第 57 回「世界平和の日」教皇メッセージ

（2024 年 1 月 1 日）

人工知能（AI）と平和

新年の初め、主がわたしたち一人ひとりに与えてくださる恵みの時にあたり、神の民、諸国民の皆さん、各国首脳、他宗教指導者、市民社会の代表者、現代に生きるすべての人に、わたしから平和を祈るごあいさつを申し上げます。

1. 科学技術の進歩を平和への道に

聖書によれば、神が人にご自分の霊をお与えになったのは、「どのような工芸にも知恵と英知と知識」（出エジプト 35・31）をもたせるためでした。知性は、わたしたちをご自分の似姿に造られた（創世記 1・26 参照）創造主から授かった尊厳の表れであり、それによってわたしたちは、そのかたの愛に、自由意志と理解をもってこたえられるのです。科学と技術は、人間の知性が本質的に備えるこの関係性という性質の特別な表出であり、どちらも知性の創造的な潜在能力による、傑出した産物です。

第二バチカン公会議は『現代世界憲章』においてこの真理を再確認し、「人間は、労働と才能をもって自分の生活の向上を目指し、つねに努力してきた」と明言します[1]。人間は「技術を用いて」、大地が「全人類家族のよい住みか」[2]となるよう努力するとき、神の計画にかなう振る舞いをし、創造の完成をもたらし、諸民族の間に平和を広めようという、神のみ旨に協力しているのです。同じように科学技術の進歩は、それが人間社会のよりよい秩序に貢献し、自由と友愛ある交わりの拡大に寄与するかぎり、人間の向上と世界の変革へと結びつくのです。

人間の生活を悩ませ、大きな苦しみを引き起こしてきた無数の不幸からの救済を可能にした、科学技術の驚異的な進展をわたしたちは率直に喜び感謝しています。同時に、テクノサイエンスの進歩は、かつてないほど現実を操作することを可能にし、人類の手に膨大な可能性をもたらしていますが、その中には人類の生存を脅かし、共通の家を危険にさらすものもあります[3]。

新たな情報技術の目覚ましい進歩はとりわけデジタル領域で著しく、それゆえ、諸民族間の正義と和合の追求にとって重大な意味をもつ、胸躍るチャンス、と同時に深刻なリスクをもたらしています。ですから、いくつかの差し迫った問いに向き合う必要があります。新しいデジタル技術がもたらす中長期的な影響とはいかなるものなのか。それは、個人の生活や社会生活に、国際社会の安定と平和に、どのような影響を与えるのか。

2. 期待とリスクのはざまにある人工知能の未来

ここ数十年間の情報技術の進歩とデジタル技術の発展は、すでにグローバル社会とその動向に大きな変革をもたらし始めています。新しいデジタルツールは、コミュニケーション、行政、教育、消費、人的交流、その他日常生活の無数の側面に変化をもたらしています。

さらに、インターネット上に残されたデジタル記録から、さまざまなアルゴリズムを採用するテクノロジーによって、多くの場合利用者が知らないうちに、商業目的や政治的意図によって、人々のメンタリティーや人間関係における習性の操作を可能にするデータが抽出され、自由な選択を自覚的に行うことが制限されています。実際、情報の氾濫を特徴とするウェブのような空間では、利用者が認識するとは限らない選択基準に従って、データの流れが構成されうるのです。

科学研究と技術革新は、現実から切り離されたものでもなければ「中立」[4]でもなく、文化的な影響を受けるものであることを忘れてはなりません。ひたすら人間による活動である以上、その方向性は各時代の個人的、社会的、文化的価値観によって条件づけられた選択を反映しています。その結果についても同様です。それこそが、取り巻く世界に対する、まさしく人間からのアプローチの成果として必ず倫理的側面を有し、実験を計画して特定の目的へと生産を向かわせる人間の決断に密接に結ばれているのです。

