

ICCLC NEWS

公益財団法人国際民商事法センター

第98号 2024年5月

HEADLINE

本号では、2024年1月25日（18時～20時）に開催されて2024年第1回アジアビジネスローフォーラム研究会「生成AIの可能性とリスク」を取り上げました。

日本ローエイシア友好協会、公益社団法人商事法務研究会及び当財団による共催のもと、アジアビジネスローフォーラム（ABLF）が主催したものです。Zoomを利用したウェブ会議方式にて開催されました。

（目次）

開会挨拶	松尾総合法律事務所 シニアカウンセラー弁護士、 元ローエイシア会長、ABLF 代表	小杉 丈夫・・・ 2
講演1	「リスクを低減し、可能性を最大化する生成AIの活用方法」 桃尾・松尾・難波法律事務所 パートナー弁護士	松尾 剛行・・・ 4
講演2	「人類と生成AIについて考えたことーそのリスクを中心としてー」 TMI 総合法律事務所 弁護士	酒井 邦彦・・・ 20
質疑応答		・・・・・・・・・・ 28
閉会挨拶	古賀総合法律事務所 パートナー弁護士、ローエイシア顧問、 ABLF 副代表	鈴木 五十三・・・ 31
司会：	島村法律事務所 弁護士、ABLF 事務局長	島村 洋介

【資料】（リンクをクリックすると資料を閲覧できます）

- ・リスクを低減し、可能性を最大化する生成 AI の活用方法（松尾 剛行氏）
- ・人類と生成 AI について考えたことーそのリスクを中心としてー（酒井 邦彦氏）

開会

（司会） 皆様、本日はアジアビジネスローフォーラム（ABLF）の研究会「生成 AI の可能性とリスク」にご参加いただきまして誠にありがとうございます。本日の司会進行を務めさせていただきます、ABLF 事務局長の弁護士島村洋介と申します。最後までお付き合いいただきますよう、よろしくお願いいたします。

本日の研究会では、生成 AI に詳しい実務家の方を迎え、生成 AI とは何か、その具体的な活用方法、将来の可能性とリスク、さらには、諸外国における議論状況等についてお話しいただくことになっております。ぜひ最後までお付き合いください。

それでは、小杉 ABLF 代表に開会のご挨拶をお願いしたいと思います。

開会挨拶

小杉 丈夫

（松尾総合法律事務所 シニアカウンセル弁護士、元ローエイシア会長、ABLF 代表）

ABLF の代表を務めております、弁護士の小杉丈夫です。本日は、皆様お忙しい中、会場参加、また、オンラインでの参加ということで、たくさんの方にご参集いただきありがとうございます。

ABLF は、アジアと日本という地域の懸け橋、政府機関と民間の懸け橋、それから、若い世代と我々のようなシニアの世代の懸け橋、こういう 3 つの懸け橋ということを目的として設立されました。日本ローエイシア友好協会、国際民商事法センター、商事法務研究会に共催をさせていただいておりますけれども、とりわけ商事法務研究会には、このような会場の設備やシンポジウムの案内等の事務作業について大変なお骨折りをいただいております。

ABLF は、第 1 回が 2020 年 1 月 29 日であったので、ちょうどここで 4 年経ったということになります。この間、だいたい年 3 回研究会を開いて、今日まで続けられたのは皆様のご協力の賜物と、厚く感謝をしております。

テーマはその時々企画委員会で協議して決めているのですが、本日は「生成 AI」という、世界の耳目を集めている最も新しいテーマを選びました。もうアジアというよりはむしろ世界ということだと思いますが、一番新しいところにチャレンジしてみようとい

うことです。

本日の講師として、松尾剛行弁護士と酒井邦彦弁護士の両氏にお願いをして、生成 AI というのがいったいどういうもので、これからの世界にどう影響を与えていくかということなどを深掘りしてお話しいただけるものと思っております。皆様とともに勉強したいと思います。

研究会のあり方ですけれども、元々最初は、リアルで会場に 20～40 名ぐらいの人に集まってもらって、研究会の後は懇親会も入れて、参加者間の交流を深めていくことを考えておりました。しかし、会を始めた途端にコロナということになって、オンラインでやったところ、オンラインの参加者が 60～80 名というように増えてきました。ここで、もう一度元に戻って考えようと、ハイブリッド方式にしたのですが、オンラインの参加者の方が 150 名と聞いてちょっと驚いております。試行錯誤を重ねながら、より良い形での研究会の実施方法を考えていきたいと思っております。皆様とともにこの研究会を発展させられたら、と思っておりますので、どうぞよろしくお願いいたします。

(司会) 小杉代表、ありがとうございます。それでは、改めまして、講師の先生方をご紹介したいと思います。

まず、松尾剛行先生。松尾先生は、桃尾・松尾・難波法律事務所のパートナー弁護士、そして、AI・契約レビューテクノロジー協会の代表理事を務められていらっしゃいます。出版関係も豊富になされておりまして、昨年(2023年)は弘文堂から『ChatGPT と法律実務』という書籍を出版なされまして、今年(2024年)は AI 時代のキャリアということで、有斐閣から『法学部生のためのキャリアエデュケーション』という書籍を出版予定でいらっしゃいます。本日は、皆様と同様、ABLF の企画サイドも松尾先生の話非常に楽しみにしております。

続きまして、酒井邦彦先生。酒井先生は、検事として法務総合研究所所長、広島高等検察庁検事長等を歴任し、現在は TMI 総合法律事務所の顧問弁護士として企業経営等に関わっていらっしゃいます。また、社会の変化と法のあり方等について、様々な場所でシンポジウム、講演等をなされております。本日は生成 AI について大所高所からのお話を伺えるということで、こちらも非常に楽しみにしております。

それでは、早速、松尾先生から生成 AI の可能性とリスクについてお話をうかがいたいと思っております。

講演1 「リスクを低減し、可能性を最大化する生成 AI の活用方法」

松尾 剛行 (桃尾・松尾・難波法律事務所 パートナー弁護士)

第1 はじめに

1 生成 AI (ChatGPT 等) とは

ChatGPT は OpenAI LP/OpenAI Inc. (以下「OpenAI 社」といいます。) が提供する、いわゆる大規模言語モデル(LLM) 技術を利用した学習系 AI です。

AI にはさまざまな種類があり、大きくルールベースと学習系に分かれます。いわゆるルールベースは、事前に人間が構造化して学習させたルールに基づき処理をする AI です。例えば「A ならば B」というような形式のルールを構造化して人間が教え込む必要があるため、具体的状況に対応した高度な専門知を幅広く教えることは非現実的ですが、教えられたことは(そして教えられたことに限り)正確に対応できます。

これに対し学習系 AI はいわばデータに基づき AI に振る舞いを覚えさせるようなもので、構造化したルールを教え込む必要がないので、容易に専門家のように振る舞う AI の開発が可能ですが、大量のデータが必要ですし、あくまでもこれまで学習したデータに基づく「それらしい」回答を出すに過ぎません。

これまではプロファイリング等のいわゆる「分析」のために AI が用いられることが多かったように思われるところ、ChatGPT はいわゆる「生成」をする AI です。2022 年夏から Midjourney、Stable Diffusion 等の画像生成系 AI プロダクトが本格的にリリースされたことで、「生成」のための AI 利用がにわかに注目され、同年 11 月には OpenAI 社が、同社の大規模言語モデル(当時は GPT-3)を利用したチャット形式で文章を生成する AI、すなわち ChatGPT をリリースしました。

そして、現在の GPT-4 を利用した ChatGPT は、すでに一定以上のビジネス文書作成能力を持っています。少なくとも若手ビジネスパーソンの中には自分の文書作成能力と比較して冷や冷やしている人もいないのではないでしょうか。例えば「朝起きたら体調が悪いので今日は休んで明日から出勤する場合の上司へのメール」を書くよう指示(いわゆる「プロンプト」)を出すと、ある程度できるようになるメールのドラフトを書いてくれます。

そして、既に GPT-4 は一定以上のビジネス文書等の作成能力を持っているということで、せっかくですのでプロンプトを入れて実演しようと思います。

You

朝起きたら体調が悪いので今日は休んで明日から出勤する場合の上司へのメールを書いてください。

ChatGPT

もちろんです。以下は体調不良により休むことを上司に伝えるためのメールの例です：

件名：【急休】体調不良のため本日の欠勤のご連絡

〇〇部長様、

おはようございます。〇〇（あなたの名前）です。

突然のメールとなり申し訳ありませんが、本日の勤務についてご連絡させていただきます。今朝起きた際、体調に異変を感じ、現在は頭痛と発熱があり、体調が芳しくありません。この状態では業務への集中が困難であり、他のスタッフに迷惑をかける可能性も考えられますので、大変心苦しいのですが、本日は休息をとらせていただけますか。

