1

Introduction

Rebecca Fiebrink, Marco Gillies, and Gonzalo Ramos

Artificial intelligence (AI) has seen huge, media-grabbing breakthroughs in the 2010s and 2020s, from the advent of deep learning, which first hit the public consciousness with the image-classification success of AlexNet (Krizhevsky et al., 2012) through the defeat of world champion Go grandmaster Lee Sedol by AlphaGo (Silver et al., 2016), to the rise of large language models (LLMs) such as GPT (OpenAI et al., 2023) that are seemingly able to generate text that is indistinguishable from human writing. There have probably been new advances reported between the time of us writing these words and you reading them (the rise of Diffusion models and LLMs, for example, has occurred during the writing of this book). In almost all cases these successes are the result of a specific approach to AI called machine learning (ML), and in particular an approach to ML called deep neural networks (DNNs; LeCun et al., 2015). The speed at which ML advances has surprised even experts, with successes that seemed impossible even a decade ago. These advances have also led to new anxieties around the technology, including the impacts of AI on jobs, bias in AI systems resulting in worsening social inequality, and even the long-term threat of human extinction due to AI. (See the works of Crawford (2022), Gebru,[1] Benjamin (2019), and Russell (2023) for discussions of these issues.)

Both the benefits and dangers stemming from ML and AI are very real. There is an urgent need to ensure that ML serves the needs of people rather than endangering them. In this book we take the position that ML is a technology and that we, as human-centered practitioners, have a long intellectual and practical tradition that aims to ensure that technologies meet human needs. In the field of computing technology, this discipline has many names, each with its own nuance: human–computer interaction (HCI), interaction design, user

[1] See the Distributed AI Research Institute. www.dair-institute.org.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 19:26:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781108872225.001

experience (UX), or human-centered computing. All these practices share a focus on how humans interact with technology rather than focusing simply on the technology itself. Moreover, researchers and practitioners in these disciplines often bring commitments to develop technology that not only surpasses requirements to be *usable*, or even merely *ethical*, but which also supports humanistic values and human well-being. The term "human-centered" can refer to all of these. Human-centered practices can not only help address the risks of ML, but they can also help make it a valuable and successful technology.

This book is for those looking to have a panoramic view of the field of human-centered ML research by providing a cross-section of the different types of works that can take place in it. We envision that the people who will benefit most from this book are students in HCI or ML, as well as researchers in adjacent fields who are looking to know more about what this emerging field is and how they might fit into it. We also envision the book being of interest to professionals working in and with technology, who can benefit from seeing what work in this space looks like, without the veneer of hype often encountered in popular media. We do this by bringing together researchers and practitioners sharing their diverse perspectives and experiences on how to apply human-centered approaches to ML.

Some of these perspectives manifest as detailed overviews and explanations of general fields of research. Others come in the form of interviews, where researchers and practitioners in the area of human-centered technologies share more personal insights into the field and what it is like to work in it. We have grouped the interviews into three chapters based on increasing scale of their primary focuses: People, Teams, and Humanity (though all interviews cover wide-ranging and diverse topics).

Before moving to the main text of the book, we will lay out a common ground for those unfamiliar with ML and human-centeredness. What is ML? What does it mean for technology to be human-centered? And, therefore, what is human-centered ML?

1.1 WHAT IS MACHINE LEARNING?

Machine learning is an approach to developing computer systems that allow people to generate a function (i.e., something that, given an input, will deliver an expected output) without explicitly providing a program or algorithm. Instead, one defines a function implicitly by providing example data. For example, an ML system for recognizing images of cats and dogs would start with a large number of images of cats and dogs. It would use "learning algorithms" to define a function that can do the recognition rather than asking a human to create a program by hand. This process of learning a function is often called "training" (and we will use this term in this chapter). The resulting function is commonly called a model. In other words, and at a very high level, ML lets us use datas to create functions that we don't know how to code.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 19:26:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781108872225.001

This definition can raise more questions than it answers, and it hides a rich complexity of different learning algorithms, working methods, and application areas. Rather than attempting to simplify ML artificially, we will draw out its different defining features using the chapters of this book as examples.

