

汎用AIエイアイのソフトローと〈法と文学〉

——SFが警告する〈強いAI／AGI〉エイジー・アイ用規範を巡る記録から——

平野 晋

科学者はとても立派な人達だから、

科学がどこまでやって良いかを決める究極的な責任を

彼等に安心して任せられる、

と思いきむことは、広く共有された誤りである。

実際には、科学者の立派さは、

他の労働者達と同程度に過ぎないのである。

……

科学者は、社会を科学から保護したいと望むよりも、寧ろ

科学の知識を進歩させたいと望むのである。

リチャード・A ポズナー 『大惨事と危険と反応』一九八〇―一九九頁

(オックスフォード大学出版二〇〇四年)(筆者拙訳)

RICHARD A. POSNER, CATASTROPHE: RISK AND RESPONSE 98-99

(Oxford Univ. Press 2004)

はじめに

I フィクションの利用を肯定する〈倫理学〉の指摘例と、〈法と文学〉の指摘例

II ソフトローの対象から〈汎用AI〉を除外すべきと主張・要求して、フィクションを用いた危険性の指摘を非難する経
済界やAI開発系ベンチャー企業

III 〈汎用AI〉もソフトローの対象にすべきと判断した有識者会議

おわりに

はじめに

新技術の登場は、一方では人類への恩恵が期待されるけれども、他方では危険性も懸念される。例えば核科学は、核医学の発展と普及のように医療に貢献したけれども、他方では核兵器も生んでしまった。そして核兵器は、急速に世界に普及(?!)——すなわち〈拡散〉——したのである。

人工知能(AI: artificial intelligence)も、〈制御不可能性〉(un-controllability)、〈不透明性〉(opacity)、及び〈差別的判断〉(discrimination)等々の危険性が懸念されている。⁽²⁾ AIは、その影響力が大きいゆえに、危険性もまた大きいと捉えられているのである。⁽³⁾ 法律家である筆者としては、危険性を無視することは出来ず、これを極小化したいと望むのは極めて自然なことであろう。

尤も、AIの具体的な危険性は、一部は明らかになってはいるものの、その全てが現時点では明らかではない——発展途上の技術だからである——。特に、既に普及し始めている〈特化型AI〉⁽⁴⁾に比べて、ときに〈強いAI〉とも

呼ばれる〈汎用AI〉の危険性についてはなおのこと不透明であり、かつ不安も大きい⁽⁵⁾。しかし将来の危険性に備えて、人々をそこから防護する為には⁽⁶⁾、将来についての、(或る程度)具体的な危険性を十分事前に予見しなければならぬ——具体的に予見できない危険性に対しては、回避策を想起できないか、又は仮に想起できても費用を掛けてまで回避策を採用する決定が下せないからである⁽⁷⁾。そして、或る程度具体的な危険性を予見する為には、〈想像力を働かせる〉必要がある。何故なら抽象的に考えを巡らせても、ヒトは危険性を実感することが難しいからである。そしてヒトが想像力を働かせる為には、〈物語〉(narratives)や〈物語の話術〉(storytelling)等を用いることが効果的である、と後掲するように指摘されている。そのような物語の例として、ディストピアを描くSFや映画作品を用いることが有用である、と筆者は以前から主張してきた⁽⁸⁾。この主張にはもちろん、主にアメリカの〈法と文学〉に代表される先行研究上の根拠も知見も存在する⁽⁹⁾。確かに、嘗ては多少なりとも実現不可能、又は荒唐無稽であると捉えられる向きがあったSF作品でも、今、鑑賞し直すと、意外にも現実感を帯びて再評価されるべき優秀なSF作品も見受けられる。例えば、獣とヒトのハイブリッドのような怪物を、マッド・サイエンティストが辺境の地で創っていた恐怖を描いたH・G・ウェルズの小説『モロー博士の島』(一八九六年)が、遺伝子組み換え技術の可能な現代では空想物語に止まらないことは自明であろう⁽¹⁰⁾。更に例えば、スタンリー・キューブリック監督作品で、アーサー・C・クラークが共同原作者である映画「2001年宇宙の旅」(MG M一九六八年)には——しばしばアンケート調査で最も偉大な映画の一つに選ばれるので⁽¹¹⁾、そもそも荒唐無稽とは一般に捉えられていない名作ではあるが——、「H・A・L 9000型」と呼ばれる、宇宙船搭載の人工知能が登場する。この人工知能は、木星探査ミッションを遂行する為には宇宙飛行士達を殺害することも厭わないという判断を実行するのだが⁽¹²⁾、この判断・行動は、AIに関して近年判

明してきた、〈制御不可能性〉や〈不透明性〉、又は〈フレーム問題〉⁽¹³⁾等の抽象的な危険性の概念を、極めて具体性を伴って視聽者に理解させてくれるのである。⁽¹⁴⁾

このように、未来の危険を予見して、これに備えて人々を防護する試みに於いてSF等の物語を用いることが正当化される動向は、左のような指摘にも端的に象徴されている。

今日の空想科学 (science fiction) は、しばしば明日の科学的事実 (scientific fact) である。⁽¹⁵⁾

今日の科学技術は、昨日の空想科学に近づいている。⁽¹⁶⁾

月旅行は、十八〜十九世紀には単なる空想科学だった。しかし「今」、月探査は四〇年以上も昔の歴史的出来事になってしまった。そして、月着陸と異なって、人工知能の開発は社会の統治に直接的影響があるのだ。⁽¹⁷⁾

(強調付加)

ところでAIの危険性、特に汎用AIのそれを、SF等の物語を用いて語る姿勢に対しては、主に経済界や、AI開発系ベンチャー企業や、一部のAI研究者等が猛烈に反発する姿勢が見受けられる。⁽¹⁸⁾ AIを利益拡大の源泉と捉えて、AIを通じて「功成り名遂げ」ようとする立場からすれば、その危険性が指摘されるだけに止まらず、物語を通じて危険性が説得力さえも増すような指摘は、非常に不都合であるから猛烈に反対し、かつ反発するのである。あるいは、AI研究者の中には、自分の研究分野だけが世界の全てであるかのごとき近視眼的な信念に固執するあまり、機械が生物のように生存本能や増殖本能を発揮することは絶対ないと指摘する傾向もあるようである——しか

し広い世界の中には、生物学と機械とのダイコトミ^二に拘ることなく、両者のハイブリッドで学際的な研究に従事する者達も居るのだから、¹⁹「機械は云々」という近視眼的視野だけで世界の安全性を語る姿勢は、短絡的過ぎるであろう。

ところで物語を用いて「汎用」AIの危険性を指摘して、その危険性に備えて人々を防護する手法は、宜しくないのか。フィクションを用いて不都合な予見をすることが一切禁じられて、危険性を含んだAIの開発や普及を手をこまぬいて看過せねばならないのか。そのような、不都合な批判を圧殺するような態度や言論に対して筆者は、明確に反対の意を表明したい。²⁰そこで本稿は、様々な文献や先行研究を具体的に例示しながら、将来の危険性、特に汎用AIの危険性、の予見と警告に於いてSF等を用いることが正当化されている事実を、記しておきたい。

以下、Ⅰでは、倫理学や〈法と文学〉の分野に於いては既にフィクションを用いることが肯定されている事実を例示する。Ⅱでは、経済界やAI開発系ベンチャー企業等がフィクションの使用を非難しつつ、〈汎用AI〉をソフトローの対象から外すように要求した事実を示す。そしてⅢでは、その要求を退けた有識者会議の判断を紹介する。

Ⅰ フィクションの利用を肯定する〈倫理学〉の指摘例と、〈法と文学〉の指摘例

1 〈倫理学〉に於ける指摘例

そもそもSF作品等の物語を使用することが肯定されるという筆者の指摘は、おそらく「少なくとも一部の」読者の反発を買うであろう。そこで、筆者の主張が全く荒唐無稽又は根拠がない訳ではない事実を理解してもらう為に、

