11 The Proletarianization of Data Science

James Steinhoff

Pick up any business-oriented publication today, and you are likely to find at least one article vaunting the economic potential of data science and artificial intelligence (AI). For instance, analysts from consulting firm PwC expect AI to increase global gross domestic product up to $15.7 trillion by 2030 and reckon it as the "biggest commercial opportunity in today's fast changing economy" (Rao and Verweij 2017). Big technology firms such as Google and Amazon, as well as a host of smaller companies, are aiming to seize this opportunity by producing AI commodities for consumers (e.g., smart speakers) and for businesses (e.g., face and object recognition software). The economic opportunity presented by AI is, however, contingent on the labor of data science workers who possess the relatively rare skills to develop said technology.

Data science workers are celebrated for their desirable working conditions and posh benefits. Data scientist has been touted as the "sexiest job of the 21st century" (Davenport and Patil 2012), and it was rated the best job in the world, based on wage, satisfaction, and number of openings, by employment site Glassdoor (Columbus 2019). Data science work is also cited as the planetary future of employment, projected to both revitalize the deindustrializing Global North and boost the Global South into the digital era (Microsoft 2018). As the bright future of the digital economy, it is contrasted both with submission to the algorithmic management of Uber-like platforms and tedious "ghost work," or digital piecework on platforms such as Amazon Mechanical Turk (Gray and Suri 2019).

This chapter shows that the data science labor force, while globally distributed, is predominantly tied to powerful firms concentrated in specific locales. In particular, I argue that the planetary data science labor force is increasingly created *by* and *for* powerful technology capital in the United States of America. The increasing efforts of large technology firms to produce their own bespoke labor force has implications for digital labor in general.

From the perspective of capital, data science labor power is a scarce commodity to be competed for. Around the world, efforts are thus being made to proletarianize data science labor power, to increase its supply and decrease its value, while capturing a competitive share of it. Even the sexiest kinds of work are, for capital, wage expenditures to be minimized. Wider distribution of the skills to perform currently rewarding and well-remunerated digital labor is often positioned as a way to close the economic gap between the Global North and South, but the distribution of such skills is accompanied by their simplification and consequent devaluation. While the proletarianization of data science labor power may make more data science jobs available outside the Global North, it will do so only insofar as it reduces labor costs for big technology firms. Rather than the elevation of less privileged laborers to the Silicon Valley stratum, the proletarianization of data science labor power suggests the coming degradation of a privileged type of labor to the status of precarious and poorly remunerated ghost work.

This chapter begins by introducing the concept of labor power. I then outline data science work in the context of the AI industry, which is concentrated foremost in the United States and China. Next, I discuss how governments around the world are attempting to increase local supplies of data science labor power by adopting national AI strategies. Then, I show that US tech firms, unsatisfied with national efforts, now operate their own private education programs in an effort to generate suitable stocks of labor power distributed across national boundaries. I also discuss efforts to automate data science work and show how firms are motivated to pursue such forms of automation by the shortage of data science labor power. I conclude by considering how the proletarianization of data science labor power could combine with the mediation of digital platforms to transform data science work into something resembling ghost work.

The Special Commodity of Labor Power

Capitalism is defined by the production of commodities: products and services intended for profitable sale. However, the production of commodities requires the prior purchase of other commodities, including materials, tools, and the ability of someone to work. Karl Marx ([1867] 1990, 270) referred to this ability to work as labor power and emphasized that it is a "special commodity." It is what is exchanged for when someone is paid for work. This exchange is unequal because labor power is unique among commodities insofar as it can produce more value than is paid for it. But Marx's theory of value is not of interest here. The relevant point is that capitalism depends on there being people who need to sell their labor power. Capitalism is incompatible with a society in which

large portions of the population can meet their own survival needs autonomously. Capital thus has an impetus to make people dependent on earning a wage.

Marxists call the process of rendering people dependent on a wage proletarianization. Marx describes it as "when great masses of men are suddenly and forcibly torn from their means of subsistence, and hurled onto the labour market as free, unprotected and rightless proletarians" ([1867] 1990, 876). Examples of this include the enclosure of common land beginning in the thirteenth century in England and the appropriation of peasant lands in China since the 1990s (Zhang 2014). In both cases, stripped of the possibility of self-subsistence, great new chunks of the population were forced to sell their labor power for a wage.

Developed capitalist economies do not simply leave the availability of the special commodity labor power to chance: "Capital itself regulates [the] production of labor power, the production of the mass of men it intends to exploit in accordance with its own needs. Hence capital not only produces capital, it produces a growing mass of men, the material through which alone it can function as additional capital. . . . [C]apital produces on a steadily increasing scale the productive wage-labourers it requires" (Marx [1867] 1990, 1061). Proletarianization thus occurs wherever new markets emerge (McNally 1993, 31).

