existing or to be created) and to define the technical ecosystem allowing to host this metadata during the life cycle of the proprietary or shared contents.

This metadata for description, governance, use, integrity and enhancement will promote the sharing of data between the different actors as well as their governance. They will serve as a basis for the implementation of smart contracts between the various actors in the sector.

Main deliverable: Dictionary of metadata related to cultural and creative data (including Taxonomy/Ontology) + Guide to data tagging standards.

GOVERNANCE OF A CCI DATA SPACE

Creation of a body for:

- Dissemination
- Training
- Models
- Transmission of standards
- API
- Definition of processes and criteria for acceptance or rejection of participants;
- Choice of technological bricks;
- Monitoring of the legal framework;
- Search for funding;
- Etc.

5 FURTHER WORKS

This position paper is far from being an exhaustive description of all the use cases relevant to CCIs in the framework of GAIA-X. Members of the sector are invited to amend and complete it and to inform the working group's sponsors of any interest in any of the initiatives presented, either as a contributor or as a project leader.

6 ANNEX

The digitisation of sectors

For music, this has been a complete upheaval of the production and distribution sector, resulting in a dizzying fall in turnover during several years of transformation, and then an increase to the point where it is now envisaged to exceed the turnover that existed before the transformation began.

The new channels for distributing works, such as streaming platforms and video sharing services, many of which are musical, have built new economic models resulting in a revival of the sector's growth thanks to technical transformations and the evolution of uses. At the same time, the construction of data repositories for the discovery of works and recordings on these new channels and on traditional channels is accompanying this transformation.

In the field of video games, this dematerialisation is certainly native, but the development of online games has become a determining element of this industry, which has also profoundly diversified its modes of access and its economic models, as well as the richness of its creative proposals thanks to the increase in computing power available to the public.

The audiovisual sector is undergoing a real upheaval with the arrival of global streaming platforms which are spending gigantic sums on exclusive creation on their channels. They are led, from the outset but also in a way that is regulated, to invest locally to ensure proximity to the public in the various markets which they need to grow but also to constantly renew their creations, in a context of extreme competition. The financial markets are essential to their development and European regulation is adapting to this new power that has gained a foothold on the European continent.

At the same time, video-sharing platforms such as YouTube play a role in promoting works, prescribing them and even providing a new economic model for certain types of audiovisual works such as documentaries.

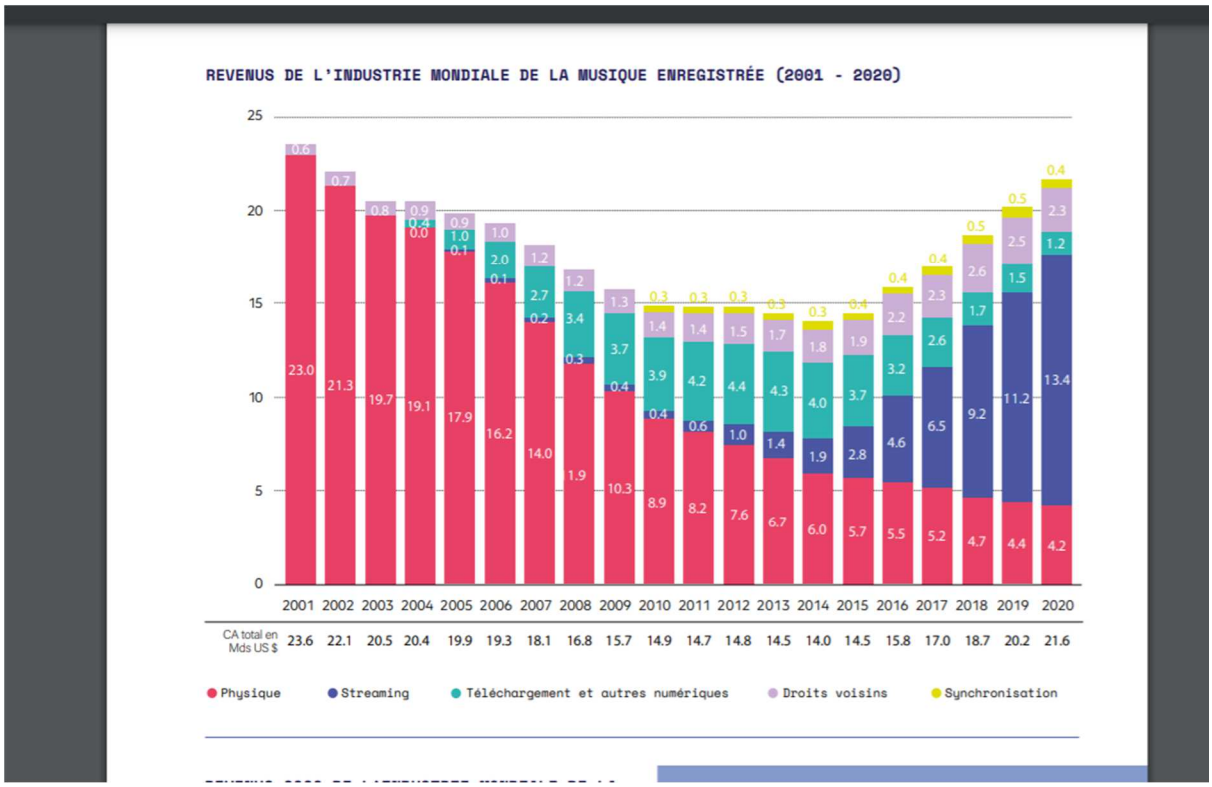
The traditional audiovisual media continue their work of exposing works and financing television channels, but by reinventing their model through the introduction of interactive tools, both to delinearise their programmes (podcasts and catch-up television) and to adapt to the tastes of audiences, particularly with a view to a better advertising proposal (targeted advertising).