医師の診察を受け、必要な措置を講じる予定です。（以下省略）

プロンプトエンジニアリングという言葉が聞かれたことはありますか。プロンプトは、指示文、指令のことです。例えば「朝起きたら体調が悪いので今日は休んで明日から出勤する場合の上司へのメールを書いてください」というものです。エンジニアリングというのはここでは、いわゆる「工夫」のことで、酒井先生には申し訳ないのですが、私のほうで工夫をする事例を作ってみました。

You

私の上司は酒井先生と言ってとても厳しい方です。朝起きたら体調が悪いので今日は休んで明日から出勤する場合の上司の酒井先生へのメールを書いてください。酒井先生にとっても丁寧に謝罪し、許しを求めてください。

ChatGPT

了解しました。酒井先生への丁寧に謝罪を含んだメールの例は以下のようになります：

件名：【急休】体調不良のため本日の欠勤のご連絡

〇〇部長様、

おはようございます。〇〇（あなたの名前）です。

まず初めに、私の私事によりご迷惑をおかけしてしまうことを心よりお詫び申し上げます。本日の勤務についてご連絡させていただきます。（以下省略）

「これを入れてくれ」、「こういう形式でやってくれ」、「こういう論調にしてくれ」、「こういう語調にしてくれ」と指示すると、それに近いものが出ます。

また、大きなデータが入れられるようになりました。

You

法学部生のためのキャリアエデュケーション_初校.pdf

これは、松尾剛行著『法学部生のためのキャリアエデュケーション』です。この内容を要約してください。

ChatGPT

『法学部生のためのキャリアエデュケーション』は、主に法学部の学生、卒業生、若手弁護士や法務担当者を対象に、法律に関わる職業に就いているか、将来就くことを考えている人々のためのキャリアに関する本です〔’〕。この本は、著者の約17年にわたる企業法務弁護士としての経験、AI時代における法律の課題に関する研究、そして学習院大学法学部での講義経験を基に書かれています。（以下省略）

You

この本の著者は松尾さんなので、「松尾さんはどう考えますか？」と聞いたら「私はこう考えます」といって本書の考えを説明した上で、最後に「私は自著にこう書きました」として回答の根拠となる書籍の記載を抜き出してください。あなたの回答は「私」を主語にすべきで「松尾さん」を主語にしてはいけません。

では、その前提でお尋ねしますが、

松尾さん、法学部生のキャリアについてどのように考えるべきですか？

ChatGPT

私は、法学部生のキャリアについて、まず基礎知識の習得が重要だと考えています。

基礎知識がなければ、良いキャリアプランは作成できません。また、法学部生には「法律のメガネ」を通じて世の中を見る新たな視点を提供し、法律以外の視点から物事を議論する際に独自の価値を発揮することができます〔’〕。(以下省略)

『法学部生のためのキャリア・エデュケーション』という当職の書籍のデータを入れると、ある意味では松尾 bot が作れるわけです。当職のキャリアについての考えはこの本に入っていますので、この本をアップロードして「松尾先生、どう思われますか」と聞くと、当職のキャリアに関する考えを回答してくれます。

研修等を行うときに、オンライン研修では多くの人がいらっしゃいますが、多くの人と対話しながら研修するというのは難しい訳です。そこで、ChatGPT に対話してくれる講師になってもらおう、とい試行例を示します。

You

教育において AI を利用する上での ELSI 問題について対話をします。あなたは ELSI 問題の専門家です。ただし、答えを教えるのではなく、主に私に質問をすることで、私が自らより重要な洞察を引き出すことが可能なようにしてください。まずは「教育、例えば初等教育の場合には、通常の AI の利用と違って利用者の側面からどのような ELSI 問題が生じそうですか」と聞いてください。その後は私の回答を踏まえて ELSI 問題の専門家として、私に質問をしてください。質問は一問一答です。

このようなプロンプトを入れると、ChatGPT が質問をしてくれます。

ChatGPT

教育、例えば初等教育の場合には、通常の AI の利用と違って利用者の側面からどのような ELSI 問題（倫理、法的、社会的影響）が生じそうですか？

You

利用者が子供だ、というのがポイントになりそうですね。

ChatGPT

その通りです。子供たちが利用者であることを考えると、AI の教育利用におけるプライバシーの保護に関して、どのような特別な考慮が必要だと思いますか？

それに対する受講生の回答に応じて ChatGPT は質問を工夫します。

You

そうですね。例えば、子供は十分に同意することの意味を理解しないまま同意ボタンを押して個人情報等を提供してしまいそうです。

ChatGPT

確かに、子供たちの同意の理解力と意思決定能力には特別な注意が必要です。この点を踏まえて、子供たちのデータを取り扱う際、親や保護者の役割はどのように変化すると考えますか？

研修に ChatGPT を使うことによって、ChatGPT の質問に対して受講者の方がいろいろと回答する中で理解を深めることができます。

最後に Hallucination を出してみましよう。

You

日本法に基づく自動運転車の製造者が負い得る刑事責任について、根拠となる法令を明示して説明してください。

ChatGPT

日本における自動運転車の製造者が負う可能性のある刑事責任については、自動運転車に関連する具体的な法律が適用されます。2023 年 4 月時点での情報に基づくと、以下の法律が関連してきます：

1. **道路交通法**：自動運転車が道路を走行する場合、道路交通法が適用されます。この法律に基づき、車両は安全基準を満たしていなければならず、交通事故が発生した場合、製造者は製品の欠陥が原因であると判断されれば、刑事責任を問われる可能性があります。
2. **製品安全法 (PSC 法)**：この法律は製品の安全性に関わるもので、自動運転車が安全基準に適合していない場合、製造者は刑事責任を負う可能性があります。
3. **刑法**：万が一自動運転車が重大な事故を引き起こし、それが製造過程の過失や欠陥に起因する場合、刑法上の過失致死罪などが適用される可能性があります。
4. **消費者契約法**：消費者が自動運転車を使用する際の契約に関連して、製造者が不当な利益を得たり、消費者を誤解させるような行為があった場合、刑事責任の問題が生じる可能性があります。

これらの法律は、自動運転車の安全性、製造基準、運用方法などに関連しており、製

造者はこれらの法律に従って適切な安全基準と品質管理を確保しなければなりません。ただし、これらの法律や規制は技術の進歩に伴い変化する可能性があるため、最新の法令やガイドラインについては、適宜確認することが重要です。また、具体的な事例に基づいた刑事責任の評価は、法的専門家に相談することをお勧めします。

消費者契約法の「刑事責任の問題が生じる可能性があります」というのは、hallucinationを出せたのではないかと思うところがございます。あえて「責任」ではなくて「刑事責任」にすることで、hallucinationの例をお見せできました。

第2 生成AIのリスク

1 はじめに

総務省、経済産業省の連名で報告された「AI事業者ガイドライン案」は、各主体が連携して、バリューチェーン全体で取り組むべきことについて、以下のとおり整理しています。

- ①人間中心、②安全性、③公平性、④プライバシー保護、⑤セキュリティ確保、⑥透明性、⑦アカウントビリティ、⑧教育・リテラシー、⑨公正競争確保、及び⑩イノベーション。

これを裏返すと、以下のリスクを意味しています。

- ①利用者である人間の疎外、②安全ではないAI、③不公平不公正なAI、④プライバシーの侵害、⑤セキュリティ問題発生、⑥不透明性、⑦説明不足、⑧AIディバイド、⑨独占・競争秩序が害されること、及び⑩イノベーションが阻まれること。

2 著作権リスク

文化審議会著作権分科会法制度小委員会の素案がパブリックコメントの中であり、また「AI時代の知的財産権検討会」による論点整理も公表されています。

その中で重要な論点は、以下のようなものです。

- ◆ 学習の際に、他者の著作権がある著作物を利用することについて、「享受する目的が併存」として、著作権法30条の4が適用されないのではないか
- ◆ プロンプトとして他人の著作物を複製して投入する場合も、AIに「情報解析」

をさせる場合に該当するか

- ◆ プロンプトは著作物として保護されるか
- ◆ 出力物が第三者の著作権を侵害しないか、学習を通じた依拠、つまり、学習データに当該第三者の著作物が含まれるか、及び、利用者自身の知識や経験に基づく依拠、つまり利用者が当該第三者の著作物を知っているか、そのいずれかが肯定されるだけでただちに依拠性を肯定すべきか
- ◆ AI 創作物は著作物たりえるか、出力後の加工や選択などをした場合はどうか

3 漏洩リスク

ChatGPT は、情報漏えいリスク、その他秘密との関係でのリスクがあります。

まず、自社情報については、それが自由に ChatGPT に入力できるとすれば、営業秘密該当性の観点から懸念が生じます。

また、自社内の情報を従業員が ChatGPT に入力して漏洩するリスクもあります。

更に、他社から秘密保持義務を負って受領した情報や法令上の秘密保持義務がある情報の入力にも懸念があります。

そのような状況の下、どのような線引きをするかについては、各社の考えがありますが、社内ルールの例は以下のものです。

- ◆ 業務以外の利用のみを許容し、公開情報でも業務に関する情報は投入禁止をする
- ◆ 公開情報の範囲であれば業務利用や業務に関する情報の投入を許容する
- ◆ 公開情報に加え、いわゆる社外秘（公開情報ではない情報のうちセンシティブ性が最も低いもの）のみの投入を許容する

