Preliminary. Please do not cite or circulate

Machine Learning and Occupational Change

Erik Brynjolfsson, MIT Sloan School of Management and NBER
Tom Mitchell, Carnegie Mellon University
Daniel Rock, MIT Sloan School of Management

May 30, 2018

Note to Ajay, Avi and Joshua: This paper is still a work in progress. Here's our timeline: As of today, we have completed the development of all three rubrics (SML, Rule-based and Robotics) and have completed contracting with a crowdsourcing platform. The rubrics are being validated by experts in ML, robotics and rule-based systems. Current empirical results are based on a first wave of data using only the SML rubric. We expect to have three massive new datasets available for our analysis by the end of June. Before the end of July, we will complete the new analyses, and use August to fine tune them, write up the results and get some additional feedback from colleagues. In parallel, we are refining the theory, and that seems relatively straightforward to update. We hope to use the Toronto workshop to unveil our results.

Acknowledgments: We thank participants in seminars at MIT, Harvard Business School and the American Economic Association annual meeting (2018) for useful comments on this research and the MIT Initiative on the Digital Economy for generous funding.

Until very recently, the idea of automated prediction technologies guiding motor vehicle traffic without human intervention was little more than science fiction. Now, new developments in machine learning have led researchers to find this and many other AI challenges within reach. The basic neural net models and algorithms thriving today are decades old (White and Rosenblatt 1963; Rumelhart, Hinton, and Williams 1986; LeCun et al. 1998). But now the widespread availability of inexpensive, scalable computing resources and the proliferation of large scale training data to build Deep Neural Nets (DNNs) with viable commercial applications has stirred rapid growth in DNN research and investment.[1] According to the 2017 AI Index Annual Report, by 2016 the number of AI papers had increased nine-fold over its 1996 level, the number of AI systems startups had increased by a factor of 14 since 2000, and venture capitalists had invested six times the capital in AI startups as in 2000.[2] AI is poised to transform occupations across the economy much as rules-driven information technology did in prior years.

AI is a general-purpose technology (GPTs). Like all GPTs, AI is pervasive, improves over time, and spawns complementary innovations (Bresnahan and Trajtenberg 1995; Bresnahan 2010; Brynjolfsson, Rock, and Syverson 2017). In particular, the recent advances in AI depend on improvements in machine learning (ML). ML systems are increasingly adopted in virtually every industry. Perhaps more than any other technology, the essence of ML is that it improves with better training data and better algorithms. The practical application of of AI and ML almost always requires complementary investments in intangible capital – new business processes, redesign of jobs, training and human capital investments – and enables the development of new goods and services like digital concierges, data center energy optimization, and even fake news filtration.[3] As noted by Furman and Seamans (2018), recent work on the effects of AI on innovation (Cockburn, Henderson, and Stern 2018), economic growth (Aghion, Jones, and Jones 2017; Agrawal, McHale, and Oettl 2018), productivity (Brynjolfsson, Rock, and Syverson 2017), and public policy (Furman and Seamans 2018;

[1] Deep learning (or Deep Neural Nets) describes a class of models within the field of Machine Learning (ML), which is itself situated within the broader field of Artificial Intelligence (AI).
[2] http://cdn.aiindex.org/2017-report.pdf
[3] https://ai.googleblog.com/2018/05/duplex-ai-system-for-natural-conversation.html; https://deepmind.com/blog/deepmind-ai-reduces-google-data-centre-cooling-bill-40/; https://www.technologyreview.com/s/609717/can-ai-win-the-war-against-fake-news/

Trajtenberg 2018; Goolsbee 2018) suggest that the coming impact of AI will be widespread across the economy.[4]

Progress over the past decade has been especially fast using techniques of deep learning. These fall within the machine learning paradigm of supervised learning, where a model is constructed to map observable input features into observable output labels or values (Hinton, Bengio, and LeCun 2015). Deep learning is not a unique solution concept for this kind of task. Linear regression, for example, is commonly used for the same purpose. Similar to linear regression models, a deep learning model will apply different weights and functional transformations to the input values in order to predict or classify the outputs. These models are complicated and it can be difficult to explain why they make a particular prediction.[5]

The key advantage of this approach over rules-driven software though is that they do not require a pre-determined recipe or series of codified instructions to map input values to outputs. The mapping function is *learned* automatically via an optimization algorithm.[6] This has at least two major economic implications. First, DNNs help overcome Polanyi's Paradox that "we know more than we can tell". This has been a major obstacle to automation by software in the past (Polanyi 1966; Autor 2014). For instance, while humans can easily recognize faces, it is difficult to write formal rules for how to do it. Until recently, that made it difficult for machines to do this task. However, now, using DNNs, machines have learned to do it very well. Furthermore, the machine learning approach has proven surprisingly effective for a wide variety of other tasks, from recommending products to consumers and placing ads on websites, to translating speech and diagnosing cancer. In many cases, machine learn systems now match or even surpass humans doing the same tasks (see e.g. Esteva et al. (2017); Graves, Mohamed, and Hinton (2013)).[7] One can think of these applications of AI as reducing the cost of various forms of prediction, leaving judgment tasks to humans (Agrawal, Gans, and Goldfarb 2017, 2018). The widespread applicability of "prediction machines" has potentially sizeable implications for changing the nature of economic activity. In particular, because tasks like these are important parts of many occupations, machine learning is likely to have significant effects on content, wages and demand of for many types of workers.

---

[4] Furman and Seamans (2018) point out in particular that much of the current research on the Economics of AI is in NBER *Economics of Artificial Intelligence* volume (edited by Agrawal, Gans, and Goldfarb).
[5] For a brief overview, Appendix A describes the types of machine learning paradigms.
[6] Typically, Stochastic Gradient Descent (SGD) or some variant of SGD.
[7] Also see the AI Index at http://www.aiindex.org

At the same time, there are also many tasks people do that cannot be done by machines at the current frontier of automation. Artificial General Intelligence (or AGI), which seeks to match humans in all tasks, remains a distant goal. Therefore, it is useful to distinguish between the set of tasks mostly like to be affected by machine learning and those which are not likely to be affected, and assess the likely implications for tasks, occupations, and the economy.

Accordingly, in this paper we seek to assess how and where ML will affect the workforce, extending the ideas of Brynjolfsson and Mitchell (2017) and Brynjolfsson, Mitchell, and Rock (2018). Brynjolfsson and Mitchell (2017) outline a series of factors that relate machine learning and labor demand, offering a 23-question rubric to evaluate each task in the economy for its "Suitability for Machine Learning" or SML. Brynjolfsson, Mitchell, and Rock (2018) provide a first implementation of that rubric using the O*NET database from the U.S. Bureau of Labor Statistics.

This paper expands and improves that implementation in three ways. First, we build a significantly larger data sample using a new crowdsourcing platform. This enables us to we create a variety of measures of SML and SML-related concepts at the level of regions, industries, and firms (in addition to occupations). Second, we extend the rubric concept and implementation with two new rubric editions: rules-based software and robotics, in addition to an improved machine learning rubric. Because the rules-based approach has been affecting occupations for over 30 years, the implementation of this rubric makes it possible for us to benchmark actual historical changes against our rubric's predictions. The robotics rubric extends the set of tasks and occupations analyzed to those that involve dexterity, mobility and other physical skills. Thirdly, we develop an economic model of the occupation-level production process, modeling the complementarities among tasks, rather than simply looking at weighted averages of tasks within occupation. Specifically, present a model based on Autor, Levy, and Murnane (2003) and Holmstrom and Milgrom (1991) which describes some of the task-level effects of occupational re-design following ML automation. We supplement the original six factors relating machine learning and labor demand in Brynjolfsson and Mitchell (2017) with the findings from that model.

While building on the ideas in Brynjolfsson, Mitchell, and Rock (2018) and related work by Felten, Raj, and Seamans (2018), these additional measures help contextualize concerns about occupations which are generally exposed to automation technologies. Our results for the SML rubric suggest that very few (if any) occupations can be fully automated by ML, but most occupations have at least one task that is exposed to a possible ML implementation. This suggests that realizing the potential gains to the ML GPT will require substantial investment in the redesign of occupations and

business processes. Our rubrics are designed to evaluate technological potential in occupations in the intermediate term (a few years to a decade or so in the future). Nevertheless, some predictions based on which types of tasks are suitable for machine learning are possible.

An important limitation of our approach is that while we assess the effects of ML, rules-based systems, and robotics on existing tasks, we do not predict the *new* tasks which will be created as technology advances. As stressed by Brynjolfsson and McAfee (2011) and Acemoglu and Restrepo (2017), humans do not simply race against machines, but also work with them to complete tasks they could never have solved before. However, imagining and anticipating these tasks before they exist is far more difficult than analyzing existing tasks. For the purposes of this paper, we leave this challenge to savvy entrepreneurs, creative managers, and perhaps more skilled researchers.

Automation, Artificial Intelligence, and Labor Demand

Impressive as they are, the recent growth trends in AI-related investment and research constitute a relatively small fraction of the economy. It is too soon to conclusively measure any major economic effects of AI technologies, much less declare any kind of transformational change. Here we borrow from the economics and management literature on the economic effects of more general trends in information technology and automation. The base unit of analysis in our technological rubrics is the task. Jobs are then bundles of tasks, and those bundles are a design choice by employers. This follows the work of Autor, Levy, and Murnane (2003) (ALM), which explains shifts in labor demand by analyzing the impact of IT-related automation at the task-level. In particular, ALM offers a model whereby some tasks can be substituted for by capital and others cannot. They find that computer capital has reduced the relative demand for routine types of work tasks. Routineness in this context is a measure of how possible it is to describe a given task with rules, indicating a viable automation opportunity for standard types of software. That is, these tasks were not subject to Polanyi's Paradox.

