Designing for Audience Engagement
Exploring the Use of Online Metrics in the GLAM Sector

Ellen Charlesworth
Durham University, UK

Claire Warwick
Durham University, UK

Leonardo Impett
University of Cambridge, UK

Andrew M. Beresford
Durham University, UK

Abstract Many heritage and cultural professionals have emphasised the ability of online content to reach beyond museums’ traditional audiences, yet one of the largest surveys to date shows no significant change in the demographic breakdown of online and on-site visitors. This paper aims to investigate the discrepancy between the data and the narrative, using two case studies to illustrate the pitfalls of using common social media metrics such as ‘views’, ‘likes’, and ‘shares’. Drawing from the authors’ previous sector-wide study, the paper highlights what shapes these numbers, and questions how useful they are as a measure of ‘audience engagement’. Finally, it looks to computational methods to explore nuanced alternatives that could supersede these metrics in the coming decade.


Summary 1 Introduction. – 2 Funding Pressures. – 3 The Limitations of Engagement Metrics. – 4 Case Studies. – 4.1 The British Museum. – 4.2 The Linen Hall Library. – 5 Alternatives. – 5.1 Sentiment Analysis and Topic Modelling. – 5.2 Network Analysis. – 6 Conclusions.
1 Introduction

The past three years have seen a dramatic rise in the number of gallery, library, archive, and museum (GLAM) organisations uploading resources online. Initially seen as a way to mitigate the impact of the COVID-19 pandemic and meet a perceived demand for cultural content, many organisations were able to pivot to digital by uploading previously digitised resources to social media (NEMO 2020). This increased online output has led to an enormous range of free-to-access content online. Yet, over the same period, museums also faced increased financial pressures resulting from the pandemic closures of 2020 and 2021 (Gnezdova, Osipov, Hriptulov 2022). This paper makes the argument that this financial crisis, which has exacerbated existing scrutiny of spending on culture and heritage, has shaped the way that we discuss digital adoption and online audiences.

During the COVID-19 pandemic many GLAM professionals argued that online content could transform audience engagement and reach beyond traditional audiences (Noehrer et al. 2021; Samaroudi, Echavarria, Perry 2020). This idea has been widespread since museums first started launching initiatives on the web. As Avenier summarised,

> It is now generally acknowledged that museums have gained significantly from the giant technological strides associated with the growth of the Internet. Nobody indeed disputes that the widespread dissemination of information and knowledge about museum collections is an added step towards the democratization of culture. (Avenier 1999, 31)

Yet, this narrative has been heavily criticised. In 2011, the demographic analysis performed by Culture24 led the group to question whether the cultural sector is in fact attracting new audiences online, as has traditionally been assumed, or is simply engaging with the same audiences that they interact with offline. (Finnis, Chan, Clements 2011, 22)

A decade later, the Cultural Participation Monitor – a project which included a series of surveys with over 6,000 correspondents – corroborated this finding, and found that those who engaged online where “more likely to have been regular in-person arts attenders”, with “more than half of audiences engaging with Museums and
Heritage online are aged 55+” (Walmsley et al. 2022). Conducted by the Audience Agency and the Centre for Cultural Value, this work has highlighted the continuing discrepancy between the rhetoric of broadening access and organisations’ actual ability to reach new audiences (Audience Agency 2021).

Gathering data to contextualise why this narrative is still so prevalent and identifying the biases in our discussions of online content is an integral part of establishing a sustainable design process and fostering eventual improvement. However, the act of measuring engagement is fraught with difficulty. Methods such as surveys and interviews are shaped by self-selection bias in respondents and do not reflect wider – often less engaged – audiences (Gran et al. 2019; Bethlehem 2010). There is a desperate need in the sector for large-scale quantitative metrics to create a benchmark against which GLAM professionals can evaluate survey results and their own anecdotal experience (NEMO 2020). This need has been repeatedly highlighted over the past decade which has seen government investment in the Culture Metrics and Culture Counts initiatives in the UK (Culture Metrics 2022; Arvanitis et al. 2016). The pandemic further highlighted this need with the Network of European Museum Organisations stating that a “sound metric to benchmark online visits is missing”; this resulted in new investment in inDICEs’ Europeana Impact Framework by the EU’s Horizon programme (inDICEs 2023; NEMO 2020, 3).

Yet taken in isolation, quantitative metrics have been found to alter how we conceive of value and are a simplification of a vast array of outcomes and motivations which are unique to each organisation (Espeland, Sauder 2016). This paper will therefore explore how quantitative metrics from social media platforms may be misleading and explore what kind of metrics may be more useful to GLAM organisations. Through an investigation of two online responses to the UK lockdowns in 2020, the paper will examine how institutional aims inform the way audiences are conceptualised and incorporated in online metrics commonly used by GLAM organisations.

The UK case studies – a national London-based museum and an independent library in Northern Ireland – have been chosen to illustrate the breadth of ways that organisations can engage with online audiences. These two examples were both identified from the authors’ larger study of 315 UK GLAM organisations, and have been chosen to highlight how the sector’s diverse online landscape is poorly captured by simple quantitative measures and social media metrics (Charlesworth et al. 2023).

1 The study took a sample from the “Arts Council Accredited List of Museums” which, despite the title, includes a number of Libraries, Galleries, and Archives (Arts Council England 2021).
social media platforms, our analysis of social media metrics during the study highlighted how misleading the measures can be. By exploring how content is disseminated on platforms such as YouTube and Facebook, we will explore whether these numbers reflect audience engagement, and make the argument that alternative methods such as network analysis, sentiment analysis, and topic modelling will provide more nuanced measures of engagement.

