



# Leveraging Big Data, Linked (Open) Data and (Multilingual) Language Technology Standards in Global Enterprises



**MultilingualWeb-LT Workshop**  
**June 11-13, 2012**  
**Dublin, Ireland**

**Prof. Dr. Jörg Schütz**  
**biloom group**

# Today's Agenda

---

## Overview

- *Application scenario*
- *Needs and Requirements*

## Challenges

- *Data, Processes and Workflows*
- *Tools, Curation and Management*

## Solutions

- *Standards*
- *Existing Gaps*
- *Proprietary Solution*
- *Future(s)*

# Enterprise related Data

RDB  
Data  
(internal)

Multiple  
Data  
Streams  
(external)

## Data and Meaning

Semantics

Metadata

## Storage / Tools

Vocabularies

Processing

# Needs and Requirements

---

## Social Media

- Identify
- Extract
- Analyze
- Categorize
- Channel

## Core Language

- Assure Quality
- Curate Rules, Styles, Meta and Vocabularies
- Monitor and Optimize

## Corporate Languages

- Assure Quality
- Curate Rules, Styles, Meta, TMs and Vocabularies
- Monitor and Optimize

**Dynamic Streams**

**Multiple Languages**

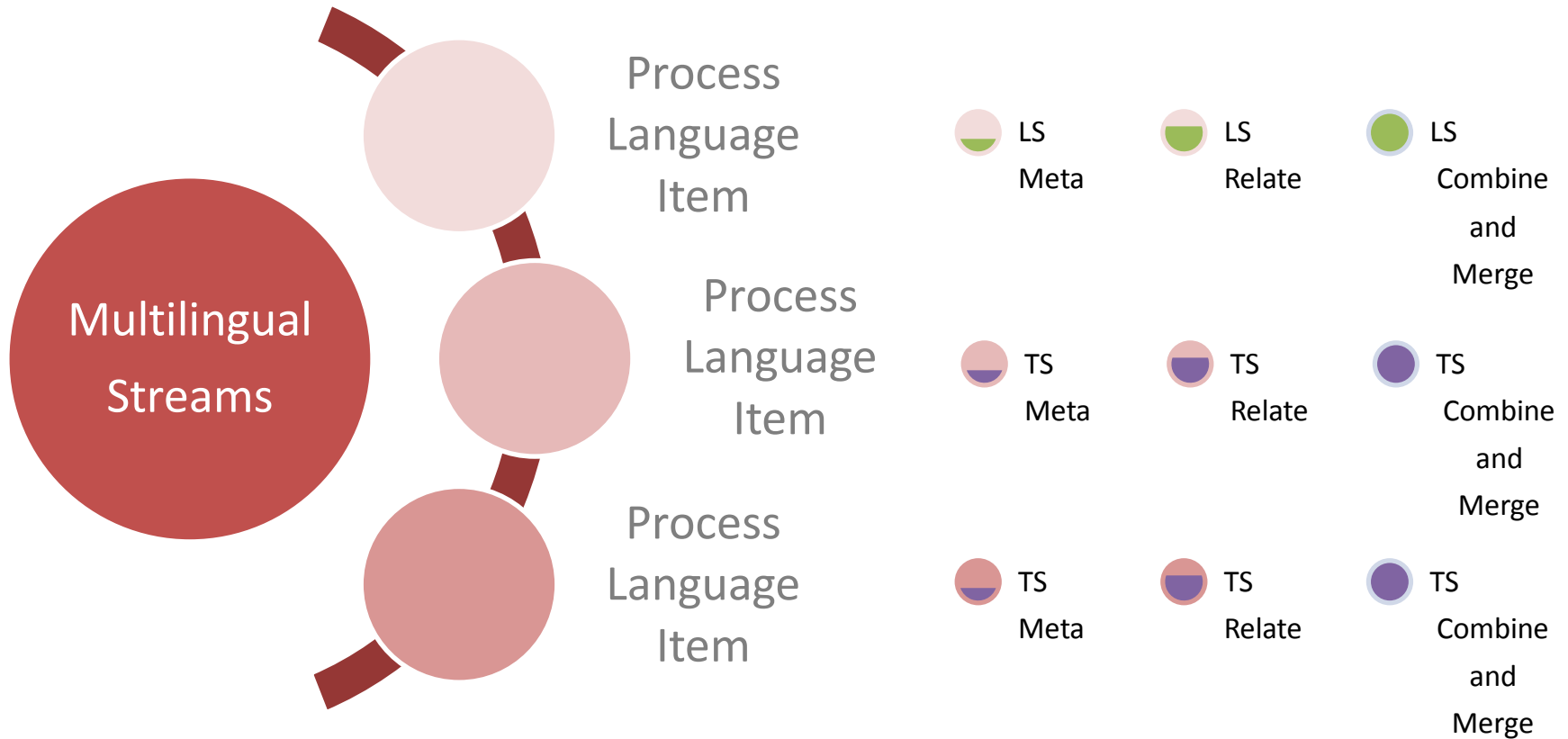


**Entailed Knowledge**

**Relate to Curated Repositories**

# Data, Processes and Workflows

---



