

Schema Matching

Discovering potential correspondences between attributes of different relations with various schemata.

Pitfalls of existing methods



Syntax-based



External Knowledge



Human Help

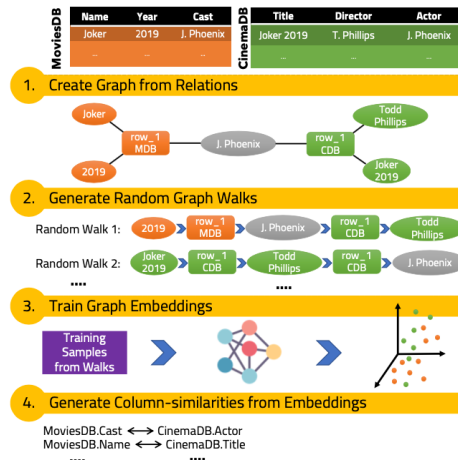
Relational Embeddings MATCHer (REMA)

An automated domain-agnostic schema matching approach, relying only on the information conveyed from the input datasets, without the need of external knowledge.

REMA's Pipeline

Relational data from several tables are transformed to a **non-directed graph**.

Node embeddings are trained, so that relevant nodes have similar vector representations.



The graph is processed to create a set of **random-walks**, which will serve as **training data**.

We use **embeddings** to calculate **similarity** between **columns** of different relations.

Work In Progress



Further configuration and enhancements for REMA



Tackle storage and scalability issues for large datasets



Dataset fabrication to enable a better evaluation of Schema Matching methods



REMA towards capturing human-understandable relationships