

潮流を読む

AIが社会全体に及ぼすリスクへの懸念の高まり

生成AIという高度AIテクノロジーによって、人間のように自然な会話ができるAIチャットサービス「ChatGPT」が、オープンAI社から2022年11月に公開され、無料で利用できる革新的なサービスとして注目された。さらに23年3月には、より高度な性能を有する「GPT-4」が発売され、さまざまな企業・個人がChatGPTのAPIを活用してサービス開発を試みている。

その一方、AIが社会全体に及ぼすリスクへの懸念が急速に高まっている。23年5月にはサンフランシスコに拠点を置くAIの安全性に関する非営利団体CAIS (Center for AI Safety) が、「AIによる（人類）絶滅のリスクを軽減することは、パンデミックや核戦争などの他の社会規模のリスクと並んで世界的な優先事項であるべきです」との声明文を出した。そこには前述のオープンAI社のCEO、グーグル社のCEOなど、多くのAIに携わる著名な関係者が署名している。

他方、23年5月に開催された広島でのG7サミットの共同宣言（G7広島首脳コミュニケ、外務省仮訳）では「我々は、急速な技術革新が社会と経済を強化してきた一方で、新しいデジタル技術の国際的なガバナンスが必ずしも追いついていないことを認識する」とした。さらに「我々は、関係閣僚に対し、生成AIに関する議論のために、包括的な方法で、OECD及びGPAI[注1]と協力しつつ、G7の作業部会を通じた、広島AIプロセスを年内に創設するよう指示する」としている。具体的なアクションが国際協調の中でどのように実行に移されていくのかが注目される。

この点について欧州はかなり先んじた動きをしている。6月14日には欧州議会は、AI法に関するEU加盟国（欧州委員会と欧州閣僚理事会）との交渉上の立場について、同法に賛成することを採択した。同法に対して、賛成499票、反対28票と賛成票が圧倒的であった。しかし、棄権も93票

あり、他国とのAI開発競争で劣後することなど、複雑な議員の胸中が読み取れる。この理由として、同法は欧州で開発および使用されるAIに限定していることが挙げられよう。その一方、現実のAIシステムの導入では「人間の監視、安全、プライバシー、透明性、無差別、社会的および環境的福祉を含むEUの権利と価値観」（当該AI法）に反するケースが多く見られることも事実である。これらの事例が権利と価値観に完全に合致している保証を受ければ、欧州では正しいAIの導入という側面で先んじて優位に立てると期待できる。

さらに欧州議会の先導的な施策としては、当該AI法に定められる規則はリスクベースのアプローチを採用していることも挙げられる。つまりそのリスクのレベルに応じて、プロバイダー（AIシステムの提供者）とそれを導入する個人・企業などの当事者の義務を定めているということだ。既に欧州議会は、禁止リストを列举[注2]して、AIによるプライバシー侵害および差別的使用の禁止行為を列举している。

ただし、これらの禁止リストは、裏を返せば、他国では既にAIシステムが導入されている事例といえる。グローバルにシステムがオープン化されている現在、リスク・アプローチとはいえ、他国でAIシステムが導入されているかどうかの判断は難しく、規制することは困難を極めると考えられる。やはり国際協調がなければ抜け穴だらけの規制になりかねない。世界は一段と分断の道を進み、AIによる社会全体に及ぼすリスクへの懸念が高まっている中、G7の広島AIプロセスの早急な具体的なアクションが待たれる。他方、AIが人類を滅ぼすとは人類のおごりであり、結局、それを利用する人類の倫理・道徳の欠如が人類を滅ぼすといえるのではないだろうか。

（6月19日執筆）

[注1] 人工知能グローバルパートナーシップ（GPAI）

[注2] 「公的にアクセス可能な空間における『リアルタイム』遠隔生体認証システム」「『事後』遠隔生体認証システム。ただし、重大な犯罪の訴追のための法執行機関を除き、司法の許可があった場合に限定」「センシティブな特性（性別、人種、民族、市民権ステータス、宗教、政治的指向など）を使用した生体認証分類システム」「予測警察システム（プロファイリング、場所、または過去の犯