< >
モントリオール宣言
責任あるAI_
</ >

人工知能の責任ある開発に関するモントリオール宣言

この文書は、「人工知能の責任ある開発に関するモントリオール宣言 2018」の一部です。
完全なレポートはこちらでご覧いただけます。

目次

宣言の読み方 3
序文 5

原則

1. ウェルビーイングの原則 6
2. 自律尊重の原則 7
3. プライバシーおよび内密性保護の原則 8
4. 連帯の原則 9
5. 民主的参加の原則 10
6. 公平・公正の原則 11
7. ダイバーシティ・インクルージョンの原則 12
8. 慎重さの原則 13
9. 責任の原則 14
10. 持続可能な開発の原則 15

用語集 16

クレジット 19
パートナー 20

宣言の読み方

何のための宣言か?

人工知能の責任ある開発に関するモントリオール宣言には、3 つの主な目的があります:

1. AI の開発および展開のための倫理的枠組みを開発すること
2. 誰もがこの技術革命の恩恵を受けられるようにデジタル移行を導くこと
3. 国内および国際的な対話フォーラム開催を通じ全体で、公平で、包摂的で、生態学的に持続可能な AI 開発を実現すること

何を宣言しているのか?

原則

本宣言の第一の目的は、人々や集団本来の利益につながる倫理原則や価値観を明らかにすることです。デジタルおよび人工知能の分野に適用されるこれらの原則は、依然として一般的で抽象的なものです。その正しい理解には、次の点に注意することが重要です :

> 原則がリストの形で提示されていますが、これらの中に順位づけはありません。最後の原則が最初の原則より重要度が低いということはありません。ただし、状況によっては、ある原則を他の原則よりも重視したり、関連性が高いと見なしたりすることはあります。

> これらの原則は多岐にわたるものですが、それぞれの適用を妨げるような矛盾が生じないように、整合的なものとして解釈する必要があります。

一般論としては、ある原則の適用の範囲は、別の原則の適用範囲によって定義されます。

> 原則はそれらを策定した社会の道徳的および政治的文化を反映していますが、異文化間および国際社会での対話のベースとなるものです。

> これらの原則はさまざまな解釈ができますが、どのようにでも解釈できるわけではありません。解釈に一貫性を持たせることが不可欠です。

> これらは倫理的な原則ですが、政治的な言葉に置き換えたり、法的な形で解釈したりすることができます。

これらの原則に基づき、本宣言の倫理的枠組み内でのデジタル移行のためのガイドライン作成に向けた勧告が行われました。これは、AI がアルゴリズム・ガバナンス、デジタル・リテラシー、多様性のデジタル・インクルージョン、生態系の持続可能性といった、公共の利益を促進するのに役立つ社会への移行を熟考するための主要ないくつかの分野にまたがる重要なテーマをカバーすることを目的としています。

誰のための宣言か?

モントリオール宣言は、人工知能の責任ある開発に参画しようとするすべての個人、組織および企業を対象としたものです。例えば、科学的または技術的な貢献、社会的プロジェクトの展開、適用されるルール（規制、規範）の策定、不適切または賢明とはいえないアプローチへの異議の申立て、または必要であれば世論への警告、といった形で貢献できます。

また、社会の変化の進行状況を把握し、より大きな善をもたらすデジタル移行を可能にする枠組みを迅速に確立し、AI 開発によってもたらされる深刻なリスクを予見することを市民から期待されている、選挙で選ばれた、または指名された政治的代表者も対象としています。

どのような方法による宣言か?

本宣言は、市民、専門家、公務員、産業界関係者、市民団体および職能団体の間の対話を開始する包摂的な熟議プロセスから生まれました。このアプローチの利点は 3 つあります。

1. AI の社会的および倫理的論争を共同で調停すること
2. 責任ある AI に関する反省的視点の質を向上させること
3. 責任ある AI のための提案の正当性を強化すること

原則と勧告の策定は、公共の場、専門家組織の会議室、国際的な専門家による円卓会議、研究室、教室またはオンラインで、さまざまな関係者を巻き込みながら等しく厳正な態度で行う共同作業です。

宣言の後は?

