

## Better data on the cultural economy

Scoping Study: Availability, coverage, gaps  
and limitations of data representing economic  
activity in the cultural sector - and how to improve it

By MyCake, in association with The Audience Agency  
For DCMS, November 2021

 the audience agency

## MyCake

MyCake specialises in financial benchmarking, especially for third sector organisations. Our core skill is finding and analysing organisational and financial data from organisations across a sector, and interpreting it. The point of doing so is to create actionable insights.

We go through data meticulously, line by line, learning from what others have done. From that, we create powerful insights, relevant insights and flexible products and services. We marry up data sourced via the APIs of large national datasets with manually acquired highly nuanced, detailed and often heterogeneous data on individual organisations.

The benchmarks we create help funders and policy experts to make decisions, and help organisations to be more successful. To be resilient. Even to innovate.

## The Audience Agency Group

The Audience Agency Group is a mission-led social enterprise based in the UK and operating internationally. We believe that taking part in cultural and creative activities supports people's wellbeing and resilience. Our mission is to help drive greater reach and more equitable access to such activities, to grow the scale of relevant activities - both funded and commercial - and to build more creative places. We do this through insight, people-centred approaches, innovation and collaboration.

We gather evidence - joining up our own Audience Finder cultural sector intelligence and audience insight platform, datasets from partners such as MyCake and bespoke research - and use this to advise on strategy and policy in the arts, culture and heritage and wider creative sectors.

Our Createch Innovation Unit leads and supports Research & Development in the two focus areas of cultural analytics and the user-centred design of new creative content and experiences.

## Contents

Glossary	4
Purpose of this report	5
Executive Summary	6
Headline recommendations	9
1. Introduction	12
2. How financial value is created in the cultural sector	16
2.1. What drives the generation of data on the financial value of culture?	17
2.2. Which subsectors are collecting data on value?	17
3. What data do we want to help us demonstrate value?	20
4. Fulfilling the data specification	23
4.1. Financial data	23
4.2. Significant gaps in the financial data	26
4.3. Detailed financial data: recommendations	28
5. The cultural sector data platform: phase one	29
5.1. The cultural sector data platform: core data	29
5.2. Data querying & extraction	32
6. The cultural sector data platform: phase two	33
6.1. The cultural sector data platform: additional data	33
6.2. Phase two: benefits and limitations	35
7. Cultural sector data: challenges and issues	36
7.1. Access	36
7.2. Completeness, accuracy & timeliness	36
7.3. Unique identifiers	37
7.4. Data quality & definitions	37
7.5. Standardisation & regulation	38
7.6. Key metrics	38
8. Conclusion	40
9. Detailed recommendations	42
Appendix 1: Methodology	48
Appendix 2: Mapping data flows into government	49
Appendix 3: Sources of value creation	53
Appendix 4: Linking datasets using URNs	56

## Glossary

ACE	Arts Council England
ALBs	Arm's Length Bodies
API	Application Programming Interface
BFI	British Film Institute
CCEW	Charity Commission for England and Wales
CH	Companies House
CRF	Culture Recovery Fund
CSV	Comma Separated Values
DCMS	Department for Culture, Media and Sport
DLUHC	Department for Levelling Up, Housing and Communities
FCA	Financial Conduct Authority
GVA	Gross Value Added
HESA	Higher Education Statistics Agency
HMRC	Her Majesty's Revenue and Customs
ICNPO	International Classification of Non-Profit Organisations
IDBR	Inter-Departmental Business Register
iXBRL	Inline eXtensible Business Reporting Language
MTD	Making Tax Digital
LUNA	Liquid Unrestricted Net Assets
NLHF	National Lottery Heritage Fund
P&L	Profit and Loss
PAYE	Pay As You Earn
SIC	Standard Industrial Classification
SORP	Statement of Recommended Practice
TAA	The Audience Agency
URN	Unique Reference Number
UTR	Unique Tax Reference
VAT	Value Added Tax
VCSE	Voluntary, Community and Social Enterprise (sector/organisations)

## Purpose of this report

- 0.0.1 Through a competitive process, DCMS commissioned MyCake in association with The Audience Agency (TAA) to review existing data from a variety of sources representing cultural activities and containing financial information. This report is the result.
- 0.0.2 We were asked to consider the extent to which such data could be joined up, the challenges of doing so and the limitations or gaps that would continue to exist afterwards. The purpose of the resulting joined up data would be to support economic modelling of the cultural sector to inform policy development.
- 0.0.3 Further, we were asked to recommend:
  - 1. how best to join up existing data – the data sources and data to use, the necessary infrastructure, data standardisation and changes to data collection processes
  - 2. future research, methodologies and approaches to address remaining limitations or gaps in the resulting joined up data.
- 0.0.4 Finally, we were to incorporate feedback from academic researchers with knowledge of these data and data sources.
- 0.0.5 The focus of this project was in meeting the evidence needs of DCMS and central government policymaking. However, it became clear through consultation with data sources and other stakeholders that many others have similar requirements.
- 0.0.6 As with other cross-sectoral initiatives, a partnership approach is required to succeed. Therefore, we also set out recommendations for the governance of such an initiative as a key precondition for success. Whilst DCMS is not formally responsible for leading data standards across the cultural sector it will be more likely to meet its own and central government's needs by playing such a role.
- 0.0.7 This report is therefore also relevant to those other potential beneficiaries and organisations whose contribution will be required – including local authorities and Arms Length Bodies, sector representative bodies and other organisations funding, supporting or developing policy for the cultural sector in England.
- 0.0.8 We hope this report is a useful prompt to progress discussions on how to create a substantially better evidence base for England's cultural economy.

## Executive Summary

- 0.0.9 **The cultural sector needs better data to understand its value and contribution to the economy and society.** Covid-19 exposed the need for organisations that support the cultural sector - arts, museums and heritage organisations - to have access to better evidence about the cultural economy. This includes central government, arms-length bodies, sector bodies and local government, all of whom have specific needs for data on the cultural sector. While national-level estimates for the sector’s economic and financial performance are available, more granular data is needed for most decision-making.
- 0.0.10 To assess the financial value generated by the cultural sector’s multiple supply chains, DCMS and other interested parties need data with more comprehensive coverage, greater detail and improved accuracy. Better data will enable a more nuanced segmentation of the sector’s diverse activities. This will facilitate a more accurate representation of the cultural sector’s value to the economy. In the context of this report we are focussing on the market value as measured by key financial data such as the income and expenditure of organisations operating within the sector. We are not addressing questions of the measurement of total economic value of the sector as this is covered by the work on Culture & Heritage Capital.<sup>1</sup>
- 0.0.11 **Several problems characterise cultural sector datasets.** Our audit concluded that gaps in coverage and quality-related problems characterise currently available datasets. They are neither current nor updated in real-time. Standard systems for the classification of organisations, such as the Standard Industrial Classification (SIC), are not used. Significant data gaps prohibit our ability to account for the value of the cultural sector accurately. However, **there are opportunities to collect better data.** Our audit concluded that there are five opportunity areas for collecting better data:
1. exploiting *administrative data* - this represents the best opportunity for collecting data to report on the cultural sector’s impact;
  2. *standardisation* - systematising how organisations collect, segment and analyse data is most likely to generate new insight;
  3. *linking* - connecting different datasets is most likely to add value to the data we collect;
  4. *annual, organisation-level* data is the most valuable data building-block; and
  5. *modernising* how we collect data, such as moving away from submission using PDFs, so generating new insights is sustainable.
- 0.0.12 **There is a significant volume of administrative data about the cultural sector.** Administrative data is information created when people or organisations interact with services or other organisations.<sup>2</sup> It includes information collected at the point of a transaction, whether a subsidy payment such as a grant or tax relief, or a ticket purchase.

<sup>1</sup> See <https://www.gov.uk/guidance/culture-and-heritage-capital-portal>

<sup>2</sup> “Administrative data can be generally described as data which are derived from the operation of administrative systems (e.g. data collected by government agencies for the purposes of registration, transaction and record keeping)” See also ADRUK - What is administrative data?

- 0.0.13 This scoping study identifies data sources that, if linked, would cover most registered organisations in the cultural sector. This will ensure that sector intelligence goes beyond organisations receiving government grants or other subsidies. It also identifies the mechanisms by which ‘administrative data’ is generated, compiled, and made available - and how other data sources can fill gaps and add value. This is based upon a review of statutory and regulatory data sources in England, such as Companies House, and data held by intermediaries, including funders and membership bodies.
- 0.0.14 **National level administrative data is often preferable to survey data or research.** Our review highlights the potential for synthesising administrative data collected routinely from activities such as ticket purchases. Data from annual reports and accounts are equally important. It concludes that national administrative data should be prioritised as the core element of a new platform to strengthen the cultural sector data infrastructure.
- 0.0.15 **A standardised, core dataset will transform our understanding of the cultural sector’s scope and value.** This can be addressed through a phased development programme of work to strengthen the sector’s data infrastructure. This should start with the relatively modest goal of collecting primary descriptive data about the cultural sector: subsector of operation;
1. location;
  2. total income and expenditure;
  3. total asset value;
  4. expenditure on staff costs;
  5. profit/contribution to reserves; and
  6. liquid unrestricted net assets.
- 0.0.16 **We propose building a cultural sector data platform to collate and host data from a range of sources.** Many agencies hold the descriptive data outlined above. Our audit found that the cultural sector is diverse and complex, comprising multiple supply chains, and therefore this data is not held in one place. To produce regular (in the first instance, annual), accurate statistics or analysis on the cultural sector and its subsectors, a data platform to combine and regularly update data firstly from statutory and regulatory sources, and later funders and other sector bodies offers the best solution. This is a complex undertaking, so our proposals for a phased approach start with a minimum viable product based on existing available data from statutory and regulatory sources with a focus on Charity Commission for England & Wales, Companies House and Interdepartmental Business Register (IDBR). We envisage that this iterative process should learn from the planned changes to the IDBR.
- 0.0.17 **Detailed financial information is harder to collect, but necessary.** The proposed data platform will be an essential building block in the sector’s data infrastructure, but it will not address all requirements. For detailed financial modelling of the cultural sector and its subsectors, in-depth financial data from constituent organisations is required. Our audit found that this still requires extensive manual collection of data, with problems including access, timeliness and accuracy. Data is too often stranded in non-machine-readable formats such as PDFs. Longer-term, we concluded that better solutions are needed, such as the ability to machine-read PDFs - until we reach a point where data can be collected in machine-readable formats at the point of submission.

- 0.0.18 **Linking datasets will create new data and add value – but this requires common standards.** Data about cultural sector organisations appears in many datasets. Linking these together will create new data and insights, but this will require common standards and identifiers such as company numbers. However, there is no Unique Reference Number (URN) system that works across all datasets or has universal coverage across the cultural sector. We found that solutions are available, such as the ‘master data’<sup>3</sup> approach adopted by 360Giving. Linking datasets represents a substantial opportunity to improve our knowledge of the sector.
- 0.0.19 **Segmentation, segmentation, segmentation.** Our audit found that segmenting the cultural sector into functional clusters or policy areas remains a significant challenge. Systems such as the Standard Industrial Classification remain too broad for helpful analysis of the cultural sector. We concluded that data sources such as membership directories and registers of cultural provision could provide the basis for addressing this, given some standardisation and alignment. Bodies holding data on the sector favour standardisation and are willing to participate in efforts to align activities. New approaches such as <https://charityclassification.org.uk> (based on keyword search of company descriptions plus existing classification schemes) demonstrate what is possible, but this remains challenging to implement.
- 0.0.20 **Governance and leadership.** We concluded that longer-term development work would likely require government and cultural sector bodies to collaborate around standards, data sharing and data collection. Whilst governance was not the focus of this audit, we concluded that the Lottery Forum may provide a useful basis for the cultural sector and government to develop the sector’s data infrastructure and would be worth exploring further.
- 0.0.21 **Longer-term development.** Our audit found that strengthening the cultural sector’s data infrastructure would substantially benefit government and the sector. Developing a core data platform is an essential building block in that infrastructure. Still, we identified several longer-term challenges around accessibility, quality and accuracy, standardisation, and analysis and reporting. As such, we have made recommendations covering the short, medium and long term.

<sup>3</sup> The Government Data Quality Framework identifies as a key risk the “failure to specify use of master or reference data”. Gartner defines master data as “the consistent and uniform set of identifiers and extended attributes that describes the core entities... including customers... citizens, suppliers, sites, hierarchies and chart of accounts.”



## Headline recommendations

0.0.22 We propose eleven main recommendations to strengthen cultural sector data and the infrastructure needed to deliver this. These are set out in more detail in the conclusion. Our short-term, practical recommendations centre on the development of a cultural sector data platform. Our medium and long-term recommendations to modernise how data is collected and used are more aspirational. They will require engagement and agreement beyond the cultural sector. Nevertheless, we think that these are important points worth making.

### Short term recommendations

0.0.23 **Build a cultural sector data platform.** Data about the cultural sector is fragmented, often inaccessible and incomplete, making it difficult to assess its value, particularly for subsectors. Our first main recommendation is to address these issues by building a cultural sector data platform which takes existing data feeds as its start point:

1. Using a phased approach, build a platform comprising a secure, linked data repository and associated reporting and analysis layer;
2. Use the platform to aggregate and link data from selected data sources, starting with core statutory and regulatory data from national bodies;
3. In a second phase, add additional data from sources including funders, membership organisations and ALBs; and
4. Use the completed national data structure to develop a classification system for subsectors and use these as sample frames for collecting detailed financial data.

0.0.24 **Build shared leadership around cultural sector data.** To make data-driven decision making the norm across the cultural sector, government and sector bodies need to help ‘build the field’ – growing and convening a community of skilled researchers and analysts working with cultural sector data and evidence. The proposed remit of the leadership working group would be to improve the quality and availability of cultural sector data. The working group would lead the standardisation and alignment of data formats and drive improvements across the sector in data collection, analysis and reporting. We suggest that this working group is a ‘coalition of the willing’: interested parties who want to strengthen the field, rather than a working group with an official role. We recommend:

1. Expanding the existing ‘Lottery Forum’ role to create a cultural sector data infrastructure working group;
2. Using this working group to help to address the data sources, specifications and standards underpinning the platform proposed in recommendation one; and
3. Coordinating action across government and interested parties who want to strengthen cultural sector data.

0.0.25 **Develop common standards and indicators.** Building the field also involves setting norms and common language and standards around data. We recommend this work starts by:

1. Setting standards for data collection (such as the 360Giving standard for funders) and analysis, including segmentation of the cultural sector into subsectors; and
2. Defining a set of core indicators for reporting on the cultural sector.

0.0.26 **Linking datasets.** To maximise the value of existing administrative data on the cultural sector and minimise the compliance burden from collecting new data, we propose that research and development work is undertaken to:

1. Develop the technical approach required to link together datasets, including the development of a Unique Reference Number (URN) system for the cultural sector;
2. Agree on an approach to encourage the use and integration of widely used URNs, including company and charity numbers, into cultural sector datasets.

### **Medium-term recommendations**

0.0.27 As noted above, our medium and long-term recommendations are more aspirational. Further details are contained within the conclusion.