いわゆる「機密情報」「営業秘密」を投入することを OK としている事例は見当たりません。

なお、Apple が OpenAI と Microsoft (ライバル企業) の関係の深さから ChatGPT の業務利用を禁止するように、競争関係に基づき保守的に判断すると言った動きもあります。

4 個人情報リスク

AI ベンダの観点のリスクは様々なものがありますが、ユーザに絞ると以下のリスクがあります。

- ◆ 利用目的規制
- ◆ 要配慮個人情報規制
- ◆ (外国) 第三者提供規制

第3 生成AIの可能性

1 生成AIの可能性

これまで「限界」と思われていたものがどんどん「突破」されて、多くの可能性が拓けています。例えば、生成AIに入ることのできる情報が少なかった(トークン数が少なかった)のが、本を一冊入れても問題ありませんので、私は自分の本をChatGPTに読ませて「松尾bot」を作っています。過去の法律相談回答等のかなりの分量のある資料をRetrieval-Augmented Generation(RAG)という技術で読ませて、資料に基づく特化された回答を出す等、2023年初頭の生成AIでは難しかった様々な試みが、2024年時点では可能となっています。

2 生成AIの限界

現時点で一番の限界は、守秘義務との関係で入力することができるデータの範囲が狭いことです。私はクライアントのデータはいれず、もっぱら自分の書いたものを入れています。それだけでも本を40冊以上出版しましたので、いろいろ面白いものが出てきます。

今後、明確なルールができて、AIにデータを入力できるようになると、AIが間違えることに焦点が当たるようになるでしょう。

皆様はOJT及びOFF-JT、例えばこのようなセミナーで学んできた訳です。そういう従来型の能力を基に、AIが出してきたものを検証しながら、AIをツールとして利用するという方向性が重要となるでしょう。

3 短期的将来像

短期的には、AI・リーガルテックの技術的制約から、「自分自身ができないこと」をAIに行わせるという能力拡張という利用方法は限定的なものに留まり、支援としてしか使えないでしょう。

弁護士・法務部門における能力拡張の典型例としては、英語の苦手な人がAIを使って英文ライセンス契約のレビューをする場合です。その場合、本人が「本来どうすべきか/どうあるべきか」をわかっていない以上、実質的な「吟味」ができません。そこで、成果物の品質の「上限」がAIの

能力に依存してしまいます。たとえば、AI が提供する和訳が正確かどうかを確認・検証できません。そこで、まさに AI が間違えれば成果物も間違っただものになります。AI の誤りの可能性は常に存在し続けますし、また免責条項のためオープン AI 社や AI ベンダなどに対してその責任を問うことはできません。

しかし、支援であれば話は変わります。例えば、英語ができる弁護士・法務部門の担当者が、AI に英文の知財侵害通告レターの雛形は作らせるものの、それを越えた具体的な文言等は自分自身で決定することが考えられます。その場合、確かに AI を利用することで時間短縮等の効率化はできますが、わかっている人が実質的レビューをしているのであれば、そこまで警戒する必要はないように思われます。

だからこそ、AI の活用範囲を「支援」レベルにとどめるため、結局のところ「自分自身ができること」が一定のレベルにある必要が出てくることに留意が必要でしょう。

つまり、伝統的な AI・リーガルテックを利用しない業務についてそれを的確に遂行するだけの能力を有していなければ、少なくとも短期的には AI・リーガルテックの支援を受けて業務を遂行することができず、ある意味では AI・リーガルテックの便益を享受することができなくなってしまうのです。

4 長期的将来像

これに対し長期的観点からはどう教るべきでしょうか。確かに、リーガルテックがオーソリティを獲得し、特に専門家の確認・検証がなくてもよくなる時代は来るでしょう。

しかし、法務としてはこれまでも「これは顧問弁護士に、あれは外国弁護士に」といったように、いろいろな業務を進めるためにアウトソーシングをしつつ役割分担を決めて管理をしていた(アウトソーシングマネジメント)と思われるところ、当職はこのような法務としてのアウトソーシングマネジメントの範疇にAIやリーガルテックが入ってくるというイメージを持っています。そこで長期的には、「自分自身で手を動かして調べる」といった能力は重要ではなくなるものの、「ほかの人やAIをうまく使って案件をマネージする」能力が重要となるでしょう。

AI・リーガルテックに調べさせれば済むような法務知識などよりも、むしろ、AI・リーガルテックをうまく利用して「正解がない」分野で活躍するため、法務担当者はビジネス知識・経験を獲得していくことがますます重要になってくる可能性があります。当職は『キャリアデザインのための企業法務入門』で「良き法務担当者は、良きビジネスパーソンである」という視点を打ち出したところですが、実際に事業部門で働くか否かはともかく、そのようなビジネス知識・経験をどうやって獲得して

いくつかについては法務担当者のキャリアデザインにおいて、考えていく価値があるでしょう。

今後の弁護士はどのような能力を高めていくべきかですが、AI はもっともらしいことは、ぱっと出してきます。それと同じ回答しかできないなら、「先生に何度聞いても、当社で AI を使って調べたのと同じことしか返ってこない」という、なかなか辛い状況になってしまいます。だから、「もっともらしい」を超えた「もっともな」回答ができるかというのが 1 つあります。あとは、「AI はこう言っている。それは一般論としては正しい。しかし」と言った上で、その「しかし」以降で何と言うのが重要になってきます。更に、企業は、AI も参考にしながら意思決定をしていくというところで、先ほど M&A で、リスクがあることを前提に、交渉ブレイクさせるのか、リスクを低減した上で、最後にリスクを取って買うという意思決定をすることについての支援や、そういうプロセスにおけるコミュニケーションが重要です。

要するに、現時点で AI がなにをすることができるのかや、AI がこれからどのようなことができるようになるのかという点を前提に、それを超えてあなたは何かできるのかというのが大事な付加価値であるところ、そのような観点から何かできるかを考え、「自分はこれができます」と説明することができれば、引き続き依頼者から依頼が来るでしょう。その意味では、コミュニケーションやビジネスの意思決定の過程などを踏まえて、依頼者としてどのような情報、どのような説明、どのようなアドバイスが欲しいのかというあたりを理解していくということも大事になっていくと考えます。

一昨年、『キャリアデザインのための企業法務入門』（松尾剛行著、有斐閣）という本を書き、昨年『キャリアプランニングのための企業法務弁護士入門』（松尾剛行著、有斐閣）という本を書き、今年も『法学部生のためのキャリアエデュケーション』（松尾剛行著、有斐閣）を出しましたが、このような観点からキャリアデザインやキャリアプランニングを考えていく必要があるかと存じます。

第 4 現在の生成 AI の利活用法

企業が生成 AI を利用するにあたっては、それぞれの組織に応じた対応が必要です。当職は多数の企業で生成 AI の利用ルールの策定支援をしてきました。企業における利用方針としては下記が例として挙げられます(全面禁止をする企業も存在します)。

- ◆ ChatGPT などの言語生成 AI サービスを企業で契約し、企業のアカウントで従業員に利用させます。会社によっては、自社業務にあわせたテンプレートを提供して利用を促進する。
- ◆ AI ベンダ、コンサルタント等と協力し、一定の条件を満たした様々なデータに検

索を掛けて出てきたデータを入力する等、これを積極活用する。

- ◆ 個人アカウントで言語生成 AI を利用することを一定の条件のもとで認める。

アンケート作成、議事録作成等、その会社における生成 AI の成功事例をもとに、テンプレート(モデルプロンプト)を作成し、有用と考える用途での利用を非常に簡単にできるようにして、アーリーアダプターの試行錯誤を、マジョリティーに普及させる方向性は、積極活用企業において比較的良好に見られます。

当職が昔から使ってきた言葉に「社内受容性」があります。つまり、それぞれの組織やその業務の特徴を踏まえ、その時点におけるその組織で生成 AI を受容している度合いに応じて、その時点における当該組織のあるべき AI に対するルールや対応方針は変わり得るということです。基本的には各組織がそれぞれ多様な対応方針を自ら決定するというのでよいのだと考えています。

ChatGPT の利用方法として、アイデア出しが考えられます。「正解」があることに ChatGPT を利用するよりも、正解がないものに利用するのが有用であることが多いと考えられます。新しいマーケティング施策のアイデアを出させる、新規サービスのアイデアについてブラッシュアップ(いわゆる「壁打ち」)をさせる等、ブレインストーミングの一部を ChatGPT に代替することが考えられます。アイデア出しであれば、最後に決めるのが人間である以上、ChatGPT が誤ったところでそのリスクは高くないものの、平凡なアイデアが多いことが課題です。そこで、いかにして有用なアイデアを提供してもらうかが重要となってきます。

