Detecting Dark Matter Data: data gaps for innovation and R&D activity in the creative industries

PLEASE CITE THE PUBLISHED VERSION

https://doi.org/10.5281/zenodo.7418481

VERSION

VoR (Version of Record)

PUBLISHER STATEMENT

LICENCE

REPOSITORY RECORD

McDonald, Caitlin, and Jennie Jordan. 2023. "Detecting Dark Matter Data: Data Gaps for Innovation and R&D Activity in the Creative Industries". Loughborough University. https://hdl.handle.net/2134/22117118.v1.

Detecting Dark Matter Data:

data gaps for innovation and R&D
activity in the creative industries

Caitlin McDonald, Creative Informatics
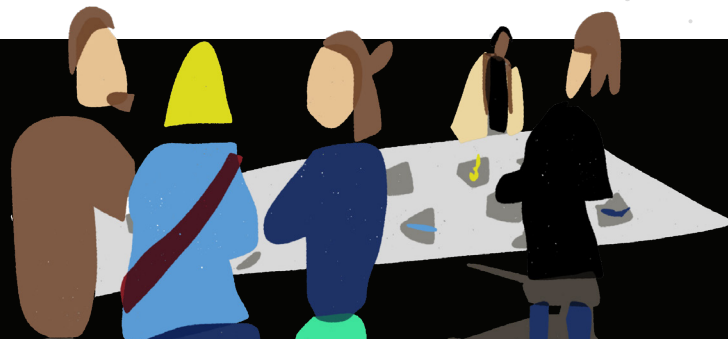Jennie Jordan, Creative Research and Innovation Centre

CREATIVE
INFORMATICS

Contents

Executive Summary

How can we collect data about the creative industries for processing and analysis in a way which is more useful for data consumers (such as policymakers and funders) and for data producers (such as businesses and individual creatives)? Although a wide variety of quantitative and qualitative data sources on the creative industries are available, policymakers and creative practitioners alike still struggle to use data effectively as a decision support tool in their strategic thinking and planning. In particular, there is tacit knowledge that sector activity occurs which is not well-captured through traditional economic analysis mechanisms like Companies House data or the Office for National Statistics' Annual Business Survey. One research participant referred to this data gap as the 'dark matter' of the sector. Our research focused especially on data gaps for innovation and R&D activity in the creative industries.

To support better decision-making for innovation in the creative industries, we need to shine a light on this 'dark matter', improve existing data sources with richer and more frequently updated information, streamline data collection processes to make them less onerous, particularly for small businesses, and encourage greater transparency about how and where data is used and shared after it is collected. In this paper we outline our findings on how different groups of stakeholders think about innovation and data with respect to the creative industries. We look at where different viewpoints on these concepts create challenges for devising appropriate data capture, sharing and analysis mechanisms for the creative industries. We provide four core recommendations to key stakeholder groups working in and around the creative industries: policymakers and policy advisors, membership organisations and trade bodies, individual creatives and creative companies, and data platform providers. We close with a series of provocation questions to spark reflection and open opportunities for change from each of these perspectives.

Figure 1: Participants discussing data challenges for the creative industries. Live drawing from Policy Hack Day, September 2022. [Credit: Elspeth Murray, 2022.]

KEY RECOMMENDATIONS

1. Develop a common data standard for easier data collection, processing & sharing.

2. Collaborate on longitudinal tracking for better cross-sectoral comparisons.

3. Improve data literacy for better strategic decision-making.

4. Encourage interoperability to combine existing and new data sets in ways that add value.

Introduction

Discovering more useful ways we can collect, process, analyse and share data about innovation in the creative industries both for data consumers and data producers in the creative ecosystem (policymakers, funders, individual creatives and creative businesses, etc) is the core focus of this research. This paper discusses our findings and makes key recommendations for improving the creative industries data ecosystem.

This research project is a collaboration between **Creative Informatics**, Edinburgh's **Creative Industries Cluster** breaking new ground in how data-driven innovation can benefit the creative industries, **CRAIC**, the Creative Research and Innovation Centre at Loughborough University London, and **The Data City**, a platform using machine learning to add value to data on innovation in the UK economy. The project makes four key recommendations on moving towards better and more ethical economic analysis frameworks to support creative industries research and development, which you can find at the end of this paper.