Capital can also create new proletarians by "extending its rule to sections of the population not previously subject to itself, such as women or children" (Marx [1867] 1990, 1061). Here capital does not create new proletarians by dispossessing those with access to the means of subsistence. Instead, women and children, who depend on the wage of an already proletarianized man, are compelled to sell their labor power as well. This hints at the diversity of ways in which capital can create new supplies of labor power. This chapter focuses on a scenario in which capital requires labor power possessing a rare, technical, and highly profitable skill set. These are data science workers. Their proletarianization proceeds by different means.

Data Science Work

Roles in the nascent field of data science have yet to settle into rigid categories. So rather than try to draw hard lines between data scientists, data engineers, and machine learning scientists, for example, I follow Muller et al. (2019, 126) and employ the general term *data science worker*.[1]

Data science work occupies a position near the apex of hype around the so-called "Fourth Industrial Revolution" (Schwab 2016). Data science draws on statistics, computer science, and mathematics to perform operations of "data collection, data engineering,

data analytics and data architecture" (Getoor et al. 2016). Data science is usually, but not necessarily, associated with "big" data, defined in terms of high velocity, variety, and volume (Zikopoulos and Eaton 2011). One notable component of data science is machine learning, a subfield of AI based on statistical pattern recognition in large data-sets. Machine learning involves "extract[ing] patterns from data" (Kaplan 2016, 27) and automatically constructing an algorithm, called a model, from those patterns that can be used to analyze or make predictions on new data (Alpaydin 2014, 2–3). The automatic production of a model from data is the source of much excitement for machine learning and data science.

Data science work is appearing in all sectors of the economy as businesses strive to incorporate data-driven analysis and decision making into their operations. However, this chapter focuses specifically on the data science workers involved in the production of machine learning systems. As such, its context is the AI industry, a subset of the larger technology industry.

In 2017, both Google and Microsoft declared themselves "AI first," and the rest of the tech industry followed (Agrawal, Gans, and Goldfarb 2018). Today's AI industry is global, but it is concentrated in the Global North and China. One study of governmental AI "readiness" concludes that "countries in the Global North are better placed to take advantage of these gains [produced by AI] than those in the Global South" (Miller and Stirling 2019, 5). However, China is second only to the US in AI stature, and all of the biggest AI companies are from either the United States (Google, Apple, Microsoft, Amazon, Facebook, Intel, and IBM) or China (Baidu, Alibaba, and Tencent). Smaller AI companies are also concentrated in these two countries. The *China AI Development Report 2018* shows that the US leads the AI start-up tally at 2,028 companies, China takes second place with 1,011, and the United Kingdom and Canada come in third and fourth with 392 and 285 companies, respectively (CISTP 2018, 46).

One study, based on LinkedIn data, reports that in 2018 there were 36,524 "self-reported AI specialists" in the world, a 66 percent increase from the previous year (Gagne 2019). According to this data, the majority of AI specialists work in the United States (43 percent), the United Kingdom (9.8 percent), Canada (4.1 percent), France (3.9 percent), and Germany (3.7 percent). However, as the authors point out, LinkedIn is not used uniformly around the world (44 percent of Americans use it, compared to only 3 percent of Chinese), so these figures leave much to be desired. Another analysis (data source unspecified) is provided by the *China AI Development Report, which* puts the global AI talent pool at the end of 2017 at 204,575, with the majority concentrated in the United States (13.9 percent), China (8.9 percent), India (8.5 percent), Germany (4.6 percent), and the United Kingdom (3.9 percent) (CISTP 2018, 34).

The dominance of the United States and China is clear, but the picture becomes more nuanced when one looks at where data science workers are educated and where they end up finding employment. One analysis of "top-tier" AI researchers reports that while 29 percent received their undergraduate degree in China, 20 percent in the United States, 18 percent in Europe, 8 percent in India, 5 percent in Canada, and the rest in the United Kingdom, Iran, Israel and other countries, 59 percent end up working in the United States, 11 percent in China, 10 percent in Europe, 6 percent in Canada, and 4 percent in the United Kingdom (Banerjee and Sheehan 2020). While China is the largest producer of data science workers, the United States ultimately retains most of this labor power.

The lead held by the United States and China in AI is such that several analyses describe them as a duopoly of nearly uncontestable AI "superpowers" (see, for example, Lee 2018). Both countries benefit from large populations of users essential for generating the data necessary for training machine learning models, regulations amenable to business, and wide availability of venture capital. Some analyses argue that if other countries fail to invest aggressively in AI soon, they will fall irrevocably behind as exponential productivity gains, powered by AI, will boost the economies of China and the United States beyond all competition (Cummings et al. 2018, vi). Lee (2018, 169) speculates that not only will AI research remain out of reach for less developed countries, but they will also "lose the one competitive edge that their predecessors used to kick-start development: low-wage factory labor" because the AI superpowers will run automated factories at lower costs. According to such accounts, the current lead on AI research will evolve into a new global economic divide in which whole countries will be rendered surplus to the market. However, there are more optimistic appraisals of the current situation. Analysts from consulting firm PwC suggest that since AI is at a "very early stage of development overall . . . there are therefore opportunities for emerging markets to leapfrog more developed counterparts" (Rao and Verweij 2017, 3). Below, I will discuss how many governments are attempting to do just that.