これはまた、各種人工知能にも当てはまります。今のところ科学技術の世界では、その一義的な定義はなされていません。現在では一般化されたこの語自体は、人間の認知能力を機械に再現・模倣させることを目的とした、さまざまな科学、理論、技術を含んでいます。「各種知能」と複数形で語ることで、そうした圧巻的で強力なシステムと人間との間にある、埋めようのない隔たりを強調することができます。それらは、人間の知能の一部の機能を単に模倣・再現するにすぎないという意味で、結局のところ「半端」なものなのです。複数形を用いることは、多種多様なこれらのデバイスは、つねに「社会技術的（ソシオテクニカル）なシステム」として捉えなければならないことを明確にします。実際、基盤となっている技術とは関係なく、それらからどのような影響を受けるのかは、設計によるばかりでなく、その技術の保有者や開発者の目的と関心、そして使用される状況によっても異なります。

したがって人工知能は、さまざまな現実の集合体として理解されるべきもので、その開発が人類の未来や民族間の平和に有益な貢献をなすのは自明であると決めつけることはできません。そうした肯定的な結果が生じうるのは、わたしたちが責任をもって行動し、「包摂性、透明性、安全性、公平性、プライバシー、信頼性」[5]といった人間の基本的価値観を尊重できる場合のみです。

アルゴリズムやデジタル技術の設計を企図する側の、倫理的かつ責任ある行動の確約を前提とするだけでは十分ではありません。新たな倫理課題の検討や、各種人工知能を使用する人、また人口知能

の影響を受ける人の権利保護を管轄する機関を強化し、必要であれば新設することが求められています[6]。

それゆえ技術の大幅な拡充には、その後の進展に対する責任についての十分な養成が伴われるべきです。人間が利己主義、私利私欲、利益追求、権力欲の誘惑に負けたとき、自由で平和な共存は脅かされます。ですからわたしたちには、広く目を配り、人間と共同体の包括的発展に寄与するため、科学技術研究を平和と共通善の追求に向かわせる義務があるのです[7]。

各人が本性的に備える尊厳と、わたしたちを唯一の人類家族として結びつける兄弟愛が、新技術の開発の基盤であるべきで、その実用化にあたっての評価の厳然たる基準とならなければなりません。正義を重んじつつデジタル技術が進歩し、平和に貢献するようにです。人類全体の生活の質を向上させず、逆に格差や争いを悪化させるような技術開発は、真の進歩とはいえません[8]。

人工知能は、ますます重要なものとなるでしょう。そこにある課題は技術的なものばかりでなく、人間学的、教育的、社会的、政治的なものにも及びます。たとえば、手間の削減、効率的な生産、輸送負担軽減、市場の活性化、さらにはデータの収集・整理・分析プロセスにおける革命が約束されています。起きている急速な変化を認識し、基本的人権を守り、かつ人間の全人的発展を促進する制度と法律に沿って、その変化をうまく管理しなければなりません。人工知能は、人間の比類なき潜在能力や、より高い志に仕えるべきで、それらと競合するものであってはなりません。

3. 未来の技術――自ら学習する装置

機械学習（ML=machine learning）技術に基づく人工知能は、さまざまな形態を有し、いまだ開拓段階にあるとはいえすでに社会構造に顕著な変化をもたらしており、文化や社会行動、平和構築といったものに多大な影響を及ぼしています。

機械学習やディープラーニングなどの開発は、技術や工学の領域を超えた、人間のいのちの意味、基本的な知の歩み、真理に至るための思考力、これらと強く結ばれた理解力に関する問いを提起しています。

たとえば何らかのデバイスに、構文的にも意味的にもまとまりのある文章を生成する能力があったとしても、それがその信憑性を保証するわけではありません。人工知能は「欺く」ことがある、つまり一見もっともらしくても、実際はでたらめだったり偏向性があったりする主張を生成することがあるといわれています。人工知能が、フェイクニュースを拡散してメディア不信を高めるデマ活動に利用されれば、深刻な問題が引き起こされます。秘密保持、データ所有権、知的財産なども、問題とされるテクノロジーが深刻なリスクを招くまた別の分野であり、それが不適切に用いられると、差別、選挙活動への介入、監視・管理社会の定着、デジタル・デバイド（情報格差）、共同体から分離した個人主義の極端

