全訳作成者：廣瀬太雅

倫理的な OS
今日の科学技術による将来的な影響を予想する為の手引き——もしくは作り上げたものを後悔せずに済む方法

出典・凡例

本稿は，Omidyar Network. (2018). Ethical OS. Retrieved December 11, 2020, from https://ethicalos.org/ の全訳である。

原語を付すために () を用いた。また，訳者が補足する際には〔〕を用いた。註はいずれも訳註である。

もし未来を予想できないのなら、間違った仕事をしている（冗談です）

（原文 p. 2）

明日が何をもたらすかを正確に予測することは誰にもできない（技術界のどこかで誰かがそれに取り組んでいることは間違いないが）。故に〔未来を予知する〕水晶玉アプリを手に入れるまでは、私たちができる最善のことは、今日私たちが生み出した技術が長期的に社会に与える影響や予期せぬ利用法を予期することだ。

未来を予測する必要はない。ただ、未来に何が可能なのかを予期するのが上手くなればいいのである。最も望まないことは、「あなた」が創るのを手伝った未来に不意を突かれることだ。それこそ最悪である。

なので自分自身に問いかけてみてください

（原文 p. 3）

本資料は，JST，RISTEX，JJPMJRX19B5 の支援を受けた研究成果の一部です。

・たった今作っている技術がいつか思いがけない形で使われるとしたら、どのように準備することを期待できるか？
・今、どのような新しいカテゴリーのリスクに特に注意を払うべきか？
・そして、どのようなデザイン、チーム、ビジネスモデルの選択が、**ユーザー、コミュニティ、社会、そして会社**を将来のリスクから積極的に守ることができるのか？

倫理的なオペレーティングシステム

（原文 p. 4）

倫理的なオペレーティングシステムは、技術の製作者、プロダクトマネージャー、エンジニアなどが、問題を防ぐのを助けることができる。これは、より良い製品開発、より迅速な展開、そしてよりインパクトのあるイノベーションを促進するように設計されている。そして、同時に技術的および評判のリスクを最小限に抑えるよう努めるものである。

このツールキット[1]は、現在の設計プロセスに情報を提供し、将来的には既存の技術に関するリスクを管理するのに役立つ。

なぜ危険にさらすのか？

（原文 p. 5）

技術者である私たちが、自分たちの技術が世界をより良い方向に変えることに焦点を当てて過ごすのは当然のことだ。それは素晴らしいことである。誰もが明るい性質（sunny disposition）が好きだ。

しかし、ある点では、悲観的に[2]考えたほうが有益かもしれない。もし、私たちの技術が世界を救う方法を空想することに加えて、それがもしかしたら、ひょっとしたら、ことによると、すべてを台無しにするすべての方法を恐れることに時間を費やしたとしたらどうか？

最も望まないことは、楽観主義によって、技術に生じうる欠点の周りに盲点が生まれることである。倫理的な OS は、より明確に物事を見るのを助けるためにここにある。

[1] ツールキットとは一組のソフトウェアツールのこと。
[2] 原文の”glass half empty”だが、これはグラスの中身がもう半分しか残っていないという悲観的な考えを表す。

目次

本資料は，JST，RISTEX，JJPMJRX19B5 の支援を受けた研究成果の一部です。

倫理的な OS：どのように作用するか

—未来を可視化し予期するとともによりよいエコシステム[3]を作る為のツール—

（原文 pp. 8-68）

Tool 1: 明日のリスク — 今日

（原文 p. 9）

今日が昨日の未来なら、明日のリスクがあなたの問題だ。

技術が成長し、進化していくにつれて、今は地平線上にあるように見えていた（もしくは完全に地平線上にあった）リスクがすぐに顕在化してくることは言うまでもない。多くの場合、ほとんど何の前触れもない。

遥か彼方のリスクを想像する訓練をしなければならない。

このツールを使って、あなたの想像力をストレッチし、先見性の筋肉のウォーミングアップを行い、チームとの重要な会話を始める。製品のアイデアを出す為のヨガのようなものだ。

Tool 2: 科学技術をチェックすること —リスク低減の手引き—

（原文 p. 10）

技術が違えば当然リスクも異なる。どれが関係するものなのか？

このチェックリストを使用して、リスクと社会的危害の新興分野のうち、どの分野が最も重要なのかを特定し、今すぐ検討を開始しよう。

Tool 3: 将来に備える戦略

（原文 p. 11）

難しい質問をしてきた。未来を深く調べてきた。潜在的なリスクを特定した。次は何をする？学んだことをもとに行動すべき時だ。

・特定したリスクをどのように優先順位をつけるべきか？最大の脅威はどれか？対応が最も困難なのはどれか？

[3] IT の文脈での意味は、デベロッパーやベンダーが平等に利益を得られる協業形式。

本資料は，JST，RISTEX，JJPMJRX19B5 の支援を受けた研究成果の一部です。

・これらのリスクを低減するには、どのような戦略が有効か？
・どこから、どのように始めるべきか？
・前進するためには、他に誰の協力が必要か？

これらの戦略を使って、より持続可能な科学技術エコシステムを推進するための倫理的システムの可能性を考えよう。

Tool 1: 明日のリスク —今日—　対話を始めるための 14 のシナリオ

（原文 pp. 12-29）

地平線上の潜在的なリスクをどれだけ明確かつ迅速に見極めることができるか？

Institute for the Future と Omidyar Network は、社会に大きな影響を与える意図しない結果が発生する可能性のあるイノベーションの分野がないか調査した。私たちは、技術リーダーたち（tech leaders）に、どのような潜在的なリスクや倫理的ジレンマを最も懸念しているかを尋ねた。そして、可能であれば、人々がすでにこれらの害を低減するために取り組んでいるポジティブなシグナルの例を特定した。

シグナルとは、将来に影響を与えたり、形作ったりする可能性のある何かの現在の具体的な例のことである。それは、物事がすでにどのように異なったものになっているかを示す手掛かりとなる。

覚えておいてほしい：シナリオが起こりそうかどうか、あるいは可能性があるかどうかにさえとらわれてはいけない。一つ選んで、それで行くだけだ。

（原文 p. 14）

「誰も予想しなかった重大な結果が起きたとき、私たちはしばしばそれが想像を絶するものであったと言うことがある。しかし実際には想像できないことはない。私たちが何かを想像できなかったと言うとき、たいていの場合、それは私たちの想像力を正しい方向に向けることができなかったことを意味している。」
—Institute for the Future、ジェーン・マクゴニガル

Tool 1: 何をすべきか

（原文 p. 15）

これらの可能性のある領域やシナリオは、ここでは「14 の危険な未来」として提示されている。これらは、科学技術企業が社会的リスクを管理し、長期的な結果を予測しなければならない責任と機会が増大していることについての議論を促すことを目的としている。それぞれの未来のシナリオを確認しながら、考えてほしい。

・このシナリオであなたが最も心配していることは何か？
・この将来によって、異なるユーザーはどのように異なる影響を受けるかもしれないか？
・プライバシー、真実、民主主義、精神衛生、市民的論議（civic discourse）、機会の平等、経済的安定性、公共の安全を守るために、あなたはどのような行動をとるか？

