Smart, P. R. (in press) Predicting Me: The Route to Digital Immortality? In R. W. Clowes, K. Gärtner & I. Hipólito (Eds.), *The Mind-Technology Problem: Investigating Minds, Selves and 21st Century Artifacts*. Springer, Berlin, Germany.

Predicting Me: The Route to Digital Immortality?

Paul Smart

Abstract An emerging consensus in cognitive science views the biological brain as a hierarchically-organized predictive processing system that relies on generative models to predict the structure of sensory information. Such a view resonates with a body of work in machine learning that has explored the problem-solving capabilities of hierarchically-organized, multi-layer (i.e., deep) neural networks, many of which acquire and deploy generative models of their training data. The present chapter explores the extent to which the ostensible convergence on a common neurocomputational architecture (centred on predictive processing schemes, hierarchical organization, and generative models) might provide inroads into the problem of digital immortality. In contrast to approaches that seek to recapitulate the physical structure of the human brain, the present chapter advocates an approach that is rooted in the use of machine learning algorithms. The claim is that a future form of deep learning system could be used to acquire generative models of a given individual or (alternatively) the sensory data that is processed by the brain of a given individual during the course of their biological life. The differences between these two forms of digital immortality are explored, as are some of the options for digital resurrection.

To die,—to sleep,
To sleep! perchance to dream. . .
For in that sleep of death what dreams may come. . .

—Hamlet, William Shakespeare

Paul Smart
Electronics & Computer Science, University of Southampton, Southampton, SO17 1BJ, UK
e-mail: ps02v@ecs.soton.ac.uk

1 Introduction

Do you want to live forever? If so, then the 21st century may be the perfect time to die. Digital immortality has long been a source of fascination for the transhumanist movement, yielding many proposals as to how a given individual might be 'serialized' to a digital medium and then 'resurrected' as part of some digital afterlife. Such accounts have often been the target of philosophical criticism, inspiring more in the way of philosophical invective than they have technological innovation. But is all this about to change? Recently, there has been a renewal of interest in the notion of digital immortality, with an increasing number of companies now offering some form of digital afterlife. "Become virtually immortal," reads the slogan of one such company, Eternime.[1] It is, of course, unclear whether this particular commercial offering relates to a virtual form of immortality or a form of immortality that is virtually possible. But notwithstanding these ambiguities, the long-term aims of the company are relatively clear:

> Eternime collects your thoughts, stories and memories, curates them and creates an intelligent avatar that looks like you. This avatar will live forever and allow other people in the future to access your memories.

Discussions of digital immortality typically go hand-in-hand with an appeal to some form of *mind uploading*. According to Goertzel and Ikle' (2012), mind loading is:

> ...an informal term referring...to the (as yet hypothetical) process of transferring the totality or considerable majority of the mental contents from a particular human brain into a different substrate, most commonly an engineered substrate such as a digital, analogue or quantum computer. (Goertzel and Ikle', 2012, p. 1)

Perhaps unsurprisingly, there are many different proposals as to how this rather vague objective might be realized. Most proposals advocate the use of advanced technology to record information about the structure of an individual's biological brain. Hayworth (2012), for example, discusses how an advanced imaging technique (Focused Ion Beam Scanning Electron Microscopy) might be used to map the structure of whole brain neural circuits, yielding a more-or-less complete model of the human connectome—i.e., the connection matrix of the human brain (Sporns et al, 2005). Inasmuch as it is this structural description of the brain that defines who and what we are (see Seung, 2012), then such approaches have an obvious appeal, even if they are still recognized as being impractical or beyond the limits of current technology.

The present chapter describes an approach to digital immortality that is similar, at least in spirit, to many forms of mind uploading. Where it departs from previous accounts is with respect to the approach taken to model (and re-generate) the functional dynamics of the human brain. Instead of trying to directly map the detailed microstructure of the biological brain using imaging or tracing techniques, the present approach is rooted in the use of machine learning techniques, especially

[1] See `http://eterni.me/` (accessed: 7th March 2018).

those forms of machine learning whose styles of computation, representation, and overall architecture share some similarity with recent models of brain-based (or at any rate, cortical) processing. The inspiration for this approach is based on a recent neurocomputational model of brain function that depicts the biological brain as a hierarchically-organized predictive processing system, constantly engaged in the attempt to predict its own activity[2] at a variety of (increasingly abstract) spatial and temporal scales (Clark, 2016, 2013b; Friston, 2010). This account of brain function has been the target of considerable scientific and philosophical interest, at least in part because the account is deemed to be relevant to a broad swath of seemingly disparate psychological phenomena, including learning, attention, perception, action, emotion, imagination, memory, and various forms of mental illness (Clark, 2016; Friston et al, 2014). Beyond this, however, the vision of the brain as a hierarchically-organized predictive processing system is one that is reflected in recent approaches to machine learning, especially those associated with deep learning systems (Bengio, 2009; LeCun et al, 2015). Such forms of convergence are interesting given the recent successes of deep learning on a number of Artificial Intelligence (AI) problems, and they may even mark the beginnings of a still somewhat ill-defined path towards experientially-potent forms of machine cognition. The aim of the present chapter is to describe these two areas of research (i.e., predictive processing models of brain function and deep machine learning) with a view to outlining an approach to digital immortality that highlights the relevance of research into deep learning, virtual reality, and big data processing. The technologies associated with these research areas are likely to have a substantial impact on various spheres of human activity throughout the 21st century.

The chapter is structured as follows: Section 2 outlines the broad shape of the predictive processing (PP) framework, as discussed by Clark (2016), Friston (2010), and others. It focuses on a key feature of the PP account, namely, the use of generative models to construct the sensory signal ‘from the top down’. Section 3 aims to draw attention to some of the similarities between the PP account of brain function and deep learning systems. Section 4 then goes on to suggest that the link between deep learning systems and PP serves as the basis for a particular approach to digital immortality: one that is rooted in the idea that deep learning systems might be used to recreate the generative models that are acquired by biological brains as a result of prediction-oriented learning. Section 5 reflects a shift in focus, from machine learning to big data. It discusses some of the problems confronting the approach to digital immortality presented in Section 4. In particular, Section 5 raises questions about the sort of data that ought to be collected, as well as some of the technical, social, and ethical challenges that are likely to confront the data collection effort. Section 6 explores the role of virtual reality technologies in establishing some sort of digital afterlife, with a particular emphasis on the notion of embodiment. Finally, Section 7 concludes the chapter.

[2] In essence, the brain is viewed as a multi-layered prediction machine, with ‘higher’ layers attempting to predict the activity of ‘lower’ layers. It is in this sense that the brain can be seen to predict its own activity, i.e., to predict the activity of its constituent neural elements.

2 Predictive Processing

Over the course of the past decade, a particular view of the brain has become increasingly popular, both in cognitive neuroscience and the philosophy of mind. This is a view that sees the biological brain as a hierarchically-organized system that is constantly striving to predict its own internal activity, relative to the play of energy across the organism's sensory surfaces (Clark, 2016). The most popular version of this account is known as the predictive processing account of cognition (PP for short).

Some insight into the general flavour of PP can be gleaned by comparing the PP approach to perception with its more traditional counterpart (see Figure 1). On the traditional view, perception occurs via the stepwise analysis of incoming sensory information, with more abstract features being detected at progressively higher levels of the cortical hierarchy (see Figure 1a). The aim, in this case, is to analyse the upward/forward-flowing stream of information to the point where action can be coordinated with respect to abstract properties of the perceptual scene.

