The Computer & Internet Lawyer

Volume 40 ▲ Number 7 ▲ July-August 2023

Ronald L. Johnston, Arnold & Porter, Editor-in-Chief

Dos and Don'ts for Developing, Extending and Using Generative AI Models

By Barath R. Chari, Scott A. McKinney, Gary R. Greenstein, Laura De Boel, Maneesha Mithal, Stefan Geirhofer, Kristina Wang, Mary O'Brien and Emily Chan

1. INTRODUCTION

Generative AI (GenAI) refers to a category of artificial intelligence (AI) models capable of generating text, images, music, or other content in response to a user's input prompts and based on training data embodied in the AI model.

While GenAI models show great promise, companies should consider the legal, commercial, and ethical risks posed by this technology. All companies that intend to use GenAI tools for day-to-day business tasks or to develop their own GenAI applications should evaluate the risk/reward tradeoffs and develop an AI policy that provides guidance to employees about the acceptable use of GenAI tools and the risks that could result from unauthorized use.

This article is organized as follows:

- Section 2 discusses the relationship among training data, input prompts, and output content by treating the GenAI model as a "black box" in order to identify key legal and commercial risks;
- Section 3 analyzes the dos and don'ts for typical AI uses cases, ranging from developing GenAI models from scratch and extending foundational models for specific applications to simply using GenAI tools for day-to-day business tasks; and
- Section 4 summarizes key provisions that should form part of an organization's AI usage policy.

The authors, attorneys with Wilson Sonsini Goodrich & Rosati, may be contacted at bchari@wsgr.com, smckinney@wsgr.com, ggreenstein@wsgr.com, ldeboel@wsgr.com, mmithal@wsgr.com, sgeirhofer@wsgr.com, kristina.wang@wsgr.com, mobrien@wsgr.com and emily.chan@wsgr.com, respectively.

2. SPOTTING KEY LEGAL ISSUES RELATED TO TRAINING DATA, INPUT PROMPTS, AND GENERATED CONTENT

At a Glance:

- AI developers should be judicious in their selection of training data and evaluate whether it is appropriately licensed or otherwise available for training purposes.
- Users should understand whether their input prompts may be fed back into the GenAI model and could thus potentially "leak" sensitive information to third parties.
- AI developers should monitor risks resulting from the use of generated output from an infringement, breach-of-contract, reputational, and regulatory perspective.

GenAI models employ machine learning techniques to generate content based on input prompts and probabilistic model parameters "learned" by the GenAI model through exposure to vast amounts of sample or training data. Because GenAI models are not programmed to make decisions in an "if this, then that" fashion, the inner workings of a GenAI model can remain opaque even to the developers of the GenAI models. For purposes of this article, we view GenAI models as a "black box," and focus on the models' inputs and outputs:

- Training data used to train the GenAI model;
- Input prompts submitted by users of the model; and
- Output generated by the model.

Throughout this article, we use the term "AI developers" to refer to entities that develop pre-trained, general purpose "foundational" GenAI models and "AI application developers" to refer to entities that build on foundational models developed by others.

2.1 Training Data

GenAI models are trained on vast sets of training data. The availability of massive amounts of computing power and storage through the proliferation of cloud computing has made sophisticated models and new AI applications possible. GenAI models typically rely on machine learning, a process through which computer programs are able to "learn" through exposure to large quantities of training data without being explicitly programmed. Large corpuses of training data are often required to achieve adequate GenAI model performance, and may be refined by using human interactions with the GenAI model.

While developers of GenAI models need large training datasets, they should be judicious in their selection of appropriate training data. As a general matter, AI developers should evaluate the types of data they use for training purposes and determine whether it is appropriately licensed or otherwise available for training purposes. For AI application developers that build on top of foundational GenAI models offered by third parties, it is important to seek as much information as possible on the types of data that the foundational model was trained on for the same reason.

Training data that contains personal data (which is a broad concept, both in the U.S. and the EU) could raise significant privacy law compliance obligations. This includes U.S. and EU legal requirements governing notice, consent, and data subject rights (e.g., right of individuals to access, delete, or correct information). In addition to complying with these general obligations, companies should evaluate whether they are using any particularly sensitive personal data in their training sets, which may include voice data, biometric data, location information, children's information, and information about race or ethnicity. For these categories of data, heightened requirements will likely apply. For example, developers may be required to conduct data protection impact assessments.

