

Converted dictionaries

OASIS DMLex 1.0 (XML + JSON)

Slovenian lexicographic datasets from `dict-conversions`. Every dictionary is provided in **both** DMLex 1.0 serializations: `<CODE>.xml` and `<CODE>.json`. Two collections:

intrinsic/ — Core collection (faithful conversion; the project's own sources only)

All 15 converted dictionaries: DVANSL en->sl, DVFRSL fr->sl, DVITSL it->sl, DVRUSL ru->sl, DVSHSL sh->sl, DVSPSL es->sl, DVSLAN sl->en, DVSLFR sl->fr, DVSLNE sl->de, DVSLSH sl->sh, DVSLSP sl->es, DRSLAN sl->en, VSIS sl->it, LAT_AZ la->sl, KNAUR sl (monolingual encyclopedia).

extrinsic/ — Enriched collection (external resources)

KNAUR.xml + KNAUR.json — the monolingual encyclopedia re-serialized with sloWNet-derived **antonym (142)** and **synonym (2,655)** `relation`s. Each carries its sloWNet provenance (ILI / synset id) in `relation/description`, and the relation types link to the Global WordNet vocabulary via `relationType/sameAs`.

KNAUR is the ONLY resource whose external (sloWNet) enrichment is expressible as *DMLex*: DMLex 1.0 allows external `sameAs` links only on tag definitions, not on senses/entries/relations. The other enrichment layers — CLASSLA silver lemma/UPOS/MSD, per-lemma synset/ILI links, imported antonyms, and candidate scoring — are tabular and ship as Parquet in the `dist/enriched/` collection (not in this archive). `intrinsic/KNAUR.*` is the base KNAUR (no sloWNet relations); `extrinsic/KNAUR.*` is the enriched version — diff them to see the added relations.

Revision #2

Created 23 June 2026 08:27:41 by Simon Krek

Updated 23 June 2026 08:29:24 by Simon Krek