MALMÖ UNIVERSITET

Neural Novelty

How Machine Learning Does Interactive Generative Literature

Love Lagerkvist

Interaction Design
Two-Year Master
15 Credits
Spring Semester 2020
Supervisor: Anne-Marie Hansen

Abstract

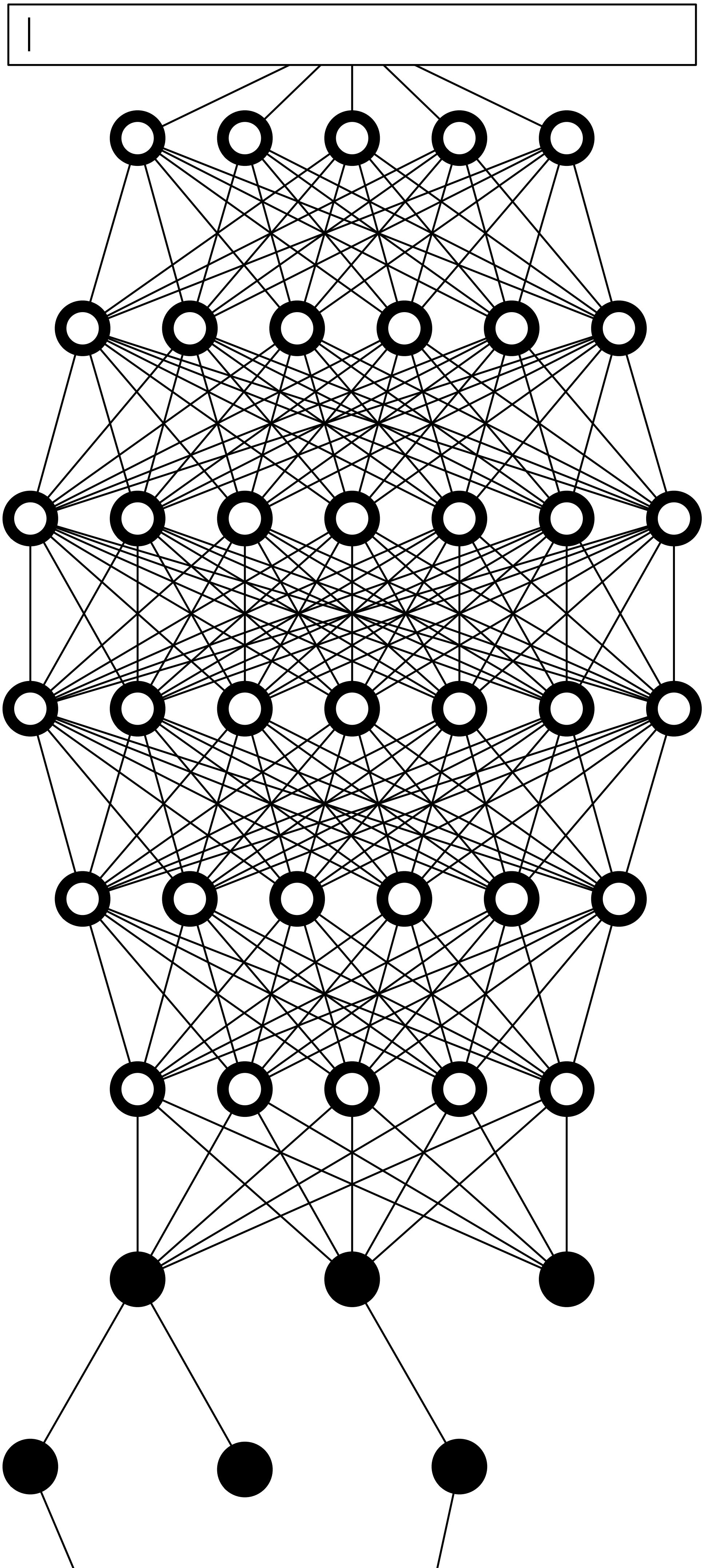
Every day, machine learning (ML) and artificial intelligence (AI) embeds itself further into domestic and industrial technologies. Interaction designers have historically struggled to engage directly with the subject, facing a shortage of appropriate methods and abstractions. There is a need to find ways though which interaction design practitioners might integrate ML into their work, in order to democratize and diversify the field. This thesis proposes a mode of inquiry that considers the interactive qualities of *what machine learning does*, as opposed the technical specifications of what machine learning *is*. A shift in focus from the technicality of ML to the artifacts it creates allows the interaction designer to situate its existing skill set, affording it to engage with *machine learning as a design material*. A Research-through-Design process explores different methodological adaptions, evaluated through user feedback and an-in depth case analysis. An elaborated design experiment, *Multiverse*, examines the novel, non-anthropomorphic aesthetic qualities of generative literature. It prototypes interactions with bidirectional literature and studies how these transform the reader into a cybertextual “user-reader”. The thesis ends with a discussion on the implications of machine written literature and proposes a number of future investigations into the research space unfolded through the prototype.

If the writer is a sorcerer it is because writing is a becoming

Deleuze-Guattari

In the multiplicity of writing, everything is to be disentangled, nothing deciphered; the structure can be followed, “run” (like the thread of a stocking) at every point and at every level, but there is nothing beneath: the space of writing is to be ranged over, not pierced

Roland Barthes



Introduction

Interaction designers are no strangers to the seemingly infinite hype surrounding artificial intelligence (AI). As cottage industries get founded on public confusion and insecurity, springing up on what appears to be a daily basis (Kelly, 2017), world governments leverage advanced computer vision and big data to further accelerate mass surveillance (Kuo, 2020). As the technologies embed themselves deeper into everyday life, we keep talking about them in terms of “magic” and other immaterial potentials. Historian Rasmus Fleischer likens this type of vague discourse, centered around a floating signifier rather than the material conditions, to the discussion of *the machine* that occupied interwar western society (Fleischer, 2018). The technologies of AI/*the machine* move from the actual into the magical, fostering a discussion either of reactionary fear or blind techno-optimism. Studies have shown that interaction designers relate to the subject in much the same mode, repeating statements of “magic” and “mystery” (Dove et al., 2017). There is little knowledge about the potentials of these new technologies, leading to sub optimal implementations or entirely missed opportunities (Yang, Scuito, et al., 2018). In order to have a productive, critical debate and practice, I argue that thinkers and designers must necessarily engage with the real, material conditions of artificial intelligence and its different applied forms, like machine learning (ML). This way in which we a accomplish this is though two main concepts: The designerly engagement with *what machine learning does* and *the use of machine learning as a design material*.

The rapid application and development of machine learning in the field of natural language processing (NLP) affords anyone with a computer and some basic domain knowledge to leverage sophisticated artificial neural networks which are able to generate bodies of text at a level almost indistinguishable human performance (Quach, 2019). The packaging of pre-trained model into accessible software libraries (Wolf et al., 2019) or services (Cloud, 2020) makes it possible for a wider set of practitioners to engage with this technology, a development this thesis argues is vital for the field to grow and, more importantly, diversify (Leavy, 2018). As technologists in general and system designers in particular, it is important that interaction designers to engage with, and diversify (Leavy, 2018), emergent technologies relevant to our field. We might in turn build new kinds of applications that afford an even wider audience engagement with these tools, repeating the process.

In order to effectively act in such a vast and complex field as machine learning, I have situated the research of the thesis in the prototyping of novel interactions with generative literature, as natural language processing has proven to be one of the more exiting applications of ML technology (Y. Liu & Lin, 2019). I understand *literature* in the broad sense of the term, as any collection of written artifacts, with the recognition that such definition might be insufficient or controversial in relation to literary theory proper (Eagleton, 2011).

The result of the thesis is not a specific design artifact, but the development of methods and abstractions through which interaction designers can work with the new generation of generative technologies, currently being applied in areas such as architecture (Spacemarker, 2020) and media production (Inc., 2020). Methodological explorations must necessarily include evaluation of their application, here manifested in an elaborated prototype of an interactive system for generative literature (Section 4.5).

The task of approaching machine learning as a design material is thankfully not without precedence. Rich studies on the use of computation (Landin, 2005) and Bluetooth as materials (Sundström et al., 2011) are open for reference. Landin's engagement with the aesthetics of computation is especially relevant, as she approaches the subject not as a "neutral technical medium", but an "expressive design material", a sentiment I argue applies equally to technologies of machine learning. She identifies *fragility* and *magic* as material properties of computation, a radical gesture that goes beyond the technical specifications that otherwise premeditates its discussion. These attributes are often interpreted as negative, but Landin re-appropriates them, imbuing these terms with agency of their own. In this thesis I aim to perform the same move, through the shift of concern from what a machine learning technically *is* to what it machine learning expressively *does*. I argue for an open, non-speciesist view that considers the sometimes odd and in-human outcomes of literary generation as an essential quality of a new form of expression, in line with the tradition of postmodern literature and stochastic art. The properties and attributes of computational materials affords the system emergent behavior without human intervention, a phenomenon I argue expresses a novel of form of cybernetic creative agency.

1.1 Research Focus

By 2018, there where only 9 DIS papers that mentioned "machine learning" (Yang, Banovic, et al., 2018). Contrasted to claims that "ML is the new UX" (Yang et al., 2016), the need for interaction design researchers to engage with and diversity the field becomes very apparent. This thesis explores approaches that might afford engagement with machine learning, through the notion of *machine learning as a de-*

sign material (Dove et al., 2017). A response to the call of Dove et al. to explore ways in which we as interaction designers can work and prototype with machine learning and related technologies, it acts in the tradition of interaction design research as a fertile ground for exploring new notions and materials (Buxton, 2010), through experimentation and drifting (Krogh et al., 2015). By "reflecting in action", designers get a tacit understanding of materials that are otherwise foreign (Yang, Banovic, et al., 2018). This mode of inquiry is summarized as concern of what machine learning *does*, a contrast to most literature on the subject, which is largely technical in nature and pertains to what machine learning *is*.

The research in this thesis does not investigate ways by which interaction designers might write or improve the underlying machine learning algorithms or systems. I consider the designer a subject who assembles and applies technologies, and as such I have chosen to explore the tools that enables such synthesizing engagement (Billewicz, 2017). This might entail direct material engagement with machine learning technology in the form of pre-packaged software libraries, possible interactive interfaces with artificial intelligence or, as in this thesis, methods that appropriate and situate existing interaction design practice in machine learning contexts.

Artificial intelligence and machine learning are huge subjects, with histories spanning almost a century. It is would be impossible for a lone researcher to take on an investigation on the field as a whole. As such, I have limited to the investigation of possible interactions with generative literature, understood as any system that creates or manipulates literary texts by means through the application of set rules or an algorithm. In this specific case, the subject that generates the textual artifacts is a state-of-the-art artificial neural network, like OpenAI's GPT-2 (Radford et al., 2019) or Google Corporation's BERT (Devlin et al., 2018). This new generation of model come "pre-trained", meaning that they can be used out of the box, with little-to-no additional work or programming required. It is the ambition of this research to exploit this approachability as a tool through which it is possible to explore novel interactions artificial neural networks. Through documentation and reflection of a bottom-up approach, I hope to be able to extrapolate new ways of sketching with machine learning as a design material.

Of the 9 DIS papers that mentioned machine learning, none touched on the interactive aspects to generative literature. A majority of the generative literature that can be found online is written it in the same static, read-only mode as conventional literature (van Stegeren & Theune, 2019). Situating the research in an impoverished area like interactive literature will hopefully foster interest in a topic, which I believe will grow all the more relevant as these technologies get more and more sophisticated.

A majority of the research into machine learning/natural language processing pursuits an ever higher fidelity in the generation of text, with an ideal of human-like performance (Devlin et al., 2018). Such a task is better suited for a researcher in computer science, not a designers who assembles and applies. Instead, through a situation in *what ML does*, I take the stance that there is an unique aesthetic quality in the literature generated by artificial neural networks, which is fundamentally different to the those written by a human agent. New design materials require new understandings of their aesthetics. By attempting an exploration of generative literature through machine learning in a mode free of speciesism, it might be possible to identify novel qualities one would otherwise discard as faulty.

1.2 Research Questions

1.2.1 Primary Question

How can interaction designers work with ML as a design material?

The primary research question investigates the ways through which interaction designers can use machine learning as a design material (Dove et al., 2017). In this thesis this manifests partially as an investigation of suitable methodological approaches (both new and appropriated), partially as a survey of the new types of tools available to artificial intelligence practitioners. The research looks at machine learning in the limited context of natural language processing, specifically generative literature, in order to accommodate for the relatively short time frame of the project. Instead of proposing brand new tools, it investigates how one might go about in re-contextualizing existing methods in this new field by situating it in a creative and researched practice.

1.2.2 Secondary Question

Which novel interactions are possible with generative literature?

The secondary research question concerns the case interactive generative literature. The current state-of-the-art in natural language processing and machine learning afford cheap and sophisticated text generation to a large audience. With this democratization comes questions about how we relate to machine made art and how it differs from human authorship. Finding novel interactions for these types of systems might lead to better insight on the role they could play in the future. I explore this through a series of interactive prototypes that use bidirectional storytelling to highlight the shift generative literature crates in the reader-author relation. This is done through design choices that highlight the unique aesthetics of generative literature through machine learning, manifested in a prototype of an interactive system for generative literature (Section 4.5).

1.3 Knowledge Contribution

As machine learning graduates from niche field in computer since to one of the primary modes of computation (Shapiro et al., 2018), there is a need for a diversity in research and approaches. The lack of applied machine learning in contemporary interaction design education has been identified as one of the major issues for incorporating the technologies into wider practice (Dove et al., 2017). By conducting, documenting and sharing a project that deals with the technology in a limited scope and highly material fashion, I aim to produce both personal and collegiate knowledge. This goal is reflected in the form of the thesis as an object in itself, which assumes a reader with little to no previous knowledge of machine learning. A thorough glossary (Section 8) is provided and serves as a general, easy to digest reference on topics in the filed. Practice and de-mystification are essential in achieving an increased understanding of machine learning, and as a result, a more productive and critical discussion on the subject.

The thesis formulates an attempt to find tools for interaction designers to engage with machine learning, much due to necessity. Early in the process of this thesis I attempted, and failed, to apply traditional interaction design methods to the field of generative literature. The lack of tools that afforded me to engage with machine learning as a design material became painfully clear, which shifted my focus into methodological explorations. Instead of attempting a malformed traditional design project with machine learning, this research aims instead to contribute knowledge about possible ways to appropriate the interaction designer's existing toolbox in the domain of machine learning. The work situates itself in the relatively young tradition of interaction design inquiry about machine learning, following previous theses from the same institution, like Agnieszka Billewicz's *Study of a relationship: Designerly explorations of machine learning algorithms* (2017).

1.4 Nomenclature

This thesis avoids sweeping terms like *artificial intelligence* in an attempt to find a balance between the highly domain specific nomenclature and the language used in contemporary interaction design discourse (Yang, Banovic, et al., 2018). It primarily uses the term machine learning when taking about these technologies and their research field. Specific subtopic such as deep learning, now the primary approach for research in the area, but not a field in itself (Alpaydin, 2014), or specific architectures are not refereed to unless explicitly relevant to an understanding of the aesthetic or interactive qualities of machine learning systems. For an exhaustive list of the technical terms used in this thesis, consult the *Glossary* (Chapter 8). Included words are hyperlinked throughout the text. For a breakdown on artificial intelligence and how it is used in the thesis, see Section 2.1.

1.5 Ethical Concerns

The public debate around AI has gravitated towards two topics: Implicit bias in models, particularly concerning gender and race (Leavy, 2018), and the loss of employment opportunities as an effect of increased automation (Arntz et al., 2017). In the area of natural language processing, there is much concern about the potential misuse of human-level text generation in relation to fake news and forgery (Knight, 2019).

Such concerns are, necessarily, very much present in contemporary machine learning research. In the announcement of Open AI's GPT-2, the main machine learning model used in this thesis, the authors state their intention not to release the entire pre-trained model "due to our concerns about malicious applications of the technology". Instead, they opted to publish it as progressively larger chucks (Alec et al., 2019), carefully observing their application. GPT-2 is "casually trained" on a massive (in the billions) corpus mixed of source. This means that the model is capable of re-producing almost any kind of literary style, with nothing but a small prompt to kick it off. If one has a more specific use case, they can easily tailor the network via transfer learning, a type of training that leverages the connections already formed in causal training by fine tuning them to a much smaller, specialized data set. The researchers at Open AI touch on this in said announcement when assessing the possibilities of malicious applications, such as the production of highly specialized fake news. However, they argue that "the underlying technical innovations inherent to these systems are core to fundamental artificial intelligence research", instead urging policy makers to "consider introducing penalties for the misuse of such systems".

This thesis explores the potentially productive aspects of state-of-the-art machine learning, arguing that education on the material conditions that shape these systems is essential for an informed and productive debate on the subject. It acknowledges that the use of GPT-2 and other pre-trained models is potentially problematic, as the downstream practitioner has little insight in the data set used to train it. machine learning has been criticized for producing systems with inherent bias inadvertently inherited from its selection of training material (Zou & Schiebinger, 2018) (Leavy, 2018). Without direct insight in the training data set, it is impossible for the user to rule out such behaviors. Even though this research work deals with ML and generative literature as such, and not strictly a specific implementation or artificial neural network, it is always very important to stay aware and critical of the social contexts that situate the creation of technologies.

1.6 Sustainability

Before the proliferation of à la carte pre-trained model, users had to spend a lot of time and computing power going through the error-prone process of training a model before they could start using it in their projects. The large amounts of compute required to train machine learning models has been identified as be one of the main factors prohibiting interaction designers from working with machine learning as a design material (Yang, 2018). Such massive computation requires expensive and often specialized hardware that consumes a large amount of electrical power (García-Martín et al., 2019). The full GPT-2 model has an estimated training cost of $256 per hour, while the similar XL-Net model clocks in at a grand total around $61,000 (Peng & Sarazen, 2019). By using pre-trained models for the thesis, this (redundant) work can be avoided, resulting in substantial power savings.

Background

This chapter describes the theoretical background of the thesis: a mix of interaction design studies with computer science and literary theory. It also contains a prior art relevant to the research, ranging from installation art to interactive fiction.

2.1 Artificial Intelligence & Machine Learning

Even a brief history of artificial intelligence lies beyond the scope of this thesis. This section covers the main strands of artificial intelligence research, supplemented with the historical background on the present day proliferation of of machine learning approaches in relation to generative literature.

