

2017.11.09 AIネットワーク社会推進会議 環境整備分科会・影響評価分科会 合同研究会

メディア論ならびに情報倫理の視座からの論点検討

河島茂生(青山学院女子短期大学, 理化学研究所)

社会のなかでの技術の形成(1)

- ▶ 技術はいくつもの可能的様態をはらんでおり、歴史社会的な文脈のなかではじめてメディアのありようが決まり、その影響力が定まっていく。メディアは「国家や資本の編制力から、市民、あるいは大衆の想像力にいたる、複合的で重層的な社会の諸力の錯綜した結果として、今日のような姿に固定化されてき」(水越, 1996:186-187)ている。

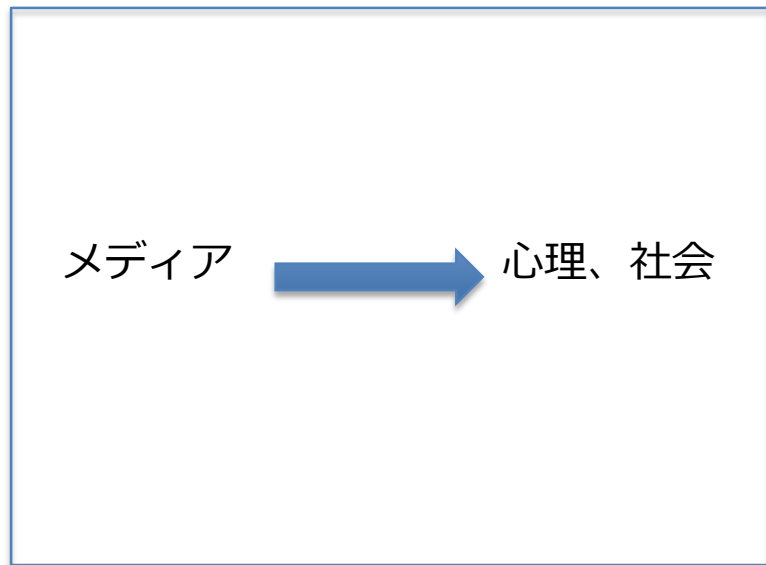


図 技術決定論

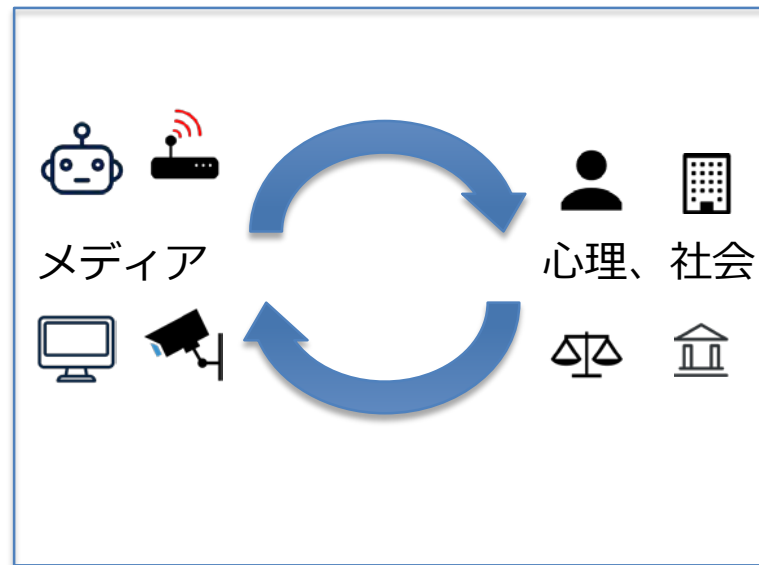
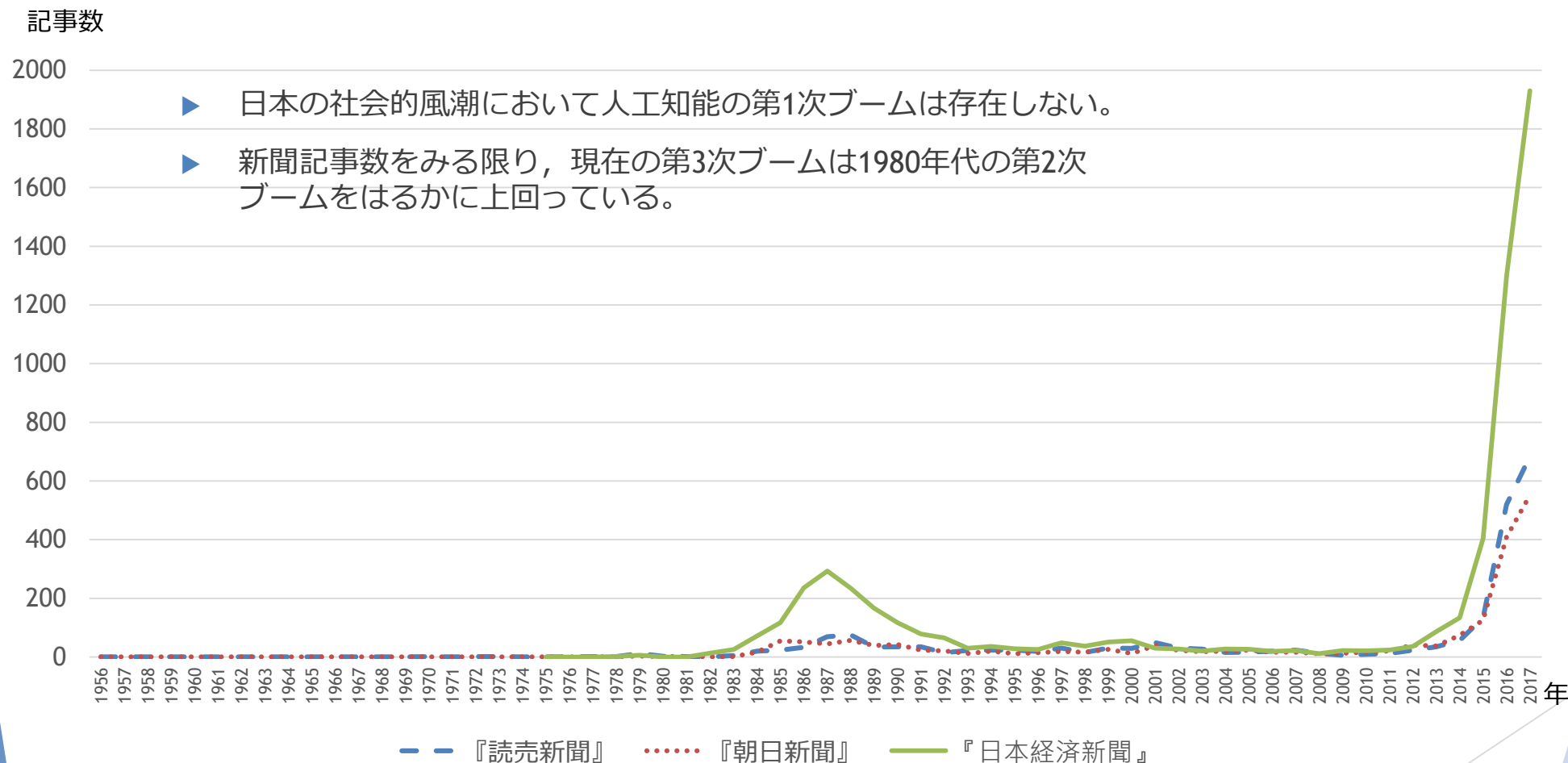
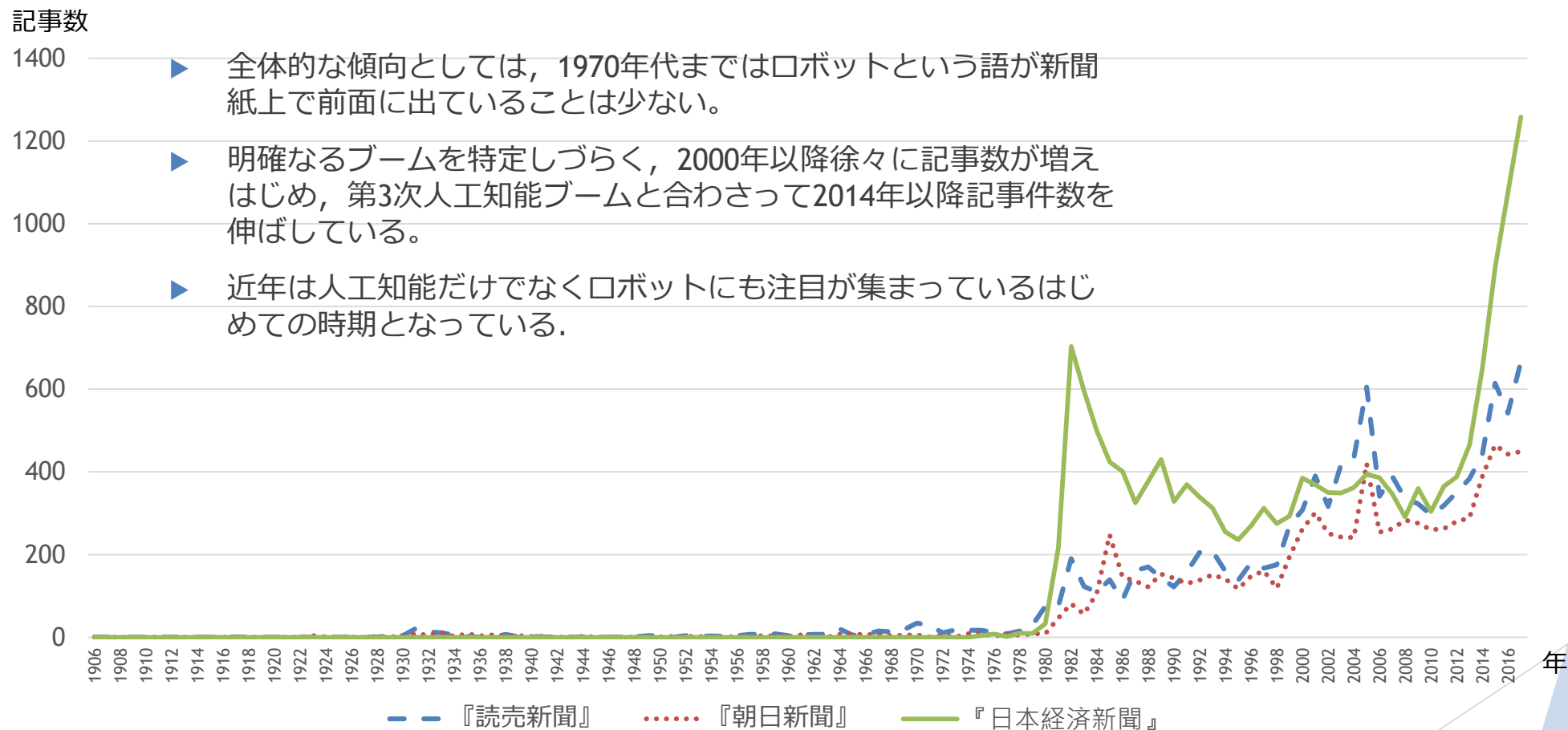