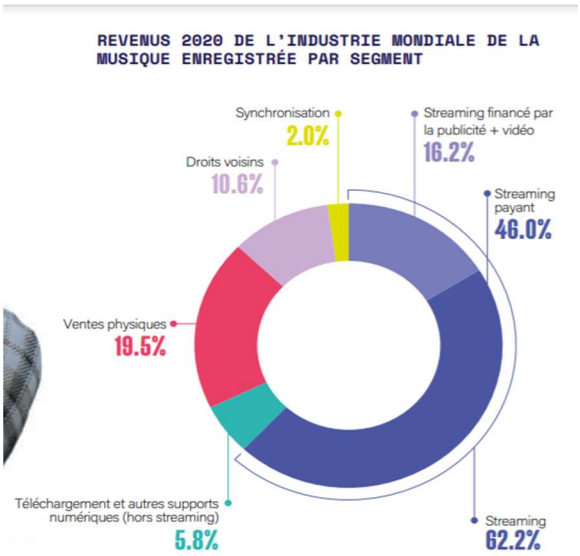
Nevertheless, in all cases, data on works and uses are essential for their discoverability, whether for referencing or for understanding the tastes of the public and the most relevant modes of distribution and promotion according to the works.

Examples:

i) Music

(Source : SNEP – IPFI)





Additional figures :