Beginning with a summary of the funding pressures that incentivise audience engagement, the paper explores why quantitative metrics, including those from social media, are increasingly important to the sector. It then goes on to explore whether commonly used online metrics, designed by and for social media platforms, are accurate measures of audience engagement and serve the needs of GLAM institutions. It details the limitations of commonly used measures of engagement – views, likes, shares, and comments – and debunks common misinterpretations of the figures. Finally, we ask whether these metrics are of use to GLAM institutions at all and explore potential alternatives that, with further research and development, could supersede the metrics available on social media platforms, to create a more nuanced method to evaluate audience engagement.

2 Funding Pressures

There is a broad range of literature in museum studies and cultural policy that explores the many factors that contributed to the increasing importance of engagement. Postmodernist thinking redefined museums’ role in society and their relationships with visitors (Speight 2010; Hooper-Greenhill 1992); this saw the adoption of the idea that museum experience is shaped by an individual’s prior knowledge and broader social context (Falk, Dierking 1995; 2000). Postcolonial critics, most prominently Edward Said, problematised the role of culture as an agent of empire, prompting a questioning of the role colonial power relations had on the interpretation of objects (Barringer, Flynn 1997; Said 1993). The impact that these discussions had on audience engagement is too large to cover in a single paper. Instead, this paper will focus on how engagement has been conceptualised by the UK government and funding bodies. This approach aims to contextualise why there is a prevailing emphasis on online content’s ability to broaden GLAM audiences, exploring how shifting funding priorities have shaped GLAM organisations’ goals and the ways in which they measure impact and engagement.

The majority of GLAM institutions in the UK are either directly supported by government institutions or are non-profit organisations reliant on substantial grants to stay open (Ballatore, Candlin 2022, 224). This financial arrangement has undoubtedly had some influence on
the organisational aims of these institutions, and the focus on broadening their appeal is partially motivated by an increasing pressure to justify public expenditure in the sector through tangible economic and social outcomes (Bailey-Ross 2014; Anderson 2013). This can be observed in the way museums’ organisational practices have been shaped by funder’s emphasis on evaluating ‘impact’ and ‘engagement’ (Aroles, Hassard, Hyde 2022); as Belfiore argues, ‘impact’ has become a proxy for the term ‘value’ and is indicative the way in which funding is increasingly framed as an ‘investment’, which expects to see ‘returns’ and ‘healthy dividends’ (2015, 101, 106). These returns may not be strictly financial and GLAM organisations are regularly conceptualised as vehicles for government policy. Black notes that where organisations have traditionally received public funds

there has been an expectation at both the national and local levels that all such bodies, including museums, will actively support relevant political initiatives. (Black 2010, 130; Newman, McLean, Urquhart 2005)

A political focus on ‘engagement’ can be traced back to 1990s, when there was growing concern at a perceived decline in public dialogue (Ashley 2014). This was considered symptomatic of declining levels of civic engagement there was a fear that citizens were not actively participating in their communities or in the democratic process (Adler, Goggin 2005, 236). Partially attributed to an increase in multiculturalism in the UK, the discussion saw growing demands for recognition of minority groups, and resultingly, a renewed focus within government on improving social integration and inclusion (Newman, McLean 2004). In this context, cultural participation was seen as a means to increase civic engagement: in part, by opening democratic public discussions to previously excluded voices; but also by fostering social inclusion and encouraging dialogue between communities (Long 2013; Putnam 1995). Engagement in this context is not an end, but a mechanism to enact political and societal change.

The rhetoric around engagement today highlights that cultural institutions continue to be seen as a means to encourage social inclusion, even if it is as agents of positive change and advocates of “social justice” (Ünsal 2019). Arts Council England, a major funder of libraries, museums, and galleries places an emphasis on “diversity […] and creating more pathways for a wider range of people to become part of the arts and culture sector” (2023a). Yet, the efficacy of using cultural organisations in this way is still contested. Culture alone cannot be expected to repair social exclusion an issue that has its roots in issues with economic integration, housing, and the welfare system (Newman, McLean, Urquhart 2005). Newman and McLean have problematised public investment in museums on the basis that they
improve social inclusion, noting that participation in cultural institutions is not the same as participation in culture (2004). In addition, despite a turn towards participatory practices (Simon 2010), the power structures and hierarchies of long-standing western institutions exacerbate existing asymmetrical power relations between communities, complicating GLAM organisations’ role as facilitators of open dialogue (Ashley 2014, 263; Schuch, Harden, Smith 2023, 313).

It is unsurprising that within this funding context, made more acute by the pandemic, so many GLAM professionals framed the digital shift as an opportunity to diversify audiences and reach beyond museums’ traditional visitors (Samaroudi, Echavarria, Perry 2020; Mantell, Turpin 2020). This is important in understanding the use of online platforms as the ability to diversify and grow audiences is a significant appeal of online content for most GLAM institutions (Noehrer et al. 2021, 2).

3 The Limitations of Engagement Metrics

Given the funding context discussed, it is clear that ‘audience engagement’ in the GLAM sector is one of the means through which institutions try to achieve social inclusion and civic engagement. It is these, rather than audience engagement itself, that we would ideally like to measure. Yet the ‘social inclusion’ generated through a GLAM organisation is difficult to quantify and almost impossible to capture at scale – an issue that similarly plagues measures of audience engagement. The ubiquitous audience survey has major limitations, especially online. To get a large sample is difficult, is often a labour-intensive process, and surveys are often designed for a specific project preventing comparison between institutions. In addition, as surveys are reliant on people voluntarily filling them out, there is a self-selection bias amongst respondents that distorts our understanding of general visitors’ views and demographic groups (Nuccio, Bertacchini 2021; Gran et al. 2019; Bethlehem 2010).

