ERC Advanced Grant 2023
Part B2
Deep Culture – Living with Difference in the Age of Deep Learning

Section a. State-of-the-art and objectives

Introduction: The Challenge of Deep Learning

A new era in the relation between humans and machines dawned in the latter half of 2022 with the public emergence of consumer-centred artificial intelligence (AI), which has dazzled observers with the ability to generate text (ChatGPT/GPT-3), images (DALL-E/Stable Diffusion) or music (MusicLM) (AI Demos, 2023). Companies offer completely new AI services and showcase how far AI has come for cultural production, circulation and consumption. Users can now develop and share 'prompting' strategies to create digital art, music and writing, translating 'words' into 'tokens' that AI can operationalise (Oppenlaender, Linder and Silvennoinen, 2023). We have all come to realise that digital culture has fundamentally changed, brought about not by a general AI but by deep learning, a specific subset of machine learning that has defined AI developments in the past decade. The media has dubbed it a 'new era of machine learning' (Waters, 2023), where 'old' giants like Google can quickly lose billions of dollars if they are seen to fall behind in the latest AI arms race (Milmo, 2023).

Until recently, AI advancements have not been so directly interacted with and hidden in very large systems for Internet search or facial recognition. Many of these systems have proven to be highly controversial. Facial recognition has become a stand-in for the global expansion of surveillance and the end of privacy (Roussi, 2020), while search engines have struggled to keep, for instance, Holocaust-denying sites from their top rankings (Makhortykh, Urman and Ulloa, 2021). In 2022, it did not take long for AI image-generators like Stable Diffusion to replicate racist and sexist stereotypes, which they learned from extracting Internet material (Rose, 2022). Language models, which produce text out of prompts, easily reproduce Holocaust denial (Gordon, 2022) or allow chatting with Hitler (Ingram, 2023). ChatGPT seems to pass the 'Nazi test' (Kantrowitz, 2022), as it has human-made guardrails against Holocaust denial, producing curated answers for anything its creators see as ethico-politically complicated. Prompting for style modifications, one can, however, still produce horrific accounts of the history of the Holocaust, bordering on denial and relativisation (Gaumond and Wittes, 2023). Between prompting and curation for the worst effects of deep learning, our new relationship with these machines veers between imaginaries of AI magic and destruction.

Yet despite such strong imaginaries and talks about 'sparks' of Artificial General Intelligence (Bubeck *et al.*, 2023), much is missing from this kind of machine learning that would make it intelligent – including classical goals of AI research like the capability to think differently from what is learned or a clear ability to fully address questions of causality (Jordan, 2019). However, this has not stopped the global expansion of deep learning into cultural consumption, circulation and production, driven by platform and surveillance capitalism (Srnicek, 2017; Van Dijck, Poell and Waal, 2018; Zuboff, 2019). Deep learning is perfectly suited for multi-billion-dollar Big Tech companies with their focus on one-size-fits-all solutions to improve search or recommendations, to be applied to millions of users for very large profits. Every day, new cultural meanings, relations and objects are formed in exchanges with commercial AI algorithms built with deep learning. These algorithms write essays for us, organise our timelines on social media and provide suggestions on Netflix, but are also used by governments to detect crime or create risk profiles to save money on social security. We are not necessarily witnessing a crisis, but rather a protracted situation where our lives and cultures have become entangled with developments in deep learning. How these latest AI algorithms reason and why they do what they do is opaque and unclear, when not secret (Knight, 2017). They are generally not made from specific pre-planned designs but are fed large amounts of data to learn patterns that fit their one-size-fits-all applications (Arora, 2020). The danger is that a different kind of culture that addresses the heterogeneity and diversity of cultural expressions is lost together with humanistic values and methodologies.

Ground-breaking Contribution and Objectives

The DEEP CULTURE project offers a novel approach to what we term emergent 'deep culture' in order to conceptualise the wide-ranging transformations of culture brought about by deep learning. It claims that the relationship between culture and deep learning can be productive if humanistic values are included, which address the multiplicities and complexities of cultures. To investigate this claim, the project advances deep culture **as a conceptual and methodological framework**; firstly, to understand emerging deep culture and, secondly to develop new productive ways of interacting with it. It is the first project to develop deep culture

as an **object of study as well as a method and capacity to attend to this new historical moment and go beyond it**. We aim to create a pathway from what we call the current **deep culture of uniformity** that is produced with universal data extraction and technical solutionism to new **deep cultures of difference** shaped by humanistic data, concepts and approaches. The project will explore how, starting from ideas and values of difference such as detail, relationality and imagination, we can remake deep culture away from producing sameness and regularity towards considering the heterogeneity and complexity of culture and living with difference (Hall, 2021). It investigates how ideas like the situatedness of knowledge or rediscovering lost voices can be employed for new productive relations with deep learning. Charting and transforming deep culture with critical analysis and practice-led methodologies is a ground-breaking contribution that leads to new ways of interacting with and researching digital cultural material (Figure 1).

Figure 1: Transforming Deep Culture
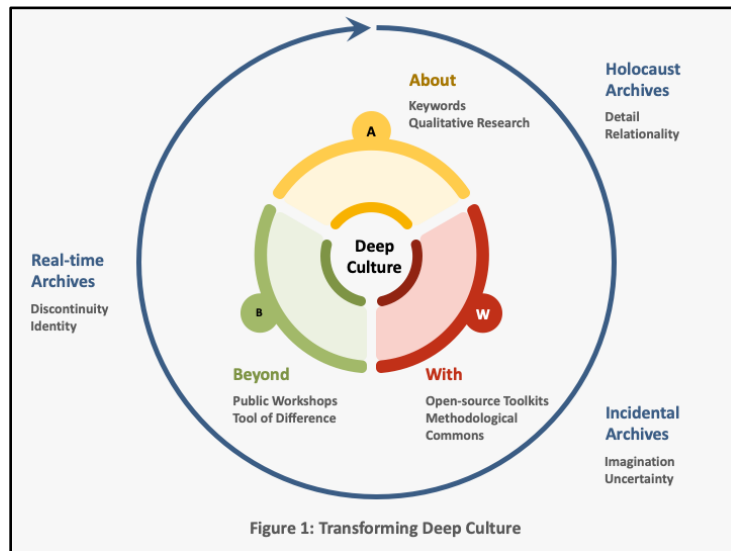
DEEP CULTURE builds on the tradition of digital humanities, which have already begun to work with deep learning methods in order to understand, transform and critique the production, circulation and consumption of digital culture. This undertaking has not been without risks, as some have argued that computational methods more generally have subordinated humanities research to computing logics, positivism and scientism (Liu, 2012; Kirschenbaum, 2014). The risk is to get lost in ad-hoc experiments that put us in awe at the technologies' potential without a critical, systematic investigation (Prescott, 2011). DEEP CULTURE goes beyond these limits and shows how a critique of affordances of deep learning as it is currently used for cultural production and consumption can be the starting point for new methodologies and relations with deep culture.

To research and transform deep culture, the project adopts a three-dimensional understanding of culture. Cultural studies theorist Raymond Williams highlighted the complexities of culture in his seminal two-pronged articulation of culture as a 'whole way of life' through shared meanings and practices as well as arts and 'creative practices' (Williams, 1958, p. 3). Deep learning already enacts and is part of 'ordinary', everyday culture (Cultural AI Lab, 2021; Born, 2022). As a third dimension, a material culture of AI should be added, 'what goes on inside the technology' (Born, 2022). DEEP CULTURE attends to all three dimensions of culture as shared meanings, creative processes and materiality – how these can be traced and transformed. It advances three shifts in the relations between deep learning and all these dimensions of culture through three main objectives at the centre of Figure 1:

i) The first objective is **a critical inquiry about deep culture** starting from keywords and epistemic translations. The dominant deep learning epistemology and practices are translated to understand how they might render, ignore or speak to humanistic ideas and values.

ii) The second objective is **methodological, building critical inquiry with deep culture** in order to develop new data and methods commons informed by humanistic ideas and values. Deep culture currently binds us into the cycle of ever larger data extraction and solutionism. In new deep cultures of difference, data is processed in a culturally sensitive way, the methods target humanistic values, and we become active enactors.

iii) Thirdly, the project works towards **new productive, critical relationships with and beyond deep culture** with diverse publics, breaking down the boundaries of experts and non-experts, data haves and have-nots, leading to a more evenly circulated AI. This **ethico-political** objective will focus on new types of human-machine agencies and relations.

With these objectives, the project is set apart from a non-grounded and exceptionalist approach to technology innovations that is currently found in the discourse of the race to 'god-like AI' through deep learning (Hogarth, 2023) and an overly strong rhetoric of computational modelling in the humanities that ignores the limits of all computational capacities when it comes to researching culture (Bode, 2017; Da, 2019). Each objective also defines a step in **DEEP CULTURE's pathway (Figure 1) from the deep culture of uniformity to new deep cultures of difference**, from mapping the cultural enactments of deep learning to reshaping its data and methods and co-producing new public cultures and human-machine agencies.

The project will trace this pathway towards new deep cultures of difference through experiments at **three archival sites (Figure 1): historical archives (Holocaust archives), real-time archives (web**

archives) and 'incidental archives' (archives that are a by-product of other cultural/social interactions). This selection builds on the PI's experience in digital humanities research with archives in transformative, large-scale international initiatives like the European Holocaust Research Infrastructure (Blanke and Kristel, 2013) and in analysing web archive materials (Brügger and Finnemann, 2013) and other online sources, often hidden away in the repositories of the deep web (Noordegraaf *et al.*, 2021; Valdivia *et al.*, 2022). While the PI's work has up to now focussed mainly on developing new digital research with archives, DEEP CULTURE will also query their position in current relations of deep learning with culture and how this can change. Archives have been key sites for critical research on culture as well as for innovations on methodologies of historical-cultural 'archival research' that have gained traction beyond the humanities (Moore *et al.*, 2016; Jo and Gebru, 2020). They preserve and develop humanities traditions of critical work with sources, which counter the problems of seemingly abundant cultural data and online misinformation and misrepresentation (Blanke, 2020). Maybe most importantly, archives, both old and new as well as official and informal, are the crucial body of knowledge current deep learning uses to encode 'the full extent of human knowledge (and then some)' in ever larger language and image models (Welsh, 2022, p. 35). While archives have been a key part of humanities and cultural practices, little is known how deep learning produces and circulates deep culture from cultural collections. Expanding this knowledge will allow us to make archives the key site for reimagining cultural meanings and materialities of deep culture as well as for developing historical-cultural methods and ethico-political sensibilities. Through its conceptual and methodological innovations, DEEP CULTURE will reclaim these sites of crucial cultural collections (and others) from culturally insensitive deep learning practices through **six in-depth case studies that go beyond the current state-of-the-art by integrating humanistic critique and advanced digital methodologies**.

State-of-the-art: Interdisciplinary Research on AI
The last decades have been incredibly productive for the study of digital culture and the changing ways of communicating and engaging with digital media (Gere, 2009; Hjorth, 2018; Beer, 2019). The large-scale digitisation of cultural collections and fast growth of born-digital materials in archives and elsewhere have led to the global expansion of digital humanities (Schreibman, Siemens and Unsworth, 2008; O'Sullivan, 2022). Digital humanities have contributed to research on digital cultures their own research agenda on novel human-machine interactions. They innovated new 'collaborative' places and infrastructures (Svensson, 2013; Edmond, 2015; Kristel, Blanke and Romary, 2015), whereas 'distant reading' (Moretti, 2013; Underwood, 2017) and 'algorithmic reading' (Esposito, 2022) are widely cited beyond digital humanities as new ways of thinking about the relations between human and machine interpreters of texts. However, in the ubiquity of digital cultures and its global manifestations, one of the most important changes, if not the most important one since the re-emergence of AI from its so-called last funding 'winter', has been until very recently almost completely missed and is still met mainly with surprise and a large amount of defensiveness. Deep learning is not just an incremental step in AI. It has taken over much of the academic and industry interests – also replacing the former with the latter (Whittaker, 2021). As *Science* magazine laments, with deep learning, '[i]ndustry is gaining control over the technology's future' (Ahmed, Wahed and Thompson, 2023).

Deep learning builds on existing machine-learning achievements, but it is also a paradigm shift (LeCun, Bengio and Hinton, 2015; Arora, 2020). Randomly initialised, its systems search for and find their own representations of the input data they are given (Chollet, 2017). They are end-to-end systems, that have unified machine-learning frameworks into a common paradigm, working equally with texts and images (Cornia *et al.*, 2020; van Noord, 2022) as with structured data (Ryan, 2020). By adding layers of alternating linear and non-linear processing units and harmonizing divergent programming practices into block-based architectures, deep learning has achieved unprecedented levels of performance in several areas that until very recently seemed not amenable to computational reasoning. These new digital capacities are highly relevant to digital culture and media, beginning with images (LeCun and Bengio, 1998; Krizhevsky, Sutskever and Hinton, 2017) and later for texts (Hochreiter and Schmidhuber, 1997; Vaswani *et al.*, 2017). Even minor changes to deep learning have great impact. The worldwide amazement after the release of ChatGPT from the language model GPT3 (Floridi and Chiriatti, 2020) and subsequent anxiety about the effects it has on cultural production contrasts with the fact that GPT3 was based on a largely unchanged previous version (Brown *et al.*, 2020). GPT4 has also been primarily focussed on including more (cultural) data to empower an even larger network of parameters as well as to improve 'model alignment' with user input and to reduce the probabilities of undesirable outputs (Liu *et al.*, 2023).

The research that has driven AI in computing has been accompanied by a fast-growing critical literature, which has analysed the effects of the new technologies on society, economics and politics. They have rendered their transformations through many diagnoses such as 'algorithmic culture' (Striphas, 2015), 'algorithmic reason' (Aradau and Blanke, 2022), 'black box society' (Pasquale, 2015), 'dataveillance' (Van Dijck, 2014), 'platform capitalism' (Srnicek, 2017) or 'algorithmic violence' (Bellanova *et al.*, 2021).