The PP view of perceptual processing is somewhat different (see Figure 1b). Here, the stream of incoming sensory information is met with a downward/backward cascade of predictions that emanate from progressively higher layers of the processing hierarchy. The purpose of this downward-flowing stream is to predict the activity of neural circuits at each layer in the hierarchy, with the forward/upward-flowing stream of information being used to communicate the mismatch between actual and predicted activity. The overall aim of the system, in this case, is to minimize prediction error and thus suppress the forward flow of information. From an information-theoretic standpoint, prediction error is seen to provide a measure of "free energy," which is defined as the "difference between an organism's predictions about its sensory inputs (embodied in its models of the world) and the sensations it actually encounters" (Friston et al, 2012, p. 1). Reductions in prediction error therefore correspond to reductions in free energy, which, over the longer term, equates to a form of entropy minimization (see Friston, 2010). As noted by Clark (2016):

> ... good [predictive] models... are those that help us successfully engage the world and hence help us to maintain our structure and organization so that we appear—over extended but finite timescales—to resist increases in entropy and (hence) the second law of thermodynamics. (Clark, 2016, p. 306)

The means by which predictive capabilities are acquired by the brain is often depicted as a form of perceptual learning. In essence, prediction error is seen to promote changes in synaptic strength that reconfigure the structure of neural circuits, enabling higher-level neural regions to better predict the activity of lower-level regions. Crucially, one of the upshots of this particular form of prediction-oriented learning is the installation of a hierarchically-organized model that (by virtue of the organism's sensory contact with the world) tracks the hidden causes (or latent variables) that govern the statistical structure of incoming sensory information. An important feature of these models is that they are *generative* in nature. That is to say, the models encoded by the neural circuits at each layer in the hierarchy must

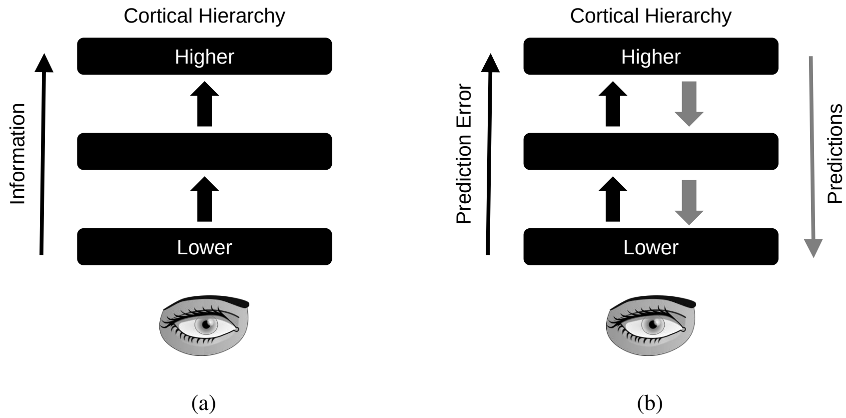


Fig. 1 Two approaches to perceptual processing. (a) The traditional approach to perceptual processing is characterized by an upward/forward-flow of information through a succession of cortical regions. (b) This contrasts with the PP approach, which emphasizes the role of backward/downward-flowing predictions in suppressing the upward/forward flow of information.

be such as to allow each layer to predict activity in the layer below. This means, in effect, that each layer is able to generate the information encoded by the lower layer, and this extends all the way out to the lowest levels of the hierarchy, i.e., to the point where the torrent of downward-flowing predictions meets the incoming tide of sensory information.

It is widely assumed that the overall result of this hierarchical organization is a multi-layer generative model that embodies the causal structure of the environment. In particular, it is assumed that by virtue of the attempt to recapitulate the activity of lower levels, and thus accommodate the incoming sensory signal, the brain is attempting to model the interacting set of worldly causes that give rise to particular kinds of sensory stimulation. In a sense, successful prediction is like a form of 'understanding', where the understanding in question concerns the causal forces and factors that shape whatever bodies of sensory information exist within the organism's local environment (see Clark, 2013a). It is at this point that the commitment to multi-layer, hierarchically-organized neural architectures takes on a special significance, for it seems that such an organization is ideally suited to the kind of world in which we humans live—a world built around a structured nexus of interacting, and often deeply nested, causal forces. The goal of perception, according to PP, is to invert this casual structure by using a generative model to infer the causes of sensory input (see Figure 2). "The hierarchical structure of the real world," Friston (2002) suggests, "literally comes to be reflected by the hierarchical architectures trying to minimize prediction error, not just at the level of sensory input but at all levels of the hierarchy" (pp. 237–238).

As noted by Clark (2016), generative models may lie at heart of a number of cognitive phenomena, including our capacity for hallucination, dreaming, fantasy, and the

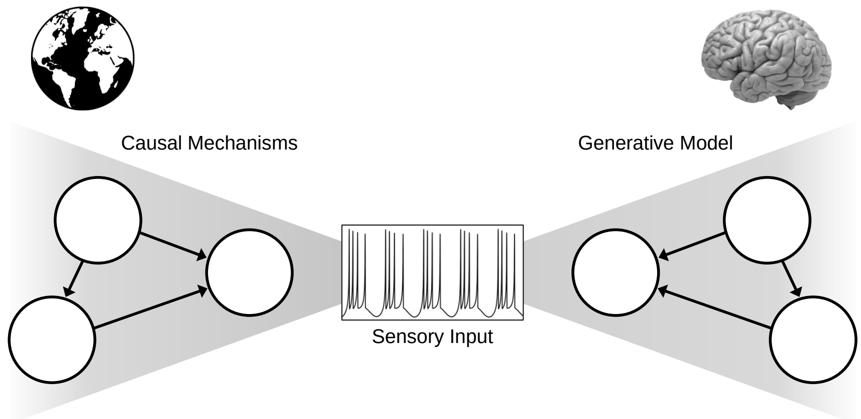


Fig. 2 A generative model describes how variables or causes in the environment conspire to produce sensory input. Perception then corresponds to the inverse mapping from sensations to their causes.

potential for self-generated forms of mental imagery. In acquiring a generative model, the brain develops a capacity to drive patterns of neural activity from the top-down, and in doing so it is able to recreate the patterns of activity that would be instantiated if the organism were to be confronted with a particular pattern of sensory input. "Probabilistic generative model based systems that can learn to visually perceive a cat," suggests Clark (2013b), "...are, ipso facto, systems that can deploy a top-down cascade to bring about many of the activity patterns that would ensue in the visual presence of an actual cat" (p. 198). This seems to be of profound importance relative to our understanding of many aspects of our mental lives. Generative capacities thus seem relevant to our ability to dream the non-existent, imagine the counterfactual, anticipate the future, and (via memory) reconstruct the past.

In addition to causing us to rethink the relationship between seemingly distinct cognitive phenomena, the PP account also informs our view of what is perhaps the most elusive part of our cognitive economy: conscious experience. In particular, the appeal to generative models is sometimes seen to reinforce an approach to consciousness that portrays it as a form of virtual reality (see Revonsuo, 1995) or controlled hallucination. Metzinger (2003), for example, suggests that a fruitful way of looking at the brain is to view it:

> ...as a system which, even in ordinary waking states, constantly hallucinates at the world, as a system that constantly lets its internal autonomous simulational dynamics collide with the ongoing flow of sensory input, vigorously dreaming at the world and thereby generating the content of phenomenal experience. (Metzinger, 2003, p. 52)

As a means of establishing a better grip on this idea of phenomenal experience as a form of controlled hallucination or actively constructed virtual reality, it may help to consider a hypothetical situation in which the predictability of the environment (or the predictive power of the brain's generative model) is such that the activity at each layer in the neural processing hierarchy is perfectly predicted by the layer above. In such a situation, the incoming flow of sensory information is met by a cascade of downward-flowing predictions, which perfectly captures the activity at each and every layer of the processing hierarchy. Such a state-of-affairs is interesting, for the predictive successes of each layer eliminate any forward (or upward) flow of information (i.e., from lower to higher layers in the hierarchy). According to the PP account, recall, the forward flow of information corresponds to prediction error—the mismatch between actual and predicted activity at each level in the hierarchy. But in this particular case—let us call it the *no-prediction-error-case*—there is no prediction error, and thus there is no forward/upward flow of information. What we are left with, therefore, is a purely backward or downward flow of information, from higher cortical regions all the way out to the point of sensory input.