The following types of training data are often encountered in practice:

- Training data collected through "web scraping," of websites that are in front of a paywall or not otherwise access restricted, could result in a number of potential claims, including:
  - (a) Breach of contract for violation of the terms of use governing the scraped website;
  - (b) Trespass to chattels, if the web scraping causes system degradation or performance issues;
  - (c) Copyright infringement for unauthorized duplication of copyrighted material on the website;
  - (d) Unfair competition and misappropriation due to "free riding" on the work and expense of website operators that devote substantial

efforts and incur significant costs in developing and operating a website; and

(e) Anti-circumvention claims, if the scraping bots circumvent access control mechanics in violation of the Digital Millennium Copyright Act (DMCA);[1]

- Research datasets licensed from universities or research institutions can be a valuable starting point, but it is important to examine whether the license terms permit nonresearch or commercial use of the dataset, and if so, under what terms and conditions;

- Training data collected from a company's own end users or customers (including employees and contractors) is likely subject to the company's contracts with its end users or customers (such as terms of use or a subscription agreement) as well as its privacy policy and other disclosures. Companies should be aware of any restrictions and limitations on the use of data in its own contracts, disclosures, and policies and make prompt updates if necessary, taking care not to apply these policies retroactively and to obtain appropriate user consent as necessary; and

- Training data subject to open source licenses may impose additional restrictions on how the data can be used. For example, as discussed in Section 2.3.3 below, open source licenses may require attribution or copyright notices, especially if recognizable portions of the training data find their way into generated content.

In addition to the specific concerns listed above, to the extent the training process involves making copies of the training data (for example, during pre-processing for training neural networks), the duplication of training data may constitute copyright infringement if the developer is not authorized or permitted to reproduce the data. Companies may claim that such copying is a fair use, as is outlined in Section 2.3.1 below.

2.2 Input Prompts

Input prompts are submitted to and processed by GenAI models to create responsive output. In addition to text, input prompts could also include images, source code, audio, video, or other forms of data.

Users of GenAI models, including AI application developers, should understand the terms under which the GenAI model is licensed. Some GenAI models require end users to provide rights to use their input for the purpose of improving the model. In this manner, input prompts become part of the training set, which could result in portions of the input being provided to other users of the GenAI model. Particularly sensitive information in this context could include the following types of data:

- If the input data constitutes a company's or organization's trade secret, then providing that data to a GenAI model operated by a third party could result in that information losing its trade secret protection under applicable law. A company's or organization's trade secrets may include any nonpublic information that constitutes or relates to the company's or organization's proprietary technology, code base, client list, business strategies, or research plans;

- Similarly, information subject to the attorney-client privilege or work product doctrine could lose its privileged status if it is deemed to have been disclosed to others by virtue of its submission to the GenAI model. Waiver of privilege not only risks disclosure of what is likely sensitive information of a legal nature but would also require producing that information if discovery were sought in connection with a lawsuit;

- Companies should carefully evaluate if information submitted to a GenAI model might constitute material nonpublic information (for securities law purposes). This information could include, for example, earnings information, business plans, strategies, or information related to M&A and other corporate transactions. Similarly, information relating to potential or existing litigation, or to internal or government investigations should be afforded extra caution;

- If the input data constitutes the confidential information of a third party or is otherwise subject to confidentiality obligations under an existing contract between the company and a third party, then uploading or submitting that information to a GenAI platform could breach the user's confidentiality obligations, entitling the third party to seek injunctive relief or other remedies that are available under the existing contract with the third party;

- Data licensed from a third party could also be subject to terms that restrict how the data can be used.

Submitting third-party data to a GenAI model could inadvertently breach applicable license terms and subject the company to liability; and

- It will be important to explain to individual users how any personal data they provide in queries to AI-based systems will be used in the model and comply with relevant privacy laws on this issue.

In view of this potential "leakage" of sensitive information, users of GenAI models should carefully review the applicable terms of use and understand how their input prompts are used. Conversely, the providers of GenAI models should clearly specify in their terms of use or disclosures how input prompts are used and whether users have the ability to exercise choices over whether their input prompts are used to further train the GenAI model. Commercially available and emerging enterprise GenAI applications may be very different in their approach compared to free-to-use or consumer facing GenAI applications when balancing these interests.

