OUKA
The University of Osaka
Institutional Knowledge Archive

Title	「人間以上」と「人間の条件」人工知能時代におけるヨーロッパ的政治制度の行方
Author(s)	常盤, 成紀
Citation	Co＊Design. 2018, 3, p. 71-78
Version Type	VoR
URL	https://doi.org/10.18910/67894
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

https://ir.library.osaka-u.ac.jp/

The University of Osaka

「人間以上」と「人間の条件」
人工知能時代におけるヨーロッパ的政治制度の行方

常盤成紀（大阪大学大学院 法学研究科 博士後期課程）

“More than Human” and“ Human Condition”
How should we design European political systems in AI age?

Masanori Tokiwa (Doctor’s Course, Graduate School of Law and Politics, Osaka University)

本稿は、近年みられる、政治の現場における人工知能への期待の高まりがいかなる意味をもつのかについて、ハンナ・アーレントの『人間の条件』を手掛かりとして、政治哲学的に考察したものである。人工知能に政治をゆだねることにより、人々は公正で合理的な判断（基準）を手に入れることが期待できる。しかし政治的判断とはアウトプットに過ぎず、社会は本来、そのアウトプットが出るまでに練り上げられた、あらゆる価値観が織りなす交渉のプロセスによって作られる。そのプロセスをカットしてただ帰結を機械に求めることは、アーレント的な「活動」を無意味化するものである。その結果、人間のアイデンティティは無効化されながら、人工知能は、圧倒的情報量に基づいて、ただ私たちに判断（基準）を提供する。拒絶しようのない人工知能社会の到来の中で、私たちが人工知能に対して、必要に応じて「それは違う！」と抗うためには、議会、あるいは公的空間は、正義や平等についてもっぱら議論する空間として機能するように、制度設計する必要がある。そのために参考になるのが、不確かな人間社会を機能させるための「約束」と「許し」という、アーレントのアイデアである。

This Article discuss how we can design political systems in AI (Artificial Intelligence) age, based on Hannah Arendt *The Human Condition*. If we entrust activities about politics to AI, it would promise the equal and rational judgements (or criteria of judgements). But political judgements are merely output, and the society is made by the communication among many values that have orchestrated by the time when that output is produced. To cut this process means making “Action” meaningless. As a result of that human identity would be undermined, and AI would suggest the criteria of judgements strongly with using overwhelming amount of information. But AI society could come, and we cannot reject that. In order to protest the suggestion of AI as appropriate, we should design the parliament or public spheres as the space where people discuss about justice or equality. Arendt’s “Promise” and “Forgiveness” that make the uncertain human society work is very helpful idea for this design.

キーワード ―― 政治制度、人工知能、人間の条件

Keyword ―― political system, AI, Human Condition

2014年、人工知能（AI）の世界的権威である数学博士ベン・ゲーツェルが、いわゆるAI政治家ともいうべきシステム、“ROBAMA”（Robot + Obama）の開発に着手した。政治の担い手である人間は、限界合理的で私利私欲にからめとられ、かつ往々にして知識不足である。その中で、あらゆる情報をもとに合理的な政治的判断を下すことのできる人工知能を作ることができれば、人類はもっとも公正で不正腐敗のない社会を実現させることができるだろうというのが、彼のアイデアである。2006年にディープラーニングが発明され、第3次AIブームが起こって以降、人工知能に対する期待が政治の現場にも持ち込まれるのは、すでに時間の問題であったといえよう。ゲーツェルは、ROBAMAを汎用型AI[1)]の形で、2025年までに完成させることを目指しているという[2)]。

本稿は、こうした政治の現場における人工知能への期待の高まりの中で、従来の政治制度は時代とどう向き合うべきなのかについて、ハンナ・アーレント（Hannah Arendt, 1906-1975）の『人間の条件』を手掛かりとして、政治哲学的に考察したものである。なおここでの政治制度とは、君主制、民主制などの統治形態に加えて、議会制や三権分立などの統治構造を指している。これまでヨーロッパの政治哲学は多くの場面において、政治的意思決定と制度との関係を論じてきた。それは、不完全であるものとしての人間が、政治社会を正しい方向性に導くための仕組として、制度を考えてきた歴史である。この、意思決定を行う上での人間の不完全さを補うものとして考えられてきた制度は、人工知能の導入可能性を受けて、人間をほぼ代替するものとなることが期待されており、人間はその制度のいわば「緊急停止装置」として存在することで、手続き的正統性を担保する役割が与えられつつある。このいわば主従の逆転ともいえなくない未来は、私たち人間にとってどういう意味を持つのだろうか。

