

Title	AI倫理分野における人文・社会科学の「輸入学問」から「発信学問」への転換 / JST・RISTEX「人と情報のエコシステム」領域の取り組み
Author(s)	茅, 明子
Citation	年次学術大会講演要旨集, 34: 138-143
Issue Date	2019-10-26
Type	Conference Paper
Text version	publisher
URL	http://hdl.handle.net/10119/16512
Rights	本著作物は研究・イノベーション学会の許可のもとに掲載するものです。This material is posted here with permission of the Japan Society for Research Policy and Innovation Management.
Description	一般講演要旨

AI 倫理分野における人文・社会科学の「輸入学問」から「発信学問」への転換/JST・RISTEX「人と情報のエコシステム」領域の取り組み

○茅明子（科学技術振興機構・社会技術研究開発センター）

akiko.kaya@jst.go.jp

1. はじめに

日本の人文・社会科学分野は、近代化の過程で西洋の学問を受容したという歴史的経緯があり、現在に至るまで例外はあるものの多くの分野で西洋の研究者の研究成果を学習・紹介する研究、つまり「輸入学問」のスタイルが主流とされてきた。

しかしながら、近年活発に議論されているAI倫理における検討においてはそれとは若干様相が異なっている。AIの社会での活用は欧州・米国・中国など地域によって戦略が大きく異なっており、1 国内で閉じるのではなくグローバルな視点での考察を実施し、その成果を発信することが欠かせないからだ。また、AIをはじめとする情報技術は、人間を介在せずに「学習する」という性質上、事故が起こった際の責任の所在の解決に深刻な問題を投げかけている。つまり、AIは西洋の近代哲学が基盤とする外界からの影響から完全に独立した「自律的な個人」による統制という図式におさまりきらない性質を持ち合わせているということであり、これまでの「輸入学問」の形式だけでは使いこなすことができない技術であるともいえる。その観点からは、今後のAI倫理は、主体と客体の分離が厳格ではない、それがゆえに批判も多かった日本社会の人間と機械との関係性を基に検討していくほうが、西洋のそれを基にするよりも親和性が高いとも見なすこともでき、日本社会を西洋社会から「遅れた」ものとして捉えるのではない新たな視点を世界に提示することの価値はこれから高まっていく可能性も考えられる。

このような状況を踏まえ、科学技術振興機構(JST)・社会技術研究開発センター(RISTEX)「人と情報のエコシステム(HITE)」研究開発領域は、2019年度の公募を英国のリサーチカウンスルUKRI(UK Research and Innovation)との連携にしぼり、AI倫理分野における日本の人文・社会科学研究の「発信学問」への転換を試みている。本報告では、その取り組みの一端を紹介する。

2. 「近代社会」における「輸入学問」としての人文・社会科学

2.1. 人文・社会科学の「輸入学問」としての役割

日本は、明治維新以降150年あまりにわたって西洋文化を貪欲に取り入れ、近代民主国家と近代市民社会を創出し、経済的な繁栄を成し遂げてきた。日本が非西洋圏にあり、西洋の言語と全く異なる日本語というローカルな言語体系をもちながらも、例外的に西洋圏以外から「近代化」を成し遂げることができた一つの要因は、人文・社会科学分野に係わる研究者達が高度な翻訳文化を確立し、母国語で基本文献や最新文献を読み思考することができる環境、すなわち大学教育をつくりだしたことであるといわれている[1][2]。世界では独自の言語を使って高等教育を実践できない国々も多い中で、日本では母国語で容易に先端の知見に触れることができ、独自の思考を深めることができてきた。このアドバンテージは我が国が「近代化」を成し遂げるにあたり、非常に大きかったと考えられる。

2.2. 「輸入学問」への批判

しかしながら、これら人文・社会科学研究者の翻訳による貢献は、時に「輸入学問」として継続的に批判されてきている[2][3]。代表的な意見としては、日本の人文・社会科学研究者の研究成果は、その水準はきわめて高いにもかかわらず日本語の障壁に遮られているため、研究の質・量を世界標準で評価することができておらず、国際的な評価を勝ち得ていないのではないのか。日本の人文・社会科学研究者の有力な研究スタイルは欧米の研究者の研究成果を学習したり紹介したりするタイプであり、それゆえ