0.0.28 **Develop standard approaches to segmenting/classifying cultural sector entities.** The ability to segment the sector into standardised, widely recognised subsectors will enable consistent reporting and analysis. We recommend:

1. Development of a single, standard model to segment the cultural sector into subsectors
2. Working with regulators to improve how they classify data, building upon systems to categorise organisations such as the Standard Industrial Classification

0.0.29 **Collect financial data from subsector surveys.** Detailed financial data is critical to demonstrating value, but this is unlikely to come from existing administrative data. To address this gap, we recommend:

1. The development of representative subsector samples and collection of financial data for these entities, typically using surveys
2. Longer-term, working with funders and sector bodies to develop a standardised, short profit and loss template for collecting financial data

0.0.30 **Improve regulatory data.** Administrative returns submitted to regulators such as the Charity Commission for England and Wales generally provide a consistent and reliable data source. However, the accuracy and consistency of this data requires improvement. We recommend:

1. Supporting regulators to enhance the quality of financial data obtained via mechanisms such as annual returns, including better data validation.

0.0.31 **Modernise data acquisition/analysis to support sustainable cultural data.** To reduce the costs of data collection and modernise processes, we recommend:

1. Increasing the use of machine-readable data. We recommend undertaking further research to explore the barriers and opportunities for machine-readable data/iXBRL
2. Making access to support organisations' longitudinal data easier. Recommended development work includes building partnerships with support organisations and sector bodies that hold data on the cultural sector to facilitate sharing and understanding of how data is collected and used.

### Long-term recommendations

- 0.0.32 **Encourage common reporting formats.** To make reporting less burdensome and improve quality, we recommend the standardisation of application and reporting processes. This could be achieved by:
1. Encouraging ALBs to standardise application and reporting processes
  2. Piloting the production of fixed-format accounts
- 0.0.33 **Generate more timely data for shorter periods.** Policymakers require more up-to-date data than is currently available.<sup>4</sup> To generate more up-to-date data, we recommend research and development to experiment whether:
1. Shorter, more frequent reporting periods are possible
  2. It is possible to reduce the lag between reporting and publication of regulatory data, such as through innovation around data structures and feeds
- 0.0.34 **Collect data on freelancers and the labour market data.** Administrative data on the cultural sector workforce is inconsistently reported. Given the prevalence of freelancers, a better understanding of their role is critical to understanding the cultural sector as a whole. We recommend:
1. Building our sense of the scale and role of the self-employed by collecting better data on freelancers, potentially including new sources of data
  2. Collect more accurate, consistent data on staff numbers in the cultural sector

---

<sup>4</sup> The demand for more timely business demographic data is not limited to the cultural sector. The Office for Statistics Regulation has highlighted this is an area requiring more development work.

## 1. Introduction

- 1.0.1 The cultural sector is increasingly recognised as a significant contributor of value to the UK economy. DCMS estimates that in 2021 the cultural sector employed 696,000 people, generating £29.6 bn in Gross Value Added (GVA).<sup>5</sup>
- 1.0.2 The cultural sector is integral to the social and economic life of the UK, so it is no surprise, therefore, that it is increasingly the focus of government policymaking. As the sector's presence and role in all parts of the UK rises, it is more important than ever that government creates an environment that supports innovation, encourages growth where appropriate, and financial and operational resilience where needed. Creating and maintaining such an environment requires high-quality evidence and data about the cultural sector.
- 1.0.3 Understanding the sector's scope and contribution, and role and resilience, is the basis for effective policy development. More granular analysis of cultural activity will increase understanding of resilience, productivity and impact. In turn, this will enable more effective targeting and evaluation of policy interventions.
- 1.0.4 It is just as important for the people and organisations working in the sector - data and insight is a significant driver of value creation in the modern economy. As the recently launched National Data Strategy highlights, better data will open up new markets and drive demand for a more highly skilled workforce.<sup>6</sup> Collecting better - and sometimes more - data is not a compliance burden on the arts and cultural sector. It is an investment in its future. This report - an audit of existing data and data infrastructure - explores where that investment is needed to strengthen our understanding of where and how the sector makes a difference. We think that this work will also increase the capability of the sector to contribute to our cultural, social and economic lives.

### **The cultural sector needs better data to achieve a bigger impact**

- 1.0.5 The starting point for this report is the reality identified in the National Data Strategy - that too much data is locked away, unused or fragmented; that our data infrastructure needs to reflect the importance of data in the modern economy and society; and that better use of data will drive improvement in our sector. For the cultural sector, this reality has been described by the Centre for Cultural Value as a 'data deficit', who argue:

*There's a data deficit in the cultural sector which can make it challenging to evaluate the differences culture makes to people and places. Data is often poor quality and there is no standardised approach to collecting and managing it. As a result, policy and funding decisions aren't always informed by evidence and can feel disconnected from the people and places they have an impact on.<sup>7</sup>*

<sup>5</sup> <https://www.gov.uk/government/statistics/dcms-sectors-economic-estimates-monthly-gva-to-december-2021>

<sup>6</sup> <https://www.gov.uk/government/publications/uk-national-data-strategy/national-data-strategy>

<sup>7</sup> <https://www.culturalvalue.org.uk/our-work/making-data-work/>

- 1.0.6 On this reading, addressing the cultural sector’s data deficit is critical – not simply to enable better decision making in government or to support better management and strategy in cultural organisations and funders – but because the sector’s legitimacy and public support depends on being able to give an account of how it changes lives. If we can agree on the centrality of data to the success of the modern cultural sector, the question then turns to what data is needed, where data might be available, and how we can make best use of it. These questions are the focus of this report.

### **Improving data for the cultural sector**

- 1.0.7 For those working in and with the cultural sector, easier access to better data is increasingly central to how value is created, delivered and communicated. We believe that this mindset requires new tools and approaches that enable timely and accurate analysis to inform decision-making about the cultural sector. This will also require access to a broader spectrum of more detailed data than is currently possible. And where data is available, it needs to yield more value if we are to justify its collection and generate better insight from its analysis. This implies greater standardisation of measures and insights, linking different, related datasets, and better use of existing data routinely collected for administrative purposes.

### **What data about the cultural sector can we collect – and what will this tell us?**

- 1.0.8 The table below sets out the primary data types we suggest are accessible and sufficiently robust to justify collecting to assess the high-level economic impact of the cultural sector. The data outlined here will provide an overview of the cultural sector’s overall shape, scale, and total financial value.
- 1.0.9 To enable detailed economic analysis of the cultural sector, more detailed data is required. The most readily accessible data from sources such as Companies House lacks sufficient detail on the finances of cultural sector organisations. In the longer term, it will be possible to generate data for detailed financial modelling by manually processing data from annual reports and accounts, drawn from representative samples for each subsector.

Data	What information can this data be used to provide?
Annual income & expenditure (per organisation/site of cultural supply)	<p>Mapping of cultural providers and provision – with low granularity and gaps where provision is not via separate organisations identified to be ‘cultural’</p> <p>Partial indicators of scale and value of cultural activities</p> <p>Broad financial income and expenditure data</p> <p><b>Achieves: ability to determine total income for the sector and relevant subsectors, geographies and turnover bands.</b></p>
Balance sheet data (per organisation)	<p>Working capital and reserves levels as indicators of financial resilience and health of cultural organisations</p> <p>Partial value of some types of assets held by cultural organisations or on which cultural supply depends</p> <p><b>Achieves: better understanding of levels of LUNA across the sector</b></p>
Directories of organisations and cultural assets	<p>Identification of organisations or assets not represented in financial reporting or not identified there as cultural</p> <p>Greater granularity of different types of cultural provision or activity</p> <p><b>Achieves: broadens perspective on the sector and helps work out what is not included in financial accounts data</b></p>
Categorisation of organisations/ assets/ provision using common subsector typologies	<p>Financial data can itself be divided into subsectors</p> <p>Value of economic impact can be better modelled so that fundamentally different parts of the cultural economy are not averaged together</p> <p><b>Achieves: more nuanced economic models which are specific to key cultural subsectors/stages of growth</b></p>
Geographical location data of organisations/ assets/provision	<p>Mapping of the economic value of culture against indices of deprivation, urban/rural/coastal etc</p> <p>Comparisons of the economic value of culture between localities - against attributes of cultural provision in that locality</p> <p>Investment case and choice of policy levers for place-based interventions</p> <p><b>Achieves: ability to benchmark both cultural provision and economic viability of cultural organisations in relation to place</b></p>
Other sources, including annual reports and accounts	<p>Directories enable sampling of organisations - and detailed financial data can then be acquired from annual reports, though this requires manual extraction. Other sources may include commercial databases such as Fame.</p> <p><b>Achieves: detailed financial data where detailed modelling is required</b></p>

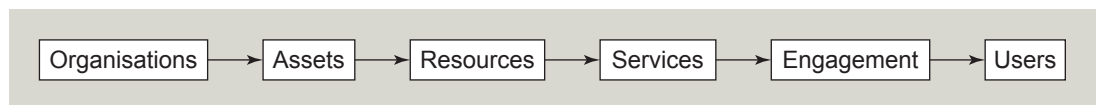
## About this report

- 1.0.10 This is a wide-ranging, long-term agenda. The report aims to map out a way forward, setting out recommendations for strengthening the data infrastructure for the cultural sector, including practical suggestions for how interested parties might collaborate and build the requisite structures, standards and tools. It is also clear that any attempts to strengthen the infrastructure should be collaborative, including government, regulators, funders, and cultural organisations.
- 1.0.11 **Section two** outlines our thinking about how and where the cultural sector creates value and why capturing evidence of this value is useful. This provides a framework for the remainder of this report. It then sets out what types of data feasibly can be collected to demonstrate the role and value of the cultural sector.
- 1.0.12 In **section three**, we set out the specification for the data required. In section four, we identify and audit the primary data sources about the cultural sector in England to establish how we can meet this specification. The priority is to locate datasets generated in the ordinary course of supply chain operations or ‘administrative data’. This is an important basis for future work to bring together and link data about the sector.
- 1.0.13 The second half of this report turns to our proposals for strengthening the data infrastructure of the cultural sector. In section five, we set out proposals for the first phase of building a data platform for the cultural sector. The platform aims to bring together a core data set on the cultural sector that, over the medium term, contains granular information on individual organisations, sites and venues in which cultural activity occurs.
- 1.0.14 **Section six** outlines more comprehensive development activities that we believe are necessary for the second phase of development to turn cultural sector data into intelligence and insight on impact.
- 1.0.15 Strengthening the data infrastructure for the cultural sector will, we have concluded, require more than increasing our capacity to collect and pool data. This includes work on linking datasets, standardisation of data and indicators, and the development of tools to access and query data, the subject of **section seven**.
- 1.0.16 The **conclusion** to this report offers thoughts about next steps. This is followed by a section on **detailed, technical recommendations**.
- 1.0.17 The **appendices** contain a note on the methods underpinning our audit and schematics to illustrate how data about the cultural sector could be collected by DCMS, a key central data repository. There is also a discussion of data matching and the role of URNs, using several worked examples.

## 2. How financial value is created in the cultural sector

- 2.0.1 In order to audit data on the financial value of the cultural sector, it is helpful first to set out an analytical framework for where value is created. This section explores how value is created in each stage. Our framework is a model of the production and consumption of cultural products and services, illustrated in the value chain below.

### Value creation around cultural services



- 2.0.2 Organisations combine assets and resources to develop services, with which users engage. We think each step represents a measurement point where there is potential to capture data about the value and impact of cultural activities. Appendix three contains a detailed discussion of each point in the chain, which includes:

#### Organisations

- 2.0.3 There are many types of arts, culture and heritage organisations providing cultural services across all sectors. It is worth noting these include universities, local authorities and unincorporated community and voluntary groups.

#### Assets

- 2.0.4 These are either tangible, such as buildings, land, artwork, costumes, or measurable intangible assets. The latter includes productions (such as theatre or TV productions), recorded music, exhibitions or software, and related intellectual property rights.

#### Resources

- 2.0.4 Organisations apply resources to assets to devise and deliver cultural services. Creative products require creative people, and also those with technical, financial, management and leadership skills. People - whether employed, freelance or contracted - are for most cultural subsectors the dominant cost.

#### Services

- 2.0.5 The supply of services accounts for most of the activity generated by the cultural sector. This starts with business-to-business cultural services, such as ticketing, marketing and distribution. Business-to-consumer services could include a film screening, exhibition or play, or a digital product are defined as a cultural service. User engagement data, although incomplete, may provide the best national picture of levels of service provision.

#### Engagements

- 2.0.6 Engagement data is a helpful indicator of the scale of economic activity. We refer to 'engagements' rather than 'sales' because some activities are non-market based<sup>8</sup> - that is, the end-user does not pay for them. Engagements include providing business to consumer ancillary services such as hospitality or retail.

<sup>8</sup> For further details on non-market-based valuation see <https://www.gov.uk/guidance/culture-and-heritage-capital-portal>.



### 2.0.7 Users

The cultural sector provides services to a wide range of users, from individuals to other cultural organisations. Each category generates different direct and indirect economic benefits: a local family will likely purchase different cultural services than an overnight business visitor.

## 2.1. What drives the generation of data on the financial value of culture?

2.1.1 The section above highlighted the locations in the value chain where value is created, implying that these points in the chain are where data on value can be captured. This section builds on this analysis by exploring what drives or triggers data generation through the supply chain.

2.1.2 On the supply side, there are three main drivers of data generation that can be used to map and track goods and services supplied across a supply chain. These are:

1. **regulation and traceability** – of goods/services manufactured, of quality, of liability risk and responsibility;
2. **market data** – production, sales, pricing, volumes traded in key marketplaces or points in the supply chain; and
3. **management of ownership and rights associated with intellectual or physical property** - ensuring that rights are appropriately licensed so that they deliver a financial return when used by others.

2.1.3 The cultural sector is not highly regulated. Similarly, traceability through the supply chain is not required. There is little by way of data generated above and beyond the financial data already covered above.

2.1.4 Market data, however, is generated by a variety of commercial and non-profit entities, mainly on a subsector basis. The volume and quality of market data are better in subsectors with a substantial element of commercial activity. The most common reasons for harnessing supply chain data for other purposes are policy and lobbying activities.

## 2.2. Which subsectors are collecting data on value?

2.2.1 The diagram below identifies key data sources on cultural supply in the main subsectors, collected from the sector's supply chains. It distinguishes data generated in the course of operations, such as sales data, from survey data. Our audit prioritised routine administrative data collection but included survey data where this administrative data is missing. We have sought to identify datasets with the best reach, quality and consistency of definitions to result in the shortest practicable list of priorities.

