The Problem

Practical Motivation

Analysts need:
- **Information-on-demand**: efficient tracking of business outcomes and access to high-quality complete business information.
- **Agility**: ability to react to **real-time changes** - business policies, objectives and processes - in a timely fashion.
- Current BI tools mostly handle data integration and cleansing of IT-supplied data in isolation from any form of business analysis and interpretation, and ignoring incomplete and inconsistent data.
- In practice, there is a continuous social discovery process between IT and business users where they routinely introduce assumptions to coerce data into a form that is complete and adequate for business querying.

One-size-fits-all interpretations of data are rarely sufficient for business analysts.

Example

Proposed Solution

- Support on demand business-driven integration of data.
- Use a business model to capture what information is needed and, in what form.
- Use a semi-automatic schema mapping tool to turn raw data into complete high-quality data, specifically aligned with a particular business application.
- Openly record vivification assumptions and strategies for dealing with incompleteness and uncertainty in business mapping rules.
- Build a framework for resolving incompleteness and inconsistency by repairing the data or the mappings.

Types of Vivification Rules

1. **Entity Identification Rules**: “patient name, dob, address”.
2. **Missing Value Rules**: “assume all missing values on key positions are distinct for counting purposes”.
3. **Conflict Resolution Rules**: “count a patient only once even though he might have seen a cardiologist and a nephrologist on the same date”.
4. **Exception Rules**: “assume two patients with distinct Patient Hospital Ids are the same if their phone numbers match”.

References