What are the data infrastructure and data adoption barriers to innovation in the creative industries? To find out where the biggest challenges and opportunities were, we conducted 28 interviews with individuals and ran five interactive research workshops with stakeholders interested in various aspects of creative industries data. These included creative practitioners and organisations, trade bodies, NGOs, government departments, policymakers, and market and academic researchers in Scotland and across the UK. Everyone we spoke to provides, collates, or uses data within the creative industries. We asked participants how they defined data and innovation, how they engaged with and used data, and what problems they associated with data in the sector. Our interview series culminated in a collaborative 'Policy Hack Day' where we invited research participants to view our early conclusions from the research and work together to identify potential solutions for barriers to innovation in the creative industries. Our core findings follow, with recommendations tailored to policymakers, creative membership organisations, advocacy groups and trade bodies, individuals working in the creative sector, and data platform providers.

But of course creative practitioners work with different sectors at one time!

Figure 2: Fatima Garcia of The Data City describing a common data challenge. Live drawing from Policy Hack Day, September 2022. [Credit: Elspeth Murray, 2022.]

What is innovation?

The most pertinent standard definition of innovation for this research is found in the Oslo Manual, the OECD's guide for collecting, reporting and using data on innovation. According to the **Oslo Manual**, innovation is:

> a new or improved product or process (or combination thereof) that differs significantly from the unit's previous products or processes and that has been made available to potential users (product) or brought into use by the unit (process). This definition uses the generic term "unit" to describe the actor responsible for innovations. It refers to any institutional unit in any sector, including households and their individual members (OECD and Eurostat, 2018, p. 20).

When we asked what innovation was, not one of our research participants referred to the Oslo Manual. Instead, they told us...

- "Innovation isn't a word we use – it's a word that seems like it's from a different discipline."
- "Innovation is a bit of a red herring."
- "I spend about half my life explaining to people what innovation is!"

The Oslo Manual has its uses for standardising definitions of how innovation is operationalised through R&D resources and activities, IP generation, workforce skills, technological change, knowledge flows and supply chains. However, its focus is skewed towards larger firms with in-house R&D and data teams, and towards innovative products, as these are easiest to identify and measure. Hassan Bakhshi, Director of the Creative Industries Policy and Evidence Centre, has argued that this leaves a gap in understanding around innovation processes for the creative industries. Frameworks focused on registered companies, which have the infrastructure and resources to tabulate their R&D activity within formalised requirements for tax credits, industrial surveys, and other purposes, are poorly adapted for sectors such as the creative industries which rely on many small, agile companies and ad-hoc creative collaborations that arise and disband on a project-by-project basis. These smaller, agile organisations include the 32% of the creative industries workforce which is self-employed (compared to 16% of the UK workforce overall), as stated in the 2021 Creative Industries Policy and Evidence Centre report, **'Freelancers in the Creative Industries'.** As we discuss later, ad-hoc creative coalitions and the movement of freelancers from job to job result in the formation of important networks of skills and knowledge between disparate organisations. These networks catalyse innovation across the whole sector, a mechanism Bakhshi addresses in his 2008 work on 'creative linkages' in **'Creating Innovation: Do the creative industries support innovation in the wider economy?'**

What is innovation? Contd.

In other words, the day-to-day work of creative practitioners, policymakers and trade bodies working on behalf of the creative industries involves a range of different innovation definitions, informal conceptual models, and formal frameworks in practice. A local authority might be particularly concerned with social or cultural innovation, for example, while the Department for Business, Energy and Industrial Strategy is far more likely to take an interest in economic innovation.
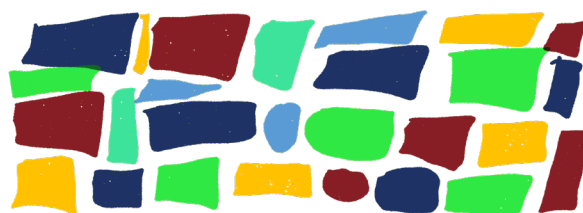
Figure 3