少々長めの引用には成るものの、まずは倫理学に於いて（も）フィクションの使用が肯定される事実を、左の引用から直接読み取っていただきたい。

……、リアリティを感じるためにフィクションを用いる、というやり方に違和感を感じる人がいてもおかしくない。……。「しかし」たとえば、……、戦争の経験をしたことのない人は、……優れた映画（たとえばステイヴン・スピルバーグ監督の……『プライベート・ライアン』（一九九八年）……を見ることで、道徳的問題を具体的に思い描き、リアリティをもって思考できる、という効果はさまざまに見られる。

私たちが、人間とは何かを問うたり、自らの生き方について反省的に思考したりする実に強烈な瞬間は、……フィクションの鑑賞を通じて現れることが少なくない。……、フィクションの作中人物の場合、その人物の経験は……、しばしば私にも起こるかもしれないものとして受け止められる。……。

2 SF映画の効用

仮想的状況と思考実験——あなたはどうか考えるだろうか

いかにもありそうではあるが、現実には生じてはいないこと——SFが「思考実験」という倫理学に特有な方法に向いているという事情がある。「思考実験」においては、「もし〜だったら」という仮想的な状況設定のもとで、その場合に、あなたはどうか思うだろうか、どうするべきだろうか、などを考えるように求められる。……。

思考実験は、極端で非現実的な状況を想定させることで、かえって、私たちがどういう風に物事を見ているのか、その直感を明らかにする。たとえば、……「ノーベル文学賞を受賞した日系英国人カズオ・イシグロ著の」「わたしを離さないで」……。そこでは、クローン人間たちが将来臓器提供するために寄宿制の施設で育てられているというSF的状况……で、ク

ローン人間を、移植用臓器を得るための手段として利用し尽くして良いのか、が問われた。それによって、現実には私たちが面している臓器移植をめぐる道徳的問題をありありと考えるきっかけが、提供されたのである。

仮想的な状況設定の意味解釈——私たちは何を考えるべきなのか

SF映画の効果は思考実験を促して、私たちの道徳的直観を明らかにするのに役立つ、というだけではない。一体、その空想上のシーンを設定することで、世界や人間についてその映画は何を言わんとしているのか。その仮想的状況は何を描きだしているのか。「もし」であれば「の」の部分自体を私たちに解釈させることにより、SFは、私たちに、今、あるいは今後、何がシリアスな問題になりうるのかを示唆する。あるいは、何を今考えるべきなのか、あるいは何を今後考えていくべきなのか、そういう問い自体を強いる。

(強調付加)

いささか長い引用で恐縮だったが、読者には是非とも、SF作品のような〈フィクション〉から学ぶことが、倫理学では学術的にも正当化されている事実を、倫理学の分析から直接理解して欲しかったので、敢えて長文にわたる引用を試みた。筆者が付加した強調部分から特に読み取って欲しい事実の一つは、SF作品を通じてAIの将来の危険性を具体的に予見できる点である。すなわち、「もし」だったら」という仮想的な状況設定のもとで、その場合に、あなたはと思うだろうか、どうするべきだろうか、などを考えるように求められる。」という点が重要である。⁽²²⁾ 例えば、映画「エリジウム」(ソニー・ピクチャーズエンタテインメント二〇一三年)では、マット・デイモン演じるスラム街の青年が、工場労働者としてまっとうな日々を送っているけれども、少年時代の前科を理由にロボット警官やロボット保護観察官から、不当な職務質問や保護観察期間延長を言い渡されるような差別を受ける。その結果、負のスパイラルから抜け出られないという、後掲「バーチャル・スラム」が現実化したディストピアな未来を描

かれている。マット・デイモンに感情移入する観客は、ロボットが警官や保護観察官等々の法執行機関に利用される未来に於いては、〈慈悲の心〉や〈寄り添う感情〉等を欠いた冷酷な法執行がなされないように、と自然に思いを馳せると思われる。すなわち、前掲倫理学の文献引用文が云うところの、「今後、「AIの」何がシリアスな問題になりうるのか」「あるいは、何を今考えるべきなのか、あるいは何を今後考えていくべきなのか、そういう問い自体を「映画「エリジウム」は」強いる」のである。

更にこの法執行上の論点は、既に今、有罪が決定した被告人の量刑や収監者の仮釈放を決定する際に、AIを用いることでアフリカ系アメリカ人に不当で差別的な判断・法執行がなされることへの批判にも通じるのではあるまいか。そのようなディストピアにならないように、AIやロボットの利活用では配慮が必要であるという考えに、良質なSF映画は導いてくれる。これは倫理学が有用であると捉える「思考実験」がもたらしてくれる、望ましい成果であると評すべきである。

ところで、このように有用な思考実験は、危機管理等の分野に於いて「シナリオ分析」という文言で言い換えることが出来る手法に近いものかもしれない。なお「シナリオ（物語）」と云えば、法学に於いても実際のところ、憲法学の立場からAIの危険性を指摘する慶大・山本龍彦教授も、未来のSF的な仮想「シナリオ」を用いて、「バーチャル・スラム」というディストピアな社会を現実化させるAIの危険性に対して警笛を鳴らして⁽²⁵⁾おり、非常に説得的である。このように日本の法律学に於いても既に、SF作品的な物語を用いての、具体的な危険性の指摘・分析が実際に表明され、かつ説得力を伴っている事実を、ここで指摘しておきたい。

ところで、引用文中の「仮想的な状況設定」という文言から、倫理学の「思考実験」と法律学の教育手法との近似

性——特にアメリカの法律学教育手法と倫理学との親和性——を筆者は感じている。すなわち、アメリカのロースクールに於ける教育では、「hyp^{ハイ}o^ホ」(hypotheticalの略語)と呼ばれる〈假想事例〉を用いた事例研究(case studies)を用いて法的思考力(think-like-a-lawyer)を養成する手法が、伝統的に採用されて来た。⁽²⁶⁾先例から帰納法的に導き出させた規範(rules)を、ハイボ上の争点(issues)にあてはめた場合にどのような結論に至るのか。そのような、教室内に於ける〈ソクラテス方式〉と呼ばれる問答を通じて、演繹的な〈あてはめ：application of the rules〉と呼ばれる法的思考力も法学生に養成されるのである。他方、倫理学に於ける「思考実験」に於いても、「假想的な状況設定」に自らが置かれた場合に、どのように判断・行動するかが問われるので、その「判断・行動」を採る基本原則としてどのような(例えば功利主義のような)倫理的指導原理——すなわち規範——をあなたが採用するのか・支持するのか、明らかにするのである。双方共に、〈假想的〉である〈事例〉——すなわち〈フィクション〉である〈物語〉——を道具として用いることにより、抽象的な規範が有する具体的な意味・効果の理解を深化させる点に於いて、非常に似通っていると筆者には感じられる。

なおアメリカの法学教育では実際に、〈物語的思考〉(narrative thinking)や〈物語の話術〉(storytelling)が法曹として備えるべき技能として重要であると指摘されて来ている。⁽²⁷⁾これは、陪審裁判制度が今でも重視され、かつ実践されているアメリカの裁判制度に於いて裁判実務弁護士(trial lawyers)に不可欠な技能は、陪審員及び裁判官を説得する能力であることを考えれば納得がいく——〈物語的思考〉や〈物語の話術〉を知りかつその技能を高めれば、陪審員や裁判官に対する〈説得力〉の向上に繋がるからである——。

