6–7 OCTOBER 2021

REGULATING BIG TECH
CONFERENCE

Research session
Data in digital markets and money

Thursday 7 October

15:00 CEST (UTC +2)

BIS

#BISResearchConference
#RegulatingBigTech

"Harms of AI"

Daron Acemoglu
Institute Professor
Massachusetts Institute of Technology

"The Digitalization of Money"

Markus Brunnermeier
Edwards S Sanford Professor of Economics
and Director of the Bendheim Center for Finance,
Princeton University

Chair: **Jean-Charles Rochet** Full Professor, Geneva Finance Research Institute

Harms of AI

Daron Acemoglu

MIT

October, 2021.

Promises and Perils of AI

- ▶ Huge excitement about AI.
  - ▶ AI, in this context, means the current approach to machine intelligence, based on using machine learning and other statistical techniques to very large, unstructured data sets in order to learn how to perform "narrow tasks".
- ▶ No doubt that these techniques are powerful and have started to change the way we trade, work, learn, live, and communicate.
- ▶ Its enthusiasts (of whom there are many) think that AI will transform humanity for the better.
- ▶ But there are also concerns about many of its unintended (and some of its intended) consequences.
- ▶ We do not currently have enough data to adjudicate these claims empirically.
- ▶ This makes it perhaps even more valuable to think about what some of the consequences of AI will be, which is my objective.

Roadmap

- What's wrong with AI nutshell—based on its current trajectory.
- Harms in product markets.
- Harms in labor markets.
- Harms in social communication and politics.
- Futility of relying on competition or yet more innovations to solve these problems, and elements of the regulatory framework.
- Some comments on the open future of AI.

What's Wrong with AI in a Nutshell

- Every technology creates winners and losers, and some technologies greatly empower some actors in society, while disempowering others.

- These social, economic, and political implications often depend on whose vision and interests shape the technology's trajectory.

- Despite claims to the contrary, AI is an unusually centralized technology, dominated by a few companies and a very monolithic vision.

- As such, it is empowering large corporations and some governments.

- This has myriad negative economic, social, and political implications, some of which I will discuss next.

CONTROL OF INFORMATION: Data and AI

- ▶ Data is the lifeblood of AI.
- ▶ Almost universal agreement in the industry and most of academia that pooling data across users is hugely beneficial.
- ▶ Privacy concerns are recognized, but argued to be small for most people and platforms can compensate users with free services or in the future with payments.
- ▶ Does it stack up?

Data and AI (continued)

- ▶ Not necessarily.
- ▶ Whatever makes data so useful also makes it potentially harmful when shared—data is social and interlinked across users.
- ▶ Let me give an example to illustrate this point (full model, with much greater generality in Acemoglu, Makhdoumi, Malekian and Ozdaglar, 2020).
- ▶ Each user i has a privacy value v_i per unit of data—either intrinsic or in order to protect his consumer surplus—and also enjoys payments or free services from the platform.
- ▶ Normalize the value of data to the platform to 1
  - ▶ More formally this can be done in terms of mean square error reductions or entropy.
- ▶ In this example, suppose there are two users.
- ▶ The social nature of data is captured by the fact that the two users have correlated data, and suppose everything is normally distributed and the correlation coefficient between the two users $\rho > 0$.

Data and AI (continued)

- Suppose $v_1 < 1 < v_2$.
- Then the platform will always acquire user 1's data.
- But this creates negative externalities on user 2 because of the correlation, and if v_2 is large, then data sharing by user 1 can be socially harmful.
- Worse, this correlation empowers the platform.
  - Suppose $\rho \simeq 1$. Then when user 1 shares her data, she reveals almost everything about user 2. This means that user 2 has no value for protecting her own data anymore. She would be happy to sell it for very cheap.
  - But once the platform knows it can purchase user 2's data, then user 1's data becomes 'dispensable', and it can buy that one for very cheap as well. In the limit both users can be induced to sell their data for 0 price.
- With the same logic, even when $v_1 = v_2 > 1$, the platform can induce both sellers to sell for very cheap.

Data and AI: General Lessons

1. The social nature of data — enabling companies to use an individual's data for predicting others' behavior or preferences — creates externalities, which can be positive or negative. When negative externalities are important, there will tend to be too much use of data by corporations and platforms.

2. The social nature of data additionally makes each individual less willing to protect their data when others are sharing theirs. This effect ("submodularity") adds to the negative externalities, but even more importantly, it implies that data prices will be depressed and will not reflect users' value of data and/or privacy.

3. In addition to leading to excessive use of data, both of these economic forces have first-order distributional consequences: they shift surplus from users to platforms and companies.

