Artificial Intelligence (AI) Privacy Considerations: US State Laws, HIPAA, EU/UK GDPR

*Contributed by **David Peloquin** & **Rohan Massey**, Ropes & Gray*

Editor's Note: This checklist provides a list of privacy-related legal considerations for practitioners to ask when evaluating artificial intelligence (AI) products systems (collectively, AI Services), whether developed internally within an organization or by a third-party vendor who makes the AI Service available to its customers. This checklist is focused exclusively on US and EU/UK privacy- and data protection-related matters, including some that are particularly relevant to the health care industry. However, many other industry and legal considerations could apply that are governed by other bodies of law. For example, employment issues, the regulation of software as a medical device (SaMD) under US Food and Drug Administration authorities or EU/UK legislation, intellectual property laws, and restrictions related to avoiding bias. This is a rapidly evolving area and new laws and regulatory requirements are emerging in the US and across Europe that companies should take into account when assessing and understanding their AI obligations. To keep track of the latest developments related to AI across jurisdictions, visit **In Focus: Artificial Intelligence (AI)**.

The left-hand column provides a practical checklist of privacy- and data protection-related considerations to use when evaluating an AI Service, whereas the right-hand column provides commentary on the applicable US and EU/UK legal privacy and data protection framework. In the US, compliance with the **Health Insurance Portability and Accountability Act of 1996** (HIPAA) may be relevant if the AI Service is developed or used by a HIPAA “**covered entity**” or “**business associate**,” or involves the use of protected health information (PHI). To the extent that applicable information is not PHI but could nonetheless be linked to an identifiable individual, the **California Consumer Privacy Act** of 2018 (CCPA), as amended by the **California Privacy Rights Act** of 2020 (CPRA), and other similar **comprehensive state privacy laws** may apply. States have similar yet nuanced definitions for such information, which this checklist collectively refers to as “Personal Information”. Uses of certain “sensitive” categories of information, such as genetic information or HIV/AIDS test results, may be subject to heightened protection under state laws. Those heightened protections are not discussed here but should be taken into account when evaluating AI. In the EU/UK, the **EU General Data Protection Regulation** (GDPR) and the **UK GDPR** (as defined in section 3(10) of the Data Protection Act 2018), which govern the processing of “**personal data**” and “special category personal data”, may also apply. For more details on the requirements of GDPR and state privacy laws, see **Comparison Table - GDPR vs. State Comprehensive Consumer Privacy Laws**.

Privacy Considerations for Legal Evaluation of AI Systems / Products	
Considerations	**Applicable Legal Framework**
1. Data Inputs; Privacy & Cybersecurity	
1.1. What end user inputs that are PHI, Personal Information (US), personal data or special category personal data (EU/UK) (together, Input Data) are entered into the AI Service?	US: It is important to understand what data must be put into the AI Service to identify whether any inputs contain PHI or Personal Information, and if so, to assure compliance with HIPAA, CCPA, and other state data privacy laws. CCPA and other state privacy laws also require providing notice to individuals as to how their Personal Information is used. EU/UK: Understanding whether the Input Data contains personal data that is being put into the AI Service, as well as more generally considering how it has been collected, used, stored, and shared, in addition to how this has been communicated to individuals, is essential to ensuring compliance with the GDPR/UK GDPR. An information notice provided under **Article 13/14** of the GDPR/UK GDPR (commonly referred to as a Privacy Notice) must be provided to individuals in order to inform them about the categories of personal data being used, the purposes of the processing and their associated lawful bases, the rights that individuals have over their personal data, how long the data will be stored, and how it will be shared. There are also broader GDPR/UK GDPR principles to consider, such as demonstrating accountability, collecting data lawfully, fairly, and transparently, for specified, explicit, and legitimate purposes, and processing in a manner that ensures security of data.