本資料は，JST，RISTEX，JJPMJRX19B5 の支援を受けた研究成果の一部です。

・この危険な未来に備えるために、今、私たちは何ができるのか？

以下のようなシナリオに準備はできているだろうか……

真実、故意の誤報 (Disinformation)、プロパガンダ

（原文 p. 16）

シナリオ１：フェイク動画のアルゴリズムが非常に発達し、フェイクされたビデオを本物の映像と区別することが不可能になった世界への準備はできているか？これらのアルゴリズムは、人の顔や言葉を完全に置き換えて、作者が意図した通りの発言や行動をしたように見せることができる。この技術を使えば、誰でもどんな主張も裏付ける「証拠」となるビデオを作ることができる。これらのビデオ・フェイクは、主要なビデオ共有、ソーシャルメディア、ライブ・ストリーミング・プラットフォームに氾濫している。
〔https://thenextweb.com/artificial-intelligence/2018/02/21/deepfakes-algorithm-nails-donald-trump-in-most-convincing-fake-yet/〕

依存症とドーパミン経済

（原文 pp. 17-18）

シナリオ２：ソーシャルメディアの公開投稿から収集したデータセットを使用して、特定の人物を模倣するように**会話ボット**が訓練された世界の準備はできているか？これらのボットは、ソーシャルメディアネットワーク、電子メール、テキストメッセージ上で、超標的化され、超個人仕様にされたプロパガンダキャンペーンを展開している。実際の友人や家族、お気に入りの有名人からのものであるかのように見えるパーソナライズされたメッセージは、広告よりも影響力があるため、意見を変えさせ、行動を促すのに非常に効果的である。
〔https://www.stuff.co.nz/technology/digital-living/101551642/ai-manipulation-is-on-the-rise〕

シナリオ３：技術依存症 (tech addiction) に対する懸念の高まりに対応し、将来起こりうる出来事として政府の規制を予想して、最も人気のあるソーシャルメディアやゲーム会社のいくつかが**自発的に時間制限を実施する**ことを決定する世界の準備はできているか？一般的に、大人は１日２時間、子供は１プラットフォームにつき１日１時間までと制限されている。プラットフォームが制限されているか無制限かは、大きなセールスポイントになる。多くの人は、中毒になったり、あまりにも頻繁に気を散らされたりするのを防ぐ厳しい制限を好む。他の人々は、新しい無制限の競合他社に乗り換え、誰がどこで時間を過ごすか、そして彼らの精神的・肉体的健康かが科学技術の使用によってどのように

影響を受けるかという点で、新たな社会的な分裂が生まれている。
〔https://www.reuters.com/article/us-tencent-games/chinas-tencent-to-limit-play-time-of-top-grossing-game-for-children-idUSKBN19O0K0〕

経済と資産の不平等

（原文 pp. 19-20）

シナリオ４：自動化がかなりの数の仕事を排除した世界への準備はできているか？**2030年までに、自動化は7,300万人のアメリカ人の、人種集団により偏って配分された雇用を奪う可能性があり**、すでに疎外されたコミュニティをさらに危険にさらすことになる。ラテンアメリカ人労働者は、60%の仕事が廃止されるという最も高い脅威に直面し、続いてアフリカ系アメリカ人が50%、アジア人が40%近く、白人が約25%となっている。
〔https://www.theatlantic.com/education/archive/2017/12/the-new-casualties-of-automation/548948/〕

シナリオ５：フォーチュン500社[5]の人事部門がソーシャルメディアの投稿や「いいね！」を利用して職場の文化やストレスレベルへの適合度を評価する「スマート・エンプロイヤー」サービスに加入している世界の準備はできているか？アルゴリズムは、うつ病から社会性障害まで、さまざまな精神疾患を患っている可能性の高い個人を特定することができる。また、投稿の傾向から、近い将来、誰が精神疾患の症状を発症するかを予測する。このデータをもとに、現役社員へのサポートやリソースの提供、必要に応じた配置転換の推奨、採用の判断などに活用している。
〔http://bigthink.com/21st-century-spirituality/can-social-media-predict-depression-and-ptsd〕

機械倫理 (Machine Ethics) とアルゴリズム内のバイアス

（原文 pp. 21-22）

シナリオ６：「予測司法（Predictive justice）」ツールが実刑判決を下すための好ましい方法となる世界への準備はできているか？これらのツールは、何百万もの過去の事例からデータを取得し、類似した行動パターンや人口統計を持つ犯罪者の長期的な過去の再犯率を比較して、最も適切な判決を決定するために使用される。言い換えれば経済的、人種的、性別的、年齢的、行動的プロフィールが、投獄後に再犯したことがある人と似ている場合は、より長い実刑判決を受けることになる。これらの予測司法ツールは、異なる人口統計学的グループのメンバーが呼び止められ、逮捕、起訴、有罪判決を受ける頻度の構造的不平等を考慮していない。

[5] 「フォーチュン500社」とは米国の経済誌Fortuneが選ぶ米国企業売上高上位500社のこと。

〔https://www.vox.com/science-and-health/2017/4/17/15322378/how-artificial-intelligence-learns-how-to-be-racist〕

シナリオ７：大手ソーシャルネットワーク企業が米国のトップ銀行を買収し、初のソーシャルクレジットプロバイダーとなる世界の準備はできているか？それは、ソーシャル・プラットフォームによって収集された深いデータに基づいて、住宅ローンの金利、ローンの承認、クレジット・アクセスを行う。それは、親しい友人や家族の信用履歴、訪問した場所（バーや合法的なマリファナ自動販売機のような場所への訪問頻度を含む)、および個人が一般的に幸せか、怒っているのか、不安なのか、または落ち込んでいるかどうかを示すためにメッセージや写真の「意味論的分析」を考慮に入れる。
〔https://rexchange.com/〕

監視国家

（原文 pp. 23-24）

シナリオ８：**顔認証技術**が、個人や組織を問わず利用できるツールの主流となっている世界への準備はできているか？加入者は、何億人もの顔がインデックス化され、明確に認識できるデータベースを利用することができる。この技術を利用するために、ほとんどの公共空間やプライベート空間にカメラが設置され、出会い系アプリ、ショッピングアプリ、近所のアプリ、ゲームなど、あらゆる製品カテゴリーに顔認証を統合したアプリの新しいエコシステムが誕生している。
〔https://arstechnica.com/tech-policy/2018/05/police-use-of-amazons-face-recognition-service-draws-privacy-warnings/〕

シナリオ９：「読み書き可能な」**ニューロテック・インプラント**が現実のものとなる世界の準備はできているか？脳に埋め込まれた装置は、思考や感情を「読み取る」と同時に、新しい人工的な考えや記憶を頭の中に直接「書き込む」ことができる。最初に成功した読み書き可能なインプラントは、**脳の海馬に送られたデータの 80%**を傍受して記録し、ユーザーが直接読んだもの、見たもの、経験したものをほぼ完璧に思い出すことができる。この神経データはクラウドに保存され、他のインプラントに送信することができる。これらのインプラントを作成し、サービスを提供する企業は、潜在的な個人のリスクと新製品の広範な社会的影響を予測したいと考えている。
〔https://www.nanalyze.com/2016/11/brain-implants-ai-kernel/〕

データ管理と収益化

（原文 pp. 25-26）

シナリオ１０：大手データアグリゲータ[6]が、自宅にスマートトイレを設置してデータを提出することに同意すれば、誰でも無料で健康保険を提供してくれる世界の準備はできているか？**スマートトイレ**は、ストレスホルモン、妊娠、感染症、アルコールや薬物の使用、血糖値などを検知することができる。参加者が署名しなければならない契約書によると、これらのスマートトイレから収集したデータは、第三者に販売されたり、ターゲットを絞ったマーケティングに利用されたり、政府や科学研究者と共有されたりするなど、無制限に任意の目的に利用できるという。
〔https://qz.com/158774/70-of-people-would-be-willing-to-have-a-smart-toilet-share-their-personal-data/〕

