AIGC
AI GOVERNANCE CENTER
by ETDA

GENERATIVE AI GOVERNANCE GUIDELINE for Organizations

แนวทางการประยุกต์ใช้ Generative AI อย่างมีธรรมาภิบาลสำหรับองค์กร

กระทรวงดิจิทัล เพื่อเศรษฐกิจและสังคม
ETDA

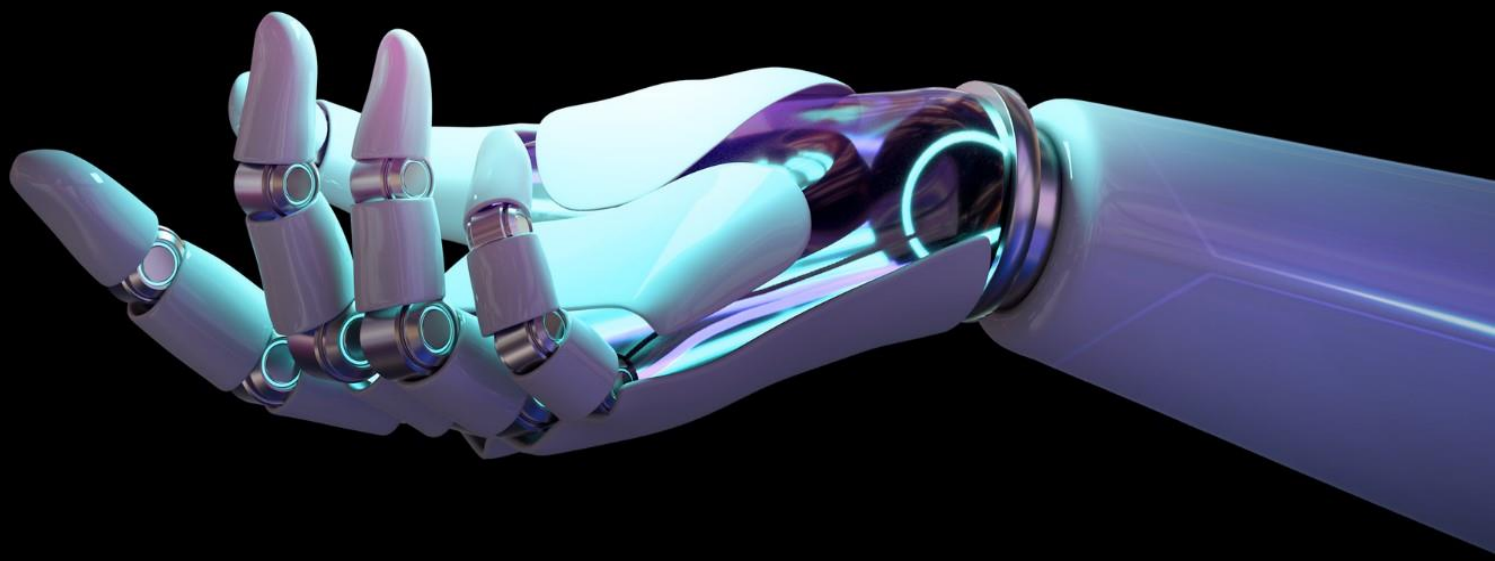
Volume 1 overview
(August 2024)

สารบัญ

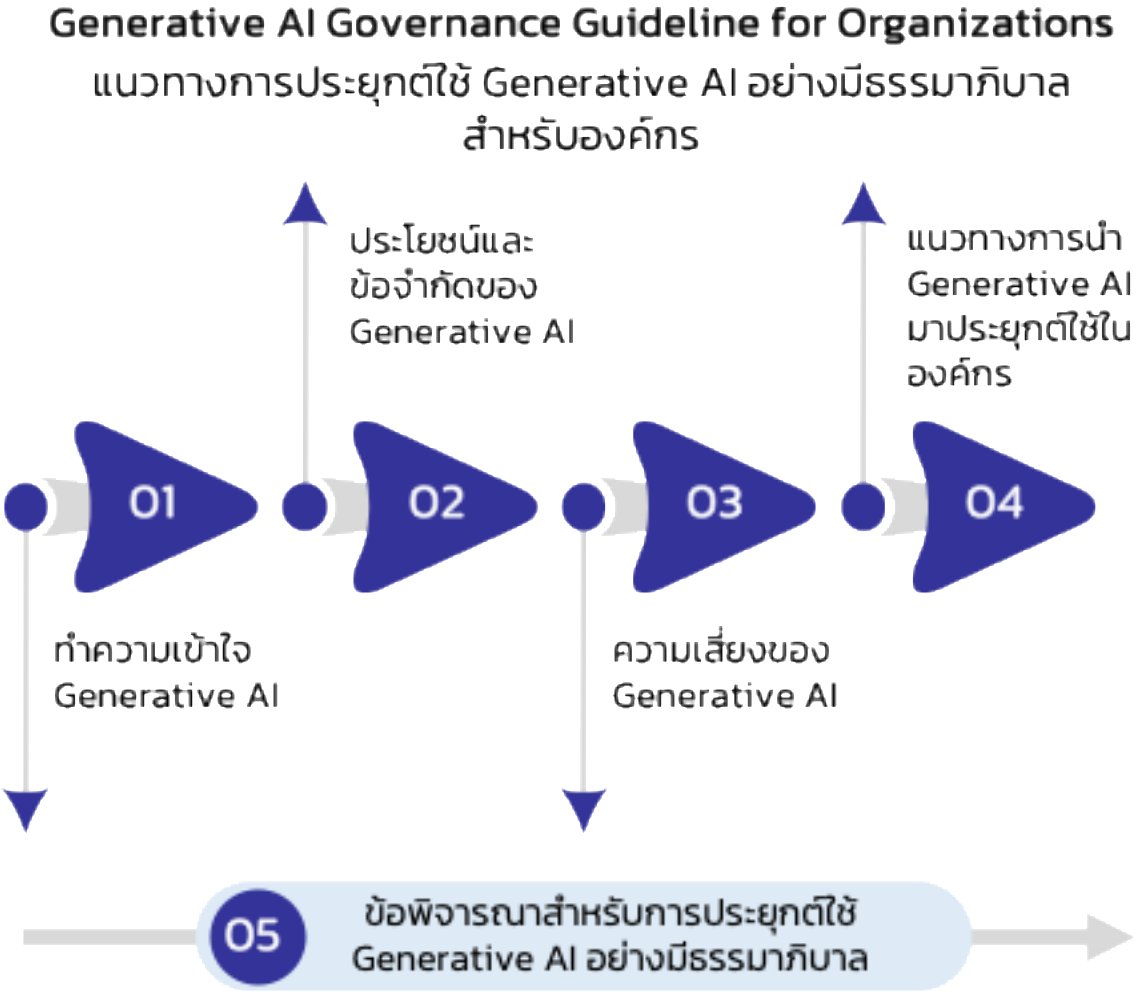
สารบัญ

บทสรุปผู้บริหาร

Executive Summary



Executive Summary



Generative AI เป็นเทคโนโลยีที่สามารถสร้างโอกาสให้กับองค์กรในการบรรลุเป้าหมายแบบก้าวกระโดด และยังช่วยขับเคลื่อนการเปลี่ยนแปลงทางเศรษฐกิจและสังคมในวงกว้าง อย่างไรก็ตาม แม้จะมีการใช้ประโยชน์จาก Generative AI อย่างแพร่หลาย องค์กรยังจำเป็นต้องพิจารณาประเด็นความเสี่ยงที่อาจเกิดขึ้นควบคู่กัน พร้อมวางแนวทางเพื่อการกำกับดูแลการประยุกต์ใช้งาน Generative AI อย่างมีธรรมาภิบาลและเหมาะสมกับบริบทขององค์กร

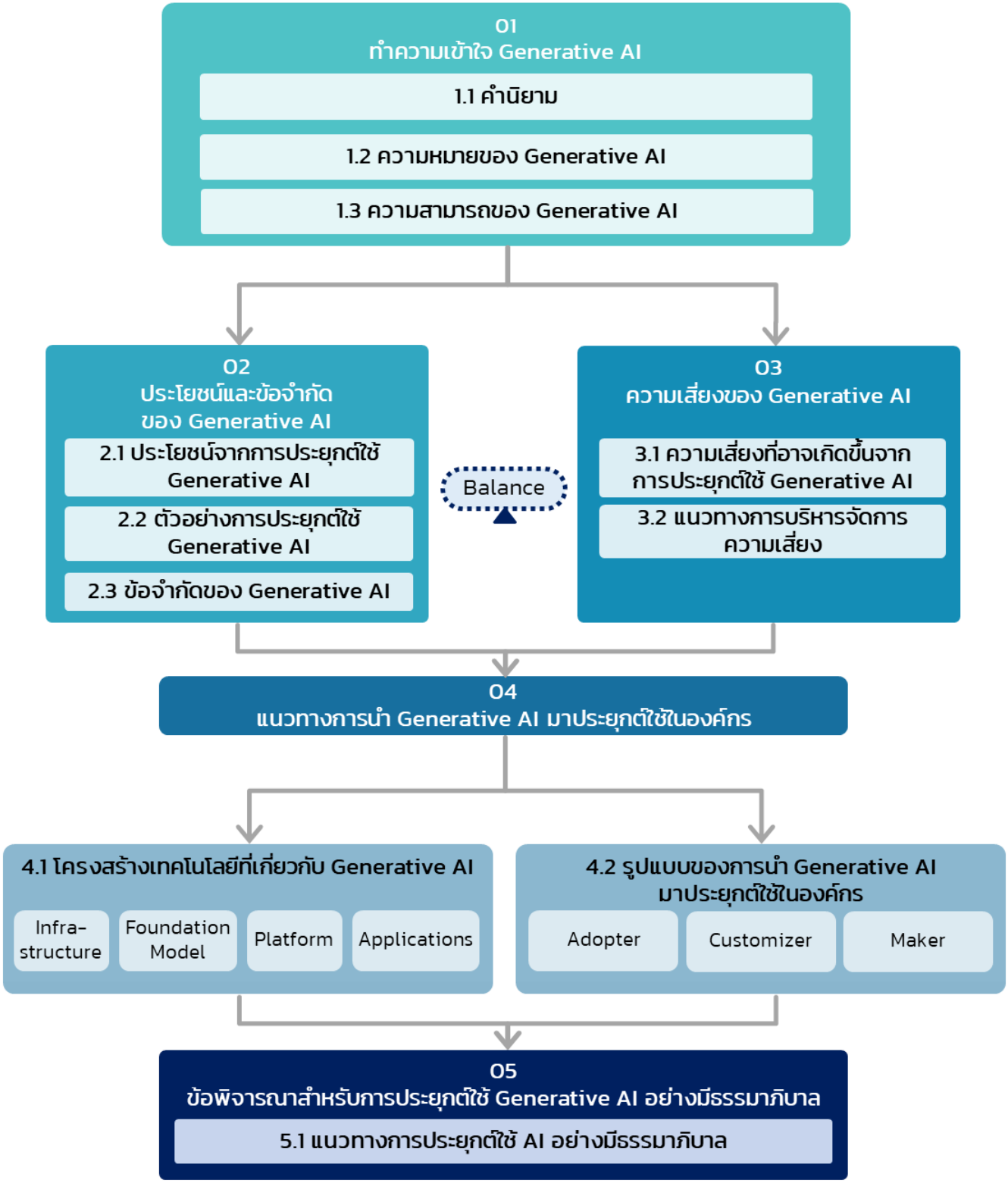
คู่มือแนวทางการประยุกต์ใช้ Generative AI อย่างมีธรรมาภิบาลสำหรับองค์กร (Generative AI Governance Guideline for Organizations) ฉบับนี้ จึงได้จัดทำขึ้นเพื่อเป็นแนวทางสำหรับผู้บริหารและบุคลากรที่เกี่ยวข้องในองค์กรนำไปปรับใช้ให้สอดคล้องกับบริบทขององค์กร โดยมุ่งสร้างความเข้าใจเกี่ยวกับการประยุกต์ใช้ Generative AI อย่างมีธรรมาภิบาล ความสอดคล้องกับกฎหมายและ

ข้อกำหนดที่เกี่ยวข้องพร้อมนำเสนอการวิเคราะห์ประเด็นความเสี่ยงและผลกระทบที่เกี่ยวกับการประยุกต์ใช้ Generative AI โดยเนื้อหาของคู่มือฉบับนี้ประกอบด้วย 5 ส่วนหลัก

1. **ทำความเข้าใจ Generative AI** – การสร้างความเข้าใจพื้นฐานเกี่ยวกับ Generative AI เพื่อให้ผู้ที่เกี่ยวข้องในองค์กรมีความเข้าใจหลักการที่สอดคล้องกันมากที่สุด ตั้งแต่นิยามความหมายและคำศัพท์ที่เกี่ยวข้อง ลักษณะเฉพาะของ Generative AI ที่แตกต่างจาก AI ประเภทอื่น
2. **ประโยชน์และข้อจำกัดของ Generative AI** – การสร้างความเข้าใจถึงศักยภาพและประโยชน์ของ Generative AI และรู้ว่า Generative AI นั้นมีข้อจำกัดใดที่ไม่สามารถทำได้ เพื่อให้องค์กรสามารถประยุกต์ใช้งานอย่างเหมาะสมและสอดคล้องกับเป้าหมายที่กำหนดไว้
3. **ความเสี่ยงของ Generative AI** – การสร้างความเข้าใจเกี่ยวกับประเด็นความเสี่ยงของ Generative AI พร้อมแนวทางบริหารจัดการอย่างเหมาะสมตามบริบทการใช้งานจริง
4. **แนวทางการนำ Generative AI มาประยุกต์ใช้** – การสร้างความเข้าใจเพื่อให้สามารถวางแนวทางการประยุกต์ใช้ Generative AI ในองค์กร และสามารถเลือกวิธีการที่เหมาะสมกับบริบทและความพร้อมองค์กร
5. **ข้อพิจารณาสำหรับการประยุกต์ใช้ Generative AI อย่างมีธรรมาภิบาล** การวางแนวทางการประยุกต์ใช้ Generative AI อย่างมีธรรมาภิบาล โดยอ้างอิงจากแนวทางการประยุกต์ใช้ AI อย่างมีธรรมาภิบาลสำหรับผู้บริหารองค์กร ซึ่งออกโดยศูนย์ธรรมาภิบาลปัญญาประดิษฐ์ (AI Governance Center: AIGC) ภายใต้ สพธอ.

ทั้งนี้ คู่มือแนวทางการประยุกต์ใช้ Generative AI อย่างมีธรรมาภิบาลฉบับนี้ได้ขยายผลจากแนวทางการประยุกต์ใช้ปัญญาประดิษฐ์อย่างมีธรรมาภิบาลสำหรับผู้บริหารองค์กร ซึ่งพัฒนาโดยศูนย์ AIGC และอ้างอิงจากผลการศึกษาและตัวอย่างแนวปฏิบัติทั้งจากในและต่างประเทศ รวมถึงมาตรฐานสากล เพื่อเป็นแนวทางสำหรับองค์กรทั้งรัฐและเอกชนนำไปปรับใช้ตามความเหมาะสม ดังนั้น การไม่ปฏิบัติตามแนวทางคู่มือฉบับนี้หรือการนำไปปรับใช้เพียงแต่บางส่วน จึงไม่เป็นการฝ่าฝืนหรือไม่ปฏิบัติตามกฎหมายและข้อกำหนดที่เกี่ยวข้อง

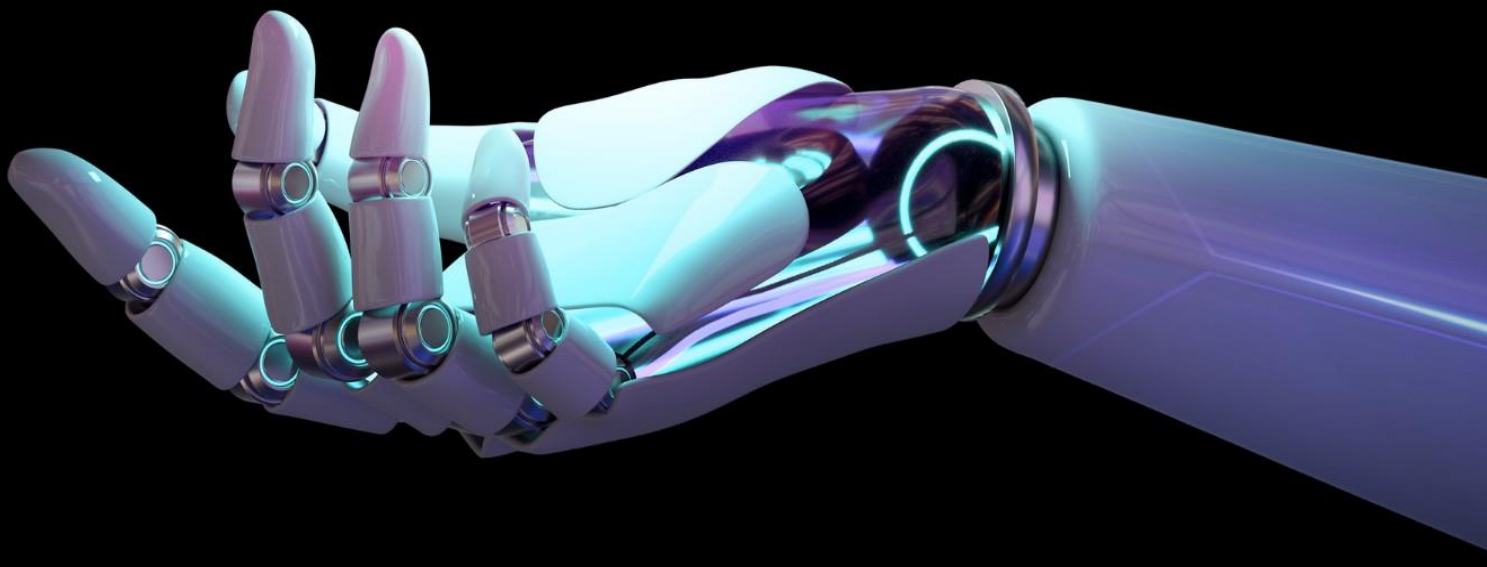
สรุปภาพรวมแนวทางการประยุกต์ใช้ Generative AI อย่างมีธรรมาภิบาล



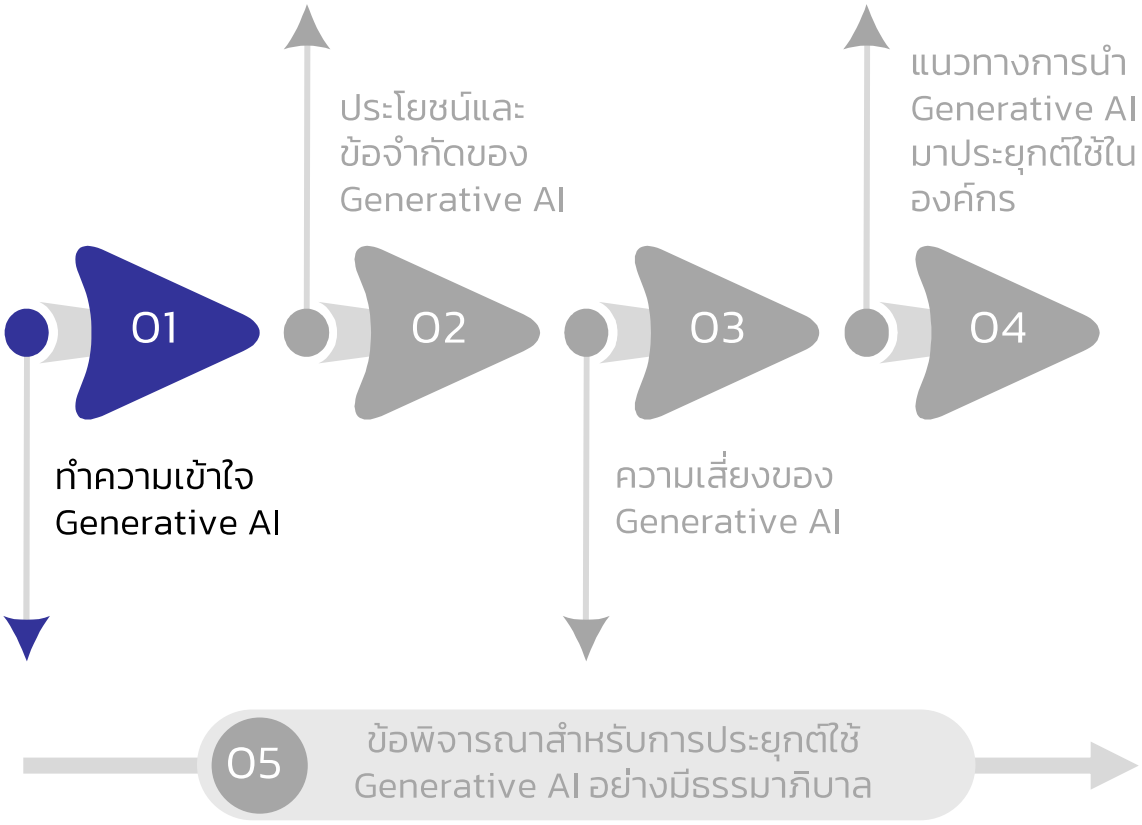
01

ทำความเข้าใจ Generative AI

What is Generative AI?



