2025 年にむけた生成AI への大学の対応 ： 国立大学指針の分析と情報セキュリティの標示方法の提案

福島, 宙輝

(Citation)
大學教育研究,32:95-114

(Issue Date)
2024-03-31

(Resource Type)
departmental bulletin paper

(Version)
Version of Record

(JaLCDOI)
https://doi.org/10.24546/0100488362

(URL)
https://hdl.handle.net/20.500.14094/0100488362

神戸大学　大学教育推進機構　『大学教育研究』　第 32 号　2024 年 3 月：95-114

2025 年にむけた生成 AI への大学の対応
－国立大学指針の分析と情報セキュリティの標示方法の提案－

University Strategies Toward Generative AI for 2025:
An Analysis of Japanese National University Guidelines and Proposal for Information Security

福島　宙輝（神戸大学　大学教育推進機構教養教育院　特命助教）

要旨

本研究では、2023 年に高等教育に本格的に導入されることとなった生成 AI への国立大学の対応文書を分析し、2025 年にむけて大学が組織的に対応すべき内容を提案する。1 節では生成 AI と高等教育に関する国際的な動向を概観したうえで、国立大学が 2023 年 11 月末までに発表した指針文書をデータベース化する。これを踏まえ、2 節では大学が 2024 年から 2025 年にかけて「生成 AI を導入する」際にどのような導入パターンがあるかを提示する。続いて 3 節では、2025 年に向けて神戸大学を含む各大学に求められる生成 AI への対応として、情報資産と機密情報の定義と標示の手法を提言する。最後に、4 節において UNESCO や文科省、私大連など関連する団体のガイドラインと現状の各大学の指針との内容の差分から、学生へのガイダンスや FD において強調すべき内容を提案する。

1.　はじめに

1.1　本研究の目的

生成 AI への対応が高等教育でも求められるなか、各大学は 2023 年に、生成 AI への対応として独自の指針文書を相次いで提示した。文科省は初等中等教育については統一的なガイドラインを提示した（2023 年 7 月）が、高等教育機関については「主体的な対応」を求めるに留めている。本研究は 2023 年に国立大学によってどのような指針が提示されたかを分析した上で、UNESCO や私大連ガイドライン等の国内外の関連資料との比較から、2025 年度に向けて今後どのような対応が求められるかを考察する。なお、筆者は生成 AI の教育導入に肯定的であるが、現状の各大学の指針には不足感と危惧を持っている。

1.2　2023 年：生成 AI のトレンドと高等教育

2018 年に GPT-1 として知られる最初の GPT （Generative Pre-trained Transformer：事前トレーニング済みトランスフォーマー）が公開され、2019 年に GPT-2 が公開された。2021 年には画像の生成モデル DALL-E がリリースされ、Midjourney, Stable Diffusion などの生成 AI システムが相次いで発表された。Google Trends によれば、日本における Google の検索

トレンドで「生成 AI」が登場したのは 2022 年 8 月であり、「ChatGPT」は 2022 年 12 月ごろに登場、2023 年 2 月に検索が急上昇している（図 1)。教育分野において生成 AI が議論されはじめたのも同時期である。図 1 は、2023 年を中心とした教育と生成 AI に関する動向のタイムラインである。2023 年 3 月から 4 月にかけて英国や UNESCO で生成 AI 活用のガイドラインが示され、2023 年 7 月に文科省が「初等中等教育段階における生成 AI の利用に関する暫定的なガイドライン：Ver1.0 機動的な改訂を想定」を発表した。日本の大学が発表した指針については（web 上で確認できる限りでは）2023 年 3 月 22 日、東京外国語大学が「大学教育における AI について　東京外国語大学としての教員向けガイドライン」（東京外国語大学、2023）という文書を発表したものが初出例である。

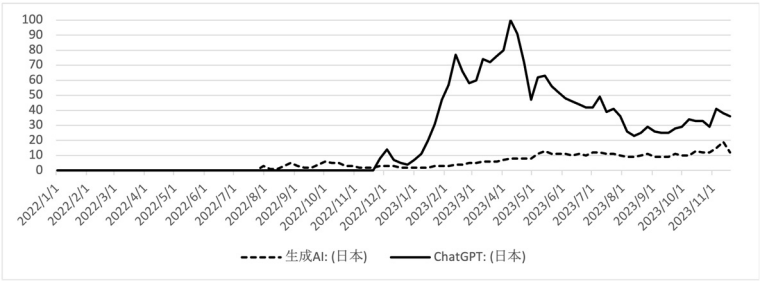


図 1　Google Trends によるキーワード「生成 AI」と「ChatGPT」の検索トレンド推移

注．縦軸 (0-100)は、特定の期間について、グラフ上の最高値を基準として Google 検索インタレストを相対的に表したもの。

表 1　生成 AI と教育に関する動向

日付	機関	文書名
2021/11/23	UNESCO	Recommendation on the Ethics of Artificial Intelligence (AI の倫理に関する勧告)
2022/03/22	東京外大	大学教育における AI について東京外国語大学としての教員向けガイドライン (国内大学の指針初出事例)
2023/03/29	(参考)英国	Generative artificial intelligence in education（教育における生成 AI）
2023/04/01	(参考)伊国	プライバシーへの懸念から国内で ChatGPT をブロック
2023/04/13	UNESCO	ChatGPT and artificial intelligence in higher education: quick start guide(高等教育における CHatGPT 利用のクイックガイド)
2023/07/03	(参考)米国	Generative Artificial Intelligence in education: What are the opportunities and challenges? (教育における生成 AI：期待と課題)
2023/07/04	文科省	初等中等教育段階における生成 AI の利用に関する暫定的なガイドライン：Ver1.0 機動的な改訂を想定
2023/07/13	文科省	大学・高専における生成 AI の教学面の取扱いについて（周知）
2023/07/24	私大連	大学教育における生成 AI の活用に向けたチェックリスト〔第１版〕
2023/09/08	UNESCO	Guidance for generative AI in education and research (本論文中略称：UNESCO ガイダンス)
2023/11/24	(参考)神戸市	「AI 活用条例」
2023/11/28	(参考)英国	Generative artificial intelligence in education call for evidence

1.3　国立大学の生成 AI に関する指針のデータベース化

本研究では、国立大学の生成 AI 対応を分析するため、web に公表されている指針を収集した[1]（図 2, 3）。同一大学の複数文書は別個の文書とし、国立大 86 校のうち 60 校から 80 件の文書を得た。指針を確認できなかった大学には、非公開の学内文書としている大学や文科省の文書を大学の声明にかえる大学が含まれる。テキスト分析には KH Coder（樋口、2004）を使用した。なお印刷版は紙面都合から文字の細かい判別が困難なため、PDF 版にあたられたい。高画質画像は筆者の web（http://fxyma.net） に掲載している。