Norwig and Russel (2009) identify three main takes on artificial intelligence research: *logicist*, *symbolic* and *connectionist*. The logisict approach, championed by pioneers like John McCarthy in his hypothetical *Advice Taker* system, which uses axiomatic knowledge of the world in order to infer solutions to problems. The system would be able to accept new axioms during runtime, expanding its knowledge to cover more areas without the need for reprograming. Logicist tenancies later evolved into complex, rule-based *Expert Systems*, which saw limited use industry as tools for decision making. Combined, these ideas embody a formalist understanding of the world, whose explicit representation is enough for a correct deductive process.

Symbolic approaches have their background in linguistics and semiotics, centering around the manipulation of structures of symbols. The landmark *physical symbol system* hypothesis states that “a physical symbol system has the necessary and sufficient means for general intelligent action.” (Newell & Simon, 1976). Here, *physical* refers to an object bound by the laws of physics, while the *symbols* are defined as the physical patterns that make up expressions — a structure of symbols. A physical symbol system is thus a machine that “produces through time an evolving collectin of symbol structures”. While considerate criticism has been leveraged against the hypothesis (Nilsson, 2007), it is still core to much of the field.

The connectionist stance models intelligence systems from the structure of the human brain, in particular with concern to how it learns and remembers. The primary tool of the connectionist is the artificial neural network, a network composed of individual “neurons” that are able to transmit signals in among each other in a synapse-like matter. As connectionist founder McCulloch succinctly put it, “What we thought we were doing (and I think we succeeded fairly well) was treating the brain as a Turing machine.” (McCulloch & Pitts, 1943).

These neurons are created not by hand, but through a process called *training*, where the network uses different mathematical processes to find patterns and connections in the given *training data*. There are three fundamental kinds of training: supervised learning, unsupervised learning and reinforcement learning. During supervised learning networks are given labeled input-output pairs, from which they try to extrapolate a way to re-produce the mapping. The time consuming processes of labeling a data set is often conducted by a human agent, either manually or with the help of some kind of automated system. This labor can be alleviated through the use of unsupervised learning, a process that works on unlabeled data on which it looks for patterns without human intervention. Finally, reinforcement learning pits a network to perform a task on which is given automated positive or negative feedback (reinforcement), through which the network adjusts its model accordingly (Norvig & Russell, 2009). All three approaches have their use cases and are sometimes combined into composite methods like *semi-supervised learning*. The transformer based artificial neural networks used in thesis, like GPT-2, are trained using an unsupervised approach.

The potent combination of the exponential increase in computational resources per Moore’s Law and the rise of big data collection late ’00s (Bughin et al., 2010) has enabled major advances in connectionist machine learning. The rise of deep learning techniques expands on traditional artificial neural network by forming multiple layers between the input and output, convoluting the connections and so vastly enhancing the precision of the calculations performed within (Bengio, 2009). These models thrive on massive datasets, with recent transformer based models like OpenAI’s GPT-2 using over 1.5 billion parameters of an extended Web Text corpus (V. Liu & Curran, 2006). Such a training process requires massive amounts of compute and energy, rather problematic from a sustainability standpoint (García-Martín et al., 2019). The massive scales of these casually trained models means they have been able to outperforming task-specialized networks at their own game (Radford et al., 2019), a paradigmatic shift that has both technical and ethical implications (see Section 1.5).

Compared to symbolic approaches (fig. 2.1), artificial neural networks are a highly black boxed affair. The network has no symbolic understanding of the relationship between its neurons. For the black box, it all comes down to mathematical relationship between the new input and the pattern of data it is trained on. Drawing a comparison to semiotics, I make the argument that the artificial neural network exists purely on a plane of signifiers, disconnected in absolute from that which is signified (fig. 2.2). This state of hyperreality (Baudrillard, 1994)

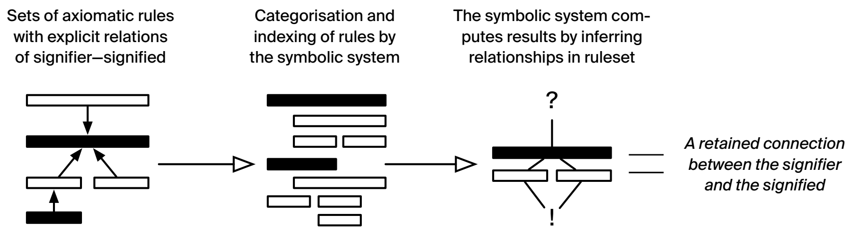


Figure 2.1: The conceptual diagram of a symbolic system, retaning the connection between signifier and the signified.

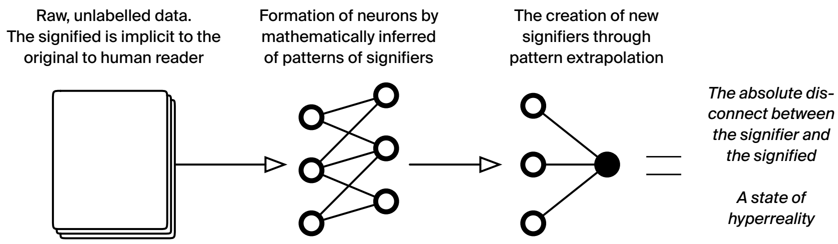


Figure 2.2: The conceptual diagram of a neural network, illustrating the disconnection of signifier and the signified.

makes it impossible for the human user to "open the black box", for there is nothing to be seen.

With the advent of modern machine learning models, it seems that, for the moment, the connectionist approach has triumphed. Criticism has been raised on the basis of lost symbolic qualities (Norvig & Russell, 2009) or the exclusion non-human of brain modeling (Zador, 2019). In spite of these claims, the rapid development and application of machine learning shows no signs of slowing down, with the new generation of pre-trained models leading way to further democratized access.

2.2 Machine Learning in Interaction Design

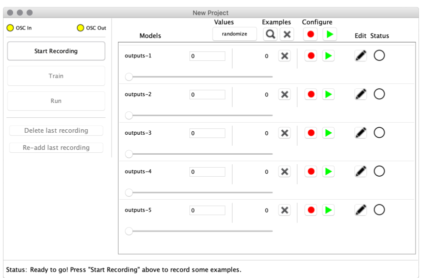


Figure 2.3: The UI of Wekinator.

Despite its growing proliferation in the IT sector as a whole, machine learning has yet to break into interaction design practice or research. As mentioned during the introduction, by 2018 there had only been 9 DIS papers that mentioning "machine learning" (Yang, Banovic, et al., 2018). As such, it becomes important to acknowledge and understand the few existing ways through which interaction designers have been able to leverage machine learning as a design material.

Rebecca Fiebrink's *Wekinator* (Fiebrink, 2010) is the closest one gets to a canonical machine learning/interaction design crossover. An early example of interactive machine learning (IML), it was originally conceived as a tool for music composition and performance. *Wekinator* affords interaction designers a graphical interface for the supervised learning of artificial neural networks. It takes as input any kind of real time data (ex. mouse movements or the capture of a camera) which it feeds into a black boxed model, where it is processed and mapped to the desired output (ex. audio synthesis or physical actuators). The labeling of the input-output mapping is supplied by the user through the same interactive process by providing real time examples.

Unlike machine learning systems aimed at computer- or data scientists, Wekinator concerns itself with what machine learning *does*. The implementation details of the models themselves are completely black boxed, wrapped in an interface so slim that it takes short of two minutes to explain and demonstrate the core functionally (Fiebrink, 2020). This is much due the fact that it is built on elements already familiar to users of "creative software", like a "recording" interaction similar to digital music creation or real time video editing (fig. 2.3). The novel use of a fully interactive interface with instantaneous feedback makes *Wekinator* very different, and for some tasks more efficient (Laguna & Fiebrink, 2014), when compared to most other machine learning systems, which often require at least some degree programming.

Fiebrink continuous to be on the forefront of the area with the *Gesture Mapper*, an elaborated system for sketching digital music instruments. By visualizing elements of the artificial neural networks in an interactive form (Laguna & Fiebrink, 2014), it enables rich inquiry in the emerging field of Human-Centered Machine Learning (HCML) (Ramos et al., 2019). The ladder is highly relevant for interaction designers as it concerns itself with the construction of interactive machine learning systems where the human is "more than a source of labels". By transforming machine learning into something interactive and responsive, it becomes a tangible material imbued with creative potential.

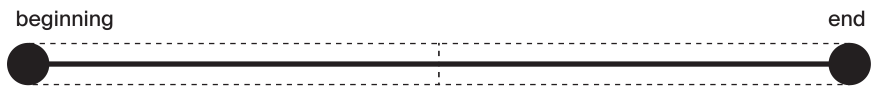


Figure 2.4
The diegetic axis of conventional literature in the book medium, with a linear flow and a defined beginning/end.

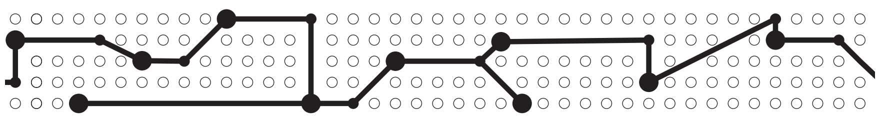


Figure 2.5
The diegetic axis of generative literature, a space of alepsis without defined terminals.

2.3 Generative Literature

French poet and scholar Jean-Pierre Balpe defines generative literature as "the production of continuously changing literary texts by means of a dictionary, a set of rules, and/or the use of algorithms" (Balpe, 2007). He notes that the author in generative literature becomes a "meta-author", a designer of the generative system rather than the text itself. This definition traces a clear lineage back to generative or systems art, similarly defined as a “practice where the artist uses a system set in motion with some degree of autonomy resulting in a work of art” (Galanter, 2003). The disciplines share much of the same history, with movements like Fluxus, Dadaism and conceptualism underpinning modern practice in both fields (Howe & Soderman, 2009). Variations on the now famous cut-up technique as described by Tzara in 1920 (Tzara, 1920) can be regarded as an early form of generative literature, as might the systematic use of the classical Chinese text *I Ching* (Aarseth, 1997) by authors like Philip K. Dick (Cover, 1974) and Jorge Luis Borges (Borges, 1992). These practices differ from conventional types of creative restrictions in the sense that they shift the agency of making from the author to the system as such.

Balpe employs Gérard Genette’s *diegetic axis* (Genette, 1983) as a device for understanding the differences between conventional and generative literature (Balpe, 2007). The *diegesis* of the axis implies the narrative as something spanning the axis through a beginning into an end. Balpe traces this logic to the linear medium of the book, with its matters front to end (fig 2.4). To break from or rewind reading are not actions afforded by the book, unless for detached reflection or referential clarification. A generative literature reverses this logic through its mode of production. At any given moment in generative work, the current passage is only a “temporary specimen of an infinite family of virtual texts” (Balpe, 2007). Every point in its axis is equal, simultaneously the start and end of an infinity of texts, a situation Balpe calls *alepsis* (fig 2.5). The act of choosing which axes to explore has to be made somewhere, either by the system, the meta-author or the reader. The existence of alepsis thus makes possible, and necessitates, a reader with agency, transforming it into the *user-reader* as described in the medium of cybertext (Aarseth, 1997). The user-reader has the liberty to materially affect the narrative, with each reading resulting in a *microfiction*, a local diegesis for that reader-user at this point in time. As predisposed structures vanish, what is left is a virtual microfiction built by the act of reading as such. This necessitates a new medium, as the solid state of the book depends on a fixed diegetic axis.

The move from reader to user-reader and author to meta-author (fig. 2.6 & 2.7) recalls a passage from literary theorist and philosopher Roland Barthes’ landmark text *The Death of the Author*, where “the birth of the reader must be ransomed by the death of the author” (Barthes, 1967). Barthes shifts the agency of literature away from the writing author by arguing that text is not a set unity, but a constantly shifting assemblage of citations as held by the reader. This definition maps closely to the attributes of interactive generative literature, understood in relation to the concept of the “cybertext”. I explore this relation in the experimental *Multiverse* system detailed in Sections 4.5 and 5.2. This discussion about generative literature continued in Section 5.3.

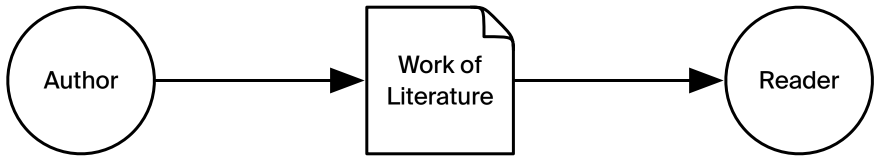


Figure 2.6
The linear author→ work → user relation in conventional literature.

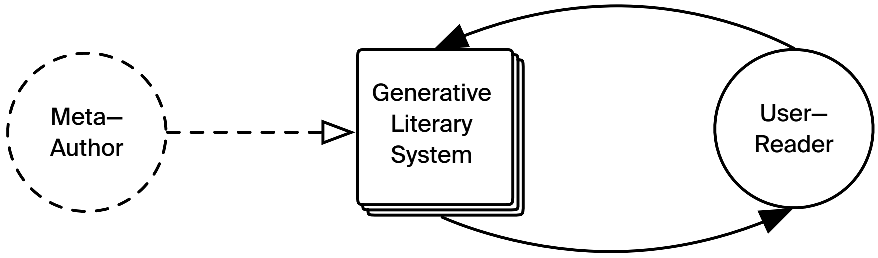


Figure 2.7
The complex, cybernetic (meta-author→) system ↔ user–reader relation in generative literature.

2.4 Cybertext and Ergodic literature

In 1997, poet and author Espen Aarseth's formulates the concept of the "cybertext" (Aarseth, 1997). Influenced by the accelerating subsumption of printed typographic culture into the digital, he proposes a new mode of engaging with literature through an understanding of the how medium *as such* constitutes the literary experience. The cybertext is not inherently digital†, rather, a cybertextual work is defined by how the literary message is delivered. Contrary to conventional literature, where all aspects of the story are given to the reader by an author, the act of cybertextual reviving requires significant work work of the reader. Aarseth (1997) dubs this concept "ergodic",

† Aarseth identifies the *I Ching* as an early example of cybertext, as it requires and contains rules for how to be read.

> *In ergodic literature, nontrivial effort is required to allow the reader to traverse the text. If ergodic literature is to make sense as a concept, there must also be nonergodic literature, where the effort to traverse the text is trivial, with no extranoematic responsibilities placed on the reader except (for example) eye movement and the periodic or arbitrary turning of pages.*

The cybertext is central to this thesis in how it affords an analysis of how an ergodic medium transforms the reader into a user-reader. This parallels the understanding of generative literature as outlined in Section 2.3, which details the transformation of the author into meta-author and the work of literature into a generative system. Together, these two perspectives constitute the mode by which this thesis understands interactive generative literature, i.e. *a cybertextual medium centered around the post-typographic feedback between generative system and user-reader, instanced by a meta-author* (see fig. 2.6 & 2.7).

2.5 Prior Art

This section details a set of prior works deemed relevant to the topic of research. Criteria included citations, historical use in the field or state-of-the-art examples. The sections on Diemut Strabe's *The Prayer* and *AI-Dungeon 2* are kept short, as they are the subject to a comparative study in Section 4.4.

2.5.1 The Prayer

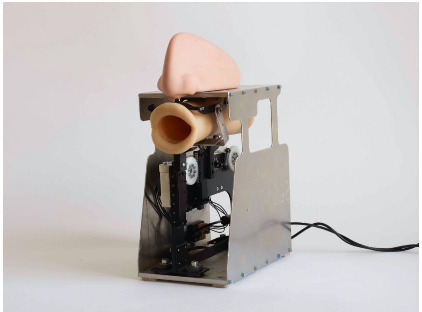


Figure 2.8: Diemut Strebe's The Prayer. © Diemut Strebe, 2020.

German artist Diemut Strebe's 2020 piece *The Prayer* (Strebe, 2020b) is not interactive, but still serves a prime example of machine learning as a design material. In her a complex multimedia installation, a GPT-2 model is trained to generate prayers, informed by a large corpus of sacred texts using transfer learning. These are in turn fed into a text-to-speech model that recites the words as they are being generated through a speaker mounted in a disembodied, silicone mouthpiece, with expressive mechatronic articulation (Strebe, 2020a). The project exemplifies the power of transfer learning, generating prayers that are distinctively "prayer-like" yet not distinct to a specific religion.

2.5.2 AI-Dungeon 2

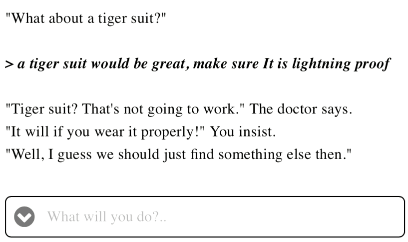


Figure 2.9: The UI of AI Dungeon 2.

The 2019 video game *AI Dungeon 2* (Walton, 2019) applies GPT-2 model to the text adventure style of interactive fiction. It plays much like the archetypal game in its genre, with the exception that it is entirely generated on-the-fly. Unlike its predecessors in AI Dungeon 2 does not limit the player to a closed vocabulary. Instead, the network correctly interprets any arbitrary string its is passed, and is able to react in often comical, yet surprisingly coherent ways.

2.5.3 HuggingFace Transformers & Write With Transformer

Advances in machine learning and natural language processing, such as the maturation of pre-trained transformer architectures, has done a lot in affording non-specials access to highly sophisticated technologies (Wolf et al., 2019). This radical shift in accessibility is one of the primary developments that might allow for interaction designers to engage with ML as a design material.

This thesis uses the *HuggingFace Transfomer* project from Hugging Face Inc. as a case for this technological shift. Built on tenants of "ease of access, diversity and sharing" (Wolf et al., 2019), it provides a sleek library for the Python programming language that interfaces with ma-

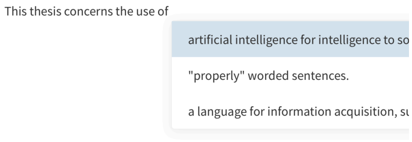


Figure 2.10: The user interface of Write With Transformer.

chine learning tools *Tensorflow* (by Google Corporation) and *PyTorch* (from Facebook's AI Research lab). The library provides several pre-trained models to choose from, including the previously mentioned GPT-2, which is suitable for text generation (Radford et al., 2019).

Users can try out the text generation capabilities of the HuggingFace transformers without writing a single line of code through the online tool *Write With Transformer* (HuggingFace, 2020b). The interface (fig. 2.10) is that of conventional text editor, with the added feature of ML powered text auto-completion. Hitting the `tab` key pops up a selection with three possible continuations, selecting one inserts it into the document. The completion is near instantaneous and generates believable suggestions more often than not, allowing anyone to experience the power, speed and ease of use afforded by this new generation of pre-trained transformer based models.

Methodology

The chapter describes and argues for the choice of methodologies employed throughout the thesis, informed by the Covid-19 pandemic.