2.3 Generated Content

The output of GenAI models, while responsive to input prompts, necessarily reflects a combination of a myriad of sample data points found in the training data set. This section examines risks arising from the use of this output.

2.3.1 Copyright Infringement

In some cases, the output of GenAI tools includes recognizable or identifiable portions of training data. This creates a risk of infringement liability for violating third-party copyrights by reproducing copyrighted material without authorization or permission. In addition, content generated by GenAI models may constitute an unauthorized derivative work of the copyrighted material used in training the underlying GenAI model, which could also create a risk of infringement liability.

AI service providers could try to rely on the "fair use" doctrine as a potential defense against infringement claims. At a high level, a fair use analysis examines four factors:

- The purpose and character of the use;
- The nature of the copyrighted work;
- The amount and substantiality of the portion used; and
- The effect of the use upon the copyrighted work's potential market.

Fair use analyses are notoriously fact-intensive, and decisions often reflect the unique facts of a given case, which makes generalizations difficult. This area of the law is also likely to evolve rapidly over the coming months and years, and reliance on fair use can often be an impractical approach from a business perspective.

2.3.2 Other Copyright Violations – Copyright Management Information

AI developers and AI application developers may also be subject to a claim of altering copyright management information in violation of Section 1202 of the Copyright Act. "Copyright management information" (CMI) includes the title and other identifying information of a work, and the name and identifying information of a work's author or copyright owner.[2]

Where generated content is substantially similar to a copyrighted work, the copyright owner may demand that the generated content retain any CMI included with the works used to train the GenAI model and claim that removal of the CMI violates the Copyright Act. While such a claim requires an intent to induce, enable, facilitate, or conceal infringement, developers of GenAI models should consider monitoring this developing area of the law to ensure that their service does not facilitate infringement in any way. When using inputs that could contain copyrighted works to train GenAI models, developers may consider programming their GenAI model to look for any CMI attached to the copyrighted work, and to include required attributions to any copyright owners.

2.3.3 Breach of Contract and Open Source Software

In addition to infringement risks, AI service providers may also violate contractual limitations on their use of training data if those limitations were not properly followed in generating the output. For example, if developers of GenAI models use open source code repositories as part of the models' training data, then those models or specific output generated from the models may become subject to the open source license terms associated with those repositories. These terms may require the developers to provide attribution or copyright notices as part of or in connection with the output, which could be viewed as a reproduction or derivative work of the original open source code. These open source license terms may be in addition to other contractual obligations imposed by the third party offering the source code repositories.

Similarly, users of GenAI models could also be subject to a breach of contract claim if, for example, proper attribution was not given in connection with their use or redistribution of an output that was originally licensed under open source terms that require attribution.

Many GenAI models have been trained on open source software code, so companies or organizations that use those GenAI models to help produce software code run the risk of having "copyleft" code being included as output.[3] If copyleft code is incorporated into the companies' or organizations' own product or code base, it may expose the companies or organizations to copyleft risk in addition to any infringement or breach of contract risk. One way to address this risk is to use open source auditing tools to scan AI-generated output against known open source databases.

2.3.4 Reputational Risks

While GenAI models may excel at certain tasks, their output often contains errors, cites to inappropriate or nonexistent sources, or is otherwise inaccurate, misleading, biased, or false.

Because the quality of generated content varies and cannot reliably meet quality-control criteria, AI model providers, and AI application developers who build on foundational models, should include appropriate disclaimer language in their terms of use to alert customers that generated content should be reviewed by individuals before it can be shared or relied upon in a different way. Conversely, users of GenAI models should understand that high-quality results are not always the norm, and companies that permit their employees to use GenAI models for their work should establish an AI usage policy that sets forth guidelines for seeking approval and appropriately vetting output.

2.3.5 Regulatory Risks

Finally, the use of GenAI models in certain fields may also raise regulatory concerns. For example, the European Commission proposed a draft AI Act in April 2021, which is intended to apply to standalone AI systems and components in products. The Federal Trade Commission (FTC) also provided guidance on practices that the FTC may consider to be "unfair" or "deceptive," including, for example, not taking steps to mitigate risks of fraud associated with an AI tool.[4] California law requires disclosures when a chatbot is being used to incentivize a purchase or influence a vote. Detailed discussions on regulatory risks associated with GenAI models are beyond the scope of this article.