Critical 'algorithm studies' (Seaver, 2017) and 'data studies' (Iliadis and Russo, 2016; Hepp, Jarke and Kramp, 2022) have been especially productive as new fields to examine digital culture and society. However, global analyses of data and algorithms risk confusing types of algorithms and machine-learning architectures, merging them all together. The historical moment of deep learning is either missed or appears unintelligible, 'an occult power that cannot be studied' (Pasquinelli and Joler, 2021, p. 1265). Making all algorithms and data the same, we quickly arrive at dystopian and catastrophic views. Epistemically and methodologically, only the concerns about a 'new positivism' (Kitchin, 2014) and 'inductivism' (Leonelli, 2020) dominate, while an 'all-dominant' platformisation is seen to take over cultural production (Nieborg and Poell, 2018). Perceived as applying to everything equally, the 'rules' of algorithms are seen as biased towards what can be quantified and appears frequently in data. Moreover, understood as 'thin' rules (Daston, 2022), machine-learning algorithms seem to be the antinomy of 'thick' cultural meanings, relations and contexts. This project recognises the specificity of this historical moment and what Science and Technology Studies (STS) scholar Helga Nowotny has called the beginning of the 'co-evolution of humans and their digital machines' (Nowotny, 2021). It asks how we can begin to shape it from the humanities, what it means for us as humans and especially for our understanding of culture.

Deep learning has fast-tracked this co-evolution because it has fundamentally reorganised programming principles and design workflows, leading to the complete industrialisation of machine learning. Through deep learning, programming becomes end-to-end non-parametric modelling on an industrial scale, moving away from AI's beginnings in dedicated research labs and machine learning as part of a wider data science workflow. This means that cultural production through deep learning is just one example of similarly industrialised workflows, starting generally not with a detailed design but with data and its strategic 'appropriation'. Commercial deep culture is intrinsically linked to 'surveillance capitalism' (Zuboff, 2019) and the global extraction of bigger and bigger data (Couldry and Mejias, 2019; Anwar and Graham, 2022), making the development of deep culture applications the same as creating shoe-recommendation algorithms and generally ignoring humanistic values. Produced through industrialised deep learning, current deep culture tends to identify cultural shared meanings with identities and similarities, creative processes with large-scale patterns and democratisation with trust in algorithmic solutionism. Against these tendencies, the project aims to recover alternative potentials for humanistic thought about, with and beyond deep culture.

To address the challenges of the global expansion of deep learning, new long-term interdisciplinary collaborations have been promised both in science and industry, which are supposed to reignite AI as the 'study of intelligence' including humanities (Rich, 1985). The deep learning powerhouse DeepMind has as its mission to 'solve intelligence to advance science and benefit humanity' (DeepMind, 2023). However, social sciences and humanities research is still largely ignored by those who drive the industrial development of deep learning. This also applies to areas that are close to social and cultural research, when, for instance, AI-based decisions are supposed to be explained to humans (Miller, 2019) or the collection of training data could be improved with archival research on data provenance and accountability (Colavizza *et al.*, 2021). While some critical computer science approaches attempt to change this (Jo and Gebru, 2020; Graziani *et al.*, 2023), claims to larger interdisciplinary research on AI made by industry and academia are exaggerated. Particularly research that takes seriously the different traditions and objectives of social and cultural analysis and what they contribute to work on deep learning, is completely missing. There are no projects that bring together the computing and critical side of deep learning. If deep learning is engaged in the humanities, it is either as an object of critical study (Floridi and Chiriatti, 2020; McQuillan, 2022; Weatherby and Justie, 2022) or as a methodology (Colavizza *et al.*, 2021; Suissa, Elmalech and Zhitomirsky-Geffet, 2022). Yet, these computational and critical dimensions have remained disparate. DEEP CULTURE changes this by bringing together digital humanities with critical studies of AI.

To develop new practices of deep learning from its critique and vice versa, the project sets off from the rich critical literature on machine learning in general and further small inroads into deep learning. Cultural analysis (Mackenzie, 2017) has demonstrated how to explore the inner workings of machine learning (though not yet deep learning), while the history of science has traced how AI is now about large-scale systems and experimentation rather than committed to simulating human reasoning (Dick, 2019). Many more fields such as the emerging political economy of AI (Srnicek, 2018; Prainsack, 2020; Luitse and Denkena, 2021) or the ever-faster expanding AI ethics (Taddeo and Floridi, 2018; Hagerty and Rubinov, 2019; Jobin, Ienca and Vayena, 2019; Nat Mach Intell, 2022) will help us define a new critical analysis of deep learning's current epistemologies and practices by combining qualitative research with a focus on enacted values (Law and Mol, 2002; Jensen and Gad, 2009). Safiya Noble (2018) and Ruha Benjamin (2019) have offered exemplary analyses combining contemporary philosophy with qualitative investigations into how AI technologies reproduce and intensify racism. For Weatherby and Justie (2022), a critical epistemic analysis of neural networks reveals a reductionist 'social faith', while Jaton (2021a, 2021b) highlights the morality of machine learning from moments of 'hesitation' during its design when genuine choices are made

to engage 'different possible futures' (see also Amoore, 2019). All these new studies have the advantage that they focus on how its various technological elements are actually working together in the production of knowledge and values. They work on the moments in machine learning that show that things 'could be otherwise' (Jaton, 2021a), which the project will especially target. Valuations diverge in practice (Heuts and Mol, 2013), because the actions that sustain deep learning differ. The project will begin from practices, relations and how things emerge in contingent fragile assemblages, as different histories are materialised in them (Law and Mol, 2002). We want to avoid, however, that critical work is split into a meta-normative part and a social research part (Osborne, 2013). In DEEP CULTURE, they feed together further practice-led research and new critical methods.

Whereas critical studies of AI still largely ignore alternative affordances of and practices with the new AI technologies, digital humanities interventions generally do not work with insights from the critical studies of AI. They are concentrated on applying the new technologies across cultural fields. There are notable exceptions that have shown that digital humanities interventions can be sensitive to critical work on digital cultures (Risam, 2018; D'Ignazio and Klein, 2020; O'Sullivan, 2022). However, these calls to engage with wider critical studies are at the moment ignored where digital humanities deal with deep learning, which are fixed to ad-hoc and domain-bound methods (Suissa, Elmalech and Zhitomirsky-Geffet, 2022). One of the more recent success stories of applied deep learning in the digital humanities is the Transkribus tool for hand-written text recognition (Muehlberger *et al.*, 2019), which has revolutionised digitisation efforts. In *Nature*, DeepMind (Assael *et al.*, 2022) presents the Ithaca system to make Greek inscriptions legible. DEEP CULTURE wants to show how the inquiry with deep culture can be more than often ad-hoc individualised success stories using yet another technology to work with large cultural data.

Working with deep learning methods and taking its critique seriously, we can address long-term challenges in digital humanities. For instance, the promise of small cultural data is to allow for better answers to specific questions (Kitchin and Lauriault, 2015) and to uncover 'tiny clues' (Lindstrom, 2017). Cultural expression and relations are often sparse and small and not recorded in systematic fashion (Schöch, 2013). Deep learning works well with small data under certain conditions (Howard and Gugger, 2020), but we have only begun to explore these conditions and what might change for cultural data analysis. Deep learning perfects distributed digital memory practices in the weights of its networks through its 'encoder-decoder' and 'transformer' architectures (Weatherby and Justie, 2022). 'Transfer learning', using the encoded information in other domains, is a breakthrough of deep learning and a very active research field in digital humanities (Banar, Daelemans and Kestemont, 2020; Inbasekaran, Gnanasekaran and Marciano, 2021; Suissa, Elmalech and Zhitomirsky-Geffet, 2022). It is currently, however, used to develop applications in the wider cultural domain and not to advance the project's objective of critical inquiry with deep learning. The critique of the current practices of deep culture is not taken as a starting point of new deep cultures.

Deep learning can analytically discriminate as well as creatively generate (Goodfellow *et al.*, 2014; Dhariwal and Nichol, 2021). As outlined in the Introduction, 'generative AI' has been one of the biggest break-through technology stories of the last year with billion-dollar investments (Criddle and Bradshaw, 2022), setting up AI for the first time as a global cultural consumption experience with tools like ChatGPT or apps like Midjourney for AI-generated art. For the project, consumer AI raises vital questions about how deep culture seemingly easily conjoins what until now has been considered fundamentally different media, from video to text. Deep learning's strengths are multimodal, being able to create multimedia from text and vice versa. Generative AI has been picked up in the field of 'generative digital humanities' (Offert and Bell, 2020). With deep learning, machines develop their own cultural imaginations, which can be productively employed. They can imagine different cultural expressions and realities to overcome selection and survival bias in archives (Kim, 2022), or they can synthesise more data for the many periods of human history where we have none (Assael *et al.*, 2022). Generative digital humanities and transfer learning for cultural data are just two examples of creative methodological experimentations in digital humanities the project will address.

To develop possibilities for different deep cultures, this project works through archives as its sites of empirical, methodological and political investigations. It will counteract deep culture's 'wholesale appropriation of existing culture' (Bridle, 2023) by redefining the relationship between cultural collections and deep learning. The creators and curators of cultural materials have been ignored by Big Tech, but have recently begun to fight back (Arkin, 2023). Cultural data has often been consumed and (re-)created without much long-term thinking about consequences, which has led to the famous examples of societally harmful datasets deep learning has been built upon (Paullada *et al.*, 2021) like seminal studies for facial images (Buolamwini and Gebru, 2018) or text (Bolukbasi *et al.*, 2016). In this new research field addressing the harms from data and algorithms, (digital) humanities play an increasingly active role (Wachter, Mittelstadt and Floridi, 2017; Morley *et al.*, 2020; Berry, 2022; Prescott, 2022). A few computer scientists have already recognised the potential of (digital) humanities and how the big collection of machine-learning data should learn from a critical tradition of archival research (Jo and Gebru, 2020; Scheuerman, Hanna and Denton,

2021). 'Datasheets', for instance, have been suggested for documenting data creation and transformation (Gebru *et al.*, 2021). DEEP CULTURE re-records and recodes cultural data for new culturally sensitive deep learning across its three archival sites as an important pillar of its work on critical methods.

Only with a holistic project addressing deep culture practices and methodologies together, can we tackle the project's final objective of new human-machine relations and countering monopolies on deep culture productions in public and creative spheres. A top-down applied ethics is no countermeasure against corporate AI power and asymmetries, as has been shown by critics of ethics-washing (Bietti, 2020) and of global AI imbalances furthering North-South divides (Jobin, Ienca and Vayena, 2019). AI ethics can also not be reduced to computational questions of different optimizations and fairness as a new metrics, and it requires a plurality of ideas (Hagerty and Rubinov, 2019; Raji, Scheuerman and Amironesei, 2021). Humanities research is necessary for deep learning, because 'qualitative decisions are made about what metrics to optimise for, which categories to use, how to define their bounds, who applies the labels' (Bartolo and Thomas, 2022). Tracing these decisions is hard and requires the new methodologies this project will develop, as deep learning systems are not designed directly but are created with permanent experimentations. Deep nets are initialised randomly and then learn their own parameters from data (Arora, 2020). As the PI has argued, although there is now a vast literature on AI ethics, it generally speaks to experts like engineers, designers or planners or criticises ethics as simply a cynical strategy by Big Tech companies (Aradau and Blanke, 2022). DEEP CULTURE proposes a novel approach to the problems AI ethics speaks to, which is not separate from, but entwined with political questions about publics and contestation.

Bonnie Honig links democratic, ethico-political actions to fragile, contested public things, which 'bind citizens within the complicated affective circuitries of democratic life' (Honig, 2017, p. 49; Aradau and Blanke, 2022). They are not just out there but need to be reconstituted all the time through action in concert. DEEP CULTURE asks whether we can make some of deep learning's underlying principles like encoding and decoding publicly contestable by developing the project's toolkits into public things. Humans are co-working and co-researching with deep learning things all the time. The project makes permanent involuntary crowdsourcing visible and algorithmic decisions contestable by everyone, and not just citizens with the right 'skills and free time' (Birchall, 2021, p. 50). It offers otherwise passive 'crowds' strategies to challenge deep learning and the dominance of commercial deep culture, where Big Tech leverages their own algorithmic intransparencies 'to undermine users' confidence in what they know about algorithms and destabilise credible criticism' (Cotter, 2021, p. 1). As the PI has shown in earlier work, hackathons, e.g., can be a part of this strategy. They can be 'inverted' from practices of Silicon Valley to efficiently produce more technology and can become means to collectively explore the materiality and fragility of technologies (Pybus, Coté and Blanke, 2015). Such digital humanities approaches facilitate new digital action in concert (Berry *et al.*, 2015; Lodato and DiSalvo, 2016; Svensson, 2016) and contribute to a new ethico-politics that changes how dominant deep learning 'things mediate publics' (Marres, 2016, p. 23). An engagement of digital humanities practices with AI ethics avoids falling back onto relations of human control and sovereignty, mirroring technological imaginaries of mastery, which often inform engagements with AI ethics and politics.

Dominating research and industry, science and society, deep learning accelerates the 'convergence' (Jin, 2021) of AI, digital platforms and digital culture into a deep culture that is driven by a desire to automate shared meanings and creative processes further than anything we have seen before. Deep learning is so valuable to industry because it does not require as much expensive prior human domain expertise. It learns to find edges in images and does not care that for centuries the basic token of many natural languages has been the word, which is transformed into numbers that neural nets can run calculations on. In fact, deep learning gains much of its strengths by breaking with research traditions the humanities (and other sciences) are built on. Domain-dependent pre-processing of collections, requiring knowledge about culture, is replaced by trust into systems that seem to be able to do it all. The danger is that in the flurry of new data, algorithms and infrastructures, core human(-istic) insights and interests in culture are marginalised, together with traditions of responsible recording of cultural practices. The chance is that we can remake deep culture and learn from its knowledge productions to develop new interactions with our digital cultural worlds.

Section b. Conceptual and Methodological Innovation

Breaking out of Deep Culture of Uniformity
While the effects of deep learning on culture and its records have been very much in the public eye since the emergence of consumer-facing AI with large-scale appropriation of culture, they are under-researched in the itself nascent research on AI's impact on culture (Roberge and Castelle, 2021; Born, 2022; NeurIPS, 2022). Building on the PI's previous work on critical datafication (Blanke *et al.*, 2014) and 'algorithmic reason' (Aradau and Blanke, 2022), DEEP CULTURE addresses the new challenges of deep culture through an interdisciplinary framework emerging across **agonistic encounters** (Barry, Born and Weszkalnys, 2008)

between digital humanities, cultural studies and computer science. The project's main claim is that deep culture is more than a potential threat to cultural productions and can be remade by harnessing interdisciplinary research for new digital humanities. DEEP CULTURE is the first project to define and analyse deep culture and develop alternative meanings, creative processes and materialities. Given the global ascendency of deep learning, the project is urgent.