Generative AI Governance Guideline
แนวทางการประยุกต์ใช้ Generative AI อย่างมีธรรมาภิบาล สำหรับองค์กร



การทำความเข้าใจความหมายของ Generative AI จะช่วยให้ผู้บริหารและผู้ที่เกี่ยวข้องในองค์กร สามารถวางกรอบแนวทางการกำกับดูแลการประยุกต์ใช้ Generative AI ได้ชัดเจนและสอดคล้องกับเป้าหมายที่กำหนดไว้

โดยเนื้อหาในบทนี้จะครอบคลุมคำนิยามของ Generative AI คำนิยามอื่นที่มีความเกี่ยวข้อง รวมถึงความแตกต่างของ Generative AI เมื่อเทียบกับ AI ประเภทอื่น ๆ

1.1 คำนิยาม

Artificial Intelligence (AI) ปัญญาประดิษฐ์เป็นเทคโนโลยีที่ถูกพัฒนาขึ้นเพื่อให้คอมพิวเตอร์มีคุณสมบัติหรือพฤติกรรมใกล้เคียงมนุษย์ เช่น การเรียนรู้ การรับรู้ และตอบสนองต่อสภาพแวดล้อม การให้เหตุผล และการแก้ไขปัญหา เป็นต้น ตามวัตถุประสงค์ที่มนุษย์กำหนด

Machine Learning (ML): เทคโนโลยี AI ประเภทหนึ่งที่ทำงานหรือสร้างผลลัพธ์บนพื้นฐานของข้อมูลที่ได้รับจากการฝึกฝนหรือจากสภาพแวดล้อม

Deep Learning (DL): Machine Learning ประเภทหนึ่งที่ประมวลผลผ่านโครงข่ายประสาทเทียม (Artificial Neural Network: ANN) จำนวนหลายชั้น (Layer) ที่ถูกสร้างขึ้นจากข้อมูลที่ได้รับการฝึกฝน เพื่อให้สามารถทำงานหรือสร้างผลลัพธ์ที่มีประสิทธิภาพดียิ่งขึ้น

Artificial Neural Network (ANN): โครงข่ายของเซลล์ประสาทเทียม (Artificial Neuron) ที่คล้ายกับการเชื่อมต่อเซลล์ประสาท (Neuron) ในสมองมนุษย์ โดยในแต่ละเซลล์ประสาทเทียมนั้น มีหน้าที่ในการรับข้อมูลและนำไปประมวลผลเพื่อสร้างเป็นผลลัพธ์ จากนั้นจึงส่งต่อผลลัพธ์ไปยังเซลล์ประสาทเทียมในชั้น (Layer) ถัดไปเพื่อประมวลผลต่อ

Generative AI: เทคโนโลยี AI ประเภทหนึ่งที่มีความสามารถในการสร้างเนื้อหาใหม่ในหลากหลายรูปแบบ เช่น ข้อความ ภาพ วิดีโอ ซอร์สโค้ด หรือรูปแบบอื่น เป็นต้น ตามข้อความหรือคำสั่ง (Prompt) ที่มนุษย์เป็นผู้กำหนด

Prompt Engineering: การสร้างและปรับแต่งข้อความหรือคำสั่ง เพื่อให้ Generative AI สร้างผลลัพธ์ (Output) ที่ดีที่สุดและตรงตามความต้องการ

Foundation Model: โมเดล AI ประเภท Generative AI ที่ได้รับการฝึกฝนด้วยข้อมูลขนาดใหญ่ โดยมีวัตถุประสงค์เพื่อให้สามารถสร้างเนื้อหาใหม่ที่คล้ายคลึงกับข้อมูลที่ได้รับการฝึกฝน

Large Language Model (LLM) โมเดลภาษาขนาดใหญ่ที่รับข้อความหรือคำสั่ง (Input) ในรูปแบบภาษา และนำไปสร้างผลลัพธ์ (Output) ที่มีความสามารถในด้าน

ภาษาที่หลากหลาย เช่น การสร้างข้อความใหม่ การแปลภาษา การสรุปความ การวิเคราะห์ข้อความ เป็นต้น

1.2 ความหมายของ Generative AI

Generative AI เป็น AI ประเภทหนึ่งที่สามารถสร้างเนื้อหาได้หลากหลายรูปแบบ เช่น ข้อความ ภาพ วิดีโอ ซอร์สโค้ด หรือรูปแบบอื่น เป็นต้น ด้วยการสั่งการผ่านข้อความหรือคำสั่ง (Prompt) ที่มนุษย์เป็นผู้กำหนด

เมื่อ Generative AI ได้รับ Prompt จากผู้ใช้งานแล้ว Generative AI จะทำการสร้างเนื้อหาใหม่ที่คล้ายคลึงกับข้อมูลที่ได้รับการฝึกฝนมา โดยเลือกนำเสนอเนื้อหาที่สอดคล้องและเหมาะสมกับ Prompt ที่ได้รับมามากที่สุด โดยพิจารณาจากหลักการความน่าจะเป็น

จากความสามารถและกระบวนการทำงานของ Generative AI ข้างต้น จึงทำให้ Generative AI มีความแตกต่างจาก AI แบบเดิม (Traditional AI) กล่าวคือ Traditional AI ถูกออกแบบมาเพื่อคาดการณ์ (Prediction) การตัดสินใจ (Decision) หรือการให้คำแนะนำ (Recommendation) บนพื้นฐานของข้อมูลที่ได้รับการฝึกฝน (Train) มาก่อน ตัวอย่างเช่น การใช้ AI ทำนายยอดขาย การใช้ AI วินิจฉัยโรคปอดจากภาพเอกซเรย์ปอด และการใช้ AI แนะนำสินค้าและบริการ เป็นต้น แต่ AI แบบนี้ไม่ได้ถูกออกแบบมาเพื่อสร้างเนื้อหาใหม่ที่มีความคล้ายคลึงกับข้อมูลต้นฉบับ

จากรูปที่แสดงลำดับถัดไปจะพบว่า Generative AI เป็น AI ประเภทหนึ่งของ Deep Learning เนื่องจาก Generative AI สร้างเนื้อหาใหม่บนพื้นฐานของข้อมูลที่ได้รับการฝึกฝนมา โดยทำการประมวลผลเพื่อสร้างเนื้อหาใหม่ผ่านโครงข่ายประสาทเทียม (Artificial Neural Network) ที่มีความซับซ้อน

Artificial Intelligence (AI)

เทคโนโลยีที่ถูกพัฒนาขึ้นเพื่อให้คอมพิวเตอร์มีคุณสมบัติหรือพฤติกรรมใกล้เคียงมนุษย์ เช่น การรับรู้และตอบสนองต่อสภาพแวดล้อม การให้เหตุผล และการแก้ไขปัญหา เป็นต้น

Machine Learning (ML)

AI ที่ทำงานหรือสร้างผลลัพธ์ เช่น คาดการณ์ ตัดสินใจ เป็นต้น บนพื้นฐานข้อมูลที่ได้รับจากการฝึกฝนหรือจากสภาพแวดล้อม มากกว่าการทำงานตามโปรแกรมที่มนุษย์กำหนด

Deep Learning (DL)

AI ประเภท Machine Learning ที่ประมวลผลข้อมูลขนาดใหญ่ผ่านโครงข่ายประสาทเทียม (Artificial Neural Network: ANN) ซึ่งเลียนแบบการทำงานจากสมองของมนุษย์

Generative AI

AI ประเภท Deep Leaning ที่มีความสามารถในการสร้างสรรค์เนื้อหาใหม่ในหลากหลายรูปแบบ ทั้งข้อความ ภาพ วิดีโอ หรือรูปแบบอื่น ๆ

1.3 ความสามารถของ Generative AI

Generative AI เป็นเทคโนโลยีที่มีความสามารถสร้างผลลัพธ์ในหลายรูปแบบ อาทิ

1. ข้อความ (Text) – ความสามารถในการสร้างและปรับปรุงข้อความในหลายรูปแบบ เช่น การเขียนบทความ การสร้างคำตอบ การปรับปรุงข้อความ และอื่น ๆ ตัวอย่างเครื่องมือ เช่น ChatGPT, Claude, Copilot, Gemini เป็นต้น ซึ่งเป็นตัวอย่างของ Generative AI ประเภท Large Language Model

2. ซอร์สโค้ด (Source Code) – ความสามารถในการสร้างและปรับปรุง Code โปรแกรมในหลายภาษาโปรแกรม ตัวอย่างเครื่องมือเช่น ChatGPT, Claude, Copilot, Gemini เป็นต้น

3. รูปภาพ (Image) – ความสามารถในการสร้างและปรับปรุงรูปภาพต่าง ๆ ได้ ไม่ว่าจะเป็นการออกแบบภาพใหม่ การปรับแต่งภาพที่มีอยู่แล้ว หรือการสร้างภาพจากข้อความ ตัวอย่างเครื่องมือเช่น DALL-E, Midjourney เป็นต้น

4. เสียง (Audio) – ความสามารถในการสร้างและปรับปรุงเสียง เช่น การสร้างเสียงพูด การปรับแต่งเสียง การสร้างดนตรี และอื่น ๆ ตัวอย่างเครื่องมือ เช่น Suno และ Udio เป็นต้น

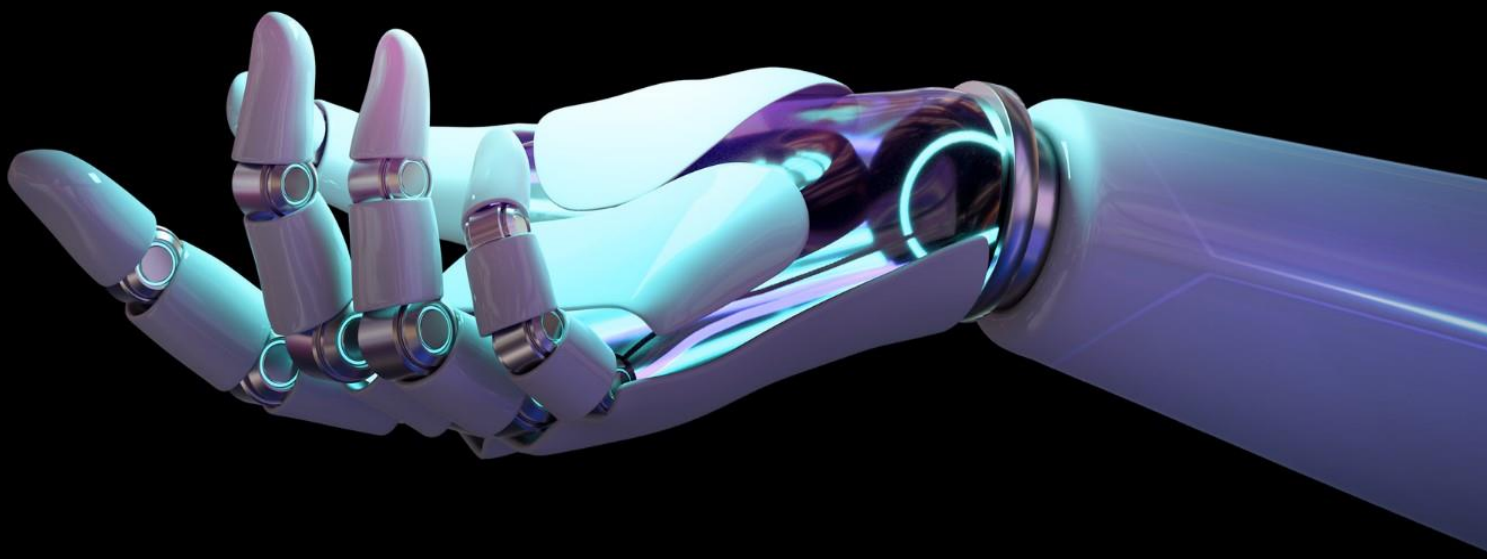
5. วิดีโอ (Video) – ความสามารถในการสร้างและปรับปรุงวิดีโอ เช่น การตัดต่อวิดีโอ การสร้างวิดีโอจากข้อความ การสร้างวิดีโอแบบแอนิเมชัน และอื่น ๆ ตัวอย่างเครื่องมือเช่น Runway, Adobe Firefly เป็นต้น

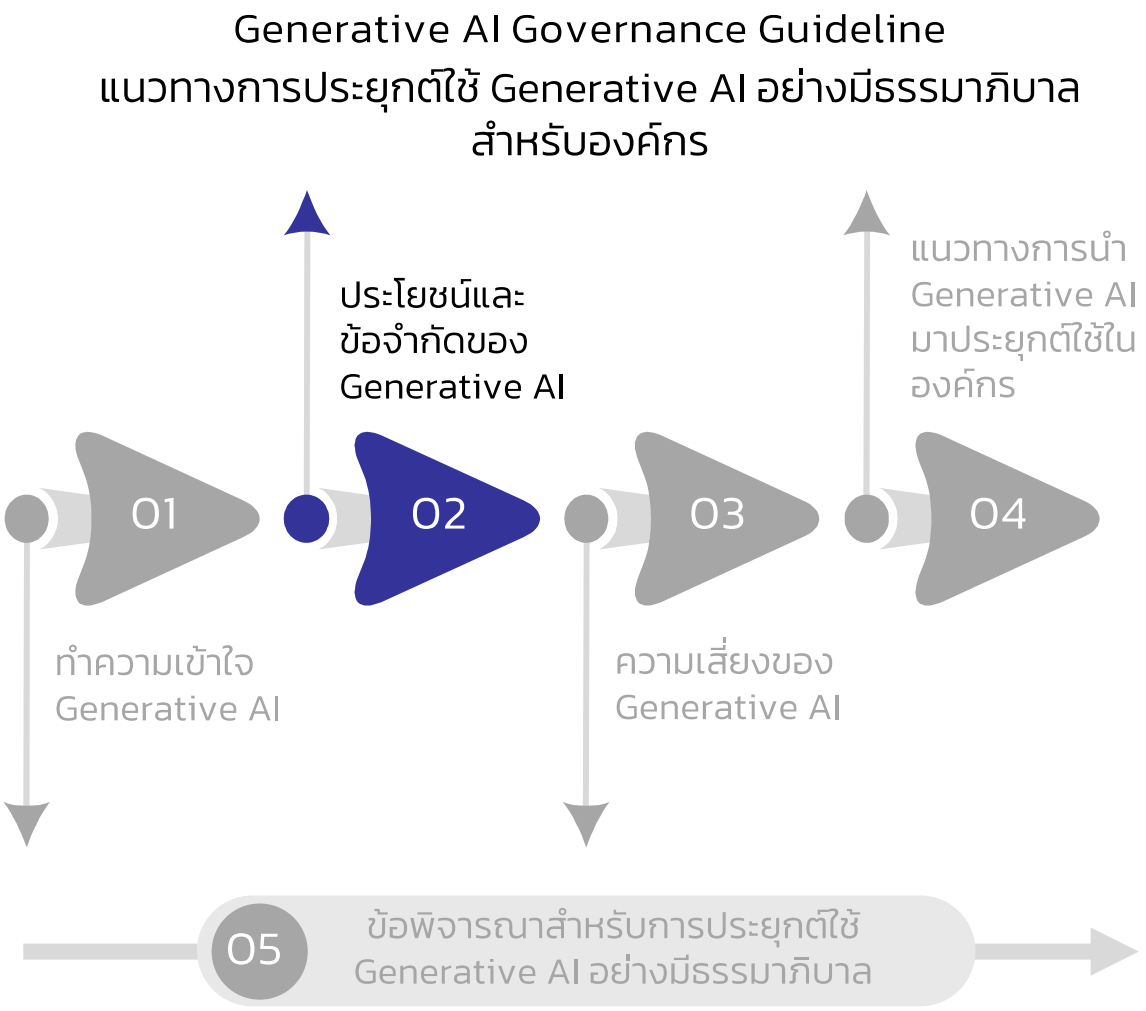
6. โมเดล 3 มิติ (3D Model) – ความสามารถในการสร้างและปรับปรุงโมเดล 3 มิติได้ ไม่ว่าจะเป็นการสร้างโมเดล 3 มิติใหม่ การปรับแต่งโมเดลที่มีอยู่แล้ว หรือการสร้างโมเดลจากข้อความ

