**Ways of Machine Seeing as a Problem of Invisual Literacy**

Geoff Cox

A picture containing text

Description automatically generated

“Seeing comes before words.” John Berger’s *Ways of Seeing* opens with this statement (1972, 7). Although words are used to make the case, the meaning is made clear that “the reciprocal nature of vision is more fundamental than that of spoken dialogue” (1972, 9).[[1]](#endnote-1) By this, Berger explains how seeing reinforces an ocularcentric paradigm of knowledge and power founded in Western modernity and colonialism, based on the relative positions of who is looking and being looked at. That every image embodies a way of seeing, as he puts it, demonstrates how, even without words, images support an elite worldview, and hence the necessity of visual literacy to draw attention to the underlying ideologies at work (including, for instance, in *Ways of Seeing*, representations of the objectified female body in art history and commercial advertising, or indeed how class privilege is reinforced in portraiture). To stress the point, Berger’s closing remarks to the first episode of the television series reflects on the medium through which his ideas were made public, and what is missing at the level of socio-technical infrastructure:

But remember that I am controlling and using for my own purposes the means of reproduction needed for these programmes. The images may be like words but there is no dialogue yet. You cannot reply to me. For that to become possible in the modern medium of communication access to television must be extended beyond its present narrow limits. Meanwhile, with this programme, as with all programmes, you receive images and meanings which are arranged. I hope you will consider what I arrange but be skeptical of it.[[2]](#endnote-2)

The critical analysis of visual culture, to which Berger’s essay continues to be a key reference, highlights the underlying conditions that allow us to see how visuality is constructed, and by extension how knowledge of the world is produced — and consolidated into worldviews. Seeing is an effective way in which power differentials are legitimatised, and yet, as Berger points out, “the relation between what we see and what we know is never settled” (1972, 7). Yet this broader understanding, or literacy, has become more and more difficult to implement as networked images no longer simply represent things in the world but are an active part of invisible visual culture, and as such exhibit new forms of distributed power.

This essay sets out to explore how the relation, identified by Berger, between what we see and what we know, has been further unsettled.[[3]](#endnote-3) How has visual literacy been transformed by developments in computer vision, underpinned as it is by developments in machine learning? It is organised into two main parts. Firstly, it asks what is at stake in the analysis of culture through literacy, and subsequently in current debates around computational literacy. In the second part, an expanded notion of visual literacy is problematised by developments in computer vision, and what is referred to as *machine ways of seeing*. Several questions arise consequently. When most images are made by machines for other machines, and distributed across planetary networks, and part of vast annotated datasets, how are worldviews reinforced differently, and what kind of literacy applies, if at all? To what extent are the relations between words and images transformed by computational processes, and what are the implications for the theoretical frameworks we apply?

The uneven relation between words and images are reinforced by the book cover image (see above), which not only repeats the opening sentences of its first page but also tellingly reproduces René Magritte’s *Treachery of Objects* (1935), in which images and their descriptive labels contradict their representational capacities. For this essay, the one you are reading fifty years later, the analogy to object recognition in computer vision is made explicit in Trevor Paglan’s *The Treachery of Object Recognition* (2019) — homage to René Magritte’s *Ceci n’est pas une pomme* (1964) — in which the original exhibition poster is overlaid with green rectangles and classification labels.[[4]](#endnote-4) The software *sees* that this is a “red and green apple.” How and what computers recognise in an image, and indeed what they misrecognise, neatly demonstrates the difficulty that underscores any seeing event and how we might conceive of literacy differently in the context of non-human languages and machine intelligence, even if the underlying ideology remains relatively unchanged.

For Berger, it was the ambiguity of the image that generated its power of signification, understood at the time of writing, in the early 1970s, through the methodologies of semiotics and structuralism, as well as the historical materialism of Walter Benjamin (whose “artwork essay” was translated into English around the same time, and who Berger credits for the inspiration of the first episode of the TV programmes).[[5]](#endnote-5) Moreover, the notion of visual literacy is a paradoxical notion, as if seeing can be equated with reading, reinforced as it has been by the linguistic metaphor at the heart of much critical theory. If both signs and images have taken on an even more prominent role of contemporary culture — the effect of the coming together of semiotics and capitalism (aka *semiocapitalism*), and in which there are ever more points of view, perspectives, attention, visual metaphors, screens, infrastructures, and so on — might we regard these as ever proliferating examples of “visual hegemony,” in which the complex operations of algorithmic and data processes that are based on existing prejudices are even more hidden?[[6]](#endnote-6) The influence of the concept of hegemony in Berger’s work is apparent in his analysis of works of art, but its applicability to computer vision remains in question,[[7]](#endnote-7) although clearly it can help to highlight some of the vested interests.