Innovation types & audiences
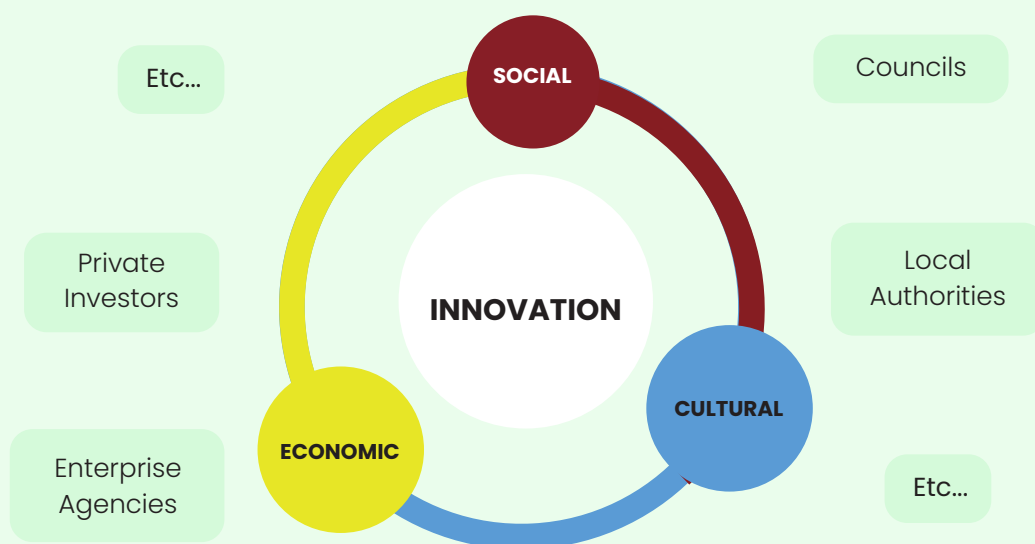
Figure 4

These disparate conceptualisations of innovation require different ways of describing and measuring what innovation is: different types of data.

Figure 3: The data (adoption) barriers to innovation. Live drawing from Policy Hack Day, September 2022. [Credit: Elspeth Murray, 2022.]

Figure 4: Innovation Types and Audiences. [Credit: Caitlin McDonald, Creative Informatics. Licensed under Creative Commons Attribution 4.0 International.

What is data?

Just as with innovation, when we asked about data, we received widely varying responses as to what data is and how it is used for decision-making.

- "If you've got good data it backs up the possibility for innovation" – Research participant 28.

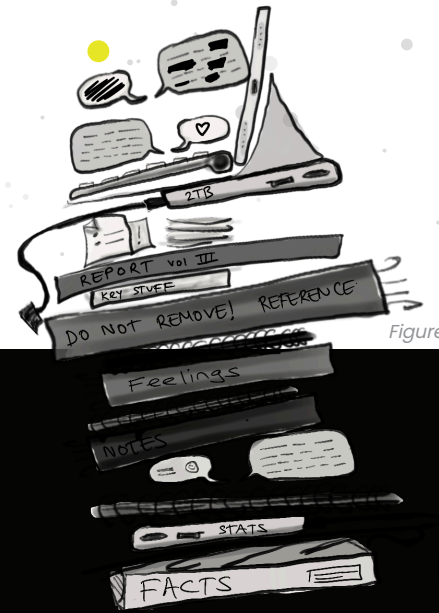- "0s and 1s in different boxes that make our work abstract in some way. It's an intended simplification that ends up being a reduction" – Research participant 17.

- "Key points that help people to understand, create new knowledge or inform them if they are making decisions. It's really always related to that, because data without a purpose is nothing that really matters, in my opinion" – Research participant 9.

When asked, participants typically responded to the idea of 'data' as machine-readable, numerical or categorical data about a business, broader industry segment, or other key topic relevant to the creative industries (e.g., information that might help a business find a collaborator with skills they themselves lacked, identify market trends, or understand the competitive landscape).

However, participants also recognised other forms of data, such as narrative descriptions of organisational activities (case studies), or personal knowledge and networking data. Policymakers recognised three categories of data that are frequently used to support a strategic policy decision: statistics, case studies, and wider evidence.

Robust policy narrative uses...

Statistics

Case studies

Wider evidence

Figure 5: Precarious piles of knowledge. [Credit: Elspeth Murray, 2022.]

