座談会「情報法制の可能性について―AIをめぐる動向を中心に―」

九州大学大学院経済学研究院教授（当時。2017年4月より中央大学総合政策学部教授）　実　積　寿　也
東京大学大学院工学系研究科准教授　鳥　海　不二夫
（司会）　東京大学大学院法学政治学研究科教授　宍　戸　常　寿

1　はじめに
2　AI研究の現状
3　意識とプログラミング
4　強いAIと弱いAIの区別
5　AIと法的責任
6　総務省のAI開発ガイドライン
7　ガイドラインのグローバル化は可能か
8　予測に基づくガイドラインの限界
9　研究・開発と利用の区別
10　AIネットワークシステムとは
11　AIネットワークシステムの制限可能性
12　研究と社会システムの対話，AIの定義
13　透明性，セキュリティ，プライバシーの原則
14　利用者支援の原則
15　アカウンタビリティの原則
16　連携の原則
17　むすびに代えて

1　はじめに

宍戸：『情報法制研究』創刊号の企画として，今回は，実積さん，鳥海さんと私で座談会をさせていただきます。まず簡単な読者向けの自己紹介として，情報政策との関わりを中心に，実積さんからお願いいたします。

実積：九州大学の実積です。経済学部で通信政策を中心に教えていまして，最近は，ネットワーク中立性という問題に興味を持っています。文系人間なもので技術的な話は正直門外漢なのですが，経済活動の観点から妥当な政策形成を情報通信分野で実現することに貢献していければと思っています。

宍戸：ありがとうございます。次に，鳥海さんからお願いいたします。

鳥海：東京大学の鳥海です。私は完全に理系側の人間でして，計算社会科学と呼ばれているコンピュターサイエンスと社会学を結び付ける研究をメーンでやっています。たとえば，未成年者の安心安全なネットワーク利用に向けたコミュニティサイトのデータ分析や未成年者のネットワーク利用におけるリスク管理といった研究を情報法制研究所理事の江口清貴さんと一緒にやっていたりします。また，AI関係の活動としては，人工知能学会の編集委員をやっております。

宍戸　常寿

宍戸：ありがとうございます。最後に私ですけれども，東京大学で，法律，特に憲法と情報法を担当しています。表現の自由と通信の秘密，プライバシーの保護が，それぞれ憲法の基本価値としてあり，その具体化という観点から情報法に関わっています。基本的には，総務省関係の研究会とか，事業者・民間団体に有識者として呼んでいただいて，耳学問をしています。

情報法制学会では，このように，法律，経済，ICTの技術に関わる研究者，事業者がそれぞれの知見を持ち寄って，情報法制の健全な発展に資することを目的にしていますが，今回はAIをめぐる動向を中心に議論したいと思います。

実 積　寿 也

2　AI研究の現状

宍戸：そこでまずは現在のAI研究の状況について，鳥海さんから，話題提供をお願いします。

鳥海：はい，分かりました。現在のAI研究の現状について簡単にお話しします。今は第3次人工知能ブームと呼ばれています。その前の第2次というのが30年ぐらい前ですかね，第5世代コンピューティングとかをやっていた時期なんですけが，そのころ人工知能がいったんブームになり，人工知能あまり使えないじゃないかということで冬の時代と呼ばれるのが何年か続きました。そして今，ディープラーニングあるいは深層学習と呼ばれるものが使えるぞということが分かってきて，3回目の人工知能のブームになっています。その中で，AIがいろいろ問題を起こすんじゃないかという危惧があり，こういった情報法制に関係するような話が出てきてるんじゃないかなと思っています。

そもそも，この人工知能って何だ？というのが，技術者や研究者と，そうでない人たちの間でかなり乖離がある気がするんですね。どうも皆さん，自分のイメージするAIで全て語ってしまうので，それは結構問題かなと思っています。

人工知能と呼ばれるものには大きく2種類あって，1つはただの技術であり，今までの技術の延長上にあるものです。そして，もう1つはドラえもんや鉄腕アトムといった夢のAIです。この夢のAIと今の技術との間には，ものすごい乖離があり，実は今の技術が発展していってもドラえもんはできないだろうというのが，人工知能研究者の意見なんですね。

にもかかわらず，人工知能の問題というと，ドラえもん的なものができたときにどうしよう？といった議論がされるんですね。そうすると，10年後の世界の話をしてるのに，突然300年後の世界の話を混ぜてきていることになってしまう。なので，その辺を整理すると，いろいろな議論がうまくできるようになるんじゃないでしょうか。

夢のAIと現在のAIは，「強いAI」と「弱いAI」と呼ばれ，区別されています。これは別に，殴り合って強い弱いではありません。

強いAIというのは「意識を持っているAI」という定義になります。人工知能自身が自分で意識を持って活動する自律型になって，人間の手を完全に離れて活動する。まさにドラえもんとか鉄腕アトムなどが強いAIになります。

一方，弱いAIというのは，人間が知的に作業していることを賢く代替する技術です。たとえば車の運転をするAIは明らかに意識を持っていません。あるいは，囲碁で世界チャンピオンを破ったAlphaGo（アルファ碁）がありますが，これも人間よりも賢いんですが「囲碁ができたから，次は将棋を指そう」という意思は持ってないわけですね。そのため，AlphaGoも弱いAIです。

この，強いAIなのか，弱いAIなのかということを混同すると議論がうまくいかなくなります。今現在できるのは弱いAIだけです。強いAIは人工知能研究者の中でも実現への道は遠いと考えられています。ですので，自律型のAIというのは当分出ないといえるでしょう。

3　意識とプログラミング

実積：「強いAI」と「弱いAI」の区別の基準となる「意識」とはどういうものとして考えているんですか？

鳥海：いい質問ですね。それはかなり定義が難しく，全然解決していない問題です。だから，まだそこの，意識とは何か，何ができたら強いAIになるのかということも分からないぐらいのレベルです。

実積：ルンバは，たぶん弱いAIという話だと思うんですけど。物にぶつかると避ける方向を機械自身が決めているという意味では，意識があるよ

うにも見えますが。

鳥海：でも，あれはプログラミングですからね。

実積：意識とプログラミングは違うってことなんですね。

鳥海：まあそうですね。そこの話は，ほんと難しいんですよ。人間は意識を持っています。じゃあ，イヌは意識を持ってるか。たぶんイヌは持ってるような気がする。すると，セミは意識を持っているのか。セミって持ってるんですかね？セミが土から出てきて蛹になるという行動は考えているわけではなく，DNAにプログラミングされているだけですよね。同じように，赤ちゃんが生まれてきてすぐ泣くのは，泣こうと思って泣いてるわけじゃないから，あれは意識じゃなくて，もうプログラミングであるというように考えることもできるわけですね。そういう意味では，うちの1歳の息子は意識を持ってるのか。最近なんか持ちだした気がするんですけど。

実積：そうすると，「意識がある」と評価されるためには，「人間並み」という基準に合格する必要があるという意味ですか？

鳥海：人間並みというと，サルは意識持ってないことになっちゃうじゃないですか。イヌもそうですけど。

宍戸：今の話は非常に重要だと思うんですけれども，たとえば自動走行で，よく議論されるトロッコ問題も，基本的にはプログラミングの問題であって，意識の問題ではないということですか。

鳥海：トロッコ問題に関しては，意識を持ったところで人間が解決できていない問題なので，AIに解決させるべき問題ではないと思います。哲学の話なので，人間がまず解決してから，プログラミングすればよいのではないでしょうか。

宍戸：そうすると，意識も，かなりの部分哲学問題ではないでしょうか。本当に意識があるのかは，哲学で独我論的な立場に立てば分からないわけですよ。私から見ると，鳥海さんや実積さんが意識主体なのか分からない。