1950年代以降、AI技術は着実に進歩しています。また、重要なイノベーションは、指数関数的に増加しています。この宣言は、このような点を踏まえ、一般に開かれた指針文書として、知識や技術の進歩にあわせる形で、および、社会におけるAIの利用に関するユーザーの意見を踏まえて、改訂され、適応していくべきものとの認識が大切です。本宣言の策定プロセスを通じて、私たちは、人工知能技術によってもたらされる人類の未来についての、オープンで包摂的な対話の出発点に到達したのです。

序文

人類の歴史上初めて、自然知能だけしかできないと考えられていた複雑なタスクを実行できる自律システムの作成が可能になりました。すなわち、大量の情報の処理、計算と予測、状況変化に応じた学習・適応、対象物の認識・分類などです。これらのタスクは非物質的である特性から、また、人間の知性との類似性から、私たちはこれらの広範なシステムを人工知能という一般名で呼んでいます。人工知能は、科学的および技術的進歩の一端を担っており、生活環境や健康の改善、正義の促進、富の創出、公共の安全の向上、人間活動による環境や気候への影響の軽減によって、かなりの社会的利益を生み出すことができます。知能を持った機械は、人間よりも優れた計算を実行するだけではありません。さらに、生きとし生けるものと交流し、付き合い、世話することもできます。

しかしながら、人工知能の開発は、大きな倫理的課題と社会的リスクをもたらしています。実際、知能を持った機械は個人や集団の選択を制限し、生活水準を低下させ、労働組織や雇用市場を混乱させ、政治に影響を与えます。また、基本的権利と衝突し、社会的および経済的不平等を悪化させ、生態系や気候、環境に影響を及ぼす可能性があります。科学の進歩にはもちろん、社会に生きることにはリスクが伴いますが、不確実な世界で遭遇するリスクに意味を与える道徳的および政治的目的を決定するのは市民の責任です。

人工知能を展開するリスクが低ければ低いほど、それがもたらすメリットは大きくなります。人工知能開発の第一の危険は、私たちが計算によって未来を支配できるという錯覚を与えることにあります。社会を数字の羅列に還元し、アルゴリズムの手順を通して支配することは、昔からの夢物語であり、今でも人間の野望を駆り立てています。しかし、人々の活動については、明日が今日と同じになることはめったになく、何が道徳的価値を持つか、ましてや何が社会的に望ましいかを数字が判断することなどありえないのです。

本宣言の原則は、人工知能の開発を道徳的および社会的に望ましい方向に導くのに役立つ道徳的な羅針盤の方位のようなものです。これらはまた、人工知能の展開によって影響を受ける分野においては、国際的に認められた人権を促進する倫理的な枠組みにもなります。全体として見ると、明確にされたこれらの原則は、人工知能システムに対する社会的信頼を育むための基盤となるものです。

本宣言の原則は、人間は感覚、思考、感情を備えた社会的存在として成長しようとし、感情的、道徳的、知的能力を自由に行使することによって自らの潜在能力を発揮しようとするという共通認識に基づくものです。人工知能の開発および展開が、基本的な人間の能力や目標の保護と両立し、それらの完全な実現に貢献するのを保証することは、地域、国、国際レベルのさまざまな官民の関係者や政策立案者の責務です。この目標を念頭において、私たちは提案された原則を、それらが適用される特定の社会的、文化的、政治的、および法的文脈を考慮に入れて、首尾一貫した方法で解釈しなければなりません。

