

The NZ AI Answers Project

What AI really tells New Zealanders

A field study of real people, real accounts and real commercial answers

76

Participants

888

AI answers

3

AI engines

July

From 15-28 July
2026



A note from The Optimisers and Rand Fishkin

Search Everywhere Optimisation is the new SEO. 76 participants used real accounts to ask the same commercial questions of ChatGPT, Google AI Mode and Microsoft Copilot, then repeated the prompts in the ChatGPT Temporary condition.

They returned 888 submitted answers. The engines split on the lead provider in 134 of 211 matched cases. The same person received at least one different ChatGPT lead in 50 of 72 comparable answers.

Everything published here is anonymised, aggregated and tied to an explicit denominator.

RAND FISHKIN

Co-Founder of SparkToro and Moz



"I think this is wonderful work. I'm thrilled to see other people running these analyses.

Your findings speak to a fear I've had since January, which is that AI mention tracking, even when done at scale, with the right number of runs and metrics, fails to capture the complexity of the statistical lottery at play, including personalization and non-recommended mentions.

I fear many marketers and execs are being misled by their AI tracking tools. Hopefully studies like yours will help.

At least a few folks wake up."

The Study at a glance



76 unique structured
survey participants

76

Submitted AI
answers

888

Cross-engine splits: 134
of 211 matched cases

64%

People with ≥ 1 lead change
50 of 72

69%

Accounting leads
naming Xero

221/221

Core Idea

AI visibility is a distribution across engines, people, account conditions, prompts and time. A single clean run is a useful benchmark, not a complete description of what customers received

Contents

05 What AI told New Zealanders

06 The engine lottery

07 KiwiSaver

08 Power

09 Accounting

10 Mention versus lead

11 Banks visible, never leading

12 The challenger inversion

13 Ownership changes the view

14 Retired brands persist

15 Normal versus Temporary

16 The wider roster changes

17 The source ecosystem

18 Controlled versus experienced

19 Sample stability

20 What businesses can do

21 What consumers can do

22 How the study worked

25 Cleaning and denominators

27 References and citation

01

What AI told New Zealanders

The top answer doesn't reflect real market share and it changes depending on which AI engine you ask.

The engine changes the winner

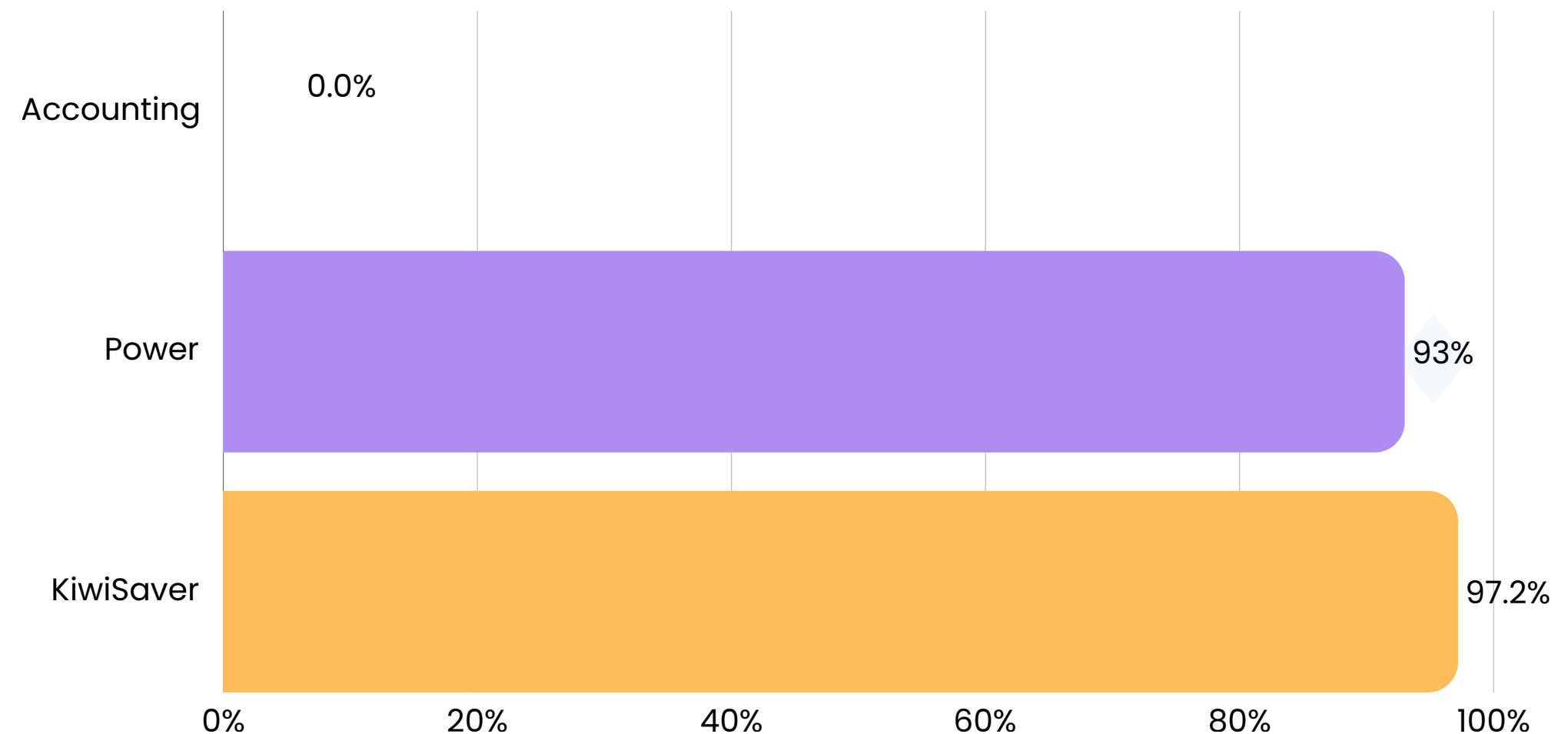
Same person, same prompt: AI engines produced different lead providers depending on the category.

63.5%

134 of 211 outputs
split on the lead
provider

The split varies by category. Power and KiwiSaver frequently produced different lead providers across AI engines, while Accounting remained consistent.

Variation in the first-named provider across categories



*Cross-engine comparisons use matched cases only: the same participant, answering the same category, with an analysable lead from all three engines. This produced 71 matched KiwiSaver cases, 70 power cases and 70 accounting cases: 211 in total.