更に次項で触れるアメリカの〈法と映画研究〉分野に於いては、法律に関する映画が、ロースクールに於ける法曹

教育でも有用であると指摘されて、フィクションである映画の使用が肯定的に捉えられている事実も記しておこう。

3 フィクションの利用を肯定する〈法と文学〉の指摘例

アメリカの法律学に於いては古くから、文学に学ぶという学際法学 (Law and Literature) 研究の伝統が存在する。⁽²⁹⁾ 例えば、〈証拠法〉で高名なジョン・H・ウィグモアは、文学の重要性を説き、指摘しつつ、江戸時代の法制度に興味を抱いて貴重な研究成果をあげた事実は、⁽³⁰⁾ 特に法制史又は比較法研究者の間では有名な話であろう。もつと近年になってからも、例えば現在急激に脚光を浴びて隆盛の極みにあるプライバシーや個人情報保護法の分野で云えば、ジョージ・オーウエルのSF小説『1984』(一九四九年)に登場する「ビッグ・ブラザー」に言及することは、いわば枕詞として常識と云える程にその慣行——SF作品を引き合いに出す慣行——に疑いを差し挟むことさえも見受けられない程に、研究者仲間に受容されているのである。(実際「ビッグ・ブラザー」に象徴される、最新技術を用いた高度に全体主義的監視社会が、最早フィクションではなくファクトである現状は、隣国を見れば明らかである。) 古^{いにしへ}から大学に於ける三大専門分野として、〈神学〉及び〈医学〉と共に学問を深化させて来た〈法学〉が、文学への言及や比喻を軽視して来なかった事実を、経済界とAI開発系ベンチャー企業は改めて学ぶべきであろう。

ところで近年では、アメリカの法律学に於いて、〈法と文学〉(Law and Literature)という独立した学際法学さえ存在する。⁽³¹⁾ この分野は多くの場合〈文学の中の法〉(Law in Literature)と〈文学としての法〉(Law as Literature)の二つの中分類に分けられる。〈文学としての法〉は、更に次の二つの小分類に分けられる——小分類の一つ目は、文学に於ける文芸批評の手法を法解釈に援用しようとするもの。そして二つ目の小分類は、文学の〈修辞法〉(Theoric)や、

〈物語の話術としての法〉(law as storytelling) や 〈法的な物語の話術〉(legal storytelling) や 〈ナラティブ法学〉⁽³²⁾ 等と呼ばれるものである——。ところでこの分類は絶対的なものではなく、諸説が存在する⁽³³⁾。しかし〈法と文学〉研究の共通点を大まかに言い表せば、皆、物語に関心を寄せていると要約することが出来る⁽³⁴⁾。なお〈法と文学〉から分派した研究分野には、〈法と映画研究〉(law and film studies) と呼ばれる分野も見受けられる⁽³⁵⁾。更にアメリカ法学に於いては、SFを題材とすることを肯定する指摘も見受けられる⁽³⁶⁾。

紙面と時間の制約から、本稿では法と文学の詳細に触れることは敢えて避けて、物語に関心を寄せるその姿勢が如何に有用であるかの例を挙げてみることにしたい。例えばスケプリは、「物語 (Stories) [を用いれば] ……想像力 (imagination) と新しい視点 (new points of view) を導入することが出来る」(強調付加) と指摘する⁽³⁷⁾。例えばアフリカ系アメリカ人に対する差別的な現状が治癒されることなく、当たり前のように受容されている風潮に対して、〈物語〉を用いればその風潮を改めさせることに資すると指摘されている⁽³⁸⁾。これは、当原稿を書いている同時代的事件に鑑みても理解できる指摘である、と筆者は思う。すなわち、白人警官がアフリカ系アメリカ人に対して過剰な暴力を用いて死に至らしめてしまう〈具体的な事件〉——つまり「物語」——を通じて、「BLM」(Blacks Lives Matter) と呼ばれる反対運動が全米に広がった現状に鑑みれば、私達にも理解できる指摘であろう。単に抽象的に、〈差別が存在する〉と云うよりも、具体的に白人警官が権力を濫用して不当にアフリカ系アメリカ人を殺害したミネアポリスのジョージ・フロイド事件を例示した方が、人々の心に深く訴え掛けて、社会の不正を正す運動に繋がる効果が生じたのである。すなわち物語は、その説得力に於いて、抽象よりも勝ると云える。

この論理を、AI・ロボットの未来の危険性議論に類推適用してみれば、物語を用いることが正当化されると読者

にも理解してもらえらると思う。例えば、AIの危険性の一つとして、AIには制御可能性が欠けていると指摘されている。しかし、例えば私が国際情報学部にて担当する二年生科目の「AI・ロボット法」に於いて、学生に対し、単に「AIは制御可能性を欠いている」と抽象的に教えても、学生はその瑕疵の重要性を認識しにくいかもしれない。記憶力に長けた学生ならば、やみくもにAIは「制御可能性を欠く」と記憶して試験に臨むかもしれないけれども、制御可能性を欠くことがどのように大きな危険性を顕在化させて、だからこそその瑕疵は費用を掛けて迄も事前^{proactive}に抑え込まねばならないという視点を、俄^{にわ}かには抱けないかもしれないのである。他方、映画「ロボコップ」オリジナル版（ワーナー・ブラザーズ一九八七年）の冒頭部分に登場する試作品の警備ロボットが、その製造業者たるオムニ社の重役会議のプレゼン会場で突如誤作動して若手重役を銃殺する場面を観てもらえば、制御可能性を欠く危険性の大きさを実感してもらえらるはずである。⁽⁴⁰⁾この例のように、SF作品を含む〈物語〉は、説得力を伴って問題の大きさの理解に資するのである。

なお右の例は、認知科学的にも実証されている事実を、ここに記しておきたい——すなわち、ヒトは抽象的に世界を理解することが不得手であるけれども、物語を通じれば理解も記憶も深まり、効果的な学習方法であり、かつ説得力が高いと指摘されているのである。⁽⁴¹⁾

尤も筆者は、物語の利活用をやみくもに主張している訳ではない。例えば物語の利活用は、⁽⁴²⁾〈偏見〉という弊害を伴う場合もあり得るので、その弊害は考慮せねばならないと筆者は考えている。具体的には、『連邦証拠規則』第四〇三条（FEDERAL RULES OF EVIDENCE 403）の考え方を類推して、物語の使用はそれが生む不当な偏見（unfair prejudice）等よりも証明力（probative value）が上回る場合に限定すべきである、と既に他の拙稿に於いて提案済みなので、こ

ここで繰り返し返すことは差し控えたい。

II ソフトローの対象から（汎用AI）を除外すべきと主張・要求して、フィクションを用いた危険性の指摘を非難する経済界やAI開発系ベンチャー企業

既に拙書にて紹介した通り、AI等を用いた自律型「ロボット兵器」を禁止すべきか否かの国際的な議論に於いては、国際的な人権団体等が、「ターミネーター」等々のSF映画を例示しながら反対論を展開する事例が複数見受けられる——そしてそのようなSF映画を用いた反対論に対しては、ロボット兵器賛成論者から、SF映画を例えに用いることへの批判が散見される——⁽⁴⁴⁾。

このようにフィクションを毛嫌いし忌み嫌う傾向は、ロボット兵器賛成論者に限られた現象ではない。広く民生用に普及するであろうAIやロボットを巡る議論一般に於いても、主要関連業界は「Partnership on AI」と呼ばれる企業連合を形成して、SF作品が描くAIのディストピアな印象を払拭すべく正の面を強調している、と指摘されている。⁽⁴⁵⁾

1 経団連からの反発・主張の具体例

主導的グローバル巨大ICT企業連合「Partnership on AI」に呼応するように、国内の経済界も、AIの危険性、特に汎用AIの危険性をSF等を通じて語る姿勢に対し反発している。例えば（一社）日本経済団体連合会のAI活

