

# Could open data help Arts and Culture listings?

**A discovery report.**

**By:**

Rob Ashelford, Nesta

Luke Piper The Satori Lab

Ben Proctor The Satori Lab

Esko Reinikainen, The Satori Lab

Amy Richards, Nesta

**nesta**

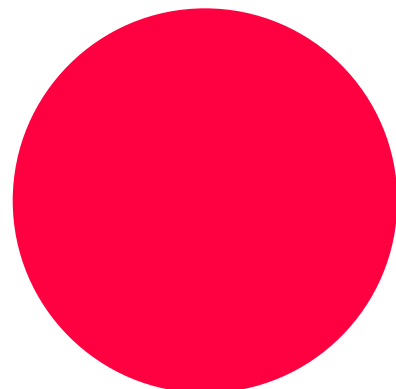
**satori**  
culture hackers **lab**



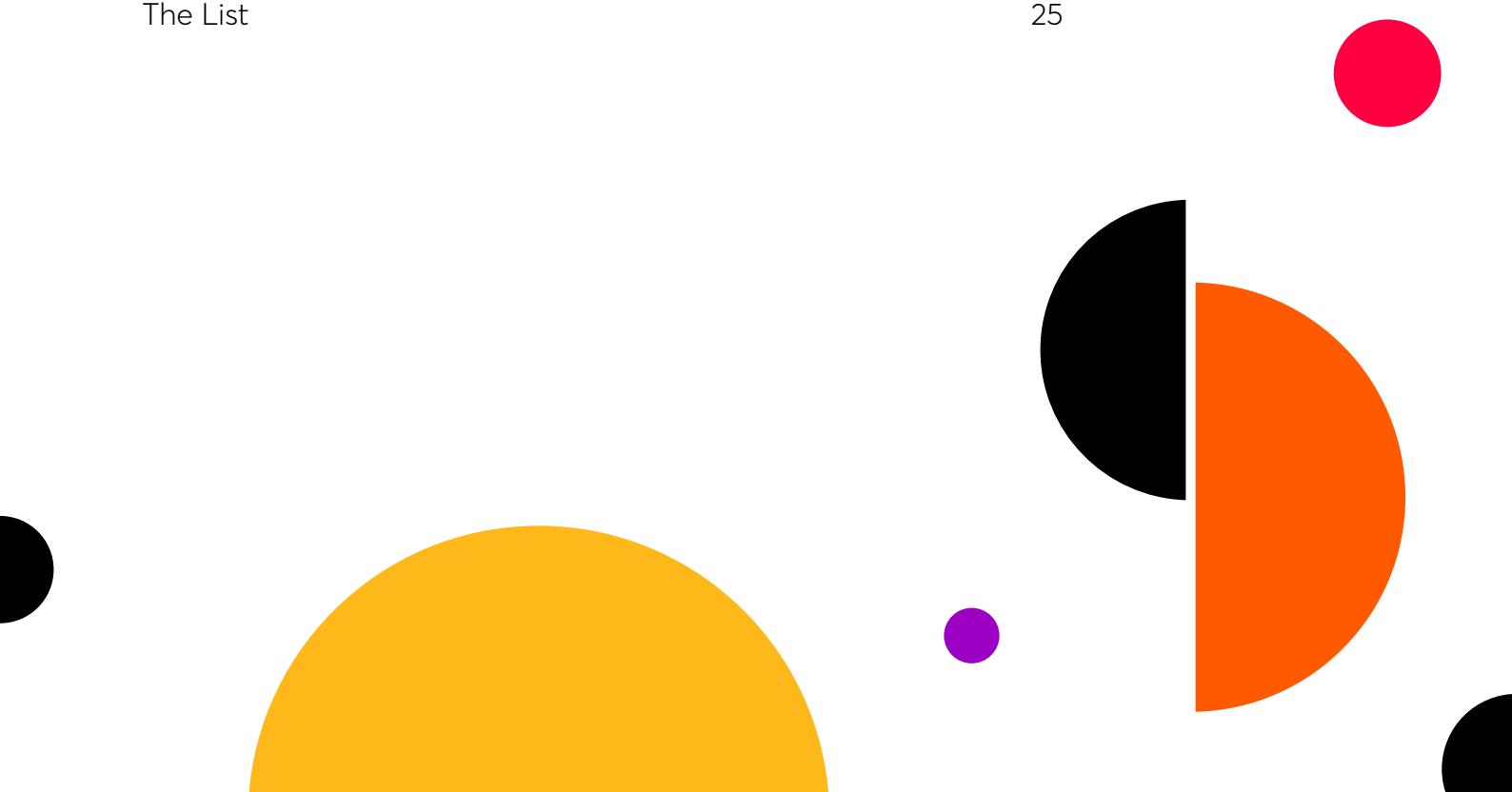
**CARDIFF**

# Contents

<b>Executive Summary</b>	<b>4</b>
<b>Background</b>	<b>5</b>
The starting problem	5
The approach	6
<b>An overview of the listings ecosystem</b>	<b>7</b>
Data flows	7
Data producers and data users	8
A simple data flow	8
Events data flowing to the theatre	8
Events data flowing to the listings magazine	9
Events data flowing to the audience	9
Data flow characteristics	10
Many hands	10
Multiple records	10
What audiences need	10
What venues need	11
Inefficiencies	12
Box office and ticketing systems	12
<b>Data infrastructure</b>	<b>13</b>
Data standards	13
Arts and culture listings standards	14
APIs	14
Scraping	15
Open data	15
The Data Spectrum	16
The listings pipeline	16
<b>The problem space as we now understand it</b>	<b>17</b>



<b>Recommendations</b>	<b>18</b>
For The Arts and Cultural Sector as a Whole	18
Use Schema.Org	18
For listings publishers	19
Listings publishers include festival organisers, venues and potentially artists, promoters and producers with their own listings sites	19
When commissioning websites specify an API and allow third parties access	19
Work with artists to reduce the inefficiencies of providing you with data	19
Commit to providing listings under an open licence	19
For policy makers	20
Interventions to improve data and digital literacy	20
Investment in data infrastructure work	20
<b>Acknowledgements</b>	<b>21</b>
<b>Appendices</b>	<b>22</b>
Appendix One: A worked example of data flows	22
Appendix two: Case studies	23
OpenActive	23
Edinburgh Festivals	23
Schema.org	24
Gigs.guide	24
The List	25



# Executive Summary

**We undertook a discovery project to better understand the problem space around arts and culture listings.**

**We discovered:**

- There are costs and inefficiencies along the data pipeline from the artist to the audience.
- There is no commonly or widely used data standard for publishing arts and culture listings.
- Reaching audiences is sometimes a problem for venues.
- Discovering relevant events is sufficient of a problem that a number of commercial services have been developed to help audiences discover events.