## Data holders & types of sectoral data

Sectors	Organisations	Assets & Resources	Services	Engagement	Users
Film	<b>BFI</b> • Own data • Commercial data feeds	<b>BFI</b> • Funded projects • Film tax relief	<b>BFI</b> • Own data • Commercial data feeds	<b>ComScore</b> <b>Omdia</b>	<b>BFI</b> • Own data • Commercial data feeds
Theatre & Performing Arts	<b>ACE</b> (portfolio) <b>Theatre Trust</b> <b>UK Theatre</b>	<b>ACE</b> (portfolio) <b>NLHF</b> (heritage buildings)	<b>Audience Finder</b> (funded) <b>TRG/Purple 7</b> (commercial funded) <b>Ticketing Services</b>		
Music	<b>UK Music</b> <b>ABO</b> <b>Music Venue Trust</b>	<b>Audience Finder</b> (funded) <b>TRG/Purple 7</b> (commercial funded) <b>PRS for Music</b> <b>Ticketing Services</b>			
Heritage (excl. Mus/Gal)	<b>Historic England</b> (in development) <b>NLHF</b> (project) <b>Heritage Sector Bodies</b>	<b>NLHF</b> (project) <b>Historic England</b> (listed) <b>Heritage Sector Bodies</b>	<b>ALVA</b>		
Museums & Galleries	<b>NLHF</b> (project) <b>ACE</b> (portfolio) <b>Music Sector Bodies</b>	<b>ACE</b> (portfolio/capital/projects) <b>NLHF</b> (portfolio)	<b>Audience Finder</b> (funded)		

- 2.2.2 In recorded and live performance, the supply chain stages covering funded organisations, assets & resources and services have the best datasets. The film subsector has relatively strong supply-side data across funded and commercial organisations. This is due to the particular nature of film financing and the combination of a non-profit lead body and effective commercial market intelligence agencies. It has excellent admissions and sales data but relatively less user intelligence data.
- 2.2.3 Funded arts have an unusual and helpful level of aggregation of engagement and user data. This is due to the history and centralisation of funding and related reporting. Ticketing services such as TicketMaster and ticketing system vendors such as Spektrix control significant additional engagement and user data, but this is often inaccessible due to commercial confidentiality and data protection requirements.
- 2.2.4 Data sharing across the supply chain and in regional clusters is more prevalent in the funded arts, thanks to Arts Council England requirements and a less competitive environment. In the television industry, where the national broadcasters and new entrants such as Netflix and Amazon hold substantial datasets on users and engagement levels, this is used for commercial advantage, such as predicting demand for a particular product or genre. This data is rarely made available to content producers.

- 2.2.5 Museums and galleries are similar to the funded arts in the structure of engagement and user data. However, the lower level of ticketed events means a key data source is missing for this sector. Looking ahead, the National Lottery Heritage Fund and Historic England are emphasising public engagement and diversity of access in their strategies. This more joined-up, national approach among the non-museum and gallery heritage category may result in more demand-side intelligence and customer insight.
- 2.2.6 Some organisation and resource data may be available from museums and galleries with regular revenue funding from an arts funder, such as Arts Council England. However, this is unlikely to be comprehensive as there remains a financially ‘hidden’ level of provision. This is within local authorities, universities and a range of privately and company-managed services, where spending on cultural provision forms part of an organisation’s broader activities and is not separated in any financial reporting.
- 2.2.7 Where organisations receive capital funding, data may provide some insight into assets, but little organisational finance data is typically collected. When the National Lottery Heritage Fund supports projects, it collects financial data on a rolling basis on its current and recent portfolio of organisations. Historic England is attempting to map the whole sector in terms of organisations, even where no associated financial or provision data is available.
- 2.2.8 This section has set out the analytical framework that we used to undertake our audit. We think this is a useful way to approach data collection in the future and interrogate data across the different subsectors.
- 2.2.9 Our audit has suggested that while some subsectors collect data on the financial value of their activities, there is significant variance. Therefore, the following sections propose the phased development of the cultural sector’s data infrastructure, starting with developing a data platform. Section three sets out the specification for this platform.

### 3. What data do we want to help us demonstrate value?

3.0.1 The previous sections have identified how and where data is created in the value chain. Given the availability and quality of existing data, this section now addresses what data we think should be collected to demonstrate the value and impact of the cultural sector. It begins by setting out a specification for what data is needed and is followed by a review of data types.

3.0.2 As noted in the introduction, we have prioritised administrative data generated in the ordinary course of supply chain operations. We believe that administrative data is more comprehensive in coverage and a more reliable source of trends data than cross-sectional surveys.

#### **The data specification**

3.0.3 To report on the financial value and impact of the cultural sector consistently and robustly, we think that any dataset should meet a precise specification. The dataset should:

1. cover the whole of England (and ideally include Wales, Scotland, N. Ireland);
2. be updated annually or, preferably, in real-time;
3. use unique identifiers for critical entities such as organisations or sites that enable traceability through the supply chain and linkage between data sources;<sup>9</sup>
4. map the supply side of the value chain covering the sites at which cultural activities occur, outlined above in section two;
5. include all subsectors in the DCMS definition of the cultural sector;
6. make possible the identification of organisations according to the service they provide. For example, it should be possible to distinguish between an arts centre, a theatre or a music venue;
7. include activity within the cultural sector with no physical point of supply, such as digital content, experiences, services;
8. include organisations not permanently based in consumer-facing locations, such as production companies without a permanent venue or individual creators;
9. include business-to-business transactions such as commissions or outsourcing of production services;
10. contain administrative data created in the course of the operation and regulation of the cultural sector supply chain, distinct from data collected as a result of surveys or research;
11. contain organisational/site level financial data with coverage of income & expenditure, capital & revenue, assets, working capital, reserves, employment – and ideally other expenditure accruing to the cultural sector. This includes spending on freelancers;
12. include employment data, containing headcount estimates and employment costs) for Pay as You Earn (PAYE) staff and freelancers/self-employed;
13. contain data on the key assets held by cultural organisations. This should range from buildings and collections to creative works of all kinds and their related IP. Examples include film, theatre and other performance productions, music and other published works, exhibitions and other craft or visual artworks;
14. be stored and accessible in a machine-readable format. We recommend the development and use of an API for the source data with definitions for each data point; and
15. contain a machine-readable licence that enables the data to be used where it is required while including any necessary protections.

<sup>9</sup> An example of how URNs are used by funders is 360Giving, which allows matching, traceability and aggregate analysis.

3.0.4 Much of this data is published in various forms by organisations in the sector describing complex supply chains. However, three immediate challenges are apparent from our audit:

1. **fragmentation:** there is no single source of data for the cultural sector;
2. **harmonisation:** data definitions are not harmonised across these various sources, making difficult comparison and aggregation; and
3. **coverage:** there are undoubtedly gaps in the data coverage.

3.0.5 These are challenging, though not insurmountable, problems. We believe that an iterative process should be undertaken, whereby any initial repository is built with the minimum number of data sources required to give adequate geographical and sectoral coverage. This will mean that work to harmonise definitions and check data quality is manageable.

3.0.6 Over the longer term, it will be necessary to improve the reporting of supply chain data in terms of quality, accuracy, and reach. This will require agreement on a set of key metrics and related data standards which deliver value at both the organisational level and when aggregated. Where data is limited, this can be augmented from survey data, though it is likely that new surveys may need to be commissioned. The membership of an umbrella body or organisations in a supply chain can act as a sample frame.

3.0.7 It will also be essential to take an informed view about licensing. This should balance the potential gains from data sharing while protecting ownership and confidentiality where necessary. Data held by DCMS would need to adhere to any restrictions, commitments or confidentiality set up at the point of data collection. There is also value in making aggregate data publicly available across the sector to maximise benefit, utility, confidentiality, and privacy. We recommend that making aggregate data available online, using a mix of dashboards and standard query tools, will widen the use and increase the value of data, encouraging its use by the sector in performance management and strategy development.

**The data specification: data units**

3.0.8 The data specification also needs to set out the granularity of the data we are seeking to collect to maximise our ability to drill downwards and aggregate upwards.<sup>10</sup> The key data types for which the unit of measurement needs to be agreed are:

1. **cultural suppliers** - the database should take the legal entity and site of cultural supply as the starting point for any data framework;
2. **time** - some feeds, such as ticketing data, update daily and report activity over 24 hours. Others will only contain annual data. The starting point for any data on time periods should be aggregation into a financial year, where the start and end date are based on the organisation's tax year. It is not the goal to look at fluctuations within a single year;<sup>11</sup>
3. **money** - the financial data unit is pounds sterling; and
4. **staffing** - the unit of staffing is ideally FTE. However, IDBR data relating to employment only includes the number of employees and people employed. For data on FTEs, survey data is likely to be required. Regulatory data for some organisations is meant to report FTEs, but data quality is problematic.

<sup>10</sup> Our principles for such data structures are set out in our paper titled 'Is your data inside out or outside in?' [www.mycake.org/news/is-your-data-inside-out-or-outside-in](http://www.mycake.org/news/is-your-data-inside-out-or-outside-in)

<sup>11</sup> April-March and January-December are the most common year start/end dates

3.0.9 In addition to setting out the units of measurement, we also need to determine the most common data segments against which DCMS wishes to report. This type of profile data will be held for each organisation and site and is used to set up reporting slices. We suggest that this should include:

1. Sector segmentation
  - a. Standard Industrial Classification (SIC) code
  - b. Postcode – enabling allocation by geographic area (regions, cities, local authorities)
  - c. Charity Commission for England & Wales (CCEW) activity coding
  - d. International Classification of Non-Profit Organisations (ICNPO) code
2. Source of sector/membership data (Arts Council England, British Film Institute, Historic England etc.)

## 4. Fulfilling the data specification

- 4.0.1 The cultural sector comprises a complex range of organisations and individuals spanning several subsectors and a wide range of activities. Different and distinct supply chains are in operation, making consistent data collection and synthesis complicated.
- 4.0.2 We therefore recommend that the central spine of the data platform is based on organisation-level data collected by government bodies for statutory and regulatory purposes. These include annual tax returns, PAYE data and annual accounts submissions. This approach would maximise coverage from a small number of sources while minimising challenges such as harmonisation.
- 4.0.3 These data feeds should then be overlaid with data that reflects the diversity of activities and business models in the cultural sector. This will enable the construction of sectoral samples which provide a sufficiently detailed taxonomy of cultural supply activities.

### 4.1. Financial data

- 4.1.1 Financial data is separately reported to central government (HMRC) and regulators (Companies House (CH), Charity Commission for England and Wales (CCEW) and the Financial Conduct Authority (FCA)). These are separate data supply chains that are not always connected. Data is rarely supplied in a machine-readable format.<sup>12</sup>
- 4.1.2 Income and expenditure data are required to summarise the cultural sector's scale, structure, and value. The same applies to subsectors. Historic trends data is necessary to track revenue growth, annual profitability, and contribution to reserves.
- 4.1.3 HMRC data offers the most comprehensive coverage. It is also machine-readable. However, legal constraints mean that it is likely to be unavailable to DCMS in the foreseeable future. The nearest alternative is the Inter-Departmental Business Register (IDBR), a survey population frame held by the Office for National Statistics (ONS). The IDBR is based upon data for organisations registered with HMRC for PAYE and VAT and is augmented by information from Companies House. This provides a feed of total income, expenditure, and total spending on salaries on a per organisation site level.<sup>13</sup>
- 4.1.4 The IDBR is an essential source of data. It has the most comprehensive coverage of legal forms, from sole traders and unincorporated associations to entities that also report to Companies House, CCEW and FCA.<sup>14</sup>

<sup>12</sup> The Financial Conduct Authority (FCA) is the registrar for mutual organisations and operates the mutuals register.

<sup>13</sup> An organisation is defined as a legally registered entity which reports to HMRC, and/or one or more of CH, CCEW or FCA. A site is defined as a geographic location which may or may not also be a legally registered organisation and may also include sites which are held under the umbrella of a larger organisation whether that is a holding company, Local Authority, School or University.

<sup>14</sup> The IDBR covers only those sole traders who have registered for PAYE or VAT.

- 4.1.5 However, the IDBR does not provide a detailed breakdown of data on income and expenditure. It does not distinguish between capital and revenue income, nor between different types of expenditure. It has no data on assets or reserves. Therefore, a separate feed of profit and loss and balance sheet data is required to analyse financial health and resilience.
- 4.1.6 In the absence of HMRC data, regulators – Companies House and CCEW - provide the next best solution. Neither provides data on sole traders or unincorporated associations or data on cultural suppliers where the provider is part of a larger entity such as a Local Authority, University or School.
- 4.1.7 Machine-readable data in iXBRL<sup>15</sup> format is available from Companies House via an application programming interface (API)<sup>16</sup>. When combined with IDBR data, this feed provides data to enable analysis of working capital, reserves and other key metrics based on the balance sheet. Data on organisations with charitable status and a turnover of at least £500k per annum is available from the Charity Commission for England and Wales, again via API.
- 4.1.8 It is possible to collect income and expenditure for registered charities with a turnover below £500k from the CCEW.<sup>17</sup> We can also identify what data is missing for this group. It is also possible to undertake a similar analysis for mutual organisations reporting to the FCA and establish whether data gaps are material.
- 4.1.9 Table 1 below sets out the various types of legal entities found across the commercial and non-profit elements of the culture and heritage sectors. This is based upon their legal form. The table sets out the regulator, agency or government department to which each type of entity reports for each class. The table also sets out the Unique Reference Number (URN) systems in use, such as Unique Tax Reference (UTR) number, and whether data is machine-readable.
- 4.1.10 The purpose of this table is to help inform decisions about which data feeds should be prioritised for any future platform. It is again worth noting that HMRC data is unlikely to be available.

<sup>15</sup> <https://ewf.companieshouse.gov.uk/xbml>

<sup>16</sup> An application programming interface (API) is a connection between computers or between computer programs. See <https://en.wikipedia.org/wiki/API>

<sup>17</sup> <https://register-of-charities.charitycommission.gov.uk/register/full-register-download>



**Table 1: Administrative data submitted for legal and regulatory purposes**

	Submits data to								
	HMRC				Companies House	Charity Commission E&W	Financial Conduct Authority	DLUHC	HESA <sup>19</sup>
	Council Tax Office <sup>20</sup>	PAYE & NI Office <sup>21</sup>	VAT Office <sup>22</sup>	Tax relief application <sup>23</sup>					
URN system	UTR no.	Accts. Ref no.	VAT no.	UTR no.	Company no.	Charity no.	Mutuals no.		
Machine-readable	✓	✓	✓		✓/x <sup>24</sup>	✓/x <sup>25</sup>	x <sup>26</sup>		
Entity Type									
Companies Limited by Share	✓	✓	✓	✓	✓				
Companies Limited by Guarantee	✓	✓	✓	✓	✓				
Community Interest Company	✓	✓	✓	✓	✓				
Limited Liability Partnership	✓	✓	✓	✓	✓				
Charitable Trust	✓	✓	✓	✓		✓			
Charitable Incorporated Org.	✓	✓	✓	✓		✓			
Co-operative	✓	✓	✓	✓			✓		
Registered Society	✓	✓	✓	✓			✓		
Community Business Society	✓	✓	✓	✓			✓		
Local Authority		✓						✓	
University		✓	✓						✓
Self-employed/ Freelancer		✓	✓						

## 4.2. Significant gaps in the financial data

### Subsector classification

- 4.2.1 Current approaches to sectoral classification by regulators and government are of limited utility for describing the cultural sector and its constituent parts or activities. Systems such as the Standard Industrial Classification (SIC) are currently insufficiently granular to identify subsectors such as theatres, although SIC is currently under review. As such, it is not possible to segment into groups that align with each supply chain or subsector in the cultural sector: additional classification data is required to report on the component parts of the cultural sector.
- 4.2.2 This can be addressed by using data held by organisations including funders, membership bodies and ALBs and matching this with the regulatory and statutory data outlined above. This data typically includes membership lists, entities reporting ticketing data and other data reporting organisations and sites. Recent work to develop methods for bulk classification of charities highlights advances in using keyword searching for classification, but this may still require manual labour.<sup>18</sup> Data acquisition should again focus on the minimum number of sources with the best coverage at a suitable level of granularity.

