Putting Ethics into Practice with Your SAP Security Strategy

How to Ensure an Ethical Implementation of Artificial Intelligence

by **Guido Wagner**, Chief Development Architect, SAP SE

With any new technology, it is important to consider the potential risks and threats it can bring, and to develop a solid strategy for avoiding them before that technology is widely used. As artificial intelligence (AI) moves toward becoming a standard technology in daily business — such as the use of conversational user interface technology in call centers and financial close processes driven by machine learning — it is critical that organizations ensure they have a well-thought-out plan in place for mitigating any potential issues.

While some visionary minds, such as Elon Musk and Stephen Hawking, have issued warnings about the longer-term existential risks AI can pose to society,[1] in the near term, organizations must consider the general impact AI can have on business operations. For example, automated decision making that doesn't consider extenuating circumstances, the perception of misuse of personal data, and machine

[1] Bernard Marr, "Is Artificial Intelligence Dangerous? 6 AI Risks Everyone Should Know About" (Forbes, November 19, 2018; http://bit.ly/6AIRisks).

learning based on faulty data can significantly affect the security of a company's business model and reputation. With AI, companies increasingly need to balance the safety and privacy of people with the pursuit of growth and success.

This article shows security and risk managers and decision makers who are considering the use of AI in their organization's software system landscape how to mitigate the risks posed by AI software by expanding existing security standards with a clearly defined set of digital ethics. It first looks at some typical ethical challenges businesses using AI-based software face and how these challenges can be mitigated. It then provides an overview of SAP's own guiding principles for developing and deploying AI-based software and offers advice for how to get started developing a strategy for ethical AI in your own organization.

Security Meets Ethics

So what are some of the ethical challenges businesses face when it comes to AI software? Here, we look at a few typical examples — specifically, ethical challenges around safety, privacy, and bias when using AI software — and some ways you can start to think about mitigating these risks.

Ethical Challenges Around Safety and Privacy

One business area that involves ethical challenges is the balance between profit and safety risks. Imagine you are on vacation and a hurricane is on its way to your location. You want to fly home, but the airfare is much higher than usual because some pricing software detected a spike in bookings without factoring in the potential impact on your safety. Similar scenarios have occurred when it comes to the price of drinking water during times of water shortage. In addition to creating safety risks, these scenarios can also tarnish the company's public image. Mitigating measures for these scenarios would be to include the event (the hurricane or water shortage, for instance) in the price-finding algorithm (via automated rules or a rule that triggers an alert for human intervention, for example) and provide unambiguous instructions for handling the situation.

Another area of ethical challenges that can influence how a company is perceived is data privacy. There are many existing regulations about the usage of personal data — such as the EU's General Data Protection Regulation (GDPR)[2] — that do permit the use of anonymized data that cannot be connected to an identifiable person. However, many people do not want even anonymized data about their behavior used for any purpose. While this is not generally a legal compliance issue, it is a good example of the concept of farther-reaching "ethical compliance," and of how even just the perception that a company is violating people's privacy can cause a damaging lack of trust in that business. This type of scenario must be addressed by additional standards, and may include some academic research to understand how factors such as cultural backgrounds can impact ethical requirements, and how technological solutions can mitigate the impact.

[2] For more on GDPR, see the article "Meeting Modern Data Protection Requirements" in the July-September 2017 issue of *SAPinsider*, available at SAPinsiderOnline.com.

"Automated decision making that doesn't consider extenuating circumstances, the perception of misuse of personal data, and machine learning based on faulty data can significantly affect the security of a company's business model and reputation."

— **Guido Wagner**, Chief Development Architect, SAP SE

Ethical Challenges Around Bias

Other types of ethical challenges have more technical roots, such as machine-learning software that bases its behavior on a certain selection of data. One well-known example is the chatbot "Tay," which learned slang from reading posts on Twitter and eventually began to post inflammatory and offensive tweets.[3] The reason this happened is simple: the bot was trained using the data of only a small sample of human culture — the language typically used on Twitter by people with specific interests and goals. Another example would be a hiring app that is trained using data that consists almost exclusively of profiles of male applicants and employees — in this scenario, the system (and hiring process) will be weighted against anyone who doesn't fit that same profile.