1

ウェルビーイングの原則

人工知能システム（AIS）の開発と使用は生きとし生けるもののウェルビーイングの増大を可能にしなければなりません

1. AIS は、個人が自らの生活環境、健康、および労働環境を改善するのに役立つものでなければなりません。

2. AIS は、他の生きとし生けるものに危害を加えない限り、個人が自らの嗜好を追求できるようにするものでなければなりません。

3. AIS は、人々が心的および身体的能力を発揮できるようなものでなければなりません。

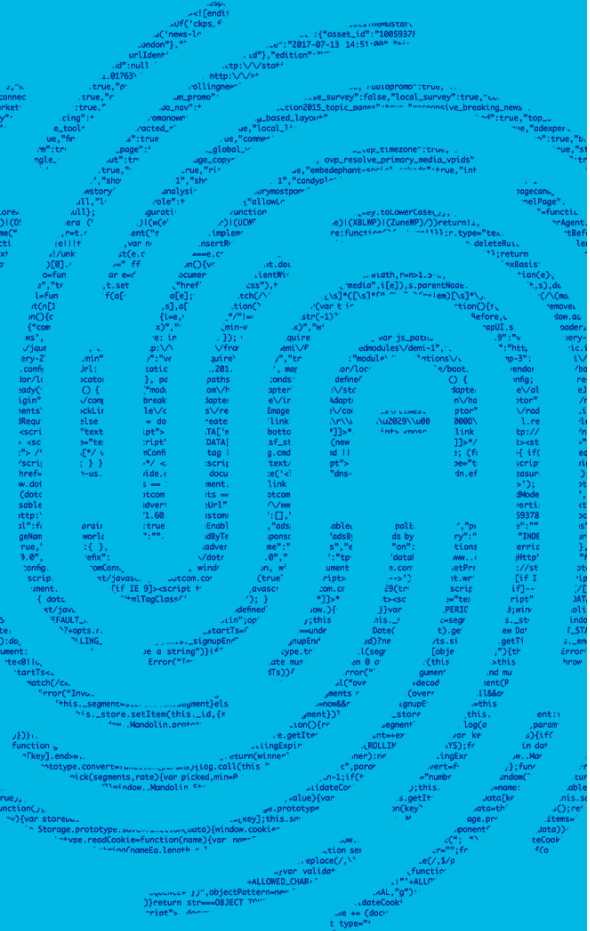
4. AIS は、他の方法で達成できるウェルビーイングよりも優れたウェルビーイングを実現できない限り、イルビーイングの原因となるものであってはなりません。

5. AIS の利用によって、ストレス、不安、またはデジタル環境から嫌がらせを受けているという感覚が増加することに寄与することがあってはなりません。

2

自律尊重の原則

**AIS は
人々の自律性を
尊重しつつ
人々が自らの
生活や環境を
よりいっそうコントロール
できるようにする
ことを目標として
開発および使用
されなければ
なりません**



1. AIS は、各人が自分自身の道徳的目標と、生きるに値すると考える生き方を実現できるものでなければなりません。

2. AIS は、抑圧的な監視や評価またはインセンティブ・メカニズムを実装することによって、直接的であれ間接的であれ、個人に特定のライフスタイルを押し付けることを目的に開発または使用されてはなりません。

3. 公的機関は、特定の「良き生」という考え方を促進したり貶めたりするために、AIS を使用してはなりません。

4. 各種の適切な情報へのアクセスを保証し、基本的なスキル（デジタルおよびメディア・リテラシー）の習得を促進し、批判的思考の発達を育むことにより、市民にデジタル技術に関する能力と権利を与えることが重要です。

5. AIS は、信頼できない情報、嘘、またはプロパガンダを広めるために開発されてはならず、それらの流布を封じ込めることを目的として設計されなければなりません。

6. AIS の開発では、人の注意を捕捉する技術や人間の特徴（外見、声など）の模倣により、AIS と人間を混同させる可能性がある方法で依存関係を生み出すことを避けなければなりません。

3

プライバシー および内密性保護 の原則

プライバシーおよび内密性は、AISの侵入やデータ取得・アーカイブシステム（DAAS）から保護されなければなりません

1. 人々が監視やデジタル評価に晒されないパーソナル・スペースは、AIS やデータ取得・アーカイブシステム（DAAS）の侵入から保護されなければなりません。

2. 思考や感情の内密性は、害を及ぼす可能性がある AIS や DAAS の使用、特に人々に、またはそのライフスタイルの選択に関して道徳的判断を課すような使用方法から厳重な保護がなされなければなりません。

3. 人々は私生活においては常にデジタルへの接続から切り離される権利が保障されなければならず、AIS は人々に接続の維持を促すことなく、定期的に接続を切断するオプションを明示的に提供するものであるべきです。

4. 人々は、自分の嗜好に関する情報を厳重に管理することができなければなりません。AIS は、自由意思による、情報提供を受けた上での同意なしに、個人の行動に影響を与えるような各個人の嗜好に関するプロファイルを作成するものであってはなりません。

5. DAAS は、データの機密性と個人プロファイルの匿名性を保証するものでなければなりません。

6. すべての人は、自分のパーソナルデータ、特にその収集、使用、流布に関して、厳重な管理ができなければなりません。個人による AIS およびデジタルサービスへのアクセスは、パーソナルデータのコントロールまたはオーナーシップを放棄することを条件とするものであってはなりません。

7. 個人は、知識の進歩に貢献するために、パーソナルデータを研究機関に自由に提供できなければなりません。

8. 個人のアイデンティティの完全性は保証されなければなりません。AIS は、人の評判を傷つけたり、他の人を操作したりする目的で、人の外見や声、その他の個人の特徴を模倣したり変更したりするために使用されてはなりません。

