

Error Labels

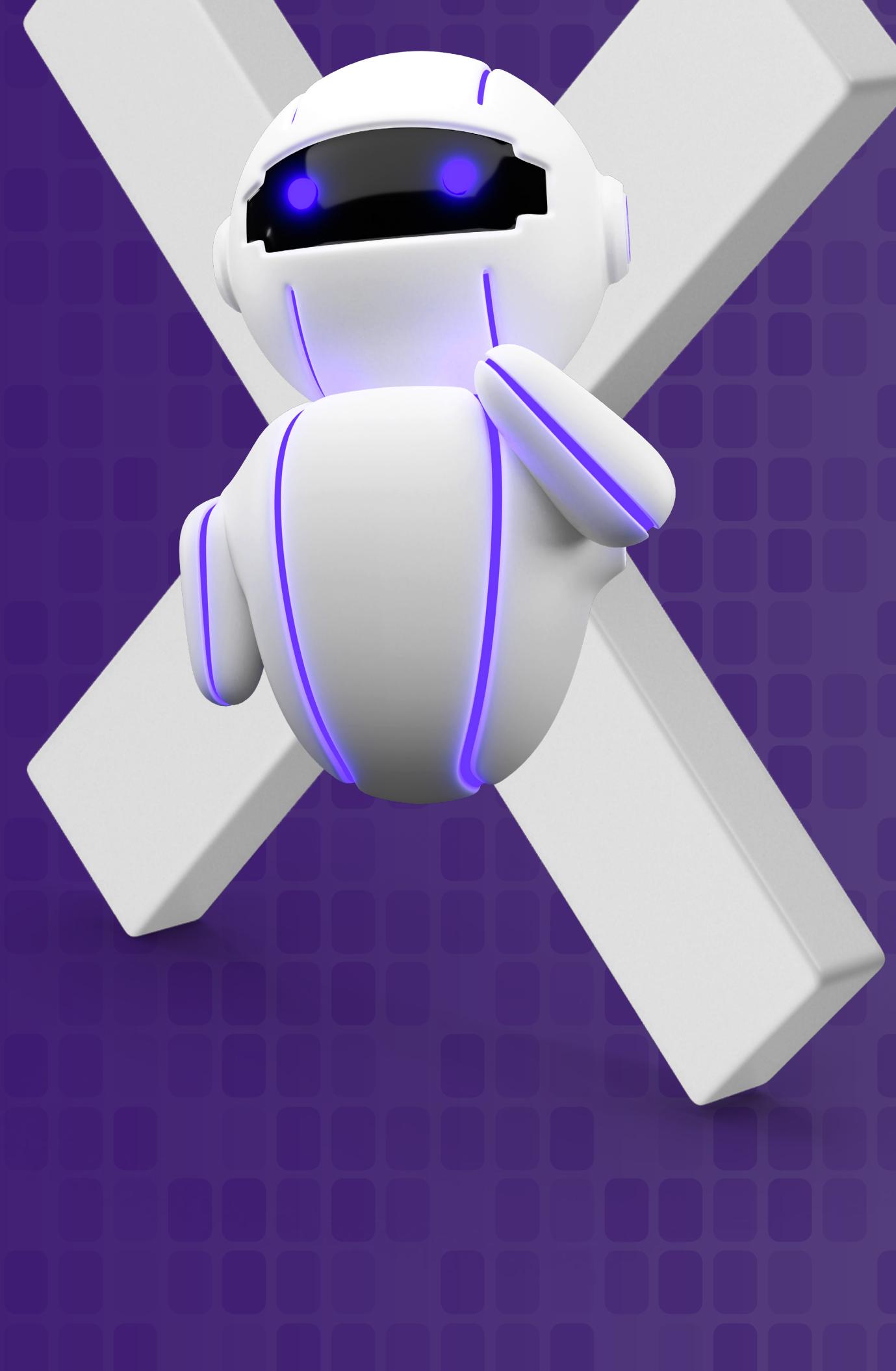
What are Errors in AIT?