1.1.1 Types of ML

Part of the power of ML is that it can be applied to so many domains and problems. For example, Mimno and Thompson (Chapter 3) apply it to understanding humanities archives; Bevilacqua et al. (Chapter 4) to movement interaction for the arts; Vázquez et al. (Chapter 9) to understanding human social interaction; and Prabhu and Birhane (Chapter 13) study its application to image cropping.

These different applications are generally divided into a number of types of ML task:

- **Supervised learning** is a common type of ML, where the training set consists of labeled data. Using the earlier definition of ML, the training set consists of (input, output) pairs. One of the key activities that can make up part of a supervised learning process is the definition of *features*, which informally refers to what the learning algorithm needs to know about the training data, which in turn will be used by the model to compute its output. The most basic application of supervised learning is classification, in which a learned model assigns a label to some object or phenomenon (e.g., a document or a part of it). For example, Mimno and Thompson (Chapter 3) classify pages from arts periodicals as either Dada or non-Dada. In other variations of supervised learning, a model might output a real value instead of a discrete label (as in regression), or it might output a set of probabilities associated with different labels. For example, Bevilacqua et al. (Chapter 4) describe the use of regression to recognize movement features, and Prabhu and Birhane (Chapter 13) study how ML is used to measure the saliency of different parts of an image (and reveal that just because ML can measure a concept, that does not mean it is meaningful). Supervised learning is normally trained with large datasets that include a label indicating the correct output for each input. This process of labeling can often involve people as sources of truth, and hence can be labor-intensive as described by Yin and Law (Chapter 6).
- **Unsupervised learning**, on the other hand, does not include any explicit labels or "correct" output. Instead, the algorithm learns about the structure of the data without any explicit labeling of data, though, like supervised learning, it still requires a choice of learning architecture and hand-tuning of hyperparameters (the parameters of a model that must be set, not learned). A classic unsupervised learning application is clustering, where one is looking to organize data into groups of (semantically) similar items. Mimno and Thompson (Chapter 3) use unsupervised learning to gain insights into the diverse themes present in science-fiction writing.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 19:26:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781108872225.001

- **Reinforcement learning** is a different approach in which an AI agent learns how to act in an environment by receiving rewards or punishments for successful or unsuccessful behavior, much as we might train a pet. During reinforcement learning, the learning algorithm tries to arrive at the best behavior *policy* that specifies how the AI agent should act given a particular state in its environment. Krening (Chapter 8) describes the challenges of humans training agents using reinforcement learning.
- **Generative learning** does not simply learn to classify or group data items, but how to create completely new data items such as images or text, often using a combination of supervised, unsupervised, and reinforcement learning. Some of the biggest successes of the early 2020s have been in generative learning, including image generators such as Stable Diffusion and text generators such as GPT. Akten, in his interview with us (Chapter 2), discusses his use of generative ML for artistic practice.

As a process, ML can involve a number of steps, from gathering and preparing data, to training and testing a model, to using the output of that model. The following sections will describe each step of that process.

1.1.2 Datasets

Data is central to ML. In order to train an ML algorithm, that is, produce a model, one needs training data, often called a dataset. Different learning algorithms and learning tasks have different dataset size needs. The dataset size can have a direct influence on a model's performance, but so can the dataset's quality. The phrase "garbage in, garbage out" is not only an apt metaphor; it also is an apt reflection on the factors that affect the training of a model.

The importance of the quality of datasets means that creating datasets can be one of the most work-intensive parts of the ML process. Many of our chapters discuss the importance of data. For example, Mimno and Thompson (Chapter 3) describe the work of curating datasets in the humanities, and Law and Yin (Chapter 6) discuss the role of crowdsourcing in dataset creation.

1.1.3 Features

For data to be used in ML, they must be represented in a suitable numerical form. These representations are called "features," and they may be constructed and/or determined by people, or they may be learned by the ML algorithm. For certain types of problems where data is plentiful, large deep-learning models can learn effective features from the data rather than needing them to be specified by hand. However, other types of data can require more thought in the learning process; for example, Bevilacqua et al. (Chapter 4) discuss how data from complex movement sensors can be processed and represented as features.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 19:26:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781108872225.001

1.1.4 Model Training

Once a dataset has been prepared and potentially represented with suitable features, it is ready to be used to train an ML model. DNNs can be very powerful, but they have drawbacks of requiring large datasets and being very difficult to interpret, that is, a model's self-learned features might not be semantically sensical or they may be difficult to understand. Also, DNNs present many architectural and parametric choices that are far from trivial. Bevilacqua et al. (Chapter 4) and Vázquez et al. (Chapter 9) describe other training methods used, often in situations in which large datasets are not available. Liao and Varshney (Chapter 7) also discuss the problem of interpretability, including the use of more interpretable learning methods than deep learning. The process of training a model is itself a complex piece of work, which Krening (Chapter 8) discusses in the context of reinforcement learning from a human-centered perspective.