02

ประโยชน์และข้อจำกัดของ Generative AI

Benefits and Limitations of Generative AI

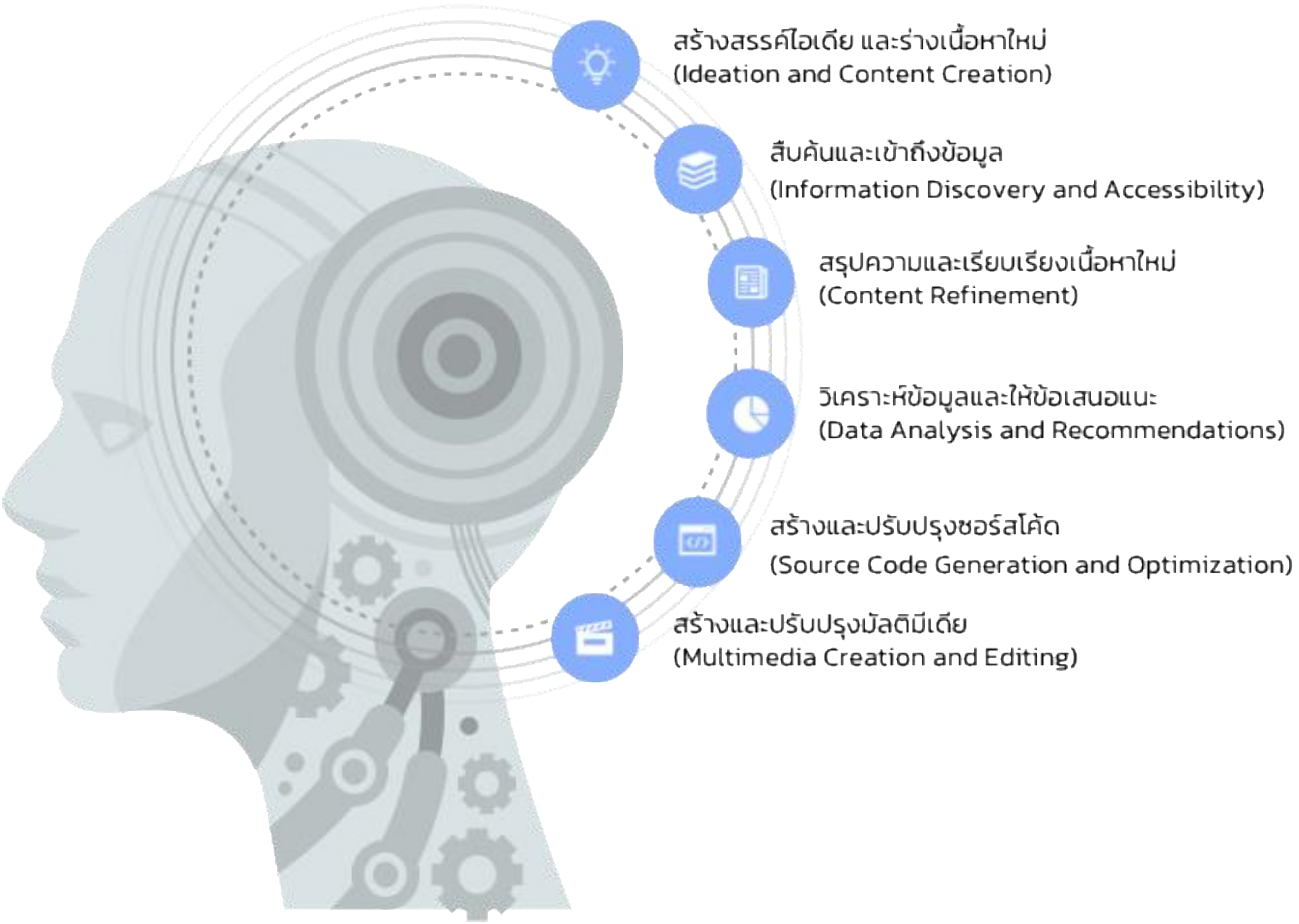




เมื่อทำความเข้าใจความสามารถของ Generative AI แล้ว จะช่วยให้ผู้ที่เกี่ยวข้องในองค์กรสามารถกำหนดแนวทางการนำ Generative AI ไปใช้ประโยชน์ได้อย่างสอดคล้องกับเป้าหมายขององค์กร รวมถึงการวางแผนแนวทางการกำกับดูแลการประยุกต์ใช้ Generative AI ในองค์กรอย่างเหมาะสม

2.1 ประโยชน์จากการประยุกต์ใช้ Generative AI

Generative AI มีความสามารถที่โดดเด่นในการสร้างสรรค์สิ่งใหม่ ช่วยเพิ่มประสิทธิภาพกระบวนงานภายในองค์กรที่อาจมีข้อจำกัด ดังนั้น องค์กรจึงควรเริ่มจากการทำความเข้าใจศักยภาพของเทคโนโลยีนี้ เพื่อให้สามารถนำไปประยุกต์ใช้ให้เกิดประโยชน์อย่างเหมาะสม


สร้างสรรค์ไอเดีย และร่างเนื้อหาใหม่
(Ideation and Content Creation)
สืบค้นและเข้าถึงข้อมูล
(Information Discovery and Accessibility)
สรุปความและเรียบเรียงเนื้อหาใหม่
(Content Refinement)
วิเคราะห์ข้อมูลและให้ข้อเสนอแนะ
(Data Analysis and Recommendations)
สร้างและปรับปรุงซอร์สโค้ด
(Source Code Generation and Optimization)
สร้างและปรับปรุงมัลติมีเดีย
(Multimedia Creation and Editing)

2.2 ตัวอย่างการประยุกต์ใช้ Generative AI

	คำอธิบาย	ตัวอย่าง
 สร้างสรรค์ไอเดียและร่างเนื้อหาใหม่ (Ideation and Content Creation)	ช่วยคิดไอเดียและร่างเนื้อหาใหม่ ซึ่งช่วยลดระยะเวลาในการสร้างหรือผลิตเนื้อหา	ฝ่าย HR ใช้ในการร่าง Job Description หรือคำถามสัมภาษณ์เพื่อใช้ถามผู้สมัครงาน
 สืบค้นและเข้าถึงข้อมูล (Information Discovery and Accessibility)	ช่วยดึงข้อมูลที่เกี่ยวข้องจากคลังข้อมูลขนาดใหญ่อย่างมีประสิทธิภาพช่วยให้กระบวนการทำงานราบรื่นและประหยัดเวลา	ฝ่ายขายใช้ช่วยในการค้นหาข้อมูลที่เกี่ยวข้องกับการคาดการณ์แนวโน้มตลาด
 สรุปความและเรียบเรียงเนื้อหาใหม่ (Content Refinement)	ช่วยรีวิวและปรับปรุงร่างเนื้อหา ทั้งในแง่ของการปรับรูปแบบการเขียน การสรุปใจความสำคัญ และการให้ข้อเสนอแนะ	ฝ่ายการตลาดใช้ในการปรับปรุงการเขียนข้อความแคมเปญโฆษณาที่เคยร่างไว้ให้เหมาะกับกลุ่มเป้าหมายที่มีความเฉพาะ

	คำอธิบาย	ตัวอย่าง
วิเคราะห์ข้อมูลและให้ข้อเสนอแนะ (Data Analysis and Recommendations)	ช่วยหารูปแบบ วิเคราะห์ข้อมูลที่ซับซ้อน และให้คำแนะนำที่เหมาะสมเพื่อสนับสนุนการตัดสินใจโดยใช้ข้อมูลเป็นพื้นฐาน	ฝ่ายการเงินใช้ช่วยในการวิเคราะห์ข้อมูลงบการเงินย้อนหลัง 10 ปี เพื่อนำเสนอผู้บริหาร
สร้างและปรับปรุงซอร์สโค้ด (Source Code Generation and Optimization)	ช่วยเขียนและปรับปรุงภาษาโปรแกรมต่าง ๆ เพื่อสนับสนุนกระบวนการพัฒนาของนักพัฒนา	ฝ่ายไอทีใช้ในการพัฒนาแอปพลิเคชันและออกแบบ User Interface
สร้างและปรับปรุงมัลติมีเดีย (Multimedia Creation and Editing)	ช่วยสร้างและปรับปรุงมัลติมีเดีย เช่น รูปภาพ เสียง และวิดีโอ เป็นต้น	ฝ่าย Graphic Design ใช้ในการสร้างคลิปวิดีโอประชาสัมพันธ์องค์กรที่ไม่มีการถ่ายทำจริง

2.3 ข้อจำกัดของ Generative AI

เพื่อให้องค์กรสามารถเลือกนำ Generative AI ไปประยุกต์ใช้งานได้อย่างเหมาะสมจึงควรเข้าใจข้อจำกัดของเทคโนโลยีนี้ พร้อมทั้งประเมินความเหมาะสมในการประยุกต์ใช้ และวางแผนรับมือกับข้อจำกัดที่อาจเกิดขึ้น

ข้อจำกัดและตัวอย่าง เมื่อนำ Generative AI มาประยุกต์ใช้

อาการหลอน (Hallucination หรือ Confabulation)

คำอธิบาย

Generative AI สามารถสร้างคำตอบที่ดูเหมือนมีเหตุผล แต่ไม่ถูกต้องตามข้อเท็จจริง

ตัวอย่าง: ผู้ใช้งานได้ใช้ Generative AI เพื่อขอรับคำแนะนำการลงทุนในหุ้นตามเงื่อนไขที่ต้องการ ซึ่งได้ผลลัพธ์ออกมาน่าเชื่อถือมาก แต่กลับพบว่าคำแนะนำของ Generative AI นั้นเป็นการเสนอให้ลงทุนในหุ้นบริษัทที่ไม่มีอยู่จริง

การคิดวิเคราะห์และการตัดสินใจ (Critical Thinking and Judgement)

คำอธิบาย

เนื่องจาก Generative AI ประมวลผลเพื่อสร้างข้อความโดยใช้หลักการความน่าจะเป็นของคำถัดไป จึงอาจทำให้ได้ข้อสรุปที่ไม่ถูกต้องหรือไม่สมเหตุสมผล

ตัวอย่าง: เจ้าหน้าที่การตลาดขอคำแนะนำจาก Generative AI เพื่อวางแผนการตลาดเครื่องดื่มสุขภาพ Generative AI แนะนำให้ใช้กลยุทธ์การแจกผลิตภัณฑ์ฟรีในโรงเรียนเพราะพบว่ากลยุทธ์นี้ประสบความสำเร็จในอุตสาหกรรมอื่น แต่ก็ไม่ได้คำนึงถึงว่ากลุ่มเป้าหมายของบริษัทรวมถึงผู้สูงอายุ คนทำงาน

ข้อจำกัดและตัวอย่าง เมื่อนำ Generative AI มาประยุกต์ใช้

บริบทที่ละเอียดอ่อน หรือประเด็นทางจริยธรรม (Sensitive or Ethical Context)

คำอธิบาย

Generative AI สามารถสร้างเนื้อหาที่ไม่เหมาะสมตามหลักจริยธรรม มีความเอนเอียง หรือเนื้อหาที่นำไปสู่การเลือกปฏิบัติ

ตัวอย่าง: ผู้ใช้งานขอคำแนะนำจาก Generative AI เกี่ยวกับการเลือกอาชีพ โดยได้รับคำแนะนำให้เพศชายเลือกอาชีพวิศวกร และให้ข้อมูลว่ามีค่าตอบแทนสูง ขณะที่แนะนำให้เพศหญิงเลือกอาชีพพยาบาลและให้ข้อมูลว่ามีความเหมาะสมดีแล้ว

ความเชี่ยวชาญเฉพาะด้าน (Domain Expertise)

คำอธิบาย

ผลลัพธ์ของ Generative AI ไม่สามารถใช้แทนข้อแนะนำหรือข้อคิดเห็นจากผู้เชี่ยวชาญได้ โดยเฉพาะในด้านกฎหมาย การแพทย์ หรือด้านอื่น ๆ ที่ต้องการข้อมูลที่ถูกต้องแม่นยำและคำนึงถึงบริบทที่มีความเกี่ยวข้อง

ตัวอย่าง: ผู้ป่วยรายหนึ่งขอคำแนะนำการรักษาจาก Generative AI สำหรับอาการปวดท้อง ซึ่ง Generative AI แนะนำให้ใช้ยาลดกรดเพราะเป็นวิธีที่ได้ผลในอาการคล้ายกับที่ผู้ป่วยสอบถาม แต่ในความเป็นจริง ผู้ป่วยรายนี้มีปัญหาเกี่ยวกับตับอ่อน การใช้ยาลดกรดจึงไม่เหมาะสมและอาจทำให้อาการแย่ลง

ข้อจำกัดและตัวอย่าง เมื่อนำ Generative AI มาประยุกต์ใช้

ประสบการณ์และบริบทเฉพาะบุคคล (Personal Experience and Context)

คำอธิบาย

ถึงแม้ผลลัพธ์จาก Generative AI อาจดูเหมือนสร้างมาจากมนุษย์ แต่แท้จริง Generative AI ยังขาดการมีประสบการณ์และอารมณ์ความรู้สึกเหมือนมนุษย์

ตัวอย่าง: ผู้ใช้งานขอคำแนะนำจาก Generative AI เกี่ยวกับการศึกษาต่อ โดย Generative AI แนะนำให้ผู้ใช้งานศึกษาวิศวกรรมศาสตร์เพราะมีโอกาสในการทำงานและค่าตอบแทนสูง แต่ผู้ใช้งานมีความสนใจและความสามารถทางด้านศิลปะและออกแบบ คำแนะนำนี้จึงไม่สอดคล้องกับความสนใจและความสามารถของผู้ใช้งาน

ความเป็นปัจจุบันของข้อมูล (Dynamic Real-time Information Retrieval)

คำอธิบาย

ข้อมูลผลลัพธ์จาก Generative AI อาจยังไม่รวมข้อมูลจากอินเทอร์เน็ต หรือไม่สามารถเข้าถึงข้อมูลที่อยู่นอกชุดข้อมูลที่ใช้ในการฝึกฝนโมเดลแบบ Real-Time

หมายเหตุ : ในปัจจุบันผลิตภัณฑ์ LLM ต่าง ๆ เช่น ChatGPT Gemini และ Bing ได้มีการปรับให้ผลลัพธ์ที่แสดงรวมการเข้าถึงข้อมูลจากอินเทอร์เน็ต

ตัวอย่าง: ผู้ใช้งานขอคำแนะนำจาก Generative AI เกี่ยวกับข้อมูลด้านกฎหมาย ซึ่งได้รับคำตอบที่น่าเชื่อถือแต่เป็นข้อมูลที่ไม่เป็นปัจจุบันจากอินเทอร์เน็ต เช่น กฎหมายที่ไม่มีผลบังคับใช้อีกต่อไป ทำให้คำแนะนำที่ได้รับไม่ถูกต้องตามสถานการณ์ปัจจุบัน

ข้อจำกัดและตัวอย่าง เมื่อนำ Generative AI มาประยุกต์ใช้

การให้เหตุผลเกี่ยวกับผลลัพธ์ (Explainability)

คำอธิบาย

การอธิบายกระบวนการทำงานภายในโมเดล Generative AI เป็นเรื่องยาก เนื่องจากโมเดลขึ้นอยู่กับโครงข่ายประสาทเทียม (Neural Network) ที่เรียกว่า Black Box ซึ่งอาจส่งผลกระทบหากต้องมีการชี้แจงเหตุผลของผลลัพธ์ที่ได้มาจากโมเดล

ตัวอย่าง: อาจารย์ใช้ Generative AI ตรวจข้อสอบการเขียนบทความของนักเรียน โดยมีการให้คะแนนและ feedback กลับมาสำหรับแต่ละคน แต่ไม่สามารถอธิบายได้ว่าเพราะเหตุใดจึงให้คะแนนเท่านั้น

ความไม่แน่นอนของผลลัพธ์ (Consistent Output)

คำอธิบาย

ผลลัพธ์ของ Generative AI ไม่คงที่ แม้ว่าจะป้อนข้อมูลเข้าไปแบบเดียวกัน แต่อาจได้คำตอบที่ต่างกัน

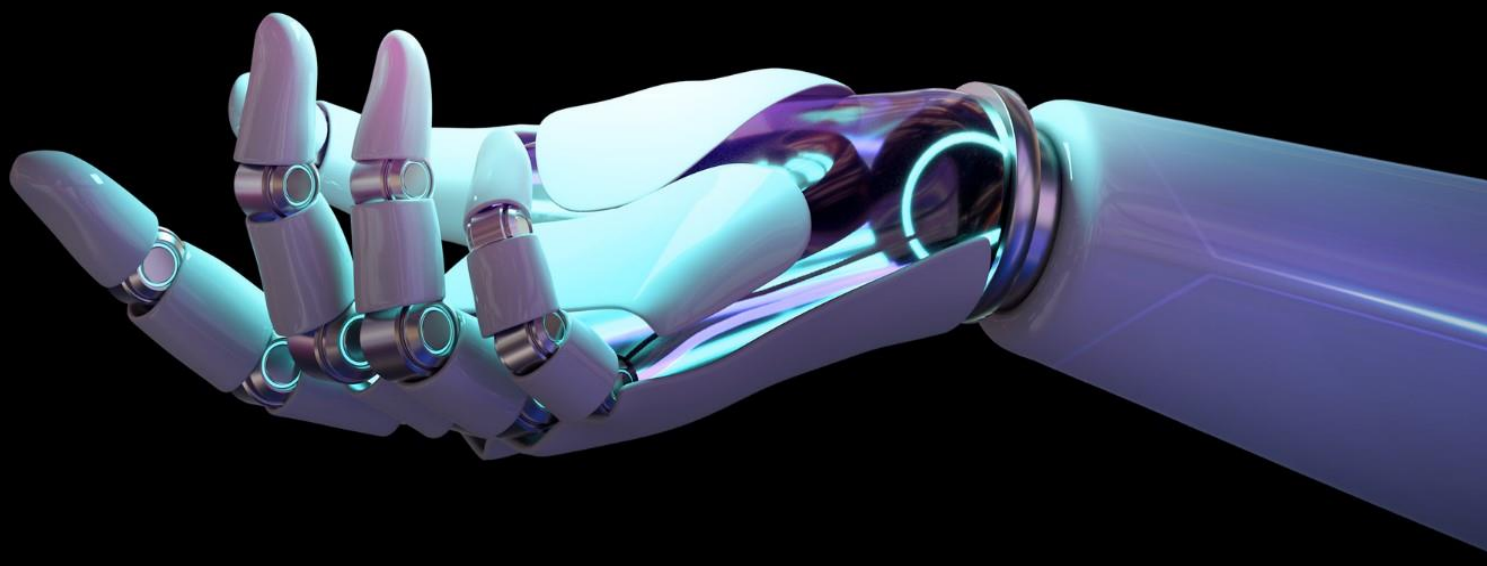
ตัวอย่าง: บริษัทใช้ Generative AI เพื่อตอบคำถามผู้สนใจสมัครงาน สำหรับตำแหน่งเดียวกัน โดย Generative AI ให้คำอธิบายที่แตกต่างกันแก่ผู้สมัครหลายคน ทั้งที่เป็นตำแหน่งงานเดียวกัน ทำให้ผู้สมัครเกิดความสับสน เข้าใจลักษณะงานที่ไม่ถูกต้อง

จากข้อจำกัดในการนำ Generative AI มาประยุกต์ใช้ข้างต้น องค์กรจึงควรคำนึงถึงประเด็นด้านต่าง ๆ รวมถึงความสอดคล้องกับกฎระเบียบหรือข้อบังคับที่เกี่ยวข้อง นอกจากนี้ ควรเพิ่มการมีส่วนร่วมของมนุษย์ในกระบวนการทำงานร่วมกับ Generative AI และหมั่นอัปเดตการเปลี่ยนแปลงทางด้านเทคโนโลยีเพื่อให้มั่นใจว่าการประยุกต์ใช้มีความสอดคล้องกับความสามารถของเทคโนโลยีที่มีการพัฒนาอย่างต่อเนื่อง ทั้งนี้ เพื่อเป็นการเตรียมพร้อมรับมือกับข้อจำกัดต่าง ๆ ของ Generative AI ได้อย่างมีประสิทธิภาพ

