Content Authenticities: A Discussion on the Values of Provenance Data for Creatives and Their Audiences

Caterina Moruzzi
Institute for Design Informatics
University of Edinburgh
Edinburgh, United Kingdom
cmoruzzi@ed.ac.uk

Ella Tallyn
Institute for Design Informatics
University of Edinburgh
Edinburgh, United Kingdom
e.tallyn@ed.ac.uk

Frances Liddell
Institute for Design Informatics
University of Edinburgh
Edinburgh, United Kingdom
fliddell@ed.ac.uk

Billy Dixon
Institute for Design Informatics
University of Edinburgh
Edinburgh, United Kingdom
billy.dixon@ed.ac.uk

John Collomosse
CVSSP
University of Surrey
Guildford, Surrey, United Kingdom
j.collomosse@surrey.ac.uk

Chris Elsden
Institute for Design Informatics
University of Edinburgh
Edinburgh, United Kingdom
celsden@ed.ac.uk

Abstract

The proliferation of AI-generated digital content has intensified the user demand for accurate provenance information to ensure content authenticity. Technical advancements now provide tools to make the digital media content supply chain more transparent through the use of provenance data. This paper foregrounds the importance of understanding how the situated nature of user-content engagement influences perceptions and uses of this data. Insights from a workshop with experts in the creative media sector suggest that, as the adoption of provenance data becomes more common, users need richer and more nuanced information. We suggest that analyzing the increasing demand for content authenticity through the lens of multiple "authenticities", each reflecting different user needs and contexts, can help identify and address the needs for, and uses of, provenance data by creators and audiences alike.

CCS Concepts

• **Human-centered computing → HCI theory, concepts and models**.

Keywords

Provenance, Content Authenticity, Media, Creative Industry, Creativity Support, HCI, Design

ACM Reference Format:
Caterina Moruzzi, Ella Tallyn, Frances Liddell, Billy Dixon, John Collomosse, and Chris Elsden. 2025. Content Authenticities: A Discussion on the Values of Provenance Data for Creatives and Their Audiences. In *Creativity and Cognition (C&C '25), June 23–25, 2025, Virtual, United Kingdom.* ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3698061.3726918

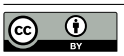
1 Introduction

Determining the originality, authenticity, and the history of how media or artworks have been handled has long been a critical expert practice. In the current digital landscape, where an estimated 34 million images are created each day with the use of Artificial Intelligence (AI) tools [55], the need for reliable provenance data, which tracks the origin and history of content, has intensified [22, 43, 48]. Content is often remixed, edited, and shared without clear attribution, thereby diluting the creator's control over their work and complicating efforts to verify its authenticity.

Based on the development of cryptographic techniques by which provenance data can be preserved, regardless of how media is subsequently reused or manipulated [15], various industry solutions are being proposed to securely track the creation, publication and sharing of creative work (Section 2.1). Applications and design interventions based on provenance data have been primarily studied to reduce the spread of disinformation and protect audiences from being misled [23, 24, 52] (Section 2.2). However, the design space, potential value, and implications of provenance data stretch further, especially when we consider media ecosystems beyond journalism and news. Although the need for "provenance facts" [15] is becoming commonplace, how these tools can and should be deployed across a creative media ecosystem — from original content creators, editors and aggregators, through to end-user audiences — remains under-researched.

In this paper, we aim to address these gaps by **exploring how the contextual and situated nature of audience engagement with content shapes the need for, and value of provenance data**. We do so by discussing findings from a workshop where we engaged 34 expert users spanning different roles in the creative media supply chain, from content production to consumption (Section 3).

Over the past years, there has been significant attention directed toward the authenticity of media content, particularly in the context of deepfakes and their proliferation on social media platforms [9, 34, 41]. However, within this research, the concept of "authenticity" is rarely defined or explored in its multiple nuances, overlooking its complexity and the diverse ways in which authenticity can be related to representations of provenance and interpreted across different contexts, cultures, and media forms. Insights from the workshop we conducted suggest that users' need for authenticity cannot be understood in a single, universal way (Section 4). Instead, it should be viewed through the lens of multiple "authenticities", each reflecting different user needs and contexts. In discussing the



C&C '25, Virtual, United Kingdom

ACM ISBN 979-8-4007-1289-0/25/06
https://doi.org/10.1145/3698061.3726918

workshop findings, we adopt the **"authenticities" lens as a tool** to reflect on the diverse demands, expectations, and contexts in which users engage with content (Section 5).

In summary the contributions of this paper are:

C1: Uncovering context-dependent expectations and needs for provenance data through the results from a workshop with experts across various stages of the creative supply chain.

C2: Broadening the value of provenance data by acknowledging how, while establishing authenticity and enhancing transparency are essential goals, provenance data should also support broader needs, such as ensuring accurate attribution, clarifying creative intent, and supporting the proper management of rights of creative workers.

C3: Acknowledging the nuances of provenance and suggesting the lens of multiple "authenticities" as a useful tool to reflect on how to design flexible provenance systems that can support diverse user needs.

2 Related Work

2.1 Technical Tooling for Media Authenticity

There is an increasing need to find tools and solutions to check and verify digital content on behalf of users [19]. **Media Provenance** tools help identify the origins and operations performed on content to help contextualize users' decisions on authenticity. This information can also enable creators to assert authorship over content. A significant development is the creation of the Content Provenance and Authenticity (C2PA) [12] technical standard by a broad coalition of technology companies like Adobe, Microsoft, Intel, BBC, OpenAI, and camera manufacturers such as Sony, Nikon, Canon, and Leica. More recently, Collomosse and Parsons [15] have argued that these technical solutions are not enough in isolation to support permanent provenance documentation. For example, information is often lost as content is remixed, reformatted and republished across different digital platforms, such as is the case with social media. Therefore, the authors propose a multi-faceted approach to provide *"three pillars of provenance"* formed of metadata, fingerprints, and watermarking. *Metadata* refers to information about provenance carried within the media file; *fingerprinting* uses hashes to search for a copy of the metadata in a trusted media repository should the original be stripped away when media is shared; and *watermarks* are invisible signals embedded into the image which can be used for easily identifying content.

Recent C2PA adoption has been driven by the need to indicate the extent of generative AI (GenAI) use in media content. However, this indication is often presented to users in a binary way. Recent implementations by LinkedIn and Meta initially signaled entire images to be AI-generated, even if only a small region was edited with a GenAI tool [57]. C2PA includes a user experience (UX) specification that describes *how* provenance information should be exposed to users at selectable levels of detail [11]. These solutions require a degree of media literacy to ensure that the provenance information presented is meaningful to the user, highlighting a need for socio-technical solutions and an HCI and Design-oriented focus.

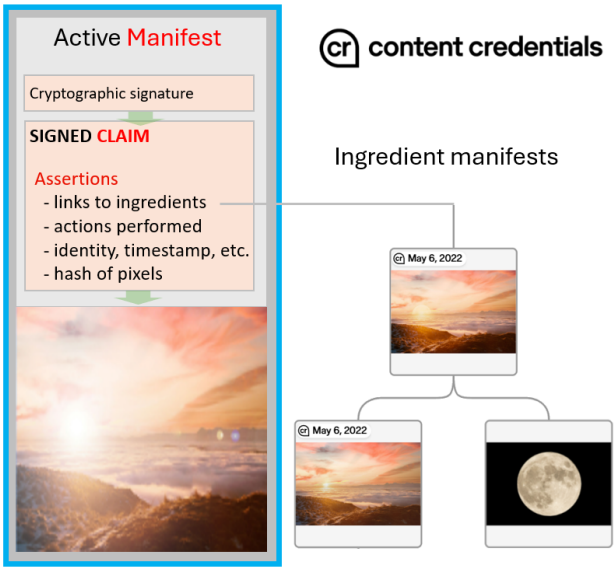


Figure 1: C2PA Provenance Standard: An asset, such an image, carries C2PA metadata (a 'manifest') within its XMP metadata. The manifest contains a cryptographically signed 'claim' that in turn contains facts ('assertions') about the provenance of the asset. This can include reference to 'ingredient' assets used in its creation, so forming a graph of its provenance history.