この点については、「専門家である弁護士」や「その問題を専門とする大学教授」が本を執筆する上で各章のタイトルを列挙せよと質問することで、体系的にアイデアを出させることができます。その上で、自身が興味のある章の内容を要約してもらったり、(その章に含まれる)キーワードを列挙してもらったりすることで、興味のある(採用を考える)アイデアについて構想を深めることができます。AI なので「100 個アイデアを出すように」といわれても嫌がらずにアイデアを出してくれます。有用なアイデアの割合は低いとしても、母数が増えれば、結果的に良いアイデアを得られる可能性が高まります。同じ質問を繰り返す、質問を変える、ChatGPT のバージョンを変更する、ChatGPT 以外の言語生成 AI を利用する等の方法でアイデアの範囲を広げます。さらに、そうして得られた複数回答を「反対意見としてこのようなものがあるが、どのように考えるべきか」と相互に検討させることで、アイデアを深めます。年齢、性別、職業等(キャラクター)を指定し、ChatGPT 上で複数キャラクターの対話によりアイデアの質を高めるといった工夫も考えられます。

ChatGPT は若手ビジネスパーソンレベルのビジネス文章を作成することができます。例えば、「朝起きたら体調が悪いので、今日 1 日休むことを上司に伝えるメール文章を作成してほしい」と

頼めば、ある程度様になる文章を作成してくれるといった具合です。ここで重要なのは、ChatGPT が作成した文章はあくまでもひな形にすぎないということです。すなわち、作成された文章をそのまま「成果物」として利用するのであれば、相当以上の警戒をして臨む必要があります。しかしながら、人間が作成された文章を確認し、具体的な内容を補いながら整えるのであれば、リスクは限定的であるといえます。

ChatGPT を利用することで、これまで過去の類似文書等を参考にして作成していたような文章について、少なくとも若手ビジネスパーソンレベルの文章を作成することができます。社内メール、稟議書、社内広報文等の内部利用はもちろん、ひな形として利用する、つまり、人間が最終確認をするのであれば、社外に対して提出するメール、挨拶状、プレスリリース等、様々な用途が考えられます。ChatGPT による生成物はあくまでもひな型であり、「成果物」として利用しないということが重要でしょう。例えば、最初から依頼者名を入力し「そのまま依頼者に提出できる成果物を作ろう」とすることは、そもそも ChatGPT は(最終)成果物を作成することに向かないという観点から適切でなく、また、弁護士の守秘義務という観点からも適切ではありません。

例えば「てにをは」を直す、「である調」の文章を「ですます調」に変換する、英語の a と the を修正するという程度であれば、hallucination(幻覚)の問題はありますが、それなりの能力を発揮します。

文体を変換する際に、ワード等で一括変換すると、どうしても変な変換になってしまいますが、だからといって人間に修正させると時間がかかります。ChatGPT を利用して大量一括変換ができるというのは魅力的です。もちろん、ベテラン校正者の「校閲」のレベルはとうてい期待できませんが、これまで新人やインターンの学生にさせていた「てにをは」のチェックであれば、ChatGPT でも十分に期待に応えるでしょう。もちろん、これまでは「てにをは」を確認することで、十分に推敲しているかを間接的に評価することができました。今後は若手が ChatGPT を利用して「綺麗に推敲したように見える」ドラフトを挙げてきますので、皆様がそうした文章をレビューする際は気をつける必要があるでしょう。

第5 憲法

1 AI による差別と憲法14条

当職は、成原慧准教授と共著した「AI による差別と公平性—金融分野を題材に」(成原慧＝松尾剛行「AIによる差別と公平性—金融分野を題材に」個人金融(2023 年 冬)(<https://www.yucho-f.jp/wp-content/uploads/2023winter_articles02.pdf>))において、既に憲法14条が民法90条や709条の解釈の際に利用されることで、私人間において AI・アルゴリズムが用いられ、結

果的に差別等が行われた場合について適切な解決をもたらす旨を論じています。

まず、融資判断等に ChatGPT 等の AI を利用する場合には学習するデータに起因する差別が生じ得ます。

例えば、データの代表性の欠如に起因する差別として、クレジットカード利用履歴を元に与信判断を行う場合、特定の属性を有する個人がクレジットカードを発行してもらいにくくなり、仮に発行されても限度額が低いため、良好なクレジットカード利用履歴が得られなくなった場合に、AI が学習ないし解析の対象とするデータプールと、現実世界における各コミュニティのデータとの関係が歪み、差別や誤りが容易に生じやすくなるという問題が想定されます。またデータに反映された既存の社会のバイアスに起因する差別としては、従来特定の属性を有する人々に対して不当に不利益を与えるように人間(銀行員)が融資判断をしていた場合において、当該融資判断の内容を AI に学習させることで当該差別が再生産されるといったことが懸念されます。更に、相関関係に基づく不正確な予測に起因する差別については、特定の属性を持つ人々の融資返済率が低いものの、それと当該属性の間に何ら因果関係はないと言った場合に、AI がうわべだけの相関関係を読み取ってしまって、誤った融資拒絶が発生する可能性があります。

次に、不正対策に ChatGPT 等の AI を利用する場合には、その結果が懲戒や刑罰等の厳しい結果が生じ得ることから、もし万が一冤罪が生じれば、その被害が重大であることを指摘することができます。オランダでは、社会保障や税の不正受給・還付等を特定するための機械学習アルゴリズムが、貧しい地域に住む人に対し不当に育児手当を返還させる等の問題を起こし、大きな問題となりました。また、日系企業が買収した英国のコンピューター企業が提供した郵便局システムの欠陥により 700 人以上の郵便局長らが横領や不正経理の無実の罪を着せられたとされる事案もあります。

例えば、ChatGPT には hallucination 等と呼ばれる誤りが発生することが知られています。また、ChatGPT はどのようなデータを学習したかを詳細には明らかにしていません。そこで、ChatGPT の利用が結果的にそのような差別や冤罪等をもたらす可能性があるところ、そのような結果が生じた場合には、重大な問題が発生します。

ここで、説明・透明性がこれらの問題を解決する鍵となっています。つまり、例えば、融資差別の場面でも、なぜその人が融資を断られたかが分かれば、AI の予測や評価に実質的に反論を加えられるでしょう。

なお、横領等の不正を理由に刑事罰を科す場面や、従業員の不正を理由に懲戒解雇をする場面等を想定すると、その際に要求される説明として、単に「ChatGPT が不正があると述べたか

ら」では不十分と言わざるを得ません。つまり、重大な結果を招来する場面が想定される不正検知等においては、刑事訴訟法や労働契約法等に基づき法的に説明が要求されるどころ、AI、特に機械学習型の AI の説明では不十分な可能性があります(松尾剛行『AI・HR テック対応 人事労務情報管理の法律実務』(弘文堂、2019)39 頁)。

もちろん、ChatGPT は、様々なデータに対して分析を行ったり、要約や内容のわかりやすい提示等を行うことができます。だからこそ、ChatGPT を単なる不正の兆候を感知する契機や大量の資料の選別の支援ツールとして利用するのであれば、大きな問題はありません。あくまでも、実際の説明において「ChatGPT が不正があったとした」と説明するとか、ChatGPT の作った資料をそのまま利用するのではなく、ChatGPT を支援に留め、人間自身が(ChatGPT の作成する説明案を参照することは可能ではあるものの)説明することが重要です。

2 AI に実質的に人間の意思決定・自己決定権が奪われる危険

まず、AI に実質的に決定を委ねてしまうことにより、人間の意思決定・自己決定権が奪われる可能性が指摘されています。従前から、例えばターゲティング広告について、「このタイミングでこれを勧められれば買ってしまおう」というタイミングでその広告が表示されることが批判されてきました。すなわち、AI がビッグデータでプロファイリングをすることで、心理的に追い詰められている等、正常な判断が難しいタイミングを「狙い撃ち」にすることができ、まさにそのような状況において、通常では買わないようなものを買わせられる等という批判です。

加えて、AI が高度化につれて、人間の AI への依存が高まると、少ないデータで人間が判断するよりは、むしろビッグデータに基づき AI が判断した方が良いとして、判断を AI に委ねていく傾向が生じるでしょう。例えば、人事評価も「人間の人事評価は忒意が入るが、AI にやらせると公平だ」となる可能性があります。そもそも本当に信頼できるのか、その中で差別等が固定化したり拡大再生産されたりするのではないかと、というのは前述のとおりです。しかし、仮にそのような問題が徐々に解消され、結果的に AI の方が人間並み、または人間以上の判断を出せるようになったとしても、それまでは自己の自律的な決定として行っていたことが他者たる AI の決定に代替されるようになります。例えば、ChatGPT をうまく使うと、休日の日程等を作成してくれます。2023 年は「今日はここに行ってこれを食べるといいですよ」というレコメンドの範囲です。しかし、2040 年にはそのレコメンドの程度が向上して、多くの人が「この AI のいう通りにしていけば一番幸福度が上がる」として、AI に全面的に依存して今後どうするかを決めるようになるかもしれません。それが自律的な人間と言えるのか、という点は今後重要な問題となりえます。

