South Korea's Election Deepfake Detection — a State-Built AI Model Behind 10,000+ Takedown Requests per Campaign

AI in the Public Sector · Good practice · South Korea

Quality score: 53/100

Evidence strength: Moderate

Summary

Under a Dec 2023 ban on realistic AI candidate content 90 days before polls, Korea's NEC sought 388 deepfake takedowns in 2024 and 10,510 in 2025; a ministry-built detector (92% accuracy in validation, self-reported) backed 10,319 requests before the June 2026 local elections.

Key dimensions

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

Evaluated by C-NAPSE

Data sources

C-NAPSE web research — accessed 2026-11-10

Ministry of the Interior and Safety; National Forensic Service; National Election Commission — accessed 2026-11-10

Kyunghyang Shinmun — Government unveils AI deepfake detection model for elections (Mar 2026) — accessed 2026-11-10

The Star / AFP — AI disinfo tests South Korean laws ahead of local elections (May 2026) — accessed 2026-11-10

Segye Ilbo — NEC deepfake takedown requests before the 3 June local elections (May 2026) — accessed 2026-11-10

Herald K — NEC files first complaint over deepfake false-information (Feb 2026) — accessed 2026-11-10

Korea JoongAng Daily — Authorities vow strict punishment for election deepfakes (2024) — accessed 2026-11-10

Cite this — Evidoria. *South Korea's Election Deepfake Detection — a State-Built AI Model Behind 10,000+ Takedown Requests per Campaign*. Persistent ID: 6e7fe826-0eaa-4be0-a12c-8532f1534007.