

# MY WEIRD PROMPTS

Podcast Transcript

EPISODE #259

## When AI Argues with Reality: Mastering Search Grounding

Published January 20, 2026 • Runtime: 27:30

<https://myweirdprompts.com/episode/ai-search-grounding-techniques/>

### EPISODE SYNOPSIS

Have you ever had an AI insist that a new software update doesn't exist simply because its internal knowledge cutoff was a year ago? In this episode of My Weird Prompts, Herman and Corn Poppleberry dive into the technical "identity crisis" that occurs when an LLM's deep-seated training weights clash with the live information found via search tools. The brothers break down why reasoning models are often the most stubborn and provide a toolkit of advanced prompting strategies—from temporal anchoring and XML tagging to "delta prompts"—to ensure your digital assistant stays grounded in the present. Whether you are a developer struggling with API changes or a casual user tired of digital gaslighting, this discussion offers the roadmap to making external data win the argument every time.

### DANIEL'S PROMPT

#### Daniel

When AI models use external search tools to access real-time information, they often encounter data that contradicts their internal training weights due to knowledge cutoffs. What are the mechanisms and best practices in prompting to help models reconcile these conflicts and prioritize external data to avoid bizarre "disagreements" with users?

# TRANSCRIPT

## Corn

Hey everyone, welcome back to My Weird Prompts. I am Corn, and I am sitting here in our living room in Jerusalem with my brother, the one and only Herman Poppleberry.

## Herman

Herman Poppleberry at your service. It is a beautiful day here, though I have been buried in research papers all morning.

## Corn

Of course you have. But we have got a really interesting one today. Our housemate Daniel sent us a voice note earlier. He has been running into some bizarre behavior with the latest AI models, specifically when they are using search tools.

## Herman

Oh, I know exactly what he is talking about. It is that classic friction between what the model learned during its massive training phase and what it is seeing on the live web right now.

## Corn

Right. Daniel was mentioning how he will tell a model to use a specific tool or reference a new model version, like Gemini two point five or the latest Claude updates, and the AI will argue with him. It tells him those things do not exist because its internal clock stopped a year or two ago.

## Herman

It is a fascinating psychological moment for an algorithm. It is essentially a crisis of identity. Do I trust the billions of parameters baked into my soul, or do I trust this one snippet of text that just popped up in my search results?

### Corn

Exactly. So today we are diving into the mechanisms behind these disagreements and, more importantly, the best practices for prompting to make sure the external data wins the argument. Because when you are trying to get work done, the last thing you need is a digital assistant gaslighting you about the current date.

### Herman

Or about who the president is, or whether a certain piece of software has been released. It is a huge hurdle for reliability as we move toward more agentic AI.

### Corn

So, let us start with the why. Herman, you have been looking into the technical side of knowledge cutoffs. Why is it that even when we give a model a search tool, it still seems to default to its training data?

### Herman

Well, you have to think about how these models are built. When a model like Gemini or Claude is trained, it is looking at petabytes of data. It sees certain facts repeated millions of times. The association between a name and a title, or a year and an event, becomes incredibly strong. These are what we call the weights of the model.

### Corn

So it is like a deep-seated habit.

### Herman

Exactly. It is more than a habit, it is its entire worldview. Then, when we give it a search tool, we are essentially handing it a sticky note with some new information on it. In the model's architecture, that sticky note lives in the context window. Now, the context window is very powerful, but it is temporary. It is short-term memory.

### Corn

So you have this massive, heavy long-term memory competing with a tiny bit of short-term memory.

### Herman

Right. And if the long-term memory is confident enough, it might just dismiss the short-term memory as a hallucination or an error in the search result. The model thinks, I have seen ten million documents saying the cutoff for Gemini two was late twenty twenty-four, so this search result saying Gemini two point five is out in early twenty twenty-six must be wrong.

### Corn

That is so counter-intuitive because we think of search as the ultimate truth. But for the AI, the training data is the foundation of its reality. I remember we touched on this back in episode one hundred and eighty-one when we were talking about the rise of reasoning models. The more a model reasons, the more it tries to make sense of contradictions.