3.1 Research-through-Design

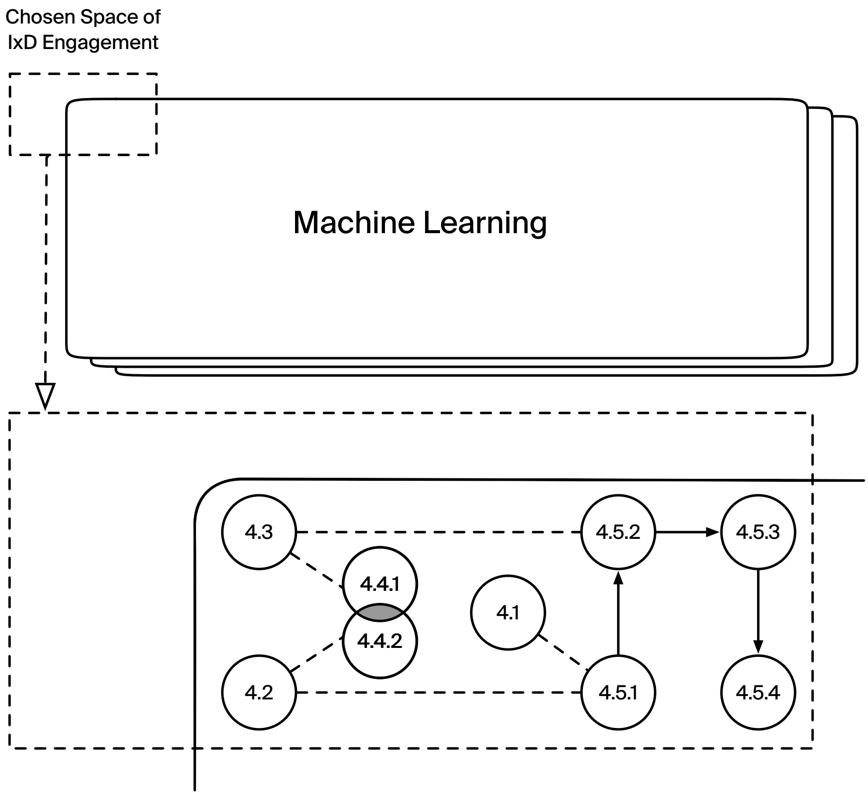
The thesis is framed through a Research-through-design (RtT) methodology. The RtD framework provides useful methods that allow for designers to produce knowledge through the novel results that come out of the design process (Zimmerman et al., 2007). The questions raised lend themselves to a such a process, as the expected outcome is not a design artifact as such, but an inquiry on how interaction designers can work with machine learning as a design material. The fluid practice of the designer has the potential to discover novel answers to a hypothesis through its high tolerance of “drifting” (Krogh et al., 2015).

3.2 RtD Typologies

Krogh et al. (2015) write about the “drifting” afforded by the design process through the question of how this impacts the quality of results. Through a study of ten different PhD theses, the authors formulate five RtD “typologies”. These interpolated methodological approaches are shaped through the varying degrees to which RtD attributes and methods, like drifting and sketching, are employed. These typologies are meant both as tools to retrospectively frame a RtD process and as methods to structure new research.

This thesis was primarily researched following the typology Korgh et al. identify as *probing*. This approach follows threads as they emerge naturally in the design process. Korgh et al. argue that this meandering mode is useful when engaging with a research field larger than what an individual project is able to cover. This suits the focus and questions of this thesis well, as it only embeds itself in the outer, interactive slimmers of the huge machine learning space (see fig. 3.1). By following a probing approach, this research aims to identity possible openings in an unknown topic and then approach these from different angles.

Figure 3.1: Visualization of the design process of this thesis, the numbers refer to the sections in the text. The research was conducted in a probing mode, with a comparative (4.4) and serial (4.5.n) detours. The numbering correspond to sections in the thesis: **4.1** Generating Text with HuggingFace Transformers **4.2** Sketching with Neural Plumbing **4.3** Articulating Machine Learning Attributes **4.4** Comparative Study: The Prayer & AI-Dungeon 2 **4.3** Prototyping the Cybertextual Multiverse



Supplementing the probing approach are two specialized detours into a comparative and serial mode. The former was employed in Section 4.4, where two different approaches to generative literature are examined using the tools developed in Section 4.2 and 4.3. Deep examination not only uncovered previously unseen qualities in the examined artifacts, but also in the methods used. The serial typology was used during the creation of the "Multiverse" prototype in Section 4.5, as an RtD approximation of an iterative design process.

3.3 Auto-Ethnography

The research process of this thesis employs auto-ethnographic elements as a secondary mode of inquiry. Such practice recognizes the researchers self-reflection as a tool for generating qualitative knowledge (Cunningham & Jones, 2005). These techniques can be useful during fast paced design processes or in situations that for whatever reason (in this case, the exceptional situation of the Covid-19 pandemic) prohibit the use of external subjects. In the case of this thesis, the author is in a similar situation to the imagined reader: the interaction designers that wishes to engage with machine learning, but lacks formal education on the subject. Taking this personal experience into account can help inform for the pedagogy of new methods.

3.4 Limitations & User Research

Due to the Covid-19 pandemic, with its policies of social distancing and quarantine, the decision was made to minimize human user research. For experiments where such input was required, like experience prototyping of the *Multiverse* exploration (Blashki, 2013), exceptions where made but limited to family and friends in the form of micro usability tests (Goodman et al., 2012). As the project employs a Research-through-design approach, as opposed to a more feedback heavy user-centric-design process, this should invalidate the results.

3.5 Ethical Considerations

No personal data was collected at any point during this research. Any activity involving humans was left unrecorded, limited to ephemeral notes and the memory of the discussions.

The source of this thesis and the code referenced is stored in a decentralized git repository that is available openly on Github, a service owned and maintained by Microsoft Corporation. It contains no personal data or other kinds of sensitive information.

Design Process

This chapter describes the main design process undertaken during this thesis. It details a series of prodding explorations, following the Research-through-Design methodology as outlined Chapter 3. The process entailed initial explorations and the development of two methods, culminating in an interactive prototype. Section 4.1 and 4.5 deal directly with generative literature, while Section 4.2 through 4.4 constitute a methodological exploration of tools that will allow for interaction designers to engage with the machine learning space. See fig. 3.1 for a visualization of the design process.

4.1 Generating Text with HuggingFace Transformers

The initial experiment of the design process investigates the validity of the claims that pre-trained machine learning models afford non-specialists the ability to engage with these technologies. As stated in Section 2.5.3, *HuggingFace Transformers and Write With Transformer*, this thesis employs the HuggingFace Transformers library as a primary example this new generation of tools.

After consulting the HuggingFace documentation (HuggingFace, 2020a), I was able to write the tiny program below, using a Python interoperability library for the Clojure programming language.

```
;; Initializing the HuggingFace library
(def tokenizer (py. transformers/GPT2Tokenizer from_pretrained "gpt-2"))
(def model (py. transformers/GPT2LMHeadModel from_pretrained "gpt-2"))

;; Simple functions that interface with the HF Transformers
(defn ->tokens [input] (py. tokenizer encode input))
(defn decode [tokens] (map #(py. tokenizer decode %) tokens))
(defn generate [tokens] (py. model generate tokens))
```

This is the minimal amount of code required to build a system that uses the GPT-2 model to generate text from an input prompt. There are multiple ways to make the text generation more sophisticated using different types of search and sampling techniques (von Platen, 2020), most of which are easily accessible as optional parameters to the functions provided by the HuggingFace Transformers library.

Below is an example of the program running a string through a pipeline of functions that performs tokenization on the input, generates new text and then decodes it back to alphanumerical form. The => denotes the resulting output.

```
(-> "This thesis concerns the use of machine learning as a design material"
    ->tokens
    generate
    decode)

=>  "This thesis concerns the use of machine learning as a design material.
    In our study, we found that machine learning applications can be used
    to provide a number of useful skills, such as the generalization of
    machine learning and the construction of a set of applications,
    such as the generalization of machine learning, or the construction of a network."
```

This might seem like a banal exploration, following documentation to interface with a library is far from cutting edge research. Yet, it is quite profound that an amateur programmer is now able to use the same state-of-the-art machine learning technologies as leading IT-companies, no prerequisite knowledge required. Not all interaction designers program, but it is not all that difficult to imagine the jump from an example like the one above to a Wekinator-like† software packaging with a graphical user interface.

† For background on Wekinator, see Section 2.2

As a quick proof of concept, a lightweight command line interface was written to enable elementary interactions with the model (fig 4.1). The program prompts the user-reader for an initial sentence, which is required by the artificial neural network to generate text. After input text processing and generation is complete, the user-reader gets to choose between one of three possible continuations, the selected of which get appended to the current “story” and fed back into the network. The decision to provide a moment of choice is motivated by both practicalities and concept. From the pragmatic standpoint, it is sometimes necessary to give the network more than one try to generate something the user-reader deems meaningful. As the experiment employs random sampling (von Platen, 2020), it has the ability create multiple outputs from the same input. The selection process also affords the user-reader to explore initiate space of alepsis, engaging with the unique materials of generative literature and machine learning.

```
I was alone in my study, contemplating the nature of intelligence.

1. I found the number of times I had seen someone that had no interest in what they were doing.
2. I felt that for some reason, I was able to discern that in my experience,
   there were many different ways of experiencing the nature of intelligence.
3.  I was surprised to learn that the very definition of intelligence is defined in our minds.
>
```

Figure 4.1: The simple command line interface to the text generating artificial neural network, showing the state after the user-reader has entered a prompt. The user can choose between one of the three continuations of the text.

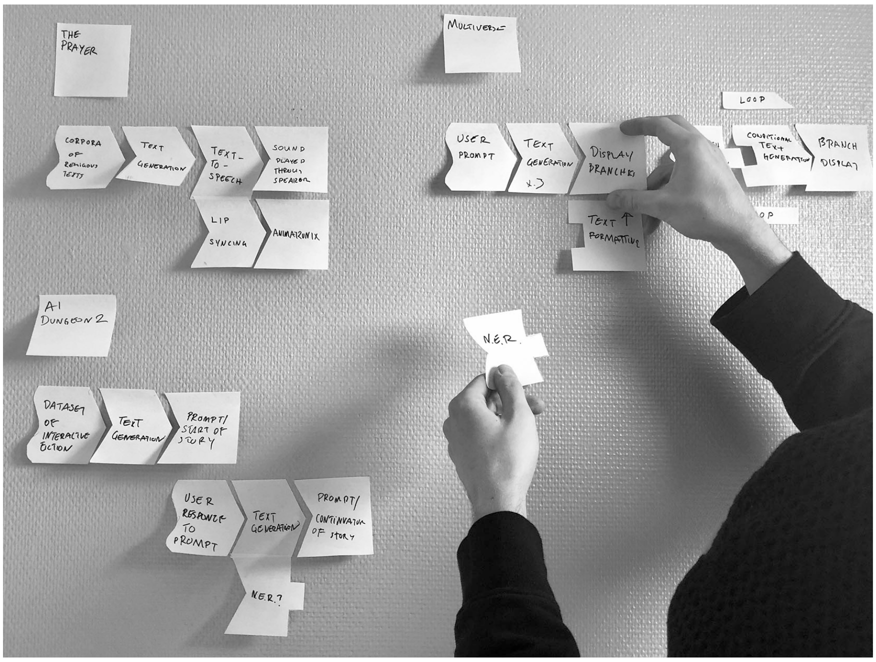


Figure 4.2: A session of the work analyzing canonical examples and sketching a system for interactive literature.

4.2 Sketching with Neural Plumbing

Yang, Scuito, et al. (2018) identify a combined lack of education and absence specialized methods as major constraints for enabling interaction designers to work with machine learning as a design material. They suggest that design researchers should try to provide “abstractions, exemplars and new tools and methods” for practitioners. This exploration answers the call by proposing and exploring a domain situated workshop dubbed *Neural Plumbing*. It aims to afford interaction designers and other non-technical stakeholders ways to sketch with machine learning, by focusing on with what machine learning *does*. This keeps in line with the call of Yang, Scuito, et al. to develop new abstraction for designers to work with machine learning, while still anchoring it in the material properties of the actual technology. The method described here is not specific to the case of generative literature and is intended to be used as a stand-in for other brainstorming activities in any design process that involves machine learning.

The star of *Neural Plumbing* is the *pipeline*, a common pattern in data processing and machine learning. Conceptually, the pipeline denotes a sequence of data processing elements through which output of one part is threaded as input to the next. This is encountered at multiple scales of machine learning systems, from the high level of how data flows between its parts, to the internal plumbing that chains different artificial neural networks together to provide complex output. During the workshop, participants cooperatively build pipelines from a set of given elements, each representing a common application of machine learning. Technically, it acts on a relatively high level of abstraction, exposing only the relevant parts from multiple levels of the stack.

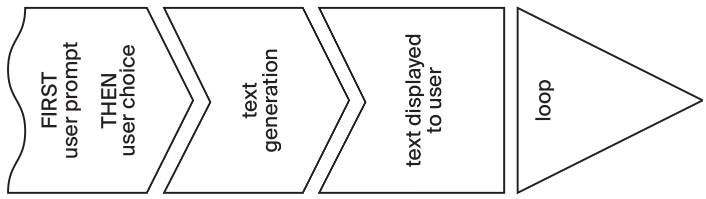


Figure 4.4: The simple prototype from Section 4.1 visualized through the neural pipeline.

Pipelines are constructed by arranging shapes, each corresponding to a single part of the system. These puzzle piece-like elements have a semantic form that denote their type and use, allowing for easy visual re-organization of process. This abstraction acts on the logic of *what machine learning does*, affording designers and other stakeholders to freely sketch the implications of the parts, disregarding implementation. It builds upon familiar techniques like brainstorming, with influences from more specialized methods like the lightweight yet complex paper prototyping of "Business Origami" (BO) (Hanington & Martin, 2019). In BO, participants use paper-cutout tokens to represent the parts that make up a system, arranged in what Hanington and Martin call *sets* or *stages*. The act of bringing otherwise abstract elements into a psychical dimension makes their relations explicit, while the jigsaw-shape of each element suggests a degree of plasticity. The shapes where designed using a button-up approach, where I applied the technique to explain existing design examples. The material from this impromptu session was then re-worked and simplified into the contents of Appendix A.

Figure 4.4 visualizes the command line-prototype of Section 4.1 through a neural pipeline. The elements are able to communicate what the simple program does at both a high and material level. The very abstract shape of the generated material becomes clear, while staying agnostic to implementation. The diagram similarly schematizes the control flow of the program to an near 1:1 mapping, thanks to the use of high level constructs like the HuggingFace Transformers library (see: Section 2.5.3). An argument could be made that the pipeline is *too* abstracted, lacking qualities essential to the experience of using the actual prototype, such as literary style. These shortcomings are further discussed in Section 5.1.

In addition to serving as a tool for sketching and prototyping, the exercise pedagogically explores the differences and integration points between the two main types of input-output in machine learning: *vector* and *sequence*. A sequence represents a variable amount of data with a pre-defined ordering, while a vector is of fixed length in its space. Different artificial neural networks require/produce either one or the other, which becomes important in the construction of pipelines. The varying edges of the elements make up a jigsaw-like systematic set (Sanders & Stappers, 2012) that help reinforce the relation between elements though a clear visual language.

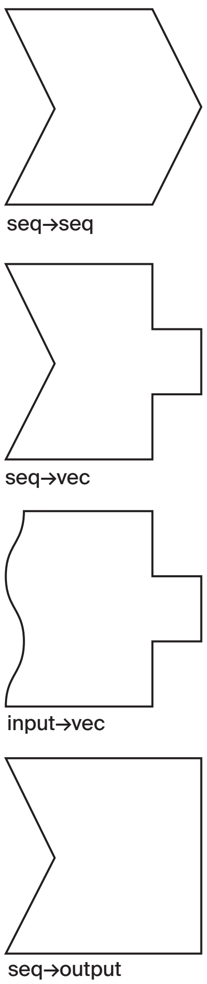


Figure 4.3: Examples of pipeline elements.

The goal of this exercise is to help get the participants closer to the material basis of machine learning, while still keeping it at a non-technical level of abstraction. It achieves this through a focus on content and transformation, rather than underlying technologies as such. For example, the decision between using GPT-2 or BERT might have a large impact in a production environment, but not during sketching. If more precision is required, it is fairly traction free to extend the set of symbols by cutting the post-it into another shapes. The workshop is imagined as a fairly stand-alone exercise, but could also be used as a preface for interactive machine learning or similar activities.

Neural Plumbing is used in Section 4.4 as a tool to examine two cases where machine learning is used as a design material. The methods is discussed in Section 5.1. For the materials created for use in the workshop, including a complete list of available elements, see Appendix A.

4.3 Articulating Machine Learning Attributes

The lack of a situated terminology has been identified as one of the primary obstacles preventing interaction designers from engaging with machine learning (Dove et al., 2017). This problem is not unique to machine learning, the problems of formulating a domain-specific language is something that applies to most underdeveloped research fields. This section explores a possible terminology describing what machine learning *does*, unlike the existing technical terms that describe what machine learning *is*.

The broader field of interaction design has approached the problem of a shared, appropriate terminology through multiple angles. Löwgren articulates what he calls *interaction aesthetics*, a framework to characterize the temporal aesthetics qualities of interaction (Löwgren, 2009). Though the vivid concepts of *pliability*, *rhythm*, *dramaturgical structure* and *fluency*, he attempts to “initiate a practice of criticism” within/about interaction design, using a terminology that is capable of describing the unique experiential qualities of interactive artifacts.

Löwgren’s vivid terminology can be contrasted with the tuabualted approach of Lim et al. (2009). They propose a system of *interactivity attributes*, a way to describe immaterial interactions in the same concrete fashion as one “describes and perceives physical material”. Their proposal manifests itself as a set of attributes represented by a “prototype pair”. For example, the *movement* attribute exists on a scale from *slow* to *fast*, while *concurrency* moves between the axis of *concurrent* and *sequential*. They decouple interactivity *as such* from the material that situates it, retaining the ability to maintain a rich discourse about “the invisible quality” of interactions through this set of attributes, what they call the shape of interactivity (Lim et al., 2009).

After considerations and discussions with colleges, I found the terse conceptual framework of Löwgren's interactive aesthetics better suited for a context of an already established practice, while the simple attributes proposed by Lim et al. could be more helpful in aiding the articulation of new field. The act of decoupling the shape of interactivity with its material medium is similar to the move from to discourse around what machine learning *does*.

