FROM ALGORITHMS TO ARTIFICIAL INTELLIGENCE

Since around the 2010s, the development of AI has intensified (Russell & Norvig 2022). Advancements in the area of machine learning and neural networks have been key drivers in propelling the field forward (LeCun et al. 2015), and leading to various forms of AI-driven systems and solutions becoming implemented in a growing number of areas in society (Brynjolfsson & McAfee 2016). This means that in today's world, AI has left the realm of science fiction to become an integral part of our daily lives, for example by being used in different forms of decision-making systems, healthcare applications, logistical operations, financial services, as well as in the creative industries (Agrawal et al. 2018; Anantrasirichai & Bull 2022; Chui et al. 2016; Topol 2019). Naturally, of course, the ever more advanced capabilities of AI also impact on the realm of digital media. One of the key dimensions here relates to the use of various forms of trace data for training machine learning models that can be used in content recommendation systems, such as various 'what to listen to/watch next' features on digital media platforms, in crafting personalised social media feeds (with all of the good and bad that entails), and in dynamically targeting advertisements that are tailored to individual users (Gomez-Uribe & Hunt 2016; Shumanov et al. 2021).

The datafication of society is indeed a crucial driving force for this development, as the traces that digital media users leave behind through clicks, views, shares, and other forms of interaction provide AI corporations with a rich repository of data that can be used in machine learning to discern patterns, preferences, and behaviours (see, the notion of 'data commodity' in Chapter 8). In the next step, the trained models can be put to use in increasingly sophisticated ways, enabling platforms to not only predict user preferences with quite remarkable accuracy but also to potentially influence those preferences through deploying subtle nudges and recommendations, thereby shaping user experiences and behaviours in profound ways (Zuboff 2019).

In other words, then, one of the main ways in which AI impacts on the landscape of digital media is through its use for personalising the distribution of content. In this context, the increased refinement and black-boxedness that the evolving, and highly commercialised, AI systems bring to the table adds an extra layer of complexity to the ways in which algorithms – as discussed earlier – impact on how content is curated and presented to individual users, with the double-edged consequences of increased convenience and efficiency on the one hand, and echo chambers, polarisation, and potential manipulation on the other. We should note, however, that the term 'algorithms' is not equal to 'AI', and there is significant conceptual complexity here. When we speak of algorithms in popular debate and everyday discussions, it is generally taken to refer to the automated processes that determine what content we see on platforms like search engines, social media, and streaming services. It is important to note, though, that all such algorithms were not always driven by machine learning and AI. For example, the early forms of Google's search algorithm (known as PageRank) in the late 1990s primarily relied on a system that ranked web pages based on how many links pointed to them and what 'quality', in terms of relevance and authority, those links carried. The algorithm evolved over time to take into account several other factors beyond link analysis, but it was only much later, around 2015, that machine learning came to play a major role in Google's search algorithm. Similarly, many other platforms' algorithms have progressed over the years to incorporate increasingly AI-driven approaches. When it comes to the difference between mere 'algorithms' and (AI-driven) 'machine learning algorithms', computer scientist Pedro Domingos (2015: xi) writes that:

> Traditionally, the only way to get a computer to do something – from adding two numbers to flying an airplane – was to write down an *algorithm* explaining how, in painstaking detail. But *machine-learning algorithms*, also known as learners, are different: they figure it out on their own, by making inferences from data. And the more data they have, the better they get. Now we don't have to program computers; they program themselves. (Emphasis added)

From the earlier discussion, we can remind ourselves that the term 'algorithm' generally refers to a set of rules or a process that has been designed to perform a specific task. In the context of social media platforms, there will be many different forms of algorithms at play. Some of them can be very basic and rule-based, ones that are there to carry out the most straightforward tasks, such as sorting posts in chronological order or filtering content based on static and user-defined preferences (for example, blocking out the posts of some other accounts). Such algorithms will largely not be based on AI, but they can still be 'engineered' in certain ways as any algorithm will always be coloured by choices made by its designer. So-called heuristic algorithms are a bit more complex as they will involve some form of evaluation function to decide the course of action (Pearl 1984). They can involve elements of recognising patterns and learning from earlier data and can thus, in the context of digital media platforms, be used to make decisions that are more informed than those of simple rule-based algorithms. We can imagine, for example, a friend recommendation algorithm that relies on a set of pre-defined heuristics based on principles such as that users are more likely to know and connect with friends of their friends, that users with similar interests (in terms of pages liked, hashtags used, groups joined, etc.) are more likely to connect, that people in the same geographical area are more likely to know each other, and so on. Those of us familiar with platforms such as LinkedIn or Facebook will recognise from the friend recommendations offered to us that heuristics like these are probably at work behind the scenes here. Heuristic algorithms are, by definition, based on rules of thumb or experience-based techniques, which means that they are not inherently AI-based. However, their operation involves elements of pattern recognition and problem solving which means that they can incorporate elements of AI, especially in cases when they adapt and learn from new data over time.

Finally, there are the explicitly AI-based algorithms that draw on machine learning. Some of these also use 'deep learning' which is a particularly advanced form of machine learning that uses so-called neural networks – simulations of the structure and function of the human brain – to enable learning through multiple layers (hence 'deep') and thereby increase their capacity to discern complex patterns and relationships within data.