用原則タスクフォース（TF）は、左のような主張を公表している⁽⁴⁶⁾。

今まで発表されているAI原則を俯瞰すると、それらは基本的にAIのリスクや負の側面に焦点を当てたものになっている。確かに、AI技術は、その影響の大きさから勘案し、その潜在リスクを認識し、適切な対策を講じておくことは重要である。同時に、その対策がAIの社会問題の解決への導入を阻害することになっては、本末転倒となる危険性もある。例えば、基礎医学領域においての「ヒトES細胞の樹立及び使用に関する指針」の場合にみられるように規制の存在によって国内での研究が遅延するなどの事例が見られている。AIに関しても、実現のめどが全く立っていない技術を前提とした脅威論が展開されるなど、不正確なサイエンスとテクノロジーの理解に基づいた議論も散見される。本TFでは、正確な技術の理解と将来の見通しのもとで、バランスのとれた議論を展開したい。

（強調付加）

なお右の文書の二頁目では、引用部分に続けて、識者がAIの〈正〉の面を以下のように指摘していると書かれている。「……、Google DeepMindのファウンダー&CEOデミス・ハサビス博士は、人工知能は人類の存続のために必須であると主張している」（強調付加）と。しかしこの博士は、強調付加部分が示すように、前掲「Partnership on AI」連合に所属してAI開発を推進したいグローバル巨大企業の人物である。更に、この文書が指摘を怠る複数の世界的に著名な識者達は、この博士とは真逆にAIの危険性を指摘していて、後者の方が実は有名な事実であることを、ここで指摘しておく——すなわち、自動運転車で有名なテスラ・モーターズのイーロン・マスク氏や、ビル・ゲイツ氏や、理論物理学者の故ステイヴン・W・ホーキングズ博士も含めて、複数の識者達がAIの危険性を指摘している事実の方が非常に有名である⁽⁴⁷⁾。しかし右の文書は、その指摘を全く怠っているのである。

2 AI開発系ベンチャー企業からの反発・主張の具体例

AI開発系ベンチャー企業の代表者も、左のように指摘していた。⁽⁴⁸⁾

一つの考えは、技術的な実現可能性が見えていない、汎用人工知能を対象とするというものである。例えば総務省「未来デザインチーム」が二〇三〇年～二〇四〇年を想定して創作した小説「新時代家族」[□]には家庭で使われる擬人化されたAIロボットの登場している。残念ながらこのような擬人化されたAI、あるいは未知の問題を臨機応変かつ創造的に解くことのできるAI（汎用AI、強いAIなどと呼ばれることもある）の存在はあくまで仮説であり、現時点では技術的な実現可能性への道筋が見えていない。そのような技術ができるという仮定の下での議論は……、議論の枠組みが現実的なリスクやベネフィットを反映しているかどうか、十分な注意が必要である。特に、フィクション作品を下敷きにしたような議論は題材として平易であるがゆえに、実際の社会に起こりうる現実的リスクからかえって目を背ける結果につながりかねない。より具体的な個々のシナリオに基づいた議論が必要であろう。

（強調付加）

なお筆者は引用された「未来デザインチーム」に参画していないし、その内容については不知であるが、筆者も座長代理又は幹事・分科会長を務めていた（る）総務省の他の有識者会議「AIネットワーク化検討会議」や「AIネットワーク社会推進会議」の名譽の為に付言すれば、同有識者会議は「実際の社会に起こりうる現実的リスクからかえって目を背ける」ような検討はしていない——その成果は公開しているばかりか、OECDやG20等を通じた国際標準の形成にも貢献したので、AI開発関係者はそれを是非勉強して欲しい——。

3 グローバルな巨大コンピュータ企業からの反発・主張の具体例

A I の開発に対しては、ハードローによる事前規制が開発を萎縮させるという強い反発が経済界から強く、それならばソフトロー（A I 開発／利活用「原則」や「ガイドライン」等）の構築に、主導的グローバル巨大 I C T 系諸企業等にも自主的に参画してもらって遵守してもらった方が得策であることから、総務省に於いて、ソフトロー構築の有識者会議——前掲「A I ネットワーク化検討会議」や「A I ネットワーク社会推進会議」——が二〇一六年頃から現在まで継続的に開催されている。その議論の中で、或るグローバル巨大コンピュータ企業も、前記と同様な態度を執っていたことが、左の記録からも読み取れよう。⁽⁵⁾

【●●「人名」●●「企業名」特別顧問（●●「人名」構成員代理）】

・ ガイドライン案が……、国際的な議論において、どのように扱われるか不透明であるとともに、非規制的・非拘束的であるとしても、その内容を遵守するとなると、事実上、研究開発の自由度を制約する、あるいは、萎縮させる可能性があることを研究する立場からは懸念する。

・ 汎用 A I が、どのようなものであるか分からない、どのようなリスクがあるか分からない段階で汎用 A I を「含み得る」とすること、すなわち、現時点で存在していないもの、すぐには実現しそうにないものを対象にすることの方が強く疑問がある。実現が十分に見込まれる段階になってから検討することで問題ないのではないか。

・ ……。

（強調付加）

Ⅲ 〈汎用AI〉もソフトローの対象にすべきと判断した有識者会議

1 未だ実現していない（SFのような？）汎用AIも対象にすべき

汎用AIのように未だに実現化されていない新技術「の可能性」は、たとえ強制力を欠くソフトローのような緩やかなルールであっても、その対象にすべきではないのだろうか。特化型AIよりも未知であるだけに、大衆の不安も大きく、かつ危険性が大きいかもしれない危険源を無視することが正しい政策であろうか。筆者が参加した有識者会議は、汎用AIをソフトローの対象から除外すべしという主張を退けている⁽³²⁾。左の引用は、再び少し長くなって恐縮ではある。しかしその引用から、たとえ実現化されていない新技術の可能性であってもソフトローの対象にすべきであると有識者会議が結論した理由を、広く本誌読者にも理解していただきたい。

【平野幹事「中大・筆者」（開発原則分科会長）】

● 国際的な議論のためのAI開発ガイドライン案において、AIの定義に汎用AIを含めるか否かについて、……● 社の主張を記述するのであれば、七月二〇日に開催された合同分科会において示された「ガイドライン案は、利用者からみても納得できる常識的なものである。このような内容のガイドラインに賛成できないという開発者の姿勢は、利用者に対する安を与えることになる」という● 構成員「消費者団体代表」⁽³⁴⁾の反対意見も記述すべきである。

● 汎用AIを対象に含まないとすると、汎用AIに関しては、AI開発原則（九原則）を守らなくてもよいということになってしまうので、● 社の主張は内容的にも受け入れられるものではない。

● ……。

【大屋構成員「慶大・法哲学」（影響評価分科会長代理）】

- ・……、汎用AIが現時点で存在していないからガイドライン案の対象から外すというのは一般的な考え方ではなく、むしろ、不確定要因が大きいのであれば、基本的にガイドライン案の対象に含めておいて、安全性が確認できたところで対象から外すというのが一般的な考え方である。さらに指摘があったのは、汎用AIを対象から外すことにより、偽汎用AI、自称汎用AIが、AI開発原則を守らなくなるおそれが生ずるということである。
- ・汎用AIが現時点で存在しないからといって、ガイドライン案の対象に含めないとする合理的な理由はない。他方、過剰な規制、不適切な規制を避けるために、対象とする範囲、内容を不断に見直すことは必要であり、それを前提として汎用AIも対象に含めるべきである。
- ・……。

【堀幹事「東大・AI研究」（開発原則分科会技術顧問）】

- ・米国と欧州との中間的な立ち位置で、非規制的なガイドライン案を作成しようと議論を積み重ねてきた。ここに至って、汎用AIを対象から外すという後退した印象を与えることは望ましくない。……。

【中川幹事「理研・AI研究」】

- ・AI開発原則の内容は非常に常識的なものであり、汎用AIでも特化型AIでも守ってもらいたいものである。
- ・I E E E「アイ・イー・イー・イー」「55」の議論では、当然汎用AIが射程に入ってくることを意識しているものと思われる。したがって、技術開発のスピードや諸外国の議論を踏まえても、汎用AIと特化型AIとの線引き「を」することが妥当かどうか、取り立てて汎用AIを区分して考える必要はないのではないか。

