

## LinkWise: A Modern Record Linkage Software Application

Izakian, H<sup>1</sup>

<sup>1</sup>PolicyWise for Children & Families

### Introduction

Because of a lack of unique identifiers among datasets, and different data collection standards, record linkage is challenging. Thus, despite the importance of record linkage in unleashing the power of data, there are few software applications built for this purpose. Each software application has unique strengths and weaknesses.

### Objectives and Approach

Data linkage comprises various steps such as selecting linkage identifiers, data cleaning, data pre-processing, calculating the linkage weights for identifiers, and estimating similarity thresholds to decide if two records are true matches. These steps require expertise and are costly for organizations interested in data sharing. Although data linkage software applications have been developed, there are drawbacks with these applications. They are either costly, difficult to use, not able to preserve the privacy of individuals, not able to handle big datasets, or perform poorly in terms of specificity and sensitivity. LinkWise is a software application developed to resolve these issues.

### Results

LinkWise is a probabilistic modern linkage software implemented using Microsoft C#.Net. The following features are implemented in this software: automated all data linkage steps, a simple and user friendly interface, ability to link both unencrypted and encrypted data (privacy preserving record linkage), transparent linkage algorithm (not a black box), ability to perform incremental linkage (linking new data to previously linked data), ability to handle millions of records, ability to run on multiple processors to reduce run time, and high specificity and sensitivity. The software was tested over many datasets with various characteristics (e.g., different data fields, data formats, number of records, various amount of noise etc.). Results show that it is able to link data with a high specificity and sensitivity in a reasonable time.

### Conclusion/Implications

LinkWise is a software application designed to address many issues arising in the process of data linkage. The software automated all steps of data linkage and preserves the privacy of individuals. It is very easy to use and technical background knowledge is not required to work with this software.