Figure 6: Robust policy narrative uses... [Credit: Caitlin McDonald, Creative Informatics. Licensed under Creative Commons Attribution 4.0 International.]

What is data?
Contd.

These categories can vary in importance depending on the types of data available: for example, although sector statistics are often the most trusted and preferred decision support mechanisms by policymakers, the creative industries have unique challenges which make it difficult to collect robust statistics that enable comparisons to other industries. Case studies are often used to complement sector statistics to enrich and humanise the data available, but in the absence of any statistics they are sometimes used as primary data points in themselves. 'Wider evidence' here refers to comparative statistics from other countries, regions, or sometimes adjacent industrial sectors like hospitality and tourism which have some overlap with the creative industries, and are assumed to be at least usefully indicative if not an exact representation of what is happening within the UK's creative industries. In the absence of governmental statistical data, policymakers informally ranked data according to its source as much as its type. Organisations such as UNESCO and universities were highly trusted, while trade bodies were considered to have important insights, but be less reliably objective.

How do policymakers use this data?

Ideally, they go through an evidence-informed decision-making cycle, which one of our research participants articulated as an iterative process starting with initial data and assumptions that become a hypothesis for testing, leading to a proposed policy or strategy narrative, requiring more data to evidence and demonstrate the narrative.
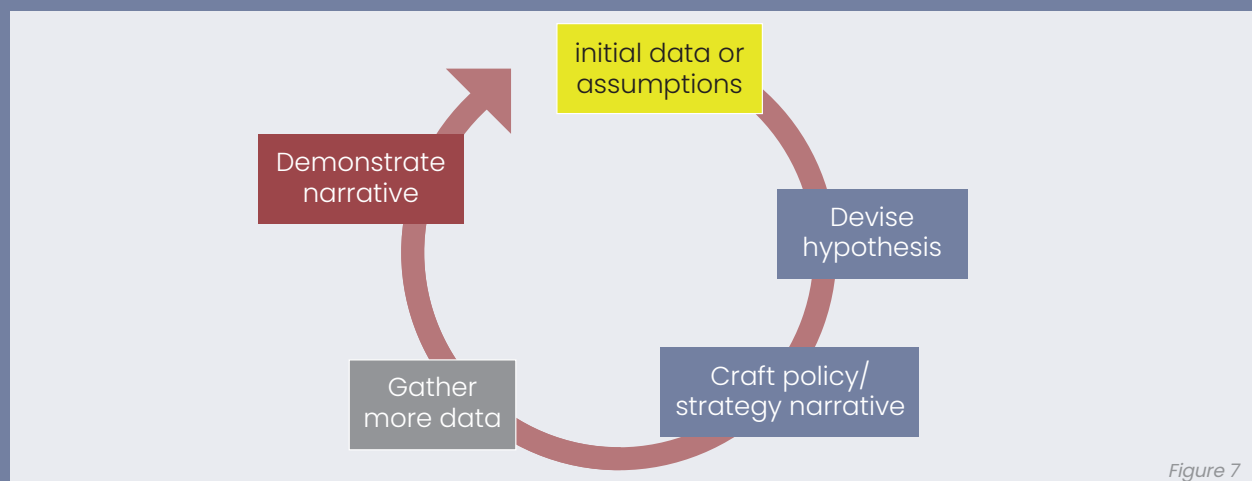
Policymaking decision cycle

Figure 7

Figure 7: Policymaking decision cycle. [Credit: Caitlin McDonald,
Creative Informatics. Licensed under Creative Commons Attribution 4.0 International.]

What is data?
Contd.

One of the challenges we identified in our research is that for the creative industries, rather than collecting data in a continuous learning cycle which uses data in near-real-time to inform future decisions, data is often provided at the end of a grant monitoring cycle, for example, or through infrequent mechanisms such as annual reports. One participant described this as the difference between data for 'monitoring' and that for 'analysis':

Monitoring vs Analysis

Monitoring
- Theory of change
- Assess against predefined goals
- Impact evaluation

Analysis
- The 'So what?' factor
- "See something that others cannot" – Research participant 21
- Set new goals

Figure 8

When data is used solely for monitoring or at the end of an evaluation cycle, it diminishes its value as a decision-informing tool. On the data supplier side, creatives criticised the fact that these monitoring cycles impose a burden of capturing and supplying data to stakeholders like funders, but that this data is a one-way street:

Data flow is currently a one-way street

Audiences → Organisation → Stakeholder

Figure 9

Figure 8: Monitoring vs analysis. [Credit: Caitlin McDonald, Creative Informatics. Licensed under Creative Commons Attribution 4.0 International.]