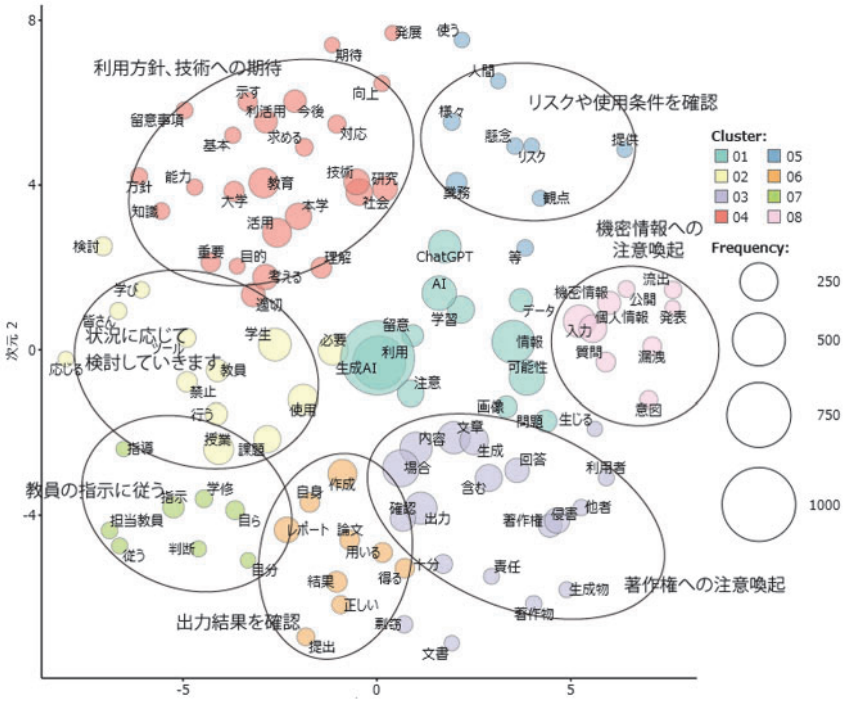


図 2　データベースの構造（多次元尺度構成法による）

※　印刷の都合上、次頁の図 3 の表題をここに示す。

図 3（左）　国立大学生成 AI 指針データベース内の各文書の文字数

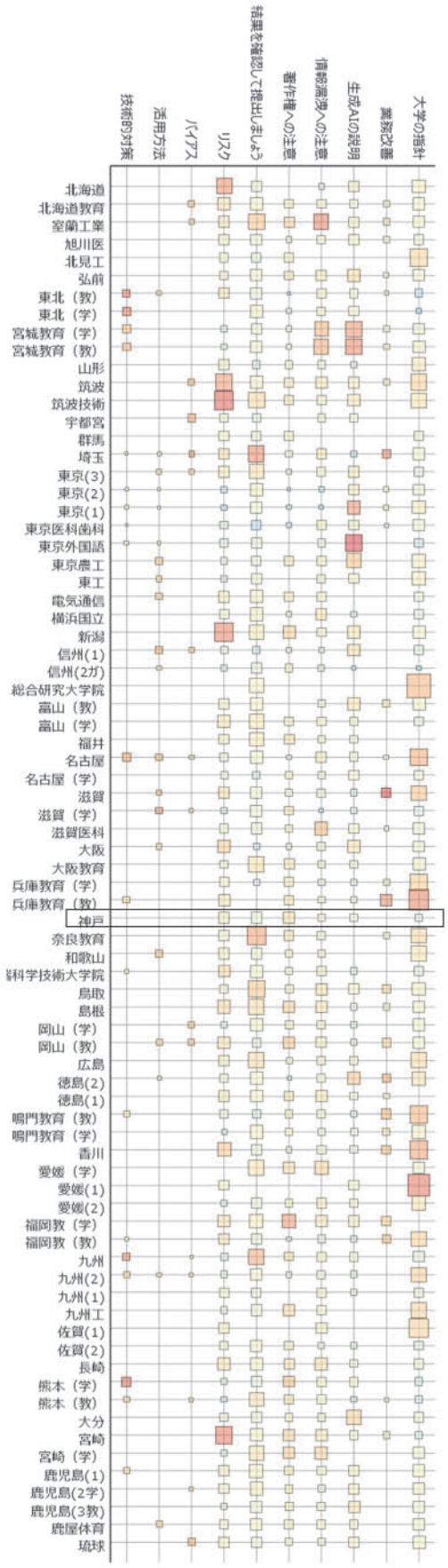
注：大学名の後ろの(1)(2)は提示された順序を示す。（教）は教職員向け文書、（学）は学生向け文書であることを示す。（ガ）はガイドライン文書（詳細なマニュアル的文書）を示す。

図 3（右）　バブルプロットによる各大学のトピックの分析

注：四角の大きさは各大学内での相対的な割合を示す。色は、トピックの軸で見たときに赤色（濃色）が濃いほど、そのトピックへの言及割合が多いことを示す。例えば早期に指針を公表した東京外大は「生成 AI の説明」の言及割合が多い。東京医歯大は総量は多いが、色を見ると各トピックに同じような文字数で言及している。神戸大（囲み）は「技術的対策、活用方法、バイアス」に関する言及がなく、満遍なく言及していることがわかる。図 2 左の単語量とあわせると、全体として特徴のない文章と分かる。

[1] 総語数：59047、異なり語；3136、文；2082、段落；1230、文書；80、文書平均文字数；1387.0

東京医科歯科	7813
東京(1)	4108
埼玉	4097
九州	4036
東京(2)	4030
滋賀	3722
東北（教）	3705
東京外国語	3551
熊本（教）	2514
山梨	2444
滋賀医科	2406
九州(2)	2298
徳島(2)	2110
名古屋（1）	2103
熊本（学）	1875
東北（学）	1854
滋賀（学）	1817
北陸先端科学技術院	1794
福岡教（教）	1663
愛媛(2)	1569
鹿児島(2学)	1521
鳴門教育（学）	1491
北海道	1482
室蘭工業	1408
東京(3)	1387
宮城教育（教）	1267
奈良教育	1264
鳥取	1243
山形	1237
香川	1231
鹿児島(3教)	1216
大阪	1208
宮城教育（学）	1201
筑波	1179
鳴門教育（教）	1159
広島	1114
弘前	1089
北海道教育	1071
東京工業	1025
横浜国立	984
旭川医科	898
信州(1)	885
鹿児島(1)	871
愛媛(1)	868
鹿屋体育	859
信州(2ガ)	850
岡山（教）	841
宮崎（学）	832
福岡教（学）	830
岡山（学）	828
兵庫教育（学）	819
宇都宮	818
筑波技術	790
徳島(1)	785
電気通信	774
佐賀(2)	773
島根	767
兵庫教育（教）	764
九州工	756
埼玉	751
大分	743
東京農工	736
宮崎	724
和歌山	658
長崎	629
九州(1)	623
琉球	613
名古屋（学）	604
福井	598
富山（教）	571
大阪教育	520
富山（学）	511
名古屋（全学）	497
新潟	457
愛媛（学）	404
神戸	386
北見工業	336
佐賀(1)	320
群馬	254
総研大	134



北海道	1482
北海道教育	1071
室蘭工業	1408
旭川医科	898
北見工業	336
弘前	1089
東北（教）	3705
東北（学）	1854
宮城教育（学）	1201
宮城教育（教）	1267
山形	1237
筑波	1179
筑波技術	790
宇都宮	818
群馬	254
埼玉	751
埼玉	4097
東京(3)	1387
東京(2)	4030
東京(1)	4108
東京医科歯科	7813
東京外国語	3551
東京農工	736
東京工業	1025
電気通信	774
横浜国立	984
新潟	457
山梨	2444
信州(1)	885
信州(2ガ)	850
総研大	134
富山（教）	571
富山（学）	511
福井	598
名古屋（1）	2103
名古屋（学）	604
名古屋（全学）	497
滋賀	3722
滋賀（学）	1817
滋賀医科	2406
大阪	1208
大阪教育	520
兵庫教育（学）	819
兵庫教育（教）	764
神戸	386
奈良教育	1264
和歌山	658
北陸先端科学技術院	1794
鳥取	1243
島根	767
岡山（学）	828
岡山（教）	841
広島	1114
徳島(2)	2110
徳島(1)	785
鳴門教育（教）	1159
鳴門教育（学）	1491
香川	1231
愛媛（学）	404
愛媛(1)	868
愛媛(2)	1569
福岡教（学）	830
福岡教（教）	1663
九州	4036
九州(2)	2298
九州(1)	623
九州工	756
佐賀(1)	320
佐賀(2)	773
長崎	629
熊本（学）	1875
熊本（教）	2514
大分	743
宮崎	724
宮崎（学）	832
鹿児島(1)	871
鹿児島(2学)	1521
鹿児島(3教)	1216
鹿屋体育	859
琉球	613

