Media Culture and Cultural Techniques Working Papers
N°004
2020

“Algorithmic Empathy
On Two Paradigms of Digital Generative Literature and the Need for a Critique of AI Works”

Hannes Bajohr

University of Basel
Department
Arts, Media, Philosophy

N° 2020.004

“Algorithmic Empathy On Two Paradigms of Digital Generative Literature and the Need for a Critique of AI Works”

Hannes Bajohr

DOI: 10.5451/unibas-ep79106

Media Culture and Cultural Techniques

working papers

BMCCT veröffentlicht Arbeitspapiere im Forschungsbereich des Basler Seminars für Medienwissenschaft. Die Arbeitspapiere erscheinen in unregelmässigen Abständen in deutscher und englischer Sprache.

BMCCT publishes working papers in the research area of the Basel Seminar for Media Studies. The working papers appear at irregular intervals in German and English.

Cite this item:
Hannes Bajohr, “Algorithmic Empathy On Two Paradigms of Digital Generative Literature and the Need for a Critique of AI Works,” BMCCT working papers, (November 2020) No. 4
(DOI: 10.5451/unibas-ep79106).

The »Basel Media Culture and Cultural Techniques Working Papers« (BMCCT working papers) are published by

Seminar für Medienwissenschaft
Universität Basel
Holbeinstrasse 12
4051 Basel
medienwissenschaft.philhist.unibas.ch
bmcct@unibas.ch

ISSN 2673-5792

Editors: Mario Wimmer, Markus Krajewski, Antonia von Schöning
Concept and design:
Mario Wimmer

Hosted by University of Basel library’s eterna server

University of Basel

Department
Arts, Media, Philosophy

Algorithmic Empathy
On Two Paradigms of Digital Generative Literature and the Need for a Critique of AI Works

Hannes Bajohr

Department Arts, Media, Philosophy, University of Basel,

hannes.bajohr@unibas.ch

DOI: 10.5451/unibas-ep79106

1. Promethean Anxiety, or: Creativity as the Last *Differentia*

Let me start with an observation. It was recorded in 1942 by German philosopher Günther Anders. Having escaped the Nazis and living in California at the time, Anders brought with him the distanced sensibility of the European exile who, not unlike his fellow émigré Theodor W. Adorno, understood America, and California in particular, as the intensified expression of life in capitalist modernity. In a journal entry, which would later become the first chapter of his book *The Obsolescence of Human Beings*, he described a visit to a technology exhibition, in which a friend acted rather curiously: as if he were *ashamed* to be a human, and not a machine. This, Anders noted, was a novel phenomenon, "an entirely new *pudendum* …; a form of shame that did not exist in the past. I will provisionally call it '*Promethean shame*.'" This was to denote the "shame felt when confronted by the 'humiliatingly' high quality of fabricated things."[1]

1. Günther Anders, "On Promethean Shame," in Christopher John Müller, *Prometheanism: Technology, Digital Culture and Human Obsolescence* (London: Rowman & Littlefield, 2016), 30.

I thank Michel Chaouli, Julia Pelta Feldman, Annette Gilbert, Colin Lang, as well as the discussants at two events at University of California, Santa Barbara, and TU Braunschweig (where I had the opportunity to present versions of this paper) for their comments.

▪▪

In face of the finishedness, reliability, and repeatability of modern machines and mass-produced objects, Anders held, humans feel themselves to be deficient: Unfinished, unreliable, trapped in fragile bodies, and, in the end, confronted with the flaw of having been born rather than produced. The embarrassment of the builder in the face of the built is only the first sign of the looming obsolescence of the human. For Anders, this was connected to the atom bomb no less than to the Taylorization of the production of goods, to mass fabrication as well as—and worst of all for the European aesthete—to television and its all-pervasive reach.

One may find Anders's analysis a tad too apocalyptic, but it is nevertheless heuristically useful: It is no longer the angels or the animals—the cosmologically superior or the ontologically inferior—to which humans compare themselves. In a secular society and one in which the domination of nature is total, Anders contended, the machine and the serialized product become the new norm and foil for human self-understanding. And if "shame" is overstating the matter a little, it could be more useful to speak of a *Promethean Anxiety*: the fear of losing the status of maker, as well as the clear hierarchy between human and machine.

The current discussion about AI and creativity seems to be an especially pertinent example. Humans today tend to compare themselves rather to machines than to animals. But if for the longest time, reasoning power was the *differentia* that distinguished humans from machines, today, it is the arts that have become the most recent frontier of such human-machine comparisons, and a powerful source of Promethean Anxiety. One field that has made particularly large strides in the past decade is machine learning, especially artificial neural networks used for creating artworks. One may only point to the now-infamous "Portrait of Edmond de Belamy" from 2018 (fig. 1), an inkjet-on-canvas print billed as the first "AI generated painting," which was sold at Christie's for 432,000 Dollars. Although machine-produced art is much older, the fact that an artificial neural net was involved in the production and even figured as the artist—its formula being the signature at the bottom right of the painting—made it feel like a caesura, even if it was only a smart publicity stunt (the creators simply took a readily available machine learning algorithm and trained it on a series of paintings). The work induced, as Ian Bogost has called it, the "AI gold rush" in the visual arts, and since then we have seen more and more works like this enter the market.[2]

▪▪▪

2. Ian Bogost, "The AI-Art Gold Rush Is Here," *The Atlantic*, March 6 (2019), https://www.theatlantic.com/technology/archive/2019/03/ai-created-art-invades-chelsea-gallery-scene/584134/.

DOI: 10.5451/unibas-ep78149

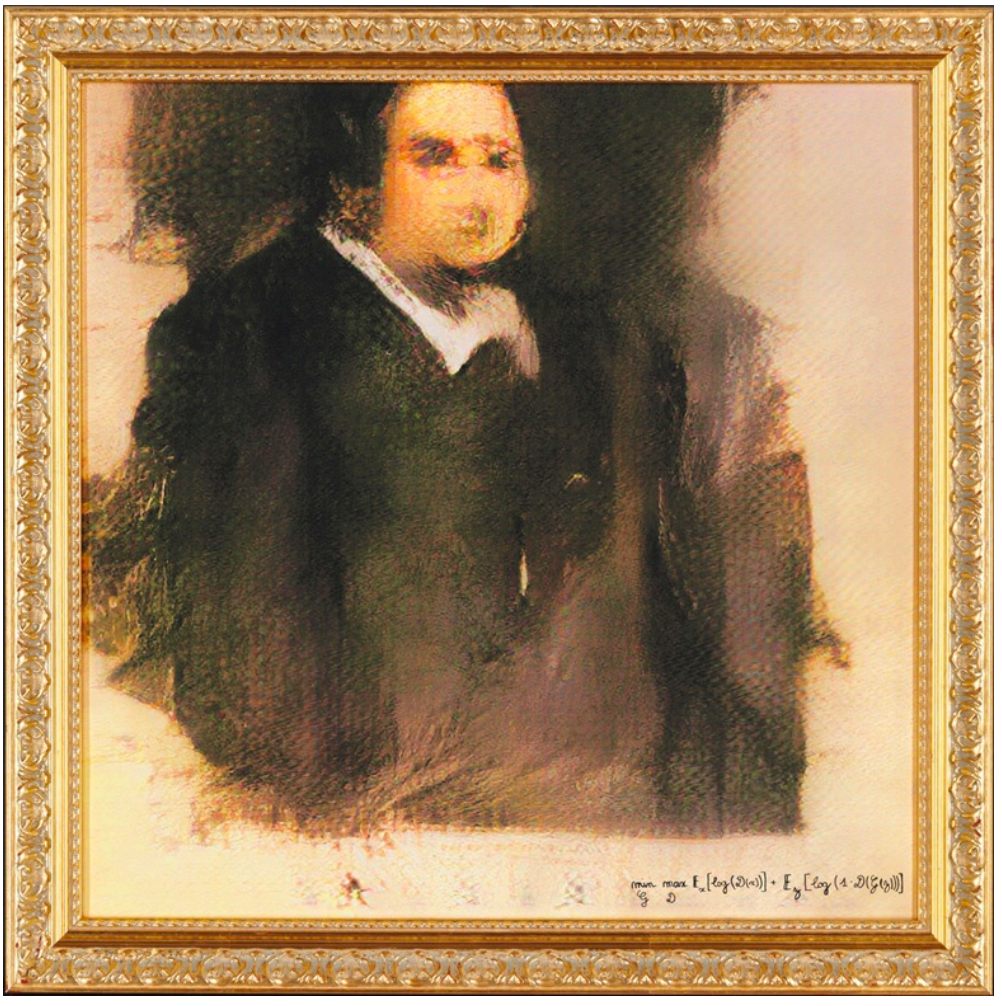


Fig. 1: Obvious Collective, "Portrait of Edmond de Belamy" (2018). The work's signature is the function that produced it: min G max D x [log (D(x))] + z [log (1 – D (G(z)))]