1.2. What are the sources of the Input Data? How is notice provided to individuals regarding the use and disclosure of their Input Data, and are there any exemptions to providing such information?	US: CCPA and other state privacy laws may require providing notice to individuals if their Personal Information is used as an input. Understanding the source of the data also helps indicate which laws may apply. EU/UK: Where Input Data is collected directly or indirectly from individuals, the controller of personal data (i.e., the organization that decides why and how data will be processed) is required to provide a Privacy Notice. Where the personal data is collected directly, the Privacy Notice should be given to individuals before or at the time of collection, but where the data is obtained from another party or source, the Privacy Notice should be given at the latest within one month of obtaining the data.

1.3. Will the AI Service use Input Data for any automated decision-making?	US: State privacy laws increasingly govern the use of automated decision-making in specific contexts. For example, CCPA draft regulations addressing automated decision-making remain in process, but the statute would consider any algorithm that uses computation to make or execute a decision and that processes Personal Information, including through "**profiling**," to be an "automated decision". Under most state privacy laws, in certain use cases (e.g., related to finance, banking, insurance, housing, etc.), consumers must be informed of this process and given the opportunity to opt-out, and businesses must conduct data protection assessments. Another example is a **regulation** that took effect in New York City in July 2023 that regulates the use of "automated employment decision tools" in the employment context. EU/UK: Under **Article 13(2)(f)** of the GDPR/UK GDPR, the controller of personal data must inform individuals about the use of any automated decision-making, including **profiling**, and provide meaningful information about the logic involved in it, as well as the significance and the potential consequences of such processing for the individual. Under **Article 22** of the GDPR/UK GDPR, individuals also have the right not to be subject to decisions based solely on automated processing where it produces legal or similar effects on individuals (e.g., relating to financial decisions or employment opportunities). There are exceptions to this prohibition, such as when individuals have explicitly consented to the processing. In these cases, the controller must implement suitable measures to safeguard the individual's rights, freedoms, and legitimate interests, and at least provide the right to obtain human intervention. Further, **Article 35(3)(a)** of the GDPR/UK GDPR requires a data protection impact assessment (DPIA) to be completed where systematic and extensive evaluation of individuals' personal aspects is based on automated processing and which produces legal or similar effects on those individuals. Regulator **guidance** issued by the UK Information Commissioner's Office (ICO) on automated decision-making suggests that the DPIA should specifically include: (i) a description of the processing activities, including data flows and the stages when AI processes and automated decisions may produce effects on individuals; (ii) information on margins of error in the performance of the AI Service that may affect fairness of personal data processing; (iii) a thorough description of the scope and context of the processing, including what data is being processed, the number of individuals, the source of the data and the extent to which individuals are likely to expect the processing; and (iv) the degree of human involvement in the decision-making process and at what stage humans are involved.
1.4. Is Input Data limited to the information that is truly needed to train and/or operate the AI Service?	US: HIPAA requires covered entities and business associates to make reasonable efforts to limit use, disclosure of, and requests for protected health information to the minimum necessary to accomplish the intended purpose. Several US state privacy laws take a similar approach, following a principle of "data minimization" akin to that found in the GDPR/UK GDPR. EU/UK: Under **Article 5(1)(c)** of the GDPR/UK GDPR, the principle of data minimization requires personal data to be limited to what is adequate, relevant, and necessary for the purpose of the AI Service. ICO **guidance** on training AI acknowledges there are competing interests, such as: (i) training a sufficiently accurate AI Service at the same time as reducing the quantity of personal data processed to train that system; or (ii) producing an AI Service that is sufficiently statistically accurate and which avoids discrimination. The balance depends on the specific sector and is a matter of judgment; however, the entity making this assessment must justify any decision reached. Examples of how to weigh this balance include: (i) assessing current or potential trade-offs when designing or procuring an AI Service, and considering the impact that it may have on individuals versus what is proportionately needed; (ii) implementing a clear criteria and lines of accountability about the final trade-off decisions, including a robust, risk-based and independent approval process; and (iii) reviewing trade-offs that are decided on a regular basis, including considering the trade-offs from the viewpoint of individuals whose personal data is likely to be processed by the AI Service. These processes should be documented as part of a controller's responsibility under **Article 24** of the GDPR/UK GDPR and the accountability principle at **Article 5(2)** of the GDPR/UK GDPR, to demonstrate that processing is fair, necessary, proportionate, and limited. This can be done as part of

	a DPIA or, where legitimate interests are being relied on as a lawful basis, by conducting a Legitimate Interests Assessment.