By all accounts, the AI industry is growing rapidly. The global revenue of the AI industry was around $5 billion in 2015, and by 2019 it had increased to somewhere between $15 billion and $37 billion (Statista 2020). From 2010 to 2018, investment in AI start-ups increased from $1.3 billion to over $40.4 billion (Perrault et al. 2019, 6). The burgeoning number of AI companies creates a large demand for data science labor power, which continues to outpace supply. According to the AI Index (Shoham et al. 2017), AI jobs in the United States grew by 4.5 times between 2013 and 2017. Fresh graduates find themselves in a job seeker's market where companies compete to offer lucrative employment offers (Saphir 2018). A survey by Stack Overflow (2018) puts the global median salary for a data scientist at $60,000, and at $102,000 in the United

States. Salaries can reach into the millions for acclaimed experts (Metz 2018). In general, capital rewards data science workers much better than academia does, resulting in a "brain drain" to industry (Kunze 2019; Perrault et al. 2019, 6).

Despite the enormous demand for data science work, this sector is marked by a "diversity crisis" (Snow 2018; see also West, Whittaker, and Crawford 2019). Only 22 percent of AI professionals in the world are women (Duke 2018), and sexism is rampant in the industry (Vassallo et al. n.d.). Global data about the racial identities of data science workers has yet to be produced, but the demographics of the workforces of the largest US tech companies give an idea of the broader picture. In 2018, at Google, Facebook, Microsoft, and Apple, white people made up around 40–51 percent of the workforce and Asians between 35 and 52 percent, while Latinx, Black, and Native people represented only single digits (Harrison 2019).

Bucking a history of apoliticism (Hyde 2003), tech workers have recently begun mobilizing around issues of gender and race, as well as around the militarization of AI. In 2018, over 20,000 Google employees walked out after the company gave a $90-million severance package to a senior employee after he was fired for sexual assault (Canon 2018). The same year, Google employees forced their employer to drop a contract to develop drone vision technology for the Pentagon (Harwell 2018).

In sum, while their situations are not identical the world over, data science workers occupy a relatively privileged position in the digital economy. One might therefore reasonably describe data science workers, especially the white and Asian men working for large US firms, as a "labor aristocracy" or a group of workers who enjoy a relatively privileged position of power vis-à-vis capital (Fuchs 2014, 229–230). However, processes are underway that aim to undermine such aristocratic privilege, to proletarianize data science labor power by both reducing its scarcity and devaluing its content. The following sections show how this is occurring in three ways: through the deployment of national AI strategies aiming to create and retain skilled data science workers, through the creation of proprietary data science education programs by large tech companies, and through the automation of data science work.

National AI Strategies

Faced with the prospect of USA/China AI hegemony, nearly 20 countries have released national AI strategies since 2017 (Dutton, Barron, and Boskovic 2018). A central component of the global race for AI is data science labor power. While national AI strategies vary in their details, all involve increasing the national supply of data science labor power, along with other objectives such as boosting academic research output

and speeding up the commercialization of AI. To achieve these goals, governments are establishing AI institutes in partnership with industry and academia. These institutes provide lucrative positions for experts in AI and data science, and some provide education and training to increase the supply of data science labor power. To take two examples, in 2017 the Canadian government established three AI institutes across the country (Finance Canada 2017), with founding sponsorship from Google, and in 2019 the Brazilian government announced the creation of eight AI research labs and an institute in partnership with IBM (Henriques 2020).

National AI plans also frequently include the allocation of funds to universities to increase their capacities to provide data science education as well as new regulations to streamline immigration processes and the dispensation of work visas. In connection with the 2017 Pan-Canadian Artificial Intelligence Strategy, Canada has adopted high-speed, uncapped pathways to citizenship for foreign students and workers with data science skills (Huang and Arnold 2020).

The proliferation of national AI strategies is significant because it is the "first time that governments around the world have almost simultaneously released national plans to develop the same technology" (Dutton, Barron, and Boskovic 2018, 4). This wave of investment is also distinguished from previous ones in that "AI is seen as strategic technology by many governments" (Walch 2020). This does not mean that AI is being sought solely for military capacities, although this is certainly of interest to some governments.[2] More broadly, the strategic interest in AI is economic.

Competing national AI strategies may aim to draw data science labor power away from US capital, but the US government also has its own plans.

The Trump administration sought to make it harder for highly skilled labor to enter the country. On June 22, 2020, President Trump released an executive order temporarily suspending the H-1B work visa widely used by high-tech workers coming from China, India, and Iran and by foreign students seeking to remain in the United States after completing their studies. The purported motive was to keep jobs for American citizens. The ban was lifted in April 2021, but at the time of writing in 2021, the Biden administration aims to replace the previous lottery system for the H-1B with a system giving preference to workers with the highest previous salaries (Anderson 2021). This will make it harder for students to stay in the country and presumably is intended to increase the supply of experienced, rather than novice, labor power.

