

# Edmonton Data Management Meetup

201702

YEG Data Management Project Kickoff



# Who We Are

## **Rob Kellington**

Long time data geek with lots of experience in “data management” projects

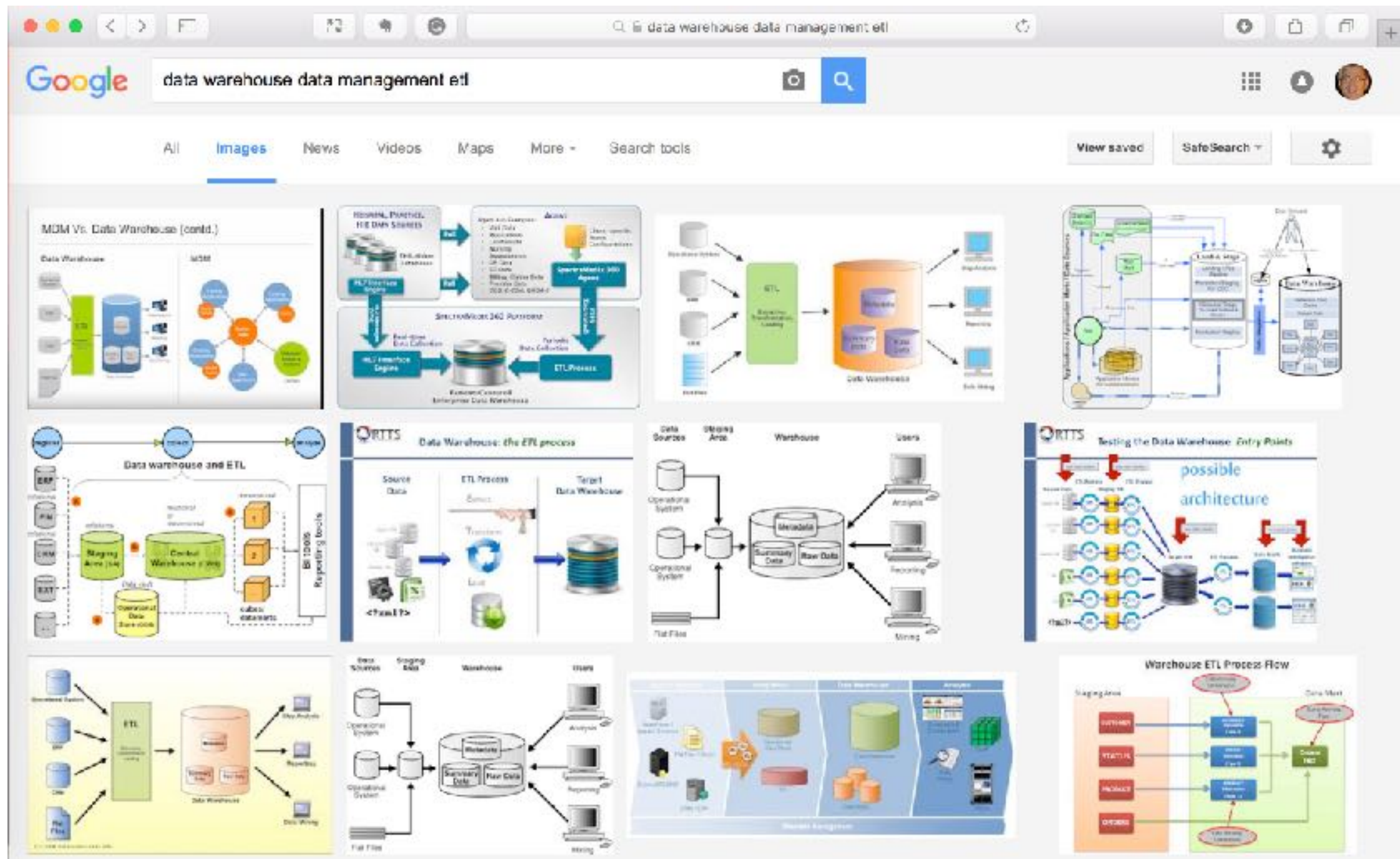
## **Alec Forest**

Computer Science Student

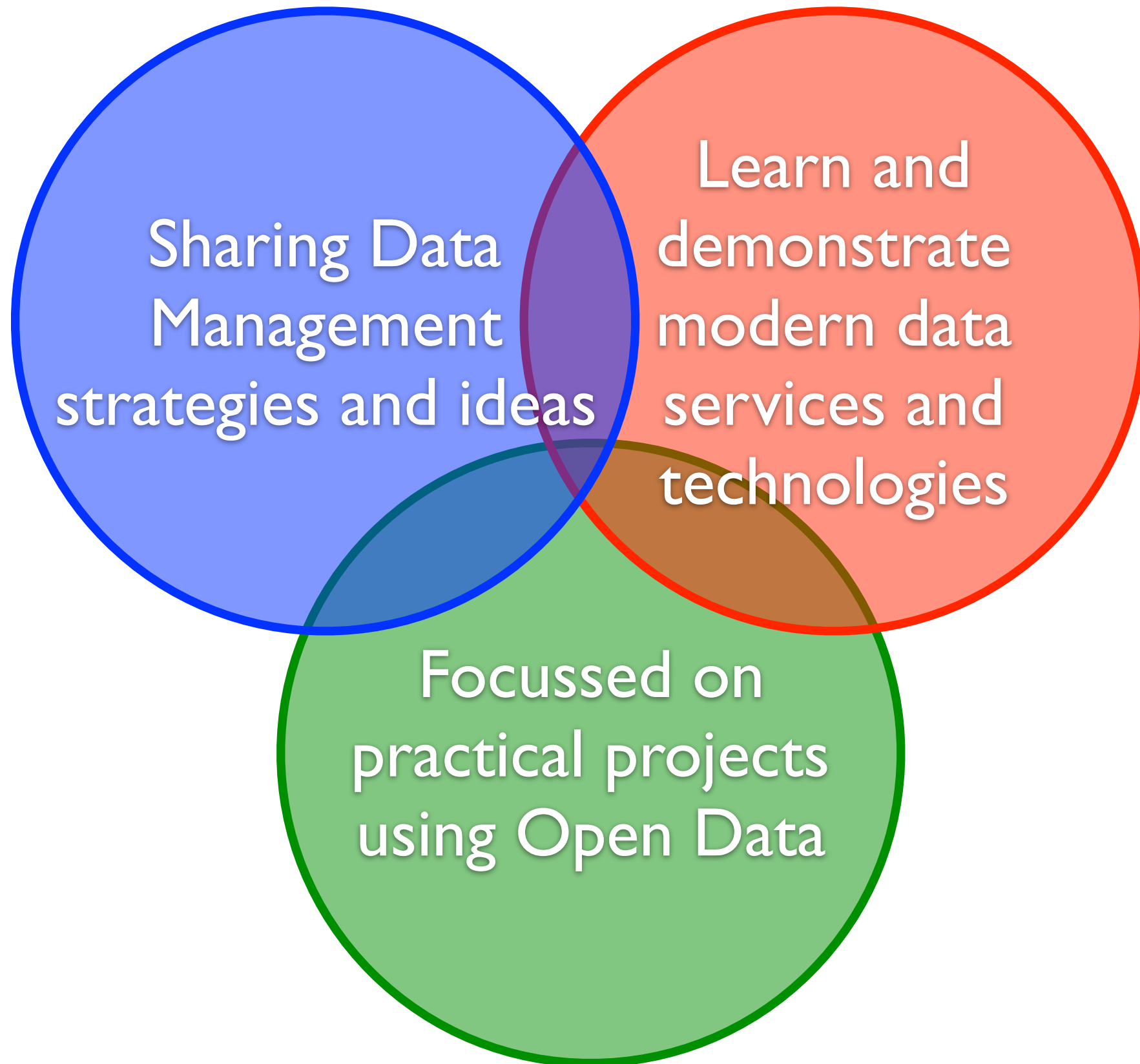
# Who Are You?