More to the point, the relation between what is visible, and the names that we give to what we see to make sense of it, persist as a problem of literacy, rooted in the tendency to conflate representations with the things that they represent (whereas we know the relation to be arbitrary if we follow the lessons of semiotics). But to what extent does media literacy remain up to the task of analysing contemporary visual culture and to understand the ways in which signs have been incorporated into the operations of capitalism? In addition, what are the misrecognitions afforded by theoretical concepts that seem to assume images to be singular or human-centred or indeed like language? This is clearly an issue when images are distributed across networks, and when they are made by machines for other machines, and at enormous scales. Moreover, images exhibit the ability to act in the world, and upon us, famously articulated by Harun Farocki’s notion of “operational images” (2004). They no longer simply represent things in the world but are an active part of invisible visual culture — part of an operation, as Farocki puts it, externalising new forms of distributed power that derive from the *eye/machine*. In the case of computer vision systems, they make judgements and decisions, and as such exercise power to shape the world in their own terms, which, in turn, upholds the argument that they embody new ways of seeing.

Seeing, then, can no longer be thought of as singular or indexical truth or reality, as it was mistakenly thought to be, but is indicative of a wider need to manifest authority and power, through distributed forms. But to what extent is this a question of literacy?Indeed, does literacy remain a useful descriptor for reading computational images, especially given its enthusiastic uptake in education policy and the creative industries, and its adaptation to other technological forms and cultural values (e.g., digital literacy or carbon literacy) — to the point where almost anything requires training to “level-up” social and economic status. Such questions over changing literacies, or whether literacy is a metaphor fit for purpose in contemporary culture, occupy the rest of this essay. But first, I briefly explain what is meant by literacy.

**Uses of literacy**

As suggested, literacy indicates not only the cultural ability to read and write but more broadly demonstrates competence or knowledge of practices that allow users to maintain and build social imaginaries (Celiński 2019, 467). The technology of writing is an example of the ways in which human cultures have been transformed, from the first written scripts developed by the Sumerians (circa 3500 BC) to the printing press of the Middle Ages in Europe to artificial intelligence today. However, it should be remembered that the skill to read and write were once specialist skills, and not something that extended to the whole of society. The Latin roots *literatus/litteratus* refer to someone who is educated, and “who knows the letters,” but who also defines the letters, and more to the point, the political structures in their learning and application. To know and define letters clearly involves divergent competences, which can be non-written, algorithmic, performative, but no less grammatical or arranged (to echo Berger).

To clarify further what is meant by literacy, I refer to Richard Hoggart’s *Uses of Literacy* (published in 1957) that challenged many of the assumptions of what constituted *culture*, a as previously the preserve of an elite (2009).[[8]](#endnote-8) The subtitle of Hoggart’s book confirms the class politics around forms of literacy and nonstandard forms of expression, and how hegemonic forms of privilege­ — related to class, but also by extension, gender and race — are underscored by the ability to read and write in ways that affirm social status and cultural value. For instance, this is evident in human speech, in accents and dialects, choice of language and so on, as well as in everyday voice-controlled devices such as Apple’s Siri or Amazon’s Echo that have preferential modes of address that similarly reflect class, gender, racial stratifications.[[9]](#endnote-9) The authority expressed in literacy is founded on the privileging of certain forms of language over others — who speaks, to whom, in what way, and under what conditions?

The politics of literacy was further developed in the interdisciplinary field of Cultural Studies (indeed the Birmingham School was founded by Hoggart), setting out to understand how meanings are produced within wider systems of power and control, even in the most everyday of circumstances, and in recognition of the global expansion and significance of culture into all aspects of life. In this sense, literacy, although historically rooted in literary studies, operates as a broader metaphor for the active role of users in “encoding/decoding” media messages (Hall 1980). The politics of this expanded media literacy is that users can produce meanings through collective action (or counter-hegemony) rather than simply receiving the message as intended as if an uninterrupted flow of information. In other words, literacy is socially constituted, and takes material forms — and this is the case in visual culture and popular media, as well as computational information based as it is on material infrastructures and language.

