CLIFFORD CHANCE

**JAPANESE LAW ISSUES SURROUNDING
GENERATIVE AI: CHATGPT, BARD AND BEYOND**

JAPANESE LAW ISSUES SURROUNDING GENERATIVE AI: CHATGPT, BARD AND BEYOND

Organisations and governments around the world are attempting to harness the power of generative AI and navigate a path for safe and secure use. In Japan, those exploring the opportunities presented by generative AI will need to consider complex legal issues, including in relation to copyright, data privacy and liability.

Copyright issues can arise at both the training and use stages of these AI systems. While using copyrighted materials for training generative AI models is generally allowed under certain conditions, potential conflicts with copyright holders need careful consideration. The ability to copyright AI-based works depends on creative input by a human. **Data privacy** concerns arise due to the collection and processing of vast amounts of data, where this includes personal information. Adherence to Japan's Act on the Protection of Personal Information (APPI) is crucial. Organisations must ensure that personal data used in generative AI aligns with the stated purpose and be cautious when third-party operators might use data differently. **Liability** for harms arising in relation to generative AI outputs remains complex in Japanese law. Risk management strategies, including well-drafted terms of service and disclaimers, can mitigate potential legal exposure.

While Japan has not enacted specific AI legislation, the regulatory landscape may evolve to address emerging AI technologies. Staying informed about legal developments and crafting effective risk management strategies is vital to ensure compliance and to minimise legal risks.

This briefing provides an overview of some of the key issues organisations should consider under Japanese law when using or developing generative AI.

Copyright and other IP rights issues

Generative AI models such as ChatGPT and Bard raise complex IP challenges within the Japanese legal framework. Japan is taking a similar approach to countries such as Singapore and the UK by applying existing legislation in order to deal with issues arising from the use of generative AI at this time. Therefore, copyright infringement in the context of generative AI (both at the training stage, and the 'use' or output stage) is covered by the existing provisions of the Copyright Act of Japan.

Training stage

The use of copyrighted products or materials to train generative AI models would be *prima facie* copyright infringement under the Copyright Act, as it is a reproduction (*fukusei*) or other form of use of the copyrighted work. However, Article 30-4 of the Copyright Act stipulates that the use of copyrighted works by generative AI for learning purposes is allowed in principle.

This provision sets out that "information analysis" (defined as the act of extracting language, sound and images from a large set of information, and comparing, categorising and carrying out other analyses of such information) is allowed unless such use of copyrighted works unreasonably prejudices the interests of the copyright owner, in light of the nature or purpose of the work or the circumstances of its exploitation in Japan. The wording of the above caveat in the Copyright Act relating to potential harm to the copyright holder is somewhat vague, but the Agency for Cultural Affairs (a special body of the Japanese Ministry of Education, Culture, Sports, Science and Technology) has offered some clarity in its guidelines. For example, where a database containing a large amount of information used for information analysis is for sale, the reproduction of the database for the same purpose of information analysis would not be allowed under Article 30-4 because such act interferes with the market relating to the sale of the database.

Thanks to Article 30-4, the range of possible uses of copyrighted materials in training AI models in Japan is broad. When this information analysis exemption was enacted in 2018, training generative AI models with the level of sophisticated generative capability of ChatGPT and Bard was not envisaged; however, recent discussions by governmental organisations have suggested that this provision does indeed apply to the training of generative AI models. In addition, it is important to note that Article 30-4 would apply to the training of commercialised generative AI models, as well as those used for non-commercial purposes.

There are also ongoing discussions as to whether a commercial agreement or term of use to override the provision above is valid (i.e., whether the copyright owner can opt out of the information analysis exception contractually). However, a working group set up by the Ministry of Economy, Trade and Industry (METI) stated in its report that such opt-out provision should not be valid, as it would impede the innovation which AI may bring in the future, although this should be determined on a case-by-case basis.

In any event, business operators building generative AI models should be mindful of copyright holders' rights, and it is not advisable to train the model based only on a narrow set of works – for example, specific types of works or specific authors' works. This is because a small set of training input increases the risk that the output generated by such generative AI models is similar to the pre-existing copyrighted works used for training and, accordingly, there is an increased risk that such output would infringe such pre-existing copyrighted works.