図 ソシオ・メディア論

「人工知能」の新聞記事数の変化

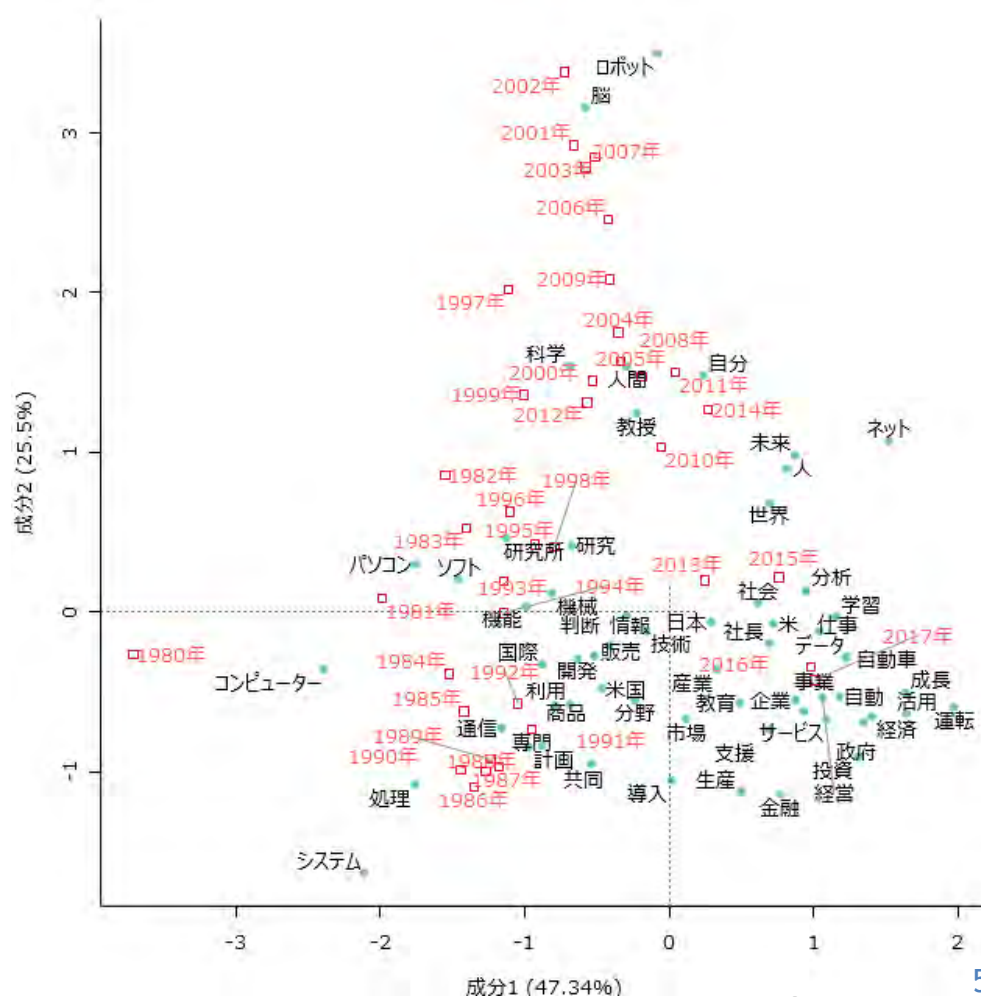


「ロボット」の新聞記事数の変化



「人工知能」の新聞記事の対応分析

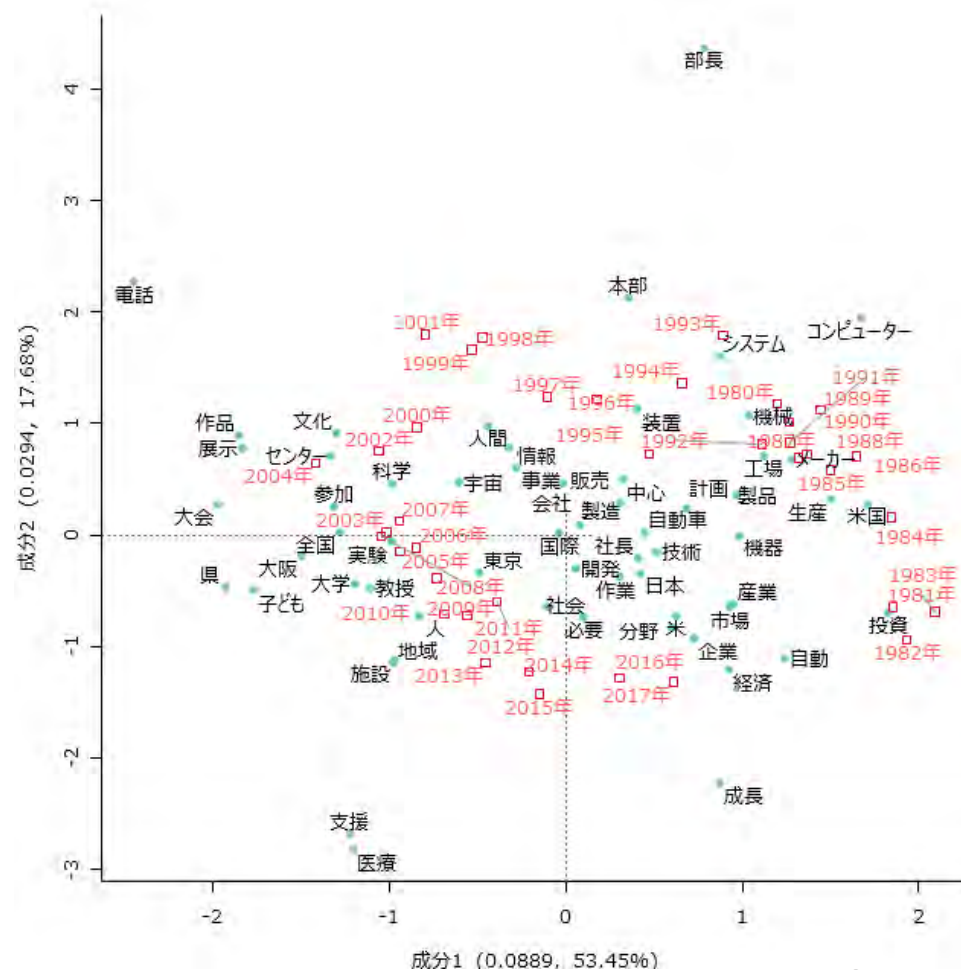
- ▶ 1980年代は、専門知識を入れ込んだエキスパート・システムが流行ったため、「専門」がこの時期を特徴づける語に挙がっている。また、第五世代コンピュータの「計画」がしばしば伝えられたため、その語も当時の動向を反映する語として算出されている。
- ▶ 第3次ブームの2015年頃は、人工知能の社会的影響が頻繁に取り上げられている。つまり、「経済」「経営」「企業」「投資」「成長」などが特徴語に挙がっている。また、ビッグデータを学習データとして用い解析していくため「データ」という語が頻出し、自動車をめぐっては自動運転が話題の中心にあるため「自動」「運転」が頻出している傾向が明確に現れている。



(河島, 2017a)

「ロボット」の新聞記事の対応分析

- ▶ 1980年代の前半に「投資」が特徴的な語として抽出されており、1980年代半ばから「機材」「上場」「メーカー」「製品」「生産」など産業用ロボットに関連した語が目立つ。
- ▶ 近年は、「市場」「企業」「経済」といった語が頻繁に使われており、工場内のロボットにかぎらず広く経済活動全般での活用が話題に挙がっていることがみえてくる。



