Subject: Almost Timely News: 🗞️ Where AI is Going in 2026 (2025-12-21)
Date: Sunday, December 21, 2025 at 6:33:27 AM Eastern Standard Time
From: Christopher S. Penn from Almost Timely Newsletter
To: Tor D. Wager

Almost Timely News: 🗞️ Where AI is Going in 2026 (2025-12-21)

Plus the ONE thing you must learn

CHRISTOPHER S. PENN
DEC 21

READ IN APP

Almost Timely News: 🗞️ Where AI is Going in 2026 (2025-12-21) :: View in Browser

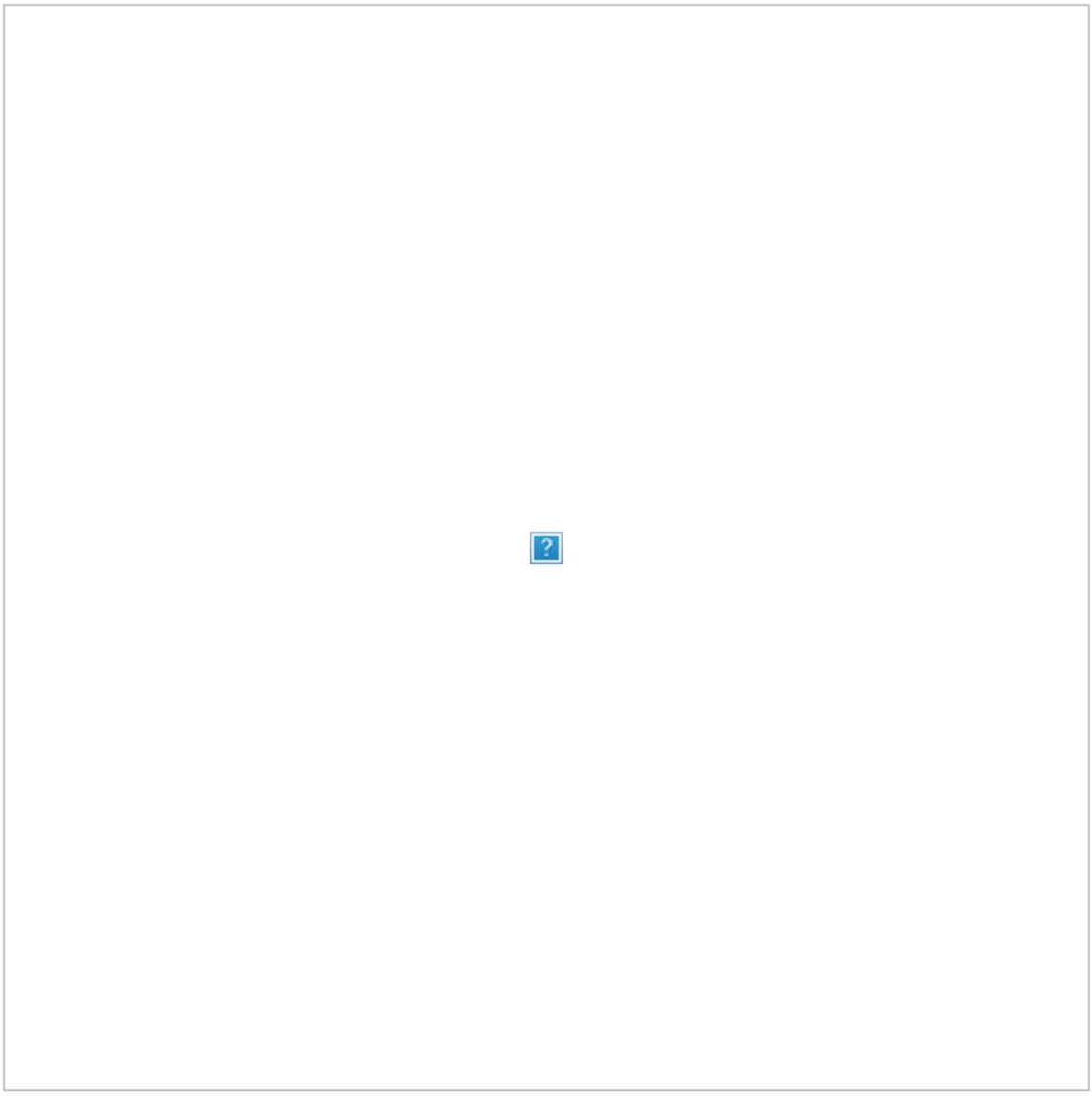
The Big Plug

🚨 ***Read new research from Trust Insights: What If LinkedIn's AI Treats Your Profile Differently Based on Your Name? Free, no info required, no forms to fill out.***

Content Authenticity Statement

100% of this week's newsletter was generated by me, the human. Learn why this kind of disclosure is a good idea and might be required for anyone doing business in any capacity with the EU in the near future.

Watch This Newsletter On YouTube 📺



Click here for the video 📺 version of this newsletter on YouTube »

Click here for an MP3 audio 🎧 only version »

What's On My Mind: Where AI is Going in 2026

Where is AI going in 2026?

Lots of people are prognosticating about what's going to happen in AI in 2026. As with so many forecasts and trends, they tend to generally be "more of whatever's happening now".

The obvious trends will continue - AI models will generally get smarter, agents will more easily connect to systems we care about, connectors will do the same, etc. AI usage will continue to increase as companies seek more and more cost savings in the short term, even when those savings are bad for the company in the long term.

None of that is a surprise and none of that really qualifies for a "look ahead"; few things make you go off the road faster than trying to drive a car by only looking out the rear view mirror.

If what was is a poor fit for what could be, then how should we be thinking about the future? How could we predict what's going to happen in the weeks and months ahead when it comes to AI?

The good news is that AI, unlike other disciplines, takes time for advancements to make their way from lab to laptop. What happened in 2025 in the labs is likely to show up on our laptops and phones in 2026. If we know what's happening in labs now, we can know what's likely to happen in production next year.

The challenge is knowing what's happening. So in this issue, I'll walk you through my thought process and creation process to answer this question. I don't have any inside information. I don't have any secrets. No one is DMing me confidential information. But with what is publicly available, I think we have enough information to predict the future.