This is one aspect of what is called "bias," and while avoiding it can be a significant challenge for software developers and AI trainers, it is critical from a security perspective. Imagine what might happen if AI-based software is trained with an unbalanced data set, whether unintentional or with fraudulent intent. While this might not be a huge problem in the case of simple image recognition, what happens if it's a neural network that a business uses to run prediction routines, and that managers rely on to make decisions? Recommendations from such a system will be misleading, even if the algorithms are correct, which can affect not only individuals, but also business operations and the company's public image.

If there is fraudulent intent involved, it can be hard to identify or prove, since technologies such as neural networks do not work as transparently as previously used analytical algorithms — neural networks do not contain a well-defined set of algorithms, so it is nearly impossible to tell how exactly they calculated a result. Because of this missing transparency, it is often not possible to reproduce how a trained deep-learning system comes to a particular result.

The best chance to mitigate the risk of bias is to deeply understand the use cases of a system when designing it, and to ensure the completeness of the training data. Systems designed in this way will have limitations — that is, boundaries within which they work to ensure a low bias level — that are clearly communicated by the manufacturer or trainer of the system. Another possibility is to create qualifiers — such as a "bias index of learned data" — that could be helpful for evaluating the quality of AI recommendations. This is an idea that needs research and standardization to become a valid approach, however.

A Framework for Digital Ethics

The ethical challenges surrounding AI-based software is a key consideration for any organization that is moving forward into a digital business model, including SAP. One way SAP is taking action to address these challenges is by participating in broad discussions about establishing norms around the use of AI-based software that range across industries, academics, and politics. For example, SAP is engaged with the Partnership on AI,[4] the European Commission's High-Level Expert Group on Artificial Intelligence,[5] and the Council on the Responsible Use of AI.[6]

Another way SAP is addressing these challenges is by creating a framework of ethical values that goes beyond legal compliance to take into account the effect AI-based software can have on people's quality of life. As a starting point, in the summer of 2018, SAP founded an AI Ethics Steering Committee and created a set of seven guiding principles for developing and deploying AI-based software, such as its SAP Leonardo Machine Learning portfolio. See the sidebar "SAP's Guiding Principles for AI" on the next page for an overview of these principles.

These principles are maintained and executed by SAP's AI Ethics Steering Committee, which consists of nine executive managers, including the heads of design, machine learning, legal, data protection, and sustainability. The AI Ethics Steering Committee is supported by an external AI Ethics Advisory Panel where academic experts contribute

[3] Jane Wakefield, "Microsoft Chatbot Is Taught to Swear on Twitter" (BBC News, March 24, 2016; http://bit.ly/TayChatbot).

[4] Learn more about the Partnership on AI at https://www.partnershiponai.org.

[5] Learn more about the European Commission's High-Level Expert Group on Artificial Intelligence at http://bit.ly/ECExpertGroupAI.

[6] Learn more about the Council on the Responsible Use of AI at http://bit.ly/AICouncil.

not only from the point of view of IT, but also from the perspective of biology, theology, legal, and other sciences and areas of study.

The AI Ethics Steering Committee is supported internally by a diverse expert group and collaborates with an internal "think cell" that handles questions about business ethics in digitalization scenarios. This type of setup has several advantages. For example, it facilitates a close connection to the board and direct access to employees who define, build, and implement AI. It also connects the committee with relevant discussions on ethics outside of SAP and ensures that upcoming questions around business ethics are involved in all AI-related work.

Developing Principles and Putting Them into Practice

Developing a set of guiding principles for digital ethics is an important first step for any software provider working with AI technology. There are three tasks in particular that are key to paving the way to ethical, sustainable AI practices based on guiding principles similar to the ones outlined by SAP: gathering requirements for ensuring the implementation of ethical AI, adding ethical AI information to your existing standards and policies, and monitoring and auditing AI activities. Let's take a closer look.

