

https://dspace.org/

A generic framework for the matching of similiar names.

Schay, Warren Marc.

2012-01-30
http://hdl.handle.net/10539/11201

Downloaded from a DSpace repository. https://dspace.org/

Abstract

Name matching is a common requirement in modern business systems, wherein fuzzy matching techniques are employed to overcome variations between names. The purpose of this dissertation was the development of a framework, which is capable of implementing various fuzzy matching algorithms, while abstracting the name matching process away from external business systems. Through a study of existing fuzzy matching algorithms and frameworks, several design requirements were identified; the maintaining of name relationships, non-algorithm specific logic, abstraction of the matching process, user configured matching logic, consistent external interface and performance considerations. The deployment to a production environment and a series of tests, demonstrated that the framework fulfilled all but one of its design requirements, as certain algorithm implementations yielded excessive search times. The cause and remedy of this shortcoming were identified. Finally, based on an evaluation of the design's strengths and weaknesses, recommendations for future developments were suggested.