Acemoglu and Restrepo (2017) adopt a task-level approach to understanding how the incentives to automate or employ humans are contingent on the available supply of labor for doing tasks (and the prevailing wage), the research costs to develop automation capital for existing tasks, and the research costs to develop new tasks for human workers. These trade-offs might help to understand, for example, the Kaldor (1961) stylized fact that, until recently, the labor and capital shares of income were fairly stable even as more and more tasks were automated. We are concerned

principally with the piece describing the research costs of machine learning-based automation and less on the creation of new tasks, though this is obviously a crucial element of the impact of ML technology on labor demand. Much of the impact of IT, another GPT, resulted from the follow-on complementary innovation and reconfiguration of economic activity (Bresnahan et al. 1996; Greenstein and Prince 2005). Our analysis centers on the task components which are most suitable for ML, for rule-based systems, and for robotics.

While ML and robotics can automate a different set of tasks compared to rules-based IT, some of the mechanisms driving shifts in labor demand may be similar. Indeed, earlier waves of automation also had a major impact on productivity, wages, employment, and investment (Acemoglu and Restrepo 2017a; Acemoglu and Autor 2011; Autor and Dorn 2013). While we might not see the same direction of skill-biased technological change and polarization that resulted from earlier waves of automation, it is likely that demand for human tasks easily substituted by ML capital will decline relative to demand for less suitable for machine learning tasks. Furthermore, because ML is differentially suited to different purposes, the distribution of the technology's impact across occupations will likely be uneven (Levy 2017).

Concern about the effect of AI on the "Future of Work" and how to deal with the impending automation of huge swaths of the workforce are growing. Some studies suggest that nearly half of the work activities in the U.S. could be automated with existing technology (Bughin et al. 2017). If workers could be completely automated, it is possible that the economic effects would immiserate laborers in the long run (Benzell et al. 2016).[8] Yet seemingly in spite of the pace of technological innovation in the past century, the vast majority of prime-age workers are still employed in every economy around the world. This is largely because of the strong complementarities between automation technologies and the tasks that they do not replace (Autor 2015). For example, Bessen (2016) shows that employment of bank tellers initially *increased* after the adoption of the ATM. The tellers were put to work emphasizing tasks that the machines were poorly suited to do. Bessen (2018) argues that the elasticity of labor demand is a major factor too in the overall effects of an automation technology. With elastic demand for the non-automated tasks, we might expect employment in non-automated tasks to increase as the costs of complementary automated tasks drop.

[8] Acemoglu and Restrepo (2017) refer to this as the "horse equilibrium".

Brynjolfsson and Mitchell (2017) identify demand elasticity as one of six determinants of how ML and labor demand and wages are related, adding as well

2) complementarities,

3) substitution,

4) supply elasticity,

5) income elasticity, and

6) reengineering.

Substitution effects are the primary worry of the neo-Luddites, and the complementarity effects are similar to those in Autor (2015). Brynjolfsson and Mitchell (2017) point out that cheaper production via ML technology has the effect of increasing wealth as well. The effects of greater income change the production function directly through employer demand for labor and indirectly through consumer demand for products. ML can also affect labor supply elasticity; an inelastic labor supply for complementary non-automatable work could lead to higher wages in that task. Lastly, reengineering the task bundles into new jobs may be an important source of innovation. Firms that take advantage of better job designs that allow machines and humans to specialize will generate output more efficiently. In accordance with the LeChatelier Principle (Milgrom and Roberts 1996), production inputs are typically more elastic in the long-run.

Why might this time be different?

DNNs and similar architectures[9] are appropriate for solving types of problems which other AI and machine learning approaches perform poorly.[10] The reason that these structures are so useful in prediction tasks is because the learned parameters in these models can approximate nonlinear functions with high dimensionality. Input features are (in many cases) automatically transformed to represent data in better ways, making it possible to use DNNs in the case where "it is easier to collect the data than to explicitly write the program".[11]

Models of this type with human-level (and in some cases superhuman) performance have been trained to classify images, translate speech and text, and even to prove mathematical theorems (e.g. Loos et al. 2017). Often DNN algorithms are combined with reinforcement learning

[9] Including but not limited to convolutional neural nets (CNNs), recurrent neural nets (RNNs), long short-term memory (LSTMs), generative adversarial networks (GANs), and ordinary artificial neural nets (ANNs).
[10] For example, linear regression, support vector machines, or tree-based models, let alone rule-base systems.
[11] https://medium.com/@karpathy/software-2-0-a64152b37c35

frameworks and other types of machine learning to build high performance composite models. Google DeepMind's AlphaZero algorithm achieved superhuman performance in chess, shogi, and Go within 24 hours of training combining a number of different techniques (Silver et al. 2017). Given an appropriate domain and favorable contextual factors, these algorithms can automate many work activities where the goal is to map inputs to outputs. Recent advances (see e.g. Goodfellow et al. 2016) have implications for prediction and classification tasks in fields ranging from cryptography (Shokri and Shmatikov 2016), to medical imaging (Litjens et al. 2017), and even to archaeology (Guyot, Hubert-Moy, and Lorho 2018). A particular combination of DNNs with reinforcement learning techniques called "inverse" reinforcement learning means that robots can be programmed with videos of the tasks they are meant to perform (Abbeel and Ng 2004; Finn et al. 2016). Machine learning is therefore a method for discovering better programs (or as Tesla Director of AI Andrej Karpathy calls it, "Software 2.0").[12] Though particular trained models might not generalize to different tasks, the class of models is potentially useful for a large contingent of oft-repeated but hard to codify tasks which so far have eluded automation.

A Model of Machine Learning Automation

The model in Autor, Levy, and Murnane (2003) is a useful starting point for understanding the incentives to re-bundle tasks into new occupations as automation technologies improve. We contrast the ALM model with a constrained version of the same firm problem in the constant elasticity of substitution case. Following ALM, assume production Q (equation 1) is a constant elasticity of substitution function with total factor productivity parameter A, capital inputs K (including intangible capital used in machine learning – like data), labor in efficiency units devoted to machine learnable (SML) tasks L_{ML}, labor in efficiency units devoted to other (non-SML) tasks L_o, share parameter α where $1 \geq \alpha \geq 0$ and elasticity of substitution $1/(1-\sigma)$. Capital and labor suitable for machine learning are assumed to be perfect substitutes. Workers have productivity endowments in ML tasks and other tasks respectively of $E_i = [ML_i, O_i]$ where $1 \geq ML_i, O_i > 0 \ \forall i$ and the labor supplied is $L_i = [\lambda ML_i, (1-\lambda)O_i]$, $\lambda \in [0,1]$. Our production function Q is then:

$$(1) \quad Q = A(\alpha(K + L_{ML})^{\sigma} + (1-\alpha)L_o^{\sigma})^{\frac{1}{\sigma}}$$

[12] In the blogpost from the previous footnote.

where at the margin income-maximizing workers are indifferent between offering ML-able services and other services. If the marginal user cost of capital (Jorgenson 1963) to develop ML applications r sets the wage for ML-able labor W_{ML}, we have wage equations (2) and (3) from first order conditions:

$$(2) \quad W_{ML} = \alpha A^{\sigma} \left(\frac{Q}{K + L_{ML}} \right)^{1-\sigma} = r$$

$$(3) \quad W_o = (1-\alpha) A^{\sigma} \left(\frac{Q}{L_o} \right)^{1-\sigma}$$

and then the ratio of capital and ML labor to other labor is:

$$(4) \quad \theta \equiv \frac{K + L_{ML}}{L_o} = \left(\frac{(1-\alpha) W_{ML}}{\alpha W_o} \right)^{\frac{1}{\sigma - 1}}$$

Letting σ approach negative infinity (i.e. a Leontief production function) with constraints on task groupings illustrates how suboptimal bundling of tasks can block potential productivity gains. Under the conditions presented above, if the rental cost of ML capital went to a marginal value of 0, the workers would prefer to all switch into doing only the tasks that ML cannot. While nonzero in general, marginal costs of building these models need not exceed the costs of computation in specific cases, a cost which is already low and continues to fall very rapidly. The final changes in the ratio of capital and total labor employed, however, depend on the factor shares of different types of inputs and the elasticity of substitution between the different inputs. The Leontief production case demonstrates a version of extreme complementarity, but this need not be the case. Further, while more tractable and useful for this illustration, the elasticity of substitution in multifactor production functions need not be the same across inputs.

