Centre for Information Policy Leadership
Hunton Andrews Kurth LLP

ico.
Information Commissioner's Office

CIPL and UK Information Commissioner's Office (ICO)

Joint Roundtable on the ICO Auditing Framework for AI

18 September 2019 - London

10:00 am – 4pm

Agenda

- 10:00 **Registration and Coffee**
- 10:15 **Opening Remarks**
- 10:30 **Session I: Accountable AI Procurement**
- 12:00 **Lunch**
- 13:00 **Session II: AI Data Protection Impact Assessments (DPIAs)**
- 14:30 **Session III: Data Minimization and Purpose Limitation**
- 16:00 **End of Roundtable**

Opening Remarks

Bojana Bellamy
President, CIPL

Ali Shah
Head of Technology Policy, ICO

Centre for Information Policy Leadership
Hunton Andrews Kurth LLP

A Global Privacy and Security Think Tank

BRIDGING REGIONS
BRIDGING INDUSTRY & REGULATORS
BRIDGING PRIVACY AND DATA DRIVEN INNOVATION

ACTIVE GLOBAL REACH

75+ Member Companies

We **INFORM** through publications and events

We **NETWORK** with global industry and government leaders

5+ Active Projects & Initiatives

We **SHAPE** privacy policy, law and practice

We **CREATE** and implement best practices

20+ Events annually

15+ Principals and Advisors

ABOUT US

- The Centre for Information Policy Leadership (CIPL) is a global privacy and security think tank
- Based in Washington, DC, Brussels and London
- Founded in 2001 by leading companies and Hunton Andrews Kurth LLP
- CIPL works with industry leaders, regulatory authorities and policy makers to develop global solutions and best practices for data privacy and responsible use of data to enable the modern information age

Twitter.com/the_cipl
https://www.linkedin.com/company/centre-for-information-policy-leadership
www.informationpolicycentre.com

2200 Pennsylvania Ave NW
Washington, DC 20037
Park Atrium, Rue des Colonies 11
1000 Brussels, Belgium
30 St Mary Axe
London EC3A 8EP

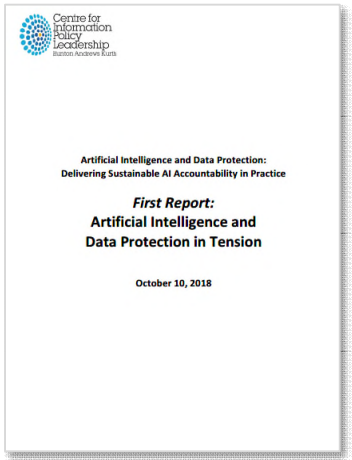
CIPL Project on Accountable AI

CIPL Project on Artificial Intelligence and Data Protection:
Delivering Sustainable AI Accountability in Practice

https://www.informationpolicycentre.com/ai-project.html

Status:

- First Report - **Artificial Intelligence and Data Protection in Tension** (October 2018) https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_ai_first_report_-_artificial_intelligence_and_data_protection_in_te....pdf
- Second Report in progress (estimated release October 2019)


Artificial Intelligence and Data Protection:
Delivering Sustainable AI Accountability in Practice

First Report:
Artificial Intelligence and Data Protection in Tension

October 10, 2018

First Report

Describes in clear and understandable terms:

(1) What AI is and how it is being used all around us today;

(2) The role that personal data plays in the development, deployment and oversight of AI; and

(3) The opportunities and challenges presented by AI to data protection laws and norms.

Disclaimer *- These slides have been prepared by CIPL. They are not endorsed by, and do not represent the views of, the Information Commissioner's Office.*

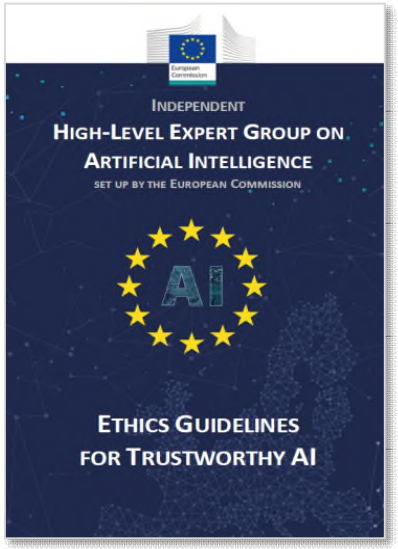
AI and Machine Learning: Challenges and Tensions with Data Protection Principles