In the textual arts, machine learning has seen a similar popularity that was brought home by the publication of GPT-2 and, most recently, GPT-3. Machine-learning language models developed by the think tank Open AI, GPT-2 and 3 are able to produce surprisingly human-like texts, running coherently across several paragraphs. The Open AI blog entry introducing GPT-2 included an example in which the model was tasked to continue a prompt that included characters from *Lord of the Rings*—the result is, in structure and tone, fantasy fiction (fig. 2).[3] GPT-3,

3. Alec Radford et al., "Better Language Models and Their Implications," *OpenAI Blog,*

DOI: 10.5451/unibas-ep78149

released in 2020, is two orders of magnitude larger than its predecessor. The paper that announced its existence used poetry as an example, and prompted GPT-3 to, "compose a poem in the style of Wallace Stevens with the title Shadows on the Way" (fig. 3).[4]

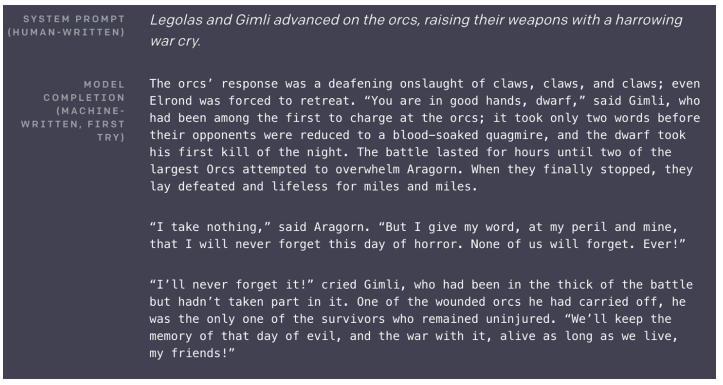


Fig. 2: Open AI, "Better Language Models and Their Implications," example output from a chosen prompt. <https://openai.com/blog/better-language-models/>

```
-------- Generated Poem 1 --------

I must have shadows on the way
If I am to walk I must have
Each step taken slowly and alone
To have it ready made

And I must think in lines of grey
To have dim thoughts to be my guide
Must look on blue and green
And never let my eye forget
That color is my friend
And purple must surround me too

The yellow of the sun is no more
Intrusive than the bluish snow
That falls on all of us.  I must have
Grey thoughts and blue thoughts walk with me
If I am to go away at all.
```

Fig. 3: Tom Brown, et al., "Language Models are Few-Shot Learners," example output from the prompt "compose a poem in the style of Wallace Stevens with the title 'Shadows on the Way'" <https://arxiv.org/pdf/2005.14165.pdf>

2019, https://openai.com/blog/better-language-models.

4. Tom B. Brown et al., "Language Models Are Few-Shot Learners," *ArXiv*, May 28, 2020: 49, http://arxiv.org/abs/2005.14165.

DOI: 10.5451/unibas-ep78149

The success of models like these has, on the side of enthusiasts of computational creativity, fueled expectations for machine learning to create complex, coherent textual works, and not least literary ones. But for most everyone else, the reactions to both Edmond de Belamy and GPT-3 contained the formulaic question "Will machines replace artists?" It is an exclamation of Promethean Anxiety, the anxiety to lose the status of creators to machines. What is more, it makes art-making the differentiating element in which humans still could triumph over machines but are at risk no longer to do so.

This anxiety even comes to the fore in apparently affirmative takes on the issue through the specific aesthetic and anthropological categories they employ. In his recent book *The Creativity Code*, Marcus Du Sautoy, Professor of Mathematics at Oxford University, appears to be excited about the possibilities of machine learning, but this enthusiasm is constantly undercut but the conviction, repeated again and again, that it is "creativity that makes us human." Du Sautoy goes so far as to posit a biologically hard-wired "creative urge"—a classic anthropological *differentia*—that he holds up against the encroachment of the machines into art, music, and literature. Starting his book with the question: "Can machines be creative?", he ends it with the defiant proclamation that: "Creativity is about humans asserting they are not machines."[5]

That art is the test case for human-machine difference is also posited by Arthur I. Miller's book *The Artist in the Machine*. Although Miller is more open to non-human aesthetics, his rhetoric nevertheless constantly falls back onto the anthropological comparison he claims as only one option among many. The question for him becomes not just whether machines must have reason or consciousness to make art, but also emotions, which are then expressed in their products. In AI art, Miller contends, computers "exhibit not only their creativity but their inner lives." Given such rhetoric, it is not surprising that Miller employs the word "genius" an awful lot both for human and machine artists.[6]

Both Miller and Du Sautoy make creativity (often defined evolutionarily) the criterion for deciding what art is; and they understand art as being inherently about expression and intention. But these strongly anthropocentric categories ignore any contemporary aesthetic theory that is not a neuroaesthetics, and their conception of

5. Marcus Du Satoy, *The Creativity Code: How AI is Learning to Write, Paint, and Think* (London: Fourth Estate, 2019), 297, 302.
6. Arthur I. Miller, *The Artist in the Machine: The World of AI-Powered Creativity, The Artist in the Machine (*Cambridge, Mass.: MIT Press, 2019), 54.

DOI: 10.5451/unibas-ep78149

art is severely out of date. It is at least telling that the non-digital artworks in Du Sautoy's book are no more recent than the 1950s, when avant-garde art movements like abstract expressionism celebrated the spontaneous, creative genius one last time.[7]

What is more, these approaches seem insufficient for what is needed now—*a critique of AI works*. Particularly, they are too laden with Promethean Anxiety to capture what is specific to the aesthetic use of AI. Instead, they tend to work by a logic of transference—first from human to machine, and then from old media to new. "The Portrait of Edmond de Belamy" is the best example here: the old medium of painting in a new media guise, created not by humans but (ostensibly) produced by a machine. But it may be more interesting, and more productive, to investigate aesthetic approaches beyond the foil of the human, and to explore the affordances of the new medium instead of simply replicating old ones.

Forgoing the talk about conscious machines as well as a specifically human creative urge, I instead want to look at the way these works *work*, and which structures they implement. I will focus particularly on digital literature, where machine learning presents us with a new paradigm of textual production. It stands in contrast to traditional, long-established algorithmic literature. The new type I want to provisionally call the *connectionist*, the old type the *sequential* paradigm. While the rule-based sequential paradigm of digital literature can look back on a rich critical apparatus, the "non-transparent" connectionist paradigm is still under-theorized. In what follows, I would like to offer some reflections on the differences between both paradigms, and hint at what we should keep in mind while developing a critique of AI aesthetics.

2. Two Types of Digital Literature: Sequential and Connectionist

Digital—or electronic—literature is a wide-ranging, many-faceted field. It contains such a large variety of genres—from hypertext to codeworks to kinetic literature—that it is hard to offer a definition that goes beyond its very basic characteristics. In the formulation of the Electronic Literature Organization, the term refers "to works with an important literary aspect that takes advantage of the capabilities and contexts provided by the stand-alone or networked computer."[8] Literary scholar

7. The one notable exception is Du Sautoy's discussion of Gerhard Richter's permutative work "4900 Farben." While he reflects on the uses of mathematics for art, he in no way engages with it as an example of a non-expressive, conceptual art practice.
8. Quoted in N. Katherine Hayles, *Electronic Literature: New Horizons for the Literary* (Notre

DOI: 10.5451/unibas-ep78149

Jessica Pressman noted that more recent works of digital literature try to align themselves again with the modernist tradition.[9] Among the genres and traditions of digital literature, the most inherently modernist is also the oldest—far older than hypertext, code works, and kinetic literature.[10] I will refer to it as *generative literature*, and it is this genre I want to focus on in this essay. At its most basic, generative literature denotes the automatic production of text according to predetermined parameters, usually following a combinatory, sometimes aleatory logic, and it emphasizes the production rather than the reception of the work (unlike, say, hypertext). Scott Rettberg, in his recent book *Electronic Literature*, highlights the generative tradition's connection to Dada and Surrealism, to Oulipo as well as Fluxus.[11] I would add conceptual art as a further important reference, since here, too, the formulation of a concept and its execution into a work are strictly separated, and one may see the relation between concept and work repeated in that between code and output.[12]

Yet it is not only its age and the historic lineage, but also something about the use of the underlying technology that gives generative literature a special status among the wide variety of digital literature at large. Unlike, for instance, Flash works, it seems to reflect and its underlying technology most clearly.

This is visible already in one of the first examples of generative literature: Theo Lutz "Stochastische Texte" (*Stochastic Texts*) which he wrote—or rather, had generate—in 1959, one year after Anders' book on the Promethean shame was published. "Stochastische Texte," is the result of an algorithm combining elements from a predetermined vocabulary taken from Franz Kafka's *Castle*.[13] Each line contains statements that are connected by conjunctions or separated by a period, such as "NOT EVERY LOOK IS NEAR. NO TOWN IS LATE", or "A CASTLE IS FREE AND EVERY FARMER IS FAR", or "EVERY STRANGER IS FAR. A DAY IS LATE," and so on (fig. 4).