Table 4.1 contains an early sketch of what a set of machine learning attributes could look like. The attributes/prototype pairs were formulated bottom-up, through a casual analysis of that featured both interactive and non-interactive machine learning projects. For example, the simple command line-prototype from Section 4.1 is fairly *surprising* despite its high level simplicity. This *novel* application of machine learning often produced output that was well above expectations, while still remaining quite *transparent*, as it is easy to see how the user supplied prompt is used to generate a continuation of the text. The interaction is *two-way*, as the user is in control of the alepsis of the literature. The prototype is somewhere in the middle of *technomorphic* and *anthropomorphic*, in the sense it does not shy away from the bewildering nature of raw machine generated texts, while also writing a story — an activity we usually attribute to humans.

The set of machine learning attributes presented here is no way complete. Like the other experiments, it is an attempt to identify possible approaches for interaction designers to engage with machine learning. The method is further discussed in Section 5.1.

Table 4.1: Proposed machine learning attributes, complete with prototype pairs.

ML Attributes	Prototype Pairs	
authenticity	*predictable*, output is predictable/spurious	*surprising*, output is surprising/authentic
features	*novel*, adds new behavior	*enhancing*, modifies existing behavior
graspability	*transparent*, affords outside understanding	*black box*, prohibits outside understanding
interaction	*one-way*, output-only	*two-way*, input-output loop with user
integration	*tangiable*, obvious integreation point	*intangiable*, non-obvious integration point
locality	*local*, tangibly present data/processing	*foreign*, intangiable or distant data/processing
subjectivity	*technopomorphic*, ML is a subject in itself	*anthropomorphic*, ML imitates a human subject

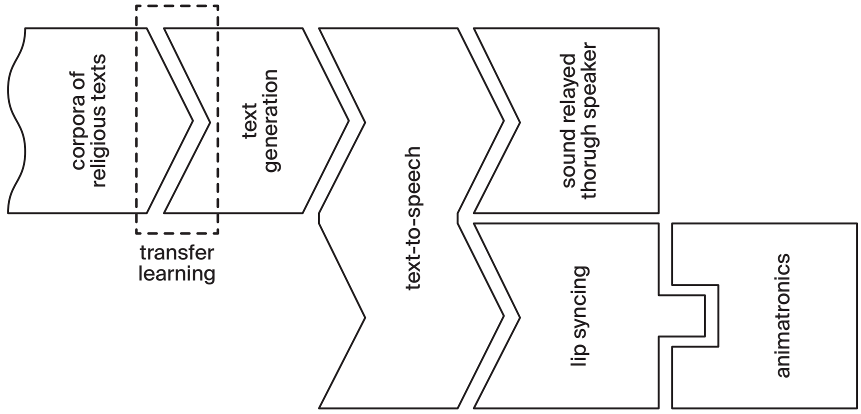


Figure 4.5: Diemut Strabe's The Prayer visualized through the neural pipeline.

4.4 Comparative Study: *The Prayer* & *AI-Dungeon 2*

This study compares two projects, *The Prayer* and *AI-Dungeon 2*, that work with machine learning as a design material. Deeply examining two succinct cases can reveal complexities and additional qualities not found when seen only in themselves (Krogh et al., 2015). It uses the two tools proposed earlier in this chapter: *Neural plumbing* (Section 4.2) and *Machine Learning Attributes* (Section 4.3).

4.4.1 The Prayer

Unlike *AI-Dungeon 2*, Diemut Strebe's 2020 piece *The Prayer* is not strictly interactive. Closer to art, it manifests itself as a multimedia installation, first exhibited at the show *Neurons, Simulated Intelligence* at Centre Pompidou, Paris in 2020 (Strebe, 2020b). It uses a GPT-2 model specialized through transfer learning, a process where one applies a niche corpus to the existing connections in a casually trained general artificial neural network. When done correctly, this leverages the progress made in a generic domain to gain a head start when tasked with a related problem. In this case, applying ability to "correctly" predict the next word to the job of predicting the next word of a religious text. As a result, the system generates sermons that sound succinctly "prayer like", without belonging to one single faith (in some sense, they belong to all input religions simultaneously). If one wants to generate a specific type of prayer, one would have to tune either the supplied input or the data set used during the transformer. The piece is accompanied by a display visualizing the "sound-mechanical interface" using the software Max MSP (Strebe, 2020a).

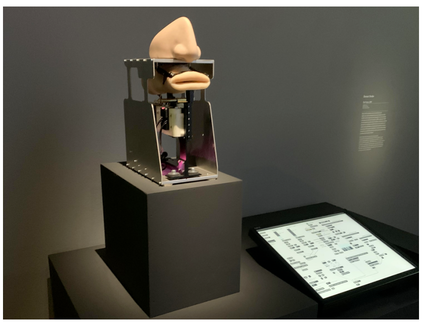


Figure 4.6: Diemut Strebe's The Prayer exhibited at Centre Pompidou with display. © Diemut Strebe, 2020.

The Prayer explores machine learning as a design material in multiple theoretical modes. The installation employs deep learning not only as a technical tool, but also as a device to illustrate what Strabe calls "'the noumenal' as the mystery of 'the unknown'" (Strebe, 2020b). The very concept of an artificial neural network's mimicry of the human brain inadvertently creates a noumenal black box, whose mathematical divinity repeatedly subverts understanding. She argues that self learning programs constitute a state where scientific knowledge and belief coexist in harmonious paradox. Materially, artificial neural networks are nothing but probabilistic mathematical matrices, and yet, an outside observer has no way of either knowing or understanding which/why these artificial synapses have formed. It makes a point by hiding the alepsis, re-articulating it as attribute of the supernatural.

Materially, *The Prayer* serves as a textbook example of a machine learning pipeline, the sequential processing of an input through discrete elements. A GPT-2 model generates the base prayer, which then gets processed (likely by other artificial neural networks or natural language processing software) to clean structural errors and introduce intonation indicators. These are required by *Aamzon Polly*, a neural text-to-speech service that outputs audio to generate animatronic lip movements. See fig 4.5.

4.4.2 AI-Dungeon 2

AI-Dungeon 2 applies machine learning to the medium of interactive fiction, the cybertextual genre of software where the user-reader is able to influence the progression of a story through textual commands or a graphical interface (Montfort, 2005). It replaces the human author with GPT-2 powered text generation, resulting in a system that is able to tell any of story the user-reader might (or might not) imagine. Instead of parsing a fixed grammar of responses, it feeds the model with any of free text the user supplies. This makes for a profoundly interactive experience that blurs the line between reader and author. The generated stories are often surprising in both content and form. If the user starts writing a narrative about, for example, the municipal policies of Umeå in 2014, *AI-Dungeon 2* happily obliges. The results are not always fully coherent, but often rather humorous.

Figure 4.7: AI-Dungeon 2 visualized through the neural pipeline.

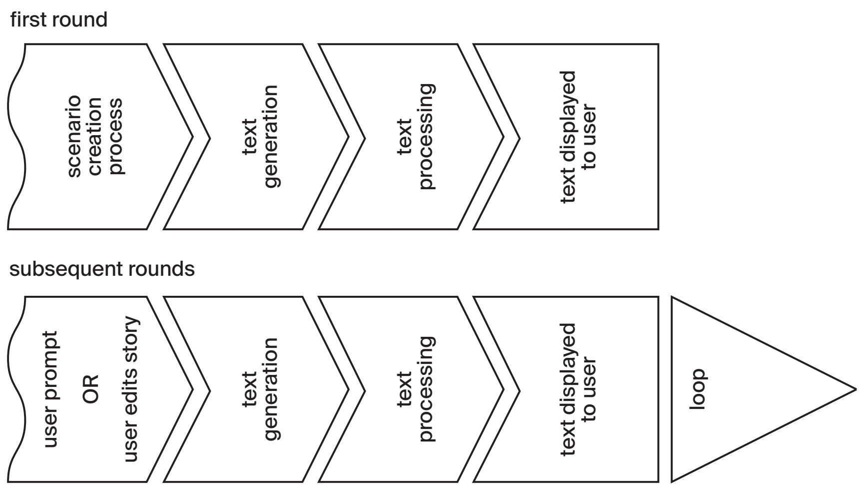
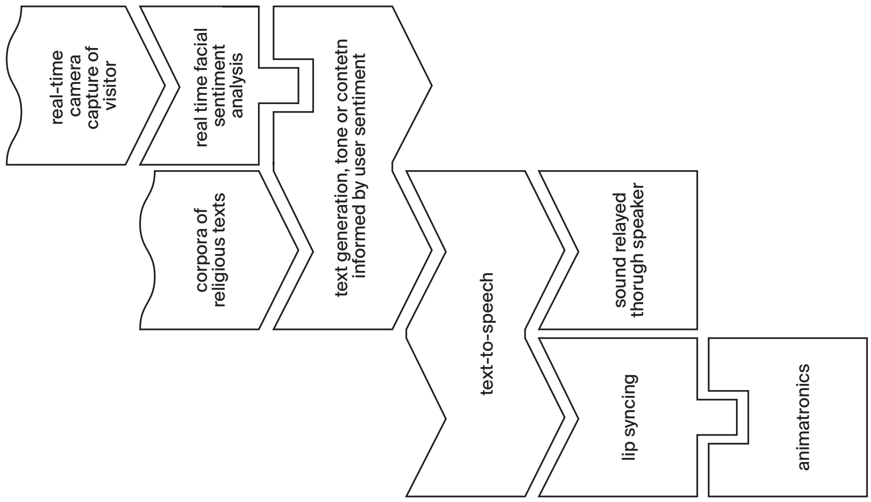


Figure 4.8: A proposed extension of *The Prayer*, using real time video sentiment analysis to change the tone and content of the generated prayers.



Some clever technical moves are required in order to support such freedom. The sole function of GPT-2 is to predict the next word in a sentence, as such, free form text generation requires some kind of initial prompt. *AI-Dungeon 2* deals with this restriction by starting each new session with process where the user-reader picks from a setting, enters a character name and other miscellaneous information relevant in the chosen context. Once this process is complete, the story begins by prompting the network using a template introduction informed by the users choices. From here, the network and user-reader take turns in progressing the story, prompting each other in a two-way communication. The user-reader replies in one of three modes: `do` (an action), `say` (a reply) or `story` (describing progression). In addition to appending a continuation of the story, the user-reader might edit any previous segment, shifting the alepsis.

The ability to edit any part of the story at any point in time is perhaps the most fascinating aspect of *AI-Dungeon 2*. It represents one of the most direct interface with the alepsis of a story imaginable, smoothing the strained mode of conventional literature into something more akin to verbal culture. The stoicism of the typographic medium is transgressed, making way for a fluid new logic.

4.4.3 Comparison

The designs distinctly articulate the aesthetics of generative literature through different approaches to the alepsis. The literary process of *The Prayer* is intentionally obscured, transforming the black box into a design material uniquely able to expresses the noumenal qualities of religious text. Strebe questions the interactions between technology and faith through a literal deus ex machina. If we are unable to understand the artificial neural network, then it only follows that machine learning is a machine of transcendence. This is a sharp contrast to the tangible alepsis of *AI-Dungeon 2*, which turns the black box inside out, giving the user-reader direct access to a smoothed typographic space.

As interaction designers we might consider ways to introduce the interactivity of *AI-Dungeon 2* into *The Prayer*, without losing its conceptual core. Using the library of elements included in the material of Neural Plumbing (see Appendix A), we can start to sketch an extension of the pipeline. A possible development might be to introduce video processing into the installation, maybe in the form of a camera that reads the sentiment of the viewers facial expressions. This value could then be used to control the tone and content of the generated prayers. For example, the concentrated look of a visitor might entail a meandering, quiet explanation, while a skeptic or tired face might trigger a flurry of glossolalic conviction. This introduces interaction with the alepsis in a subtle but interesting way, mirroring the way a preacher adapts a sermon to the atmosphere and need of the room.

Section 4.3 describes an adaption of *interactivity attributes* (Lim et al., 2009) to machine learning as a way to afford interaction designers a language to discuss the aesthetic qualities of such technologies. Table 4.2 provides a full comparison between the two cases, highlighting a shared understanding of machine learning as a tangiable design material, albeit from widely differing angles.

Table 4.2: Comparing the machine learning attributes of the two cases.

	The Prayer	**AI-Dungeon 2**
alepsis	intangible, “noumenal”	tangible, interactive
authenticity	surprising	surprising
features	novel	novel
interaction	one-way, user-observer	two-way, user-reader
graspability	black box	tangible, editable
locality	local, tangiable	foreign, cloud
subjectivity	technopomorphic	anthropomorphic

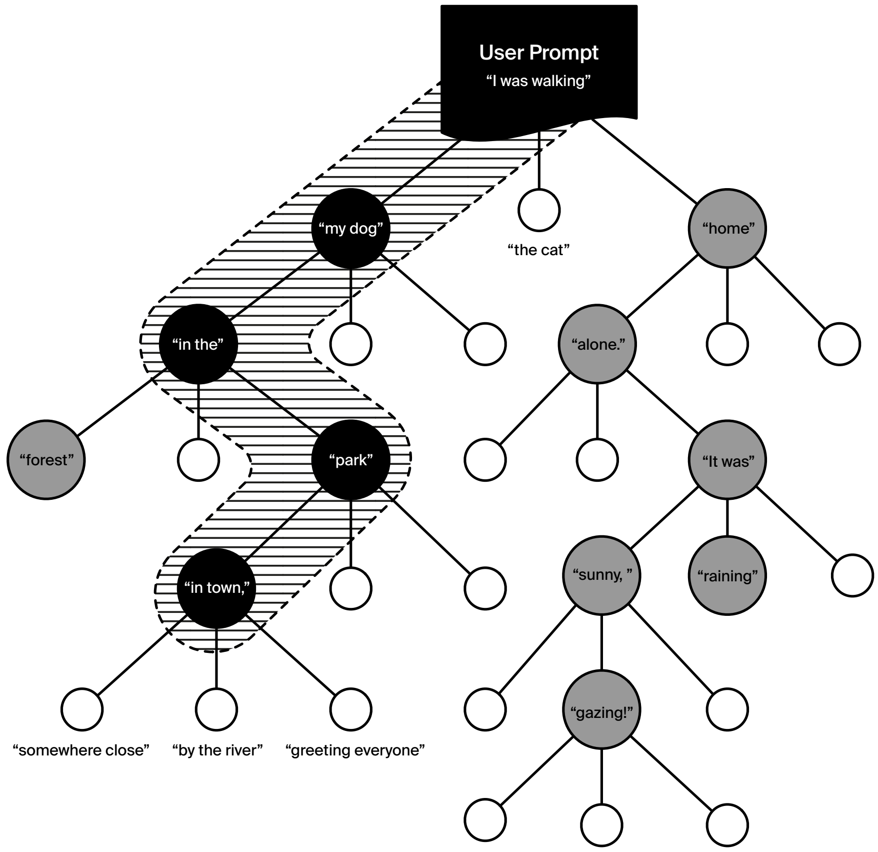


Figure 4.9: The literary space of a story visualized as a tree. The initial user prompt makes up the root, proceeded by a set of nodes with hierarchical parent/child relationships. Explored nodes are filled, while unexplored ones remain empty. The explored nodes with a black fill from a literary axis, highlighted here with a dashed field. Traversing down an axis returns set of sentences. The words attached to the nodes of this diagram illustrates how a story might branch out from an initial prompt. In the real design, each nodes is two sentences long. This illustration is simplified to a single linguistic component for sake of brevity.

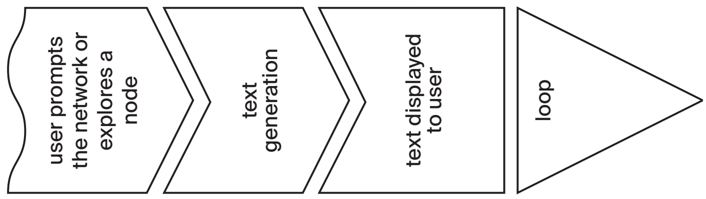
4.5 Prototyping the Cybertextual Multiverse

The concluding part of this chapter details the process of designing a fully functioning prototype of a cybertextual system for generative literature. The goal with this exercise is to synthesize the literary theory discussed in Sections 2.3 and 2.4 with the text generation of the HuggingFace Transformers (Section 4.1), using the methods proposed in 4.2 and 4.3. It moves into what Krogh dubs a "serial" mode, a series of sequential investigations more in kind to a regular design process (Krogh et al., 2015). This allows the designers to thoroughly explore the chosen topic using methods unique to its profession, allowing for inquiry into the potentials of machine learning as a design material.

The name of the prototype, *Multiverse*, refers to the idea of the multiverse as conceptualized in the many-worlds (or Everett) interpretation in quantum physics, which implies the existence of a large, perhaps infinite, number of universes (DeWitt & Graham, 1973). The theory sees time as a multiversal tree from where every possible quantum outcome is actualized. This parallels Balpe's definition of generative literature as "infinite family of virtual texts" (Balpe, 2007).

This exploration employs the device of the design prototype, as understood by Houde and Hill (1997). In order to deal with the complexity inherent in interactive computer artifacts, massively amplified by the use of machine learning, they chart the attributes of a specific prototype in a variable space of "what a prototype prototypes". This space

Figure 4.11: The pipeline of the first iteration, extracted from the prototype described in Section 4.1.



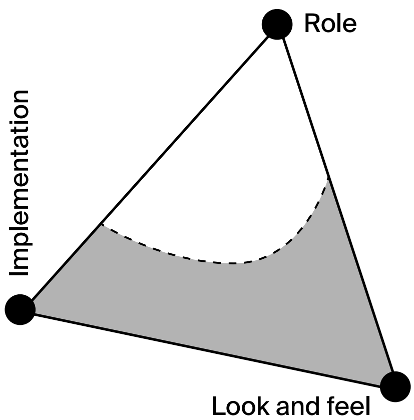


Figure 4.10: The prototype-space investigated through the *Multiverse* prototype (Houde & Hill, 1997).

is delinted by three poles: Role (in the user's life), implementation and look and feel. Situated in this framework, the *Multiverse* prototype can be understood as an exploration in how the implementation and look-and feel of a cybertextual system for generative literature might be like. There is an emphasis on the type of interactions that afford an intuitive navigation and spatial understanding of complex, bidirectional texts. Implementation is investigated by programming a fully functional system that employs a "real" artificial neural network for text generation. As the goal of the prototype is to serve as a device for a RtD investigation into the ways interaction designers might engage with machine learning as a design material. Aspects of feature discoverability and presentation as such where considered secondary.

Due to time constraints and the 2020 Covid-19 epidemic there was no in-depth user testing sessions done during, or after, the development of the prototype. To partially remedy this shortcoming, I conducted casual testing sessions with friends and family. These users where generally familiar with the premise and scope of the project, so we where able to perform a series of micro usability tests (Goodman et al., 2012). These where informal in nature and not documented, serving mainly as a way to get continuous feedback on the prototype as it developed, as a sort of lightweight co-design process.

