Trustworthy AI in HMRC

Built on understanding, not bureaucracy

Anna Kwiatkowska

Chief Data Scientist | HMRC

IOTA Forum Oct 2026



HM Revenue & Customs

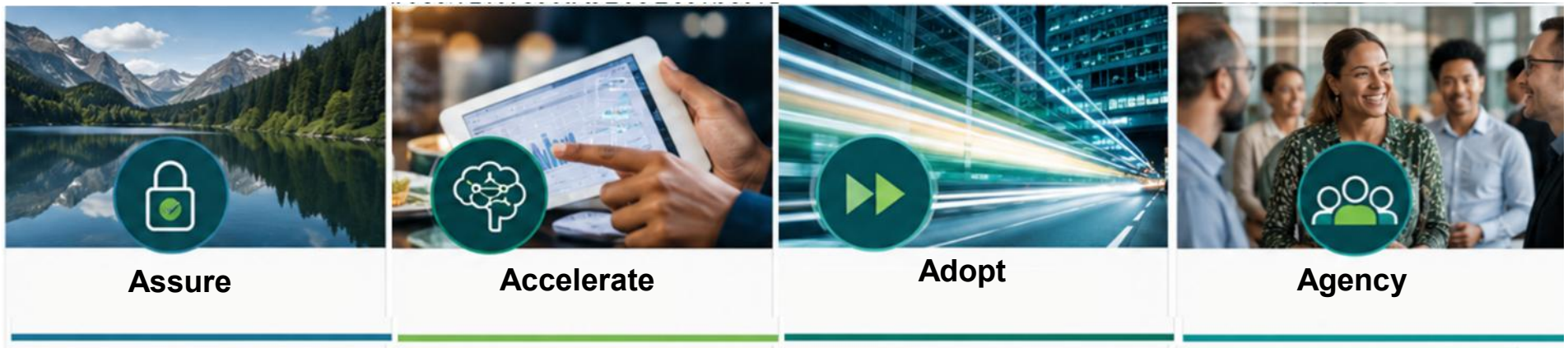
Introduction: Trust and Transparency

HMRC Turning AI ambition into responsible delivery

HMRC is not new to AI. We have decades of experience using advanced analytics, machine learning and algorithmic techniques to tackle **fraud, reduce error, target compliance activity, support debt management** and improve customer outcomes.

Our ambition is to be the world's most AI-enabled tax, customs and valuation authority, using AI **responsibly**, at pace and at scale to improve outcomes for customers, colleagues and the public.

Focus and federation



OFFICIAL

AI is already supporting colleagues and customers



Customers

6.3m

Ask HMRC interactions

Helping customers get their tax right during 2025-26

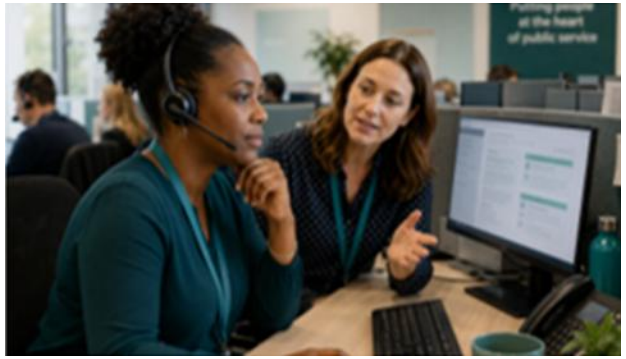


Colleagues

60k+

Copilot licences deployed

Extending access, with people checking outputs and retaining accountability



Skills

80k

AI learning activities

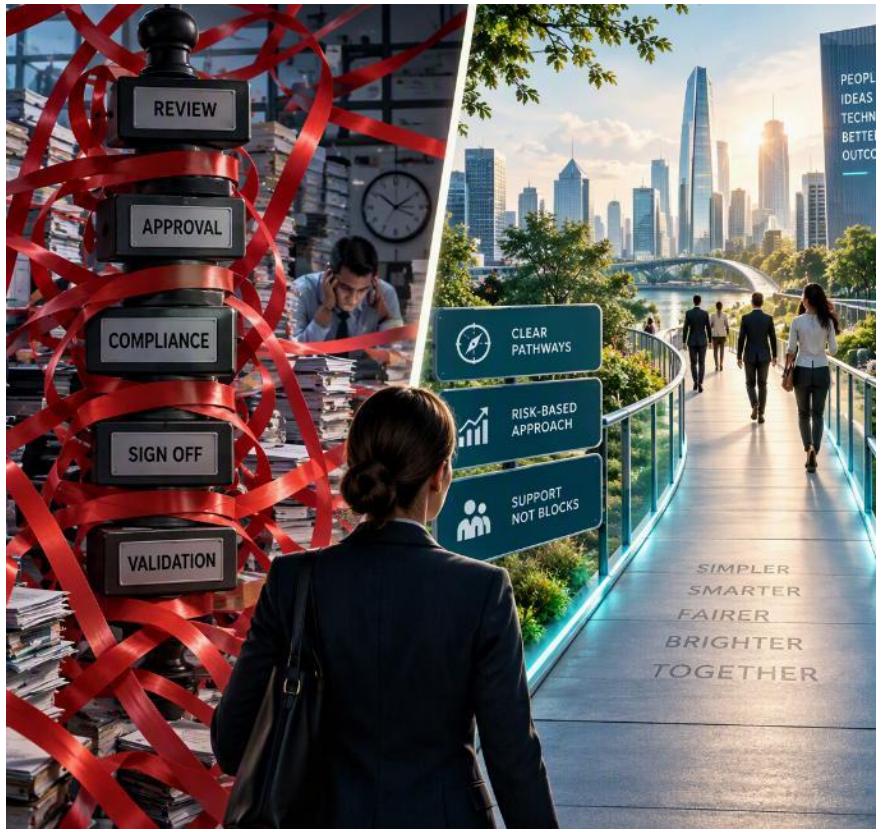
Completed in 2025-26, alongside practical role-based support

But how do you do trustworthy at this scale and without stifling innovation ?

