



SASC Guidance on
The Use of Artificial Intelligence (AI)
in
Diagnostic Assessment and Report Writing

May 2026

author:

Peter Thomas, Distinguished Engineer and Principal Architect, Information Technology

with the support of the SASC Chair and members of the board:

Lynn Greenwold

Alice Thomas

Introduction & Principles

- **Purpose:** To establish a framework for the ethical, safe, and professional use of Generative AI Applications (e.g., ChatGPT, Google Gemini, Microsoft Copilot) in Specific Learning Difficulties (SpLD) assessments.
- **The Human-First Principle:** AI is recognised as an administrative tool, not a clinical agent. The diagnostic decision, interpretation of data, and final recommendations remain the sole responsibility of the qualified assessor.

Data Protection

- **Prohibition of Personally Identifiable Information:** Under no circumstances should Personally Identifiable Information be entered into applications containing AI functionality. This includes:
 - Full names (Individual, Parents, School Staff).
 - Dates of Birth.
 - Specific Addresses/Postcodes.
 - Sensitive data about individuals (e.g. medical history)
- **Anonymisation Protocols:** Assessors must strip all identifying data *before* inputting text for rephrasing (e.g., using "Learner X" or "The candidate").
- **Data Training Opt-Out:** Assessors are expected to configure AI tool settings to prevent input data from being used to train future models.

Professional Responsibilities

- By signing the report, the assessor accepts full legal and professional liability for every word, regardless of its origin.
- **Hallucination Risk:** Assessors must manually verify all AI-generated content, particularly:
 - Definitions of test subtests (e.g., CTOPP-2 or TOMAL-2 descriptors).
 - Statistical explanations.
- **Assessment Validity:** AI cannot observe an individual. Qualitative descriptions of behaviour must be rooted in the assessor's actual notes, not generated by AI based on a score.
- **Disclaimer about the use of AI:** Assessors should be transparent about their use of AI within the report. For example: 'While AI was used in the drafting of this document, all diagnostic decisions and profile analysis remains the sole responsibility of the assessor'. Assessors may want to include this type of statement in their terms and conditions.
- **Expertise:** Do not rely upon AI's specific expertise on a topic. For example, it is not an expert on tests and does not have the knowledge contained in test manuals.
Use of the report: Assessors may want to have a statement within their terms and conditions discouraging the recipient from using AI instead of the assessor to explain or summarise the report because of the risk of inaccuracy and data privacy concerns with the following wording "The report content is confidential and should be considered in its entirety. No part may be copied, quoted or manipulated through AI without prior consent of the author."

Acceptable Use (Low Risk)

- **Drafting & Refinement:** Using AI to improve the flow, grammar, and clarity of the assessor's own rough notes.
- **Generic Explanations:** Generating standard definitions for the Appendix (e.g., "Explain what a Confidence Interval is in simple terms").
- **Idea Generation:** Requesting some initial ideas of strategies for specific difficulties (e.g., "List 5 classroom strategies for weak verbal working memory in Year 4")—*provided these are vetted and selected by the assessor.*

Unacceptable Use (High Risk)

- **Diagnostic Decision Making:** AI must not be asked to determine the outcome (e.g., "Based on these scores, is this child dyslexic?").
- **Score Calculation:** AI must not be used to calculate chronological ages or convert raw scores to standard scores (due to high error rates in AI's mathematical logic).
- **Unverified Copy-Paste:** Copying large blocks of text without reading/editing is considered professional negligence.
- **Sensitive Information:** Inputting any personally identifiable information or sensitive information.
- **Disabling data usage:** Using an AI solution without understanding and minimising the data usage settings for training purposes.
- **Training or grounding the AI Application:** Whatever kind of document has been uploaded, an assessor should not rely on AI's analysis without review.
- **Breaking terms of use:** Uploading a document which breaches copyright or terms of use.

Assessor's Quality Assurance Checklist

- Has Personally Identifiable Information been removed from all prompts?
- Have all test scores been calculated manually or via publisher software (not AI)?
- Has the assessor reviewed in detail the entire AI produced content for correctness and accuracy?
- Have specific recommendations been tailored to the individual?
- Has an AI policy been shared with the person commissioning the report?

Critical Considerations of AI Usage

- **Accuracy:** How accurate will the AI be in producing the output?
- **Impact:** What is the impact of incorrect output?
- **Privacy:** What data privacy risks could be relevant?

Assessors should review the risks of using AI, and consider mitigating factors which may help in reducing risk.

Case study Examples

Scenario A: "Here are the scores: VR 110, PR 85, WM 70. Write a conclusion about why they have dyslexia."
(Delegating the clinical judgment to the machine).

- Accuracy: **RED**
 - It is unlikely that a Generative AI solution will be able to accurately analyse the data necessary to determine the cognitive profile, this would require a specialised model trained on many example diagnoses to form a Discriminative model.
- Impact: **RED**
 - Inaccuracies in the analysis or diagnosis will likely have a large impact on an individual's future.
- Privacy: **RED**
 - Personal and confidential data of the individual may have to be used within the model's analysis causing increased risk of data leakage
- Mitigating Factors: No specific mitigation could be performed, other than parallel analysis by a trained assessor to compare to the model's results.

Scenario B: "Rewrite this observation to be more professional: 'Child kept wriggling, dropped pencil 3 times, looked tired.'" *AI Output:* "The learner demonstrated motor restlessness, frequently dropping writing implements, and appeared fatigued."

