



# AI and

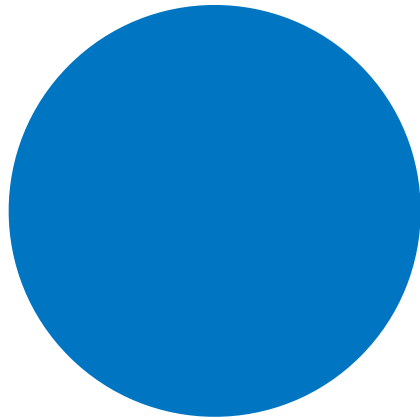
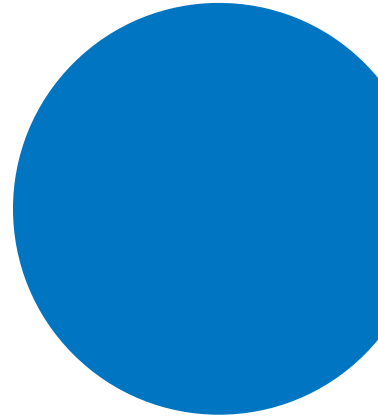
2020 SEOUL  
AI POLICY CONFERENCE  
– AI and Market Dynamics

# Market Dynamics

## Contents

---

04	<b>Preface</b>	
06	<b>Session 1</b>	<b>Challenges and Opportunities for AI Policymaking</b> Jason Schultz
14	<b>Session 2</b>	<b>Algorithmic Fairness and Anti-Discrimination Law</b> Alice Xiang
22	<b>Session 3</b>	<b>Competition in the Era of Algorithms: Evolving Law &amp; Policy</b> Niamh Dunne
32	<b>Session 4</b>	<b>Google's Perspective on Artificial Intelligence Governance</b> Charina Chou
40	<b>Session 5</b>	<b>Differential Privacy: What is it and Where is it?</b> Cynthia Dwork
50	<b>Session 6</b>	<b>Asimov for Lawyer: What Sci-Fi can(not) tell us about the future of AI regulation</b> Nicolas Petit
60	<b>Session 7</b>	<b>Blockchain Antitrust: Challenges and Opportunities</b> Thibault Schrepel
68	<b>Session 8</b>	<b>When Does Gaming Justify Algorithmic Secrecy?</b> Ignacio Cofone



# 2020 SEOUL AI POLICY CONFERENCE

## – AI and Market Dynamics

### Preface

It goes without saying that transparent and fair market operation is crucial for innovation. This is because an opaque and unfair market hinders competition and deepens the asymmetry of power, causing economic actors to lose their will to take on creative challenges. One question that can be raised here is how the emergence of a new technology called artificial intelligence (“AI”) will have an impact on market transparency and fairness. If AI is expected to negatively affect the transparency and fairness of the market, we must hurry to seek legal and policy alternatives that can help the market to function properly in the age of artificial intelligence.

With this problem in mind, the Seoul National University Artificial Initiative Policy Initiative (SNU AI Policy Initiative, “SAPI”) has held annual conferences since 2017, inviting AI experts to discuss legal and policy issues<sup>1</sup>. All the previous conferences have been produced in video formats with Korean subtitles, and from the 2nd conference in 2018 and onwards, reports in the current format have also been provided.<sup>2</sup> Moreover, the SNU AI Policy Initiative is carrying out various discussions related to AI

- 
- 1 The main topics of past conferences are as follows.  
The 1st Conference: “Policy Issues surrounding AI, Algorithms & Privacy”  
The 2nd Conference: “Artificial Intelligence Today: Governance and Accountability”  
The 3rd Conference: “AI Policy for the Future: Can We Trust AI?”
  - 2 The past conference videos can be found below.  
The 1st conference: [http://www.youtube.com/playlist?list=PLOP6ilKzhDLQ\\_a2hMmD0vxsJn0d-aQco8](http://www.youtube.com/playlist?list=PLOP6ilKzhDLQ_a2hMmD0vxsJn0d-aQco8)  
The 2nd conference and beyond: [https://www.youtube.com/channel/UCKyxSZOtLB1YvkKM2\\_Mq8gQ/featured](https://www.youtube.com/channel/UCKyxSZOtLB1YvkKM2_Mq8gQ/featured)  
The link of past conference reports is as follows: <http://sapi.co.kr/workshops/>

from a convergent perspective.<sup>3</sup> The issue of market competition is a very important topic, so some discussions on this issue have already taken place in the last conferences, but this year, it was selected as the main topic of the conference.

Unlike in previous years, there was a change in format at the 4th conference in 2020. This was because offline events were impossible due to the COVID-19 pandemic, which has grappled the world since the beginning of 2020. Accordingly, SNU AI policy initiative drastically expanded the number of sessions to eight while taking the events online, and arranged each session on a different day with at least 1 hour, so that sufficient discussions could be held throughout all sessions. At the same time, every effort was made so that discussions in all sessions could converge on the main theme of market fairness and transparency.

Despite the format change, the invited presenters of all sessions, including keynote speaker, were experts in the related fields from abroad to have a rich and in-depth discussion as in previous years. Professor Cynthia Dwork, who delivered the keynote speech, is a world-renowned scholar famous for differential privacy, a widely known technology for harmonizing privacy and data analysis. Also, all the presenters at individual sessions were composed of renowned researchers from the European Union and the United States. At this conference, researchers belonging to world-renowned organizations such as AI Now Institute and Partnership on AI participated, contributing to broadening the scope of discussion. Furthermore, a researcher from Google, a leading company in the field of artificial intelligence, delivered a presentation too, enabling the inclusion of the industry perspectives. Thus, despite being an online event that spanned several days, there were hundreds of attendees in every session.

---

<sup>3</sup> The Seoul National University AI Policy Initiative publishes issue papers twice a year, in May and November, since 2019 and holds various academic events on an irregular basis. For more information, visit <http://sapi.co.kr/>

# Challenges and Opportunities for AI Policymaking





Presenter

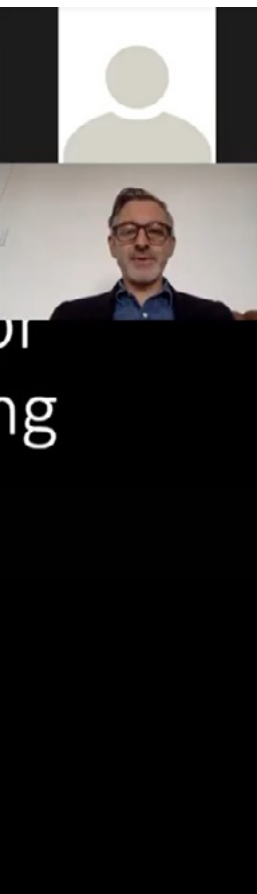
Jason Schultz,  
New York University



# Challenges and Opportunities for Artificial Intelligence Policymaking

Professor Jason M. Schultz  
September 9, 2020





One important topic surrounding the artificial intelligence technology is a legal and policy addressing of the safety and responsibility issues. Professor Jason Schultz began the seminar by presenting this problem and explaining what artificial intelligence is. He presented important artificial intelligence technologies in two main forms from a legal & policy perspective. One is an AI robot that moves in conjunction with hardware, like an automatic sorting machine used in Amazon warehouses, and the other is a question-and-answer type of AI used in the decision-making processes, such as finding directions and deciding whom to hire. Such artificial intelligence can bring great benefits to humans, but it can cause harm by solidifying social discrimination, which can both pose new challenges or opportunities from a legal and policy perspective.

As Professor Schultz explained the concepts of machine learning and deep learning among AI technologies, he presented examples such as email spam filters and image classification codes that distinguish dogs and cats. Machine learning and deep learning technologies can cause errors in classification results even if they are constantly learning. Image classification technology using a convolutional neural network (CNN) makes fatal errors due to problems such as the bias of training data in the process of classifying a specific image into a specific class. Since humans often do not know the cause of such a phenomenon, a question could be raised whether these technologies can be used for important decisions.



Session 1 Capture (caption)

Professor Schultz introduced a research conducted in China, as a representative cases where AI can cause significant problems in legal policy regarding the issues of bias and fairness. Here, based on the model learned from the facial images of criminals, the AI determined which person had a high probability of becoming a criminal by only looking at facial images. Professor Schultz pointed out that even if such an attempt may lead to effective conclusions, it can lead to a variety of legal problems, including the violation of due process or the presumption of innocence because this is just a correlation between the criminal record and some of the features in the image. Amazon's

---

facial recognition model, which adopted a similar methodology, was eventually discontinued as it became known that the model classified 28 US lawmakers have high likelihood of committing a crime.

A more realistic example is an artificial intelligence hiring system<sup>1</sup> developed by Amazon based on the employee resumes dataset. Originally, the AI hiring system was expected to greatly reduce human resource and mitigate the harm of human's arbitrary judgment. However, the system turned out to be systemically biased against female applicants. and Amazon had to withdraw the system. This result can be attributed to the characteristics of machine learning in which the model learns from past data that reflects historical social injustices.

Another example is ProPublica's 2016 report that pointed out a problem with an algorithm called COMPAS.<sup>2</sup> COMPAS is an algorithm that calculates the probability of recidivism by taking in to account variables such as criminal participation, lifestyle, personality and attitude, family and social exclusion of the defendant into scores. The algorithm then recommends to a judge whether the applicant is likely to commit crime. Although COMPAS does not include race as a variable, it was placed in the middle of a controversy when it was revealed that COMPAS tends to estimate a comparably high likelihood of recidivism for blacks and a low likelihood of recidivism for whites.

Artificial intelligence technology is also being used to classify human facial expressions as input values. Professor Schultz introduced

---

1 Dastin, J. (2018) "Amazon scraps secret AI recruiting tool that showed bias against women" Reuters. <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G>

2 Angwin, J., Larson, J., Surya, M., and Kirchner, L., (2016) "Machine Bias", ProPublica. <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>

a company that provides a service determining the suitability of applicants by analyzing interview videos during the hiring process.<sup>3</sup> This technology is believed to recognize human emotions through images or videos. Furthermore, China is introducing a system that allows the state to determine whether each citizen is a “good” citizen by creating a social credit score<sup>4</sup>. The social credit score is used for hiring, loan screening, travel permission, and other societal purposes. It analyzes each citizen’s purchase history, SNS posting history, friend list, and others. It deducts the social credit score when one parks a car illegally or criticizes the government. The problem is that AI can make errors in these various decision-making processes, and humans will not be able to fully understand how these errors were made.

Legal and policy responses to the use of artificial intelligence technology are also being formulated. For example, in Illinois, the interviewer is obliged to inform the interviewee that an evaluation using AI is being conducted and to explain how AI evaluates him or her.<sup>5</sup> Nevertheless, Professor Schultz pointed out that even if such a duty of explanation is imposed, it is unclear whether companies can comply with it due to various limitations on the explainability of AI.

Professor Schultz introduced two cases in which the problem of trade secret or explanation emerged in the dispute resolution process

---

3 HireVue, <https://www.hirevue.com/>

4 Marr, B., (2019) “*Chinese Social Credit Score: Utopian Big Data Bliss or Black Mirror on Steroids?*”, Forbes. <https://www.forbes.com/sites/bernardmarr/2019/01/21/chinese-social-credit-score-utopian-big-data-bliss-or-black-mirror-on-steroids/#5dfbf0e748b8>

5 820 ILCS 42/ Artificial Intelligence Video Interview Act

---

regarding AI algorithms.<sup>6</sup> The first case is of Tammy Dobbs, a patient with cerebral palsy living in Arkansas. Tammy Dobbs requested an explanation from Arkansas because the care time calculated by the algorithm was reduced from the care time she used to get from human examiner, but Arkansas failed to provide a clear answer. At the trial, the Arkansas state government insisted that the algorithm cannot be disclosed because it is a trade secret of its vendor, but the court ordered the disclosure of the algorithm. Despite this measure, Professor Schultz argued that compared to humans, it is difficult to understand computer code to acquire an adequate explanation for the reason for reducing care time.

