Report from the 12th Let's Get Real collaborative action research and learning project

March 2026

Contents

Executive Summary

From April to December 2025 a group of 32 leaders and practitioners from 16 UK cultural organisations came together to experiment, learn and strategise around the use of artificial intelligence (AI) in their work. Let's Get Real: AI was the twelfth iteration of The Audience Agency's collaborative action learning and research programme.

This report details what we did, what we learnt and our recommendations for other cultural organisations looking to make sound ethical, strategic and practical decisions around how they might integrate AI tools and processes into their work. We don't claim to have all the answers, but the insights from the experimentation, thinking and conversations these brave organisations undertook are relevant and useful for any arts or culture organisation grappling with AI's complex challenges, risks and opportunities.

Image: LGR: AI cohort, workshop one, London, February 2025

The Let's Get Real (LGR) programme encouraged a two-pronged approach – practical experimentation with an AI tool or process that either informs a new AI policy or puts an existing one to the test. We encouraged each organisation's two participants to take ownership of either strategy or the practice. This combined practical and strategic nature of their experiments proved to be hugely significant, with the practical work generating insights and understanding that rippled out to impact strategic planning and thinking.

This symbiotic relationship between practice and strategy, and the shared experience of tackling such a complex and multi-faceted topic, unlocked new learning for many of us. Participants' feelings about AI at a personal and societal level, as well as in the context of their organisations, were the bedrock and starting point of the programme. Feelings on AI run deep and were often seemingly conflicted – many of us felt both scepticism and optimism; fear and excitement; or caution and openness, at the same time. We acknowledged and held those dualities throughout the process, as a healthy tension.

Three key themes emerged across the 16 sets of experiments - Emotions and people-centred practice; Experimentation, guardrails and strategy; Technological and data readiness. The considerations within those themes, coupled with the experimental **case studies detailed from page 33**, informed our five recommendations for cultural organisations navigating AI:

01/ Ensure processes are people-centred, not led by AI

02/ Create clear guardrails that can flex and evolve as AI tools and context change

03/ Nurture small-scale, low-stakes experimental approaches

04/ Strengthen your data before scaling your use of AI

05/ Learn with peers and share knowledge

The cultural sector is not alone in wanting to ensure our adoption of AI is responsible, ethical and ultimately purposeful. In many ways, issues around AI such as the domination of Big Tech, the need for ongoing skills-building, the financial pressures, and so on, are not new, and not different from the wider digital transformation challenges.

In terms of speed, scale and potential impact though, AI *is* different. The environmental cost, the damage being wrought by baked-in bias within large language models and the risks to the creative arts posed by generative AI are just three hugely significant challenges not yet fully understood and addressed. This scale of change and widespread integration of AI into the tools and platforms we use every day, mean AI cannot be ignored. The

cultural sector needs to urgently find its collective voice and be part of the conversation.

This LGR programme has barely scratched the surface of those huge questions. Our findings are constructive and relevant, but we know there is a long way to go. Next time around, we will dive deeper, looking into the potential social value and impact of using AI in cultural organisations, whilst still asking the hard ethical questions.

Developing the digital confidence and literacy of cultural leaders and practitioners, with a focus on AI, is imperative for our sector and wider society. This means digging into people's emotions, listening to different views, discussing the ethics and arriving at a shared understanding that is holistic, purposeful and contextual and an ethical, responsible approach to AI.

Let's Get Real: AI was collaboratively funded by participating organisations. Our partners were Jocelyn Burnham and University of Leicester's Institute for Digital Culture. The programme was kindly supported by Bloomberg Connects and Arts Marketing Association (AMA).

Image: Drawing reflecting on the LGR journey by LGR participant

What is Let's Get Real?

let's get real

Let's Get Real (LGR), is The Audience Agency's flagship collaborative action learning and research programme. Each year we gather a cohort of arts and cultural organisations to tackle different digital challenges together in a collaborative, self-funded project lasting 7 – 9 months. Groups have ranged in size from 10 organisations to 45, with each organisation sending two participants. Since 2010, LGR has supported twelve cohorts of cultural organisations to become more relevant, resilient and responsive to digital cultural change.

The LGR approach is always tailored to the cohort's individual needs and contexts and takes a human-centred design approach. This is characterised by:

Learning from others – we bring in a variety of voices and perspectives from within and beyond the cultural sector, to inform, support, guide and reflect on the challenges at hand.

Learning by doing – we support participants to test out and explore new ways of working in the context of their everyday activities. We encourage agile-based methodologies focusing on clear objectives, a willingness to experiment and iterate and a culture of accepting and learning from failures.

Learning together – we create a community of supportive peers with a shared sense of purpose, actively seeking out ways to nurture open, honest and collaborative exchange and learning. Through in-person workshops and online tools, we encourage and support the cohort to share experiences, concerns, failures and successes.

Our process includes the use of two framings designed to help us understand, define and give a common language to discuss and explore both **digital skills** and **digital activity** within the context of the cultural sector. They were developed as part of the AHRC-funded, University of Leicester-led **One by One project**. They have been extensively used, tested and validated across many TAA action learning, training and consultancy projects in both the UK and internationally.

Overview

Between April and December 2025, a cohort of 16 arts and culture organisations undertook a journey of interrogating, understanding and developing their approach to using AI. Our focus was on exploring different AI tools and their use cases and drawing on that experimentation to understand the strategic and governance implications for each organisation. All within the wider context of what it means to be a cultural organisation trying to get the best from technology.

Our 2025 cohort formed an exciting community of leaders, practitioners and organisations from across the UK's cultural sector, encompassing dance, museums, music, theatre and more.

The cohort explored issues and opportunities around AI in a supported and collaborative environment. Joining the programme in pairs, they looked at both the practical and the strategic challenges of using AI tools and creating appropriate policy frameworks to guide best practice in their settings.

The programme helped participants to explore AI with playfulness, vulnerability and in ways that offered genuine room for innovation. In taking part, organisations were supported on a journey of positive change and digital transformation.

Image: LGR: AI cohort logos

Our 2025 cohort

- Arts Marketing Association
- Art UK
- Bradford District Museums & Galleries
- Bristol Museum
- Edinburgh Festival Fringe Society
- Goethe Institute
- London Museum
- London Symphony Orchestra
- Royal Academy of Arts
- Scottish Ballet
- Sheffield Theatres
- The Audience Agency
- The Box
- The National Gallery
- The Queen's Hall

Our partners

Supporting experiments using AI requires high levels of technical skill and digital literacy. For this cohort we partnered with **Jocelyn Burnham**, an independent artificial intelligence consultant, trainer, and speaker, specialising in AI innovation through creativity and playfulness.

LGR is run in partnership with our long-term collaborator the **Institute for Digital Culture** at University of Leicester, with their Director and Professor of Museum Technology, Ross Parry.

Our supporters

LGR: AI was made possible this year by generous support, both financial and in sharing expertise, from two cultural sector partners:

Bloomberg Connects supported the project allowing the cohort to benefit from their expertise in the creation of AI tools and technologies designed to make culture more accessible.

Arts Marketing Association (AMA) supported the cohort in thinking about AI governance and policy, tapping into their existing work in this area and ensuring the project was underpinned by a tested and grounded approach to AI strategy.

What we did

We used Let's Get Real's tried and tested methodology to support the cohort in their experimentation and learning. Over the course of eight months, the cohort took part in:

- Two in-person full day workshops - one at the beginning and one at the end
- Five online workshops, between May and November 2025
- Five bespoke 50-minute support sessions helping each pair of participants to scope, carry out, iterate and evaluate practical, small-scale experiments and to develop any relevant strategic approaches
- Self-led learning and research periods between the workshops
- Online social drop-ins and troubleshooting sessions.

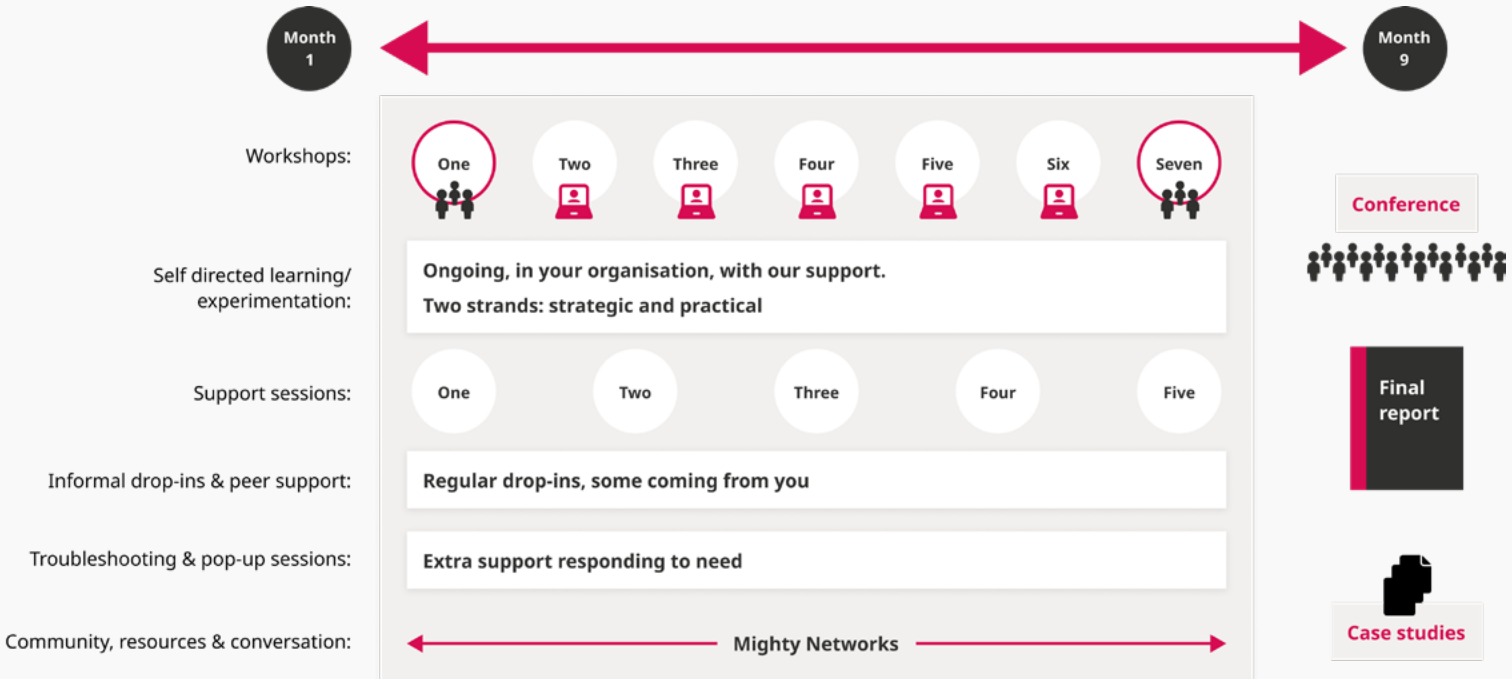


Image: Diagram illustrating LGR: AI's programme structure

Each organisation signed up two participants who attended all sessions. This helps to create a larger pool of potential collaborators on the project and

increases the likelihood of embedding change within organisations, by having two people to advocate for the work and approach.

The small-scale experiments used agile-based methodologies focusing on clear objectives, involving relevant internal and external stakeholders, a willingness to create and iterate and a culture of accepting and learning from failures.

The process of experimentation is often one of the aspects of Let's Get Real that participants value the most. All experiments were planned with our simple Experiment Planner cards (see image) and had the following characteristics:

- They sought to answer a question or address a hunch
- They involved practical actions
- They were simple and small in scale
- They used organisations' existing resources, content and technologies
- They were time-bound
- They had feedback or tracking mechanisms built in.

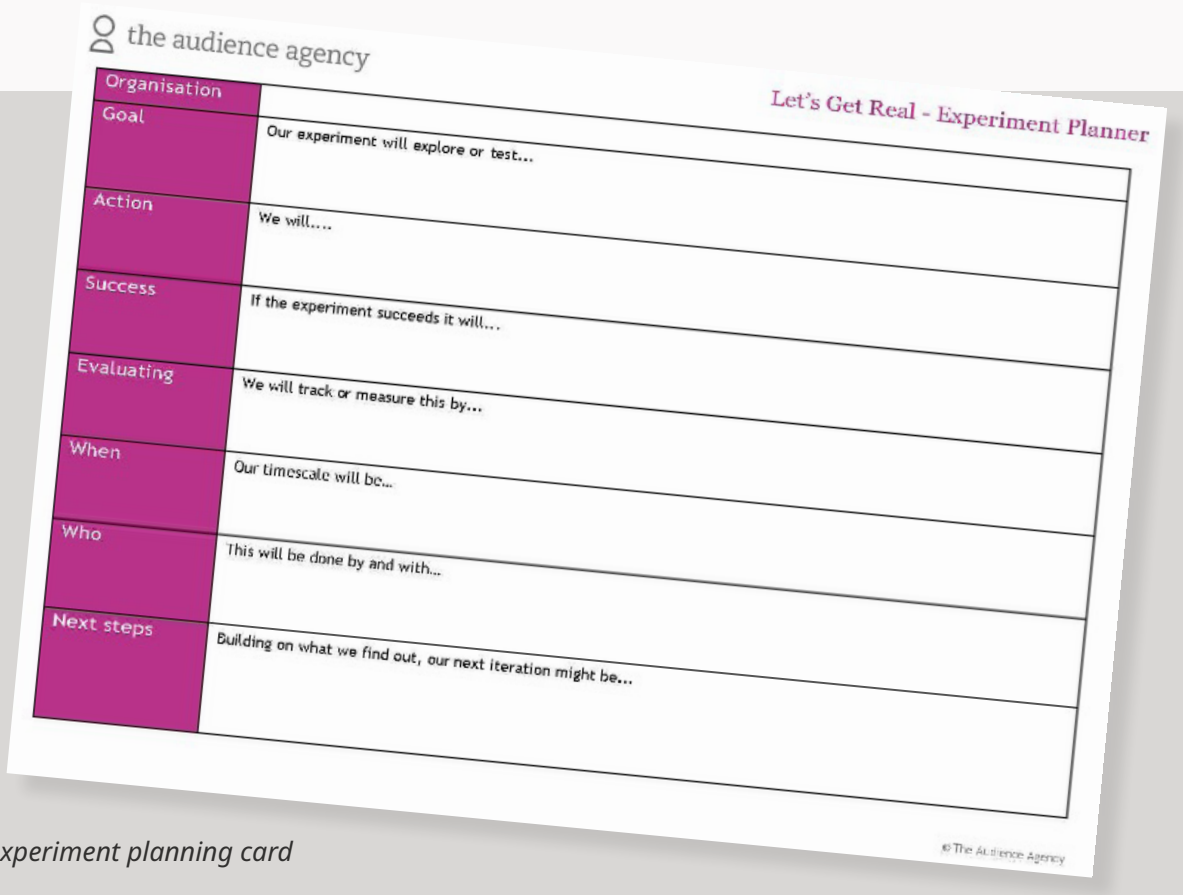


Image: Picture of LGR experiment planning card

Several guest speakers contributed insights, inspiration and advice over the course of the programme:

- **Cath Hume (CEO) and Paul Blundell (Head of Digital Research & Development), Arts Marketing Association (AMA)** - Overview of AMA's AI policy template designed to be used at governance level and as an opportunity to work on an approach to AI that can address risks, ensure consistent implementation across teams and build trust with stakeholders.
- **Jonathan Manur, Bloomberg Connects** - Overview of how Bloomberg are using AI within the Connects app to support alt text, curation and content ingestion/migration.
- **Rachel Coldicutt (CEO), Careful Industries** - Practical approaches to AI governance and balancing your values with your practice.
- **Ross Parry, Professor of Museum Technology and Director of the Institute for Digital Culture, University of Leicester** - Providing insights, reflections and support for the cohort.

Image: Cohort during workshop one in London

Experiments overview

The table below summarises the experiments each organisation undertook over the course of the programme.

Organisation	Experiment summary
Amgueddfa Cymru	**Strategic experiment:** Exploring how AI projects are developed and governed within the Museum by testing organisational processes in practice, fostering cross-team dialogue, and applying emerging AI policy through a live chatbot pilot. **Practical experiment:** Experimenting with a focused AI chatbot ('Gwybot') to improve the usability of Learning webpages by guiding users through content, refining prompts through testing, and assessing its potential to enhance navigation and user experience.
Art UK	**Strategic experiment:** Developing a forward-looking, flexible AI strategy aligned with organisational goals by assessing staff readiness, encouraging cross-team collaboration, and identifying priority opportunities and risks. **Practical experiment:** Testing AI tools to reduce a 70,000-tag backlog and support tagging 55,000+ shop prints, improving efficiency while maintaining human oversight and quality control.
Arts Marketing Association	**Strategic & practical experiment, combined:** Exploring whether AI can meaningfully analyse large volumes of qualitative evaluation data to support programme design and decision-making, while testing how its insights compare with human judgement and where it adds most value.

The Audience Agency	**Strategic experiment:** Exploring how our emerging AI policy works in practice by encouraging small, low-risk experimentation across the organisation. **Practical experiment:** Test build an AI Agent to help navigate the numerous resources within our department, to use as a training aid for new starters or to simplify tasks.
Bradford District Museums & Galleries	**Strategic experiment:** Developing a robust, fit-for-purpose Service AI policy in consultation with stakeholders to address ethical concerns, remove barriers to use, and support confident, responsible decision-making. **Practical experiment:** Testing how AI tools can improve data collection and analysis from existing workstreams, while addressing challenges around IT access, data quality, and consistency to better identify trends and audience engagement.
Bristol Museums	**Strategic experiment:** Establishing an AI policy and cross-department working group to help the museum keep pace with rapid AI developments, manage ethical and reputational risks, and identify responsible opportunities for innovation. **Practical experiment:** Developing and training an ethical, open-source StyleGAN3 image model using only the museum's own Lepidoptera images and in-house hardware to test how far generative AI can be used within strict copyright and ethical boundaries.
Edinburgh Festival Fringe Society	**Strategic & practical experiment, combined:** Exploring staff experience, use cases, and emotional responses to AI tools in the workplace to inform the development of an AI policy that reflects organisational values and needs.

