Navigating the Future: Practical tips for HR professionals in navigating the use of AI and algorithms

Oliver Patel
Enterprise AI Governance Lead, AstraZeneca

Connect on LinkedIn

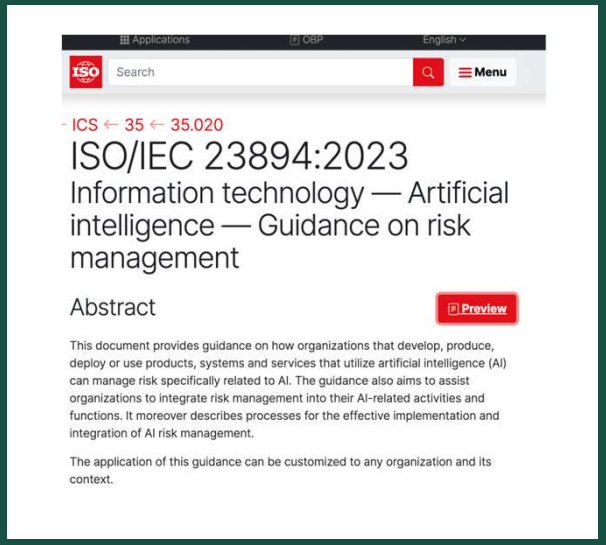
Overview

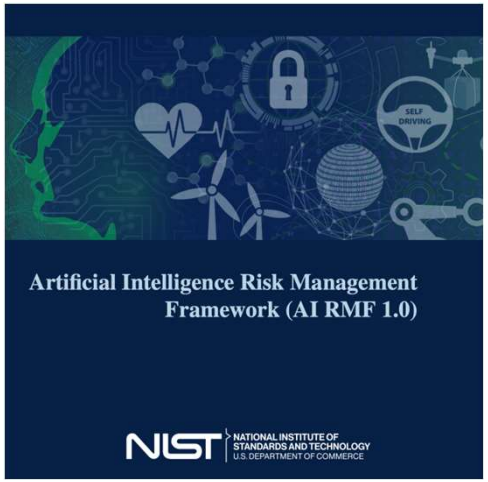
1. AI risk management (NIST; EU AI Act)
2. Enterprise AI governance
3. AI testing, monitoring and quality assurance
4. Developing AI literacy

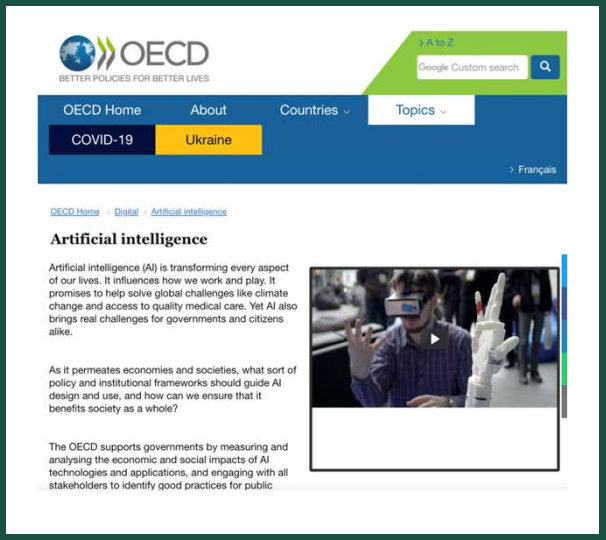
Practical tips

1. Don’t reinvent the wheel, draw upon existing policy resources and toolkits
2. Understand the EU AI Act - are you a ‘provider’ or a ‘user’?
3. Engineers want to build tools, not manage risks
4. You are exposed to your AI vendor’s worst practices
5. There are advanced testing and mitigation techniques for each responsible AI pillar
6. AI quality assurance, auditing and risk management is a growing industry
7. Understand and use ChatGPT, because your candidates and employees are
8. Become as AI literate as possible and ensure you have a seat at the table!

1. AI risk management







Tip 1: Don’t reinvent the wheel, draw upon existing resources


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

Source: NIST

How do AI risks differ?