1.1.5 Testing

Once a model is trained it should be tested to confirm that it works as intended, that is, given a particular input, it produces the expected (or a reasonable) output. Testing normally leads to an iterative process of improvement and further testing. A traditional way of testing a model consists of dividing the training dataset into training, validation (to help refine a model's parameters), and testing (to evaluate a model on data that was not used to train it) subsets.

This testing strategy leans on quantitative metrics such as precision and recall, and does not consider human-centered factors such as those measuring the quality of the user experience using an ML model. Several authors, but particularly Stumpf and Smith-Renner (Chapter 5) enhance traditional test metrics with human-centered approaches.

1.1.6 Output

The outputs of a learned model can take many forms and play a significant part in the way people interact with or are impacted by ML. Examples include Mimno and Thompson (Chapter 3) who use the outputs directly as part of their humanities research, while Bevilacqua et al. (Chapter 4) map the outputs from performers' movements onto other media such as audio and graphics. Sometimes the outputs are used in a purely automated way but one that impacts people, ranging from the image-cropping methods discussed by Prabhu and Birhane (Chapter 13) to the forms of data injustice discussed by Aal et al. (Chapter 12).

1.2 What Does It Mean for Technology and Research to Be Human-Centered?

Like many other technologies, ML is often thought of and designed from a primarily technical point of view. This means its developers or practitioners

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 19:26:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781108872225.001

work on it either purely for technical interest or aiming to solely improve some technical efficiency. This type of practice, because it does not adequately consider the needs and well-being of people using or impacted by ML, can often lead to negative outcomes ranging from frustrating experiences to the introduction of significant societal harms, such as the accelerated spread of misinformation, hate speech, and manipulation of messages in social media during the Tigray conflict, as described by Hadgu in his interview (Chapter 14).

In other areas of technology, there has been a growing acceptance of the need to take account of, and improve, technology's impact on people. In assembling this book, we have considered approaches to ML research and/or development work as "human-centered" if they involve taking actions to better understand and improve technology's impact on people. Such an approach broadly often draws upon the practices, methodologies, and values of HCI and adjacent disciplines (including interaction design, UX research, and human-centered computing). Historically, work focused on human-centered technology research and development began with an aim of studying the ergonomics of work software: how human users can interact with software efficiently and effectively. However, over the years it has become obvious that human-centeredness expands beyond this starting point in many ways:

- **Human-centered thinking expands beyond efficiency:** We need to consider many different factors related to people, including their physical and mental well-being, as well as their relationships with others and other social factors.
- **Human-centered thinking expands beyond work scenarios:** HCI methods have been applied to many social and leisure scenarios such as games, social media, creative expression, and communication.
- **Human-centered thinking expands beyond users:** The use and existence of a technology, solution, or tool can impact people far beyond those directly using it. Human-centered methods should take account of all stakeholders and impacted communities. They should take account of the whole social context.

Human-centered methods alone do not prevent the development of highly unethical technology that will have negative impacts on people. For example, one can use human-centered methods to better understand people's strengths and weaknesses in order to engage in social manipulation or other dark patterns. Barraza, in his interview (Chapter 14) describes how such methods are used to create highly addictive social media and game experiences that can have very negative societal impacts. He calls for a "Humanity-Centered Computing," which is ethical at heart, working for the good of humanity, not simply for narrow profit-maximizing goals. In this book, it is our aim to explore human-centered approaches that align with Barraza's "humanity-centered" vision rather than focusing narrowly on considerations such as usability.

Human-centered work in technology includes a number of distinct activities, many of which are reflected in the chapters of this book:

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 19:26:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781108872225.001

1.2.1 Understanding People

The first step in human-centered work is to understand the stakeholders impacted by an existing or proposed technology. This would certainly include direct users of a solution, tool, or technology, but should also include wider community members. While stakeholders can be "studied" in a distanced way as research subjects, Neiss et al. (Chapter 12) favor a more participatory approach, particularly when working with marginalized communities. They describe case studies of working directly with stakeholders to understand their needs and to design software that benefits them.