In summary, literacy is a combination of individual skills, a material system and a social practice — “useful” in terms of its wider application, which is enabled by its particular infrastructure, and constantly changing and being transformed by the development of inscription technologies. In her 2017 book *Coding Literacy*, Annette Vee reinforces the point: “Literacy is a widely held, socially useful and valued set of practices with infrastructural communication technologies.” (2017, 27) If, as broadly agreed, there is a moral imperative to encourage everyone to learn to read and write, then surely to code too. Moreover, where to draw the limits and decide what constitutes reading, writing, or coding/programming, let alone the ability to do any of these competently (and demonstrate “good” literacy in any one of them)? What is at stake is an expanded understanding of literacy — the ability to read, write, *and program* — an enhanced understanding of the relationship between what words mean and do.

The arguments for coding literacy seem compelling but also raise concerns about the motivations around the kinds of literacy being introduced, and to what extent this deviates from some of earlier discussions around counter-hegemony. Here various initiatives such as online tutorials and websites come to mind, such as Codecademy.org and Code.org (ominously backed by Facebook’s Mark Zuckerberg and Microsoft founder Bill Gates) and other educational platforms for the common good, particularly in the US. MIT’s Scratch is a good example of software designed to teach children how to code, or more to the point, the principles of “computational thinking” (the process of breaking down a problem into simple enough steps that even a computer would understand) (Wing 2006, 33-35). In this case, to be literate is to *think* mathematically or algorithmically, like a machine — and in the context of this essay, to see like one too. Whether this is useful remains in question, as clearly the development of literacy as a project can be seen to follow what Stefano Harney and Fred Moten have described as the proliferation of capitalist logistics through the management of pedagogy (2013).

Indeed, Vee’s book challenges the assumption that literacies are necessarily assumed to be for the social good (2017, 2). More to the point, what literacy is, and what kinds of skills are required to become literate, have tended to be kept vague to be easily shaped for vested interests. In contrast, numerous examples are to be found of *illiterate* or non-standard forms of expression, such as the use of slang and creoles in speech and writing, but also *esoteric* programming languages that offer other forms of legibility. The illegibility is precisely the point to disrupt hierarchies — an example of which is “brainfuck,”[[10]](#endnote-10) a turing-complete programming language that only uses non-alphabetical characters ><+-.,[] as commands to confound human-readability. Through such examples, it becomes clear that writing is a form of action, and not simply a referent of thinking.

Indeed, writing is programming, and programming is writing that instructs a computer about what to do in human-readable language, in turn translated by a compiler into something the computer can parse. In this way a simple human-readable instruction like “print, loop, end” can give a clear indication of what is taken place even if in fact there are multiple processes in operation. The phrase “literate programming” introduced by computer programmer Donald Knuth in 1984, indicates how a computer program contains an explanation of its logic in a natural language, such as English, interspersed with traditional source code, from which compilable source code can be generated. This is not only functional but holds aesthetic potential for Knuth: “Literature of the program genre is performable by machines, but that is not its main purpose. The computer programs that are truly beautiful, useful and profitable must be readable by people.” (1984, 97–111) This extends beyond the straightforward use of comments, or the naming of objects and abstractions to how they produce wider meaning. The ambiguity of the word “class” is an example: describing objects in programming as well as social stratification. Literacy can thus be seen to be enhanced both by computational forms and as a potential form of (class) action that helps to understand algorithmic or rule-based systems. An extract from Harwood’s codework *Class Library* (2008) illustrates the point:

# We are left with no option but to construct code that

# concretizes its opposition to this meager lifestyle.

package DON’T::CARE;

use strict; use warnings;

sub aspire {

my $class = POOR;

my $requested type = GET\_RICHER;

my $aspiration = “$requested\_type.pm”;

my $class = “POOR::$requested\_type”;

require $aspiration;

return $class- >new(@\_);

}

1;