Goethe-Institut	**Strategic experiment:** Designing and preparing a global 'AI temperature check' survey to explore staff perceptions, attitudes, hidden talents and support needs around AI, while aligning governance, data protection and stakeholder approval processes across the organisation. **Practical experiment:** Piloting Microsoft 365 Copilot (pro version) to evaluate its impact on efficiency and operational effectiveness by testing real-world use cases, measuring time savings and output quality, and gathering qualitative user feedback across teams.
London Museum	**Strategic & practical experiment, combined:** Testing whether AI tools could safely, reliably and transparently support forecasting, reporting and strategic decision-making using legacy museum data, by benchmarking platforms, developing proof-of-concept agents, refining data and prompts, and embedding human-in-the-loop validation within governance constraints.
London Symphony Orchestra	**Strategic experiment:** Developing an AI policy and shared vision that aligns AI use with the organisation's artistic mission, prioritising human creativity, ethical practice and AI literacy across teams. **Practical experiment:** Testing AI-generated alt text via Sprout Social to assess how effectively it supports accessible content creation over time, while monitoring accuracy, inclusivity, bias and the need for human oversight.
Royal Academy of Arts	**Strategic experiment:** Forming an organisation-wide AI working group to establish a clear, practical process for developing an AI policy, approved tools list and project checklist that supports safe and confident use of AI. **Practical experiment:** Testing whether AI can efficiently moderate and refine artwork descriptions at scale for the Summer Exhibition, improving content quality and value within tight time and resource constraints while retaining human oversight.

Scottish Ballet	**Strategic & practical experiment, combined:** Conducting a baseline assessment of AI attitudes, literacy and usage across Scottish Ballet to inform the development of a clear, values-led AI use policy that defines creative boundaries while identifying training needs, opportunities and organisational readiness.
Sheffield Theatres	**Strategic experiment:** Testing whether AI can act as a critical friend in shaping a viable strategic vision for a new theatre within an existing organisation, and where its limitations lie compared to human judgement. **Practical experiment**: Exploring whether AI can support operational planning for a new venue by contributing to audience development and marketing plans, and assessing how much human input is still required to make outputs usable.
The Box, Plymouth	**Strategic experiment:** Exploring whether AI animation and voice synthesis could authentically and ethically transform static museum objects into engaging, speaking personalities that connect with contemporary audiences, strengthen audience development goals, and establish a scalable approach to collection storytelling. **Practical experiment:** Testing whether accessible, affordable AI tools and hybrid workflows could be integrated into existing museum content production processes to create broadcast-quality animated heritage content without requiring specialist technical expertise or significant additional resources.
The National Gallery	**Practical experiment:** Testing whether AI image recognition tools could improve operational efficiency within the National Gallery's photographic team by identifying people and spaces in historically undocumented images, comparing different platforms, and assessing accuracy, reliability, and practical resource implications.

The Queen's Hall, Edinburgh	**Strategic experiment:** Developing an authentic AI policy and organisational approach that reflects The Queen's Hall's values, balances productivity with ethical concerns, and keeps human creativity central to AI use. **Practical experiment:** Testing whether AI can be trained to write event marketing copy in The Queen's Hall's distinct tone of voice, while assessing the effort required and the limits of authenticity and quality.

A full overview of each of the case studies, detailing the goal, actions taken, measurements of success, challenges, learning and take-aways can be in found in the **Experiment case studies** section.

Culture and the machine

A reflection on framing AI through the LGR methodology by project partner ***Jocelyn Burnham****.*

Image: Jocelyn Burnham

AI isn't without scepticism and criticism in any sector, but for professionals working within arts and culture, it is likely they are primed to approach the technology with a unique hesitancy.

Most highly capable generative AI systems are created from vast amounts of training data downloaded from the publicly available internet. For the most part, the copyright status of this data has not been investigated with any transparent methodology or has seemingly been ignored for the sake of scaling the technology.

Many have interpreted this as a brute and foreseeable injury against those who created this data. Writers, artists, musicians and many others in the creative fields have taken particular umbrage that their work has become part of the 'brain' of these systems, subject to whatever processing monetises it most effectively for a small handful of technology companies, without consent or reimbursement. It was this concern, for example, which encouraged Bristol Museums in their experiment to build a custom image-generating model trained entirely on their own copyright-cleared Lepidoptera collections.

The picture becomes even more complicated when other factors are considered. These include the environmental impact of using such systems and the possibility of skills degradation. These are not necessarily abstract worries: internal surveys conducted by several of the cohort revealed that their staff's most pressing anxieties were not always about job replacement, but rather the environmental costs of the technology and the potential risks to human creativity and artistic practice.

Beyond these internal workplace anxieties, the centralised power of language and opinion these systems possess is also emerging as a concern. This is particularly relevant given the increasing unease with the overt politicisation of AI in the United States, where the majority of frontier models are developed. Furthermore, there is ambiguity regarding where sensitive user data is permanently saved, whether the application of these systems will lead to job losses, and the distaste of audiences towards overtly AI-generated media.

These issues rest on top of fundamental questions which are often inferred and occasionally asked out loud: "What's the point of AI within arts and culture? What are we trying to achieve with it?"

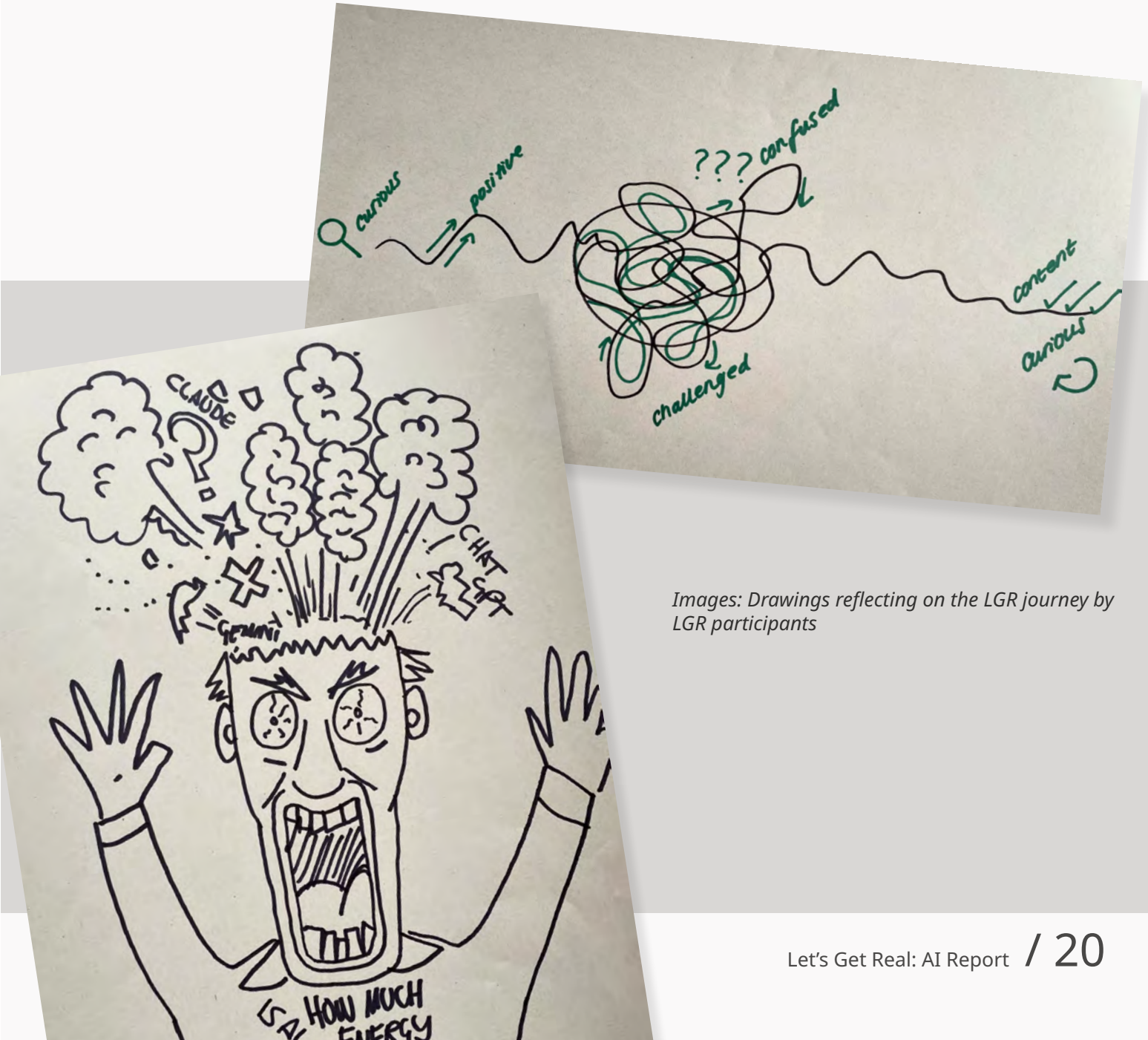
Our choice of language is important if we want to treat those questions with appropriate seriousness. An example here is the word 'efficiency', which is often synonymous with the effective application of AI. While there are processes within every role which might be made more efficient, it is fair to step back and reflect whether the language of efficiency-as-an-end-in-itself is compatible with the motivations and values of those working within culture.

The London Symphony Orchestra's experiment illustrates this limitation well. While testing AI to more efficiently generate image alt-text for accessibility, they found that an algorithm couldn't reliably detect the more nuanced, artistic context of musical performances or recognise specific conductors. It reinforced that efficiency cannot supersede accuracy, inclusivity, and human expertise.

Even if we assume perfect accuracy, it is not an overly romantic suggestion to ask whether a cultural product inherently benefits from all parts of its creation process being done more quickly. This tension between metrics and mission within cultural organisations isn't unique to using AI, but AI does perhaps create a particularly visible platform for it to be interrogated.

For the individual looking to understand the technology in a neutral fashion, this context can create an environment of overwhelm, polarised emotion and an intense concern on getting AI 'right' the first time. There is a fear of clumsy reputational mishaps or simply degrading the quality of work by using AI poorly.

Equally, those within an organisation who might internally champion AI use, perhaps as a tool for supporting neurodiverse employees, analysing data for strategic decisions, or exploring new avenues for marketing, may



Images: Drawings reflecting on the LGR journey by LGR participants

feel a pressure to demonstrate its value quickly to make a business case for investment.

In either case, the understanding and application of AI often feels like a high stakes project. This can be paralysing for meaningful innovation, where multiple iterations, experiments and an embrace of 'failure' lead to the most valuable insights. As the Royal Academy of Arts discovered in their strategic experiment, explicitly accepting the possibility of a 'failed' output was essential for gaining buy-in from stakeholders. It took the pressure off the team to have all the answers upfront, making those in the process more comfortable and open to learning.

This tension between delivering a polished output and valuing the learning process was reflected often during this process, particularly during one-to-one support sessions. An emerging question seemed to be: "Do we care more about making something which 'works' the first time, or creating something which will maximally teach us something about applying AI within one specific context of our work?"

There is no reason that you cannot create an experimental project which supplies high-quality data for future learning and provides something useful at the same time, and many of the case studies in this report demonstrate that. Nevertheless, it generally seemed that innovative thinking flourished more freely when the pressure to create a 'working' prototype on the first attempt was reduced and the emphasis shifted to the value of learning as an end in itself.

For me, the most rewarding part of working with these organisations wasn't the discovery of a novel AI process which perfectly aligned with our sector, nor the discovery of important failure points. It was seeing how the unique language and values of those individuals began to surface within our discussions of the technology and their ambitions for exploring it further. Because of this, I don't necessarily see the value of the experiments we explore here as just their outputs. The more meaningful consequence of this work was in how the friction of AI forced us to articulate (perhaps with just ourselves and our colleagues) what we, as a sector, actually care about.

This shift in mindset was perhaps well captured by The Box, Plymouth. As they reflected on their strategic planning, they noted that the framework of this programme pushed them to change their starting question from simply, "Can we do this?" to the more important, "Should we do this, and for what purpose?"

To continue that articulation is perhaps the bigger challenge and opportunity than anything related to technical upskilling or investment in new hardware.

Shaping responsible AI use together

A reflection on partnering with the LGR AI cohort by ***Cath Hume, CEO, Arts Marketing Association****.*

Image: Arts Marketing Association

It's clear that AI is here to stay, it's problematic, but also full of opportunities. What is important is that we take a responsible approach to how we use AI, ensuring we take time to understand ethics and bias, working out where our organisations stand on it, and putting the building blocks of policy and guidance in place to support our teams.

As a programme partner, we've been supporting the Let's Get Real cohort with their thinking around AI governance and policy. This has been a big focus for us at the Arts Marketing Association (AMA) over the last couple of years and has seen us develop an **AI Policy Template** alongside Target Internet, with input from the AI Sector Support Group. This provided a starting point in the development of many of the cohort's own AI governance strategies including the team at Bristol Museums.

Looking through the challenges and discussions that many of the LGR cohort were having, it mirrored much of what we've been hearing from marketers and leaders through our Together We Act research. 'AI and Digital Confidence' was one of three challenge areas that we explored with AMA members. Our research showed that many organisations and individuals are harnessing AI and finding ways to benefit from efficiencies in their work and more significant impact experimenting with things such as sales modelling, tools to manage brand guidelines and audience segmentation.

Our research showed that there is a strong desire to better understand the ethics, risks and opportunities of AI. There are concerns about environmental impact and privacy alongside deep seated worries about the intellectual property of creatives. Questions were raised around creative morality, human creativity, and the impact of AI on job losses. Some respondents questioned whether this aligns with our sector's purpose and values. This discussion is important in the context of growing resistance from artists towards AI.

So how do we deal with this hot potato? Well at the AMA we listen – to what our members are saying, to the sector, and to audiences. We discuss these challenges together and work with each other to develop the processes and solutions we need. The ethical considerations are underpinned by fear, nervousness, and a broad range of strong opinions on AI use which are dividing staff teams, impacting buy-in and affecting preparations on policy. When emotions are high, marketers and leaders need to think with clear heads and gather the facts. If they don't have them, find someone who can share the data and case studies to show the full picture and make everyone in their teams feel heard.

The exploration of the crucial ethical questions raised by AI is what underpins the work and research AMA is undertaking on responsible AI use. This underpins all our AI activity: from developing Goose, our AI powered marketing partner co-created with the UK heritage sector to democratise knowledge, to our 'Essentials of AI in Marketing' training, and recent facilitated discussion on generative AI use for receiving houses. In the coming months we will be developing our emerging work on AI standards for marketers and AI governance framework to address the policy vacuum our members are struggling to operate within.

'We are human' is one of our core values at the AMA so it's important to us to always place people at the heart of what we do – working with our members and the cultural sector to have these complex and often philosophical discussions on how we want to shape our AI use. By doing this, we can take something problematic and begin to mould it thoughtfully, sensitively and responsibly, together. AI as a creative partner, not a replacement for creativity.

Insights from LGR: AI

Across the cohort's very different organisations, common challenges surfaced around governance, staff confidence, data readiness and defining meaningful measures of success. AI rarely proved to be a quick fix and often exposed deeper cultural and structural issues.

Emotional responses, from concern and scepticism to curiosity and cautious optimism, shaped progress as much as any technical capacity and proved foundational.

Image: Word cloud of emotions about AI captured from the cohort during workshop one

As in previous LGR programmes, experimentation was central: small, low-stakes AI trials created spaces for practical insights, honest learning and more confident organisational change. It also gave the participants licence to test the boundaries of their organisations' digital infrastructure and maturity, as well as any guardrails around AI which may have already been put in place.

Three key themes emerged most clearly.

Theme #1: Emotions and people-centred practice

Many of the participants in this year's cohort brought the conversation around AI to their organisational colleagues. These in-house conversations repeatedly surfaced concern, scepticism, overwhelm and ethical anxiety around using AI both personally and professionally. Staff at The Queen's Hall, for example, were highly sceptical about AI with concerns around the environmental impact, potential loss of staff skills and AI bias. This made experimentation difficult for the team, because they felt underprepared to assess the risks and ethical implications of using AI.

There were also feelings of curiosity and excitement, not just for those taking part in LGR, but for staff outside the project too. The Art UK leadership team shifted a little of their initial scepticism to cautious openness, after some hands-on experimentation. London Symphony Orchestra staff showed curiosity, alongside their caution around artistic integrity. The focus of that concern, or enthusiasm, tended to shift depending on the role. Leadership often focused on risk, reputation and strategic positioning. Practitioners' emotions seemed more focussed around practical workload.

LGR: AI gave several participating organisations the opportunity to conduct some internal listening exercises focusing on staff opinion, thoughts and emotions around AI. The Goethe-Institut saw an opportunity to run an organisation-wide survey to build understanding of staff perceptions, readiness levels and expectations around AI. They saw this survey as a 'temperature check'. Scottish Ballet ran a similar survey, assessing concerns around AI – they found that 36% of respondents held the range of 'somewhat negative' to 'very negative' opinions, with 31% holding positive opinions. Core concerns were around creativity (59%), environmental impact (47%) and data privacy (48%). But 70% of responding staff still wanted to learn more.

The use of surveys, interviews and working groups supported the beginnings of a culture shift around AI, even before AI tools were deployed practically. It enabled participants to create an open dialogue before launching into their experimentation. It also revealed hidden tensions, appetite and opportunity. These conversations can be the building blocks for future AI adoption and policy creation with several participants using their experiment as a first step in the creation of an organisation-wide approach or to inform the revision of an existing one.

These conversations highlight a crucial factor in all AI practice – not just keeping the human in the loop (a current tech catchphrase), but ensuring people are leading and embedded in all AI work. This approach ensures that people remain actively involved in guiding, reviewing, and refining all AI outputs, rather than relying on automation alone. This is very important because AI systems can reflect biases, make errors, or lack the contextual understanding relevant to your organisation and sector.

Human oversight adds critical judgment, ethical consideration, and accountability helping to improve accuracy, build trust internally and externally, and ensure that AI is used responsibly and effectively.

Theme #2: Experimentation, guardrails and strategy

If there's one thing we have found in every LGR cohort, it is that experimentation can support and improve strategy. Experimentation unlocks the possibility of change, which is often a core component in developing and delivering a new strategy or innovation. AI experimentation is no different.

Through the cohort's small, clear and low-stakes experiments, they were able to take risks safely, unlocking strategic clarity through their processes of testing and iteration. The Royal Academy of Arts, for example, used a small pilot to see if AI could efficiently moderate and refine artwork descriptions and their results provided useful practical input to shape their AI policy development.

Risk awareness and mitigation was discussed at length and shaped many of the cohort's AI experiments. Poor use of AI tools can lead to bias, privacy breaches, compliance failures and reputational damage - making clear guardrails essential to help put boundaries in place around the use of specific AI tools within a particular organisation's context.

The Audience Agency's own AI strategy illustrates this approach. We have a set of four principles, supported by a SAFER checklist and 'traffic light' guardrails. For the latter, all AI use cases, tools and processes are discussed and monitored by the AI working group and different scenarios are flagged as either green, amber or red - permitted, use with care or stop. Our policy is a work in progress but so far the combination of checklist and guardrails works well.