1.2.2 Design

Once the needs and contexts of stakeholders are understood, it is possible to design technologies or solutions to address these needs. Human-Centered Design places humans at the center of this process. Designs stemming from this process are based on the needs and contexts identified in the first step, but in most cases, stakeholders continue to be involved, either as testers, or as co-designers in the kind of participatory processes described by Neiss et al. (Chapter 12). Forlizzi et al. (Chapter 15) talk about the design process in detail and how it can be applied in the context of building new ML-based systems.

1.2.3 Evaluation

Humans must also be at the heart of the evaluation of technologies. While traditional evaluation of ML systems focuses solely on performance metrics such as accuracy, precision or recall, these remain abstracted from the context of actual use. In contrast, human-centered evaluation focuses on how people use a technology, and more broadly how the technology may have broader consequences within a human context. This type of evaluation involves a range of methods of study with human participants ranging from controlled laboratory experiments to longitudinal ethnographies of use in real-life contexts. Stumpf and Smith-Renner (Chapter 5) give a detailed overview of human-centered evaluation methods in the context of their work on ML. Prabhu and Birhane (Chapter 13) take a more critical approach and focus on the technology they are evaluating rather than running human studies, but they do so with a detailed critique that exposes the impact of the technology in a human context.

1.3 What is Human-Centered ML?

Finally, we come to a central question for this book: what is human-centered ML? A simple definition is to say that it is the application of a human-centered mindset and human-centered methods to the practices of designing, using, and evaluating ML. This definition covers most of what we would want to call

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 19:26:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781108872225.001

human-centered ML, and it is also broad, encompassing many different approaches, methods, technologies and application areas. This is also a still-emerging, diverse and multifaceted field.

In this book we aim to provide a number of illustrations of what human-centered ML can mean in practice. Below, we summarize the different ways in which the authors of the chapters in this book have engaged in human-centered ML practices, and in cases engaged explicitly with the question of how human-centered ML should be defined.

1.3.1 Applying HCI Approaches to ML

At the center of human-centered ML practice is applying human-centered methods to ML, including understanding people, design, and evaluation. All of the authors in this book do this in some way, but there are some who focus particularly on these methods. Neiss et al. (Chapter 12) examine how participatory methods can be used to understand the needs of marginalized communities. This can help create better ML systems that serve everyone's needs. Working with marginalized communities can be difficult and making community members an integral part of the process can help build trust. Forlizzi et al. (Chapter 15) study how designers work with ML and the challenges they face. Applying ML to the design of everyday solutions is a relatively new challenge for most designers and many are not familiar with the technologies involved. Designing solutions that use ML requires new types of tools and working patterns. Stumpf and Smith-Renner (Chapter 5) explain how human-centered evaluation methods can be applied when working with ML. This includes using standard measures of user experience, such as effectiveness, efficiency and satisfaction, but also other measures that are more specific to ML such as intelligibility, predictability and trust. Over the years HCI researchers have developed a very wide range of methods for improving the human-centeredness of technology. These methods have been very effective in creating highly successful technology products. Applying them to the development and use of ML is a key part of making this technology and its outcomes more human-centered.

1.3.2 Understanding the Role of Humans in the ML Process

Many of our authors aim to understand the roles people can take within the process of creating ML models. These people could be highly skilled (and potentially highly paid) individuals who are essential to the design process, for example, the designers studied by Forlizzi et al. (Chapter 15). These people can also be the low-paid cloud workers studied by Law and Yin (Chapter 6). Knowing how these individuals understand the ML process, and how it is (or could be) designed around their needs, not only can impact the quality and effectiveness of ML products, but also contribute to a sustainable environment where the well-being and dignity of people involved in creating ML systems are

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 19:26:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781108872225.001

prioritized. Krening (Chapter 8) looks at the process of training ML models and how human norms and conventions can differ greatly from what is expected by the algorithms. Some of our authors also reflect as practitioners using ML from outside of the traditional technology disciplines. For example, Thompson and Mimno (Chapter 3) show how their work as humanities scholars elucidates the process of ML, while Bevilacqua et al. (Chapter 4) show how ML can be integrated into artistic practice.