Now assume that firms, perhaps because of historical business processes, implementation frictions, and/or adjustment costs, only offer labor contracts that have some preset mixture of SML and non-SML tasks. Formally we have a modification to the first order condition such that it is now constrained (this constrained problem is not in the ALM model). The firm solves:

$$(5) \quad \max_{L_{ML}, L_o, K} Q(L_{ML}, L_o, K) - w_{ML} L_{ML} - w_O L_O - rK$$

$$s.t. \; L_{ML} - \beta L_o = 0 \; and \; \beta \geq 0$$

This carries the additional constraint (in addition to feasible production) that the ratio of machine learnable labor and non-SML labor is fixed at a constant β. We assume that the costs to reorganize

job task bundles in this case might be prohibitively costly, so the firm in the short-run is bound by the constraint. The new wage conditions with Lagrange multiplier μ are:

$$(6) \quad \frac{\partial Q}{\partial L_{ML}} - W_{ML} - \mu = 0 = \alpha A^{\sigma} \left(\frac{Q}{K + L_{ML}}\right)^{1-\sigma} - W_{ML} - \mu$$

$$(7) \quad \frac{\partial Q}{\partial L_o} - W_o + \beta\mu = 0 = (1-\alpha) A^{\sigma} \left(\frac{Q}{L_o}\right)^{1-\sigma} - W_o + \beta\mu$$

$$(8) \quad L_{ML} = \beta L_o$$

This leads to an intermediate step that shows that in the presence of no automation, wages and marginal products are the same. But differences between marginal products and wages in one type of labor must be compensated for in the other type of labor:

$$(9) \quad \beta \left(\frac{\partial Q}{\partial L_{ML}} - W_{ML}\right) = -\left(\frac{\partial Q}{\partial L_o} - W_o\right)$$

Leading to

$$(10) \quad W_{ML} = r$$

when wages in SML tasks are set by the user cost of capital *r* and

$$(11) \quad W_o = \beta \left(\alpha A^{\sigma} \left(\frac{Q}{K + L_{ML}}\right)^{1-\sigma} - r\right) + (1-\alpha) A^{\sigma} \left(\frac{Q}{L_o}\right)^{1-\sigma} = \frac{\partial Q}{\partial L_o} + \beta \left(\frac{\partial Q}{\partial L_{ML}} - r\right)$$

Equation (11) illustrates that the non-SML task is compensated at a (weakly) lower rate since $\beta > 0$ and $r \geq$ the marginal product of SML tasks, or it would not be automated. Total compensation is then:

$$(12) \quad W_{ML} L_{ML} + W_o L_o = r L_{ML} + \left(\frac{\partial Q}{\partial L_o} + \beta \left(\frac{\partial Q}{\partial L_{ML}} - r\right)\right) L_o$$

The last term in equation (12) shows that each additional unit of non-SML labor hired has an opportunity cost in terms of the difference in marginal product between SML labor and capital. In the case of full automation, wages paid to SML labor would be zero and the wage paid to non-SML labor would equal the marginal product of non-SML labor (See Benzell and Brynjolfsson (2018) for a related model and some macroeconomic implications). This illustrates a shift in the labor demand curve: the elimination of the constraint to hire SML labor results in higher employment for non-SML labor at all wages. Thus, re-bundling tasks in new occupations will reward clever entrepreneurs with the difference between their new labor demand curve and the demand they would face under the constraint of the old job design.

In the Leontief case all of the labor effort put toward SML tasks has an opportunity cost in efficiency unit output of the share-adjusted non-SML labor. This is because production capability is the minimum of the adjusted input efficiency units. ML could be doing those tasks, and the firm could increase profit if it were able to reorganize their job design. Note that this provides and incentive for lower average O_i productivity workers to take on the non-SML task, leading to cost disease effects similar to those explored in greater detail in Baumol (1967) and specifically for AI in Aghion, Jones, and Jones (2017) and Benzell and Brynjolfsson (2018).

Job-Task Bundle Design and Measurability

In some cases, suitability for machine learning is also a question of measurability and monitoring. This is especially true for reinforcement learning algorithms (RL), which can be combined with deep learning to build AI systems that take actions autonomously. Instead of Karpathy's insight for supervised learning, we would characterize the heuristic for RL as "if you can measure the outcomes, the program can write itself".[13] This has implications for the types of tasks that remain as building blocks for non-SML occupations once ML systems have been deployed. Defining SML-as-measurability is restrictive, but is helpful for illustrating the risk transfer effects we might expect from ML. With ML systems (deep learning, reinforcement learning, or some combination) substituting for the easiest to measure tasks, the residual task bundles will 1) in the short-run diminish the outside option for task attention of the worker and 2) possibly cause a risk transfer between firms and workers because workers will now focus on tasks where performance is noisier. Put differently, when the highest signal-to-noise ratio tasks are taken by machines, the marginal task to automate is better suited to low-powered incentives.

We describe some of the basic information effects of machines on job design below, using the setup of Holmstrom and Milgrom (1991). Holmstrom and Milgrom (1991) (henceforth HM) show that it is suboptimal to make two agents jointly responsible for the same noisily measured task. What if one of the agents, however, is a machine? The machine will not care about the risk in its performance. And we know from ALM and the above model that firms are incentivized to automate where they can on the margin. Following equations 16-24 in Holmstrom and Milgrom (1991) (henceforth HM), let the realized performance in a task k be measured as:

[13] https://medium.com/@karpathy/software-2-0-a64152b37c35

$$(13) \quad x_k(t) = \mu(t,k) + \epsilon_k$$

The measurement is a function of the signal μ(k, t), itself a function of the task k and effort t, with noise epsilon. The normalized performance is defined below in equation (14) and on the margin the compensation for a completed task given a performance measurement must be equivalent across tasks.

$$(14) \quad \hat{x}_k(t) = \frac{\mu(t,k) + \epsilon_k}{\mu'(t(k),k)}$$

If the total effort t (bar) that an agent i puts into their assigned tasks (following equation 13 in HM) is:

$$(15) \quad \bar{t}_i = \int t_i(k)dk$$

And for any task in which the agent is assigned work by the principal there is the condition that the per unit of output for task k normalized compensation rate α(k) multiplied by the marginal increase in signal value (the marginal benefit to the agent) is equal to the marginal cost of total effort C (following equation 15 in HM):

$$(16) \quad \alpha_i(k)\mu'(t(k),k) = C'(\bar{t}_i) \; if \; t_i(k) > 0, k \in [0,1]$$

We take the solution to the program in HM equations 16-20. Effort is costly, but the agent will take effort to overcome a noisy signal if the compensation rate is high enough. Herein we find a negative first-order effect if we consider the HM model in light of our adaptation of the ALM model. If some region of the possible k tasks is removed by a machine such that we have a truncated interval, then the set of tasks to which human labor is assigned diminishes. This is similar to the effects in (Acemoglu and Restrepo 2017b) when automation starts to squeeze human labor into a smaller region of the available tasks. Here, however, human labor is squeezed into the task region wherein the tasks are least measurable.

Effectively humans begin to specialize in those tasks which have *low* normalized commission rates (α(k)/μ'(t(k),k)). HM points in proposition 7 to an optimal solution where the human agent (in our case) would work less hard and specialize in low measurability tasks. In this case, at every price labor supply would increase as humans formerly working on measurable tasks begin to compete in hard-to-measure tasks. This is the kind of change we would expect for a reinforcement learning (RL) system, especially an RL system augmented by embedded DNNs. Both humans and machines have better performance when it is easy to map effort and action into rewards. Within the technological frontier for RL, the tasks and incentive schemes which RL can manage well will generate effects of

the style presented in the ALM model above. Effectively the elimination of less noisy work by machines leads to greater competition for noisy, poorly measured work. Even if the costs of effort were identical (i.e. automation costs were similar to labor effort costs on the margin), the risk-averse agents choose a (weakly) lower optimal total effort quantity, and also receive less for their work. On the margin, the risk-neutral machine competitor has a lower reservation wage for effort in tasks with less noisy outcomes. The main effect, however, is that of ALM: for some tasks, the machine capital can do the same task at a lower price.

Measuring Suitability for Machine Learning at the Level of Tasks

As in Brynjolfsson, Mitchell, and Rock (2018), we measure the suitability for machine learning (SML) using the SML rubric in Brynjolfsson and Mitchell (2017), although for this paper we improve and extend it somewhat.[14] The degree to which there is an opportunity for machine learning (and DNNs in particular) to automate a task relies on the characteristics of the particular work activity and the context in which it is done. The first iteration of the SML rubric contained 23 distinct evaluative statements designed to measure the ML potential in a given work activity's context and characteristics. The main idea behind these rubrics is to give primacy to the economic relevant features of the technology in question, whether that be ML, IT, or robotics. Each of these rubrics is customized to reflect the focal "affordances and constraints" of that particular style of automation capital (Pickering 1993).

[Note: We plan to have the next iteration complete in the coming weeks, along with first runs of rubric evaluations for suitability for rules-based software (SRBS) and suitability for robotic systems (SRS). Furthermore, we plan to simplify some of the statements, bringing them closer to the concepts in the nascent economics of AI literature. For example, building on Agrawal, Gans, and Goldfarb (2018), we will include the following in the next survey round:

New Q1: This task is primarily about predicting a value or classifying an object
New Q2: This task requires use of good judgement in choosing the correct course of action

[14] With a slight modification to add two statements describing the physical requirements of the activity. The rubric is included in Appendix B.

Questions like these will help us validate our approach and determine how different models of the impact of AI on labor demand might lead to disparate conclusions.]

The current data includes evaluation statements such as:

Q1: Information need to complete the task (inputs) and outputs can be explicitly specified in machine-readable format
Q3: Task feedback (on the success of outputs) is immediate
Q8: The task is principally concerned with matching information to concepts, labels, predictions, or actions
Q11: The task is describable with rules
Q12: There is no need to explain decisions during task execution
Q16: The task requires working with image/video data or might require working with image/video data in the future

In our most recent iteration, we applied the SML rubric to 2,069 detailed work activities (DWAs) in the Bureau of Labor Statistics O*NET content model for all of the occupations in the U.S. We use CrowdFlower[15], a Human Intelligence Task (HIT) platform, to implement a crowdsourced evaluation of the 2,069 DWAs, rating each DWA for each of the 23 statements on a 5-point scale varying from "Strongly Disagree (1)" to "Strongly Agree (5)". To improve the quality of our responses, each DWA is rated by at least 7 distinct respondents. Respondents for whom there is no variance in responses across all 23 statements are dropped, as are respondents who self-report that they do not understand the DWA that they are rating (without penalty). For the remaining responses, we take the median score for each statement as the value for that DWA-statement tuple.

Higher values correspond to more suitability for machine learning, and vice-versa. These DWA scores for each of the 23 statements in our rubric are many-to-many mapped to 18,156 specific occupation-level tasks (with equal weighting within a task). The tasks in O*NET are weighed by the normalized occupational importance scores provided in the O*NET database and aggregated to form our occupation-level measures.