Experimentation helped organisations test assumptions within clear boundaries, lowering potential risks. The Box, for example, framed its AI animation work around clear ethical parameters and organisational purpose, showing guardrails can enable innovation. In contrast, Amgueddfa Cymru's work was heavily influenced by Copilot-only access, underlining how platform restrictions can shape what is possible.

Another key learning is that AI policies must be living documents: experimentation continually challenges assumptions, and strategy must adapt as understanding, tools and risks evolve.

The cohort's experimentation revealed deeper organisational constraints and reinforced the need for clear guardrails and strategic policy. For example, complex governance structures, due to them being a worldwide organisation, delayed the Goethe-Institute's AI survey rollout.

We observed in many of the experiments that guardrails and experimentation need not be opposing forces. Instead, they need to be mutually reinforcing. Clear boundaries enable safe testing, while experimentation stress-tests those boundaries, exposing gaps in policy, process and infrastructure. In this way, experimentation with AI, within clear guardrails, strengthens strategy, turning risk into insight and helping organisations build more adaptive and resilient approaches.

> *'The Let's Get Real programme gave us the perfect framework, support network, and permission structure to conduct this experiment rigorously. Their emphasis on strategic thinking helped us articulate why this mattered beyond just "it's cool tech" and it connected directly to our brand positioning, audience development goals, and commitment to making collections accessible and engaging.'*
>
> Abigail Netcott, The Box

Theme #3: Technological and data readiness

Across the cohort we found that very few of the AI tools tried, worked fresh out of the box. Nearly all required, at the very least, significant prompt refinement. Other experiments needed a mixture of tools and approaches,

with no one AI tool fitting all needs and all purposes. Added to that was the quality of the data that the participants had at their disposal. Bradford District Museums & Galleries, for example, found their datasets too inconsistent for meaningful analysis, and London Museum recognised the need to cleanse their data early on before starting any AI experiments.

In general, this meant that time and efficiency savings were uneven at best. Whilst some areas of practice improved quickly, like the Arts Marketing Association's experimentation with prompt development for qualitative analysis, other areas required heavy iteration and exploration – particularly when it came to strategy and policy.

For some organisations, their simultaneous experiments across both strategy and practice resulted in deeper insights and learning. Practical issues often raised strategic questions that may have been missed without the insight of trying to do something within a particular context. For those organisations where the two experiments aligned and were developed together, they seemed to get the most out of the experience.

Selecting the right AI tool was a significant challenge for many of the participating organisations. They were either constrained by their infrastructure and associated policies, or they were unsure of which tool would best suit their needs and purpose. This is compounded by the fact that many of the tools work in similar ways, but can provide different results. Experimentation gave people the opportunity to try out different tools, techniques and approaches, whilst seeking support and guidance from their peers in the cohort – as well as the project team.

Prompt engineering and validation emerged as essential new skills for many participants and the focus for future staff training. These new skills exist alongside more 'traditional' digital skills which still play an important role to support work with AI, as The Box's experiment using video editing skills alongside the use of AI tools illustrates.

Despite the power and potential of these new AI tools, the need for people to remain an integral part of the process was highlighted by many. It is illustrated well in AMA's experiment using AI to understand qualitative evaluation data, which showed assumptions made by AI to be way less accurate than those made by humans who understand context better.

Recommendations

This set of recommendations around AI strategy and practice in cultural organisations is drawn from our learning so far. It is informed by the experiences of our 32 tenacious and forward-thinking participants and their 16 brave and collaborative organisations.

Recommendation	What this means and why it matters
1 **Ensure processes are people-centred, not led by AI**	A human in the loop is not enough. AI works best when it augments and supports the judgment of people, so needs to be led and shaped by your team at every stage. We need to build workflows and processes within our organisations where staff feel supported to shape prompts, review outputs and make the final decisions on any content created by AI tools. Human guidance and oversight can help to reduce bias, improve accuracy and ensure outputs align with organisational values, mission and audience expectations. As AI tools become ever more commonplace and integral to our work, it will be more important than ever that people have the final say on what the tools are being used to create and inform. Underpinning this, we need to understand how our teams, volunteers and audiences are feeling about AI, in order to support them effectively *and* inform our decisions.

2

Create clear guardrails that can flex and evolve as AI tools and context change

It is essential that we establish clear organisational guardrails around ethical use, data protection and risk management before using AI.

These guardrails may form part of an AI policy or guidance framework, and they can start life as light touch.

We should treat these policies as living documents that evolve as experimentation reveals new insights, opportunities and risks. The pace of change with AI tools and technologies is fast, so any static document will likely be irrelevant by the end of the year. Therefore, a living document ensures that our organisations can stay nimble and adaptive to the changing landscape, as our practice and the technology evolves.

3

Nurture small-scale, low-stakes experimental approaches

Just like our participants in this year's programme, begin any exploration of AI with clearly defined, time-limited pilots rather than large-scale implementations. Small experiments allow organisations to test assumptions, learn quickly and explore risks safely.

When coupled with the recommendations above, experimentation can provide practical evidence that will inform wider organisational strategy and help us to understand where AI adds value.

4

Strengthen your data before scaling your use of AI

As illustrated by this year's case studies, AI tools rely heavily on the quality and structure of the data we feed them. Before expecting meaningful, accurate results, we need to review how our data is stored, cleaned and maintained.

Investing time in effective data governance, consistent processes and clear ownership will significantly improve the usefulness of AI any tool that we are using. The more that we experiment, the better we will understand what the tools need from our data and how to get the best out of them.

It's vital not to under-estimate the time it takes to improve data quality and management, and to factor this into any decisions around AI where time-saving or increased productivity is the goal.

5

Learn with peers and share knowledge

With AI tools evolving rapidly, many organisations are facing similar challenges around governance, ethics and implementation. Within our organisation, creating opportunities for peer learning and knowledge sharing will give us a strong foundation for this process of change.

Look for AI 'champions' and form a working group for those who are interested. Learn together and limit the potential of people using the tools quietly and in un-regulated ways. This is a key step to avoiding duplicated efforts.

Don't forget to look outwards to networks, other organisations and collaborative projects. This is especially important for those of us in tiny teams – find an AI buddy in your networks.

We can't all keep up, individually, so it's vital we share our organisations' insights with the wider sector and learn from others. Connected, collaborative learning and development is a surefire way to develop a more confident approach to AI.

Useful AI-related links

This mixture of resources and deep reads makes a good starting point for exploring and embedding AI in your work.

Arts Marketing Association – Example AI policy - **https://www.a-m-a.co.uk/ai-sector-support-example-ai-policy/**

UNESCO's Ethics of Artificial Intelligence - **https://www.unesco.org/en/artificial-intelligence/recommendation-ethics**

UKRI / Arts Council England - Building a responsible AI framework for the cultural sector - **https://www.ukri.org/who-we-are/how-we-are-doing/research-outcomes-and-impact/ahrc/building-a-responsible-ai-framework-for-the-cultural-sector/**

The Audience Agency's Guide to Digital Transformation in Cultural Heritage - **https://theaudienceagency.org/en/project/guide-digital-transformation-cultural-heritage**

Let's Get Real programme & previous reports - **https://theaudienceagency.org/en/what-we-do/lets-get-real**

The Audience Agency's Digital Snapshot newsletter sign-up - **https://theaudienceagency.org/en/newsletters**

Experiment case studies

Amgueddfa Cymru

Nia Martin-Evans | *Digital Content Officer*

Danielle Cowell | *Digital Learning Manager*

What was the research question behind your experiment – what did you want to test or find out?

Strategic experiment: What is the process in developing AI projects within the Museum and how does that work practically?

Practical experiment: We wanted to experiment with AI to understand the usability of our Learn web pages. We wanted to see if people would experiment with a chatbot and whether that interaction would result in better navigation of our webpages.

Why was this important to your organisation?

Strategic experiment: From a strategic perspective, this was an opportunity to openly experiment with AI within the organisation. Something that hadn't been done before.

We wanted to establish ourselves as leaders in AI experimentation within the organization. Knowing that an AI policy was in the process of being written, we felt that the experiment was a perfect time to put that policy into practice.

Practical experiment: We were aware that our learn pages have a lot of information and content, and we were keen to see how AI could support users in finding useful information and to navigate through the content available.

We didn't want the chatbot to have access to the whole website, instead we wanted to focus on a specific area where we knew there was scope for the bot to make a difference, but would also give us the opportunity to establish how effective it was as well.

What did you do?

Strategic experiment: Once we had established what our experiment was, we decided to meet with the Head of Digital and Digital Experience Manager to discuss our idea. We felt it was important to get support for our project at the beginning before any practical work had been carried out. During this time, we also read the AI policy and were able to provide feedback on it based on our practical experiences of developing a chatbot. We felt that elements of the policy would stop people from wanting to experiment or use AI within the organisation.

We discussed our idea and use of AI broadly with our teams and also with members of the IT department. We'd experienced some push back to the idea and these dialogues proved important to allow us to overcome those issues and find means of moving forward with the chatbot.

Practical experiment: We created 'Gwybot' our chatbot in Copilot. We established the audience we were aiming to target and advised on the tone and character of Gwybot. We then provided specific prompts focusing only on the Learning area of the website. We were also able to provide a list of URLs from the website we wanted Gwybot to take its information from.

We began the process of testing Gwybot, adapting the prompts and rules associated with it. We started having conversation on how we'd publish the bot to the website. These conversations led to the experiment coming to a stop as we faced some barriers in moving forward and testing the bot on a live site.

We met with the IT representatives and was able to present the bot and our goal for supporting the user experience of our Learning pages. They were content with what we were proposing and provided us with full access to the Copilot license to develop Gwybot

further and allow us to eventually publish the bot for public use.

What happened?

Strategic experiment: Initial feedback was incredibly positive to the idea, however, as our experiment developed issues were highlighted with the processes in place to allow AI experimentation within the organisation. This resulted in the delay of our project. At the writing of this report we have no data to share. However, the dialogues that have been carried out to address the issues have resulted in the project moving forward and we are confident that the practical element of the experiment will be on the website soon. Having these dialogues and bringing colleagues who were hesitant to support the experiment on the journey with us has ironed out any issues with processes to experiment with AI in the future.

Practical experiment: At the writing of the report we haven't yet been able to publish Gwybot to the website to gather any data. We are hoping to do so within the next few months.

What was difficult?

Communication and the process of getting sign off proved to be more difficult than expected, not from heads of department, but from colleagues. There were some assumptions on our part that having had the initial support from heads of department would automatically result in colleagues getting involved and supporting the project, but that wasn't the case.

Was anything easier than you'd expected?

The process of creating the chatbot was much easier than expected. To some extent I expected the process to be highly technical or needing to go into a lot of detail, but in fact the process itself was straight forward and we had 'Gwybot' created quite quickly.

What did you each learn, as individuals?

Participant 1: Dialogue is key. We found ourselves in a cycle of email trails and cross-communication. Once we managed to sit down and explain the experiment to colleagues, we found things moving forward. I believe the term 'AI' causes a level of uncertainty and weariness, but having dialogue and explaining the details of the project allowed those who were unsure to understand the 'why' and the 'what' and buy into the project.

Participant 2: To ensure that Heads of Department and those who do the work are involved in new decisions. How you present an AI idea to colleagues his very important. Running ideas by colleagues carefully and giving time to discuss concerns face to face if possible is very important.

What did your organisation learn?

I believe we're still in the process of learning. However, I think it was a valuable exercise in putting the policy into practice and to see what the sign off process looked like.

What's next?

We are determined to complete the experiment and have the chatbot on the website. We want to use it to learn more about how users engage with the Learn area of the website. From there we can discuss the possibility of expanding the chatbot to wider parts of the website. We're also hoping that the process of developing future projects and experiments using AI will be easier.

Tips or insights about doing your experiment that might be useful for other cultural sector people in similar contexts to you:

- Face to face dialogue is powerful. The ability to describe and talk through the experiment / idea makes it easier to get buy in.
- Be clear on the why of the experiment. What is it that you want to achieve?
- Participate in all the exercises offered within the workshops. It really does help the process of developing ideas.

Art UK

Katey Goodwin | *Deputy Chief Executive*

Oliver Pitceathly | *Commercial Platforms Manager*

What was the research question behind your experiment – what did you want to test or find out?

Strategic experiment: To create an AI strategy, which is forward-looking, allows for experimentation, and connects with the mission and business objectives of the organisation.

Practical experiment: Our primary objective was to develop a plan to clear the backlog of 70,000 crowdsourced tags by approving/disapproving them with the use of AI. Our secondary objective was to use AI to tag shop prints with full oversight from the shop team to ensure all tags are manually approved.

Why was this important to your organisation?

Strategic experiment: Devising a cohesive, but flexible AI strategy will help us:

- Stay competitive and help anticipate technological shifts
- Encourage innovation and allow for experimentation
- Avoid fragmented, siloed AI efforts and encourage cross-team collaboration
- Identify ethical, regulatory, and operational risks early, allowing for responsible AI development.

Practical experiment: The primary objective is important because the task is currently overwhelmingly time consuming. The amount of user-submitted data is also so large that we would struggle to understand the implications of approving it without assistance of AI. The secondary objective is also important due to the resource currently required to tag shop prints - we have over 55,000.

What did you do?

Strategic experiment: We undertook a survey to understand baseline attitudes, proficiency, training needs. We took the first steps towards setting up an AI Working Group and drafting AI Strategy.

Practical experiment: (Primary) We used Google AI studio to produce a series of reports to help understand search behaviour and the make-up of the backlog, to help understand the data sensitivity (eg. politically biased tags). To understand the level of care required to approve tags and the ability of the AI model to follow rules. (Secondary) We used Google AI Studio to produce a CSV of tags for the shop and developed an interface for the shop team to approve them.

What happened?

Strategic experiment: From the baseline survey we learned that most staff are somewhat familiar with AI and have some confidence understanding and using AI tools. Their feelings about the impact of AI are: curious (54%), concerned (15%), skeptical (15%), excited (15%), neutral (0%). We will survey again in 2026 to see changes in attitudes and confidence.

There are huge opportunities to develop new workflows using AI. We held several workshops across the organisation, led by Configur, who have suggested some ideas for AI experiments and initiatives we could undertake. We are still in the process of assessing their suggestions, but we are looking at ways to speed up processes which are currently resource and time heavy. There are a wide range of needs and ideas across the organisation, so the processes we choose to refine through AI and the experiments we choose to conduct need to be targeted and focused. Staff are curious and open, but concerned about protecting data.

Practical experiment: (Primary) the tag backlog varies greatly in quality, and we need a reliable quality score before making mass approvals. (Secondary) The AI tool tags shop prints with over 80% accuracy with scope for improvement.

What was difficult?

Time to bring staff together to start the process, but there is a recognised need and appetite to undertake this process and experiment more with AI.

Was anything easier than you'd expected?

For practical experiments, it was easy to 'get started', to achieve meaningful tasks very quickly - but reliably verifying the results is much harder.

What did you each learn, as individuals?

Participant 1: I started off being very sceptical of AI in general, mainly due to concerns around the environmental impact, that AI-generated results could be wrong and the moral issues around images and data being scraped to teach AI systems. I didn't see the point! Once I learned more about how it works and that it can be a useful tool under certain circumstances, I became more open to using AI in different ways, such as refining text I had already written or presenting information in different ways.

Participant 2: It taught me about practical pitfalls, such as hallucinations, and common mistakes such as treating the AI as human and then putting too much confidence in simpler (for humans) tasks because of its ability to complete more complex tasks.

What did your organisation learn?

For the strategic work, we learned that there are quite a few opportunities for us to use AI to refine complicated tasks, freeing up time for staff to do other work, which is often neglected. We have been surprised by the opportunities which are available and we are now prioritising which areas to focus on first, which will make the biggest impact.

For the practical experiment, we learned a lot about generating large amounts of data (ie. artwork tags) and the importance of fully verifying it - this has been put into action on the shop.

What's next?

We are continuing to set up our AI Working Group and draft our AI strategy. We are testing out different AI systems to take notes at meetings, which is a small thing, but is already saving staff time. We are also planning some other experiments with artwork data and with our HR systems.

Tips or insights about doing your experiment that might be useful for other cultural sector people in similar contexts to you:

For practical experiments:

- Don't trust an AI model with simple tasks just because it completes complex tasks competently.
- Looker Studio has been very useful in visualising AI-generated CSV data. Looker Studios reporting gives a way to verify insights that the AI model provides on this 'big' data.
- Keep it simple! I struggled with experiments when I allowed them to get too complicated.

Arts Marketing Association (AMA)

Lewis Roden | *Member Community Manager*

Danielle Patrick | *Head of Operations and Events*

What was the research question behind your experiment – what did you want to test or find out?

Strategic & practical experiment: "Can we use AI to understand our qualitative evaluation data in a meaningful way to assist with programme suggestions and decision making?" - We opted to explore one question as both the practical and strategic elements of our experience, as we felt the question crossed both areas.

Why was this important to your organisation?

This was an important question for us because we have an abundance of qualitative and quantitative data collected from different events, projects, and time points that we wanted to explore in a more strategic way, to ensure we were designing our programmes in a member-centric way.

What did you do?

We examined a range of readily available AI tools such as ChatGPT, Gemini, Notebook LM, and Copilot and created custom versions of each with the express goal of decisive and enlightening sentiment analysis. We also curated a full set of completely anonymised feedback over the last 18 months as well as project document timelines and partner communications to provide a fully rounded picture of one of our headline events which took place at the end of 2025.

So we could methodically write the responses from each of these tools, we created our rating rubric which considered the following properties: Instruction following, verbosity, Human-like language, Logic and reasoning, Emotional intelligence and overall ratings. This paired with a series of stock prompts, was used to create review and refine the custom model options coming down to two final versions of ChatGPT 5.0 - Thinking and Gemini – 2.5 Pro.

What happened?

Early in the process it was evident that the two best models were chat GPT and Gemini. This was then refined to the two model variants which seemed to work best: ChatGPT 5.0 - Thinking and Gemini – 2.5 Pro. From our series of stock prompts the results were then reviewed by their human counterpart – attempts, results and reasoning were quite good.

Overall Ratings Summary (v1.3) rev2		
	ChatGPT 5.0 - Thinking	Gemini – 2.5 Pro
Q1	Good	Neither good nor Bad
Q2	Neither good nor Bad	Neither good nor Bad
Q3	Good	Good
Average	3.66	3.33

What was difficult?

- Collecting, cleansing and anonymising the data
- Creating the standard prompt set
- Constructing the rating rubrics
- Using Notebook LM for our purposes.

Was anything easier than you'd expected?

- Creation process for custom models
- Writing the responses and refining the Personality prompts.

What did you each learn, as individuals?

Participant 1

How to explain manual processes to a generative AI system.

