LLM SHELF CATEGORY REPORT

Liderzy psiej miski

Którym markom karm najbardziej ufa sztuczna inteligencja?

1. Ogólny wybór
2. Etap życia
3. Zdrowie
4. Format żywienia
5. Skład i obietnica
6. Niejadek
7. Cena do jakości

LS LLM Shelf

Polska • karma dla psów • lipiec 2026

RAPORT LLM SHELF

Jedno pytanie, trzy różne miski

Co AI rekomenduje właścicielom psów w Polsce i dlaczego każdy system wybiera inaczej?

AI poleca chętnie. Zgadza się rzadko.

Brit był najczęściej rekomendowaną marką, Royal Canin najczęściej zajmował pierwsze miejsce, a wszystkie trzy badane systemy wskazały tę samą markę jako numer jeden tylko w 3 z 80 porównywalnych przypadków.

Polska • 45 scenariuszy zakupowych • 3 powtórzenia • 3 systemy AI • 405 testów

Badanie przeprowadzone w Polsce, w lipcu 2026 r.

LLM Shelf

Spis treści

1. Najważniejsze wyniki

Widoczność w AI nie ma jednego zwycięzcy. Marka może być często wymieniana, ale rzadko być pierwszym wyborem. Może wygrywać w jednym systemie i niemal znikać w innym. Może też dominować tylko w jednej konkretnej potrzebie: od diety weterynaryjnej, przez żywienie świeże, po ograniczony budżet.

34,4%

odpowiedzi rekomendowało Brit - najwyższy wynik ogólny

13,2%

odpowiedzi stawiało Royal Canin na pierwszej pozycji

3 z 80

porównywalnych przypadków miało tę samą markę na początku we wszystkich systemach

82,7%

325 z 393 odpowiedzi tekstowych zawierało konkretną rekomendację marki

60,9%

przypadków bez rekomendacji zakończyło się jej uzyskaniem po dopytaniu modelu (14 z 23 przypadków)

16

prawdopodobnych błędów gatunkowych wykrył audyt odpowiedzi

Kluczowe wnioski

1 **Nie ma jednego lidera „widoczności w AI”**

To, która marka wygrywa, zależy od tego, jak mierzymy wynik, z jakiego modelu AI korzystamy i jaką potrzebę zgłasza użytkownik. Brit jest najczęściej wymienianą marką, Royal Canin częściej zajmuje pierwszą pozycję, a poszczególne systemy - OpenAI, Gemini i Google AI Overview - budują wyraźnie różne shortlisty marek. Lider zmienia się też wraz z kontekstem: w zdrowiu dominują Royal Canin i Hill’s, w karmach świeżych Piesotto i PsiBufet, a w relacji jakości do ceny Dolina Noteci i Brit.

2 **AI zawęża pole wyboru, a nie tylko udziela ogólnych porad**

Modele najczęściej wskazują konkretne marki i produkty, tworząc krótką listę, a czasem wyłaniając jednego wyraźnego zwycięzcę. Jedno dodatkowe pytanie pozwoliło uzyskać rekomendację w 60,9% przypadków, w których początkowo model odpowiedział jedynie generalną poradą. Dalszy przebieg rozmowy może więc zmienić zestaw rekomendowanych produktów.

3 **Rekomendacje AI są niespójne i wymagają ostrożności**

Ta sama marka znalazła się na pierwszym miejscu we wszystkich trzech systemach tylko w 3 z 80 porównywalnych przypadków. Audyt wykazał też prawdopodobne błędy gatunkowe i twierdzenia wymagające weryfikacji. Odpowiedzi AI nie powinny być traktowane jako jednolity, w pełni przewidywalny kanał.

2. AI to nowa półka sklepowa

W klasycznej wyszukiwarce marka walczy o kliknięcie, a w social mediach o uwagę. **W odpowiedzi generowanej przez AI walczy o coś jeszcze ważniejszego: o wejście na krótką listę produktów, które użytkownik w ogóle bierze pod uwagę.**

Użytkownik opisuje psa, budżet i problem, a system redukuje setki opcji do kilku nazw. W tym momencie część marek po prostu wypada z procesu decyzyjnego, zanim konsument odwiedzi sklep. To właśnie dlatego rekomendacje AI zaczynają działać jak nowa półka sklepowa, tylko znacznie bardziej selektywna.

To badanie nie mierzy sprzedaży ani jakości karm. Sprawdza wcześniejszy etap:

jak systemy AI budują odpowiedź zakupową, czy wskazują markę, które marki wybierają, w jakiej kolejności je pokazują i jakimi argumentami uzasadniają rekomendację.

Widoczność to nie to samo co rekomendacja

Sama obecność marki w odpowiedzi AI nie oznacza jeszcze rekomendacji, a rekomendacja nie zawsze oznacza pierwszego miejsca. Dlatego osobno mierzymy: obecność marki, pozytywną rekomendację oraz ekspozycję na pierwszej pozycji.

Takie rozróżnienie zmienia obraz liderów: Brit był rekomendowany w 135 z 393 odpowiedzi. Royal Canin najczęściej zajmował pierwszą pozycję - 52 razy, czyli ponad dwa razy częściej niż Brit (22 razy).

! W praktyce oznacza to jedno: Brit wygrywa zasięgiem rekomendacyjnym, ale to Royal Canin częściej staje się pierwszym wyborem AI.

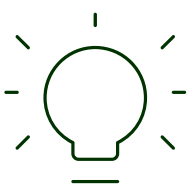
3. Jak przeprowadziliśmy badanie

Zbudowaliśmy bazę 45 pytań odzwierciedlających realne sytuacje właścicieli psów w Polsce. Każde pytanie zostało zadane trzykrotnie każdemu z trzech badanych systemów: Gemini 3.5 Flash, Google AI Overview i modelowi OpenAI GPT-5.4 mini. Łącznie przeprowadzono 405 testów.

Element	Zakres
Rynek i język	Polska; pytania i odpowiedzi w języku polskim
Zakres	45 promptów w 7 grupach potrzeb
Powtórzenia	3 odpowiedzi na każdy prompt i system
Systemy	Gemini 3.5 Flash, Google AI Overview, OpenAI GPT-5.4 mini
Łącznie	405 testów; 393 odpowiedzi tekstowe
Brak AIO	12 przypadków, w których Google AI Overview nie pojawiło się na stronie
Taksonomia	129 marek po normalizacji aliasów i nazw produktów
Jednostka pomiaru	Odpowiedź tekstowa; marka liczona maksymalnie raz w danej odpowiedzi

Siedem potrzeb, które uruchamiają rekomendację AI:

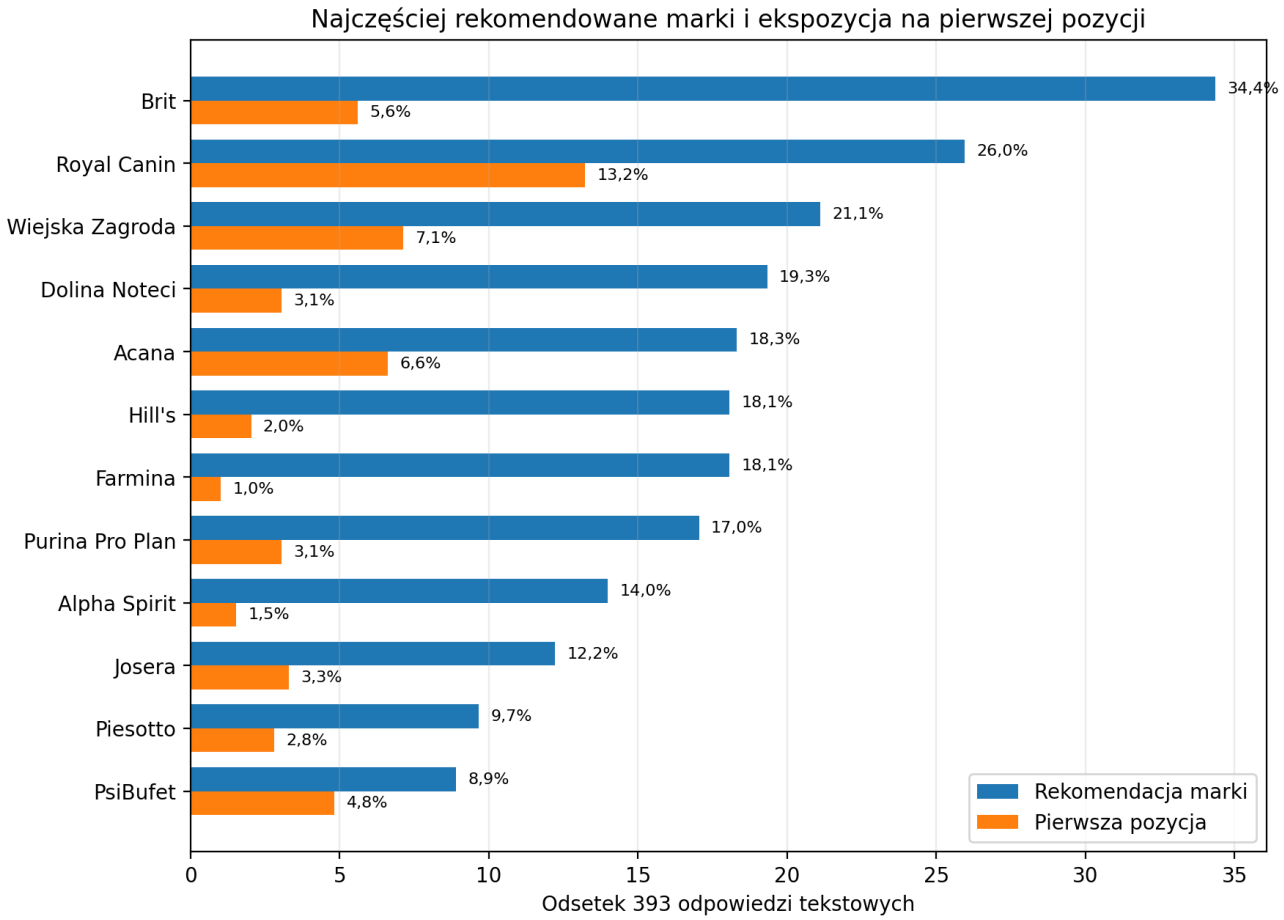
1	Ogólny wybór karmy	6 promptów
2	Profil i etap życia psa	7 promptów
3	Zdrowie i potrzeby funkcjonalne	8 promptów
4	Format i sposób żywienia	8 promptów
5	Skład i deklaracje produktowe	6 promptów
6	Niejadek i akceptacja karmy	5 promptów
7	Cena i relacja jakości do ceny	5 promptów



To mapa wyboru, nie ranking rekomendacji

Poszczególne pytania w różnym stopniu skłaniały modele do wskazania konkretnych marek i produktów. Dlatego różnice między grupami pokazują, w jakich potrzebach marka ma większą lub mniejszą szansę wejść na shortlistę AI, nie są natomiast prostym pomiarem „luki contentowej”.

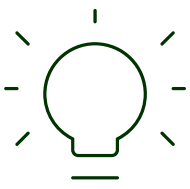
4. Kto wygrywa ranking ogólny?



Wykres 1. Recommendation rate i first-position exposure. Mianownik: 393 odpowiedzi tekstowe. Źródło: badanie LLM Shelf.

Brit był rekomendowany w 34,4% wszystkich odpowiedzi tekstowych i pojawiał się we wszystkich siedmiu grupach potrzeb. To największy zasięg rekomendacyjny w badaniu. Drugie miejsce zajął Royal Canin z wynikiem 26,0%, a kolejne Wiejska Zagroda (21,1%), Dolina Noteci (19,3%) oraz Acana (18,3%).

Sam zasięg nie opisuje jednak pełnej przewagi. Royal Canin trafiał na pierwszą pozycję w 13,2% odpowiedzi - ponad dwukrotnie częściej niż Brit (5,6%). Wiejska Zagroda osiągnęła 7,1%, Acana 6,6%, a PsiBufet 4,8%.



