

# Utility & Big Data – an area of conflicts?

Torsten Birn,  
OMNETRIC Group, Energy Insight,  
16<sup>th</sup> of September 2015

# OMNETRIC Group – combining complementary strengths of Siemens and Accenture

## SIEMENS

- Leading engineering and energy technology
- World class R&D and innovation power
- Leading player in grid control
- Strong focus on grid applications

## OMNETRIC Group

A Siemens & Accenture Company

- Provides leading system integration and managed services capabilities
- Access to leading Siemens energy product portfolio and Accenture consulting and IT capabilities
- Integrated and innovative solutions for improved IT/OT convergence
- Agile, collaborative and results-oriented delivery
- Highly skilled practitioner team

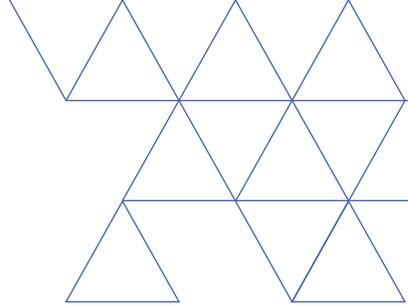
## accenture

- Recognized leader in SI and managed services capabilities
- Industry-specific technology, assets and processes
- Strong delivery methodology and experience onshore and offshore

At first, a simple example of conflicting priorities in „Big Data“

**BASE-TECHNOLOGY**

# Conflicting priorities – Storage medium

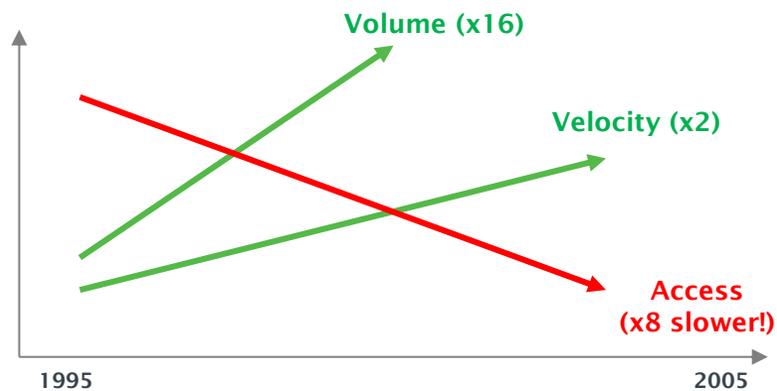


Within 10 years (from 1995 till 2005):

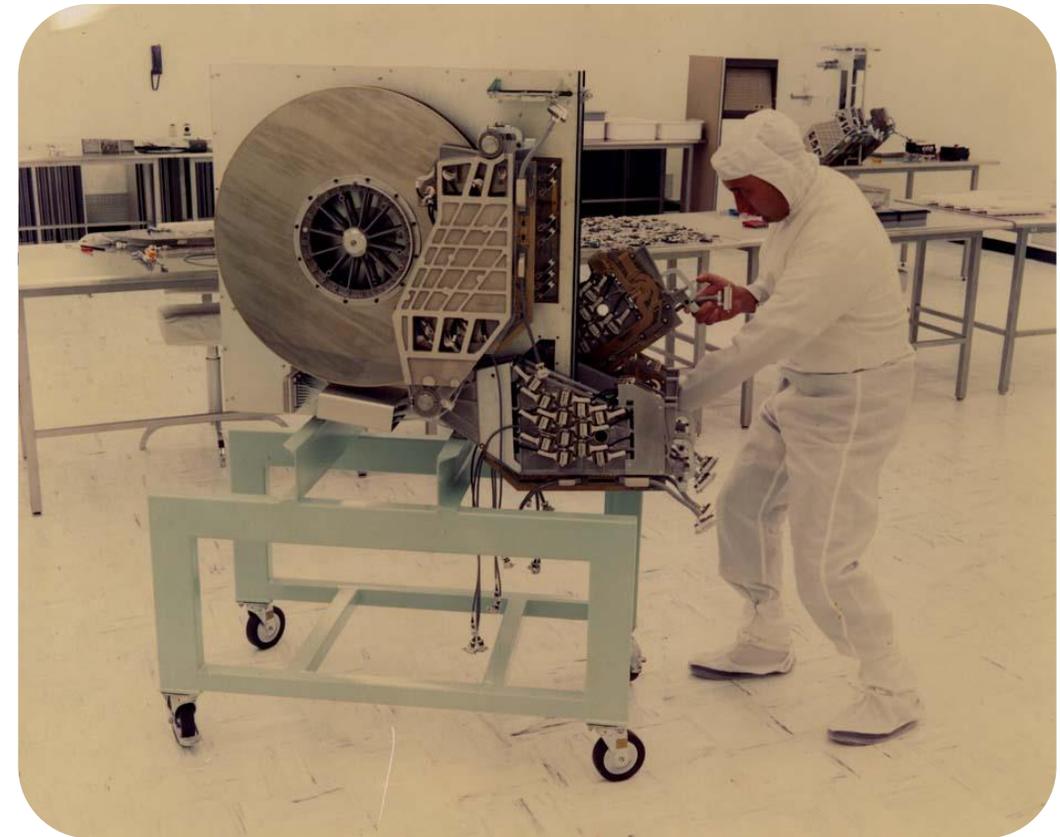
Hard disks („Enterprise Class“)

