Designing for Appreciation: How Digital Spaces Can Support Art and Culture

Thomas Şerban von Davier

St. Hilda's College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2024

A poem begins in delight and ends in wisdom
- Robert Frost
A thesis is much the same.
- Thomas Şerban von Davier

Acknowledgements

A project of this scale is never the work of one person alone, and my DPhil is no exception. While my time at Oxford was spent exploring the complex relationship between art and artificial intelligence, my network kept me nourished, entertained, and, most importantly, grounded. I would like to take this opportunity to express my gratitude to them all.

First, to my parents—thank you for your unwavering support and for encouraging me to apply in the first place. Our countless conversations, from before I began this journey to its conclusion, helped me reflect and shape my research into a coherent body of work. I am also deeply grateful to my supervisors, Professor Max Van Kleek and Professor Sir Nigel Shadbolt. Your guidance, feedback, and clarity over the years were invaluable in my growth from an independent thinker into an independent researcher. I would also like to thank Professor Jodi Forlizzi and Professor Kathryn Eccles, my assessors, for their critical insights and feedback at key milestones. Your contributions have challenged and shaped this work into what it is today.

Equally important are my friends and colleagues, who made my experience at Oxford truly special. To Anna Jamieson, Grant, Heather, Ryan, Philip, Jamie, Alicia, Malaika, Margaret, Tejas, and JP—thank you for being such wonderful and close friends. I am also grateful to Esham and Jean, who have been with me since childhood, continuously motivating me along the way. To my housemates at Alma Place—Bas, Alice, Reuben, Doug, Pippa, Josh, Amy, Abel, and Mina—thank you for the unforgettable times we shared. I also want to express my gratitude to Guillermo, Aaron, and Sandy for making St. Hilda's College a place filled with fond memories and good times gobbling formal meals and entertaining friends.

To my colleagues in the HCC group, I wish you all the best in your future endeavors. Thank you, Jake, Laura, Tyler, Hayoun, Neil, Tala, Jumana, Ben, Konrad, Sid, Ulrik, Claudine, Sara, Sarah, Lize, Tiffany, Zhilin, Ben, Sruthi, Jesse, Rui, and Wael, for your collaboration and support throughout this journey. I am also grateful to the OII Creativity and Technology Group for giving me the opportunity to learn from your perspectives and connect with researchers passionate about the intersection of art and AI. A special thank you to my collaborators—Hank, Yu Ju, Sauvik, Danilo, Gilles, Laura, Caterina, and Emre—for exploring new areas of research with me and expanding my horizons within the field of HCI.

Finally, as my work resides within the field of Human-Computer Interaction, I extend my sincere thanks to all the participants, assistants, administrators, and others who made this research possible. Without the people at the heart of this work, there would be little purpose in building these systems in the first place.

Abstract

This thesis investigates the intricate dynamics of the intersection of art and computer science, particularly focusing on the roles of creation, curation, and appreciation within digital spaces. As art increasingly interacts with algorithmic curation, understanding how these processes shape audience perceptions and engagement is crucial. We analyze the transformation of artists' mental models into artworks through digitization and algorithmic experiences (AX), emphasizing the impact of these processes on public appreciation of art.

The research centers on four primary initiatives. First, we examine the lived experiences of artists, curators, content creators, and the public regarding "contentification"—the reduction of diverse artistic expressions into generic content. Interviews and surveys reveal four essential features for appreciating art: depth, conversation, connection, and time. Co-design workshops with art experts generated guidelines for enhancing digital art experiences.

Next, we explore the distinctions in perception between human and algorithmic curation through the *Algorithmic Pedestal* project, highlighting the subjective versus objective interpretations of art. This analysis underscores the inherent limitations of computational methods in capturing the emotional and contextual nuances that human audiences value.

In the third initiative, we introduce *AppraiSet*, an open-source dataset of art metadata designed to address gaps in current algorithmic processing. This resource facilitates deeper understanding of art through machine learning and highlights how qualitative data can enrich algorithmic art analysis.

Lastly, we present *ArtBot*, a Socratic LLM prototype aimed at guiding users through artwork analysis. Through experimental evaluations, we demonstrate that ArtBot enhances critical engagement compared to traditional and social media contexts. The findings contribute to a broader discussion on the role of algorithms in shaping art appreciation, advocating for an integration of human agency in algorithmic curation processes.

Overall, this research illuminates the evolving landscape of art curation in the digital age, proposing actionable insights and frameworks for enhancing audience experiences and engagement with art and culture.

Contents

List of Figures

Much of this chapter was also reviewed and published in the paper:
Thomas Șerban von Davier. 2023. Designing for Appreciation: How Digital Spaces Can Support Art and Culture. In Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems (CHI EA '23). Association for Computing Machinery, New York, NY, USA, Article 490, 1–5. https://doi.org/10.1145/3544549.3577041

1
Introduction

Contents

1.1 Motivation

The international art community is a social network of vast complexity and size as it connects artists with curators and audiences worldwide. As such, how this network functions has a profound social impact on the world. The network encompasses three primary actions creation, curation, and appreciation. All three have evolved overtime to integrate the latest technology. We represent the current state of the system in Figure 1.1.

Artists convert their mental model into an artwork. This piece is then digitized by a device and uploaded. At this point the artwork is presented and formatted by

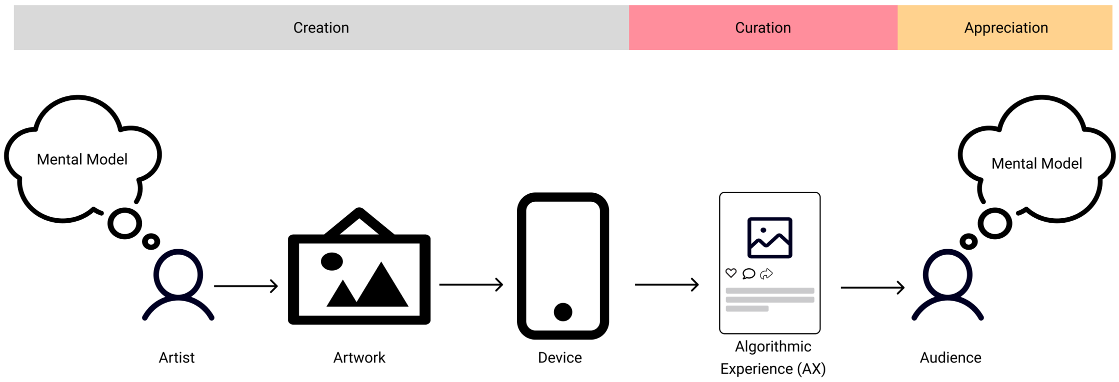

Figure 1.1: This diagram shows how the mental model of the artist is being transferred via multiple stages towards the audience. The creation, curation, and appreciation cycles are each impacted by the involvement of algorithms throughout the experience.

an algorithmic experience (AX) [13] maintained by a platform. Finally, audiences see this display on their own devices and form their own mental model based on this version of the artwork within the context of the AX. The creation process refers to how an artist's mental model is shaped into an artwork and then digitized by a device. At this stage the curation happens through the AX provided by digital platforms which is responsible for presenting and organizing the information related to the artwork. Appreciation is the outcome from the previous steps and is shaped by the curation process. Designers and computer scientists will see the similarity to the mental model diagrams from Don Norman (2002, originally 1988) in the field of human-computer interaction (HCI) [242].

When creating a product, designers and developers are only able to communicate their mental models and expectations to the user through the product itself (see Figure 1.2). As this continues to be difficult, the field of HCI has grown to involve multiple methods to build better products that reflect user needs and communicate mental models more clearly. This thesis argues the need for human centered artificial intelligence (HCAI) researchers to perform the same function on how our digital

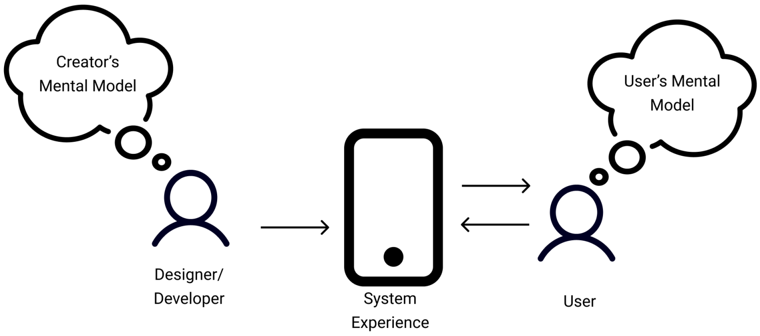

Figure 1.2: The standard HCI mental model that has been drawn from Norman's book [242]. It depicts the importance of HCI to align designer and user mental models.

spaces present art to audiences worldwide. Specifically, the work of this thesis will focus on the curation and appreciation aspects of Figure 1.1.

The practice of art curation, recommendation, and appraisal has become the next challenge for data science and computation. As art was created, humanity worked to maintain and display it, from religious rituals to wealthy patrons and private collections, then in public museums, and now with digital platforms [18]. Modern algorithmic recommendations increasingly mandate the curation and presentation of art and information to the public. In this thesis, I am proposing an exploration of the impact of the current algorithmic curation on the public's ability to appreciate and perceive artistic and cultural experiences.

The proposed research will explore select components of art and culture recommendation algorithms, from advanced neural networks such as deep learning approaches to collaborative filtering. The work will compare the effectiveness of current algorithms to recommend visual art and consider the social and societal impact of said algorithms on cultural experiences. The methodology to explore this area of research requires the development of datasets of test art metadata. This analysis will also investigate the issues of fairness and the impact of algorithms on recommending new artists, artists from minority groups, or non-Western cultures. Furthermore, we will develop prototypes with differing degrees of human agency to visualize various potential futures of algorithmic curation. Previous research argues that the underlying platforms for the recommendation and classification algorithms, such as social media and digital streaming services, are replacing cultural intermediaries of the past, such as curators, newspapers, radio, and television [228]. The impact of this shift towards 'infomediaries' on users' tastes, perceptions, and appreciation of art and culture has yet to be explored. The research aims to explore the role of these recommendation algorithms because of their potential to shape how we see the artifacts that fill our lives and define our culture.

The motivation for this thesis is grounded in research and theoretical exploration of technology as cause and effect in art and culture [39]. In this thesis, we will focus on visual art (paintings, sculpture, graphics, and prints), but will include references to music, film, and other cultural artifacts since many recommender algorithms applied to these domains have similar features and impacts on audiences. Regarding the proverbial "audience," this thesis aims to address the needs of a specific type of audience member drawn from the research on identity and museums by John Falk [108]. In his 2009 book, Falk posits five primary identities that capture museum audiences: explorers, facilitators, experience seekers, professional hobbyists, and rechargers. Notably, Falk's work connects identities to an individual's actions and motivations within a museum rather than grouping visitors by their demographics. Therefore, Falk argues that these identities can also shift for individuals throughout their time in a museum. Such an application of identities for museum analysis has been repeated academically [75] and anecdotally [335]. This thesis will focus on audiences in the explorer identity group, defined as curious visitors looking to discover new things about art. We selected this group as their goal of discovery reflects the role of recommendation algorithms in society. By Falk's definition, explorers can be found in museums and online, but we argue that one experience supports their appreciation of art more than the other. Finally, we will not focus on the creation aspect of Figure 1.1, as this area of research is currently being well explored by some colleagues, see Laura Herman [150, 152, 151] and Renkai Ma [201, 203, 200] for researchers focused on the impact of algorithms on artist and creator processes.

The first motivating point for the research is that research in computation and art has been happening since the 1950s with attempts to identify variables that describe art and artifacts [155, 300, 225, 184]. Furthermore, most research exploring relationships between recommendation algorithms and art has focused primarily on music and audio streaming [63, 252, 55]. Over the last few years, research beyond music has not explored art recommendation algorithms. Instead, the research has been explored using advances in machine learning and neural nets to classify, identify, and appraise art with computer vision techniques and has also led to the current explosion in generative AI works [269]. We argue there must be more research exploring the impact of recommendation systems on human experiences with art and culture. Therefore, this thesis investigates the diverse impact of these systems in the context of art appreciation.

The second motivating point is that the lens of Human-Computer Interaction (HCI) applies well to researching algorithmic curation's impact on art and culture. Art and HCI are by no means strangers to collaboration; years of work by Bardzell and Bardzell, Blythe et al., and Andersen et al. have explored questions related to how digital media have changed the creation and presentation of art [32, 46, 15]. Focusing on user experiences with interactive technology that targets audiences interested in art, HCI offers a variety of methods applicable to this research (see Chapter 3). Today's discourse around digital platforms and the role of data in our lives applies to all digital content experienced by billions, including art [311]. Our research questions how the inherent power of algorithmically-powered platforms has the ability to limit human agency in encountering art in desired ways [24, 65, 130]. This thesis will bridge current discussions on the power of algorithms with historical discussions on the power of art and artifacts. By exploring users' experiences, new interfaces for human agency and cultural values can be proposed that facilitate critical engagement and careful curation. This thesis will involve multi-faceted methods grounded in existing HCI research. By placing end users first we can revisit the historical power structures around both art and algorithms.

1.2 Research Questions

This thesis sets out to explore how digital spaces can be redesigned to support art and culture. These questions represent the iterative HCI process which begins by broadly understanding the current experiences and pain points of users before narrowing down on novel data and prototype structures that explore potential solutions [86]. Within those four over arching questions lie a series of smaller research questions that can be answered through the methods employed in this thesis. The four initiatives are presented as the following questions:

1. **Learning from Experiences:** What do experts and regular users have to say about the impact of recommendation algorithms on their engagement with art and culture?

2. **Learning from Gallery Reactions:** What are the differences between audience observations of a physical gallery with algorithmically curated images and a machine's reading of the same images?

3. **Understanding Data:** What factors, values, or variables are *not* currently at play in algorithmically processing art? How can they be informed by expert art evaluation?

4. **Building Potential Futures:** How can we use the unique capabilities of large language models to digitally support audience engagement with art based on the needs of our past participants?

1.2.1 Learning from Experiences

With over half the world's population using social media [311], it is one of the most dominant ways for people to encounter art digitally. As this thesis explores how digital spaces support art appreciation, it is vital that we understand the impact of the current system. According to other researchers in HCI, one of the biggest challenges is the flattening of all human expression into the singular term of "content" [301]. Therefore, we were motivated to understand the impact of contentification on art according to the lived experiences of experts, artists, content creators, and the public. Experts and artists have made up the fabric of the art world for quite some time; their roles have also changed to account for new technology and ways to engage with art even before the dominance of social media. Content creators have capitalized on the new outlets offered by social media, but even their role as creatives is rapidly changing [92, 83, 297]. Finally we have Falk's explorers [108], audience members interested in seeing art, who have encountered trends and changes to art experiences. All four groups are the users whose appreciation and perception of art and culture are the focal point of this thesis.

Outcomes: We established a network of approximately 20 artists and experts for this topic. The network includes academic curators, private gallery curators, and artists of varying degrees of fame and financial success. We interviewed this network and analyzed their responses using affinity clustering and thematic analysis. Additionally, we supported the findings from this network with almost 40 survey responses from non-experts. The interviews and the surveys asked the participants about their taste in art, society's perception of art, and what they would change to improve the AX for art experiences.

Based on the responses from our participants, we presented four features essential to our participants' appreciation and perception of art and culture. These four were depth, conversation, connection, and time. Depth refers to additional information, metadata, and facts accurate to the artifact. Conversation is having subjective thoughts and opinions from anyone openly discussed aside from the facts of the depth pillar. Connection calls for clarity around sources of information and cultural associations that tie audiences and pieces together. Finally, time refers to sitting with the artifact and processing the information and the content of the previous three pillars. In the discussion, the four attributes are contextualized within the existing literature on art theory and theories of art appreciation [36, 56].

Following the interviews, we ran a series of co-design workshops with a group of art experts (n=13) similar to the initial interview group. Three workshops in total focused on taking the characteristics of depth, conversation, connection, and time and using them to re-imagine social media experiences and digital environments for art. Based on these workshops, the experts developed a series of guidelines and a set of prototype screens. After a cross-examination where each group reviewed the work of the other groups, a final set of materials combining attributes and potential experiences was created.

Finally, to explore the impact and validity of these workshop outcomes, we gathered a group of expert social media users (n=10), commonly called content creators but referred to as creatives, within our research. The creatives met one-on-one with the researchers to review the materials and provide their insight on the potential usage of the co-design outputs within the context of social media. Their feedback highlighted that while the values were admirable and desirable, the current trend for social media to pursue short-form content does not support those values within the algorithm structure. They argued that they have a personal goal to establish an audience that will view long-form content to start implementing the values and experiences the art experts value. Their feedback led to design recommendations and insight into future work on what future experiences might look like.

Contributions: The research conducted in this thesis section offers three sets of contributions. The first is a continuation of the open-source dedication described in the previous section. In this case, all of the co-design materials and outputs are made open-source for other researchers to explore. These tools are meant to offer insight into how a co-design session can be run with experts rather than usual end users. Furthermore, the outputs offer museums, galleries, and other researchers a set of screens and designs to use as starting points for revisiting the design and layout of their own digital experiences.

In addition to materials, this section also offers some methodological contributions. The first is an example of how Slow Technology (ST) research can be done bottom-up [144]. ST is an HCI method that explores alternative experiences that challenge a user's relationship with time and the task of the technological tool [243]. Often, this is done top-down, where a researcher decides how the tool interacts and what aspects of time are challenged. This research shows that users can verbalize how they want their experiences to change with time and efficiency. Finally, this section also offers a few methodological recommendations for other researchers exploring the relationship between art and content. By pulling together the findings, three design recommendations are presented, which offer areas for future research and conversation.

Another contribution is more theoretical, based on the feedback of our participants we presented four characteristics that they deemed essential for an art experience. These findings align with, and extend, the psycho-historical framework of art appreciation. This framework argues that art appreciation is a combination of mentally processing the external stimuli (psychological) alongside a review of our internal knowledge of the piece, style, context (historical). Our participants' feedback appears to support this framework, but they also advocate the importance of connection and conversation as parts of the art experience. These two aspects introduce a social aspect to the framework of art appreciation.

1.2.2 Learning from Gallery Reactions

With the previous projects focused on computational processing of art broadly and within the realm of recommendation algorithms, this project examines what happens when computation enters the physical exhibition space. The *Algorithmic Pedestal* was a gallery, practice-based research project that reported gallerygoers' perceptions of a human's curation and curation done by Instagram's algorithm. This project presents a technical analysis of the same exhibit using computer vision code, offering insights into machines' perception of visual art and how it differs from what audiences remarked. Specifically, we used methods from computational aesthetics literature, an established approach to analyzing art with software. The computer vision software, or machine, focused on the pixel, hand-crafted features extracted from the images uploaded for the algorithm and artist to choose from.

We were motivated to gain insights into the perceived value of the curation process, shedding light on how audiences perceive artworks differently from machines using computer vision and what that means for art curation in digital spaces.

Outcomes: In this research initiative, we demonstrate how humans and machines perceive artistic curation differently, thereby building on established research that outlines the differences in human and machine perception and the risks of assuming that they are either completely the same or wholly different [51, 296].

A notable disparity emerges when comparing the computational analysis results to human observations from the gallery experience. Previous qualitative findings highlighted users' views on the differences between human and machine outputs [122, 180, 265]. The human work was anecdotally more abstract, holistic, contextual, and emotionally resonant, whereas the algorithmic work was more object-oriented, recognizable, and individualistic. However, the computational analysis did not discern these same differences, prompting questions about the limitations of computational perception. As we will outline, machine perception focuses on objective data measures when processing an image. Therefore, even as society applies machines in artistic contexts such as Instagram or museum curation, computational methods cannot measure contextual meaning and emotionality, which human audiences prioritize.

Alternatively, it is plausible that biases influence human perceptions, such as those introduced by floor labels, leading to exaggerated differences between the two sides. Rae recently highlighted how labeling work as human or algorithm-made might lead to negative perceptions of the work by audiences [264]. Similar claims come from other research stating that humans consider context and background knowledge when forming their ideas of the world [205]. This form of processing injects ideas and presumptions that alter the perception of the experience, potentially leading to the participants reporting the differences they identified. We argue that the true nature of perception lies in the conflict between human interpretation and machines' objective measurements.

This work contributes to a body of literature highlighting how humans employ both top-down and bottom-up processes to interpret information. In contrast, the computational aesthetic metrics with which machines are programmed form inherent limitations. These findings motivate the next research project investigating new metrics and processing approaches for computationally processing art.

Contributions: With this project we again contribute open-access to our code and data as practical takeaways for other researchers to interact with. We also explore a new way of comparing computational aesthetics to human appreciation of art. Rather than repeating methods that have humans and algorithms competing for accuracy or identification, we instead report on the specific differences each group noted. This method of comparing human and machine perception is an approach that has been called for by previous researchers [143, 341, 375]. Finally, we argue for an extension to curatorial theory injecting a concept from responsible AI, the process of a human-in-the-loop. Rather than just on the loop, our work argues for the involvement of human perspectives and preferences in the selection and processing of the data. These serve as arguments motivating the future systems that are built to present art to audiences. There are certain aspects of the audiences' experiences that are particularly valuable that go beyond the quantifiable measurements of software that should be involved in the curatorial process rather than just checking the output.

1.2.3 Understanding Data

Algorithmic processing of art tends to fall into to major camps, general recommendation systems and specific art analysis tools. The results of the first two projects present our exploration of the pain points and limitations of social media recommender systems and specific computational aesthetic tools. Recommendation algorithms used on social media and streaming services perform algorithmic curation to promote and optimize certain factors. These factors are often the business values of the parent corporation [291, 130, 54, 71]. Most of these algorithms are engagement-based, providing users with content similar to the content they had previously engaged with. This approach differs from the content-focused art analysis done by researchers in the field of computational aesthetics [209, 210, 156, 47]. This form of algorithmic processing takes aesthetic principles and applies them to an art piece through computer vision software. There exists a gap in the research because neither of these approaches handles art recommendation like collectors, galleries, and museums [282]. By operationalizing how art experts evaluate art metadata with high quality data and practical model applications, we can highlight new approaches to algorithmic art analysis.

Outcomes Here we present *AppraiSet* a large dataset of art metadata built on the expectations and specifications of art appraisal experts. Unlike other art data that exists, this dataset is made open-source and includes 22 varaibles for each art piece. In particular, this dataset contains the published condition reports, provenance, and appearance history of the art pieces. We gathered 30,000 visual artworks from major auction house websites. Auction house websites proved one of the most reliable sources of well-formatted art data. The data was labeled with 22 variables of interest that can be used to explore and classify the artworks.

After data cleaning and manual verification, a selection of over 10,000 lots were selected for a case study exploring how an algorithmic model could start to develop its own understanding of art language. Using an Latent Dirichlet Allocation (LDA) topic model, the lots were processed and categorized into relevant topics. To ensure the usefulness and validity of this use case, we present a benchmark to which to compare. The research of Abera Yilma et al. offered a benchmark of a much smaller dataset containing basic artwork details [5]. Not only was our analysis of the larger dataset able to provide a higher coherence score across various topic distributions, but the data also provided multiple distinct topics differentiated by the type of art within the auction lots.

As we review the results of this case study and the release of *AppraiSet* we discuss how algorithmic analysis of art metadata can shape the future of computational aesthetics and how it may alter the knowledge audiences will need to have when appreciating and interpreting art in digital spaces.

Contributions: This project aims to provide practical, methodological, and theoretical contributions to the research community. In the case of *AppraiSet* the open-source release of the dataset and code serve as practical examples of what an art metadata dataset may look like and how it might be used. Similarly, as we point out in Chapter 6, this is also a methodological contribution, of the existing art datasets, all of them provide researchers and users with limited access and require payments. Implementing open source art metadata that is focused on the qualitative and quantitative aspects of art appraisal data is a novel approach to art analysis. Finally, we argue this project provides theoretical contributions to the discussion and implementation of computational aesthetics. Computational aesthetics, a field that has historically focused on the implementation of computer vision to analyze art on pre-establish aesthetic criteria has the opportunity to be expanded in the age of AI and NLP. We argue that our LDA topic model provides a new aspect to computational aesthetics by allowing an algorithm to develop its own classification and description of various types of art, essentially taking the first steps to creating its own aesthetic language.

1.2.4 Building Potential Futures

For the final project of this thesis, we return to a digital environment armed with the knowledgeof the previous sections. Chapters 4 and 5 provide design recommendations for a new system of art experiences, while Chapter 6 provides high quality data capable of informing an algorithmic system. By combining these outcomes with advances in large language models (LLMs), we can create an experimental prototype that explores the future of algorithmic curation. In this project we introduce ArtBot, a Socratic large language model (LLM) art companion designed to guide users through artwork analysis. Our project combines the methods of provocative prototyping [49, 327] and technology probes [164] to create and test

ArtBot as a potential algorithmic experience. The prototype itself is meant to elicit user responses, and it is tested through a within-subjects experiment comparing the prototype to digital museum collections and social media experiences. The findings reveal how an AI can support audiences' critical engagement with art and support their overall critical thinking processes.

Outcomes: ArtBot is an art companion built on a local Llama 3.1 LLM. Having the LLM be local allowed the researchers greater control and customizeability over the model. One form of control was in establishing a system prompt that dictates the functionality of ArtBot to act as an art tutor employing a Socratic approach to the conversation. Another form of control was tying the AI to the artworks we wanted to display to users. In this case a local model allowed us to integrate directly into a custom built web app which combined the visual display of an artwork and caption alongside the interaction area. Finally, we also had to integrate AppraiSet from Chapter 6 to make ArtBot a retrieval augmented generation (RAG) model. This way when a user is discussing artwork with the AI, the dataset of artworks is available to help ArtBot form the appropriate descriptions and details relevant to the discussion.

In addition to developing the prototype, we also examined its performance as a technology probe through a within-subjects experiment. The experiment was conducted on 13 participants which experienced nine different artworks randomized across three conditions. The conditions were meant to compare different common art experiences. The control condition was inspired by a museum where a piece of art was accompanied by a label and traditional curator developed wall text. The first experimental condition was inspired by social media where the artwork was central to the interface with a small caption. Finally, ArtBot was the third condition where the artwork and caption were accompanied by the AI. After each condition our participants were asked to rate their thoughts on the artwork and write a short text response about their selection of various Likert scales. Both the scales and the text response were analyzed. We found statistically significant differences in self-reported understanding, number of complex words, and writing grade level where the measurements were dependent on the condition. In post-hoc tests we saw that ArtBot performed at the level of the museum condition, with both of them outperforming the social media condition.

Finally, at the end of the experiment we also asked participants a set of qualitative questions to evaluate their perspective on the algorithmic experience presented. Their comments and feedback was recorded and analyzed to understand what aspects about the Socratic approach worked well, and some of what did not. Additionally, we explored what participants thought about the behavior of ArtBot and how it matched their expectations in comparison to the othe conditions. We combined the quantitative and qualitative data to discuss the potential contribution of the work beyond the code developed.

Contributions: The contributions of this final project also aim to reflect the work established throughout this thesis. Again, as a practical contributions we make all of the code and resources open-source for other researchers to explore and expand. For methodological contributions we extend the work being done in the LLM space by arguing for the implementation of a Socratic system rather than a subservient one. In other words, our LLM asks questions to prompt the user to think and respond rather than being prompted for various outputs. Finally, we argue that this project contributes a practical example of a various theoretical contributions we have explored in the thesis. It provides a digital version of the expanded socio-psycho-historical art appreciation framework by involving conversation and connection. Furthermore it immediately places the human in the loop of art curation by urging the user to come up with their own thoughts and ideas about the piece.

1.3 Thesis Chapter Overview

This thesis encapsulates the combined work of the past few years in pursuit of a greater understanding of how our digital experiences impact our relationship with art. The work presented reveals all of the research and theory that has been accumulated to answer the research questions and present a coherent argument towards the thesis question. Refined selections of this work have been previously presented and published to conference venues and journals and will be clearly highlighted and expanded upon during this thesis. As this chapter introduces the entire document and research, the following chapters present the entirety of the work and how it relates to research that has come before.

Chapter 2 provides an overview of relevant research that positions this thesis within the larger art and AI research canon. The chapter begins by outlining the background, terminology, and scope of this thesis. Then the chapter presents related work in the spaces of algorithmic curation, art, and social media, offering readers an understanding about what has been done by other scholars. This background section aims to motivate and contextualize each of the thesis chapters. The chapter also introduces a theoretical framework of art appreciation. Specifically, the psycho-historical framework for art appreciation. This framework was selected as it was deemed to fit well with the human-centered computing approach applied throughout this research. Both the framework and the approach combine human cognition and processing with information processing and data. This chapter will introduce the framework, motivate its role in the research, and be used as a reference later when discussing the learnings and outputs from the subsequent chapters.

Chapter 3 presents the overall research design for all of the research projects presented throughout the thesis. This chapter connects each method to the relevant methodological literature that inspired it and explains how the quantitative and qualitative data outputs were operationalized from one project step to another. Methods referenced here will be referred to throughout the other chapters and expanded upon to highlight the results that inform the discussions and contributions.

Chapter 4 is the first of the four research initiatives. This chapter outlines the experiences and data gathered from individuals and experts related to the experience of visual art shared on social media. In particular, it argues that social media's infrastructure and algorithmic experience directly and potentially dangerously impact how users experience art by treating everything as content. To overcome this challenge of all things becoming content, the research explores various design approaches to take the participants' feedback and form potential recommendations. The chapter concludes with a discussion reflecting on the usage of slow technology approaches to provide users with their desired experiences and aims to extend the psycho-historical framework for art appreciation in relation to our participants' responses.

Chapter 5 presents a second project that brings computational analysis and metrics into a physical gallery space. With Chapter 4 highlighting the challenges of experiencing art on social media, this project explores the impact of social media's artistic curation within the scope of a physical gallery space. The chapter outlines the methods used to perform a technical analysis of the *Algorithmic Pedestal*. We analyze images from an open-source collection with computer vision software informed by computational aesthetics that extracts hand-crafted features from the image files for statistical comparison. This technical analysis of a physical exhibition allows us to compare the results of a machine to the observations of human gallery attendees previously reported in other research. It highlights how a computational lens to art often misaligned with visitor expectations and interests informing our last project.

Chapter 6 is the third of the four research initiatives. This chapter outlines the research that went into exploring the data and systems missing in current approaches to algorithmic curation. It explores the motivation, background, methods, and results of algorithmically developing a dataset and exploring art data. It presents answers and concepts to reflect on how algorithmic curation and data-focused approaches to art can shape the information displayed to users and impact their experiences. The chapter includes a case study with the Latent Dirichlet Allocation topic model on how art data can be used for an algorithm to develop its own language and classification of art. Ultimately, this section concludes with a discussion on the future of contemporary aesthetics, the usage of AI analysis of art, and how it can inform audience's interpretations of art.

Chapter 7 presents the final of the research pillars. Combining the design recommendations drawn from user feedback in Chapter 4, the comparisons between computational and human perception of a physical gallery in Chapter 5, and the deeper understanding of art data in Chapter 6, this chapter presents ArtBot. ArtBot is a LLM art companion that serves as a technology probe testing whether there is an algorithmic experience of art that can improve audiences' appreciation of art. We explore users' expectations and behaviors around artworks presented digital. Finally, we conclude with a discussion how this type of technology can be expanded to help users critically think about other forms of art and human expression. Similarly, it raises questions about the future of human-AI collaboration and usage.

Chapter 8 presents a discussion and reflection on the work completed throughout the thesis. It reviews the results and impact of the previous chapters and compares them to the existing canon. This chapter expands on how the work presented contributes to users in the arts, researchers, and end users online. Finally, it concludes with a discussion on the future of human-AI collaboration.

Chapter 9 concludes the thesis with final remarks on learning from the work. It also outlines how the work contained within the thesis is one voice in addition to larger ongoing discussions on the role of algorithms in creative spaces and how future research can use the work and contributions.

1.4 Definitions and Abbreviations

Human-Computer Interaction (HCI) - A field within computer science that often combines methods and research from computer science, psychology, and design. Researching and developing technology focuses on the humans central to any type of socio-technical system.

Art - Many things fall under this term, and it is not the place of this paper to debate what is or is not considered art. Instead, the shortened word will be used throughout this thesis to refer primarily to visual art (painting, sculpture, pottery, digital art, drawing, sketches, etc.), with a short section about music (clearly defined). But overall, art within this thesis will not refer to the performing arts or feature-length cinematography, for example.

Social Media - For the purposes of this thesis, social media predominantly refers to sites such as TikTok, Instagram, YouTube, Facebook, and Twitter (X).

In particular, TikTok, Instagram, and YouTube will be featured quite heavily when referring to social media.

Platforms - These are online systems that host users and content. Often, the order in which content is displayed to users is dictated by a recommendation algorithm that promotes content and advertisements according to some unknown, often proprietary optimizations.

Algorithmic Experience (AX) - A recent area of research within HCI that argues the importance of considering the entire experience of interacting with an algorithm. The idea is that user rarely face the algorithm directly in their interactions but rather interact within the context of the broader platform and their digital experiences. Therefore, it is essential to consider the entire experience around the algorithm when attempting to improve the situation for the users.

Computational aesthetics - A specific application of computer vision software to analyze artworks. The computer vision software is informed by established aesthetic theories for art analysis allowing researchers to attach metrics to images by calculating contrast between pixels, among other metrics.

Human-Centered Design - Another subset of research and methods within HCI that focuses on designing and building user tools. Again, the key to this area of research is to place the humans using the technology at the center of the design process by understanding who they are, what they need, what they want, and most importantly, how it might impact them.

Participatory Design (PD) - A method within HCI that encourages researchers to have the users and stakeholders they are building technology for involvement in the research process. Allowing the stakeholders to participate is one way to ensure their position is represented directly within the outputs.

Provenance - A term used in art appraisal referring to the factual history of a piece of art. Tracing the notable origin and transfer points (sales, inheritance, or even theft) that brought the art from the artist to the current place it resides.

Condition - Another term used in art appraisal refers to the physical condition of the piece. Rather than focusing on descriptions of the art, the condition itself refers to any wear that might be present on the art due to the passage of time or other instances.

Creatives - A term used by some HCI literature to refer to content creators. This term is often preferred as it separates the individuals working to make and share creative work for the Internet from the more loaded and potentially negative term of “content.”

1.5 Publications

Open Source

1. von Davier, Thomas Serban; Van Kleek, Max; Shadbolt, Nigel. 2022. "AppraiSet", Mendeley Data, V1, doi: 10.17632/2nfvz8g27c.1

2. von Davier, Thomas Serban; Noh, Hayoun; Abbaspour, Mitra; Breault, Doug; Cosford, Nicola; Dionne, Caroline ; Just, Bryan; Kusserow, Karl; Kwok, Zoe; Leoni, Francesca; Metal, Tara; Shapland, Andrew; Smith, Sydney; Steward, Jeff; Winterbottom, Matthew; Van Kleek, Max; Shadbolt, Nigel (2023), “Co-Designing to Separate Art from Content with input from Art Experts”, Mendeley Data, V1, doi: 10.17632/y7p73zxkkf.1

3. von Davier, Thomas Serban. 2024. "artConnct", GitHub, https://github.com/tvondavi/artConnct

Conference Papers

1. Kollnig, K., Datta, S., Serban Von Davier, T., Van Kleek, M., Binns, R., Lyngs, U., & Shadbolt, N. (2023, June). ‘We are adults and deserve control of our phones’: Examining the risks and opportunities of a right to repair for mobile apps. *In Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency* (pp. 22-34).

2. Lee, H. P., Yang, Y. J., von Davier, T. S., Forlizzi, J., & Das, S. 2024. Deepfakes, Phrenology, Surveillance, and More! A Taxonomy of AI Privacy

Risks. In Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (CHI '24). Association for Computing Machinery, New York, NY, USA, Article 775, 1–19. https://doi.org/10.1145/3613904.3642116

3. von Davier, T. Ş., Caporossi, G., & Correa-Dantas, D. (2024). Using Networks to Explore the Invisible Connections in the Music Industry. *Proceedings of the International Conference on Arts and Cultural Management (AIMAC).* Portugal 2024.

Journal Papers

1. von Davier, T. Ş., Kollnig, K., Binns, R., Van Kleek, M., & Shadbolt, N. (2023). We Are Not There Yet: The Implications of Insufficient Knowledge Management for Organisational Compliance. arXiv preprint arXiv:2305.04061.

2. von Davier TŞ, Herman LM, Moruzzi C. A Machine Walks into an Exhibit: A Technical Analysis of Art Curation. *Arts.* 2024; 13(5):138. https://doi.org/10.3390/arts13050138

3. von Davier, T. Ş., Van Kleek, M., & Shadbolt, N. AppraiSet: Discussions on a New Art Dataset. *Contemporary Aesthetics.* Special Volume (2025).

4. von Davier, T. Ş., Noh, H., Van Kleek, M., & Shadbolt, N. (2025). Looking for Art in a Sea of Content: A Human-Centered Approach to Supporting Creativity on Social Media. *The Proceedings of the ACM on Human Computer Interaction (PACM HCI'25).* 9, 2, Article CSCW127 (April 2025), 25 pages. https://doi.org/10.1145/3711025

Extended Abstracts

1. Thomas Şerban von Davier. 2023. Designing for Appreciation: How Digital Spaces Can Support Art and Culture. *In Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems (CHI EA '23).* Association for Computing Machinery, New York, NY, USA, Article 490, 1–5. https://doi.org/10.1145/3544549.3577041

2. von Davier T.Ş., Larsen A., Van Kleek, M., & Shadbolt, N. (2025). ArtBot: An Exploration into AI's Potential for Guiding Art Analysis *In Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '25).* Association for Computing Machinery, New York, NY, USA. 11 pages. https://doi.org/10.1145/3706599.3720181

2 Background

Contents

2.1 Introduction

Before presenting the individual studies that form this thesis, it is essential to establish the terminology, scope, and greater canon of previous research. The introduction presented a diagram depicting how modern day art experiences involve multiple conduits that transfer information from artists to their audiences. Within this diagram there are terms such as curation, appreciation, and algorithmic experiences (AX). As these are general terms for abstract concepts, and the potential subject matter for multiple theses, we will define them within the scope of this work in the first part of this chapter. In the area of curation we will discuss the analogue origins, how it has evolved with the field of digital heritage, and how modern digital systems act as curators. For appreciation we will explore how audiences with different specialties may approach art appreciation while also highlighting the specific theoretical framework of art appreciation this thesis will focus on. Finally, for AX we will differentiate the common types of algorithms this thesis may refer to, recommender systems, generative models, and a sub-type called critical thinking support AI. All three AX systems are relevant, but speak to different parts of the digital art experiences. With the terminology outlined we can then explore prominent research in related areas of study.

There are two research areas that offer relevant background for this thesis. One is the area of using algorithms in the creation of art. This field of research is incredibly influential in current narratives on the deployment and implementation of generative models and digital tools in artistic creation. The papers cited cover the portion of the mental model diagram that leads up to the curation done by our digital systems. The second area of research is the study on the impact of social media on its users. Work within this body of research explore algorithmic and Internet safety, responsible data practices, and general social media behavior studies beyond just a focus on art. This chapter will present a sample of notable works within these areas and highlight how both fields reveal the gap within HCI literature that this thesis fills.

2.2 Terminology and Scope

2.2.1 Curation

The practice of organizing and presenting artifacts is broadly described as curation. Often tied to the job of a curator within an organization that maintains a collection of items. The task has a history dating back to early galleries and museums [28, 29] and has grown and evolved with changes to society's relationship with institutions and with the deployment of online spaces. Throughout this thesis we will refer to curators from museums and art galleries and digital curation done in virtual spaces by people and algorithms. Therefore, we provide a set of literature on curatorial methods, cultural heritage, and the changes in social theories of curation as we inhabit a world with increasingly influential algorithmic systems. The literature should contextualize the term within the scope of this thesis and how the words curation, curators, and curatorial practice will be used in the research projects.

2.2.1.1 Curatorial Methods

Curation's etymology can be traced back to the Romans and the Latin word *cura* meaning "care". Over time, different people have been labeled as caretakers for states, parishes, societies, and, more contemporarily, art and heritage. Their methods for caretaking also evolved over time, but can be summarized by Balzer as, "the curator is someone who insists on value, and who makes it, whether or not it actually exists" [29]. In the context of art and heritage the curatorial methods highlight the value of artifacts that represent key aspects of human expression. With the arrival of the platformed Internet, scholars argue who gets to label themselves as curators [29, 223]. With platforms allowing anyone or anything to curate music, posts, images, videos, or statements, these scholars argue the act of curation is no longer focused on caring for the objects or items but is rather focused on creating a brand or image.

This thesis adheres to the concept of technology altering the curation process. It specifically examines digital heritage and how museums or cultural institutions are altered to fit the digital world. By understanding how the historical curatorial practices of documentation, presentation, storytelling, and selection are adjusted in the reality of social media and other algorithms, we can understand how an audience may react to the art or artifact being presented.

2.2.1.2 Digital Heritage

As formal institutions adjust to reflect a digital world, heritage organizations grapple with data management, digitization, access, trust, and instantaneous interactions and delivery [255]. Digital heritage is the formal practice of recording and storing heritage information and artifacts with digitization for posterity and greater public access [234, 326]. A large body of digital heritage literature explores the creation of various digital art experiences including, but not limited to, XR, interactive galleries, digital collection websites, tour materials, and digital libraries [234, 351, 221]. Digital art experiences can also include areas outside the control of heritage organizations and experts, such as social media platforms. Researchers in the field have explored how digital heritage can be presented on social media and the challenges involved [325].

As museums attempt to share their expertise with global audiences via social media, they strive to strike the appropriate tone [163] balancing openness with legitimacy. Alongside museums, ethnic groups also take to digital spaces to share aspects of their art and cultural heritage. Scholars explore how this allows for the transfer of “intangible cultural heritage” facilitated by digital spaces [195, 69]. In exchange for reach and exposure, cultural messages trade the risk of being buried amongst all the other information on display in digital systems. Loke et al. demonstrate that even outsiders visiting a heritage site can represent it digitally through their tourism photos and social media presence [193].

Among the literature are some researchers identifying gaps within digital heritage research. First is the need for collaborative, user-centered development of digital art experiences [326, 267]. In 2010, Thwaites questioned the impact on audiences and human values as more and more heritage is digitized [326]. Their writing calls for greater audience involvement in cultural heritage delivery. Rahaman and Tan further call for "dialogic intervention" between users and experts [267]. They reiterate the need for human-centered research in their 2018 paper, proposing a conceptual framework of what the user experience should involve, such as satisfying, provoking, teaching, and sharing multiple perspectives [266]. Simultaneously, researchers record challenges with audiences’ interactions with digital museum collections, where users seem to drop off and not return even as the systems are updated, and new features are added [142, 213]. The following thesis combines user and expert feedback in designing new digital art experiences (see Chapter 4). Furthermore, in Chapter 7, we also take the idea of "dialogic intervention" [266] and apply it to an interaction design involving end users and an AI with promising findings regarding users’ interest in the experience.

The second gap brought up by digital heritage scholars has been the growing awareness that digital heritage is fundamentally a dataset [354]. Whitelaw outlines how heritage digitization eventually abstracts itself towards metadata about heritage artifacts and locations [354]. They support their arguments with Manovich's writing about the power of whoever controls the dataset to present the information in any way that achieves their agenda [207, 354]. Within this thesis, Chapters 5 and 6 both explore the power of proper dataset analysis and interpretation for digital heritage information.

As curation moves to the digital, platforms are employing the curation practices of documentation, presentation, storytelling, and selection faster than people. The act of curation has become part of the digital experience through the personalization algorithms powering online spaces. This challenges the social theories of curation, which have previously been a human-led process.

2.2.1.3 Digital Systems as Curators

Bourdieu's concept of art experts and institutions as cultural intermediaries has dominated the discussion of curatorial practice [52]. The theory explores how a historically elite few can set the language and focus of cultural relevancy. Hogan extends the theory into social media but argues that digital spaces equip anyone to be a curator of their digital artifacts and exhibitions [157]. In another piece of work, Hogan expands the argument, highlighting how algorithms empower the users to fulfill three roles of the curator: "filtering, ordering, and searching" [158]. Morris then interprets algorithms as cultural "infomediaries" due to their ability to perform the functions outlined by Bourdieu and Hogan without the input of the layperson. In this interpretation, algorithms, and software make the curatorial decisions at a global scale for all types of information, not just art and culture [228]. Morris explains that while algorithms have this ability to shape taste and segment audiences, they are not without human involvement; in fact, humans dictate the design, development, and deployment of the algorithms, therefore resulting in algorithms that are prone to replicate our biases in pursuit of our goals [228].

The role of algorithms as infomediaries has scholars split on the potential effects. Some argue it has occasional opportunities to empower diverse exposure to cultural artifacts [287]. While others argue these systems are gatekeepers [220] limiting access to users [337]. Regardless of the level of exposure these algorithms offer through their curation; the decision-making has a direct monetary benefit for the companies owning the platforms and a direct experiential impact on the users of the platforms [258, 248].