鳥海：哲学的ゾンビの話ですね。

宍戸：そうです。私がこれは意識がある存在だと思うから，人間，あるいは同じような意識を持った生命体として扱うわけですね。同じことが実は人工知能にもいえるのではないか，と文系的には

鳥 海　不二夫

思う。だんだん人間と挙動が同じようになってきて，厳密には分からないけれども，一定程度了解可能な形で，行動の選択肢を選んでいるらしいという存在が，人間とあまり変わらないレベルであれば，もうすでに意識があると人間が思って，そのように扱うことはできるんじゃないか。そうすると意識は，技術的あるいは科学的な水準でなくて，いわば相対的にというか，社会の受容性の問題ではないかという気がするんですが，AI研究者としてはどう思われますか。

鳥海：まさにそのとおりだと思います。だから，そこはかなり難しい問題で，どこまでいったら強いAIなのか議論というのは，実は先ほども言いましたけど，全く解決してない問題なんですね。だから，クラゲに意識があるのかを定義できない限り，AIに意識があるのかは定義できない。

たとえばドラえもんも，内部機構的にはたぶんパラメーターがあって，プログラミングされていて，ディープラーニングを使ってるかは分かんないですけど，神経系ネットワークがあって，ちゃんと解析すればすべての行動が説明できるかもしれない。ただ，それは人間の脳も同じですよね。だから，ある程度いくと，そこから先はもう哲学的な問題にしかならないんですね。

ただ，現在のAIはどうかというと，やはり意識というには程遠いですよね。たとえばクラゲが意識を持っているとはちょっと思い難い。クラゲは意識を持っているという人もたまにいるので，インフルエンザウイルスだとどうか考えてみると，さすがにちょっと厳しい。あれは機械的に構造を変えていくだけなので，意識的にやってるわけで

はないですよね。それと同じように，グレーゾーンがあるにせよ，今は意識がない状態だというふうにいえるだろう。

実積：意識のあり・なしは明確な閾値があって峻別できるものではなくて，中間的な状況があるということですね。

鳥海：今のところ濃淡ですね。閾値がないので。ここから先はさすがにあるってみんな認めるけど，ここから下はさすがにないって認めるだろう，そういう話です。

実積：チューリングテストを通ったら人工知能として認めるという考え方は，いまでもあるんですか。

鳥海：数年前にチューリングテストを合格したっていうAIが初めて出ました。

実積：たとえば「りんな」はチューリングテストに合格するレベルなんでしょうか？

鳥海：あれはたぶんチューリングテストの大会に参加していないのではないでしょうか。ちなみにチューリングテストに合格したAIは「ウクライナ人の6歳の少年です」っていう設定でやってる。それで人間かどうか区別がつかなかったということらしいです。

実積：それは意識あるAIなんですか。

鳥海：意識はないですよ。だから弱いAIだと思います。

実積：強いAIと弱いAIを区別するチューリングテストみたいなものはありますか。

鳥海：少なくとも私が知る限りはないですね。要は前の段階なんで。人間っぽく見せることすらできてないわけですから。

実積：でも，先ほどの話だと，息子さんが1歳のときは意識なかったけど，今意識出てきたっていうことは，鳥海さんの中で，意識の有無を決める閾値を持っているわけですよね。

鳥海：そうですね。ただ，それは前提条件として，この物体は人間であるというのが分かってますからね（笑）。

4　強いAIと弱いAIの区別

宍戸：制度側から見ると，強いAIと弱いAIを区別することによって，たとえば規制なり何なり税法上の取扱いなり，商品サービスとしての格の違いなり，何か意味があると分かるんですけれども，技術の研究者の側から見たときに，強いAIと弱いAIを区別する実益は何ですか。

鳥海：いや，実益はおそらくない。何しろ夢の物語の話をしているので。「新幹線」と「どこでもドア」の区別をしてくださいと言われているようなものです。弱いAIの話の延長上に，強いAIはたぶん来ないです。今の段階で「どこでもドア」規制法を作る意味があるのかは疑問です。

宍戸：もう少し教えていただきたいのは，スタンドアローンのコンピューターと，インターネットと，今のAIとで，それぞれ飛躍的な変化があるように見えます。あとは強いAIも弱いAIも同じようななだらかな変化だと思いがちですけれども，鳥海さんの話からすると，むしろ弱いAIを含む今までの技術と強いAIとの間に大きな断絶があるということでしょうか。

鳥海：あります。だからArtificial Intelligenceという言葉がよくなかったのではないかという議論もありますね。違う言葉にしたほうが良い，と。

実積：僕は，AIを独立のスタンドアローンで使う状況があまり想像できません。車の自動運転にしても自動運転しろと命じる人がいる。強いAIが仮にできたとしても，その後ろには，それに対して命令を下したり，あるいは，キル（停止）スイッチを持っている人がいるだろうと思えるんですけれども。また，十分に進化したAIはたぶん人と区別がつかなくなるはずなので，人間を相手にしていると思っていたのに実はAIだったという状況が生じてくると思います。たとえば，強いAIが自律的に行なっている行為と，人が弱いAIの助けを借りつつ行なっている行為と，完全に人間が行なっている行為の区別がつかなくなるのではないかと思います。

鳥海：Watson（ワトソン）の，何か病気を発見したというのがありますよね。あれも結局，AIが発見したのを医者が判断したっていうことになります。そのうち，判断しない医者が出てくるでしょうね。それは弱いAIの話ですし，直近であり得る話です。

実積：その場合，AIが単体として意識を持つ，持たないというのは，利用者からみて意味がないですね。

鳥海：その場合はそうですね。

5　AIと法的責任

実積：人間はAIの言うとおりにしました，AIのほうはプログラムどおりやりましたっていうケースでは誰が意思決定したことになるんでしょうか？たとえば契約法の文脈だと個人の意思の合致で契約成立ですよね。AIが絡むと法的行為は現行法でどのように取り扱うんですか？

鳥海：弱いAIは意識を持ってない道具ですので，今とほとんど変わらないのではないでしょうか。

実積：契約法は個人の意思を尊重するので，AIの意識の有無を問うテストが，そのうち必要になるんだろうなと思います。

鳥海：AIに責任を持たせるかどうかを決定するときに，それは必要になってくるかもしれません。ただ，個人的な感覚としては，ドラえもんが悪いことをしたから，じゃあドラえもんを罰しましょうということを人間や社会がどう捉えるのかは，私は今の段階ではちょっと想像ができないですね。

実積：ジャイアンはドラえもんを殴りますよね。

鳥海：まあ，そうですけどね（笑）。たとえば，ある悪いことをしたロボットを処刑しました。だけどそのロボットの人工知能のコピーはどこかにあって，ドラえもん2やドラミちゃんに，クッと入れたら同じものがピョンって出てくる状態になったとき，果たしてロボットに罰を与えることにみんなが何かを感じるのかは，今はわからないですよね。おそらく，もっと議論が続いたあとの話になる気がします。法制度の問題以前に，AIを罰するかどうかという重要な，社会的な問題になってくるんじゃないでしょうか。

ただ，自動運転の話で責任の話はもう出てますよね。そこに関しては，法制度のほうでAI開発者も自動運転を使った人も守らなきゃいけない。

宍戸：私には，逆に法律家の傲慢がかえって恐ろしいですね。新しい技術が出てきても，法律家は，技術あるいは物それ自体よりも，結局それを使ったり生み出したりした人間が，何か新しい問題を起こすはずだということが分かっていたのに，それを止めなかったと考えて，その人を罰したり責任を負わせたがる。