4.5.1 Bidirectional Text Navigation

The simple command line prototype in Section 4.1 uses the HuggingFace Transformers to generate continuations of an initial prompt. The user reader is presented with three possible sentences, choosing one starts the loop anew. As such, it navigates the text in a unidirectional fashion, not all too different from the diegetic axis of conventional literature. A novel bidirectional type of navigation is necessary in order to afford exploration of the wider literary space of generative texts. Non-linear structures require a fundamentally different mental and computational representation than the sequential model used in the first prototype. Accepting that GPT-2 and other artificial neural networks require some kind of initial input in order to generate text means there is a descending, hierarchical relationship between the prompt and its continuations. Such relationships can be effectively modeled using trees, a design choice with plenty of prior art in the related genre of interactive fiction (Montfort, 2005).

The shift in representation to a bidirectional mode warrants a similar move in terminology. For the remainder of this thesis, I will refer the tree-story structure of the using the nomenclature below (see fig. 4.9).

story A specific tree/literary space.

node The atomic part of the story, contains a fixed amount of text (in this implementation, two sentences). Has a single parent and some amount of children.

axis A specific path in the space of the story. Using an axis as a map to traverse the story, one looks at each node in sequence to extrapolate that specific narrative.

exploration A node has two states: explored or unexplored. An explored node has been visited when traversing an axis, which has triggered the creation of three children, each containing its own possible continuation of said axis. Unexplored nodes have not been traversed as part of an axis and so are without children. Unlike the leaf of a conventional tree (which marks an end point of a structure), an unexplored node represents just that, a conceptually tangible part of Balpe's infinite virtual text, unrealized only in computation. Unexplored nodes are visible to the user-reader, otherwise they would not be able to progress, so they are unknown in the sense that it has yet to reveal its children.

In order to make a new kind of interaction like bidirectional navigation manageable, it has to be as simple as possible. The first screen greets the user-reader with a typewriter-like text prompt, fig. 4.12. Upon submission, the screen changes to display the axis that the user-reader is currently exploring. Continuations are drawn with a dimmed color in a separate area below the current axis. When the user-reader hovers a sentence, it changes its appearance to match the other explored sentences, fig. 4.13. Selecting a continuation moves the user-reader down a node in the tree, and the process repeats, fig. 4.14. If instead the user-reader hover over a sentence that is part of the current axis, its style changes to signal a vertical movement in the tree, highlight only that node and its ancestors, fig. 4.15. If the user-reader selects an earlier node, the axis changes appropriately, fig. 4.16. From here, the user-reader is free to explore new literary axes, fig. 4.17.

Although different in navigation and interaction, this iteration still shares the same basic neural pipeline as the command line example, fig. 4.11. It re-uses the text-generation core, but moves away from a terminal interface to a richer web-based environment. It is built using a conventional client-server architecture, employing a full stack written in the Clojure programming language, with a reactive Clojurescript front-end. The source for the prototype can be found in the project repository and totals less than 700 lines of code.

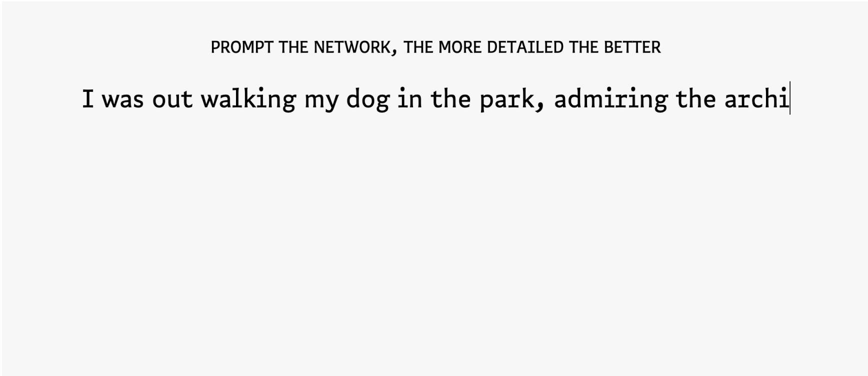


Figure 4.12: The first screen, prompting the user for an introduction to the story.

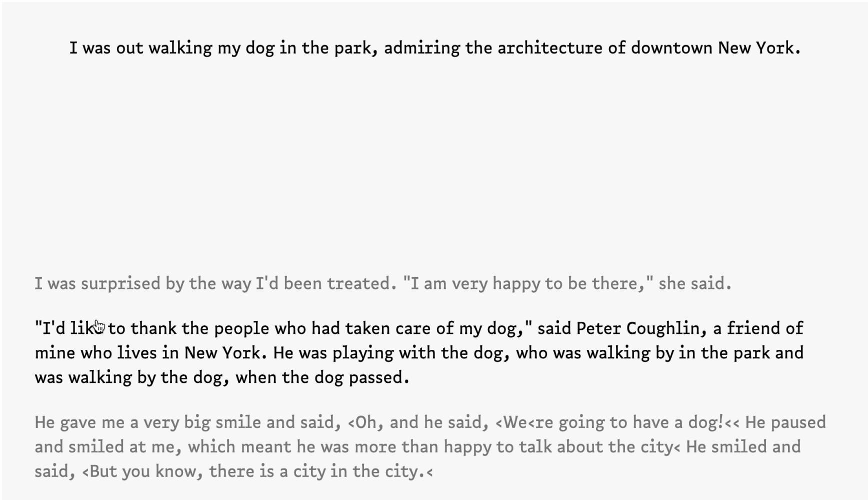


Figure 4.13: The initial story screen, showing the root node and three possible continuations, the second of which is highlighted.

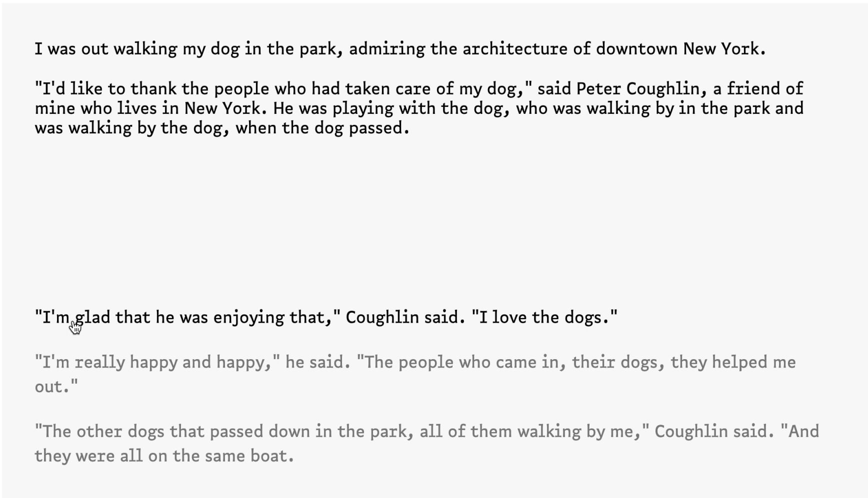


Figure 4.14: The story after the user-reader has explored a single continuation. This process repeats as it explores this axis further.

Figure 4.15: Navigation back up the axis is done by selecting the desired parent node. Here, the user is highlighting the second part of the axis, indicated by a dimming of its children.

Figure 4.16: After selecting the desired ancestor node, the user-reader has now navigated vertically upwards in the axis.

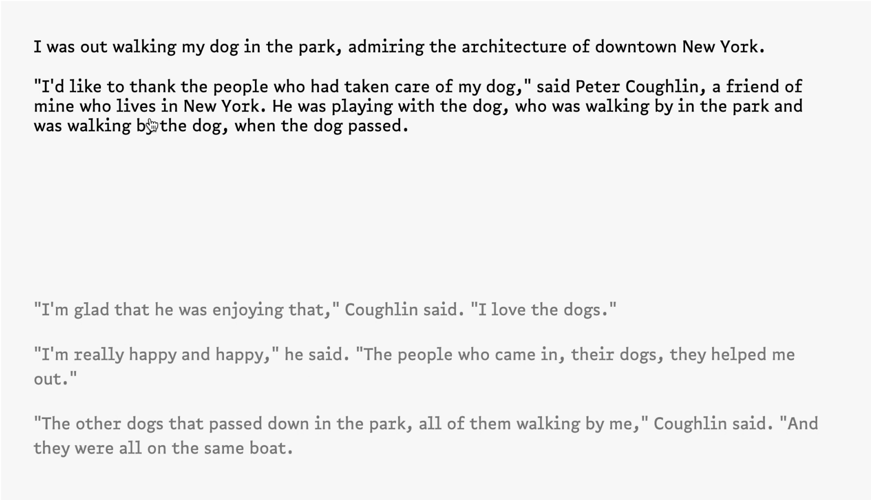


Figure 4.17: This time, the user-reader decided to explore another possible continuation of the previously explored axis, forming a new one.

4.5.2 The Aesthetics of a Story

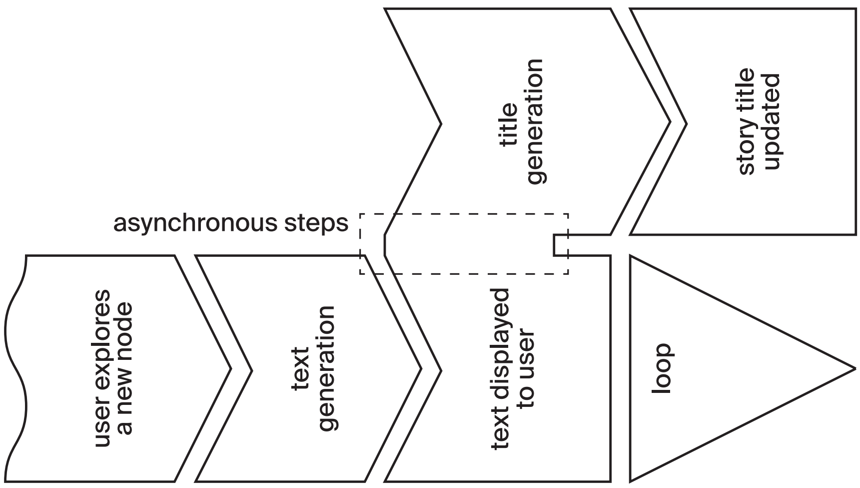
The initial *Multiverse* prototype was bare bones in order to focus on the core interaction of how to do bidirectional navigation of a cybertextual story. Testers found the navigation to be intuitive enough, but had a difficult time relating to the work as a piece of literature. There are plenty of reason for such a reaction. As readers, we are not used to working with bidirectional cybertexts, especially when they come one sentence at a time. Testers also expressed that the prototype lacked of the unique aesthetic qualities that together comprise a literary work.

A book has a physical presence, a vivid title and the bearing of the author, attributes that affords the reader to situate the work in a wider context, as something more than the sum of its words. The first prototype lacks most of these qualities, a void accentuated by a highly ephemeral implementation that discards the work the moment the browser tab is closed. Once this shortcoming had been identified, it might be turned into a design opportunity.

This iteration of Multiverse adds a library that automatically saves the stories that the user-reader has explored in a separate page. It is carried over across sessions and browser tabs, giving each new story a context and home. While it is tempting to fill such a place with all sorts of skeuomorphic designs, I settled for a simple gird of stories, see fig. 4.19. It would be difficult to properly store and organize the kind of anonymous cybertexts produced by the first prototype, so in adding a library we must necessarily add a title and author.

The title generation is done through a process called *summarization*. This classic natural language processing task deals with the problem of finding a subset of a presumably longer text that still retains its essence, i.e. summary (Alpaydin, 2014). Google's BERT model excels at this task (Miller, 2019), so to generate a title we use it through the HuggingFace Transformers summarization pipeline, tweaking the parameters to produce a five to ten word sentence as output, thus forming a kind of ad-hoc title. This appropriation of summarization is far

Figure 4.18: An expanded neural pipeline, including the summarization/title generation.



from a perfect generator, but it adds the legitimacy of a title and an exiting micro-interaction where the user-reader can see the results of their explorations change the context of the story. This addition extends our neural pipeline, see fig. 4.18.

This approach to title generation hosts a comparatively small, but rather important design choice: Should each unique axis have its own title or does a story have a single name that reflects the entirety of the explored literary space? The first option certainly has its allure, accentuating the literary qualities found only in that particular axis. On the other hand, having a single monolithic title contributes to a larger sense of the story as a contained literary work. After experimenting with both, I decided to go for the ladder, as it aligned closer to this iteration's goal of solidifying the identity of an individual story. The title is refreshed every time a new node is explored, thus re-defining any pre-conceptions of what a name contains.

The question of authorship as such in cybertexts is a thought provoking topic it itself and is covered in more detail in the discussion (Section 5.3). In the name of moving the design forward, *Multiverse* contributes a dual authorship to whatever name the user-reader enters and the artificial neural network it is currently writing with.

These features are supported by an expanded interface, fashioned in a language familiar to that of a web app. Aside from the necessity of navigation between the library and the story view, the elaborated user interface was experienced to give more "heft" to the rest of the interactions by situating them in a coherent digital space.

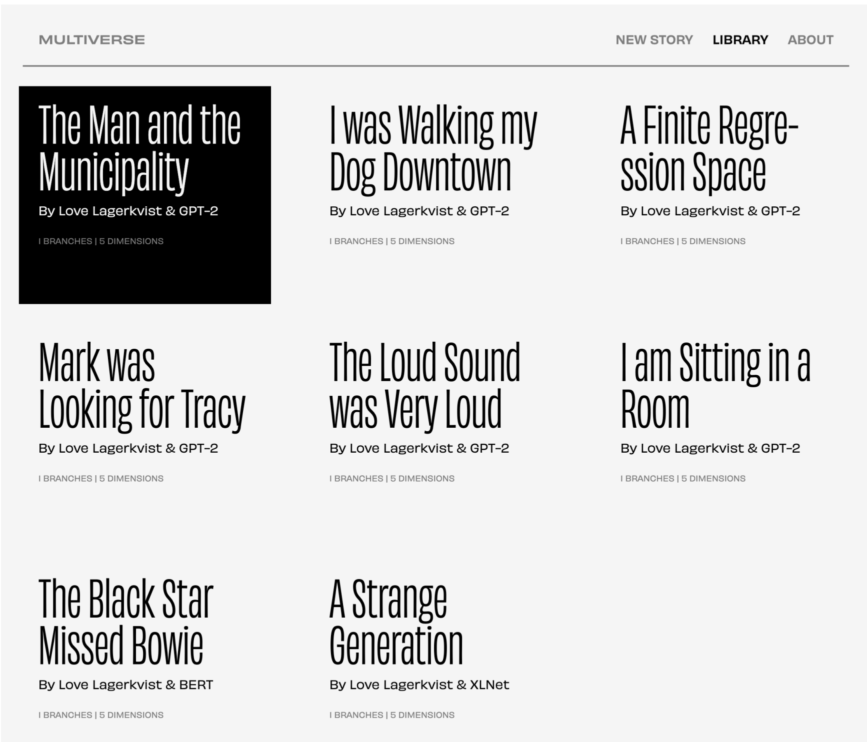


Figure 4.19: The Library view, showing a number of stories. The first story is highlighted.

Figure 4.20: The elaborated story view, showing the navigational menu (top) and title/author/date metadata in a sidebar.

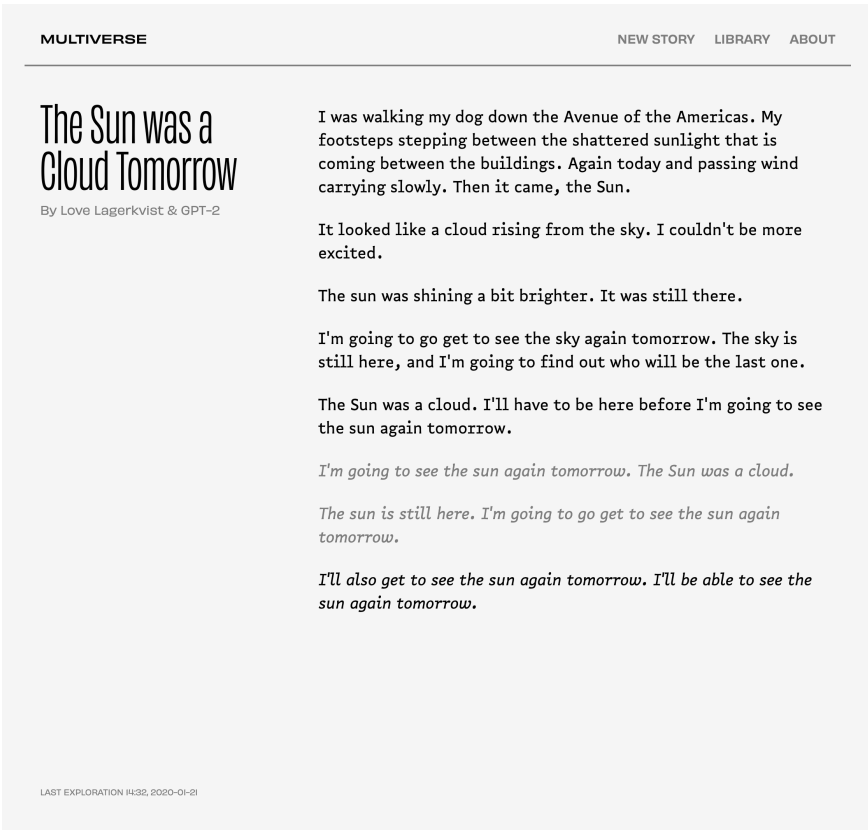


Figure 4.21: Deeper down the axis from fig. 4.20. As the story pro-gresse, the title is updated to reflect the new content.

4.5.3 Navigating and Visualizing Bidirectional Literature

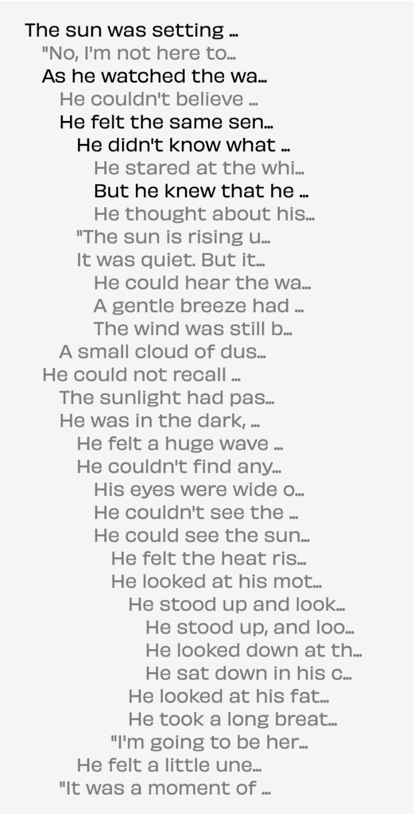


Figure 4.22: List tree visualization with the start of a node's as label.

