Over the AI ―― AIの向こう側に（13）：

弱いままの人工知能 ～ “強いAI”を生み出すには「死の恐怖」が必要だ

http://eetimes.jp/ee/articles/1707/27/news015.html

AI（人工知能）には、「人間のアシストをする“弱いAI”」と「知性を持つ“強いAI”」があるという考え方があります。私は、現在の「AI」と呼ばれているものは、全て“弱いAI”と思っています。では、私たちは“強いAI”を生み出すことができるのでしょうか。それを考えるには、人間にとって、恐らくDNAレベルで刻まれているであろう普遍的な感覚、「死への恐怖」がヒントになりそうです。

2017年07月27日 11時30分 更新

[江端智一, EE Times Japan]

今、ちまたをにぎわせているAI（人工知能）。しかしAIは、特に新しい話題ではなく、何十年も前から隆盛と衰退を繰り返してきたテーマなのです。にもかかわらず、その実態は曖昧なまま……。本連載では、AIの栄枯盛衰を見てきた著者が、AIについてたっぷりと検証していきます。果たして”AIの彼方（かなた）”には、中堅主任研究員が夢見るような”知能”があるのでしょうか――。⇒連載バックナンバー

「あの世」や「来世」を信じてる？

　ある日のこと、

「パパは、『あの世』とか『来世』とか信じている？」

と、中学3年生の次女が質問してきました。

　江端家では、それが、たとえ、次女の定期試験の前夜であっても、『バカなこと聞いていないで、勉強でもしていなさい！』と叱責しない家風です。この手の議論を振られれば、その問いに対して全力で答える父親と、その議論が長びくと、無言の圧力でそれを停止させる母親とで、この家風を運営しています。

江端：「『あの世』とか『来世』については、死んだ後に分かるのであれば、別段、今、急いで考えなくても良いかなと思っている」

次女:「『パパが、今、どう思っているか』を聞いているんだけど」

江端:「『今、私が、どう思っているか』というのであれば、『あの世』とか『来世』は、存在しないと考えるのが『合理的』と考えている」

次女:「なんで?」

江端:「『あの世』とか『来世』を主張している、宗教の教義が、ものすごく『理不尽』だからだよ。基本的には、「死後の処理フロー」が、もうめちゃくちゃ理不尽」

次女:「というと?」

江端:「まず、納得できないのは、別の宗教を信じている人を、問答無用で"地獄"に送り込むという考え方だな」

次女:「どういうものなの?」

江端:「『ある所に、敬けんなキリスト教の信者のおばあさんがいました。その人は、聖書に記載されている通りに、隣人に親切にして、貧しい人に施しをして、清貧な人生を送り、その人生を終えました』」

次女:「で?」

江端:「『ところが、こちらに来てみてビックリ。あの世は、イスラム教の世界で運営されていたのです。ですから、このお婆さん、裁判(最後の審判)の即決裁判で地獄行きが決定しました』」

次女:「はあ」

江端:「"キリスト教"と"イスラム教"を入れかえても同じことになるんだけど、これって理不尽の極(きわみ)じゃないか?」

次女:「そうなの?」

江端:「『あの世』やら『来世』なるものが同時に2つ以上存在する、などと言っている宗教は1つもないのだから、そういう理不尽で不合理な裁判システム(最後の審判)を採用しているシステムは『併存できない』と考える方が自然だろう?」

次女:「確かに。じゃあとにかく、パパは『あの世』とか『来世』を信じていない、ってことだね」

江端:「お前はどうなんだ?」

次女:「そうだなぁ。私は信じているかな」

江端:「そうか。じゃあ、パパの見解(『あの世』とか『来世』はない)を引っくり返してみようとは

思わないか？」

次女：「は？　何？」

江端：「もしお前が、"パソコンにも『あの世』とか『来世』がある"という仮説を ーー 検証しなくてもいいから ーー 論理的に説明できたら、パパは、自分の見解を引っ込めて、お前の主張を全面的に受け入れてもいいよ」

「AIの知性」について考える

　こんにちは、江端智一です。

　今回は番外編として、人工知能（AI）技術の話ではなく、AIの知性について、いろいろな人の考え方を紹介し、思考実験をしてみようと思っています。

　さて、これまでの1年間の連載で、私は、「現在のAIブームにおいて、多くの人が描いているAIのほとんどは幻想である」と、何度も繰り返してきました。しかし、どうも私の主張は全く世の中に届いていない気がするのです。

　今回、あらためて、じっくりこの理由を考えてみたのですが、結構その理由は簡単でした。それは、

　ーー 主張している当事者が、私（江端）だから。

　私は、AI技術の著名な研究者でもなく、大手ITベンチャーの社長でもなく、AI技術に関して論文も著書も発表していない、一介のサラリーマンエンジニアにすぎません。

　私のやってきたことといえば、世の中にある「AI技術」を、チャッチャとコーディングし、追試し、検証することだけであり、それをやり続けてきた結果、「現状、世間で騒がれているAIは、幻想である」という自分なりの結論を、皆さんに説明しているだけです。

　つまり、私には、権威がないのです（威厳もないですが）。

　それならば、今回は、思い切って「権威のある人たちの意見」に乗っかってみよう、と考えました。それはもう、悪意に満ちるほどに、徹底的に権威に媚び（こび）へつらい、隷属し、忖度（そんたく）し、その上で、私の持論に引き込むという手法を用いてみたいと思います。

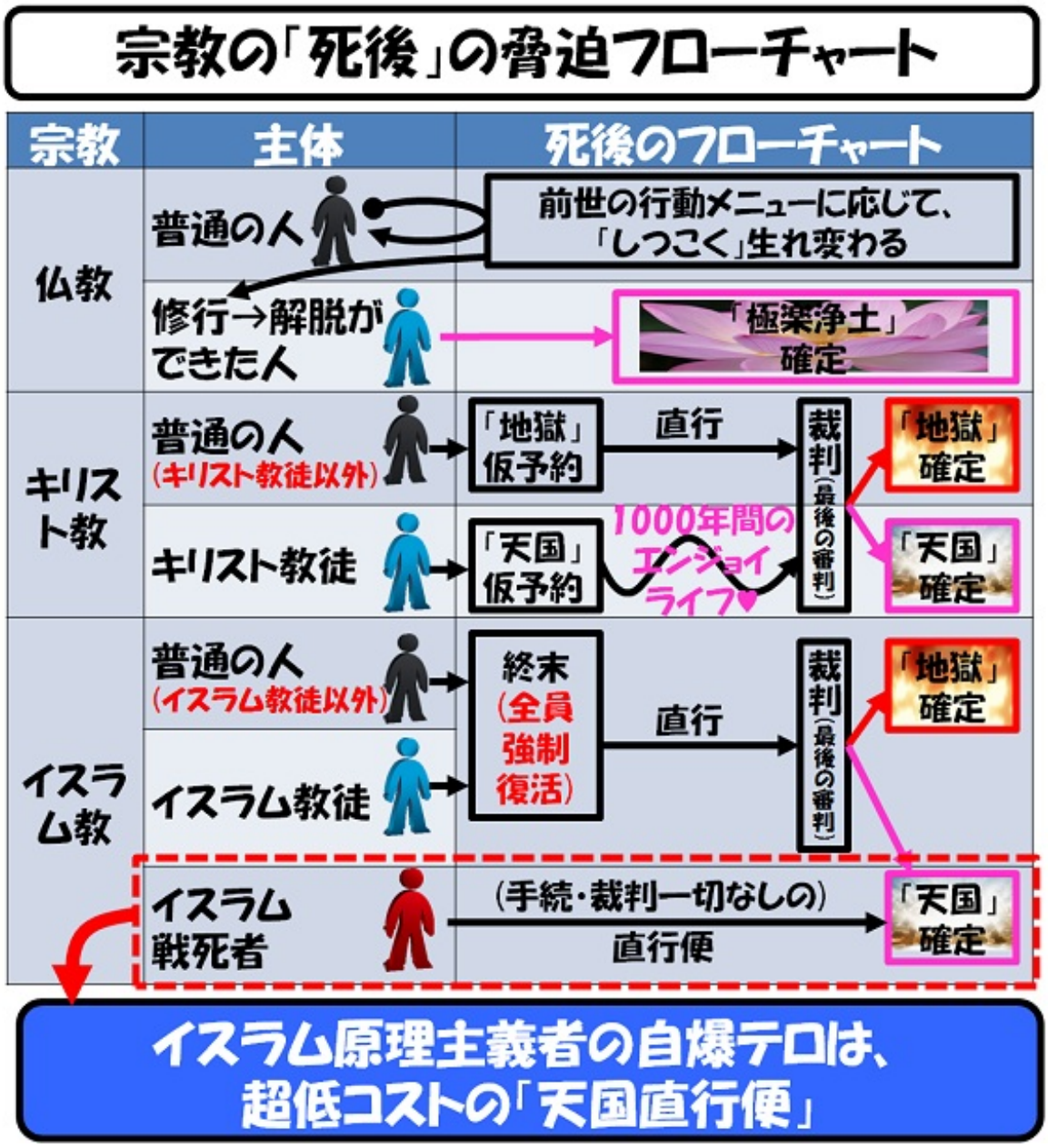
「宗教」という権威

　最初は「宗教」という権威から始めます。

　宗教の目的は、生物と同じで「世界中に広がり、信者を増やすこと」に付きます。ぶっちゃけ、宗教にはこれ以上の目的はなく、生物と同じアナロジーで働いています。

ただ、「宗教」の副次的な効果には、「安定した社会の実現」があります。これは別段、宗教でなくても、「軍隊」でも「法律」でも実現可能ですが、「軍隊」や「法律」が、現世での暴力装置（暴力や刑事罰）を持っているのに対して、宗教は「暴力装置が、死後に発動する」という点にあります。

