【CCRC インタビュー】 法学部　大屋雄裕教授 「法哲学の視点で AI を考える」

まだないものに対する法のあり方を考える

2022 年 11 月の ChatGPT の公開以降、AI に関するニュースを見聞きする機会は一気に増え、政府関係者を始め多くの有識者らが AI に対する見解を述べている。AI すなわち人工知能自体は何年も前に登場し、すでに私たちの生活に深く浸透しているが、一昨年登場した生成 AI はそれまでの AI と大きく異なり、私たちの想像を大きく超える能力を示してきた。できることが限定されない AI を「万能 AI」や「汎用性 AI」と呼ぶが、ChatGPT を代表とする生成 AI は完全な万能 AI ではないにせよ、その要素を大きく持ち合わせているという。来る将来、登場するであろう万能 AI への期待とリスクへの懸念が世界中で膨らんでいる。

そうした中、日本 AI に対しどのような姿勢をとるかについては、AI 事業者だけでなく、1ユーザーとして誰もが注目すべき事項だろう。今回インタビューをさせていただいた大屋雄裕教授は、総務省「AI ネットワーク社会推進会議」構成員、内閣府「人間中心の AI 社会原則検討会議」構成員など、AI の利活用に関する政府レベルの会議に数多く関わってこられた法学者である。大屋教授の専門分野は法哲学で、まだ世に出てきていない万能 AI を対象とする規制のあり方については、法律が無いところから概念形成を行う必要があることから法哲学的なアプローチが求められたということになる。

多くの人にとって聞き慣れない言葉である「法哲学」とはどのような学問であるのか。法学にも哲学にも疎い自分が理解できるか不安な状態で臨んだインタビューだったが、お話を伺い「そもそも」から始まる問いを起点に、法律に関連した用語や概念を考え直す学問であることを知り、身構える必要はないことに気付いた。万能 AI の登場と言う大きなパラダイムシフトの前段階にいる私たちにとって、哲学の手法である「概念を考え直すこと」は必然的に求められていくのではないだろうか。
今回は AI 規制策定の最前線にいる大屋教授に自身のご専門である法哲学についてまずお聞きしし、後半からは日本における AI 規制の状況について伺った。

根本を考え直す法哲学の手法

（聞き手：鈴木）
法哲学とはどのような学問なのでしょうか。

（大屋教授）
カテゴリーとしては法律学の中に含まれます。法律学は実定法学と基礎法学から成り立っていますが、実定法学では「法律はある」という前提からスタートします。国会で成立されたすでにある法律に対し、それを個々の実際の事例にどうあてはめられるか、解釈をどうすべきかについて議論するのが実定法学です。

実定法学には憲法や民法、商法が含まれ、これらはすでに存在する法律をもとに、その解釈を争う学問として成立していますが、法は時代によっても国によっても大きく異なります。そのため理学部や文学部の方々からすれば、実定法学は普遍性のある真理の探究ではないから学問ではない、ということになるかもしれません。

法律の解釈に関して、この国で何が正しいかは最高裁判所が決めます。学会の大半において「この解釈が正しい」という結論が出ていたとしても、最高裁が異なる判決を出せば、最高裁が出したその答えが「正しい」とされます。法律家が異論を唱え続けることはできても、現実に機能するのは最高裁による決定だということです。その意味では確かに、実定法学は学問的ではないのかもしれません。

一方の基礎法学についてですが、その中でも法社会学が典型的な基礎法学にあたります。社会学の手法を使って、法的な現象の解明に取り組むのが法社会学です。例えば、ある意図をもって法律を改正した場合、それが意図した通りに実際なったか否かについては実定法学ではわかりませんが、社会学の手法である統計分析などを使うことで明らかにできます。ここで導き出される結論は普遍性のあるものなので、真理を探究する学問としての条件は満たしていることになります。法哲学も基礎法学のひとつで、哲学の手法である概念分析を主に使うことで、法に関連した用語や概念を考え直します。法とは何なのか、法解釈とはどういうことなのか、権利とは何を意味するのか、そういった「そもそも」の問いをもとに考えを深めていきます。