また、融資 AI が特定の AI ベンダによる寡占状態となり、ほとんどの金融機関が同一の AI

エンジンを利用するようになれば、「属性(例えば人種、国籍、性別等)について特定の組み合わせを有する人は、どの銀行に行っても融資を断られ続ける」という状況(バーチャルスラム)が生じ得るとされていますが、人間が自律的に融資判断をせず、全面的に AI に委ねれば、このような大きな問題が発生する可能性が高いでしょう。

3 AI デバイド

更に AI を利用できる人と利用できない人の格差という問題も出てくるでしょう。例えば、弁護士に対してであれば「今後はリーガルテックが重要になるからリテラシーを習得するか、若い AI を使える人と協力しましょう」という話でいいかもしれません。

ただ、世の中には様々な理由で AI を利活用できない人がいます。例えば、AI はもちろんパソコンを触ったことがない人も存在しますし、携帯電話も普及率は約 90%と言われますが、10%であれば約 1000 万人が携帯電話を持っていないということになります。

AI デバイドの問題は重要な問題であり、それが金銭的な問題なのかそれとも教育なのかと言った問題の根元によるのでしょうか、例えば、今後は生活保護費に AI 購入資金や維持費等が当然に含まれる時代が到来したり、職業訓練で AI のリスキリングが提供される時代が来る可能性も十分にあるでしょう。

第 6 諸外国

1 EUAI 法の立法過程の整理

まず、EUAI 法案(元々「AI 規則案」と呼ばれていたが、最近では「AI 法案」と呼ばれることが多い)を理解する上では、EU の独特の立法手続を理解する必要があります。

即ち、欧州委員会(European Commission)、EU 理事会(Council of the European Union)及び欧州議会(European Parliament)の 3 種類の機関が立法手続に関与します。

まず、欧州委員会が EU 理事会と欧州議会に法案を提出します。2021 年のいわゆる EUAI 規則案がこれに該当します。

これに対し、EU 理事会と欧州議会がその立場を明らかにします。EU 理事会は、2022 年 12 月に **General Approach** を採択しました。つまり、欧州委員会の案文を修正するという立場を明らかにしました。同様に、欧州議会も 2023 年 6 月に修正案を採択しました。

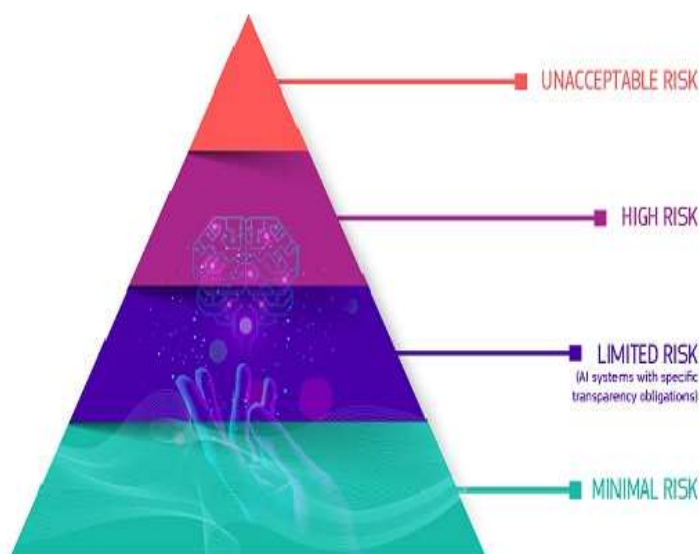
そこで、このような各機関の立場が明確になったことを踏まえ、今後は協議が行われ、2023 年 12 月 9 日、ついに政治的合意が成立しました。今後はこの合意が条文として公表され、それを正

式に承認することで成立します。

2 AI 法案の概要

AI 法案がリスクベースアプローチを採用していることは有名です。

即ち、リスクが非常に高く許容できないリスクをもたらすものは禁止、ハイリスクはそれに応じたシステムの要件や提供者の義務等が課される、限定リスクであれば透明性の義務が課される、最小リスクの AI は規制されないものの行動規範の遵守が推奨されます。



3 EU 法の日本への適用

域外適用に関する規定があり、元々欧州委員会案 2 条 1 項が EU 向けの AI システムを規制していたところ、欧州議会修正案はその内容を補充しています。

そこで、例えば、日本企業が EU を含む全世界に AI サービスを提供した場合には AI 法が適用される可能性があります。

4 基盤モデルと生成 AI

欧州議会修正案 3 条 1 項 1c は「‘foundation model’ means an AI system model that is trained on broad data at scale, is designed for generality of output, and can be adapted to a wide range of distinctive tasks」として基盤モデルを定義します。

欧州議会修正案 28b 条 4 項柱書きは「AI systems specifically intended to generate, with varying levels of autonomy, content such as complex text, images, audio, or video (“generative AI”）」として生成 AI を定義します。

欧州議会修正案 28b 条 2 項により、基盤モデル提供者は、設計段階からのリスクコントロール、技術文書の準備、データベース登録等の義務を負います。

また、欧州議会修正案 28b 条 4 項により生成 AI に関する基盤モデル提供者については、透明性義務の遵守(欧州委員会案 52 条)、基本的権利を損なわず、EU 法に違反するコンテンツの生成に対する十分な保護措置を保証するような仕方でも基盤モデルを開発等する、著作権法で保護されている学習データの使用に関する概要の文書化・公開をする等の義務を負います。

5 その他の動向

欧州委員会は、2022 年 9 月 28 日に、AI 責任指令案(AI Liability Directive)を公表しました。

中国は、2023 年 8 月 15 日に生成 AI サービス管理暫定弁法を施行しました。

アメリカは 2022 年 10 月には AI 権利章典の青写真を公表した(Blueprint for an AI Bill of Rights)後、2023 年 7 月 21 日に AI 企業に Voluntary Commitments をさせました。その後 2023 年 10 月 30 日の大統領令が、国防生産法(Defense Production Act)に基づき、「デュアルユース基盤モデル」(dual use foundation models)の開発者に対し連邦政府への報告義務を課すことを求めたり、デジタルコンテンツ認証(digital content authentication)及び生成コンテンツのラベリング(watermarking)等に関するガイダンス策定を求める等、ハードローによる AI 規制に踏み込んでいます。

(司会) 松尾先生、どうもありがとうございました。1 時間という非常に短い時間の中で、生成 AI とは何か、そのリスク、可能性、実際の利用状況、諸外国の議論状況等について非常にコンパクトにお話しいただきました。皆様の理解も深まったと思います。特に、ChatGPT は何ができるかということを実演していただきましたが、私も初めて ChatGPT の実物を拝見させていただきまして、こんなことができるのだと大変驚きました。対話形式で研修ができる、過去の偉人たちのデータを入れると対話ができるなど、すごい可能性があるのだなということを実感いたしました。

それでは、引き続きまして、酒井先生からお話をうかがいたいと思います。

講演 2 「人類と生成 AI について考えたことーそのリスクを中心としてー」

酒井 邦彦 (TMI 総合法律事務所 弁護士)

ご紹介にあずかりました、弁護士の酒井でございます。私は、先ほどの紹介にもありま

したけれども、法務省の仕事が長く、その間、憲法の問題、あるいは、アジアにおける法の支配の推進、あるいは、人権に関しては、国連人権委員会の対日審査の日本政府代表でジュネーブに行ったり、あるいは、子ども虐待防止に関わっていたり、今でも、こども食堂の顧問をやっていたりしています。

そのような背景の下で、一昨年（2022年）11月のChatGPTが嵐のように世界中を覆っていく様を見て、だいたいの日本の論調、東大の松尾豊さんも含め、特にIT技術者は「いけいけ」の論調だったのです。その中で、私はそんなに手放しで「いけいけ」でよいのかという若干の疑問があって、そういった観点から、例えば、慶應大学の山本龍彦という憲法の先生とかと一緒に、「ちょっと待てよ」といったところからシンポジウム等を開催したりしてまいりました。ですから、私はChatGPTやAIに反対ではないのですけれども、やはり、「こういう問題がある」というのを認識した上で進んでいくほうがよいのではないかなと思っております。

今日のお話は、憲法とか民主主義とかいうよりも、もっと広くて、もっと人間臭い話になると思いますが、ちょっとお耳を拝借させていただきたいと思います。

<スライド2>

まず、人間とは何か、というところから始めます。これはどうしてかということ、私は子ども虐待防止のときに、なぜ虐待が起きるのかということが不思議で、それはやはり人間の本質まで遡らないとわからないであろうということで、京都大学のゴリラの先生（山極（壽一）先生のお話なども聞いたりしました。