Part 1: Where Does The Knowledge Live?

Obviously, we can't know what's going on at the most current levels in AI labs around the world because that work is quite secret - and the nature of science means that there may be lots of things that they're working on which just won't pan out.

The next best thing to being in the labs, however, is being at the forefront of the labs' outputs, and that means conference papers. The word conference papers and the word conference mean different things in different contexts. For marketers and business folks outside of academia, conferences are where you go to see a glimpse of yesterday. It's often a joke at events like affiliate marketing conferences or SEO conferences that by the time someone shows a technique on stage that works, it stopped working a year ago.

At those kinds of events, the real action is in the hallway conversations and at the bar or at the restaurant.

As much fun as events are in the marketing and business world, what's on stage is rarely novel. That's not true at academic evens like ICLR, NeurIPS, and many other top-tier academic events. At these events, stuff that's just been proven in the labs tends to make it on stage long before there's ever a commercial use case for something discovered. Some discoveries may never have a practical application but advance the knowledge of the field overall.

That's where we go to figure out what's happening in AI right now. As of late 2025, those events are ones like AAAI (Association for Advancement of Artificial Intelligence), ACL (Association of Computational Linguistics), ICLR (International Conference on Learning Representations), ICML (International Conference on Machine Learning), and NeurIPS (Conference on Neural Information Processing Systems).

These flagship academic conferences are where literally thousands of papers are published about findings in AI and machine learning labs around the

world. This is where the good stuff comes out. But how do we get a hold of it?

Part 2: How Do We Get The Knowledge?

Almost every academic event in the AI space publishes its papers openly; many are on services like arXiv.org, OpenScience, OpenReview, etc. Those papers are typically fairly dense academic treatises but contain tons of valuable information about specific techniques and technologies that the different labs are working on.