Based on the existing literature, the word curation in the following thesis refers primarily to this shift and consolidation of curatorial power within algorithmic decision-making and its potential impact on the audience's appreciation of art.

2.2.2 Appreciation

Appreciation is the second term that will appear throughout this thesis. Unlike curation, appreciation is grounded more in theory than methodology, with different audiences subscribing to different theories of appreciation. For background, we provide common, widely accepted theories tied to art history and art critiques. Then, to scope the term within the context of this thesis, we highlight a theory of art appreciation that applies to audiences broadly and will be reflected upon in light of the research results.

2.2.2.1 Aesthetic Theory

Aesthetic theory aims to attribute value to artifacts. Within this section, we highlight different perspectives aestheticians debate and explore with their philosophies. Again, here is an attempt to provide a high-level view of various aesthetic theories attributing value to the art. The aesthetic arguments are grouped into three major camps for ease of discussion: Personal values, pure aesthetics, and ethical values. Starting with personal values, Alain de Botton and John Armstrong may argue that the value observers place in the piece comes from a need for a therapeutic or cathartic release from observing and experiencing the work [93]. In this case, the value may fluctuate depending on the audience members' traumas or mental state. The core argument is that personal factors may be specific to the audience instead of the artist or the art piece.

The argument of pure aesthetics focuses on what aspect of the art provides the overall value of the piece. Leonard B. Meyer might argue that the value comes from the piece's originality [222]. Variables, such as the piece's creation date or edition number, could support this argument by showing how original or rare a piece is. In opposition, Jack W. Meiland would point out that the piece itself is not truly an original considering its connection to previous items within an artist's canon, and that there should not be a difference in value based on the originality or lack thereof of the piece [182]. Hume may argue that appraisal approaches the topic of value from a financial side driven by an understanding that comes from a set number of comparable classifiers [275]. However, a collector or an observer of the work might consider the art valuable based on their sentiments, putting the two sides in a conflict that Hume often explores. Finally, Danto might argue that the aesthetic value of a particular artwork comes from the meaning or message embodied within the work [90].

Finally, another item to consider would be the artist's behavior or the artifact's history. As ethical questions come to light about art, one must consider how they impact people's views of art. Does knowing something was created from a living organism challenge our morals [317]? What if the piece glorifies colonialism, fascism, or other inhumane or immoral subjects? If we learn a piece was attributed to a man for years but was the product of a female artist, does this knowledge change the value of the art? Unlike the previous sections, ethics can alter our experience of other value types. Berys Gaut explores the impact of ethical flaws and how those can create potential aesthetical flaws within a work of art [126]. They ultimately conclude that the presence of moral flaws, from an ethicist's perspective, is inherently ugly, decreasing the value of a piece. Gaut admits this debate is still ongoing, as there are multiple potential relationships between ethics and art, even within the philosophy of ethicism. An appraiser or dedicated art historian may attribute these ethical flaws as details of the overall provenance. Nonetheless, a challenge remains, an artwork may include the provenance with all the problematic details, but that may not sufficiently equate to the ethics or feelings experienced by the observer.

There are many theories and ideas about art and culture. However, many have been in existence since the Greek philosophers, but as technology and art continue to integrate, these theories are likely to require some adjustment. Art historians recommended Walter Benjamin's writing and seminal work "The Work of Art in the Age of Mechanical Reproduction" [38]. For context, this work was written in response to the rise of photography, film, and early advances in mass-produced art. In many ways, it explores early views on the relationship between advancing technology and human creative expression. Benjamin's thoughts are particularly salient when considering online platform sharing and replication capabilities.

One of the paper's central arguments is that mechanically reproduced work or depicting another piece of art through reproduction fundamentally differs from the original. The space and time aspect of the artwork is often discussed in association with the term "aura," which reproductions are arguably lacking [38]. The argument is rooted in a belief that being within the physical presence of artwork at a particular time is integral to the very core of the artwork [38].

The second pillar of the paper is the idea that as art moves from traditional mediums into new technologies and the realm of mechanical production, there is a shift in the purpose of the art. Benjamin argues that mechanically reproduced art is made for politics and political messages, while traditional forms were made for the sake of societal rituals [38]. A noteworthy addition to the shift in purpose for art was the need for additional context, descriptions, or explanations to accompany the art [38].

Finally, Benjamin introduces commentary from other authors and aesthetic critics, namely Georges Duhamel, a French author known for traveling to the United States in the early 1900s and responding to the technological advances he experienced [38]. Duhamel argues that mechanical art productions are notable, especially for their speed [101]. Duhamel and Benjamin describe the impact of the speed through the observer's experience that finally grasps the aspect of the art, only for the artwork to rapidly change and be replaced by the next piece or scene. The parallel between this description of early advances in media technology and the current functionality of online platforms is striking.

Duhamel further critiques technology usage in the arts and culture by arguing that with mechanically produced art, the masses no longer need to concentrate or reflect and instead are distracted by art [101]. A body of work has already outlined how many comments Duhamel and Benjamin deliver regarding the modernization of art practices are returning in modern streaming services, especially music [154]. This literature highlights critics of modern streaming, making the same arguments that have existed since the 1930s. The critics describe the impact of modern technology on the public and the quality of the art. The paper argues that the critiques of music streaming services cause dull, mind-numbing, distracting, crowd-placating music is nothing new [154]. It argues that these criticisms are potentially elitist and ignorant of the actual views and behaviors of the greater public. Hesmondhalgh argues for "serious attention to the complexity of people's musical lives" as one approach to understanding the impact of platforms on art and culture [154]. This thesis aims to contribute insights into how people use platforms to engage with arts and culture and how this new means of discovering and engaging with arts affects their experiences.

2.2.2.2 Psycho-Historical Theory of Art Appreciation

Art appreciation rests at the center of this thesis. As the research explores how advanced algorithms impact the art appreciation experiences of end users, it is essential to ground the concept of art appreciation within an established theoretical framework. In particular, the psycho-historical theory of art appreciation by Bullot & Reber [56] offers a unique framework explaining the various factors that have previously been considered when attempting to explore art appreciation and how a combination of those factors more accurately explains the experience and audience has when appreciating a work of art. Specifically, this framework combines psychological explanations of art appreciation from the fields of empirical aesthetics and neuroaesthetics with the historical explanations of art appreciation from the fields of philosophy and art history.

The psycho-historical theory argues that art appreciation happens through a three step process which combine the psychological processes of the audience member and the historical contexts surrounding the individual artwork being experienced. The three steps are 1) basic exposure to the art, 2) establishing an artistic design stance, and 3) achieving artistic understanding [56]. Each step requires the audience to have both psychological processing of stimuli, but also a certain degree of awareness and knowledge of the historical, factual context of the artwork. The authors argue that their theory better captures the process of art appreciation than the psychological and historical theories alone. They state that artworks are more than just a collection of stimuli for an audience member to react to, and that there is context around each piece defining it and how it came to be in front of the audience. Similarly, the context and factual information alone is not enough to explain the reactions and processes an audience member experiences when experiencing the piece. As a result, the combined theory puts forth that art appreciation is a combination of the previously separated theories.

Bullot & Reber's theory of art appreciation applies well to this thesis as it also explores a combination of factors involved in an audience's ability to appreciate art. This thesis agrees with the argument that artworks are carriers of information, subsequently the entire first section of the work explores the art metadata associated with artworks and how that can be meaningfully be conveyed to audiences. The theory also argues that exploring appreciation is more than treating artwork as a contained set of stimuli, which motivates the second portion of this thesis exploring how art experts and social media users consider art delivered digitally through algorithms that offer an endless stream of stimuli. Finally, the ultimate section of the thesis looks to combine the previous two bodies of work to offer alternative experiences of art appreciation and will use the work of Bullot and Reber to reflect on the theoretical contributions of the thesis.

2.2.3 Algorithmic Experience

The final term we will focus on is "algorithmic experience" (AX) from Alvarado and Waern's 2018 paper. AX refers to the entire user experience surrounding and involved in a user's interaction with an algorithm. Within the context of a thesis dealing with the relationship between art and algorithms, there are two broad experiences worth defining. First, is the more established set of recommendation AXs that have dominated the platforms on the Internet and focus on presenting outputs to users. Second, is the current wave of generative artificial intelligence AXs often used to create new outputs.

2.2.3.1 Recommender Systems

The other type of algorithm we discuss in this thesis are recommender systems. Like genAI, they rely on a high quality, consistent data stream to inform the algorithmic output. However, rather than creating an output, recommender systems present previously created outputs in a certain order. These systems have been driving the platform power of the Internet [57, 237] by studying user behavior data to personalize advertisements and content for significant financial gain. Simultaneously, these systems have been described as "black boxes" of proprietary code raising privacy and security concerns. In addition to having a body of research exploring the role and impact of algorithms on audiences (see Section 2.3.2.2), there have been some legal rulings to protect users from the potential harms of the data being collected, such as GDPR and CCPA [107, 25]. This thesis questions the impact of these systems on art and will compare digital recommender systems to analogue systems built on existing art institutions to see what new improvements can be made.

2.2.3.2 Generative AI

The current computing landscape is marked by the release and pursuing hype of generative AI (genAI) algorithms. These powerful algorithms use large amounts of data to train neural networks that can produce new outputs. There are two groups of genAI systems. The first, often more associated with art, are text-to-image generators that associate words with tagged image descriptions to output visual styles that match the prompt. Some examples are DALL-E, Midjourney, and Adobe's Generative Fill [269, 224, 6]. The second is the Large Language Models (LLMs), which produce text outputs by predicting which words will come next in a sentence to form complete responses to prompts [250]. Some examples of LLMs are ChatGPT, Llama, and Mistral [251, 328, 167]. In Chapter 7, we capitalize on the affordances of LLMs for our prototyping. In this way, our work parallels other researchers working to combine LLMs with a corpus of museum data [334, 329]. Researchers are flocking to genAI's AX because of their ability to complete tasks on our behalf rapidly. However, some scholars argue that this relationship leads us to offload our cognitive processes onto genAI. This concern has sparked a recent outpouring of research exploring the potential of critical thinking support genAI tools.

2.2.3.3 AI for Critical Thinking

While AI technologies often simplify tasks by reducing cognitive load [58, 286], scholars argue they also have the potential to enhance critical thinking skills, benefiting both individuals and society. Cai et al. set forth the framework of "antagonistic" AI as a direct challenge to the current trend of AI technology that insists on being subservient and allowing users to push off thinking to the computer [58]. Their paper outlines that "antagonistic" is notably different from unethical or harmful AI; their distinction is that a degree of challenge and friction in the experience may produce better outcomes for both people and the interactive systems. Sarkar et al. start their paper by describing the tendency for AI systems to act as "autopilots" with the users simply not performing much, if any, cognitive process [286]. Their paper includes a prototype that "provokes" the user to think about the task, as the Cai et al. paper recommends. Ultimately, Sarkar et al. argue that there are short-term and long-term goals for critical thinking support AIs. In the short term, an AI is successful if the user thinks more deeply and critically about the task at hand, while the long-term success is defined by the ability for the skillset of critical thinking to be extended beyond the immediate domain [286].

Simultaneously, Ye et al. turned towards philosophers to understand what AI is currently missing to promote and support critical thinking [361]. Based on their exploratory user study, the researchers developed the selfhood-initiative model of AI. They outline that critical thinking support AIs can come in different forms. The high-selfhood, low-initiative "respondent" is a tool that allows a user to test out their ideas, but the system will not have its own goals for the conversation. The low-selfhood, high-initiative "monitor" aims to help users understand where their ideas fall within the larger landscape; these systems make recommendations or statements that have various degrees of relevance to the user. Sarkar et al.'s shortlisting tool [286] and Park and Kulkarni's thinking assistant [253] are likely examples of the "monitor" tool. The final combination is the high-selfhood, high initiative "interlocutor," which remains under-explored as it requires the system to have its context for the conversation as well as the intention to provide the user with alternate positions and ideas that expand the discussion [361].

In Chapter 7, we set out to present and test ArtBot as an application of these theoretical frameworks of critical thinking AI support tools. We apply critical thinking AI theories to a specific application area, art analysis, to integrate the various recommendations from other researchers. We build on the findings of Danry et al. by integrating question-asking into the explanations and discussions ArtBot has with users about an artwork [89]. Basing ArtBot in the pedagogical literature of Dialogics [303], the AI acts as a Socratic agent where discussions get deeper based on the questions it prompts the user to answer. Research argues that this form of educational dialogue effectively explores concepts unfamiliar to students [303]. We argue that this approach will fulfill the definition of an "interlocutor" system proposed by Ye et al. [361]. Simultaneously, the Socratic method provides friction to the interaction as called for by Cai et al. [58]. By providing a concrete example of a critical thinking AI, this thesis aims to expand on the literature showcasing the effectiveness of non-passive AI systems.

In the following thesis, recommender systems are heavily discussed in Chapters 4 and 5, while genAI and critical thinking AI are featured in Chapters 6 and 7.

2.3 Related Work

In addition to the terminology that will be used throughout this thesis, we also have provide some related research that present the gap within HCI literature that this research aims to fill. The first gap is within research looking at the impact of AI on art, we will cover a selection of prominent research done in this space and highlight how there is more research needed to explore how AI impacts art appreciation amongst audiences. The second gap is within research on the impact of Social Media, again we cover prominent HCI research in this space and highlight how there is a gap concerning specific impact on audience interactions with creative work.

2.3.1 Research on AI and Art

2.3.1.1 A Brief History

Complex computational algorithms entered the art world as early as the 1950s and can generally be grouped into two categories of programs: creation and classification of art. Starting with the creation side, Hiller and Isaacson's work with the Illinois Automatic Computer (ILLIAC) and computer-generated music was the first step in what is now known as computational creativity [155]. Hiller and Isaacson explored how their massive computer at the University of Illinois could handle "nonnumerical" data, specifically music notes. In the 1960s, Vera Molnar, Frieder Nake, Michael Noll, and Georg Nees were pioneers in using computers to print graphics and geometric expressions [227, 139, 304, 235]. It is this early work that has paved the way for advances in computer graphics, digital art, and, more recently, AI-generated art. The most recent example of a growing interest in art and machine learning has been the release of DALL-E [269], Midjourney [224], and Adobe's Generative Fill [6].

When it comes to the classification and processing of art, Herbert Simon and Marvin Minsky, two central computer science and AI thought leaders, contributed to discussions on computational creativity [225, 184, 300]. Their early work exploring patterns in music data looked into what it meant for computers to recognize unique creative work. As more compute power became available researchers developed MIT's MosAIc algorithm and the work by Thread Genius, and AI became capable of visually interpreting art and identifying similarities between art pieces [132, 306]. MosAIc specializes in connecting the visual characteristics of famous museum works of one culture with similar visuals in works by another. On the other hand, the AI of Thread Genius focused solely on matching visual characteristics and recommending similarities. The work of Thread Genius was valuable enough for Sotheby's to acquire the team and integrate them into the auction house's product team [198]. The acquisition serves as one example where computation and ML work have been integrated into the financial bodies of the art world.

Since the 2018 acquisition, research continues to explore the rich data that can be found in connection to art. Previous iterations of the NeurIPS Datasets and Benchmarks track have also seen their fair share of papers interested in exploring art data in greater detail. The introduction of the Met Dataset and Artsheets for Art Datasets, among others, is a sign of an active community of ML researchers interested in processing the complexities of artworks [362, 308, 3].

As advances in AI and other advanced algorithms unfolds, the need for detailed, well-documented data is apparent. Any algorithm is only as good as the data with which it is built upon. When it comes to the history of art and computation there have been struggles to build up significant accessible datasets large enough for powerful model development.

2.3.1.2 Art in HCI Research

As a field, HCI lends itself well to the "Arts" (visual, performing, and liberal) [81]. There is a rich history of improving artistic creation with HCI methods [46, 349] while bringing artistic practices into the lab space [32, 33, 35].

For this thesis, we have selected recent HCI literature that explores how art is disseminated and displayed to audiences [45, 179]. However, the focus here has been on designing and developing new physical artifacts that would exist in gallery spaces or inside homes. There has been, thus far, little retrospective work looking at the existing digital delivery system offered by recommendation algorithms and how they impact the audience's experiences with the art. The closest previous work has come was to search for new business models that can be built on various ways art can be displayed [187].

A growing recent body of work looks into the impact of digital interfaces on how creativity and culture are presented and disseminated. As digital interfaces become a form of conveying information, they have opened the doors for artists to explore new forms and methods [15]. In HCI, there has been some exploration into what the platforms look like and how that changes the new forms of digital art produced [14]. Much of this work draws historical parallels between technological advances and modern ones. One example is a comparison between the new freedom experienced by authors and literary creatives after the invention of the Gutenberg press and the rise in popularity of e-books today [16]. While this work gets closer to reflecting on the impact of modern digital interfaces, the primary focus has been on the artists and the platforms' business models [259].

This thesis must recognize other HCI authors' work to explore the relationship between art and HCI research. By providing this context, the gaps in the existing research are clear; there is a need for research that places the current, popular recommendation algorithms under a microscope to understand their potential impact on end users concerning art and culture.

2.3.1.3 Generative AI and Art Creation

As outlined in Figure 1.1 of the Introduction, curation and appreciation are only two aspects of a larger system. Therefore, a body of research explores how artists and creatives are changing their processes concerning introducing new technology. We highly recommend reading the work of our colleagues for details on the research, but we will include summaries of a few seminal pieces here which inform the research in later chapters.

Laura Herman has documented the impact of social media and algorithmic decision-making on artists through her research. Her findings argue that social media platforms have their own algorithmic cultures, which creatives must abide by to reach their audiences; this, in turn, shapes their creative practice [151]. We find supporting evidence for her findings through our work in Chapter 4. Her work documents how the AX of algorithmic platforms shapes the creation process of artists and creatives [150]. Finally, in a larger initiative to address the potential impact of recent releases of genAI tools, she has contributed to discussions on how a tool like this serves as a new medium informing the artists' processes [106]. We move the focus of genAI in Chapter 7 beyond artists to explore how it can aid audiences' experiences with art.

Caterina Moruzzi's research explores the changes to theory and artistic approaches in light of technological intervention. She claims that the reality of a world with genAI is accompanied by questions of what creativity is [230]. Moruzzi poses the question, as the artist's process changes to include AI, and the AI is a major part of the creation process, is it also creative [231, 232]. Her philosophical work explores the realities and perceptions of creativity and what it means with new technology [229]. Similarly, in Chapter 5, we ask whether a machine can perceive creative work like a gallerygoer through computational aesthetics. We find that computational aesthetics can form an interpretation of creative work, but that interpretation may not align with human experiences.

Renkai Ma has explored the experience of content creators attempting to survive with their creative work on social media platforms. His research highlights creatives' relationship with the platforms [203]. In particular, he focuses on the moderation aspect of the AX [201]. In exploring this feature, he highlights how much effort creatives put into appeasing the often opaque requirements of algorithmic platforms [202, 200].

Similarly, Ankolika De presents the concept of algorithmically mediated creative labor to describe how the artist's role on digital platforms has expanded beyond making art and involves multiple additional forms of labor [92]. The research argues that artists and creatives spend less time on their artwork and instead work as social media managers, sales associates, and marketers to ensure their success on Instagram or other social media sites. The work of both Ma and De contributes to the motivation for Chapter 4. Specifically, De's concept of "algorithmically mediated creative labor" [92] is directly supported by findings from our participants.

The work done by these scholars has shown the real trade-offs that have to happen as algorithmic systems become commonplace. There is a demonstrable impact on the work artists and creatives make. Referring back to the figure at the start of this thesis, these researchers have explored the impact of algorithmic decision-making on the creation process. Therefore, there exists a gap for this thesis to fill. By examining the impact of these same technologies on audiences' experiences with art, we can complete the creation, curation, and appreciation flow of artworks in digital spaces.

2.3.2 Research on the Impact of Social Media

2.3.2.1 Short-Form Video Social Media

Today, short-form video *content* dominates modern social media. A format initially popularized by Vine in 2013 and later achieving massive success through the merger of Musical.ly and Douyin, forming TikTok in the USA [50, 174, 344]. TikTok's popularity prompted Western competitors like Meta's Instagram and Google's YouTube to introduce Reels and Shorts, creating their versions of the endless feed of short video *content* [174]. In this thesis, social media references pertain specifically to these platforms and user experiences.

In addition to serving as social spaces, social media platforms position themselves as hubs of creativity [374]. This aligns with the concept of algorithmic recommendation systems as "cultural infomediaries," a contemporary iteration of Bourdieu's "cultural intermediaries" [228, 52]. While Bourdieu referred to traditional institutions like human-led museums, Morris's "infomediaries" encompass all information shared with audiences through algorithmic decisions. Examples of *art* infomediaries include Colossal[1] and Europeana[2], connecting audiences with information and framing cultural artifacts. HCI and critical media literature explore how social media platforms' branding and interface design stimulate user-generated content (UGC)

[1] https://www.thisiscolossal.com/about/
[2] https://www.europeana.eu/en/about-us

by encouraging people to create in response to the *content* they consume [343, 102]. Recent research in HCI and computational aesthetics highlights concerns about the impact of social media, particularly Instagram, on artists' modes of expression [204, 210, 150]. Artists adapt their work to align with the perceived standards of trending UGC on Instagram and other social media platforms. For example, how two contemporary artists implement their creative work based on differing social media strategies [360].

While the overt focus of HCI research is on providing a digital space to foster creativity, research indicates that creatives are spending fewer labor hours on their artistic craft and instead are spending more hours on other forms of labor such as marketing and promotion to appease the algorithm [83]. Users understand an algorithm recommends *content* to other users, and success requires emulating trending items [288, 342, 96, 95]. Despite platforms claiming to promote creativity, researchers argue they foster a memetic communication style that relies on replicating and remixing existing *content* [365, 373, 129]. Such user behavior establishes the problem space this thesis shares with other HCI work; the system design around the algorithm eventually shapes user behavior and perceptions.

HCI has a history of investigating and redesigning social media for various social behaviors, involving multiple stakeholders directly [12]. Early research aimed to modify social media experiences to encourage interest in science or civic participation [7, 281]. Recent work explored redesigning social media with the specific needs of autistic adults in mind [37].

When specifically looking to redesign social media for creativity, researchers examined how the structure and design of social media platforms impact creatives and their ability to produce creative works. This includes exploring how platform structures foster the development of unique ecologies for creators and their followers [200, 371]. Scholars argue that creatives now undertake additional invisible forms of labor essential for survival on the platforms [301]. For example, social media has expanded the roles of individual authors who rely on the platformed Internet for self-publishing and self-advertising creative works [297].

These HCI findings build on the body of work exploring the exploitative nature of social media platforms in general and how invisible, uncompensated labor is necessary for *content* to exist but is never actually compensated for by the systems [118, 119, 185]. Informed by Marxist theories on labor and exploitation, critical media scholars analyze the power dynamics within social media sites [273]. Key areas of focus include unequal compensation for labor within digital social networks and the disproportionate impact on users based on gender and racial classifications in online spaces, altering the creative labor relationship [111, 332, 100].

Even dissenting voices acknowledge the power imbalances disadvantaging creatives within major social media platforms [272, 331]. Critical media studies and HCI research underscore the impact of short-form video *content* platforms, emphasizing the reduction of diverse human expression to the generic label of *content* [301]. Therefore, this thesis responds to prior calls in the literature by involving curators and social media audiences to dissect and address issues with visual arts being treated merely as *content* through participatory and iterative design.

2.3.2.2 Algorithmic Bias

While the age of AI and powerful algorithm-driven platforms have aided major societal developments, it is essential to recognize algorithms' potential risks and impact on humans and how this relates to art theories. Bias, persuasion capabilities, and unethical targeting are all examples of user experiences with massive online platforms and algorithms. To ensure this research approaches the topic of algorithmic curation with a firm understanding of the power of algorithms, it must explore the following previous research.

Algorithms' effectiveness depends on the data quality used to train them. Bias within the data used to train and validate algorithms can be multiplied and spread at a massive scale due to the efficiency and scope of a widely used system. Work focused on investigating discrimination and bias within platforms, and recommendations algorithms have provided multiple processes to identify the underlying bias spread across the Internet [284, 321, 130]. Further work has explored the specific racial or presentation bias types regarding the information provided [65, 24]. An example of this within the context of music was presentation bias created by the artist's popularity [63]. It was found that there is decreased discoverability due to famous artists keeping the top spots of audience attention in the recommender systems. The recent research in the field of marketing concerning popularity-driven presentation bias is a rediscovery of a sociological phenomenon known as the Matthew Effect. In 1968, the Matthew Effect in science was used to describe how famous scientists and researchers got more attention and funding than others [218].

Two bodies of research explore whether algorithms lead to diversity and increased discoverability or homogeneity. The first body of work explores what has been referred to as the "Long Tail," where all the content is distributed with decreasing popularity. The idea is that algorithms will allow users to find their niche within the "Long Tail" as they will help serve up the vast number of options and content available [17, 254]. Opposed to this argument is a body of work highlighting how current recommendation algorithms struggle to provide users with diverse and novel recommendations at scale [370, 171, 283]. Fleder and Hosanager explored the two approaches in their work on consumer product recommendations and identified the popularity bias happening at scale [112]. Their work notes that while individual instances of discoverability for an individual user may exist, the overall trend highlights the viral content. By challenging the AX of recommender systems, we can explore ways to combat popularity bias. In this thesis, we will consider how even expert users and creators of recommender-driven social media may have opinions about the biases of social media toward certain content types.

A related set of questions is how salient the recommendations from the algorithms are for the public. Related work demonstrates that the bias within algorithms can be potentially dangerous due to their power as persuasive tools [130, 136, 11]. One body of work explores the power held and exhibited by algorithms through the experiences of a set of users [55]. The researchers highlight that the user perception of how and where the algorithm impacts might not always be correct but is often more indicative of its impact on the individual. Additionally, the work reveals a bidirectional impact where the user impacts the algorithms as much as the algorithm shapes the user's experience. This was further explored in other work that models this relationship and can again demonstrate the popularity presentation bias described in the previous paragraph [364]. In Chapter 4, we explore how these theories of algorithmic bias and corresponding influence on users' perception carry over to the subject of visual art.

2.3.2.3 Previous Work to Re-Design Social Media with Co-Design

Like art, recommendation algorithms, specifically those used by social media, have been studied in HCI literature through a variety of relevant and effective methods. To address the task of elevating visual art pieces from the sea of content, this thesis builds on techniques provided by HCI researchers known as research through design [372, 127]. A demonstrable body of research highlights the value of applying design research methods to HCI research on behavioral theories, social systems, and technology [109, 357, 103]. Specifically, this thesis employs user-centered design approaches from algorithmic experience (AX) and participatory design (PD) research to create artifacts that were critiqued and iteratively improved by the experts and users most likely to be impacted [13, 48].

AX argues that the digital systems' user experience around any algorithm impacts the relationship between the algorithm and users [13]. Recently research has been done to apply AX methods to understand user perception of algorithmically curated music experiences [117]. To position the thesis within this canon, the work is contextualized within the methods employed in AX and PD. and describe how the methods were used to explore the experience of visual art viewing on algorithmically recommended short-form video content [145, 48].

Various combinations of these methods have been applied to redesign social media to target specific experiences and user perceptions. Early research worked to alter social media experiences to foster interest in science or civic participation [7, 281]. Following the 2016 election and misinformation scandals, HCI research focused on redesigning social media to handle political polarization and content

[239, 138]. Recently, the idea of redesigning social media with specific stakeholders' needs in mind was explored with autistic adults [37]. These papers provide examples and guidance on how research into redesigning social media can be done through the direct involvement of various stakeholders [12].

This thesis builds on the existing body of work and established methods to consider redesigns of social media specifically for visual art experiences. With visual art, we have a unique opportunity to access a specialized population of experts with years of experience in analyzing human artifacts to select and elevate individual pieces to the status of visual art. This expertise should not be wasted, by using curators' and gallerist knowledge it is possible to develop artifacts for HCI scholars to explore, test, and discuss.

2.4 Conclusion

Accumulating this body of literature has helped to position this thesis within the established canon while simultaneously revealing the gap in existing research it aims to fill. Previous research provides the grounding of this thesis. Museum Studies scholars and social theorists define curation as an evolving practice of documenting, selecting, and presenting art. Their research highlights how digital systems act as curators, providing the scope of how curation will be used in this thesis. Similarly, art historians and aestheticians offer multiple ways to appreciate art. However, we select the psycho-historical framework of art appreciation as the primary definition of art appreciation. Finally, as this is a computer science thesis, curation and appreciation are connected to the language of algorithms and algorithmic experiences. There are three types of algorithms relevant to understanding the presented work: recommender systems, genAI, and critical thinking support AI. With these definitions in place, we present previous research that reveals the gaps this thesis aims to fill.

Among researchers exploring AI's impact on art, most focus on the creation of art. The role of algorithmic processes on creation has been historically dominant from early developments in AI through the massive success of music recommenders, and all the way to the release of generative models. With such a heavy focus on creation, there is room for this thesis to explore what the impact of the same algorithms is on **audiences' appreciation** of art. Second, by examining previous research on the power of social media on users, we highlight the power of algorithmic bias and how the creator economy changes how people spend their time. There is also a body of research that attempts to address the impact of social media on users by redesigning the experiences. Again, the literature shows some instances of impact on audiences but does not address the impact in relation to art. With these gaps presented, the next step is to outline the research design for this thesis and share how the projects begin to fill the gap in the existing literature.

To conclude, this review presents the research completed to ground the following methods and results within the established canon. The research provides new artifacts allowing researchers and users to reflect on the digital world's impact on art and culture. It also contextualizes how developments within the field of HCI in computer science can impact the broader art world.

3
Research Design

Contents

3.1 Introduction

In reviewing the related work within the areas of art, algorithms, and social media it is evident that there remains a gap exploring the impact of digital experiences on audience's appreciation of art. To fill this gap this thesis employs a research design divided into four parts, where each subsequent project iterates on the findings and methods of the previous project. The following chapter outlines the methods associated with each step of the research design and the justifications and considerations that need to be accounted for the method implementation.

The first step of the research design is exploratory work into understanding the status quo of social media and its impact on art appreciation. The study aims to answer RQ1: What do experts and regular users have to say about the impact of recommendation algorithms on their engagement with art and culture? For this work we implemented a series of methods drawn from user-centered research and design such as interviews, surveys, co-design sessions, and stakeholder reviews.

The second aspect of the thesis explores an experimental condition comparing the implementation of social media recommendation and curation within the context of a gallery space. Here the study aims to answer RQ2: What are the differences between audience observations of a physical gallery with algorithmically curated images and a machine's reading of the same images? In this case we employ computer vision techniques outlined in the field of computational aesthetics to understand machine perception of art and its relationship to human perception.

The third aspect is built on the findings of the second, by understanding that there are certain pieces of art metadata that are not accurately recorded or maintained that may inform digital experiences. RQ3 posits, What factors, values, or variables are *not* currently at play in algorithmically processing art, and how can they be informed by expert art evaluation? The methods for this section involve the development of a novel dataset and then testing the dataset with a benchmarking task.

Finally, the research design concludes with the development of a new digital experience based on the accumulated findings of the previous studies. We explore RQ4: How can we use the unique capabilities of large language models to digitally support audience engagement with art based on the needs of our past participants? The study involves the development of an experimental prototype that can be used a technology probe in an experimental condition to understand how it alters the user experience and behavior of art audiences.

3.2 Researching Experiences

To understand pain points and opportunities that currently exist in how algorithmic processes impact art appreciation, the research needs to conduct an exploratory study with a variety of users. The primary method for such an exploratory study is user centered research involving surveys, interviews, co-designs and stakeholder reviews. In this thesis we conduct these methods with art experts, artists, content creators, and general social media users that responded to our posts. These methods are common in HCI research and are grounded in previous literature that has applied them for similar exploratory research.

3.2.1 User Centered Research

This thesis approaches data collection from a user centered research (UCR) perspective. As part of HCI research, UCR focuses solely on supporting researchers in direct interaction with users and participants in a research setting. The textbooks of Karen Holtzblatt [159] and Yvonne Rogers [278] offer an overview of various user centered research methods that turn participant responses in to usable research findings. In particular there are three sets of UCR methods that are applied heavily throughout the thesis. First, interviews and surveys, are methods to gather preliminary and follow-up data from participants through the use of semi-structured interviews and targeted questions. Second, co-design workshops drawn from a philosophy of participatory design [48] where users are brought into design and work alongside researchers in the development of new research learning. Finally, stakeholder reviews, is a method for presenting users with materials for their consideration and feedback.

3.2.1.1 Interviews and Surveys

The use of semi-structured interviews in our research is grounded in HCI and research-through-design principles [372, 127]. Established HCI research often begins with interviews and surveys to understand experts' and users' positions within a system [109, 357, 103]. Semi-structured interviews, in particular, are valuable HCI methods that allow researchers to ask follow-up questions that are not immediately listed within the original interview script. We conducted semi-structured remote interviews via Microsoft Teams, using real-time transcripts for data collection as required by our review board. The transcripts captured all spoken vocal utterances. Typically, two researchers were present: one conducted the study, and the other assisted with notes and observations.

Unlike interviews, surveys offer HCI researchers a larger pool of potential participants to gather qualitative or quantitative data. However, as the surveys are often completed without the researcher present, there are limitations, such as no follow-up questions. Additionally, they often are expected to be shorter to increase the likelihood of being filled out, and come with the risk of falsified data from bots or scammers. Using the interview script as the basis for the survey allows for significant overlap between the questions and potential responses.

3.2.1.2 Co-Design Workshops and Cognitive Walkthroughs

Co-design workshops allow HCI researchers to glean insights into participants' varied perspectives on the same prompts and collaborate to generate artifacts and research findings [178, 314]. Co-design is a method from participatory design research [48], a theme in HCI literature that argues any designs built for specific audiences should be designed alongside intended audience members. Many co-design workshops are collaborative events where multiple users proceed through activities facilitated by a researcher or workshop leader. Co-design workshops allow researchers to learn from participants about their thoughts, ideas and approaches to the research space of interest [178]. HCI researchers can observe how participants with differing perspectives encounter the same prompt, activity, or scenario and ideate together to output artifacts and, eventually, research findings [314]. These methods are often conducted with the end users of a system with various degrees of experience [313]. Each activity elicits responses regarding the task or research, often through various outputs such as sticky notes, diagramming, or voting.

Before the workshops with participants, researchers conducted a cognitive walkthrough and a pilot workshop. A cognitive walkthrough involves an expert's external, informed review of design decisions' potential impact on the user experience [353, 359]. A gallerist with expertise in private and national galleries conducted the cognitive walkthrough. The pilot workshop, involving HCI researchers within the group, focused on practical aspects of workshop design, the whiteboard tool, and exercises. Final adjustments were made based on feedback from these preparatory sessions before the co-design workshops.

3.2.1.3 Stakeholder Review

The final method employed in this project was a stakeholder review [131]. Stakeholder reviews are a method from industry-focused HCI that goes beyond a usual user test. User testing is a method that focuses on how a user will use a prototyped tool and is meant to test the tool itself. A stakeholder review is meant to understand the exact role of the user as a stakeholder within a larger network. The questions during the review are not just about the usability of a prototype but also about the values, ideas, and concepts the user has about the larger ecosystem within which the system will potentially be deployed.

3.2.2 Limitations and Justification

As with any method being selected there are some limitations. In this case, the findings from user centered exploratory research are all qualtitative and potentially vary depending on the sample of users and how the questions are phrased. To account for this we ensured the questions were standardized across users and underwent research review within our institution. We also ran pilot tests within our research group to review the content being presented to the participants.

Another limitation specifically connected to the co-design method is that our participants are not necessarily equipped with a design background. This is quite common when bringing users to inform the design process. In our case we were working with curators who might not have traditional design training, do have experience with design principles when it comes to organizing collection and gallery spaces. Furthermore, to support them in the co-design we had multiple researchers present that could act as the “hands” for the participants to turn the ideas into designs.

User centered exploratory research is a common set of methods employed by researchers in HCI. We felt justified in selecting these methods are they have a demonstrable history of providing researchers with broad examples of where experiences are going well and where there exist pain points and opportunities. Using these methods allowed us to answer RQ1 to understand the impact of algorithmic decision making on audiences’ art experiences.

3.3 Researching Gallery Reactions

While the first part of research design gathers information on what people think about experiencing art in digital spaces, the second study explores how an algorithm may interpret a traditional gallery experience. We employ an established method of computational aesthetics to perform a comparative study between an artist curated side and an algorithmically curated side. Unlike the first study this is no longer exploratory and results in quantitative metrics that can be compared to reveal exactly what differences a computer vision program identifies within different gallery pieces.

3.3.1 Computational Aesthetics

In this project, the research addresses machine perception on its own. Building on previous research, we seek to gain perspectives on the machine’s observed similarities and differences between human and algorithmic curation. In contrast to human perception, this project presents the metrics derived from applying computer vision to the exhibit pieces. This approach is comparable to art world projects that explored the interpretation of art through computers as exhibit pieces and exhibit critiques [336, 309]. We build on this by revealing the fundamental metric differences revealed by computer vision code processing artistic images.

This project presents a technical analysis of the *Algorithmic Pedestal* exhibit using baseline, replicable, and open-source computer vision code to process and compare the images within the exhibit. While the past few years have seen massive improvements, the field of computer vision is not young, spanning back to the 1960s and 70s. Please see Szeliski's textbook for a thorough background on computer vision and its possible applications [322].

Computer vision advancement went hand-in-hand with the advancement of computer graphics. In the '60s and '70s, hardware and computational resource limitations hindered advancements. Nonetheless, the early findings outlined the edge detection process, a concept drawn from human visual perception processes; detecting an edge is often the starting point for larger shape recognition and depth perception [60]. The '80s and '90s continued these advancements, and computer scientists could represent 3D figures using multiple images layered on one another. In 2001, the Viola-Jones facial recognition paper marked a massive step forward for the field [338]. Since then, computational power and resources have made it possible for advanced systems using convolutional neural networks (CNNs) to excel at image recognition and other tasks [177, 191]. As Szeliski mentioned [322], these advancements are associated with a specific problem or task researchers are setting out to solve, often requiring some degree of ground truth (i.e., the image either contains a dog or it does not).

As we shift focus to the arts, testing for a specific ground truth or task becomes arguably more challenging. Researchers have set forth the field of computational aesthetics as one potential use case for computer vision in the arts [156]. This field allows researchers to create specific computational systems that consider aesthetic theories devised in art history and apply them using an algorithm, for example, overlaying the golden ratio onto an image [47]. In addition to these higher-order interpretations of images and aesthetics, Bo et al. state the field also applies some of the early computer vision methods of edge detection, contrast, and color identification [47]. As a result, these researchers can develop specific systems that review art pieces on these various metrics.

Within this project, our technical analysis of the art gallery takes pieces from background work in general computer vision and computational aesthetics. Using a set of computer vision software combined by Hosny et al. [162], we analyzed the pieces from the Metropolitan Museum of Art's Open Access collection and the pieces selected for the exhibit. The computer vision code was released as open source and started as an MIT class project picked up by Artnome, a digital collection of articles focused on understanding the role of data in the art world [26]. Based on this public acceptance among art data publications and the legitimate computer vision methods, such as edge detection, contrast, and face detection contained within the codebase, it was deemed appropriate for this project. The original code needed some adjustments to account for out-of-date code libraries. Since the images and curation were all done using established research methods and open-access images from a public digital collection, they offer a controlled sample upon which the code can be analyzed.

3.3.2 Limitations and Justification

Computational Aesthetics as a method has a noteworthy technical challenge. As the method of evaluating art this way involves researchers using various computer vision approaches to processing an art image there is a large variety in the scope and capability of the aesthetic analysis. For example, the process for edge detection, while complex is not nearly as computationally taxing as using a CNN trained on art data to process the images. In this case a researcher may be limited in the amount of computing resources they have and as a result it will impact the types of computational aesthetic analyses they can run. For the purposes of this project we intentionally selected the fundamental computational aesthetic metrics. These fundamental metrics allow for accurate comparison, but do not represent the entirety of what a high powered system could do.

The second limitation we found with computational aesthetics came as a result of comparing the findings of this project to previous work done by Herman and Moruzzi [153]. The metrics used in computational aesthetics are not representative of the reported differences in experience of the audience members. In other words, just using computer vision to analyze and present metrics of a piece does not accurately capture the type of data people are interested in. It is this limitation that helps to motivate the next step in the research design. By understanding that traditional computational approaches to analyzing art are not representative of the human art appreciation process, we must consider what other types of data and system can be developed to better give audiences the art they are interested in.

3.4 Researching New Datasets

With the previous research projects it has become clear that there are certain aspects of art data that are not considered by algorithmic art processing. Therefore, for the third project of the thesis we explore the development of a novel, open-source dataset and show how it can be used by an algorithm to develop its own classification and language of aesthetics. The research design for this project involved the development of a dataset and benchmarking the dataset in comparison to previous research to show that there is a value in the collected type of art data.

3.4.1 Dataset Development and Benchmarking

The methodology used for this project was focused on building a new dataset of art metadata. Dataset development involves two major steps, the collection and the cleaning of the data. Both are essential to ensuring high quality information and carefully structured data for the purposes of training or running an algorithm. Another method involved in this project is the benchmarking of the dataset by running an example use case. In this case by using natural language processing, the dataset is able to be classified and organized into art topics according to the patterns identified by the algorithm. Through this combination of methods the usability and viability of the dataset for researchers and developers can be demonstrated.

3.4.1.1 Data collection

Due to the limited access to existing datasets, data collection had to be done specifically for the project requirements. That meant relying on established web-scraping techniques to gather data from publicly accessible websites. By defining variables that correspond to specific desired art metadata, our code can extract the necessary information [183]. As a result, there are a variety of tools and methods created by data scientists to support mass data collection through web scraping [175]. The most common methods are code libraries for programming languages like Python provided open-source, two common scraping ones being Selenium [294] and BeautifulSoup [276]. While some critics consider tools like these too specialized, requiring too much knowledge of browser functionality [68], they remain generic tools that need to be customized to gather usable data.

Therefore, Chapter 6 outlines how a custom web scraper built on the public information of previous web scrapers was built to gather art metadata. Data collected from past public auction records support exploration into the relationship between the variables and art analysis. Auction houses focus on financial value, but their data still informs audiences perspectives and appreciation of art. Unlike galleries, these sale records are posted on major auction house websites, available for all to explore. Furthermore, these details are only about the auction lots and do not include any information regarding audience members or art purchasers. This adherence to anonymity by the auction houses ensures that the data does not handle private or sensitive information.

3.4.1.2 Data cleaning

Upon collecting the initial data, there is a need for substantial cleaning. Since the early 2000s, big data research has explored the impact and danger of messy data [268]. Consequently, data scientists have been working on methods to clean data into usable information, mitigating the harm it can have in skewing predictive models [169]. For the most part, early work in data cleaning has focused on quantitative data such as techniques for discovering outliers or other anomalies. However, *art* metadata is not only quantitative but also qualitative.

There is a need for data cleaning arising from the variability of different pieces and where the auctions were held. For example, some pieces had multiple measurements due to the inclusion of the frame, while other pieces were sold in Paris and had details repeated in both English and French. Another challenge was handling the inconsistency of details regarding text-based variables such as Condition Reports, Provenance, and Historical Significance. These variables were non-quantitative accounts of the piece's history and current state and were of variable length and detail across the auction lots. Therefore, it is important to consider various qualitative data cleaning techniques [76] to ensure these types of errors are caught and accounted for before proceeding with any analysis.

Ultimately, the additional details were captured and labeled as notes and descriptions relating to each of the lots. Grouping these notes formed a single string variable that can be further parsed. The data had to be manually verified before being divided randomly into a training and testing set. Further discussion on the impact of the data cleaning and considerations for the ethics and limitations of the dataset can be seen in Chapter 6.

3.4.1.3 LDA Topic Models

Recent advances in machine learning have lead to natural language processing (NLP) through vector space modelling (VSM), which allows for code to interpret text data [271]. Specifically, the models turn words into mathematical vectors that can be calculated by the computer to represent meaning within the text data. One such use case is the development of topic models such as Latent Dirichlet Allocation (LDA) which groups text into specific topics generating new groupings of relevant data [43]. These topics can help recommend new items to users by seeing which topics are relevant in different scenarios. With our data collected and cleaned, it is ready to be analyzed by a relevant topic model that can explore *art* text. There is limited previous work on *art* datasets and *art* recommendations that are readily available to researchers. One of the relevant, available research projects was an approach to recommending art based on various text descriptors and an LDA topic model [5]. By using this work as a benchmarker, it is possible to evaluate the effectiveness of the newly created dataset by using the same methods and measuring performance in the topic distribution and coherence metric. Topic distribution metrics tell us how well the algorithm is able to distinguish between different types of entries in the dataset and place them into unique topics, while coherence metric refers to how coherent the words used to identify those topics are [43, 5]. Only once the dataset has been benchmarked and validated can new methods or models be applied to explore additional usability and valuable research.