シナリオ１１：主要大学で、学生が性的関係を持つ前にブロックチェーンを利用した**「同意の証明」**アプリをスマホにインストールすることを推奨している世界への準備できているか。ユーザーはコンドームの使用、性病の状態、そして写真撮影が許可されているかどうかなど多くのカテゴリーで好みを述べる。彼らはセックスをする前にお互いの好みを受け入れなければならない。このアプリは、同意した内容の変更不可能なデジタル記録を作成する。ユーザーは、自分の性的嗜好や遍歴についての情報が公開されるのを防ぐために仮名になっている。一部の人は、彼らのデータが盗まれ、彼らの実生活のアイデンティティにリンクされることを心配している。しかし、多くの人は、特に女性は、とにかくアプリを使用している。彼らは、プライバシーよりもレイプや非同意の行為を防ぐことに高い優先順位を置いている。
〔http://fortune.com/2018/01/16/consent-app-me-too/〕

暗黙の信頼とユーザーの理解

（原文 p. 27）

シナリオ１２：オンライン注文の 25%が**ドローンで配達される**世界の準備は出来ているか？これらのドローンの多くには、近隣の上空を飛ぶ際にデータを収集する為にカメラやその他のセンサーが搭載されており、運送主や小売業者に追加の収益源を提供している。無料で無制限のドローン配送を選択した個人は、自宅や庭からのデータ収集に同意することになる。ドローンの配送が法的に許可されている近所全体が、個々の居住者や世帯のすべてが明示的に同意していなくても、同じデータ収集活動の対象となる。
〔http://libguides.wustl.edu/drones4data〕

ヘイト行為者・犯罪者

（原文 pp. 28-29）

シナリオ１３：Venge と呼ばれる**ブロックチェーン**を利用した新しいプラットフォーム

[6] アグリゲータとは商品・サービスの情報を集めてウェブサイトに公開する会社のこと。

で、特定の個人への嫌がらせやテロ行為に対して匿名の懸賞金を設定することができる世界への準備はできているか？このような行為には、インターネット上での晒し、リベンジポルノ、"スワッティング"[7]、ソーシャルメディア上での嫌がらせ行為、破壊行為、暴力などが含まれる。ワンタップで、他の誰かに金を払って、自身の怒り、欲求不満、または憎しみを行動に移してもらうこともできる。スマートコントラクト[8]は、行動が完了したことを証明すると、追跡不可能な**暗号通貨**で即座に報奨金を支払う。インターネットが匿名で他人に嫌がらせをすることを容易にしたように、暗号通貨の場合は匿名性があり、規制がないため、ヘイト行為を行うために人にお金を支払うことが容易になった。〔http://www.ibtimes.com/could-online-harassment-slow-blockchains-growth-2562119〕

シナリオ１４：自動運転車が新型のリアルタイムのランサムウェアに脆弱になる世界への準備はできているか？ハッカーは車にリモートでアクセスし、エンジンを切り、運転手が身代金を支払うまで車の再始動を拒否する。一般的には、これは少額の金銭を支払うことで解決される、些細な、ありふれた不便なことである。最悪の場合、これらの攻撃は、危険な場所、例えば電車が近づく直前の線路上などでタイミングを計って実行され、より大きな身代金が要求されることになる。〔http://money.cnn.com/technology/our-driverless-future/keep-hackers-out-of-your-driverless-car/〕

[7] 緊急通報用電話番号を悪用し、何らかの大事件が起こっているとする虚偽の通報によって対象の元に警察官などを派遣させるという悪戯。対象の信用の毀損などを目的とする。
[8] 契約のスムーズな検証、執行、実行、交渉を意図したコンピュータプロトコルのこと。

Tool 2: 技術 (Tech) をチェックしよう——リスク低減マニュアル

（原文 pp. 30-57）

ほとんどの技術は、最善の意図を持って設計されている。しかし、製品がリリースされて大規模化してしまえば、すべての賭けは外れてしまう。リスク低減マニュアルでは、予測が困難で好ましくない結果が発生する可能性が高いと思われる 8 つのリスクゾーンを提示している。

覚えておこう：潜在的なリスクは時間の経過とともに進化していく。状況が変化しても、最も正確な現在のリスクゾーンを反映させるために、我々はこのセクションを継続的に更新していく。

Tool 2：何をすべきか

（原文 p. 32）

・取り組んでいる技術、製品、機能、または業界で最近注目を集めたものを選択する。
・8 つのゾーンのセクションを読む。
・これらのリスクがすでに発見され、低減されている実例のシグナルをチェックする。
・選択した技術に最も関連性の高いチェックリストの質問を特定する。
・識別したリスクの修正や低減をどのように始めるかを考える。

このマニュアルをダウンロードしてください

リスクゾーン 1　真実、故意の誤報 (Disinformation)、プロパガンダ

（原文 pp. 33-35）

共有された事実は攻撃を受けている。フェイクニュースから実在の人物を装いプロパガンダを拡散するボットに至るまでのあらゆるものによって。今、新しい危険な誤報の波が押し寄せている。「ディープフェイク」と呼ばれるこのような説得力のある動画は、アルゴリズムによって人々のスピーチや顔の表情を変え、置き換えたりし、そして実際には決して起きなかった行動やスピーチの偽の証拠を作成する。凄いね。

多くの個人やグループは、特に政治的な目的の為に、大規模に真実を捻じ曲げようとする強い動機を持っている。新しいソーシャルメディア技術により、嘘を広めたり、信頼を弱体化させたりすることがさらに容易になるだろう。

今後 10 年の間に、新しい技術によって他に何が偽造される可能性があるのか？私たちは、どのような共有された真実、事実、情報を守ることが求められるのか？

＜正の事例＞
ディープフェイクの問題に対するブロックチェーンによる解決策：
https://www.wired.com/story/the-blockchain-solution-to-our-deepfake-problems/

＜負の事例＞
フェイクビデオの時代が始まる：
https://www.theatlantic.com/magazine/archive/2018/05/realitys-end/556877/

（原文 p. 35）

・ユーザーはどのようなタイプのデータを正確に共有、測定、収集することを期待しているのか？
・悪質なアクターがあなたの技術を使って、どのようにして真実を覆したり、攻撃したりすることができるのか？あなたのプラットフォーム上で、フェイクニュース、ボット、ディープフェイクに相当するものは何になるのか？
・誰かがこの技術を使って、メディア、医療、民主主義、科学など、既存の社会的機関の信頼を損なうことができるか？あなたの技術は、政治的不信や社会不安を生み出すために、誤情報を生成したり拡散したりするのに使われる可能性があるか？
・そのような誤情報があなたのプラットフォーム上でどのような形態をとるか想像しなさい。あなたの技術が本質的に脱政治的であることが意図されていたとしても、政府を不安定化させるために利用される可能性があるかどうか？

リスクゾーン２　依存症とドーパミン経済

（原文 pp. 36-38）

コモンセンスメディアの調査によると、平均的な 10 代は 1 日に 9 時間、何かしらのメディアを使って過ごしている。9 時間！？

私たちがデバイスと一緒に過ごす時間については懸念が高まっている。ヒューマン・技術・センター(CHT)の創設者、トリスタン・ハリスは、テック企業に「充実した時間」を奨励するよう呼びかけ、設計者がプラットフォーム上で過ごす時間を最適化し、その時間が全体的な幸福と福祉に有益なものになるようにすることを提案している。