Figure 9: The one-way street of data flow. [Credit: Caitlin McDonald, Creative Informatics. Licensed under Creative Commons Attribution 4.0 International.]

What is data?
Contd.

This means that some stakeholders in the ecosystem are not getting value out of the data they produce: there are few mechanisms to review or challenge decisions made with this data, nor can individual creatives or creative organisations make use of the bigger picture derived from data for their own strategic planning and decision-making. They do not get to see the full jigsaw puzzle.
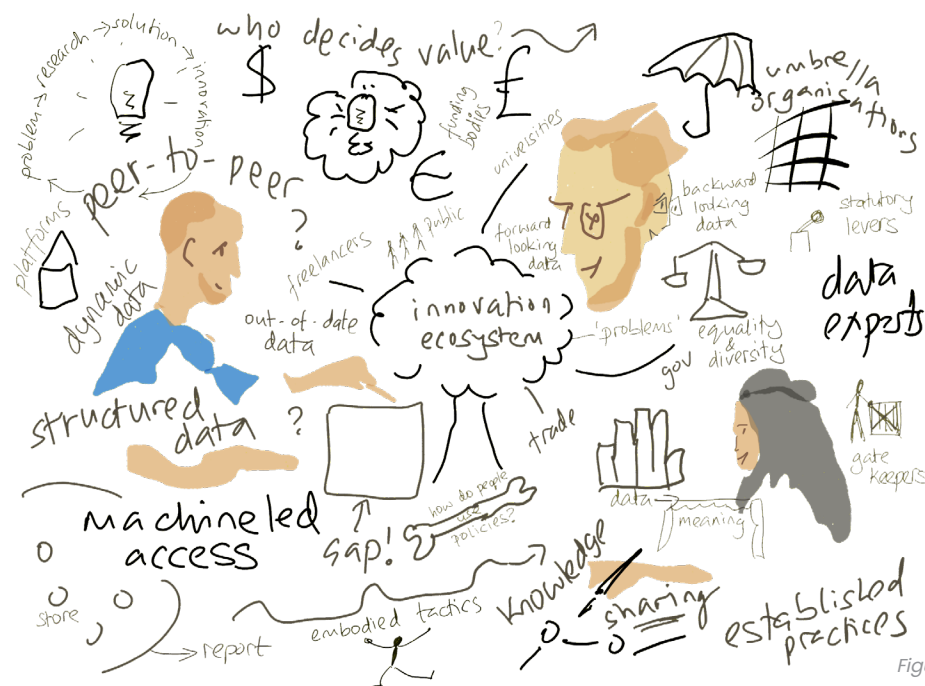
Figure 10

On the other side of the data coin, data platform providers expressed some frustration that the organisations they work with do not make as much use of their own data as they might for creating strategic and tactical plans. Rather than analysing their own ticketing or sales data to investigate changes in audience behaviour or to test theories about what impact proposed changes might have, many organisations asked platforms to provide 'one size fits all' monitoring dashboards and templates that fall more in the 'monitoring/ evaluation' than the 'analysis'/decision-support category. There appears to be a learned helplessness even for medium to large organisations working in the creative industries, particularly in relation to numerical data.

Figure 10: Innovation ecosystems discussion among participants. Live drawing from Policy Hack Day, September 2022. [Credit: Elspeth Murray, 2022.]

Dark Matter
know-how and know-who

The 'dark matter' of the creative industries

Another contributing factor to the difficulty of seeing the full jigsaw puzzle is that the statistical models used for understanding innovation overlook know-how, know-who and knowledge exchange. One participant referred to this lack of knowledge as the 'dark matter' of the sector: "all this stuff that's going on, this activity, but we don't see it, we don't understand it" -- Research participant 16.

Figure 11: The creative industries jigsaw puzzle.
[Credit: Elspeth Murray, 2022.]

