

Matthew Vollo

*Book Review: Atlas of AI by Kate Crawford*

At the core of *Atlas of AI* is the concept of power, and how AI is used as a tool to buy those in power in order to reinforce the structures that keep them at the top, at the expense of everyone else in society. In pursuing this notion, Crawford demystifies AI, stripping it bare and showing how AI is not something “alien” or disembodied from humanity, but rather is a product of social, political and economic forces. She states this succinctly in her intro, where she states that there is nothing *artificial* about artificial intelligence, in the sense of how integrated and dependent it is on human creation and human labor. The book was given the title of Atlas, and does its job as a chronicle of the topic of AI: from the creation of microchips and electronics in lithium mines to artificial intelligence algorithms being used in surveillance and warfare, all is shown in great, well researched detail. Though there were some sections, I felt, that could have made their points a bit more strongly (this is mainly a structuring problem), the book is extremely effective at what it aims to do, and that is to educate the reader about the role that AI plays in our interconnected modern world.

Crawford starts the book by showcasing the example of *Clever Hans*, a horse from the late nineteenth century that could apparently solve math problems, spell out words and display all sorts of other feats that, to an unknowing observer, could be described as “magical”. However, it was revealed that *Clever Hans* wasn’t actually as clever as his namesake would have one believe, and that all of his feats relied on unconscious directioning from the questioners. Hence, the horse was producing answers the owners, and the questioners wanted to see. This

example starts off the book perfectly, as *Clever Hans* serves as a microcosm of the topic of modern artificial intelligence. In this regard, this is when the demystification of AI begins in this book. The introduction is spent dispelling some common myths and false notions about AI that are pervasive in our society, with some even being propagated by AI companies themselves. After offering this microcosm, Crawford then goes to outline the rest of her text, chapter by chapter (though the final chapter entitled *Coda* is not included in this, which leads one to assume that it was added onto more recently). In this outline, Crawford outlines what is essentially the main point of each chapter, with each point contributing to or being a part of the larger message of the book as a whole. In the introduction, Crawford makes it clear to dispel the notion of AI as a spectral or “unfathomable” force. She grounds the notion of AI by focusing on the problems it poses on society at the current moment, and she does not waste time even entertaining the notion of the sci-fi-esque idea of AI usurping the position of humans in a destructive sense. In this opening, she succeeds in laying the groundwork for her text by clearly defining what she means by “Artificial Intelligence”, and she outlines how, chapter by chapter, she goes about explaining and dissecting information in order to present a clear message to the reader.

The first chapter of the text is entitled *Earth* and is focused on the environmental toll that AI produces. It is the goal of this part to hammer home the notion of AI being something constructed by humans, as opposed to being something abstract or “magical”. She does this by showing how AI is dependent on technology, and how that technology is dependent on manual labor for its extraction and production. The construction of AI is done at the cost of a myriad of invisible hands putting together what they can from finite resources. In stressing the notion of finite resources, Crawford connects the construction of AI to climate change by focusing heavily on how much energy consumption is needed to produce and maintain the technology that

Artificial Intelligence runs on. In doing so, Crawford brings up the notion of colonization, referring to how developed countries take resources from developing countries, and this theme of colonization, of an imposing imperial extracting force, is prevalent throughout the text. It is wise, in this regard, that the theme of colonialism is brought up in the chapter that deals with resource extraction, and Crawford stresses that “The greatest benefits of extraction have been captured by the few.” (Pg 28). To give the reader a better comprehension about the size and scale of resource extraction, Crawford utilizes historical and visual features that command the reader's attention. For example, on page 27, in order to emphasize the relation between cities and the labor that sustains them, she brings up the example of San Francisco, by saying “we should think of the skyscrapers of San Francisco as inverted minescapes”. This image is striking, and it serves a crucial function in expecting the reader to directly envision the impact that mining has had on sociological development. The chapter is anchored around Crawford's visit to a mine called Silver Peak, a mine that was once thriving during the 1800s, which faded after the mining boom ended, only to rise again in modernity due to the lithium deposits next to it. The chapter ends with the image of Silver Peak, an empty shell, a discarded piece of land that has been extracted and forgotten about. Striking images like this show the result of resource extraction, and such destitute, as is the main point of this chapter (and a central tenet of the book), will go unnoticed by the general population.

