AU0009 / LEVEL 5

AI leadership – AI strategy and opportunity

# Lead responsible AI adoption

People, ethics, risk appetite, human oversight and information security

Skills England Apprenticeship Unit

**Version 1.0 • 30-hour programme**

CURRICULUM READY

**03**

**BLOCK**

**180 minutes**

+ 10-minute break

# What you will be able to do

**01** Evaluate workforce, ethical, reputational and information-security implications.

**02** Set appropriate levels of human oversight.

**03** Describe a practical organisational risk appetite for AI use.

**LO1, LO3, LO4 • K1, K2, K3, K7 • S3, S4**

INDIVIDUAL → PAIR • 4 MIN

# “A human approves it, so the risk is solved.”

**01** Decide true or false privately.

**02** Tell a partner what the reviewer would need.

**03** Identify one way approval could become a rubber stamp.

**1 min: decide. 2 min: compare. 1 min: share one missing condition.**

WHOLE GROUP • REVEAL • 5 MIN

# Human approval needs the means to challenge

**01** Time to examine the source and the output

**02** Capability to spot an error or unfair inference

**03** Authority to override, stop and escalate

**04** Clear accountability for what happens next

**Keep your first answer. We will revisit it after the case.**

TUTOR INPUT → INDIVIDUAL • 5 MIN

# Responsible adoption is a leadership decision

**01** Purpose: whose problem should improve?

**02** People: who benefits, who bears risk and who can challenge?

**03** Information: what may enter, leave or be retained?

**04** Value: what costs and consequences sit outside the headline?

**3 min input. 2 min privately: identify one overlooked stakeholder.**

SMALL GROUP • RISK RADAR • 7 MIN

# What could change for people?

**01** Fictional case: an assistant drafts internal service advice.

**02** Staff must check every answer while keeping existing targets.

**03** The team expects faster service but has not measured checking work.

**Groups: map effects on workload, wellbeing, trust and fairness. Name the most material risk.**

GROUP • CONTROL DESIGN • 8 MIN

# A risk needs an owner and a workable control

**01** Consider confidentiality, access and inappropriate reuse.

**02** Consider wrong advice, reputation and customer consequences.

**03** Consider ongoing resource use, sustainability and dependency.

**04** Choose two material risks. State a control and who operates it.

**Output: two risk–control–owner links, including how you will know the control works.**

WHOLE GROUP • CHALLENGE • 5 MIN

# A policy sentence is not proof of a control

**01** Each group offers one control.

**02** Another group names a way it could fail.

**03** The proposing group improves its control or marks an unknown.

**Give a concrete check, an accountable owner and a remaining uncertainty.**

# Human in, on or out of the loop

**01** IN: a person checks before the action takes effect.

**02** ON: the system acts within limits and a person monitors and can intervene.

**03** OUT: no routine human intervention in that step.

**04** Match oversight to consequences, speed and ability to recover.

**Name the exact action being supervised. Oversight labels alone do not establish safety.**

GROUP • SCENARIO SORT • 8 MIN

# Where should a person intervene?

**01** A: draft a routine internal message from approved text.

**02** B: route a time-critical service alert.

**03** C: recommend a decision affecting an employee.

**04** D: copy approved structured fields using fixed rules.

**Choose an oversight position for each fictional case. State an assumption and a stop condition.**

PAIR • OVERSIGHT TEST • 7 MIN

# Can the reviewer actually intervene?

**01** Your reviewer sees 100 suggestions before a deadline.

**02** The source evidence is hidden behind another system.

**03** Managers reward speed and rarely record overrides.

**Design a better workflow: what changes before, during and after a questionable output?**

WHOLE GROUP • EXPLAIN • 5 MIN

# An accountable workflow

**01** Who owns the decision and the source information?

**02** Who can challenge, override and pause?

**03** Who handles an affected person’s concern?

**04** What record lets the organisation learn from mistakes?

**Explain one revised workflow. Others test whether intervention is practical.**

TUTOR INPUT → PAIR • 5 MIN

# Risk appetite becomes a boundary

**01** GREEN: proceed with normal controls.

**02** AMBER: extra approval, testing or oversight is required.

**03** RED: do not proceed under current conditions.

**04** A category depends on context, evidence and consequences.

**Pairs: explain why the same tool could fall into different categories.**

GROUP • BOUNDARY SETTING • 8 MIN

# What would you allow, condition or stop?

**01** Fictional task: draft public-event wording from approved public text.

**02** Variation: add unpublished customer details.

**03** Variation: publish automatically without a review route.

**Agree green, amber or red for each. Record the evidence or control that would change your choice.**

PAIR → WHOLE GROUP • 7 MIN

# The boundary that matters most

**01** Choose the most consequential variation.

**02** One partner argues for proceeding with controls.

**03** The other tests whether those controls are credible.

**2 min each to argue. Final 3 min: agree a boundary and share an unresolved question.**

BREAK • ADDITIONAL TO THE 180 MINUTES

# 10-MINUTE BREAK

When we return

## NORTHSTAR REVEAL 03

Use the evidence. Make a decision.
Be ready to explain your conditions.

