Making AI GDPR Compliant

As one of the strictest regulations related to data protection and privacy, the EU General Data Protection Regulation (GDPR) has attracted attention from countries all over the world. At the same time, the financial industry, especially Internet finance organizations, has started to explore and invest in artificial intelligence (AI), specifically, a method of designing algorithms that automatically improves and optimizes the services or products provided to customers (e.g., applying facial-recognition technology [FRT] when opening a bank account and machine learning [ML] when evaluating credit ratings).

How to make AI GDPR compliant has become a major discussion. Will GDPR result in the prohibition of AI for use with EU individuals' data? How does one obtain informed consent for an AI algorithm that cannot explain its decision-making criteria? If a user opts out, is an alternative human-based decision system available?

Conflicts Between AI and GDPR

On 8 April 2019, the European Commission released "Ethics Guidelines for Trustworthy AI,"[1] which provides guidance on four ethical principles (i.e., respect for human autonomy, prevention of harm, fairness and explicability) and seven key requirements that AI systems should implement for a trustworthy environment. The principles of "fairness" and "explicability" in the guideline are in accordance with principles of "lawfulness, fairness and transparency" in GDPR. The requirements of "privacy and data governance" and "transparency" in the guideline meet GDPR provisions as well. Therefore, the guideline is worthy to refer to when considering the solutions for AI GDPR compliance.

Consider an organization adopting AI represented as a self-driving car. Data serve as gasoline, which provides the driving force to the car; ML is the automobile engine, which determines the performance of the car; and AI operates as the role of the sensor in the car, contributing to the process of automatic decision-making. A self-driving car with good performance requires more data input to obtain continuous driving force to become more competitive and make more accurate analysis and predictions. However, especially for an Internet finance organization, multiple relational data sets can easily result in "isolated islands of information,"[2] which make it difficult to connect the data sets so they can talk to each other. How to implement data sharing effectively without violating GDPR provisions becomes one of the biggest concerns of AI GDPR compliance.

GDPR's AI-Limiting Provisions

The adoption of AI may violate GDPR provisions with respect to two data subject rights and two GDPR principles. The rights that may be violated include the right to not be subject to automated decision-making and right to erasure. The two GDPR principles that may restrict the use of AI are transparency and data minimization.

Andrea Tang, ISO 27001 LA
Works at one of the Big Four organizations and has experience in data security and privacy for financial institutions. Tang previously worked at Oracle and has IT development experience. She is an active volunteer in the ISACA® China Hong Kong Chapter. Having attended the 2018 and 2019 Asia-Pacific CACS events, Tang has a passion for ISACA events and has organized successful knowledge-sharing events in her region.

Right to Not Be Subject to Automated Decision-Making

The right to not be subject to automated decision-making is prohibited only if the decision-making is based solely on automated processing and produces legal effects concerning the data subject or similarly significantly affects them.[3] However, it is allowed if the process is done with the data subject's explicit consent or the controller has put sufficient safeguards in place.[4] In this scenario, the safeguards provided by the newly released trustworthy AI assessment list include:

- Obtaining human intervention in an unintended way. The way in which AI systems are developed incorporates unfair biases. This could be counteracted by putting in place human oversight processes to analyze and address the system's purpose, constraints, requirements and decisions in a clear and transparent manner. Involving human intervention is necessary to comply with GDPR and to improve the accuracy of the results of AI. Take Internet finance organizations for example. Because of their heavy reliance on AI to provide services or products to customers, it is a must for them to perform credit checking manually to offer customers appropriate credit ratings.
- Explaining the automated decision-making clearly in the privacy policy to notify clients before processing their data. Article 29 Working Party Guidelines on Transparency Under Regulation 2016/679 (WP29) provides guidance on information that must be provided to a data subject under GDPR Article 13 or Article 14.[5] WP29 requires that the existence of automated decision-making including profiling, meaningful information about the logic involved, and the significance and envisaged consequences of such processing for the data subject be in the privacy policy. In addition, clear signage should be present on websites or in mobile applications to highlight where this detailed information can be obtained.
- Obtaining explicit consent to notify a data subject that a decision is the result of an algorithm decision and they are interacting with an AI agent (e.g., a chatbot or robot) or other conversational system.[6] For example, a pop-up window used to collect customers' explicit consent before using AI services is a recommended approach.