「人権とは本当に権利なのか」について考えてみます。人権と言うと権利っぽいニュアンスがあるので、ほとんどの人は人権を権利と思っていますが、実は他の権利とは全く異なる性質がいくつかあります。大きな違いとして挙げられるのが、通常の権利では名宛人[i]が存在しますが、人権の場合は存在しないということです。

権利と義務は通常対になっており、私の権利は相手の義務で、相手の義務は私の権利であるという関係になっています。売買関係をイメージしてもらえるとわかるでしょう。自分が代金を払う義務を果たしていないのに、商品を渡す義務を果たせと相手に要求することはできません。

ところが人権の場合は対応する義務がありません。「自分がまず義務を果たさないと人権が保障されない」ということにもなりません。人権は万人が無条件に尊重し、尊重されなければならないものだとされています。また、通常の権利は放棄したり譲渡したりできるものですが、人権は放棄することも譲ることもできません。このように異なる性質を持つにも関わらず人権を権利と捉えて話を進めると、話がどん詰まることが多いので、そもそも人権を権利と捉えないほうがよいという考え方を私は主張しています。

（鈴木）
当たり前に使っていた用語や概念について考え直す。確かに哲学的ですね。

（大屋教授）
はい。ですので、こうした「そもそも」の話を扱うのが法哲学と言えます。また、「そもそも権利とはこういうものであるから、私たちの社会にはこういった法が必要だ」といった主張をすることも可能で、すでにある法をもとに議論する実定法学の人たちとは違い、あるべき法の姿を議論することも法哲学の役目です。

（鈴木）
理想を掲げても良いのですね。

急速に発展する AI を前に

（大屋教授）
はい。近年私は AI 倫理関係で忙しくしていますが、その理由の一つがこれで、AI に関する法律がまだないということです。AI 自体はすでに何十年も前から開発されていますが、AGI と呼ばれる万能 AI についてはまだ出てきていません。しかし 2040 年か 2050 年には出てくるだろうと言われており、その時に万能 AI もカバーできる法律がない状態で大丈夫なのか、どのような法を用意しておくべきかといった話し合いが現在進められています。まだ現れていないものについて、どのような対策を立てるのが正しいのかという議論をしなくてはならないので、まだないものについて考えて概念を規定する法哲学的なアプローチが求め

られたことになります。

（鈴木）
先生は、AI への規制についての話し合いの場にも携わっていらっしゃるのでしょうか。

（大屋教授）
はい。日本でも、AI に対して何かしらの規制が必要であることは、政府レベルでの共通理解になっています。5 年くらい前の OECD（経済協力開発機構）[ii]の場で、AI に対し、ガイドラインとして緩やかな規制を始めることについて日米欧による合意がなされました。そこから EU は近年、ガイドラインでは不十分として AI 法制定を推し進め、もうすぐ可決という段階に来ています。日本は、去年の広島 G7 サミットにおいて、日米欧で手を取り合って AI 規制を進めていきましょうという合意をしましたが、法としてではなく、とりあえずガイドラインとして規制を進めていくという意向を示しました。

（鈴木）
法とガイドラインはどう違うのでしょうか。

（大屋教授）
いわゆるハードローとされる法は、議会で制定される強制力を持つもので、違反すると罰金などの罰則が課されます。一方のソフトローとされるガイドラインには強制力はなく、みんなで守りましょうねと言い合っている規定にすぎません。違反しても世間から指を指されて怒られる程度です。日本での AI 規制に関する議論ではソフトロー派が多い印象です。その理由として技術の発展が早すぎることが挙げられます。日本において法律は、一度作ったらなかなか変えられないものなので、まだ具体的にどのような技術かわからないのに法律をつくって的外れなことになったらどうするのか、そういった意見が強くあります。

（鈴木）
近年登場した Chat GPT は万能 AI ではないのでしょうか。

（大屋教授）
すでに広く使われている AI は特化型 AI といって何かしらの用途に特化した AI ですが、Chat GPT は特化型からやや飛躍した立ち位置にあります。Chat GPT は文章を生成するという機能しかありませんが、プログラミング含めありとあらゆる文章を書けるので、応用範囲が極めて広く、さらにマルチモーダル[iii]といって画像入力や画像出力もできるようになると、特化型の性質を持ちながらも非常に汎用性が高い AI と言