こう論じたときに想定されるいくつかの意見について、予め答えておきたい。まず、質の悪いSFの見すぎではないのか、実際に人工知能に完全に依存する社会など到来しないであろう、という意見があるかと思う。これに対しては、未来は誰にも分らない以上、イエスともノーともいえない（分かりやすい未来予想は大体外れるようにも思えるが）。ただ政治哲学として取り上げることができるのは、今、すでに述べたような期待が出現しているという事実である。そうした時代が来たらどうするか、ではなく、そういう期待にどのように意味を与えるか、も十分に議論の対象となる。

くわえて、またありていの技術批判か、という意見もあるだろう。しかし、筆者はシンギュラリティ（技術的特異点、後述）の到来を拒絶することは無意味だと思っている。ただ、シンギュラリティが到来するであろうということは、その技術に人間の仕事を最大限まかせるべきだということにはならないし、ローテク的な制度は依然として残るだろう。それは現在、グローバル化の時代においてもなお、国民国家や国境が残っているのと似ている。いや、残っているどころか、政治を語るうえで重要なテーマですらある。それを思うと、これから先も、これまでの制度を抜きにして政治を語ることはできないであろう。本稿は、来るべき新しい技術への期待に湧きがちな現在にあって、これまでの制度が依然として持つ重要性について、焦点を当ててゆきたい。

1957年10月4日、旧ソ連からスプートニク1号が打ち上がった。この事件を現代の契機のひとつとし

て語ったのが、ハンナ・アーレントである。彼女は『人間の条件』において、このスプートニク1号への人々の歓喜は、「『地球に縛り付けられている人間がようやく地球の脱出する第一歩』という信念」[3]、「人間の条件を脱出したいという望み」[4]に基づいていると述べた。彼女にとって「人間の条件」とは、地球そのものであり、人間が生きていくうえで必ず依存する自然そのものである。彼女によれば、人間とは、アテナイ・ポリス以来、他の動物とは異なり、言語を用いてコミュニケーションを行うことで、自分がいかなる存在であるかを人々の間で共有させ、物語として残していく生き物である。この行為を彼女は「活動」と呼び、生きる必要に迫られて行う「労働」や、使い物になるかならないかで価値が決められる「仕事」とは区別された、人間の最も高貴な行為であるとした。そして、この「活動」は言語を媒介とする以上、語る対象にはリアリティが必要となってくる。世界は語れるものでないと言語の意味がないからである。ところが、彼女によれば、この宇宙への旅は、地球の外部から地球を見る「引証点／a point of reference」を作ってしまった。この引証点とは、宇宙空間において地球や太陽の動きを把握するのに設定する、とりあえずの基準である。テクノロジーの進歩によって、人間の認知を超えた機械の演算が任意の引証点をもとに、星々の動きを割り出す。これは地球が天動説的でも地動説的でもなくなり、完全に相対化されたことを意味する。ここにあって地球はリアリティを失い、代数的に処理されることになる。アーレントはこのことを、「世界疎外」と呼び[5]、この「引証点」のことを「アルキメデスの点」と呼んだ[6]。

テクノロジーによって地球という条件が無意味化し、世界が疎外され、リアリティを失うこと。これが、アーレントが問題とした現代の傾向である。彼女はこのスプートニクショック以外にも、試験管ベビーや核実験などを念頭に置いており、特に現代世界は「最初の原子爆発で生まれた」としている[7]。これは、人間が地球に手を加え、操作可能な対象とし始めた1950年頃以降の時代を、人新世（Anthropocene／アントロポセン）と名付けて議論しようとする地質学とも、考えを共にしているようにも思われる――もっとも地質学は、単に政治的な意図でもってこの時代区分を受け入れようとしているわけではないが、他方多くの人文学の分野がこのアイデアに触発され、地質学上の決着がついていないにもかかわらず、概念としてこの区分用語を用いて議論を展開している。