### Freelancers and self-employed

- 4.2.3 Freelancers are a substantial proportion of the workforce.<sup>19</sup> DCMS estimates that 1.6m self-employed people work in the creative industries, or 71% of the total workforce. In the cultural sector, DCMS estimates that 321,000 people are self employed, equivalent to 54% of the workforce.
- 4.2.4 The employment status of freelancers varies. Some are self-employed 'sole traders'; others operate through service companies; some are paid via PAYE. A graphic designer might be a sole trader, while an architect is more likely to be part of a limited company.
- 4.2.5 Many people working in the cultural sector provide services to industries outside the cultural sector (and vice versa).<sup>20</sup> These activities represent a significant slice of the cultural sector. An accurate picture of cultural sector supply chains should reflect these different economic activities.

<sup>18</sup> <https://charityclassification.org.uk/data/charity-classification-report.pdf>

<sup>19</sup> Creative Industries Federation (2017) Creative Freelancers. Freelance roles include: "editors...photographers... producers, publishers, technicians, musicians, managers, composers, live performers...artists, education technicians, gallery educators, evaluators, external curators, historians, project coordinators, film-makers, musicians, fabricators, designers... choreographer, animator, video designer...guest teachers...orchestral musicians" (pp 34-6)

<sup>20</sup> Such as photographers exhibiting and selling their own 'cultural' work alongside commercial work for companies and advertising agencies; cinematographers involved in filmed performance or 'cultural' film as well as high-end TV drama; hair or make-up specialists working on theatre and dance projects in tandem with fashion and publishing.

- 4.2.6 In the absence of HMRC data, no data on freelancers' income and expenditure is available if they are not VAT registered, which describes the majority. There is no available data on reserves.<sup>21</sup>
- 4.2.7 Many freelancers are contracted by organisations. This may make it possible to build a model to illustrate the ratio and relative income and spending of freelancers and employees by examining the expenditure of the organisations that employ them. A small proportion of freelancers provide services directly to consumers, except in subsectors such as craft and visual arts. Information on the value of services supplied directly to consumers by freelancers is not available as there is no means of identifying or aggregating data should it exist.
- 4.2.8 Should HMRC data become available, the taxonomy of activities undertaken by the self-employed is too broad to identify those operating in the cultural sector. One alternative is data from the Covid Self-Employment Income Support Scheme.<sup>22</sup> This identifies arts, entertainment and recreation and is generated by HMRC. This is, however, a one-off, time-limited dataset. It also includes activities outside the scope of the cultural sector, such as gambling, and as such would require considerable manual work.

### **Tangible assets**

- 4.2.9 There is a substantial gap in any data on the cultural assets available to create benefits for customers and audiences. Organisations that hold cultural assets will often maintain internal registers of the items in a collection.<sup>23</sup> However, these may not be valued and recorded on the balance sheet, such as in the case of inalienable assets.<sup>24</sup>
- 4.2.10 Performance productions such as theatre or opera involve the creation of physical items such as costumes or sets. Many are critical to an organisation's ability to 'remount' the production and to continue to generate value in future years, whether reused for other productions or hired out to create a recurrent revenue stream. The cost of creation of these, nor a carefully considered net book value is – in our experience – rarely accurately recorded in balance sheets in non-profit organisations in the cultural sector.

---

<sup>21</sup> We expect that some sole traders will maintain substantial assets or working capital. Their financial resilience would be better measured by looking at pension provision or levels of savings held. Access to income & expenditure data is a greater priority than the equivalent of balance sheet data for sole traders.

<sup>22</sup> Covid Self-Employment Income Support Scheme guidance

<sup>23</sup> A 'collection' is the assets held by a cultural organisation – which will be defined by the purpose of that organisation, so they may be artefacts held by a museum, works of art held by a gallery, books held by a library etc.

<sup>24</sup> Valuations are undertaken for insurance purposes, but for some organisations insurance is unaffordable or inappropriate: the organisation would cease to exist if the collection was lost or sold.

- 4.2.11 Land and buildings are usually listed in an asset register. Their valuations are updated periodically. Organisations that secure long-term debt against such assets are more likely to update valuations regularly. However, valuation approaches range from the cost of building replacement to open market value. Some organisations distinguish ‘heritage assets’ from ‘investment buildings’. We expect these to be the most consistently recorded of all the assets included in an asset register or balance sheet.

**Intangible assets**

- 4.2.12 Intangible assets are those assets without any physical substance, such as copyrights or software. Very few intangible assets in the cultural sector are identified and capitalised. Standard financial reporting does not directly help us estimate their value or levels of investment in intangibles.

**4.3. Detailed financial data: recommendations**

- 4.3.1 The machine-readable data available offers insufficient detail on income and cost types, making additional manual data acquisition a likely requirement. This is expensive on a per-unit basis, necessitating the use of representative samples for each subsector.
- 4.3.2 It is not yet possible to estimate accurately the total number of organisations needed to build a representative sample for each subsector or other clusters based on turnover, location, or deprivation level.
- 4.3.3 Detailed financial data on commercial organisations whose turnover is below £10m or 50 employees is minimal from sources other than HMRC. Companies below this threshold are not obliged to make public their profit & loss sheet via Companies House. Most companies limited by share make use of this exemption. Therefore, access to IDBR data is critical.
- 4.3.4 It may be possible to collect longitudinal data by developing partnerships with key sector bodies to obtain raw data from individual organisations. A likely mechanism is the creation of annual surveys to gather further information from the organisations they represent.
- 4.3.5 We recommend the following in relation to collecting detailed financial data:
1. build subsector samples that are statistically representative of the subsector population;
  2. agree on a taxonomy of income and cost types for which data is required;
  3. source 1-3 years of annual financial data and input manually into the database;
  4. partner with appropriate arms-length bodies and other sector bodies to standardise a short format P&L account so that data conforms to a common template, irrespective of the sector body;<sup>25</sup> and
  5. align this shortened P&L account with the CCEW ‘Part B’ format and the splits of types of grants and contracts for which data is published in the CCEW API.

<sup>25</sup> An example of where such an approach has been delivered can be seen with cultural funding in Canada. See <https://thecadac.ca/cms/en/reports.html>

## 5. The cultural sector data platform: phase one

### Recommendations

- 5.0.1
1. Goal: build a minimum viable product (MVP) based on core national datasets
  2. Development of single secure online database containing both data from key sources and integration of the APIs to secure ongoing data feed
  3. Integration of key government data sources to supply a continuing feed of data
  4. Cultural supply list of entities/sites which covers all key subsectors and all key types of supplier, except for freelancers
- 5.0.2
- Previous sections have set out the results of our work to audit data in the cultural sector. This section now proposes a way forward to bring these different data sources together. The first phase of development involves constructing a cultural sector data platform designed to the specification set out in section four, drawing upon the statutory and regulatory data outlined in section five. We believe that this minimum viable product is the basis to test whether the concept is correct and whether additional development work in phase two is justifiable.
- 5.0.3
- As our audit has noted, building a platform is the first stage of a more comprehensive programme of development required to strengthen cultural sector data. Further development work is discussed in section seven.

### 5.1. The cultural sector data platform: core data

- 5.1.1
- We envisage that the phase one platform should be built around the core data from the disclosed IDBR and public Charity Commission for England and Wales (CCEW) and Companies House (CH) data feeds. This process includes identifying and selecting organisations that have self-reported as suppliers of culture, using the classification systems in each system. This will generate a stream of annual income, expenditure and balance sheet data.
- 5.1.2
- As noted already in section 5.2, the classification data in these three sources lacks granularity. Additional sectoral data is required to add granularity and fill key gaps. As the supply chains in each area of the cultural sector are distinct and separate, several datasets will be needed, in the following order of priority:

Financial data source & content	Coverage achieved	Notes
IDBR – disclosed level data	Site/organisation level data on income & expenditure (VAT data) and staffing spend (PAYE). Segmentation by SIC code (single code per site/org from PAYE).	Does not include self-employed/freelancers. Organisations/sites not included if no PAYE staff or below VAT threshold. URN: IDBR number
Companies House - iXBRL	iXBRL provides balance sheet data on all organisations which report to Companies House	Multiple SIC codes per organisation. URN: Company number.
Charity Commission for England and Wales – API & Part B	Income & expenditure data for each financial year. Government grants/contracts. Balance sheet data for organisations >£500k turnover	No balance sheet data if <£500k turnover URN: Charity Commission number.

### Priority cultural supply data sources

Cultural supply data source	Coverage achieved	Notes
Arts Council England	NPOs	
BFI	Funded projects (BFI) Screenings (Omdia) Sites (ComShare/BFI)	Commercial licensing required from providers
Historic England	Register of heritage organisations	In development - aiming to include data from NLHF, Historic Houses, Heritage Trust Network etc
National Lottery Heritage Fund	Organisations, Assets	Financial data is limited project information and only relates to organisations that have received grant funding recently.
The Audience Agency	Provision (granularity to individual sites/services/events), Ticketing transactions, engagement (granularity to provision against individual households)	Performing arts (including film and music), museums, galleries. Most are funded, but can be extended to cover commercial providers.
Purple7	Performing arts venues and music venues. There is a greater weighting of commercial sites than funded sites.	Performing arts venues with a focus on commercial Holds ticket and provision data.

### **A live data platform for the cultural sector**

- 5.1.3 This first phase of work would deliver a working and live data stream. It will enable high-level reporting on both the scale and shape of the cultural sector. The cultural supply data would be up to date and could be refreshed annually. The financial data would use annual accounts data which has a lag of 9 months after year-end for submission to CH/CCEW/FCA.
- 5.1.4 It will also provide the basis for segmenting the cultural sector into subsectors. Current estimates for the sector are mainly based on funded organisations - a potentially biased sample, as unfunded organisations may have different characteristics or operating models. Once we can map the whole sector and have lists of organisations from sector bodies, we can randomly sample from these population frames to gather full profit and loss information.

### **Technical issues**

- 5.1.5 The focus of this project is to set out what data can be accessed. It does not explore how data would be received, processed, or transformed into a state which could be easily manipulated for analysis. Secure data services to hold the data in one or more databases overlaid with analytical tools will be required. Depending on the source of data and level of detail disclosed, there may well be stipulations about minimum data management standards.
- 5.1.6 We recommend that the data feeds be created to allow the APIs and other sources of machine-readable data to be queried by organisation or site and by year. A further requirement is the ability to segment the data by filters such as geography, subsector, turnover band and Index of Multiple Deprivation (IMD) decile.
- 5.1.7 Finally, in the partial absence of Unique Reference Numbers, manual work will likely be required to match the datasets.

### **Remaining gaps**

- 5.1.8 The proposals for phase one would not address all of the needs outlined earlier in the data specification. Phase two, or later phases, will need to address:
1. data on the cultural supply from freelancers;
  2. data gaps around cultural assets where collections are not included on balance sheets, and when cultural assets are not capitalised on balance sheets, with the resultant loss in any link between expenditure on annual cultural production costs and long-term residual value; and
  3. little or no breakdown of the income & expenditure data, limiting any analysis of different cultural sector business models.

## 5.2. Data querying & extraction

- 5.2.1 It will be necessary to build a secure, online database where financial data is structured by organisation/site and by year and where sectoral segmentation matches the distinct supply chains in operation across each subsector. The minimum requirements for this database include:
1. capability to house multiple data feeds/sources;
  2. updatable in line with the periodicity of data from each source;
  3. match up the various unique identifiers in use across the multiple data feeds to enable the individual organisation/site to be the basic unit against which data is held;
  4. provide the means to download segments of data as determined by the structure of the organisation/site profile data;
  5. track the source of the data which feeds into each element of data held on a single organisation/site;
  6. accessible via an online interface;
  7. enable raw data download into CSV format so that data can be transferred into specialist data analysis tools such as Tableau;
  8. enable a level of data visualisation for the occasions on which specialist data tools are not required;
  9. track data downloads by each user of the system; and
  10. be capable of enabling several different levels of access depending on the profile of each user of the system
- 5.2.2 As the platform will incorporate disclosed level data, it will need to be held under conditions that meet the data security requirements set out by IDBR. Compliance with regulations such as GDPR is also a minimum requirement.



## 6. The cultural sector data platform: phase two

- 6.0.1 The establishment of the data platform and integration of the core statutory and regulatory data will provide total income and expenditure data for a set of cultural suppliers across England. However, for any financial modelling of different cultural sector business models (such as trading, contracts, IP, grants, and donations), additional data is required to distinguish between separate income and expenditure categories. Phase two sets out how to build on the core data collected in the first phase by collecting more granular financial information from the cultural sector.

### 6.1. The cultural sector data platform: additional data

- 6.1.1 The starting point for collecting additional data will be to:
1. review the consistency and completeness of the Companies House iXBRL feed and Charity Commission for England & Wales part B data to establish whether it can be improved. Current Companies House data is incomplete, and for organisations below £10m, turnover is frequently limited to balance sheet data. CCEW part B data is more detailed but still lacks the required data, such as grants data. A review of these sources is a prerequisite to collecting additional data, ensuring we minimise any new activity;
  2. review transaction-level data feeds such as ticketing data on individual income types; what the data definition harmonisation challenges would be regarding financial data sources such as CH etc.; what the data confidentiality challenges would be, and the extent to which rules around data aggregation/dis-aggregation could overcome this
  3. review the extent to which sectoral data sources could be either harnessed or adjusted to work to a universal shortened P&L account; establish what level of work is required with arms-length bodies and other sector bodies to achieve this;
  4. provide data on the cost of production of tangible and intangible cultural assets and the general costs of producing cultural products and services; and
  5. develop a method for improving the data feeds on cultural assets utilised across multiple years but not recorded in balance sheet data. This is a significant proportion of museum collections and a substantial level of intangible cultural assets such as touring theatre productions.
- 6.1.2 Supply chain data at either the organisational reporting or transaction level may not be available directly. In that case, it will be necessary to build representative data samples for each culture subsector. The sector-wide data feeds developed in phase one would also enable representative samples to be built and adjusted over time.
- 6.1.3 Concerns about the confidentiality of transaction-level data can be addressed by reference to the innovations in open-banking data that have been developed in recent years. These suggest that there are ways to both maintain confidentiality whilst achieving granularity.
- 6.1.4 In a second stage of development, the essential tasks would be to achieve a greater granularity in the income and expenditure data at an organisational level and evaluate the potential to harness transaction-level data. Ideally, transaction-level data would split out the various types of income.
- 6.1.5 This is likely to be simpler when we have a relatively small number of high-value transactions, e.g. grants awarded, and harder to achieve when they are retail transactions in high volumes as the latter tend to require active participation from individual organisations. In contrast, the former can be achieved by working with a sector partner.