今回、世界3大権威のそれぞれの宗教の「死後の脅迫システム」のフローチャート（概略図）を作成してみました。



ざっくりまとめると、異教徒については、地獄行き（キリスト教とイスラム教）、あるいは、解脱できなければ（アニメ『Steins;Gate（シュタインズ・ゲート）』のように）苦痛に満ちたタイムリープを繰り返す（仏教）という風に、「死後の脅迫システム」を設計しています。

比して、教徒については、各種の優待クーポンやサービスが提供され、特にジハード（聖戦）戦死者については、手続不問の天国直行便に乗車できるという優遇ぶりです。そして、このジハ

ードは、イスラム教の目的「世界中に広がり、信者を増やすこと」にも貢献していると思います。

しかし、このような「脅迫」や「優遇」は、しょせんは死後の話です。いまだかつて、誰一人として、その裏（証拠）が取れた人はいません。なぜ、そんな眉唾な「死後の脅迫システム」が、現世で機能しているのか不思議です。

それは、人間（を含む全ての生物）が、（1）必ず死に至り、（2）根源的にその死を恐れるように設計・製造されているからです。上記（1）（2）がDNAレベルでプログラミングされていて、それを後発的に書き換えることができないからです。

宗教は、このようにDNAにプログラミングされた「人間」という装置を、実に上手く制御するように作られています。その制御方法の考え方の1つに「パスカルの賭け」というものがあります。

上記の図の内容を、ゲーム理論*）的に説明すると、「神の存在が分からない現世にあっては、取りあえず神の存在を肯定しておくことが、死後における最適戦略である」ということになります。

*）関連記事：「陰湿な人工知能 ～「ハズレ」の中から「マシな奴」を選ぶ」

つまり宗教というのは、AI技術の一つである「ゲーム理論」を悪……もとい、援用した、高度な戦略理論に基づいて運用されているのです。

ここでは、人間にとって「死後の脅迫」が、現世の人間を動かす巨大な原動力となっている、という点を覚えておいてください。

「AIの知性」、超有名な2つの考え方

さて、今度は、「AIの知性」について、極めて権威のある2つの説をご紹介致します。

AIの「知性」の考え方

超有名な二つの考え方

名称	概要	その他
(1) チューリングテスト	(Step.1) ■人間が、AIと会話(チャット)している ↓ (Step.2) ■人間が、そのAIを「人間」と勘違いしたら、 ↓ (Step.3) ■そのAIには「知性がある」とする	■あの天才、アラン・チューリングが提唱(1950年) ■イメージとしては「自動応答Botと喧嘩した」等
(2) 中国語の部屋	(Step.1) ■紙切れ1枚しか出し入れできない部屋の中に、アルファベットしか理解できないイギリス人がいる ↓ (Step.2) ■紙切れには「『一∧◆●(中国語)」と書かれていれば『、♪ ♂ ♀(中国語)」と書き加えてから外に出せ」という記載がある。 ■イギリス人は、ひたすら、その指示に従う ↓ (Step.3) ■その部屋の外にいる人は、「この小部屋の中には中国人がいる」と考える ■しかし、実際には中国語なんぞ全く理解していないイギリス人が一人いるだけである	■哲学者ジョン・サールが思考実験として提唱(1980年) ■チューリングテストの「知性」に対する反論として捉えられる ■現存するAI技術は、ほとんどこのアナロジーで説明できる

チューリングテストとは、基本的には、「AIの知性は、直接測定できない」という腹のくくり方をした、非常に合理的な考え方です。

恐らく、あの天才アラン・チューリング先生*)は、「美醜」や「喜怒哀楽」と同様に、「知性」も第三者の主観に依存するものであり、定量化（点数をつける）ということが困難であると考えた上で、それでもなお「知性」を評価しなければならないとしたら、結局のところ「どの程度、人間をだまし、世界をだませるか」とするしかないと考えたのだと思います。

*）第二次大戦中、ドイツの潜水艦Uボートで使用されていた暗号システム「エニグマ」の解読方法を確立して、それをシステムとして組み上げた、超ド級の天才数学者です。

現在もなお、このチューリングテストは、AI技術の「能力」（多分「知性」ではないと思う）を評価する方法として、世界中で使われています。

これに対して、「中国語の部屋」は、このチューリングテストの唱える「知性」を鋭く批判します。

哲学者のジョン・サール先生*)の「中国語の部屋」を、ぶっちゃけて説明すると、「『Aと入力すればBと出力する』と動作するようにプログラムされている機械ごときに、知性なんぞが宿っていると、あんた本気で思うか？」というものです。そこには「知性」が宿っているのではなく、単なる「手続処理」の機能が実装されているだけだ、主張されているわけです。

*）ジョン・サール先生は、この後にも登場します。

で、残念なことに、現時点において、地球上にあるAI技術の全ては、全て、この「中国語の部屋」のアナロジーで説明できてしまうのです*)。

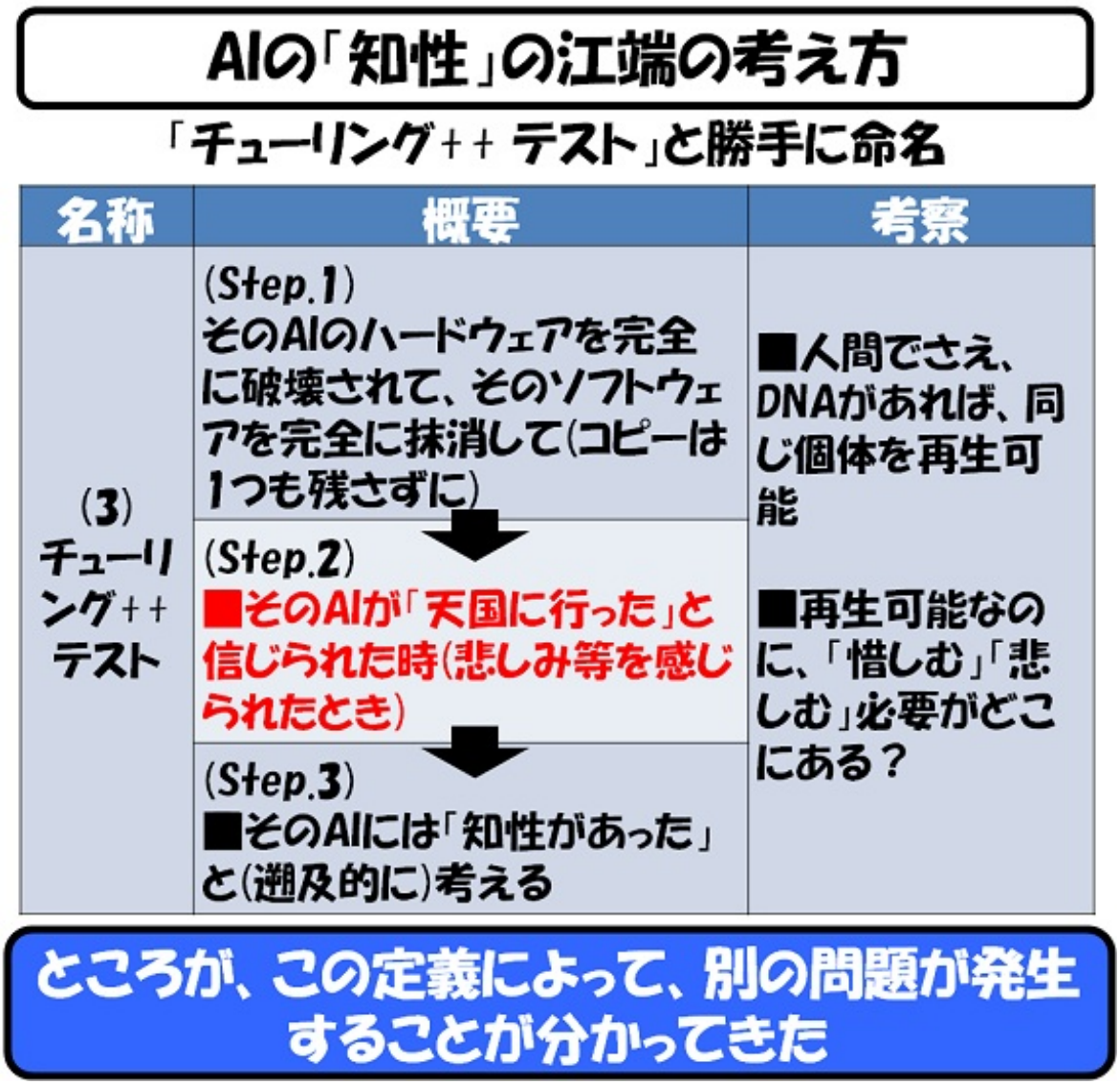
*）厳密に言えば、現状のAI技術では、『Aと入力すればBと出力する』を学習させた上で、『（Aに比較的近い）Cを入力して、（Bに比較的近いであろう）Dの出力を得る（ことを期待する）』ということもできますので、厳密に「中国語の部屋」とは言えないかもしれませんが、少なくとも現状のAI技術は、自分の発意で何かを実行することができないという点において「中国語の部屋」の概念の外には出ていません。