Another core feature of AI that the general public is unaware of is the labor that goes into, not only producing, but into manufacturing the illusion of seamless AI. This illusion of AI is the one propagated to the public, where a disembodied intelligence seamlessly solves problems. In the second chapter, entitled *Labor*, Crawford seeks to further deconstruct this point, by highlighting all the hidden labor that goes into the production and maintenance of Ai systems

and algorithms. Crawford here leans heavily on examples of underpaid labor that fuels a deceptive image of AI. This ranges from bringing up Amazon's Mechanical Turk to bringing up the example of companies such as x.ai, which claim to showcase "magical" (page 65) instances of AI, which were actually faked by human employees. One critique I had about this chapter was that I felt the line between genuine AI algorithms, those that work but need maintenance by human coders, and those "algorithms" which were only propped up by human coders wasn't made clear enough, at least when the concept was first introduced. Despite this, Crawford makes an effort to include examples that deconstruct the notion of AI acting independently and seamlessly, and making sure to say that there is always a human involved in the process. An interesting feature of this chapter is the inclusion of tech companies trying to manipulate time, and specifically the concept of "True Time". This inclusion of the concept of the manipulation of time also blends into themes that will be explored later in the book, such as the tech companies controlling or desiring to have the "truth" (or rather, controlling the narrative) and the use of time in order to subjugate workers. Indeed, the concept of further controlling workers, and the desire to increase control in general is a central tenet of Crawford's work. This is bolstered by a mini-chronological summary Crawford provides of how labor has changed due to industrialization, first starting with the Chicago meat-packing industry of the 1870s (page 73) whose practices and philosophies (and Crawford delves extensively into the philosophies of control, and specifically the control of employees to maximize production) evolved into the systems that produce technology today. Here we have another instance of Crawford using scale and scope in order to illustrate a greater point. By putting a chronological timeline here, as readers we can follow Crawford as she makes her points in a logical manner.

The third chapter is entitled *Data*, and delves into the practice of data harvesting. Rather than simply mention that data extraction is indeed happening, Crawford delves into the very philosophies that underlie the practice of data collection from private companies (she also delves into the concept of the state extracting data later on in chapter six). Now, the two previous chapters serve as background information into the topic, while chapter three seemed to highlight the beginning of a crucial part of the book. From here, the themes of chapters three through six seem to be connected in a way that chapters one and two were not. Chapters one and two built on top of each other well, but chapters three through six seem to directly and logically follow each other.

Chapter three, as stated before, focuses on the unethical ways that tech companies extract data, with particular attention (backed up by photographic evidence) to the notion that prison mugshots were being used to attune facial and affect recognition systems. This takes up a large part of chapter three, where Crawford states that the images were taken and put into a database without the person in the photo's consent, and that despite the fact that the mugshot can be the result of a myriad of crimes (either ranging from a false conviction, a minor sentence or something major), that this mugshot will forever remain on the database as an example of training data. Essentially, data harvested unethically is being used without the consent of those who the data was harvested from to fine tune algorithms (which the next chapter brings up the fact that some of these algorithms are flawed on a conceptual level). Crawford tracks the changing opinions and views on data, where data was collected from willing participants (like college students) to unwilling, unconsenting individuals. She also clearly states here (which will also be developed in further chapters) of tech companies desire to map out and classify things. This effort to map out the "entire world of objects" (page 107) is wishful thinking at best and

unintentionally harmful at worst. Crawford discusses how the imperial notion of classification is used here, the desire to oversimplify an overly complex world to make it more machine readable (and thus people here, with all their complexities are simplified). Crawford brings up a major point here with regards to the datasets themselves, and how data was collected, makes the algorithm inherently flawed. Crawford utilizes here the examples of WordNet and ImageNet to highlight the false objectivity of algorithms (since they are made with human efforts, and thus human biases) and shows how the algorithms are flawed by design to the extent that minorities (LGBT, african americans, undocumented people, etc) are disproportionately affected by changes like this. This, combined with the modern concept of data being anything that can be extracted, paints a dark picture of the worldwide tech industries.