その根本は西欧のテキストにあり、研究が日本の歴史や文化・社会から乖離したところで営まれてきているのではないかと、などが挙げられる。英語での発表が主であり、常に世界的評価を念頭に研究を行う自然科学分野とはこの点で大きな相違がみられる。他方で、社会科学の中でも経済学は最も国際化になじみやすい学問といわれており、国際的学術雑誌に掲載される論文数や引用数も近年増加傾向にあるとされており [4]、人文・社会科学分野全体に上記の批判が当てはまるわけではないということも明記しておきたい。

2.3. 人文・社会科学の国際化のためのこれまでの取り組み

こうした状況に対し、人文・社会科学の国際化を促進するために様々な取り組みがこれまでなされてきた。東京大学社会科学研究所は、日本語で発表済みの優れた論文に英語読者向けの改稿を施し掲載する形式の英語ジャーナル Social Science Japan Journal (OUP)を発行しているし、日本国際政治学会も英文雑誌を年1回発刊している(International Relations of the Asia-Pacific) [4]。文化庁や国際交流基金、国際文化会館や企業による翻訳出版事業も地道に実践されている。日本学術会議からは、国家プロジェクトとして日本語で書かれた人文社会系学術研究図書を恒常的に翻訳出版し海外発信するために、日本語人文社会系学術図書翻訳出版センター（仮称）を立ち上げる試みも提案されている [5]。

しかしながら、これらの取り組みは規模として小さく、日本学術会議の提案した翻訳センターは提言として採択されることはなかった。また、人文学・社会科学研究者からは、言語表現自体が研究上の特性のため研究言語として日本語の軽視はなされるべきではないという意見や、異動や昇進などの評価において海外発信が重視されていないため英語で論文を書き発信するインセンティブがないとの意見も表明されており、それなりの質と量を伴った人文・社会科学分野の国際発信、つまり「発信学問」への転換は今もって十分にはなされてきてはいないというのが実情である。

3. AI 倫理をめぐる地政学的考察

3.1. 世界の AI ガイドラインの例

ビッグデータを活用したディープラーニング技術の発達により、2010年代初頭より第3次 AI ブームが勃興している。過去2度のブームとは異なり、今回のブームに特徴的な事柄としては、AI がもはや研究フェーズに留まらず社会のあらゆる分野に実際に実装されるフェーズに入ってきているということが挙げられる。それゆえ、AI は社会のインフラ技術となり、その社会的影響は幅広い分野に及ぶため、世界の様々な組織において AI の倫理的・法的・社会的影響を検討しガイドラインを制定する試みがなされており [6]、そこには AI 研究者のみならず多くの人文・社会科学研究者が多数参加している。

日本においてもそれは例外ではなく、2016年頃から総務省や経産省、経団連や人工知能学会などで AI 開発及び利活用ガイドラインについて人文・社会学者を交えながら活発に議論がなされてきた。これらの議論のとりまとめとして、2018年には内閣府に「人間中心の AI 社会原則検討会議」が設置され、2019年3月に「人間中心の AI 社会原則」が定められた [7]。

表1 世界の AI ガイドラインの例

規程・指針等の名称	Recommendation of the Council on Artificial Intelligence	Ethics Guideline for Trustworthy AI	Ethically Aligned Design	Asilomar AI Principles	Tenets	人工知能学会倫理指針	人間中心の AI 社会原則
組織・団体	OECD	European Commission (High Level Expert Group on AI (HLEG))	IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems	Future of Life Institute (FLI)	Partnership on AI	人工知能学会 (JSAI)/Japan	統合イノベーション戦略推進会議（人間中心の AI 社会原則会議）/Japan
公表時期	2019/5/22	2019/4/8	2019/3/25 (1st edition)	2017/2/	2016/9/28	2017/2/28	2019/3/29