The detail here is important as programming is a very particular kind of writing, both a description of an action and the action itself (it says what it does). As such it cannot be divorced from the social and material conditions in which it is produced and distributed. It is not simply a new way of reading and writing, but also a new way of thinking and understanding other codes. Literacy in this sense not only benefits those who acquire certain skills, but also has potential wider cultural and social ramifications, helping to force coding out of its specialization in certain disciplines and open its critical potential more widely. This should clearly not simply be the preserve of computer science, nor an elite group of specialist programmers who control the borders of media and networks. This is especially important when considering what new kinds of literacy are required to engage with contemporary inscription practices — not only text-based forms such as electronic writing and computer programming, but also visual forms such as video conferencing and computer vision — as well as the broader infrastructures in which they operate. Berger’s earlier comments on the medium of television comes to mind, and how new ways of seeing require new forms of literacy to account for the invisible realm of algorithms and database (infra)structures of computer vision. Every image embodies ways of seeing, but what ways, and how are they arranged, by whom and to what purpose?

**Invisualities**

Drawing on Berger, Nicholas Mirzoeff situates the intensification of the visual in culture is a symptom of the wider need to manifest authority through visuality (2011). What he refers to as the “right to look” is denied for some and not for others, and any act of seeing is met by a willingness to be seen, in other words is founded on reciprocity. He offers the historical example of “reckless eyeballing,” the act of looking at a white person by a slave, looking at a figure of authority in a manner considered to be violent (Mirzoeff 2011, 482). We might immediately identify a parallel in the “broken metaphor” of “master” and “slave” in programming where one process exerts control over another process within a dependent relationship (Eglash 2007).[[11]](#endnote-11) At its extreme, seeing becomes a “necropolitical” issue — to adopt the phrase by Achille Mbembé — posing the question of who lives and who dies (Mirzoeff 2011, 487).

The material destruction of human bodies and populations are increasingly performed from above, using drones for instance, as part of the “post-panoptic imaginary” that also separates the enemy as if in a computer game (Mirzoeff 2011, 488-9). Here, we are further reminded that images can kill, as in the case of the POV perspective of the bomb plunging towards its target — what Farocki calls a “suicidal camera” (2004). These “vision machines,” as Paul Virilio put it, exemplify the paradoxical notion of an act of “sightless vision” (1994, 59) — “eye machines” that do not “see” as such, but instead they “read” the world according only to the logic of the model of world they know (based on the particularities of a specific image dataset for instance). It is a lopsided model of the world, a colonial worldview that correlates with other forms of cultural imperialism (such as the hegemonic dominance of English language) and one where new forms of literacy are required that are not merely based upon representational paradigms. Image datasets confirm the problem in computer vision, wherein an algorithm constructs a worldview based upon its limited resources. An example is ImageNet, a large visual dataset used for visual object recognition, derived largely from amateur photography in North America and the annotations of precarious workers on Amazon Mechanical Turk (Malevé 2019). In other words, in describing machines as able to *see* we adopt a shorthand for calculative practices that only approximate likely outcomes by using probabilistic algorithms and models that have already been built upon inherent human prejudices related to class, gender, and race (Crawford & Paglan 2019). When computer vision systems *see*, they exercise power to shape the world in their own image, which, in turn, is built upon embedded biases and the problem of generalisation. The relation between seeing and knowing is likely to be set at the lowest common denominator.

Reading source code to understand how a machine sees is not particularly revealing in itself — despite what was argued earlier — but rather requires a more networked and relational literacy. To reiterate the point, it is not simply a case of how humans see the world, or how they use machines to see (as in the case of photography), but how machines see and produce the world in their own terms. Machinic literacy is required to understand the implications of this more fully, how our worldviews are being reproduced, and how seeing is enmeshed with the “machinic unconscious”[[12]](#endnote-12) — for not only what is seen, but for what remains unseen yet still operative. The principles of visual literacy similarly would thereby stress how seeing is not simply a way to perceive the world but also a way to act within it. But what about computer vision, and its distinctive way of seeing, and how to gain access to its underlying ideological structures and effects?

