PRISMA-style flow of the WP44 mapping corpus

Identified (OpenAlex works with a CS topic, 2018–2026, snapshot 2026-10-14)
16,592,966

− Outside window (E1)
59,453

− Non-research types (E2)
3,994,472

− XPAC expansion records, not DBLP-linked (E3)
1,430,851

− No title (E4)
1,193

− Preprint duplicates merged into the published version (E5)
114,421

− Generic deposit platforms without DBLP/arXiv support (E7)
473,549

After exclusions
10,519,027

CS core — mapping corpus (DBLP ∨ primary CS ∨ CS topic ∧ arXiv cs.*)
7,110,289

CS penumbra — secondary CS topic only (sensitivity set)
3,402,475