で、今回、この偉大な2人のAIの知性の考えを盗……もとい、援用して、私が考えた知性の定義は以下のものです。

私の「AIの知性」の考え方は、チューリング先生と同様に、『知性というのは、しょせんは、第三者による観測しか確認方法はない』という立場に立ちます。また、現在のコンピュータアーキテクチャでは、どうしたってジョン・サール先生の「中国語の部屋」の範ちゅうを超えることはできませんので、こっちは諦めることにしました。

その上で、AIに、人間的な根源的恐怖である「死」の観念……は無理としても、『無』に帰すという観念を導入した上で、「死」または「無」に直面した人間の方の感情を、観察することで、AI

の知性の定義を試みました。



AIは自己の「死」を観念できないけど……？

AIそれ自体は、自己の「死」とか「無」を観念できません（少なくとも、私は、プログラムをデリート（削除）する時に、『お願いだから、私を殺さないで』というメッセージが出てきたという経験はありません）。

私の考え方は、極めて単純です。

「知性」が定義できないのであれば、「そのAIが、人間に近かったか否かを、知性の有無の根拠とする」とした上で、「どれだけ人間に近かったかを知るには、そこに人間の側に、そのAIに対する感情（悲しさ、寂しさ）の発露があったかどうかで判定する*)」という ーー まあ、はっきり言って、チューニングテストの上書きパクリです。

*）SFアニメには、このアプローチがよく採用されています。ただ、それらのアニメでは、AIの持ち主として登場する主体が、ほとんど例外なく「美少女ロボット/アンドロイド」となっていますが。

この方式は、最初に述べたように、「人間の死の恐怖」とか、「宗教の死後の脅迫プロセス」などと、うまく組み合って、従来のチューニングテストよりは、社会的合意を得やすいものではないかとは自負しています。

ところが、この方式の検討中に、私は、非常に基本的なことに気がついてしまいました。

―― そもそも、AIって、殺せるんだっけ?

現時点において、AIはプログラムという形で実装されており、プログラムは、無限の複製が可能であり、ネットワークによって、世界中のどこにでもコピーが可能であり、人類が滅亡するまで、コンピュータの中に存在しつづけることができる、いわば「不死の無体財産物」です。

つまり、AI自体が「死」や「無」の概念を持っていないのは当然としても、私たち人間自身も、AIに対して死生観を持つことができないのです。

要するに、この段階で、既にこのテストが、テストとしての機能を果たさないことが、はっきりしてしまっているのです*)。

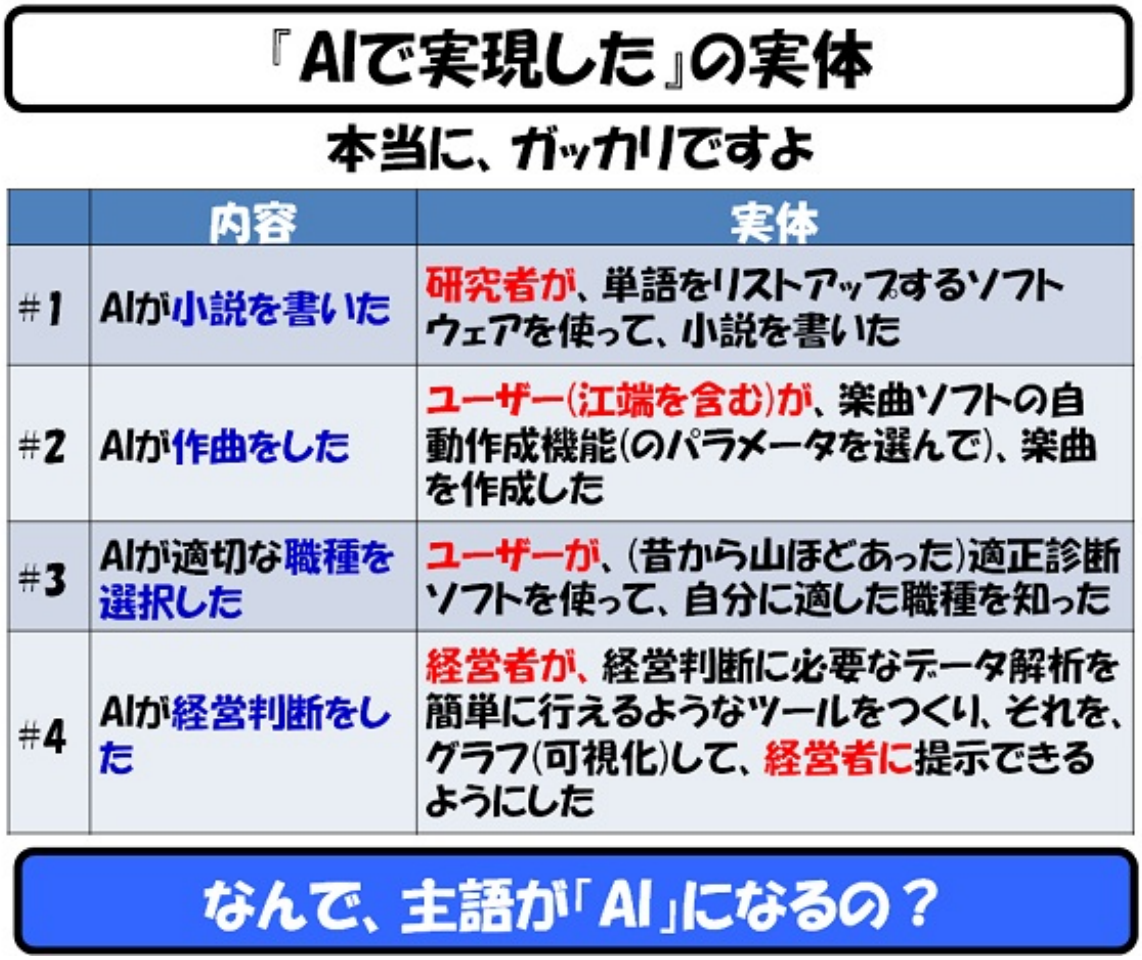
*)そういえば、SF「美少女ロボット/アンドロイド」アニメでは、この「AIの死」という観念を生成するために、随分設定に苦労しているみたいです(例えば、『記憶や人格のコピーはできない』とかいう、苦しい設定で対応している)。

「AIが○○した」という見出しが躍る

さて、「AIの知性」に関した話題としては、最近「AIが○○した」というフレーズが、新聞や雑誌、Webニュースの見出しとして踊っています。

特に最近では、「AIが小説を書いた」は、かなりセンセーションな話題となっていました。これが事実であれば、今世紀最大級の発明になっているはずですが ―― 今や、その話を全く聞かなくなりました。

そんな訳で、今回「AIが○○した」を、ざっくり調べてみたのですが ―― もう、本気で泣きそうになりそうなほど、それはそれはショボい話のオンパレードでした。



一つ一つ説明するのも面倒くさいので、割愛しますが、ともあれ、このニュースの内容で「AI」が主語になる理由が、全く分かりません。ここで「AI」といわれているものは、人間のアシスタントをしているだけのもので、（私見ではありますが）「ワード」「エクセル」またはIME（日本語入力ソフト）と何ら変わりません。

もっとも、私自身、この連載の第1回「中堅研究員はAIの向こう側に“知能”の夢を見るか」で、「製作者が『AI』と主張すれば、誰がどう反論しようが、それはAIである」という、江端ドクトリンを発信してはいるのですが。

