TOWARDS
A NATIONAL
COLLECTION

UKRI | Arts and Humanities Research Council

SECOND REPORT
DISCOVERY PROJECTS

**Transforming Collections:
Reimagining Art, Nation
and Heritage**

University of the Arts London | Tate

MAY 2023

Table of Contents

Authors

susan pui san lok (UAL); Mick Grierson, Rebecca Fiebrink, Anjalie Dalal-Clayton, Athanasios Velios (UAL); Christopher Griffin, Hannah Barton (Tate); Jon Gillick, Ananda Rutherford, Charlotte Webb, Kit Bower-Morris (UAL); Jerneja Rebernak, Fleur Kaminska (UAL)

Executive Summary

Transforming Collections is led by University of the Arts London (UAL) Decolonising Arts Institute in collaboration with UAL Creative Computing Institute, and close partnership with Tate as an Independent Research Organisation. Our additional 15 project partners are: Arts Council Collection, Art Fund, Art UK, Birmingham Museums Trust, British Council Collection, Contemporary Art Society, Glasgow Museums, Government Art Collection, iniva (Institute of International Visual Art), JISC Archives Hub, Manchester Art Gallery, Middlesbrough Institute of Modern Art (MIMA), National Museums Liverpool, Wellcome Collection, and the Van Abbemuseum, Eindhoven.

Transforming Collections is underpinned by the belief that a national collection cannot be imagined without addressing structural inequalities, engaging debates around contested heritage, and revealing contentious histories embedded in objects. In 1999, the late sociologist and cultural theorist, Stuart Hall, posed the question 'Whose heritage?'. Hall called for the 'unsettling' and 'reimagining' of heritage and nation.[1] More than 20 years on, the need to critically question and transform notions of 'heritage' and 'nation' remain as urgent as ever.

The project combines critical art historical and museological research with participatory and interactive machine learning (ML) design. A series of artistic research residencies will engage with emerging findings and embed creative activations of the research within the project's public programme. The project seeks to surface suppressed histories, amplify marginalised voices, and re-evaluate artists and artworks long ignored or side-lined by dominant narratives and institutional practices. It aims to enable search across collections, to surface patterns of bias and relations of power, and interrogate policies and practices of classification, categorisation, description, and display. Critically and creatively connecting, interrupting, and disrupting collections, this interdisciplinary collaboration will open up new interpretative frames to imagine an evolving 'national collection', at once connected yet distributed. *Transforming Collections* aims to 'unsettle' and enrich existing knowledge with multiple and multivocal narratives, and 'potential histories' of art, nation and heritage.[2]

The project is designed around five parallel and interweaving work strands:

- Strand A: Surfacing Bias Across Collections (critical case studies)
- Strand B: Resurfacing Artists and Artworks Across Collections (critical case studies)
- Strand C: Participatory Design of ML (ideation and refinement workshops)
- Strand D: Interactive ML Technology Development (prototyping and testing)
- Strand E: Public Engagement Programme (in person and online events)

---

[1] Keynote speech by Stuart Hall given on November 1, 1999, at the National Conference, Whose Heritage? The Impact of Cultural Diversity on Britain's Living Heritage held in Manchester, England.
[2] Ariella Aïsha Azoulay, Potential History: Unlearning Imperialism (London; Brooklyn, NY: Verso, 2019).

Progress to Date

The first year of *Transforming Collections* focused on recruitment, building a diverse team of researchers and practitioners of colour (with over 50% representation to support research capability, capacity and early careers into the future); developing collaborative interdisciplinary working processes; developing data sharing guidance and ethical working principles; undertaking initial partner workshops; identifying early case studies; gathering diverse datasets from Tate and our various partners and organisations; and initial public programme development including a call for practice researchers in residence.

The initial 12 to 18 months of activity across Strands A and B of the project saw the extended and expanded audit of selected UK public art collections (building on audit data from the AHRC Black Artists and Modernism project, 2015-18), to produce an updated indicative snapshot of acquisitions over a specific period from 1900 to 2022. Strands A and B also saw the auditing of Tate Collection's subject index tagging since 2000 and the digitisation of iniva's uncatalogued archive of c.3,000 slides and c.180 artists' files. Between the UK collections audit, Tate Collection audit and iniva digitisation, and ongoing data gathering from partner collections, the project is bringing existing and new datasets into critical relation for the project's research purposes. The iniva dataset makes searchable hitherto hard-to-access documents relating to artists of colour[3]; and offers potential counterpoints to objects and data both present and absent in the project's UK collections partners. Digitisation within the project also represents a significant step in iniva's ongoing development of this vital resource within the Stuart Hall Library, located at UAL.

Over the same period, activity across Strands C and D of the project has focused on consultation, co-design and testing of early ML prototypes with 15 of our project partners through an iterative participatory and interactive process. Regular cross-strands meetings ensure that critical research and ML design processes are mutually informed and driven. Participatory design and feedback mechanisms are being established across the project partnerships to embed and sustain collaborative and ethical approaches. As the project unfolds, invited workshop participants have so far included diverse academic researchers, staff from partner institutions, as well as representatives from concurrent Towards a National Collection (TaNC) programme Discovery Projects. A cross-TaNC Projects workshop on data ethics is currently in development and will be delivered in autumn 2023.

The second year of the project saw the expansion of the team with six Postdoctoral Research Fellows joining between September 2022 and January 2023, and four artists taking up 15-month practice research residencies to engage with the project's emerging findings. At least 10 collections research case studies are underway, progressing in tandem with the interactive machine learning development. In April 2023, the project delivered an international conference to engage general and specialist audiences in the Netherlands, building on UAL's existing partnership with Van

---

[3] *Transforming Collections* refers to 'Black artists' for consistency with the term used in a collections audit initiated by the AHRC Black Artists & Modernism (BAM) project, significantly extended and expanded by the AHRC TaNC project. The project also refers to 'artists of colour' while recognising that artists may choose to identify by neither or different terms, as suggested by the #BAMEover statement for the UK.

Abbemuseum, Eindhoven and developing research networks, including with The Nieuwe Instituut in Rotterdam and the Research Center for Material Culture at the Tropenmuseum in Amsterdam. The conference took place in the context of the Rewinding Internationalism exhibition to foreground practice as research, where in-progress findings and prototypes were shared. Ongoing research case studies, interactive machine learning prototype testing and continued co-design with partner organisations will culminate in a third All Partners Workshop in November 2023. Planning for the public programme due to take place in collaboration with Tate Learning and Art UK in autumn 2024 continues.

Interim Recommendations

At the mid-point of the *Transforming Collections* project, we have several interim recommendations for consolidating the work of the TaNC programme and maximising the potential impacts of the five Discovery Projects. These recommendations also underline the impacts of recruitment delays over the first year of the project, the time required to establish effective interdisciplinary and collaborative working across the team, and the value of iterative looped processes:

- Further resource to deepen and extend critical engagement and evaluation activities beyond the end of the current funded three-year period, triangulating and underpinning heritage information as data and human/digital interactions with ethical practice, e.g. by examining approaches to language across all Discovery Projects.
- Further resource for an extended programme to roll-out durational testing of our adaptable, interactive ML tools designed to work with existing infrastructures, to enable sustained engagement and gather feedback over time from cross-partner, cross-project, cross-sector and cross-disciplinary and international stakeholders.
- Emphasis on the need for flexible interactive tools to enhance existing and future infrastructures with critical prompts and reflexive valves that encourage the rethinking of habitual formulations, hierarchies, values and vocabularies, and enable dynamic categorisations or tags to be created and refined by users, 'delinking and relinking' data to surface unexpected relations.
- Investment in long-term cataloguing work across project and programme partnerships informed by subject specialist researchers and our interactive ML tools, to enrich digital records by foregrounding transparency, context, and absences, not to situate the digital in opposition or progressive relation to the analogue; e.g. making explicit discrepancies with analogue records, or omissions across digital and analogue data, as part of the holistic record.

Inclusive collaborative practices must underpin and drive inclusive innovation in research. As such, we recommend that future funding criteria should require projects to:

1. Recruit a diverse, inclusive and representative research team with lived personal and professional experiences relevant to the wider questions, aims, objectives and intended audiences and beneficiaries of the project.

2. Engage diverse and directly implicated stakeholders as early as possible in bid development, project conceptualisation, and iterative design and delivery stages, including researchers, artists, estates and archival donors.
3. Value and embed practice-led or practice-based research and close interdisciplinary collaboration; to paraphrase Horst Hörtner: Science creates technologies; but it's the art that helps us to understand their impacts.
4. Value and embed active engagement with ethics and ethical implications of the research in principle and in practice.

Abstract

Transforming Collections aims to enable digital search across collections, to uncover patterns of bias in collections systems and narratives, to reveal hidden connections, and to open up new interpretative frames and 'potential histories' of art, nation and heritage[4]. Whose voices, bodies and experiences are centred and privileged in collections? This project is underpinned by the belief that a national collection cannot be imagined without addressing structural inequalities, contested heritage and contentious histories embedded in objects. In 1999, the late sociologist and cultural theorist Stuart Hall posed the question 'Whose heritage?'. Hall called for the 'unsettling' and 'reimagining' of heritage and nation. Nearly 25 years on, the need to critically question and transform notions of 'heritage' and 'nation' remain as urgent as ever.

Led by UAL in close partnership with Tate among our 16 partners across the UK, the project seeks to surface suppressed histories, amplify marginalised voices, and re-evaluate artists and artworks long ignored or side-lined by dominant narratives and institutional practices. We want to imagine an inclusive, evolving, (re)distributed 'national collection' that builds on and enriches existing knowledge with multiple and multivocal narratives, to critically connect and imaginatively disrupt collections, and transform them.

Our approach brings together academic and artistic research into collections and museum practices with participatory machine learning (ML) design, working with smaller, 'messy' datasets as well as larger collections' data. The interactive ML development will lead to lightweight adaptable tools that can both function as critical prompts and generate connective patterns: critical prompts to rethink habitual formulations, hierarchies and values expressed in collections' text-based digital records; and connective patterns through the bespoke creation of dynamic categorisations or tags refined by the user (that would not otherwise be made visible through standard search functions within collections databases), thereby surfacing unexpected relations between disparate objects or makers to shape new research. A series of artistic residencies will lead to new works that critically and creatively activate the emerging research and ML tools. A major public programme with Tate Learning will generate insights and understandings of the ways in which the *Transforming Collections* research can enable new stories to be told.

---

[4] Azoulay, Potential History.

Aims and Objectives

Transforming Collections aims to dissolve barriers between collections by addressing digital search and research capability as entwined priorities, with public engagement conceived as integral to the research development, design, and dissemination. Bringing together decolonial intersectional feminist praxes and creative machine learning, the project's approach is driven by the following questions:

1. Whose heritage? Whose voices, bodies and experiences are centred and privileged in collections? What are the problematic concepts used to label, classify, categorise and describe objects, and how do these perpetuate racist, xenophobic, misogynist, heteronormative, white supremacist and ableist views?

2. How can we surface and transform the architectures, algorithms and relations of oppression that structure and define collections' narratives? Can decolonial feminist approaches to machine learning help to counter or resist the replication of colonial and imperial modes of perception embedded in cataloguing and interpretation practices? How can we ethically engage and empower diverse stakeholders in transforming the politics and processes of representation?

3. What could an equitable, inclusive, distributed, connected, evolving 'national collection' look like? How might we look across collections to create multivocal narratives that enrich knowledge within collections? What inspiring, unexpected or uncomfortable stories could such a national collection tell?

The project builds on the insights and findings of previous research projects led by UAL, namely the AHRC Black Artists and Modernism project (BAM, 2015-18) and the UKRI MIMIC project (Musically Intelligent Machines Interacting Creatively, 2018-21), as well as the recent Tate-led TaNC Foundation Project, Provisional Semantics (2020-22).

The project's objectives are to develop, model and test critical, creative and ethically grounded and inclusive technological approaches that enable cross-search of collections and increase research capability, by surfacing patterns of bias, revealing hidden or unexpected connections, and opening up new interpretative frames and potential narratives of art, nation and heritage.

So far, *Transforming Collections* has:

1. Recruited diverse researchers and practitioners of colour, ensuring more than 50% representation in the project team, to support research capability, capacity and early careers into the future.

2. Completed an extended and expanded audit of selected UK public art collections, to produce an updated indicative snapshot of acquisitions over a specific period from 1900 to 2022.

3. Digitized iniva's uncatalogued archive of c.3,000 slides and c.180 artists' files, to establish an important new dataset for the project's research purposes. The dataset makes searchable hitherto hard-to-access documents relating to artists of colour; and offers potential critical counterpoints to objects and data both present and absent in the

project's UK collections partners. Digitisation also represents a significant step in iniva's development of this vital resource within the Stuart Hall Library, located at UAL.

