# Rage against the machine

WHAT ARE OUR PROSPECTS AS ADVANCES IN AI CHANGE THE LABOUR EQUATION? SELMER BRINGSJORD AND JOE JOHNSON SEE TROUBLE AHEAD

Unemployment rates are at disturbingly high levels, even in the so-called "technologised" economies. When President Obama requested funds for "stimulating" the US economy after the economic turbulence of 2008, he and his economic advisors (who have now departed DC and returned to their comfortable universities) confidently predicted that spending this money would prevent the unemployment rate from exceeding 8%. Soon thereafter the rate exceeded not only 8%, but 10%. The fact is that only about half of the US working-age population is currently employed. Moreover, the rest of technologised world isn't exactly swimming along: Greece, Spain, Ireland; these countries are in deep economic trouble; and other countries, for example Italy, apparently aren't far behind. In

sum, things aren't great these days, economically speaking; and everyone knows it and is worried. Yet, if we are right, recent and current woes are a foretaste of much darker things to come.

Everything might somehow miraculously turn around very soon, but even if that happens, in the longer term, in light of AI, the employment picture, as far as we can see, will *ceteris paribus* turn positively dismal and dark. The reason is simply that AI will produce technology that leads to excess labour that will make today's travails seem like a walk in the park. That is, AI will produce machines able to do the jobs of most human beings, and therefore unemployment will dramatically rise. To make matters worse, as we'll briefly explore in this paper, recent developments in AI and their long-term effects are not even being acknowledged
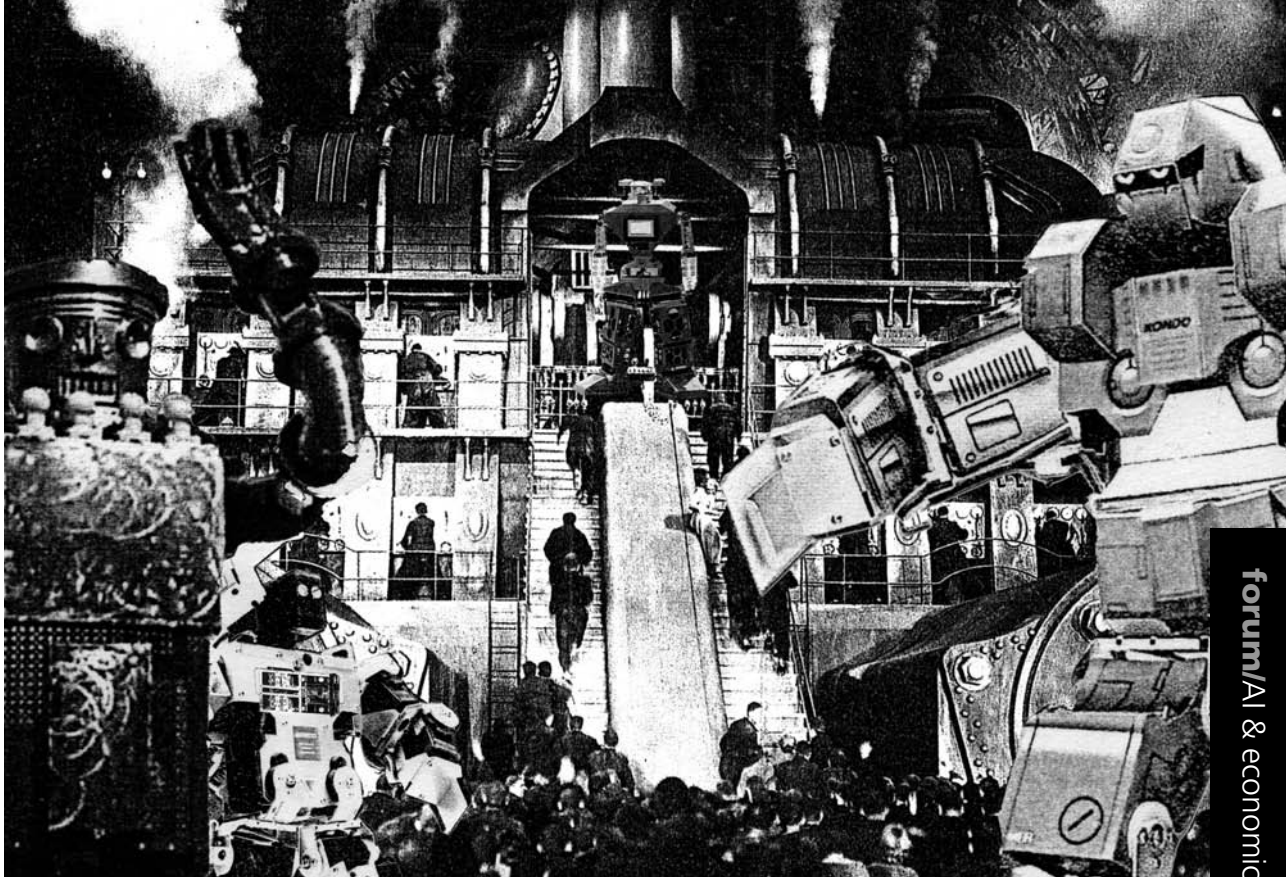
(much less addressed) by economists, and by our economic policy makers.