# *Tools, Curation and Management*

---

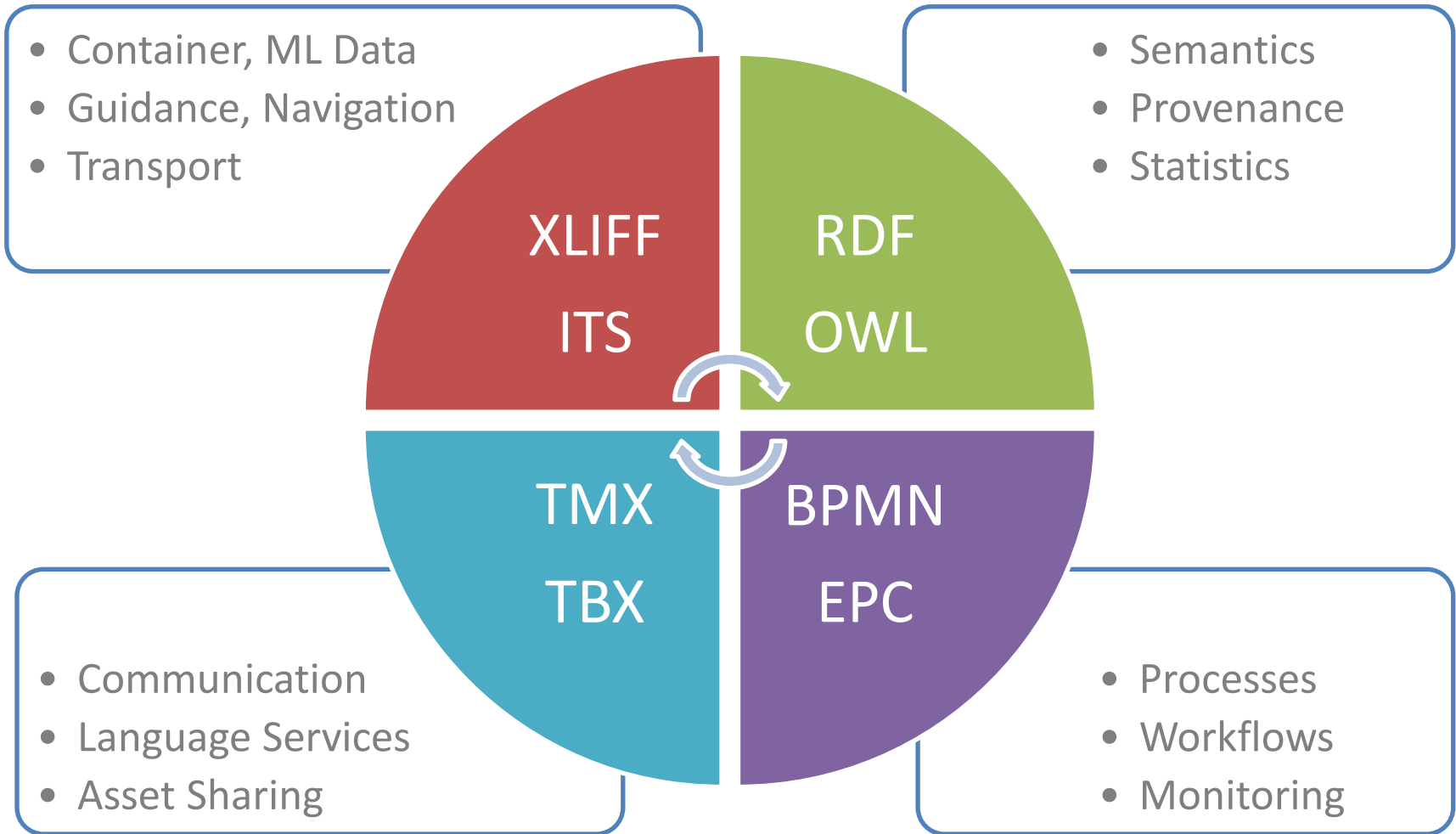
## Corporate Communication

- Relational Database Systems
- Content Management
- BPM (BPN, EPC, ...)
- IT Governance and Compliance
- ...

## Language Technologies

- Terminology Database Systems
- Translation Memories
- Checking Tools
- Crawler, Parser, ..., Machine Translation
- ...

# Combining Standards



## *Existing Gaps ...*

---

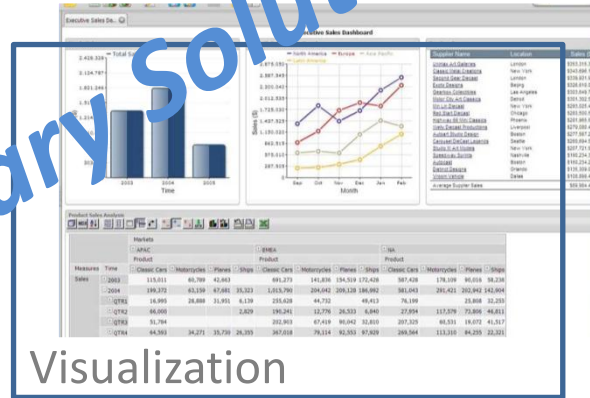
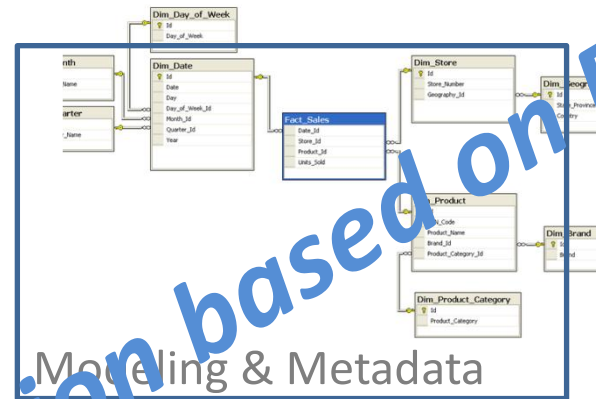
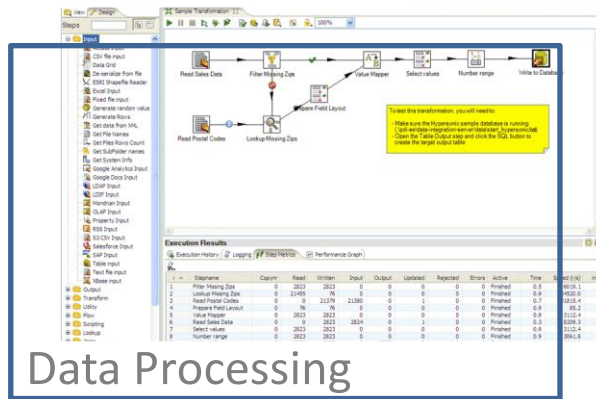
Too complex, not intuitive, not fitting, ...

