NATIONAL SOCIAL STUDIES ASSOCIATION CONFERENCE

April 2-4, Las Vegas, NV

"DEVELOPING GUIDELINES AND BEST PRACTICES FOR ENSURING ETHICAL AND RESPONSIBLE USE OF CHATGPT AND OTHER LARGE LANGUAGE MODELS IN SCHOOLS."

Presented by Dr. Greg Levitt & Dr. Steven Grubaugh – UNLV

THREE (3) ASSUMPTIONS

- Students will use ChatGPT/BARD/Bing/Dall-e and other AI resources.

- Teachers will use ChatGPT and other AI Resources.

- AI is here to stay so we need to learn to get the best use out of all AI tools.

1. PRIORITIZE TRANSPARENCY

1.

• Clearly communicate the use of ChatGPT/BARD/Bing/Dall-e and other AI in the educational setting to students, parents, and educators.

• Outline its purpose, benefits, limitations, and potential risks to ensure informed use.

• Require students to show or label, 1) Any text/images generated by AI, 2) their own work/changes to AI generated text/images, and 3) describe how their own work improved, corrected any bias, or information in AI generated text or images.

Example 1: Include a section in the school's handbook or website explaining the role of ChatGPT in the learning process, its capabilities, and limitations.

Example 2: Hold informational sessions or workshops for parents and students to demonstrate how ChatGPT is used in classrooms and address any concerns.

2. PROTECT PRIVACY AND DATA

- Implement robust data privacy policies, handling student and educator data securely.

- Use encryption, anonymization, and data access restrictions to prevent unauthorized access and data breaches.

Example 1: Regularly review and update the school's data privacy policy, ensuring it adheres to relevant data protection regulations.

Example 2: Employ data protection officers to oversee the management of data and to ensure compliance with privacy regulations.

3. ENSURE INCLUSIVITY

- Actively promote inclusivity and address potential biases in AI-generated content.

- Provide training for educators on identifying and addressing biases and promote diverse perspectives in the learning environment.

Example 1: Offer professional development workshops for teachers that focus on understanding and mitigating biases in AI-generated content.

Example 2: Encourage students to engage in discussions about biases they encounter in AI responses, fostering a culture of critical thinking and inclusivity.

See examples of bias on the last slide.

4. MAINTAIN SUPERVISION AND ACCOUNTABILITY

- Assign educators the responsibility to monitor AI-generated content and intervene when necessary.

- Establish a clear chain of accountability for AI tool usage in the school.

- Establish clear rules and consequences for student AI tool usage in school.

Example 1: Create a system for educators to report and address issues or concerns related to ChatGPT or BARD-generated content, ensuring timely intervention and resolution.

Example 2: Implement a review process where teachers and administrators evaluate the effectiveness of ChatGPT usage in the classroom, discussing successes and areas for improvement.

5. ENCOURAGE RESPONSIBLE USE

• Teach students about digital citizenship and the responsible use of AI tools like ChatGPT/BARD/Bing/Dall-e/Etc and others.

• Promote critical thinking and independent verification of AI-generated content.

• Require student use of plagiarism checkers before turning in assignments.

Example 1: Integrate digital citizenship and AI literacy lessons into the curriculum, emphasizing the ethical use of AI tools.

Example 2: Design classroom activities that require students to verify the information provided by ChatGPT, fostering a habit of critical thinking and fact-checking.

6. EVALUATE EFFECTIVENESS

● Regularly assess the impact of ChatGPT and BARD on student learning outcomes.

● Adjust its implementation based on evaluation results to improve effectiveness.

Example 1: Conduct semesterly surveys to gather feedback from students and educators about their experiences with ChatGPT and its impact on learning.

Example 2: Analyze student performance data in correlation with ChatGPT usage to identify areas where the tool is most effective or requires adjustments.

7. PROVIDE TECHNICAL SUPPORT

- Regularly assess the impact of ChatGPT and other AI on student learning outcomes.