Data and Market Power

- Another domain in which control of data can be harmful is competition.
- Consider two firms competing for users.
- One of them acquires much better data about its user base.
- This can enable it to offer better products, but also to extract more surplus from its users.
- Economists sometimes bank on competition to control such things. But better data relaxes product market competition—the firm without data cannot really discipline the data-rich firm.
- Worse, when the data-rich firm charges higher prices to its consumer base, this may increase the ability of the other firm to charge higher prices (even if it doesn't have the same ability to acquire data about its own user base).

Data and Market Power: General Lessons

1. The use of AI technologies and detailed consumer data for prediction may improve the ability of firms to customize products for consumers, potentially improving overall surplus.
2. However, it also increases the power of (some) companies over consumers.
3. This has direct distributional implications, enabling AI-intensive firms to capture more of the consumer surplus.
4. The indirect effect of the better collection and processing of data by one firm is to relax price competition in the market, increasing prices and amplifying the direct distributional effects.

AI and Behavioral Manipulation

> *"Once one accepts that individuals systematically behave in non-rational ways, it follows from an economic perspective that others will exploit those tendencies for gain."* Hanson and Kysar (1999, p. 630).

- More data may enable platforms and companies using their data to offer better products to users.
- Or it can enable them to manipulate the users.
- This is particularly true when people have behavioral biases they do not recognize (say, an inability to see future costs in full) or for "vulnerable populations".
    - e.g., Target successfully forecasting whether women are pregnant and sending them hidden ads for baby products.
- The two opposing faces of advertising have long been recognized in economics. AI supercharges these issues.

AI and Behavioral Manipulation: General Lessons

- Analysis in Acemoglu, Makhdoumi, Malekian and Ozdaglar (2021):

1. AI technologies can enable platforms to know more about consumers' preferences than they themselves do.
2. This opens the way for potential behavioral manipulation, whereby the platform can offer products that may temporarily appear as higher-quality than they truly are.
3. This type of behavioral manipulation tends to do more than just shift surplus from consumers to the platform; it also distorts the composition of consumption, creating new inefficiencies.

LABOR MARKET EFFECTS: Automation

- ▶ Many economists view any technology that increases productivity as ultimately beneficial to labor—the tide that lifts all boats.

- ▶ Not so in reality (e.g., Acemoglu and Restrepo, 2018, for theory; Acemoglu and Restrepo, 2020, for empirical evidence).

- ▶ Automation technologies, which have been central since the beginning of the Industrial Revolution, reduce the labor share and may reduce labor demand.

- ▶ They need to be accompanied with and counterbalanced by other more "human-friendly" technologies in order to contribute to wage and employment growth.

- ▶ AI is not responsible for intensive automation in the US and other industrialized nations starting around 1980. But looks like it will become a powerful sequel to it.

A Framework for Labor Market Analysis

- Imagine a single good in the economy, Y, whose production requires the combination of a measure 1 of tasks:

$$Y = \left(\int_{N-1}^{N} Y(z)^{\frac{\sigma-1}{\sigma}} \, dz \right)^{\frac{\sigma}{\sigma-1}}.$$

- Tasks $z \leq I$ can be automated given the current level of automation technology. This implies:

$$Y(z) = \begin{cases} A^L \gamma^L(z) l(z) + A^K \gamma^K(z) k(z) & \text{if } z \in [N-1, I] \\ A^L \gamma^L(z) l(z) & \text{if } z \in (I, N]. \end{cases}$$

Here $\gamma^L(z)$ and $\gamma^K(z)$ are task-specific productivities.

- Suppose also that there is a wage floor due to labor market imperfections or minimum wages:

$$w = \max \left\{ \underline{w}, \mathrm{MPL}(L) \right\},$$

where $\mathrm{MPL}(L)$ is the marginal product of labor when there is full employment at L.

Effects of Automation

- What does automation do?

$$\frac{\partial \ln \text{MPL}(L)}{\partial I} = \frac{\partial \ln Y(L, K)}{\partial I} \qquad \text{(Productivity effect)}$$

$$+ \frac{1}{\sigma} \frac{1 - s^L}{1 - \Gamma(N, I)} \frac{\partial \ln \Gamma(N, I)}{\partial I} \quad \text{(Displacement effect)}$$

  where s^L denotes the labor share and $\Gamma(N, I)$ is a measure of the *labor's importance* — tasks allocated to labor.

- In the special cases where $\sigma = 1$ or where $\gamma^K(z) = \gamma^L(z)$, we have $\Gamma(N, I) = N - I$. In general it is decreasing in I.