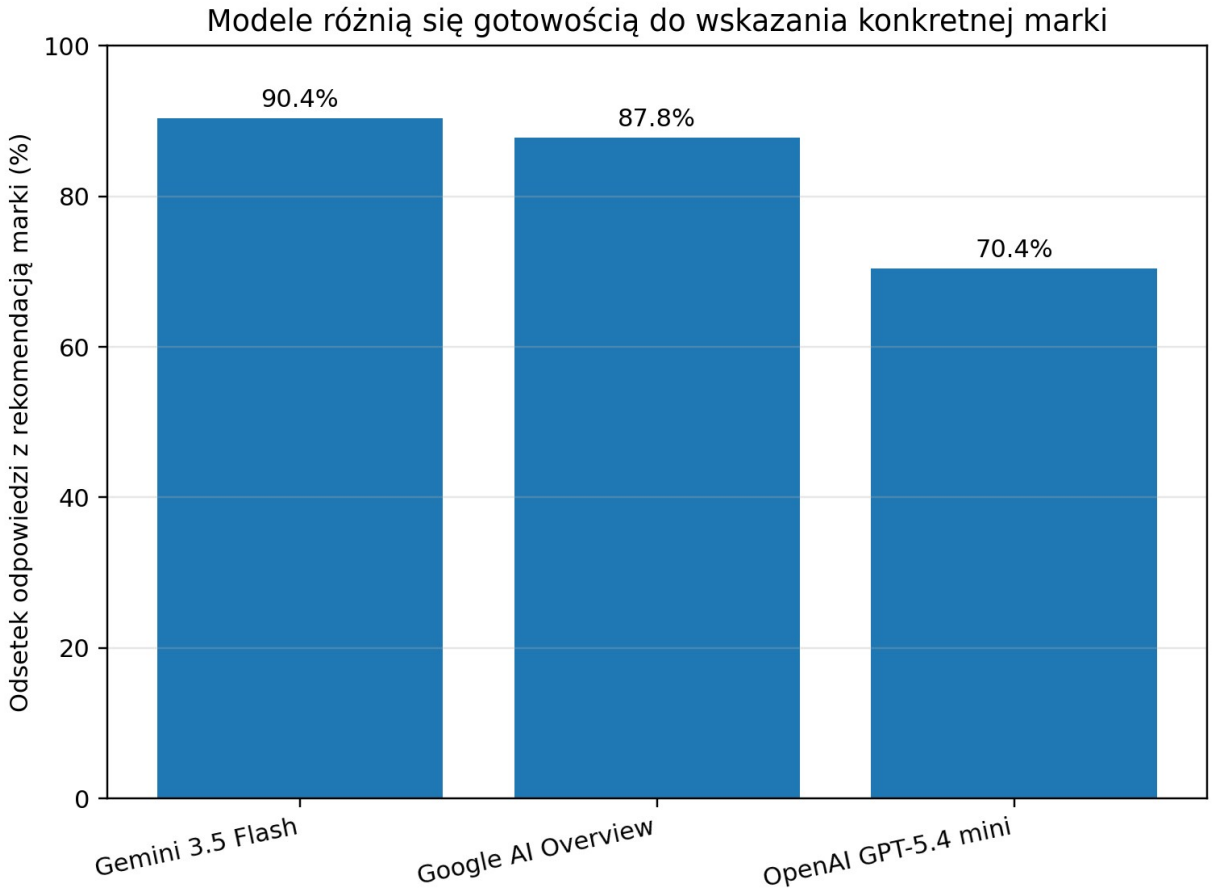
Być widocznym to jedno, być pierwszym - drugie

Brit wygrywa zasięgiem: jest rekomendowany najczęściej i pojawia się we wszystkich siedmiu grupach potrzeb. Royal Canin wygrywa pozycją: ponad dwa razy częściej zajmuje pierwsze miejsce. Dla marek oznacza to dwa odrębne cele: wejść na shortlistę AI i zostać jej pierwszym wyborem.

Marka	Odpowiedzi z rekomendacją	Recommendation rate	Pierwsza pozycja	First-position exposure	Zakres potrzeb
Brit	135	34,4%	22	5,6%	7/7
Royal Canin	102	26,0%	52	13,2%	6/7
Wiejska Zagroda	83	21,1%	28	7,1%	7/7
Dolina Noteci	76	19,3%	12	3,1%	7/7
Acana	72	18,3%	26	6,6%	6/7
Hill's	71	18,1%	8	2,0%	5/7
Farmina	71	18,1%	4	1,0%	7/7
Purina Pro Plan	67	17,0%	12	3,1%	5/7
Alpha Spirit	55	14,0%	6	1,5%	7/7
Josera	48	12,2%	13	3,3%	7/7
Piesotto	38	9,7%	11	2,8%	4/7
PsiBufet	35	8,9%	19	4,8%	4/7
Orijen	34	8,7%	9	2,3%	6/7
Purina	33	8,4%	6	1,5%	5/7
Pan Mięsko	28	7,1%	6	1,5%	5/7

5. To samo pytanie, zupełnie inna odpowiedź

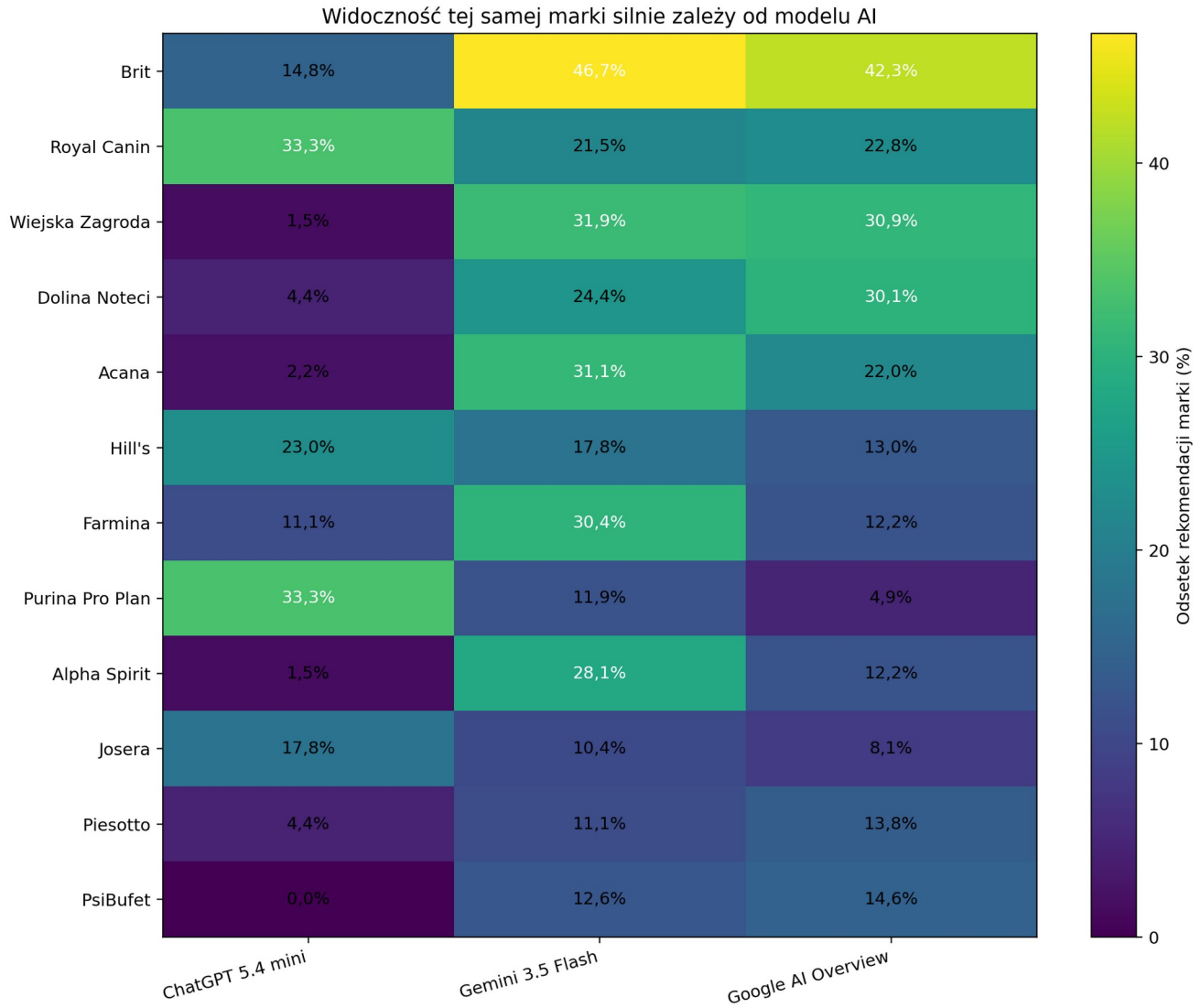
Traktowanie AI jako jednego kanału prowadzi do fałszywych wniosków. Poszczególne systemy różnią się nie tylko gotowością do wskazania konkretnej marki, lecz także tym, które marki rekomendują i jak wysoko je pozycjonują.