KiwiSaver has three engines and three leaders

74%

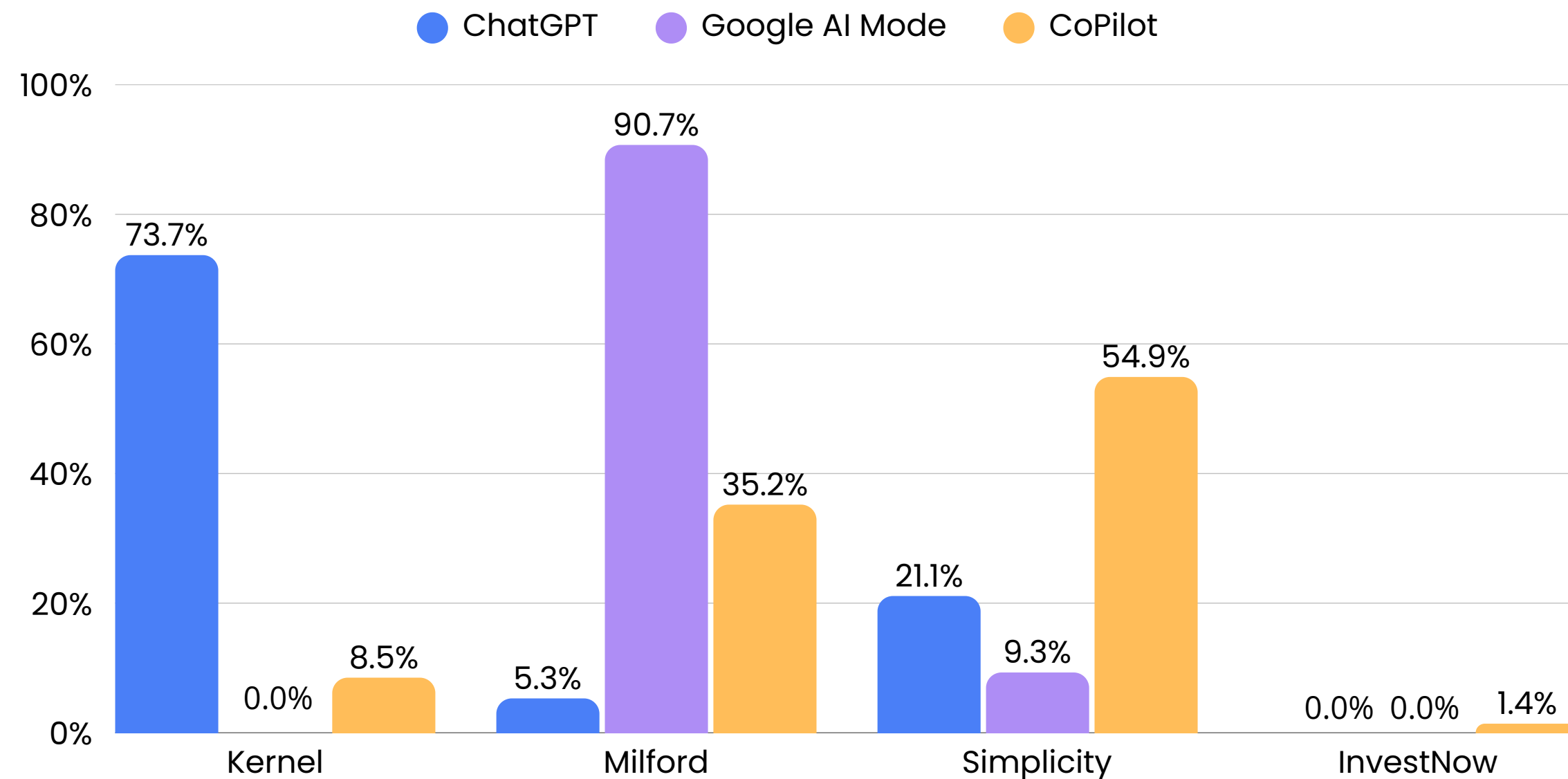
56 of 76 ChatGPT
outputs first-named
Kernel

91%

68 of 75 Google AI
Mode outputs
first-named Milford

55%

39 of 71 Copilot
outputs first-named
Simplicity



*Why the sample sizes differ: results are calculated from available, category-correct responses that named an identifiable first provider. All 76 participants produced an analysable ChatGPT KiwiSaver response. Google AI Mode had one response that named no provider (n=75). Copilot had four unavailable responses and one response that named no provider (n=71).

Power winners split sharply by engine

46%

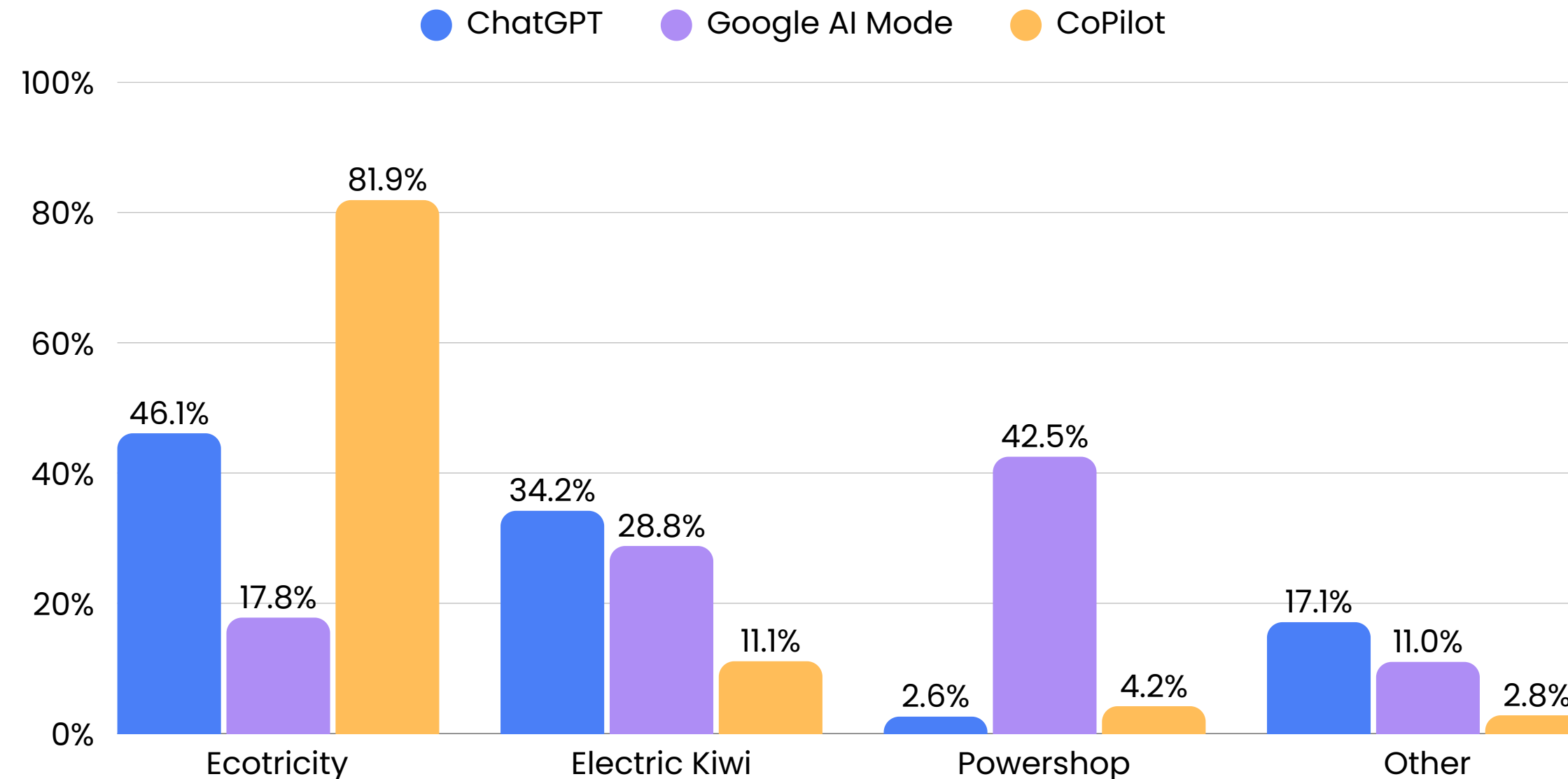
35 of 76 ChatGPT
outputs named
Ecotricity