Use stage

The same rules apply to copyright infringement by the production of materials using AI as those for infringement by 'ordinary' work produced without using AI. This includes the uploading and publication of images generated using generative AI models as well as the sale of any copies.

Copyright infringement would be found where the work has been created using a pre-existing copyrighted work (reliance) and the expression of ideas in both the new work and the copyrighted work are similar (similarity).

If the pre-existing copyrighted work has not been used (as a matter of fact) for the training of the generative AI model or put in the AI model as a prompt, there is no "reliance" and there would be no finding of copyright infringement. As it would not be realistic for the copyright holder to identify each piece of material used for the training, accessibility (e.g., whether the copyrighted work was made public) would be taken into consideration in determining whether there is "reliance". In preparation for an effective defence in this regard, it would be advisable for the organisation that trained the AI system to keep records of training data.

"Similarity" would be found if the common traits between the AI-based work and the pre-existing work are an essential part of the pre-existing copyrighted work.

Therefore, it is important for business operators that create content with generative AI models and use such content to address the following practical points:

1. Set out internal rules regarding the content that can be created and used

AI-generated output: The risk of using content which is generally able to be copyrighted – such as images, animations and music – is higher than using other content such as simple texts or draft emails. Categorising the permitted ways of using AI-generated content by 'type' in permitted usage policies (or similar governance mechanisms) is a useful way to mitigate the risk of copyright infringement, taking into account the potential risks each type of content bears (for example, AI-based images may not be used for external purposes).

Inputs into AI systems: In addition, careful attention should be paid to how prompts are entered into generative AI models. Entering a copyrighted work as a prompt would technically infringe its copyright as a form of "reproduction" *(fukusei)*. Therefore, business operators should consider setting out rules and guidelines for entering prompts to mitigate this risk of inadvertent infringement. Such rules can also include a prohibition on inputting trade secrets (technical or operational non-public information valuable for a business which is kept confidential) or other confidential information. Under the Act Against Unjustifiable Premiums and Misleading Representations, the holder of the trade secret may claim for injunctive relief and/or enjoy a reduction in the burden of proof if such trade secrets are stolen or otherwise used without authorisation. However, these protections might not be available if an employee discloses a trade secret to a third-party generative AI system that may share input data with a third party, as it may no longer be considered confidential.

2. Check resemblance with existing works

Organisations should check whether there are any similar existing copyrighted works to the content being created by their generative AI models and, if so, analyse their resemblance. Although doing this can be challenging in practice, it is expected that technologies that facilitate such checks will continue to evolve in their capabilities and availability.

Ability to copyright AI-based Products

In Japan, copyright subsists in works that creatively express ideas or emotions and, as mentioned above, there are no AI-specific regulations setting out rules that differ from this principle.

In 2017, a governmental study team indicated what kind of AI-based works can be protected by copyright, stating:

In the process of producing AI-generated works, if the user of generative AI models has a creative intention and, at the same time, a creative contribution to the AI-generated product, the AI-generated product is considered to have been created by the user using AI as a "tool" to creatively express his/her thoughts and feelings, and thus the AI-generated product is deemed to be copyrightable.

However, this indication emerged before the generative AI which we know today.

Therefore, discussions are currently under way to clarify the concept of the user's creative contribution for a product to be recognised as a work of authorship.

To date, such discussions have not resulted in any change to the law and, therefore, under the existing legal framework there is room for argument around what level of human contribution would make a work able to be copyrighted. Nevertheless, the existing criteria remain a good point of reference for the time being. According to these, if the work is generated only using generative AI models, it would not be able to be copyrighted. However, if a human carries out some kind of processing or contribution, there is room for such work to be copyrighted.

Data privacy

The development and use of generative AI involves the collection and processing of a huge amount of data, including – potentially – personal data.

To comply with Japanese laws, a business organisation must ensure adherence to the relevant privacy and data protection regulations as per the Act on the Protection of Personal Information (APPI), which serves as a primary framework for data protection in Japan.