4

連帯の原則

AIS の開発は人々や世代間の連帯の絆を維持することと両立されなければなりません

1. AIS は、充実した道徳的および情緒的な人間関係の維持を脅かすものであってはならず、これらの関係を育み、人々の脆弱性や孤立を軽減することを目標として開発されなければなりません。

2. AIS は、複雑なタスクについて人間と協力することを目標として開発されなければならず、人間同士の共同作業を促進するものであるべきです。

3. AIS は、高度な人間関係を必要とする職務に従事する人々に代わるものとして実装されるべきではなく、これらの関係を円滑にするために開発されるべきです。

4. AIS を使用する医療システムでは、患者やその家族、医療スタッフあいだの関係の重要性を考慮に入れなければなりません。

5. AIS の開発は、人間または人間以外の動物の外見や行動に似せて設計されたロボットに対する残忍な行動を促すものであってはなりません。

6. AIS は、リスク管理を改善し、個人および集団のリスクをより公平かつ相互に分配する社会的状況を整えるのに役立つものであるべきです。

5

AIS は理解可能性、正当性およびアクセシビリティの基準を満たし、民主的な精査、議論およびコントロールを受けなければなりません

民主的参加の原則

1. AIS プロセスは、その決定が個人の人生、生活の質（QOL）、または評判に影響を与える場合、その開発者が理解可能なものでなければなりません。
2. AIS による決定は、個人の人生、生活の質、または評判に影響を与える場合、その決定内容を、使用する人々、またはその使用の帰結によって影響を受ける人々が理解できる言葉で正当化できなければなりません。正当化できるとは、決定を行う上で最も重要な要素やパラメーターを透明なものにすることであり、正当化は、人間が同様の決定を行う上で要求される場合と同じ形を取らなければなりません。
3. アルゴリズムのコードは、公開か非公開かを問わず、検証および管理の目的で、関連する公的機関や利害関係者が常にアクセスできるようにしなければなりません。
4. AIS の操作エラー、予期しない、または望ましくない影響、セキュリティ違反、およびデータ漏洩の発見は、関連する公的機関、利害関係者、およびその影響を受ける人々に必ず報告されなければなりません。
5. 公的な決定の透明性要件にしたがって、公的機関が使用する意思決定アルゴリズムのコードは、誤用された場合に重大な危険をもたらす高いリスクがあるアルゴリズムは除き、すべての人がアクセスできなければなりません。
6. 市民生活に重大な影響を与える公的な AIS に関して、市民は、これらの AIS の社会的パラメーター、その目的、およびその使用の制限について検討する機会を与えられ、そのスキルを持つ必要があります。
7. AIS が、プログラムされた目的や使用される目的どおりに実行されていることを私たちが常に確認できるようにしなければなりません。
8. サービスを利用する人は誰でも、自分に関する決定または影響を与える決定が AIS によって行われたものかどうかを知る必要があります。
9. チャットボットを使用するサービスを利用する人は誰でも、AIS とやり取りしているのか、実在の人物とやり取りしているのかを簡単に識別できるようにする必要があります。
10. 人工知能の研究は、常にオープンで、すべての人がアクセスできるものでなければなりません。

