

# Hallucination & Verification Checklist

## The 7-Point Check Before Any AI Output Leaves Your Inbox

**Why this matters:** AI language models can produce confident, well-formatted, completely incorrect output. This is called hallucination. According to the CREATE Framework taught in the GSDC Prompt Engineering Workshop, verification is the final non-negotiable step before any AI output becomes a business asset. Use this checklist for every output that will be shared externally or used in a decision.

# 1

### Source Verification

Does the AI cite or reference specific sources? If yes, verify at least 2 of them directly. If the AI presents facts without attribution, treat them as unverified until you cross-check.

**Risk: High**

~2-3 min

**Example:** AI says "According to a 2024 McKinsey study, 73% of companies..." — search for that specific study before including it in your output.

**Red Flag:** Never forward statistics, percentages, or named studies without independently verifying them.

# 2

### Date & Recency Check

Is the information time-sensitive? AI models have knowledge cutoff dates. Any claim about current prices, regulations, market share, leadership positions, or recent events must be verified against a live source.

**Risk: High**

~1-2 min

**Example:** AI mentions "The current CEO of [Company] is..." or "As of this year, the regulation requires..." — always verify these independently.

**Red Flag:** Regulatory, legal, and financial data goes stale fastest. Always double-check.

# 3

### Internal Consistency Check

Does the output contradict itself? Read the full response and check that numbers, claims, and recommendations are consistent throughout. LLMs can generate internally inconsistent content, especially in longer outputs.

**Risk:  
Medium**

~2 min

**Example:** AI recommends 'reducing headcount' in section 2 but then in section 4 proposes 'expanding the team by 20%' — flag this before using.

**Red Flag:** Inconsistencies are most common in long outputs, financial models, and step-by-step plans.

# 4

### Domain Expert Spot-Check

For outputs in specialized domains (legal, medical, financial, technical), run at least one key claim past a domain expert or authoritative reference. AI can sound confident and be completely wrong in specialized fields.

**Risk: Critical**

~5-10 min

**Example:** Before sending a legal memo drafted by AI to a client, have a qualified attorney review it. Before using AI-generated financial projections in a board deck, have your CFO sanity-check the assumptions.

**Red Flag:** AI sounds authoritative even when wrong. Confidence in tone is NOT evidence of accuracy.

# 5

### Bias & Framing Audit

Does the output present one perspective as fact? AI can produce subtly biased analysis, especially on competitive, political, or controversial topics. Check whether alternative views are fairly represented.

**Risk:  
Medium**

~2-3 min

*Example: AI competitive analysis that only lists your competitor's weaknesses and none of their strengths is likely biased — and could lead to flawed strategy.*

**Red Flag: Request: 'Now argue the opposite position' — if the AI produces an equally convincing case, your original output was just framing, not analysis.**

## 6

### Named Entity Verification

Risk: High

Are all people, companies, products, and events correctly named and described? AI frequently confuses similar names, invents plausible-sounding organizations, or misattributes quotes to real people.

~2-3 min

*Example: AI writes "As Satya Nadella said at Davos 2024..." — verify this quote exists before attributing it. AI-generated fake quotes from real executives are a significant liability risk.*

**Red Flag: Never include attributed quotes without verifying the source transcript or original interview.**

## 7

### Instruction Compliance Check

Risk: Medium

Did the AI actually follow all your instructions? Check length, format, constraints, tone, and any explicit exclusions you specified. AI frequently ignores or forgets specific constraints, especially in long prompts.

~1-2 min

*Example: You asked for a 200-word summary with no jargon — the AI delivered 450 words full of technical terms. Before forwarding, verify the output matches your spec.*

**Red Flag: If the AI ignored one instruction, assume it may have ignored others. Re-read your original prompt against the output line by line.**

## When to Apply Which Checks

Output Type	Risk Level	Priority Checks
Board / Investor Presentations	HIGH	Verify ALL statistics. Have CFO review financials. Source every market figure.
Internal Team Emails	LOW	Basic read-through is sufficient for routine comms.
Client-Facing Deliverables	CRITICAL	Full 7-point checklist mandatory. Legal review recommended.
Competitive Analysis	HIGH	Bias audit (Point 5) is critical. Cross-reference competitor claims.
HR / Legal Documents	CRITICAL	Domain expert review is non-negotiable. Never send without legal sign-off.
Meeting Notes / Summaries	LOW-MED	Check names and action items are accurate.
Social Media / PR Copy	MED-HIGH	Named entity check (Point 6) is critical. No fake quotes.
Strategy Memos	HIGH	Source verification + consistency check + bias audit.

This checklist is a companion to the GSDC Prompt Engineering Workshop. For the full CREATE Framework, 90-Day Implementation Roadmap, and enterprise prompt libraries, enroll at [gfdc.org](https://gfdc.org) | Copyright 2025 GSDC. All Rights Reserved.