自分が何か不条理な事態に陥ったときに，それがただ偶然起きたのではなく，誰かが意図的に，あるいは過失によって悪いことをしたからだ，だからこの人に対して責任を負わせるんだという議論を社会が求めて，それを法律家がルールにして運用してしまうという問題が起きやすいように思います。それで本当にAIの研究や社会実装が進むのかどうかが，やはり懸念としてあります。むしろAI研究者の側から，社会なり法がどこで線を引くべきなのか，発信が実は欲しいんですね。たとえば原発にしても，あるいは宇宙衛星にしても，これ以上は責任を負えないということと，リスクはあるけれども技術やサービスの発展から見たときに，責任を限定するとか，あるいは客観責任にするけれども保険で担保するとか，いろんな仕組みを考えるにしても，何かアドバイスを聞かせてください。

鳥海：法律で罰を与えるのは，社会の感情を満足させるためなのか，それとも二度とそういうことが起きないようにするための抑止力なのか，2つの観点があるのではないかと思っています。社会の感情のほうはわかりませんが，抑止力の点から考えると，あまり技術者を罰するということに意味がないと思うんですね。

技術者は基本的には精いっぱい頑張ってやっていて，ただ，予想外のことが，どうしても発生することはある。そのときに，予想外のことが発生したからおまえが悪いんだって罰しても，それは抑止力には全くなりません。AI研究者は別にそうしようと思って作ったわけではないので。たとえば，事故率が何パーセント以下であれば，技術者に責任はないという風にしていただきたいなと思います。

実積：今の話をお伺いしていると，法律家の観点からいくと，AIに意識があるって認定はできない，すべきではないという結論になりますね。AIは意識ある人間に使われる道具にすぎず，人間が最終的に責任を負うべきだという立場ですから。そうすると，技術者であっても完全にはコントロールができないものについて，利用者の責任を問えるんだろうかという疑問が生まれます。その場合，PL法みたいな枠組みは機能するでしょうか？

宍戸：ペットが人にかみついた場合に，このワン

ちゃんに何らかの意識があるけれども，法律的には物ないし道具としてひとまず取り扱って，飼い主の責任を問える場合は問う，問わないこともある，と考えるわけです。本当に意識があるかどうかではなく，社会の側で，法的な責任を問い得る主体と思うかどうかによって線を引くわけです。その線を引く基準は，まさに鳥海さんがおっしゃったように，責任を問うことで社会をどうしたいのかによる。

刑罰についていえば，「目には目を，歯には歯を」という応報，こいつにはもう二度と悪いことをさせないという特別予防，こういうことをした者に対して罰を与えれば他の人たちも気を付けるという一般予防の3種類ぐらいの考え方がありますが，刑事法以外でも似たような考え方をしています。

問題はそこから先で，実積さんのいうPL法は，製造物に着目して，消費者と製造者の間の格差を前提にして，消費者を守るという観点から責任を負わせているわけです。AIの問題も，法律家の中でも2通りの考え方が分かれるだろうと思います。1つは，AIは物ではなくて，情報あるいはプログラムの問題であり物とは違い，問題が生じたときに限度がないという側面も含めて，製造物責任と同じように全部の責任を事業者に求めることは難しいというように，AIと物の違いに寄せることです。

もう1つの考え方は，事業者と消費者の地位の非対等性に着目する。AI研究者と普通の消費者，要するに鳥海さんと実積さん及び私との間には，AIについての理解の差が，あるいはどちらが問題を止めやすいかについての圧倒的な差があるわけです。そうすると，製造物責任類似の発想で，開発者に責任を負わせたほうが，AIに由来する問題がより生じにくいという整理になってくるのではないでしょうか。

鳥海：ペットと同じで考えるというのは，かなり良いと思います。イヌと同じ扱いというのは，ああ，なるほどなと思いました。要は，人工知能には使っているうちに学習するっていう機構が入るので，使い方によって動作が変わってくる。そのときに，製造者は，どう使われるのか分からない以上，責任は負えないですよね。逆に育てているほうも，中身が分からないので，どう育てたらどうなるか，やっぱり分からないという状態ではないでしょうか。

となると，イヌと同じく，ブリーダー，ペットショップに責任があるのかという問題のように考えるのがよさそうですね。たとえばこのイヌとこのイヌを掛け合わせて，こういうイヌを作った人と，それを育てた人がいました。そのイヌが誰かにかみつきました。いや，このイヌとこのイヌを組み合わせたら，どう猛になるでしょうという考え方もあるし，育て方によってはどう猛にならないよねっていう考え方もある。これにかなり近いんじゃないかと，今の話を伺っていて思いました。

実積：たとえば，自動車学校って運転者を大量生産していて，そのうち何人かはひどい事故を起こすわけですよ。そのとき，自動車学校に責任をとれという人はいないです。あくまでも運転者本人にいきます。一方，テスラが自動運転中に事故を起こした場合は，テスラが責任を負うべきだという議論があります。この差は，意識の有無，コントロールの有無に由来するのではないでしょうか。そうするとAIはどっちにいくんだろうか。今のAIはたぶんコントロールができるんだろうけれど，そのうちコントロールができない，何考えているか理解できないAIができたときは，誰が責任をとってくれるのか。それがはっきりと分からないと，社会的に，みんなが安心して使ってくださいと言うのには時期尚早かもしれないと感じます。

鳥海：責任問題の話なら，安心して使えないという話になりますね。

実積：いや，責任問題よりも，AIを改善して事故発生を防止するというインセンティブが十分には生じないという点に危機感を覚えています。AIが関係して発生する事故の責任を研究者が負わないし，利用者にも問えない，もちろんAIも負わないとすると，誰が改善しようというインセンティブを持つのでしょうか。この点について，考えていただきたいなと思っています。

鳥海：確かにそれはそのとおりですね。

実積：インセンティブを与えた結果，資源配分が改善したように見えるのが，二酸化炭素排出規制や公害規制のケースです。悪い結果が発生した場

合の責任を設定することでインセンティブを設計した結果，適切な方向に技術開発が進むという側面があります。だれに責任があるかを認識させるということは，やっぱり必要ですね。

鳥海：インセンティブがあると動くということは，納得するのですけれども，インセンティブがないと動かないには結び付かないんじゃないかと思います。直接的なインセンティブがないと動かないではなくて，他の仕組みによって動かすほうが，建設的な気がします。インセンティブについては実積さんのおっしゃる通りなんですが，現状を考えていると，インセンティブを与えるのはかなり難しいんじゃないか。だから，違う仕組みを持ってくるほうが，AIを今後どうするかに関しては現実的だと個人的には思っています。

6　総務省のAI開発ガイドライン案

宍戸：その違う仕組みの候補になり得るかという観点から，総務省で検討されている開発原則ガイドラインの話に移りたいと思います。現在，政府でもあちこちでAIについての研究がなされています。

たまたま実積さんと私は総務省の「AIネットワーク化検討会議」，その後継組織である「AIネットワーク社会推進会議」のメンバーとして一緒に議論しています。この会議では，ただのAIではなく，ネットワークにつながっている，つながり得るAIという観点からさまざまな議論をしておりますが，AIの研究開発についてのガイドラインを提案して国際的な標準にできないかという方向を目指している点が，総務省での検討の1つの特徴です。

そこで，この開発ガイドラインの中身について少し議論をさせていただきますが，まずは実積さんから，ガイドラインの内容のご紹介および問題提起を頂けますでしょうか。

実積：まず，2016年2月から6月まで，AIネットワーク化検討会議がありまして，そこで，AIが複数社会に存在し，お互い共同作業を行うという社会が，今後実現されるだろうという議論をしました。その上で，安全・安心な社会を作るためには，どういった制度的なセーフガードを張っていくべきかという話を扱うのが，今回のAIネットワーク社会推進会議で，昨年12月から議論を始めています。その議論の一環として，1月末を締め切りにして，AI開発ガイドラインの意見照会をしているところです。