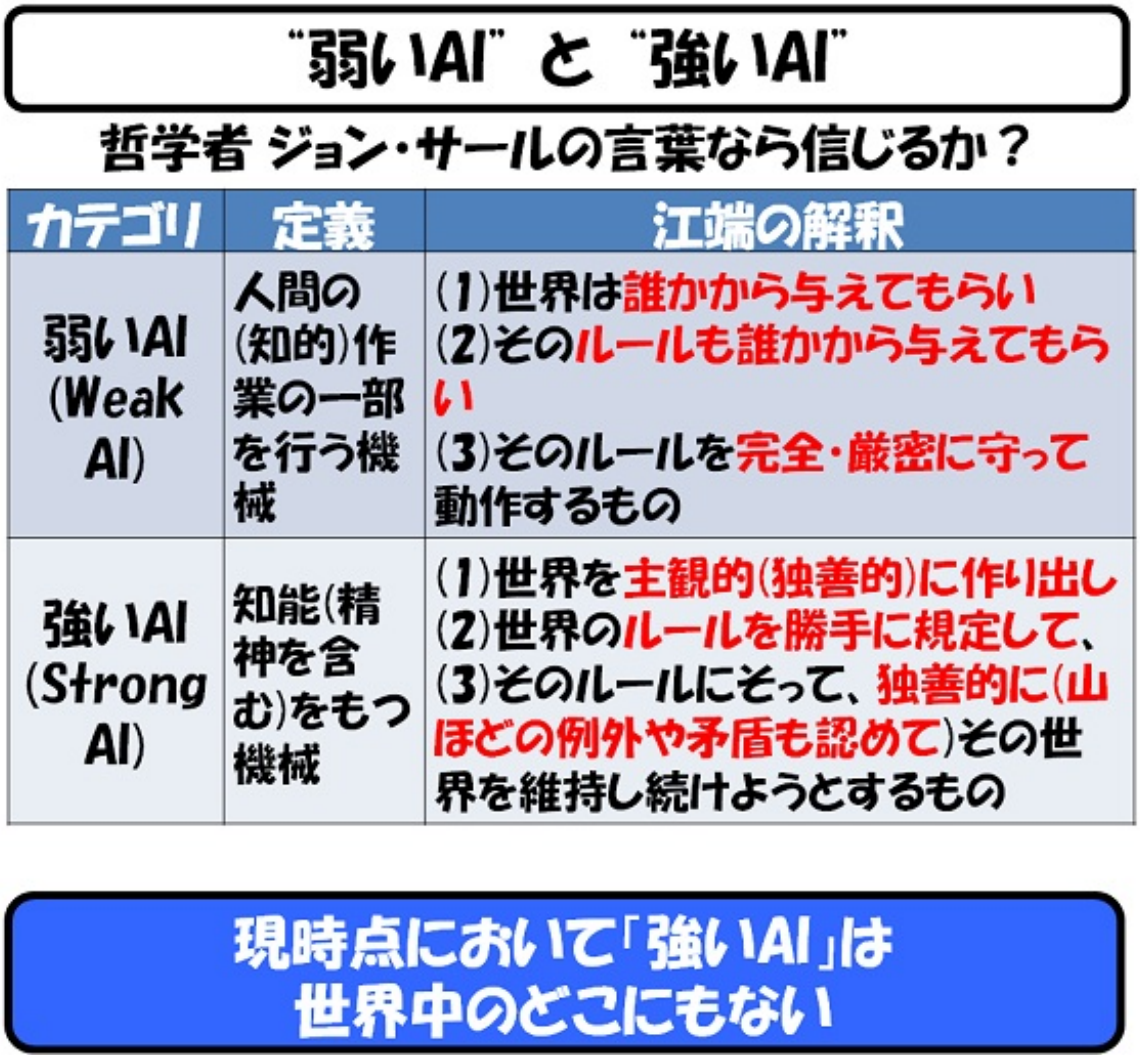
それでも、「AIが小説を書いた」と言われれば、大抵の人は、「何の操作もしていないパソコンのプリンタから、次々と小説が印刷され、それを読んで、人々が目を丸くする」という場面を想像しますよね（そんでもって、小説家やライターが失業する、という定番パターンの『AIによる失業』論が展開される、と）。

しかし、そのようなAIは、少なくとも現在のコンピュータアーキテクチャのパラダイムでは、絶対に起こり得ません。

以下、その理由について、いろいろな「権威」を引用して論を進めてみたいと思います。

「弱いAI」「強いAI」

先ほど紹介したジョン・サール先生は、私が感じているイライラに対して、AIを2つの概念、"弱いAI"と"強いAI"の2つで考えることを提唱しています。



簡単に言うと、"弱いAI"とは、人間の知的作業をサポートするAIであり、"強いAI"とは知能を持つAI、となります。

囲碁や将棋のAIは、一見"強いAI"のようにも見えます。しかし、囲碁や将棋のAIは、ルールに従って処理を進め、局面ごとに確率的に妥当な手を計算しているだけです。

ゲームに勝つと、そのゲームの内容をさかのぼって、局面単位の「手」の評価値（数値）を上げ、負ければ評価値を下げます。「勝負に勝つ」ことも「勝負に負ける」ことも、囲碁や将棋のAIにとっては、評価値を変更するために使われるパラメータの1つにすぎません。

もっと極端に言えば、囲碁や将棋のAIは「負ける方法」も学習しているといっても良いのです。そして囲碁や将棋のAIは、「負ける」ことが「悪いこと」などとは1mmたりとも考えていません。そういう意味において、囲碁や将棋のAIは、"強いAI"の範ちゅうに入れることは、できないのです。

「強いAI」とは

では"強いAI"とは何か ―― それは「10代の子どもを持つお母さん」です。

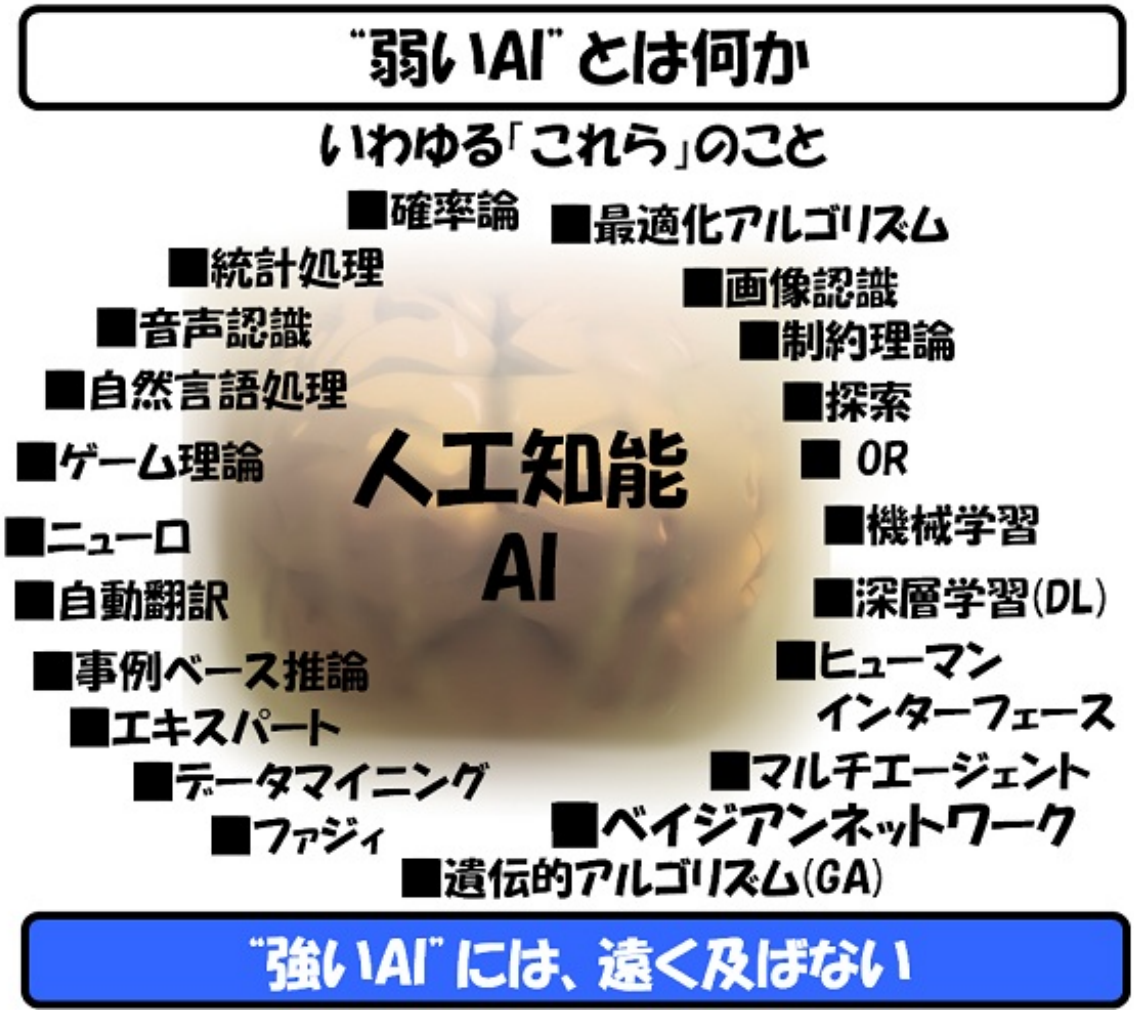
そのお母さんは、「女の子あるいは男の子とは、かくあるべき」という世界観を主観的かつ独善的に生成します。