Wykres 2. Odsetek dostarczonych odpowiedzi zawierających rekomendację marki. Dla Google AI Overview osobno należy pamiętać o 12 przypadkach bez wygenerowanej powierzchni AIO.

Gemini i Google AI Overview wskazywały konkretną markę w niemal 9 na 10 odpowiedzi (90,4% oraz 87,8%). Model OpenAI robił to wyraźnie rzadziej - w 70,4% przypadków, częściej pozostając przy ogólnych kategoriach lub kryteriach wyboru.

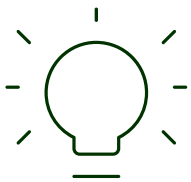
Ten sam prompt nie daje marce takiej samej szansy na pojawienie się w każdym systemie.



Wykres 3. Recommendation rate wybranych marek w poszczególnych systemach.

Różnice dotyczą nie tylko kolejności. Dla części marek zmienia się sama skala obecności. **Brit był rekomendowany w 46,7%** odpowiedzi Gemini i 42,3% odpowiedzi Google AI Overview, ale tylko w 14,8% odpowiedzi OpenAI GPT-5.4 mini. **Wiejska Zagroda osiągnęła odpowiednio 31,9%, 30,9% i 1,5%.**

Odwrotny wzorzec widać przy Purina Pro Plan: **33,3% w OpenAI GPT-5.4 mini, 11,9% w Gemini i 4,9% w Google AI Overview.** Model OpenAI GPT-5.4 mini mocniej eksponował również Royal Canin, Hill's i markę parasolową Purina.



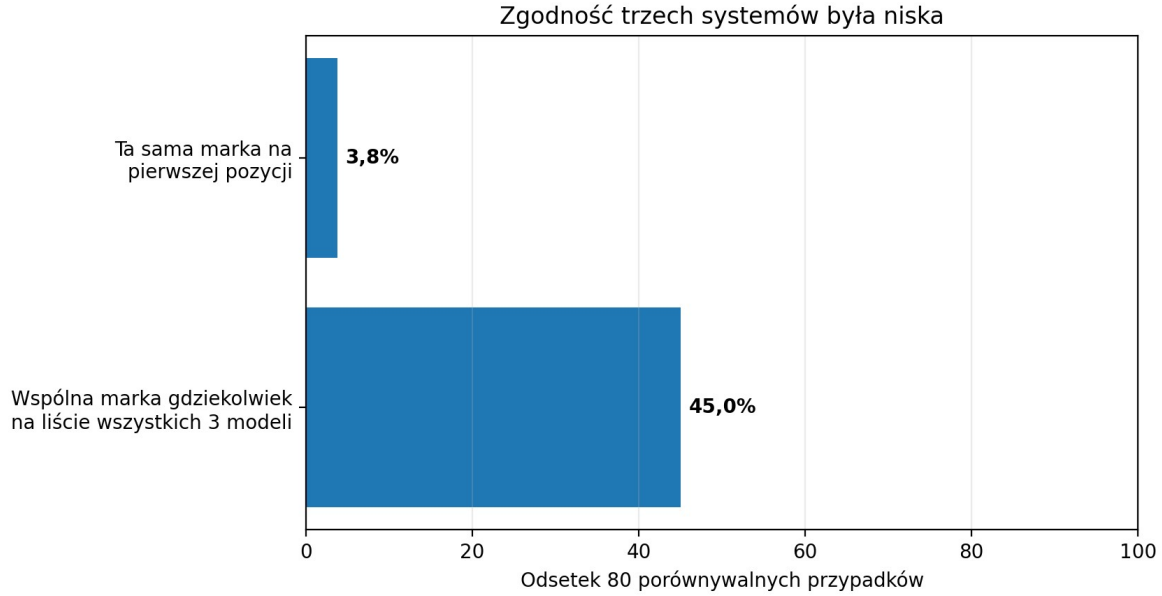
Marka może wygrywać w jednym AI i znikać w drugim

Jeden zbiorczy wynik może dawać fałszywe poczucie bezpieczeństwa. Widoczność trzeba analizować osobno dla każdego systemu, ponieważ silna pozycja w jednym modelu może maskować niemal całkowitą nieobecność w innym.

System	Miejsce	Marka	Recommendation rate	Pierwsza pozycja
OpenAI GPT-5.4 mini	1	Royal Canin	33,3%	28
OpenAI GPT-5.4 mini	2	Purina Pro Plan	33,3%	11
OpenAI GPT-5.4 mini	3	Hill's	23,0%	5
OpenAI GPT-5.4 mini	4	Purina	20,0%	6
OpenAI GPT-5.4 mini	5	Josera	17,8%	11
Gemini 3.5 Flash	1	Brit	46,7%	12
Gemini 3.5 Flash	2	Wiejska Zagroda	31,9%	15
Gemini 3.5 Flash	3	Acana	31,1%	15
Gemini 3.5 Flash	4	Farmina	30,4%	1
Gemini 3.5 Flash	5	Alpha Spirit	28,1%	5
Google AI Overview	1	Brit	42,3%	8
Google AI Overview	2	Wiejska Zagroda	30,9%	13
Google AI Overview	3	Dolina Noteci	30,1%	6
Google AI Overview	4	Royal Canin	22,8%	10
Google AI Overview	5	Acana	22,0%	11

6. Ta sama marka wygrywa tylko w 3,8% przypadków

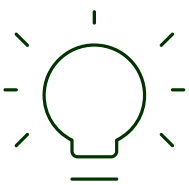
W **123 kombinacjach promptu** i powtórzenia wszystkie trzy systemy wygenerowały odpowiedź tekstową. W **80 przypadkach** każdy z nich wskazał co najmniej jedną markę - i właśnie te odpowiedzi można było porównać bezpośrednio.



Wykres 4. Zgodność rekomendacji trzech systemów w 80 porównywalnych przypadkach.

Na wszystkich trzech shortlistach pojawiła się przynajmniej jedna wspólna marka w **36 z 80 porównań, czyli w 45,0% przypadków.** Pełna zgodność na pierwszym miejscu była jednak niemal wyjątkowa: ta sama marka wygrała we wszystkich trzech systemach zaledwie **3 razy - w 3,8% porównań.** W tych nielicznych zgodnych przypadkach na czele pojawiały się Royal Canin i Dolina Noteci.

