December 2019

Skadden

Privacy & Cybersecurity Update

ICO and Turing Institute Publish Draft Guidance on Explaining Artificial Intelligence Decisions

In response to a directive from the U.K. government, on December 2, 2019, the U.K. Information Commissioner's Office (ICO) and The Alan Turing Institute (Turing Institute) published a three-part draft guidance for consultation explaining decisions made with artificial intelligence (AI). The consultation period is open until January 24, 2020. The guidance makes clear that it is not a binding statutory code of practice, but rather a practical instruction that sets out good practices and can aid in compliance.

Part 1: The Basics of Explaining AI

Part 1 includes an analysis of the benefits and risks to organizations of explaining AI-assisted decisions to those that are affected by these decisions. Key benefits noted include: (1) legal compliance (for example, with the General Data Protection Regulation's (GDPR) information requirements for solely automated decision making); (2) the building of trust with an organization's use of AI; and (3) improved internal governance on the basis that having to explain AI-assisted decisions means that organizations will have a better understanding of what these systems do, which may in turn lead to better outcomes. Key risks of explaining AI-assisted decisions include: (1) commercial sensitivities, such as source code or algorithmic trade secrets that an organization would not want to share; (2) inappropriate disclosure of personal data as a result of how AI is trained; and (3) gaming of an organization's AI model if too much about the reasoning underlying the model is shared.

This part of the guidance also sets out the six explanations that the ICO and Turing Institute believe can be used to clarify AI-assisted decisions: (1) a "rationale explanation" provides the reasoning behind a decision; (2) a "responsibility explanation" concerns who is involved in the development and management of the AI model and who can be contacted for a human review of the decision; (3) a "data explanation" gives information on what data has been used with regard to a particular decision, which may involve looking at the data used to train and test the AI model itself; (4) a "fairness explanation" describes the steps taken to ensure that AI-assisted decisions are unbiased and fair; (5) a "safety and performance explanation" helps people to understand the steps taken to maximize the accuracy, reliability, security and robustness of an AI system's decisions and behaviors; and (6) an "impact explanation" considers the impact an AI system in a decision-making process has or may have on an individual and on general society. The guidance suggests that these can be combined into one explanation depending on the decision at hand. For example, if an AI-assisted decision produces a cancer diagnosis, the guidance suggests that the rationale, responsibility, impact, and safety and performance explanation types should be prioritized.

The basics of explaining AI also includes four guiding principles to provide a "broad steer" to assist in explaining AI-assisted decisions to a wide range of individuals given that AI-assisted decisions are not unique to a single sector or type of organization. These principles aid in helping organizations to: (1) be transparent about the use of AI for decision-making; (2) be accountable for AI-assisted decision systems if challenged internally or by external bodies, such as regulators or individuals; (3) consider the context in which the AI-assisted decisions are made; and (4) reflect on the impact of the organization's use of AI on the individuals affected and on wider society. The guidance makes clear that these principles are complementary to the principles outlined under Article 5 of the GDPR.

Part 2: Explaining AI in Practice

Part 2 of the guidance sets out a seven-step process to help guide organizations explain their AI-assisted decisions.

First, organizations should select which of the six explanation types established in Part 1 should be prioritized while keeping in mind the context in which the AI-assisted decision is to be made. The guidance notes that rationale and responsibility explanations will likely always be priority explanations given the need to know what an AI system is doing and who is responsible for it in order to be able to provide an accurate explanation.

Second, organizations should collect the information for each explanation type that they have identified as relevant. For instance, with regard to the responsibility explanation, information on both those responsible within an organization as a whole for the use of the AI, as well as those actually involved in an individual's case, will need to be recorded.

Third, organizations may want to ensure that the underlying logic of the AI system being used can be understood. Of particular concern are the use of "black box" AI systems, which include systems whose workings are unclear to human understanding. The guidance states that maximally interpretable AI systems ought to be used. However, where this is not possible and "black box" models are implemented, the potential impacts and risks should be "thoroughly considered" in advance and supplemental interpretability tools should be put in place.