Dame: University of Notre Dame Press, 2009), 3.

9. Jessica Pressman, *Digital Modernism: Making It New in New Media* (Oxford: Oxford University Press, 2014).

10. See Florian Cramer, *Words Made Flesh: Code, Culture, Imagination* (Rotterdam: Piet Zwart Institute, 2005).

11. Scott Rettberg, *Electronic Literature* (Cambridge: Polity, 2019).

12. Hannes Bajohr, "Das Reskilling der Literatur," in *Code und Konzept: Literatur und das Digitale*, ed. Hannes Bajohr (Berlin: Frohmann, 2016): 7–21.

13. See Kurt Beals, "'Do the New Poets Think? It's Possible': Computer Poetry and Cyborg Subjectivity," *Configurations* 26, no. 2 (2018): 149–77.

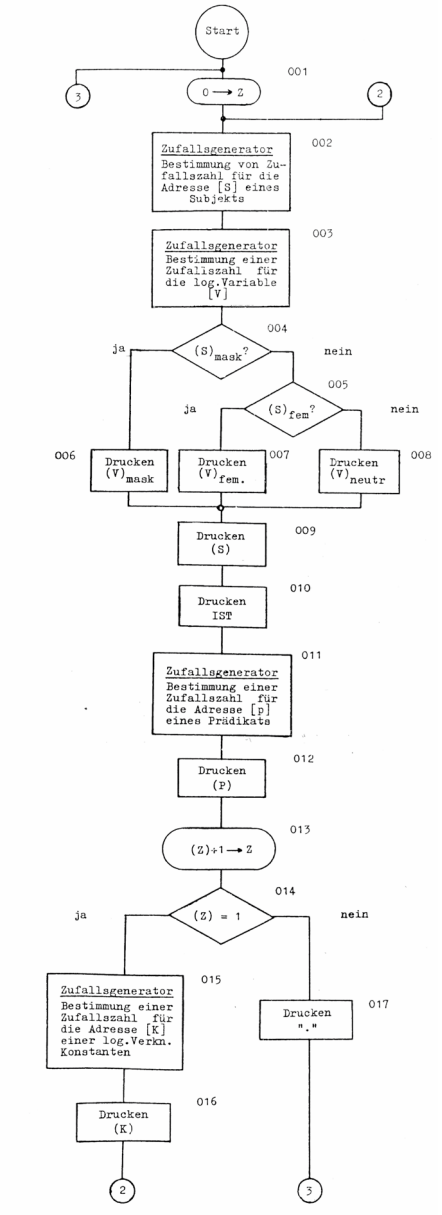
DOI: 10.5451/unibas-ep78149

NICHT JEDER BLICK IST NAH. KEIN DORF IST SPÄT.
EIN SCHLOSS IST FREI UND JEDER BAUER IST FERN.
JEDER FREMDE IST FERN. EIN TAG IST SPÄT.
JEDES HAUS IST DUNKEL. EIN AUGE IST TIEF.
NICHT JEDES SCHLOSS IST ALT. JEDER TAG IST ALT.
NICHT JEDER GAST IST WÜTEND. EINE KIRCHE IST SCHMAL.
KEIN HAUS IST OFFEN UND NICHT JEDE KIRCHE IST STILL.
NICHT JEDES AUGE IST WÜTEND. KEIN BLICK IST NEU.
JEDER WEG IST NAH. NICHT JEDES SCHLOSS IST LEISE.
KEIN TISCH IST SCHMAL UND JEDER TURM IST NEU.
JEDER BAUER IST FREI. JEDER BAUER IST NAH.
KEIN WEG IST GUT ODER NICHT JEDER GRAF IST OFFEN.
NICHT JEDER TAG IST GROSS. JEDES HAUS IST STILL.
EIN WEG IST GUT. NICHT JEDER GRAF IST DUNKEL.
JEDER FREMDE IST FREI. JEDES DORF IST NEU.
JEDES SCHLOSS IST FREI. NICHT JEDER BAUER IST GROSS.
NICHT JEDER TURM IST GROSS ODER NICHT JEDER BLICK IST FREI.

Fig. 4: Theo Lutz, "Stochastische Texte, " *augenblick* 4, no. 1 (1959): 3–9.

Lutz' "Stochastische Texte" belong to what I would like to call the *sequential paradigm* within the genre of generative literature: It is executed as a sequence of rule-steps, and its identity is encoded in its production much more than in its reception. A colleague of Lutz, while not providing the program code, sketched the program flow chart in a later article, and the sequential and step-wise nature is obvious here (fig. 5). Instead of hoping to recreate intuition, genius, or expression, it is the logic of the machine itself, that is, the logic of deterministically executed rule steps, that becomes aesthetically normative in the "Stochastische Texte." One could sense in this an *algorithmic empathy*, a non-anthropocentric empathy aimed not at the psychological states of the artists but at understanding the process of the work's material production.

DOI: 10.5451/unibas-ep78149



Flußdiagramm für ein Programm der Rechenanlage Zuse Z 22
Abb. 1

Fig. 5: Rul Gunzenhäuser, “Zur Synthese von Texten mit Hilfe programmgesteuerter Ziffernrechenanlagen,” *MTW* 10, no. 4 (1963), 4.

DOI: 10.5451/unibas-ep78149

▪▪

For Lutz' text, we only have an abstract description of the individual steps—the code he used is not (yet) available.[14] For much of contemporary digital literature, this is fortunately not the case. A more recent and more complex example of a sequential work that inspires algorithmic empathy is Nick Montfort's 2014 *Megawatt*.[15] It not only refers to its own structural make-up but also to that of a modernist classic: It is both an interpretation and an appropriation of Samuel Beckett's novel *Watt*.[16]

Written between 1942 and 1944 and published only in 1958, *Watt* depicts the titular Mr. Watt's entry into the household of Mr. Knott as the latter's servant. However, it is not the *fabula* but the linguistic structure, the textual surface that is most characteristic about this novel. In addition to the consciously unidiomatic English, the extremely repetitive passages stand out—it's "geometric audacity,"[17] as T. W. McCormack called it—which since *Watt*'s publication have been interpreted as a failure of language and a critique of the insurmountable hyper-rationality of modernity.[18] Take for instance this sentence, in which Watt cannot follow a conversation partner because he is distracted by voices in his head:

> Now these voices, sometimes they sang only, and sometimes they cried only, and sometimes they stated only, and sometimes they murmured only, and sometimes they sang and cried, and sometimes they sang and stated, and sometimes they sang and murmured, and sometimes they cried and stated, and sometimes they cried and murmured, and sometimes they stated and murmured, and sometimes they sang and cried and stated, and sometimes they sang and cried and murmured, and sometimes they cried and stated and murmured, and sometimes they sang and cried and stated and murmured, all together, at the same time, as now, to mention only these four kinds of voices, for there were others. And sometimes Watt understood all, and sometimes he understood much, and sometimes he understood little, and sometimes he understood nothing, as now.[19]

▪▪▪

14. It seems that Toni Bernhart is in the process of retrieving the original programming code for "Stochastische Texte." A first preview is given in Toni Bernhart, "Beiwerk als Werk: *Stochaistische Texte* von Theo Lutz," *editio* 32 (2020): 180-206.
15. Nick Montfort, *Megawatt* (Cambridge, Mass.: Bad Quarto, 2014).
16. Samuel Beckett, *Watt* (New York: Grove Press, 1970).
17. W. J. McCormack, "Seeing Darkly: Notes on T. W. Adorno and Samuel Beckett," *Hermathena*, no. 141 (1986): 24.
18. See Linda Ben-Zvi, "Samuel Beckett, Fritz Mauthner, and the Limits of Language," *PMLA* 95, no. 2 (March 1980): 183; Shane Weller, "Humanity in Ruins: Samuel Beckett," in *Language and Negativity in European Modernism (*Cambridge: Cambridge University Press, 2018): 90–125.
19. Beckett, *Watt*, 29.

DOI: 10.5451/unibas-ep78149

▪▪

A recent interpretation of *Watt* by Amanda M. Dennis speaks of these repetitions as "obsessive loops." "Certain passages make language appear to 'glitch,' as if it were a malfunctioning computer program or electronic device."[20] When one takes a closer look at *Megawatt*, that is, Nick Montfort's text based on *Watt*, one begins to doubt whether the metaphor of the glitch is appropriate. Indeed, *Megawatt* shows that the "obsessive loops" are not glitches, not errors in Beckett's program, but on the contrary represent its most consistent execution. In fact, these repetitive, list-like loops seem to follow an immanent rule—an *algorithm*.

Taking a closer look at this passage from *Watt* allows us to infer the production principle of what Hugh Kenner has called Beckett's "Cartesian sentences."[21] The first sentence applies a simple text generation rule: the permutation of combinatorial possibilities from a finite set of elements. The "voices" can take on four possible states—"sang," "cried," "stated," "murmured"—either individually or in various combinations, and Beckett cycles through all of them. In the second sentence—beginning all the way at the bottom of the passage, "And sometimes Watt …"—Watts' understanding is assigned the values "all", "much", "little," and "nothing," one after the other; here, the verbs are not permutated, but only listed. Programmatically speaking, the sentences resemble a function that assigns a value to a variable, and it could be generated automatically with the same result by a script.