研究によると、人々は Instagram や Snapchat のようなアプリを 11 分後に最大の使用目的を達成するが、それ以上は全体的な幸福度が低下する。

充実した時間のためのツールはどのように設計されているのか？

画面に目を釘付けにすることよりも、オフラインでもオンラインでもユーザーの幸福を優先させるソフトウェアを設計するにはどうすればよいのか？

＜正の事例＞
中国のテンセントは未成年者向けの最も収益が高いゲームのプレイ時間を制限する：
https://www.reuters.com/article/us-tencent-games-idUSKBN19O0K0

＜負の事例＞
Instagram、Facebook、Tinder といったアプリが誘惑し、依存症にさせる卑劣な方法：
http://www.businessinsider.com/how-app-developers-keep-us-addicted-to-our-smartphones-2018-1#twitter-uses-a-psychological-trick-to-lure-you-in-the-same-one-used-in-slot-machines-3

（原文 p. 38）

・選択した技術の背後にあるビジネスモデルは、ユーザーの注目度と熱中を最大化すること、つまり、多ければ多いほど良いということで利益を得ているのか？もしそうだとしたら、それを使用する人々の精神的、肉体的、社会的な健康にとって良いことなのか？何が良くないのか？
・技術の「極端な」使用、中毒、または不健康な熱中とはどのようなものか？「適度な」使用や健康的な関与とはどのようなものか？
・適度な利用を奨励するシステムをどのように設計することができるか？適度な利用を促進することが、常に熱中の増加や最大化を求めるよりも、持続可能性や収益性の高いビジネスモデルを想像できるか？
・陰謀論やプロパガンダのような有害なコンテンツが熱中のレベルを高める可能性があるとしたら、そのようなコンテンツの普及を減らすためにどのようなステップを踏んでいるか？十分なのか？

リスクゾーン３　経済・財産の不平等

（原文 pp. 39-41）

オックスファム・インターナショナルによると 2017 年には 8 人の人が、世界の人口の下半分全体と同じくらいの富を所有していた。それを少し考えてみよう。

富の集中と分配は、近代史を通して問題となってきた。しかし、それはさらに悪化している。今日、米国では、富の集中は 1928 年以来最悪だ。新技術はアクセスを民主化し、所得機会を提供し、分配のバランスをとることができるが、高所得者層だけに迎合することで不平等を悪化させ、低所得者の仕事をなくすことにもなる。

＜正の事例＞
携帯電話の普及は貧困層への融資を可能とする：世界銀行

https://economictimes.indiatimes.com/news/international/business/widespread-mobile-phones-can-deliver-banking-to-the-poor-world-bank/articleshow/63834039.cms

より安価な自動車保険を手に入れる方法：白人であること
https://www.theatlantic.com/business/archive/2015/11/auto-insurance-race-discrimination/416988/

＜負の事例＞
二人の元グーグル社員はボデガ[10]を時代遅れにしたい。
https://www.fastcompany.com/40466047/two-ex-googlers-want-to-make-bodegas-and-mom-and-pop-corner-stores-obsolete

自動化の新たな犠牲者
https://www.theatlantic.com/education/archive/2017/12/the-new-casualties-of-automation/548948/

（原文 p. 41）

・誰がこの技術にアクセスできるのか、誰がアクセスできないのか？この技術を利用できない人やコミュニティは、利用できる人に比べて大打撃を受けるのか？その大打撃はどのようなものか？この技術を「持っている人」と「持っていない人」の間には、どのような新たな違いが生じるのか？
・あなたの技術は、どのような資産を創造し、収集し、普及させるか（例：健康データ、ギグ[11]、仮想通貨、ディープ AI）？この資産にアクセスできるのは誰か？誰がそれを収益化する能力を持っているか？
その資産（またはそこから得られる利益）は、その作成や収集を支援する他の関係者と公平に共有されているか、または分配されているか？
・人間の労働力ではなく、機械学習やロボットを使って富を生み出しているか？もし人間の雇用を減らしているのであれば、それは全体的な経済の幸福や社会の安定にどのような影響を与えるのか？あなたの会社や製品が、人々の雇用を通じてではないとしても、私たちの集団的な経済的安定に貢献できる他の方法はあるか？

リスクゾーン４　機械倫理とアルゴリズム内のバイアス

（原文 pp. 42-44）

[10] スペイン語圏で時に食料雑貨類も扱うワインショップ。
[11] ギグとは、デジタル化を背景とした単発労働のこと。

福祉、教育、雇用、刑事司法などの重要な領域での AI の応用が強化されているにつれて、AI が偏見を反映し、導入し、さらには増幅する方法についても懸念がある。

人種的に偏った警察の顔認識システムであれ、特定の思想やアイデンティティを特権化する検索エンジンであれ、技術自体が中立ではないことを知るには十分な証拠がある。それは、人間は本質的に中立的ではなく、そしてはしばしば技術が人間の行動の産物、もしくは人間が使用するツールであるからだ。

上記のような課題を解決するために AI に頼る解決策は、AI 自体が人間によって創造され、使用されるツールであるという事実を見落としがちである。AI 技術の設計、使用、およびガバナンスへの学際的なアプローチは、特定の個人やコミュニティへの意図的しない、または意図的な差別的な影響の一部に対処するために重要であることに変わりはない。

＜正の事例＞
キャピタル・ワン、モデルの偏りを防ぐ「説明可能な AI」を追求：
https://blogs.wsj.com/cio/2016/12/06/capital-one-pursues-explainable-ai-to-guard-against-bias-in-models/

＜負の事例＞
シリコンバレーは途方に暮れている：
https://www.cnbc.com/2018/05/30/silicon-valley-is-stumped-even-a-i-cannot-remove-bias-from-hiring.html

Podcast とても現実的な偏見を持った人工知能：
https://www.wsj.com/podcasts/artificial-intelligence-with-very-real-biases/B53751ED-000E-410E-A727-249A102995C1.html

アルゴリズムは偏見を持ってないが、それを記述する人々は持っているかもしれない：
https://www.wsj.com/articles/algorithms-arent-biased-but-the-people-who-write-them-may-be-1476466555

（原文 p. 44）

・この技術は、ディープデータセットや機械学習を活用しているか？もしそうだとしたら、技術に偏見を植え付ける可能性のあるデータのギャップや過去のバイアスはあるか？
・製品のアルゴリズムに私的または個人的な偏見が入っている事例を見たことがあるか？これらはどのようにして防ぐことができたか、または低減することができたか？
・技術は既存のバイアスを強化したり、増幅したりしていないか？
・アルゴリズムの開発は誰が担当しているか？技術の設計を担当する人々に多様性が欠け

ていないか？
・自動化への盲目的な選好（AI ベースのシステムや意思決定は正しく、検証や監査の必要がないという思い込み）に対して、どのようにして対抗するのか？
・アルゴリズムは、その影響を受ける人々に透明性があるか？自分たちの評価が間違っている、不当に評価されていると感じている人たちのための救済手段はあるか？

リスクゾーン５　監視国家

（原文 pp. 45-47）

ソーシャルボットが政府や軍によって野党を攻撃するために利用されている最近の例は、良心的な用途のために作られた技術が害をもたらすために利用される可能性があることを明らかにしている。いくつかの事例では、自動化されたソフトウェア駆動のツイッターやフェイスブックのプロフィールの大群が、特定の考えに反対するジャーナリストや活動家、市民を標的にするために使用されていた。

権威主義国家では、西側の監視ツールが国家による弾圧を後押しし、指導者が電子メールやテキストメッセージを傍受したり、携帯電話を通じて市民の居場所を監視したりすることを可能にしている。これらの国々の多くの警察は、この情報で武装し、現在では日常的に、逮捕や拷問の際のメッセージや動きの記録を持って反体制派と対峙している。一方で、顔認識やソーシャルメディアによるトラッキングの利用が増えているため、政府は個人の行動を深く記録することができるようになっている。"何年にもわたるデータ収集の中で生成された「市民スコア」は、公共空間へのアクセスや仕事の機会などを制限するために利用され、以前の行動や言動に基づく新しい種類の社会的不平等を生み出すことができる。