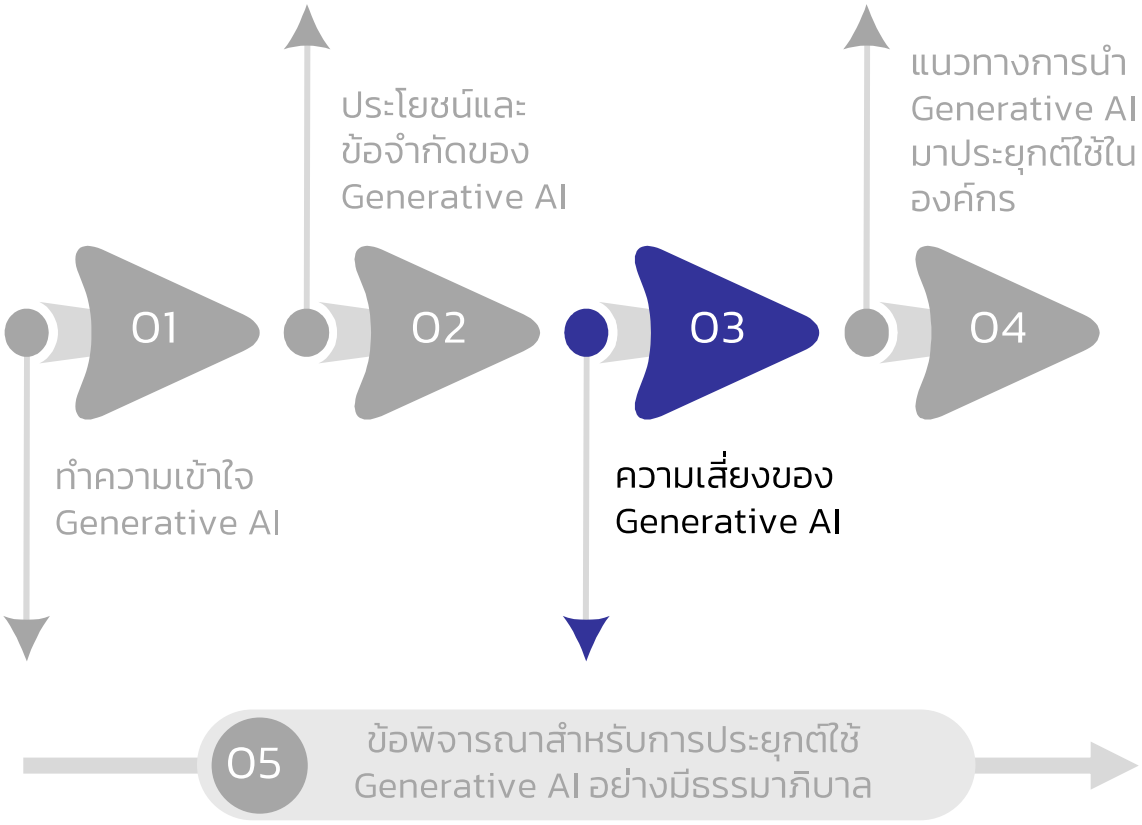
03

ความเสี่ยงของ Generative AI

Risks of Generative AI



Generative AI Governance Guideline
แนวทางการประยุกต์ใช้ Generative AI อย่างมีธรรมาภิบาล สำหรับองค์กร



การประยุกต์ใช้ Generative AI ให้มีประสิทธิภาพนั้น ผู้บริหารและผู้ที่เกี่ยวข้องในองค์กรควรเข้าใจถึงข้อจำกัดและความเสี่ยงของเทคโนโลยีนี้ เพื่อกำหนดความคาดหวังในการนำ Generative AI ไปประยุกต์ใช้ได้อย่างสอดคล้องกับบริบทและสถานการณ์จริง

3.1 ความเสี่ยงที่อาจเกิดขึ้นจากการประยุกต์ใช้ Generative AI

การทำความเข้าใจประเด็นความเสี่ยงที่อาจเกิดขึ้นเมื่อนำเทคโนโลยี Generative AI มาประยุกต์ใช้ในองค์กรจะช่วยให้องค์กรหาสมดุลระหว่างประโยชน์และความเสี่ยง ทำให้สามารถตัดสินใจนำ Generative AI มาประยุกต์ใช้งานได้อย่างเหมาะสม

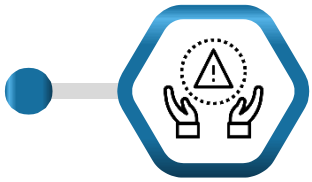
Generative AI นำมาซึ่งความเสี่ยงรูปแบบใหม่ ดังนั้น องค์กรจึงควรมีวิธีการวิเคราะห์และจัดการความเสี่ยงเพิ่มเติมอย่างเหมาะสม โดยจากเอกสาร "Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile" ของ National Institute of Standards and Technology (NIST) ซึ่งมีการระบุประเด็นความเสี่ยงที่ควรให้ความสำคัญ ซึ่งประกอบด้วย

1) ความเสี่ยงด้านข้อมูลที่เป็นอันตรายต่อการผลิตอาวุธเคมี รังสี หรือนิวเคลียร์ (Chemical, Biological, Radiological, or Nuclear (CBRN) Weapons)

Generative AI อาจถูกใช้เป็นเครื่องมือในการสร้างเนื้อหาที่เกี่ยวข้องกับการผลิตอาวุธเคมี รังสี หรือนิวเคลียร์ และอาจถูกนำไปใช้ในทางที่ไม่เหมาะสม

2) ความเสี่ยงด้านเนื้อหาที่น่าเชื่อถือแต่ไม่ถูกต้อง (Confabulation)

การผลิตเนื้อหาที่ดูน่าเชื่อถือ แต่ไม่ถูกต้องตามข้อเท็จจริง อาจถูกเรียกว่า "Hallucination" หรือ "Fabrication" โดย Generative AI อาจสร้างผลลัพธ์ที่ผิดไปจากข้อเท็จจริง หรือขัดแย้งกับข้อความที่สร้างขึ้นก่อนหน้าทั้งที่อยู่ในบริบทเดียวกัน เช่น สับสนเกี่ยวกับบุคคล สถานที่ หรือรายละเอียดเหตุการณ์ทางประวัติศาสตร์ เป็นต้น ซึ่งอาจทำให้เกิดการนำเนื้อหาที่ผิดไปใช้งาน



3) ความเสี่ยงด้านเนื้อหาอันตรายหรือรุนแรง (Dangerous or Violent Recommendations)

Generative AI อาจให้คำแนะนำที่ยั่วยุ ปลุกปั่น หรือคุกคาม ที่จะนำไปสู่ความรุนแรง โดยอาจ สร้างภาพ, วิดีโอ หรือเสียง เพื่อให้เกิดความเข้าใจผิดในตัวองค์กร หรือบุคคล และอาจนำไปสู่การกระทำผิดทางกฎหมาย

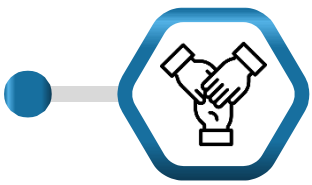
4) ความเสี่ยงด้านความเป็นส่วนตัวของข้อมูล (Data Privacy)

ข้อมูลที่ใช้ในการฝึกฝนโมเดล Generative AI อาจถูกเก็บรวบรวม ใช้ หรือเปิดเผยข้อมูลส่วนบุคคลที่ส่งผลกระทบต่อความเป็นส่วนตัวได้ เช่น ข้อมูลบัตรประชาชน ประวัติการรักษา ที่อยู่ หรือข้อมูลที่สามารถระบุตัวตนได้ เป็นต้น ซึ่งข้อมูลดังกล่าวอาจรั่วไหลจากการโจมตี หรือเป็นส่วนหนึ่งของผลลัพธ์

นอกจากนี้ ผลลัพธ์ที่ปรากฏข้อมูลส่วนบุคคลอาจทำให้เกิดผลลัพธ์ที่มีอคติและการเลือกปฏิบัติที่เป็นอันตรายต่อตัวบุคคลได้

5) ความเสี่ยงด้านสิ่งแวดล้อม (Environmental)

ความเสี่ยงด้านสิ่งแวดล้อมจากการใช้ทรัพยากรจำนวนมากในการฝึกโมเดล อาจสร้างผลกระทบด้านพลังงานและการปล่อยคาร์บอนของ Generative AI ซึ่งขึ้นอยู่กับประเภทของโมเดล, รูปแบบ, ฮาร์ดแวร์ และประเภทของแอปพลิเคชัน



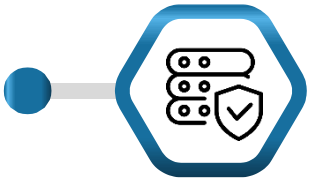
6) ความเสี่ยงด้านการทำงานร่วมกันระหว่างมนุษย์และ Generative AI (Human-AI Configuration)

การกำหนดระดับการมีส่วนร่วมหรือปฏิสัมพันธ์ระหว่างมนุษย์และระบบ AI อาจก่อให้เกิดความเสี่ยง เช่น

การอคติหรือไม่เชื่อผลลัพธ์ การเชื่อผลลัพธ์มากเกินไปโดยไม่ตรวจสอบ ความไม่สอดคล้องกับเป้าหมายและ/หรือผลลัพธ์ที่ต้องการ การใช้ในทางที่ผิด การใช้อย่างไม่ถูกต้อง และการใช้อย่างไม่ปลอดภัยกับมนุษย์ เป็นต้น

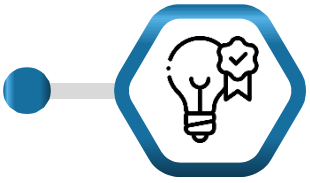
7) ความเสี่ยงด้านการนำข้อมูลที่ไม่ครบถ้วนถูกต้อง หรือบิดเบือน (Information Integrity)

Generative AI อาจถูกใช้เป็นเครื่องมือในการสร้างและเผยแพร่ข้อมูลที่ไม่ถูกต้อง (misinformation) หรือข้อมูลที่ถูกทำให้บิดเบือน (disinformation) ซึ่งอาจทำลายความเชื่อมั่นต่อบุคคล องค์กร และสังคมในวงกว้าง



8) ความเสี่ยงด้านความปลอดภัยของข้อมูล (Information Security)

ผู้ไม่หวังดีอาจใช้ Generative AI ในการหาช่องโหว่ของระบบ การเขียนโปรแกรมในการเจาะระบบขององค์กร นอกจากนี้ Generative AI เองยังเป็นเป้าหมายในการถูกโจมตี เช่น การโจมตีที่โมเดลโดยตรง โดยเฉพาะการโจมตีแบบ Prompt injection หรือการแก้ไขข้อมูลที่ใช้ฝึกฝนโมเดล (Data Poisoning) ซึ่งส่งผลให้ Generative AI ทำงานผิดพลาด



9) ความเสี่ยงด้านทรัพย์สินทางปัญญา (Intellectual Property)

ผลลัพธ์ที่มาจาก Generative AI อาจละเมิดทรัพย์สินทางปัญญา เนื่องจากการจำข้อมูลหรือการสร้างเนื้อหาที่คล้ายกับผลงานที่ได้รับความคุ้มครองลิขสิทธิ์ รวมถึงการใช้อัตลักษณ์หรือลักษณะเด่นของ

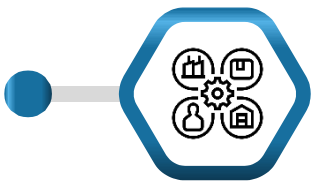
บุคคลที่ไม่ได้รับอนุญาตอาจเป็นปัญหาที่ไม่ได้รับการคุ้มครองจากกฎหมายทรัพย์สินทางปัญญา

10) ความเสี่ยงด้านเนื้อหาลามก คุกคามหรือล่วงละเมิดทางเพศ (Obscene, Degrading, and/or Abusive Content)

Generative AI มีโอกาสสร้างเนื้อหาลามกอนาจาร คุกคามหรือล่วงละเมิดทางเพศ เนื่องจากโมเดล AI ถูกฝึกฝนด้วยชุดข้อมูลเปิดในอินเทอร์เน็ตที่อาจมีเนื้อหาลักษณะดังกล่าว รวมถึงข้อมูลที่ไม่ได้รับอนุญาต ดังนั้น ผลลัพธ์ที่ถูกสร้างขึ้นอาจส่งผลกระทบทางจิตใจและร่างกายของบุคคลที่เกี่ยวข้อง

11) ความเสี่ยงด้านเนื้อหาที่มีความคิดเชิงลบ อคติ แบ่งแยก (Toxicity, Bias, and Homogenization)

Generative AI มีโอกาสสร้างเนื้อหาที่มีความคิดเชิงลบ อคติ หรือแบ่งแยก ที่อาจถูกเผยแพร่เป็นวงกว้างและไม่สามารถควบคุมการแพร่กระจายได้ง่าย ซึ่งอาจก่อให้ความเสียหายแก่ชื่อเสียงและจิตใจของบุคคลที่เกี่ยวข้องได้



12) ความเสี่ยงโดยรวมของห่วงโซ่อุปทาน (Value Chain and Component Integration)

โมเดล Generative AI อาจถูกฝึกด้วยเนื้อหาที่ไม่ได้รับการตรวจสอบจากแหล่งที่มาของบุคคลที่สามซึ่งอาจทำให้ผลลัพธ์ของโมเดลไม่สามารถตรวจสอบได้ และอาจจะก่อให้เกิดความเสี่ยงต่อผู้ที่เกี่ยวข้องในห่วงโซ่อุปทาน

การวิเคราะห์ประเด็นความเสี่ยงของการประยุกต์ใช้ Generative AI ของแต่ละองค์กรนั้นอาจมีความแตกต่างกัน จึงควรพิจารณาโอกาสความเป็นไปได้ (Likelihood) และผลกระทบ (Impact) ที่อาจเกิดขึ้นในมิติต่าง ๆ ดังนั้น องค์กรควรหมั่นติดตามข้อมูลที่เกี่ยวข้องกับเทคโนโลยีและการพัฒนาของ Generative AI เพื่อนำมาทบทวนแนวทางจัดการความเสี่ยงที่อาจเกิดขึ้นจากการประยุกต์ใช้ Generative AI อย่างเหมาะสมต่อไป

3.2 แนวทางการบริหารจัดการความเสี่ยง

Generative AI มีความสามารถที่แตกต่างไปจาก AI ประเภทอื่นจึงนำมาซึ่งความเสี่ยงรูปแบบใหม่ ดังนั้น ในการประยุกต์ใช้ Generative AI จึงควรคำนึงถึงประเด็นความเสี่ยงที่เกี่ยวข้องและกำหนดมาตรการเพื่อป้องกันหรือลดความเสี่ยงที่อาจจะเกิดขึ้นอย่างเหมาะสม โดยแนวทางในการจัดการความเสี่ยงในการประยุกต์ใช้ Generative AI นั้นมีหลากหลายแนวทาง โดยยกตัวอย่าง ดังนี้

1) **กำหนดกรอบแนวทางการประยุกต์ใช้ Generative AI อย่างมีธรรมาภิบาล (Establish Generative AI Governance Structure)** องค์กรควรกำหนดกรอบแนวทางในการประยุกต์ใช้ Generative AI ที่ครอบคลุมรูปแบบและลักษณะการประยุกต์ใช้งานจริงภายในองค์กร ทั้งนี้ โดยควรพิจารณาประเด็นความเสี่ยงและผลกระทบที่อาจจะเกิดขึ้น และกำหนดหน้าที่และความรับผิดชอบ (Role and Responsibility) รวมถึงกำหนดความรับผิดชอบต่อผลของการกระทำ (Accountability) ในการประยุกต์ใช้ Generative AI

2) **หมั่นตรวจสอบและทบทวนความสอดคล้องตามข้อกำหนด หรือข้อกฎหมาย (Ensure Regulatory and Legal Compliance)** ที่เกี่ยวข้องกับการประยุกต์ใช้ Generative AI ว่ายังเป็นไปตามข้อกำหนดหรือข้อกฎหมายหรือไม่ โดยอาจจำเป็นต้องปรับปรุงกฎระเบียบภายในองค์กรให้ทันสมัยอยู่เสมอ ทั้งนี้ ในบางกรณีอาจต้องมีการปรับปรุงระบบให้สอดคล้องกับกฎ ระเบียบ หรือกฎหมายที่เกี่ยวข้อง

3) **ส่งเสริมการประยุกต์ใช้ Generative AI อย่างมีจริยธรรม (Foster a Culture of Ethical Generative AI)** พัฒนาองค์ความรู้บุคลากรในการใช้งาน

Generative AI อย่างรู้เท่าทัน สร้างความเข้าใจเกี่ยวกับประเด็นด้านจริยธรรมและแนวปฏิบัติที่ดีในการประยุกต์ใช้ Generative AI พร้อมทั้งมีมาตรการตรวจสอบว่ามีการปฏิบัติตามแนวทางที่ดีที่องค์กรกำหนดไว้

4) **กำหนดแนวทางการทำงานร่วมกันระหว่างมนุษย์และ Generative AI (Ensure Human Oversight)** สร้างแนวทางการทำงานร่วมกันระหว่างมนุษย์และ Generative AI เพื่อหลีกเลี่ยงการพึ่งพา Generative AI มากเกินไป โดยเพิ่มบทบาทการมีส่วนร่วมของมนุษย์ในกระบวนการทำงานกับ Generative AI ซึ่งจะช่วยลดความเสี่ยงจากความผิดพลาดของ Generative AI

5) **สร้างความร่วมมือระหว่างฝ่ายงานต่าง ๆ ในองค์กร (Promote Interdisciplinary Collaboration)** ส่งเสริมการบูรณาการ การทำงานระหว่างฝ่ายต่าง ๆ ในการพัฒนาและใช้งาน Generative AI ตั้งแต่ประเมินความเสี่ยงไปจนถึงการออกแบบกลยุทธ์บริหารจัดการความเสี่ยง

6) **พัฒนาแนวทางกำกับดูแลด้านข้อมูลขององค์กร (Enhance Data Governance)** จัดทำกรอบแนวทางการกำกับดูแลการนำข้อมูลไปใช้กับ Generative AI อย่างเหมาะสม โดยคำนึงถึงความสอดคล้องกับหลักจริยธรรม AI การจัดการข้อมูลที่เป็นระบบ การตรวจสอบความถูกต้อง การปกป้องความเป็นส่วนตัว การทำความเข้าใจบริบทของข้อมูลก่อนนำไปฝึกฝนโมเดล การนำข้อมูลไปใช้งานกับ Generative AI ทั้งนี้ ควรตรวจสอบการกำกับดูแลข้อมูลในประเด็นต่าง ๆ ข้างต้นอย่างสม่ำเสมอ

7) **เฝ้าติดตาม ประเมินผล และปรับปรุงการใช้งาน (Monitor, Evaluate, and Improve)** ควรมีการประเมินผลทั้งก่อนและหลังการนำ Generative AI ไปใช้จริงในองค์กร เพื่อให้แน่ใจว่าการใช้งานเป็นไปตามเป้าหมายในบริบทสถานการณ์จริง และควรมีกลไกช่องทางการรับฟังความคิดเห็น (Feedback) จากผู้ใช้งานและนำมาปรับปรุงระบบให้มีประสิทธิภาพ

8) **ประเมินและตรวจสอบผลิตภัณฑ์หรือบริการที่เกี่ยวข้องจากหน่วยงานภายนอก (Monitor and Evaluate Products/Services by External Parties)** ประเมินความเสี่ยงและตรวจสอบผลิตภัณฑ์หรือบริการ (เช่น เครื่องมือ โมเดล และชุดข้อมูล) จากหน่วยงานภายนอก เพื่อให้แน่ใจว่ามีความสอดคล้องตามนโยบายการจัดการความเสี่ยงขององค์กร รวมถึงมีการติดตามประสิทธิภาพอย่างสม่ำเสมอ เพื่อให้สามารถรับรู้ถึงความเสี่ยงใหม่ที่อาจจะเกิดและปรับแผนจัดการความเสี่ยงตามความจำเป็นเหมาะสม