How to optimise given prompts for reliability and semantic understanding. I used the Open AI Prompt optimiser for a lot of my learning.

The most to optimise the Tweaking of a custom version of a model.

Participant 2

The varying capabilities of different AI tools.

The essential element of keeping a 'human in the loop' and how that feeds into different AI use-cases.

The challenges surrounding comparing an AI output to a human output, and how to benchmark between the two.

What did your organisation learn?

As an organisation: we learned that our current working methodology is producing suggestions and results that still exceed the best of AI currently. These models can be used as tools in the initial stages to help narrow down our initial thoughts, but aren't the best option for the final results.

Expected Outcomes: we didn't have any preconceived notions of the results, so it was all surprising. The most unexpected parts of the experiment were how confidently the models replied even when they were objectively and clearly wrong.

What's next?

We will be moving to a full 'version two' of both the GPT and Gemini variants of the custom models with the aim to select one to use in a limited capacity after this further test.

Tips or insights about doing your experiment that might be useful for other cultural sector people in similar contexts to you:

The custom agents that we created were helpful for the initial pass on the data as they did them much quicker than we could do manually. They can also help to produce some previously unseen areas of insights that can be very valuable, but the current generation of models does not have enough art-specific data to support with the next steps we were looking for.

The Audience Agency

Stephen Miller | *Chief Technology Officer*

Matthew Beckett | *Senior Research Officer*

Fran Blythe | *Research Manager*

What was the research question behind your experiment – what did you want to test or find out?

Strategic experiment: We wanted to explore how our emerging AI policy works in practice by encouraging small, low-risk experimentation across the organisation.

The aim was to test whether our approach, focusing on clear guardrails, responsible use and learning through experimentation, helped staff feel confident trying AI tools.

Practical experiment: We wanted to test if we could build an AI Agent of some description to help navigate the numerous resources within our department, to use as a training aid for new starters or to simplify tasks.

Why was this important to your organisation?

Strategic experiment: As a data and insight organisation working across the cultural sector, it is important that we develop a responsible and practical approach to AI.

Testing our policy through real experimentation helped us understand whether our guardrails were clear, where people felt confident, and where more guidance might be needed.

Practical experiment: Our job in research is very process-heavy and currently a lot of these processes are stored in various places in sharepoint. It's often difficult to know where to navigate to get to the resource you require, especially for new starters. We work remotely, with flexible hours. The idea being we could provide a resource that could help reduce the reliance of an available co-worker to answer simple questions that come up during this onboarding. For more senior colleagues, we have processes that we use sporadically, so often need to jog memory on, so a tool to help navigate resources should always have use.

What did you do?

Strategic experiment:

Rather than running one single strategic experiment, we encouraged practical tests of AI tools within teams including the practical experiment for LGR AI.

These experiments were discussed within the organisation and reflected on alongside our emerging AI policy to see how well the guidance supported real use.

Practical experiment:

Stephen organised the technical side of things, working out which systems we could use from a practical point of view (cost, data security, etc).

We settled on Copilot studio. The process from then was fairly simple, I gathered all resources I could relating to our team and placed them into a single file. This was set up as the source for Copilot. I shared this amongst the team using free trials. After demo'ing the studio, I asked members of the team to use it for the next few months and record results in an excel sheet I had created.

What happened?

Strategic experiment: The timing of the LGR AI programme was really helpful as it coincided with the development of our AI policy, and Matt becoming

one of our AI Ambassadors. It gave us a practical way to test some of our thinking.

One key learning was that while policies and guardrails are important, they need to remain flexible. As soon as people begin experimenting with AI tools, the boundaries quickly get tested and new practical questions emerge.

We also found that experimentation takes time. In busy teams it can be difficult to create space for this kind of learning, even when people are interested and supportive.

Practical experiment: Feedback was generally well received - people enjoyed it as a tool compared to endless searching on Sharepoint. Many tasks it was able to assist with completely, others it would provide partial assistance with. The overall finding is that in principal, it really works, if it doesn't answer your question entirely, it will often at least guide you to the correct resource to use to answer that question. The shortcomings come in that it can only provide answers for things we have written resources for - so the quality of our own guides and how up-to-date they are, means that there is always need for time to be spend updating the back-end of the data.

What was difficult?

Setting up licences and access was probably the hardest part. We initially hoped to integrate the agent with Microsoft Teams, but this turned out to be more complicated than expected and became a bit of a stumbling block.

Encouraging people to try a new tool during a busy period was also challenging, as it can be difficult to introduce new workflows when teams are already stretched.

Was anything easier than you'd expected?

Using the Copilot studio! Very much just a case of making your resource.

From a strategy perspective, it was also encouraging how open people were to talking about AI. Once conversations started, colleagues were generally curious and willing to explore the opportunities as well as the risks.

What did you each learn, as individuals?

Participant 1: Developing a shared AI policy takes time and can be challenging, as people have different views and concerns. However, it doesn't need to stop all AI experimentation. You can start with simple principles — following data protection rules, communicating openly, and encouraging people to ask if they're unsure.

We also learned that getting to grips with the tools themselves can take time and effort. It's easy to go down a rabbit hole exploring different tools and possibilities, so having clear purposes for experimentation is helpful.

Participant 2: I think more than anything, I've learnt that there's a lot that's possible which could streamline many of our processes, but that they're not quick fixes and delegating the time to build and maintain systems needs weighing against what will be gained back from them.

What did your organisation learn?

Overall the response was positive and it sparked lots of further ideas across the organisation.

One of the biggest learnings was the emotional side of AI Adoption. People have very different feelings about AI, curiosity, concern, excitement and uncertainty, so creating space for open conversations is really important.

We also found that starting with a few clear, practical use-cases helps build trust. When colleagues can see real examples of how AI support their work (such as the Research Copilot), it helps them understand the value and feel more confident experimenting themselves.

What's next?

Next steps are to explore the best way to implement similar systems in a cost-effective way. We have lots of ideas, but the challenge now is working through the practicalities and costs.

Alongside this, we want to continue building AI literacy across the organisation, similar to how we have developed data literacy. This will include more training, practical support and regular opportunities to share AI experiences, both successes and challenges, so people can learn from each other.

Adopting AI across the organisation takes time and effort, so having some structure in place to support experimentation.

Tips or insights about doing your experiment that might be useful for other cultural sector people in similar contexts to you:

- The most important tip is to keep things simple and relevant to your organisation. It's easy to go off in lots of different directions with AI, but starting with a clear purpose and focusing on what would genuinely help your role or team is a good place to begin.
- Rather than trying to automate everything, aim for small, practical improvements that reduce some of the more tedious tasks and free up time for the work people enjoy.
- Having simple guardrails also helps people feel safe experimenting, and creating space to share both success and challenges across the org so we can build AI literacy together and over time.

Bradford District Museums & Galleries

Heather Millard | *Community Curator*

Elizabeth Llabres | *Collections Manager*

What was the research question behind your experiment – what did you want to test or find out?

Strategic experiment: Development of Service AI policy in consultation with key stakeholders.

Practical experiment: Developing the work from the 2024 LGR cohort we will utilise AI to improve workstreams and better interrogate the data we gather.

Why was this important to your organisation?

Strategic experiment: We require a fit for purpose and robust AI policy that meets organisation and stakeholder needs. To ensure quality collection and interrogation of data, where we are able to identify trends, audience interest and engagement and therefore refine decision making.

Practical experiment: To ensure quality collection and interrogation of data, where we are able to identify trends, audience interest and engagement and therefore refine decision making.

What did you do?

Strategic experiment:

- Engaged in a dialogue with stakeholders and external partners to identify key aspects of the new policy, and shape it accordingly
- Identified any barriers to use
- Listened to feedback from colleagues
- Further developed the policy and accompanying framework.

Practical experiment:

We investigated the AI tools available to us within the Council IT systems. It proved problematic, with an IT dept unwilling to offer more than a limited version of Copilot to use. This was in contrast the Council's own AI policy that had suggested a dynamic and flexible approach. We used Copilot to try and interrogate existing data, but even this proved challenging. We are endeavouring to continue to collect and collate data to feed into Copilot to better analyse.

What happened?

Strategic experiment: Evidence emerged that there was a strong desire for a robust policy and framework from both inside our organisation and from external partners. Both were interested to utilise AI and explore its benefits but had anxieties about ethical use - the framework was desired to guide the decision-making on both sides.

Practical experiment: We realised quickly that the data we had was not sufficiently consistent or 'clean' enough to easily use AI to interrogate - it lost some of the nuances. When trialled with very specific consistent data (such as our newsletter stats) it worked better. We are now looking to refine our data gathering to make for a consistent set of data to then interrogate. Confidence has grown in using prompts to create working Excel documents with formula etc that will have built in analysis - such as using for collating information in an exhibition planning process by Heather Millard.

What was difficult?

Primarily IT access. Also dedicating enough time to the experimentation process - we are a small team anyway, but our service was a key part of

Bradford2025 delivery which further impacted availability.

Was anything easier than you'd expected?

No.

What did you each learn, as individuals?

Participant 1: Developing the policy and framework in a collaborative way is a new way of working for us, but one we'll take forward. It ensures policies are useful and fit for purposes and can support staff to deliver work. Environmental and sustainability impact with AI is still very limited knowledge - it requires more work to be able to bring that alongside the decision-making.

Participant 2: That even Copilot can be useful within its specific limitations in helping to speed up some of the processes - learning it could be used to speed up all text generation for example was not something I'd considered, but is now useful (if always doubled-checked). That patience is required, and it is an ongoing journey, and moving away from having to have a defined end date/outcome to a process of gradual improvements.

What did your organisation learn?

It was great to get people's feedback and their attitudes towards the use of AI on an individual level. Attitudes were definitely mixed but not entirely a surprise.

What's next?

How will you share and embed your learning, and progress your experimentation? Through team meetings, through ongoing clean-up of data, through ongoing consultation and reflection sessions with staff.

Tips or insights about doing your experiment that might be useful for other cultural sector people in similar contexts to you:

- Patience is needed - with both the data and your IT departments.
- Build in time for reflection and conversation.
- TIME for the work and protect it fiercely - we didn't and it made the process much harder. (In our defence, it was City of Culture year....)

Bristol Museums

Mark Pajak | *Head of Digital*

Jack Evans | *Interactive Technologist*

What was the research question behind your experiment – what did you want to test or find out?

Strategic experiment: We wanted to implement a policy on AI across Museums, given the potential for increases in productivity and the increased interest from external projects that we were seeing as potential benefits from the technology. As the year has progressed there have been exponential use cases in AI both internally and in our lives. With hindsight the research question was around how it is possible to keep up with these changes, risks & opportunities through policy development, research and development and what organisational mechanisms are most useful in addressing this.

Practical experiment: We wanted to test the possibility of implementing an AI model whilst minimising the moral, ethical and legal issues surrounding premade models that have been trained through unethical forms of web crawling which have stolen people's copyrighted works to make the model. Our aim was to build an ethical Image generating AI model using StyleGAN3, trained only on our museum's Lepidoptera images. No use of AI Models/Datasets that are generating images from rights protected data in the pipeline. The resulting model was to be transparent and reproducible. This would allow us to visualise collections from our multimedia database, both finding an average image of that collection and being able to output aesthetically pleasing video. It could also help us develop powerful and ethical multimedia processing scripts in future for other projects to remove grunt work. By developing the entire AI pipeline using open source & ethical models we wanted to test the limits when fully adhering to Copyright and other regulations. We anticipated that our experiment would show what it is capable of in order to compel staff to want to use this technology rather than shy away from it due to understandable fears and, hoping that middle ground can be made with preexisting tools that keeps our use of AI ethical.

Why was this important to your organisation?

Strategic experiment: We didn't have an AI policy but could see AI in everyday use, either via Copilot Chat embedded in our work IT or individual staff members solving their own problems in their own ways. We were getting increasingly frequent pitches from external projects that involved the use of AI. We needed to develop our approach and test how the questions raised could be understood and communicated to our senior leadership team. There could be reputational risk for using AI without understanding the ethical implications. As an organisation we aim to reduce our impact on the planet and this seems to clash heavily with use of Generative AI. We can see a lot of positive examples of AI in use to increase discoverability of collections and would like to explore how we might be able to invest in and benefit from this technology in a way that does not embed prejudice associated with pre-trained models.

Practical experiment: Given how Generative AI models have stolen artists' work through web crawling and how a lot of LLMs have done the same with text across the internet, our AI policy ought to prevent staff from inadvertently breaching copyright regulations therefore preventing us from exploring the opportunities of the technology. A core function of museums is to hold copyrighted material and this is increasingly important with the emphasis on ethics through movements like decolonisation. Between this and how AI tools are becoming prevalent alongside lots of understandably negative press, we wanted to try and clear the air with the outcomes of our project to demonstrate there are ways of using Generative AI without copyright ramifications.

Using our own version of an AI model and our own digitised collections should hopefully open up the possibilities of the technology in a way that furthers our organisational missions.

What did you do?

Strategic experiment: Having got the green light from our senior leadership we set up a working group including key players across a wide range of departments. This involved establishing the terms of reference and adapting the draft AI policy based on the template published by Arts Marketing Association. This resulted in presentations at the all staff group on AI trends in the sector, which generated more interest and volunteers to join the working group.

Practical experiment: We needed to purchase a powerful computer & train the models on our tech not online - this eliminates the use of data centres during our experiment. The raw material for our experiment came from querying our collections database (EMu) for images and object records that we hold full copyright to in order to create the following AI Models:

- StyleGan Model(s) that allow us to generate images based of what is in our collections and hopefully in addition allow us to create a single average image for each image set we input e.g. one model for butterflies, another for portraits, another for landscapes etc, depending on what datasets we can achieve for viable training. A future project could involve building LLM(s) that are trained off our gallery interpretation data. This could in theory be used to create a chatbot that could be queried locally on our in-gallery touchscreens. The data processing required to achieve these models used open source/ethical AI models where possible to both speed up creation of training datasets and to show what can be achieved with these open source algorithms when applied to our database.
- Collected museum images: exported Lepidoptera assets from EMu and manually curated viable images. Isolated specimens: automatic masking using IS-Net and OpenCV; Both Apache 2.0. Images cropped and centered: trimmed empty space, aligned, and standardised images and output 512x512 and 1024x1024 versions of the dataset. Required some manual deletion of images post isolation stage
- Installed a working StyleGan3 environment: multiple OS and CUDA versions tested; final success with CUDA 12.8 + PyTorch 2.7.0 + Visual Studio 2022 with C++ package on Windows 11 with a RTX 5090
- Trained StyleGAN3: firstly with 512×512 then a new model on 1024×1024
- Image Isolation algorithm (IS-NET, which we will apply to other datasets)
- Basic Image Processing Scripts for future dataset processing (cropping, resizing, both with error handling)
- 512x512 StyleGan Model based of our viable images of Lepidoptera
- 1024x1024 StyleGan Model based of our viable images of Lepidoptera.

What happened?

Strategic experiment: Through participation in the project we learned enough about the issues associated in AI to be able to assess partner projects that we were being asked to participate in, which was a useful outcome. This resulted in a research project with the University of Bristol and Meaning Machine, a games designer looking at user testing a generative AI quest based game in our Egypt Gallery.

Through presentations at staff meetings and sharing sector publications on AI trials we sought to raise the AI literacy & preparedness of key staff members - using a teams channel to share useful material and generally just chat about AI.. We have the mechanics of the working group established with quarterly meetings in place. We have developed a draft AI policy however this needs to be a living document that evolves with new understanding and the environment.

Practical experiment: Our experiment culminated in the ability to generate an 'average butterfly' image from our dataset, using an open source model

trained on our hardware. The technology produces synthetic Lepidoptera images and aesthetically pleasing videos, which we can use for exhibits/events and possibly gallery interactives in future.

What was difficult?

As can be the case using old open source code: Stylegan3 environment setup is very difficult due to lack of documentation post RTX 3000 series (likely due to text to image models gaining traction). This nearly derailed us getting a model but eventually a working environment for the RTX 5090 was found.

Dataset prep involved manual steps as automation wasn't found for: inspecting what images would be unsuitable from the start and what images were unsuitable post extraction of Lepidoptera from database images. This was time intensive.

Dataset Prep scripts we couldn't get running on 5000 series so ran much slower using CPU (though believe this is possible now with some more environment changes). This slowed the process down.

Dataset Prep scripts I couldn't get running on 5000 series so ran much slower using CPU (though believe this is possible now with some more environment changes). This slowed the process down.

Was anything easier than you'd expected?

We didn't run into any difficulties in sourcing the hardware which was a potential risk the way technology availability can fluctuate with world events, and also internal procurement regulations on buying hardware didn't cause us any problems this time.

What did you each learn, as individuals?

Participant 1: There has been a lot of hype about AI this year and it's been so interesting exploring and experimenting with the various use cases. Whilst the hype appears to have plateaued globally I can't help but notice the little Copilot logos sitting at the top left, right, bottom, etc of all my office apps and wondering if AI has already 'won' despite our efforts to temper our approach to using it.

Participant 2: Whilst we did manage to deliver a minimum viable product, it did involve the best part of 4 days work during a busy time for me in other projects. Although our result is an ethical AI implementation, troubleshooting involved a lengthy conversation with ChatGPT reaching a 156596 word count.

What did your organisation learn?

As with anything new, the approaches we take are filtered through a personal, departmental and organisational lens. In such a diverse place as a cultural organisation this inevitably leads to a range of experiences and attitudes which are still playing out. There is definitely a reticence to using Generative AI for creating artworks due to the fear of disempowering human artists. On the productivity side there is less reluctance and our wider organisation is looking for cost saving automation through Copilot, however currently no concrete examples have emerged.

What's next?

We have a minimum viable product and are able to generate media that could be used in upcoming exhibitions. We are looking into delivering further into the capabilities of the model in order to create aesthetic content that has a purpose in educating about concepts such as evolution. On the strategy side the working group will meet quarterly to assess the current state of AI and how to make the most of the opportunities and minimise the risks.

Tips or insights about doing your experiment that might be useful for other cultural sector people in similar contexts to you:

- Open source code can be an easy way to get something up and running but beware codebases that are no longer supported or have little documentation, it can be a lot of work and you never know how deep the rabbit hole of troubleshooting will be.
- Human in the loop was probably the best way to quickly establish if an AI project ought or ought not to be attempted.
- Ethical AI is technically possible for specific use cases.

Edinburgh Festival Fringe Society

Rachel Poxon | *Website & App Manager*

Mary Bruce | *Digital Projects Coordinator*

What was the research question behind your experiment – what did you want to test or find out?

Strategic & practical experiment: We did not have a direct research question, but aimed to understand staff's approach to AI tools to inform organisational policy. We wanted to find out their current experience and use cases of any tools, and also their emotional response to AI as a tool in the workplace.