- Open data is not well understood in the sector and does not form a normal part of the discourse. In fact, most of the data publishers we spoke to as part of this discovery want to retain control over who can access and share their events listings data. Some organisations believe that they are offering Open Data but, in fact, provide Shared Data.
- There is evidence that where data is published against clear standards and via APIs that this will enable new and innovative services.
- There are examples of services and companies that are seeking to address aspects of this problem space.

**As a result of our discovery work we make the following recommendations:**

- For The Arts and Cultural Sector as a Whole
  - Agree a common standard for what a listing is, maintain it and use it
  - Use Schema.Org
- For listings publishers
  - When commissioning websites specify an API and allow third parties access
  - Work with artists to reduce the inefficiencies of providing you with data
  - Commit to providing listings under an open licence
- For policy makers
  - Invest in improving data and digital literacy in the sector
  - Invest in data infrastructure work across the sector

# Background

## The starting problem

This project started with an apparently simple question:

**"Can open data help to make listings more efficient whilst reaching new audiences?"**

Nesta approached The Satori Lab with a thesis.

They said:

- the current system for publishing arts and cultural events listings is broken
- there is a lack of common standards and a lack of technical competence or will around the use and potential of data across cultural institutions
- event listings are not published as open data and this creates barriers to innovation, it also reduces opportunities available in an otherwise data driven world.
- current events listing and publication relies on too many manual steps and is, accordingly, labour intensive, expensive and non interoperable
- this in turn acts as a barrier for small, local or innovative services to enter the listings market and makes it harder for potential audience members to find out what events are actually taking place

Nesta's starting point was that an effective solution to this problem would involve the publishing of events listings in open formats against commonly agreed standards.