The good news? Services like OpenReview have APIs, and that means we can programmatically extract all the papers from an event.

Even better news? This is already a largely solved problem. On sites like Github, enterprising folks have already built software to extract this information for us. If you're comfortable working in the command line (one of the top skills I recommend everyone get good at if you want to take your AI capabilities to the next level), there's a good chance someone already invented the thing you want.

In this case, a person named Miroslav Lžičař built a simple Python script to download papers from NeurIPS. I liked it, and noticed it was released under the Apache 2.0 license which basically says you can take that software and do with it what you want. So I made a copy of it and customized the heck out of it to do what I wanted.

This let me download all 5,800 papers from NeurIPS.

Of course, 5,800 papers is a bit much to read.

Part 3: How Do We Get The Specific Knowledge We Want?

Miroslav's original software was simple and focused: grab all the papers from the event. For it to be useful to me, I needed it to have things like search

functionality. I loaded up the original in Google Antigravity, their newest coding tool, powered by Gemini 3, and gave Gemini 3 a laundry list of the things I wanted it to include:

- Better rate limiting
- The ability to pause and resume downloads
- A search feature with fuzzy searching
- A UI

After all was said and done, I dictated aloud my requirements to Antigravity and it made all the changes in the code, transforming it into a piece of software that suited my needs specifically. Rather than invent the wheel from scratch, I was - with Miroslav's implicit permission via the Apache license - able to take it and remake it to fit what I wanted.

Once that was built by Antigravity, I was able to find all the papers submitted by Google, Microsoft, Alibaba, etc. from their respective labs. Now, instead of 5,000 papers, I had more like 500 papers to review.

The next step, unsurprisingly, is to put all those papers into an AI system capable of not only reading them, but drawing connections among them. That system is Google NotebookLM. In the free version, you can load up to 50 data sources. In the paid version, you can load up to 300, more than enough for our purposes.

The trick is to determine what you want to know. In this case, I wanted to focus on what each lab was working on individually. So papers from Google DeepMind I'd want to look at separately than papers from Alibaba Group, and that's how I partitioned the massive corpus into manageable chunks.

Part 4: So What Does the Future Look Like?

At the broadest, highest levels, the underlying architecture of generative AI underwent huge changes in 2025. The legacy format of a dense model like

what we had in 2023 and 2024 largely fell by the wayside as companies embraced hybrid architectures. There are a bunch of fancy names for these, but the bottom line is that we've moved past the single thread, single turn AIs of yesteryear. If you have been using AI for a while and used Chat GPT in 2023 and 2024, you remember how single track, one task, one thing at a time the tool used to be.

In 2026, that is not going to be the case.

For example, in several papers from NeurIPS, Google DeepMind showed how models have even gone past tool handling into creating parallizable sub-tasks within a chain of thought. If that sentence sounds like it means absolutely nothing, imagine you're cooking in the kitchen. Maybe you're making breakfast for your family. You could, and people do, cook exactly one dish at a time.

But you only have so much time, so naturally you start to multitask. While bread is in the toaster, you can be doing something else. You can get the coffee maker brewing. While the coffee starts brewing, you can be frying up eggs. As long as you time things well and stay organized, you can have multiple dishes and courses operating in parallel.

That is what Google's paper says they were able to do in a lab. Instead of having single turn one at a time, they created AI that could truly multitask and stay coherent for all of its sub-tasks.

These kinds of advancements are what will power the next generation of AI Agents. Today's agents require a LOT of orchestration, a lot of compute power to get the individual agents working. They're slow and error prone as a result. Discoveries like this kind of parallel execution framework means agents will be smarter, faster, and use fewer resources.

Another major advancement from Google is the Titans model family, a new kind of AI that overcomes many of the issues with today's models, most

critically that of long term memory or lack thereof. The paper showcases how models can not only learn, but remember very large context, like a friend who can remember key details after months of conversations.