4. Consulted, co-designed and tested early ML prototypes with 15 of our project partners through an iterative participatory and interactive process.
5. Delivered an international conference to engage general and specialist audiences in the Netherlands, building on an existing partnership with Van Abbemuseum and developing research networks, including with The Nieuwe Instituut in Rotterdam and the Research Center for Material Culture at the Tropenmuseum in Amsterdam.

Transforming Collections will:

6. Produce a series of critical case studies underpinned by decolonial and intersectional feminist approaches that inform and are informed by the interactive ML development. These will be disseminated through workshops/seminars, a conference, online/peer-reviewed publications, and a project website.
7. Develop adaptable and transferable supervised ML software through participatory design processes, with interactive ML methods. These lightweight tools will enable researchers, curators and wider audiences to cross-search collections and diverse datasets; uncover unexpected connections; surface hidden works and histories; and empower diverse users to control the process of determining significance in search terms.
8. Build, host and maintain the supervised ML software (algorithm and interface) at UAL for a minimum of 10 years on the CCI public GitHub, with resources to support dissemination to partners for beta testing and potential wider roll-out.
9. Develop and launch an accessible project website for general audiences.
10. Develop, present and prepare for manuscript submission a series of papers/essays towards an edited journal issue and a book publication.
11. Develop and deliver a major culminating public engagement programme as part of Tate Learning Projects and Programmes, including the online and/or physical display of artists' commissions, an international symposium, and an extensive programme of participatory workshops/seminars and talks.
12. Collaborate with Art UK and partners to showcase selected artists and artworks online.

Partnership Structure

Lead Institutions

University of the Arts London's (UAL) Decolonising Arts Institute (DeAI) is leading on the project's conceptual and strategic direction and delivery, including project management, planning, coordination, administration, and communications. DeAI is working in close collaboration with UAL's Creative Computing Institute (CCI), whose team is leading on the participatory co-design and technological development strands of the project. As UAL's main partner and Independent Research Organisation, Tate co-leads, co-develops and co-delivers the project, contributing expertise, research, insights, facilitating access to Tate Collection data, and co-curating the project's public engagement programme.

Project Partners and Collaborating Organisations

Initially launching with 15 partners including Tate, *Transforming Collections* now has a total of 16 partners including the Government Art Collection who joined the project in July 2022. The *Transforming Collections* partnership reflects the nationwide scope and ambition of the project to engage, benefit and positively impact collections across the UK, large and small, and builds on UAL's long history of collaboration with Tate, as well as previous projects with 12 of our partners. These include the major AHRC Black Artists & Modernism project (2015-18) led by UAL in partnership with Middlesex University, and several DeAI projects: the Decolonising British Art seminar series (2020-21, supported by British Art Network and Paul Mellon Centre); the Decolonising Collections research residencies (2021-22, supported by Art Fund); the Digital Artist Residency (2021-22, in collaboration with iniva); and the workshop/seminar series, Doing the Work: (2021-22, in collaboration with Contemporary Arts Society).

Partner organisations across museum, collection, arts, charity and archive sectors will contribute collections management and curatorial research insights, support access to collections databases, participate in ML co-design workshops, as well as engagement and dissemination activities. The partnership includes four Collaborating Organisations who are direct beneficiaries of the project funding. Selected collections will also host and support artists in digital residence starting from January 2023.

- The Arts Council Collection, British Council Collection, Birmingham Museums Trust, Glasgow Museums, Government Art Collection, Manchester Art Gallery, National Museums Liverpool, Middlesbrough Institute of Modern Art and the Wellcome Collection are facilitating access to their collections management data and records, engaging in curatorial support and participating in selected workshops/seminars. JISC Archives Hub will enable access to its numerous art collections among its 350 repositories.
- The major arts charities, Art Fund and Contemporary Art Society, are providing data on their UK-wide collection acquisitions and donations since the 1900s, including grants

awarded, gifts and bequests, artists and exhibitions profiled through the magazine *Art Quarterly*, and organizational histories.

- Art UK is providing analytics on its aggregated collections data on c.48,000 artists, c.250,000 artworks, user-generated tags, descriptions, and artworks online.
- iniva is digitising its unique artists' archive, which is physically housed within iniva's Stuart Hall Library and located at UAL's Chelsea College of Art.
- Van Abbemuseum in Eindhoven co-hosted an international conference in April 2023 to situate the project's key questions, emerging research and ML development, in relation to artistic and curatorial interventions into 'national' and 'international' collection praxes within the context of the Netherlands and discourses around Dutch colonialism and its legacies.

Staffing Structure

Project Management and Administration

The Principal Investigator (PI) **Prof susan pui san lok** and the project team of 18 Co-Investigators and researchers are supported by Project and Partnerships Manager **Jerneja Rebernak** and Project Administrator **Fleur Kaminska**. The PI is further supported by a **Management Board** which meets every six months to track the project's progress against identified objectives and milestones. The Management Board members are **Benjamin Stopher** (Dean of UAL CCI, on parental leave since January 2023), **Prof Malcolm Quinn** and **Prof Pratap Rughani** (UAL Associate Deans of Research), and **Emily Pringle** (Tate Head of Research until December 2022, replacement tbc). Standing in for Benjamin Stopher on a temporary basis are Dr Vali Lalioti (CCI Director of Programmes) and **Prof David Mba**, Deputy Vice-Chancellor for Research, Knowledge Exchange and Enterprise. An external **Advisory Board** meets formally once a year and informally as needed, to contribute cross-sector expertise, and ensure the relevance and impact of the project. The Advisory Board members are: **Melanie Keen** (Director, Wellcome Collection), **Prof Dorothy Price** (Professor of Modern and Contemporary Art and Critical Race Art History, Courtauld Institute), **Sara Wajid** (co-CEO, Birmingham Museums Trust), **Rachael Minott** (Joint Head of Diversity Equity and Inclusion, National Archives and Joint Head of Participation, Birmingham Museums Trust), **Prof Stephanie Dinkins** (Transmedia Artist and Kusama Endowed Chair of Art, Stony Brook University), **Dr Alexandre White** (Assistant Prof of Sociology and History of Medicine, John Hopkins University).

Principal Investigator (PI)

[Professor susan pui san lok](#) (Professor of Contemporary Art and Director of the UAL Decolonising Arts Institute) is leading the *Transforming Collections* project, directing all research strands, co-leading the public engagement programme, with overall responsibility for the delivery of the project.

Co-Investigators (Co-Is)

[Dr Anjalie Dalal-Clayton](#) (Research Fellow, UAL Decolonising Arts Institute) and Christopher Griffin (Senior Curator: *Transforming Collections*, Tate) are co-leading on Strands A&B of the project, focused on surfacing bias across collections and resurfacing artists and artworks. Dalal-Clayton has undertaken an extended and expanded audit of works by black artists in UK public collections acquired since 1900. Griffin is developing Tate-focused historiographic case studies.

Hilary Knight (Digital Director, Tate; Co-I until June 2022) and **Hannah Barton** (Senior Project Manager, Tate Digital; Co-I from July 2022 until April 2023) led on Tate Collections' Subject Index tagging research, uncovering problematic language in explicit and euphemistic forms, exploring the ways in which this has been addressed over time, and implications for the project research and diverse users' experience (Strands A&B). **Liam Darbon** (Director of Digital and Innovation, Tate) will be taking on this Co-I role from May 2023.

[Professor Sonia Boyce](#) (Chair in Black Art and Design, UAL; Co-I from July 2022) is contributing to the development of case studies focused on iniva's archive (Strands A&B).

[Professor Mick Grierson](#) (Professor and Research Leader, UAL Creative Computing Institute) and [Professor Rebecca Fiebrink](#) (Professor, UAL Creative Computing Institute) are co-leading on Strands C&D of the project, focused on participatory co-design workshops and interactive ML technology development.

[Dr Athanasios Velios](#) (Reader in Digital Documentation at UAL Camberwell, Chelsea and Wimbledon colleges) is contributing to the project technological development with a focus on possibilities, problematics and limitations of data integration, ontologies and Linked Data, including the potential to record collections' uses of the ML tool (Strands C&D).

[Dr Peaks Krafft](#) (Senior Lecturer and MA Internet Equalities Course Leader, UAL Creative Computing Institute) is contributing to ethical considerations of artificial intelligence in terms of conceptualisations of data, and the critical impacts and implications for the technological design, engagement and outputs of the project (Strands C&D).

Mark Miller (Director of Tate Learning) is co-leading Strand E with the PI, focused on public engagement through the project's culminating public programme.

Research Fellows

Ananda Rutherford is developing collaborative case studies on collections information systems and practices. She supported Dalal-Clayton on the collections audit (Strands A&B), and is working with [Dr Charlotte Webb](#) (Senior Lecturer Online courses, UAL Creative Computing Institute) on developing and applying data ethics principles in practice across project strands.

Dr Jon Gillick (from July 2022) and **Dr Ireti Olowe** (from January 2023) are focused on the ML development and user experience, from programming and visualisation perspectives respectively (Strands C&D).

Five further **Post-Doctoral Research Fellows** (from Sept 2022) are developing critical case studies with selected partner collections and organisations, engaging with the emerging data and with the evolving ML tool. They are: **Dr Tiffany Boyle, Dr Alice Correia, Dr Andrew Cummings, Dr Tehmina Goskar** and **Dr Ian Sergeant**.

Research Assistants

Kit Bower-Morris is contributing to the early ML prototype development by coordinating and delivering participatory co-design workshops with the team and partners (Strand C).

Veera Jussila was seconded to UAL CCI (from Sept 2022 until April 2023) and contributed to the early ML prototype development with a focus on image search.

iniva Archivist

Kaitlene Koranteng (Archivist and Engagement Producer, iniva) is leading on the digitisation of iniva's artists' archive.

Overall Programme

	Year 1 2021/22				Year 2 2022/23				Year 3 2023/24			
	Quarters				Quarters				Quarters			
Extension beyond original timeline ▨	1	2	3	4	1	2	3	4	1	2	3	4
Strand A: Surfacing Bias Across Collections												
Collections audit extended and expanded (1900-2022)	■	■	▨	▨	▨							
iniva artists' archive digitised		■	■	▨	▨	▨						
Develop case studies on structural and systemic biases		■	■	■	■	■	■					
Case studies presented at conference and/or workshops				■	▨	▨						
Develop papers into articles/essays for edited journal/book									■	■	■	■
Strand B: Resurfacing Artists and Artworks												
Close readings of selected artworks		■	■	■	■	■						
Develop case studies exploring new connections & relations between works across collections		■	■	■	■	■	■	■				
Case studies presented at conference and/or workshops				■	■	■						
Develop papers into articles/essays for edited journal/book									■	■	■	■
Strand C: Participatory Design Workshops												
Ideation & refinement workshops with stakeholders	■	■	■	■								
ML v1 UI prototyping, user studies, testing, evaluation				■	■	■	■					
ML v2 UI prototyping, user studies, testing, evaluation							■	■	■			
Develop papers into articles/essays for edited journal/book									■	■	■	■
Strand D: ML Technology Development												
Develop ML adapter layer to interact with collections data			■	■	■	■	■	■	■			
Provide ontological modelling to partners & collaborators					■	■	■	■	■			
Ongoing refinement, collaborative testing & feedback			■	■	■	■	■	■	■	■	■	
Develop papers into articles/essays for edited journal/book									■	■	■	■
Strand E: Public Engagement and Impact												
Establish & maintain project resources on GitHub	■	■	■	■	■	■	■	■	■	■	■	■
Establish & maintain accessible public project website					■	■	■	■	■	■	■	■
Conference & workshops with partners & stakeholders						■			■			
Publish workshop, conference & seminar recordings							■	■		■		
Prepare edited book contents for publication									■	■	■	■

Task												
Tate Learning programme R&D, production, installation		■	■	■	■	■	■	■	■	■	■	■
Artists practice research residencies in progress					■	■	■	■	■			
Artists ML commissions digital build / physical install / live										■	■	■
Tate installation and public programme live & documented												■
Art UK online showcase development, design, production							■	■	■	■		
Art UK online showcase live											■	■

