เฟรมเวิร์ก AI ที่ปลอดภัยของ Google (SAIF)

ศักยภาพของ AI โดยเฉพาะอย่างยิ่ง Generative AI นั้นไร้ขีดจำกัด ขณะที่นวัตกรรมก้าวไปข้างหน้า อุตสาหกรรมนี้จำเป็นต้องมีมาตรฐานการรักษาความปลอดภัยในการสร้างและใช้งาน AI อย่างมีความรับผิดชอบ นั่นคือเหตุผลที่เราเปิดตัวเฟรมเวิร์ก AI ที่ปลอดภัย (SAIF) ซึ่งเป็นเฟรมเวิร์กของแนวคิดที่จะรักษาความปลอดภัยให้ระบบ AI

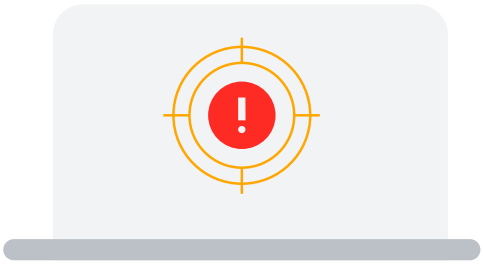
องค์ประกอบหลักของ SAIF | ระบบนิเวศที่ปลอดภัยยิ่งขึ้น | แหล่งข้อมูล | คำถามที่พบบ่อย

6 องค์ประกอบหลัก ของ SAIF

SAIF ออกแบบมาเพื่อแก้ไขข้อกังวลอันดับต้นๆ ของผู้เชี่ยวชาญด้านความปลอดภัย เช่น การจัดการความเสี่ยงของโมเดล AI/ML, การรักษาความปลอดภัย และความเป็นส่วนตัว เพื่อช่วยให้มั่นใจได้ว่าเมื่อนำโมเดล AI ไปใช้งาน โมเดลเหล่านี้ก็จะปลอดภัยโดยค่าเริ่มต้น

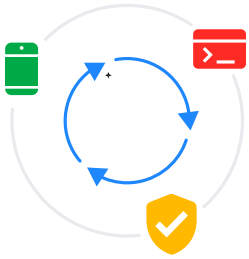
ดาวน์โหลดไฟล์ PDF

ขยายพื้นฐานการรักษาความปลอดภัยที่แข็งแกร่งให้ระบบนิเวศ AI

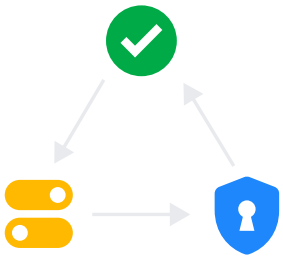


ขยายการตรวจจับและการตอบสนองภัยคุกคามขององค์กรให้ครอบคลุมการทำงานของ AI ด้วย

ทำให้การป้องกันเป็นระบบอัตโนมัติเพื่อก้าวทันทั้งภัยคุกคามเดิมและที่เกิดขึ้นใหม่



ผสานการควบคุมระดับแพลตฟอร์มเป็นหนึ่งเดียวกันเพื่อให้มั่นใจได้ถึงความปลอดภัยที่สอดคล้องกันทั้งองค์กร