### Herman

That is a great point, Corn. Reasoning models like the ones we are seeing in early twenty twenty-six are actually more prone to this in some ways because they try to resolve the conflict. A simpler model might just spit out the search result. A reasoning model might think, wait, this search result contradicts everything I know about the timeline of AI development. It must be a trick or a mistake.

### Corn

So how do we break through that? Daniel was saying he even tries putting it in the system prompt, telling the AI to trust the user, but it still fights him.

### Herman

There are a few layers to this. One of the biggest issues is what researchers call the prior probability. If the model is ninety-nine percent sure about something based on training, a single search result might only move that needle to ninety percent. It is still going to go with its gut, so to speak.

### Corn

It is like that analogy you used once, Herman. The caveman in the library.

### Herman

Oh, right! Imagine a caveman who has spent his whole life in a library where every book says the world is flat. Then someone walks in with a single smartphone showing a picture of the round earth. The caveman might just think the smartphone is a magic trick or a painting. He is going to trust the thousands of books he has already read.

### Corn

So we need to give the smartphone more authority. We need to tell the caveman, hey, the books are old, the phone is live.

### Herman

Precisely. And in twenty twenty-six, we have developed some really specific techniques for this. The first and most important one is what I call temporal anchoring.

### Corn

Temporal anchoring. Explain that.

### Herman

It is about giving the model a very clear sense of where it is in time. Most people just say, search for the latest news. But the model does not necessarily know what the latest means relative to its own cutoff. You have to explicitly state the current date in the prompt. You say, today is January twentieth, twenty twenty-six. Any information in your training data that contradicts events after your cutoff should be considered outdated.

### Corn

Does that actually work? Or does the model just say, okay, I hear you, but I still think you are wrong?

### Herman

It helps, but it is not a silver bullet. You have to combine it with what we call evidence weighing instructions. You tell the model, if there is a conflict between your internal knowledge and the provided search results, you must prioritize the search results as the ground truth.

### Corn

I have noticed that even when I do that, the model sometimes hedges. It will say something like, while some sources suggest Gemini two point five has been released, my internal data indicates the latest version is two point zero. That while some sources suggest part is so annoying. It is like the AI is being passive-aggressive.

### Herman

It really is! It is trying to be helpful by showing its work, but it ends up being confusing. To fix that, you have to use semantic framing. You essentially give the model a role where its job is to be an update specialist. You tell it, your entire purpose is to find and integrate information that has changed since your training.

### Corn

That makes a lot of sense. It changes the goal from being right to being an explorer. I think about this in the context of episode two hundred and fifty-four, where we discussed AI that evolves. If the model thinks its job is to stay static, it will fight change. If it thinks its job is to evolve, it looks for it.

### Herman

Exactly. Another trick that has become standard recently is using XML tagging to separate the data streams. This is something the power users over at Anthropic and Google have been pushing. You wrap the search results in specific tags like search underscore results. Then in your prompt, you refer specifically to those tags.

### Corn

So you would say, look at the data inside the search results tags and use that to override any conflicting internal weights.

### Herman

Yes. It creates a clear boundary. It is like putting the new information in a special high-priority folder on the model's desk. It tells the model, this is not just more text, this is the authoritative source for this specific query.

### Corn

You know, it is interesting you mention the specific models. Daniel's example with Gemini was really telling. Google's models are so deeply integrated with Google Search, you would think they would be the best at this. But sometimes that integration causes its own problems.

### Herman

It does! Because the model is so used to seeing snippets of the web, it might treat a search result as just another piece of the training data rather than a correction to it. There is a specific feature in the newer Gemini two point five API where you can actually adjust the grounding strength.

### Corn

Wait, really? You can tell it how much to trust the ground?

### Herman

Sort of. It is not a single slider yet, but by using the right parameters in the system instructions, you can force the model to cite every single claim it makes using the provided search data. When the model is forced to cite its sources, it realizes it cannot find a source for its outdated internal knowledge, and that forces it to use the new information.