そしてその人新世が始まったとされる頃と同じ時期に創始された考え方が、「人工知能（Artificial Intelligence）」であった。1956年のダートマス会議でジョン・マッカーシーによって提唱されたこの概念は、幾度かの流行の浮き沈みを経たのち、2006年にジェフリー・ヒントンによるディープラーニングの発明を受け急速に注目が集まるようになる。これを一般に第3次AIブームと呼ぶが、これと前後して2005年に提唱された「シンギュラリティ（技術的特異点）」（レイ・カーツワイル）という概念は、当初こそ荒唐無稽なものとあしらわれていたが、ここにきて一挙に議論の対象となった。シンギュラリティとは、コンピュータ技術の発達によって人類の変化曲線の傾きが無限大に達する瞬間（特異点）のことを指し[8]、それはまた、「われわれの生物としての思考と存在が、みずからの作りだしたテクノロジーと融合する臨界点」[9]であるとされる。カーツワイルは、これが人工知能の開発によって2045年に訪れるとして、物議をかもした。具体的には、1台のコンピュータのスペックが全人類の知能の合計を超えるタイミングが、2045年に起こるとする。これを受けて世間では、人工知能が多くの仕事を代替する社会が到来するのではない

かといわれはじめ、これについて多くの関連文献が出版されていることは、人々のよく知るところである。

この人工知能について、アーレントは言及しているわけではない。しかしながら、この全人類の知能の合計を超えるスペックをもつコンピュータの社会的実装は、いうなれば「人間以上（More than Human）」の存在が（政治的な）意思決定を行う／補助することを意味する。そしてそれは、リアリティなき社会形成が行われる時代の到来を予期させる。このことについてはもう少し詳しく見てゆくつもりであるが、そうであるとすれば、このことは私たちに「アルキメデスの点」を確実なものにさせるのではないか。『人間の条件』が書かれた1958年当時、アーレントは、「私たちは、アルキメデスが立ちたいと願った地点に実際に立っているわけではないし、依然として人間の条件によって地球に拘束されている」[10]としていたが、シンギュラリティの到来がまことしやかにいわれる今日、人類は本当にアルキメデスの点を手に入れるかもしれない――もっとも、まだそうなると決まったわけではなく、実際には技術的課題も多いわけであるが。「人間以上」は「人間の条件」にいかに影響を与えるのか。だがこれを論じる前に、そもそも意思決定システムとしての政治制度がこれまで、ヨーロッパの政治哲学の文脈でどのように語られてきたのかを確認しておきたい。そうする理由は、なによりもこの人工知能への期待自体が欧米発祥であり、欧米におけるこの新潮流を理解するのに、一度欧米の政治哲学的ルーツを踏まえておく必要があると考えるゆえである[11]。

――政治を担うものとしての人間は、完全ではない。これはヨーロッパであれ、アジア（中華圏）であれ、共通した考え方である。しかしながら、これに対する政治的応答は両者の間で異なっていたといえよう。アジア（中華圏）の場合、人間が完全ではないことへの応答として、「最もすぐれた人」に頼るのが基本的な発想であった。「徳治」や「天子」といった考え方はこれを端的に表している。他方ヨーロッパの場合、人間が完全ではないならば、頼るべきは「最もすぐれた人」ではなく、「制度」であり、その制度とは、被治者としての人民だけではなく、政治の担い手としての人間をコントロールする存在であった。世襲王政の堕落が大統領制によって、貴族の腐敗と民衆の暴走が二院制によって乗り越えようとされてきた歴史は、これを物語っている。わが国における代表的な制度論者である荻生徂徠についても、彼にとっての制度（「道」）とは、「『聡明叡智』なる『聖人』たち〔＝為政者〕が『建立』し『制作』した」[12]、被治者を統治するための手段であって、為政者自身をコントロールするためのものではなかった。また反対にプラトンは、『国家』で哲人王政治を志向しながらも、そのプロジェクトは挫折し、『法律』において法制度に次善の期待をかけていた。もっともこのような東西比較論が粗雑な一般化であることは自覚している。しかしここでは、ことわが国にとってヨーロッパ由来といって差し支えないであろう政治制度一般が、おおまかにどういう性格をもっているかを素描するという意味で、こう論じたと理解されたい。またアジア（中華圏）との比較という大鉈ぶりも、この限りでのみ意味をもつものであることも断っておきたい。