Fourth, the AI system's results and the underlying logic behind them should be converted to language that is easy to understand. This could, for example, involve explaining what is meant by a probability that an AI system has produced.

Fifth, human decision-makers involved in providing an AI-assisted decisions should receive appropriate training that should provide a basic knowledge of machine learning and its limitations.

Sixth, the organization should consider the context in which the explanation is being delivered. The ICO and the Turing Institute, identified five contextual factors that affect why people may want explanations of AI-assisted decisions: (1) the setting in which the decision is being made (*e.g.*, criminal justice or health care) (domain factor); (2) the impact that the decision can have on an individual and wider society (impact factor); (3) the data used to train and test the AI model and the input data at the point of the decision, as this may impact an individual's willingness to accept or contest an AI-assisted decision (data factor); (4) the importance of receiving or acting upon the outcome of a decision within a short timeframe (urgency factor); and (5) who the recipient of the explanation will be (*e.g.*, what is their level of expertise) (audience factor).

Seventh, organizations should consider how to present their explanation. The guidance notes in particular that explanations may be layered so that priority explanations are given first (*e.g.*, in person), while other explanations could be provided at a later point or in a different form (*e.g.*, a leaflet).

Annex 1 of Part 2 provides an example of AI being used to aid in the explanation of a cancer diagnosis while going through the above seven steps.

Part 3: What Explaining AI Means for Organizations

Part 3 of the guidance considers the roles, policies, procedures and documentation that should be put in place to ensure that an organization is set up to provide explanations to affected individuals. Regarding organizational roles, the guidance notes that anyone involved in the decision-making pipeline has a role to play in explaining an AI-assisted decision. The policies should set out what the rules are, why they are in place and to whom they apply, while the procedures should provide directions on how to implement the rules. Finally, a documentation system should be in place to demonstrate how an AI system can be explained. This should cover both the design and implementation of the AI system, as well as the actual explanation of the AI-assisted decision that is given to an individual.

The ICO has made clear that the guidance[1] is intended to be applicable in the real world and hopes that the consultation period will adhere to that goal. The final version of the guidance is expected to be published in 2020.

Return to Table of Contents

Email Management Company Settles FTC Allegations Over Alleged Misleading Statements

Pursuant to a settlement reached with the Federal Trade Commission (FTC), email management company Unrollme is enjoined, for the next 20 years, from making any misleading statements regarding how it accesses and shares users' emails. The FTC had alleged that the company misled consumers who had initially declined to grant access to their emails during the signup process for the company's services.

Background

Unrollme, Inc. is a company that offers two services to help consumers organize and manage daily subscription emails. First, it helps users unsubscribe from unwanted subscription emails. Second, to declutter users' inboxes, it consolidates all the wanted subscription emails into a single, daily email called the "Rollup." To provide these services, Unrollme requests full access to users' email accounts.

Upon obtaining full access to users' email inboxes, Unrollme allows its parent company Slice Technologies, Inc., a market research company, to generate anonymized market research from users' e-receipts (*i.e*., purchase or order confirmation emails). Slice uses an automated crawler to capture the entire body of e-receipts — which results in the capture of personally identifiable or sensitive information, such as the user's name, billing and shipping addresses, credit card information, and the identity of purchased products or services that reveal, for example, a user's medical condition. Slice then uses a "parser" to extract data from the e-receipts and create a database of anonymous purchase information, which it sells as part of its market research analytics products. Slice stores the data until the user deletes his or her Unrollme account.

Unrollme did not inform users on its homepage, FAQ page or any screen displayed to users during the signup process that it or Slice would collect, store and sell information from e-receipts. Unrollme's privacy policy, however, disclosed that it may collect and sell "data from and about [user's] commercial electronic email messages" and "transactions or relationship messages" as defined by the CAN-SPAM Act. Although Unrollme users generally were required to click a box agreeing to the privacy policy, they were not required to actually view it.

When consumers declined to grant Unrollme permission to access their email accounts, a message would appear encouraging them to change their minds and continue with the signup process. The messages changed over time. From approximately January to November 2015, Unrollme's message stated:

> "It looks like you clicked No thanks. In order to use Unrollme, you need to tell [your email service provider] to allow us to monitor your emails. Don't worry, we won't touch your personal stuff."