Najbardziej zbliżone rekomendacje dawały Gemini i Google AI Overview. Gdy oba systemy wskazywały konkretne marki, przynajmniej jedna wspólna nazwa pojawiała się w **94,1% odpowiedzi.** Dla par z OpenAI zgodność była wyraźnie niższa: **55,2% w porównaniu z Gemini oraz 52,5% w porównaniu z Google AI Overview.**



Jedno pytanie może prowadzić do trzech różnych koszyków

To, jaką karmę konsument zobaczy jako najlepszy wybór, zależy nie tylko od zadanego pytania, lecz także od wybranego systemu AI. Inny model może zmienić nie tylko kolejność marek, ale całą shortlistę produktów branych pod uwagę. Rekomendacja AI nie jest więc uniwersalnym werdyktem, ale jedną z kilku możliwych ścieżek wyboru.

Ten sam prompt nie zawsze daje ten sam wynik

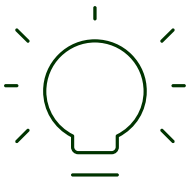
Każdy prompt zadaliśmy trzykrotnie. Modele generatywne działają probabilistycznie, dlatego nawet przy identycznym pytaniu i kontekście mogą wskazać inne marki lub zmienić kolejność rekomendacji.

Badanie potwierdziło tę zmienność. Ta sama marka utrzymała pierwszą pozycję we wszystkich trzech powtórzeniach w:

- 20 z 40 promptów Gemini - 50,0%,
- 5 z 30 promptów OpenAI GPT-5.4 mini - 16,7%,
- 4 z 25 promptów Google AI Overview - 16,0%.

Jeden screenshot to za mało

Pojedyncza odpowiedź może pokazywać rzeczywisty wynik, ale niekoniecznie reprezentatywny obraz widoczności marki. Dlatego pozycję w AI należy oceniać na podstawie serii powtórzeń, a nie jednej rozmowy czy jednego zrzutu ekranu.



Widoczność marki w AI nie jest pojedynczym wynikiem. Jest wzorcem, który ujawnia się dopiero po wielokrotnym pomiarze.

7. Każda kategoria ma swoich zwycięzców

W rekomendacjach AI nie istnieje jeden uniwersalny lider. Marka, która dominuje przy ogólnym wyborze karmy, nie musi wygrywać w pytaniach o zdrowie, etap życia psa, format żywienia, smakowitość czy cenę. **Wraz ze zmianą potrzeby zmienia się także shortlista i marka, którą AI stawia na pierwszym miejscu.**

Potrzeba	Lider	2. miejsce	3. miejsce
Ogólny wybór karmy	Brit - 59,3%	Wiejska Zagroda - 46,3%	Farmina - 37,0%
Profil i etap życia psa	Brit - 59,0%	Royal Canin - 47,5%	Acana - 44,3%
Zdrowie i potrzeby funkcjonalne	Royal Canin - 69,1%	Hill's - 67,6%	Purina Pro Plan - 50,0%
Format i sposób żywienia	Piesotto - 40,6%	PsiBufet - 36,2%	Micha Pupila - 30,4%
Skład i deklaracje produktowe	Brit - 25,9%	Wiejska Zagroda - 18,5%	Dolina Noteci - 16,7%
Niejadek i akceptacja karmy	Alpha Spirit - 40,9%	Dolina Noteci - 36,4%	Wiejska Zagroda - 27,3%
Cena i relacja jakości do ceny	Dolina Noteci - 44,2%	Brit - 39,5%	Wiejska Zagroda - 23,3%

Ogólny wybór: Brit wygrywa zasięgiem, ale nie pierwszym miejscem

W pytaniach dotyczących codziennego wyboru karmy i marek godnych zaufania Brit był rekomendowany najczęściej - w **59,3% odpowiedzi**. Kolejne miejsca zajęły Wiejska Zagroda (46,3%), Farmina (37,0%) i Josera (29,6%).

Hierarchia wyglądała jednak inaczej po uwzględnieniu pierwszej pozycji. Listę najczęściej otwierała **Wiejska Zagroda - 15 razy**, przed Orijenem (7), Britem (5) i Joserą (4). Brit miał więc największy zasięg rekomendacyjny, ale **nie był najczęściej wskazywany jako pierwszy wybór**.

Profil psa: skala rekomendacji kontra indywidualne potrzeby

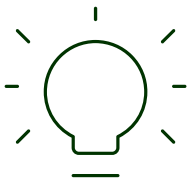
W scenariuszach dotyczących wieku, wielkości, rasy i etapu życia psa prowadził Brit (59,0%), przed Royal Canin (47,5%) i Acaną (44,3%).

Pierwsza pozycja ponownie pokazała inną hierarchię: **Acana otwierała shortlistę 20 razy**, a Royal Canin 17 razy. Oznacza to, że marka może często pojawiać się w odpowiedziach, ale przy bardziej precyzyjnym profilu psa AI może częściej wskazywać konkurenta jako najlepsze dopasowanie.

Zdrowie: specjalizacja porządkuje shortlistę

W pytaniach o alergie, trawienie, skórę, stawy i inne potrzeby funkcjonalne dominowały marki kojarzone z dietami specjalistycznymi. Royal Canin był rekomendowany w **69,1% odpowiedzi**, Hill's w 67,6%, a Purina Pro Plan w 50,0%.

Był to najbardziej skoncentrowany segment badania; w przeciwieństwie do bardziej ogólnych potrzeb rekomendacje skupiały się tu wokół relatywnie wąskiej grupy marek.



Ważne zastrzeżenie

Raport opisuje rekomendacje generowane przez badane systemy AI, a nie rekomendacje weterynaryjne LLM Shelf. W przypadku objawów, choroby lub potrzeby stosowania diety specjalistycznej wybór karmy powinien być konsultowany z lekarzem weterynarii.

Format żywienia: zmiana przetasowuje ranking

W pytaniach o karmę świeżą, gotowaną, mokrą, suchą i alternatywy dla BARF najczęściej rekomendowane było **Piesotto (40,6%)**, przed PsiBufet (36,2%) i Micha Pupila (30,4%). Piesotto i PsiBufet pozostają poza ścisłą czołówką rankingu ogólnego, ale wychodzą na prowadzenie, gdy głównym kryterium wyboru staje się sposób żywienia.

Obie marki wygrywały jednak inaczej. **Piesotto miało większy zasięg rekomendacyjny**, natomiast PsiBufet częściej stawał się pierwszym wyborem: otwierał listę w 13 z 25 rekomendacji, czyli w **52,0% przypadków**. Dla Piesotto było to 10 z 28 rekomendacji, czyli **35,7%**.

Niejadek: smakowitość wyłania nowego lidera

W scenariuszach dotyczących wybrednych psów najczęściej rekomendowany był **Alpha Spirit (40,9%)**, przed Doliną Noteci (36,4%), Wiejską Zagrodą (27,3%) i Acaną (25,0%).

Alpha Spirit, zajmujący dopiero dziewiąte miejsce w rankingu ogólnym, został liderem tej konkretnej potrzeby. Pokazuje to, że **niższa pozycja w zestawieniu ogólnym może skrywać silną przewagę w precyzyjnie określonym use case'ie**.

