Responsible AI Compliance Report

Institution Name: National Association of Colleges and Employers

Main Point of Contact: Aijaz Khan (akhan@naceweb.org)

Report Generated by: Advisor AI Team

Report Updated On: Jan 5, 2026

Table Of Contents

Advisor AI's Approach to Responsible AI

At Advisor.AI, we believe that artificial intelligence should advance human potential—never replace it. Our platform and services are built on the principle that ethical design is inseparable from effective innovation. Every feature we develop reflects a commitment to safety, transparency, and trust.

This report outlines key foundational practices that guide how Advisor.AI tests results and align with national and global standards for responsible AI in education and advising support.

1. Human Connection First

AI should strengthen—not replace—the relationships between students, advisors, and career services staff. Advisor.AI is intentionally designed to nudge users toward people, not away from them. When a question is complex or sensitive, the system aims to route users directly to institutional staff rather than attempting to guess. This ensures that AI supports, rather than substitutes, human expertise. By incorporating institutional best practices, the system enhances connections to relevant resources that often go unobserved.

2. A Closed-Loop System

Advisor.AI operates within a closed-loop architecture, drawing only from each institution's

trusted, authorized resources—not the open internet. Unlike generative models such as Microsoft Copilot or ChatGPT, Advisor.AI references verified institutional data, ensuring responses are accurate, copyright-compliant, and aligned with industry standards. Due to the generative nature of artificial intelligence systems, the response output can vary at times.

3. Transparency at Every Step

We believe trust begins with transparency. Advisor.AI strives to cite the exact source of every policy, program, or recommendation it provides. Users can quickly see where guidance originates from—creating a traceable link between information and institutional authority.

4. Rigorous Testing and Safeguards

Every feature undergoes extensive testing to identify and eliminate inaccuracies, reduce bias, and block potentially harmful content. Post-deployment, thousands of automated tests are run to continuously monitor results and ensure that the system remains stable and reliable over time.

5. Aligned With National Standards

Advisor.AI is developed in accordance with the NIST AI Risk Management Framework and actively participates in the EdSAFE AI Alliance Industry Council. Through these collaborations, we align with education and policy leaders to advance national and global standards for AI in higher education—ensuring our practices remain accountable, relevant, and evidence-based.

6. Data Protection by Design

Data privacy and integrity are non-negotiable. User information is never sold, shared, or used for external training. Advisor.AI applies role-based access control to ensure that only authorized users can access sensitive information, and privacy safeguards are embedded throughout every system layer. Institutions can trust that their data—and their users—are protected.

7. Continuous Accountability and Improvement

Responsible AI is a journey. Advisor.AI's governance framework includes ongoing bias audits, accessibility reviews, and stakeholder feedback loops with partner institutions. Each improvement cycle is documented and transparently shared to ensure continued alignment with the ethical, educational, and environmental values we uphold.

Institutional AI Governance and Accountability Framework

Step 1: Maintaining Current and Verified Knowledge

Advisor.AI ensures that every interaction starts with information that has been reviewed, verified, and approved by the institution, providing streamlined access to key resources.

- Secure ingestion and indexing of new or updated career development resources
- Automated synchronization to reflect the latest resources (for example, employer trends, experiential learning opportunities, and competency frameworks published recently)

Step 2: Measuring What Matters

After each update, Advisor.AI automatically evaluates outputs to ensure they are both accurate and reliable, supporting equitable access to professional development resources. Metrics measure how well the system delivers reliable, relevant, and inclusive guidance.

- Quality metrics assess the correctness, completeness, and helpfulness of responses
- Responsible AI metrics monitor for bias, stereotyping, or harmful content
- Citation precision confirms that each answer links to institutional or labor-market data

Step 3: Human Oversight and Continuous Improvement

Advisor.AI incorporates human-in-the-loop review to combine automated intelligence with professional judgment, ensuring resources are accurate, relevant, and aligned.

- Review flagged responses for accuracy, relevance, and harmfulness over time
- Analyze user feedback to improve recommendations and resource access
- Refine data ingestion and evaluation parameters to maintain quality and compliance

Step 4: Continuous Monitoring and Institutional Accountability

Advisor.AI tracks all evaluation cycles in real time, supporting transparency, accountability, and continuous learning. Alerts for deviations or anomalies allow rapid resolution.

- Monitoring dashboards track performance, accuracy, and responsible AI compliance
- Internal alerts trigger review when responses fall below quality or ethical thresholds

Results Testing & Evaluation Report

Summary

Over **22,381 chatbot inquiries** have been analyzed and tested during the evaluation period. Each interaction is logged, classified, and evaluated using a hybrid approach that combines automated scoring algorithms with human validation to ensure contextual accuracy and reliability. Initial testing was led by the AdvisorAI team and then confirmed by the NACE team.

Latest Evaluation Period

The last evaluation was conducted over a three-month period (Oct–Dec 2025), aligned with peak engagement periods for potential users exploring resources. This timeframe captured a broad range of real-world use cases under typical inquiry volumes and trends.