In a different set of papers from Alibaba Group, the makers of the Qwen models, show that reinforcement learning using checklist feedback enables better instruction following, particularly for complex tasks with multiple constraints.

All AI models are trained to be helpful, but often their training for good and bad responses is unstructured and generic. Alibaba's paper shows that you can get better performance out of AI by having it learn how to read and write checklists as its native form of checking its own work, rather than plain prose, to enable better compliance.

No surprise that Alibaba will probably build their models to do this particular type of checking its own work. If they train their models that way, their models will behave better than Meta's models that don't do that.

In a separate paper, they demonstrate how to generate realistic background sounds using a set of prompts as part of integrated video and audio generation. You can guess how that research is going to be used.

One of the key things to remember about conference papers like this is that all these papers are open science, open findings. Many of the papers are accompanied by code and data, which in turn means that if one lab discovers something highly effective, other labs will inevitably take those findings and apply them to their own models. Back in January of this year, researchers carefully examined the papers around DeepSeek's R1 model and some of its advancements, like GRPO, became standard procedure for similar types of AI at a bunch of different model makers.

What does this mean for AI in 2026? Despite rumors and opinions to the contrary (I love those LinkedIn hot takes about how AI has reached its limits),

today's research labs are demonstrating that we still haven't really scratched the surface of these technologies. Advancements like MIRAS and Titans, complex math like we see in Qwen, and thousands of other research discoveries point to us just barely beginning to make this technology powerful. We are in the earliest days of this technology, still three years later.

What you can expect in 2026 are dramatically faster, smarter, and more powerful AI models, and lots more in the agent space. I said in a recent keynote that if you have to place a bet in one area of AI for 2026, place that bet in learning how to make agents work for you. It's going to be THE skill to learn, especially as the underlying models get even smarter at working with agents.

Part 5: Other Benefits Of This Data

Outside of the big labs, there were still thousands and thousand of other papers from academic institutions, smaller labs, independent researchers, and more. Conferences often break up papers into 4 major categories - accepted, meaning a paper was reviewed and found to be valid. Spotlight, meaning a paper is worth paying attention to and was deemed a cut above. Oral means the paper authors are invited to give a talk about the paper. Given the limited number of speaking slots at a major event, a paper that earns stage time is considered fairly important. The final bar is award - papers that win awards, a handful out of thousands.

Here's the thing - by nature, we pay attention to the top tiers, oral and award. For sure, we want to read those papers and understand them deeply. Alibaba's Qwen Group's most recent paper on sigmoid gates as a way to stop the attention clogs at the beginning of a model won an award. Now, is that something that you're going to use in your day to day? No, but if you use Alibaba's models, you're gonna see dramatically better results for certain types of tasks than you will see out of, say, ChatGPT.

But ANY paper accepted for a major academic conference is worth having on

hand, even if it's not flashy.

Suppose there's a practical part of AI you want to focus on, like how to prompt better. A big lab like Alibaba or Google probably isn't going to have a paper on prompt engineering because they're well past that in terms of what they're working on. But other labs and researchers are still working on how to squeeze performance out of the AI in the field today. At NeurIPS alone, there were a couple hundred papers on prompting and prompt engineering that were accepted.

And those papers are where you'll find useful techniques for yourself, for practical use cases, rather than cutting edge research. Not everyone is or needs to be on the cutting edge; solid research with abundant, peer-reviewed data about the basics is just as valuable, if not even more valuable, for the day to day person who wants AI to work better.

For example, a paper on Role Identity Activation from Baidu in China showed how AI models often get distracted by task-specific instructions, forgetting what role they're supposed to play. The paper explains how to use role identity activation techniques to remind the model of WHO it is supposed to be and not just what it's supposed to be doing, enabling richer roleplay. How is this practical? For those folks who are using things like Ideal Customer Profiles and Buyer Personas with AI, this paper would be a gold mine for making those techniques work better. Is it cutting edge tech? No. Is it valuable? Yes.