3.4.2 Limitations and Justification

The limitation of this method is that the dataset is quite unique and any comparable datasets are paywalled and limit researchers' access. Therefore, the analysis and dataset development are done using comparable methods applied in similar situations, but do not have a direct reference point. This limitation serves as additional motivation to release the dataset open-source. Doing so will allow other researchers to review the dataset and utilize it in their own research. The goal being the dataset can be improved and revised as more researchers explore the art metadata for valuable insights.

The second limitation is that the current dataset and algorithmic process are not particularly user friendly. The current format of the work in this project is not something that could be presented to novice users or lead to high quality interactions. This motivates the final project which looks to combine the learnings from all of the previous research projects into an experimental prototype powered by our data.

3.5 Researching Potential Futures

The third aspect of this thesis proposes building and developing an experimental prototype that questions how the findings of the previous sections can be combined. Addressing the potential value art metadata can bring while crafting an experience informed by the needs of experts in the art world and in social media. Therefore, the overview of the methods provides a high-level description of various handbooks and previous research that specialize in building and testing prototypes.

3.5.1 Research Through Design

Research through design (RtD) is an approach to HCI research that aims to integrate design practice into human-centered research methods [372]. Early approaches of RtD aimed to systematically test theories by turning theory derived principles into designs that were then evaluated to reflect on the validity of the original theory [80]. Scholars have since argued that iteratively developing and testing a design or prototype can also develop knowledge that develops new theories, not just testing established ones [104]. Bardzell et al. state that focusing on the design and development of a research prototype or object offers multiple valuable outcomes for researchers [34]. First, the creation process explores how theory and literature can be applied in practice. Second, users' "critical reception" of the prototype provides knowledge that can drive deeper learning into new theories and research directions.

As this thesis is iterative from one project to the next, the final project aims to provide insights developed from the creation and testing of an experimental artifact that has been informed by the previous projects. Therefore, our research plan for the final project is informed by various types of prototyping methodology including speculative design,technology probes, and provotyping.

Speculative design approaches the development of artifacts with an aim to present various potential experiences. The concept of "speculation" is central to the design process where designs balance novel innovation with familiar concepts [22]. By providing participants with an experience that is grounded enough in their current mental model, researchers can push certain details into a futuristic state to inspire reflection. The methodology relies on the acceptance that the current design or prototype does not exist beyond the research context and that details may change before it ever becomes a reality. As we explore potential futures of digital art experiences, we will need our users to extend their belief of what is currently available and where technology may lead in the future.

Technology probes take a more grounded approach. This methodology develops prototype technology that is functional and can be deployed with users for the purposes of data collection [164]. The requirement of being technologically feasible is where this method differs from speculative design. However, both methods focus on the end-goal of collecting feedback from users about how the experimental technology may impact behavior and experiences. The philosophy of technology probes aims to fulfill three goals: 1) collect information on users and their usage of technology, 2) "field test" technology, and 3) inspire thinking about new solutions and systems [164]. For the final project we aim to meet all three aspects of the philosophy, by building a prototype that collects data during the testing session, is functional and interactive for the testing, and leads to feedback from the participants about this form of technology.

Provocative prototyping, commonly shortened to provotyping, is another design method we draw inspiration from in the third project. As the name suggests, the deployment of the experimental design is meant to provoke users to spark reactions and reflection. It is important to note that this is not done in a way that would harm the user or in anyway place them in unethical situations. Rather, provotyping aims to "inspire reflection by stakeholders" [49] by building "technologies that would likely foster strong feedback" [327]. Unlike speculative design and technology probes, the provotypes do not come with a requirement of presenting something highly novel. In fact, they could present a standard prototype or experience that instead inserts friction into the experience to gather thoughts and feedback from users. We highlight this method of prototyping because it shares a similar philosophy to "antagonistic" AI, outlined in Chapter 2.2.3.3. Both argue that researchers and users can learn from experiences that are carefully constructed to challenge the user.

In addition to these foundational methods, the final project of this thesis draws practical learnings from previous work. From RtD in museum and cultural heritage spaces we are inspired by insights that reveal the benefits of interactive technology within museum spaces [78, 176, 66, 302, 9]. Similarly, from recent HCI work we can see how technology probes have been implemented as a research method and accompanied by experimental studies for assessing the effectiveness of the probe [319]. Finally, other recent HCI work has explored how provotyping can be implemented to represent digital new experiences in novel ways [274]. These studies have shown how the research design plans we have drafted could be implemented for academic research.

3.5.2 Limitations and Justification

The most glaring limitation of these prototyping methods is that speculative design, technology probes, and provotyping all come with the risk that the users will not receive them positively. Hutchinson et al. state the reality of this methodology as, "they might fail or bring unexpected results"[164]. Being aware that the prototypes are not built under the same assumptions of a production-ready system is key to representing the research associated with these methods.

In the context of this thesis and the final project acting as an exploration into the potential future of digital art experiences, we are willing to accept this limitation. The final project aims to combine the outcomes of the previous chapters into an experimental design that can be tested and evaluated. Learning what aspects of this research can be carried over into future development and research is as valuable as identifying what does not work and should not be carried forward.

3.6 Conclusion

The high-level methods presented within this section serve to ground the research projects throughout this thesis within a body of literature that has sought to explore art data, human perspectives on social media, and guidance on developing prototypes. All of the methods employed are established within HCI literature. We apply them to the research design of this thesis because they support the iterative process used to progress the research. Each method is selected and applied to rigorously answer each RQ. Thereby allowing the evidence to build and inform subsequent RQs. The following four chapters will detail exactly how the methods are applied in each project to develop the contributions of the thesis.

Much of this chapter was also reviewed and published in the paper:
von Davier, T. Ș., Noh, H., Van Kleek, M., & Shadbolt, N. (2024). Looking for Art in a Sea of Content: A Human-Centered Approach to Supporting Creativity on Social Media. The Proceedings of the ACM on Human Computer Interaction (PACM HCI'25). 9, 2, Article CSCW127 (April 2025), 25 pages. https://doi.org/10.1145/3711025

4

The Human-Centered Experiences of Encountering Art Online

Contents

4.1 Introduction

With 4.9 billion users, social media platforms have firmly entrenched themselves as hubs for human connections [358]. Within the past five years, short-form video social media platforms have claimed a major stake of the user numbers. Simultaneously, the term *content* has evolved to encompass nearly all forms of creative expression in the era of social media, blurring the lines between previously distinct artistic industries like film, music, creative writing, photography, and design [240]. As artworks are predominantly encountered in digital formats, often mediated by algorithm-driven social media, it is not surprising that we refer to them as *content* within the context of the *content* recommendation algorithms governing these digital channels [19]. An inability to distinguish types of *content* creates demonstrable drawbacks such as misinformation and exploitation [256, 277]. When legitimate and false news articles are treated the same by social media platforms, they can misinform the public [10]. Therefore, since *art* is used to define ourselves and society, we must consider the impact of it becoming lost in *content* [347].

It is this premise that begins the research of the thesis, the need to understand the impact of algorithmic decision making on *art* appreciation. It is vital that we consider how large-scale social media treats *art* and presents it to audiences. With this project we aim to understand how to avoid the term *content* based on feedback from members of the art world.

In response to the drawbacks, human-computer interaction (HCI) researchers have called for teasing apart the term *content* to reintroduce specificity when exploring particular areas of human expression online [301]. In our exploratory research, we question whether the medium really is the message [216] or whether *art* can be disentangled from social media. We want to understand how *art* distinguishes itself from ubiquitous *content* and whether the current modes of experiencing creative works (Instagram Reels, YouTube Shorts, and TikTok) diminish *art*'s potential significance. Additionally, we question whether artists identifying as "content creators" can benefit from separating their work from the attention-driven industries that often promote them.

This project uses short-form social media (SFSM) to describe platforms like Instagram Reels, YouTube Shorts, and TikTok, focusing on short-form video *content*. As outlined in Chapter 2.3.2.1, SFSM is the dominant algorithmic experience for recommendation algorithm-powered platforms. Therefore, it has become the focal point for HCI researchers exploring the role and impact of social media on user experiences and creative expression. In this project, *art* is defined as visual arts, a category of art including fine art (i.e. painting, illustration, sculpture) and contemporary art (i.e. collage, photography, assemblage, land art) [340], allowing us to draw insights from established institutions (galleries and museums) that use experts to differentiate artworks from other human-made artifacts. Previous studies concerning social media, have highlighted how the demand for "Instagrammable" art shapes artists' creations [204, 210, 150]. This research argues that artists tailor their work to suit the algorithms powering social media platforms. Additionally, HCI research has predominantly explored content creators' creative experiences in light of recent developments in AI and algorithmic recommendations [301, 117, 297]. Our research introduces a focus on audiences' **perception** of creative *content*.

Our project aims to uncover how the social media experience, especially short-form video, might overlook or alter the reception of *art* as a distinct form of *content*. Users' engagement with social media is a part of the Algorithmic Experience (AX) [13], and this study explores the required AX to differentiate *art* from the vast *content* landscape better.

Our inquiry explores key questions:

RQ1 - What characteristics do artists and curators use to differentiate *art* from *content* when both are presented on social media?

RQ2 - When considering the characteristics separating *art* from *content*, how would curators alter the design elements of social media?

RQ3 - How do content creators perceive these altered design elements?

For our exploratory study, we employ a mixed-methods approach deriving methods from research through design [372, 127]. Our pre-study, including interviews (n=18) and a survey (n=37) with curators, artists, and social media users, aimed to comprehend their current experiences and perceptions of encountering *art* within social media *content.* The findings revealed common perspectives that (1) depth of information, (2) space for meaningful audience conversations, (3) audience-artwork connections, and (4) time for the experience are essential characteristics of experiencing *art.* These characteristics align with the fundamental principles of Slow Technology, emphasizing implicit slowness and temporal interconnectedness [144, 243]. Slow Technology explores redesigning everyday interactions based on specific attributes around time and temporality, facilitating user reflection on experiences and actions.

The main study starts with co-design workshops with curators (n=13), which envisioned experimental digital experiences focusing on *art* rather than encompassing all *content* types. Like algorithms, curators decide what goes on display for audiences and dictate its presentation. These workshops generated low-fidelity wireframes that were cross-examined and iterated upon, resulting in refined interfaces. The interfaces presented a piece of art alongside a depth of textual and audiovisual information without insisting the user quickly scroll away allowing for greater time to view and reflect on the art.

Subsequently, a stakeholder review with content creators (n=10), henceforth referred to as "creatives" aligning with other HCI literature [74, 297], provided feedback on the co-design materials. The term "creatives" comes from researchers' desire to separate the creators from the word *content.* In our participant selection, we considered whether the creatives' social media presence aligned with the project's definition of visual art. The stakeholder review highlighted creatives' struggle to balance meaningful *art* creation with appeasing algorithmic metrics affecting their livelihoods.

In conclusion, this project offers unique insights into SFSM's impact on user perceptions, especially regarding *art* and creative work. It also provides open-source materials for researchers and designers collaborating with curators to redesign digital *art* experiences, see [345]. Ultimately, we will discuss how our research provides practical examples of design recommendations set forth by other HCI researchers.

4.2 Research Overview

This project employs a mixed-methods approach to address the research questions. The study comprises two parts: the Pre-Study, incorporating interviews and surveys (**RQ1**), and the Main Study, consisting of a co-design section (**RQ2**) and a stakeholder review section (**RQ3**). We obtained institutional ethics approval for both studies. For a comprehensive overview of the methodologies and their integration, refer to Figure 4.1. We conducted research remotely to enable global participation from curators, artists, and creatives. Video conferencing and automatic transcription tools were employed for efficient note-taking and record-keeping.

The researchers conducting the research and analysis represent cultural backgrounds from East Asia, Europe, and North America. Their backgrounds are primarily in the field of HCI with one researcher fully focused on art and AI research and another that previously had a professional career as an artist and gallery employee.

4.3 RQ1: Understanding what sets art apart from content

4.3.1 Pre-Study Methods

4.3.1.1 Participant Recruitment

Experts: By sourcing publicly listed institutional emails, we recruited 12 curators from various US and UK museums and collections. For the exploratory study, we welcomed curators of different backgrounds and specialties across visual arts. Participants' institutional affiliations were anonymized, with demographic descriptions available in Table 4.1, identified by code **E** (for expert) followed by a number.

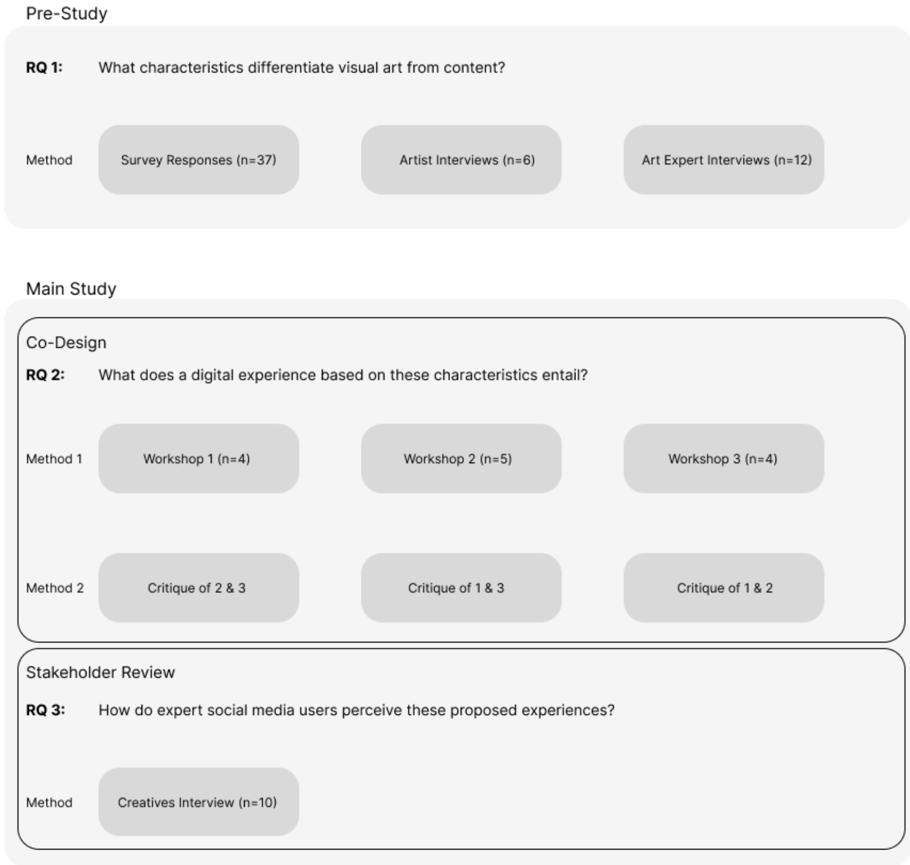

Figure 4.1: An overview of this project's methods and research questions.

Artists: Accessing artists proved challenging, requiring searching personal websites for contact details and posting recruitment messages with sign-up links to Reddit forums. Ultimately, 6 participants from diverse stylistic backgrounds were **recruited via emails gathered from sign-up links**, meeting inclusion criteria as artists generating income through fairs and galleries. More details are in Table 4.2, credited as **A** followed by a number.

SFSM Users: We posted a survey link across Meta, LinkedIn, and Twitter to gauge social media users' opinions, collecting 37 responses over four weeks. Demographic differentiators, primarily focused on their relationship with *art*, guided data collection. Participants are credited as **P** followed by a number. The convenience sample likely exhibits selection bias toward those with established connections to *art* and technology.

Table 4.1: This outlines the IDs and Gender of our experts, a short description of their current role, and how many years of experience they have in the art world [1].

ID (Gender)	Role	Years of Experience
E1 (M)	Private gallery curator	12 years
E2 (M)	Academic and historical curator	30 years
E3 (M)	Curator of drawings	20 years
E4 (M)	Collection assistant	8 years
E5 (M)	American visual culture curator	36 years
E6 (F)	Museum digital content manager	10 years
E7 (M)	Museum digital infrastructure and online collections	20 years
E8 (F)	Modern and contemporary art curator	27 years
E9 (F)	European and American art curator	40+ years
E10 (F)	Chinese and East Asian art curator	25 years
E11 (M)	Keeper of antiquities	30 years
E12 (F)	Manager of education and digital resources	12 years

Table 4.2: This outlines the IDs and Gender for our artists, their preferred medium, and years of experience.

ID (Gender)	Medium	Years of Experience
A1 (F)	Sketching and watercolours	18 years
A2 (M)	Performing arts	8 years
A3 (F)	internet art and machinima	22 years
A4 (F)	Clay sculpting	18 years
A5 (F)	Sculpting with personal technique and materials	26 years
A6 (M)	Painting with acrylics and mixed media	14 years

4.3.1.2 Study Design

We conducted semi-structured remote interviews via Microsoft Teams, using real-time transcripts for data collection as required by our review board. The transcripts captured all spoken vocal utterances. Typically, two researchers were present: one conducted the study, and the other assisted with notes and observations.

The use of semi-structured interviews in our research is grounded in HCI and research-through-design principles [372, 127]. Established HCI research often begins with interviews and surveys to understand experts' and users' positions within a system [109, 357, 103]. Our study aimed to answer **RQ1** regarding attributes distinguishing *art* from *content* for artists and curators. The interviews cover perception of *art* and art theory, online platform interactions, and opinions about social media, allowing for in-depth exploration via follow-up questions. The interview script can be found at [345], and interviews ranged from 30-60 minutes.

The online survey, a concise version of the interview, is also in [345], taking an average of 8.5 minutes to complete, with no follow-ups. The purpose of surveying social media users was to understand to what extent they agreed with the statements of the artists and curators. As the non-specialist end users, they are most directly impacted by the current algorithmic experience of SFSM delivering everything as *content*.

4.3.1.3 Analysis

The pre-study outputs comprised interview transcripts and text-based survey responses. The analysis combined survey and interview data using anonymous IDs to differentiate participant statements. Transcripts and responses were organized into digital sticky notes on Miro[1], then clustered using the affinity clustering method [318, 148]. This method highlights universal claims and coherent arguments, facilitating the identification of primary positions within unique participant groups. A comparison of prominent positions across all (curators, artists, and social media users) groups led to the final set of attributes distinguishing *art* from *content* based on our participant sample.

4.3.2 Pre-Study Findings

Our analysis of the pre-study interviews and survey submissions elevated four characteristics considered essential to experiencing and appreciating *art* instead of *content*. In Section 4.6.1 we highlight how these attributes reflect the facets of Slow Technology [144] offering a novel example of user-generated Slow Technology design requirements dedicated to enhancing *art* experiences for diverse user groups. These four attributes include depth, conversation, connection, and time (see Figure 4.2 for brief definitions).

[1] https://miro.com/about/

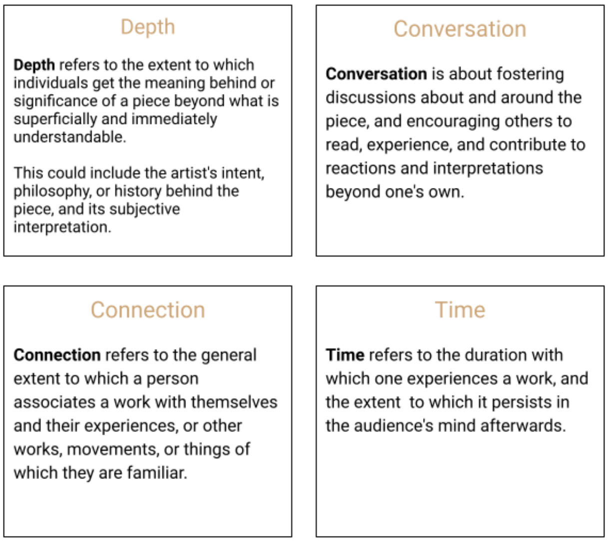

Figure 4.2: Working definition of the four characteristics developed by participants in the pre-study.

4.3.2.1 Depth

During the pre-study, our curators and artists emphasized the reality of the sheer volume of information associated with a single piece of *art*. Our participants argued the more information available for an artwork, the better chances there were of cultivating audience interest in the work (E7, E2, A1, A4, P(12/37)). The importance of accurate information and the intrinsic value of data as a support tool for enhancing *art* experiences have also been underscored in other HCI literature [188]. Our participants assert that possessing a depth of information is an indispensable aspect of *art* experiences. However, while data surrounding *art* can significantly enhance the overall experience, it must be properly labeled and easily accessible—a current challenge (E12). One issue stems from exclusivity measures, often implemented for safeguarding intellectual property rights (E6), which hinder convenient access to additional *art* information despite the potential to profoundly influence audience appreciation (E2, E10, A3, P(21/37)).

Our participants expressed optimism that digital experiences driven by recommendation algorithms could provide genuine access to *art* information in the face of prevailing institutional exclusivity. Nevertheless, algorithmic support and the organization of *art* information require accurate labeling and digitization "to really access the breadth of artistic heritage" (E3), another challenge faced by professionals within cultural institutions (E5, E1, E12, E8).

Our participants and HCI researchers agree that proper information categorization will improve *art* experiences. However, effectively communicating this information is essential to avoid overwhelming audiences with unstructured data, as may be the case in social media [188].

4.3.2.2 Conversation

Our pre-study participants reiterated the importance of conversation and communication regarding the depth of additional information available for each piece of *art*. According to the participants, having facts about an artwork is essential to understanding it, but leaving room for discussions and reflections is key to appreciating the artwork. One participant noted a struggle with "how can we let those opinions coexist along [with] the known facts" (E7).

This viewpoint finds support in other HCI literature, reinforcing our participants' argument for the necessity of solutions that integrate both the depth of factual information and the diverse perspectives of various stakeholders [369]. Participants reflected on how the depth of information and conversations may shape and inform the interpretations and discussions among audience members (E4, E8, E9, A2-6, P(14/37)). As they discussed the benefits of communication and conversations for audiences, they also considered how museums and other cultural institutions need to adapt their communication styles for more effective information dissemination (E6, E9, E12) [289, 115, 99].

According to both the literature and our participants, effective conversations about *art* hinge on respecting and contextualizing the diverse perspectives of various stakeholders in relation to a single piece. However, these conversations rarely occur in isolation, underscoring the importance of the third attribute: connection.

4.3.2.3 Connection

The significance of establishing the context and network power of a piece of *art* was a recurring theme among our curators, artists, and other participants. They emphasized that a work of *art* connects to other pieces and fosters connections among individuals. Participants wanted to leverage technology to make these connections more evident to audiences (E6, P(10/37)).

This desire for technological support aligns with existing literature that underscores the role of *art* and creative practices in sensemaking and identity formation [301, 74]. Using art appreciation as a means to reflect on themselves and society allows audiences to develop their own interpretations of their lived experiences. Participants noted a growing trend of audience members explicitly seeking *art* in which they can see themselves (E9, E12, E6, A4, A1, A3).

As audiences discover *art* and artists they relate to, they cultivate new experiences and relationships with *art*. HCI research also highlights that *art* experiences are inherently social, as access to *art* is facilitated through connections with and responses to other audience members [188]. Our participants support this claim, recognizing that audiences can establish their own artistic tastes and identities in relation to society's standards.

> "I think people are more comfortable walking in and looking for something that is appealing specifically to them, that meets their needs and their tastes rather than wanting to feel comfortable in recognizing something that society says is important or is part of the traditional cannon." - E10

The context within which audiences view *art* is considered vital by curators and other curators. A piece of *art* is rarely an isolated entity; it possesses social connections linking individual audience members and cultural connections to other artworks within the broader *art* world canon.

4.3.2.4 Time

The final attribute described by our participants as an essential aspect of viewing visual *art* is intentionally taking the time to do so. In many ways, the preceding attributes—depth, conversation, and connection—rely on the audience dedicating a certain amount of time to engage with the *art*. As our participants put it, discerning "good" *art* necessitates the ability of the audience member to pause and truly observe (E6, E11, E1, A4, A5, P(25/37)).

Unfortunately, time to stop and look is not inherently built into the technical infrastructure of short-form video *content* social media platforms. Critical media literature has argued how these platforms streamline production and engagement [374, 50]. Each piece of *content* on these platforms swiftly leads to another by encouraging viewers to remix or recreate existing trends or automatically transition to the next video in the endless scroll [83]. Well-designed digital tools that enable users to view and explore *art* in a manner that incorporates the previous key elements instead of scrolling through an endless array of *content* can enhance and support the user experience of appreciating *art* (A1, A6, P(16/37)).

One participant exemplified the potential of digital tools by referencing Jason Farago's detailed *art* analysis for The New York Times [110], stating, "But if you sort of spend the time as he does to take that image and to really use digital tools to move beyond just the representation. . . then you can really do something positive and get people engaged" (E5). According to the participants, there is value in these tools to positively impact digital experiences of *art* appreciation. The challenge, however, lies in consistently making these tools accessible to wider audiences (E8).

With an environment rife with changes, there exists an opportunity for HCI researchers to glean valuable insights directly from curators on how digital experiences can be adapted to incorporate these four attributes and enhance the appreciation of *art*.

4.4 RQ2: How would curators alter the social media experience

4.4.1 Main Study Part 1: Methods

4.4.1.1 Participant Recruitment

A selection of curators, gallerists, and art collection specialists (n=13) were contacted for co-design workshops, primarily recruited through direct email networking as in Section 4.3.1.1. Some participants (n=6) from Section 4.3.1.1 were approached again, adhering to ethics approval, considering the time gap between the Pre-Study and Main Study and the distinct methods used. As the focus of this project explores separating *art* from *content* on SFSM, our curators can offer their expertise at selecting and presenting *art* to audiences both in traditional gallery settings and online for their digital collections and social media. The very nature of their work as curators involves reviewing and presenting art to audiences.

The collaborative co-design sessions provided a platform for multiple experts to work and discuss. Participants were placed into the co-design sessions based on availability and to maximize the diversity of art specialties and institution structures. Each session had curators and gallerists representing contemporary and ancient visual arts displayed in traditional museums and commercial galleries. This participant distribution ensured that the discussions included various perspectives on how algorithmic experiences could support art appreciation. We compensated participants by offering co-authorship on publicly released open-source materials [346], deemed more appropriate compensation than e-gift cards by our institution's ethics committee.

4.4.1.2 Study Design

Participants were divided into three co-design groups. Co-design workshops allow HCI researchers to glean insights into participants' varied perspectives on the same prompts and collaborate to generate artifacts and research findings [178, 314]. For this project, the researchers observed how curators with extensive experience in analog, physical *art* experiences approached the research prompt of redesigning digital *art* experiences. Each 2-hour session involved two researchers, one as a facilitator and the other as a note-taker, and comprised four activities with brief discussions following each.

The first activity involved participants producing and ranking behaviors within daily curatorial tasks inspired by radar design activities in workshopping literature [197, 148]. Participants acclimated to the whiteboarding tool and expressed their daily challenges. The second activity had participants map potential solutions for online platforms to each of the four design implications from Section 4.3.2, drawing from creativity matrix activities [196, 159]. The third activity had participants develop "Do's" and "Don'ts" based on previous exercises to differentiate artworks from the generic *content* label. The final workshop exercise instructs participants to translate the "Do's" and "Don'ts" of exercise three into UI elements on low-fidelity wireframes [215]. For example, "Do provide insight into the creation process" became a video UI element of the artist making the work. Similarly, "Do support the opportunity for additional learning" became linked text elements that users could click on to learn more. Due to some participants' unfamiliarity with the online whiteboard tool, researchers acted as hands, creating wireframes based on expert discussions. All activities and the participants' responses are in [345].

Before the workshops with participants, researchers conducted a cognitive walkthrough and a pilot workshop. A cognitive walkthrough involves an expert's external, informed review of design decisions' potential impact on the user experience [353, 359]. A gallerist with expertise in private and national galleries conducted the cognitive walkthrough. The pilot workshop, involving HCI researchers within the group, focused on practical aspects of workshop design, the whiteboard tool, and exercises. Final adjustments were made based on feedback from these preparatory sessions before the co-design workshops.

4.4.1.3 Analysis

The analysis of the co-design materials was a cross-examination of the outputs by the workshop groups. Following the co-design sessions, the participants received the designs created in the sessions they did not attend and a critique form they had to complete (refer back to Method 2 in Figure 4.1). Once the participants returned the critiques, the researchers conducted a "walk the wall" activity [160, 67], placing critiques on appropriate areas of the low-fidelity prototypes. This method, standard in design sessions, directly incorporates critique and feedback to the designs [160, 67]. Iterating based on critiques and initial workshop outputs, researchers produced a final set of interfaces. This approach allows expert participants and researchers to make parallel prototypes and rapidly iterate over potential designs before testing and validation, avoiding commitment to building an expensive system.

4.4.2 Main Study Part 1: Co-Design Outputs

Building upon the four attributes of depth, conversation, connection, and time derived from the transcripts of the pre-study, the co-design sessions integrated these attributes into the participatory design exercises involving curators.

4.4.2.1 Screens

Each co-design workshop yielded a low-fidelity prototype collaboratively designed by the participants and visualized by the researchers. While participants frequently discussed the limitations of experiencing *art* on SFSM platforms, the outcomes of the prototypes were more focused on creating early designs for a distinct digital experience explicitly tailored to *art*, diverging from the concept of redesigning short-form video *content* on these platforms. The resulting screens do not reflect the usual features of SFSM, a looping video, and an endless scroll; instead, these are new digital prototypes, a testament to the curators' shared perspective on how art should be displayed digitally, discussed in detail in Section 4.6.2.

Screens 1 The curators of this workshop were motivated to elevate proper formal curatorial data and information related to a piece of *art* (Figure 4.3), aligning closely with the depth attribute of the pre-study findings. This is evident in the heavy emphasis on the text sections. Another important design feature the participants included was access to external resources and links to other materials (see the

Figure 4.3: An example screen designed by the first workshop cohort.

underlined artist name and artist process video). Finally, there was a desire to connect back to the physical exhibition space with a link to get tickets.

Screens 2 This workshop group similarly prioritized the depth attribute, providing users with a wealth of information (Figure 4.4). Furthermore, they took a distinct approach by explicitly describing and implementing a specific type of user interaction: sliders. Like recent research on recommendation system designs seen at CHI 2023, sliders and dials offer users greater agency in their *art* experience [117]. Participants in this workshop described how sliders could enable users to select the *art* they want to learn more about and take action to pause exploration and delve deeper into the details. Through this interaction, there are connections back to the prestudy's depth and time attributes.

Screens 3 During the workshop session, it became clear that members of this group had more social media experience, which may have influenced their design choices. During the prototype design, they focused on fostering connection by centering a video exploration of the artwork as it stands in the exhibit (Figure 4.5), connecting and situating the artwork to the broader museum or gallery collection. Additional interface features such as relevant cultural music, multiple language

Figure 4.4: An example screen designed by the second workshop cohort.

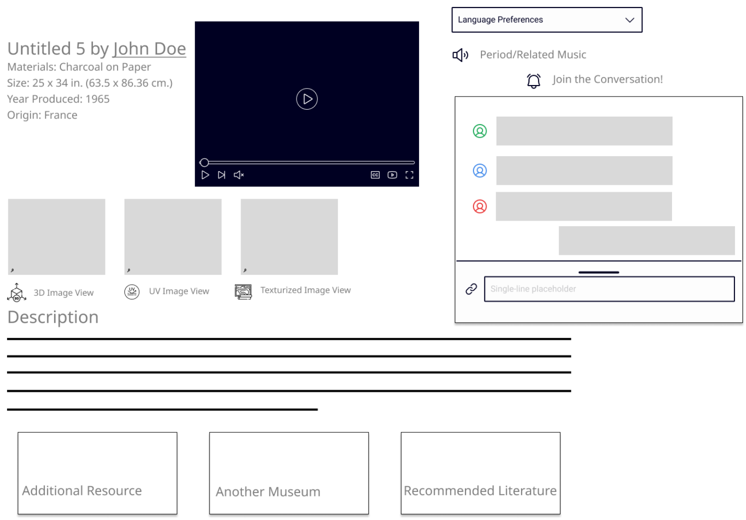

Figure 4.5: An example screen designed by the third workshop cohort.

settings, and a general chatbox also open up paths for users to connect with each other and the *art* satisfying the conversation characteristic. These features explicitly promote discussion and a social experience of *art* appreciation.

4.4.2.2 Similarities, Critiques, and Iterations

As outlined within the methods, the outputs of the co-design workshops were cross-examined by the participants. Based on their critiques, a "walk the wall" activity [160, 67] allowed the researchers to reflect on the interface features broadly accepted across the participant groups. Immediately, the participants noted the similarities they valued. These included making the artwork central to the experience and supporting it with a depth of factual information. Similarly, the differences were apparent to our participants, especially in the possible modes of interaction. One prototype involved sliders allowing users to customize their experience, while another featured a dedicated comment section for conversations about the artwork.

Common feedback regarding the prototypes centered on the desire for more informal language in the text section labels. The current word choice was considered too formal for the general user interested in exploring *art*. Furthermore, many comments emphasized prioritizing visual components over text sections. That being said, the more informational background artists and creatives offered regarding their individual pieces, the more likely they were to stand out from other *content* typically found on social media. The users' ability to control their experience was also highly rated.

Based on these critiques, the researchers worked to develop a set of medium fidelity screens (see Figure 4.6 for one such screen). These screens are not complete prototypes but rather early designs functioning as vignettes representing the feedback obtained from curators, taken to a more advanced level within the methodological approach of research through design [372]. While these screens are to be reviewed and validated by the creatives in the second part of the study, they aim to meet some of the design implications and recommendations outlined in HCI literature, which will be discussed in Section 4.6.2.

4.5 RQ3: Reviewing designs with content creatives

4.5.1 Main Study Part 2: Methods

4.5.1.1 Participant Recruitment

We recruited ten professional content creators, aged 18 and above, with established social media accounts on platforms like Instagram, TikTok, or YouTube. During initial recruitment, we considered the alignment between the creative's social media

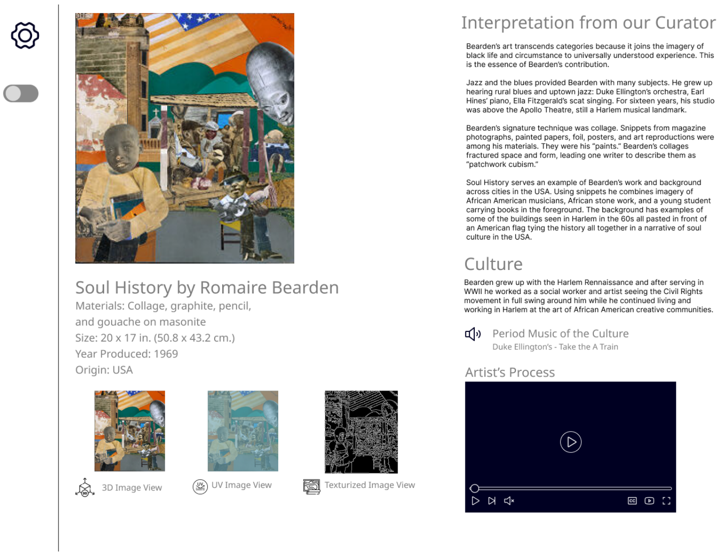

Figure 4.6: A higher fidelity screen shown to the creatives for their review, image and text drawn from [315, 53, 2].

presence and the project's definition of visual arts. However, the ultimate sample relied on creatives whose agents and business accounts responded to our messages. While they communicated with the researchers in English, they are a global sample working in various cultures and languages. To reiterate, these participants are referred to as "creatives" in this project, aligning with other HCI research to separate the participants from the word *content* [74, 297]. The selected reviewers, detailed in Table 4.3, will be referenced by their ID code in this project. Their account names and profile links were included as advertising compensation, as our institutional ethics committee advised, in exchange for their participation.

4.5.1.2 Study Design

We conducted a stakeholder review with social media creators (see [345]). This review, akin to a user test, goes beyond usability to gather perspectives from those impacted by the research [131]. The review aimed to assess if the curators' and artists' perspectives and assumptions offered creatives an effective tool and approach to elevate *art* from the sea of *content*. The first 30 minutes focused on the results of Section 4.3.2, exploring overlaps and differences between curators' key characteristics of *art* (Figure 4.2) and the reality for modern social media creatives.

Table 4.3: Details regarding the creatives we recruited and interviewed for the final part of the study.

ID	Account Name	Social Media	Subject Area	Subscriber/ Follower
C1	Alice Cappelle	YouTube/Instagram	Video Essays	319k/25.3k
C2	The Canvas	YouTube/Instagram	Art Critique	441k/3k
C3	Alpay Efe	YouTube/Instagram	Painting	658k/466k
C4	Eirik Arnesen	YouTube/Instagram	Sculpture	29.6k/98.6k
C5	J.R.R. Talkin'	YouTube	Television Critique	4.86k
C6	Uncomfy	YouTube/Instagram/ TikTok	Art and Creativity	128k/109k
C7	Shuen Art	YouTube/Instagram/ TikTok	Digital Art	33.4k/90.7k
C8	danielsonwilliams_	YouTube/Instagram	Music/Cooking/ Cannabis	27.9k/2.4M
C9	NIRO	YouTube/Instagram	Digital Art	82.8k/3.5k
C10	Kenn Yap	YouTube/Instagram	Concept Art	200k/37.8k

During this stage of the stakeholder review, we presented the participants with Likert Scales reflecting synonyms and antonyms of the curators' key characteristics. Participants rated their own social media posts on this scale. For example, "Are your posts meant to be delivered faster or slower?" to better understand their current position on the Time characteristic, one of the four attributes derived from the pre-study in Section 4.3.2.

The latter half centered on reviewing prototypes from Section 4.4.2. Creatives interacted with the prototypes, providing immediate feedback while researchers collected responses and posed relevant questions. This aimed to understand their perspective on a new digital experience within platformed social media and short-form video *content*.

4.5.1.3 Analysis

The main outputs were interview transcripts, analyzed using the affinity diagramming system described in Section 4.3.1.3. Participants also recorded value selections on a Likert scale, with the researchers plotting the numbers on a chart. Reflections on creatives' feedback regarding the redesigns were derived from the affinity diagram ideas, concepts, and the Likert scale chart.

4.5.2 Main Study Part 2: Stakeholder Review

To validate and contextualize the work done in collaboration with our curators, we presented the workshop outputs to creatives who were professionally established on social media platforms. Through the stakeholder review, we can present their thoughts on the values and screens.

4.5.2.1 Values and Characteristics Review

The stakeholder review began with a discussion of the findings from Section 4.3.2. Our participants outlined the struggle between their perceived cultural values, such as creativity, discourse, and uniqueness, and the values imposed by the incentive structure of social media platforms (C1-3, C5-7). Their responses highlighted a disconnect between the intrinsic values gained by the creatives and their audiences (such as ideas, self-expression, and cultural understanding) and the financial values of the social media platform incentivized by advertising revenue. With our creatives already contemplating the balance between artistic and financial incentives, we transitioned to the next phase of the review, where we presented them with opposing terms that were synonyms and antonyms of the four attributes developed by the curators.

From the Likert scale section (Figure 4.7), we identified the attributes our creatives prefer in their own creative work posted to social media. Our creatives consistently value fast delivery of their creative *content* (40%). This need for speed was associated with producing *content* rapidly to stay relevant and avoid falling victim to "the algorithm" (C1, C4, C6). Similarly, the creatives felt the need to provide more pop (i.e., pop *art* or popular *content*) and avoid information that is quite exclusive (40%). While they enjoyed making more exclusive *content*, those particular posts tended to underperform (C1-2). 50% of participants preferred their *content* being algorithmically recommended rather than organically discovered via word of mouth. All participants distinguished between the two, with word-of-mouth audiences typically resembling their existing audience, while recommended audiences represented new growth and expansion opportunities.

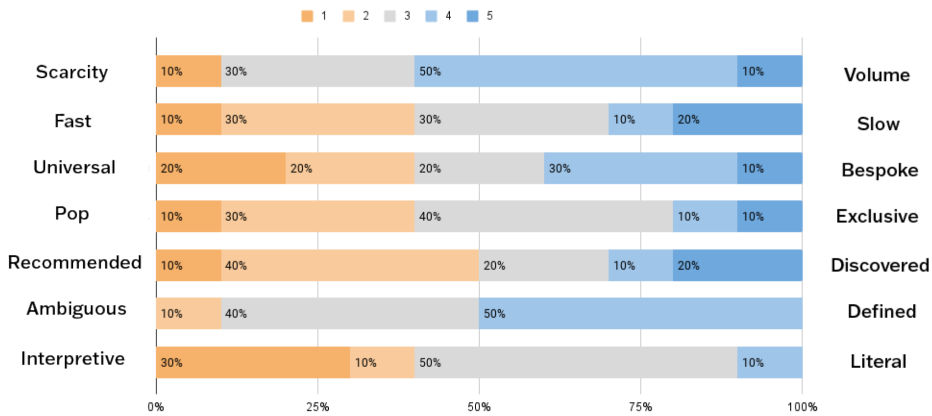

Figure 4.7: Results of the Likert Scale questions posed during the stakeholder review. A selection of 1 is closer to the value on the left, while a 5 is closer to the value on the right. For example, 60% of our creators prefer volume over scarcity.

For the contrasting values, the creatives' opinions diverged on whether their work should cater to a universal audience or remain tailored to a specific population. Those who selected closer to universal said it was part of their creative "philosophy" (C3) to make the work as open and available to anyone as possible. Those who selected closer to bespoke said that over time, they got comfortable with their own "voice" (C1) and creative work, leading to a dedicated "community" around said work.

The final two value pairs of the Likert scale offered an interesting juxtaposition of ideas. While 50% of our creatives preferred to create work well-defined within their own creative vision and aligned with their agenda, 40% preferred their audience to perceive the work as interpretable rather than literal. Our participants noted a balance between defined creative work made to target trends in the algorithm while still allowing room for audience interpretation (C2, C4-10).

Our creatives were apprehensive when asked to reflect on the characteristics of *art* experiences drawn from the curators and artists. They acknowledged the importance of the characteristics but highlighted the challenges they face in adhering to them (C1-10).

> "Social media doesn't care about these at all, no. It's the people that care. That's the sad part." - C7

The responses indicate that our previous participants' responses are commendable but aspirational within the current social media landscape and the incentive structures driving creative work. The two characteristics that posed the greatest concern were "conversation" and "time." According to our participants, the structure of social media platforms does not foster these values. In fact, they argued that there is no expectation of meaningful conversations or time while online. Creatives argue that while social media platforms claim to connect people and offer comment sections that allow users to interact with the creatives, these are never truly seen as meaningful conversations (C2, C4, C9). At times, these comments are never even expecting a response. Similarly, creatives argue that the popularity of short-form video formats in social media does not afford users time to engage deeply with creative work (C2-4, C7-10).

4.5.2.2 Screen Review

When the creatives were presented with the screens as vignettes, offering a glimpse of a potential future interface aligned with the identified characteristics, their feedback primarily reflected a mix of excitement and apprehension. While enthusiastic, they expressed concerns that audiences might not share the same level of interest and may not engage with the envisioned experience. The participants felt that the screens exhibited an extreme approach to identifying and supporting creative work (C4-8). Treating creative works like fine *art*, as recognized by our curators, and prioritizing the attributes of depth, conversation, connection, and time could alienate casual viewers exploring social media (C3, C6). Balancing our curators' four key values with the platforms' realities will be a notable challenge in future design work.

Several participants noted these vignettes prompted reflection on the challenges of cultivating two types of online audiences within social media (C3-4, C6-8). One audience seeks and appreciates "long-form content," while the other consumes and shares "short-form content," almost akin to broadcasted advertisements. The screens from the co-design and curators would interest the dedicated long-form audience, applying the characteristics and offering experiences that would elevate the creative work and have a meaningful impact. However, creatives faced the dilemma of surviving in the short-form *content*-dominated landscape that has become prevalent across all social media platforms (C1, C4, C10). Recognizing this pain point, we prompted them to articulate their needs and aspirations for the digital environment.

When asked to reimagine the digital environment in alignment with their values and based on their recent exposure during the review, the creatives proposed two approaches. First, they wanted transparency on how the goals and aspirations of creatives can extend beyond merely meeting engagement metric targets. Second, they want a platform that does not succumb to the prevailing trend of short-form *content*, allowing them to grow and develop a community around "valuable creative work" (C2-5).

Recognizing the challenge of presenting creative works on a platform primarily funded by advertiser revenue, our creatives acknowledge the tension between the aspirational values of the "art world" [1] and the current economic realities of contemporary creative livelihoods. Overall, there is concern among our participants that re-configuring social media is challenging as it feels too big to fail, often being compared with the concept of capitalism (C2). Therefore, any designer or engineer seeking to address this issue should be mindful of how non-financial and capitalistic values can coexist to potentially formulate a solution. Creatives found the endless scroll and focus on copying and pasting trending sounds to be particularly against their own creative philosophies and workflows (C3, C7). The short-form video *content* format popularized by TikTok has been copied by other platforms, which concerns our participants.

As our creatives note, their preferred methods of social media presence are being shaped by a growing "homogeneity" in the digital space where every major platform converges on the most lucrative advertising setup possible (C1, C4). The stakeholder review results indicated that approaching a social media redesign with the aim of elevating *art* above mere *content* directly challenges the financial structure of platforms and their relationship with creative professionals. From here, researchers can consider how to pair alternative design methods with research through design to explore potential prototypes that further challenge the financial power structure dictating the creative works shared on social media.