Gathering Requirements for Ensuring the Implementation of Ethical AI

To build a successful set of principles, it is a good practice to first gather requirements based on customer feedback and academic discussions. Questions might include: In which cases must a system involve a human decision maker? What should the human-machine interaction look like? Which processes must be logged or monitored? Which parameters must be customizable to enable ethical system behavior? Within the purchasing process,

SAP's Guiding Principles for Artificial Intelligence

SAP organizes its policies for developing and deploying software based on artificial intelligence (AI) around seven guiding principles, which are summarized here (the complete text of the principles is available at **http://bit.ly/SAPAIPrinciples**):

1. We are driven by our values. SAP actively supports a variety of UN-defined ethical precepts, including the Guiding Principles on Business and Human Rights,* and developed the SAP Global Human Rights Commitment Statement** that spells out its commitment. Furthermore, SAP's AI Ethics Steering Committee operates based on this principle.

2. We design for people. SAP strives to make the user experience of its software human-centered, where users can interact with systems just as they interact with other humans. SAP seeks to be inclusive with its AI software, and to empower and augment the talents of its diverse users. To achieve this goal, SAP actively drives co-innovation with customers.

3. We enable businesses beyond bias. Bias must be seen as a major risk when creating AI-based advisory or decision-making systems. SAP asks its teams to gain a deep understanding of the business problems they are trying to solve, and the data quality this demands. SAP is investigating new technical methods for mitigating biases and is open to co-innovation opportunities in this area.

4. We strive for transparency and integrity in all that we do. SAP customers will always remain in control of the deployment of SAP products, and SAP will clearly communicate the intended purpose of its products, their capabilities, and their limitations.

5. We uphold quality and safety standards. Ensuring the quality of AI-based products and the safety of humans when using them is at least as important as for any other product. AI-based products will be subject to quality assurance processes, which will be reviewed and adapted on a regular basis.

6. We place data protection and privacy at our core. Data protection is relevant to ethical behavior beyond legal compliance requirements — it also includes consideration of the impact AI can have on people's quality of life. SAP will communicate clearly how, why, where, and when customer and anonymized user data is used in its AI-based software. Together with partners, SAP will continue to research the development of the next generation of privacy-enhancing technologies.

7. We engage with the wider societal challenges of AI. SAP is aware that AI is one of the major drivers of what is widely known as the "future of work" discussion, and that AI has the potential to cause ethical dilemmas in business software. SAP will continue to consider and discuss the social impact and economics of AI across industries, borders, cultures, and philosophical and religious traditions.

* The UN's Guiding Principles on Business and Human Rights is available at http://bit.ly/UNHumanRightsGuide.

** The SAP Global Human Rights Commitment Statement is available at http://bit.ly/SAPHumanRightsGuide.

Learn More

- The Partnership on AI
  https://www.partnershiponai.org
- The European Commission's High-Level Expert Group on Artificial Intelligence
  http://bit.ly/ECExpertGroupAI
- The Council on the Responsible Use of AI
  http://bit.ly/AICouncil
- SAP's Guiding Principles for Artificial Intelligence
  http://bit.ly/SAPAIPrinciples
- UN Guiding Principles on Business and Human Rights
  http://bit.ly/UNHumanRightsGuide
- SAP Global Human Rights Commitment Statement
  http://bit.ly/SAPHumanRightsGuide
- SAP Leonardo Machine Learning: AI Ethics and Society
  https://www.sap.com/products/leonardo/machine-learning/ai-ethics.html

for example, it could be a requirement to define a certain level of fair-traded goods and instruct the AI-based software to choose vendors accordingly. It could also be a requirement to ask users before using their personal data, even if the data is anonymized. These types of requirements must be gathered in close collaboration with customers, AI providers, and people who handle business ethics questions (executive leaders, portfolio or compliance managers, and sustainability departments, for instance).