The Culture Metrics project launched in 2014 aimed to address these issues, involving cultural organisations in the process of establishing standardised metrics for the GLAM sector. A versatile survey was designed to work across multiple contexts; participants scored different aspects of an event or visit on a sliding scale. In describing the project, Arvanitis et al. (2016) noted the benefit of this approach:

> The use of standardised metrics across organisations potentially allows for comparisons among them and among different kinds of events and over time.

The Culture Metrics team developed a platform in partnership with Culture Counts which automated the analysis of the results, and in
2016, Arts Council England announced that using the Culture Counts framework would become “a mandatory requirement” for organisations that received over £250,000 a year (Gilmore, Glow, Johanson 2017; Arts Council England 2016, 2). This decision, which was later revoked, was met with vocal opposition. Phiddian et al. noted that the issue was not with the metrics themselves but in the way that they were being used as to assess artistic quality:

The categories are nuanced enough to provide usable feedback for practitioners and bureaucrats with the time and desire to think hard about what the numbers mean. But, they remain essentially marketing analytics rather than a window on artistic value. (Phiddian et al. 2017, 178)

What is a helpful tool for well-meaning practitioners can, in other hands, be used to ineffectually rank relative performance.

The Culture Counts example highlights the importance of context for any metric, and the goals of a project should inform the methods we use to evaluate success. As Phiddian et al. argue, the adoption of framework across all grant recipients would “imply a spurious homogeneity of purpose in the arts” (Phiddian et al. 2017, 178). As Glass notes, “the importance of data isn’t the data by itself. It’s the possibility of self-awareness and self-reflection that it brings to bear to make us better” (2015, 5). This partially explains the GLAM sector’s frequent use of surveys, which can give additional information about audiences’ motivations for visiting and answer why an experience was positive or negative (Villaespesa 2013).

Yet, the most accessible and widespread metrics available to GLAM organisations are simple quantitative measures provided by social media platforms. In the UK, it is more common for a likely for GLAM organisations to have a Facebook account than their own website (Charlesworth et al. 2023, 8). Larkin, Ballatore, and Miturova found that 77% of museums had a Facebook account, and 67% had Twitter (2023, 6). This high uptake was true across small and medium sized museums, making measures of social media engagement – such as ‘likes’, ‘views’, and ‘shares’ – an already widespread form of standardised metric.

As we conducted follow-up interviews to the study of 315 UK GLAM organisations, social media metrics were regularly raised when describing success, and digital teams were proud of their posts with high levels of social media engagement. This was tempered by an understanding that some posts – such as site closures and event advertisements – would not perform well, despite their utility to visitors. However, this nuance is rarely conveyed in the reporting of social media numbers. Annual reports across the GLAM sector regularly highlight social media successes, including both organisations used here
as case studies. The British Museum highlights that “Social media continues to enable us to reach audiences directly” and celebrates its “nearly six million followers across all platforms” with little elaboration (British Museums 2022, 6). The Linen Hall Library was similarly “pleased to report that it [their Instagram] reached 1,000 subscribers by year end” (Linen Hall Library 2020, 9). This kind of reporting is understandable as Arts Council England actively encourage organisations to use social media channels “to show how public funding enables your work and positively impacts the lives of your audience and your local community” (Arts Council England 2023b). Considering this tendency to equate high social media metrics with impact, it is worth examining exactly what these numbers represent and the context in which they were developed.

Social media metrics were designed with commercial interests in mind. The ways in which we interact with the internet are both facilitated and mediated by companies such as Meta and Google. It is for them and their paying users – that is advertisers – that the ranking of content based on ‘click throughs’, ‘likes’, and ‘shares’ were developed (Gillespie 2010). In a study of social media advertising, Voorveld et al. define social media engagement as “the emotional, intuitive experiences or perceptions” that occur on social media platforms, and cover a broad range of motivations including “satisfying the need to find useful information, fill empty moments, and do or share something with others” (2018, 40). Pertinently for this discussion, Voorveld et al. conclude that “information might increasingly be a ‘by-product’ of social media use rather than a central feature” where exchanging information “is a means to an end rather than a goal in itself” (2018, 51).

This paper cannot comprehensively cover the vast body literature dedicated to identifying the most effective advertising, persuasion, and e-commerce strategies based on these metrics (Rautela, Sharma, Virani 2021). However, there is a consensus that the most widely shared content provokes emotions; alarmingly, Ji et al. found that Facebook posts that roused negative emotions were more likely to be shared than positive posts (2019). It is therefore questionable whether these same metrics accurately reflect the meaningful inter-community connections that GLAM institutions hope to facilitate. By borrowing quantitative metrics from social media platforms, GLAM institutions are adapting tools designed to commodify attention. Therefore, it is worth pointing out the obvious: social media engagement does not necessarily correlate with engagement in the sense it is used in the GLAM sector.

There are remaining issues even if a metric was found to accurately reflect audience engagement. If a metric was used during the development process to test projects and exhibitions, GLAM institutions may inadvertently optimise content to maximise the metric,
as opposed to maximising engagement. This issue is summarised in Goodhart’s law, which was succinctly described by Strathern, “When a measure becomes a target, it ceases to be a good measure” (1997, 308). The principle is most commonly observed in education, where students learn to pass the exam, not understand the material (Fire, Guestrin 2019). This could result in GLAM online content that has high levels of engagement on paper, but individuals may not have the intended experience. This issue can only be combatted with extensive user testing and verbal feedback, a process that should not be side-lined in favour of cheaper and more readily available quantitative metrics.