【●●「氏名」●●「企業名」取締役常務執行役員（●●「氏名」構成員代理）】

- ・ガイドライン案は、非規制的・非拘束的なものであることが謳われており、また、不断に見直すこととなっているので、汎用AIを対象に含めてよいのではないかと考えている。

・ AIについては、使っているうちに振る舞いが自律的に変化するため、現実的には利活用の段階で問題が生じる可能性が高いものと考えられる。今後は、利活用を検討する中で、汎用AIについてきちんと議論する必要がある。

【山本構成員「慶大・憲法」】

・ 汎用AIが今後どのようなことになるのか予測できないが、だからこそガイドラインという基本的な考え方となるものが必要であり、そのことによって、汎用AIの研究・開発がさらに進められるのではないかと考えている。

・ 汎用AIをガイドライン案の対象に含めたとしても、汎用AIの将来性を危惧するものではない。是非、日本の主張として世界に発信すべきである。

(強調付加)

ところで特化型のみならず汎用AIについても危険性に留意すべきであると指摘するのは、右のような国内の識者達だけではなく、左の例のように海外でも見受けられる事実も指摘しておきたい。⁽⁵⁶⁾

一九八八年にノーベル賞受賞物理学者リチャード・フェインマンが亡くなったとき、黑板には次のように書かれていた。『私が創れないものを、私は理解できない』と□。すなわち、人類が汎用AIを創るまでは、汎用AIは人類の理解の範疇を超えることになる。このことは実際、皮肉な運命を人類に提示している。汎用AIを管理する為には、人類が先ずは汎用AIを理解しなければならぬ。しかし汎用AIが創られるまでは、人類はそれを理解できないのだから。更に、マックス・テグマークによれば『もし人類が人間と同等なAIを創り出すことに成功した場合に、何が起こるかは全く分からない』□。従って、もし汎用AIが創られた場合に、その結果が正である(the outcome will be positive) 当たり前前に捉えるべきではない□。

(強調付加)

右の引用は、前掲した総務省の有識者会議に於いて、某グローバル巨大コンピュータ企業が、汎用AIはソフトローの対象外にせよと主張した要求が、やはり的外れであった根拠にも成る。すなわち同企業は、「汎用AIが、どのようなものであるか分からない、……、すなわち、現時点で存在していないもの、すぐには実現しそうにないものを対象にすることの方が強く疑問がある。実現が十分に見込まれる段階になってから検討することで問題ないのではないか。」と主張していた。しかしこれに対して、右で例示引用した海外での指摘は、現時点では汎用AIの特性が分からない上に、実現されてその特性が判明した段階では〈負〉の影響が出る可能性を否定できないと分析している。すなわち、仮に手をこまぬいて汎用AIの出現を待つて、出現真近に成つてから考えろという態度を採つた場合には、安全対策は〈出たとこ勝負〉となる訳である。しかし〈出たとこ勝負〉で対策が間に合うとは、筆者には思われない。〈負〉の影響が出る可能性がある以上、慎重に、事前in advanceに出来るだけの留意は払っておくべきである。前掲した総務省の有識者会議に於いて識者達が指摘したように、特に常識的な内容のソフトローを汎用AIにも遵守してもらうことは、最低限必要であろう。

2 エルシー——AI開発者も習得すべき人文・社会科学の素養

AIの開発や利活用の促進に於いては、理数工学系教育が重要であり、「Science, Technology, Engineering, and Mathematics」の頭字語である「STEM」教育を重視するような記述を近年散見する。しかし、AIは社会への負の影響も懸念されているから、「STEM」教育だけでは不十分である。負の影響を極小化すべく、倫理や法や社会への影響に関する教育・研究も非常に重要である。この分野は「Ethical, Legal, and Social Implications」の頭字語

から「ELSI」と呼ばれる⁽⁵⁷⁾。

その「ELSI」が重要であることは、内閣府の有識者会議——「人間中心のAI社会原則検討会議」——に於いても、例えば次のように指摘されている。

……、AIに関わる……経営者は、……、AIの……、社会的に正しい利用ができる知識と倫理を持つていなければならない。
……。一方、AIの開発者側は、AI技術の基礎を習得していることが当然必要であるが、それに加えて、社会で役立つAIの開発の観点から、……規範意識を含む社会科学や倫理等、人文科学に関する素養を習得していることが重要になる。

(強調付加)

右に引用した文言、すなわち「AI開発者側は……規範意識を含む社会科学や倫理等、人文科学に関する素養を習得していることが重要である」部分の挿入を、有識者会議「の構成員起草者達の間」で先ず提案したのは、実は筆者自身である。そこに込めた起草者の意図は、法学や倫理学の知見の一部である、〈法と文学〉、〈物語〉、及び〈思考実験〉等の知見もAI開発者は学ばねばならないという願い——すなわちフィクションからも学んで欲しいという願い——である。

新技術が生む未来の危険性に対しては、前述の通りSF作品を含むフィクションと物語を用いる手法が、人文科学と社会科学に於いては認められている。例えばAIの倫理を含む〈機械倫理〉(machine ethics)の分野に於いても、アメリカの法学に於いても、SF作家アイザック・アシモフが考案した「ロボット工学三原則」に言及する例は多い⁽⁵⁸⁾。そして、右の引用部が勧告しているように、AIの開発者側はそのような人文科学と社会科学の「素養を習得し

ていることが重要」になるから、SF作品を含むフィクションと物語を用いて危険性を予見する手法も認めることが、重要になる。

更に、前掲引用した総務省有識者会議の議論や結論が指摘している通り、AI開発系ベンチャー企業や経済界は、人文科学及び社会科学の研究成果が警告する危険性に対して真摯に向き合わねばならない。経済発展と利益追求にとっては不都合だからといって、人文科学及び社会科学の知見を無視又は軽視する姿勢は、嘗て開発者達に核兵器の開発を許して人類に対する大きな罪を犯させた失敗を再度繰り返させることにもなりかねない。そのような罪を再度犯させる前に、人文・社会科学から指摘される危険性に鑑みて、そもそもそのような危険性を治癒させないままAIの開発又は利活用が許容されるか否かをAI開発系ベンチャー企業や経済界は、検討しなければならぬ。

筆者のこの思いは、実は彼らが忌み嫌うSF作品からも学ぶことが出来る事実を、左の〈法と文学〉の業績上の指摘から紹介しておこう。⁽⁵⁹⁾

科学者達も、「ファウスト博士、フランケンシュタイン博士、モロー博士、……、「及び」異常な愛情博士 (Dr. Strangelove)」
達に連なる民間伝承の「警告と正面から向き合う」必要がある。

引用された架空の科学者達に共通する問題は、皆、非倫理的に行動した点である(本稿冒頭部のR・ボズナーからの引用を想起して欲しい)。彼等、博士達——ファウスト、フランケンシュタイン、モロー、及び異常な愛情博士——は皆、経済界やAI開発系ベンチャー企業が忌み嫌う〈フィクション〉上の人物ではあるけれども、科学者が倫理的に行動することの重要性を教えてくれる。〈法と文学〉という研究分野の効用の一つは、〈倫理的な判断〉(moral judgment)

を育成する為の有効な手段の一つになる点であると指摘されている。⁽⁶⁰⁾従って、モロー博士達のような非倫理的な姿勢は厳に慎むべきとの教訓が、〈フイクション〉には含まれている事実を、AI開発者系ベンチャー企業や経済界にも学んでいただき、「倫理的な判断」を下せる存在に成って欲しい、と筆者は願うばかりである。