Of course, actual instances of the no-prediction-error-case are unlikely to be found in the real world. Real-world environments are seldom perfectly predictable, and the activity of neural circuits is often characterized by the presence of neuronal noise (e.g., Faisal et al, 2008). Thus even if organisms were to seek out a dark, unchanging chamber as per the worries raised by the "Dark-Room Problem" (Friston et al, 2012), it is unlikely that such organisms would be completely free of prediction error. As a simple thought experiment, however, the no-prediction-error-case is useful in helping us get to grips with the idea of perceptual experience (and perhaps phenomenal experience, more generally) as something that is generated from the 'inside out'. Assuming that perceptual experience occurs as a consequence of the formation of stable neural states (i.e., those that successfully predict the activity of other neural regions and are thus unperturbed by any form of prediction error), then the no-prediction-error-case is one in which the experience of (e.g.) seeing a scene occurs in the absence of any forward flow of information through the brain. The result is a view of conscious experience as something that is actively generated by the brain as part of its attempt to model the causal structure of the sensorium and thus predict its own (sensorially-shaped) neural activity.

The upshot of all this is a vision of the brain as a form of virtual reality generator—the biological equivalent of technologies that render virtual objects and virtual worlds. According to this vision, aspects of our daily conscious experience are tied to the brain's attempt to acquire and deploy generative models that track the causal structure of the external environment. In particular, it seems that conscious experience might be linked to the activation of representations corresponding to an interacting set of external causes that are acquired as the result of the attempt to predict the structure of incoming sensory information.

This is a compelling, although still somewhat puzzling, vision. It is a vision that depicts our phenomenal experience as something akin to a simulation of reality, and it is a vision that blurs the distinction between ostensibly distinct cognitive phenomena, such as perception, imagination, dreaming, and fantasy. Relative to such a vision, it is perhaps easy to think of life as nothing more than a dream. To echo the views of Metzinger, what we call waking life may be nothing more than a form of "online dreaming" (Metzinger, 2003, p. 140).

3 Dream Machines

The PP account bears some interesting similarities to recent work in machine learning, particularly that which focuses on so-called deep learning systems (Bengio, 2009; LeCun et al, 2015). As with the PP model of brain function, deep learning emphasizes the importance of multi-layer architectures, with 'higher' layers yielding more abstract representations of the response patterns exhibited by 'lower' layers (at least in some systems). The notion of generative models also marks a point of commonality between PP and at least some strands of deep learning research. Here, the attention of the machine learning community has shifted away from a traditional focus on the discriminative capacities of neural networks (e.g., their ability to discriminate between objects of different types) towards a better understanding of their generative capabilities (i.e., their ability to re-create bodies of training data).

This shift in focus—from discriminative to generative capacities—is important, for it highlights the potential relevance of deep learning systems to debates about digital immortality. Inasmuch as deep learning systems are able to emulate the functionality of the brain with respect to the acquisition and deployment of generative models (and inasmuch as generative models are revealed to be a cornerstone of the human cognitive economy), then we might wonder whether a deep learning system could be used to re-create the generative models embodied by a biological brain. Perhaps if we could capture the streams of sensory data against which brain-based generative models take shape, we could then use this data to train a deep learning system and thereby reinstate (some of) the cognitive properties of a given human individual. This is the essence of an approach to digital immortality that highlights the potential relevance of two key 21st century technologies—deep learning systems and big data technologies—to issues of digital immortality.[3] We will explore some of the implications (and problems) associated with this approach in subsequent sections. For now, however, let us direct our attention to the nature of the (putative) link between deep learning systems and the PP account of brain function.

When it comes to deep learning systems with generative capacities, a number of systems have been the focus of recent research attention. These include Deep Belief Networks (DBNs) (Hinton, 2007a,b), variational autoencoders, and Generative Adversarial Networks (GANs) (Goodfellow et al, 2014). DBNs are a particular kind of deep learning system. They are composed of multiple layers of what are called Restricted Boltzmann Machines (RBMs). These RBMs are a type of neural network consisting of two layers: a visible layer and a hidden layer. The nodes within each layer are connected to nodes in adjacent layers; however, there are no intra-layer connections (i.e., nodes within a particular layer are not connected to nodes in the

[3] It is doubtful whether the current state-of-the-art in deep learning is sufficient to achieve the sort of digital immortality vision being proposed here. Nevertheless, it is worth bearing in mind that deep learning research is likely to be a prominent focus of global research attention over the next 10–20 years. Given the amount of time, effort, and money that is likely to be devoted to deep learning systems in the coming years, it is likely that we will see significant changes in their capabilities during the course of the 21st century.

same layer).[4] The nodes in the visible layer represent the data that is presented to the RBM, and the goal of the hidden layer is to capture higher-order correlations between the data that is represented at the visible layer. Typically, all the nodes in the RBM are binary, with two states represented by the digits '0' and '1'. This means that in cases where the RBM is presented with a black and white image, the nodes at the visible layer will represent the image using a binary data vector whose elements represent the individual pixel intensities of the image (e.g., '1' for white and '0' for black). Relative to this case, the nodes in the hidden layer now function as binary feature detectors that seek to model the higher-order correlations between pixel values at the visible layer. In particular, the aim during learning is to configure the weights associated with the top-down connections (from hidden layer to visible layer) such that the hidden layer is able to recreate the training data represented by the nodes of the visible layer. It is in this sense that the RBM's model of the training data is said to reside in its top-down connections.

Clearly, a single layer of binary feature detectors is unlikely to capture all the latent structure that exists within a complex set of images, especially when we reflect on the complexity of the interacting causal processes that conspire to generate individual pixel intensities (see Horn, 1977). For this reason, additional layers are added to a base RBM to expand its representational capabilities. It is at this point that a single, two-layer RBM begins to morph into a multi-layer DBN (see Figure 3a). As each layer is added, the new layer is treated as the hidden layer of a new RBM, while the erstwhile hidden layer of the original RBM now functions as the visible layer. In effect, the representations of the original hidden layer now become the training data (or 'sensory' input) for the new layer, and the goal of the new layer is to learn a suite of more abstract representations that capture the dynamics of the layer below. Perhaps unsurprisingly, the addition of each new layer enhances the system's ability to model abstract structural regularities, thereby improving its capacity to generate the training data at the lowest layer in the hierarchy (i.e., the visible layer of the original RBM).

The upshot of this kind of (incrementally-oriented) learning regime is a multi-layer neural network that is, in effect, a composite of multiple RBMs (see Figure 3b). This system is what is typically dubbed a DBN. At this point, the system possesses a good generative model of the target domain, as represented by the training data, but it isn't necessarily well-suited to other kinds of tasks, such as the classification of images into particular classes. Nevertheless, in being forced to recreate the training data, the network has learned a great deal about the hidden causes or latent variables that structure the training data. What the network has learned, in effect, is a way of 'explaining' each input vector (each sample of sensory data) in terms of a nexus of interacting and deeply nested (hidden) causes, where the notion of a good explanation corresponds to "a binary state vector for each layer that is both likely to cause the binary state vector in the layer below and likely to be caused by the binary state vector in the layer above" (Hinton, 2010, p. 179). It turns out that this 'explanatory'

[4] This highlights one of the differences between DBNs and PP accounts of cognition. In PP, it is typically assumed that elements within the same layer of the processing hierarchy engage in some form of lateral processing.