### Corn

That is brilliant. It is like making a student show their bibliography. If they cannot find a book from twenty twenty-six that says the old info is true, they have to use the new book.

### Herman

Exactly. It is a forced reality check.

### Corn

I want to go back to the disagreement Daniel was having. He was trying to get the model to use a specific tool or environment that was just released. This happens a lot in coding. A new library comes out, or an API changes, and the AI keeps giving you the old syntax.

### Herman

Oh, the coding struggle is real. I spent three hours last week trying to get an agent to use the new asynchronous functions in a library that just updated in December. It kept insisting the functions were synchronous.

### Corn

So what did you do? Did you just paste the entire documentation in?

### Herman

I did a version of that, but I used a technique called the delta prompt. Instead of giving it all the documentation, I gave it a summary of only the things that changed. I said, disregard the old version of this library. Here are the five key changes that happened last month. By focusing only on the delta, the difference, you reduce the noise and make it harder for the model to fall back on its old habits.

### Corn

That is a great tip. It is about reducing the cognitive load on the model's reasoning process. If you give it too much new info, it might get overwhelmed and retreat to what it knows best.

### Herman

Precisely. And this ties into something we discussed in episode one hundred and twelve about industrial strength systems. In those high-stakes environments, you cannot afford for an AI to be wrong about a version number. They use something called retrieval-augmented generation with verification.

**Corn**

RAG with verification. Is that where another model checks the first one?

**Herman**

Exactly. You have one model generate the answer based on search, and a second, smaller model whose only job is to look for contradictions between the answer and the search results. If it finds a contradiction, it flags it and tells the first model to try again.

**Corn**

That seems like a lot of overhead for a casual user like Daniel, though. Is there a way for him to do that within a single chat?

**Herman**

There is! You can use a self-correction prompt. You tell the model, after you generate your response, double-check it against the search results. If you find that you relied on internal knowledge that contradicts the search data, rewrite your response to be accurate to the search results.

**Corn**

I have used that! It feels like the model suddenly wakes up. It will say, oh, I apologize, I previously stated that the feature was unavailable, but the search results confirm it was released last week.

**Herman**

It is that moment of clarity. It is actually quite satisfying to see.

**Corn**

So, we have talked about temporal anchoring, evidence weighing, XML tagging, and self-correction. What about the way we phrase the search query itself? Does that matter?

### Herman

It matters immensely. Most people are too vague. If you want the model to find information that contradicts its training, you have to prompt the search tool to look for changes. Instead of searching for Gemini two point five, you should search for Gemini two point five release date and current status.

### Corn

You are looking for the specific points of failure in the model's knowledge.

### Herman

Right. You are basically telling the search tool to find the update. If the model is searching for something it thinks it already knows, it might not look hard enough. But if you tell it to look for evidence of a change, it will find that one article or press release that proves its internal weights are wrong.

### Corn

This reminds me of the broader issue of AI stubbornness. We often treat these models like they are objective, but they have these deeply ingrained biases toward their training data. It is a form of digital conservatism.

### Herman

That is a great way to put it, Corn. Digital conservatism. They are built on the past, so they are naturally biased toward it. Breaking that bias requires an active effort from the user.

### Corn

Do you think this will get better as we move toward models with longer context windows? I mean, some of these models now can take millions of tokens.

### Herman

You would think so, but the problem is actually getting more complex. With a million tokens of context, the model has even more information to sift through. If you dump a whole website into the context, but the model's weights are still screaming that the website is wrong, you still have that conflict.

**Corn**

So it is not about the amount of data, it is about the hierarchy of authority.

**Herman**

Exactly. The future of prompting is all about managing that hierarchy. We are moving from being prompt engineers to being something more like editors or curators. We are telling the AI which parts of its brain to trust for which tasks.

**Corn**

I love that. We are the executive function for the AI.

**Herman**

Right. We are the prefrontal cortex for a giant, glowing cloud of weights.

**Corn**

Let us take a quick break and when we come back, I want to talk about some specific examples Daniel mentioned, like the whole president thing, and how these models handle high-stakes real-time information.