[15] Now known as Figure 8: https://www.figure-eight.com/

The final measure is constructed via a simple process. We ask about 4 different types of data (image/video, speech, text, and structured/sensor data) for each one of the activities. These values are unlikely to all be rated highly for each activity, but individually a high value might be sufficient to make a task high SML. We therefore create a measure of data intensity as the maximum of the responses to statements 15 through 18 which correspond to these data types. This data score, the scores from statements 1-14, and scores for statements 19-23 are averaged to generate our primary SML measure. Table 1 below summarizes the primary SML measure for DWAs, tasks, and jobs; Figure 1 below shows a histogram of the SML scores. The mean SML measure in DWAs, tasks, and jobs is close to 3.47, with a maximal SML score at the DWA-level of 4.475 and a minimal SML score at 2.375. The standard deviation of SML scores for occupations is about 0.115, approximately 1/3 of the standard deviation of DWA SML. This reflects the diversification effect of bundling multiple activities into the same occupation. The range between the maximal and minimal SML score for the occupational averages is less than that of DWAs and tasks as well. Table 2 shows the top and bottom 15 occupations ranked by SML.

TABLE 1 — SUITABILITY FOR MACHINE LEARNING: SUMMARY STATISTICS

	Occupations (966)	Tasks (19,612)	Detailed Work Activities (2,069)
Mean Suitability for Machine Learning (SML)	3.4656	3.467	3.468
Standard Deviation of SML	0.1148	0.3094	0.321
Minimum SML	2.7805	2.375	2.375
25th Percentile SML	3.3972	3.25	3.25
75th Percentile SML	3.536	3.675	3.7
Max SML	3.9018	4.475	4.475

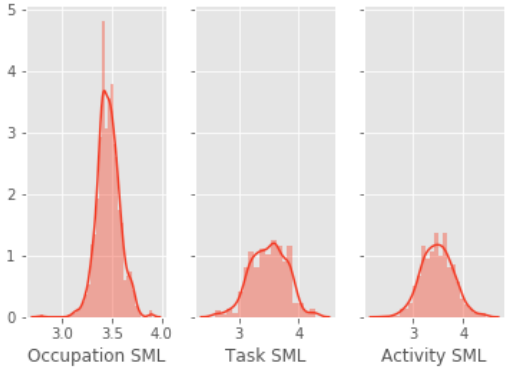


Figure 1- Distribution of Counts of Suitability for Machine Learning (SML) Score for Occupations, Tasks, and Detailed Work Activities

TABLE 2 — LOWEST AND HIGHEST 15 SML SCORE OCCUPATIONS

Rank	Lowest SML Ranked	SML	Highest SML Ranked Occupations	SML

	Occupations			
1	Massage Therapists	2.78	Concierges	3.90
2	Animal Scientists	3.09	Mechanical Drafters	3.90
3	Archeologists	3.11	Morticians, Undertakers, and Funeral Directors	3.89
4	Public Address System and Other Announcers	3.13	Credit Authorizers	3.78
5	Plasterers and Stucco Masons	3.14	Brokerage Clerks	3.78
6	Dancers	3.15	Gaming Cage Workers	3.77
7	Slaughterers and Meat Packers	3.15	Gem and Diamond Workers	3.75
8	Logging Equipment Operators	3.15	Office Machine Operators, Except Computer	3.74
9	Stonemasons	3.18	Multimedia Artists and Animators	3.74
10	Hoist and Winch Operators	3.18	Ushers, Lobby Attendants, and Ticket Takers	3.74
11	Brickmasons and Blockmasons	3.19	Insurance Claims Clerks	3.74
12	Physical Therapist Assistants	3.20	File Clerks	3.73
13	Animal Trainers	3.20	Fiberglass Laminators and Fabricators	3.72
14	Segmental Pavers	3.21	Bookkeeping, Accounting, and Auditing Clerks	3.72
15	Athletes and Sports Competitors	3.21	Document Management Specialists	3.72

The highest SML occupation in the scores from our first implementation of the rubric is concierge, with a SML score of 3.9. Indeed, companies like Google are already testing automated concierges.[16] A close second is mechanical drafter. Here complementarities may be as important as substitution effects.[17] The typical occupation with a high SML score in the first run tends to be data-intensive and clerical in nature, with morticians, credit authorizers, brokerage clerks, gaming cage workers, insurance claims clerks, file clerks, accounting clerks, and document management specialists all showing up in the top 15. In contrast, the lowest SML occupations often have some highly-skilled physical component to their work. Massage therapists are the least SML job, but dancers, meat packers, stonemasons, brickmasons, physical therapists, and professional athletes also making the low SML list. Economists, for the interested, are relatively close to the center of the SML distribution.

An important point, however, is that most occupations do have some component activity or task that is highly rated for SML. This indicates a potential for reorganization. We measure the within-occupation standard deviation of SML as one proxy measure for the "suitability for job re-design" (sdSML). Table 3 below has the results from our first evaluation run of the SML rubric.

TABLE 3 — LOWEST AND HIGHEST 15 STANDARD DEVIATION SML SCORE OCCUPATIONS

Rank	Lowest SD SML Occupations	sdSML	Highest SD Occupations	sdSML

[16] See e.g. Duplex: https://ai.googleblog.com/2018/05/duplex-ai-system-for-natural-conversation.html
[17] See for instance the discussion of Autodesk software in McAfee and Brynjolfsson (2017)

1	Dredge Operators	0.31	Public Address System and Other Announcers	0.88
2	Fabric Menders, Except Garment	0.32	Neuropsychologists and Clinical Neuropsychologists	0.81
3	Insulation Workers, Floor, Ceiling, and Wall	0.37	Archeologists	0.80
4	Court Reporters	0.38	Farm and Home Management Advisors	0.80
5	Music Composers and Arrangers	0.38	Farm Labor Contractors	0.79
6	Carpet Installers	0.39	Radio and Television Announcers	0.79
7	Umpires, Referees, and Other Sports Officials	0.39	Eligibility Interviewers, Government Programs	0.79
8	Agricultural Inspectors	0.39	Statistical Assistants	0.79
9	Motorcycle Mechanics	0.40	Human Resources Assistants, Except Payroll and Timekeeping	0.78
10	Gaming Surveillance Officers and Gaming Investigators	0.40	Financial Examiners	0.77
11	Marking Clerks	0.41	Electronic Drafters	0.77
12	Medical Transcriptionists	0.42	Epidemiologists	0.77
13	Counter and Rental Clerks	0.43	Payroll and Timekeeping Clerks	0.77
14	Water and Wastewater Treatment Plant and System Operators	0.43	Layout Workers, Metal and Plastic	0.76
15	Biofuels Processing Technicians	0.43	Oral and Maxillofacial Surgeons	0.75

In Table 3 we can see where there might be opportunity. PA system announcers have high variability in the SML values of their constituent tasks, as do archeologists and statistical assistants. Also, of note is that oral surgeons have high variability in their task-level SML. These occupations, where it is difficult to imagine full automation by ML, may instead be candidate examples of where the effects of complementarities or incentives to re-bundle tasks will be strong instead. On the other hand, marking clerks, medical transcriptionists, and court reporters all have low variability in their task-level SML values.

Were the constituent tasks of these jobs to be high SML, these would be possible candidates for occupations where we might expect automation substitution effects to dominate. But it seems that few occupations, if any, can be entirely automated away by ML technologies. Figure 2 shows a histogram of the proportion of tasks within an occupation for which SML exceeds the 90th and 50th percentiles. Some occupations have many tasks above the 50th percentile, but none have over 80% of their constituent tasks exceeding the 90th percentile of task SML.

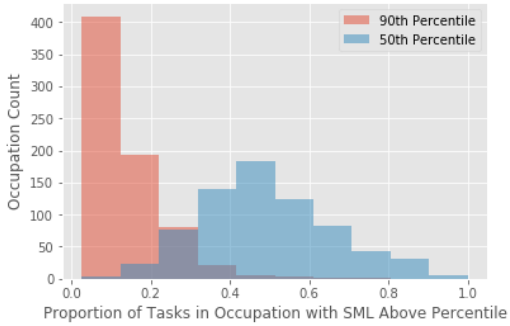


Figure 2 – Histogram of Occupations by Proportion of Tasks with SML Larger than 90th and 50th Percentile Thresholds

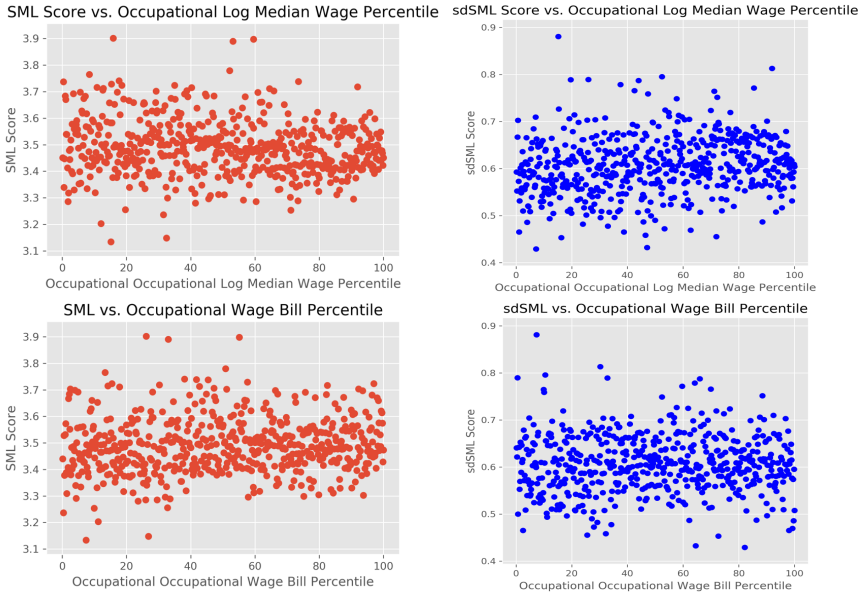
Strikingly, our analysis suggests that ML is unlikely to have the same effects on wages and inequality as earlier waves of automation. This can be seen by examining the correlation between SML and wages. Figures 3 and 4 show the scatterplots of wage percentile and 2016 total wage bill percentile with respect to SML and sdSML scores. The correlations between SML and sdSML with wages are low but statistically significant (-0.14** and 0.17** respectively)[18], suggesting that ML can potentially affect occupations at any skill level. The correlation of SML measures with the total occupational wage bill (wages times total employment in that occupation) is also close to zero (0.099** and 0.002 respectively for SML and sdSML). We see that high wage occupations are slightly negatively correlated with SML and positively correlated with sdSML; total wage bill is also slightly positively correlated with SML. There are not large differences across all percentiles.