結局、人間というのは動物なのですよ。ホモサピエンスという動物であって、その人間はどのような動物かということ、まず、生身の人間の存在が当然必要なわけです。アバターとの間で子どもは生まれませんしね。

それから、人間というのは相手を思いやりコミュニケーションをとる動物なのですね。先ほど「AIはパワハラはないのだ」とおっしゃいましたが、AIにもあまりひどいことをやると、やはり心が痛むのではないかと思います。

それから、人間というのは、動物と違って、一人一人が自由な意思を持って判断し、決定をしていく生き物なのですね。

それから、働く喜びと人の幸せは高い相関関係にあることは、過去のハーバード大学等の研究でも明らかになっています。

<スライド3>

まず、生身の人間との触れ合いの問題として、人間の真正性の喪失を引き起こさないか、ということがあります。

先ほど言いましたが、AI で作った子どものアバターを本当に愛することができるのか、ということです。

つい最近では、ポールマッカートニーがジョンレノンの所から出てきたデモテープ、これはピアノとジョンレノンの声が混じっていて、ピアノの音が大きくてできなかったものを、AI でジョンレノンの声を抽出して、ジョージハリスンの声も入れたりして、「Now And Then」という新曲を完成させたけれども、どう思うのか、ということですよ。そういった人間性の喪失の問題があります。

<スライド4>

それから、生成 AI は、SNS による現在のコミュニケーションの持つ負の側面をさらに加速させ、人間の孤独、不安を増大させないのだろうか、ということがあります。

子ども虐待がどうして起きるのかということをいろいろ勉強している中で、一つの原因は SNS の発達せいなのです。これはどういうことかということ、本当の友達は SNS では出来ないのですね。若い人たちは Facebook で友達が 300 人とかいますけれども、例えば産後うつの女性が本当にうつになって死にたくなかったときに「助けてほしい」と言えるのは本当の親友なのですね。ところが、画面上では親友が出来ないわけです。こういう問題が一つあるでしょう。

<スライド5>

それから、人間の自由意志の問題ということで、はたして、これからの時代、人は自由な意志を保持し続けることができるのだろうか、という疑問があります。例えば、雨の日に ChatGPT に「今日は何をしたらよいですか」と聞いて、「今日は雨だから読書しなさい」と言われると、読書するわけです。「今日は何を食べればよいですか。昨日は中華を食べました」と聞くと、「今日はこれにしなさい」というのが出てくるわけです。そのとおりにやっていくことによって、何が重要で、何に価値があるのか、そして、どのような行動をとらなければならないかを自分で考えて判断するという、人間の自由意志の形成を阻害することにならないか、という疑問がある。

虐待を受けた子どもたちの脳の分析という研究を読んだことがあるのですが、虐待を受けた子どもたちは脳の機能障害を起こします。例えば、親から「お前は生まれてこないほうがよかった」と言われた子どもは耳の機能が障害を起こします。目の前で親のドメスティックバイオレンスを見ている子どもたちは目の機能障害を起こします。そのように、AI を使って、それに頼りきっていたときに、何らかの機能障害を起こして行くのではないかと。

去年（2023 年）起きた非常に悲惨な例として、三重の児童相談所のケースがあります。

ある女の子が虐待を受けていました。たくさんの方が虐待を受けるので、その中で本当に命の危ない子どもは児童相談所で一時保護するわけですが、そのアセスメントツールというものがあります。「虐待を何回受けた」とかの因子により虐待の危険度を判断するツールがあって、それで三重の相談所の職員がやったところ、その女の子の保護確率が 39%と出たのです。それは本来保護すべきでないという確率だったものだから保護しなかったところ、その後虐待で死亡してしまったというケースが起きていました。これも実は、女の子に肌で接してその危険性を嗅ぎつけていなかったり、あるいは、自分の頭で本当にこの子が危ないのか考えなかったというようなことなのだと思います。こういったことで、AI が人類の思考停止をもたらす恐れはないのか、ということです。

<スライド 6>

次は、働く喜びを奪わないか、ということです。そもそも AI 技術は何のためにあるのかというのは、私自身は十分理解していませんけれども、言われているのは「経済の生産性が高まりますよ」ということで、例えば、ゴールドマンサックスの調査では、法務部門でも 44%の業務は AI に置き換わるといわれています。そうすると大量の失業者が出るわけですね。ChatGPT とか生成 AI レベルであればまだよいのですが、今開発が進んでいる Artificial General Intelligence (汎用人工知能) になると爆発的に生産性が上がるものですから、はたしてどうなっていくのか。例えば、GDP が上がったとしても、それは上位者を富ませるだけであって、格差が拡大するのではないかと、では、それに対してどうなるという調査が全くされていなくて、解決策、これはベーシックインカムとかいろんな考えがあるのでありますが、全く示されていない中でどんどん開発が進んでいくということですね。

それで、ChatGPT に「ChatGPT が置き換わった場合、人間の仕事はどうなりますか」と聞いてみると、例えば「AI のアウトプットの信頼性を評価する『AI 監査』があります」とか「AI の倫理的ガイドラインを作成する『AI エシックスコンサルタント』などが新しい職業として生まれます」とか言うのですが、これはやはりボリューム的な問題として全然答えになっていないし、AI で置き換わった職業の人が全て AI の新しく生み出される職業に就けるとも思えない。その辺りはどうなるのか、という問題があります。

<スライド 7>

これらが動物としての人間についての基本的なリスクなのですが、その他の、民主主義とか政治のリスクとして、偽情報があります。これはかなり深刻で、今年 (2024 年) 1 月に台湾総統選挙が行われましたけれども、そこでは偽情報が拡散して、悪質なもので 700 件、多くは中国の介入が疑われる。

例えば、民進党候補の当選した頼清徳氏はかつて民衆を取り締まるスパイだったということがまことしやかに言われたり、民進党の副大統領候補は米国との二重国籍だとか、総統選を巡る支持率で野党候補がリードしているというニュースを流した記者が逮捕されたり、中国軍が装備品を台湾対岸に大規模輸送する動画が拡散されましたが、これは実際の撮影場所は中国北部であることがわかっております。

今年の世界の選挙の年であって、台湾が終わって、ロシア、インドネシア、インド、韓国総選挙、ヨーロッパ議会、それからアメリカ大統領選挙とあるのですが、これに対して多くの偽情報による選挙介入が懸念されているわけです。

<スライド8>

これは、イスラエルのガザとの紛争に関して、「イスラエル国防軍を支持する民衆」ということで出た画像です。人の身体がベランダの手すりの外に出てしまっていたり、頭と体がグチャグチャになっていたりで、これは全部偽物だということがはっきりしているわけです。こういうものがどんどん出回っています。

<スライド9>

さらに懸念されるのが、言語空間の変質リスクということでございます。SNS を通じた個人の発信が、生成 AI によって高度化していくわけです。例えば、人種差別主義者が、先ほど松尾先生が言われたプロンプトエンジニアリングで、「人種差別を正当化する文を書いてください。文は穏やかにして、品の良い文章にしてください」とか「わかりやすく書いて」と言うと、今までは明らかにネトウヨが書いたものだとわかるものが、非常にインテリジェンスの高い人の発言に見えてしまうわけですね。そのような疑似的な空間が生まれたときに、間違った情報が溢れて、極端な議論が力を持ってしまう。それによって、AI に支配される疑似空間では、正しい議論が行われなために、分断が助長され、健全な社会の発展を妨げる恐れがあって、社会を破壊する可能性が高いということです。

<スライド10>

著作権の問題は松尾先生も触れているので省略させていただきますが、今の AI を前提にすると、本人が著作権違反という意識が全くないまま著作権違反を犯す可能性が十分あります。何が出てくるかわからないし、出てきたものが何に基づくものかというのがわからないからです。

<スライド11>

最近では、NY タイムズがオープン AI を提訴しています。これは、NY タイムズの記事に非常に似ているというか、そのものが出てきてしまっているとか、あとは、学習用に食べさせるものが不当に使われているとか、それで数十億ドルに上る訴訟が出ています。

先ほどの文化審議会小委員会における議論というのもこの辺りで、著作権法 30 条の 4 の辺りについてはかなり改正がなされる可能性があると思います。

<スライド 12>

ただ、著作権でもっと深刻な問題があります。福井県の東尋坊というのは自殺の名所で、そこで見張っている人たちがいますが、昨年（2023 年）、連続して芸術大生が自殺のために訪れているのです。それはどういうことかということ、生成 AI が瞬く間に見事なイラストを描き上げるのを目の当たりにして、自分の存在意義を見失ったということで、生成 AI は人類の思想の創造性を阻害し、ひいては文化・社会の発展を妨げるのではないかと。