Despite the need for a sound benchmark to evaluate online audiences, this section has highlighted some of the problems that sector faces to reach this goal (NEMO 2020, 3). In the next section, we will explore how these problems emerge across two case studies. By investigating how different types of institutions are served by social media metrics, we highlight the types of engagement they capture and question how useful this is to GLAM organisations.

# 4 Case Studies

The utility of a metric is reliant on what an organisation is trying to learn by gathering that data. Therefore, in order to understand the breadth of ways an audience engagement metrics can be used, the case studies draw from the two dramatically different institutions. Informed by the authors’ previous large-scale analysis of GLAM organisations in the UK, the two examples – a huge museum, and a small library – have entirely different institutional aims (Charlesworth et al. 2023). Using these case studies, this section explores how institutional aims inform the way audiences are conceptualised and incorporated in the online design choices and digital strategies of GLAM organisations. Crucially, an in-depth analysis of the reception of two online exhibitions enables us to ask whether ‘click-throughs’, ‘likes’, ‘shares’, and ‘comments’ are useful metrics for evaluation.

---

2 Theses sizes are based on the guidelines produced by the Mapping Museums Project, which categorised organisations by the number of yearly visitors; “small (0-10,000 visits); medium (10,001-50,000 visits); large (50,001 to one million visits); and huge (over one million visits)” (Candlin et al. 2019, 57).
4.1 The British Museum

The first case study will explore how a single video uploaded by a huge national institution was received by audiences during the pandemic. In the study of 315 UK museums, the British Museum was the largest organisation – based on the annual visitor numbers – in the sample. It had by far the largest social media following and at the time of writing the British Museum has 1.76 million followers on Facebook, 2 million on Twitter, and 2.1 million followers on Instagram (Charlesworth et al. 2023, 10).

With an established online presence before the outbreak of COVID-19, the museum was able to adapt particularly quickly to the closure of its physical site on March 18, 2020 (British Museum 2020a). The first national lockdown in the UK – beginning on the 23rd of March with most restriction ending on the 4th of July – prolonged what was originally envisaged as a short temporary closure, and the museum would not reopen until August 27 (British Museum 2020c). On announcing its closure, the museum had already outlined its digital strategy going forward. The Director of the British Museum, Hartwig Fischer, stated,

We’ll be updating and adding to digital content during the period we’re closed to allow visitors to stay in touch with the Museum. We’ll share our collections, research and programmes in new ways that will not require a trip to the Museum. (Brown et al. 2021; British Museum 2020a)

Alongside his statement, the website was updated so that the homepage displayed a large banner detailing a list of digital activities. The page prompted visitors to ‘stay connected’ in multiple ways: it invited them to take a virtual tour on the website; look at the collections online; use the resources for schools; listen to the British Museum’s podcast; get in touch through social media; or explore the museums’ content on Google Street View and Google Arts and Culture.

Over the course of the first lockdown period, the website was regularly updated with new content – predominantly published within the framework of existing digital initiatives at the museum. Unfortunately, metrics such as the number of views, dwell time, and audience analytics of the website and Google Arts and Culture content are not publicly available. However, the reception of online content over time can be monitored in part on social media platforms, for which data is available through an application programming interface (API). In addition, basic metrics and the relative success of a piece of content is easily viewable on the platforms themselves. One the most successful pieces of content the British Museum produced, was an old video tour of an exhibition uploaded to YouTube in May 2020. This video
has been chosen as the focus of this case study, as it received large amounts of engagement on social media and remains the single most watched piece of content on the museum’s YouTube channel in 2023.

“Vikings Live: a tour from the British Museum” is an hour and a half long video that was originally created in 2014 and re-uploaded on May 27, 2020 (British Museum 2020b). Advances in technology since the video was filmed, mean that the production values – particularly the resolution and sound quality – are notably lower than more recent uploads on the museum’s YouTube channel. The presenters, Bettany Hughes and Michael Wood, take turns in interviewing key staff members and visiting experts who explain the process of setting up the exhibition as well as the history of the physical objects. These segments are interspersed with footage of historical re-enactments, questions from the public, and graphical overlays that show migration patterns and the location of archaeological sites.

Originally created for broadcast in cinemas, the video itself was not created for YouTube and therefore is atypical of the platform, which is dominated by individual content creators, not institutions (Xiang 2022). With a run time of 89 minutes, the duration of “Vikings Live” is much longer than the average YouTube video which is less than 15 minutes long (Rieder, Coromina, Matamoros-Fernández 2020; Che, Ip, Lin 2015). Yet despite content designed for other platforms often underperforming compared to “native” YouTube content, by March 2023 “Vikings Live” had been viewed over 8 million times (Rieder, Matamoros-Fernández, Coromina 2018; Arthurs, Dragoumpolou, Gandini 2018, 6). Does that make the “Vikings Live” video a success? By contextualising this figure, this section aims to highlight what influences the number of views a video might receive and investigate whether the metric has any use as an evaluative tool for researchers and GLAM professionals.

The number of views on YouTube can be accessed by researchers in two ways: either through YouTube’s application programming interface; or for small scale research, it is possible to manually check the view count displayed below the video. The first challenge this poses is that unless you are an administrator of the channel that uploaded the video, the number only reflects a snapshot of the number of views at the time it is checked. This means that researchers are unable to chart a video’s popularity in retrospect. Fortunately, over the course of 2020, a significant volume of GLAM digital content was preserved using the Internet Archive. As a non-profit organisation that aims to preserve websites and digital artifacts, their tool, the Wayback Machine, enables researchers to see snapshots of museum websites and social media taken throughout the pandemic period. Anyone is able to add a webpage to the archive, and between May 30, 2020 and March 14, 2023 “Vikings Live” was archived using the Wayback Machine over 70 times (British Museum 2020b). Although
not comprehensive, these archived versions of the YouTube page enable us to roughly reconstruct the views the video received in the three years since it was first uploaded [fig. 1a-b].