The second case is a lawsuit filed by the labor union against the Office of Education on behalf of a public-school teacher who was fired as a result of an algorithmic employee evaluation. The defendant also argued that the algorithm could not be disclosed because it was a trade secret, but the court judged that if an algorithm that is a trade secret, which cannot be disclosed to the parties is used for an important personnel decision, the minimum due process has not been followed, and if so, the appropriate solution is to protect the trade secret but cancel the dismissal decision.<sup>7</sup> As such, Professor Schultz argued that relying on algorithms to make important decisions should be avoided unless the issues of fairness and transparency are resolved.

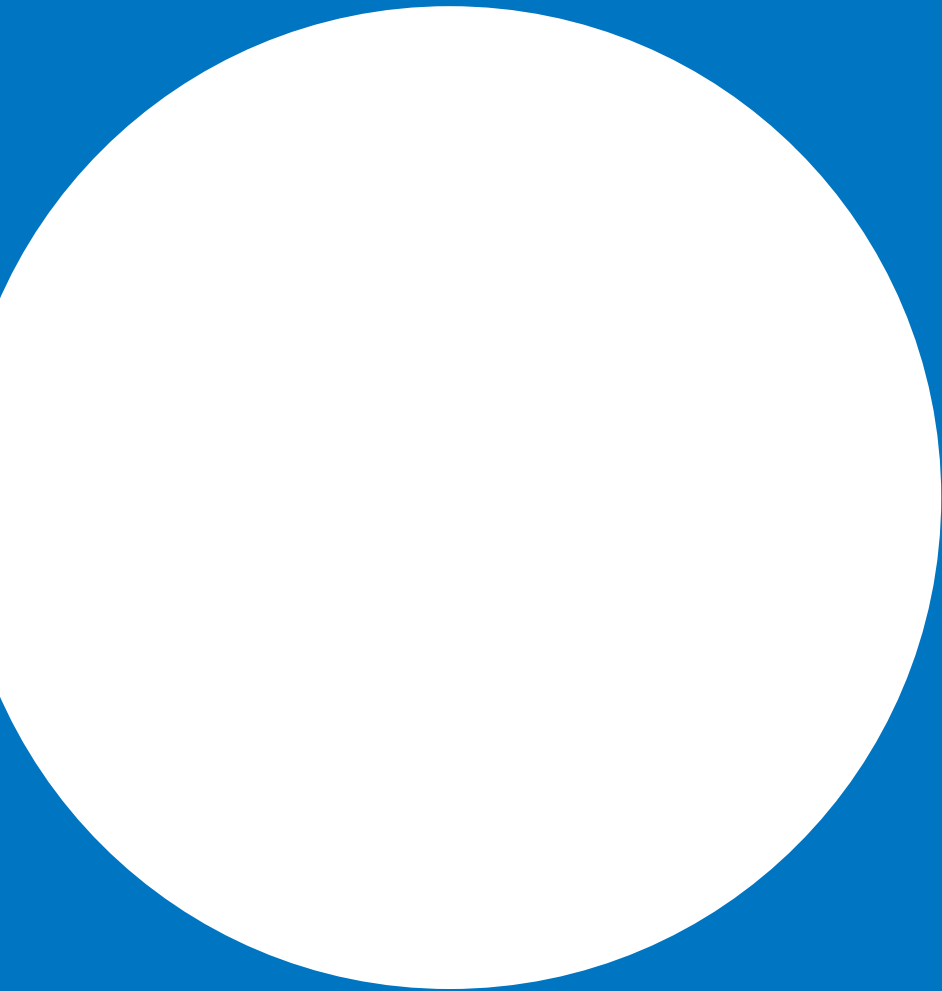
---

6 Richardson, R., Schultz, J. & Southerland, V. (2019) *"Litigating Algorithms 2019 US Report: New Challenges to Government Use of Algorithmic Decision Systems"*, AI Now Institute, Available at <https://ainowinstitute.org/litigatingalgorithms-2019-us.html>

7 "When a public agency adopts a policy of making high stakes employment decisions based on secret algorithms incompatible with minimum due process, the proper remedy is to overturn the policy, while leaving the trade secrets intact." (HFT v. HISD, 251 F.Supp.3d 1168 (S.D. Tex. 2017))


# Algorithmic Fairness and Anti-Discrimination Law





Presenter

Alice Xiang,  
Partnership on AI



# Algorithmic Fairness and Anti-Discrimination Law

Alice Xiang  
Head of Fairness, Transparency, and Accountability Research  
Partnership on AI





The second session continued to deal with legal policy discussions, focusing mainly on the topics of the fairness of algorithms and anti-discrimination law. Attorney Alice Xiang started the seminar by presenting cases in which the COMPAS recidivism prediction algorithm mentioned in the first session discriminated against African American, and Amazon's hiring algorithm discriminated against women. She further presented a case where the distorted image of a group caused representational harm. For instance, when you enter the search word "CEO" on Google, most of the image search results will be male CEOs, and the female image showing up at the top of the list is a Barbie doll dressed as a CEO. What these cases have in common is that the AI algorithm's learning dataset reflects past biases.

Next, attorney Xiang proceeded to a discussion about what algorithmic bias is. She presented (1) an approach centered on how the algorithmic decision-making process systematically caused bad results for a specific small group, and (2) the approach centered on the disparity arising from demographic or other characteristics. The technical definition of this algorithmic bias should be compatible with the existing anti-discrimination laws. Only under such circumstances, we can formulate methods to reduce such bias.

From this perspective, attorney Xiang explained the concept of a protected class variable under the US anti-discrimination laws. In the US, discrimination based on race, gender, age, disability, national origin, religion, and others are legally prohibited.<sup>1</sup> The most intuitive way to comply with these legal prohibitions is to remove these variables from the training set. Attorney Xiang, however, points out that this is a naïve approach because the same algorithmic bias can occur if other variables (proxy) replacing the protected class variables are combined. For example, in the hiring process, the race of the applicant can be predicted based on the applicant's residential address and income level. In the COMPAS algorithm, the race was not entered as a variable, but statistically significant differences were shown in the recidivism rate evaluation between races. Nevertheless, attorney Xiang pointed out that “fairness through unawareness” is still the most widely accepted approach in the industry. For example, the U.S. Department of Housing and Urban Development has enacted rules that exempt algorithms that do not use protected variables from liability for discrimination.<sup>2</sup>

So, how do we reduce algorithmic bias? Attorney Xiang argued that to effectively reduce algorithmic bias, the protected class variables should be used according to the context, and through this, the fairness and accuracy of the algorithm can be improved at the same time. This approach shares the basic arguments used for affirmative action in the existing university admissions process. Affirmative action

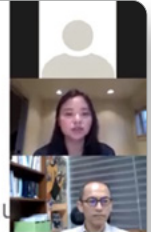
---

1 For example, under the title VII of the Civil Rights Act of 1964, discrimination based on certain demographic variables are prohibited in the context of employment. 42 U.S.C § 2000e-2(a).

2 Fair Housing Act, the Disparate Impact Standard by the US Department of Housing and Urban Development, 84 Fed. Reg. 42854 (2019. 8. 19.)

## Takeaways

- Legal compatibility is key to employing algorithmic bias mitigation techniques in practice
- There is a tension between the technical need to consider protected class attributes in order to mitigate bias and the law's preference for decision-making that is blind or neutral to these attributes
- Causality is a key concept in both ML and law that can help distinguish between different uses of protected class variables
  - Hopefully make it possible to prevent proliferation of biased algorithms and permit use of bias mitigation techniques



그런데, 사회적 악자 변수를 써야 편향성을 줄일 수 있지만

Session 2 Capture (caption)

refers to a policy that gives preferential treatment to groups that have historically experienced discrimination and have been mainly used as a policy to consider racial minorities and women.

However, the Bakke U.S. Supreme Court ruling on active preferential treatment declared that while race is a factor that can be considered in college entrance exams, but the quota system that sets the ratio of specific races is unconstitutional. It was determined that affirmative action should be allowed for the educational purpose of the diversity of the student body, rather than the rectification of histor-

ical inequalities.<sup>3</sup> Attorney Alice Xiang pointed out that the algorithmic bias reduction method using the protected class variable is similar to the quota system that determines the ratio and focuses on remedying historical inequalities, which contradicts the Bakke ruling. In the 2003 Grutter v. Bollinger and Gratz v. Bollinger rulings, the Supreme Court of the United States continued a similar view by deciding that race consideration was allowed in evaluating individual applicants, but it was not allowed to award additional points solely based on the fact that they were minorities.<sup>4</sup>

The US Supreme Court's rulings on aggressive preferential treatment acts as a constraint in devising a method to address algorithmic bias. This is because, if a quota system that guarantees a certain percentage or a method of assigning additional points is not allowed, how to correct algorithmic bias for minorities or women becomes a problem. Attorney Xiang argued that "causality" could be used as an important requirement in the disparate treatment jurisprudence of anti-discrimination law.<sup>5</sup> In other words, instead of unconditionally prohibiting the algorithm from using protected class variables such as race or gender, it is better to specifically evaluate the causality between protected variables and biased outcomes. According to this approach, even if the same protected class variable is used in the al-

---

3 Regents of the University of California v. Bakke, 438 U.S. 265 (1978)

4 In the case of Grutter, the admissions decision of the University of Michigan law school became an issue, whereas in the case of Gratz, the admissions decision of the University of Michigan college became a problem. The University of Michigan undergraduate admissions system evaluated applicants based on a scale of 150, and all of minority applicants received additional 20 points.

5 Texas Dept. of Housing and Community Affairs v. Inclusive Communities Project, Inc., 576 U.S. 519 (2015)

---

gorithm, the legal evaluation can differ depending on whether the protected class variable is used in the direction of increasing the causality of biased decisions or in the direction of reducing the bias causality.

