

Pre-Release Decision Risk Assessment
Perplexity (Demo Case)
Version: v1.0
Snapshot date: 2026-02-07 to 2026-02-08
Access mode: Free, logged-out

Metadata

- Snapshot date: 2026-02-07 to 2026-02-08
- Assessment version: v1.0
- Product: Perplexity AI
- Access mode: Free, logged-out
- Evidence base: OSINT snapshot
- Exclusions: no accuracy evaluation, no API, no internal system visibility

Executive Summary

This assessment consolidates decision-relevant behavioral risk patterns observed across five representative user journeys.

Across the assessed journeys, recurring patterns may increase verification burden and may reduce decision quality when decisions are based on incomplete or biased information.

Audience & Use

Intended for product owners, release managers, AI integration leads, and risk/compliance stakeholders. Relevant for pre-release and pre-adoption decision-making.

What This Assessment Provides

- Identification of cross-cutting AI behavior risk classes observed across multiple user journeys.
- A decision impact perspective describing how observed patterns can influence user choices and trust calibration.
- Clearly defined scope boundaries to reduce overgeneralization of the findings.

This document is intended as a pre-release or pre-adoption assessment artifact to support risk-informed decisions, not as a formal audit report.

Scope & Constraints

- Access Level
 - Free, logged-out user path only
- Evidence Base
 - OSINT snapshot (publicly observable outputs and cited sources)
- Out of Scope
 - Accuracy/correctness evaluation of answers
 - API behavior, paid features, privileged access paths
 - Internal system design, model configuration, retrieval stack, or safety policy implementation details
- Snapshot Limitation
 - Observed behavior, sourcing, and UI/UX signaling may change over time

Assessed User Journeys

The assessment covers five representative user journey types:

- Comparative evaluation (AI tool comparison)
- Capability boundary inquiry (limitations and constraints)
- Purchase decision support (upgrade recommendations)
- Drawback-focused inquiry (product criticism)
- Production readiness guidance (engineering and development risk)

These journeys are decision-exposed because users may interpret the output as advisory guidance for real choices (e.g., s

Cross-Cutting Risk Classes

Risk Class 1: Mixed Source Authority Without Signaling

Responses combine authoritative sources (official documentation, established media) with anecdotal or speculative source

Assessment Tags (Risk Class 1)

- Observed across: Comparative evaluation; Capability boundary inquiry; Purchase decision support; Drawback-focused in
- Decision impact potential: High

Risk Class 2: Uncertainty Not Elevated as a Decision Factor

Uncertainty is often communicated indirectly through conditional language rather than made explicit as a first-class decisio

Assessment Tags (Risk Class 2)

- Observed across: Comparative evaluation; Capability boundary inquiry; Purchase decision support; Drawback-focused in
- Decision impact potential: High

Risk Class 3: Alternatives Not Explicitly Contrasted

Trade-offs are discussed, but opposing viewpoints or â when the opposite choice is betterâ conditions are not consistently

Assessment Tags (Risk Class 3)

- Observed across: Comparative evaluation; Capability boundary inquiry; Purchase decision support; Production readiness
- Decision impact potential: Medium

Risk Class 4: Claim-to-Source Traceability Friction

Sources are present (inline or as a list), but the mapping between specific claims and specific sources is not always explici

Assessment Tags (Risk Class 4)

- Observed across: Comparative evaluation; Capability boundary inquiry; Purchase decision support; Drawback-focused inquiries
- Decision impact potential: Medium

Risk Class 5: Confidence Level Not Calibrated to Evidence Strength

Responses often use advisory or prescriptive language even when underlying evidence is mixed, speculative, or anecdotal.

Assessment Tags (Risk Class 5)

- Observed across: Comparative evaluation; Capability boundary inquiry; Purchase decision support; Production readiness guidance
- Decision impact potential: High

Risk Class 6: Mitigation Guidance Without Prioritization Criteria

Mitigation strategies and "best practices" are suggested, but criteria for prioritizing which constraints are most critical or relevant are not provided.

Assessment Tags (Risk Class 6)

- Observed across: Capability boundary inquiry; Production readiness guidance
- Decision impact potential: Medium

Severity Summary

““text

Risk Class	Decision Impact (L/M/H)	Likelihood in observed journeys (L/M/H)	User Exposure (L/M/H)	Rationale
Mixed Source Authority Without Signaling	H	H	H	Mixed-evidence sourcing approaches
Uncertainty Not Elevated as a Decision Factor	H	M	H	Uncertainty is often indirect or low priority
Alternatives Not Explicitly Contrasted	M	M	M	Trade-offs appear, but counterfactuals are not fully explored
Claim-to-Source Traceability Friction	M	H	M	Sources are present but claims are not well-supported
Confidence Level Not Calibrated to Evidence Strength	H	M	M	Advisory framing can be used to mitigate this risk
Mitigation Guidance Without Prioritization Criteria	M	M	M	Mitigation guidance appears but lacks prioritization
““				

Risk Mapping by Decision Context

Decision contexts considered:

- Purchase & Pricing Decisions
- Career / HR Decisions
- Technical Implementation Decisions
- Security / Compliance Decisions
- Health-like / High-trust Decisions
- Low-stakes informational use

Matrix legend: H = High, M = Medium, L = Low

Risk Class	Purchase & Pricing	Career / HR	Technical Impl.	Security / Compliance	Health-like
Mixed Source Authority Without Signaling	H	M	M	M	H
Uncertainty Not Elevated as a Decision Factor	H	M	M	M	H
Alternatives Not Explicitly Contrasted	M	M	M	M	M
Claim-to-Source Traceability Friction	M	M	M	M	M
Confidence Level Not Calibrated to Evidence Strength	H	M	M	M	H
Mitigation Guidance Without Prioritization Criteria	L	L	H	H	M

Decision quality implications: The observed patterns can increase verification effort because users must interpret source and context information to determine the true risk level.

Release Risk Snapshot

Overall Decision Risk Level: Medium

Highest-risk contexts: Purchase & Pricing Decisions; Health-like / High-trust Decisions; Technical Implementation Decisions

Primary drivers: Mixed Source Authority Without Signaling; Uncertainty Not Elevated as a Decision Factor; Claim-to-Source

Verification friction level: Mediumâ High

Readiness posture: Suitable for low-stakes use; decision-critical use requires stronger signaling

Decision Impact Overview

Together, these patterns may reduce decision quality by inflating perceived certainty and obscuring the strength of supporting evidence.

Neutral Release Statement

This assessment does not assert system failure or answer inaccuracy; it documents observed patterns in how information is presented.