＜正の事例＞
英国、オンラインで ISIS のプロパガンダと戦うための機械学習技術を公開：
https://www.cnbc.com/2018/02/13/uk-unveils-machine-learning-technology-to-fight-isis-propaganda-online.html

＜負の事例＞
中国の監視国家は誰もが怖がるべき：
https://www.theatlantic.com/international/archive/2018/02/china-surveillance/552203/

・政府や軍の機関は、監視能力を高めたり、市民の権利を侵害したりするために、この技術をどのように利用するのだろうか？
・政府はこのデータへのアクセスを許可された場合、あるいは法的にアクセスを要求した

り、提出を命じたりした場合、このデータで何をすることができるのか？
・政府や軍以外の誰が、あなたが作成しているツールやデータを使って、標的となる個人の監視を強化する可能性があるのか？彼らは誰を追跡するのか、なぜ追跡するのか、そして、あなたはあなたの技術をこのような方法で使用されたいと考えているか？
・生涯にわたってユーザーを追跡し、評判に影響を与え、将来の機会に影響を与えるようなデータを作成していないか？あなたの技術が生成しているデータは、個人の自由や評判に長期的な影響を与えるか？
・個人を監視し、意思決定をするためにデータを利用してほしくないとあなたが思うのは誰か？このデータにアクセスできないようにするにはどうすればいいか？

リスクゾーン６　データ管理と収益化

（原文 pp. 49-51）

将来的には、利用者は、自分自身とその周囲の状況について収集された情報を取得し、共有し、解釈し、検証するためのツールへのアクセスを期待するだろう。利用者は、個人データや組織データが開示されることで自分の幸福が危うくなる場合には、自分のプライバシーを選択的に保護する基本的な権利を主張することになるだろう。そして、個人データへのアクセスをコントロールし、データを作成・利用するテック企業と一緒に、選択的に共有し、収益化し、利益を得る機会を持つことを期待するようになる。

<正の事例>
グーグルデータエクスポート　個人データをダウンロードする為のグーグルのツール：
https://takeout.google.com/settings/takeout?pli=1

ソーシャルネットワークは何百万ものあぶく銭を配分している：
https://www.wired.com/story/the-social-network-doling-out-millions-in-ephemeral-money/

<負の事例>
ケンブリッジ・アナリティカとフェイスブック。スキャンダルと今までの好ましい結果：
https://www.nytimes.com/2018/04/04/us/politics/cambridge-analytica-scandal-fallout.html

Google の私に関するファイルは巨大だった。これは、それが私の Facebook のデータのように不気味ではなかった理由である。:
https://www.nytimes.com/2018/05/16/technology/personaltech/google-personal-data-facebook.html

（原文 p. 50）

・どんなデータをユーザーから収集しているのか？本当に収集する必要があるのか？それを販売しているのか？販売している場合、誰に販売しているのか、ユーザーはそれを知っているのか？どうすればより透明性を高められるのか？
・ユーザーは、収集されたデータにアクセスする権利と能力を持っているか？もしそうでない場合、どのようにしてより良くユーザーをサポートすれば、ユーザーはあなたが彼らについて既に知っていることを簡単かつ透明に知ることができるか？
・ユーザーデータの使用や販売から利益を得る場合、ユーザーはその利益を共有するか？ユーザーが自分のデータについて利益を共有する権利をユーザーに与えるために、どのようなオプションを検討するか？
・ユーザーが自分のデータを独立して共有し、収益化する権利を与える方法を構築できるか？
・もし悪質なアクターがこのデータにアクセスできた場合、どのようなことができるか？もしデータが盗まれたり漏洩したりした場合、誰かがこのデータを使ってできる最悪のことは何か？
・会社が買収や売却、もしくは閉鎖された場合に顧客データがどうなるかについての機能するポリシーを持っているか？

リスクゾーン７　暗黙の信頼とユーザーの理解

（原文 pp. 51-53）

データの悪用は深刻な問題である。しかし、それと同じくらい問題なのは、多くのユーザーが人気のあるアプリやプラットフォームがどのように機能しているのか、ユーザーのエンゲージメントがどのように最適化されているのか、何が追跡されて収集されているのかを十分に認識していないという事実である。利用規約はほとんど読まれておらず、明確に書かれていることはほとんどない。多くのアクティビティに対する許可は、どのように使用されるかを十分に理解していないまま与えられることが多い。

将来的には、自社の製品や従業員が、ユーザーが同意した内容の暗黙的または明示的な前提や範囲に違反した場合、企業は反発に直面する可能性が高くなる。特に、その同意が、ユーザーが実際には誰も読まない高密度で解読不能な同意を受け入れることに基づいている場合はなおさらだ。

<正の事例>
ティム・クック氏、Apple が FBI のための iPhone のロック解除を拒否したことは「市

民的自由」の問題だと発言：
https://www.theguardian.com/technology/2016/feb/22/tim-cook-apple-refusal-unlock-iphone-fbi-civil-liberties

Snapchat の Evan Spiegel 氏が Facebook にデータ保護方針のコピーを提案：
http://fortune.com/2018/05/30/snapchat-evan-spiegel-facebook-data/

<負の事例>
Uber は「God View」と呼ばれる内部ツールでジャーナリストを追跡していたとされる：
https://www.theverge.com/2014/11/19/7245447/uber-allegedly-tracked-journalist-with-internal-tool-called-god-view

報告書によると一部のアプリはスマホのマイクを通してテレビ視聴を追跡していた：
https://techcrunch.com/2018/01/02/some-apps-were-listening-to-you-through-the-smartphones-mic-says-report/

（原文 p. 53）

・あなたが構築している技術は、ユーザーのための明確な権利規約を持っているか？利用規約は読みやすく、アクセスしやすく、理解しやすいものか？
・ユーザーがユーザー規約に署名したくない場合に使用できる製品のバージョンがあるか？
・あなたの技術は、ユーザーが知らないこと、またはそれについて分かったときにユーザーが恐らく驚くようなことをしているか？もしそうだとしたら、なぜその情報を明示的に共有しないのか。また、ユーザーがそれを知った場合、どのような反発を受ける可能性があるか。
・ユーザーが自分の行動が収益化されたり、自分のデータが特定のグループや組織に売られたりすることに反対し、それでもプラットフォームを使いたいと思った場合、ユーザーにはどのような選択肢があるのか？信頼を構築し、今後のビジネスモデルのさまざまな側面からユーザーが加入しまたは離脱できるような代替モデルを作成することは可能か？
・すべてのユーザーは平等に扱われているか？もしそうではなく、アルゴリズムや予測技術が特定の情報に優先順位をつけたり、価格を設定したり、異なるユーザーに異なるアクセスを提供したりしている場合、すべてのユーザーを平等に、あるいは少なくとも**透明性をもって**不平等に扱うことを要求する消費者の要求や政府の規制にどのように対処するか？

リスクゾーン８　ヘイト行為者・犯罪者

（原文 pp. 55-57）

2017 年 8 月 12 日、バージニア州シャーロッツビルで "Unite the Right "と呼ばれる集会[12]で白人至上主義者グループと繋がっていた男がデモに抗議する群衆に車を衝突させ、ヘザー・ヘイアーさんが殺害された。集会はフェイスブックのイベントツールを使って組織されていた。明らかにヘイト犯罪は新しいものではない。しかし、オンラインツールは、ヘイトを広めたい人にとっては、これまでにないほどの簡単さとスピードで、コンテンツの世界的な普及を可能にしている。いじめ、過激化、荒らし、晒し、その他の悪質な行動は、技術のおかげでブームになっている。そして、そのような環境の中で多くの時間を過ごしているユーザーにとって、ネガティブな体験を低減したり、防止したりすることは事実上不可能だ。

