



CREATING A DYNAMIC ARCHIVE OF RESPONSIBLE ECOSYSTEMS IN THE CONTEXT OF CREATIVE AI (CRADLE)

AHRC/UKRI BRAID Programme 2024

Output 2: Workshop Report 2: Cat Royale¹

CAT ROYALE WORKSHOP: 25th March 2024

ORGANISERS: Lydia Farina, Helena Webb

¹ To cite this report use: Drury, M.R.F., Miles, O., Brundell, P., Farina, L., Webb, H., Giannachi, G., Moore, J., Benford, S., Jordan, S., Stahl, B., Perez Vallejos, E., Vear, C. 2004, Workshop Report 2, BRAID/ AHRC OPP18206, 2.

FUNDERS: Bridging Responsible Artificial Intelligence Divides (BRAID) – an Arts and Humanities Research Council (AHRC) project.

AUTHORS: Megan R. F. Drury, Oliver Miles, Pat Brundell, Lydia Farina, Helena Webb, Gabriella Giannachi, John Moore, Spencer Jordan, Steve Benford, Bernd Stahl, Elvira Perez Vallejos, Craig Vear

Table of Contents Overview 4

OBJECTIVES	4
PARTICIPANTS	5
AGENDA	6
§1: CHARACTERISTICS OF CAT ROYALE	7
§1.1: CONTEXT: FUNDING AND COLLABORATION	7
§1.2: CONTENT: MAKING AND FILMING THE ‘CAT UTOPIA’	8
§1.3: MAPPING AN ECOLOGY OF TRUST	10
THE ENCLOSURE	11
THE CONTROL ROOM	12
ARTIST’S STUDIO	12
PUBLIC CONTEXT	12
SUMMARY	13
§1.4: GOVERNANCE CONTEXT: ETHICS AND MULTI-DISCIPLINARY COMMITTEES	13
§2: DYNAMIC ARCHIVES: DISCUSSIONS AND QUERIES	17
§2.1: WHAT IS HAPPINESS?	17
§2.2: ARCHIVING CAT ROYALE	16
§2.3: PRACTICALITIES AND SUMMARY	198
§3: ANALYSING CAT ROYALE: COMPARISONS WITH JESS+	219
§3.1: GROUP A	21
§3.2: GROUP B	23
§3.3: GROUP C	24
§4: SUMMARY OF FINDINGS AND NEXT STEPS	275
§4.1: REFLECTING ON OBJECTIVES	27
§4.2: NEXT STEPS	286
WORKSHOPS AND REPORTS	286
STRUCTURE OF DYNAMIC ARCHIVE.....	28
THE DEMONSTRATOR PROJECT	29

Overview

The following report documents proceedings from the second of three workshops set up to address the Arts and Humanities Research Council (AHRC) funded project ‘CREating a Dynamic archive of responsibLe Ecosystems in the context of Creative AI (CRADLE) theme. The event was co-organised and convened on Monday 25th March 2024 at the University of Nottingham (UoN) by project lead Dr Lydia Farina, Department of Philosophy, and project Deputy lead, Dr Helena Webb, Department of Computer Science. It was attended by 11 additional team members representing interdisciplinary expertise from Computer Science and Human Computer Interaction (HCI) to English and Creative Writing, Philosophy, Psychology, Music, and Archiving.

The report contains four sections:

[§1](#) – A report of the presentation given by the collaborative team from the Mixed Reality Lab (University of Nottingham) and Blast Theory (artist group) on their artwork research case study, Cat Royale.

[§2](#) – A discussion of important themes and applications of Cat Royale’s archive which are relevant to the main research questions of the CRADLE project.

[§3](#) – A summary of group-based conversations regarding insights gained from the Jess + and Cat Royale archives which can be used to structure a dynamic archive.

[§4](#) – Recounts the key findings, directs readers to relevant sections, and summarises information for the third workshop and preliminary information regarding a follow up Demonstrator Project.

Objectives

1. Get information on Cat Royale from the project team.
2. Identify main stakeholders in this project.
3. Is the ecosystem inspired by this project compatible with RAI (Responsible AI) priorities and policies?
4. What insights can we extract on the relation between creativity and responsibility?
5. Does the project provide any insights on the relation between creativity and authenticity?
6. Were there any worries / issues during the project relating to responsibility?

Participants

Name	Initials	Role	Affiliation
Lydia Farina	LF	Project lead, UoN	Philosophy
Helena Webb	HW	Deputy Project lead, UoN	Computer Science
Steve Benford	SB	Co-Investigator, UoN and Cat Royale	Computer Science
Nick Tandavanitj	NT	Artist, Cat Royale,	Blast Theory
Alan Chamberlain	AC	Co- Investigator, Cat Royale	Computer Science
Eike Schneiders	ES	Co-Investigator, Cat Royale	Computer Science
Gabriella Giannachi	GG	Co-Investigator, Exeter	English / Archiving
John Moore	JM	Co-Investigator, TNA (The National Archives)	Emerging Technologies, TNA
Oliver Miles	OM	Research Assistant, UoN	Computer Science
Pat Brundell	PB	Research Fellow, UoN	Computer Science
Elvira Perez	EP	Co-Investigator, UoN	Psychology / Computer Science
Spencer Jordan	SJ	Co-Investigator, UoN	English / Creative Writing
Megan Drury	MD	Research Assistant, UoN	Philosophy

Agenda

Time:	Activity:	Initials:
12:00 – 12:15	Arrivals, coffee, and introductions.	
12:15 – 12:45	Project updates.	LF
12:45 – 13:30	Lunch.	
13:30 – 15:30	Presentation of key characteristics of Cat Royale.	SB, AC, NT
15:30 – 15:45	Coffee break.	
15:45 – 16:45	Analysing Cat Royale and comparison with Jess+. <ol style="list-style-type: none"> 1. <i>Who are the main stakeholders in this context?</i> 2. <i>Are there any insights on the relationship between creativity and authenticity or creativity and responsibility in this context?</i> 3. <i>Are there any insights which can contribute to the debate on the concept of responsibility?</i> 4. <i>Is the model of AI ecosystem inspired from this context applicable to other contexts? Can we create a blueprint map of AI ecosystems which is aligned with RAI principles and priorities?</i> 5. <i>What would a prototype design of a dynamic archive of Cat Royale look like? How similar or different would it be to a prototype of a dynamic archive based on Jess+? How might it relate to the idea of an open AI dataset?</i> 	Group Discussion
16:45 – 17:00	Summary findings and next steps.	Group Discussion, LF

§1: Characteristics of Cat Royale

§1.1: Context: Funding and Collaboration

SB and NT introduced Cat Royale as the product of a collaboration between the University of Nottingham's [Mixed Reality Lab](#) (MRL) which is “grounded in the disciplinary field of Human-Computer Interaction” (HCI) and the Brighton-based artist group, [Blast Theory](#) (BT), which focuses primarily on “creating groundbreaking new forms of performance and interactive art” that “explores interactivity and the social and political aspects of technology”.