The bottom line here is pay attention to any accepted paper at one of these major academic conferences because there may be hidden nuggets that are going to be helpful for your specific use cases of AI.

Part 6: Wrapping Up

As we look forward to the year ahead, you'll see and hear a lot of navel gazing based on... well, mostly opinion. If you really want to know what's

going to happen with AI, I've given you a rich blueprint in this newsletter to help you find the advancements happening now that will show up later.

When they find their way into commercial products like ChatGPT in the weeks and months to come, you'll have had advance notice of them and will be ready to use them on day 1 to their maximum effect.

There is so much going on in the AI research space right now that it's difficult to keep up, but there's a lot of value to be had from all of these different papers.

When people ask me, "What conferences do you go to?" or "Who do you follow to learn about AI?" I often have to say, yeah, I don't. Not really. Because a lot of the people that you and I know in this space, popular names and things, they're not researchers, they're not PhDs, they're not hardcore coders, they're not the people at these labs who are building the technology of the future.

And so while I care about and respect and enjoy the takes of my friends and colleagues in the space, I don't follow their work because their work is based on what is in market now and not where the market is going.

I would encourage you as well to take a peek under the hood at what's happening at some of these labs. Go to their engineering pages, go to their research portals. With today's generative AI tools, there is absolutely no excuse for any of us to not know what's going on. Because even if you don't understand what is a sigmoid gate put before a scaled dot-product attention matrix, AI can explain it to you.

AI can explain it to you like pizza. In that particular example, pizzas are coming off the line and no one's quality checking them. So you put a quality checking chef at the end of the line saying, "Why did you make 8 pepperoni pizzas when the person only ordered 1? Try again." That's essentially what a sigmoid gate does for this particular use case.

We have so much knowledge, good knowledge, high quality knowledge we could be taking advantage of. It's out there. It's waiting for you to use it.

How Was This Issue?

Rate this week's newsletter issue with a single click/tap. Your feedback over time helps me figure out what content to create for you.

- 😄 Good
- 😐 Neutral
- 😦 Bad

Here's The Unsubscribe

It took me a while to find a convenient way to link it up, but here's how to get to the unsubscribe.

[](https://almosttimely.substack.com/action/disable_email?utm_source=almost-timely-newsletter&utm_medium=email&utm_campaign=almost-timely-2024- 12-21)

If you don't see anything, here's the text link to copy and paste:

https://almosttimely.substack.com/action/disable_email

Share With a Friend or Colleague

If you enjoy this newsletter and want to share it with a friend/colleague, please do. Send this URL to your friend/colleague:

https://www.christopherspenn.com/newsletter

For enrolled subscribers on Substack, there are referral rewards if you refer 100, 200, or 300 other readers. Visit the Leaderboard here.

Advertisement: The Unofficial LinkedIn Algorithm Guide

If you're wondering whether the LinkedIn 'algorithm' has changed, the entire system has changed.

I refreshed the Trust Insights Unofficial LinkedIn Algorithm Guide with the latest technical papers, blog posts, and data from LinkedIn Engineering.

The big news is that not only has the system changed since our last version of the paper (back in May), it's changed MASSIVELY. It behaves very differently now because there's all new technology under the hood that's very clever but focuses much more heavily on relevance than recency, courtesy of a custom-tuned LLM under the hood.

In the updated guide, you'll learn what the system is, how it works, and most important, what you should do with your profile, content, and engagement to align with the technical aspects of the system, derived from LinkedIn's own engineering content.

👉 Here's where to get it, free of financial cost (but with a form fill)

12 Days of AI Use Cases

The annual series has debuted on the Trust Insights blog!