This thesis described a problem space that is multi-faceted and there are many aspects of the space that are uncertain. Nesta partnered with The Satori Lab on a rapid discovery project. The aim of the discovery was to reduce as much uncertainty as possible and to build an evidence base for future action and further investment.

# Background

## The approach

A team, formed from Nesta and The Satori Lab staff, undertook a range of actions:

- we organised two workshops in Cardiff bringing together a range of stakeholders from the arts and culture sectors in South Wales
- we undertook desktop research into the problem space in Wales and other parts of the world
- members of the team undertook a series of semi-structured interviews with stakeholders with relevant experience from across the UK and other parts of the world

We present this report which summarises:

- what we found out about the problem space
- examples of organisations that are already tackling some of the issues we discovered in the problem space
- some next steps which, it seemed to us, would be helpful in addressing aspects of the problem space

# An overview of the listings ecosystem

## Data flows

Because we brought together stakeholders with expertise in different parts of the listings ecosystem we were able to gain a good understanding of the data flows between organisations and individuals.

Ultimately the purpose of arts and culture listings is to enable people who would like to attend a particular event to find out about that event and, if appropriate, book.

Audiences find out about events via a range of sources such as:

- the marketing channels of a venue that they prefer or have visited before
- the marketing channels of a ticketing or event marketing service
- the marketing channels of an artist
- relevant news or entertainment media
- friends and family
- Other advertising

The final event listing typically contains data that originated with the artist (or the producer) combined with data that originated with the venue.

# An overview of the listings ecosystem

## Data producers and data users

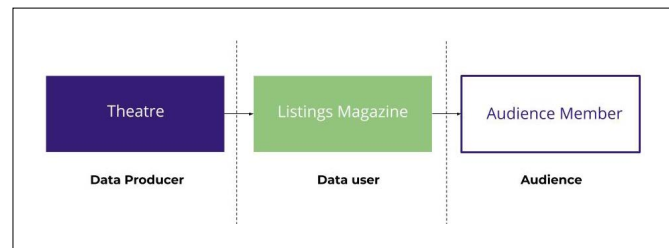
It can be helpful to think about data producers who own or control events data and data users who want to access this data to create services for audiences.

In this example the theatre is a data producer: it holds and controls event information.

The listings magazine is a data user: it provides a service to its readers but it relies on the data from the theatre to deliver this service.

In Arts and Culture listings the data flows are often much more complicated.

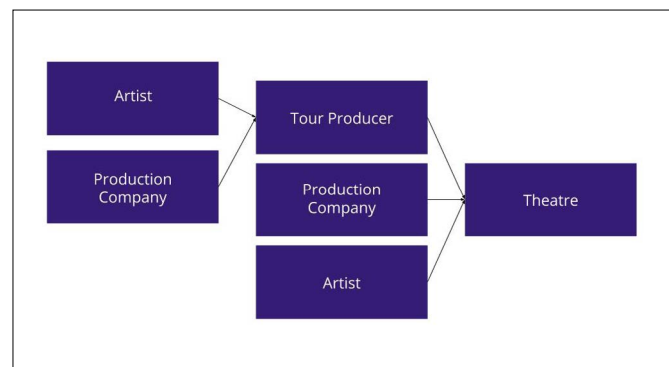
### A simple data flow



## Events data flowing to the theatre

In this example the event data that the theatre controls (for a season) may have originated from several sources. The theatre asked a production company for data about their dance production, a tour promoter for data about their show with a national TV tie-in and a local artist for data about their music set.

When we talk about data here we are including all of the information and content that someone might find useful when thinking about whether to book.



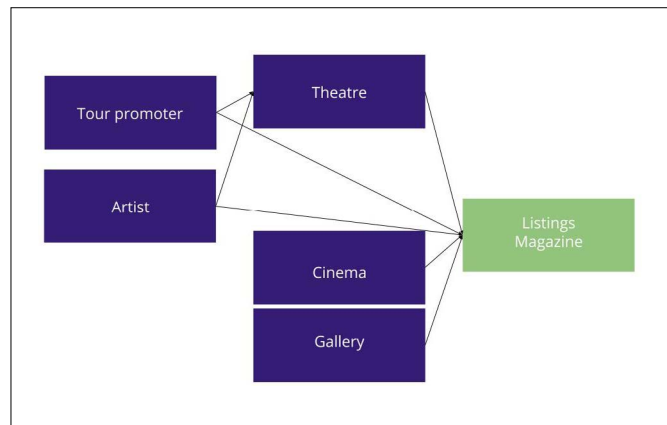


# An overview of the listings ecosystem

## Events data flowing to the listings magazine

The listings magazine receives listings data from a range of venues. In this example a cinema, gallery as well as the theatre provide data.