MRL and BT are longstanding collaborators, and Cat Royale constitutes their most recent joint enterprise that concerns taking emergent technologies in unexpected directions to provoke the public into considering their relationship with both developing and entrenched technologies, typically with sociopolitical inflections. Though initially, MRL and BT had considered the use of Generative AI, NT told us, it was too ephemeral and abstract. To make the art more engaging, more able to pique the interest of the public (intended audience) MRL and BT turned to robotics as a way of making the virtual actual. Cat Royale was focused on the interaction between human and non-human animals with Artificially Intelligent systems.² The key themes of the Cat Royale project were: 1) the investigation of the role of care, 2) AI-managed welfare, 3) human wellbeing in a world of AI-mediated living.

In terms of funding and further collaboration, the context for the design and deployment of Cat Royale can be traced to the £30m UKRI funding of the University of Nottingham's Trustworthy Autonomous Systems (TAS) hub, which has a remit inclusive of creative programmes of artistic work. In this landmark project for the TAS hub, Cat Royale was the result of initial exploratory work with the University of Nottingham's CoBot (collaborative robot) Maker Space (CMS). Importantly for the CRADLE project, as Cat Royale included a variety of different perspectives from all collaborators at different stages, this created a complex ecosystem including a variety of stakeholder perspectives.

In terms of timelines, from inception to filming of the cats using the space, the project took 18 months. Major interim work packages included filming, editing, and presenting the film in a final, touring theatre piece. The Cat Royale project highlighted several EDI risks and considerations as it included animal participants. Notably, the ethics process was thorough and rigorous, taking 15 months to completion due in part to the interdisciplinary nature of the work, with three separate ethics committees involved. The overall concept was to create a vivid and tangible experience that could engage the public to debate several important and contested concepts such as happiness along with practical and ethical considerations attaching to using AI applications to improve happiness.

² We interpret Artificial Intelligence Systems in a broad sense to include systems performing tasks which are typically associated with intelligence e.g. decision making, memory, data processing, model formation etc.

§1.2: Content: Making and Filming the ‘Cat Utopia’

The project itself consisted in, broadly, three phases:

- FIRST PHASE: Planning, designing, and building the ‘cat utopia’ (The Enclosure, see [§1.3](#))
- SECOND PHASE: Filming and editing the three cats’ daily lives interacting with the AI-power robotic arm.
- THIRD PHASE: Disseminating the 12 daily logs ([YouTube](#)) and touring the [7-hour slow film](#).

The key question underpinning Cat Royale was ‘*can AI make us (humans) happier?*’. By using cats as proxies for human animals, the project wanted to raise questions on whether we can increase happiness using AI applications. Happiness was evidenced as increased engagement with the AI applications. Thus, artists set about building an enclosure designed for the maximalisation of feline comfort and luxury where three cats would spend six hours a day for twelve days. During this time, the cats would interact with a robotic arm that could perform a series of actions determined via an AI system connected to a computer vision system. The arm/AI system would monitor the behaviour of cats (measured against the Participation and Play scale, see [§2.1](#)) and respond by ‘learning’ which activities led to the greatest observable happiness.

To gain further understanding, we watched three YouTube video excerpts from Blast Theory’s promotional work: Screenshots from videos are shown in Figure 1

Intro day 1: Cat Royale is introduced as a ‘utopia for cats’. In this utopia, a robot arm is central to interactions in the system. The overarching goal is to ‘increase and maintain the cat’s happiness’. We were introduced to the cats – Ghostbuster, Pumpkin, and Clover – unique actors with their own stories and characteristics. Artefacts included 36 toys and games for the cats to interact with. Questions revealed a concern for responsible research and innovation (RRI) such as ‘is it best for the cats if the AI learns they love snacks, and acts accordingly by offering more?’, or ‘should the human operators intervene?’

Highlights: The highlight video from day 10 drew attention to the number of appearances of artefacts as meaningful touchpoints for narrating interaction. This evoked questions such as ‘are the cats aware of what they are doing? And ‘are their interactions significant beyond the immediate and obvious?’

End video: The final video summarised the filming of Cat Royal and included final ‘measures, such as the provision of a game ‘rating’ by the AI. This provoked questions such as ‘has the AI or the cat led the interaction in terms of provision of favourite toy?’. In total, 7500 video clips were tagged from 8 iPhones positioned around the room. The robot operator controlled the arm, and automation was

continually tweaked and altered. Anecdotally, further contextual data was captured here. For instance, Clover and Ghostbuster were observed to ‘queue to get in’ the enclosure. In summary, many voices were included in the contingency of making Cat Royale happen, such as vets and animal behavioural psychologists. In terms of the artist’s perspective, the analogy was that the cats in Cat Royale represent the human user. The enduring thought was, ‘what are the costs of instant gratification’?



Figure 1: Cat Royale Video Screenshots

AC observed that the end-products (the video logs and the film) are particularly powerful because the audience’s perspective is restricted to the 8 iPhones that had been

positioned by the researchers to document life in The Enclosure. The audience comes to feel themselves as *embodied* inside the project as though invited into the world of Ghostbuster, Pumpkin, and Clover (the cats).

What is not observable is how Cat Royale depended on the collaboration of specialists from multiple disciplines (vets, computer scientists, roboticists, computer vision specialists, and artists). What made Cat Royale work were the highly specified roles given to each member of the team as ‘moving parts’ that came together to generate this quasi-performative artwork.

Prior to answering the planned questions asked of all our case studies, we began by unpacking the various characteristics of Cat Royale in a way that set up critical discussion about ecosystems. For instance, we noted that that using AI in this case study was laborious and expensive. Behind the scenes interaction, despite being crucial to the performance of Cat Royale, is not overtly depicted in the video artefacts; however, a robot operator controlled every move of the robot arm, artists were making directions and decisions whilst the cat welfare officers were constantly watching.

When the cats were not in the enclosure, humans would clean and prepare the space for their return. Between the sessions with the cats, humans were making the enclosure safe and appropriate for them. AC underscored the importance of conceptualising Cat Royale as an embodied process, suggesting that when it comes to archiving, there is a need to cultivate a sense of ‘being there in the moment’ rather than merely recording the disembodied perspective of the 8 iPhone cameras. In focusing the archive on this aim, users of the archive can gain a different kind of interaction with the performance, where that is not a repetition of the ‘end-product’ (the film/vlogs) but an understanding of Cat Royale’s ecology *in toto* (including an ‘in the moment’ element of interaction).