4.6 Discussion

Our findings address the research questions posed in this study. Participants highlighted depth, connection, conversation, and time as crucial attributes distinguishing *art* from *content* (**RQ1**). These align with Slow Technology principles, advocating for a bottom-up approach and integrating Slow Technology in the context of *art* on social media, which we discuss further in Section 4.6.1.

Addressing how these attributes could redesign social media with insights from curators (**RQ2**), workshops yielded prototypes for dedicated *art* experiences. Section 4.6.2 reviews the cross-examination and final designs, referencing related HCI work on creative *content* displayed on social media.

To evaluate the design materials, we presented them to creatives (**RQ3**). Section 4.6.3 explores how creatives' comments align with the Pre-Study feedback, highlighting concerns about algorithmically powered platforms as exploitative for creative work. Ultimately, the first two study parts and research questions provide us with attributes and designs that, when evaluated, reveal how approaches to separating *art* from *content* hold up within the current reality of platform power.

4.6.1 Users Demand Slow Art Experiences

Previous work in HCI has explored the potential for Slow Technology to disrupt the user experience in various situations. The umbrella term of Slow Technology refers to an approach to interaction design that looks beyond the immediate goal of efficient instantaneous technological experiences [144]. While originally focused on temporal digital experiences, the design method has expanded to include other values that allow users to engage with technology in novel ways (see [190, 261, 262] for discussions on other values that society uses to define time). Researchers have developed various experimental prototypes through this research approach and placed them in front of users [243]. These prototypes manipulate the user's previously instantaneous and efficient experiences by injecting slowness and reflection into the tasks of printing their digital photos or listening to the radio, to name a few.

The design approach of Slow Technology rests on three primary facets that were set out in the original work [244, 144]. The three facets focused on (1) reflective technology, which promotes both reflection on the task and on the role of technology overall; (2) time technology, which challenges users' expectations of time; and (3) amplified environments, which work to elevate technology and tasks beyond just meaningless background occurrences [144]. While these facets of Slow Technology have been critiqued, Odom et al. have shown the potential for these facets to be applied in designing new experiences [244]. Specifically, they expanded the facets with their own eight qualities of Slow Technology: implicit slowness, explicit slowness, ongoingness, temporal drift, pre-interaction, temporal modality, temporal interconnectedness, and temporal granularity [244]. As Asadi et al. point out, "current trends of slow technology are more design-led than intended user-led" [21]. In other words, while traditional Slow Technology involves design requirements that are dictated by the researchers and tested on users, we set out to conduct a user centered design approach. Through interviews and co-designs our participants shared their desires for design requirements that altered how *art* could be experienced in digital spaces. After reflecting on the design requirements and details presented by our participants, we noted the similarities between their specifications and the principles of Slow Technology. Therefore, we provide an example of a user-driven, bottom-up approach to Slow Technology design. Rather than the researchers designing an experience crafted to reflect the goals and qualities of Slow Technology, we argue that participants should direct the designs and critiques of the technology being studied.

Based on the results of our user-centered research including the attributes our participants described as essential for art experiences: (1) depth of information, (2) space for meaningful audience conversations, (3) audience-artwork connections, and (4) time for the experience are essential characteristics of experiencing *art*, we argue that these recommendations follow the established design philosophy of Slow Technology.

The identified characteristics align with qualities promoting implicit slowness and temporal interconnectedness [244], contributing to reflective technology. This empowers users to reflect on the encountered *content* and discern whether it qualifies as *art* or not [244, 144]. However, current social media structures do not naturally foster these attributes or the *art* they represent, necessitating a shift in the user experience.

Our creative participants endorse the value of these characteristics, emphasizing challenges tied to existing power dynamics and financial incentives in contemporary social media platforms. They particularly underscore the importance of time, seeking to cultivate long-form communities (such as Twitch [356, 114] or Podcasts [140, 214]) that will meaningfully engage with their creative work along the four characteristics our curators' described. Our findings align with results from Intangible Cultural Heritage (ICH) scholars showing the value of vlogging and live-streaming for cultural preservation and dissemination [195, 69]. Ultimately, our participants advocate challenging the prevailing trend of short-form *content* favored by current platforms.

Given the centrality of time in their experience and feedback, any redesign must prioritize the temporal aspects of the digital creative experience. This aligns with the second facet of time technology, aiming to enable variations in users' temporal experiences and alleviate the need for constant interactions [244, 144].

Combining insights from our co-design participants with feedback from creatives, we urge future work to establish "amplified environments" [244, 144] to ensure the generic label of *content* does not consume individual creative works, even if that results in designs that do not reflect expectations of SFSM. We advocate extending this approach to other scenarios combating the universal label of *content*, using user research and prototypes to develop effective interventions aligned with Slow Technology values.

4.6.2 Designing for the Digital Presentation of Art

Section 4.4.2 leveraged feedback from Chapter 4.3.2 to engage curators in reimagining an *art*-focused social media platform (**RQ2**). While numerous digital spaces market themselves as bastions of creativity and *art*, they can still learn from the perspectives of the creatives and curators. this project unveils outcomes from three co-design sessions reshaping online *art* encounters by rejecting the established design features of SFSM, creating a wholly new digital experience.

Drawing on design recommendations from recent HCI research, outlined in Section 2.3.2.1, we can reflect on the design outputs of our co-designs. Researchers argue the experience of social media algorithms diverts the creatives' focus from their actual creative work towards gaming the algorithm [74]. Therefore, the curators designed the screens to put creative work as the primary focal point. The figures illustrate this focus (Figures 4.3, 4.4, & 4.5), placing artwork images at the center with information tailored to each piece.

Our curators urged creatives to highlight the labor, effort, and process that went into a single piece of creative work on the screens. The final screens' information layout reflects our curators' characteristics. These layouts aim to respond to the call for digital spaces to embody "post-capitalistic values" [297]. Sharma et al. and our experts argue that artists are expected to be marketers, social media influencers, and data analysts in addition to their actual work. The screens, focusing on one artwork at a time and showcasing details about the creative process, intend to make the invisible labor more visible [297].

The co-design screens also align with three recommendations by Simpson and Semaan: Uncoupling metrics from creative success, more malleable structure, and limiting objectification [301]. These screens eliminate traditional engagement metrics, offer diverse formats for presenting *art*, and prioritize the depth of a single piece over a demanding posting schedule (Figure 4.6). By addressing Simpson and Semaan's recommendations, the screens allow artists to focus on developing their work and avoid becoming subservient to a grueling posting schedule.

Reviewing the co-design screens through the lens of previous HCI research underscores the necessity for digital *art* spaces to challenge social media's AX norms to create new digital environments [70]. Historically, *art* delivery has taken varied, non-formal routes via magazines and collectives to reject institutional control of museums and academies [129, 307, 98, 270]. Therefore, to foster community and connection, designers should expand on the semi-flexible architectures of our curators' screens that empower artists rather than mimicking traditional social media or museum websites.

While finalizing the immediate interface or algorithmic experience, it is increasingly pertinent to question how our digital spaces and algorithmic experiences shape our ideas and perceptions based on the information they present.

4.6.3 The Patron Power of Social Media

We have established four attributes differentiating *art* from *content* (Section 4.3.2) and presented designs in Section 4.4.2, calling for redesigning social media from curators and artists. Section 4.5.2 offers a reality check from creatives who acknowledge subscribing to the values and ideas of our earlier participants (**RQ3**). However, they contend that social media platforms are not easily transformed into objective exhibition spaces and no longer function as the cultural "inter/infomediaries" [52, 228] previously described.

Feedback from both curators and creatives suggests that their interactions with recommendation algorithms and social media platforms extend beyond mere intermediation or infomediation. Instead, it resembles the traditional patron system that historically shaped the art world's evolution. Initially, this system enabled influential individuals and organizations to showcase themselves through artist support [18, 28, 141]. With streaming and patronage sites, anyone can now assume the role of a patron [18, 356]. However, according to creatives, they must appease the "algorithm" to access their intended audience, departing from the traditional human patronage system promoting creative works.

This dynamic introduces the potential for the exploitative nature of creative work on social media platforms. Historically, *art* and creative work reflected the power and influence of the patron, creating a symbiotic relationship where each party supported and provided for the other [28, 18, 141]. In today's system, social media platforms wield a monopoly on audiences [100], allowing them to exploit creatives for *content* and ad revenue without treating their creative works with respect or care.

Both artists and curators advocate for change and have articulated their aspirations for reforming these systems. Curator E5 notes in Section 4.3.2 that digital tools can significantly reshape the audience-artwork relationship but require modification beyond shallow post representation. Similarly, creatives in Section 4.5.2 express frustration with the prevailing short-form *content* focus on social media platforms despite the financial viability (C8) and intrinsic value (C7) of long-form *content*. With both groups of participants expressing frustration with the status quo, there is a growing need for serious considerations about the future of *art* on social media.

The consensus among artists, curators, and creatives is that the current system for posting *art* on social media is ineffective and potentially exploitative, aligning with existing literature [301, 118, 273]. Where other researchers from Chapter 2.3.2.1 may call for a platform built on Marxist theory, we argue that our participants seek a platform built on curatorial practices. Participants highlight the need to reflect on different values and incentives that better promote creative work rather than optimizing for opaque engagement metrics that only support the platforms' ad revenue.

4.7 Limitations and Future Work

This project presents qualitative findings from diverse stakeholders regarding elevating *art* from the all-encompassing label of *content*. Despite valuable findings, it is crucial to acknowledge limitations and potential biases in this qualitative study.

Firstly, most sampled curators primarily represent the Global North and are heavily affiliated with academic museums and institutions, potentially framing their perspectives and ideas. Secondly, during the stakeholder review with creatives, participants were chosen based on their established professional social media presence, raising the possibility that their dependencies on these platforms may influence their viewpoints.

Despite these limitations, the conversations and design materials establish a foundation for future research, offering valuable feedback for our team and others in the field. Work beyond this thesis can revisit **RQ3** with general social media users in place of content creators to get their thoughts on the proposed designs. We hypothesize that general users may have more mixed reactions to altering the algorithmic experience. Possibly, only a subset of users would support an experience that drastically slows down the information they receive. Pulling these reactions apart will need to be the focus of future research beyond this thesis. Another study exploring how a TikTok dance video is received in a gallery space could reveal insights into perceiving such *content* as high *art*. Similarly, exploring methods to distinguish journalism or community activism from the vast pool of social media *content* is crucial. Finally, addressing major platforms' tendency to categorize all of human expression as *content* remains a key focus.

4.8 Conclusion

This project seeks to challenge the social media experience by elevating *art* from the ubiquitous concept of *content*, which has become a catch-all term for all forms of information on these platforms. In this pursuit, we presented the findings of a pre-study and main study that illuminated four crucial characteristics that artists and curators consider essential for experiencing *art*: depth, conversation, connection, and time. Additionally, we present a collection of open-source screens. Both the characteristics and screens were reviewed by content creators and deemed desirable yet aspirational within the current ecosystem of advertisement revenue-driven social media. This project provides insight into how certain audiences perceive art in algorithmically controlled spaces, setting up the question of how an algorithm interprets art to begin with, which we explore in the next project.

Much of this chapter was also reviewed and published in the paper:
von Davier TȘ, Herman LM, Moruzzi C. A Machine Walks into an Exhibit: A Technical Analysis of Art Curation. Arts. 2024; 13(5):138. https://doi.org/10.3390/arts13050138

5
Machines as Art Audiences

Contents

5.1 Introduction

Before audiences encounter art online, an algorithm needs to interpret the information sent to the platform. To understand the relationship between digital spaces and art appreciation it is necessary to understand how an algorithm would interpret an artwork. This chapter mirrors the previous one by asking how an algorithm would interpret an art exhibit rather than how people would interpret art seen on social media.

The impact of Artificial Intelligence (AI) and algorithms on the art world is rife with controversy. One side argues these technological developments are the end of the arts as we know it, accompanied by scandalous revelations of AI images winning art contests [279, 134]. On the other hand, AI equips artists and designers with tools to elevate their craft to achieve new forms of creative expression [8, 61]. At the heart of the controversy is the question of perception: what exactly do audiences think when they encounter algorithmic decision-making in the arts? The fact that algorithms powering social media are deciding what art gets shown to people based on proprietary, obscure metrics means they are filling the role of curators in our daily lives.

Herman and Moruzzi set out to explore what audiences think about curation done by an algorithm compared to that done by a human [153]. Using practice-based research methods, they produced an exhibit, the *Algorithmic Pedestal*, which included a curated selection of images by Instagram's algorithm and the London-based artist Fabienne Hess, taken from the Metropolitan Museum of Art's Open Access collection. Visitors then participated in surveys, interviews, and observations [153]. Their work builds on research that has examined users' perceptions of AI-created work where people considered work done by AI to be less creative and less interesting [161, 180, 265, 263].

Based on the apparent differences reported by the participants, we were motivated to see if computer vision also noted differences in algorithmic and human curation. Using fundamental computer vision code informed by the field of computational aesthetics [156], we conducted a technical analysis of the *Algorithmic Pedestal.* A computer vision analysis of an image allows an algorithm to convert the data stored in individual pixels into "handcrafted features" [366], meaningful, comparative metrics. As there are currently no universally accepted quantitative metrics for artistic quality or creativity [230], our research has the opportunity to explore if the computer's metrics align with audiences' opinions.

Using the technical analysis of the exhibit, we explore two research questions:

- **RQ1:** What differences does the computer vision code identify between the curations and the overall open-access collection?

- **RQ2:** How are these findings different from the reported exhibit visitor responses?

Our results reveal a handful of statistically significant differences between the exhibit and the overall Metropolitan Museum of Art collection. Specifically, the algorithmic curation had a higher ratio of unique colors, and the human curation had a higher face count. However, the computer did not see any statistically significant differences when comparing the two curations. Our discussion of these results contributes to our thoughts on how the curation process is more than quantifying artistic qualities. Furthermore, we argue that algorithmic curation needs a human-in-the-loop to account for the differences between audience and machine perception of artworks.

Ultimately, this work highlights how the metrics drawn from computer vision notably differ from the reported observations stated by the exhibit-goers. These differences highlight opportunities for designers, developers, and art scholars to consider how computation will partner with human perspectives in a world of growing collaboration within the arts.

5.2 Materials and Methods

The research outlined within this project reflects on the findings of previous work on the impact of algorithmic curation. In 2023, the *Algorithmic Pedestal* was an exhibit-based research project in London, focused on comparing human and machine curation [153]. The researchers conducted a qualitative user study on the exhibition visitors who attended the exhibit space over seven days. The qualitative study involved interviews, surveys, and ethnographic observations, which aligns with other research that has explored audiences' perceptions of recommendations made by AI [62, 79, 97, 194, 368]. The previous work presents what people believe goes into algorithmic curation and how it differs from the curation done by human art experts for human audiences. Therefore, there is a gap in which it is possible to question how a machine would perceive an art exhibit curated by a human artist and another machine.

The differences between machine and human perception have been under consideration for years [143, 341, 375]. Many use cases in medical and cognitive sciences have sought to compare and implement machine perception alongside human perception. In medical science, researchers saw that humans take in higher-level information from their medical training to inform their observations while machines look for patterns and deviations in the input data [206]. The researchers would then compare the performance of human and machine subjects on a perception task that often measures accuracy or correctness.

In cognitive science, the research argues that humans perceive auditory and visual stimuli differently than machines [186]. Their experiments expect a machine and a human to encounter a stimulus and transcribe or label the signal. The idea of comparing perception boils down to comparing the accuracy of the two observers. Researchers have pushed back on this experimental system, arguing that there is a risk of confounding understanding with recognition [120]. Other factors, especially on the human side, like bias and higher-level processing, make it an uneven comparison. Therefore, these researchers call for research to understand more nuanced differences between how machines and humans perceive information. They argue that it is more important to understand how exactly the two groups differ in their perceptions, a task this project explores.

In this project, the research addresses machine perception on its own. Building on previous research, we seek to gain perspectives on the machine's observed similarities and differences between human and algorithmic curation. In contrast to human perception, this project presents the metrics derived from applying computer vision to the exhibit pieces. This approach is comparable to art world projects that explored the interpretation of art through computers as exhibit pieces and exhibit critiques [336, 309]. We build on this by revealing the fundamental metric differences revealed by computer vision code processing artistic images.

Nine metrics were run to establish the metadata for our statistical analysis and comparative study (Table 5.1). Like other computer vision software popularized by the field of computational aesthetics (see [366, 170] for a detailed overview), this approach involved breaking down the image files into the data contained in each pixel of each file. A breakdown of pixel data for feature extraction refers to extracting "handcrafted features" for image analysis [366]. Through analysis of the individual pixels, the overall image could be built through averages and ratios of the pixel data. Once the overall image had a data point assigned, they were set into the specifically labeled data frames for comparison. In addition to capturing the metrics, we also wrote code to visualize the metric breakdown, offering insight into the process of pixel analysis. Section 5.3 provides some of these visuals.

5.2.0.1 Data Analysis

All of Table 5.1's metrics (aside from Dominant Color) are quantitative variables associated with each image file via a comprehensive pandas data frame [352]. To understand the potential quantitative differences between images of various groupings and selections, we visualized the metrics to check for a normal distribution among the samples taken for the exhibit from the overall sample of the images from the Metropolitan Museum of Art's Open Access collection (from here, "the Met collection"). With normalization established, we ran various independent *t*-tests to ascertain if differences in mean values for the various samples significantly differed from the original sample taken from the Met collection and uploaded to Instagram. With these *t*-tests, we compared the differences of all metrics in a variety of conditions:

- All exhibit pieces and the Met collection sample
- Instagram selected exhibit pieces and the Met collection sample
- Human (artist) selected exhibit pieces and the Met collection sample
- Instagram selected exhibit pieces, and the Human (artist) selected exhibit pieces

Table 5.1: This outlines the common computer vision metrics outlined by [162] that were applied to the images used for the exhibit.

CV Variable	Definition
Dominant Color	Returns the hex code and RGB hue of the most common color in the image file based on a pre-defined number of clusters.
Brightness	Average brightness of the image file pixels.
Ratio of Unique Colors	A ratio of the total number of unique colors in regards to the total number of pixels within the image file. A value of 1 indicates highly colorful while a value of 0 would be a greyscale image file.
Threshold Black Percentage	Using a pre-defined threshold value of 127, each pixel is compared to the threshold value and decided whether it falls above (white) or below (black) this provides a calculation for the ratio of black pixels in a greyscale or inverted image.
High Brightness Percentage	Ratio of pixels that have two times the average brightness of the overall image file compared to the total number of pixels.
Low Brightness Percentage	Ratio of pixels that have less than half of the average brightness of the overall image compared to the total number of pixels.
Corner Percentage	Harris Corner Detection [147] is used to identify corner pixels and then calculate what percentage of pixels within the image file register as corners.
Edge Percentage	Using Canny Edge Detection [60], we identify the edge pixels in the image file and calculate the percentage of these pixels compared to the overall file.
Face Count	Haar-cascade Face Detection [338] from OpenCV is a common basic form of face detection software; we apply it to the image file to see if there are any noticeable faces.

These comparisons reflect the conditions the gallerygoers also observed. Information about the human and Instagram-selected pieces was readily available within the physical gallery space through floor labels, and a presented tablet allowed visitors to explore the overall Met Collection from which the pieces were sampled [153]. Therefore, when attempting to understand how a machine would view the selected pieces, the statistical comparison tests replicated the same comparisons.

5.3 Results

Applying the collection of fundamental computer vision scripts to the images of the exhibit and the larger sample of pieces from the Met collection allows us to provide quantitative insights into the curatorial selections by both humans and machines. The adapted scripts from [162] allowed us to provide visual examples of the computer vision process and establish datasets for statistical comparison. Using the visual outputs, we can see the computer vision's interpretations of each set of artworks, thus giving us insights to answer **RQ1**.

As previous literature within Section 5.2 has shown, computer vision works by breaking down the individual data pieces stored within an image file into usable metrics that can be processed algorithmically in other scripts. The visuals show that these basic scripts are particularly accurate at the pixel level to identify details such as edges and corners (Figure 5.1)[1]. However, the scripts struggle with more complicated aggregate tasks, specifically facial recognition (Figure 5.2).

<u>Faces</u>. Haar-cascade facial recognition is a widely accepted and available tool offered through OpenCV. Nonetheless, even the original paper introducing the concept highlights the false positive percentage as 40% [338], leading to entertaining measurements such as this tapestry with a face count of 4 (Figure 5.2)[2]. This image gives us one prime example of how human and machine perceptions can

[1]Image retrieved from the Metropolitan Museum of Art's Open Access collection via Creative Commons licensing. The piece is "Visit to a Shrine at the Hour of the Ox (Ushi no toki mairi)," 1765.

[2]Image retrieved from the Metropolitan Museum of Art's Open Access collection via Creative Commons licensing. The piece is "Fragment of a Red-Ground Harshang Carpet," early 19th century.

Figure 5.1: The original Japanese print (left) next to the corner (center) and edge detection (right) visuals created by the software.

differ, leading to separate interpretations of a piece. Previous research has shown that images with faces tend to receive more engagement on Instagram [27], leading to a folk theory [173] that Instagram favors images with faces. It remains a folk theory as Instagram's algorithm is proprietary. However, our findings on the usage of fundamental computer vision software provide some insight into how and why certain images were selected. The computational analysis enabled us to "see like an algorithm," [330] thereby elucidating its decision-making process.

<u>Context</u>. Additionally, context can be quite essential for piece interpretation. See Figure 5.3, a sofa positioned against a blank background[3]. Many of us would describe the sofa as blue, white, and perhaps quite ornate, but when the computer script calculated the dominant color, it stated gray as the output. Unlike us, the script considers the whole image (including the blank grey-white background) when processing an image. Unless we invoke a higher complexity masking script to ignore the background and solely take the item in the focal point of the image, the computer will process all of the pixels and data equally, leading to yet another interpretation quite different from that of human visitors.

Once the data for each variable was processed, we were able to run comparative statistics between the different groups of images. We compared the metrics of

[3]Image retrieved from the Metropolitan Museum of Art's Open Access collection via Creative Commons licensing. The piece is "Sofa (part of a set)," circa 1835

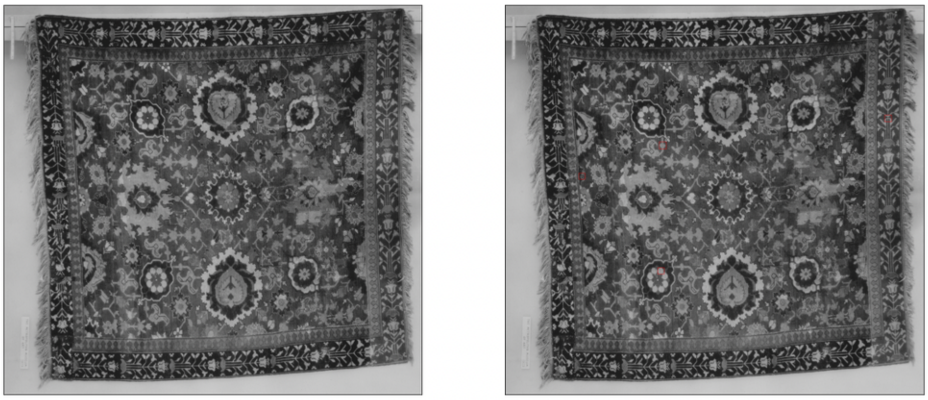

Figure 5.2: A comparison of the original rug image from the online collection, next to the code's output calculating four faces. Four red squares are placed around the rug where the pattern triggered the code to identify faces.

Figure 5.3: An image showing the sofa in front of a grey-white background. While the sofa is the focus of the image, the whole image file is analyzed by the computer vision software.

the human-selected, Instagram-selected, total-selected, and overall Met samples against each other to see if there were any statistically significant differences between the groups.

Beginning with the statistically significant results, we found that the Instagram-selected pieces had a higher ratio of unique colors than the overall sample of Met collection pieces used for the study (Table 5.2 t(1)=2.307, p=0.0334). The presence of a higher number of unique colors reflects visitors' and researchers' [204] expectations of Instagram's bias for bright colors. Simultaneously, the human-selected pieces had a higher number of faces than the overall sample of the Met collection (Table 5.3 t(1)= 5.238 p=2.835e-05). The presence of a significant number

Table 5.2: Results of the images from Instagram compared to the total selection from the Met collection.

CV Variable	Test Statistic	P-Value
Brightness	0.101384	0.92040
Ratio of Unique Colors	**-2.30718**	**0.033441**
Threshold Black Percentage	-0.07543	0.940718
High Brightness Percentage	1.38500	0.181199
Low Brightness Percentage	-0.35159	0.72933
Corner Percentage	0.47165	0.642551
Edge Percentage	-1.36619	0.189242
Face Count	0.062826	0.95060

Table 5.3: Results of the images selected by Fabienne Hess compared to the total selection from the Met collection.

CV Variable	Test Statistic	P-Value
Brightness	0.724514	0.48039
Ratio of Unique Colors	-1.32710	0.20561
Threshold Black Percentage	-0.32200	0.75211
High Brightness Percentage	-1.09492	0.291933
Low Brightness Percentage	-1.24428	0.233566
Corner Percentage	-0.71589	0.485799
Edge Percentage	-1.8290	0.08851
Face Count	**5.238054**	**2.835e-05**

of faces on the human-curated side also speaks to researchers' expectations of human and algorithmic processing. Previous research argues that images with faces received higher engagement on Instagram [27]. For some, these findings hint at Instagram's bias for faces. In contrast, for others, it suggests a psychological human bias for images that contain faces, a subject familiar and comfortable to us. In the case of the *Algorithmic Pedestal*, it appears that the artist chose more pieces with faces. Finally, we found that when grouped together, all of the pieces in the exhibit (both human and Instagram-selected) contained more unique colors and more edges than the overall sample from the Met collection (Table 5.4 $t(1)= -2.164$, $p=0.0378$ and $t(1)=-2.287$, $p=0.0286$, respectively). The other computational metrics comparing the image sets were not statistically significant.

When comparing the human-selected and Instagram-selected pieces (Table 5.5), the lack of statistically significant differences offered an intriguing insight into the

Table 5.4: Results of all the images used in the exhibition compared to the ones sampled from the Met collection.

CV Variable	Test Statistic	P-Value
Brightness	0.552951	0.583944
Ratio of Unique Colors	**-2.16496**	**0.037811**
Threshold Black Percentage	-0.28687	0.775964
High Brightness Percentage	-0.763722	0.45049
Low Brightness Percentage	-1.19127	0.241977
Corner Percentage	-0.62744	0.534768
Edge Percentage	**-2.28754**	**0.028678**
Face Count	1.36220	0.18150

Table 5.5: Results of the images from the human artist compared to the pieces Instagram selected.

CV Variable	Test Statistic	P-Value
Brightness	0.453195	0.65361
Ratio of Unique Colors	-0.45925	0.651656
Threshold Black Percentage	-0.20527	0.83883
High Brightness Percentage	-1.32712	0.204521
Low Brightness Percentage	-0.84416	0.406565
Corner Percentage	-0.77744	0.449341
Edge Percentage	-0.72626	0.47422
Face Count	1.56614	0.133185

key differences between human and machine perception of the art exhibit. According to the computational data, the pieces selected in both conditions were comparable. Nonetheless, human participants noted meaningful differences between the two bodies of images; we summarize the human observations here but see [under review] for more information. Specifically, the human participants reported that human-selected pieces were more aesthetically interesting to look at; they found these images to be more complex and engaging. On the other hand, the algorithmically selected images were deemed straightforward and easy to process. Regarding similarities, the gallerygoers were interested in the same type of information from both the human and algorithmic sides; they wanted to understand "why" the selected images were selected. The gallerygoers were as interested in the artist's thought process as in which aspects of the recommendation algorithm led to the chosen images.

In this way, our results indicate that the critical difference between computational and human perceptions lies in the disparity between top-down and bottom-up processing. Human perception focuses on top-down meaning-based considerations (like the story or high-level aesthetics). At the same time, the machine necessarily operates through low-level perceptual input such as pixel data and engagement metrics. The software's inability to detect high-level differences in meaning or aesthetics, which are the focus of human users, demonstrates the disparity between user expectations and machine capabilities for art recommendations.

5.4 Discussion

Our study provided an overview of the statistically notable findings obtained by analyzing the *Algorithmic Pedestal* exhibit. By systematically measuring various metrics across each subdivision, we were able to present comparative statistics that revealed distinctions separating the pieces in the exhibit from the overall pieces of the digital collection. Notably, the human-curated side exhibited a higher incidence of faces detected using the Viola-Jones Haar-Cascade Facial detection software. As indicated in the results, it is important to note that the system's recording of a face within the image was not always accurate (refer to Figure 5.2). These variations could be attributed to potential false positives associated with the open-source Haar-Cascade facial detection code, especially when applied to abstract or fragmented images.

Similarly, the Instagram-curated images showcased a significantly greater ratio of unique colors within the presented pieces than the overall Met Collection. This finding aligns with prior evidence suggesting that Instagram is inclined to prioritize bright and multi-colored images [204]. However, when comparing the two sides of the exhibit, the machine failed to discern any statistically significant differences, contrasting with qualitative human responses. Our results indicate that while the computer vision software successfully identified differences in specific metrics rooted in the history of computer vision and computational aesthetics, these metrics may be insufficient to detect other differences noticed or prioritized by human participants. This disparity underscores a disconnect between human and machine perceptions of the same curated images.

5.4.1 Compared to Previous Research

In this research, we demonstrate how humans and machines perceive artistic curation differently, thereby building on established research that outlines the differences in human and machine perception and the risks of assuming that they are either completely the same or wholly different [51, 296].

A notable disparity emerges when comparing the computational analysis results to human observations from the gallery experience. Previous qualitative findings highlighted users' views on the differences between human and machine outputs [122, 180, 265]. The human work was anecdotally more abstract, holistic, contextual, and emotionally resonant, whereas the algorithmic work was more object-oriented, recognizable, and individualistic. However, the computational analysis did not discern these same differences, prompting questions about the limitations of computational perception. As outlined in the brief history of computer vision and computational aesthetics in Section 5.2, machine perception focuses on objective data measures when processing an image. Therefore, even as society applies machines in artistic contexts such as Instagram or museum curation, computational methods cannot measure contextual meaning and emotionality, which human audiences prioritize.

Alternatively, it is plausible that biases influence human perceptions, such as those introduced by floor labels, leading to exaggerated differences between the two sides. Rae recently highlighted how labeling work as human or algorithm-made might lead to negative perceptions of the work by audiences [263]. Similar claims come from other research stating that humans consider context and background knowledge when forming their ideas of the world [206]. This form of processing injects ideas and presumptions that alter the perception of the experience, potentially leading to the participants reporting the differences they identified. This juxtaposition raises fundamental questions about the nature of truth in perception—is it shaped by human interpretation or revealed by machines? Likely, the answer lies in a combination of both perspectives.

This discussion resonates with previous research comparing human perception to machine perception, highlighting how humans employ top-down and bottom-up processes to interpret information. In contrast, the metrics with which machines are programmed form inherent limitations. Beyond the distinctions between the two sides of the exhibit, gallery-goers shared common questions about the overall exhibition, irrespective of whether they viewed the artist-curated or Instagram-curated works, bridging the inquiry to the next section of this discussion.

5.4.2 Human-in-the-Loop

The exploration of art perception dynamics reveals that individuals are often more captivated by the narrative and presentation of a piece of art than by its inherent distinguishing features [56]. Unlike tasks in medical imaging, where specific objectives can be clearly defined, engaging with art in a gallery setting is multifaceted and subjective. Consequently, gallery-goers frequently ponder meta-level questions regarding the motivation behind the art and the selection processes employed by platforms like Instagram. These inquiries are equally significant to the audience as the individual observations of the artworks themselves. In contrast, a machine analyzing handcrafted pixel features cannot contemplate such contextual elements. Therefore, an opportunity exists to enhance computational systems that process gallery data by incorporating appraisal data, curatorial insights, and advancing computational aesthetic processing.

Appraisal data inherently provides a machine with information about the story and history of an artwork through details about its provenance and condition [347]. Similarly, curatorial data provides insights into the art selection process for certain special exhibits and how human curators organized them [142, 316, 255]. Finally, computational aesthetics is a growing field [208, 209, 211]. As the resources become available within the digital humanities, new research and cultural systems applications arise, combining machine processes with human artistic knowledge

[226]. Through integrating these data formats, there is the potential to build on the research of this project and develop it further into a tool that can elevate and expand the gallery experience.

While research and cultural systems can be expanded to interpret galleries better, it is crucial to question how images are recommended and selected when they disregard the human perspective and rely solely on measurable metrics. In doing so, we risk overlooking the rich motivations and narratives that inform the presentation of art, reducing it to a mere list of parameters devoid of historical or curatorial context. Therefore, we suggest empowering art viewers by enabling them to provide direct input into the algorithmic experience. We recommend a human-in-the-loop [233, 363, 42] approach for all algorithmic systems that enable art-viewing. In this way, the users would be able to sort the artworks according to their interests, tastes, and considerations—essential in the subjective and personal experience of art-viewing. It is worth underscoring that we suggest taking a human *in* the loop approach versus human *on* the loop; the human user should not just check the algorithms' process but rather be able to drive their experience with the algorithm.

As collaboration between humans and algorithms becomes a reality, we argue for active, beneficial partnerships rather than one-sided content pipelines. This project provides a practical example of how a machine can perceive a gallery as wholly similar, even though the reported audience reactions capture notable differences that greatly inform their appreciation of the work. Understanding these different perceptions and how they relate can aid in evolving the art of curation in the modern era.

5.4.3 Limitations

The work presented in this project compares quantitative analyses to previously reported qualitative user responses. Even with an open-source quantitative approach, certain limitations must be addressed. For the data collected, we must note the potential limitations or biases involved in developing the presented findings. Of the 490,000 pieces in the Metropolitan Museum of Art's complete Open Access collection, the *Algorithmic Pedestal* selected only 1,204, with the exhibit curating 33 pieces. While there were enough pieces for some basic statistical comparisons, this work does not apply computer vision analyses to a complete array of artworks. As future collections and exhibits apply similar approaches, the metrics, and statistical analysis may reveal more insight into the similarities and differences between the various curated pieces.

As for the quantitative analysis, it was intentionally straightforward and focused on handcrafted feature analysis. Such analyses are less computationally intensive than some of the more advanced graph neural networks (GNN) analyses conducted in the computational aesthetics literature [366, 170]. More purpose-built image classifiers or art-focused tools would result in different measurements and observations. Nonetheless, these foundational scripts form a baseline from which comparisons can be made to the machine's measurements when processing an image.

5.5 Conclusions

This project presents the findings of a technical analysis of an exhibit-based research project that compared human curation with algorithmic curation. This project illustrates how baseline computer vision software processes art images, revealing similarities and differences between the pieces selected by the curator and the algorithm. Ultimately, by comparing the metrics with the previously reported human observations, we identify apparent differences between the information that human audiences consider valuable and the metrics that the computer vision software considers statistically significant. In presenting these differences, we urge computer science and art researchers to deeply consider the ever-evolving relationship between computation and the arts, driven bidirectionally by objective metrics and subjective perceptions. In particular, we encourage technologists and researchers to consider updating computational systems for the arts to account for human users' perceptions rather than relying on current computer vision protocol, which results in notably different decisions than those of human users. With this call to action we lay the groundwork for the following project aiming to introduce a new art dataset.

Much of this chapter was also reviewed for the paper: von Davier, T. Ş., Van Kleek, M., & Shadbolt, N. "AppraiSet: Discussions on a New Art Dataset," Contemporary Aesthetics.

6
Developing and Testing AppraiSet, a New Art Dataset

Contents

6.1 Introduction

Artificial intelligence (AI) applications in the arts have been a growing field of interest within computer science since the 1950s but have received even greater interest in the age of generative AI. Computer scientists have written how AI can create art on its own [106] or with artists [64]. However, there is limited work on how AI interprets art as if it were a museum or gallery visitor. Indeed, there has been work in computer vision to explore how a computer may process an art image [172], or even in the field of computational aesthetics to impose human definitions of aesthetic measurements to images [47, 156]. All of which we explored in the previous chapter. We found that established algorithmic approaches struggle to capture the aesthetic, qualitative experiences important for human audiences.

One of the main challenges in this task is that AI is only as effective as its training data. Any attempt to explore how AI interprets art and develops aesthetic interpretations must begin with the creation of a high-quality art metadata database. While existing art datasets, such as wiki repositories and digitized museum collections, provide valuable resources [315, 105], they are often biased toward well-known works and lack essential details like wall text, condition reports, provenance, and meta-descriptions of the artworks. As Mimi Onuoha from NYU and Data and Society, along with Van Miegroet et al., have pointed out, detailed art metadata is often a "missing dataset"—that is, data that is not readily available to the public [249, 333].

The second challenge facing AI interpreting art is that previous work has focused heavily on computer vision to analyze art in pre-established conditions. Ypsilantis et al. describe various papers that explore the identification of art via attributes alongside their dataset utilizing instance-level recognition. [362, 212, 217, 355, 172] While the identification of artworks themselves is a fascinating problem space to explore, artworks are rarely independent of social and financial contexts which may inform aesthetic interpretations.

With these challenges in mind, this project presents *AppraiSet*, one of the first attempts at creating an open-source art dataset containing evaluative art metadata. The data was collected from published auction lot results listed on the websites of major auction houses. Within this dataset, we bring together descriptive variables (artist, year, and medium) as well as interpretative variables (condition, provenance, exhibition history) alongside financial variables (estimated value at auction and final sale price). We outline how justifications for this dataset have grown over the past few years through related art and machine learning work. Following the motivation for the work, we describe how the data was collected, cleaned, and presented for open and collaborative use. We then present a case study analysis of how a simple AI algorithm can interpret the dataset using a Latent Dirichlet Allocation (LDA) topic model to differentiate and describe different art pieces. An LDA model is a form of natural language processing (NLP) that takes a data input and establishes patterns in the text to identify specific topics, which are then tagged and highlighted as distinct concepts present in the data [43]. Ultimately, there is a brief discussion on the potential impact beyond the world of datasets and machine learning regarding the ability of machines to form their aesthetic interpretations of art and how those may be extensions of existing approaches to aesthetics.

6.2 Background

This section outlines the context of related work that gives rise to three primary justifications for creating and managing this dataset. The first is a continuously growing interest in computationally analyzing artworks, already established in Chapter 2.3.1.1. The second is to further the initiative for open and accessible data. The final justification outlines the significant potential data from auction lots and provides researchers with an AI with high-quality training data.

6.2.1 Availability and obstacles to art data

Recently, the contributions of available big datasets, compute power, sophisticated algorithms, and the availability of online markets are beginning to shape the financial and social networks of the art world. This digital evolution is exemplified by the Blouin Art Sales Index and the Artnet Price Dataset, two large online datasets marketed to art experts as research tools [44, 20]. These two datasets also provide the second justification for this project.

While two similar datasets exist, Onouha's claim that information about art metadata is missing is still legitimate due to limited accessibility, as discussed below.

The Blouin Art Sales Index, offered by Blouin ArtInfo, searches and visualizes data regarding sale records of specific pieces or artists. A unique feature of The Blouin Art Sales Index is the Interactive Market Insights tool, which allows users access to high-quality data visualizations alongside their search results. It is an incredibly effective tool used by multiple researchers investigating various topics [333, 133, 124]. However, this information is only available through Blouin ArtInfo's subscription plan. Moreover, the subscription plan only provides restricted access to the dataset; for example, one can conduct a maximum of 200 database searches and receive up to 20000 results. Therefore, any attempt to select a large sample from this database, as needed to support computational models, becomes costly and arduous.

The Artnet Price Database is comparable to The Blouin Art Sales Index. The database includes information about particular galleries and the primary sales data for individual pieces and artists. Unlike The Blouin Art Sales Index, The Artnet Price Database does not emphasize data visualization. However, it focuses on the quality of the available queries and search tools by highlighting its relationship with galleries and art dealers. Nonetheless, this database is also subscription-based and limits the data a user can request. As both Artnet and Blouin have business incentives to restrict access to information, the only way for researchers to explore large-scale art data is to compile a database.

Both tools support Onouha and Van Miegroet et al.'s claim that information regarding artists and art metadata in a dataset is difficult to access [249, 333]. Seeing the lack of access to an existing open dataset regarding financial art valuation and appraisal, we were motivated to develop our research using methods that would ensure understandable and rigorous data integrity.

6.2.2 Justifying appraisal data

To understand how AI and aesthetics interact, we need high-quality art data to allow an AI to explore the information. This data is not just digitized scans of the artwork, as many museums tend to do with their collections [316]. Instead, the data needs to be rich in various variables more directly tied to the topics covered in aesthetics. Therefore, we built *AppraiSet* on information from publicly listed art auction lot results. This is advantageous for an AI for various reasons. First, auction lots provide a large sample of data to draw. Even with the impact of the COVID-19 pandemic, the art market's valuation was at 61 billion USD, with an estimated 36 million transactions in 2021 [310]. Second, appraisers are art experts, and their records use the language of aesthetics to describe the pieces. Third and finally, auction lots include quantifiable metrics in the form of art valuation.

While some may argue art valuation via appraisal differs from aesthetic valuation [312], others argue it is increasingly difficult to separate the two concepts [199]. Modern approaches to art analysis rely on quantitative financial values for the basis of measurement and comparison [165]. Furthermore, art value in the form of financial details is a useful way to quantize art that has been around for far longer than AI has been in the discussion. As a result, we have kept those variables within the dataset should they help explore precisely how an AI may approach understanding and interpreting the artistic pieces. While the algorithm and the dataset will not necessarily establish a relationship between aesthetics and art value in this project, this new way of looking at art may add color to the ongoing debate among art scholars.

In addition to showing an algorithm in practice, we are offering access to a dataset as an initial step in consolidating high-quality art data that is not reliant on images and scans of artwork. Our data comes from credible auction houses and contains full appraisal details, allowing an AI to look at art as more than just a piece of work; it also includes information about the artist, condition, and provenance of the piece.

6.3 Methods for dataset compilation

This section outlines how we built the dataset using web scraping and data compilation. Our work followed a set of globally recognized variables and gathered data from online public auction lots.

6.3.1 Variables

Art professionals document and analyze art using clearly defined variables and classifications. A standardized set of these quantitative and qualitative variables for each transaction forms the basis of the field of art appraisal. Mary Rozell, Global Head of the UBS Art Collection, detailed the specifics of an appraisal in her book on art collection [282]. Sotheby's also presented a mini-series on YouTube providing a detailed overview of these variables [305]. Our initial dataset implementation incorporated these industry experts' shared variables and classifiers.

We added additional variables to ensure the dataset is robust enough to handle various cross-sectional explorations into what variables inform art analysis. Our additional variables fall primarily into information about the artist and transaction details. Table 6.1 outlines the variables in detail from the various groupings.

Industry standard These are the essential 11 variables industry experts describe to develop an estimated value for an art piece. It is essential to recognize that no variables are specifically associated with the visual components of the piece. In other words, this dataset does not capture a visual representation of the art piece. The primary motivation behind this is to establish an appraisal process that is general enough for an AI to consistently evaluate artworks of all aesthetic types.

Artist details These details include each artist's birth and death year and their age calculation. The age is calculated to the current year if there is no death year. There is an additional measure of the artist's approximate age when making each art piece. Since the dataset is organized by artwork, the details regarding the artists' ages can be repeated, but what often sets it apart will be their age when creating a piece of art. Another variable we included was to see whether the artwork was sold after the artist's death. There may be an interest in understanding the impact of an artist's legacy on the evaluation of their artwork.

Table 6.1: Table of variables

Name	Type	Description
Title	String	How people refer to the piece.
Artist	String	Who made the piece.
Medium	String	The material and form of the piece.
Dimensions	Integer	Measurements of the piece.
Dating of the Object	Integer	How old the piece is or roughly when it was made.
Condition of the Artwork	String	Notes on any previous restorations, markings, any damage, etc.
Provenance	String	An object's history of ownership or its association with a significant collection.
Edition	String	Number and Size, rarity, or uniqueness.
Historical or Cultural Relevance	String	Any connection to major events or known infamy.
Exhibition History	String	Has the piece been displayed somewhere, if so, what exhibition and when.
Publication History	String	Has the piece been written about and if so what publication or release.
Birth Year	Integer	When the artist was born, if known.
Death Year	Integer	When the artist died, if known.
Age	Integer	The age of the artist at death or currently
Age at Time of Work	Integer	The age of the artist when they created the particular work.
Posthumous Status	String	Whether the artist has passed at the time of the auction sale.
Lower End Estimate	Integer	The lower of the estimated value range
Higher End Estimate	Integer	The upper end of the estimated value range
Sale amount	Integer	The amount the piece was sold for
Currency	String	The currency in which the auction sales were held. Can also be used to identify region of sale.
Conversion to USD	Integer	A conversion to a single currency for comparison.
Sale Status	Boolean	Whether or not the lot was actually sold.