AI開発ガイドラインとは，研究開発に携わる人たちが，こういった原則に従って開発するのが望ましいというものを9つ挙げ，それを3つのグループに分けて記述しています。中身については総務省の資料をご覧いただければいいと思うんですけども，透明性とか，制御可能性，セキュリティのように，こういったものは当然守るべきだろうなというところが入っていて，項目立て自体については，それほど問題はないと思います。

AIの開発は，政府のナショナルプロジェクトの形よりも，大学とか研究機関，あるいは企業が独自にどんどん先に進んでいく，しかもそれが国境を越えてグローバルに展開するという形で進んでいくと想像しています。そこでは，市場メカニズムをベースとした経済的なインセンティブで作業が進んでいきます。その過程で，先ほど言ったように，誰も責任を取らないというか，公害問題のような外部経済が発生するのであれば，社会的に望ましい方向に開発が進むような仕組みを作っておくべきだろうと思います。開発原則にはそのための指針を提示するものとなってほしいと思います。

ただ，注意しなければいけないのは，開発原則の原案を作成したのは法律家が主体の集団ですから，最新の技術知識に対して必ずしも完璧な知見を持っているわけではありません。その意味で，今回のパブリックコメントにおいて技術者コミュニティからの意見が多く出てくることに私は期待しています。

今回この開発原則で，AIの開発者に対して，こういうところを守ってほしいということを9項目挙げていますが，先ほど鳥海さんのほうからありましたけども，AIの開発者あるいはそれをマーケットに出した人が想定できない使い方，組み合わせが出てくると想定されます。そうすると，利用の原則というか，モラルというか，利用者教育というか，そういったものもセットにならないと，今回この開発原則が狙っている目的が果たされないだろうなとも感じています。

対象となる利用者にも２種類あって，１つは本当のエンドユーザー，もう１つは開発されたAIを組み合わせて，利用者に対して新しい価値を提供するような人たち。この２種類の利用者のAI利用原則を，先ほどの責任問題を含めて，ちゃんと決めておくことが，社会に必要な制度づくりではないかと思います。

AIの開発は民間を主体にグローバルに進んでいきますから，この原則を日本国内にとどまらず，世界のガイドラインに持っていこうという総務省の方針には基本的に賛成します。ただ，その中で，将来の民間活用とか研究開発に関するトライアル・アンド・エラーのために，できるだけ大きな余地を見込んでいただきたいと思います。起こるべきリスクに対して予防的な態度に徹すると，せっかくの技術開発の可能性を，特にICTの分野で，摘むことになりはしないかということが一番懸念するところです。AIは学習するものなので，当初，限定的な範囲では失敗を認め，失敗をその次の開発に生かすような試みに対して寛容なガイドラインになればいいなと思います。

7　ガイドラインのグローバル化は可能か

宍戸：ありがとうございました。鳥海さんいかがでしょうか。

鳥海：まず，このガイドラインの目的の話です。これはグローバルに展開して，世界中がこれに従うということを一応目指すのでしょうか。となると疑問なのは，これを外国にガイドラインとして出したときに，果たしてアメリカや中国といった国々が，これに批准するのかどうか。

今AIの開発はほとんど中国とアメリカが行っています。その上で，日本もこれからどんどん発展させていかないといけないわけです。そのときに，一番危惧しているのは，こういったガイドラインを設けることによって，日本での開発に制約を受けてしまうわけですよね。一方で，アメリカ，中国がこれを全く無視するというのが，日本にとっては一番うれしくない状況です。それに関して，このガイドラインを策定する上でどうお考えなのかをお聞きしたいです。

宍戸：私の理解している限りで２点申しますと，第１に総務省としては，アメリカも中国も載れるように，条約のように厳密に拘束的なルールではなく，グローバルな研究開発のコミュニティが共有できるようなソフト・ローとして，ガイドラインを考えていると理解しています。ですから，日本特有の話はかなりそぎ落として，ユニバーサルに受け入れ得るようなたたき台として考えているだろうと思っていますが，もちろんそれで本当に大丈夫かということは問題ですので，今後もさまざまなレベルで議論されていくはずです。

２点目は，ガイドラインが研究開発を縛るのか促進するのかですけれども，事柄にも両面ありますが，少なくとも検討している側は，促進するために必要な社会的合意の一側面だと思っているのでしょう。日本社会によくありがちなことですけれども，たとえば自動走行で，何の手当てもなく社会的な前提もなく事故が起きた場合，メディアや政治が炎上して，そんな研究はやめろという展開が，一番恐れられている。一定のガイドラインという形で，前もって一定の社会的合意を調達しておいて研究開発を守るということが，このガイドラインが目指している理想だと思います。

ただ，この種のガイドラインについて付き物ですが，そういった理想どおりに本当に機能するかどうかは問題です。遺伝子研究など，様々な研究分野でガイドラインがありますけれども，幅があり研究開発を守るものとして運用されるよりは，書いてあることを杓子定規に適用したり，違反がなくてもガイドラインを拡大解釈したり補ったりして，とにかく研究開発の手足を縛っていくという方向になる恐れは，確かにあると思っています。

一般的な研究開発の原則指針，基本理念の部分で，関係する価値，利益のバランスを確保しなければいけないことを示し，弾力的な比較衡量を行いながら，このガイドラインは具体化され，適用・運用されるべきものであることを明らかにしているのは，その趣旨だと思いますが，それで本当に大丈夫か，さらにご議論いただきたいと思います。

8　予測に基づくガイドラインの限界

鳥海：おそらく中国なんかは，取りあえずやって，問題が起きたときにどうするという対処法になるわけじゃないですか。アメリカやヨーロッパ系は

みんなそうですよね。それならば，いろいろ試してみても，おそらくありがたいんですね。もちろん，たとえば人のプライバシーをバリバリ外に出してはいけないというように，最低限はあるとは思います。ただ，それは別にAIの話ではないところで，すでに抑制されている部分ですよね。そこはいいんですけども，そうでないところで，しかもわれわれも法律家の方も何も予想できない状況下にもかかわらず，何か予想して縛るというのは，おそらく縛る方向にしか行かないのではないか。

逆に，失敗したときには何らかの形で対処をしましょう，きちんとしましょうねというほうであれば，まだよい。しかも，その責任を，たとえば研究者ではなくて，国が取りますからと言ってくれるんだったら，すごく研究が進むと思うんですが，。そうではないですね。研究者がきちんと気を付けなさいね，でしかあり得ないわけですよね。ガイドラインなので，当然それは当たり前なんですけれども，それでいて，実はこれは縛るものではないのですと言うのは，やっぱり無理があるのではないかと思います。

宍戸：事後的に失敗したときに気を付けましょうねといっても，人間が悪用してそうなったのか，当該AI自体の中に内在的な欠陥があったのか学習によって生じたのかとか，逆にバイ・デザインといいますか，事前に研究者の方でも最低限気を付けていただくときに，おそらくこのガイドラインの要として想定されているのが，9番目のアカウンタビリティの確保に関する原則，「AIの開発者がステークホルダーに対し果たすべき責任に関する原則」です。

開発者は何を自分がしているか，あるいはどこまでは責任を負えないかが分かり，それをあらかじめ説明していただくと同時に，事後的に問題が起きたときに原因を追いかけるために必要なものとして入っているのだと思いますが，これは研究者の視点からは難しいでしょうか。

