

Gemeentelijke Chatbots Getest — an Independent Audit Finds Dutch Municipal AI Chatbots Answer Only 1 in 10 Questions Correctly

AI in the Public Sector · Good practice · Netherlands

Quality score: 33/100

Evidence strength: Observational / pre–post

Summary

An independent second-year audit of 39 Dutch municipal chatbots found only 10% of citizen questions answered correctly, with 64% wrong and identical errors repeated across ~30 municipalities sharing one 'Gem' chatbot; VNG called the tools 'not yet finished products.'

Key dimensions

Dimension	Score
Evidence of impact / measured public value	0/3
Transparency, fairness & accountability	1/3
Transferability / demonstrated replication	2/3
Scalability beyond pilot	1/3
Governance, capability & sustainability	1/3

Evaluated by C-NAPSE

Data sources

[C-NAPSE web research](#) — accessed 2026-08-13

[Independent researchers Wiep Hamstra and Jules Ernst; response from VNG \(Vereniging van Nederlandse Gemeenten\) and the Ministry of the Interior and Kingdom Relations \(BZK\)](#) — accessed 2026-08-13

[Digimonitor methodology & data](#) — accessed 2026-08-13

[Binnenlands Bestuur coverage](#) — accessed 2026-08-13

[Informatieprofessional coverage](#) — accessed 2026-08-13

Cite this — Evidoria. *Gemeentelijke Chatbots Getest — an Independent Audit Finds Dutch Municipal AI Chatbots Answer Only 1 in 10 Questions Correctly*. Persistent ID: e28e6d9f-1128-463a-816f-30ac364d9a35.