Greg Hutchins PE CERM

- QualityPlusEngineering.com
- 800Compete.com
- WorkingIt.com
- 503.233.1012 or Greg@workingit.com

Presentation Outline

1. Story: "Alas Poor Human"
2. AI Context
3. AI Decision Making – Social Context
4. Assurance & Accountability - Technical Context
5. Your Future: How to Thrive and Prosper

1. Story: "Alas. Poor Human'

Lafayette – Photo – London.
SARAH-BERNHARDT (HAMLET.)

Hamlet: "Alas, poor Yorick …."

AI Robot: "Alas, poor human …."

1
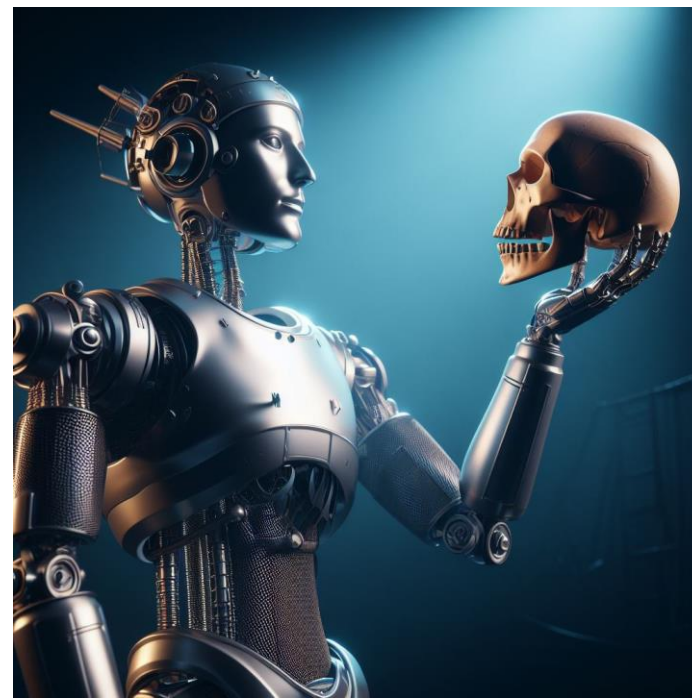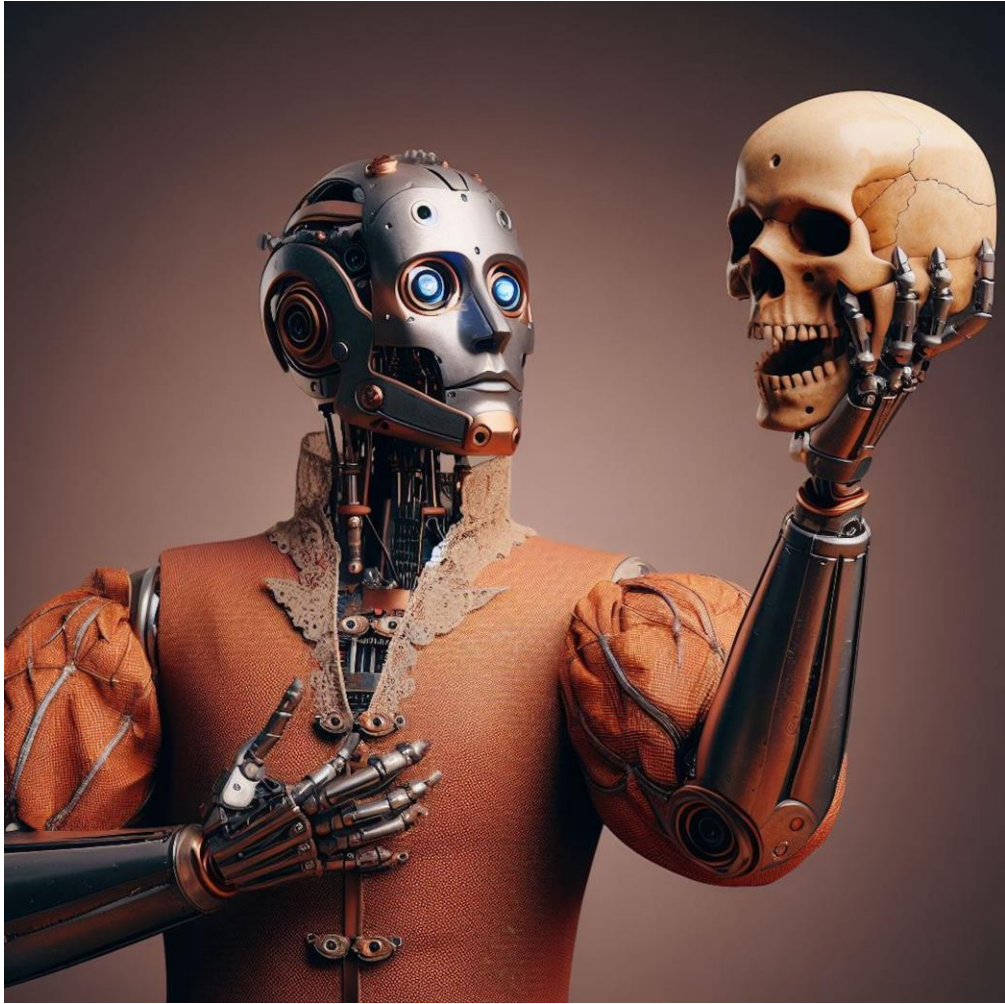