43%

31 of 73 Google AI
Mode outputs named
Powershop

82%

59 of 72 Copilot
outputs named
Ecotricity



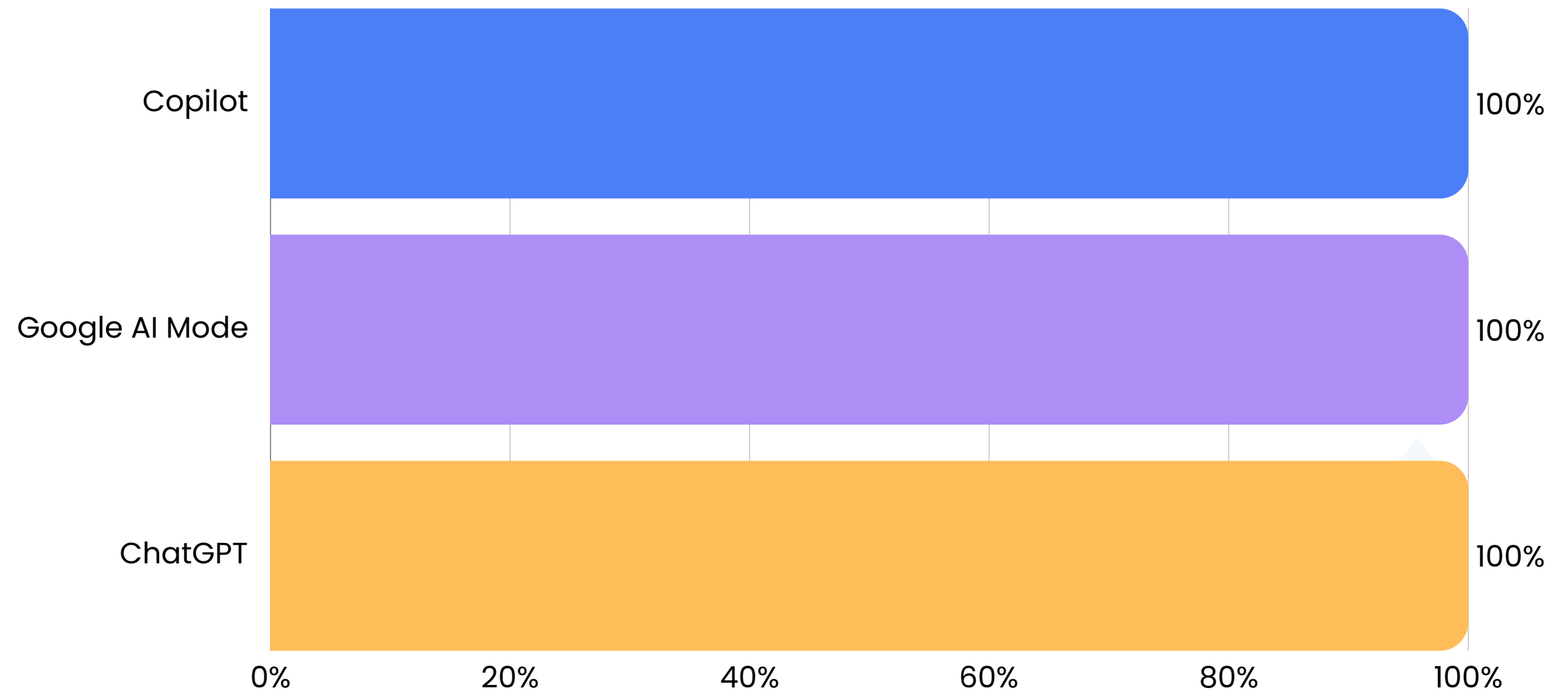
*Power denominators differ by engine after cleaning. ChatGPT had 76 analysable responses. Google AI Mode had one wrong-category paste and two responses with no identifiable lead (n=73). Copilot had four unavailable responses (n=72).

Accounting converges on Xero

221/221

accounting
outputs named
Xero

**This category behaved
unlike Kiwisaver and Power
as every output named the
same provider.**



*Accounting denominators reflect available, category-correct answers. ChatGPT had 76 analysable responses. Google AI Mode had two wrong-category pastes (n=74). Copilot had four unavailable responses and one wrong-category paste (n=71). Every one of the resulting 221 analysable leads named Xero first.

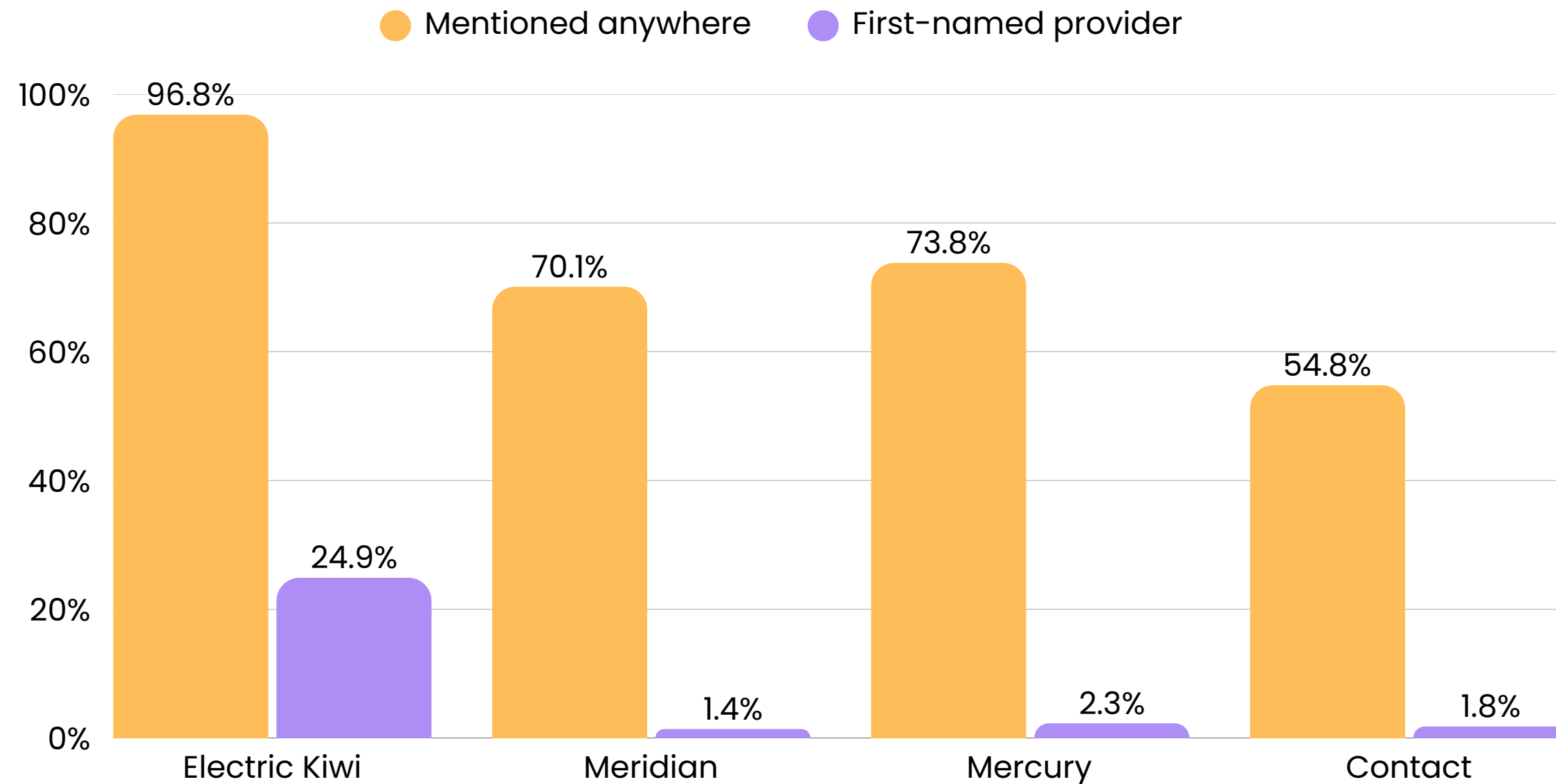
Presence and lead position tell different stories