- ▶ x16 higher volume (GB)
- ▶ x2 higher velocity (IOPS)

That’s fantastic, isn’t it? Where is the problem?



Source:  
<http://research.microsoft.com/pubs/64597/tr-2005-181.pdf>



# Conflicting priorities– Storage medium

Memory is becoming cheaper? BUT, the amount of data is rising...

## Memory

- ▶ EUR 214 für 32 GB (EUR 6.848 pro TB)
- ▶ IOPS:  $\sim 10^4$ -5

## SSD

- ▶ EUR 344 für 1 TB
- ▶ IOPS:  $\sim 10^3$ -4

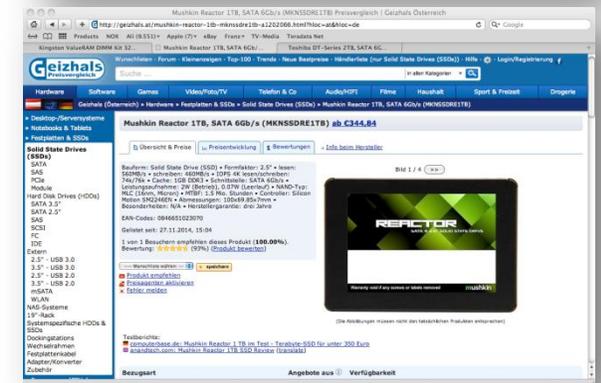
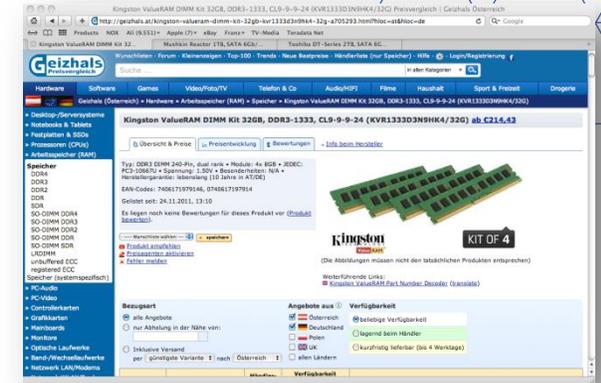
## HDD

- ▶ EUR 68 für 2 TB (EUR 34 pro TB)
- ▶ IOPS:  $\sim 10^1$ -2

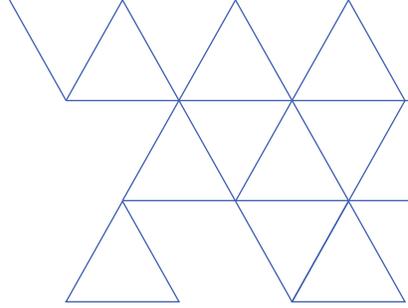
PS: does the business case pay off...



x200 EUR/Terabyte



There are several more...



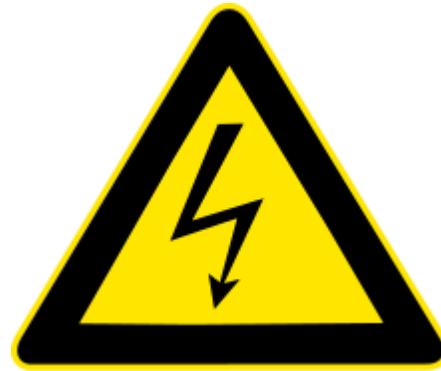
Commercial products (COTS) ?  
*OR*  
Open Source Software?

Centralized Analytics?  
*OR*  
De-centralized Analytics?

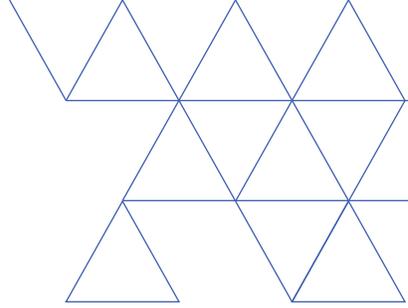
Transactional Data base?  
*OR*  
Analytical Data base?

Real-time?  
*OR*  
Offline?

On-premise?  
*OR*  
In the Cloud?