A machine learning system is centred around a 'model'. The model is created by training it on a dataset, which is most often of a very large size. The role of the dataset is to provide examples from which the model can learn patterns, relationships, and underlying structures within the data. For example, if the dataset consisted of a large number of Instagram user accounts with information about their posting patterns, interactions with other users, and engagement with different forms of content, the model would analyse those examples to understand how different factors may contribute to user engagement. The model, simply put, is then a mathematical representation of patterns in the data. After training, the model can be used to make predictions based on new data that it has not yet seen, based on the patterns that it has learned. In the Instagram example, the trained model could, for example, be leveraged to predict which types of posts are likely to generate the highest levels of engagement from users, or predict which users are the most likely to interact with certain types of content. By extension, then, those predictions could be used as input for various decisions that could be automated within the platform, such as tailoring content recommendations, optimising advertisement placements for maximum profits, or in other ways personalising user experiences. Machine learning systems, especially those based on deep learning, are able to identify subtle and complex patterns in data that it would be impossible for a human programmer to know about beforehand. Furthermore, as machine learning systems improve their performance continuously as they are exposed to more and more data over time, models developed in this way support algorithms that are much more flexible and efficient – but also much less transparent and scrutiniseable – than static, rule-based ones.

At a more concrete level, for example that of YouTube or Netflix recommender systems that use machine learning to analyse viewing histories, preferences, and other user behaviours, while the content recommendations and curation might become much more dynamic and capable than those performed through simpler heuristic algorithms, critical questions raised throughout this book and chapter also come to the fore. These are, for example, relating to privacy and data security, as in the discussions around data commodities and surveillance capitalism, and furthermore to issues surrounding the transparency of the processes that guide the recommendations. If we broaden the perspective to the social and political arena beyond the sphere of entertainment on streaming platforms, the potential consequences of the machine learning systems for democracy in the digital society also come into view. It has been widely debated that AI models often carry embedded values and ideologies within them (Crawford 2021; Eubanks 2017; Lindgren 2024). Examples of this include algorithmic perpetuation of discrimination and societal inequalities (Benjamin 2019; Noble 2018), algorithmic reinforcement of echo chambers, algorithmic amplification of misinformation, and other processes that limit the diversity of information that users are exposed to.

Something seems to happen when computers are put in control of discursive and political processes, Birhane (2020) for example argues that the tech industry, despite its celebration of disruption, tends to serve existing power relations. This may have to do with the fact that machine learning always looks backwards. It can learn from our history of patriarchy, racism, colonialism, wars, and heteronormativity, but knows nothing about alternative futures. To put it clearly: if an AI system is trained on data that contains skewed, biased, stereotypical, prejudiced patterns, and that underrepresents certain groups, it will perpetuate those things in its outputs, such as news articles or content recommendations. As argued by Bender et al. (2021: 617) in a paper on 'stochastic parrots', an AI language model (LM)

> is a system for haphazardly stitching together sequences of linguistic forms it has observed in its vast training data, according to probabilistic information about how they combine, but without any reference to meaning: a stochastic parrot. […] The ersatz fluency and coherence of [the models] raises several risks, [e.g.] the risks that follow from the LMs absorbing the hegemonic worldview from their training data.

In other words, the potential political consequences of using AI-based models on digital media platforms are profound. When automated systems of providing efficient content recommendations and for 'improving the user experience' spill over the boundaries of pure entertainment platforms to flood into critical areas such as news dissemination, public discourse, and civic engagement, the stakes become significantly higher (Lindgren 2023; 2024). Aside from indirect effects in terms of 'algorithmic bias' (Kordzadeh & Ghasemaghaei 2022), AI-based models can also be used for more malicious ends such as to manipulate public opinion by amplifying certain messages while suppressing others with severe implications for democratic processes as it can affect how people perceive various political issues and candidates. The manipulation of AI models can be quite deliberate as various actors, ranging from state-sponsored entities to independent interest groups, may exploit these technologies to steer public discourse directions that align with their ideologies and goals. Woolley and Howard (2019: 5) define the notion of *computational propaganda* as a socio-technical complex around variables, codes, algorithms, and models.

A number of different ethical challenges surround the uses of AI in relation to digital media. One of these relates to the lack of transparency in AI algorithms. As argued by Pasquale (2015: 3), we live now in a 'black box society' where the typical system is one where 'we can observe its inputs and outputs, but we cannot tell how one becomes the other'. This opacity comes with a wide range of problems for politics and democracy. Second, there are the challenges that relate to the issues of privacy and surveillance that have been repeatedly discussed above. Such ethical dilemmas have given rise to an entire area of its own around issues of ethics in AI (Dignum 2019; Steinhoff 2023; Vieweg 2021; Waelen 2022). For us as digital media researchers, AI also presents some significant challenges. The commercialised and proprietary nature of many of the dominant AI systems and models means that the opportunities to get access to the data and algorithms that drive the technologies are non-existent or severely limited. Such difficulties can, however, be overcome as methodological approaches that capture the social and political impacts of the systems – without the need for direct access to their technological elements – can be devised. Examples of such methods include, for example, algorithm ethnography and critical data studies, which allow researchers to infer the workings and effects of these opaque systems by examining their outputs, impacts, and the contexts in which they operate (boyd & Crawford 2012; Lindgren & Eriksson Krutrök 2024; Seaver 2017).

CHAPTER SUMMARY

This chapter has built further on some of the issues about the increased role of data in the evolving digital society (Chapters 1, 3, and 4), as well on the issues of exploitation (Chapter 8) and resistance (Chapter 9). We have looked more closely at the particular type of power that is exerted through the process of datafication, but also at how it can be challenged. I have introduced the notion that, following datafication, we live in an age of 'surveillance capitalism', and have thereby laid the foundation for critically analysing so-called big data. This chapter has also dug more deeply into the phenomenon of algorithms, trying to understand what they are and how these technological phenomena increasingly and sometimes covertly affect our social lives. There is a certain kind of algorithmic power being exerted, and we have looked at how some scholars have pointed out the need for increased algorithmic literacy. In the name of