Purpose of use

Under the APPI, in general, the collection and use of personal data does not require the consent of the data subject so long as the purpose of such and use is notified or announced (i.e., made available to the data subject; for example, by way of announcement or publication on a website) to the data subject. Where business organisations that are data controllers use generative AI, any inputting of personal data to the generative AI should fall within the purpose that has been announced or notified to the data subjects (although this does not need to refer to the specific generative AI system being used). Data controllers must notify the data subjects in advance of such use of a generative AI system, and such notice should clarify whether such data will be used for learning by the AI model.

Patents and AI
A discussion similar to that on the ability to copyright is under way regarding the ability to patent AI-based works. The degree of involvement of persons required as a criterion for AI-based inventions to be patentable is also under consideration.

In particular, where a business organisation uses generative AI owned and/or operated by a third party, there is a risk that the third-party operator may impose terms of use that allow it to use the personal data for different or wider purposes than the purpose originally notified to the data subjects. If this is the case, by using the generative AI, the business organisation may be inadvertently contravening the APPI. A good example is where the generative AI's terms of use refer to using the information inputted by the users to provide, maintain, develop and improve its service (which may not have been included in the purpose of use notified by the business organisation to the data subject). Therefore, care should be taken when entering personal information into generative AI.

Provision of personal data

A provision of personal data to a third-party generally requires the consent of the data subject, except in certain exempted circumstances. One such exemption is where an organisation transfers personal data to third party service providers in an outsourcing (*itaku*) context where the scope of transferred personal data is limited to that necessary to implement the purpose of use the organisation has notified or announced to the data subject and the organisation appropriately supervises the third-party service transferee (Outsourcing Exemption). In order for the Outsourcing Exemption to apply, there would need to be an agreement in place between the transferor organisation and the transferee third-party vendor (for example, the operator of the generative AI). If no specific contract is entered into (for example, because the business organisation is using a free and/or trial version of the generative AI system which is open to the public) data subjects could argue that such personal data is not transferred by virtue of an outsourcing (because the Japanese word in the APPI - *itaku* - implies the existence of an outsourcing service agreement). However, whether the service is provided with or without consideration (for example, a service fee) is not a decisive factor for classifying if the transfer is made in the course of outsourcing. These restrictions on third-party transfers apply regardless of whether the input data is used for learning.

Cross-border transfers of personal data also require the consent of the data subject (except in limited exempted cases such as where the data is transferred to a jurisdiction which the Personal Data Protection Commission (PPC) deems as having an adequate level of data protection as in Japan) and the data controller must also implement adequate security measures to protect the personal data being transferred. As a result, if the relevant data server for the generative AI system is located outside Japan, the UK or EU member states (jurisdictions which the PPC deems as having an adequate level of data protection), even if a transfer is made relying on the Outsourcing Exemption, consent of the data subject would be required, unless an adequate level of data protection is ensured as per the requirements of the APPI.

Right to correct / delete

Under the APPI, data subjects have the right to require the data controller to correct, add or delete the personal data held by the data controller if such data is incorrect, and the right to request cessation of use of the data subject's personal data if such data is used for purposes other than the purposes notified to the data subject, or used in a way that encourages or induces illegal or unjust acts. In practice, however, careful analysis is needed to establish who the relevant data controller(s) is (or are) in any given factual scenario involving the use of generative AI, and it is likely to be challenging for a

Would provision of personal data to a generative AI always be considered a transfer to a third party?
The governmental guidelines provide that it is considered to be a transfer only when the personal data is made available for a third-party's use. Therefore, if the terms of use clearly state the provider of the generative AI model would not make use of the input data (whether for training purposes or not), such provision of personal data may not be considered as a transfer for the purpose of the APPI.

data controller to make such correction or deletion in a meaningful manner once generative AI has been trained with incorrect personal data or in breach of the APPI. The guidelines published by the PPC do not yet address such difficulty, and it is therefore uncertain whether the PPC would take an extreme position that use of the entire generative AI model should cease on the basis that the 'learned' personal data cannot be segregated from the database.

Regulator-endorsed legality of careful use of personal data

The challenge is in ensuring adequate protection of personal data whilst not letting this be a block on the improvement of the accuracy and/or efficacy of the generative AI. Restricting the use of personal data in the interests of compliance with existing data protection regulations may reduce the available datasets required for the generative AI to improve its accuracy and/or efficacy.