As discussed above, digital humanities have started to include critical humanities work like data feminism (D'Ignazio and Klein, 2020), post-colonialism (Risam, 2019) and expanded to smaller data and practices like thinking through visualisation or data narrations (Rezai, 2022). These interventions have been fostered by a critique of digital humanities that concentrate on large-scale, universalist projects and relegate 'difference to its margins' (Risam, 2015). They are signs of the fundamental transformation of digital humanities, which does not only take computational transformations as inputs and as objects of analysis but reconfigures them as part of scholarly, ethical and political interventions. Building on these approaches and ideas, **DEEP CULTURE re-centres digital humanities around the idea of 'difference'**. We want to move away from a deep culture of uniformity to deep cultures of difference where, in the words of Stuart Hall, 'differences refuse to disappear' and 'homogeneity cannot be assumed' (Hall, 2021, p. 411). In light of critical work addressing formations of identity and difference in culture and society, it is surprising that theories and practices of 'computational modelling and digitisation in the cultural sphere' have up to now paid little attention to **how differences, marginal temporalities and ambiguities could be approached computationally**. Difference is a core principle of a pluralistic humanistic **epistemology** (Derrida, 1982; Haraway, 1991) that starts from a tendency to specify, aims to situate knowledge and is interested in rediscovering lost voices in privileged knowledge (Stoler, 2002). **Ontologically**, difference points to the multiplicities and contingency of human cultures and how another everyday is possible (Williams, 1976; Risam, 2015; Hall, 2021), while **methodologically** it implies a reorientation to focus on the noise that remains in current systems of identifications (Chang and DeDeo, 2020), 'exposing complexities and ambiguities' (Dobson, 2020). The more hyped and futuristic the digital capacity appears, the further away we seem from humanistic ideals of 'difference'. There is a widely shared assumption that difference is absent in deep learning. Yet, it can be rearticulated in many modulations in deep learning and its data representations, which the project will develop for new deep cultures where differences continue to matter.

DEEP CULTURE investigates deep learning's interactions with culture and the possibilities and limitations of (re-)casting deep cultures through empirical explorations at three distinct archival sites: historical, real-time and 'incidental' archives. These archival sites are often at the borderlines of what has traditionally been seen as an archive – oral histories of ignored victims or irrelevant governmental documents, which would have been thrown away. Yet, as they are online today, many have enabled the easy access and 'cheap data' revolution to train large-scale deep learning systems (Halevy, Norvig and Pereira, 2009). Archives (and special collections) have long been the 'labs' of (digital) humanities (Manoff, 2004) but have also become a key part of the big data (Marciano *et al.*, 2018) that deep learning needs. These project's archival sites have been selected for their importance in existing deep culture and their diverse temporalities and global localities. They all stem from work that crosses multiple scientific disciplines and industries, and they are transnational. They integrate multi-lingual, multi-modal collections (text, images and videos), and are often cited as emerging 'infinite archives' (Goldstein, 2004). Thus, they form the perfect 'laboratory' for high-risk, high-gain studies on deep culture.

The first **archival site of transnational historical Holocaust** is special for many reasons. Owing to memory traditions of the victims as well as large-scale political support, it has been at the forefront of historical 'big data' in the humanities (Hand, 2011). The beginnings of deep learning applications in the field have quickly received broad attention across media and academia (BBC, 2022). Given the highly dispersed nature of Holocaust violence, the collections have also been widely distributed, which means that they benefit from a cross-domain, end-to-end approach that is interested in new types of data contextualisation. Holocaust studies cannot be apolitical, and its digital archives were among the first to experience the dangers of online misrepresentations. Holocaust deniers try to use online materials for their perverted version of history (Allington, 2017; Walden, 2022). Because of online misrepresentations, data-extractivist deep learning still easily misses the 'Nazi test' (Metz, 2018). ChatGPT struggled at the beginning with negations and produced Holocaust denial out of prompts to negate the 'truth' of the Holocaust (Miguel, 2023). At the same time, projects like the European Holocaust Research Infrastructure (EHRI) also show how to enable the localisation and distribution of Holocaust knowledge. The PI has co-led EHRI for over ten years to develop new digital research (Blanke and Kristel, 2013) and provide unique access to Holocaust collections in a rich collections graph (Blanke, Bryant and Speck, 2015). Holocaust collections are both born-digital and digitised, ranging from government records to oral narrations in several multimedia formats.

DEEP CULTURE's **second archival site consists of real-time archives such as web archives or social media sources for training machine learning** – Flickr, Twitter, news records, etc. Web-based

archives and training sources are often easy to access and have become a key focus of critical AI research (Paullada *et al.*, 2021), because they continue to be the single most important site for dominant deep culture data extraction. Although research has early on shown how web materials are plagued by severe capture and label biases (Torralba and Efros, 2011), they are connected, more than any other data, to the accumulation of seemingly cheap and context-free training data (Paullada *et al.*, 2021; Scheuerman, Hanna and Denton, 2021; Ciston, 2023). Organised national and international web archives are at the same time at the centre of the critique of senseless data extraction and attempt to break out of these practices by reintroducing critical archival knowledge of transparency and authenticity. Being real-time, these archives also stand for a new regime of algorithmic mediation of culture (Blanke, 2014). The PI has experimented with various social media based deep learning training collections and worked with web archives both nationally and internationally (RESAW, 2013; Institute of Historical Research, 2017; Westerhof *et al.*, 2021).

The third archival site is about what can be called **'incidental archives', collections that are online as a (minor) supplement to something else. They are mainly secondary effects of other cultural/social interactions**, have often not been designed as archives and would be non-archives following a traditional interpretation where they would be perceived as not having 'enduring value' (Cox and Samuels, 1988). Traditionally, their collections might have been discarded, but many existing deep learning algorithms have also been trained using their online materials. They are a 'strong metaphor for any corpus of selective forgettings and collections' (Stoler, 2002, p. 94), because they are often 'indirect sources' that also speak of those who do not have the power and resources to preserve their records (Kim, 2022). Incidental archives can be seen as the ultimate 'dark (…) archives' (Guldi and Armitage, 2014) and are often hidden away in the deep web, requiring dedicated access and discovery methods. Yet, they have been harvested to train deep learning models. Google Translate, for instance, is rumoured to have employed EU documents during its training (Regner, 2010) though details are kept secret. Government documents lingering in abandoned online repositories are a typical example of incidental archives, as are archives of app codes used for training code language models or online diaries of Covid-19 experiences. But there are many more and of numerous different types. For the incidental archives especially, we need to glean 'cultural practices across multiple locations' (Seaver, 2017, p. 6), rewrite them from their original purpose and 'counter-archive' them (Stoler, 2002), rediscovering lost voices against the privileged knowledge that has often created them.

While situated at these archival sites, the project does not start from a particular collection or technique, as it is common in digital humanities. Its ambition is to advance **a critical inquiry into deep culture and the potential for its transformation, starting from keywords that can foster epistemic translations between deep learning and humanities**. This is inspired by Raymond Williams's work to understand how we talk about culture. Keywords are 'significant, binding words' for activities in culture and indicative of 'certain forms of thought' (Williams, 1976, p. 15). The project focuses on keywords binding together 'ways of seeing culture' (Ibid.) across deep learning and humanistic questions of difference. They are our main 'tools of interdisciplinarity' and our 'heuristic and methodological basis' (Bal, 2009, p. 14). The project innovates through epistemic translations that relate the vocabularies and practices of deep learning to ontological, methodological and epistemological ideas of difference.

Since Williams, keywords have been used in many adaptations for the study of (digital) cultures (Bennett, Grossberg and Morris, 2013; Striphas, 2015; Peters, 2016; AI Now Institute, 2021; Thylstrup and Agostinho, 2021), which demonstrate their interdisciplinary potential. The PI has also shown the importance of interdisciplinary translations through the idea of 'scholarly primitives' and how they relate computing practices to humanities research (Blanke and Hedges, 2013). While these primitives were set, keywords in DEEP CULTURE transcend boundaries if they are not seen to be fixed and do not resolve seemingly opposing propositions. They are open to interpretation and supposed to inspire different research, extending both the boundaries of how deep learning currently enacts a deep culture of uniformity and developing new deep cultures. The approach is productive for the interdisciplinarity of AI precisely when it focuses on terms that escape an agreed-upon definition and require deliberation. In the critical tradition of translation (Benjamin, 2009), deep learning keywords are not just applied to new contexts, but their meaning is also transformed in the process. For example, errors in deep learning applications change from something to be avoided and overcome to something actively investigated, as they might indicate interesting threshold cases for cultural analysis (Munk, Olesen and Jacomy, 2022). These words are the 'keys' for epistemic translations, thought of as a crossover from both sides (Hall, 2017) and ensure at the same time that methodological innovations are grounded in practice. Through keywords, we can show that deep learning practices need not be as alien to humanities research as they currently seem.

The possibilities of epistemic translation with keywords will be first explored through a close, qualitative reading of seminal deep learning literature (see also Mackenzie, 2017; Amoore *et al.*, 2023) such as AlexNet (Krizhevsky, Sutskever and Hinton, 2017) or 'Attention is all you need' (Vaswani *et al.*, 2017). Experts from diverse disciplines will be invited to contribute to close cultural readings of deep learning.

There are also recent attempts to develop a new humanities epistemology for deep learning (Fazi, 2021; Weatherby and Justie, 2022), although they remain focussed on specific architectures or applications. As we are interested in the application and materialisation of deep learning for cultural production and consumption, we analyse computer science textbooks and survey papers, following a methodology developed to understand opaque algorithmic 'security practices' (Aradau and Blanke, 2018). This close-reading effort will map the keywords that presently circulate in deep learning theories and methodologies.

These conceptual innovations underpin investigations of how these keywords of deep learning play out in practice, how they gain meaning, enable new processes and become materialized through techniques. DEEP CULTURE explores the current deep learning practices behind the keywords in data, algorithms and infrastructures by engaging scientists and practitioners working with similar data as found at the project's archival sites, but in other application domains. A deep learning engineer working for DeepMind with Instagram images or a non-governmental organization working through forgotten government data with the latest GPT4-Bing extensions to understand surveillance practices might all be seen to enact deep culture but have very different perceptions and attach distinct values to it. Qualitative research methodologies like semi-structured interviews or participant observations (Star, 1999), practice-based collaborations with scientists and practitioners, software studies methods (Fuller, 2018) and the analysis of documents and media representations in desk studies allow us to research the brittleness and knowledge regimes of dominant deep culture and its social boundaries. We examine the deep culture connections that emerge from places of its current production, how data, algorithms and people relate to each other, what kind of changes are afforded or constrained, shifting attention away from approaches that make deep culture uniform and surfacing new deep cultures of difference. Iteratively, the perceptions and values of experts and practitioners enrich our epistemic translations. We have already begun to experiment with these methodologies in diverse digital culture settings, from working with teenagers to reclaim their 'mobile ecosystems' (Pybus, Coté and Blanke, 2015) to new work on deep learning processes in health (Luitse, Blanke and Poell, 2023).

Keywords	Dominant Deep Culture	Deep Cultures of Difference	Archival Site
Detail	Profile	Narrative	Holocaust
Relationality	Vector	Context	
Discontinuity	Anomaly	Contingency	Real-time
Identity	Bias	Ambiguity	
Imagination	Generativity	Creativity	Incidental
Uncertainty	Probability	Doubt	

Table 1: Epistemic Translations

Preparing Deep Cultures of Difference

Equipped with the analysis of keywords and their practices, DEEP CULTURE addresses its other objectives of critical inquiry with and beyond deep culture (Figure 1). The project focuses on keywords in deep learning that resonate with humanities thinking and practices but seem to currently stand against them in the way they are operationalised and materialised. We explore how they can be recast in **six case studies based on the keywords** (Flyvbjerg, 2011). **Table 1 summarises the currently selected keywords – detail, relationality, discontinuity, identity, imagination and uncertainty –,** all related to ideas of thinking of and through difference. They are derived from rich debates in digital humanities about the intersections between digital humanities and cultural critique (e.g., Drucker, 2011; Prescott, 2011; Liu, 2012; Risam, 2015; Hayles, 2017; So, 2020). The second column of Table 1 shows their current operationalisation in deep learning, while the third indicates how they can be redefined. Given how fast deep learning is changing, we will need to be flexible in the selection of the keywords. Until Month 32, we will review changes to existing entries and, if necessary, add new keywords or modify existing ones based on the results of the epistemic translations.

None of the six cases eliminates the need for human interpretation and critical synthesis, and all will be jointly supervised by the PI and experts in the cases' respective fields from the historical and cultural research schools in Amsterdam. Research at the archival sites will also be supported through the PI's existing collaborations with experts in digital history and Holocaust (Martijn Eickhoff and Danielle van den Heuvel), social media and Internet studies (Thomas Poell, Stefania Milan and Richard Rogers) or incidental collections research (Charles Jeurgens and Julia Noordegraaf). They will be engaged in the project's PhD student supervision to ensure a productive, critical dialogue with machine-readings of cultural materials.

Table 1 lists the epistemic translations that underpin the six case studies and their archival sites. At the site of Holocaust archives, the first case explores how hierarchical deep representation learning facilitates concentrating on **details**, with the generic problem taken care off by larger language or image models. We reconfigure deep learning practices to '**profile**' individuals across diverse socio-historical datasets into means for detailed **narratives**, both textual and visual, to retell marginalised stories in

Holocaust collections. A focus on smaller collections and individual details is constitutive of humanities research, while also being a key digital humanities challenge. For Katherine Hayles (2017), digital methods in cultural research are relevant today as they specify meanings in digital sources and do not simply replicate what has been done in other disciplines. This would mean attending to what the humanities are good at: the marginal and overlooked. Holocaust research has been at the forefront of giving voice to individual victims. We will work with the multi-national archives of 'Jewish councils' (Jewish Museum in Prague, 2015; NIOD, 2023) and other sources like the Jewish Councils to Combat Fascism (USHMM, 2023). These and related collections contain accounts of the deportation of Jewish people and looting of their possessions across Europe. For this case study, we concentrate on often overlooked stories of resistance in 'close distant reading' (Jänicke *et al.*, 2015; Fan and Presner, 2022) and analyse the many forms of resistance that are sometimes forgotten in grand narratives of war. Investigating deep learning practices of detailing, social-media analysts will be interviewed about how they join up 'noisy' data in their historical collections with hybrid human-artificial intelligence (Wang *et al.*, 2022). At the moment, research largely ignores the historical collections involved in marketing. We will connect with the analysts through the Dutch Association of Marketing, where the PI has given talks on AI ethics (NIMA, 2023). To research and visualise new networked narratives of forgotten resistance in Holocaust sources, we aim to employ novel deep learning methods to construct detailed knowledge graphs with relationship extraction (Nigam *et al.*, 2020; Huguet Cabot and Navigli, 2021). To this end, we will also generate new language models of relevant EHRI collections and analyse the position of Holocaust collections in existing large language models like GPT4. Similar analyses will be conducted for all the project's collections, and new culturally sensitive language and image models are compiled in each case study. This way, critical cultural data can become a starting point for a deep learning that does not ignore existing curation steps.