2.2 HCI Research on Supporting Media Creativity and Consumption with Provenance Data

2.2.1 HCI and Creativity Support Tools. A decade after the inaugural symposium on Creativity and Cognition in 1993, Candy and Hori encouraged the field towards an increased focus on researching **how Creativity Support Tools (CSTs) can benefit people in a wide range of domains** [13]. The last 20 years saw an increased interest in the HCI community for research around CSTs, an interest that has recently extended to modes of interaction between users and CSTs powered by AI technologies (AI-CSTs). Numerous studies have been conducted in this domain, from the sociology and overall motivations in the use of AI-CSTs [51], to studies on more specific applications, such as large language models, and their effectiveness in supporting users in their creative work [14], to studies on how to modulate the level of agency and control between users and AI in co-creative processes [45]. Rather than focusing on particular use cases of AI-CSTs, the work we present here contributes to the literature which is aimed at capturing the **perceptions, challenges, and benefits from both creatives and users when creating or consuming content that has been generated with AI** [49, 50], and the potential design intervention that can be taken to protect creatives' sense of agency and ownership on their creations [35].

2.2.2 ***Historical Perspectives on Provenance and Authenticity.*** Concerns over the authenticity and provenance of creative

content are not new but part of a broader pattern where technological shifts prompt renewed discussions about authenticity and ownership. Walter Benjamin's seminal 1935 essay, *The Work of Art in the Age of Mechanical Reproduction* [4], argues that technological advances, from the printing press to photography and film, have continually disrupted traditional notions of artistic authenticity and aura. This challenge to authenticity has continued with recent advances in technology such as augmented reality or 3-D scanning which have prompted questions over the 'migratable' nature of aura in art and cultural digital replicas, and challenging to what extent authenticity can be seen as an intrinsic value existing in a singular material work [36, 38]. The question of who can make such judgments on authenticity is also important, the philosopher David Hume [31] posed this question, challenging who are the "true judges" who can determine the authenticity of artworks. Who makes these judgments and how permeates contemporary debates on authentication practices in the art market, and technologies are increasingly being used in the triangulation of art historical documentation, scientific analysis, and connoisseurship that are used to make claims of authenticity [7]. Media and cultural studies have also examined how emerging technologies complicate provenance, authorship, and originality through the remediation of digital media as one form of media is represented in a new form [6, 42]. The challenges that AI development poses to authenticity and content provenance are situated within this **longer trajectory of media transformations** and of a corresponding evolution of the discourse on authenticity alongside emerging technologies.

2.2.3 ***Misinformation interventions.*** Recent years have seen a growing body of work within HCI that seeks to address the harms of misinformation propagated across online media, for example online news, social media networks and search engines [10, 23, 24, 28, 29, 37, 52, 60, 61]. This research highlights the need to effectively engage media consumers in combating misinformation. It explores how misinformation spreads, the impact of media design on this process, and proposes design interventions to mitigate harms. Some **design interventions** focus on tackling misinformation within specific media types, for example video only [28, 52], or visual media (video and still images) [23, 60], both text and image in combination, as is common in news and social media sites [24, 37], and text only (though supporting image) [61]. Hartwig et al. [29] offer a systematic review of misinformation literature and find that there is currently less work focused on non-textual media and video in particular. They also provide a useful taxonomy of interventions to date, describing them in relation to three categories: (i) intervention design, e.g., correction or "fact-checkers" [61], showing indicators, binary labeling [23, 24, 52], visibility reduction [37], (ii) user interaction (active/passive), and (iii) timing, e.g., pre-exposure, post-exposure. This taxonomy provides a way to view both the range, gaps and limitations of current design interventions.

2.2.4 ***The binary approach.*** Intervention design work in HCI also explores the **user experience**, and Feng et al. [23] and Zavolokina et al. [61] in particular study how to maintain a smooth media consumption experience in order to sustain user acceptance. This work considers Kahneman's "thinking fast, thinking slow" [33] to better understand the two distinctive modes of thought – **intuitive (fast) and reflective (slow)** – in relation to their designs. This helps make sense of the tension between providing detailed, transparent information to assist users in developing their own critical thinking (slow), and enabling users to maintain a fluid consumption experience (fast). When users are employing a fast-thinking approach, it makes sense to maintain positive consumer acceptance by using a binary approach, which simply warns the user if something is amiss, rather than expecting them to work it out for themselves. However, user groups actually involved in the creation and production of media may be working in a critical or slow-thinking mode when parsing or employing provenance information, so may have different tolerances and needs in this regard.

2.2.5 ***The unintended consequences.*** The simple binary design approach to misinformation design interventions has been found effective at improving credibility assessment and in particular reducing trust in misleading information [24, 37]. However, studies have also revealed that this approach can have unintended consequences. Most notably, users may **overgeneralize** the meaning of indicators, so that raising awareness of misinformation increases suspicion and **skepticism** of factual information [23, 30]. The same effect has been found in approaches that have studied binary labeling of content as AI [60]. Sherman et al. [52] found that some users overgeneralized positive confirmation signals leading to trust in potential misinformation. These effects arise from the delivery of definitive indicators to users in fast, or what Hartwig et al. [29] classify as a passive, mode of consumption, where critical thinking is not emphasized.

2.2.6 ***Promising ways forward.*** Provenance or source indicators have been demonstrated to be a valuable heuristic for users in establishing credibility [52] and providing transparency [29]. In addition, Sherman et al. [52] found that presenting the content's source was effective for identifying false information regardless of media type. The provision of explanations of signals has found to be both useful and desirable [37] with Feng et al. [23] demonstrating that presenting a journey of changes made to content has notable beneficial effects on both user trust and accuracy. There is a delicate **trade-off between simplicity and transparency** in this space. To achieve greater clarity on what makes a design intervention effective, Wittenberg et al. [60] state that clearly established goals for any design intervention are vital. To address issues around establishing what constitutes efficacy and good practice, in 2023 Guay et al. [27] set out appropriate research methods for this work, and in 2024 Hartwig et al. [29] found that much of the prior work in this area provides contradictory findings. It is clear that this work is still in its early stages.

2.2.7 ***Beyond misinformation.*** Beyond the focus on misinformation we find other examples of designing with provenance data in HCI. Underpinned by the development of novel technologies to support trusted provenance data, **design projects** such as the Bitbarista [53] and Tales of Things [16] explore the use and perceptions of provenance data displayed as part of coffee consumption (to choose a country of origin for coffee beans) [53] and second-hand goods purchases (to tell stories about objects prior lives) [16]. These projects tell us about the potential purpose, value and experience

of integrating provenance data into consumption processes. Within **data visualization**, Vancisin et al. [56] recognize the potential complexity of visualizing long histories of provenance data for historic documents, and explore solutions to some of these hard problems. Whilst these projects offer insights into the potentials and challenges of designing with provenance data, they are disparate. In this paper we begin to pull these threads together, to advocate for making the design of provenance data a central focus in future HCI and design research for creativity support tools.

3 Workshop on the Value of Provenance Data

We engaged a group of 34 experts in the exploratory online workshop "The Value of Provenance Data in the Age of AI" on 17th July 2024. Our objectives were to (i) **reflect** on the value that users throughout the creative media supply chain attribute to having access to diverse provenance data regarding media content, (ii) **discuss** which key signals they perceive as most relevant, (iii) **gain insights** into how AI is transforming users' expectations and requirements for these signals, and (iv) **identify context-sensitive perspectives** on the use and significance of provenance information across different stages of the creative life cycle (C1, C3).

3.1 Participant Recruitment

Participants were recruited using a combination of convenience sampling and snowballing techniques. The event was advertised on Eventbrite and promoted through targeted channels, such as newsletters and networks connected to the authors. The goal of this recruitment strategy was to attract individuals with a pre-existing interest in content provenance, without enforcing strict selection criteria. This approach resulted in a participant pool that was both engaged and informed, reflecting a medium to high level of media literacy. The expert nature of the participants is crucial to our findings. Unlike studies that focus solely on end-users [29], this research prioritized the perspectives of **experts** and professionals whose interactions with provenance data span the entire life-cycle of content, **from its creation and distribution to its final consumption**. As emerged from the discussions, participants' background ranged from creative industry policy and research, human-data interaction, media studies, user experience, service design, ethics, digital humanities, communication design, and games and immersive experience. Data collection for this study was conducted in compliance with the ethical guidelines and regulations set forth by the University of Edinburgh. A total of **34** participants attended the workshop (referred to anonymously with numbers within parentheses).

3.2 Workshop Design

The workshop was hosted online on Microsoft Teams, and the collaborative sessions were recorded and transcribed. The workshop opened with two 10-minutes talks by Laura Ellis, Head of Technology Forecasting, BBC and Mansoor Ahmed-Rengers, Founder of OpenOrigins and researcher at the Department of Computer Science and Technology at University of Cambridge. The aim of the talks was to provide an overview of the domain and sensitize participants to the focus of the workshop without delving into context specific values and expectations of provenance data, which was the focus of discussions in the subsequent participatory activities.

