MASRI — Les corpus ouverts de reconnaissance vocale de l'Université de Malte pour une langue d'environ 500 000 locuteurs, diffusés uniquement pour un usage non commercial

Digital Inclusion · Bonne pratique · Malte

Score de qualité : 72/100

Niveau de preuve : Solides

Résumé

MASRI a créé les premiers corpus de reconnaissance vocale en maltais, pour des interfaces vocales au-delà de l'anglais : un corpus de 8 heures évalué par les pairs (25 locuteurs) a atteint environ 142 heures. La licence non commerciale freine l'usage.

Dimensions clés

Dimension	Score
Innovation level	3/5
Sustainability (ongoing)	1/1
Evaluation evidence	1/1
Demonstrated reach	0/1
Documented learning	1/1

Évalué par C-NAPSE

Sources de données

C-NAPSE web research — consulté le 2026-11-19

University of Malta — Institute of Linguistics and Language Technology — consulté le 2026-11-19

University of Malta — MASRI project page — consulté le 2026-11-19

Mena et al. (2020), MASRI-HEADSET: A Maltese Corpus for Speech Recognition, LREC 2020 (arXiv) — consulté le 2026-11-19

MASRI-Dev evaluation set on Hugging Face (MLRS) — consulté le 2026-11-19

Citer — Evidoria. *MASRI — Les corpus ouverts de reconnaissance vocale de l'Université de Malte pour une langue d'environ 500 000 locuteurs, diffusés uniquement pour un usage non commercial*. ID persistant : 20dcd978-f906-46c6-9479-e19936878200.