Also working with the Holocaust archives, a second case study addresses the idea of **relationality** (Bourdieu, 1998), which brings together multiple humanistic interests from feminism (Keller, 1997; D'Ignazio and Klein, 2020) to post-colonialism (Glissant, 1997; Sullivan and Tuana, 2007) and has been suggested as a key orientation for new digital humanities (So, 2020). Deep learning's underlying **vectorisation** (Howard and Gugger, 2020; Rieder, 2020), i.e. the conversion of all data into numerical vectors, promises to support new multi-dimensional relationality. In particular, 'attention layers' can decode long complex spatial and temporal dependencies and allow different parts of a neural net to communicate (Vaswani *et al.*, 2017). We can use vectors to link EHRI's historical collections and their **contexts** better together (Gefen, Saint-Raymond and Venturini, 2021). Vectorisations, however, also evoke concerns for a total quantification of knowledge (McQuillan, 2022; Schaffer, 2023) and reasoning without context in dominant deep culture. For this study, we will organise software-studies inspired walkthroughs (Light, Burgess and Duguay, 2018; Jaton, 2021a) with Natural Language Processing (NLP) researchers working on hate speech and online antisemitism at the Institute for Logic, Language and Computation (ILLC) and elsewhere (Zannettou *et al.*, 2019; Kiela *et al.*, 2021). How do they employ techniques of 'language numericalisation' and how do embedding vectors combine and split up contexts at different stages of deep learning? We are especially interested in where linguistically trained deep learning engineers are surprised by how their neural nets communicate subtle differences and relations that are typical for hate speech. It is a key promise of EHRI to counter hate speech with profound contextual knowledge from historical archives. The EHRI collection graph is thus the perfect site to investigate how spatial and temporal vectorisation and attention layers might help integrate critical archival contextual knowledge from large national archives to smaller community- and micro-archives that are typical for the Holocaust.

Discontinuity is the third keyword, which will be studied with real-time archives. Following Michel Foucault's genealogical methods, the PI has previously suggested that traditional machine-learning methods can be reconfigured to 'seek out discontinuities' and 'small details, minor shifts, and subtle contours' (Dreyfus and Rabinow, 2014, p. 106). Building on a critical examination of how others are algorithmically produced in policing and security governance (Aradau and Blanke, 2018), time series **anomaly** detection found in security applications can be recast as a 'computational genealogy' that recognises temporal differences and discontinuities in history and is attentive to **contingency** and emergence (Blanke and Aradau, 2019). While the PI's earlier research produced promising first insights, it is limited to specific data types and longer timespans because of temporal biases in language. With deep learning, it becomes possible to analyse heterogenous, overlapping materials in web archives for temporal differences and small changes even during short time spans and for non-elites. Deep learning's optimisation is used here not to remove but to detect the unexpected. Where it struggles, we can detect a wide range of anomalies, from contextual ones and discords to point anomalies (Choi *et al.*, 2021; Pang *et al.*, 2021). For the qualitative part of the study, we will interview practitioners who work on open web materials in the field of 'Open Source Intelligence', detecting the unexpected for defence and security (Aradau and Blanke, 2018). The Turing Institute's work with the UK's signals intelligence agency GCHQ is one its most established collaborations (Janjeva, Harris

and Byrne, 2022). We will focus on the practitioners' own critique of these approaches when applied to messy data as well as their social impact – visiting their conferences, conducting interviews and encouraging them to analyse 'discontinuity' with us. Building on insights about theories and practices of deep anomaly detection, the project will set out to research contingencies in the political campaigning of grassroots civil society actors against mass surveillance according to multi-national web archive collections. Looking at the discontinuities and contingencies in their political campaigning, we can attend to minor shifts in mobilisations, reactions to contingent events and how strategies compare internationally.

The fourth case turns to real-time archives to examine how deep learning addresses '**identity**' – the key opposition to difference. In critical cultural research, identities are 'multiple', 'incomplete, in process' and make differences from an 'other' (Grossberg, 1996, p. 89). For deep learning, identity is about resolving similarities between automatically derived latent features of individuals while at the same time avoiding systematically prejudiced **biases** that correspond to outside, observable features (Singh *et al.*, 2020; Wehrli *et al.*, 2022). An identification should not be partial, and individuals should neither dominate the group they are supposed to belong to nor fall out of a correct group identification. In this move, groups are always present, and identities are assumed to be always complete. To decode what is not in this presentness and completeness, 'variational autoencoders' can reveal where the integration of data into a deep learning distribution fails because it contains biases that are hidden and intersecting across many different observable features (Amini *et al.*, 2019; Kingma and Welling, 2019). The case study analyses what kind of **ambiguities** in the data fall out of deep learning, when it is applied to integrate heterogenous real-time image collections. Where deep culture practices currently aim to identify faces across social media sites to pull more and more personal information into Big Tech ecosystems, the case study asks what is left ambiguous in these identifications. What is seen, also seen and what is not seen by deep learning's facial identifications? For the qualitative part, we will interview developers working on a social media ecosystem about how they identify faces across their collections and work around biases. We then turn this task around and develop methodologies of making ambiguities explicit. Facial recognition has been highly controversial for over-bearing collection techniques from social media by private companies and researchers alike (Van Noorden, 2020). Even if a dataset is removed by one source because of concerns, it can still be found at other sites. Harvested from photo-sharing sites, the People in Photo Albums dataset, e.g., is still accessible from servers in Germany despite being used by the Chinese surveillance company SenseTime (Harvey, 2021). The case study will survey these collections and compare them for hidden ambiguities. Reclaiming ambiguities and fragilities in facial identifications enables critical analysis and resistance against overbearing data collections.

The remaining two cases work with 'incidental archives'. The first one explores different **imaginations** with **generative** AI. Text-to-image models are criticised for not being **creative** because they ignore hundreds of years of experience of artists and make just another form of consumerist 'wizbang images' (Trevor Paglen cited in MoMA, 2023). For Noam Chomsky et al., generative AI fails to distinguish the 'possible from the impossible' (Chomsky, Roberts and Watumull, 2023). It leads to ignoring 'risks' involved in its productions, as its synthesised data seems to be not 'real' (Jacobsen, 2023, p. 6). Behind every generative data, there is, however, very 'real' data, as our first study at an incidental archive explores. We will concentrate on incidental collections that are known to feed generative AI according to tools like 'Have I been trained' (NN, 2022). With this search engine, an artist found out that her private medical records had become training data, when they left the custody of her doctor after his sudden death (Wiggers, 2022). For practices of feeding generative AI, we will interview scientists and practitioners who are working on overseeing generative AI and are linked to our European SOLARIS project (2023). We want to comprehend the criteria used to determine what generative AI systems can and cannot generate, and to discriminate the possible from the impossible. Then, we will develop new methodologies to represent absences and losses in the dominant ethico-politics of generative AI and visualise the latent spaces that sit in-between these dominant spaces, in-between what is allowed and not allowed in generative AI. To visually re-imagine these spaces, we collaborate with artists in the ARIAS network (2023) to create a public exhibition.

For the second case at incidental archives and the final keyword, we expand on the PI's earlier work to decompose algorithmic predictions into their **uncertainties** and underlying **probabilities** (Blanke, 2018; Blanke and Venturini, 2022). Critical work in the humanities understands its own research as open-ended, is comfortable with remaining **doubts** and generally embraces uncertainty (Hecht, 2004; Panagiotidou *et al.*, 2023), while algorithmic decision-making must always provide an answer (Aradau and Blanke, 2022). We unpack how machine-learning uncertainties are resolved into doubt-free judgements at human-made decision thresholds. Because deep learning can establish its own representations of its input, but it is not always clear how it does so, thresholds become a key means of communicating uncertainties in its decision-making. A model exhibits 'epistemic uncertainties' when it cannot relate new inputs to its distribution. This can be tested by rerunning the predictions with numerous slightly modified models (Lakshminarayanan, Pritzel and Blundell, 2017) and is a key part of new research on 'Robust AI' to make AI 'trustworthy'. To understand

practices of teaching a model to know when it is uncertain and should doubt itself, we conduct participant observations with the researchers and practitioners in the only recently funded ROBUST AI Dutch Long-term National Program (NWO, 2023), for which the PI sits on the social sciences and humanities committee. We turn tests for epistemic uncertainties around into a protocol to target especially those inputs that are in doubt and at the human-selected threshold, therefore deserving detailed critical attention. The decisions of the UK's Upper Tribunal on asylum and immigration (2023) will be our incidental archive. It is currently a simple repository of tribunal reports and just one example of countless similar repositories that can be found at government Internet sites around the world. By developing deep learning models to decode the outcome of asylum appeals and test underlying uncertainties, we want to understand how remaining reasonable doubts are resolved and at which thresholds. Which evidence is made to count, and which one is discounted when refugees struggle to make their case against a myriad of legal and administrative problems? Their (political) lives are in these reports, and the tribunal's decisions split these lives into legally relevant and non-relevant, evidential and non-evidential to arrive at a final judgement.

DEEP CULTURE thus sets out with the **ambitious claim that current deep culture should be not just criticised to deliver some socio-technical improvements but can also be fundamentally transformed**. New deep cultures of difference cannot be addressed by a new conceptual and methodological toolkit for research alone, but this toolbox has to be also remodelled into what political theorist Bonnie Honig has called 'public things' (Honig, 2017; Aradau and Blanke, 2022). The project questions how digital humanities can contribute to the democratisation of deep culture by making its toolkits into public things to engage diverse publics in critical analysis and action in concert. Digital humanities excel at their commitment to public participation and collaboration across disciplinary boundaries and between academic and non-academic institutions. They broaden digital methods like 'hacking', 'programming', 'tinkering', etc. to include more than just an elite of technical experts. Drawing on the PI's work with 'hackathons in concert' (Aradau and Blanke, 2022) and 'techno-cultural workshops' (Coté and Pybus, 2016), the project will advance public digital humanities to also cover the 'dark secret at the heart of AI' (Knight, 2017) of deep learning intransparency. The aim is not necessarily to develop 'better' deep culture technologies but to reclaim them as human-non-human composites that can be contested in practice. Key components of deep culture become tools of contestation and agonism rather than optimisation, efficiency and harmonisation.

Through a series of public workshops, an app for contesting deep learning will be co-developed in joint work with non-experts. We will co-produce it together into a '**little tool of difference**' (Law, 2017) that surfaces the ambiguity and contingency at the heart of deep culture and helps 'recognize and articulate difference' (Ibid.). Data from the project's archival sites can be gathered in the app, a model trained and deployed, or other cultural data can be explored like selfies or personal music collections – all without writing any code. The app will recognise texts, images, sounds, etc. and can be deployed in websites via JavaScript (Laborde, 2021) to be reused in educational programmes. To this end, we plan to re-engage with Public Data Lab and Tactical Tech. We previously built an app used to explore mobile datafications collectively in the 'Glass Room' exhibition (Tactical Tech, 2017) and demonstrated the potential of such a tool for overcoming divisions between experts and non-experts in public workshops (OrgCon, 2019). The new app can be used in similar ways by Tactical Tech and others. Non-experts will be able to explore the fragility of deep learning vectorisation processes, the qualitative decisions made to enable the quantifications but also how the project makes a difference with new deep cultures. They can find out about the kinds of remote Cloud communications involved in the training of deep learning and how this could be reduced with localised models, or how individuals across the archival sites are rendered as vectors and how this could be different by adding critical contexts. The app will be a tool for non-experts and data have-nots and allow working with deep learning in an easily accessible way.

High-gain/High-risk Dimensions
With its threefold focus on critical inquiry about, with and beyond deep culture, DEEP CULTURE aims to develop a new agenda for digital humanities in the age of deep learning, which departs fundamentally from an over-emphasis on computation by starting from epistemic relations and recasting deep culture from existing practices in one-size-fits-all commercial solutions. **This is very high risk, as we do not know whether it is even possible to reappropriate a technology in such a way and extend it to values of differences expressed in ideas such as relationality or ambiguity.** It is also very uncertain whether some of the possible transformations of deep learning resonate with existing keywords of humanities research yet, and new vocabularies might need to be found for Table 1. We might fail for some keyword-based transformations and succeed for others. **Integrating successes and failures is part of our ambitious high-risk aim to move beyond simply importing methods from the sciences in the digital humanities**, which is frequently demanded but not often attempted (Drucker, 2011; Hayles, 2017). Our practice-led research starting from epistemic translations and case studies **aims to stimulate a fundamental shift of digital**

humanities to be led by a critical analysis and by interdisciplinary interactions and not by imported technologies.

Another high risk of the project is that the age of deep learning is unfolding at unprecedented pace. The project intervenes into an ever-faster acceleration that is transforming the world around us in real time. ChatGPT has been the fastest growing app in history (Hu, 2023), 'prompting' has developed only while designing and writing this project but is now generating jobs better paid than some in machine learning (Thier, 2023). Soon to be released successors to GPT4 might well do hand-written character recognition better than Transkribus. In many ways, the project can only be a live investigation about fast changes that fundamentally transform the everyday world around us. While this speed shows how important this project is, it poses unique conceptual and methodological challenges and means that we need to be flexible in terms of starting points of our case studies and materials. They need to be pro-actively reviewed at regular intervals to adjust the direction of the project. The **gain** is, however, even higher than the risks and highly necessary, as this new machine intelligence producing novel cultural representations and encodings will not go away. **The question of deep learning and its impact on culture has become central to critical thinking about futures where we continue to live with differences** (Hall, 2021).