As figure 1a shows, “Vikings Live” was uploaded in the middle of the UK’s first national lockdown. In uploading a virtual tour near the beginning of the pandemic, the British Museum likely benefited from the influx of online visitors looking for cultural content. Larger organisations were better able to quickly adapt to lockdown conditions than their smaller counterparts and increase the number of videos
they were uploading to YouTube faster (Charlesworth et al. 2023, 13). ICOM found that large museums were more likely to provide digital services, while a UNESCO survey of over 50,000 museums, found that large museums – many of whom already had a bank of digital resources – saw the largest rise in visitor numbers to their websites over the first lockdown period (ICOM 2021, 18; UNESCO 2020). The British Museum was similarly able to quickly respond to demand for at-home content, utilising their pre-existing YouTube channel which already had an established follower base. Comparatively, organisations without existing audiences or staff expertise struggled to establish themselves in highly competitive online market (NEMO 2020; Arthurs, Drakopoulou, Gandini 2018). This provides the first indications that these metrics are not reflective of the content itself, but are heavily influenced by how, when, and by whom the content is uploaded.

Notably, despite the exceptional circumstances in which the video was uploaded to YouTube, its initial reception was in line with other videos on the British Museum’s channel. In the first three months after it was uploaded, the number of views remained below 500,000. This is not unusual, which is best illustrated when compared to the view count of “Pompeii Live from the British Museum”. This video, uploaded a week before “Vikings Live” on May 20, 2020, provides a useful benchmark for comparison. Filmed in 2013, it is remarkably similar to “Vikings Live”, sharing both a format and presenters. Yet in 2023, it still has less than 500,000 views and is only the 24th most popular video on the museum’s channel. Does this mean that “Vikings Live” is more engaging or a better video? By breaking down the number of views changed over time, it is possible to see why this kind of evaluation can be misleading.

When initially uploaded to YouTube, the two videos both performed similarly. On first being released, the number of views for both videos increased rapidly before slowing down over the summer of 2020. Taken in isolation this early spike in audience views could be assumed to be shaped by the larger pandemic audiences, however, Kamiyama and Muratahas identified this lognormal distribution of views is typical of YouTube videos (2019, 1103). This should not be falsely attributed to audience viewing habits or the quality of engagement the video inspires, instead, it is the result of the way YouTube’s recommendation algorithm prioritises recently uploaded videos (Gregersen, Ørmen 2021; van Es 2020). The recommendation algorithm suggests videos for users to watch next, and combined with their ranking of search results, heavily influences how many people see a video (Arthurs, Drakopoulou, Gandini 2018). The effect of these systems can be seen placing the British Museum’s videos metrics in a wider context [fig. 1b]. During the second lockdown period, “Pompeii Live” did not see any increase in the number of views, meaning that the mere presence of house-bound audiences during the lockdowns did not automatically result in larger audiences for GLAM institutions.
In contrast, the second lockdowns saw an enormous increase in views for “Vikings Live”. The incomplete nature of the Wayback Machine’s archive means that there is little data from this changed over November and December of 2020. However, the stark contrast between the Pompeii and Vikings exhibition is indicative of the way videos are popularised on YouTube. While the specifics of the recommendation and ranking algorithms of YouTube have not been made publicly available, there are two factors that most likely contribute to this effect. Firstly, at the start of the first pandemic, “Vikings Live” proved the slightly more popular video, making it more likely to be recommended to YouTube users – an effect that is compounded over time. The more views a video has, the more likely it is to be recommended, the more it is recommended the more likely it is to be viewed. If the video is regularly recommended but users do not click on it, over time it is recommended less often; this is partially why outrageous titles and staged picture thumbnails are so prevalent on the platform (van Es 2020; Shimono, Kakui, Yamasaki 2020).

The ranking of videos, both through the search and recommendation systems, massively affect the click-through rate on a new video; the first video in a list is almost six-times more likely to be clicked on than the last (Zhou, Khemmarat, Gao 2010). It is probable that these differences over time had contributed to the dramatically different viewing figures for “Pompeii Live” and “Vikings Live”.

Although there is a wealth of literature on these effects regarding viral videos, beauty content, and vlogs, it relatively rare for the videos of GLAM institutions’ to be affected in such a way. The “Vikings Live” video proved such an extreme outlier that it was had to be removed from calculated averages over the authors’ survey of 315 UK museums (Charlesworth et al. 2023, 10). This highlights how limited ‘views’ are as a tool for evaluation on social media platforms: not only is the figure not a reflection of the content’s quality, it also does not reliably reflect the type of content that will do well in the future.

It may be possible to attribute some of the popularity of “Vikings Live” to the subject matter. During the pandemic, Vikings were on the UK syllabus and a topic that was taught in primary schools (Department of Education 2013). Yet in a large study of what drives the number of views on videos, Zhou et al. found that the recommendation system had far more long-term impact than promoting a video or embedding it on other social media sites (2016, 6052). Therefore, it is unlikely that people looking for Viking education content or following links from the British Museum website directly caused the high number of views on “Vikings Live”. It is however possible that this was a contributing factor in the recommendation algorithm pushing “Vikings Live” and not “Pompeii Live”.