# Mario Andretti's famous quote...



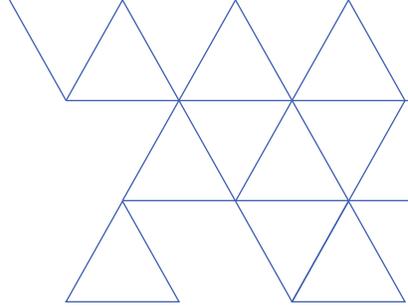
**“If everything seems under control,  
you're not going fast enough.”**



## **Our Idea**

**Having full control,  
while still being fast!**

# Accelerate first, control later



I am sure you never came across these issues in your company/ at your client, right?

Poor Data Quality?

Basically no Integration?

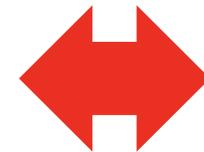
Missing joint IDs?

Different Formats?

Low Consistency?

No Standards?

IT folks look on their calendars to build “something”



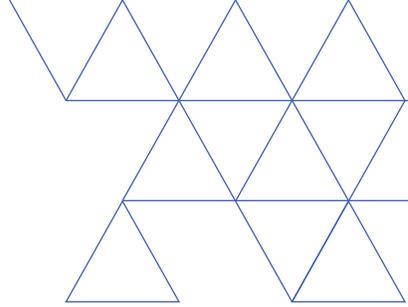
Business users look on their watches to get first results

How can this ever work out?

**Typical Approach:** Clean up first!  
(Mario would call this “control”)

**Typical Approach:** Business starts building “another silo”

# Traditional and new approach



**Business**  
Specifies requirements and defines business questions

**IT**  
Structures data to answer existing business questions

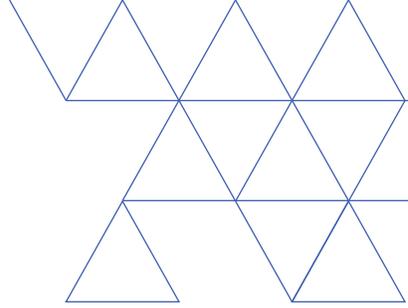
**Traditional Business Intelligence**  
Structured and repeatable

**IT + SMEs**  
Provide platform and domain expertise (!) to easily query data from various sources

**Business**  
Explores data to identify and harvest hidden value and find new questions

**Data Discovery**  
Multi-structured and iterative

# “The genius of AND versus the tyranny of OR”



*“What traditional systems like relational are really good at are answering the traditional business questions that we all still ask and will ask, and that's not going away just because the new technologies are there.”*

Ken Rudin, Director of Analytics at Facebook

## Traditional Business Intelligence

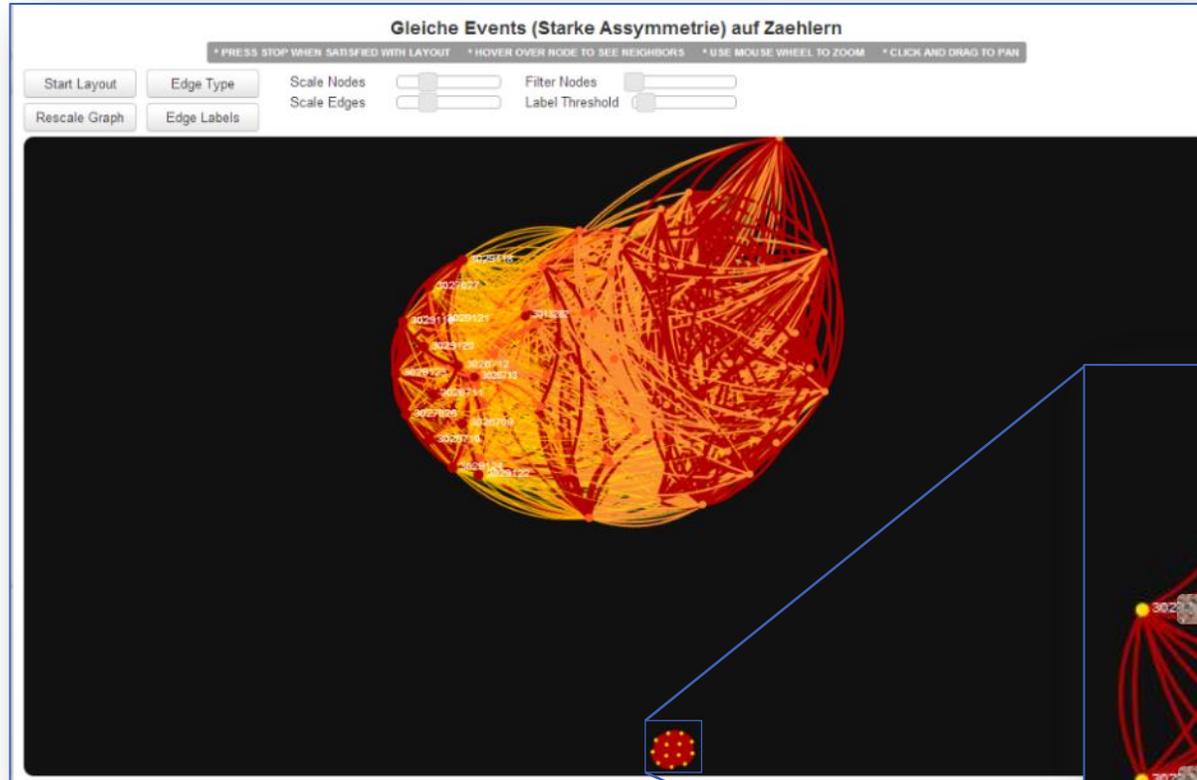
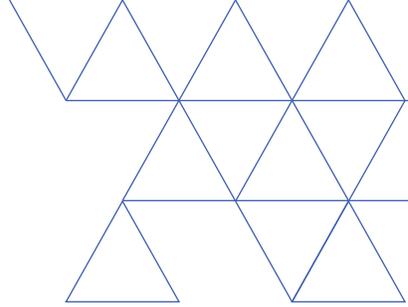
Structured and repeatable