Section c. Team Composition, Collaboration, Project Plan and Outputs

Team: The team consists of the PI, a UvA-provided project manager, a research assistant, 2 postdocs (PDs) and 3 PhDs with co-supervisors from Amsterdam's Institute for Humanities Research – especially the School for Cultural Analysis, the School for Heritage, Memory and Material Culture, and the ILLC. The supervisors are committed as further senior researchers and engaged in the project as an advisory board. The three PhDs are each assigned to one of the archival sites; each conducting two studies individually and collaborating on two more with the other PhDs and PDs. Each PhD project (article-based) will consist of two articles on the outcomes of the qualitative research and two articles on new deep cultures at the archival sites. The PDs will focus on the interdisciplinary work and joint publications, also ensuring that the data and methods of the project are based on open-science principles. The first PD will have a humanities specialisation and co-lead with the PI on the epistemic translations and qualitative research, while the second one will have a computer science background and co-lead on the app and digital methods. Both will support the supervision of the PhDs in their respective areas. The whole team will collaborate on joint publications and events.

Collaboration: Progress in each case study is iterative and organised around the keywords. We start with a literature review and securing access where necessary. Then, we will iteratively investigate how a keyword can be redefined based on the concerns of practitioners, different valuations of developers and challenges in data. Progress across these iterations will be reported in the project blog, and each PhD will also have dedicated articles on the qualitative research outcomes. Engaging deep culture productively is a long-term project, but the choice of specific sites and cases guarantees the feasibility of this pioneering project. Access to the archival sites is ensured by open-access materials and agreements with EHRI and national web archives. For the qualitative analysis of practices, the locations depend on the study. Should we not gain access to some sites, we will use alternative qualitative research such as desk-based studies, (online) media representations or conference visits. The PI has ample experience with the methods of the project as well as working with external partners and will train other team members. In his current role, he is a key partner in several collaborations that can aid the project like Amsterdam's Research Priority Area (RPA) Humane AI or nation-wide programmes like the ELSA (Ethical, Legal and Societal Aspects) AI Labs and ROBUST AI.

Project Plan:

	Task	Owner	Year 1				Year 2				Year 3				Year 4				Year 5			
			Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
WP1: Translations	Keywords Reading Group	PD1+PhDs																				
	Deep Culture Epistemology	PD1+PhDs																				
	Qualitative Studies	PD1+PhDs																				
	Opening Conference	PD1+PI																				
	Midterm Workshop	PD1+PI																				
WP4: Case Studies	Deep Culture Methods	PhDs+all																				
	Deep Culture Data	PD2+PhDs																				
	Collections White Paper	PD2																				
	Method Labs	PI																				
	Data Labs	PI																				
WP5: Public Things	Little Tool of Difference	PD2																				
	Training	PD2																				
	Public Workshops	PDs+PI																				
WP6: Composition	Website/SocialMedia	PI+PM																				
	Conference Presentations	All																				
	Edited Book on Keywords	PDs																				
	Monograph	PI																				
	Final Conference	PI																				
	Writing Sprints	PI																				
Key Outputs	Articles, Books, Software, etc.			D1							D2		D3-7				D8-10					D11

Work Packages: DEEP CULTURE is organised in four work packages (WPs) that integrate its methodological and conceptual innovations. They correspond to the objectives of studying dominant deep culture (WP 1), of reconfiguring existing deep learning methodologies for the new critical inquiry with deep culture (WP 2), and of moving beyond monopolised deep culture through public things (WP 3+all). WP 4 addresses project ethics as well as joint work, dissemination and management.

(WP1) Epistemic Translations: This WP groups the tasks exploring epistemic relations of deep learning using close readings of seminal texts and interactions with experts. Its qualitative research situates the methodologies, epistemologies and ontologies of deep learning in their practices. **Outputs:** 2 articles on epistemic translations of deep learning (PD1+PI, key output D3, delivered by M36); Bibliography and survey (PD1, D2, M24); Online reading group of keyword texts with outside experts (PI, M36); 6 reports on observations and methods in the qualitative research (PD1+PhDs, M48); Opening conference on epistemic translations (PI+PD1, M10); Mid-term workshop on deep culture practices (M24).

(WP2) Keyword-based Case Studies: The WP will cluster the work and outputs of the case studies, as described above. The new methods from the studies will be assembled in a 'methodological commons' (Anderson, Blanke and Dunn, 2010), and culturally sensitive language and image models are shared in open repositories. In reports for the cultural collections community, we will detail the steps involved to create culturally sensitive language and image models, and the position of cultural collections in deployed large language models like GPT4. The project's main digital product will be an open-source software package and data toolkit for research in the typical low-resource, low-cost environment of humanities. They will come with extensive training materials and events as well as direct user collaborations. We will organise data and methods labs, showcasing a wide range of related approaches and promoting interoperability between different technologies and methodologies. **Outputs:** 9 articles on the results from the case studies (PhDs+all, D8, M48); 2 articles on the case study approach (PI+all, D4, M36); 6 data publications (PhDs+all, D9, M48); 2 reports for the collections community (PDs, D5, M36); Bi-monthly open data and methods labs (PI+all, M48); Contributions to methods and data markets like Hugging Face and EU-SSHOC (all, M48).

(WP3) Public Things: This WP brings together the tasks to engage diverse publics in moving beyond current deep culture. Non-experts will be invited to discuss the keywords and methods and in three public workshops co-develop the project's app. This will be done in collaboration with UvA's Humanities and Society initiative and its Humanities Lab with the City of Amsterdam; the Cultural AI lab as well as the CREATE initiative, which works across cultural institutions like the Rijksmuseum, the Royal Library, etc. Broader societal partners will be engaged with the RPA Humane AI and organisations like Tactical Tech, Public Data Lab and the European Citizen Science Association. **Outputs:** 2 Articles on the little tool of difference (PI+PD2, D7, M36); 1 Co-created app (PD2+all, D6, M36); 2 App training events (PD2, M40); 3 Public workshops (PD2+all, M48).

(WP4) Composing Deep Cultures of Difference: The WP synthesises the project's insights and organises the joint work including project management and communications. Furthermore, it addresses the ethical aspects of the project, holding regular reviews. The PI will write a monograph on deep culture and organise monthly writing sprints. The PDs will lead on a collective book from the case studies to understand how we can analyse and remake deep culture – inspired by Williams but including investigations on how deep culture can be reclaimed for both humanistic purposes and the archival sites it has been made from. Each chapter will present the results from the qualitative work on practices of identity, discontinuity, etc. in current deep culture and how they can be recast for and with our archival sites. **Outputs:** Website/social media including open-source code and data repositories (PI+all, D1, M6); Monthly writing springs (PI+all, M46); 1 edited volume (PDs+all, D10, M48); 1 monograph (PI, D11, M60); Final conference on Deep Cultures of Difference (PDs+all, M48).

Facilities/Mitigating Climate Change: The team will be trained by the PI in using the Dutch national supercomputing infrastructure (SURF, 2023). The project budget includes funding for additional dedicated computing cycles. The PI is also on the UvA board for High-Performance Computing, whose team provides additional training, as well as for the Data Science Centre, that also runs regular training workshops. We follow practices of Green Deep Learning (Xu *et al.*, 2021) including model compression, virtualisation and energy-efficient training and inference.

Open Access and Science: All publications are open access, and events will be hybrid to maximise participation possibilities. We support and encourage highest levels open access wherever possible. The curated multilingual bibliography on deep culture will be made shareable with Zotero. The app/little tool of difference and the recompiled deep culture datasets will be published with descriptions and training reports on the project's GitHub pages and the university's repositories for free download. A software and data toolkit, which makes it easy to interact with deep learning techniques for culture and embed them in research methodologies, will be released as an open-source Python library and Jupyter notebooks for reuse.

B2 References:

Ahmed, N., Wahed, M. and Thompson, N.C. (2023) 'The Growing Influence of Industry in AI Research', *Science*, 379(6635), pp. 884–886. Available at: https://doi.org/10.1126/science.ade2420.

AI Demos (2023) *Homepage*. Available at: https://www.aidemos.info/ (Accessed: 30 April 2023).

AI Now Institute (2021) *A New AI Lexicon*. Available at: https://medium.com/a-new-ai-lexicon (Accessed: 11 February 2023).

Allington, W. (2017) 'Holocaust Denial Online: The Rise of Pseudo-Academic Antisemitism on the Early Internet', *Journal of Contemporary Antisemitism*, 1(1), pp. 33–54. Available at: https://doi.org/10.26613/jca/1.1.4.

Amini, A. *et al.* (2019) 'Uncovering and Mitigating Algorithmic Bias through Learned Latent Structure', in *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*. New York, NY, USA: Association for Computing Machinery (AIES '19), pp. 289–295. Available at: https://doi.org/10.1145/3306618.3314243.

Amoore, L. (2019) 'Doubt and the Algorithm: On the Partial Accounts of Machine Learning', *Theory, Culture & Society*, 36(6), pp. 147–169. Available at: https://doi.org/10.1177/0263276419851846.

Amoore, L. *et al.* (2023) 'Machine Learning, Meaning Making: On Reading Computer Science Texts', *Big Data & Society*, 10(1). Available at: https://doi.org/10.1177/20539517231166887.

Anderson, S., Blanke, T. and Dunn, S. (2010) 'Methodological Commons: Arts and Humanities E-Science Fundamentals', *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 368(1925), pp. 3779–3796.

Anwar, M.A. and Graham, M. (2022) *The Digital Continent: Placing Africa in Planetary Networks of Work*. Oxford: Oxford University Press.

Aradau, C. and Blanke, T. (2018) 'Governing Others: Anomaly and the Algorithmic Subject of Security', *European Journal of International Security*, 3(1), pp. 1–21.

Aradau, C. and Blanke, T. (2022) *Algorithmic Reason: The New Government of Self and Other*. Oxford: Oxford University Press.

ARIAS (2023) *Homepage*. Available at: https://arias.amsterdam/ (Accessed: 17 April 2023).

Arkin, D. (2023) *Artificial Intelligence is a Sticking Point in the Hollywood Writers Strike*, *NBC News*. Available at: https://www.nbcnews.com/news/writers-strike-2023-hollywood-screenwriters-ai-concerns-rcna82543 (Accessed: 5 May 2023).

Arora, S. (2020) *Solving the Mysteries of Deep Learning - Ideas | Institute for Advanced Study*. Available at: https://www.ias.edu/ideas/arora-deep-learning-mysteries (Accessed: 7 January 2023).

Assael, Y. *et al.* (2022) 'Restoring and Attributing Ancient Texts Using Deep Neural Networks', *Nature*, 603(7900), pp. 280–283. Available at: https://doi.org/10.1038/s41586-022-04448-z.

Bal, M. (2009) 'Working with Concepts', *European Journal of English Studies*, 13(1), pp. 13–23. Available at: https://doi.org/10.1080/13825570802708121.

Banar, N., Daelemans, W. and Kestemont, M. (2020) 'Neural Machine Translation of Artwork Titles Using Iconclass Codes', in *Proceedings of the The 4th Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature. CLFL-COLING-LaTeCH-LaTeCHCLfL 2020*, Online: International Committee on Computational Linguistics, pp. 42–51. Available at: https://aclanthology.org/2020.latechclfl-1.5 (Accessed: 26 May 2022).

Barry, A., Born, G. and Weszkalnys, G. (2008) 'Logics of Interdisciplinarity', *Economy and Society*, 37(1), pp. 20–49. Available at: https://doi.org/10.1080/03085140701760841.

Bartolo, L. and Thomas, R. (2022) *Qualitative Humanities Research is Crucial to AI*. Available at: https://rachel.fast.ai/posts/2022-06-01-qualitative/ (Accessed: 21 February 2023).

BBC (2022) *AI Reunites Holocaust Survivor with Childhood Photos*, *BBC News*. Available at: https://www.bbc.com/news/technology-63483694 (Accessed: 4 January 2023).

Beer, D. (2019) *The Quirks of Digital Culture*. Bingley: Emerald Group Publishing.

Bellanova, R. *et al.* (2021) 'Toward a Critique of Algorithmic Violence', *International Political Sociology*, 15(1), pp. 121–150. Available at: https://doi.org/10.1093/ips/olab003.

Benjamin, R. (2019) *Race After Technology: Abolitionist Tools for the New Jim Code*. London: Polity.

Benjamin, W. (2009) 'The Task of the Translator: An Introduction to the Translation of Baudelaire's Tableaux Parisiens', in *Readings in the Theory of Religion*. London: Routledge, pp. 131–140.

Bennett, T., Grossberg, L. and Morris, M. (2013) *New Keywords: A Revised Vocabulary of Culture and Society*. Hoboken, NJ: John Wiley & Sons.

Berry, D. (2022) 'AI, Ethics, and Digital Humanities', in J. O'Sullivan (ed.) *The Bloomsbury Handbook to the Digital Humanities*. London: Bloomsbury, pp. 445–458.

Berry, D.M. *et al.* (2015) 'The Data Sprint Approach: Exploring the Field of Digital Humanities Through Amazon's Application Programming Interface', *Digital Humanities Quarterly*, 9(4).

Bietti, E. (2020) 'From Ethics Washing to Ethics Bashing: A View on Tech Ethics from Within Moral Philosophy', in *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*. New York, NY, USA: Association for Computing Machinery (FAT* '20), pp. 210–219. Available at: https://doi.org/10.1145/3351095.3372860.

Birchall, C. (2021) *Radical Secrecy: The Ends of Transparency in Datafied America*. Minneapolis, MN: University of Minnesota Press.

Blanke, T. (2014) *Digital Asset Ecosystems: Rethinking Crowds and Cloud*. London: Elsevier.

Blanke, T. *et al.* (2014) 'Mining Mobile Youth Cultures', in *2014 IEEE International Conference on Big Data (Big Data)*. Los Alamitos, CA, USA: IEEE Computer Society, pp. 14–17. Available at: https://doi.org/10.1109/BigData.2014.7004447.

Blanke, T. (2018) 'Predicting the Past', *Digital Humanities Quarterly*, 12(2).

Blanke, T. (2020) *AI and Humanities: New Intelligence with Attention to Detail*, *University of Amsterdam*. Available at: https://www.uva.nl/en/about-the-uva/organisation/faculties/faculty-of-humanities/research/faces-of-humanities/tobias-blanke/interview-tobias-blanke.html (Accessed: 19 March 2023).

Blanke, T. and Aradau, C. (2019) 'Computational Genealogy: Continuities and Discontinuities in the Political Rhetoric of US Presidents', *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, pp. 1–15.

Blanke, T., Bryant, M. and Speck, R. (2015) 'Developing the Collection Graph', *Library Hi Tech*, 33(4), pp. 610–623.

Blanke, T. and Hedges, M. (2013) 'Scholarly Primitives: Building Institutional Infrastructure for Humanities E-Science', *Future Generation Computer Systems*, 29(2), pp. 654–661.

Blanke, T. and Kristel, C. (2013) 'Integrating Holocaust Research', *International Journal of Humanities and Arts Computing*, 7(1–2), pp. 41–57.