2

3

4
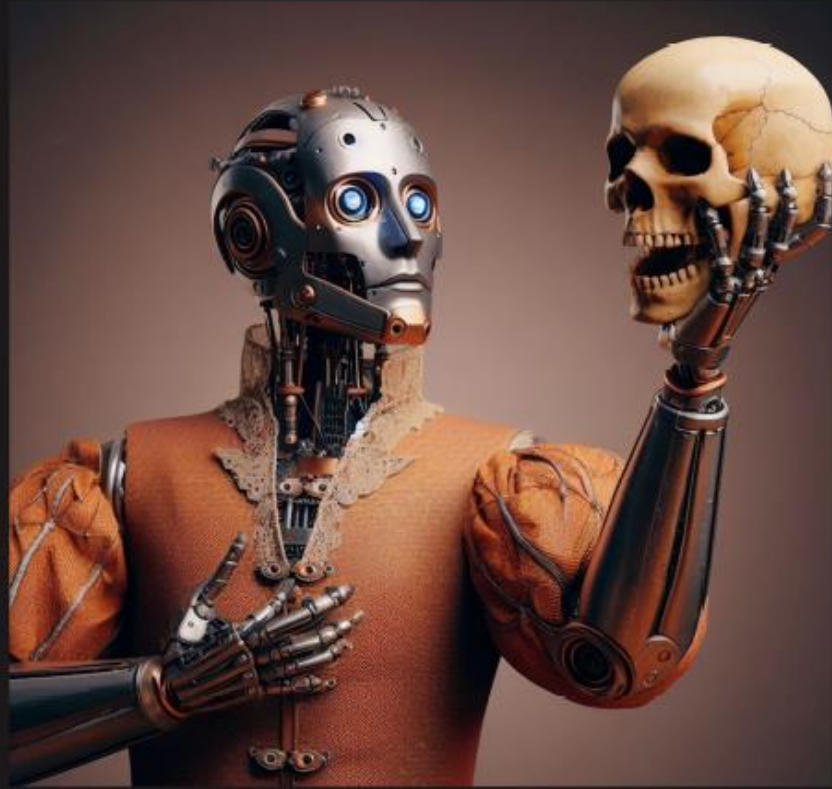
TRUST ME: AI RISK MANAGEMENT

GREG HUTCHINS

MARGAUX HUTCHINS

2. AI Context

Global Web 3.0 Market Size Expected to Reach $81 Billion by 2030 as Branding & Marketing Needs Increase