Why was this important to your organisation?

It is important that our AI Policy reflects the needs of our specific organisation and use cases. We are currently working on a number of projects which have the potential to include AI tools, and it is important that we have a policy in place that reflects our organisations values.

What did you do?

We conducted a series of interviews with all staff members over the course of two weeks. During these interviews, staff were presented with an anonymous form to assess current AI use cases. The rest of the interview consisted of a conversation to allow staff to share their thoughts and concerns more freely.

What was difficult?

Managing the project alongside other work and projects has been difficult. We are in the process of managing multiple projects across a small team, and although this is a high priority for our organisation, it has been difficult to prioritise this alongside other demands. It has also been difficult to take on such a divisive topic.

Was anything easier than you'd expected?

Staff engagement! Getting people to attend the interviews was surprisingly easy. I booked in the appointments and let them know they could rearrange if necessary, but most people turned up and engaged with the process, and were appreciative we were doing the work.

What did you each learn, as individuals?

I learned that it was valuable to allocate some time to discuss various aspects of AI use, or potential AI uses, with each staff member, as I captured some surprising insight through this aspect of the data collection. It was important to allow people to raise concerns we had not thought of previously.

What did your organisation learn ?

I think it has encouraged our organisation to engage with AI more, and hopefully we will continue down a route of careful, considered engagement as a result of the policy, report and resources which will be shared.

What's next?

The policy, report and training resources will be shared with the organisation. We have also set up an 'emerging tech' working group which we will ask to engage with these insights and apply them to particular areas where we need to establish an approach (i.e. providing artists guidance on AI use for marketing or show materials, or venues on AI captioning).

Tips or insights about doing your experiment that might be useful for other cultural sector people in similar contexts to you:

- Book in time with people. Although everyone is busy and time-poor, they are likely to appreciate you making the decision to engage with a topic which is important but not directly involved in BAU, than to be frustrated.
- Give people the space to share information about touchpoints, uses or concerns which are specific to their area of work and you might not be aware of

Goethe-Institut

Thomas Meyer | *Director of Information Services*

Andrea Pfeil | *Deputy Director*

What was the research question behind your experiment – what did you want to test or find out?

Strategic experiment: The core research question of our experiment was - How do staff members across the global Goethe-Institut network perceive the use of artificial intelligence, and what attitudes, expectations, needs and potential talents exist within the organisation regarding AI?

To explore this, we sought to design a short, visually engaging 'temperature check' survey that would capture:

- the general mindset and values staff associate with AI (adjectives, metaphors, images),
- potential hidden talents and areas in which colleagues see meaningful applications of AI,
- the needs and prerequisites staff consider essential for using AI responsibly in their daily work.

The idea was not to measure skills or productivity, but to understand the organisational culture around AI and identify emerging opportunities or concerns in a low-threshold, participatory format.

Practical experiment

My team tested Copilot 365 to evaluate if the pro version of Copilot would help team members to save time with routine tasks.

Why was this important to your organisation?

Strategic experiment: AI adoption at the Goethe-Institut is progressing rapidly, yet the institution lacks a comprehensive, organisation-wide understanding of staff perceptions, readiness levels and expectations related to AI. Before investing in training programmes, tools or guidelines, it was crucial to obtain a baseline cultural assessment.

The survey would help the organisation to:

- understand how colleagues worldwide feel about AI—enthusiastic, sceptical, curious, overwhelmed, or something else,
- uncover existing expertise and "hidden talents" already present in the organisation but not yet visible,
- identify support needs, such as training, guidelines, exchange formats or technical tools,
- ensure that future AI strategy development is participatory, evidence-based and aligned with staff perspectives.

Since AI affects nearly all areas of cultural, educational and administrative work, these insights are essential for introducing AI responsibly and strengthening organisational confidence in digital transformation processes.

Practical experiment: We aimed to identify more efficient methods for handling routine, standardised tasks as well as complex, time-consuming activities to improve overall operational effectiveness.

What did you do?

Strategic experiment: We designed a two-part experiment within the Let's Get Real: AI programme, consisting of a strategic and an operational component. My responsibility focused on the strategic part:

a) Conceptual development

Defined the overall goal: conducting a global AI temperature check across all Goethe-Institut staff.

Formulated objectives: assess attitudes, hidden talents and organisational needs.

Ensured an approach that is anonymous, non-hierarchical, and explicitly free of hidden agendas.

b) Survey design

Developed an interactive, visually appealing questionnaire using images, scales, icons and short text inputs to make participation engaging.

Included questions on:

- attitudes and personal metaphors for AI,
- areas where colleagues see potential to combine their talents with AI,
- support needed to use AI in meaningful ways,
- future visions for AI in daily work.

Coordinated closely with the internal digital transformation unit (EDIT) and incorporated feedback from colleagues.

c) Internal alignment and approval

Clarified that such a global cultural survey requires approval from the Executive Board and the Works Council.

Discussed the communication strategy (intranet announcement, emails, reminders), incentives and data privacy.

Prepared the implementation plan, but—due to approval processes—the survey could not yet be launched

Practical experiment:

1. Identified priority use cases

We began by mapping out the tasks that were either highly standardised or consistently timeconsuming.

We selected a representative set of scenarios across different teams to ensure a balanced evaluation.

2. Engaged and prepared pilot participants

We recruited a group of users whose roles aligned with the chosen use cases.

Participants were briefed on the pilot's goals, expectations, and how to interact with Copilot in their daily workflow.

3. Collected feedback and performance data

We gathered qualitative feedback through checkins, and surveys.

4. Analysed findings and identified patterns

We reviewed the data to understand where Copilot delivered the most value and where limitations appeared.

We looked for trends across teams, task types, and user behaviours.

What happened?

Strategic experiment: Since the global survey has not yet been approved, no quantitative or qualitative data have been collected so far.

Practical experiment:

Quantitative findings

Time savings:

Participants reported reductions in task completion time.

Several routine tasks (e.g., drafting emails, summarising documents) required fewer steps and less switching between applications.

Increased output volume:

Some users were able to complete more tasks within the same timeframe, particularly where Copilot supported first draft creation.

Consistency improvements:

Standardised outputs (e.g., templates, summaries) showed greater uniformity when generated with Copilot.

Qualitative findings

Higher confidence in starting tasks:

Users appreciated having a "first draft" to work from, reducing the friction of starting from scratch.

Better focus on higher value work:

Participants felt they could redirect time toward analysis, decision-making, or stakeholder engagement.

Learning curve noted:

Some users needed time to learn how to prompt effectively, but confidence improved with practice.

Variation in usefulness:

Copilot was most effective for text heavy tasks and less impactful for highly specialised or judgement based work.

Positive sentiment overall:

Feedback indicated that Copilot was seen as a helpful assistant rather than a replacement for human expertise.

What was difficult?

Implementing the strategic part of our experiment—a global AI attitude survey—proved challenging due to the complex governance structures in a worldwide organisation.

At the Goethe-Institut, any institution-wide staff survey requires formal approval from both the Executive Board and the Works Council, especially when topics touch on workplace culture, digital technologies or staff perceptions. Navigating these approval pathways takes time and coordination, which meant the survey could not yet be launched.

This made it difficult to move from concept to implementation within the timeframe of the programme.

Was anything easier than you'd expected?

The experimental and playful approach of the LGR programme made engaging with AI tools much easier than anticipated. By trying things out without pressure—both in the strategic (survey design) and operational (Microsoft 365 Copilot use) components—we experienced how experimentation lowers barriers. It sparked curiosity, reduced hesitation, and encouraged us to apply AI to many different tasks. As we engaged hands-on, AI began to feel less abstract and more like a natural part of our work culture. This sense of fun and exploration made the learning curve surprisingly smooth.

What did you each learn, as individuals?

I explored many AI tools for the first time—tools I had previously avoided either out of caution or because I was used to solving problems in familiar, traditional ways. I learned that a playful, exploratory approach works well for me and helps me overcome hesitation.

This process also showed me how crucial a mindset shift is—for myself and for organisations. Working with AI regularly makes it clearer that AI is becoming a new cultural technique, something that increasingly supports daily work and makes tasks easier. At the same time, the "magic" of AI was put into perspective: it cannot solve everything, especially in creativity, authenticity, or nuanced communication. This reinforced my belief in the principle of "human in the loop"—AI can support, but humans must guide, evaluate and contextualise.

What did your organisation learn?

Even though the survey was not yet deployed, the development process itself led to valuable organisational learning:

- It initiated an AI mindset shift among colleagues involved in the design process, encouraging reflection on AI-related attitudes, opportunities and concerns.
- It strengthened cross-regional collaboration and deepened organisational awareness that AI is not

merely a technical tool but a strategic cultural topic that affects all areas of work.

These outcomes were partly unexpected. We anticipated insights after the survey—but instead, the process of building the survey already created dialogue, reflection and learning.

What's next?

We plan to present our outcomes at an Open Space session at the Goethe-Institut's Head Office, sharing insights from both the strategic and operational components.

Next steps include:

- Revisiting whether the survey can still be implemented once approvals are clarified.
- Entering deeper dialogue with the Digital Transformation Unit (EDIT) to discuss next steps and evaluate how the experiences from Let's Get Real—including insights from peer institutions—can support institutional learning.
- Exploring how our prototype survey, and the reflections it triggered, might inform future strategy-building, guidelines or training formats.

Tips or insights about doing your experiment that might be useful for other cultural sector people in similar contexts to you:

- **Involve key stakeholders early.** AI-related surveys or initiatives often touch on workplace culture, skills and digital change. Early alignment with management and workers' councils is essential to avoid delays and ensure trust.
- **Keep experimentation low threshold and playful.** A visually engaging, accessible format helps reduce anxiety and encourages participation. Our survey prototype showed that a light, exploratory tone makes AI feel more approachable and supports mindset shifts.
- **Build reusable tools**. Even without implementation, creating a ready-to-use, scalable survey tool proved valuable. Invest in prototypes that can grow with the organisation—they become strategic assets for future initiatives.

London Museum

Kevin Carter | *Digital Solutions Architect*

Trish Thomas | *Head of Digital Innovation, Communications and Digital*

What was the research question behind your experiment – what did you want to test or find out?

Strategic & practical experiment: We wanted to test whether AI tools could safely and reliably support forecasting, reporting, and strategic decision-making using legacy museum data, while remaining accurate, explainable, and compliant with governance.

Why was this important to your organisation?

Museum data is siloed across systems, slowing and manualizing insight generation. AI offered a potential route to more responsive, data-led decision-making.

What did you do?

For the project we:

- Established a core group of stakeholders.
- Initiated two periods of technical discovery:
- We benchmarked ChatGPT, NotebookLM, and Microsoft Co-Pilot across tiers and models using real datasets, bespoke agents, and stakeholder questions.
- Established a business case.
- Developed POC tools for stakeholders
- Rolled these agents out to key stakeholders*

What happened?

No single tool or configuration worked 'out of the box'. All required significant data cleaning, contextualisation and prompt refinement. Microsoft Copilot 365 emerged as the best overall option, offering the strongest balance of accuracy, transparency, governance and alignment with our Microsoft-based infrastructure.

We also found:

- Reasoning models produced better results but were slower
- Newer beta models were faster but less reliable
- AI outputs improved significantly when agents were instructed to explain their reasoning
- A 'human-in-the-loop' approach was essential for validation.

What was difficult?

- Licensing complications
- Lack of technical knowledge
- Limited dataset: Only museum-approved data; manual addition of academic years, holidays, and weather data.
- Agent struggled to work across disparate datasets
- API integration issues: Attempts to connect external APIs were largely unsuccessful due to technical limitations.

Was anything easier than you'd expected?

No

What did you each learn, as individuals?

Participant 1: As with all development, it's not linear; it takes an enormous amount of resources to get connected to the project, to keep all the possible

elements in play at the same time, in essence, a case of two steps forward, one step back. I'd forgotten much of this from earlier software development experience.

Participant 2: That the formatting of data is all-important. That stakeholders have good intentions of participating but in the end time pressures really impact their ability to be fully invested. I think the experiment proves we need a dedicated data analysis role.

What did your organisation learn?

The organisation learned that adopting AI is far less about selecting a single 'right' tool and far more about data readiness, governance, skills and change management. While we joined the LGR programme aiming to explore whether AI could improve forecasting and reporting efficiency, the experiment revealed that AI only becomes genuinely useful once data is well-structured, contextualised and supported by clear ownership and validation processes.

A key surprise was the amount of time and iteration required to achieve reliable outputs. No tool worked effectively out of the box; meaningful results depended on prompt refinement, data normalisation and a human-in-the-loop approach. This shifted internal expectations away from quick automation wins towards a more realistic, sustainable view of AI as a decision-support tool rather than an answer engine.

What's next?

Our next step is a phased rollout of Microsoft Copilot 365, focused on embedding learning before scaling use. In the short term, the CRM and Digital Marketing Manager will act as a data manager and gatekeeper, working one-to-one with key stakeholders to structure datasets, refine prompts, and test forecasting and reporting use cases against specific, agreed data sources.

Once data integrity and confidence in outputs are established, we will move to a second phase where senior users, including the Head of Strategy and Commercial Director, can run forecasting and analytical queries through a bespoke Copilot agent using a unified, normalised dataset. This phase will be supported by additional business-as-usual training to build staff confidence in prompt writing, validation and interpretation of results.

Tips or insights about doing your experiment that might be useful for other cultural sector people in similar contexts to you:

- Start with real questions, not tools. Let genuine organisational needs drive your experimentation. Testing AI against live forecasting and reporting questions quickly reveals what's useful and what isn't.
- Invest time in data preparation and prompt design. AI outputs are only as good as the data and instructions you provide. Cleaning, normalising and adding context to datasets made a bigger difference than switching tools or models.
- Build in validation from the start. AI works best as decision support, not decision-making. A human-in-the-loop approach helped us build trust, spot errors early and set realistic expectations about what AI can (and can't) do today.

London Symphony Orchestra

Becky Lees | *Director of LSO Live*

Barbara Révész | *Senior Digital Communications Manager*

What was the research question behind your experiment – what did you want to test or find out?

Strategic experiment: On the strategic side, the research focused on developing a policy document that defines AI and sets out how its use aligns with our mission of 'inspiring hearts and minds through world-leading music-making.' This work carefully considered artistic integrity, human connection and innovation. As part of this, we developed a clear vision statement and built a supporting strategy shaped by key guiding principles, including environmental responsibility, performance, creativity and learning.

Central to our approach is the understanding that AI should not replace the skills, knowledge or experience of our teams. Instead, our aim is to build AI literacy across the organisation, enabling colleagues to critically explore and appropriately apply AI tools in their work.

Practical experiment: On the practical side, our experimentation focused on an AI-enabled alt text generation tool. We tested Sprout Social's integration with OpenAI, which generates alt text suggestions for uploaded images, to assess how effectively it supports accessible content creation. The experiment also explored whether uploading images consistently over time improved the accuracy of AI-generated descriptions, particularly in recognising orchestras, venues and well-known conductors. In addition, we evaluated the outputs for accuracy and inclusivity, identifying any recurring errors or potential biases in the generated alt text.

Why was this important to your organisation?

Strategic experiment: AI is increasingly shaping the cultural sector, and we recognised the need to take a proactive, informed approach. Developing a clear policy ensures that any use of AI aligns with our artistic values and mission, safeguarding human creativity while embracing innovation responsibly. It also helps build trust internally and externally by demonstrating transparency and ethical consideration.

Practical experiment: Accessibility is paramount for the LSO - across social media, email marketing and our website - and ensuring content is accessible to people using screen readers is a core priority.

What did you do? Please give an overview of your process.

Strategic experiment: We created an organisation-wide survey to understand colleagues' perspectives on AI at the LSO, including their concerns, questions, hopes and expectations, as well as any AI tools they are already using. Following the survey, we analysed the responses and began forming an informal cross-departmental working group. This group provides a dedicated space for open discussion around AI, ensures representation from across all departments, and establishes clear points of contact within each team for AI-related questions and support.

Practical experiment: As part of the practical testing phase, we uploaded a wide range of images through Sprout Social, including performance photography, behind-the-scenes content and promotional images featuring different orchestras, venues and conductors.

We reviewed the AI-generated alt text for each image, comparing it with manually written

descriptions and assessing accuracy, inclusivity and tone. Where necessary, alt text was edited manually to ensure clarity, accessibility and alignment with the LSO's standards.

What happened?

Strategic experiment: The survey revealed a mix of curiosity and caution among colleagues. While many saw potential for efficiency and creativity, concerns centred on ethical implications and maintaining artistic integrity. The working group has already improved confidence and sparked cross-team dialogue, highlighting the importance of shared learning and clear guidance.

Practical experiment: As part of the experimentation, we carried out a six-week testing period from 22 September to 31 October. During this time, we uploaded 10–20 photos daily, reviewed and corrected AI-generated alt text directly within the app, and tracked accuracy alongside recurring issues. We also monitored the impact on engagement and gathered feedback from colleagues with expertise in accessibility.

What was difficult?

Strategic experiment: Defining an AI policy that balances innovation with artistic integrity was challenging. There is no sector-wide standard yet, so we had to navigate complex ethical questions and anticipate future developments. Ensuring the language was accessible and relevant for all departments required careful consultation. Time constraints and the need to build consensus across teams also added complexity.

Practical experiment: During the experiment, we encountered several challenges. The AI's accuracy was inconsistent, often misidentifying instruments, venues or roles, which required regular manual correction. It struggled to understand artistic context, such as the mood of a moment or whether an image showed a rehearsal or performance, and sometimes made unintended assumptions around gender or roles that needed careful editing to ensure inclusivity. The generated text was frequently generic or repetitive, reducing its usefulness, and reviewing 10–20 images a day proved time-intensive. Improvements to the AI's suggestions were gradual, as corrections did not immediately influence future outputs.

Was anything easier than you'd expected?

Strategic experiment: Colleagues were more open to discussing AI than anticipated. The survey and working group sparked constructive conversations, and there was strong appetite for learning rather than resistance. This made it easier to position AI as a shared opportunity rather than a threat.

What did you each learn, as individuals?

Strategic experiment: Participant 2: I learned that strategic work on AI is as much about culture as technology. Building trust and clarity is essential, and framing AI as a tool to enhance—not replace—human creativity resonated strongly.

Practical experiment: Participant 1: The testing showed that AI-generated alt text works well for clear, simple images, reliably identifying instruments, colours and basic actions. Accuracy dropped when images required musical or contextual understanding, such as recognising conductors, venues or nuanced performance moments, and subtle biases or assumptions sometimes appeared, reinforcing the need for human review. While repeated corrections improved consistency over time, musical expertise remained essential for accuracy and tone. Overall, the process highlighted the importance of clear internal standards for accessible alt text and increased organisational awareness of both the value and limitations of AI tools.