- Speed, scale, amplification, autonomy and constant change
- Unrepresentative or incomplete training data
- AI system scale and complexity (trillions of parameters / decision points)
- Privacy risk due to enhanced data aggregation
- AI systems require continuous monitoring and more frequent maintenance (due to model, data or concept drift)
- New field: underdeveloped AI software testing standards

GOVERN – cultivate a culture of AI risk management

- Establish and implement policies, processes, procedures and practices across the organisation
  - AI development and deployment
  - Third party risk
- Accountability and governance structures are put in place for mapping, measuring, managing and owning AI risks
- Structures are in place for robust engagement and collaboration between AI teams and business teams

MAP – context-driven approach to risk

- Establish the goals and purpose for the organisation’s use of AI and the corresponding risk appetite
- Map the risks and benefits for each AI system
- Categorise and document the AI systems in development and production across the organisation (AI inventory)
-

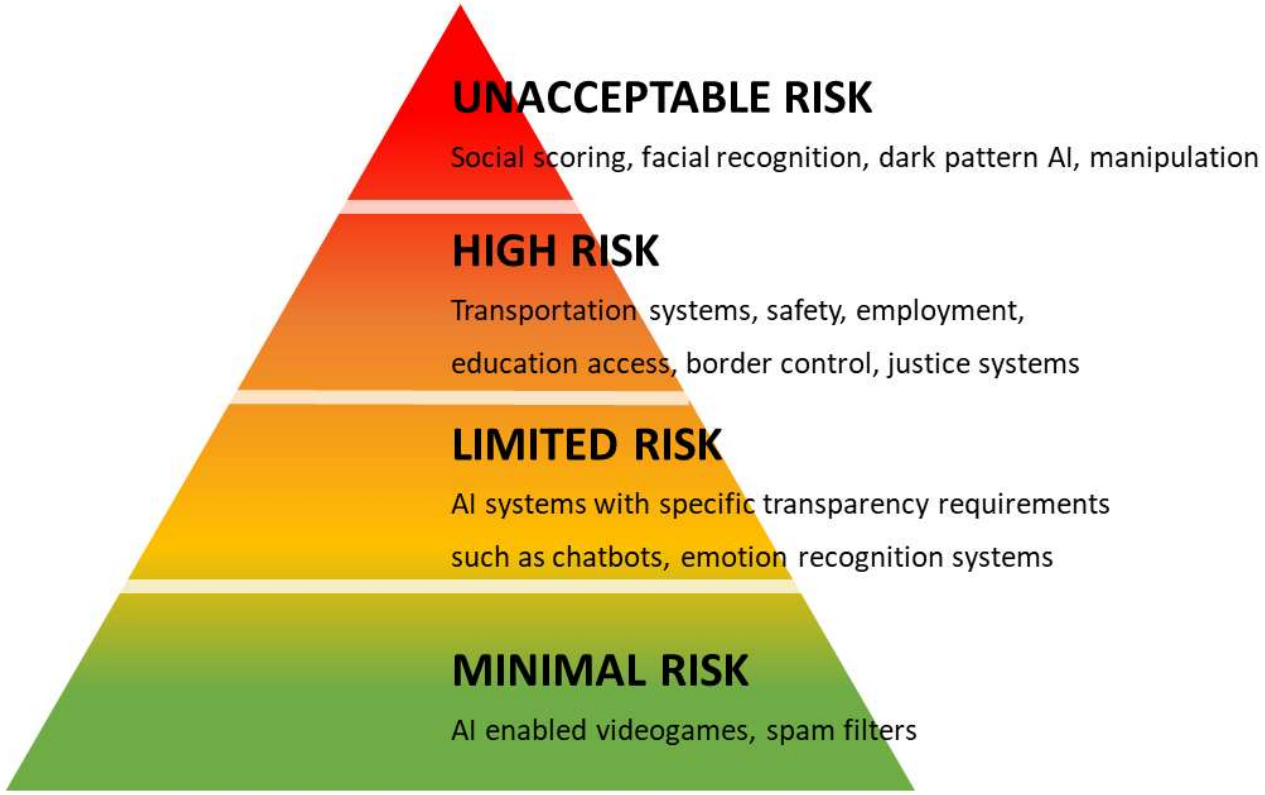
MEASURE – analyse, assess and monitor AI risks

- Identify appropriate metrics and methods
- Evaluate AI systems for responsible AI characteristics
  - Quantitative and qualitative assessment / testing for performance, functionality, robustness, security, privacy, bias, explainability etc.
- Establish mechanisms for tracking AI risks