Challenges associated with AI

Fairness | Ethical Issues | Public Trust | Legal Compliance | Tensions

Data Protection Requirements	Tensions To Resolve	Artificial Intelligence
Collection limitation / Data minimisation		Needs sufficient volumes of data for research, analysis, operation, training and to avoid bias
Purpose specification & Use limitation		Uses data for new and unforeseen purposes beyond original scope
Legal basis for processing		Insufficient/limited variety of legal bases may undermine full range of AI applications
Retention limitation		Needs to retain for AI training, deployment and oversight
Transparency		Operates in a black box and may produce unexplainable and unanticipated outcomes
Individual rights		Cannot always facilitate access, correction or explanation of the logic involved
Rules on ADM		Based on ADM & No human involvement

Guidance from the European Commission High Level Expert Group on AI



Identifies the **ethical principles** that must be respected in the development, deployment and use of AI systems:

- Respect for human autonomy, prevention of harm, fairness and explicability
- Pay attention to more vulnerable groups (children, disabled individuals, employees, consumers)
- Acknowledge that in spite of substantial benefits, AI systems also pose certain risks and wider impacts on society

Provides a **Trustworthy AI assessment list** to operationalize key requirements – test through piloting process until December 1st

Provides **seven requirements** to realize Trustworthy AI (technical and non-technical means):

- Human agency and oversight
- Technical robustness and safety
- Privacy and data governance
- Transparency
- Diversity, non-discrimination and fairness
- Environmental and societal well-being
- Accountability

Centre for Information Policy Leadership
Hunton Andrews Kurth LLP

The HLEG Guidelines and GDPR

Key requirements of Trustworthy AI	Overlap with GDPR provisions
Human Agency and Oversight	Legitimate interest balancing test (**art. 6(1)(f)**)/ Transparency (**art. 13** & **14**)/ ADM (**art. 22**) and Right to obtain human intervention (**art. 22(3)**) / Risk assessment and DPIA (**art. 35**)
Technical Robustness and Safety	Security (**art. 32**) / Risk assessment and DPIA (**art. 35**) / Data accuracy (**art. 5(1)(d)**)
Privacy and Data Governance	Data protection principles (**art. 5**) / Legal grounds for processing (**art. 6**)/ Legal grounds for sensitive data (**art. 9**)/ Rights of the data subject (Chapter III) and in particular Transparency (**art. 13** & **14**) and Right to information on ADM and logic involved (**art. 15(1)(h)**) and Right not to be subject to an ADM decision (**art. 22**) and right to human intervention (**art. 22(3)**) / Accountability (**art 24(3)**) / Data protection by design (art. 25)/Processor due diligence (**art. 28(1)**) / Security (**art. 32**) / DPO (**art. 37 & 38**)
Transparency	Transparency (**art. 13 & 14**)/ ADM (**art. 22**)
Diversity, Non-Discrimination and Fairness	Fairness Data protection principle (**art. 5.1(a)**) / Risk assessment and DPIA (**art. 35**) / Right to information on ADM and logic involved (**art. 15(1)(h)**)
Societal and environmental wellbeing	Risk assessment and DPIA (**art. 35**) / Transparency (**art. 13** & **14**)
Accountability	Accountability (**art 5(2)** & **24(3)**) / Risk assessment and DPIA (**art. 35**) / Processor due diligence (**art. 28(1)**) / DPO (**art. 37** & **38**)

Implementing Accountability

Organizations must be able to demonstrate accountability – internally and externally

Accountability is not static, but dynamic, reiterative and a constant journey

Accountability requires comprehensive privacy programs that translate legal requirements into risk-based, verifiable and enforceable corporate practices and controls

Company values and business ethics shape accountability

What Could an Accountable AI Governance Model Look Like?

Note: this AI framework sits on top of legal data protection norms and companies' ethics and values

- Public commitment and tone from the top to respect ethic, values, specific principles in AI development
- Institutionalised AI processes and decision-making
- Internal Code of Ethics rules
- AI / Ethics/ Review /Oversight Boards, Councils, Committees (internal or external)
- Appointing Board member for AI oversight
- Appointing Responsible AI Lead/ Officer
- Privacy/ RAI engineers and champions

What Could an Accountable AI Governance Model Look Like?