ともかく、ヨーロッパにとって政治制度とは、政治を担うものとしての人間が政治的意思決定を行う上で、誤った方向に行かないようにするための知恵であったということができる。これはなによりも、人間を意思決定の主体としながら、一部の意思決定について、それをできなくする制御装置的役割を、制度

が担っていたということである。

この伝統に照らした時に、いわゆるAI政治家への期待はどういう意味を持つのだろうか。そもそもこの技術が政治過程にどのように組み込まれるかについては、まだ明らかになっていない。今のところ分かっているのは、政治的事象について公平で合理的な意志決定ができる人工知能を開発しようとする動きがあるということである。なおこれは人工知能に政治的意思決定をゆだねる以外に、人工知能のサジェスチョンをもとに人間が政治的意思決定を下す場合も含んでいる。

もっとも、こうした期待は分からなくもない。こと先進国では、経済的繁栄はほとんど共通の価値であり、それを達成してくれるのが人工知能なのであれば、歓迎しない手はないだろう。技術者たちのいう「公平で」「合理的な」人工知能がそれをもたらすかどうかは別としても、冒頭にも述べたように、限界合理的で私利私欲にからめとられ、かつ往々にして知識不足である人間よりかはまともな社会を実現してくれるだろうという期待が、人々の中にあることは確かだろう。保守が暴力的になり、リベラルが教条主義的になる時代、もっともらしい答えを誰か、できれば頼れる存在に任せたくなる気持ちはよく理解できる。

だがこの帰結至上主義ともいうべき考え方には、重要な点を補う必要があるだろう。確かに、「正しい答え」や「最善の決断」は政治社会のかじ取りをするうえで大切ではある。しかしそれらはアウトプットにすぎず、社会は、そのアウトプットが出るまでに練り上げられた、あらゆる価値観が織りなす交渉のプロセスによって作られてゆくのである。このコミュニケーションプロセスと、生成的ともいえるネットワーキング的な社会構成プロセスこそ、アーレントが「活動」と呼んで重要視した人間の行為に外ならない。もっともらしさを機械に依存した結果、そこから演繹される意思決定はリアリティを欠き、同時にコミュニケーションプロセスは一気に無意味なものとなってしまう。これにより人々は、社会というネットワークの中に自身を物語として編み込んでいくという機会を失い、自分たちの間にある差異（アイデンティティ）すら、たちまち意味を失ってゆく。

この指向性は、「エビデンシャリズム」（千葉雅也）に支えられている。エビデンシャリズムとは、昨今蔓延する、いかなる事実の説明の際にも常に（過剰に）エビデンスを要求する態度のことをいう[13)]。ここでは言葉ではなく、ただエビデンスが重要な説得、納得の材料となってくる。反対に、ビッグデータを解析にかけたり、機械学習にかけたりした結果に対して、私たちが直ちに反論を加えることは非常に難しい。それは、人間では処理できないほどの量のデータに基づいて演算しているからである。ここにきて私たちは、とやかく議論する前に突き出されたデータに「ぐうの音もでない」という事態に陥る。こう考えたとき、人工知能の演算の結果は、政治的意思決定の少なくとも最重要参考資料となる可能性がある。ここにあって人間には、ふたつの予防策が残されているだろう。ひとつは、よほど受け入れがたいサジェスチョンが提出された際に、それを拒否する「拒否権」である。もうひとつは、誤作動が起こった際にも人間に危害を加えないように一定の帰結をコントロールする、フェイルセーフ機能の搭載である。前者は自動運転車（コネクティッドカー）における「レベル3」の考え方であり、後者は同じく自動車についての安全性規格ISO26262をはじめとする多くの規格において採用されている、人間中心主義の機械設計思想である。