鳥海：一般的な製品に関しては，アカウンタビリティの原則はあるんですか。

宍戸：一般に，現在では消費者保護の観点から見て，商品・サービスを売るときに，これはどういう機能を持っており，あなたにどういうことが起きうるかについて，いわば説明責任が要求される場合はあります。一番典型的なのは整形手術ですけれども，最近だと電気通信事業法が改正されて，適合性の原則が置かれています。

AIは文系から見るとまさに魔法の世界で，一般の消費者もそうだろう。そこで，一体これはどういうものなのか，イメージができる説明が欲しいということですね。

鳥海：だとすると，従来のレベルであればAIでも可能だと思います。ただし，AIの場合は，開発者も予期できない動きをする可能性がある。そこまで説明責任を求めると，もうAIは作れないので，想定し得る危険性について書くのは当たり前だとしても，それ以上について書け，あるいは説明せよ，というのは無理ですよね。

9　研究・開発と利用の区別

実積：そこは利用原則が必要で，開発者にそこまで責任を負わせるのはたぶん無理だろうと思います。開発者が使い方を完全にコントロールできるわけではありません。

その意味で，この開発原則がその方向に技術開発をしてくれというふうなインセンティブとして働けば良いと思っています。

たとえば，車のエンジンを作る人は良いエンジンを目指します。でも，利用者目線ではエンジン単体ではなく，システムとしての車が欲しいわけです。アクセル踏んだら進め，ブレーキ踏んだら止まれ，ハンドル回したら曲がれ，エアコンはちゃんと利け，その上で，説明書がちゃんとあり，保険も適用されている，そういうパッケージでわれわれは車を欲しいわけです。AIもそういうパッケージとして提供してもらわないと使えません。その段階まで，つまりパッケージにするまでは研究者コミュニティの領分です。

そのとき，個々の研究者に全システムの責任を任せることは無謀です。「研究者同士が連携を取って，AIというシステムを世の中に出してください。そのあとは利用側が考えますよ。」というふうにしないと，上手く回らないだろうなと思います。AI研究者に，この原則を全部満たすような研究以外は認めませんよと言うと，いつまでたっても製品が出てこないかもしれません。

鳥海：今のお話を聞いていて，気付いたんですが，これは開発原則なのか，研究原則なのかどちらなのでしょうか。今おっしゃっていたのは，研究側ですよね。セキュリティの原則は研究のレベルの話であって，それを組み合わせて開発した車を販売すると，たとえばトヨタならトヨタ，日産なら日産が，これを全部カバーしなさいという話になるわけですよね。そういう意味では，最後に商品を出す人がこの原則を守った商品を出しなさいねというガイドラインになるのでしょうか。

宍戸：この研究開発ガイドラインは，私の理解では，研究開発の全部を一応包含しているイメージですが，エンドユーザーに提供する局面だけじゃなくて，その前段階のメーカーへの提供も含む。AIを組み込んだ空気清浄機を製造して販売するときに，空気清浄機のメーカーにとってはAIが分からないのであれば，その前段階までが研究開発ですね。このガイドラインの適用範囲はかなり相対的だろうと思います。なればこそ，さっき実積さんが強調されたように，利活用ガイドラインは，また別に作らないと駄目だ，という状況ですね。

鳥海：たとえば透明性の原則，制御可能性の原則を個々の部品に組み込む必要性があるかないか。今のお話ですと，AI搭載エアコンの個々の部品が透明性と制御可能性を持っていなければいけないことになります。でも，エアコンの中にそのAIを入れるときには，そのAIの外側で制御可能なことはすごく多いですよね。そうすると，そのAI自身に制御可能性は必要ない。にもかかわらず，制御可能性がないAIを作っては駄目となると，非常に開発が難しくなります。

もっと上の段階でバタンとふたを閉じちゃったほうが全然楽なのに，個々のAIに対して，いちいち制御可能性の原則を取り入れたものを開発するっていうのは，コストが2倍3倍に膨れ上がる可能性がある。

実積：なので，AIネットワークシステムに着目しているんだと理解しています。今言った，制御可能性を担保するキルスイッチのようなものは，AIと外部からつないで使いましょうよという提案になるのかなと思っています。

鳥海：たとえばエアコンの中に入ってるAIも，ネットワークシステムの可能性は十分あります。たとえば隣の部屋のエアコンと通信して何か温度調整するように。AIネットワークシステムのネットワークは何なのかを，きちんと定義したいですね。

実積：さっきの質問の答えとすれば，AIだけで全ての原則を満たすことはないと僕は思っています。たとえばリスク対処についても，保険のシステムを導入すればよいと思っています。いずれにせよ，「対処している」「対処していない」の二元論的な対応ではなく，さまざまな強度の対処方法があるべきだと思います。透明性の原則についても同じです。AIの本体部分に求めるんではなくて，最終的に提供されるパッケージとしてこれらの原則が保証されればよいのではないかというのが僕の意見です。その意味で，この開発原則は，制御可能性を担当する部分，いざとなったら止めてくれる部分などを含めてAIシステムの開発を進めるべきだ，というふうに読みたいと思っています。

鳥海：それだと，最後に商品レベルになったときの，全体システムの制御可能性であり透明性でありという話になってくると思うんですね。

そうなると，システムって何だという話が今度は出てきます。実積さんがイメージされているシステムは，引っこ抜けば止まるものだと思います。一方で，車のシステムについて考えてみると，AI搭載型の車による「交通システム」が出てきます。

その場合は，もはやコンセントを引っこ抜くことはできないんですね。個々の車が勝手に動いてるので，なんかポチッとボタンを押したら全ての車がピタッと止まる仕組みは導入できない。山手線は，今日来るときも止まってましたけども（笑），そういうシステムにできるのかというと，それは難しい。

実積：利用者や利用企業が責任を持つ範囲と，開発者が責任を持つ範囲が分かれる境界線はアプリオリには決められません。なので，利用原則と開発原則は全く別ものではなく，同じ項目は幾つか入るべきなんだろうし，入らければ駄目なんだろうなと思っています。

宍戸：私も同じ理解で，鳥海さんには本当の研究

開発ガイドラインしか適用がない，AIを実装し売る人には研究開発と利活用の両方がかかって，エンドユーザーには利活用ガイドラインだけがかかるというイメージだと思いますね。

実積：全部研究者に責任をおっかぶせて，ということにはしたくない。

鳥海：ええ。AIは基本的にはロジックですから，AIを作るということはロジックを作ることにすぎません。これは私が研究者だからそう思うのかもしれませんけれども，ロジックを作るところに制限を置くのは非常に難しい。それをやると，日本はアメリカや中国に勝てなくなる。ロジックの部分は自由に作らないと，たぶん駄目ですね。研究は自由にして，開発のあたりから徐々に制限が入り，商品化のときにはきちっとした制限が必要だと思います。

10　AIネットワークシステムとは

鳥海：さっきも言ったとおり，車自身がAI化されて，それが他と通信して動くときに，この製品に対してガイドライン作りましょうというのは分かります。ところが，AIネットワークシステムとして考えた場合，複数台の車がネットワークでつながり，それが1つのシステム，AIネットワークシステムになります。それに関して，透明性，制御可能性を担保しようとすると，それは誰が担保するのでしょうか。

実積：それは，ある意味，ビジネスチャンスですね。そういう情報を出す車に乗れば保険料が安くなりますよという企業が出てくるだろうと思います。

鳥海：ただ，ガイドラインの場合は，担保しなさいと書くわけですよね。

実積：今回のガイドラインは拘束的なものではないので，準拠すれば何らかしらメリットが発生するものになってほしいと思っています。ガイドラインに書いているからといって，これに準拠しないシステムを作ることは止めようがないし，そういうことをする企業は現れる。ただし，このガイドラインに準拠すれば，たとえば，透明性とか制御可能性の原則に準拠するように作れば，補助金がもらえるというメリットが付くというイメージです。何らかのインセンティブを与えることができるなら，時間はかかるでしょうが，望ましい原則に準拠するための技術が出るだろうなという期待はしてますね。