# Proposal A: employee information in an unapproved tool

**01** Absence comments, engagement comments, exit notes and performance information

**02** Aim: identify workforce trends and target support

**03** Approval, retention, reuse and processing arrangements are unclear

**04** No upload has happened. No approved data-flow assessment or purpose boundary.

**Read the proposal. Separate the useful workforce question from the proposed processing method.**

# Approve, modify, pause or reject proposal A?

**01** Comments may identify people or reveal sensitive circumstances.

**02** Themes could become unsupported individual or small-group inferences.

**03** Staff have not been told how information may be repurposed.

**04** A person reading the report does not resolve these concerns.

**Record a decision, purpose/data boundary, approval owner and excluded downstream decisions.**

# Proposal B: a public-information chatbot

**01** Possible benefit: common answers outside staffed hours.

**02** Public source pages conflict and ownership is weak.

**03** Human access, accessibility and complaint handling need a workable route.

**04** Account-specific answers are a later request, not an approved extension.

**Read separately from proposal A. What scope could be tested responsibly?**

# A separate decision for the chatbot

**01** Opening-hours question: who owns and checks the source?

**02** Disputed answer: how does the customer reach a person?

**03** Urgent request or complaint: what stops automated advice?

**04** Wrong answer discovered later: who disables and investigates?

**Agree approve / modify / pause / reject, with scope, test evidence and hand-off conditions.**

NORTHSTAR • WHOLE GROUP • 4 MIN

# What changed between the two proposals?

**01** Compare the consequences, information and oversight required.

**02** Name a control that enables a bounded investigation.

**03** Name a boundary that remains firm under current conditions.

**Explain the reasoning. Seek specialist advice where required; these are leadership decisions, not legal conclusions.**

YOUR TURN • INDIVIDUAL • 10 MIN

# Part C: your material risks

**01** Strategy Builder C: Responsible adoption.
Find “Your material risks” under YOUR ORGANISATION.

**02** Key risk / issue and who or what could be affected

**03** Likelihood / impact and a proportionate control

**04** Human oversight / owner

**Complete the risk rows. Use non-sensitive information and record material issues only.**

PAIR COACHING • 5 MIN

# Does the control address the risk?

**01** Explain one risk and its consequence.

**02** Partner asks who operates the control and what evidence it leaves.

**03** Identify any control that depends on unavailable time or authority.

**2 min each. Final 1 min: capture one improvement in Part C.**

INDIVIDUAL • APPLY • 8 MIN

# Your organisation’s risk appetite

**01** GREEN – normal controls

**02** AMBER – extra approval, testing or oversight

**03** RED – do not proceed under current conditions

**Complete Our boundary / example for each position. Use your own scope and context.**

TRIOS • BOUNDARY CHALLENGE • 7 MIN

# Where would you pause?

**01** Each person explains one amber or red boundary.

**02** Peers change one assumption and test whether the position still holds.

**03** Do not remove a material safeguard just to make the idea proceed.

**2 min each. Final 1 min: record the boundary or uncertainty you will refine.**

INDIVIDUAL • APPLY • 10 MIN

# Responsible AI principles that guide action

**01** Complete Principle and What it means in practice.

**02** Use one principle to review a procedure or the tutor’s fictional policy.

**03** Record the gap, one amendment and approval role in an existing field.

**Refine principles and risk rows. Keep workforce, information and accountability explicit.**

PAIR • WORKFLOW WALK-THROUGH • 8 MIN

# Can your oversight work on a busy day?

**01** Walk through one action from input to consequence.

**02** Test source access, reviewer time, capability and authority.

**03** Check who can stop, escalate and handle a challenge.

**4 min each. Update Human oversight / owner and the linked control.**

INDIVIDUAL • REFINE • 7 MIN

# Your strongest pause or reject boundary

**01** Review the risk record, appetite and principles together.

**02** Complete the issue most likely to make you pause or reject.

**03** Mark what requires specialist input or further evidence.

**Output: a coherent Part C, with one explicit pause/reject boundary.**

DECISION REVISIT • INDIVIDUAL → PAIR • 6 MIN

# Is human approval enough now?

**01** Return to your opening true/false judgement.

**02** Write what you would require before relying on approval.

**03** Compare how your reasoning changed.

**2 min privately, 2 min in pairs, 2 min whole group: share the most important change.**

GROUP • TEACH-BACK • 8 MIN

# A boundary colleagues can use

**01** Choose one fictional use case from today.

**02** Explain allowed scope, responsible owner and meaningful oversight.

**03** Invite a challenge about workforce or information consequences.

**Prepare for 3 min, explain for 2 min, respond and refine for 3 min.**

REFLECTION → BRIDGE • 6 MIN

# Next: what deserves priority?

**01** Choose one boundary to carry into prioritisation.

**02** Name one opportunity that may need more evidence first.

**03** Block 04: Prioritise what matters.

**Record your next evidence question. Bring Parts B and C to the prioritisation work.**