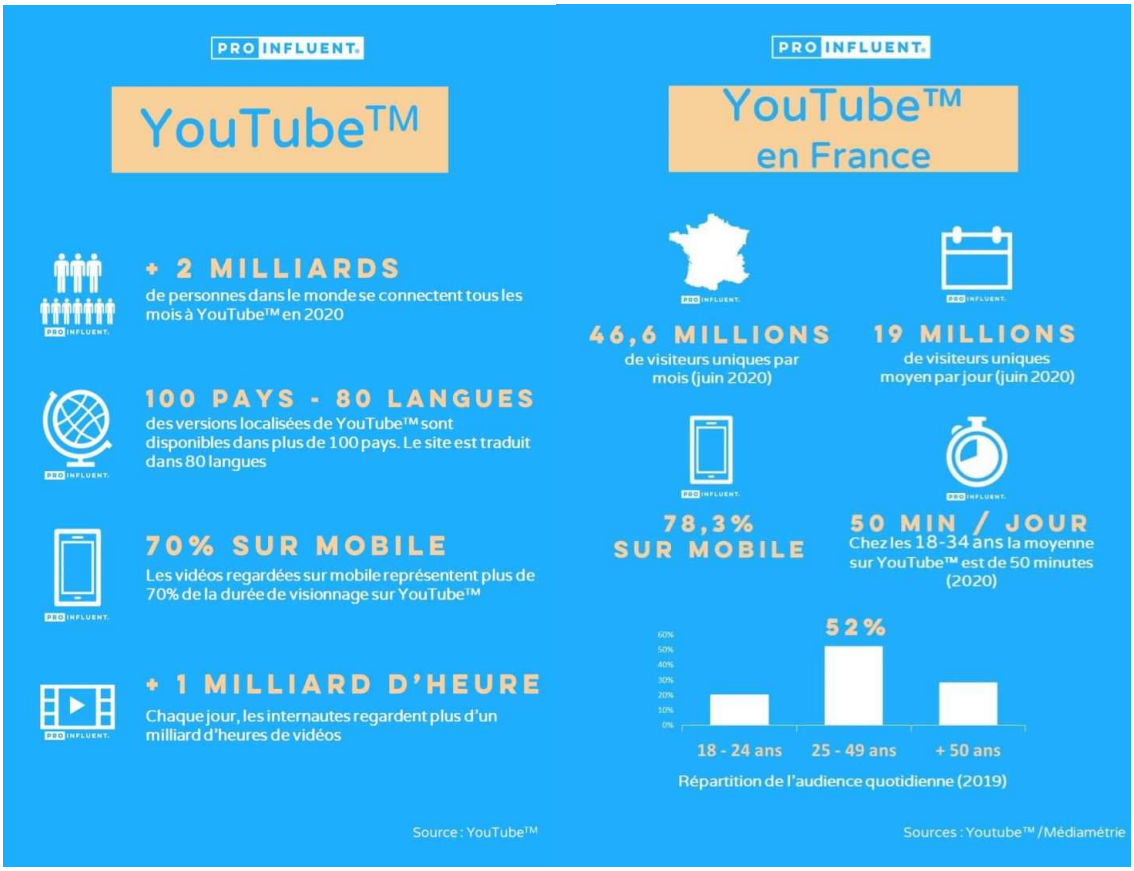
- Income:
  - Since 2016, digital has accounted for more than 50% of recorded music market revenues (Unesco 2020).
  - Digital revenue growth is driving recorded music revenue growth: respectively +19.9% and +7.4% in 2020 (IFPI)

- Users:
    - There will be 443 million users of paid music streaming accounts in 2020, up 30% in one year (IFPI).
    - There will be 345 million monthly active users on Spotify in 2021 (Les Echos).

ii) Audiovisual

Additional figures:

- Netflix has over 200 million subscribers worldwide and could reach 286 million subscribers by 2026 according to Digital TV Research,
- Amazon Prime has more than 200 million subscribers worldwide with at least 175 million watching Amazon Prime Video in 2020,
- Disney+ has nearly 120 million subscribers worldwide and, according to Digital TV Research, could overtake Netflix in subscribers by 2026 by reaching 294 million,
- YouTube 2021 figures (for 2020) :



Additional information on data definitions

Metadata

In its broadest sense, metadata "is data used to define or describe other data, regardless of its medium (paper or electronic)" (wikipedia).
The “Grand dictionnaire terminologique de l'Office québécois de la langue française” is relevant for the definition of digital metadata: "Structured set of data accompanying a work and used in particular to describe its content and format, to ensure its indexing in search engines and databases, and to facilitate the management of related copyright. […] From the perspective of data warehouses, metadata is a key element and is intended for various categories of users. In particular, metadata allows users to know the origin and nature of the data stored in the warehouse, to understand how it is structured, to know how to access and interpret it, to know the different data models involved and the rules for managing this data.
In the context of a report on cultural metadata produced for the French Ministry of Culture and Communications, Pierre Lescure provides some details on cultural metadata, which he describes as "all the information describing any resource, digital or not. Their primary function is to describe the content of the resource, while making it possible to identify, qualify and enrich it; they are not necessarily contained in the document itself. […][Metadata] is the DNA of cultural content".

Usage data

According to the definition provided by Canadian and French reports on the subject, usage data is data on the consumption of a particular cultural content, automatically constituted by the activity of Internet users frequenting this online resource.