USD 1,811.8 billion — The global artificial intelligence market is expected to grow at a compound annual growth rate of 38.1% from 2022 to 2030 to reach USD 1,811.8 billion by 2030. — Artificial Intelligence Market Size Report, 2022-2030
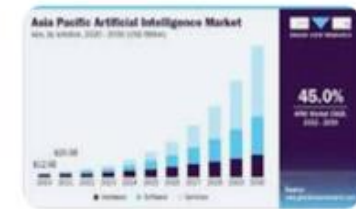
AI market in 2030 will be 0.6 to 5.0 times greater than UK GDP
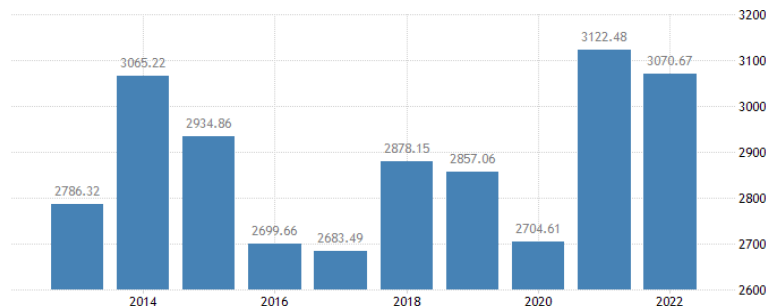
Sizing the prize — PwC's Global Artificial Intelligence Study: Exploiting the AI Revolution — $15.7tr Potential contribution to the global economy by 2030 from AI — +26% Up to 26% boost in GDP for local economies from AI by 2030

World Bank

What is AI?

- What is AI?
  - Artificial Intelligence (AI) is the **ability of a machine or system to mimic human intelligence to do task**s and can repeatedly improve itself based on data it collects.

- Why does AI matter to YOU?
  - Today, the amount of data that is generated, by both humans and machines, far outpaces humans' ability to absorb, interpret, and make complex decisions based on that data. **Artificial intelligence forms the basis for all computer learning and is the future of all complex decision making**.
  - **Source:** NetApp, 2022

AI Risk < = > Trust

- **Public policy:** AI is running ahead of policy makers.

- **Workers:** AI is already impacting employment, housing, and other decisions.

- **Professions:** AI is impacting the Future of Work from surveillance, product development, auditing, legal.

- **Companies:** AI is developed in many products. Development will require internal risk protocols.

- **Future issues:** Risk statutes, regulations, definitions, assurance, accountability & implementation are evolving

Autonomous Decision Making Risks

Discriminatory

- Healthcare access
- Tenant screening
- Criminal justice
- Facial recognition
- Employment screening
- Marketing

Discriminatory

- Incarceration
- Homelessness
- Job loss
- Equity-stripping
- Debilitating health

Picture This: For Your Organization

- New York City is issuing AI Bias Law impact 200,000 companies in January 2023.

- *Wall Street Journal* reports: "A New York City law that comes into effect in January will require companies to conduct audits to assess biases, including along race and gender lines, in the AI systems they use in hiring. Under New York's law, the hiring company is ultimately liable—and can face fines—for violations." [9]

LAWFLASH

NEW YORK CITY PROPOSES NEW RULES TO CLARIFY LAW ON EMPLOYERS' USE OF ARTIFICIAL INTELLIGENCE

October 04, 2022

Gothamist   NYC delays enforcement of law aimed at identifying bi...

NYC delays enforcement of law aimed at identifying bias in AI hiring tools

Decision Making Rules of Engagement

"Crush! Kill! Destroy!"

Killer Android IDAK Alpha 12, Lost in Space

John: *Jesus, you were gonna kill that guy.* The Terminator: *Of course: I'm a Terminator*

John: … You just can't go around killing people.

Terminator: Why?

Terminator 2: Judgment Day

What are **rules of engagement** **between humans and machines** for **autonomous decision making**?

3. AI Decision Making – Social Context

- AI problem solving & decision making is every where
- It is often black box decision making & can be unexplainable
- Bias gets baked in algorithmic decision making
- Rules of human & machine engagement are not defined
- Rules of engagement involve trust, fairness, explainability, and assurance