鳥海：要は，これは自律分散システムになるので，制御する人がいない。

実積：制御する人がいなくても問題ない。われわれは制御する人がいなくても自分の意志で保険に入ったりするじゃないですか。何らかの行動にメリットを与えるというフレームワークを作れば，長期的にはその方向に動くんじゃないか，という期待です。

鳥海：つまり，ガイドラインに入っている原則，たとえば制御可能性の原則は，満たされなくなりますね。

実積：制御可能性の原則を満たす方法が技術的なものだけとは，僕は思っていないです。制御可能になるものをシステムに組み込むことによって，目的を達成する。どんなものでも乱用される危険はあるし，暴走する危も険あるかもしれないから。そうすると，制御可能性を満たしたほうが得ですよというシステムができると思っています。制御可能な自動運転車とそうじゃない自動運転車があって，制御可能な自動運転車をみんな使ったほうが税金も安くなるかもしれないし，事故も少なくなるかもしれないし，保険が安くなるかもしれないよというような社会システムを作ることで望ましい状況が実現されていくんだろうなと思います。

11　AIネットワークシステムの制御可能性

宍戸：総務省のガイドライン案では，制御可能性について，一応対象をAIとした上で，「制御不能となるリスクにつき，その蓋然性が高い又は不確実と考えられるAIについては，一般社会で利用される前に，実験室等閉鎖された空間において，当該空間外につながる情報通信ネットワークシステムに接続せずに，AIの制御可能性について実験を行い，リスク評価を行うことにより，制御可能性を確保すべきとしてはどうか」と書いているのですが，どう思われますか。

鳥海：一般社会で利用される前に，実験するのは，今でも当たり前にやりますね。なので，別にそんなに目新しい話ではありません。実験室で，ちゃんと動くことを確認しないで外に出すっていうの

はあり得ないです。なので，これは普通のICTシステムと同じことを言ってるだけだと思います。

宍戸：それで，さっきの車の話に戻るのですけれども，制御可能性が順守されれば，ネットワークシステム化された車の自動走行も，自律分散で問題があったときに制御できるというイメージで書かれているだろうと思うのです。

鳥海：これを書かれた方は，複雑系システムについて考慮していないのだと思います。基本的に複雑系システムは制御不可能です。ネットワークシステムを考えたときに，それが複雑系であった時点で，それを完全に制御することは不可能です。

宍戸：これも完全制御までは考えていないと思うのですけれども，何か問題があったときに，すぐそのスイッチを切るということもできない，ということでしょうか。

鳥海：日本全国で車が1万台走っているとしましょう。このとき，個々のAIが問題を起こすのではなくて，それらが組み合わさって初めて起きる問題が存在します。たとえば，今の人間による運転でも，事故がなくても大渋滞が起きることはよくあります。別に個々の人間に問題がなく，みんなきちんと法律を守って普通に運転している。それでも渋滞が起きることが分かっています。それはシステム的な問題なわけです。

同様に，AIならではのシステム的な問題を起こしたとすると，そこに制御可能性はなくなります。

実積：混雑問題の場合は，ピーク時には高く，オフピーク時には低く設定した混雑料金を導入するという解決策があります。AIの評価関数の中に高速道路の利用料金を安くするということを入れれば，システム全体として，混雑回避性が実現できます。制御可能性に関しても，ここでいう混雑料金のような社会システムを導入することで，社会的に最適な制御可能性を実現できるはずです。

鳥海：実積さんの言うとおりの方法が，おそらく唯一の解決策だと思います。ただ，これは複雑系なので，事前には分からないんですよね。ですので，そうなったときに逐次的に問題を解決するための社会的な基盤を作っていくことができる状態になっていなければいけない。

ただ，それはAIの話なのか，社会のシステムの話なのかというと，社会システムの話だと思います。だからAIネットワークシステムに導入するというよりは，社会システムのほうにそれを導入するという考えがよいのではないでしょうか。それを，AIネットワークのほうでそれをなんとかしろというのは難しい。

宍戸：AI研究開発のシステムの側に，全部責任を寄せるということは，私はもともとあり得ないと思いますね。

12　研究と社会システムの対話，AIの定義

宍戸：お話を伺っていて私も見えてきたのは，こういう問題だろうと思います。社会システムの側が，AIの研究開発と孤立して適当にルールを作っても，全く機能しない。逆に，AIを研究開発して，そのアウトプットとして出てくる，AIの利活用が，制度，法とか経済的インセンティブから全く孤立して作られるのも，全く機能しない。だから，鳥海さんのおっしゃる社会システムの問題と，AIの研究開発の双方が，システム間でカップリングできるように，お互いにその視点を内在していないと，有効に機能しないと思うのです。

総務省での議論は，社会システムの側から研究開発について，最低限こういうことをしておいてくれると，あとで問題があったときに介入というか修正ができるという，その取っかかりとなる足場を，研究開発の側の中に築くということだと思いますね。制御可能性のようなAIの研究開発の原則が何もないところで，社会システムの側で全部やろうとすると，とにかくAIは研究開発するなという話になってしまう。少なくとも，問題が起きるとしてもとんでもない問題，社会システム全体を破壊するところまではいかない程度であるとか，何かあったときに一定の手が打てる程度の枠内で一歩一歩進めてくださいということを，社会システムの側は要求せざるを得ないと思うんですね。

鳥海：そういうの全てを包括するようなガイドラインとなるとものすごくふわっとしたガイドラインにしかならないような気がします。今の法律をちゃんと守りましょうねというレベルになってしまうのではないでしょうか。1回世に放ったら，もう二度と変更できないようなシステムは作らな

いようにしましょうねとか，なんとも当たり前なガイドラインになります。

実積：制度を作る側と，研究者の間で意見交換の機会を定期的に持つのが大事だと思います。その嚆矢が今回の総務省による意見照会という位置づけですね。この一回の意見照会で全てが決まって，法律になるという流れではありえない。

宍戸：そもそも，すぐに法律にすることは考えられてないものですよね。

鳥海：それは非常にいいと思うんですが，技術者側にコメントすることのインセンティブを明確に出さないといけないと思います。少なくとも，AIロジックレベルの研究者にとって，あまりインセンティブにはならないですよね。

宍戸：研究者側から見ると，制度屋がいろんなことを言っている，何か言わないと問題が起きるという，ディスインセンティブのほうですよね。

鳥海：だとすると，その段階でOECDに出のは意味がない行動をしているように見えるんですよね。研究者側からすると，われわれの意見を全くくみ取らないで，勝手に作って勝手に出しちゃった，という感じになります。

宍戸：私のような構成員も十分分からないで議論している自覚はありますので，総務省の研究会にしても，きちんとした研究者のコミュニティと議論する機会が必要だろうと思います。

鳥海：そうだと思います。ただ，それは1～2時間議論して済む話ではなくて，1年間かけて議論する話だと思うんです。

実積：1年間かけて議論することにした場合，何を議論しますか。たとえば，今のこの9項目で，これはあり得ないというものはありますか？

鳥海：まず，定義が全然なされてないですよね。やっぱり理系なので定義は重要なんですよ。

実積・宍戸：いや，文系も重要です。

鳥海：おっと，すみません（笑）。定義がないので議論できないということが，とりあえず指摘したいことです。4のガイドラインの目的と理念に関しては，定義が変わるとがらっと変わると思います。適用範囲が利用者と開発者，研究者が明確に分かれていなければ駄目ですね。