Transparent Processing

Both GDPR[7] and the guideline[8] have the common principle of transparency, which requires that data subjects should be informed of the existence and purpose of the processing,[9] especially when the data processing activities involve automated decision-making. Meaningful information about the logic, significance and envisaged consequences of such processing should be explicitly transparent to users. AI service providers should avoid "black boxes"[10] by providing all involved users an explanation as to why a model has generated a particular output or decision (and what combination of input factors contributed to that), which is not always possible. Another advantage of increasing the interpretability of AI algorithms used in data processing is to increase the credibility from users.

> "DE-IDENTIFICATION (PSEUDONYMIZATION) ALLOWS MORE DATA TO BE USED, PROCESSED AND ANALYZED IN AI."

The Right to Erasure

Since data sharing and data openness are core concepts of AI, a large amount of data may be stored in a third party's data server or deployed in the cloud. As a result, it is hard for data controllers to ensure that the server implements the deleting operation or if the data required to be erased are deleted completely from other joint controllers or data processors. The British House of Lords points out that the data link terminal (DLT) contributes to compliance with GDPR regulations by authorizing or accessing the specified DLT.[11] Corda network, which is a public network composed of nodes operated by the participants under strict policies with defined rules for personal data handling by different network services, could contribute to GDPR compliance with respect to the right to erasure.[12]

Data Minimization
De-identification (pseudonymization) allows more data to be used, processed and analyzed in AI. Pseudonymization meets the GDPR principle of data minimization, unlike anonymization, which means that the data subject is no longer or not identifiable. If an Internet finance organization wants to recommend customized financial products without having interest in knowing the actual identities of individuals, it can pseudonymize customers' personal data by replacing the obvious identifiers such as name or email address with a simple reference number. GDPR promotes pseudonymization as an appropriate safeguard for organizations to repurpose data without additional consent.[13] As a result, pseudonymization is an approach to give AI providers flexibility in processing personal data by undergoing different levels of de-identification.

GDPR Compliance Approaches for AI Application in an Internet Finance Organization

In an Internet finance organization, there are many AI applications adopted such as FRT when opening a bank account and ML in evaluating personal loan applications. The application of AI adds a number of risk factors including data accuracy, hackers/bad actors and intended use case. To mitigate the risk, organizations can implement the measures of privacy by design, security by design and trust by design; data protection impact assessment (DPIA); and trustworthy AI assessment.

The following scenarios involve AI applications and detail how each scenario can comply with GDPR.

Scenario 1: Facial-Recognition Technology
Facial- and image-recognition technology can help analyze huge volumes of crime video footage, narrowing down the search and showing where people should focus their attention. Risk factors involved in FRT include:

- **Data accuracy**—The risk of data accuracy may include, but is not limited to, the following scenarios:
  - A photo taken with a similar background or similar facial expression may be recognized as the same person.
  - The system has difficulty recognizing facial features in low light conditions.
  - The system has an inherent racial or gender bias.

"GDPR COMPLIANCE APPROACHES INCLUDE ENSURING THAT PROCEDURES ARE IN PLACE TO MONITOR THE AI AGENT PERFORMANCE AND RESPOND TO DEVIATIONS FROM THE EXPECTED PERFORMANCE."

These risk factors are often caused by a small training data set scope, inappropriate training data collection approaches or inappropriate training data labeling. The risk can be mitigated by collecting more data. However, it may violate the GDPR principle of data minimization, which states that data controllers must only collect and process personal data that are relevant, necessary and adequate to accomplish the purposes for which they are processed.[14] GDPR does not promote data sharing with third parties unless adequate safeguards exist. Article 28(3) has strict regulations on what information must be included in the contract with the outsourcer processors.[15] To solve this problem, federated learning provides some technical support to improve data openness without unsafely transferring data among different organizations.[16]

GDPR compliance approaches include ensuring that procedures are in place to monitor the AI agent performance and respond to deviations from the expected performance including adversarial inputs, out-of-distribution errors, errors in the learning process, unexpected rapid capability gain and other large context changes. They also involve employing strict safety and control measures to prevent uncontrolled evolution of the agent.

- **Hackers/bad actors**—Another risk is brought by hackers/bad actors in which careful manipulation of real-life scenarios and objects

can lead to unexpected outcomes, even in cases where the adversary does not have access to the underlying data or assumptions used in training. Because of the strict regulations of GDPR, beyond encryption, there are many other *de facto* mandatory security-enhancing technologies such as antivirus, antispam, firewalls, identity and access management, and data loss prevention (DLP), to ensure that training and input data are stored in a secure environment.