そのような、人文・社会科学のな考慮や検討を欠きつつ、危険性を包含したままAI開発又は利活用に向けて盲目的に突き進む姿勢は、有識者のみならず消費者からも受け入れられないことは、総務省有識者会議の前掲引用議事概要からも自明であろう。AI開発又は利活用が経済発展と利益追求に好都合だからといっても、社会全体から〈信頼〉(Trust)を得ることが出来なければ、そのような商品は市場に受け入れられず廃れるであろうし、また、廃れるべきであろう。

国際機関である経済協力開発機構(OECD)が、「信頼に値するAI」(Trustworthy AI)を目指すべきであると、日本を含む先進加盟各国とこれに同調する合計四二か国とそれ以外の非加盟諸国に対しても勧告する内容は、社会が抱く不安の払拭にAI関係者達が努めるべきであるというメッセージである。⁽⁶¹⁾AI開発系ベンチャー企業や経済界は、AI開発と利活用、そしてその先にある利益と経済発展に通じる理数工学系STEM教育ばかりを強調する姿勢を厳に改めて、AIの危険性を抑制又は極小化することにより、社会がディストピアに繋がるおそれを払拭して、「人間中心な」AIを目指すべく、人文科学と社会科学の知見とELSIの重要性を認め、これを真摯に学び、かつ尊重すべきである。

前掲にて少しだけ触れた、アメリカに於ける〈法と映画研究〉⁽⁶²⁾は、映画等の大衆文化に法曹や法制度に対して大衆が抱くイメージが反映されていると指摘する。その上で、映画等の大衆文化に反映されている大衆イメージを法曹は真摯に受け止めるべき、と〈法と映画研究〉は主張する。何故なら大衆は、法曹にとつての将来の依頼人か又は陪審員の候補者達であり、その候補者達が法曹と法制度を如何に捉えているかを知りかつ理解しなければ、「裁判で陪審員を説得したり、依頼人にとつて最善の役務を提供するという弁護士」職務を全う出来ないからである。

このような〈法と映画研究〉の教えを、経済界とAI開発系ベンチャー企業も真摯に受け止めるべきである。すなわち、彼らのAI商品を購買又は使用することになる将来の顧客は、延いては消費者や市民達になるのであるから、そのような顧客が抱く不安を取り除いて〈信頼〉を得られなければ、彼らのAI商品は結局売れないことになる。大衆の不安が全てフィクションに左右された虚構である等とも読めるような主張に固執して、大衆の不安を無視したり、軽視したり、又は圧殺するような態度や行動が、大衆の支持を得られるとは到底思われない。寧ろ、アメリカ法学の知見の一つである〈法と映画研究〉の教えも真摯に受け止めて、丁寧不安を取り除き、欠点を改善して、安心・安全を追求・向上させる姿勢こそが、大衆から求められていると捉えるべきである。

おわりに

〈はじめに〉で紹介した「2001年宇宙の旅」の話題に戻って、本稿をしめくくってみよう。コンピュータ・サイエンスの研究で有名な、あのMIT、マサチューセッツ工科大学は、『HAL'S LEGACY: 2001's COMPUTER AS DREAM AND REALITY』(David D. Stark ed. 1997)と題する本を出版している。すなわちAI開発系ベンチャー企業と経済界が忌み嫌う〈フィクション〉を、あのMITさえも(！)語っている事実を、彼らはおそらく知らないからこそ、フィクションを毛嫌いするような主張を展開したのかもしれない。AIの開発や利活用の促進に於いて重要な教育は、STEMだけではない。AIの暴走や負の側面を極小化して、ヒトがAIによって不幸になることを避けつつ、「人間中心の」AIや、「信頼に値する」AIが不可欠である。その為に必要な教育・研究は、人文科学及び社会科学であり、それはELSIと呼ばれる。この私の主張は、日本及び国際社会に於いて次々とELSIを重視したAI諸原則やガイドライン等々がとりまとめられて公表されて来た近年の動向で裏付けられている。

従ってAI開発系ベンチャー企業と経済界は、人文科学や社会科学の知見や先達の業績——もちろんそこには倫理学が重視する〈思考実験〉や、法律学が重視する〈法と文学〉も含まれる——を軽視・無視する姿勢を改めて、大衆イメージ——すなわち民の声——に真摯に耳を傾けて、その不安を取り除き、安心・安全なAI開発や利活用の普及に励んで欲しい。

- (1) SF作品が核兵器の危険性を指摘する」として、社会的に有用であると指摘する文献としては、*see, e.g.,* Bruce L. Rockwood, *Law, Literature, and Science Fiction*, 23 *LEGAL STUD.* F. 267, 270 (1999) (「[W]riters like Robert A. Heinlein and Philip K. Dick helped readers understand the threat of the Atom Bomb and the risk of the Cold War turning hot」)と指摘)及び拙書『ロボット法(増補版)〜A Iとヒトの共生にむけて〜』四九―五〇頁(弘文堂二〇一九年)(スタンリー・キューブリック監督作品映画「博士の異常な愛情または私は如何にして心配するのを止めて水爆を愛するようになったか」(コロムビア映画一九六三年)を例示している)。
- (2) 例えば、総務省・A Iネットワーク社会推進会議「報告書二〇一七」七頁(平成二九年七月二八日) https://www.soumu.go.jp/main_content/000499624.pdf (最終アクセス令和二年七月三〇日)、及び同「報告書二〇一九」三六頁(令和元年八月九日) https://www.soumu.go.jp/main_content/000637096.pdf (最終アクセス令和二年七月三〇日)等参照。
- (3) *See, e.g.,* Brandon W. Jackson, *Artificial Intelligence and the Fog of Innovation: A Deep-Dive on Governance and the Liability of Autonomous Systems*, 35 *SANTA CLARA HIGH TECH. L.J.* 35, 41 (2019); Mathew U. Schere, *Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies*, 29 *HARV. J. L. & TECH.* 353, 362 (2016).
- (4) そもそも「A I」の定義は定まっていなければならないけれども、一般に、特定の仕用に特化した〈特化型A I〉(Artificial Narrow Intelligence: A N I)と、ヒトのような汎用性のある〈汎用A I〉(Artificial General Intelligence: A G I)に分けられる。*See, e.g.,* Michael Guihot et al., *Nudging Robotics: Innovative Solutions to Regulate Artificial Intelligence*, 20 *VAND. J. ENT. & TECH. L.* 385, 395-96 (2017). 拙書『ロボット法』前掲注(一)・一四―二四九―五〇頁も参照。更に〈超知能〉(Artificial Superintelligence: A S I)も加えて、A G IとA S Iを次のように区別する例が興味深い——ヒトに似せた認知能力を伴った設計のプログラムがA G Iで、合理性やヒトの定義を超えた存在がA S Iである——。David T. Laton, *Manhattan Project: Exec: A Nuclear Option for the Digital Age*, 25 *CATH. U. J. L. & TECH.* 94, 94-95 (2016).
- (5) 総務省「報告書二〇一七」前掲注(2)・二七頁&脚注六六参照。
- (6) 〈A I・ロボット法〉という学術分野が必要なる理由は、将来の危険性を予測し(Predict)、これに備え(Provide)、かつ人々を防護する(Protect) 為である——すなわち「予防法務の3 P s」——という筆者の主張については、『国際情報学概論』一三三―四五頁(ミネルヴァ書房二〇二〇年)等参照。