これら制定されたガイドラインからは、それぞれの国・地域によって人間と技術のあり方についての考え方に大きな違いがあることが読みとれる。米国では、合衆国憲法に「自由の恵沢」が掲げられていることから、個人の自由を最大限に尊重した原則が好まれる [8]。例えば、Amazon, Apple, Facebook などの民間企業が創立メンバーに加わった米国の団体・Partnership on AI が制定したガイドラインでは、AI の便益をできるだけ多くの人々に行き渡らせることが重要であるとされており、そのために様々なステークホルダーと開かれた対話を実施し、そこから随時あがってくる懸念を技術にフィードバックしていくというオープンでアダプティブな姿勢を重要視することが明記されている [9]。一方、EU が掲げる AI 倫理ガイドラインからは、AI は人間の自律性 (autonomy) を損なわせてはならず、人間の選択における自己決定を支援するものでなければならないという「人間の尊厳」を至上の価値とする確固とした思想があることが読み取れる [10]。実際、EU は AI に関する政策や法律、企業のあり方をヨーロッパの憲法が保障する基本的人権を基にした価値規範を基に形成していこうとしており [11]、世界で最も厳格なデータ規制といわれる欧州議会が制定した General Data Protection Regulation (GDPR) でも、データの自己コントロール権の保護が重要項目として定められている [12]。

AI の倫理における地域差は、西洋諸国圏内の違いだけに留まらない。ここ数年で AI 大国としての存在感を示し始めている中国は、その倫理観においても世界で特異なポジションを築き始めている。中国の憲法の序言にはアヘン戦争以来の停滞の歴史が総括されており、半植民地・半封建的な状況から脱するために、中国政府主導で AI をはじめとする情報科学を用いたいわゆる「第 4 次産業革命」を強力に推し進めようとしている [13]。中国は情報化の推進を経済面のみならず、社会管理の両面でも進めており、インターネット上の検閲や「信用スコア」を用いた独自の管理システムを構築し始めている。官と民とが密に連携することにより生み出される中国の「監視社会」は、西洋社会の「個人の自由」や「個人の尊厳」という基本理念からは明らかに異物であるが、その人口大国のネットワーク外部性の相乗効果も相まって、「デジタルレーニン主義」 [14] と西洋諸国から揶揄されるほどの脅威となっている。

このように AI の開発・利活用は、それぞれの国や地域の文化・社会の根本的なあり方と密接にリンクしているため、今後 AI がますます社会に浸透していくにつれて、それぞれの社会における「人間」や「人間と機械の関係性」のあり方、「プライバシー」の考え方、データの活用の仕方などの違いを強く生み出していくに違いない。そしてそれは、その国や地域の経済的な発展度合いにも影響を及ぼし、様々な地政学的緊張を生み出していくことが予想される。

3.2. AI ガイドライン策定において EU との連携を選択した日本

このような状況の中、本邦で制定された「人間中心の AI 社会原則」は、当初より日本国内で完結するのではなく、OECD や G20 など発信され世界の議論に積極的に関与していくことが期待されていた。AI 開発の方向性に関しては、1 国で完結できることではなく国際協調の中で検討される必要があると考えられるからである。それをうけ、この AI 原則は 2019 年 5 月にパリで開催された経済協力開発機構 (OECD) 閣僚理事会にて英語で提案され、最終的に「国際的な AI 原則の勧告」 [15] としてまとめていく際の議論に日本は大きく貢献したといわれている。原則のような抽象的な提案を日本から国際会議に提案することは過去にはあまりなかったといわれる中で、今回は日本からガイドラインの素案を発信し議論を進めた立場となったわけであり、特筆に値するといえるだろう。AI ガイドライン制定における日本の人文・社会学者が果たした役割を鑑みると、そこではこれまでの「輸入学問」の批判は当てはまらな

いと考えられる。しかしながら、この原則が成立する過程の議論からは、日本が日本の文化や社会の現場に基づいて AI 原則を立ち上げたというよりは、AI 技術を「個人の尊厳」というヨーロッパの理念から強く規制しルール作りの面で世界を主導することを試みる EU との連携を深めていこうとする姿勢がうかがえた [16]。実際、総合科学技術・イノベーション会議からは「人工知能・研究イノベーション分野における日 EU 協力の強化」と名されたプレスリリースが 2019 年 5 月に発表され [17]、欧州委員会が掲げた AI の開発