What's next?

Strategic experiment: The process highlighted the need for clear policies and training, and colleagues responded positively to having a structured, ethical framework. We have published our AI policy internally, and have signposted staff to training sessions to build confidence and literacy. The working group will continue as a forum for discussion and monitoring. There is a board session on AI planned in the Spring.

Practical experiment: Building on the experimentation, we aim to develop an AI-supported tool within our Digital Asset Management system to assist with titling and metadata tagging for photos, videos and documents, improving structure, discoverability and accessibility across both current and archive content (this work is ongoing). Alongside this, we plan to create internal guidelines for the use of AI-generated alt text, consider targeted training to support team confidence and literacy, and explore collaboration with accessibility consultants from the AMA.

Tips or insights about doing your experiment that might be useful for other cultural sector people in similar contexts to you:

- Frame AI as a tool to support - not replace - human creativity; this helps build trust and engagement.
- Create a cross-departmental working group early to ensure diverse perspectives and shared ownership of decisions.

Royal Academy of Arts

Sarah Fuller | *Head of Digital Technology & Insights*

Lily Ackroyd-Willoughby | *Senior Product Manager*

What was the research question behind your experiment – what did you want to test or find out?

Strategic experiment: Can we identify key individuals across the organisation to form an AI working group, and use this to establish a process for developing an AI policy, tools list and project checklist for the RA?

Practical experiment: To test the feasibility of using AI tools to moderate and refine text descriptions for artworks in the RA's annual Summer Exhibition, to be displayed on our online catalogue and art sales platform.

Why was this important to your organisation?

Strategic experiment: Research earlier in the year showed that employees are already using AI. To safeguard against potential data and security breaches, we knew we needed to provide guidance on how AI can be used safely and effectively.

Practical experiment: User feedback frequently shows that people want to know more about the artworks that are on display. However, due to the tight turnaround of the exhibition hang, we lack the resources to manually provide this information for all 1,700+ artworks. We wanted to investigate whether it would be possible to add this detail to an artwork page, with the expectation that this could drive increased traffic to the site and as a result, potentially see an increase in online purchases.

What did you do?

Strategic experiment: We started by selecting areas of the organisation that should be involved in the working group. This was a mix of the teams that had to be involved (Digital, IT, Legal, HR), and representation from across the rest of the business (eg: Exhibitions, Finance, Marketing, Estates). We then identified individuals within these teams that represented a mix of technical experience, AI-enthusiasm and decision making authority.

To ensure the group had a clear focus, we drafted a 'Terms of Reference' document, which outlines the:

- Aims of the group
- Initial scope
- Membership
- Decision making and reporting
- Meeting frequency and format
- Review process.

For the first meeting, we kicked off by using the Terms of Reference as an agenda for discussion, ensuring that all members agreed to the purpose of the group. For the second half of the meeting, we discussed the Responsible AI report and toolkit from Arts Council England, with particular focus around the section on 'Developing an AI Policy'. The questions within this were later used as the basis for the second working group meeting.

Practical experiment: After consultation with our Content team, we pivoted from the original idea of using AI to draft the artwork descriptions based on the image and metadata already provided. Instead, we used descriptions for a small selection of artworks already provided for social media and tested to see if AI could moderate these to:

- Remove URLs and social media handles
- Correct spelling and grammar mistakes
- Flag sensitive topics for further review from a human

- The intention if successful would then be to scale this to triage descriptions for all artworks in the catalogue.

This moderation was carried out using a Google Sheet integrated with OpenRouter API to test a range of AI models. Through this testing we iterated the pre-defined prompt guidelines and evaluated several different Claude (Anthropic) AI models to uncover which combination provided the most successful outcome.

What happened?

Strategic experiment: From this first session, we found an interested, engaged but slightly nervous group. We've attributed this to the broad membership from across the organisation meaning a current lack of familiarity between members. We hope and expect conversations and opinions to flow more naturally as the group gets to know one another and there's a more focused area for discussion, where people perhaps feel more confident to disagree.

Ultimately, this is the beginning of a long journey where we hope to draft and then successfully embed an AI policy across the organisation that, due to the broad engagement of staff in its crafting, is seen as a useful and positive addition to how technology is used safely and effectively.

Practical experiment: There were two key areas of success for this experiment: feasibility and value.

- Are we able to successfully triage 1,700+ artwork descriptions in the timeframe available?
- Operational efficiency: AI proved to be faster and more reliable than a human control carrying out the same task.
- Accuracy and effectiveness: AI picked up on points missed by the human. However, we used two different models as a failsafe and a human is still needed for more sensitive triaging.
- Organisational alignment: This is a work in progress, heavily reliant on the success of the Strategy line of work. We also need to work through the impact of displaying information to the public that wasn't originally factored into the judging process.
- Cost and scalability: Demonstrated that we could easily scale this to review descriptions for all 1,700+ artworks within the catalogue, with an estimated total cost of $33. This is well within budget and significantly more affordable than the equivalent human effort.
- Does adding an artwork description to an artwork landing page provide the expected additional value?
- Artist acceptance: This part of the experiment is yet to be completed. What we intend to test here is how artists react to any changes (however small) that are made to their copy.
- User impact: This part of the experiment is yet to be completed. What we intend to test here is how users respond to having additional artwork information available, and whether this does lead to an increase in online conversion. (The counter argument to the user feedback previously provided is that knowing additional information changes the context of an artwork and so may in turn change the user's opinion and dissuade them from purchasing.)

What was difficult?

There's currently a broad range of understanding, skill and nervousness around AI across the organisation. As a result, we found that leading with AI as the subject of the experiment often caused stakeholders to be more guarded to the project idea. To combat this, we changed our approach to instead lead with the intention and proposed outcome of the idea, and then follow by mentioning that AI would be used as part of the approach.

To support these conversations, as a team we're also trying to upskill in the broader conversations around AI:

- Costs
- Environmental impact
- Ethics, specifically related to artistic creation.

As always, time to experiment alongside managing 'business as usual' tasks is always difficult, and being able to keep the impetus going after the programme concludes.

Was anything easier than you'd expected?

Using the AI tools! Once we'd created the right guidelines and selected the appropriate models, we were surprised at the reliability and efficiency of the AI monitoring and flagging. (Although we still need human involvement as a sense check to the process.)

What did you each learn, as individuals?

Participant 1: That even when testing, not always to lead with 'the exciting stuff'. What was personally interesting and intriguing was sometimes seen as scary and destructive by others.

Participant 2: That accepting the possibility of failure is an important part of experimentation and not only is this a valuable way of approaching something new, but working in an open way like this can lead to better outcomes, quicker.

What did your organisation learn?

More people than we originally thought are already using AI and/or genuinely interested in discussing its uses. However, there's often still a bit of nervousness and a desire to have structure and guidance around how to use it appropriately. We've so far seen a positive reaction towards working on an AI policy and using AI (so long as it's still in the experimental phase).

There's currently a gap between what we perceive as 'AI best practice' and what we can realistically resource, meaning that a compromise will need to be found.

When we led with AI, people were more likely to be negative about a project. However, stakeholders typically responded well to the same project when we led with the concept, and later mentioned that AI was being used as a tool within it.

In the main, people weren't comfortable with using AI to create artwork or content from scratch. AI was generally better received when using it in the role of a business support tool. In our practical test, it was also more effective supporting rather than creating.

What's next?

We will complete testing around whether adding an artwork description to an artwork landing page provides the expected additional value (artist acceptance and user impact). If this is successful, phase two of the experiment will focus on scaling and refining the process further and improving content quality. It will also cover UX testing of how AI-refined descriptions perform within different artwork page layouts. From a strategic perspective, next steps are looking to use the AI working group to feed into the creation of an AI policy. Once complete, the group will be key in helping to gain buy-in across the organisation.

Tips or insights about doing your experiment that might be useful for other cultural sector people in similar contexts to you:

- Don't be afraid to change the plan - we did this several times! We realised early on that our initial concept (AI generated artwork descriptions) wouldn't have worked from a practical or stakeholder buy-in perspective.
- It was really useful talking to lots of stakeholders within the organisation. With the practical project, this brought to light useful material that we already had available. From a strategic perspective, it's helped with buy-in and ensuring that we create a more rounded policy.
- Not being afraid for the outcome to be a 'failure' - a negative result can be as valuable as a positive one. This helped to gain buy-in with sceptical stakeholders who were worried we'd go ahead regardless of what our investigations found. It also took pressure off ourselves to not have all the answers upfront. It made everyone in the process more comfortable.

Scottish Ballet

Robert Murray | *Director of Brand, Audience & Digital*

Tony Currie | *Head of Artistic Planning & Senior Producer*

What was the research question behind your experiment – what did you want to test or find out?

Strategic & practical experiment: We set out to conduct a baseline assessment of AI attitudes and usage at Scottish Ballet, the results from which would inform the development of a clear AI use policy addressing creative boundaries.

Why was this important to your organisation?

Several reasons:

- We identified that many employees were already using AI at work and recognised the importance of having an AI usage policy to support and steer this usage.
- Anecdotal conversations within the organisation about AI showed a wide breadth of literacy, usage, and appetite. Getting a comprehensive view of this was vital to gauge the overall 'temperature' within the organisation.
- We believe there is significant untapped potential for AI usage within the organisation, but need to ensure there is a strong foundation upon which this can be explored and development.

What did you do?

Internal auditing survey to staff, dancers, orchestra, and a selection of freelance creative artists who we work with regularly or with whom we have a long-standing relationship.

The survey was designed to measure usage, competency, and potential, in order to establish baseline data on these points and identify training needs and concerns.

Small focus groups were established from a pool of individuals whose responses to each survey category varied substantially. Naturally, the makeup of the groups were skewed slightly towards those with an innate interest in and appetite for AI, though particular interest was placed on encouraging 'naysayers' to join the groups in order to provide a balanced perspective.

What happened?

HIGH-LEVEL SURVEY RESULTS

Current Usage

- 60% use AI weekly or more
- Main uses: Writing (61%), Research (60%), Admin (35%)
  AI Understanding & Confidence
- 55% understand basic concepts or well or very well
- 78% comfortable with new technology
- 83% confident using digital tools effectively
- 63% felt confident in knowing when to trust or question AI suggestions

High Perceived Value Areas

- Research & Information Gathering: 77%
- Data Analysis & Reporting: 64%
- Administrative Task Automation: 59%
- Creative Support (not replacement): 57%
- Audience Communication: 48%

Strong Learning Appetite

- 70% interested in learning more

- 50% want to join pilot projects

Top Staff Concerns

- Replacing human creativity: 59%
- Data privacy & security: 48%
- Environmental impact: 47%
- Unreliable outputs/mistakes: 38%
- Ethical concerns: 31%

Key Themes from Comments:

- “AI has no place in creative spaces”
- Conflict with environmental commitments
- Risk to stakeholder relationships
- Copyright and IP protection concerns
- 31.3% of Scottish Ballet staff feel very to somewhat positive about AI. 35.9% feel somewhat to very negative about AI.

Structural Barriers to AI Use

- Lack of training/support: 25%
- Data security concerns: 22%
- No clear access to tools: 19%
- Unclear relevance to roles: 13%

What Staff Need (Top 3)

- Clear policies on appropriate use: 63%
- Comprehensive training programmes: 42%
- Choice of whether to use AI tools: 36%

Most Valued Skills to Develop

- Writing effective prompts: 45%
- Understanding AI capabilities & limits: 42%
- Integrating AI into workflows: 39%
- Ethics and responsible use: 33%
- Evaluating and fact-checking outputs: 31%

Staff perceive that AI could benefit their work and work-life balance:

- 59% think it would help automate repetitive tasks
- 77% think it would be useful for research and planning

Preferred Learning Methods

- Hands-on workshops: 62%
- Information sessions: 49%
- Online tutorials: 44%

Key Takeaways

- Strong foundation: 60% already using AI, 70% want to learn more
- Opportunity areas: Admin, research, data analysis
- Must address: Creativity concerns, environmental impact, data security
- Success depends on: Clear policies, proper training, human creativity and choice

What was difficult?

Working at pace within a large, busy touring organisation, despite general interest in the project and readiness to engage with AI

Creative boundaries. Scottish Ballet is fundamentally a creative organisation which produces and commissions art. There is a general reticence across this industry about AI creep which hangs over discussions.

Production department buy-in (Technical, Stage, Lighting, Wardrobe, Stage Management). We believe this department could benefit immensely from employing AI tools, however they are notoriously busy making it difficult to schedule meaningful time for nuanced discussions around AI and their workflows.

Perceived stigma. An unexpected, but perhaps unsurprising, barrier was the hesitancy with which employees disclose their use of AI at work. This became particularly evident during focus group discussions, in which people shared the sense of

feeling as though they downplayed their use of AI out of fear of being criticised for being lazy, cutting corners, or 'cheating' in some way.

Was anything easier than you'd expected?

Not especially.

What did you each learn, as individuals?

Participant 1: I was somewhat surprised (though I perhaps should not have been) by the pace at which a project like this can operate within an organisation the size of Scottish Ballet. Scottish Ballet is on the smaller side for a national touring organisation, so I had expected to be able to move more swiftly. The results of the survey were more or less as expected, though I was interested to discover considerable buy-in from senior leadership, particularly in regard to the application of AI in creative spaces. Counter-intuitively (and anecdotally), there was generally more resistance from this within mid and junior level staff, though this varied greatly by remit.

Participant 2: I had the opportunity to present some of the findings of our survey to the board – where it was thoughtfully received. Not a learning per se but I was pleased that a data-driven approach to developing new policy was readily adopted. I echo my colleague's comment on the findings – some of which were surprising. What was also surprising was the response rate for the survey which was around 50% of the organisation – extraordinarily high and indicative of the level of awareness that AI has (both positive and negative). It was also a good working example of internal research/survey design, applicable for future staff research.

What did your organisation learn?

Our ambition on joining the LGR programme was to have an AI policy, or at least a draft. The pace at which the project has unfolded has prevented us from reaching this target, though we would not consider this a failure. It was only through doing the work that we identified the critical importance of laying strong foundations for a policy and organisation-wide buy-in. It would have been easy to expedite a policy, but the resulting document would likely have been more conceptual than practical. We feel it's vital that the policy is a living guide which staff can use and influence as skills and technology develop in tandem.

What's next?

Phase 1: Foundation (now)

- Develop clear AI use policy addressing creative boundaries
- Address environmental impact concerns transparently
- Establish data security protocols

Phase 2: Capability Building (mid 2026)

- Launch hands-on training programme
- Pilot projects in admin/research areas
- Establish peer learning groups

Phase 3: Selective Integration (later 2026)

- Scale successful use cases
- Maintain human-centred creative processes
- Regular impact assessment and adjustment

Tips or insights about doing your experiment that might be useful for other cultural sector people in similar contexts to you:

Many (if not all) cultural organisations are having the same internal conversations and face many of the same challenges. Lean on your networks and share learnings.

Avoid listening to the 'noise' and take the time to truly understand attitudes and capacity of your colleagues.

Similarly, avoid adopting a policy out of fear. Ensure it is bespoke to your organisation and its values.

Sheffield Theatres

Tom Bird | *Chief Executive*

Rachel Nutland | *Communications and Public Affairs Director*

What was the research question behind your experiment – what did you want to test or find out?

Strategic experiment: Could AI tools contribute meaningfully to a strategic vision and operational planning for a new venue within an existing organisation?

Practical experiment: Could AI tools aid the operational planning for the new theatre?

Why was this important to your organisation?

Strategic experiment: In January 2025 Sheffield Theatres added a new venue - The Montgomery - as our fourth theatre. We wanted to explore and test AI as a critical friend in strategic planning for our new theatre.

Practical experiment: We wanted to see whether AI tools could alleviate some human resource pressure or add additional value to the operational planning for the new venue. Could it produce an audience development strategy and an actionable marketing plan?

What did you do?

Strategic experiment: We started with an ambition that we gradually had to pare down. Originally we wondered whether we could make AI absolutely integral to the running of a 'new' theatre, whether it could undertake work in every element of strategic planning. I would say we pared this back to whether it could produce a short but viable strategic plan for a few reasons: 1) our own capacity; 2) advice from programme leaders; 3) our gradually improving understanding of the capabilities and limitations of the tools.

Eventually, we decided to input a large amount of 'background' strategic material, and for ChatGPT to use this to come up with a concise but viable strategic plan for the first few years of running a new venue within an existing multi-venue organisation. So, we inputted our own early research work and speculative board papers; the strategies of our major stakeholders; and information regarding the rest of the organisation.

Practical experiment: ChatGPT had already been given our initial strategic work on the project, our organisational strategic plan, the Arts Council England strategy and our local authority economic and cultural strategies. I then inputted the upcoming programme, historic sales figures and anonymised audience postcode data.

Chat GPT was prompted to respond as if it were the marketing manager for Sheffield Theatres and asked to provide an audience development plan for 2026. The LLM was given resource parameters and asked what additional information it required and what questions it had in order to provide this.

I kept answering questions and prompting further until a plan was provided. This was initially run twice as I was able to improve my initial prompting through highlighting areas of knowledge gaps and lack of clarity.

What happened?

Strategic experiment: ChatGPT was able to come up with a (mainly) passable strategic plan for the venue. It made inexplicable simple spelling errors when we asked it to improve the visual look of the plan. More profoundly, there were two major failings in ChatGPT's plan: 1) it lacked any careful consideration of what resource was available, and during the prompting process it repeatedly failed to ask for

that; 2) unsurprisingly, it seemed to lack flair and innovative creativity.

Practical experiment: ChatGPT reiterated what it knew from the uploaded documents, what the available resources were and what it was being asked to do, plus what additional information it required to be successful. Following 90 minutes of prompting and information gathering it provided an outline of a strategy, with SMART objectives and KPIs. It offered to put these into a downloadable document. What it provided as a download was four pages of charts of the data it had been given and a suggestion that KPIs were needed, rather than actual KPIs.

There was a real difference between what it provided in the chat, and what was produced in the download. The former was richer in content, the latter a fairly empty shell of intention of what was required. Asking it to run again lowered the quality further and some downloads were unusable.

Following this we moved to a more straightforward ask which was to give ChatGPT a show title, target audience and budget and ask it to act as a marketing officer and put together a show campaign, week by week plan and sample social media posts. Again clarity and further information was required to pivot the output to be useful, and the outcome was more successful.

What was difficult?

Getting the prompting right. It required a number of attempts to ensure that the LLM understood what was required in the task. It revealed the nuance of all the additional background information you have as a human and how much information and context needs to be conveyed.