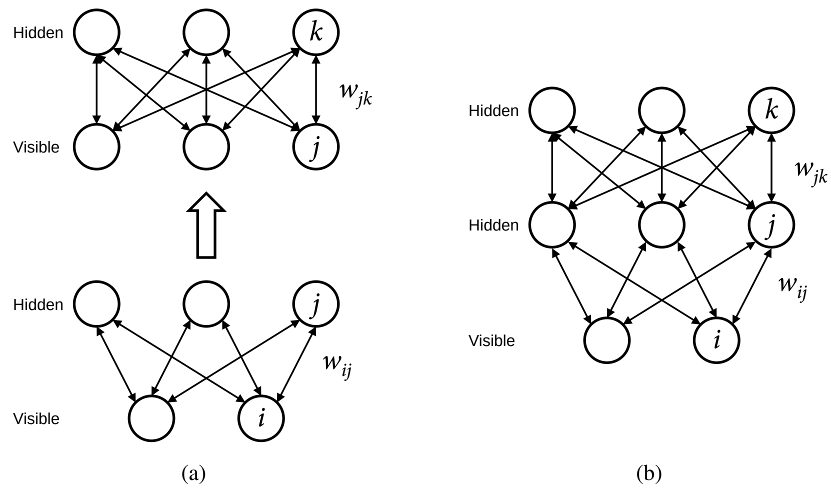
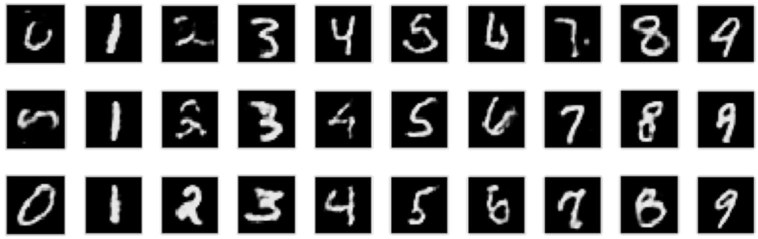


Fig. 3 Restricted Boltzmann Machines and Deep Belief Networks. (a) Two separate RBMs. The stochastic binary variables in the hidden layer of each RBM are symmetrically connected to the stochastic binary variables in the visible layer. There are no connections within a layer. The higher-level RBM is trained using the hidden activities of the lower RBM as data. (b) A DBN formed from the merger of two RBMs.

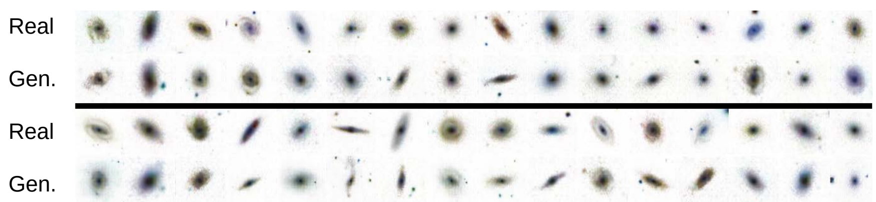
capability serves as the basis for enhanced performance in a variety of task contexts. In the case of Hinton's early work with DBNs, for example, a DBN was trained with images of handwritten digits taken from the MNIST data set. The DBN's subsequent discriminative performance was tested by extending the network with a set of 'label' nodes corresponding to the kind of conceptual distinctions that we humans make when dealing with the realm of handwritten digits (i.e., our ability to recognize a particular image as representing a particular number). A variety of studies have shown that this kind of approach—essentially treating digit classification as something of a post-processing step relative to the primary goal of acquiring a generative model—is able to deliver superior performance compared to networks that are trained with conventional back propagation techniques and random initial weights (Hinton and Salakhutdinov, 2006).

In addition to highlighting the performance benefits of DBNs when it comes to the recognition of handwritten digits, Hinton's work also provides a compelling demonstration of the generative capacity of such architectures (see Hinton, 2007b). With the addition of a set of label nodes representing digit types, Hinton was able to selectively activate particular label nodes and then observe the data vector (the sensory output) produced by the network at the visible layer. Figure 4a illustrates some of the images generated using this method.

A further demonstration of the generative capacity of deep learning systems comes from a study by Ravanbakhsh et al. (2017). The purpose of Ravanbakhsh et



(a) Digit generation.



(b) Galaxy generation.

Fig. 4 Examples of sensory data generated by deep learning systems. (a) The output from a generative model trained with handwritten digits. Each row shows 10 samples from the generative model with a particular digit label clamped on. The top-level associative memory is run for 1000 iterations of alternating Gibbs sampling between samples (source: `http://www.cs.toronto.edu/~hinton/digits.html`). (b) Actual ('Real') versus generated ('Gen.') galaxy images. Actual images are taken from the Galaxy Zoo data set; generated images were created using a conditional generative adversarial network (source: Ravanbakhsh et al, 2017).

al.'s (2017) study was to generate images of galaxies for the purpose of calibrating astronomical equipment. As part of the study, Ravanbakhsh et al. exposed two kinds of deep learning system—namely, variational autoencoders and GANs—to images of real galaxies taken from the Galaxy Zoo data set.[5] As a result of training with respect to these images, the deep learning systems acquired a generative model of the target domain, enabling them to produce galaxy images similar to those contained in the real-world data set (see Figure 4b). In a sense, of course, these images are 'fakes', since the galaxies they represent do not exist. At the same time, however, there is surely something compelling about the generative abilities exhibited by these systems. Their abilities remind us of our own human ability for creativity—our ability to use our daily waking experience as the point of departure for excursions into a realm of purely imagined objects, ideas, and other fantastical constructions. As was mentioned above, it is here that recent work into generative models provides us with a potential insight into what is perhaps the most elusive aspect of the human cognitive economy—namely, our capacity to imagine, to fantasize, and to dream. We may, of

[5] The Galaxy Zoo data set consists of $\approx 900,000$ galaxy images, which were collected as part of the Sloan Digital Sky Survey. The data set was originally used as part of a citizen science project investigating the distribution of galaxies with particular morphologies (see Lintott et al, 2008). It has since been used as the basis for a number of studies exploring the capacities of both conventional and deep neural networks (Dieleman et al, 2015; Banerji et al, 2014).

course, never know whether deep neural networks should count as *bona fide* 'dream factories'—systems that are genuine dreamers and imaginers—but we do know that they (like us) are creative engines—systems capable of using high-level abstract representations for the purpose of creating digital artefacts (e.g., images). This should perhaps give us pause when we wonder (and sometimes worry) about the capacity of AI to encroach on the oft cherished realm of human creativity. For perhaps the early successes of generative deep learning systems intimate at something more profound: an ability to emulate (and perhaps even surpass) our best intellectual triumphs and artistic accomplishments. If a deep neural network can generate an image of a galaxy (which is, after all, just a sequence of binary digits), then why should it not be able to create other kinds of digital artefact, such as a 3D object, a movie, or the design for a deep neural network capable of delivering state-of-the-art advances in machine-based generative capabilities?