[FirstName, Interest in Data, Objective, Occupation, Fun Fact]

# What is “Data Management”?



# Why this Meetup?



# What isn't DM?

- Data Creation - application development
- Data Visualization - dashboards and presentations
- Data Science - detailed stats and machine learning

Similar skills (Data Modellers, DBAs) but different focus

But they are **\*\*related\*\*** and many of the skills and techniques are applicable

# How Do We Fit?

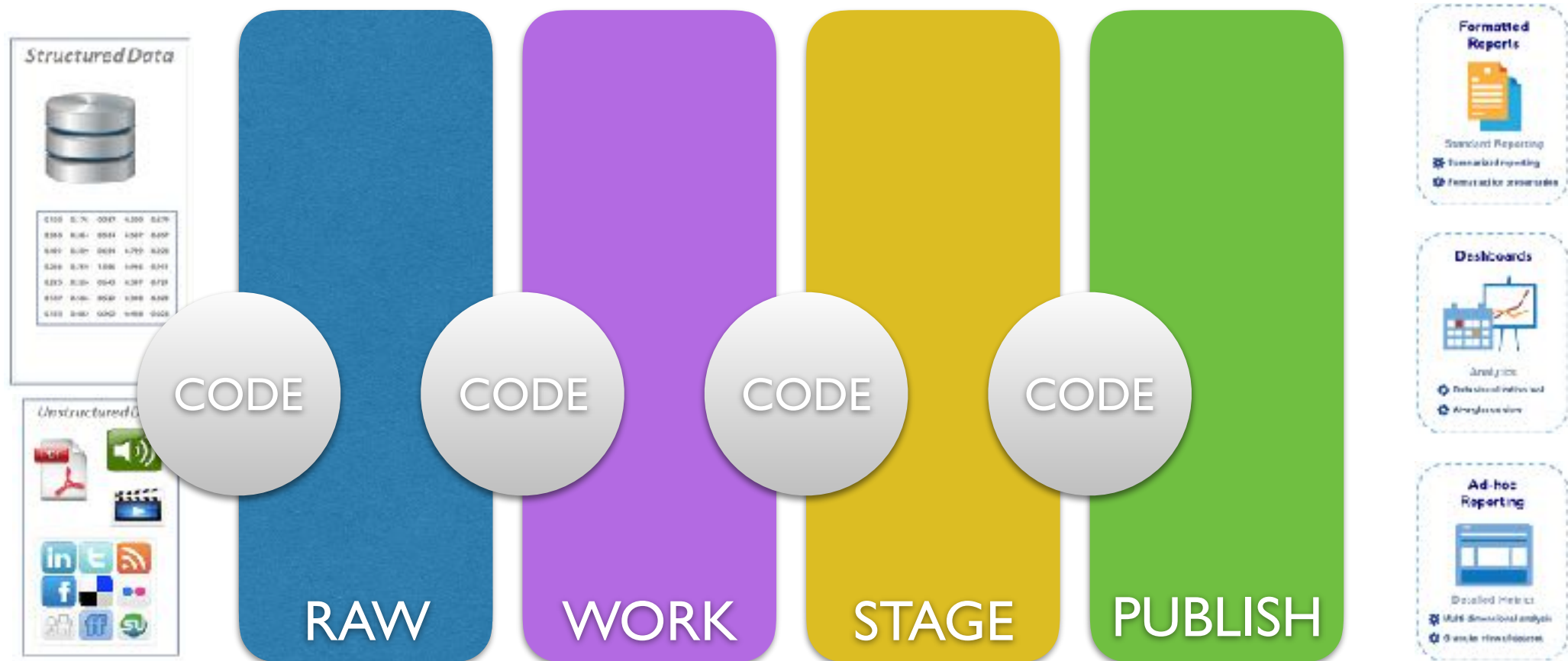


# What Will We Do?

- Some actual data management
  - Not just presentations
- Pick specific projects over the year that we can use as models to try different styles, methods and techniques
- Share ideas
- Teach each other some new skills



