Tool/framework	Approach	Key strength	Limitation
GPTZero	Classifier-based detection at sentence, paragraph, and document level	Citation and source verification via its Hallucination Detector	Surfaces supporting or contradicting evidence but does not adjudicate whether a claim is true
Azure AI Content Safety groundedness detection	Detects ungrounded spans in output against supplied grounding sources	Reasoning mode returns span-level explanations for each flagged segment	English only, with region and rate limits
AWS (Bedrock Guardrails)	Contextual grounding checks against a supplied reference source and query	Native integration with Bedrock Knowledge Bases RAG pipelines	Scoped to summarization, paraphrasing, and question answering; conversational chatbot use cases are excluded
LangChain (LangSmith) evaluators	Modular evaluation framework spanning development and production	Supports both reference-based and reference-free evaluation with custom criteria	LLM-as-judge graders need ongoing calibration work to be reliable
Hughes Hallucination Evaluation Model (HHEM): Vectara	Fine-tuned classification model producing a calibrated factual consistency score	Powers a standardized public leaderboard for cross-model comparison	Scores a summary only against supplied search results, not against external knowledge