Of course, these weak correlations we find are the case *before* the technology has been extensively implemented. To the extent technological change is directed, or more precisely, to the extent the adoption and implementation of technology is directed, there will be both wage and market size effects driving incentives to automate (Acemoglu 2002; Acemoglu and Restrepo 2017b). In that case, the occupations near the 100th percentile for total wage bill with high SML scores might be some of the most attractive to automate with ML applications. These jobs, like cashiers and accounting clerks, might later drive job polarization trends similar to those in Autor and Dorn (2013). Such trends would only be observable after the ML technologies are adopted.

This also means that our *ex ante* measures can be useful for testing theories of directed technological change. If technological change and automation are mostly directable by firms, high SML tasks in high wage (or wage bill) occupations could be more attractive automation candidates.

[18] ** denotes statistical significance at the 5% level.

In this early stage, sdSML (the standard deviation of task-level SML within a task) is neither correlated with wages nor wage bills. But we would expect if technological change is highly directable that high sdSML jobs would be attractive to re-design into new task bundles. This could lead to greater year-to-year variation in wages for the occupations due to the variety of aforementioned factors that affect labor demand. Our new method of measuring the technological content of occupations facilitates testing these kinds of hypotheses. Future studies could use the SRBS and SRS rubrics to test how directable technological change is in technologies that have already diffused, while the SML scores will be more useful for that kind of study once ML investment reaches its steady state.



Figures 3 and 4 – Scatterplot of SML Scores and sdSML Scores vs. 2016 Log Median Wage Percentile and Occupational Wage Bill Percentile

It is further worth noting that the SML measures differ from the occupational automation scores already present in the O*NET database. Figure 5 shows the scatterplot of the O*NET automation measure against the total BLS employment for 2014. Some occupations which O*NET would suggest have low levels of automation, like truck drivers, are some of the occupations for which there are some tasks with high SML scores. Further work with the results for the suitability for rules-based software scores and suitability for robotic systems scores will investigate further how much of the O*NET automation rankings we can explain using the composite scores from our three rubrics.

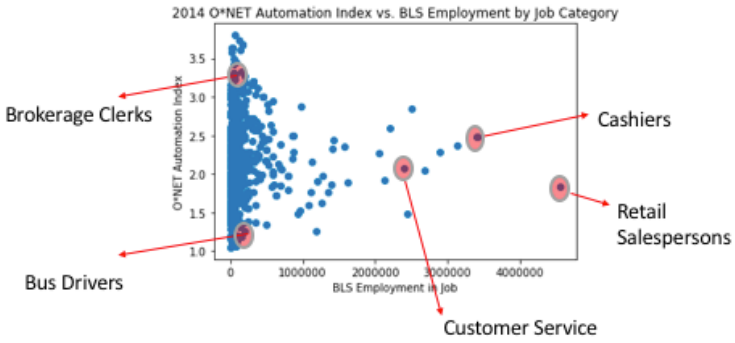
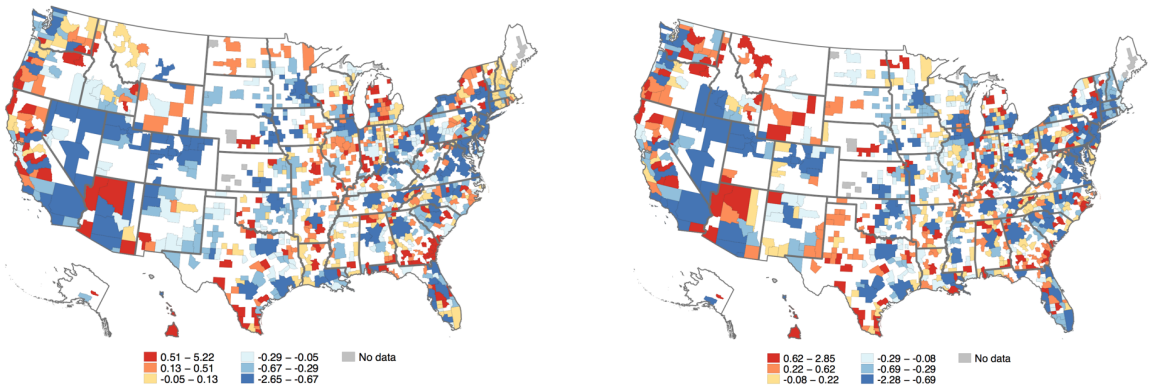


Figure 5 – O*NET Automation Scores

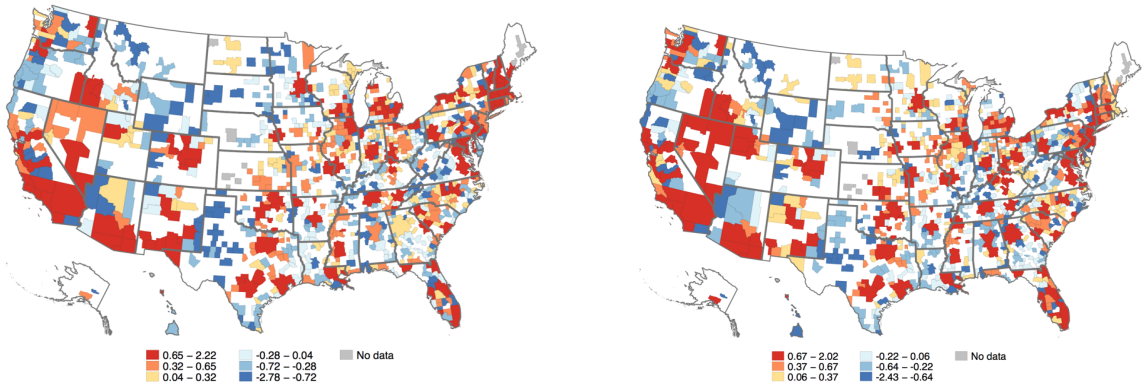
The SML scores vary by region as well, as we might expect from the economic specialization of different regions of the U.S. We join the SML scores to the regional employment data from the BLS Occupational Employment Survey and then build composite SML scores as the employment weighted average by region. Figures 6 and 7 show the standardized SML scores and standardized sdSML scores (respectively) by core-based statistical area (CBSA). We have a few missing CBSA values, which are grey in the map. Large cities tend to have lower average SML scores across all of their occupations, while rural areas are more heterogeneous. CBSAs in the upper Midwest and parts of New England tend to have higher SML scores on average with lower sdSML scores, indicating a relative preponderance of occupations with higher machine learning suitability and less potential for reorganization. In contrast, parts of Texas, Arizona, and the Pacific Northwest have high SML scores on average with high sdSML scores. These trends differ somewhat from the overall automation risks discussed in (Frank et al. 2017).

We can also look into the separate components of the SML scores and see how they are distributed by region. This will be important if various some types of ML flourish sooner than others. Figures 9 and 10 show the regional distribution of the standardized text data and image data scores (respectively) that serve as components in our overall data measure described above. Here we see a starkly different picture to the SML and sdSML regional distributions. The sources of tasks with high intensity in text and image/video data are concentrated in urban centers. The two sets of regional plots are almost mirror images of each other, suggesting that the gains to ML implementation may benefit some regions while affecting labor demand in others. One prediction of our model is that if one type of ML data, e.g. image recognition, starts affecting occupations and regions that score high on our rubric, it will soon start affecting other occupations and regions that

also score high, even if they distant geography or in very different industries. Future work will investigate the regional dynamics of occupational SML in greater detail and will incorporate the SRBS and SRS rubric results.



Figures 7 and 8 – SML and sdSML scores by Core-Based Statistical Area (CBSA)



Figures 9 and 10 – Text and Image/Video Data Intensity by Core-Based Statistical Area (CBSA)

The incentives to invest in ML, while pervasive throughout the economy, will vary by industry. As in ALM, the factor shares and elasticities of substitution of tasks will influence capital investment and employment decisions. This is the case both within occupations and between occupations in the same firm-level production function. Looking at occupation-level SML averages by industry, we see uneven potential impact of ML. Figure 11 shows 2016 employment-weighted industry SML scores for 2-digit NAICS Codes. Here we see some early evidence for the kind of potential job polarization that machine learning investments might generate. Industries with many clerks and knowledge workers like the finance, insurance, retail, real estate, and information industries have relatively high SML scores. Forestry, mining and oil and gas, accommodation and food services, and manufacturing have lower SML scores. At this point, we cannot how labor demand will change by industry. It is too soon, for instance, to tell whether the complementary applications of machine learning will dominate the potential for labor-saving automation in finance.