4 Generating Me

It should by now be clear that the representational and computational profile of at least some forms of deep learning system echo those that lie at the heart of the PP approach to cognition. There are, of course, important differences. Deep learning systems seldom feature any form of intra-layer processing and important elements of the PP model (e.g., real-time prediction error estimation and the precision-weighted modulation of information flows) are absent (or at least under-represented) in current deep learning systems. Nevertheless, the basic commitment to hierarchically-organized processing schemes and the emphasis on generative capabilities has paid substantial dividends, yielding important advances in problem areas that have long stymied the efforts of the AI community (LeCun et al, 2015; Najafabadi et al, 2015). In addition to notable successes in the areas of language processing and computer vision, deep learning research has yielded interesting results in a number of more 'niche' areas, helping us understand the generative mechanisms responsible for the structure of neural circuits (Betzel et al, 2016) and providing us with predictive models of drug molecule activity (Ma et al, 2015). The question, of course, is whether these early successes help to reveal the beginnings of a path that makes inroads into the problem of digital immortality.

It is at this point that I will opt to bite the bullet, so to speak, and suggest that some form of deep learning system—specifically, a system whose computational profile is closely aligned with the PP model—is, indeed, a technology that ought to be considered by the proponents of digital immortality. The basic idea is that a particular form of deep learning system—let us call it a Synthetic Predictive Processing (SPP) system to emphasize the overlap with the PP model—is tasked with the objective of acquiring generative models that resemble those acquired by an individual's biological brain. The idea, in essence, is for a SPP system to exist as a form of constant digital companion that accompanies the individual across the course of their life and participates in a form of lifelong learning. The ultimate goal of such a system is to deal with the same sort of predictive challenges that confront the biological brain. The hope is that in dealing with these challenges, the SPP system will come to acquire a set of probabilistic generative models whose generative capacities are functionally similar to those acquired by its biological counterpart. Inasmuch as our 'online dreams' are produced by generative models in the biological case, is there any reason to think that the general shape of those dreams could not be manufactured by a synthetic predictive processing machine with more-or-less the same generative capabilities?

There are multiple possible variations on this sort of idea. The proposal just outlined is what I will dub the *first-person proposal*. It seeks to situate a SPP system within the same sensory environment as that in which the biological brain is itself embedded. An alternative proposal comes in the form of what I will call the *third-person proposal*. In this case, the human agent is the primary target of prediction-oriented learning. The goal is thus to monitor the responses of the human individual and predict those responses relative to features of the individual's local environment. The upshot, perhaps, is a generative model that embodies something about the character of an individual—their propensity to act in particular ways in certain situations, their tendency to spend their hard-earned money on certain commodities, and their likely linguistic responses to certain kinds of conversational context.

These two proposals differ with respect to the kind of 'sensory' data over which generative models are formed. In both cases, however, we see a commitment to uncovering the deep structuring causes or latent variables that best explain the shape of the evolving sensory signal. Of the two proposals, the first-person proposal is likely to be the one that best serves the interests of the digital immortality agenda. But this does not mean that the third-person proposal is entirely without merit. It is, indeed, the third-person proposal that probably best reflects the current interests of the technological community in preserving some trace of an individual (e.g., by maintaining repositories of an individual's social media posts). Crucially, in thinking about the *kind* of data that can be used to shape the dynamical profile of PP systems, we are provided with some insight into the shape of what might be called the 'morphospace of the mind' (i.e., the universe of all possible minds) (see Mitteroecker and Huttegger, 2009),[6] as well as (perhaps) additional means of achieving digital immortality. Consider, for example, the way in which current deep learning systems are being used to create generative models of the human connectome (Betzel et al, 2016), or the way in which so-called 'brain reading' devices are yielding new opportunities to model the real-time response profile of the biological brain using prediction-oriented learning and generative modelling techniques (see van Gerven et al, 2010).

---

[6] The general idea, here, is that different kinds of data environment provide the basis for different kinds of mind, with phenomenological differences linked to the causal mechanisms that operate in each environment. A generative model of the human social environment, for example, may yield a *mind of society* that tracks the hidden causal structure of social mechanisms. Inasmuch as subjective experiences are tied to the properties of generative models, then such a mind may yield a subjective reality that is profoundly different from the sort of 'reality' we know (or could, perhaps, imagine).

5 Memories For Life (And Beyond)

The form of digital immortality envisioned in the previous section relies on the availability of substantial bodies of data about the individual. From this perspective, the process of 'generating me' emerges as a particular form of big data challenge—one that dovetails with the interests and concerns of a number of research communities. In yielding the data for prediction-oriented learning processes, it should thus be clear that the vision of digital immortality scouted above establishes a natural point of contact with the emerging discipline of data science (Committee on the Analysis of Massive Data, 2013), especially when it comes to the acquisition, analysis, and storage of big data. Of particular interest is work that seeks to use big data for the purpose of enhanced prediction, typically by drawing on a capacity for predictive modelling (Dhar, 2013). Another prominent point of interest concerns the application of deep learning methods to big data assets. As noted by Najafabadi et al. (2015), the ability to detect statistical regularities in data using unsupervised learning techniques makes deep learning particularly well-suited to dealing with some of the analytic challenges associated with big data science.

It has to be said, of course, that the data sets targeted by contemporary data science are unlike those that form the basis of the aforementioned first-person proposal. For the most part, the term "big data" is often applied to bodies of data pertaining to scientific observations (e.g., astronomical data) or data that tracks the properties of *multiple* human individuals (e.g., epidemiological data). There is, however, a growing interest in the analysis of data that is gleaned from individual human agents. Indeed, the analysis of individual (or personal) data forms the basis of research into so-called *personal informatics*, which is being driven, at least in part, by the availability of new digital recording devices, such as wearable cameras, smartphones, and activity monitors. The use of technology to record or track information about individual human subjects is also a central element of work that goes under the heading of the *quantified-self* (Lupton, 2013; Swan, 2013),[7] which has highlighted the way in which self-tracking technologies can be used to record data relating to (e.g.) body weight, energy levels, time usage, heart rate, body temperature, exercise patterns, sleep quality, sexual activity, diet, dreams, and blood chemistry. It is at this point, perhaps, that some of the data-oriented challenges associated with the aforementioned vision of digital immortality—specifically those pertaining to the acquisition of training data—look to be a little less formidable. There is no doubt a sense in which our current modes of self-related data acquisition remain deficient relative to the requirements of the digital immortality vision. But there is, I suggest, no reason to doubt the overall feasibility of the data tracking effort: our current technologies are already yielding ample data about a range of physiological and behavioural variables, and such data is already recognized as a valuable source of information, with the use of machine learning and other forms of big data analysis

[7] A variety of other terms are sometimes used to refer to the same phenomenon. These include "lifelogging," "measured me," "self-tracking," and "self-surveillance."

poised to reveal hidden structure in our self-generated digital trails (Fawcett, 2015; Phan et al, 2017; Hoogendoorn and Funk, 2018).

Future tracking technologies are likely to expand both the volume and variety of data that can be acquired from individuals during the course of their lives, altering the opportunities for prediction-oriented learning and the acquisition of generative models. When it comes to our ability to monitor individual movements, for example, research into so-called artificial skin (or e-skin) devices (Yokota et al, 2016; Someya et al, 2004) and smart fabrics (Foroughi et al, 2016; Wang et al, 2014) is likely to be particularly significant. In supporting the acquisition of data about physical movement, such technologies may provide a means to track information of a proprioceptive nature. This looks to be important inasmuch as we see the route to digital immortality as predicated on the attempt to recapitulate the sorts of predictive processing undertaken by the biological brain. From this perspective, digital recording technologies function as a substitute for biological sensors, recreating the kind of information streams that characterize the sensorium of the biological individual. In this respect, there seems to be ample cause for optimism. In addition to the monitoring of information of a (broadly) proprioceptive nature, technological advances are likely to improve our ability to record information from both the corporeal and extra-corporeal environment (i.e., information of a broadly interoceptive and exteroceptive nature). The advantage of this particular approach to data acquisition is that it depicts a SPP system as in more or less the same position as the biological brain. Both the brain and its synthetic counterpart are thus attempting to establish a predictive grip on common bodies of sensory information, and they are thus under more-or-less the same pressure to acquire and deploy similar generative models. It is this particular approach—the approach mandated by the first-person proposal in Section 4—that is perhaps best placed to deliver the most potent forms of digital immortality. For the goal of digital immortality is not merely to learn about a specific individual as an object of study (as per the third-person proposal); it is rather to duplicate the generative structures that make a particular individual the person they are. Inasmuch as we see the elements of the self—our memories, our personalities, our hopes, our fears, and our dreams—as inhering in the structure of a complex multi-layer network that is progressively shaped by our contact with the sensorium, then the quest for digital immortality is perhaps best served by presenting SPP systems with the same bodies of sensory data that confront their neurobiological counterparts.

