

Background

UCL

- UCL Computer Science
 - National AI Certification Service
 - AI Audit and Assurance
 - Piloting
- Publications
 - [Towards Algorithm Auditing](#)
 - [AI Standards \[...\] Through Experimentation](#)
 - [AI Audit in Financial Services & Recruitment](#)
 - Etc..
- Collaborations

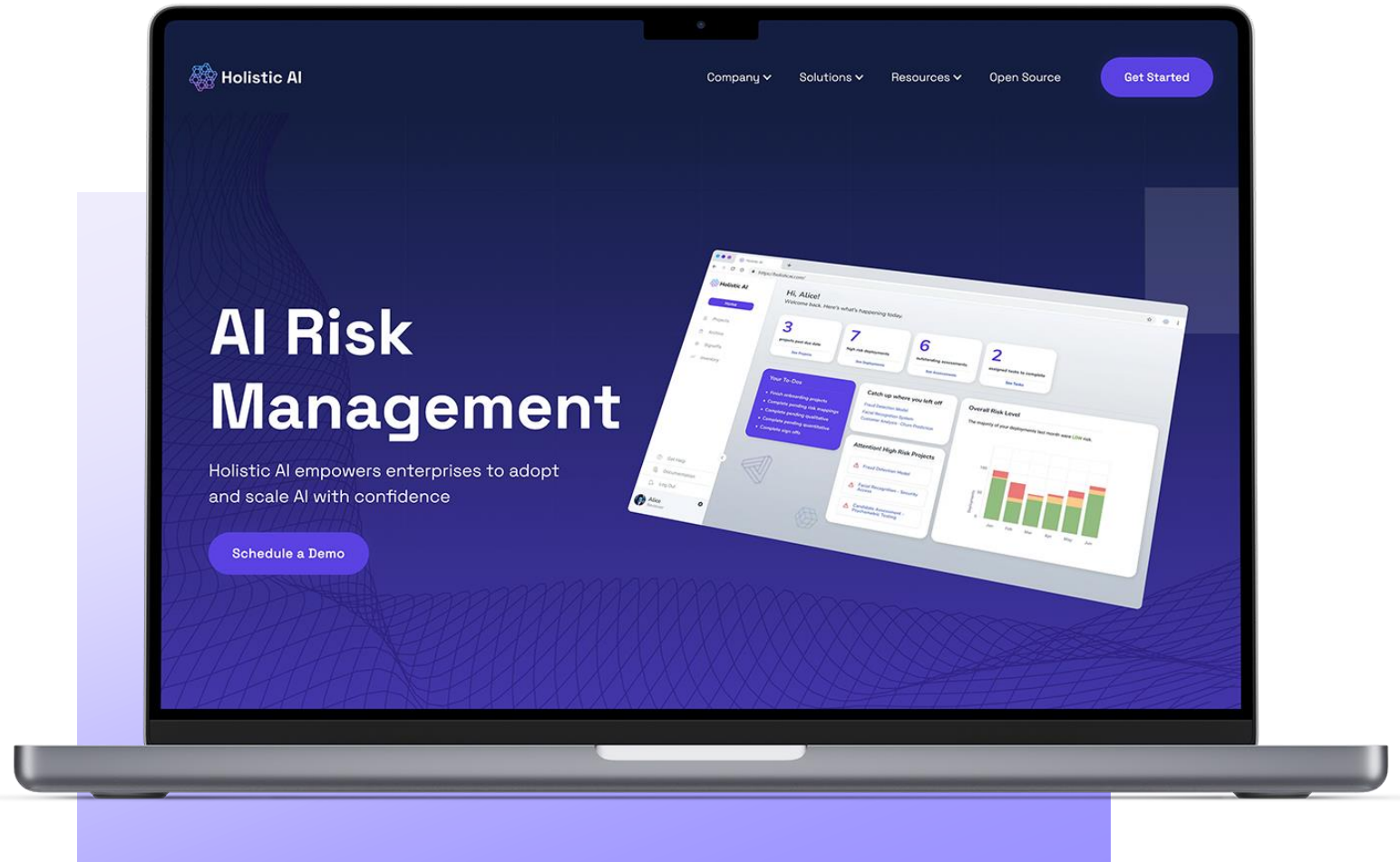


The
Alan Turing
Institute





Holistic AI



From Harm to Regulation



Harm

- High profile harm
- Moral Panic?
 - Opportunity vs harms

Knight Capital Says Trading Glitch Cost It \$440 Million

BY NATHANIEL POPPER AUGUST 2, 2012 9:07 AM 356

Runaway Trades Spread Turmoil Across Wall St.