For the discussion, participants were randomly divided into three breakout rooms, moderated by two workshop organizers. These discussions were facilitated on a board on the collaborative online platform Miro and split into the following two parts.

In **Part 1**, participants were asked to watch **three short video clips** given as provocations, rather than test objects, for participants to respond to (Appendix A). Each group was presented with different clips, which ranged from political campaign ads and news reports to creative concept videos and event coverage and included both AI-mediated (meaning AI-generated or AI-edited) footage and unedited content. This mix was intentional, as it could allow participants to assess the differences in their needs and expectations about provenance data across a range of media content.

Time was provided for participants to individually answer **five questions** (Appendix B) and record their answers on sticky notes before moving to a group discussion. The first three questions were designed to examine participants' attitudes toward how the involvement of AI in content creation influences the perceived need for provenance information, trust, and verification processes. Here, participants were invited to position themselves as **consumers**, tapping into their habitual way of consuming and interpreting media. The last two questions asked them to imagine themselves as **producers**, in particular as editors at a news organization responsible for verifying the provenance of content before publication. They were challenged to think critically about the processes, tools, and information needed to verify the provenance of AI-mediated content compared to more traditional, human-generated content.

Part 2 of the workshop encouraged participants to **identify the types of information they considered crucial for assessing the background and authenticity of media content**. Participants were invited to add notes to predefined options they felt were most valuable, and provide justifications for their choices. They were also able to create new options (Figure 2 and Appendix B).

3.3 Workshop Constraints

This workshop convened a large number of experts spanning different areas of relevant expertise and interest. While this had clear benefits in broadening the discussion, we also recognize drawbacks to this approach. Acknowledging the need to gather consistent feedback from a large number of contributors for feasible analysis [21], the conversations were allowed to flow freely, while the participatory activities involving sticky notes were carefully predefined and strictly timed. Due to the numbers of participants and limited time of the workshop it was not possible for all participants to discuss content most important to them or their specific work. Finally, in Part 2, we are conscious that the given examples may have influenced participants' choices, as they responded to these first.

3.4 Data Collection and Analysis

Data collection during the workshop encompassed different methods including participants recording their responses on sticky notes or through the commenting function, along with gathering data from discussions. The primary data for this study consisted of the participants' sticky notes (275 notes) and transcriptions from the

three breakout rooms (221 minutes of recordings transcribed). Transcriptions were automatically generated by Microsoft Teams and subsequently reviewed by Author 1 for accuracy.

Following the workshop, these transcripts were analyzed using a **reflexive thematic analysis approach** [8, 46], starting more inductively, and moving to a more deductive approach later, when applying the authenticities reflective lens (subsection 3.5). In the initial phase, Author 1 identified codes within the data. After this first iteration of coding, a collaborative discussion took place among Authors 1, 2, 3, and 6. During this meeting, the team reviewed the initial findings, debated interpretations, and considered additional themes that may not have been immediately apparent in the first round of analysis. Following the discussion, Author 1 conducted a second iteration of coding, revisiting the transcripts with the new insights gained from the group discussion. The key themes developed through the second iteration are described in detail in Section 4.

3.5 Content Authenticities as a Reflective Lens

Efforts to create widely adopted and inter-operable provenance standards, such as those led by C2PA, speak to users' needs to verify that the content they encounter online is "authentic". But the need that consumers have for authenticity extends beyond the wish to seek transparency regarding where content comes from, reflecting a broader desire for integrity, genuineness, and meaningful connections to the products and content they engage with [26].

The complex nature of authenticity makes it a rich area for research exploration. The discussions within the research team that followed the thematic analysis by Author 1, led us to consider the concept of **multiple content "authenticities"** [44] as a further analytical lens to reflect on our initial findings and categorize the different ways authenticity is understood and valued by users (Section 5). In doing so, we drew from established theories of authenticity in art, media, and consumer theory by Dutton [17], Newman and Smith [47], and Grayson and Martinec [26], whose core arguments we briefly present here.

Dutton refers to the identification of "the origins, authorship, or provenance of an object" as "**nominal authenticity**" [17]. This type of authenticity focuses on the correct attribution of an object or content, thereby enhancing its perceived legitimacy. Newman and Smith's concept of "**historical authenticity**" involves "the evaluation of an object's unique spatio-temporal history" [47, p. 612]. Importantly, such judgments tend to be binary in nature (Is it the thing or not?) and seem to assume that authenticity is something that can be verified via an objective, external source such as notes of provenance and expert evaluation. For Newman and Smith's "**categorical authenticity**," instead, "judgments may be graded and, critically, do not require an external source of validation. In fact, this type of authenticity seems related to the notion that authenticity judgments are dictated by observers' own expectations and, therefore, may be much more subjective in nature." [47, p. 613] Similarly, through their notion of "**iconic authenticity**", Grayson and Martinec note how our interpretation of indicators may be "highly influenced by our personal predilections and perceptual imperfections" [26, p. 299]. Rather than "indicators", they become "signs" or "icons", whose meaning is not objective, but instead dependent on the interpretation that it generates in the users of the sign [2]. For Dutton's "**expressive authenticity**," an object's authenticity is partly defined by its ability to express the values and beliefs of its creator or society [17]. Similarly, Beverland et al.'s [5] concept of "**moral authenticity**" shifts the focus from the end product to the creators' intentions, and Newman and Smith's concept of "**value authenticity**" emphasizes normative considerations, such as whether creators are motivated by intrinsic values rather than greed, and whether their behavior reflects the cultural values of given community [47].

We applied these different kinds of authenticity post-hoc, when reflecting on the workshop findings, in order to bring greater clarity and nuance to the insights gained, particularly in exploring the expectations and demands expressed by participants. We use these different categories later in the paper to structure our discussion (Section 5).

4 Findings

In this section we develop key insights about the contextual nature of provenance data (Section 4.1), the challenges of binary labeling of AI inputs (4.2), and the risk of excessive transparency (4.3). We also share participants views on how provenance signals may be tailored to different users and forms of engagement (4.4).

4.1 The Need for Provenance Signals is Context and Content-Dependent

The workshop discussions revealed how the need for provenance signals varied significantly depending on both the context in which the content is consumed and the nature of the content itself. Participants emphasized that **provenance is not a one-size-fits-all solution**. Rather, its importance varies on the basis of different factors, depending upon: the **type of content** (e.g., news vs fiction), the **platform** on which it is presented (e.g., social media vs traditional news outlets), the **reputation** of the source, and the level of the user's **media literacy**. These points are illustrated when participants compared how they might react differently to labeling or content depending on the context for example: *"I tend to trust certain sources for factual information."* (24); "*if I saw it on a documentary for some reason, I would assume that historians have been consulted with the visuals and they would have some level of accuracy, whereas if I'm on social media, I think I'd be very hesitant to believe anything that was in it. Context is everything.*" (7)

There are also cases where provenance signals may pose risks to individuals, such as journalists acting in environments where freedom of the press is limited or working with confidential sources. However it remains crucial to have access to the provenance or history of the document itself, while maintaining privacy: "*there are definitely times where you wouldn't want that person's individual identity to be exposed, but you would want to know the kind of provenance or history of that document.*" (10)

Participants expressed a higher demand for provenance signals when the content is vague or lesser-known. This applies also to cases where verification is less easily attainable due to the anonymity of content: *"if I recognize the person in the video, it's helpful to know that it's AI generated but I can check up if I doubt it. In the last video*

clip, it's a crowd scene and everything is essentially anonymous, and there's no way to check and in that context I thought it would be more significant to know how that was generated." (33)

The abundance of different contexts of content creation and consumption mentioned by participants presents a complex picture: it ranges from political content, to propaganda, news, social media feeds, fiction, entertainment, tourism, education, documentary, sports story, and artistic explorations, among others. The "context of creation" was also the second most selected among the detail options when assessing the authenticity of media content in Part 2 of the workshop, reflecting the need for more contextual details in provenance signals (Figure 2).

