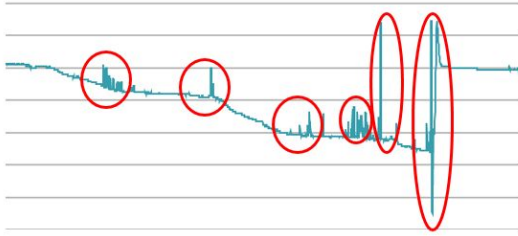# Uncertainty in Data: A Survey on Probabilistic Databases

Desh Raj & Kanhaiya Rathi

The need to store Uncertain Data and Levels of
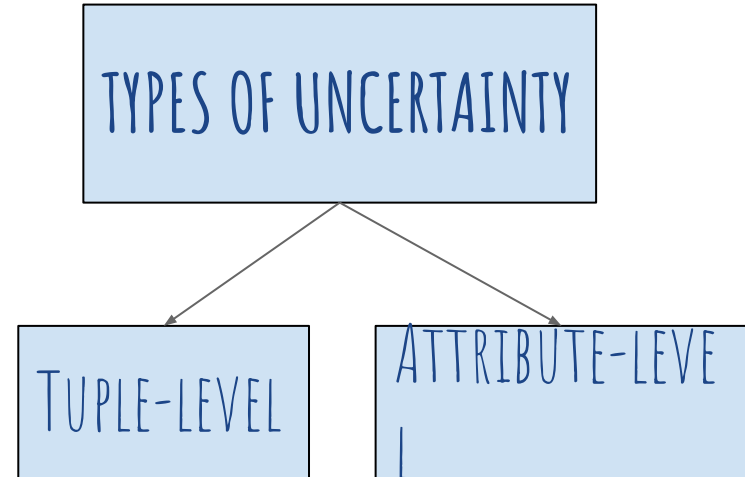

Noisy Sensor Data

Information Retrieval

Security and Privacy

TYPES OF UNCERTAINTY

Tuple-level

Attribute-level

What consists a Query and How is it Processed?

- Projection
- Selection
- Join
- Pooling
- Extension
- Threshold

Probabilistic Inference and Lineage Evaluation – Together vs Separate

Database **Conditioning** is done to make probability computation **faster**.

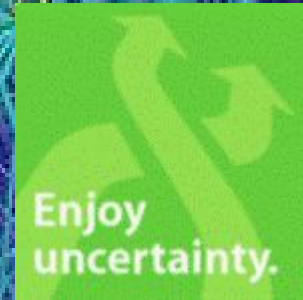By **removing branches** at every step.

- Monte Carlo Simulation
- David Putnam Method

Probabilistic databases are not popular!

Scaling has not been successfully implemented.

MayBMS (Cornell, Oxford) and Trio (Stanford) are only prototypes.

THANK YOU