Cross-Platform AI-Visibility Validation

Profound vs. Peec AI vs. Otterly: what agreement across three tools actually proves

Néstor G. Trillo | SEO / AEO / GEO Strategist

Cross-Platform AI-Visibility Validation

Profound vs. Peec AI vs. Otterly: what agreement across three tools actually proves

Néstor G. Trillo | SEO / AEO / GEO Strategist

> *A single dashboard percentage is one vendor's opinion. Three tools, tested separately, agreeing anyway… that's evidence. This note covers two things: where Profound, Peec AI, and Otterly converge. And what each one is actually built to do.*

The Convergence Finding

Case Study 2 (Bank of Hawaii, independent research) ran the same fixed 5-question set across all three tools, each on its own pass. Checking and Savings converged as BOH's strongest topics. Mortgage converged as the clear weakest overall topic across the full metric set in all three tools, not just one headline number.

Topic	Profound	Peec AI	Otterly	Verdict
Checking	93.3% Visibility (#1 rank)	100% Visibility / 41% SoV	100% Brand Coverage	Strongest topic, agrees all three
Savings	73.3% Visibility (#1 rank)	67% Visibility / 33% SoV	100% Brand / 100% Domain Coverage	Strong, agrees all three
Mortgage	26.7% Visibility (#4), 2.7% Citation Share (#10, last)	33% Visibility / 7% SoV / 0.00 citation rate	25% Brand / 0% Domain Coverage	Weakest, agrees all three, full metric set checked

Here's the near-miss worth showing. Peec's raw Mortgage visibility (33%) ties with HELOC's (33%)… taken alone, that reads as "no longer uniquely worst." It isn't. Peec's own Share of Voice (7% vs. HELOC's 30%) and citation rate (0.00 vs. 0.33) both show Mortgage clearly worse. Otterly's independent pass confirms it. Stopping at the first matching number would have been the wrong call. Checking the full metric set is what the finding actually required.

Peec's and Otterly's gap-analysis features flagged the same root-cause candidate on their own: Hawaii's state first-time-homebuyer program pages (HHFDC/DBEDT participating-lender lists) dominate the mortgage evidence chain, and BOH is absent from them. Two tools, same hypothesis, arrived at independently. Call it a strong candidate, not yet a confirmed finding.

What Each Tool Is Actually Built For

Profound: enterprise measurement platform. The richest metric surface of the three (visibility, share of voice, position, citation share, sentiment, and query fanouts). Also the highest analytical overhead. The raw export's headline "147 answers" didn't match the 75 answers actually driving the dashboard's core metrics, so that denominator had to be manually reconciled before any number from it could be trusted.

Peec AI: diagnostic workbench. It distinguishes retrieval (a source was found) from citation (explicitly referenced) from brand inclusion (the brand actually got picked). That's the sharpest tool of the three for tracing why a brand lost, not just that it lost. Its Gap Analysis feature is what surfaced the HHFDC/DBEDT hypothesis directly, pointing at the exact competitor-cited pages BOH doesn't appear on.

Otterly: lightweight monitoring, clean stakeholder snapshot. Fastest to set up. Fastest to explain. Its own math even reconciled cleanly on inspection: 13 of 20 eligible executions[1] produced its headline 65% Brand Coverage figure exactly. No reconstruction needed, unlike Profound's denominator.

1. The 20 is 5 core prompts × 4 AI services Otterly tracks per prompt, one more service than the three engines (ChatGPT, Google AI Overviews, Perplexity) this piece and Case Study 2 otherwise focus on.

Known Quirks Worth Knowing Before Trusting a Dashboard

Peec's Crawlability check reported boh.com had no accessible robots.txt. A direct fetch confirmed the file exists and is open to all major AI crawlers. Dashboard claims about a site's technical health need independent verification, not acceptance at face value.

Profound's entity name-matching is retroactive, but doesn't auto-merge historical matrix rows. A competitor tracked under two names (HawaiiUSA Federal Credit Union / HawaiiUSA FCU) kept showing as two separate rows until an asset-hierarchy merge was applied. A name match alone wasn't enough.

AI-generated query fanouts, in both Peec and Profound, occasionally invented plausible-sounding entities. One fanout string referenced "Hawaii Island Financial Federal Credit Union." Turned out not to be a real institution. Fanout-derived findings are directional, not a citable fact on their own.

None of this means the tools are unreliable. It means a dashboard summary is a starting point for investigation, not a citable finding on its own.

Which Tool Earns Its Keep, and When

Peec's retrieval-vs-citation distinction makes it the strongest day-to-day diagnostic tool. It gets you from "this topic underperformed" to "here is the exact source gap" faster than the other two. Otterly is the fastest, cleanest complement for stakeholder-facing snapshots and a second, independent cross-check. Profound earns its complexity once a program has real scale (multiple teams, formal governance, and API/data-pipeline needs) to use its full measurement surface. Skip that scale, and it's the tool most likely to hand you a confidently wrong number if nobody checks the denominator first.

The value isn't picking a favorite tool. It's knowing precisely what each one is (and isn't) telling you, and using more than one when a finding actually matters.