化などにさらなる負の影響を来します。これら要因にはどれも、対立を助長し、平和を阻害する危険があるのです。

4. 技術主義的（テクノクラティック）パラダイムの限界を意識する

この世界はあまりにも広大で、多様であり、複雑で、完全に把握することも分類することもできません。人間の頭脳をもってしてでは、たとえ最先端のアルゴリズムの助けを借りたとしても、その豊かさを使い尽くすことはできないでしょう。事実、そうしたアルゴリズムは、確実な未来予想を提供するのではなく、統計的な近似値を提示するにすぎないのです。すべてを予測できるはずはなく、すべてを計算できるわけでもありません。結局のところ「現実は理念に勝る」[9]のであり、わたしたちの計算能力がどれほど驚異的であろうとも、いかなる測定の試みからもこぼれ落ち、手の届かない部分はつねに残るのです。

さらに、人工知能が分析する膨大なデータは、公平性をそのまま保証するものではありません。アルゴリズムが情報を割り出す際には、それを歪曲させる危険性がつねにあり、情報が生み出された環境にある不正義や偏見を再現してしまうのです。アルゴリズムが高速化、複雑化すればするほど、なぜそうした結果が得られたかを理解するのは難しくなります。

「知能」を搭載した機器は、課された仕事をこなすごとに効率を上げていくでしょうが、その作業の目的や意味は、固有の価値領域を有している人間の側が決定し意義をもたせることに変わりないでしょう。何らかの意思決定の基盤となっている基準が不明瞭になり、意思決定の責任の所在が分からなくなり、製作者が共同体の善益のために行動する義務を逃れようとする危険が存在します。ある意味これは、経済がテクノロジーと結託し、効率という基準を優先させ、目先の利益に結びつかないものはことごとく切り捨てる傾向のある、技術主義的なシステムによって助長されています[10]。

ここからわたしたちは、今日の技術主義的でコスト重視の考え方では軽視されがちな、しかし個人と社会の発展にとって決定的な意味をもつ「限界の意識」について考えさせられます。実際、死ぬことを免れえない人間が、あらゆる限界をテクノロジーによって突破しようと考えれば、すべてを支配しようという考えに取りつかれ、自己を制御できなくなる危険があります。際限のない自由を求めて、技術主義の独裁の渦に飲み込まれてしまうのです。被造物として、人間には限界があると認識しそれを受け入れることは、充満に至るため、さらにいえば贈り物として充足を受け取るために、欠いてはならない条件です。逆に、技術主義的パラダイムが思考の土台となる中では、自己充足というプロメテウス的思い上がりが勢いづき、格差が悲劇的に拡大し、情報と富は少数者の手の中で増大し、民主主義社会と平和的な共生は深刻な危機に陥りかねません[11]。

5. 喫緊の倫理的課題

将来的には、ローン申請者の信用調査、個々人の職業適性、有罪判決を受けた人の再犯可能性、政治亡命や社会的援助受給の適格性などが、人工知能システムによって決定されるようになるかもしれません。こうしたシステムによる仲立ちは多角性を欠いているため、さまざまな偏見や差別を露呈しがちです。システムエラーはたちまち増殖し、個々の事案に不正義を来すだけでなく、ドミノ効果によって、社会的不平等の現実形態を生み出しうるのです。

さらに、各種人工知能が、促進や抑止につながるあらかじめ定められた選択肢によって、あるいは情報整理に基づいて個人の選択を規制するシステムによって、個人の意思決定に影響力をもつように思えることもあります。こうした各種操作や社会的統制は細心の注意をもって監視する必要があり、製作者、利用者、政府当局にはそれぞれ明確な法的責任が課されます。

監視システムの拡大や、社会信用システムの導入などによる、個人の分類の自動処理を当てにすれば、市民の間に不当な格付けを定め、市民社会に深刻な影響を及ぼすおそれもあります。そしてこうした人為的な格付けの流れは、仮想の対象者だけでなく生身の人間も巻き込んだ権力闘争にもつながりかねません。人間の尊厳を根本から守るには、その人の唯一無二性が一連のデータに置き換えられることを否定しなければなりません。アルゴリズムには、わたしたちが人権をどのように理解するかを決めさせたり、共感、いつくしみ、ゆるしといった欠かすことのできない価値を無視させたり、個々人が変化して過去と決別する可能性を排除したりさせてはなりません。