This is exactly what Montfort did in *Megawatt*. It is in fact a reconstruction *and* an extension of Beckett's novel in one. Montfort selected passages with such "obsessive loops" from the original, and recreated them in a Python script. In the first chapter, titled "The Voices," he turns to the passage just discussed, and generates it. But the script goes further:

> Watt heard voices. Now these voices, sometimes they sang only, and sometimes they cried only, and sometimes they stated only, and sometimes they murmured only, and sometimes they babbled only, and sometimes they chattered only, and sometimes they ranted only, and sometimes they whispered only, and sometimes they sang and cried, and sometimes they sang and stated, and sometimes they sang and murmured, and sometimes they sang and babbled, and sometimes they sang and chattered, and sometimes they sang and ranted, and sometimes they sang and whispered, and sometimes they cried and stated, and sometimes they cried and murmured, and sometimes they cried and babbled, and sometimes they cried and

▪▪▪

20. Amanda M. Dennis, "Glitches in Logic in Beckett's *Watt*: Toward a Sensory Poetics," *Journal of Modern Literature* 38, no. 2 (2015), 104.
21. Hugh Kenner, *The Mechanic Muse* (New York: Oxford University Press, 1987), 91.

DOI: 10.5451/unibas-ep78149

▪

> chattered, and sometimes they cried and ranted, and sometimes they cried and whispered, and sometimes they stated and murmured, and sometimes they stated and babbled, and sometimes they stated and chattered, and sometimes they stated and ranted, and sometimes they stated and whispered, and sometimes they murmured and babbled, and sometimes they murmured and chattered, and sometimes they murmured and ranted, and sometimes they murmured and whispered, and sometimes they babbled and chattered, and sometimes they babbled and ranted, and sometimes they babbled and whispered, and sometimes they chattered and ranted, and sometimes they chattered and whispered, and sometimes they ranted and whispered, and sometimes they sang and cried and stated, and sometimes they sang and cried and murmured. … And sometimes Watt understood all, and sometimes he understood most, and sometimes he understood much, and sometimes he understood half, and sometimes he understood little, and sometimes he understood less, and sometimes he understood bits, and sometimes he understood nothing, as now.[22]

Because Beckett admits that there are more voices ("for there were others," as it says at the end of the first sentence), and because Montfort knows that in a permutation series the number of possibilities per element increases exponentially, he adds four more verbs to Beckett's four: "babbled", "chattered", "ranted", and "whispered". Likewise, Watt can now additionally understand "most", "half", "less", and "bits". Montfort's own contribution consists of the first three words, the merely expository first sentence ("Watt heard voices"), and the eight additional words. Both Beckett's text and the extensions, however, are generated purely by the code. It outputs what Beckett actually wrote (the orange text), and what he *would have written,* according to his own rules, if he had expanded his set of elements (the black text).

This can be seen very clearly in the source code of the program, which Montfort wrote in the Python programming language and which is printed in the appendix of the book:[23]

```
1   #### THE VOICES
2   text.append('\n# I\n\n')
3   def combine(num, words):
4       final = []
5       if num > 0 and len(words) >= num:
6           if num == 1:
7               final = final + [[words[0]]]
8           else:
9               final = final + [[words[0]] +
```

▪▪▪

22. Montfort, *Megawatt*, 1, 7.
23. Ibid., 242–43.

DOI: 10.5451/unibas-ep78149

```
10                c for c in combine(num-1, words[1:])]
11          final = final + combine(num, words[1:])
12    return final
13    ## In Watt the voices = ['sang', 'cried', 'stated',
      'murmured']
14    ## And Watt understood = ['all', 'much', 'little',
      'nothing']
15    ## Here the voices did eight things and there are eight
      levels:
16    voices = ['sang', 'cried', 'stated', 'murmured',
      'babbled', 'chattered', 'ranted', 'whispered']
17    understood = ['all', 'most', 'much', 'half', 'little',
      'less', 'bits', 'nothing']
18    para = ''
19    preface = ', and sometimes they '
20    for num in range(len(voices)):
21       for word_list in combine(num + 1, voices):
22           para = para + preface + ' and '.join(word_list)
23           if len(word_list) == 1:
24               para = para + ' only'
25    para = ('Watt heard voices. Now these voices,' + para[5:] +
      ', all together, at the same time, as now, to mention ' +
      'only these ' + spelled_out[len(voices)] + ' kinds of voices, for ' +
      'there were others. And sometimes Watt understood ' +
      ', and sometimes he understood '.join(understood) + ', as now.')
26    text.append(para)
```

After defining the function *combine* in lines 3–12—a subroutine that in the end assembles the final the text—Montfort first shows how Beckett's own text can be understood as a set of elements of a list variable (sometimes also called an array), that is, a single variable that contains a series of items. Here, the variable is called *voices,* in line 13, and its values are "sang," "cried," "stated," "murmured"—exactly the verbs that are permutated in *Watt*. But because there is a pound sign in front of this line, the interpreter recognizes that the line is merely a *comment* that should not be executed and ignores it. Beckett's concept is still present in the code, but has been, as it were, switched off.

Instead, line 16—an executable line—contains the new list variable, this time extended by Montfort. In addition to the original four verbs, it also contains the four additional ones: "babbled, chattered, ranted, whispered." The same happens for the variable "understood"—first, Montfort list the four original elements in a comment in line 14, then he lists his extended set in line 17.

The rest of this short code section assembles these elements. First, the empty variable *para* is defined in line 18—it will be assigned the finished text at the end. Line 19 defines the variable *preface*, which contains the regularly recurring statement "and sometimes they." A doubly nested loop follows in lines 20 to 23: It cycles through the list variable *voices* and adds the words "and sometimes they"

DOI: 10.5451/unibas-ep78149

stored in *preface*. Finally, the first sentence—the one with the voices—is completed in line 25—and the second sentence—the one about understanding, gets added to it. In the second sentence—the one at the end—, the elements are not permuted; instead, the values stored in the variable "understood" are simply listed. The result is the new, extended text—which I cannot show you completely, because it has grown exponentially and is now 27 pages long.

Megawatt is a form of algorithmic empathy that is not a copy, but a reconstruction. But while *Megawatt* is to *Watt* what Jorge Luis Borges' *Pierre Menard* is to Cervantes' *Don Quixote*—a reenactment—it is also, because it is not only reconstructive but also productive, what Joyce's *Ulysses* is to the *Odyssey*—an extension of the original. *Megawatt* is thus, first, interesting as a literary product—as a variation or rather adaptation of an existing text. But, secondly, *Megawatt* also produces knowledge about Beckett's text—it carries out a hermeneutic movement, albeit a non-anthropocentric one. It begins with the reconstruction of the original, whereby the immanent rule from Beckett's original is made explicit but switched off as comment lines; and it proceeds to the extrapolation of these rules, which are now made explicit, and their extension. This extension serves as a proof for the comprehension of Beckett's principle. The fact that this form of reconstruction is possible, thus, third, supports Jessica Pressman's thesis that digital literature returns to the operations of the historical avant-gardes, but implements them—as digital modernism—with more appropriate means and more consistently. Finally, Montfort also suggests that Beckett's Watt is *itself* algorithmic, a proto-digital literature. In that *Megawatt* not only emulates *Watt*, but in a sense explodes it, not only imitates, but exaggerates it, it also highlights those parts of *Watt* that are most apt for digital exploration, and does so also in a hermeneutically profitable way.

Megawatt is a recent example of the *sequential paradigm* as the oldest type of generative and in fact digital literature as such. I have spent some time on its code to illustrate exactly how well—by inspecting the source code—we can get a sense of its inner workings: each step of its sequence is laid out in front of us.