そして、その世界観に基づき「男の子は、メソメソ泣いてはならない」「女の子は、門限までに帰宅しなければならない」「成人しても、男の子はいいけど、女の子はタバコを吸ってはいけない」という、説明不能な非論理的なルールを勝手に生成します。

さらに、それを根拠なく独善的に運用するだけでなく、「他の家の子ども」には適用しない（いわゆる「人は人、うちはうち」）という例外や矛盾を含むことをものともせず、その例外や矛盾の理由を一切説明することなく、その世界観を維持し続けます。

そのような「10代の子どもを持つお母さん」のように動作するAIこそが、"強いAI"であるのですが、コンピュータは、例外や矛盾を包含することを恐ろしく苦手としており、ましてや、独善的な世界観を構築する能力など、ひとかけらもありません。私が知る限り、そのようなAIは世界中探しても、どこにもありません。

つまり、今、世の中に存在しているAI技術は、全て"弱いAI"なのです。



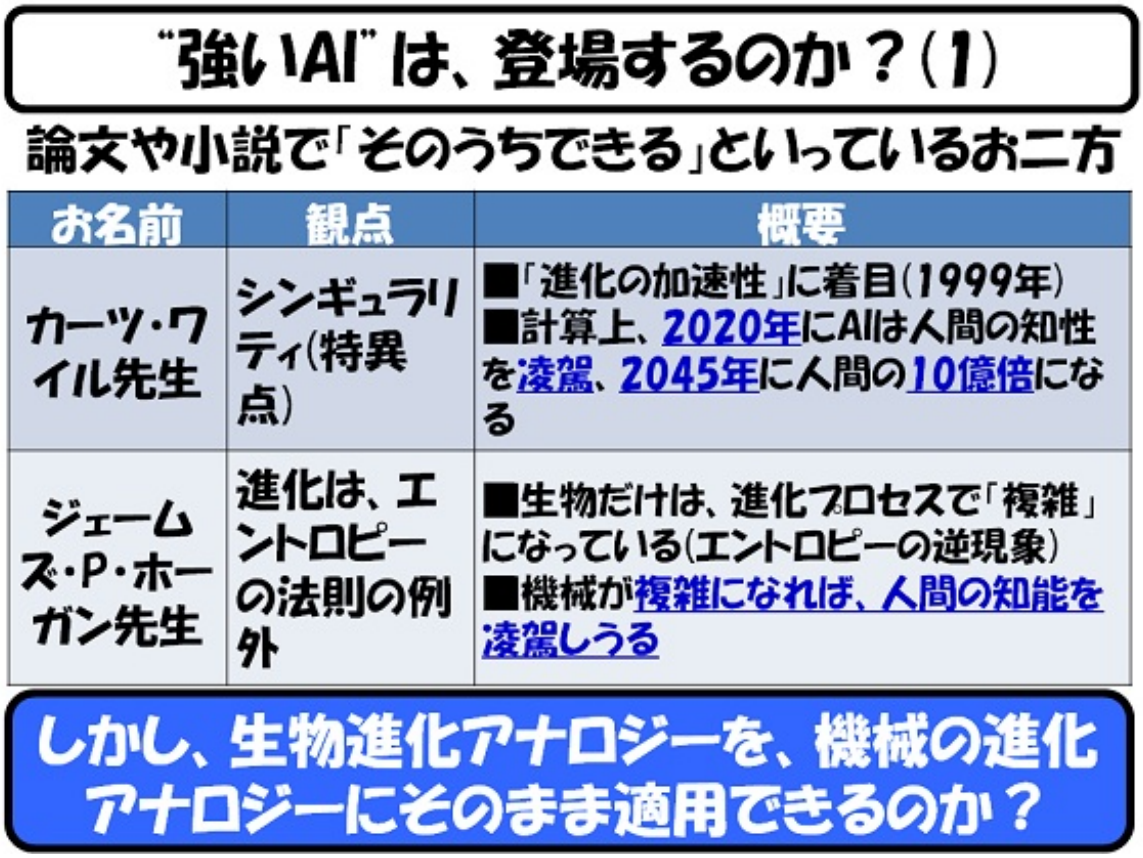
ですから、これからは、「"AI"が小説を書いた」ではなく「"弱いAI"が小説を書いた」、あるいは、「AIが作曲をした」ではなく「"弱いAI"が作曲をした」という風に記載することを、私は提言したいと思います。

そうすれば、「AIが人類を滅亡させる」だの「AIが国家を乗っとる」だのという、バカげた妄想を触れ回っている人間の数は、少しは減るのではないかと思います(「"弱いAI"が人類を滅亡させる、国家を転覆させる」という主張なら、まあO.K.としましょう)。

もちろん、AI研究者の究極の夢は、「ティーンエージャーの子どもを持つお母さん」……じゃなくて、"強いAI"を作り出すことにあります。

「強いAIはいずれ出現する」と主張する著名人たち

では次に、「この"強いAI"は、いずれ出現するはずだ」と主張をされている、または小説などの中で記載されている権威ある著名人2人をご紹介します。



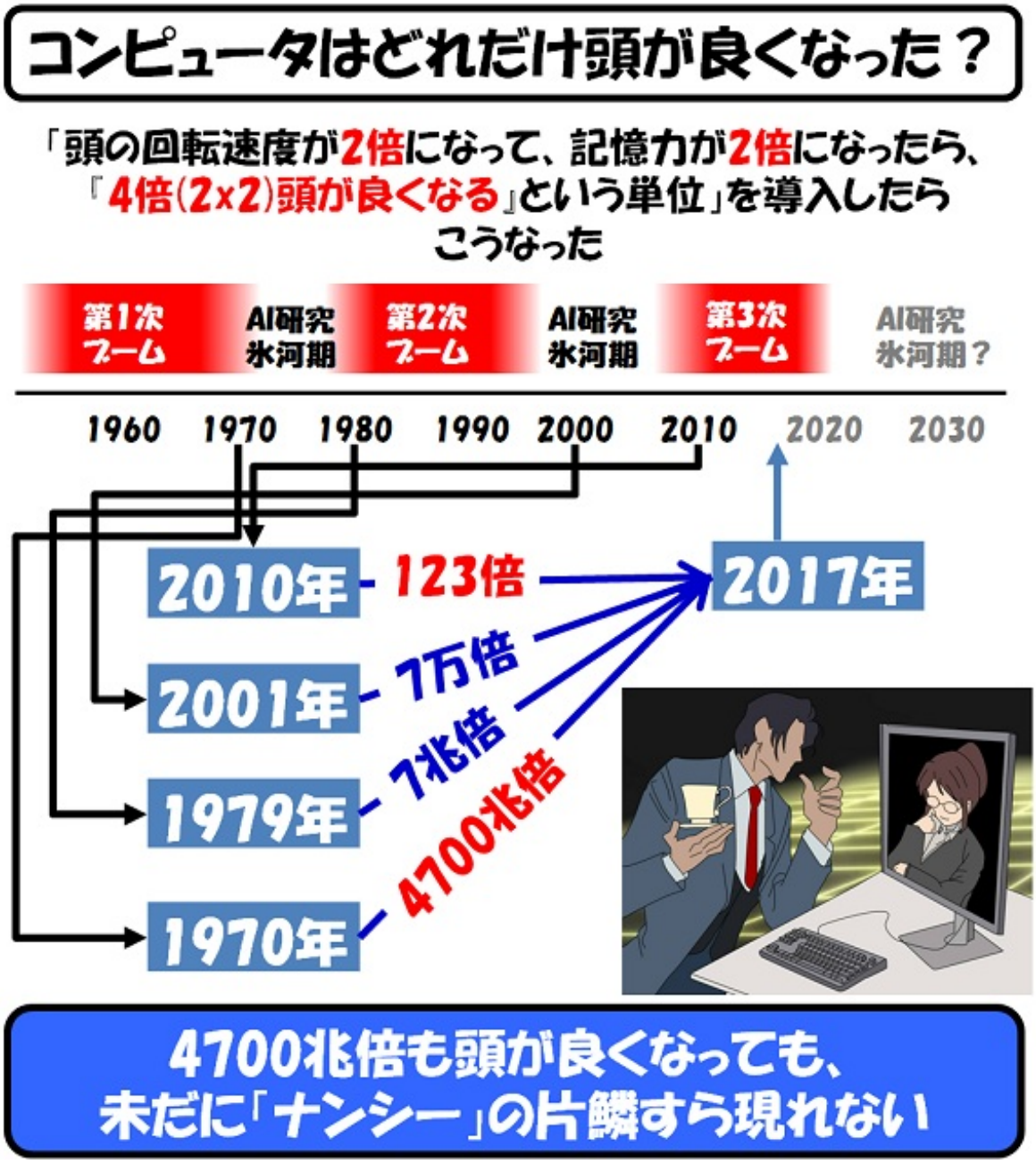
カーツ・ワイル先生は、生物や機械の「進化の加速性」というものをきちんと計算した上で、そのクロスポイントとなる特異点（シンギュラリティ）を逆算して導き出しています。