We reflected on the likelihood of there being competing agendas in terms of the various expert stakeholders, and the extent to which this is captured: How many times do humans come into the space, and on what grounds? This ability to be in the room is of course exemplified by the ‘room within a room’ arrangement. Despite these critiques, we returned to the central notion that this was the cat’s own performance, which was maintained through the quick opening and closing of doors, changing and replacing of toys, and so forth. The more disruptive work of cleaning and maintenance of environment was done when cats were not present. Implications for the ecosystem emerge here in that: 1. The focus remains on the cats, but we should be clear about context and: 2. The crew is never shown but remain an established point of reference (for example on the website). This enabled response to unforeseen issues – such as destroyed toys.

Responsibility, creativity, and authenticity were mentioned as key themes for Cat Royale; trust, in particular, takes centre stage for the Cat Royale researchers. We turn to these themes in more detail below.

§1.3: Mapping an Ecology of Trust

SB provided the workshop with early sketches of [ecologies of trust](#); a framework consisting of concentric circles (see Figure 2) in which the original performance is

depicted as central, moving out through layers of ecological zones and diverse stakeholders. Cat Royale depended on a large number of human and non-human animals, and artificial agents operating successfully. By mapping the entities and environments in this manner it is possible to identify main stakeholders and relevant ‘trust trajectories’ (i.e., the circuits of trust firing within the ecosystem between stakeholders).

The resource became central to discussing what needs to be put in place so that interactions are maintained. This cartographic approach to responsibility enables the identification of key stakeholders or foci of responsibility at different ‘levels’ or ‘layers’, working from the ‘nucleus’ outwards.

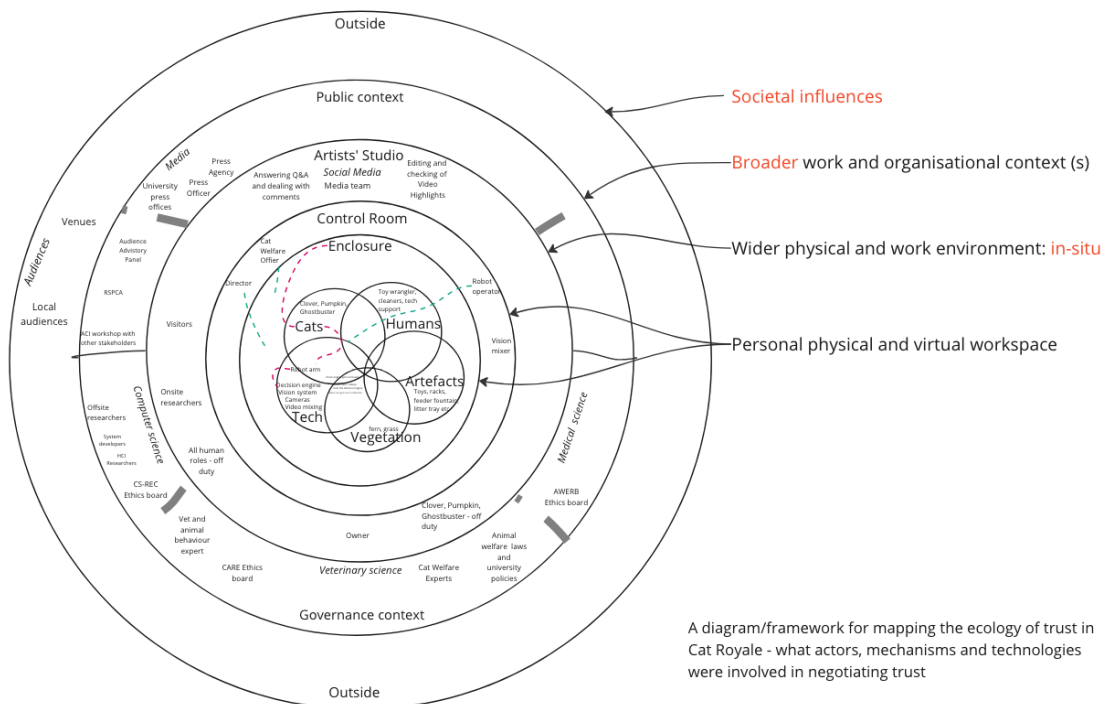


Figure 2: Ecology of Trust in Cat Royale

The Enclosure

The inmost sphere – The Enclosure – is the ‘environment’ of the primary stakeholders (i.e., the cats, humans, artefacts, vegetation, and technologies). Specifically, other than Ghostbuster, Clover, and Pumpkin, the robot arm was pre-programmed to move and manipulate toys, with 8 iPhones capturing interaction inside of this ‘layer’. The Enclosure was the space that could be affected by the actions taken from the next ‘layer’, the Control Room, which housed the director, welfare officer, robot operator and vision mixer.

The Control Room

As indicated already, in The Control Room, human operators have a responsibility to ensure the effective management and safety of those within The Enclosure. This included the robot arm operator, the cat behaviour specialist, vets, computer scientists, and director. The systems in place allowed for the monitoring and overriding of the AI decision engine and the operation of the robot arm (e.g., to prevent over-feeding, underfeeding, feline stress) which were sometimes guided by the decisions and recommendations of the cat behaviour specialist (measuring feline scores against the play participation scale).

Tasks such as the completion of a cat stress score (participation in play scale) were attended to, yielding a measure of ‘happiness’ which could be fed into a decision engine resulting in recommendations. There were two modes for the algorithm; random and exploitation (using things learned by the algorithm). Critical questions at this juncture include ‘who should handle and hold executive decisions over recommendations?’, and what objectively constitutes ‘too stressed/too relaxed’?

The first showing of Cat Royale was at the Brisbane festival ([Blast Theory](#), n.d.) which was the closest to the live performance (24 hours delayed). We reflected on this latency, both from the perspective of the audience, and from the perspective of any stakeholder who might wish to analyse the original performance.

Artist’s Studio

In terms of the Artist’s Studio, various stakeholders were present including visitors, the onsite researchers, the owner, the cats, and generally all ‘off duty’ human roles. This space is important because many different agendas were operating in tandem in an attempt to innovate in a responsible manner. Practically, a meeting took place in this space once a day to discuss welfare, daily agendas and what – if anything – should be changed. This element was not within public view and represented a ‘boundary within boundaries’.

The Studio was used to reunite the three cats with their owner. It also allowed space for ethnographers and other researchers to interact with the project team. Social media teams also present in the Studio managed the public face of the project

Public Context

The public context represents the first domain located completely ‘outside’ of the Cat Royale enclosure. Stakeholders encompassed third-party actors such as social media and overseas audiences. This space crucially includes the gallery and the theatre and audiences of current exhibition.