The magazine also receives data from the tour promoter and the local artist directly. This duplicates much of the information that has been sent to the theatre (and will come from there to the magazine) but it might not be identical.

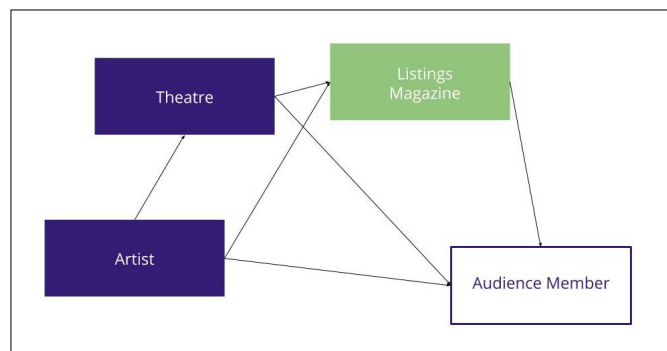


## Events data flowing to the audience

Audience members access details of events from multiple sources.

In this example the audience reads the listings magazine as well as visit the website of the theatre and receiving an email directly from the artist. The artist also provides their data to the theatre and to the magazine. The theatre also provides information to the listings magazine.

Overall this builds up to a complicated data ecosystem. Appendix One describes this from the perspective of the artist.



# An overview of the listings ecosystem

## Data flow characteristics

### Many hands

The data that comprise listings may pass through several hands from the artist planning her tour to the audience member finding out about it. In South Wales these hand-offs typically require manual intervention. Commonly information has to be re-keyed as it passes between stages in the process. Re-keying data means the risk of errors is higher, it causes delays and increases cost.

### Multiple records

This process leads to multiple listings of the same event on different platforms. Each platform may have different data requirements and limitations and human error may introduce variation in the way performances, artists and venues are described. Competition for ranking on search engines means different sites can describe the same event in different ways.

As one listings expert we spoke to described it:

events data, as you probably already realised is pretty nasty, unpleasant stuff, it has to be collected from a huge different range of providers, all of them use entirely different systems none of which talk to each other, it changes all the time, as soon as an event's gone then in many respects, it's actually kind of useless. And it changes incredibly quickly if a singer has a sore throat. [for example]

## What audiences need

There was agreement in our workshops with stakeholders from the arts and culture sector that audiences (or potential audiences) are missing out on events because the listings ecosystem is not working effectively.

It is hard to measure directly the degree to which this is happening. It is also likely to be strongly affected by a range of other factors: taste, social and cultural capital, age, wealth, ability to use technology and so on.

We have spoken to a number of companies that have business models predicated on making it easier for audiences to find events. This suggests that there is a problem for at least some people.

Several stakeholders suggested that listings is only one of a range of datapoint that people need to decide whether they will attend a performance. Other significant datapoint include transport costs and options, recommendations and reviews, the speed and efficiency of the purchase transaction and so on.

# An overview of the listings ecosystem

## What venues need

One might imagine that if there are potential audience members out there who struggle to find arts and cultural events that they would like to attend that this would be a problem for the venues that offer the events.

In our workshops we found that the situation is more complicated. Venues we spoke to sometimes have problems marketing certain types of performance or show but most venues will offer a mixed programme. They will offer shows that they are confident will sell well alongside performances that they believe will be harder to fill. So many venues have a more narrow problem: reaching audiences for specific events within their programmes.

Of course every venue is different and each venue has a distinct approach and set of problems. Most stakeholders we spoke to agreed that venues do want to get their listings out to a wider audience but this was rarely identified as the most significant marketing issue. Issues like managing dynamic pricing or understanding what other significant events have been scheduled were more likely to be raised before reaching wider audiences.

One stakeholder told us:

marketeers [want more complex] ways of discounting and pricing shows and different things like that

# An overview of the listings ecosystem

## Inefficiencies

There was widespread agreement existing arts and culture data flows are inefficient. Many people reported having to rekey data into new systems on a regular basis as well as passing data in unstructured formats (press releases, emails etc). The same data has to pass to multiple systems and this process is rarely automated.