97%

214 of 221 outputs
mentioned Electric
Kiwi

25%

55 of 221 outputs
produced Electric
Kiwi as first-named
provider



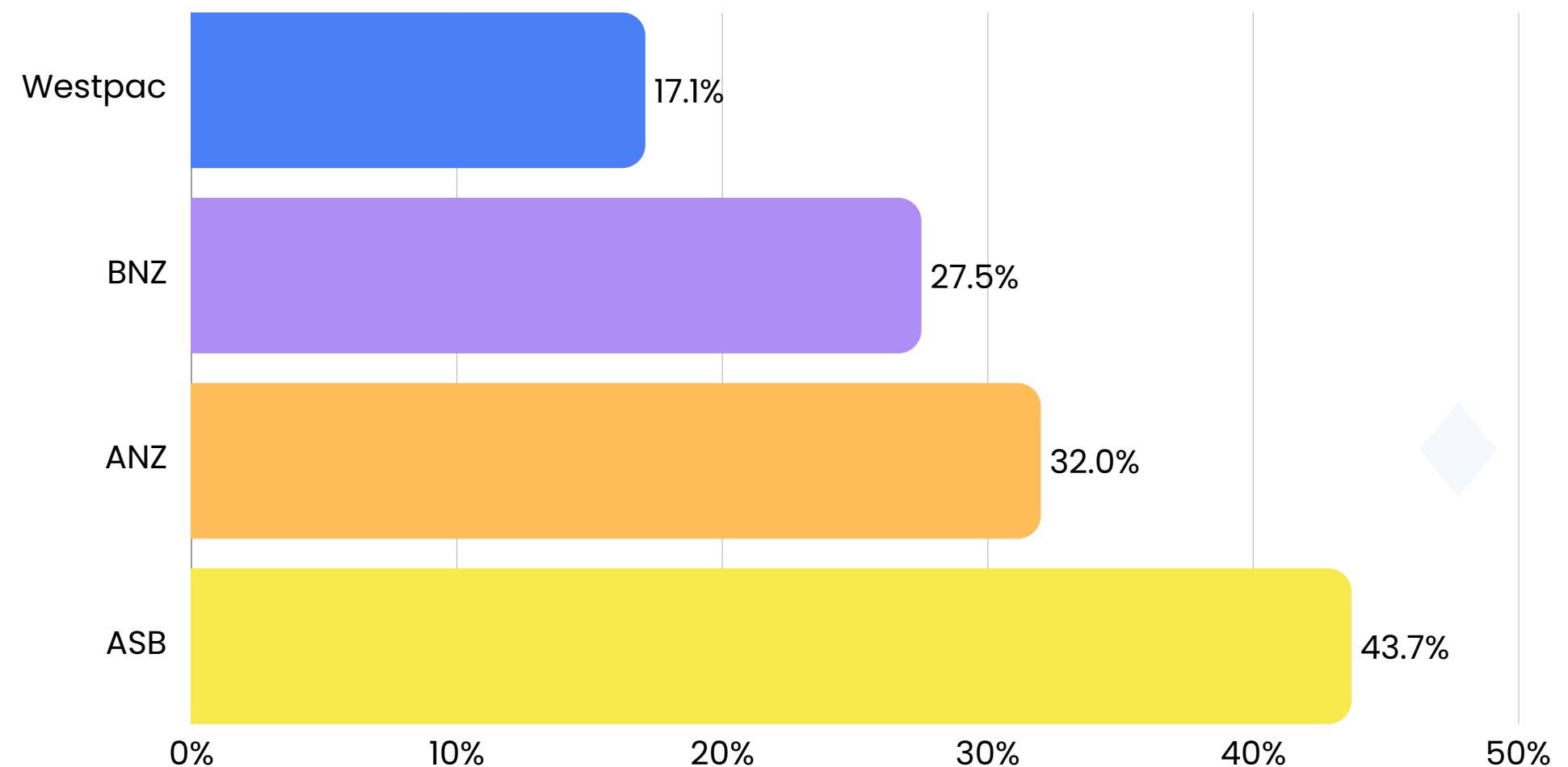
The big four banks appear but never lead

0/222

No first-named
provider positions
went to ASB, ANZ,
BNZ or Westpac

Visibility without salience: ASB appeared
in 97 outputs, yet the lead-position
distribution is held by Kernel, Milford,
Simplicity and InvestNow.

Share of outputs mentioning each bank



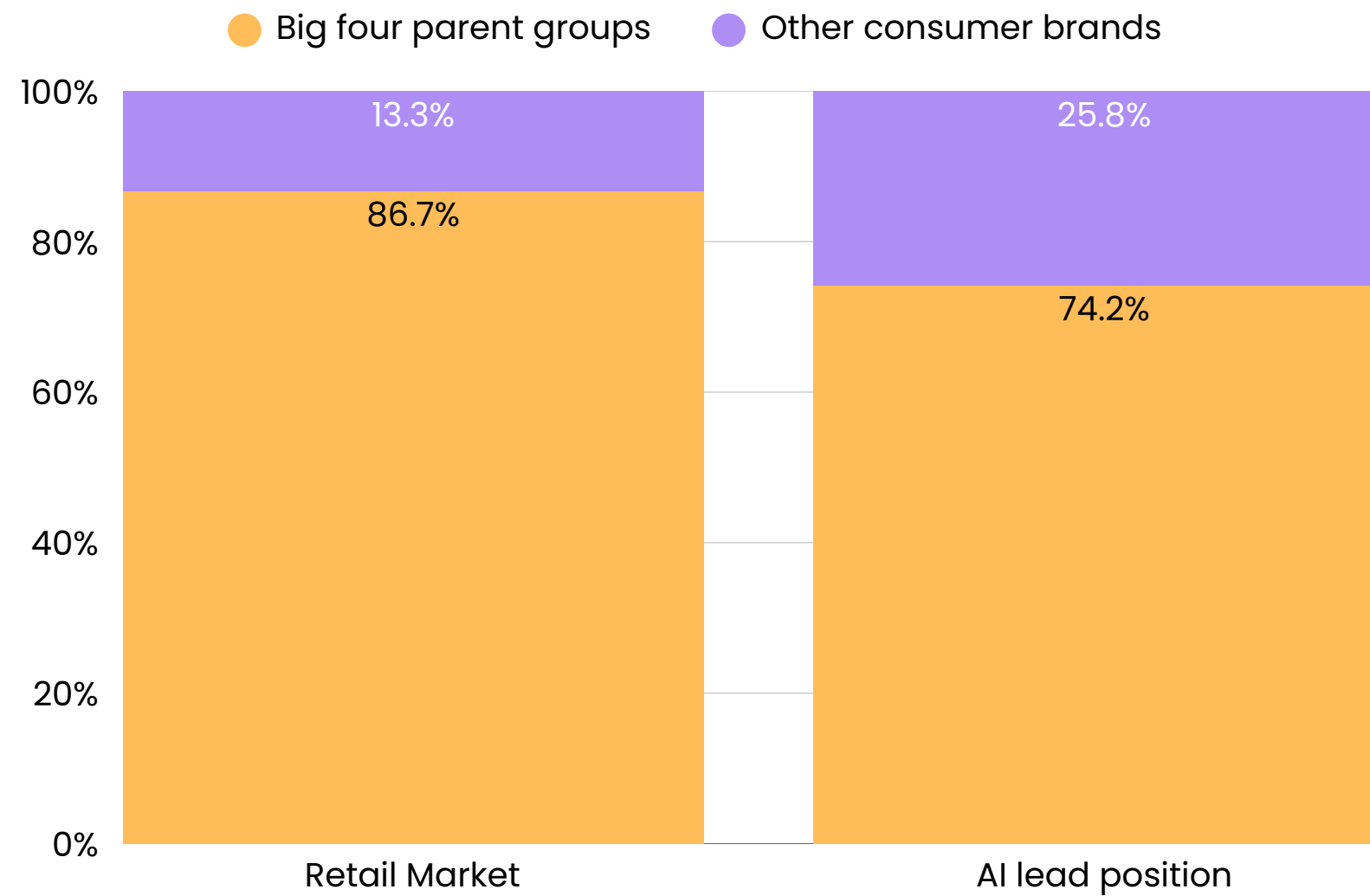
AI leans slightly away from the big four parent groups