We are well aware of the fact that there are notoriously negative views of the future in light of AI and robotics. The co-founder of Sun Microsystems, Bill Joy, writing in *Wired*, contemplates the prospect of humanity either losing control of its future to machines or finds itself "reduced to the status of domestic animals" by them. Such views are ones we find frivolous, for they presuppose revolutionary advances in AI that may or may not come to pass.

It is of course also well known that some seemingly smart people are dead sure that "The Singularity" will soon happen. (The original argument for the sudden and world-changing arrival of ultra-intelligent machines was made by I J Good in 1965.) If they're right, unemployment could instantly go to 100%, because the brilliant computing machines that leave us in the dust might not have any use for us whatsoever. But once again, the formal and empirical evidence doesn't support the view that The Singularity will happen. It's 2012, a whole decade past the point when, according to Turing, computing machines would be linguistically indistinguishable from humans – and yet a normal toddler can make a mockery of a machine in this regard.

Our concern arises not from worries about an unlikely leap in machine intelligence, but from assuming only that AI will manage to accomplish what we can already see that it will accomplish. In fact, we don't even assume that there will be an artificial agent with general intelligence. Moreover, we'll even cheerfully pin our case to just two research-and-development projects that

are already underway, one in Japan, and one in the US.

So, Japan first. At the national level, because of its aging population and demographic calculations showing that there won't be enough young people to care for the elderly, Japan has formulated a strategy to use robots to look after the elderly in the future. There can be no ques-

Current economic woes are a foretaste of darker things to come

tion that Japan will succeed in this. We have not only seen the demonstrations, but pondered the underlying technical challenges; this looks like it will work. Now, success here implies that robots of the future will be able to excel in domains where there isn't a dearth of human workers. And if a robot can care for an elderly person, then surely a robot can be built that is able to work as a janitor. But we take it as a fact that in the foreseeable future AI can produce artificial agents and robots capable of replacing not just nurses and elder-care workers and janitors – but also social workers, primary-care physicians, and so on.

Now consider work on the driverless car in the US. As recently as 2004, this technology was relatively undeveloped, and indeed pronounced a pipe dream. However, it was widely reported, for example in the *New York Times* in November 2011, that Google has succeeded in making this concept reality, and has set a goal of having its cars log upwards of a million miles on US roadways, before building and marketing them in the US. Despite the fact that this invention may well lead to an improvement to the quality of life for

many people, particularly those who spend much of their day car-commuting, what does it mean for those who drive for a living? Any reasonably alert taxi cab driver or truck driver should be concerned.

One might think that economists and policy makers are concerned about developments like these and their implications for employment. Apparently, they aren't. There is no mention of AI whenever the topic of long-term prospects for the economy is broached. Recent work by different branches of the US Federal Government charged with maintaining balance between inflation and unemployment indicate blindness to what's happening – not by what they say, but by what they don't say. The implications of technological advancement of the sort we have in mind is not mentioned anywhere.

Academic economists are missing this trend too. In all the mainstream formal treatments of the effects of automation on employment that we are aware of, the future that we see is by definition excluded. For example, in a paper called "Technological Progress, Job Creation, and Job Destruction", published in *Review of Economic Dynamics*, Mortensen and Pissarides offer a formal model of the effects of automation on firms, but there's only one small problem. There's nothing in the model that corresponds to a machine literally replacing a human worker. It's thus not a great surprise that the future we envisage can't even be expressed in the model.