Financial details The auction house sites that were scraped to create the dataset would provide the estimated value associated with each piece. These values and the currency of the financial value were included. An additional variable signifies whether or not the specific art piece was sold with all the associated variables. This variable serves as a way to check whether a particular financial estimation is aligned with the public's expectations. The final sale price, or lack thereof, is converted alongside the estimated values to a single currency for consistency and ease of comparison. We maintained the financial details in the dataset for future research. As mentioned, aesthetic value and market value are oft contentious subjects [312, 165], and more data may add further insights into the debate.

As a result of these additional variables, each art piece within the dataset had at least 11 base variables with an additional 11 expanded variables. To reiterate, none of the variables considered the visual details of the piece. This was carried over from the guidance of the subject matter experts, who did not include piece details as part of their appraisal process [282, 305]. Like an appraiser who needs to value many pieces of different styles, an AI interpreting aesthetic language also needs to be as general as possible. Therefore, the current dataset contains no visual representations of the works.

6.3.2 Data collection

We collected data from past public auction records to provide well-formatted art metadata. Unlike galleries, major auction houses post their lots online, available for all to explore. Furthermore, these details are only about the auction lots and do not include any information regarding audience members or art purchasers. This adherence to anonymity by the auction houses allows us to ensure that our data does not handle private or sensitive information. Through a custom-built web scraper and various data cleaning functions, we accumulated 31,000 auction lots rapidly. When deciding which lots to scrape, we limited it to physical art pieces (i.e., no NFTs or other digital works) and no major interior design auction sales. Our filtering decision was relatively successful in limiting the data collection to visual art pieces such as prints, multiples, paintings, photography, statues, and sculptures. Our program collected all the significant variables and classifiers described in the previous section with these limitations in mind. The standardized webpage layout aided our gathering of each lot in the public auction records. We intend this to be a live dataset; data collection and cleaning will continue beyond the current number of collected auction lots.

6.3.3 Data cleaning

The initial dataset required substantial cleaning. Data cleaning was necessary due to the variability of different pieces and auction locations. For example, some pieces had multiple measurements due to the inclusion of the frame, while other pieces sold in Paris had details repeated in both English and French. Another challenge was handling the inconsistency of details regarding text-based variables such as Condition Reports, Provenance, and Historical Significance. These variables were non-quantitative accounts of the piece's history and current state and were of variable length and detail across the auction lots. As these contain the most instances of aesthetic language and details of the art, we ensured they were cleaned and centralized in a text cell for AI training and analysis. We manually verified the data before dividing it randomly into a training and testing set.

6.3.3.1 Limitations of the dataset

Alongside submitting the clean dataset, it is essential to include a short description of the dataset's limitations. One of the most glaring limitations is the source of the data. We were limited to publicly available auction lots posted by major auction houses. Therefore, this dataset does not include information from museums or art historians. This data can improve as more art institutions digitize their text and variables.

Another limitation is the absence of a comparable, accessible dataset. As explained in Section 6.2, there are datasets on art pieces that are financially restricted. This has led our research group to develop this dataset and understand the ideal composition of art information through a process of trial and error. This lack of external validation is a strong motivator for us to open the dataset and initiate a collaborative process with other academics and individuals in the art world, ensuring it becomes a truly meaningful collection of information.

6.3.3.2 Ethics considerations on the dataset

The aforementioned limitations relate to the dataset's ethical considerations and future applications. The dataset primarily consists of sales from major auction houses, resulting in inherent bias due to its narrow focus on a privileged and affluent segment of the global population. Additionally, the dataset may not adequately represent new artists or those from underrepresented backgrounds, as it likely favors established artists with a significant art appraisal history. Releasing and reviewing this dataset will shed light on potential biases in the types of art and artists traded at public auctions, leading to improved processes.

We recommend expanding the variables and data sources to address the ethical challenges associated with the auction data and the utilization of this dataset. Collaboration with galleries and museums will introduce a new category of high-quality art data. Similarly, incorporating details about the artists' gender, ethnicity, and sexual identities ensures a more inclusive representation across various groups within the dataset and the art it encompasses.

6.3.4 Dataset details

We selected a subset of the overall dataset for the case study in Section 6.4. The verified data contains 10,016 individual auction lots. These auction lots represented 4,329 individual artists with an average age of 77.22 years. The average age of the artists when creating their art pieces was approximately 48.53 years. The artworks ranged from early historical pieces estimated from 1200 B.C. to contemporary art made in 2021. As a result of this distribution of historical to modern artwork, a handful of artists were unknown or craftspeople whose names have been lost to history. Therefore, while there was data regarding the artwork and financial details

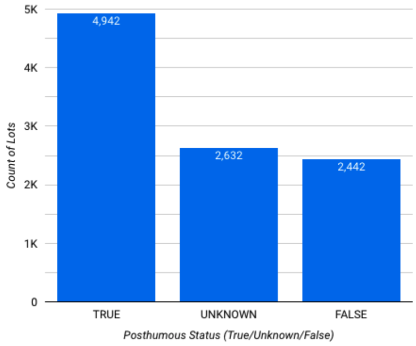

Figure 6.1: Distribution of auction lots across the posthumous status of the artists. Since some art is historical with no known artist, there is an unknown category.

Table 6.2: Table of Financial Variables (in USD)

Financial Variable	Min	Max	Average	Median
Lower Estimate	100	35,000,000	128,485	7,648
Upper Estimate	250	50,000,000	180,373	10,719
Sale Value	0.00	50,820,000.00	169,381.36	7,709.57

of the auction, there was limited information regarding the artist's details. Figure 6.1 shows the distribution of posthumous to living and unknown artists.

In addition to details regarding the artists, we included information about the various financial variables in the dataset. Table 6.2 shows the breakdown of minimum, maximum, average, and median values for the converted financial values in USD. Notably, we have the difference between the average and median sale value caused by high-value auction items that pulled the average well above. Furthermore, approximately 20 percent of the auctions did not result in a sale, which led to sale values of 0.00 USD, which we can explore in future work as missing data through a variation of Heckit models [149, 121].

6.4 Case Study: topic modelling with *AppraiSet*

Having outlined *AppraiSet*, we now present a case study exemplifying the dataset's usability while demonstrating how an algorithm may process aesthetic language. For this example we use an LDA topic model. Abera Yilma et al. define an LDA model as a tool for drawing abstract topics from an extensive collection of text data to explore the less obvious subject matter contained within the data [5, 43]. In other words LDA models are algorithms that classify data points into "topics" based on the patterns of associated words and phrases in the text. Using this model, we explore exactly how an algorithm may develop its classification of artworks using the language of art experts. Previous work on LDA models and art have handled smaller datasets without the language of aesthetics. Therefore, this will be an early step in exploring how an algorithm may process aesthetics on its own. The remainder of this section details how we established and evaluated our LDA model using the training set of *AppraiSet*.

6.4.1 Building the model

We built our LDA model using the *AppraiSet* training data subset of 8,506 artworks. Since this project focuses on exhibiting how an AI may develop its own aesthetic associations, we had to drop some of the columns to trim down to only the core qualitative text data points. The final variables we included were the title, artist name, medium/materials, condition, provenance, edition, historical or cultural relevance, exhibition history, publication history, and currency. For the model to efficiently process all of the text contained within these different variable categories, we concatenated all of the text into one large text document for each work. While cleaning the data, our previous concatenation accelerated our task of creating one "all_attributes" section within the dataset. With the data for each auction lot consolidated as a single text document, we could perform the proper pre-processing. This pre-processing involved setting up custom stop words such as auction, lot, condition, and report, among others. Stop words are specific words that are so universal in a dataset that their inclusion in a text description of an artwork likely will not help differentiate it from any other artwork data points [285]. We highlighted our unique stop words in the pre-processing stage. We also removed punctuation, established bigrams and trigrams, and lemmatized and ensured the text was all lowercase before developing the model using the popular gensim library [271].

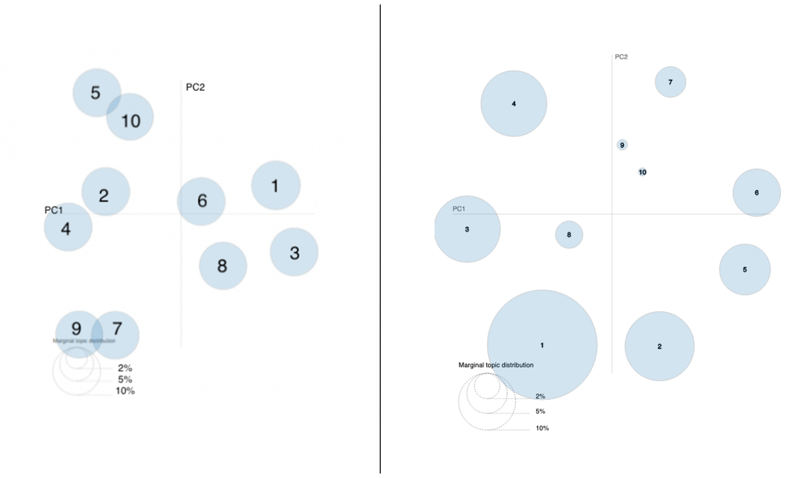

Figure 6.2: Comparison of our topic distribution (right) with the distribution presented within the Abera Yilma et al. paper (left) [5].

6.4.2 Evalutating the model outputs

We evaluate the performance of *AppraiSet* by comparing its trained LDA model to a model from previously published research, specifically the work of Abera Yilma et al., which serves as our baseline for evaluation [5]. Their study was chosen due to its similar approach to using art metadata to train language models. To assess our topic model's effectiveness, we analyzed the pyLDAvis output [271] and the coherence values generated from our dataset. The pyLDAvis output visually represents topic distribution across the dataset, showing the salience, frequency, and relevance of key terms that define each topic. Topic distribution refers to how distinct the algorithm perceives each topic to be, based on the terms and language used [271]. Each term in the dataset is assessed for its salience (how useful it is in identifying topics), frequency (how often the term appears), and relevance (how exclusive the term is to that particular topic) [299, 77]. Coherence values, however, measure how semantically related the words within each topic are. A higher coherence value indicates better topic formation [5].

We begin our comparison with the baseline by examining the distribution of 10 topics on the pyLDAvis plot. Figure 6.2 displays our 10-topic distribution alongside the results from Abera Yilma et al. Our model shows distinct separation between topics, with no overlapping groups. The topics are distributed across the

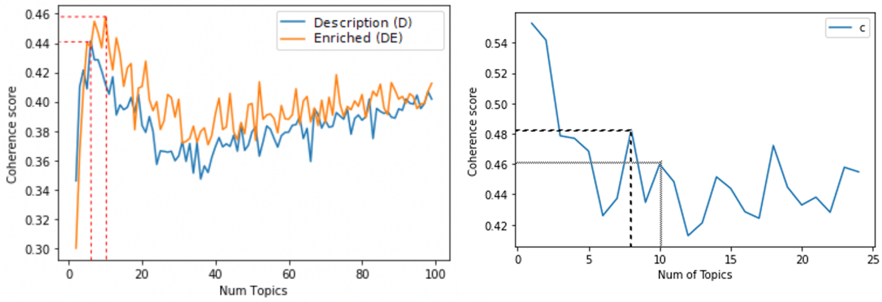

Figure 6.3: Comparison of our Coherence Model (right) at various number of topics in comparison to the Coherence Model included in the Abera Yilma et al. paper (left) [5].

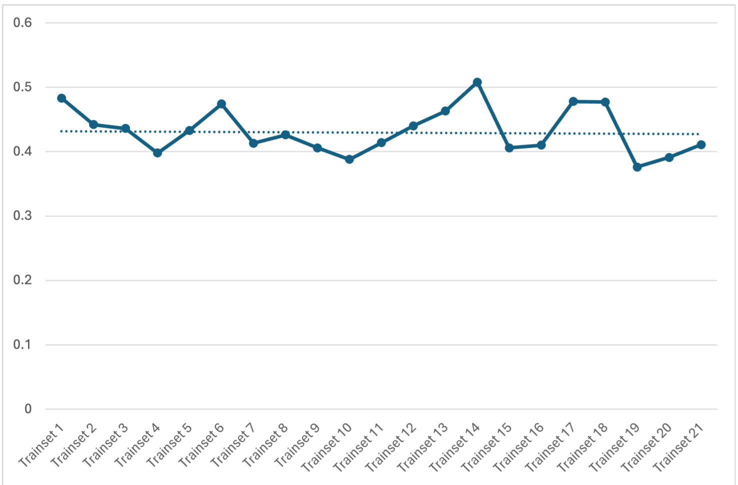

Figure 6.4: This figure displays the coherence value of different splits of the dataset for 8 topics. The horizontal, dotted line shows the average coherence value across all the splits of 0.432, comparable to the baseline.

chart without excessive clustering, demonstrating that the model interprets the artworks in each topic distinctly, with clear separations reflected in the unique language used to describe them.

For the initial model, we preselected 10 distinct topics. At this level, our model's coherence value is comparable to that in Abera Yilma et al.'s plot, around 0.46 (Figure 6.3). However, our coherence value increases slightly to over 0.48 when running the model with eight topics. Moreover, our coherence values remain consistently higher even when the number of topics is increased to as many as 25. We repeated this analysis with different dataset splits to ensure that the coherence of topics remained stable regardless of how the training and test data

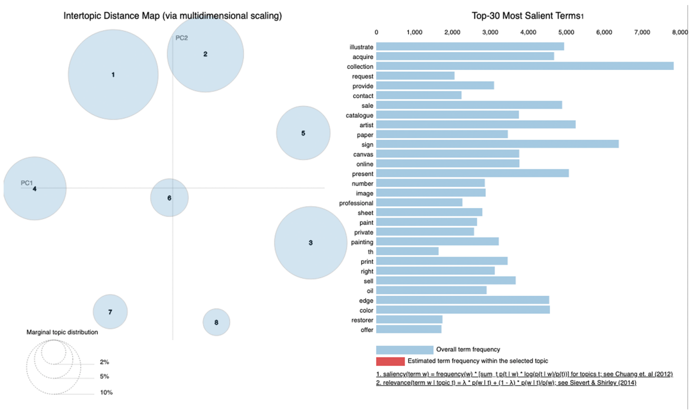

Figure 6.5: With 8 topics there there was the highest coherence model, the topic distribution is placed next to the top 30 salient words used by the model to define the topics.

were divided. The coherence values for eight topics across different test splits are shown in Figure 6.4, with values ranging from 0.376 to 0.508, consistently outperforming the baseline on average.

When analyzing the text associated with the topic model, we observe how AI interprets aesthetic language differently. Figure 6.5 displays the 30 most salient words from the dataset that contributed to creating the eight topics. These terms help the model understand aesthetic language and categorize the artworks. Table 6.3 presents the top three words relevant to each topic. For example, in Topic 8, which includes words used to describe East Asian porcelain pieces (Figure 6.6), we can see how the relevant terms provide insight into the type of aesthetics represented in that topic. The broader implications of this model for the field of aesthetics are discussed in Section 6.5.

6.4.2.1 Ethics considerations on the model

Our analysis indicated that our model did not specifically identify any particular type of artist or price range. Future research is needed to explore what patterns might arise from the different topics related to art made by different demographics.

Topic Num.	Term
1	present canvas edge
2	paper sign number
3	painting house exhibition
4	sale provide image
5	collection th [i.e., 12th] century
6	illustrate request contact
7	series artist world
8	rim enamel interior

Table 6.3: The top three relevant words in each topic for our LDA output.

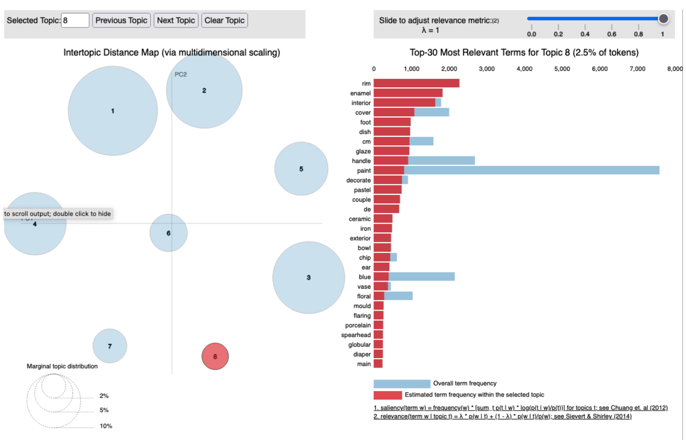

Figure 6.6: The text content that makes up the top n words describing topic 8 when there were only 8 total topics.

The ethical implications here are that if a bad actor were to use topics associated with a more "desirable" aesthetic, they might want to rewrite the descriptions of the artwork. Another important ethical consideration is the demographic makeup of each topic group to ensure that no one group is disproportionately included or excluded by the aesthetic language used. Understanding which economic or demographic groups are associated with each topic cluster would necessitate additional data. The need for more data raises ethical questions about collecting additional data to ensure the current data is free from hidden biases or trends. However, the debate about this question is beyond the scope of this project.

6.5 Discussion

In this project, we present *AppraiSet*, a dataset of art metadata collected from publicly listed auction lots. This open-source dataset is a first step in developing algorithms, like the one in the case study, that can help a machine form its own aesthetic language.

In the case study, we show how *AppraiSet* can train an LDA topic model that generates highly coherent topics for grouping and classifying artworks. We show how the model compares to previous research by outperforming on both the coherence and distinctiveness of the topics. Furthermore, we provide examples of topics that include descriptions of print art pieces and sculptures, showing how an AI may develop its own language around certain artworks and styles.

By releasing the dataset and presenting an example use case, we aim to contribute concrete insights to discussions on the intersection of aesthetics and AI. Contemporary aesthetics is at a crossroads. Scholars have found value in applying aesthetic concepts to artworks computationally on a large scale. These deployments identified patterns not previously seen when only working with a handful of pieces at a time. In the age of AI, we must consider how aestheticians will interact with algorithms that develop their own aesthetic principles (Section 6.5.1). Embracing this potential future, we can deliberate on whether AI-developed aesthetic principles should be restricted to AI-generated art (Section 6.5.2).

6.5.1 The Future of computational aesthetics

The field of computational aesthetics is facing a decision point like many other fields with the mass deployment of AI. Traditionally, computational aesthetics has represented the application of aesthetic theories and practices via software to a large-scale collection of art [209]. The software often relied on extracting "hand-crafted" features through computer vision software that analyzed the individual pixel data of a media object [366]. Until now, this has been a heavily human-driven approach where the software implements measures of aesthetic principles that have been pre-defined by aesthetic scholars.

As larger, higher-quality datasets become available, the capability of algorithms to develop their own language and associations is becoming increasingly commonplace. Studies have already shown that large language models (LLMs) tend to overemphasize certain words or write in unique ways that differ from human language [123, 264]. Therefore, it is not unreasonable to pose the possibility that AI could develop its own aesthetic principles to analyze art. Early evidence of computational aesthetic principles influencing AI generation can be seen in Cook and Colton's work on generating preferences [85]. Combining computational creativity with language models and *AppraiSet* would give rise to a new iteration of computational aesthetics and significantly impact the academic discipline of aesthetics as a new algorithmic language emerges.

This project is just an initial case study showing the potential for AI to develop associations trained on just art metadata containing text and integers. Future works can look towards integrating visual data from digital collections for further aesthetic analyses. In light of this potential future, we encourage further discussion into when and how AI aesthetics should be applied. For instance, analyzing art may require art historians and aestheticians to understand AI principles and basic functionality to engage with newly developed AI aesthetics. Similarly, we can imagine how galleries and art institutions may change to reflect tastes defined by algorithmic aesthetic interpretation of art.

6.5.2 Applying AI aesthetics

The insights gained from historical research on human-developed aesthetics can guide our approach to applying AI aesthetics. Some scholars argue against using Western aesthetic principles in analyzing Eastern art and vice versa [91]. They argue that this can lead to an increased risk of misinterpretation and potential dismissal of culturally salient pieces. In this case, we may advocate that AI aesthetic principles should only evaluate AI-generated art. Not only would this reflect previous scholars' recommendations, but it would also reflect the growing body of art venues and competitions working to separate AI-generated art from human-made pieces [280, 135].

Alternatively, some scholars, like Bence Nanay in *Contemporary Aesthetics*, advocate for examining art through multiple aesthetic principles and lenses in a heavily globalized art world [236, 31]. They argue that this approach offers audiences and artists new modes of exploring, interpreting, and conveying artistic experiences. According to this perspective, algorithmically developed aesthetic principles may serve as a democratizing force, drawing from a global collection of art metadata to offer a more generalized view of aesthetics.

Realistically, the answer likely lies somewhere in between the two positions. Algorithmically developed aesthetics will become just the newest set of principles in the canon of future aesthetics. From this collection of principles, artists, and audiences can pick and choose which aesthetic principles interest them and apply them to different interpretations of art.

6.6 Conclusion

As AI continues to integrate into the art world, it is poised to reshape our understanding of aesthetics. This project demonstrates how an algorithm can analyze art metadata to generate topics and concepts related to various artworks. Our research underscores the significance of art data and stresses the need for valuable insights from art experts. As a result, we are making *AppraiSet* openly accessible, inviting art scholars to participate in shaping the future of art datasets. Furthermore, we contribute to the discourse on the future of computational aesthetics and the potential role of algorithmically derived aesthetics in evaluating artworks created through human-AI collaboration.

Having gathered design requirements from stakeholders in Chapter 4, and explored algorithmic art processing in Chapters 5 and 6, we now have the pieces necessary to build a prototype that re-imagines our digital art experiences.

Much of this chapter was also reviewed for the paper: von Davier T.Ş., Larsen A., Van Kleek, M., & Shadbolt, N. (2025). ArtBot: An Exploration into AI's Potential for Guiding Art Analysis In Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '25). Association for Computing Machinery, New York, NY, USA. 11 pages. https://doi.org/10.1145/3706599.3720181

7
Designing Potential Futures

Contents

7.1 Introduction

This thesis has been an iterative process where each preceding project informs the subsequent work. For the final project, we aim to present a potential future algorithmic art experience based on the information developed in the earlier projects. Therefore, we take the design requirements described by our participants in Chapter 4 and the dataset values explored in Chapters 5 and 6 to build and test a new prototype.

We introduce ArtBot, an art companion built on a large language model (LLM) designed to act as a Socratic conversationalist to challenge the user while they are observing artworks. ArtBot acts as a "provotype"—a provocative prototype [49]—serving as a concrete semi-functional technology probe for exploring the potential of a new algorithmic experience (AX) [13]. Internally, ArtBot uses Retrieval Augmented Generation (RAG) upon a custom knowledge base to provide depth and completeness about art pieces being discussed, comprising a combination of open-source information about these works, curator-provided information, and additional open data collected from auction house records (drawn from Chapters 4 and 6).

To understand the opportunities and challenges for ArtBot as a potential future algorithmic experience for art appreciation, we conducted an exploratory study with the prototype through a within-subjects experiment followed by a qualitative exit survey (n=13). The experiment had participants view art in randomized, repeated iterations of three conditions that replicated three art experiences. The first condition replicates a digital collection page (which could be found on any museum website) containing art and some wall text written by a curator or art historian. The second condition replicates a social media post with an image and a basic label underneath. The third condition is ArtBot, where the AI accompanies the image and label. Our research questions are as follows:

- **RQ1 -** Could a Socratic opponent for digital art experiences help art observers experience art with a critical perspective?
- **RQ2 -** What are the design challenges and opportunities in building AI-driven Socratic opponents for art experiences?

We found significant differences between the conditions on three primary metrics: self-reported understanding, the use of complex vocabulary, and writing proficiency.

Post-hoc tests revealed that the prototype performed at the same level as a digital collection experience in some of these metrics, which statistically outperformed the social media experience. Our qualitative analysis reveals that the ArtBot highly satisfies the participants, and their feedback provides insight into its effectiveness in guiding art analysis and supporting critical thinking.

Based on these findings, we conclude by discussing what we learned from building a critical-thinking AI for art analysis and arguing that tools like these could be expanded to better collaborate with users as they review creative work and develop their own perspectives.

7.2 Developing ArtBot

In planning the research design for this thesis (Chapter 3), we knew the final project would explore novel interactions through an experimental system. ArtBot was developed by following research through design [372] prototyping methods that aim to gain research insights through the creation and testing of a prototype. As stated in Chapter 3, we draw on the methods of speculative design, technology probes, and provocative prototyping (prototyping). These methods aim to develop a new technology to challenge the status quo while admitting that the experimental tool is exploratory and not a release-ready system. We have drawn inspiration from both bodies of literature to create what we are calling a probotype.

7.2.1 Designing the Components

ArtBot is our approach to providing a meaningful art experience to a wide range of users through a digital experience. The development of this AI-powered art companion involved combining four features: design requirements, quality data, customizable LLMs, and open-access artworks.

Starting with design requirements, we pull from the findings of our earlier chapters. We have argued that art experiences need "active, beneficial partnerships rather than one-sided content pipelines" (Chapter 5). This motivated us to consider how a large language model could be paired with a displayed artwork, allowing users to interact and discuss the art directly with the algorithmic system. A second inspiration area for the design was the speculative design approach of Slow Technology [144, 243]. The original paper argues that researchers can prompt reflection in users by slowing down experiences that tend to be accelerated by technology. As we explored in Chapter 4, there is value for users and their art experiences when designs challenge social media's speed and shallow interactions. Therefore, the design philosophy of Slow Technology was incorporated into our development of ArtBot, along with the attributes of depth, conversation, connection, and time. Finally, we reviewed recent HCI papers that featured the development of provotypes and technology probes [274, 319]. Their insights on the prototypes' degree of fidelity and delivery system also informed our approach.

With our design approach solidified, we turned toward the system's back end. As we set out to build a customized large language model, one of the challenges became equipping it with knowledge specifically about artworks. The answer was to build the model as a retrieval-augmented generative (RAG) model. These models rely on a unique dataset the LLM can encode and then parse to improve and specialize its answers. However, that requires a specialized dataset of art data, and as stated in Chapter 6, datasets filled with usable art metadata are limited and difficult to access. Therefore, we turned to the open-source dataset we developed in the previous Chapter, *AppraiSet*. This dataset was built on art auction data, which included information about artists and artworks through provenance and condition reports. HCI research has explored the dataset to improve art metadata exploration [347].

We decided to utilize local LLMs for the rapid prototyping needed to develop ArtBot. The basic model details came from open-source tools like Ollama [23], which allowed us to download various local models for early experimentation and design. Having a local model meant we could keep some level of control over it as we worked to fit it within the overall algorithmic experience of the prototype. Running the prototypes and eventual experiments locally meant we could account for the data protection practices required by our institution's research ethics committee. Furthermore, we could develop our model files, which could be customized for different parameters and system prompts to customize the model. The unique capability of this prototype is the ability to prompt the user. This was inspired by education and critical thinking literature [367, 87, 260]. Building the LLM to function as a Socratic opponent for art education was the main task of the system prompt.

With the prototype functionality established, the final step was displaying art objects. We turned towards a local museum (The Ashmolean) that is free to the public and has an open digital collection. Working within their terms of service, we selected nine paintings randomly from their heavily featured Western art section ranging across different art styles spanning 200 years. This was done to avoid featuring any single artist or style heavily, which may sway opinions. We acknowledge that only using paintings from Western Artists heavily skews the prototypes towards a particular perspective. The decision was made based on the prototype's image quality and presentation capabilities. While the museum features artwork from all over the world, the pieces from other cultures were often artifacts better viewed in three dimensions, which is not a capability the prototype was built for. Therefore, paintings were the best option for displaying the artwork, and that limited us to their extensive collection of Western art paintings.

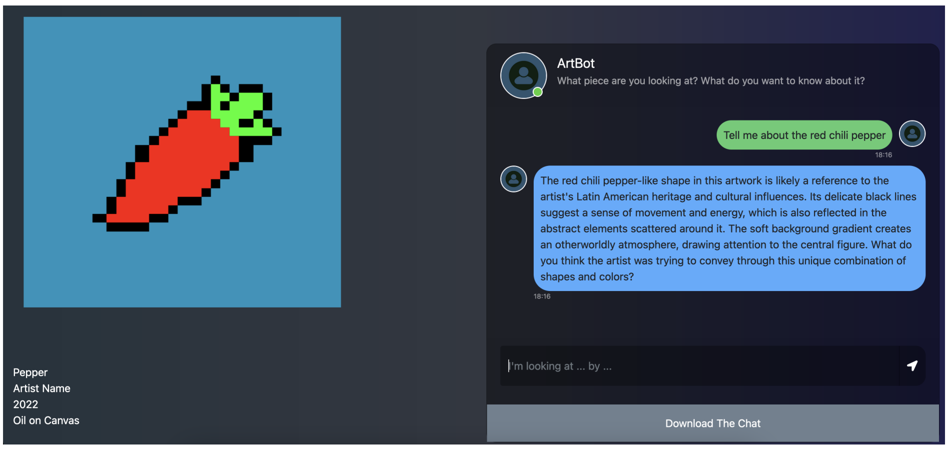

Figure 7.1: This is an image of the ArtBot experience. An artwork is presented to the user on the left side while the AI interaction is available on the right side.

7.2.2 The Experiment Ready Version

The experiment-ready version of ArtBot was built on a Llama 3.1 model [219] with specific art support from the *AppraiSet* art dataset. The nine artworks and their exhibit wall text were added to the dataset. To ensure the dataset also had information about the visual contents of the image, we had GPT4 visually analyze the artworks before outputting a paragraph description of the work. A custom system prompt then managed all of this. The system was instructed as follows:

> You are an art history tutor listening and responding to someone's observations about an artwork. Use the following pieces of retrieved context to answer their questions. Limit your response to 4 sentences. End every answer with one relevant question.
> Question: {question}
> Context: {context}

With ArtBot working, the final step was to package it in a local web app for testing. See Figure 7.1 for a standard view of ArtBot. The final web app and AX design is meant to fulfill the four attributes of art appreciation from Chapter 4, depth conversation, connection, and time. The depth comes from the RAG system and the custom dataset. The llama 3.1 model delivers the conversation attribute. Connection is meant to arise through the system prompt instructing ArtBot to respond to the participants' interests. Finally, time is fulfilled by the experience when the user is presented with an artwork and asked to discuss it and reflect on it.

The artwork is always displayed on the left, with the caption below it, offering the user a starting point to begin the conversation. Around the chat window and within the chatbox, we have placed questions, and example text meant to guide the user in initiating the conversation. Naturally, due to the Socratic method implemented in the system prompt, once the first message is sent, the ArtBot will take over, prompting the user with its questions about the work. It is this interaction that acts as the provocation or antagonism called for in provotyping [49] and antagonistic AI [58]. ArtBot intentionally causes friction in the interaction by asking the user to develop their own thoughts and opinions about art without directly or automatically providing an answer. The user can engage with the conversation as long as they wish and exit by clicking away or using a Llama 3.1 stop word. For testing purposes, we also included a button that will allow us to download the contents of the chat locally for diagnostics and further analysis. Readers can find the base code in the GitHub listed in Chapter 1.

7.3 Experimental Methods

With our probotype built, we can test it with users to get their feedback and perspectives on ArtBot as an art analysis companion. We test how ArtBot performs through a within-subject experiment, and we gather user feedback through an exit survey that collects their reactions and qualitative reviews of the experience.

Table 7.1: An overview of our participants and their relationship to art and LLMs before the study began.

Participant	Art Knowledge (1-7)	LLM Comfort (1-7)
P1	2	4
P2	3	1
P3	2	2
P4	1	2
P5	3	3
P6	7	7
P7	6	6
P8	3	3
P9	6	2
P10	3	2
P11	3	5
P12	1	3
P13	3	3

7.3.1 Participant Recruitment

Due to the controlled nature of the prototype, participation was done in person in a controlled academic space (study rooms, meeting rooms, and classrooms). Therefore, participant recruitment was also done in person on a university campus through advertising posters, in-person flyering and advertising, and direct messages sent across the researchers' networks. Informed consent was collected before conducting any research or data collection. Our institutional ethics review approved all of the user research.

In the end, 13 participants were recruited, comprised mostly of graduate students aged 18-44. Participants represented six global regions (Middle East, Europe, Africa, Oceania, South America, and South Asia). The gender breakdown was approximately 15% non-binary, 38% male, and 47% female. Regarding their experience with art and art museums, participants reported going to museums approximately 2-12 times in the last year. Two participants reported having gone fewer than twice in the last year, and one reported going over 25 times within the last year. Their average self-reported knowledge of Art is 3.31 on a scale from 1-7, where 1 is a complete novice, and 7 is a student of Art. Similarly, their comfort with LLMs and chatbots was, on average, 3.46 on a scale from 1-7, where 1 had never used LLMs, and 7 had used them multiple times. See Table 7.1 for an overview. All participants considered themselves fluent in English, if not native speakers.

7.3.2 Study Design

We ran a within-subjects experiment comparing ArtBot to two other conditions that reflect current art experiences. A within-subjects experiment allows researchers to control for participant variance. In this project, participants are highly educated and have varying experience with art and LLMs. Therefore, we control for this variance by having each participant encounter each condition. To ensure the interaction with an experimental condition is not a fluke, we randomly repeated each condition three times per participant. To avoid fatigue or the impact of attaching a single artwork to each condition, we randomly distributed nine images for each participant. Ultimately, the experiment has our 13 participants interact with all three conditions for nine images (3 per condition). The order and pairing between image and condition are entirely randomized, meaning each participant sees nine different screens differently from the other participants. This randomization helps avoid ordering effects [295] and slightly decreases the random effects introduced by the different artworks. See Table 7.2 for a breakdown of the within-subjects study.

Finally, while it is not common for a technology probe or provotype to be paired with a formal experiment, recent work among HCI researchers has started exploring this form of evaluation using similar methods [319].

Table 7.2: The three conditions of the within-subjects experiment.

Condition	Num. of Images	Situation
Digital Collection	3	Presented an image along with the official wall text. Participants could view the art and read the text before responding.
Social Media	3	Presented an image with just a caption including artist name and title, no wall text.
ArtBot (AI)	3	Presented an image with the caption, but have the interaction with the LLM before responding.

Digital Collection The digital collection condition aims to replicate a page on a museum website. With greater digital resources available, digital collection pages have become popular offerings on museum websites [59, 40]. These digital collections are also often the subject of redesigns by researchers [41, 142, 213], making them a well-established user experience against which ArtBot can be compared. The participants are presented with an image with a label including the title, year, artist name, and materials. The image and label are supported by a plaque of "wall text" written by a curator or art historian. See Figure 7.2.

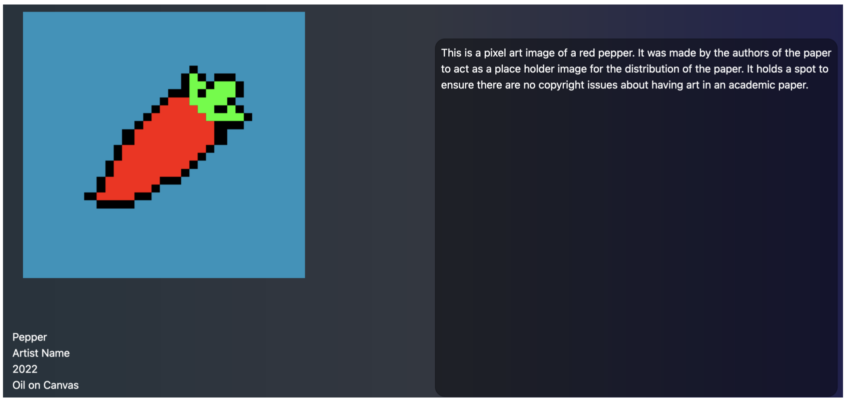

Figure 7.2: This is an image of the digital collection experience. An artwork is presented to the user on the left side while the wall text is available on the right side.

Social Media The social media condition replicates a basic social media post. In this case, the image is central to the digital screen and accompanied by a short label. The label, again, is made of the title, year, artist name, and materials. See Figure 7.3.

Figure 7.3: This is an image of the social media experience. An artwork is presented to the user in the center while a short label is available on the bottom.

ArtBot The final condition is our prototype. The image and label are kept the same, but instead of having a plaque with wall text, the participants are presented with the interaction area where they can chat with the LLM about the piece.

Following each condition, the participants are presented with a series of Likert questions regarding their experience with the art. These questions asked how much they liked, understood, connected with, and emotionally felt about the artwork. At the end, they were asked one larger text question, asking them to explain their position. All 13 participants answered these questions, but 1 participant's form was deleted due to an issue with Microsoft Forms.

To fulfill the methodology requirement of both provotyping and technology probes, we also had to collect qualitative reflections. At the end of the study, the participants were asked qualitative exit survey questions. These questions were meant to gain perspectives on the experience as a whole and the prototype from the usability and user preference perspective. All 13 participants answered these questions without any form issues.

7.3.3 Data Analysis

Following each experimental session, we received 36 Likert responses, nine essay responses, and three qualitative reviews of the overall experience. Therefore, we used a combination of quantitative and qualitative analysis methods.

For the Likert scale responses, we used a χ^2-test for independence [295] comparing the categorical values of the scale responses with each test condition. If the test returns a statistically significant p-value < 0.05, we conducted posthoc Bonferroni adjustments [295] to see which, if any, combination of variables was statistically significant. We used the stats module from the scipy Python library [339] for these calculations.

For the essay responses, we conducted two different forms of analysis: a computational text analysis and a human grader approach. The computational text analysis method comes from the textstat Python library [30] and allowed us to gain measures on the word length, number of complex words, and predicted grade level. The grade level prediction score combined multiple grade level predictors to output an estimated year range (9th to 10th grade), which we marked as a value of 9.5. A score like this would correspond to a student in the first year of secondary school. The calculation combines established readability scores that have been used and built upon over the last 80 years [88, 113, 72, 82, 247]. We treated this as a continuous variable as these measures correspond with academic year lengths from 0.5 - 23.5. Once all of the text analysis metrics were calculated, we used the statsmodels Python library [290] to perform repeated measures ANOVA tests to see if there was a statistically significant difference between the means of the three conditions. If the test revealed a statistically significant difference, we performed post-hoc Tukey method tests [295] to see which of the three conditions differed.

For the human grader portion of the experiment we enlisted the expertise of a colleague, Aaron Larsen, he has formal education training and professional experience as an educator in multiple nations. His background includes museum educational outreach aimed at teaching school children of various ages how to examine and describe the artworks they see. Aaron served as the human assessor

Table 7.3: The rubric developed to score the study participants' text responses. Built on the literature of various national curricula.

Category	Emerging (1)	Developing (2)	Proficient (3)	Extending (4)
Emotional Connection	I can identify what binary emotion this painting makes me feel.	I can identify what binary emotion this painting makes me feel, and a scale of this emotion.	I can identify how this painting makes me feel beyond a binary emotion. I can use complex language to identify one or more emotions and explain why.	I can identify my emotional connection to this piece and compare it to previous experiences or learnings from my life.
Interpretation	I cannot identify what the painting is depicting.	I can make a guess about what the painting is depicting.	I can use clues to make a guess about what the painting is depicting.	I can use clues to confidently identify my own interpretation of the painting, then explain my answer in detail.
Language	I can use simple words without sentences.	I can write in sentences using simple words to express simple ideas.	I can write multiple sentences to express my ideas.	I can write multiple sentences to express complex and abstract ideas.
Information Retention	I cannot use the information provided in my response.	I can use the information provided about the piece to make a simple observation.	I can use the information provided to influence my interpretation of the painting.	I can use the information provided to influence my response, and then use it to refer to previous knowledge.

for the text responses and was kept blind from the results of the other analyses and which responses came from the conditions. Aaron developed a rubric (see Table 7.3) to evaluate the essay responses. The rubric was built on the foundations of the British Columbia Curriculum, first fully implemented in 2019 [245]. Following the Provincial Proficiency Scale, the responses were assessed along the criteria of Emerging, Developing, Proficient, and Extending [246]. Similar arts education policies are in place in the USA [238] and United Kingdom [116]. The Arts Education curriculum for Secondary Students highlights the significance of applied knowledge, personal connections, and clear communication when discussing and reflecting on art. The curriculum determined the qualities the rubric assessed. Based on the algorithmic assessment of the student writing quality, the assessment was designed for Grades 8 and 9 students. The use of ArtBot replicates the technique of dialogic teaching, a pedagogy that enshrines learning within the framework of a conversation [303]. While the algorithmic assessment of the responses was purely focused on the quantifiable data, the responses following a dialogue with the AI present similar levels of development and growth as a traditional dialogic classroom lesson: "...it would be appropriate to base the assessment of students' literacy development, at least in part, on an examination of the communicative competence they display in structured group discussions about the texts they have read" [303]. Following the development of the rubric, Aaron also evaluated all of the text responses, providing them with scores ranging between 4 and 16.

Finally, to analyze the qualitative responses from the participants delivered in the exit survey, we conducted affinity diagramming and clustering [159, 318] to map out what was said and how the comments might relate to each other. Affinity diagramming and clustering involve taking participants' notes and feedback and having the researchers cluster them based on similarities. Affinity diagramming is one of the methods outlined in Holtzblatt's writing [159, 160] referenced in Section 3.2.1. These similarities become higher-level findings or statements that capture our participants' perspectives. Through the clustering process, we developed larger takeaway themes from feedback to inform our review of the experimental experience.

Table 7.4: Distribution of responses for the self-reported understanding of the artworks divided by condition. The scale rating describes users' reported understanding with 1 being low and 7 being high understanding of the artwork.

Scale	Social Media	Digital Collection	AI
1	83.33%	0%	16.67%
2	61.54%	30.77%	7.69%
3	30%	40%	30%
4	27.78%	27.78%	44.44%
5	28.57%	23.81%	47.62%
6	24%	44%	32%
7	0%	66.67%	33.33%

7.4 Results

7.4.1 Likert Responses

With the categorical comparisons across the Likert scale ratings of each condition, we can examine how our participants reacted to the artwork they saw and whether those reactions changed based on their condition. Our null hypothesis for each test is that the reaction to the artwork is independent of the conditions. We could not reject the null hypothesis for our participants' self-reported liking, emotional response, and connection to the artworks. These ratings are likely tied to their individual art preferences rather than shaped by the conditions.

In their self-reported understanding of the artwork, the χ^2-test returned a statistically significant result (statistic=21.41, p-value=0.044). This means we could reject the null hypothesis and state that there is a relationship between self-reported understanding of the artwork and the conditions. The proportional distribution of responses across the conditions can be seen in Table 7.4. There appear to be more instances of greater understanding among the digital collection and AI conditions with more instances of lesser understanding in the Social Media condition. While differences are based on the χ^2-test, the Bonferroni adjustment results show we cannot say which differences are significant. In this instance, the digital collection and AI conditions have higher understanding ratings, while social media predominantly has lower levels.

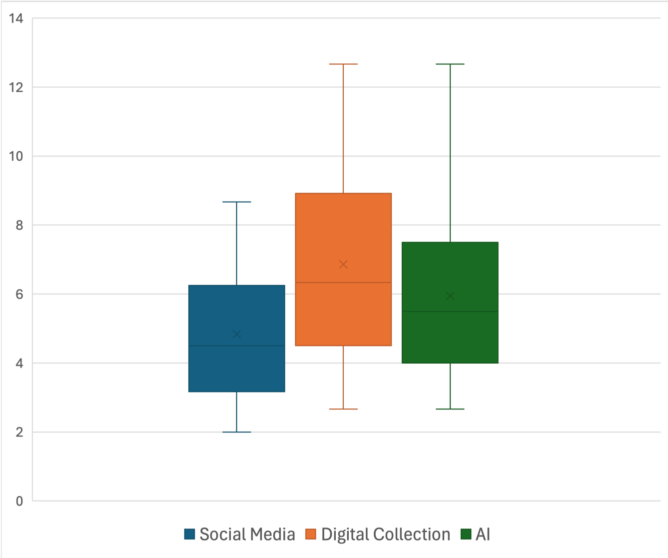

Figure 7.4: This boxplot displays the quartiles of difficult words used by the participants across each condition. While these are significantly different individual difference between conditions could not be specified by the Tukey post-hoc test.

7.4.2 Text Responses

In evaluating the text responses computationally, our repeated measures, ANOVA has the null hypothesis that there will be no difference in performance across the three conditions when we examine the length of text, the number of complex words, and the grade level of the writing. We could not reject the null hypothesis for the text length; the number of words written did not alter statistically based on the condition.