- The productivity effect is:

$$\frac{\partial \ln Y(L, K)}{\partial I} = \frac{1}{\sigma - 1} \left[\left(\frac{R}{A^K \gamma^K(I)} \right)^{1-\sigma} - \left(\frac{\text{MPL}(L)}{A^L \gamma^L(I)} \right)^{1-\sigma} \right].$$

- This will be very small for "so-so" technologies, which do not increase productivity much: $\underline{w}/A^L \gamma^L(I) \approx R/A^K \gamma^K(I)$

- Then automation reduces the marginal product and when the wage floor is binding, employment and also *welfare*.

AI and Automation

- Current path of AI going more in automation direction (e.g., Acemoglu, Autor, Hazell and Restrepo, 2021).
- Moreover AI does not seem to increase human productivity by much in most tasks (the reason why we discussed below).
- Then double whammy from automation for labor (Acemoglu and Restrepo, 2019).

AI and Automation: Some Preliminary Lessons

1. Automation reduces the labor share and may also reduce the (average) wage and/or employment, and this latter outcome is more likely when productivity gains from automation are small.

2. When labor market imperfections are present, automation tends to be excessive and welfare-reducing, particularly when it impacts employment negatively as well—again particularly when its productivity effects are small. Same considerations apply when there are non-market reasons for preferring high levels of employment (e.g., employed workers contribute more to their families, communities or society).

3. Because it increases the capital share and reduces the labor share and because it boosts inequality among workers, automation may also be excessive from a welfare point of view due to distributional concerns

4. If AI is used predominantly for automation, it will have similar effects to other automation technologies, and depending on its productivity effects and relevant welfare criteria, it may have a negative impact on social welfare.

AI and New Tasks

- ▶ In principle, AI can be used for producing new tasks or other human-friendly technologies.
- ▶ These could counterbalance automation.
- ▶ In the framework above, this corresponds to an increase in N:

$$\frac{\partial \ln \mathrm{MPL}(L)}{\partial N} = \frac{\partial \ln Y(L, K)}{\partial N} \qquad \text{(Productivity effect)}$$
$$+ \frac{1}{\sigma} \frac{1 - s^L}{1 - \Gamma} \frac{\partial \ln \Gamma(N, I)}{\partial N}. \quad \text{(Reinstatement effect)}$$

- ▶ The reinstatement effect is also positive, thus helping wages, employment and welfare.
- ▶ However, the current evidence is that AI is not going in the direction of generating human-friendly technologies — presumably related for the same reasons as were its origin and interests of its enthusiasts come from.

AI and New Tasks: General Lessons

1. AI could in principle be used for increasing worker productivity and expanding the set of tasks in which humans have a comparative advantage, rather than focusing mainly on automation. If it is used in this way, it may counterbalance some of the negative effects of automation on labor and may generate more positive welfare effects and beneficial distributional consequences

2. But there is no guarantee that the composition of technological change in general and the balance of AI between automation and more human-friendly activities should be optimal. In fact, there are many possible distortions, some of them economic and some of them social, encouraging excessive automation using AI.

AI and Human Judgment

- One counter argument is that AI can take over routine tasks and leave humans focus on the tasks that require greater human judgment.
- But there is not much evidence that this is taking place in the labor market at the moment.
- Conceptually, there are counterarguments.
- Particularly an issue if there are "economies of scope", so that humans learn from routine tasks how to better perform those that require judgment (e.g., knowledge of algebra and calculus help us engage in better mathematical problem-solving).
- Then excessive automation can reduce human judgment as well.

AI and Human Judgment: General Lessons

1. In addition to the costs of worker displacement discussed earlier in this section, economies of scope across tasks may create additional costs from the use of AI technologies. In particular, the deployment of AI in various cognitive tasks that do not require a high degree of human judgment and creativity may enable workers to reallocate their time towards tasks that involve judgment and creativity. But if economies of scope are important for human productivity, AI may have additional costs.

2. Cost-minimization incentives of firms may encourage them to use AI technologies in inefficient ways, when there are such economies of scope.

AI and Excessive Monitoring

- AI technologies are currently used extensively for monitoring workers—e.g., Amazon workhouses or transport and delivery industries.
- This could be a useful technological application.
- But economic theory implies that there will be incentives for excessive monitoring (e.g., Acemoglu and Newman, 2001).
- At the margin, monitoring is a *rent-shifting activity*—it enables firms to reduce worker rents.
- Then it will be used excessively from the viewpoint of social welfare.