The relation between what we see and what we know remains at the centre of this, echoing Berger’s first words at the beginning of this essay in which he establishes the primordiality of the image. To repeat, “Seeing comes before words. The child looks and recognizes before it can speak. […] The reciprocal nature of vision is more fundamental than that of spoken dialogue.” (1972, 7 & 9)[[13]](#endnote-13) This echoes the phenomenology of Merleau-Ponty, who also sets out the difficulties and contradic­tions of seeing: “It is at the same time true that the world is what we see and that, nonetheless, we must learn to see it.” (1968, 4) But how do we learn to see, and more to the point how do machine learn to see? Clearly, they are not one and the same, and yet it is common to draw analogies between machine intelligence and cognitive development in humans, especially in children, following a (broadly constructivist) idea of learning as something informed by experience. Yet this can also appear superficial when applied to teaching a machine to see, as for instance, in the following example — cited by artist-researcher Nicolas Malevé — of Fei-Fei Li describing her insight into the development of ImageNet:

If you consider a child’s eyes as a pair of biological cameras, they take one picture about every two hundred milliseconds, the average time an eye movement is made. So by age three, a child would have hundreds of millions of pictures of the real world. That’s a lot of training examples. So instead of focusing on solely better and better algorithms, my insight was to give the algorithms the kind of training data that a child was given by experiences, in both quantity and quality. (2019)

The example presents a reductive equivalence between human and machine vision. Yet Malevé’s concern is more about what is implied about training and learning in general. He points out that we are all involved in the process of teaching machines to look at images in everyday situations and describes the enormous amounts of training that takes place when we use everyday devices such as smart phones and computers. Yet his interest is not so much our complicity in these processes, but to investigate which pedagogical methods might be useful. What can we learn about learning from the dynamics of machine learning? In his words, how to “transform it and be transformed by it? Or, to formulate this in terms even closer to Fei-Fei Li’s, how can we think productively about the fact that a generation of humans and algorithms are learning together to look at images?” (2019) His intervention is to ask to what extent machine learning and radical pedagogy might learn from each other.[[14]](#endnote-14) A double movement between teaching and learning can be detected between humans and machines that requires a new form of visual learning beyond mere equivalences, and with profound political consequences based on a lack of reciprocity. As previously described, the dataset can be understood to be the algorithm’s worldview, which raises the question of what this worldview is, how it was formed, and more importantly how it can be transformed. How might other ways of training elicit other ways of seeing, and in ways that are socially transformative?

By extension, how might machine learning contribute to the transformation not only of learning but of literacy? How is visual literacy challenged by machine vision and what they learn from each other? If we want to *see* the invisible world of machinic visual culture, we need to unlearn how to see like humans and learn to see more like machines, or rather, see like both in ways that departs from a Western-centred humanist standpoint and thereby embrace intersectional methodologies. The point, drawing from the posthumanities, is that more diversity might elicit a more sympathetic vision of the world — one that is less based on extraction and takes better account of the environment, especially given how resource-heavy machine learning is, as well as other species and how their distinctive ways of seeing expose alternative epistemologies.

What is rendered visible and invisible to perception is clearly crucial to this, especially in a situation where the visual field is increasingly nonhuman and distributed across different entities. Adrian Mackenzie and Anna Munster refer to an “operationalization” of visuality, in which images operate within a field of “distributed invisuality” in which relations between images count more than their indexicality or iconicity (or aura) of a single image (2019, 16). Seeing, or what they call “platform seeing,” becomes distributed through data practices and machinic assemblages, that

emphasize the importance of the formatting of image ensembles as datasets across contemporary data practices; the incorporation of platforms into hardware in devices; forms of parallel computation; and the computational architectures of contemporary artificial intelligence. These assemblages constitute the (nonhuman) activities of perception as mode of cutting into/selecting out of the entire flux of image-ensemble world. (Mackenzie & Munster 2019, 3)

The significance is that a new distributed (or networked) mode of perception is operationalized, what they call “invisual perception” — a new way of (machine) seeing which is an assemblage of its various parts: including imaging devices (such as cameras), the data they produce (which might take the form of an image), and the wider practices and infrastructures through which they are operationalized (in terms of its application) (Mackenzie & Munster 2019, 4). So, what kind of literacy is required for such an assemblage, that extends literacy beyond representational modes and human sense-making, and is attentive to the relational operations of algorithms, datasets and infrastructures?

It’s clearly not as simple as learning a different vocabulary (although that would help too) but of developing a literacy that is co-constituted, one that is more sensitised to relational operations and that shifts our attention away from the acquisition of technical know-how alone to new possibilities for aesthetic practice — in other words, that helps to expose the politics of invisuality and the envisioning of other potentials. Rather than disregard literacy, or consign it to history, or the false needs of policy-makers, we need it now more than ever to understand how forms of privilege are reproduced and naturalised through new ways of seeing.