Events and Consultations

Event / Consultation	Date	Location(s)	Attendees
Audit Consultation	04.02.2022	Online	Tate Information Management team
Audit Consultation	23.02.2022	Online	Cartwright Hall, Bradford
All Day Workshop: Strands C&D	29.03.2022	UAL Decolonising Arts Institute (DeAI)	13 (project team + 1 additional Tate staff)
Tate Collection Research Workshop	30.03.2022	Tate Britain	14 (12 Tate staff including 2 Tate Co-Is + 2 TaNC Project team members)
Audit Consultation	22.04.2022	Online	Wellcome Collection
All Day Workshop: Strands A&B	27.04.2022	Tate Britain, iniva, UAL DeAI	16 (project team + 3 additional Tate staff + 2 additional iniva staff)
ML Co-design Conversation	03.05.2022	Online	Wellcome Collection
Audit Consultation	05.05.2022	Online	Sheffield Museums
ML Co-design Conversation	10.05.2022	Online	Arts Council Collection
ML Co-design Conversation	10.05.2022	Online	JISC Archives Hub
ML Co-design Conversation	11.05.2022	Online	British Council Collection
ML Co-design Conversation	11.05.2022	Online	Art UK
Audit Consultation	16.05.2022	Online	Leicester
All Day Workshop: All Partners	18.05.2022	UAL Creative Computing Institute (CCI)	34 (project team, UAL stakeholders, 18 partner organisation representatives)
Audit Consultation	23.05.2022	Online	British Museum
Audit Consultation	23.05.2022	Online	Herbert Art Gallery, Coventry
Audit Consultation	24.05.2022	Online	National Portrait Gallery
Audit Consultation	24.05.2022	Online	Bristol
ML Co-design Conversation	11.06.2022	Online	Contemporary Art Society
Audit Consultation	20.06.2022	Online	Wolverhampton
ML Co-design Conversation	27.06.2022	Online	Manchester Art Gallery
ML Co-design Conversation	01.08.2022	Online	Birmingham Museums Trust

International Partner Visit	15.09.2022 16.09.2022	VAM	Van Abbemuseum (VAM) collections and curatorial staff
ML Co-design Conversation	21.09.2022	Online	Government Art Collection
ML Co-design Workshop	25.10.2022	Online	8 (representatives from 4 project partner organisations: BCC, ACC, iniva, GAC)
All Day Workshop: All Partners	23.11.2022	UAL CCI	46 (project team, UAL stakeholders, 25 representatives from partner organisations)
Artists Welcome Workshop	17.01.2023	Tate Britain	23 (4 artists and 19 project team members
Prototypes Conversation	01.02.2023	Online	Tate
Data Conversation	02.03.2023	Online	Birmingham Museums Trust
Two Day International Conference & Workshop	20.04.2022 21.04.2023	Van Abbemuseum	c.65 delegates, public-facing event
Planned			
Artists Mid-Residency Seminar	04.07.2023	Tate Modern	TBC
Data Ethics Workshops	09.2023	In-person / online	c.25 (all TaNC Projects)
All Day Workshop: All Partners	08.11.2023	In-person	TBC
Artists End of Residency Seminar	27.03.2024	TBC	TBC
Public Programme: Displays, Workshops & Talks	10-11.2024	TBC	TBC

In addition to the above events and consultations, one-to-one co-design consultations and conversations will continue to the end of the project. Selected collections site visits and interviews are taking place in Year 2 to support the development of case studies.

Research Approach

Context

Transforming Collections seeks to complicate the 'digital cultural record' by embedding decolonial voices and decolonizing impulses to interrogate the structures and narratives of art, nation and heritage. Building on the findings of the AHRC Black Artists and Modernism project (2015-18, led by Prof Boyce with Prof Iok as Co-I and Dr Dalal-Clayton as Research Fellow) and the UKRI MIMIC project (Musically Intelligent Machines Interacting Creatively, 2018-21, led by Prof Grierson), the project also reflects on the recent TaNC Foundation project, *Provisional Semantics* (led by Tate with Dr Dalal-Clayton as Co-I and Rutherford as Research Fellow). The project also builds on UAL and Tate's past and current work on the challenges facing collections. These include 'Tate Encounters' (Dewdney, Dibosa, Walsh, 2007-12) and the HLF-funded 'Archives and Access' project involving digital access, participation and learning with archives; revisions to the 'Art and Artists' database; and an ongoing audit of legacies of slavery in the historic collection (2012-17). Tate's 'Accounts and Accountability' project (led by Griffin) aims to address the history of collection interpretation and the ways in which it has perpetuated racism, imperialism and white supremacy, to ensure that all new written accounts are representative of Tate's commitment to anti-racism, institutional accountability and inclusive language.

Transforming Collections takes up the problem of persistent colonial perspectives and narratives, whereby public institutions (including collections) contribute to and perpetuate an oppressive shared cultural archive[5]. Structural and systemic bias can be discerned across the cultural and technological spaces of the project. Researchers at the intersection of social justice and artificial intelligence (AI) have shown that AI systems can amplify racism, sexism, ableism, and other forms of discrimination; ML can be used to detect and balance bias in text collections[6]; and data augmentation can help to prevent discrimination and disambiguate data structures[7]; while data feminism presumes that oppressive systems of power harm all and hinder the possibility of creating of lasting social impact with data science[8]. Artists, activists and academics (such as Stephanie Dinkins' transmediale practice) continue to build the movement towards equitable and accountable AI (such as the Algorithmic Justice League, founded by Boulamwini, 2016), and ethical treatment of data to combat racial injustice (such as Black Beyond Data, led by Jessica Marie Johnson and Yomaira Figueroa Vasquez).

---

[5] Gloria Wekker, White Innocence: Paradoxes of Colonialism and Race, 2016.

[6] Lucas Dixon et al., 'Measuring and Mitigating Unintended Bias in Text Classification', in Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society (AIES '18: AAAI/ACM Conference on AI, Ethics, and Society, New Orleans LA USA: ACM, 2018), 67–73, https://doi.org/10.1145/3278721.3278729.

[7] Shubham Sharma et al., 'Data Augmentation for Discrimination Prevention and Bias Disambiguation', in Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society (AIES '20: AAAI/ACM Conference on AI, Ethics, and Society, New York NY USA: ACM, 2020), 358–64, https://doi.org/10.1145/3375627.3375865.

[8] Catherine D'Ignazio and Lauren F. Klein, Data Feminism, Strong Ideas Series (Cambridge, Massachusetts: The MIT Press, 2020), https://data-feminism.mitpress.mit.edu/.

Approach

Our approach is to evolve an ethical and interdisciplinary braided approach underpinned by decolonial feminist praxes by 1) folding critical art historical and museological research with 2) creative ML development and participatory design and 3) embedding creative practice-based research in collections through artists' activations of interactive ML. We aim to address barriers between collections, audiences and artworks by:

- Enabling cross-search of collections with adaptable ML to enhance the connectivity, discoverability and accessibility of collections for specialist and general audiences alike;
- Enabling researchers and the general public to surface bias and (re)discover those artists and artworks most marginalised within collections, while also ethically engaging and empowering those most marginalised by the institutions that seek to represent them;
- Building on existing systems to enable fluid connections between and across collections;
- Ethically engaging and empowering diverse stakeholders in designing and testing the ML model;
- Generating wide-ranging user and research case studies, and a free, open lightweight ML tool, framework and resources; to be made available to all TaNC projects and related institutions, with potential to roll out beyond the TaNC programme timeframe;
- Establishing an accessible project website to share discoveries with wider audiences;
- Engaging wide-ranging publics through artists' digital commissions and online showcases;
- Opening up new ways of discovering the sometimes uncomfortable stories collections can tell.

Whether a future 'national collection' is realised as a single unified entity, or as a decentralised, distributed and dynamic 'whole', *Transforming Collections* aims to deepen and complexify connections and engender new relationships between objects, collections and the publics they seek to engage and represent. We propose an ethical and inclusive model of cross-collection search and engagement that begins to democratise and decolonise relations of power – not only in terms of how collections' data is defined, held and accessed, but also in relation to the fluid and collaborative production, expansion and contestation of knowledge. The project has the potential to significantly transform the galleries, libraries, archives and museums (GLAM) sector, and thus impact on related disciplines and practices, such as art history and museology. Mobilising ML to surface at scale and at pace the artworks long 'hidden in plain sight', we aim to demonstrate the radical possibilities for transforming collections and reimagining art, heritage and nation.

Impact on Digital Search

Transforming Collections is not attempting to solve the decades-long challenge of developing universal or unifying database and search standards. Rather, we start from the assumption that information systems are never neutral or objective: that there will always be problems of bias, 'misclassification' and inconsistencies in data; and that the so-called 'truth' of labels or categories can always be contested, whether they appear to be factual or speculative, conventional or controversial.

We approach linguistic data as both bridge and barrier to visibility, searchability and value. We seek to dissolve barriers by testing ML's capacity to uncover patterns in data, combined with data science

techniques, to reveal and examine latent biases: e.g. the persistence of problematic and offensive language and terminologies that may reflect inherited colonial views and imperialist values, thus perpetuating systemic discrimination and exclusion; or patterns in purchase data that expose networks of power and influence. We aim to build bridges by training ML to search different objects, formats (e.g. texts, images, sounds) and data management systems, thereby enabling unexpected connections between collections, and new interpretative frames, narratives and histories to emerge.

We are working with partners and collaborating organisations who hold or support distinct UK public art collections of varying scales, from large collections deemed nationally significant to smaller yet arguably equally significant and influential municipal or civic collections. Regardless of profile, all hold a strikingly low percentage of works by artists of African and Asian heritage (less than 5%), compared to the representation of the UK population from 'minority ethnic' or 'non-white' backgrounds (c.13%, 2021 Census). Our partners share a commitment to understanding how artists and works have come to be under- or misrepresented, classified, categorized, 'highlighted' or hidden, described, narrativised, and de/valorised within collections and the wider culture. We will:

- Develop adaptable supervised ML software that can connect and augment different datasets, regardless of formats and locations, and thus enable digital cross-search across collections;
- Use selected IML (Interactive Machine Learning) methods and participatory design approaches to work with diverse academic and non-academic stakeholders including TaNC portfolio projects and general audiences, to control the process of determining significance according to individual experience and expertise, e.g. defining labels, designating values and generating bespoke metadata;
- Train interactive ML models to reveal patterns in data within and across disparate datasets; to see what is 'highlighted' or hidden, 'similar' or 'different', under-explored or absent within collections; to surface old/new relationships and create interconnections;
- Discover new and unforeseen links between disparate objects, stories, histories and locations.

Impact on Research Capability

We are bringing decolonial, critical race theory and intersectional feminist approaches[9] to both the ML development and design, and to the analyses of connections surfaced using the evolving ML framework, as braided work strands. A series of case studies focused on works by 'artists of colour' will demonstrate the significantly enhanced and expanded research capability for multi- and cross-disciplinary academics (e.g. engaged in art historical, museological, artistic and curatorial research); as well as the potential benefit to museum and heritage studies, information and archival sciences more widely. The patterns and connections surfaced through ML will enable researchers to identify, contextualise and interrogate the latent colonialism, racism and imperialism of collection terms,

---

[9] Margaret L. Andersen and Patricia Hill Collins, 'Why Race, Class, and Gender Matter', in Inequality in the 21st Century (Routledge, 2018); Christina Dunbar-Hester, Hacking Diversity: The Politics of Inclusion in Open Technology Cultures, Princeton Studies in Culture and Technology (Princeton, New Jersey: Princeton University Press, 2020).

concepts and information hierarchies; situate language usage and understandings through time and across locations; and both generate and accelerate new interpretations and radical re-framings of works and collections. In addition to research case studies, a series of digital residencies undertaken by artists of colour will result in commissions that activate the critical and creative potential of ML in transforming how we access, navigate and understand collections. As such, the project seeks to demonstrate the potential to breach physical boundaries and broach new understandings that may fundamentally transform disciplinary methods and understandings of the objects that collections hold.

Crucially, the project has undertaken the strategic digitisation of iniva's uncatalogued archive of just under 3,000 slides and 180 artists' files. This activity has established a rich dataset for the research purposes of the project. Together with the data on artworks and artists uncovered by the extended and expanded audit of UK public art collections (originally undertaken as part of the AHRC BAM project), iniva's digitised archive offers a valuable dataset for critical comparison with partners' collections data, which together inform a range of ongoing case studies, as well as the ongoing interactive ML development. The digitisation process has surfaced questions and ambiguities around archival donation policies and processes, and the need to address artists rights within a changing archive. This work forms part of iniva's ongoing strategy and lies beyond the scope of the *Transforming Collections* project. Nevertheless, the digitisation of iniva materials within the context of the project represents a vital step towards iniva's creation of a unique research resource, which will have a significant impact on research capability in this area when it becomes publicly accessible. The value of iniva's digital archive for the project and beyond lies in iniva's historic championing of 'artists from around the world whose work and ideas… provide new perspectives for Britain's [still] predominantly western-centric view of the visual arts', as well as iniva's continuing work 'predominantly with British-born and British-based visual artists of African and Asian descent'.