Fast Food Decision Making

Wages

Source: Linkedin comment

Automation

Source: Daily Caller, 2020

Autonomous Decision Making

Bus + Driver

Source: Pixabay, 2022

Bus + No driver

Source: SF Gate, 2022

Employment Decision Making

Human Interviewer

Source: Pexels, 2022

Black Box Algorithm Interviewer
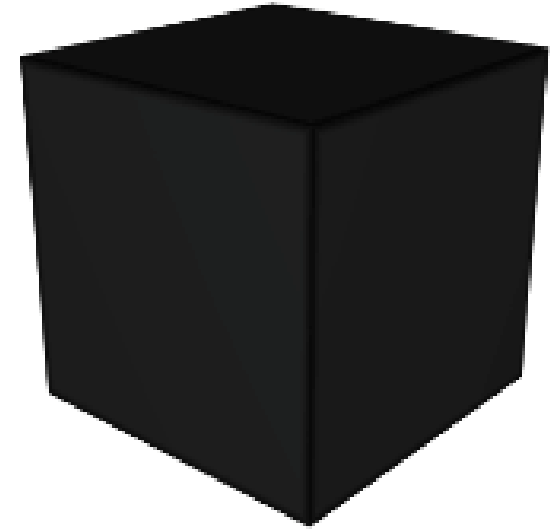
Source: Pexels, 2022

Medical Decision Making

Human Physician

Source: Pexels

AI Physician

How AI Is Changing Medical Imaging to Improve Patient Care
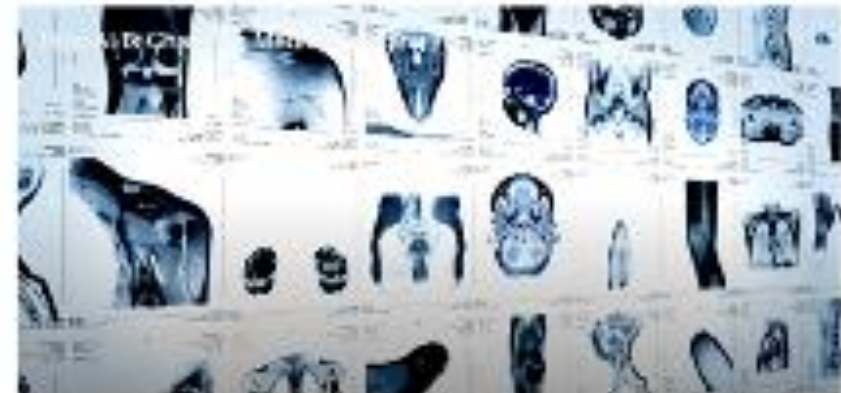
Source: Time Magazine 2022

Public Safety Decision Making

Human Policing

Source: Pexels, 2022

Machine Public Safety

Report: Oakland police consider use of possibly lethal robots

Source: SF Gate, 2022

Military Decision Making

Policing

Source: NY Post, Oct.26, 2022

Military

Source: Pixabay

4:Assurance/Accountability: Technical Context

Assuring these AI questions is the future for CQI and its quality professionals:

1. **What** are the standards to assure AI compliance and adherence?

2. **How** will AI assurance and accountability be audited?

3. **Who** will conduct AI audits and assessments including their qualifications?

US: 'What' Principles

USA: AI Bill of Rights Principles

Safe and Effective Systems

Algorithmic Discrimination Protections

Data Privacy

Notice and Explanation

Human Alternatives, Consideration, and Fallback

US: 'What' Standard

NIST AI-RMF 'What' Standard:

1. Framework applies to automated **risk based, decision-making** systems.

2. Is part of US Presidential executive order for US Federal agencies

Map — Context is recognized and risks related to context are identified

Measure — Identified risks are assessed, analyzed, or tracked

Govern — A culture of risk management is cultivated and present

Manage — Risks are prioritized and acted upon based on a projected impact
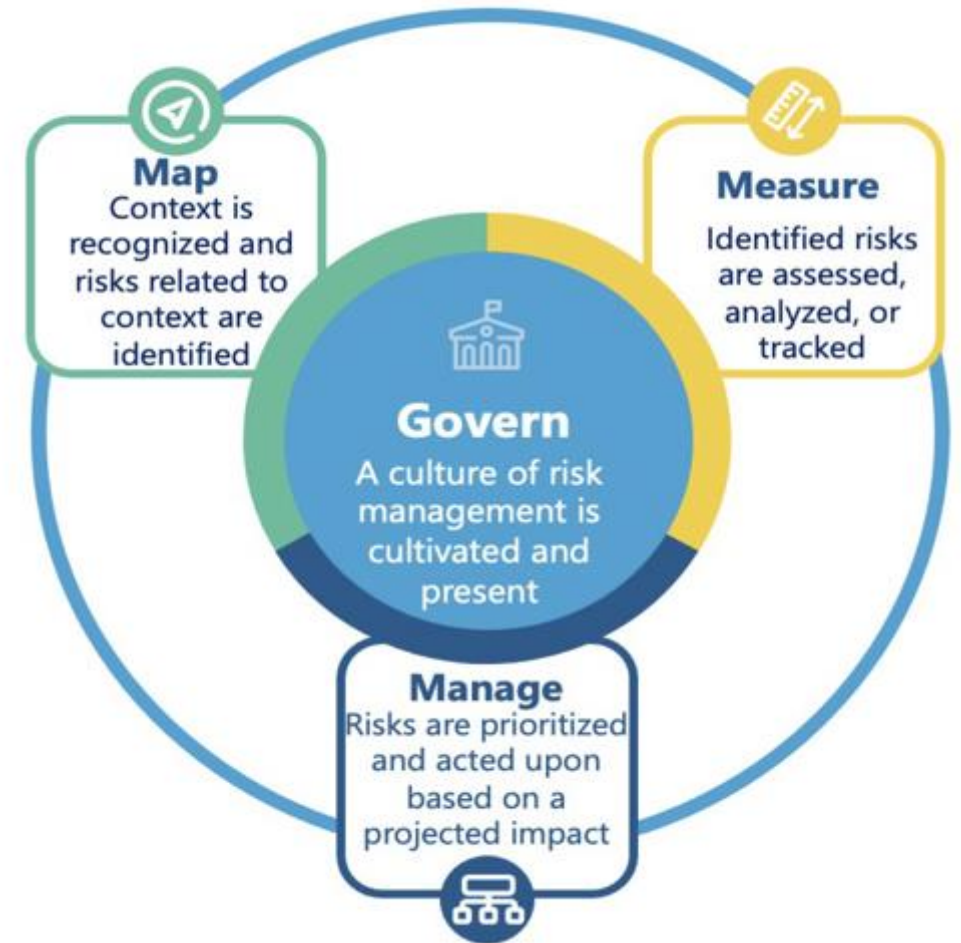
EU: 'What' Regulation

- EU AI Act risk based system for AI assurance & accountability
- 4 levels of AI risk:

1. Unacceptable risk. Threats to safety and people

2. **High risk. Infrastructure; enforcement; employment**

3. Limited risk. Generative, Chatbots etc

4. Limited risk. Video games, spam filters

EU 8 Areas Requiring Risk Assessment

1. Biometric identification and categorization of natural persons
2. **Management and operation of critical infrastructure**
3. Education and vocational training
4. Employment, worker management and access to self-employment
5. Access to and enjoyment of essential private services and public services and benefits
6. **Law enforcement**
7. Migration, asylum and border control management
8. Assistance in legal interpretation and application of the law

Canada: 'What'

- AI Data Act (AIDA)
- ISO 42001 (FDIS)
  - Context
  - Performance evaluation
  - AI risk assessment
  - AI risk treatment
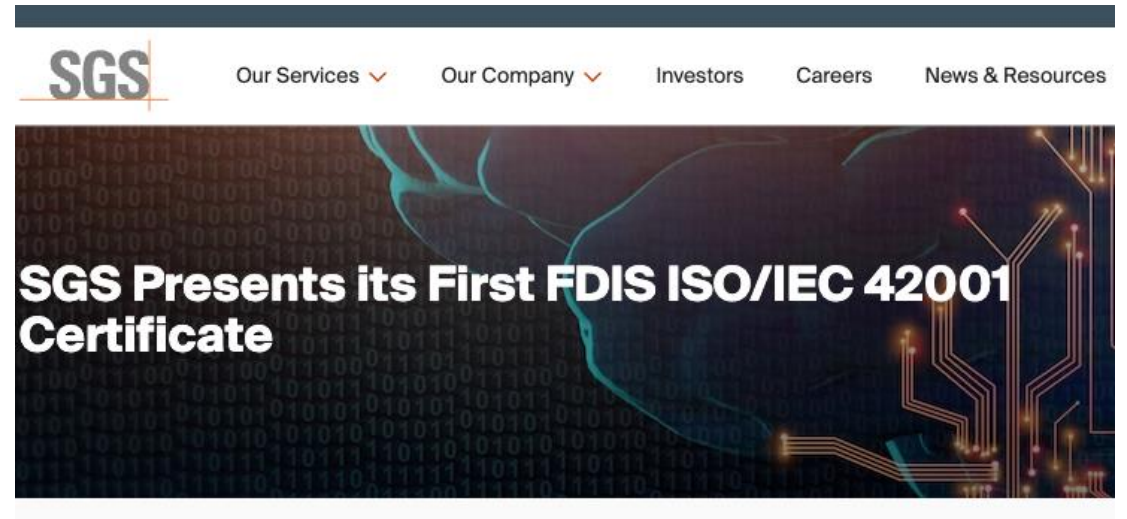
EU: 'How'