The secondary research questions of this thesis (Section 1.2.2) concerns the formulation of "novel interactions" for generative literature. The *Multiverse* prototype explores this through a bidirectional, multi axis story space. Testers found vertical navigation of the tree rather intuitive, they struggled with understanding the lateral and parallel movements made possible by the system. To afford a better understanding of the story as a larger literary space, I investigated ways to dynamically visualize stories as the user was exploring them.

Trees very common structures when visualizing hierarchical relationships. There exists a rich body of work in regard to their representation, ranging from abstract three dimensional spaces to literal trees (Schulz, 2011). In order to make an informed selection, a set of necessary attributes was identified. A successful visualization has to:

- be compact enough to fit in sidebar-area of the user interface.
- be intuitive and at least a vaguely familiar to the user.
- work not only as a visual element, but also as a tool for navigation.

Evaluating the multitude of ways in which one can draw a tree against this criteria rules out many of the more elaborate representations. After consulting testers, a decision was made to explore the humble list tree, a simple two-dimensional representation than can be made rather compact. Users should already be familiar with the gist of their appearance and interaction, as they are a common element of other computer interfaces, like file explorers (Song et al., 2006).

When the information of what one wants to represent is to large to coherently fit in view, visualizations often use shorter labels in place of the thing as such. Early explorations showed that the nodes of a story in *Multiverse* can not be reduced to an obvious label. Drawing the entire sentence occupies to much space, and steals visual attention from the relations that the list tree aims to express (fig. 4.22).

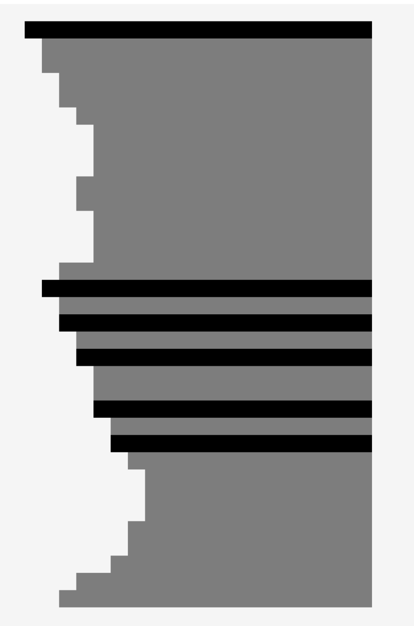


Figure 4.23: List tree visualization with bars representing node hierarchy. The back bars show the current axis.

Recognizing that the hierarchical relation between the nodes is our main carrier of information, one might strip away any unrelated elements. This leaves us with something along the lines of an indentation tree (Knuth, 1968), where the levels are drawn as receding horizontal bars (fig. 4.23). This visualization fits our requirement of being compact and succinct, but the lack of labels. Additionally, the small number of different visual node states confused testers. Due to the low contrast between levels, one easily loses sense of place.

The final iteration (fig. 4.24) represents instead represents each node with a circle, occupying one of four states:

filled, black	Explored and part of the current axis.
filled, dimmed	Explored, but not part of the current axis.
empty	Unexplored.
semi-filled	Unexplored and highlighted.

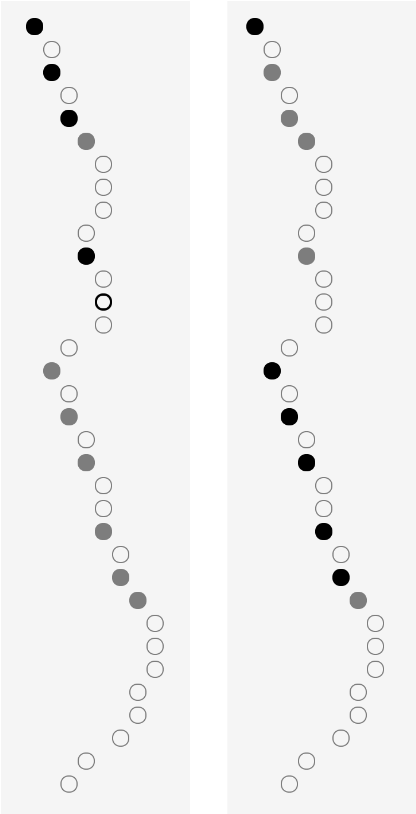


Figure 4.24: List tree visualization using circles. Black circles show the active axis, gray solids explored nodes empty circles unexplored children. The map to the left is highlighting an unexplored child node, while the one on the right shows the user peering down a deep axis.

The circles are indented based on their place in the hierarchy, resulting in an “inward-downward” type motion as more nodes in the story are explored (fig. 4.24). The user-reader interacts with the map by hovering over the circles, highlighting the axis it terminates. The interactions with the tree are reflected in the story view, allowing the user-reader to “scrub” the explored space in the of the story. This navigational flow goes both ways — vertical navigation in the main text window predictably highlight the corresponding axis in the lager context of the story. Clicking on a node in either story or map selects that axis.

This visualization is still very abstract, but testers found it far more intuitive once the highlighted axis was reflected in the story view, and vice versa. This interaction is hard to show in static images alone, a large part of why it works is the feedback is immediate and synchronized over the two story areas.

That said, this iteration also has a lot of room for improvement. The lack of proper labels was still a hurdle for testers. The depth of the map can also get out of control very quickly, due to the fact that every explored node spans three children, resulting in large vertical growth and an overflow of empty circles. On one hand, this successfully serves the purpose of illustrating the infinite space of alepsis unique to generative cybertextual literature. On the other, it can be experienced as rather noisy, to one tester even slightly stressful. Recognizing that the interactive visualization of bidirectional literature is a topic worthy a master thesis of its own, I had to stop the iteration at some, fairly arbitrary, point. Possible future directions are discussed in Section 5.2.1.

Figure 4.25: The full story UI, complete with the final tree list.

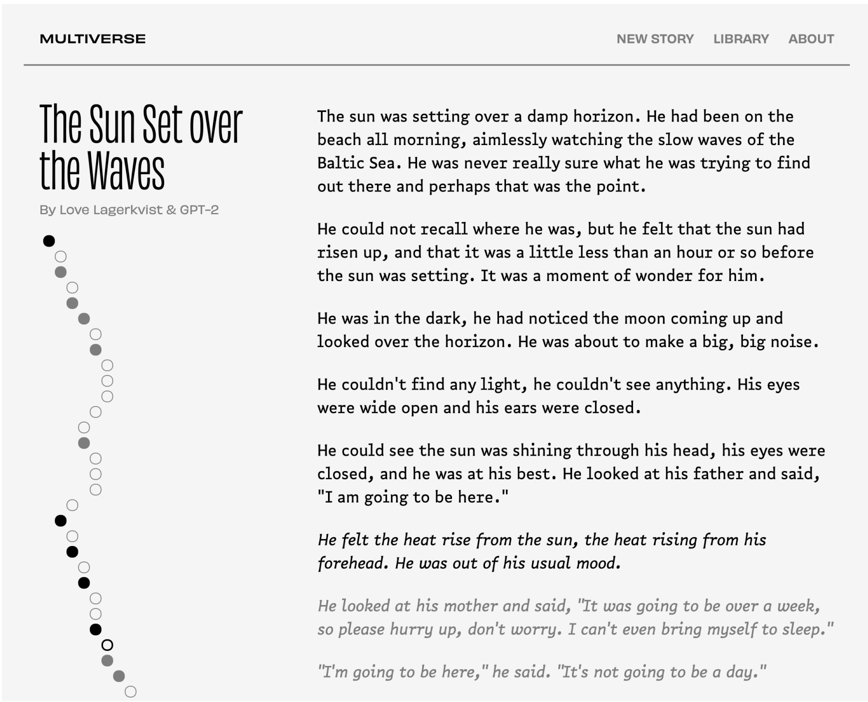
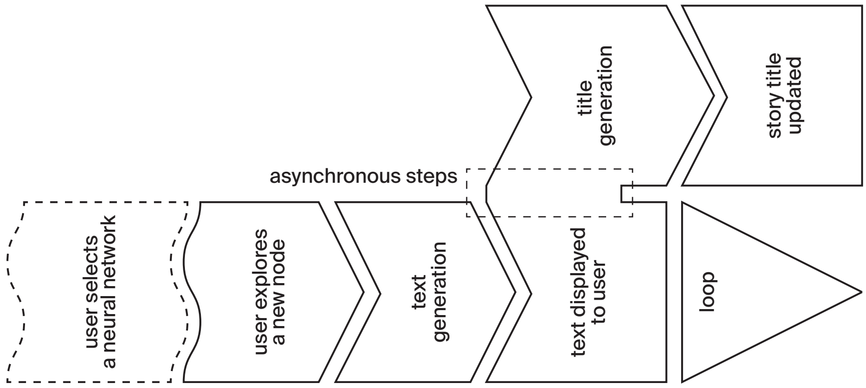


Figure 4.26: The pipeline of the final iteration, with selection over artificial neural network.



4.5.4 Nerual Networks as Design Material

The last iteration of the *Multiverse* design process deals with ways to further explore the use of machine learning and artificial neural networks as design material. The earlier prototypes treated whatever generated the texts as a black box, opaquely supping the user-reader with continuations from *somewhere*. Here, I investigate ways of opening the black box by giving the user-reader agency over the current model.

The first part of this neural unpacking is done though an elaborated "New Story" screen that features input fields for the user-readers name, the writing prompt and a selection of the artificial neural networks to act as ones "writing partner" (fig. 4.28). This choice is not final, as the user-reader is able to change their partner at any time during the text generation, through a similar interface to the on in the New Story screen. After discussing the idea with testers, a decision made not to allow for users to "re-roll" already explored nodes, even with another artificial neural network. Such a feature devalues the decision to explore and the legitimacy of the generated text. Affording such powers runs the risk of transforming the prototype into a toy where the user-reader keeps generating new continuations until they get something that fits their expectation, in stark opposition to the goal of being a tool for cybertextual co-exploration literary space.

Testers expressed that the simple act of picking the generatory model gave them a greatly increased sense of agency over the process, peeling away some of the outer layers of the black box. This adds another dimension to the neural pipeline, as visualized in fig. 4.26.

One could easily imagine a future version that affords the user-reader with a much richer selection of artificial neural networks, including the possibles to browse models specialized on different writing styles or genres through transfer learning. If none of the options felt satisfactory, there could be a sub-interface for user-readers to train their own models, using some kind of interactive machine learning.

As the prototype has grown, it felt appropriate to add a landing page as a way to tie together the interactions into a coherent system. When the user-reader enters *Multiverse*, it is faced with a choice between staring a new story or resuming progress on something in the Library (fig. 4.27). This feature further contributes to the sense of place as discussed in Section 4.5.2. It was also intended to assist in self guided online based usability tests, had the time allowed such activities.

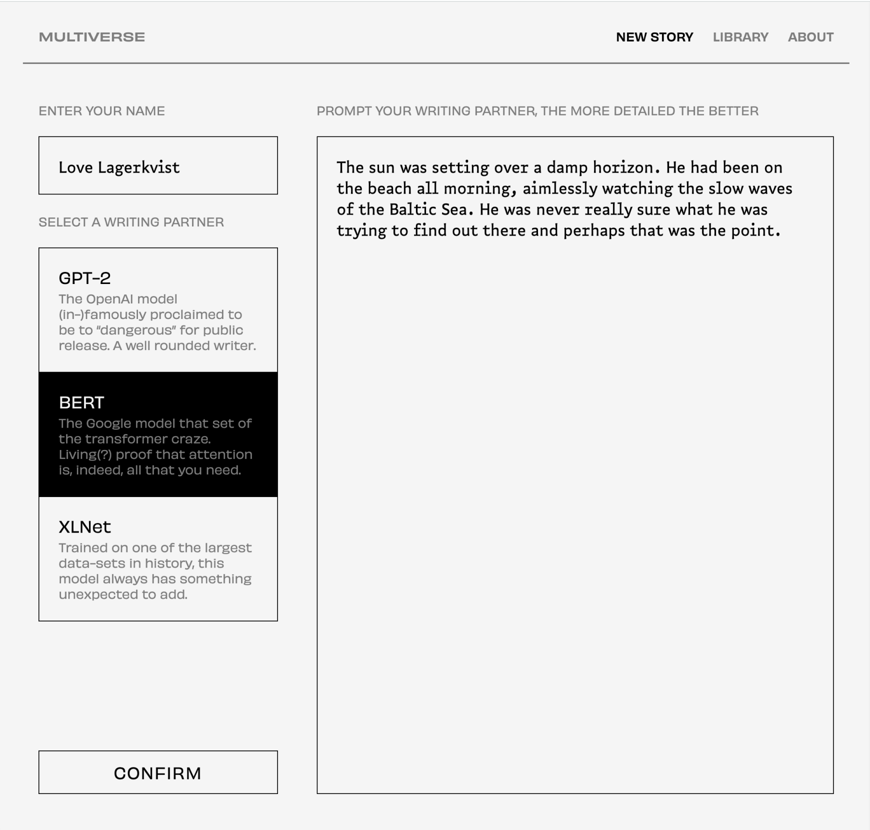
The final prototype serves as good test case for the analysis of machine learning attributes sketched in Section 4.3. Cross-referencing with Table 4.2 makes apparent the influences-of and relation-to it *AI-Dungeon 2*, another web-based, text driven application. The two works are differentiated by the way in which they model the user-reader — artificial neural network relation. In *AI-Dungeon 2*, the writing partner is hidden behind a black box, intangible to the user. This is very different from the final iteration of *Multiverse*, which affords the user-reader control of who they want to tell the story with.

Table 4.3: The ML attributes of the final *Multiverse* prototype.

alepsis	tangible, visualized
authenticity	surprising
features	novel
interaction	two-way, user-reader
graspability	editable, some control over the artificial neural network
locality	distant, intangiable
subjectivity	technopomorphic

Figure 4.27: The landing page of the prototype, guiding user-readers. The “New Story” option is highlighted.

Figure 4.28: The “New story” screen, prompting the user for required information and their choice of artificial neural network.



Discussion

This chapter reviews the Research through Design-process presented in Chapter 4. It examines the two methods proposed, reflects upon the Multiverse prototype through different conceptual lenses and then ends with a discussion on the implications of applied machine learning on literature.

5.1 Evaluating Neural Plumbing & ML-Attributes

The two proposed methods, Neural Plumbing and machine learning-attributes, were developed in part as an investigation into the potential ways of appropriating existing interaction design methods to the area of machine learning. In addition to this grander goal, they were a very helpful tool in articulating and developing my own understanding of the subject, at best affording a means to communicate these ideas with other potential stakeholders. Further to this regard, they would have been much strengthened by a more rigorous external user research. I had originally planned for a workshop session to take place with a set of identified expert users (fig. 5.1), but changes in the participants scheduling sadly rendered this opportunity void.

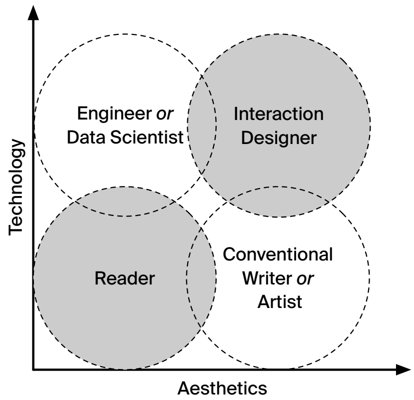


Figure 5.1: Diagram showing the spectrum of identified expert users.

The limited amount of feedback I did get from users indicated that participants found it difficult to sketch with the pipelines from a completely blank slate. A reason for this might be a result of the bottom-up, case-based methodology I used to construct the workshop. In order to formulate a stronger selection of elements, future researchers might apply the neural pipelines to a large set of cases and from that extrapolate a better selection of the most frequent elements.

These methods are still early in their development, dull tools that require a user to know exactly where to cut. Despite this limitation, I found them to be useful in analyzing and communicating what machine learning does, and I do believe that they, or some variations thereof, might play an important role in the quest to find ways for interaction designers to engage with machine learning as a design material.

5.2 Reflections on the Multiverse

The *Multiverse* prototype is the centerpiece of the Research-through-design process. Conceived as a functional synthesis between the literary theories discussed in Sections 2.3 and 2.4 and the design methods discussed in the section above, it explores the both possible ways thorough which interaction designers might engage with machine learning as a design material by applying them to a context of interactive generative literature. This section evaluates four of the main concepts explored: Visualization of bidirectional stories, the mental and computational structures that afford an understanding of the subject, possible future extensions of the prototype and the use of machine learning as a design material.

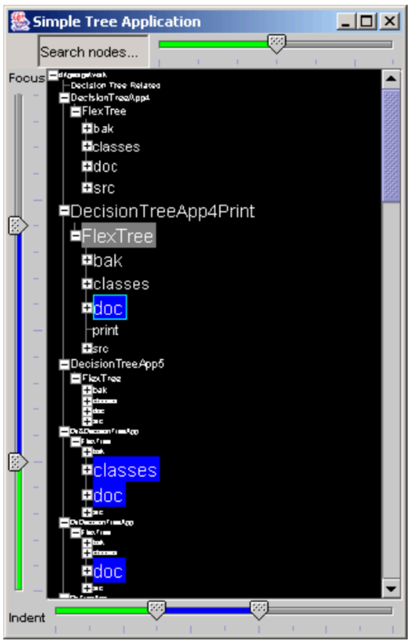


Figure 5.2: A version of the lens tree.
© Song et al., 2006.

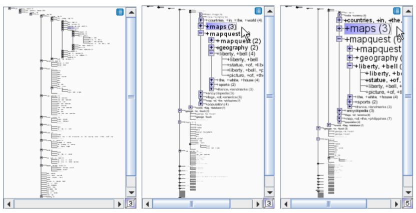


Figure 5.3: A prototype of the fisheye tree.
© Tominski et al., 2006.

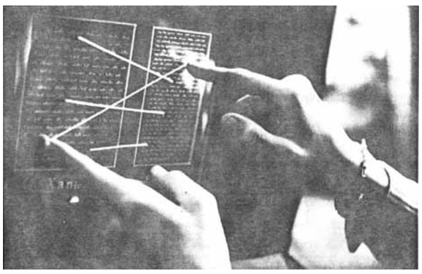


Figure 5.4: A mock-up of "transpointing" the relations of content in two parallel Xanadu hypertexts (Nelson, 1999).
© Ted Nelson, 1972.

