Ethical AI HR Prompter

AI-Powered Solution

A bias-aware prompt rewriting tool designed to support inclusive decision-making in HR workflows with a beginner friendly UX even a non technical user could use. It is a purpose driven innovation to meet:

- **Ethical hiring**: AI is increasingly used in hiring, yet bias often goes unchecked. The tool intervenes at the prompt level,where decisions begin.
- **Accessibility**: Many HR professionals aren't technical. This solution makes ethical AI usable and understandable for them.

Built with **fallback logic** to ensure reliability under real-world constraints, this prototype with resilience by pivoting from Relevance AI to mock logic, to show that ethical design is about delivering impact even when systems fail.

- Detects exclusionary or biased language in HR prompts
- Rewrites them using inclusive, ethical phrasing
- Transparently communicates bias detection and fallback behavior
- Offers a beginner-friendly UX for non-technical users
- Initially integrated Relevance AI for bias detection, but pivoted to mock logic due to API downtime .

🔗 **Live Demo**: Streamlit Cloud

Ethical Design Notes

This project was built with ethical transparency, user trust, and practical resilience at its core. Each design decision reflects a commitment to responsible AI, especially in high-stakes HR contexts.

Bias Detection Goals

The tool was originally designed to integrate with Relevance AI to detect biased or exclusionary language in HR prompts — particularly around age, gender, and cultural assumptions. The goal was to:

- Surface subtle forms of bias that often go unnoticed in hiring workflows
- Provide actionable feedback to help users reflect on their language choices
- Encourage inclusive phrasing of candidates

Although the Relevance AI agent was initially integrated, its API became unavailable during development. This led to a strategic pivot toward mock logic — simulating bias detection while preserving the educational intent.

Inclusive Rewriting Logic

Rather than simply flagging bias, the tool actively rewrites prompts using inclusive, ethical phrasing. This logic is:

- **Modular**: Rewriting functions are designed to be easily extended or customized
- **Context-aware**: Rewrites aim to preserve the original intent while removing exclusionary language
- **Empowering**: Users receive rewritten prompts alongside short explanations, helping them learn and improve

Examples include replacing "young and energetic team" with "collaborative and dynamic team," or "must be native English speaker" with "strong written communication skills."

Fallback Transparency

When the Relevance AI API became unavailable, mock logic was implemented to simulate bias detection. Rather than hiding this pivot, the tool:

- Clearly informs users when mock logic is being used
- Labels simulated responses as such, maintaining ethical transparency
- Ensures that users understand the limitations of the demo without undermining its core message

This approach reinforces trust and models responsible behavior in AI development — especially under time constraints.

Real-World Reliability Under API Constraints

Hackathon environments often involve unstable APIs, limited time, and unpredictable bugs. To ensure the demo remained resilient and judge-ready:

- Mock logic was used to simulate bias detection and rewriting
- Environment variables were carefully managed to avoid runtime errors
- The UX was simplified to reduce friction for non-technical users
- All fallback behavior was transparently communicated to users

This reliability-first mindset ensured that the ethical goals of the project were upheld, even when technical blockers emerged.

README:

Ethical AI HR Prompter

A bias-aware prompt rewriting tool designed to support ethical decision-making in HR workflows. Built for the Employment Hero Hackathon to showcase responsible AI in action.

What It Does

- Detects biased or exclusionary language in HR prompts
- Rewrites prompts using inclusive, ethical phrasing
- Handles API outages with mock logic for demo reliability
- Offers transparent feedback to users about bias and fallback behavior

Live Demo

Try it now on [Streamlit Cloud](https://ethical-ai-hr-prompter-jtstt7mstljwvrgurlarfz.streamlit.app/)

Tech Stack

- Python + Streamlit
- Mock logic for fallback reliability
- Modular prompt rewriting functions
- Beginner-friendly UX for non-technical users

How to Run Locally

```bash
git clone git@github.com:divrajgit/ethical-ai-hr-prompter.git
cd ethical-ai-hr-prompter
pip install -r requirements.txt
streamlit run app.py
```

Lessons Learned

- Resilience Over Perfection: Real-world demos rarely go as planned. When Relevance AI's API went down, pivoting to mock logic ensured the tool remained functional without compromising the core ethical message.
- Transparency Builds Trust: Communicating fallback behavior openly (e.g., "This is a simulated response due to API outage") helped reinforce ethical design principles and user trust.
- Simplicity Wins: A beginner-friendly UX with one input box and clear feedback proved more impactful than complex features. It made ethical AI accessible to non-technical users, especially in HR contexts.
- Ethics Is a Design Layer: Bias mitigation isn't just a backend task — it's a UX decision, a communication strategy, and a reliability safeguard. Every design choice reflected that.
- Hackathons Are Labs for Integrity: Constraints like time pressure, API instability, and unfamiliar tools pushed creative problem-solving. But they also reinforced the importance of building responsibly, even in prototype form.