Rigorous testing is also required. Other possible approaches to make AI GDPR compliant include restricting access to the AI agent production environment to authorized users and conducting penetration tests and cybersecurity control assessments of the AI agent. To "fight fire with fire," organizations should use technology as the answer to cybersecurity concerns that surface amid widespread technological innovation.

- **Intended use case (e.g., direct marketing)**—When processing an individual's personal data in the context of direct marketing activities, GDPR requires that data controllers satisfy all their compliance responsibilities under the regulation, including lawful processing and the transparency requirement. Possible approaches to make AI GDPR compliant include that the AI model designed or selected is commensurate with the interpretability or interrogation requirements of the AI agent considering its objectives, environment, legal or regulatory requirements, and stakeholder expectations. A process is in place to obtain consent from the data subject or group to use the data for the proposed AI agent as required.

Scenario 2: ML Applied in Personal Loan
The loan amount granted by the Internet finance organizations is decided based on customized scoring models, business algorithms and predictive analytics. AI is used to monitor risk by utilizing algorithms and the collection of different sources of data directly from borrowers or third parties. GDPR requires explicit consent (e.g., a privacy policy) to inform customers of the legal basis of their personal data use. Organizations can implement the following AI-GDPR compliance measures:

- **Privacy by design, security by design and trust by design**—GDPR requires the IT system that collects, processes and stores personal data addresses the ongoing operation and management of developments to enable organizations to effectively deal with the data's entire life cycle.[17] The ethics guideline requires AI be secure in its processes, data and outcomes, and it should be designed to be robust against adversarial data and attacks.[18]
- **Data protection impact assessment (DPIA) and trustworthy AI assessment**—GDPR requires that a DPIA is the process by which enterprises systematically assess and identify the privacy and data protection impacts of any products they offer and services they provide.[19] Similarly, the US Congress wrote its concerns about Amazon's facial recognition technology and requested written responses to a list of questions as part of a recommended trustworthy AI assessment.[20] The Amazon assessment includes testing for accuracy and bias, a mechanism for automatically deleting unused data, etc.

> "MORE ORGANIZATIONS ARE BEGINNING TO TAKE APPROPRIATE MEASURES TO COMPLY WITH GDPR USING NEW METADATA MANAGEMENT TOOLS."

GDPR—AI Compliance at Google, Microsoft and SAP

The newly released ethics guidelines require technical safety, privacy, and data governance transparency and fairness.[21] Google, Microsoft and SAP mention fairness (avoid creating or reinforcing unfair bias) and safety in their AI principles.

Google will offer its users an option to automatically delete their search and location history after three months.[22] Microsoft has been a leader in applying innovative techniques for protecting privacy, such as

differential privacy, homomorphic encryption and techniques to separate data from identifying information about individuals.[23] In 2018, SAP Innovative Center Network launched guiding principles on AI, which places data protection and privacy at its core by applying homomorphic encryption.

> MORE ORGANIZATIONS ARE BEGINNING TO TAKE APPROPRIATE MEASURES TO COMPLY WITH GDPR USING NEW METADATA MANAGEMENT TOOLS.

Emerging technologies, such as homomorphic encryption, can help organizations analyze a shared pool of encrypted data without having to disclose information to each other, taking advantages of ML and cryptography to protect special category of data. More organizations are beginning to take appropriate measures to comply with GDPR using new metadata management tools.

There is about to be an explosion in AI adoption. Although two data subject rights and two main GDPR principles may limit the use of AI, the related proposed suggestions are provided in this article. As long as enterprises take adequate safeguards (**figure 1**), implement new technologies, and refer to the Ethics Guidelines for Trustworthy AI and GDPR provisions to protect privacy, it is possible for the best utilization of AI under the strict GDPR regulations.

Figure 1 shows some potential conflicts between AI initiatives and GDPR and suggests remediation efforts.

Conclusion

AI and GDPR regulations may appear to conflict with each other. Does that mean AI use has to be restricted in EU countries? As disruptive technologies advance in the big data era, modern enterprises cannot avoid embracing new technologies. Enterprises should use advanced metadata management tools and technologies to try to minimize the risk of AI in a post-GDPR era.