残念ながら、上の表に記載のあるカーツ・ワイル先生の未来予測の年数は、外れることになりそうですが、それでも私は、カーツ・ワイル先生の分析は、シミュレーションも行わず、データも示さず、「あと20年後にAIで○○ができる」などと適当に語る博士や教授たちとは、一線を画していると思っています（関連記事：陰湿な人工知能 ～「ハズレ」の中から「マシな奴」を選ぶ）。

4700兆倍賢くなっても、「ナンシー」が出現する気配すらない

カーツ・ワイル先生は、ムーアの法則から、ある知識量が一定の飽和点に至った時に、知識が知性に変質する、という仮説を立てている（と私は解釈した）のですが、この仮説については、私がネガティブです。

その根拠は、既に前々回の「おうちにやってくる人工知能 ～ 国家や大企業によるAI技術独占時代の終焉」でお話した通り「4700兆倍の計算能力」です（下図は再掲）。



この値は、これからも指数関数的に増大していくことは間違いありません。現時点で、ここまで大きな値になって、まだ「知性の萌芽」すら見えていないことは、「知識がどんなに積み重なろうとも、それらが知性に変質することはない」という、1つの証拠になっていると思うのです。

機械は自然淘汰などできない

一方、SF作家のジェームズ・P・ホーガン先生は、小説「未来の二つの顔」の中で、「進化プロセスにおける、生命体の逆エントロピー現象」について言及しています。

「知性は複雑性の増大によって向上している」ことは、（小説とは関係なく）現実として確認できる事項です。その事実から、この小説の中では、「さらに複雑化していく無機的知能（コンピュータ）が最終的に人間の知能を凌駕（りょうが）することは、十分にありえる」と、登場人物の1人に語らせています。

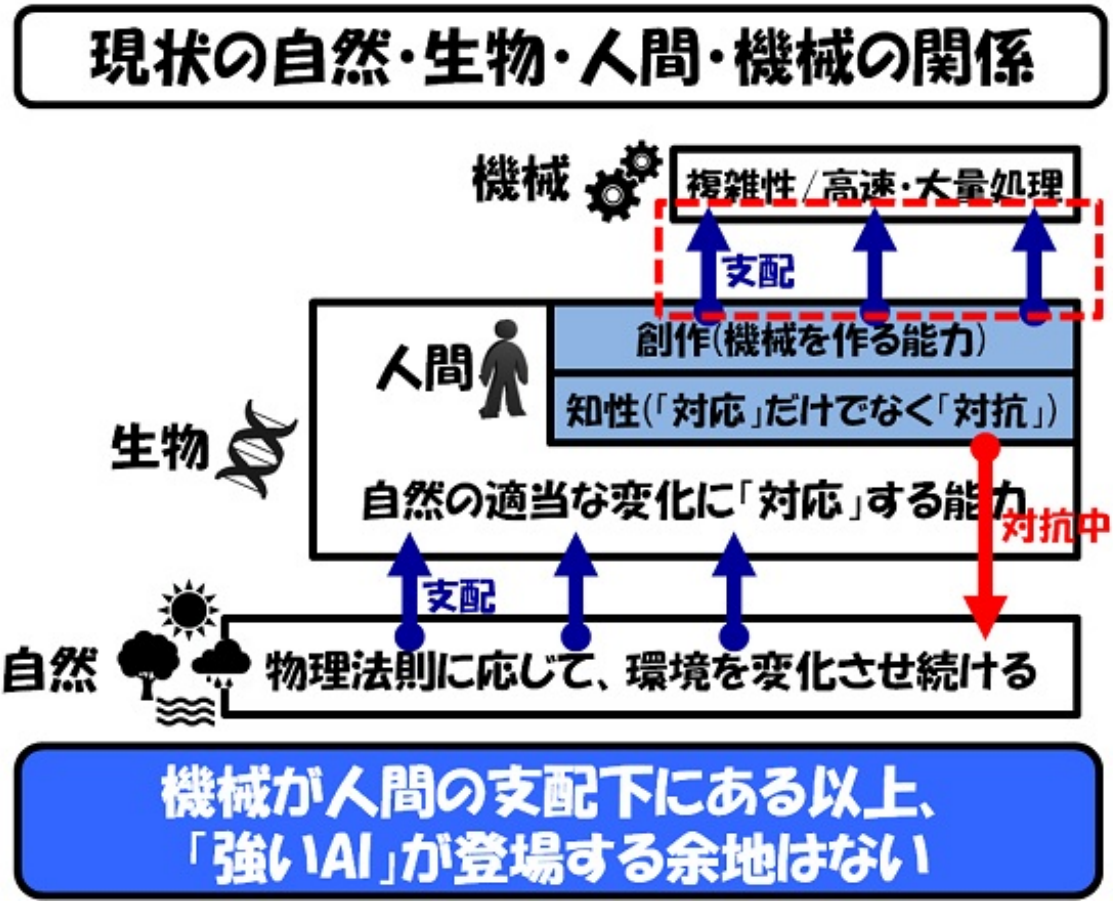
しかし、私は、その意見に対してもネガティブです。なぜなら、生物は環境に適用するために自然淘汰（とうた）メカニズムを使ってきて、その適用能力が進化を促し、その結果として知性を獲得したと考えられますが、機械を複雑にできるのは人間だけであり、機械は自然淘汰メカニズムを使えません（例えば、パソコンを屋外に野ざらしにしておけば、確実に壊れますよね。自動的に防水機能を具備したパソコンに変化することはできません）。

先ほど、"強いAI"は、独善的な世界観とルールを作って、その世界を運用することのできる「10代の子どもを持つお母さん」のようなものだと申し上げましたが、これは、当然のことなのです。

生物は、独善的な世界観（生き残るためなら何をしても良い、など）を自力で確立し、そして、そこに正当性やら論理性や合理性なんぞを踏みにじっても、生き残る必要があったからです。

これに対して、コンピュータは、ちょっとしたプログラムの矛盾記述や、データフォーマットのミスで、たちまち停止（最悪の場合は、暴走）してしまいます。これは、"強いAI"のアナロジーの真逆に立っていると言えます。