1.5. How long is Input Data stored and/or retained?	US: Many state data privacy laws require notifying individuals from whom Personal Information is collected of the anticipated retention time of that information. CCPA regulations require that the collection, use, retention, and/or sharing of a consumer's Personal Information to achieve that purpose be reasonably necessary and proportionate. EU/UK: **Article 5(1)(e)** of the GDPR/UK GDPR requires that Input Data is not kept in a form that can identify individuals for longer than is necessary for the purposes for which it is being processed. As described above, the controller must in its Privacy Notice inform individuals about the period for which their personal data will be stored, or if that is not possible, the criteria used to determine that period.
1.6. Where is Input Data stored? Are secure environments used to protect Input Data?	US: HIPAA requires compliance with security standards if Input Data involves PHI. If Input Data involves other Personal Information, state privacy laws require the implementation of reasonable data security practices. EU/UK: **Article 5(1)(f)** of the GDPR/UK GDPR requires that the Input Data is processed in a way that ensures appropriate security of the personal data. This includes putting in place appropriate technical and organizational measures to protect data against unauthorized or unlawful processing, accidental loss, deletion, or damage. Further information on the measures that can be used to ensure security appropriate to the level of risk is set out in **Article 32** of the GDPR/UK GDPR. ICO **guidance** recommends that technical teams record and document how personal data is moved and stored in order to appropriately apply security risk controls and monitor their effectiveness. Having this clear audit trail further demonstrates attention to the accountability principle of **Article 5(2)** of the GDPR/UK GDPR.
1.7. How is Input Data protected? Has a privacy and cybersecurity risk assessment been performed on the AI Service?	US: HIPAA and CCPA require performance of regular security/cybersecurity audits. EU/UK: See also the answer to Section 1.6 for the security requirements under the GDPR/UK GDPR and the specific considerations around security in the context of data storage and transit. In addition, security measures to be adopted will depend on the level and type of risks that arise from the relevant processing activities of the AI Service; there is no one-size-fits-all, and these measures must therefore be considered on a case-by-case basis. Information on the measures to ensure security appropriate to the level of risk are set out in **Article 32** of the GDPR/UK GDPR, and include pseudonymization, encryption and a process for testing, assessing, and evaluating the effectiveness of measures to ensure security of personal data. As AI Services are potentially more advanced than “traditional” IT systems due to the complexities and the way the technology may be built and deployed, security should be actively monitored and state-of-the-art security practices should be considered. A DPIA can be used to document any technical safeguards put in place to reduce risks to security and accuracy of personal data processed in the AI Service.
1.8. Is Input Data used to train the model associated with the AI Service?	US: If the AI Service is used by a HIPAA covered entity or business associate, there must be a legal basis under HIPAA for use of any PHI in connection with the AI Service (e.g., use of the AI Service in connection with treatment, payment, or health care operations). If for a permissible use, the covered entity should ensure that such use is addressed in its Notice of Privacy Practices. See also the answer to Section 1.4 for considerations around the principle of minimum necessary use. Further, CCPA and other state privacy laws require that consumers be informed of the purpose for which their personal information is collected and categories of third parties to whom their personal information is disclosed, among other things. Use of Input Data to train an AI model may therefore need to be disclosed in the privacy notice required to be provided to consumers pursuant to these state laws. EU/UK: See also the answer to Section 1.4 for considerations around the principle of data minimization and trade-off considerations to ensure that processing is fair, necessary, proportionate,

	and limited. Where Input Data is being used to train the AI model, this must be disclosed as part of the Privacy Notice to individuals. In addition, the lawful basis being relied on to train the AI Service must be considered. The ICO **guidance** on AI and legal bases can help controllers to determine which basis or bases might be most appropriate. It is worth considering whether the different stages, such as the AI Service development and the AI Service deployment, may be using separate lawful bases, as there are different circumstances and risks to be considered throughout the lifecycle of the product or service.