AI and Excessive Monitoring: General Lessons

1. AI technologies also create new opportunities for improved monitoring of workers. These technologies have first-order distributional consequences, because they enable better monitoring and thus lower efficiency wages for workers.

2. Because at the margin the use of monitoring technologies transfers rents from workers to firms, monitoring will be excessive in equilibrium. By expanding monitoring opportunities, AI may thus create an additional social cost.

AI AND POLITICS: General Issues

- AI also influences communication, learning, social discourse, indust politics in a number of ways.
- There are many interesting and complex issues, but since time is short, I will be brief, and give the bottom line of four different types of channels via which AI influences politics—in each case, potentially very harmfully.

AI and Echo Chambers

> *"many or most citizens should have a range of common experiences. Without shared experiences, a heterogeneous society will have a much more difficult time in addressing social problems."* (Sunstein, 2001)
>
> *"falsehoods diffusing significantly farther, faster, deeper, and more broadly than the truth in all categories of information"* Vosoughi et al. (2018)

1. AI-based social media distorts individuals' willingness to share unreliable information. When social media creates echo chambers, individuals become less careful in inspecting news items consistent with their existing views and more willing to allow the circulation of misinformation.

2. Social media platforms that are focused on maximizing engagement have an incentive to create echo chambers (or "filter bubbles"), because inspection and interruptions of the circulation of news items with unreliable messages reduces engagement. As a result, platform incentives are diametrically opposed to social objectives.

Perils of Online Communication

- ▶ Effects of AI-based online communication platforms may be much more fundamental.
- ▶ Online communication is devoid of the context of other social interactions.

1. Bilateral, off-line communication, especially when the subject matter is political or social, relies on trust between parties. Naturally-existing trust in in-person social networks may enable this type of communication.
2. When communication is taking place online and in multi-lateral settings, such as in modern social media platforms powered by AI technologies, this type of trust-based communication becomes harder. This may favor non-political messages, such as gossip, which then drive out political communication.
3. Barrier to online communication is exacerbated when there is competition for attention, which is encouraged by the broadcast or multi-lateral nature of online communication.

Big Brother Effects

- ▶ AI is also a powerful tool in the hands of governments, both nondemocratic and democratic.
- ▶ Nondemocratic governments, such as China and Iran, are already using it to suppress civil society and any political opposition. But increasingly, it can also start corrupting democratic politics.
- ▶ General lessons:

1. AI technologies can be used for improving government monitoring against protest activities.
2. Since the threat of protests has a disciplining role on nondemocratic governments, and even on some democratic governments, the shift of power away from civil society towards governments will weaken democracy and aggravate policy distortions.

Irrelevance of Labor and Democracy

- ▶ Automation also makes labor increasingly irrelevant.
- ▶ If labor's power and society comes partly from democracy and partly from their role in workplaces, then AI can also start undermining democracy via completely new channel.

1. Automation can also generate an indirect negative impact on democracy and redistributive politics when ensuring cooperation from labor in workplaces is an important motivation for elites to make concessions to labor.
2. When automation brings only small productivity gains, it encourages the elite to reduce redistribution and make fewer democratic concessions. This will make policies less responsive to the majority's wishes and may further raise inequality.
3. Productivity benefits of automation may soften this effect as an automation-driven increase in output raises the opportunity cost of losing labor's cooperation. But, there exists a sufficiently high level of automation such that once we reach this level, labor becomes sufficiently irrelevant for production, with harmful effects on democracy, redistribution and social cohesion.

Regulating AI

- If some of these concerns are real, how can we deal with them? Two non-starters:
  1. Innovation will take care of it: no reason to believe this. Technological progress can go in wrong directions, especially when there are alternative, competing ways of using our collective knowledge.
  2. Competition will take care of it: even less believable. Some of the models, which I did not have time to present in detail, show that competition can make things worse.

- *Regulation* is key, and the exact way in which this regulation should be carried out depends on the details of the problem.

- But there are two general principles:
  1. Regulation to reduce the power of be corporations and governments is particularly useful.
  2. *Precautionary regulation principle*: more regulation if once harms are revealed it becomes harder to reverse things. Particularly true when we have irreversible automation and even more so, when AI undermines democracy.

Conclusion: Open Future of AI

- In concluding, let me emphasize one other reason why these issues are important.
- AI is a very broad technological platform.
- There has been huge amount of interest machine intelligence for almost 70 years now.
- The current path of AI is the result of very particular vision.
- For example, the vision of early pioneers such as Norbert Wiener, Douglas Engelbert, and JCR Licklider was very different much more based on machine intelligence empowering humans.
- Current regulations and countervailing powers may be necessary to redirect AI in this more socially useful and inclusive direction.