# Generic Terms

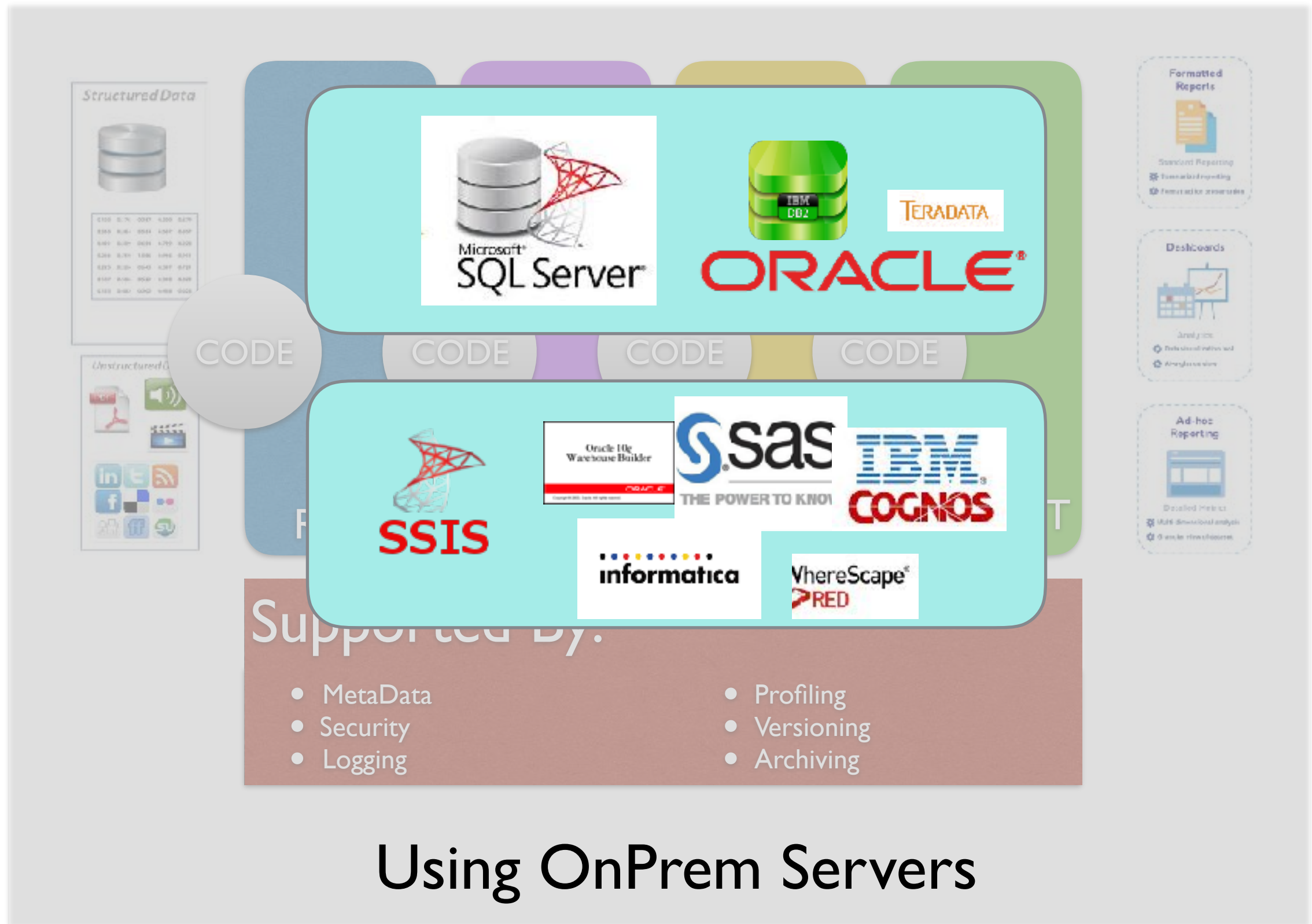


## Supported By:

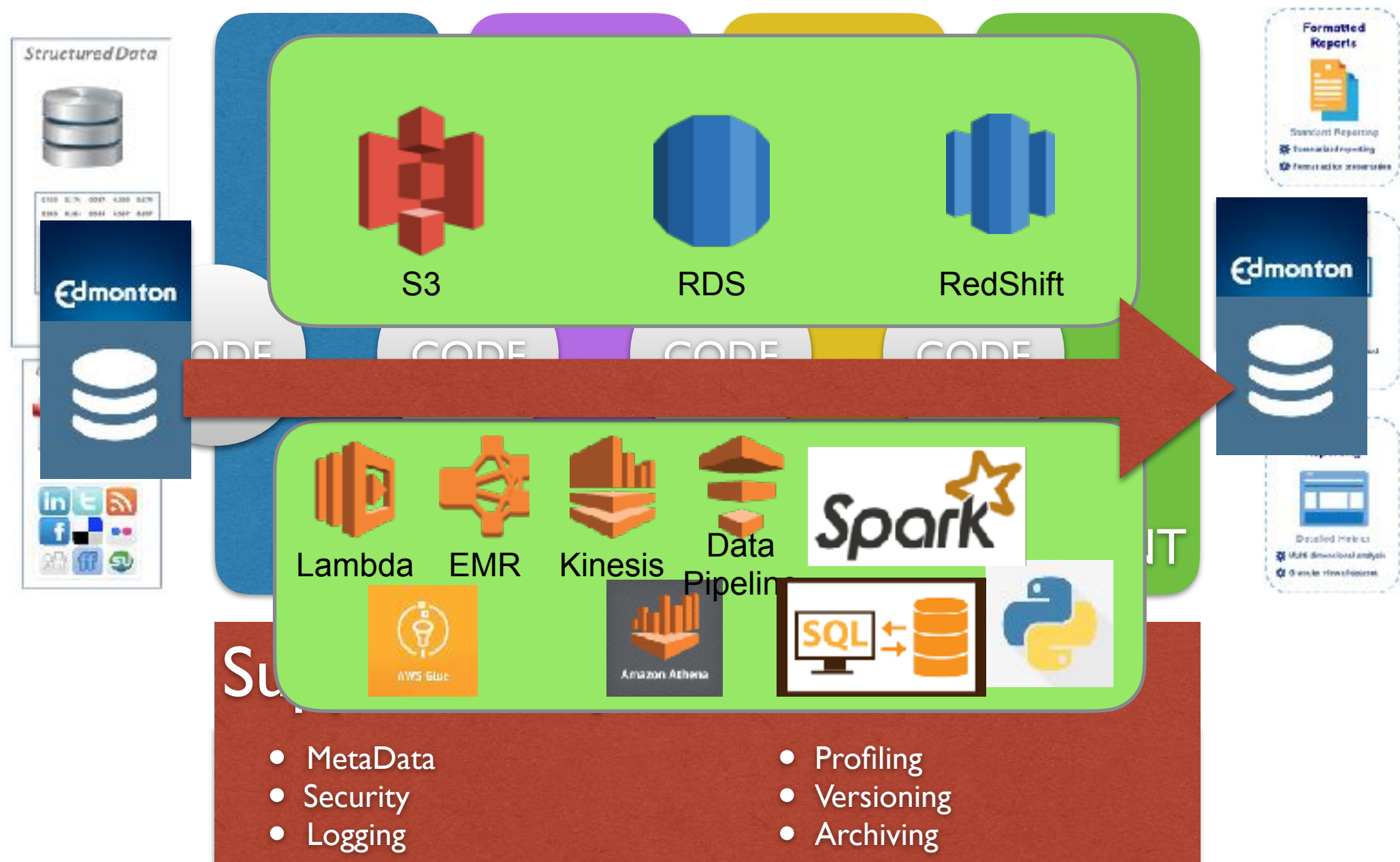
- MetaData
- Security
- Logging
- Profiling
- Versioning
- Archiving
- Documentation
- Testing
- Auditability
- Governance



# Typical (Alberta) Tools



# Proposed Environment



Using AWS Cloud Services

# What's The Plan?

Over the next few months ....

- Pick a subject area from Edmonton Open Data
- Do some Data Profiling to understand
- Modelling of proposed solution
- Design and develop some ETL processes
- Do any "Non-Functional" tasks we decide on 😊

# What About Tonight?

What are you wanting to do or learn?

*This is the part of the meetup where we can find out what folks are hoping to get out of this meetup 😊*

# What Can I Do?

## Data Management

- Look through some of the datasets
  - The YEG OpenData web site
  - Socrata based - cool features
- Try some data profiling
  - Excel
  - Azure ML
  - Jupyter Notebook
  - Tableau
  - Trifacta Wrangler
  - RapidMiner
- At Next Meetup ...
  - Tell/Show your learnings and ideas

## Other Stuff ...

- Website?

# House Keeping

- Github - <https://github.com/DataManagementYEG>
  - Website - [datamanagementyeg.github.io](http://datamanagementyeg.github.io)
  - YEG OpenData site - <https://data.edmonton.ca>
  - Twitter - @DataManageYEG
- 
- Canadian Open Data Summit 2017
    - June 12-14





# Discussion