1.9. Will the Input Data be monetized? If so, how?	US: The CCPA and other state privacy laws require providing notice to the individuals from whom Personal Information is collected of the categories of third parties to whom Personal Information will be sold or shared and the type of information that will be sold or shared. HIPAA restricts the sale of PHI without the individual's prior written authorization. EU/UK: Where Input Data is being used in a monetized product, the recipients (or categories of recipients, which could identify potential customers) receiving such Input Data must be disclosed as part of the Privacy Notice. Where personal data is being shared (e.g., an entity is selling training data sets), de-identification techniques may need to be applied, or privacy enhancing technologies may need to be introduced. If personal data in a training set will be transferred to or otherwise processed by a third party, these activities also need to be treated as data sharing and suitable contractual measures under **Article 28** (sharing between controllers and processors), **Article 26** (sharing between joint controllers) and (where applicable) **Chapter 5** (transfers of personal data to third countries or international organizations) of the GDPR/UK GDPR should be considered.
1.10. Will the Input Data be licensed, shared with a third party, or used for the benefit of a third party? If so, how?	US: CCPA and other state privacy laws require providing notice to the individuals from whom Personal Information is collected of categories of third parties to whom Personal Information will be disclosed, sold or shared, the type of information that will be disclosed, sold or shared, and the purpose for disclosure. EU/UK: Where Input Data is being licensed or shared, the recipients or categories of recipients receiving such Input Data must be disclosed as part of the Privacy Notice to individuals. Where an additional party becomes a controller through this sharing or licensing (i.e., it, alone or jointly, determines the purposes and means of the processing of personal data), it must comply with the Privacy Notice requirements of **Article 14** of the GDPR/UK GDPR. See Section 1.9 in relation to sharing personal data, as similar considerations apply.
1.11. [*For AI models that will receive PHI*] Has a business associate agreement been signed?	US: If an AI service provider will receive PHI, it is important to understand whether they operate as a business associate for the purposes of HIPAA. EU/UK: N/A
1.12. [*For AI models that will receive PHI*] Has compliance with the HIPAA **Privacy**, **Security** and **Breach Notification** Rules been performed for business associates with respect to the AI Service? What measures have been implemented to prevent PHI from being re-disclosed to other customers of the AI Service?	US: In addition to the willingness to sign a business associate agreement, an AI service provider that receives PHI should be able to explain how it safeguards PHI in accordance with the HIPAA regulations and how it prevents unauthorized disclosures of PHI. EU/UK: N/A
1.13. [*For AI models that will receive de-identified data derived from PHI*] Ensure that information that has been de-identified for purposes of HIPAA is not re-identified when input into the AI Service.	US: For de-identified data to remain out of the scope of HIPAA, it is important to ensure that the data is not re-identified after being put into the AI Service. EU/UK: N/A
2. Training Data	

2.1. Understand whether any training data contains Input Data.	US: If the AI service provider uses PHI or Personal Information to train their AI algorithm it is important to make sure that they have gained the necessary consents and permissions from relevant data subjects. It is also important to identify the lawful grounds relied upon for any use of personal data, including PHI, in the training of an algorithm. EU/UK: Recital 71(6) of the GDPR/UK GDPR provides clarity as to why **Article 22** of the GDPR/UK GDPR restricts processing of special category personal data based on automated decision-making, i.e., that to ensure “fair and transparent processing” specific circumstances and context need to be considered so as to (among other things) prevent discrimination on the basis of special category personal data. To process special category personal data (e.g., data concerning health or biometric data), in addition to a lawful basis under **Article 6** of the GDPR/UK GDPR, a further condition of **Article 9** of the GDPR/UK GDPR must be met. The specifics of the processing should be considered so that the controller can determine the correct **Article 9** GDPR/UK GDPR lawful basis, or bases, to be used. Caution should be taken where special category personal data can also be inferred from the AI Service output or as an intermediate step in the process. Even where this result is incidental, or special category personal data is not being used in the first place, ICO **guidance** states that it is possible to use combinations of features that are sufficiently revealing of a special category and which could trigger **Article 9** of the GDPR/UK GDPR. In addition, if an AI Service is being used with the intention of inferring special category personal data, the use of the AI Service in this way means that the data is treated as special category personal data, irrespective of whether the inferences are incorrect.