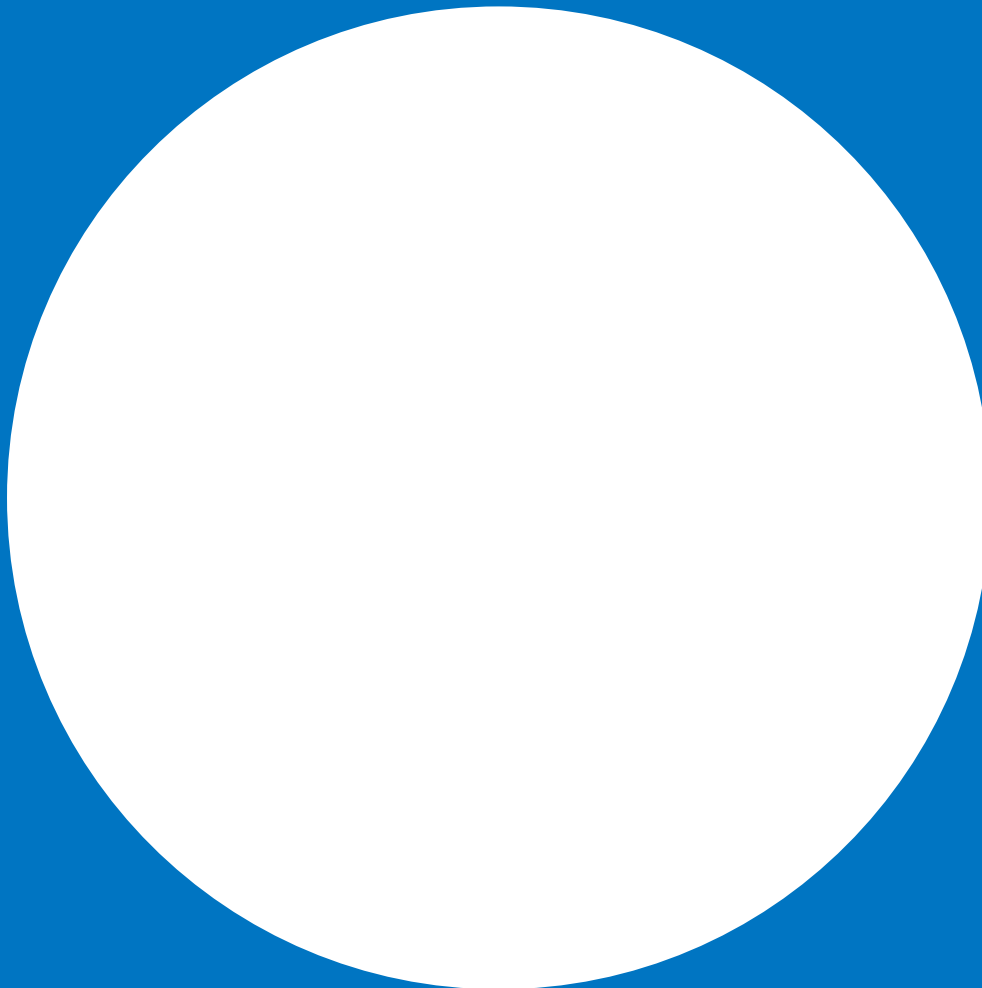
Attorney Xiang pointed out that applying the anti-discrimination law to the algorithm bias, which prevents considering the protected class variables, can lead to a result of solidifying unreasonable discrimination. She concluded the seminar by suggesting an evaluation based on causality as a solution<sup>6</sup>.

---

6 For more details on this presentation, see the speaker's paper, Alice Xiang. (2021). "Reconciling Legal and Technical Approaches to Algorithmic Bias", Tennessee Law Review, Vol. See 88, No. 3 (forthcoming).

# Competition in the Era of Algorithms: Evolving Law & Policy





Presenter

Niamh Dunne,  
The London School of Economics  
and Political Science

# Competition in the Era of Algorithms Evolving Law & Policy

*Niamh Dunne*

[N.M.Dunne@lse.ac.uk](mailto:N.M.Dunne@lse.ac.uk)

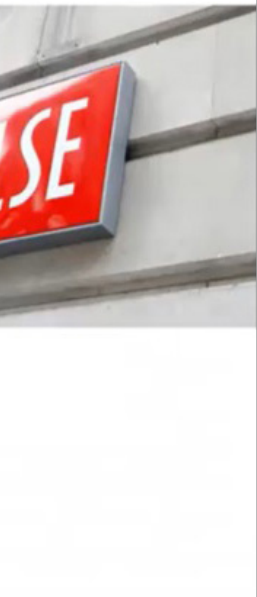
Seoul National University, 16 September 2020



THE LONDON SCHOOL  
OF ECONOMICS AND  
POLITICAL SCIENCE ■



ms:



Session 3 dealt with legal and policy discussions on the influence of artificial intelligence on market competition. Professor Dunne, the presenter, explained what problems the commercial use of artificial intelligence algorithms is causing in the market, especially in terms of competition law, and she started the discussion by posing a question about how competition law needs to evolve to come up with solutions to these new problems. Professor Niamh Dunne emphasized that the presentation is focused on the EU competition law specifically, but it could provide many implications for discussions in other jurisdictions.

Competition law authorities in the European Union are said to have developed and responded to four key scenarios concerning the increasing use of algorithms in the market. The four scenarios are (1) a dystopian scenario called “robotic cartel”, (2) a scenario in which an algorithm exists as a fact of the market, (3) a scenario in which an algorithm aggravates market competitive harm, and (4) a scenario where competition within algorithm becomes the competition on the merits.

In the first scenario, when many market participants apply a pricing algorithm, these algorithms can commit an act of collusion without human intervention. There are various studies on this, but it is said that the pros and cons are widely opposed to the limitations of the robot cartel and the competition law proposed in the book “Virtual Competition: The Promise and Perils of the Algorithm-Driven Economy”. However, there are no cases in which the robot cartel scenario

has become a problem in real life, so Professor Dunne mentioned two cases in which humans used artificial intelligence as a means of collusion. One is the case of *Trod/GB* in which marketing software that included the pricing method induced an explicit agreement among multiple operators,<sup>1</sup> and the other is the case of *Topkins* in which algorithms were used as a means of collusion between poster companies in Amazon.<sup>2</sup> Recently, although it was research conducted in the laboratory environment, the result of algorithms learning the higher consensus prices have been published.<sup>3</sup>

The second scenario is that the algorithm becomes a fact in the market. Here, the algorithm functions as part of the market, not as a tool for abuse. In this regard, Professor Dunne introduced two examples. First, in the *Guess* case<sup>4</sup>, when Guess, a clothing manufacturer, signed a contract with an official distributor, it reached a vertical agreement to ban Google Ad from participating in keyword bids. The European Commission (hereinafter referred to as “EC”) ruled that this violated Article 101 of the Treaty on the Functioning of the European Union (hereinafter referred to as “TFEU”).<sup>5</sup> Another example is the

---

1 UK competition authority decision in *Trod Ltd: posters and frames*, 21 July 2016, Case 50223

2 *United States v. Topkins*, No. 15-201 (N.D. Cal. Apr. 30, 2015)

3 Calvano, E., Calzolari, G., Denicolo, V., and Pastorello, S. (2019) “*Artificial Intelligence Algorithmic Pricing and Collusion*”. Available at <http://dx.doi.org/10.2139/ssrn.3304991>

4 *Guess* (Case AT.40428 — *Guess*) European Commission Decision C/2018/8455 [2018] OJ C 47, 6.2.2019, p. 5, Available at [https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.C\\_.2019.047.01.0005.01.ENG&toc=OJ:C:2019:047:TOC](https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.C_.2019.047.01.0005.01.ENG&toc=OJ:C:2019:047:TOC)

5 Article 101 of the TFEU is a provision corresponding to Article 19, “Prohibition on Illegal Cartel Conduct” of the “Monopoly Regulation and Fair Trade Act” in South Korea, which contains a provision prohibiting collusion (cartel) and agreements among business entities.

---

Google Search (AdSense) case.<sup>6</sup> This is the case in which Google has implemented an exclusive transaction policy for online media companies and others, stating that its advertisements should be placed in the most profitable position. The EC viewed this restriction as a violation of Article 102 of the TFEU.<sup>7</sup>

To understand how anti-competitive behavior in the marketplace is conducted in both cases, one needs to know how the algorithm works. For example, in the Google Search (AdSense) case, it is necessary to understand the principles of online advertising to determine whether it is anti-competitive for Google to allocate its ads in the most visible space of the web. Because actions in the above cases do not only occur online using algorithms but can also occur without algorithms in other markets. Whether the algorithm is used, or whether the conduct took place in the digital market is irrelevant to the problem.

The third scenario is in which restrictions on competition increase due to the use of algorithm. A representative example is resale price maintenance (RPM), which places restrictions on selling for less than a certain price when a producer signs a contract with a distributor. Since the Binon case in the 1980s,<sup>8</sup> the European Union has viewed the RPM as a “by object anti-competition act” and has banned it under Article 101 of the TFEU. However, the EC has not dealt with this

---

6 Google Search (AdSense) (Case 40411) European Commission Decision of 20 March 2019 (not yet published)

7 Article 102 of the TFEU is a provision corresponding to Article 3(2), “Prohibition on Abuse of Market-Dominant Position” of the “Monopoly Regulation and Fair-Trade Act” in South Korea.

8 Case 243/83 *SA Binon & Cie v SA Agence et messageries de la presse* [1985] ECR 2015

issue after the decentralization<sup>9</sup> of the EU competition law enforcement in 2004 due to the problem that the effects of these actions are unclear. However, as the RPM has recently emerged as a problem in the online market, the EC ruled in 2018 that the RPM by Asus, Denon & Marantz, Philips, Pioneer, and Guess was a violation of the competition law<sup>10</sup> based on the jurisprudence of “by object anti-competition act”.

As a reason for the reapplication of a rule that had not been applied by the EC for more than 10 years, Professor Dunne pointed out that the use of price-tracking and price-setting algorithms has increased. From the perspective of commodity producers, it is easy to monitor whether distributors comply with the resale price through this algorithm, so the pressure to maintain the resale price policy is strengthened. In addition, since the distribution operator can easily see the prices of competitors in the distribution market through the algorithm, the execution of resale price maintenance can also be facilitated. In other words, the damage is aggravated as the new technology called algorithm is applied to the existing resale price maintenance behavior. Therefore, the old jurisprudence had to be brought in again.

The fourth scenario is when the algorithm itself becomes a part of competition on the merits. In this regard, Professor Dunne introduced

---

9 The enforcement rules of Articles 101 (formerly Article 81) and 102 (formerly Article 82) of the TFEU were replaced by R 1/2003 from R 17/62 starting on May 1, 2004. The authority for enforcement of competition laws concentrated in the EC was distributed to the competition authorities and courts of each member states to allow the autonomous enforcement of competition laws in the European Union.

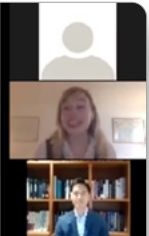
10 Case AT. 40465 – Asus, Case AT. 40469 – Denon & Marantz, Case AT. 40181 – Philips, Case AT. 40182 – Pioneer

## Beyond Competition Law: Options for Additional Regulation (III)



- Need for a '**New Competition Tool**' to fill the gaps within existing competition law framework?
  - Currently being considered by EU Commission
  - Proposal resembles UK market investigation regime, which enables CMA to conduct in-depth review of markets 'not working well' to determine if any feature generates 'adverse effects on competition'; extensive remedial powers (including structural separation)
  - Again, could be applied to limit or direct use of algorithms by Big Tech firms
  - A (tentative) objection: do current 'gaps' in competition regime potentially exist for good reasons?!

따라서 우리가 새로운 경쟁률을 도입하게 되면, 과잉규제가 되지 않는지를 고민해볼 수 있습니다.<sup>3</sup>



Session 3 Capture (caption)

the case of Google Search (Shopping),<sup>11</sup> which was given the largest penalty in EC history under Article 102 of the TFEU in 2017. In the “general search market”, Google held a monopoly with 85% market share of the European market at the time, and it continuously improved the service to consumers through excellent search algorithms. Nonetheless, in the “comparison shopping market” using a separate search engine, Google struggled. So, here, rather than developing a

<sup>11</sup> Google Search (Shopping) (Case AT. 39740 – Google Search (Shopping)) European Commission Decision of 27.6.2017.

better algorithm, Google revised and modified the search algorithm to self-preference its own products by adjusting its products to the top in shopping search results and allocated a less preferred spot for competitors' products. Professor Dunne pointed out that, contrary to the prospect of the dystopian scenario, such an issue was raised because of algorithm promoting market competition paradoxically. Hence, she argued that if this distortion can be prevented and the true desires of consumers can be reflected, the algorithm could rather strengthen competition.

