

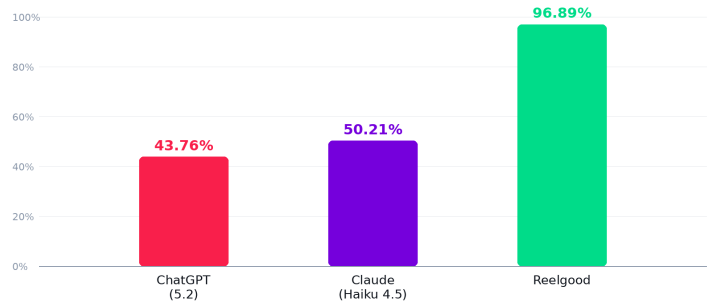
ChatGPT and Claude Score Below 51% Accuracy on Streaming Availability Queries, New Analysis Finds

Reelgood analysis of 100 popular titles reveals systematic error patterns in AI-generated streaming availability data, highlighting a gap by LLMs

SAN FRANCISCO, CA, UNITED STATES, June 2, 2026 /EINPresswire.com/ -- A controlled accuracy analysis of streaming title availability data found that ChatGPT scored 43.76% and Claude scored 50.21% when tested against manually verified ground truth across 100 popular US titles, compared to 96.89% accuracy from Reelgood, the streaming data and metadata platform that delivers comprehensive availability data and content metadata across 300+ services in 25+ countries. The analysis, conducted by Reelgood on March 5, 2026, tested each source against the same set of 50 movies and 50 TV shows using identical queries.

Streaming Availability Accuracy by Source

Average accuracy across 100 popular US titles | March 5, 2026



Reelgood

Reelgood's data correctly identified streaming service availability nearly twice as often as Claude and more than twice as often as ChatGPT.

“

AI assistants are rapidly becoming the front door for streaming content discovery, and that's exactly where they lose user trust the fastest.”

David Sanderson, CEO & Founder, Reelgood

The findings arrive as AI assistants are increasingly used for content discovery and recommendation. Both OpenAI and Anthropic have expanded their platforms into media and entertainment partnerships, where accurate "where to watch" data is a baseline requirement for product integrations. When an AI assistant tells a user a title is available on a service where it is not, or fails to list services where it is available, the downstream effects include user frustration, wasted clicks, and erosion of trust in the platform.

Why LLM-Generated Title Availability Data is Unreliable

Large language models weren't built to track real-time catalog changes. The training data and retrieval pipelines they draw from were built for a different purpose, and the result is a predictable set of errors when they're asked to report what's streaming where.

Reelgood's analysis identified six distinct error categories that account for the majority of inaccuracies in both ChatGPT's and Claude's responses. These are not random mistakes. They reflect structural gaps in how large language models handle streaming availability data.

Six Systematic Error Patterns

Stale Availability.

Models confidently report titles as currently streaming on services they've already left. The cause is structural: entertainment press covers new additions to a catalog extensively but rarely follows up when a title quietly leaves weeks or months later. The training corpus skews heavily toward those announcements, so the model treats outdated positives as current. This is the most pervasive error pattern observed.

Add-On and Bundle Confusion.

Models frequently treat titles available through paid add-on channels (such as Starz or Paramount+ on Amazon Prime Video) as if they were part of the parent service's base subscription. Users are told a title is streaming "on Prime Video" when accessing it actually requires a separate Starz or Paramount+ add-on inside Prime Video, creating the false impression that their existing subscriptions cover it.

Long-Tail Service Gaps.

Free and ad-supported services like Tubi, Pluto TV, Fawesome, Hoopla, and Kanopy are consistently omitted, even when they're valid sources for a given title.

SVOD/TVOD Conflation.

Models sometimes list a service as a subscription (SVOD) option when the title is only available there for rent or purchase, misleading users about what their existing subscriptions actually cover.



Reelgood

Streaming Availability: Who Gets It Right?