Dark Matter
know-how and know-who

Figure 12: The 'dark matter' of the creative industries.
[Credit: Elspeth Murray, 2022.]

The 'dark matter' of the creative industries

Our research highlighted the importance of iterative and interactive processes of knowledge exchange, collaboration and skill development among creatives, the **'doing, using and interacting'** (DUI) mode of innovation. Interviewees talked about their learning processes in terms that highlighted the importance of DUI in three especially important ways: learning by doing to develop skills and know-how; learning by using – a key element in design thinking and user experience (UX) approaches to innovation; and learning by interacting with other companies and creative individuals both in and out of the sector, for example, with suppliers and competitors and informally with former colleagues. The high proportion of freelancers working in the sector, compared to other industries, means that networks are especially important for driving innovation. These interactions happen through individuals moving from company to company and project to project, building expertise, and sharing that expertise in their new roles— something that one participant referred to as "pockets of knowledge that get moved through supply chains or across networks":

Dark Matter

Pockets of knowledge

Further, innovation and 'pockets of knowledge' exchange is spurred by facilitators and co-working hubs such as artist studios, technicians in specialist workshops which creative practitioners consult for specialist advice, festivals, and university art and design courses. The importance of these places for 'pockets of knowledge' exchange is corroborated by recent research from Creative Edinburgh and Creative Informatics, reported in **'Connecting, Collaborating, Creating: The Experiences of Creative Freelancers in Edinburgh in 2022'**. This research showed that Edinburgh creatives particularly value membership organisations like **CodeBase** and **The Melting Pot** as hubs for creating informal and formal knowledge networks among peers.

Methods of capturing innovation as process – as **knowledge** flow, rather than simply through inputs and outputs – are widely discussed in policy and academic innovation literature, with several proposals for capturing the dynamism of knowledge and **creative** innovation. However, putting those models into practice in ways that impact policy decisions, or individual creative companies' and creatives' strategic choices, is a different story. On a positive note, interviewees also told us that "recognition of the invisibility of sole trader creatives is growing", and there

was a desire to know how "to navigate 'dense' data networks to find what connects me to other creatives" outside existing industry niches. We chose **The Data City** as a partner in this research because they are currently moving beyond SIC and SOC codes to provide novel methods for measuring innovation and industrial impact. We found The Data City's analytical tools, while primarily aimed at policymakers, also have potential to help creative businesses and sector advocacy organisations better understand innovation networks for the creative industries.

Figure 13: Pockets of knowledge.
[Credit: Elspeth Murray, 2022.]

Dark Matter

Capturing and structuring knowledge flow data (the 'Doing, Using and Interacting' model previously described) remains a challenge for the creative industries. The Data City's work in this area has potential, but through our work together we were able to identify gaps that need to be addressed for their data aggregation and analytics platform to be considered a truly worthwhile decision-making support tool for the creative industries. These gaps included missing data on any entity operating without a registered company number (including freelancers), the absence of information on charitable funding sources such as the Arts Council, and a paucity of information on cultural activities. Some of this data is publicly available elsewhere, such as Arts Council funding data.

By bringing together a number of data platforms working on data for the creative industries, like audiences, events and market data, we were able to spark a conversation which could lead to future common data standards that will add value for policymakers and creative practitioners. Recent research in this area includes a 2021 **DCMS scoping study carried out by MyCake and The Audience Agency** providing short-, medium- and long-term recommendations to improve data infrastructure and sharing mechanisms for the creative economy, and the Centre for Cultural Value's 2022 report