However, most analyses argue that the United States is so far from domestically meeting demand for data science labor power that making it harder for foreign students to stay in the country can only undermine its industry (Zwetsloot et al. 2019; Hao 2020).

Proprietary Data Science Education Programs

Amid international AI strategizing, US technology capital is taking the problem of the shortage of data science labor power into its own hands. Big US tech firms are now establishing their own education programs to generate the skilled labor power that they need, without regard for national boundaries or the strictures of post-secondary education.

Amazon has developed AWS [Amazon Web Services] Educate, a free education service that provides "Cloud Career Pathways" in specializations including machine learning, data science, and cloud computing.[3] The service is available for people affiliated with educational institutions partnered with Amazon around the world. Salesforce similarly maintains Trailhead, a globally available gamified education platform for machine learning, coding, and business skills.[4] In mid-2020, Google announced the debut of "Google Career Certificates," for digital jobs including data analyst, UX designer, and IT [information technology] support specialist. These certificates cost a mere $39 per month and can be completed in less than six months of part-time study (Coursera n.d.). Positioning these certificates as a response to the COVID-19 pandemic, Google asserts that "we need new, accessible job-training solutions . . . to help America recover and rebuild" (Walker 2020). In their beneficence, Google also offers need-based scholarships and provides apprenticeship opportunities. The most striking aspect of these certificates is that Google will reportedly treat them as equivalent to a bachelor's degree (Bariso 2020). While currently there is no data science certificate available from Google, already there are a wealth of less prestigious options available on the Internet in the form of data science "boot camps" that typically last between two and six months.

The history of the AI industry contains a precedent for this situation. In the early 1980s, at the height of enthusiasm for expert systems, the first commercialized application of AI, Edward Feigenbaum, perhaps the first AI entrepreneur, developed an applied masters of science in artificial intelligence at Stanford University. The explicit goal of the degree was to "train students in enough AI to get them out pursuing practical applications, but not enough to make them career academics and researchers. It was one more attempt to get this new field out of the ivory tower and into the marketplace" (Roland and Shiman 2002, 196). Today, the AI industry is making a renewed push for this, which will take on a global scope.

Now that Internet access is more widely available across the Global South, US technology firms desire to outsource not only ghost work but also data science work there. Google, Facebook, and Amazon sponsored a Data Science Africa 2018 event in Abuja, Nigeria, where "building world-class capacity" was a central topic (Data Science Nigeria 2019, 25). A recent report on AI in Africa by policy-influencing firm Access Partnership,

in conjunction with Microsoft, sums up the perspective of industry: "The private sector has a critical role to play in elevating the skills of African citizens" (Access Partnership 2019, 26). Such efforts frequently draw on the rhetoric of democratization.

Democratization and Automation of Data Science

Since 2017, big US technology firms have announced programs for the so-called democratization of AI and data science. Behind its political trappings and promises to allow "every person" to produce AI (Microsoft News Center 2016), the democratization of AI refers to the diffusion of basic data science skills combined with the proliferation of automated data science tools, such that data science functions can be performed by nonspecialists or "citizen data scientists" (O'Connell 2018).

Data science work serves the interests of capital in several ways. I will discuss two. By increasing the speed of production and volume of output, automation increases the productivity of labor power. But automation also benefits capital by reducing the value of labor power. Automation allows the same products to be produced with less skill as machines take over certain tasks. In this respect, automation represents a culmination of the division of labor, in which complex labor processes are broken up into their constituent pieces. When a labor process can be "separated into elements some of which are simpler than others and each of which is simpler than the whole . . . the labor power capable of performing the process may be purchased more cheaply as dissociated elements than as a capacity integrated in a single worker" (Braverman 1998, 81). With automation, some pieces can be wholly given over to machines.

Perspectives on automation have changed in recent years. While historically it was assumed that manual labor was subject to automation while cognitive labor was less so, the sophistication of computing technology has led to a new distinction between automatable routine labor, whether manual or cognitive, and nonroutine labor, which is more difficult to automate (Autor, Levy, and Murnane 2003). Although celebratory discourse around data science work, as well as the scarcity of its requisite labor power, might lead one to believe that it is unautomatable, in fact the automation of data science work has been underway since around 2015. Efforts to automate programming, which contains large amounts of routine tasks, go back to the earliest days of computing (Chun 2011) and continue today, with boosted incentives from the shortage of skilled data science workers (Campbell 2020) and the COVID-19 pandemic (Salisbury 2020).

Data science work is a broad field, but if we consider machine learning in particular, we can schematize the labor process for its production as three stages (Dong 2017). The first stage, data processing, involves collecting, cleaning, labeling, and otherwise

preparing data. In the second stage, model building, data is input into a learning algorithm, which produces a machine learning model (itself an algorithm) based on patterns in the data. The third stage is deployment, in which the model is integrated into a business environment. Typically, this is carried out by embedding it into a website or application through a cloud-based application programming interface (API).