3. BEST PRACTICES FOR EVALUATING AND MITIGATING RISKS

At a Glance:

- "Foundational" GenAI models help to speed up the development of AI applications, but developers should perform due diligence to understand how the models were built and trained.
- Risk shifting mechanisms, including representations and warranties and indemnification provisions, may help allocate risk and mitigate potential exposure.
- Companies should also be mindful of regulatory risks and monitor developments in key markets.

3.1 Developing GenAI Models from Scratch

Companies determined to build their own GenAI models from scratch have the benefit of being able to tailor the GenAI model to the intended application or service by carefully selecting appropriate sources of training data. But GenAI models built from the ground up generally require vast amounts of training data, such as the datasets summarized in Section 2.1. While the selection of training datasets will likely be guided by the quality and technical applicability of the datasets, AI developers should also ensure that the datasets are either explicitly licensed or otherwise available for training purposes without violating any applicable terms and conditions or applicable laws.

Once trained, GenAI models cannot "unlearn" specific instances of training data if they later turn out to be problematic from a licensing perspective. GenAI models largely operate as a "black box" with training data finding its way into millions of model parameters, and most GenAI models cannot be "restored" to a state corresponding to all but the one instance of offending data. Excising offending data essentially means retraining the model from scratch on all of the remaining training data. But retraining would, in most cases, be time consuming and cost prohibitive. For this reason, companies should ensure that they have received adequate licenses to use training data (including, if relevant, input prompts) and, if in doubt, steer clear of data sources that may later turn out to be problematic.

While some commentators have suggested that AI developers maintain records of datasets that were used to train a given GenAI model, companies should balance the benefits and drawbacks of this proposal. On the one hand, contemporaneous records may be helpful

to determine whether a GenAI model was trained on certain data in the event that an infringement claim arose or regulators inquired about the source of training data. On the other hand, maintaining accurate records can be burdensome, and potential plaintiffs could obtain those records in discovery, and potentially seek to use them as putative evidence of breaches of contract, willful infringement, or other legal claims.

3.2 Building on Top of Foundational GenAI Models

Developing GenAI models from scratch is costly and requires significant computing power, storage capacity, and machine-learning expertise. To help smaller companies target specific use cases and applications without having to incur this substantial upfront investment, "foundational" GenAI models have emerged that permit companies to build on top of general-purpose GenAI models that have already been pretrained on large datasets and are intended to enable an array of different use cases with further customization. Foundational GenAI models come in two flavors: (a) proprietary models that can be in-licensed on commercial terms, and (b) open source models, which are licensed on open source terms.

While proprietary foundational models permit companies to extend the in-licensed GenAI model and provide documentation on how to do so, a complete picture of the inner workings of the model and the origin of all its training data remains confidential. Despite this lack of knowledge, companies that build on foundational models could still be liable for copyright infringement and other strict-liability claims. As a result, developers should not only evaluate foundational models from the perspective of their technical capabilities but also analyze the model from a legal perspective. Similar to the last section, this analysis will likely focus on the origin of training data and associated license provisions.

Based on the outcome of the diligence process, AI application developers building on foundational models should consider negotiating risk-shifting mechanisms that can help alleviate potential third-party claims:

- AI application developers could seek representations and warranties with respect to the development and training of the foundational model to ensure that the owner of the foundational model has obtained all necessary rights to build and use the foundational model and to license it to others;

- To mitigate liability for intellectual property infringement claims, AI application developers could seek indemnification from infringement claims resulting from their use of the foundational model. Application developers should be prepared, however, to provide similar infringement indemnities with respect to its customizations and any additional training of the foundational model; and

- AI application developers are also encouraged to consider similar contractual assurances for any reputational harms or regulatory actions taken against the application developers to the extent those claims are rooted in the foundational model.

3.3 Using GenAI Models for Day-to-Day Business Tasks

The rapid pace of AI development has prompted interest in using GenAI tools for a wide array of business tasks ranging from speeding up clerical tasks and providing outlines and checklists about topics of interest to generating marketing ideas and materials. While a company may choose to explicitly authorize the use of GenAI tools in some instances, employees frequently leverage existing GenAI tools without seeking permission.