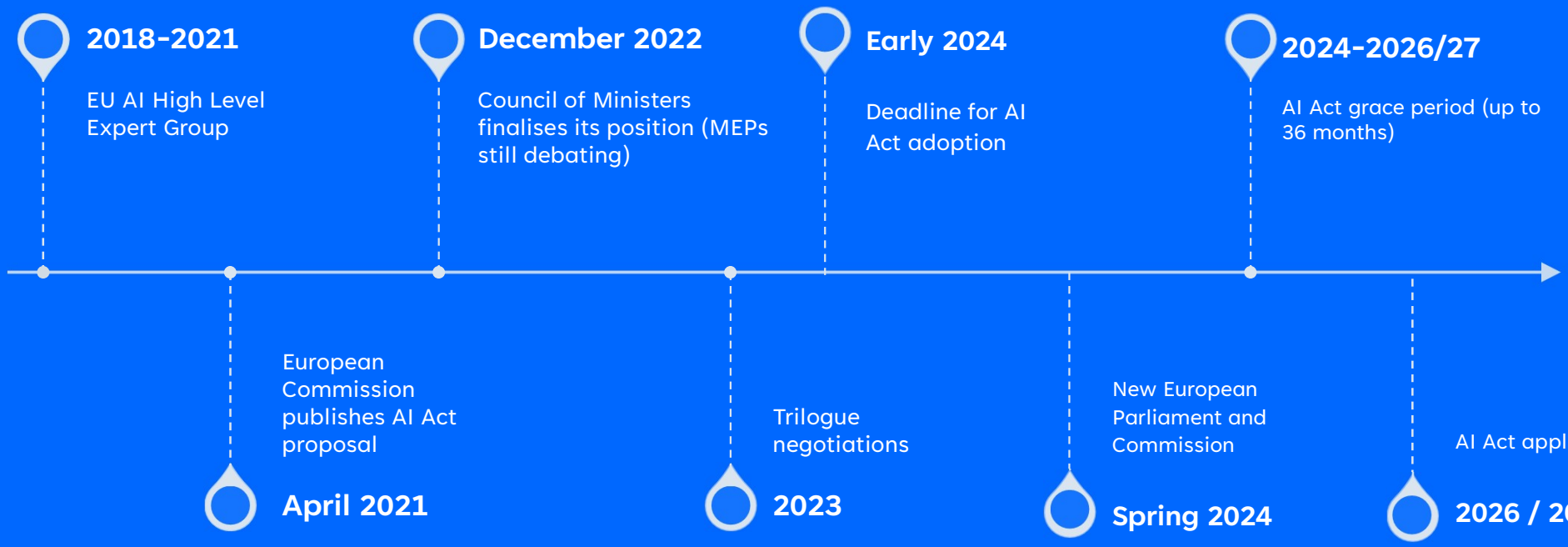
MANAGE – prioritise risks and mitigate their impact

- AI risk assessments, prioritization of risks and heatmaps / tracking
- Continuous monitoring and management of third-party AI risk
- Continuous post-deployment AI monitoring, intervention and mitigation

Tip 2: Understand the AI Act – are you a ‘provider’ or a ‘user’



EU AI Act Timeline



High-risk AI

Employment, workers management and access to self-employment:

a) AI systems intended to be used for **recruitment** or selection of natural persons, notably to place targeted job advertisements, to **analyse and filter job applications**, and to **evaluate candidates**;

b) AI intended to be used to **make decisions on promotion and termination** of work-related contractual relationships, to **allocate tasks** based on individual behavior or personal traits or characteristics and to **monitor and evaluate performance** and behavior of persons in such relationships.

Requirements for PROVIDERS of high-risk AI systems

- Set up a risk management framework
- Undertake a conformity assessment
- Quality assurance and testing
- Appropriate data governance and quality (e.g., representative training data)
- Technical documentation and record keeping
- Transparency and provision of information to users
- Accuracy, robustness and cybersecurity
- Human oversight and explainability
- Post-market / deployment monitoring

Requirements for USERS of high-risk AI systems

- Use AI system in accordance with the provider's 'instructions for use'
- Data governance and quality (if the user controls input data for training, testing and validation)
- Monitoring, record and log keeping
- Cooperation with regulators
- All GDPR provisions apply – you could be the user of an AI system and either the data controller or processor

NYC Bias Audit Law: key requirements

- Addresses bias in automated decision-making tools used in the HR context
- The law is in force from 15 April 2023
- Companies which use AI tools to inform / make decisions about hiring candidates and promoting employees must conduct bias audits
- The tool cannot be used, unless an audit has been conducted
- The audit must be conducted by an independent auditor