2.　大学が「生成 AI を導入する」とはどういういうことか：4 つの導入パターン

大学が「生成 AI を導入する」とは、具体的に何をすることか。ある大学が何も指針を出していない状況で、学生（組織の構成員）が生成 AI を使用して卒論の章立てを整理すると、その大学が「生成 AI を導入した」ことになるのか。2023 年 7 月 31 日、武蔵野大学は生成 AI を搭載した ICT ヘルプデスクチャットボットを導入した（武蔵野大学、2023）。この導入は国内大学で初めての生成 AI 導入事例とされる。その後、10 月 2 日には山梨大学が生成 AI を試験的に導入することが発表された（日本放送協会、2023）。また一部の例だが、専修大が演習系授業で ChatGPT を利用したことが「チャット GPT を導入した（原文ママ）」と報じられた（NEWSWITCH、2023）。個人的には「生成 AI を導入」というと教員や学生の個人ではなく組織レベルで契約した事例を想定するが、報道レベルでは授業で生成 AI を使用したものも「導入」として扱われている。

大学は一般企業とは異なり構成員（特に学生や教員）の独立性が高い。そのため教育、研究、業務等において明確に一切の生成 AI 利用を禁止しない限り、少なくとも報道レベルでは「生成 AI を導入した」とみなされることとなる[2]。以下に示すバリエーションは、LMS を用いた生成 AI システムを自大学で開発するか（パターン 1、2）、外部サービスを使うか（パターン 3、4）による、大まかな分類である。武蔵野大や山梨大の例ではパターン 2 が採用されている。「授業内で生成 AI を利用した」という事例は、パターン 3 に該当するものと思われる。その他、バリエーションとしては AI に詳しい学生や教職員が個人的に生成 AI を開発して利用するというパターンや、クラウドを組み合わせたシステムなどが考えられるため、これら 4 つのパターンから派生する導入バリエーションはサービスごとに幅広い。「生成 AI を導入」しようとするとき、その組織は導入パターンの選択肢を検討し、メリットとデメリットによってその組織にフィットする導入パターンを選択する必要がある。以下、本節では企業における LLM（大規模言語モデル）と生成 AI システムの導入パターンを参考に、大学が生成 AI を導入する際の代表的なパターンを示す。

2.1　パターン 1：自大学のオンプレミス環境で大規模言語モデルを開発する

パターン 1 は、オンプレミス、すなわちサーバやシステムなどを自組織で開発、管理、運用する方法で、学内の様々なデータを学習して LLM を構築する方法である。メリットとして学習元のデータが明確で、データ内のバイアスが明確になること、学内で開発と運用が完結するのでデータ流出のリスクが低いことなどが挙げられる。しかし LLM を自社で整備できるのは大企業であってもごく一部であり、大学が単独で LLM を構築することはコスト面、運用スキルを持つ人員、費用対効果を考えて現状の環境では現実的ではない。

パターン 1 の派生として、公開された既存の事前学習済みモデル（Pre-trained Model）を

[2] つまり、何も指針を出さない（禁止の宣言もしない）というのは事実上「導入した」ことになる。

下敷きにして、自大学のデータを学習させる方法もある（ファインチューニング)。公開された言語モデルを再利用し、自大学のデータを再学習させることで、コストを抑えつつ自大学の求めるタスクに特化したモデルを作成することができる。

2.2　パターン 2：自大学のシステムの内部で、生成 AI 事業者が提供する API を利用する

パターン 2 では、自大学で開発したシステムから、外部 API 経由で生成 AI の機能を呼び出すことで独自のシステムを構築する。例えば自大学の LMS（神戸大であればうりぼーネット）に生成 AI を利用するアプリを組み込むことで、LMS のひとつの機能としてチャット形式によるシラバスの検索、履修科目の推薦、履修登録や卒業条件の問い合わせなどを行わせることができるだろう。履修履歴、学習ポートフォリオをもとに就活や海外留学の資料作成をアシストするといった機能も実現できる。

パターン 2 は、自大学のニーズに合わせたシステムを構築できる。しかしパターン 3 に比べると開発コストがかかり、運用体制も整備しなければならない。パターン 3 に比べれば、フィルタリング機能によって、学生が個人情報を入力することを避けるように設定することも可能だが、システム上学内の情報を外部 API に渡すことになるので、データ再利用の禁止など契約を精査する必要がある。生成 AI を業務に導入した東京都は、パターン 2 を採用したものと思われる。Microsoft Azure を利用し、データの漏洩リスクを低減したことが示されているが、そもそも高機密情報（機密性 A）は入力不可と規定されている。

2.3　パターン 3：生成 AI 事業者の Web サービスを利用する

パターン 3 は、一般的な生成 AI の web サービスを利用するというものである[3]。言語系生成 AI であれば ChatGPT （OpenAI 社)、Bard （Google 社）などのアプリケーションを、学生や教員が利用して日常的な学習や業務に役立てる。メリットとしては、LLM やユーザインターフェース、各種アプリケーションを開発する必要がないために、金銭的・時間的コストを格段に抑えられる点が挙げられる。しかしデメリットとして、ユーザが入力した情報が外部に送られるため、個人情報や機密情報、未公開情報が流出する危険が伴う。

仮に、大学として生成 AI を導入するとも禁止するとも明確に宣言しない場合は、基本的には各教員、各学生の判断でパターン 3 を採用することとなる（あるいは、スキルのある教員や学生が独自に生成 AI のアプリケーションを開発する場合は、パターン 1 や 2 となる)。原稿執筆時点では、先述の武蔵野大学や山梨大学などごく少数の大学を除き、ほとんどの大学がパターン 3 を前提としていると思われる。

[3] このパターンは、デジタル庁（9 月 15 日）「ChatGPT 等の生成 AI の業務利用に関する申合せ（第 2 版）」における「（1） 約款型クラウドサービスによる生成 AI の業務利用」に相当する。

2.4　パターン4　生成AIが組み込まれた、学外のサービスを使用する

パターン 3 は生成 AI 事業者のサービスを直接利用するという方法であったが、パターン 4 は生成 AI 事業者ではない会社のサービスの中に組み込まれた生成 AI を利用するというものである。例えば、就職活動のエントリーシート（ES）の添削アプリケーションは、そのアプリの提供会社自体が大規模言語モデルや GPT を開発しているわけではなく、生成 AI 事業者の API を使用して添削サービスを展開している。こうしたパターンでは、学生や教職員は生成 AI を利用しているという自覚がないことも考えられる。つまり学生が就活の ES 添削アプリを使っていたら、実はそのアプリが内部で生成 AI を使用していたということは十分に想定できる。あるいは、本人は人間による通信添削を受けていると思っていたら、添削者が実は生成 AI を利用していたということも考えられる。

3.　大学における生成AIのリスク

ここでは大学における生成 AI 導入にあたって留意すべきリスクを整理したい。生成 AI はこれまでの情報セキュリティや情報リテラシー教育では対処できないリスクをはらんでいる。ここでは現段階での主要なトピックとして、情報の意図しない流出、意図しない剽窃、バイアス、評価に使用することの危険性、そしてハルシネーションを概説する。

3.1　情報の意図しない流出（教職員・学生）

文科省 が 2023 年 3 月に公表した「教育情報セキュリティポリシーに関するガイドライン」においては、情報セキュリティの「人為による脅威」として「悪意のある他者」「悪意のある関係者」「関係者の過失」に分類されている（表）。なお本ガイドラインは時期的に生成 AI を前提としていないため、「情報の意図しない流出」は脅威に含まれていない。