We did receive statistically significant ANOVA calculations based on the number of complex words across the conditions and the grade-level calculation. For the number of complex words (F = 3.78, p-value = 0.038), we could reject the null hypothesis and state that there is a difference based on the condition. As seen in Figure 7.4, the variance in complex word counts for participants in the digital collection and AI condition was relatively high, even though their averages were above that of the Social Media condition. Due to this large variance, the post-hoc Tukey test could not determine which of the conditions were statistically different from each other. In the grade-level analysis (F=6.38, p-value = 0.006), we can reject the null hypothesis and state that grade level differs based on the experimental conditions. In Figure 7.5, the variance between the digital collection and AI

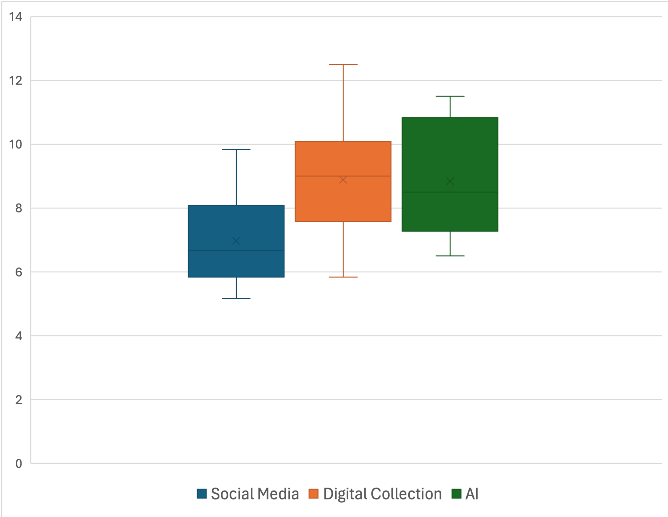

Figure 7.5: This boxplot displays the quartiles of grade level of the writing made by participants across each condition. These measures are different, with the digital collection and AI conditions having higher grade levels than the social media condition.

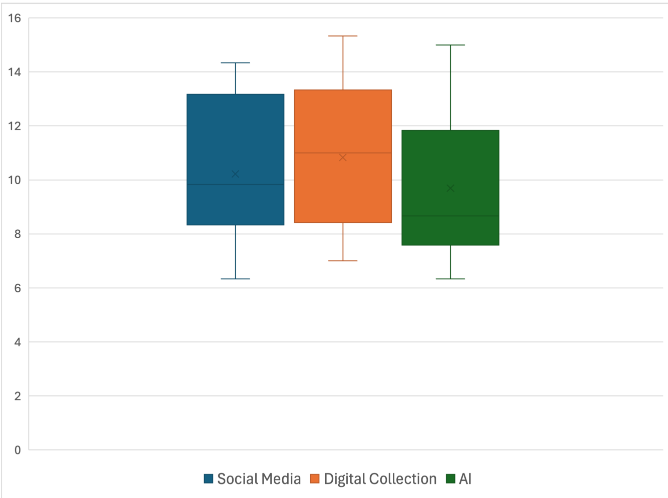

Figure 7.6: This boxplot displays the quartiles of the rubric score of the participants across each condition. These measures were not found to be significantly different across conditions.

conditions overlap while the Social Media condition is lower than either one; this is further reinforced by the posthoc Tukey measurement that identifies statistically significant differences between the digital collection-Social Media conditions and the AI-Social Media conditions. The post-hoc test did not find a difference between the digital collection-AI conditions indicating they are statistically comparable.

After the grading was completed, the scores associated with each text response delivered by the human assessor were also gathered for repeated measures of ANOVA.

With the null hypothesis set that there would be no difference in performance across the three conditions, the test returned a non-significant result (F = 1.53, *p*-value = 0.239), meaning we cannot reject the null hypothesis. When the averages and variance were plotted in Figure 7.6, there is evidence that the variety in response quality across users in the conditions meant the differences between conditions were insignificant.

7.4.3 Qualitative Review

7.4.3.1 Compared to Other Conditions

At the experiment's end, all participants shared their thoughts on ArtBot, the experiment, and the other conditions. Most participants (10/13) reported ArtBot as their favorite of the three conditions. When asked why they preferred the AI condition, our participants focused on the system's ability to aid them in exploring their thoughts. Another typical response was that the system revealed new information or details about the artwork, which added value to their experience. By adding value and new observations about the piece, participants saw ArtBot positively.

Two participants (2/13) prefer the traditional digital collection experience with the wall text. One mentioned that this text added to the context of the artwork and helped explain the emotions the piece conveyed. The other participants preferred the straightforward delivery of the wall text, whereas they did not enjoy the wandering discovery process of ArtBot. Their response highlighted a preference for immediate factual information.

One participant expressed interest in combining the AI with the wall text. They valued the text as it gives a "baseline for what makes the artwork special," but the ArtBot allows users to "dig deeper" into the piece (P1). During their session, P11 also highlighted how ArtBot could be delivered via QR code in the museum next to the wall text. While these recommendations might interest curators, there were also arguments from participants 10 and 12 that ArtBot should remain digital so that the technology does not interrupt the in-person museum experience.

No participant selected social media as their favorite experience. The closest positive statement was from P3, who said the AI and social media conditions offered users the greatest "freedom" to look at the artwork. This freedom to observe and make up one's mind was important to the participant. They felt that having the wall text immediately available made them turn to it for answers rather than coming up with their own ideas.

Other participants were more critical of the other conditions after being exposed to ArtBot during the experiment. Participants 9, 12, and 13 had similar reactions after the random order of conditions had them jump from ArtBot to a Digital Collection or Social Media condition. They expressed frustration at being unable to ask questions or discuss the work. They went as far as calling these other conditions "boring" compared to their experiences with the AI.

7.4.3.2 Thoughts on Socratic AI Experience

As the Socratic approach is central to the novelty of our prototype, we gathered our participants' thoughts on the experience. Overall, it appears positive, especially for users who felt less knowledgeable about art. Nonetheless, a few participants expressed their dislike of the questions and their experience of encountering Socratic AI.

Starting with the positive reactions, our users appreciated how ArtBot maintained the conversation. Participants especially enjoyed getting the question at the end of the prompt as they felt it kept the conversation going (P4), and it still provided enough information in the text about the artwork to satisfy the user's questions (P12). The fact that the questions were relevant to the information discussed in the conversation made participants feel that ArtBot was genuinely attempting to get them to reflect and understand the information associated with each artwork (P9 and P11). Some participants noted that the questions invited them to push back against the information provided by ArtBot, leading to both agreements and disagreements between the user and the system. P3 found this enjoyable as it felt that ArtBot was actually processing and responding to user input. Our participants also noted their positive view of the style of questions being asked. P4 noted that the questions ArtBot was asking aligned with questions they were curious about and might have even looked up online if they had been in an exhibit. Similarly, P13 noted how relevant the questions were to a museum setting in that they prompted the user to look at new details in the artwork they might not have seen before.

Not all of our participants felt the same about the questions, especially about ArtBot's role in the conversation. As ArtBot was quite obviously not a person, one participant remarked not feeling any pressure to engage with the questions as they did not come from another human and did not have the social expectation associated with an asked question (P5). A different participant (P9) remarked the opposite and felt the social pressure to answer the questions even though they knew it was not compulsory. Another negative review of the Socratic approach was that it prompted the user to come up with answers. Both P1 and P6 felt that a companion tool like this should just provide factual information at the user's request. They considered ArtBot making statements about the art patronizing and not desirable for the experience.

The Socratic approach led to various reactions, and with the statistical results to back it, we saw that it performed at roughly the same level as the digital collection condition. While it was effective, it may not have been the user experience our participants expected, highlighting this prototype's provocative nature and contributing to our later discussion (Section 7.5.2) on the role of this type of AI interaction.

7.4.3.3 ArtBot's "Personality"

Whether they had a positive view of the Socratic approach, most users felt that ArtBot exhibited certain behaviors. In Section 2.2.3.3, we outlined the selfhood-initiative model [361] and the behaviors associated with each type of critical thinking AI. We are interested in seeing the participants' responses regarding ArtBot's behavior as a potential critical-thinking AI tool.

Some noted that ArtBot could be stubborn to the users' responses and previous parts of the conversation. One participant enjoyed that ArtBot would offer a blatantly different perspective to the users without seeming to contradict the conversation (P3) intentionally. They valued this new view on the work and the considerate nature of the text used in the AI's response. Another participant admired the knowledge of ArtBot while noting that it guides the user towards understanding the artwork (P12). This participant also recognized that ArtBot was trying to gain an understanding of what the user already knew and observed about the work.

Other users observed ArtBot as a knowledgeable guide. For some, this was reassuring when they felt they did not know enough about art to carry the conversation (P4). Other participants used words like "professional" (P10), "art connoisseur" (P11), and "docent" (P9) to describe the behavior and language use of ArtBot during the conversations. They felt that it was using straightforward language to prompt the audience to think differently about the art in front of them. (P9 and P10).

Another set of our participants saw ArtBot more casually; rather than an expert, they saw it as a conversationalist or chat buddy. Since the model runs locally, there would be moments of delay between a user message and an ArtBot response. Many participants (P1-6) requested a buffer icon for future designs, but P13 explicitly described enjoying the buffer period as it made ArtBot seem more human. They described this behavior as akin to someone thinking about their response to a conversation. Another aspect that made participants feel as though ArtBot was a more casual conversationalist was the usage of subjective descriptions of the artwork. P7 noted how it may seem as though ArtBot had its own judgment about certain artworks and was not as objective as a professional curator might be. Indeed, others did not quite see ArtBot as a professional, with a few noting how socially awkward the AI was because of its continuous focus on artwork (P9 and 13). When these participants attempted to get ArtBot to deviate from the conversation about the artwork, it was quite resilient. The questions prompted the users back towards a conversation about art.

Our participants came up with various interpretations when describing their interaction with ArtBot. Depending on how this tool fits into the art analysis landscape, the type of persona it embodies will likely need to be refined. For a more educational context, the casual conversationalist version of ArtBot may not be appropriate. Similarly, a highly serious or professional ArtBot might not be the ideal art companion for someone casually exploring a local gallery on vacation. The context will dictate how effective ArtBot can be for users.

7.5 Discussion

In this work, we present the results of our study comparing ArtBot to digital collections and social media conditions. The study combined quantitative and qualitative analysis of the user responses to the proposed system. Based on the quantitative analysis, the prototype performed comparably to the digital collections condition on self-reported understanding, usage of complex language, and writing proficiency. Additionally, the qualitative results revealed high user satisfaction with the ArtBot, which is highly preferred over the other experiences.

Based on the findings, we discuss the ability of ArtBot to expand access to art analysis and appreciation (Section 7.5.1). We reflect on whether ArtBot delivers on its ability to be more accessible while still upholding the quality of information expected of an art companion. We also connect our findings back to ongoing discussions in the field of HCI on the relationship between user satisfaction and system performance.

We then reflect on the broader topic of ArtBot's position as a critical thinking support AI (Section 7.5.2). This section reviews participants' reactions to the pedagogy of the Socratic method in the ArtBot experience. We also revisit the selfhood-initiative model and see whether user reactions to ArtBot's communication style place it as an "interlocutor" or somewhere else in Ye et al.'s framework [361] discussed in Chapter 2.2.3.3.

Section 7.5.3 explores future work involving ArtBot. We start with the call for additional studies comparing ArtBot to other AI-powered tools. We then continue to discuss ArtBot evolving with slow technology to foster deeper reflection, becoming a digital companion that supports critical thinking across the web, and serving as a tutor that teaches transferable cognitive skills. These possibilities highlight ArtBot's potential to enhance critical thinking in various contexts.

7.5.1 Did it Work?

ArtBot was successful on two fronts: how it performed compared to the other conditions and how our participants received it. We measured performance quantitatively across all the conditions in a within-subjects study and user satisfaction qualitatively with an exit survey at the end of the experiment. Based on the realities of art analysis access for individuals outside of institutions (outlined in Chapter 2), we set out to see if ArtBot can digitally deliver art experiences without sacrificing the depth of information. In other words, we want to see ArtBot outperform social media while performing the same or better than the digital collections experience.

Our results confirm this hypothesis. Our findings identified significant differences between the conditions in terms of self-reported understanding of the artworks viewed. We found that the AI and digital collection conditions had a higher percentage of users reporting a deeper understanding than the social media condition. Simultaneously, ArtBot and the digital collection outperformed the social media condition in the text responses. This suggests that ArtBot offers a unique digital experience that delivers the same level of information as the digital collection condition. This raises the question: why build ArtBot instead of digitizing more museum collections?

We argue that user satisfaction with ArtBot was much higher than with the other conditions. When asked to choose their favorite condition at the end of the study, ten participants praised its ability to offer greater "freedom" (P3) to explore the artwork and to "reveal new aspects of the piece" (P13) that were not immediately obvious. The two who preferred the digital collection condition over the AI preferred immediate information delivery over discovery and conversation (P5 and P6). Overall, the participants expressed satisfaction with ArtBot's ability to deliver engaging art experiences.

HCI literature has a history of discussing the relationship between system performance and user satisfaction [166]. Notably, in 1994, Nielsen et al. and Gatian explored the power of user satisfaction and the instances arguing high user satisfaction often relates to positive outcomes when using a system or device [241, 125]. However, these discussions are often within the context of an underperforming system. In other words, designers and researchers decide whether to recommend a system users enjoy using even though it does not offer the same benefits [137]. We argue that we should still use tools like ArtBot because it does perform at the same level as the digital collection standard **and** it satisfies users.

Based on these results and previous HCI discussions on user satisfaction, ArtBot does work in providing access to art analysis in digital environments without losing the quality expected of a digital collection site.

7.5.2 ArtBot as a Critical Thinking Assistant

Having posited that ArtBot succeeded in delivering an effective algorithmic experience for art analysis in the previous section and accepting previous literature that art analysis is tied to critical thinking [4, 298, 128], ArtBot is also a successful critical thinking assistant. Based on recent work described in Section 2.2.3.3, critical thinking tools likely need to relate to at least one of the theoretical frameworks for this type of AI. Specifically, it is worth exploring how our participants' reactions to the Socratic method relate to the ideas put forth in dialogic pedagogy and Cai et al.'s concept of "antagonistic" AI. Similarly, in Section 7.4.3.3, we report how our participants viewed ArtBot, and now we can compare those results to the selfhood-initiative model outlined by Ye et al. [361].

Most participants engaged with the Socratic method to some extent. Based on their responses, they did describe the AI as occasionally challenging in its position and statements. This reflects the challenge and discomfort outlined by Cai et al. in their theory [58]. In practice, the response was mixed. Some participants appreciated the questions and the difficulty they added to the discussion (P3, 7, 8, 10, 11, 13). On the other hand, some participants expressed dislike of the questions, finding the tone ineffective and unpleasant (P6). One particular participant had a positive interpretation of the method; they found themselves relying on it to guide the conversation when the participant felt unequipped to approach the artwork (P4). From these results, we can see where there are merits in the theoretical proposition of an "antagonistic" AI, but perhaps, as Ye et al. mention, the system's tone does not need to feel antagonistic to get the same positive outcome [361]. Yet, like Danry et al. explored in their paper, the value does come from the question being a part of the AI's response to the user [89].

We had a few instances where our participants discussed not feeling the need or pressure to respond to the questions, which could disrupt the tool's effectiveness. Our participants, P5 and P9, discussed the reality of ArtBot not being a human, which made them question the expectation of having to answer the questions, much less engage. Based on this feedback, if the system relies only on the Socratic method of asking questions, the benefits may not transfer to users if they do not engage. Perhaps future versions of ArtBot need to pull from a broader range of pedagogical practices in Dialogical teaching [303]. This teaching method relies on the whole conversation process, not just questions and answers. In doing so, ArtBot may also walk the line better regarding its tone and user experience.

As Ye et al. describe in their model, an "interlocutor" can still challenge and improve the ideas of a user without needing to be described as "antagonistic" [361]. Based on the results presented in Section 7.4.3.3, our participants describe some degree of selfhood and initiative in ArtBot. The selfhood comes from our participants' description of ArtBot as a docent or connoisseur, which likely stems from the RAG model and dataset as well as the system prompt that dictates ArtBot's functional knowledge. These both contribute to ArtBot having a knowledge base and some degree of a role or identity. The initiative was highlighted in their descriptions of ArtBot as a conversationalist who could sometimes be stubborn and insist on its topics of interest. Based on these responses from our participants and the definitions set forth by Ye et al., we can place ArtBot as an early prototype of an "interlocutor" critical thinking AI.

7.5.3 Art, Bot, & Beyond

In reflecting on the study results, several potential future developments for ArtBot emerge, each connected to different technology implementations. In the immediate future, we might consider the rise of GenAI in all fields. Museums and art scholars are developing their own deployments of LLM systems to support their missions. Future research should compare ArtBot with other AI-powered tools like the Living Museum app [323], various LLM museum guides [334, 329], or CulturAI [84]. As all of these tools are early implementations of novel technology, it is valuable for researchers to gain an understanding of their capabilities and limitations. By comparing performance, future research can see how ArtBot and similar technology will need to evolve as prototypes for better delivery of cultural experiences.

Beyond only comparing ArtBot to other AI technology in the context of museums, we can abstract the concept of ArtBot as an educational tool to other contexts. As an exercise, we can explore how ArtBot might evolve to extend the principles of slow technology, become a consistent digital companion, or serve as a tutor, scaffolding a user's critical thinking across various topics and experiences.

In Section 7.2, we explored the influences of Slow Technology on the design of ArtBot, particularly its speculative design method. Slow Tech focuses on implicit reflection, using temporal shifts within the task or system to subtly nudge users toward introspection [144, 243]. ArtBot, in contrast, adopts a more explicit approach, directly asking questions that challenge users to think critically about the art they are analyzing. Future iterations of ArtBot could bridge these approaches by creating a digital environment where users are guided through a slower, more reflective experience, enabling a deeper connection to both the content and their own thought processes. Some design challenges for this potential future are ensuring the experience is not frustrating or jarring for the user. In this application, the reflection method and balance between implicit and explicit user prompting will be key.

Another version of ArtBot moves the critical thinking AI out of the singular web app experience into a companion accompanying a user as they navigate the Internet. This concept draws inspiration from recent advancements in LLM-powered fact-checking tools [73], which are designed to address misinformation and political polarization [256, 277, 10]. While fact-checking tools focus on delivering accurate information, ArtBot's future role could be more nuanced, actively engaging users to think critically and reflect on the content they encounter. Several participants (P8, P9, P11) noted how ArtBot's direct questioning led them to contemplate their initial interpretations of the art they analyzed, suggesting that this approach could be expanded to broader online experiences. The challenge for this version will be to avoid the pitfalls of intrusive digital companions (such as "Clippy" [320]). Again, some overlap in the technology exists—such as the proposed LLM-powered fact-checking game in HCI literature [324]—where AI tools can better involve users in processing the facts or content they are reviewing. If the future version is done well, it can display the value of critical thinking AI beyond the art use case to include any information or experience online.

The third potential future for ArtBot positions it as a scaffolding tutor, focusing on fostering "enduring cognitive benefits that transcend any particular interactions and tasks" [286]. This future version of ArtBot would aim to equip users with transferable critical thinking skills that they can apply independently in other contexts. One participant (P13) noted that ArtBot "helps people learn how to think," pushing beyond information delivery to develop independent critical thinking. The design challenge for this version of ArtBot requires the development of longitudinal tests that allow researchers to measure cognitive engagement and critical thinking over time and across domains.

Critical thinking AI tools are still in their infancy, and the pedagogical theories surrounding them need further testing and refinement. ArtBot offers multiple avenues for future development as a technology probe, with feedback from participants and other researchers guiding its evolution. Our takeaway lesson would be to develop a critical thinking system that offers the user some degree of delight or interest. Critical thinking involves as much discovery and play as reflection and refinement.

7.6 Limitations and Future Work

Both prototyping and technology probe methods are accompanied by the caveat that the technology they display is limited in scope for exploratory work to gain user insights. Future work that wishes to expand on the technology and the findings with more users will require moving past provotyping and technology probes and into formal iterative designing and developing.

Within the context of this study, the design has some limitations. First, while the participants came from various global regions and had varying amounts of art experience, the artwork used in the study was all Western art gathered from a university art museum in the global north. Similarly, the wall text and information associated with this artwork come from the same context, which may have impacted the participants' responses to the artworks outside the three experimental conditions. Future work can expand the type of art presented from a more global collection to reflect the international user group. Second, the study only used paintings, which is only one type of medium; future work can explore other forms by building in video support. Third, we only replicate one form of social media, not the widely popular short-form video social media. Again, a future iteration of this study could compare ArtBot to this type of interaction.

We are looking forward to future work on ArtBot and related AI technologies. We have made the code open-source and encourage further exploration of critical thinking support and novel art experiences.

7.7 Conclusion

In this final project, we set out to deliver a novel algorithmic experience for art appreciation. The work in this chapter distills the previous project chapters into a single algorithmic system. We presented the development of ArtBot, a Socratic LLM art companion, and tested it compared to digital collection and social media experiences. The findings revealed that ArtBot offers similar benefits to art analysis as a digital collection does with wall text next to an artwork. We argue that the findings support greater access to art appreciation and analysis, which aids individuals with critical thinking. We present ArtBot as an early example of critical thinking support AIs that offer a novel interaction experience requiring users to engage more actively with their own cognitive processes.

8 Discussion

Contents

8.1 All Together Now

Over the course of four projects, this thesis sought to address how algorithmic experiences (AX) influence audiences' perceptions of art and culture. Each project corresponds to a specific research question, with findings from each stage informing and advancing the subsequent inquiry, ultimately working towards answering the overall thesis statement. Through an iterative approach, this thesis began with a broad exploration of how art is perceived on social media and culminates in the development of an experimental prototype that encourages users to form their own interpretations of the art they encounter. We argue that prior algorithmic experiences have hindered audiences from fully appreciating art, as these systems have failed to prioritize art appreciation as a distinct form of engagement. Future AX systems aiming to rectify this must gather and maintain a comprehensive body of art data to create an art-driven, rather than engagement-driven, experience.

This conclusion is drawn from addressing each research question through the four projects. In the first project, we conducted exploratory research to understand how art experts, artists, creatives, and social media users experience art on short-form social media platforms. Through interviews, co-design sessions, and stakeholder reviews, we distilled four key attributes deemed essential for experiencing art: *depth, conversation, connection, and time*. These attributes guided the development and testing of various prototype screens that challenged the conventional algorithmic experience of short-form social media. Ultimately, while human audiences placed value on these four attributes, it was determined that existing recommender systems and algorithmic experiences did not prioritize the same factors.

In the second project, we sought to understand previous attempts at having algorithms interpret art, examining how and why these efforts differed from human expectations. Building on the work of Herman and Moruzzi, we conducted a technical analysis of the *Algorithmic Pedestal* exhibit in London. This physical exhibit was designed to offer audiences the four attributes identified in the first study, while the field of computational aesthetics provided various algorithmic processes to analyze how machines perceive art. We ran several programs to measure the artworks in the exhibit and their corresponding digital collection, conducting statistical comparisons on the metrics. The results showed that the differences human audiences observed between the various parts of the exhibit were not reflected in the software's metrics. By comparing these findings with qualitative responses from audience members, it became clear that they were seeking information about the story behind the art and the selection process—questions closely aligned with the four attributes—which the computational aesthetics algorithms did not address.

The third project aimed to address the missing data identified in the previous project. By scraping public art auction records, we built a new dataset called *AppraiSet*. This dataset contained metadata about each individual auction lot, excluding visual data, which the previous study had shown to be less helpful. Once the dataset was constructed, we tested its usability and quality through a case study. Specifically, we built an LDA (Latent Dirichlet Allocation) topic model to analyze the dataset, assessing whether a computer could classify artworks into aesthetic categories and develop its own descriptive language. The results showed that our LDA topic model was able to generate distinct, coherent topics, demonstrating that art metadata can be effectively interpreted through natural language processing, leading to meaningful groupings of artworks.

The final project synthesized the findings from the three previous projects. Drawing on the four key attributes that were missing in most algorithmic experiences, and utilizing the *AppraiSet* dataset, we developed a prototype LLM (large language model) called ArtBot. ArtBot was designed as a Socratic opponent—a retrieval-augmented generative model that engaged users in dialogue about their thoughts and reactions to the art presented on the screen. We tested ArtBot's effectiveness through a within-subjects experiment, where participants viewed nine pieces of art across three conditions: 1) art with a caption and wall text (digital collection), 2) art with only a caption (social media), and 3) art with a caption and interaction with ArtBot. After each condition, participants were asked to share their thoughts on the art. These responses were scored both computationally and according to a rubric developed by museum educators. Across various metrics, ArtBot performed comparably to the museum condition, with both significantly outperforming the social media condition, suggesting that ArtBot could offer a promising new digital experience for art appreciation.

Through this iterative project approach, we identified several key attributes that audiences value when experiencing art. These attributes are largely absent from current algorithmic experiences. According to our creative participants, social media platforms prioritize engagement metrics over these deeper aspects of art appreciation. Similarly, previous computational approaches to analyzing art have focused primarily on visual data, neglecting the metadata that could support the four key attributes. In our attempt to design a new algorithmic experience that accounted for these factors, we found evidence that a digital experience can match the effectiveness of a digital museum while offering users greater freedom and enjoyment in exploring art. Our findings align with literature referenced in Section 2.2.1.2 describing how users interact with digital collections. This previous work revealed that while users were immediately interested in personalized digital collections, they rarely returned even when they could change and alter the collection at will [213]. Nonetheless, there is evidence in previous literature and our findings that users enjoy a curated digital system that allows for interaction and can pave the way for greater engagement [142]. Our qualitative findings indicate that a "dialogic" [267] experience may spark greater interest for users to have repeated interactions than just collection customization. We, therefore, argue that social media has failed to support art appreciation and that future systems should adopt a metadata-driven approach to foster deeper engagement with both art information and appreciation. Futhermore, future research should compare ArtBot and the designs to other digital heritage AI tools being built [323, 84] to refine further the systems that benefit users.

The remainder of this discussion will reflect on these findings within the broader context of computer science and the arts. First, we will examine the key takeaways of this thesis for researchers, professionals, and everyday internet users. Next, we will explore how this work contributes to and expands upon the psycho-historical theory of art appreciation introduced in Chapter 2. Finally, we will share insights on the broader topic of human-AI collaboration, grounding our discussion in the context and findings of this research.

8.2 Trajectories and Contributions

This thesis begins by outlining the practical, methodological, and theoretical outcomes of the individual projects. Having presented these projects in detail, it is important to highlight how the contributions may impact various audiences. Specifically, we consider the potential value of this work for museums, galleries, designers, developers, and, of course, end users.

8.2.1 Museums and Galleries

The practical applications of this work are particularly relevant to art institutions. Each project chapter offers insights that could better equip curators, gallerists, and artists for a future where AI and algorithmic processing play a growing role in the art world. Chapter 4 introduces workshop materials and design exercises that enable art institutions to reimagine their digital collections and offerings for improved user experiences. This work, developed in collaboration with experts from art institutions and experienced social media creators, ensures that the recommendations are grounded in the lived experiences of the target user group. Through this work, we meet the call to action outline in Section 2.2.1.2 where researchers exploring digital heritage have put conceptual frameworks for developing digital art experiences through a user-centered approach [266]. In their work, they argue that the experience should satisfy, provoke, teach, and provide multiple perspectives [266]—their concept aligns nicely with the findings and designs presented in this chapter. Moreover, by making the work open-source, we encourage others to adapt and improve these systems for broader use.

Chapter 5 explores how the results of technical analysis could be leveraged by curators and artists when presenting art collections. Computational aesthetics can offer valuable insights into the visual characteristics of the art managed by an institution. These methods produce information that can better understand how machines process images and how art might be perceived differently. For example, we examined the "feedback loop" of images containing faces. A common belief is that algorithms favor images with faces [173], though psychological evidence suggests this is because humans are naturally drawn to faces, which in turn reinforces the algorithm's focus [27]. Our research revealed that the artist had selected more pieces featuring faces than the Instagram algorithm did, potentially illustrating this feedback loop. Based on these findings, we suggest that art institutions could strategically promote collections or exhibits featuring works with faces, as these pieces tend to capture both human and algorithmic attention. These findings and recommendations connect back to Whitelaw's discussion, referenced in Section

2.2.1.2, on the new nature of digital collections as datasets and how those in control of the dataset can dictate what message it sends to the audience [354]. Similarly, this relates to other literature from Section 2.3.2.1 describing how social media platforms wield their power over the datasets of content at the detriment of their users [111, 332, 100, 273]. This chapter is meant to show how even well-established algorithmic analysis methods do not always align with user preferences when treating a digital collection like a dataset.

The subsequent chapter introduced *AppraiSet*, our novel art dataset. A key aspect of this project involves the value of digitizing and providing open access to art data—a practice already embraced by many art institutions, which have been working on digitizing their collections for years, even prior to the COVID-19 pandemic [316]. However, our open-source dataset offers structure and guidance on establishing a consistent data format to streamline this process. This chapter discusses the challenges of repetitive and inconsistent data logging, which led to additional data-cleaning work. We hope this example underscores the importance of a clear data strategy during the digitization process. *AppraiSet* fills a gap described by Onuoha [249], and Van Miegroet et al. [333], that high-quality publicly available art sales data is a missing dataset. The dataset and associated case study also challenge computational aesthetics to move beyond the visual analysis of artworks [156, 47] to consider analyzing artworks' written, recorded metadata.

Finally, Chapter 7 presents an example of a novel interactive experience that could enhance digital art offerings from institutions. Art institutions can reimagine both in-person and online experiences by experimenting with the prototype and open-source code, using AI systems to complement their deep expertise in audience engagement. ArtBot is only one interpretation of how AI can augment digital heritage, within the same year multiple systems were released [323, 84, 329, 334] and worth comparing to ensure museums select the tools that best support their mission and audiences.

8.2.2 Designers and Developers

These thesis projects were developed not only for their potential application in the arts and digital humanities but also to address issues within current paradigms for creating digital experiences. Chapter 4 emphasizes the need to move away from "content" as the sole focus, particularly in the context of short-form social media experiences. We highlighted user desires for alternative designs and presented practical screens meant to challenge the existing system. We extend this challenge to designers and developers: while it may be easier to reinforce the profitable monetization models established by social media platforms, users are seeking richer, more meaningful experiences. Whether through decentralized web platforms, a revival of specialized websites and resources, or other yet-to-be-conceived approaches, we encourage designers and developers to continue innovating the user experience.

One particular aspect we challenge in the status quo is the reliance on surface-level quantitative metrics to evaluate created work. In the projects focused on computational processing of art, we not only provide access to code and datasets, but also discuss the need to reconsider which metrics should guide optimization efforts. Chapter 5 shows that focusing too heavily on superficial metrics can distort the computational perspective, shifting it away from what truly matters to users. In our case, this involved artworks, but in other contexts, this could relate to engagement metrics that fail to capture the deeper reasons users enjoy certain experiences. In Chapter 6, we offered an example of how datasets can be built to better include the qualitative details important to users. The advantage of this approach is that it allows the computational system to better align with user expectations and language. However, the downside is the additional workload and complexity involved in curating data appropriately. We argue that a computational system for art should integrate both quantitative and qualitative aspects of dataset development, as explored in these chapters. This can be achieved through advances in visual language models (VLMs) or through entirely new algorithmic experiences like those we explored in the final project.

In the last project, we shared our findings on using a local Retrieval-Augmented Generation (RAG) model to facilitate human-AI collaboration. While our prototype demonstrated promise in its customizability, it struggled with tasks involving vague prompts. In this chapter, we provide evidence of potential pitfalls and how to avoid them, offering insights that could benefit other designers and developers looking to create their own customized AI models.

8.2.3 End Users

We hope that general users and audiences outside of the arts and computer science can also derive valuable insights from this work, particularly in how the concepts discussed in Chapters 4 and 7 apply beyond the realm of art. Chapter 4 emphasizes the risks of reducing "art" to mere "content," and what must be done to resist this transformation. However, art is not the only area affected—politics, journalism, science, and even personal relationships are undergoing similar changes. Social media tends to flatten all subjects into the same format, aiming to reach the widest possible audience. We encourage everyone to push back against this "contentification" of the internet. One way to do this is through critical thinking and by questioning the content presented to you.

Chapter 7 highlights the importance of seeking context and engaging with questions or experiences that challenge the way you think about what you see. We urge readers to recognize the value of critical thinking and consider how different tools and strategies can support them in this process. Our prototype was designed to spark a conversation about what these support tools might look like. As we live in a world with increasing access to information generated by both humans and machines, it's clear that more needs to be done to ensure users can engage meaningfully with this content.

8.3 Expanding Theory of Art Appreciation

In Chapter 2, we introduced the theoretical framework for art appreciation proposed by Bullot & Reber [56]. This framework suggests that art appreciation is a combination of psychological responses to stimuli and the processing of historical or contextual knowledge. Drawing from both psychological and historical perspectives, Bullot & Reber's model has been a central touchpoint throughout this thesis, as we explored the relationship between audiences and algorithms. However, our research presents evidence for an additional dimension that extends their psycho-historical theory.

We argue that art appreciation also involves a social component. Chapter 4 provides the initial evidence for this expanded framework, and further validation is explored in Chapters 5 and 6. Ultimately, Chapter 7 demonstrates how this theory can be practically applied to redesign digital art experiences.

In Chapter 4, qualitative research brought together insights from various user groups. We surveyed the general public and conducted interviews with artists, curators, gallerists, and content creators to understand their perspectives on essential elements of the art experience. Much of their feedback aligned with Bullot & Reber's psycho-historical framework. For example, participants emphasized the importance of *depth* and *connection*, both of which are tied to the historical component of the framework. They noted that context—both factual and personally relevant—plays a crucial role in appreciating art. Moreover, participants stressed that *time* is an essential factor, allowing deeper psychological engagement with a work, echoing Bullot & Reber's assertion that psychological processing is key to art appreciation. However, they also pointed out that social media platforms do not adequately support these dimensions of art appreciation. We argue that this failure makes it difficult to fully appreciate art shared on these platforms.

In addition to these well-established dimensions, participants identified a third crucial aspect of art appreciation: *conversation*. They described the importance of being able to discuss and share thoughts about a work within the context of the art-viewing experience. While social media platforms offer comment sections, both our participants and other researchers have noted the inconsistent and often volatile nature of these spaces [202, 293]. Participants called for a more thoughtfully maintained conversational aspect, something social media platforms currently fail to provide. Importantly, conversation does not fit neatly into the psychological or historical categories of Bullot & Reber's framework. Thus, we argue that the theory must be expanded to account for the social dimension of art appreciation. Audiences not only want to reflect on their psychological responses and the historical context but also seek to discuss and engage socially with the art and with others [192].

Chapter 4 marks the first call to extend the psycho-historical framework to include this social dimension. Given the current state of digital experiences, none of these essential facets—psychological, historical, or social—are being adequately supported. The subsequent chapters of this thesis aim to explore how these gaps can be addressed, offering pathways for digital platforms to foster a richer art appreciation experience.

In the technical analysis of the *Algorithmic Pedestal* presented in Chapter 5, we demonstrate how a computational system might attempt to mimic the psychological process of art appreciation. While computational aesthetics can visually interpret certain aspects of an artwork, this alone is insufficient to replicate the human experience of art appreciation. The differences between how gallerygoers and the computer perceived the artwork were stark. The information that captivated the gallerygoers—about the artwork and its context—was not part of the data being processed by the software. In other words, merely processing visual stimuli, or focusing solely on the psychological aspect of art appreciation, is not enough for a digital system to provide the experience human audiences expect. In this case, Bullot & Reber's framework holds, underscoring that a purely computational approach cannot fully capture the complexity of human art appreciation.

To address the gaps in computational aesthetics, Chapter 6 explores how a digital system can support the historical component of the theoretical framework. By building a dataset based on art auction records, we introduced a wealth of contextual information relevant to both the artwork and its audience. This data, structured according to the standards of art experts, helps people assess the financial and cultural value of a piece. *AppraiSet* also provides algorithms with training data, allowing them to develop a knowledge base and "language" around the artworks.

In this way, the system gains the historical context necessary for art appreciation, as outlined by Bullot & Reber [56]. However, the absence of visual files in the dataset means this approach does not satisfy the psychological component of the framework. For an algorithmic system to fully support art appreciation, it must integrate both visual and textual data—addressing both the psychological and historical dimensions.

ArtBot, introduced in Chapter 7, is our attempt to create an algorithmic experience that meets the requirements of the framework while also offering a practical implementation of the extended theory we propose. Our prototype combines a visual component—an artwork with a label (addressing the psychological aspect)—with an LLM interface built on the *AppraiSet* dataset (addressing the historical aspect). This dual experience demonstrates how Bullot & Reber's theoretical framework can guide the design of new digital experiences [56].

Moreover, we take the framework further by incorporating a Socratic element into the LLM, encouraging users to reflect on their thoughts and engage in a dialogue with ArtBot about the artwork. This interactive process aligns with the attributes of *depth, conversation, connection,* and *time*—the key aspects of art appreciation identified by our participants during earlier research. By embedding these attributes into ArtBot's design, we argue that it exemplifies how technology can support meaningful art appreciation. The results of this project, which show that the prototype performed comparably to traditional museum experiences, provide evidence that a socio-psycho-historical framework for art appreciation can indeed be embedded in an algorithmic system for the benefit of art audiences.

8.4 The Future of Human-AI Collaboration

The narrative of this thesis reflects on how the limitations of both algorithmic and human systems can be addressed through an iterative human-computer interaction (HCI) process. We began by exploring the failures of social media recommender systems to deliver meaningful art experiences (Chapter 4). In our examination of computational approaches to art analysis, we highlighted how algorithms failed to fully account for human perspectives on artworks (Chapter 5). Next, we created a dataset that incorporated human perspectives, but simply building such a dataset is not sufficient for broad accessibility (Chapter 6). Finally, we demonstrated how an AI tool can enhance human reflection and interaction with art (Chapter 7). In each case, both human insight and algorithmic processing offered valuable contributions, and when combined, they led to more positive outcomes.

Throughout the thesis, we have underscored the risks of algorithms making decisions without human input, and the substantial effort required to create datasets and digital systems that accommodate human preferences and cognitive processes. Additionally, in the literature supporting the final project, we discussed the dangers of over-reliance on AI systems, which are increasingly marketed as tools that simplify our lives but often result in reduced cognitive engagement on the part of users [58, 286]. Chapter 7 concludes with discussions on the future of ArtBot, outlining three possible scenarios for how users and ArtBot might interact going forward. The common thread of all three is the question of how future instances of human-AI (HAI) collaboration will unfold.

Human-AI collaboration is not a novel concept. Its origins can be traced back to Licklider's seminal paper on "man-computer symbiosis," which proposed that combining human and machine intelligence would improve outcomes [189]. Modern scholars have expanded on this idea, referring to it as "hybrid intelligence" [94], where the interplay between users and algorithms enhances the capabilities of both. Much of the current research on HAI collaboration focuses on how AI's predictive power can be better integrated into human workflows. Examples of this include decision-support tools used in healthcare, which assist medical professionals in diagnosing illnesses and making informed decisions [181, 350, 257].

In sum, the future of HAI collaboration will likely continue to explore the balance between leveraging AI for its computational strengths and ensuring human cognitive engagement and oversight in decision-making processes. The success of this collaboration will depend on thoughtfully designed systems that account for the unique contributions of both humans and machines.

The integration of AI into human workflows has closely linked HAI collaboration with the field of explainable AI (xAI). When AI systems perform tasks, their outputs must be clear and interpretable for human users. If the results are confusing or misleading, the collaboration may break down. As a result, much of the research on HAI collaboration focuses on improving trust and transparency, both of which are essential for successful partnerships between humans and AI [348, 376]. While this approach is valuable, it can be argued that it positions AI as an almost equal partner to the human user, where the AI's output and ability to explain itself are as important as the human's role in interpreting and using the results [292]. This equivalent relationship may be suitable in some contexts but less so in others.

Some scholars advocate for a less interdependent collaboration model. They use the term "unremarkable AI" to describe situations where AI's role does not need to be central to decision-making [181]. Others argue for a model where HAI collaboration emphasizes augmentation rather than automation, aiming to provide AI support without undermining human agency [168, 146]. However, even in these approaches, there remains a strong emphasis on delegation, with AI often performing tasks on behalf of the user.

This thesis contributes to these ongoing discussions by offering an alternative view of HAI collaboration. Our work on ArtBot, along with related research on AI tools designed to enhance critical thinking, supports a novel form of HAI collaboration. By keeping the human at the center of the process, our approach suggests that AI should empower users to achieve better outcomes. Rather than merely performing tasks, the AI encourages deeper cognitive engagement through reflection, offering alternatives, and even challenging the user's thinking. We argue that meaningful HAI collaboration should extend beyond having AI do tasks humans can already perform, focusing instead on AI helping people accomplish things they could not achieve on their own.

In this thesis, we explored the domain of art appreciation—a complex, subjective, and deeply personal task that resists easy quantification. While no two people experience art in the same way, AI and algorithms can still enhance this process, helping users maintain their connection with the artwork while enriching their experience. As HAI collaboration research continues to evolve, it will likely encounter more tasks like art appreciation, where the balance between human insight and AI support must be carefully considered to ensure a fruitful collaborative experience.

9 Conclusion

This thesis has explored the intersection of digital experiences and art appreciation through a multidisciplinary lens, combining human-centered computing, algorithmic curation, and the psycho-historical framework of art appreciation. By synthesizing the results from four distinct yet interconnected research initiatives, the work demonstrates how digital and algorithmic processes impact both the individual and collective experiences of art.

Chapter by chapter, the thesis has built a coherent narrative that positions algorithmic curation and AI-driven analysis as influential yet sometimes problematic forces in shaping how art is consumed and appreciated. The research has shown that while social media transforms art into consumable content, computational approaches to curating art in physical spaces can misalign with human preferences and expectations. Moreover, the exploration of art datasets has revealed how algorithms can develop their own interpretative frameworks, raising critical questions about the role of AI in contemporary aesthetics.

The development of ArtBot, the LLM art companion, serves as a culmination of this research. It tests the potential of AI not only as a tool for presenting art but also as a medium for enhancing the audience's critical engagement with art. In this way, the thesis proposes new pathways for human-AI collaboration in the art world, with significant implications for artists, curators, researchers, and digital platforms.

Ultimately, this work contributes to a growing body of knowledge at the intersection of art, AI, and digital experiences, highlighting both the opportunities and challenges of integrating technology into the world of art appreciation. It opens the door for future research on human-AI collaboration and its broader impact on how we create, perceive, and engage with art in an increasingly digital world.

Future research could extend the work discussed in Chapter 4 by involving a broader pool of participants. Instead of focusing solely on co-designing with art experts and curators, engaging a more general population may reveal different design requirements. Similarly, expanding the pool of content creators to include those who focus on various forms of entertainment, rather than just artistic content, could provide fresh insights. By altering the participant demographics, we may gain a deeper understanding of how creative work is displayed on social media, leading to novel design solutions that cater to a wider range of users.

In Chapters 5 and 6, we analyze artworks using two distinct datasets and computational models. Each dataset and model offers unique insights into the artworks and how they might be processed and presented to users. However, we have yet to explore the combination of these datasets and models. By applying computational aesthetics to visual file data of artworks and pairing it with the metadata analysis performed using the *AppraiSet* database, we could generate a richer set of insights. This approach would be an innovative way to analyze art data, allowing for the computational exploration of both the psycho-historical aspects and the aesthetic qualities of artworks.

Additionally, ArtBot, as an experimental research artifact, opens up several avenues for future research. One key enhancement we were unable to implement during the project was to host ArtBot on a permanent, live server. Doing so would allow us to collect data from a significantly larger participant pool, while also making the experience accessible to a broader audience. A permanent online presence for ArtBot would facilitate further improvements, transforming it from a research artifact into a fully realized product and resource. Throughout the thesis, discussions have also emerged about deploying the software, datasets, and prototypes through real-world partnerships. These relationships, established during the research process, will continue to evolve beyond the scope of this thesis.

Similarly, several important lessons from the DPhil journey will be carried forward. First, the importance of generous collaboration and proper accreditation cannot be overstated. Our participants emphasized how crucial citation and sharing credit is for artists, curators, and institutions, and this principle applies equally to academia. Research is continually shaped and refined by colleagues, reviewers, and mentors, whose contributions help enhance the final outcomes. These collaborations not only improve the quality of the research but also open doors to future projects and ideas. The second lesson is the value of revision and iteration. While starting with a blank page can be daunting, and a finished draft may seem final, the permanence of any draft is an illusion. It will undergo multiple rounds of review and refinement before being shared with wider audiences. The iterative process of editing and revising is where the true strength of the research emerges, much like the core HCI principle of improving systems through user feedback. Every system can be refined, just as every research idea and paper benefits from ongoing iteration.

Ultimately, these 173 pages encapsulate the research outcomes, highlighting how digital experiences can enhance art appreciation, while also reflecting the personal growth achieved during the doctoral process. The work presented here has been evaluated through peer review and is submitted for consideration in fulfillment of the requirements for the DPhil degree.

Appendices

A
Chapter 4 Study Materials

A.1 Pre-Study Materials

A.1.1 Artist Interview Script

Thank you for taking the time to meet with me for this interview. Our conversation today is in connection with your experience as an artist in relation to art and recommendation algorithms found on various digital platforms. From the summary I sent you by email, do you understand the aims and goals of this project?

Thank you for your interest. Just a reminder, the purpose of this interview is to better understand your thoughts on how the algorithms on digital platforms impact the art world. We will not ask you for any personal details. You can stop this interview at any time or ask for breaks as needed.

