# 中川裕志 @ 学際情報学府 の研究テーマ

● 将来の社会、経済、文化、産業の発展の決め手はビッグデータの利活用にあります。ビッグデータのうちでもビジネス的価値が高いものの多くはパーソナルデータと呼ばれるものです。この中には、個人の鉄道や車による移動履歴、購買履歴、さらには医療データなどが含まれます。交通システムの設計、販売戦略、商品推薦、さらには在宅医療などでの活用も視野に入ってきます。

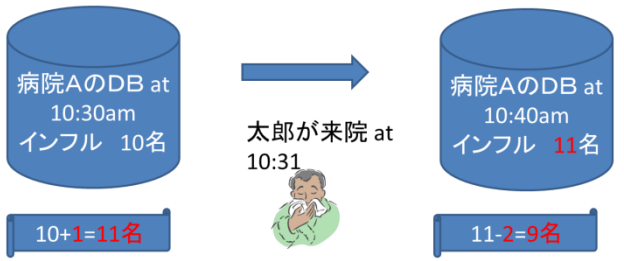
● しかし、パーソナルデータは、個人が特定される形で流出してしまうと大きな被害が予想されます。例えば、病歴や行動履歴、商品の購買履歴などが流出してしまうと個人の社会生活に悪影響が及びかねません。そこで、元データから氏名などの個人情報を消去して匿名化データにするのですが、それだけでは安全とは言い切れません。例えば、長期間にわたる行動履歴が知られると、個人をユニークに特定されかねません。こういった危険性を防ぐ技術としてK-匿名化、暗号化、雑音加算などのプライバシーを守る手法が開発されてきています。

● 中川研究室では、ビッグデータからのデータマイニングにおけるプライバシー保護の方策として、K-匿名化や差分プライバシーなどの数理モデルと機械学習手法とシステム設計、法制度との関連性、などの研究を行います。

● 研究テーマの一例として、差分プライバシーを説明しましょう。右の図にあるようにデータベース内容が変化したときに、その変化を見破られないように、質問への応答に雑音を加算する方法です。我々の研究室では数学的なモデル化、機械学習との組み合わせなどを研究しています。

● 実は機械学習、データマイニングの研究もしています。

## 差分プライバシーの有効なケース




DBにインフル患者数を質問し、10:30am で10，10:40am で11と分かると、太郎の来院を知っていた人は太郎がインフルだと分かってしまう。→ 質問への結果に雑音加算すると回避できる。