こうした状況下では、新たなテクノロジーが就労の世界に与える影響を考えないわけにはいきません。かつては人間の労働力が独占的に担っていた仕事が、人工知能の産業への導入によって急速に奪われています。ここにもまた、多数の人の貧困化を代償にして少数の人が過多な利益を手にする、格差のおそれがあります。労働者の尊厳の尊重、個人・家庭・社会の経済的安定のための雇用の重視、雇用の安定、公正な賃金、これらはこうした各種テクノロジーの職場への導入が深く浸透する中にあって、国際社会の最優先事項とすべきです。

6. 剣（つるぎ）を鋤（すき）にできるだろうか

昨今、わたしたちを取り巻いている世界に目を向ければ、軍需産業にまつわる深刻な倫理問題は避けて通れません。遠隔操作システムによる軍事作戦が可能になったことで、それらが引き起こす破壊やその使用責任に対する意識が薄れ、戦争という重い悲劇に対し、冷淡で人ごとのような姿勢が生じています。人工知能の軍事利用を含む、いわゆる「自律型致死兵器システム」の分野における新規技術の研究は、重大な倫理的懸念となっています。自律型兵器システムは、道義的責任の主体にはなりえません。人間だけが有する道徳的判断力や倫理的意思決定能力は、複雑に集積されたアルゴリズムが及ぶものではなく、その能力をマシーンのプログラミングに落とし込むことは不可能です。マシーンに「知能」が

搭載されているとしてもです。ですから兵器システムに対し、人間による適切で有意の一貫した監視を確保することが不可欠なのです。

高性能の兵器が正しくない人の手に渡り、たとえばテロ攻撃や、正当な政府機関を攪乱する活動を誘発する可能性も無視できません。つまり世界には、戦争の狂気を助長し、武器市場と武器取引の不当な拡大に加担する新規技術など必要ないのです。そんなことになれば、人間の知能ばかりか、心そのものが、いっそう「人工的」になるおそれがあります。最先端技術の応用は、暴力による紛争解決に加担するために用いられるべきではなく、平和への道を舗装するために使われるべきです。

肯定的な面からいえば、人工知能が全人的発展を促進するために用いられるならば、農業、教育、文化において重要な革新をもたらし、国や民族レベルにおいての生活水準を向上させ、人類の兄弟愛と社会的友愛を広げられるはずです。結局のところ、最後に回される人たち、つまりもっとも弱く助けを必要としている兄弟姉妹との共生のためにいかに人工知能を活用するのかが、わたしたちの人間性を明るみに出す尺度となるのです。

世界に対する人道的な視点とよりよい未来への願いから、アルゴリズムにおける倫理、すなわち新技術の進路を規定する価値観となる――アルゴレシックス（algor-ethics）――の確立を目指した学際的対話の必要性が叫ばれています[12]。倫理的懸案については、研究の初期段階から、また実験、設計、製作、流通、販売の各段階でも考察されなければなりません。これは企図に対しての倫理的なアプローチであり、そこにおいては、教育機関や意思決定の責任者が重要な役割を担うのです。

7. 教育の課題

人間の尊厳を尊重し、それに寄与する技術の開発は、教育機関や文化界にとって大きな意味があります。コミュニケーションの可能性を広げることで、デジタル技術は新しい形の出会いをかなえました。ただし、新技術によってわたしたちが引き入れられる関係性とはどういったものであるのか、継続して考察する必要があります。若者は技術が浸透した文化環境で育っていますから、指導や養成のあり方について審議しないわけにはいきません。

各種人工知能の活用に関する教育は、批判的思考の養成に重点を置くべきです。インターネット上で収集された、あるいは人工知能システムが生成した、データやコンテンツを利用する際の識別能力はどの年代においても養われなければなりませんが、とりわけ若者にはそれが求められています。学校、大学、学術団体には、学生と研究者がテクノロジーの開発と活用における社会的・倫理的視点を習得できるよう支援することが求められます。

新しいコミュニケーションツールの利用についての養成は、偽情報やフェイクニュースについてだけでなく、「新しい科学技術を隠れ蓑にして、力をつけていた」「先祖伝来の恐怖」[13]が不穏にも再燃して