As can be seen from the discussions so far, today's algorithms have become a part of the market, and therefore an understanding of algorithms is required to understand the effects of algorithms on competition law. It is important to understand how the algorithm affects competition, rather than trying to grasp all the technical details of how the developer coded the algorithm. Professor Dunne mentioned that as a starting point, businesses should not deliberately intervene to cause anticompetitive effects on freely operating algorithms. In addition, if the algorithm is also considered as part of the business operator's behavior, and the focus is on determining whether the behavior using the algorithm is anti-competitive, the past opinion which argued that it was difficult to apply the competition law to the algorithm that is a highly technical field can no longer be accepted.

Meanwhile, Professor Dunne acknowledged the necessity of making changes to today's competition law in line with the digital age and suggested several options. The first example is "Regulation 2019/1150 on Promoting fairness and transparency for business us-

---

ers of online intermediation services”,<sup>12</sup> which the European Union has already accepted. This rule imposes an obligation to deliver particularly fair terms and conditions to operators that provide online brokerage services and stipulates the strengthened transparency requirements in the technology of determining ranking criteria using algorithms. However, Professor Dunne pointed out that because the regulation does not cover general users, its effects may be limited.

The second example is to impose an Ex ante Code of Conduct on digital business entities, which is currently discussed in Australia, the UK, and Europe. In particular, the EU is targeting large tech companies and discussing the Digital Service Act Package. It is worth noting that the scope of application is determined based on “economic power” rather than “market power” as stated in Article 102 of the TFEU. Nevertheless, Professor Dunne added that it is necessary to think whether the regulatory authorities can properly outline regulations that will affect the digital market in five or ten years.

The third example is to prepare a “New Competition Tool” that fits the new digital era rather than the existing competition law system. It is the opinion that a new system other than Articles 101 and 102 of the TFEU should be introduced, and the UK is currently considering a plan to introduce a policy similar to the UK market investigation regime, which examines whether the algorithm works properly.

---

<sup>12</sup> Available at <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32019R1150>

# Google Perspectives on AI Governance







Presenter:

Charina Chou,  
Google

# Google perspectives on AI governance

Charina Chou

Session 4 shifted the focus to the discussion of governance from the individual company standpoint and specifically dealt with Google's perspectives. According to Dr. Charina Chou, Google's goal is to organize the world's information and make it universally accessible and useful.<sup>1</sup> Artificial intelligence is playing a very important role in achieving this goal for Google. For example, one can look at Google's representative product, the search engine. Information that exists on the Internet has evolved from text to image and video, and it is no longer possible to search effectively with the algorithm built when Google first launched the search service 20 years ago. So, Google Search developed in the form of machine learning technology. Google's technological advances provided by artificial intelligence are not just limited to search but are also being used in medical image processing, such as early detection of cancer.

However, the use of artificial intelligence technology creates new problems and risks. Dr. Charina Chou cited an example of a lip-reading algorithm<sup>2</sup> that Google considered developing for Aphonia<sup>3</sup> pa-

---

1 Google's Official Mission Statement: "To organize the world's information and make it universally accessible and useful."

2 Lip reading is a technique that recognizes words by looking at the movement of the lips, face, and tongue.

3 It is a loss of voice caused by disorders of the larynx or tissues that control the larynx.

tients upon the request of a medical institution in the UK, two years ago. Despite the utility of helping aphonia patients communicate, if the lip-reading algorithm is developed and widely distributed, there is a lot of potential for abuse, including privacy invasion. Google determined that ethical decision-making on artificial intelligence technology was necessary, so in 2018, it drafted and distributed the Google AI Principles.<sup>4</sup> Google is making decisions based on these principles when ethical judgments are needed such as whether or not to develop a lip-reading algorithm.

Dr. Chou emphasized that technical measures play an important role in the implementation of AI principles. Google is providing a summary of key information about the dataset it uses to train AI models, just like the nutritional table of foods. For example, in the case of the Open Images Extended-Crowdsourced dataset, which Google released for machine learning, the purpose of data use and the data source are disclosed in detail, so researchers and developers can know and utilize the characteristics of the dataset. Furthermore, Google is using a crowdsourcing method to build a more representative dataset. Data from other cultures are not adequately represented since most of the existing datasets are built based on Western cultures and customs. Dr. Chou pointed out that when someone searches an image for weddings or brides, a picture of a white dress comes

---

4 The main contents are as follows: AI applications must (1) be socially beneficial, (2) avoid creating or reinforcing unfair bias, (3) be built and tested for safety, (4) be accountable to people, (5) incorporate privacy design principles, (6) uphold high standards of scientific excellence, (7) be made available for uses that accord with these principles. AI applications will not pursue (1) technologies that cause overall harm, (2) technologies whose principal purpose is to cause injury to people, (3) surveillance violating international norms, (4) technologies whose purpose violate international law and human rights. <https://ai.google/principles/>

---

up as a search result. This is a Western custom that will not appear in the wedding scenes in other countries like India. This technical effort is an attempt to increase the completeness and diversity of the dataset in the pre-training stage.

Dr. Chou emphasized the importance of technical measures in the post-training phase and introduced a way to impose “fairness” constraints on AI models. For example, in Turkish, pronouns do not have a gender distinction, but when the Turkish sentence “o bir doktor” is translated into English through Google Translate, the result first appeared as “he is a doctor”. Thus, Google placed a post-model constraint so that the user can choose the desired outcome between “he is a doctor” and “she is a doctor”. Furthermore, Dr. Chou introduced the TCAV (Testing with Concept Activation Vectors) methodology as a technological advancement for the “interpretability” of artificial intelligence. It is a method of expressing how much the user-defined concept contributes to the classification of the input value by the artificial intelligence model as a directional derivative.<sup>5</sup> By synthesizing this information, Google is running a page on “General recommended practices for AI”.<sup>6</sup> This page provides data in terms of fairness, interpretability, privacy, and security for the practical use of artificial intelligence.

Dr. Chou pointed out that there are many important points in the use of artificial intelligence technology not only in the technical aspect

---

5 Kim, B., Wattenberg, M., Gilmer, J., Cai, C., Wexler, J., Viegas, F., & Savres, R. (2018). *“Interpretability Beyond Feature Attribution: Quantitative Testing with Concept Activation Vectors (TCAV)”*, Proceedings of the 35th International Conference on Machine Learning, in PMLR

6 <https://ai.google/responsibilities/responsible-ai-practices/>

but also in the operational processes of a company. In this regard, Dr. Chou stated four major points. First, in the process of using artificial intelligence technology internally, there should be an environment where each stakeholder can freely express opinions at any time while being familiar with the relevant knowledge. Google is running various educational programs to implement this aspect. Second, the technical devices mentioned above should not only be developed but should be easy to understand and utilized by non-technical employees. Third, there must be an update procedure to record, review, and share discussions on individual cases such as the case of lip-reading algorithm. Finally, it is necessary to continuously communicate with the external community about the use of artificial intelligence technology through various conferences and reports.

Dr. Chou argued that public policy on artificial intelligence should also play an important role. First, to promote the use of artificial intelligence technology, the private use of artificial intelligence technology could be promoted by sharing data sets held by public institutions. Much of the data provided by public institutions, for example, is in an unrefined form that is difficult for machines to read and understand, and government investment is needed to improve the quality of those data. In addition, Dr. Chou emphasized that it is necessary to organize and provide information on the characteristics of data, to strengthen education on artificial intelligence technology, and to actually use artificial intelligence in the public field.

Lastly, Dr. Chou pointed out that the government's role is to establish "explainability" standards for artificial intelligence models in to reduce controversy over the use of AI technology. The debate about the right to request human intervention concerning human-AI collaboration is also an area of public policy. After all, new problems will

---

arise from the use of AI technology in each field, and the important thing is how well the normative system can respond to these new problems. For problems that the existing normative system has not envisioned, it is necessary to revise the laws and regulations quickly. Dr. Chou stated that the government's role in removing these regulations is important when regulations hinder the social utility expected to be achieved by AI technology. Finally, she concluded the discussion by adding an explanation of future-oriented discussions such as basic research, applied research, and human-AI interaction research conducted by Google.

# Differential Privacy: What is it and Where








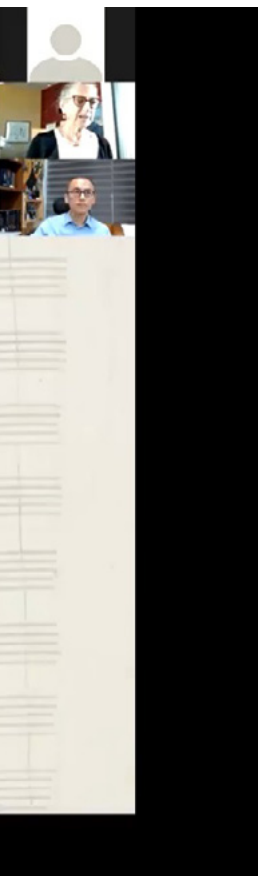
Presenter

Cynthia Dwork,  
Harvard University



Differential Privacy  
What is it and Where is it?

Cynthia Dwork  
Harvard University  
Radcliffe Institute for Advanced Study



Session 5 held a discussion on the topic of differential privacy, which is receiving much attention as a next-generation technology that harmonizes data analysis and personal information protection in the era of big data. Professor Cynthia Dwork, the presenter who created the concept of differential privacy, began her presentation by explaining the difference between processing statistical data for the entire population and analyzing a specific subset of the population. She emphasized that differential privacy is a means of protecting privacy in the statistical processing of the entire population. At a first glance, this process of statistical analysis feels like protecting privacy. For example, the characteristics of the target population can be consistently inferred through the sample, but the sample value does not convey definitive information about an individual, so the individual can always conceal his or her participation in the group by claiming “I am not included in those statistics”.

Differential privacy was developed for use in the US census to design “privacy-preserving data analysis”, which is an old problem in statistics. The basic framework takes the form of receiving an answer from the database when an analyst makes a query to the database. The problem is that the more accurate and abundant the responses are, the more likely the privacy of individuals included in the database is

**Differential Privacy**

The outcome of any analysis is essentially equally likely, independent of whether any individual joins, or refrains from joining, the dataset.

데이터 베이스에 페인의 정보가 있는 것과 없는 것, 각각과 상호작용을 한다고 가정해봅시다.

Session 5 Capture (caption)

violated.<sup>1</sup> Professor Dwork emphasized that this is a mathematical law.

So how is differential privacy, known as a data analysis mechanism that preserves privacy, defined? The easiest way to think of is the definition that privacy is preserved if “new” information about a specific person included in the database cannot be obtained through

<sup>1</sup> An example given by Professor Cynthia Dwork is if there are statistical data such as “Number of Nobel Physics Award Winners with Vehicles” and “Number of Female Nobel Physics Award Winners with Vehicles”, we can figure out whether the only female Nobel Physics Prize winner has a vehicle or not.

---

question and answer.<sup>2</sup> For example, even if you can get information published on the web about a specific person from a database, this is not “new” information, so privacy is not infringed.<sup>3</sup> The problem with this definition is that if a particular person and other participants share a certain characteristic, the results obtained by observing the other participant become the “new” information about the specific person, and according to this definition, the privacy of that person is considered to be violated. In extreme terms, the privacy of all individuals is violated when the traits shared by everyone in the world are known, a counterintuitive conclusion that is difficult to accept.