- Understand AI purpose and use case in business/ processes – for decision making, or input into decision, or other
- Understand impact on individuals
- Algorithmic Impact Assessment
- Fairness assessment tools
- Ethics Impact Assessment
- Broader human rights impact assessment
- DPIA for high risk processing
- Consider anonymisation techniques
- Document trade-offs (e.g. accuracy- data minimisation, security – transparency, impact on few – benefit to society)

What Could an Accountable AI Governance Model Look Like?

- High level principles for AI - how to design, use, sell
- Assessment questions and procedures
- Accountability measures for 2 stages – training and decision taking
- White, black and gray lists of AI use
- Evaluate the data and against the purpose - quality, provenence, personal or not, syntethic, in-house or external sources,
- Verification of data input and output;
- Algorithmic bias – tools to identify, monitor and test, including sensitive data in data sets to avoid bias
- Pilot testing AI models before release
- Use of encrypted data or synthetic data in some AI / ML models
- Use of high quality but smaller data sets
- Federated AI learning models (data doesn't leave device)
- Special considerations for companies creating and selling AI models, software, applications
- Checklists for business partners using AI tech and tools
- Using external tools, guidelines, self-assessment checklists

What Could an Accountable AI Governance Model Look Like?

- Different needs for transparency to individuals, regulators, business /data partners and internally to engineers, leadership
- Explainability is part of transparency and fairness
- Transparency trail – explainability of decision and broad workings of algorithm + more about the process than the technology + what factors + what testing to be fair + accountability for impact of decisions on a person's life + what extent of human oversight
- Explain that it is a AI/ ML decision, if possibility for confusion (Turing test)
- Provide counterfactual information
- Differentiated and flexible transparency - linked to context, audience/users, purpose of explainability and risk, severity of harm - prescriptive lists of transparency elements is not helpful
- Understand customers' expectations and deploy based on their readiness to embrace AI – tiered transparency
- From black box to glass box – looking at the data as well as algorithm /model; aspiration of explainability helps understand the black box and builds trust

What Could an Accountable AI Governance Model Look Like?

- Data scientist training, including how to avoid and address bias
- Cross functional training – privacy professionals and engineers
- Ad hoc and functional training
- Fairness training
- Ethics training
- Uses cases where problematic AI deployment has been halted
- Role of "Translators" in organisations, esplaining impact and workings of AI

What Could an Accountable AI Governance Model Look Like?

- Purpose of AI determines how much human intervention is required
- Human in the loop - in design, in oversight, in redress
- Human understanding of the business and processes using AI
- Human development of software and processes
- Human audit of input and output
- Human review of individual decisions
- Ongoing monitoring, validation and checks
- Oversight committees even in design stage
- Redress to a human, not to a bot
- Monitoring the eco-system from data flow in, data process and data out
- Reliance on different audit techniques
- Version control and model drift , tracking of black box, algirthms by engeeers
- RACI models for human and AI interaction

What Could an Accountable AI Governance Model Look Like?

- Complaints-handling
- Redress mechanisms for individuals to remedy AI decision
- Feedback channel
- Internal supervision of AI deployment

Session I : Accountable AI Procurement

Session I – Speakers & Discussion Questions

Moderator

Bojana Bellamy
CIPL President

Topic Introduction

ICO

Provocateur

Daniel Schoenberger
Head of Legal Switzerland & Austria, Google

❖ Session I: Accountable AI Procurement

- *What are the risks and challenges of using third party developed and trained AI?*
- *How do the concepts of controller and processor fit in the context of AI and with the notions of developer/vendor? user/buyer?*
- *What does supplier due diligence look like in the AI supply chain?*
- *How to address the risks of relying on third-party developed AI systems?*
- *What is the chain of accountability ?*