いずれにしてもこれらは、従来のヨーロッパの政治制度の考え方と違うことが分かるだろう。すでにみたように、これまでヨーロッパでは、人間を意思決定の主体としながら、一部の意思決定について、それをできなくする制御装置的役割を、制度が担ってきた。しかしここでは、主役と制御役が逆転している。シンギュラリティ時代においては、社会を作るのは機械であり、人間はただ拒否権をもつ存在となる。井上智洋は、人工知能の普及によって多くの「労働」から人間が解放され、より「活動」に専念できるようになるというユートピアを思い描いている[14)]。だが、こういった技術的楽観主義は、「たとえばテーブルや椅子を作るように、制度や法律を「作る」とか、人間を「よく」あるいは「悪く」作るとか、いずれにしても、人間事象の領域でなにか「作る」ことができるという幻想にもとづく純粋な迷信にすぎない。さもなければ、このような信仰は、政治的非政治的を問わず、一切の活動にたいする意識的絶望である。その場合、この絶望は、他の「材料」を扱うことができるというユートピア的希望と一対になっているのである」[15)]。そしてまたこれは、「人間事象の領域から偶然性を取り除き、同時に、行為者が多数いることから必ず生じる道徳的無責任を取り除くこと」を期待してのことである。人間の「活動」に基づく政治はもろく、不確実である。それは人間の能力に限界があるからであり、それゆえに政治は常に失敗の危険と隣り合わせである。また多くの人間の物語の織りなすものとして生成された意思決定には、本質的に責任者はいない（あるいは社会の全員が責任者である）。この不安定さを取り除き、理想の社会を「作りたい」というユートピア的希望が、人工知能への期待からは見え隠れする。確かに、「労働」は限りなくなくなるかもしれない。しかし、同時に「活動」すら消し去る可能性を、人工知能は持っている。主従逆転の「ユートピア」では、人間存在は疎外され、アイデンティティは無効化されながら、人間の理解を超えたもっともらしさが、抗えない形で、ただ私たちにサジェスチョンを提供する。この事実と私たちはどう向き合えばよいのだろうか。

だが、本当に抗えないのだろうか。すでに申し添えたように、筆者は、シンギュラリティの到来を拒絶することは無意味だと思っている。そして当然のことだが、ある技術の導入は、それに完全に依存するか、拒絶するかの二者択一ではない。この二者の間の幅広いグラデーションの中で、どういった導入の仕方を採用するかが重要となってくる。この中で、人工知能からのサジェスチョンを、「それは嫌だ!」と拒否するだけではなく、「それは違う!」といって抗うことができるための、制度的実装が必要となってくるはずである。そう考えると、シンギュラリティ時代の議会、あるいは公的空間は、正義や平等についてもっぱら議論する空間として機能するように、制度設計する必要があるだろう。ここにあって、なぜ私たちの政治社会には官僚だけでなく、政治家がわざわざ存在しているかは、考えるに値する。単に技術的議論に終始するのであれば、それこそ政治家／政治的動物としての人間はいらなくなる。だが、政治哲学が手続的正義に後退し、あらゆる価値観を「個人的なもの」として私化し、相対化した今、事態はそうなりつつある。そして、個々が持つ価値観が相対的にしか意味を持たないのであれば、そもそも個々の言論にも意味はなくなり、社会はリアリティを維持できない。

他方この「抵抗」が実装された制度の下では、私たちは「正解」や「確実性」を追いかけることをあき

らめなければならない。政治的意思決定が人間事象に委ねられるからである。ここでは、個々それぞれの行いがお互いに影響し合い、それによって常に生成的に社会が形成されてゆく以上、誰も自らの行為の帰結を予測することができないばかりでなく、自らの行為の主人となることもできない。だがそれでいて自らの行為を取り消すこともできず、にもかかわらず皆、意図せず予見しなかった帰結の責任者となる。この限りなく不安定な社会システムの中でアーレントが提供した処方箋が、「約束」と「許し」であった[16)]。自らの行為の帰結が予言できない以上、行為者は自ら他者との間に「約束」を課すことによって、一定の方向に行為が帰結するように努力する。他方「約束」を受けた他者は、その帰結がどのようなものになろうとも、その行為者を「許す」ことが求められる。この「許し」という保険によってのみ、人間は安全に行為を行うことができるとする。これは一見すると性善説すぎる考え方かもしれない。だが例えば選挙がいま、候補者の過去の仕事に対する評価という、ある意味で受動的でともすれば後ろ向きな営みになりがちな中で、「許し」を制度化し、選挙を「約束／許し」モデルとして考えることができれば、マニフェストをめぐる政治家と市民との間で積極的なダイナミズムが起こることが期待できるかもしれない。