Data science work is a new field. Since there is not yet a mature ecosystem, data science workers "generally manage their workflows in ad-hoc ways" (Toews 2020), and overall, there "isn't a codified set of strategies" for machine learning (Theuwissen 2015, 13). This leads some to describe it as an art (McClure 2018). The incipience of data science has not, however, prevented its automation. Parts of all three stages of the machine learning labor process are being automated with a technology called automated machine learning (AutoML).

AutoML tools being developed by tech giants include Google Cloud AutoML, Microsoft Azure Automated ML, and Amazon SageMaker. In general, AutoML works by recursively applying the capacity of machine learning for the extraction of patterns from large datasets to the production of machine learning models. For instance, deciding on the precise architecture for a neural network is a time-consuming task, which requires trial-and-error based adaptation to each particular case. AutoML can be used to automatically generate and compare thousands of candidate architectures and can produce outputs that exceed the performance of handcrafted architectures (Zoph et al. 2017). One review of AutoML techniques reports that a team of three data scientists worked for weeks to produce a machine learning model for a predictive maintenance application, but only months later, an AutoML tool was used to automatically produce, from the same dataset, a better performing model (Tuggener et al. 2019, 2). Another survey of AutoML techniques concludes that currently available options enable one to build "reasonably well performing ML pipelines without knowledge about ML or statistics" and enable skilled data scientists to "profit from the automation of tedious manual tasks," even if approaches that aim to automate not just one task but the whole process are "still very basic and are not able to beat human experts yet" (Zöller and Huber 2021, 448).

Regardless, AutoML presents the prospect of devaluing expensive data science labor power by making the ability to produce machine learning models available to novices. While prediction of the future is not the intent of this chapter, it seems safe to agree that if AutoML is "truly able to automate the data science workflow, [it] may be instrumental in bridging the gap between the high demand and low supply of data scientists" (Wang et al. 2019, 211).

Conclusion

A global competition for scarce data science labor power is underway. While many countries are deploying national plans to boost AI and data science capacities, they must face off with big technology firms in the United States, which have developed proprietary data science education programs and tools for the automation of data science work. I have argued that these phenomena can be interpreted as parts of a process of proletarianization whereby capital seeks to generate an abundant and cheap supply of the scarce but highly profitable commodity of data science labor power. Education programs will increase overall supply, while AutoML will lower skill requirements for some data science tasks. Both will contribute to the devaluation of data science labor power.

AutoML may also, by fragmenting the data science labor process into automatable and nonautomatable components, render data science work increasingly amenable to outsourcing via digital platforms such as Amazon Mechanical Turk. Today's data science jobs could be broken up into piecework tasks performed by workers wherever ghost work wages constitute sufficient incentive. This will likely provide jobs for the Global South but not only there. Substantial ghost work contributions come from India but also from the United States and Europe (Berg et al. 2018). Regardless, the aristocratic status of data science workers at large US tech firms will be diminished by the emergence of a globally distributed pool of devalued data science labor power.

In the wake of the 2020 COVID-19 pandemic and concurrent recession, the business incentive to economize on labor costs will be exacerbated (Blit 2020). High-skill jobs are not immune from transformation into something like ghost work, and techniques of digital proletarianization, outlined here in the context of data science work, are likely to play an important role in the future of work in the planetary market.

Notes

1. *Data science worker*, as I use the term here, does *not* refer to ghost workers, who do perform operations essential to data science, such as data labeling.

2. As Haner and Garcia (2019, 331) point out, autonomous weapons research is "advancing rapidly and without sufficient public debate or accountability."

3. See https://aws.amazon.com/education/awseducate/.

4. See https://trailhead.salesforce.com/.

References

Access Partnership. 2019. *Artificial Intelligence for Africa: An Opportunity for Growth, Development, and Democratisation*. Nairobi: Microsoft Policy Innovation Centre, Strathmore University. https://pic.strathmore.edu/wp-content/uploads/2019/03/PIC_AI_for_Africa_Whitepaper.pdf.

Agrawal, Ajay, Joshua Gans, and Avi Goldfarb. 2018. "Companies Are Suddenly Declaring Themselves 'AI First.' Why It's a Problem for Their Current Customers." LinkedIn *Weekend Essay*, April 27. https://www.linkedin.com/pulse/companies-suddenly-declaring-themselves-ai-first-why-its-joshua-gans/.

Alpaydin, Ethem. 2014. *Introduction to Machine Learning*. Cambridge, MA: MIT Press.

Anderson, Stuart. 2021. "Biden Administration Defends Trump H-1B Visa Rule." *Forbes*, July 27. https://www.forbes.com/sites/stuartanderson/2021/07/27/biden-administration-defends-trump-h-1b-visa-rule/?sh=2b72f1321602.

Autor, David H., Frank Levy, and Richard J. Murnane. 2003. "The Skill Content of Recent Technological Change: An Empirical Exploration." *Quarterly Journal of Economics* 118 (4): 1279–1333.

Banerjee, Ishan, and Matt Sheeran. 2020. "America's Got AI Talent: US' Big Lead in AI Research Is Built on Importing Researchers." *Macro Polo*, June 9. https://macropolo.org/americas-got-ai-talent-us-big-lead-in-ai-research-is-built-on-importing-researchers/?rp=m.