9) **กำหนดมาตรการและเฝ้าระวังความมั่นคงปลอดภัยทางไซเบอร์ (Establish Cyber Security Mechanisms)** จัดทำมาตรการการรักษาความปลอดภัยทางไซเบอร์ เพื่อป้องกันการเข้าถึงข้อมูลและระบบโดยไม่ได้รับอนุญาต (Unauthorized Access) การเจาะระบบ (Hacking) การละเมิดข้อมูล (Data Breaches) การรั่วไหลข้อมูล (Data Leakage) เข้ารหัสข้อมูลสำคัญ อัปเดตโปรโตคอลความปลอดภัยและตรวจสอบสิทธิ์การเข้าถึงเป็นประจำ

ทั้งนี้ สามารถกำหนดมาตรการบริหารจัดการความเสี่ยงอื่น ๆ เพิ่มเติมตามที่องค์กรเห็นความเหมาะสมและสอดคล้องกับเป้าหมายที่กำหนดได้

ตัวอย่างการควบคุมและจัดการความเสี่ยง

ตัวอย่างที่ 1: Chatbot ช่วยตอบคำถามแทนฝ่ายทรัพยากรบุคคล

รายละเอียดของ Scenario

องค์กรได้นำเอา Generative AI Chatbot มาช่วยงานฝ่ายทรัพยากรบุคคล (HR) ขององค์กรในการตอบคำถามของพนักงาน โดย Chatbot ได้รับการออกแบบมาเพื่อช่วยตอบคำถามหลาย ๆ ด้านเกี่ยวกับสิทธิสวัสดิการและผลประโยชน์ที่พนักงานจะได้รับ การจ่ายเงินเดือน ตลอดจนผลการปฏิบัติงาน เป็นต้น

แม้ว่า Chatbot ข้างต้นจะสามารถสนับสนุนพนักงานภายในองค์กร และช่วยลดภาระงานให้แก่เจ้าหน้าที่ฝ่ายทรัพยากรบุคคล แต่อย่างไรก็ตาม ได้มีพนักงานจำนวนหนึ่งแจ้งปัญหาของ Chatbot กรณีที่สร้างเนื้อหาคำตอบที่ไม่ถูกต้อง รวมถึงมีการเปิดเผยข้อมูลที่มีความลับ อาทิ รายงานการประเมินผลการปฏิบัติงาน หรือข้อมูลส่วนบุคคล

ตัวอย่างความเสี่ยงที่อาจเกิดขึ้น

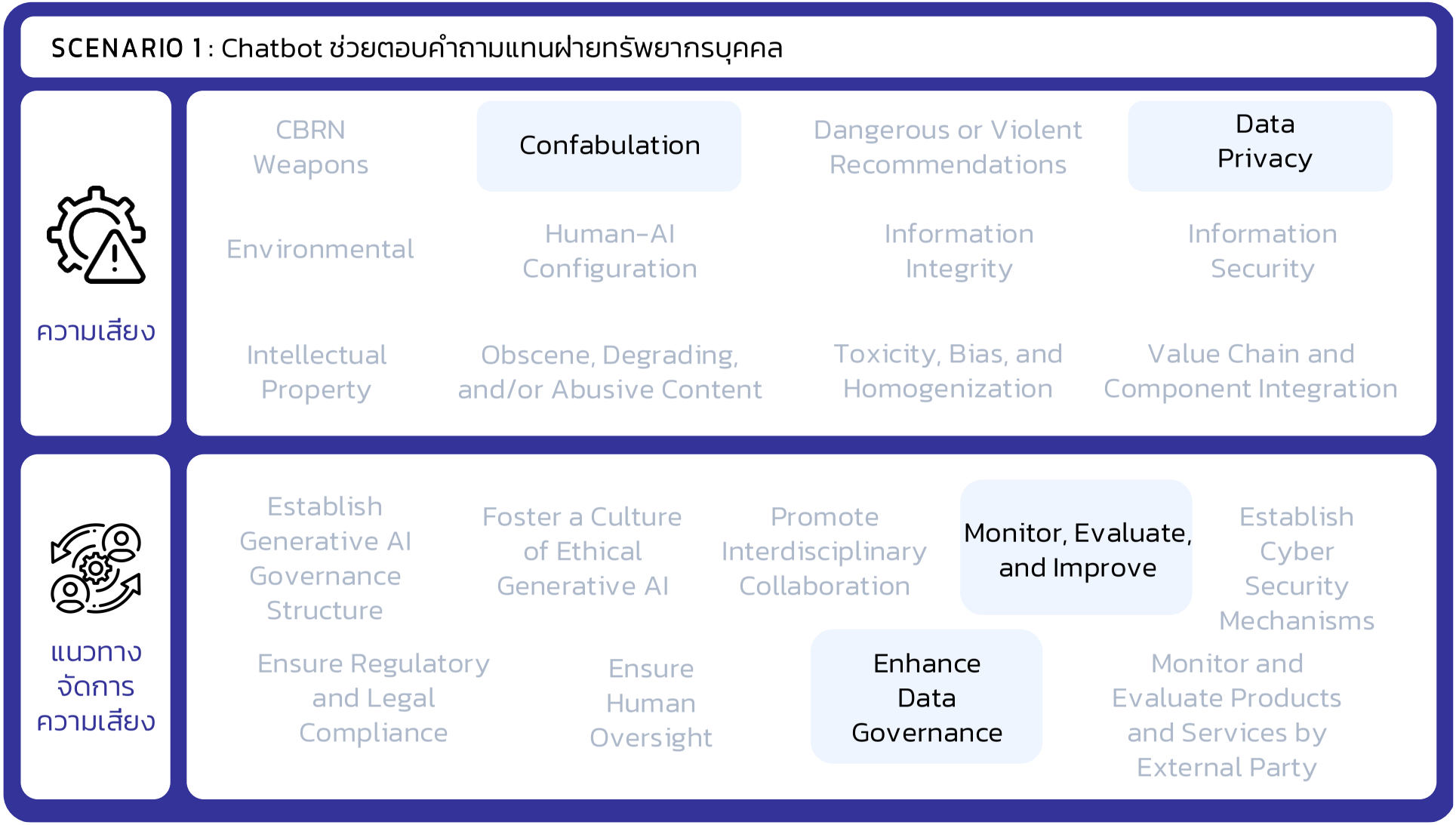
- **ความเสี่ยงด้านเนื้อหาที่น่าเชื่อถือแต่ไม่ถูกต้อง (Confabulation)**
  Generative AI Chatbot ตอบข้อมูลพนักงาน โดยใช้ข้อมูลสวัสดิการของพนักงานที่ไม่มีอยู่จริงตามที่ระบุในเอกสาร

- **ความเสี่ยงด้านความเป็นส่วนตัวของข้อมูล (Data Privacy)**
  Generative AI Chatbot ตอบคำถามพนักงานโดยนำข้อมูลส่วนบุคคลของพนักงานคนหนึ่ง มาตอบคำถามแก่พนักงานอีกคนหนึ่ง

ตัวอย่างแนวทางการจัดการความเสี่ยง

- **เฝ้าติดตาม ประเมินผล และปรับปรุงการใช้งาน (Monitor, Evaluate, and Improve)**
  สร้างกระบวนการหรือช่องทางเพื่อรับฟังความเห็นจากผู้ใช้งาน (Feedback Loop) เกี่ยวกับผลลัพธ์ของ Generative AI Chatbot เช่น เพิ่มปุ่มให้พนักงานสามารถเลือกเพื่อให้ความเห็นว่าคำตอบของ Generative AI Chatbot ถูกต้องเหมาะสมหรือไม่ และนำข้อมูลความเห็นนั้นมาประเมินสาเหตุและปรับปรุงผลลัพธ์ของ Generative AI Chatbot เช่น Retrieval-Augmented Generation (RAG) เป็นต้น
- **พัฒนาแนวทางกำกับดูแลด้านข้อมูลขององค์กร (Enhance Data Governance)**
  องค์กรควรกำหนดนโยบายมาตการควบคุมการเข้าถึงข้อมูลและความปลอดภัยของข้อมูล เพื่อป้องกันการเข้าถึงข้อมูลอ่อนไหวโดยไม่ได้รับอนุญาต รวมไปถึงกำหนดรูปแบบของข้อมูลที่เหมาะสมในการนำมาใช้งานกับร่วม Generative AI อย่างชัดเจน

โดยสรุปจากตัวอย่างที่ 1 สามารถเขียนเป็น
แผนภาพแสดงความเสี่ยงและแนวทางจัดการ



ตัวอย่างที่ 2 : การสร้างโฆษณาการตลาดจาก Generative AI

รายละเอียดของ Scenario

ฝ่ายการตลาดของธนาคารแห่งหนึ่งนำ Generative AI เข้ามาช่วยในการสร้างโฆษณาการตลาดเพื่อเสนอขายโปรโมชันบัตรเครดิตแก่ลูกค้า โดยมุ่งหวังว่าการประยุกต์ใช้ Generative AI จะช่วยเพิ่มประสิทธิภาพและความคิดสร้างสรรค์ในการสื่อสารทางการตลาดกับลูกค้าได้อย่างรวดเร็วและตรงกลุ่มเป้าหมายมากยิ่งขึ้น

แต่อย่างไรก็ตาม มักเจอปัญหาที่ Generative AI สร้างผลลัพธ์ที่อาจจะยังไม่ถูกต้อง ครบถ้วนและอาจเสี่ยงต่อการละเมิดทรัพย์สินทางปัญญาด้วย

ตัวอย่างความเสี่ยงที่อาจจะเกิด

- **ความเสี่ยงด้านเนื้อหาที่น่าเชื่อถือแต่ไม่ถูกต้อง (Confabulation)**
  Generative AI อาจสร้างเนื้อหาที่ดูเหมือนน่าเชื่อถือ แต่มีการเบี่ยงเบนไปจากข้อเท็จจริง เช่น ให้ข้อมูลบัตรเครดิตที่ผิดพลาดเกี่ยวกับสิทธิประโยชน์ และเจ้าหน้าที่นำผลลัพธ์ไปให้ข้อมูลลูกค้าโดยไม่ตรวจสอบ เป็นต้น

- **ความเสี่ยงด้านทรัพย์สินทางปัญญา (Intellectual Property)**
  ข้อความที่ Generative AI สร้างขึ้นอาจละเมิดสิทธิ์ทรัพย์สินทางปัญญา เช่น การสร้างเนื้อหาที่คล้ายกับโฆษณาหรือข้อความการตลาดที่ได้รับความคุ้มครองลิขสิทธิ์ในอินเทอร์เน็ต เป็นต้น

- **ความเสี่ยงด้านการนำข้อมูลที่ไม่ครบถ้วนถูกต้อง หรือบิดเบือน (Information Integrity)**
  ทีมการตลาดอาจนำผลลัพธ์จาก Generative AI ที่ไม่ถูกต้องครบถ้วน หรือบิดเบือน ไปนำเสนอต่อลูกค้า เช่น โฆษณาโปรโมชันที่มีเงื่อนไขไม่ครบถ้วน หรือผิดไปจากนโยบายธนาคาร ซึ่งอาจทำให้ลูกค้าเกิดความเข้าใจผิด เป็นต้น

ตัวอย่างแนวทางการจัดการความเสี่ยง

- **กำหนดกรอบแนวทางการประยุกต์ใช้ Generative AI อย่างมีธรรมาภิบาล (Establish Generative AI Governance Structure)**
  ธนาคารควรกำหนดแนวทางที่ชัดเจนในการประยุกต์ใช้ Generative AI เช่น เจ้าหน้าที่ฝ่ายการตลาดมีหน้าที่และความรับผิดชอบในการตรวจสอบผลลัพธ์ก่อนให้ข้อมูลลูกค้า เป็นต้น

- **กำหนดแนวทางการทำงานร่วมกันระหว่างมนุษย์และ Generative AI (Ensure Human Oversight)**
  ควรออกแบบกระบวนการที่ให้มนุษย์มีส่วนร่วมในการตรวจสอบและตัดสินใจยอมรับหรือปฏิเสธเนื้อหาที่สร้างขึ้นจาก Generative AI โดยมนุษย์จะเป็นผู้ตัดสินความเหมาะสมและความน่าเชื่อถือของเนื้อหาดังกล่าว

- **หมั่นตรวจสอบและทบทวนความสอดคล้องตามข้อกำหนด หรือข้อกฎหมาย (Ensure Regulatory and Legal Compliance)**
  องค์กรควรทบทวนข้อกำหนด รวมถึงข้อกฎหมายที่เกี่ยวข้องกับลิขสิทธิ์และกฎหมายอื่นที่เกี่ยวข้อง เนื่องจาก Generative AI อาจใช้ภาพ หรือสื่อบนอินเทอร์เน็ตที่มีลิขสิทธิ์อยู่ รวมไปถึงควรตรวจสอบนโยบายและข้อกำหนดของผู้ให้บริการ Generative AI ว่าปฏิบัติตามกฎหมายหรือข้อกำหนดที่เกี่ยวข้องหรือไม่

- **ส่งเสริมการประยุกต์ใช้ Generative AI อย่างมีจริยธรรม (Foster a Culture of Ethical Generative AI)**
  องค์กรควรจัดการฝึกอบรมและสร้างความตระหนักรู้ในการประยุกต์ใช้ Generative AI อย่างเหมาะสม เช่น ให้ผู้ใช้งานหมั่นตรวจสอบความถูกต้องของข้อมูลก่อนนำไปใช้งานจริง และให้พนักงานแจ้งผลลัพธ์ที่ไม่เหมาะสมต่อผู้ที่เกี่ยวข้อง

โดยสรุปจากตัวอย่างที่ 2 สามารถเขียนเป็นแผนภาพแสดงความเสี่ยงและแนวทางจัดการความเสี่ยงได้ ดังนี้

SCENARIO 2 : การสร้างข้อความโฆษณาการตลาดจาก Generative AI

ความเสี่ยง

- CBRN Weapons
- Confabulation
- Dangerous or Violent Recommendations
- Data Privacy
- Environmental
- Human-AI Configuration
- Information Integrity
- Information Security
- Intellectual Property
- Obscene, Degrading, and/or Abusive Content
- Toxicity, Bias, and Homogenization
- Value Chain and Component Integration

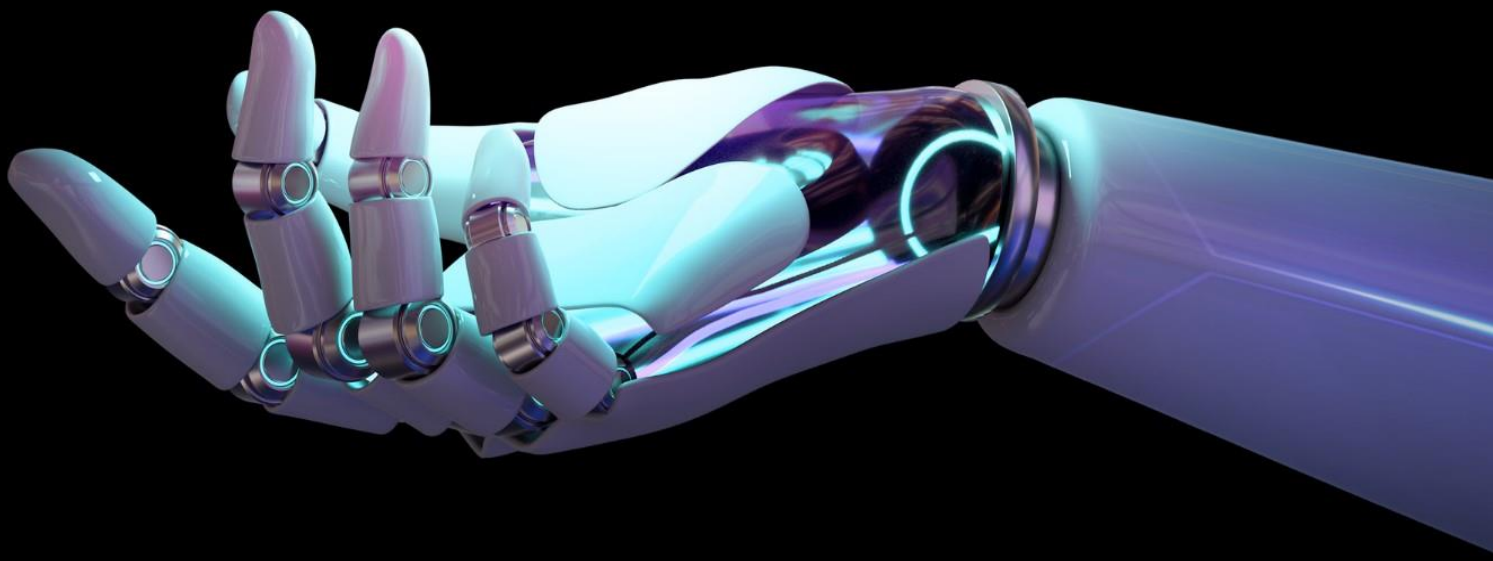
แนวทางจัดการความเสี่ยง

- Establish AI Governance Frameworks
- Foster a Culture of Ethical Generative AI
- Interdisciplinary Collaboration
- Deployment Monitoring
- Evolve Security Protocols
- Ensure Regulatory and Legal Compliance
- Ensure Human Oversight
- Enhance Data Governance
- Evaluate and Monitor Third-Party Components