CIPL and ICO Roundtable on the ICO Auditing Framework for AI

Session 1: Accountable AI Procurement
London, 18 September 2019

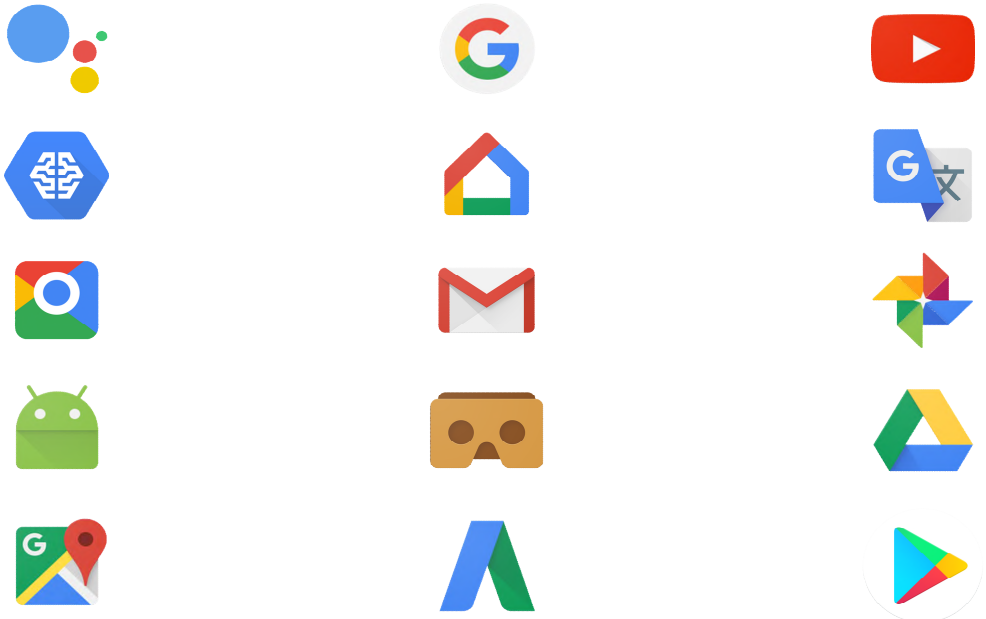
Daniel Schönberger

Attorney at Law, LLM (Edinburgh)

Head of Legal Google Switzerland & Austria

Google Making AI

AI is used across Google products



Google AI Principles

AI should:

1. be socially beneficial
2. **avoid** creating or reinforcing **unfair bias**
3. be built and tested for safety
4. be **accountable** to people
5. incorporate **privacy** design principles
6. uphold high standards of scientific excellence
7. **be made available** for uses that accord with these principles

applications we will not pursue:

1. likely to cause overall harm
2. principal purpose to direct injury
3. surveillance violating internationally accepted norms
4. purpose contravenes international law and human rights

Responsible AI Practices

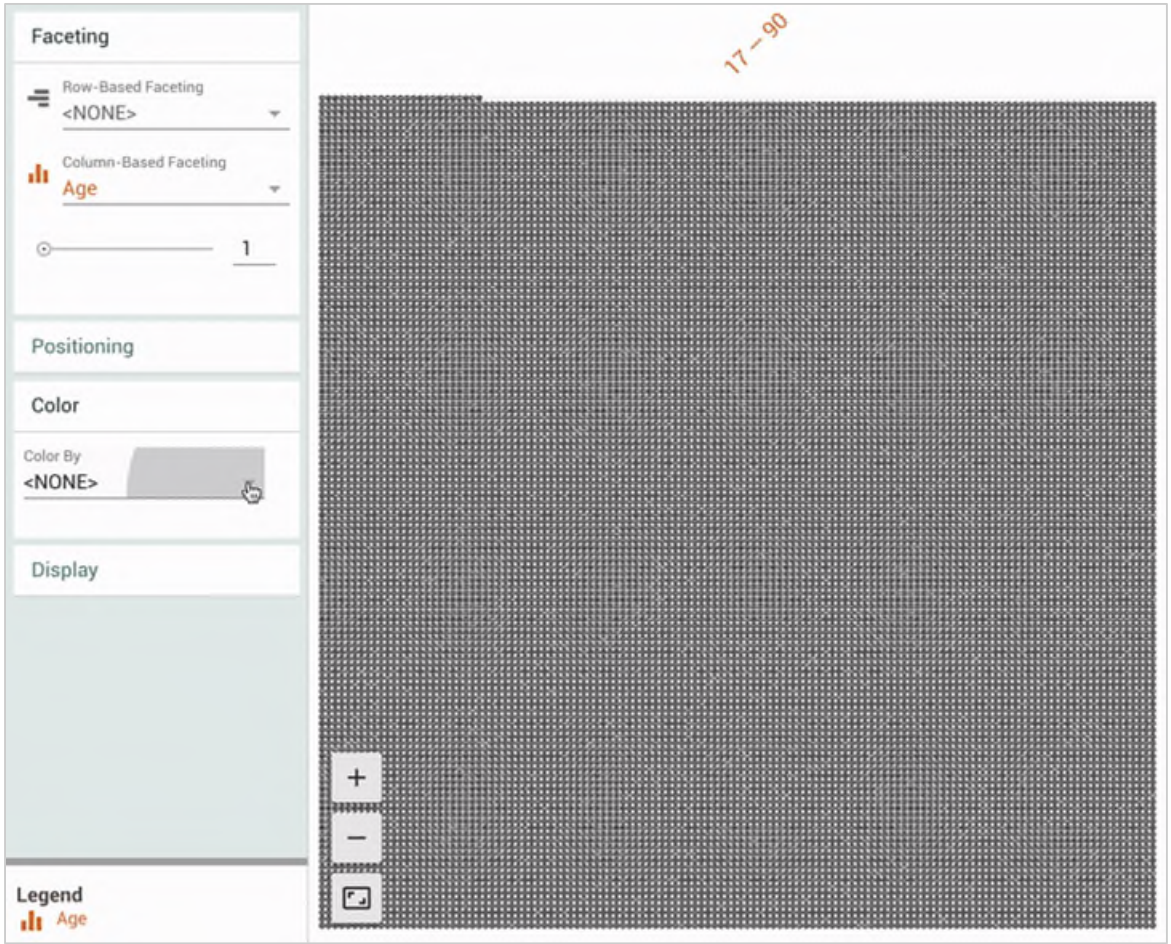
- Fairness
- Interpretability
- Privacy
- Security