In terms of the governance context, there were three components: Discussion in computer science, in the veterinary school, and the cross-university animal committee (see [§1.4](#)). This should be viewed as a positive, constructive process. However, and with regards to creating a dynamic archive, there is a question of where this process sits. Clara

Mancini's (2011; 2017) work regarding animal computer interaction (ACI) and the distinction between *contingent and mediated consent* is of relevance here. Contingent and mediated consent is a concept denoting a model for garnering consent in the context of human-animal-computer interaction research studies. The *mediated* dimension of this concept refers to the way in which consent is given on behalf of an animal (e.g., a companion animal, animals in zoos, or animals in shelters) by their relevant human companion and/or their legal guardian (Mancini 2017, 230). The *contingent* dimension refers to the conceptualisation of the act of giving consent not as a singular event, but a sustained process whereby it is possible for the relevant human mediator to withdraw consent on the behalf of the animal participant at any time (Mancini 2017, 230).

Summary

In view of the various processes across these domains that constitute the Cat Royale ecosystem, it was suggested that we might use the HCI model of trajectories to conceptualise a dynamic archive tracking, for example, a trajectory of 'trustworthiness'. Given that documentation exists in the artist's studio and the control room pertaining to trustworthy interactions - such as in risk registries and contingencies for adverse events - trustworthiness can effectively be retraced. However, we also considered whether trajectories are sufficient for a truly 'dynamic' archive: While individual artefacts on which trustworthiness can be considered contingent are situated within ecosystem boundaries, trustworthiness likely transcends these boundaries. Essentially, the challenge here is how to retain a sense of original boundaries while enabling future interactions to 'cut across' domains. As such, framing Cat Royale on the basis of a single trajectory may be too linear.

The final 'ring' or 'layer' of the ecology of trust set out by the Cat Royale team and presented to the CRADLE team by SB and NT concerned *governance*. The governance context concerns the responsibility of the Cat Royale team to their funding bodies and to institutional fora for ethical review. In the next section we discuss the key themes, difficulties, and means of resolution as recorded both during the workshop and in [Benford et al. \(2024\)](#).

§1.4: Governance Context: Ethics and Multi-Disciplinary Committees

A key ethical issue was *consent*, as noted above. Though further discussion turned to the question *who benefits?* (i.e., is it the human beings learning a lesson, or the cats having a utopic experience?); we can also ask who is at risk (the cats, primarily, insofar as the robot arm may malfunction and they may experience significant distress or fight each other during the project).

Owing to the multi-disciplinary character of Cat Royale, which was composed of artists, ethnographers, Human-Computer Interaction researchers, animal behavioural specialists, and robotics experts, the project required multiple ethics committees, each rooted in a particular field of research ([Benford et al., 2024, 8](#)). As such, SB and NT highlighted that governance bodies had difficulty evaluating Cat Royale as it defied

assessment from a single disciplinary perspective and had to negotiate “underlying differences between the diverse disciplines that became involved” (Benford et al., 2024, 8). This difficulty was accelerated by the involvement of animal participants. The tensions that emerged derived from the often ‘invisible’ values that determine how researchers in different fields approach the question of what constitutes “valid knowledge and [legitimate] methods for attaining it” (Benford et al., 2024, 8). See Fig. 5 for Benford et al.’s (2024) diagram of the relation between the ethical committees, epistemic assumptions, and Cat Royale.

The key institutional entities that Cat Royale navigated were:

1. *Animal Welfare and Ethical Review Body* (AWERB)
2. *Computer Science Ethical Review Committee* (CSREC)
3. *Committee for Animal Research Ethics* (CARE)

The key epistemological tension concerned the clash between what may broadly be termed *scientific research*, which structures knowledge according to the values of *positivism* or *objectivism*, and *artistic research* (or research through design) that structures knowledge through subjective and interpretative means (Benford et al., 2024, 9). AWERB identified a lack of clear aims or research questions as a concern. In response, the Cat Royale team highlighted that their epistemology and methodology was more aligned with explorative co-design than with the testing of hypotheses regarding feline behaviour and psychology through experiments involving autonomous systems (Benford et al., 2024, 7). CSREC, by contrast, focused on the “human-facing aspects of the project” (Benford et al., 2024, 7). Rather than the specific design methods, the concerns raised were more directly related to the use of animals in a research project and the potential damage to the reputation of the University of Nottingham. The decision was made to more precisely detail the “artistic framing of the project” and to deliberately evoke ambiguity to prompt reflection on human-AI interaction, whilst underscoring the provision of cats with “the most positive enriching play experience” the project team could provide. Finally, CARE focused on the wellbeing of the feline participants providing the project team with a Cat Welfare Protocol. Much of CARE’s requests for clarification concerned the habituation of the cats to the enclosure, the ability of cats to leave the enclosure if showing signs of distress, and whether cats should be allowed to leave should they show signs of boredom and/or fatigue. The project team uniformly agreed that should feline distress occur, they should be removed from the enclosure, but that should signs of boredom and fatigue be displayed the cats ought to be left in the enclosure to prevent them from wandering through spaces that were “narrow, dark, full of cables and computers, and [...] a door to the wider world that might be left open for ventilation” (Benford et al., 2024, 8).

Across all disciplinary/epistemic domains – the methodological concerns of AWERB, the reputation concerns of CSREC, and the feline concerns of CARE – the key to unlocking ethical approval was being able to answer the question, *were the cats being used to benefit humans, or humans to benefit cats?* As the project team began to consider who

was at risk and by which factors, the different disciplinary perspectives regarding legitimate approaches to acquire valid knowledge focused on identifying benefits, rather than risks per se. This was largely owing to the disparity in clarity: risks are simpler to identify and mitigate, whilst benefits can become seemingly amorphous when a multi-disciplinary project is viewed from singular disciplinary perspectives.