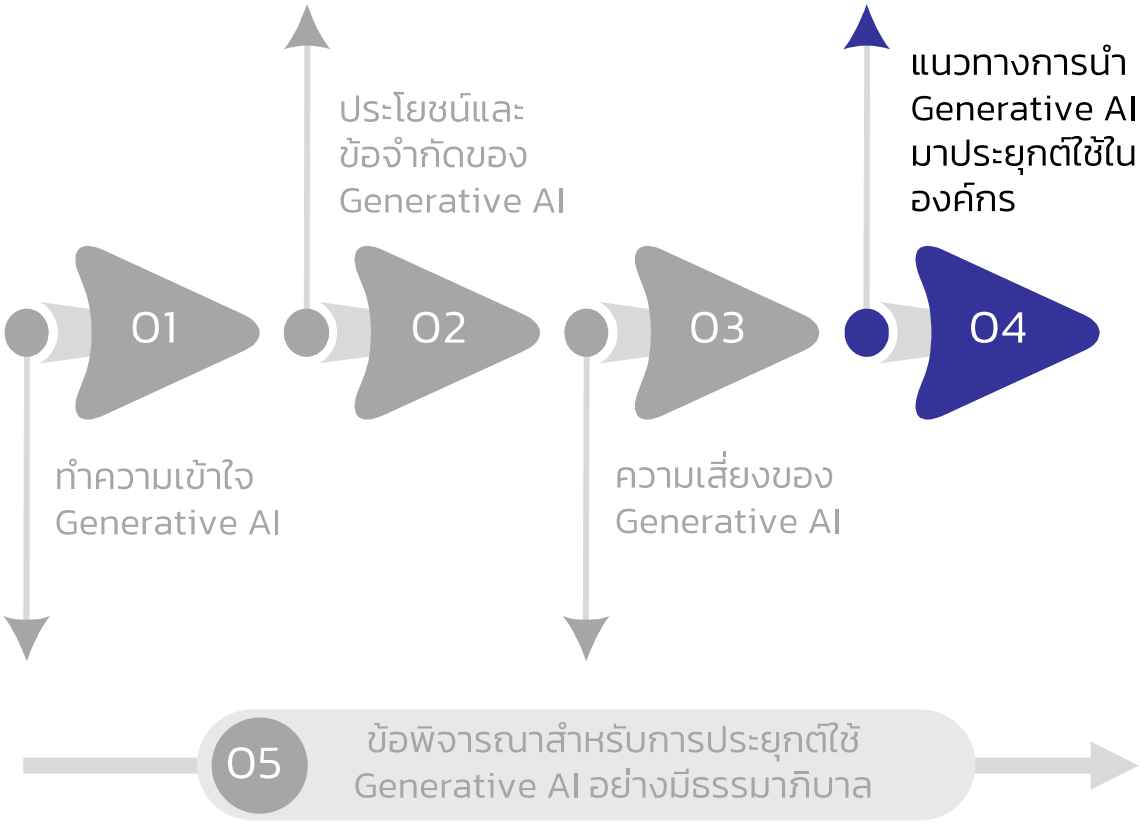
04

แนวทางการนำ Generative AI มาประยุกต์ใช้ในองค์กร

Deploying Generative AI in an Organizations



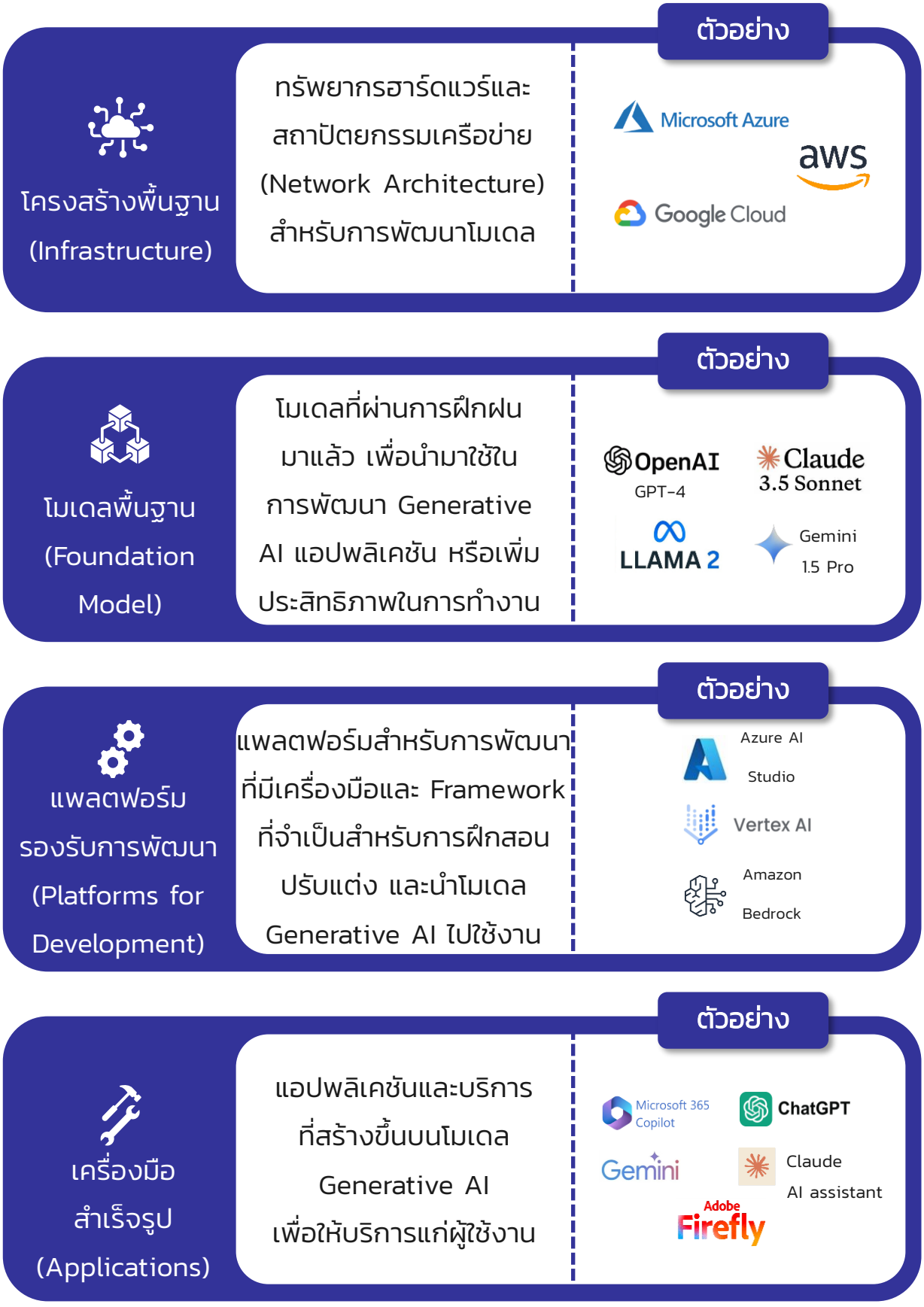
Generative AI Governance Guideline
แนวทางการประยุกต์ใช้ Generative AI อย่างมีธรรมาภิบาล
สำหรับองค์กร



เพื่อให้องค์กรเข้าใจและสามารถตัดสินใจเลือกรูปแบบของการนำ Generative AI มาประยุกต์ใช้ให้ตรงกับเป้าหมายและกลยุทธ์องค์กร จึงควรทำความเข้าใจองค์ประกอบต่าง ๆ ของ Generative AI รวมถึงรูปแบบต่าง ๆ ในการประยุกต์ใช้ Generative AI ในองค์กร

4.1 โครงสร้างเทคโนโลยีที่เกี่ยวกับ Generative AI

การทำความเข้าใจองค์ประกอบของโครงสร้างเทคโนโลยีที่เกี่ยวกับการพัฒนา Generative AI เป็นรากฐานสำคัญสำหรับการพิจารณาดำเนินโครงการเกี่ยวกับ Generative AI สำหรับองค์กร รวมถึงการออกแบบโซลูชันทางเทคนิค (Technical Solution) และการกำกับดูแลการประยุกต์ใช้ Generative AI อย่างมีธรรมาภิบาล


โครงสร้างพื้นฐาน
(Infrastructure)
ทรัพยากรฮาร์ดแวร์และ
สถาปัตยกรรมเครือข่าย
(Network Architecture)
สำหรับการพัฒนาโมเดล
ตัวอย่าง
Microsoft Azure
aws
Google Cloud
โมเดลพื้นฐาน
(Foundation
Model)
โมเดลที่ผ่านการฝึกฝน
มาแล้ว เพื่อนำมาใช้ใน
การพัฒนา Generative
AI แอปพลิเคชัน หรือเพิ่ม
ประสิทธิภาพในการทำงาน
ตัวอย่าง
OpenAI
GPT-4
Claude
3.5 Sonnet
LLAMA 2
Gemini
1.5 Pro
แพลตฟอร์ม
รองรับการพัฒนา
(Platforms for
Development)
แพลตฟอร์มสำหรับการพัฒนา
ที่มีเครื่องมือและ Framework
ที่จำเป็นสำหรับการฝึกสอน
ปรับแต่ง และนำโมเดล
Generative AI ไปใช้งาน
ตัวอย่าง
Azure AI
Studio
Vertex AI
Amazon
Bedrock
เครื่องมือ
สำเร็จรูป
(Applications)
แอปพลิเคชันและบริการ
ที่สร้างขึ้นบนโมเดล
Generative AI
เพื่อให้บริการแก่ผู้ใช้งาน
ตัวอย่าง
Microsoft 365
Copilot
ChatGPT
Gemini
Claude
AI assistant
Adobe
Firefly

4.2 รูปแบบของการนำ Generative AI มาประยุกต์ใช้ในองค์กร

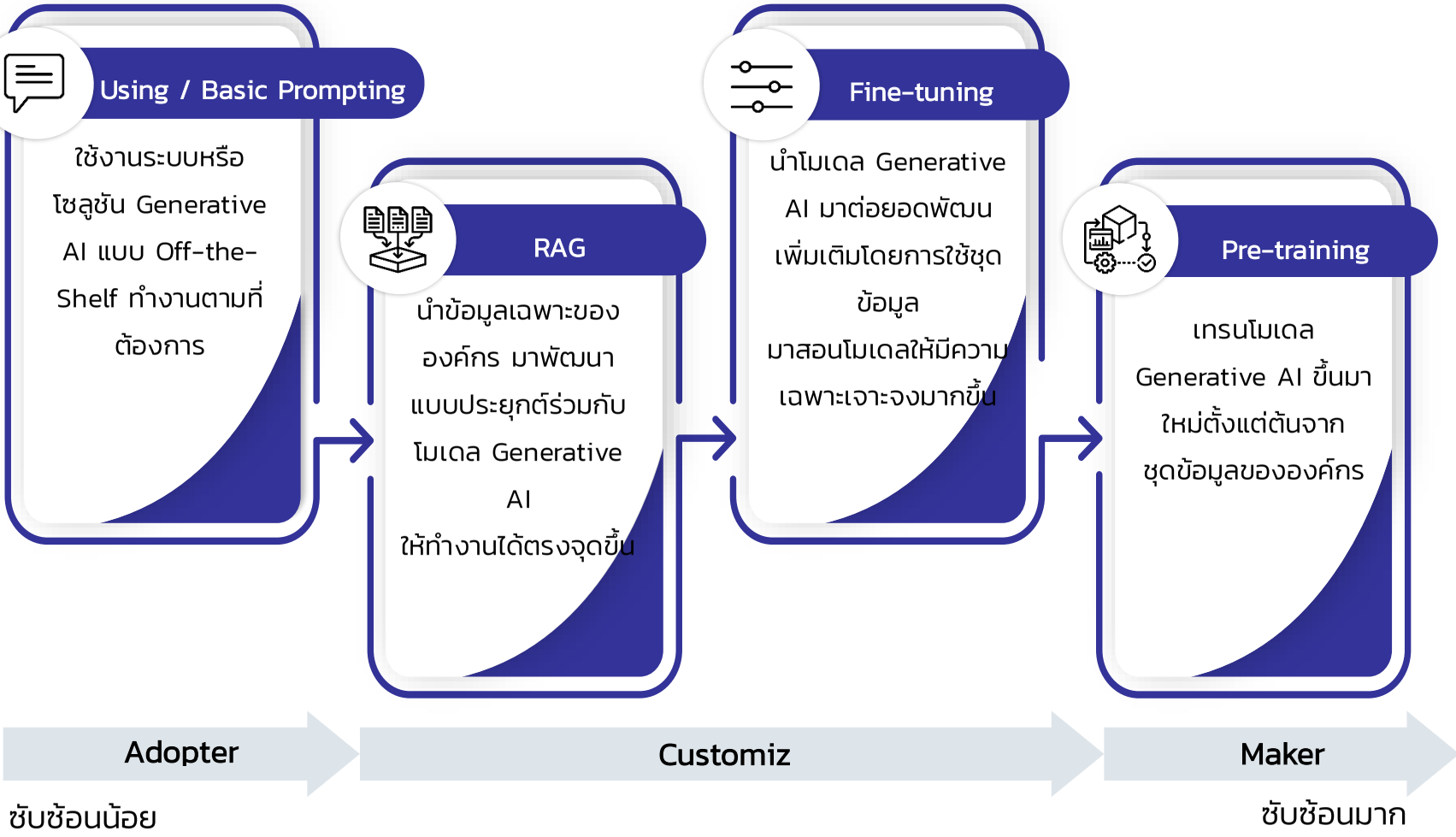
ก่อนที่องค์กรจะเริ่มนำ Generative AI ไปใช้งานจริง คำถามสำคัญที่ควรพิจารณาคือ **"องค์กรควรเลือกใช้ Generative AI สำเร็จรูปหรือพัฒนาเอง จึงจะเหมาะกับบริบทขององค์กรมากที่สุด?"** ดังนั้น ผู้บริหารหรือผู้ที่เกี่ยวข้องในองค์กรควรมีความเข้าใจว่าการนำ Generative AI มาประยุกต์ใช้ในองค์กรมีรูปแบบใดบ้าง

รูปแบบของการนำ Generative AI มาประยุกต์ใช้งานในองค์กรสามารถแบ่งได้ตามความซับซ้อนของการนำไปปรับใช้ โดยเรียงจากความซับซ้อนน้อยไปมาก

1) การประยุกต์ใช้แบบผู้นำไปใช้ (Adopter) ที่มีความซับซ้อนน้อย
2) การประยุกต์ใช้แบบผู้ปรับแต่ง (Customizer) ที่มีความซับซ้อนปานกลาง
3) การประยุกต์ใช้แบบผู้สร้าง (Maker) ที่มีความซับซ้อนมากที่สุด

ทั้งนี้ แต่ละองค์กรสามารถมีรูปแบบการนำ Generative AI ไปประยุกต์ใช้ได้หลายรูปแบบ ซึ่งหัวใจสำคัญใน**การเลือกรูปแบบการใช้ประโยชน์จาก Generative AI ควรเริ่มต้นจากการกำหนดเป้าหมายในการนำ Generative AI ไปใช้งานที่ชัดเจนก่อน** เพื่อให้สามารถเลือกรูปแบบในการประยุกต์ใช้งานอย่างเหมาะสม

ตัวอย่างรูปแบบการประยุกต์ใช้ Generative AI ประเภท LLM



ผู้นำไปใช้ (Adopter)

การประยุกต์ใช้แบบผู้นำไปใช้ (Adopter) สามารถทำได้ง่ายและมีความซับซ้อนน้อยที่สุดในการนำ Generative AI มาประยุกต์ใช้ในองค์กร โดยการประยุกต์ใช้รูปแบบนี้องค์กรจะใช้แอปพลิเคชันและบริการที่สร้างขึ้นบนโมเดล Generative AI ที่เป็นโซลูชันแบบพร้อมใช้งาน (Off-the-Shelf)

> **ตัวอย่าง** ทีมการตลาดใช้ Microsoft 365 Copilot ช่วยเขียนข้อความการตลาด หรือทีมออกแบบใช้ DALL-E ในการสร้างรูปภาพ ซึ่งเป็นโซลูชันแบบ Off-the-Shelf ที่มีในตลาด

ผู้ปรับแต่ง (Customizer)

การประยุกต์ใช้แบบผู้ปรับแต่ง (Customizer) จะเหมาะกับองค์กรที่ต้องการพัฒนาโซลูชันตอบโจทย์ความต้องการเฉพาะที่ซับซ้อนกว่ารูปแบบ Adopter โดยองค์กรอาจใช้งานผ่านการสร้างโซลูชันแบบ Retrieval-Augmented Generation (RAG) ซึ่งเป็นเทคนิคที่ช่วยองค์กรสามารถดึงข้อมูลเฉพาะขององค์กรหรือแหล่งข้อมูลที่น่าเชื่อถือ มาพัฒนาแบบประยุกต์ร่วมกับโมเดล Generative AI ที่ Pre-Trained มาแล้ว เพื่อปรับแต่งให้โซลูชันทำงานได้ตรงจุดมากขึ้น

สำหรับองค์กรที่มีการประยุกต์ใช้ที่ซับซ้อนขึ้น อาจมีการนำโมเดล Generative AI ที่ Pre-Trained ระดับหนึ่งแล้วมาต่อยอดเพิ่มเติม (Fine-Tuning) โดยการใช้ชุดข้อมูลขนาดใหญ่ เพื่อสอนโมเดลเดิมให้มีความเฉพาะเจาะจงมากขึ้น

> **ตัวอย่าง** ฝ่ายบริการลูกค้าสร้างแชทบอทสำหรับช่วยตอบคำถามลูกค้า โดยใช้เทคนิค RAG เพื่อนำข้อมูลสินค้าและคำถามที่พบบ่อย (FAQ) ขององค์กรมาช่วยในการตอบคำถาม

ผู้สร้าง (Maker)

รูปแบบผู้สร้าง (Maker) เป็นการประยุกต์ใช้ Generative AI ที่มีความซับซ้อนมากที่สุด เนื่องจากเป็นการพัฒนา Foundation Model ขึ้นมาใหม่ตั้งแต่ต้น เพื่อตอบสนองการประยุกต์ใช้ Generative AI สำหรับเป้าหมายของธุรกิจที่มี

ความเฉพาะเจาะจง การประยุกต์ใช้รูปแบบนี้ ต้องมีการเตรียมความพร้อมขององค์กรในด้านต่าง ๆ ทั้งงบประมาณ เครื่องมือและบุคลากร

> **ตัวอย่าง** **นักพัฒนาของบริษัทหลักทรัพย์จัดการกองทุน (บลจ.) พัฒนาโมเดลซึ่งมีความถนัดภาคการเงินและเชี่ยวชาญภาษาไทย สำหรับให้บริการแนะนำผลิตภัณฑ์ทางการเงินแก่ลูกค้า**

โดยสรุปรูปแบบของการนำ Generative AI มาประยุกต์ใช้ในองค์กรแต่ละแบบนั้นมีความเกี่ยวข้องกับโครงสร้างเทคโนโลยีเกี่ยวกับ Generative AI ที่แตกต่างกัน

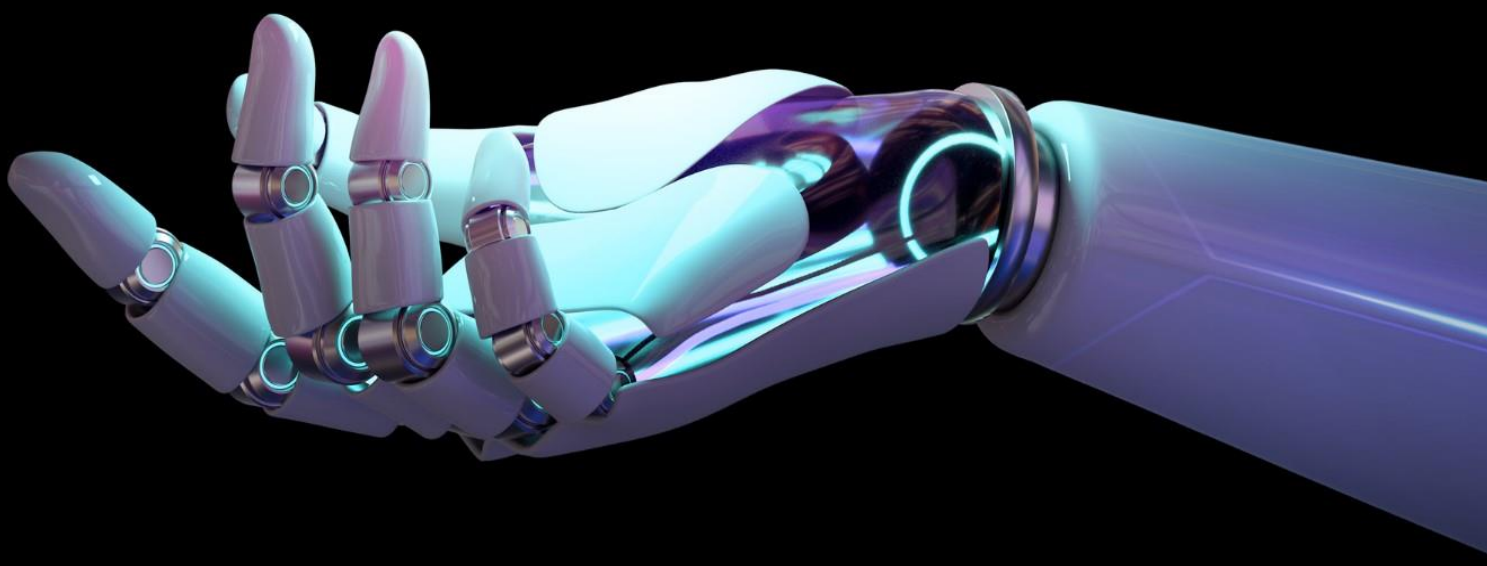
- องค์กรที่ใช้งานในรูปแบบ **Adopter** ควรเน้นทำความเข้าใจองค์ประกอบที่เป็น **เครื่องมือสำเร็จรูป**
- องค์กรที่ใช้งานในรูปแบบ Customizer ควรเน้นทำความเข้าใจองค์ประกอบที่เป็น **โมเดลพื้นฐาน แพลตฟอร์มรองรับการพัฒนา และเครื่องมือสำเร็จรูป**
- องค์กรที่ใช้งานในรูปแบบ **Maker** ควรเน้นทำความเข้าใจถึงโครงสร้างเทคโนโลยีที่เกี่ยวข้องตั้งแต่ **โครงสร้างพื้นฐาน โมเดลพื้นฐาน แพลตฟอร์มรองรับการพัฒนา**