人工知能技術の普及を恐れる人は多い。だが人類の歴史は、技術と人間の相互作用に基づく進歩の歴史である。その昔人類が火を恐れながらも使いこなし、近年では飛行機を恐れながらも使いこなしたように、今人工知能を恐れている私たちも、人工知能と共存する日が来るかもしれない。そして「シンギュラリティの到来」とは本来そういう意味であり、人工知能が人間を追い越すのではなく、人工知能によって人間（文明）がより発展するという意味である[17)]。未来は意外とシンプルかもしれない。飛行機がなぜ飛ぶのかわかっていなくとも皆飛行機を利用するように、機械学習の中身がブラックボックスでも、皆人工知能を使いこなす日が来るだろう。あとはそれをいかに「疎外」の方向に向かわないかの制度実装の問題である。

また、機械の暴走を恐れるにしても、それと同じくらいか、あるいはそれ以上に、人間こそ暴走しうる存在だということは忘れてはならないだろう。そしてすでに素描したように、その人間の暴走を抑えることこそ、制度一般の本質的な役割であったはずである。自動車の事故はフェイルセーフ機能である程度防げるかもしれないが、そもそも飲酒運転をすればほぼ確実に事故を引き起こしてしまう。人間の暴走を防ぐのはルールなのかアーキテクチャなのか、それとも社会的合意なのか。技術と制度との関係は、こういったこと抜きにしては考えることができない。

とはいえ、技術の進歩は加速度的であり、それこそシンギュラリティ（特異点）に向かって突き進んでいるように感じる。その中で来るべき未来を素朴に恐れないための最初の一歩として、そもそも人間とはどういう生き物かについてアイデアを提供するアーレントの議論は、今なお有効であるし、むしろ今だからこそ確認しておきたいものだといえるのではないだろうか。

註

1）人工知能のうち、あらゆる領域においてパフォーマンスを発揮する人工知能のこと。これとは反対に、

一部の領域においてパフォーマンスを発揮する人工知能のことを「特化型AI」といい、囲碁におけるAlphaGoなどがそれに該当する。

2) 河鐘基「AIの権威ベン・ゲーツェル氏「人工知能で政治的な腐敗を正せる」」ROBOTEER、閲覧日2017年11月30日
https://roboteer-tokyo.com/archives/6037

3) Arendt, Hannah (1958) *The Human Condition*, Chicago: the University of Chicago Press. = (1973) 志水速雄 (訳) 『人間の条件』ちくま学芸文庫 (以下*HC*) : 10.

4) *HC*: 11.

5) *HC*: 411.

6) *HC*: 415.

7) *HC*: 16.

8) Kurzweil, Ray (2006) *The Singularity Is Near: When Humans Transcend Biology*, Penguin Books. = (2016) 井上健ほか (訳) 『シンギュラリティは近い : 人類が生命を超越するとき』NHK出版.

9) 同上: 15.

10) *HC*: 421.

11) 念のため申し添えておくが、本稿は、アルキメデスの「点」と技術的特異「点」とを掛けて表現しているわけではない。あくまで人工知能の発達が、人間事象を相対化する引証点をもたらすという意味において、両者を引き合わせるものである。

12) 渡辺浩 (2010) 『日本政治思想史〔十七～十九世紀〕』東京大学出版会:180. 〔　〕内は引用者加筆.

13) 千葉雅也「アンチ・エビデンス : 90年代的ストリートの終焉と柑橘系の匂い」10+1、閲覧日2017年11月30日
http://10plus1.jp/monthly/2015/04/index03.php

14) 井上智洋 (2017) 「人工知能とベーシックインカムによる革命:ハンナ・アーレントの望む世界へ」『atプラス』32: 112-125.

15) *HC*: 305.

16) *HC*: 370-386.

17) 佐保圭 (文) 中島秀之、松原仁 (対談) 「シンギュラリティで人類はどうなるのか」日経BPネット、閲覧日2017年11月30日
http://www.nikkeibp.co.jp/atcl/column/16/ai/080300003/?P=1

(投稿日 : 2017年11月30日)
(受理日 : 2018年1月10日)