https://ai.google/education/responsible-ai-practices

Explore and test for inclusion

Facets (open source):

“Debug your data before you debug your model”



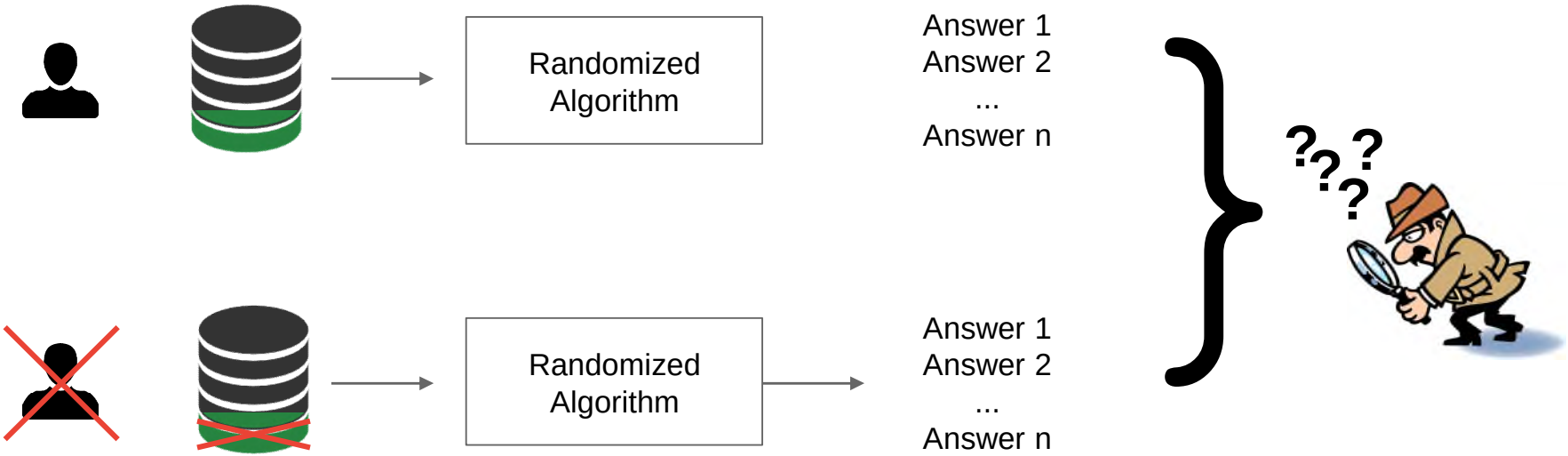
Visualisation

Safer Analytics

Tools to redact data for safer analytics or sharing

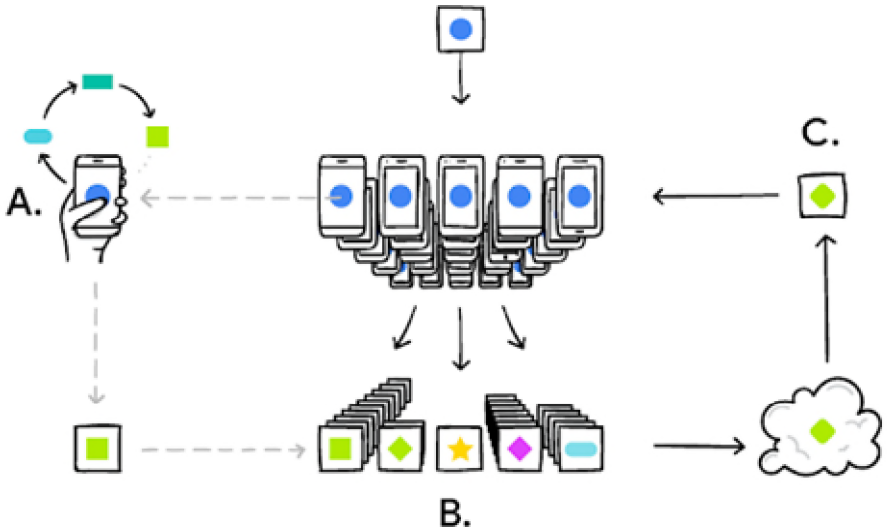
ID	Job Title	Phone	Comments
359740	Senior Engineer	307-964-0673	Please email them at jane@imadethisup.com
981587	VP, Engineer	713-910-6787	none
394091	Lawyer	692-398-4146	Updated phone to: 692-398-4146
986941	Senior Ops Manager	294-967-5508	none
490456	Junior Ops Manager	791-954-3281	Tried to verify account with their SSN 222-44-5555

Differential Privacy



Federated Learning

Collaborative machine learning without centrally stored training data



7 AI should be made available for uses that accord with these principles

Google will work to limit potentially harmful or abusive applications of AI. When considering selling or distributing AI technology that could foreseeably be misused, we take into account:

- whether the technology is *generally available or unique* to Google
- how *adaptable* it could be to harmful use, and *scale* of impact
- *nature of Google's involvement* eg: providing general-purpose tools vs developing custom solutions

☑ Cities using AI to create smarter traffic management applications

☑ Use of AI to detect IEDs, landmines or in search-and-rescue operations

☑ Use of AI to allow military to improve treatment of combat wounds

☒ Use of AI for predictive policing or sentencing without appropriate safeguards

Governance

Culture

Engrain the spirit of the Principles into everyday behaviour

- Custom training and ethics case studies; discussion of best practices

Process

Integration into existing launch and deal review processes