第3次ブームとそれ以前の言説の共通性

- ▶ 2013年以降の人工知能・ロボットへ向けられる言説は、過去と比べてそれほど特別なものではない(河島, 2017a)。

- ▶ 介護現場での活用

- ▶ 「看護ロボット登場 入浴から, お茶まで」『読売新聞』(1982.9.21)
- ▶ 「[からだと機械の握手] 進む医用工学 介助ロボット」『読売新聞』(1982.12.6)

- ▶ 技術的失業

- ▶ 「[科学]の未来は? ロボットに職を奪われた数百万の失業群」『読売新聞』(1982.11.27)
- ▶ 「「ロボットで失業」の脅威 日本生産性本部“ME革命白書”」『読売新聞』(1982.4.6)
- ▶ 「'84 世界の情景 ニューライフ 「打ち壊し運動」の再来」『読売新聞』(1984.1.23)

- ▶ ロボット税

- ▶ 「公明がロボット税提案」『読売新聞』(1982.6.5)

第3次ブームとそれ以前の言説の共通性

- ▶ 2013年以降の人工知能・ロボットへ向けられる言説は、過去と比べてそれほど特別なものではない(河島, 2017a)。

▶ 金融

- ▶ 「第2部AI特集 AI時代本番、産業界を変える勢い、市場規模「数十兆円」の声も。」『日本経済新聞』(1988.2.24)

▶ 自動運転車

- ▶ 「文字通りの“自動車” ロボット運転で事故なくす」『読売新聞』(1958.12.18)
- ▶ 「動く映像 自在処理」『読売新聞』(1986.1.10)