えます。この 1 年で AI 技術は急速に成長しており、将来万能 AI が登場することは夢物語ではなくなっています。

AI に対する国レベルでの認識の違い

（鈴木）
EU の AI 法は、万能 AI を見据えたものなのでしょうか。

（大屋教授）
そこまではいかないと思います。EU の AI 法は、AI の種類を 4 段階にランク分けし、ランクごとに規制度合いを設定しているのが特徴です。サブリミナルに相手の思考をコントロールするような AI は 4 段階の一番上に位置付けられて禁止対象となり、人の行動に影響を与える可能性が高いとされた AI は高リスク AI としてその下に位置付けられ、適合性評価が課されます。その下に限定リスク AI、最小リスク AI と続きます。ただ、明らかに禁止 AI と見られるものはすぐに判断がついて問題ないでしょうが、それ以外の AI について、適合性評価で危険性を事前に判断するのはそう簡単ではないのではと見ています。

（鈴木）
まだ事故などの問題が起きていない時点で、その AI の安全性や危険性を評価するのは難しいということでしょうか。

（大屋教授）
はい、ChatGPT のような対話型 AI でも、それに誘導されて自殺したとされるケースが出ています。サブリミナルに相手を誘導する AI の危険性は明らかなので直ちに禁止されるでしょうが、そうではない AI でも使い方や使う状況によっては危険な事態を生み出しうるので、それらが低リスク AI と評価されてしまわないか懸念が残ります。また公共機関が個人をスコアリング（点数評価）するものは禁止 AI、重要な公的・民間サービスの提供に結び付くものは高リスク AI に位置付けられていますが、これもスコアリングの種類によるのではと疑問視しています。EU が禁止を想定しているのは、中国の信用スコアでしょう。利用者を AI による判断で 350～950 点に点数付けするものですが、あれが中国で一気に普及したのは納得がいきます。

（鈴木）

なぜでしょうか？

（大屋教授）
中国では基本的に他者というものが信用されていないと言われます。それは、国家が社会の安全を確保してこなかったため、見知らぬ他者でも信頼できるという状態が成立しなかったからです。このため、親族だから、長い付き合いの知人だから信用するといったように、人的ネットワークを通じてしか評価が行なえませんでした。ところが公的な機関は、こうした私的なネットワークに頼れず、一人一人の国民の信用度を把握できないため、公共図書館で本を貸す場合にも持ち逃げされることを恐れて利用者にデポジットを課していましたし、病院も患者に対して治療前にデポジットを要求するのが一般的でした。デポジットとして払ったお金は何もなければ戻ってきますが、何をするにもまずお金を積まないといけない社会でした。そこに登場したのが信用スコアで、これにより「あなたは 850 点だからデポジットはいりません」といったサービスが可能になったのです。

（鈴木）
その信用スコアは公的なものなのでしょうか？

（大屋教授）
もともとはアリババという中国企業が、グループ会社内でお金を貸す際の金利を決める判断材料として始めた私的なシステムでした。しかしスコアがスマホの画面に表示されるシステムだったので、図書館や病院を始め、あらゆる機関や企業が「XX 点あればデポジット不要」「XX 点あれば入会可能」というように利用し始め、アリババグループ以外でも一般的に使えるツールとして一気に普及しました。それまでお金を積んで「自分は信用できる人間ですよ」と証明しなければならない社会だったのですが、一定の信用がある人々であればそこから離脱できるようになったわけです。中流階級であれば通常 600 点は取れるようで、そうすると楽な方に多くの人が流れるのは当然でしょう。こうした背景があったために中国では信用スコアが一気に普及しましたが、ホテルに泊るときにデポジットを要求されることがほとんどないような日本社会では導入するメリットがありません。また信用スコアというのは低ランクの人を阻害するシステムですから、それが公共サービスに結び付くと危険です。EU の AI 法が禁止としたのもそれが理由でしょう。

