Method for Comparing Identity Resolution Technology

Elizabeth Schroeder, Dr. Keith Miller, Dr. Karine Megerdoomian, James Finley, Sarah McLeod, Dr. Andre Milota, Dr. Ken Samuel, Gail Hamilton, Azar Ulrich, Dr. Sherri Condon

The MITRE Corporation

{eschroeder, keith, karine, jtfinley, smcleod, amilota, samuel, ghamilton, aulrich, scondon}@mitre.org

With vast amounts of information currently available digitally, it is difficult to sift through the data and discover important details. One particularly difficult task is identifying "who is who" in different data sources and being able to recognize when the same person appears multiple times throughout. Various systems claim to offer this as an automatic capability, known as identity matching or resolution, yet it can be difficult to determine exactly how well those systems perform the reported functions. Building on previous work in evaluating and comparing name and record matching tools, this study addresses the problems involved with the evaluation of identity matching and resolution technologies. To develop an objective means of comparing alternative identity resolution technologies, this research focuses on determining appropriate evaluation metrics. The research also involves creating an infrastructure and developing test data that can be brought to bear on performing such evaluations. The evaluation infrastructure is extensible to multiple use cases, such as sharing data across agencies or resolving identities in a single data source. As a result, a user is able to define their mission, run each identity resolution system using applicable test data, and then calculate the evaluation metrics for each system to determine which produced the best score.