- Accuracy: **AMBER/GREEN**
 - If the information is merely copied and pasted the risk increases. If it is reviewed and edited if necessary by the assessor, the risk reduces.
- Impact: **GREEN**
 - Inaccuracies in the text may not have a large impact on the individual's future.
- Privacy: **GREEN**
 - As long as personally identifiable information is not included, as in the example, the risk is low.
- Mitigating Factors: Mitigation will be through review of any content generated by AI using a human in the loop approach.

Scenario C: "Give recommendations for an individual aged 10 with difficulties in.."

- Accuracy: **AMBER**
 - The AI system may be able to produce some suitable starting points for recommendations which would be reviewed and edited by the assessor, but if the recommendations were taken as-is, then the risk increases.
- Impact: **AMBER**
 - As above, if the recommendations are taken as-is then the risk increases.
- Privacy: **AMBER**
 - If personal confidential data is omitted, the risk is reduced.
- Mitigating Factors: Thorough review of any content generated by AI using human in the loop approaches could reduce the risk of incorrect content being produced.

Scenario D: “Summarise an individual’s report into bullet points”

- Accuracy: **AMBER**
 - AI solutions can be effective at summarisation, but the assessor should thoroughly review the output for accuracy.
- Impact: **AMBER**
 - As above, if the summary is taken as-is then the risk increases.
- Privacy: **RED**
 - Personal and confidential data of the individual used within the report causes increased risk of data leakage
- Mitigating Factors
 - Anonymising and removing personal and confidential data or summarising specific sections of the report omitting personal or confidential data will reduce privacy risk.

Scenario E: “Explain a particular dyslexia topic in simple terms to share that with a parent.”

- Accuracy: **AMBER/GREEN**
 - AI can be good at simplifying or rewording complex topics, if an expert then reviews the output this can reduce any risk.
- Impact: **AMBER/GREEN**
 - If the output is taken as-is, then the risk is greater.
- Privacy: **GREEN**
 - No privacy risk if general information is requested.
- Mitigating Factors:
 - Thorough review of any content generated by AI using human in the loop approaches could reduce the risk of incorrect content being produced.

APPENDIX: Explanation of Key AI Concepts

- **Artificial Intelligence:** The theory and development of computer systems able to perform tasks normally requiring human intelligence.
- **Types of AI Application:**
 - **Predictive AI:** Identifies patterns in past events to anticipate future behaviour. E.g. predicting stock prices, predicting weather.
 - **Visual AI:** Interprets images or videos. E.g. analysing road conditions for a self-driving car, identifying tumours in medical scans.
 - **Narrow AI:** Focuses on a very specific task. E.g. recommendations in Netflix, diagnosing a specific medical condition
 - **Discriminative AI:** Used to classify or predict based on input data. Trained on a specific dataset of labelled data. Learns the relationship between the features of the data points and the labels.
 - **Generative AI:** Generates content in various forms – text, audio, image or video. E.g. generating a document on a topic, creating a piece of music of a particular style. Types of generative AI include ChatGPT and Google Gemini.
- **AI Model:** A computer application which is designed to mimic human intelligence by applying algorithms to input data to find relationships and generate outputs. For Generative AI solutions, the AI model is called a Large Language Model (LLM). Common LLMs are Google’s Gemini, OpenAI’s GPT or Meta’s Llama.
- **AI Application:** AI models are integrated into applications which provide AI functionality to an end user. E.g. OpenAI’s ChatGPT is an AI Application which provides a chat style interface to the GPT Large Language Model; Microsoft CoPilot is an AI Application which provides AI features within the Microsoft Office Suite and is based on the OpenAI GPT Large Language Model.
- **Bias:** AI can increase discrimination if trained on non-representative datasets. E.g. a skin cancer diagnosis might be biased if the training data lacked diverse skin colours.
- **Data risks:**
 - **Training Data Exposure:** AI systems may record input and output data and use this to further train a model resulting in that data being embedded in the model. This means that the data can be later queried by others, revealing sensitive information from the training data set.
 - **Malicious Access:** Hackers could access the model or records revealing confidential data.
 - **Privacy Violations:** Privacy regulations, such as data protection regulations, strictly control how personal data can be collected, processed and stored. Sending personal data to an AI Application for processing can constitute unauthorised processing, unless explicit consent has been obtained.
 - **Intellectual Property Risks:** Use of certain data within input to an AI Application may breach copyright, terms of use or confidentiality agreements.
- **Grounding:** A term in Generative AI which focuses the AI Application on using a verifiable set of data within the process. In essence, grounding can use the AI as a ‘language assistant’ rather than a ‘knowledge base’ by forcing the AI to use the data set as its primary data input.
- **Hallucination:** AI Hallucination refers to instances where an AI Application generates plausible-sounding but incorrect or fabricated output information as if it was factual. This is particularly an issue for Generative AI applications where the application is trying to respond convincingly based on its training data even when it lacks accurate information.

- **Human in the loop:** Human in the loop (HITL) is a way of designing and interacting with AI Applications where the human is the ultimate decision maker. Instead of letting the AI work unsupervised, it is the human who guides, supervises, corrects and reviews the AI Application at key points. The AI does a lot of the work, but the human is ultimately responsible for the outcome.

HITL is important because:

- Validates and corrects the output generated by the AI Application.
 - Handles edge cases which the AI Application could struggle with.
 - Ensures fairness and ethics to avoid biased or harmful outcomes.
 - Builds trust in the accuracy of the AI output.
 - Maintains accountability.
- **Non deterministic:** Generative AI models don't follow pre-defined rules, they are opaque as to how they arrive at the output, and it cannot be guaranteed that they will give the same output if asked the same question a second time.