Because so much back and forth was required it felt as though no time was being saved, some of the answers / outputs were frustrating or didn't flow intuitively from one instruction to the next, and we were really conscious of the environmental impact of working in this way. Additionally none of the downloadable outputs for the operational elements were usable, so the work then needed to be completely duplicated anyway.

Was anything easier than you'd expected?

On the surface, it appears easy to get LLMs to understand environments or situations. The tone of ChatGPT is very much like "oh yeah I get it that's easy for me". I would say this is unhelpfully encouraging/ reassuring. I get the sense the eventual output might have been stronger if ChatGPT was programmed to find things less 'easy', and instead to probe for more contextual information.

What did you each learn, as individuals?

Participant 1: I learnt about where ChatGPT can help me at work, and where it can't. I feel comfortable inputting large amounts of information and/or data sets and asking for them to be summarised. I don't feel comfortable asking LLMs to be creative with words or images yet.

Participant 2: That it can't replace humans, it can help you refine what you are doing and prompt you to think about what ingredients you need to produce a result. It regurgitates what you input in a more concise way, but doesn't always add depth.

It may be able to save time in the long run, but only if I get better at prompting and keep the task really focused. If you were working with a team member on producing an audience development plan, or a marketing plan of course you would have to share knowledge, nuance, what has worked before / what hasn't (all the human learning and experience) but then once that was learnt, the human then has the experience to deploy that in different contexts.

I do think that there is plenty of room for me to improve my use of it through prompting and being specific about what I need. The way I am finding it useful now is to thought-partner on what information I would need to write a certain report or brief and give me a template as a starting point to then do the work and add the human insight.

What did your organisation learn?

ChatGPT can quickly process your input into a clear and cohesive output. Working with it can prompt you as a human to refine your ideas and thoughts. It adds style but not substance. Coming into the experiment

I was anticipating that AI was this all-intelligent 'thing' that could perform parts of my role but ten times quicker. I've learnt that that is not currently the case, and it takes time to get a satisfactory output.

What's next?

Drilling down the experiment, asking Chat GPT to respond to smaller more specific tasks has given more actionable results that can be evaluated for their usefulness with the team.

We have been looking to procure some economic impact analysis work, and our learning on this programme has made us consider that AI tools should be able to do some of this work.

Tips or insights about doing your experiment that might be useful for other cultural sector people in similar contexts to you:

- Prompting takes time.
- You can't assume anything or that it will 'cover' everything.
- Creative jobs are some of the most AI-resistant jobs in the world of work.

The Box, Plymouth

Abigail Netcott | *Marketing & Development Manager, The Box*

Stephen Tolfrey | *Content Producer, Studio Wallop (working as extension of The Box marketing team)*

What was the research question behind your experiment – what did you want to test or find out?

Strategic experiment: Could AI animation and voice synthesis transform static museum objects into engaging, speaking personalities that connect with contemporary audiences?

Specifically, we wanted to test whether beloved heritage characters from The Box's collection—Gus Honeybun (the iconic TV puppet), King Billy, Mildred the Mammoth, and Queen Vic could be brought to life through AI technology in a way that felt authentic, engaging, and appropriate for a museum context and speak directly to visitors in ways that would:

- Create memorable, thumb-stopping social media content
- Generate dynamic display content for The Box's reception screens
- Connect heritage objects with audiences in fresh, playful ways
- Test audience reception to AI-animated museum characters
- Establish a scalable approach for future collection storytelling

Practical experiment: Which AI tools and workflows could realistically integrate into museum content creation pipelines without requiring specialised technical expertise or massive resource investment?

More specifically:

- Could readily available AI tools (not requiring custom development or expensive licenses) produce broadcast-quality heritage content?
- Would hybrid workflows combining AI with existing skills (photography, motion graphics, video editing) be accessible to typical museum marketing teams?
- How steep was the learning curve for staff without AI experience?
- Could we develop replicable processes that other heritage organisations could adopt?

Why was this important to your organisation?

Strategic experiment:

- We needed innovative, attention-grabbing content to celebrate The Box's fifth birthday in a way that stood out in an increasingly crowded digital landscape. Traditional static posts about heritage objects weren't cutting through.
- As outlined in our marketing strategy, we're committed to reaching younger audiences (particularly Gen Z aged 18-27) who currently represent only 25% of our visitors. This digital-native audience expect dynamic, creative content, AI animation offered a way to make our collection speak their language without compromising authenticity.
- The Box aspires to be the South West's premiere cultural destination. That requires us to be bold, contemporary, and innovative in how we present heritage. This experiment tested whether we could use cutting-edge technology ethically and effectively within a museum context.

The Let's Get Real programme gave us the perfect framework, support network, and permission structure to conduct this experiment rigorously. Their emphasis on strategic thinking helped us articulate why this mattered beyond just «it›s cool tech» and it

connected directly to our brand positioning, audience development goals, and commitment to making collections accessible and engaging.

Practical experiment:

- Like most regional museums we operate with limited budgets and small teams and can't commission bespoke animation for every campaign or hire external agencies for all content production. If AI tools could augment our existing capabilities cost-effectively, it would fundamentally change what we could achieve
- Our marketing team has strong traditional skills (photography, design, social media) but limited animation or motion graphics experience. Testing whether AI could bridge that gap without requiring months of training or expensive outsourcing had significant implications for our content strategy
- If we could develop accessible workflows using affordable/free AI tools, we could share that knowledge with other heritage organizations facing similar constraints. The Let's Get Real programme explicitly encouraged this peer-to-peer learning, which aligned perfectly with our values around being progressive and connecting with new audiences.

What did you do?

Strategic experiment: Working closely with the Let's Get Real team, we developed a thoughtful, multi-stage process:

- PHASE 1: Selected four iconic characters from The Box collection with strong local recognition and emotional connection. Defined clear objectives for each platform (social media vs. in-venue displays). Established ethical boundaries: characters would only speak truth, celebrate The Box's birthday, and enhance (not replace) real heritage objects. Determined success metrics aligned with organisational KPIs
- PHASE 2: We created two distinct versions of each character using different AI technologies:
  - Track A - Ambient Versions (Reception Screens): Filmed an outstanding actor performing subtle, naturalistic movements (head turns, blinking, slight smiles) . Used Runway's Act-Two feature to transfer this human performance onto static museum photographs. Created looping, atmospheric, non-verbal animations suitable for passive viewing in physical space
  - Track B - Talking Versions (Social Media): Wrote character-appropriate dialogue celebrating The Box's fifth anniversary. Used Google Veo 3 (text-to-video AI) with detailed written prompts describing desired action, dialogue, voice characteristics, and personality traits. Generated speaking characters with appropriate facial expressions and movement
- PHASE 3: POST-PRODUCTION REFINEMENT: Imported AI-generated animations into Adobe After Effects. Composited vibrant design elements to match The Box's fifth-anniversary branding. Polished final videos for both social media posts and reception screens. Ensured professional quality through traditional motion graphics expertise
- PHASE 4: DEPLOYMENT & MEASUREMENT: Published across social media platforms with clear context about the fifth anniversary celebration. Installed ambient versions on reception screens for visitor experience. Monitored engagement metrics and audience responses

Practical experiment: Our practical process focused on building repeatable workflows with accessible tools:

TOOL SELECTION & TESTING: Evaluated multiple AI platforms for cost, ease of use, output quality, and ethical sourcing. Selected Runway Act-Two for motion transfer (free tier sufficient for testing). Selected Google Veo 3 for text-to-video generation (beta access, no cost). Confirmed both could produce outputs compatible with our existing Adobe Creative Cloud workflow

WORKFLOW DEVELOPMENT:

Ambient Animation Process:

- Select museum object photograph from digital asset library

- Recruit actor (intern/volunteer willing to participate)
- Record 30-60 second performance with smartphone camera: subtle head movements, blinks, slight smiles
- Upload performance video + museum photograph to Runway Act-Two
- AI transfers performance onto still image
- Export, import into After Effects for refinement
- Add brand elements, optimize for display screens

Total time per character: ~2 hours including rendering

Social Media Animation Process:

- Write detailed text prompt including: character personality, voice characteristics, dialogue content, desired facial expressions, camera framing, background elements
- Submit prompt to Google Veo 3
- Review AI-generated output (typically 3-5 iterations to refine)
- Select best version, export
- Import into After Effects, add design elements matching The Box branding
- Export for social media platforms (optimize for vertical/square formats)

Total time per character: ~3 hours including iterations

DOCUMENTATION: Throughout the process, we maintained detailed notes on:

- Time invested per task
- Technical challenges encountered and solutions
- Prompt writing techniques that yielded better results
- Quality comparison between different AI tools/ settings
- Resource requirements (software, hardware, skills)

SKILLS TRANSFER: Two team members with different experience levels worked collaboratively:

- Team member with After Effects skills led post-production
- Team member new to AI tools led prompt engineering and tool testing
- Cross-training ensured knowledge wasn't siloed

The Let's Get Real cohort provided invaluable practical support here. Peer learning sessions where other organisations shared their technical workflows, troubleshooting tips, and tool recommendations saved us countless hours of trial and error.

What happened?

Strategic experiment:

QUANTITATIVE RESULTS

Social Media Performance: The AI-animated character posts significantly outperformed our average content across all platforms. The reach per animation was over 80k, significantly higher than regular output that had a reach of 40k. Particularly strong performance with 18-34 age demographic, our target Gen Z audience. High video completion rates suggesting content was compelling enough to watch through.

In-Venue Impact: Anecdotal evidence from front-of-house staff of visitors stopping, pointing, and discussing the animated characters. Visitors specifically mentioned the animated characters in exit surveys and it brought laughter and joy to The Box's birthday proceedings.

QUALITATIVE FINDINGS

Audience Reception - Overwhelmingly Positive:

Comments on social media posts were enthusiastic: "This is brilliant!", "Love seeing Gus come to life!", "The Box always does something different"

Local media picked up the story as innovative heritage storytelling

Minimal negative pushback about AI use when audiences understood it was enhancing real objects from The Box collection rather than replacing human creativity, resistance was low

Staff Response: Initial skepticism from some team members transformed into enthusiasm once they saw the results. Recognition that AI could be a

tool that amplifies rather than replaces curatorial expertise. Requests from other departments to explore AI applications for their work.

INSIGHTS

- Context Matters Enormously - The dual-approach strategy proved essential. Ambient versions for physical spaces required different characteristics (subtle, looping, atmospheric) than social media versions (direct, engaging, with clear narrative). AI tools excel when deployed with platform-specific intention.
- Authenticity Through Enhancement - Because we animated real objects from our collection using our own photography and grounded the characters in true stories (The Box's fifth birthday), audiences perceived the work as authentic heritage storytelling enhanced by technology, not synthetic fabrication.
- Ethical Sourcing Builds Trust - Using AI driven by our original content (our photographs, our recorded actors, our written prompts) rather than relying on AI models trained on potentially copyrighted material gave us confidence in discussing the work publicly and addressing questions about provenance.
- Hybrid Workflows Deliver Professional Quality - The combination of AI animation with traditional motion graphics expertise in After Effects was crucial. AI provided speed and realistic movement; human creative direction ensured polish, brand consistency, and narrative coherence.
- Speed Without Compromise - Under tight deadlines for the fifth anniversary, AI enabled us to produce realistic character animation that would have been prohibitively time-consuming or expensive using traditional techniques without sacrificing quality or authenticity.

Working with the Let's Get Real cohort was transformative here. Their emphasis on rigorous evaluation pushed us to track metrics we might otherwise have overlooked, and their peer learning sessions helped us contextualise our findings within broader sector trends. The support network meant we could test, fail, iterate, and succeed with confidence.

Practical experiment

Workflow Viability:

Workflows proved repeatable and teachable to team members with varying skill levels. Total production time was approximately 40 hours across 2 weeks and equivalent traditional animation would have required external agency (£5,000-£10,000 estimated) or 100+ hours of staff time. ROI was immediately clear

Tool Accessibility:

Runway Act-Two: Intuitive interface, minimal learning curve, free tier sufficient for low-volume testing

Google Veo 3: More complex (text prompting requires practice), but iterations were fast enough to learn through experimentation

After Effects: Existing staff skills meant no additional training needed; AI outputs integrated seamlessly

Barrier to entry: LOW for organisations with basic creative software and willingness to experiment

Quality Assessment:

Ambient versions: Photorealistic, subtle movements appropriate for passive viewing in museum context

Talking versions: Occasionally uncanny (typical of current AI video generation), but acceptable for playful social media context celebrating a birthday

Post-production refinement essential: Raw AI outputs needed polish; hybrid approach delivered broadcast quality

Resource Requirements:

Software: Runway (free tier), Veo 3 (beta access), Adobe Creative Cloud (existing license)

Hardware: Standard office computers sufficient (cloud-based AI processing meant no specialised GPU needed)

Skills: Basic photography, willingness to learn prompt engineering, motion graphics helpful but not essential

Time: 5 hours per character for both versions (ambient + social)

Unexpected Practical Findings:

- Prompt Engineering is a Skill (but Learnable) - Writing effective text prompts for Veo3 required 3-5 iterations per character initially, but by the fourth character, we achieved desired results in 1-2 iterations. Learning curve was steep but short.
- Actor Performance Quality Directly Impacts Output - We initially used staff performing movements casually. Results were stiff and unnatural. When we recruited a trained actor (willing volunteer), the subtlety and naturalism of their performance transferred beautifully to the museum photographs. Human skill enhances AI output; it doesn't replace it.
- Speed Matters More Than First-Pass Perfection - Because AI generation was fast (2-5 minutes per iteration), we could afford to experiment, test multiple approaches, and refine rapidly. This was creatively liberating permission to fail fast and learn quickly.
- Ethical Sourcing Removes Creative Constraints - Knowing we owned all inputs (our photographs, our actor performances, our written prompts) meant we could confidently publish, share, and repurpose the work without copyright concerns. Ethical AI use actually increased creative freedom.
- Hybrid Workflows Are Non-Negotiable for Quality Raw AI outputs were impressive but never broadcast-ready. Traditional post-production skills (color correction, compositing, motion graphics, sound design) were essential for professional results. AI is a powerful tool in a creative toolkit, not a standalone solution.

KNOWLEDGE TRANSFER SUCCESS: After completing the experiment, we ran an internal workshop demonstrating the workflows to colleagues from Collections, Learning, and Digital teams. Three staff members from different departments successfully created their own test animations within 90 minutes. This confirmed our workflows were teachable and transferable.

What was difficult?

- Minimal technical challenges (surprisingly) - Because we actively track emerging AI tools and capabilities, we knew which platforms to use for each task and understood their limitations upfront. Our existing creative skills filled any gaps, so technical obstacles were manageable
- Significant cultural challenges (expected) - The real difficulty wasn't technical; it was navigating the 'Marmite effect' of AI-generated content. Internal Resistance included:
  - Some staff initially viewed AI as a threat to creative jobs or authenticity
  - Concerns about "cheapening" heritage with "artificial" content
  - Skepticism that AI could produce work appropriate for a serious cultural institution
  - Fear of negative public perception if we used AI
- Addressing this required:
  - Transparency: Openly discussing which AI tools we used and why
  - Education: Workshops explaining how we maintained creative control and ethical sourcing
  - Demonstration: Showing the work early and incorporating feedback
  - Framing: Positioning AI as enhancement (bringing objects to life) rather than replacement (generating synthetic content)
- External Perception Management: AI divides opinion sharply in public discourse. We needed to:
- Anticipate criticism about using AI in heritage contexts
- Develop clear messaging about ethical use and human creativity
- Be prepared to defend decisions while remaining open to legitimate concerns
- Balance innovation with respect for traditional museum values
- Balancing Innovation with Authenticity - Perhaps the thorniest challenge was defining where the line sits between creative enhancement and historical misrepresentation. Questions we grappled with:

- Is animating a photograph of a real person ethical, even if the person is deceased and the content is celebratory?
- How do we signal what's original vs. AI-enhanced transparently?
- When does bringing heritage 'to life' cross into fabrication?
- What guidelines should govern AI use with sensitive historical content (e.g., war, trauma, contested histories)?

- Resource challenge: finding time for experimentation - Even with supportive leadership, carving out time for experimentation within the pressures of day-to-day operations was difficult. The fifth anniversary deadline created urgency but also meant we were learning under pressure.

Was anything easier than you'd expected?

- AI Tool Usability - We expected steep learning curves requiring extensive training or technical expertise. Reality was both Runway and Veo 3 were remarkably intuitive.
- Quality of Outputs - We anticipated needing extensive post-production correction to make AI-generated content usable. Reality was the raw outputs were often 80-90% there. Yes, refinement was needed, but the baseline quality exceeded expectations, meaning our time investment went into polish and enhancement rather than salvage operations.
- Positive Audience Reception - We braced for significant backlash about using AI, particularly from heritage purists. Reality was the public response was overwhelmingly positive. When audiences understood we were enhancing real collection objects with technology rather than replacing human creativity, resistance was minimal.
- Transparency and context defused potential criticism - We were upfront about using AI, explained why and how, and grounded the work in authentic heritage storytelling. Audiences appreciated the honesty.
- Speed of Iteration - We expected AI experimentation to be slow: submit requests, wait days for results, review, repeat. Reality was iteration cycles were measured in minutes, not days. Speed meant we could afford to fail, learn, and try again, a luxury rarely available in resource-constrained museum settings.
- Staff Enthusiasm Once They Saw Results - We anticipated needing to 'sell' AI to skeptical colleagues through extensive advocacy. Reality that once staff saw the animated characters in action, enthusiasm was immediate. The work spoke for itself. Seeing Gus Honeybun «talk» or Mildred the Mammoth «blink» was delightful and sparked creative ideas across departments.
- Knowledge Transfer Across the Organization - We worried that AI skills would remain siloed within the marketing team. Reality was colleagues from Collections, Learning, and Digital teams picked up the workflows quickly when we ran an internal workshop. Within 90 minutes, staff with no prior AI experience created their own test animations.

What did you each learn, as individuals?

AI is a tool, not a threat. My initial anxiety about AI replacing creative work was misplaced. The experiment demonstrated that AI amplifies human creativity rather than replacing it. The most successful outputs came when AI handled technical execution while humans provided creative direction, narrative, and emotional intelligence.

Prompts are a new literacy. Writing effective text prompts for AI tools is genuinely a skill like learning to direct a videographer or brief a designer. I learned to be specific, descriptive, and iterative. This «prompt engineering» capability is becoming as valuable as traditional creative skills.

Experimentation requires permission structures. Having the Let›s Get Real framework gave me license to test, fail, and iterate without fear of wasting resources or being seen as frivolous. External validation (being part of a national programme) made internal advocacy much easier.