（鈴木）
AI 開発における中国の勢いが感じられます。ところで日本は AI 対して、法ではなくガイドラインで規制をする方向にあるということでしたが、ガイドラインはどのくらい出来上がっていますか。

（大屋教授）

私が今関わっている統合ガイドライン[iv]はほぼ出来上がっています。もともと AI に対するガイドラインは 2～3 年前に総務省と経産省がすでにそれぞれ別のものを作っていました。経産省の方は、AI が関係する契約に対するガイドラインを、総務省の方は AI の利活用に関するガイドラインを完成させていましたが、昨年この二つのガイドラインを統合することが決まり、その上で欧米との相互運用性があるものに直すための話し合いが進められてきました。

（鈴木）
日本も今後 EU のように法律で AI を規制する可能性はありますか？

（大屋教授）
近い将来ではそれはないと思います。ただ、EU が AI 法を作ったので歩調を合わせるために法を作る必要性が出てくるかもしれません。EU が GDPR（一般データ法規則）を作ったあと、日本でも同じレベルの規制をしていることを EU 諸国に認めてもらわないと個人データの越境が制限されることもあり、日本は個人情報保護法を改正して規制レベルを引き上げました。その時と同じで、EU の AI 法の域外への影響が出てきた場合には、日本でも法律化しないといけなくなるかもしれません。一方のアメリカでは、AI に対して産業としての新興を望む勢力の方が強い状況です。アメリカは AI 開発において世界トップを走っており、自由に開発を進めたいためです。

AI になにをどれだけ任せるか

（鈴木）
先生は AI 規制のあり方を考える時、AI に判断を委ねることと、人間が決めることのバランスをどう考えていらっしゃいますか。

（大屋教授）
現代では私たちがしなければいけない決定が増えすぎているので、重要な場面で人間の判断能力を最大限に発揮させるためには、さほど重要でない場面で AI に判断を委ねたほうがいいと考えています。レントゲン画像を診る医師が、何千枚も一日中見続けていたら目がかすんで判断が鈍る可能性があります。そこで、医師による目視の前段階で AI を使い、明らかに問題ないと判断された画像を抜いておけば、医師のパフォーマンスを最大限に発揮できるでしょう。この複雑な社会においては AI による代替を用いて、人間の判断が必要な領域に人間の力を集中させるべきと見ています。

行政の現場において AI をどれくらい使って良いかについても問題になっています。国税庁ではすでに AI が使われており、膨大な税務申告書類の中から、脱税の可能性があるとして調査する対象を抜き出す作業に AI が活用されています。しかし、仮にその判断基準に申告者の職業が用いられていたとしたら、それが差別にあたらないかといった問題が出てくるでしょう。

これまで、行政の業務は人間が担うものという価値観が強くあり、音声の自動文字起こしすら及び腰の自治体が少なくありませんでした。他方、非常に野心的に取り組んでいるのはいいとして、そこに潜んでいるかもしれない問題性に気付いていないような例も見られます。そのため、どの業務なら AI に任せて良くて、どの業務は AI に任せてはだめだということを明らかにしないと、いつまでたっても及び腰の自治体は及び腰のままで、逆に AI を利用しすぎる自治体はその傾向が助長されるでしょう。

近年、保育園の入所決定において AI を使う自治体が増えています。これまでは何人かの職員が何日もかけて行っていた作業を AI に任せたらあっという間にできたそうです。この保育園の入所決定において AI 使用が許されているのは、それが確定的なプロセスに基づいたものだからです。ここには明確な判定基準があって、片親だったらプラス X 点、仕事をしている親ならプラス X 点と点数が加算され、高い点数の家庭から優先的に希望の保育園が割り振られます。この点数をもとにして、どのようにすれば入所希望家庭全員の満足度を高く割り振れるか、これをマッチング問題と言いますが、この最適解を導くには大量の計算をする必要があります。そのため職員が計算していた時は何日もかかっていたわけですが、計算はコンピューターが最も得意とする作業なので、AI に任せるのが有効な業務と言えます。

