by metadata attached to songs when they are uploaded to Spotify by artists and recording labels. This metadata will include the song title, the artist's name and perhaps the genre. All of this information enables the model to find prospective listeners, but a collaborative filtering model needs many users to listen to a song before it can begin to determine patterns and make predictions about who might like it. Spotify has experimented with unsupervised deep-learning approaches to complement their collaborative filtering model (Popper 2015).

Engineers working at Spotify experimented with a deep-learning approach to identify patterns in the sonic structure of the songs themselves (Dieleman 2014). The deep neural network they engineered was able to cluster together songs that shared similar sonic qualities. These included songs with similar chord progressions, vocal melodies and rhythms. Many of these clusters corresponded with cultural genres and subgenres. The unsupervised model was able to identify sonic similarities between songs that humans had not labelled. Rather than 'label' the songs, an unsupervised model learns patterns in the data that enables it to cluster songs based on how 'proximate' they are to each other. The model makes a decision by building a 'latent space' where all the songs are clustered as schematically near or far from each other. If a user likes a song, then the model recommends a song that has been clustered 'nearby'.

The automated models that shape the digital platforms we use are comprised of multiple intersecting models. Seaver (2022: 49–50) explains that on digital platforms

> 'the algorithm' is not one algorithm at all, but 'dozens and dozens of algorithms,' which parse various signals: What does a song sound like? What device is a user listening on? What has a user listened to in the past? All those signals are orchestrated together into an 'ensemble' by another algorithm, which is tuned to a listener's particular listening style.

Platforms describe this ensemble of algorithmic models as a 'vast experimental apparatus', and Seaver describes them as 'traps' that are 'fine-tuned to capture the attention of users' who are their 'prey' (Seaver 2022: 51). Importantly, though, as we will discuss in the next chapter, this is not a linear and deterministic relationship. Traps are relational. Just as they attempt to design an apparatus that will 'capture' our attention, the 'prey' learns over time about the trap, how to avoid it, outwit it and even turn it to its advantage.

Generative AI

The approaches to classification, prediction and decision making developed in supervised and unsupervised machine learning models underpin the move to generative artificial intelligence over the past decade. While the launch of ChatGPT in late 2022 signalled the moment generative AI swept into the public imagination and became a technology of fascination and concern, machine learning models that could synthesize, augment and generate 'original' text, audio and images had been emerging for many years. In 2017, Justin Hendrix, the Director of the NYC Media Lab, posted to Twitter the following scenario that sounded an ominous warning about 'artificial content':

> Trust in the media is extremely low right now, but I think it may have a lot further to go, driven by new technologies. In the next few years technologies for simulating images, video and audio will emerge that will flood the internet with artificial content. Quite a lot of it will be very difficult to discern – photorealistic simulations. Combined with microtargeting it will be very powerful. After a few high-profile hoaxes, the public will get the message – none of what we see can be trusted. It might all be doctored in some way. Researchers will race to produce technologies for verification, but will not be able to keep up with the flood of machine generated content. The platforms will attempt to solve the problem, but it will be vexing. Some governments will look for ways to arbitrate what is real. The only way out of this now is to spend as much time trying to understand the externalities of these technologies as we do creating them. But this will not happen, because there is no market mechanism for it.

We present Hendrix's prediction here because, nearly a decade on, it might be helpful to ask ourselves: did any of this turn out to be true? It helps to remind ourselves, too, that generative AI didn't just miraculously appear, it was being worked on, developed and anticipated in the decade before ChatGPT's launch. Importantly, we can understand his point here a bit differently from the way he puts it; that is, if we thought more about the possible social and political impacts of new technologies as they were being developed, rather than thinking only about their possible markets and economic value, we might understand that many people have much to gain from careful consideration of 'the externalities' of technology.

A month after Hendrix's prediction, computer scientists at the University of Washington reported that they had produced an AI model that could create an extremely 'believable' video of Barack Obama appearing to say things he had said in another context:

> Given audio of President Barack Obama, we synthesize a high-quality video of him speaking with accurate lip sync, composited into a target video clip. Trained on many hours of his weekly address footage, a recurrent neural network learns the mapping from raw audio features to mouth shapes. Given the mouth shape at each time instant, we synthesize high quality mouth texture, and composite it with proper 3D pose matching to change what he appears to be saying in a target video to match the input audio track. Our approach produces photorealistic results. (Suwajanakorn 2017: 1)

In the late 2010s we began to see the first signs of a shift from neural networks that could classify and cluster text, images and video to neural networks that could synthetically generate text, images and videos from a range of prompts.

Technology companies like OpenAI, MidJourney and Stability AI began to develop models that can use enormous training sets of text and images to generate 'original'

outputs. They do this by learning the patterns and relationships within texts and images, for instance, words that appear together in texts or patterns of pixels in images. They then reproduce these patterns as new outputs in response to prompts. These generative AI models appear to produce 'new' content, but because they are trained on, and bound by, the archives of data they have access to, what they really do is recompose and remix existing ideas. They tend to reproduce the ideas and styles that already exist in our written and visual cultures.

Generative models signal an important shift in our digital cultures where our experience of algorithmic, platform media will no longer be simply the automated recommendation of content but also the automated synthesis, generation and augmentation of content. This shift will change our sense of reality in ways that are hard to anticipate. For instance, generative AI will increasingly be used to create 'synthetic data' (Steinhoff 2022), data not collected from something but generated by software to supplement missing parts in existing datasets. As other machine learning algorithms incorporate synthetic data into their datasets as the 'ground truth', or foundation, for them to be trained on, their suggestions, and our decisions based on these suggestions, will no longer be fully rooted in observable phenomena collected from the real world.

The main reason that platforms are key sites for the development of machine learning approaches is that they collect and control enormous sets of labelled and unlabelled social and cultural data. They are commercially motivated to invest significant capital in the development of new techniques for processing that data as part of their effort to automate the creation and circulation of content and products.

Algorithmic culture

Cultures are, in some fundamental ways, systems of judgement and decision making. Cultural values, preferences and tastes are all systems of judging ideas, objects, practices and performances: as good or bad, cool or uncool, pleasant or disgusting, and so on. When we are taught to speak, we are learning to correctly assign labels in our culture. Striphas (2015: 396) argues that 'over the last 30 years or so, human beings have been delegating the work of culture – the sorting, classifying and hierarchizing of people, places, objects and ideas – increasingly to computational processes'. This is remarkable in part because in Eurocentric traditions, we have long understood culture, and its systems of judgement, to be confined to the human experience. We are now creating computational machines that can simulate these forms of judgement. An algorithmic culture involves 'the use of computational processes to sort, classify, and hierarchize people, places, objects, and ideas, and also the habits of thought, conduct and expression that arise in relationship to those processes' (Hallinan and Striphas 2016: 119). There are two important parts of this definition. Algorithmic culture involves both machines learning to make decisions about culture and humans learning to address those machines. Writing about video games in 2006, Alexander Galloway explained that a game is a kind of cultural text where we are not so much the 'reader' as the 'operator'.