の 7 つの必要条件 [10] は、日本の 7 つの原則 [7] と軌を一にするものであり、「日本と EU は、G7 や G20 のような国際枠組において AI 倫理/原則に関する国際的な共通理解を促すために共に緊密に取り組み続ける」との方針が示されていた。

日 EU のそれぞれのガイドラインでは「人間中心の原則」という基本概念が共有されている。その中でも前述したように EU が掲げる AI 倫理ガイドラインには、AI は人間の自律性 (autonomy) を損なわせてはならず、「個人の尊厳」を守るべきとの思想が貫かれている。一方、日本のガイドラインでは、直接的に「人間の自律性 (autonomy)」という言葉はでてこないものの、「人間の尊厳が尊重される社会」を基本理念としており、「人間が AI に過度に依存したり、人間の行動をコントロールすることに AI が利用される社会を構築するのではなく、人間が AI を道具として使いこなす」[7] との記載があることから、EU と同じ理念を共有しているように読みとれる。

4. JST・RISTEX「人と情報のエコシステム」研究開発領域の取り組み

4.1. AI は「近代西洋」社会の価値規範では統制しきれない

JST・RISTEX では、2016 年より「人と情報のエコシステム (通称 HITE)」研究開発領域 [18] をスタートさせている。HITE 領域は、AI 研究者のみならず多数の人文・社会学者が参画し、哲学・経済・法律など複数の視点から AI をはじめとする情報技術の倫理的・法的・社会的影響評価を実施し、そこからアジェンダ化された課題への対応策を検討することを主旨としており、現在までに 18 プロジェクトを採択している。前述したように、本邦においては EU のガイドラインと足並みを揃え「個人の尊厳」を重要視し AI を人間の道具として捉えた「人間中心の AI 社会原則」が AI 倫理のスタンダードとして定められた。しかしながら、HITE 領域の幾つかのプロジェクトからは、「人間中心の AI 社会原則」が前提とする人間観を基にした倫理や法の限界を指摘し、それとは異なる人間と AI の関係性を基にした提案が成果として導き出されてきている。

AI が事故を起こした際の「責任」分配について根源的な部分から考察を行うプロジェクト [19] によれば、近代の西洋哲学を基盤とする近代刑事法は、主体と客体の厳格な分離を前提に、外界の影響から独立した自由意志をもった自律性 (autonomy) のある人間が事物を統制し、そこから外れた事柄があった場合、事物を統制しきれなかった個人に対して刑事罰を科す、という図式を採用している。しかしながら、ディープラーニングのような学習型の AI は開発者の意図を離れて継続的に学習し発展していくことが明らかになっており、また近年の脳神経科学からは人間は常に外的環境の影響下にあるためそもそも確固とした自由意志の存在は疑わしいとのエビデンスが導き出されている。つまり、近代刑事法が前提とする主体と客体の厳格な分離を前提とした自律性をもった人間による事物の統制という価値規範は根本から揺るがされており、この中で思考している限り AI が引き起こす事故の「責任の空白」という事態の解決は難しいというのだ。このプロジェクトでは、人間と AI が相互作用しながら善なるものを目指す新しい「人間観」を基にした倫理や法律の構築を目指している。

つまり日本が制定した「人間中心の AI 社会原則」は、英語で「発信」し国際会議での日本のプレゼンスを高めることには貢献したものの、日本社会の文化や慣習に基づいた「人間観」というよりは、EU のガイドラインの影響を強く受け、「自律性」をもった「人間」が AI を道具として制御するという主体と客体の厳格な分離を前提として採用したため、AI の社会への実装が深まるにつれその社会受容の阻害要因となってしまう可能性があるともみならずこともできるのではないだろうか。それゆえ、今後の AI 倫理は、主体と客体の分離が曖昧で厳格ではない日本社会の人間と機械との関係性を基に検討していくほうが、西洋のそれを基にするよりも親和性が高いとも見なすこともでき、日本社会を西洋社会から「遅れた」ものとして捉えるのではない新たな視点を世界に提示することの重要性がこれから高まっていく可能性も考えられる [20]。これまでも複数の研究結果から、日本と西洋では人間とロボットや AI の関係性の捉え方に違いがあり、日本において人々がロボットや AI に抱くイメージは、西洋のそれと比較して融和的であることが明らかになっている [21] が、HITE プロジェクトからは西洋の近代的な人間観の基では今後の AI の社会受容が阻害される可能性が指摘されているのである。実際、アバターロボット