2.2. Understand what training data set(s) that include Input Data were used to train the AI Service and where the data sets came from.	US: It is important to understand the source of the training data to determine the risks associated with outputs from the AI Service, including ownership rights and reliability of the data. EU/UK: Confirming what data sets are used to train any AI Service is key to ensuring any Input Data can be relied on and has been obtained lawfully. If you are a controller, one or more DPIAs should be conducted where there is a high risk to individuals’ rights and freedoms, and appropriate Privacy Notices should be provided to individuals when the personal data is collected (these requirements are described in Sections 1.3, 1.4 and 3.2). If you are a processor receiving the Input Data from a third party (or a controller sharing Input Data with a processor), a data processing agreement that complies with **Article 28** of the GDPR/UK GDPR will likely be required (along with, potentially, measures to ensure the lawful transfers of personal data to third countries or international organizations).
2.3. Understand whether any training data containing Input Data is in-licensed from a third party.	EU, UK, and US: If the provider uses third-party data for training the AI Service, then it is important to verify that the provider has the appropriate rights from the third party.
2.4. Understand whether any training data containing Input Data is obtained through web scraping or other automated data harvesting (e.g., deep-link, scrape, crawl or use robots, spiders, or other automated programs, devices, algorithms or methods to collect data from websites).	EU, UK, and US: If data were obtained through web scraping, then the AI Service will likely come with more legal risks since the training data may not have been obtained through authorized means. Some concerns with training data obtained by web and server scraping include: (i) potential contract and IP liability if the AI Service uses input/training data without obtaining necessary rights (e.g., if the terms of service of a site prohibit scraping); and (ii) violation of anti-hacking statutes (Federal **Computer Fraud and Abuse Act** and state statutes in the US, the **Computer Misuse Act 1990** in the UK and various laws across Europe) that are intended to protect computer systems against unlawful access. In this situation, consider obtaining data in alternative ways, such as by way of commercial agreements. In the EU/UK, where personal data is anonymized, the data will not be subject to the GDPR/UK GDPR.
2.5. Consider how to verify that the training data sets that include Input Data are accurate and complete.	EU, UK, and US: The entity creating the AI should have a way to test and maintain accuracy of the training data sets. If a third-party entity creating the AI does not have such methods, consider allocating risk in an appropriate manner in the commercial agreement (e.g., through representations, warranties, and indemnities).

2.6. Consider whether there are protections in place to prevent the AI Service from outputting sensitive information (e.g., Personal Information, PHI (US), special category personal data (EU/UK), confidential information)?	EU, UK, and US: Privacy obligations under HIPAA, CCPA, and other state privacy laws in the US, GDPR/UK GDPR in the EU/UK, as well as contractual arrangements in these jurisdictions, require protection against unauthorized disclosures.
3. Oversight	
3.1. Consider the system in place to oversee the functioning of the AI Service, including the nature and qualifications of the technical team that will perform the oversight.	EU, UK, and US: To evaluate the ongoing reliability of an AI Service, and to maintain quality, it is important to keep a human involved, especially where the resulting content is shared externally. Training and processes should be implemented for staff involved with the oversight of AI Services or those that are involved in intervention with automated decisions, to ensure a thorough understanding and ability to assess the accuracy and reliability of the AI Service output.
3.2. Consider how the reliability of the AI Service's output is tested and verified.	US: Incorrect responses confidently asserted as fact by a generative AI tool (frequently referred to as "hallucinations") remain a key issue in the industry. Consider how the AI service provider's control (or lack thereof) over hallucinations will affect use of the AI Service and how it may affect the particular use case. EU/UK: The concept of data accuracy is a fundamental GDPR/UK GDPR principle and requires controllers to ensure that personal data is accurate and, where necessary, kept up to date. Accuracy in the AI context can include assessing statistical accuracy and how often an AI Service will guess the correct answer, which is then likely determined by measuring it against correctly labelled data to reach a percentage success rate. ICO **guidance** confirms that to satisfy the accuracy principle, the AI Service does not need to be 100% statistically accurate. Recital 71 of the GDPR/UK GDPR suggests that "appropriate mathematical and statistical procedures" should be put in place for the profiling of individuals as part of technical measures utilized. From a data protection perspective, where any factors that may result in inaccuracies in personal data are identified, these should then be removed to reduce the risk of errors. ICO **guidance** recommends that a controller should assess in a DPIA: (i) the technical measures designed to heighten accuracy of personal data processed by the AI Service; (ii) how the project might compare human and algorithmic accuracy side-by-side to better justify the AI use; and (iii) whether any trade-offs are made, such as between statistical accuracy and data minimization, along with the rationale for these.