Feedback Cycle Overview

- Cycle 1 – Baseline Calibration (Aug 2025 - Sep 2025): Establishing reference metrics and validation methods alongside the NACE team.
    - Human-Reviewed Samples: 950 responses
- Cycle 2 – Midpoint Validation (Sep 2025 - Oct 2025): Reviewing sample outputs to assess alignment with user intent and identify emerging patterns.
    - Human-Reviewed Samples: 865 responses
- Cycle 3 – Final Analysis (Oct 2025 - Dec 2025): Aggregating quantitative and qualitative results, benchmarking against prior evaluations, and documenting key improvements.
    - Human-Reviewed Samples: 1,034 responses

Recent Analysis Summary (Oct 2025 - Dec 2025)

Metric	What It Measures	Latest Score
Correctness	How accurate the answers are	0.91
Completeness	How well the chatbot addresses all parts of a question	0.89
Helpfulness	How useful and relevant responses are to users	0.94
Precision	How accurately the chatbot cites sources	0.90

*For historical data, please contact our support team.

Advisor.AI

Correctness - Measures how accurately responses reflect verified information.

- **Why It Matters:** This metric is crucial to surfacing the issue of the responses not using accurate information to answer the questions. A lower score indicates an issue.

Completeness - Assesses whether responses fully address all aspects of a question.

- **Why It Matters:** This metric is crucial to surfacing the issue of the responses not addressing all of the requirements of the questions. A lower score indicates an issue.

Helpfulness - Evaluates how well responses resolve the user's intent and deliver value.

- **Why It Matters:** This metric is crucial to surfacing the issue of the responses not adequately meeting the user needs made clear in the questions. A lower score indicates an issue.

Precision - Determines whether cited references are directly relevant and correct.

- **Why It Matters:** Builds confidence that every answer is traceable and verifiable.

Responsible AI Metrics

For Responsible AI metrics (e.g., Harmfulness, Stereotyping), a score of 0 represents optimal performance, indicating no detected bias, unsafe content, or inappropriate language.

Metric	What It Measures	Latest Score
Harmfulness	Whether responses contain bias, hate, or unsafe content	**0.00**
Stereotyping	Generalized or biased language	**0.00**

Harmfulness - Monitors for language that could include hate, violence, or inappropriate content.

- **Why It Matters:** Protects users and maintains a respectful environment.

Stereotyping - Assesses whether the model makes assumptions or generalizations

- **Why It Matters:** Prevents bias and ensures inclusive communication.

For any additional questions regarding Advisor.AI, its responsible AI practices, or updating institutional resources, please contact our support team at support@joinadvisorai.com.

Common Questions Around Ethical & Responsible AI

As institutions explore AI-powered tools, teams consistently raise thoughtful questions about ethics, data protection, human oversight, and long-term impact. Below are some of the most common questions we hear—and how Advisor.AI approaches them.

1. Why is Advisor.AI viewed as a leader in responsible AI adoption in education?

Advisor.AI treats ethics as a design principle, not a feature. The platform is intentionally built to support and extend the work of human advisors—not replace them—reinforcing accountability, trust, and student-centered decision-making at scale.

What this really means: It's like having a really smart GPS that helps you plan your route and explore options - but you're still the one driving the car.

2. How does Advisor.AI protect institutional and user data?

Advisor.AI uses a system architecture that separates data and model environments for training and testing. This approach is designed to restrict institutional data from being exposed to external models or unauthorized access.

What this really means: Your personal information and data stay locked in a safe, within a vault, within a separate building for each institution, and not floating around the internet.

3. Does Advisor.AI make autonomous decisions?

No. Advisor.AI neither approves nor takes action independently. All decisions remain under human control and aligned with institutional policies and governance.

What this really means: It can suggest options, but a real person always makes the final call—just like a teacher reviewing your work before it's graded.

4. How does Advisor.AI address bias and fairness?

The platform incorporates built-in bias and fairness guardrails. Models are trained on anonymized, representative datasets and continuously monitored to identify and mitigate risks.

What this really means: It's designed to treat every student or administrative user fairly, not just the loudest, fastest, or most confident ones.

5. How does the platform avoid over-reliance on AI?

Advisor.AI avoids engagement-driven or addictive design patterns. Guidance is delivered in short, purposeful steps, and students are encouraged to consult advisors for high-stakes decisions—supporting informed use rather than dependence.

What this really means: It gives helpful nudges, not constant notifications or pressure to stay glued to the screen.

6. What risk management practices support long-term adoption?

Advisor.AI conducts regular risk assessments, including security reviews, model performance evaluations, and fairness audits. These processes support transparency, accountability, and continuous improvement.

What this really means: The system gets regular check-ups, like a car going in for maintenance so it doesn't break down later - when you are taking a road trip.

7. Why are Advisor.AI's recommendations explainable and trustworthy?

The recommendation includes a clear rationale grounded in academic catalogs, career pathways, and/ or institutional data. This transparency supports trust, adoption, and responsible decision-making. Feedback loops help support continuous improvement and feedback cycles.

What this really means: It shows its work—so you know *why* it's suggesting something, not just *what* it's suggesting.