Standards (partly) immature, too leaky, too flexible,  
too error-prone, too non-existent, ...

General lack of interoperability



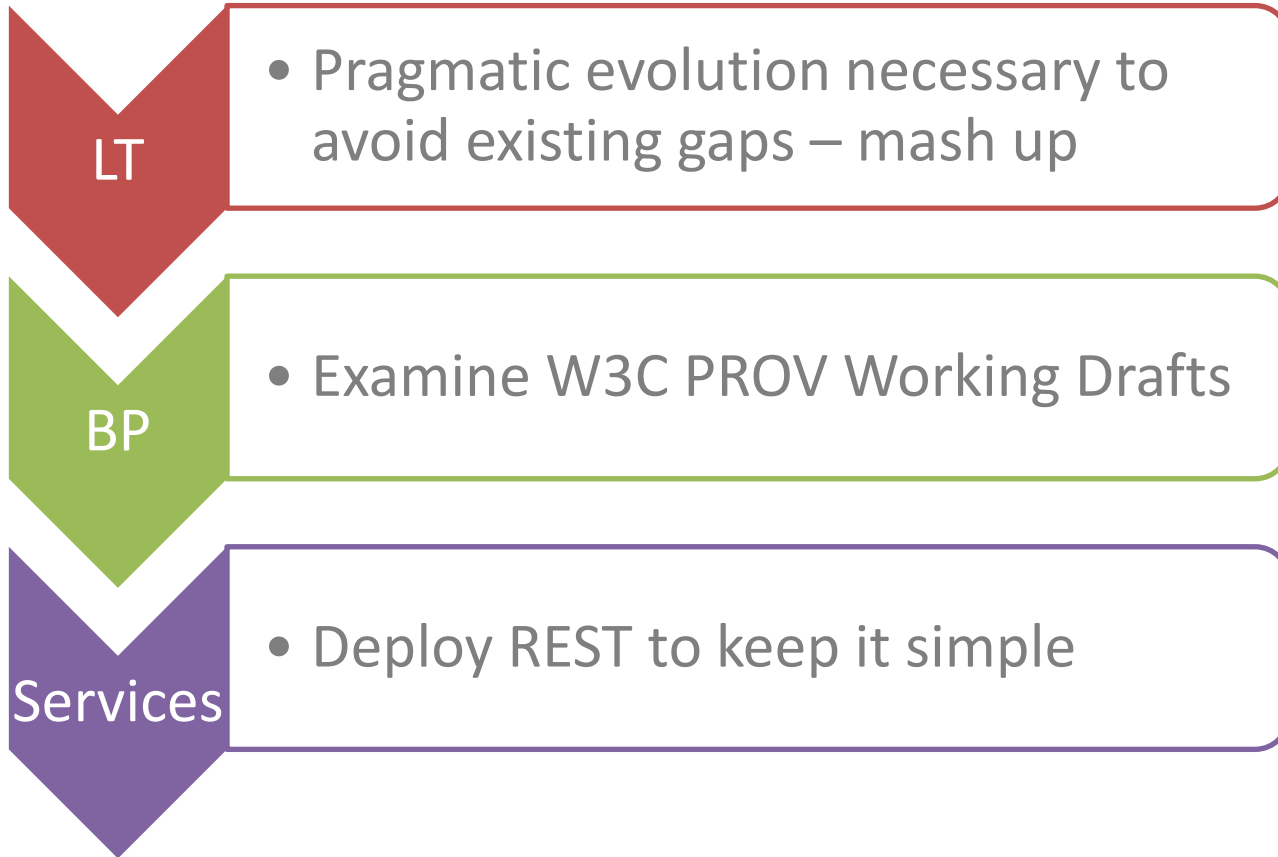
# How everything fits together



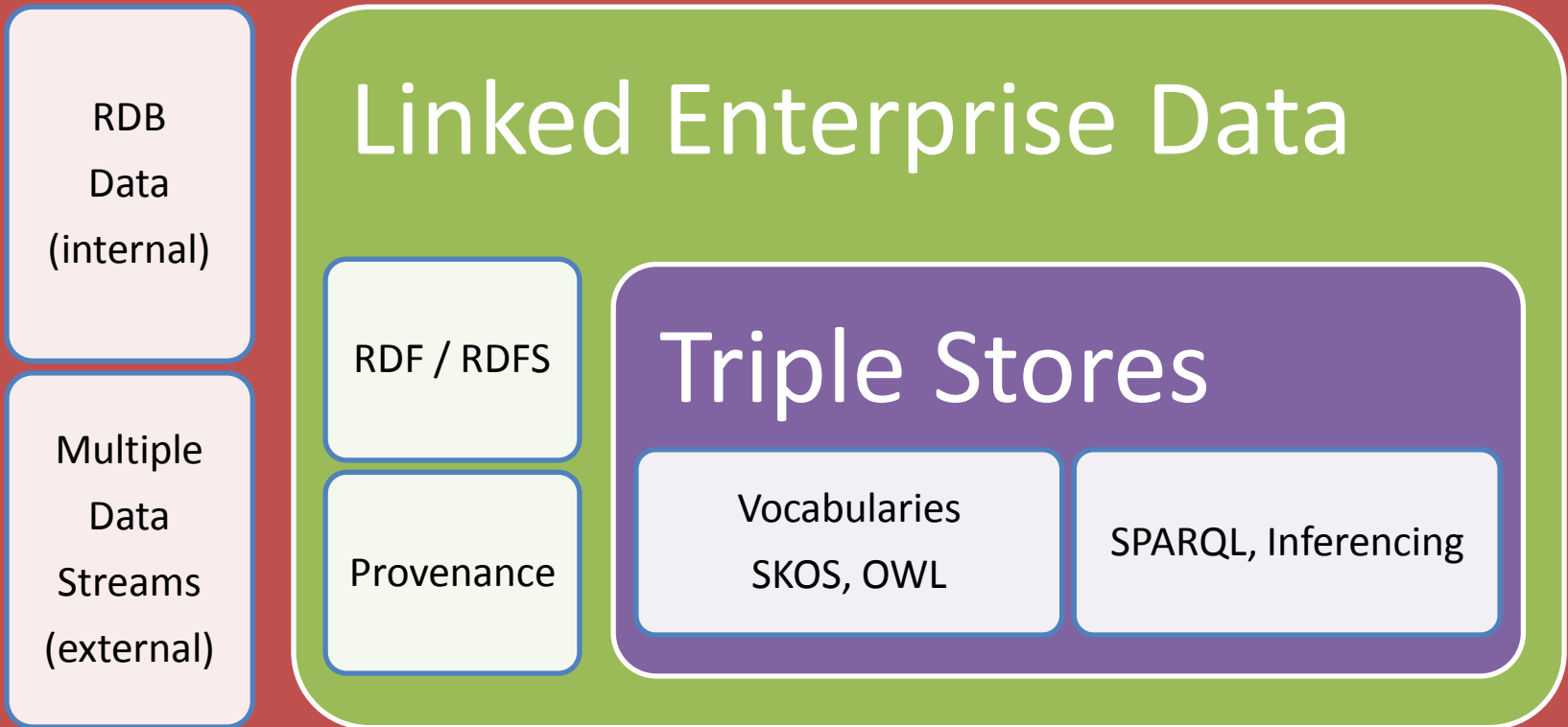
Proprietary Solution based on RDF

# *Extend / Refit or Mash up ?*

---