Still, the SML gradient by industry now suggests that the transformative effects of ML will in the long-run will not be evenly spread. The differential returns to ML investment may have important productivity implications in the long-run. Many of the low SML sectors are services industries, suggesting that ML may increase the rate at which less productive industries grow as a proportion of total factor inputs. This is further explored in Baumol (1967) and Aghion, Jones, and Jones (2017).

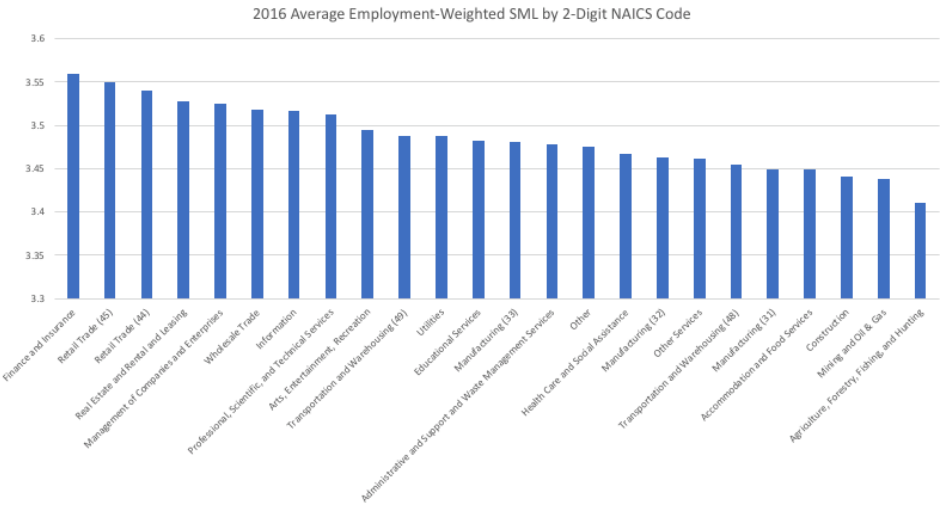


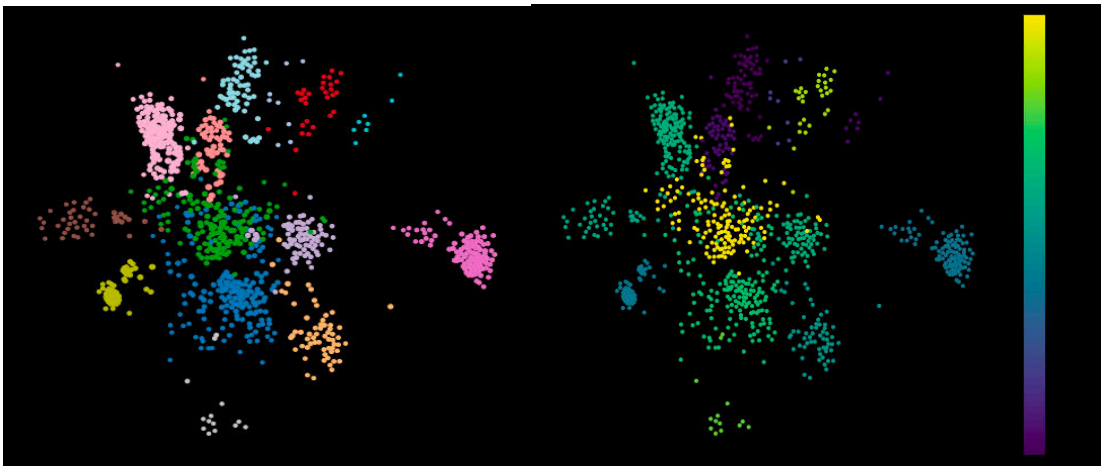
Figure 11 – Employment-Weighted Average SML by 2-Digit NAICS Industry

Augmenting the regional and industrial assessments, the fact that detailed work activities in the O*NET taxonomy are shared between occupations means we can use network analysis to understand the technological content of occupations as well. Each occupation in O*NET is connected to the others via the activities they share. This network is highly modular, meaning that community detection algorithms can easily sort different occupation nodes into clusters based on the activity edges (De Meo et al. 2011). Figure 12 shows the network of occupations grouped by the communities that the Louvain clustering algorithm detected. While these groupings have no official "category", the green cluster, for example, contains many clerical jobs and the pink cluster contains many jobs in the medical field. Applying the SML scores to the clustering, we see some clusters (like the green clerical one) have high SML on average. Others, like the pink cluster, have relatively lower SML. This SML heatmap on the same Job-Job network as in Figure 12 is represented in Figure 13. In Figure 13, high SML scores are in yellow and low SML scores are in dark purple.

It is perhaps of greater interest to represent this network differently, where nodes are the work activities and the edges are the occupations instead. This is useful to see if the SML characteristics cluster too; that provides a useful lens with which to consider the impact of ML as

well as a type of validation of the approach. If the high SML activities are often grouped into the same communities, then we have one type of indication that the dataset is internally consistent. That is, occupations with similar types of activities have similar exposure to machine learning (again, without differentiating between complements and substitutes for human labor). Figure 14 shows the network of activities.

This network is also highly modular. Many of the highest SML activities are in the red cluster, which contains clerical tasks. Ongoing work will further investigate occupational connections in activities and how they might relate to the diffusion of technology.



Figures 12 and 13 – The Job-Job Network Colored by Community (Left) and by SML Score (Right) High (Low) Average SML is pictured in Yellow (Dark Purple)

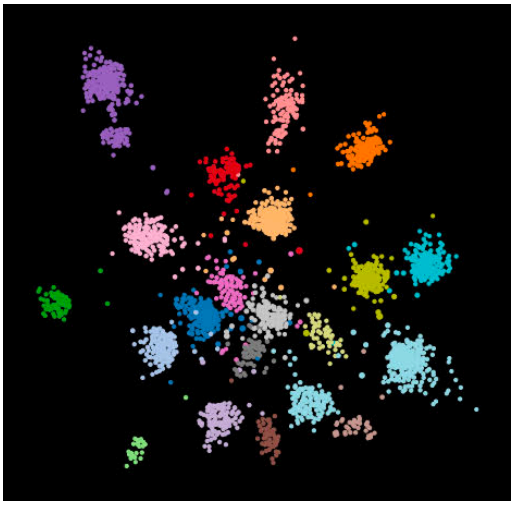


Figure 14 – The Activity-Activity Network (Connected by Occupations) Colored by Activity Community (Highest SML Community in Red – Clerical Activities)

Conclusion

The principal takeaways from all of our empirical analyses are that 1) few occupations can be completely automated using machine learning technology, 2) nearly all occupations have at least one task which is suitable for machine learning, and 3) these circumstances suggest there will be economic gains to re-designing jobs into new bundles of tasks.

There are a number of important caveats to our rubric-based analysis. Most importantly, we leave the most important work of inventing new business processes, re-organizing work, and thinking of new occupational tasks to the entrepreneurs, managers, and researchers of the future. This work focuses narrowly on what is already going on in the economy; it is therefore necessarily constrained to silence on change to occur in longer than just a few years. The rubric(s) can and should be updated if the technology changes (for example, if quantum computing becomes viable).

The emphasis on *technological* possibility also belies the myriad *social* facilitators and obstacles for adopting and implementing a GPT. In the case of ML, there are many social, legal, and ethical concerns which must be resolved (Bolukbasi et al. 2016; Bonnefon, Shariff, and Rahwan 2016; Remus and Levy 2017; Kleinberg, Mullainathan, and Raghavan 2016). Our rubric does not grapple with algorithmic bias, legal challenges, or other policy concerns in ML. So far, the analysis has concentrated on ML as well. Nevertheless, the rubric structure, as a new technologically-focused tool, provides one possible roadmap for social scientists seeking to predict where future GPTs will have the greatest effects.

[Note: We look forward to implementing the SRBS and SRS rubrics as well as building out the larger, improved SML study over the summer of 2018 to fill in the possible gaps in understanding automation that result from a narrow lens.[19]]

[19] One obvious improvement from using the SRS rubric, for instance, will be incorporation of physical motion activities into the set of tasks that could be automated. Currently, the ML rubric considers physical dexterity as an impediment to automation.

References

Abbeel, Pieter, and Andrew Y. Ng. 2004. "Apprenticeship Learning via Inverse Reinforcement Learning." In *Twenty-First International Conference on Machine Learning - ICML '04*, 1.

Acemoglu, Daron. 2002. "Directed Technical Change." *Review of Economic Studies* 69 (4):781–809.

Acemoglu, Daron, and David Autor. 2011. *Skills, Tasks and Technologies: Implications for Employment and Earnings. Handbook of Labor Economics*. Vol. 4.

Acemoglu, Daron, and Pascual Restrepo. 2017a. "Robots and Jobs: Evidence from US Labor Markets." *MIT Working Paper*.

———. 2017b. "The Race between Machine and Man: Implications of Technology for Growth, Factor Shares and Employment." *SSRN Electronic Journal*.

Aghion, Philippe, Benjamin F Jones, and C Jones. 2017. "Artificial Intelligence and Economic Growth." *National Bureau of Economic Research*, no. w23928.

Agrawal, Ajay, Joshua S Gans, and Avi Goldfarb. 2017. "Prediction, Judgment, and Uncertainty." In *Economics of Artificial Intelligence*. University of Chicago Press.

———. 2018. *Prediction Machines: The Simple Economics of Artificial Intelligence*. Boston: Harvard Business Review Press.

Agrawal, Ajay, John McHale, and Alex Oettl. 2018. "Finding Needles in Haystacks: Artificial Intelligence and Recombinant Growth." In *Economics of Artificial Intelligence*. University of Chicago Press. http://www.nber.org/papers/w24541.

Autor, David H. 2014. "Polanyi's Paradox and the Shape of Employment Growth." *Working Paper*, 129–78.