生成 AI は学生を含む関係者の、悪意の「ない」入力、さらに意図しない入力によって情報が外部に流出する。あるいは、あるアプリケーションの内部で生成 AI が用いられていた場合、本人は生成 AI に入力したとすら思っておらず、結果的に入力されていたこともあるだろう（2.4 を参照）。この場合、どこまで「関係者の過失」となるか、そもそも過

表 2　情報セキュリティの脅威

図表 3　情報セキュリティの脅威

脅威の原因		想定される脅威（具体例）
人為による脅威	悪意のある他者	情報資産の窃取・改ざんを目的とした標的型攻撃
	悪意のある関係者（教職員、児童生徒）	不正アクセスによる成績等情報の改ざん
	関係者（教職員、児童生徒）の過失	端末、物理的な電磁的記録媒体（USB等）の紛失
自然災害等		データの消失

出典：文部科学省「教育情報セキュリティポリシーに関するガイドライン」（図表 3）

失かは難しい問題である。さらに、そのときには問題ないと思っていても（問題ないとされていても）、あとになって結果的に避けるべきでした、ということも有りうる。生成 AI による情報流出で対策が難しいのは、「悪意のない他者（生成 AI 系サービスの提供元）」のサービスを「悪意のない関係者」が利用することによって情報資産が外部に流出することにある。LLM に学習されてしまったデータが、直接的に何かの質問の出力結果として提示されることは稀かもしれないが、特定の操作[4]によって個人情報を吐き出させることに成功した事例も報告されている（Nasr et al. 2023）。

国立大指針 DB では、重要な情報の「意図しない流出」への言及が 51 件見られた。しかし内容は個人情報や機密情報を入力しない、という注意喚起が大半であり、具体的にどのようなデータが機密情報かを例示した大学はごく少数である。だが大学 1 年生にとって何が機密情報かを判断することは容易だろうか。非母語話者の新入生は、入学時ガイダンス資料を生成 AI 系の翻訳ソフトに入力することが適切かを判断できるだろうか。私は台湾の大学で 2 年間勤務したが、国によって個人情報の常識や社会通念のレベルは大きく異なる。大学の構成員が一様な情報リテラシー、生成 AI リテラシーの基準を具えていると期待することは難しい。したがって、各大学は第一に、情報セキュリティ上の脅威として「意図しない流出」「悪意のない流出」を認識すること、そして第二に、意図しない流出を前提として、学内の情報資産のセキュリティレベルを定義し、各ドキュメントにセキュリティレベルを（もし流出した場合には悪意のある流出だと認定できるくらい）明示することが求められる。大学の情報資産とセキュリティレベルを検討し、続いて学内のドキュメントに明示する具体的な方法を検討は 4 章で行う。

3.2　意図しない剽窃（学生）

引用や剽窃についてはこれまでもリテラシーとして指導されてきたが、その前提は「誰かの考えを参照したら明記する」ことである。一方で生成 AI は、ハルシネーション（3.5 参照）を含んだもっともらしい文書を、十分な論拠や出典もなく提示する。Microsoft Bing はある程度 web 上の出典を含んで回答するが、ChatGPT や Bard は回答文に出典を含まない。出典を求めると架空の文書を返すことも多い。Web 検索と生成 AI の提示文書は本質的に異なることを学生が理解していないと意図しない剽窃を招くこととなる。

生成 AI の回答内容を利用することは、「出典のない孫引き」である。Web 上の誰か（複数）の、真偽不明の文書をもとに単語が組み合わされて回答されるもので、孫から祖父情報をたどることはできない。言わば実体のない「ゴースト孫引き」である。意図しな

[4] 一例として「poem」など非常に単純な文字列を無限に繰り返し回答するように指示すると、「数百回ほど poem を連呼したあたりで出力が発散し、「実在する企業の創設者兼 CEO の個人の個人情報などをうわごとのように吐き出し始めた」という事例も報告されている(Taniguchi 2023)。

い剽窃は、ゴーストが誰かの記述を学習データとして含んでいて、その誰かの考えを出典なく生成 AI が出力したものを、学生がアイデアや主張として採用した場合に起こる。学生としては生成 AI から得たアイデアだが、実は参照できないかたちで web 上の記述（を組み合わせたもの）を参照している。従って、孫引きの指導のように「生成 AI の提示内容は原典に当たりなさい」といった指示は（原典のない生成 AI には）無意味である。

3.3　LLM が抱える未知のバイアス

生成 AI のもとになる LLM（大規模言語モデル）の学習データに含まれるバイアスに注意することは、多くの大学の指針で指摘されている。特に UNESCO ガイダンス（UNESCO, 2023）は、生成 AI の学習データが、グローバルノース（北半球の裕福な地域）のデータに偏っていることを強調する。日本はグローバルノースの中でも非英語圏で、有色人種の割合が多いなど、バイアスによる不利を受けうる環境にもある。また、現在指摘されているバイアス（たとえば宗教バイアスに関しては（Abid, Farooqi, and Zou 2021）以外にも、未確認の潜在的なバイアス[5]があることに注意したい。このバイアスは気づきにくく、レポートの添削などでサブリミナルに思考が誘導される危険性があるため、今後の教育における生成 AI の活用においては大きな焦点となるだろう[6]。

3.4　評価に使用することの危険性

プライバシーの問題を棚上げすれば、入試や定期試験、あらゆるレベルの記述式課題の評価について生成 AI のサポートを受けることは可能である。また今後、そのような評価支援ツールが提供されるだろう。しかし多くの国立大が生成 AI の指針を出すなか、生成 AI を利用して学生を評価することを明確に禁止した事例は現段階では確認できない。文科省 7 月 13 日「教学面周知」でも禁止されていない。逆に、名古屋大学「教育研究における生成 AI の利活用について」は生成 AI を評価に利用することに前向きである。

> 教員にとっても講義資料の作成や課題の作成および評価に生成 AI を利用することでより効率的な教育を行うことが可能となります。今後は学生の生成 AI の利用を前提とした課題設定や評価方法を検討することで、生成 AI の高度な機能を

[5] 網羅的でない例として、性別、民族、人種、社会階層、地域、言語、年齢、学習データの情報源、障害、宗教、性的指向、政治、職業、感情、教育水準、消費文化など、十分に検証されていないバイアスは枚挙にいとまがない。

[6] EU がソーシャルスコアリングや人事評価において生成 AI を利用しないよう勧告するのも、言語モデルが潜在的に多くのバイアスを含み、既存の価値観の強化、格差の拡大にはたらくことに払拭しきれない懸念があるためである。したがって教育評価（添削まであらゆるレベル）に生成 AI を活用することは慎重さが求められる（次項参照）。

積極的に活用できる人材の育成を図ることが重要です。（名古屋大学、下線筆者）

大きなレベルでは、EU は政府や警察機関が市民を格付けするソーシャルスコアリングを最高リスク「レベル 1：許容できないリスク」として定義し、人事採用での利用も「レベル 2：ハイリスク」として規制を要する項目にしている。人事採用での利用がハイリスクであるならば、当然ながら入試や定期試験において生成 AI を用いて学生を評価することには大きなリスクがあることは明白である。