'Making Data Work: A scoping survey to develop a mixed-methods evaluation framework for culture.' More can be done in this area, as new platforms emerge that are collecting data on highly innovative sectors such as alleviating environmental impacts across live events and film production, and diversity within the creative workforce. A core recommendation of our research, therefore, is to support future initiatives for developing a common data standard to facilitate an overall better picture of the jigsaw puzzle for all stakeholders, leading to better decision-making all around.
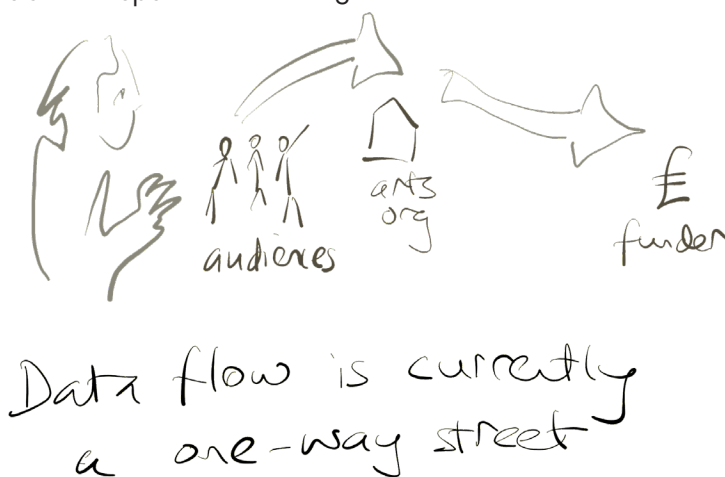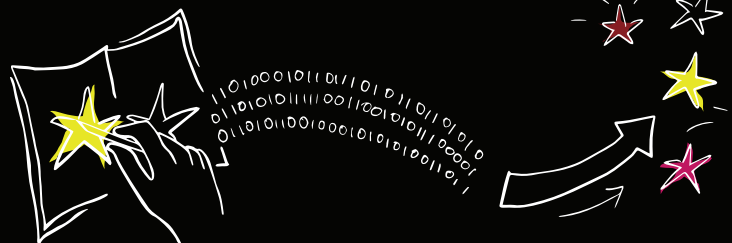

Figure 14: The one-way street of data flow. Live drawing from Policy Hack Day, September 2022. [Credit: Elspeth Murray, 2022.]

Bridge of Change

One of our participants told us that "Data is the bridge between change and art". Whether we are policymakers, creative practitioners, advocacy bodies or academics, data allows us to make better decisions, and it helps us to understand the impacts that our decisions make. The creative industries deserve a data ecosystem that is as rich and informative as those in other sectors. To achieve this, we provide four core recommendations:

1. **Develop a common data standard for easier data collection, processing & sharing.** Policymakers should develop a common data standard for creative industries data that can be collected from individual creative practitioners as well as companies. This will reduce the skews of 'dark matter' in the sector. A common data standard adopted by multiple funding bodies may also help to reduce the repetitive data burden on creative organisations and individuals of reporting the same information repeatedly to multiple places.

2. **Collaborate on longitudinal tracking for better cross-sectoral comparisons.** Creative membership organisations, trade bodies, and sector advocacy organisations should collaborate to conduct better longitudinal tracking of individual creatives' career development. This would facilitate cross-sectoral comparison of career outcomes, and a better understanding of how innovation networks develop over time, especially where technological change is driving synergies between creative sectors and radically changing working practices.

3. **Improve data literacy for better strategic decision-making.** Individual creatives and creative organisations would benefit from improved data literacy, which would enable them to use statistical data, case studies, and wider evidence to better understand the rich creative landscape in which they operate. This would help them to make strategic decisions about their careers and companies, and articulate their innovation value to funders, policymakers, and other stakeholders.

4. **Encourage interoperability to combine existing and new data sets in ways that add value.** Data providers can play a key role in supporting policymakers to develop a common data standard and encourage more interoperability across platforms. Combining existing data sets in new ways that reveal novel insights to all players in the creative industries ecosystem also has potential to add value and facilitate better decision-making across the sector.

Figure 15: Data - the bridge between change and art. [Credit: Elspeth Murray, 2022.]

Provocation Questions

We offer the following reflection questions and prompts as provocations to continue the conversations started through this research. We have devised a specific set of questions for each interest group we worked with during the research, but you may find that looking outside your own specific area of focus helps you see a more holistic picture of the overall creative industries data jigsaw puzzle.

Robust policy narrative uses...

Statistics **Case Studies** **Wider Evidence**

For Policymakers

Conduct a review of your data-driven decision-making:

1. What data do you need to make meaningful decisions?

2. How do you collect that data? Where do you keep it? Who has access to it and how do they use it? What possibilities are you exploring for automation in your data collection, matching and analysis?

3. Consider the 'three bubbles' for robust policy narrative:
   a. Which of these do you rely on most, and why?
   b. What might you learn from trying a different approach?