- 12 Days of AI Use Cases Day 1: Documentation Knowledge Blocks
- 12 Days of AI Use Cases Day 2: Sales Lead Scoring
- 12 Days of AI Use Cases Day 3: Product Feedback
- 12 Days of AI Use Cases Day 4: Landing Page Optimization
- 12 Days of AI Use Cases Day 5: Quarterly Reviews
- 12 Days of AI Use Cases Day 6: RFP Requirements

- 12 Days of AI Use Cases Day 7: Making a Video Game
- 12 Days of AI Use Cases Day 8: Business Card Extraction
- 12 Days of AI Use Cases Day 9: Content Performance
- 12 Days of AI Use Cases Day 10: Press Coverage Analysis
- 12 Days of AI Use Cases Day 11: Google Ad Campaigns
- 12 Days of AI Use Cases Day 12: Generating a Holiday Song

ICYMI: In Case You Missed It

Here's content from the last week in case things fell through the cracks:

- Never Run Out of LinkedIn Content Ideas Again: A Simple AI-Powered Strategy
- How to Stand Out When Everyone Has AI in Their Toolbox
- Discover Why MarketingProfs B2B Forum Feels Like Home for Marketing Professionals
- Why Passive Voice Kills Clarity and How AI Can Save Your Writing
- Unlocking Disney-Level Event Success: Why MAICON's Small Touches Create Magic
- Almost Timely News: 🔪 How to Update Old Content With AI (2025-12-07)
- In-Ear Insights: 2025 Year In Review
- INBOX INSIGHTS, December 17, 2025: What We Learned About AI, Humanity at Risk
- So What? How to make your own 2025 wrap-up

On The Tubes

Here's what debuted on my YouTube channel this week:

- So What? How to Make Your Own 2025 Wrap Up
- You Ask, I Answer: How to Detect AI Generated Content?
- You Ask, I Answer: How to Stop AI Writing Cliches?
- You Ask, I Answer: Will AI Replace My Job?
- You Ask, I Answer: Why You Need an Owned Audience?
- Vesper of the Dry Ravine - AI Generated Music
- Mind Readings: The Risks Of Cognitive De-skilling
- Mind Readings: Stop Reinventing The Wheel With AI

Skill Up With Classes

These are just a few of the classes I have available over at the Trust Insights website that you can take.

Premium

- 🧭 **New! The AI-Ready Strategist**
- 💥 Generative AI Use Cases for Marketers
- 👉 Mastering Prompt Engineering for Marketers
- 🦾 Generative AI for Marketers
- 👉 Google Analytics 4 for Marketers
- 👉 Google Search Console for Marketers (🚨 just updated with AI SEO stuff! 🚨)

Free

- 👉 ***New! From Text to Video in Seconds, a session on AI video generation!***
- Never Think Alone: How AI Has Changed Marketing Forever (AMA 2025)

- Generative AI for Tourism and Destination Marketing
- Powering Up Your LinkedIn Profile (For Job Hunters) 2023 Edition
- Building the Data-Driven, AI-Powered Customer Journey for Retail and Ecommerce, 2024 Edition
- The Marketing Singularity: How Generative AI Means the End of Marketing As We Knew It

Advertisement: New AI Book!

In Almost Timeless, generative AI expert Christopher Penn provides the definitive playbook. Drawing on 18 months of in-the-trenches work and insights from thousands of real-world questions, Penn distills the noise into 48 foundational principles—durable mental models that give you a more permanent, strategic understanding of this transformative technology.

In this book, you will learn to:

- **Master the Machine**: Finally understand why AI acts like a "brilliant but forgetful intern" and turn its quirks into your greatest strength.
- **Deploy the Playbook**: Move from theory to practice with frameworks for driving real, measurable business value with AI.
- **Secure Your Human Advantage**: Discover why your creativity, judgment, and ethics are more valuable than ever—and how to leverage them to win.

Stop feeling overwhelmed. Start leading with confidence. By the time you finish Almost Timeless, you won't just know what to do; you will understand why you are doing it. And in an age of constant change, that understanding is the only real competitive advantage.

👉 Order your copy of Almost Timeless: 48 Foundation Principles of Generative AI today!

Get Back to Work

Folks who post jobs in the free Analytics for Marketers Slack community may have those jobs shared here, too. If you're looking for work, check out these recent open positions, and check out the Slack group for the comprehensive list.

