Re.ai Connecting Low-Skilled Adults to Remote Opportunities

LESSON 4.4 · 1.5 HOURS

AI ethics, bias, and your legal responsibilities under EU law

IN THIS LESSON

- Why ethics protects you and others
- How bias creeps into AI
- The EU AI Act risk tiers and your duties
- Running an ethics audit

L4.4 Learning Objectives

By the end of this lesson you will be able to:

- Explain where AI bias comes from and name four types of bias
- Place an AI tool in the correct EU AI Act risk tier
- Apply transparency, GDPR and human-oversight duties to your own AI work
- Run a five-step ethics audit using a 20-point checklist

THIS LESSON

L4.4

Duration: 1.5 hours

Practice: EX 4.4

Quiz: 5 questions, 70% to pass

Key Terms

The words you will meet in L4.4, in plain language.

Bias

When an AI treats some groups unfairly, usually because of skewed training data.

EU AI Act

The EU law (Regulation (EU) 2024/1689) that regulates AI by level of risk.

Risk tier

The AI Act's classification: unacceptable, high, limited or minimal risk.

Transparency

Being open that AI is involved, and labelling AI-generated content.

GDPR

The EU law protecting personal data: lawful use, data minimisation and people's rights.

Human oversight

Keeping a person responsible who can review and override AI decisions.

Why ethics matters to you

TRUST IS YOUR ASSET

Using AI responsibly protects your customers, your reputation and you: legally and commercially, not just morally.

Fair, transparent, privacy-respecting AI keeps customers. Source: OECD AI Principles (2024).

What is bias in AI?

Bias means an AI treats some groups unfairly. It usually comes from the data, not from bad intentions.

AI learns from the past, so it can repeat the past's unfairness.

It starts in the data

AI learns from human-made examples. If they reflect past unfairness, the AI can learn and repeat it.

A simple example

A hiring AI trained mostly on one group's CVs may quietly score similar candidates higher, disadvantaging others.

Why it's dangerous

AI is fast and confident, so biased decisions can be made at scale before anyone notices.

Types of bias to watch for

Bias wears several disguises. Recognising them is the first step to catching them.

Data bias

The training data over- or under-represents certain groups, skewing results against them.

Sampling bias

The data was collected in a way that misses part of the real population.

Historical bias

The data reflects past inequalities, which the AI then treats as 'normal'.

Automation bias

People trust the machine too much and stop questioning it. A human failing, not the AI's.

Ask 'who might this treat unfairly?' for every AI output you rely on.

The EU AI Act: four risk tiers

The EU AI Act sorts AI by risk: the higher the risk, the stricter the rules. Place your project before you build.

Risk tier	What it means	Examples
Unacceptable	Banned outright	Government social scoring of citizens; manipulative AI that exploits vulnerable people
High	Strict duties: risk checks, human oversight, records	AI used in hiring, credit scoring, education grading or essential public services
Limited (transparency)	You must tell users it's AI	Customer-service chatbots; AI-generated content
Minimal	No special rules	Spam filters, AI in games, smart suggestions: most everyday AI

Source: EU AI Act, Regulation (EU) 2024/1689.

Where do your projects sit?

Most tools you build in this course, such as a customer-service chatbot, are limited risk. Your main legal duty is transparency, plus the GDPR whenever personal data is involved.

- Clearly say an AI is responding
- Label AI-generated content where the law requires it
- Check outputs for bias and protect personal data under the GDPR
- If your AI helps decide on jobs, credit or grades, it is high risk

Transparency: AI Act Article 50

The duty you'll meet most often. In plain terms: be honest that AI is involved. It applies from 2 August 2026.

Say it's AI

Chatbots and AI assistants must make clear that users are talking to a machine, unless it is obvious.

Label AI-made content

Deepfakes, and AI text published to inform the public, must be disclosed as AI-made. AI tools must mark their output.

Why it matters

Transparency lets people judge AI output for themselves. It is the foundation of trustworthy AI.

Source: EU AI Act, Art. 50 (transparency); applies from 2 August 2026.

Your responsibilities in plain terms

Your four duties

A

Be transparent

Tell people clearly when they deal with AI or AI-generated content.

B

Check for bias

Test that your AI treats different groups fairly, not just on average.

C

Protect data

Collect only what you need and store personal data securely (GDPR).

D

Keep a human in the loop

A person reviews and can override important AI decisions before they take effect.

GDPR basics for AI builders

If your AI touches personal data, the GDPR applies. Collect the least you can: what you don't hold can't leak.

Have a lawful reason

You need a valid basis to use someone's data, usually clear consent or a genuine business need.

Collect only what you need

'Data minimisation' means gathering the minimum required, not everything you could.

Respect people's rights

People can ask what data you hold, correct it, or have it deleted. Be ready to honour that.

Decisions by machine

If AI alone decides something important, like a loan, people can ask for a human to review it (GDPR Art. 22).

Reference: GDPR, Regulation (EU) 2016/679.

Run an AI ethics audit

An ethics audit is a simple, repeatable check. Five steps make it manageable.

1

Set the scope

Pick one AI tool, such as your chatbot, and define what it does and who it affects.

2

Check the data

Ask what data it uses and whether any group is missing or misrepresented.

3

Test for bias

Try inputs from different groups and look for unfair differences in the output.

4

Check transparency

Confirm it discloses that it's AI and protects personal data properly.

5

Log gaps and fix

Write down what's wrong, plan changes, and re-check after fixing.

In EX 4.4 you apply a 20-point checklist to your EX 4.1 chatbot.

Responsible vs irresponsible AI use

The line between the two is mostly honesty, fairness and care.

Responsible use

- Tells people clearly when AI is involved
- Tests outputs for bias and fixes them
- Collects and protects data carefully
- Keeps a human accountable for decisions

Irresponsible use

- Hides that AI made the decision
- Ignores who might be treated unfairly
- Hoards personal data carelessly
- Lets the machine decide with no oversight

Aim for the left column in everything you build.

Check Your Understanding

Try answering these in your own words before moving on.

In your own words, where does bias in AI usually come from?

What are the four risk tiers of the EU AI Act, from highest to lowest?

What must you disclose under Article 50, and from when?

Name the five steps of a simple AI ethics audit.

Learning Resources

Want to go further? These are free and trusted.

European Commission: AI Act regulatory framework

https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai

EU AI Act explorer (full text by article)

https://artificialintelligenceact.eu/

OECD AI Principles

https://oecd.ai/en/ai-principles

L4.4 Key Takeaways

Before the next lesson, hold on to these five points.

- Ethics protects people and you: trust, fairness and the law all depend on it
- Bias usually comes from the training data, so test outputs across groups
- The EU AI Act ranks AI by risk; most tools you build are limited risk
- From 2 Aug 2026 you must disclose AI and label AI-made content (Art. 50)
- Four duties: be transparent, check bias, protect data, keep a human in the loop

Now Put It Into Practice

Head to the Re-AI learning platform to finish this lesson.

INTERACTIVE ACTIVITY

Which risk tier, which duty?

On the platform · about 10 minutes

HANDS-ON EXERCISE EX 4.4

Run an AI ethics audit

1 hour · Tools: 20-point AI Ethics Checklist (in this worksheet), Your EX 4.1 chatbot, Google Docs or Word

LESSON QUIZ

5 multiple-choice questions

70% to pass · unlimited retakes