

[P5-37] 時間横断的同一性判定のための 機械学習方式

小山 聡(北大) / 鹿島 久嗣(東大)

我々は、異なる時点で観測されたデータを
名寄せするための方法を考えます

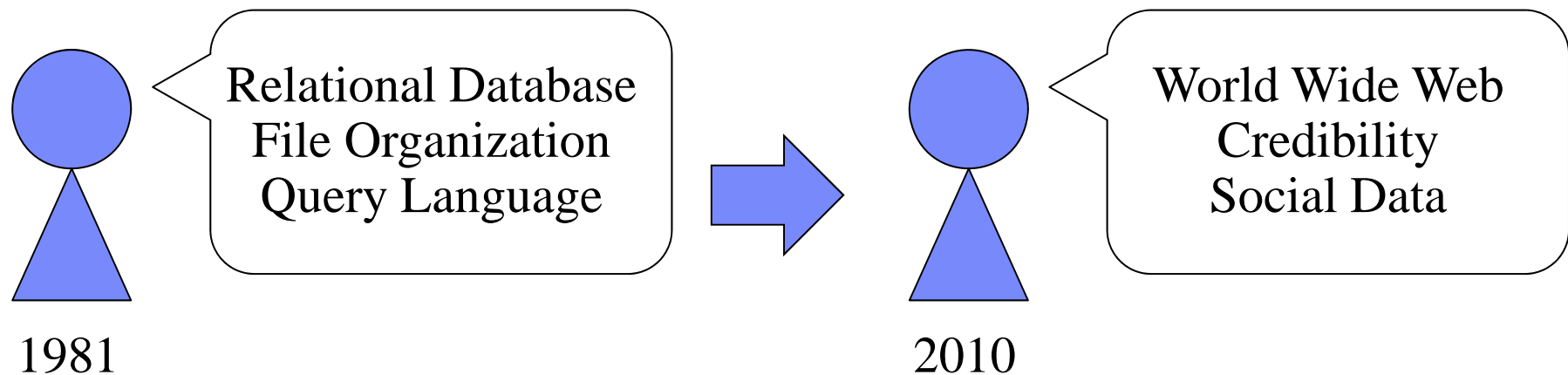
K. Tanaka: Web Information Credibility. WAIM 2010

同一人物？

K. Tanaka, Y. Kambayashi: Logical Integration of Locally Independent
Relational Databases into a Distributed Database. VLDB 1981

「局所保存射影」を時間を含む形に拡張することで 属性の時間変化の問題を解決します

- 時間の経過とともに属性が変化すると、ベクトルを直接比較しての名寄せが困難になります



- 異なる時点でのデータを、共通の低次元空間に射影することで、比較を可能にします
- 今回は、局所保存射影を時間を含む形に拡張し、各時点での射影ベクトルを学習する方法を提案します
- 行列の一般化固有値問題を解くことで、複数の射影ベクトルを一度に求めることができます