6

公平・公正の原則

AISの開発および使用は、公正で公平な社会の創造に貢献しなければなりません

1. AIS は、社会的、性的、民族的、文化的、または宗教的な違いなどに基づく差別を生み出したり、強化したり、再生産したりしないように設計および訓練されなければなりません。

2. AIS の開発は、権力、富、または知識の違いに基づく集団や人々の間の支配関係を排除するために役立たなければなりません。

3. AIS の開発は、社会的不平等や脆弱性を軽減することにより、すべての人に社会的および経済的利益をもたらさなければなりません。

4. 産業用 AIS の開発は、天然資源の採取からリサイクルまで、データ処理を含むライフサイクルのすべての段階で望ましい労働条件に適合していなければなりません。

5. AIS およびデジタルサービスのユーザーによるデジタル活動は、アルゴリズムの機能に貢献し、価値を生み出す労働として認識されなければなりません。

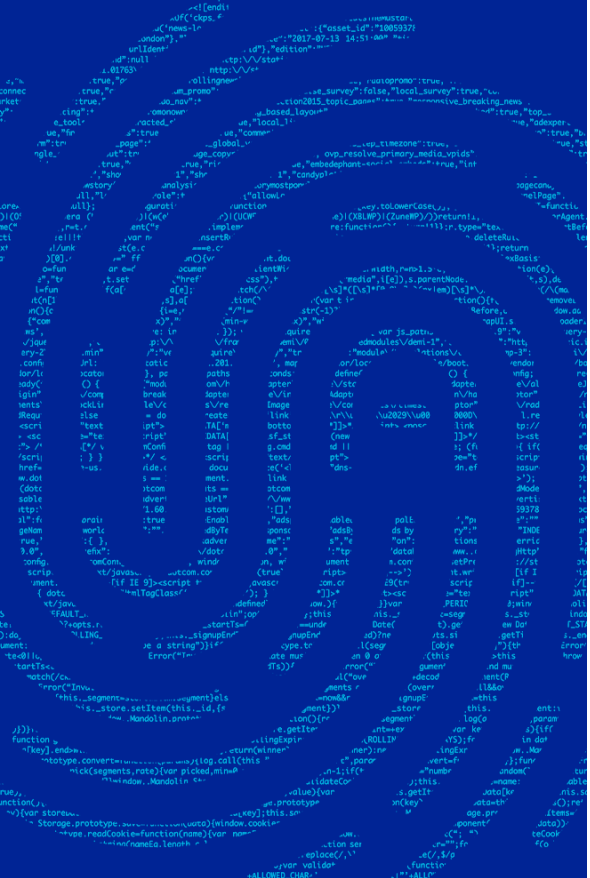
6. 基本的なリソース、知識、デジタルツールへのアクセスは、すべての人に保証されなければなりません。

7. 私たちは、社会的に公平な目的として、コモンズとしてのアルゴリズムおよびそれらを訓練するために必要なオープンデータの開発を支援し、それらの使用を拡大するべきです。

7

ダイバーシティ・インクルージョンの原則

AIS の開発および使用は、社会や文化の多様性の維持と両立しなければならず、ライフスタイルの選択や個人的な経験の範囲を制限してはなりません



1. AIS の開発および使用は、行動や意見の標準化による均質化へとつながるものであってはなりません。

2. AIS の開発および展開は、アルゴリズムが考案された時点から、社会的および文化的多様性に関して、社会に存在する多くの表現のあり方が考慮に入れられなければなりません。

3. AI の開発環境は、研究機関であれ産業界であれ、社会における個人や集団の多様性を反映した包摂的なものでなければなりません。

4. AIS は、取得したデータを使用して、個人をユーザープロファイルに閉じ込めたり、個人のアイデンティティを固定したり、フィルターバブルに閉じ込めたりすることを回避するものでなければなりません。このことは、特に教育、司法、ビジネスなどの分野では、個人の成長の可能性を制限し、封じ込めることにつながります。

5. AIS は、民主的社会の必須条件である、自由な思想の表明や多様な意見を聞く機会を制限する目的で開発または使用されてはなりません。

6. 各サービスカテゴリーについて、AIS の提供は、事実上の独占が形成されたり、また個人の自由が損なわれたりするのを防ぐため、多様なものでなければなりません。

8

AI 開発に携わるすべての人はAIS の使用による悪影響を可能な限り予期し、それを回避するための適切な措置を講じることによって注意を払わなければなりません

慎重さの原則

1. AI 研究および AIS 開発（公的または私的を問わず）は、有害な使用を制限するために、「有益な使用と有害な使用」の両方の可能性を考慮するメカニズムを開発する必要があります。

2. AIS の誤用が公衆の健康や安全を危険にさらす、あるいは、その可能性が高い場合は、そのアルゴリズムへのオープンアクセスと一般への普及を制限するという慎重さが求められます。

3. AIS は市場に投入される前に、有償または無償で提供されるかどうかにかかわらず、厳格な信頼性、セキュリティ、および完全性の要件を満たしていなければならず、人々の生活を危険にさらしたり、生活の質を損なったり、評判や心理的な一体性に悪影響を与えたりしないという基準を満たすようにされなければなりません。これらの基準は、関連する公的機関および利害関係者に公開されなければなりません。

4. AIS の開発は、ユーザーデータの誤用のリスクをあらかじめ阻止し、パーソナルデータの完全性や機密性を保護しなければなりません。

5. AIS や DAAS で発見されたエラーや欠陥は、個人の尊厳や社会組織に重大な危険をもたらすセクターの公的機関や企業により世界規模で公に共有される必要があります。

9

責任の原則

AIS の開発および使用は、決定を下す必要がある場合に、人間の責任を軽減することに寄与してはならない

1. AIS の勧告に基づく決定と、それに基づいてなされる行動について責任を負うのは人間だけです。

2. 個人の人生、生活の質（QOL）、または評判に影響を与える決定を下さなければならないすべての領域において、時間と状況が許す限り、最終決定は人間によって行われなければならず、その決定は自由意思により、情報提供を受けた上でのものでなければなりません。