MELTANT- α を開発した粕谷昌宏氏は、「日本でのポジティブな反応に対して、CNNに取り上げられた際には、「怖い」とか「仕事を奪われる」といった声が多く見られました。原子力発電所や宇宙空間などでも人間が危険を犯すことなく作業できるようにしたり、人間の仕事に伴うハードルをより低くすることで各個人がより活躍できるようにするための技術なのに、文化的な背景によって受け止められ方が大きく変わってしまう。技術だけでなく、ビジョンを伝えることの重要性や難しさを痛感しましたね。」と、自身が開発した技術の文化的受容性の違いとビジョンをもって技術を伝えることの重要性を述べている [22]。

4.2. 英国リサーチカウンスル UKRI との共同公募の実施

前項でみたように、AI や脳神経科学などの発達は西洋の近代哲学を基盤とした社会システムを根本から揺るがし始めている。また、中国は西洋とは異なる価値規範の基でAI をアグレッシブに活用しこれまでの世界秩序を揺るがしつつある。そしてこれから「データ保護主義」を採用するインドや新興国の台頭により、既存の西洋近代哲学を基盤とする社会システムは益々挑戦をうけていくことになるであろう。世界がこのような大変革期をむかえつつある中、明治維新以降西洋の学問を輸入することが主流であった日本の人文・社会科学のあり方も、これから反省的に再検討される局面が到来する可能性もあると予想される。

実際、AI 倫理分野では日本の人文・社会科学の知見が海外から求められるという事態が到来している。Brexit に揺れる英国は新たな海外パートナーの 1 国として日本を選択し、英国と日本の共同研究の可能性を積極的に模索している。2018 年 7 月には、日英人文・社会科学研究者協力グラント (ESRC-AHRC UK-Japan SSH Connections grants) を設定し、日英の研究者のネットワーキングを拡大する施策を打ち出し、2018 年 9 月には英国の人文・社会科学研究者と HITE 領域の研究者の AI 倫理研究における人材交流のためにワークショップを東京で開催している。

このような中、2019 年度 HITE 領域は、英国の UK リサーチ・イノベーション (UK Research and Innovation, UKRI) 傘下の Economic and Social Research Council (ESRC) と Arts and Humanities Research Council (AHRC) と連携し、AI 倫理分野において日本と英国との共同プロジェクトの公募をはじめた [23]。日本の研究者は JST から、英国側の研究者は UKRI より支援を受けるという形式の公募であり、提案書もメインとなる部分は日英共通で英語での記載が必須となっている。公募期間は 2019 年 5 月 11 日～7 月 11 日であったが、この期間には英国の研究者から JST に日本の人文・社会科学研究に関して多くの問い合わせがあった。現在採択プロジェクトは審査中ではあるが、当初の想定を超える数の応募があり、日英双方の担当者を驚かせている。

4.3. 「発信学問」にむけてこれからの取り組みへの期待

明治維新以降、日本の人文・社会科学研究者が翻訳という作業を通じて母国語での高等教育を可能にしたことは大変な功績であり、今後もこれらの営みが継続することが重要であるということはない。一方で、西洋近代哲学の価値規範では統制しきれない恐れのある AI をはじめとした情報技術の倫理の検討には、日本の文化や社会に根ざした「人間観」や「人間と機械の関係」の知見が活かされる可能性が多いにあると考えられ、我が国の知見を積極的に海外に発信していくことが求められる時代が到来しているといえる。実際、JST・RISTEX の日英共同公募の取り組みからは、AI の倫理分野においては若手～中堅の日本の人文・社会科学研究者が海外からの知見を取り入れるだけでなく、日本の知見を海外に英語で積極的に発信しようという気運があること、また英国側から日本の人文・社会科学研究者の知に対して大きな期待があることが可視化され始めている。