Bariso, Justin. 2020. "Google Has a Plan to Disrupt the College Degree." *Inc.*, August 19. https://www.inc.com/justin-bariso/google-plan-disrupt-college-degree-university-higher-education-certificate-project-management-data-analyst.html.

Berg, Janine, Marianne Furrer, Ellie Harmon, Uma Rani, and M. Six Silberman. 2018. *Digital Labour Platforms and the Future of Work: Towards Decent Work in the Online World*. Geneva: International Labour Organization. www.ilo.org/wcmsp5/groups/public/---dgreports/---dcomm/---publ/documents/publication/wcms_645337.pdf.

Blit, Joel. 2020. "Automation and Reallocation: Will COVID-19 Usher in the Future of Work?" *Canadian Public Policy* 46 (S2): S192–S202.

Braverman, Harry. 1998. *Labor and Monopoly Capital: The Degradation of Work in the Twentieth Century*. New York: New York University Press.

Campbell, Mark. 2020. "Automated Coding: The Quest to Develop Programs that Develop Programs." *Computer* 53 (2): 80–82.

Canon, Gabrielle. 2018. "Google Gave Top Executive $90m Payoff but Kept Sexual Misconduct Claim Quiet—Report." *The Guardian*, October 25. https://www.theguardian.com/technology/2018/oct/25/google-andy-rubin-android-creator-payoff-sexual-misconduct-report.

CISTP. 2018. *China AI Development Report 2018*. Beijing: China Institute for Science and Technology Policy at Tsinghua University. http://www.sppm.tsinghua.edu.cn/eWebEditor/UploadFile/China_AI_development_report_2018.pdf.

Chun, Wendy Hui Kyong. 2011. *Programmed Visions: Software and Memory*. Cambridge, MA: MIT Press.

Columbus, Louis. 2019. "Data Scientist Leads 50 Best Jobs in America for 2019, According to Glassdoor." *Forbes*, January 23. https://www.forbes.com/sites/louiscolumbus/2019/01/23/data-sci entist-leads-50-best-jobs-in-america-for-2019-according-to-glassdoor/#12abe5667474.

Coursera. n.d. "Google IT Support Professional Certificate." Accessed September 10, 2021. https://www.coursera.org/professional-certificates/google-it-support?action=enroll.

Cummings, M. L., Heather M. Roff, Kenneth Cukier, Jacob Parakilas, and Hannah Bryce. 2018. *Artificial Intelligence and International Affairs: Disruption Anticipated*. London: Chatham House. https://www.chathamhouse.org/sites/default/files/publications/research/2018-06-14-artificial -intelligence-international-affairs-cummings-roff-cukier-parakilas-bryce.pdf.

Data Science Nigeria. 2019. *Annual Report July 2018–June 2019: Building One Million AI Talents in 10 Years*. Lagos: Data Science Nigeria. https://www.datasciencenigeria.org/wp-content/uploads /2019/08/annual-report-final.pdf.

Davenport, Thomas H., and D. J. Patil. 2012. "Data Scientist: The Sexiest Job of the 21st Century." *Harvard Business Review*, October. https://hbr.org/2012/10/data-scientist-the-sexiest-job-of -the-21st-century.

Dong, Catherine. 2017. "The Evolution of Machine Learning." *TechCrunch*, August 8. https://techcrunch.com/2017/08/08/the-evolution-of-machine-learning/.

Duke, Sue. 2018. "Will AI Make the Gender Gap in the Workplace Harder to Close?" World Economic Forum (website), December 21. https://www.weforum.org/agenda/2018/12/artificial -intelligence-ai-gender-gap-workplace/.

Dutton, Tim, Brent Barron, and Gaga Boskovic. 2018. *Building an AI World: Report on National and Regional AI Strategies*. Toronto: CIFAR. https://cifar.ca/wp-content/uploads/2020/05/buildinganai world_eng.pdf.

Finance Canada. 2017. "Growing Canada's Advantage in Artificial Intelligence." Ottawa: Government of Canada. https://www.canada.ca/en/department-finance/news/2017/03/growing_canada_ sadvantageinartificialintelligence.html.

Fuchs, Christian. 2014. *Digital Labour and Karl Marx*. New York: Routledge.

Gagne, JF. 2019. *Global AI Talent Report 2019*. https://jfgagne.ai/talent-2019/.

Getoor, Lise, David Culler, Eric de Sturler, David Ebert, Mike Franklin, and H. V. Jagadish. 2016. "Computing Research and the Emerging Field of Data Science." *CRA Bulletin*, October 7. https://cra.org/data-science/.

Gray, Mary L., and Siddharth Suri. 2019. *Ghost Work: How to Stop Silicon Valley from Building a New Global Underclass*. Boston: Houghton Mifflin Harcourt.

Haner, Justin, and Denise Garcia. 2019. "The Artificial Intelligence Arms Race: Trends and World Leaders in Autonomous Weapons Development." *Global Policy* 10 (3): 331–337.