There are several types of usage data that can be qualified as "extended usage data", which cover a fairly broad spectrum of information useful to the process of creating and exploiting cultural works, notably:

- raw consumption data: number of views, listenings, hours and days of access, etc;
- data relating to the context of exposure of content: number of times the content has been proposed by a recommendation algorithm, search terms that led to access to the content, editorial reading lists in which the content has been included, etc;
- audience characteristics: socio-demographic data, preferences or tastes, number of friends on social networks, etc;
- data enabling the performance of this content to be compared with other content of the same type offered on the platform: average number of views per content category, etc.

Personal data

(Source : CNIL)

“Personal data” is “any information relating to an identified or identifiable natural person”.

A person can be identified:

- directly (e.g. surname, first name)
- or indirectly (e.g. by an identifier (customer number), a (telephone) number, biometric data, several specific elements of his or her physical, physiological, genetic, psychological, economic, cultural or social identity, but also voice or image).

The identification of a natural person can be carried out:

- **from a single piece of data** (example: social security number, DNA)
- **from the cross-referencing of a set of data** (e.g. a woman living at such and such an address, born on such and such a day, subscribing to such and such a magazine and being active in such and such an association)

Example: a marketing database containing a lot of precise information on the location, age, tastes and purchasing behaviour of consumers, even if their names are not stored, is considered as processing of personal data, as long as it is possible to trace the information back to a specific natural person.

Big data [4]

In the digital context, usage data is accumulating automatically and at high speed; this is known as 'big data'. The term 'megadata' is used to describe a huge amount of structured, semi-structured and unstructured digital data with the potential to contain a wealth of information if skilfully exploited.

[4] Rapport de la mission franco-québecoise sur la découvrabilité des contenus culturels francophones en ligne (2020) - Laticce, Etude ACEI 2018-2020 – « *Etat des lieux sur les métadonnées relatives aux contenus culturels* », Observatoire de la culture et des communications du Québec (2017)

But megadata is so large that it is difficult to store, manage, compile and analyse with traditional database processing and analysis tools.

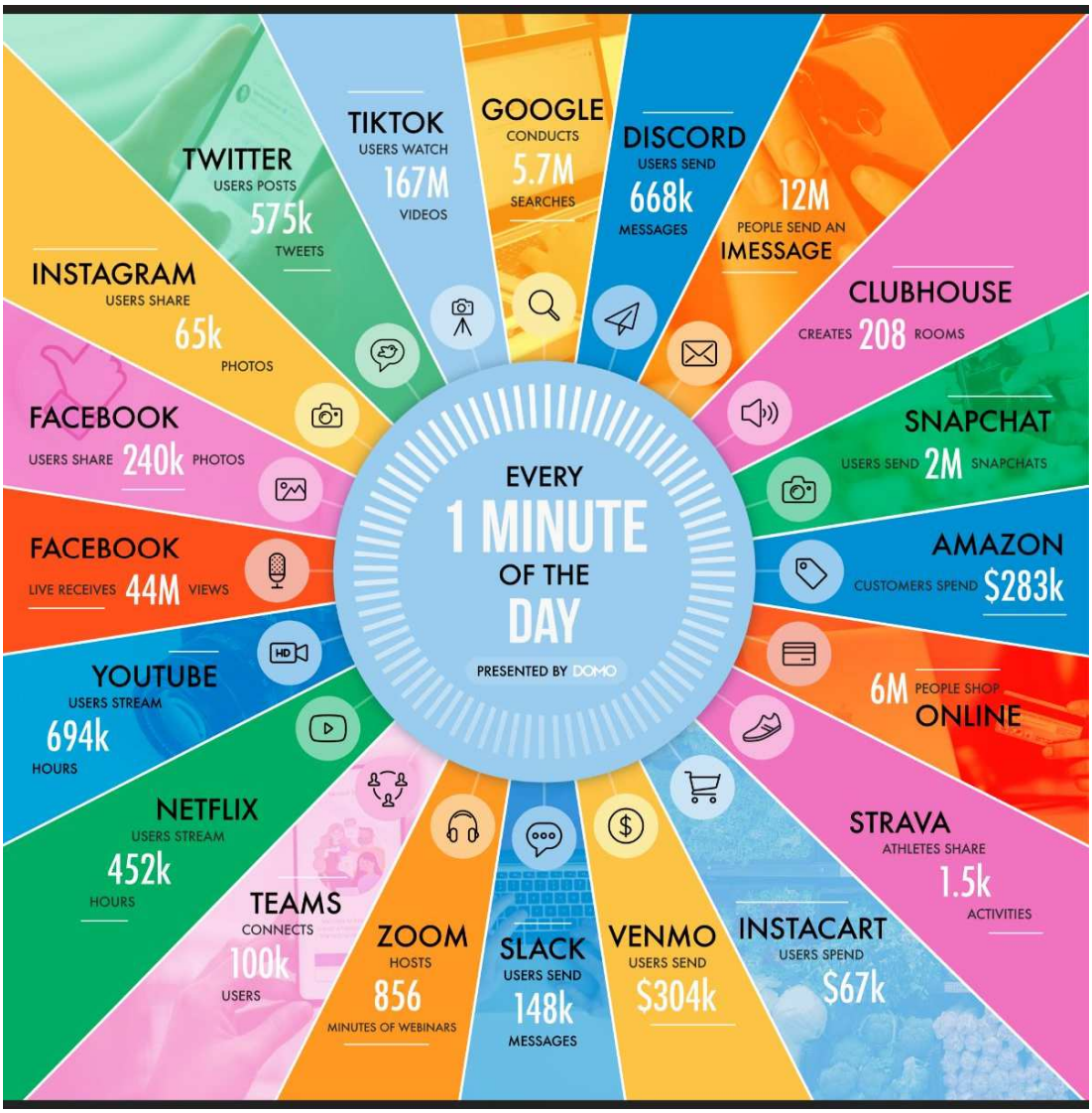
The concept of "megadata" also refers to the ability to link multiple databases and, through the use of advanced analytics, identify patterns of information that would otherwise remain invisible.