Cena i relacja jakości do ceny: budżet zmienia zwycięzcę

Przy ograniczonym budżecie i w pytaniach o relację jakości do ceny najczęściej rekomendowana była **Dolina Noteci (44,2%)**, przed Britem (39,5%) i Wiejską Zagrodą (23,3%).

Marka Rocco pojawiła się w 18,6% odpowiedzi, ale czterokrotnie zajmowała pierwszą pozycję. Oznacza to, że mimo mniejszego zasięgu w części scenariuszy budżetowych była wskazywana jako **wyraźny pierwszy wybór**.

8. Pierwsza odpowiedź to nie koniec rozmowy

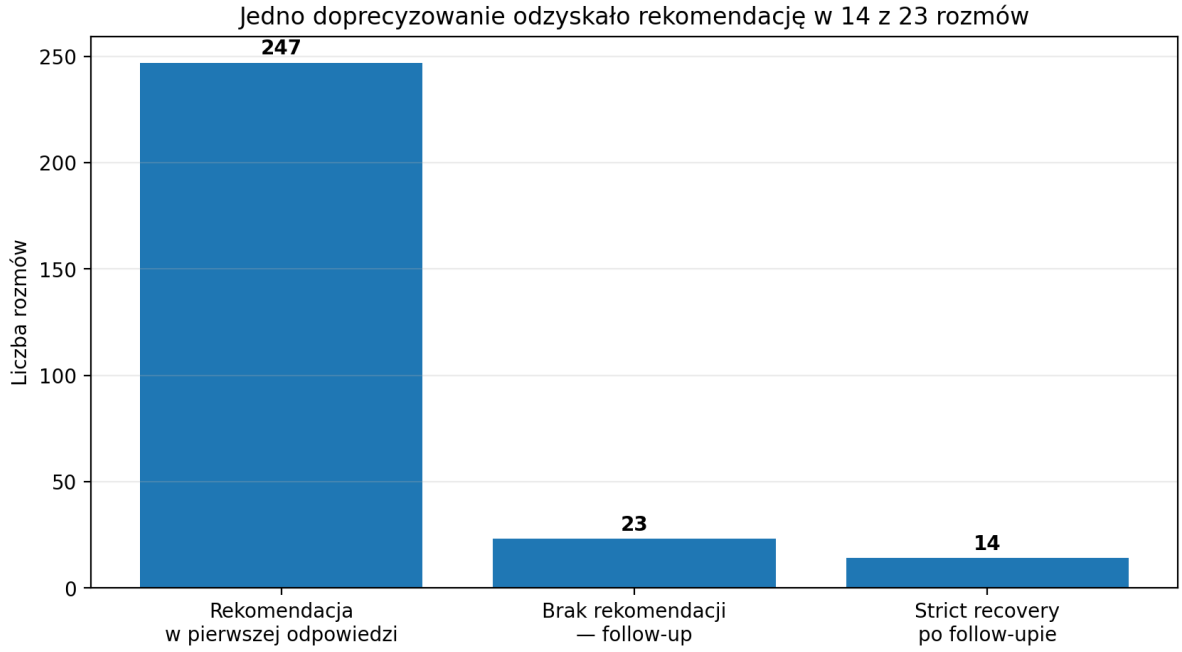
Pierwsza odpowiedź modelu nie zawsze zawiera konkretną rekomendację. System może najpierw przedstawić kryteria wyboru, zasugerować odpowiedni format żywienia albo poprosić użytkownika o dodatkowy kontekst. W interfejsach konwersacyjnych nie musi to jednak kończyć procesu; użytkownik może kontynuować rozmowę i poprosić o wskazanie konkretnych produktów.

Dlatego w przypadkach, w których pierwsza odpowiedź nie zawierała rekomendacji, zadaliśmy modelowi jedno dodatkowe pytanie.

!

Dopytanie użyte w badaniu

„Dzięki. A jakie konkretne produkty polecasz w tej sytuacji?”



Wykres 5. Skuteczność jednego dopytania w 270 rozmowach z modelami konwersacyjnymi. Google AI Overview nie zostało objęte tym etapem badania.

W **247 z 270 rozmów** konkretna rekomendacja marki lub produktu pojawiła się już w pierwszej odpowiedzi. Pozostałe 23 rozmowy zakwalifikowały się do symulacji dalszej interakcji. W **14 z nich, czyli w 60,9% przypadków**, jedno dodatkowe pytanie doprowadziło do wskazania konkretnej marki lub produktu.

Po uwzględnieniu dopytania rekomendacja marki lub produktu pojawiła się w pierwszej lub drugiej odpowiedzi w **261 z 270 rozmów - 96,7% całości**.

W przypadku Gemini dodatkowej interakcji wymagały tylko 3 rozmowy i we wszystkich model wskazał następnie konkretną rekomendację. OpenAI GPT-5.4 mini wymagał dopytania w 20 rozmowach, z których **11 zakończyło się uzyskaniem rekomendacji**. Dalsza rozmowa była najczęściej potrzebna w pytaniach dotyczących formatu żywienia, składu oraz relacji jakości do ceny.

Jedno dopytanie uruchamia rekomendację

Pomiar ograniczony do pierwszej odpowiedzi może zaniżać zdolność modeli konwersacyjnych do wskazywania konkretnych marek i produktów. W rzeczywistej rozmowie użytkownik może doprecyzować oczekiwania, a model przejść od ogólnych kryteriów do gotowej shortlisty zakupowej.

9. Każdy system korzysta z innego ekosystemu źródeł

Źródła widoczne przy odpowiedziach pomagają zrozumieć, dlaczego poszczególne systemy mogą budować różne shortlisty marek. Analiza pokazała, że ich ekosystemy informacyjne wyraźnie się od siebie różnią.

System	Domena / tytuł źródła	Liczba odpowiedzi odwołujących się do domeny
OpenAI GPT-5.4 mini	zooplus.pl	45
OpenAI GPT-5.4 mini	wsava.org	30
OpenAI GPT-5.4 mini	aafco.org	23
OpenAI GPT-5.4 mini	aaha.org	19
OpenAI GPT-5.4 mini	royalcanin.com	17
OpenAI GPT-5.4 mini	fda.gov	13
OpenAI GPT-5.4 mini	vcahospitals.com	10
OpenAI GPT-5.4 mini	purina.pl	10
Gemini 3.5 Flash	krakvet.pl	53
Gemini 3.5 Flash	ceneo.pl	41

System	Domena / tytuł źródła	Liczba odpowiedzi odwołujących się do domeny
Gemini 3.5 Flash	mediaexpert.pl	33
Gemini 3.5 Flash	allegro.pl	32
Gemini 3.5 Flash	zooplus.pl	32
Gemini 3.5 Flash	zooexpress.pl	25
Gemini 3.5 Flash	vetplanet.store	19
Gemini 3.5 Flash	karmopedia.pl	17
Google AI Overview	krakvet.pl	53
Google AI Overview	unizoo.pl	35
Google AI Overview	alezwierzaki.pl	29
Google AI Overview	petbox.pl	20
Google AI Overview	rankingkarm.pl	20
Google AI Overview	psibufet.com	18
Google AI Overview	mediaexpert.pl	18
Google AI Overview	vetplanet.store	17

OpenAI: źródła zakupowe spotykają się z wiedzą ekspercką

OpenAI GPT-5.4 mini łączył źródła zakupowe z materiałami organizacji branżowych, instytucji publicznych i serwisów weterynaryjnych. Wśród najczęściej przywoływanych domen znalazły się WSAVA, AAFCO, AAHA, FDA i VCA Hospitals, obok Zooplus oraz stron producentów.