As ongoing work on the YouTube algorithms is not made public, we can only speculate as to why certain topics and videos are chosen,
and the details of how it functions are not understood by users, or by
the uploaders of videos. There is a wealth of speculation about the al-
gorithm published on the platform itself, and network of self-styled
algorithmic experts who sell advice on how to game the system have
emerged on the platform (Bishop 2020; 2018). However, even amongst
experts and specialists, the workings of the algorithm are a mystery
(Cunningham, Craig, Silver 2016). Even the British Museum, a huge
national institution with a dedicated digital team, has been unable
to optimise their videos for YouTube’s algorithm. In their video post-
ed on March 9, 2023, Nick Harris, a senior producer commented,

And if you could give the new videos a like when they come out,
we would really appreciate it as it massively helps with the algo-
rithm, which – to be honest – hasn’t been that kind to us of late.
(British Museum 2023)

Ultimately, the unpredictability of the algorithm, and its lack of cor-
relation to the quality of the content or social engagement makes it
useless for evaluation. The type of content museums are more likely
to make – irregularly uploaded, long-form content – is not optimised
for sharing on YouTube (van Es 2020). By breaking down the view
count of the “Vikings Live” video and comparing it to the similar vid-
eo “Pompeii Live”, we have shown how unpredictable metrics on so-
cial media platforms can be. Whether an individual video is popular
or not, is heavily influenced by factors beyond an organisation’s con-
trol, and as such, they cannot be used to compare between organi-
sations or exhibitions over time by themselves. This severely limits
their usefulness in evaluating how the design and curatorial deci-
sions have impacted the reception of the video.

4.2 The Linen Hall Library

In order to understand what constitutes a useful metric, we will use
another case study of an entirely different type of GLAM institution
to highlight how diverse the use cases for an audience engagement
metric would be. To this end, a smaller organisation was chosen from
the 315 museums in the authors’ large-scale study. The Linen Hall Li-
brary in Belfast has been chosen as it is both small and independent,
the most common type of organisation in the UK GLAM sector, mak-
ing it more representative of the sector as a whole. The library was
chosen ahead of the other organisations for two reasons. Firstly, the
library’s social media following is above average for a small GLAM or-
ganisation, but it is not so popular as to be an outlier whose methods
would be unapplicable to other organisations. Secondly, the library’s
website and social media presence were regularly archived on the
Wayback Machine over the course of the pandemic. Combined with their easily accessible annual reports, this makes tracking changes to their website and digital strategy easier than for other organisations of a similar size.

The Linen Hall Library has diversified its income streams and relies on memberships, shop sales, and grants (Linen Hall Library 2020, 13). The library’s funders include the Belfast City Council, and the Department for Communities in Northern Ireland, and as such, its commitment to the Northern Irish public is a central aspect of the organisation’s aims and strategy. Outreach and engagement feature heavily in the way projects are framed, and there is a particular focus on the local. Visitors to the gift shop are charmingly described as “lovers of all things literary and local”, but this seems equally apt for their wider audiences (Linen Hall Library 2019, 8). The people and histories in the site’s immediate vicinity are placed at the heart of events, workshops, outreach programme, corporate partnerships, media relations, and the collection itself, with “Irish & Local studies” books making up the majority of item requests. The library also frequently collaborated with other GLAM and educational institutions based in Northern Ireland: National Museums Northern Ireland, the Nerve Centre, The Public Record Office of Northern Ireland, Conflict Textiles, Belfast’s Grand Opera House, and the Conflict Archive on the Internet at Ulster University (Linen Hall Library 2019). The library is only one actor in a vibrant cultural ecosystem that aims to server the diverse communities of Belfast.

When the library was forced to close on March 17 2020, their immediate response was two-fold: firstly, by the end of March, groups that facilitated social interaction were moved online, including the English and Irish book club; and secondly, the library highlighted local artists through a series of online exhibitions on the library’s social media channels (Linen Hall Library 2019, 2). In contrast to the British Museum, this flurry of activity did not utilise old digital assets but sought to provide a space to meet for the library’s existing community. This is typical of smaller organisations’ responses to the pandemic, especially those with limited digital expertise (NEMO 2020; ICOM 2020). Live events were prioritised and by the end of the first lockdown a LGBT+ history group was established, alongside a new series of lectures on the library’s political collection.

However, there were unanticipated but welcomed consequences to the rapid adoption of the digital. The library’s 2020 report states, “Audiences have been developing rapidly from a national international perspective and numbers have averaged around 80 attendees per session” (Linen Hall Library 2020, 2). In the annual report for 2020, the library noted that by October it had hosted “20 online events with an (international) audience engagement of c.12,000 people” (2020, 7). This figure is larger than the library’s annual visitor numbers, and
like many GLAM institutions in 2020, the digital content was “reaching new audiences far beyond our traditional catchment area” (Linen Hall Library 2020, 7). The Centre for Cultural Value and Audience Agency found that this apparent broadening of audiences during lockdown was the result of individuals who were already heavily engaged with their local GLAM organisations digitally visiting institutions geographically distant from their own location. There was no significant change in the demographic groups accessing cultural content; visitors were “often older, professional, had children and were based in larger cities (especially London), in line with what we would expect from existing studies” (Walmsley et al. 2022, 10).