Impact on Public Engagement

Transforming Collections will embed public engagement at key points in its design and delivery. We are building an open-source GitHub to host resources as they develop (e.g. videos, presentations, case studies, walk-throughs, and tutorials). This is currently project-facing only, and a public-facing GitHub will be launched in the third and final year of the project. In Year 3, we will also launch an accessible public-facing website to engage general audiences in the project's aims and activities, and the contested and 'potential histories' of collections beginning to emerge through our research case studies and artists' practice research residencies.

At the mid-point of the project in Year 2, we held an international conference to share work-in-progress and gain feedback from specialist audiences already engaged and committed to interrogating museums and collections from decolonial perspectives.

Transforming Collections will culminate in a major public programme in collaboration with Tate Learning, which will take place towards the end of Year 3. The programme aims to engage wide-ranging academic and institutional stakeholders, as well as specific communities invested or implicated in particular collections. Featuring a large-scale symposium, workshops/seminars, and a series of talks and events, we are also exploring public installations, interventions, and participatory events in the galleries and online. The public programme will be a significant and celebratory

extended moment for sharing and reflecting on the project's innovations and discoveries. Featuring four artists' commissions as physical and/or virtual interventions on site and online, engagement activities will explore critical, experimental, creative ways in which ML can be ethically engaged and activated to transform collections. We will also be working with Art UK to showcase selected artists and artworks online, and mobilising their Art Detective initiative to invite the wider UK public to contribute to existing knowledge around works in collections. We aim to demonstrate what an emerging, distributed yet connected and evolving digital national collection can look like, and in so doing, to empower as many people as possible in the collective reimagining of art, heritage and nation.

Project Design and Methodologies

Transforming Collections is designed around five parallel and interweaving work strands.

> Strand A: Surfacing Bias Across Collections – Critical Case Studies
> Strand B: Resurfacing Artists and Artworks Across Collections – Critical Case Studies
> Strand C: Participatory Design (PD) of ML – Ideation and Refinement Workshops
> Strand D: Interactive ML Technology Development – Prototyping and Testing
> Strand E: *Transforming Collections* – Public Engagement and Impact

Collections research case studies (Strands A and B) and ML technology (Strand D) are developing in tandem, ensuring that critical research and design processes are mutually informed and driven. Embedded participatory design and feedback mechanisms (Strand C) have been established to sustain a collaborative ethical and interdisciplinary approach. Project strands will draw on the collections audit data, and the new digitised data from iniva's archive, along with diverse datasets provided by Tate and our 15 UK partners and organisations. Strands will address the outlined project questions as follows:

In Years 1 and 2, work strands are converging through a series of up to 20 workshops/seminars/consultations interweaving critical research case studies and ML design, following a participatory action methodology to avoid extractive relationships.[10] Project partners, collaborating organisations and local stakeholders are being invited to participate in selected workshops, along with representatives from all the TaNC Discovery Projects; members of the public will be invited to participate further on in Years 2 and 3.

Year 2 has seen four artists of colour joining the project to undertake digital practice research residences, engaging with selected partner collections including Tate. After an initial period of research, artists are developing proposals for critical and creative interventions in response to the connections surfacing through the research case studies, and the potentialities and risks of ML in the context of collections.

In April 2023, the mid-point of the project was marked by an international conference co-hosted with Van Abbemuseum (VAM), coinciding with their major exhibition, 'Rewinding Internationalism' (November 2022 to May 2023). The conference provided an international context for reflecting on

---

[10] Sasha Costanza-Chock, Design Justice: Community-Led Practices to Build the Worlds We Need, Information Policy (Cambridge, Massachusetts: The MIT Press, 2020).

the project's questions, methodologies and progress, situating the notion of a 'national collection' and questions of access, inclusion and decolonisation in relation to other local-international initiatives and practices (e.g. Collecting Otherwise at the Nieuwe Instituut). Importantly, the context foregrounded the potential of exhibition-making as a critical site of creative research, featuring new commissions and works from numerous European museum collections across ten galleries, and a keynote from the transdisciplinary American artist and academic, Professor Stephanie Dinkins.

In spring 2024 (early Year 3) in addition to an open-source project-facing GitHub hosting resources (e.g. datasets, code, workshop materials, tutorials, walk-throughs), we aim to have established an accessible public-facing project website, featuring video, audio, visual and text documentation. Drafted research case studies will shape an edited journal proposal and an edited book proposal, as well as alternative forms of publication.

Towards the end of Year 3, *Transforming Collections* will culminate in high-profile public engagement activity including a major public programme co-curated with Tate Learning, and an online showcase with Art UK. An edited book illuminating artists' digital residencies and artworks surfaced through the project will be prepared for online and print publication, to extend the project's impact beyond the funded period.

Research Results

Strands A&B: Surfacing Bias Across Collections / Resurfacing Artists and Artworks

Strands A and B have focused on four areas of activity: a collections audit (see Annex); a Tate case study; selected digitisation of iniva's artists' archive; and a range of emerging and developing case studies shaped and informed by this evidential knowledge base.

The purpose of the collections audit was to create a dataset of currently held information on works by Black artists in a UK-wide selection of 31 public collections.[11] The first audit stage updated the lists of modern and contemporary works (initially identified as part of the AHRC Black Artists & Modernism project) by extending the audit period from 1900 to 2016, to 1900 to 2022. Data collecting and processing was completed for the 31 collections in November 2022: 3,767 artworks within the scope of the audit, by 578 individual artists and artist groups were identified. The second stage, which expanded on the basic inventory and acquisition data to gather related information, such as biographical data, cataloguing records, descriptive and interpretive texts, was completed in January 2023. The resulting dataset evidences which artists are currently represented (or not) within a sample of UK collections, and the extent and distribution of their representation over place and time. It further indicates the volume, variety and variability of available (and absent) information, and related interpretation material currently held.

A range of ethical considerations emerged while conducting the audit: these relate to the racialised classifications employed to identify artists and their potential to perpetuate divisions; how this audit might be mobilised and instrumentalised outside of the project; and how this in turn might be detrimental to the artists identified through the audit. The project stresses the limitations of the audit methodology and discourage its wholesale application by other organisations. As with the earlier Black Artists & Modernism audit, the data should be regarded only as broadly indicative of acquisition practices in a specific selection of collections within the specific period under review. As such, the data should not be interpreted as a comprehensive or definitive assessment of collection holdings, but rather used as a 'jumping off point' for further research. Project case studies currently in process that take the data as points of departure are beginning to reveal:

- Patterns of acquisition that have shaped public art collections and structural biases in documentation, visibility and funding; display, and interpretation practices over time and place and in relation to particular artists
- Differences in collecting practices within civic and national collections, or between Scotland, Wales, Northern Ireland, and England more broadly, and engagement with the diversity, equality, and inclusion agenda

---

[11] For continuity with the AHRC BAM project, the term 'Black artist' is defined as an artist with African, Caribbean, Asian or MENA-region heritage.

- What, where, when and how problematic concepts have been used to label, classify, categorise, describe, and valorise (or devalorise) objects in collections; and how such concepts can perpetuate racist, xenophobic, misogynist, heteronormative, white supremacist and ableist views.

The Tate case study addresses the project's research questions in several ways, engaging Tate Research, Interpretation and Digital teams. Firstly, by surveying and consolidating all existing collection information that is known to perpetuate bias, including critical commentaries from staff and audiences (e.g. in the form of offensive or euphemistic language, the absence or erasure of certain histories, and the under-representation or misrepresentation of identities). Secondly, by analysing editorial approaches and interventions within collection texts (e.g. catalogue entries, gallery captions) to address instances of discrimination. Thirdly, by scrutinizing the infrastructures that determine how Tate produces, stores and presents information and makes discoverable the 150,000 collection items on http://www.tate.org.uk/, namely through a process called Subject Index tagging.

An internal review of the ways in which descriptive terms are applied to collection items during the acquisition and cataloguing process, and subsequently data processed and linked, has revealed variances in approaches to Subject Index tagging since its introduction over 20 years ago. Examples of incorrect, problematic and inconsistent tagging also represent risks of reproducing biases and perpetuating harmful experiences for Tate's audiences. By mapping the numerous departments, disciplines, systems and platforms implicated in knowledge capture and production, information management, and the online presentation of Tate's collection, the review makes evident the contingency of cultural knowledge production. The ongoing research will address questions of authorship, ownership and responsibility over the Subject Index tagging process; the extent and approach to any remediation work on tags and texts, examples of which demonstrate the ethical considerations to be embedded in the project's ML development. The Tate case study is key to *Transforming Collections* in addressing how system architectures come to define information practices and collection narratives. As home to 'the national collection of British art from 1500 to the present day', Tate represents a microcosm of the wider GLAM information landscape and makes a significant contribution to the wider cultural digital record.

In addition to a series of micro-historiographic case studies undertaken by the project researchers, Tate will commission subject-specialist researchers to examine the interpretation of particular artworks in the collection with respect to the historical and cultural specificity of each work. These short historiographic accounts of each work will also be published on the individual artwork pages of Tate's website, offering reflections on the way in which the artwork has been understood over time. Similarities and differences in interpretation across different kinds of artworks will be analysed to reveal patterns of bias. These analyses will also inform the project's interactive machine learning development.

Digitization of iniva's archive of approximately 200 'artists from around the world' was completed in spring 2023. Founded in 1994 in response to the lack of 'Black and Asian' artists within the mainstream visual arts sector, the digitisation of the archive is a strategic step towards enriching iniva's offer as a unique research resource. Pragmatically addressing the question of how we might surface suppressed histories, amplify marginalised voices, and re-evaluate artists and artworks ignored or sidelined by dominant narratives, the digitisation process has established a dataset that

enables comparative critical analyses. Research case studies are beginning to explore the relative presence, absence, representation and interpretation of particular artists and works within iniva, and across audited collections, as well as the multiplication, duplication or disruption of connections between. The cataloguing of the digitised archive is an ongoing activity beyond the scope of the *Transforming Collections* project, which iniva is continuing in order to make the digital archive accessible as a public resource.

A range of critical case studies are currently in development, variously informed by oral histories, including:

Dr Hannah Barton on the conditions, challenges, inconsistencies and contingencies of subject indexing as a critical digital data process of identifying and describing items held in Tate's collection begun in 2000.

Prof Sonia Boyce on the questions raised by digitising iniva's artist's archive around its constitution, care and relation to other archives and collections.

Dr Tiffany Boyle on the disbandment of the Scottish Arts Council Collection, and the acquisition of works by the artist Maud Sulter in museum collections across Scotland.

Dr Alice Correia on the acquisition, documentation, and contexts of display of paintings by artist Gurminder Sikand, whose work is held in six project partner and other collections.

Dr Andrew Cummings on the acquisition, description, and display of Hamad Butt's works Transmission (1990) and Familiars (1992) by Tate between 2014 and 2020.

Dr Anjalie Dalal-Clayton on peak acquisition periods revealed by the collections audit; the acquisition and interpretation of work by the artist Tam Joseph across several collections; and the controversy and redressing of the Rex Whistler mural at Tate Britain through a new commission from the artist Keith Piper.

Dr Tehmina Goskar on 'patterns of patronage', looking at what collections data can tell us about the nature of bias in acquisitions, and what policies and decisions of art funders can tell us about the shaping of public art collections.

Christopher Griffin on discrepancies between the comparative breadth of acquisitions and depth of knowledge around particular works, revealing that while Tate has broadened its representation of artists in the collection over the last 20 years, cataloguing has not kept pace to ensure that knowledge is also deepened.

Prof susan pui san lok on the acquisition and interpretation of works by Vong Phaophanit (and Clare Oboussier) at Tate, and by Lesley Sanderson across four municipal collections; and on questions raised around the constitution, care and relation between archives and collections, through practice research into the iniva archive and Van Abbemuseum's Gate Foundation archive.

Ananda Rutherford on the process and ethics of the collections audit with Dr Anjalie Dalal-Clayton; embedding data ethics within the methodology of the project, with Dr Charlotte Webb, and problematising catalogue texts, editing and machine learning with Dr Jon Gillick.

Dr Ian Sergeant on the acquisitions and narratives around the work of the late artist Donald Rodney since 1989, held across seven collections.

Dr Charlotte Webb on the interdisciplinary formulation, reflection, testing and embedding of data ethics within the methodology of the project, towards a data ethos.