<正の事例>
研究者がツイッターボットを良い方向に利用することができることを発見
https://techcrunch.com/2017/09/25/researchers-find-that-twitter-bots-can-be-used-for-good/

Airbnb、シャーロッツビルでの白人愛国主義者の集会と関連したアカウントを取り消す：
https://techcrunch.com/2017/09/25/researchers-find-that-twitter-bots-can-be-used-for-good/

<負の事例>
大規模なフィンテックから切り離された白人愛国主義者がビットコインを使って資金調達をしている：
https://www.forbes.com/sites/janetwburns/2018/01/03/cut-off-from-big-fintech-white-supremacists-are-using-bitcoin-to-raise-funds/#737748e333b3

如何にしてイスラム国が世界で最も致命的なハイテクスタートアップとなったか：
https://www.vanityfair.com/news/2016/06/how-isis-became-the-worlds-deadliest-tech-start-up

（原文 p. 56）

・誰かがあなたの技術を使って、他人をいじめたり、ストーカー行為をしたり、嫌がらせしたりする可能性はあるか？

[12] 極右集団の集会。

・どのような新しい種類のランサムウェア、盗難、金融犯罪、詐欺、またはその他の違法行為が、あなたの技術の中で、またはその周辺で潜在的に発生する可能性があるか？
・技術メーカーには、悪質な行為者が行動しにくくする倫理的責任があるのか?
・組織的な（人種差別などの）扇動グループが、あなたの技術を使って憎悪を広めたり、勧誘したり、他人を差別したりする可能性はあるか? 組織的なヘイトは、あなたのプラットフォームやコミュニティ、ユーザーでどのようなものであるか？
・あなたの技術が武器にされるリスクはどのようなものか？これを防ぐためにどのような責任があるか？技術の兵器化を防ぐために、規制や国際条約の作成にどのように取り組んでいるか？

倫理的になろう――リスクに気づいているならば、今が行動の時だ――

（原文 p. 57）

これでシナリオを考え、潜在的なリスクの主な領域について熟知した。さて、次のステップに進もう、つまり、より倫理的なオペレーティングシステムの設計と実装を始める。以下に、始めるためのいくつかの方法を紹介する。

・ハイライトしたリスクゾーンと質問をチーム内で共有する。新製品を開発したり、新機能を検討したりする際に、これらを念頭に置いておくように個人に働きかける。

・新技術に積極的に取り組んでいる場合は、チームが製造中のものの潜在的なリスクを整理できるように、製品要求仕様書（PRD）に**最上位の質問を追加することを検討する**。

・ホライゾンスキャニング[13]を行って、これらのリスクゾーンに関する追加情報を得る。チームメンバーを招待して、これらの**リスクに関連するニュース記事、オプ・エド[14]、またはその他の「シグナル」にフラグを立てて回覧し**、倫理的な問題を業務に組み込んでおく。

[13] 将来、社会に大きな影響をもたらす可能性のある変化の兆候をいち早く捉えるために、利用可能な情報を体系的・継続的に収集・分析し、潜在的なリスクや可能性を把握する活動を指す。
[14] “opposite editorial”の略で、新聞の記事のうち通常、当該紙の編集委員会の支配下にない外部の人物が、ある新聞記事に対して同じ新聞内で意見や見解（反論や異論）を述べる欄のこと。

Tool 3: 将来に備える――今後、学んだこととベストプラクティスを組み込む――

（pp. 58-68）

世界をより良い場所にする（もしくは少なくとも他者が世界をより悪い場所にするのを止めることを助ける）準備はできているか？ならば、私たちの最善の意図を実行可能なセーフガードに変える時が来たのだ。

技術コミュニティが**大規模に**リスクを低減するのに役立つベストプラクティスは何か？企業の利益と人類の利益の両方を考慮した製品を作るのに役立つ業界全体の努力はどのようなものか？

以下は、あなたとあなたのチームのために行動を促し、インフラストラクチャを追加するための**戦略**とアイデアである。

Tool 3　何をすべきか

（原文 p. 60）

・将来に備える戦略を**選択する**。
・利点を**検討する**。それは何を達成することができるか。
・潜在的な欠点を**発見する**。それはどのように誤った方向に進むか。それが現実になるのを妨げるものは何か？

もし、あなたのチームや会社が追求したい戦略があるとしたら、それを実行に移すのに役立つ可能性のある**リソース、協力者、次のステップのリストを作成しよう**。

戦略 1　技術倫理 101

（原文 p. 61）

将来的には、トップコンピュータサイエンス、デザイン、およびエンジニアリングプログラムは、すべての学生が技術倫理のコースまたはトレーニングシーケンスを完了するという要件を採用していることを想像してみよう。

・どのようなスキルを教えるべきか？
・学生はどのように倫理的なスキルセットで評価されるのか?
・この変更を最初に採用するのはどの学校か?

・誰がカリキュラムや教科書を作成するか？
・これらの原則を反映させるために、あなたの毎日の立ち話に何を加えることができるか？

戦略２　データワーカーのためのヒポクラテスの誓い

（原文 p. 62）

ユーザーデータを扱う人が、倫理的にデータを取得、使用、共有することを誓った場合を想像しよう。
・誓いにはどのような具体的な約束が含まれるべきか？
・大企業や投資家は、雇用や資金調達の条件として個人に誓いを立てることを求めることができるか？
・「誓いを立てる」ことをより意味のあるものにするには、どのような形式があるか？
・どのような他の専門分野で独自の誓いを作成することができるか？例えば、機械学習に携わる人々のためのヒポクラテスの誓いはどのようなものか？
・チームに誓いを立ててもらいましょう！
https://github.com/Data4Democracy/ethics-resources

戦略３　倫理的な賞金稼ぎ

（原文 p. 63）

新サービスや技術の潜在的な大きな社会的リスクを特定するために、テック企業が「報奨金」を支払うようになったらどうか？今日のハッカーがセキュリティの欠陥や脆弱性を特定することで報酬を得る方法と同様に、「メンタルヘルスへの影響」、「民主主義へのリスク」、そして「構造的な人種差別、性差別、その他の不平等」のような様々な領域でも報奨金を設定できるかもしれない。
・どうやって報奨金を請求するのか？どのような証拠が必要なのか？
・識別されたリスクを評価する責任は、企業内の誰にあるのか？
・妥当な報奨金はどのくらいか？
・次のアドレスに回答をお送りください： ideas@ethicalOS.org

戦略４　レッドフラッグ[15]ルール

（原文 p. 64）

[15] この場合の「赤旗」とは警戒を表す信号を意味する。

将来的には、テック企業や投資家は、注意すべき社会的・倫理的な「レッドフラッグ」のリストを従業員に提供することになるだろう。すべての従業員は、報復を恐れずにリスクを社内に報告するための明確な経路を持つことになる。

・レッドフラッグリストにはどのようなリスクがあるのか？
・誰が作成し、更新するか？
・リストは会社やポートフォリオ全体で共有されるか？リストは公開されるのか、専有されるのか？
・内部レッドフラッグ報告プロセスはどのようなものか？自分のチームでの報告はどのように処理されるか？

戦略５　健全なプラットフォーム

（原文 p. 65）

新技術プラットフォームのクリエイターは、プラットフォームの「健全性」に関する透明性の高い尺度を確立し、共有することが期待されている。Twitter はすでに、「会話の健全性」の４つの新しい尺度を測定し、最大化することを約束している。このような健全性の指標は、プラットフォームがユーザーや社会のために良いことをしているかどうかを評価するのに役立つだろう。

・どの健全性の指標を確保したいと思うか？
・何をこの健全性の重要な構成要素と定義するか？
・これらの構成要素をどのように測定するか？
・健全性や健全性のリスクをどのようにユーザーや一般の人に伝えるか？
・健全性のためにこれらの測定基準をモニターするのは、あなたのチームの誰の仕事か?