From approximately November 2015 through October 2016, Unrollme's message stated:

> "Authorization Declined. In order to use Unrollme, you need to authorize us to access your emails. Don't worry, this is just to watch for those pesky newsletters, we'll never touch your personal stuff."

And from October 2016 through September 2018, Unrollme's message stated:

> "Oops! Looks like you declined access. Unrollme requires access to your inbox so we can scan for subscriptions and allow you to begin clearing out your inbox."

More than 55,000 users completed the sign-up process after viewing one of these messages.

In addition, Unrollme's customer service representatives responded to questions about the company's privacy policy — which certain users found confusing — by stating that Unrollme only accesses users' emails to provide its free services, making no mention of Slice accessing and collecting the entire body of e-receipts.

The FTC's Allegations

The FTC's complaint alleged that Unrollme violated Section 5(a) of the Federal Trade Commission Act, 15 U.S.C. § 45(a) by making material, false and deceptive statements designed to encourage users to grant access to their email accounts and continue the sign-up process for Unrollme's services.[2]

[1] You can access the full draft guidance here.

[2] The complaint can be read here.

The first count alleged that the messages in effect from approximately January 2015 through October 2016 were false or misleading because Unrollme granted Slice "access to its users' inboxes, including personal emails in the form of e-receipts, which is then used to collect and sell purchase information contained therein to third parties."

The second count alleged that the message in effect from approximately October 2016 through September 2018 failed to disclose (or adequately disclose) that Unrollme also granted Slice access to its users' inboxes and that "fact would be material to consumers in their decision to use Unrollme's services."

The Settlement

Under the FTC's consent order, Unrollme is prohibited, for the next 20 years, from misrepresenting the "extent to which it accesses, collects, uses, stores, or shares" users' emails and personally identifiable information in regard to any product, service or software requiring access to users' emails. Unrollme also must take affirmative steps in regard to current users who enrolled in the company's services after viewing the allegedly misleading messages. First, Unrollme must send those users an email notice, prescribed in the consent order itself, informing users that the company or its parent "access or collect email purchase receipts for use in market research products that are sold to third parties." Second, Unrollme must delete all stored e-receipts and all personally identifiable information obtained from those receipts. The consent order also imposes reporting and compliance obligations, which include keeping consumer complaints and screenshots of the user enrollment interface. The consent order carries the force of law and each violation may result in a civil penalty of up to $42,530.[3]

Key Takeaways

Companies must provide clear and complete disclosures on how they access and use their customers' information, including in their privacy policies and any instance in which they ask customers for access to accounts or other repositories that contain personally identifiable or sensitive information. Failure to do so could result in potential liability and reputational harm, particularly for companies whose services require consumers to make a trade-off between privacy and using the service. Companies also cannot simply rely on broad statements in a privacy policy regarding data usage to insulate themselves from liability for misleading consumers in other communications.

Return to Table of Contents

National Association of Insurance Commissioners Releases Draft Principles for the Use of AI in Insurance

In early December 2019, the National Association of Insurance Commissioners (NAIC) released draft principles on artificial intelligence, which are intended to guide insurance companies and persons or entities facilitating the business of insurance on the "responsible stewardship of trustworthy AI" in the insurance industry. The high-level principles are organized under five headings that form the acronym "FACTS," which state that the use of AI should be Fair and Ethical; Accountable; Compliant; Transparent; and Secure, Safe and Robust.

The NAIC Draft Principles on Artificial Intelligence[4] state that all insurance companies and employees that use or design AI systems (AI actors) should adhere to the following principles:

Fair and Ethical

"AI actors should respect the rule of law" and "should proactively engage in responsible stewardship of trustworthy AI in pursuit of beneficial outcomes for consumers and society," while respecting "cultural, social and legal norms where they operate."

Accountable

AI actors "should be accountable for the proper functioning of AI systems" that they design and use, even when those AI systems create unintended outcomes. AI actors should "implement mechanisms and safeguards" to ensure that AI systems are subject to "ongoing human monitoring" and "human intervention." Stakeholders should have access to "plain, easy-to-understand" information about how AI systems use their insurance data, as well as ways to "seek recourse of AI-driven decisions."