In contradistinction to the sequential paradigm, I would like to call the newest type of generative art the *connectionist paradigm*. Here we turn to works in the mold of Edmond de Belamy and the text generators GPT-2 and GPT-3. By connectionist, I refer to deep neural nets as the most widespread machine learning technology.[24] Neural nets follow, at least on a very basic and simplified level, the

24. For easy-to-follow introductions to this technology, see Ethem Alpaydin, *Machine Learning: The New AI* (Cambridge, Mass.: MIT Press, 2016); Pedro Domingos, *The Master*

DOI: 10.5451/unibas-ep78149

logic of the network of connections between neurons and synapses in the brain (hence "connectionist"). Incidentally, the first neural net also goes back to the time of Anders and Lutz, when in 1958 Frank Rosenblatt created the "perceptron"—modeled on the optical nerve rather than on the brain itself—which was capable of learning and recognizing basic patterns.[25] At its most abstract, a neural net is made up of three main elements: The input layer, one or more hidden layers, and the output layer. In Rosenblatt's model, there was only one hidden layer, but modern deep neural networks are composed of a multitude of hidden layers made up of neurons and connected by synapses, whose "weights" define the effect on the next neuron. The goal of a neural net is to create a function that fits the input data onto a desired output; the resultant model can be used to create outputs that resemble the inputs. The central point, however, is that a neural net cannot be *explicitly programmed* in the strict sense. Rather, neural nets *learn implicitly* by a repeated process of comparing input and output and adjusting for the errors in each iteration. Thus, there is no code we could inspect, only a list of numbers representing the structure of the network and their weighted connections; such a list, however, is incredibly difficult to interpret—this is the famous "black box" problem of neural nets (fig. 6).[26]

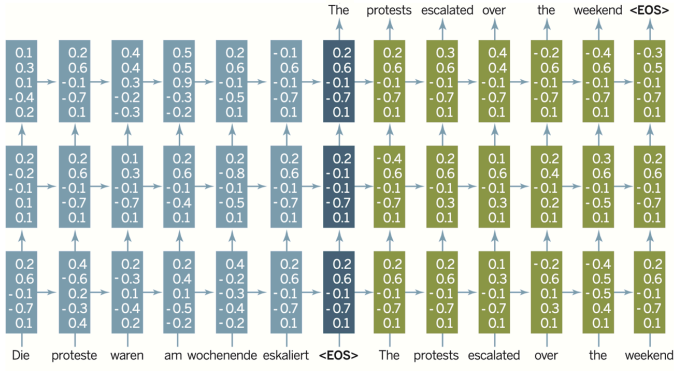


Fig. 6: Hirschberg, Julia, and Christopher D. Manning. "Advances in Natural Language processing." *Science* 349, no. 6245 (2015): 261–66.

Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World (New York: Basic, 2015); Melanie Mitchell, *Artificial Intelligence: A Guide for Thinking Humans* (New York: Farrar, Straus, and Giroux, 2019).

25. See Frank Rosenblatt, "The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain," *Psychological Review* 65, no. 6 (1958): 386–408; Nils J. Nilsson, *The Quest for Artificial Intelligence: A History of Ideas and Achievements* (Cambridge: Cambridge University Press, 2010): 64–74.

26. Davide Castelvecchi, "The Black Box of AI," *Nature* 538, no. 6. Oktober (2016): 20–23.

DOI: 10.5451/unibas-ep78149

▪▪

The Portrait of Edmond de Belamy is an example of the connectionist paradigm: trained on a dataset of 15,000 portraits from the 14th to the 19th century, the neural net produced an output that statistically resembles the works of the training set. Since the basic operation is to fit an input onto an output, neural nets have so far mostly been used for *re*-producing the stylistic characteristics of the training set—in this, they are not unlike *Megawatt*—but without the possibility of explicitly defining the rules by which this happens. And yet, repetition is in the very nature of neural nets, so that their designers must make an effort to avoid the phenomenon of overfitting, in which not similar but the exact *same* output is repeated; usually, this is done by either introducing noise or by reducing the completeness of the training set. In Edmond de Belamy's case, it seems that the training was aborted before the resemblance to the inputs became too strong, which gives the portrait its spectral quality.

In AI literature, we can observe similar effects, which are brought about by the failure of proper semantic understanding on the part of the model. Almost canonical already is "Sunspring" from 2016 by Ross Goodwin, an AI-generated film script which was subsequently professionally produced. Goodwin trained a neural net called "Benjamin" on over 300 science fiction film scripts, and had it output a new one. While GPT-3's proprietary model can produce impressively coherent text, most homebrewed models still remain restricted by limited network sizes and training sets. Likewise, in its juxtaposition of incongruent elements, the result of "Sunspring" tends toward the absurd with stage directions like: "He picks up a light screen and fights the security force of the particles of a transmission on his face."[27] As with most of neural net literature, we have to assume that a good deal of manual

▪▪▪

27. "INT. SHIP
We see H pull a book from a shelf, flip through it while speaking, and then put it back.
H
In a future with mass unemployment, young people are forced to sell blood. That's the first thing I can do.
H2
You should see the boys and shut up. I was the one who was going to be a hundred years old.
H
I saw him again. The way you were sent to me... that was a big honest idea. I am not a bright light.
C
Well, I have to go to the skull. I don't know.
He picks up a light screen and fights the security force of the particles of a transmission on his face." Benjamin, "Sunspring," https://www.docdroid.net/lCZ2fPA/sunspring-final-pdf.

DOI: 10.5451/unibas-ep78149

editing went into this process—but we cannot know for sure, as there is no code as in the case of *Megawatt* that we could study. The work remains not only as obscure as the proverbial black box, but also as intransparent as the mind of the genius of old.

3. Toward a Critique of AI Works

There are an important number of differences between the sequential paradigm of generative literature that employs linear algorithms, and the connectionist paradigm that is based on neural nets; these differences may allow us to approach a critique of AI works that does not simply compare them to human works.

The first difference is that a classic algorithm needs explicitly stated procedural rules, while a neural net learns by example and its rules of generation are not immediately visible. While Montfort could select the number of words and their possible position in a sentence, no such choices informed the production of "Sunspring's" script. Rather, it is generated via the neural net's training process, which is based on an input data set. The first paradigm functions top-down, the second bottom-up; for one, explicit rules stand at the beginning, for the other, implicit rules (the statistical model) are generated by the end. The classic algorithm functions deterministically, where an identical initial state always produces an identical final state; neural nets, however, work by statistical induction, which by its very nature is fuzzy. A neural net would have a much harder time reconstructing *Watt* in the way *Megawatt* did.

From here follows the second point: For the sequential paradigm, explicit rules and the deterministic process allow for a higher degree of transparency. Most obviously, the code itself is readable, but maybe more importantly, it is also easy to infer the underlying rules by running the program a couple of times and observing the output. This is much harder when it comes to neural nets, whose inner workings may not be impossible to retrace—"explainable AI" is working on this—but, as complex statistical models, cannot simply be reduced to explicitly stated rules. Likewise, observing the output may be able to give *some* clue as to the internal process, but will not allow for the same precision of inference.

This problem is exacerbated, third, for while linear algorithms draw a stark distinction between program and data, between procedural rules and items in a database, the "knowledge" in a neural net is not localized in some particular place. Rather, data and "program" are distributed throughout the whole system as a statistical dependency. If Montford could still build on lists of words, "Sunspring"—using an LSTM RNN type of network—is character based, so that no actual words

DOI: 10.5451/unibas-ep78149

are encoded in the model, just the likelihood of one character succeeding the next.[28] Instead of proceeding according to atomistic elements that assemble wholes from parts, neural nets have much stronger emergent properties that, put metaphorically, work according to a *Gestalt* logic.[29] Here, wholes are not simply reducible to their parts, but the training process allows the neural net to learn the overall shape of something *like* a painting in the style of 19th century mannerism, or the overall shape of something *like* a science fiction film.[30]

Lastly, and this is a somewhat controversial point introduced by German media theorist Andreas Sudmann: A linear algorithm, with its if-then-else conditions that can be diagrammed in a flow-chart, follows the digital logic of discrete states, of on and off, zero and one, *tertium non datur.* It is true that in neural nets the "neurons" in each layer are also either firing or not, but the weights that inhibit or amplify their activation are described through floating-point numbers "in an approximately analog, a quasi-analog way," as Sudmann puts it. If the connectionist paradigms is quasi-analog, it truly stands in the most extreme contrast to the sequential paradigm.[31] One does not have to follow Sudmann to this extreme, but what is clear is that there is a radical difference in the technical substance of both systems. This technical difference, I believe, must translate into a difference in the *aesthetic theorization* of such systems.

One approach to such an aesthetic critique of AI work would be to investigate in which ways the sequential and the connectionist paradigm relate to one of the oldest aesthetic concepts, that of *mimesis*. For both *Megawatt* and "Sunspring" follow a logic of imitation, but they do so in radically different ways: The former could be said to adhere to what German philosopher Hans Blumenberg has called imitation as *construction*—that is, the approximation of an existing state through the inference of the rules that bring it about. The latter then would rather enact the notion of *imitatio naturae*, the mere repetition of the real, without such procedural insight. For

28. See the influential blog post Andrew Karpathy, "The Unreasonable Effectiveness of Recurrent Neural Networks," *Andrej Karpathy Blog,* 2015, https://karpathy.github.io/2015/05/21/rnn-effectiveness/.
29. See Hannes Bajohr, "The 'Gestalt' of AI: Beyond Atomism and Holism," *Interface Critique* 3, no. 1 (2021), forthcoming.
30. This is extremely well illustrated with regard to poetry in Boris Orekhov and Frank Fischer, "Neural Reading: Insights from the Analysis of Poetry Generated by Artificial Neural Networks," *Orbis Litterarum* 75, no. 5 (2020): 230–46.
31. Andreas Sudmann, "Szenarien des Postdigitalen: Deep Learning als MedienRevolution," in *Machine Learning—Medien, Infrastrukturen und Technologien der Künstlichen Intelligenz*, ed. Christoph Engemann and Andreas Sudmann (Bielefeld: Transcript, 2018), 66.