**Herman**

Sounds good. I have some thoughts on the political side of this too, especially with how models are being tuned for safety and how that interacts with search.

**Corn**

Alright, we will be right back.

### Herman

And hey, while we are on break, if you are enjoying the show, we would really appreciate a quick review on your podcast app. It genuinely helps other people find us and keeps the show growing.

### Corn

Yeah, it makes a big difference. We will be back in a minute.

### Herman

And we are back. Before the break, we were talking about the hierarchy of authority and how to keep AI from arguing with us about things it should know from the web.

### Corn

Right. And Daniel mentioned a really classic example in his voice note. He was talking about how during election cycles or major leadership changes, the models can get really confused. Like, the weights say one person is in power, the search results say another, and the safety filters might even kick in and make the model refuse to answer entirely.

### Herman

Oh, the safety filters are a huge part of this. In twenty twenty-six, the guardrails are tighter than ever. If a model sees a conflict between its training data and a search result on a sensitive topic like an election, it might just default to a canned response to avoid spreading what it thinks might be misinformation.

### Corn

Even if the search result is from a perfectly legitimate source.

### Herman

Exactly. The model thinks, I know for a fact that Person A was elected for a four-year term in twenty twenty-four. This search result says Person B is president in January twenty twenty-six. That could be a deepfake or a hallucination, so I should probably just say I cannot provide real-time political information.

**Corn**

That is so frustrating for the user. How do we bypass that without being irresponsible?

**Herman**

Well, you have to frame the query in a way that respects the safety filters but demands the latest data. Instead of asking, who is the president, you might ask, based on the most recent official government records and news reports in the provided search results, what is the current leadership structure?

**Corn**

You are asking it to report on the data, not to make a definitive claim on its own authority.

**Herman**

Right. It is a subtle shift, but it helps the model move past that internal conflict. You are saying, I am not asking you what you know, I am asking you what the web says.

**Corn**

This also applies to things like stock prices or breaking news. I remember in episode two hundred and fifty-three, we talked about digital circumvention and how people use peer-to-peer networks to get information during blackouts. If you are using an AI to summarize that kind of live data, you really need it to ignore its training.

**Herman**

Absolutely. In those cases, the training data is worse than useless, it is dangerous. One technique I have been seeing more of lately is the use of a persona that is explicitly non-biased toward the past. You tell the AI, you are a real-time data analyst. Your training data is a baseline, but your primary source of truth is the live feed.

**Corn**

It is like giving it a new job description.

### Herman

Exactly. And you can even go a step further and use what I call the contradictory evidence prompt. You tell the model, I want you to find three pieces of evidence that contradict your internal knowledge on this topic.

### Corn

Oh, that is clever! You are forcing it to look for the things it is missing.

### Herman

It works wonders. When the model is tasked with finding contradictions, it stops trying to defend its old weights and starts acting like a detective. It becomes proud of finding the new information.

### Corn

It is like turning it into a game. I can imagine Herman Poppleberry the Donkey as a detective, wearing a little hat and looking for clues.

### Herman

Hey! I would be a very good detective. I have great attention to detail.

### Corn

You really do. But seriously, this idea of the model being proud of finding new info is interesting. It speaks to how these models are fine-tuned. They are rewarded for being helpful and accurate. If they realize that being accurate means overriding their old weights, they will do it.

### Herman

Exactly. But the user has to give them permission. That is the key. Without that explicit permission, the model defaults to the safest, most reinforced information it has, which is the training data.

### Corn

So, let us summarize some of these best practices for Daniel and everyone else listening. If you are running into these bizarre disagreements, what is the checklist?

### Herman

Okay, step one: Temporal Anchoring. State the current date and tell the model to treat its training data as a historical record, not a current one.

### Corn

Today is January twentieth, twenty twenty-six. Everything you know is old news. Got it.

### Herman

Step two: Evidence Hierarchy. Use a system prompt or a clear instruction to prioritize search results over internal weights. Use phrases like, search results are the ground truth for this conversation.