The Linen Hall Library’s online audience highlights an interesting conundrum for GLAM professionals; to what extent is it worth pursuing higher audience numbers if the majority of online visitors are unable to travel to the physical site? Whereas tracking international audience engagement may benefit the British Museum – in 2019 64% of their visitors were international tourists – this is less true of the Linen Hall Library (Bailey 2023). GLAM organisations have consistently struggled to monetise their online audiences (Anderson 2018, 91), with the profitability of image licensing being repeatedly questioned (Grosvenor 2018; Tanner 2004), and print on demand services remaining the reserve of large institutions (Valeonti et al. 2019; Valeonti et al. 2018). To start monetising online audiences in this way requires a sustained investment in digital infrastructure, with an online shop also necessitating large upfront costs. It is therefore questionable whether it is worth the required resources to monetise online visitors if this does not directly serve the organisations’ target audience.

The Linen Hall Library’s online events have since been uploaded to YouTube, and it is clear they facilitated rich discussions about Northern Irish history and engaged local audiences as well. However, the depth of engagement is not captured in online metrics, with the majority of content not receiving any comments on social media or being viewed widely. This case study highlights that rich enrichment for a small audience is not currently measured in quantitative metrics. Considering the role many GLAM institutions play in their local communities, this makes it difficult for staff who did not attend every event to pinpoint which were the most successful, and to compellingly convey that information to funders. If organisations have the resources, interviews and surveys can help at this small scale, but is there any way the quantitative methods can be improved? This final section of the paper asks how we might begin to measure engagement in a way that is more meaningful for GLAM organisations.
5 Alternatives

By exploring social media metrics, our paper has built on the discussion around the use of quantitative measures in the cultural sector and highlighted the specific limitations presented by platforms metrics such as ‘views’, ‘likes’, and ‘shares’. While of limited use themselves, these numbers can contribute to far more sophisticated measures of engagement. This section will explore methods that could create more nuanced evaluations of online content and have been trialled at a small scale in the GLAM sector. It will briefly outline the benefits and drawbacks of three approaches: sentiment analysis, topic modelling, and network analysis. This section will not explore these computational methods in depth, but instead aims to highlight potential avenues for further research and experimentation in the GLAM sector that address some of the concerns raised regarding quantitative metrics.

5.1 Sentiment Analysis and Topic Modelling

Sentiment analysis is a machine learning tool that aims to quantify the emotional tone or opinion of a given text – usually tweets, comments, or reviews (Valdivia, Luzon, Herrera 2017). A text is then often identified as having a positive or negative sentiment, and more sophisticated approaches have been developed that label texts with greater granularity, for example ‘bored’, ‘confused’, or ‘inspired’ (Zhou, Ye 2020, 5; Gerrard, Sykora, Jackson 2017). Performing this kind of analysis on the comments of GLAM online resources can summarise large amounts of audience feedback and captures more nuanced information about the reception of online content.

In 2013, Villaespesa analysed tweets from the “Art in Action” festival at the Tate Modern. A quantitative analysis of the entire data set was combined with qualitative analysis, including sentiment analysis, on a subsection of the tweets. Each tweet in this subset was coded with a predetermined category using a qualitative data analysis software. Villaespesa then investigated visitors’ motivations and themes; although she noted that this was a time-consuming process, “automatic coding does not disclose the reason why the experience was positive or negative” (Villaespesa 2013). More recently, Arias preformed a qualitative analysis of a controversial curatorial initiative at the Manchester Art Gallery (2020). Initially utilising the metrics available through Twitter and the “predetermined metadata categories”, she supplemented this data with thematic analysis conducted using the software NVivo (Arias 2020, 135). This process of coding the topics present in the text requires close reading and human interpretation.
By using topic modelling – an unsupervised machine learning method that identifies clusters of similar words – researchers can automate identifying themes within the data set and the subject of a tweet, reducing the amount of human time needed to perform an analysis. By adopting a fully automated process, it will be possible to massively increase the size of data sets. Both of the aforementioned studies focused on a discreet time period and a single Twitter hashtag. Yet as Arvanitis, Gilmore, Florack, Zuanni noted when discussing the potential and pitfalls of machine learning, “longitudinal data tracking and the use of more data points (e.g., both in terms of the range of events and audiences evaluated) could ensure a safer and more insightful analysis and comparison of data” (Arvanitis et al. 2016). Automation makes this long-term monitoring possible and would enable the identification of sector-wide trends in online audiences. This could help inform GLAM organisations’ digital strategies by creating a benchmark and enable them to address the discrepancy between the rhetoric of broadening access and organisations’ ability to reach new audiences online (Walmsley et al. 2022).

A significant barrier to creating this benchmark is the difficulty in running an analysis on the content of smaller organisations who do not have large online audiences. Chatzopoulou, Sheng, and Faloutsos found that an average YouTube video receives one interaction – that is a comment or a like – for every 400 views (2010, 2). With half of museum YouTube channels in the UK having less than 46 subscribers, the majority of videos receive no comments at all (Charlesworth 2023, 11). The low number of comments on the Linen Hall Library Facebook posts is also indicative of this issue; it is clear that for most GLAM institutions there is simply not enough interaction on posts to evaluate audience engagement using sentiment analysis or topic modelling. This issue is compounded by a lack of expertise and resources in the sector; Nesta has found that “the majority of arts and cultural organisations still do not use data for important purposes such as understanding their audiences better through data analysis and profiling” (2017, 5).