Blanke, T. and Venturini, T. (2022) 'A Network View on Reliability: Using Machine Learning to Understand How We Assess News Websites', *Journal of Computational Social Science*, 5(1), pp. 69–88. Available at: https://doi.org/10.1007/s42001-021-00116-w.

Bode, K. (2017) 'The Equivalence of "Close" and "Distant" Reading; or, Toward a New Object for Data-Rich Literary History', *Modern Language Quarterly*, 78(1), pp. 77–106. Available at: https://doi.org/10.1215/00267929-3699787.

Bolukbasi, T. *et al.* (2016) 'Man Is to Computer Programmer as Woman Is to Homemaker? Debiasing Word Embeddings', in *Proceedings of the 30th International Conference on Neural Information Processing Systems*. Red Hook, NY, USA: Curran Associates Inc. (NIPS'16), pp. 4356–4364.

Born, G. (2022) 'Recommender Systems as Cultural Technologies: Steps to an Expanded Definition of Culture in/and AI', in. *NeurIPS - Workshop on Cultures in AI/AI in Culture*, New Orleans. Available at: https://ai-cultures.github.io/papers/recommender_systems_as_cultura.pdf (Accessed: 4 April 2023).

Bourdieu, P. (1998) *Practical Reason: On the Theory of Action*. Cambridge: Polity.

Bridle, J. (2023) 'The Stupidity of AI', *The Guardian*, 16 March. Available at: https://www.theguardian.com/technology/2023/mar/16/the-stupidity-of-ai-artificial-intelligence-dall-e-chatgpt (Accessed: 16 March 2023).

Brown, T.B. *et al.* (2020) 'Language Models are Few-Shot Learners'. arXiv. Available at: https://doi.org/10.48550/arXiv.2005.14165.

Brügger, N. and Finnemann, N.O. (2013) 'The Web and Digital Humanities: Theoretical and Methodological Concerns', *Journal of Broadcasting & Electronic Media*, 57(1), pp. 66–80.

Bubeck, S. *et al.* (2023) 'Sparks of Artificial General Intelligence: Early Experiments with GPT-4'. arXiv. Available at: https://doi.org/10.48550/arXiv.2303.12712.

Buolamwini, J. and Gebru, T. (2018) 'Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification', in *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*. *Conference on Fairness, Accountability and Transparency*, PMLR, pp. 77–91. Available at: https://proceedings.mlr.press/v81/buolamwini18a.html (Accessed: 4 January 2023).

Chang, K.K. and DeDeo, S. (2020) 'Divergence and the Complexity of Difference in Text and Culture', *Journal of Cultural Analytics*, 4(11), pp. 1–36.

Choi, K. *et al.* (2021) 'Deep Learning for Anomaly Detection in Time-Series Data: Review, Analysis, and Guidelines', *IEEE Access*, 9, pp. 120043–120065. Available at: https://doi.org/10.1109/ACCESS.2021.3107975.

Chollet, F. (2017) *Deep Learning with Python*. New York: Manning Publications.

Chomsky, N., Roberts, I. and Watumull, J. (2023) 'Opinion | Noam Chomsky: The False Promise of ChatGPT', *The New York Times*. Available at: https://www.nytimes.com/2023/03/08/opinion/noam-chomsky-chatgpt-ai.html (Accessed: 17 April 2023).

Ciston, S. (2023) *A Critical Field Guide for Working with Machine Learning Datasets*. Available at: https://knowingmachines.org/critical-field-guide (Accessed: 11 February 2023).

Colavizza, G. *et al.* (2021) 'Archives and AI: An Overview of Current Debates and Future Perspectives', *Journal on Computing and Cultural Heritage*, 15(1), p. 4:1-4:15. Available at: https://doi.org/10.1145/3479010.

Cornia, M. *et al.* (2020) 'Explaining Digital Humanities by Aligning Images and Textual Descriptions', *Pattern Recognition Letters*, 129, pp. 166–172. Available at: https://doi.org/10.1016/j.patrec.2019.11.018.

Coté, M. and Pybus, J. (2016) 'Simondon on Datafication. A Techno-Cultural Method', *Digital Culture & Society*, 2(2), pp. 75–92. Available at: https://doi.org/10.14361/dcs-2016-0206.

Cotter, K. (2021) '"Shadowbanning is not a thing": Black Box Gaslighting and the Power to independently know and credibly critique Algorithms', *Information, Communication & Society* [Preprint]. Available at: https://doi.org/10.1080/1369118X.2021.1994624.

Couldry, N. and Mejias, U.A. (2019) 'Data Colonialism: Rethinking Big Data's Relation to the Contemporary Subject', *Television & New Media*, 20(4), pp. 336–349.

Cox, R. and Samuels, H. (1988) 'The Archivist's first Responsibility: A Research Agenda to improve the Identification and Retention of Records of Enduring Value', *The American Archivist*, 51(1–2), pp. 28–42.

Criddle, C. and Bradshaw, T. (2022) 'Investors Seek to Profit from Groundbreaking "Generative AI" Start-Ups', *Financial Times*. Available at: https://www.ft.com/content/9c5f7154-5222-4be3-a6a9-f23879fd0d6a.

Cultural AI Lab (2021) *Homepage*. Available at: https://www.cultural-ai.nl (Accessed: 20 February 2023).

Da, N.Z. (2019) 'The Computational Case against Computational Literary Studies', *Critical Inquiry*, 45(3), pp. 601–639. Available at: https://doi.org/10.1086/702594.

Daston, L. (2022) *Rules: A Short History of What We Live By*. Princeton: Princeton University Press.

DeepMind (2023) *Homepage*. Available at: https://www.deepmind.com/ (Accessed: 11 February 2023).

Derrida, J. (1982) *Margins of Philosophy*. Chicago: University of Chicago Press.

Dhariwal, P. and Nichol, A. (2021) 'Diffusion Models Beat GANs on Image Synthesis'. arXiv. Available at: https://doi.org/10.48550/arXiv.2105.05233.

Dick, S. (2019) 'Artificial Intelligence', *Harvard Data Science Review*, 1(1). Available at: https://doi.org/10.1162/99608f92.92fe150c.

D'Ignazio, C. and Klein, L.F. (2020) *Data Feminism*. Cambridge, MA: MIT Press.

Dobson, J. (2020) 'Interpretable Outputs: Criteria for Machine Learning in the Humanities', *Digital Humanities Quarterly*, 15(2).

Dreyfus, H.L. and Rabinow, P. (2014) *Michel Foucault: Beyond Structuralism and Hermeneutics*. Chicago: University of Chicago Press.

Drucker, J. (2011) 'Humanities Approaches to Graphical Display', *Digital Humanities Quarterly*, 5(1), pp. 1–21.

Edmond, J. (2015) 'Collaboration and Infrastructure', in S. Schreibman, R. Siemens, and J. Unsworth (eds) *A New Companion to Digital Humanities*. Hoboken, NJ: John Wiley & Sons, Ltd, pp. 54–65. Available at: https://doi.org/10.1002/9781118680605.ch4.

Esposito, E. (2022) *Artificial Communication: How Algorithms Produce Social Intelligence*. Cambridge, MA: MIT Press.

Fan, L. and Presner, T. (2022) 'Algorithmic Close Reading: Using Semantic Triplets to Index and Analyze Agency in Holocaust Testimonies', *Digital Humanities Quarterly*, 16(3).

Fazi, M.B. (2021) 'Beyond Human: Deep Learning, Explainability and Representation', *Theory, Culture & Society*, 38(7–8), pp. 55–77.

Floridi, L. and Chiriatti, M. (2020) 'GPT-3: Its Nature, Scope, Limits, and Consequences', *Minds and Machines*, 30(4), pp. 681–694. Available at: https://doi.org/10.1007/s11023-020-09548-1.

Flyvbjerg, B. (2011) 'Case Study', in N.K. Denzin and Y.S. Lincoln (eds) *The SAGE Handbook of Qualitative Research*. Thousand Oaks: SAGE, pp. 301–316.

Fuller, M. (2018) 'Software Studies Methods', in J. Sayers (ed.) *The Routledge Companion to Media Studies and Digital Humanities*. London: Routledge. Available at: https://www.routledge.com/The-Routledge-Companion-to-Media-Studies-and-Digital-Humanities/Sayers/p/book/9781138844308 (Accessed: 20 April 2023).

Gaumond, E. and Wittes, B. (2023) *It Was Smart for an AI*. Available at: https://www.lawfareblog.com/it-was-smart-ai (Accessed: 11 February 2023).

Gebru, T. *et al.* (2021) 'Datasheets for Datasets', *Communications of the ACM*, 64(12), pp. 86–92. Available at: https://doi.org/10.1145/3458723.

Gefen, A., Saint-Raymond, L. and Venturini, T. (2021) 'AI for Digital Humanities and Computational Social Sciences', in B. Braunschweig and M. Ghallab (eds) *Reflections on Artificial Intelligence for Humanity*. Cham: Springer International Publishing (Lecture Notes in Computer Science), pp. 191–202. Available at: https://doi.org/10.1007/978-3-030-69128-8_12.

Gere, C. (2009) *Digital Culture*. London: Reaktion Books.

Glissant, E. (1997) *Poetics of Relation*. Translated by B. Wing. Ann Arbor: University of Michigan Press.

Goldstein, H. (2004) 'The Infinite Archive', *IEEE Spectrum*, 41(1), pp. 66–66.

Goodfellow, I.J. *et al.* (2014) 'Generative Adversarial Networks'. arXiv. Available at: https://doi.org/10.48550/arXiv.1406.2661.

Gordon, T. (2022) *ChatGPT's Threat to Google is that it Makes Search Worse, Not Better*, *Medium*. Available at: https://medium.com/@tim_23050/chatgpts-threat-to-google-is-that-it-makes-search-worse-not-better-76ace95e9427 (Accessed: 11 February 2023).

Graziani, M. *et al.* (2023) 'A Global Taxonomy of Interpretable AI: Unifying the Terminology for the Technical and Social Sciences', *Artificial Intelligence Review*, 56(4), pp. 3473–3504. Available at: https://doi.org/10.1007/s10462-022-10256-8.

Grossberg, L. (1996) 'Identity and Cultural Studies: Is That All There Is?', in P. du Gay and S. Hall (eds) *Questions of Cultural Identity*. London: Sage Publications, pp. 87–107.

Guldi, J. and Armitage, D. (2014) *The History Manifesto*. Cambridge: Cambridge University Press.

Hagerty, A. and Rubinov, I. (2019) 'Global AI Ethics: A Review of the Social Impacts and Ethical Implications of Artificial Intelligence', *arXiv preprint arXiv:1907.07892* [Preprint].

Halevy, A., Norvig, P. and Pereira, F. (2009) 'The Unreasonable Effectiveness of Data', *IEEE Intelligent Systems*, 24(2), pp. 8–12.

Hall, S. (2017) *The Fateful Triangle: Race, Ethnicity, Nation*. Cambridge, MA: Harvard University Press.

Hall, S. (2021) *Selected Writings on Race and Difference*. Durham, North Carolina: Duke University Press.

Hand, E. (2011) 'Culturomics: Word play', *Nature*, 474(7352), pp. 436–440.

Haraway, D. (1991) *Simians, Cyborgs and Women: The Reinvention of Nature*. London: Free Association Books.

Harvey, A. (2021) *Exposing.ai: PIPA*. Available at: https://exposing.ai/datasets/pipa/ (Accessed: 9 May 2023).

Hayles, N.K. (2017) *Unthought: The Power of the Cognitive Nonconscious*. Chicago: University of Chicago Press.

Hecht, J.M. (2004) *Doubt: A History: The Great Doubters and Their Legacy of Innovation from Socrates and Jesus to Thomas Jefferson and Emily Dickinson*. New York: HarperOne.

Hepp, A., Jarke, J. and Kramp, L. (2022) *New Perspectives in Critical Data Studies*. Cham: Palgrave Macmillan.

Heuts, F. and Mol, A. (2013) 'What Is a Good Tomato? A Case of Valuing in Practice', *Valuation Studies*, 1(2), pp. 125–146. Available at: https://doi.org/10.3384/vs.2001-5992.1312125.

Hjorth, L. (2018) 'Digital Cultures and Critical Studies', in J. Nussbaum (ed.) *Oxford Research Encyclopedia of Communication*. Oxford: Oxford University Press, pp. 110–121. Available at: https://doi.org/10.1093/acrefore/9780190228613.013.648.

Hochreiter, S. and Schmidhuber, J. (1997) 'Long Short-Term Memory', *Neural Computation*, 9(8), pp. 1735–1780. Available at: https://doi.org/10.1162/neco.1997.9.8.1735.

Hogarth, I. (2023) 'We Must Slow down the Race to God-Like AI', *Financial Times*. Available at: https://www.ft.com/content/03895dc4-a3b7-481e-95cc-336a524f2ac2 (Accessed: 18 April 2023).

Honig, B. (2017) *Public Things: Democracy in Disrepair*. New York: Fordham University Press.

Howard, J. and Gugger, S. (2020) *Deep Learning for Coders with fastai and PyTorch*. Sebastopol, CA: O'Reilly Media, Inc.

Hu, K. (2023) *Chatgpt Sets Record for Fastest-Growing User Base*, *Reuters*. Available at: https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/ (Accessed: 25 March 2023).

Huguet Cabot, P.-L. and Navigli, R. (2021) 'REBEL: Relation Extraction By End-to-end Language Generation', in *Findings of the Association for Computational Linguistics: EMNLP 2021. Findings 2021*, Punta Cana, Dominican Republic: Association for Computational Linguistics, pp. 2370–2381. Available at: https://doi.org/10.18653/v1/2021.findings-emnlp.204.

Iliadis, A. and Russo, F. (2016) 'Critical Data Studies: An Introduction', *Big Data & Society*, 3(2), p. 2053951716674238. Available at: https://doi.org/10.1177/2053951716674238.

Inbasekaran, A., Gnanasekaran, R.K. and Marciano, R. (2021) 'Using Transfer Learning to Contextually Optimize Optical Character Recognition Output and Perform New Feature Extraction on a Digitized Cultural and Historical Dataset', in *2021 IEEE International Conference on Big Data (Big Data). 2021 IEEE*

International Conference on Big Data (Big Data), pp. 2224–2230. Available at: https://doi.org/10.1109/BigData52589.2021.9671586.