———. 2015. "Why Are There Still So Many Jobs? The History and Future of Workplace Automation." *Journal of Economic Perspectives* 29 (3):3–30.

Autor, David H., and David Dorn. 2013. "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market." *American Economic Review* 103 (5):1553–97.

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

Baumol, William J. 1967. "Macroeconomics of Unbalanced Growth: The Anatomy of an Urban Crisis." *American Economic Review* 57 (3):415–26.

Benzell, Seth G., and Erik Brynjolfsson. 2018. "Digital Abundance and Scarce Genius: Implications for Interest Rates and Wages." *Unpublished Working Paper. MIT*.

Benzell, Seth G, Laurence J Kotlikoff, Guillermo Lagarda, and Jeffrey D Sachs. 2016. "Robots Are Us: Some Economics of Human Replacement." *NBER Working Paper Series*, no. Working Paper 20941.

Bessen, James. 2018. "AI and Jobs: The Role of Demand." In *Economics of Artificial Intelligence*. University of Chicago Press.

Bessen, James E. 2016. "How Computer Automation Affects Occupations: Technology, Jobs, and Skills." *SSRN Electronic Journal* 16 (2):293–321.

Bolukbasi, Tolga, Kai-Wei Chang, James Zou, Venkatesh Saligrama, and Adam Kalai. 2016. "Man Is to Computer Programmer as Woman Is to Homemaker? Debiasing Word Embeddings," 1–25.

Bonnefon, Jean-François, Azim Shariff, and Iyad Rahwan. 2016. "The Social Dilemma of Autonomous Vehicles." *Science* 352 (6293):1573–76.

Bresnahan, Timothy. 2010. "General Purpose Technologies." *Handbook of the Economics of Innovation* 2 (1):761–91.

Bresnahan, Timothy F., and M. Trajtenberg. 1995. "General Purpose Technologies 'Engines of Growth'?" *Journal of Econometrics* 65 (1):83–108.

Bresnahan, Timothy F, Shane Greenstein, David Brownstone, and Ken Flamm. 1996. "Technical Progress and Co-Invention in Computing and in the Uses of Computers." *Brookings Papers on Economic Activity: Microeconomics*, 1–83.

Brynjolfsson, Erik, and Tom Mitchell. 2017. "What Can Machine Learning Do? Workforce Implications." *Science* 358 (6370):1530–34.

Brynjolfsson, Erik, Tom Mitchell, and Daniel Rock. 2018. "What Can Machines Learn, and What Does It Mean for Occupations and the Economy?" *AEA Papers and Proceedings*, 43–47.

Brynjolfsson, Erik, Daniel Rock, and Chad Syverson. 2017. "Artificial Intelligence and the Modern Productivity Paradox: A Clash of Expectations and Statistics." In *Economics of Artificial Intelligence*. University of Chicago Press.

Bughin, Jacques, James Manyika, Jonathan Woetzel, Frank Michael Mattern, Susan Chui, Anu Lund, Sree Madgavkar, et al. 2017. "A Future That Works: Automation, Employment, and Productivity." *McKinsey Global Institute*, no. January:1–28.

Cockburn, Iain M, Rebecca Henderson, and Scott Stern. 2018. "The Impact of Artificial Intelligence on Innovation: An Exploratory Analysis." *National Bureau of Economic Research* No. w24449 (September).

Esteva, Andre, Brett Kuprel, Roberto A. Novoa, Justin Ko, Susan M. Swetter, Helen M. Blau, and

Sebastian Thrun. 2017. "Dermatologist-Level Classification of Skin Cancer with Deep Neural Networks." *Nature* 542 (7639):115–18.

Felten, Edward W., Manav Raj, and Robert Seamans. 2018. "A Method to Link Advances in Artificial Intelligence to Occupational Abilities." *AEA Papers and Proceedings* 108:54–57.

Finn, Chelsea, Paul Christiano, Pieter Abbeel, and Sergey Levine. 2016. "A Connection between Generative Adversarial Networks, Inverse Reinforcement Learning, and Energy-Based Models."

Frank, Morgan R., Lijun Sun, Manuel Cebrian, Hyejin Youn, and Iyad Rahwan. 2017. "Small Cities Face Greater Impact from Automation."

Furman, Jason, and Robert Seamans. 2018. "AI and the Economy," 1–33. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3186591.

Goodfellow, Ian, Bengio, Yoshua, Courville, Aaron. 2016. "Deep Learning." *MIT Press*.

Goolsbee, Austan. 2018. "Public Policy in an AI Economy." In *Economics of Artificial Intelligence*. University of Chicago Press.

Graves, A, A.-R. Mohamed, and G Hinton. 2013. "Speech Recognition with Deep Recurrent Neural Networks." *2013 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, no. 6:6645–49.

Greenstein, Shane, and Jeff Prince. 2005. *The Diffusion of the Internet and the Geography of the Digital Divide in the United States*. *Applied Economics*.

Guyot, Alexandre, Laurence Hubert-Moy, and Thierry Lorho. 2018. "Detecting Neolithic Burial Mounds from LiDAR-Derived Elevation Data Using a Multi-Scale Approach and Machine Learning Techniques." *Remote Sensing* 10 (2).

Hinton, Geoff, Yoshua Bengio, and Yann Lecun. 2015. "Deep Learning Tutorial." *Nips*.

Holmstrom, Bengt, and Paul Milgrom. 1991. "Multitask Principal-Agent Analyses : Incentive Contracts, Asset Ownership, and Job Design." *Journal of Law, Economics, & Organization* 7:24–52.

Jorgenson, Dale W. 1963. "Capital Theory and Investment Behavior." *American Economic Review* 53 (2):247–59.

Kaldor, Nicholas. 1961. "Capital Accumulation and Economic Growth." In *The Theory of Capital: Proceedings of a Conference Held by the International Economic Association*, 177–222.

Kleinberg, Jon, Sendhil Mullainathan, and Manish Raghavan. 2016. "Inherent Trade-Offs in the Fair Determination of Risk Scores," 1–23.

LeCun, Yann, Léon Bottou, Yoshua Bengio, and Patrick Haffner. 1998. "Gradient-Based Learning

Applied to Document Recognition." *Proceedings of the IEEE* 86 (11):2278–2323.

Levy, Frank. 2017. "Computers and Populism: Artificial Intelligence, Jobs, and Politics in the Near Term." *Oxford Review of Economic Policy*, Forthcoming.

Litjens, Geert, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen A.W.M. van der Laak, Bram van Ginneken, and Clara I. Sánchez. 2017. "A Survey on Deep Learning in Medical Image Analysis." *Medical Image Analysis*.

Loos, Sarah, Geoffrey Irving, Christian Szegedy, and Cezary Kaliszyk. 2017. "Deep Network Guided Proof Search," 1–20. http://arxiv.org/abs/1701.06972.

McAfee, Andrew, and Erik Brynjolfsson. 2017. *Machine, Platform, Crowd : Harnessing Our Digital Future*. *Harnessing Our Digital Future*.

Meo, Pasquale De, Emilio Ferrara, Giacomo Fiumara, and Alessandro Provetti. 2011. "Generalized Louvain Method for Community Detection in Large Networks." In *International Conference on Intelligent Systems Design and Applications, ISDA*, 88–93.

Milgrom, Paul, and John Roberts. 1996. "The LeChatelier Principle" 86 (1):173–79.

Pickering, Andrew. 1993. "The Mangle of Practice: Agency and Emergence in the Sociology of Science." *American Journal of Sociology* 99 (3):559–89.

Polanyi, Michael. 1966. "The Logic of Tacit Inference." *Philosophy* 41 (155):1–18.

Remus, D, and F Levy. 2017. "Can Robots Be Lawyers? Coomputers, Lawyers, and the Practice of Law." *Geo. J. Legal Ethics* 30:501.

Rumelhart, David E., Geoffrey E. Hinton, and Ronald J. Williams. 1986. "Learning Representations by Back-Propagating Errors." *Nature* 323 (6088):533–36.

Shokri, Reza, and Vitaly Shmatikov. 2016. "Privacy-Preserving Deep Learning." In *2015 53rd Annual Allerton Conference on Communication, Control, and Computing, Allerton 2015*, 909–10.

Silver, David, Thomas Hubert, Julian Schrittwieser, Ioannis Antonoglou, Matthew Lai, Arthur Guez, Marc Lanctot, et al. 2017. "Mastering Chess and Shogi by Self-Play with a General Reinforcement Learning Algorithm," 1–19.

Trajtenberg, Manuel. 2018. "AI as the Next GPT: A Political-Economy Perspective." 24245. Working Paper Series.

White, B. W., and Frank Rosenblatt. 1963. "Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms." *The American Journal of Psychology* 76 (4):705.

Appendix A – Machine Learning Paradigms

Artificial Intelligence as a field includes approaches to problem solving like symbolic manipulation, knowledge-based systems, and cybernetics in addition to ML. Within ML, there are three main paradigms which are useful for solving different types of problems:

Supervised Learning — This is the area of machine learning that currently has the most widespread commercial use. Algorithms are designed to learn a mapping function between input and output data where both the inputs and the outputs are known (and paired together). Classification and regression are common objectives. Standard deep and convolutional neural nets fall under this category.

Reinforcement Learning — A class of machine learning algorithms which seek to learn an optimal policy function given a set of possible actions and representation of the environment. In training, actions are chosen, the environment changes, and a reward is assigned. The goal of the algorithm is to learn an action policy conditioned on the environmental state that maximizes the (discounted) sum of rewards over time. Reinforcement Learning techniques can often be combined with deep learning methods.

Unsupervised Learning — A class of machine learning algorithms where structure in a set of inputs is computed without having an explicit output target to predict or classify. Examples of applications include clustering and dimension reduction.