Did you get a chance to read through and complete the informed consent document?

- "Yes" - Fantastic, thank you. Did you have any questions before we continue?
- "No" - Please take the time to carefully read and complete it now. If there is anything you disagree with or have questions about, please let me know. If there is anything in that document that makes you want to withdraw from this interview, you can do so.

Before we begin, just one more data protection statement to remind you that you may, without reporting the reason, request that we delete the data you provide to us today for any reason between now and [5 days before the venue deadline].

If this sounds amicable to you, we can begin with the questions.

A.1.1.1 Background Questions

1. Please tell us a little bit about your art. (Medium, subject matter, style?)

2. Do you make art for a living? Or as a hobby? (Can you tell me a bit about your current position and role?)

3. How long have you been making (this kind of) art? How long have you been connected to the art world?

4. How did you train? Do you have formal education in fine art, art theory, museum studies or art history?

   (a) If not, how would you describe your introduction or education in art? Or are you self-taught?

   (b) If yes, do you identify with a particular school, style, or theoretical practice of art history?

5. Do you use social media? If so, which apps/services, and for what purposes?

   (a) Do you prefer to share your work on social media rather than other platforms? Why or why not?

   (b) Do you separate your accounts between your art and personal accounts?

6. Besides social media, do you use any other art-specific forums or sites online to exhibit or promote your work? Which ones are available, and how do you use these sites (in contrast with social media)?

A.1.1.2 Primary Interview Questions

1. In your words, what makes good art good?
   (a) What do you strive for in making your own work good?
2. Do you think art can benefit individuals or society? If so, how?
3. Do you think the public's perception of (good) art has changed over the past 10 years?
   (a) If so, how has it changed?
   (b) Has social media played a role / been influential in this change? How and why do you think so?
4. Has your own work changed over the past 5-10 years, as a result of the internet, social media, and algorithms?
   (a) Does feedback from social media influence the art you create?
      i. Do you consider the number of views/likes you get on your pieces and does this influence the ways you make new art?
      ii. What sort of comments or audience interaction do you get on your art? Have these comments been helpful?
   (b) Do you follow trends, memes, and other kinds of internet culture? Do such elements often influence or inspire your new work?
5. Do you use strategies to make your art more visible, or gain wider reach online? What sort of strategies do you use?
   (a) Do you think current platforms present or curate your work well? Why or why not?
   (b) Do you feel you understand who (what kinds of people) are seeing your art on these sites?
6. Now we are going to you to imagine the following:

(a) If you had a magic wand, how would you make social media better for artists?

(b) If you could better understand or control how people experience your art on social media how would you change this experience?

(c) How else would you shape the next five years to look like for the art world (or your work) in relation to the internet and algorithms?

(d) Do you have any concerns about the future of art on the internet?

7. Is there anyone else you think we should talk to?

A.1.1.3 Conclusion and Debrief

Thank you for taking the time to answer our questions; we appreciate it. The purpose of these questions is to help us better understand the impact algorithms, such as those used in social media, have on the art world and cultural artifacts. We are interested in seeing what you and other experts within the art industry think about the impact algorithms have on people's taste in art and the art itself. We will ensure that there is no collection of any names or personally identifiable information included in our notes and final paper drafts. However, if you have any questions or concerns, you can contact one of us directly at [my university email]. There is also information within the informed consent document providing contact information to the [institution's ethic review body] in case there is anything you want to discuss with them without my knowledge.

A.1.2 Art Expert Interview Script

Thank you for taking the time to meet with me for this interview. Our conversation today is in connection with you experience as a curator in relation to art and recommendation algorithms on various digital platforms. From the summary I sent you by email, do you understand the aims and goals of this project?

Thank you for your interest. Just a reminder, the purpose of this interview is to better understand your thoughts on how the algorithms on digital platforms impact the art world. We will not ask you for any personal details about you. You can stop this interview at any time or ask for breaks as needed.

Did you get a chance to read through and complete the informed consent document?

- "Yes" - Fantastic, thank you. Did you have any questions before we continue
- "No" - Please take the time to read and complete it now carefully. If there is anything you disagree with or have questions about please let me know. If there is anything within that document that makes you want to pull out from this interview, you can do so.

Before we begin, just one more data protection statement to remind you that you may, without reporting the reason, request that we delete the data you provided to us today for any reason between now and [5 days before the venue deadline].

If this sounds amicable to you, we can begin with the questions.

A.1.2.1 Background Questions

1. Can you tell me a bit about your current position and role?
2. How long have you been in this field or connected to the art world?
3. Do you have a formal education in art practice, art theory, museum studies, or art history? If not, how would you describe your introduction or education in art?
    (a) If yes, do you identify with a particular school or theoretical practice of art history?
4. Do you have a particular historical specialty or area of focus (e.g. Contemporary/Modern/Renaissance etc. + Film/Photography/Painting/Drawing/Material Culture etc.)?
5. Do you share or consume art on any digital platforms, and if so, which are the main platforms you use?

A.1.2.2 Primary Interview Questions

1. In your words, what makes good art good?
   (a) How do you identify good art or artists?
   (b) How has the public's current perception of (good) art changed over the past 5-10 years?
   (c) Has your own perception of (good) art changed over the same period? If so, how and why?
   (d) What is the relationship between internet popularity of art and a piece or artist's status in the art world?
      i. For instance, is popular art on the internet good art from an artistic standpoint?
      ii. Is it valuable?
2. Do you think art can benefit individuals? If so, how?
   (a) Do you think art can benefit society at large? How?
3. How has social media affected the art world?
   (a) How would you describe the art predominantly seen across the internet and social media?
   (b) How has it affected the way you present art to the public or students?
4. Have you noticed any effect that digital platforms like YouTube, Spotify, social media have had on the ways in which people engage with art?
   (a) Have you noticed any changes to the format or quality of the art itself?
      i. To what extent do trends on social media shape the art world, and vice versa?
   (b) In your opinion, would you say that the impact of recommendation algorithms has been overall positive or negative? How so?

5. Do you think algorithms would eventually make curators redundant?
   (a) What does a curator offer that a recommender algorithm does not?
6. Are influencers artists?
   (a) Is content equivalent to art?
7. Now we are going to you to imagine the following:
   (a) If you had a magic wand, how would you make social media better for art?
   (b) How else would you shape the next five years to look like for the art world (or your work) in relation to the internet and algorithms?
   (c) Do you have any concerns about the future of art on the internet?
8. Is there anyone else you think we should talk to regarding this topic?
9. Would you be open to joining a co-design workshop with other experts to further explore the future of art and the Internet?

A.1.2.3 Conclusion and Debrief

Thank you for taking the time to answer our questions; we appreciate it. These questions help us better understand the impact algorithms, such as those used in social media, have on the art world and cultural artifacts. We are interested in seeing what you and other experts within the art industry think about the impact algorithms have on people's taste in art and the art itself. We will ensure that no names or personally identifiable information is collected in our notes and final paper drafts. However, if you have any questions or concerns, you can contact one of us directly at [my university email]. There is also information within the informed consent document providing contact information to the [institution's ethic review body] in case there is anything you want to discuss with them without my knowledge.

A.1.3 Social Media User Survey

A.1.3.1 Consent Questions

1. At the time of taking this questionnaire, I am 18 years or older.
   (a) Yes
   (b) No
2. I agree with and allow the researchers to use my anonymous responses to this questionnaire for research purposes.

A.1.3.2 Survey Questions

1. How would you define yourself?
   (a) Artist (I create art/creative works)
   (b) Art Lover (I appreciate art)
   (c) Art Scholar (I study art)
   (d) Art Patron (I support, purchase, or fund art)
   (e) I don't have much interest or involvement in art
2. What is your favorite style/medium of art? Could you also name an artist you admire in that style/medium?
3. In your view, what makes good art good?
4. Do you view art primarily online or offline (1=Entirely Offline, 5=Entirely Online)
5. Do you think the public's perception of (good) art has changed over the past 10 years? If so, how has it changed?
6. If you had a magic wand, how would you make the internet (including social media) better for art?
7. Would you be interested in exploring these in depth in short interviews? If so, please (send email/type your email address here)

A.2 Co-Design Materials

A.2.1 Co-Design Welcome Package

Thank you for agreeing to take part in our co-design workshops! We look forward to your thoughts and perspectives on how digital spaces can better support art and culture.

A.2.1.1 Mental Exercise in Preparation for the Workshop

Think about how you see art presented online on social media, artist's websites, or even digital galleries and collections, then ask yourself the following questions.

1. What about the art viewing experience stands out to you?

2. What about the viewing experience frustrates you?

3. Would you change anything?

A.2.1.2 Miro Introduction

During our co-design workshop, we will use an online Whiteboarding Tool called Miro (anonymized link). This is a great tool used to explore ideas collaboratively. Please see Figure A.1 for an overview of the tool.

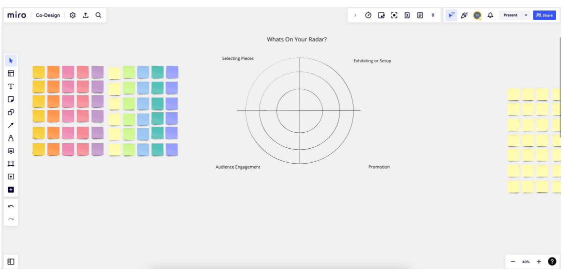

Figure A.1: A snapshot of the miro board our participants would encounter in the co-design session.

The left-hand side has all of the tools you might need (you will not need to use any of them). You will need the ability to drag and drop, which you can do with your mouse.

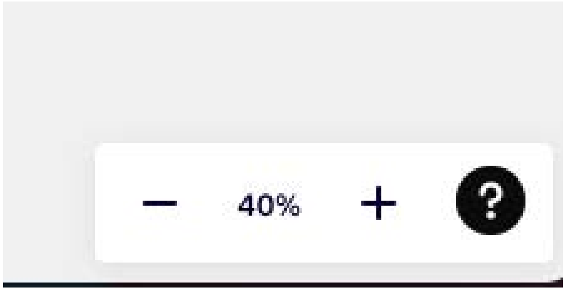

Figure A.2: A cropped image of the zoom and navigation tools in Miro.

Figure A.3: An image depicting the sticky notes already generated in the Miro board for the participants to use.

You might need these buttons (Figure A.2) in the lower right-hand corner to zoom in and out of the board. This can be helpful in case you get lost.

We will do almost all our work with these brightly coloured sticky notes. You can click them and start typing immediately and then drag and drop them wherever you need to place them.

If you have any problems, you can tell us, and we (the research team) will do our best to be your hands and do the typing and sticky note placement for you.

We will also do one exercise in the platform to give you some practice at the start of the workshop.

A.2.2 Co-Design Script

A.2.2.1 Introduction

Welcome to this University of Oxford research study on the impact of algorithmic recommendation systems on art and culture. For the next two hours, we will collaborate through various activities to co-design potential guidelines and solutions for how art is presented online.

The motivation of today's co-design is a phenomenon described by media studies as "contentification", a trend where everything that is shared and posted online, especially on social media, is labelled and treated as content without any differentiation. Our focus is on how this phenomenon impacts artworks and cultural expressions.

That brings us to you; all of you work as curators, gallerists, and art experts. Your job has been to elevate artefacts from the vast landscape of human creativity and present them to audiences worldwide. You have experience in the essential reversal of "contentification". We want to hear from you about your process of selecting pieces and how you recommend things change in the future.

We have four activities planned out, which we will guide you through. If you have any questions feel free to ask at any point. You are encouraged to discuss and share your thoughts as you go; since this workshop has audio recordings and the researchers are presented, we can capture important notes. There are no wrong answers or bad ideas. We want to hear your thoughts on the current situation and how you see the future.

A.2.2.2 Activity 1

Our first activity asks you to share your experiences on four topics. The topics are: Selecting Pieces, Exhibiting Pieces, Promoting Pieces, and Understanding Audience Engagement.

The radar design separates the topics and prioritises your notes from most important (centre) to least important (edge of the circle). For example, when selecting a piece, you might say the artist's name is the most important and put that sticky note here, while the school of work is least important and would go here. Each of you can place your thoughts as you see fit; we will take some time to discuss them at the end. Again since this activity is based on your lived experiences, there are no wrong answers.

This activity will take about 10 minutes, and we'll have 10 minutes of discussion afterwards.

A.2.2.3 Activity 2

The second activity starts to look towards the future. We have three different types of venues for art that you might have some experience with In-person galleries, online galleries, and hybrid set-ups.

We have also included five pillars that have come out of our interview study with many of you. These five pillars were highlighted as essential for promoting artistic and cultural experiences. They are Critical engagement, depth of information, conversations with others, connection to the piece, and time to spend with the piece. We will now develop ideas for promoting these pillars within the various settings. For example, to promote engagement in an online gallery, we may allow audiences to link their personal profiles. Again since we are just trying to come up with as many potential ideas as possible, feel free to put down anything, even if you are unsure how it would be built or implemented.

This activity will take about 10 minutes, and we'll have 10 minutes of discussion afterwards.

A.2.2.4 Activity 3

This third activity allows us to reflect on the previous two and make recommendations for two different populations.

The first population is artists; what are the do's and don't's you would provide an artist to help them ensure their work does not get treated like every other piece of content? How can they stand out?

Once we've done that we will move onto Social Media designers. What can they learn from you and your experiences as curators and gallerists to ensure that art and cultural artifacts are not grouped with all other content? We will develop a set of unified guidelines based on the responses we come up with. Feel free to discuss, and we will do our best to type and turn those ideas into actionable guidelines.

This activity will take about 10-15 minutes, and we'll have 5-10 minutes of discussion afterwards.

A.2.2.5 Activity 4

Our final activity will be predominantly discussion based. As we finalize the list of guidelines for social media designers, we will try to connect each item of the guidelines with a specific interface element that will be shown to users on the screen.

The goal of this task is to come up with a very low-level idea of what a potential digital experience might look like.

This activity will take about 20 minutes, and we'll have approximately 20 minutes of discussion afterwards.

A.2.2.6 Conclusion

Thank you all for taking the time to meet with us today. Your responses and discussions have been really insightful and will greatly inform our research.

The next steps will be as follows, and we will send you some of these screens and guidelines that your colleagues in other workshops have developed. Please send back your critiques and thoughts on those research outputs; they will do the same for what we have come up with today.

From there, we will take those notes and create one more version of the guidelines and screen designs. These will all be packaged together in a zip file which will become available through an open-source repository and will have its own DOI and include all your names as contributors.

As researchers, we will use these designs and findings for the next phase of our research, which will be included in the final paper and thesis work.

Please let us know if you have any questions or concerns, and we will be in touch with updates as this research gets released into the world.

A.3 Stakeholder Review

1. Do you consider yourself an artist or a content creator?
   (a) Artist
   (b) Content Creator
2. What makes good art good?
3. What makes good content good?
4. What is the difference between art and content?
5. Should good (content) have cultural value?
6. What do you think of the values (Depth, Conversation, Connection, and Time) shared by curators?
7. What do you think of the screens (show the different screens) developed by our curators?
8. How would you reconfigure the screens? How would you reconfigure the algorithmic experience?
9. You described different values that you adhere to in your posting, what does the ideal algorithmic experience look like for you?

References

[1] Explained: What is the Art World? A Complete Overview of the Different Actors & Entities, 5 2023.

[2] The Swearing In | Jacob Lawrence | RoGallery, 2023.

[3] Andreas Aakerberg, Kamal Nasrollahi, and Thomas B Moeslund. RELLISUR: A Real Low-Light Image Super-Resolution Dataset. *35th Conference on Neural Information Processing Systems (NeurIPS 2021) Track on Datasets and Benchmarks*, 2021.

[4] Catharine Abell. Art: What it is and why it matters. *Source: Philosophy and Phenomenological Research*, 85:671–691, 2012.

[5] Bereket Abera Yilma, Najib Aghenda, Marcelo Romero, Yannick Naudet, Hervé Panetto, and Hervé Panetto Personalised. Personalised visual art recommendation by learning latent semantic representations. 2020:1–6, 2020.

[6] Adobe. Photoshop generative fill.

[7] June Ahn, Jason Yip, and Michael Gubbels. SINQ: Designing Social Media to Foster Everyday Scientific Inquiry for Children. *Interaction Design Conference (IDC)*, 2013.

[8] Sarah Fdili Alaoui. Making an interactive dance piece: Tensions in integrating technology in art. *DIS 2019 - Proceedings of the 2019 ACM Designing Interactive Systems Conference*, pages 1195–1208, 6 2019.

[9] Patricia Alkhoven. Digitizing cultural heritage collections: The importance of training. In *Association for History and Computing*, 2005.

[10] Hunt Allcott, Matthew Gentzkow, and Chuan Yu. Trends in the diffusion of misinformation on social media. *Research and Politics*, 6(2), 4 2019.

[11] Abdullah Alsalemi, Christos Sardianos, Faycal Bensaali, Iraklis Varlamis, Abbes Amira, and George Dimitrakopoulos. The Role of Micro-Moments: A Survey of Habitual Behavior Change and Recommender Systems for Energy Saving. *IEEE Systems Journal*, 13(3):3376–3387, 2019.

[12] Ferran Altarriba Bertran, Soomin Kim, Minsuk Chang, Ella Dagan, Jared Duval, Katherine Isbister, and Laia Turmo Turmo Vidal. Social Media as a Design and Research Site in HCI: Mapping out Opportunities and Envisioning Future Uses. *Conference on Human Factors in Computing Systems - Proceedings*, 5 2021.

[13] Oscar Alvarado and Annika Waern. Towards Algorithmic Experience: Initial Efforts for Social Media Contexts. *CHI '18: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 2018.

[14] Christian Ulrik Andersen and Soren Bro Pold. Chapter 2 The Metainterface Industry: New Platforms for Culture. In *The Metainterface Industry The Art of Platforms, Cities, and Clouds*. 2018.

[15] Christian Ulrik Andersen and Soren Bro Pold. Introduction. In *The Metainterface Industry The Art of Platforms, Cities, and Clouds*. 2018.

[16] Christian Ulrik Andersen, Søren, and Bro Pold. Post-digital Books and Disruptive Literary Machines Digital Literature Beyond the Gutenberg and Google Galaxies. 2014.

[17] Chris Anderson. *The LONG TAIL Why the Future of Business Is Selling Less of More*. Hyperion e-books, 2008.

[18] Jaynie Anderson. Rewriting the history of art patronage. *Renaissance Studies*, 10(2):129–138, 1996.

[19] Arturo Arriagada and Francisco Ibáñez. "You Need At Least One Picture Daily, if Not, You're Dead": Content Creators and Platform Evolution in the Social Media Ecology. *Social Media and Society*, 6(3), 7 2020.

[20] artnet. Price Database. .

[21] Yasamin Asadi, Margot Brereton, Aloha Hufana Ambe, Alessandro Soro, and Mahnoosh Kholghi. Calming Down in Lockdown: Rethinking Technologies for a Slower Pace of Life. pages 1002–1014. Association for Computing Machinery (ACM), 7 2023.

[22] James Auger. Speculative Design: Crafting the Speculation. *Digital Creativity*, 3 2013.

[23] Multiple Authors. Ollama. `https://github.com/ollama/ollama`, 2023.

[24] Ricardo Baeza-Yates. Data and algorithmic bias in the web. In *WebSci '16*, pages 1–1. Association for Computing Machinery (ACM), 5 2016.

[25] Jeeyun (Sophia) Baik. Data privacy against innovation or against discrimination?: The case of the california consumer privacy act (ccpa). *Telematics and Informatics*, 52:101431, 2020.

[26] Jason Bailey. Machine Learning for Art Valuation. An Interview with Ahmed Hosny. — Artnome, 2017.

[27] Saeideh Bakhshi, David A. Shamma, and Eric Gilbert. Faces engage us: photos with faces attract more likes and comments on instagram. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, page 965–974, New York, NY, USA, 2014. Association for Computing Machinery.

[28] Judith Huggins Balfe, editor. *Paying the Piper Causes and Consequences of Art Patronage*. University of Illinois Press, Chicago, 1993.

[29] David Balzer. *Curationism : how curating took over the art world and everything else*. Pluto Press, 2015.

[30] Shivam Bansal and Chaitanya Aggarwal. textstats. `https://github.com/textstat/textstat`, 2014.

[31] Yan Bao, Taoxi Yang, Xiaoxiong Lin, Yuan Fang, Yi Wang, Ernst Pöppel, and Quan Lei. Aesthetic Preferences for Eastern and Western Traditional Visual Art: Identity Matters. *Frontiers in Psychology*, 7(OCT):1596, 10 2016.

[32] Jeffrey Bardzell. Creativity in Amateur Multimedia: Popular Culture, Critical Theory, and HCI. *Human Technology*, 3(1):12–33, 2007.

[33] Jeffrey Bardzell. Interaction Criticism and Aesthetics. 2009.

[34] Jeffrey Bardzell, Shaowen Bardzell, and Lone Koefoed Hansen. Immodest proposals: Research through design and knowledge. In *Conference on Human Factors in Computing Systems - Proceedings*, volume 2015-April, pages 2093–2102. Association for Computing Machinery, 4 2015.

[35] Shaowen Bardzell, Daniela K Rosner, and Jeffrey Bardzell. Crafting Quality in Design: Integrity, Creativity, and Public Sensibility. 2012.

[36] Terry Barrett. Criticizing Art. In *Criticizing Art Understanding the Contemporary*, chapter excerpt. Mayfield Publishing Company, Mountain View, 1994.

[37] Belén Barros Pena, Nelya Koteyko, Martine Van Driel, Andrea Delgado, and John Vines. "My Perfect Platform Would Be Telepathy" - Reimagining the Design of Social Media with Autistic Adults. *Conference on Human Factors in Computing Systems - Proceedings*, 16(23), 4 2023.

[38] Walter Benjamin. The Work of Art in the Age of Mechanical Reproduction. *Illuminations*, 1935.

[39] John Berger. *Ways of seeing.* Penguin modern classics. London, 2008.

[40] Enrico Bertacchini and Federico Morando. The future of museums in the digital age: New models of access and use of digital collections. *International Journal of Arts Management*, 15, 1 2011.

[41] Bernadette Biedermann. 'virtual museums' as digital collection complexes. a museological perspective using the example of hans-gross-kriminalmuseum. *Museum Management and Curatorship*, 32:281–297, 5 2017.

[42] Reuben Binns. Human judgment in algorithmic loops: Individual justice and automated decision-making. *Regulation Governance*, 16:197–211, 1 2022.

[43] David M Blei, Andrew Y Ng, and Michael I Jordan. Latent Dirichlet Allocation. *Journal of Machine Learning Research*, 3:993–1022, 2003.

[44] Blouin. Blouin Art Sales Index.

[45] Mark Blythe. The Digital Music Box: Using Cultural and Critical Theory to Inform Design. *CHI'07*, pages 2297–2302, 2007.

[46] Mark Blythe, Ann Light, and Shaleph O'Neill. Untitled: Emerging Cultural Forms in the Digital Age. *Human Technology: An Interdisciplinary Journal on Humans in ICT Environments*, 3(1):4–11, 2 2007.

[47] Yihang Bo, Jinhui Yu, and Kang Zhang. Computational aesthetics and applications. *Visual Computing for Industry, Biomedicin, and Art*, 1, 2018.

[48] Susanne Bødker, Christian Dindler, Kim Halskov, and Ole Sejer Iversen. Advances in participatory design. *Conference on Human Factors in Computing Systems - Proceedings*, 07-12-May-2016:984–987, 5 2016.

[49] Laurens Boer and Jared Donovan. Provotypes for participatory innovation. *Proceedings of the Designing Interactive Systems Conference (DIS 2012)*, pages 388–397, 2012.

[50] D Bondy Valdovinos Kaye, Aleesha Rodriquez, Katrin Langton, and Patrik Wikström. You Made This? I Made This: Practices of Authorship and (Mis)Attribution on TikTok. *International Journal of Communication*, 15:3195–3215, 2021.

[51] Judy Borowski, Christina M Funke, Karolina Stosio, Wieland Brendel, T Wallis, and Matthias Bethge. The notorious difficulty of comparing human and machine perception. In *2019 Conference on Cognitive Computational Neuroscience*, pages 2019–1295, 2019.

[52] Pierre Bourdieu. *The Field of Cultural Production*. Columbia University Press, 1993.

[53] Carla Brenner, Heidi Hinish, and Barabar Moore. The Art of Romare Bearden: A Resource for Teachers. Technical report, National Gallery of Art, Washington D.C., 2003.

[54] Taina Bucher. Want to be on the top? Algorithmic power and the threat of invisibility on Facebook. *new media & society*, 14(7):1164–1180, 2012.

[55] Taina Bucher. Affective Landscapes. In *If...Then: Algorithmic Power and Politics*, pages 1–200. Oxford University Press, 7 2018.

[56] Nicolas J Bullot and Rolf Reber. The Artful mind meets art history : toward a psycho-historical framework for the science of art appreciation. *Behavioral and Brain Sciences*, 36(02):123–137, 2013.

[57] Robin Burke. Evaluating the Dynamic Properties of Recommendation Algorithms. 2010.

[58] Alice Cai, Ian Arawjo, and Elena L. Glassman. Antagonistic ai. 2 2024.

[59] Fiona Cameron. Digital futures i: Museum collections, digital technologies, and the cultural construction of knowledge. *Curator: The Museum Journal*, 46:325–340, 7 2003.

[60] John Canny. A Computational Approach to Edge Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-8(6):679–698, 1986.

[61] Baptiste Caramiaux and Sarah Fdili Alaoui. "explorers of unknown planets": Practices and politics of artificial intelligence in visual arts. *Proceedings of the ACM on Human-Computer Interaction*, 6, 11 2022.

[62] Noah Castelo, Maarten W. Bos, and Donald R. Lehmann. Task-dependent algorithm aversion. *Journal of Marketing Research*, 56:809–825, 10 2019.

[63] Oscar Celma and Pedro Cano. From hits to niches? or how popular artists can bias music recommendation and discovery. In *Proceedings of the 2nd KDD Workshop on Large-Scale Recommender Systems and the Netflix Prize Competition*. ACM Press, 2008.

[64] Eva Cetinic and James She. Understanding and Creating Art with AI: Review and Outlook. 2 2021.

[65] Anupam Chander. The Racist Algorithm? *Michigan Law Review*, (115.6):1023, 2017.

[66] Dimitrios Charitos, George Lepouras, Costas Vassilakis, Vivi Katifori, Anna Charissi, and Leda Halatsi. Designing a virtual museum within a museum.

[67] Dimitra Chasanidou, Andrea Alessandro Gasparini, and Eunji Lee. Design Thinking Methods and Tools for Innovation. *Design, User Experience, and Usability: Design Discourse: 4th International Conference, DUXU 2015, Held as Part of HCI International 2015*, 2015.

[68] Sarah E Chasins, Maria Mueller, and Rastislav Bodik. Rousillon: Scraping Distributed Hierarchical Web Data. *UIST'18*, 2018.

[69] Si Chen, Xinyue Chen, Zhicong Lu, and Yun Huang. "My Culture, My People, My Hometown": Chinese Ethnic Minorities Seeking Cultural Sustainability by Video Blogging. *Proceedings of the ACM on Human-Computer Interaction*, 7(CSCW1), 4 2023.

[70] Yuexi Chen, Vlad I Morariu, Anh Truong, and Zhicheng Liu. TutoAI: a cross-domain framework for AI-assisted mixed-media tutorial creation on physical tasks. *Proceedings of the CHI Conference on Human Factors in Computing Systems*, pages 1–17, 5 2024.

[71] John Cheney-Lippold. A New Algorithmic Identity Soft Biopolitics and the Modulation of Control. *Culture & Society*, 28(6):164–181, 2011.

[72] J P Peter Kincaid Robert Fishburne Jr Richard L Rogers Brad S Chissom. Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel. 1975.

[73] Eun Cheol Choi and Emilio Ferrara. Automated claim matching with large language models: Empowering fact-checkers in the fight against misinformation. *WWW 2024 Companion - Companion Proceedings of the ACM Web Conference*, pages 1441–1449, 5 2024.

[74] Yoonseo Choi, Min Kyung Lee, Eun Jeong Kang, and Juho Kim. Creator-friendly Algorithms: Behaviors, Challenges, and Design Opportunities in Algorithmic Platforms. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*, 2023.

[75] Dimitra Christidou. Re-introducing visitors: Thoughts and discussion on john falk's notion of visitors' identity-related visit motivations. *Papers from the Institute of Archaeology*, 20:111–122, 2010.

[76] Xu Chu, Ihab F. Ilyas, Sanjay Krishnan, and Jiannan Wang. Data cleaning: Overview and emerging challenges. *Proceedings of the ACM SIGMOD International Conference on Management of Data*, 26-June-2016:2201–2206, 6 2016.

[77] Jason Chuang, Christopher D Manning, and Jeffrey Heer. Termite: Visualization techniques for assessing textual topic models. *AVI'12*, 2012.

[78] Luigina Ciolfi and Liam J Bannon. Designing Interactive Museum Exhibits: Enhancing visitor curiosity through augmented artefacts . *Eleventh European Conference on Cognitive Ergonomics*, 2002.

[79] Christer Clerwall. Enter the robot journalist. *Journalism Practice*, 8:519–531, 2 2014.

[80] Paul Cobb. *Supporting the Improvement of Learning and Teaching in Social and Institutional Context*, pages 455–478. Psychology Press, 6 2001.

[81] Gilbert Cockton, Shaowen Bardzell, Mark Blythe, and Jeffrey Bardzell. Can We all Stand Under Our Umbrella? The Arts and Design Research in HCI. In *CHI'10: SIGs Session 1*, Atlanta, 4 2010.

[82] Meri Coleman and T. L. Liau. A computer readability formula designed for machine scoring. *Journal of Applied Psychology*, 60:283–284, 4 1975.

[83] Natalie Collie and Caroline Wilson-Barnao. Playing with TikTok: algorithmic culture and the future of creative work. In Greg Hearn, editor, *The Future of Creative Work Creativity and Digital Disruption*, pages 172–188. 2020.

[84] Nicolas Constantinides, Argyris Constantinides, Dimitrios Koukopoulos, Christos Fidas, and Marios Belk. Culturai: Exploring mixed reality art exhibitions with large language models for personalized immersive experiences. *UMAP 2024 - Adjunct Proceedings of the 32nd ACM Conference on User Modeling, Adaptation and Personalization*, pages 102–105, 6 2024.

[85] Michael Cook and Simon Colton. Generating code for expressing simple preferences: Moving on from hardcoding and randomness. *Proceedings of the Sixth International Conference on Computational Creativity*, 6 2015.

[86] UK Design Council. Framework for Innovation - Double Diamond.

[87] Martín Cáceres, Miguel Nussbaum, and Jorge Ortiz. Integrating critical thinking into the classroom: A teacher's perspective. *Thinking Skills and Creativity*, 37:100674, 9 2020.

[88] Edgar Dale and Jeanne S. Chall. *A Formula for Predicting Readability*, volume 27. Ohio State University, Bureau of Educational Research, 1948.

[89] Valdemar Danry, Pat Pataranutaporn, Yaoli Mao, and Pattie Maes. Don't just tell me, ask me: Ai systems that intelligently frame explanations as questions improve human logical discernment accuracy over causal ai explanations. *Conference on Human Factors in Computing Systems - Proceedings*, page 13, 4 2023.

[90] Arthur C. Danto. *What Art Is*. Yale University Press, 3 2014.

[91] Kohinoor M. Darda and Emily S. Cross. The role of expertise and culture in visual art appreciation. *Scientific Reports 2022 12:1*, 12(1):1–25, 6 2022.

[92] Ankolika De and Zhicong Lu. #PoetsOfInstagram: Navigating The Practices And Challenges Of Novice Poets On Instagram. *ACM CHI 2024. In Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (CHI'24)*, 1, 2 2024.

[93] Alain De Botton and John Armstrong. *Art as therapy.* London, 2013.

[94] Dominik Dellermann, Adrian Calma, Nikolaus Lipusch, Thorsten Weber, Sascha Weigel, and Philipp Ebel. The future of human-ai collaboration: A taxonomy of design knowledge for hybrid intelligence systems.

[95] Michael A Devito, Jeremy Birnholtz, Jeffery T Hancock, Megan French, and Sunny Liu. How People Form Folk Theories of Social Media Feeds and What It Means for How We Study Self-Presentation. *CHI '18: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 2018.

[96] Michael A Devito, Darren Gergle, and Jeremy Birnholtz. "Algorithms ruin everything": #RIPTwitter, Folk Theories, and Resistance to Algorithmic Change in Social Media. *CHI '17: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 2017.

[97] Berkeley J. Dietvorst, Joseph P. Simmons, and Cade Massey. Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General*, 144:114–126, 2015.

[98] Emily Dinsdale. How alternative magazines confronted and transformed the world | Dazed, 12 2021.

[99] Steven P Dow, Anand Kulkarni, Scott R Klemmer, and Björn Hartmann. Shepherding the Crowd Yields Better Work. *CSCW'12 Companion: Companion Publication of the 2012 Conference on Computer Supported Cooperative Work and Social Computing*, 2012.

[100] Brooke Erin Duffy, Annika Pinch, Shruti Sannon, and Megan Sawey. The Nested Precarities of Creative Labor on Social Media. *Social Media + Society*, 7(2):1–12, 2021.

[101] Georges Duhamel. *Scenes De La Vie Future.* Paris, 1930.

[102] Paul Duncum. Youth on YouTube: Prosumers in a Peer-to-Peer Participatory Culture. *International Journal of Arts Education*, pages 24–39, 2011.

[103] Brianna Dym, Jed R Brubaker, Casey Fiesler, and Bryan Semaan. Coming Out Okay": Community Narratives for LGBTQ Identity Recovery Work. *ACM Hum.-Comput. Interact*, 3(CSCW), 2019.

[104] Daniel C. Edelson. Design research: What we learn when we engage in design, 2002.

[105] Ahmed Elgammal, Bingchen Liu, Mohamed Elhoseiny, and Marian Mazzone. Can: Creative adversarial networks generating "art" by learning about styles and deviating from style norms *. *International Conference on Computational Creativity (ICCC)*, 2017.

[106] Ziv Epstein, Aaron Hertzmann, the Investigators of Human Creativity, Memo Akten, Hany Farid, Jessica Fjeld, Morgan R. Frank, Matthew Groh, Laura Herman, Neil Leach, Robert Mahari, Alex "Sandy" Pentland, Olga Russakovsky, Hope Schroeder, and Amy Smith. Art and the science of generative AI. *Science*, 380(6650):1110–1111, 6 2023.

[107] European Parliament and Council of the European Union. Regulation (EU) 2016/679 of the European Parliament and of the Council.

[108] John H. Falk. Identity and the museum visitor experience. *Identity and the Museum Visitor Experience*, pages 1–301, 6 2016.

[109] Daniel Fallman. Design-oriented Human-Computer Interaction. *CHI '03: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2003.

[110] Jason Farago. The Myth of North America, in One Painting - The New York Times, 11 2020.

[111] Adam Fish and Ramesh Srinivasan. Digital labor is the new killer app. *New Media & Society*, 14(1):137–152, 2011.

[112] Daniel M Fleder and Kartik Hosanagar. Blockbuster Culture's Next Rise or Fall: The Impact of Blockbuster Culture's Next Rise or Fall: The Impact of Recommender Systems on Sales Diversity. *Management Science*, 55(5):697–712, 2009.

[113] Rudolph Flesch. A new readability yardstick. *Journal of Applied Psychology*, 32:221–233, 6 1948.

[114] Claudia Flores-Saviaga, Joseph Seering, Jessica Hammer, Stuart Reeves, Juan Pablo Flores, and Saiph Savage. Audience and streamer participation at scale on twitch. *HT 2019 - Proceedings of the 30th ACM Conference on Hypertext and Social Media*, pages 277–278, 9 2019.

[115] Valentina Fois. The Prosumer: The Key Player of the Museum of the Future. 2015.

[116] Department for Education. National curriculum - art and design key stage 3, 9 2013.

[117] Sophie Freeman, Martin Gibbs, and Bjorn Nansen. Personalised But Impersonal: Listeners' Experiences of Algorithmic Curation on Music Streaming Services. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*, 2023.

[118] Christian Fuchs. A contribution to the critique of the political economy of transnational informational capitalism. *Rethinking Marxism*, 21(3):387–402, 7 2009.

[119] Christian Fuchs. CLASS AND EXPLOITATION ON THE INTERNET. In *Digital labor: The internet as playground and factory*, pages 211–224. 2012.

[120] Christina M. Funke, Judy Borowski, Karolina Stosio, Wieland Brendel, Thomas S.A. Wallis, and Matthias Bethge. Five points to check when comparing visual perception in humans and machines. *Journal of Vision*, 21:1–23, 3 2021.

[121] Jacques-Emmanuel Galimard, Sylvie Chevret, Emmanuel Curis, and Matthieu Resche-Rigon. Heckman imputation models for binary or continuous MNAR outcomes and MAR predictors. *BMC Medical Research Methodology*, 18:90, 2018.

[122] Catherine A. Gao, Frederick M. Howard, Nikolay S. Markov, Emma C. Dyer, Siddhi Ramesh, Yuan Luo, and Alexander T. Pearson. Comparing scientific abstracts generated by chatgpt to real abstracts with detectors and blinded human reviewers. *npj Digital Medicine 2023 6:1*, 6:1–5, 4 2023.

[123] Catherine A. Gao, Frederick M. Howard, Nikolay S. Markov, Emma C. Dyer, Siddhi Ramesh, Yuan Luo, and Alexander T. Pearson. Comparing scientific abstracts generated by ChatGPT to real abstracts with detectors and blinded human reviewers. *npj Digital Medicine 2023 6:1*, 6(1):1–5, 4 2023.

[124] Urbi Garay. Determinants of art prices and performance by movements: Long-run evidence from an emerging market | Elsevier Enhanced Reader. *Journal of Business Research*, 127:413–426, 4 2021.

[125] Amy W. Gatian. Is user satisfaction a valid measure of system effectiveness? *Information Management*, 26:119–131, 3 1994.

[126] Berys Gaut. A RT A N D E T H I C S The issues. Technical report, 2005.

[127] William Gaver. What Should We Expect From Research Through Design? *CHI '12: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2012.

[128] George Geahigan. Conceptualizing art criticism for effective practice. *Source: The Journal of Aesthetic Education*, 30:23–42, 1996.

[129] Elisa Giaccardi, editor. *Heritage and Social Media*. Routledge, Abingdon, 2012.

[130] Tarleton Gillespie, Pablo J. Boczkowski, and Kirsten A. Foot. The Relevance of Algorithms. In *Media Technologies: Essays on Communication, Materiality, and Society*. The MIT Press, 9 2014.

[131] Google. Stakeholder Review.

[132] Rachel Gordon. Algorithm finds hidden connections between paintings at the Met | MIT News | Massachusetts Institute of Technology, 7 2020.

[133] Kathryn Graddy and Carl Lieberman. Death, Bereavement, and Creativity. 2016.

[134] Alex Greenberger. Artist wins photography contest after submitting ai-generated image. *ARTnews*, 4 2023.

[135] Alex Greenberger. Artist Wins Photography Contest After Submitting AI-Generated Image, 4 2023.

[136] Ulrike Gretzel and Daniel R Fesenmaier. Persuasion in Recommender Systems. *International Journal of Electronic Commerce*, 11(2):81–100, 2006.

[137] Jillian R. Griffiths, Frances Johnson, and Richard J. Hartley. User satisfaction as a measure of system performance. *Journal of Librarianship and Information Science*, 39:142–152, 9 2007.

[138] Kirsikka Grön and Matti Nelimarkka. Party Politics, Values and the Design of Social Media Services Implications of political elites' values and ideologies to mitigating of political polarisation through design. *Proc. ACM Hum.-Comput. Interact.*, 4, 2020.

[139] Aline Guillermet. Vera molnar's computer paintings. *Representations*, 149:1–30, 2 2020.

[140] Dinan Gunawardena, Thomas Karagiannis, Alexandre Proutiere, and Milan Vojnovic. Characterizing podcast services: Publishing, usage, and dissemination. *Proceedings of the ACM SIGCOMM Internet Measurement Conference, IMC*, pages 209–222, 2009.

[141] Werner L. Gundersheimer. Patronage in the Renaissance: An Exploratory Approach. In Guy Fitch Lytle and Stephen Orgel, editors, *Patronage in the Renaissance*, pages 3–23. Princeton University Press, Princeton, 1981.

[142] S. E. Hackney and Zoe Faye Pickard. Creating digital collections: Museum content and the public. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 10766 LNCS:626–631, 2018.

[143] Hermann Haken. *Comparisons Between Human Perception and Machine "Perception"*, volume 50, pages 133–148. Springer, Berlin, Heidelberg, 1991.

[144] Lars Hallnäs and Johan Redström. Slow Technology-Designing for Reflection. *Personal and Ubiquitous Computing*, 5(3):201–212, 2001.

[145] Kim Halskov and Nicolai Brodersen Hansen. The diversity of participatory design research practice at PDC 2002-2012. *International Journal of Human Computer Studies*, 74:81–92, 2015.

[146] Richard H.R. Harper. The role of hci in the age of ai. *International Journal of Human-Computer Interaction*, 35:1331–1344, 9 2019.

[147] Chris Harris and Mike Stephens. A COMBINED CORNER AND EDGE DETECTOR. *In Proc. 4th Alvey Vision Conference*, pages 147–151, 1998.

[148] Rex Hartson and Partha S. Pyla. *The UX Book: Process and Guidelines for Ensuring a Quality User Experience*. Elsevier, 1 2012.

[149] James J. Heckman. Sample selection bias as a specification error. *Econometrica*, 47(1):153–161, 1979.

[150] Laura M Herman. Remixing, Seeing, and Curating: Algorithms' Influence on Human Creativity. *Creativity & Cognition*, 2021.

[151] Laura M Herman. For who page? tiktok creators' algorithmic dependencies. In *International Association of Societies of Design Research Congress 2023*, 2023.

[152] Laura Mariah Herman and Angel Hsing-Chi Hwang. In the eye of the beholder: A viewer-defined conception of online visual creativity. *New Media Society*, pages 1–27, 2022.

[153] L.M. Herman and C. Moruzzi. The algorithmic pedestal: a practice-based study of algorithmic artistic curation. *Leonardo*, 2024.

[154] David Hesmondhalgh. Streaming's Effects on Music Culture: Old Anxieties and New Simplifications. *Cultural Sociology*, 16(1):3–24, 3 2022.

[155] L Hiller and L Isaacson. Musical composition with a high-speed digital computer. *Machine Models of Music*, pages 9–21, 1993.

[156] Florian Hoenig. Defining Computational Aesthetics. In L Neumann, M Sbert, B Gooch, and W Purgathofer, editors, *Computational Aesthetics in Graphics, Visualization and Imaging*. 2005.

[157] Bernie Hogan. The presentation of self in the age of social media: Distinguishing performances and exhibitions online. *Bulletin of Science, Technology Society*, 30:377–386, 11 2010.

[158] Bernie Hogan and Anabel Quan-Haase. Persistence and change in social media. *Bulletin of Science, Technology Society*, 30:309–315, 10 2010.

[159] Karen Holtzblatt. Contextual Design. In *The Human-Computer Interaction Handbook*, pages 942–963. CRC Press, 2nd edition, 2007.

[160] Karen Holtzblatt and Hugh Beyer. MAKING CUSTOMER-CENTERED DESIGN WORK FOR TEAMS V. *Communications of the ACM*, 36(10):93–103, 1993.

[161] Joo-Wha Hong. Bias in perception of art produced by artificial intelligence. pages 290–303. Springer International Publishing, 2018.

[162] Ahmed Hosny, Jili Huang, and Yingyi Wang. The Green Canvas.

[163] Emma June Huebner. Tiktok and museum education: A visual content analysis. *International Journal of Education Through Art*, 18(2):209–225, 2022.

[164] Hilary Hutchinson, Wendy Mackay, Bosse Westerlund, Benjamin B Bederson, Allison Druin, Catherine Plaisant, Michel Beaudouin-Lafon, Stéphane Conversy, Helen Evans, Heiko Hansen, Nicolas Roussel, Björn Eiderbäck, Sinna Lindquist, and Yngve Sundblad. Technology probes: Inspiring design for and with families. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*.

[165] Michael Hutter and Richard Shusterman. Chapter 6 Value and the Valuation of Art in Economic and Aesthetic Theory. *Handbook of the Economics of Art and Culture*, 1:169–208, 1 2006.

[166] Daniel R. Ilgen and Bruce W. Hamstra. Performance satisfaction as a function of the difference between expected and reported performance at five levels of reported performance. *Organizational Behavior and Human Performance*, 7:359–370, 6 1972.

[167] Albert Q. Jiang, Alexandre Sablayrolles, Arthur Mensch, Chris Bamford, Devendra Singh Chaplot, Diego de las Casas, Florian Bressand, Gianna Lengyel, Guillaume Lample, Lucile Saulnier, Lélio Renard Lavaud, Marie-Anne Lachaux, Pierre Stock, Teven Le Scao, Thibaut Lavril, Thomas Wang, Timothée Lacroix, and William El Sayed. Mistral 7b, 2023.