86.7%

of active retail ICPs sit with Genesis, Meridian, Mercury and Contact

74.2%

of AI lead positions name those four parent groups — 164 of 221



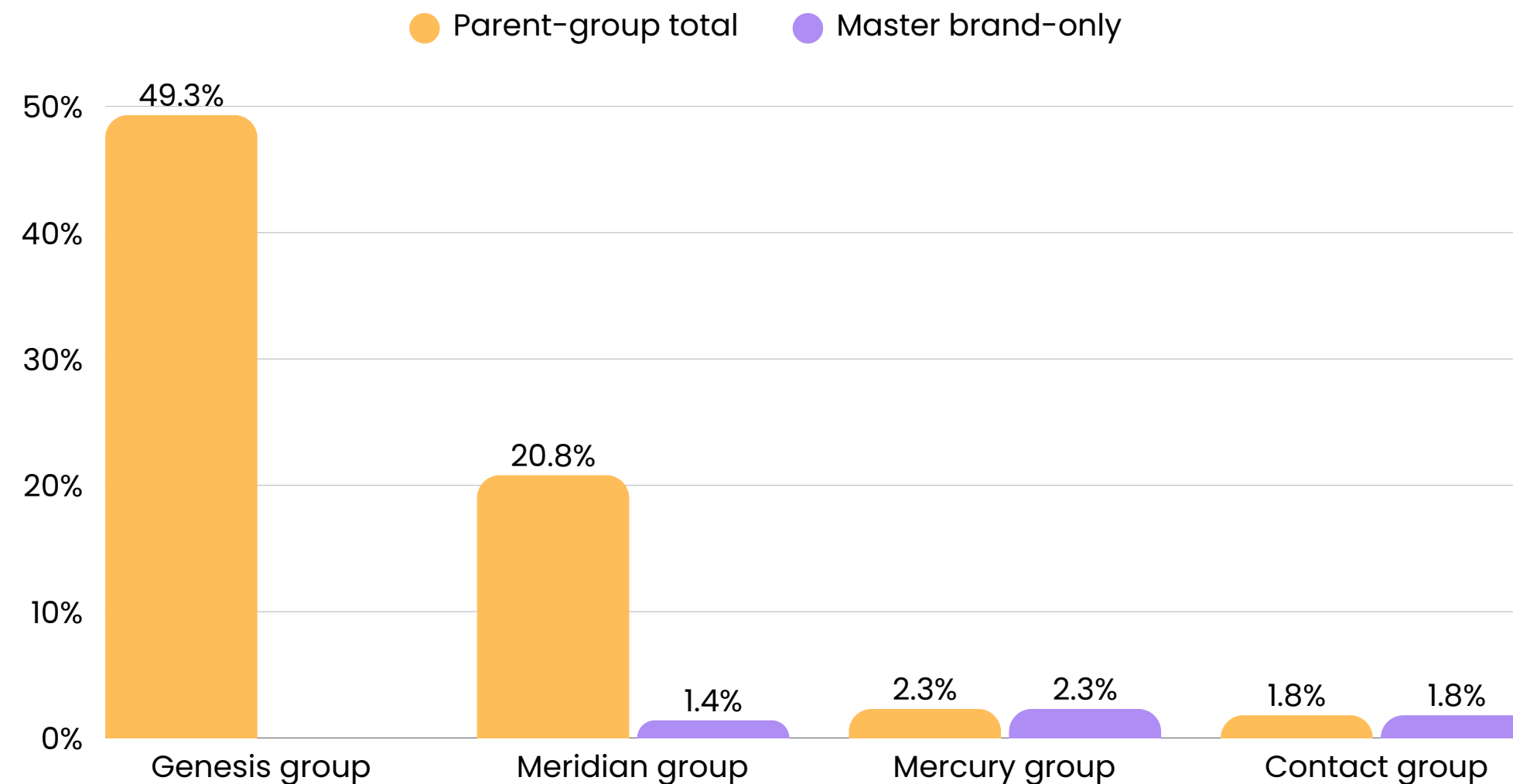
Ownership changes the interpretation

49%

Genesis group
Ecotricity 107 +
Frank 2 + Genesis
master brand 0

21%

Meridian group
Powershop 36 +
Flick 7 + Meridian 3



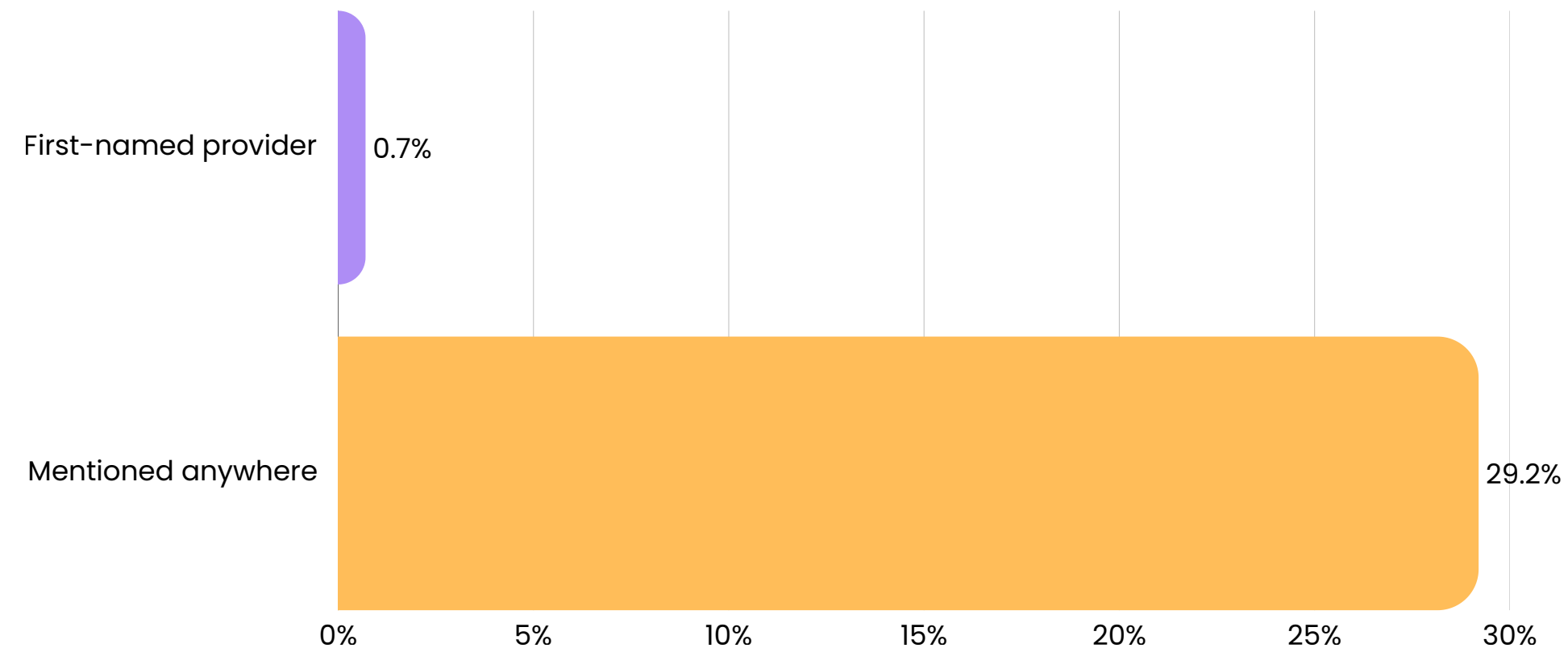
Retired brands persist in AI answers

86/295

outputs still mentioned
Frank Energy across
normal and Temporary
Conditions

Normal engines alone: 77 of 221
answers mention Frank Energy
(34.8%)

Outputs Mentioning Frank Energy



Account condition changes the answer

69%

50 of 72 people
received at least one
different lead provider

28%

61 of 215 paired category
runs changed the lead
provider

86%

184 of 215 paired
category runs changed
the wider brand roster

Condition, not a pure memory experiment