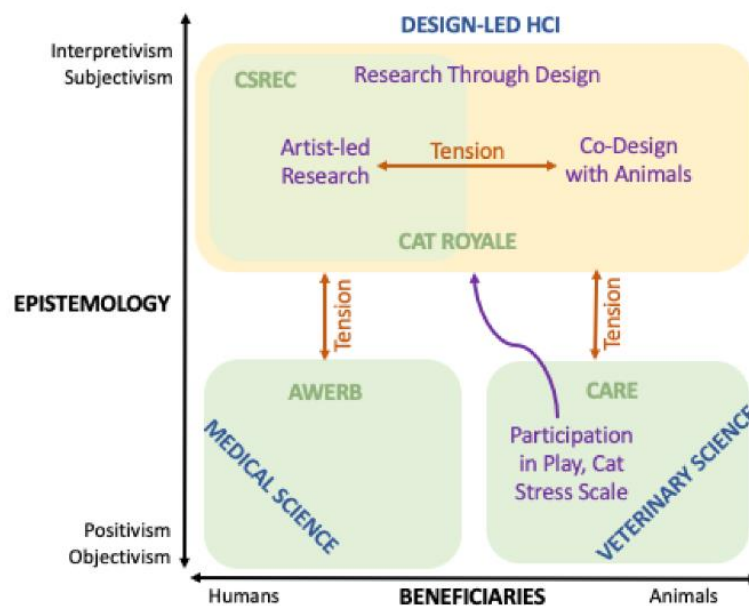


Figure 3: Epistemological Approaches

Articulating benefits was fundamentally linked to good explanations of project research questions, and specifically, the methods involved in data collection. This was because there was a reliance on tools and measures that are traditionally used in the veterinary science domain, and as such, the team became accountable for their (appropriate) usage. Epistemological differences can clearly create challenges in understanding and communicating the benefits of the project and are illustrated in Fig. 3.

Moreover, it was felt that reputational risk posed a significant factor to mitigate, and this was shared across all three ethics committees and the University of Nottingham as an institution, particularly as many had no previous experience of animal computer interaction (ACI). Consideration of reputational risk was handled as a transcendental issue that focussed on external factors beyond the immediate research environment and was as much about perceived risk as actual risk: The perception that a robot ‘might’ injure a cat was thus countered with assurances of it being ‘not that kind of robot’ in the sense that the robotic arm was always controlled by a human being

Ethical considerations concerning responsibility were broadly framed as either a question of responsibility for the cats, or responsibility for the audience. This dichotomous framing created an interesting tension, as responsibility for cats might not

align with responsibility for humans, and vice-versa . Applying this insight to the structure of a dynamic archive, this would mean that a responsibly constructed dynamic archive should operationalize this tension, by enabling continued use and interpretation by multiple stakeholders.

§2: Discussion and Queries

§2.1: What Is Happiness?

Cat Royale, as mentioned, was designed to produce a 'utopia for cats'. The conversation turned towards how researchers approached the subjective motion of 'happiness', slippery as it is for humans, let alone animals. The happiness levels of Ghostbuster, Pumpkin, and Closer were inferred by means of the participation and play score which was both narrative and provocative in function. However, it was noted that the use of the happiness scale was useless without the contextual grounding of interaction, baseline measures or scientific comparison. Thus, in this artistic context, the happiness score functioned also as a provocative vehicle for reflection. In comparison with other animals, cats have a cultural status as the self-determined domestic creature par excellence, and hence their participation is likely to signal desire, will, or enthusiasm rather than seeking, fishing, or placating an owner or overseer.

The AI algorithm used in the project operates on two modes: 1) Experimentation, where it essentially chooses activities and artefacts at random and sees how they play out with the cat. 2) Exploitation, where it uses previous experience to try and encourage more play. This also included choosing a different kind of activity for each cat, so it would go through a cycle of choosing different activities for each cat in turn. Separate scoring was created for each of the cat's current state and this was used as part of the context when for activities were chosen.

§2.2: Archiving Cat Royale

Moving towards a conceptual archive for the Cat Royale ecosystem, SB talked the group through [a mind-map depicting the various environments and materials of Cat Royale](#). These included the enclosure, the code for the robotic arm, data, artists' own video, operational documents, design documents, Zooniverse activity, proposals and funding, social media and press footage, history of touring, the artwork, and research materials. SB highlighted that the collation of materials can be of use for future research into (and revised and improved iterations of) Cat Royale. Postgraduate student, Cao, is developing a [computer vision system](#) that is able to identify *where* the cats spent their time as a post hoc development to the capacities of the original vision system which was capable only of recognising the individual cats.

It was highlighted within the conversation that the role of archival material as impetus for future projects is not the defining mark of dynamic archives; indeed, static archives tend to be deployed in just this way. The question, 'how might we make archival data accessible beyond just 'a large folder of data?', was addressed with the proposal that in addition to archiving 'Cat Royale' for its innovative use of AI systems, that dynamic archives themselves *make use of the capacities and capabilities of AI systems*.

The use of Artificial Intelligence need not be limited to the artwork being archived (in this case, Cat Royale), but it was noted during the discussion that AI could:

- (1) Facilitate accessibility.
- (2) Increase impact (beyond, for example, journal articles), and;
- (3) Give a static archive the kind of dynamism CRADLE is interested in.

By structuring the material curated for the archive with taggable keywords it would be possible to allow for an AI program to perform complex searches and draw from a wealth of data the specific content the user is looking for (e.g., in the documentation) and also provide relevant video-clips. A difficulty regarding the viability of a semantic search was raised given that much of the data may be largely visual or non-text-based in addition to the polysemic nature of data resources (for example, ‘entertainment’).

As regards Cat Royale, GG observed that insofar as Cat Royale is unique in that it is the first artwork of its kind to combine animal computer interaction (ACI) with AI, if archived dynamically, this artefact can be operationalised to preserve the way the project was undertaken, how it was approached, what deliberative and decision-making processes were used and their outcomes. One of the aims then this dynamic archive, could be to *encourage a positive response towards autonomous systems* (i.e., facilitate *trust*) that counterbalances the largely negative perspectives that are so frequent within the mass media coverage of AI applications and autonomous systems. A dynamic archive of Cat Royale by showing how to apply AI responsibly in this context, could function as a cohesive deployment package rather than as a snapshot of a past project.

When looking to archive works (such as Cat Royale) there are some metadata frameworks for archives and collections mapping on to computer science that can be instructive for the development of a structure of a dynamic archive (relevant to the question of the role, purpose, or affectivity of the archive) such as the Tate’s [Live List](#) (superseded by the [Documentation Tool](#)) and [Archiving Net Art](#). These resources contain extensive information that can help to understand how performative and conceptual artworks can be preserved for and reactivated in the future. These guidelines draw on “notions of authorship, authenticity, autonomy, documentation, memory, continuity and liveness” and can provide a good basis from where to develop a comprehensive framework for responsible use of AI in archiving

One pressing problem for archiving large-scale art projects such as Cat Royale concerns *where* the archived information is to be stored. JM suggested that web archiving is used in terms of .gov and social media domains. The [National Archives’ best practice document](#) was discussed, which offers clear guidance on the preservation of digital artefacts of value.