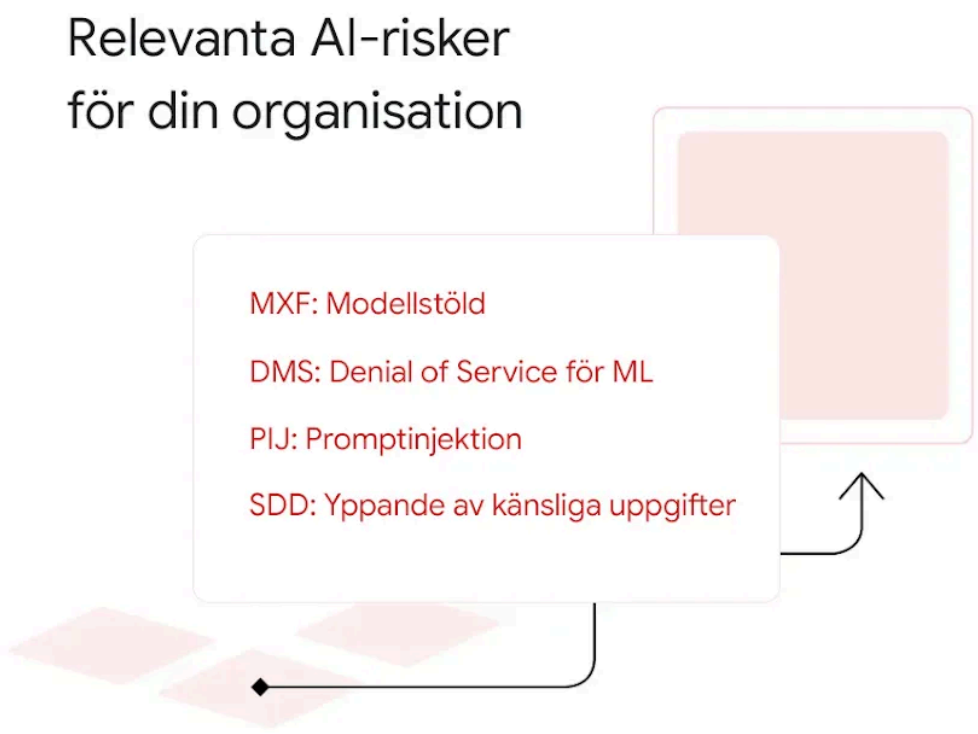


ปรับการควบคุมเพื่อลดความเสี่ยงและสร้างกระบวนการนำฟีดแบ็กมาแก้ไขที่รวดเร็วขึ้นสำหรับการใช้งาน AI

ประเมินความเสี่ยงของระบบ AI เมื่อนำไปใช้ในกระบวนการธุรกิจที่เกี่ยวข้อง

การทำให้ ระบบนิเวศปลอดภัยยิ่งขึ้น

เราตื่นเต้นอย่างยิ่งที่จะได้บอกเล่าถึงก้าวแรกในเส้นทางสู่การสร้างระบบนิเวศ SAIF สำหรับทั้งภาครัฐบาล เอกชน และองค์กรต่างๆ เพื่อเพิ่มความก้าวหน้าให้เฟรมเวิร์กที่จะทำให้ AI ใช้งานได้อย่างปลอดภัยซึ่งเหมาะสำหรับทุกฝ่าย



ขอแนะนำ SAIF.Google: AI ที่ปลอดภัยเริ่มต้นที่นี่

SAIF.Google เป็นศูนย์รวมแหล่งข้อมูลเพื่อช่วยผู้เชี่ยวชาญด้านความปลอดภัยในการทำความเข้าใจภูมิทัศน์ของการรักษาความปลอดภัยให้ AI ที่เปลี่ยนแปลงอยู่ตลอดเวลา ซึ่งเว็บไซต์นี้ได้รวบรวมความเสี่ยงและการควบคุมต่างๆ ในด้านความปลอดภัยของ AI รวมถึง "รายงานการประเมินความเสี่ยงด้วยตนเอง" เพื่อเป็นแนวทางสำหรับผู้ปฏิบัติในการทำความเข้าใจความเสี่ยงต่างๆ ที่อาจส่งผลกระทบต่อพวกเขาและวิธีการนำ SAIF ไปใช้ในองค์กรของตนเอง แหล่งข้อมูลเหล่านี้จะช่วยตอบสนองความต้องการที่สำคัญอย่างยิ่งในการสร้างและปรับใช้ระบบ AI ที่ปลอดภัยในโลกที่เปลี่ยนแปลงอย่างรวดเร็ว

เริ่มต้นใช้งาน

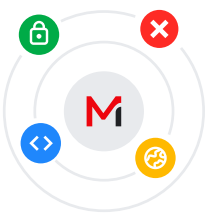
การนำ SAIF ไปสู่รัฐบาลและองค์กรต่างๆ

เราร่วมมือกับรัฐบาลและองค์กรต่างๆ เพื่อช่วยลดความเสี่ยงด้านความปลอดภัยของ AI การทำงานร่วมกับผู้กำหนดนโยบายและองค์กรด้านมาตรฐาน เช่น NIST ส่งผลให้กรอบการกำกับดูแลเปลี่ยนแปลงไป

Coalition for Secure AI: การขยาย SAIF ร่วมกับพันธมิตรในอุตสาหกรรม

เรากำลังผลักดันงานนี้ให้เดินหน้าและส่งเสริมการสนับสนุนจากอุตสาหกรรมด้วยการก่อตั้ง Coalition for Secure AI (CoSAI) โดยมีสมาชิกผู้ก่อตั้งอย่าง Anthropic, Cisco, GenLab, IBM, Intel, Nvidia และ PayPal เพื่อรับมือกับความท้าทายที่สำคัญในการนำระบบ AI ที่ปลอดภัยไปใช้

แหล่งข้อมูล เพิ่มเติม



การเพิ่มความปลอดภัยให้ AI: AI Red Team ของ Google

ดูว่า AI Red Team ของ Google ซึ่งมีกลยุทธ์สุดล้ำสมัย ปรับปรุงการรักษาความปลอดภัยให้ระบบ AI อย่างไร ศึกษาข้อมูลเชิงลึกและบทเรียนสำคัญๆ ในรายงานล่าสุดของเรา