4.2 Human vs AI: Not a Useful Distinction

The context-dependency of provenance extends to the attitudes toward AI-mediated content; why the AI was used might matter in some contexts and not in others: *"I like to know intent for sure if it is news, but for art I think that could remain a mystery."* (24) Also the **purpose for which AI tools are used** can affect the its acceptability: "*Depends on informational context — if it's being used to speed up an editing process or to create a deliberate simulation/reenactment etc. where we would ALREADY expect tools to be used for that, I don't think knowing whether AI is among the toolkit is that important. But if it's materially affecting the presentation of the content, then it is important to know this.*" (10) In particular, knowing whether the AI had been used only to **enhance the quality of the content** (1, 6, 16, 21, 24, 29), or rather to **change its semantics** is felt as a key priority: "*was it* [the AI] *changing the semantics or was it just making it sharper? That's what I would like to know.*" (34)

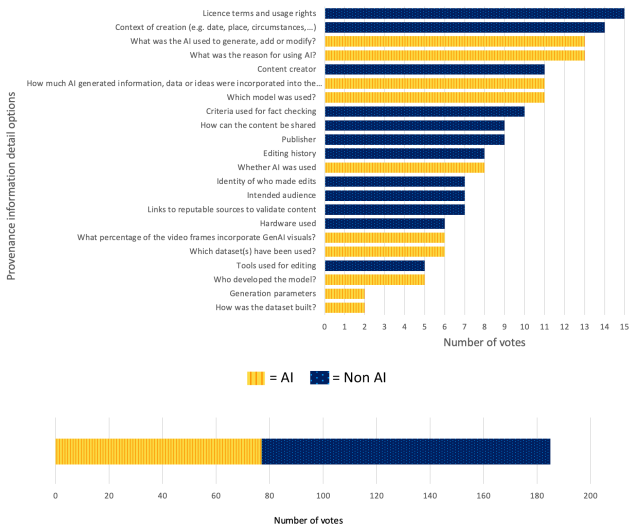


Figure 2: Provenance Detail Preferences: These charts illustrate the number of votes for each provenance detail option presented in Part 2 of the workshop and the proportion of votes for AI related provenance details versus non-AI related details.

Workshop participants expressed that the **binary distinction between human and AI** is "*dissatisfying*" (21) and not meaningful enough. This theme emerged from discussions that questioned the usefulness of labeling content as either human or AI-generated, emphasizing the need for more nuanced categories: "*We got into this idea that things are not binary. They're not necessary, altered or not altered. There might be things that don't change the sense at all, but are tidying up. There might be things that do change a sense there for good reasons or bad reasons.*" (27)

The increasing integration of AI into various stages of content production further complicates the distinction. Many participants felt that **AI is progressively being normalized in creative processes**, to the point where the line between human and AI involvement is becoming irrelevant. One participant noted: "*over time it's just going to be assumed that AI tools to some extent [are] taking labor out of this content production. I don't think the distinction between human and AI will matter as much.*" (34)

The need for more nuanced information regarding the scope and nature of AI intervention surfaced also in Part 2, where several participants indicated "What was the AI used to generate, add or modify." "What was the reason for using AI?," and "How much AI generated information, data or ideas were incorporated into the video." as highly relevant details to be included as part of provenance information (Figure 2).

4.3 The Dual Impact of Provenance Signals: Enhancing Transparency or Fostering Skepticism?

Participants noted how provenance signals are not merely indicators of how content was generated and/or edited, but also influence how audiences perceive the trustworthiness and intent behind the content itself. In particular, the labels "AI-generated" or "AI-edited" can raise suspicion or create a perception in the audience that the content might be misleading or manipulated [60]: "*For clip 1, if it were labeled as AI generated, I would instantly assume the intent was malicious and therefore it wasn't trustworthy.*" (10) Or *"At present, if I saw 'AI generated' there would be trust issues as AI has connections to 'fake news' or 'false'."* (29) This surfaces a broader reflection about algorithmic aversion, where people inherently distrust AI-mediated content. One participant summed this up by stating: "*I have an inherent mistrust of AI. I can't have a conversation with it. I can't see its eyes and know whether it's telling me the truth.*" (4)

Some participants expressed that seeing an AI label would prompt them to question the content more critically, even if they might not have done so otherwise: "*If it's not labeled, I probably wouldn't even think about it. But the fact that it is labeled, I would then be questioning: Which part of that was AI? What do I believe? What don't I believe?*" (23) This indicates that provenance signals can act as **triggers for deeper scrutiny**. Participants raised concerns that overemphasizing AI-mediated content might lead to a "*downward shift of trust*" (30) in media more broadly, and this could result in audiences dismissing legitimate media as fake, complicating trust dynamics between creators, publishers and the audience [18, 32].

On the other hand, the absence of a label might lead to the assumption that no AI was involved, which could be equally misleading. This assumption could give rise to what a participant referred

to as "**human washing**" (30). The better known term "AI washing" refers to the trend where companies overemphasize or overstate the extent to which they are using AI in their business to attract investors and companies. "Human washing" is a newer term which is usually reserved to cases in which AI is deceptively programmed and designed to look and sound like humans [59].

To balance the need for transparency with the risk of fostering excessive skepticism, some participants suggested that a more general shift in the audience's "*expectations*" (27) and familiarity toward provenance signals might be beneficial (3). Indeed, the fact that users are not yet accustomed to seeing such labels on media content might amplify their suspicion and makes them wonder which parts and why/how AI had been used.

4.4 Provenance Signals for Diverse Types of User Engagement

Participants' observations during the workshop highlighted how provenance signals need to be **tailored to the needs, literacy levels, and individual motivations** of different user groups. Some users, especially those with higher media literacy, or with knowledge in and specific interest for content provenance, may actively seek detailed provenance information in order to critically evaluate content. However, a user who is passively consuming content may overlook such information entirely. As one participant noted: "*the general viewer might not consider how things are made or not necessarily the why, but how things are made is not something we naturally think about unless you're specifically in that area of interest.*" (23)

This, however, should not be interpreted as a binary distinction between different types of users. Rather, it would be more appropriate to talk about different types of **user engagement**. Indeed, the same user might have different demands depending on the context. For example, a participant who is actively involved in research projects around content provenance said: "[provenance] *is not something I would think about if I'm watching a TV programme sometimes we just want to watch something and not think about it.*" (23)

In some cases, especially in the case of "fast" interaction with content (Section 2.2), minimal and unobtrusive signals specifically related to the reputation of the source might be sufficient, in cases of "slow" content consumption, instead, more detailed information might be required: "*I just don't think that your average person looking at news information online is conditioned to ask those kinds of questions about the content. And so for them, the hierarchy of what they care about is the source from where they're getting that information, whether that source is like the BBC, or Fox News, or from Aunt Karen or whoever.*" (6) When users lack the means or time to examine the provenance trail themselves, they may turn to expert testimony for guidance: "*I don't want to do the work myself if I'm feeling trusting (surely the whole point is that the organisation presenting it has robust fact-checking in place; it's not my job!)*" (10)

Participants in our workshop acknowledged the importance for provenance systems to consider how to **balance the need for transparency of users and their cognitive load limits**: "*I think it's cognitive load as well. If you're scrolling through news media, you're watching the news. Do you really want to stop and interrogate the metadata of everything that you go through. So I just don't see people doing that every single time.*" (33) Information overload is not the only concern. For users who are more prone to **cognitive biases**, such as confirmation bias [20, 54], even well-designed provenance signals may be ineffective if the information contradicts their ideological stance: "*I don't know what to do about the audience that wants to hear things that confirm their views, because if it doesn't agree with what they want to hear it's actually unhelpful that it's expert testimony, because that gives them a reason to refute it.*" (34)

4.5 The Value of Communicating Intent

The importance of the creator's intent was a central theme during the discussion. Participants emphasized that understanding the purpose behind the creation and dissemination of content, as well as behind the use of AI, is crucial for assessing its authenticity and trustworthiness. Immediate and easily accessible indicators of how much users can trust content are welcome, but they might not be enough if they do not provide users with details regarding the intent with which the content was generated in the first place, or subsequently edited and disseminated: *"Absent a strong trust relationship with the source, C2PA-style credential would be mechanistically-useful, but intent is still important."* (30), *"I think the label has to be thoughtful. Just knowing that it was or wasn't AI generated. It doesn't tell you anything about the intention of the person who put the AI on the content, as it were.*" (10), and *"I think it just becomes so immensely complicated that you just keep coming back to intent the whole time.*" (27) In a future where content is increasingly mediated by AI, **the intentions behind content creation might even become a marker of quality**: "*I think it comes down to a lot about intent and "I think we are going to see human-generated content become sort of premium and maybe we'll need ways to seek it out.*" (27)

Provenance signals were discussed by participants also as a means to **give creators more agency** over the communication of their intent to the audience. This is particularly relevant in a context where AI can easily replicate and modify content, with the consequence that creators may struggle to maintain control over their creations. One participant expressed this concern: "*[provenance is important] for small creators and artists to have control over what they've created and to stop broad AI pilfering and reusage.*" (1) The notion of provenance as a protective measure for creators also extends to the recognition and tracking of modifications made to their work. This could prevent situations where creators encounter altered versions of their content that they did not authorize: "*I guess I'm wondering if provenance has some really positive value there about helping them understand the ways in which their work might get, you know, to help avoid situations where a creator sees a version of their work that they don't recognise or they weren't expecting.*" (30) For creators, having a system that tracks modifications could provide reassurance and allow them to retain agency over their creations, even when these creations might be later modified to respond to personalization needs by users: "*We want them to retain agency, right? We want them to be confident to author media that might be personalised.*" (30)

Lastly, the theme of intent also intersects with the need for **contextual understanding**. Participants discussed that content,

especially if it serves purposes like satire (34), might require additional labeling to clarify its intent. The need for more information around the context and intent behind content creation is even more acute in today's media landscape, where disintermediated content — content that is consumed without traditional gatekeepers like editors or journalists — proliferates. Without these intermediaries, the loss of contextual information can make it difficult for users to assess the authenticity or original intent of the content, thereby increasing the need for clear provenance signals: "*Because we see so much disintermediated stuff, I'd like to see some kind of ways of signalling: this is satire, you know, this is made by synthetic media in a satirical way.*" (27)

4.6 Is it Authentic or is it True?

As Figure 3 shows, when discussing the value of provenance signals, the terminology used by participants is varied, and authenticity is only one of the many terms that are used, often without acknowledging the semantic distinction between them. Terms like "accurate," "credible," or "correct" might assume different meanings to different users and might also depend on the context in which these users are active.