いることについても顧慮すべきです。残念なことに、わたしたちは再び、平和的で友愛に満ちた共生の発展と「他の文化、他の人々とのこうした接触を防ぐために、壁の文化を作ろうとする……壁を築こうとする誘惑」[14]と戦わなければならない状況にあるのです。

8. 国際法を拡充させるうえでの課題

人工知能の世界規模の普及から明らかになったのは、その活用を主権国家が自国内で規制する責任と並び、多国間協定を確立し、その適用と実施を整備するうえで、国際機関が決定的な役割を果たすはずだということです[15]。これについてわたしは、多国間共同体が、さまざまな形で人工知能の開発と活用を規制する拘束力のある国際条約を採択するべく、一致して働くよう強く訴えます。当然のことながら法制化は、有害な慣行の予防に加え、好例の推奨、新たな創造的取り組みの活性化、個人や団体による新規参入の促進も目指すべきです[16]。

結局のところ、デジタルテクノロジーの開発者に倫理指針を示すはずの法規範を探るうえでは、必要な法的枠組みの策定、採択、適用という社会の責務の基礎となるべき、人道的な諸価値の割り出しが欠かせません。各種人工知能を製作するうえでの倫理的指針の起草作業では、人間の存在の意味、基本的人権の保護、正義と平和の追求、これらに関する深遠な問いの熟慮をなおざりにはできません。こうした倫理的・法的観点からの識別の過程は、個人生活や社会生活においてテクノロジーが果たすべき役割や、より公正で人間的な世界の創造にテクノロジーがいかに貢献しうるかを、ともに考える貴重な機会となるはずです。ですから人工知能の規制に関する論議においては、世界規模の意思決定プロセスから締め出されている貧困層や社会から疎外されている人々を含め、あらゆる利害関係者の声を考慮すべきなのです。

……

各種人工知能の開発の進展が、最終的には、人類の兄弟愛と平和の大義への確実な貢献となるべく、以上の考察が励ましとなるよう願っています。これは一部の人の責任ではなく、人類家族全体の責任です。平和はまさに、侵すことのできない尊厳を有するものとして他者を認め受け入れるというかかわりの果実であり、すべての人とあらゆる民族の包括的発展を求める協働と努力の果実なのです。

新年の初めにあたり祈ります。各種人工知能の急速な発展が、世界にすでにある、あまりに多くの格差と不正義を増大させることなく、戦争や紛争を終わらせ、人類家族に及ぶさまざまな形態の苦しみを減らす助けとなりますように。キリスト信者、さまざまな宗教の信者、そして善意ある人たちが、デジタル革命による機会をとらえ、そこにある課題に取り組み、連帯のある公正で平和な世界を次世代に残すため、調和のうちに協力していくことができますように。

バチカンにて
2023 年 12 月 8 日
フランシスコ

---

[1] 33。
[2] 同 57。
[3] 教皇フランシスコ回勅『ラウダート・シ（2015 年 5 月 24 日)』104 参照。
[4] 同 114 参照。
[5] 教皇フランシスコ「ミネルバ・ダイアローグの参加者へのあいさつ（2023 年 3 月 27 日)」(訳注：ミネルバ・ダイアローグは、教皇庁文化教育省主催でバチカンにて行われた、デジタルテクノロジーについての科学者や起業家および神学者・哲学者による会議)。
[6] 同参照。
[7] 教皇フランシスコ「ダボスでの『世界経済フォーラム』議長へのメッセージ（2018 年 1 月 12 日)」参照。
[8] 『ラウダート・シ』194、「教皇庁主催『デジタル時代の共通善』セミナー参加者へのあいさつ（2019 年 9 月 27 日)」参照。
[9] 教皇フランシスコ使徒的勧告『福音の喜び（2013 年 11 月 24 日)』233。
[10] 『ラウダート・シ』54 参照。
[11] 教皇フランシスコ「教皇庁生命アカデミー総会に寄せたメッセージ（2020 年 2 月 28 日、代読)」参照。
[12] 同参照。
[13] 教皇フランシスコ回勅『兄弟の皆さん（2020 年 10 月 3 日)』27。
[14] 同。
[15] 同 170－175 参照。
[16] 『ラウダート・シ』177 参照。

（カトリック中央協議会事務局訳）