Appendix B – The Suitability for Machine Learning Rubric

1. Information needed to complete the task (inputs) and outputs can be explicitly specified in machine-readable format

1: It is very difficult or impossible to identify particular inputs and outputs (e.g. emotions, ideas, impressions)
3: It is possible to create rankings or partial representability of inputs and outputs
5: It is easy to quantify results on a machine/computer (e.g. calculations, concrete inputs and outputs)

2. Task information is recorded or recordable by computer

1: Input and training outputs in separate, incompatible databases, or not online at all (e.g. no dataset exists at all)
3: Some inputs and training outputs online/can be observed by computer (e.g. sensors can be installed at low cost to generate necessary data - includes places where images, video, or sound data can be recorded but may not be yet).
5: Inputs and training outputs already online in consistent format (e.g. queries can be made to a database to get necessary information)

3. Task does not require a wide range of complex outputs (mental and/or physical)

1: Task requires a range of different mental ***or*** *physical abilities (e.g. performing a difficult surgery, completing an obstacle course, acting in a film, creating a plan, negotiating)*
3: Tasks requires moderate variety in responding and acting (e.g. walking a dog, supervising a construction site, inspecting a building)
5: Task requires very little variety in thinking or action/movement (e.g. sorting mail, doing taxes, labeling an image)

4. Task feedback (on the success of outputs) is immediate

1: Feedback is never received or takes a long very long time (e.g. predictions may take months or years until accuracy can be proven, it takes years often, to determine the success of writing a business plan)

3: Feedback is received but response time is inconsistent/unclear and/or unclear on if it is beneficial to progress (e.g. teaching quality can be measured through standardized tests and observations, which are neither immediate nor long-term)
5: Instant results/feedback when the action is executed (e.g. calculations can be immediately categorized as either correct or incorrect)

5. The task output is error tolerant

1: A mistake could lead to serious harm, injury, or death to those involved, or could lead to lasting negative consequences (e.g. mistake during surgery, mistake at a nuclear facility)
3: A mistake will have negative consequences, but can fixed with some work (e.g. a construction mistake, or human resources slip-up will be noticed and reprimanded, but would not result in termination of employment or injury)
5: A mistake can be easily fixed, and holds few, if any negative consequences (e.g. a slip up in factory work or mail sorting mistakes could go potentially unnoticed)

6. It is not important that outputs are perceived to come from a human

1: Task fundamentally requires human connection (e.g. therapy, teaching, making a speech, delivering bad news like a diagnosis)
3: Task could be done by a non-human, but might cause frustration or inefficiency (e.g. customer service)
5: Task requires little human connection, empathy, or emotional intelligence (e.g. telemarketing, preparing taxes, performing calculations, lifting boxes)

7. Task does not require complex, abstract reasoning

1: Task requires intuition or highly involved reasoning (e.g. writing a plan, administering therapy, coming up with a research proposal/plan, teaching)
3: Task requires some reasoning, but can mostly be broken into well-defined rules (e.g. playing chess, sorting mail)
5: Task is mainly perception and does not require complex reasoning skills (i.e. can be done in less than 1 second)

8. Task is principally concerned with matching information to concepts, labels, predictions, or actions

1: The task does not have clear, consistent inputs or outputs (e.g. open-ended travel vacation)
3: The task potentially has well-defined inputs and outputs, but does not require mapping of the two (e.g. middle management, camp counselor, kindergarten teacher)
5: The task has clear inputs and outputs. The purpose of the task is to determine how the inputs affect the outputs (e.g. translating one language to another, matching an image to words describing the image)

9. Task does not require detailed, wide-ranging conversational interaction with a customer or other person

1: Task requires explaining something deeply to another person, or having a deep conversation that cannot be predicted in advance (e.g. therapy, negotiation)
3: Task involves communicating but about a relatively small, pre-set range of topics (e.g. giving instructions or directions, answering/asking specific/common questions)
5: Task doesn't require any form of communication/conversation with another person (e.g. solving equations, lifting objects, observing)

10. Task is highly routine and repeated frequently

1: Task is not routine, and different approaches must be taken every time (e.g. negotiating contracts, fighting a fire, treating rare and specific issues)

3: Task is routine, but might be done differently each time (e.g. waiting tables, operating a multi-purpose machine, teaching a class)
5: Task is very repetitive, and must be done in the same way each time (e.g. working in an assembly line, delivering things along a route, being a cashier)

11. Task is describable with rules

1: The task has no clear, well-known set of rules on what is and is not effective (e.g. inventing a product, dealing with exceptions, painting a painting, writing a book)
3: The task may be subject to some general rules (e.g. cutting hair, making a chair, overseeing a construction site, driving a car)
5: The task can be fully described by a detailed set of rules (e.g. folding origami, following a recipe, determining loan eligibility via formula)

12. There is no need to explain decisions during task execution

1: Decisions highly impact the lives of others and require justification (e.g. persuasion, long term planning, law-making, courtroom decisions)
3: There is some need to explain decisions, particularly when people ask questions (e.g. doctors performing checkups, operating machinery)
5: There is no need to explain decisions. The task is only concerned with having the correct output and does not depend on the process through which the output is determined. (e.g. correctly predicting the weather, performing the correct calculation, optimizing allocation of resources, determining quickest route)

13. Task can be converted to answering multiple choice questions, ranking alternative options, predicting a number, or grouping similar objects

1: Task output does not have to do with any of these options (e.g. lifting objects, collecting things, making things)
3: Task output may have something to do with categorizing or identifying things but the rules or criteria are not clear in advance and must be discovered (e.g. Supervising others, making sure that a team stays on track, recommending a plan)
5: Task is focused on applying a rule or pattern, particularly related to sorting or grouping things (e.g. grading food quality, recognizing faces, diagnosing common conditions, sorting mail)

14. Long term planning is not required to successfully complete the task

1: The task is concerned with planning around a timeline of months or years (e.g. supervising research projects, constructing complex buildings, entrepreneurship, crafting long term cancer treatment plans)
3: The task is concerned with a timeline in the range of weeks or days or an indeterminate amount of time (e.g. managing others' workloads, teaching students a specific set of lessons)
5: The task involves an immediate response and isn't concerned with a future impact (e.g. answer questions in a call center, lifting objects, performing calculations)

15. The task requires working with text data or might require working with text in the future:

1: Task does not include working with any text (e.g. making a hamburger, operating machinery)
3: Task may include some light reading (e.g. reading labels, occasionally reading directions)
5: Task includes heavy text processing, reading, or writing (e.g. reading documents, writing a letter)

16. The task requires working with image/video data or might require working with image/video data in the future:

1: Task does not require looking at images or videos, or otherwise using your eyes (e.g. having a phone conversation)

3: Task may occasionally require utilizing images and video (e.g. greeting customers, knowing whether you're alone before entering a password)
5: Task requires analyzing images and videos (e.g. finding defects in products, looking at surveillance footage, classifying objects in pictures, facial recognition)

17. The task requires working with speech data or might require working with speech data in the future:
1: Task does not require listening to or communicating with speech (e.g. independent tasks such as lifting objects, repetitive assembly work)
3: Task may require occasional listening, talking, or communicating (e.g. construction work, being a cashier, financial analyst)
5: Task requires heavy speech processing, or communicating with speech (e.g. telemarketing, translating between languages, giving a speech, having a conversation)

18. The task requires working with other types of data (other than text, image/video, and speech):
1: Task does not require working with digital data in any form (e.g. making handmade art)
3: Task requires working with some types of data at a low frequency (e.g. performing restocking tasks at a grocery store, testing machines for maintenance)
5: Task requires constant interaction with machine or sensor data. (e.g. monitoring temperature/weather, analyzing pricing data, monitoring machine performance)

19. Many components of the task can be completed in a second or less
1: Task takes a long time to complete (e.g. making a plan, writing a book)
3: Task cannot be done instantly, but also does not involve much long-term planning (e.g. performing a surgery, delivering food)
5: Task can be done instantly, or can be broken up into smaller choices that can be done instantly (e.g. Identifying a picture)

20. Each instance, completion, or execution of the task is similar to the other instances in how it is done and these actions can be measured
1: Task primarily involves rare or unique situations that cannot be summarized easily with machine-readable data (e.g. making strategic decisions for a company)
3: Data can be collected but the data output structure is highly varied (e.g. performing different types of surgery will generate different kinds of feedback)
5: Data are already available or can be easily collected (e.g. customer service transcripts, text translation, image classification, stock price movements)

21. Actions in the task must be completed in a very specific order, and practicing the task to get better is easy
1: Task involves many rare-occurring or unique situations that make the task hard to practice (e.g. therapy, negotiating)
3: Task involves some components for which a reward function can be defined (e.g. shipping/receiving in a warehouse, designing mechanical components)
5: Sequences can be repeated and tested over and over again, and there are "right" moves that can be used to generate rewards (e.g. video games, industrial process optimization, investment decision-making)

-- The following 2 questions were conducted in a separate survey --

22. The task requires physical labor and movement

1: The task requires very little physical work (e.g. drafting a report, talking with a customer on the phone)
3: Some of the task requires physical work (e.g. preparing mail, watching a store as a security guard)
5: The task is mostly intensive physical work (e.g. picking up or moving objects, assembling things)

23. The task requires highly dexterous, polished, or skillful physical movement

1: Either little movement is required or required physical work can be clumsy and imprecise (e.g. packaging items in boxes, clearing areas of debris, disposing trash, drafting letters)
3: There are some precise movements required to complete the task (e.g. washing dishes, cutting meat products)
5: There are many highly precise movements required to complete the task (e.g. repairing a watch, choreographing dances, installing machinery)