Collecting information in line with these models may provide the means to give the dynamic archive greater potential by shifting the interest in gathering relevant information from a static archive to a living collection. This discussion concerning the *audience* in conjunction with the question of using *AI* to bring the archive ‘to life’ led to questioning the *practicalities* of such a project, and what might be termed the *practical values* of dynamic archives. These important questions can be included in a Demonstrator project focusing in on responsible use of AI in archiving.

§2.3: Practicalities and Summary

To summarise this section, metadata and indexing – as well as raw materials – play a crucial role in making artworks using AI applications accessible and searchable. We should also consider tools for different audiences, for example, semantic search. This presented a dilemma of snapshots (lots of data) vs. retrospective search. Moreover, there are cultural heritage standards that differ from social media standards. A contemporary example was given of the Flickr Foundation (JM) exploring the possibility of ‘Creative Archives’ that can [last 100 years](#), surviving their host organisation.

We turned to discussion of the structure of a dynamic archive, and queried whether the ‘dynamic’ element is, or should be, a provocative element. We considered value-based structuring of information systems, and the example of distinction between types of value categories; for instance, hedonic (sensory) and eudaimonic (extra-sensory) values in fast-moving consumer goods (FMCG) personalization. We considered user-based tagging of positive-negative values, based on questions such as ‘what is important in this data for me?’ Moreover, we discussed provenance data as a potential source.

Cat Royale is unique in that it is one of the first artworks including interaction of non-human animals with AI applications, and so we asked how we might ensure this element of originality and innovation is retained in the archive. We suggested that Cat Royale could be used as an effective ‘tool’ to direct a positive narrative towards interacting with AI applications and autonomous systems, particularly as it contrasts with popular media narratives of negative ACI. Thus, it is exiting and interesting in that it challenges public perceptions. We considered the potential for Cat Royale to act as a catalyst for other artists more broadly.

In terms of the practicality of using AI applications currently, many AI applications are designed to be specific (and are valued for being specific) short of ‘general AI’. In addition, AI applications work best when dealing with companies with a long, online existence. However, anything created over the past 3 years is inherently poor due to a restricted remit of AI tools such as GPT (prone to hallucinations), and of course, there is no offline remit.

At this point we also discussed values, with the example that medical research serves a utilitarian value beyond that which directly benefits the individual research participant. In the same way, we should consider which parts of the archive are of most benefit to future stakeholders, and therefore perhaps thinking about archives in valued terms is practically useful. This includes how we might design a curated archive on stakeholder terms, as well as accommodating how artists want their work to be represented in an archive. Central to this was the idea that sometimes, a constellation of artefacts which might seem unrelated to each other are in fact integral to a new use case.

§3: Analysing Cat Royale: Comparisons with Jess+

The group was split into three sub-groups to consider how Cat Royale might inform the prototype dynamic archive, as in the previous workshop 1 (JESS+). This ensured a thorough discussion and analysis of Cat Royale such that underlying features and functions might become more explicitly connected with dynamic archiving, rather than remaining static qualities of a closed ecosystem. The below subsections have been kept separate from one another to preserve the organic nature of the discussions and the individual development of each group's trajectory during this section of the workshop. Each author made use of their own writing style to convey the discussions that they recorded.

GROUP	INITIALS	SECTION
Group A	PB, SB, SJ	§3.1
Group B	MD, LF, GG, NT, HW	§3.2
Group C	OM, JM, EP, ES	§3.3

§3.1: Group A

Discussion of how to reveal different perspectives in the context of archiving multi-species artistic performances. TA dynamic archive should capture the interaction between humans, AI, and cats. This archive should include not only the performance (in the form of video and other materials), but also documentation of the human labour involved in creating the work.

There is a question of granularity and detail in the archive – potentially a balance between practical work involved in tagging, and otherwise creating metadata and creating enough contextual material to allow use by all interested parties. Does responsibility relate to granularity?

Perspectives must include *stakeholders* who wish to access an archive e.g. animal behaviourists, computer scientists, artists, cultural historians and so on, but also perspectives of participants in the case study.

Discussion focused on what an AI 'perspective' on the performance might look like in the context of interaction between humans, cats, and AI. At times, there was an interaction between cats and AI which humans did not understand. This may be revealed by an understanding of the AI and the perspective it is taking, including making mistakes. At any point in time, the AI has a model of the context (perhaps because of computer vision) which might include merged objects, multiple animals and other "confused" views. Representing the AI view through the use of probabilities may perhaps be useful. Archiving the decision processes of the AI might facilitate discussion such as "how would the final product have looked like if we had decided this instead of that?". The archive may be used for mapping future things, but also for looking at what could have been had

different decisions been taken by cats, AI or humans. This may reveal the importance of these decisions and these key shifts made at particular moments in time. This notion is illustrated by the images in Figure 4 and Figure 5.



Figure 4: Alexei Efros: Why Computer Vision is a Hard Problem for AI



Figure 5: Alexei Efros: Why Computer Vision is a Hard Problem for AI

§3.2: Group B

MAIN STAKEHOLDERS. Drawing on SB's presentation of the [ecology of trust](#) that is capable of charting the trust trajectories between the main stakeholders, group members were clear that identifying the key stakeholders within The Enclosure could include not only Ghostbuster, Clover, and Pumpkin, but also the robotic arm, the vegetation, and other non-human actants. LF asked NT whether during Cat Royale it had become apparent that any main stakeholders had not been accounted for (i.e., lack of oversight). In response, NT highlighted that, as a multi-media artwork, anyone who engages with artwork – as a consumer, critic, or producer – *and* who has an interest in or affinity with cats or AI applications is likely to find themselves influenced in one way or another. One of the advantages of the ecology of trust is that *it allows for the individual to locate themselves on the map, identifying themselves against that field of relations mapped out by the archive*, assuming that they do find a field they identify with.

LF and HW inquired as to whether any team member had removed themselves from the project. NT clarified that those that had left (e.g., the cat behaviour specialists and different computer vision specialists) did not do so because of ethical qualms or because of a lack of faith in the project. During the project conflicts arose, but these were largely the product of an interdisciplinary team (e.g., animal behaviour studies, computer science, robotics, etc.) working within the more fluid context of an artwork. This precipitated ethical *and* creative differences. Part of the role of the director was to allow for these differences as they were part of the creative process without losing overall control of the project itself.

AUTHENTICITY, CREATIVITY, AND RESPONSIBILITY. Cat Royale, as stated in [§1.1](#), was the product of ~18 months of multidisciplinary interactive teamwork. NT highlighted that the kinds of visions or deliverables that are familiar to academics or those working within well-defined fields (unlike Human-Computer Interaction, see [§1.4](#)) are not always compatible with the artistic process which is predominantly *explorative*. The kinds of benefits and deliverables that are to be derived from a work are not plainly identifiable from the beginning but emerge in due course. For example, the capacities and limitations of the computer vision process were not fully determined until the project was fully underway; multiple computer vision specialists left to work on other projects, and the feline behavioural expert also took up other work. Consequently, the vision system could not identify the emotions of the cats from their faces, though it was able to identify the cats and locate them in space.