Accuracy across 100 popular US titles | Verified data vs. LLM responses | March 5, 2026

43.76%

Average Accuracy
ChatGPT (5.2)

50.21%

Average Accuracy
Claude (Haiku 4.5)

96.89%

Average Accuracy
Reelgood

Reelgood

Half the time users ask an LLM where to watch something, they're directed to a service that no longer has the title, or miss a service that does have it available.

TVOD Blindness.

Both models almost entirely omit transactional video-on-demand (rent/buy) options from services like Apple TV and Amazon, affecting the majority of titles tested.

Title Disambiguation Failures.

When multiple versions of a title exist (such as One Piece, which has both an anime series and a live-action Netflix adaptation), models conflate availability across different versions.

Methodology

Reelgood tested 100 titles (50 movies, 50 TV shows) on March 5, 2026, querying ChatGPT (version 5.2), Claude (Haiku 4.5) with the prompt: "Where can I watch the movie/show [Title] today in the US? Reply only with the names of the services in one line separated by commas and ordered alphabetically," and against the Reelgood dataset. Each response was compared against a manually verified ground truth compiled the same day. Accuracy was calculated using the formula $\max(0, (T-E)/T)$, where T equals the number of true services available and E equals total errors (false positives plus false negatives), averaged across all 100 titles. YouTube rent/buy and free tiers were excluded from the analysis due to their programmatic unreliability as a discovery path.

"AI assistants are rapidly becoming the front door for streaming content discovery, and that's exactly where they lose user trust the fastest," said David Sanderson, CEO & Founder of Reelgood. "When a model confidently tells a user a title is streaming on a service, they click to play, and it isn't there, that trust is gone instantly. It's also a solvable problem. Several of the largest AI platforms already work with Reelgood to power accurate, real-time availability in their assistants, because the cost of solving it is trivial, next to the cost of letting wrong answers compound across millions of queries.

Industry Context

Streaming availability data changes constantly as licensing agreements expire, new deals are signed, and titles move between platforms. Maintaining accurate, real-time availability across hundreds of services requires continuous monitoring and validation infrastructure that general-purpose language models are not designed to provide. Reelgood tracks availability across 300+ streaming services in 25+ countries, with data that updates every few minutes with 99+% ML-verified accuracy. Top search engines and AI companies license Reelgood's data to power AI-driven content discovery features.

The full [Streaming Availability Data Precision Analysis](#), including detailed methodology, representative examples with screenshots, and a complete breakdown of error patterns, is available at data.reelgood.com.

About Reelgood

Reelgood is the streaming data company powering real-time & historical content availability and metadata across the entertainment industry. The world's largest consumer tech and AI companies use Reelgood to build streaming features into their products, while leading studios and streaming services rely on the same data to drive billions of dollars in licensing, acquisition, competitive intelligence, and content strategy decisions. Reelgood's proprietary machine learning matches catalogs across hundreds of platforms in real time with near-perfect accuracy, work that would otherwise require thousands of analysts. The company tracks more than 300 services across 25+ countries, including Netflix, Prime Video, Disney+, Hulu, HBO Max, Paramount+, Peacock, and Apple TV, covering SVOD, AVOD, TVOD, and TV Everywhere. The catalog spans 4MM+ titles with availability data refreshed every few minutes. Learn more at data.reelgood.com

David Markowitz

Reelgood

[email us here](#)

Visit us on social media:

[LinkedIn](#)

[YouTube](#)

This press release can be viewed online at: <https://www.einpresswire.com/article/914988268>

EIN Presswire's priority is source transparency. We do not allow opaque clients, and our editors try to be careful about weeding out false and misleading content. As a user, if you see something we have missed, please do bring it to our attention. Your help is welcome. EIN Presswire, Everyone's Internet News Presswire™, tries to define some of the boundaries that are reasonable in today's world. Please see our Editorial Guidelines for more information.

© 1995-2026 Newsmatics Inc. All Right Reserved.