Companies that intend to use GenAI tools in their business without developing their own GenAI tools should vet AI vendors in a similar manner as other suppliers or partners. First and foremost, because of the possible AI leakage summarized in Section 2.2, companies should ensure that they understand whether input prompts and data submitted to the GenAI tool will be incorporated into the model or kept separate from the GenAI model. The answer to this question will likely inform a company's usage policies about what types of data should be permitted to be submitted to the GenAI model. But even where the GenAI model segregates a company's data from the model and input prompts of other users, companies may wish to nevertheless prohibit the upload of any sensitive or privileged data to further minimize risk. As part of the vetting process, companies should also consider the provider's data security models and whether to monitor employee use of the GenAI model consistent with applicable law.

Even businesses that choose not to use GenAI tools for the time being should consider establishing an internal AI policy to ensure that employees are aware that GenAI tools may not be used.

3.4 European Union Regulatory Risks

In Europe, companies must consider the application of the General Data Protection Regulation (GDPR) when using GenAI tools. The use of personal data to

train, develop or deploy GenAI is subject to the GDPR. The GDPR sets forth principles such as lawfulness, transparency, purpose limitation, data minimization, and risk assessment. It also requires including certain data protection terms in contracts, and implementing certain internal policies, procedures, and documentation. Noncompliance with the GDPR can lead to fines of up to 20 million EUR or four percent of a company's global turnover, whichever is higher.

Regulators including the Spanish AEDP,[5] the French CNIL[6] and the UK ICO[7] have issued (draft) guidance on the use of AI. The ICO also released a blog[8] outlining specific questions that companies should consider when developing or using GenAI. In terms of regulatory enforcement, the Italian Garante has taken the lead with several enforcement actions against GenAI services. These enforcement actions highlight the complexity of aligning GenAI with GDPR obligations, e.g., legal basis, transparency, and individuals' rights.

Meanwhile, a draft AI Act is making its way through the EU legislative process. If adopted, it will ban certain AI systems because of their (assumed) unacceptable risks, and it will impose strict requirements (such as providing human oversight, training and testing, risk assessment transparency, security) for AI systems that are considered "high-risk."

4. INTERNAL AI USAGE POLICIES FOR EMPLOYEES

At a Glance:

- Developing an AI usage policy is an important step in evaluating the risks and benefits of using AI tools and providing guidance to internal stakeholders.
- Permitted AI usage will likely vary across use cases and involve different levels of supervision and approval.

As emphasized throughout this article, companies should proactively consider the benefits and risks of AI tools. Establishing an AI usage policy is an important first step in this process because doing so encourages an organization's management and internal stakeholders to evaluate the implications of use of AI and permits the communication of policies throughout the organization. Employees should understand that they generally cannot use GenAI tools for their work unless expressly permitted and properly vetted by the employer.

This section first addresses the potential leakage of business data that could result from the unauthorized uploading of sensitive business information and then turns to how a company could go about vetting output generated by the approved use of GenAI tools for content creation of images or writings as well as special considerations related to AI-assisted software development.

4.1 Avoiding Leakage of Business Data

As discussed in Section 2.2, if employees submit sensitive data of a confidential or privileged nature to the GenAI model, it could be incorporated into the GenAI model and eventually be disclosed to other parties downstream, waiving privilege or risking the trade secret status of confidential information.

Some GenAI models offer enterprise licenses whereby the GenAI model or application developer agrees that information uploaded by an organization's employees will remain accessible only to users of that organization. Companies that take advantage of these restricted deployments should nonetheless develop guidelines about the types of content that employees may upload, involve the legal department in this process, and consider including additional restrictions with respect to any content that may be subject to attorney-client privilege or the work product doctrine. An organization should also consider undertaking technical due diligence to ensure that proposed restrictions are both feasible and properly implemented.

Organizations that permit their employees to upload customer information to a GenAI model should take special care to ensure that the organization's customers consent to that use. Moreover, if companies incorporate a third-party's GenAI model into their platform, they should ensure that they understand and clearly communicate to their customers how the customers' data will be used.

To enforce the AI usage policy, companies should consider monitoring the use of popular websites that offer AI chatbot solutions and potentially block access to those websites. If feasible, companies could also consider monitoring and logging the upload of confidential business information to third-party websites or platforms.

4.2 Content Creation

If a company permits its employees to use GenAI tools for content creation, it should separately establish protocols for the responsible use of GenAI tools and ensure that generated content is reviewed before it is used either internally or externally. Companies are well-advised to consistently review and carefully vet the output of any GenAI tool to safeguard against "AI malfunction."