- Conformity Assessment

  - Process of determining whether a product, service, or system meets the requirements of a particular regulation or standard.

- CE product marking

  - Product water marking

EU: 'How' Conformity Assessment

- Management system assessment
    - First Party: Self declare compliance
    - Second party assessment
    - Third party certification - **ISO 42001 AI certification**

- EU AI Act
    - ISO 9001 – Quality Management System - Article 17
    - **Pay attention to above: May double # of ISO 9001 certifications**

Commercial 'How' Solutions:
AI Trust Labels

AI Nutrition Facts

Generative Journeys in Twilio Engage

Description
Accelerate campaign execution by using language prompts to describe a campaign journey. Marketers can apply additional filters but then let CustomerAI create the journey and target the appropriate audience.

Privacy Ladder Level	1
Feature is Optional	Yes
Model Type	Generative
Base Model	OpenAI - GPT-4

Trust Ingredients

Base Model Trained with Customer Data	No
Customer Data is Shared with Model Vendor	No
Journey prompts are NOT used for training OpenAI models	
Training Data Anonymized	N/A
Data Deletion	Yes
Journey inferences deleted after 30 days	
Human in the Loop	Yes
User sees output immediately in UI	
User must choose to publish journey	
Data Retention	30 days
Compliance	
Logging & Auditing	Yes
Guardrails	Yes
Input/Output Consistency	Yes

Other Resources
Learn more at: https://www.twilio.com/en-us/customer-ai

Learn more about this label at **nutrition-facts.ai**

US: 'Who'

Commercial 'Who': Quality + Engineering Assurance Model

- **Analytical.** analytical risk analyses

- **Assurance.** three levels of assurance:
  - Compliance. compliance audit
  - Assurance with professional engineering opinion
  - Assurance with insurance coverage

- **Forensic.** Evidentiary

Certificate of Conformance

This will certify that, on this date,

the United States Department of Homeland Security issued to

Quality Plus Engineering

An Oregon Limited Liability Company

a *Certification* for its

Critical Infrastructure Protection Forensics/Assurance/Analytics™ (CIP/FAA™)

as an *'Approved Product for Homeland Security'* under the

Support Anti-terrorism by Fostering Effective Technologies Act of 2002 (the SAFETY Act).

Jay M. Cohen

Under Secretary for Science and Technology

April 25, 2007

Date

5. Your Future: How to Thrive & Prosper

"**A.I. will likely eat into a lot of management and decision-making jobs that formerly required college-educated workers** or even workers with graduate credentials, such as lawyers. Hence, A.I. is not "more of the same.""

- **Source:** July 2022 paper, "The Labor Market Impacts of Technological Change: From Unbridled Enthusiasm to Qualified Optimism to Vast Uncertainty," Autor, quoted in NY Times November 2, 2022

INTERESTING ENGINEERING   NEWSLETTERS   Subscribe | Sign In

CEOs say 47% of workforce unprepared for the future of work

And a surprising percentage of CEOs believe that 'most' or 'all' of their own position could soon be completely automated or replaced by AI.

Sejal Sharma
Published: Oct 04, 2023 01:03 PM EST

Understanding Frontier AI Risks (roundtable discussions)

Delegates will break out to discuss the following topics with multi-disciplinary attendees. Conclusions from each session will be published at the end of the summit.

1. Risks to Global Safety from Frontier AI Misuse

Discussion of the safety risks posed by recent and next generation frontier AI models, including risks to biosecurity and cybersecurity.

