生成AIのプライバシー侵害リスクと規制

ChatGPTに代表される生成AIは様々な領域で生産性向上に寄与すると言われているが、一方でその能力が悪用される可能性も高い。特にプライバシーに関してAIが悪用されることのリスクへの対処は急務である。そして各国で「AI規制」の議論が進んでいる。AI利用企業はグローバルの規制動向に適切に対応する必要があるだろう。

生成AIの急速な進歩による負の側面

ChatGPTに代表される生成AIが登場してそろそろ一年になる。この間、様々な領域で生成AIは驚異的な進歩を示した。文章生成・要約、画像生成、音声認識、音声生成（作曲含む）、さらには動画認識といった領域での各種の生成AIの能力向上には目を見張るものがある。

これらの生成AIの能力向上は、生産性に劇的な向上をもたらす可能性が示されている[1)]。一方で、この生成AIをマイナスの方向に利用することもできる。

マイナス方向の活用として、特に懸念されるのが生成AIによるプライバシー侵害リスクだ。そしてこのリスクに対して様々な対策が検討されている。

生成AIによるプライバシー侵害リスク

生成AIによるプライバシー侵害リスクには主に3つが考えられる。一つは「大規模なデータの収集・分析によるデータの個人特定につながる紐づけ」、そして「収集されたデータの漏洩・不正使用」、最後に「データに内在するバイアス・偏見の増幅」だ。

基本的に生成AIはネット上に存在する様々なデータを学習用のデータとして活用する（これらのデータの利活用に関する著作権の取り扱いについての懸念もあるがここでは一旦措く）。一見関係がないようなデータでも、特定の関連要素を推測し、それらの雑多なデータを紐づけることで、個人の行動や特性が推測される可能性がある。例えば、個人の購買履歴とSNSでの発信が紐づいた場合、勝手にAIがSNS上であたかも本人がやっているようなインフルエンサー的行動に出ることも考えられる。

そして、このように収集されたデータが適切に保護されているのかも問題となる。現在の生成AIが収集・分析しているデータセットは膨大な量だが、この膨大なデータのうちに個人データが含まれないとは言えない。実際、我々の個人データを大量に保有しているGAFAによる生成AIはすでに存在している。これらのデータセットが漏洩した場合、大規模な個人データ侵害が発生しかねない。

そして、最も懸念されるのは個人データが様々に紐づけられることによって、そのデータに内在するバイアスや偏見が生成AIによってブラックボックス化した上で社会に影響を及ぼすリスクだ。以前、Amazonが人事採用の意思決定にAIを活用したところ、顕著に男性に有利な判定をしたことが問題となった[2)]。この人事採用AIは過去のAmazonの実際の採用実績のデータをもとに学習を行っていたため、「男性従業員が多い」という現状を「男性は優秀である」と誤って判定していた可能性が高いといわれている（Amazonはすぐに利用を停止した）。

既存のデータを学習データとして活用するAIには、このようなバイアスをブラックボックス化するリスクが常に存在する。

AI規制を巡る各国の動向

このような生成AIが生み出しかねないリスクを未然に防ぐために各国でAI規制論が活発化している。特に活発なのがEUだろう。

過去にEUは個人情報保護に対して、EU一般データ保

NOTE

1) 「生成AIは世界経済に年間620兆円の価値を加えるとマッキンゼーが報告」GIGAZINE
https://gigazine.net/news/20230615-mckinsey-report-generative-ai-global-economy/
2) 「焦点：アマゾンがAI採用打ち切り、「女性差別」の欠陥露呈で」ロイター
https://jp.reuters.com/article/amazon-jobs-ai-analysis-idJPKCN1ML0DN
3) 「EUのAI規制法案の概要」総務省
https://www.soumu.go.jp/main_content/000826707.pdf
4) 「乗り遅れた日本、生成AIを巡る日米欧中の規制動向：日米／欧中で方向性の違いが明確に」EE Times Japan
https://eetimes.itmedia.co.jp/ee/articles/2307/27/news088.html
5) 文化庁令和5年度著作権セミナー「AIと著作権」文化庁著作権課 (pp.30-40)
https://www.bunka.go.jp/seisaku/chosakuken/seidokaisetsu/seminar/2023/pdf/93903601_01.pdf

護規則（GDPR）を制定することで、個人情報保護のグローバル規制の主導的立場に立つことになった。今回、EUはAIに関する規制でも同様の動きを見せているように思われる。

現在、EUで議論されている「AI規制法案」は、大きく3つの特徴を持つ[3]。一つはリスクベースアプローチ（リスクの大小によって規制の軽重を決める）、2つ目はGDPRでも採用した域外適用（EU市民に影響があればEU域外にも規制が適用される）、最後に違反した場合、多額の制裁金やEU域内での事業停止といった大きな規制逸脱リスクが存在することだ。

アメリカでは一部の巨大ITプレーヤーによるネット上の情報空間の独占・寡占がそもそも問題視される中、AIによるさらなる「囲い込み」が生じることを問題視する議論が起きている。さらに、ChatGPTを開発したOpenAIのCEOサム・アルトマン氏をはじめとして、経済界やアカデミアからも「AIの暴走を抑止する世界的な規制が必要だ」と訴える意見も出てきている。

中国では「共産党の価値観に沿ったAI発展」を目的とした規制案が議論されている。中国にとってAIは「国家主導のテクノロジー政策」であるべきとの意見が根強い。2023年8月に施行された「生成人工知能サービス管理暫行弁法」では、生成AIの利用目的の開示と国への届け出が義務付けられた。違反すれば高額な罰金と制裁金が課される可能性がある[4]。

一方、日本ではAIに対する直接的な規制は現時点では具体化されていない。また、日本では著作権法第30条の4において、AIが著作物を学習データとして利用する場合は著作権侵害とはみなさない、という規定を設けている[5]。そのため、日本の著作権の保護下にある著作物をAIの学習データに用いることは世界的に見ても「緩め」の規定であるとも言える。

岸田総理は2023年10月に、生成AIの規制のあり方などを議論する「広島AIプロセス」を提唱している。この提案は、比較的自由度の高いAI規制の枠組みを提案しているのだろうと筆者は考えている。

説明可能AI（XAI）とAI倫理の重要性

AIの進化はトータルで見れば人類にとってプラスになることは間違いないだろうが、あらゆる技術には負の側面が付随する。そのような認識のもと、AI領域では「説明可能AI（Explainable AI：XAI）」というコンセプトが注目されている。AIが学習したデータセットを明らかにし、その学習結果をブラックボックス化せずに、AIが返してきた返答に対する判断基準を明示するような枠組みを設けようという機運が高まっている。AIによる「回答」に対する「結果責任」を説明できる仕組み・枠組みが今後要求されるようになる可能性は高い。

このような背景を踏まえると、AIを活用する企業は、自社のAI利活用ガイドラインを策定することはもちろんのこと、ガイドラインの遵守を徹底することが求められる。さらに規制の動向をにらみつつ、ガイドラインを随時アップデートしていくことも必要となるだろう。

Writer's Profile

柏木 亮二 Ryoji Kashiwagi
金融ITイノベーション事業本部
エキスパートリサーチャー
専門はIT事業戦略分析
focus@nri.co.jp