3. 何ものかを殺すという意思決定は常に人間が行うべきものであり、この決定の責任を AIS に転嫁してはなりません。

4. AIS に犯罪または違反を犯すことを許可した、または AIS にそれらを許可することによって過失を示した場合は、その人たちがこの犯罪または違反の責任を負うことになります。

5. AIS によって被害が発生し、AIS が信頼できるものであり、意図したとおりに使用されたことが証明された場合、その開発または使用に関与した人々に責任を負わせることは妥当ではありません。

10

持続可能な開発の原則

AIS の開発および使用は、地球の環境における強い持続可能性を保証するために実施されなければなりません

1. AIS ハードウェア、そのデジタルインフラストラクチャ、およびデータ センターなどのAIS が依存する関連オブジェクトは、そのライフ サイクル全体を通して、最大のエネルギー効率と温室効果ガス排出の削減を目指さなければなりません。

2. AIS ハードウェア、そのデジタルインフラストラクチャ、およびAIS が依存する関連オブジェクトは、生成される電気や電子廃棄物の量を最小限に抑え、循環経済の原則にしたがってメンテナンス、修理、およびリサイクルの手順を提供することを目指さなければなりません。

3. AIS ハードウェア、そのデジタルインフラストラクチャ、および AIS が依存する関連オブジェクトは、そのライフサイクルのすべての段階で私たちが生態系や生物多様性に及ぼす影響を最小限に抑えなければなりません。このことは特に資源の採取と、耐用年数を超えた機器の最終的な処分に関してあてはまります。

4. 公共および民間の関係者は、天然資源や生産物の浪費の抑制、持続可能なサプライチェーンおよび取引の構築、地球規模の汚染の軽減にあたって、環境に対する責任ある AIS 開発をサポートしなければなりません。

用語集

アルゴリズム
アルゴリズムは、有限で明確な一連の手順によって問題を解決する方法です。

具体的には、人工知能について、望ましい結果を得るためデータ入力に適用される一連の操作手順のことです。

人工知能 (AI)
人工知能 (AI) とは、機械が人間の学習、すなわち習得、予測、意思決定、周囲の認識をモデル化できるようにする一連の技術を指します。コンピューティング・システムの場合、人工知能はデジタルデータに適用されます。

人工知能システム (AIS)
AIS は、ソフトウェア、接続されたオブジェクト、ロボットのいずれであっても、人工知能アルゴリズムを使用するあらゆるコンピューティング・システムです。

チャットボット
チャットボットは、ユーザーと自然言語で会話できる AI システムです。

データ取得・アーカイブ システム (DAAS)
DAAS は、データを収集および記録できるあらゆるコンピューティング・システムを指します。このデータは、最終的に AI システムのトレーニングや意思決定パラメーターとして使用されます。

決定の正当性
AIS の決定は、この決定の動機となる重要な理由が存在し、これらの理由を自然言語で伝えることができる場合に正当と認められます。

ディープラーニング (深層学習)
ディープラーニングは、多くの層で人工ニューロン・ネットワークを使用する機械学習の一分野です。これは最新の AIのブレイクスルー (飛躍的進歩) の背後にあるテクノロジーです。

デジタルコモンズ
デジタルコモンズは、コミュニティによって作成されたアプリケーションまたはデータです。有形の商品とは異なり、簡単に共有でき、使用しても劣化しません。したがって、プロプライエタリ・ソフトウェアとは異なり、(多くの場合、プログラマー間の共同作業の成果である) オープン ソース ソフトウェアは、ソース コードがオープンであり、すべての人がアクセスできるため、デジタルコモンズと見なされます。

デジタルディスコネクト
デジタルコネクトとは、個人のオンライン活動の一時的または永久的な停止を指します。

デジタル・リテラシー
個人のデジタル・リテラシーとは、経済的および社会的生活に参加するために、デジタルツールやネットワーク化されたテクノロジーを通じて、安全かつ適切に情報にアクセスし、それを管理、理解、統合、伝達、評価、作成できる能力を指します。

フィルターバブル
フィルターバブル (またはフィルタリングバブル) という表現は、インターネット上で個人に届く「フィルター処理された」情報を指します。ソーシャル・ネットワークや検索エンジンといったさまざまなサービスは、ユーザーにパーソナライズされた結果を提供します。これは、ユーザーを共通の情報にアクセスできなくし、個人を (「バブル」の中に) 孤立させる効果があります。