こうした状況の中、研究資金配分機関は、日本の人文・社会科学研究者から海外の研究者と共同で研究を実施し日本の文化や社会に根ざした知を発信していくというボトムアップの気運が立ち上がっていることを見逃すべきではなく、さらにこの動きを大きなうねりにするためにも、海外のリサーチカウンスルとの連携を強化し、その地域の拡充も含めて「発信学問」を確立するための継続的な取り組みを実

施していくことが期待される。

参考文献

- [1] 三成美保, 人文・社会科学研究評価の課題と展望—日本学術会議の成果をふまえて, 学術の動向 (2018/10)
- [2] 荻谷剛彦, 誰のための, 何のための研究評価か—文系研究の日本的特徴, 学術の動向 (2018/10)
- [3] 科学技術・学術審議会学術分科会 学術研究推進部会 人文学及び社会科学の振興に関する委員会 学術研究推進部会, 人文学及び社会科学の振興に関する委員会 (第 12 回) 配付資料 (2009)
- [4] 独立行政法人日本学術振興会 人文・社会科学の国際化に関する研究会, 人文学・社会科学の国際化について (2011)
- [5] 日本学術会議哲学委員会芸術と文化環境分科会, 第 22 期 哲学委員会芸術と文化環境分科会記録 (2014)
- [6] 総務省 AI ネットワーク社会推進会議, 総務省 AI ネットワーク社会推進会議報告書 2019 AI ガイドライン比較表 (2019)
- [7] 統合イノベーション戦略推進会議, 人間中心の AI 社会原則 (2019)
- [8] 宮下紘, 人間とデータのあいだ—プライバシーの居場所, chuoOnline (2017)
<https://yab.yomiuri.co.jp/adv/chuo/research/20171214.html> (参照 2019/9/24)
- [9] The Partnership on AI, Tenets (2016)
- [10] European Commission (High Level Expert Group on AI (HLEG)), Ethics Guideline for Trustworthy AI (2019)
- [11] BDVA, Joint Vision Paper for an Artificial Intelligence Public Private Partnership (AI PPP) (2019)
- [12] European Commission, EU data protection rules (2018)
- [13] 伊藤亜聖, デジタルチャイナ 「第四次産業革命」の中国的展開, 『東亜』 (2019/2)
- [14] Sebastian Heilmann, Leninism Upgraded: Xi Jinping's Authoritarian Innovations (2016)
- [15] OECD Legal Instruments, Recommendation of the Council on Artificial Intelligence (2019)
- [16] 人間中心の AI 社会原則検討会議, 人間中心の AI 社会原則検討会議第 7 回議事録 (2018)
- [17] 総合科学技術・イノベーション会議, 共同プレスリリース: 人工知能・研究イノベーション分野における日 EU 協力の強化 (2019/5/3)
- [18] 科学技術振興機構・社会技術研究開発センター, 「人と情報のエコシステム」研究開発領域
<https://www.jst.go.jp/ristex/hite/> (参照 2019/9/24)
- [19] 稲谷龍彦, 人工知能搭載機器に関する新たな刑事法規制について, 法律時報, 日本評論社 (2019/4)
- [20] 松浦和也・葭田貴子・稲谷龍彦, 人工知能時代の責任と主体とは?, HITE 冊子 vol. 02 (2018/3)
<https://www.jst.go.jp/ristex/hite/topics/img/book-vol2-7.pdf> (参照 2019/9/24)
- [21] Frederic Kaplan, Who is Afraid of the Humanoid? Investigating Cultural Differences in the Acceptance of Robots (2004)
- [22] 粕谷昌宏 (MELTIN MMI 代表取締役), 人類進化を導くサイボーグ技術のインパクト (2019)
https://www.excite.co.jp/news/article/Finders_1176/ (参照 2019/9/24)
- [23] ESRC, UKRI-JST Joint Call on Artificial Intelligence and Society (2019)
<https://esrc.ukri.org/funding/funding-opportunities/ukri-jst-joint-call-on-artificial-intelligence-and-society/> (参照 2019/9/24)