Strands C&D: Participatory Design and Interactive ML Technology Development

The technological approach within Strands C and D of the project has focused on human-centred participatory co-design as method.[12] Individual partner consultations and collective participatory workshops have been held to share early research activities, explore design ideas, user needs and experiences, understand challenges and 'pain points' for partner organisations relating to the goals of the project, and provide hands-on demos of our first lo-fi ML prototypes.

Informed by Strands A and B, the aim is to not only identify patterns in problematic language, but also to recognise the influence and inscription of colonial contexts and ideologies through the way that certain information is included, withheld, or prioritised. In Year 2, the focus has expanded to explore visual modes of search and multimodal analyses across text and image. With initial ML prototypes under ongoing development, the next stage of the project will focus on building more flexible and robust prototypes, testing with larger and diverse datasets, creating accessible explanations and feedback mechanisms, exploring the possibilities and limits of using Linked Data to record the use of the ML tools, and co-facilitating working sessions that explore data ethics in practice.

The technological development and design of the project has been guided by the need for transparency, context and attention to absences as fundamental principles. Transparency in terms of methodologies – ensuring clarity on how data will be shared, as well as when, where, and how specific ML techniques are deployed, explaining processes and enabling users to explore alternatives. Transparency also refers to the research aim of surfacing the histories and historiographies of objects, which may demonstrate systemic oppression or marginalisation, and how certain voices, bodies and experiences have remained unheard, unseen or erased. Identifying a lack of transparency and accountability (e.g. through de-historicised, 'neutral', anonymous, authorial voices) could also demonstrate how racist, xenophobic, misogynist, heteronormative white supremacists and ablest views can be inadvertently perpetuated. Comparative analyses of the narrative structure of texts, may reveal shifting hierarchies of value bestowed on different artists and works within collections, as can the mapping of absences – that is, the missing, unrecorded or hidden information, or material that remains undigitised. Context refers, then, to the need to make explicit the often-Western Eurocentric socio-political, historical and theoretical views reproduced in the interpretation and analysis of data and information framed as knowledge. Context also points the implicit question of the positionality and partiality of project researchers and participants.

Early discussions within the project foregrounded the need for critical, ethical approaches to data. These have led to the creation of two documents, both of which have been shared with our partners: a Data Sharing Working Guide explaining how we intend to use data within the project; and a set of Data Ethics Working Commitments. The data ethics commitments comprise a set of principles and prompts to guide the ongoing research, and structure a series of planned internal and partner-facing

---

[12] Peter M. Asaro, 'Transforming Society by Transforming Technology: The Science and Politics of Participatory Design', Accounting, Management and Information Technologies 10, no. 4 (October 2000): 257–90, https://doi.org/10.1016/S0959-8022(00)00004-7.

workshop sessions focused on testing and embedding the principles within practice. We are planning a data ethics workshop in autumn 2023 in collaboration with other TaNC Discovery Projects.

Strand E: Public Engagement and Impact

Strand E has focused on early conversations and initial development around the *Transforming Collections x Tate Learning* public programme which will conclude the project. Year 1 saw the shaping of the call for practice researchers in residencies which launched in autumn 2022; and Year 2 was marked by an international conference that took place in spring 2023 in collaboration with Van Abbemuseum.

The four artists selected for the practice research residencies are: Evan Ifekoya, Christina Peake, Erika Tan and Yu-Chen Wang. The artists are undertaking 15-month virtual residencies, engaging with the project's emerging findings, exploring connections across collections, and the existing and emerging interactive ML software as potential investigative tool and creative medium. The resulting physical and/or virtual commissions will critically and playfully intervene in and disrupt partner collections and displays. Works created will be exhibited digitally and/or physically as part of a major public programme curated with Tate Learning and showcased online by Art UK.

Transforming Collections, Rewinding Internationalism was our mid-project public-facing conference held at the Van Abbemuseum in the Netherlands on the 20th and 21st April 2023. Bringing artistic, curatorial, art historical and museological practices into critical dialogue with machine learning development, the conference shared ongoing research and emerging project findings, and took place in the context of the closing days of the exhibition, [Rewinding Internationalism](#) at the Van Abbemuseum, conceived as a generative site of speculation and experimentation. This conference was addressed to artists, curators, researchers, academics, archivists and creative computing technologists interested in exploring decolonial feminist praxes through the interrogative, interruptive and disruptive potential of practice-led research in and with museums. It attracted a specialist audience of c.65 delegates, including representatives from organisations and institutions in Amsterdam, Arhus, Rotterdam, Stockholm, Utrecht, and organisations from across the UK.

Convened by susan pui san lok, Nick Aikens (Curator, Van Abbemuseum) and Scarlett Lösch (Digital Archivist, Van Abbemuseum), the programme opened with exhibition viewings, a performance, and a keynote by transdisciplinary artist and academic, Stephanie Dinkins. Speakers included Amal Alhaag, Sonia Boyce, Alice Correia, Rebecca Fiebrink, Irene Fubara-Manuel, Jon Gillick, Tehmina Goskar, Mick Grierson, Christopher Griffin, Setareh Noorani and Ananda Rutherford. Sessions explored ongoing case studies research on architectures and archaeologies of absence, decolonising data and unlearning erasures, and approaches to disrupting collections. The conference also included a hands-on participatory interactive ML workshop with c.35 active participants.

Edited conference recordings will be published on the Van Abbemuseum YouTube channel in due course.

Project Outputs

Transforming Collections will generate a range of outputs, including:

- Digitisation of iniva's artists archive towards an eventual public research resource;
- Novel adaptable ML software for dissemination across the TaNC programme;
- Critical and interdisciplinary case studies (papers, articles, book chapters);
- An international conference in partnership with Van Abbemuseum;
- A public programme curated with Tate Learning which will include artists' digital/ physical practice research commissions with selected partner collections;
- Selected artists and artworks showcased online with Art UK;
- An accessible project website for general audiences;
- An edited book showcasing artists' digital residencies and artworks.

Year 1 led to the digitisation of iniva's artists archive, a vital step towards iniva's ongoing work the establish this as a new digital research resource. Year 2 has been marked by an international conference in partnership with Van Abbemuseum. Years 2 and 3 will see various critical and interdisciplinary collections case studies presented / in preparation (including conference papers, journal articles, book chapters), and the launch of an accessible project website with video and audio recordings, case studies summaries and reflections. In Year 3, novel adaptable ML software will be disseminated to our project partners and collaborators for beta testing, as well as TaNC projects across the programme. The project will conclude with a public programme curated with Tate Learning which will include artists physical/digital practice research commissions engaging with partner collections. Selected artists and artworks will also be featured in and online showcase with Art UK.

Transforming Collections will culminate in a major public programme featuring four artists' commissions by practice researchers in residency as physical and/or virtual interventions, in gallery and online. Developed as part of Tate Learning's Projects and Programmes, the engagement programme will include participatory workshops and discussions on the ethical, critical, experimental, creative ways in which ML can be activated to transform collections. The proposed public programme will be structured in three parts: A Symposium / Seminar comprising talks, debates and presentations delivered by key stakeholders, to highlight the key ideas, outcomes and outputs of the research project. The event will be aimed at academics, partners, artists within national and international cultural sectors.

The symposium/seminar will be followed by a Large-Scale Evening Programme conceived as a multi-disciplinary event across platforms, with media and digital material installations, performance, sonic based work and discursive components. Artists and creatives will contribute to the programme under guiding themes. The event will represent a celebratory launch moment to increase visibility and advocacy for *Transforming Collections'* broader research ideas, propositions and findings, and the interconnected and collaborative nature of the partnership across institutions and networks.

A Weekend Series of discursive, performance and participatory events will explore the questions and themes of the *Transforming Collections'* research project, with the aim of engaging a broad and diverse audience. Contributors and audiences will be drawn from a range of research, artistic, academic, public and cultural interest and professions to offer a range of perspectives. The series will focus on connection, sharing critical thinking, exploring potential areas of impact and application, including for future museum practice. Both the large-scale evening programme and weekend series will be aimed at the general public – local, national and international.

Cross Project Collaboration

A Cross Project Workshop proposal addressing data ethics is in development. The proposed workshop will focus on introducing the project's data ethics principles and exploring how these can be applied in practice, addressing, for example, positionality, algorithmic bias and injustice, transparency and extraction. Exploratory conversations have taken place with all five Discovery Projects and the workshop is due to take place in autumn 2023.

Sustainability and Infrastructure

The *Transforming Collections* project will generate a range of digital outputs including through data acquisition and processing, research and user studies, technical experiments, archive research outcomes, and the creation of new ML tools and models. These will be hosted on the UAL Creative Computing Institute (CCI) public GitHub for a minimum of 10 years, with resources to support dissemination to partners for beta testing and potential wider roll out. UAL CCI commits to maintaining all publicly viewable information in line with w3c and uk.gov guidance, maintaining open and freely accessible data where appropriate, and acting as a resource for future projects.

The CIDOC-CRM ontology[13] will be tested for coverage against emerging data structures and if proved adequate, it will be used to map datasets for ingestion at the adapter layer of the system. If the CIDOC-CRM is not considered adequate, then a custom global schema will be created to accommodate variances not represented within the existing common set of fields. Mapping will be done either through 3M Editor (https://github.com/isl/3Meditor) or using custom Python scripts. Prior to transforming datasets, an evaluation of entity reconciliation within a ML context will take place as well as a similar evaluation for vocabulary alignment. The transformed data will be tokenised and ingested into the ML system. Freely accessible open data may include CIDOC-CRM classes and properties as semantically valid tokens to be used in ML experimentation. After augmenting collection data with new tokens featuring ontologically rich statements, the resulting data will be communicated back to the partners and audiences as Linked Data based on the CIDOC-CRM global schema using popular serialisations.

Guidance will be provided on how the CRM can record uses and changes prompted by the ML system, stressing the limitations of both the CRM and ML. Long term data storage will also be through UAL's GitHub enterprise facility. This will mean that research can continue beyond project end, including through further funding and with new institutions and partners, subject to appropriate agreements and available data.

---

[13] Chryssoula Bekiari et al., 'Definition of the CIDOC Conceptual Reference Model', 2022, 238.

Interim Recommendations

The Towards A National Collection conference held on 26 April 2023 at the British Museum highlighted shared institutional impulses towards standardisation, diversification, and a sense of opportunity and urgency around the potential for the cumulative efforts and learnings of 30 years of investment into the digitisation of collections so far, to converge through the TaNC programme and consolidate into a digital strategy for dissolving barriers, broadening access, knowledge and representation. There was also a welcome questioning of "what a true digital landscape really means", a call to "truly unlock potential" and for "honest[y] about challenges".

Sessions structured around the themes of 'Innovation', 'Inclusion' and 'Collaboration' belied their mutual entanglement and dependency as necessary conditions for meaningful engagement and world-changing research, that might not only critically challenge understandings of what the 'challenges' are, but also open up and transforms understandings of what a 'national collection' might be or become. For *Transforming Collections*, inclusive practices are fundamental to genuine collaboration and innovation in research that might afford or generate alternative world views and 'worldings'.[14] The simplistic and binaristic characterisation of stakeholders as 'researchers and communities', 'collections and audiences', or 'institutions and individuals' (misrepresented, marginalised or excluded historically and in the present), or the recognition of excluded subjects after the fact, fails to address structural power differentials. 'Barriers' are often invoked as a technological challenge to be addressed in the domain of 'digital innovation', while social and cultural barriers are more likely to considered at the 'end point' of project delivery and dissemination, addressed through a belated 'bringing in' of so-called 'hard to reach' audiences. 'Collaboration' can disguise extractive and instrumentalizing practices, unless frameworks and mechanisms for reflexivity, accountability, attribution are embedded, and the lines between choices, decisions, actions and consequences are drawn.

There is also, perhaps, a tendency to understand "true" in this context as "pure", as in a "pure digital landscape". Rather than seeing the 'digital landscape' as a future space to be 'cleaned up' and removed from the analogue, we might lean into the empowering notion of 'honesty' (if it is reciprocated and sustained) – to embrace the messy complexities and contradictions of the digital and analogue artefacts and collections in the past, present and future, inhabiting a shared and contested continuum. We might then approach the inconsistencies and gaps between analogue and digital records not as problems to be resolved and closed, but as opportunities to expose and explore how partial 'truths' and knowledges from privileged perspectives are constructed and perpetuated through particular periods, contexts and practices. This includes the truth of the hidden labour, entangled exploitations and ecological impacts powerfully demonstrated by the international guest speaker, Horst Hörtner, with the *Anatomy of an AI System*, a map of human labour, data and planetary resources using the example of Amazon Echo to show the countless components and

---

[14] Gayatri Chakravorty Spivak, 'The Rani of Sirmur: An Essay in Reading the Archives', History and Theory, v.24, n.3, October 1985, pp. 247-272.

factors behind the production of artificial intelligence systems.[15] If 'standardisation' does not enable heterogeneities, differences and gaps and hidden relations to remain or become visible, it risks replicating existing harms.