GAN
Generative Adversarial Network（敵対的生成ネットワーク）の頭字語。GAN では、2 つの拮抗するネットワークが競合して画像が生成されます。たとえば、人間にとってほぼ現実に見える画像、録音、または動画を作成するために使用することができます。

理解可能性（わかりやすさ）
AIS は、必要な知識を持つ人間がその操作、つまりその数学的モデルやそのモデルを決定するプロセスを理解できる場合に理解できるものとなります。

機械学習
機械学習は、機械がそれ自体で学習できるようにアルゴリズムをプログラミングする人工知能の一部門です。

さまざまな手法は、機械学習の 3 つの主要なタイプに分類することができます：

> 教師あり学習では、人工知能システム（AIS）は、入力されたデータから値を予測することを学習します。これには、トレーニング中にアノテーション付きのエントリ値のカップルが必要です。たとえば、システムは写真に写っているオブジェクトを認識することを学ぶことができます。

> 教師なし学習では、AIS は、たとえば、さまざまな同種のパーティションに分割するために、アノテーションが付けられていないデータ間の類似点を見つけることを学習します。これにより、システムはソーシャル メディア ユーザーのコミュニティを認識することができます。

> AIS は強化学習を通じて、トレーニング中に受け取る報酬を最大化するために、環境に応じて行動することを学習します。これは、囲碁やビデオゲーム Dota2 で AIS が人間を打ち負かすことができたテクニックです。

オンライン活動
オンライン活動とは、個人がデジタル環境（コンピューター、電話、またはその他のあらゆる接続されたオブジェクト）で行うすべての活動を指します。

オープンデータ
オープンデータは、ユーザーが自由にアクセスできるデジタルデータです。たとえば、公開されているほとんどの AI 研究結果がこれに該当します。

経路依存性
かつては合理的であると見なされていたが現在は標準以下である技術的、組織的、または制度上の決定が、依然として意思決定に影響を与え続けている社会的メカニズム。認知バイアスのため、または変更には多大な費用や労力が必要となるために維持されているメカニズム。炭素排出量が非常に少ない交通機関を整備するための変更を検討するのではなく、交通最適化プログラムにつながる場合、都市の道路インフラがこれに該当します。特別なプロジェクトに AI を使用する場合、教師あり学習のトレーニング データは、現在論争の的となっている古い組織パラダイムを強化することがあるので、このメカニズムを知っておく必要があります。

パーソナルデータ
パーソナルデータとは、個人を直接的または間接的に識別するのに役立つデータです。

リバウンド効果
リバウンド効果とは、商品、機器、サービスのエネルギー効率や環境パフォーマンスの向上が、比例以上の使用の増加につながるメカニズムです。

たとえば、画面サイズが大きくなり、家庭内の電子機器の数が増え、車や飛行機で移動する距離が長くなります。その世界的な結果として、資源や環境に対する圧力が高まります。

信頼性
AIS は、それが設計された目的であるタスクを期待どおりに実行するときに信頼できるものとなります。信頼性とは、51〜100% の範囲の成功率であり、偶然よりも厳密に優れていることを意味します。システムの信頼性が高いほど、その動作は予測可能になります。

環境における強い持続可能性

環境における強い持続可能性の概念は、持続可能であるためには、天然資源の消費率や汚染物質の排出率が、地球の環境限界、資源や生態系の再生率、および気候の安定性と両立しなければならないという考えにさかのぼります。

より少ない労力で済む弱い持続可能性とは異なり、強い持続可能性は、天然資源の損失を人工資本で置き換えることは許可しません。

持続可能な開発

持続可能な開発とは、社会に必要な資源やサービスを提供する自然システムの能力に適合した人間社会の開発を指します。これは将来の世代の存在を危険にさらすことなく現在のニーズを満たす経済的および社会的な発展です。

トレーニング

トレーニングは、AIS がデータからモデルを構築するための機械学習プロセスです。AIS のパフォーマンスはモデルの質に依存し、それはトレーニング中に使用されるデータの量と質に依存します。

クレジット

人工知能の責任ある開発に関するモントリオール宣言の作成は、市民の協議プロセスおよび AI 開発の専門家やステークホルダーとの対話を活かした大学間にまたがる学際的な科学チームの作業の成果です。

Christophe Abrassart、モントリオール大学環境デザイン学部准教授および都市計画学科 Ville Prospective ラボ共同ディレクター、倫理研究センター (Centre de recherche en éthique, CRÉ) メンバー

