Luis Videgaray

The Race to Regulate AI

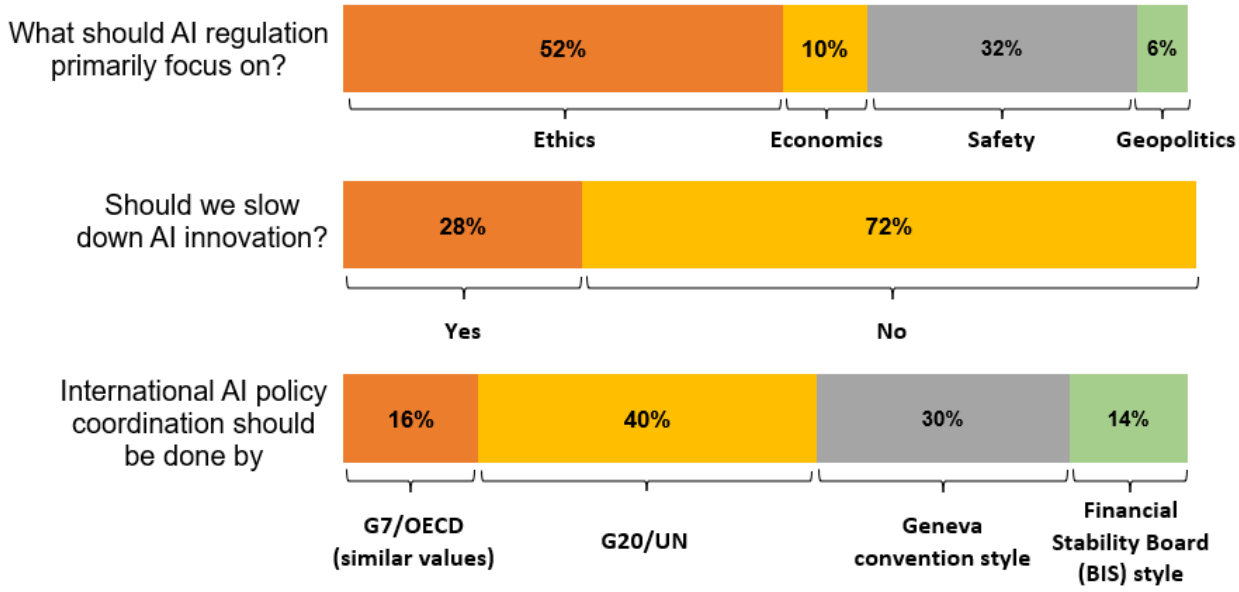
On Thursday, April 25, Luis Videgaray joined Markus' Academy for a conversation on "The Race to Regulate AI." Luis Videgaray is a Senior Lecturer at MIT Sloan and the Director of MIT AI Policy for the World Project. He is also the former Foreign Minister and Finance Minister of Mexico.

A few highlights from the discussion.

- **A summary in four bullets**
    - AI is hard to regulate because of the speed of innovation and because it is a general purpose technology, so it is difficult to envision the different use (and misuse) cases
    - When designing AI laws, policymakers should: choose between a horizontal or a vertical approach, decide whether to establish a new AI regulator, and (most critically) decide whether to regulate the inputs to models or their outcomes
    - After recognizing the difficulty of developing LLMs, emerging markets should prioritize regulating their application and customization. The priority should be to demand guarantees from LLM makers and to establish clear liability rules
    - We should be optimistic about the power of AI to improve the workings of government in emerging markets, for example by improving processes like the allocation of construction permits or the distribution of welfare benefits

- **[0:00] Markus' introduction and poll questions**
    - AI regulation is distinct for two main reasons. Firstly, as AI becomes integrated into the economy it introduces systemic risks similar to those seen in the financial sector. Secondly, the challenge of the singularity is about a self-perpetuating risk that could become existential
    - Whether a society is resilient depends on the difference between its speed of innovation and the speed of adaptability of laws and social norms. There are hence two potential policy responses: slowing down innovation or speeding up adaptability
    - Specific tools to regulate AI include moratoria, sandboxes, explainability requirements, or steering innovation to make AI safer or more labor augmenting (as Acemoglu argued in a recent episode)
    - There are two views on the industrial organization of AI. We may want to promote competition to reduce rent extraction, or we may want to control proliferation to control existential risk (similar to nuclear technology)
    - The industry can be structured in layers, with a relatively standardized first layer of base/foundation LLMs (dominated by few regulated players). Subsequent layers are more competitive and involve more specialized bespoke technology

- Should LLM makers pay content creators since they have used the existing internet to train their models? From a rent sharing perspective there should be compensation, but this may increase investment costs and reduce competition. If the goal of compensation is to incentivize the creation of more content, only new (not past) content should be remunerated



