

# NDLOCR — Japan's National Diet Library Doubles Historical-Text Recognition Accuracy With AI

AI in the Public Sector · Good practice · Japan

Quality score: 80/100

Evidence strength: Observational / pre–post

## Summary

Japan's National Diet Library open-sourced an AI OCR engine that lifted recognition accuracy on pre-1945 texts from about 40% to over 90%, then used it to convert 2.47 million digitized books and periodicals — 223 million page images — into searchable text.

## Key dimensions

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

Evaluated by C-NAPSE

## Data sources

[C-NAPSE web research](#) — accessed 2026-08-19

[National Diet Library, with Morpho AI Solutions, Toppan Inc. and LINE Corporation](#) — accessed 2026-08-19

[NDL Lab — FY2021 OCR text conversion dataset](#) — accessed 2026-08-19

[Morpho AI Solutions — NDLOCR development announcement](#) — accessed 2026-08-19

---

Cite this — Evidoria. *NDLOCR — Japan's National Diet Library Doubles Historical-Text Recognition Accuracy With AI*. Persistent ID: 5f915873-fa73-4704-8535-410b858ee2da.