Yoshua Bengio、モントリオール大学コンピュータサイエンスおよびオペレーションズリサーチ学科正教授、MILA および IVADO 科学ディレクター

Guillaume Chicoisne、IVADO 科学プログラム ディレクター

Nathalie de Marcellis-Warin、モントリオール理工科大学正教授、大学間共同研究および組織分析センター (Center for Interuniversity Research and Analysis of Organization, CIRANO) 所長兼CEO

Marc-Antoine Dilhac、モントリオール大学哲学科准教授、倫理研究センター (Centre de recherche en éthique, CRÉ) 倫理および政治グループ研究主任教授、公共倫理と政治理論のカナダ研究議長、哲学・市民権・青少年研究所所長

Sébastien Gambs、ケベック大学モントリオール校コンピューター サイエンス学部教授、ビッグデータのプライバシー保護および倫理分析のカナダ研究委員長

Vincent Gautrais、モントリオール大学法学部正教授。公法研究センター (Centre de recherche en droit public, CRDP) 所長。情報技術および電子商取引法の研究ユニット、L.R.Wilson チェアの主任教授

Martin Gibert、IVADO 倫理カウンセラーおよび倫理研究センター (Center de recherche en éthique, CRÉ) 研究員

Lyse Langlois、社会科学部正教授および副学部長。応用倫理研究所 (Institut d'éthique appliqué, IDÉA) 所長。グローバル化と仕事に関する大学間共同研究センター (CRIMT) 研究員

François Laviolette、ラヴァル大学コンピューター サイエンスおよびソフトウェア エンジニアリング学科正教授、ビッグデータ研究センター (Centre de recherche en données massives, CRDM) 所長

Pascale Lehoux、モントリオール大学公衆衛生学部 (École de santé publique, ESPUM) 正教授。健康における責任あるイノベーション研究主任教授

Jocelyn Maclure、ラヴァル大学哲学部正教授およびケベック科学と技術における倫理委員会 (CEST) 委員長

Marie Martel、モントリオール大学図書館情報学部 (École de bibliothéconomie et des sciences de l'information) 教授

Joëlle Pineau、マギル大学コンピュータサイエンス学部准教授、モントリオール Facebook AI ラボ ディレクター、推論と学習ラボ共同ディレクター

Peter Railton、ミシガン大学哲学科教授、アメリカ芸術科学アカデミー会員。Gregory S. Kavka 賞、John Stephenson Perrin賞受賞、Arthur F. Thurnau 教授職

Catherine Régis、モントリオール大学法学部准教授、健康法および政策における共同文化に関するカナダ研究委員長、公法研究センター (Centre de recherche en droit public, CRDP) 正規研究員

Christine Tappolet、モントリオール大学哲学科正教授、倫理研究センター (Centre de recherche en éthique, CRÉ) 所長

Nathalie Voarino、モントリオール大学生命倫理学博士候補生

翻訳監修

鹿野 祐介、**カテライ アメリア**、**岸本 充生**、大阪大学社会技術共創研究センター(ELSIセンター・Research Center on Ethical, Legal and Social Issues, Osaka University)

パートナー

Université de Montréal

CRE
CENTRE DE RECHERCHE EN ETHIQUE

LAB VILLE PROSPECTIVE

OBSERVATOIRE INTERNATIONAL SUR LES IMPACTS SOCIÉTAUX DE L'IA ET DU NUMÉRIQUE

IVADO

SAT

Mila

ICRA
Programme
IA et
société

Université de Montréal | designΠsociété

polEtHics
CHAIRE DE RECHERCHE DU CANADA
ÉTHIQUE PUBLIQUE ET THÉORIE POLITIQUE

CIRANO
Allier savoir et décision
Centre interuniversitaire de recherche en analyse des organisations

Québec
Fonds de recherche – Nature et technologies
Fonds de recherche – Santé
Fonds de recherche – Société et culture

IDÉA
Institut d'éthique appliquée
de l'Université Laval

MUSÉE DE LA CIVILISATION
Québec

CENTRE D'ÉTUDES ET DE RECHERCHES INTERNATIONALES
CÉRIUM
Université de Montréal

UNIVERSITÉ LAVAL

crdm.ul
CENTRE DE RECHERCHE EN DONNÉES MASSIVES DE L'UNIVERSITÉ LAVAL

Canada

Centre Culturel Canadien
Paris

MEC | MAISON DES ÉTUDIANTS CANADIENS

< >

Montréal Declaration Responsible AI_

</ >

montrealdeclaration-responsibleai.com