It was notable that the people and organisations that are most likely to experience these inefficiencies are those least likely to have power in the system. Artists and production companies must repeatedly complete online forms and submit content by email against unique templates if they want their work to be promoted.

Within the marketing chains the people who are rekeying information tend to be the more junior employees.

## Box office and ticketing systems

Venues typically use third party services to handle the purchase of tickets. The suppliers that we spoke to emphasised that they served the needs of their customer base (their customers being the sales or marketing functions within venues). Any changes or developments within these services would be developed in response to changing customer demands.

# Data infrastructure

This discovery project was prompted by an idea that open data could help make listings more efficient and find new audiences. It is worth spending a few moments explaining the principle of how that might work.

## Data standards

Listings can be described in a wide range of ways. Take something as simple as the date of a performance. It can be written in any of these ways:

- 2 April 2019
- April 2, 2019
- 04-02-19
- 02/04/2019
- 04/02/2019
- 4/2/19

A data standard would lay out the way in which dates of performances will be provided (for example DD/MM/YYYY). Then anyone providing listings using that data standard would always publish their dates in that way (02/04/2019). A data standard lays out not just how dates are formatted but also how venues are described, how events are described and so on. A US Government agency (USGS) describes data standards this way<sup>1</sup>:

Using standards makes using things easier. For example, let's say you need a AAA battery for your flashlight. You don't need to worry about the make of the battery, since all AAA batteries are the same size - because they are produced to a standard. You don't need to worry about getting a specific brand of AAA battery, since all AAA batteries will work in your flashlight.

Without data standards moving data between people and organisations can be slow and inefficient. Any data publisher is free to define its own data standard. Data standards typically become more useful when a single standard is used by many publishers who collaborate on developing and improving the standard. This approach is known as an open data standard. The Open Data Institute has published a guide to open data standards which is free to access<sup>2</sup>.

Publishing listings that comply with an agreed standard makes it much easier for services that depend on the listings.

<sup>1</sup> <https://www.usgs.gov/products/data-and-tools/data-management/data-standards>

<sup>2</sup> <http://standards.theodi.org>

# Data infrastructure

## Arts and culture listings standards

Services that take listings data from several places and create new offerings based on the data understand the importance of data standards. Attempts have been made to encourage the adoption of a data standard for Arts and Culture listings. For example the IVES standard (though this is no longer being maintained).

## APIs

An API is a way to pass data from one system to another. Unlike, for example, moving a complete set of listings as a spreadsheet, an API allows another computer to ask a series of questions like:

- what events do you have planned for 02/04/2019?
- what music performances do you have planned in April?
- what are the full details of the Open Mic Night on 02/04/2019?

APIs are used in the Arts and Culture sectors.

We found that many of the commonly used box office systems will provide an API if clients request the service.

The List provides listings data via an API<sup>3</sup>, as does CultureHosts<sup>4</sup> and the Edinburgh Festivals offer an API providing details on all events across 11 festivals in the city<sup>5</sup>.

<sup>3</sup> <https://api.list.co.uk/>

<sup>4</sup> <https://api-docs.culturehosts.co.uk/>

<sup>5</sup> <https://api.edinburghfestivalcity.com/>

# Data infrastructure

## Scraping

If an organisation publishes listings on web pages but does not provide an API third party organisations can run applications that read the web pages and turn the content into data.

This process is often described as “web scraping” or just “scraping”.

We are aware of several services that use web scraping to some degree.

Scraping is problematic for several reasons:

- web pages are often not well structured so errors and omissions can be made in interpreting the content
- web pages are, in general, static so changes to listings: updates, cancellations, sold-out notices and so on may not find their way to the scraping organisation easily
- the rights over the use of the data may be unclear or may even forbid its use in this way which creates a risk for the scraping organisation

## Open Data

Open data is data that’s available to everyone to access, use and share<sup>6</sup>.

When data, including listings data, is published there is a copyright in the data. People and organisations who wish to take the data and use it to build new services must ensure they have the consent of the copyright owner.

To make data open, the copyright owner must provide a licence with the data that permits anyone to use the data for any purpose. They can require the user to acknowledge the source of the data but any more restrictions than this prevent the data from being free for everyone to access, use and share.