- **[10:59] The race to regulate AI**
  - A Mexican initiative at the UN to coordinate AI policy in 2017 received little interest, and the topic was thought to be not serious. Today, even if we are cognizant of the problems, nobody really knows how to best regulate AI
  - AI deployment is happening through a layered supply chain: foundation models of the first layer are becoming highly capital intensive and with high economies of scale. As a result only big tech and governments can participate in their development
  - There are 3 categories of AI policy: (1) government spending (e.g. R&D investments or subsidies), (2) using AI in government administration and in the military, (3) laws and regulations, (4) international agreements
  - AI policy has four key objectives: (1) ethics, (2) safety, (3) economics, (4) geopolitics
  - Ethics encompasses issues like privacy, discrimination, explainability, and fake content. Safety is about existential risks and the potential for AI to self-replicate. Safety issues have been gathering an increasing amount of attention despite some ethics advocates arguing they are a distraction
  - AI's four objectives are not orthogonal, creating trade-offs even within each objective. For example, privacy regulations may exacerbate AI's biases, while in economics there is a trade-off between innovation and consumer protection

    - With generative AI the US has taken the lead over China. Each country is now attempting to hinder the other's progress with measures such as restrictions on semiconductor exports and reduced academic collaboration
    - However the geopolitical divide is also between the makers and takers of AI
    - The first national AI strategies were published in 2017, and by now more than 50 countries have strategies. They are non-binding high level documents that establish priorities and goals, but they have had an impact on government spending and the adoption of AI in the public sector
    - Establishing principles for the regulation of AI is about agreeing on its desired technical attributes. Most influential was the OECD/G20's principles (2019): (1) inclusive growth, (2) human-centered values like rule of law and human rights, (3) explainability, (4) safety, and (5) accountability

- **[28:19] Designing AI Laws**
    - China was the first country to enact an AI law. It is a misconception that its lack of laws and regulations are an advantage, it is just that their laws are meant to protect a different form of government
    - The EU has established the most comprehensive regulations, ranging from its AI Act to rules on data privacy and tech competition, and inspiring countries like Brazil to adopt a similarly comprehensive approach. However the lengthy legislative process was surpassed by the speed of innovation, and the AI Act was even amended before it was enacted
    - The EU's efforts to regulate AI have not been driven by the fact that they do not have a lot of AI companies. Although the AI Act is not perfect, it is based on profound convictions around human rights and the role of the state
    - Despite some progress at the state and municipal level, it is unlikely that we will see a federal regulation in the US any time soon. Ultimately regulation will be decided by regulators (e.g. FDA, SEC, FTC) and by courts
    - When designing an AI law, the first decision is whether to have any binding rules at all: India, South Korea and Israel have explicitly stated they will not regulate AI
    - One must then choose between horizontal or vertical rules. Under a horizontal approach one develops different regulations for each industry and recognizes that the context of AI's application matters. Horizontal rules may be easier politically, but vertical ones may be more effective. One must also decide whether to establish a new separate AI regulator, as the EU has done
    - A critical decision is whether to regulate the inputs of models or their outcomes. The EU's approach has been to regulate inputs, ensuring that the data used for training models is good and balanced
    - On the other hand the Consumer Financial Protection Bureau in the US has focused on outcomes, for example ensuring there is no discrimination in loan underwriting, regardless of whether or how AI is used
    - Sometimes the outcome you want to ensure is explainability. For example preserving the right of loan applicants to know the reason for the denial of a loan

- **[37:28] Why is it so hard to regulate AI?**
    - As we have seen in the EU, it is hard to regulate AI because the innovation is fast while the lawmaking process is not. It is also hard to regulate because it is a general-purpose technology. It is very difficult to envision the different use (and misuse) cases, and even to envision what models will be like in a few years
    - Scissors are another example of a general-purpose technology. They are extremely useful but can also be used to injure someone else. There is no specific law regulating them. Instead, we regulate the outcomes of their use through measures like medical liability laws or laws against violence
    - It's important to recognize that AI is not the first human invention whose workings are not immediately understood. For another example, the battery, invented in 1800 by Volta, was not fully understood until 1860 by Maxwell
    - Emerging markets should recognize that the AI supply chain is completely global. Many countries are focused on ensuring data and computational sovereignty, but it's not clear this is efficient given the economies of scale and network effects. Instead the priority should be to demand guarantees from LLM makers and to establish clear liability rules
    - Although AI-making countries have less of an incentive to regulate at the global scale, the majority of the world would benefit since most countries are AI-takers
    - It is just difficult to get it done. For example, efforts to amend the Geneva Conventions since 2018 to regulate the use of autonomous lethal weapons systems have failed to secure an agreement

- **[47:55] Q&A**
    - After acknowledging that developing base layers (LLMs) will be very difficult due to the economies of scale, emerging markets should focus on regulating their application and customization. Social norms and preferences can be achieved through fine tuning in the second or third layers
    - We should be optimistic about the power of AI to improve the workings of government, especially in emerging markets. Examples include improving the allocation of construction permits or the distribution of welfare benefits
    - Looking to the econ. profession, everyone should be using AI to use and process data. From a policy perspective economists are not thinking about AI beyond jobs. A lot of industrial organization work is needed to understand the AI production function, as well as issues of concentration and vertical integration

Timestamps: