AIAND DATA ETHICS AND REGULATIONS

ASST PROF PEERAPAT CHOKESUWATTANASKUL, DBA (JDBA), PHD (CAMBRIDGE)

DATA DILEMMAS

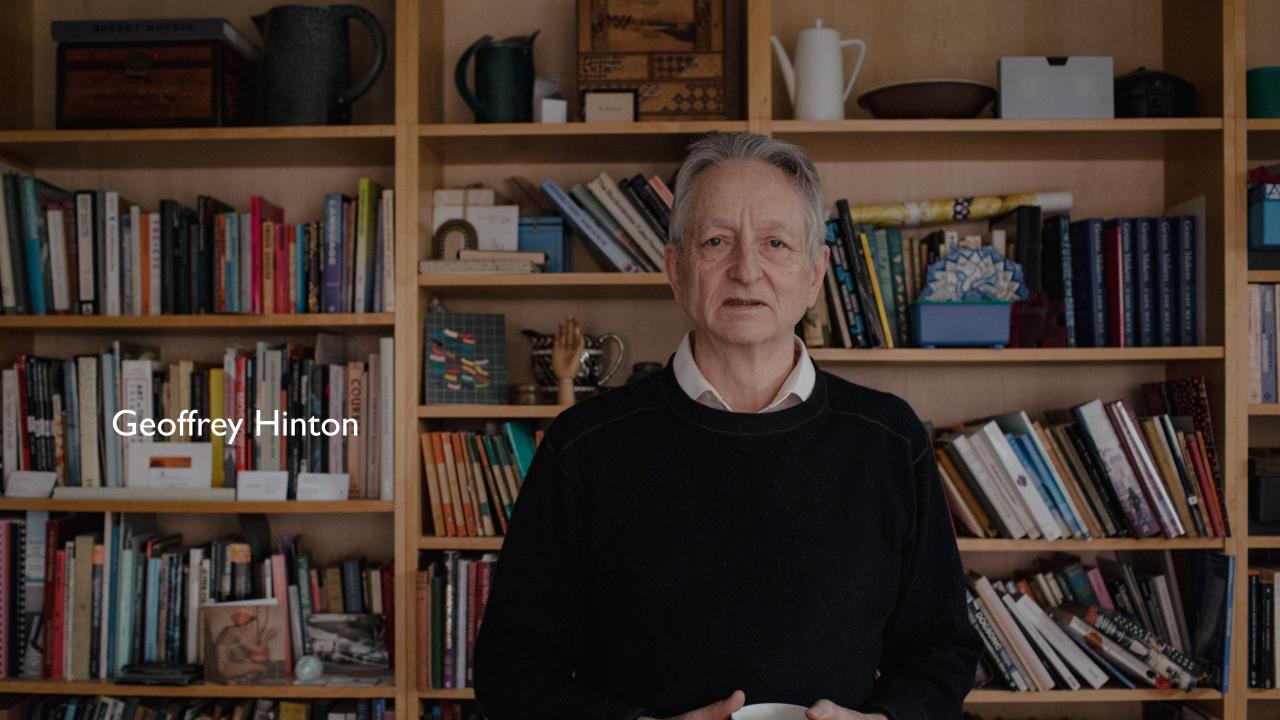
Privacy v. Utility

Accuracy v. Bias

Transparency v. Security

Ownership?

The limit of consent



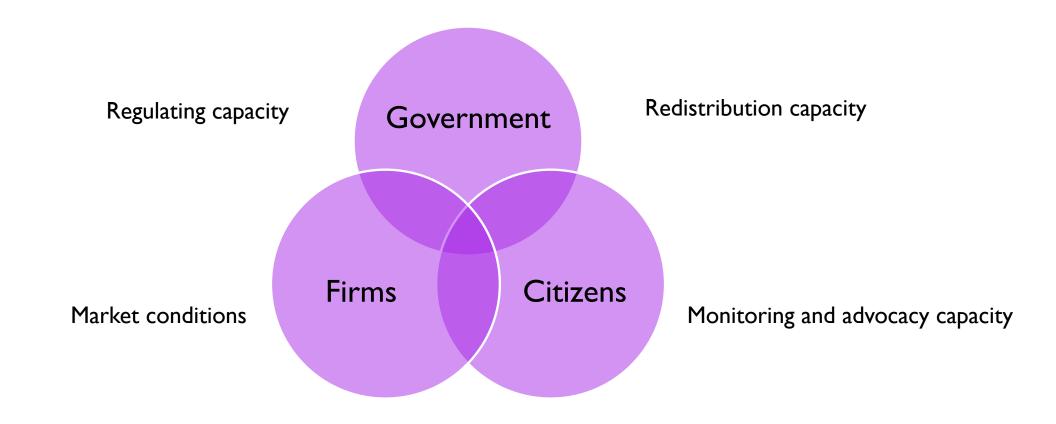
3 KEY TAKEAWAYS

Different aspects of ethics

Not all ethical aspects are equal.

Many ways to make people 'care': Risk-based

UNDERSTAND THE INSTITUTIONAL SETTINGS



INCENTIVES ALIGNMENT?

- One size doesn't fit all. Each mechanism has its unique benefits. However, it depends on 'whom' (not all firms are the same) and 'action' (what kind of problems) we are aiming at.
 - Standard/Guidelines: For those who have other forces in place.
 - Regulation: For those who need some push.
 - Competition: Of course, will enhance all above.
 - Public awareness and advocacy: again, boost them all.
 - Government intervention: Can it be done in the process (China?)?
- Ultimately, we need to consider whether the concerns are 'internalised' into the objective functions of the firms or not. How?
 - Competition might be so effective as consumers care about the 'harms' of AI e.g. biases. Showing that you have considered these aspects will increase the value perceived by consumers and competition will drive you to 'care'.
 - If no competition, regulations will be needed.

IMPACTS ON...

01

People as

- Labours
- Privacy-entitled
- Consumers
- Citizens (e.g., biases)

02

Economic growth

03

Economic distribution

DIFFERENT WAYS TO TACKLE THE PROBLEMS: CHOOSE THE RIGHT ONE

Risk-based regulation and compliance (EU)

- Ex-ante
- Risk-based
- e.g. GDPR, Al Act

Market-driven approach (US)

- Information-based
 - Benchmarks
- Ex-post
- Existing laws e.g.
 - Consumer protection
 - Competition
 - Data security

Ethics and social norms-based approach (China)

- Government were there at the beginning of development.
- Ethics and social norms to set the mindset of the obedient.

AI MAY INVOLVE IN DIFFERENT 'PURPOSES'

AI + purpose

AI regulation

Data + purpose

Personal Data Protection

Specific purposes

Digital Platform

RISK IDENTIFIERS (DE MINIMIS)

Pre-market parameters

Data source

Computing resources required

Investment

Third-party reviewers accessibility

Post-deployment parameters

Numbers of users/downloaders

Types of users

Applications

Revenues

LET'S HAVE A LOOK AT THE MOST CONCRETE ONE: REGULATIONS THAT FORCE YOU TO THINK.

PERSONAL DATA PROTECTION PDPA B.E. 2562 (THAILAND)





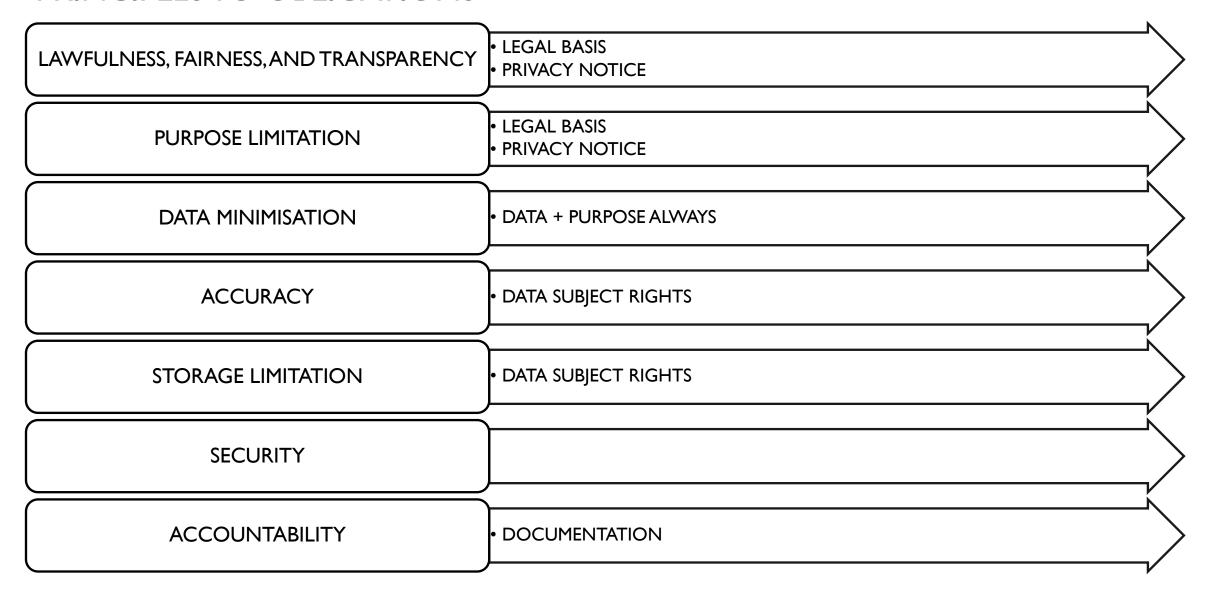








PRINCIPLES TO OBLIGATIONS



KEY ELEMENTS

Object

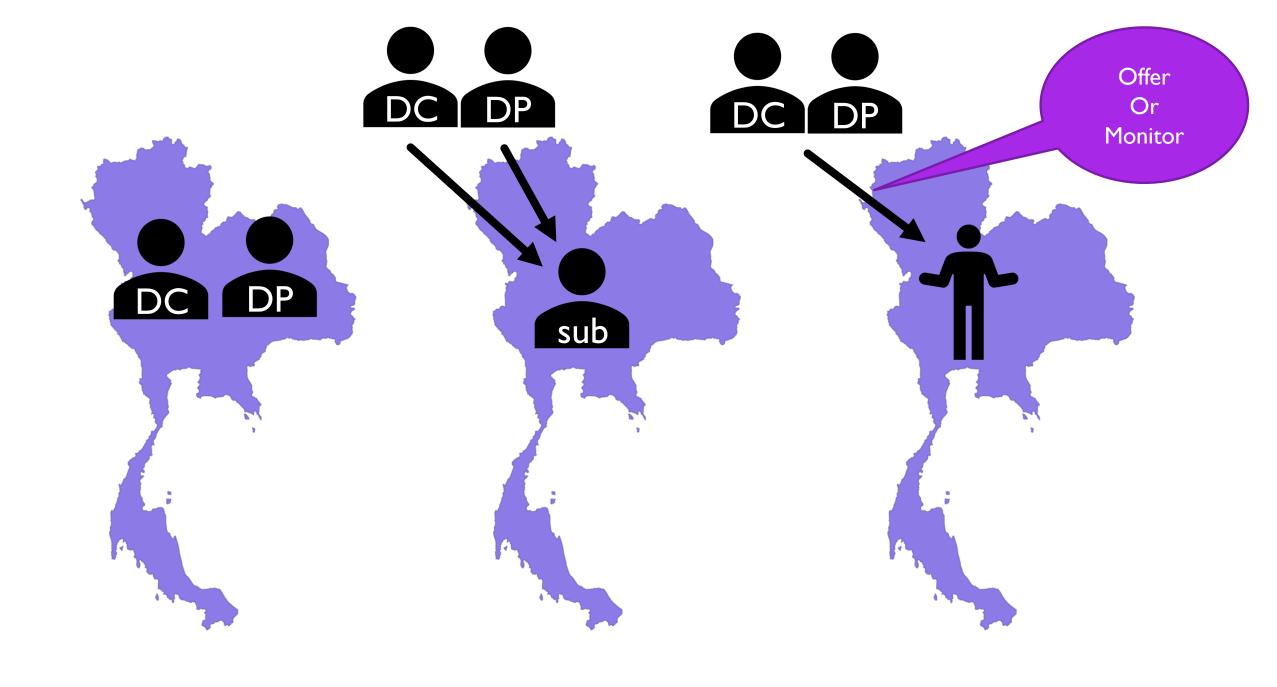
Personal Data => Data subject

Actors

Data Controller, Data Processor

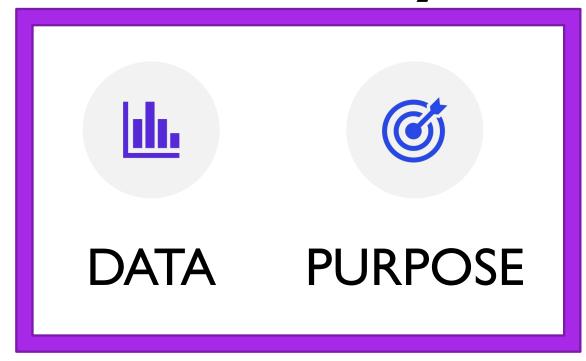
Action

Processing (Collect, Use, Disclose)



UNIT OF ANALYSIS IS NEVER DATA ALONE: 'ACTIVITY'

Activity



TWO PRONGES OF **PDPA**

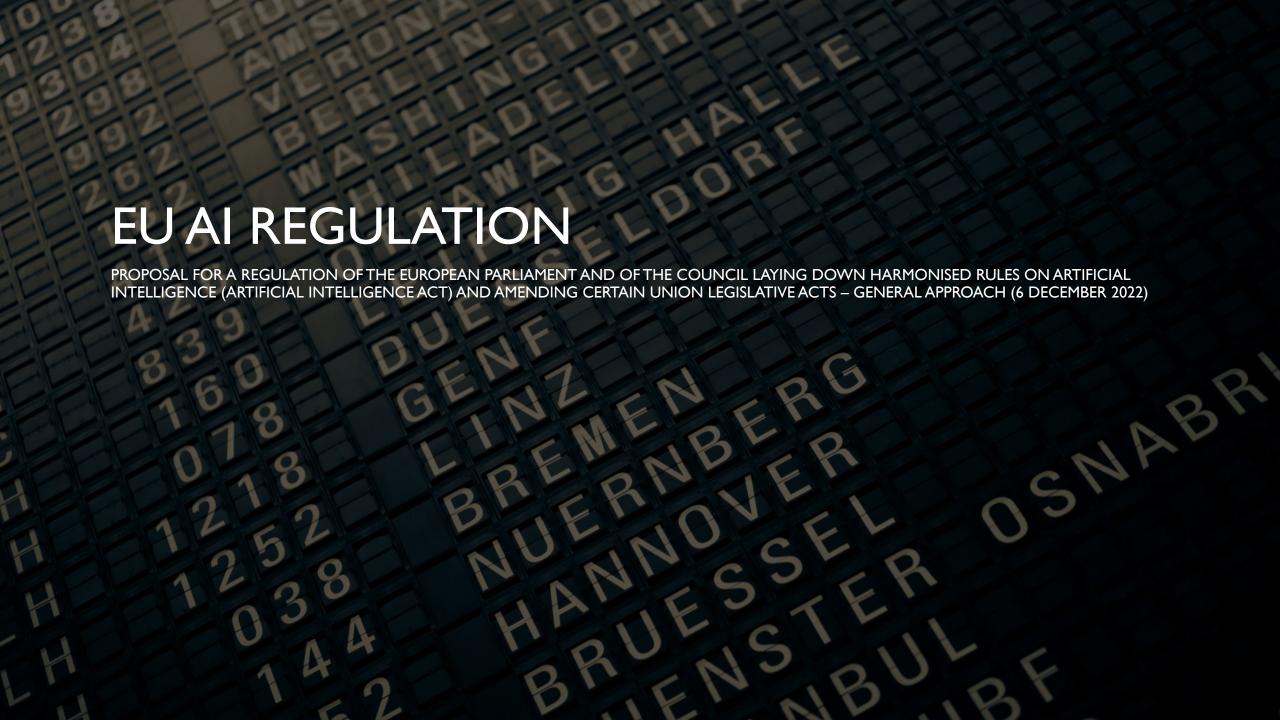




PRIVACY SECURITY

Why? Informed?

CIA



Faculty of Law, Chulalongkorn University

THAILAND ARTIFICIAL INTELLIGENCE

GUIDELINES 1.0

แนวปฏิบัติเกี่ยวกับมาตรฐาน การใช้ปัญญาประดิษฐ์



RESPONSIBLE TRUSTWORTHY ETHICAL HUMAN-CENTRIC

Privacy

Accountability

Safety and Security

Transparency and Explainability

Fairness and Non-discrimination

Human Control of Technology

Professional Responsibility

Promotion of Human Values

A BIG LIST TO SHOP...

Reliability Resilience Privacy Security Robustness Transparency Accountability Usability Interoperability **Fairness** Controllability Well-being

Israel

- ความเป็นธรรม (Fairness)
- ความรับผิดชอบ (Accountability)
- การปกป้องสิทธิมนุษยชน (Protecting human rights)
- ความปลอดภัยทางไซเบอร์และ ข้อมูล (Cyber and information security)
- ความมั่นคงปลอดภัย (Safety)
- การรักษาตลาดการแข่งขันและกฎการ ปฏิบัติที่เอื้อต่อการแข่งขัน (Maintaining a competitive market and rules of conduct that facilitate competition)

China

- ความสอดคล้องและความเป็นมิตร
- ความเป็นธรรมและความยุติธรรม
- ความครอบคลุมและการแบ่งปัน
- การเคารพความเป็นส่วนตัว
- ความปลอดภัยและความสามารถในการควบคุม
- ความรับผิดชอบร่วมกัน
- ความร่วมมือแบบเปิด
- ธรรมาภิบาลที่คล่องตัว

Japan

- หลักการให้มนุษย์เป็นศูนย์กลาง (Human-Centric Principle)
- หลักการให้ความรู้
 (Principle of Education/Literacy)
- หลักการให้ความคุ้มครองความเป็นส่วนตัว (Principle of Privacy Protection)
- หลักการสร้างความมั่นใจด้านความ ปลอดภัย
 (Principle of Ensuring Security)
- หลักการแข่งขันที่เป็นธรรม
 (Principle of Fair Competition)
- หลักความเป็นธรรม ความรับผิดชอบ และ ความ
 โปร่งใส (Principle of Fairness, Account ability, and Transparency)

THAILAND

- 💶 (1) ความสามารถในการแข่งขันและพัฒนาอย่างยั่งยืน (Competitiveness and Sustainability Development)
- 💶 (2) ความสอดคล้องกับกฎหมายจริยธรรมและมาตรฐาน (Law Ethics and International standards)
- (3) ความโปร่งใสและความรับผิดชอบ (Transparency and Accountability)
- (4) ความมั่นคงปลอดภัยและความเป็นส่วนตัว (Security and Privacy)
- (5) ความเท่าเทียม หลากหลาย ครอบคลุมและเป็นธรรม (Fairness)
- (6) ความน่าเชื่อถือ (Reliability)

STATUS QUO

- The initial draft was proposed by the EU Commission in April 2021.
- The Council of the EU adopted the amended draft in December 2022
- The European Parliament has adopted its negotiation position on 14 June 2023.
- The draft is now back with the Council of the European Union to negotiate its position.
- Finally, the interinstitutional (trilogue) negotiations will finalise the process.

CHANGES IN THE LATEST AMENDMENTS

More limited scope

- Exclude Military (previously)
- Introduce the 'general purpose Al' and some duties for the providers.
- Exclude more of the followings:
 - R&D
 - Defense and national security
 - Used for purely personal, non-professional activities => just transparency
- SMEs are exempted from/flexible in many obligations (e.g. tech doc)

CHANGES IN THE LATEST AMENDMENTS

Some stricter rules

- Facial recognition
 - Initially: only real-time remote biometric identification systems in publicly accessible spaces
 - Latest: include 'post' process as well. Explicit consent will be needed.

Stronger safeguards

- Right to privacy, non-discrimination, and fair trial
- Create EU Al Board to oversee the implementation and investigation

EU AI Regulation (initial draft)

'artificial intelligence system' (Al system) means **software** that is developed with one or more of the techniques and approaches listed in Annex I and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with;

EU AI Regulation (newest draft)

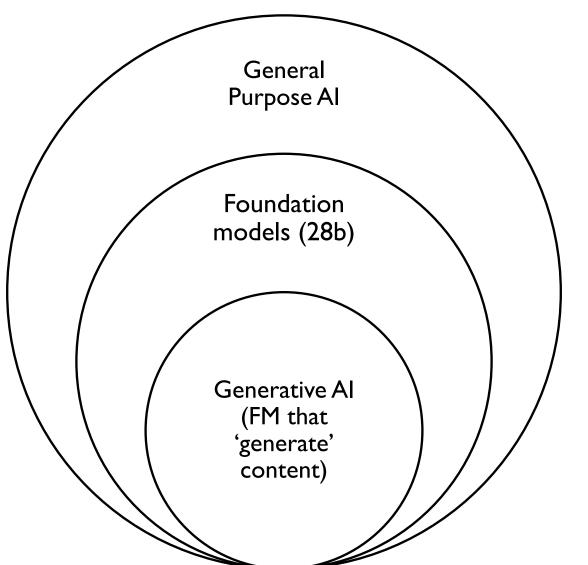
"artificial intelligence system" (Al system) means a machine-based system that is designed to operate with varying levels of autonomy and that can, for explicit or implicit objectives, generate outputs such as predictions, recommendations, or decisions, that influence physical or virtual environments;

ADDED TERMS

• 'Life cycle of an Al system' means the duration of an Al system, from design through retirement. Without prejudice to the powers of the market surveillance authorities, such retirement may happen at any point in time during the post-market monitoring phase upon the decision of the provider and implies that the system may not be used further. An Al system lifecycle is also ended by a substantial modification to the Al system made by the provider or any other natural or legal person, in which case the substantially modified Al system shall be considered as a new Al system.

ADDED TERMS

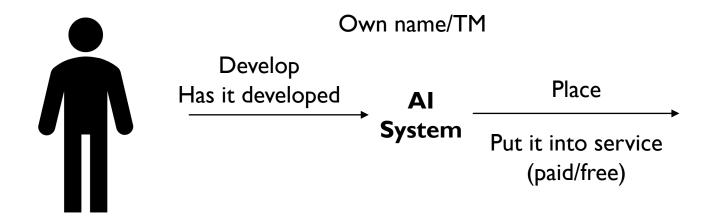
• 'General-purpose Al system' means an Al system that irrespective of how it is placed on the market or put into service, including as open source software is intended by the provider to perform generally applicable functions such as image and speech recognition, audio and video generation, pattern detection, question answering, translation and others; a general purpose Al system may be used in a plurality of contexts and be integrated in a plurality of other Al systems;



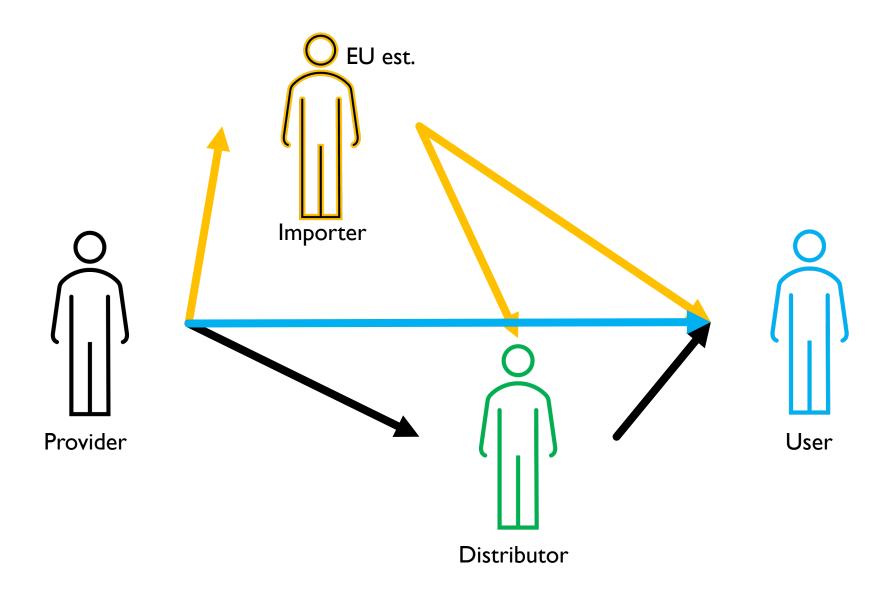
- Risk assessment and mitigation
- Data governance
- Design and develop with standards (performance, predictability, interpretability, corrigibility and sustainability)
- Technical docs
- Get Registered

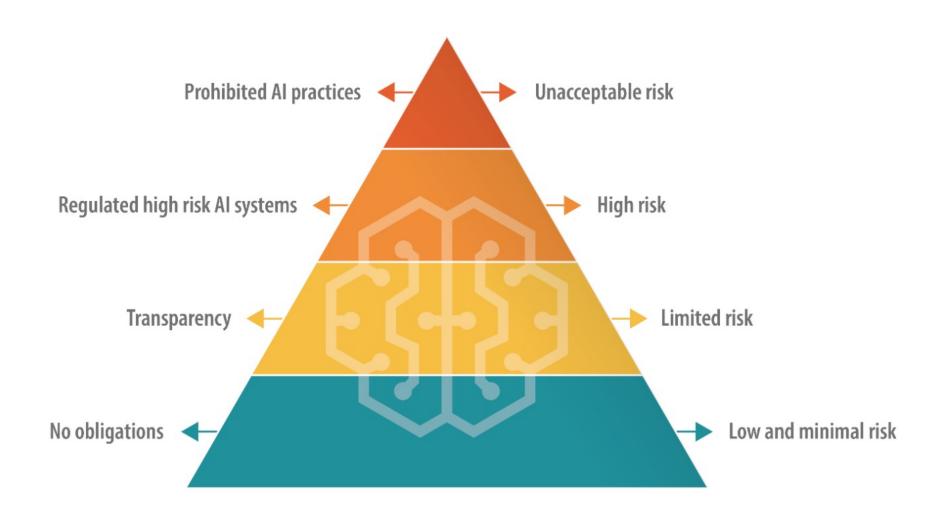
- Transparency
- Content-moderation safeguards
- Copyright data acknowledge

UNDERTAKINGS: PROVIDERS



- a natural or legal person,
- public authority,
- agency or
- other body





UNACCEPTABLE RISK



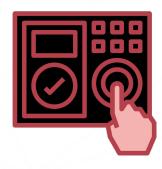
ระบบปัญญาประดิษฐ์ที่ใช้เทคนิคที่ มนุษย์ไม่รู้ตัวเพื่อที่จะเปลี่ยนแปลง พฤติกรรมของมนุษย์อันจะก่อให้เกิด อันตรายในทางร่างกายหรือจิตใจ



ระบบปัญญาประดิษฐ์ที่ใช้ประโยชน์จากความ เปราะบางของกลุ่มบุคคล ไม่ว่าจะเป็น อายุ หรือความทุพพลภาพทางร่างกายหรือจิตใจ เพื่อที่จะเปลี่ยนแปลงพฤติกรรมของบุคคลใน กลุ่มนั้นอันจะก่อให้เกิดอันตรายแก่ตัวเขาเอง หรือบุคคลอื่นในทางร่างกายหรือจิตใจ



ระบบปัญญาประดิษฐ์ที่เพื่อประเมินหรือ จำแนกความน่าเชื่อถือของบุคคลในช่วง ระยะเวลาหนึ่งตามพฤติกรรมทางสังคม หรือคุณลักษณะส่วนบุคคล โดยใช้วิธีการ ให้คะแนนทางสังคม (social scoring) อันนำไปสู่อย่างใดอย่างหนึ่งต่อไปนี้



ระบบปัญญาประดิษฐ์ที่ใช้ระบบการพิสูจน์และ ยืนยันตัวบุคคลด้วยช้อมูลชีวมิติโดยระยะทางแบบ เวลาปัจจุบัน (real-time remote biometric identification system) ในพื้นที่สาธารณะเพื่อ วัตถุประสงค์ในการบังคับใช้กฎหมาย เว้นแต่เพื่อ วัตถุประสงค์ตามที่กฎหมายกำหนดไว้อย่างชัดเจน เท่านั้น

THE UNACCEPTABLES ARE STILL NEGOTIATED.

- Post" remote biometric identification systems, with the only exception of law enforcement for the prosecution of serious crimes and only after judicial authorization;
- Biometric categorisation systems using sensitive characteristics (e.g. gender, race, ethnicity, citizenship status, religion, political orientation);
- Predictive policing systems (based on profiling, location or past criminal behaviour);
- Emotion recognition systems in law enforcement, border management, workplace, and educational institutions; and
- Indiscriminate scraping of biometric data from social media or CCTV footage to create facial recognition databases (violating human rights and right to privacy).

HIGH RISK (5-15% OF EU AI)



โครงสร้างพื้นฐาน ที่สำคัญ (Safety Component of Critical Infrastructure)



การศึกษา และฝึกอบรม (access & evaluation)



การระบุตัวบุคคลโดย ข้อมูลชีวมาตรจาก ระยะไกล (Remote biometric ID system)



การจ้างงานและเลื่อน ตำแหน่ง



การเข้าถึงและใช้งาน บริการที่สำคัญและ ประโยชน์ต่างๆ (รัฐ& เอกชน) e.g. ประกัน



กระบวนการยุติธรรม และการบังคับใช้ กฎหมาย



ตรวจคนเข้าเมือง และการลี้ภัย



การจัดการการ เลือกตั้ง และการ ตัดสินคดีของศาล



ส่งผลต่อการเลือกตั้ง และการแนะนำ (recommender system ของ GK

THE HIGH-RISK ARE ALSO STILL NEGOTIATED.

- Al systems to influence voters in political campaigns
- Al systems in recommender systems used by social media platforms (with more than 45 million users under the <u>Digital Services Act</u>).

LIMITED RISK







HIGH-RISK AI OBLIGATIONS



ธรรมาภิบาลข้อมูล (Data Governance)



เอกสารทางเทคนิค (Technical Documentation)



มีการประเมินการปฏิบัติตามกฎหมาย (Conformity Assessment)



มีการจัดเก็บข้อมูลอัตโนมัติ (Automatically Generated Logs)



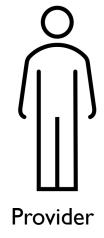
มีความโปร่งใส (Transparency)

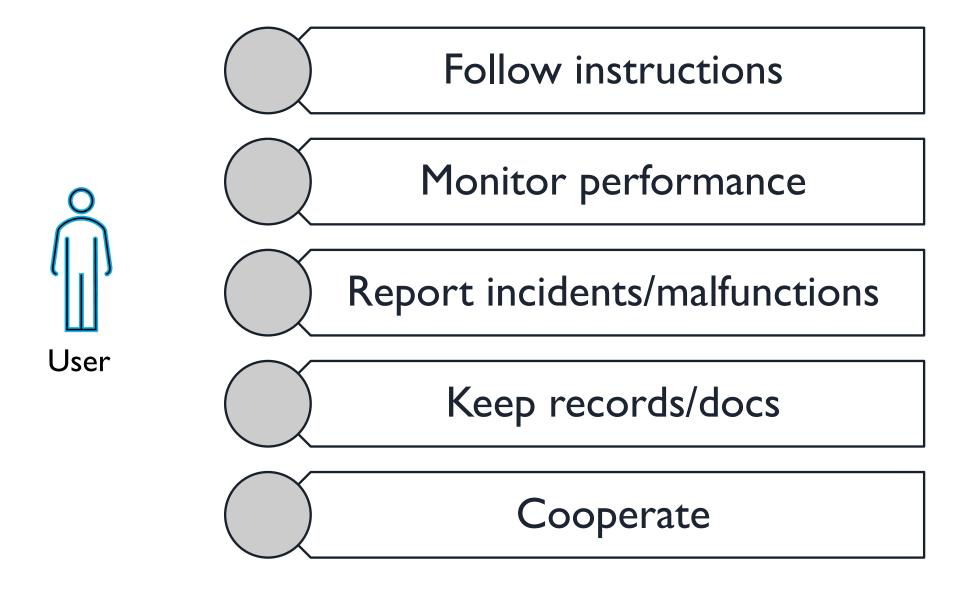


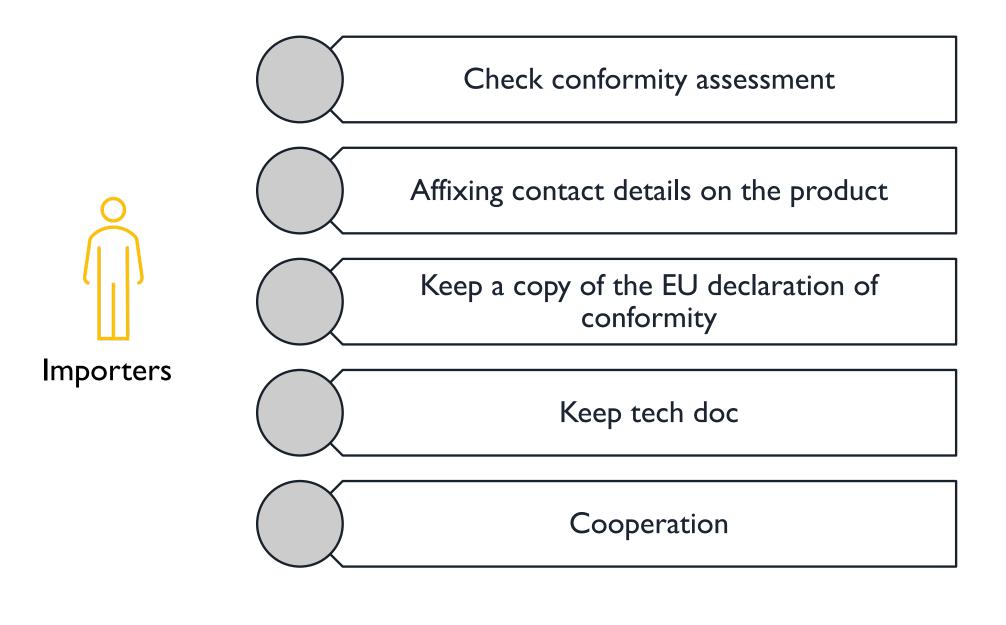
มีช่องทางสั่งการโดยมนุษย์ที่เหมาะสม (Human Oversight)

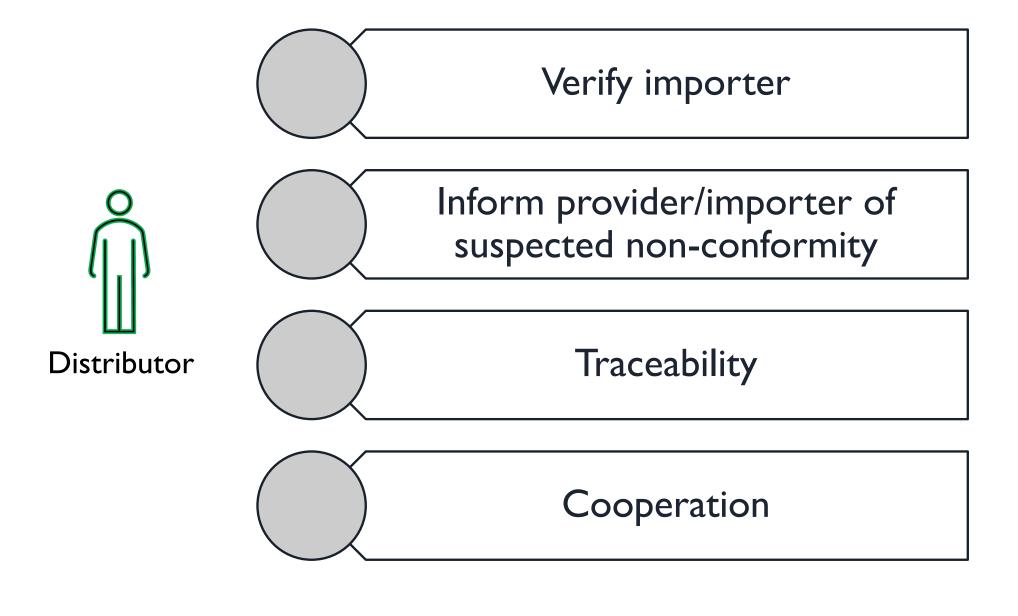


มีความแม่นยำ ผ่านการทดสอบ และความมั่งคงทางไซเบอร์ที่เหมาะสม (Accuracy, Robustness and Cybersecurity)









PUNISHMENT

Fine < EUR 30M / 6% worldwide annual turnover

Max only when it's prohibited Al

SMEs: Fine < 3% Annual turnover

<250 employees or < EUR 50M

3 KEY TAKEAWAYS

Different aspects of ethics

Not all ethical aspects are equal.

Many ways to make people 'care': Risk-based

THANK YOU

PEERAPAT.CHOKESUWATTANASKUL@GMAIL.COM