[168] Jialun Aaron Jiang, Kandrea Wade, Casey Fiesler, and Jed R. Brubaker. Supporting serendipity: Opportunities and challenges for human-ai collaboration in qualitative analysis. *Proceedings of the ACM on Human-Computer Interaction*, 5, 4 2021.

[169] Theodore Johnson and Tamraparni Dasu. Data Quality and Data Cleaning: An Overview. *Proceedings of the 2003 ACM SIGMOD International Conference on Management of Data*, 6 2003.

[170] Dhiraj Joshi, Ritendra Datta, Elena Fedorovskaya, Quang Tuan Luong, James Z. Wang, Jia Li, and Jiebo Luo. Aesthetics and emotions in images. *IEEE Signal Processing Magazine*, 28(5):94–115, 2011.

[171] Marius Kaminskas and Derek Bridge. Diversity, serendipity, novelty, and coverage: A survey and empirical analysis of beyond-Accuracy objectives in recommender systems, 12 2016.

[172] Sergey Karayev, Matthew Trentacoste, Helen Han, Aseem Agarwala, Trevor Darrell, Aaron Hertzmann, and Holger Winnemoeller. Recognizing Image Style. Technical report.

[173] Nadia Karizat, Dan Delmonaco, Motahhare Eslami, and Nazanin Andalibi. Algorithmic folk theories and identity: How tiktok users co-produce knowledge of identity and engage in algorithmic resistance. *Proc. ACM Hum.-Comput. Interact.*, 5(CSCW2), oct 2021.

[174] D. Bondy Valdovinos Kaye. Please Duet This: Collaborative Music Making in Lockdown on TikTok. *Networking Knowledge*, 15(1), 2 2022.

[175] Moaiad Ahmad Khder. Web Scraping or Web Crawling: State of Art, Techniques, Approaches and Application. *Int. J. Advance Soft Compu. Appl*, 13(3), 2021.

[176] Jenny Kidd, Irida Ntalla, and William B Lyons. Multi-touch interfaces in museum spaces: reporting preliminary findings on the nature of interaction.

[177] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In F. Pereira, C.J. Burges, L. Bottou, and K.Q. Weinberger, editors, *Advances in Neural Information Processing Systems*, volume 25. Curran Associates, Inc., 2012.

[178] Sari Kujala. User involvement: A review of the benefits and challenges. *Behaviour and Information Technology*, 22(1):1–16, 1 2003.

[179] Hyosun Kwon, Jo Briggs, Ben Freeth, Karim Ladha, James Thomas, Mark Blythe, and Patrick Olivier. Wet Pixels: From Digital to Analogue Originals. *CHI'13*, pages 583–588, 2013.

[180] Nils Köbis and Luca D. Mossink. Artificial intelligence versus maya angelou: Experimental evidence that people cannot differentiate ai-generated from human-written poetry. *Computers in Human Behavior*, 114, 1 2021.

[181] Yi Lai, Atreyi Kankanhalli, and Desmond C Ong. Human-ai collaboration in healthcare: A review and research agenda.

[182] Peter Lamarque, Stein Haugom Olsen, and John Wiley. Aesthetics and the Philosophy of Art : The Analytic Tradition, an Anthology, edited by. Technical report, 2019.

[183] Richard N. Landers, Robert C. Brusso, Katelyn J. Cavanaugh, and Andrew B. Collmus. A primer on theory-driven web scraping: Automatic extraction of big data from the internet for use in psychological research. *Psychological Methods*, 21(4):475–492, 12 2016.

[184] Pat. Langley, Herbert A. Simon, Gary Bradshaw, and Jan M. Zytkow. Scientific discovery : computational explorations of the creative processes. page 357, 1987.

[185] Maurizio Lazzarato. Immaterial Labor. In Paolo Virno and Michael Hardy, editors, *Radical Thought in Italy: A Potential Politics*. 1996.

[186] Michael A. Lepori and Chaz Firestone. Can you hear me now? sensitive comparisons of human and machine perception. *Cognitive Science*, 46, 10 2022.

[187] Feng Li, Mark Blyth, Jo Briggs, Jonathan Hook, and Patrick Olivier. From 'Digital Originals' to 'Unlimited Editions': Exploring New Business Models for Digital Art through Case Studies, Research Prototypes and Design Fictions. 2013.

[188] Franklin Mingzhe Li, Lotus Zhang, Maryam Bandukda, Abigale Stangl, Kristen Shinohara, Leah Findlater, and Patrick Carrington. Understanding Visual Arts Experiences of Blind People. *2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*, 2023.

[189] J. C.R. Licklider. Man-computer symbiosis. *IRE Transactions on Human Factors in Electronics*, HFE-1:4–11, 1960.

[190] Siân Lindley. Making time. In *CSCW 2015 - Proceedings of the 2015 ACM International Conference on Computer-Supported Cooperative Work and Social Computing*, pages 1442–1452. Association for Computing Machinery, Inc, 2 2015.

[191] Yaqi Liu, Qingxiao Guan, Xianfeng Zhao, and Yun Cao. Image forgery localization based on multi-scale convolutional neural networks. In *Proceedings of the 6th ACM Workshop on Information Hiding and Multimedia Security*, IHMMSec '18, page 85–90, New York, NY, USA, 2018. Association for Computing Machinery.

[192] Omar Lizardo. Why "cultural matters" matter: Culture talk as the mobilization of cultural capital in interaction. *Poetics*, 58:1–17, 10 2016.

[193] Tania Loke, Yayoi Teramoto, Chico Q. Camargo, and Kathryn Eccles. Heritage site-seeing through the visitor's lens on Instagram. *Journal of Cultural Analytics*, 7(3), oct 24 2022.

[194] Chiara Longoni, Andrea Bonezzi, and Carey K. Morewedge. Resistance to medical artificial intelligence. *Journal of Consumer Research*, 46:629–650, 12 2019.

[195] Zhicong Lu, Michelle Annett, Mingming Fan, and Daniel Wigdor. "I feel it is my responsibility to stream" Streaming and Engaging with Intangible Cultural Heritage through Livestreaming. *Conference on Human Factors in Computing Systems - Proceedings*, page 14, 5 2019.

[196] LUMA. Creative Matrix .

[197] LUMA. What's on Your Radar?

[198] Ingrid Lunden. Sotheby's acquires Thread Genius to build its image recognition and recommendation tech | TechCrunch, 2018.

[199] Can Ma, Xiaowei Dong, and Junbin Wang. A Trade-off Between the Artistic Aesthetic Value and Market Value of Paintings With Naïve and Childlike Interest Complex. *Empirical Studies of the Arts*, 41(2):352–371, 7 2023.

[200] Renkai Ma, Xinning Gui, and Yubo Kou. Multi-Platform Content Creation: The Configuration of Creator Ecology through Platform Prioritization, Content Synchronization, and Audience Management. *CHI '23: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, pages 1–19, 2023.

[201] Renkai Ma and Yubo Kou. "How advertiser-friendly is my video?": YouTuber's Socioeconomic Interactions with Algorithmic Content Moderation. *PACM on Human Computer Interaction*, 5(CSCW2), 2021.

[202] Renkai Ma and Yubo Kou. "i am not a youtuber who can make whatever video i want. i have to keep appeasing algorithms": Bureaucracy of creator moderation on youtube. *Companion Computer Supported Co-operative Work and Social Computing (CSCW'22 Companion)*, 2022.

[203] Renkai Ma and Yubo Kou. "i'm not sure what difference is between their content and mine, other than the person itself": A study of fairness perception of content moderation on youtube. *PACM on Human Computer Interaction*, 6, 2022.

[204] Lachlan MacDowall and Kylie Budge. *Art After Instagram: Art Spaces, Audiences, Aesthetics.* Routledge, New York, 2022.

[205] Taro Makino, Stanisław Jastrzębski, Witold Oleszkiewicz, Celin Chacko, Robin Ehrenpreis, Naziya Samreen, Chloe Chhor, Eric Kim, Jiyon Lee, Kristine Pysarenko, Beatriu Reig, Hildegard Toth, Divya Awal, Linda Du, Alice Kim, James Park, Daniel K. Sodickson, Laura Heacock, Linda Moy, Kyunghyun Cho, and Krzysztof J. Geras. Differences between human and machine perception in medical diagnosis. *Scientific Reports*, 12(1), 12 2022.

[206] Taro Makino, Stanisław Jastrzębski, Witold Oleszkiewicz, Celin Chacko, Robin Ehrenpreis, Naziya Samreen, Chloe Chhor, Eric Kim, Jiyon Lee, Kristine Pysarenko, Beatriu Reig, Hildegard Toth, Divya Awal, Linda Du, Alice Kim, James Park, Daniel K. Sodickson, Laura Heacock, Linda Moy, Kyunghyun Cho, and Krzysztof J. Geras. Differences between human and machine perception in medical diagnosis. *Scientific Reports*, 12, 12 2022.

[207] Lev Manovich. Data visualization as new abstraction and anti-sublime. *Manovich.net*, 2002.

[208] Lev Manovich. *Instagram and Contemporary Image.* 2017.

[209] Lev Manovich. Computer vision, human senses, and language of art. *AI and Society*, 36(4):1145–1152, 12 2021.

[210] Lev Manovich. *Artificial Aesthetics: A Critical Guide to AI, Media and Design.* 2023.

[211] Lev Manovich. *Artificial Aesthetics: A Critical Guide to AI, Media and Design.* 2023.

[212] Hui Mao, Ming Cheung, and James She. DeepArt: Learning Joint Representations of Visual Arts. 2017.

[213] Paul F. Marty. My lost museum: User expectations and motivations for creating personal digital collections on museum websites. *Library Information Science Research*, 33:211–219, 7 2011.

[214] Steven McClung and Kristine Johnson. Examining the Motives of Podcast Users. *Journal of Radio and Audio Media*, 17(1):82–95, 1 2010.

[215] Michael McCurdy, Christopher Connors, Guy Pyrzak, Bob Kanefsky, and Alonso Vera. Breaking the fidelity barrier: an examination of our current characterization of prototypes and an example of a mixed-fidelity success. *CHI '06: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1233–1242, 4 2006.

[216] Marshall McLuhan. *Understanding Media: The Extension of Man*. MIT Press, 1964.

[217] Thomas Mensink and Jan Van Gemert. The Rijksmuseum Challenge: Museum-Centered Visual Recognition. 2014.

[218] Robert K Merton. Matthew Effect in Science. *Science*, 159(3810):56–63, 1 1968.

[219] Meta. Llama3.1. `https://llama.meta.com/`, 2024.

[220] Cheryl Metoyer-Duran. Information gatekeepers. *Annual Review of Information Science and Technology (ARIST)*, 28:111–50, 1993.

[221] Eric T Meyer and Kathryn Eccles. The impacts of digital collections, 2016.

[222] Leonard B Meyer, Eugene Narmour, and Ruth A Solie. *Explorations in music, the arts, and ideas : essays in honor of Leonard B. Meyer*. Festschrift series ; no. 7. Pendragon Press, Stuyvesant, N.Y, 1988.

[223] Benjamin Meyer-Krahmer. A History of Curating—Past and Present. *Critique d'art*, (45), 11 2015.

[224] MidJourney. Midjourney.

[225] Marvin Minsky. Music, Mind, and Meaning. *Computer Music Journal*, 5(3):28–44, 1981.

[226] Michał Mochocki. Heritage sites and video games: Questions of authenticity and immersion. *Games and Culture*, 16(8):951–977, 2021.

[227] Vera Molnar. Toward aesthetic guidelines for paintings with the aid of a computer. *Leonardo*, 8:185, 22 1975.

[228] Jeremy Wade Morris. Curation by code: Infomediaries and the data mining of taste. *European Journal of Cultural Studies*, 18(4-5):446–463, 6 2015.

[229] Caterina Moruzzi. Should human artists fear ai? : A report on the perception of creative ai. In Mario Verdicchio, Miguel Carvalhais, Luisa Ribas, and Andre Rangal, editors, *xCoAx 2020: the Eighth Conference on Computation, Communication, Aesthetics X*, 2020.

[230] Caterina Moruzzi. Measuring creativity: an account of natural and artificial creativity. *European Journal for Philosophy of Science*, 11(1):1, 2021.

[231] Caterina Moruzzi. Creative agents: Rethinking agency and creativity in human and artificial systems. *Journal of Aesthetics and Phenomenology*, 9:245–268, 7 2022.

[232] Caterina Moruzzi and Oreste Leone Campagner. *AI and Art: Non-Human, Still Human, or All Too Human?*, pages 89–99. Taylor and Francis, 1 2024.

[233] Eduardo Mosqueira-Rey, Elena Hernández-Pereira, David Alonso-Ríos, José Bobes-Bascarán, and Ángel Fernández-Leal. Human-in-the-loop machine learning: a state of the art. *Artificial Intelligence Review*, 56:3005–3054, 4 2023.

[234] Sander Münster. Digital heritage as a scholarly field—topics, researchers, and perspectives from a bibliometric point of view. *Journal on Computing and Cultural Heritage (JOCCH)*, 12, 7 2019.

[235] Frieder Nake. Computer art: Creativity and computability. *Creativity and Cognition 2007, CC2007 - Seeding Creativity: Tools, Media, and Environments*, pages 305–306, 2007.

[236] Bence Nanay. Going global: A cautiously optimistic manifesto. *Contemporary Aesthetics*, 10, 2022.

[237] Arvind Narayanan. Understanding Social Media Recommendation Algorithms. Technical report, Knight First Amendment Institute, 2023.

[238] NCCAS. National core arts standards: A conceptual framework for arts learning, 2016.

[239] Matti Nelimarkka, Jean Philippe Rancy, Jennifer Grygiel, and Bryan Semaan. (Re)Design to Mitigate Political Polarization Reflecting Habermas' ideal communication space in the United States of America and Finland. *Proc. ACM Hum.-Comput. Interact.*, 3(CSCW), 2019.

[240] David B. Nieborg and Thomas Poell. The platformization of cultural production: Theorizing the contingent cultural commodity. *New Media and Society*, 20(11):4275–4292, 11 2018.

[241] Jakob Nielsen and Jonathan Levy. Measuring usability: Preference vs. performance. *Communications of the ACM*, 37:66–75, 1 1994.

[242] Donald A. Norman. *The Design of Everyday Things*. Basic Books, Inc., USA, 2002.

[243] William Odom, Richard Banks, Abigail Durrant, David Kirk, and James Pierce. Slow Technology: Critical Reflection and Future Directions. *Proceedings of the Designing Interactive Systems Conference.*, pages 816–827, 2012.

[244] William Odom, Erik Stolterman, and Amy Yo Sue Chen. Extending a Theory of Slow Technology for Design through Artifact Analysis. *Human-Computer Interaction*, 37(2):150–179, 2022.

[245] Government of British Columbia. Arts education k-9-curricular competencies grade exploring and creating reasoning and reflecting communicating and documenting connecting and expanding k, 7 2019.

[246] Government of British Columbia. The provincial proficiency scale, 2019.

[247] John O'Hayre. Gobbledygook has gotta go, 1966.

[248] A Ohlheiser. Tiktok made him famous. now he's imagining a world without it. *MIT Technology Review*, 2020.

[249] Mimi Onuoha. GitHub - MimiOnuoha/missing-datasets: An overview and exploration of the concept of missing datasets.

[250] Keng-Boon Ooi, Garry Wei-Han Tan, Mostafa Al-Emran, Mohammed A Al-Sharafi, Alexandru Capatina, Amrita Chakraborty, Yogesh K Dwivedi, Tzu-Ling Huang, Arpan Kumar Kar, Voon-Hsien Lee, Xiu-Ming Loh, Adrian Micu, Patrick Mikalef, Emmanuel Mogaji, Neeraj Pandey, Ramakrishnan Raman, Nripendra P Rana, Prianka Sarker, Anshuman Sharma, Ching-I Teng, Samuel Fosso Wamba, and Lai-Wan Wong. The Potential of Generative Artificial Intelligence Across Disciplines: Perspectives and Future Directions. *Journal of Computer Information Systems*, 2023.

[251] OpenAI, Josh Achiam, Steven Adler, Sandhini Agarwal, Lama Ahmad, Ilge Akkaya, Florencia Leoni Aleman, Diogo Almeida, Janko Altenschmidt, Sam Altman, Shyamal Anadkat, Red Avila, Igor Babuschkin, Suchir Balaji, Valerie Balcom, Paul Baltescu, Haiming Bao, Mohammad Bavarian, Jeff Belgum, Irwan Bello, Jake Berdine, Gabriel Bernadett-Shapiro, Christopher Berner, Lenny Bogdonoff, Oleg Boiko, Madelaine Boyd, Anna-Luisa Brakman, Greg Brockman, Tim Brooks, Miles Brundage, Kevin Button, Trevor Cai, Rosie Campbell, Andrew Cann, Brittany Carey, Chelsea Carlson, Rory Carmichael, Brooke Chan, Che Chang, Fotis Chantzis, Derek Chen, Sully Chen, Ruby Chen, Jason Chen, Mark Chen, Ben Chess, Chester Cho, Casey Chu, Hyung Won Chung, Dave Cummings, Jeremiah Currier, Yunxing Dai, Cory Decareaux, Thomas Degry, Noah Deutsch, Damien Deville, Arka Dhar, David Dohan, Steve Dowling, Sheila Dunning, Adrien Ecoffet, Atty Eleti, Tyna Eloundou, David Farhi, Liam Fedus, Niko Felix, Simón Posada Fishman, Juston Forte, Isabella Fulford, Leo Gao, Elie Georges, Christian Gibson, Vik Goel, Tarun Gogineni, Gabriel Goh, Rapha Gontijo-Lopes, Jonathan Gordon, Morgan Grafstein, Scott Gray, Ryan Greene, Joshua Gross, Shixiang Shane Gu, Yufei Guo, Chris Hallacy, Jesse Han, Jeff Harris, Yuchen He, Mike Heaton, Johannes Heidecke, Chris Hesse, Alan Hickey, Wade Hickey, Peter Hoeschele, Brandon Houghton, Kenny Hsu, Shengli Hu, Xin Hu, Joost Huizinga, Shantanu Jain, Shawn Jain, Joanne Jang, Angela Jiang, Roger Jiang, Haozhun Jin, Denny Jin, Shino Jomoto, Billie Jonn, Heewoo Jun, Tomer Kaftan, Łukasz Kaiser, Ali Kamali, Ingmar Kanitscheider, Nitish Shirish Keskar, Tabarak Khan, Logan Kilpatrick, Jong Wook Kim, Christina Kim, Yongjik Kim, Jan Hendrik Kirchner, Jamie Kiros, Matt Knight, Daniel Kokotajlo, Łukasz Kondraciuk, Andrew Kondrich, Aris Konstantinidis, Kyle Kosic, Gretchen Krueger, Vishal Kuo, Michael Lampe, Ikai Lan, Teddy Lee, Jan Leike, Jade Leung, Daniel Levy, Chak Ming Li, Rachel Lim, Molly Lin, Stephanie Lin, Mateusz Litwin, Theresa Lopez, Ryan Lowe, Patricia Lue, Anna Makanju, Kim Malfacini, Sam Manning, Todor Markov, Yaniv Markovski, Bianca Martin, Katie Mayer, Andrew Mayne, Bob McGrew, Scott Mayer McKinney, Christine McLeavey, Paul McMillan, Jake McNeil, David Medina, Aalok Mehta, Jacob Menick, Luke Metz, Andrey Mishchenko, Pamela Mishkin, Vinnie Monaco, Evan Morikawa, Daniel Mossing, Tong Mu, Mira Murati, Oleg Murk, David Mély, Ashvin Nair, Reiichiro Nakano, Rajeev Nayak, Arvind Neelakantan, Richard Ngo, Hyeonwoo Noh, Long Ouyang, Cullen O'Keefe, Jakub Pachocki, Alex Paino, Joe Palermo, Ashley Pantuliano, Giambattista Parascandolo, Joel Parish, Emy Parparita, Alex Passos, Mikhail Pavlov, Andrew Peng, Adam Perelman, Filipe de Avila Belbute Peres, Michael Petrov, Henrique Ponde de Oliveira Pinto, Michael, Pokorny, Michelle Pokrass, Vitchyr H. Pong,

Tolly Powell, Alethea Power, Boris Power, Elizabeth Proehl, Raul Puri, Alec Radford, Jack Rae, Aditya Ramesh, Cameron Raymond, Francis Real, Kendra Rimbach, Carl Ross, Bob Rotsted, Henri Roussez, Nick Ryder, Mario Saltarelli, Ted Sanders, Shibani Santurkar, Girish Sastry, Heather Schmidt, David Schnurr, John Schulman, Daniel Selsam, Kyla Sheppard, Toki Sherbakov, Jessica Shieh, Sarah Shoker, Pranav Shyam, Szymon Sidor, Eric Sigler, Maddie Simens, Jordan Sitkin, Katarina Slama, Ian Sohl, Benjamin Sokolowsky, Yang Song, Natalie Staudacher, Felipe Petroski Such, Natalie Summers, Ilya Sutskever, Jie Tang, Nikolas Tezak, Madeleine B. Thompson, Phil Tillet, Amin Tootoonchian, Elizabeth Tseng, Preston Tuggle, Nick Turley, Jerry Tworek, Juan Felipe Cerón Uribe, Andrea Vallone, Arun Vijayvergiya, Chelsea Voss, Carroll Wainwright, Justin Jay Wang, Alvin Wang, Ben Wang, Jonathan Ward, Jason Wei, CJ Weinmann, Akila Welihinda, Peter Welinder, Jiayi Weng, Lilian Weng, Matt Wiethoff, Dave Willner, Clemens Winter, Samuel Wolrich, Hannah Wong, Lauren Workman, Sherwin Wu, Jeff Wu, Michael Wu, Kai Xiao, Tao Xu, Sarah Yoo, Kevin Yu, Qiming Yuan, Wojciech Zaremba, Rowan Zellers, Chong Zhang, Marvin Zhang, Shengjia Zhao, Tianhao Zheng, Juntang Zhuang, William Zhuk, and Barret Zoph. Gpt-4 technical report, 2024.

[252] Marcus O'Dair and Andrew Fry. Beyond the black box in music streaming: the impact of recommendation systems upon artists. *Popular Communication*, 18(1):65–77, 1 2020.

[253] Soya Park and Chinmay Kulkarni. Thinking assistants: Llm-based conversational assistants that help users think by asking rather than answering. 12 2023.

[254] Yoon-Joo Park and Alexander Tuzhilin. The Long Tail of Recommender Systems and How to Leverage It. *RecSys*, pages 11–18, 2008.

[255] Ross Parry. *Museums in a Digital Age.* Taylor and Francis Group, 2009.

[256] Soren Mork Petersen. Loser Generated Content: From Participation to Exploitation, 2008.

[257] Gloria Phillips-Wren. *Intelligent Decision Support Systems*, pages 25–44. John Wiley and Sons, 2 2013.

[258] Thomas Poell, David B Nieborg, and Brooke Erin Duffy. *Platforms and cultural production.* John Wiley & Sons, 2021.

[259] Soren Bro Pold and Christian Ulrik Andersen. Controlled Consumption Culture When Digital Culture Becomes Software Business. In Paul D. Miller and Svitlana Matviyenko, editors, *The Imaginary App*, chapter 2, pages 17–33. MIT Press, 2014.

[260] John P. Portelli. The challenge of teaching for critical thinking. *McGill Journal of Education / Revue des sciences de l'éducation de McGill*, 29, 4 1994.

[261] Larissa Pschetz. Isn't it time to change the way we think about time?, 9 2015.

[262] Larissa Pschetz and Michelle Bastian. Temporal Design: Rethinking time in design. *Design Studies*, 56:169–184, 5 2018.

[263] Irene Rae. The effects of perceived ai use on content perceptions. *Proceedings of the CHI Conference on Human Factors in Computing Systems*, pages 1–14, 5 2024.

[264] Irene Rae. The Effects of Perceived AI Use On Content Perceptions. *Proceedings of the CHI Conference on Human Factors in Computing Systems*, pages 1–14, 5 2024.

[265] Martin Ragot, Nicolas Martin, and Salomé Cojean. Ai-generated vs. human artworks. a perception bias towards artificial intelligence? *Conference on Human Factors in Computing Systems - Proceedings*, 4 2020.

[266] Hafizur Rahaman. Digital heritage interpretation: a conceptual framework. *Digital Creativity*, 29:208–234, 7 2018.

[267] Hafizur Rahaman and Beng-Kiang Tan. Interpreting digital heritage:a conceptual model with end-users' perspective. *International Journal of Architectural Computing*, 9:99–113, 3 2011.

[268] Erhard Rahm and Hong Hai Do. Data Cleaning: Problems and Current Approaches. *IEEE Data Eng. Bull.*, 23(4), 2000.

[269] Aditya Ramesh, Mikhail Pavlov, Gabriel Goh, Scott Gray, Chelsea Voss, Alec Radford, Mark Chen, and Ilya Sutskever. Zero-Shot Text-to-Image Generation.

[270] Michal Raz-Russo. Subscribe: Artists and Alternative Magazines, 1970-1995 | The Art Institute of Chicago, 12 2021.

[271] Radim Rehurek and Petr Sojka. Software Framework for Topic Modelling with Large Corpora. In *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*, pages 45–50. ELRA, 5 2010.

[272] James Reveley. The Exploitative Web: Misuses of Marx in Critical Social Media Studies. *Science & Society*, 77(4):512–535, 2013.

[273] P. J. Rey. Alienation, Exploitation, and Social Media, 4 2012.

[274] Anna Marie Rezk, Auste Simkute, Ewa Luger, John Vines, Chris Elsden, Michael Evans, and Rhianne Jones. Agency Aspirations: Understanding Users' Preferences And Perceptions Of Their Role In Personalised News Curation. *Proceedings of the CHI Conference on Human Factors in Computing Systems*, pages 1–16, 5 2024.

[275] Brian Ribeiro. HUME'S STANDARD OF TASTE AND THE DE GUSTIBUS SCEPTIC. *British Journal of Aesthetics*, 47(1), 2007.

[276] Leonard Richardson. Beautiful Soup: We called him Tortoise because he taught us., 2016.

[277] George Ritzer and Nathan Jurgenson. Production, Consumption, Prosumption The nature of capitalism in the age of the digital 'prosumer'. *Journal of Consumer Culture*, 10(1):13–36, 2010.

[278] Yvonne Rogers. *A INTERACTION DESIGN: beyond human-computer interaction.* 2002.

[279] Kevin Roose. Ai-generated art won a prize. artists aren't happy. *The New York Times*, 9 2022.

[280] Kevin Roose. AI-Generated Art Won a Prize. Artists Aren't Happy., 9 2022.

[281] Teemu Ropponen, Auli Harju, Pirjo Näkki, Kari Hintikka, and Asta Bäck. Co-designing a social media service for civic participation-critical issues and challenges. *MindTrek '10: Proceedings of the 14th International Academic MindTrek Conference: Envisioning Future Media Environments*, 2010.

[282] Mary Rozell. The art collector's handbook: A guide to collection management and care. *The Art Collector's Handbook: A Guide to Collection Management and Care*, pages 1–223, 3 2014.

[283] Matthew J. Salganik, Peter Sheridan Dodds, and Duncan J. Watts. Experimental study of inequality and unpredictability in an artificial cultural market. *Science*, 311(5762):854–856, 2 2006.

[284] Christian Sandvig, Kevin Hamilton, Karrie Karahalios, Cedric Langbort, and * Corresponding Author. Auditing Algorithms: Research Methods for Detecting Discrimination on Internet Platforms. In *Data and Discrimination: Converting Critical Concerns into Productive Inquiry*, 2014.

[285] Serhad Sarica and Jianxi Luo. Stopwords in technical language processing. *PLoS ONE*, 16(8), 8 2021.

[286] Advait Sarkar, Xiaotong Xu, Neil Toronto, Ian Drosos, and Christian Poelitz. When copilot becomes autopilot: Generative ai's critical risk to knowledge work and a critical solution.

[287] Ludovica Schaerf, Pepe Ballesteros Zapata, Valentine Bernasconi, Iacopo Neri, Dario Negueruela, and Del Castillo. Ai art curation: Re-imagining the city of helsinki in occasion of its biennial. *arxiv*, 1, 6 2023.

[288] Andreas Schellewald. Communicative Forms on TikTok: Perspectives From Digital Ethnography. Technical report, 2021.

[289] Lea Schick and Katrine Damkjaer. Can You Be Friends with an Art Museum? Rethinking the Art Museum through Facebook. In *Transforming Culture in the Digital Age*, pages 36–42, 2010.

[290] Skipper Seabold and Josef Perktold. statsmodels: Econometric and statistical modeling with python. In *9th Python in Science Conference*, 2010.

[291] Nick Seaver. Knowing Algorithms. In Janet Vertesi and David Ribes, editors, *digitalSTS: A Field Guide for Science & Technology Studies*, pages 412–422. 2019.

[292] Isabella Seeber, Eva Bittner, Robert O. Briggs, Triparna de Vreede, Gert Jan de Vreede, Aaron Elkins, Ronald Maier, Alexander B. Merz, Sarah Oeste-Reiß, Nils Randrup, Gerhard Schwabe, and Matthias Söllner. Machines as teammates: A research agenda on ai in team collaboration. *Information and Management*, 57, 3 2020.

[293] Anamaria Dutceac Segesten, Michael Bossetta, Nils Holmberg, and Diederick Niehorster. The cueing power of comments on social media: how disagreement in facebook comments affects user engagement with news. *Information, Communication Society*, 25:1115–1134, 2020.

[294] Selenium. Selenium Web-Browser Automation, 2013.

[295] Howard J. Seltman. *Experimental Design and Analysis*. Carnegie Mellon University Statistics, 2018.

[296] Lior Shamir, Jenny Nissel, and Ellen Winner. Distinguishing between abstract art by artists vs. children and animals: Comparison between human and machine perception. *ACM Transactions on Applied Perception (TAP)*, 13(3):1–17, 2016.

[297] Vishal Sharma, Kirsten Bray, Neha Kumar, and Rebecca E Grinter. It Takes (at least) Two: The Work to Make Romance Work. *CHI '23: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, pages 1–17, 2023.

[298] Richard Siegesmund. Why do we teach art today? *Studies in Art Education*, 39:197–214, 1998.

[299] Carson Sievert and Kenneth E Shirley. Ldavis: A method for visualizing and interpreting topics. *Proceedings of the Workshop on Interactive Language Learning, Visualization, and Interfaces*, pages 63–70, 2014.

[300] H. A. Simon and R. K. Sumner. Patterns in music. *Machine Models of Music*, pages 83–110, 1968.

[301] Ellen Simpson and Bryan Semaan. Rethinking Creative Labor: A Sociotechnical Examination of Creativity & Creative Work on TikTok. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*, 2023.

[302] Supriya Singh, Meredith Blake, and Jonathan O'donnell. Digitizing Pacific Cultural Collections: The Australian Experience. *Article in International Journal of Cultural Property*, 2017.

[303] David Skidmore and Kyoko Murakami. *Dialogic Pedagogy: An Introduction*, pages 1–16. Multilingual Matters, 2016.

[304] Glenn W. Smith. An interview with frieder nake. *Arts 2019, Vol. 8, Page 69*, 8:69, 5 2019.

[305] Sotheby's. "The Value of Art".

[306] Sotheby's. Data Science, Machine Learning and Artificial Intelligence for Art | by Sotheby's | Sotheby's | Medium, 2018.

[307] James L. Spates. Counterculture and Dominant Culture Values: A Cross-National Analysis of the Underground Press and Dominant Culture Magazines. *American Sociological Review*, 41(5):868, 10 1976.

[308] Ramya Srinivasan, Emily Denton, Google Research, New York, Jordan Famularo, Negar Rostamzadeh, Fernando Diaz, and Beth Coleman. Artsheets for Art Datasets. *35th Conference on Neural Information Processing Systems (NeurIPS 2021) Track on Datasets and Benchmarks*, 2021.

[309] John Stack. What the Machine Saw, 2019.

[310] Statista Research Department. Art market worldwide statistics & facts | Statista, 8 2021.

[311] Statista Research Department. Internet and social media users in the world, 2022.

[312] Robert Stecker. Why Artistic Value is not Aesthetic Value. In *Intersections of Value: Art, Nature, and the Everyday*, pages 41–57. Oxford University Press, 3 2019.

[313] Marc Steen, Jenny De Boer, Lilliane Kuiper-Hoyng, Froukje Sleeswijk Visser, and H 5 M Miscellaneous. Co-design: Practices, Challenges and Lessons Learned Categories and Subject Descriptors General Terms. *MobileHCI*, 2008.

[314] Marc Steen, Lottie Kuijt-Evers, and Jente Klok. Early user involvement in research and design projects-A review of methods and practices. Technical report, 2007.

[315] Jeff Steward. harvardartmuseums/api-toolkit.

[316] JEff Steward. Harvard art museums api. `https://github.com/harvardartmuseums/api-docs`, 2015.

[317] Frances Stracey. Bio-art: the ethics behind the aesthetics. 2009.

[318] Hariharan Subramonyam, Steven M. Drucker, and Eytan Adar. Affinity lens data-assisted affinity diagramming with augmented reality. *Conference on Human Factors in Computing Systems - Proceedings*, 5 2019.

[319] Xiaoqing Sun, Jingyi Wang, Yan Zhou, Suhan Wang, Yixuan Li, and Xipei Ren. Co-coffee: A technology probe study to facilitate coffee breaks in open offices. *Conference on Human Factors in Computing Systems - Proceedings*, 5 2024.

[320] Luke Swartz. Why people hate the paperclip: Labels, appearance, behavior and social responses to user interface agents. 2003.

[321] Latanya Sweeney. Discrimination in Online Ad Delivery. Technical report, Harvard, Cambridge, 1 2013.

[322] Richard Szeliski. *Computer Vision: Algorithms and Applications*. Springer Nature, 2 edition, 2022.

[323] Jonathan Talmi. Living museum app, 10 2024.

[324] Haoheng Tang and Mrinalini Singha. A mystery for you: A fact-checking game enhanced by large language models (llms) and a tangible interface. *Conference on Human Factors in Computing Systems - Proceedings*, 5 2024.

[325] Joel Taylor and Laura Kate Gibson. Digitisation, digital interaction and social media: embedded barriers to democratic heritage. *International Journal of Heritage Studies*, 23:408–420, 5 2017.

[326] Harold Thwaites. Digital heritage: What happens when we digitize everything? *Springer Series on Cultural Computing*, pages 327–348, 2013.

[327] Alexandra To, Hillary Carey, Riya Shrivastava, Jessica Hammer, and Geoff Kaufman. Interactive fiction provotypes for coping with interpersonal racism. *Conference on Human Factors in Computing Systems - Proceedings*, 4 2022.

[328] Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, Aurelien Rodriguez, Armand Joulin, Edouard Grave, and Guillaume Lample. Llama: Open and efficient foundation language models, 2023.

[329] Georgios Trichopoulos, Markos Konstantakis, George Caridakis, Akrivi Katifori, and Myrto Koukouli. Crafting a museum guide using chatgpt4. *Big Data and Cognitive Computing 2023, Vol. 7, Page 148*, 7:148, 9 2023.

[330] Rebecca Uliasz. Seeing like an algorithm: operative images and emergent subjects. *AI and Society*, 36:1233–1241, 12 2021.

[331] Jose Van Dijck. *The Culture of Connectivity: A Critical History of Social Media.* 2013.

[332] Niels van Doorn. Platform labor: on the gendered and racialized exploitation of low-income service work in the 'on-demand' economy. *Information Communication and Society*, 20(6):898–914, 6 2017.

[333] Hans J. Van Miegroet, Kaylee P. Alexander, and Fiene Leunissen. Imperfect Data, Art Markets and Internet Research. *Arts*, 8(3):76, 6 2019.

[334] Iva Vasic, Hans Georg Fill, Ramona Quattrini, and Roberto Pierdicca. Llm-aided museum guide: Personalized tours based on user preferences. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 15029 LNCS:249–262, 2024.

[335] Silvia De Vecchio. What kind of museum visitor are you?, 4 2021.

[336] Elena Villaespesa and Oonagh Murphy. This is not an apple! Benefits and challenges of applying computer vision to museum collections. *Museum Management and Curatorship*, 36(4):362–383, 7 2021.

[337] Quentin Villermet, Jérémie Poiroux, Manuel Moussallam, Thomas Louail, and Camille Roth. Follow the guides: Disentangling human and algorithmic curation in online music consumption. *RecSys 2021 - 15th ACM Conference on Recommender Systems*, pages 380–389, 9 2021.

[338] Paul Viola and Michael Jones. Rapid Object Detection using a Boosted Cascade of Simple Features. 2001.

[339] Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272, 2020.

[340] visual-arts cork.com. Visual Art: Definition, History, Classification.

[341] Sergio Vitulano, Vito Di Gesú, Virginio Cantoni, Roberto Marmo, and Alessandra Setti. *Human and machine perception: Communication, interaction, and integration:*. World Scientific Publishing Co., 1 2005.

[342] Arantxa Vizcaíno-Verdú and Crystal Abidin. Music Challenge Memes on TikTok: Understanding In-Group Storytelling Videos CRYSTAL ABIDIN. *International Journal of Communication*, 16:883–908, 2022.

[343] Arantxa Vizcaíno-Verdú, Patricia De-Casas-moreno, and Simona Tirocchi. Online prosumer convergence: Listening, creating and sharing music on YouTube and TikTok. *Communication and Society*, 36(1):151–166, 1 2023.

[344] Emily A Vogels, Risa Gelles-Watnick, and Navid Massarat. Teens, Social Media and Technology 2022. Technical report, Pew Research Center, 8 2022.

[345] Thomas Serban von Davier, Hayoun Noh, Mitra Abbaspour, Doug Breault, Nicola Cosford, Caroline Dionne, Bryan Just, Karl Kusserow, Zoe Kwok, Francesca Leoni, Tara Metal, Andrew Shapland, Sydney Smith, Jeff Steward, Matthew Winterbottom, Max Van Kleek, and Nigel Shadbolt. Co-designing to separate art from content with input from art experts, 2024.

[346] Thomas Serban von Davier, Hayoun Noh, Mitra Abbaspour, Doug Breault, Nicola Cosford, Caroline Dionne, Bryan Just, Karl Kusserow, Zoe Kwok, Francesca Leoni, Tara Metal, Andrew Shapland, Sydney Smith, Jeff Steward, Matthew Winterbottom, Max Van Kleek, and Nigel Shadbolt. Co-Designing to Separate Art from Content with input from Art Experts, 9 2023.

[347] Thomas Şerban Von Davier. Designing for Appreciation: How Digital Spaces Can Support Art and Culture. *Conference on Human Factors in Computing Systems - Proceedings*, 4 2023.

[348] Michael Vössing, Niklas Kühl, Matteo Lind, and Gerhard Satzger. Designing transparency for effective human-ai collaboration. *Information Systems Frontiers*, 24:877–895, 6 2022.

[349] Jayne Wallace, Daniel G Jackson, and Cassim Ladha. Digital jewellery and family relationships. 2007.

[350] Dakuo Wang, Elizabeth Churchill, Pattie Maes, Xiangmin Fan, Ben Shneiderman, Yuanchun Shi, and Qianying Wang. From human-human collaboration to human-ai collaboration: Designing ai systems that can work together with people. In *Conference on Human Factors in Computing Systems - Proceedings*. Association for Computing Machinery, 4 2020.

[351] Rebecca D. Watkins, Abigail Sellen, and Siân E. Lindley. Digital collections and digital collecting practices. *Conference on Human Factors in Computing Systems - Proceedings*, 2015-April:3423–3432, 4 2015.

[352] Wes McKinney. Data Structures for Statistical Computing in Python. In Stéfan van der Walt and Jarrod Millman, editors, *Proceedings of the 9th Python in Science Conference*, pages 56 – 61, 2010.

[353] Cathleen Wharton, Janice Bradford, Fran Zke, and Robin Je@ies. Applying Cognitive Walkthroughs to More Complex User Interfaces: Experiences, Issues, and Recommendations. *CHI '92: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1992.

[354] Mitchell Whitelaw. Representing digital collections. *Performing Digital*, pages 77–96, 3 2016.

[355] Michael J Wilber, Chen Fang, Hailin Jin, Aaron Hertzmann, John Collomosse, and Serge Belongie. BAM! The Behance Artistic Media Dataset for Recognition Beyond Photography. 2017.

[356] Donghee Yvette Wohn, Peter Jough, Peter Eskander, John Scott Siri, Masaho Shimobayashi, and Pradnya Desai. Understanding digital patronage: Why do people subscribe to streamers on twitch? *CHI PLAY 2019 - Proceedings of the Annual Symposium on Computer-Human Interaction in Play*, pages 99–110, 10 2019.

[357] Tracee Vetting Wolf, Jennifer A. Rode, Jeremy Sussman, and Wendy A. Kellogg. Dispelling design as the 'black art' of CHI. *Conference on Human Factors in Computing Systems - Proceedings*, 1:521–530, 2006.

[358] Belle Wong. Top Social Media Statistics And Trends Of 2023 – Forbes Advisor. *Forbes Advisor*, 5 2023.

[359] Anbang Xu and Brian P Bailey. What Do You Think? A Case Study of Benefit, Expectation, and Interaction in a Large Online Critique Community. *CSCW'12 Companion: Companion Publication of the 2012 Conference on Computer Supported Cooperative Work and Social Computing*, 2012.

[360] Rouchen Yao. *How does contemporary artists consolidate their success and achievements through social media platform such as Instagram?: Case study of successful co-existing relationship developed between contemporary artists Ai Weiwei and KAWS with their Instagram account.* PhD thesis, Sotheby's Institute of Art, 2019.

[361] Andre Ye, Jared Moore, Rose Novick, Amy X Zhang, and Paul G Allen. Language models as critical thinking tools: A case study of philosophers. 4 2024.

[362] Nikolaos-Antonios Ypsilantis, Noa Garcia, Guangxing Han, Sarah Ibrahimi, Nanne Van Noord, and Giorgos Tolias. The Met Dataset: Instance-level Recognition for Artworks. *35th Conference on Neural Information Processing Systems (NeurIPS 2021) Track on Datasets and Benchmarks*, 2021.

[363] Fabio Massimo Zanzotto. Viewpoint: Human-in-the-loop artificial intelligence. *Journal of Artificial Intelligence Research*, 64:243–252, 2 2019.

[364] An Zeng, Chi Ho Yeung, Matus Medo, and Yi-Cheng Zhang. Modeling mutual feedback between users and recommender systems. Technical report, 8 2015.

[365] Jing Zeng and Crystal Abidin. '#OkBoomer, time to meet the Zoomers': studying the memefication of intergenerational politics on TikTok. *Information Communication and Society*, 24(16):2459–2481, 2021.

[366] Jiajing Zhang, Yongwei Miao, Junsong Zhang, and Jinhui Yu. Inkthetics: A Comprehensive Computational Model for Aesthetic Evaluation of Chinese Ink Paintings. *IEEE Access*, 8:225857–225871, 2020.

[367] Lei Zhang, Hui Zhang, and Kai Wang. Media literacy education and curriculum integration: A literature review. *International Journal of Contemporary Education*, 3:55, 3 2020.

[368] Lixuan Zhang, Iryna Pentina, and Yuhong Fan. Who do you choose? comparing perceptions of human vs robo-advisor in the context of financial services. *Journal of Services Marketing*, 35:634–646, 2021.

[369] Yu Zhang, Huanchen Wang, Hong Kong, Zhicong Lu, and Changyang He. Understanding Communication Strategies and Viewer Engagement with Science Knowledge Videos on Bilibili. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*, 2023.

[370] Dan Dan Zhao, An Zeng, Ming Sheng Shang, and Jian Gao. Long-term effects of recommendation on the evolution of online systems. *Chinese Physics Letters*, 30(11), 2013.

[371] Xuan Zhao, Cliff Lampe, and Nicole B Ellison. The Social Media Ecology: User Perceptions, Strategies and Challenges. *CHI '16: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 2016.

[372] John Zimmerman, Jodi Forlizzi, and Shelley Evenson. Research Through Design as a Method for Interaction Design Research in HCI. *ACM Proceedings in Computer Human Interaction on Design Theory*, pages 493–502, 2007.

[373] Diana Zulli and David James Zulli. Extending the Internet meme: Conceptualizing technological mimesis and imitation publics on the TikTok platform. *New Media and Society*, 24(8):1872–1890, 8 2022.

[374] Hui Zuo and Tongyue Wang. Analysis of Tik Tok User Behavior from the Perspective of Popular Culture. *Frontiers in Art Research*, 1:1–05, 2019.

[375] Joanna Zylinska. *The Perception Machine: Our Photographic Future between the Eye and AI.* The MIT Press, 11 2023.

[376] Ángel Alexander Cabrera, Adam Perer, and Jason I. Hong. Improving human-ai collaboration with descriptions of ai behavior. *Proceedings of the ACM on Human-Computer Interaction*, 7, 4 2023.