

When AI Chatbots Hallucinate, Infrastructure Pays

Written by: Max Mahdi Roozbahani, Senior Lecturer, Georgia Tech

Imagine a team takes a photo near a jobsite. In the background, a work truck on the highway flashes a red beacon.

A human understands the context: It is likely a utility vehicle. A generative artificial intelligence (AI) system might label the beacon as “police presence,” infer there was an accident, and summarize the day as an incident. That one incorrect label can appear in a daily report, a safety log, or a claim file months later, complete with a timestamp and a confident narrative that no one intended to create.

Construction firms are rapidly adopting generative AI copilots to search for and summarize project documents, emails, and schedules. The goal is speed. With tight margins and chronic labor shortages, the industry is desperate for leverage.

While construction is ready to benefit from AI, it should not treat AI outputs as gospel or rely on them for final sign off.

FLUENCY IS NOT PROOF

I teach machine learning at Georgia Tech, and my background spans computer science and civil engineering across both academia and industry. That vantage point makes one risk apparent: Teams conflate well written answers with ground truth. In construction, ground truth is what is physically installed and supported by field evidence, including photos, timestamps, and locations. But not text alone.

In finance or law, text often is the reality. A contract is the deal. But in construction, text is merely a proxy. A daily report



can be wrong. A submittal can be outdated. An invoice can be premature. A clean summary can hide a qualifier that matters. A large language model can read that record and still be wrong, because the record itself may be incorrect, incomplete, or outdated.

The danger is most significant in work that becomes invisible once covered: foundations, reinforcing steel, post-tensioning, fireproofing, and critical mechanical, electrical, and plumbing routing. These are the places where errors become catastrophic.

Consider seismic risk. A geotechnical report warns about liquefaction and recommends a specific pile count. Later, additional borings update those assumptions, and the design is revised to reflect an even higher pile count. If a generative AI assistant is asked, “Is the foundation plan compliant?” it might retrieve an older revision, omit a conditional statement, and reduce engineering nuance to a definitive “Yes.”

If that sentence influences a decision, the project drifts from engineering judgment toward automated optimism. The first time those assumptions are tested may be after a disaster.

This is not a theoretical concern.

In my research, my collaborators and I built EIDSeg, a dataset designed to assess post-earthquake damage from social media photos. The work required pixel-level labeling, a multi-phase protocol, and sustained attention to consistency, because even highly motivated human annotators disagree when the evidence is messy.

The central lesson is uncomfortable but essential: It takes significant time and rigor to make visual AI reliable even for a narrowly defined task. A live construction project is far more complex, with more ambiguity, more missing context, and higher consequences.

Multimodal AI, meaning models that interpret photos and text together, is a promising direction, but it introduces new failure modes. Context is thin in a single frame.

From the example above, a utility truck's red beacon is interpreted as police presence. Police can imply an accident. A temporary condition is mistaken for a permanent installation. A blocked egress path is missed due to a shadow.

The most dangerous pattern is not that AI makes errors. It is that AI produces errors with an authoritative tone and speed.

HOW IS AI STILL USEFUL?

A reasonable objection is clear: If humans must verify every output, what is the point of AI?

The point is leverage. Many tasks in construction are repetitive: drafting routine communications, summarizing meetings, mapping submittals, and flagging missing attachments. If AI reduces time spent on repetitive work, engineers can spend more time on field checks and verification.

AI should accelerate preparation so professionals can certify reality. This means that construction requires training, not just software. Responsible AI is now an engineering skill. Teams need to understand data provenance, audit trails, and failure modes. They also need to learn how to ask better questions.

Prompting matters. Vague questions invite vague answers. Safer prompts require citations, list assumptions, and force the system to say what it cannot confirm. These practices reduce hallucinations, but they do not eliminate them. That is precisely why human review remains essential.

Leaders should set a basic procurement standard for any AI used in compliance, safety, or claims. Every answer must cite specific sources with dates and revision identifiers. Field status answers must surface supporting photos and inspection records. When proof is missing or contradictory, the system must abstain from drawing a conclusion and flag for review.

Generative AI can make construction faster. It can help teams find information buried across folders and emails. But the built environment depends on proof, not prose. When chatbots hallucinate, infrastructure pays. 🛠️



About the Author

Max Mahdi Roozbahani is a senior lecturer teaching machine learning and natural language processing at Georgia Tech.

About the Article

This article is republished from [Construction Dive](#) online. Construction Dive is a leading industry publication operated by Industry Dive. Their business journalists spark ideas and shape agendas for 10+ million decision makers in the most competitive industries. The daily email newsletter and website cover topics such as commercial building, residential building, green building, design, deals, regulations, and more.

Any views and opinions expressed in this article may or may not reflect the views and opinions of the Construction Management Association of America (CMAA). By publishing this piece, CMAA is not expressing endorsement of the individual, the article, or their association, organization, or company.