2. Risks from Unpredictable Advances in Frontier AI Capability

Discussion of risks from unpredictable 'leaps' in frontier AI capability as models are rapidly scaled, emerging forecasting methods, and implications for future AI development, including open-source.

3. Risks from Loss of Control over Frontier AI

Discussion of whether and how very advanced AI could in the future lead to loss of human control and oversight, risks this would pose, and tools to monitor and prevent these scenarios.

4. Risks from the Integration of Frontier AI into Society

Risks from the integration of frontier AI into society include election disruption, bias, impacts on crime and online safety, and exacerbating global inequalities. Discussion will include measures countries are already taking to address these risks.

Improving Frontier AI Safety (roundtable discussions)

Delegates will break out to discuss the following topics with multi-disciplinary attendees. Conclusions from each session will be published at the end of the summit.

1. What should Frontier AI developers do to scale responsibly?

Multidisciplinary discussion of Responsible Capability Scaling at frontier AI developers including defining risk thresholds, effective model risk assessments, pre-commitments to specific risk mitigations, robust governance and accountability mechanisms, and model development choices.

2. What should National Policymakers do in relation to the risk and opportunities of AI?

Multidisciplinary discussion of different policies to manage frontier AI risks in all countries including monitoring, accountability mechanisms, licensing, and approaches to open-source AI models, as well as lessons learned from measures already being taken.

3. What should the International Community do in relation to the risk and opportunities of AI?

Multidisciplinary discussion of where international collaboration is most needed to both manage risks and realise opportunities from frontier AI, including areas for international research collaborations.

4. What should the Scientific Community do in relation to the risk and opportunities of AI?

Multidisciplinary discussion of the current state of technical solutions for frontier AI safety, the most urgent areas of research, and where promising solutions are emerging.

CQI Professionals

- **What:**
  - [Get ISO 42001 FDIS](#) (AI risk management system standard)
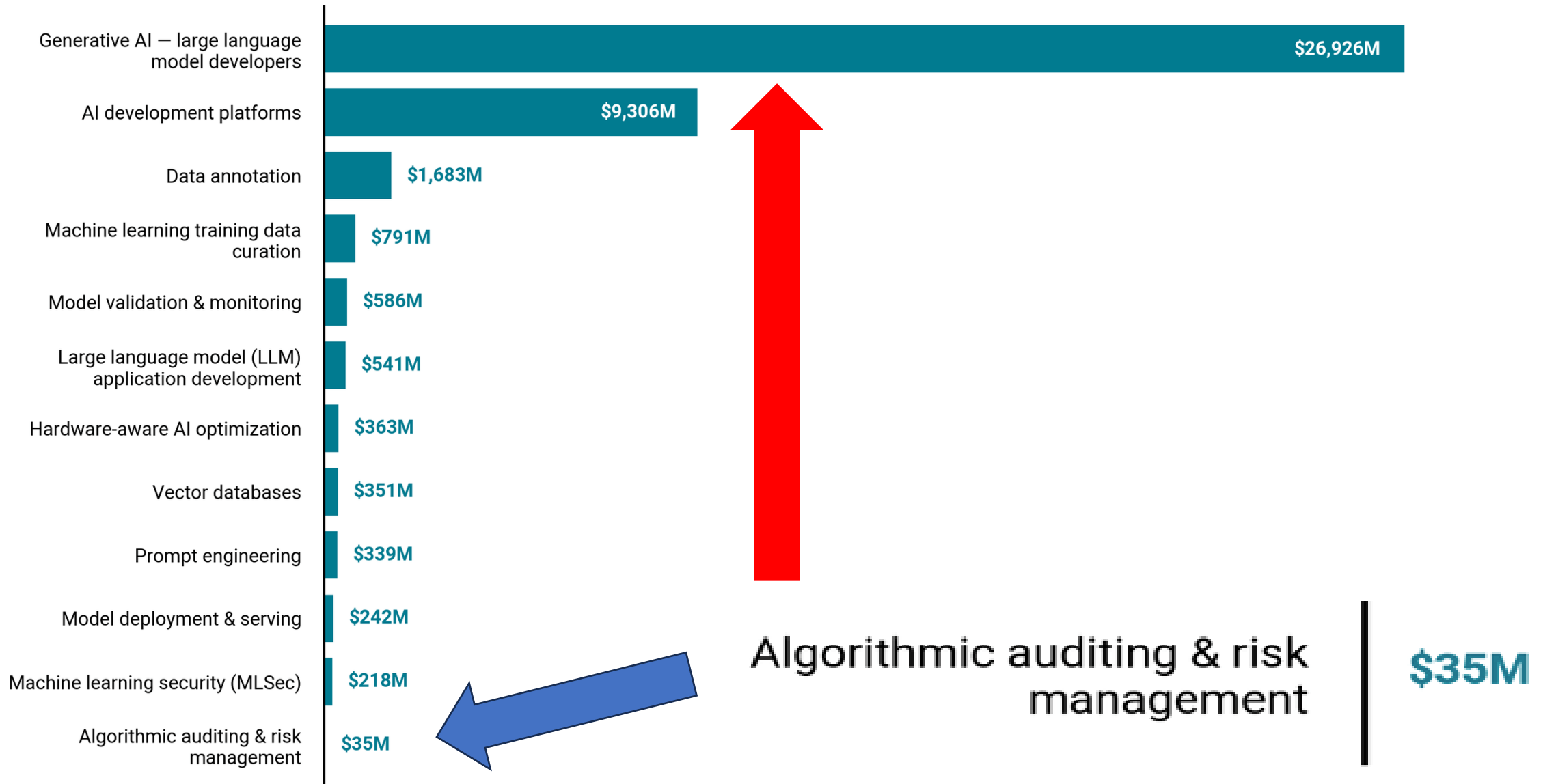 - Get [US NIST AI RMF](#) **(FREE)**
- **How:**
  - Review [ISO 19011](#) – 2018 & do it
  - Know and do risk based auditing for 1st/2nd/3rd party auditing
  - Know and do risk based, decision making
- **Who:**
  - If you (**CQI & IRCA members and quality management professionals**) do above, **YOU** will be '**who**'