Even among experts the term "authenticity" is rarely used. Rather, users seem to prioritize the "**truth**" or "**integrity**" of the content they consume. There is a high risk that users misunderstand provenance signals as signals of the "truthfulness" of the content they are referring to, rather than of its origin and historical trail. For example, in Part 2 of the workshop, when asked to indicate which provenance information detail they found most relevant, one participant was mostly interested in provenance indicators to signal whether the clips they were watching were "real": *"Are these images real, do they correspond to real things?*" (30) While authenticity relates to whether something is what it claims to be, truth goes further, representing the ultimate goal of accuracy and factual integrity of the content. One participant in particular highlighted the importance of recognizing this distinction: "*I think we need to be careful that we're not conflating authentic footage that hasn't been generated with AI, with content that is like correct or true.*" (33)

5 Discussion

The variety of diverse user demands that are expressed by the workshop participants surface multiple reasons why users might value having access to provenance signals. Many of the considerations made during the workshop resonate with prior studies, (e.g., Burrus et al. [10]), which point to the influence of context in shaping the value and uses of provenance data (Section 4.1). Our findings extend this understanding by offering nuanced insights into how context shapes the relevance and need for provenance data from the creation and production as well as consumption viewpoint (C1, C2). We do this, for example, by reflecting on how the specific nature and extent of AI contribution in the creation or alteration of content might be deemed relevant in some contexts of applications but not in others (Section 4.2), and on how the same user might engage with content differently and have different demands according to the context of use (Section 4.4).

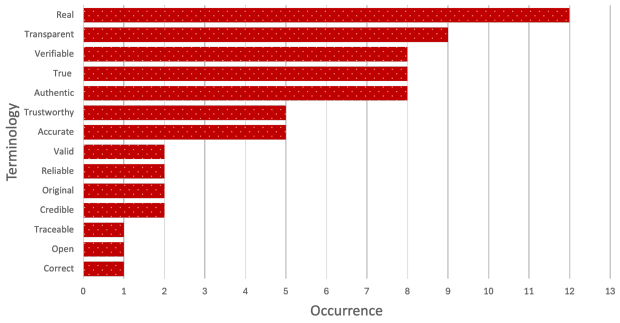


Figure 3: Authenticity-related Terminology: this bar graph displays the frequency of terminology associated with the concept of authenticity that was used by participants. The y-axis lists key terms associated with 'authenticity' that were used during the workshop, while the x-axis shows how frequently each term appeared in participants' notes and breakout room discussions.

We believe that this variety of experiences, contexts, and needs can best be understood not under a single understanding of authenticity but rather considering many different "**authenticities**". In our discussion, we use the theories of authenticity introduced in Section 3.5 to discuss workshop findings, bringing a new dimension and clarity to these insights in exploring expectations and demands expressed by participants. In particular: the need for insights into the **content origin** and the tools used in its creation (Section 5.1), the necessity for provenance signals to adapt to the **different kinds of user engagement** (Section 5.2), and the users' wish to know more about the creator's **intent** (Section 5.3). One of the key aims of this work is to **acknowledge the multifaceted nature of provenance by unpacking its nuances** (C3). The concept of multiple "authenticities" serves as a lens that is not only theoretically interesting, but can also be practically valuable in designing provenance systems that support diverse conceptions of authenticity, ensuring that these systems are flexible and relevant to various user needs and contexts.

5.1 I Want to Know Where the Content "Comes From": Nominal and Historical Authenticity

One of the themes discussed in Section 4 was the association of authenticity with notions of "truth" and "reality" (Section 4.6). Participants frequently expressed the desire to be able to assess the correspondence of the content with "facts". **Truth as correspondence** between statements and facts has a long history, tracing back to the works by Aristotle and Aquinas [25] and the access to provenance "facts" seem to be interpreted by users as a way to establish this connection. As highlighted by one of the workshop participants, however, **authenticity does not necessarily equate to truth**. For example, a video can be authentic in the sense that it has not been manipulated by AI or another human individual, but it may still be misleading or framed in a way that distorts the facts. Similarly, labels and provenance signals should not be seen

as guarantees of truth but rather as tools that help users make informed trust decisions.

Undoubtedly, a primary need for users is to know **where the content they are consuming comes from**, thereby speaking to interpretations of authenticity which link it to the objective verification of an object's unique history [17, 47]. Dutton's "**nominal authenticity**" [17] and its focus on the origins and authorship of content resonates with the participants' concerns about accessing information regarding, for instance, the identity of the content creator, the hardware and tools that were used, and the details about the AI model if the content is AI-mediated (Figure 2). Having access to the raw files is perceived as important for knowing where the content was sourced from and for verifying the source itself. Factual indicators like these function as attestation of content origin.

However, assessing authenticity is not always straightforward. Previous literature highlighted how when users adopt a fast-thinking approach in consuming content, it is more appropriate to provide them with simple and accessible provenance signals [23, 61]. An insight that resulted from the workshop is that, sometimes, when the content's provenance is fragmented or when users don't have the time, or wish, to interrogate the provenance trail themselves, they might rely on expert testimony (Section 4.4). This idea aligns with Newman and Smith's concept of "**historical authenticity**", which involves the assessment of an object's distinct history [47].

5.2 "Who Are We Doing This For?": Iconic and Categorical Authenticity

Users' expectations can be expertly manipulated by letting "fake" content appear as "authentic", exploiting the influence that "personal predilections and perceptual imperfection" – following Grayson and Martinec's [26] notion of "**iconic authenticity**" – can have on the interpretation of indicators. Trusted sources lose their role and effectiveness in establishing provenance if users are not willing to let their "personal predilections" and already established assumptions be challenged by external sources, as noted in previous work [19, 58]. This consideration underscores the challenge of designing provenance systems that are universally trusted and that take individual biases into account.

The complex notion of 'truth' has become further problematized with the rise of misinformation, disinformation, and fake news. The long-standing correspondence theory of truth loses its hold in a context where public opinion is more influenced by emotional appeals and personal beliefs than by objective facts and evidence. In this context, the notion of "**post-truth**" emerges [1]. Just as "post-truth" challenges the reliance on factual accuracy, we can use the term "**post-authenticity**" to refer to a context which challenges the idea that something being "authentic" inherently holds value or truth. In a "post-authenticity" context, the traditional markers of authenticity, such as originality, provenance, or the "aura" of a work [4, 39], may no longer hold the same significance. Particularly with the rise of AI-mediated content, the boundaries between the original and the manipulated, the genuine and the artificial, are increasingly blurred. The notion of "authenticity" becomes fluid and the focus shifts from verifying the content's origin and history, to understanding its **context, purpose, and impact**.

This is in line with Newman and Smith's "**categorical authenticity**" and its emphasis on the central role that the viewer can play in identifying and interpreting the authenticity of content [47]. As Dutton [17] notes, discussions of authenticity often overlook the role played by the audience in establishing the context for creative content. This observation directly ties into one of the themes developed in the findings: "***Who are we doing this for?***" (27) (Section 4.4). Interpretations which center authenticity on the subjective observer's expectations, like Newman and Smith's [47], can be more suited to understand how to tailor the display of provenance signals to different types of users and their modes of engagement with content. The question "Who are we doing this for?" must also address the idiosyncratic needs of both **consumers** and **producers** regarding provenance data. For consumers, provenance signals can serve to reduce uncertainty or enhance trust, but these needs are not uniform and often depend on the specific use-case or context. For producers, who create and/or disseminate content, provenance data might serve as a tool for asserting credibility, protecting intellectual property, or communicating value to targeted audiences.