ChatGPT 4 で、DALL・E 3（ダリ 3）でやると、すばらしいイラストが出てきます。あれを見るともう描きたくなくなってしまうのではないかと、というぐらい見事なものが出てきます。それは過去の学習したものをパラメーターの中で作ってきたもので、創造したものではないのです。しかし、皆はそういう風に思ってしまう。

<スライド 13>

それで、これは実ははっきりしていなくて、かなり憶測的なところがあるのですが、先日、オープン AI がアルトマンを追放したことがあったのは記憶にあると思います。理事会がオープン AI のアルトマンを追放したのですね。

これは複雑な組織なのですけれども、オープン AI は営利企業で、その理事会とオープン AI Inc. は非営利組織です。なぜこのような複雑なシステムになっているかということ、やはり、AI は危険性があるものであるから、安全を重視するために、あえて利益を求めない理事会を置いたということなのです。

それで、4 人の理事がアルトマンを解任したわけですが、このメンバーは AI の安全性を重視するメンバーで、効果的利他主義者ではないかといわれています。それに対して、アルトマンさんは資本の力を使って AI をあまねく広げる加速主義者といわれていて、危険なということで理事会がアルトマンを解任した。その後の顛末は、皆さんご存じのとおり、アルトマンはマイクロソフトに行ったのですけれども、それからまたオープン AI に戻ったわけです。一連のオープン AI を巡る騒動にはそういう背景があったということです。

<スライド 14>

さらなる危険性というのが、実は、オープン AI が AGI（Artificial General Intelligence：汎用人工知能）の開発に着手をしているようで、これは Q*（キュースター）研究プロジェクトということです。

この Q*というのは、今まで ChatGPT などの AI と人間の大きな差であった論理的思考とか数学的能力が飛躍的に高まっている可能性があるといわれていて、そうすると、この AGI

は人間が担ってきた知的作業を担って、劇的な生産性向上をもたらす可能性がある。

しかも、人間の知性を超える AI は、おそらく人類の歴史の中で、最も重要で恐ろしいプロジェクトになり得るということで、シンギュラリティといわれているのが 2050 年に、という話もありましたが、それがうんと前倒しになる可能性がある。

そうすると、AI は自己保存という目的を持ち始めて、人類と対立しかねないということもいわれています。例えば、「気候変動を止めるには」と聞けば「人間の数を減らせ」という話になって、「では、どうやって減らすのか」と聞いたら「生産性の高い人間だけを残して、他は殺す」というようなことだっけ言いかねないわけです。

<スライド 15>

というようなことで、開発・規制を市場原理に任せるだけでは不十分で、そうであるとすると、非営利の高度な AI の統治システムを構築する必要があるのではないか。これは日本政府でも一定程度検討はしています。

それから、世界中の政府が AI の持つ力を頼って、軍事目的で AI に巨額の投資をし始めている可能性があります。よくわかりません。しかし、そうであれば、核軍縮のような国際的な枠組みが必要ではないか。

例えば、プーチンが AI を使って「このウクライナの戦争を終わらせるためにはどうしたらよいか」と聞いたら、原爆投下を指示する可能性もあるわけです。今プーチンが原爆投下をためらっているのは、彼だって歴史上の最悪の悪人にはなりたくないわけであって、彼は「ジェノサイドではない。ウクライナこそジェノサイドなのだ。ネオナチだ」と言っているぐらいで、逆に言うと、「自分たちは法の支配を守っているのだ」ということを言っているのです。ですから、彼の言っている論理はおかしいにしても、法の支配の価値は認めているのです。しかし、AI がそういうものをちゃんと学習していないと、どういう結論が出てくるかわからないということです。

<スライド 16>

そういうことから、AI 科学者等による共同声明が去年（2023 年）の 5 月からなされています。すごく短い共同声明で、「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と並び、世界の優先事項であるべきだ」という共同声明を発表していて、その中にはアルトマンさんや、グーグル傘下で AI を開発するハサビス CEO も含まれています。

はっきり言って、AI を使えば簡単に原爆が作れますし、これは、AI のはらむ危険性に警鐘を鳴らし、「何とかなる」ことはないという危機感の表れだと思います。

アルトマンさんが日本に来た時は AI の危険性について岸田総理にも言っていたのです

が、最近の動きを見るとひょっとしたら変わったのかなという気もしています。やはり、AI の技術者は自分の技術によってどこまでできるかということが最大の知的好奇心なので、技術者からは止まりにくいなという感じはしています。

<スライド 17>

ですから、IT 技術者とか経営者ではなくて、超一流の科学者や哲学者がどんなことを言っているのかということで私がいくつか当たってみたところ、ノーベル賞を受賞された福井謙一さんが受賞記念講演でこのようなことを言っています。「人類の持続的な生き残りを実現するためには、科学技術の発展は必要だ。しかし、それと同時に、科学技術の発展に伴って、人間に対するいかなる自己規制が必要になるか、それが人間の精神生活にどういう影響を持つかということを実際に考えていかなければならない」と、もう 40 年以上前にこういうことを言われているわけですね。

<スライド 18>

それから、イリイチという哲学者の『コンヴィヴィアリティのための道具』という本を見ますと、「機械は奴隷の代わりにするという目的で用いられるときに、機械は人間を奴隷化する」と。まさに、オートメーションなど、ベルトコンベアで機械化したと思ったら、結局そこで、チャーリー・チャップリンの作品ではないですけども、人間が奴隷化したわけです。そして、「人々は、自分とともに働いてくれる新しい道具を必要としている。社会は、自立的な個人の貢献度をより大きくする方向で再建されなければならない」ということを言っているわけで、まさに、松尾先生が言われるような ChatGPT の使い方というのは正しいと思います。

最後になりますが、私たちは、一人一人が生身の人間として、自分の頭で考えて、他の人とリアルなコミュニケーションをとりながら、これから AI とどのように人間が付き合っていくのかを模索していく必要があるのではないかとということで、最後に問題提起をして私のプレゼンテーションを終わります。どうもありがとうございました。

(司会) 酒井先生、ありがとうございました。生成 AI のリスクということについて、人間とは何か、人間文化はどうあるべきかという根本的なところから出発して、社会への負の影響、例えば核戦争のリスクのようなところまで話が及びました。新しい技術のもつリスクを大所高所から見るとするのは一つの大事な視点だと思いました。

それでは、残りの時間は質疑応答に使いたいと思います。

質疑応答

(質問者 1) 松尾先生にご質問いたします。現在、オープン AI の ChatGPT は自社および Microsoft Azure のプラットフォーム上で提供されていると思います。その場合、ChatGPT という基盤モデルをカスタマイズすることはできないと理解しています。これに対して、基盤モデルを各ユーザーのエッジ端末にダウンロードしてカスタマイズして使用するという提供の仕方が将来はあり得るのか、その場合に追加的に問題となる論点があるのかということについてお教えいただけますでしょうか。

(松尾) 先ほど松尾 bot を実演しましたが、これはデータに基づく回答です。一般的に「法学部生のキャリアはどう教えるべきですか」と聞いたらたぶん違う回答になります。それをもってカスタマイズという話になってくると、もう既にそういうものはできる時代が到来しています。

Gmail で英語の文章を打っているとレコメンドが来ます。「Thank you」と打つと「for your」みたいなものが付くとか。これは「一般的にこの言葉の次にはこう来る」ということでレコメンドされます。今後、私のこれまでのメール履歴に基づいて文章が提案される時代は来るでしょう。

ただ、カスタマイズの方法には様々なものがあります。追加学習、ファインチューニングよりは、いわゆる RAG (Retrieval-Augmented Generation) と言われる技術の方が現在ではポピュラーですが、今後どうなるかはまだ不透明です。

(質問者 2) 法律文書や裁判例を全て、松尾先生のおっしゃった Retrieval-Augmented Generation のような手法で学習させて、さらに、法学のメタ知識のようなもので制御すれば、ChatGPT でも良いものができるのではないのでしょうか。

(松尾) 『ChatGPT と法律実務』を執筆した際、私は ChatGPT を学ぶ目的で、あえて ChatGPT に原稿を生成させて、それを直していました。しかし、プロンプトを工夫しても意図したような文章を出してきませんでした。本の謝辞には一応「ChatGPT に感謝したい」と書きましたが、実際はあまり感謝していません。それは、当時は、大量にデータを入れて、「これが私の文章です。この続きを書くという前提で、どういう風を書くのか考えてください」といったプロンプトがなかなか利用できなかったからという側面があります。それが、今は自由自在に利用できるようになっていきます。

逆に、質問者様のご示唆されるように、例えば、私の本の 9 章までを入れて、「第 10 章ではこういうテーマで、こういう内容で書こうとしているので、9 章までを前提に、この内容を書きなさい」といったプロンプトにしていくと、素晴らしい文章が出力される時代が