13　透明性，セキュリティ，プライバシーの原則

鳥海：透明性の原則については，透明性とは何か，からスタートして，透明性のレベルを指定する必要があります。本当の透明性は，逆に出せるんですよ。プログラムとデータとを全部公開すればいい。ただ，解釈できない。おそらく透明性とは解釈可能性の問題なので，そこをきちんと明確にしたほうが良い。

実積：全く不透明であるのはまずいっていう意味での透明性原則であれば大丈夫なんでしょう。

鳥海：どのレベルの透明性なのかは，1年ぐらいかけて検討するべき点だと思っています。制御可能性も同様ですね。特にネットワークシステムが複雑系であるということが意識されてないので，複雑系の専門家を入れていかないと議論できないと思います。

　セキュリティに関しては，今後議論すればいいと思うので，あまり言うことはない。安全保障に関しては，どうなんでしょう。保障しなきゃいけないのは分かりますが「バグを出さないようにしましょうね」と同レベルの話になっている気がします。100パーセントは無理なので，どこまでと考えるか。

実積：この辺って，アカウンタビリティの話と似てきますね。誰がこの措置を決めたのか，何が原因でこれが発生したのかを明らかにするよう努力しようっていうのは必要ですね。

鳥海：そうですね。できるだけ努力をしていますということを見せるとか，車だったら人がいたらよけるとかぐらいの話にしかならないですが。

宍戸：私は，これはそのぐらいの話なんだろうと，もともと思っていました。

鳥海：プライバシーに関しては，プライバシーデータを全く使っちゃいけませんということは，おそらくできない。できないというよりもほかの国に勝てなくなるので，そこもかなり注意して書かないといけないところだと思います。

14　利用者支援の原則

鳥海：利用者支援の原則は，よく分からないところです。

宍戸：これは最終的なネットワークシステムのこ

とを念頭に置いた上で，その構成要素たるAIがユーザーフレンドリーなものであることを求めるものですね。

鳥海：操作しやすいインターフェースを設計すべきではないかというと，操作しやすいインターフェースとは何かというところから始まりますね。実積さんの理論がまさにこれで，利用しづらいものは誰も使わないという経済的な話じゃないでしょうか。だから，これはあまり意味がない気がします。

実積：利用者支援原則の利用者は，エンジンメーカーに対するトヨタに該当するんだと思います。部品としてのエンジンの細かい話については運転者にではなく，組み立てメーカーであるトヨタに説明しないと通常は仕方がない。AIについても同じで，利用者支援の対象はAI利用企業でないと実効性がない気がします。

鳥海：ユニバーサル・デザインの話だったらありかなと思います。社会的弱者に気を使いなさいという話ですね。

宍戸：私の理解では，AIを使ったサービスが進んできたときに，デジタルデバイドのより強力なバージョンが出てくることを，懸念していると思います。

実積：その担保をAI開発者に求めるのは難しいと思いませんか。

宍戸：この場合の開発者は最後に市場に出す人だと思います。

鳥海：生産者ですね。

実積：問題となるのは利用局面だから，利用ガイドラインに入れるべき話ですね。

宍戸：だから利用企業は，真ん中にいて，研究開発ガイドラインと利用ガイドラインの両方の側面があるわけです。ここではエンドユーザーに提供する開発者としての，AI込みの製品サービスを開発する人について利用者支援を要求しているのではないでしょうか。

実積：車を買うという例で考えましょう。宍戸先生がおっしゃっているのは新車を開発するチームに利用者支援を求めるのと同じですよ。僕は，利用者支援は販売店の人に求めるべきだと思います。

宍戸：なるほど販売店は利活用の対象ですが，それは両方あるのではないですか。

実積：エンジンの開発者は，販売店の人に説明しろという義務は負わせてもいいかもしれないと思いますが。

宍戸：販売店の人が，車の中に入っているインターフェース，いじれないですよね。

実積：そこは，B to Bの局面での利用者支援に関わる話ですね。

鳥海：ここはいろいろな解釈の仕方があると思うので，深く議論するところではないでしょうか。「利用者支援の原則」はたぶん解体していろんな書き方に変えるべきですね。

実積：利用者の定義がないから，難しいんじゃないかな。

15　アカウンタビリティの原則

宍戸：先ほどもアカウンタビリティはある程度議論しましたけれども，これはどうですか。

鳥海：アカウンタビリティに関しては，法制度でどこまでを求めるのかと，技術者がどこまで応えられるのかの擦り合わせの話だと思います。

実積：これは2種類たぶんあって，企業間であれば，リテラシーが十分にあるから，何もしなくても必要な部分は個別にやると思う。もう1つは，最終利用者に絡む話で，事故が発生したあとで損害賠償うんぬんの話になったときに裁判所に提出する必要があるから，一定の形式でログを記録しておきなさい，という話かと思っていました。

鳥海：そこはアカウンタビリティとして用意しておくのではなく，第三者機関が入ってきて，データを集めて説明するようになるのではないかと思っています。事前にここまで用意しておくというよりは，その都度，何が起きたのかを調べなければいけないのではないでしょうか。

宍戸：事故調査委員会のイメージですね。

実積：航空機事故が起きるとブラックボックスの解析をして全容解明を試みるわけですよね。AI開発においても同様のシステムが必要だ，という話だと思います。

鳥海：そこに対して説明責任はあると思います。開発者側が，事故が起きたときに全部説明する責任っていうのを負わせるということではない。

実積：そういう意味ではないと思います。ログを取っておきなさい，ログを取っておいたら，アカ

ウンタビリティ責任を果たすことになります，と意味かと理解しています。

宍戸：私は，これももう少し広いように思います。最初からガイドライン自体は拘束的に事細かに決め切らないことを前提に，研究が進んでいく中で，研究者がやりたいことについて社会的に合意があるかどうか分からないときに，世間に対して議論をせずに研究を進めるのはやめて欲しい。しっかり逐次に，多様なステークホルダーと対話を行って，意見を聴取する仕組みが必要ではないか，というものではないでしょうか。

鳥海：それは，実運用するときに必要なので，開発と研究というよりは，運用する段階までいったときに，初めて必要になる話ですよね。その手前段階では，可能性の芽をつぶしているような気がします。

実積：アカウンタビリティの技術を開発してくれる人に対して報償が与えられるというシステムがないと難しいよね。飛行機のブラックボックスを作る人や企業に対してお金を出しますというようなものではなくて，飛行機を作る人にブラックボックスも作れというのは，確かに可能性の芽をつぶすかもしれない。

鳥海：お金などの作るためのリソースがちゃんと渡されるのであればできると思うんですけど，現状はそうではないですね。だとすると，研究開発段階では，ある程度自由にやってもいいような気がするんですよね。最終的に社会に出すときには，飛行機のブラックボックスのようなものを組み込もうという話になると思います。つまり，研究開発というよりは，利用者の話になるんじゃないでしょうか。

宍戸：いい喩えか分からないですけれども，遺伝子研究の場合，治験の段階もありますけれども，その前にこういう研究を進めていいのどうか判断する例もありますよね。

鳥海：ヒト胚の問題とか，その辺ですね。そこまでの危険性がAIで生まれるのかどうかが未知な段階で，制限を加えるべきかどうかは疑問です。危険そうなものが出てきたときに，それを抑えることができれば十分ではないでしょうか。