- Eg: Privacy working group and ML fairness teams will assess relevant issues in context of new tools that incorporate AI

Integration of oversight mechanisms into operating processes

- Eg: Google's Trust & Safety team pilot initiative to provide expert hands-on help in checking for possible bias is now being rolled out company-wide

AI Review Councils

Formally designated bodies to assess challenging issues

- Advanced Technology Review Council
- A diverse group representing international, cross-functional points of view that can provide a perspective beyond immediate commercial considerations

Thank you!

Centre for Information Policy Leadership
Hunton Andrews Kurth LLP

The HLEG Guidelines and Procuring AI responsibly

Guidelines on Trustworthy AI	Assessment List of Trustworthy AI
• Trustworthy AI requires a holistic and systemic approach, encompassing the trustworthiness of **all actors and processes** • Deployers of AI should ensure that the systems they use and procure to developers **meet the requirements of trustworthy AI** • Top management **evaluates the AI systems' procurement** and serves as an **escalation board** when critical concerns are detected • Procurement department ensures that the process to procure AI-based products or services includes **a check of Trustworthy AI**.	• In case of AI system's development, did you clearly communicate characteristics, limitations and potential shortcomings of the AI system to whoever is deploying it into a product or service? • Did you establish processes for third parties (e.g. suppliers, consumers, distributors/vendors) to report potential vulnerabilities, risks or biases in the AI system? • Did you establish mechanisms that facilitate the system's auditability, such as ensuring traceability and logging of the AI system's processes and outcomes?