Another international guest, Setareh Noorani, spoke eloquently to the collective collaborative work at the Nieuwe Instituut on 'collecting otherwise', articulating the urgency of committing to collective care and accountability while testing modes of activation, contestation, and negotiation, to invigorate relations with archives and collections. Questions of heritage necessarily prompt questions of provenance, ownership, restitution and reparation too. As such, in the process of "unlocking potential" for greater connectivity and accessibility, we must also ensure greater care in respecting the rights of the living, deceased, and their ancestors, identified and otherwise, whose objects have been variously 'gifted', donated or dubiously acquired. Attention to records and rights necessitates time and human-centred care, to rethink, enact, re-dress, re-visit, and human-scale resistance to uncritical acceleration at scale. Here, 'standardisation', should refer to the standard embedding of ethical considerations, including the Nieuwe Instituut's notion of an 'archival care rider', which might be expanded into an 'archival and collections care rider.'

At the mid-point of the *Transforming Collections* project, we therefore have several interim recommendations for consolidating the work of the TaNC programme and maximising the potential impacts of the five Discovery Projects. These recommendations also underline the impacts of recruitment delays over the first year of the project, the time required to establish effective interdisciplinary and collaborative working across the team, and the value of iterative looped processes:

1. Further resource to deepen and extend critical engagement and evaluation activities beyond the end of the current funded three-year period, triangulating and underpinning heritage information as data and human/digital interactions with ethical practice, e.g. by examining approaches to language across all Discovery Projects.
2. Further resource for an extended programme to roll-out durational testing of our adaptable, interactive ML tools designed to work with existing infrastructures, to enable sustained engagement and gather feedback over time from cross-partner, cross-project, cross-sector and cross-disciplinary and international stakeholders.
3. Emphasis on the need for flexible interactive tools to enhance existing and future infrastructures with critical prompts and reflexive valves that encourage the rethinking of habitual formulations, hierarchies, values and vocabularies, and enable dynamic categorisations or tags to be created and refined by users, 'delinking and relinking' data to surface unexpected relations.
4. Investment in long-term cataloguing work across project and programme partnerships informed by subject specialist researchers and our interactive ML tools, to enrich digital records by foregrounding transparency, context, and absences, not to situate the digital in opposition or progressive relation to the analogue; e.g. making explicit discrepancies with analogue records, or omissions across digital and analogue data, as part of the holistic record.

---

[15] Kate Crawford/AI Now Institute and Vladan Joler/SHARE Lab. https://anatomyof.ai/

Inclusive collaborative practices must underpin and drive inclusive innovation in research. As such, we recommend that future funding criteria should require projects to:

5. Recruit a diverse, inclusive and representative research team with lived personal and professional experiences relevant to the wider questions, aims, objectives and intended audiences and beneficiaries of the project.

6. Engage diverse and directly implicated stakeholders as early as possible in bid development, project conceptualisation, and iterative design and delivery stages, including researchers, artists, estates and archival donors.

7. Value and embed practice-led or practice-based research and close interdisciplinary collaboration; to paraphrase Horst Hörtner: Science creates technologies; but it's the art that helps us to understand their impacts.

8. Value and embed active engagement with ethics and ethical implications of the research in principle and in practice.

Contacts

General enquiries

ContactTransformingCollections@arts.ac.uk

Professor susan pui san lok

Transforming Collections, Principal Investigator

Professor of Contemporary Art and Director, UAL Decolonising Arts Institute

s.lok@arts.ac.uk

Professor Mick Grierson

Transforming Collections, Co-Investigator

Professor of Creative Computing and Research Lead, UAL Creative Computing Institute

m.grierson@arts.ac.uk

Jerneja Rebernak

Transforming Collections, Project & Partnerships Manager, UAL Decolonising Arts Institute

j.rebernak@arts.ac.uk

Fleur Kaminska

Transforming Collections, Project Administrator, UAL Decolonising Arts Institute
f.kaminska@arts.ac.uk

Links

Transforming Collections Project Page:

https://www.arts.ac.uk/ual-decolonising-arts-institute/projects/transforming-collections

Transforming Collections Zotero Library:

https://www.zotero.org/groups/4566230/transforming_collections/item-list

References

Andersen, Margaret L., and Patricia Hill Collins. 'Why Race, Class, and Gender Matter'. In *Inequality in the 21st Century*. Routledge, 2018.

Asaro, Peter M. 'Transforming Society by Transforming Technology: The Science and Politics of Participatory Design'. *Accounting, Management and Information Technologies* 10, no. 4 (October 2000): 257–90. https://doi.org/10.1016/S0959-8022(00)00004-7.

Azoulay, Ariella A. *Potential History: Unlearning Imperialism*. London; Brooklyn, NY: Verso, 2019.

Bekiari, Chryssoula, George Bruseker, Erin Canning, Martin Doerr, Philippe Michon, Christian-Emil Ore, Stephen Stead, and Athanasios Velios. 'Definition of the CIDOC Conceptual Reference Model', 2022, 238. https://www.cidoc-crm.org/sites/default/files/cidoc_crm_version_7.2.1.pdf

Chun, Wendy Hui Kyong, and Alex Barnett. *Discriminating Data: Correlation, Neighborhoods, and the New Politics of Recognition*. Cambridge, Massachusetts: The MIT Press, 2021.

Costanza-Chock, Sasha. *Design Justice: Community-Led Practices to Build the Worlds We Need*. Information Policy. Cambridge, Massachusetts: The MIT Press, 2020.

D'Ignazio, Catherine, and Lauren F. Klein. *Data Feminism*. Strong Ideas Series. Cambridge, Massachusetts: The MIT Press, 2020.

Dixon, Lucas, John Li, Jeffrey Sorensen, Nithum Thain, and Lucy Vasserman. 'Measuring and Mitigating Unintended Bias in Text Classification'. In *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society*, 67–73. New Orleans LA USA: ACM, 2018. https://doi.org/10.1145/3278721.3278729.

Dunbar-Hester, Christina. *Hacking Diversity: The Politics of Inclusion in Open Technology Cultures*. Princeton Studies in Culture and Technology. Princeton, New Jersey: Princeton University Press, 2020.

Risam, Roopika. New Digital Worlds: Postcolonial Digital Humanities in Theory Praxis and Pedagogy. Evanston Illinois: Northwestern University Press, 2019.

Sharma, Shubham, Yunfeng Zhang, Jesús M. Ríos Aliaga, Djallel Bouneffouf, Vinod Muthusamy, and Kush R. Varshney. 'Data Augmentation for Discrimination Prevention and Bias Disambiguation'. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 358–64. New York NY USA: ACM, 2020. https://doi.org/10.1145/3375627.3375865.

Smith, Chris, Stuart Hall, Maya Jaggi, and Arts Council of England. Whose Heritage?: The Impact of Cultural Diversity on Britain's Living Heritage : National Conference, Manchester, 1st-3rd November, 1999 : Keynote Addresses. London: Arts Council of England, 1999.

Spivak, Gayatri Chakravorty. 'The Rani of Sirmur: An Essay in Reading the Archives'. *History and Theory* 24, no. 3 (1985): 247–72. https://doi.org/10.2307/2505169.

Wekker, Gloria. *White Innocence: Paradoxes of Colonialism and Race*, Durham: Duke University Press 2016.

Further References

Aikens, Nick, susan pui san lok, and Sophie Orlando, eds. *Conceptualism - Intersectional Readings, International Framings*. Eindhoven, The Netherlands: Van Abbemuseum, 2019, https://mediabank.vanabbemuseum.nl/vam/files/alexandria/publiciteit/e-pub/2019/vam_epub%20BAM_DEC19_V5_final3.pdf

Boyce, Sonia and Dorothy Price eds. Art History; Journal of the Association for Art History. 'Rethinking British Art: Black Artists and Modernism', June 2021.   https://arthistoryjournal.org.uk/issues/june-2021/.

Dalal-Clayton, Anjalie, and Ilaria Puri Purini. *Doing the Work: Embedding Anti-Racism and Decolonisation in Museum Practice*. London, UK: Contemporary Art Society and UAL Decolonising Arts Institute, 2022. https://www.arts.ac.uk/ual-decolonising-arts-institute/publications-and-podcasts.

Dalal-Clayton, Anjalie, and Ananda Rutherford. 'Against a New Orthodoxy: Decolonised "Objectivity" in the Cataloguing and Description of Artworks – Features – Paul Mellon Centre'. Accessed 2 May 2023. https://photoarchive.paul-mellon-centre.ac.uk/groups/against-a-new-orthodoxy.

lok, susan pui san. 'Through the Gate / An(g)archivery' in *Deviant Practice*, pp. 176-209. Eindhoven, The Netherlands: Van Abbemuseum, 2018.

Annex

Collections Audit Report: Dr Anjalie Dalal-Clayton & Ananda Rutherford

Stage 1

Introduction

This report outlines the findings of the Transforming Collections project's audit of artworks by 'B/black' artists (defined here as racialized and minoritised artists of African, Caribbean, Asian, or MENA-region descent) in public art collections in the United Kingdom. The audit was originally carried out as part of the AHRC-funded Black Artists & Modernism project (BAM, 2015-2018), and obtained basic inventory data on works by Black artists in 31 national and civic collections.

The audit has been subsequently updated and expanded as part of the Transforming Collections project and this comprised two stages. In stage 1, the data gathered for the BAM audit was updated to account for artworks acquired since 2016 (when the BAM audit ended). In stage 2, the scope of the audit was extended to include descriptive and interpretive information held in files, records, and other documents on a selection of artists and artworks from each collection.

This report presents the data obtained in, and an analysis of, the Transforming Collections audit. The initial findings relate to data gathered from 31 collections, 11 of which are held by major national institutions, and 20 by civic museums and galleries situated outside the four capitals. Eight of the audit participants were project partners.

Purpose of Research for Stage 1

The audit sits within strands A and B of Transforming Collections, and specifically responds to the research question: whose voices, bodies and experiences are centred and privileged in collections, and whose are unheard, unseen or erased? The basic inventory data gathered in stage 1 of the audit answers the 'who' and 'what' of this question, whilst the object history files, and descriptive and interpretive texts obtained in stage 2 answer 'how'. Further, the data gathered in both stages has been used as a jumping-off point for researchers working within strands A and B, in their in-depth investigations into particular artists and collections, and as a dataset upon which researchers working on strands C and D can test their machine learning (ML) prototypes.

Scope for Stage 1

Collections Geography

Within the BAM project, the collections were selected for the audit based on their significance within the exhibitions and collecting history of the British Black Arts Movement, and, with at least one collection based in every region of the nation, the project team considered the selection as potentially indicative of collecting histories and practices across the UK. This same selection provided the basis for the *Transforming Collections* audit, with the addition of project partner collections that had not participated in the original BAM audit, but without the participation of one non-partner collection with which contact could not be successfully made.

List of Audited Collections (*project partner)

Arts Council Collection*
Birmingham Museum and Art Gallery
(Birmingham Museums Trust)*
Bristol Museum and Art Gallery
The British Museum
Cartwright Hall, Bradford
Glasgow Museums*
Glynn Vivian Art Gallery, Swansea
The Government Art*
Graves Art Gallery, Sheffield Museums
The Harris Museum and Art Gallery, Preston
Imperial War Museums
Leeds Art Gallery
Leicester Museum and Art Gallery
Manchester Art Gallery*
Middlesbrough Institute of Modern Art *

National Museums Northern Ireland
National Museum, Wales
National Portrait Gallery, London
Newport Museum and Art Gallery
Norwich Castle Museum and Art Gallery
Nottingham City Museums and Galleries
Pallant House, Chichester
The Royal Academy, London
The Scottish National Gallery of Modern Art
The Scottish National Portrait Gallery
Southampton City Art Gallery
Tate*
The Herbert Gallery, Coventry
The McManus, Dundee
The Walker Art Gallery, Liverpool*
Wolverhampton Art Gallery

Efforts were made to audit the Wellcome Collection, which is a project partner, but due to the particular organisation of their collections database, this was not possible.

Artists

There were three criteria for inclusion: artists whose work in the collections dates between 1900 and the present day, and who were either born, raised, worked, or studied in the UK, and who are of African, Caribbean, Asian, or MENA-region descent.

These criteria were developed from, and premised on, the BAM project's use of the political term 'Black,' as conceived and mobilised during the British Black Arts Movement of the 1980s (for a critical discussion of the ethical dimensions of this ethnic framing, see section 5).