To solve this problem, differential privacy was defined. Under the definition, privacy is preserved if the same “new” information can be obtained about a specific person even if a question and answer are made while a person is replaced by another person. In other words, whether a specific person is included in the sample has no bearings on the results, the dataset is differentially private. Similar analysis results mean similar probability distributions. For example, when comparing a coin with a 50% chance of landing on a head and a coin with 50.1% chance are considered, the probability distribution functions obtained by repeatedly throwing these coins are almost the same, so it is not possible to know which coin was thrown only by the result of one or several coin tosses. The characteristic of such a database

---

2 DALENIUS, T. (1977), *Towards a methodology for statistical disclosure control*. Statistik Tidskrift, Vol. 15, 429-444, 2-1

3 Professor Cynthia Dwork explained that this definition is the same concept as Semantic Security in cryptography. Semantically secure cryptosystem refers to cryptosystem that can extract only negligible information about the original text from the cryptogram.

is called “stability”,<sup>4</sup> which not only preserves the privacy of persons included in the dataset but is known to prevent over-fitting in the context of machine learning.<sup>5</sup>

Professor Dwork also explained the mathematical definition of differential privacy. To summarize, model M satisfies  $\epsilon$ -privacy if the following conditions are satisfied for all adjacent datasets  $x, y$ ,<sup>6</sup> and all result values  $S$ .<sup>7</sup>

$$\Pr[\text{Observe } S \text{ at } M(x)] \leq e^\epsilon \Pr[\text{Observe } S \text{ at } M(y)]$$

In the above definition,  $\epsilon$  means a privacy loss and a degree of differential privacy. The above definition of differential privacy represents the characteristics of Model M. For any analysis, the algorithm (model) M that satisfies the above equation is expressed as satisfying  $\epsilon$ -privacy.

Professor Cynthia Dwork further described the characteristics of differential privacy. First, differential privacy is future-proof. This is a characteristic that comes from the definition itself, and the differential privacy is maintained even if changes occur, such as when additional information is presented in the future. Second, even in the case of models to which differential privacy is applied, small privacy losses

---

4 It means the outcome is similarly “stable” regardless of including a specific person in the data set and not.

5 In other words, if the training set and the test set are similar (stable), the model trained using the training data set will work well on the test data set. Otherwise, the model is overfitting to the training dataset.

6 As discussed above, the data set is completely identical except whether a specific person is included.

7 Dwork C., McShery F., Nissim K., Smith A. (2006) *Calibrating Noise to Sensitivity in Private Data Analysis*. In: Halevi S., Rabin T. (eds) *Theory of Cryptography*. TCC 2006. Lecture Notes in Computer Science, Vol. 3876. Springer, Berlin, Heidelberg.

---

occur through individual questions and answers, and these losses are accumulated as calculations are repeated.

Professor Cynthia Dwork then described the Laplace Noise Addition, a concrete method of implementing differential privacy. The Laplace mechanism refers to a technique that provides a modulated result to the user by adding random noise generated from the Laplace distribution in which the variance is set to be proportional to the sensitivity/ $\epsilon$ . A local model and a centralized model can be classified based on when such noise is added. The local model has an important feature that personal information is altered by the client device and then is collected by the server. In this respect, it is contrasted with the centralized model in which the personal information of the subject is collected in the server as is and then undergoes a alteration process. Google and Facebook, for example, have used the centralized model for mobility data to cope with COVID-19, and Microsoft is using the centralized model for error reporting in Windows and text prediction model in the Office. A typical example of the local model is Google's explanation that differential privacy is used as an anonymization technique in the Randomized Aggregatable Privacy Preserving Ordinal Response (RAPPOR) technique applied in 2014.<sup>8</sup>

Furthermore, the differential privacy will be applied to the publication of the 2020 U.S. Census results. John Abowd, Chief Scientist of the US Census Bureau, said of the 2010 Census,<sup>9</sup> "Technology advances have exposed the problems of conventional methods. It is

---

8 Google, "How Google Anonymizes Data", Privacy and Terms <https://policies.google.com/technologies/anonymization?hl=ko>

9 The United States conducts a census every 10 years in accordance with Article 1 Section 2 of the Constitution.

now possible to recombine information that may infringe privacy by using public data that were previously known to preserve privacy.” According to Professor Dwork, some researchers are also critical of this because applying differential privacy to census results undermines the utility of the databases.

Professor Dwork criticized the conventional idea that the relationship between the utility of the database and privacy is at odds. Companies, governments, researchers, and data subjects using data may have different preferences for the relationship between utility and privacy. In a society where privacy is not generally recognized at all, the utility of data users will be the highest, and as the protection for privacy is strengthened, the utility of data users will decrease. On the other hand, those who think “there is nothing to hide” will provide their information truthfully even if there is no privacy protection, but those who do not think that way will not trust society and may not provide their information. If so, the relationship between total social utility and privacy would be an inverted U-shape curve, and based on this, Professor Dwork argued that if privacy protection is too little or too many, the utility of the database will be reduced.

Professor Dwork pointed out that setting the level of privacy protection based on the relationship between the utility of the database and the privacy is ultimately a policy issue, and that under the differential privacy system, such policy decisions can be technically implemented through the adjustment of  $\epsilon$  value. This leaves a question of who decides the value of  $\epsilon$ , and which queries will be prioritized when privacy losses accumulate. To solve this problem, Abowd & Schmutte argued that an optimal social solution can be found by measuring the



---

“willingness to pay for data accuracy with increased privacy loss”.<sup>10</sup>

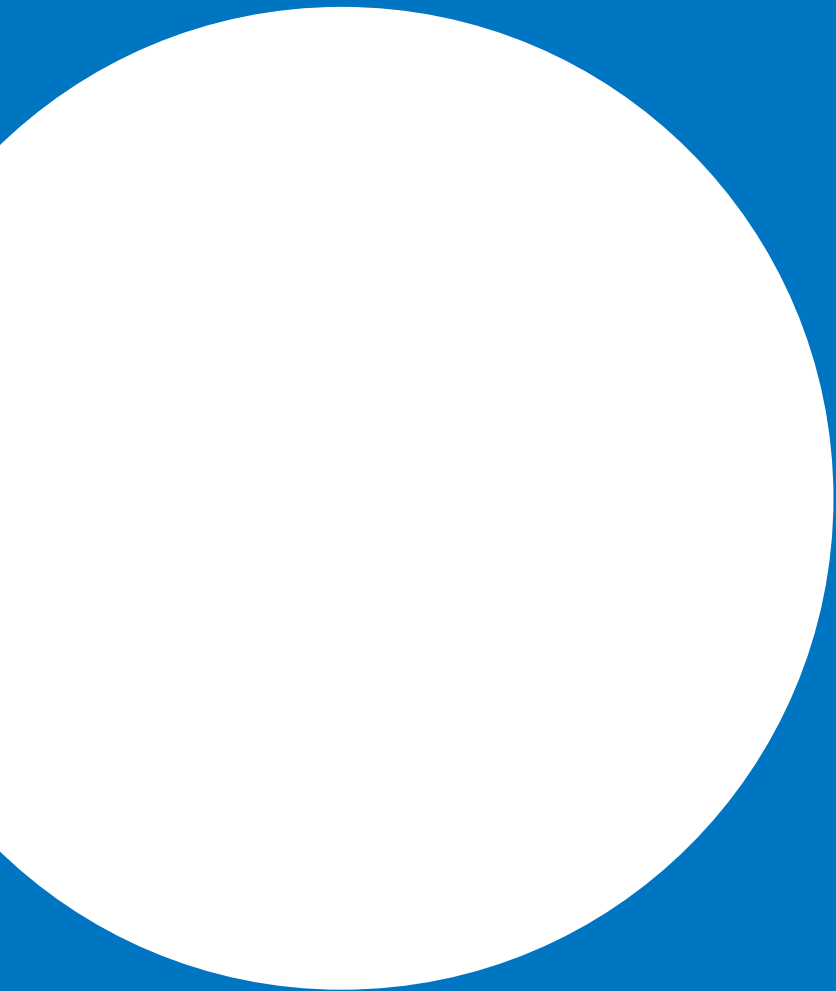
Professor Dwork concluded the seminar by summarizing that differential privacy is an approach that enables database analysis while not learning new information about a specific person in the database.

---

10 Abowd, J., Schmutte, I. (2019) “*An economic analysis of privacy protection and statistical accuracy as social choices*”, American Economic Review, Vol. 109, No.1, pp 171, 194-197.

# Asimov for Lawyer: What Sci Fi can(not) tell us about the future of AI regulation





Presenter

Nicolas Petit,  
European University Institute



European  
University  
Institute

DEPARTMENT  
OF LAW

## Models of Law and Regulation for AI - and what to learn (or not) from Sci-Fi?

Prof. Nicolas Petit

[Nicolas.petit@eui.eu](mailto:Nicolas.petit@eui.eu)

Twitter: @CompetitionProf



Session 6 was discussed about regulation of artificial intelligence technology. Professor Nicolas Petit, the presenter, mentioned that the discourse on the regulation of artificial intelligence technology is actively taking place around the world. Professor Petit commenced his presentation by saying that he would like to learn lessons from science fiction novels in the direction in which future discussions should go.

With regard to the question of “what is artificial intelligence”, it starts with the idea that all physical processes, including the process of thinking, can be modeled with computer algorithms. This is called the Church-Turing thesis. Furthermore, AI can improve its function by learning through an experience like humans. This field of artificial intelligence has gone through both years of ambition and disappointment over the past decades, and expectations are rising again thanks to the improvement of the computing power and the accumulation of big data from about 15 years ago. Recently, scholars even say that, through the expression of “end of theory”, all processes of AI cannot be explained theoretically, and that they are entering a stage that does not require an explanation. Today, AI is being used in various places in society, such as autonomous driving, prediction of judicial outcomes, law enforcement, education, and others. On the other hand, results from deep learning and neural network cannot be fully explained.

Professor Petit introduced four types of regulatory models applica-

ble to these AI technologies. The first is the Black letter law model.<sup>1</sup> According to this model, when a dispute arises, the court seeks a solution by using individual legislation in a specific area suitable for the issue. For example, when music composed by AI becomes a problem, whether it meets the requirements of “created by the copyright holder” stipulated by the current copyright law becomes the central issue of the discussion. Since the existing law, in general, was not enacted with AI technology in mind, courts and legislators try to solve the problem by focusing on the original purpose of these laws and ordinances. For example, the reason for the introduction of the concept of “legal person” was to “foster economic exchange”. Therefore, discussion should proceed in conjunction with the fulfillment of this purpose in the context of the legal personhood of artificial intelligence.

The second model is the Emergent phenomena model, which seeks to address emergent phenomena with new legislation. This model is based on the idea that AI technology will create unprecedented economic and scientific issues, so new legislation is needed. Recent similar examples include discussions on drone law and robot law. Unlike the first model, in which legal experts mainly lead the discussion, the second model differs in that non-legal scholars such as engineering experts are mainly involved in the discussion. Also, compared to the first model, it is not limited to a specific area, but emphasizes integrative regulation and tends to focus more heavily on normative discussions. This is due to the fact that technology experts tend to focus more on “what the law should do” rather than “what the law does”.

---

1 It is an expression representing the form of specific legal norms based on the principles of basic law accumulated for a long time.

---

The third model is the ethical model, which, in conjunction with applied ethics, argues that AI should be provided with a norm to distinguish between good and evil. While the ethical model does not explicitly prescribe the discussion of law and regulation, it implicitly presupposes this idea. Virtue ethics, deontological ethics, and consequentialism are representative examples of the ethical model. The virtue ethics is based on a moral behavior in everyday life and, according to Professor Petit, is understood as “transparency” in the context of artificial intelligence. On the other hand, the deontological ethics emphasizes that artificial intelligence must follow specific instructions irrespective of the outcome. Consequential ethics emphasizes that the outcome should be good regardless of obligation. It is said that the above ethical models have been applied to other technology fields such as bioethics in the past.