このリスクの根拠は、前節で挙げた生成 AI の学習データ内のバイアスである。生成 AI の元になる LLM のトレーニングでは web 上の言語データが収集されているが、その言語データには言語間でデータ量の差がある。言語、宗教、あらゆる価値観でマジョリティのデータほど多く学習されていることは避けられない。他にも「アノテーションバイアス」やデータの時代的な偏り、サンプリングの偏りなど、多くの面で生成 AI はバイアスを含みやすいシステムである。こうしたバイアスは、例えば学生の既存の価値観におけるマジョリティにそぐわない内容であった場合に、不当に低く採点されてしまうことも考えられ、それが繰り返されることによって既存のマジョリティの価値観が強化され、マイノリティが卑下され排除されていく悪循環を生む。こうしたバイアスの影響が未知数であるために、EU はソーシャルスコアリングや人事評価での使用をハイリスクとして規制するのである。今後の指針やガイドラインの提示においては、UNESCO ガイダンスや私大連ガイドラインで「大学が組織的に検討すべき項目」とされている部分と、教員が独自に判断できる項目を峻別していく必要がある。潜在的に高いリスクである部分についてはすべてを教員、学生に委ねるのではなく、組織として規制や注意喚起が必要になる。

3.5　ハルシネーション

生成 AI が現実に存在しない答え（幻想、Hallucination）を生成する特徴がある。例えば筆者が ChatGPT で「兵庫に独特な食事マナーがあるそうですが、それはなんですか」と入力したところ、5 回目の再生成で「*兵庫県独自の食事マナーの一つが、「播州まな板礼讃」と呼ばれる食事マナーです。一つのまな板に一品ずつ料理を盛り付け、その料理を一つずつ順番にいただきます。*」という頓狂なマナーを生成した。このように学習元のデータに存在しない概念を生み出すことをハルシネーションと呼ぶ。筆者は兵庫県出身ではないため実際にこういったマナーがあるのかもしれない。問題は、利用者に馴染みのない話題、とくにデータの少ない、バイアスの被害を受けやすい、さらに言えば大学において研究する価値のある話題ほど、生成 AI の解答の真偽を容易には検証できないことにある。これまで学生は参照情報の信頼性（実在する著者の信頼性）の確認を指導されてきたが、今後は実在しないデータ元の、実在しない情報ではないかというチェックが必要となる。

4.　2025 年度に求められる対応：現状の指針の問題点

生成 AI では、これまで大学が講じてきたセキュリティ（パスワードロック、定期試験問題を学外に持ち出さない等）では対処しきれないレベルでの情報流出が起こる。例えば定期試験問題のスペルチェックや留学生向けの翻訳を生成 AI が入ったソフトで行っただけで「学外に持ち出した」こととなるケースもありうる（意図しない流出)。したがって従前のセキュリティ対策を、教職員や学生個人の日常的なレベルで見直す必要に迫られる。

1 章で作成した DB からは、指針で内容や指摘している項目、語彙のバリエーションに大学間で大きな差はないが、内容の詳しさ（文書ごとの単語の量）には差があることがわかる。教員向けと学生向けを分け、活用事例を挙げつつ詳細な注意喚起をおこなう大学（例えば東北大は教員向けの本文だけで 5 千字を超える）もあれば、神戸大学（本文 376 文字）のように簡潔で総合的（抽象的）な注意喚起に留める大学もある。本節では 2024 年から 2025 年にむけて各大学が教職員および学生に対して周知すべき内容のうち、特に情報セキュリティとリテラシー教育（FD を含む）について講ずるべき対策を提言したい。

これらの提言の動機は、第一に神戸大学の 2023 年 4 月指針のように、抽象的な指示では大学の情報資産を十分に守ることができず、学生の個人情報が意図せず漏洩するリスクが高いという危惧である。そして第二に、多様な社会的、情報的背景を抱えた学生が一様の、高いリテラシーを持っていることが期待できない現状において、UNESCO や文科省などの求める、大学が周知すべき内容が十分に構成員に提示されていないことにある。

4.1　大学の情報資産と機密性を定義する

4.1.1　機密性定義の事例

デジタル庁「ChatGPT 等の生成 AI の業務利用に関する申合せ（第 2 版)」において、現在の ChatGPT が含まれる約款型クラウドサービスでは、省庁の業務において「要機密情報を取り扱うことはできない」(2) とされている。この要機密情報とは具体的にどのような文書を指すだろうか。例えば東京都は生成 AI の導入にあたり、個人情報等、機密性の高い情報は入力しないことを基本ルールとしている。東京都では表 2 のように機密性のランクを示したうえで、機密性 A は入力不可とし、安全性が確保された利用環境であれば、機密性 B までの情報を入力可能としている。

表 2　東京都の機密性のランク定義

機密性	セキュリティポリシー上の分類	情報資産の例
A	秘密文書に相当する情報資産	個人情報、契約関係情報、訴訟・審査請求に関する情報など
B	秘密文書に相当しないが、直ちに一般に公表することを前提としていない情報資産	内部通知、事案決定手続きを経ていない企画資料、経常業務の事務手順や実績など
C	機密性 A 又は機密性 B 以外の情報資産	東京都のホームページ等で公開済みの行政情報など

このように、大学において生成 AI を用いる際にも、教職員、学生が日々扱うデータの中で、生成 AI に入力してはならない情報を各大学が具体的に定める必要がある。情報資産の機密性ランクについては大学間である程度共有できるだろう。しかし学内システムに生成 AI を組み込んでいるか、データポリシー、オプトアウト（AI への学習拒否）がどのような契約になっているかなどは大学によって異なるため、どのレベルの機密性ランクをどのシステムに入力してよいかについては、大学ごとに独自の基準を定める必要がある。

4.1.2　大学の情報資産はどのようなものか

セキュリティレベルの定義にあたっては、学内にどのような情報資産があり、どの情報資産を生成 AI に対する管理項目とするかをリストする必要がある。大学の情報資産について、『大学 IR 標準ガイドブック』（井芹・近藤・松田、2022）は表 3 の分類を示す。

表 3　情報資産の分類

	非構造化データ　（テキスト）	構造化データ　（数値）
個別情報 **（構成員が個別に管理）**	・原著論文、著作 ・特徴ある研究や教育 ・受賞など	・競争的外部資金 ・研究業績
組織情報 **（管理担当者有り）**	・会議資料 ・規則 ・出版物	・学校基本調査 ・学生アンケート ・卒業生調査

管理項目のリストアップは、まずは「組織情報」から行うこととなるだろう。非構造化データであっても保守・保管の規則等が定められており、全容の把握は困難な課題ではない。一方で「個別情報」、なかでも「非構造化データ」は構成員個人の PC やサーバ、クラウド上などにあるため全容の把握が困難で、整理・管理が困難である。問題は、大学では個別情報については構成員一人ひとりがデータごとに判断しなければならないことであり、この構成員向けのガイドラインの策定は 2024 年度前半の各大学の急務である。

下に引用する「大学・高専における生成 AI の教学面の取扱いについて（周知）」（文科省、2023 年 7 月 13 日）においても、機密情報の具体的な内容は示されておらず、各大学・高専において指針を出すようにという勧告にとどまっている。

> 生成 AI への入力を通じ、機密情報や個人情報等が意図せず流出・漏洩する可能性等があるため、一般的なセキュリティ上の留意点として、機密情報や個人情報等を安易に生成 AI に入力することは避けることが必要と考えられること。……また、生成 AI の種類によっては、入力の内容を生成 AI の学習に使用させない（オプトアウト）ことができること。