Endnotes

1 European Commission, "Ethics Guidelines for Trustworthy AI," Belgium, 8 April 2019, *https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai*

Figure 1—The Main Conflicts Between AI and GDPR, and Proposed Remediation

AI vs. GDPR	Proposed Suggestions
Accuracy of automated decision-making	• Obtain human intervention and do not rely solely on a machine. • Use data accuracy analysis technology; monitor the AI agent performance and use ML to increase the accuracy. • Conduct a DPIA and trustworthy AI assessment. • Conduct rigorous testing, e.g., penetration tests and cybersecurity control assessments. • Implement traceability, auditability and transparent communication on system capabilities.
The right to erasure	• Utilize easy removal of information, such as Google's option of automatic deletion of their search and location history.
Data minimization	• Pseudonymize data. • Use data distortion processing technology; keep the property of data for statistics use in AI. • Apply federated ML and transfer learning when there is a need to collect personal data.
Transparency principle	• Use metadata management tools: data governance to authorize specific person accessing the specified DLT. • Have a specific privacy notice and explicit consent. • Use a differential privacy model; delete personally identifiable information without modifying the meaning of datasets.

2 Wallace, N.; D. Castro; "The Impact of the EU's New Data Protection Regulation on AI," Information Technology & Innovation Foundation, 26 March 2018, *https://itif.org/publications/2018/03/26/impact-eu-new-data-protection-regulation-ai*

3 General Data Protection Regulation (GDPR), "Art. 22 GDPR: Automated Individual Decision-Making, Including Profiling," Intersoft Consulting, *https://gdpr-info.eu/art-22-gdpr/*

4 Ustaron, E.; "GDPR—A New Age for Data Protection," Privacy Perspectives, 23 May 2019, *https://iapp.org/news/a/gdpr-a-new-age-for-data-protection/*

5 European Commission, Guidelines on Transparency Under Regulation 2016/679 (wp260rev.01), Belgium, 22 August 2018, *https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=622227*

6 Europa, "Ethics Guidelines for Trustworthy AI," European Commission, 8 April 2019, *https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai*

7 GDPR, "Art. 5 GDPR: Principles Relating to Processing of Personal Data," Intersoft Consulting, *https://gdpr-info.eu/art-5-gdpr/*

8 *Op cit* Europa

9 GDPR, "Art. 13-15, Recital 60, 71," Intersoft Consulting, *https://gdpr-info.eu/art-13-gdpr/*

10 Zhang, J.; *Artificial Intelligence Manager Practice Brochure at the Era of AI*, Electronic Industry Press, China, June 2018

11 Information Security and Communication Security Magazine, "Blockchain Application and Law Risk Analysis," 5 May 2019, *http://www.btb8.com/blockchain/1905/47094.html*

12 Elec Fans, "Relationship Between DLT and Blockchain," 22 April 2019, *www.elecfans.com/blockchain/916000.html*

13 GDPR, "Art. 6: Lawfulness of Processing," Intersoft Consulting, *https://gdpr-info.eu/art-6-gdpr/*

14 *Op cit* GDPR, "Art. 5: Principles Relating to Processing of Personal Data"

15 GDPR, "Art. 28(3) GDPR: Processor," Intersoft Consulting, *http://gdpr-info.eu/art-28-gdpr/*

16 Yang, Q.; "GDPR Issued a Challenge to AI and Federated Transfer Learning," FedAI Ecosystem, 31 August 2018, *https://www.fedai.org.cn/achiList/*

17 GDPR, "Art. 25: Data Protection by Design and by Default," Intersoft Consulting, *https://gdpr-info.eu/art-25-gdpr/*

18 *Op cit* Europa

19 GDPR, "Art. 35: Data Protection Impact Assessment," Intersoft Consulting, *https://gdpr-info.eu/art-35-gdpr/*

20 Suppe, R.; "Amazon's Facial Recognition Tool Misidentified 28 Members of Congress in ACLU Test," *USA Today*, 30 July 2018, *https://amp.usatoday.com/amp/843169002*

21 *Op cit* Europa

22 Simon, M.; "How to Automatically Delete the Web Activity and Location History Data in Your Google Account," *PCWorld*, 9 July 2019, *https://www.pcworld.com/article/3405850/how-to-automatically-delete-the-web-activity-and-location-history-data-in-your-google-account.amp.html*

23 Microsoft Corporation, *The Future Computed—Artificial Intelligence and Its Role in Society*, USA, 2018