



School Psych AI Ethics Council Guiding Document

Overview

Welcome to the Guiding Document of the School Psych AI Ethics Council.

This document establishes a clear framework for the ethical use and development of AI in school psychology. Our primary aim is to ensure that AI technology supports students in a way that is fair, safe, and respects their privacy. We recognize the unique position we are in as leaders in integrating AI into school psychology. With this comes the responsibility to set a high ethical standard. Our guiding principles are crafted to help professionals navigate this new territory, ensuring that the introduction of AI tools into educational settings is done thoughtfully and responsibly.

Our focus on fairness and equity is about actively reducing biases in AI to ensure every student benefits, regardless of their background. This involves careful consideration of the data and algorithms we use, striving for representation and fairness across all demographics.

Safety and reliability are non-negotiable. We commit to rigorous testing and validation of AI tools to guarantee their effectiveness and safety for student assessments. Privacy is also paramount; we adhere strictly to laws and standards to protect student data.

Inclusivity guides our development process. We design AI tools to meet the varied needs of students, including those often marginalized. This commitment extends to our approach to transparency. We believe in clear communication about how our AI tools work, their purposes, and their limitations, ensuring that all stakeholders, especially those without technical backgrounds, are well informed.

Accountability underpins our use of AI. We emphasize the critical role of professional judgment alongside AI recommendations, ensuring that decisions are always made with the student's best interests in mind.

School psychologists maintain a key role in this balance, ensuring AI tools are used responsibly and effectively. **This guidance is a living document, subject to continuous refinement as we learn to integrate AI in school psychology ethically.** We invite feedback and collaboration from all stakeholders to ensure these guiding principles meet the needs of our field.

SPAI Ethics Council Special Acknowledgment

A special acknowledgement to those who helped lead the way in developing the School Psych AI Ethics Guiding Document:

Terence Jackson, Chief Security Advisor, Microsoft

Dan Florell, Ph.D., NCSP – Eastern Kentucky University

Dr. Soye Zaid-Muhammad, Author of The Power of A Thousand Heroes (The PATH)

Micah Reid, Machine Learning Engineer

Rachele Teson, School Psychologist

Juan Felipe Gomez, PhD candidate, Harvard university, Calmon Lab

Jessica Markham-Anderson, School Psychologist, Boston, MA

Amber D Gibson M.A., Ed.S School Psychologist Certified Trauma Professional

Dr. Byron McClure, NCSP, Founder, School Psych AI

School Psych AI Six Guiding Principles

Fairness and Equity:

- **Adapt AI Tools for Equal Accessibility:** Ensure that AI tools for use by school psychologists and school administrators are designed with the best interest of students in mind. By minimizing demographic bias for all groups of students to the extent technically possible. This includes being mindful of biases in AI algorithms, especially concerning gender, race, or socioeconomic status.
- **Incorporate Diverse Data Sets:** Use diverse data sets that have enough data to fairly represent all major demographics we serve in AI development to reduce biases and ensure that tools are equitable for all students across the US, including but not limited to those with disabilities, those with underrepresented and varied identities, and from various racial, cultural, religious, and geographic backgrounds.
- **When developing AI tools, include a diverse team from varied backgrounds, especially regarding gender and race. Include domain experts, including but not limited to practicing school psychologists, when developing or making major changes to AI tools.**
- **Considerations**
 - **Bias Assessment and Correction:** Regularly assess AI tools for algorithm biases and correct them to ensure fair treatment across all student groups, using metrics collaboratively chosen by ML researchers and domain experts in school psychology.
 - **Diverse Training Data:** Use training data representing various demographics, including race, gender, and socio-economic backgrounds, to ensure AI tools do not perpetuate existing inequalities. Monitor datasets regularly for demographic gaps or other instances of disproportionality and take swift corrective steps to rebalance datasets when necessary. Create specific guidelines for each AI tool to take corrective actions for when imbalances are detected.
 - **Inclusive Design and Development:** To ensure diverse perspectives are considered, a diverse group of stakeholders, including educators from various backgrounds, should be involved in the design and development of AI tools.

Reliability and Safety:

- **Ensure Consistency in Psychological Evaluations:** Utilize AI tools that have been rigorously tested and validated for their reliability in psychological assessments, ensuring consistent and accurate support. Regularly compare with best in class non AI solutions to ensure equal or greater efficacy.

- **Prioritize Student Well-being:** Always prioritize student safety and well-being in the use of AI tools, particularly in sensitive areas of psychological evaluation and intervention.
- **Create procedures utilizing human-in-the-loop training** to ensure consistent quality and formatting for generative output specific to each generative AI tool. Use human auditing by hiring domain experts to audit random samples of real-life output, or utilize user-friendly customer feedback tools for each output. Use this human-labeled and audited data continuously to ensure long-term model performance and identify and address emerging problem areas.
- **Considerations**
 - **Evidence-based Validation:** Ensure that AI tools are based on evidence-based practices and validated through rigorous scientific research for their effectiveness and safety in psychological evaluations.
 - **If we choose to consciously object to an existing practice in school psychology, especially in the case of choosing not to uphold already biased practices,** enlist help from domain experts to justify and communicate these decisions to the internal team and all users.
 - **Regular Performance Monitoring:** Continuously monitor and update AI systems to maintain their accuracy and reliability over time.
 - **Risk Management Strategies:** Develop and implement risk management strategies to address potential safety concerns related to using AI in psychological assessments.