- B2b Marketing Analytics Lead at Talener
- Chief Marketing Officer at AlignityX
- Director Of Growth & Marketing at Brightsmith
- Director Of Growth Marketing at Octave
- Director, Ai Marketing Analytics at Experian
- Director, Career Strategy at DELTA-T
- Head Of Marketing at Prologue
- Part-Time Head Of Digital (Seo, Paid Search & Operations Leader) at Stryde - Ecommerce Search Marketing Agency
- Senior Manager, Data Science at Hoxton Circle
- Vice President, Research & Analytics at Spectrum Science
- Vp Of Data And Ai at Solomon Page
- Vp Of Martech at Harnham

Advertisement: New AI Strategy Course

Almost every AI course is the same, conceptually. They show you how to prompt, how to set things up - the cooking equivalents of how to use a blender or how to cook a dish. These are foundation skills, and while they're good and important, you know what's missing from all of them? How to run a restaurant successfully. That's the big miss. We're so focused on the how that we completely lose sight of the why and the what.

This is why our new course, the AI-Ready Strategist, is different. It's not a collection of prompting techniques or a set of recipes; it's about why we do things with AI. AI strategy has nothing to do with prompting or the shiny object of the day — it has everything to do with extracting value from AI and avoiding preventable disasters. This course is for everyone in a decision-making capacity because it answers the questions almost every AI hype artist ignores: Why are you even considering AI in the first place? What will you do with it? If your AI strategy is the equivalent of obsessing over blenders while your steakhouse goes out of business, this is the course to get you back on course.

👉 ***Take the course now!***

How to Stay in Touch

Let's make sure we're connected in the places it suits you best. Here's where you can find different content:

- My blog - daily videos, blog posts, and podcast episodes
- My YouTube channel - daily videos, conference talks, and all things video
- My company, Trust Insights - marketing analytics help
- My podcast, Marketing over Coffee - weekly episodes of what's worth noting in marketing
- My second podcast, In-Ear Insights - the Trust Insights weekly podcast focused on data and analytics
- On Bluesky - random personal stuff and chaos
- On LinkedIn - daily videos and news
- On Instagram - personal photos and travels
- My free Slack discussion forum, Analytics for Marketers - open conversations about marketing and analytics

Listen to my theme song as a new single:

- Apple Music
- Spotify
- YouTube

Advertisement: Ukraine 🇺🇦 Humanitarian Fund

The war to free Ukraine continues. If you'd like to support humanitarian efforts in Ukraine, the Ukrainian government has set up a special portal, United24, to help make contributing easy. The effort to free Ukraine from Russia's illegal invasion needs your ongoing support.

👉 **Donate today to the Ukraine Humanitarian Relief Fund »**

Events I'll Be At

Here are the public events where I'm speaking and attending. Say hi if you're at an event also:

- Social Media Marketing World, Anaheim, April 2026

There are also private events that aren't open to the public.

If you're an event organizer, let me help your event shine. Visit my speaking page for more details.

Can't be at an event? Stop by my private Slack group instead, Analytics for Marketers.

Required Disclosures

Events with links have purchased sponsorships in this newsletter and as a

result, I receive direct financial compensation for promoting them.

Advertisements in this newsletter have paid to be promoted, and as a result, I receive direct financial compensation for promoting them.

My company, Trust Insights, maintains business partnerships with companies including, but not limited to, IBM, Cisco Systems, Amazon, Talkwalker, MarketingProfs, MarketMuse, Agorapulse, Hubspot, Informa, Demandbase, The Marketing AI Institute, and others. While links shared from partners are not explicit endorsements, nor do they directly financially benefit Trust Insights, a commercial relationship exists for which Trust Insights may receive indirect financial benefit, and thus I may receive indirect financial benefit from them as well.

Thank You

Thanks for subscribing and reading this far. I appreciate it. As always, thank you for your support, your attention, and your kindness.

See you next week,

Christopher S. Penn