Is any of this remotely feasible? There are, to be sure, plenty of challenges that confront the attempt to record personal data and use it for the purpose of recreating the functional profile of a brain-based generative model. To my mind, the challenge of emulating the biological brain's predictive and generative capabilities is one of the more daunting challenges, and certainly one that is more daunting than the challenge of acquiring large-scale bodies of personal data. Although research into predictive processing and deep learning exhibits a degree of convergence on hierarchical organizations, generative models, and (to a lesser extent) prediction-oriented learning, there remains a somewhat worrying gap in our understanding of how to emulate the representational and computational wherewithal of the biological brain. Inasmuch as progress in this area lags behind our capacity to capture and store the data that will ultimately be used to train future forms of deep learning system, perhaps there is room within the emerging discipline of data science for a program of research devoted to *data cryogenics*—a field of scientific and engineering research that seeks to preserve bodies of digital data until such times as deep learning systems are deemed able to raise the dead. To me at least, this idea seems no less plausible, and no more outlandish, than the forms of resurrection that are envisioned by those advocating conventional forms of cryogenic preservation.

There is no doubt a further worry raised by all this talk of digital tracking and data monitoring—one that is already felt by those who are ever-more intimately connected to a surrounding penumbra of digital devices. The concern is that the PP route to digital immortality is one that feeds directly into existing fears about digital surveillance and privacy violation. There is clearly a sense in which such fears, at least as they relate to the present analysis, are justified. Personally, I do not doubt that the sort of vision outlined here will require some degree of privacy violation, and I doubt whether technological advances will do much to assuage such fears, especially if such data is to be 'handed' over to third parties for safekeeping. Perhaps this is something that individuals will need to decide for themselves. Ultimately, it may be the case that privacy is just the price we pay for the possibility of everlasting life.

6 The Afterlife

Your biological life has come to an end. You have done your best to weave a digital fabric that tracks the defining moments of your existence. You have, you hope, created some nice memories for your SPP system to capture and model, and you hope that in the process of recreating those moments, some part of you will be preserved. There have, of course, been ups and downs, disasters as well as triumphs, brief moments of happiness punctuated with perhaps longer periods of despair. It does not matter now. Your life is over. Time to die.

But is that necessarily the end of the story? Does the departure of your biological body mean that you yourself are gone, irrevocably lost to those you loved and to those who loved you in return? If anything I have said thus far is anywhere near the mark, then it should be clear that the tale is not quite over. There is, it seems, space for a few pages more.

What, then, is the final part of this vision of digital immortality? What happens once the biological body has fulfilled its purpose and been laid to rest? Arguably, no form of digital immortality is complete without a corresponding form of digital resurrection. But what is the nature of this resurrection relative to the present vision of a SPP system that seeks to emulate the generative capabilities of the biological brain? There are, I suspect, many options available here, but I will choose to limit my attention to the role of virtual reality technologies in supporting a digital afterlife.

Consider first the idea that advances in holographic computing could be used to render an individual in holographic form. This idea is perhaps best exemplified by the character, Joi, in the movie *Blade Runner 2049*. Joi is a virtual companion for the movie's main protagonist, K, who is a replicant blade runner. Unlike K, Joi has no substantive physical presence in the world. She is instead a being made of light; a cinematic entity projected *into* an onscreen (physical, albeit fictional) world. There are, of course, no real-world counterparts to Joi at the time of writing—for the time being at least, she exists solely in the realms of fiction and fantasy. There are, however, reasons to think that something akin to a Joi-like entity might be possible once we reach the midpoint of the 21st century—once our own timeline coincides with the timeline of the *Blade Runner 2049* universe. In this respect, it is worth noting the recent progress that has been made in the development of mixed reality devices. One example is the Microsoft HoloLens, which renders virtual objects (called holograms) within the local physical environment of a human user. Other research establishes an even closer alignment with the *Blade Runner 2049* vision. Consider, for example, research into so-called volumetric displays, which render virtual objects as three-dimensional light displays that can be viewed by multiple users (from multiple angles) without the use of headsets or other user-worn technology (e.g., Smalley et al, 2018). Indeed, in some respects, the capabilities of today's holographic technologies have already surpassed that depicted in *Blade Runner 2049*. In the movie, Joi is a character who can be seen but not touched; her status as a hologrammatic entity precludes the possibility of physical contact and this complicates the nature of her relationship with her physical companion, K. Relative to the thematic structure of the movie, of course, Joi's ethereality is important, for it encourages us to reflect (*inter alia*) on the 'reality' of relationships that transcend the physical/virtual divide. In the real world, however, our interactions with holograms may be far less intangible affairs. Recent research has thus already demonstrated the possibility of so-called touchable or haptic holograms—holograms that can not just be seen, but touched, felt, and even moved (see Kugler, 2015).

Here, then, is one of the possibilities for a digital afterlife: individuals will be resurrected as hologrammatic entities—entities that 'live' among us as virtual 'ghosts'. These will be beings whose perceptuo-motor exchanges with the real world are driven by whatever generative models were acquired as part of a biological life within that very same world. Such beings will sense the world (via technological sensors) and implement actions within that world (via changes in photonic rendering technology). Whether they will be able to interact with us, in the sense of being able to touch us, remains to be seen. We may, however, still be 'moved' by the presence of these virtual souls, even if the more tactile elements of a human relationship should fail to survive the transition to holographic 'heaven'.

Of course, one way of dealing with the problems thrown up physical reality is to retreat from it altogether. Perhaps, then, a second possibility for the digital afterlife is to situate an individual's SPP system within a purely virtual environment, similar to those built around the use of contemporary game engines. This scenario will probably require little in the way of an introduction, for the idea of a life within a virtual (sometimes self-created) world is one that has been explored by a number of cultural products. Movies such as *The Matrix* and *The Thirteenth Floor* deal with the more general notion of life within a purely virtual environment, while the post-mortem possibilities of virtual environments are explored by the movie *Vanilla*

Sky and (my personal favourite) the *San Junipero* episode of the *Black Mirror* sci-fi series.

Both these scenarios rely on the use of virtual reality technologies to address the challenges posed by the demise of an individual's biological body.[8] The technological challenges associated with these scenarios are no doubt immense, but the appeal to virtual reality is also apt to raise a host of philosophical concerns and worries. Perhaps one of the more pressing concerns comes from a consideration of what is lost during the process of biological death. The loss of the biological body is particularly worrisome, since there are reasons to think that the body is a crucial component of the human cognitive system, yielding a range of opportunities for intelligent action (Clark, 2008) and mediating our emotional responses to both actual and counterfactual states-of-affairs (Damasio, 1996). Such insights are not lost on those who work from within the PP camp. Seth (2013), for example, suggests that the processing of interoceptive information deriving from the non-neural bodily environment is relevant to some aspects of conscious experience, with a variety of "subjective feeling states (emotions)... arising from actively-inferred generative (predictive) models of the causes of interoceptive afferents" (p. 565). Similarly, Hohwy and Michael (2017) present an intricate and intriguing account of the role of the biological body in giving rise to a sense of self.