5.3 "Intent is Key": Expressive and Value Authenticity

Section 4.5 surfaced an original insight that resulted from our workshop: **users have a specific interest in understanding the creators' underlying motivations and how they shape the content origins and history**. Insights from the production perspectives shared by participants have revealed how provenance signals that include information about the creator's intent can enhance the accuracy and reliability of content interpretation, addressing potential ambiguities and ensuring that content is evaluated in its proper context. Knowing more about the intent behind engaging with AI tools for content creation or editing might also address the risk of a generalized downward shift in user trust which might concern both trustworthy and untrustworthy content (Section 4.3). The relevance of intent for assessing the authenticity of content aligns with theoretical frameworks, such as Beverland et al.'s concept of "**moral authenticity**" [5], Dutton's "**expressive authenticity**" [17], and Newman and Smith's "**value authenticity**" [47], which are all interested in shifting the attention from the end product of creation to the creators' intentions.

By preserving and communicating their intent, provenance signals can also support creators in protecting their authorship over content and preventing unauthorized alterations or misrepresentations. For example, recent studies on media tokenization and C2PA under the ORA (Ownership, Rights, and Attribution) framework demonstrate this possible functionality [3]. Tokenization refers to the use of blockchains and non-fungible tokens to create distinct digital content, and by employing ORA in their work, creators are able to tokenize both their work and bespoke licenses of their work, each of which contain embedded metadata and provenance data [40]. Such systems hold opportunities for maintaining expressive authenticity in digital media as, once work is shared and reused using ORA, creators are able to trace their contributions. These contributions are marked out in provenance signals, embedding a kind of 'DNA' element to work that is reflective of "**expressive authenticity**" [17] and "**value authenticity**" [47]. Therefore, provenance

signals have potential benefits not only for users, but also for creators who can regain their agency over their work as it is shared online.

The emphasis on the creator's intent speaks to the paper's aim of **broadening the value of provenance** (C2). Beyond providing signals of trustworthiness, workshop discussions stressed that provenance data hold significant potential for asserting the creator's authorship and intent. This additional role of provenance data can be fully appreciated and researched only by widening the focus to include production alongside consumption.

5.4 Future Work

Through the exploratory approach of the workshop, we aimed at understanding the situated response of participants in a new technological domain. Discussions often shifted to **production-specific requirements**, suggesting that production roles may require targeted approaches to meet unique needs. We anticipate that follow-up studies will deepen these insights by focusing on the production aspects that emerged as particularly relevant, e.g., the importance of the creator's intent.

In this paper, we also discussed how the "content authenticities" lens can help to categorize diverse user behavior and preferences when interacting with provenance data, recognizing the different mindsets (fast and slow) and cognitive biases with which users relate to media content, as either producers and/or consumers. A more practical and detailed framework could be developed to support the design and understanding of varied user experiences with provenance data on this basis. For instance, this might include a **tiered system of provenance information**, tailored to different user groups — ranging from those who require only basic indicators to those who seek more detailed provenance information. Resources such as the C2PA UX specifications [11] are promising, though additional work remains essential to fully address the complexity of users' needs and interactions with provenance data. Although this paper does not delve into specific design testing, future work could also consider applying the taxonomy by Hartwig et al. [29], which classifies different types of indicators, to analyze and categorize the types of provenance signals that better respond to the demands of each user type and the forms of authenticity which are significant in that context.

Beyond traditional uses, provenance data also holds potential as a **medium for artistic expression**, enabling creators to share their process and intent while shaping audience engagement. Critically assessing the social, cultural, and creative implications of provenance tools, is crucial to explore how they can be responsibly implemented to benefit creatives, and their audiences in the current AI-driven digital economy.

6 Conclusion

In this paper, we integrated insights from a workshop on the value of provenance signals with prior research on provenance within HCI. Findings from the workshop validate prior work but also extend it, indicating the need for broadening the value of provenance data beyond its focus on combating misinformation. As provenance data becomes increasingly standardized and common in everyday interactions with online content, users seek more detailed and nuanced information. We suggest that reflecting on the growing demand for content authenticity through the lens of multiple "authenticities" can be useful to identify and address the diverse user needs for, and uses of, provenance data.

Acknowledgments

We are grateful to all participants who have been involved in the workshop. This research was supported by the UK EPSRC Next Stage Digital Economy Centre in the Decentralised Digital Economy (DECaDE) (EP/T022485/1).

A Video Clips

Table 1: Video Clips Details: the table provides details regarding the video clips that were shown to participants during the workshop.

Room	Clip	Title	Link
1	1	Time for change, Scottish Labour Party	https://www.facebook.com/ScottishLabourParty/videos/1495128451423022/ [Full video]
	2	New Mexico Wildfire Claims 1,400 Structures, The Weather Channel	https://weather.com/news/video/1-dead-1400-structures-lost-in-raging-new-mexico-wildfire [Full video]
	3	AI-generated climate change in Alberta CA, Rachael Roland	https://www.youtube.com/watch?v=89eKh9sMSpU [1:07-1:31]
2	1	Photo Finish, @paralympics	https://www.youtube.com/watch?v=WNG_iQ7TIRM&t=31s [0:31-1:05]
	2	Satellite shows extent of terrible destruction to the planet, BBC	https://www.youtube.com/watch?v=L9zWDtDKDS8 [1:03-1:42]
	3	Roman Empire: Restored Footage concept video, @midaiartwork	https://www.instagram.com/reel/C8pwqaPKrJc/?igsh=cjZvb2lqcm0xcnds [Full video]
3	1	Thousands of activists across Europe protest climate change, Washington Post	https://www.youtube.com/watch?v=TltRbhC-v-c [0:00-0:34]
	2	From the river to the sea. Roger Waters at Piers Morgan show, @edworldview986	https://www.youtube.com/shorts/pdkmJB0drw8 [Full video]
	3	AfD protests: demonstrators clash with police at start of far-right congress, Guardian News	https://www.youtube.com/watch?v=_DemRBQ3QuU [0:00-0:38]

B Workshop Questions

Part 1

Questions:

1a. Imagine that these clips have been generated with the help of AI tools. Does the potential for AI involvement in creating this content affect your need for provenance information? How so?

1b. Imagine that Clip 1 is clearly labelled as "AI-generated" with some information about the AI tools used and the creation process. Clip 2 has no such labelling, and it's not immediately clear whether it was generated by a human or an AI. How does the presence or 1ence of labelling affect your perception and trust in the content?

2. Imagine that Clip 1 is labelled as "AI generated" and Clip 2 as "Human generated". How important is it for you to know the steps and tools used in generating the content in Clip 1? And in Clip 2? Tell us your thinking.

3a. Imagine you are an editor at a news organisation responsible for verifying the provenance of content before publication. You receive two clips: Clip 2 is labelled as "human-generated," and Clip 3 as "AI-generated" Consider the steps you would take to verify the provenance of each clip. How would your approach differ between the human-produced clip and the AI-generated one?

3b. Think about any additional information you might need to verify the AI-generated content. What specific challenges do you foresee in this process? How might the AI generation process complicate the verification of its provenance compared to traditional content?

Part 2

Instruction: "What information is important to know about the video in the centre's background and authenticity? Put a note on each aspect that you think is important (in the section titled 'Why is it important?')"

Instruction: "Look at the different detail options and add a sticky note to the ones you agree with. Optional: add why you think this is important inside the note."

Detail options:

- Content creator
- Context of creation (e.g., date, place, circumstances,...)
- Hardware used
- Intended audience
- Licence terms and usage rights
- Publisher
- Links to reputable sources to validate content
- Editing history
- Identity of who made edits
- How can the content be shared
- Tools used for editing
- Criteria used for fact checking
- Whether AI was used
- What percentage of the video frames incorporate GenAI visuals?
- How much AI generated information, data or ideas were incorporated into the video?
- What was the AI used to generate, add or modify?
- What was the reason for using AI?
- Which model was used?
- Who developed the model?
- Which dataset(s) have been used?
- How was the dataset built?
- Add your own here:

References

[1] John Ambrosio. 2022. Problematizing Truth-Telling in a Post-Truth World: Foucault, *Parrhesia*, and the Psycho-Social Subject. *Educational Philosophy and Theory* 54, 12 (2022), 2133–2144. https://doi.org/10.1080/00131857.2022.2034619

[2] Albert Atkin. 2023. Peirce's Theory of Signs. In *The Stanford Encyclopedia of Philosophy* (Spring 2023 ed.), Edward N. Zalta and Uri Nodelman (Eds.). Metaphysics Research Lab, Stanford University, Stanford.