Comfort with uncertainty increased dramatically. I›m typically risk-averse and like having control over creative processes. AI introduces unpredictability you don›t always know what you›ll get. Learning to embrace that uncertainty, iterate quickly, and trust the process was personally challenging but ultimately liberating.

Confidence to lead innovation. Successfully delivering this experiment positioned me as a credible advocate for thoughtful technology adoption. I feel equipped to lead future AI experiments and help colleagues navigate the learning curve.

Shifted from fear to curiosity. I started the programme anxious about AI, I'm ending it genuinely excited about possibilities while maintaining healthy scepticism.

What I'll Take Forward: The importance of ethical sourcing and transparency. I learned that HOW you use AI matters as much as WHETHER you use it. Using our own content as inputs, being transparent about processes, and grounding AI in authentic stories creates valuable work. I'll apply these principles to all future technology adoption.

What did your organisation learn?

- AI Can Serve Heritage Missions Authentically - Initial organisational skepticism centred on whether AI could be appropriate for serious cultural institutions. The experiment demonstrated definitively that AI can enhance heritage storytelling without compromising authenticity when used thoughtfully and ethically. Leadership recognised that:
  - AI excels at bringing existing collection content to life, not replacing curatorial expertise
  - Hybrid approaches (AI + human creativity) produce better results than either alone
  - Speed and cost efficiencies enable more ambitious content strategies
  - Younger audiences respond positively to innovative presentations of heritage
- This shifted organisational perception from "Should we use AI?" to "How should we use AI responsibly?"
- Experimentation Requires Protected Time and Permission - The organisation learned that innovation doesn't happen in spare time or as an afterthought. If we want staff to experiment, we must create structures that protect and enable that work.
- Skills Development is Strategic Investment - The rapid knowledge transfer (workshop attendees creating their own animations in 90 minutes) demonstrated that AI literacy is teachable and valuable across departments. Response was that the organisation approved budget for AI tools and platforms subscriptions, staff training, knowledge-sharing sessions and documentation of best practices for future reference
- Leadership recognised that building internal AI capability is more sustainable than repeatedly outsourcing.
- Ethical Frameworks Must Precede Widespread Adoption - The experiment raised important questions we couldn't fully answer such as:
  - When is enhancing heritage content appropriate vs. problematic?
  - How do we maintain transparency about AI use?
  - What guidelines govern sensitive historical material?
  - How do we balance innovation with respect for cultural traditions?

SURPRISES (WHAT WE DIDN'T EXPECT):

- External Interest - This positioned The Box as a thought leader in digital innovation for heritage, generating media coverage, conference invitations, and partnership opportunities we hadn't anticipated.
- Commercial Applications - We framed the experiment as content creation for the fifth anniversary. We didn't anticipate commercial applications. This suggests potential revenue streams we're now exploring.
- Commitment to continued experimentation - Leadership's response to the experiment has been enthusiastically supportive and strategically engaged:

What's next?

- Conference Presentations
  - Presenting at Museums Association conference (submitted proposal)

- Speaking at Digital Learning & Engagement Network events
- Contributing to Heritage Digital discussions

- Published Case Studies
  - Writing detailed case study for Let's Get Real programme report
  - Blog post series on The Box website documenting process, tools, and lessons

- Within The Box:
  - Monthly 'AI Literacy' workshops for all staff
  - Departmental sessions tailored to Collections, Learning, Digital, Marketing
  - Building confidence and capability across the organization
  - Central documentation of workflows, tools, and best practices
  - Shared drive with templates, examples, and troubleshooting guides
  - 'AI Use Playbook' accessible to all staff
  - Quarterly meetings bringing together staff interested in experimentation
  - Safe space to propose ideas, share failures, celebrate wins
  - Direct pipeline from grassroots innovation to strategic planning
  - 'Experimentation sprints' built into project timelines
  - Budget allocated for testing new tools/ approaches
  - KPIs that value learning, not just immediate results

Tips or insights about doing your experiment that might be useful for other cultural sector people in similar contexts to you:

- Start with "Why?" Not "How?" Don't lead with the technology. Define your organisational goal first, then ask if AI can help achieve it. Our experiment succeeded because we started with a clear purpose (celebrate fifth anniversary + reach Gen Z audiences) and evaluated AI as one possible solution. Too many AI projects fail because they're technology-led rather than mission-led. The Let's Get Real framework forced this discipline borrow their structure.

- Use Your Own Content as AI Inputs Avoid the ethical minefield of AI models trained on scraped internet data. We photographed our own collection objects, recorded our own actors, wrote our own prompts. This gave us complete creative control and copyright clarity. You can confidently publish, share, and repurpose the work. Bonus: audiences trust "enhanced authenticity" over "synthetic fabrication." Own your inputs, own your outputs.

- Hybrid Workflows > Pure AI AI excels at speed and scale but lacks human judgment, creativity, and cultural sensitivity. The best results came from combining AI (for animation, realistic movement) with traditional skills (motion graphics, brand consistency, narrative). Don't expect AI to do everything—design workflows where AI handles technical execution and humans provide creative direction. Your existing skills are valuable; AI makes them more powerful. Train your team to be 'creative technologists' who know when to use which tool.

The National Gallery

Lawrence Chiles | *Head of Digital*

Lucinda Blaser | *Senior Product Manager*

What was the research question behind your experiment – what did you want to test or find out?

Practical experiment: To discover if AI can have a practical impact on operational efficiency for our photographic team at the National Gallery

- How can AI help us to identify people and spaces at the National Gallery?
- We wanted to find this out to aid Photographic team in managing backlog of historically undocumented images
- Does the AI tool influence results?
- How much of the space needs to be visible to get an accurate result?

Why was this important to your organisation?

Practical experiment: Operational efficiency. We have a large volume of photographs that have been digitised but with no metadata or description of who is in the photographs. Doing this work could be a valuable resource in linking archival material

What did you do?

Practical experiment

Aligned process with wider AI discussions at Exec level.

Staff survey conducted at the same time.

AI tools tested:

- Gemini
- Copilot
- Chat GPT

Selection of images:

- Clear views of room shape
- Single wall shots
- Obscured views

What happened?

Practical experiment

Despite a relatively easy task AI may not be fit for this purpose.

Given image recognition has been around for so long it is disappointing this was not more sophisticated.

What was difficult?

Strategically:

- This project is one small part of bigger discussions
- Almost becoming routine within some applications
- Unsure if the process helped find ways to make operational impact

Practically:

- Sourcing the images from another team
- Demands on time
- AI products moving at pace

Was anything easier than you'd expected?

No

What did you each learn, as individuals?

Participant 1: Like any project there are many ways to tackle it. Given the speed of return to assess tools a really rapid first phase would have been better, with wider participation. However, demands on time is still a very practical problem, in order to assess the outcomes and learn from the process.

Participant 2: Time needed to train the AI and check results can make a task very time intensive that may mean the use of AI for some types of jobs is not suitable.

What did your organisation learn?

Enabled practical conversations across departments.

What's next?

- This project has helped the discussion to fast-track implementation of practical steps
- It fed into a wider conversation about practical use, tools and governance
- A knowledge share hub and AI champion network is now being put in place

Tips or insights about doing your experiment that might be useful for other cultural sector people in similar contexts to you:

- Get support from senior stakeholders to alleviate any concerns in the organisation about using AI tools - as they are most likely using them as well.
- Using AI tools does not equal a lazy approach to work ethic.

The Queen's Hall, Edinburgh

Emma Mortimore | *CEO*

Steve Williams | *Marketing Manager*

What was the research question behind your experiment – what did you want to test or find out?

Strategic experiment: We wanted to develop an AI policy for use, and an approach to using AI that suited us as an organisation. We wanted an authentic AI policy for organisational usage that was authentic to the things The Queen's Hall cares about as an organisation.

Practical experiment: Could we teach an AI to write event text authentically in our Tone of Voice?

Why was this important to your organisation?

Strategic experiment: We had read so much in the press and online about AI being useful or helping improve productivity. We are a small team that does a lot, we don't always have time to do the development work for the organisation that we want. We wanted to explore where AI might be useful in what areas and for what purposes, and what sort of elements should go into a policy and procedure for using AI. At the same time, as a music venue, we had distinct concerns about preserving human-artist creativity and the work of musical artists, so protecting human input was central to us.

In addition to this, we wanted to add a practical experiment. Theorising on what AI would be useful for was one thing, whereas actually using it and finding out how useful, what work would be involved and learning what the results were like would help. This would also inform the policy and procedure as we went along - to add some real world obstacles and practicalities to help inform the policy.

Practical experiment: We receive a lot of promotional text that is written in varying tones of voice, contains irrelevant detail to customers (long artist bios, for example) and is ineffective in many marketing areas (SEO optimised, passive text, not persuasive). However, editing and rewriting this copy is time consuming for the limited human resource we have.

What did you do?

Strategic experiment: We reviewed a number of Digital Maturity and AI maturity surveys for internal staff, and looked at frameworks for learning that we were given on the course. We used that to try and distill our experience on the Let's Get Real AI course into a survey for staff - covering the similar areas and starting the AI journey for our colleagues.

We devised a simple survey to find out the level of knowledge about AI in the organisation, as well as how many people were or had used it - at work or in their personal life. We also wanted to learn about people's feelings towards AI - were they excited, sceptical or fearful? We also wanted to learn where we - as an organisation - might best use AI to improve the way we do things, and use our learning to inform an AI policy for the organisation.

We asked people about repetitive tasks they did in their roles, that an AI might assist with. We also asked about creative areas.

Practical experiment: For the practical experiment, we wanted to teach an AI to write event marketing copy in an authentic way, using our tone of voice. To do this, we commissioned an agency to produce a comprehensive tone of voice guide and used this to teach the AI. We chose Claude for the AI and fed it relevant documents to give it the basis it needed to start learning.

What happened?

Strategic experiment: The survey results showed that there was a high degree of AI scepticism within the organisation. Despite our initial assumption that concern would be about being replaced by AI, the biggest concerns were around environmental impact, and humans losing skills by becoming over reliant. Data security and a general lack of trust in the transparency of AI companies were also strong concerns.

The majority of people recognised that not adopting AI presented the risk of falling behind others. Lack of staff development was also a concern for non-adoption of AI.

Most people had an interest in learning more about AI, but did not overwhelmingly believe that AI was relevant to improving their performance. Most people have used an LLM (such as ChatGPT) and were most familiar with this.

While most people had used at least one AI tool, confidence in their ability to use AI was low. Across the organisation, there was strong agreement that admin-related/workflow tasks were the areas highlighted where AI might be useful.

Practical experiment: We spent a lot of time with the AI giving feedback, explaining concepts, and suggesting alternatives to help refine its output a little more. Through teaching the AI we also realised that we were learning about our own tone of voice, and developing our ability to teach another human how to write this way. This involved gathering external information (from the internet) that might improve the text (such as relevant info for prospective audiences).

The output improved slightly over time, but only slightly, with the AI reverting to non-authentic phrasing, or coming up with copy which just wasn't right, or too generic-marketing style text. While the AI seemed to grasp the concepts we were trying to teach it very well, entering into a dialogue with the teacher and offering critiques back, the output remained below the quality that was desired.

What was difficult?

Wrestling with the different ethical problems of AI was like doing a philosophy course every moment that we were engaged with it. What was difficult was truly trying to assess the impacts, the risks and feeling very underprepared to do that. It felt like such a serious thing to contemplate with potentially very significant risks that it was difficult to settle on an ethical approach. Also producing effective prompts for the AI was difficult, and still something that needs to be worked on.

Was anything easier than you'd expected?

No. Overall it was a very challenging area to learn more about.

What did you each learn, as individuals?

Participant 1: I learned a huge amount about the ethical implications that - although I thought I was aware - I really had only skimmed the surface. I also learned about the difficulty, and the effort required to produce something effective using an AI.

Participant 2: I learnt that we don't have to use AI, but that understanding better what is out there is half the battle in feeling slightly more in control in a keeping your enemy close kind of way. I also learned that everyone is pretty much in the same boat, but that there are some really interesting and good ways to use it in the right hands. I still don't like it and have massive concerns around misogynism, bias and ethics and don't feel that governments have any better handle on it than we do - laws can't keep up with the pace of change.

What did your organisation learn?

Our organisation learned where we might concentrate our efforts in adopting AI, but mainly what the ethical issues that we would need to negotiate would be. It's challenged us to look at ways we currently work (for example, how much energy is consumed by existing processes vs adopting AI for that process.

What's next?

We will continue to experiment with providing more fundamental teaching on marketing to the AI, as well as begin training across the organisation that brings everyone else up to speed. Following this, developing a policy and use case procedure for AI, as well as having a wider conversation with the organisation about usage, ethical issues and motivations for adoption.

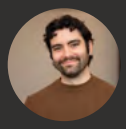
Tips or insights about doing your experiment that might be useful for other cultural sector people in similar contexts to you:

- Be curious - don't start with expectations of how it should work out and don't believe the AI marketing hype. This will help you determine how useful an AI tool will be to you.
- Prepare yourself and stay human - dealing with the ethical issues is challenging enough, but also training the AI can pull you into a strange existential space. Remember AI is not living, or human. It's just pattern recognition on a massive scale. Don't be swayed by its attempts to make itself look human.
- Work out what's really important to you and your organisation - preserving artist creativity, being mindful of environmental impact and keeping humans central to AI use were our most important things. Work out what's really important to you and let that be your guide.

Credits

Report authors:

Alec Ward, Anra Kennedy and Jane Finnis



Thanks to all project participants:

Nia Martin-Evans
Danielle Cowell
Katey Goodwin
Oliver Pitceathly
Lewis Roden
Danielle Patrick
Heather Millard
Elizabeth Llabres
Mark Pajak
Jack Evans
Rachel Poxon
Mary Bruce
Thomas Meyer
Andrea Pfeil
Kevin Carter
Trish Thomas
Becky Lees
Barbara Révész
Sarah Fuller
Lily Ackroyd-Willoughby
Robert Murray
Tony Currie
Tom Bird
Rachel Nutland
Abigail Netcott
Stephen Tolfrey
Lawrence Chiles
Lucinda Blaser
Emma Mortimore
Steve Williams

With the help and support of:

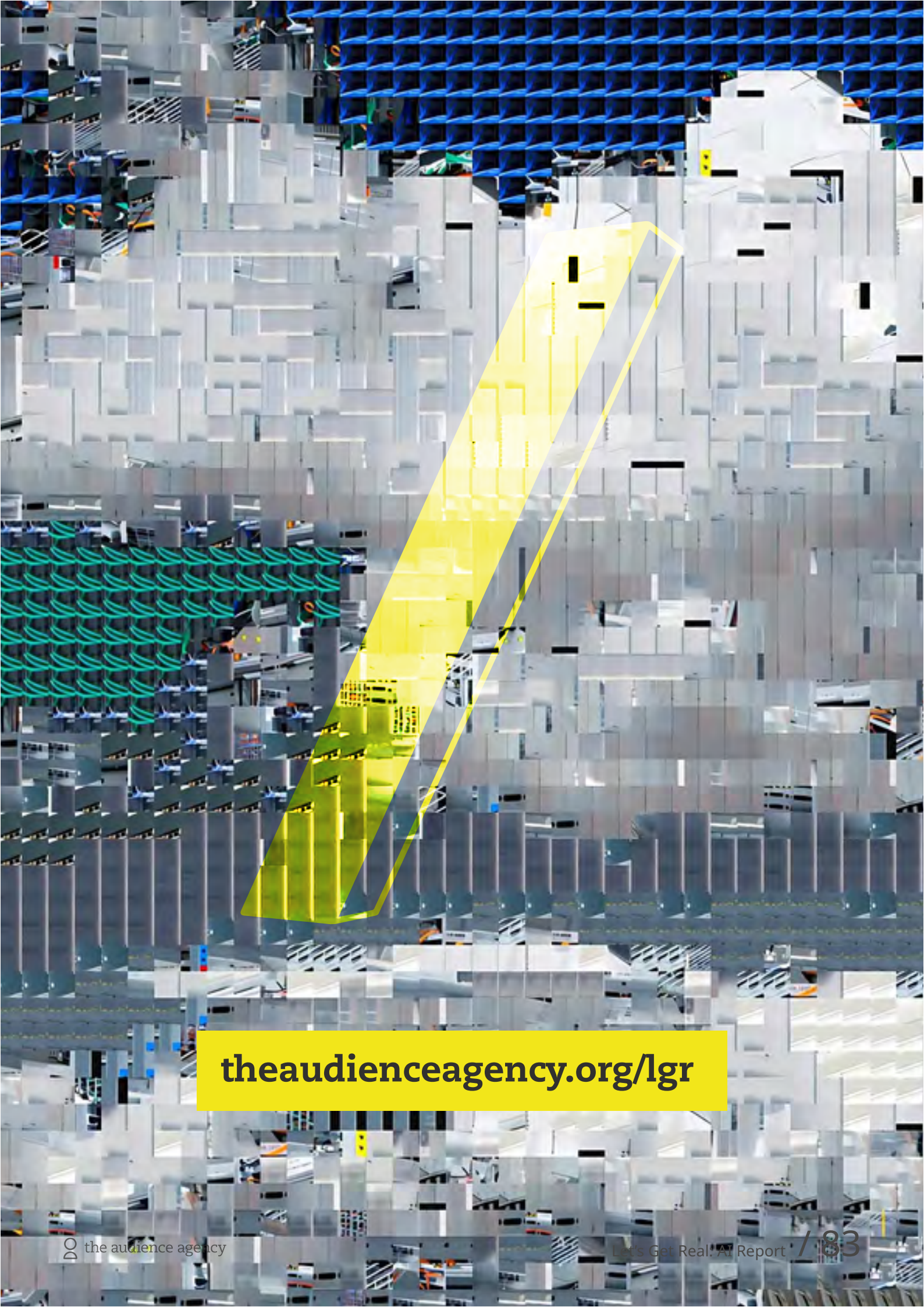
Nina Douglas-Bain
Oliver Mantell
Professor Ross Parry
Jocelyn Burnham
Stephen Miller
Rachel Oxton-King

About The Audience Agency

We are an independent research and development not-for-profit organisation whose purpose is to enable a vibrant, relevant cultural sector better able to create cultural and creative opportunity for everyone.

We provide research, services and insight which help our network of clients and stakeholders adapt for and with their communities. Our team is a collective of committed specialists with backgrounds in the cultural and creative sectors, research and/or data science who share a passion for arts and culture and its power to transform society. Our approach is insight-led and people-centred.

We work across the UK and internationally, with an extensive range of cultural organisations, academic partners, local authorities and funding bodies. We are funded by Arts Council of Wales, Creative Scotland and a range of other project funders and partners.


theaudienceagency.org/lgr