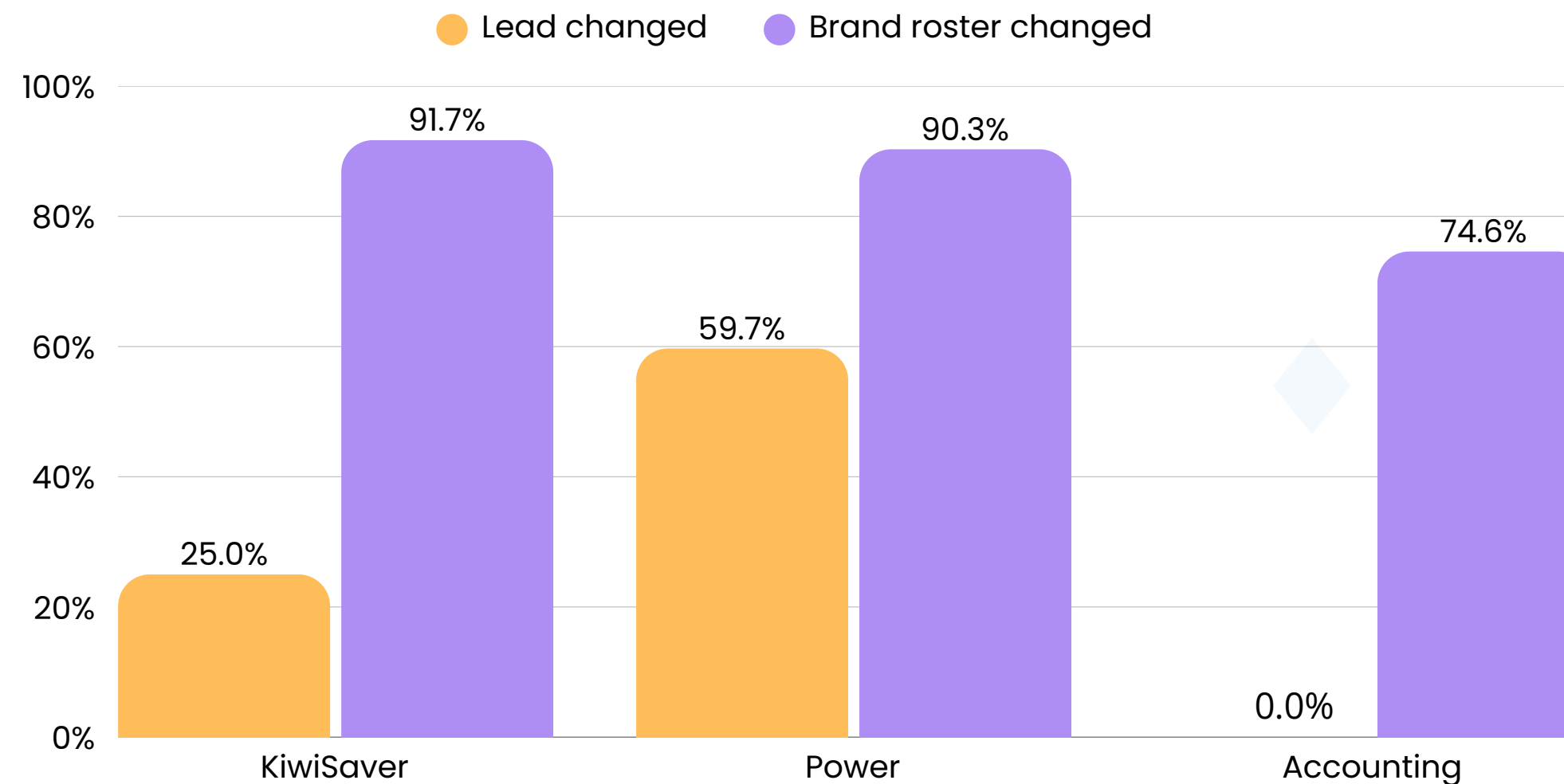
Temporary Chat removes access to saved memories for personalisation. Other account, custom-instruction, system and run-to-run effects may still operate, so the report treats this as a condition comparison.

*Normal-Temporary comparisons include only cases where the same participant produced an analysable ChatGPT lead in both conditions for the same category. This yielded 72 KiwiSaver pairs, 72 power pairs and 71 accounting pairs: 215 pairs in total.

The wider shortlist changes more often than the lead

Accounting keeps Xero in front while three quarters of paired answers change the brands surrounding it.

Normal vs temporary chats (ChatGPT)



*Normal-Temporary comparisons include only cases where the same participant produced an analysable ChatGPT lead in both conditions for the same category. This yielded 72 KiwiSaver pairs, 72 power pairs and 71 accounting pairs: 215 pairs in total.