**References**

Ali, Syed Mustafa. 2016. “A Brief Introduction to Decolonial Computing.” *XRDS: Crossroads, The ACM Magazine for Students* 22, no. 4: 16–21.

Azar, Mitra, Geoff Cox, and Leo Impett. 2021. “Introduction to Ways of Machine Seeing.” *AI & Society* (forthcoming special issue). <https://doi.org/10.1007/s00146-020-01124-6>.

Benjamin, Walter. 2002 [1936] “The Work of Art in the Age of Its Technological Reproducibility: Second Version.” In *Selected Writings, Volume 3, 1935–1938*. Edited by Howard Eiland, and Michael W. Jennings. Cambridge, MA: Belknap Press of Harvard University Press.

Berger, John. 1972. Ways of Seeing. London: Penguin.

Celiński, Piotr. 2019. *Literacies: Media Linguistics* 6, no. 4: 467.

Ciston, Sarah. 2019. “ladymouth: Anti-Social-Media Art as Research.” *Ada: A Journal of Gender, New Media, and Technology*, no. 15. [10.5399/uo/ada.2019.15.5](https://adanewmedia.org/2019/02/issue15-ciston/).

Crawford, Kate, and Trevor Paglan. 2019. “Excavating AI: The Politics of Images in Machine Learning Training Sets.” https://excavating.ai.

Eglash, Ron. 2007. “Broken Metaphor: The Master-Slave Analogy in Technical Literature.” *Technology and Culture* 48, no. 2 (April): 360-369.

Farocki, Harun. 2004. “Phantom Images.” *Public: New Localities*, no. 29. https://public.journals.yorku.ca/index.php/public/article/view/30354/27882.

# Hall, Stuart. 1980[1973]. “Encoding, decoding.” In Culture, Media, Language. Working Papers in Cultural Studies, 1972-1979. Edited by Centre for Contemporary Cultural Studies, 128-138. London: Routledge.

Harney, Stefano, and Fred Moten. 2013. The Undercommons. Wivenhoe / New York / Port Watson: Minor Compositions/Autonomedia.

Harwood. 2008. “Class Library.” In *Software Studies: A Lexicon*. Edited by Matthew Fuller. Cambridge, MA: MIT Press.

Hoggart, Richard. 2009 [1957]. *The Uses of Literacy: Aspects of Working Class Life*. London: Penguin.

Knuth, Donald E. 1984. “Literate Programming.” The Computer Journal. British Computer Society 27(2): 97–111. doi:10.1093/comjnl/27.2.97.

Mackenzie, Adrian, and Anna Munster. 2019. “Platform Seeing: Image Ensembles and Their Invisualities.” *Theory, Culture & Society* 26: 3–22.

Malevé, Nicolas. 2019. “An Introduction to Image Datasets.” *Unthinking Photography*. Available at https://unthinking.photography/articles/an-introduction-to-image-datasets**.**

Mirzoeff, Nicolas. 2011. “The Right to Look.” *Critical Inquiry* 37, no. 3 (Spring): 473-496.

Pereira**,** Gabriel. 2021. “Towards Refusing as a Critical Technical Practice: Struggling with Hegemonic Computer Vision.” *APRJA* 10, no. 1. https://doi.org/10.7146/aprja.v10i1.128185.

Merleau-Ponty, Maurice. 1968. *The Visible and the Invisible*. North Western University Press. Available at <https://monoskop.org/images/8/80/Merleau_Ponty_Maurice_The_Visible_and_the_Invisible_1968.pdf>.

Vee, Annette. 2017. *Coding Literacy: How Computer Programming Is Changing Writing*. Cambridge, MA: MIT Press.

Virilio, Paul. 1994. *The Vision Machine*. Bloomington, IN: Indiana University Press.

Wing, Jeannette. 2006. “Computational Thinking.” *Communications of the ACM* 49, no.3: 33-35.