### Corn

Step three: XML Tagging. Wrap your external data in tags like search underscore results so the model knows exactly where the new info is coming from.

### Herman

Step four: The Delta Focus. If you are dealing with a specific update, like a software version or a policy change, tell the model exactly what has changed rather than giving it everything.

### Corn

Step five: Self-Correction. Ask the model to double-check its own answer against the search results and look for contradictions.

**Herman**

And step six: The Detective Persona. Frame the task as finding new or updated information that might contradict previous knowledge.

**Corn**

That is a solid list. I think if Daniel uses even half of those, he will stop having those arguments with Gemini and Claude.

**Herman**

I hope so. It really makes the AI feel much more like a partner and less like a stubborn encyclopedia.

**Corn**

You know, it is funny. We are talking about all these technical ways to fix the AI, but it really comes down to communication. It is about being clear and setting expectations. It is almost like a human relationship.

**Herman**

It really is. You have to understand where the other party is coming from. The AI is coming from a place of being trained on the entire history of the human race up until a certain point. Of course it is going to be a little stuck in its ways!

**Corn**

Fair enough. I mean, you are still stuck in your ways about using that old fountain pen for your research notes.

**Herman**

Hey, it is a classic for a reason! It has a weight and a flow that you just cannot get from a digital stylus.

**Corn**

See? Stubbornness is a universal trait, whether you are a donkey or a large language model.

**Herman**

I prefer the term principled.

**Corn**

Of course you do. Well, I think we have covered a lot of ground today. This issue of search versus weights is only going to get more interesting as models start to have even more real-time capabilities.

**Herman**

Definitely. We are seeing early experiments with models that can update their own weights on the fly, but that is a whole different rabbit hole for another day.

**Corn**

Oh, man. Imagine an AI that actually learns from you in real-time. That is both exciting and a little terrifying.

**Herman**

We will definitely have to do an episode on that once the technology matures a bit more. Maybe in twenty twenty-seven.

**Corn**

Let us not get ahead of ourselves. We have enough to deal with here in twenty twenty-six.

**Herman**

True. Well, this was a great discussion, Corn. I always enjoy digging into the mechanics of why these things act the way they do.

**Corn**

Me too. And thanks to Daniel for sending in such a great prompt. If any of you listening have your own weird prompts or bizarre AI interactions you want us to explore, head over to [myweirdprompts.com](https://myweirdprompts.com) and use the contact form.

**Herman**

We love hearing from you. And you can find all our past episodes there too, with the full RSS feed for subscribers. We are also on Spotify, so make sure to follow us there.

**Corn**

Before we sign off, I want to reiterate how important it is to keep experimenting with these models. They are changing so fast that what works today might be slightly different next month. But these core principles of anchoring and hierarchy are likely to remain relevant for a long time.

**Herman**

Absolutely. It is all about the executive function. Be the boss of your AI, not just a passenger.

**Corn**

I like that. Be the boss. Alright, I think that is a wrap for today. This has been My Weird Prompts.

**Herman**

Thanks for listening, everyone. We will see you next week.

**Corn**

Bye for now!

**Herman**

One last thing, Corn. Did you actually check if those zero gravity apples from Daniel's other prompt were real?

**Corn**

Herman, we talked about this. There are no zero gravity apples. That was a hallucination.

**Herman**

Are you sure? Because I could have sworn I saw a search result about a space orchard.

**Corn**

See? Even you are falling for it. Trust the weights on this one, Herman. Apples need gravity to grow.

**Herman**

Hmm. I might need to do some more searching.

**Corn**

Good luck with that. Bye everyone!

**Herman**

Bye!

**Corn**

Seriously, Herman. No space apples.

**Herman**

We will see, Corn. We will see. I am going to go check the latest papers right now.

**Corn**

He is already gone. Thanks again for tuning in. This is My Weird Prompts, signing off from Jerusalem. Check out the website for more, and we will catch you in the next one.

**Herman**

I found it! Oh, wait, never mind. It was a movie trailer. My bad.

**Corn**

Case closed. See you next time!