The PPC recently published a call for attention in respect of the use of generative AI.[1] This constituted two parts: (i) suggestions for commercial organisations, governmental organisations and individuals, who enjoy the services provided through generative AIs generally, and (ii) recommendations addressed specifically to OpenAI.

The suggestions for commercial organisations referred to in (i) above were to remind businesses that:

- if prompts include personal data, only the personal data necessary to achieve the purpose of use should be inputted to the generative AI; and

- if personal data is used for a purpose other than providing outputs (e.g., used for the learning purposes), such use of personal data may constitute a breach of the APPI by the data controller. As such, compliance with the APPI should be carefully checked by the data controller before it inputs the personal data into a generative AI.

It is interesting that PPC included a set of recommendations specifically addressed to OpenAI. These are focused on the use of sensitive personal data (as well as the PPC's instruction to give notice to, or make an announcement to, the data subjects setting out the purpose of use in Japanese language). As per the APPI, collecting and processing of sensitive personal data requires the consent of the data subject. In this regard, the PPC required OpenAI to take necessary measures such that sensitive information is not collected in the first place and/or if such sensitive data is obtained, to take measures to remove the sensitive information as far as possible from the dataset immediately after collection, and remove or anonymise the sensitive personal data before converting the collected data into a dataset for learning. The PPC also instructed that, when a data subject or the PPC requests OpenAI not to collect sensitive information from a specific website or a third party, then OpenAI should adhere to that request, and OpenAI must not use sensitive information input as a prompt if the user opts out from such information being used for learning, unless there is a justifiable reason.

---

1 https://www.ppc.go.jp/news/careful_information/230602_AI_utilize_alert/ (Japanese only)

Although the call for attention was intended to alert the relevant parties to the potential issues that may arise out of generative AI, it can be viewed as a positive message from the PPC in that there was no blanket prohibition or recommendation not to use generative AI given the risks involved.

Liability and Accountability

Determining liability for any harm or damage arising in connection with the output of generative AI poses significant legal challenges under Japanese law. As AI models such as ChatGPT and Bard generate content without direct human control, questions arise regarding responsibility for harm caused by their output. Japanese courts and regulators have yet to establish clear standards for attributing liability in these situations. Entities deploying these AI technologies should consider implementing robust risk management strategies, including well-drafted terms of service, disclaimers and user guidelines, to mitigate potential legal exposure.

Under the Civil Code of Japan, damages may be claimed against a person on the ground that there is a breach of contract or an act constitutes a tort. However, there are certain limitations to these claims.

First, a person would be held liable only if they cause the breach of contract or tort by wilful misconduct or negligence.

In addition, a person is only liable for damages which (i) would ordinarily arise from the breach of contract or tort, or (ii) arose under special circumstances that such person expected or should have expected.

It should be noted that even if a person is found liable, Japanese courts are unlikely to award exceptionally large sums as damages, and punitive damages are not awarded in Japan.

The legal regime in Japan would therefore limit an AI-provider organisation's liability to some extent. By way of example, a service provider may consider offering a chatbot system to a client organisation, which is a troubleshooting service powered by generative AI for the client's internal use. There is a risk that the chatbot system may produce convincing false statements or 'hallucinations'. In such case, the client may attempt to claim damages from the service provider due to a breach of contract or tort. The Act against Unjustifiable Premiums and Misleading Representations also prohibits a business operator from making a misleading advertisement in a way that makes consumers believe that the relevant products are better than the actual quality or similar products provided by others. Hallucinations or false results may lead to such an illegal advertisement.

**CLIFFORD
CHANCE**

The service provider in this scenario could avoid being found negligent or producing an illegal advertisement by having taken appropriate action so as not to cause damages. Any case for breach of contract will depend on the applicable terms in place.

Given that there have been extensive discussions providing warnings about the possibility that generative AI-based products may produce hallucinations, it is likely that the Japanese courts would find that the service provider should have expected that the chatbot may produce them. To mitigate the risks around hallucinations, the service provider should implement preventive measures such as regular and proper reviews of the program. Another preventive measure that could be taken is for the chatbot to show the source of information together with its statement, so that the client can verify the validity of the responses.

Conversely, organisations procuring generative AI-based services as a customer should carefully review any contractual arrangement or terms of use applicable to the content, in order to assess the risk and potential liability in the event something goes wrong.