NYC Bias Audit Law: transparency provisions

- Employers must publish the results of their bias audits
  - Information about the employment outcomes of different demographic groups
- Must inform users that an automated decision-making tool is being used
- Fines for non-compliance, but greater reputational risk

2. Enterprise AI governance

Tip 3: Engineers want to build tools, not manage risks

- Don’t expect the technical teams to be proactive about implementing responsible AI
- However, responsible AI can only be implemented with technical support
- Responsible AI by design is the best approach

Enterprise AI governance best practice

- Appropriate governance structures and forums need to be established
  - Three lines of defence
- Standards, policies and procedures need to be in place for AI development and procurement
- Roles and responsibilities: there needs to be accountability for and ownership of risks
- You need an updated inventory of your AI systems: SSOT
  - How many AI systems are in development and deployment?

Tip 4: You are exposed to your AI vendor's worst practices

- Most AI systems used in HR will be procured from third party vendors. Due diligence is key!
- The AI vendor landscape is a wild west – many companies won't have prioritised responsible AI.
- Establish and implement robust AI procurement and third party risk policies and procedures.

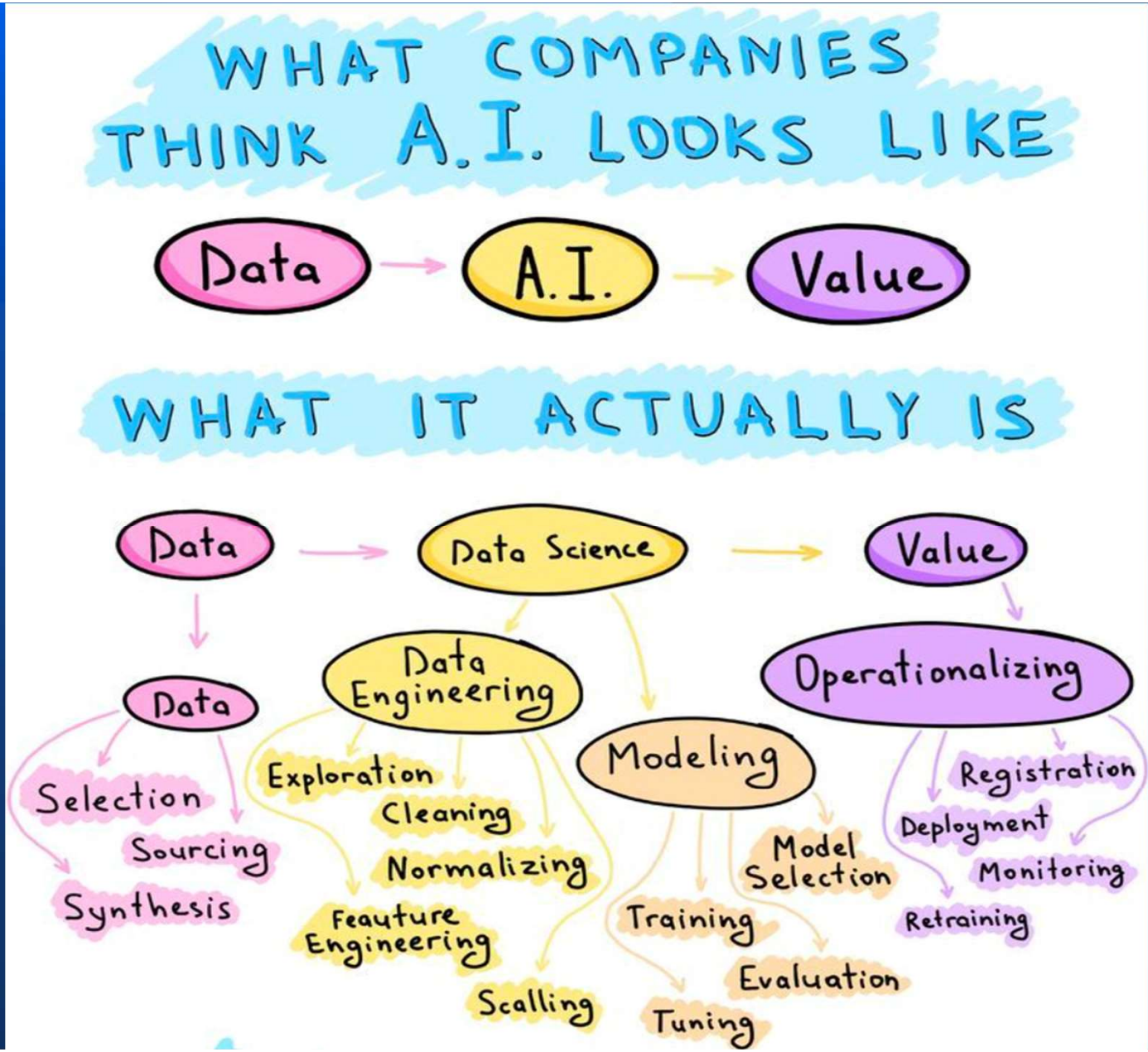
Third party risk management best practice