A widely used open licence is the [Creative Commons Attribution Licence](#).

If someone has to register or ask permission in order to access the data, that would not normally meet the standard of being available for everyone to access, use and share.

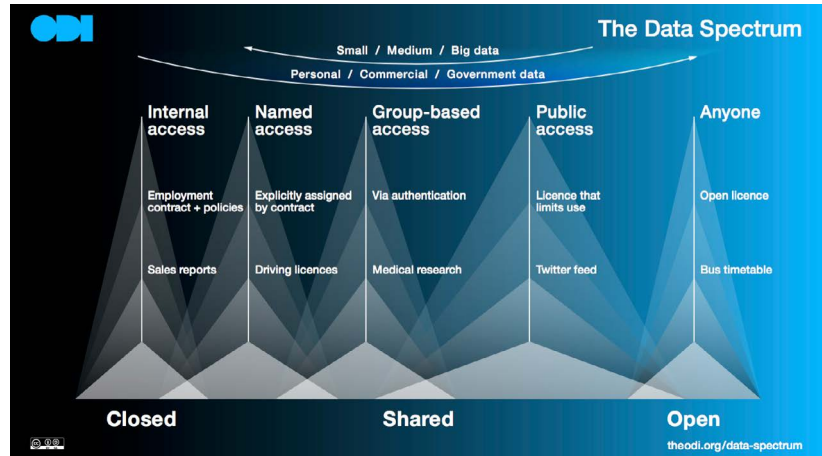
There is a range of different restrictions on the availability and use of data ranging from closed, through shared, to open. The ODI describes this as the Data Spectrum<sup>7</sup>.

<sup>6</sup> <https://theodi.org/article/what-is-open-data-and-why-should-we-care/>

<sup>7</sup> <https://theodi.org/about-the-odi/the-data-spectrum/>

# Data infrastructure

## The Data Spectrum



We did not find any examples of arts and culture listings being made available as open data, although there are a few examples of shared data.

## The listings pipeline

Listings are typically compiled from more than one source. Each of those sources of data potentially has copyright in the data that they provide.

Accordingly for the final publisher of the listing (for example the venue) to be able to apply an open licence to their listings, they would need to ensure that the data they received (for example from an artist) was provided under an open licence.



# The problem space as we now understand it

There are a number of widely recognised problems around the publishing of and use of arts and culture listings data. The arts and culture listings data ecosystem is complex and multi-faceted. Different stakeholders within the ecosystem have different problems and different perspectives on the problem space.

- There are costs and inefficiencies along the data pipeline from the artist to the audience.
- There is no commonly or widely used data standard for publishing arts and culture listings.
- Reaching audiences is sometimes a problem for venues.
- Discovering relevant events is sufficient of a problem that a number of commercial services have been developed to help audiences discover events.
- Open data is not well understood in the sector and does not form a normal part of the discourse. In fact, most of the data publishers we spoke to as part of this discovery want to retain control over who can access and share their events listings data. Some organisations believe that they are offering Open Data but, in fact, provide Shared Data
- There is evidence that where data is published against clear standards and via APIs that this will enable new and innovative services.
- There are examples of services and companies that are seeking to address aspects of this problem space.

# Recommendations

We believe that there is an opportunity to address a number of issues in the problem space but it will require action by a wide range of stakeholders. Luckily many of these interventions will be low cost and will provide benefits in and of themselves. In English local government the idea of “fixing the plumbing” has become widely used. We think that the challenge here is a similar “plumbing” problem. Like plumbing in buildings it is hidden and rarely considered (unless it goes wrong) but it is very important to the future health and success of the sector.

As more organisations make these interventions the collective benefit is likely to be very significant.

## For The Arts and Cultural Sector as a Whole

Below, we list specific actions that different actors in the sector can take to improve the way in which this data is published and used. There is, however, one recommendation that applies to the whole sector.

### **Agree a common standard for what a listing is, maintain it and use it.**

It is (comparatively) easy to define a data standard but a standard is only useful when it is widely used. Where a community of organisations forms around a standard, they make it more useful by creating new versions that are better suited to that community and they make it more widely accepted by using it themselves.

### **Use Schema.Org**

When listing events on web pages publishers can improve the visibility and usefulness of their listings by marking up the listings based on Schema.Org. This requires a degree of technical understanding and it should be factored in when commissioning new websites or when redesigning sites.