Hao, Karen. 2020. "Trump's Freeze on New Visas Could Threaten US Dominance in AI." *MIT Technology Review*, June 28. https://www.technologyreview.com/2020/06/26/1004520/trump-executive-order-h1b-visa-threatens-us-ai/.

Harrison, Sara. 2019. "Five Years of Tech Diversity Reports—and Little Progress." *Wired*, October 1. https://www.wired.com/story/five-years-tech-diversity-reports-little-progress/.

Harwell, Drew. 2018. "Google to Drop Pentagon AI Contract after Employee Objections to the 'Business of War.'" *Washington Post*, June 1. https://www.washingtonpost.com/news/the-switch/wp/2018/06/01/google-to-drop-pentagon-ai-contract-after-employees-called-it-the-business-of-war/?utm_term=.f27036684980.

Henriques, Bruno. 2020. "Brazil Is Emerging as a World Class AI Innovation Hub." *Venture Beat*, January 12. https://venturebeat.com/2020/01/12/brazil-is-emerging-as-a-world-class-ai-innovation-hub/.

Huang, Tina, and Zachary Arnold. 2020. *Immigration Policy and the Global Competition for AI Talent*. Washington, DC: Center for Security and Emerging Technology. https://cset.georgetown.edu/publication/immigration-policy-and-the-global-competition-for-ai-talent/.

Hyde, Alan. 2003. *Working in Silicon Valley: Economic and Legal Analysis of a High-Velocity Labor Market*. Armonk, NY: ME Sharpe.

Kaplan, Jerry. 2016. *Artificial Intelligence: What Everyone Needs to Know*. Oxford: Oxford University Press.

Kunze, Lars. 2019. "Can We Stop the Academic AI Brain Drain?" *Künstlich Intelligenz* 33: 1–3.

Lee, Kai-Fu. 2018. *AI Superpowers: China, Silicon Valley, and the New World Order*. New York: Houghton Mifflin Harcourt.

Marx, Karl. (1867) 1990. *Capital Volume I*. New York: Penguin.

McClure, Sean. 2018. "GUI-fying the Machine Learning Workflow: Towards Rapid Discovery of Viable Pipelines." *Towards Data Science*, June 25. https://towardsdatascience.com/gui-fying-the-machine-learning-workflow-towards-rapid-discovery-of-viable-pipelines-cab2552c909f.

McNally, David. 1993. *Against the Market: Political Economy, Market Socialism and the Marxist Critique*. London: Verso.

Metz, Cade. 2018. "A.I. Researchers Are Making More than $1 Million, Even at a Nonprofit." *New York Times*, April 19. https://www.nytimes.com/2018/04/19/technology/artificial-intelligence-salaries-openai.html.

Microsoft. 2018. *The Future Computed: Artificial Intelligence and Its Role in Society*. Redmond, WA: Microsoft Corporation.

Microsoft News Center. 2016. "Democratizing AI: For Every Person and Every Organization." Microsoft Features, September 26. https://news.microsoft.com/features/democratizing-ai/.

Miller, Hannah, and Richard Stirling. 2019. *Government Artificial Intelligence Readiness Index 2019*. Malvern, UK: Oxford Insights.

Muller, Michael, Ingrid Lange, Dakuo Wang, David Piorkowski, Jason Tsay, Q. Vera Liao, Casey Dugan, and Thomas Erickson. 2019. "How Data Science Workers Work with Data: Discovery, Capture, Curation, Design, Creation." In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–15. New York: Association for Computing Machinery.

O'Connell, Michael. 2018. "Citizen Data Science and the Democratization of Analytics." *Information Week*, August 28. https://www.informationweek.com/big-data/ai-machine-learning/citizen -data-science-and-the-democratization-of-analytics/a/d-id/1332679.

Perrault, Raymond, Yoav Shoham, Erik Brynjolfsson, Jack Clark, John Etchemendy, Barbara Grosz, Terah Lyons, James Manyika, Saurabh Mishra, and Juan Carlos Niebles. 2019. *The AI Index 2019 Annual Report*. Human-Centered AI Institute. Stanford, CA: Stanford University.

Rao, Anand S., and Gerard Verweij. 2017. "Sizing the Prize: What's the Real Value of AI for Your Business and How Can You Capitalise?" PwC. https://www.pwc.com/gx/en/issues/analytics/assets /pwc-ai-analysis-sizing-the-prize-report.pdf.

Roland, Alex, and Philip Shiman. 2002. *Strategic Computing: DARPA and the Quest for Machine Intelligence, 1983–1993*. Cambridge, MA: MIT Press.

Salisbury, Allison Dulin. 2020. "COVID-19 May Become 'An Automation Forcing Event': Already Vulnerable Look to Reskilling for Path Forward." *Forbes*, May 7. https://www.forbes.com/sites /allisondulinsalisbury/2020/05/07/covid-19-may-become-an-automation-forcing-event-already -vulnerable-workers-look-to-reskilling-for-path-forward/.