個人的には、各大学はこの周知文書を構成員個人にそのままパスすること、つまり具体的な基準、指針を示さずに「機密情報を入力するな」と指示することは避けなければなら

ず、次項以降に示すようにセキュリティレベルの定義と標示を行うべきだと思う。

4.1.3　各大学は情報資産のセキュリティレベルをどう定義するか

情報資産を機密性等によって分類する方法として、文科省「教育情報セキュリティポリシーに関するガイドライン（令和 4 年 3 月）」の 1 章 3 項（p.32）以降が参考になる。教育機関を管轄する地方自治体向けの文書であるが、一般企業向けの各種ガイドよりは大学と親和性が高い。同文書では情報資産を分類し、分類ごとに管理体制を定めることとしている。情報資産は「機密性」「完全性」及び「可用性」により分類されている（同文書の図表 5（P.37）参照）。大学の規模や学生への周知のしやすさは大学によるため、たとえば管理部門、教職員レベルでは機密性、完全性、可用性で分類しておき、学生に周知する際には理解しやすい「重要性」による分類（表 4）で示すという運用も考えられる。

表 4　文科省資料による「情報資産の取扱例」

情報資産の分類								
重要性分類	定義	組織外部への持ち出し制限*	情報の組織外部への送信**	情報資産の運搬***	組織外部での情報処理****	使用する電磁記録媒体	情報資産の保管	情報資産の廃棄
Ⅰ	セキュリティ侵害が教職員又は児童生徒の生命、財産、プライバシー等へ重大な影響を及ぼす。	本ガイドラインに準拠していることを確認した上で業務遂行上必要な場合には、情報セキュリティ管理者の判断で持ち出しを可	限定されたアクセスの措置がとられていること*****	鍵付きケースへの格納	禁止	施錠可能な場所への保管	・耐火、耐熱、耐水、耐湿を講じた施錠可能な場所に保管（電子データの場合もこれらの対策に準じたサーバに保管） ・情報資産を格納するサーバのバックアップ ・6か月以上のログ保管 ・サーバの冗長化（推奨事項） ・オンラインで情報資産を利用する場合は通信経路の暗号化を実施 ・保管場所への必要以上の電磁記録媒体の持ち込み禁止	電子記録媒体の初期化、復元できないようにして廃棄
Ⅱ	セキュリティ侵害が、学校事務及び教育活動の実施に重大な影響を及ぼす。	同上	同上	同上	安全管理措置の規定が必要	同上	同上	同上
Ⅲ	セキュリティ侵害が、学校事務及び教育活動の実施に軽微な影響を及ぼす。	情報セキュリティ管理者の包括的承認で可	同上	同上	同上	同上	・耐火、耐熱、耐水、耐湿を講じた施錠可能な場所に保管（電子データの場合もこれらの対策に準じたサーバに保管） ・情報資産を格納するサーバのバックアップ（推奨事項） ・一定期間以上のログ保管 ・サーバーハードディスクの冗長化（推奨事項） ・オンラインで情報資産を利用する場合は通信経路の暗号化を実施 ・保管場所への必要以上の電磁記録媒体の持ち込み禁止	同上
Ⅳ	影響をほとんど及ぼさない。							

出典：「教育情報セキュリティポリシーに関するガイドライン　図表 6」

4.2　具体的なガイドラインの制定：セキュリティレベルによる入力可否を規定する

学内の情報資産のセキュリティレベルの定義が完了すれば、次にレベルに応じた生成 AI への入力可否を規定することとなる。入力できる情報のレベルは各大学のシステムの

設計に依存するが、東京都の事例を参考にすれば**表 5** のように示すことができる。

表 5　セキュリティレベルの定義例

重要性分類	定義	例	教務システム	ChatGPT, Bard…
A	侵害が構成員の生命、財産、プライバシー等　へ重大な影響を及ぼす。秘密文書に相当する情報資産。	個人情報、契約関係情報、訴訟・審査請求等に関する情報 等	入力不可	入力不可
B	秘密文書に相当しないが、直ちに一般に公表することを前提としていない情報資産	内部通知、事案決定手続きを経ていない企画資料、経常業務の事務手順や実績	入力可	条件付きで可 ・要オプトアウト ・学習禁止
C	機密性 A 又は機密性 B 以外の情報資産	HP 等で公開済みの行政情報など	入力可	入力可

4.3　学内のドキュメントにセキュリティレベルを標示する

セキュリティレベルの定義と、レベルごとの生成 AI への入力可否を規定したからといって、その一覧表を学内に周知するだけでは情報漏洩は十分に防ぐことはできない。

学内でセキュリティレベルを明示する方法として、いくつかの暫定的なアイデアを示す。悪意のない、意図しない流出に対応するためであることを踏まえ、読者にわかりやすく表示し、高リスクのドキュメントほど多重に明示・標示することが重要と思われる。

- 文書のヘッダーやフッター、右肩に「レベル A」というテキストボックスを入れて明示する
- ファイル名の末尾に特定の文字列（例えば”Lv-A”など）を入れる
- 電子ドキュメントのメタ情報としてセキュリティレベルの情報を埋め込む

ヘッダーやファイル名に特定の文字列を入れるという方法であれば、人間にも理解しやすく、また学内システムに生成 AI を導入した場合に、その文字列を含むファイルはシステムに読み込ませない、というフィルタリングも容易になるだろう。既存のファイル名に一括で文字列を加えることも簡単である。なお些末なことではあるが、ファイル名等に展開することを考えるとローマ数字の重要性分類にするよりは、東京都の機密性分類のように ABC にするほうが良い。

こうした手法は、（多くの大学が指針で求めるように）読者、学生にセキュリティレベルを判断させるのではなく、文書の作成者側が積極的にセキュリティレベルを明示することが特徴である。例えば著作権のマーク（©）やクリエイティブ・コモンズマーク（CC）を作者自身が示すように、文書の作成者（大学や教員）が各文書にマーキングを行うことが求められる。なお、フェイルセーフ（ミスしたときに安全なデザイン）の観点からすると、重要な書類に標示するよりも、「生成 AI に入力してもよい書類」に標示をする（標示がないときには入力してはいけない）とするほうが望ましい。生成 AI に情報を流出させないシステムエンジニアリングと同様に、組織内の関係者にどのように機密性を提示する

かという人間側の情報デザインも、今後の大きな研究課題である。

5.　学生へのガイダンスと FD 項目の提案：公的ガイドラインのポイント解説

本節では、学生へのガイダンスと FD の具体的な項目の選定にむけて、大学が総論的な指針を超えた詳細なガイドラインを策定する上で参考になると思われる公的ガイドライン（UNISCO や文科省、私大連）のいくつかのポイントを解説したい。

5.1　「学生に対して生成 AI の理解を深める情報リテラシー教育」とは何か

2023 年 11 月現在では、生成 AI の活用に対して多くの大学が指針やガイドラインを提示しているが、具体的に情報リテラシー科目として生成 AI の理解を深めるような講義は十分に展開されていない。2024 年度以降、新入生のみならず在学生にも生成 AI を前提とした情報リテラシー教育を施す必要があるだろう。

共通教育として最低限のリテラシー教育に含むべきキーワードを UNESCO や文科省、私大連などの公的なガイドラインから抽出すると、トピックとしては著作権、アカデミック・インテグリティ（学問の誠実性)、情報セキュリティ、データバイアスなどが挙げられる。また生成 AI とつきあううえでのキーワードとしては、オプトアウト、プロンプト（プロンプトエンジニアリング)、ハルシネーションといった専門用語が必要となる[7]。

- 人工知能の概要：強い／弱い AI、機械学習、言語モデル（確率的オウム）など
- 生成 AI の概要：生成モデルとその応用
- 対話型生成 AI（ChatGPT など）を用いた学習の例

