

JudgeGPT — a Nationwide Randomized Trial of Generative AI for Pakistan's Trial-Court Judges

AI in the Public Sector · Good practice · Pakistan

Quality score: 80/100

Evidence strength: Randomised controlled trial

Summary

A randomized trial gave ~1,559 Pakistani trial-court judges (about half the bench) access to JudgeGPT, a GPT-4 tool searching 129,235 rulings and laws. Trained judges' districts resolved 6.3% more cases, rulings rated better 59% of the time, and no rise in bias.

Key dimensions

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

Evaluated by C-NAPSE

Data sources

[C-NAPSE web research](#) — accessed 2026-07-30

[Pakistan trial court judiciary, with ETH Zurich, Imperial College London and the New Economic School](#) — accessed 2026-07-30

[Working paper — Courts of Tomorrow: Evidence from a Nationwide Rollout of Generative AI](#) — accessed 2026-07-30

[The Decoder — An AI system helped Pakistani judges clear massive backlogs](#) — accessed 2026-07-30

[The Friday Times — Can AI Fix Pakistan's Broken Courts?](#) — accessed 2026-07-30

Cite this — Evidoria. *JudgeGPT — a Nationwide Randomized Trial of Generative AI for Pakistan's Trial-Court Judges*. Persistent ID: acd0ea4b-66af-4908-b7c1-1cf7788e38a6.