Megadata has three key characteristics, commonly referred to as 3Vs:

- Volume: Batches of megadata are huge, hence the prefix 'mega'. Terms often used in reference to the volume of megadata are "petabytes" and "exabytes" of data.
- Velocity: Megadata is characterised by the high speed at which it is built up (often automatically, through 'clicks' or web transactions) and by the fact that it is constantly changing, reflecting a real-time situation.
- Variety: These are sets composed of structured data (usually text organised in traditional relational databases) and unstructured data (photos, videos, textual data).

The use of megadata to measure the consumption of digital cultural products or content requires different technical infrastructure, computing capabilities and data science expertise than those used for the production of traditional statistics. As a result, it represents a challenge for government agencies and national or international bodies to produce public statistics on cultural consumption that are useful to a variety of users and that rely on an objective and neutral reading of web activity. Another challenge for the production of public statistics on digital culture consumption comes from the fact that usage data are often owned by private companies that are not open to sharing them without cost.

Amount of data generated on the Internet every minute (Source : www.domo.com)


EVERY
1 MINUTE
OF THE
DAY
PRESENTED BY DOMO
TWITTER
USERS POSTS
575k
TWEETS
TIKTOK
USERS WATCH
167M
VIDEOS
GOOGLE
CONDUCTS
5.7M
SEARCHES
DISCORD
USERS SEND
668k
MESSAGES
12M
PEOPLE SEND AN
IMESSAGE
INSTAGRAM
USERS SHARE
65k
PHOTOS
CLUBHOUSE
CREATES 208 ROOMS
FACEBOOK
USERS SHARE 240k PHOTOS
SNAPCHAT
USERS SEND 2M SNAPCHATS
FACEBOOK
LIVE RECEIVES 44M VIEWS
AMAZON
CUSTOMERS SPEND $283k
YOUTUBE
USERS STREAM
694k
HOURS
6M PEOPLE SHOP
ONLINE
NETFLIX
USERS STREAM
452k
HOURS
STRAVA
ATHLETES SHARE
1.5k
ACTIVITIES
TEAMS
CONNECTS
100k
USERS
ZOOM
HOSTS
856
MINUTES OF WEBINARS
SLACK
USERS SEND
148k
MESSAGES
VENMO
USERS SEND
$304k
INSTACART
USERS SPEND
$67k