Ingram, D. (2023) *Chatbot That Lets You Talk to Jesus and Hitler Is Latest AI Controversy*, *NBC News*. Available at: https://www.nbcnews.com/tech/tech-news/chatgpt-gpt-chat-bot-ai-hitler-historical-figures-open-rcna66531 (Accessed: 11 February 2023).

Institute of Historical Research (2017) *Born Digital Big Data and Methods for History and the Humanities | Institute of Historical Research*. Available at: https://archives.history.ac.uk/ihrcms/digital/projects/born-digital-big-data-and-methods-history-and-humanities.html (Accessed: 11 February 2023).

Jacobsen, B.N. (2023) 'Machine Learning and the Politics of Synthetic Data', *Big Data & Society*, 10(1). Available at: https://doi.org/10.1177/20539517221145372.

Jänicke, S. *et al.* (2015) 'On Close and Distant Reading in Digital Humanities: A Survey and Future Challenges.', in R. Borgo, F. Ganovelli, and I. Viola (eds). *Eurographics Conference on Visualization (EuroVis) - STARs*, The Eurographics Association, pp. 83–103. Available at: https://doi.org/10.2312/eurovisstar.20151113.

Janjeva, A., Harris, A. and Byrne, A. (2022) *The Future of Open-Source Intelligence for UK National Security*. Available at: https://cetas.turing.ac.uk/sites/default/files/2022-06/330_op_futureofopensourceintelligence_finalweb.pdf.

Jaton, F. (2021a) 'Assessing Biases, Relaxing Moralism: On Ground-Truthing Practices in Machine Learning Design and Application', *Big Data & Society*, 8(1). Available at: https://doi.org/10.1177/20539517211013569.

Jaton, F. (2021b) *The Constitution of Algorithms: Ground-Truthing, Programming, Formulating*. Cambridge, MA: MIT Press.

Jensen, C.B. and Gad, C. (2009) 'Philosophy of Technology as Empirical Philosophy: Comparing Technological Scales in Practice', in J.K.B. Olsen, E. Selinger, and S. Riis (eds) *New Waves in Philosophy of Technology*. London: Palgrave Macmillan UK (New Waves in Philosophy), pp. 292–314. Available at: https://doi.org/10.1057/9780230227279_14.

Jewish Museum in Prague (2015) *Grant for Digitization Project Jewish Council Archives in Europe*. Available at: https://www.jewishmuseum.cz/news-detail/90-430/grant-for-digitization-project-jewish-council-archives-in-europe/ (Accessed: 17 April 2023).

Jin, D.Y. (2021) *Artificial Intelligence in Cultural Production: Critical Perspectives on Digital Platforms*. London: Routledge.

Jo, E.S. and Gebru, T. (2020) 'Lessons from Archives: Strategies for Collecting Sociocultural Data in Machine Learning', in *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*. New York, NY, USA: Association for Computing Machinery (FAT* '20), pp. 306–316. Available at: https://doi.org/10.1145/3351095.3372829.

Jobin, A., Ienca, M. and Vayena, E. (2019) 'The Global Landscape of AI Ethics Guidelines', *Nature Machine Intelligence*, 1(9), pp. 389–399.

Jordan, M.I. (2019) 'Artificial Intelligence—The Revolution Hasn't Happened Yet', *Harvard Data Science Review*, 1(1). Available at: https://doi.org/10.1162/99608f92.f06c6e61.

Kantrowitz, A. (2022) 'Finally, an A.I. Chatbot That Reliably Passes "the Nazi Test"', *Slate*, 2 December. Available at: https://slate.com/technology/2022/12/chatgpt-openai-artificial-intelligence-chatbot-whoa.html (Accessed: 11 February 2023).

Keller, J. (1997) 'Autonomy, Relationality, and Feminist Ethics', *Hypatia*, 12(2), pp. 152–164. Available at: https://doi.org/10.1111/j.1527-2001.1997.tb00024.x.

Kiela, D. *et al.* (2021) 'The Hateful Memes Challenge: Competition Report', in *Proceedings of the NeurIPS 2020 Competition and Demonstration Track. NeurIPS 2020 Competition and Demonstration Track*, PMLR, pp. 344–360. Available at: https://proceedings.mlr.press/v133/kiela21a.html (Accessed: 18 April 2023).

Kim, D.S. (2022) 'Taming Abundance: Doing Digital Archival Research (as Political Scientists)', *PS: Political Science & Politics*, 55(3), pp. 530–538. Available at: https://doi.org/10.1017/S104909652100192X.

Kingma, D.P. and Welling, M. (2019) 'An Introduction to Variational Autoencoders', *Foundations and Trends in Machine Learning*, 12(4), pp. 307–392. Available at: https://doi.org/10.1561/2200000056.

Kirschenbaum, M. (2014) 'What Is "Digital Humanities," And Why Are They Saying Such Terrible Things About It?', *Differences*, 25(1), pp. 46–63. Available at: https://doi.org/10.1215/10407391-2419997.

Kitchin, R. (2014) 'Big Data, New Epistemologies and Paradigm Shifts', *Big Data & Society*, 1(1).

Kitchin, R. and Lauriault, T.P. (2015) 'Small Data in the Era of Big Data', *GeoJournal*, 80(4), pp. 463–475. Available at: https://doi.org/10.1007/s10708-014-9601-7.

Knight, W. (2017) *The Dark Secret at the Heart of AI*. Available at: https://www.technologyreview.com/2017/04/11/5113/the-dark-secret-at-the-heart-of-ai/ (Accessed: 7 January 2023).

Kristel, C., Blanke, T. and Romary, L. (2015) 'Crowds for Clouds: Recent Trends in Humanities Research Infrastructures'. arXiv. Available at: https://doi.org/10.48550/arXiv.1601.00533.

Krizhevsky, A., Sutskever, I. and Hinton, G.E. (2017) 'Imagenet Classification with Deep Convolutional Neural Networks', *Communications of the ACM*, 60(6), pp. 84–90. Available at: https://doi.org/10.1145/3065386.

Laborde, G. (2021) *Learning TensorFlow.js - Powerful Machine Learning in JavaScript*. New York: O'Reilly Media, Inc. Available at: https://www.oreilly.com/library/view/learning-tensorflowjs/9781492090786/ (Accessed: 23 February 2023).

Lakshminarayanan, B., Pritzel, A. and Blundell, C. (2017) 'Simple and Scalable Predictive Uncertainty Estimation using Deep Ensembles', in *Advances in Neural Information Processing Systems*. Curran Associates, Inc. Available at: https://proceedings.neurips.cc/paper/2017/hash/9ef2ed4b7fd2c810847ffa5fa85bce38-Abstract.html (Accessed: 17 April 2023).

Law, J. (2017) 'The Little Tools of Difference', *EASST Review*. Available at: https://www.easst.net/article/the-little-tools-of-difference/ (Accessed: 24 February 2023).

Law, J. and Mol, A. (2002) 'Local Entanglements or Utopian Moves: An Inquiry into Train Accidents', *The Sociological Review*, 50(1_suppl), pp. 82–105.

LeCun, Y. and Bengio, Y. (1998) 'Convolutional Networks for Images, Speech, and Time Series', in M.E. Arbib (ed.) *The Handbook of Brain Theory and Neural Networks*. Cambridge, MA: MIT Press, pp. 255–258.

LeCun, Y., Bengio, Y. and Hinton, G. (2015) 'Deep Learning', *Nature*, 521(7553), pp. 436–444. Available at: https://doi.org/10.1038/nature14539.

Leonelli, S. (2020) 'Scientific Research and Big Data', in E.N. Zalta (ed.) *The Stanford Encyclopedia of Philosophy*. Summer 2020. Redwood City, CA: Stanford University. Available at: https://plato.stanford.edu/archives/sum2020/entries/science-big-data/ (Accessed: 1 March 2023).

Light, B., Burgess, J. and Duguay, S. (2018) 'The Walkthrough Method: An Approach to the Study of Apps', *New Media & Society*, 20(3), pp. 881–900. Available at: https://doi.org/10.1177/1461444816675438.

Lindstrom, M. (2017) *Small Data*. London: John Murray Press.

Liu, A. (2012) 'Where Is Cultural Criticism in the Digital Humanities?', in M.K. Gold (ed.) *Debates in the Digital Humanities*. Minneapolis, MN: University of Minnesota Press. Available at: https://dhdebates.gc.cuny.edu/read/untitled-88c11800-9446-469b-a3be-3fdb36bfbd1e/section/896742e7-5218-42c5-89b0-0c3c75682a2f#ch29.

Liu, Y. *et al.* (2023) 'Summary of ChatGPT/GPT-4 Research and Perspective Towards the Future of Large Language Models'. arXiv. Available at: https://doi.org/10.48550/arXiv.2304.01852.

Lodato, T.J. and DiSalvo, C. (2016) 'Issue-Oriented Hackathons as Material Participation', *New Media & Society*, 18(4), pp. 539–557. Available at: https://doi.org/10.1177/1461444816629467.

Luitse, D., Blanke, T. and Poell, T. (2023) *AI Competitions as Infrastructures*, *The 23rd Annual Conference of the Association of Internet Researchers*. Available at: https://spir.aoir.org/ojs/index.php/spir/article/view/13044/10924 (Accessed: 2 November 2022).

Luitse, D. and Denkena, W. (2021) 'The Great Transformer: Examining the Role of Large Language Models in the Political Economy of AI', *Big Data & Society*, 8(2). Available at: https://doi.org/10.1177/20539517211047734.

Mackenzie, A. (2017) *Machine Learners: Archaeology of a Data Practice*. Cambridge, MA: MIT Press.

Makhortykh, M., Urman, A. and Ulloa, R. (2021) 'Hey, Google, Is It What the Holocaust Looked Like?', *First Monday* [Preprint]. Available at: https://doi.org/10.5210/fm.v26i10.11562.

Manoff, M. (2004) 'Theories of the Archive from Across the Disciplines', *portal: Libraries and the Academy*, 4(1), pp. 9–25.

Marciano, R. *et al.* (2018) 'Archival Records and Training in the Age of Big Data', in J. Percell et al. (eds) *Re-envisioning the MLS: Perspectives on the Future of Library and Information Science Education*. Bingley: Emerald Publishing Limited (Advances in Librarianship), pp. 179–199. Available at: https://doi.org/10.1108/S0065-28302018000044B010.

Marres, N. (2016) *Material Participation: Technology, the Environment and Everyday Publics*. London: Palgrave Macmillan.

McQuillan, D. (2022) *Resisting AI: An Anti-Fascist Approach to Artificial Intelligence*. Bristol: Policy Press.

Metz, R. (2018) *Microsoft's Neo-Nazi Sexbot Was a Great Lesson for Makers of AI Assistants*, *MIT Technology Review*. Available at: https://www.technologyreview.com/2018/03/27/144290/microsofts-neo-nazi-sexbot-was-a-great-lesson-for-makers-of-ai-assistants/ (Accessed: 11 February 2023).

Miguel, D. (2023) *Probing Negation in ChatGPT*, *Medium*. Available at: https://betterprogramming.pub/probing-negation-in-chatgpt-eb8e99cf0a9f (Accessed: 11 February 2023).

Miller, T. (2019) 'Explanation in Artificial Intelligence: Insights from the Social Sciences', *Artificial Intelligence*, 267, pp. 1–38. Available at: https://doi.org/10.1016/j.artint.2018.07.007.

Milmo, D. (2023) 'Google AI Chatbot Bard Sends Shares Plummeting After It Gives Wrong Answer', *The Guardian*. Available at: https://www.theguardian.com/technology/2023/feb/09/google-ai-chatbot-bard-error-sends-shares-plummeting-in-battle-with-microsoft (Accessed: 11 February 2023).

MoMA (2023) *How to See like a Machine*. Available at: https://www.youtube.com/watch?v=G2XdZIC3AM8 (Accessed: 1 April 2023).

Moore, N. *et al.* (2016) *The Archive Project: Archival Research in the Social Sciences*. London: Routledge.

Moretti, F. (2013) *Distant Reading*. London: Verso.

Morley, J. *et al.* (2020) 'From What to How: An Initial Review of Publicly Available AI Ethics Tools, Methods and Research to Translate Principles into Practices', *Science and Engineering Ethics*, 26(4), pp. 2141–2168. Available at: https://doi.org/10.1007/s11948-019-00165-5.

Muehlberger, G. *et al.* (2019) 'Transforming Scholarship in the Archives Through Handwritten Text Recognition: Transkribus as a Case Study', *Journal of Documentation*, 75(5), pp. 954–976. Available at: https://doi.org/10.1108/JD-07-2018-0114.

Munk, A.K., Olesen, A.G. and Jacomy, M. (2022) 'The Thick Machine: Anthropological AI between Explanation and Explication', *Big Data & Society*, 9(1). Available at: https://doi.org/10.1177/20539517211069891.

Nat Mach Intell (2022) 'Much to Discuss in AI Ethics', *Nature Machine Intelligence*, 4(12), pp. 1055–1056. Available at: https://doi.org/10.1038/s42256-022-00598-x.

NeurIPS (2022) *Cultures in AI/AI in Culture*. Available at: https://ai-cultures.github.io/ (Accessed: 7 January 2023).

Nieborg, D.B. and Poell, T. (2018) 'The Platformization of Cultural Production: Theorizing the Contingent Cultural Commodity', *New Media & Society*, 20(11), pp. 4275–4292. Available at: https://doi.org/10.1177/1461444818769694.

Nigam, V.V. *et al.* (2020) 'A Review Paper On The Application Of Knowledge Graphs On Various Service Providing Platforms', in *2020 10th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, pp. 716–720. Available at: https://doi.org/10.1109/Confluence47617.2020.9058298.

NIMA (2023) *NIMA | Nederlands Instituut voor Marketing Beroepsvereniging*, *NIMA*. Available at: https://www.nima.nl/ (Accessed: 2 May 2023).

NIOD (2023) *The 'Jewish Councils' of Western Europe*. Available at: https://www.niod.nl/en/projects/jewish-councils-western-europe (Accessed: 13 March 2023).

NN (2022) *Have I Been Trained?* Available at: https://haveibeentrained.com/ (Accessed: 17 April 2023).

Noble, S.U. (2018) *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: New York University Press.

van Noord, N. (2022) 'A Survey of Computational Methods for Iconic Image Analysis', *Digital Scholarship in the Humanities*, 37(4), pp. 1316–1338. Available at: https://doi.org/10.1093/llc/fqac003.

Noordegraaf, J. *et al.* (2021) 'Microscopic Views on a Global Pandemic: Social and Cultural Effects of the Covid-19 Pandemic as Documented in Two Dutch Community Archives', *Journal of Open Humanities Data*, 7.