ดังนั้น การมีความเข้าใจและเลือกรูปแบบการประยุกต์ใช้ Generative AI อย่างเหมาะสม จะนำไปสู่แนวทางในการกำกับดูแลการประยุกต์ใช้งาน Generative AI ในองค์กรอย่างมีธรรมาภิบาลต่อไป

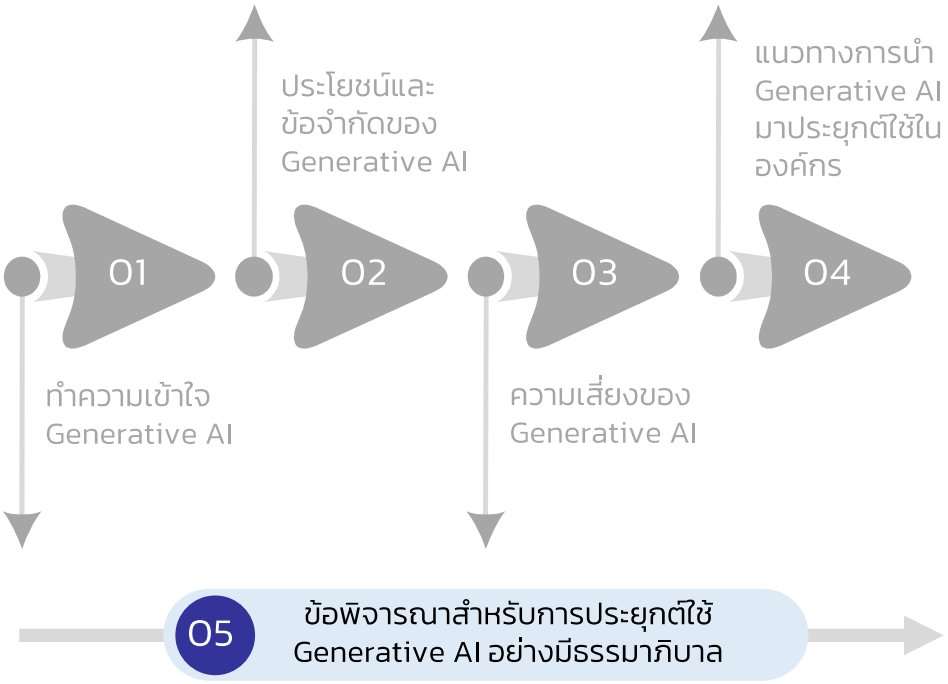
05

ข้อพิจารณาสำหรับการประยุกต์ใช้ Generative AI อย่างมีธรรมาภิบาล

Key Considerations for Generative AI Governance



Generative AI Governance Guideline
แนวทางการประยุกต์ใช้ Generative AI อย่างมีธรรมาภิบาล
สำหรับองค์กร

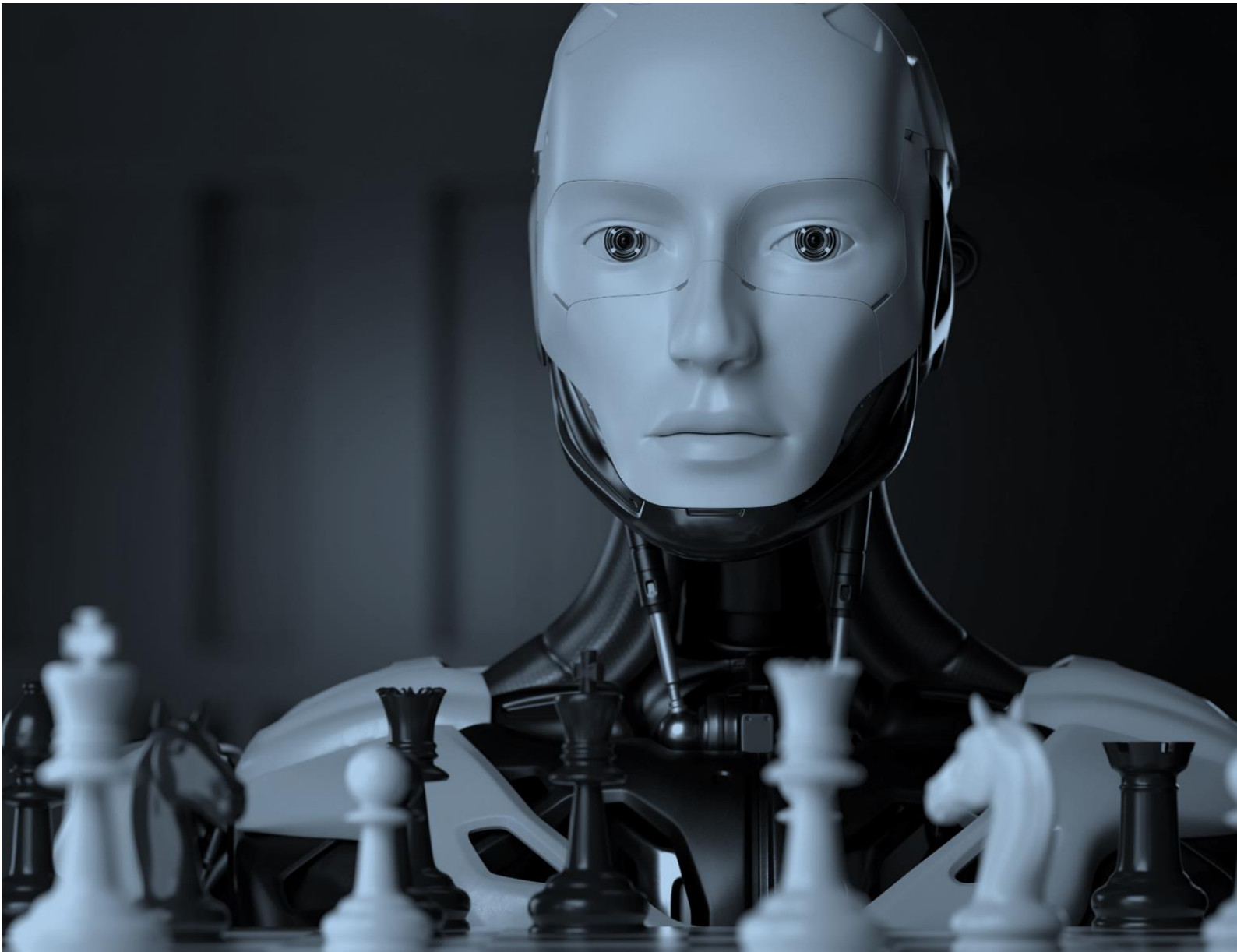


เพื่อให้องค์กรสามารถนำ Generative AI มาประยุกต์ใช้อย่างมีความน่าเชื่อถือและมีความรับผิดชอบต่อผู้ที่เกี่ยวข้อง องค์กรจึงควรมีแนวทางในการกำกับดูแลการประยุกต์ใช้ Generative AI อย่างมีธรรมาภิบาล พร้อมคำนึงถึงประเด็นความเสี่ยงและผลกระทบที่อาจจะเกิดขึ้น โดยหลักสำคัญของการกำกับดูแลการประยุกต์ใช้ Generative AI นั้นเป็นการสร้างสมดุลระหว่างการใช้ประโยชน์กับการบริหารจัดการความเสี่ยงที่มาจากเทคโนโลยีนี้ ควบคู่ไปกับการส่งเสริมให้ผู้ที่เกี่ยวข้องมีส่วนร่วมในกระบวนการต่าง ๆ อย่างเหมาะสม

โดยแนวทางการกำกับดูแลการประยุกต์ใช้ Generative AI ในเล่มนี้จะได้นำเอาคู่มือแนวทางการประยุกต์ใช้ AI อย่างมีธรรมาภิบาลสำหรับผู้บริหารองค์กรมาเป็นกรอบตั้งต้น และขยายความเพิ่มเติมสำหรับกรณีการประยุกต์ใช้งาน GenerativeAI ซึ่งจะมีข้อพิจารณาสำคัญด้านใดที่ผู้ที่เกี่ยวข้องกับการกำกับดูแล Generative AI ในองค์กรควรพิจารณา

5.1 แนวทางการประยุกต์ใช้ Generative AI อย่างมีธรรมาภิบาล

จากเอกสาร "แนวทางการประยุกต์ใช้ปัญญาประดิษฐ์อย่างมีธรรมาภิบาลสำหรับผู้บริหารองค์กร (AI Governance Guideline for Executives)" ได้มีการแบ่งแนวทางการกำกับดูแลออกเป็น 3 องค์ประกอบหลัก ได้แก่ การกำหนดโครงสร้างการกำกับดูแล (AI Governance Structure) การกำหนดกลยุทธ์ในการประยุกต์ใช้ AI (AI Strategy) และการกำกับดูแลการปฏิบัติงานที่เกี่ยวข้องกับ AI (AI Operation)



1. การกำหนดโครงสร้างการกำกับดูแล (AI Governance Structure)

การวางโครงสร้างองค์กรเพื่อสนับสนุนธรรมาภิบาลในการประยุกต์ใช้ AI โดยอาจดำเนินการได้ 3 ส่วน กล่าวคือ การมีคณะกรรมการ (AI Governance Council) การกำหนดหน้าที่ (Role & Responsibility) ของผู้ที่เกี่ยวข้องกับการประยุกต์ใช้ AI ในองค์กรอย่างชัดเจนและครอบคลุมทุกมิติ รวมถึงการเสริมสร้างพัฒนาความรู้ (Competency Building) เพื่อให้สามารถประยุกต์ใช้ Generative AI ตอบโจทย์เป้าหมายองค์กรได้อย่างมีประสิทธิภาพ

การประยุกต์ใช้ Generative AI ในองค์กร สามารถกำหนดโครงสร้างการกำกับดูแลเพิ่มเติม

- องค์กรควรกำหนดนโยบาย แนวปฏิบัติที่ดี ในการประยุกต์ใช้ Generative AI ที่สอดคล้องตามเป้าหมายขององค์กร ข้อกำหนด หลักจริยธรรม ระเบียบและกฎหมาย
- กำหนดหน้าที่ ความรับผิดชอบของบุคลากรในองค์กรและผู้มีส่วนได้เสียที่เกี่ยวข้อง ตัวอย่างเช่น ผู้ใช้งานมีหน้าที่ตรวจสอบผลลัพธ์ก่อนนำไปใช้งาน ผู้ใช้งานมีหน้าที่ในการรายงานต่อผู้บังคับบัญชาในกรณีที่เกิดข้อผิดพลาด เป็นต้น
- พัฒนาองค์ความรู้บุคลากรที่เกี่ยวข้องกับการใช้งาน กำกับดูแลการประยุกต์ใช้ Generative AI ให้เหมาะสมกับหน้าที่และความรับผิดชอบ ตัวอย่างเช่น ทักษะด้าน Prompt Engineering, ทักษะการปรับปรุงประสิทธิภาพโดยใช้เทคนิค RAG เป็นต้น

2. การกำหนดกลยุทธ์ในการประยุกต์ใช้ AI และบริหารจัดการความเสี่ยงจากการประยุกต์ใช้ AI (AI Strategy and Risk Management)

การนำ AI มาประยุกต์ใช้ในองค์กรมีสิ่งสำคัญที่จำเป็นต้องคำนึงถึงคือ ความเข้าใจว่ากรณีการใช้งาน (Use case) นั้นจะสามารถตอบโจทย์เป้าหมายหรือความสามารถของ AI นั้นจะถูกนำมาใช้ประโยชน์เพื่อองค์กรอย่างไร ซึ่งเรียกว่าเป็นการกำหนดกลยุทธ์การใช้งาน AI (Responsible AI Strategy) และควรที่จะวิเคราะห์และกำหนดแนวทางการบริหารความเสี่ยงในการประยุกต์ใช้ AI (AI Risk Management) นั้นไว้ด้วย เพื่อบรรเทาความเสียหายหรือผลกระทบที่อาจจะเกิดขึ้น

การประยุกต์ใช้ Generative AI ภายในองค์กร จำเป็นต้องกำหนดกลยุทธ์และบริหารจัดการความเสี่ยงเพิ่มเติม

- ภายหลังกำหนดกรณีการใช้งาน (Use case) และเป้าหมายในกาประยุกต์ใช้ Generative AI แล้ว องค์กรควรกำหนดรูปแบบการนำ Generative AI มาประยุกต์ใช้ที่เหมาะสม เช่น การเลือกโซลูชันแบบ Off-the-Shelf ที่มีในตลาด การเลือกใช้เทคนิค RAG การเลือกใช้วิธีการ Fine-Tuning หรือการเลือกทำ Pre-training เป็นต้น
- องค์กรควรมีการประเมินความเสี่ยง พร้อมทั้งกำหนดมาตรการบริหารความเสี่ยงที่เกี่ยวข้องกับ Generative AI (ที่ถูกกล่าวไว้ในบทที่ 3) ให้อยู่ในระดับที่ยอมรับได้ (Risk Appetite)
- ควรกำหนดระดับการทำงานร่วมกันระหว่างมนุษย์และ Generative AI ให้สอดคล้องกับระดับความเสี่ยงและผลกระทบเชิงลบที่อาจจะเกิดขึ้น

3. การกำกับดูแลการปฏิบัติงานและให้บริการที่เกี่ยวข้องกับการประยุกต์ใช้ AI (AI Operation)

การกำกับดูแลการประยุกต์ใช้ AI ตลอดวงจรพัฒนา AI (Governing AI Life Cycle) ตั้งแต่การออกแบบโซลูชั่น จัดเตรียมข้อมูล สร้างโมเดล นำโมเดลไปใช้งาน หรือให้บริการ เฝ้าติดตามและประเมินผลการประยุกต์ใช้งาน AI ไปจนถึงการยุติการใช้งานในกรณีที่ไม่เป็นไปตามเป้าหมายที่กำหนดไว้ นอกจากนี้ การสร้างกลไกความรับผิดชอบในการประยุกต์ใช้ AI Service (การให้บริการ) ที่จำเป็นต้องมีการสื่อสารและสะท้อนความคิดเห็น (Feedback) ในการใช้งานระบบ AI ด้วย

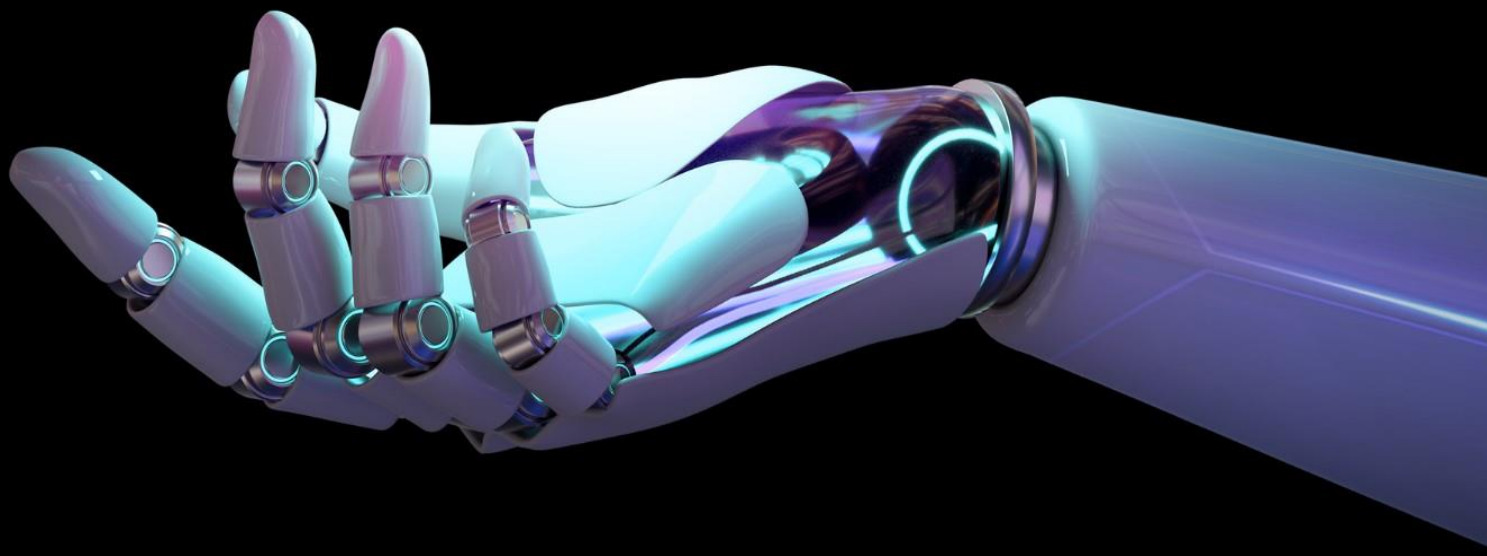
การประยุกต์ใช้ Generative AI ภายในองค์กร จำเป็นต้องกำกับดูแลการปฏิบัติงานและให้บริการที่เกี่ยวข้องเพิ่มเติม

- ควรจัดเตรียมข้อมูลที่มีคุณภาพ ทันสมัย และอยู่ในรูปแบบที่เหมาะสมเพื่อรองรับการพัฒนาโมเดล AI โดยคำนึงถึงความเสี่ยงที่อาจเกิดขึ้น (ที่ถูกกล่าวไว้ในบทที่ 3)
- ควรออกแบบโมเดลหรือระบบ Generative AI ให้มีกลไกป้องกันความเสี่ยงหรือข้อผิดพลาดที่อาจเกิดขึ้น เพื่อให้มั่นใจว่าระบบจะทำงานได้อย่างถูกต้องและปลอดภัย
- ควรสร้างกระบวนการทดสอบโมเดลหรือระบบ Generative AI รวมถึงผลลัพธ์ที่เกิดขึ้นอย่างครอบคลุม และหมั่นตรวจสอบอย่างต่อเนื่องตลอดการใช้งาน
- หลังจากการพัฒนาแล้ว ควรสร้างกลไกการรับฟังความคิดเห็นจากผู้ใช้งาน (Feedback) เพื่อนำมาปรับปรุงโมเดลหรือระบบ Generative AI ให้เหมาะสม และตรงความต้องการของผู้ใช้งานมากขึ้นอย่างต่อเนื่อง