- What has the vendor done to test for and mitigate issues relating to bias, robustness, transparency, privacy, security etc.?
- Why is it better to buy versus build?
- What is the track record of this vendor with respect to regulatory compliance?
- What technical documentation can the vendor provide about responsible AI?
  - Auditing? Certification?
  - Preparation for the EU AI Act?

3. AI testing, monitoring and quality assurance

AI / ML development Lifecycle

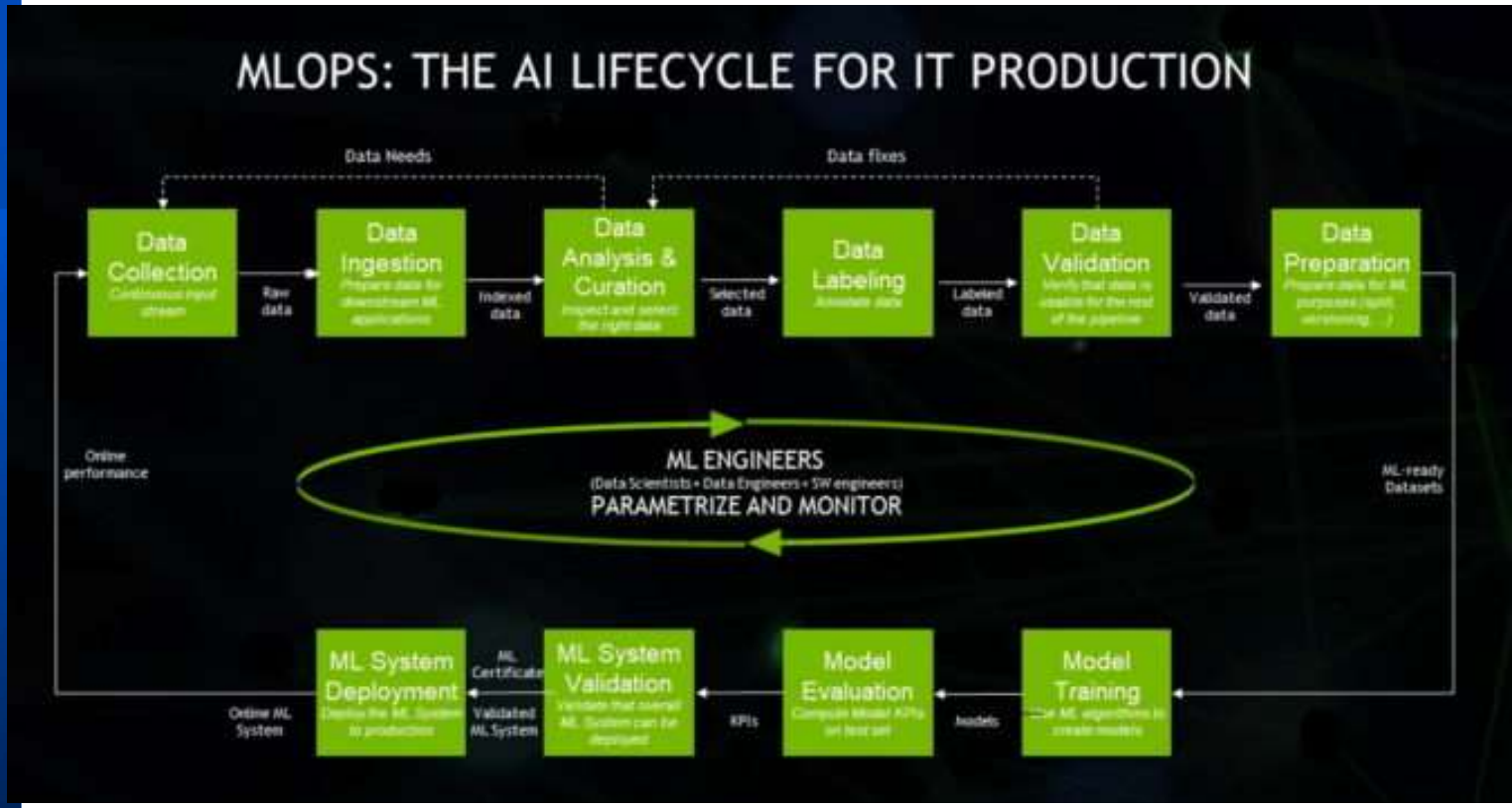
Source: Alex Wang, LinkedIn



MLOps

A set of practices that aims to deploy and maintain machine learning models in production reliably and efficiently.

Source: Nvidia



Tip 5: There are advanced testing and mitigation techniques for each responsible AI pillar

Robustness

Performance over time

Model accuracy

Model reproducibility / drift

Adversarial attacks

Fairness

Training data quality

Biased / discriminatory decision-making

Varied performance across different groups

Transparency

Model interpretability and explainability

Notification and disclosure

Technical documentation and record keeping

AI vendor documentation

Accountability

Human in the loop

Right to object / challenge decision-making

AI liability

Enterprise governance: roles and responsibilities

Privacy

Data minimisation

Automated decision-making (GDPR Article 22)

