DIARIO DIGITAL DEV / RECURSO DE LECTURA

IA para limpiar catálogos: cómo evitar falsos duplicados

Un nombre parecido no demuestra que dos registros correspondan al mismo producto. A partir del proyecto DAQU en Chile, revisamos cómo separar normalización, identidad y aprobación de cambios.

Ideas clave del artículo

- DAQU es un proyecto de desarrollo anunciado; sus objetivos no equivalen a resultados productivos medidos.

- Separa la normalización de texto, la comprobación de equivalencias y la fusión de registros.

- Evalúa falsos duplicados, casos omitidos y tiempo de revisión, conservando el historial de cambios.

Notas para tu equipo

_____

_____

_____

_____

Artículo y fuentes

https://www.digital-dev.cl/blog/ia-catalogos-duplicados-datos-chile/

CENIA: CENIA se adjudica dos proyectos FONDEF IDeA I+D 2026 para avanzar en IA aplicada a la fiscalización vial y a la calidad de los datos industriales
https://cenia.cl/2026/09/22/cenia-se-adjudica-dos-proyectos-fondef-idea-id-2026-para-avanzar-en-ia-aplicada-a-la-fiscalizacion-vial-y-a-la-calidad-de-los-datos-industriales/

OpenRefine Documentation: Clustering Methods In-depth
https://openrefine.org/docs/technical-reference/clustering-in-depth

GS1: GTIN Management Standard — Release 1.1, septiembre de 2023
https://ref.gs1.org/standards/gtin-management/