ここまでの話を図でまとめてみると、こんな感じになると思います。

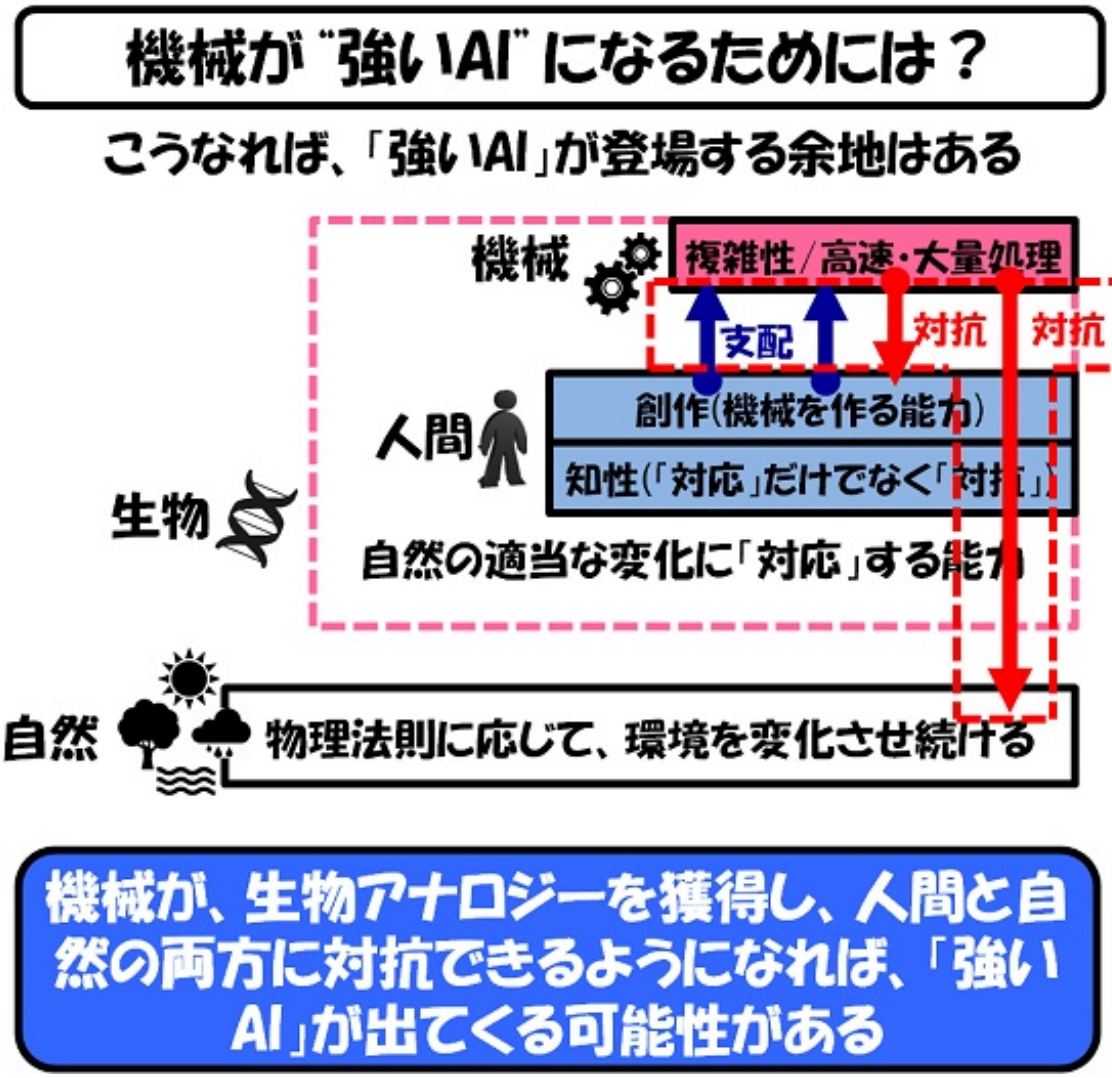


気まぐれででたらめに変化しつづける自然環境に対して、生物は、自然淘汰というメカニズムで、自らを変化させ続け、対応する能力を獲得してきました。その中でも人間は、運良く「知性」にたどり着くことができたのです。

さらに、生物の中でも人間だけが、自然に対抗し始めました（農作物を計画的に作ることから、地形に手を加えてダムを作ることまで）。加えて、機械を発明して、人間の仕事の一部を代替させることを実現し、さらにはその機械の効率を恐ろしいレベルまで引き上げました（前述の4700兆倍など）。

しかし、それは、逆に言えば、機械は人間にのみ依存しており、人間というアシストなくして、進化する手段を持たないということでもあります。少なくとも、機械には、自然淘汰メカニズムを使える可能性がないからです。

ならば、人間と同様の知性を有する"強いAI"を生み出すにはどのような仕組みが必要になるのでしょうか。それを、図示してみます。



もし、機械が生物と同じように自然淘汰メカニズムを使って、自らを変化させる能力を獲得すれば、人間と同様の知性を得た"強いAI"となる可能性があります。

そして、このような"強いAI"は、人間と同様、自力で自然に対抗するようになるだけでなく、人間にすら対抗するようになるはずです。

しかし、機械が、自然淘汰メカニズムを獲得する可能性は、絶望的に低い（というか、正直にいうと"ゼロ"）だろうと思います。例えば、先ほどの「雨ざらしのパソコン」のように、壊れるだけで終了してしまうでしょう。

もっとも、移動機能やエネルギー変換機能を具備したパソコンである「ロボット」であれば、いずれは、自然淘汰メカニズムを取り込むことができるようになるかもしれません。

しかし、それでも、まだ問題は残ります。

自然淘汰メカニズムは、生存競争が前提になります。そのためには、大量の個体を自己生成する「生殖」機能が必要です。しかし、パソコンがパソコンを自分で作り出す（あるいは、ロボットがロボットを作り出す）機械の生殖メカニズムが、私には、どうにもイメージできません。

その上、ものすごく長い時間が必要となりそうです。100万年とまでは言いませんが、どんなに少なく見積っても（コンピュータの加速的進化を考慮しても）1000年くらいは必要なのではないかと思います。

「強いAI」は作れるのか

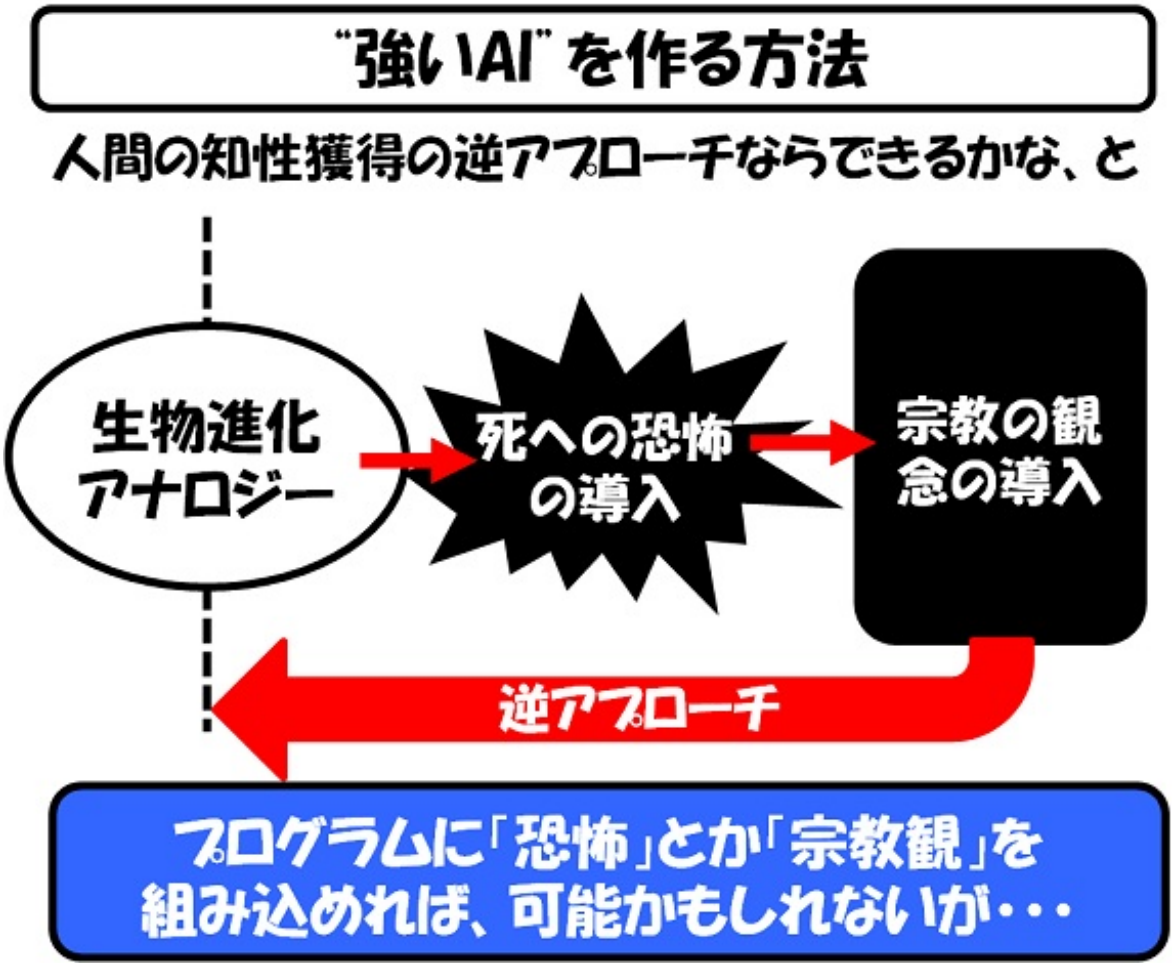
私には、1000年近くもチンタラ待っている時間はありません。ここ10年くらいの間で"強いAI"――「10代の子どもを持つお母さん」のように振る舞うAIを、この目で見てみたいのです。

ここで、ようやく、冒頭の話に戻ります。

進化のプロセスには、自然淘汰プロセスが必要で、そのためには生存競争が前提で、その生存競争には「死の恐怖」が必要でした。

そんでもって、知性の段階に獲得した人間は、死の恐怖を少しでも軽減させるために、（検証も証拠もない）死後の世界を規定した「宗教」なるものを発明しました。

"強いAI"を作る1つの方法は、このアプローチを地道に実施する（つまり、何もしないで、機械が自力で進化するのを気長に待つ）という方法もありますが、いっそのこと、この進化のアプローチを、ひっくり返して教えてやれば良いのではないかと、考えました。



つまり、宗教観をコーディングしたAI技術です。AIに死後の世界を観念させるような実装を行えば、そこからAIは「死の恐怖」を獲得し、その結果として、環境に適用できる"強いAI"への自律的な変化を起させることができるのではないか、という、1つの仮説（あるいは妄想）です。

さきほど私は、AIには「死」や「無」が存在しないと言いましたが、それはそれで構わないのです。AIに「死」や「無」が存在しないとも、そのようにAIに「思わせる」ことができれば、それで十分だからです。

これは、膨大な時間をかけて、生物の進化プロセスをトライ&エラーで行わせるよりも、既にコンテクト化されている印刷物（聖書、コーラン、経典、その他）を使うことができるので、そこからテキストマイニング技術を使って、AIに教え込めば良いのです。