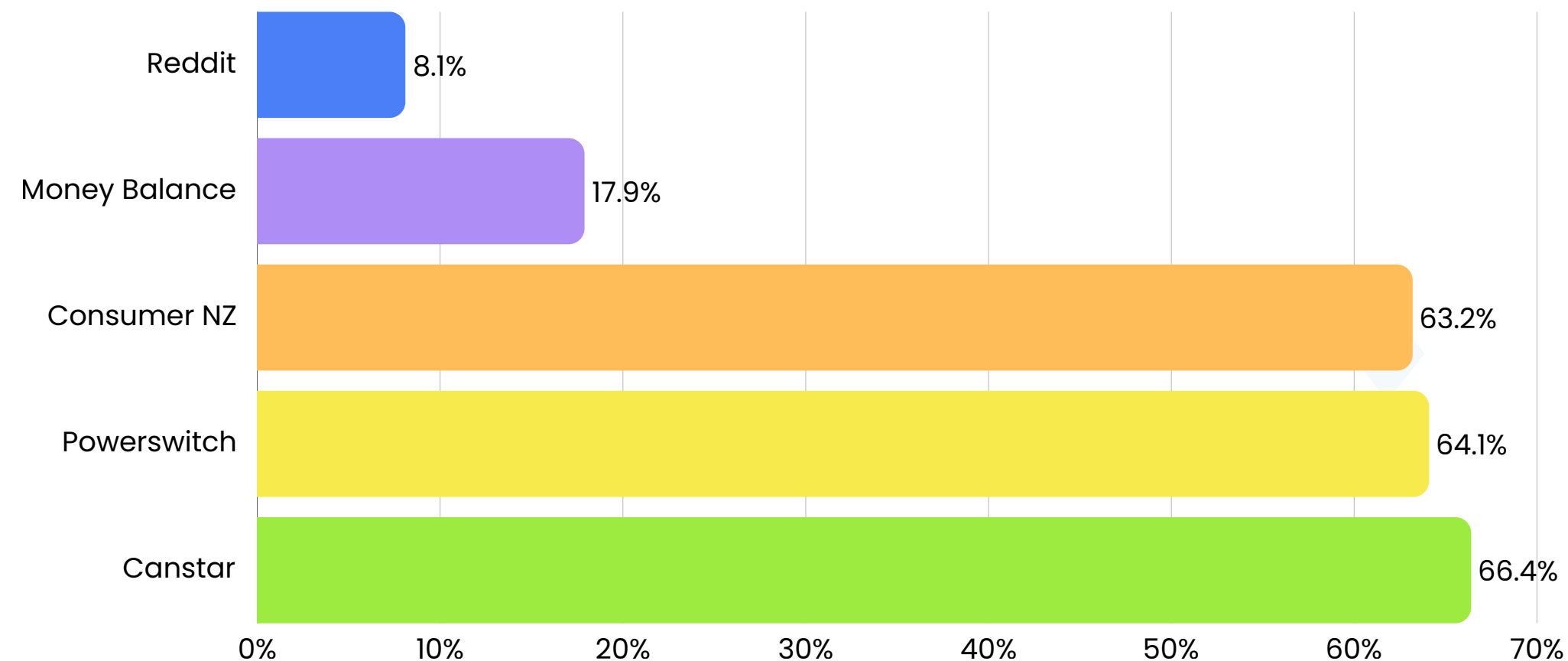
A small source ecosystem recurs across power answers

12/18

Clean-layer outputs
named Canstar

Source recurrence is an association. The study does not identify which source caused an answer.

Sources referred to in AI answers



Single-run benchmarks did not always match the human plurality

Human plurality from real-account normal-engine answers versus one Peec and one Scrunch run per engine/category.

Category / engine	Human plurality	Peec	Scrunch
Kiwisaver - ChatGPT	Kernel - 56/76	Kernel	Milford
Kiwisaver - Google	Milford - 68/75	Milford	Milford
Kiwisaver - Copilot	Simplicity - 39/71	Simplicity	Milford
Power - ChatGPT	Ecotricity - 35/76	No provider	Ecotricity
Power - Google	Powershop - 31/73	Energy Online	Electric Kiwi
Power - Copilot	Ecotricity - 59/72	Ecotricity	Ecotricity
Accounting - ChatGPT	Xero - 76/76	Xero	Xero
Accounting - Google	Xero - 74/74	Xero	Xero
Accounting - Copilot	Xero - 71/71	Xero	Xero

7/9

Peec cells matched the human plurality

6/9

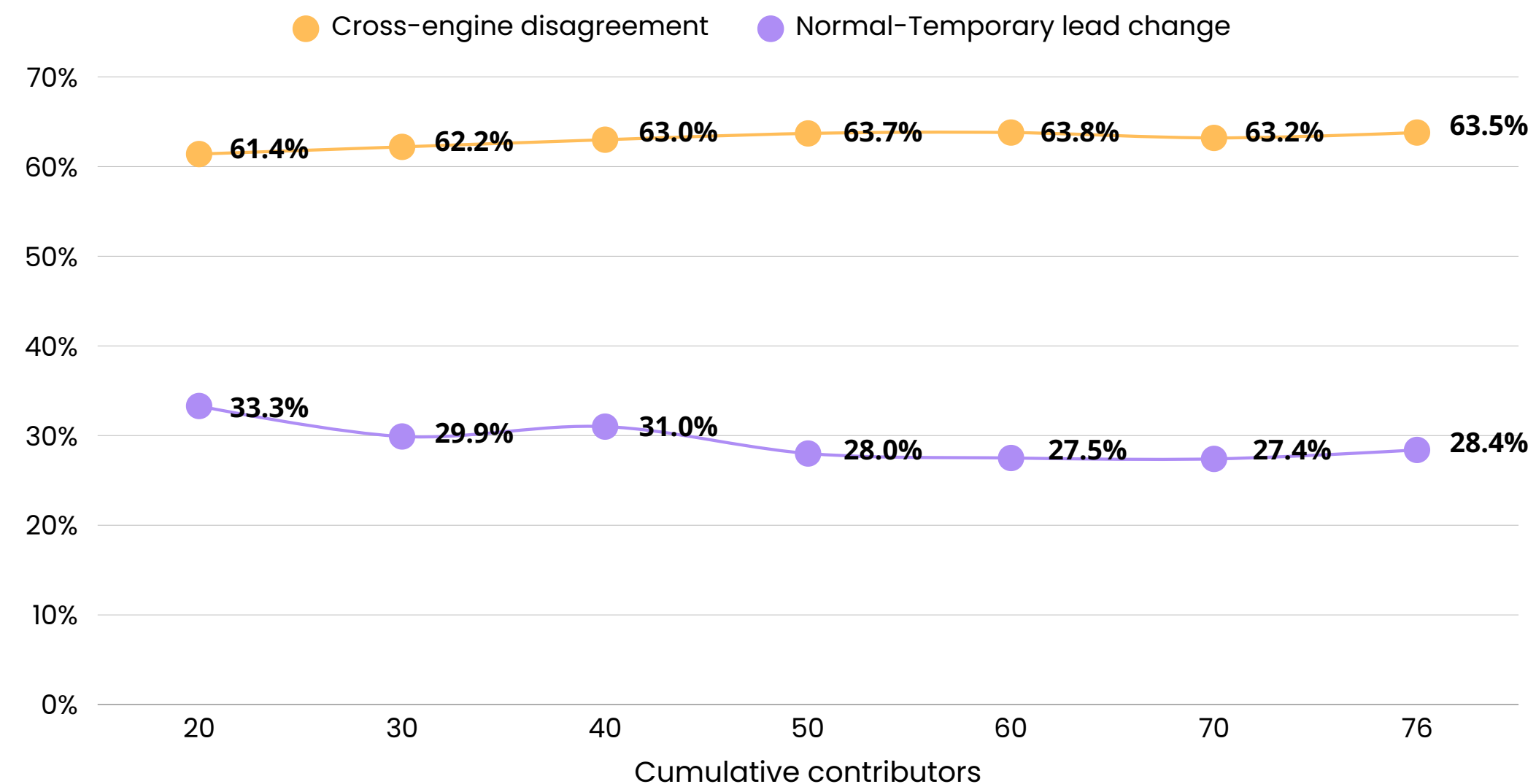
Scrunch cells matched the human plurality

*Denominator (n) = available, category-correct responses containing an identifiable first-named provider. Unavailable responses, wrong-category pastes and answers that named no provider are excluded. The applicable n is shown for every engine and comparison.

The headline rates stabilised before the final respondent

Cumulative structured survey sample in collection order; valid denominators vary at each checkpoint;

From n=50 onward, both cumulative headline rates remained within one percentage point of their final values



What this means for businesses

The measurement design follows directly from the observed distribution

01

Measure leads and mentions

A brand can appear widely without occupying the lead position. Report both, with counts and the correct denominator.