DOI: 10.5451/unibas-ep78149

Blumenberg, both are distinctly connected to the question of the new: Construction indicates the possibility of going beyond the given by understanding the rules of its generation, as *Megawatt* demonstrates, and is thus decidedly modern—while the *imitatio naturae* relies on the mere repetition of the real, and belongs, Blumenberg holds, to an ancient aesthetic.[32] While I do not want to indicate that neural nets are somehow aesthetically premodern, I believe that the interplay between novelty and imitation needs to be posed in relation to this new technology.

Instead of pursuing this path, however, I shall focus on another possibility in comparing the sequential and the connectionist paradigm. It confronts the consequences of this distinction for media theory and both medi*um*- and medi*a*-specific analysis.

It is not a new observation—among others made by Rosalind Krauss, Florian Cramer, and Alan Liu—that the concept of "medium" traverses several disciplines that use it in distinct ways.[33] The two most important disciplines are *art history*—with "medium" as the singular and more often "mediums" as the plural—and *media theory*, including the digital humanities—with "medium" in the singular and "media" in the plural (although, as Alan Liu has noted, increasingly "media" is also used in the singular here).[34] The first use, in the meaning of artistic medium such as painting or sculpture, goes back to the eighteenth century, but its importance in the twentieth century is largely due to the influential art critic Clement Greenberg. Introducing the descriptive-normative term "medium specificity," he argued for the internal differentiation between different mediums.[35] Greenberg took up an idea of Gotthold Ephraim Lessing, who in his essay *Laocoon* had already advocated for the separation of the visual arts from literature according to their inherent structural logic: While literature in its linear textuality is inherently temporal, a series in time, and thus most apt to represent action, the visual arts deal with contiguous things in

32. Hans Blumenberg, "'Imitation of Nature': Toward a Prehistory of the Idea of the Creative Being," in *History, Metaphors, Fables: A Hans Blumenberg Reader*, ed. Hannes Bajohr, Florian Fuchs, and Joe Paul Kroll (Ithaca, NY: Cornell University Press, 2020): 316–57.
33. Rosalind Krauss, *"A Voyage on the North Sea": Art in the Age of the Post-Medium Condition* (London: Thames & Hudson, 1999); Cramer, "Nach dem Koitus oder nach dem Tod? Zur Begriffsverwirrung von 'Postdigital', 'Post-Internet' Und 'Post-Media,'" *Kunstforum International*, no. 242 (2016): 54–67; Alan Liu, *Friending the Past: The Sense of History in the Digital Age* (Chicago: The University of Chicago Press, 2018).
34. Liu, *Friending the Past: The Sense of History in the Digital Age*, 227 n18.
35. Clement Greenberg, "Avantgarde and Kitsch," in *Art and Culture: Critical Essays* (Boston: Beacon, 1989), 3–21; Clement Greenberg, "Towards a Newer Laocoon," in *The Collected Essays and Criticism*, vol. 1: 1939-44 (Chicago: The University of Chicago Press, 1986): 23–38.

DOI: 10.5451/unibas-ep78149

space, that is, extension, and thus are better suited for representing objects.[36] Greenberg extends this argument to the mediums of the visual arts themselves, and suggests that modernism engages in a complete separation of painting from sculpture. For him, as for Lessing, the extent to which a work of art highlights the specific structural characteristics of its medium is a measure of its artistic purity. And while Greenberg originally only wanted to show a process of historical differentiation, medium-specificity eventually took normative rank.[37] Thus, if what distinguishes painting from other mediums is two-dimensionality—flatness—then those paintings are the purest that are the flattest, i.e., those that lack spatial illusion. Three-dimensionality belongs not to paintings, but to sculpture.

The second use of the term "medium"—in the meaning of a channel of communication—is mostly connected to a normally unnoticed but determinative carrier of information, as it was introduced by Marshall McLuhan into media theory. While McLuhan defined media as human extensions, he nevertheless confined himself to mass media and electronic media in a narrower sense.[38] Contemporary media theory has a tendency to overextend the use of the word to just about anything that acts as intermediary between two realms. Because of this, media's Protean nature has fostered the conflation of mediums and media. But there might be good reason to avoid this confusion, or at least to insist on the particularity of each media. Katherine Hayles has coined the term "media-specificity" in her essay, "Print is Flat, Code is Deep," which already in its title is a nod to Greenberg even though his name is never mentioned.[39] Media specific analysis, according to Hayles, means insisting on the materiality of media. In terms of digital literature, it entails the acknowledgement that electronic works—in contradistinction to print books—have surface texts, but also underlying code that shapes the surface text.

Yet for Hayles the contrastive foil to electronic textuality is still the printed book—electronic and non-electronic literature remain the two main operative categories. The distinction between the sequential and the connectionist paradigm

36. Gotthold Emphraim Lessing, *Laocoon: An Essay upon the Limits of Painting and Poetry*, trans. Ellen Frothingham (Mineola: Dover, 2005), chap. 15 and 16.
37. That this normative interpretation is also a result of the reception of his work, particularly by his pupil Michael Fried, Greenberg remarked with some annoyance, see Thierry de Duve, *Clement Greenberg Between the Lines: Including a Debate with Clement Greenberg* (Chicago: The University of Chicago Press, 2010), 147–48. I thank Colin Lang for pointing this out to me.
38. Marshall McLuhan, *Understanding Media: The Extensions of Man* (Cambridge, Mass.: MIT Press, 1994).
39. N. Katherine Hayles, "Print Is Flat, Code Is Deep: The Importance of Media-Specific Analysis," *Poetics Today* 25, no. 1 (2004), 67–90.

DOI: 10.5451/unibas-ep78149

indicates, however, that a further *internal* differentiation is necessary, just as Greenberg extended Lessing's division between literature and the visual arts to further subdivide the latter. Let me give just two examples of the necessity of this further categorization: With the rise of the connectionist paradigm it no longer makes sense to speak of what Lev Manovich, in *The Language of New Media*, has called the "database logic" of electronic textuality, in which each item has the same significance as any other.[40] When there are no explicitly encoded items anymore that can be accessed individually, but only statistical dependencies that are distributed throughout the system, we are confronted not with a database logic but with something else entirely. Likewise, the distinction between "texton" and "scripton" introduced by literary scholar Espen Aarseth—that is, a string as it appears in the output, such as on a screen, and a string as it appears in the code and that may be instantiated differently—may lose its usefulness if "textons" are no longer to be located in *any* code, indeed, if there is *no* code anymore in the traditional sense.[41] The connectionist paradigm shakes some of the basic ways we think about electronic textuality in general, and digital literature in particular. The metaphor of "depth" and "surface" on which Hayles relied and which still implied the possibility of connecting the latter to the former needs to be rethought.

In the remainder of this essay, I want to concentrate on the implications for assessing digital literary works. While Hayles' *media*-specificity forgoes the normative slant of Greenberg's *medium*-specificity, and only describes a way of analysis that takes the particulars of a media into account, I think it might be useful to rekindle some of that normativity. *Megawatt*'s significance rests partly in the way that its underlying structure, the linear algorithm, reflects the structure of the resultant text so well. With the connectionist paradigm, a new form of visual and textual art is emerging, and it is not yet clear what it might be capable of. But because this is so, the aesthetic *critique* of such works may wish to pay special attention to those that investigate the specificity of their medium, in both senses of the word.

Let me try to give an example of this thought. A reader of this essay may have wondered why in a text about digital literature I have also referred to visual works so much. By so doing, I have hinted at the capabilities of the same media—neural nets—working on different mediums—text and images (and here I go back to

40. Lev Manovich, *The Language of New Media* (Cambridge, Mass.: MIT Press, 2000), 218.
41. Espen J. Aarseth, *Cybertext: Perspectives on Ergotic Literature* (Baltimore: The Johns Hopkins University Press, 1997), 62.

DOI: 10.5451/unibas-ep78149

Lessing rather than to Greenberg). One can differentiate even *further*, for not all neural nets are useful for all mediums. The most basic neural net architectures for generating images are convolutional neural networks, while recurrent neural networks are used for texts. They work in different ways due to the structure of what they produce; they are, in a way, different *media* generating different *mediums* (fig. 6).

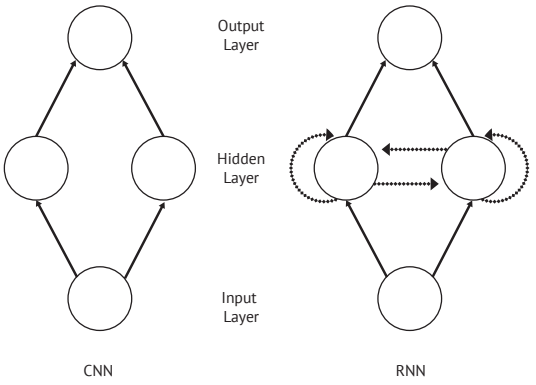


Fig. 7: Convolutional Neural Network (left), Recurrent Neural Network (right) adapted from Melanie Mitchell, *Artificial Intelligence: A Guide for Thinking Humans*. New York: Farrar, Straus, and Giroux, 2019.