Privacy and Security:

- **Safeguard Student Information:** Adhere strictly to FERPA, HIPAA, and other relevant privacy laws in managing student and health data, ensuring that AI tools comply with these regulations.
- **Implement Robust Security Protocols:** Use secure data storage and processing methods, and maintain transparency in how student data is used and protected.
- **Considerations**
 - **Compliance with Privacy Laws:** Strictly adhere to laws like FERPA and HIPAA, ensuring all AI tools comply with legal standards for student data privacy.
 - **Data Encryption and Security Protocols:** Implement strong encryption and security protocols to protect sensitive student data from unauthorized access or breaches. Comply with standards such as SOC2 and hire third-party penetration testers and/or security auditors.

- **Controlled Data Access:** Limit access to sensitive data strictly to authorized personnel and establish clear protocols for data sharing and usage.
- **Create a published shared responsibility model** with customers outlining the protective measures we take, and the protective measures we expect them to take for full compliance.

Inclusiveness:

- **Accommodate Diverse Learning Needs:** Ensure that AI tools are accessible and beneficial to all students, including those with special education needs or from underrepresented communities.
 - **Ensure AI tools for school psychologists are inclusive** and serve all students, especially those with special needs or from underrepresented communities.
- **Customize AI Applications:** Tailor AI tools to address the specific needs and contexts of different student groups, ensuring inclusivity in their design and application.
 - **Use the simplest tools necessary** to meet client needs to ensure maximum transparency and prevent unpredictable outcomes.
- **Considerations**
 - **Accessibility Features:** Ensure AI tools have accessibility features for students with disabilities or special needs.
 - **Cultural and Contextual Relevance:** Adapt AI tools to be culturally and contextually relevant for students from various backgrounds.
 - **User-Centered Design:** Involve students, educators, and psychologists in the design process to ensure the tools meet the diverse needs of all users.

Transparency:

- **Communicate Clearly About AI Tools:** Be transparent about how AI tools function, their data processing methods, and the basis for their recommendations. This helps in building trust among educators, students, and parents.
 - **Follow best practices around consent and transparency** as defined by competitors, and exceed their standards for transparency. If we forgo a common consent process, explain why this decision was made both internally and externally.
- **Involve Stakeholders in AI Integration:** Engage educators, parents, and students in discussions about the use of AI in school psychology to ensure transparency and address any concerns.
 - **Provide clear and understandable explanations and infographics** of how AI tools work and make decisions, suitable for non-technical audiences.

- Openly communicate the limitations and potential errors of AI tools to users to set realistic expectations.
 - User Training and Support: Offer comprehensive training and support to users for them to understand and effectively use AI tools.
- Own as much technical tooling as possible (typically by fine-tuning open-source models) so that all design choices are explainable internally and externally. Use third-party APIs as a last resort, disclose their usage, and disclose all additional data protection and sanitization methods to ensure safety when sending data to external providers.
- Considerations
 - Most of our stakeholders do not have technical backgrounds, take extra steps to ensure that ML models are explainable at a high level to school staff and parents of any background
 - Many prebuilt ML tools are not transparent, even to those with a technical background. Strive to use and build upon existing tools with high transparency whenever possible.
 - The highest priorities for stakeholder communications are being able to explain bias metrics and performance, data sourcing, and third-party handling of data. Model creation, architecture, and training details are a very low priority due to their highly technical nature and low direct impact on customer experience, reliability, and safety.

Accountability:

- Balance AI Recommendations with Human Judgment: While utilizing AI tools for insights, maintain a balance with professional human judgment, ensuring that decisions are ethically sound and in the student's best interest.
 - Clearly highlight where humans need to provide input, proofreading, and judgment, to prevent overreliance on tools that may provide inaccurate results.
- Maintain Professional Oversight: Establish clear lines of responsibility and accountability in the use of AI tools, with school psychologists playing a key role in monitoring and evaluating their impact on student outcomes.
- Considerations
 - Professional Responsibility: Ensure that school psychologists remain ultimately responsible for decisions made with the assistance of AI, not solely relying on AI-driven recommendations. Express this through messaging and a user experience, making this an easy and obvious step in utilizing any AI-generated artifacts.

- **Ethical Decision-Making Framework:** Develop a framework for ethical decision-making when using AI tools, including guidelines on when and how to use these tools, focused on student outcomes.
- **Ongoing Evaluation and Feedback Mechanism:** Implement mechanisms for ongoing evaluation of AI tools' impact on student outcomes and solicit regular feedback from users for continuous improvement.

References

[Microsoft 6 Principles](#)

[Microsoft Detailed Approach](#)

[Powerschool 6 Principles](#)

[NASP Social Justice](#)

[Aequitas](#)

[IBM AI Fairness 360](#)