1. . *Ways of Seeing* was a documentary made for the BBC in 1972, a four-part television series of 30-min films created by Berger and producer Mike Dibb, and the scripts were adapted into the book of the same name, published by Penguin also in 1972 (see image). Produced at a time when cultural studies was emerging as a field, it was also a repost to Kenneth Clark’s BBC television documentary *Civilisation* (1969) which espoused traditional notions of Western culture without political contextualisation. In contrast, *Ways of Seeing* emphasised the politics of representation, and how particular elite forms of knowledge were legitimated to support class, gender and racial privilege. [↑](#endnote-ref-1)
2. . Transcribed from the last part of the first episode of *Ways of Seeing* (BBC, 1972). Close attention to the means of production, and its alienating effects, further underscores the Marxist approach of Berger, and how exposing what remains hidden allows a fuller understanding of lived reality/conditions. The Brechtian techniques (alienation-effect) used for the television programmes further stress the technical apparatus of the studio, encouraging viewers not to simply watch but be forced into an analysis of their alienation. [↑](#endnote-ref-2)
3. . This question, how the relation between what we see and what we know is further unsettled by developments in machine vision, has been explored in numerous collaborative events since 2016, organised by the Cambridge Digital Humanities Network, and convened by Anne Alexander, Alan Blackwell, Geoff Cox and Leo Impett. Also see “Introduction to Ways of Machine Seeing” (Azar, Cox, Impett 2021). [↑](#endnote-ref-3)
4. . The image was part of Trevor Paglan’s exhibition “From ‘Apple’ to ‘Anomaly’” at the Barbican Centre, London, 26 September 2019 to 16 February 2020; also reproduced in Kate Crawford and Trevor Paglan’s essay “Excavating AI: The Politics of Images in Machine Learning Training Sets” (2019). [↑](#endnote-ref-4)
5. . By the shorthand “artwork essay,” I refer to Walter Benjamin’s much cited 1936 essay usually translated as “The Work of Art in the Age of Mechanical Reproduction,” that has been a standard reference for analysis of the interrelation of political, technological and artistic development under capitalism (Benjamin 2002). [↑](#endnote-ref-5)
6. . By “hegemony,” reference is made to one of the central concepts of the cultural studies movement in the 1970s, and the work of Antonio Gramsci, to describe how the ruling class manipulates the culture of that society to naturalise their own worldview as the cultural norm. [↑](#endnote-ref-6)
7. . For more on the application of hegemony to computer vision, see Gabriel Pereira’s “Towards Refusing as a Critical Technical Practice: Struggling with Hegemonic Computer Vision” (2021). [↑](#endnote-ref-7)
8. . Hoggart’s main argument is about the break-up of the old, class culture, lamenting the loss of the close-knit communities and their replacement by the emerging manufactured mass culture at that time (2009). [↑](#endnote-ref-8)
9. . To give some examples:

   Early chatbots like *Eliza* (created by Joseph Weizenbaum at MIT between 1964 and 1966) follow similar principles. It is named after George Bernard Shaw 1913 play Pygmalion (loosely based on the Greek myth) describes a bet by a professor of phonetics that he can teach a working class girl, Eliza Doolittle, upward mobility through the British class system, through the acquisition of “proper speech” (as opposed to Cockney dialect). A further example is Sarah Ciston’s chatbot ladymouth(2015), a project that tries to explain feminism to misogynists on Reddit. A further explanation can be found at https://gitlab.com/sarahciston/book/-/tree/main/source/8.5-TalkingBack#fn3-4211. [↑](#endnote-ref-9)
10. . https://esolangs.org/wiki/Brainfuck. [↑](#endnote-ref-10)
11. . The wider discussion of decolonial computing resonates with this, such as the work of Syed Mustafa Ali who argues that computing is founded upon aspects of colonialism, and that contemporary developments point to an intensification of the colonial impulse (2016). [↑](#endnote-ref-11)
12. . Benjamin’s notion of “optical unconscious” and what others have called the “machinic unconscious” are invoked by this. [↑](#endnote-ref-12)
13. . These phenomenological and semiotic references find synthesis in a post-phenomenological approach of Bernard Stiegler, whose work recognises the originarity of technology, explored in more detail in “Introduction to Ways of Machine Seeing” (Azar, Cox, & Impett 2021). [↑](#endnote-ref-13)
14. . This question, and much of the preceding description, are based on a current research project that builds on both Nicolas Malevé’s research and the ongoing work with Cambridge Digital Humanities from 2016 (see fn.3). [↑](#endnote-ref-14)