Relative to these claims, it is far from clear that the attempt to replicate the generative capacities of the biological brain will be enough for the digital afterlife, especially if what we want to achieve is a state-of-affairs in which a given individual is resurrected as a sentient being, capable of enjoying (and enduring) the rich panoply of emotional states and conscious experiences that characterized their biological life. It is in this sense, perhaps, that the present account of digital immortality may be seen to be inadequate, focusing, as it does, on the biological brain at the expense of a larger, material fabric that includes the individual's biological body and certain aspects of their local extra-organismic environment (see Clark, 2008). The critic will no doubt want to highlight the indispensable role of the biological body in realizing certain aspects of the human cognitive economy, with a disembodied form of intelligence perhaps counting as no form of intelligence at all. They will also perhaps be inclined to view the afterlife options listed above as failing to address this problem. Talk of insubstantial hologrammatic ghosts and virtual world simulations are unlikely to do justice, they will say, to the role the biological body plays in shaping (and perhaps even realizing) the complex array of experientially-potent states that to a large extent make our lives worth preserving in the first place. From this perspective, perhaps the very best we could hope for would be some form of experientially-diminished afterlife—one in which our continued existence (if we care to call it that) comes at the expense of an ability to experience emotional states or to even have a sense of oneself as an entity that continues to exist. Perhaps a hologrammatic ghost, for

[8] These scenarios do not, of course, exhaust the possibilities for digital resurrection. In addition to virtual reality technologies, the 21st century is likely to see significant advances in the development of biomimetic materials, 3D printing technology, and robotic systems. These may open the door to a more concrete form of digital afterlife, one in which the biological body is substituted with a synthetic, but no less substantial, corporeal presence.

example, is a prime candidate for a virtual version of what is dubbed the Cotard delusion—a psychiatric disorder in which the affected individual holds the delusional belief that they are dead or do not exist (Young and Leafhead, 1996). This seems particularly likely in the wake of recent analyses of the Cotard delusion, which link the delusion to anomalies in bodily experience (Gerrans, 2015) or aberrations in the processing of interoceptive information (Seth, 2013). Of course, a delusion is only a delusion if the convictions of the relevant individual do not align themselves with reality. In this sense, it is doubtful that are any genuine cases of the Cotard delusion in the afterlife. Believing you are dead in the afterlife is not delusional; what is delusional is to believe that you are alive when you are, in fact, dead. (No one said that the discipline of thanato-psychiatry would be straightforward!)

There is no doubt much here that is contentious, and I will not have the space to cover (let alone resolve) all the issues that are likely to animate future discussions in this area. It is worth noting, however, that nothing in the PP approach to digital immortality seeks to deny the importance of the body in mediating our cognitive engagements with the world, shaping our emotional responses, or, indeed, realizing aspects of conscious experience. In this sense, the biological body remains an important aspect of the human cognitive economy and a relevant target of generative models—hence the emphasis on tracking body-related information in Section 5. What is perhaps more problematic is the extent to which the various forms of afterlife I have described—holograms, virtual characters, and so forth—are properly characterized as lacking a body. It is here, I suggest, that it pays to make a distinction between what Wheeler (2013) dubs *implementational materiality* (which involves a commitment to the idea that the body is no more than a material realizer of functionally specified cognitive roles) and *vital materiality* (according to which the body makes a non-substitutable contribution to cognitive states and processes). It should be clear that given the choice between these two options, it is only the commitment to vital materiality that poses any real threat to the prospect of virtual forms of embodiment. From a functional standpoint, therefore, I suggest that there is no real reason to regard a holographic entity or the inhabitant of a purely virtual world as congenitally condemned to a disembodied existence. Providing the functional contributions of the biological body can be replicated in virtual form, a hologram, I submit, counts as just as much an embodied entity as does an agent that has a more substantive physical presence.

7 Conclusion

The present chapter outlines an approach to digital immortality that is rooted in recent advances in theoretical neuroscience and machine learning. In line with other approaches to digital immortality, the present proposal highlights the importance of collecting and storing data about an individual, with a view to using that data for the purpose of digital resurrection. The difference between the present proposal and other accounts relates to the kind of data that is acquired, the way in which the data is analysed, and the kind of computational substructure that is deemed relevant to the digital immortality agenda. The claim is that some hierarchically-organized PP system—some variant of today's deep learning systems—could engage in a form of lifelong learning, attempting to build generative models that tackle the same sort of predictive challenges as those confronting the biological brain. Such models, it is suggested, will—by dint of the attempt to minimize prediction error—resemble those acquired by the biological brain as it attempts to secure a predictive grip on the sensorium. Inasmuch as we see these synthetic generative models as capturing the essential elements of who and what we are—models whose generative capabilities reflect our own biological capacity to render our realities, recall our pasts, and create our futures—then they may provide the means by which some aspect of ourselves is able to persist long after the biological body has withered away. The claim, in short, is that a hierarchically-organized predictive processing machine may serve as a vehicle that sustains our dreams as we inexorably succumb to the "sleep of death."

And what of poor Hamlet and his post-existential woes? Hamlet wonders whether it is better for him to die than to face up to his earthly troubles. But he worries that his death will be occasioned by dreams that merely serve to prolong his suffering. It is at this point, of course, that issues of technical feasibility come face-to-face with a host of more normative concerns. Just because digital immortality is possible (if, indeed, it is possible), does this mean that we should seek to make it actual? As a species we have done our best to preserve human life, and we have, I suppose, become somewhat good at it (even if many other biological species have had to pay the price). But have we done enough to ensure that the world in which we live is one that is worth living, as opposed to one that is worth leaving? Perhaps, then, the issue that lies at the heart of debates about digital immortality is not so much the technical obstacles that lie on the road ahead, as whether the route to digital immortality is one that is itself worth pursuing. Should we rage, rage against the dying of the light and resist the rule of the second law? Or should we accept that all dreams must end in a darkened room? To dream? To die? To be, or not to be? Ay, there's the rub.

Acknowledgements I would like to thank two anonymous referees for their helpful comments on an earlier draft of this material. This work is supported under SOCIAM: The Theory and Practice of Social Machines. The SOCIAM Project is funded by the UK Engineering and Physical Sciences Research Council (EPSRC) under grant number EP/J017728/1 and comprises the Universities of Southampton, Oxford, and Edinburgh.

References

Banerji M, Lahav O, Lintott CJ, Abdalla FB, Schawinski K, Bamford SP, Andreescu D, Murray P, Raddick MJ, Slosar A (2014) Galaxy Zoo: Reproducing galaxy morphologies via machine learning. Monthly Notices of the Royal Astronomical Society 406(1):342–353

Bengio Y (2009) Learning deep architectures for AI. Foundations and Trends in Machine Learning 2(1):1–127