▶ 入学試験

- ▶ 「人工知能「考える葦」に機械は迫れるか」『朝日新聞』(1985.1.1)

▶ パワード・スーツ及びダイレクト・ブレイン・インターフェース

- ▶ 「人間の脳波で動く」『読売新聞』(1958.3.10)

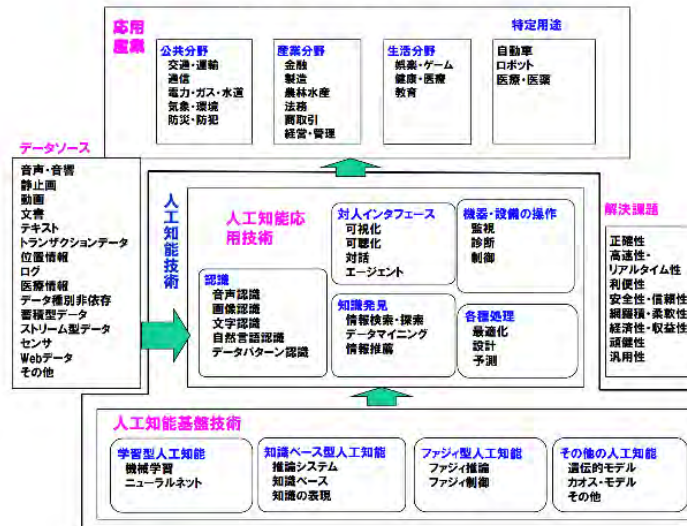
▶ 人工知能の創造性

- ▶ 「バッハ風賛美歌電算機で作曲、IBM研究者がシステム開発」『日本経済新聞』(1988.8.19)

AIネットワーク社会の将来像

- ▶ 社会に組み込まれた技術のありかたや影響力は、きわめて多数のファクターが絡む。技術者や企業の活動，国家や地方自治体の政策，一般の人々の意思や利用形態，予期せぬ事故等がグローバルな規模で連動していく。AIネットワーク社会の将来像の描出は，困難を極め複雑多岐にならざるをえない。
- ▶ しかし，AIネットワーク社会の将来像を描かなければ，前もって社会制度上の準備ができず，種々の予想を行うことは必要不可欠である。「困難な作業だけでも，外れることも十二分にあることを踏まえて将来像を描く」という姿勢が求められる。
- ▶ 複数のシナリオを描き，技術的・社会的な大きな分岐点を挙げる。どのような課題が克服されれば社会が大きく変わり，どのような課題が克服されなければ停滞を招くのかを描くことが望ましい。

図 1-1 人工知能技術の俯瞰図



利活用原則・ガイドラインに向けて(1)

▶ 人間の領分

- ▶ 利活用の場面においても、AIシステムが自己制作(人間の手を離れてハードウェアを含めてAIシステム自身が独自にハードウェアやソフトウェアの自己開発を行うこと)しないかぎり、人間が作り使うものであるという前提のもとで議論すべきである。AIシステムは、自然ではなく、たとえ問題が生じたとしても、AIシステム自体の瑕疵とすることはできない(河島, 2016)。

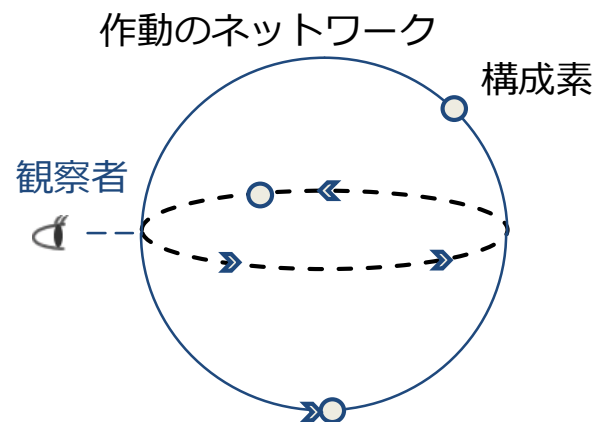


図 オートポイエティック・システム

生物や社会は、構成素を産出しながら自己制作している。たとえば生物の場合、神経系の発火パターンが構成素であり、社会の場合はコミュニケーションが構成素である。

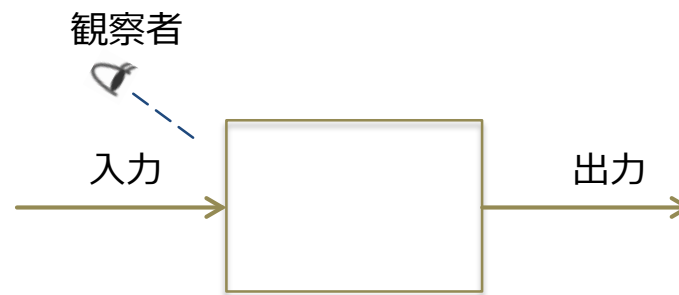
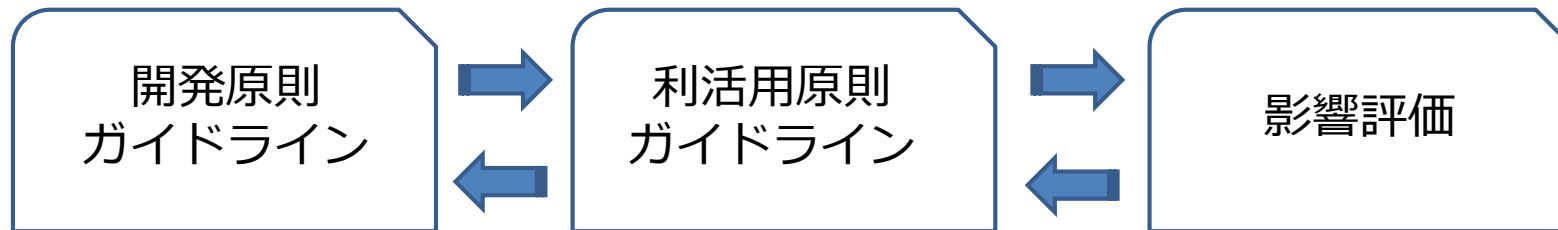