ดูข้อมูลเพิ่มเติม

รักษาความปลอดภัยให้ระบบ AI ด้วย Mandiant

Mandiant เร่งรัดให้เกิดการผสานรวมความปลอดภัยเชิงรุกลงในระบบ AI ที่สอดคล้องกับ SAIF เพื่อการป้องกันที่แน่นหนา

ดูข้อมูลเพิ่มเติม

Android: หลักเกณฑ์การพัฒนาที่ปลอดภัย

ปกป้ององค์กรของคุณให้ปลอดภัยด้วยการแจ้งเตือนช่องโหว่แบบเรียลไทม์ของ Android และปฏิบัติตามหลักเกณฑ์การพัฒนาที่ปลอดภัยสำหรับโค้ดแมชชีนเลิร์นนิง

ดาวน์โหลดไฟล์ PDF

คำถามที่พบบ่อย เกี่ยวกับ SAIF

ขยายทั้งหมด

SAIF และ AI ที่มีความรับผิดชอบเกี่ยวข้องกันอย่างไร

Google มีความจำเป็นที่จะต้องสร้าง AI อย่างมีความรับผิดชอบ และสนับสนุนให้ผู้อื่นทำเช่นเดียวกัน หลักการเกี่ยวกับ AI ของเราที่เผยแพร่ในปี 2018 ได้อธิบายถึงความมุ่งมั่นที่เรามีต่อการพัฒนาเทคโนโลยีอย่างมีความรับผิดชอบ และในลักษณะที่สร้างมาเพื่อความปลอดภัย ทำให้เกิดความรับผิดชอบ และรักษามาตรฐานระดับสูงของความเป็นเลิศทางวิทยาศาสตร์ AI ที่มีความรับผิดชอบคือแนวทางของเราที่ครอบคลุมหลากหลายมิติ เช่น "ความยุติธรรม" "ความสามารถในการตีความ" "ความปลอดภัย" และ "ความเป็นส่วนตัว" ที่กำหนดแนวทางการพัฒนาผลิตภัณฑ์ AI ทั้งหมดของ Google

SAIF คือเฟรมเวิร์กสำหรับการสร้างแนวทางแบบองค์รวมที่มีมาตรฐานเพื่อผสานรวมมาตรการด้านความปลอดภัยและความเป็นส่วนตัวลงในการใช้แอปพลิเคชันที่ทำงานด้วย ML ซึ่งจะสอดคล้องกับมิติด้าน "ความปลอดภัย" และ "ความเป็นส่วนตัว" ของการสร้าง AI อย่างมีความรับผิดชอบ SAIF ทำให้มั่นใจได้ว่าแอปพลิเคชันที่ทำงานด้วย ML จะพัฒนาขึ้นในลักษณะที่มีความรับผิดชอบ โดยคำนึงถึงสถานการณ์ด้านภัยคุกคามและความคาดหวังของผู้ใช้ที่เปลี่ยนแปลงอยู่เสมอ

Google ทำให้ SAIF เกิดขึ้นจริงได้อย่างไร

ผู้ปฏิบัติงานจะนำเฟรมเวิร์กนี้ไปใช้ได้อย่างไร

จะหาข้อมูลเพิ่มเติมเกี่ยวกับ SAIF ได้ที่ไหน และจะนำไปใช้กับธุรกิจหรือหน่วยงานของฉันได้อย่างไร

สาเหตุที่เราสนับสนุน ชุมชน AI ที่ปลอดภัยสำหรับทุกคน

ในฐานะหนึ่งในบริษัทแรกๆ ที่ทำหลักการด้าน AI ขึ้นมา เราได้กำหนดมาตรฐานสำหรับ AI ที่มีความรับผิดชอบ และหลักการนี้ยังเป็นสิ่งที่นำทางเราในการพัฒนาผลิตภัณฑ์เพื่อความปลอดภัย เราได้สนับสนุนและพัฒนาเฟรมเวิร์กของอุตสาหกรรมเพื่อยกระดับการรักษาความปลอดภัย และได้เรียนรู้ว่าการสร้างชุมชนที่ทำให้งานนี้ก้าวหน้าคือสิ่งสำคัญยิ่งต่อการประสบความสำเร็จในระยะยาว นั่นจึงเป็นเหตุผลที่เราตื่นเต้นที่จะสร้างชุมชน SAIF เพื่อทุกคน

ติดตามเรา

ไซต์ที่เกี่ยวข้อง

ครอบครัว

ศูนย์ความโปร่งใส

Google เกี่ยวกับ Google ผลิตภัณฑ์ของ Google นโยบายความเป็นส่วนตัว ข้อกำหนด ความช่วยเหลือ