Compliant

AI actors are required to have "specific knowledge of all applicable federal and state insurance laws and regulations," and ensure that AI systems are compliant. AI systems must be subject to "consistent monitoring" for legal compliance, particularly to guard against outcomes that are "unfairly discriminatory" or violate "cultural, social and legal standards."

[3] The order can be read here.

[4] The draft principles are available here.

Transparent

To increase public confidence, AI actors must be transparent about their use of AI and make "proactive disclosures to stakeholders" about the data being used by AI systems, the purpose of that use and the potential consequences. Transparency should be achieved "while maintaining the ability to protect confidentiality of proprietary algorithms and adherence to individual state regulations in all states where AI is deployed."

Secure, Safe and Robust

AI systems should remain "robust, secure and safe" so that they can "function accurately and appropriately" during "normal use and reasonably foreseeable use or misuse." AI systems should be designed to "ensure traceability in relation to datasets, processes and decisions made" by the system and to enable analysis of the system's actions and responses. AI actors should employ a "systematic risk management approach" to AI systems to properly address risks related to privacy, security and bias.

Key Takeaways

It is anticipated that the FACTS AI principles should be finalized and adopted by the NAIC in 2020, after which the association likely will follow with more specific guidance or model law. In light of the continued increase in use of AI in the insurance industry, including in areas such as customer analysis, claims management, operations and fraud prevention, the NAIC's work may meaningfully impact consumer protection and privacy, marketplace dynamics and the state-based insurance regulatory framework in this ever-evolving arena.

Return to Table of Contents

Privacy & Cybersecurity Update

Contacts

Stuart D. Levi
Partner / New York
212.735.2750
stuart.levi@skadden.com

James Carroll
Partner / Boston
617.573.4801
james.carroll@skadden.com

Brian Duwe
Partner / Chicago
312.407.0816
brian.duwe@skadden.com

David Eisman
Partner / Los Angeles
213.687.5010
david.eisman@skadden.com

Patrick Fitzgerald
Partner / Chicago
312.407.0508
patrick.fitzgerald@skadden.com

Todd E. Freed
Partner / New York
212.735.3714
todd.freed@skadden.com

Marc S. Gerber
Partner / Washington, D.C.
202.371.7233
marc.gerber@skadden.com

Rich Grossman
Partner / New York
212.735.2116
richard.grossman@skadden.com

Michael E. Leiter
Partner / Washington, D.C.
202.371.7540
michael.leiter@skadden.com

Amy Park
Partner / Palo Alto
650.470.4511
amy.park@skadden.com

William Ridgway
Partner / Chicago
312.407.0449
william.ridgway@skadden.com

Jason D. Russell
Partner / Los Angeles
213.687.5328
jason.russell@skadden.com

Ivan Schlager
Partner / Washington, D.C.
202.371.7810
ivan.schlager@skadden.com

David Schwartz
Partner / New York
212.735.2473
david.schwartz@skadden.com

Jen Spaziano
Partner / Washington, D.C.
202.371.7872
jen.spaziano@skadden.com

Ingrid Vandenborre
Partner / Brussels
32.2.639.0336
ingrid.vandenborre@skadden.com

Helena Derbyshire
Of Counsel / London
44.20.7519.7086
helena.derbyshire@skadden.com

Jessica N. Cohen
Counsel / New York
212.735.2793
jessica.cohen@skadden.com

Peter Luneau
Counsel / New York
212.735.2917
peter.luneau@skadden.com

James S. Talbot
Counsel / New York
212.735.4133
james.talbot@skadden.com

Eve-Christie Vermynck
Counsel / London
44.20.7519.7097
eve-christie.vermynck@skadden.com

This communication is provided by Skadden, Arps, Slate, Meagher & Flom LLP and its affiliates for educational and informational purposes only and is not intended and should not be construed as legal advice. This communication is considered advertising under applicable state laws.

Skadden, Arps, Slate, Meagher & Flom LLP
Four Times Square
New York, NY 10036
212.735.3000