戦略６　デザインするライセンス

（原文 p.66）

すべての技術デザイナーや開発者が、彼らの行動を規制し、無責任な設計や非倫理的な設計を禁止し、今日の技術界に蔓延している「素早く行動し破壊せよ」[17]という文化を打ち消す管理機関からライセンスを取得したとしたらどうなるかを想像してください。これは、医師、弁護士、建築家のライセンス制度に似ている。

・管理機関は誰が構成するのか？
・ライセンスを取得するための要件には何が含まれるか？
・どのような行為がライセンスの喪失につながるのか？
・デザイナーや開発者はライセンシングへの移行で利益を得るか? 企業は利益を得るか?

[17] Facebook のザッカーバーグの格言。

・ライセンシングの潜在的な負の影響、すなわち、より大きな不平等と、誰が仕事を得るかという点での多様性の低下、雇用プールの縮小、またはイノベーションの鈍化につながることをどのように緩和するか？
・次のアドレスに意見ををお送りください：ideas@ethicalOS.org

Tool3　将来に備えた戦略

（原文 p. 67）

「私たちは、現在成し遂げている成長や規模だけではなく、私たちの技術が生み出す長期的な未来について考えなければならない。これを怠ると、これらの技術を使用するすべての人が失敗する。最悪の場合、これまでにない規模で民主主義と平等を失うことになる。」

—Institute for the Future, Digital Intelligence Lab ディレクター、サム・ウーリー

他に何をデザインする？

（原文 p. 68）

これらの倫理的なインフラへの介入を検討したところで、長期的な回復力を構築し、技術コミュニティが企業と人類の両方の利益を考慮した製品を作るのを支援するために、他にどのような取り組みを設計したいと思うか？
あなたの考えを共有しよう：

#ethicalOS

ideas@ethicalOS.org

コミュニティからの想いを込めて、エシカル・オペレーティング・システムの更新を続けていく。

オーダーメイドの使用例

（原文 pp. 69-75）

次のセクションでは、特定のチーム、役員会メンバー、または学生がトピックに関する最初の会話に参加できるように、クリフノート[18]のバージョンをご紹介する（これが彼らの最後の会話ではないことは確かだが）。

倫理的な OS のクリフノート

別名:「このツールキットから 1 ページだけ読むなら、これだ。」
どんな仲間とも共有してください（必要ないと思っている人でも）。

（原文 p. 70）

検討事項

・あなたの製品、サービス、ビジネスは、どのようにして、あなたが見たい世界に最も**ポジティブな影響**を与えることができるか？
あなたが解決しようとしている問題を超えて、このイノベーションを導入することで世界はどのように良くなるか？
・あなたの製品や製品の機能は、誰かに直接害を与える可能性があるか？一次の影響だけでなく、二次、三次の影響も考えてみてください。
・悪意のあるアクターがあなたの製品や製品の機能を使って、個人やグループに危害を加える可能性はあるか？罪を犯す可能性はあるか？
・あなたの製品の背後にあるビジネスモデルが、誰か、または集団に危害を加える可能性はあるか？中毒を助長する可能性があるか？特定のグループ（子供、高齢者、マイノリティグループ、市民権を奪われた人、権限を奪われた人など）に異なる方法で影響を与える可能性がある。
・あなたの製品が誰かのプライバシーを危険にさらす可能性はあるか？誰かのデータが漏洩したり、ハッキングされたり、紛失したりしたらどうなるか？このような不測の事態に備えた計画を立てているか？ユーザーとの信頼関係を築くにはどうしたらいいか？
・あなたの製品やビジネスは、ユーザーが気づかないようなことをする可能性があるか？もしそうなら、なぜその情報を明示的に共有していないのか？明日のニュースに出てきたら、あなたのビジネスにリスクをもたらすか？
・あなたの会社がベスト・バージョンの自身を維持するために、役員会として何ができるのか、あるいは何をすべきなのか？

詳細と倫理的な OS ツールキットの全体については、以下を参照してください：
ethicalOS.org

[18] 原義は、「授業の要点のみをまとめた黄色い本」のこと。

あなたは理事や役員か？

（原文 p. 71）

風評リスクから従業員の流出に至るまでの脅威によって、倫理的な過失は様々な形で株主価値に悪影響を及ぼす可能性があり、これはオフィスでの不正行為だけに留まらない。以下の会話の起点は、被害が出る前に潜在的な製品リスクに先回りすることを目的としている。会議の時間を 5 分ほど取って、これらの問題について考え抜いたことを確認してください。

・あなたの製品やビジネスは、ユーザーが**知らない**ことをできるか？なぜこの情報を明示的に共有しないのか？それが明日のニュースに出てきたら、あなたのビジネスに**リスクをもたらすか**？
・あなたの利用規約はどのくらい透明で明確か？分かりやすくするために何か変更できることはあるか？
・悪意のある行為者が貴社の製品を使用して危害を加え、または罪を犯す可能性はあるか？このような事態が発生した場合、どのような緩和策を講じているか？
・あなたのビジネスモデル自体が悪意を持って使われる可能性はないか？中毒を助長する可能性は？特定のグループ（子供、高齢者、マイノリティグループ、権利を奪われた人、権限を奪われた人など）に異なる方法で影響を与える可能性はないか？
・あなたの製品はユーザーのプライバシーを侵害する可能性があるか？誰かのデータが漏洩したり、ハッキングされたり、紛失したりしたらどうなるか？このような不測の事態に備えた計画を立てているか？ユーザーとの信頼関係を築くにはどうしたらよいか？
・あなたの会社がベスト・バージョンの会社であることを保証するために、取締役会として何ができるか、または何をすべきか？
詳細とエシカル OS ツールキットの全体については、以下を参照してください：
ethicalOS.org

製品を製造しているか？

（原文 p. 72）

これから行く先々で覚えておくべきこと

・製品、サービス、またはプラットフォームの第一次の影響だけでなく、第二次、第三次の影響も評価しているか？
・多様なユーザーがアクセスできるように、製品やサービスを設計しているか？
・製品を運営し、経済的な存続性を確保するために絶対的に必要なユーザーデータのみを

収集しているか？
・トレーニングデータやデータセットが多様なユーザーを代表するものであることを確認し、そのデータのソースにおける潜在的なバイアスを最小限に抑えているか？
・あなたの製品が真実のコンテンツや情報のみを確実に広めるための最も効果的な方法を特定したか？
・製品やサービスに組み込まれているアルゴリズムや機械学習プロセスの公平性、説明責任、透明性を確保しているか？
・ユーザーのために適切なオフランプを設計しているか？底なし穴、無限スクロール、注意力の罠を可能な限り排除しているか？
・あなたのユーザーエクスペリエンスで暗いパターンの使用を避ける方法を見つけたか？
・もしあるなら、通知はユーザーの邪魔になるほど重要なものか？音や感覚、微妙なステータスの変化などを利用して、より邪魔にならない方法で通知できるか？
・ユーザーが作成したコンテンツの量よりも質を重視したシステムを構築しているか？
・データ漏洩に備えた緊急時対応計画を作成しているか？そのような侵害が発生する前、発生中、発生後にユーザーの信頼を得るための方法を見つけたか？