[3] Kar Balan, Shruti Agarwal, Simon Jenni, Andy Parsons, Andrew Gilbert, and John Collomosse. 2023. EKILA: Synthetic media provenance and attribution for generative art. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE Computer Society, Vancouver, 913–922. https://doi.org/10.48550/arXiv.2304.04639

[4] Walter Benjamin. 2018. The work of art in the age of mechanical reproduction. In *A museum studies approach to heritage*. Routledge, London, 226–243.

[5] Michael Beverland, Adam Lindgreen, and Michiel Vink. 2008. Projecting Authenticity Through Advertising: Consumer Judgments of Advertisers' Claims. *Journal of Advertising - J ADVERTISING* 37 (04 2008), 5–15. https://doi.org/10.2753/JOA0091-3367370101

[6] Jay David Bolter and Richard Grusin. 2000. *Remediation: Understanding new media*. MIT Press, Cambridge, MA.

[7] Anna Bolz. 2023. *A Regulatory Framework for the Art Market? : Authenticity, Forgeries and the Role of Art Experts* (1st ed. 2023. ed.). Springer International Publishing, Cham.

[8] Virginia Braun and Victoria Clarke. 2020. *Thematic Analysis*. Springer International Publishing, Cham, 1–7. https://doi.org/10.1007/978-3-319-69909-7_3470-2

[9] Jasper David Brüns and Martin Meißner. 2024. Do you create your content yourself? Using generative artificial intelligence for social media content creation diminishes perceived brand authenticity. *Journal of Retailing and Consumer Services* 79 (2024), 103790. https://doi.org/10.1016/j.jretconser.2024.103790

[10] Olivia Burrus, Amanda Curtis, and Laura Herman. 2024. Unmasking AI: Informing Authenticity Decisions by Labeling AI-Generated Content. *Interactions* 31, 4 (jun 2024), 38–42. https://doi.org/10.1145/3665321

[11] C2PA. 2024. *C2PA User Experience Guidance for Implementers*. C2PA. Retrieved September 11, 2024 from https://c2pa.org/specifications/specifications/2.0/ux/UX_Recommendations.html

[12] C2PA. 2024. *Coalition for Content Provenance and Authenticity*. C2PA. Retrieved September 1, 2024 from https://c2pa.org/

[13] Linda Candy and Koichi Hori. 2003. The digital muse: HCI in support of creativity: "creativity and cognition" comes of age: towards a new discipline. *Interactions* 10, 4 (July 2003), 44–54. https://doi.org/10.1145/838830.838833

[14] Tuhin Chakrabarty, Vishakh Padmakumar, Faeze Brahman, and Smaranda Muresan. 2024. Creativity Support in the Age of Large Language Models: An Empirical Study Involving Professional Writers. In *Proceedings of the 16th Conference on Creativity & Cognition* (Chicago, IL, USA) *(C&C '24)*. Association for Computing Machinery, New York, NY, USA, 132–155. https://doi.org/10.1145/3635636.3656201

[15] John Collomosse, Andy Parsons, and Mike Potel. 2024. To Authenticity, and Beyond! Building Safe and Fair Generative AI Upon the Three Pillars of Provenance. *IEEE Comput. Graph. Appl.* 44, 3 (jun 2024), 82–90. https://doi.org/10.1109/MCG.2024.3380168

[16] Martin L. de Jode, Ralph Barthel, and Andrew Hudson-Smith. 2011. Tales of things: the story so far. In *Proceedings of the 2011 International Workshop on Networking and Object Memories for the Internet of Things* (Beijing, China) *(NoME-IoT '11)*. Association for Computing Machinery, New York, NY, USA, 19–20. https://doi.org/10.1145/2029932.2029940

[17] Denis Dutton. 2005. 258 Authenticity in Art. In *The Oxford Handbook of Aesthetics*. Oxford University Press, Oxford. https://doi.org/10.1093/oxfordhb/9780199279456.003.0014

[18] Cynthia Dwork and Martha Minow. 2022. Distrust of Artificial Intelligence: Sources & Responses from Computer Science & Law. *Daedalus* 151, 2 (05 2022), 309–321. https://doi.org/10.1162/daed_a_01918 arXiv:https://direct.mit.edu/daed/article-pdf/151/2/309/2060588/daed_a_01918.pdf

[19] Stephanie Edgerly, Rachel R. Mourão, Esther Thorson, and Samuel M. Tham. 2020. When Do Audiences Verify? How Perceptions About Message and Source Influence Audience Verification of News Headlines. *Journalism & Mass Communication Quarterly* 97, 1 (2020), 52–71. https://doi.org/10.1177/1077699019864680

[20] Menna El-Assady and Caterina Moruzzi. 2022. Which Biases and Reasoning Pitfalls Do Explanations Trigger? Decomposing Communication Processes in Human–AI Interaction. *IEEE Computer Graphics and Applications* 42, 06 (nov 2022), 11–23. https://doi.org/10.1109/MCG.2022.3200328

[21] Chris Elsden, Ella Tallyn, and Bettina Nissen. 2020. When Do Design Workshops Work (or Not)?. In *Companion Publication of the 2020 ACM Designing Interactive Systems Conference* (Eindhoven, Netherlands) *(DIS' 20 Companion)*. Association for Computing Machinery, New York, NY, USA, 245–250. https://doi.org/10.1145/3393914.3395856

[22] Ziv Epstein, Mengying Cathy Fang, Antonio Alonso Arechar, and David Gertler Rand. 2023. What label should be applied to content produced by generative AI? https://doi.org/10.31234/osf.io/v4mfz

[23] K. J. Kevin Feng, Nick Ritchie, Pia Blumenthal, Andy Parsons, and Amy X. Zhang. 2023. Examining the Impact of Provenance-Enabled Media on Trust and Accuracy Perceptions. *Proc. ACM Hum.-Comput. Interact.* 7, CSCW2, Article 270 (oct 2023), 42 pages. https://doi.org/10.1145/3610061

[24] Errol Francis II, Catherine Barwulor, Ayana R Monroe, Kediel O Morales, Samya Potlapalli, Kimberly Brown, Julia Jose, Emily Sidnam-Mauch, Susan E Mcgregor, and Kelly Caine. 2024. Usable News Authentication: How the Presentation and Location of Cryptographic Information Impacts the Usability of Provenance Information and Perceptions of News Articles. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 830, 20 pages. https://doi.org/10.1145/3613904.3642331

[25] Michael Glanzberg. 2023. Truth. In *The Stanford Encyclopedia of Philosophy* (Fall 2023 ed.), Edward N. Zalta and Uri Nodelman (Eds.). Metaphysics Research Lab, Stanford University, Stanford.

[26] Kent Grayson and Radan Martinec. 2004. Consumer Perceptions of Iconicity and Indexicality and Their Influence on Assessments of Authentic Market Offerings. *Journal of Consumer Research* 31, 2 (09 2004), 296–312. https://doi.org/10.1086/422109 arXiv:https://academic.oup.com/jcr/article-pdf/31/2/296/17928869/31-2-296.pdf

[27] Brian Guay, Adam J Berinsky, Gordon Pennycook, and David Rand. 2023. How to think about whether misinformation interventions work. *Nature Human Behaviour* 7, 8 (2023), 1231–1233. https://doi.org/10.1038/s41562-023-01667-w

[28] Katrin Hartwig, Tom Biselli, Franziska Schneider, and Christian Reuter. 2024. From Adolescents' Eyes: Assessing an Indicator-Based Intervention to Combat Misinformation on TikTok. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 905, 20 pages. https://doi.org/10.1145/3613904.3642264

[29] Katrin Hartwig, Frederic Doell, and Christian Reuter. 2024. The Landscape of User-centered Misinformation Interventions - A Systematic Literature Review. *ACM Comput. Surv.* 56, 11, Article 292 (jul 2024), 36 pages. https://doi.org/10.1145/3674724

[30] Emma Hoes, Brian Aitken, Jingwen Zhang, Tomasz Gackowski, and Magdalena Wojcieszak. 2024. Prominent misinformation interventions reduce misperceptions but increase scepticism. *Nature Human Behaviour* 8 (2024), 1–9. https://doi.org/10.1038/s41562-024-01884-x

[31] David Hume. 2019. *Of the Standard of Taste*. De Gruyter, Berlin, Boston, 25–40. https://doi.org/10.1515/9783110585575-003

[32] James Hutson and Daniel Plate. 2024. The Algorithm of Fear: Unpacking Prejudice Against AI and the Mistrust of Technology. *Journal of Innovation and Technology* 2024 (Dec. 2024), 17 pages. https://iuojs.intimal.edu.my/index.php/joit/article/view/625

[33] Daniel Kahneman. 2011. *Thinking, Fast and Slow*. Farrar, Straus and Giroux, New York: New York.