## Phase 2 finance & cultural supply data additions

Data source	Coverage achieved	Notes
IDBR/ONS	Total income & expenditure data via VAT returns for all organisations above VAT threshold; Salary cost data for all organisations who have PAYE data	Will not include organisations whose trading is under the VAT threshold. Will not include freelancers or subcontractor costs as these are not paid as PAYE
Companies House	Organisational level data - Balance sheet data on all organisations (iXBRL), P&L data on some (pdf only)	
Charity Commission for England & Wales	Organisational data - Part B financial data - simplified P&L and balance sheet on organisations >£500k/annum	In our opinion, there is substantial work required to improve the data quality in 'Part B'. The definitions set out in the Charities SORP are inconsistently applied by accountants, and the definitions in the Part B data don't indicate the income source types.
360Giving	Transaction level data on grants awarded	Available now, continuing to be improved. Local Authority grant data patchy.
Arts Council England	Organisational data – finance data from annual surveys, Culture Recovery Fund (CRF) data	Needs alignment of unique identifiers for matching to be automated, e.g. inclusion of company/charity numbers
BFI	Screen/film/site level data	BFI combines data to show an unparalleled picture of the screen sector. However, individual datasets will need licensing from their primary sources (e.g. Comshare). CRF data may provide a valuable one-off addition to the map of provision.
The Audience Agency	Transaction level data from ticketing/admissions	For all organisations regularly funded by Arts Council England - and others as required
Purple 7	Transaction level data from ticketing/admissions	Would require a commercial licence, so prioritise data not available from other sources
HMRC	P&L and balance sheet data for individual organisations	Unlikely to be available in the medium term.

### **Data improvements & harmonisation of definitions**

- 6.1.6 Challenges are expected around data definitions and data standards. The work around is to source high-level data only on organisations and sites. In the longer term, data quality and value will be materially enhanced by improving the interoperability of key data sources, which is discussed in section seven.

## 6.2. Phase two: benefits and limitations

- 6.2.1 The work outlined above would result in the following improvements to cultural sector data:
1. **granularity in the financial data** so that individual income types are visible at an organisational level;
  2. **clarity on the transaction-level data available**, its coverage, the cost of accessing the data and any specifics around confidentiality;
  3. **development of representative samples** for each subsector of the culture economy if required; and
  4. **clarity on the scale of work required to harmonise definitions** across a supply chain and between subsectors
- 6.2.2 Despite this additional work in phase two, several gaps are still likely upon completion:
1. **freelancers:** this work may not address the lack of data on freelancers; and
  2. **small organisations:** IDBR data is expected to fill a substantial proportion of the Companies House data gap on P&L information for organisations under £10m or 50 employees. Organisations reporting to Companies House with a turnover below the VAT threshold or no PAYE staff are not covered by the IDBR. Although the formal economic contribution of small organisations is likely to represent a small part of the aggregate contribution of the sector, they are likely to play an important role at the local level and, as such worth understanding. They are also likely to involve more significant informal resources such as volunteering, which many policymakers are keen to better understand.

## 7. Cultural sector data: challenges and issues

- 7.0.1 Our audit has identified several challenges and issues for collecting and synthesising a cultural sector data platform. The most critical issues are identified in this section.

### 7.1. Access

- 7.1.1 Securing access to IDBR data is critical to the success of this project. There are formal channels to access this data, and the process is well understood. Access to HMRC data is unlikely, and this has been deprioritised in the short term. We propose to work around this by combining IDBR, CCEW and CH data.

Access to sector datasets will need negotiation with individual data holders. Any discussions need to address:

1. **accessibility:** collection methods may not be set up to allow organisation level data to be passed on from the data collector to a third party;
2. **licensing:** this includes specifications about use, sharing and publication. This relates to the maintenance of confidentiality and protection of the commercial value of the data; and
3. **fees for use:** data is costly to collect and commercially valuable. Rightsholders will likely require fees for access and use

- 7.1.2 Access to machine-readable raw annual accounts data from HMRC should remain a longer-term goal to improve the quality of the data platform.

### 7.2. Completeness, accuracy & timeliness

- 7.2.1 The data from IDBR and CH is likely to be complete and relatively accurate. There is a legal requirement both to submit the data and to do so accurately and on time. As IDBR data combines PAYE and VAT data feeds, the data is available monthly and quarterly, respectively, though this is not available for all businesses.

- 7.2.2 Over the long-term, the move to Making Tax Digital (MTD) for all VAT registered entities and all newly registered entities means that more data will be reported to HMRC in machine-readable format every quarter. Access to MTD data may make it possible to shorten the lead time between the end of a financial period and access to data.

- 7.2.3 Data from CCEW is less accurate. MyCake estimates that these discrepancies exist between the data in published annual reports and accounts and the data reported in annual returns for approximately 5% of charities. There is also a lack of consistency in the reporting of capital income. With appropriate cleaning, this data feed can be considered accurate on a national level.

#### **Subsectors**

- 7.2.4 For data from sector bodies, clarity will be needed on the extent to which any dataset is deemed 'complete'. Few umbrella bodies count (or know) all organisations in a subsector as their members. This matters if membership lists are used as proxies that aim to be representative of an entire sector. Nevertheless, representative samples can be drawn using membership data to identify criteria for similar organisations registered with CH or CCEW.

- 7.2.5 In heritage, Historic England is attempting to compile a ‘complete’ register of organisations. Some work may be needed to align its scope and subsector categorisation with the National Lottery Heritage Fund. A fuller picture of the arts subsector can be generated by supplementing the data ACE collects on organisations it funds. Sources include sector body directories and licencing commercially available sector intelligence. This approach is used by BFI, which has an overall picture of the sector it is responsible for, not just the organisations it funds.

### **Freelancers**

- 7.2.6 There is a high degree of certainty that PAYE data will give a complete picture of salaried staff, but this will only include freelancers, temporary or self-employed workers contracted by an organisation in a minority of cases.

## 7.3. Unique identifiers

- 7.3.1 Significant value can be created by joining different datasets covering the same organisations or sites. Joining records from separate datasets is based on matching the data in a common field. This can be as simple as a name, but this is an inaccurate process. Much better is matching by using unique identifiers known as unique reference numbers (URNs).
- 7.3.2 There are some examples in the cultural sector where this works well. The British Film Institute (BFI) has a department that coordinates tax credit work on behalf of HMRC/DCMS, using a connection between a sector data set and the relevant government systems. Matching is based on URNs. This example illustrates that it is likely that HMRC has created an approach to linking Unique Tax Reference (UTR) numbers, VAT numbers, and the Accounts Reference Numbers (for PAYE).
- 7.3.3 However, as the data feed to HMRC is entirely separate from the data feed to CH, CCEW and FCA, we cannot assume that there is a master list that connects HMRC unique identifiers to the Charity, Company or FCA number. We are in the process of confirming this.
- 7.3.4 Although URNs can be shared with data users, it is not common practice for ALBs or sector bodies to hold an organisation’s company or charity number in their datasets. It is more common to record the organisation’s name only. The mismatch between the name held and the legal name is a substantial problem requiring considerable manual matching (10-20%) when joining datasets. Standards such as the 360Giving URN go some way to addressing this problem, but it is still common to have substantial error rates in data submissions. Identifiers for local authorities are also likely to be problematic as each government department uses its own scheme. Until recently gov.uk published official lists of all such registers, but this is no longer in operation.

## 7.4. Data quality & definitions

- 7.4.1 Statutory and regulatory data is sufficient to develop a high-level national picture and segment organisations into appropriate subsectors. However, because some reporting systems do not adhere to data standards or definitions or implement them consistently, it can be difficult to interrogate the data at a more granular level.

- 7.4.2 For example, Companies House, the Charity Commission for England and Wales and the Financial Conduct Authority all use different data definitions, reporting routes and levels of machine-readable data in their financial reporting. This is despite all three regulators collecting data where the submission is both legally mandated and enforced.
- 7.4.3 Several issues are in evidence across the different datasets:
1. HMRC tax return data is fixed format but currently unavailable;
  2. Companies House data does not clearly separate current assets from fixed assets. It is therefore not possible to calculate reserves or, more specifically, Liquid Unrestricted Net Assets (LUNA), an important indicator of financial health;
  3. CCEW and FCA apply definitions inconsistently;
  4. data on charities is characterised by the patchy implementation of the charities Statement of Recommended Practice (SORP). This results in data quality issues with Part B data; and
  5. arms-length bodies such as Arts Council England publish a definitions list and mandate the publication of data by organisations they fund. However, the data collected varies from year to year, creating issues when interrogating data over longer periods.
- 7.4.4 It is possible to create workarounds to minimise issues created by poor quality data by creating generic P&L accounts and using experienced data entry specialists to manually input data in as consistent a manner as possible.

## 7.5. Standardisation & regulation

- 7.5.1 Perhaps the best example of the challenges inherent in standardising datasets is illustrated by the data made available by the CCEW. It publishes two machine-readable datasets:
1. **Alpha feed:** income and expenditure every year for every organisation; and
  2. **Annual Return Part B:** completed by organisations with a turnover greater than £500k
- 7.5.2 The Part B feed reports income and expenditure using categories such as charitable activity, defined by the charities Statement of Recommended Practice (SORP). While every charity should supply their data in a P&L account in the SORP format, it is inconsistently applied in practice. The data is also published in PDF documents which are almost impossible to transpose into a machine-readable format.
- 7.5.3 The result is that much of this data is poor quality or is hard to use. A further problem with the data standard is that the ‘charitable activity’ category is too abstract. More useful would be a categorisation including grants, trading, donations, contracts etc. The same organisations are sending data about their financial activities to HMRC in a machine-readable format, so we do not see data submission in a standardised and machine-readable format as an additional data burden for charitable organisations.

## 7.6. Key metrics

- 7.6.1 High-level metrics such as GDP and GVA are clearly defined, well used and familiar terms. These measures are well established and, at an aggregate level, can be used to describe the value of the cultural sector. However, there is arguably a need to develop more detailed metrics for the cultural sector.

- 7.6.2 There is no commonly agreed set of granular metrics adopted across the various cultural supply chains. This is in part because the availability of relevant data at an organisational level is not consistently produced.
- 7.6.3 Therefore, it would be valuable to develop a set of standardised financial reporting metrics. These should tread a line between being sufficiently generic as to be widely applicable (e.g. a definition of contribution to reserves or use of EBITDA, a definition for working capital) while being sufficiently nuanced to be meaningful, both to DCMS and to individual organisations and sector bodies.<sup>26</sup> For example, there is no single correct answer as to what the ratio of grant to trading income should be across the whole of the cultural sector or even within a single art form. However, we suggest that sufficient data is available to analyse the differences between subsectors, small and large organisations, and commercial and non-profit legal forms.

---

<sup>26</sup> EBITDA: earnings before interest, taxes, depreciation, and amortization  
[https://en.wikipedia.org/wiki/Earnings\\_before\\_interest,\\_taxes,\\_depreciation\\_and\\_amortization](https://en.wikipedia.org/wiki/Earnings_before_interest,_taxes,_depreciation_and_amortization)

## 8. Conclusion

- 8.0.1 The research for this report has identified an enthusiasm across the sector to take forward development work on cultural data and address the challenges involved. There has been a positive response from the owners of statutory and regulatory data, arms-length bodies and across the sector to the questions raised in this scoping study. Interviewees recognise the benefits and practical challenges around implementation.
- 8.0.2 Our audit concluded that there are five opportunity areas for collecting better data:
1. exploiting *administrative data* - this represents the best opportunity for collecting data to regularly report on the cultural sector's impact;
  2. *standardisation* - systematising how organisations collect, segment and analyse data is most likely to generate new insight;
  3. *linking* - connecting different datasets is most likely to add value to the data we collect;
  4. *annual, organisation-level* data is the most valuable data 'building-block'; and
  5. *modernising* how we collect data so that generating new insights is sustainable.
- 8.0.3 **We recommend the setting up of a shared data platform, built and maintained by stakeholders from across the sector, in a phased approach.**
- The cultural sector data platform: phase one**
- 8.0.4 Phase one focuses on core statutory and regulatory data from national bodies. Phase two aims to collect additional data from sources including funders, membership organisations and Arms-Length Bodies.
- 8.0.5 To realise this widely shared ambition to collect better data, our report has argued that we first need to create a framework that maps the whole of the cultural sector using available data sources as they currently exist. To enable this, we have identified:
1. where data is held;
  2. who it is owned by;
  3. what taxonomies apply to each dataset;
  4. the level of detail held; and
  5. the extent to which it might reasonably be expected that these data sources could be combined.
- 8.0.6 This information can be used to map the sector and its subsectors using existing data sources in their current state. Mapping the sector will enable the production of aggregate data about its size and scope, illustrated by data on the number of organisations, their location, activity and function, and who their audiences or consumers are.
- 8.0.7 Using this data, in phase two, it will be possible to create a series of representative samples of the sector. Initially, these samples will contain a limited amount of financial or other data. However, as each sample will include fewer organisations, it will be quicker and easier to acquire complete, high-quality data on individual organisations over one or more years of activity. Extrapolation from these samples will enable the production of detailed estimates of economic activity for each subsector and, in turn, the whole cultural sector. This approach gives both reach and granularity and enables the modelling work to start as soon as possible.



**The cultural sector data platform: phase two**

- 8.0.8 This approach will inform phase two of our recommended approach, which is to build on the work of the existing Lottery Forum to address the quality, accuracy and availability of the data. This stage will look at areas where there are gaps in the data, e.g. on freelancers, as well as improving and rolling out better data definitions and standards.
- 8.0.9 This will consider how the structure of the various supply chains (and the data generated by them in the course of operations) can be matched up to the types of financial data needed for modelling purposes. At the same time, there may be challenges in harmonising data standards, especially where reporting organisations are not accountable to DCMS. Addressing these issues would at least ensure an understanding of the opportunities for and limitations of any data standards alignment work that might be pursued.
- 8.0.10 To ensure that data can be accessed and used by bodies outside government such as arms-length bodies, any data platform used to combine data feeds will need to set out how existing rights and commitments are maintained. The rights put in place should enable flexibility in data querying and the utilisation of organisational-level unaggregated data, or if it is to be sourced from multiple locations with differing ownership and confidentiality rights.

**Why a phased approach is recommended**

- 8.0.11 An advantage of taking a staged approach is that as the benefits of the analytical work materialise and can be demonstrated, it may become easier to harness some of the more granular data. This creates an opportunity for DCMS to build on its own needs for data by creating an opportunity to lead on enabling improvements in data-backed decision-making processes across the sector as a whole. The processes that create access to high-quality data for DCMS could also support methods that widen access across the sector. While commercial operators in the sector are investing in data to support decision-making, it is clear that the non-commercial cultural sector is lagging and does not yet know why it should harness administrative data.
- 8.0.12 The focus of this project is meeting the data needs of DCMS, who are not responsible for developing leadership on data standards across the cultural sector. It may help DCMS to meet its own data needs by by playing such a role. Further work to repurpose data and collecting methods will address questions that were not considered when the current data reporting processes were set up. Addressing them may future-proof subsequent phases of work.