As a creative endeavour, Cat Royale proved to be illuminating regarding the use of live non-human actors within an artwork. When we see only the final artworks we fail to see the points of analysis – the decision-making processes, the rationale for human

intervention, but also the motivation to use cats in the first place, the choice behind that specific robot arm, and so on. A great deal of complexity is smoothed over when the final artwork is taken to be a finished product rather than the constantly evolving product of an inherently dynamic process of creativity and coalition between multiple moving parts.

CONCERNING PROTOTYPICAL DYNAMIC ARCHIVES. The idea emerged that the archive itself becomes dynamic or something in use, in process when it's not understood as a repository of video and text files, but something that has its meaning structured by our tagging of terms and implementation of a semantic search. GG commented on the rhizomatic nature of the dynamic archive as something that grows and spreads from multiple points rather than a single root. Similar questions identified in relation to Jess+ (e.g., How do we tag the material? What kind of AI applications do we use to search or access it?, etc.) emerged in the context of Cat Royale.

GG raised the limitations of using a Large Language Model (LLM) like ChatGPT 3.5 in the dynamic archive. The difficulty concerns the fact that when not trained on the relevant data, then this AI system will begin 'hallucinating', connecting parallel or totally incorrect information to the prompt itself in absence of sufficient quantities of data. The question is, when fed the right data, would the AI system itself be able to string together the kind of end-product that we are interested in for the dynamic archive? It was suggested that if we could feed the videos into a multimodal LLM, then this data could be used as its subsystem enabling the end-user to search for clusters of terms or questions. What would take this vision of a dynamic archive further, it was added, concerned *reflection on the process*, for example, being able to ask the question 'How was this put together?' to the archive and receiving a unique explanation of the creative process or environmental design, and so on.

As regards the Tate Live List and Documentation Tool, the driving idea that emerged from the discussion of this group focused on modifying the idea underlying existing metadata frameworks. It was noted that interviews are a common form of archiving, and especially so when it comes to archiving works of a performative or transitive nature; the Tate Live List has a focus on prompts such as providing "[a description of the work for someone who has never seen it before](#)". a LLM can be deployed such that it is able to [mimic the voice of the artists](#) and researchers (e.g., NT, SB, AC, and others) and speak the kinds of outputs that the LLM generates in response to prompts that initiate a search of the multimedia materials selected for archiving. Furthermore, beyond the recreation of relevant investigator's voices (and even likenesses), we discussed whether an AI-generated representation of the perspectives from the cats – Ghostbuster, Clover, and Pumpkin – could be used to provide access to a standpoint not otherwise available, not even to the researchers themselves.

§3.3: Group C

Q1. WHO ARE THE MAIN STAKEHOLDERS IN THIS CONTEXT?

Group C considered three main stakeholders: the public, the researchers and funders, and the cats.

In terms of the public, it was reasoned this is likely the most impactful group, owing to Cat Royale's contemporary publicity and current exhibition presence, such as at the National Science Gallery. Given this stakeholder group, the question 'how has Cat Royale changed perceptions of AI?' is best answered in this broadest sphere of influence drawing on implicit and explicit feedback.

In terms of the researchers and funders, it was noted that this relationship should be made to be as transparent as possible. It was reasoned that this is because researchers and funders have a responsibility to acknowledge and demonstrate that AI is not an intrinsically automated process. Instead, it takes significant collaboration and a budget, in this case, of £250,000.

Finally, and in term of the cats themselves, we reasoned that a good question was 'should we recreate "utopias for cats"?' Anecdotally for instance, it was reported by the owner that the cats had a tighter bond and were more inquisitive because of participation in Cat Royale. Should such evidence 'speak for' the cats alone, should it be taken in conjunction with other measures such as the cat happiness scores, or 'happiness'?

Q2. ARE THERE ANY INSIGHTS ON THE RELATIONSHIP BETWEEN CREATIVITY AND AUTHENTICITY OR CREATIVITY AND RESPONSIBILITY IN THIS CONTEXT?

It was suggested that evidencing creativity and authenticity may be contingent on much of the metadata resources documenting adherence to various standards (for example, the cat happiness score). While authenticity in this sense was not complex from the artist's point of view, the fact that the cats could not be consulted and asked directly for consent should be acknowledged.

In terms of creativity and responsibility, this was noted as being far more effortful to evidence. Conversation focussed on how responsible use of AI in Cat Royale needed to go beyond evidencing 'benefit [to] the species [and toward] benefit of the individual cat'. While audience advisory board meetings with 15 members of public demonstrate an ongoing commitment to responsibility, again, the 'cat perspective' remains a matter of inference.

Q3. ARE THERE ANY INSIGHTS WHICH CAN CONTRIBUTE TO THE DEBATE ON THE CONCEPT OF RESPONSIBILITY?

For an archive to be dynamic in this way, it was suggested that stakeholders should consider what is being added to the archive when we talk about responsibility. These additional components were talked about in the abstract, such as 'Accessibility and Transparency' (in terms of taking precedence over reproducibility); Expertise (who defines terms and on what authority?); and *Granularity* (to what level are artefacts within the archived labelled and indexed?).

More practically, it was suggested that any dynamic archive needs to be structurally conceived as a graph or *relational model* showing relationships between stakeholder choices and nodes. Still, questions of knowing what needs to be archived and why?

remain, as well as a consideration of homogeneous versus heterogeneous stakeholder priorities and consensus. An emergent theme here concerned the question of ‘who has a vote re what is important?’ in the eventual archive.

Q4. IS THE MODEL OF AI ECOSYSTEM INSPIRED FROM THIS CONTEXT APPLICABLE TO OTHER CONTEXTS? CAN WE CREATE A BLUEPRINT MAP OF AI ECOSYSTEMS WHICH IS ALIGNED WITH RAI PRINCIPLES AND PRIORITIES?

The use of public money and issues of who should have access was cited as a common practical challenge in archiving, and consequently one which Cat Royale might contribute to informing. The issues of funding and access also inspired consideration of *maintenance*, in terms of the price of maintaining the archive beyond the original ecosystems. As well as financial considerations for an eventual model, the *environmental impact* was also considered, with *governance* in this area cited as often the biggest challenge (for example, at the National Archives). The use of an agile approach to developing RAI with a view to generating user personas was given as a practical means of addressing these and other challenges.