- (7) 例えば映画「アルマゲドン」(タッチストーン・ピクチャーズ一九九八年)のように、巨大な隕石が地球に衝突して人類が絶滅する大惨事の危険性—catastrophic risks—に対しては、その蓋然性が高くなければ、巨額の税金を使った回避策を採るという判断を人類は下やなごびである。See, e.g., David A. Koplow, *Extraterrestrial Ploushares: Using a Nuclear Explosive Device for Planetary Defense against an Incoming Asteroid*, 23 UCLA J. INT'L. & FOR. AFF. 76, 78 (2019). See generally RICHARD A. POSNER, *CATASTROPHE: RISK AND RESPONSE* (Oxford Univ. Press 2004).
- (8) 例えば拙書『ロボット法』・前掲注(1)・四一—五〇頁、及び拙稿「ロボット法と学際法学：〈物語〉が伝達する不都合なメッセージ」『情報通信学会誌』三五巻四号一〇九頁(二〇一八年)等参照。
- (9) 後掲本文1、及び拙書『ロボット法』・前掲注(1)・四一—五〇頁等参照。
- (10) 『モロー博士の島』を例示しながらSFの使用を正当化する文献としては、拙稿「ロボット法と学際法学」前掲注(8)・一〇〇—一一頁 (Mitchell Travis, *Making Space: Law and Science Fiction*, 23 L. & LIT. 241 (2011)) を紹介しながら(5)。
- (11) See, e.g., *The 100 Greatest American Films*, BBC, <https://www.bbc.com/culture/article/20150720-the-100-greatest-american-films> (last visited July 30, 2020) (第四位を獲得)。
- (12) HALが宇宙飛行士殺害を判断した(動機)は、ミッションの障害を排除したという説や、生存の為の正当防衛だったという説や、指令の矛盾説が存在する。拙書『ロボット法』前掲注(1)・三七、一八四、二二一—三四頁、及び中島秀之「HALの謀反」『人工知能学会誌』一六巻一—号八二頁、八四頁(二〇〇一年一月)参照。
- (13) 「フレーム問題」については、拙書・同右二五—二五二頁参照。
- (14) 本稿が紹介する〈法と文学〉という研究分野に於いては、〈物語〉は抽象的概念を具体化させてヒトの理解度を深める機能があると指摘されている。See, e.g., Jane Baron, *Essay, Law, Literature, and the Problems of Interdisciplinarity*, 108 YALE L.J. 1059, 1064 (1999).
- (15) POSNER, *CATASTROPHE*, *supra* note 7, at 109 (訳は筆者拙訳)。
- (16) Stephen E. Henderson, *Fourth Amendment Time Machines (and What They Might Say about Police Body Cameras)*, 18 U. PA. J. CONST. L. 933, 936 (2016) (訳は筆者拙訳) 及び拙書『ロボット法』前掲注(1)・四八—四九頁も参照。
- (17) John O. McGinnis, *Accelerating AI*, NW. U. L. REV. 1253, 1256 (2010) (訳は筆者拙訳)。

- (18) 例えば、拙書『ロボット法』前掲注(1)・四二頁参照。
- (19) 同右・一六六―六七頁、二三三―三四頁参照。See also Ryan Dowell, Note & Comment, *Fundamental Protection for Non-Biological Intelligences or: How We Learn to Stop Worrying and Love Our Robot Brethren*, 19 MINN. J. L. SCI. & TECH. 305, 317-19 (2018) (脳の模倣研究が実施されつつある事実を指摘); Douglas O. Linder, *The Other Right-to-Life Debate: When Does Fourteenth Amendment "Life" End?*, 37 ARIZ. L. REV. 1183, 1205 n.134 (1995) (脳機能がソフトウェアにダウンロードされる将来の可能性を指摘)、更に国内の報告書でも、「脳全体のアーキテクチャをモデルにして研究を行うことにより、今後十数年以内に人間のような汎用AIが開発されることになる」との見通し」と指摘されている。総務省「報告書二〇一七」前掲注(2)・二六頁脚注六四。
- (20) 既に表明した筆者の主張は、「拙書『ロボット法』前掲注(1)・五一―六、四一―五〇頁、及び拙稿「ロボット法と学際法学」前掲注(8)参照。
- (21) 吉川孝、他編著『映画で考える生命環境倫理学』一八五―一八八頁(勁草書房二〇一九年)。
- (22) なお「もしくだったら」という思考実験に於いてSF作品が有用であるという分析・指摘は、興味深いことに法律学にも見受けられる。See Rokokwood, *supra* note 1, at 272 (Science Fiction views our present situation through the lenses of either a possible future, or an alternative view of the past, using such devices as "what if" counterhistories and parallel universes.) (emphasis added) と分析・指摘している。
- (23) See Julia Angwin et al., *Machine Bias: There's Software Used across the Country to Predict Future Criminals. And It's Biased against Blacks*, ProPublica (May 23, 2016), <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing> (last visited Aug. 3, 2020).
- (24) 例えば、日本リスク研究学会編『リスク学事典』一〇〇頁(丸善出版二〇一九年)、及び山口健太郎、他「安全・安心に係る人文・社会科学分野における学術研究フロンティアの抽出」『社会技術研究論文集』第五卷一七二頁、一九一頁(二〇〇八年三月)参照。
- (25) 山本龍彦「AI時代のプライバシーとデータ保護——プロファイリングを中心に——」二一三頁(二〇一九年二月二七日) [in](https://総務省「AIネットワーク社会推進会議AIガバナンス検討会(第五回)資料1(平成三十二年二月二七日)https://) 総務省「AIネットワーク社会推進会議AIガバナンス検討会(第五回)資料1(平成三十二年二月二七日) <https://>

www.sounu.go.jp/main_sosiki/kenkyu/ai_network/02icp01_04000161_00002.html (最終アクセス二〇二〇年七月二六日)。

- (26) 例えば、『拙書』『アメリカ不法行為法―主要概念と学際法理―』四頁、五九―六二、九二頁(中央大学出版部二〇〇六年)等参照。

(27) See Nancy Levit, *Reshaping the Narrative Debate*, 34 SEATTLE U. L. REV. 751, 756–57 (2011). 更に法学教員は、教室に於ける仮想事例を考案したり、試験問題と二つの仮想事例を考案する二つを通じて、皆、物語を作っているという興味深い指摘もある。Nancy L. Cook, *Outside the Tradition: Literature as Legal Scholarship: The Call to Stories: Speaking in and about Stories*, 63 U. CIN. L. REV. 95, 98 n.14 (1994).

(28) David S. Caudill, *Idealized Images of Science in Law: The Expert Witness in Trial Movies*, 82 ST. JOHN'S L. REV. 921, 923 (2008).

(29) See, e.g., Anita L. Allen, *The Jurisprudence of Jane Eyre*, 15 HARV. WOM. L.J. 173, 178 n.17 (1992). (現在の法と文学運動は、ウイングモアやロスロウ・ハンントやメンジャーミン・カーブーン等の著名法律家が法と文学の関係性を探求した潮流の再発見である)と指摘)。

(30) See, e.g., David S. Clark, *American Participation in the Development of the International Academy of Comparative Law and Its First Two Hague Congresses*, 54 AM. J. COMP. L. 1, 10 (2006). なおウイングモアは更に、法律家が文学に親しむ重要性を説くことにより有名である。John H. Wigmore, *A List of Legal Novels*, 2 ILL. L. REV. 574 (1908).

(31) リチャード・A・ボスナー著、平野晋監訳『法と文学(原著第三版)』(木鐸社二〇一一年)参照。

(32) See, e.g., Michael Burger, *Environmental Law / Environmental Literature*, 40 ECOLOGY L.Q. 1, 13–14 (2013); Baron, *supra* note 14, at 1065–66. 和文資料としては小林史明『法と文学』研究の歴史と展開(博士学位請求論文要旨)二二頁 https://m-repo.lib.meiji.ac.jp/dspace/bitstream/10291/19657/2/759_sumpdf (最終アクセス二〇二〇年七月二八日)等参照。

(33) See, e.g., Jennifer Bard et al., *Three Ways of Looking at a Health Law and Literature Class*, 1 DREXEL L. REV. 512, 526–27, 528–29 (2009).

(34) Jennifer Reynolds, *Games, Dystopia, and ADR*, 27 OHIO ST. J. ON DISP. RESOL. 477, 483 n.11 (2012).