So, why is no one paying attention? One explanation is simply that the advance of smart machines is happening too fast for anyone to appreciate it properly. As noted in the case of driverless cars, these developments have all occurred within the past five to six years. But this
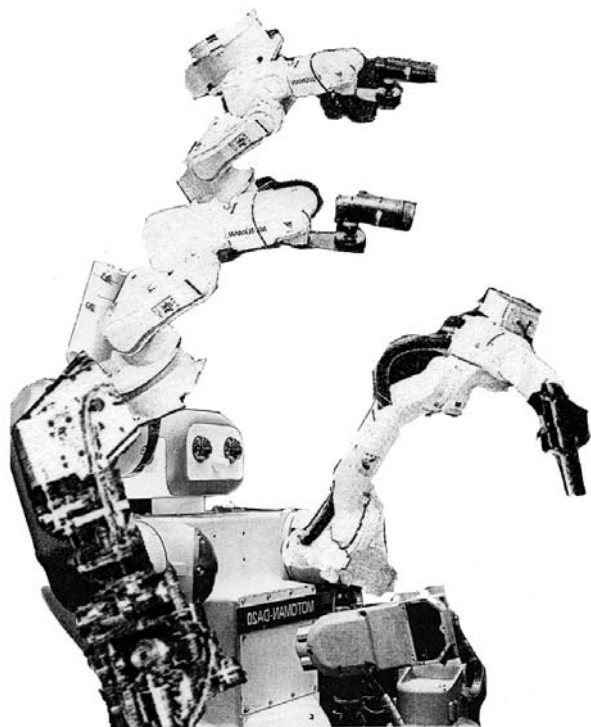
explanation is not really tenable. We've spotted the changes, and it's clear enough for anyone else to see too.

A second theory, perhaps the most likely, and one that harkens back to the blindness of professional "rigorous" economists, pertains to conventional wisdom among economists regarding automation; namely, that, in aggregate, technological advancement is always a good thing for employment. After all, automation has been with us, and steadily expanding and improving, since Aristotle first systematically considered the nature and value of commerce. And yet we see relatively flat unemployment rates. Unemployment reached 25% in the Great Depression; today, despite the rise of automation, we are well shy of that percentage. The objection might be expressed like this:

"Don't you know your history? Yes, jobs in one sector may be lost due to innovation, but new jobs will be created to support the new technology; and this often adds up to an overall gain. Moreover, the new jobs are higher skilled and therefore better paying. Consider the advent of the automobile. Luddites like you decried the effect this innovation would have on poor workers in the horse-and-buggy industry: drivers, ferriers, hay farmers, buggy producers, service people, and so on. Yes, these jobs were indeed doomed the second the Model T arrived. But these primitive jobs were soon replaced with new, plentiful ones required to produce automobiles and service them; and the new jobs, requiring expertise to operate more sophisticated machinery, brought better wages to the retrained workers. New automation should always be embraced by labour, even if it poses an immediate threat to certain labour markets in the short term. New, improved opportunities will more than offset those lost."

Before you conclude that the sceptic we have imagined here is a straw man, note that not only ivory-tower economists like Mortensen et al are blind to the future we see, but the same holds of venerated thinkers with a "real-world" grasp of economics and business. Here, for instance, is the influential philosopher, economist, and critic Henry Hazlitt, writing in 1948, on automation and employment:

"Suppose a clothing manufacturer learns of a machine that will make men's and women's overcoats for half as much labour as previously. He installs the machines and drops half his labour force. This looks at first glance like a clear loss of employment.

"But the machine itself required labour to make it; so here, as one offset, are jobs that would not otherwise have existed. At this point, it may seem, labour has suffered a net loss of employment, while it is only the manufacturer, the capitalist, who has gained.

"But it is precisely out of these extra profits that the subsequent social gains must come. The manufacturer must use these extra profits in at least one of three ways, and possibly he will use part of them in all three: (1) he will use the extra profits to expand his operations by buying more machines to make more coats; or (2) he will invest the extra profits in some other industry; or (3) he will spend the extra profits on increasing his own consumption. Whichever of these three courses he takes, he will increase employment."

The problem is that such reasoning is the weakest kind of induction. It doesn't follow from employment-boosting scenarios seen in the past that that all automation will boost employment. How can convenient cases like the automobile and the sewing machine be taken to show that in the general case automation doesn't result in a loss of employment? They can't. The only hope our optimistic opponent has is to provide some kind of across-the-board, formal rationale for his rosy view. Without this rationale, our concerns remain firmly in place.

A third and final response to our worries could be dubbed the "head-in-the-sand theory". It is assumed under this theory that economic policy makers know exactly what is going on – that advancements in AI portend a dark future for employment – but they simply choose to ignore the issue, likely due to the fact they have no idea what to do about it. This is of course the most cynical of all the possibilities, but it's not one we rule out.