## 9. Detailed recommendations

### 9.0.1 **Short term recommendations:**

**Build a cultural sector data platform.** Data about the cultural sector is fragmented, often inaccessible and incomplete, making it difficult to assess its value, particularly for subsectors. Our first main recommendation is to address these issues by building a cultural sector data platform:

1. Using a phased approach, build a platform comprising a secure, linked data repository and associated reporting and analysis layer
2. Use the platform to aggregate and link data from selected data sources, starting with core statutory and regulatory data from national bodies
3. In a second phase, add additional data from sources including funders, membership organisations and ALBs
4. Use the completed national data structure to develop a classification system for subsectors, and use these as sample frames for collecting detailed financial data
5. Evaluate the platform and test whether the data it aggregates is a viable basis for detailed economic impact analysis of the cultural sector at the subsector level

### 9.0.2 **In detail:**

**Build/commission a platform technology solution consisting of a secure, linked data repository and associated reporting and analysis layer to:**

1. Aggregate and link data from selected data sources per the technical approach developed above;
2. Implementing the agreed URN approach(es) from the options analysis above;
3. Make available reporting and analysis to DCMS and R&D team;
4. Maintain confidentiality and other restrictions imposed by data sources and expected by those contributing - or represented within - datasets;
5. Add granularity to segmentation models in the government data sources via data on cultural supply sourced from ACE, NLHF, BFI, HE, TAA, and other sector bodies.

### 9.0.3 **Test the feasibility and effectiveness of the linked data and repository in modelling economic value through:**

1. Defining a segmentation of cultural s that span the whole sector and provide an opportunity to consider different underlying economic models (e.g. subsector, turnover, location) and data availability or quality
2. Build a national (England-wide) model of economic impact within the solution and using other analysis and modelling tools alongside drawing data from the solution

### 9.0.4 **Assess the limits of development of this approach given the current state of data, quality and access**

1. Potential benefits of this approach over existing methods or alternatives to non-DCMS stakeholders (e.g. arms-length bodies, sector bodies, local authorities, LEPS)
2. Limits to feasibility with existing data sources and technologies
3. Further requirements for data standards
4. Potential impact on other data sources
5. Additional requirements on repository, reporting and analysis platform

- 9.0.5 **Build shared leadership around cultural sector data.** To make data-driven decision making the norm across the cultural sector, government and sector bodies need to help ‘build the field’ – growing and convening a community of skilled researchers and analysts working with cultural sector data and evidence. The proposed remit of the group would be to improve the quality and availability of cultural sector data. The group would lead the standardisation and alignment of data formats and drive improvements across the sector in data collection, analysis and reporting. We suggest that this working group is a ‘coalition of the willing’: interested parties who want to strengthen the field, rather than a working group with an official role. We recommend:
1. expanding the existing ‘Lottery Forum’ role to create a cultural sector data infrastructure working group;
  2. using this working group to help to address the data sources, specifications and standards underpinning the platform proposed in recommendation one; and
  3. co-ordinating action across government and interested parties who want to strengthen cultural sector data
- 9.0.6 **In detail:**  
This group should harness the expertise of key individuals from:
1. DCMS and other parts of government
  2. Arms-length bodies
  3. Cultural sector bodies
  4. Vendors of systems or solutions managing relevant data
  5. Other relevant financial, economic or technical experts.
- 9.0.7 This working group should:
1. Establish key data standards and indicators (recommendation three)
  2. Identify medium to long-term governance arrangements for the project and the ensuing management of the data.
  3. Establish whether a data institution<sup>27</sup> is required to bring stability and consistency to the project.
- 9.0.8 **Develop common standards and indicators.** Building the field also involves setting norms and common language and standards around data. We recommend this work starts by:
1. Setting standards for data collection (such as the 360Giving standard for funders) and analysis, including segmentation of the cultural sector into subsectors
  2. Defining a set of core indicators for reporting on the cultural sector
- 9.0.9 **In detail:**  
The working group set out in recommendation two should establish key data requirements and indicators, including:
1. High-level data points required to model the economic impact of the sector by turnover, subsector and location;
  2. Detailed data points required to evaluate income patterns, investment in cultural production, profitability/contribution to reserves, cultural ‘stock’ (intangible and tangible assets), organisational productivity and resilience;
  3. Minimum granularity of data - for example, at the level of a single organisation or site for no greater than 12 months - i.e. no pre-aggregated data;
  4. Help to identify subsectors based on underlying business models and economic activities.

<sup>27</sup> <https://theodi.org/project/rd-data-institutions/>

- 9.0.10 **Linking datasets.** To maximise the value of existing administrative data on the cultural sector and minimise the compliance burden from collecting new data, we propose that research and development work is undertaken to:
1. Develop the technical approach required to link together datasets, including the development of a Unique Reference Number (URN) system for the cultural sector
  2. Agree on an approach to encourage the use and integration of widely used URNs, including company and charity numbers, into cultural sector datasets
- 9.0.11 **In detail:**
1. Agree on a path to the integration of widely used URNs, including company and charity numbers currently used to identify unique entities, including organisations, sites and services
  2. Encourage and support usage by key government, arms-length body and sector data sources
- 9.0.12 **Develop a detailed technical approach to importing and linking data from the following stewards:**
1. Machine-readable data from the Inter-Departmental Business Register<sup>28</sup> (IDBR), Charity Commission for England and Wales, and Companies House
  2. Sources of cultural provision and providers (e.g. arms-length bodies, TAA, sector bodies)
- 9.0.13 **Research the viability of different options for approaches to URNs:**
1. Investigate how a universal URN can be used to aggregate data about a single entity in the sector through techniques such as matching between different data sources, mapping between different unique identifiers, pattern matching of various attributes (e.g. similarity of name and location) or other approaches (e.g. semi-automated with human intervention)
  2. Define the ways URNs are imported, stored, linked and represented
- 9.0.14 **Medium-term recommendations**
- As noted above, our medium and long-term recommendations are more aspirational. Further details are contained within the conclusion.
- 9.0.15 **Develop standard approaches to segmenting /classifying cultural sector entities.** The ability to segment the sector into standardised, widely recognised subsectors will enable consistent reporting and analysis. We recommend:
1. Development of a single, standard model to segment the cultural sector into subsectors
  2. Working with regulators to improve how they classify data, building upon systems to categorise organisations such as the Standard Industrial Classification
- 9.0.16 **In detail:**
1. Development of a single, standard model to segment the cultural sector into sub-sectors. This should:
    - a. move beyond SIC codes, which lack granularity
    - b. consider categorisations used in government data sources and data on cultural supply sourced from Arts Council England (ACE), British Film Institute (BFI), National Lottery Heritage Fund (NLHF), Historic England (HE) together with The Audience Agency (TAA) and other sector bodies

<sup>28</sup> <https://www.ons.gov.uk/aboutus/whatwedo/paidservices/interdepartmentalbusinessregisteridbr>

- c. map to existing reporting standards such as SIC codes & Charity Commission classifications
  2. Working with regulators to improve how they classify data, building upon systems to categorise organisations such as the Standard Industrial Classification.
  3. Aligning different regulators' classification schemes
  4. Improve the granularity of regulators' existing classification schemes
  5. Working with the Charity Commission to improve the granularity of their sector classification scheme, including better alignment with other taxonomies
- 9.0.17 **Collect financial data from subsector surveys.** Detailed financial data is critical to demonstrating value, but this is unlikely to come from existing administrative data. To address this gap, we recommend:
1. The development of representative subsector samples and collection of financial data for these entities, typically using surveys
  2. Longer-term, working with funders and sector bodies to develop a standardised, short profit and loss template for collecting financial data
- 9.0.18 **In detail:**
1. build subsector samples that are statistically representative of the subsector population;
  2. agree on a taxonomy of income and cost types for which data is required;
  3. source 1-3 years of annual financial data and input manually into the database;
  4. partner with appropriate arms-length bodies and other sector bodies to standardise a short format P&L account so that data shared with DCMS conforms to a common template, irrespective of the sector body;<sup>29</sup> and
  5. align this shortened P&L account with the CCEW 'Part B' format and the splits of types of grants and contracts for which data is published in the CCEW API
- 9.0.19 **Improve regulatory data.** Administrative returns submitted to regulators such as the Charity Commission for England and Wales generally provide a consistent and reliable data source. However, the accuracy and consistency of this data requires improvement. We recommend:
1. Supporting regulators to enhance the quality of financial data obtained via mechanisms such as annual returns, including better data validation.
- 9.0.20 **In detail:**
1. Work with the Charity Commission for England and Wales to improve consistency in financial reporting. In particular, improving the quality of Part B of the annual return.
  2. Consider the feasibility of working with regulators, funders or other interested parties to mandate rather than recommend reporting using standardised categories of income, expenditure, liabilities and assets
  3. Exploring the potential to acquire data on the cost of production of cultural assets
- 9.0.21 **Modernise data acquisition/analysis to support sustainable cultural data.** To reduce the costs of data collection and modernise processes, we recommend:
1. Increasing the use of machine-readable data. We recommend undertaking further research to explore the barriers and opportunities for machine-readable data/iXBRL
  2. Making access to support organisations' longitudinal data easier. Recommended development work includes building partnerships with support organisations and sector

<sup>29</sup> An example of where such an approach has been delivered can be seen with cultural funding in Canada (see CADAC for details)

bodies that hold data on the cultural sector to facilitate sharing and understanding of how data is collected and used.

9.0.22 **In detail:**

**Machine-readability & integration**

1. For all data sources and providers, map the barriers and enablers of data for machine-readable data and how much data is currently and potentially machine-readable
2. In liaison with relevant bodies including the Data Standards Authority<sup>30</sup> and Creative Industries Council Technical Working Group; following ODI and other guidance on open standards to develop a roadmap for machine-readability
3. Review the consistency & completeness of iXBRL data to see if it can be improved

9.0.23 **Long-term recommendations**

**Encourage common reporting formats.** To make reporting less burdensome and improve quality, we recommend the standardisation of application and reporting processes. This could be achieved by:

1. Encouraging ALBs to standardise application and reporting processes
2. Piloting the production of fixed-format accounts

9.0.24 **In detail:**

**Development of standard reporting formats across ALBs**

Over the longer term, we recommend Research and development to explore:

1. whether a standard could be set across regulators such as the Charity Commission for England and Wales and all arms-length bodies that report to DCMS.
2. A short (e.g. 12 lines) fixed format model for profit & loss account and balance sheet data;
3. Whether such a shortened P&L account could be harmonised between DCMS and HMRC.

9.0.25 **Generate more timely data for shorter periods.** Policymakers require more up-to-date data is currently available.<sup>31</sup> To generate more up-to-date data, we recommend:

1. Research and development to explore whether it is possible to reduce the lag between reporting and publication of regulatory data, such as through innovation around data structures and feeds

9.0.26 **In detail**

1. Reviewing the potential to structure the data in shorter periods than annual and the capability of the data feeds to fulfil this;
2. Reviewing the potential to shorten lead times on data availability, i.e. closer to live and shorter than nine months after year-end;
3. Reviewing the transaction-level data feeds such as ticketing systems to see if, over time, these could deliver further functional granularity to the data feeds. This should explore the extent to which anonymisation/pseudonymisation impacts utility.

<sup>30</sup> <https://www.gov.uk/government/groups/data-standards-authority>

<sup>31</sup> The demand for more timely business demographic data is a not limited to the cultural sector. The Office for Statistics Regulation has highlighted this is an area requiring more development work.

9.0.27 **Collect data on freelancers and the labour market data.** Administrative data on the cultural sector workforce is inconsistently reported. Given the prevalence of freelancers, a better understanding of their role is critical to understanding the cultural sector as a whole. We recommend:

1. Building our sense of the scale and role of the self-employed by collecting better data on freelancers, potentially including new sources of data
2. Collect more accurate, consistent data on staff numbers in the cultural sector

9.0.28 **In detail:**

1. Evaluate the potential to gather more consistent and detailed data on freelancer expenditure via arms-length body reporting;
2. Consider how work undertaken by HMRC on the SEISS<sup>32</sup> data could be continued on an ongoing basis;
3. Research how collections, cultural products, and other assets of cultural organisations on which cultural services depend are represented in balance sheets.
4. Longer-term: evaluate the potential for the Charity Commission for England and Wales to record data on staff as FTE rather than average monthly employees.

---

<sup>32</sup> Self-employment income support scheme.

## Appendix 1: Methodology

- 10.0.1 This scoping study sought out datasets that hold a combination of:
1. Profile information on cultural supply
  2. Financial information on cultural organisations.
- 10.0.2 We reviewed datasets owned or managed by:
1. Government departments, related entities and regulators or such as the Charity Commission for England and Wales.
  2. Arms-length bodies, sectors support organisations and commercial entities operating in the cultural sector
- 10.0.3 The method was adapted in the course of the work to accommodate findings made along the way. Most notably, it became clear that the sector bodies do not consistently collect administrative data across the whole sector and that subsector datasets are partial in their coverage. This led to a change in the focus of sector interviews to prioritise locating lists of sector organisations to augment sector segmentation work going forward.
- 10.0.4 In parallel, we conducted interviews with dataset managers in government and stakeholders in Arms-Length Bodies, sector organisations, research organisations and universities. The following organisations have been consulted:
- 10.0.5 **National & local government:** Charity Commission for England and Wales, Companies House, HM Revenue & Customs, Local Government Association, Office of National Statistics (Annual Business Survey and IDBR)
- 10.0.6 **Arms-Length Bodies:** Arts Council England, British Film Institute, Historic England, National Archives, National Lottery Community Fund, National Lottery Heritage Fund
- 10.0.7 **Sector bodies, specialist research units & commercial entities:** 360Giving, Access Foundation, Association of Independent Museums, Association of Large Visitor Attractions, Big Society Capital, CAST, Catalyst, Collections Trust, Creative England, Historic Houses, Karl Wilding, National Council for Voluntary Organisations, Pro Bono Economics, Purple Seven, Spektrix, Social Enterprise UK, Society of Women Art Dealers, Theatres Trust, UK Cinema Association, UK Theatre, University Museums Group, University of Bristol, University of Leeds.