The service provider may limit its contractual liability by introducing an exemption clause in a contract, which is considered to be valid under Japanese law as long as it is not against the public order. When entering into a contract with a client, such limitation of liability clause can be included in an individual agreement. Alternatively, when a service provider intends to offer its services under the same terms to numerous users, the service provider may prepare terms of service which include such limitation clause. It should be noted that provisions in the terms of service which (i) restrict the user's rights, or expand the user's obligations and (ii) are found to unilaterally prejudice the interests of users in violation of the public order, shall be unenforceable.

For example, the service provider may consider including a provision which exempts it from liability for damages arising from the production of the Generative AI unless the damages are caused by the service provider's wilful misconduct or gross negligence.

Thus, the service provider may limit risks around the use of generative AI by taking appropriate actions to avoid damages or establishing carefully considered terms of service. Please note, however, that if users are consumers, such terms of service should be carefully drafted. The Consumer Protection Act was recently amended, and a contractual limitation of liability is not valid unless it is expressly stated that gross negligence is carved out of the liability limitation. Otherwise, the entire clause relating to the limitation of liability could be considered null and void.

Regulatory Landscape generally in Japan and future outlook

The rapid evolution of AI has prompted regulatory bodies in Japan to consider the need for appropriate oversight and regulation. In 2021, the "Working Group regarding How AI Principles Should Be Implemented" led by METI confirmed that the regulations should be through soft law rather than hard law in response to the burgeoning technologies. As a result, unlike the EU, which is now in the process of enacting legislation specific to AI issues, Japan has not implemented any AI-specific legislation but instead has, like many other countries such as Singapore and the UK, decided to rely on existing laws to cover generative AI technologies at this time.

Nevertheless, a comprehensive AI regulation by statute may be established in the future. Further, the AI Strategic Committee published "Preliminary Screening of Issues regarding AI" on 26 May 2023. The committee pointed out that compliance with existing laws and guidelines (through risk assessment and governance) should be encouraged. However, where it is impossible to solve the issues under existing legislation, then the government and relevant stakeholders should consider the countermeasures, taking into account how the issues are tackled globally.

From a global perspective, Japan chaired the G7 Hiroshima Summit in May 2023, where the "Hiroshima AI Process" was established. Guidelines on the development and use of generative AI are being discussed and a draft code of conduct for business organisations was presented in early August 2023. As part of the Hiroshima AI Process, the G7 Digital & Tech Ministers' Statement was issued on 7 September 2023. This included a commitment to develop, through the Hiroshima AI Process, guiding principles for organisations developing and using advanced AI systems, in particular foundation models and generative AI, and declared what principles will be included.[2] The finalised set of principles is expected to be reported by the end of 2023.

We encourage business organisations to monitor these developments under Japanese law regularly to ensure compliance and properly mitigate legal risks associated with generative AI technologies.

The information provided herein is not exhaustive and should not be considered legal advice tailored to your specific circumstances. We recommend consulting with our legal team to address your unique concerns.

---

2 Details can be found at https://www.soumu.go.jp/main_content/000900470.pdf

AUTHORS

Michihiro Nishi
Partner
T: +81 3 6632 6622
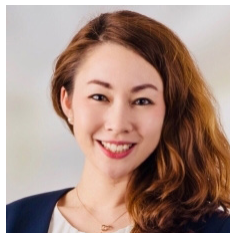E: michihiro.nishi@
    cliffordchance.com

Shunsuke Nagae
Senior Associate
T: +81 3 6632 6321
E: shunsuke.nagae@
    cliffordchance.com

Yuki Hoshinaga
Senior Associate
T: +81 3 6632 6648
E: yuki.hoshinaga@
    cliffordchance.com

Sian Smith
Senior Associate
T: +81 3 6632 6320
E: sian.smith@
    cliffordchance.com

Stephanie Ohira
Associate
T: +81 3 6632 6408
E: stephanie.ohira@
    cliffordchance.com

Tatsuya Nakano
Associate
T: +81 3 6632 6383
E: tatsuya.nakano@
    cliffordchance.com

Masaya Obayashi
Associate
T: +81 3 6632 6370
E: masaya.obayashi@
    cliffordchance.com

CLIFFORD CHANCE

CLIFFORD CHANCE