If we look at socio-economic trends over the last three decades, we see some disturbing movements in income and wealth inequality. In *Race Against the Machine*, Erik Brynjolfsson and Andrew McAfee show that the US has become increasingly segregated into the haves and have-nots. For example, in 1977 the top 1% took in 9% of GDP. However, in 2007, the take for the top 1% was 23%. And, the top 0.1% took 11%. Furthermore, 50% of the nation's wealth is now controlled by the wealthiest 3% of households.

Why is the "divide" intensifying? Whatever the explanation, it's not good news. The possibility should be clear. It will become increasingly difficult for all workers, including higher-skilled ones, to justify their employment, much less pay raises, when an inexhaustible work-for-free machine can step in and replace them. So, we believe that the trend of increasing income and wealth inequality will become more pronounced as time goes on.

Eventually, history tells us, there's a tipping point beyond which the masses, with their prospects dim, will force some form of upheaval. In recent times, the "upheaval" was relatively peaceful: strife between labour and management
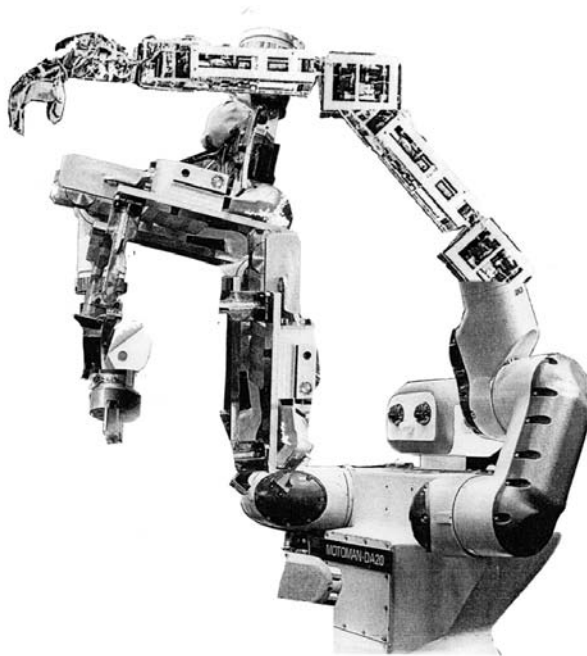
AI will produce technology
that leads to excess labour

in the end culminated in the rise of formal labour unions and government programs to dramatically redistribute wealth; for example, the introduction of federal income tax.

However, a peaceful transition might not be in the cards this time, for a number of reasons, not the least of which are the following two. First, if things progress as we envision, it won't be traditional labour, per se, that stands in opposition, but white-collar folks facing a new mechanised force. Second, technology will likely result in change, even revolution, occurring much more rapidly than past economic shifts.

Consider the Egyptian chapter of the Arab Spring, in which the citizens, particularly young, educated, male Egyptians, took to the streets and eventually forced out their country's leader, Hosni Mubarak, in response to outrage over an economic and political structure that became so stacked against them they found themselves unable to feed their families. A compelling aspect of this episode is the role the internet, specifically the globally-adopted social networking site, Facebook, had not only in helping the disenfranchised find each other and share grievances, but in facilitating organisation and collective action. This is a dimension of technology that simply didn't exist as recently as six years ago, and will likely mean those who are unhappy with their situation will be more able to push back in more organised fashion, more quickly, as they did in the Middle East.

The reader might come to the conclusion that we are making an argument against capitalism. Absolutely not. Instead, we'd like to see exactly the opposite: we'd like to see the US and other countries initiate a collective dialogue regarding how to avert the wrenching upheaval we currently foresee resulting from AI's upward march. By having the conversation now, reasonable heads can prevail.

The alternative is to wait for resentment to build while unsustainable trends continue, until public outrage prompts measures that are extreme and reactionary. The US is a country that built itself on the idea of social mobility: the idea that anyone, by working hard, can succeed. Do we have the answers to questions about how to sustain not only this idea, but how to sustain it practically? At the current time, we do not; but we are busy thinking. Regardless, the only reasonable way forward surely starts with strategising together about how to allocate the wealth created by machines, in a way that strengthens the social fabric instead of one that tears it apart.

Selmer Bringsjord is professor of logic and philosophy, cognitive science, computer science, and management and technology, and chair of cognitive science at Rensselaer Polytechnic Institute, and director of the Rensselaer AI and reasoning lab. He is the author of many papers and books, and a full bio is available at rpi.edu/~brings.

Joe Johnson was an actuary and software engineer, working for many years for financial firms and hedge funds. He has an MS in computer science from Rensselaer Polytechnic Institute, where he is pursuing a PhD on applications of AI in economics and finance.