Sensitive data

Privacy engineering / by design

Security

AI system resilience

Open source AI

Third party risk

Data poisoning

System retirement

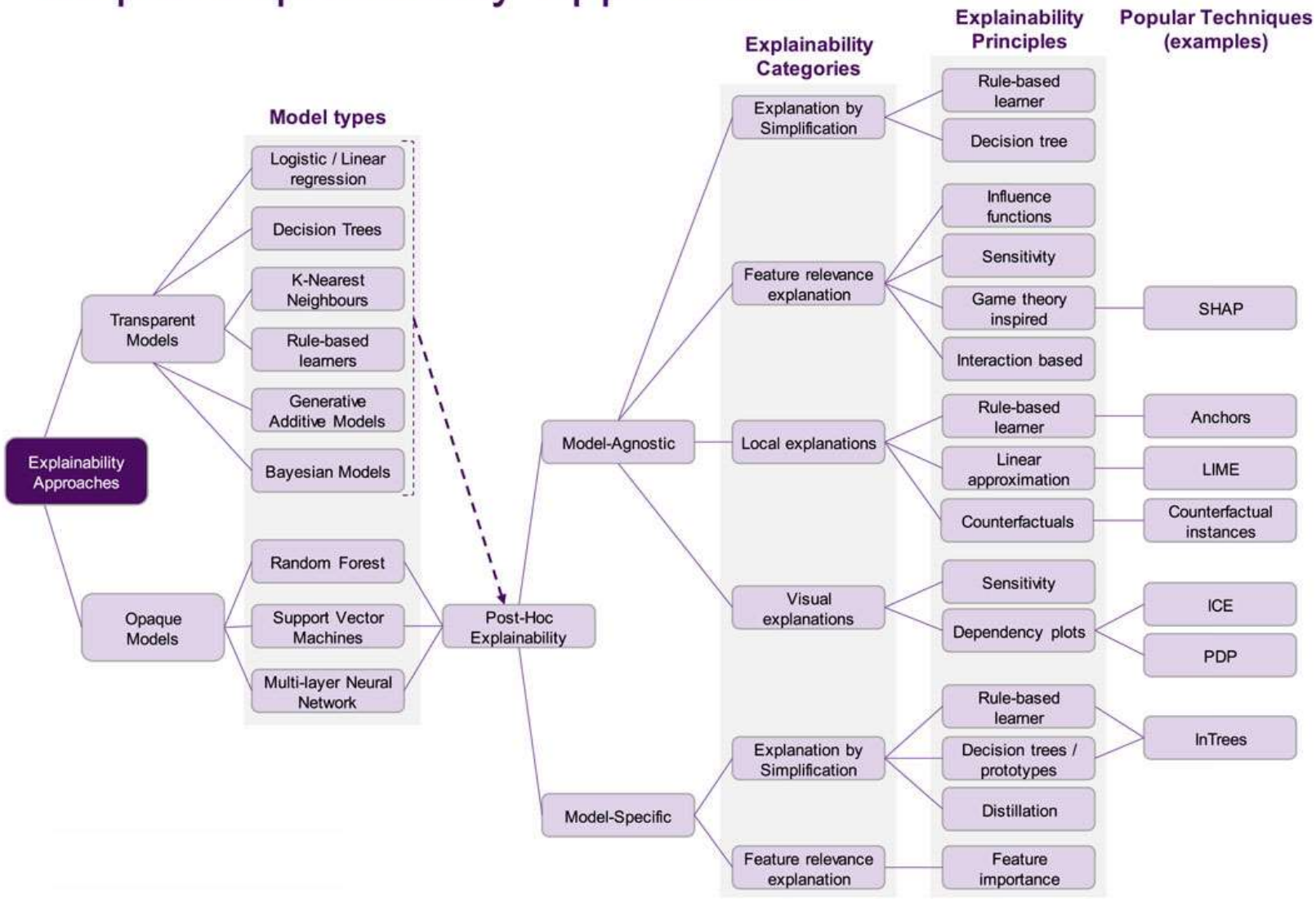
Sustainability

Carbon emissions / footprint

AI re-usability and duplication of work

AI supply chain / modern slavery

Map of Explainability Approaches



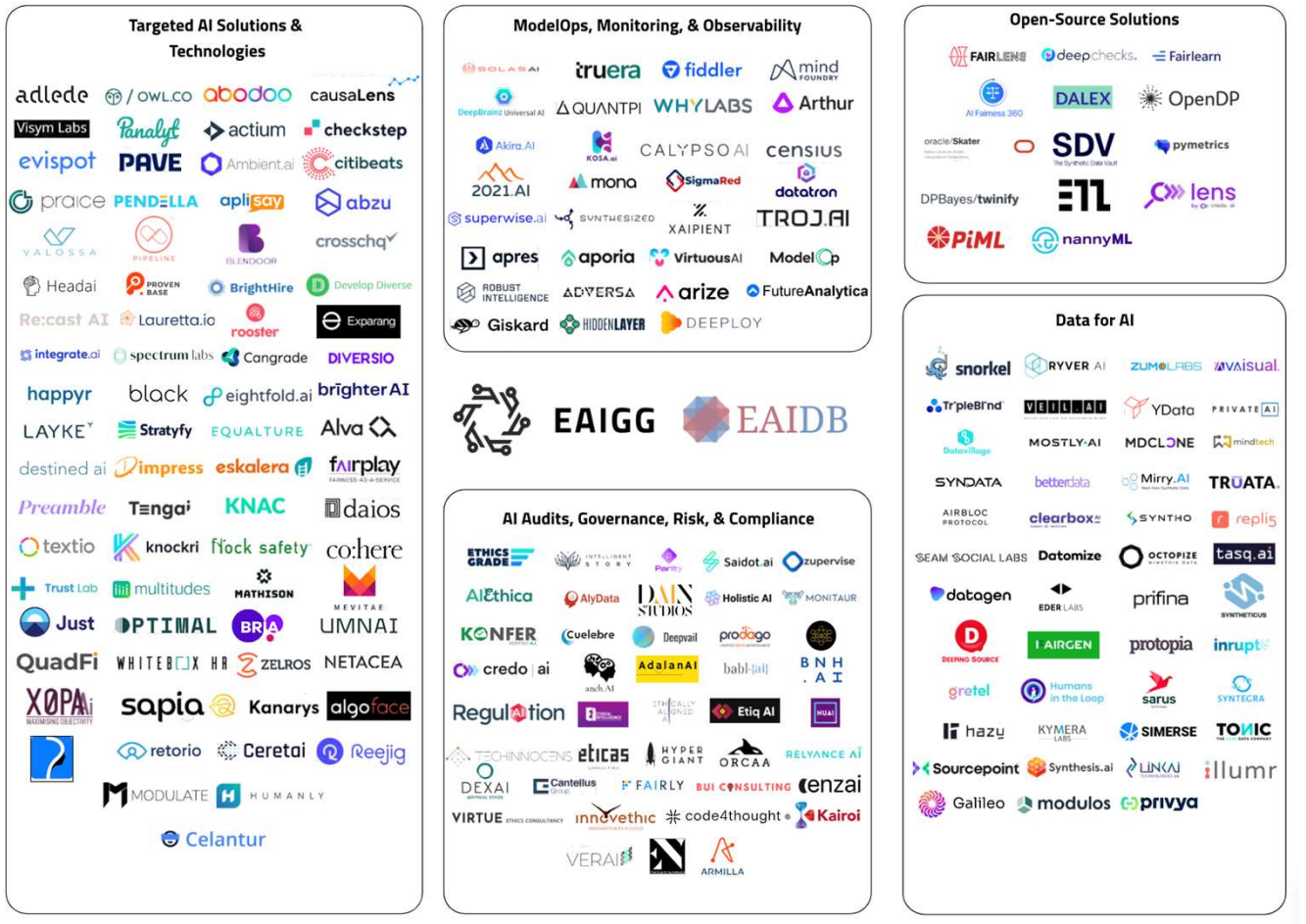
Source: Frontiers

Explainable AI explained

- Explainability is a set of techniques designed to determine which model feature or combination of features led to a specific output (e.g., decision / prediction).
- General trade off between model accuracy / performance and explainability
- SHAP (Shapley Additive exPlanations) - method to compute the contribution of each feature to the prediction in order to identify the impact of each input.
- LIME (Local Interpretable Model-Agnostic Explanations) - attempts to understand the relationship between a particular example's features and the model's prediction by training a more explainable model such as a linear model with examples derived from small changes to the original input.

Source: AI Infrastructure

Ethical AI Startup Landscape



Tip 6: AI quality assurance, auditing and risk management is a growing industry

- Ethical AI Database

Table 2. Certification Percentage, Level, and Logo

Total Score	Level Obtained	Corresponding Mark
0-49.9%	Not Certified	N/A