4. What data do you give back to the communities in your remit and how do you do that?

5. How does the data you collect represent the diversity of people, organisations and structures operating within your remit? Who is currently getting left out? How could you find ways to represent them in your policy narratives?

6. How can you support a data cycle that serves everyone in the creative industries ecosystem, rather than a one-way data street?

Provocation Questions
Contd

For Creative Organisations and Individual Creative Practitioners

Discover and tell your data-driven creative story:

1. How can you find your own 'pockets of knowledge' who contribute to your professional skills development and collaboration opportunities? What data sources can you use to discover these resources?

2. What data do you use to prepare for investment / grant / commissioning opportunities? Are you aware of how that data is being used? How do you situate yourself in the broader creative landscape through your responses?

3. What case studies could you use to help articulate your creative story? How could you use visuals (including data visualisations) to have greater impact?

Provocation Questions
Contd

For Trade Bodies, Advocacy Organisations, and Unions

Enrich the data landscape by working together:

1. How can you collaborate with other bodies working in the creative industries to identify common patterns of information in use for decision-making?

2. Who can you work with to use these common patterns to advise on a shared data standard?

3. What possibilities can you explore to combine data sets with other organisations to find new insights? What safeguards would you need to put in place?

4. What resources would you need to develop more longitudinal research on career trajectories and innovation development over time?

Provocation Questions
Contd

For Data Platforms, Providers and Aggregators

How can you leverage the power of open data to create new value for everyone?

1. Who can you collaborate with to combine data sets to uncover novel insights? How can you make those available throughout the creative industries knowledge ecosystem (policymakers, creatives, trade bodies, etc.)?

2. What interoperability and **open source** principles can you adopt to facilitate a freer flow of information?

3. What potential pitfalls might arise from combining data in new ways or using old data for new purposes? How can you mitigate these challenges ethically, fairly, and legally for all stakeholders?

4. How can you leverage the knowledge of interdisciplinary groups with a shared interest to shed new light on existing data challenges? What are the key points of data intervention that are most important for each group of stakeholders?

Bibliography

Bakhshi, H., Freeman, A. and Higgs, P. (2013) A Dynamic Mapping of the Uk's Creative Industries. Nesta. Available at: **https://eprints.qut.edu.au/57251/1/57251.pdf.**

Bakhshi, H., McVittie, E. and Simmie, J. (2008) Creating Innovation: Do the creative industries support innovation in the wider economy?
Nesta. Available at: **https://media.nesta.org.uk/documents/creative-innovation.pdf.**

Connell, L. et al. (2022) 'Connecting, Collaborating, Creating: The Experiences of Creative Freelancers in Edinburgh in 2022'. Creative Edinburgh.
Available at: **https://doi.org/10.5281/ZENODO.7106424.**

Easton, E. and Beckett, B. (2021) Freelancers in the Creative Industries. Creative Industries Policy and Evidence Centre. Available at: **https://pec.ac.uk/policy-briefings/freelancers-in-the-creative-industries**

Jensen, M. et al. (2007) 'Forms of Knowledge and Modes of Innovation',
Research Policy, 36, pp. 680–693. Available at: **https://doi.org/10.1016/j.respol.2007.01.006.**

MyCake and The Audience Agency (2021) 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.
Available at: **https://www.gov.uk/government/publications/better-data-on-the-cultural-economy-scoping-study.**

OECD and Eurostat (2018) Oslo Manual 2018: Guidelines for Collecting, Reporting and Using Data on Innovation, 4th Edition, The Measurement of Scientific, Technological and Innovation Activities. OECD.
Available at: **https://www.oecd-ilibrary.org/science-and-technology/oslo-manual-2018_9789264304604-en.**

Walmsley, B. et al. (2022) Making data work: a scoping survey to develop a mixed-methods evaluation framework for culture. Leeds.
Available at: **https://www.culturalvalue.org.uk/our-work/making-data-work/.**

How to Cite this Work

Contact us

Email:
caitlin.mcdonald@ed.ac.uk
j.jordan@lboro.ac.uk
creativeinformatics@ed.ac.uk

Twitter:
@cmcd_phd; @JennieTweetz; @CreateInf

Editorial Design by Alix @Lunncreative

Illustrations by Elspeth Murray