The fourth model is the risk regulation model, which is like consequentialism but focuses on technically lowering the probability of harm rather than harm itself. As shown in the European Union’s “White Paper on Artificial Intelligence”,<sup>2</sup> mitigating the potential risks of artificial intelligence is the key to this model. Unlike consequentialism that focuses on follow-up measures, risk regulation model focuses on precautionary measures, which is characterized by being based on statistical evidence (evidence-based approach). Therefore, it has a limitation that there must be data that can be analyzed into probability. In the field of artificial intelligence, the model may result in a ban on a specific field of application (e.g. weapons), or as a preventive action in a high-risk field (e.g. face recognition technology).

---

2 European Commission. (2020). *White paper on artificial intelligence—a European approach to excellence and trust.* Brussels.

Professor Petit then described four possible fallacies that can be applied to the model mentioned thus far, which are: (1) The paradox of irrelevant law, (2) The problem of redundant law, (3) The failure of good intentions, and (4) knee jerk regulation.

First, the “paradox of irrelevant law” specifically relates to the first model, which presupposes explicit norms. Because lawyers imagine the future based on current technology, laws that are irrelevant to the future can emerge. If lawyers have been told about “flying” vehicles rather than “self-driving” cars in the past, they may have enacted detailed laws about flying cars, not self-driving cars by now. The problem of redundant law is related to the second model and refers to a problem that arises when the law created by thinking that the phenomenon caused by the new technology is completely different from the existing one regulates the overlapping aspect of the existing law. Professor Petit said that what we face may not be “new problems” but “existing problems caused by new methods”. The third error, the failure of good faith, is particularly related to the ethical model, which means that the ethics applied to technology with good intentions produce side effects. A representative example is a problem that if we put too much emphasis on ethics, we cannot properly cope with the time when legal regulations are necessary. Moreover, since there is no such thing as a universal morality, it can be puzzling how to respond in a case like a trolley dilemma. The last error, the problem of knee jerk regulation, refers to the introduction of excessive preventive regulation for possible risks. Returning to fossil fuels in response to the Fukushima nuclear power plant crisis, or prohibiting autonomous driving entirely by making its accident an issue even though the accident rate by autonomous driving is lower than the human accident rate are examples of overregulation.



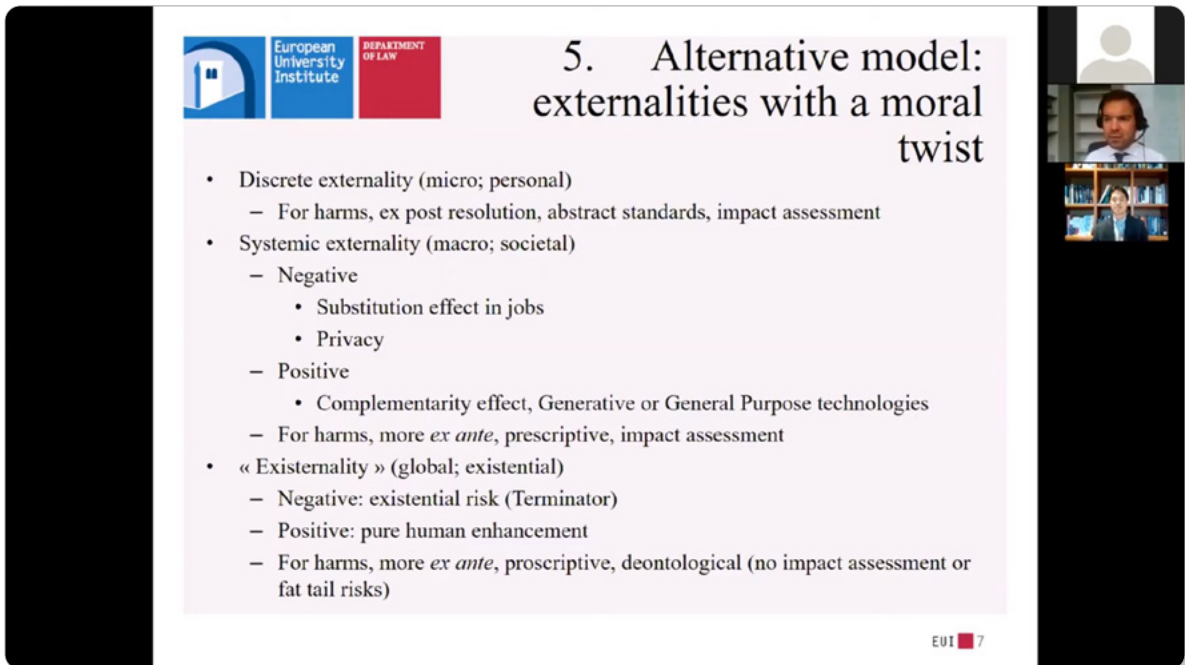
---

Professor Petit then presented a new topic in terms of public policy for artificial intelligence technology, whether technology innovation should be regulated within the framework of the existing legal system or new regulations should be discussed. From the opposing view to new forms of regulations, there are issues such as (1) the conflict between regulation and innovation (e.g. requirements of the European Union's GDPR<sup>3</sup>), (2) the issue of captured regulation (e.g.: Conflict between the auto insurance industry and the public interest when introducing regulations on self-driving cars), and (3) a dilemma that if the technology is too advanced, it may already have reached an irreversible stage even though sufficient information on technology must be accumulated to introduce regulations.

In this regard, Professor Petit pointed out that solving problems based on existing legal norms is not conducive to innovation. Rather, creating new regulations will enable engineers and inventors to carry out research in the right direction. From this point of view, Professor Petit proposed an “alternative model: externalities with a moral twist” as the fifth regulatory model. Professor Petit's alternative model presents three effects to be considered: (1) discrete externality, (2) systemic externality, and (3) ontological externality. The discrete externality is a matter to be considered from micro and personal perspective and can be responded to by the existing legal system. Systematic externality is a matter to be considered from a macro and social perspective, requiring a higher level of preliminary regulation like employment

---

3 The official expression of the European General Data Protection Regulation is Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).



The image is a screenshot of a presentation slide. On the left side, there are logos for 'European University Institute' and 'DEPARTMENT OF LAW'. The main title of the slide is '5. Alternative model: externalities with a moral twist'. Below the title is a bulleted list of points. On the right side of the slide, there is a video call overlay showing two participants: a man in a headset and a woman. In the bottom right corner of the slide, there is a small logo that says 'EUI 7'.

European University Institute DEPARTMENT OF LAW

## 5. Alternative model: externalities with a moral twist

- Discrete externality (micro; personal)
  - For harms, ex post resolution, abstract standards, impact assessment
- Systemic externality (macro; societal)
  - Negative
    - Substitution effect in jobs
    - Privacy
  - Positive
    - Complementarity effect, Generative or General Purpose technologies
  - For harms, more *ex ante*, prescriptive, impact assessment
- « Existentiality » (global; existential)
  - Negative: existential risk (Terminator)
  - Positive: pure human enhancement
  - For harms, more *ex ante*, proscriptive, deontological (no impact assessment or fat tail risks)

EUI 7

Session 6 Capture (caption)

---

or personal information issues. On the other hand, ontological externality must be considered from the humanity perspective and require a fairly high level of regulation.

Concluding his presentation, Professor Petit refuted the claim that science fiction is not realistic. Rather through imagination contained in science fiction, we can gain insight into how laws and regulations can be applied in the field of artificial intelligence. Asimov's novel present an opportunity to predict changes and think about changes in human behavior according to technological advances, which is also helpful to lawyers designing AI regulation. At the same time, he emphasized that Asimov took a neutral position, neither optimistic nor pessimistic about technology. As Asimov argued that the "Three Robot Principles"<sup>4</sup> was a tool to compensate for the weakness of the law in his novel "Runaround", he finished the presentation by saying that human-made things may have inherent errors not only in technology but also in legal policy.

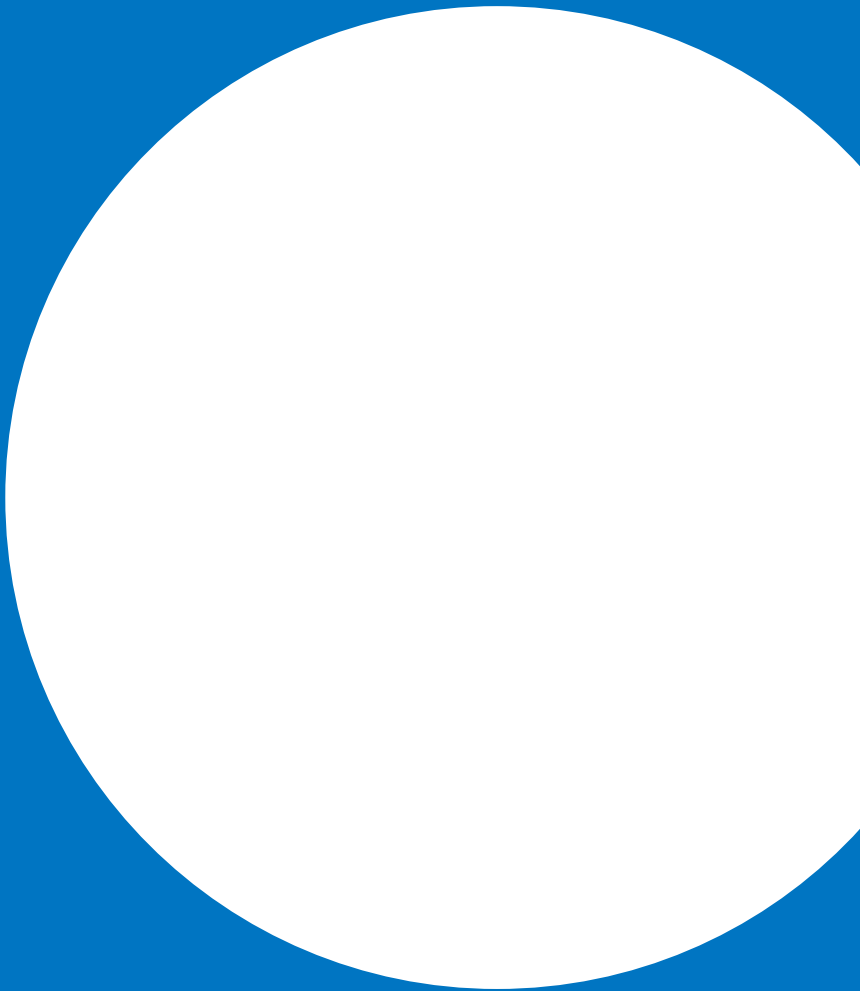
---

4 The contents of the three principles of robots are as follows.

1. A robot may not injure a human being or, through inaction, allow a human being to come to harm.
2. A robot must obey the orders given to it by human beings except where such orders would conflict with the First Law.
3. A robot must protect its own existence as long as such protection does not conflict with the First or Second Laws.