50-59.9%	Certified	
60-69.9%	Silver	SILVER
70-79.9%	Gold	GOLD
80+%	Platinum	PLATINUM

Dimensions and Subdimensions

Dimensions	Subdimensions
Data and System Operations	> System Scope and Function > Human-in-the-Loop > Model is Fit for Purpose > Data Relevance and Representativeness > Data Quality
Explainability & Interpretability	> Communication About the Outcome > Notification > Recourse > Understanding the AI System's Decisions or Functions
Accountabiltiy	> Organizational Governance > Team Governance
Consumer Protection	> Transparency to the User and Data Subject > Harms to Individuals > Protections
Bias & Fairness	> Bias Impacts

Source: Responsible AI Institute

4. Developing AI literacy

Tip 7: Understand and use ChatGPT, because your candidates and employees are

Let's start by asking the chatbot itself. UNLEASH asked ChatGPT, 'How could HR use ChatGPT?'.

The AI replied:

HR departments can use ChatGPT in several ways:

1. Recruiting: ChatGPT can be used to automate the initial screening of job applicants by answering common questions about qualifications and experience.
2. Employee engagement: ChatGPT can be used to answer common questions from employees, such as benefits and company policies, freeing up HR staff to focus on more complex issues.
3. Performance evaluations: ChatGPT can be used to generate performance evaluations based on data input, such as an employee's job responsibilities and accomplishments.
4. Employee training: ChatGPT can be used to create interactive training modules and to answer common questions from employees.
5. Employee retention: ChatGPT can be used to analyze employee sentiment and provide insights into potential issues that could lead to high turnover.
6. Onboarding: ChatGPT can be used to onboard new employees by answering common questions and providing information about company policies and procedures.
7. Compliance : ChatGPT can be used to ensure compliance with regulatory requirements by providing accurate, up-to-date information about laws and regulations that apply to the organization.

Tip 8: Become as AI literate as possible and ensure you have a seat at the table!

- Don’t wait to be invited – get involved in your company’s AI governance work
- You don’t need to code, but it’s crucial to speak the language of technical teams and understand their perspective and priorities
- Resources for AI literacy:
  - IBM AI Essentials (video series)
  - Google ML courses + glossaries
  - Empowering AI Leadership: AI C-Suite Toolkit (WEF)
  - AI Standards Hub (Alan Turing Institute)

“AI is a tool. It is neither good nor evil. The choice about how it gets used is ours.”

Oren Etzioni

Thanks for listening

Q&A

Oliver Patel
Enterprise AI Governance Lead, AstraZeneca

Connect on LinkedIn

Images from Flaticon