Saphir, Ann. 2018. "As Companies Embrace AI, It's a Job-Seekers Market." Reuters, October 15. https://www.reuters.com/article/us-usa-economy-artificialintelligence/as-companies-embrace-ai -its-a-job-seekers-market-idUSKCN1MP10D.

Schwab, Klaus. 2016. *The Fourth Industrial Revolution*. New York: Crown Business.

Shoham, Yoav, Raymond Perrault, Erik Brynjolfsson, Jack Clark, and Calvin LeGassick. *The AI Index 2017 Annual Report*. Human-Centered AI Institute. Stanford, CA: Stanford University. https://hai.stanford.edu/ai-index-2017.

Snow, Jackie. 2018. "'We're in a Diversity Crisis': Cofounder of Black in AI on What's Poisoning Algorithms in Our Lives." *MIT Technology Review*, February 14. https://www.technologyreview .com/2018/02/14/145462/were-in-a-diversity-crisis-black-in-ais-founder-on-whats-poisoning-the -algorithms-in-our/.

Stack Overflow. 2018. *Developer Survey Results 2018*. London: Stack Overflow. https://insights .stackoverflow.com/survey/2018/#overview.

Statista. 2020. "Market Size and Revenue Comparison for Artificial Intelligence Worldwide from 2015 to 2025." Accessed November 30, 2020. https://www-statista-com.proxy1.lib.uwo.ca /statistics/941835/artificial-intelligence-market-size-revenue-comparisons/.

Theuwissen, Martijn. 2015. The Different Data Science Roles in the Industry. KDnuggets, November. https://www.kdnuggets.com/2015/11/different-data-science-roles-industry.html.

Toews, Rob. 2020. "A Massive Opportunity Exists to Build 'Picks and Shovels' for Machine Learning." *Forbes*, March 22. https://www.forbes.com/sites/robtoews/2020/03/22/a-massive-opportunity-exists-to-build-picks-and-shovels-for-machine-learning/#6b3f036a7ab3.

Tuggener, Lukas, Mohammadreza Amirian, Katharina Rombach, Stefan Lörwald, Anastasia Varlet, Christian Westermann, and Thilo Stadelmann. 2019. "Automated Machine Learning in Practice: State of the Art and Recent Results." In *Proceedings of the 2019 6th Swiss Conference on Data Science*, 31–36. Bern, Switzerland: IEEE.

Vassallo, Trae, Ellen Levy, Michele Madansky, Hillary Mickell, Bennett Porter, Monica Leas, and Julie Oberweis. n.d. "Elephant in the Valley." Accessed August 18, 2021. https://www.elephantinthevalley.com/.

Walch, Kathleen. 2020. "Why the Race for AI Dominance Is More Global Than You Think." *Forbes*, February 9. https://www.forbes.com/sites/cognitiveworld/2020/02/09/why-the-race-for-ai-dominance-is-more-global-than-you-think/.

Walker, Kent. 2020. "A Digital Jobs Program to Help with America's Economic Recovery." *Grow with Google* (blog), July 13. https://blog.google/outreach-initiatives/grow-with-google/digital-jobs-program-help-americas-economic-recovery/.

Wang, Dakuo, Justin D. Weisz, Michael Muller, Parikshit Ram, Werner Geyer, Casey Dugan, Yla Tausczik, Horst Samulowitz, and Alexander Gray. 2019. "Human-AI Collaboration in Data Science: Exploring Data Scientists' Perceptions of Automated AI." In *Proceedings of the ACM on Human-Computer Interaction* 3, Issue CSCW: 1–24.

West, Sarah Myers, Meredith Whittaker, and Kate Crawford. 2019. "Discriminating Systems: Gender, Race, and Power in AI." New York: AI Now Institute, New York University. https://ainowinstitute.org/discriminatingsystems.html.

Zhang, Yulin. 2015. "Land Grabs in Contemporary China." *Nao's Blog*, January 6. Translated by Pancho Sanchez. https://libcom.org/blog/china-land-grabs.

Zikopoulos, Paul, and Chris Eaton. 2011. *Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data*. New York: McGraw-Hill Osborne Media.

Zöller, Marc-André, and Marco F. Huber. 2021. "Benchmark and Survey of Automated Machine Learning Frameworks." *Journal of Artificial Intelligence Research* 70: 409–472. https://doi.org/10.1613/jair.1.11854.

Zoph, Barret, Vijay Vasudevan, Jonathon Shlens, and Quoc Le. 2017. "AutoML for Large Scale Image Classification and Object Detection." Google AI Blog, November 2. https://ai.googleblog.com/2017/11/automl-for-large-scale-image.html.

Zwetsloot, Remco, James Dunham, Zachary Arnold, and Tina Huang. 2019. *Keeping Top AI Talent in the United States: Findings and Policy Options for International Graduate Student Retention*. Washington, DC: Center for Security and Emerging Technology. https://cset.georgetown.edu/wp-content/uploads/Keeping-Top-AI-Talent-in-the-United-States.pdf.