図 アロポイエティック・システム

外部(人間)によって作られた人工物であり、人間が指示した通りに動くことが求められる。AIシステムは、アロポイエティック・システムである。ただし、人間の介入が少なくなるように開発が進められている。

開発原則・ガイドライン案／影響評価との関係性

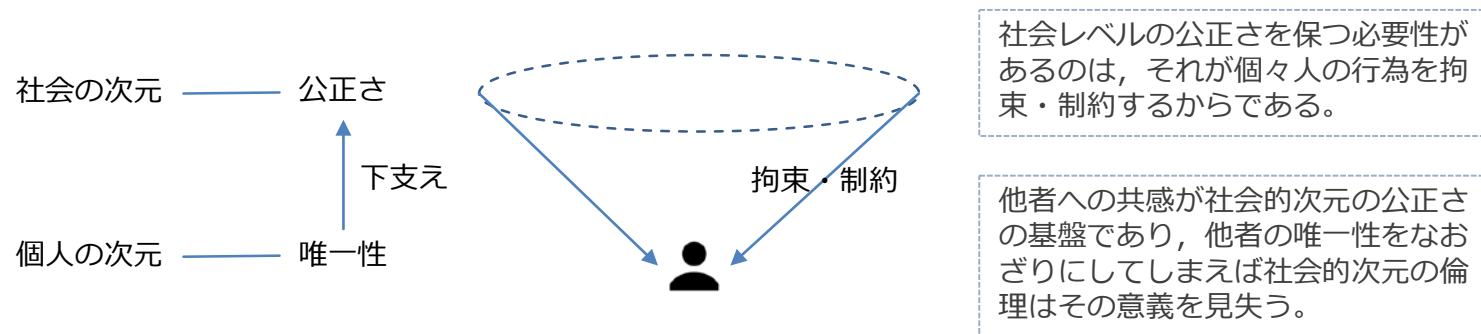
- ▶ 利活用原則・ガイドラインは，開発者と利用者が入れ替わることも多いため，あるいは両者が一緒になってシステムを開発していくこともあるため，先行的に議論した開発原則・ガイドライン案を多分に意識して策定していくことが望ましい。
- ▶ また，AIシステムの影響評価との連動も図ることが期待される。影響を予想しなければ，適切な原則・ガイドラインを作成することが難しいからである。



- ▶ AIシステムは，ネットワーク化されグローバルな規模で連携が起き，社会的な影響を生む。そのため，国際的協調が欠かせない。

倫理の原則

- ▶ 開発原則と同様，倫理の原則は欠かせない。その際，人間の尊厳と社会全体のなかでの公正さが同時に保たれるように視野に収めるべきである。
- ▶ 人間は，機能面だけでみては抜け落ちてしまう唯一性(かけがえのなさ)をもった存在であり，利活用の場面でもそのことは念頭におかれなくてはならない。
- ▶ ただし，ある人がかけがえのない存在であるからといって，その人だけの特権的に扱ってしまえば，対象とする社会全体での公平性が保たれず平等性に欠ける。
- ▶ 人間の尊厳と社会全体のなかでの公正さの両面を見据えてAIシステムは利活用されていくことが強く望まれる(河島，2017b)。



ビッグデータ型AIシステムの特徴

- ▶ ビッグデータ型AIシステムについては、ビッグデータであったとしてもデータの偏向性がともなうことや決定論的推論ではなく確率論的計算である面を踏まえて利活用されることが求められる。したがって、AIシステムがこれまでにない分析結果を提示することを理解しつつも、それが必ず「正しい」決定であると受け止めることなく、人間とは違った情報処理の一結果であるとみなすことが望ましい。
- ▶ アメリカの連邦取引委員会は、低収入や十分な社会サービスを受けられていない人々の排除につながらないように、次の項目を考慮するように推奨した (Federal Trade Commission, 2016)。これは、ビッグデータ型AIシステムを利活用する際についても留意が求められる項目である。
 - ▶ 「データの代表性」
 - ▶ 「モデルに入り込んでいる偏見」
 - ▶ 「予測の正確性」
 - ▶ 「倫理的な問題の誘発」