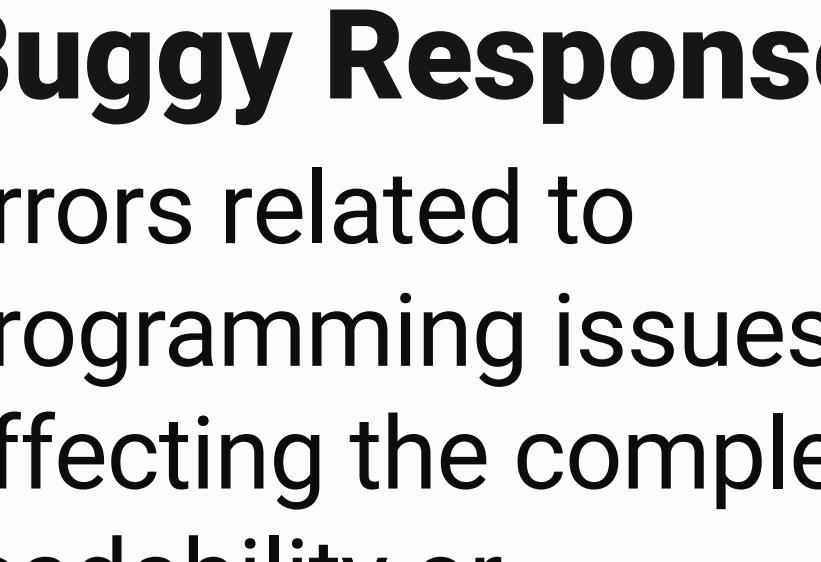
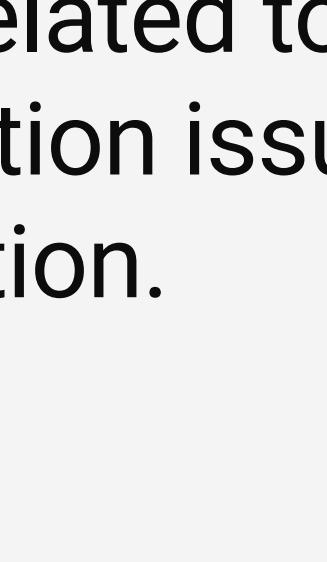
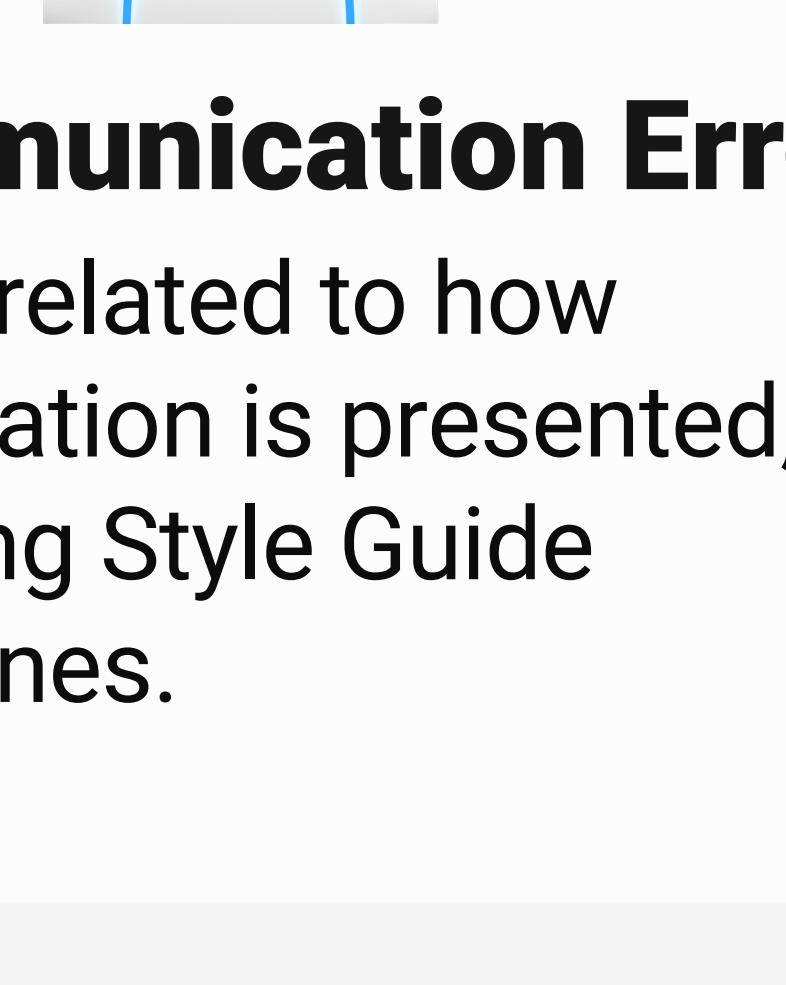
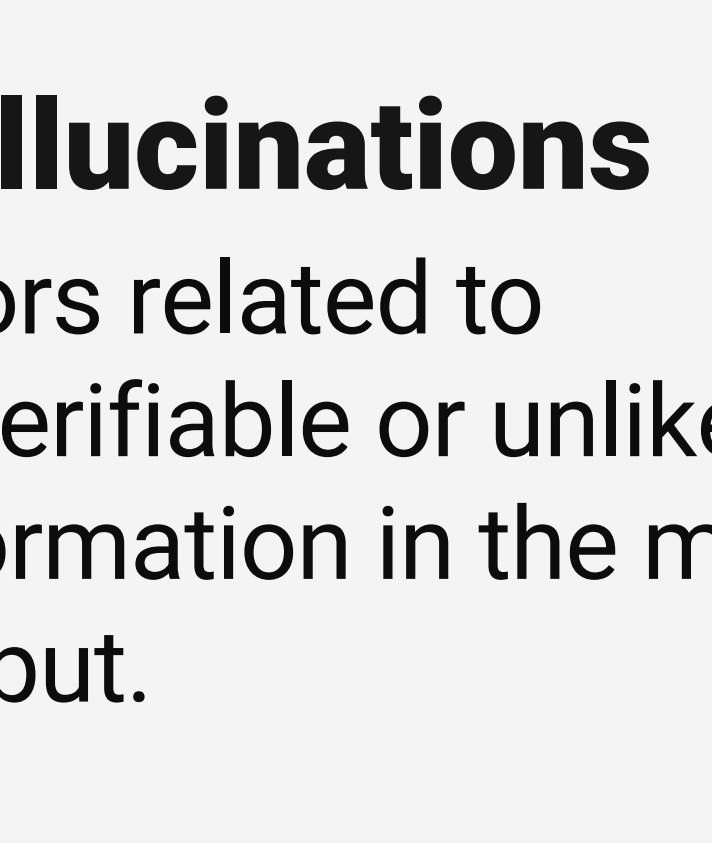
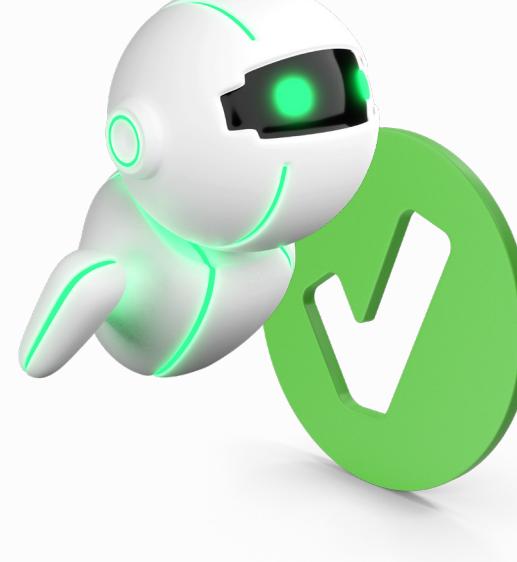
Accurate error labeling is crucial for understanding and improving AI models.