Session II : AI Data Protection Impact Assessments

Session II – Speakers & Discussion Questions

Moderator

Fred Cate

Senior Policy Advisor, CIPL

Topic Introduction

ICO

Provocateur

Steve May

EU Data Protection Officer, Microsoft

❖ Session II : AI Data Protection Impact Assessments

- *How organisations can use DPIAs in AI projects?*
- *What is the appropriate risk assessment framework?*
- *Should the GDPR DPIA be adapted to AI contexts?*
- *Should both risks and benefits be balanced and how?*
- *At which stage should a DPIA take place (training phase, use phase)?*

Centre for Information Policy Leadership
Hunton Andrews Kurth LLP

The HLEG Guidelines and AI DPIAs

Guidelines on Trustworthy AI	Assessment List of Trustworthy AI
• A **fundamental rights impact assessment** must be done prior to the system's development; It should include an evaluation of whether those risks can be reduced or justified in a democratic society. • The level of **safety measures** required depends on the magnitude of the risk posed by an AI system, which in turn depends on the system's capabilities. Where it can be foreseen that the development process or the system itself will pose particularly high risks, it is crucial for safety measures to be **developed and tested proactively**	• Did you carry out a **fundamental rights impact assessment** ? • Did you identify and document **potential trade-offs** made between the different principles and rights? • Did you put any process in place **to measure and assess risks and safety?** • Did you identify potential **safety risks** of foreseeable uses of the technology, including accidental or malicious misuse? • Did you estimate **the likely impact of a failure of your AI system when it provides wrong results**, becomes unavailable, or provides societally unacceptable results (for example discrimination)?

Session III : Data Minimisation and Purpose Limitation

Session III – Speakers & Discussion Questions

Moderator

Fred Cate
Senior Policy Advisor, CIPL

Topic Introduction

ICO

Provocateur

Vivienne Artz
Chief Privacy Officer, Refinitiv

❖ Data Minimization and Purpose Limitation

- *How do these fundamental data protection principles have to be interpreted in the AI context?*
- *How to implement these principles during the algorithmic training phase?*
- *Is the training of an AI model a specific distinct purpose?*
- *How do these principles also apply during the subsequent use of the trained algorithm?*

Centre for Information Policy Leadership
Hunton Andrews Kurth LLP

The HLEG Guidelines and Data Minimisation and Purpose Limitation

Guidelines on Trustworthy AI	Assessment List of Trustworthy AI
• Prevention of harm to privacy necessitates adequate data governance that covers **relevance of the data in light of the domain in which the AI systems will be deployed**, its access protocols and the capability to process data in a manner that **protects privacy**. • AI systems must guarantee **privacy and data protection** throughout a system's entire lifecycle (information initially provided by the user and information generated during the interaction with the system) • Data collected about individuals should not be **used to unlawfully** or unfairly discriminate against them.	• Did you assess the **type and scope of data** in your data sets (for example whether they contain personal data)? • Did you consider ways to develop the AI system or train the model **without or with minimal use of potentially sensitive or personal data**? • Did you build in mechanisms for **notice and control** over personal data depending on the use case (such as valid consent and revocation)? • Did you take measures to **enhance privacy**, such as encryption, anonymisation and aggregation? • Did you **clarify the purpose** of the AI system and who or what may benefit from the product/service?

Centre for Information Policy Leadership
Hunton Andrews Kurth LLP

Thank You

Bojana Bellamy

President

bbellamy@HuntonAK.com

Markus Heyder

Vice President & Senior Policy Advisor

mheyder@HuntonAK.com

Nathalie Laneret

Director of Privacy Policy

nlaneret@HuntonAK.com

Sam Grogan

Global Privacy Policy Analyst

sgrogan@HuntonAK.com

Matt Starr

Privacy and Public Policy Manager

mstarr@HuntonAK.com

Giovanna Carloni

Global Privacy Policy Manager

gcarloni@HuntonAK.com

Centre for Information Policy Leadership
www.informationpolicycentre.com

Hunton's Information Security Law Blog
www.huntonprivacyblog.com

@THE_CIPL

linkedin.com/company/centre-for-information-policy-leadership

Appendices

Centre for Information Policy Leadership
Hunton Andrews Kurth LLP

GDPR and AI

GDPR aims to be technology neutral and applies fully to the use of personal data in AI

Several GDPR provisions are specifically relevant for AI:

Art. 5(1)(a): Lawful, fair and transparent processing

Art. 13(2)(f): Inform individuals of existence of ADM and meaningful information about logic involved (data collected directly)

Art. 14(2)(g): Inform individuals of existence of ADM and meaningful information about logic involved (data collected indirectly)