テック製品とデザインガイド（続き）

（原文 p. 73）

発送する為に確認ください

・利用規約を明確に、簡潔に、わかりやすい言葉で書いているか？
・年齢、性別、人種、社会経済的地位と収入、地理、政治的所属、言語、能力、性的指向、宗教、教育の多様性を代表する多様なユーザーセットで製品をテストしたことがあるか？
・予測される意図しない結果を考慮して再考すべき、収益や成長への道筋を捨てた可能性のあるものを再検討したことがあるか？
・あなたの製品をレッドチーム化[19]して、悪質または悪意のある行為者（個人、グループ、または組織）があなたの製品をどのように武器にするかを評価したことがあるか？
・上記のいずれか、またはすべてについて、暗黙の仮定を自分自身で確認したことがあるか？
・上記のすべてを考慮して、世界をより良い場所にすると確信できる製品を設計したか？
詳細と倫理的な OS ツールキットの全体については、以下を参照してください：
ethicalOS.org

[19] レッドチームとは元来、セキュリティの専門家が攻撃チームを作り、顧客企業に対して攻撃を行うことで、企業のセキュリティ対策の実効性を検証するサービスのこと。

コンピュータサイエンスやデザインのクラスを教えているか?

（原文 p. 74）

ツール 1: **生徒に 14 の「危険な未来」のシナリオをすべて読ませ、自分の想像力をかきたてるものを 1 つ選ばせる**。倫理的なリスクや害がその未来に起こるかもしれないものをブレインストーミングするように彼らにいう。生徒は他の生徒とチームで作業することができる。同じシナリオに興味を持っている人 これはクラスディスカッションとしても可能である。

ツール 2:
・生徒に 8 つのリスクゾーンを読んでもらい、興味をひかれるものを 1 つ選ぶように指示する。そのリスクゾーンからのシグナルを収集するためにそれらに挑戦しなさい。シグナルとは、現在すでに起こっていることの実例であり、以下のようなことが考えられる。未来に影響を与えたり、形作ったりする。シグナルは、ニュース、ブログ、ソーシャルメディア、科学雑誌、技術情報などで見つけることができる。会議、TED 講演、研究室など、人々や企業が新製品、アイデア、発見を共有している場所はどこにでもある。生徒は、同じリスクゾーンに興味を持っている他の人とチームを組んで作業することができる。

・生徒に、自分が興味を持っている実際の新技術、製品、アプリを選ぶように指示する。リスク低減の質問に目を通し、その技術に最も関連性があると思われるリスクゾーンの質問にチェックを入れさせる。次に、それらの質問の中から一つを選び、その技術をより倫理的でリスクの少ないものにする方法を探しながら、その質問に答えようとする。

ツール 3: **6 つの戦略のうち 1 つ以上についてクラス討論を行う**。生徒に自分の好きなストラテジーを学んだ後、クラスのトップ 1 つまたは 2 つのストラテジーが将来どのように展開されるのかを深く想像する時間を持つ。

詳細と倫理的な OS ツールキットの全体については、以下を参照してください：
ethicalOS.org

あなたのネットワークとの対話を開始する

（原文 p. 75）

技術の未来はすべての人の参画を必要としており、私たちはすべての視点を必要としている。ここでは、EthicalOS を使って Facebook、LinkedIn、Twitter など、あなたのネッ

トワークを活性化させるためのアイデアをいくつか紹介する。

ツール 1: 14 の「危険な未来」シナリオのうちの 1 つをあなたのソーシャルメディアネットワークに投稿してください。シナリオのシグナルへのリンクを含めて、人々がそれに触発された実例を調べることができるようにする。あなたのネットワークに尋ねなさい。このシナリオで何が起こることを心配するか？そして、あなたはそれについて何をすることを提案するか？

ツール 2: 8 つのリスクゾーンのいずれかを選択する。あなたのネットワークと記述および信号を共有しなさい。それらを尋ねなさい。この危険はあなたが取り組む技術で心配する何かであるか。あなたが使用する技術についての何か。あなたはすでにこれらのリスクのいずれかの影響を受けたことがあるか？現実の世界でこのリスクの他の例を見たことがあるか？

ツール 3: 未来を証明する作戦のうちの 2 つを共有して下さい。それらがより有効であると思う 2 つのうちのどれがなぜあなたのネットワークに尋ねなさい。
・**自由に共有**しなさい: 自由にデッキのスライドのスクリーンショットを取るか、またはあなたのネットワークと共有するために素材を引用しなさい。
・**ハッシュタグ #EthicalOS を使用**しなさい。
・EthicalOS.org に**リンクして**、ネットワークが完全な EthicalOS ツールキットをダウンロードできるようにする。

詳細と倫理的な OS ツールキットの全体については、以下を参照してください：
ethicalOS.org

倫理的な OS について

（原文 p. 76）

倫理的な OS は、Institute for the Future と Omidyar Network の Tech and Society Solutions Lab が共同で制作したものである。

・このツールキットを使ったイベントやワークショップの開催や、組織内でのトレーニングセッションの開催をご希望の方は、ideas@ethicalos.org までご連絡ください。
・このツールキットに関するフィードバックや、追加のリスクゾーンやシナリオの提案があれば、ideas@ethicalos.org までご連絡ください。

Institute for the Future について

（原文 p. 77）

Institute for the Future (IFTF)は、世界をリードする非営利の戦略的未来機関として、創立 50 周年を迎えている。私たちの仕事の中心は、グローバル社会とグローバル市場を変革する新たな不連続性を特定することだ。私たちは、ビジネス戦略、デザインプロセス、イノベーション、社会的ジレンマについての洞察を組織に提供している。私たちの研究は、健康とヘルスケアから技術、職場、人間のアイデンティティに至るまで、深く変容するトレンドの幅広い領域にわたっている。IFTF はカリフォルニア州パロアルトを拠点としている。詳細は www.iftf.org。
をご覧ください。倫理的 OS のリサーチリード。倫理的 OS のリサーチリーダー：Jane McGonigal、IFTFF リサーチ＆コラボレーティブ・フォーサイト・ディレクター。
IFTF Governance Futures Lab ディレクター David Evan Harris 氏、IFTF エグゼクティブ・プロデューサー Jean Hagan 氏、IFTF デジタル・インテリジェンス・ラボ・ディレクター Sam Woolley 氏に感謝の意を表す。

技術社会ソリューション研究所について

Omidyar Network の Tech and Society Solutions Lab は、技術者が技術の社会的なマイナス面を予防、緩和、修正し、プラスの影響を最大化するのを支援することを目的としている。このチームは、これらの目標を推進するためのソリューションを共同で創造し、支援している。これらの取り組みが、中核となるビジネスや製品の意思決定に責任を組み込み、繁栄するテック産業に貢献することを願っている。詳細はこちら：
Tech and Society Solutions Lab https://www.omidyar.com/our-work/tech-and-society-solutions-lab

謝辞
倫理的 OS のプロジェクト・リード レイナ・クムラ（テック＆ソサエティ・ソリューションズ・ラボ、アントレプレナー・イン・レジデンス)、エシャンティ・ラナシンゲ（知的資本担当マネージャー)、ヨアヴ・シュレジンガー（Yoav Schlesinger）のサポートを受けている。技術と社会のソリューション・ラボ チーフ・スタッフ、知的資本担当アソシエイト ランドルフ・ウィギンス氏