(35) See, e.g., David S. Caudill, *Idealized Images of Science in Law*, *supra* note 28, at 921–22; Lolita K. Buckner Inniss, *Bicentennial*

- Man - The New Millennium Assimilationism and the Foreigner among Us*, 54 LUTGERS L. REV. 1101, 1101 (2002).
- (36) Rockwood, *supra* note 1.
- (37) Kim Lane Scheppele, *Legal Story Telling: Foreword: Telling Stories*, 97 MICH. L. REV. 2073, 2075 (1989) (‘‘Stories represent experience, and can introduce imagination and new points of view’’) (訳は筆者拙訳).
- (38) *Id.*
- (39) *See, e.g., Majority of Minneapolis City Council Pledges to Dismantle Police Department*, The New York Times, June 17, 2020.
- (40) なお古い映画ファンであれば、ユル・プリンナー主演映画「ウエストワールド」(MGM一九七三年)が、誤作動したロボットの恐怖という文脈で思い出されよう。あるいは本文前掲「2001年宇宙の旅」の人工知能HAL 9000型が宇宙飛行士を殺害する物語も、同様であろう。
- (41) *See, e.g., Levit, supra* note 27, at 758-59; Bret Rappaport, *Tapping the Human Adaptive Origins of Storytelling by Requiring Legal Writing Students to Read a Novel in Order to Appreciate How Character, Setting, Plot, Theme, and Tone (CSPTT) Are as Important as IRAC*, 25 T.M. COOLEY L. REV. 267, 298 (2008).
- (42) 物語や物語の物語を用いる手法については、本文で指摘した〈偏見〉という欠点以外にも、以下のような批判も存在する。例えば個人的経験に基づく物語の物語を用いる場合には客観的証拠によつて証明できないとか、個人的経験は世界全体の経験を代表する訳ではないとか、話者の一方的かつ感情的な話であつて分析的ではないとか、不合理的である等の批判も見受けられる。*See, e.g., Levit, supra* note 27, at 754. 拙稿「ロボット法と学際法学」前掲注(8)・111-112頁も参照(ボズナーによるナラトロシスト批判を紹介してゐる)。
- (43) 拙稿「ロボット法と学際法学」前掲注(8)・112頁。See FRE 403 (‘‘The court may exclude relevant evidence if its probative value is substantially outweighed by a danger of one or more of the following: unfair prejudice, confusing the issues, misleading the jury, undue delay, wasting time, or needlessly presenting cumulative evidence.’’)と規定してゐる)。
- (44) 拙書『ロボット法』前掲注(一)・七六-七七頁。
- (45) Guñot et al., *supra* note 4, at 434.

- (46) 北野宏明〔経団連〕AI活用原則TF主査「Mission Statement 経団連AI活用原則TF」一一二頁 in 内閣府「人間中心のAI社会原則検討会議 第一回」(平成三〇年五月八日) 参考資料一 <https://www8.cao.go.jp/cstp/tyousakai/humanai/ikai/sankol.pdf> (最終アクセス二〇二〇年八月三日)。
- (47) See, e.g. OECD, *Intelligent Machines, Smart Policies: Interview*, <https://www.facebook.com/watch/?v=10154858561222461> (last visited July 25, 2020) (OECDで収録され公開されている事務次長他と筆者とのインタビュー動画の冒頭部分に於いて、本文引用識者達が危険性を指摘している事実を事務次長が紹介している)。See also Schere, *supra* note 3, at 356 (ICT系著名人も規制の必要性を主張していると指摘); Laton, *supra* note 4, at 95-96 (ビル・ゲイツ等のICT系著名人が人類にとってのAI脅威論に同調と指摘); Dowell, *supra* note 19, at 309-11 (同上)。
- (48) 株式会社 Preferred Network 「人間中心のAI社会原則検討会議」に対する意見」一頁 in 内閣府「人間中心のAI社会原則検討会議 第一回」(平成三〇年五月八日) 参考資料二 <https://www8.cao.go.jp/cstp/tyousakai/humanai/ikai/sankol2.pdf> (最終アクセス二〇二〇年八月三日)。
- (49) 総務省「AIネットワーク化検討会議」https://www.soumu.go.jp/main_sosiki/kenkyu/ict/index.html (最終アクセス二〇二〇年八月三日)、及び同省「AIネットワーク社会推進会議」https://www.soumu.go.jp/main_sosiki/kenkyu/ai-network/index.html (最終アクセス二〇二〇年八月三日)。
- (50) 総務省「経済協力開発機構(OECD)における人工知能(AI)に関する取り組み」https://www.soumu.go.jp/menu_seisaku/itseisaku/oced_ai/index.html (最終アクセス二〇二〇年七月二五日)、及び同省「報告書二〇一九」前掲注(2)・二、四、一四、一八、二二、二四、五八頁。
- (51) 総務省「AIネットワーク社会推進会議 開発原則分科会・影響評価分科会合同分科会 議事概要」三頁(平成二九年七月二〇日) https://www.soumu.go.jp/main_content/000513655.pdf (最終アクセス二〇二〇年七月二五日)(企業名・個人名は伏字加工付加)。
- (52) 汎用AIがソフトウェアの対象になった結果、理由、及び根拠については、総務省「報告書二〇一七」前掲注(2)・二六一二八頁及び脚注(六三)~(六七) https://www.soumu.go.jp/main_content/000499624.pdf (最終アクセス二〇二〇年八月三日) 参照。

- (53) 総務省「AIネットワーク社会推進会議 第六回 議事概要」(平成二九年七月二五日) https://www.soumu.go.jp/main_content/000513654.pdf (最終アクセス二〇二〇年八月三日)。
- (54) 二〇一七年七月二〇日に開催された合同分科会に於いて、消費者団体の代表である構成員からは、次のような発言があった。

【●構成員「消費者団体代表」】

- ・ ガイドライン案は、とても立派で利用者からみても納得できる常識的なものである。書いてある内容は当たり前のごとく、開発していただくときに当然守っていただきたい原則である。
- ・ このような内容のガイドラインに賛成できないという開発者の姿勢は、利用者に不安を与えることになる。

(強調付加)

- 総務省「合同分科会 議事概要」前掲注(51)・四頁。
- (55) 「IEEE」は、米国電気電子学会の略称。
- (56) Brian S. Haney, *The Perils and Promises of Artificial General Intelligence*, 45 J. LEGISLATION 151, 169 (2019) (MAX TEGMARK, LIFE 30 BEING HUMAN IN THE AGE OF ARTIFICIAL INTELLIGENCE 156 (2017)) 等を引用しつつ(訳は筆者拙訳)。なお総務省「報告書二〇一七」前掲注(2)・二六―二七頁及び脚注(六三)～(六七)も参照(海外の例を紹介)。
- (57) 本文当段落の出典は、拙稿「AIの支配」と「法の支配」(日本発AI原則とガイドライン)、『法の支配』一九七号四一、四二―四三頁(日本法律家協会、二〇二〇年四月)参照。
- (58) 例えば当原稿執筆時の二〇二〇年八月四日に「Lexis Advance」データベース内の「Law Reviews & Journals」の中で「(Isaac Asimov) AND (Three Laws of Robotics)」のキーワード検索をしたところ、五二件も検出された。
- (59) David Caudill, *Law-and-Literature, Literature-and-Science, and Enhancing the Discourse of Law/Science Relations*, 27 J. Legal Prof. 1, 5 (2002/2003) (訳は筆者拙訳)。
- (60) Baron, *supra* note 14, at 1064.
- (61) See OECD, *Recommendation of the Council on Artificial Intelligence*, OECD/LEGAL/0449 (adopted on May 22, 2019),

https://www.soumu.go.jp/main_content/000642218.pdf (last visited Aug. 3, 2020).

(62) 本文中の記述については、see Caudill, *Idetified Images of Science in Law*, *supra* note 28, at 922.

(本学国際情報学部教授&学部長・米国弁護士(NY州))