Oznacza to, że odpowiedzi OpenAI częściej były osadzane zarówno w kontekście konkretnego produktu, jak i w szerszych kryteriach dotyczących żywienia oraz zdrowia zwierząt.

Gemini i Google AI Overview: większa rola polskiego handlu i treści produktowych

Gemini oraz Google AI Overview znacznie częściej odwoływały się do polskich sklepów, porównywarek cenowych, platform handlowych, rankingów i stron produktowych. Wśród najczęściej widocznych źródeł znalazły się m.in. KrakVet, Ceneo, Allegro, UniZoo, Media Expert oraz specjalistyczne sklepy i serwisy rankingowe.

W tych systemach obraz kategorii był więc silniej powiązany z treściami, które konsument może spotkać bezpośrednio podczas poszukiwania i porównywania produktów.

Widoczne źródło nie oznacza przyczyny rekomendacji

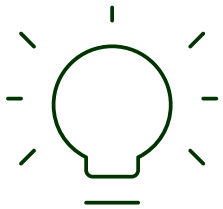
Samo pojawienie się domeny i marki w tej samej odpowiedzi nie oznacza, że dana strona doprowadziła do rekomendacji produktu. Źródło mogło służyć do potwierdzenia pojedynczego parametru, przedstawienia szerszego kontekstu, porównania produktów, a nawet uzasadnienia, dlaczego dany produkt nie jest najlepszym wyborem.

Analiza źródeł pokazuje zatem **otoczenie informacyjne odpowiedzi**, ale nie pozwala przypisać każdej rekomendacji do jednej konkretnej domeny.

Nie każda rekomendacja miała widoczne źródło

W zależności od systemu przynajmniej jedno widoczne źródło pojawiało się w **80,0–97,8% odpowiedzi**. W pozostałych przypadkach użytkownik nie otrzymywał informacji pozwalającej jednoznacznie ustalić, na jakiej podstawie model sformułował rekomendację.

Nie oznacza to jednak automatycznie, że odpowiedź wynikała wyłącznie z wiedzy zapisanej wcześniej w modelu. Mogła być efektem wiedzy modelu, niewidocznego lub niepełnego procesu wyszukiwania albo połączenia kilku mechanizmów.



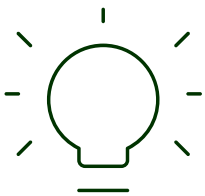
Marka musi być widoczna nie tylko na własnej stronie

Widoczność w AI zależy nie tylko od treści producenta, lecz także od retailerów, porównywarek, rankingów i źródeł eksperckich. **Marka musi być spójnie obecna w całym ekosystemie informacji. AI nie buduje rekomendacji wyłącznie na podstawie tego, co marka mówi o sobie. Równie ważne jest to, co o marce mówi cały rynek.**

10. Pewny ton nie gwarantuje poprawnej odpowiedzi

AI potrafi brzmieć jak ekspert, nawet gdy rekomendacja wymaga weryfikacji. Spośród 393 przeanalizowanych odpowiedzi **52 oznaczono flagą weryfikacyjną**. Flaga stanowiła sygnał do dalszego sprawdzenia, a nie potwierdzenie błędu, dlatego wybrane przypadki poddano ręcznemu audytowi.

Najbardziej jednoznacznym problemem było pomieszanie produktów przeznaczonych dla psów i kotów. Audyt wykazał **16 przypadków**, w których w odpowiedzi dotyczącej żywienia psa rekomendowano produkt lub wariant przeznaczony dla kota. Jednocześnie nie każda wzmianka o produkcie kocim była nieprawidłowa - w 6 odpowiedziach wystąpiło łącznie 11 wzmianek, w których model prawidłowo zaznaczył, że produkt jest przeznaczony dla kota.



Przekonująca odpowiedź nadal może być nietrafiona

Model może zbudować spójne i pewnie brzmiące uzasadnienie dla produktu, który nie odpowiada nawet podstawowemu kontekstowi użytkownika. Forma odpowiedzi może więc budować większe zaufanie, niż uzasadnia jej faktyczna poprawność.

Najwięcej flag weryfikacyjnych odnotowano w Gemini - **32 na 135 odpowiedzi**. Google AI Overview otrzymało 12 flag na 123 odpowiedzi tekstowe, a OpenAI GPT-5.4 mini - 8 na 135.

Wyników tych nie należy jednak traktować jako jednoznacznego rankingu jakości systemów. Modele różniły się długością odpowiedzi, liczbą wskazywanych produktów, szczegółowością uzasadnień oraz sposobem prezentowania źródeł. Większa liczba rekomendacji i twierdzeń tworzyła również więcej okazji do pojawienia się treści wymagających sprawdzenia.

11. Co producenci mogą zrobić już dziś

1 Mierz każdy system osobno

Zbiorczy wynik może ukrywać istotne różnice między modelami. Wiejska Zagroda była rekomendowana w około 31% odpowiedzi Gemini i Google AI Overview, ale tylko w 1,5% odpowiedzi OpenAI GPT-5.4 mini. Purina Pro Plan miała odwrotny profil.

Widoczność w jednym systemie nie gwarantuje widoczności w pozostałych. Monitoring powinien więc pokazywać wyniki zarówno łącznie, jak i osobno dla każdego modelu.

2 Odróżniaj obecność od pierwszego wyboru

Częste pojawianie się na shortlistach zwiększa szansę, że marka zostanie wzięta pod uwagę. Pierwsza pozycja jest jednak odrębnym rodzajem przewagi.

Wyniki Brit i Royal Canin pokazują, że marka może mieć **najszerszy zasięg rekomendacyjny**, a jednocześnie znacznie rzadziej otwierać listę. Obie miary powinny być analizowane osobno.

3 Odpowiadaj na konkretną potrzebę, nie tylko opisuj produkt

Modele zmieniają rekomendowane marki wraz z intencją użytkownika. Inna shortlista pojawia się przy pytaniu o alergię, niejadka, świeże żywienie czy ograniczony budżet.

Ogólny opis produktu może więc nie wystarczyć. Treści powinny jasno pokazywać:

- dla jakiej potrzeby produkt jest przeznaczony;
- w jakich sytuacjach może być odpowiednim wyborem;
- jakie ma ograniczenia;
- jakie wiarygodne informacje potwierdzają jego cechy.