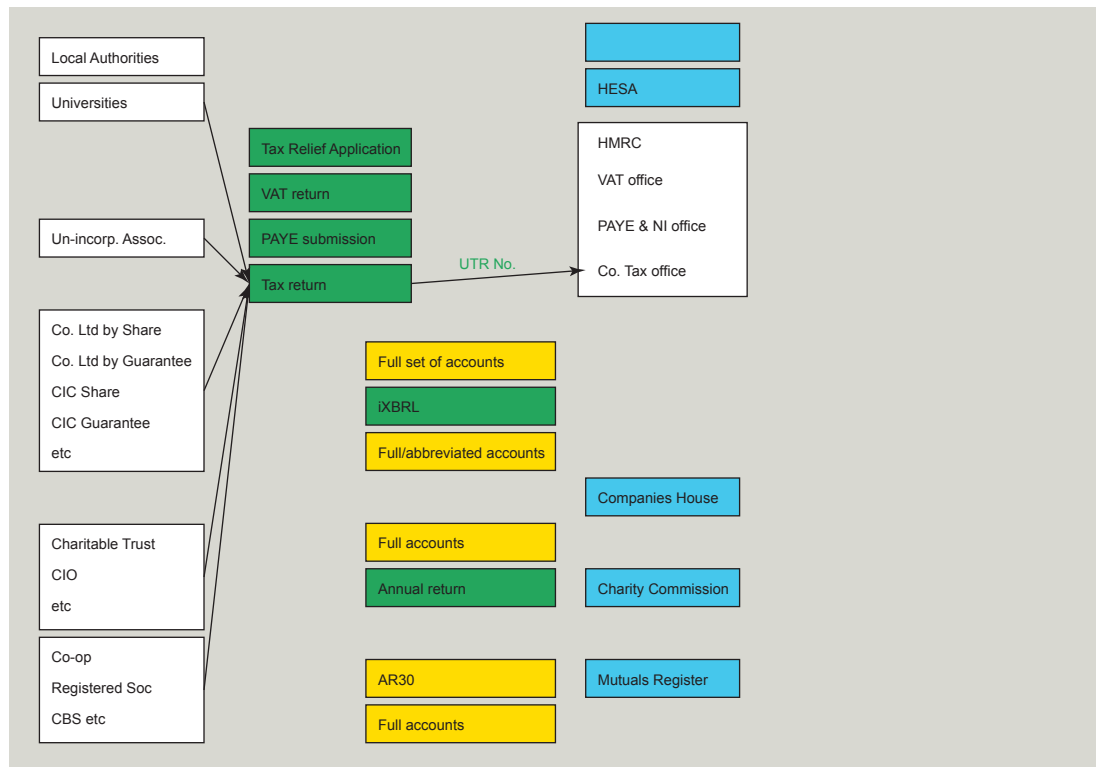


## Appendix 2: Mapping data flows into government

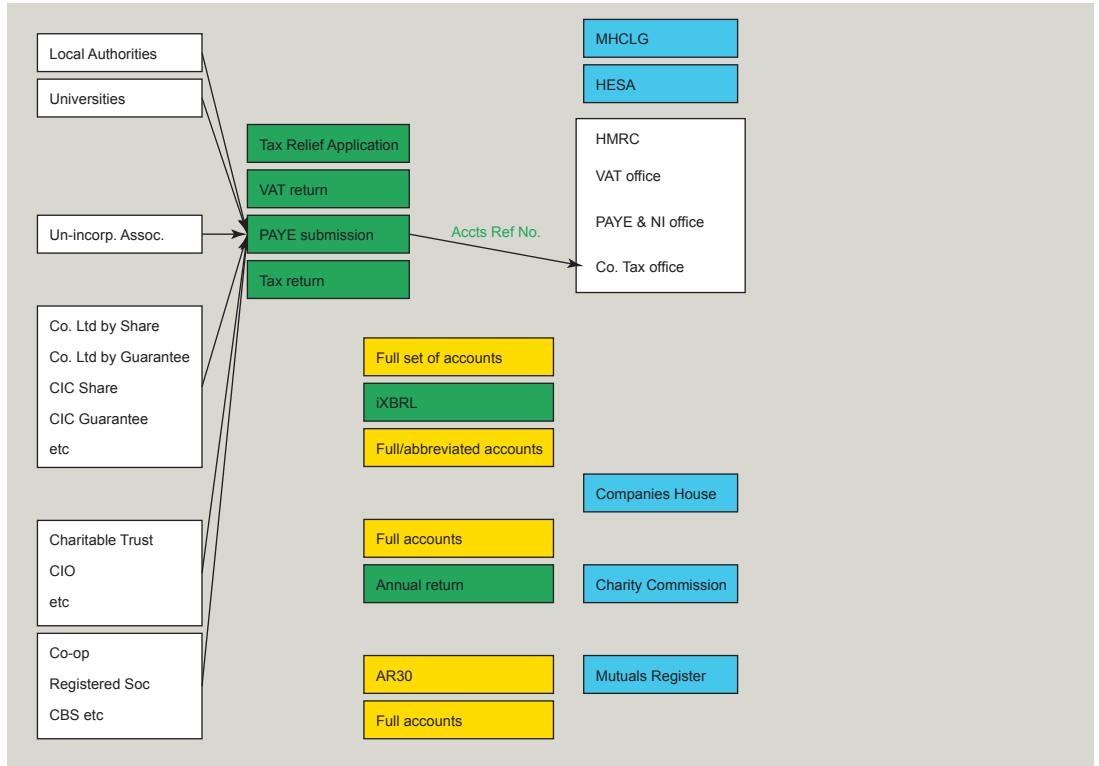
11.0.1 The following charts look at how data is supplied to government and their availability to DCMS. It covers:

1. The legal entities supplying administrative data are on the left in white.
2. Data that is supplied in a machine-readable format is coloured in green.
3. Data that is supplied as PDFs is coloured in yellow.
4. The government departments, regulators and ALBs are marked in blue.
5. The URN in use is written against the relevant arrow.

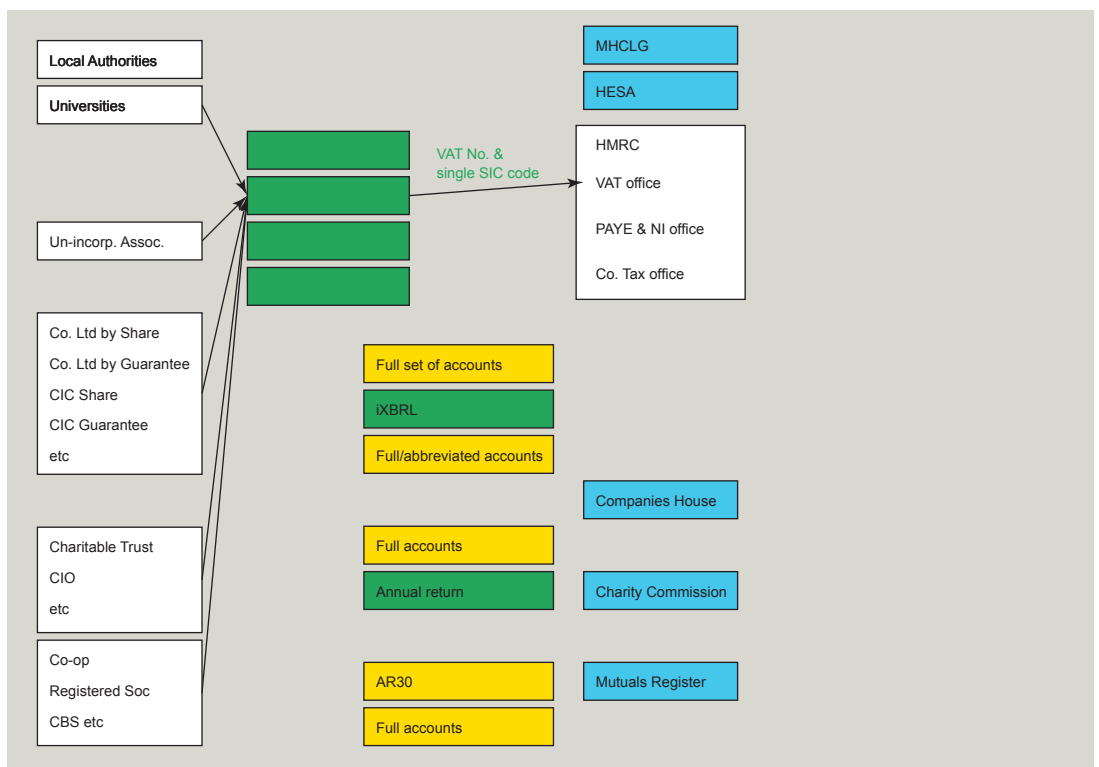
11.0.2 All organisations that engage in ‘trading’ (white boxes) must submit a tax return to HMRC. This data is machine-readable (green)



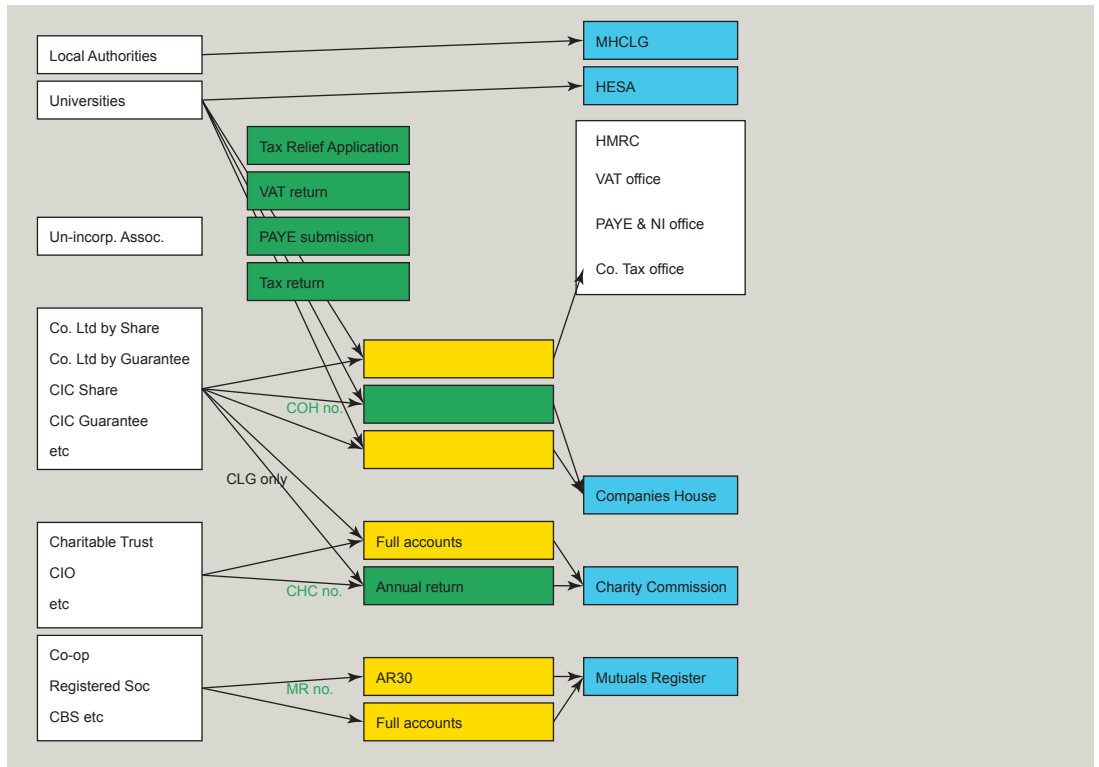
11.0.3 All organisations with PAYE staff submit data to HMRC. This data is machine-readable (green).



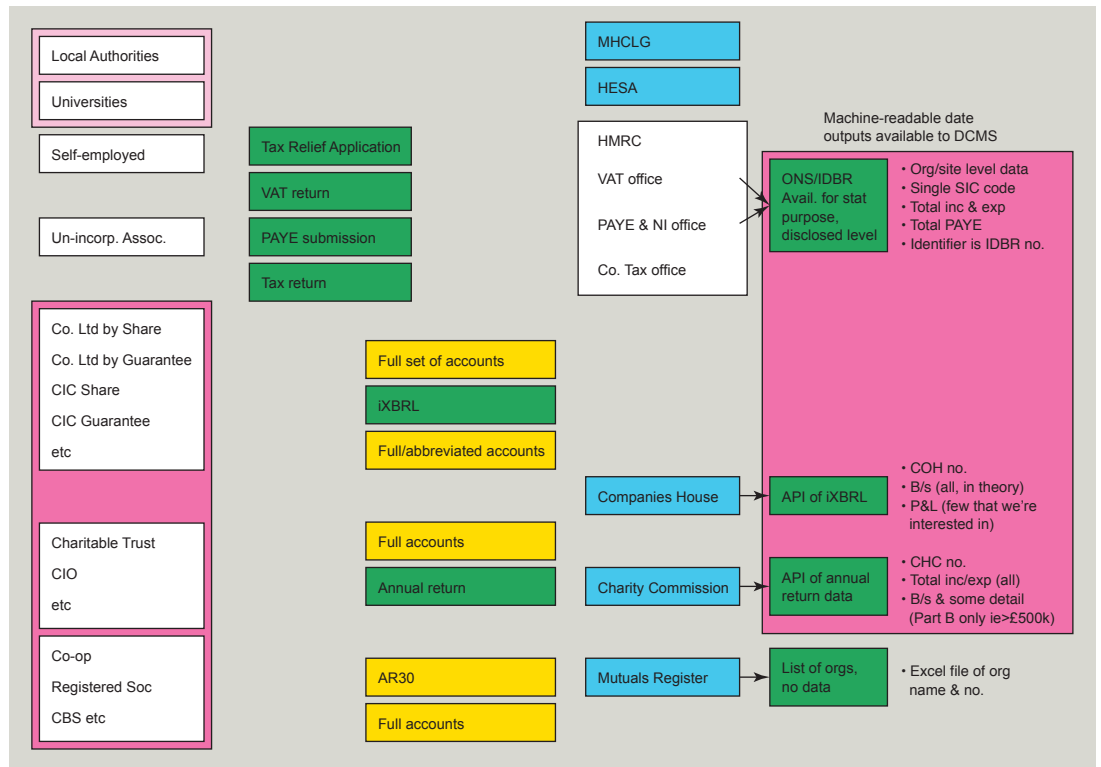
11.0.4 VAT registered organisations must submit quarterly VAT returns to HMRC. This data is machine-readable and has a single SIC code associated with it.



11.0.5 In addition to sending a PDF (yellow) to HMRC, organisations also submit data to one or more of CH, CCEW and FCA. Some of this data is machine-readable (green), but in the main, it is not (yellow).



11.0.6 The data sets that we suggest DCMS prioritises make use of machine-readable data, existing APIs<sup>33</sup> and other outputs. There are some notable gaps that these feeds cannot cover, e.g. data on freelancers. The sections coloured in pink denote the legal entities where we expect data to be accessible (left) and the data outputs via which their data can be acquired (right).



<sup>33</sup> API – application programming interface – is a computing technique that creates and defines interactions between different software applications or between hardware and software intermediaries <https://en.wikipedia.org/wiki/API>

## Appendix 3: Sources of value creation

### Organisations

- 12.0.1 There are many types of arts, culture and heritage organisations providing cultural services. Some fall within a strict cultural sector category, some not. For example, universities and local authorities hold significant cultural assets, provide resources to the cultural sector and provide cultural services to the public. More broadly, Voluntary, Community and Social Enterprise (VCSE) organisations, such as development trusts, often have a cultural strand of their work. Numerous cultural ‘organisations’ are unincorporated community and voluntary groups, particularly in heritage.
- 12.0.2 Working with such a broad range of organisations means that accurate, descriptive and classification data is a prerequisite for analysing impact or econometric modelling. The business models and economic impacts of different subsectors are fundamentally different. Location may also play a part. Therefore, collecting data to describe what organisations do, where they operate, and how their finances work is an important starting point.