宍戸：社会システムの側から言うと，その危険そうなことが起きたときに，アラートが研究者コミュニティから上がってくればありがたいのですが。

鳥海：まっとうな研究者であれば学会発表をするので，その段階で公知になると思います。その時点で問題があれば，問題があると思った人が何か言うかもしれませんし，最初はその業界のシステム（業界団体，学会など）にある程度任せたほうが良いと思います。上から押さえつけると萎縮してしまいますよね。まだ，AIに関してはそこまで危険なレベルの話はあまりない気がします。

宍戸：私も，個々の研究開発者にこれを求めるのは酷だと思います。自主規制は個々の人ではなくて，業界全体として健全に回っていればいいということのはずですよね。

鳥海：そうですね。いずれにせよ，最初は自由にやって，問題がありそうな場合はあとから止めるとやらない限り，アメリカ，中国には勝てないと思います。

16　連携の原則

宍戸：連携の原則は，いかにも総務省の研究会ですが，いかがでしょうか。

鳥海：これに関しては，2つの解釈が考えられるのでどちらにするかによって，全然話が変わってきます。1つは物理的なネットワーク，もう1つはロジカルなネットワークですが，どちらかが全くここでは議論されてない，というより混在しているんですよ。どちらを指していると定義するかで，AIネットワークの話は全然違う方向に行きますから，ここは最初に議論しないといけません。

宍戸：AIおよびAIネットワークの定義をはっきりさせてくれないと，これは困るということですか。

鳥海：そうです。連携とは何なのか。「連携の原則」にいくまでは，物理的につながっていることを指してAIネットワークといっているのかなと思ってたんですけど，連携ということはロジカルなネットワークを考えているのでしょうか。

実積：外国のAI，たとえば，右側通行の米国の自動運転車が日本に来た場合に，左側通行の日本の自動運転車との協調をどうやって達成するのかという問題が，AIネットワーク検討会議のときに念頭にあった問題点だと記憶しています。そうすると，電子的なネットワークだけを相手にして

いるだけでは済みません。情報通信網に乗らない物理空間でのネットワークを議論しないわけにはいきません。例えば，陰から人が飛び出してきたというのは，情報通信ネットワークに載っていない主体間の問題ですよね。

鳥海：実はロジカルなネットワークは，そういう意味でものすごく広いです。たとえば，カメラを搭載した車が，あの車が今どういうふうに動いてるから，速度をちょっと落としましょうというのも，この連携に入るはずです。そういった連携には，トヨタでもホンダでも関係ないんですよね。あそこに車がいるから，ちょっと遅くなりましょうといった形で連携がうまくいくんですね。その意味でAIネットワークが物理的につながっているものを指すのか，何らかの形で連携しているものを指すのかというのが，大きいポイントになるので，どちらなのか定義してほしいですね。

宍戸：私の理解は，まさに物理ネットワークの話で，相互接続，相互運用と言っているのは，それぞれのAIがどういうふうに動いてるかは知らないけど，AI間同士で情報をやりとりして，お互いに理解できるようにするということではないですか。

鳥海：それは物理的につながっていなくてもできるので，ロジカルなネットワークでいいわけですよ。電気信号が飛ぶ必要はないわけです。

実積：電気信号が飛んでるイメージだと，今回の問題にしているところとは違う話だと思います。もしネットワークで電子的につながっているなら，全体を1個のAIとして定義すれば済む話で，連携の話は必要ないと思っているんです。スーパーAIを1個置いて，みんなそこからの指示に従えばいいだけですから。連携しなければいけないという原則がわざわざ必要なのは，頭脳が複数あって，それぞれ優劣がないから，どうしましょうかという局面のはずです。

鳥海：もしそうだとすると，やはりそのシステムは複雑系になりますね。とても難しい話にどうしてもなってくるっていうところで，「連携の原則」については，1年。あるいは1年は無理でも，半年間かけて，きちんと議論すべきところだと思います。

17　むすびに代えて

宍戸：AIネットワーク社会推進会議のAI開発ガイドラインを素材に非常に厳しいご意見を頂きましたが，今のディスカッションでも，分野横断的な協力の必要性がかなり見えてきたかなと思います。この点に関して，『情報法制研究』という雑誌，あるいは学会，財団への期待も込めて，実積さんと鳥海さんから一言ずつ，頂きたいと思います。

実積：情報法制研究所が，情報法研究所にならなかったという点に注目していただきたいなと思います。法制の「制」というのは法制度の「制」じゃなくて，技術を含む社会制度の「制」です。今回のガイドラインの話でもそうですが，技術者コミュニティの意見を聞かないまま法律家を中心にガイドラインを作ると，それはうまく働かない。鳥海さんからご指摘がありましたけれども，両者の架け橋の場がやっぱり必要で，情報法制研究所，情報法制学会はそういうものになっていただきたいなと思いますね。

法律の議論はもちろん大事だし，それで決まるところもあるんですが，プレイヤーが法律に従うかインセンティブを持つか否かという点を経済学者としては着目しています。従ったほうが得というインセンティブを社会全体で確保するような体制を技術者と法律家が手を携えて作るべきだろうし，そのためには密接な意見交換を実現する必要があります。ICTの世界は国境が関係ないので，従いたいとは思わない法律を日本で作ると，プレイヤーは海外にでかけて，そこから国内にサービスを提供するだけです。これでは，みんなが損をするので，それを避けるための場づくりが必要です。

宍戸：ありがとうございました。鳥海さん，お願いします。

鳥海：この議論も含めて，いろいろ話していると，やはり相互理解が非常に不足していると感じました。法律家のほうは，やっぱり法律がこうなって，しょうがないんだから，技術者がなんとかやれよという話をするし，技術者のほうは，法律家は技術が分かってないから，どうしようもないという言い方をする。だから衝突するんですね。

お互いに，もうちょっと相互理解を深めて，お互いに尊敬し合ってやれるというのは言い方がおかしいですけれども，もっと理解し合う場を設ける必要があり。たとえば私と実積さんはいろいろ一緒にやっているんで，わりとけんかをしながらも，お互いを理解できていくようになるわけですけども，そういう場すら持ってない人たちがほとんどですよね。

そういったときに，じゃあうまくいくようにしましょう，とお題目だけ唱えても当然うまくいかない。特に情報系の研究者は，ほんとに法律家の人たちは何も分かってないぜぐらいの認識でしかない状態なので，それをなんとかするのはすごく重要なことだと思うんですね。お互いが理解し合った上で，ではどうすべきだっていう議論をしないと。ただ，お互い理解し合うことが自分たちのメリットにならない。基本的には研究者は研究してればいいし，法律家は法律を作ってればいいから，お互い相手のことを考えてやらなくても回ってしまうんですよね。だから，対話へのインセンティブがない。ここは実積さんにお願いして，経済的にお互いに協力したほうがいいシステムを考えてもらいましょう。

宍戸：法律家は無知を自覚しているから入っていかないというところもあります。法律家は傲慢に見えるところもあると思いますが，しかしそれ以上に臆病なところが正直ありまして，今後，この『情報法制研究』というジャーナルの場で，ぜひ今後も分野間の対話のきっかけになる企画を，やらせていただきたいと思います。長時間にわたりましたけれども，座談会はこれで終わりにしたいと思います。ありがとうございました。

（2017年1月16日収録）

＊　本座談会で言及された「ガイドライン（案）」とは，総務省が2017年1月末まで実施した「『AI開発ガイドライン』（仮称）の策定に向けて整理した論点に関する意見募集」（http://www.soumu.go.jp/menu news/s-news/01iicpO1 02000054.html）当時のものである。総務省はその後も「AI開発ガイドライン」（仮称）の策定に向けて引き続き検討を進めており，2017年夏頃を目途に報告書を取りまとめる予定としていることにご注意いただきたい〔宍戸注記〕。