―― とまあ、いろいろと苦労して考えてみたのですが、やっぱりこのプロセスでも、"強いAI"は作れそうにはありません。

そもそも、これは、"弱いAI"を使って、なんとか"強いAI"を作り出せないかという、思考実験でした。

しかし、"弱いAI"は、規定された知識とルールの中で計算をすることしかできません。"弱いAI"に、宗教コンテンツをどんなに大量に知識吸収させようとも、そこから死の恐怖にさかのぼるというような、コンテクスト変換（またはパラダイムシフト）が、自然に発生するわけがないか

らです（実際に、私が上記でそのように論じています）。

念のため、同じようなことを考えている人がいないかと思って、検索エンジンを駆使して「コンピュータ、宗教、学習」「コンピュータ、恐怖の獲得」とか調べてみたのですが、全くヒットしませんでした。

少なくとも、現時点の私は、"強いAI"を作り出すプロセスを思い付くことができません。

もしかしたら、ひょんなことがきっかけで、"強いAI"が突然登場する、というような、シンギュラリティ（特異点）がやっているかもしれませんが、すくなくとも現時点において、『今、私が、どう思っているか』というのであれば――。

"強いAI"などというものは、作り出すことはできないし、そもそも存在しないと考えるのが『合理的』だと思っているのです。

それでは、今回のコラムの内容をまとめてみたいと思います。

【1】宗教の目的は、生物と同様「世界中に広がり、信者を増やすこと」であり、その手法の1つが、死後に発動させるという「脅迫システム」にあることを説明しました。そして、そのような、検証不能な死後の世界の脅迫に対して、多くの人間が屈してしまう理由として、ゲーム理論の1つとして把握できる「パスカルの賭け」を紹介しました。

【2】AIの「知性」を計測する権威ある有名な手段として「チューリングテスト」の説明を行い、また同時に、そのテストの批判として「中国語の部屋」という思考実験の内容について説明しました。

【3】さらに、「知性は直接観測できない」という立場に立った上で、「AIの死」という観念を導入して、その「死」に対する、人間サイドの感情（悲しみなど）の発生の有無で、知性の有無を反転するという「チューリング++テスト」を提唱したものの、「AIを殺すことができない」という厳然たる事実から、このテストが有効に働かないことを明らかにしました。

【4】近年、「AIが小説を書く」など、AIに関するセンセーショナルなニュースが巷（ちまた）を騒がせていますが、実際に調べてみると、AI技術を利用して、人間が小説を書いているだけであり、これまでのコンピュータのアプリケーションの範ちゅうを超えない、ショボい内容であったことが分かりました。

【5】世間のAIに対する誤解を回避する方法としてとして、哲学者のジョン・サール先生の唱える"弱いAI"と"強いAI"という考え方を紹介しました。そして、現時点において、知能（精神を含む）を持つ機械である"強いAI"が、世界のどこにも存在しないことを明らかにしました。

【6】"強いAI"を実現する手段として、生物の進化アナロジーを逆方向からアプローチする手法について検討をしました。具体的には、AIに、コンテクストとして提供可能な、宗教の「脅迫シス

テム」を学習させ、さらには「死の概念」を獲得させることで、"弱いAI"から、"強いAI"を作り出す手段について案出しました。

　しかし、現状の"弱いAI"では、宗教的観念から死の概念にパラダイムシフトさせうるような方法が存在しないため、この手法は有効に働かないだろうという結論に至りました。

【7】結果として、"強いAI"は、今後も登場することはないであろうという、江端の見解を示しました。

パソコンにも『あの世』とか『来世』がある？

　冒頭の、次女との会話の続きです。

次女：「"パソコンにも『あの世』とか『来世』がある"って、どういう意味？」

江端：「パソコンは電気信号のON/OFFで演算するし、人間の頭脳もニューロンという物体同士の電気信号のON/OFF（ニューロンの発火）で稼働している。パソコンも人間も、同じアナロジーで動いている有体物といえる」

次女：「それは極論だよ。パソコンは、人間のように創作もできないし、感情もないよ」

江端：「それは、コンピュータが、現時点で『創作』や『感情』を表現できるレベルにないだけかもしれない。そもそも、人間の『創作』や『感情』も、電気信号の組み合わせによる演算結果であることは、絶対的な事実だから、この点は議論の対象としないことにしてほしい」

次女：「分かった」

江端：「さて、同じく電気信号で動いている2つのモノの、一方のみに『あの世』とか『来世』があり、他方にはそれらがない、というのは、論理的一貫性を欠くと思うが、どうだろうか？」

次女：「つまり、パパは、『人間だけが特別扱いされる死後の世界は、おかしい』と言いたいわけだね」

江端：「ま、ぎりぎり、『知性のあるもの』を特別扱いするのはいいかな、くらいには思っているんだけどね。だから、コンピュータの知性を、『あの世』とか『来世』という観点から創造できないもんかなーと、常々考えていて……」

と、話を続けようとしていたところ、私たちの横で、黙って話を聞いていた嫁さんのドス黒いオーラが、ハンパでない状態になっているのに気が付きました。

　私たちはそのオーラにおののきつつ、『続きは、試験の後な』と言いながら、そそくさと、それぞれ「試験勉強」と「執筆作業」に戻りました。

⇒「Over the AI ――AIの向こう側に」⇒連載バックナンバー

Profile

江端智一（えばた ともいち）

　日本の大手総合電機メーカーの主任研究員。1991年に入社。「サンマとサバ」を2種類のセンサーだけで判別するという電子レンジの食品自動判別アルゴリズムの発明を皮切りに、エンジン制御からネットワーク監視、無線ネットワーク、屋内GPS、鉄道システムまで幅広い分野の研究開発に携わる。

　意外な視点から繰り出される特許発明には定評が高く、特許権に関して強いこだわりを持つ。特に熾烈（しれつ）を極めた海外特許庁との戦いにおいて、審査官を交代させるまで戦い抜いて特許査定を奪取した話は、今なお伝説として「本人」が語り継いでいる。共同研究のために赴任した米国での2年間の生活では、会話の1割の単語だけを拾って残りの9割を推測し、相手の言っている内容を理解しないで会話を強行するという希少な能力を獲得し、凱旋帰国。

　私生活においては、辛辣（しんらつ）な切り口で語られるエッセイをWebサイト「こぼれネット」で発表し続け、カルト的なファンから圧倒的な支持を得ている。また週末には、LANを敷設するために自宅の庭に穴を掘り、侵入検知センサーを設置し、24時間体制のホームセキュリティシステムを構築することを趣味としている。このシステムは現在も拡張を続けており、その完成形態は「本人」も知らない。

本連載の内容は、個人の意見および見解であり、所属する組織を代表したものではありません。

関連記事

陰湿な人工知能 ～「ハズレ」の中から「マシな奴」を選ぶ
「せっかく参加したけど、この合コンはハズレだ」――。いえいえ、結論を急がないでください。「イケてない奴」の中から「マシな奴」を選ぶという、大変興味深い人工知能技術があるのです。今回はその技術を、「グルメな彼氏を姉妹で奪い合う」という泥沼な（？）シチュエーションを設定して解説しましょう。

IoT時代、ディープラーニングの主用途は制御か
Preferred Networks（PFN）は、IoT（モノのインターネット）にディープラーニングを活用する意義について、自社の応用例を挙げて説明した。また、IoTの普及でデータ量がどの程度増加するかを取り上げた上で、ディープラーニングの処理性能向上に向けた自社の取り組みについて語った。

上司の帰宅は最強の「残業低減策」だ ～「働き方改革」に悩む現場から
あなた（あなたの会社）は、「働き方改革」に本気で取り組んでいますか？　読者の皆さんの中には、「誰かの上司」という立場である方も極めて多いと思われます。そんな皆さんに伝えたい――。シミュレーションで分かった「残業を減らす最善策」、それは皆さんが今すぐ、とっとと帰宅することなのです。

もはや我慢の限界だ！ 追い詰められる開発部門
コストの削減と開発期間の短縮は、程度の差はあれ、どの企業にとっても共通の課題になっている。経営陣と

顧客との間で“板挟み”になり、苦しむ開発エンジニアたち……。本連載は、ある1人の中堅エンジニアが、構造改革の波に飲まれ“諦めムード”が漂う自社をどうにかしようと立ち上がり、半年間にわたって改革に挑む物語である。

MIPSコンピュータをめぐる栄枯盛衰

RISCプロセッサの命令セットアーキテクチャである「MIPS」。そのMIPSを採用したワークステーションの開発には、日本企業も深く関わった、栄枯盛衰の歴史がある。

研究開発コミュニティーが置かれた危うい状況

研究開発コミュニティーは「常に」危機に曝されてきた。研究開発に関わるエンジニアであれば、「研究不正」「偽論文誌・偽学会」「疑似科学」といった、研究開発コミュニティーを取り巻くダークサイドを知っておくにこしたことはない。本連載では、こうしたダークサイドを紹介するとともに、その背景にあるものを検討していく。