Nowotny, H. (2021) *In AI we Trust: Power, Illusion and Control of Predictive Algorithms*. Hoboken, NJ: John Wiley & Sons.

NWO (2023) *NWO Funds Ten Years of AI Research with 25 Million Euros*. Available at: https://www.nwo.nl/en/news/nwo-funds-ten-years-ai-research-25-million-euros (Accessed: 17 April 2023).

Offert, F. and Bell, P. (2020) 'Generative Digital Humanities', in *Proceedings of the CHR 2020: Workshop on Computational Humanities Research*. Amsterdam: CEUR-WS, pp. 202–212.

Oppenlaender, J., Linder, R. and Silvennoinen, J. (2023) 'Prompting AI Art: An Investigation into the Creative Skill of Prompt Engineering'. arXiv. Available at: https://doi.org/10.48550/arXiv.2303.13534.

OrgCon (2019) *Homepage*. Available at: https://orgcon.openrightsgroup.org/2019/ (Accessed: 2 May 2023).

Osborne, P. (2013) *Philosophy in Cultural Theory*. London: Routledge.

O'Sullivan, J. (2022) *The Bloomsbury Handbook to the Digital Humanities*. London: Bloomsbury Publishing.

Panagiotidou, G. *et al.* (2023) 'Communicating Uncertainty in Digital Humanities Visualization Research', *IEEE Transactions on Visualization and Computer Graphics*, 29(1), pp. 635–645. Available at: https://doi.org/10.1109/TVCG.2022.3209436.

Pang, G. *et al.* (2021) 'Deep Learning for Anomaly Detection: A Review', *ACM Computing Surveys*, 54(2), p. 38:1-38:38. Available at: https://doi.org/10.1145/3439950.

Pasquale, F. (2015) *The Black Box Society: The Secret Algorithms that Control Money and Information*. Cambridge, MA: Harvard University Press.

Pasquinelli, M. and Joler, V. (2021) 'The Nooscope Manifested: AI as Instrument of Knowledge Extractivism', *AI & SOCIETY*, 36(4), pp. 1263–1280. Available at: https://doi.org/10.1007/s00146-020-01097-6.

Paullada, A. *et al.* (2021) 'Data and its (Dis)contents: A Survey of Dataset Development and Use in Machine Learning Research', *Patterns*, 2(11). Available at: https://doi.org/10.1016/j.patter.2021.100336.

Peters, B. (2016) *Digital Keywords: A Vocabulary of Information Society and Culture*. Princeton: Princeton University Press.

Prainsack, B. (2020) 'The Political Economy of Digital Data: Introduction to the Special Issue', *Policy Studies*, 41(5), pp. 439–446. Available at: https://doi.org/10.1080/01442872.2020.1723519.

Prescott, A. (2011) 'Consumers, Creators or Commentators?: Problems of Audience and Mission in the Digital Humanities', *Arts and Humanities in Higher Education*, 11(1–2), pp. 61–75. Available at: https://doi.org/10.1177/1474022211428215.

Prescott, A. (2022) 'The Grand Challenges of Digital Humanities', in J. O'Sullivan (ed.) *The Bloomsbury Handbook to the Digital Humanities*. London: Bloomsbury, pp. 385–396.

Pybus, J., Coté, M. and Blanke, T. (2015) 'Hacking the Social Life of Big Data', *Big Data & Society*, 2(2), p. 2053951715616649.

Raji, I.D., Scheuerman, M.K. and Amironesei, R. (2021) 'You Can't Sit With Us: Exclusionary Pedagogy in AI Ethics Education', in *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*. New York, NY, USA: Association for Computing Machinery (FAccT '21), pp. 515–525. Available at: https://doi.org/10.1145/3442188.3445914.

Regner, E. (2010) *Parliamentary Question | Translated EU Documents Used by Google to Develop Its Language Software | E-3436/2010 | European Parliament*. Available at: https://www.europarl.europa.eu/doceo/document/E-7-2010-3436_EN.html (Accessed: 7 January 2023).

RESAW (2013) 'Resaw | a Research Infrastructure for the Study of Archived Web Materials'. Available at: https://resaw.eu/ (Accessed: 17 April 2023).

Rezai, Y. (2022) 'Data Stories for/from All: Why Data Feminism Is for Everyone', *Digital Humanities Quarterly*, 016(2).

Rich, E. (1985) 'Artificial Intelligence and the Humanities', *Computers and the Humanities*, 19(2), pp. 117–122.

Rieder, B. (2020) *Engines of Order: A Mechanology of Algorithmic Techniques*. Amsterdam: Amsterdam University Press.

Risam, R. (2015) 'Beyond the Margins: Intersectionality and the Digital Humanities', *Digital Humanities Quarterly*, 9(2).

Risam, R. (2018) 'Decolonizing the Digital Humanities in Theory and Practice', in J. Sayers (ed.) *The Routledge Companion to Media Studies and Digital Humanities*. London: Routledge, pp. 78–86.

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

Roberge, J. and Castelle, M. (2021) *The Cultural Life of Machine Learning: An Incursion into Critical AI Studies*. New York, NY: Springer.

Rose, J. (2022) 'The AI That Draws What You Type Is Very Racist, Shocking No One', *Vice*, 13 April. Available at: https://www.vice.com/en/article/wxdawn/the-ai-that-draws-what-you-type-is-very-racist-shocking-no-one (Accessed: 11 February 2023).

Roussi, A. (2020) 'Resisting the Rise of Facial Recognition', *Nature*, 587(7834), pp. 350–353. Available at: https://doi.org/10.1038/d41586-020-03188-2.

Ryan, M. (2020) *Deep Learning with Structured Data*. London: Simon and Schuster.

Schaffer, H. (2023) *Humanist Discussion Group - Volume 36*. Available at: https://dhhumanist.org/volume/36/484/ (Accessed: 1 April 2023).

Scheuerman, M.K., Hanna, A. and Denton, E. (2021) 'Do Datasets Have Politics? Disciplinary Values in Computer Vision Dataset Development', *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2), pp. 1–37.

Schöch, C. (2013) 'Big? Smart? Clean? Messy? Data in the Humanities?', *Journal of the Digital Humanities*, 2(3).

Schreibman, S., Siemens, R. and Unsworth, J. (2008) *A Companion to Digital Humanities*. Hoboken, NJ: John Wiley & Sons.

Seaver, N. (2017) 'Algorithms as Culture: Some Tactics for the Ethnography of Algorithmic Systems', *Big Data & Society*, 4(2). Available at: https://doi.org/10.1177/2053951717738104.

Singh, R. *et al.* (2020) 'On the Robustness of Face Recognition Algorithms Against Attacks and Bias', *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(09), pp. 13583–13589. Available at: https://doi.org/10.1609/aaai.v34i09.7085.

So, R.J. (2020) *Redlining Culture: A Data History of Racial Inequality and Postwar Fiction*. New York, NY: Columbia University Press.

SOLARIS (2023) *Home | SOLARIS*. Available at: https://projects.illc.uva.nl/solaris/ (Accessed: 17 April 2023).

Srnicek, N. (2017) *Platform Capitalism*. Hoboken, NJ: John Wiley & Sons.

Srnicek, N. (2018) 'Platform Monopolies and the Political Economy of AI', in J. McDonnell (ed.) *Economics for the Many*. London: Verso, pp. 152–163.

Star, S.L. (1999) 'The Ethnography of Infrastructure', *American Behavioral Scientist*, 43(3), pp. 377–391.

Stoler, A.L. (2002) 'Colonial Archives and the Arts of Governance: On the Content in the Form', in C. Hamilton et al. (eds) *Refiguring the Archive*. Dordrecht: Springer, pp. 83–102. Available at: https://doi.org/10.1007/978-94-010-0570-8_7.

Striphas, T. (2015) 'Algorithmic Culture', *European Journal of Cultural Studies*, 18(4–5), pp. 395–412. Available at: https://doi.org/10.1177/1367549415577392.

Suissa, O., Elmalech, A. and Zhitomirsky-Geffet, M. (2022) 'Text Analysis Using Deep Neural Networks in Digital Humanities and Information Science', *Journal of the Association for Information Science and Technology*, 73(2), pp. 268–287. Available at: https://doi.org/10.1002/asi.24544.

Sullivan, S. and Tuana, N. (2007) *Race and Epistemologies of Ignorance*. Albany, NY: SUNY Press.

SURF (2023) *Dutch National Supercomputer Snellius*. Available at: https://www.surf.nl/en/dutch-national-supercomputer-snellius (Accessed: 24 April 2023).

Svensson, P. (2013) 'Beyond the Big Tent', in M.K. Gold (ed.) *Debates in the Digital Humanities*. Minneapolis, MN: University of Minnesota Press, pp. 49–63.

Svensson, P. (2016) *Big Digital Humanities: Imagining a Meeting Place for the Humanities and the Digital*. Ann Arbor, MI: University of Michigan Press.

Tactical Tech (2017) *The Glass Room*. Available at: https://tacticaltech.org/projects/the-glass-room/ (Accessed: 2 May 2023).

Taddeo, M. and Floridi, L. (2018) 'How AI Can Be a Force for Good', *Science*, 361(6404), pp. 751–752. Available at: https://doi.org/10.1126/science.aat5991.

Thier, J. (2023) *Someone with a 'Hacker Spirit' Can Earn Over $300,000 for a New Kind of AI Job*, *Fortune*. Available at: https://fortune.com/2023/03/09/new-ai-jobs-chatgpt-like-assistants/ (Accessed: 25 March 2023).

Thylstrup, N.B. and Agostinho, D. (2021) *Uncertain Archives: Critical Keywords for Big Data*. Cambridge, MA: MIT Press.

Torralba, A. and Efros, A.A. (2011) 'Unbiased Look at Dataset Bias', in *CVPR 2011*, pp. 1521–1528. Available at: https://doi.org/10.1109/CVPR.2011.5995347.

Tribunal Decisions (2023) *Immigration and Asylum Chamber: Decisions on Appeals to the Upper Tribunal*. Available at: https://tribunalsdecisions.service.gov.uk/utiac (Accessed: 12 March 2023).

Underwood, T. (2017) 'A Genealogy of Distant Reading', *Digital Humanities Quarterly*, 11(2). Available at: http://www.digitalhumanities.org/dhq/vol/11/2/000317/000317.html.

USHMM (2023) 'Records from the Jewish Council to Combat Fascism and Anti-Semitism'. Available at: https://collections.ushmm.org/search/catalog/irn47345 (Accessed: 24 April 2023).

Valdivia, A. *et al.* (2022) 'Neither Opaque nor Transparent: A Transdisciplinary Methodology to Investigate Datafication at the EU Borders', *Big Data & Society*, 9(2). Available at: https://doi.org/10.1177/20539517221124586.

Van Dijck, J. (2014) 'Datafication, Dataism and Dataveillance: Big Data between Scientific Paradigm and Ideology', *Surveillance & Society*, 12(2), pp. 197–208. Available at: https://doi.org/10.24908/ss.v12i2.4776.

Van Dijck, J., Poell, T. and Waal, M. de (2018) *The Platform Society: Public Values in a Connective World*. Oxford: Oxford University Press.

Van Noorden, R. (2020) 'The Ethical Questions That Haunt Facial-Recognition Research', *Nature*, 587(7834), pp. 354–358. Available at: https://doi.org/10.1038/d41586-020-03187-3.

Vaswani, A. *et al.* (2017) 'Attention is All you Need', in *Advances in Neural Information Processing Systems*. Curran Associates, Inc. Available at: https://proceedings.neurips.cc/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html (Accessed: 11 February 2023).

Wachter, S., Mittelstadt, B. and Floridi, L. (2017) 'Transparent, Explainable, and Accountable AI for Robotics', *Science Robotics*, 2(6).

Walden, V.G. (2022) 'Understanding Holocaust Memory and Education in the Digital Age: Before and After Covid-19', *Holocaust Studies*, 28(3), pp. 257–278. Available at: https://doi.org/10.1080/17504902.2021.1979175.

Wang, W. *et al.* (2022) 'A Survey of Hybrid Human-Artificial Intelligence for Social Computing', *IEEE Transactions on Human-Machine Systems*, 52(3), pp. 468–480. Available at: https://doi.org/10.1109/THMS.2021.3131683.

Waters, R. (2023) *Generative AI: How Will the New Era of Machine Learning Affect You?*, *Financial Times*. Available at: https://www.ft.com/content/1e34f334-4e73-4677-9713-99f85eed7ba0.

Weatherby, L. and Justie, B. (2022) 'Indexical AI', *Critical Inquiry*, 48(2), pp. 381–415.

Wehrli, S. *et al.* (2022) 'Bias, Awareness, and Ignorance in Deep-Learning-Based Face Recognition', *AI and Ethics*, 2(3), pp. 509–522. Available at: https://doi.org/10.1007/s43681-021-00108-6.

Welsh, M. (2022) 'The End of Programming', *Communications of the ACM*, 66(1), pp. 34–35. Available at: https://doi.org/10.1145/3570220.

Westerhof, G. *et al.* (2021) *Navigating Stories in Times of Transition, the Covid19 Pandemic*, *Universiteit Twente*. Available at: https://tinyurl.com/2sbuuzcv (Accessed: 19 March 2023).

Whittaker, M. (2021) 'The Steep Cost of Capture'. Rochester, NY. Available at: https://papers.ssrn.com/abstract=4135581 (Accessed: 30 March 2023).

Wiggers, K. (2022) *This Site Tells You If Photos of You Were Used to Train the AI*, *TechCrunch*. Available at: https://techcrunch.com/2022/09/21/who-fed-the-ai/ (Accessed: 16 April 2023).

Williams, R. (1958) 'Culture is Ordinary', in J. McGuigan (ed.) *Raymond Williams on Culture & Society: Essential Writings*. 2014th edn. London: SAGE Publications Ltd, pp. 1–18. Available at: https://doi.org/10.4135/9781473914766.

Williams, R. (1976) *Keywords: A Vocabulary of Culture and Society*. New York, NY: Oxford University Press.

Xu, J. *et al.* (2021) 'A Survey on Green Deep Learning'. arXiv. Available at: https://doi.org/10.48550/arXiv.2111.05193.

Zannettou, S. *et al.* (2019) 'A Quantitative Approach to Understanding Online Antisemitism'. arXiv. Available at: https://doi.org/10.48550/arXiv.1809.01644.

Zuboff, S. (2019) *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. London: Profile Books.