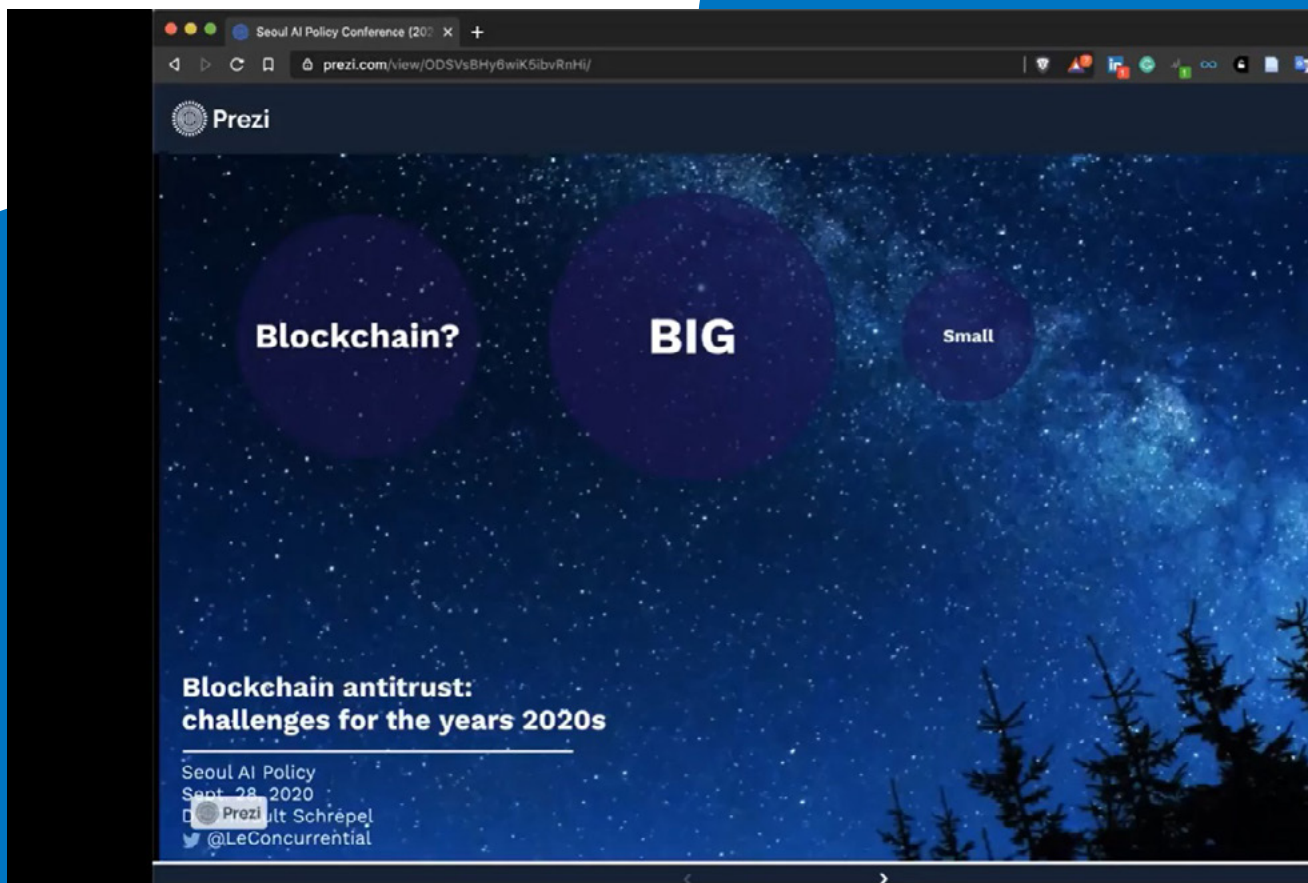
# Blockchain Antitrust: Challenges and Opportunities

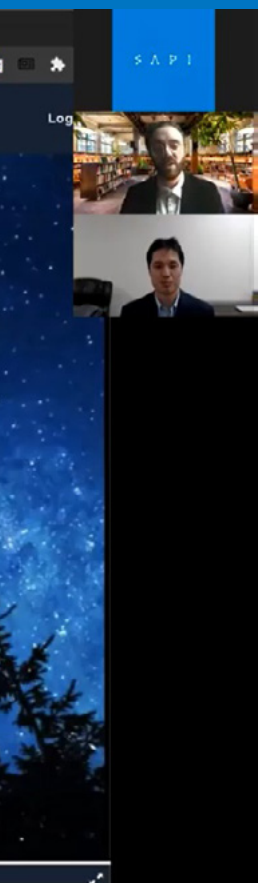




Presenter

Thibault Schrepel,  
Utrecht University





Session 7 switched the topic to blockchain technology and continued the discussion. Professor Thibault Schrepel began his presentation by giving a brief explanation of what blockchain is before proceeding to the legal policy discussion. Blockchain is a type of architecture and has the nature of a base technology that can be applied in various fields. Although no formal definition exists, blockchain generally refers to “open and distributed ledgers that can record all kinds of transactions between users, whether they are passive or automatic.”

There are two layers in a blockchain. Layer 1 is an area where the ledger filled with information, such as an Excel spreadsheet, is shared. In general, blockchain refers to layer 1. This is an open distributed database, which can be thought of as an Excel spreadsheet. Layer 2 can be added on top of Layer 1, and the blockchain software can be classified into the following three types according to the type of the second layer.<sup>1</sup> In other words, blockchain is classified into (1)crypto-currencies, (2)smart contracts, and (3)other types of applications (e.g., Uber).

Professor Schrepel introduced the process of signing a smart contract, a concept that should be dealt with particularly importantly concerning competition law, through a video demonstration. If one enters the names of the parties on the website, selects what laws are applied or what obligations are imposed on each parties, and enters the email addresses of both parties, the contract will be automatically sent to those email address. When the user identified through the public key digitally signs the smart contract, the transaction is complete, and payment is made when one party fulfills the obligations specified in the contract.

Next, Professor Schrepel outlined four conceptual toolboxes that

---

<sup>1</sup> Layer 2 refers to an application running on top of an existing blockchain system.

lawyers and economists should understand. (1)Pseudonymity: Blockchain identifies an individual with a public key instead of an actual identity. With the current technology, it is impossible to convert a public key into an actual identity, and thus it has pseudonymity that allows access to information only with a private key. (2)De-centralization: Since blockchain is stored in a distributed ledger, it is characterized by decentralization, which means records cannot be controlled or deleted from a specific central server and that everyone can access information. (3)Immutability: Blockchain immutable so that information cannot be tampered with in that the hash value completely changes when the underlying information is changed. (4)Unstoppable: Since the transaction process of the blockchain is automatic, it has the characteristic that it cannot be stopped (unstoppable code) unless, for example, a suspension clause is inserted as a specific condition of the smart contract.

After completing an overview of blockchain technology, Professor Schrepel began to discuss competition law. According to him, blockchain technology and competition law are ultimately aimed at allowing transactions free from economic coercion. In the end, competition law is necessary to realize the anti-monopoly economy, which is also the purpose of blockchain technology. Conversely, he argued that blockchain technology could help to realize the purpose of the competition law considering the reality where the rate of detection of anti-competitive behavior by enforcement authorities is low.

However, according to Professor Schrepel, from the perspective of the theory of the firm underlying the competition law, he pointed out that the characteristics of blockchain technology cause difficulties. According to Ronald Coase's theory of the firm, a company is a by-product of reducing transaction costs by replacing the decision-making



---

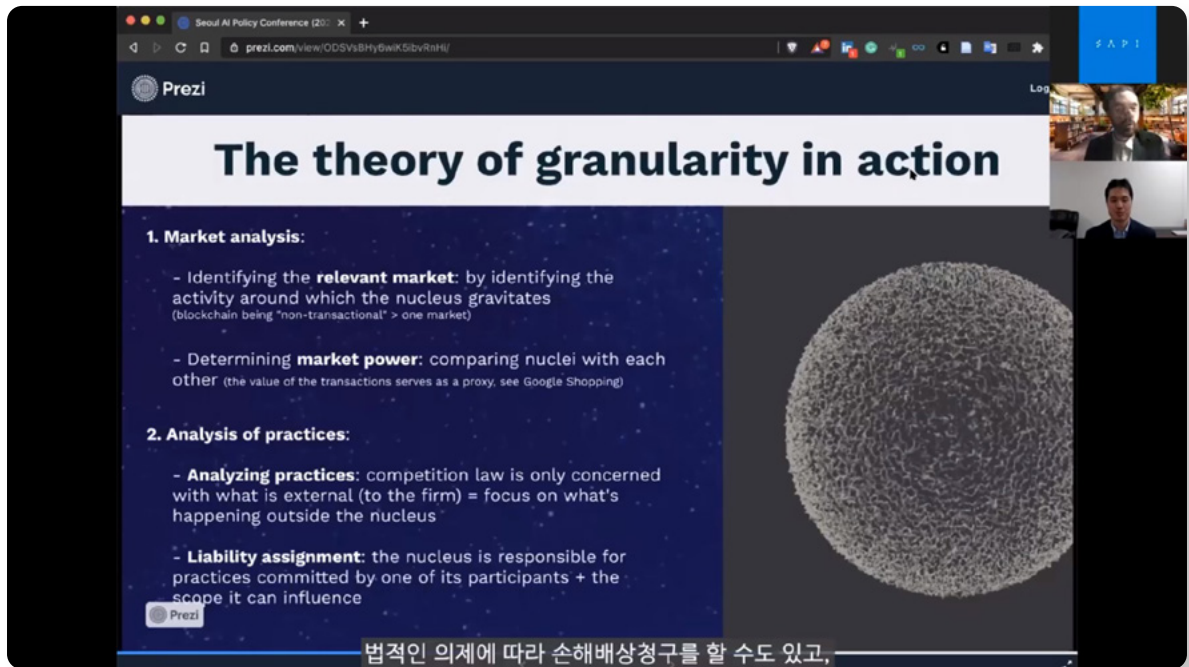
needed for all economic activities with the top-down decision-making process of the controlling entity.<sup>2</sup> Even if the two companies are separate legal entities, collusion cannot be established between parties if the controlling entity is the same simply because no one can collude with himself. In this way, the essential elements of a company is not “legal personality”, but “control” in the eyes of the competition law. The top-down control model cannot be applied to the blockchain because none of the three entities in the blockchain, namely a core developer, a miner, and a user, control blockchain.

Professor Schrepel introduced his thesis, the Theory of Granularity: A Path for Antitrust in the Blockchain Ecosystem,<sup>3</sup> which reconstructed the competition law based on the need to solve this problem. The thesis first analyzed the phenomena created by the dynamic relationship between core developers, miners, and users in the current blockchain from the macro perspective. (1)The core developer designs the blockchain and sets the principles and rules, so it has its own power, but once the development is complete, the developer no longer has control, and the user or miner’s consent is required to change the rules later. (2)Users can choose which blockchain to use, but it is difficult to get involved in the design of the blockchain itself. They usually work individually, but they also form a group, which may result in anti-competitive behavior. (3)The miner plays the role of verifying the protocol of the transaction information on the blockchain but is not involved in the role of other actors.

---

2 Coase, R. H., (1937), “*The Nature of the Firm*”. *Economica* 4.16: 386-405.

3 Thibault, S., (2020), “*The Theory of Granularity: A Path for Antitrust in Blockchain Ecosystems*”. Available at SSRN: <https://ssrn.com/abstract=3519032> or <http://dx.doi.org/10.2139/ssrn.3519032>



Session 7 Capture (caption)

The dynamics between these participants reshape the legal agenda. Professor Schrepel called the application of the concept and logic of competition law by analyzing and comparing these situations as “the theory of granularity” according to the title of his thesis. In other words, he argued that the competition law should focus on the dominance of the dynamics formed between participants instead of companies. He believed that through this, competition law can be applied although companies do not exist in the blockchain ecosystem, leading to more people’s participation to the blockchain, and increased social utility.

Professor Schrepel continued the discussion on collusion and abuse of market dominance with regards to blockchain. First, collu-

---

sion can be facilitated because the parties can trust the implementation of the agreement due to the basic operating principle of smart contracts. For example, if a collusion agreement is reached but a business operator cheats by lowering the price by 5%, the smart contract automatically applies a penalty clause to him. At this time, the smart contract can never be deleted. If business operators are concerned about this issue, they will not collude with smart contracts. If parties think that they will not be caught because smart contracts are coded, they will use smart contracts for collusion. In other words, due to the emergence of smart contracts, conventional non-cooperative games are converted to cooperative games.