一方で刑事事件の量刑を考える場面を考えてみます。裁判官には量刑を決める裁量が与えられていますが、すでにデータベースから類似判例を抜き出し、ある種の相場が見えるようにしてくれるシステムは、すでに日本の裁判所にも導入されているそうです。ただし、事件にはデータベース上に挙がらない情報というものが必ずあり、それも考慮して量刑は決められるべきです。AI に任せて人間は関与しなくていい領域と、AI に全て任せることで問題が生じうる領域とがあり、後者の場合は AI の利用を補助レベルに留めておく必要があります。そこでは AI による判断の後に、人間の判断があって最終決定となるべきで、そういった文脈も現在進めているガイドラインには盛り込まれています。ちなみに、先にお話しした経産省と総務省の統合ガイドラインは AI 事業者を主な対象としたものですが、いまお話ししたような行政を対象とする議論は総務省行政管理局で進められています。私は双方の議論に関わっていますが。

（鈴木）
現在進められている AI に関するガイドラインの骨柱となる考えはどのようなものですか。

（大屋教授）
日本では 2018 年に「人間中心の AI 社会原則」というものが作られましたが、そこにも示されているように、人間中心主義を維持すべきという考えは政府レベルでも共有されています。AI 開発に向けた技術革新の流れを止めないよう自由にさせようという姿勢と、個人の尊厳を守るためにある程度規制をしようという姿勢とのバランスをとったガイドラインと見ています。

日本ではガイドライン策定を進める際に、AI 事業者、ユーザーの代表として消費者団体の関係者、AI 技術者や研究者、法律家や経済の専門家を呼び集めて、各々から意見を聞きます。これをマルチステークホルダー・プロセスと言いますが、様々な領域の人たちを関係者として取り込んだうえで、みんなが大体納得できる落としどころで決めるというやり方が一般的です。悪く言えば中途半端な内容に留まると言えますし、良く言えば偏りのないバランスの取れた内容に収まると言えます。EU のように人間中心主義を高く掲げて厳しく規制するのも一つですが、逆に言えば柔軟性がないということになります。冒頭でも触れましたが、日本の場合は一度法律を作るとなかなか変えられない仕組みになっているので、もし変えようとなったとしても改正までに 10 年かかることは十分ありえます。そういった事情があるので、様々なステークホルダーが「まぁこれでいいだろう」と納得した上で決まったガイドラインで上手く行くのであれば、法律を作るまでのことはしなくていいというのが日本の AI 規制に関する今の姿勢です。ただガイドラインは強制力がないので、やりたい放題と批判されているアメリカの巨大 IT 企業などがどこまでそれに従うかは懸念されます。

（鈴木）
詐欺広告に対する SNS 事業者の対応について問題視する声が大きくなっていますが、海外企業がどこまで日本のガイドラインを遵守してくれるのか、私も疑問が拭えません。AI 規制において先行を走る EU の動向を見ながら、日本としての規制の動向を一ユーザーとして今後も注視していきたいと思います。

（聞き手・文章　鈴木薫）

大屋雄裕教授 Profile

慶應義塾大学法学部教授。東京大学法学部を卒業、同大学助手、名古屋大学大学院法学研究科助教授等を経て、2015 年 10 月より現職。専攻は法哲学。

著書に『自由とは何か：監視社会と「個人」の消滅』（ちくま新書、2007 年）、『自由か、さもなくば幸福か？二一世紀の〈あり得べき社会〉を問う』（筑摩選書、2014 年）、『法哲学』（共著、有斐閣、2014 年）等がある。

内閣府「人間中心のAI社会原則会議」構成員。総務省「AI ネットワーク社会推進会議」構成員。

---

[i] その義務を履行しなければならない者のこと

[ii] ヨーロッパ諸国を中心に日・米を含め 38 ヶ国の先進国が加盟する国際機関。

[iii] テキスト、音声、画像、動画、センサ情報など、2 つ以上の異なるデータの種類から情報を収集し、それらを統合して処理する機能

[iv] 「AI 事業者ガイドライン（第 1.0 版）」2024 年 4 月 19 日発表