# Recommendations

## For listings publishers

Listings publishers include festival organisers, venues and potentially artists, promoters and producers with their own listings sites.

### **When commissioning websites specify an API and allow third parties access**

The problem of discovering events that one might like to attend is commonly identified as a problem. One solution: of creating new listings services is also a commonly proposed solution.

For example: Cardiff City Council recently commissioned a study into the Music Ecosystem in the city. This made a large series of recommendations including "Create a gig listing platform online that is inclusive and representative of the wider music scene in Cardiff" (recommendation 10.2)<sup>8</sup>

If the City Council takes up this recommendation it will have a range of options over how it might approach the design of the platform. Based on our findings in discovery we would strongly encourage an approach that encourages and enables bands, venues and other listings publishers to publish their data via an open API and against a common standard.

### **Work with artists to reduce the inefficiencies of providing you with data**

We have heard that many venues ask for data in the format and structure that is most convenient for them. This means that artists and producers have to provide what is broadly the same information in slightly different ways to each venue.

Listings publishers could offer to harvest data from an API provided by the artist or publisher. This would mean the artist could create a single record of their work and share it with multiple venues.

### **Commit to providing listings under an open licence**

Though we found that listings providers worry about losing control of their listings if they are released under an open licence, we believe that the potential benefits to audiences will outweigh any risks to the venue.

Publishing listings under an open licence will give developers confidence that they can use the data without legal risk. This will encourage the development of new and innovative services making it easier for audiences to find events that they are interested in.

In order to be able to publish listings data under an open licence, listings publishers will need to understand the restrictions placed on the data provided to them by artists and producers. Listings publishers should encourage artists and producers to provide their data under open licences to improve the efficiency of this process.

<sup>8</sup> <http://cardiff.moderngov.co.uk/documents/s29331/Cabinet%2018%20April%202019%20Music%20Ecosystem%20study%20App%201a.pdf>

# Recommendations

## For policy makers

### Interventions to improve data and digital literacy

Stakeholders that we spoke to in all parts of the arts and culture ecosystem agreed that digital and data literacy across the sector is low.

This is, arguably, one of the root causes of the whole problem space. The Arts and Culture sector is full of passionate, talented, skilled people. The fact that the sector has not tackled these problems successfully itself suggests that there are not enough people with the relevant fluency with the language and concepts of data infrastructure.

Policy makers should create interventions to encourage people across the sector to develop their digital and data skills.

The sort of interventions that we recommend include:

- training and development opportunities focused around digital and data, and particular open data. These should support not simply marketing professionals but all staff working in arts and culture.
- supporting and promoting data standards and encouraging different types of organisation to contribute to a data standards community

### Investment in data infrastructure work

Participants in the workshops often framed this problem around “listings services”. Even when talking about ideas such as data standards and APIs there was a tacit assumption that the problem will be solved by someone creating a single, central service to which all organisations will contribute.

We heard of several situations where there had been attempts to create central listing services. In some cases funded arts and culture organisations had been required to provide listings data to these services. Several of these services have folded when central funding was withdrawn.

We believe that a sustainable model will rely on a vibrant mix of companies, individuals and organisations innovating around agreed standards. A sustainable solution will emerge from a wider understanding of and investment in data infrastructure within each data publisher.

There are helpful investments that could be made to stimulate the ecosystem.

OpenActive provides an interesting model for other policy makers to consider. The project aims to build up the ecosystem in stages, finding effective and reciprocal relationships between data publishers and data users. These sorts of investments would potentially be helpful in the arts and culture sector.

The level of investment required is substantial. OpenActive, for example, represents several millions of pounds over several years. We believe that the opportunity for audiences, for creators and for venues if substantial investment is made in open data is huge. Conversely the opportunity cost of a lack of investment in data infrastructure is now and will continue to be significant. Opportunity costs are hidden costs. In this report we have tried to bring these hidden costs out into full view. It is for the sector to decide how to address them.

# Acknowledgements

We would like to thank all of the people who gave up their time to attend workshops or be interviewed as part of this project. There was a lot of generosity towards the team and a great deal of interest in the project overall. Many people are already trying to address aspects of the problem space but, as far as we can see, there isn't an organisation trying to address the problem at the whole system level.