Q5. WHAT WOULD A PROTOTYPE DESIGN OF A DYNAMIC ARCHIVE OF CAT ROYALE LOOK LIKE? HOW SIMILAR OR DIFFERENT WOULD IT BE TO A PROTOTYPE OF A DYNAMIC ARCHIVE BASED ON JESS+?

Envisioning a prototype dynamic archive led to discussion of two conceptions based on previous works – concentric circles and trajectories - which were not necessarily mutually exclusive.

Cat Royale initially presented a more complex challenge than Jess+ in terms of managing a diversity of stakeholders to include ACI as well as HCI. It was felt that the presentation of the Cat Royale ecosystem via concentric circles afforded a useful framework for making sense of artefacts, in terms of framing their original context and proximity to stakeholders. Considering the ‘dynamic’ requirement, the concentric circles framework was provocative in terms of considering what might be transcendent across contexts, and therefore potentially operationalizable, such that the ecosystem ‘as a whole’ is rendered intelligible for future use, and to any stakeholder regardless of location ‘within’ or ‘without’ it.

Similarly to Jess+, the idea of user trajectories was considered as a means by which the archive could be both presented and operationalised for ongoing use.

Towards an emerging consensus?

To summarise this section and draw together some key threads, we have reviewed the discussions of each of the three groups and determined two critical points for consideration: (i) conceptual and (ii) logistical.

Group A and Group B’s discussions were particularly conceptual in focus, and the defining consensus apparent in both conversations was the potential for an Artificial Intelligence system to push an archive over the threshold transforming a static

assortment of data (e.g., textual, visual, audio, etc.) into an interactive system capable of responding to the user of the archive and curating responses to key inquiries. In short, the answer to the question, *how do we make an archive do things?* may be to bring an AI application into the fold.

Group A's discussion centred on the fruitful consideration that an AI perspective could enhance an understanding of the project's nonhuman stakeholders; in other words, by incorporating the outputs of the computer software it could become possible to allow human users of the archive to gain an understanding of the relation between the feline participants and the autonomous system. Group B's conversation, on the other hand, focused on the possibility for the AI system to parse the breadth of data in response to queries such as "show me every time the cats played [game]" or "how many times did the cats receive [treatment]?". More importantly, the use of AI could allow for the voices of the artists to be used to narrate answers to questions, perhaps the paradigm example here would be "tell me which of the decisions made by the Cat Royale team were definitive of the final artwork?". There are two key foci in the conceptual discussion: (i) using AI to enhance an understanding of stakeholder relations, and (ii) using AI to enhance the experience of the user of the archive.

Group C's discussion, in contrast to the prior two groups, gave a lot of attention to the logistics of dynamic archiving. There are two categories to consider: (i) the cost of archiving process (including personnel), and (ii) the storage location of the archive. These two challenges present practical limitations that must be overcome for the *implementation* of the structure of a dynamic archive, and hence ought to be primary research questions in the Demonstrator project.

§4: Summary of Findings and Next Steps

§4.1: Reflecting on Objectives

Structured by [six outlined objectives](#), the workshop was directed towards, broadly, two aims. First, *the analysis of Cat Royale and its prototypical archive*, and second, *the analysis of this prototypical archive in comparison with Jess+*. [§1](#) of this report detailed the key features of Cat Royale as an artwork integrated with an AI application (autonomous system). Crucially, [§1.3](#) and [§1.4](#) were dedicated to the explication of the ecology of trust model and the ethical constraints and affordances that can be used to chart the ethical relations of the actors involved and the influences of their contexts. Importantly, [§1.4](#) highlights what is likely to be an increasingly common problem for multidisciplinary projects making use of autonomous systems.

[§2](#) and [§3](#) were directed towards analysing in detail the contributions that the prototypical dynamic archive of Cat Royale makes to the conception of a dynamic archive aimed at by CRADLE, and the comparative analysis between the structure of a dynamic archive of Jess+ and that of Cat Royale, respectively. [§3](#) in particular highlights the different conversations had during the workshop which aim at illuminating the relationship between creativity, responsibility and authenticity, and the potential practical dimensions of a dynamic archive.

§4.2: Next Steps

Workshops and Reports

The workshop in review is the second of three multi-disciplinary workshops during which project members discuss research projects that make use of Artificially Intelligent systems in creative contexts. The [first workshop](#) (14/02/2024) concerned:

- (1) An overview of archiving and documentation in the context of mixed media art,
- (2) A presentation on Jess+ as the first of 8 case studies analysed by CRADLE, and
- (3) An initial discussion of what the structure of a prototypical dynamic archive would consist of.

The second workshop (this report) focused on:

- (1) A presentation of Cat Royale as the second of 8 case studies analysed by CRADLE.
- (2) A discussion of Cat Royake's prototypical model of a dynamic archive.
- (3) An analysis of dynamic archives in light of the discussions of Jess+ and Cat Royale.

The third workshop (08/07/2024) will be held at the National Archives in London. Agenda pending. Though it will focus on the five following additional case studies:

- (1) [Embodied Trust in TAS: Robots, Dance, Different Bodies.](#)
- (2) [BlueSkeye AI \(Specific Project/App pending\).](#)
- (3) [TARICS](#)
- (4) [Rider Spoke – Blast Theory](#)
- (5) [The Future Machine](#)

The reports generated for each workshop will form the basis for a co-authored journal article as the primary research output of the CRADLE project. The article should discuss the structure of a dynamic archive foregrounding the theme of responsibility as it pertains to the use of AI systems in creative practices and projects.

Structure of Dynamic Archive

As mentioned, the CRADLE project has identified the structure of a dynamic archive as a key output promised by the project team to the funding body (BRAID/AHRC).

This is discussed in detail in §2 and §3, though at the outset of the workshop LF highlighted three key characteristics:

- (1) A dynamic archive is continually evolving; in a sense it is alive.
- (2) It is iteratively developed (by others engaging in repetition, by way of the archive itself).
- (3) The modification of its composition over time.

The discussion of Cat Royale and its analysis in conjunction with Jess+ yielded key insights. These are succinctly reviewed in [§2.3](#) and developed in various directions in [§3.1](#), [§3.2](#), and [§3.3](#).

The Demonstrator Project

LF raised the topic of the demonstrator project. CRADLE is one of ten 6-month scoping projects funded by the AHRC program, Bridging Responsible AI Divides. The aim of this scoping project is to work towards securing funding for a longer term (up to 36 months) project. The deadline for submitting proposals for the demonstrator project (27/06/2024) falls prior to the completion of the current scoping project. One of the aims of the scoping project is to identify key research questions which will be the focus of a larger demonstrator project.