5.2.1 Exploring Bidirectional Stories

A highlight of the *Multiverse* processes of was the designing of the interactive visualization, described in Section 4.5.3. As the project developed, it became clear that the formulation of an intuitive, yet powerful, visual representation of the literary space constitutes the main interaction design challenge. The coupling of visualization and navigation does introduce complexity into the system, but when successful, it affords a very direct interaction with the literary material. The final assemblage of circles (fig. 4.24) still has a long way to go in achieving such direct tactility. Possible future directions include an investigation in richer tree list interfaces, following concepts like the Lens Tree (fig. 5.2) (Song et al., 2006), which introduces rich hierarchical context and focus, or the Fisheye Tree (fig. 5.3) (Tominski et al., 2006), featuring different "lenses" the users can apply to further filter their view.

These elaborated visions still carry with them the problematic high level of abstraction indebted by the separation of navigation from content. They require the user to form the semantic connections between the two on their own, even though both are known to the system. The issue of representing complex non-linearity is not unique to bidirectional generative literature. Hypertext has grappled with the same problem since its (practical) inspection in the 1960s. Far from idealistic precursors to the World Wide Web, the designers of these systems understood that we could use the boundless memory afforded us by computers to find brand new ways of thinking and reading.

Ted Nelson's now (in)famous Xanadu features a concept called "transpointing" that visualizes connections in one or more parallel documents (Nelson, 1999). This visualization allows for a "parallelism" of through the explication of interrelation and subtext between the examined items. In most Xanadu prototypes, this was achieved though a literal, interactive connection between the two items (fig 5.4). Transposing the concept unto a hypothetical version of *Multiverse* could look something like fig. 5.5, with direct connections between the nodes in the story view and the circular map. One could even imagine these connections as interactive objects in themselves, affording the ability to form new connections between previously unconnected paths or the creation of new axes though re-arrangement of the nodes. Supporting features of such a development might include facilities for “bookmarking” axes (see fig. 5.5) or the ability to track the origin of a node, similar to the “transclusions” of Xanadu (Nelson, 1982).

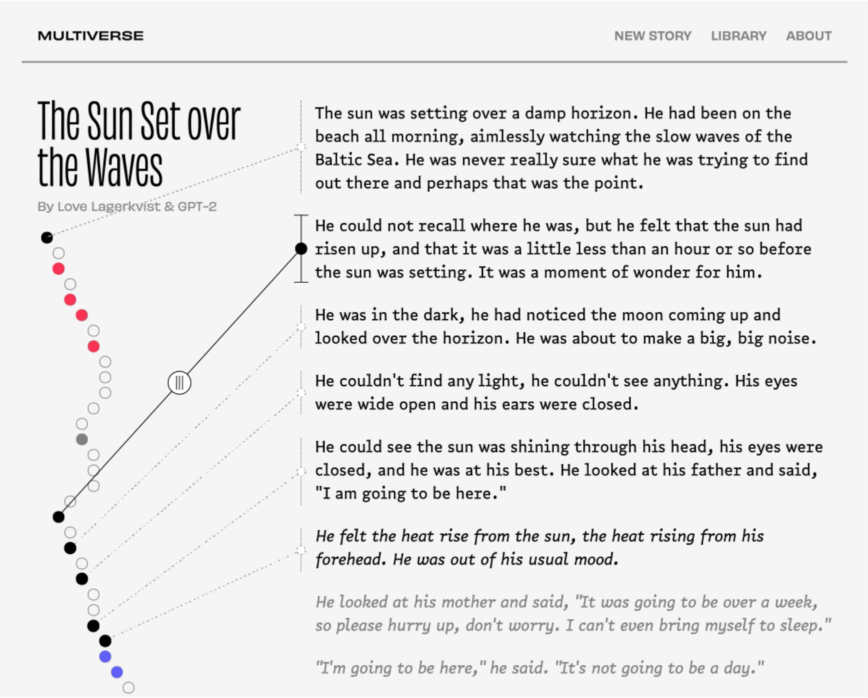


Figure 5.5: Mock-up of potential future extension of *Multiverse*. The Xanadu principle of “transpointing” applied along with a rudimentary, color coded bookmarking function (in red and blue). A bookmark uniquely labels an axis allowing the user-reader to mark notable paths through the story.

When we regard the multiversal node as something in itself, one might plot a research path that translates them into artifacts. We should not couple the principles that underlie generative literature with the traditional typographic literary realm. In fact, one could make the argument that generative literature through machine learning operates in an oral mode, in the sense Ong understands the term through his second orality — a new culture that depends on the literary form but express it though other, conversational mediums (Ong, 1971). The collaborative exploration of literary space takes place in constant dialogue with the artificial neural network, a process one could easily imagine situated in a media art installation similar to Strebe’s prayer, expanding into sound or projected visuals. Another direction could be an embodiment of the nodes themselves, physical objects though which the user-reader manipulate, re-arrange or progress the narrative, perhaps inspired by the ReacTable (fig. 5.6) (Jordà et al., 2007).

Figure 5.6: The node based, multi-touch interface of the RactTable. Image by Daniel Williams - *The Reactable*. CC BY-SA 2.0

Just as the cybertextual medium transforms the author into a meta-author, it brings with it the possibles of a meta-story, understood as parts of the narrative explored by the user-reader. This is not necessarily equal to the complete set of nodes generated by the artificial neural network. This layer is indirectly present in the current iteration of the *Multiverse* prototype, as it must be, but one could image a future development where its potentials are explored in depth. One might try and understand this subconscious process as a cognitive one, analysis (or interfacing) with the eye-movements of a reading-working user-reader. Another approach could be a social one, expanding the system to afford multiple user-readers to explore the literary space in parallel, sharing meta-stories during the process.

5.2.2 A Story is not a Tree

At a fundamental level, a representation is limited by its subject. *Multiverse* structures the literary space as a tree, a hierarchical and linear understanding of relationships. This design choice rests on seemingly uncontroversial assumptions (a story can only be written as a radiation from an ancestral prompt) and practicalities (trees are a widely understood building block of computer science). However, it is important to critically examine what other consequences such an important structural decision entail.

Many of the problems outlined with the visualization of a bidirectional story can be traced back to the fact that deeply spanning hierarchical structures grow almost exponentially. When such a span is anonymous, as opposed to labeled or measured, difficulty of representation grows in tandem. Additionally, trees falter in situations where a relation might be multi-dimensional or have multiple linearity, as is very much the case with Balpe's infinite literary space (Balpe, 2007).

Mathematician and architectural theorist Christopher Alexander identified these conceptual limitations in his 1964 essay *A City is not a Tree*. In it, he argues against the simplified, instrumentalized view of the city as something artificial that might be understood, and thus controlled, through a hierarchical mode of analysis. Instead, he argues, we should adapt the structure of a semi-lattice, whose axiom defines it as a collections of overlapping sets (Alexander, 1964). This foundational principle of the overlap allows for a far richer model of the bidirectional relationships that make up a city, or in our case, a story. If we were to allow for non-linear construction of literary paths through the generative space, we might model these as sets of sentences, whose overlaps together constitute a spatial understanding of what a story might be (fig 5.7). Such a literary space would, by all accounts, be difficult to for a human conceive of, partly due to our prior conception of a story as a linear diegetic axis. Artificial neural networks don't have such limitations and might therefore be useful materials for exploring the outer possibilities of multidimensional literary spaces.

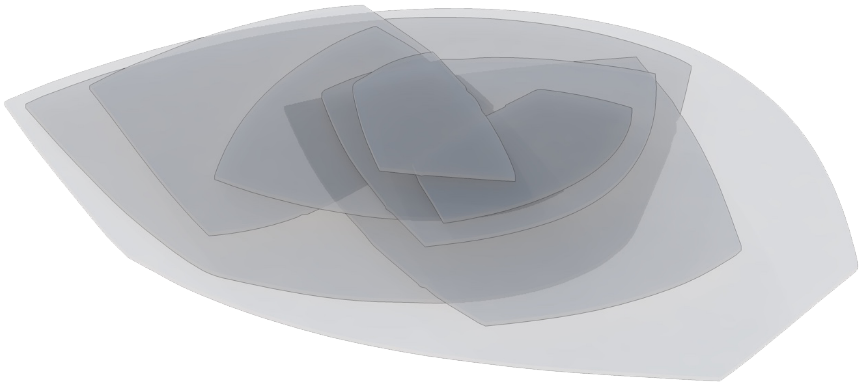


Figure 5.7: An overlapping literary space of a bidirectional story, visualized as semi-lattice.

5.2.3 Machine Learning as a Design Material

The promise of à la carte, human-level text generation delivered by the current generation of transformer style artificial neural networks instigated my investigation of how these might afford the interaction designer the opportunity to work with machine learning as a design material. As the project progressed, I tried to distill my background research and technical knowledge into formalized methods interactions designers might use when approaching the subject. However, as I discussed the project with colleges, it was always the application of machine learning found in the *Multiverse* prototype that managed to intrigue, not the methods. In a hindsight, such a result seems almost obvious. Compared to the vivid interactions of the prototype, the proposed workshops seem hopelessly dull.

Interaction design is many respects a field of application, therefore, if we want to show the potentials of machine learning as a design material, it necessarily has to go through the route of practice. By making the source code of the *Multiverse* prototype freely available, I invite other designers to further explore the many uncharted potentials of interactive generative literature. This chapter outlines a few paths forwards, so I look forward to see how other minds will fork it into completely diffident directions.

→ https://github.com/motform/multiverse

5.3 Generative Literature & The Absolute Death of the Author

One of the central questions in generative literature is the attribution of authorship, an equation that gets even more complex when factoring in the cybertexutal mode of interaction. Is "true" author the system designer, even when the literary work is realized only through the action of the user-reader?

Here, I make the argument that generative literature produced by machine learning through artificial neural networks mark an *absolute death of the author*. If Barthes affords the execution of the author on the grounds that that text is nothing but "tissue of citations, resulting from the thousand sources of culture" (Barthes, 1967), it only follows that we consider generative literature through machine learning the absolute rendition of this concept. The works produced by such artificial neural networks are literal tissues of citations, necessarily outputs of nothing but the input corpus and the rules of the model. As such, Barthes' claimed execution as sentenced by a jury of reasoning becomes *relative* to the *absolute* mathematical mode of the new literary model.

This is not as far fetched of a claim as it might first sound. Recalling the intertwined historical origins of both generative literature, *the death author* and postmodern literature found in the postwar art and theory, it is possible to trace the invariability of such a statement. The historical connection is made explicit by the fact that the first English translation of Barthes' essay was published in *Aspen*, the groundbreaking multimedia magazine, published from 1965 to 1971, that also featured contributions from likes of Marshall McLuhan, John Cage and Le Monte Young (Allen, 2011), the latter two of which are artists known for their systematic approaches to art. The comparison to a post-modern mode of writing goes further, considering the artificial neural network as a set of only signifiers, whole relation is reduced to intuitive discovery of statistical patterns. This disconnection of the signified situates any work the network produces in a state of Baudrillardian hyperreality, which he identifies as the main mode of late capitalism (Baudrillard, 1994).

Just as Barthes' relative removal of the author took with it the need of "deciphering" a text (Barthes, 1967) does its absolute ditto necessitate a new mode of reading. Through the cybertextual perspective we know this to be the birth of the user-reader, whose life is only possible through an interactive logic.

Conclusion

> *It is therefore necessary to grasp the concept of artificial general intelligence not merely as a technoscientific idea, but more fundamentally as a concept belonging to a thought or form of intelligence that treats its very possibility as an explicit opportunity to pierce through the horizon of its givenness: it does not matter what it currently is; what matters is what can be done — all relevant things considered — to expand and build on this possibility.*
>
> Reza Negarestani, *Intelligence and Spirit*

At the time of writing this thesis, the Covid-19 pandemic is sweeping the world. In its wake lies the largest wave of unemployment and precarity modern civilization has ever seen. Artificial intelligence and machine learning have certain roles to play in this new world order, as politicized vehicles for mass scale surveillance and automation — all behind a deceiving veil of "magic" and technobabble.

My hope is for this thesis to serve its small part in what necessarily must be a long and difficult process of opening up machine learning to the practice of interaction design. In order for the field to flourish, there need to be an increased diversity. Interaction designers bring with them a unique skills and sensibilities vital in introducing the human into the artificial.

I hope to leave the reader of this thesis with an increased, critical understanding of the actual attributes of machine learning, devoid of fake magic. Equally, I hope to leave the reader with thoughts and ideas for the many novel ways by which we might approach it as a design material. However, I must also leave the reader to seriously consider the new aesthetic spaces machine learning might uncover if we allow it agency to act as something novel and interesting in itself, rather than a slightly confused simulation of the human mind. Lets end this imitation game and start a playful dialogue.

Acknowledgments

Thanks to Kristina, my amazing partner who have supported me throughout the entire process. Without you, this would have not been possible.

Thanks to Anne-Marie Hansen for supervising me during this thesis, your input was invaluable for keeping the thesis grounded in practice.

Thanks to Jens Pedersen for his invaluable input and enthusiasm on the philosophical inquery of artificial intelligence.

Thanks to Kyle and Linhda of the podcast *Data Skeptic*. Without your pedagogic explanations of data science and machine learning, I would be yet another confused interaction designer.

Last but not least, a huge thanks to my all of my fellow classmates.
We have fun.

This thesis was written in Emacs 26.3 and typeset in Swiss Typefaces' *Suisse Int'l* using XƎLATEX. The source for and the software discussed can be found in the project repository, freely licensed under GPLv3.

https://github.com/motform/multiverse

Glossary

A

alepsis The situation in generative literature where there is an infinite number of possible continuations to any point of a story (Balpe, 2007).
artificial intelligence The study of intelligence in synthetic beings. Sometimes study of "rational agents" that employs data about its context to take actions that will achieve its goals (Norvig & Russell, 2009).
artificial neural network Computing systems inspired by the biological neural networks that form animal brains. Made up of units (artificial neurons) that are able to transmit signals amongst each other in a synapse like manner (McCulloch & Pitts, 1943). Characteristically able to "learn" tasks by considering examples without being taught specific rules.

C

corpus In linguistics and NLP a structured set of texts, often of generous size. Used for statistical analysis or to train artificial neural network.

D

deep learning A broad set of machine learning architectures that uses artificial neural networks. Characteristically uses multiple layers as a tool for pattern recognition and feature exaction (Alpaydin, 2014).
diegetic axis Gérard Genette's definition of textual narrative structure as organized around diegesis, necessarily with a set beginning and end (Genette, 1983).

G

generative literature A movement in literature that concerns with "the production of continuously changing literary texts by means of a dictionary, a set of rules, and/or the use of algorithms" (Balpe, 2007).
GPT-2 A pre-trained machine learning model released and developed by OpenAI in 2019. Predicts the next word based of input text. Uses a transformer architecture (Radford et al., 2019).

H

Human-Centered Machine Learning The research area that looks at ML systems in terms of human goals (Ramos et al., 2019).

I

interactive machine learning A process that allows for non-experts to apply domain knowledge to machine learning systems through some kind of user interface (Dudley & Kristensson, 2018).

M

machine learning Coined in 1959 by Arthur L. Samuel (Samuel, 1959). The sub-field in artificial intelligence that deals with "the study of computer algorithms that improve automatically through experience" (Mitchell, 1997). The process of improving the network is called training, which is performed by applying a data set to the model in one/a mixture of three modes: supervised, unsupervised or reinforced.
model In machine learning, the model is the mathematical representation of training data produced by the algorithm to make predictions (Alpaydin, 2014).

N

natural language processing The study of how computers interact with natural (human) languages. Includes sub-fields that deal with language syntax, semantics, discourse or speech recognition.

P

pipeline A sequential chain of data processing elements where the output of each part is the input to the next. Used on multiple scales in machine learning, data science and most fields of computer science.
pre-trained model A machine learning model that comes trained on an (often very large) data set. Can serve as a foundation when performing transfer learning or used as is to perform a both general and specialized tasks (Devlin et al., 2018).

R

reinforcement learning One of the three basic learning paradigms in ML. In reinforcement learning, the network learns through positive or negative feedback (reinforcement) as judged by a set of rules. The network finds the patterns that resulted in the response and adjusts accordingly (Norvig & Russell, 2009).
Research-through-design A research methodology that uses designerly methods like a design process and drifting (Krogh et al., 2015) as a form of inquiry (Zimmerman et al., 2010).

S

sequence In machine learning, sequence denotes data that has a defined ordering and a variable length. Representations of natural languages are almost always sequences (Alpaydin, 2014).
supervised learning One of the three basic learning paradigms in ML. In supervised learning, the network learns by searching for the optimal way to re-produce input-output pairs. Requires a labeled data set to extrapolate patterns, which makes it a potentially labor intensive process (Norvig & Russell, 2009). Notably used in interaction design contexts with Rebecca Fiebrink's Wekinator .

T

transfer learning The process by which the patterns derived from one learning problem is used to improve or serve as the foundation for another (Ventura & Warnick, 2007).
transformer In machine learning, an architecture used primarily in natural language processing. Notable for being able to handle large amounts or sequences data, like natural language, out of order, allowing for parallel computation. The current state-of-the-art (Vaswani et al., 2017).

U

unsupervised learning One of the three basic learning paradigms in ML. In unsupervised learning, the network learns without explicit feedback or labeling in the data set (Norvig & Russell, 2009). Used in modern transformer based architectures, like GPT-2 and BERT.

V

vector In machine learning, vector denotes data of a fixed length in a vector space. Representations of images or specific output values like sentiment are almost always vectors (Alpaydin, 2014).

Bibliography

Aarseth, E. J. (1997). *Cybertext: Perspectives on ergodic literature*. JHU Press.

Alec, R., Jeffrey, W., Dario, A., Daniela, A., Jack, C., Miles, B., & Ilya, S. (2019). *Better language models and their implications*. Retrieved April 13, 2020, from https://openai.com/blog/better-language-models/

Alexander, C. (1964). A city is not a tree. *Architectural Forum*, *124*.

Allen, G. (2011). *Artists' magazines: An alternative space for art*. MIT Press.

Alpaydin, E. (2014). *Introduction to machine learning*. MIT Press.

Arntz, M., Gregory, T., & Zierahn, U. (2017). Revisiting the risk of automation. *Economics Letters*, *159*, 157–160.

Balpe, J.-P. (2007). Principles and processes of generative literature. In P. Gendolla & J. Schäfer (Eds.), *The aesthetics of net literature: Writing, reading and playing in programmable media* (pp. 309–316). New Jersey: Transaction Publishers.

Barthes, R. (1967). The death of the author. *Aspen*, (6).

Baudrillard, J. (1994). *Simulacra and simulation* (F. S. Glaser, Trans.). University of Michigan press.