Adding to governance or adding to understanding and assurance

Effective and efficient governance requires:

- System (Data, process, users, costs and dependencies) understanding and visibility of AI
- Clear accountability (named owners and record decisions)
- Scaling with continuous evaluation



OFFICIAL

HMRC’s approach: ambition with responsibility

AI is part of our transformation plan, not a standalone programme

OUTCOMES

Start with customer, colleague and taxpayer outcomes

FEDERATED DELIVERY

Business areas own delivery; shared standards enable reuse and scale

LIFECYCLE ASSURANCE

Assess risk, test, evaluate and monitor beyond initial approval

PEOPLE & SKILLS

Build confidence, judgement and specialist capability

The organising principle: enable informed decisions at the point of delivery, while retaining a coherent enterprise view.

Before applying AI: understand the system

Business needs	What problem are we trying to solve ?
Data	Do we have access to the right data, and can it be trusted?
Process	Where does value flow, and where are the constraints, is AI the right solution
Cost	What is the total cost of ownership and change? All costs build, run, integration, tokens etc
Risk*	What could go wrong, and who is impacted?
Accountability	Who is responsible for outcomes and decisions?



* AI Risk Management Toolkit - GOV.UK

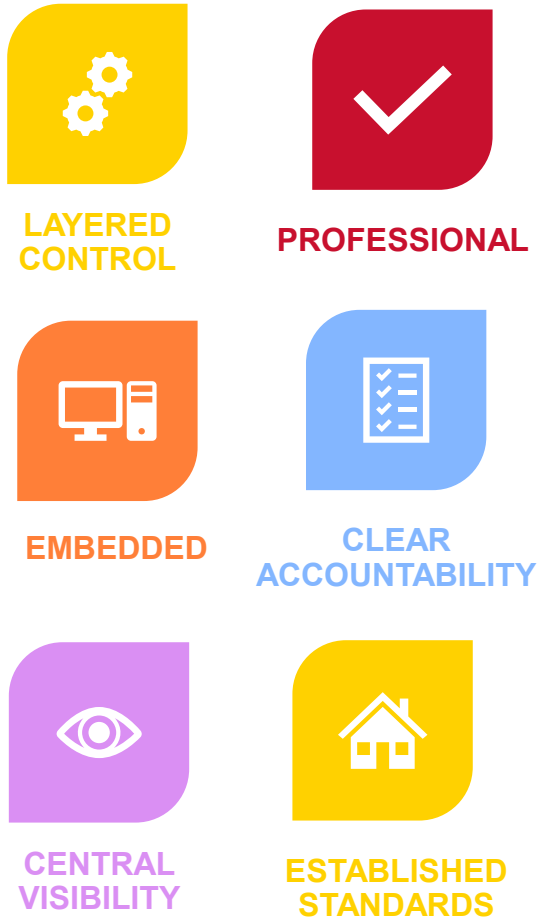
OFFICIAL

Aligned with existing governance and focused on guardrails, observability and accountability

Governance with focus:

- What is the intended outcome?
- What evidence is sufficient?
- What could go wrong, and for whom?
- What must remain under human control?
- What changes require re-assurance?

Embed in existing governance, ensuring AI is designed, delivered and operated by skilled professionals under clear accountability, ethics and assurance across the full lifecycle



OFFICIAL

Human in the loop and control



Human accountability requires:

- Understanding when AI could / should and should not be used
- Judging whether outputs are legal*, safe and appropriate
- Applying domain expertise and professional judgement
- Having the authority to challenge, override or stop AI decisions
- Ensuring outcomes can be questioned, challenged, corrected and / or redressed

The goal is not human approval of AI outputs

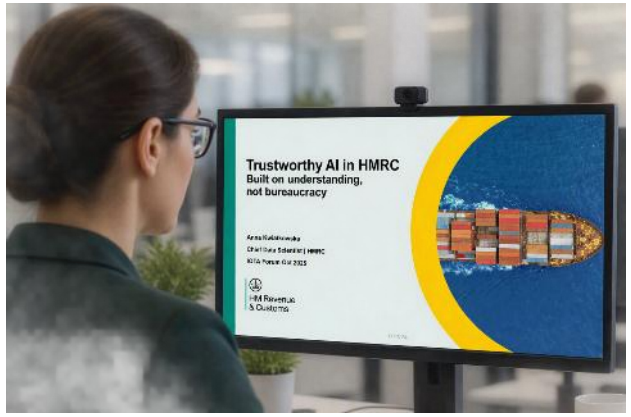
It is informed human judgement over AI-enabled decisions.

But not all decisions require the same level of scrutiny

*legal landscape is shifting

OFFICIAL

Scale assurance with consequence and uncertainty



Copilot drafting

Assistive, reversible, internal

Clear ownership | approved tools | basic evaluation



AI Assisted complaints triage

Material workflow or service influence

Documented testing | monitoring | fallback | user feedback



VAT Risk Modelling

Customer-facing or decision-supporting

Independent challenge | explainability | stronger evidence | ongoing oversight

Underpinned by understanding

OFFICIAL

Understanding enables both trust and innovation

UNDERSTAND

the data, process, costs and context

ACCOUNT

for outcomes, decisions and failures

ENABLE

innovation through proportionate guardrails and federated model

The question is not “How much governance do we have?”
The question should be “Do we understand enough to act responsibly?”

OFFICIAL



Discussion