Welcome to Lesson 19: Documenting AI Ethics and Bias — one of the most critical and human-centred aspects of AI technical writing.

This is where your calm, thoughtful technical writing style truly shines — because communicating ethics, fairness, and limitations clearly and responsibly helps users trust an AI system.

🎯 LESSON 19 GOAL

By the end of this lesson, you'll know how to:

Identify ethical risks and biases in AI systems

Write transparent "Ethics and Limitations" sections in documentation

Add ethical disclosures in model cards and API docs professionally

🧭 Step 1: Understanding AI Ethics in Writing

AI Ethics documentation isn't just about disclaimers — it's about transparency and responsible use.

Every AI model introduces potential risks such as:

⚖️ Bias — unfair outcomes for certain users or languages

🔒 Privacy — risk of exposing sensitive data

💬 Misuse — used outside of its intended purpose

🤖 Over-reliance — users trusting AI too much

As a technical writer, your job is to acknowledge these risks clearly, simply, and responsibly.

🪶 Step 2: Typical Structure of an AI Ethics Section

Here's the recommended structure used in real AI model cards (e.g., Google, Hugging Face, OpenAI):

Ethical Considerations

1. Bias and Fairness

Explain how the model may produce biased outputs, and what steps were taken to reduce it.

2. Privacy

Describe whether user data is collected, stored, or shared.

3. Intended Use

Clarify what the model is for and **what it should not** be used for.

4. Limitations

Describe known weaknesses or failure cases.

5. Mitigation Steps

Explain what ongoing actions are in place to reduce ethical risks.

🧠 Step 3: Example — Ethical Section for "TextInsight" API

Let's apply this structure to your TextInsight API (the AI summarisation tool from Lesson 18):

Ethical Considerations — TextInsight API

1. Bias and Fairness

Although TextInsight summarises text neutrally, its training data comes primarily from English-language news and academic sources.
This means summaries may reflect Western or academic writing norms, which could reduce accuracy for non-native English content.

2. Privacy

TextInsight does not store user text.
All input text is processed transiently and deleted immediately after summarisation.
However, users are strongly advised **not to send personally identifiable information (PII)**.

3. Intended Use

TextInsight is designed for:

- Summarising articles, chat transcripts, and documentation.

It is **not intended for**:

- Legal, political, or medical decision-making.
- Summarising sensitive or classified information.

4. Limitations

- May omit subtle context or sarcasm.
- Occasionally produces summaries that sound overly formal or repetitive.
- Performance decreases on non-English content.

5. Mitigation Steps

- Continuous retraining with more diverse datasets.
- Manual bias testing on region-specific content.
- Clear user guidelines on responsible use.

✅ Pro Tip: When writing ethics documentation, always use a calm, factual, and non-defensive tone. Avoid emotional language — instead, stick to what's being done to reduce risks.

Example:

❌ "We are not biased at all." ✅ "We continue to evaluate model outputs for potential linguistic and cultural bias."

🧩 Step 4: How to Add This to Your Documentation

Add this Markdown file to your /ai-docs folder as:

ethics-textinsight.md

And in your portfolio.html, add a new link under your AI section:

- ⚖️ Ethics & Bias Report

🎛️ Step 5: Example Visual Summary (Optional)

You can later visualise ethics in documentation using a simple table:

Category Description Mitigation Bias English-language bias Add multilingual data Privacy No data storage Transient data handling Misuse Used for sensitive content Add user disclaimers Transparency Explainable summaries Public model card ✍️ Your Task

Create a new Markdown file: ethics-textinsight.md

Paste the example content above

Save it and preview it with Cmd + Shift + V

Add the download link to your Portfolio AI Docs section

🌿 Reflection

Writing AI ethics documentation shows empathy, integrity, and professionalism. It's often the difference between "a writer who can document AI" and "a writer who companies trust with AI."

Would you like to continue to Lesson 20: Creating Explainable AI Documentation (XAI) — where we'll teach how to write docs that explain how AI models make decisions in a clear, non-technical way?