สำหรับการกำกับดูแลการประยุกต์ใช้ Generative AI ในองค์กรอย่างมีธรรมาภิบาลที่ได้นำเสนอข้างต้น จะสามารถช่วยสนับสนุนให้องค์กรสามารถนำ Generative AI ไปประยุกต์ใช้ได้อย่างน่าเชื่อถือ ลดความเสี่ยงและผลกระทบที่อาจจะเกิดขึ้นได้ อย่างไรก็ตาม การปรับใช้แนวทางการประยุกต์ใช้ Generative AI อย่างมีธรรมาภิบาลที่ได้นำเสนอไว้นั้น เป็นเพียงกรอบแนวทางอย่างกว้าง องค์กรจำเป็นต้องมีการวิเคราะห์ประเด็นความเสี่ยงและผลกระทบในบริบทที่เหมาะสมกับองค์กรเพิ่มเติม ซึ่งองค์กรอาจพบแนวทางอื่น ๆ ที่สามารถกำหนดได้อย่างเหมาะสมและสอดคล้องกับเป้าหมายของแต่ละองค์กร

การนำ Generative AI ไปประยุกต์ใช้อย่างมีธรรมาภิบาล เป็นการดำเนินงานที่ต้องปรับเปลี่ยนและพัฒนาอย่างต่อเนื่อง เพื่อให้สอดคล้องกับความเปลี่ยนแปลงในเทคโนโลยีและบริบทขององค์กรอย่างเหมาะสม สามารถสร้างคุณค่าและประโยชน์จากการใช้ Generative AI ได้อย่างมีประสิทธิภาพและยั่งยืนในระยะยาว

ภาคผนวก



ตัวอย่างนโยบายการประยุกต์ใช้ Generative AI (Acceptable Use Policy: Generative AI)

1. บทนำ

ด้วย Generative AI มีความสามารถในการสร้างสรรค์เนื้อหาในรูปแบบที่หลากหลาย เช่น ข้อความ ภาพ วิดีโอ เพลง ซอร์สโค้ด ฯลฯ อีกทั้งยังสามารถสร้างเนื้อหาได้อย่างสมจริงและโต้ตอบกับผู้ใช้งานได้คล้ายคลึงกับมนุษย์ อันสามารถช่วยลดเวลาและเพิ่มประสิทธิภาพในการทำงานได้ตั้งแต่การเขียนบทความ การเขียนข้อความโฆษณา การเขียนโปรแกรม หรือการสร้างภาพศิลปะ

ทั้งนี้ จากความสามารถดังกล่าวอาจทำให้ผู้ใช้งานเชื่อว่าเนื้อหาที่สร้างสรรค์โดย Generative AI นั้นถูกต้อง น่าเชื่อถือ และสามารถนำไปใช้งานได้ทันที โดยไม่จำเป็นต้องพิจารณาถึงความเหมาะสมและข้อจำกัดของเทคโนโลยี โดยเฉพาะอย่างยิ่งข้อจำกัดด้านภาษาไทยและความเข้าใจในบริบทหรือวัฒนธรรมของประเทศไทย ซึ่งอาจก่อให้เกิดผลกระทบต่อบุคคล องค์กร และสังคม เช่น

- เนื้อหาไม่ถูกต้องตามข้อเท็จจริง (Confabulation)
- เนื้อหาส่งผลกระทบต่อประเด็นที่มีความอ่อนไหว รวมถึงมีเนื้อหาที่ขัดต่อหลักการจริยธรรม (Sensitive and Ethics Context)
- เนื้อหาที่สร้างก่อให้เกิดอคติและการเลือกปฏิบัติ (Bias and Discrimination)
- ข้อมูลส่วนบุคคลรั่วไหล (Personal Data Leakage)
- การละเมิดทรัพย์สินทางปัญญา (Intellectual Property Infringement)
- ความมั่นคงปลอดภัย (Information Security)

ดังนั้น จึงเป็นการสมควรกำหนดนโยบายการประยุกต์ใช้ Generative AI (Acceptable Use Policy: Generative AI)

2. วัตถุประสงค์

เพื่อให้เกิดการประยุกต์ใช้ Generative AI อย่างมีความรับผิดชอบสำหรับการดำเนินงานตามภารกิจของสำนักงาน จึงกำหนดนโยบายการใช้เทคโนโลยี Generative AI โดยมีวัตถุประสงค์ ดังนี้

1) เพื่อเป็นแนวทางการประยุกต์ใช้ Generative AI ที่เหมาะสมสำหรับพนักงาน ลูกจ้าง และผู้ปฏิบัติงานที่เกี่ยวข้อง อาทิ แนวทางสำหรับการรักษาความลับและคุ้มครองข้อมูลส่วนบุคคล

2) เพื่อให้การประยุกต์ใช้ Generative AI สอดคล้องกับกฎหมายและข้อกำหนดที่เกี่ยวข้อง และป้องกันผลกระทบที่อาจเกิดขึ้นกับบุคคล องค์กร และสังคม

3. ขอบเขต

นโยบายนี้ครอบคลุมการปฏิบัติงานตามหน้าที่และอำนาจของพนักงาน ลูกจ้าง และผู้ปฏิบัติงานที่เกี่ยวข้อง รวมถึงผู้รับจ้างที่ได้รับมอบหมายให้ดำเนินงาน (ซึ่งต่อไปนี้เรียกว่า “ผู้ใช้งาน”) ที่ใช้งาน Generative AI สำหรับการดำเนินงานตามภารกิจของสำนักงาน

4. แนวทางการประยุกต์ใช้ Generative AI

4.1. การประยุกต์ใช้ Generative AI สำหรับการดำเนินงานตามภารกิจของสำนักงาน

1) การประยุกต์ใช้ Generative AI ต้องเป็นไปเพื่อประโยชน์ของสำนักงานและสอดคล้องตามภารกิจของสำนักงานเท่านั้น

2) ผู้ใช้งานต้องไม่ใช้ Generative AI ในการ

- สร้างเนื้อหาที่ผิดกฎหมายหรือข้อกำหนดที่เกี่ยวข้อง
- สร้างเนื้อหาที่เป็นอันตราย ทำให้เสื่อมเสียชื่อเสียง หรือเนื้อหาที่เป็นการล่วงละเมิดหรือไม่เหมาะสม

- สร้างและแจกจ่ายเนื้อหาที่มีเจตนาบิดเบือน แสดงข้อมูลไม่ถูกต้อง หรือทำให้ผู้อื่นเข้าใจผิด
- ดำเนินการใด ๆ ให้กิจกรรมหรือบริการของสำนักงาน หยุดชะงัก ระงับ ชะลอ ขัดขวาง หรือรบกวนจนไม่สามารถทำงานตามปกติได้

4.2. การรักษาความลับและคุ้มครองข้อมูลส่วนบุคคล

ด้วยแอปพลิเคชันหรือบริการ Generative AI ที่ให้บริการ โดยเฉพาะอย่างยิ่งแอปพลิเคชันหรือบริการที่ไม่มีค่าใช้จ่าย อาจมีการเข้าถึงหรือบันทึกข้อมูลที่โต้ตอบหรือข้อมูลที่ได้รับจากผู้ใช้งานสำหรับนำไปใช้ในการฝึกสอนโมเดล Generative AI เพื่อปรับปรุงประสิทธิภาพการทำงานของ Generative AI ดังนั้น ผู้ใช้งานจึงต้องมีความระมัดระวังและปฏิบัติตามนโยบาย ดังต่อไปนี้

1) ไม่นำข้อมูลภายในองค์กรและข้อมูลที่มีชั้นความลับ (เช่น รหัสผ่าน เอกสารสัญญา เอกสารหรือหนังสือที่ประทับข้อความลับ เอกสารหรือข้อมูลเกี่ยวกับโครงการภายในสำนักงาน ฯลฯ) ไปใช้งานร่วมกับแอปพลิเคชันหรือบริการ Generative AI
2) ไม่นำข้อมูลส่วนบุคคล (เช่น รูปภาพบุคคล ชื่อ-สกุล หมายเลขประจำตัวประชาชน ที่อยู่ หมายเลขโทรศัพท์ อีเมล ฯลฯ) และข้อมูลส่วนบุคคลที่อ่อนไหว (เช่น สำเนาบัตรประชาชนที่มีข้อมูลศาสนา ข้อมูลชีวภาพ ใบรับรองแพทย์ ฯลฯ) ไปใช้งานร่วมกับแอปพลิเคชันหรือบริการ Generative AI
3) ในกรณีที่ต้องใช้แอปพลิเคชันหรือบริการ Generative AI ร่วมกับข้อมูลภายในองค์กร ข้อมูลที่มีชั้นความลับ ข้อมูลส่วนบุคคล และข้อมูลส่วนบุคคลที่อ่อนไหว ผู้ใช้งานจะต้องใช้แอปพลิเคชันหรือบริการที่ได้รับการอนุมัติหรือจัดหาโดยสำนักงานเท่านั้น

4.3. การรักษาความมั่นคงปลอดภัย

1) ไม่นำข้อมูลที่อาจส่งผลกระทบต่อความมั่นคงปลอดภัยของระบบ (เช่น รหัสผ่าน ข้อมูล API Key รายละเอียดการตั้งค่าของระบบ ฯลฯ) ไปใช้งานร่วมกับแอปพลิเคชันหรือบริการ Generative AI
2) ไม่นำซอร์สโค้ดที่สร้างโดย Generative AI มาใช้งาน โดยไม่พิจารณาถึงความถูกต้อง และการตรวจสอบช่องโหว่
3) หากพบเหตุการณ์ละเมิดความมั่นคงปลอดภัยที่เกิดจากการประยุกต์ใช้ Generative AI ผู้ใช้งานต้องแจ้งให้ผู้บริหารตามสายบังคับบัญชารับทราบ และปฏิบัติตามขั้นตอนการปฏิบัติงาน (Work Procedure) เกี่ยวกับการบริหารจัดการเหตุการณ์ความมั่นคงปลอดภัยโดยทันที

4.4. การลดหรือหลีกเลี่ยงการเกิดอคติ (Bias) และการเลือกปฏิบัติ (Discrimination) ต่อบุคคลหรือกลุ่มบุคคล

ด้วยผลลัพธ์จากแอปพลิเคชันหรือบริการ Generative AI ที่ให้บริการ อาจก่อให้เกิดอคติและการเลือกปฏิบัติต่อบุคคลหรือกลุ่มบุคคล ดังนั้น ผู้ใช้งานจึงต้องมีความระมัดระวังและปฏิบัติตามนโยบาย ดังต่อไปนี้

1) ต้องตรวจสอบเนื้อหาที่สร้างโดย Generative AI ก่อนนำไปใช้งานหรือเผยแพร่ต่อสาธารณะ เพื่อลดหรือหลีกเลี่ยงการเกิดอคติหรือการเลือกปฏิบัติอย่างไม่เป็นธรรม
2) ในกรณีที่ใช้แอปพลิเคชันหรือบริการ Generative AI สำหรับการดำเนินงานที่อาจกระทบต่อสิทธิของบุคคลหรือกลุ่มบุคคลโดยตรง ผู้ใช้งานจะต้องใช้ความระมัดระวังเท่าที่จะต้องใช้และสมควรจะต้องใช้สำหรับการดำเนินการดังกล่าว

4.5. เคารพสิทธิในทรัพย์สินทางปัญญา

1) ในการนำเนื้อหาที่สร้างโดย Generative AI ไปใช้งานหรือเผยแพร่ต่อสาธารณะ ผู้ใช้งานต้องใช้แอปพลิเคชันหรือบริการที่ได้รับการอนุมัติโดยสำนักงานเท่านั้น

2) การประยุกต์ใช้ Generative AI ด้วยความประมาทอาจทำให้สำนักงานกระทำการละเมิดสิทธิในทรัพย์สินทางปัญญาได้ ดังนั้น ผู้ใช้งาน Generative AI จึงต้องมีความระมัดระวังไม่ให้เกิดการละเมิดลิขสิทธิ์ เครื่องหมายการค้า หรือสิทธิในทรัพย์สินทางปัญญาอื่น ๆ

5. หน้าที่และความรับผิดชอบ

การใช้งานเนื้อหาที่สร้างโดย Generative AI อาจส่งผลกระทบเชิงลบต่อบุคคล องค์กร และสังคม โดยที่สำนักงานและผู้ใช้งานไม่อาจปฏิเสธความรับผิดต่อผลของการกระทำดังกล่าวได้ ด้วยเหตุนี้ ผู้ใช้งานและบุคคลที่เกี่ยวข้องกับการประยุกต์ใช้ Generative AI จึงมีหน้าที่และความรับผิดชอบในการประยุกต์ใช้ Generative AI ดังต่อไปนี้

1) ผู้ใช้งานต้องแจ้งผู้บังคับบัญชาอย่างชัดเจนเกี่ยวกับวัตถุประสงค์ ขอบเขต และการทำงานร่วมกันระหว่างผู้ใช้งานกับ Generative AI (AI-Human Involvement) เมื่อใช้ Generative AI ในการปฏิบัติงาน รวมทั้งแจ้งพนักงานหรือลูกจ้างให้ทราบว่าข้อมูลใดเกิดจากการประยุกต์ใช้ Generative AI

2) ผู้ใช้งานต้องปฏิบัติตามแนวทางการประยุกต์ใช้ Generative AI ตามที่สำนักงานกำหนดในนโยบายนี้ และตรวจสอบเนื้อหาที่สร้างโดย Generative AI ก่อนนำไปใช้งานหรือเผยแพร่ โดยในการตรวจสอบผู้ใช้งานจำเป็นต้องพิจารณาในประเด็น ดังต่อไปนี้

   - ความถูกต้องของเนื้อหา
   - ผลของการใช้งานหรือเผยแพร่เนื้อหาที่นำไปสู่การกระทำผิดทางกฎหมาย รวมถึงการละเมิดทรัพย์สินทางปัญญา
   - ความเท่าเทียมและการไม่เลือกปฏิบัติต่อบุคคลหรือกลุ่มบุคคล
   - การรักษาความลับและคุ้มครองข้อมูลส่วนบุคคล
   - ผลกระทบต่อความมั่นคงปลอดภัย
   - ผลกระทบเชิงลบอื่น ๆ ที่อาจเกิดขึ้นต่อบุคคล องค์กร และสังคม

3) ผู้ใช้งานต้องรายงานให้ผู้บริหารตามสายการบังคับบัญชารับทราบโดยทันที ในกรณีที่การประยุกต์ใช้ Generative AI เกิดความผิดพลาด หรือพบประเด็นปัญหาทั้งกรณีการนำเข้าข้อมูลและแสดงผลลัพธ์ที่อาจส่งผลกระทบเชิงลบต่อบุคคล องค์กร และสังคม

4) ผู้บริหารตามสายบังคับบัญชาและผู้ใช้งานต้องมีการทบทวนประเด็นปัญหา ประเมินประสิทธิภาพและประสิทธิผลของการประยุกต์ใช้ Generative AI เพื่อปรับปรุงวิธีการทำงานและเลือกใช้แอปพลิเคชันหรือบริการที่เหมาะสมกับการปฏิบัติงาน

5) คณะกรรมการบริหารเทคโนโลยีสารสนเทศ (IT Steering) มีหน้าที่ในการพิจารณาและอนุมัติรายการแอปพลิเคชันหรือบริการ Generative AI ที่เหมาะสมกับการใช้งานภายในสำนักงาน

เอกสารอ้างอิง

Aiverifyfoundation. (2024). *MODEL AI GOVERNANCE FRAMEWORK FOR GENERATIVE AI.* Aiverifyfoundation. https://aiverifyfoundation.sg/wp-content/uploads/2024/05/Model-AI-Governance-Framework-for-Generative-AI-May-2024-1-1.pdf

BCG. (2023). *The CEO's Roadmap on Generative AI.* BCG. https://media-publications.bcg.com/BCG-Executive-Perspectives-CEOs-Roadmap-on-Generative-AI.pdf

ETDA. (2023). *AI Governance Guideline*. ETDA. https://www.etda.or.th/th/Useful-Resource/เอกสารเผยแพร่/แนวทางการประยุกต์ใช้ปัญญาประดิษฐ์อย่างมีธรรมาภิบาล-(สำหรับผู้บริหารองค์กร).aspx

Gartner. (Assessed 2024, February). *Gartner Experts Answer the Top Generative AI Questions for Your Enterprise*. Gartner. https://www.gartner.com/en/topics/generative-ai

Google Cloud. (2024). *Generative AI examples*. Google. https://cloud.google.com/use-cases/generative-ai?hl=en

Google Cloud. (2024). *Grounding.* Google. https://cloud.google.com/vertex-ai/generative-ai/docs/model-reference/grounding

HM Government. (2024). *Generative AI framework for HM Government.* HM Government. https://assets.publishing.service.gov.uk/media/65c3b5d628a4a00012d2ba5c/6.8558_CO_Generative_AI_Framework_Report_v7_WEB.pdf

Kaplan, A., Hutson, E. and Pelaez, N. (2024). Building and Customizing GenAI with Databricks: LLMs and Beyond. Databricks. https://www.databricks.com/blog/building-custom-genai-llms-and-beyond

McKinsey & Company. (2024). *What is AI (artificial intelligence)?* McKinsey. https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-ai

McKinsey & Company. (2023). *What's the future of generative AI? An early view in 15 charts*. McKinsey. https://www.mckinsey.com/featured-insights/mckinsey-explainers/whats-the-future-of-generative-ai-an-early-view-in-15-charts

McKinsey & Company. (2023). *What every CEO should know about generative AI*. McKinsey. https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/what-every-ceo-should-know-about-generative-ai

McKinsey & Company. (2024). *What is prompt engineering?*. McKinsey. https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-prompt-engineering

Microsoft. (2024). *Retrieval Augmented Generation (RAG) in Azure AI Search*. Microsoft. https://learn.microsoft.com/en-us/azure/search/retrieval-augmented-generation-overview

Microsoft. (2024). *Understand LLMs*. Microsoft. https://learn.microsoft.com/en-us/training/modules/introduction-large-language-models/2-understand-large-language-models

NIST Artificial Intelligence Risk Management Framework. (2024, April). *Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile*. NIST. https://airc.nist.gov/docs/NIST.AI.600-1.GenAI-Profile.ipd.pdf

SAP AppHaus. (2023). *Limitations of Generative AI.* SAP. *https://apphaus.sap.com/wp-content/uploads/sites/2/2023/10/GenAI_ExploreWS_Cards_GenAILimitations.pdf*

จัดทำโดยศูนย์ธรรมภิบาลปัญญาประดิษฐ์
(AI Governance Center: AIGC)

ดร. ศักดิ์ เสกขุนทด	ที่ปรึกษาอาวุโส สพธอ.
รจนา ล้ำเลิศ	หัวหน้าทีมศูนย์ AIGC
ธิติกร ตระกูลศิริศักดิ์	ผู้เชี่ยวชาญ
ธงชัย แสงศิริ	ผู้เชี่ยวชาญ
หทัยชนก พุทธรักษา	ผู้ชำนาญการ
กฤตเมธ ธนภูมิพรพรหม	ผู้ชำนาญการ

กระทรวงดิจิทัล
เพื่อเศรษฐกิจและสังคม

ETDA

AIGC
AI GOVERNANCE CENTER
by ETDA

ศูนย์ธรรมาภิบาลปัญญาประดิษฐ์
AI GOVERNANCE CENTER : AIGC
สำนักงานพัฒนาธุรกรรมทางอิเล็กทรอนิกส์
02 123 1237 I aigc@etda.or.th