Artworks

Works of fine art and photography, as well as film, metalwork, textile and clothing, ceramics and jewellery were included in the audit, if the objects were created primarily for the purpose of display (i.e., not social history objects).

Date Range

Artworks/objects made and acquired after 1900 were included in the audit, in respect of the original BAM audit's focus on modern and contemporary art.

Methodology for Stage 1

The method employed for stage 1 was systematic, data-driven, and designed to be carried out remotely using a set of search terms and parameters that reflected the audit scope. It was carried out in four steps, via database searches, by curators, registrars, or information managers at each of the participating collections:

Step A

The participating museum/gallery exported a long list of artists whose work was made and acquired after 1900 from their collections databases.

Step B

The project researchers scanned the long list to identify artists meeting the inclusion criteria. This typically involved online biographic research to determine eligibility for inclusion (for a critical discussion of the ethical dimensions of this part of the process, see section 5).

Step C

A short-list of artists meeting the inclusion criteria was returned to the museum/gallery. The curators/information managers were then invited to suggest additional names for inclusion, and to confirm a final shortlist. Once confirmed, they produced a list of concomitant artworks in their collection, containing the following basic inventory data:

- Artist name
- Artist birth and death years
- Artist place of birth/residence/nationality
- Artwork title
- Artwork date
- Artwork materials
- Artwork dimensions
- Artwork acquisition year
- Artwork acquisition circumstances – i.e., gift/bequest/purchase etc

Step D

The data was then imported into a centralised list for all collections in preparation for analysis. The data fields for Step C had been pre-determined during the BAM project and were selected in order to respond to the following questions:

- How many artists and artworks are there across the dataset and in each collection (quantity)?
- Which particular artists are most collected both across the dataset (quantity)?
- Which artists appear across the most collections (breadth)?
- What is the most collected generation of Black artists in a given collection/ across the dataset?

- What is the most collected decade/period of artmaking?
- In which decade were the most acquisitions made?

Ethical Issues

For stage 1, the ethical issues pertained to the scope and methodology:

Use of the Contested Term 'B/black' in the Audit Scope

In the BAM project, the researchers held varied perspectives and positions on the meaning of the term 'B/black' but collectively decided that the project would engage the term in a broad political sense, derived from discussions that crystalised in the late 1970s and 1980s about a shared experience of racist oppression amongst people descended from former British colonies in Africa, Asia, the Caribbean, and MENA-region. Although few British born/based people of Asian and MENA-region descent now define themselves as B/black, the geographic parameters of the term are echoed in present-day cultural diversity policies, through the similarly contested term Black and Asian Minority Ethnic (BAME). It was therefore decided in BAM that the term had relevance to, and resonances with, present-day culture and heritage sector contexts. However, applying this use of the term to both the BAM and Transforming Collections audits raised an ethical issue relating to the term's concomitant geographic parameters. A specific concern arose in relation to the geographic ambiguity of the MENA-region and the inclusion of, for example, artists with Arab-Palestinian backgrounds, and exclusion of artists with Jewish-Israeli backgrounds. This, in turn, raised concerns amongst the researchers about 'hierarchies of racisms,' and which ethnic and religious groups are perceived within the research as being racialised and minoritised and which are not.

The researchers are continuing to reflect on the ethical appropriateness of fully adopting the reductive parameters for the audit as developed for BAM, within the Transforming Collections project, which has different concerns and objectives.

Racial Categorisation within the Methodology

In order to determine if artists met the audit inclusion criteria, researchers relied on their existing subject-specialist knowledge of the artists and their ethnic backgrounds. However, since not all relevant artists were known to the researchers, further online biographic research was required. This approach was inevitably subjective and will undoubtedly have been impacted by unconscious, internalised biases, despite its affirmative intent. The inherent categorisation articulated in the audit arguably replicates wider societal perceptions of race and ethnicity which result in the exclusion and underrepresentation that the audit has sought to surface. In other circumstances, the audit could have been undertaken differently to include the informed consent of the artists impacted and with respect and acknowledgement of their preferred terms.

Repurposing and Replicating the Audit Methodology and Data by Other Organisations

Underrepresentation of racialised and minoritised artists in public collections is understandably an issue that art institutions are seeking to address. Many UK and international museums and galleries that were aware of the BAM/ Transforming Collections audit, but were not invited to participate, have requested the details

of the audit scope and methodology with a view to applying it to their own collections. Given the ethically problematic aspects of the audit, the researchers felt it irresponsible to offer instruction and have declined these requests. The researchers are also aware that because the audit was originally conceived and devised within BAM, the wider museum and gallery sector might assume that the audit has a scope and methodology that has been validated or legitimised by researchers and artists of colour and can be uncritically replicated.

Data

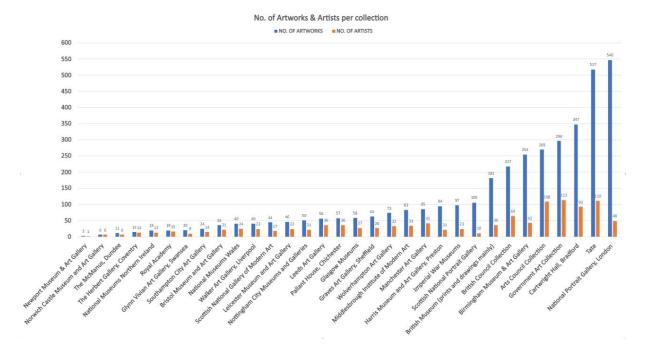All basic inventory data has been compiled within a single, large excel sheet, available on request.

Findings and Discussion

How many artists and artworks are there across the dataset and in each collection (quantity)?

The audit of 31 collections identified 3,767 artworks (2,350 in nationals and 1,417 in civics) by 578 individual artists and artist groups.

In the BAM audit, an approximate proportion of the identified artworks was presented as between <1% and 4% of all artworks made and acquired within the same date range, held by the participating collections. However, within the given time frame for conducting the audit for the *Transforming Collections* project, it was not possible or a priority to establish accurate proportionality for the updated data. The graph below shows how many artists and artworks were identified in each collection.

No. of Artworks & Artists per collection

Unsurprisingly, larger institutions which have a national remit, larger collections overall and arguably bigger acquisition budgets appear to hold a greater number of works by B/black artists than smaller regional collections which have a civic remit and arguably smaller purchasing budgets.

The number of artworks (shown in blue) may be misleading as an indicator of collecting priorities and acquisition activity; in some collections the number of individual artworks actually comprises numerous prints or prints series that have been created by a single artist (for example, of the 546 artworks held by the National Portrait Gallery, 305 are photographic prints by the artist Lewis Morley).

A better indication of collecting priorities and activity might be the numbers of artists (shown in in orange). In either case, it is interesting to note that Cartwright Hall in Bradford and Birmingham Museum and Art Gallery (both regional civic museums) appear to hold a comparable number of both artists and artworks to those at larger national collections such as Tate, the Government Art Collection, and the Arts Council Collection. This raises the question of whether the interests and concerns of individuals responsible for collections development, a civic remit and responsibility to represent local demographics has more impact on collecting activity than the more easily comparable factors of collection size and acquisition budget. What is harder to determine is the symbiotic impact of socio-political contexts and the international art market on acquisitions made by both types of collecting organisation.

Which particular artists are most collected both across the dataset (quantity)?

We organised the data into tables to show which artists have been collected most according to:
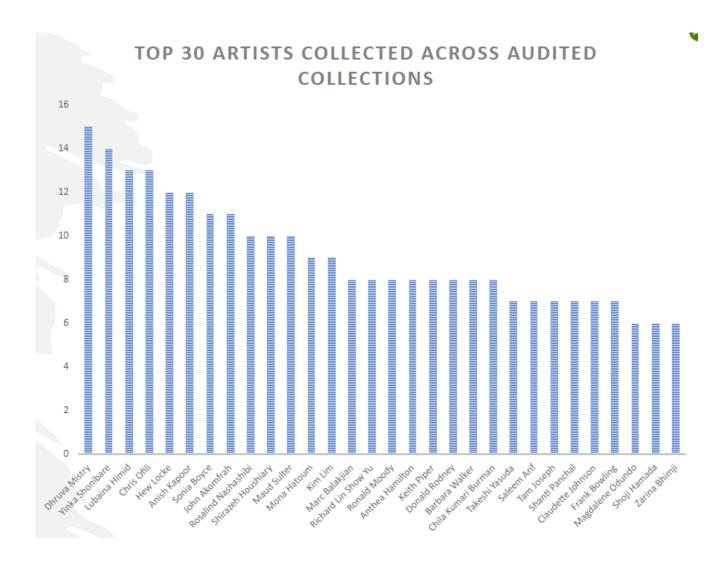
- number of works across the dataset (table 1), and
- number of collections their work has been acquired into (table 2).

Table 1 below indicates the top five most collected artists across the dataset. It is interesting to note that three of these artists are photographers, and two are sculptors, suggesting that while the use of print media can impact an artist's collectability (in terms of quantity), certain artists have been collected in comparatively large numbers for other reasons – further research could be done to explore the reasons behind this.

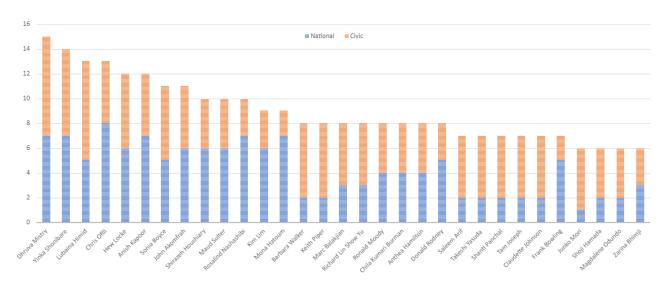| Artist by Collection	No. of works
1 Lewis Morley	**305**
National Portrait Gallery	305
2 Vanley Burke	**121**
Birmingham Museum & Art Gallery	105
Arts Council	12
Graves Art Gallery, Sheffield	2
The Herbert Gallery, Coventry	1
Nottingham City Museums and Galleries	1
3 Anish Kapoor	**120**
Tate	49
British Museum	31
British Council Collection	27
Arts Council	3
Cartwright Hall, Bradford	2
Walker Art Gallery, Liverpool	2
Wolverhampton Art Gallery	1
Scottish National Gallery of Modern Art	1
Government Art Collection	1
Pallant House, Chichester	1
Royal Academy	1
Leeds Art Gallery	1
4 Kim Lim	**93**
Tate	59
British Council Collection	17
British Museum	6
Arts Council	3
Pallant House, Chichester	2
Scottish National Gallery of Modern Art	2
Government Art Collection	2
Southampton City Council	1
Graves Art Gallery, Sheffield	1
5 Pradip Malde	**92**
Scottish National Portrait Gallery	84
Arts Council	5
The McManus, Dundee	2
British Council Collection	1

Which artists appear across the most collections (breadth)?

The table below shows the number of different audited collections that an artist's work appears in. This selection indicates the 30 most collected artists in audited collections, with Dhruva Mistry's work present in 15 of the 32 audited collections. By contrast 418 artists have a single work in a single collection.
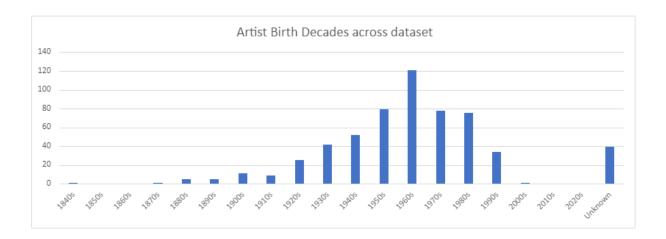
TOP 30 ARTISTS COLLECTED ACROSS AUDITED COLLECTIONS

The table below shows the same information, disaggregated to indicate the breakdown of distribution between civic and national collections.

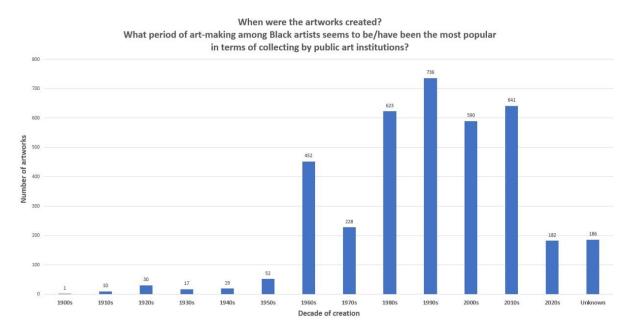## What is the most collected generation of Black artists across the dataset?

