Data Quality
A Record Linkage Toolbox

Mohamed G. Elfeky
Ahmed K. Elmagarmid
Vassilis S. Verykios

Record Linkage
Definition
To determine whether two records represent the same real-world entity

“Not a new problem, it’s similar to duplicate elimination”

Real-World Data is dirty
TAILOR: Record Linkage Toolbox
System Architecture

Data Quality Framework

(User (Data Quality Language) DQL)
(GUI (Graphical User Interface))