Betzel RF, Avena-Koenigsberger A, Goñi J, He Y, de Reus MA, Griffa A, Vértes PE, Mišic B, Thiran JP, Hagmann P, van den Heuvel M, Zuo XN, Bullmore ET, Sporns O (2016) Generative

models of the human connectome. NeuroImage 124:1054–1064
Clark A (2008) Supersizing the Mind: Embodiment, Action, and Cognitive Extension. Oxford University Press, New York, New York, USA
Clark A (2013a) Expecting the world: Perception, prediction, and the origins of human knowledge. The Journal of Philosophy 110(9):469–496
Clark A (2013b) Whatever next? Predictive brains, situated agents, and the future of cognitive science. Behavioral and Brain Sciences 36(3):181–253
Clark A (2016) Surfing Uncertainty: Prediction, Action and the Embodied Mind. Oxford University Press, New York, New York, USA
Committee on the Analysis of Massive Data (2013) Frontiers in Massive Data Analysis. The National Academies Press, Washington D.C., USA
Damasio AR (1996) Descartes' Error: Emotion, Reason and the Human Brain. Papermac, London, UK
Dhar V (2013) Data science and prediction. Communications of the ACM 56(12):64–73
Dieleman S, Willett KW, Dambre J (2015) Rotation-invariant convolutional neural networks for galaxy morphology prediction. Monthly Notices of the Royal Astronomical Society 450(2):1441–1459
Faisal AA, Selen LP, Wolpert DM (2008) Noise in the nervous system. Nature Reviews Neuroscience 9(4):292–303
Fawcett T (2015) Mining the quantified self: Personal knowledge discovery as a challenge for data science. Big Data 3(4):249–266
Foroughi J, Mitew T, Ogunbona P, Raad R, Safaei F (2016) Smart fabrics and networked clothing: Recent developments in CNT-based fibers and their continual refinement. IEEE Consumer Electronics Magazine 5(4):105–111
Friston K (2002) Beyond phrenology: What can neuroimaging tell us about distributed circuitry? Annual Review of Neuroscience 25(1):221–250
Friston K (2010) The free-energy principle: A unified brain theory? Nature Reviews Neuroscience 11(2):127–138
Friston K, Thornton C, Clark A (2012) Free-energy minimization and the dark-room problem. Frontiers in Psychology 3(130):1–7
Friston KJ, Stephan KE, Montague R, Dolan RJ (2014) Computational psychiatry: The brain as a phantastic organ. The Lancet Psychiatry 1(2):148–158
Gerrans P (2015) All the self we need. In: Metzinger TK, Windt JM (eds) Open MIND: Philosophy and the Mind Sciences in the 21st Century, MIND Group, Frankfurt am Main, Germany, pp 1–19
van Gerven MA, de Lange FP, Heskes T (2010) Neural decoding with hierarchical generative models. Neural Computation 22(12):3127–3142
Goertzel B, Ikle' M (2012) Special issue on mind uploading: Introduction. International Journal of Machine Consciousness 4(1):1–3
Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y (2014) Generative adversarial nets. In: Ghahramani Z, Welling M, Cortes C, Lawrence N, Weinberger K (eds) Advances in Neural Information Processing Systems, Montreal, Canada, vol 27, pp 2672–2680
Hayworth KJ (2012) Electron imaging technology for whole brain neural circuit mapping. International Journal of Machine Consciousness 4(1):87–108
Hinton GE (2007a) Learning multiple layers of representation. Trends in Cognitive Sciences 11(10):428–434
Hinton GE (2007b) To recognize shapes, first learn to generate images. In: Cisek P, Drew T, Kalaska J (eds) Computational Neuroscience: Theoretical Insights into Brain Function, vol 165, Elsevier, Amsterdam, The Netherlands, pp 535–547
Hinton GE (2010) Learning to represent visual input. Philosophical Transactions of the Royal Society of London B: Biological Sciences 365(1537):177–184
Hinton GE, Salakhutdinov RR (2006) Reducing the dimensionality of data with neural networks. Science 313(5786):504–507

Hohwy J, Michael J (2017) Why should any body have a self? In: de Vignemont F, Alsmith AJT (eds) The Subject's Matter: Self-Consciousness and the Body, MIT Press, Cambridge, Massachusetts, USA, pp 363–391
Hoogendoorn M, Funk B (2018) Machine Learning for the Quantified Self. Springer, Cham, Switzerland
Horn BKP (1977) Understanding image intensities. Artificial Intelligence 8(2):201–231
Kugler L (2015) Touching the virtual. Communications of the ACM 58(8):16–18
LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521(7553):436–444
Lintott CJ, Schawinski K, Slosar A, Land K, Bamford S, Thomas D, Raddick MJ, Nichol RC, Szalay A, Andreescu D, Murray P, van den Berg J (2008) Galaxy Zoo: Morphologies derived from visual inspection of galaxies from the Sloan Digital Sky Survey. Monthly Notices of the Royal Astronomical Society 389(3):1179–1189
Lupton D (2013) Understanding the human machine. IEEE Technology and Society Magazine 32(4):25–30
Ma J, Sheridan RP, Liaw A, Dahl GE, Svetnik V (2015) Deep neural nets as a method for quantitative structure–activity relationships. Journal of Chemical Information and Modeling 55(2):263–274
Metzinger T (2003) Being No One: The Self-Model Theory of Subjectivity. MIT Press, Cambridge, Massachusetts, USA
Mitteroecker P, Huttegger SM (2009) The concept of morphospaces in evolutionary and developmental biology: Mathematics and metaphors. Biological Theory 4(1):54–67
Najafabadi MM, Villanustre F, Khoshgoftaar TM, Seliya N, Wald R, Muharemagic E (2015) Deep learning applications and challenges in big data analytics. Journal of Big Data 2(1):1–21
Phan N, Dou D, Wang H, Kil D, Piniewski B (2017) Ontology-based deep learning for human behavior prediction in health social networks. Information Sciences 384:298–313
Ravanbakhsh S, Lanusse F, Mandelbaum R, Schneider JG, Poczos B (2017) Enabling dark energy science with deep generative models of galaxy images. In: Singh S, Markovitch S (eds) Thirty-First AAAI Conference on Artificial Intelligence, AAAI Press, San Francisco, California, USA, pp 1488–1494
Revonsuo A (1995) Consciousness, dreams and virtual realities. Philosophical Psychology 8(1):35–58
Seth AK (2013) Interoceptive inference, emotion, and the embodied self. Trends in Cognitive Sciences 17(11):565–573
Seung HS (2012) Connectome: How the Brain's Wiring Makes Us Who We Are. Houghton Mifflin Harcourt Publishing Company, New York, New York, USA
Smalley D, Poon TC, Gao H, Kvavle J, Qaderi K (2018) Volumetric displays: Turning 3-D inside-out. Optics and Photonics News 29(6):26–33
Someya T, Sekitani T, Iba S, Kato Y, Kawaguchi H, Sakurai T (2004) A large-area, flexible pressure sensor matrix with organic field-effect transistors for artificial skin applications. Proceedings of the National Academy of Sciences 101(27):9966–9970
Sporns O, Tononi G, Kötter R (2005) The human connectome: A structural description of the human brain. PLoS Computational Biology 1(4):e42
Swan M (2013) The quantified self: Fundamental disruption in big data science and biological discovery. Big Data 1(2):85–99
Wang Y, Wang L, Yang T, Li X, Zang X, Zhu M, Wang K, Wu D, Zhu H (2014) Wearable and highly sensitive graphene strain sensors for human motion monitoring. Advanced Functional Materials 24(29):4666–4670
Wheeler M (2013) What matters: Real bodies and virtual worlds. In: Harvey I, Cavoukian A, Tomko G, Borrett D, Kwan H, Hatzinakos D (eds) SmartData: Privacy Meets Evolutionary Robotics, Springer, New York, New York, USA, pp 69–80
Yokota T, Zalar P, Kaltenbrunner M, Jinno H, Matsuhisa N, Kitanosako H, Tachibana Y, Yukita W, Koizumi M, Someya T (2016) Ultraflexible organic photonic skin. Science Advances 2(4):e1501856

Young AW, Leafhead KM (1996) Betwixt life and death: Case studies of the Cotard delusion. In: Halligan PW, Marshall JC (eds) Method in Madness: Case Studies in Cognitive Neuropsychiatry, Psychology Press Ltd, New York, New York, USA, pp 147–171