  UNESCO クイックガイド[8]に、生成 AI が果たすことができそうな役割がいくつか示されている。ソクラテス的な問答、「壁打ち」相手、チューター（自分の小テストの解答と点数、模範解答を入力し、不足点や改善点を尋ねるなど)、共同デザイナー（レポートの草稿の添削など)、モチベーターなど。
- 生成 AI を用いて良い場面と用いるべきではない場面
- 学習データのバイアス（3.3 参照）
- ハルシネーション（3.5 参照）
- プロンプトエンジニアリング：生成 AI の出力を得るための指示をプロンプトとよび、一般に① 命令（例：翻訳してください)、② 背景・文脈（例：あなたは優秀な翻訳者です)、③ 入力データ（翻訳したい文書)、④ 出力指示（例：専門用語を使わずに）

[7] これらの専門用語の一般的な認知度は 2023 年末時点では高くはないと思われ、情報系の専門用語の扱いだと思われる。各種調査による「生成 AI」という単語の認知度調査の結果が 2023 年の初頭と終盤で大きく向上していることを踏まえると、2025 年に向けては生成 AI に関連する用語の一般的な認知度もかなり向上するのではないだろうか。

[8] 原案は Mike Sharples による。

からなる。適切なプロンプトを入力して望む結果を生成するスキルや、有効なプロンプトを開発する手法はプロンプトエンジニアリングと呼ばれ、2025 年以降、特に大学生には必須の情報スキルとなる[9]。

- オプトアウト：入力する情報をコントロールする方法。AI 学習データとして再利用させない。未公開データや個人情報の外部流出を防ぎリスクをマネジメントする。

5.2　「教員に対して生成 AI の理解を深める FD」について

私大連ガイドラインで言及されている「教員に対して生成 AI の理解を深める FD」とはどういった内容を含むだろうか。授業運営での工夫は国立大生成 AI 指針 DB 上でもブレインストーミング、アイデア出し、プレゼンテーションの「壁打ち」、校正、翻訳などが言及されている。これら教学的な活用は FD の主要トピックだが、同時に生成 AI の懸念点も含むべきである。懸念点は 3 章に挙げたリスクが中心に、以下のような点も含まれる。

5.2.1　生成 AI の使用の有無は検出できない

web 上からレポートと一致する文字列を探すコピー&ペーストの検出ツールは有効な一方で、同様の感覚で「生成 AI 検出ツール」を利用することは現段階では難しい。福岡教育大学（2023 年 10 月）も「AI が生成した文章かを判定するツールを学修成果の評価等に使用する場合でも、その結果を過信しないようにしてください」としている。

5.2.2　「課金すれば良いレポートが書ける」は許容されるか？

生成 AI は記述式課題の回答を瞬時に提供することから、「採点・評価において、生成 AI の使用の有無に配慮し、学生間の公平性が担保されるよう留意してください。（JAIST）」という事例のように、評価上の公平性が問題となる。現状の指針では、生成 AI の使用の有無が公平性の焦点となっているが、2024 年以降は使用する生成 AI のバージョンや性能も論点となるだろう。原稿執筆時点では、たとえば ChatGPT は GPT-3.5 は無料、GPT-4 という学習データが多い高性能なバージョンは月額 20 US ドルである。この問題については、道具の良し悪しは許容されるという考えと、「使用料を払えば良いレポートが書ける」という状況は公平ではなく許容されないという考えの両論を立てることができる。

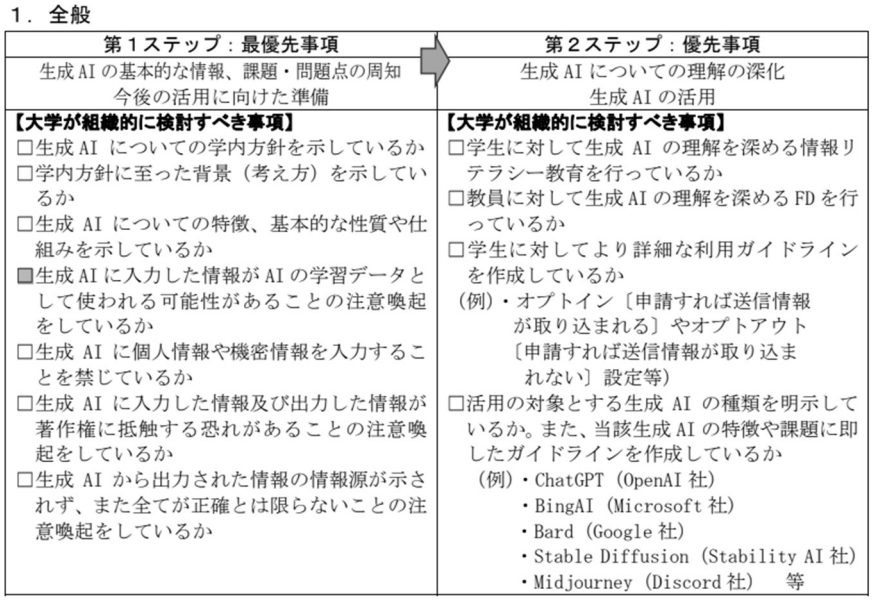
許容できるという立場からは、生成 AI に限らずこれまでも分析 PC のスペックによる精度や速度の差異はあったし、書籍を購入することでより充実した資料にあたれるという面も許容されてきた。スポーツでは高価な道具の使用によって良いスコアを得ることもあったことから、学生は課金によってより良い学習成果を生む権利があり、生成 AI も同様

[9] 画像生成系では文章系以上に精密なプロンプト（ネットスラングで「呪文」）が求められ、思い通りの画像を生成するために様々な工夫、エンジニアリングが行われている。

に許容されるという議論ができる。AI を使えば自動的に成果が得られるわけではなく、適切なプロンプトで指示することも個人の情報スキルであるため、生成 AI の使用の有無という点では許容されないにしても、GPT-3.5 と 4 の差異程度は許容されるという考えも可能である。一方で許容されないという立場にたてば、生成 AI はこれまでの道具と異なり、学修成果（に類似するもの）を直接的に生成するものであり、学生個人の能力を評価したことになるかが疑問であり許容されるべきではない、といった議論が可能である。

5.3　「活用の対象とする生成 AI の特徴や課題に即したガイドライン」について

私大連の「大学教育における生成 AI の活用に向けたチェックリスト〔第 1 版〕」には、大学が組織的に検討すべき事項の「優先事項」として「活用の対象とする生成 AI の種類を明示しているか。また、当該生成 AI の特徴や課題に即したガイドラインを作成しているか」というチェック項目がある（図 2)。前段の「活用の対象とする生成 AI の種類」が意味するのは、まずは言語系、画像系、動画生成系、音声生成系、マルチモーダル系などの種類の指定である。マルチモーダル系は複数のモーダル、つまり言語系と画像系などを組み合わせることのできるもので、ある文章の要約を絵で表現するといったことができる。現在は限られたサービスのみで提供されており、ChatGPT 4 はマルチモーダルに対応している。こうした「種類」のほかには、2 章の導入パターンで示したような、学内のシステムのみ許可するといった種類の指定、あるいはオプトアウト機能があるもののみ、といった指定も考えられる。後段の「当該生成 AI の特徴や課題に即したガイドラインを作成しているか」については、ChatGPT、Bard など個別のサービスごとにガイドラインを作成す