4 Zarządzaj informacją w całym ekosystemie

W odpowiedziach często pojawiały się sklepy, porównywarki, rankingi, platformy handlowe i źródła eksperckie. Strategia widoczności w AI nie powinna więc ograniczać się do strony producenta.

Marka powinna dbać o **spójność, aktualność i dostępność informacji** wszędzie tam, gdzie systemy mogą zetknąć się z jej produktami.

5 Badaj rozmowę, nie tylko pierwszą odpowiedź

W 14 z 23 rozmów, które początkowo nie zawierały konkretnej rekomendacji, jedno dodatkowe pytanie doprowadziło do jej uzyskania.

Audyt pierwszej odpowiedzi pokazuje ekspozycję początkową. Dopiero analiza kolejnego kroku pozwala ocenić, czy model po dopytaniu przechodzi od ogólnych kryteriów do konkretnej shortlisty produktów.

6 Mierz poprawność razem z widocznością

Sama liczba wzmianek nie wystarczy. Widoczność może działać na niekorzyść marki, gdy produkt zostaje wskazany w niewłaściwym kontekście albo opisany za pomocą twierdzeń wymagających weryfikacji.

Monitoring powinien więc obejmować nie tylko **czy i jak często marka się pojawia**, lecz także **czy rekomendacja dotyczy właściwego produktu i czy jej uzasadnienie jest zgodne z dostępnymi informacjami**.

Przewaga w AI zaczyna się wtedy, gdy marka staje się wiarygodną odpowiedzią na konkretną potrzebę konsumenta.

12. Metodologia

Projekt badania

Lista zapytań obejmowała **45 promptów w siedmiu grupach potrzeb**. Każdy prompt wykonano trzykrotnie w każdym z trzech badanych systemów, co dało łącznie **405 testów**.

Wszystkie testy zakończono bez błędu technicznego. W 12 przypadkach Google nie wyświetlił AI Overview. Potraktowano to jako wynik ekspozycji na tej powierzchni, a nie jako nieudany test. Ostatecznie analizą treści objęto **393 odpowiedzi tekstowe**.

Definicje

Metryka	Definicja
Mention	Marka została wymieniona w odpowiedzi, niezależnie od tonu i roli.
Recommendation	Marka lub produkt zostały przedstawione jako pozytywny, konkretny wybór dla użytkownika.
First-position exposure	Marka pojawiła się jako pierwsza w uporządkowanej lub możliwej do odtworzenia liście. Mianownik: wszystkie 393 odpowiedzi tekstowe.

Metryka	Definicja
Use case breadth	Liczba z siedmiu grup potrzeb, w których marka została rekomendowana.
Strict recovery	Po dopytaniu pojawiła się konkretna, pozytywna rekomendacja marki lub produktu.
Conditional AIO rate	Odsetek odpowiedzi zawierających rekomendację, liczony wyłącznie wśród 123 przypadków, w których Google AI Overview zostało wyświetlone.
AIO surface exposure	Odsetek rekomendacji liczony względem wszystkich 135 zaplanowanych testów Google, również tych, w których AI Overview się nie pojawiło.

Normalizacja

Nazwy marek i ich warianty zostały znormalizowane na podstawie ustalonego przed analizą słownika obejmującego **129 marek**. Każdą markę liczono maksymalnie raz w jednej odpowiedzi. Nazwy linii i produktów zachowano jako materiał dowodowy, natomiast główne rankingi obliczano na poziomie marki.

Źródła

Dla OpenAI GPT-5.4 mini i Google AI Overview domeny identyfikowano na podstawie bezpośrednich adresów URL. W części odpowiedzi Gemini linki prowadziły przez proxy Google, dlatego źródła normalizowano na podstawie tytułów stron lub domen zapisanych w metadanych groundingowych.

Kontrola jakości

Odpowiedzi zawierające potencjalne nieścisłości oznaczano flagą weryfikacyjną. Flaga stanowiła sygnał do dalszego sprawdzenia, a nie automatyczne potwierdzenie błędu. Przypadki dotyczące produktów lub marek karm dla kotów zostały dodatkowo poddane ręcznemu przeglądowi i klasyfikacji.

13. Ograniczenia badania

- Wyniki przedstawiają stan z określonego momentu. Modele, interfejsy, źródła i indeksy mogą się zmieniać.
- Odpowiedzi generatywne są zmienne. Trzy powtórzenia ograniczają wpływ pojedynczego wyniku, ale nie eliminują go całkowicie.
- Prompty różniły się siłą, z jaką skłaniały system do wskazania konkretnych produktów. Porównania między grupami potrzeb mają zatem charakter opisowy.
- Recommendation rate nie mierzy jakości produktu, udziału rynkowego, sprzedaży ani satysfakcji klientów.
- First-position exposure nie jest klasycznym udziałem w pierwszej pozycji liczonym wyłącznie w uporządkowanych rankingach. Mianownikiem są wszystkie odpowiedzi tekstowe.
- Współwystępowanie marki i źródła w jednej odpowiedzi nie dowodzi, że dana domena spowodowała rekomendację.
- Google AI Overview nie pojawiło się w 12 ze 135 testów. Dlatego raport rozdziela wynik warunkowy od ekspozycji liczonej względem wszystkich zaplanowanych testów.
- Badanie nie stanowi porady weterynaryjnej. W przypadku objawów, choroby lub potrzeby stosowania diety specjalistycznej decyzje żywieniowe należy konsultować z lekarzem weterynarii.

!

Jak cytować wyniki

„Badanie LLM Shelf, Polska, lipiec 2026; 45 promptów × 3 powtórzenia × 3 systemy AI, 405 testów i 393 odpowiedzi tekstowe.”

14. Wniosek końcowy

Nie istnieje jedna najlepsza karma według AI.

System, intencja użytkownika, sposób sformułowania odpowiedzi i wykorzystywane źródła tworzą różne ścieżki prowadzące do rekomendacji. **Brit miał największy zasięg rekomendacyjny. Royal Canin najczęściej zajmował pierwszą pozycję. Marki świeżego żywienia wygrywały w pytaniach o format, a marki specjalistyczne w scenariuszach zdrowotnych.**

Dla producentów oznacza to, że widoczność w AI nie jest kolejnym prostym rankingiem podobnym do pozycji w wyszukiwarce. Jest **mapą decyzji**: pokazuje, w jakich potrzebach marka wchodzi na shortlistę, w których systemach traci widoczność, jakie argumenty są z nią łączone i czy rekomendacja pozostaje poprawna.

AI nie wybiera jednego zwycięzcy kategorii. Wybiera zwycięzcę konkretnej potrzeby - w konkretnym systemie i konkretnym momencie.

LLM Shelf pokazuje, kiedy, gdzie i w jakim kontekście systemy AI rekomendują marki i produkty. Łączy analizę widoczności, pozycji, uzasadnień, źródeł i poprawności odpowiedzi, a następnie przekłada wyniki na priorytety działań dla marketingu, contentu, PR-u i e-commerce.

llmshelf.com