This theme is further added onto within the concept of *classification* (chapter four). Here, Crawford draws back into the use of *imperial* and *imperialism* to describe the act of classification, as classification is done, in her words, from a point of power in order to control and subdivide people. She starts the chapter by talking about Samuel Morton's collection of skulls, a prime example of "objective classification" that was obscured by a clear bias. In this chapter, Crawford delves into the different meanings and definitions of bias, and highlights (page 135) how there has been a failure on AI to address this. She has shown that there have been superficial attempts by tech companies that have more so addressed the processes of the algorithms as opposed to the very philosophy and the construction of those algorithms. Again here, we see the concept of the failure of tech companies address the problems they have created. This inescapability of bias is encapsulated in the quote: "every dataset used to train machine learning systems... contains a worldview" (page 135).

Chapter five adds onto the topic of affect recognition, again describing how the main problem lies not as a glitch within an algorithm, but behind the very science itself. Crawford traces the field of affect recognition back to its controversial roots, dissecting the flawed reasoning from the founder of affect detection as a means to reveal emotional inner states. She states that this technology is not scientifically sound, and algorithms based on affect recognition that claim to spot the face of a prospective target (for example, the face of a criminal or the face of a terrorist) are fundamentally flawed. I thought the reasoning in this chapter was quite sound, though I did have some issues with the organization, as I felt that the explanation of the unsoundness of affect recognition (with regards to telling details about internal states) came rather late into the chapter (152). Though I felt that the chapter as a whole was written quite well, it just takes a while for the assertion that “there is no reliable evidence that you can accurately predict someone’s emotional state from their face” to be backed up (in text, not in footnotes), despite being such a major assertion, even though it is true. This notion of affect recognition also blends soundly into chapter six, where the issues of the state utilizing these algorithms comes into play. This serves as the peak of the thesis for the book, with the state being the primary means for control. Crawford spends a portion of this chapter dispelling the jingoist notion of an “AI war” by mentioning that the means of control by method of AI are already at a global scale, with ensuring cooperation between prospective enemies and different states. Crawford also dissects here the nationalist incentive that can be uttered by tech companies as a means to justify themselves getting into wars and surveillance. A main point of the text is that someone who thinks they are doing good can actually be doing a lot of harm, and even more harm can be done when that person disregards the consequences of their own actions, as tech companies have done.

The book ends with a summary of the points that Crawford had presented in the preceding chapters. Much like the rest of the book, it is concise and well organized, and summarizes the contents well. The conclusion has the subtitle *power*, which showcases how this was a central theme of the book as a whole. As for a means for combatting or fighting this power however, Crawford does not offer much room. To me, it seems that this was more of a piece that serves more to elucidate than to call to action. Crawford states that “we must focus less on therics and more on power” as AI is fundamentally designed to reproduce forms of power (page 224). Therefore the strategy seems to be, not to try and change AI from the inside, but to change it from without. But not by simple policies, but rather by a rejection of the dogma of inevitability (or a false inevitability) (page 226), and the refusal of the current state of AI continuing, in order to create the foundations of a different society. For as detailed as the rest of the book was, I was hoping for a more detailed conclusion, or at least one that ended on more specific solutions as opposed to a simple ideological or narrative rejection. The final chapter (which reads like it was added on recently and is not as interwoven as the other chapters) ends on a somber note, where Crawford is pursued by black pickup trucks after leaving Blue Origin’s facilities. This, though foreboding, indicates a sense of urgency on the part of the reader. That the apparatus of control and power that AI has helped to prop up needs to be put in check and rejected.