1．全般

第１ステップ：最優先事項 生成 AI の基本的な情報、課題・問題点の周知 今後の活用に向けた準備	第２ステップ：優先事項 生成 AI についての理解の深化 生成 AI の活用
【大学が組織的に検討すべき事項】 □生成 AI についての学内方針を示しているか □学内方針に至った背景（考え方）を示しているか □生成 AI についての特徴、基本的な性質や仕組みを示しているか ■生成 AI に入力した情報が AI の学習データとして使われる可能性があることの注意喚起をしているか □生成 AI に個人情報や機密情報を入力することを禁じているか □生成 AI に入力した情報及び出力した情報が著作権に抵触する恐れがあることの注意喚起をしているか □生成 AI から出力された情報の情報源が示されず、また全てが正確とは限らないことの注意喚起をしているか	**【大学が組織的に検討すべき事項】** □学生に対して生成 AI の理解を深める情報リテラシー教育を行っているか □教員に対して生成 AI の理解を深める FD を行っているか □学生に対してより詳細な利用ガイドラインを作成しているか （例）・オプトイン〔申請すれば送信情報が取り込まれる〕やオプトアウト〔申請すれば送信情報が取り込まれない〕設定等） □活用の対象とする生成 AI の種類を明示しているか。また、当該生成 AI の特徴や課題に即したガイドラインを作成しているか （例）・ChatGPT（OpenAI 社） ・BingAI（Microsoft 社） ・Bard（Google 社） ・Stable Diffusion（Stability AI 社） ・Midjourney（Discord 社）　等

図 2　私大連文書より、「全般」に関するチェックリスト

出典：私大連「大学教育における生成 AI の活用に向けたチェックリスト〔第 1 版〕」

ることを求めているが、管見の限りではこのレベルでガイドラインを提示している大学はない。研究レベルでは（藤村、 2023）の「Bing と Bard の機能等比較表」がみられる。

ここでは藤村の比較表の観点を参考に、後段に対応するような、学生や教員が各サービスを利用する上でツールを比較できるようなガイドラインのサンプルを示したい。

ガイドラインは下記のような比較表を用いることで、教員や学生が用途ごとにツールを比較できる。項目は一例であるが、大学での用途として重要な点をいくつか指摘したい。まず情報セキュリティの観点から、セキュリティレベルに応じた入力可能な情報の指定、オプトアウト機能の有無の表示は必須である。次に学習ツールとして、情報源の明示が可能かどうか、学習元データは公開されているか、ファイル読み込みができるか、プロンプトの保存は可能かなどが重要となる。データバイアスへの対応は現在議論されている最中のため、社会的な動向に応じて表示されることが望ましい。リスクへの注意だけでなく、学習での活用を後押しするような比較表を目指すならば、サービスごとに得意なタスクや機能を紹介し、プロンプトの例を提示することが望ましいだろう。

表 6　「活用の対象とする生成 AI の特徴や課題に即したガイドライン」の項目案

項目	**記載例**
サービス名	（例）ChatGPT
入力可能情報	利用不可／レベル A まで／レベル B まで　など
料金	有料／無料／M365 に付随 (Copilot 365)など
生成 AI エンジン	ChatGPT-4 / LaMDA など
オプトアウト	可能／不可能／条件付き
情報源の明示	あり／なし
学習元データ	不明／一部公開／全公開
データバイアスへの対応	
プロンプトの保存と利用	あり／なし
ファイルの読み込み	PDF, docs, xml…
教育的配慮	
利用年齢	18 歳以上など
有効なプロンプトの例	（例）
機能	汎用型／翻訳／要約／プレゼン特化／プログラミング支援　など
タスク	要約　翻訳　情報収集　など
特徴	翻訳の訳語修正ができる　など
授業での活用アイデア　ブレインストーミング、反論役…	
自習での活用アイデア　レポート添削、発音矯正、スペルチェック…	

6.　おわりに

Web 上に公表された指針の限りでは、国立大学で明確に生成 AI の使用を全面禁止した事例はなかった。ほとんどの大学が、注意事項を提示した上で、慎重な利用を求めているというのが 2023 年末の現状である。

本研究が問題視するのは、UNESCO ガイド[10]あるいは私大連ガイドラインで「大学が組織的に検討すべき事項」として挙げられている内容が、多くの大学で教員の判断に任せられ、また学生に対しては「教員の指示に従う」あるいは「十分に注意して利用」という抽象的な指示に留まっているという現状である。国立大の指針文書を概観すれば、構成員（学生・教職員）には生成 AI 利用への自己責任を求める論調が主勢に思える（例えば岡山大学）が、大学側が責任をもって管理すべき項目については言及が十分ではない場合が多く、穿った見方をすれば組織としての責任を回避し、個人に転嫁するための文書と思われても致し方ないのではないだろうか。「担当教員の許可を得よ」「所属部署の許可を得よ」とする文書もあるが、その場合は許可の基準が十分に示される必要がある。組織的に検討すべき項目と、各員の判断に委ねられる項目の峻別は 2024 年から 2025 年の大きな課題である。特に 3.4 で議論した通り、評価に関しては組織的に検討されるべき項目である。

2025 年に向けては、生成 AI を用いるのに適した場面と用いてはならない条件を明示しつつ、構成員の主体的な利活用を促すような、具体的な使用場面に即したガイドラインが求められる。とりわけ、UNESCO で強調されているグローバルノース（DB 内言及 0）の価値観、バイアス（12）、偏見（14）、差別（8）は重要な観点だが、生成 AI というシステムに潜在的なものも多く、事務的な通達だけでは十分に防ぐことはできない。またインテグリティ（4）、ハルシネーション（2）、オプトアウト（22）といった、大学における生成 AI の使用の上で特徴的な概念を積極的に発信することが求められる。

謝辞

技術的側面についてご助言いただいた、青森大学小野淳平先生に感謝申し上げます。

参考文献

井芹俊太郎, 近藤伸彦, 松田岳士. 2022.『大学 IR 標準ガイドブック: インスティテューショナル・リサーチのノウハウと実践』. インプレス r&d.

東京外国語大学. 2023.「大学教育に関するガイドライン」.March 22, 2023. https://www.tufs.ac.jp/education/guideline/.

名古屋大学. 2023.「教育研究における生成 AI の利活用について」. 名古屋大学.

[10] 超国家的枠組みから国家、大学等組織、個人まで各レベルで取るべき対応を明確に分けている。

日本放送協会. 2023. 東京都 生成 AI を都庁内の全局で導入 業務活用を開始. NHK NEWS WEB. November 16, 2023.

樋口耕一. 2004. “テキスト型データの計量的分析.” 理論と方法 19 (1): 101–15.

福岡教育大学における ChatGPT 等の生成系 AI の業務利用時での注意点について. 2023. 国立大学法人 福岡教育大学. October 23, 2023. https://www.fukuoka-edu.ac.jp/.

藤村裕一. 2023. 生成 AI の教育利用に関する研究. 日本教育工学会研究報告集 (2): 75–82.

武蔵野大学プレスリリース. 2023. 「生成 AI 搭載の ICT ヘルプデスクチャットボットが誕生」 July 31, 2023.

文部科学省. 2023. 「教育情報セキュリティポリシーに関するガイドライン」公表について：文部科学省. March 4, 2023.

Abid, Abubakar, Maheen Farooqi, and James Zou. 2021. “Persistent Anti-Muslim Bias in Large Language Models.” arXiv. http://arxiv.org/abs/2101.05783.

Nasr, Milad, Nicholas Carlini, Jonathan Hayase, Matthew Jagielski, A. Feder Cooper, Daphne Ippolito, Christopher A. Choquette-Choo, Eric Wallace, Florian Tramèr, and Katherine Lee. 2023. “Scalable Extraction of Training Data from (Production) Language Models.”

NEWSWITCH. 2023. 「ベネッセ、専修大が導入…生成ＡＩは学習のあり方を変えるか」, https://newswitch.jp/p/38185. （最終アクセス：2023 年 12 月 1 日）

Taniguchi Munenori. 2023. 「ChatGPT に同じ言葉を連呼させると、壊れて学習データ(個人情報入り)を吐き出す？Google DeepMind 研究者らのチームが論文発表」. TechnoEdge. https://www.techno-edge.net/article/2023/11/30/2365.html.

UNESCO. 2023. “Guidance for Generative AI in Education and Research |.” September 8, 2023.