特に集中的に議論することが望まれる分野

- ▶ AIシステム全般にわたる利活用の場面を想定するとともに、倫理的な面から特に集中的に議論することが望まれる分野
 - ▶ 人生を左右しかねない意思決定(採用・人事評価・与信等)を支援するAIシステム



- ▶ ある社会組織・業界からの排除が連鎖的に別の社会組織・業界への排除につながるように配慮を要する。たとえば、経済的状況での困窮が教育・医療の領域まで及んでしまうことは避けられる必要がある。AIシステムの利活用によって個人が適切に社会に包摂 (inclusion) され尊厳が確保されることが求められる。
 - ▶ 直接身体に被害が加えられかねない機械(自動運転車等)のAIシステム

政府の関与(1)

- ▶ 直接身体に被害が加えられかねない機械(自動運転車等)のAIシステムの例として、自動運転車のトロッコ問題等の意識について調査した(河島・北村・柴内, 2017)。
- ▶ 完全に市場に任せることが望まれているわけではない。

▶ (問)事故が避けられない状況に直面したとき、自動運転車がとる進路を決める共通のルールを作るべきだと思いますか。そのルールは利用者個人が変更できないものとします。1～4のうち、あなたのお考えに近いもの1つに○をつけてください。

N = 468

共通ルールを作るべきと考えている人は
64.7%

共通ルール支持の選択肢

1	2	3	4
必ず作るべき	作るべき	作るべきではない	決して作るべきではない
17.7%	47.0%	27.4%	5.6%

統計的有意差($\chi^2=48.580$, $df=1$, $p=.000$)

政府の関与(2)

- ▶ 直接身体に被害が加えられかねない機械(自動運転車等)のAIシステムの例として、自動運転車のトロッコ問題等の意識について調査した(河島・北村・柴内, 2017)。
- ▶ 完全に市場に任せることが望まれているわけではない。

【前の問で「1 必ず作るべきである」「2 作るべきである」とお答えの方にはうかがいます】

- ▶ 問 どのようなかたちでのルールづくりが望ましいと思いますか。次の1～5の中からあてはまるものにいくつでも○をつけてください。(あてはまるものすべてに○)

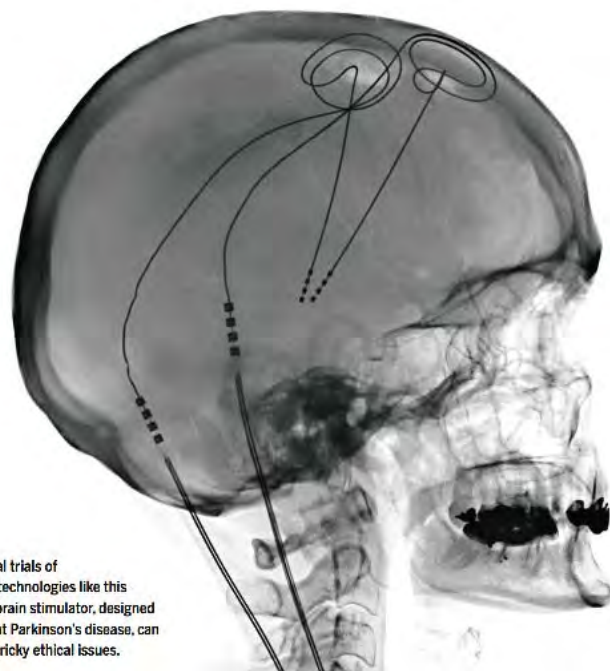
N = 314

1	国際的な条約による世界共通のルール	43.3%
2	各国政府による法的な国内共通ルール	53.8%
3	自動車業界全体が自主的に設定する共通ルール	27.7%
4	各自動車メーカーが自主的に設定するメーカー内共通ルール	7.6%
5	1～4に望ましいと思うものはない	1.6%

国家レベル以上の統一されたルールを望んでいる

長期にわたる利活用の留意点

- ▶ AIネットワーク化が「AIネットワーク化検討会議 報告書2016」「AIネットワーク社会推進会議 報告書2017」のいう4段階中の第3段階に達した際、AIシステムを身体内部に着脱が難しいかたちで組み込む場合は、長期間の利活用になることを踏まえて判断すべきである。すでに脳深部療法を長期にわたって継続的に受けることの困難さが指摘されている(Underwood, 2015)。



Clinical trials of neurotechnologies like this deep brain stimulator, designed to treat Parkinson's disease, can raise tricky ethical issues.

アカウントビリティの原則(1)

- ▶ 開発原則と同様，アカウントビリティの原則も求められる。
- ▶ 利活用する側も，どのようなシステムであるか，どのようなデータに基づいて計算されているかを抽象的にでも把握しておく必要がある。そうでなければ，AIシステムの解析結果を客観的に捉えられず，またAIシステムの解析結果によって深刻な影響を受ける人たちの理解が得られないことが想定される。結果的にAIシステムの社会的受容性が高まらない。
- ▶ 『情報通信白書』(平成28年版，pp.246-247)には，AIシステムが上司となることに日本は抵抗感を抱く人が多いことが述べられており，AIシステムの判断結果だけを見て人事評価をされると様々な軋轢が生じることが予想される。加えて，なにか問題が生じたときにまったく対応できず，開発・メンテナンス業者の言うことを批判的に捉えられない。
- ▶ アカウントビリティは透明性と関係している。