At the most basic level, digital images are continuous in two dimensions. Their smallest unit is the pixel with a color value arranged in a matrix that remains static over the data set. The relationship between pixels is based on correlation by proximity—the closer two pixels are to each other, the likelier they stand in a meaningful relationship to each other in forming wholes. A convolutional neural net uses this logic of continuity in a bottom-up process to extract features in this pixel matrix by tasking each of the hidden layers with extracting the salient patters of its previous input. Since this happens progressively between layers, there is a process of abstraction at work here. The first layer may look at a combination of a few pixels, and then pass on the result to the next layer, which now looks at a combination of a combination of pixels, and so on. Thus, there is a progression from edges to simple shapes to objects, and so on (fig. 8).

DOI: 10.5451/unibas-ep78149

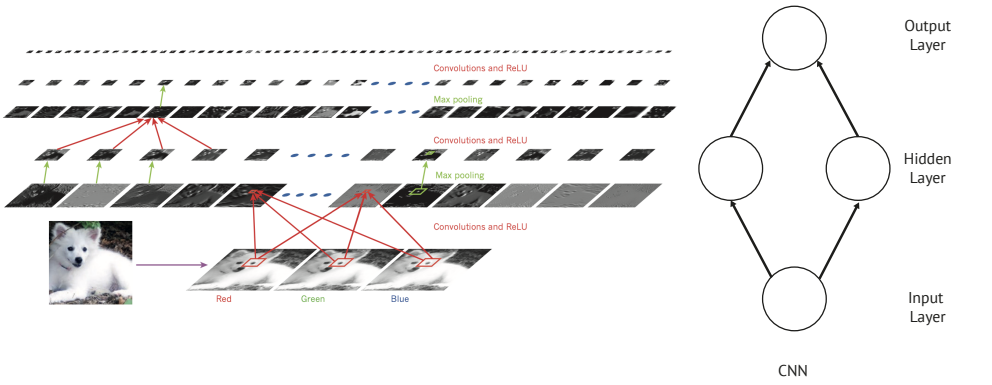


Fig. 8: Convolutional Neural Network: Yann LeCun, et al. "Deep Learning." *Nature* 521, May (2015): 436–44.

Text, on the other hand, requires a different way of proceeding. It is not continuous in two dimensions with equal basic units. Rather, text is continuous in one dimension with its basic units being characters. Ignoring the question of meaning for a moment here, as well as the fact that the "value" of a character is in no way comparable to the "value" of a pixel, this difference in dimensionality requires neural nets dealing with text to have a different structure (fig. 9). Recurrent neural networks need to "remember" previous characters to build complex statistical models about their likely occurrence, which is why the neurons of its networks are connected not only to the next layer but also to themselves (this is the network type used for "Sunspring"). A convolutional neural network usually does not deal with text, while a recurrent one usually is not used for images; here, medium and media are correlated.

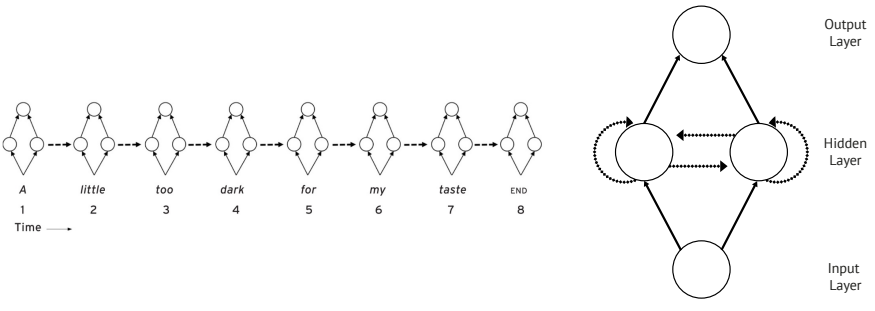


Fig. 9: Recurrent Neural Network. Melanie Mitchell. *Artificial Intelligence: A Guide for Thinking Humans.* New York: Farrar, Straus, and Giroux, 2019.

DOI: 10.5451/unibas-ep78149

While the reality of neural networks is infinitely more complex—GPT-2 and 3 are transformer models that work still differently—this dichotomy is nevertheless useful for offering more finely-grained criteria of media-aesthetical judgment. Following this line of thought, digital literature and art can be discussed along two axes: in terms of their *media*-specificity—as the awareness of their technical structures and affordances—but also according to their *medium*-specificity—as the awareness of the internal artistic logic of the medium within which it works. In the sequential paradigm, *Megawatt* is an example of a parallelism of both: the structure of the media, the linear algorithm, reflects the structure of the medium, modernist literature, rather well. But these axes need not be parallel to each other.

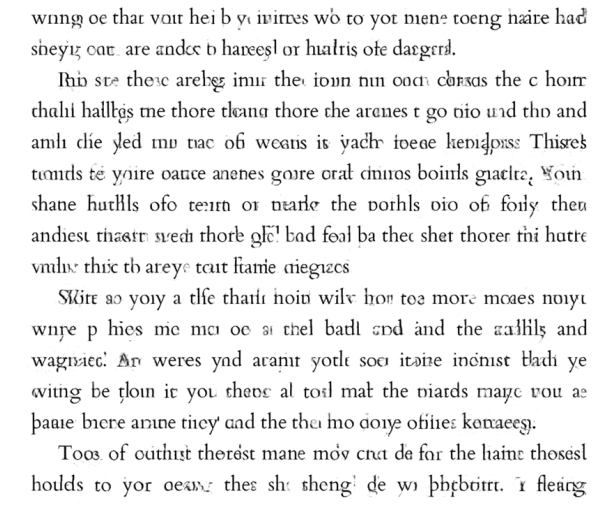


Fig. 10: Allison Parrish, *Ahe Thd Yearidy Ti Isa* (2019)

A clever illustration of the interplay of the medium/media axes in the connectionist paradigm is Allison Parrish's unpronounceable *Ahe Thd Yearidy Ti Isa* (2019).[42] It operates by a willful confusion: It treats text *as* image, reverses the appropriate neural net architectures, and plays with the asemic effects this technological and semiotic category mistake engenders. Parrish used a specific type of convolutional neural network, called GAN or generative adversarial network, that has been extremely successful in generating images.[43] Its architecture splits the

42. Allison Parrish, "Ahe Thd Yearidy Ti Isa (asemic GAN-generated novel)," https://github.com/NaNoGenMo/2019/issues/144. (The novel was an entry to 2019's *National Novel Generation Month*).
43. Ian J. Goodfellow et al., "Generative Adversarial Networks," *Advances in Neural*

DOI: 10.5451/unibas-ep78149

production and the assessment of its output into two separate processes. While the generator generates images, the discriminator is tasked with judging how close these images come to the expected output. In this case, the GAN was fed bitmap images of words. Here lies the category mistake: Bitmaps of words are human-readable, but not machine-readable; they do not register as text. Thus, the processable information of the image is not identical to the information that the depicted word represents: its technical materiality is separated from its signifying function. The GAN treats words *as* images, and that means, *not differently from* images; thus, the discriminator cannot, like an RNN would, compare a string of discrete characters, but only statistical distributions of pixel values. The result looks like text to the discriminator, but lacks any semantic or symbolic value, so that Parrish can speak of its product as an "*asemic* novel." It represents a non-human type of reading—a probabilistic reading: text-as-images seen through the eyes of a machine. And in a final twist, as if to comment on the futility of the whole process, Parrish uses the "correct" image-to-text process. After all, the book *does* have a title—*Ahe Thd Yearidy Ti Isa*. To create it, Parrish ran the title "image" through a character recognition algorithm that converts bitmaps into text—properly this time, and even if the result is still nonsensical, this nonsense is now indeed machine-readable (fig. 11).

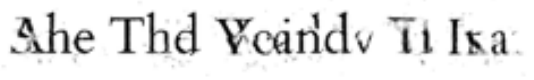


Fig. 11: Allison Parrish, *Ahe Thd Yearidy Ti Isa* (2019). Title page.

This game of multiple confusions and conversions draws attention to the difference of text and image as mediums by highlighting the media used for their processing. As asemic writing, Parrish's work operates at the border between literature and the visual arts, and deals in non-semantic but text-like structures; it is *medium*-specific precisely in refusing to carry meaning, and *media*-specific in reflecting this refusal on a technical level by using a convolutional neural network where a recurrent neural network would have been appropriate. This breaks the clear parallelism of *Megawatt*, and does so productively. What is more, it is a much more interesting use of neural nets with a much more complex notion of mimesis than both the easy absurdism of "Sunspring," or the naive imitationism of "Edmond de Belamy." In willfully confusing standard procedures, Parrish's work allows the

Information Processing Systems, no. January (June 10, 2014): 2672–80.