# Enterprise Data Web



# *Recommendations*

---

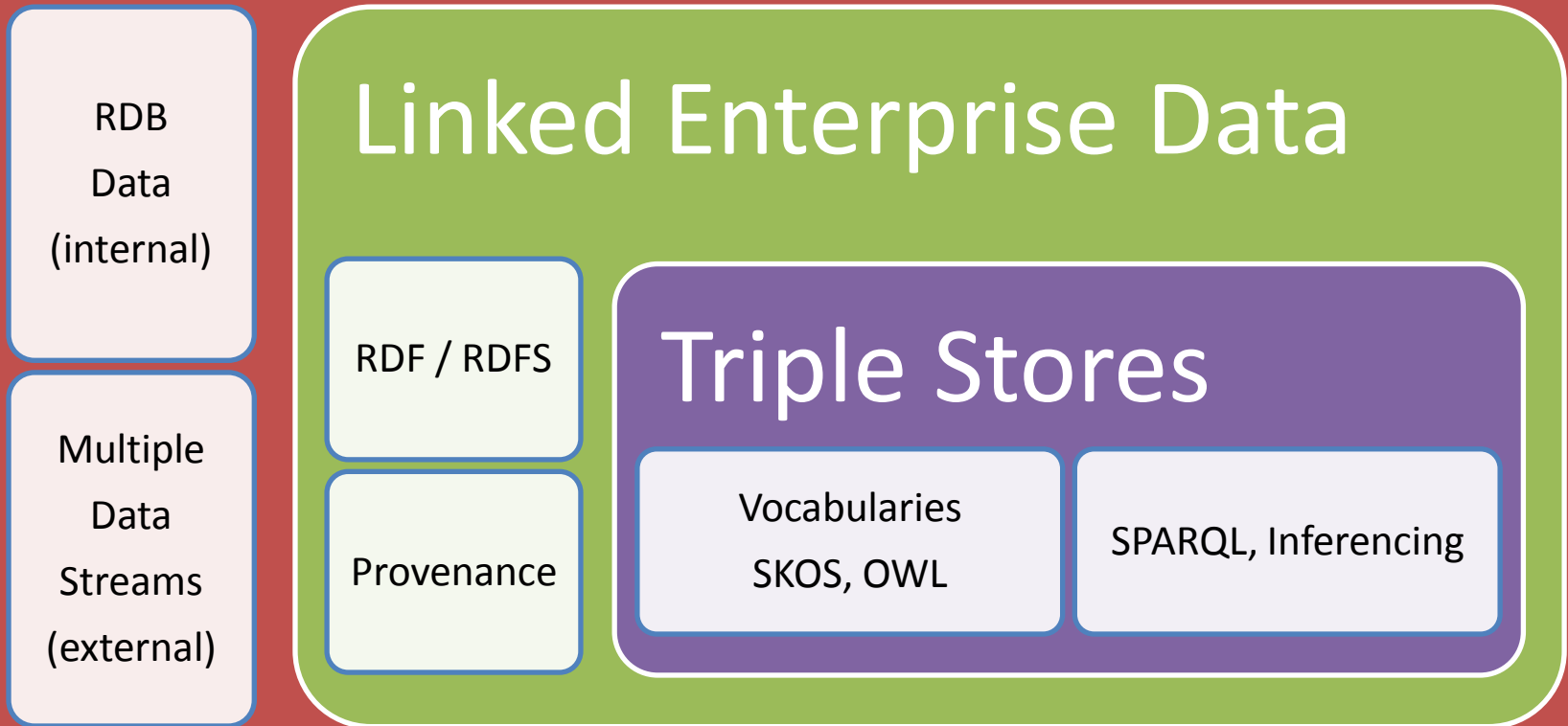
KISS Principle

Re-use and Combine

Show how it works

# World Wide Web

## Enterprise Data Web



## HOW STANDARDS PROLIFERATE:

(SEE: A/C CHARGERS, CHARACTER ENCODINGS, INSTANT MESSAGING, ETC.)

SITUATION:  
THERE ARE  
14 COMPETING  
STANDARDS.

14?! RIDICULOUS!  
WE NEED TO DEVELOP  
ONE UNIVERSAL STANDARD  
THAT COVERS EVERYONE'S  
USE CASES.



SOON:

SITUATION:  
THERE ARE  
15 COMPETING  
STANDARDS.

*Thanks for listening!*

*Additional Info: [joerg@bioloom.de](mailto:joerg@bioloom.de)*