VERNON PRATER Prior Offenses 2 armed robberies, 1 attempted armed robbery Subsequent Offenses 1 grand theft LOW RISK 3	BRISHA BORDEN Prior Offenses 4 juvenile misdemeanors Subsequent Offenses None HIGH RISK 8
DYLAN FUGETT LOW RISK 3	BERNARD PARKER HIGH RISK 10

JAMES RIVELLI Prior Offenses 1 domestic violence, aggravated assault, 1 grand theft, 1 petty theft, 1 drug trafficking Subsequent Offenses 1 grand theft LOW RISK 3	ROBERT CANNON Prior Offense 1 petty theft Subsequent Offenses None MEDIUM RISK 6
---	---





Demand for Action

- 100+ Statements of Ethics
- Significant activity



RESPONSIBILITIES >

Artificial Intelligence at Google: Our Principles

Google aspires to create technologies that solve important problems and help people in their daily lives. We are optimistic about the incredible potential for AI and other advanced technologies to empower people, widely benefit current and future generations, and work for the common good.

BLUEPRINT FOR AN AI BILL OF RIGHTS

MAKING AUTOMATED SYSTEMS WORK FOR THE AMERICAN PEOPLE

OSTP

NIST

Information Technology Laboratory

AI RISK MANAGEMENT FRAMEWORK



Policy

Responsible AI Standard

The Microsoft Responsible AI Standard is our internal playbook for responsible AI. It shapes the way in which we create AI systems, by guiding how we design, build, and test them.



Key Definitions

	Outputs Changes	Role of Humans	Automation	Technology
Information Commissioner's Office	-	-	Fully automated or with human in the loop	Algorithm-based technology
EU AI Act	Content, predictions, recommendations, or decisions	Can provide data and inputs	Elements of autonomy	Machine learning or logic- and knowledge-based approaches
OECD	Predictions, recommendations, or decisions	Define objectives	Varying levels of autonomy	Machine-based system
Canada's Artificial Intelligence and Data Act	Content, decisions, recommendation, or predictions	Human activity provides data	Autonomously or partly autonomously	Technological system that uses a genetic algorithm, neural network, machine learning, or other technique
California's proposed Amendments to employment regulations	Predictions, recommendations, or decisions	Define objectives	-	Machine learning system, where machine learning is an application of AI

Source: Lost in Transl(A)t(I)on: Differing Definitions of AI
<https://www.holisticai.com/blog/comparing-definitions-of-ai>

Some Concerns



Consistent Standards

- **Fundamental Principles**
- **Consistent Taxonomy for AI**
 - Widely accepted / shared
 - Ex. Consistent def. of AI (what falls in scope)
 - Ex. Life-cycle taxonomy
 - pre-dev., dev., deployed, monitoring

- **Generally Accepted Risk Framework**
 - Aggregated components
 - Source of risk
 - Bias, privacy, security, transparency ...
- **Ex. Facial Recognition**
 - Customer on-boarding
 - School playground



Technical Capabilities for Compliance

AI Inventory management: scalable & actionable insights

Ongoing monitoring: real-time & automated

Testing: technical assessment & risk mitigation

Market monitoring: automated AI incident tracking

Validation and testing: of business documentation

AI auditing: supplementary to enforcement action

Regulation driving Innovation



- **Market Monitoring**
 - Very difficult
 - Ex. Fraud-detection; Anti-Money Laundering
- **Compliance by Design**
 - Mandate Design Principles
 - Require Validation to check
 - Internal/External audit
 - Conformity Assessments
- **Empiricism**
 - Reporting (Non-mandatory, incentivised through recognition (implicit accreditation))
 - Evidence **base**
- **Outcomes focused and risk-based enforcement**
- **Regulatory sandboxes and pilot projects**
- **Regulatory guidance** to promote AI adoption (**ex. Mandate third-party** auditing etc.)



Holistic AI

Thanks

Dr Emre Kazim

UCL Computer Science
e.kazim@ucl.ac.uk

Holistic AI COO, Co-Founder
emre.kazim@holisticai.com