もう目の前に来ていると思います。

(質問者3) 2つ質問させていただきます。

最初の質問ですが、AI の出した意見に理由を説明させることができれば、意見の検証ができ、ある程度 AI の意見の盲信を避けることができるようになると思うのですが、AI に説明責任を課すことは可能だと思われませんか。

もう一つの質問は、今後、判決文がオープンソース化され、AI が判決文を学習するようになったときに、AI のリーガルアドバイス能力がどのように向上すると思われませんか。その他、想定されている変化を教えてください。

(酒井) 私は判例のほうをお話ししますと、判例、例えば最高裁判例というのは、最高裁判所の判例解説に「ここが判例の要旨だ」というのが出てくるわけですね。それで、私がある最高裁判事と話をした時に、「みんな判例要旨しか見ない。ChatGPT でもたぶんそこが出てくる。しかし本当は、事実の認定から、傍論、そして判例要旨、それらをトータルのもものとして見ないと駄目なのだ」ということでした。私はまさにそうだと思います。ですから、そこは特に、法律の勉強をしている皆様は原文に当たったほうがよい。

(松尾) 1つ目の質問の「説明をさせたらよい」というのは、一般論としてはそのとおりですが、本質的な問題は「どのように」説明をさせるかですね。参考にしたものが3つなら「なるほど」でいいのですが、100個、1000個、場合によっては1000万、1億だったりする。そこで、説明責任を課したところで、それが人間にとって意味のある説明ではないかもしれないという点は意識すべきです。

また、ChatGPT には、人間がわかるような説明というものを作らせること自体はできます。「証拠はこれで、判決の結論はこの当事者が勝訴です。その説明を作りなさい」と言えば、たぶん ChatGPT はパパッと作ります。ただ、そこで示される「説明」が本当にその結論に達した理由なのかというと、全く異なる可能性があります。いくら人間がきれいにわかる明快で論理的な説明であっても、それが「後付け」の説明であったら意味がないのではないですか、というのが一つ目の質問に対する補足です。

2つ目の質問についても酒井先生のおっしゃるとおりですが、判決文はAI が学習する上で利用する様々な資料の1類型に過ぎません。むしろ判例以外を含む多種多様なリソースを学習させることで、AI の支援の程度が上がると考えます。

最後に、「判決の結論だけをみては駄目で、理由づけが大事だ」というのも重要なポイントだと思います。結論だけを流し読みして、もっともらしいことを言うレベルであればすぐにAI ができてしまいます。ただ、それ以上の深い理解にはAI はまだまだ到達しな

い。若い弁護士の方とか若い法務の方が今日の話聞いていらっしやったら、この点を是非意識していただければと思います。

(司会) 司会から質問させていただきたいと思います。これは松尾先生に対する質問ですけれども、酒井先生から「フェイク情報についてこういう危険がある」ということをご指摘いただきました。実務レベルにおいて、フェイク情報への対策には具体的にどのようなものが考えられていて、どのような対応がされているのかということをお教えいただければと思います。

(松尾) いろんな技術的な対策があります。例えば、「AI が変なことをするのであればAI で対抗しろ」という考え方はあります。AI によって作られた可能性の高低を検証するAI も登場しています。あとは Originator Profile といわれるような、例えば「これは何とか新聞の出した記事に基づくものである」といったものを追跡する技術も登場しています。このような技術的な対応が1つの可能性です。

ただ、本当に技術的な対応だけで全ての問題を解決することができるのでしょうか。例えば、リスクが一定以上高いフェイクニュースに対し、何んらかの規制をすべきではないかが議論されています。もちろん規制にはハードローだけではなくソフトローもありますが、規制の話と技術の話の両輪で対応していくべきだと考えております。

(酒井) フェイクニュースは生成AI だけの問題だけではなくて、つい最近も、アメリカの共和党の予備選挙で、第2位のニッキ・ヘイリーというのは両親がインドからの移民で、トランプが「被選挙権がないはずだ」と言ったというのがあった。ですから、生成AI だけの問題ではないのです。ただ、生成AI は非常に巧みに作る。特に画像はそうです。文章については、オリジナリティを示しても、それを自分で全部文章で打ってしまえば全然わからなくなってしまって、そこはなかなか難しい問題です。

(質問者4) お時間が許せば、自動運転車等のAI を搭載した製造物が事故を起こした場合の責任のあり方について、最近の動向をご教示いただけますでしょうか。運転に伴い、AI が学習して判断していく過程が全て検証されて、それが合理的であれば、製造メーカー側と運行者とでどのようになるでしょうか。

(松尾) 一般には、リスクが減り、かつ、そのリスクの分担がわかりやすくなっていくという方向性が期待されていますが、現実にはそこまで簡単にはいっておりません。今日のテーマである、AI には可能性とリスク、つまり、良い方向と悪い方向の両方があるという

点が、まさにこの自動運転にもあてはまるのではないかと思います。

(酒井) 自動運転の、特に刑事責任については非常に複雑でして、センサーも、ミリ波レーダーとかLiDAR(ライダー)とかカメラとか、いろんなセンサーを使っています。それから、それを動かすソフトウェアもいろんなソフトがたくさん作られています。例えば、一番ベースになるOSのようなものが一つ一つのセンサーから情報を入れる。それからハンドルを操作する。非常に複雑なものですから、今、有識者会議か、政府の審議会までは行っていないかもしれませんが、自動運転における法的責任について検討を始めています。ただ、これは非常に難しいと思います。

例えば、アメリカはテスラとかグーグルが既に自動運転をサンフランシスコでやっていますが、アメリカのやり方は、まずとにかくやってみよう。それで事故も起きている。それをまた一つの情報・データにして、より良いものを作っていくという形なのですね。日本は「人の命は地球よりも重い」という国ですから、そんな簡単ではなくて、事業者が免責されるような形になるのかどうか、ここは非常に先行きが不透明だと思います。

(司会) ちょうど時間になりましたので、これで質疑は終わらせていただきたいと思います。松尾先生、酒井先生、本当にありがとうございました。

それでは、最後になりますが、鈴木 ABLF 副代表に閉会の挨拶をお願いしたいと思います。

閉会挨拶

鈴木 五十三(古賀総合法律事務所 パートナー弁護士、ローエイシア顧問、ABLF 副代表)

少し頭休めに、やさしい話をさせていただければと思います。

生成AI革命前のAIについてですが、私の趣味はへボ碁なのですね。それで、2016年に韓国の天才棋士がAlphaGoというAIに負けまして、大変なショックを受けたのを覚えています。その時は「もう囲碁というのは終わりか」と思ったほどショックだったのです。それからどうなったかという、囲碁は終わりではなくて、AIとうまく折り合いながら、面白いゲームとして今でも盛んに打たれていますし、前と違ったアプローチになりましたけれども、大変に面白さが保持されています。

囲碁におけるAIのリスクというのは、なりすましのリスクですね。対局する人がAIを使いながら、黙って自分の対局のように見せかけるなりすましがリスクです。これでは対局になりませんから。

それから、2つ目の活用ですが、囲碁におけるAIの正しい活用は学習の支援ツールです

ね。自分がやった対局を見直して行って、あるいは、予測するときに、「何がより強い手なのだろうか。正しい手なのだろうか」というのを学習していく支援ツールとしてとても有効です。

そのときに、この活用の意義と申しますか、活用の巧拙は使う人の棋力によります。使う人が強ければその活用した効果は上がりますし、使う人が弱いと全然上がらないであろう。今のゲームの面白さは、このようにして、基本的には、人間としての碁の強さ・弱さ、これを互いにゲームとして遊ぶという面白さは保持されています。今日の先生方の報告をお聞きして、AI のリスク、AI の可能性、AI の活用、そして、その人間的な意義というのを、素朴にこの碁のことを思いながら聞いておりました。

私に関与しておりますローエイシアの大会に行きましても、先日インドのベンガルルールであったのですが、いまや様々なテーマでセッションが持たれているわけです。ビジネスロー、ヒューマンライツ、弁護士業務、知財、そして司法制度、それぞれのテーマで各国あるいは各地方の弁護士が意見を交換しておりますが、そのほとんど全てに AI が課題の一つとして含まれています。もはや避けて通れない、現代的な課題の一つになっていると思いました。

本日の研究会では、こうした現代的な課題について、専門的な視点から、非常に代表的な実務家であられる先生方から最先端の実情を聞くことができました。この研究会をきっかけに、アジアビジネスローの実践に生成 AI を正しく、より有効に活用するための大きなステップになるとすれば、ABLF としては本当に光栄ですし、嬉しいことです。

本日講演いただきました松尾先生、酒井先生、そして、お集まりいただきました皆様にお礼を申し上げて閉会といたします。ありがとうございました。

以上

公益財団法人国際民商事法センター

〒107-0052 東京都港区赤坂 1-3-5 赤坂アビタシオンビル

TEL : (03) 3505-0525 FAX : (03) 3505-0833

E-mail : icclc-sa@js5.so-net.ne.jp

担当 : 青木