Art. 15(1)(h): Right to access information about existence of ADM and meaningful information about logic involved

Art. 22: Right not to be subject to a decision based on solely ADM producing legal/similarly significant effects

Art. 22(3): Right to obtain human intervention and contest decision

Art. 35: Conduct a DPIA for high risk processing, in particular when using new technology

Art. 35(3)(a): DPIA required in the case of Art. 22 ADM

Centre for Information Policy Leadership
Hunton Andrews Kurth LLP

Growing AI Regulatory Guidance

- **UK ICO**, Big Data, Artificial Intelligence, Machine Learning and Data Protection (September 2017) -https://ico.org.uk/media/for-organisations/documents/2013559/big-data-ai-ml-and-data-protection.pdf
- **CNIL**, How Can Humans Keep the Upper Hand?: The Ethical Matters Raised by Algorithms and Artificial Intelligence (December 2017) - https://www.cnil.fr/sites/default/files/atoms/files/cnil_rapport_ai_gb_web.pdf
- **Datatilsynet** (Norwegian Data Protection Authority), Artificial Intelligence and Privacy (January 2018) - https://www.datatilsynet.no/globalassets/global/english/ai-and-privacy.pdf
- **ICDPPC**, Declaration on Ethics and Data Protection in AI (October 2018) - https://icdppc.org/wp-content/uploads/2018/10/20180922_ICDPPC-40th_AI-Declaration_ADOPTED.pdf
- **Government of Canada**, AI Guiding Principles – Exploring the Future of Responsible AI in Government (November 2018) - https://www.canada.ca/en/government/system/digital-government/modern-emerging-technologies/responsible-use-ai.html#toc1
- **European Commission High Level Expert Group on AI**, Draft Ethics Guidelines for Trustworthy AI (December 2018) - https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=57112
- **Singapore PDPC**, A Proposed Model Artificial Intelligence Governance Framework (January 2019) - https://www.pdpc.gov.sg/Resources/Model-AI-Gov
- **Council of Europe**, Guidelines on Artificial Intelligence and Data Protection (January 2019) - https://rm.coe.int/guidelines-on-artificial-intelligence-and-data-protection/168091f9d8
- **OECD** Council Recommendation on Artificial Intelligence (May 2019) - https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449#_ga=2.251645126.1726117956.1559308992-1610692363.1559308992
- **UK ICO**, Project Expl*AI*n Interim Report (June 2019) - https://ico.org.uk//media/2615039/project-explain-20190603.pdf

Centre for Information Policy Leadership
Hunton Andrews Kurth LLP

Growing AI Regulatory Guidance

- **UK ICO**, AI Auditing Framework (Consultation paper expected January 2020) (multiple blog posts on various AI issues available now) - https://ai-auditingframework.blogspot.com/
- **An Overview of the AI Auditing Framework** (26 March 2019) - https://ai-auditingframework.blogspot.com/2019/03/an-overview-of-auditing-framework-for_26.html
- **Automated Decision Making: the role of meaningful human reviews** (12 April 2019) - https://ai-auditingframework.blogspot.com/2019/04/automated-decision-making-role-of.html
- **Accuracy of AI system outputs and performance measures** (2 May 2019) -https://ai-auditingframework.blogspot.com/2019/05/accuracy-of-ai-system-outputs-and.html
- **Known security risks exacerbated by AI** (23 May 2019) - https://ai-auditingframework.blogspot.com/2019/05/known-security-risks-exacerbated-by-ai.html
- **When it comes to explaining AI decisions, context matters** (3 June 2019) - https://ai-auditingframework.blogspot.com/2019/06/when-it-comes-to-explaining-ai.html
- **Human bias and discrimination in AI systems** (25 June 2019) - https://ai-auditingframework.blogspot.com/2019/06/human-bias-and-discrimination-in-ai.html
- **Trade-offs** (25 July 2019) - https://ai-auditingframework.blogspot.com/2019/07/trade-offs.html
- **Fully automated decision making AI systems: the right to human intervention and other safeguards** (5 August 2019) - https://ai-auditingframework.blogspot.com/2019/08/fully-automated-decision-making-ai.html
- **Data minimisation and privacy-preserving techniques in AI systems** (21 August 2019) - https://ai-auditingframework.blogspot.com/2019/08/data-minimisation-and-privacy_21.html