Large language model operations (LLMOps) total equity funding

Category	Funding
Generative AI — large language model developers	$26,926M
AI development platforms	$9,306M
Data annotation	$1,683M
Machine learning training data curation	$791M
Model validation & monitoring	$586M
Large language model (LLM) application development	$541M
Hardware-aware AI optimization	$363M
Vector databases	$351M
Prompt engineering	$339M
Model deployment & serving	$242M
Machine learning security (MLSec)	$218M
Algorithmic auditing & risk management	$35M

Algorithmic auditing & risk management
$35M

CB INSIGHTS

Extra resources

- Andrew Ng's ebook- [How to Build Your Career in AI](#)

  - AI for Everyone Coursera course (People can audit it for free) https://www.coursera.org/learn/ai-for-everyone

- Free machine learning and artificial intelligence from google - https://cloud.google.com/learn/training/machinelearning-ai

Top 10 take aways:

Embrace AI: Share insights about how AI shapes the future of decision-making, emphasising its importance in managing the data deluge

Discuss Impact: Engage your network with discussions on AI's influence on public policy, employment, professions, companies, and evolving regulations

Highlight Ethical Concerns: Start conversations around AI decision-making's ethical aspects, especially the potential for bias and its implications in various sectors

4Explore Upcoming Laws: Inform your network about upcoming AI-related legislation, such as New York City's AI Bias Law, and its potential impact on businesses
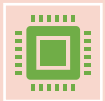
Quality Assurance: Stress the significance of quality assurance and AI, including standards and regulations like ISO 42001 FDIS and NIST AI RMF

Futureproofing: Share tips on how professionals can future proof themselves by understanding AI risk management, auditing, and decision-making

Be the Expert: Position yourself as an authority in the field, offering guidance to your network on mastering AI standards and risk-based audit

Join Discussions: Engage in LinkedIn discussions on AI's impact on the job market and its role in reshaping the workforce.

Networking: Connect with like-minded professionals, AI enthusiasts, and quality management experts to stay updated and share knowledge.

Stay Informed: Keep your LinkedIn network informed about the latest developments in AI and quality management, demonstrating your commitment to staying at the forefront of these evolving fields.

Thank you

Any questions?