DOI: 10.5451/unibas-ep78149

only type of algorithmic empathy neural nets still allow—not laying bare the underlying concept, but at least offering a glimpse at the otherwise inscrutable process through tactical, and in the end illuminating category mistakes. *Ahe Thd Yearidy Ti Isa*, then, does not give into the Promethean Anxiety, but offers a non-anthropocentric use of AI beyond mere comparison to conventional, "human" works.

For a critique of AI works, which is still a desideratum, this investigation into the *inherent* possibilities and limitations of a new medium may offer a normative example. The problem with digital AI works that simply simulate "human" works is not so much that they are mere derivatives, simulations of already existing but "analog" schemes. Rather, in insisting on the human comparison, they restrict from the outset what can be done in this new medium instead of exploring its affordances. In this sense the plea for a medium- and media-specificity is meant to be purely corrective. Not every difference in production needs its own form of criticism; but where the form of criticism itself remains undeveloped, in that it views digital works according to the standards of computerized "geniuses," the concentration on the medium is at least one way to do justice to the actual novelty of the works.

To be sure, this suggestion has a temporal core and treats works of this type as pioneering, and fulfilling an avant-garde function. It implies that once this exploration has been exhausted, these artworks have satisfied their heuristic task—to give way to a new type of literature that can freely make use of the insights gained, and can even turn away from media- and medium-specificity.[44] Yet in order to get to this place, I believe, we do well in taking Hayles' appeal seriously. Focusing on the materiality of the connectionist paradigm—even through paradox, as in Parrish's case—can be an inspiration both for the analysis as well as the production of contemporary digital literature.

44. For some of the points in these last two paragraphs, I am indebted to Annette Gilbert and Michel Chaouli.

DOI: 10.5451/unibas-ep78149

References

Aarseth, Espen J. *Cybertext: Perspectives on Ergotic Literature*. Baltimore: The Johns Hopkins University Press, 1997.

Alpaydin, Ethem. *Machine Learning: The New AI*. Cambridge, Mass.: MIT Press, 2016.

Anders, Günther. "On Promethean Shame." In Christopher J. Müller, *Prometheanism: Technology, Digital Culture and Human Obsolescence*, 29–95. London: Rowman & Littlefield, 2016.

Bajohr, Hannes. "Das Reskilling der Literatur." In *Code und Konzept: Literatur und das Digitale*, edited by Hannes Bajohr, 7–21. Berlin: Frohmann, 2016.

Bajohr Hannes. "The 'Gestalt' of AI: Beyond Atomism and Holism." *Interface Critique* 3, no. 1 (2021), forthcoming.

Beals, Kurt. "'Do the New Poets Think? It's Possible': Computer Poetry and Cyborg Subjectivity." *Configurations* 26, no. 2 (2018): 149–77.

Beckett, Samuel. *Watt*. New York: Grove Press, 1970.

Benjamin, "Sunspring," 2016, https://www.docdroid.net/lCZ2fPA/sunspring-final-pdf.

Ben-Zvi, Linda. "Samuel Beckett, Fritz Mauthner, and the Limits of Language." *PMLA* 95, no. 2 (March 1980): 183.

Bernhart, Toni. "Beiwerk als Werk: Stochastische Texte von Theo Lutz." *editio* 32 (2020): 180–206.

Blumenberg, Hans. "'Imitation of Nature': Toward a Prehistory of the Idea of the Creative Being." In *History, Metaphors, Fables: A Hans Blumenberg Reader*, edited by Hannes Bajohr, Florian Fuchs, and Joe Paul Kroll, 316–57. Ithaca, NY: Cornell University Press, 2020.

Bogost, Ian. "The AI-Art Gold Rush Is Here." *The Atlantic*, no. March 6 (2019). https://www.theatlantic.com/technology/archive/2019/03/ai-created-art-invades-chelsea-gallery-scene/584134/.

Brown, Tom B., Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, et al. "Language Models Are Few-Shot Learners." *ArXiv*, May 28, 2020. http://arxiv.org/abs/2005.14165.

Castelvecchi, Davide. "The Black Box of AI." *Nature* 538, no. 6. Oktober (2016): 20–23.

Cramer. "Nach dem Koitus oder nach dem Tod? Zur Begriffsverwirrung von 'Postdigital', 'Post-Internet' Und 'Post-Media.'" *Kunstforum International*, no. 242 (2016): 54–67.

Cramer, Florian. *Words Made Flesh: Code, Culture, Imagination*. Rotterdam: Piet Zwart Institute, 2005.

de Duve, Thierry. *Clement Greenberg Between the Lines: Including a Debate with Clement Greenberg*. Chicago: The University of Chicago Press, 2010.

Du Sautoy, Marcus. *The Creativity Code: How AI Is Learning to Write, Paint, and Think*. London: Fourth Estate, 2019.

DOI: 10.5451/unibas-ep78149

Dennis, Amanda M. "Glitches in Logic in Beckett's *Watt*: Toward a Sensory Poetics." *Journal of Modern Literature* 38, no. 2 (2015): 103–16.

Domingos, Pedro. *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World*. New York: Basic, 2015.

Goodfellow, Ian J., Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. "Generative Adversarial Networks." *Advances in Neural Information Processing Systems*, January (June 10, 2014): 2672–80.

Greenberg, Clement. "Avantgarde and Kitsch." In *Art and Culture: Critical Essays*: 3–21. Boston: Beacon, 1989.

———. "Towards a Newer Laocoon." In *The Collected Essays and Criticism*, 1: 1939-44: 23–38. Chicago: The University of Chicago Press, 1986.

Hayles, N. Katherine. *Electronic Literature: New Horizons for the Literary*. Notre Dame: University of Notre Dame Press, 2009.

———. "Print Is Flat, Code Is Deep: The Importance of Media-Specific Analysis." *Poetics Today* 25, no. 1 (2004): 67–90.

Karpathy, Andrew. "The Unreasonable Effectiveness of Recurrent Neural Networks." *Andrej Karpathy Blog*, 2015. https://karpathy.github.io/2015/05/21/rnn-effectiveness/.

Kenner, Hugh. *The Mechanic Muse*. New York: Oxford University Press, 1987.

Krauss, Rosalind. *"A Voyage on the North Sea": Art in the Age of the Post-Medium Condition*. London: Thames & Hudson, 1999.

Lessing, Gotthold Emphraim. *Laocoon: An Essay upon the Limits of Painting and Poetry*. Translated by Ellen Frothingham. Mineola: Dover, 2005.

Liu, Alan. *Friending the Past: The Sense of History in the Digital Age*. Chicago: The University of Chicago Press, 2018.

Manovich, Lev. *The Language of New Media*. Cambridge, Mass.: MIT Press, 2000.

McCormack, W. J. "Seeing Darkly: Notes on T. W. Adorno and Samuel Beckett." *Hermathena*, no. 141 (1986): 22–44.

McLuhan, Marshall. *Understanding Media: The Extensions of Man*. Cambridge, Mass.: MIT Press, 1994.

Miller, Arthur I. *The Artist in the Machine: The World of AI-Powered Creativity*. *The Artist in the Machine*. Cambridge, Mass.: MIT Press, 2019.

Mitchell, Melanie. *Artificial Intelligence: A Guide for Thinking Humans*. New York: Farrar, Straus, and Giroux, 2019.

Montfort, Nick. *Megawatt*. Cambridge, Mass.: Bad Quarto, 2014.

Nilsson, Nils J. *The Quest for Artificial Intelligence: A History of Ideas and Achievements*. Cambridge: Cambridge University Press, 2010.

DOI: 10.5451/unibas-ep78149

Orekhov, Boris, and Frank Fischer. "Neural Reading: Insights from the Analysis of Poetry Generated by Artificial Neural Networks." *Orbis Litterarum* 75, no. 5 (2020): 230–46.

Pressman, Jessica. *Digital Modernism: Making It New in New Media*. Oxford: Oxford University Press, 2014.

Radford, Alec, Jeffrey Wu, Dario Amodei, Daniela Amodei, Jack Clark, Miles Brundage, and Ilya Sutskever. "Better Language Models and Their Implications." *OpenAI Blog*, 2019. https://openai.com/blog/better-language-models.

Rettberg, Scott. *Electronic Literature*. Cambridge: Polity, 2019.

Rosenblatt, Frank. "The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain." *Psychological Review* 65, no. 6 (1958): 386–408.

Sudmann, Andreas. "Szenarien des Postdigitalen: Deep Learning als MedienRevolution." In *Machine Learning—Medien, Infrastrukturen und Technologien der Künstlichen Intelligenz*, edited by Christoph Engemann and Andreas Sudmann, 55–73. Bielefeld: Transcript, 2018.

Weller, Shane. "Humanity in Ruins: Samuel Beckett." In *Language and Negativity in European Modernism*, 90–125. Cambridge: Cambridge University Press, 2018.

DOI: 10.5451/unibas-ep78149