The graph below shows the years in which artists identified through the audit were born, by decade. In the BAM audit, the majority of artists were born in the 1950s and 60s, which corresponded with the generation of B/black artists who participated in the British Black Arts Movement of the 1980s. This pattern of acquisition has continued and grown, with increased numbers of works by this generation of artists being added to collections, but as the graph below also indicates there have been increasing numbers of works collected from artists born in the 1970s and 1980s, which at the time of the BAM audit was not the case.

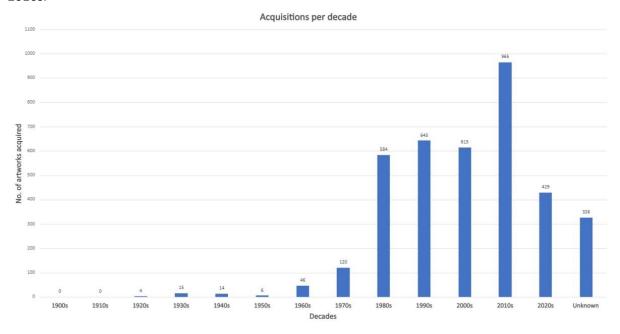## What is the most collected decade of artmaking?

The graph below shows when the works identified through the audit were created, by decade. In the BAM audit, completed in 2016, works made in the 1980s were greatest in number. However, in the

six years since the BAM audit was conducted, there are now equivalent quantities of artworks created in the 1990s, 2000s and 2010s, suggesting that museum interest in B/black artists has broadened beyond the context of the British Black Arts Movement of the 1980s.
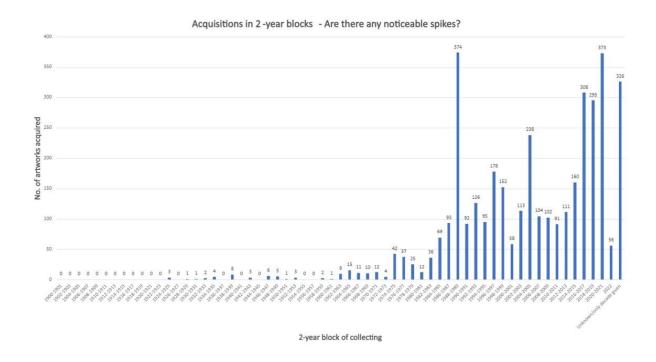

When were the artworks created?
What period of art-making among Black artists seems to be/have been the most popular in terms of collecting by public art institutions?

In which decade were the most works by Black artists collected?

The graph below shows when the artworks, identified through the audit, were acquired, by decade. It evidences that collecting activity across the collections rose significantly – by about a third - in the 2010s.

Acquisitions per decade

To gain a clearer or more specific picture of this phenomenon, the data on acquisition years was split into two-year blocks (rather than 10). In the resulting, graph, below, there are two clear spikes in collecting activity in 1989-1990 and 2004-2005. Then, looking across the decade from 2010, a

significant rise in collecting occurs after 2016, and specifically, in 2020-2021. Further, qualitative research, using the stage 2 audit data, as well as other research methods such as archival research and interviews with museum workers is required to investigate the details behind and reasons for these spikes, which may include the membership of B/black artists on acquisition committees, the impact of the BAM audit itself, and broader socio-political developments and movements, such as Black Lives Matter.

Acquisitions in 2-year blocks - Are there any noticeable spikes?

Stage 2: Beyond the Tombstone Data

Introduction

Stage 2 of the audit was designed to look beyond the inventory level data gathered for Stage 1, in order to explore more closely the information generated and held by the participating collections on the artists and works identified in Stage 1.

Further material and information were required to begin to answer the questions emerging from the patterns seen in the Stage 1 data. Why were certain works acquired and how were they funded? How have they been historically described, presented, and used? How are they understood within the institution? What do the files and records tell us beyond the highly mediated online presentation?

Purpose of Research

The purpose of Stage 2 was primarily a data-gathering exercise for collection, object, and artist information held by the participating institutions in their files and archives. Our request was for any and all material (within reason) relating to the artists identified using the Stage 1 audit criteria.

While there are increasingly significant amounts of images and information provided publicly on collections websites and via aggregators such as Europeana and Art UK, most researchers will be aware that the records presented online are usually highly edited or carefully delimited and that collections tend to hold files of information charting the life of the object within the museum or gallery as well as documents about its previous ownership, display and use. These files are usually not catalogued or digitised but can hold crucial information for research purposes. In some instances, they are not maintained once a collections database has been introduced, and in other instances collections continue to run paper-based archives in parallel to the database but hold different information.

The purpose of attempting to access the information held "behind the scenes" was:

- To provide initial data about holdings on Audit artists across the collections and to be able to interrogate stage 1 data further
- To provide initial data about holdings on artists identified in the Audit and across the collections for the Post-doctoral Research Fellows
- To be able to demonstrate to the wider team what museum records and documentation looks like, where information might be held if not digitally accessible, and to evidence the limitations of collections information
- To gather a body of relevant descriptive texts for analysis and comparison from racialised and minoritized artists
- To gather a specific body of texts as a dataset for Machine Learning training

Scope for Stage 2

26 of the 31 participating collections responded to the stage 2 data request, which was:

Paper Records and Files

- Object files and/or Artist files
- Gallery/exhibition/display files
- Card indexes other legacy systems for
- Object and artist information
- Day books and registers
- Labels
- Press releases

Collections Database Records from Current and Legacy Systems

- Object history
- Content description
- Associated people, places, and events
- Authority and selected term lists (local)

Policies and Procedures Documents

- Collection policies
- Documentation procedures, data entry manuals
- Board papers or similar that relate to decision making around the display or acquisition or other use of a work

Procedural Records

- Entry and Acquisition records and files
- Exhibition/Loans records

Web/Digital Material

- Online or other digital material relating to works or artists.

Methodology for Stage 2

Each collection was asked to supply as much information as they felt able to, on each of the artists and works identified on the audit list. Constraints to the data gathering process included:

- Physical/logistical access and retrieval – several collections held their files off site or in numerous locations, departments, or stores
- Resources – time, staff availability
- Data protection, IPR and consent
- Supply and storage – how large numbers of digital files and images might be supplied and stored

- Institutional and staff memory and knowledge – older or legacy filing or digital systems that had fallen out of use

Accessing the texts, descriptions, publications, and information that is not publicly available required negotiation with each collection and navigation through their particular systems and conventions for information management. In order to support each institution to assess and project the staff time required to complete the information request, we used a method of piloting the retrieval and digitisation of files and records for 2-5 artists/artworks.

The selection of the pilot examples was based on the following criteria

- Relevance to potential PDRF (Post-Doctoral Research Fellow) research interests
- An example of work that had been acquired in the early history of the institution
- An example of work that had been recently acquired or displayed
- Work or artists that were of particular interest to the collection

This piloting process allowed each collection to assess the types of information held, its location, volume, shareability, the redactions and restrictions required, time required for digitisation, and the resulting impact of this work on ongoing staff workloads. Several collections were unable to continue after this piloting process due to lack of resource to complete the entire data request. We remain incredibly grateful for all efforts made by the staff we worked with, who did their utmost to support our request, often in difficult circumstances.

Ethical Issues

Several ethical issues emerged in relation to the scale of this information request. The sensitivities around the collecting and use of data relating to race, nationality and ethnicity are set out in Stage 1, above. A further issue, which we attempted to mitigate in practice, related to data protection and intellectual property rights. Understanding and practice around sensitive data was variable across collections and especially problematic around research that required assessment of race, nationality, and ethnicity to achieve its aims. Copyright and GDPR (General Data Protection Regulations) were also inconsistently understood in terms of the individual artist and the rights of the holding institution. In practice organisations were cautious by policy and fear of risk, and individuals were concerned and caring about exposure or harm to individuals. It has been necessary to fully check all documents supplied and to delete, anonymise or redact a proportion of the material in order to protect all parties involved.

Data Gathered

While the process has been slower than anticipated, despite much good will and generosity, 26 out 31 of the participating collections supplied partial data (e.g. all database records with a selection of specific fields, but no further paper records), a selection of material for a pilot selection of artists, or full data for stage 2 of the audit (everything permissible to share for all artists and works on the audit list).

Three of the collections have or are in the process of supplying records for all their Black artists who feature in the audit selection.

It is difficult to determine the total number of records, images, and contextual material supplied because some single files contain multiple records and documents, but we have now gathered 1000 separate files. These files provide further information on 464 works of the 3767 works audited (just over 12%) and 119 of 593 artists (20% of the artists named in the audit).

Findings and Discussion

We are still in the process of assessing the material, but our findings fall broadly into two categories:

- findings from and about the information held
- findings about the processes and practices of information management in these collections.

Initial findings from the supplied material and its content

Several collections hold work by racialised or minoritised artists from the early part of the organisations' histories (in some cases, as early as 1920s) but there is little held on record about them (perhaps due to lost records, or information languishing in systems no longer used) and the works may not have been on display or engaged with since their acquisition.

By contrast, in cases where there has been recent engagement with a particular artist and artwork, there is often significantly more detailed information on the record/file (e.g., Barbara Walker and Maud Sulter). There were also pockets of material around a specific moment of engagement in the past (usually an exhibition), but very little material or update information since that time.

Records and information vary noticeably in volume and consistency between institutions. This is true even for well-known artists, for whom there can be a substantial body of published information. In such cases, one holding institution may have detailed and up to date information on the artist and artwork, but another institution may have little beyond an inventory record. Frequently, there is also a repetition of information across several collections, drawn from a particular exhibition or publication.

This material provides a basis for further strands of research:

- The documents and records supplied have provided evidence of bias, racialisation, perpetuation of colonial perspectives, but also changes in interpretation;
- It reveals the context for acquisition and funding practices;
- It surfaces of the work and careers of individual artists and engagement with, and reception of their work.

Findings about institutional information management

In working with a diverse range of national and civic collections, the researchers were able to see first-hand the variety of ways information is managed across the sector. The size and complexity of the larger national collections was, in some cases a barrier to further access and gathering material.

The involvement of multiple departments, people, policies, and systems, combined with a high volume of works and artists made the process prohibitively time consuming and resource heavy.

For all organisations, the issue of legacy information management systems (digital and analogue) also made locating relevant material an involved process for the researchers and the collections and documentation staff. Additionally, staff turnover, as well as recent redundancies and the impacts of the Covid Pandemic in 2020 seemed to affect both physical access, and memory and knowledge of what material was held and where. That said, the efforts made by all collections and documentation staff we engaged with was uniformly positive and generous.

While these findings do not perhaps speak directly to the main questions of the project, they remain useful as illustrative of the complex issues of access to collections information and evidence for the development of machine learning tools.

Positive Impacts from the Research Process

The remit of the research and engagement with several collections generated or complemented some interesting further work with participating organisations. In several cases collections and documentation staff expressed interest at the opportunity to engage with the work of the artists within their collections.

This resulted in:

- Collections prioritising documentation and cataloguing work around the audit artists
- Collections finding and re-engaging with artists and works
- Collections wanting to foreground this work in their own displays and research.

The question remains, however, as to why in many cases had this not happened before?

Conclusion

By auditing collections against protected characteristics museums and researchers can ascertain and evidence the degree to which particular artists and groupings of artists are represented. Such data may be essential in supporting individuals to lobby for changes in collecting policy and practice to improve representation of certain marginalised or underrepresented sections of society. However, recent global social justice movements have more than adequately argued for the need for increased representation and visibility of racialised and minoritised artists in GLAM collections. To what extent, then, can the information resulting from this type of audit prompt positive change in collecting policy and practice, and conversely how might it be mobilised or instrumentalised to the detriment of those artists included?

Furthermore, when the premise and methods of an audit such as this one are based on deeply problematic, scientifically and ethically questionable racialised classifications, can disaggregation of collections data convincingly support researchers to argue for such changes, or does it simply perpetuate division?

The *Transforming Collections* stage 1 audit data is understood by the project researchers to have significant limitations. It should be regarded only as a broad indication of collecting practices during the period under review and with a specific selection of public collections. It can be treated as a 'jumping off point' for further research, rather than as a comprehensive assessment. A key question for ongoing consideration within the project is the impact of auditing, and research more generally, on artists, and moreover, the ethical considerations that need to be made before research even begins.

Finally, there remains the question of how existing audit-based research can lead towards concrete action and change in collecting policy and practice. While the project researchers may use the stage 2 data to evidence bias in language, sporadic engagement, and the specifics of interpretation and reception, in the careers of particular artists, wholesale application of this audit's methodology is to be discouraged in favour of inclusive, collaborative, and dialogic collections evaluation and development with artists.