[34] Stamatis Karnouskos. 2020. Artificial Intelligence in Digital Media: The Era of Deepfakes. *IEEE Transactions on Technology and Society* 1, 3 (2020), 138–147. https://doi.org/10.1109/TTS.2020.3001312

[35] Reishiro Kawakami and Sukrit Venkatagiri. 2024. The Impact of Generative AI on Artists. In *Proceedings of the 16th Conference on Creativity & Cognition* (Chicago, IL, USA) *(C&C '24)*. Association for Computing Machinery, New York, NY, USA, 79–82. https://doi.org/10.1145/3635636.3664263

[36] Sarah Kenderdine and Andrew Yip. 2018. The proliferation of aura: Facsimiles, authenticity and digital objects. In *The Routledge handbook of museums, media and communication*. Routledge, London, 274–289.

[37] Jan Kirchner and Christian Reuter. 2020. Countering Fake News: A Comparison of Possible Solutions Regarding User Acceptance and Effectiveness. *Proc. ACM Hum.-Comput. Interact.* 4, CSCW2, Article 140 (oct 2020), 27 pages. https://doi.org/10.1145/3415211

[38] Bruno Latour and Adam Lowe. 2011. *The Migration of the Aura, or How to Explore the Original through Its Facsimiles*. University of Chicago Press, Chicago, 275–298. https://doi.org/10.7208/9780226038322-017

[39] Bruno Latour and Adam Lowe. 2011. The Migration of the Aura Exploring the Original Through Its Fac similes. In *Switching Codes Thinking Through Digital Technology in the Humanities and the Arts*, Thoma Bartscherer and Roderick Coover (Eds.). University of Chicago Press, Chicago, 275 – 297. https://sciencespo.hal.science/hal-02057253

[40] Frances Liddell, Ella Tallyn, Evan Morgan, Kar Balan, Martin Disley, Theodore Koterwas, Billy Dixon, Caterina Moruzzi, John Collomosse, and Chris Elsden. 2024. ORAgen: Exploring the Design of Attribution through Media Tokenisation. In *Companion Publication of the 2024 ACM Designing Interactive Systems Conference* (IT University of Copenhagen, Denmark) *(DIS '24 Companion)*. Association for Computing Machinery, New York, NY, USA, 229–233. https://doi.org/10.1145/3656156.3663693

[41] Phoebe Maares, Sandra Banjac, and Folker Hanusch. 2021. The labour of visual authenticity on social media: Exploring producers' and audiences' perceptions on Instagram. *Poetics* 84 (2021), 101502. https://doi.org/10.1016/j.poetic.2020.101502

[42] Lev Manovich. 2002. *The language of new media*. MIT Press, Cambridge, MA.

[43] Jon McCormack, Toby Gifford, and Patrick Hutchings. 2019. Autonomy, Authenticity, Authorship and Intention in Computer Generated Art. In *Computational Intelligence in Music, Sound, Art and Design*, Anikó Ekárt, Antonios Liapis, and María Luz Castro Pena (Eds.). Springer International Publishing, Cham, 35–50. https://doi.org/10.1007/978-3-030-16667-0_3

[44] Caterina Moruzzi. 2019. An Ontological Justification for Contextual Authenticity. *The British Journal of Aesthetics* 59, 4 (07 2019), 413–427. https://doi.org/10.1093/

aesthj/ayz020

[45] Caterina Moruzzi and Solange Margarido. 2024. A User-centered Framework for Human-AI Co-creativity. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI EA '24)*. Association for Computing Machinery, New York, NY, USA, Article 25, 9 pages. https://doi.org/10.1145/3613905.3650929

[46] Muhammad Naeem, Wilson Ozuem, Kerry Howell, and Silvia Ranfagni. 2023. A Step-by-Step Process of Thematic Analysis to Develop a Conceptual Model in Qualitative Research. *International Journal of Qualitative Methods* 22 (2023), 16094069231205789. https://doi.org/10.1177/16094069231205789 arXiv:https://doi.org/10.1177/16094069231205789

[47] George E. Newman and Rosanna K. Smith. 2016. Kinds of Authenticity. *Philosophy Compass* 11, 10 (2016), 609–618. https://doi.org/10.1111/phc3.12343 arXiv:https://compass.onlinelibrary.wiley.com/doi/pdf/10.1111/phc3.12343

[48] Jonas Oppenlaender. 2022. The Creativity of Text-to-Image Generation. In *Proceedings of the 25th International Academic Mindtrek Conference (Academic Mindtrek 2022)*. ACM, Tampere, 192–202. https://doi.org/10.1145/3569219.3569352

[49] Srishti Palani and Gonzalo Ramos. 2024. Evolving Roles and Workflows of Creative Practitioners in the Age of Generative AI. In *Proceedings of the 16th Conference on Creativity & Cognition* (Chicago, IL, USA) *(C&C '24)*. Association for Computing Machinery, New York, NY, USA, 170–184. https://doi.org/10.1145/3635636.3656190

[50] Jeba Rezwana and Mary Lou Maher. 2023. User Perspectives on Ethical Challenges in Human-AI Co-Creativity: A Design Fiction Study. , 13 pages. https://doi.org/10.1145/3591196.3593364

[51] Téo Sanchez. 2023. Examining the Text-to-Image Community of Practice: Why and How do People Prompt Generative AIs?. In *Proceedings of the 15th Conference on Creativity and Cognition* (Virtual Event, USA) *(C&C '23)*. Association for Computing Machinery, New York, NY, USA, 43–61. https://doi.org/10.1145/3591196.3593051

[52] Imani N. Sherman, Jack W. Stokes, and Elissa M. Redmiles. 2021. Designing Media Provenance Indicators to Combat Fake Media. In *Proceedings of the 24th International Symposium on Research in Attacks, Intrusions and Defenses* (San Sebastian, Spain) *(RAID '21)*. Association for Computing Machinery, New York, NY, USA, 324–339. https://doi.org/10.1145/3471621.3471860

[53] Ella Tallyn, Larissa Pschetz, Rory Gianni, Chris Speed, and Chris Elsden. 2018. Exploring Machine Autonomy and Provenance Data in Coffee Consumption: A Field Study of Bitbarista. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW, Article 170 (nov 2018), 25 pages. https://doi.org/10.1145/3274439

[54] Amos Tversky, Daniel Kahneman, and Paul Slovic. 1982. *Judgment under uncertainty: Heuristics and biases*. Cambridge University Press, Cambridge.

[55] Alina Valyaeva. 2023. People Are Creating an Average of 34 Million Images Per Day. Statistics for 2024. https://journal.everypixel.com/ai-image-statistics. Retrieved September 1, 2024.

[56] Tomas Vancisin, Loraine Clarke, Mary Orr, and Uta Hinrichs. 2023. Provenance visualization: Tracing people, processes, and practices through a data-driven approach to provenance. *Digital Scholarship in the Humanities* 38, 3 (04 2023), 1322–1339. https://doi.org/10.1093/llc/fqad020 arXiv:https://academic.oup.com/dsh/article-pdf/38/3/1322/51309605/fqad020.pdf

[57] Sheena Vasani. 2024. *Meta is incorrectly marking real photos as 'Made by AI'*. The Verge. Retrieved September 11, 2024 from https://www.theverge.com/2024/6/24/24184795/meta-instagram-incorrect-made-by-ai-photo-labels

[58] Emily K. Vraga. 2011. Dealing with dissonance: Responding to an incongruent test result in a new media environment. *Cyberpsychology, Behavior, and Social Networking* 14, 11 (2011), 689–694.

[59] Wired. 2024. The Blurred Reality of AI's 'Human-Washing. https://www.wired.com/story/gadget-lab-podcast-651/. Retrieved September 4, 2024.

[60] Chloe Wittenberg, Ziv Epstein, Adam J. Berinsky, and David G. Rand. 2024. Labeling AI-Generated Content: Promises, Perils, and Future Directions. https://doi.org/10.21428/e4baedd9.0319e3a6,

[61] Liudmila Zavolokina, Kilian Sprenkamp, Zoya Katashinskaya, Daniel Gordon Jones, and Gerhard Schwabe. 2024. Think Fast, Think Slow, Think Critical: Designing an Automated Propaganda Detection Tool. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 491, 24 pages. https://doi.org/10.1145/3613904.3642805