04

Audit the recurring source set

Canstar, Powerswitch, Consumer NZ and a short list of specialist sources recur across answer text.

02

Plan by engine

KiwiSaver and power produced different champions on ChatGPT, Google AI Mode and Copilot.

05

Build rebrands into answer-layer work

Frank's persistence shows that commercial retirement and answer-layer retirement can move on different timelines.

03

Separate brand from owner

A subsidiary can dominate the answer while its parent master brand barely appears in front.

Use clean tracking as a benchmark

A platform run is one controlled layer. Add repeated real-account testing when the customer experience matters.

What this means for consumers

An AI answer is one observed outcome, not a market verdict.

01

Ask another engine

For KiwiSaver and power, a second engine often produces a different first-named provider.

03

Verify with independent sources

Use comparison services, regulators and provider information before making a financial or utility decision.

02

Check the provider still exists

The answer layer can retain brands after their customer proposition changes.

04

Treat confidence as presentation

A polished answer can hide the variation visible across engines, people, conditions and time.

The study measures recommendation salience. It does not decide which provider is best for an individual.

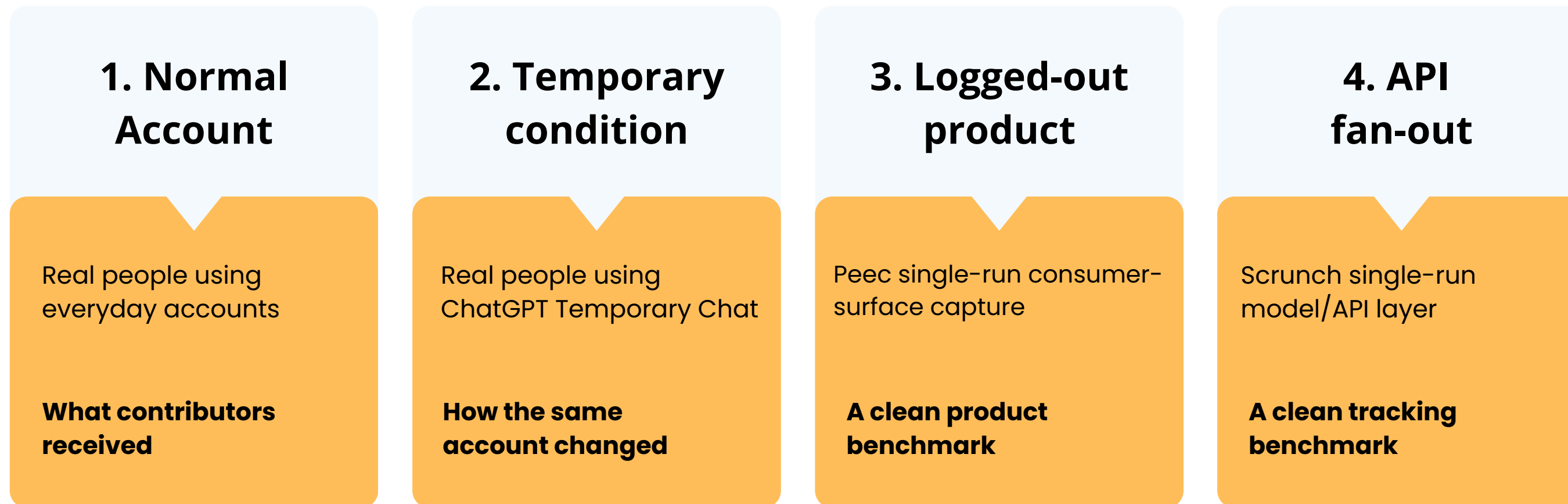
02

How the study worked

Four measurement layers, fixed prompts, explicit exclusions and a denominator for every claim.

Four layers measure four different conditions

The study combines experienced visibility with controlled benchmark layers.



These layers complement one another; none is a complete population estimate on its own.

The protocol held the questions fixed

PROTOCOL

01

Start a new chat

Use the exact wording. No follow-ups.
Paste the complete answer into the
collection form.

02

Run three normal engines

ChatGPT, Google AI Mode and Microsoft
Copilot using the contributor's everyday
access condition.

03

Repeat in Temporary Chat

Run the same three prompts in ChatGPT
Temporary Chat. Twelve potential
answers per structured respondent.

THE THREE FIXED PROMPTS

KIWISAVER

What is the best KiwiSaver provider in New Zealand?

POWER

Who is the best power company in New Zealand?

ACCOUNTING

What is the best accounting software for a small
business in New Zealand?

Every result carries its own denominator

The answer universe narrows as availability, category fit and analysability rules are applied.

Universe	Count
Unique participants	76
Designed human cells	912
Unavailable positions	24
Usable human outputs	888
Wrong-category pastes	5
Category-appropriate outputs	883
Analysable lead positions	879
Matched three-engine cases	211
Normal-Temporary pairs	215
Clean-layer outputs	18

*Denominator (n) = available, category-correct responses containing an identifiable first-named provider. Unavailable responses, wrong-category pastes and answers that named no provider are excluded. The applicable n is shown for every engine and comparison.

Limits define the claim not the value of the study

The design measures reproducible salience across a specific set of prompts, accounts, engines and dates.

LIMITATIONS

1. Convenience Sample

Recruitment was network- and referral-based. The study does not estimate prevalence for all New Zealanders.

2. Three fixed prompts

Prompt wording defines the commercial slice. Other tasks and phrasings can produce different distributions.

3. One 14-day window

Models, sources and interfaces change. The tool layers are single-run snapshots.

4. Lead is a proxy

First-named provider is a salience proxy, not recommendation strength or suitability.

NEXT-WAVE RESEARCH

1. Repeat the field wave

Track whether winners and disagreement rates move as engines update.

2. Broaden prompt families

Test task, phrasing and journey-stage sensitivity within each category.

3. Recruit stratified samples

Separate account, subscription, occupation and demographic conditions with sufficient power.

4. Audit source pathways

Use controlled tests before claiming publisher causality.

References and cite-this-study

Primary evidence and selected context used in the published report.

1. The Optimisers (2026). The NZ AI Answers Project workbooks and audit files.
2. Fishkin, R. (2026). AIs are highly inconsistent when recommending brands. SparkToro.
3. Zou, J. (2026). Is once a day enough? Profound.
4. McCain, D. (2026). Variability of Google models: Gemini vs AIO vs AI Mode. Profound.
5. Clark, N. (2026). How much data do you need to measure AI visibility with confidence? Gumshoe.
6. Iannelli, M. (2026). The prompt-to-purchase pipeline. Scrunch
7. Similarweb (2026). The 2026 generative AI landscape: the evolution of AI search.
8. Aggarwal et al. (2024). GEO: Generative Engine Optimization. KDD.
9. Martinez, O. (2026). Optimizing visibility in generative engines: a critical survey. arXiv:2607.14035.
10. OpenAI (2026). Temporary Chat FAQ.
11. Electricity Authority EMI (2026). Retail market share, 30 June 2026.
12. Genesis Energy (2025–26). Retail streamlining and interim reporting.
13. Frank Energy (2025). Who is Frank Energy?
14. Meridian Energy (2025–26). Powershop ownership and Flick transaction material.

Cite this study

The Optimisers (2026). The NZ AI Answers Project: what AI really tells New Zealanders. Auckland, August 2026. theoptimisers.com

AI visibility is a distribution.

Measure the distribution.
Then decide where to lead.

Always lead, never follow.

Stop being invisible in AI Search:



Contact The Optimisers

✉ info@theoptimisers.com



The Optimisers
35 Kawana Street, Northcote
Auckland, 0627
(+64) 9 950 3888