Bengio, Y. (2009). Learning deep architectures for ai. *Found. Trends Mach. Learn.*, *2*(1), 1–127. https://doi.org/10.1561/2200000006

Billewicz, A. (2017). *Study of a relationship: Designerly explorations of machine learning algorithms* (Master's thesis). Malmö University.

Blashki, K. (2013). *Emerging research and trends in interactivity and the human-computer interface*. IGI Global.

Borges, J.-L. (1992). For a version of the i ching (R. Mezey, Trans.). *The New York Review of Books*.

Bughin, J., Chui, M., & Manyika, J. (2010). Clouds, big data, and smart assets: Ten tech-enabled business trends to watch. *McKinsey quarterly*, *56*(1), 75–86.

Buxton, B. (2010). *Sketching user experiences: Getting the design right and the right design*. Morgan kaufmann.

Cloud, G. (2020). *Ai and machine learning products*. Retrieved April 29, 2020, from https://cloud.google.com/products/ai

Cover, A. B. (1974). Interview with philip k. dick. *Vertex*, *1*(6).

Cunningham, S. J., & Jones, M. (2005). Autoethnography: A tool for practice and education, In *Proceedings of the 6th acm sigchi new zealand chapter's international conference on computer-human interaction: Making chi natural*, Auckland, New Zealand, Association for Computing Machinery. https://doi.org/10.1145/1073943.1073944

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding.

DeWitt, B. S., & Graham, N. (1973). *The many worlds interpretation of quantum mechanics*. Princeton University Press.

Dove, G., Halskov, K., Forlizzi, J., & Zimmerman, J. (2017). Ux design innovation: Challenges for working with machine learning as a design material, In *Proceedings of the 2017 chi conference on human factors in computing systems*, Denver, Colorado, USA, Association for Computing Machinery. https://doi.org/10.1145/3025453.3025739

Dudley, J. J., & Kristensson, P. O. (2018). A review of user interface design for interactive machine learning. *ACM Trans. Interact. Intell. Syst.*, *8*(2). https://doi.org/10.1145/3185517

Eagleton, T. (2011). *Literary theory: An introduction*. John Wiley & Sons.
Fiebrink, R. (2010). Real-time interaction with supervised learning, In *Chi '10 extended abstracts on human factors in computing systems*, Atlanta, Georgia, USA, Association for Computing Machinery. https://doi.org/10.1145/1753846.1753889
Fiebrink, R. (2020). *Walkthrough | wekinator*. Retrieved April 27, 2020, from http://www.wekinator.org/walkthrough/
Fleischer, R. (2018). *Från 1930-talets maskinkultur till vad vi på 2010-talet kallar artificiell intelligens*. Retrieved April 6, 2020, from https://copyriot.se/2018/02/27/fran-1930-talets-maskinkultur-till-vad-vi-pa-2010-talet-kallar-artificiell-intelligens/
Galanter, P. (2003). What is generative art? complexity theory as a context for art theory, In *In ga2003–6th generative art conference*. Citeseer.
García-Martín, E., Rodrigues, C. F., Riley, G., & Grahn, H. (2019). Estimation of energy consumption in machine learning. *Journal of Parallel and Distributed Computing*, *134*, 75–88. https://doi.org/https://doi.org/10.1016/j.jpdc.2019.07.007
Genette, G. (1983). *Narrative discourse: An essay in method* (Vol. 3). Cornell University Press.
Goodman, E., Kuniavsky, M., & Moed, A. (2012). *Observing the user experience: A practitioner's guide to user research*. Elsevier.
Hanington, B., & Martin, B. (2019). *Universal methods of design expanded and revised: 125 ways to research complex problems, develop innovative ideas, and design effective solutions*. Rockport Publishers.
Houde, S., & Hill, C. (1997). What do prototypes prototype? In *Handbook of human-computer interaction* (pp. 367–381). Elsevier.
Howe, D., & Soderman, A. B. (2009). The aesthetics of generative literature: Lessons from a digital writing workshop. *Hyperrhiz Journal of New Media Cultures*.
HuggingFace. (2020a). *Transformers*. Retrieved April 17, 2020, from https://huggingface.co/transformers/
HuggingFace. (2020b). *Write with transformer*. Retrieved April 17, 2020, from https://transformer.huggingface.co
Inc., A. (2020). *Adobe sensei: Power incredible experiences with ai*. Retrieved May 28, 2020, from https://www.adobe.com/ie/sensei.html
Jordà, S., Geiger, G., Alonso, M., & Kaltenbrunner, M. (2007). The reactable: Exploring the synergy between live music performance and tabletop tangible interfaces, In *Proceedings of the 1st international conference on tangible and embedded interaction*, Baton Rouge, Louisiana, Association for Computing Machinery. https://doi.org/10.1145/1226969.1226998
Kelly, K. (2017). The myth of a superhuman ai. *Wired, April*.
Knight, W. (2019). *An ai that writes convincing prose risks mass-producing fake news*. Retrieved April 15, 2020, from https://www.technologyreview.com/2019/02/14/137426/an-ai-tool-auto-generates-fake-news-bogus-tweets-and-plenty-of-gibberish/
Knuth, D. E. (1968). *The art of computer programming: Fundamental algorithms* (Vol. 1). Addison-Wesley.
Krogh, P., Markussen, T., & Bang, A. (2015). Ways of drifting – 5 methods of experimentation in research through design. https://doi.org/10.1007/978-81-322-2232-3
Kuo, L. (2020). *The new normal: China's excessive coronavirus public monitoring could be here to stay*. Retrieved April 6, 2020, from https://www.theguardian.com/world/2020/mar/09/the-new-normal-chinas-excessive-coronavirus-public-monitoring-could-be-here-to-stay
Laguna, C. P., & Fiebrink, R. (2014). Improving data-driven design and exploration of digital musical instruments, In *Chi '14 extended abstracts on human factors in computing systems*, Toronto, Ontario, Canada, Association for Computing Machinery. https://doi.org/10.1145/2559206.2581327
Landin, H. (2005). Fragile and magical: Materiality of computational technology as design material, In *Proceedings of the 4th decennial conference on critical computing: Between sense and sensibility*, Aarhus, Denmark, Association for Computing Machinery. https://doi.org/10.1145/1094562.1094578
Leavy, S. (2018). Gender bias in artificial intelligence: The need for diversity and gender theory in machine learning, In *Proceedings of the 1st international workshop on gender equality in software engineering*, Gothenburg, Sweden, Association for Computing Machinery. https://doi.org/10.1145/3195570.3195580

Lim, Y.-k., Lee, S.-S., & Lee, K.-y. (2009). Interactivity attributes: A new way of thinking and describing interactivity, In *Proceedings of the sigchi conference on human factors in computing systems*, Boston, MA, USA, Association for Computing Machinery. https://doi.org/10.1145/1518701.1518719

Liu, V., & Curran, J. R. (2006). Web text corpus for natural language processing, In *11th conference of the European chapter of the association for computational linguistics*, Trento, Italy, Association for Computational Linguistics.

Liu, Y., & Lin, Z. (2019). Unsupervised pre-training for natural language generation: A literature review.

Löwgren, J. (2009). Toward an articulation of interaction esthetics. *New Review of Hypermedia and Multimedia*, *15*(2), https://doi.org/10.1080/13614560903117822, 129–146. https://doi.org/10.1080/13614560903117822

McCulloch, W., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, *5*(4), 115.

Miller, D. (2019). Leveraging bert for extractive text summarization on lectures.

Mitchell, T. M. (1997). *Machine learning*. McGraw-hill New York.

Montfort, N. (2005). *Twisty little passages: An approach to interactive fiction*. MIT Press.

Nelson, T. H. (1982). *Literary machines*. Mindful Press.

Nelson, T. H. (1999). Xanalogical structure, needed now more than ever: Parallel documents, deep links to content, deep versioning, and deep re-use. *ACM Computing Surveys (CSUR)*, *31*(4es), 33–es.

Newell, A., & Simon, H. A. (1976). Computer science as empirical inquiry: Symbols and search. *Commun. ACM*, *19*(3), 113–126. https://doi.org/10.1145/360018.360022

Nilsson, N. J. (2007). The physical symbol system hypothesis: Status and prospects. In M. Lungarella, F. Iida, J. Bongard, & R. Pfeifer (Eds.), *50 years of artificial intelligence: Essays dedicated to the 50th anniversary of artificial intelligence* (pp. 9–17). Berlin, Heidelberg, Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-77296-5_2

Norvig, P. R., & Russell, S. (2009). *Artificial intelligence: A modern approach*. Prentice Hall.

Ong, W. J. (1971). *Rhetoric, romance, and technology: Studies in the interaction of expression and culture*. Cornell University Press.

Peng, T., & Sarazen, M. (2019). *The staggering cost of training sota ai models*. Retrieved May 8, 2020, from https://syncedreview.com/2019/06/27/the-staggering-cost-of-training-sota-ai-models/

Quach, K. (2019). *This news article about the full public release of openai's 'dangerous' gpt-2 model was part written by gpt-2*. Retrieved April 9, 2020, from https://www.theregister.co.uk/2019/11/06/openai_gpt2_released/

Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language models are unsupervised multitask learners.

Ramos, G., Suh, J., Ghorashi, S., Meek, C., Banks, R., Amershi, S., Fiebrink, R., Smith-Renner, A., & Bansal, G. (2019). Emerging perspectives in human-centered machine learning, In *Extended abstracts of the 2019 chi conference on human factors in computing systems*, Glasgow, Scotland Uk, Association for Computing Machinery. https://doi.org/10.1145/3290607.3299014

Samuel, A. L. (1959). Some studies in machine learning using the game of checkers. *IBM Journal of Research and Development*, 71–105.

Sanders, E. B.-N., & Stappers, P. J. (2012). *Convivial toolbox: Generative research for the front end of design*. BIS Publishers.

Schulz, H. (2011). Treevis.net: A tree visualization reference. *IEEE Computer Graphics and Applications*, *31*(6), 11–15.

Shapiro, R. B., Fiebrink, R., & Norvig, P. (2018). How machine learning impacts the undergraduate computing curriculum. *Commun. ACM*, *61*(11), 27–29. https://doi.org/10.1145/3277567

Song, H., Qi, Y., Xiao, L., Zhu, T., & Curran, E. P. (2006). Lenstree: Browsing and navigating large hierarchical information structures, In *16th international conference on artificial reality and telexistence–workshops (icat'06)*.

Spacemarker. (2020). *Design better cities with artificial intelligence*. Retrieved May 28, 2020, from https://spacemaker.ai

Strebe, D. (2020a). The limits of the simulation of cognitive functions and the supernatural: Artificial and human perception.

Strebe, D. (2020b). *The prayer*. Retrieved April 8, 2020, from https://theprayer.diemutstrebe.com/

Sundström, P., Taylor, A. S., & O'Hara, K. (2011). Sketching in software and hardware bluetooth as a design material, In *Proceedings of the 13th international conference on human computer interaction with mobile devices and services*, Stockholm, Sweden, Association for Computing Machinery. https://doi.org/10.1145/2037373.2037434

Tominski, C., Abello, J., van Ham, F., & Schumann, H. (2006). Fisheye tree views and lenses for graph visualization, In *Tenth international conference on information visualisation (iv'06)*.

Tzara, T. (1920). Dada manifesto on feeble love and bitter love.

van Stegeren, J., & Theune, M. (2019). Narrative generation in the wild: Methods from nanogenmo, In *Proceedings of the second workshop on storytelling*.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, & R. Garnett (Eds.), *Advances in neural information processing systems 30* (pp. 5998–6008). Curran Associates, Inc.

Ventura, D., & Warnick, S. (2007). A theoretical foundation for inductive transfer. *Brigham Young University, College of Physical and Mathematical Sciences*.

von Platen, P. (2020). *How to generate text: Using different decoding methods for language generation with transformers*. Retrieved April 17, 2020, from https://huggingface.co/blog/how-to-generate

Walton, N. (2019). *Ai dungeon 2*. Retrieved April 6, 2020, from https://aidungeon.io

Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., Cistac, P., Rault, T., Louf, R., Funtowicz, M., & Brew, J. (2019). Huggingface's transformers: State-of-the-art natural language processing. *ArXiv*, *abs/1910.03771*.

Yang, Q. (2018). Machine learning as a ux design material: How can we imagine beyond automation, recommenders, and reminders?

Yang, Q., Banovic, N., & Zimmerman, J. (2018). Mapping machine learning advances from hci research to reveal starting places for design innovation, In *Proceedings of the 2018 chi conference on human factors in computing systems*, Montreal QC, Canada, Association for Computing Machinery. https://doi.org/10.1145/3173574.3173704

Yang, Q., Scuito, A., Zimmerman, J., Forlizzi, J., & Steinfeld, A. (2018). Investigating how experienced ux designers effectively work with machine learning, In *Proceedings of the 2018 designing interactive systems conference*, Hong Kong, China, Association for Computing Machinery. https://doi.org/10.1145/3196709.3196730

Yang, Q., Zimmerman, J., Steinfeld, A., & Tomasic, A. (2016). Planning adaptive mobile experiences when wireframing, In *Proceedings of the 2016 acm conference on designing interactive systems*, Brisbane, QLD, Australia, Association for Computing Machinery. https://doi.org/10.1145/2901790.2901858

Zador, A. (2019). A critique of pure learning and what artificial neural networks can learn from animal brains. *Nature Communications*, *10*. https://doi.org/10.1038/s41467-019-11786-6

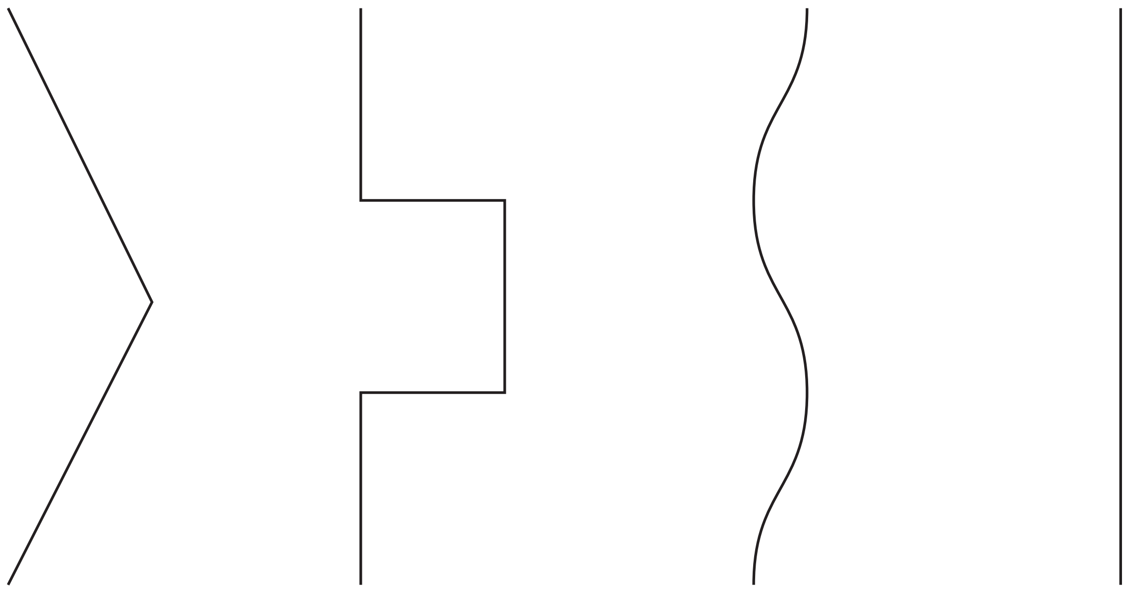
Zimmerman, J., Forlizzi, J., & Evenson, S. (2007). Research through design as a method for interaction design research in hci, In *Proceedings of the sigchi conference on human factors in computing systems*, San Jose, California, USA, Association for Computing Machinery. https://doi.org/10.1145/1240624.1240704

Zimmerman, J., Stolterman, E., & Forlizzi, J. (2010). An analysis and critique of research through design: Towards a formalization of a research approach, In *Proceedings of the 8th acm conference on designing interactive systems*, Aarhus, Denmark, Association for Computing Machinery. https://doi.org/10.1145/1858171.1858228

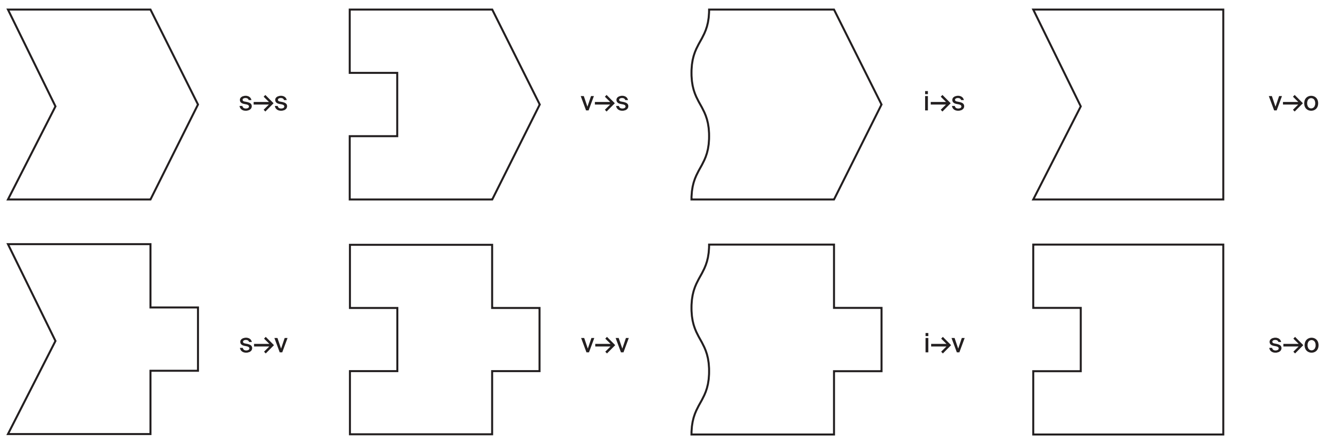
Zou, J., & Schiebinger, L. (2018). Ai can be sexist and racist—it's time to make it fair. *Nature*. https://doi.org/10.1038/d41586-018-05707-8

Appendix A
Material for Neural Plumbing

1. Cut extremes of post-its to shape using the guides
2. Annotate elements [see: complementary page]
3. Arrange and re-arrange elements to form pipelines



sequence | vector | input | output



s→s

language masking
summarization
translation
text generation
text-to-seech
question answering

s→v

named-entity-recognition
sentiment analysis
text filtering

v→v

data processing
image classification
image generation
recommendation
video object tracking
video stabilization

v→s

image captioning
recommendation
video stabilization

i→s/v

api/database query
dataset
user input

s/v→o

display
further processing
sound
video