- Adjust its implementation based on evaluation results to improve effectiveness.

Example 1: Conduct semesterly surveys to gather feedback from students and educators about their experiences with ChatGPT and its impact on learning.

Example 2: Analyze student performance data in correlation with ChatGPT usage to identify areas where the tool is most effective or requires adjustments.

8. FOSTER COLLABORATION

• Encourage collaboration between students, educators, parents, and AI experts in the development of guidelines and best practices for using ChatGPT/BARD/Bing/Dall-e and other AI in schools.

• Engage all stakeholders in discussions and decisions about AI implementation in the classroom.

Example 1: Organize focus groups or town hall meetings with students, parents, educators, and AI experts to discuss the use of ChatGPT and gather input on its implementation.

Example 2: Establish a steering committee with representatives from all stakeholder groups to guide the development and review of AI usage policies and guidelines.

9 ADDRESS ACCESSIBILITY

• Ensure that ChatGPT and other AI tools are accessible to students with diverse needs and abilities.

• Provide necessary accommodations and support to promote equal access to AI-enhanced learning experiences.

Example 1: Collaborate with special education experts to adapt ChatGPT usage for students with learning disabilities or other special needs, providing necessary accommodations and modifications.

Example 2: Train educators on using ChatGPT and other AI tools to support English language learners or students from different cultural backgrounds, promoting equitable access to learning opportunities.

10. STAY INFORMED AND ADAPTABLE

• Keep up to date with advancements in AI technology and ethical considerations.

• Be prepared to revise and update guidelines and practices as needed to ensure continued ethical and responsible use of ChatGPT and BARD in schools.

Example 1: Organize annual AI ethics training for educators, keeping them informed about the use of AI in schools.

Example 2: Assign a teacher (include students and parents) committee to review and present ethical guidelines each year or semester as needed.

TEN EXAMPLES OF POTENTIAL BIASES IN AI-GENERATED CONTENT

Data bias: AI models are only as good as the data they are trained on, and if the training data is biased or incomplete, the resulting model will be biased as well. For example, if a facial recognition algorithm is trained only on images of light-skinned people, it may not accurately recognize people with darker skin tones.

Confirmation bias: AI systems may be designed to look for patterns or correlations in data that confirm pre-existing assumptions or biases, rather than seeking out all available information. This can lead to the perpetuation of stereotypes or misinformation.

Sampling bias: AI models may be trained on data that is not representative of the larger population, leading to inaccurate or biased results. For example, a natural language processing model trained on data from a specific region may not accurately understand or generate text for speakers from other regions.

Prejudice: AI models can exhibit prejudice against certain groups of people based on their ethnicity, gender, age, or other characteristics. This can result in discriminatory outcomes, such as biased hiring decisions or credit scoring.

Algorithmic bias: Certain algorithms may be inherently biased due to their design or structure. For example, decision trees can easily fall prey to bias if the training data is not diverse enough.

Overfitting: AI models that are too closely trained to the training data may result in overfitting, where the model performs well on the training data but poorly on new, unseen data. This can lead to biased results if the training data is not diverse enough.

Lack of diversity in development teams: AI models are often developed by homogeneous teams, which can lead to a lack of consideration for diverse perspectives and experiences. This can result in biased models that do not accurately reflect the needs or experiences of underrepresented groups.

Lack of transparency: AI models may be opaque or difficult to interpret, making it difficult to understand how decisions are being made. This lack of transparency can make it difficult to detect and address biases.

Cultural bias: AI models may be developed with assumptions about cultural norms and practices that do not apply to all groups. This can result in biased outcomes, such as a chatbot that does not understand the vernacular of certain regions.

Contextual bias: AI models may not take into account the broader context in which decisions are being made. For example, a credit scoring model may not consider the fact that certain groups of people have historically faced discrimination that has affected their ability to build credit.