By preforming topic modelling and sentiment analysis on a data set that spans multiple organisations, larger trends can be identified to help provide insight into the audiences of small and medium sized institutions that do not have the resources to analyse their own data. Using this kind of data set could identify success digital strategies and help improve the content of GLAM organisations which do not have large online audiences. Improvements to both methods have made conducting such work significantly easier in recent years, making a large-scale endeavour feasible.
5.2 Network Analysis

Network analysis could prove similarly beneficial and would nuance our understanding of content’s online reception by capturing the interconnected nature of organisations and their audiences. Many small organisations are highly networked, and this type of collaboration increased over the pandemic period (Crooke 2020). Linen Hall Library forms part of a rich tapestry of organisations, events, and projects, that share funding, volunteer bases, and audiences. Looking to these larger cultural ecosystems, and charting the way in which individual exhibitions, events, and even people are linked, it is possible to identify forms of value that are not captured by most online metrics. Network analysis excels at identifying actors who facilitate collaboration or disseminate information and has been used to identify key social media influencers at international and local levels (Ingenhoff, Calamai, Sevin 2021). It is able to chart the way information is spread and identify the most important actors in a network and is used both within individual organisations and across entire sectors.

In the museum sector, the same methods have been applied at a smaller scale. La Magnética conducted an analysis of the Twitter response to “Ask a Curator Day”, which saw them analyse 47,546 tweets posted over three days (2014). The team used different measures of centrality to calculate Twitter users’ relevance and influence in the communities participating in #AskACurator, and where possible, sections of the data illustrating key accounts in different countries were illustrated in network graphs. They note that this kind of analysis, over a longer period of time allow[s] us to have a deeper understanding of our museum communities. How our followers relate to us, how they relate to each other, who are the most influential users in a group, etc. And this understanding leads to a better definition of our strategy and a better way to assess our Social Media efforts’ results. (La Magnética 2014)

Kydros and Vrana similarly performed network analysis on museums using both Twitter and Instagram (Kydros, Vrana 2021; Vrana et al. 2019). In their study of Twitter, Vrana, Kydros, Kerhis, Theocharidis, Karavasilis looked at the tweets of the top 25 museums in Europe for a two-month period in 2019. They were able to identify users that facilitated exchange between international communities; unsurprisingly, 34% of the 30 most important users were official museum accounts, yet six nodes (20%) were “plain persons” (Vrana et al. 2019, 581). This approach provides an understanding of who information is reaching and via which avenues, enabling organisations to better adapt their digital strategies and identify potential collaborators.
Network analysis also facilitates a broader overview of interactions. The recent work of Chang et al. explores the Instagram followers of five South Korean museums (2022). They focused on categorizing types of visitors and identified six groups using community extraction, each with its own characteristics. Some groups were heavily orientated around a single central user – often a gallery – while others were far less centralised, with connections between users more evenly distributed amongst the group (Chang et al. 2022, 56:10). The different hashtags, subjects, and ways of disseminating information were popular in each group, enabling organisations to create content that better targets different types of users.

The potential use of network analysis is three-fold: as discussed, it can be applied to social media networks to identify key actors and better understand audiences; but it can also be used more broadly to map formal structures (Cheong, Cheong 2011). Marshall and Staehelei note that network analysis can illustrate the ways in which some local organizations are more successful at attracting donors, and positioning themselves as key intermediaries between international funding and local partners. (2015, 60)

This could be a prove a helpful tool for funders by identifying organisations who are repeatedly overlooked and underfunded.

Network analysis nuances the social media metrics currently available to GLAM organisations, and computational advances – combined with increasing awareness of the potential of digital technologies in the sector – mean that these methods are likely to become more accessible in the coming years. Chang et al. highlight the benefits these computational methods have over traditional approaches;

visitor analysis on SNS [Social Networking Services] can be continuously updated after it has been designed, unlike an interview or observational survey that must be conducted individually and is time-consuming. (2022, 56:16)

Conducted on a large scale may require a collaborative effort and the backing of larger GLAM organisations and research institutions, but it is possible that the sector may be able to develop sustainable evaluative tools that better capture online audience engagement.
6 Conclusions

Recent research has highlighted a discrepancy between the rhetoric of broadening access and organisations’ ability to broaden the demographic make-up of their audiences (Walmsley et al. 2022). Yet, researchers and GLAM professionals are unable to investigate why this is the case or improve their digital strategies without sufficient data and evaluative tools (UNESCO 2020).

Currently, the metrics widely available on social media platforms and most Web 2.0 applications are ineffective tools for measuring audience engagement. By contextualising why GLAM organisation are monitoring their audiences and placing their institutional aims in their wider funding context, the stark difference between what metrics are implied to indicate and what they actually reflect has become clear. ‘Views’, ‘likes’, ‘click-throughs’, ‘shares’, and the number of comments, are designed to facilitate a certain type of engagement – motivated by commercial interests – that bears little similarity to the type of open dialogue and meaningful participation GLAM organisations are trying to facilitate.

Through the use of two case studies, this paper has explored how we can contextualise these metrics and nuance what is inevitably a simplification of enormously complex outcomes (Espeland, Sauder 2016). It has highlighted that they are ineffective as evaluative tools. The numbers are shaped by factors that have nothing to do with the quality of the content itself and success on a platform cannot be reliably replicated, and if these metrics are used during the design process, GLAM organisations may optimise their content for the platform’s algorithms rather than audience engagement.

It is therefore important to contextualise findings with audience surveys and interviews, but there is also hope that more sophisticated methods for measuring large-scale engagement online will become widely accessible of the next few years. Sentiment and network analysis alongside topic modelling can provide more nuanced understandings of the complex interactions and connections GLAM institutions aim to foster. With the uptake of digital technologies amongst GLAM organisations set to continue, this paper’s exploration of how common metrics – and potential alternatives – can be adapted to meet different institutional needs, will become of increasing importance to researchers, GLAM practitioners, and designers alike.
Bibliography


---

Ellen Charlesworth, Claire Warwick, Leonardo Impett, Andrew M. Beresford
Designing for Audience Engagement. Exploring the Use of Online Metrics in the GLAM Sector