In addition to having a person review AI-generated content, companies should also consider documenting

how the output was generated, either by saving the history of input prompts submitted to the GenAI tool or by downloading log files if supported. This documentation can be helpful if a third-party claim arose with respect to the generated content. Even though a lack of intent may not avoid liability for certain types of claims, a company's documented efforts to engage in a responsible use of AI may help from a reputational standpoint, and also, avoid a claim of willful or reckless conduct.

4.3 Software Development

Human review is likely needed to mitigate many concerns resulting from inappropriate visual or textual content generated by GenAI tools because quality issues or harmful content will be apparent to the reviewer. AI-generated source code, however, deserves special attention because the functional nature of source code may lead to additional risks that are not immediately apparent. For example, AI-generated source code may contain bugs or security vulnerabilities that could affect object code compiled from the source code. A malicious actor who deliberately "leaks" code with bugs or security vulnerabilities into a GenAI tool's training set could thus infect downstream source code unbeknownst to the GenAI tool's owner or the developers using the GenAI tool.

A detailed review process and functional testing of AI-generated source code may help mitigate some of the risks, but companies should consider to what extent their software developers should be permitted to have source code generated for them. For example, for short snippets of code or short segments with well-defined functions, the generated code may be more easily reviewed than if developers were permitted to have large swaths of source code generated automatically.

Companies should further consider tracking AI-generated source code segments in their overall source code repository and version control system. Without a good record keeping system, it may be difficult to later identify the location of the AI-generated source code if a party claimed infringement. By tracking the location of the code, it will be possible to remove or rewrite the implicated segments without affecting other portions of the code that were written by human software developers.

4.4 Review and Approval Chains

Across the usage scenarios outlined above, companies should implement review and approval chains to make sure that each use case is properly vetted and subject to appropriate oversight. The approval processes will likely differ across different use cases. Shorter and informal approval may be appropriate where sensitive company data is not implicated, and where any generated output is purely used within the company. For example, enlisting the help of a GenAI tool to prepare checklists or high-level outlines that are later reviewed and revised by employees poses comparable low risk. On the other end of the spectrum, where generated output is used in marketing materials that are widely distributed or where AI-generated source code is implemented into company products, the approval chain should likely consist of several steps and, in addition to the initial approval at the outset of a project, include continuous oversight by an organization's senior management and legal department.

Given the rapid pace of AI development, companies should regularly review their AI usage policy and tailor it to the current use cases and risks.

5. CONCLUSION

AI's technical and legal landscapes are rapidly evolving. We encourage readers who wish to learn more about best practices for their own use of AI technology to reach out to legal counsel.

Notes

1. The scope of permissible web scraping is a developing area of the law informed by case law surrounding the Computer Fraud and Abuse Act, the implications of technological restrictions, and unfair competition law.
2. More specifically, Section 1202(c) defines copyright management information (CMI) as certain identifying information conveyed in connection with copies or phonorecords of a work or performances or displays of a work, including in digital form, except for any personally identifying information about the user of a work, or a copy, phonorecord, performance, or display of a work.
3. Open source licenses generally require attribution and certain notices when licensees distribute copies of the licensed content. Broadly speaking, "copyleft" licenses further require that distributions be licensed under the same "copyleft" terms, including terms that require the licensee to make source materials available to recipients.
4. Fed. Trade Comm'n, Chatbots, Deepfakes, and Voice Clones: AI Deception for Sale (Mar. 20, 2023), https://www.ftc.gov/business-guidance/blog/2023/03/chatbots-deepfakes-voice-clones-ai-deception-sale; Fed. Trade Comm'n, Keep Your AI Claims Check (Feb. 27, 2023), https://www.ftc.gov/business-guidance/blog/2023/02/keep-your-ai-claims-check.
5. https://www.aepd.es/es/documento/requisitos-auditorias-tratamientos-incluyan-ia-en.pdf.
6. https://www.cnil.fr/fr/intelligence-artificielle/ia-comment-etre-en-conformite-avec-le-rgpd.
7. https://ico.org.uk/media/2617219/guidance-on-the-ai-auditing-framework-draft-for-consultation.pdf.
8. https://ico.org.uk/about-the-ico/media-centre/news-and-blogs/2023/04/generative-ai-eight-questions-that-developers-and-users-need-to-ask/.