1.3.3 ML for Understanding Human Behavior

As we underscored before, understanding people is central to human-centered methods, and understanding how people interact with ML is key to human-centered ML. It is interesting to observe that ML itself can be a tool for understanding people. For example, Vázquez et al. (Chapter 9) use ML to understand the complexity of human social behavior, while Thompson and Mimno (Chapter 3) use ML to understand features of artistic movements. The use of ML to understand people is not automatically human-centered; it could be done in a narrow, technically driven way, but we believe that to do this well requires a nuanced understanding of humanity that is a hallmark of human-centricity.

1.3.4 ML for Novel Human-Centered Interactions

ML can open up entirely novel ways of interacting with computers because it can be used to recognize far more complex and subtle human behaviors and actions than traditional user interface techniques. For example, Bevilacqua et al. (Chapter 4) show how ML can create interactive systems that respond to the complex movements and gestures involved in dance and musical performance, and Vázquez et al. (Chapter 9) discuss the creation of intelligent agents that interact based on the even more complex nuances of human social behavior. These new forms of interaction create new design challenges. Forlizzi et al. (Chapter 15) show how designers respond to the use of ML in interaction design.

1.3.5 ML Literacy

For someone to effectively use a tool or technology, they must have some level of understanding about it. This is a crucial point in the context of ML; that is, for people to interact with an ML system effectively and safely, there are particular things about it that they must understand. Acquiring this knowledge or literacy can be difficult, partially because ML is a relatively new technology to most people and it works in a very different way from existing digital technology (furthermore, of course, it is the subject of many misrepresentations in popular culture). Understanding ML is also challenging because it can be

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 19:26:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781108872225.001

opaque, acting as a black box whose behavior is very hard to interpret. There are two approaches to overcoming this problem: making ML easier to understand and educating people about ML. Liao and Varshney (Chapter 7) discuss the first, surveying the field of explainable AI, which aims to make ML and AI easier for people to interpret by providing explanations of its behavior. They argue for the importance of taking a human-centered approach to explainable AI, in contrast to many purely technical approaches that may not be helpful to people in practice.

Another approach is to educate people to better understand how ML works. For example, Neiss et al. (Chapter 12) favor working in a participatory way with marginalized communities to help them understand how to create ML systems that support their needs. Forlizzi et al. (Chapter 15) discuss how to support designers to better understand how ML can be used as a design material. Lao and Lee (Chapter 11) provide an overall framework for ML education that will help non-experts become effective users and creators of ML systems.

1.3.6 Humanity, Not Just Users

As we have already stressed, human-centricity is not simply about the needs of users of a piece of software but its impact on society, and humanity as a whole. These are issues that many of our authors grapple with in many ways. For example, Prabhu and Birhane (Chapter 13) show how a detailed critique of a particular algorithm (Saliency Cropping) can reveal often surprising biases that can have wide-ranging social impacts. Neiss et al. (Chapter 12), on the other hand, present ways of minimizing harm and maximizing benefits to marginalized communities.

1.3.7 Definitions of Human-Centered ML

During our conversations with different authors, we had an opportunity to learn what human-centered ML means. It is worthwhile to mention some of these perspectives as we bring this introduction to a close:

- Wu Qiong defines human-centered ML as making ML better meet human needs, including human–computer interaction, applications and design methods.
- For Adam Perer, human-centered ML means that a human is in the loop in all aspects of the ML process.
- Wendy MacKay describes human-centered ML as approaching the algorithm with a focus on the human, not the human-in-the-loop of the computer. It involves understanding that the same algorithm might work very differently for two different people, or two algorithms might essentially work the same for the needs of a particular user.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 19:26:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781108872225.001

- Asme Hadgu's work centers on the importance of embedding ML work within affected communities, such as his own community in Tigray, rather than simply working with individuals.
- Memo Akten's artistic work is based on ML but includes what he calls "meaningful human control." He says that the artistic result of ML should carry the artist's signature in some way.
- Rick Barraza uses the term humanity-centered ML as a practice and technology that champions human dignity, rather than using human-centered approaches to develop ultimately harmful products.
- Josh Lovejoy's answer is based on accountability. "What do the users of an eventual system need in order to take accountability?" rather than "What does the system need in order to take accountability?"
- Bill Buxton believes that human-centricity is vital to all aspects of development including the people making the engineering decisions as much as the end users. This requires interdisciplinary "renaissance teams" with different members having different specialties, but each having literacy in each other's domains.
- Phoenix Perry criticizes the term human-centered in the current Anthropocene era and instead proposes an ecological, post-human approach.