### Assets

- 12.0.3 Many cultural organisations hold or manage assets from which they create value. These are either tangible, such as buildings, land, artwork, costumes, or intangible assets. The latter includes productions (such as theatre or TV productions), recorded music, exhibitions or software, and related intellectual property rights.
- 12.0.4 The creative industries, including culture, are defined by products and services derived from unique creative assets. In 2018 investment in intangibles, or ‘artistic originals’<sup>34</sup>, stood at £6.6bn<sup>35</sup>, outstripping investment in tangible assets.
- 12.0.5 Investment levels provide a valuable view of the financial health and medium to long-term performance of a sector. Thus, the health of the UK film sector is determined at least as much by investment in production as in cinemas; the health of the UK theatre sector is determined by investment in new productions as well as in theatre venues.<sup>36</sup>

<sup>34</sup> The measurement of artistic originals in the UK, ONS

<sup>35</sup> Investment in intangible assets in the UK, 2018, ONS. This equates to 3.9% of the £169bn invested by UK businesses in intangibles in 2018. £151bn was invested in tangible assets.

<sup>36</sup> A recent survey estimated that the proportion of organisations investing in new or improved products/services falling into the broad ‘Frascati’ definition of R&D ranged from Crafts (54%) to museums/galleries (29% and performing arts organisations (39%). R&D in Creative Industries Survey, 2020, DCMS p10

12.0.6 However, statistics on the cultural sector's level and type of assets are marred by unreliable data. The capitalisation of intangibles and non-building/land, physical, cultural assets is inconsistent.<sup>37</sup> Recent estimates for the value of intangibles in 'Arts, household and other services' stood at £7.5bn, but only £2.3bn of assets were capitalised in accounts.<sup>38</sup>

12.0.7 Tangible heritage assets are often held in private hands. Trading companies providing cultural services around heritage are less likely to hold assets. Providers of capital funding may have a partial view of the primarily tangible assets they are currently funding. Data on land and buildings may be stored by the Land Registry or, where they are listed, by Historic England.

### Resources

12.0.8 Organisations apply resources to assets to devise and deliver cultural services. Creative products require creative people, and also those with technical, financial, management and leadership skills. People - whether employed, freelance or contracted - are for most cultural subsectors the dominant cost.

12.0.9 Creating a cultural product such as a theatrical production or game usually has separate development, production and delivery phases, with each phase often marked by a financial decision. The phases are:

1. **development:** the initial stage which may not result in a finished product, often characterised by several possible projects that may be investible;
2. **production:** when an organisation and other funders/investors have committed to the production such as a film or musical piece going ahead, likely resulting in a finished product; and
3. **delivery:** the phase when economic and social value is created through the consumption of a product, such as an exhibition or the retail sale of a publication.

12.0.10 These activities in this linear model take place across different organisations. For example, in film and performing arts, many organisations produce work but neither own or manage the venues where the work is exhibited or shown.

12.0.11 Non-employment, non-overhead expenditure data, with some granularity or with modelling of a whole subsector extrapolated from a sample, could provide spend levels to freelancers and other parts of the cultural supply chain. Venues will likely account for the licence or revenue split they pay out to film, theatre or other cultural producers whose work they are monetising.

### Services

12.0.12 The supply of services accounts for most of the activity generated by the cultural sector. A film screening, exhibition or play is defined as a cultural service, albeit based on an underlying cultural product. Similarly, digital 'products', filmed performance and online training are regarded as services.

<sup>37</sup> For an overview of markets in IP and intangibles – and related data sources – see Markets in IP and enabling information ecosystems, 2014, IPO pp21-22.

<sup>38</sup> The intellectual property rights related to the intangibles considered here are mostly copyright and related rights with some design rights, as opposed to patents, so they are not statutorily or even often voluntarily registered.

- 12.0.13 Services can generate direct income, but financial information alone may provide a marginal measure of the economic value created by providing services. Where activities are funded by public money or private philanthropy, the value of that activity may not correlate directly with the level of funding provided. User engagement data, although incomplete, may provide the best national picture of levels of service provision.

### **Engagements**

- 12.0.14 Engagement data is a helpful indicator of the scale of economic activity, especially where there is limited financial data available - for example, free workshops or events for schools. We refer to 'engagements' rather than 'sales' because some activities are non-market based - that is, the end-user does not pay fees: for example, Covid-19 restrictions have resulted in greater use of ticketing for free entry into museums, galleries and other heritage sites to facilitate social distancing. This has provided more data than before.
- 12.0.15 Engagements include providing business to consumer ancillary services such as hospitality, retail or space hire, and business to business services such as events management. Engagement data may relate to a single customer buying tickets for several people.
- 12.0.16 Financial data without engagement data provides only a partial picture for analysis. Together they can show categories of service on which the business models of many organisations in the sector depend (such as corporate hospitality, venue hire) that otherwise may have been hidden. This can give insights into which parts of organisations require funding and which activities, with investment, could provide the most growth in revenue and profitability. Engagement data can be a valuable proxy for missing or partial financial data or data that cannot be disaggregated. For example, where non-ticketing revenues are unreported in financial reports, ticketing engagement data could help estimate income from ticketing.

### **Users**

- 12.0.17 There is no such thing as a typical user in the cultural sector. Although visitors to an attraction or the purchasers of a ticket indicate individual users, any categorisation may need to distinguish between individuals and organisations or purchasers and users. The cultural sector provides services to a wide range of users:
1. individual physical visitors;
  2. schools and teachers;
  3. remote users of online experiences/content who do not physically visit a site;
  4. individual artists, creatives or other cultural practitioners; and
  5. other cultural organisations.
- 12.0.18 Distinctive types of users will generate different levels of direct and indirect economic benefits. For example, a local family, visiting couple, and the overnight business visitor will purchase different cultural services. All will have different secondary spend profiles.

## Appendix 4: Linking datasets using URNs

### Unique Reference Numbers

- 13.0.1 Unique Reference Numbers (URNs) are critical for matching datasets, but not all data holders track company or charity numbers. Data holders should be encouraged, where possible, to collect or append company or charity numbers as a matter of course. Where this is not the case, datasets will need to be matched. Charity and company numbers may not be a long-term solution as URNs for this data platform.
- 13.0.2 We recommend evaluating other options and using these URNs in the meantime. In the meantime, we think it is worth setting out to interested parties the value of URNs and linking data. Linking related datasets offers an opportunity to derive greater value from the data we hold about the cultural sector and support our work to understand value and impact. Linking data enables:
1. improved coverage - validating and filling gaps in data
  2. modelling of activity, reach, resource utilisation and employment - where survey and reporting information can be a proxy for financial data where financial data doesn't exist
  3. analysis of revenue and spending - data from surveys or annual reports can be cross-checked with different kinds of predominantly financial data
- 13.0.3 The following examples show how matching data achieves the coverage we are seeking overall despite it not being available in any one source on its own.

### Matching combined financial/provision data using URNs

- 13.0.4 The sample below combines Audience Finder data from TAA with MyCake data. The left-hand block shows how different datasets end up with different URNs for a single organisation. Where systems use the same URN, these can link records about the same entity from different datasets.
- 13.0.5 The middle block shows how trading and company names often differ. Different spelling or abbreviations make matching using names imperfect. Fuzzy matching technology is now sufficiently advanced to solve this problem where URNs are not present. The right-hand block contains information combined from different systems. In this case, both financial year-end date and a key attribute of provision of performing arts (whether an organisation produces its own work or whether it presents the work of others) are available. These are both required to model economic value from organisations with financially equivalent data and similar underlying business models.



## Matching and combined financial/provision data using URNs

Company Registration Number	Charity Registration Number	Trading Name	Company Name	Year End Date	Receiving, producing or both?
3908975	1080567	Corn Exchange Newbury	Corn Exchange (Newbury) Trust	3/31/21	Receiving
1874868	516673	Theatre by the Lake	Cumbria Theatre Trust (Theatre By The Lake)	3/31/21	Producing
7980160	1148432	Cast	Doncaster Performance Venue Ltd	3/31/21	Receiving
650220	234229	Everyman Theatre Cheltenham	Gloucestershire Everyman Theatre Company Ltd	3/31/21	Receiving
1681278	514719	HOME MCR	Greater Manchester Arts Centre Limited	3/31/21	Both
679960		Harrogate Theatre	Harrogate Theatre	3/31/21	Both
1210050	269645	Hull Truck	Hull Truck Theatre Company	3/31/21	Both
978161	500408	Leeds Grand Theatre	Leeds Grant Theatre & Opera House	3/31/21	Receiving
3802476	1081229	Liverpool Merseyside Theatre	Liverpool & Merseyside Theatres Trust Ltd	3/31/21	Receiving
01266053	271976	New Theatre Royal	Trustees (Portsmouth) Limited(The)	3/31/21	Both
3982202	1082016	New Wolsey Theatre	New Wolsey Theatre	3/31/21	Both
1415547	508829	Oldham Coliseum	Oldham Coliseum Theatre Ltd	3/31/21	Both
2397373	900039	Oxford Playhouse	Oxford Playhouse	3/31/21	Both
556251	233801	Theatre Royal Stratford East	Pioneer Theatres Limited	3/31/21	Producing
815227	253606	Stephen Joseph Theatre	Scarborough Theatre Trust Limited	3/31/21	Both
6002090	1118364	Sherman Cymru	Sherman Cymru	3/31/21	Producing
2053843	295178	St Georges Bristol	St George's Bristol	3/31/21	Receiving
00911924	253242	New Vic Theatre	Stoke-On-Trent And North Staffordshire Theatre Trust Ltd	3/31/21	Producing
3342581	1067869	Courtyard Hereford	The Courtyard Trust	3/31/21	Both
1173859	503597	Theatre Royal Wakefield	Theatre Royal Wakefield - Wakefield Theatre Trust	3/31/21	Both

### Provision data filling gaps in financial data

- 13.0.6 The sample below shows cultural provision sites owned by universities or local authorities. Any financial data for these sites is aggregated within the financial reports of the parent institution. In the case of this group, publicly reported data on ticket sales is unavailable. However, where TAA holds ticketing data within its Audience Finder platform, data could, in theory, be matched to these sites and used to predict this revenue stream and, potentially, total turnover. This could contribute to modelling the economic value of the cultural sector.
- 13.0.7 A cohort of 20 theatres were used to test whether TAA ticketing data was a reliable indicator for the accounting value of ticketed sales for the same financial year (2019-20). Across the cohort, there was a strong correlation between the two figures - with a mean variance of 12% and a median variance of 6%. Although a small number of theatres reported significantly different ticketing revenues, it seems likely that aggregate reporting of ticketing transactions can provide a reliable estimate of this revenue stream for sites where it is otherwise unavailable.
- 13.0.8 In practice, this will require the permission of organisations using Audience Finder. At the same time, any integration will need to respect the existing legal agreements and technical protections that prevent the loss of financial or other sensitive data about individual organisations.

## Provision & engagement data (in this case, TAA) can fill in gaps in financial data

Organisation	Status
Attenborough Arts Centre (University of Leicester)	University
Attenborough Centre for the Creative Arts (University of Sussex)	University
Derby Theatre (University of Derby)	University
Gulbenkian (University of Kent)	University
Lancaster Arts (University of Lancaster)	University
Lincoln Performing Arts Centre (University of Lincoln)	University
Nottingham Lakeside Arts (University of Nottingham)	University
Northcott Theatre (University of Exeter)	University
Royal Welsh College of Music & Drama (University of South Wales)	University
Taliesin Arts Centre (Swansea University)	University
The Edge (University of Bath)	University
Turner Sims, The Nuffield & John Hansard Gallery (University of Southampton)	University
Warwick Arts Centre (University of Warwick)	University
Mansfield Palace Theatre	Local Authority
The Core	Local Authority
Beck Theatre	Local Authority contracted out

13.0.9 **Modelling example: Economic value by service category as well as vertical subsector** Subsector categorisations available from ALBs and sector bodies can help segment economic value by different types of cultural services. With some further analysis of financial data available for individual organisations - together with metrics obtained via representative cohorts of such organisations - the economic value could be assessed for each cell in the table below, reflecting the interaction of both service category and subsector.

## Segmenting cultural activity by service category as well as vertical subsector

Organisation	Performing arts venues	Performing arts producers	Museums & galleries	Other heritage
Ticketed events	£	£	£	£
Admission charges (non events)	£	£	£	£
Hospitality, catering, facilities & space hire	£	£	£	£
Digital content / IP Licensing	£	£	£	£
Education / training / development / talent	£	£	£	£

13.0.10 This level of evidence would provide a much firmer basis for decision-making about funding and other policy interventions. This could support the recovery and growth of different subsectors in a way that is tailored to the different strands of their business models. The Digital Culture Compass activity areas<sup>39</sup> provide a useful cross-check for the full range of categories of 'cultural' services one might consider. For example, a cultural programme including:

1. Producing and programming events/exhibitions
2. Presenting and touring events/exhibitions
3. Audience/visitor experience

<sup>39</sup> Commissioned by Arts Council England with the National Lottery Heritage Fund under the DCMS Culture is Digital programme <https://digitalculturecompass.org.uk/about>

- 13.0.11 Enterprise activities including:
1. Retail
  2. Consultancy - design, creative, technical and commercial
  3. Hire/loan of physical objects

**Modelling example: estimating the value of investment into intangibles**

- 13.0.12 With the exception of some subsectors, investment by the cultural sector in intangible assets is broadly:
1. via a combination of staff time, freelancer time, sub-contracted services; and
  2. across intangible investment categories of 'artistic originals', 'Research & Development', design ('own-account' and 'purchased') and branding (product-related rather than organisational).<sup>40</sup>
- 13.0.13 A theatre or film production company will have a primarily freelance team producing a new film, play, ballet, opera or other performance pieces. A museum or gallery may have staff curators with a few specialist advisors and designers assembling, curating and creating materials around a new exhibition.
- 13.0.14 We can start to piece together what the total investment in intangibles might be. Staff and freelancer costs related to investment in intangibles can be estimated for different subsectors by more detailed research across sample cohorts of organisations from those subsectors. Sourcing staff and freelancer data is discussed under recommendation nine and elsewhere.
- 13.0.15 Direct costs in cultural sector non-profits are markedly higher than comparably sized non-profit organisations outside the cultural sector. At least some of this can be attributed to spending on intangibles - the proportion of which can again be determined by further research on sample cohorts in specific subsectors. For example, small performing arts production companies without a venue will invest comparable amounts via similar mixes of resources (e.g. staff, facilities) and bought-in assets (e.g. intellectual property rights to underlying original works). Larger and mid-sized performing arts producers will have comparable business models and investment profiles in filmed performance. Smaller galleries will create new exhibitions in one way - whereas more prominent museums and galleries will have a few 'tent pole' investments at a much larger (and riskier) scale.

<sup>40</sup> Investment in intangible assets in the UK, 2018, ONS tables 2a, 2b