3.3. Does the legal and technical ability exist to audit the AI Service in order to understand the AI Service's technical capabilities (reducing the risk of bias, maintaining appropriate quality) and legal rights and obligations?	US: If the AI service provider does not fully own the AI Service (e.g., it in-licenses a portion from a third party), then it should have the right to audit that portion of the AI Service to understand how it works and make sure the third party is maintaining appropriate quality and reducing risk of bias. EU/UK: AI Services may require audits to ensure that they are compliant with various legislation, including laws relating to data protection and information security. In particular, conducting and documenting audits is necessary to satisfy accountability and documentation requirements under the GDPR/UK GDPR. Where the entirety of an AI Service is not developed internally, audit provisions should be included in any commercial agreement to ensure that appropriate information can be accessed and obtained from the vendor or third party as and when required. Technical staff with a sound understanding of the AI Services should play a lead role in conducting any technical audits, with support from legal and other relevant departments. The ICO has produced an "AI and data protection risk toolkit" which is available on its website. The toolkit provides a good practical starting point to support organizations auditing the compliance of AI Services.
3.4. Ensure policies and procedures are in place for instances in which the AI Service is not functioning correctly, e.g., during system outages.	EU, UK, and US: The AI service provider should have business continuity procedures in place to handle outages or other disruptions in service and physical technical measures to limit such issues where possible. In the US, this is required by the **HIPAA Security Rule** as well as certain state privacy laws. In the EU/UK, there are technical and organizational measures to take account of in **Article 32** of the GDPR/UK GDPR to ensure security of personal data appropriate to the level of risk,

	which include the ability to ensure the ongoing confidentiality, integrity, availability and resilience of processing systems and services and the ability to restore the availability and access to personal data in a timely manner in the event of a physical or technical incident.
3.5. Evaluate how rules and regulations related to the AI Service are assessed and implemented. Are regulatory updates monitored and implemented?	EU, UK, and US: Providers should monitor industry and government statements, events and publications to consider probable regulatory areas and minimize future product disruptions, such as in areas like user notice, bias and sensitive applications.
3.6. Consider whether processes and procedures are in place to assist users with compliance obligations and responding to questions from regulators.	EU, UK, and US: The typical customer of generative AI tools may not have all relevant information at its disposal. Consider adding audit or information rights in any commercial agreement for a generative AI service, as well as some obligation on the provider to assist in compliance efforts.
3.7. Consider whether policies and procedures are in place to ensure compliance under laws related to privacy and data protection (i.e. special category personal data in the EU/UK or PHI and other Personal Information in the US).	US: HIPAA requires the implementation of policies and procedures regarding compliance with HIPAA **Privacy**, **Security** and **Breach Notification** Rules. Adoption of internal policies and procedures regarding compliance with CCPA and other comprehensive state privacy laws is also best practice, if applicable. EU/UK: Where the AI service provider is established in the EU/UK, or the processing activities of the AI Service involves offering goods or services to individuals in the EU/UK or the monitoring of behavior of individuals in the EU/UK, the GDPR/UK GDPR will apply and GDPR/UK GDPR-compliant policies, procedures and processes will be required to be in place. An AI-specific policy is not legally required, but any GDPR/UK GDPR policies and procedures will need to be tailored to ensure they address any specific issues that arise by using such technologies.
3.8. Consider what validation needs to be done before AI can be used.	EU, UK, and US: To evaluate the ongoing reliability of an AI Service, and to maintain quality, it is important validate the quality of Input Data before the AI System is used to develop models and insights.