As we reflect on these perspectives, we can zoom out even further, and state that human-centered ML is not simply a set of tools and practices, but a philosophy of what technology is and what it is for. If we refer to ML as AI, this implies that it is autonomous and separate from humans, perhaps even in competition with humans. If, on the other hand, we talk about human-centered ML this implies that ML is there for humans to use; it is part of human society, and it should be developed for the benefit of humanity.

1.4 STRUCTURE OF THE BOOK

This book is made up of two types of chapters. Some of the chapters are traditional academic papers that describe and summarize the research areas that make up human-centered ML. Interspersed with these chapters are interviews which feature prominent practitioners and researchers giving a more informal and opinionated view of the subject.

Part I of this book looks at applications of human-centered ML to the Arts and Humanities. This launches the book with examples of human-centered ML to give readers an idea of what it looks like in practice, particularly in areas that are very far from the scientific and engineering origins of ML. Part I starts with a series of interviews with practitioners that take a creative approach and focus on individual work with ML: artist Memo Akten, designer Josh Lovejoy and HCI researcher Wendy MacKay. It then moves to two chapters that discuss applications. Mimno and Thompson describe their research using ML for

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 19:26:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781108872225.001

humanities research and Bevilacqua et al. discuss how ML can be used for the performing arts.

Part II focuses on the details of human-centered ML in practice. Stumpf and Smith-Renner survey methods for measuring user experience aspects of ML. Law and Yin take a human-centered approach to the common practice of crowdsourcing, and Forlizzi et al. engage with AI as a design material and share how this perspective can guide the development of next generation of AI-enabled products.

Part III looks in more detail at the role of humans in the ML process. Liao and Varshney discuss explainable AI, methods used to make the outputs of ML understandable, and stress the importance of taking a human-centered approach. Krening looks at the process of teaching machines: how people train ML algorithms. Finally, V´asquez et al. look in the opposite direction: how we can train machines to understand the complexity of human behavior in context. This part ends with a second round of interviews that take a wide-ranging look at ML in the broader context of human teams and groups: designer Wu Qiong, computer scientist Adam Perer, and HCI researcher and musician Bill Buxton.

Part IV opens out to the broader social context of ML, what impacts it has on our society and how we can use ML for a better society. Lao and Lee describe a Machine Learning Education Framework that aims to create more AI literate citizens. Aal et al. look at the impact of ML on marginalized people and how to make ML that benefits those at the margins. Prabhu and Birhane look in detail at one particular case: how the cropping algorithms used by social media platforms create biased and sexist thumbnail images.

We finish the book by zooming out even further with a final set of interviews that look at humanity as a whole. Asmelash Teka Hadgu leads a start-up developing machine translation for underserved languages; Phoenix Perry is a game designer committed to social justice; and Rick Barraza is a designer who urges us to think beyond humanity and into our whole ecosystem.

REFERENCES

Benjamin, R. (2019). *Race after Technology: Abolitionist Tools for the New Jim Code*. Polity Press.

Crawford, K. (2022). *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. Yale University Press.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. Commun. ACM 60, 6 (June 2017), 84–90. https://doi.org/10.1145/3065386

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444.

OpenAI, Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F. L., Almeida, D., Altenschmidt, J., Altman, S., Anadkat, S., Avila, R., Babuschkin, I., Balaji, S., Balcom, V., Baltescu, P., Bao, H., Bavarian, M., . . . Zoph, B. (2023). *GPT-4 Technical Report*. OpenAI.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 19:26:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781108872225.001

Russell, S. (2023). *Human Compatible*. Penguin.

Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Van Den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., Dieleman, S., Grewe, D., Nham, J., Kalchbrenner, N., Sutskever, I., Lillicrap, T., Leach, M., Kavukcuoglu, K., Graepel, T., & Hassabis, D. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature*, *529*(7587). https://doi.org/10.1038/nature16961.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 19:26:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781108872225.001

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 19:26:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781108872225.001