We hope that this report might encourage some individuals or organisations to take a wider, systemic approach to tackling the problem. If our experience is anything to go by, they will be pushing at an open door.

# Appendices

## Appendix One: A worked example of data flows

Nia is a singer songwriter. She has recorded a new album and plans a tour to promote it.

Nia (or, potentially a promoter) makes arrangements with a series of venues to appear. The listing for each performance contains information about Nia, her style of music, how her music has been received, photographs from past performances etc. These data originate with Nia.

The listing for each performance also contains information about start (and end) time, location, cost, accessibility and how to purchase tickets. These data typically originate with the venue.

Once dates are agreed, Nia produces a list of her performances at all venues. This will be targeted at her fans and followers.

Nia sends details of her event to a listings "what's on service"

Nia sends details of her event to a listings "what's on" service.

Each venue promotes Nia's performance alongside other musical performances they plan within the current programme.

Different people find out in Nia's event from different sources and then go on to book.

In reality the data flows are more complicated because each artist may provide data to many venues and each venue relies on data from many artists. Different venues have different approaches to marketing channels: some may have websites, newsletters, paper programmes and advertising campaigns whereas others might focus on Facebook alone.

Venues may also provide details to the same listings services as the artists. Members of the public share details of events on social media and social media events.

# Appendices

## Appendix two: Case studies

Though we did not find any open data examples in Arts and Culture listings we did find projects with lessons that may be applicable to the open data question.

### OpenActive

Sport England and The Open Data Institute are working on a partnership project addressing a similar problem space: participation in sports and fitness events. The programme is called OpenActive.

OpenActive has approached the problem by bringing together a community of organisations wanting to use listings of sports and fitness events alongside organisations with the potential to publish these listings.

A significant step forward for the project came when Public Health England started using OpenActive data as part of the England-wide [Change for Life programme](#). This creates a strong incentive for groups and organisations to publish their data in line with the OpenActive standards: if they do their listings will appear in the national Activity Finder.

### Edinburgh Festivals

The Edinburgh Festivals created an API for their listings data specifically to improve data flows to media organisations.

Though the Joint Festivals API is not open it is free for organisations to use once they register (there is an additional registration and approval process for access to the Fringe data specifically).

Third parties have created innovative new services that take the listings data from the Joint Festivals API and combine it with other datasets. For example [Plan my Fringe](#) takes the listings data, combines it with geographical and travel data and provides a service that allows people to plan their visit to the festival based on their preferences and the distances and travel times between venues.

[This blog post lists a wide range of innovative services that have been enabled by the Joint Festivals API.](#)

# Appendices

## Schema.org

Google displays event data in search results. It typically obtains this data from the web pages of venues and producers.

Google promotes the use of a customisable data standard (Schema.org) and it provides guidance to publishers on how they can mark-up their events with against this standard<sup>9</sup>.

OpenActive has based its data standard on Schema.Org. Culture Host provides data via an API (for which registration is required) and its data standard is based on Schema.Org.

## Gigs.guide

Gigs.guide helps people find live events that match their preferences when they travel. For example, offering inspiration for trips based on where the user's favourite bands are playing.

The service works by collecting data through feeds and scraping from the major ticketing companies in Europe and North America. They then clean and organise the data, geotagging venues and adding more information about the artists. The Gigs.guide database currently maintains some 400 thousand live events at any one time.

The company's value proposition to the venues, organizers and ticketing companies is based on the opportunity to expose their events to a broader audience of travellers, at zero cost and with very little effort on their side. All that is required is reliable access to basic information about which events are planned.

However it is not all plain sailing. Reaching venues, organisers and ticketing companies takes time and many are unwilling to participate. When they do, the quality of data received is often so bad it requires extensive rework to be made consistent across different sources and ultimately usable at all.

---

<sup>9</sup> <https://developers.google.com/search/docs/data-types/event>



# Appendices

## The List

The List provides the most comprehensive set of live events data across the UK.

The service uses API or direct feeds from venues, box office systems and promoters and events are also uploaded directly to the website.

The List recognise new and innovative uses for the data they collect having developed predictive intelligence around the demand for public transport use at the time of events being held.

The company has also become expert in the field of archiving. Realising that by never deleting an event they hold a very extensive cultural record for the UK.

Collecting the data this way doesn't come without problems, inconsistent data schemas and duplication mean there is a lot of further processing still required to make the data usable.