アカウントビリティの原則(2)

- ▶ 直接身体に被害が加えられかねない機械(自動運転車等)のAIシステムの例として、自動運転車のトロッコ問題等の意識について調査した(河島・北村・柴内, 2017)。
- ▶ アカウントビリティを保つかぎりの透明性が求められている。

- ▶ (問)事故が避けられない状況に直面したときに進路を決める自動運転車内部のメカニズムは公表されるべきだと思いますか。1～4のうち、あなたのお考えに近いもの1つに○をつけてください。

N = 468

公表支持の選択肢

1	2	3	4
必ず公表されるべき	公表されるべき	公表されるべきではない	決して公表されるべきではない
39.7%	51.1%	6.4%	0.9%

統計的有意差($\chi^2=333.074$, $df=1$, $p=.000$)

公表支持の選択をした人は90.4%に上る

付言

▶ 公平性に関する付言

- ▶ ビッグデータには必ず偏向性が生じるため、完全なる公正性は半ば幻想であるといつてよい。
- ▶ T.Tatmanは、Youtubeの自動生成の字幕の精度に男女差や方言による差があることを指摘した(Tatman, 2017)。学習データは、あらゆる属性に均等に存在しているわけではなく、こうしたことは頻繁に起こり得る。
- ▶ あまりにも過度に公平性に拘泥するとビッグデータをもとにしたAIシステムのサービスを展開できなくなってしまう。どれだけ人生に深刻なダメージを与えるかを考え、必要な場合にAIシステムの非公正性を批判していくことが開発の萎縮を招かないために条件づけられる。

▶ 透明性に関する付言

- ▶ アカウンタビリティを保つかぎりの透明性が求められている。多数のAIシステムがネットワーク化されることを考慮すれば、完全なる透明性もまた半ば幻想である。
- ▶ 開発者の経済的活動や技術開発を阻害しないように、また「アルゴリズム嫌悪」(algorithm aversion)が起きないように留意していくことが求められる。
- ▶ 人間の身体に被害が及ぶ事故があった場合には、たとえ困難であっても監査証跡のような仕組みを導入して後で詳細な検証が可能なAIシステムを構築せざるをえない。

参考文献(1)

- ▶ Federal Trade Commission (2016) “Big Data”
<https://www.ftc.gov/system/files/documents/reports/big-data-tool-inclusion-or-exclusion-understanding-issues/160106big-data-rpt.pdf> アクセス日: 2017/10/10
- ▶ 河島茂生(2016)「ネオ・サイバネティクスの理論に依拠した人工知能の倫理的問題の基礎づけ」『社会情報学』 5(2) 53-69
- ▶ 河島茂生(2017a)「新聞記事に見る人工知能やロボットの言説の変化」『人工知能』 32(6) pp.935-942
- ▶ 河島茂生(2017b)「人工知能倫理に関わる社会的次元／個人的次元の峻別と交差」 2017年度人工知能学会全国大会(第31回) 2017年5月23日
- ▶ 河島茂生, 北村智, 柴内康文(2017)「自動運転車の「トロッコ問題」などに関する意識」 2017年社会情報学会(SSI) 学会大会 2017年9月17日
- ▶ 水越伸(1996)「情報化とメディアの可能的様態の行方」『メディアと情報化の社会学』岩波書店, pp.177-196

参考文献(2)

- ▶ 特許庁(2015)「平成26年度 特許出願技術動向調査報告書 (概要) 人工知能技術」
- ▶ 総務省『情報通信白書』(平成28年版)
- ▶ Tatman, Rachael(2017)“Gender and Dialect Bias in YouTube’s Automatic Captions” *Proceedings of the First Workshop on Ethics in Natural Language Processing*, pp.53-59.
- ▶ Underwood, Emily(2015)“Brain implant trials raise ethical concerns” *Science* Vol. 348, Issue 6240, pp. 1186-1187 DOI: 10.1126/science.348.6240.1186

注

- ▶ 図のアイコンは下記のものを使っている。
 - ▶ スライド2
 - ▶ メディア 左上 By Omara Abderraouf (CC BY 3.0)
 - ▶ メディア 右上 By DinosoftLabs (CC BY-NC 3.0)
 - ▶ メディア 左下 By Omara Abderraouf(CC BY 3.0)
 - ▶ メディア 右下 By DinosoftLabs(CC BY 3.0)
 - ▶ 心理, 社会 左上 By Ivan Boyko(CC BY 3.0)
 - ▶ 心理, 社会 右上 By By Dave GandySIL(Open Font License)
 - ▶ 心理, 社会 左下 By Atlas(CC BY 3.0)
 - ▶ 心理, 社会 右下 By Eva Kufulium(CC BY 3.0)
 - ▶ スライド12
 - ▶ By Ivan Boyko (CC BY 3.0)
 - ▶ スライド14
 - ▶ 右 by PICOL(CC BY-SA 3.0)
 - ▶ 左 by Designmodo(CC BY 3.0)
 - ▶ 中央 by Visual Pharm(CC BY-ND 3.0)