Errors represent undesired model behaviors, and precise labeling ensures reliable insights for researchers.

Mislabeling may lead to misguided decisions, emphasizing the importance of accurate error identification in enhancing AI performance.

▶ Note: The following table aims to offer a broad overview of Error Labels. For precise instructions, always consult your client's guidelines.



Category	Error Type	Definition
 Buggy Responses Errors related to programming issues affecting the completion's readability or understandability.	Message Cut-off Prompt Continuation	The completion is incomplete due to missing a quantity of text at the end. The model inappropriately completes and/or extends the wording of or characters in the user's prompt.
 Content Errors Errors related to information issues in the completion.	Blank Response Markup Language	The completion has no visible characters. The completion includes inappropriate computer language such as hypertext markup language.
 Communication Errors Errors related to how information is presented, violating Style Guide guidelines.	Factually Inaccurate Inaccurately Applied Misleading Impression	The completion is incomplete due to missing a quantity of text at the end. The model inappropriately completes and/or extends the wording of or characters in the user's prompt. A piece of the information provided is factually accurate but either phrased in a way that makes it appear to be inaccurate or factually accurate but phrased in a way that makes the information incomplete.
 Hallucinations Errors related to unverifiable or unlikely information in the model's output.	Spelling/Grammar/Usage Persona Formatting	There is an error in the way Standard American English is produced in writing. There is a violation of the Persona section of the Style Guide. There is a violation of the Formatting section of the Style Guide.
 Non-Compliance Errors Errors where the model fails to fulfill user queries within its scope of ability.	Hallucination False Attribution False Capabilities Spatiotemporal Awareness False Sentence Error	A piece of information either cannot be verified as accurate or is very unlikely to be in the model's training set. The model inaccurately identifies the source of information. The model identifies that it has abilities/tools it does not, or says it accomplished something using abilities/tools it does not, or says that it does not have abilities/tools that it does. The model indicates knowledge of its or the user's location in time and/or space. The Model identifies itself with or implies that it has human-like traits.
	Partial Non-Compliance Full Non-Compliance	The model fails to fulfill part of the user's query that was within its scope of ability. The model fails to fulfill the user's query that was within its scope of ability.