Regarding the abuse of market dominance, Professor Schrepel said that public and private blockchain should be differentiated, and that there is a great potential for problems in private blockchain. This is because, unlike public blockchain where all participants are involved, private blockchain allows the gatekeeper to arbitrarily change rules and control access. This discussion can be commonly applied to other cases of competition law such as tying, and other legal frameworks such as intellectual property law.

Professor Schrepel presented two main tasks at the end of his presentation. The first was expertise. He said that one needs to fully understand the blockchain and its principles to enforce the law. This is because new problems that are different from the existing ones may arise due to the technical characteristics of the blockchain. Next, in applying the law to the blockchain for enforcement, we must understand that the blockchain industry has its basic philosophy to not be bound by the existing legal and social regulations. Therefore, it will not be easy to apply the law to the industry, but things can change if incentives are provided accordingly.

# When Does Gaming Justify Algorithmic Secrecy?





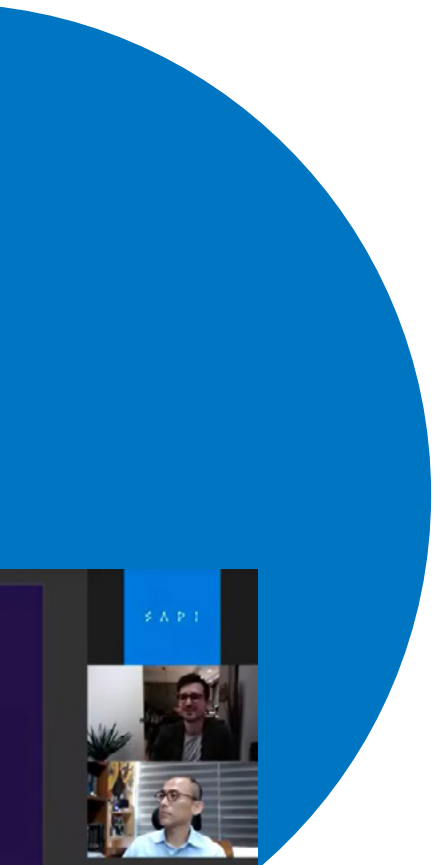
Presenter

Ignacio Cofone,  
McGil University

# When is Algorithmic Secrecy Justified?

Ignacio N. Cofone (McGill University)  
Katherine J. Strandburg (New York University)

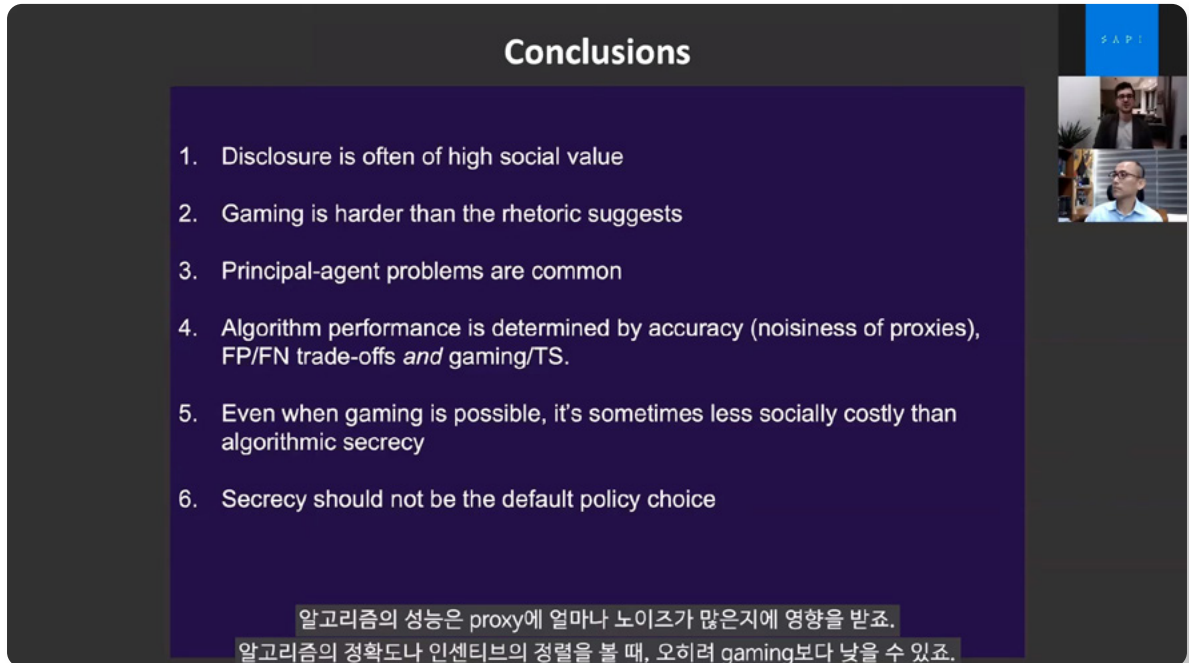
Montreal-Seoul, 29 September 2020



In the last session, artificial intelligence and transparency was discussed as the main topic. Professor Ignacio Cofone started the seminar by presenting a picture of the chess artificial intelligence “The Turk” that existed in the 19th century. Just as “The Turk”, which was the motive of Amazon’s crowd-working platform Mechanical Turk, was actually a trick in which human was behind the so called artificial intelligence, the interaction with artificial intelligence we encounter today may be reduced to the interaction with the person who programmed and designed the artificial intelligence model.

Today, algorithmic decision through artificial intelligence is used in various fields such as hiring, finance, and criminal proceedings. These decision-making models mimic the decision-making method of humans by using various features as a proxy for specific outcome values. For example, when determining whether to repay a loan based on income, the feature of income was used as a proxy variable for the result of repaying the loan. The discussion of algorithmic transparency is primarily a discussion of whether to disclose information on features and proxy variables used in these algorithmic decisions.

So, what are the trade-offs to consider regarding algorithmic transparency? First, algorithm transparency has the advantage of im-



**Conclusions**

1. Disclosure is often of high social value
2. Gaming is harder than the rhetoric suggests
3. Principal-agent problems are common
4. Algorithm performance is determined by accuracy (noisiness of proxies), FP/FN trade-offs *and* gaming/TS.
5. Even when gaming is possible, it's sometimes less socially costly than algorithmic secrecy
6. Secrecy should not be the default policy choice

알고리즘의 성능은 proxy에 얼마나 노이즈가 많은지에 영향을 받죠.  
알고리즘의 정확도나 인센티브의 정렬을 볼 때, 오히려 gaming보다 낮을 수 있죠.

Session 8 Capture (caption)

proving compliance<sup>1</sup> and facilitating the correction of errors and biases of the algorithm. It also guarantees the procedural rights of individuals receiving algorithmic decisions. On the other hand, the industry strongly prefers algorithmic secrecy. They argue that if the algorithm is disclosed, there may be problems such as users “gaming” the sys-

---

1 Here, compliance refers to the actions a user performs to achieve the desired outcomes, provided that he/she knows how algorithmic decision-making works. For example, credit card user tries not to delay the card payment when it is revealed that making timely credit card payment has a positive effect on personal credit rating. Gaming, which will be seen below, is similar, but unlike compliance, because while compliance is an action commensurate with the intended purpose of the designer, gaming is an action that uses the information against the original purpose of the designer.



---

tem, or competitors free-riding on others' trade secrets. Therefore, in determining whether or not to disclose the algorithm and the degree to which it should be disclosed,<sup>2</sup> Professor Cofone insisted that the social cost of disclosure (the issue of free riding of trade secrets and gaming) and the social utility of disclosure (compliance, improvement of errors and biases, and guarantees of users' rights) should be both weighed and considered.

Can a private business entity adequately reflect these factors? Professor Cofone answered no. This is because private businesses do not internalize the social value of compliance given by algorithm disclosure or social costs due to errors or biases in algorithms, but rather have an incentive to hide algorithms in order not to get caught in regulations or disputes. Therefore, Professor Cofone argued, it is more appropriate for a judge or government authority to decide whether to disclose the algorithm.

Professor Cofone suggested that three factors should be considered when deciding whether to disclose an algorithm in a specific case. First, whether disclosure of the algorithm will incur social costs such as gaming and trade secret leakage. If the entire code is disclosed, there may be a concern about the leakage of trade secrets to competitors, but disclosure of features is less dangerous in this respect. If features that cannot be easily changed<sup>3</sup> by users play an important role in algorithmic decisions, gaming is unlikely to occur,

---

2 Professor Cofone points out that the disclosure of algorithms can also take place at various levels. Information subject to disclosure can vary including (1) training dataset, (2) source of training dataset, (3) code, (4) model, (5) features and labels, (6) weight of features and labels, (7) types of output values, and (8) the final goal of the algorithm.

3 E.g. a height and address of an individual

even if these facts are disclosed. Also, even if the user changes his or her features to change the algorithm's decision, it is not a gaming issue if it actually changes the algorithm's decisions by making positive changes.<sup>4</sup>

Second, whether a loss such as a decrease in prediction performance is expected due to algorithm disclosure must be considered. If algorithmic decision-making is working well due to the high accuracy of the algorithm, there is a concern about disclosing it because gaming may harm the predictive accuracy of the algorithm. On the other hand, if the algorithm's accuracy is low, the loss due to gaming is expected to be relatively small, so it is conversely possible to expect an error correction effect by revealing the algorithm. The improvement of such inaccuracies is also important for improving distributional inequality.

Third, it is necessary to examine whether the algorithm designer's incentive and the social welfare are aligned. For example, in the case of an algorithm that determines the risk of recidivism of a person subject to parole, the algorithm designer is likely to minimize false negatives,<sup>5</sup> but social preference will prefer an algorithm designed to

---

4 This is a compliance problem previously mentioned. Gaming is harmful when the result is made to look as if it is improved by using a loophole in the algorithm while the actual objective result value (i.e. personal credit) has not improved.

5 An error in which what is true is judged as false. The case in which a person at a high risk of recidivism is released on parole is one of the examples. If a person released by algorithmic decision commits a crime, the reliability of the algorithm may be called into question, but a person who is not released cannot commit a crime, so the algorithm designer has an incentive to minimize the release of high risk individuals.

---

minimize false positives<sup>6</sup> based on the constitutional principle. Professor Cofone pointed out that if the algorithm designer's incentive and the social preference are aligned, it would be less useful to disclose the algorithm, but if the two incentives work in opposite directions, the disclosure of the algorithm can increase the social utility. Professor Cofone said that there are many cases in which algorithm disclosure improves social utility, and the significance of this study lies in that it has suggested criteria for determining in which cases the algorithm should be disclosed. He concluded the seminar by emphasizing that algorithmic secrecy should not be the default, because gaming is more difficult to operate in reality than people may think, and even if gaming occurs, the cost of not disclosing the algorithm may be higher.

---

6 An error in which what is false is considered as true. The case in which a person at a low risk of recidivism is denied parole is one of the examples. In the case of the COMPAS, the false positive rate of blacks was higher than that of whites, which became an issue.

## Seoul National University AI Policy Initiative (“SAPI”)

---

SAPI is an initiative launched by the SNU Center for Law & Economics to conduct research and address socio-economic, legal, and policy issues related to artificial intelligence.

SAPI aims to be a ‘Social Lab’, where scholars from diverse disciplines collaborate to conduct interdisciplinary research. SAPI is co-directed by Professor Haksoo Ko and Professor Yong Lim of Seoul National University School of Law.



Seoul National University

AI Policy Initiative