Checklists can also be helpful for identifying the requirements needed to ensure ethical AI. Checklist items should include questions related to human involvement in AI, such as the end user's cultural values, how the end user's current context is evaluated, and situations in which the end user will want AI functionality turned off. Additional checklist items should focus on AI algorithms and boundary conditions, such as how "learn and forget" processes should be monitored to detect fraudulent activities, how a minimum training status can be determined, and to what extent computational results must be reproducible. Checklist items should also consider legal compliance requirements (such as data privacy regulations), how to unveil hidden override directives, and how to assess the potential long-term impact of AI operations. Will humans — or humanity — lose knowledge or capabilities? How can behavioral changes of the AI system be detected (due to hacking activities, for instance)?[7]

The requirements you gather will help you identify the areas in which you need to operationalize your guiding principles. The next step is to transform the requirements into additions that you make to your existing product standards and corporate policies.

Adding Ethical AI Information to Existing Standards and Policies

Product standards and policies are proven to help ensure quality, including security aspects. Your organization's definition of ethical AI — and how to monitor it — can be included in implementation

[7] For a detailed exploration of how to use checklists to ensure ethical AI, see the July 16, 2018, Digitalist blog post, "A Checklist of Ethical Design Challenges for Business AI" (http://bit.ly/AIChecklists).

and operations standards as well as in security and audit policies to ensure widespread awareness and understanding across the business.

Practical instructions for everyone involved in the AI life cycle can result from adding this information to policies and standards. The information must include patterns for human-machine interaction in specific situations, customization parameters to fulfill specific cultural requirements, and procedures to overrule AI (if reasonable and secure).

Monitoring and Auditing AI Activities

Automated controls — for tracking the level of fair-traded goods in a purchasing process or the use of anonymized, human-related data, for instance — can help with monitoring AI activities and supporting audits of the AI system's behavior by ensuring that procedures are being followed. For example, automated controls could monitor price-finding algorithms for scenarios such as water shortages and apply rules for handling cases in which human health might be affected (to stop any automated price increases, for instance). You can also support audits of the AI system's behavior by reviewing any available information about why an AI system came to a particular decision. Keep in mind that evaluating a user's current situation is as important as assessing the potential risks related to alternative actions.

Another method for auditing the reasonable operability of AI is to turn off the AI algorithms and use the raw data and different calculation methods to generate the results. If the resulting data is different from the AI-based results, something obviously is wrong. Turning off AI functionality and providing more basic data to the user can also be a requirement — humans sometimes do not want to only rely on a system output, but rather support their gut feelings and come to their own decisions. Of course, using "exit doors" — that is, turning off AI algorithms — is not always possible, especially if immediate action is required, such as with high-speed trading. In cases where trend-setting decisions are required — such as adjusting a product portfolio, closing a branch office, or making a decision about investments — the ability to turn off AI algorithms at least for test purposes may help to avoid misuse or identify fraudulent changes by hackers or competitors. The results of such analyses must become part of product standards to ensure that business managers can rely on AI-based proposals.

In general, the process of AI auditing will be of high interest to insurance agents and lawyers when it comes to liabilities based on a system's decisions and proposals, but it is also relevant and useful for any organization that is planning on using AI-based software. It is important to be aware of any potential issues related to the use of AI-based systems, since these issues represent risks. It is good practice to be proactive about how to mitigate these risks using additional security-related measures, such as implementing random reviews of AI behaviors, controls, and audits, or specifying human-machine interaction schemas for situations in which someone must make a decision relevant to a person's ethical attitude.

Summary

When it comes to AI-based software, to ensure the lowest possible level of risk for human life and the highest possible level of integrity in modern enterprises, it is crucial to have security management policies in place. These policies should include standards and directives that not only are based on legal compliance regulations, but also incorporate well-thought-out guidelines for the ethical use of these solutions. Just as effective security management requires collaboration across company lines, implementing a successful digital ethics strategy for AI requires integrating ethical principles across your organizational standards, policies, and behaviors. AI is increasingly influencing people's lives, and security and digital ethics can help ensure it is a positive impact. ■

Guido Wagner (guido.wagner@sap.com) is responsible for innovation projects in SAP Design. He focuses on user experience optimization in a digitalized business environment. Preparing for the future of work through sustainable artificial intelligence that improves the way people live is his passion. Share your thoughts with Guido on LinkedIn at https://www.linkedin.com/in/guido-wagner-693965/.