

DHISOLVE x GOVERNMENT

NYC's AI Chatbot Told Citizens to Break the Law

How structural decomposition eliminates AI hallucinations in government — with real incidents, real costs, and a proven architectural solution.

AI in Government Market: \$3.316B (US Fed AI R&D, FY2025) ' \$2.52T worldwide AI spending (2026, Gartner)

THE PROBLEM

Why General-Purpose LLMs Fail in Government

REAL INCIDENTS

NYC MyCity chatbot told business owners to take workers' tips (violates NY Labor Law), told landlords to refuse housing vouchers (illegal since 2008), and denied cash-acceptance regulations exist (NYC requires cash since 2020). When tested by The Markup, it gave different answers to identical questions from 10 staff members. Remained active for months despite evidence.

3+

3+ laws the chatbot told citizens to violate

10/10 identical queries got different (wrong) answers

KEY PAIN POINTS

Government AI actively giving citizens illegal advice — and remaining active for months

Inconsistent responses: 10 identical queries produce 10 different wrong answers

AI translation errors in high-stakes asylum proceedings affecting people's lives

Public trust in government services permanently damaged by unreliable AI

THE SOLUTION

Structural Decomposition: Specialists Beat Generalists

Government queries span benefits eligibility, regulatory compliance, permit requirements, tax obligations, and legal rights — each governed by specific statutes, regulations, and agency policies.

A general LLM has no connection to actual law and no understanding of jurisdiction-specific regulations. NYC's MyCity chatbot demonstrated exactly what happens when you deploy a general-purpose AI for government services: it confidently gives illegal advice.

dhisolve routes each citizen query to a model trained on the relevant regulatory domain. Benefits eligibility is verified against specific statutory criteria. Permit requirements reference actual municipal codes. Every answer is linked to the specific regulation or statute that supports it.

Inconsistency is eliminated because each model draws from a single authoritative source — identical questions always get identical, legally grounded answers.

COST COMPARISON

DHISOLVE

\$0.10–\$0.50

per 1M tokens

BIG LLMS

\$10–\$60

per 1M tokens

OUTCOMES

Measurable Results

Zero illegal recommendations — every answer grounded in actual law

100%statute-backed answers with verifiable legal references

Zero data retention — privacy-first architecture

Consistent answers regardless of how or when the question is asked

ROI CASE

Government AI giving illegal advice destroys public trust and creates massive liability. Manual citizen services cost \$15–30 per interaction vs \$0.01–0.05 with AI. But the AI must be trustworthy — one incident like NYC's chatbot undoes years of digital transformation investment.

With dhisolve:

- Every answer legally grounded and traceable to specific statutes
- Consistency guaranteed: identical queries ' identical answers
- Citizen trust maintained through visible legal citations
- AI cost: \$0.01–0.05 per interaction with verified accuracy

REGULATORY COMPLIANCE

Built for Compliance, Not Bolted On

EU AI Act — risk-based classification, transparency requirements for high-risk AI

Executive Order 14110 on Safe, Secure, and Trustworthy AI (October 2023)

OMB M-24-10 guidance on federal AI use — accountability and transparency

FedRAMP requirements for cloud AI systems used in government

ACADEMIC & INDUSTRY BACKING

NYC MyCity chatbot: documented failure of monolithic AI in government services

Policy-grounded models structurally prevent illegal recommendations by design

Compound AI: each sub-model independently auditable for government compliance

Consistent routing: identical queries always produce identical results

MARKET OPPORTUNITY

AI in Government: \$3.316B (US Fed AI R&D, FY2025) ' \$2.52T worldwide AI spending (2026, Gartner)

Join the waitlist at dhisolve.com

<https://dhisolve.com>