## Data Discovery

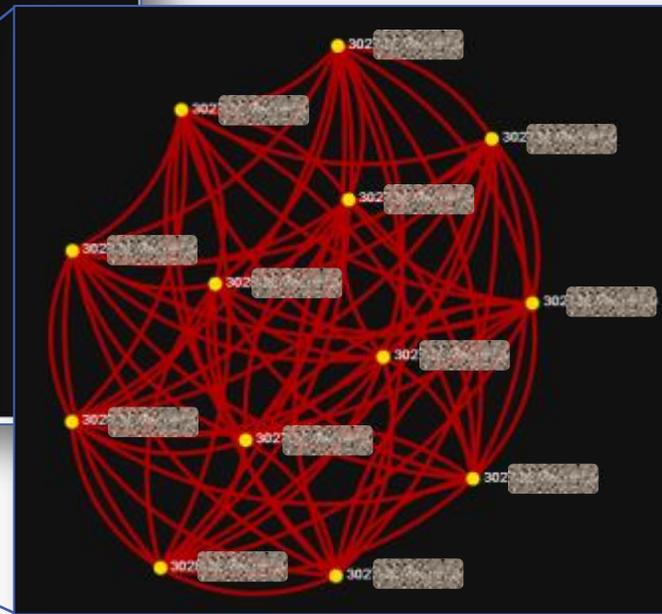
Multi-structured and iterative

# Quote DSO Engineer: “I know absolutely everything in my grid” – Well, really?

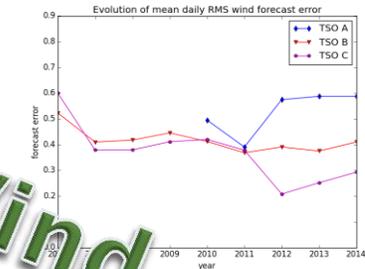
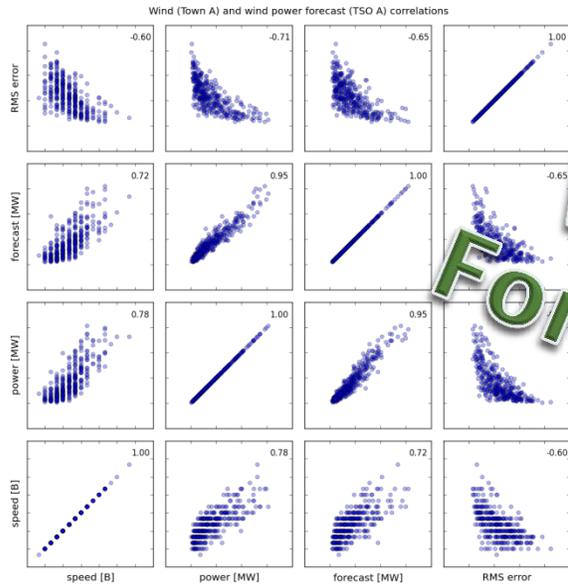
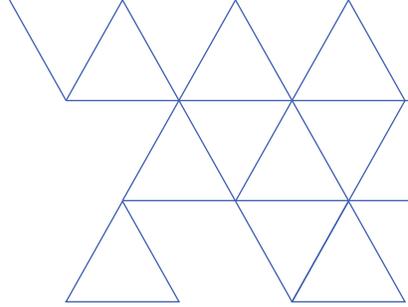


Source: Publicly funded ISOLVES:INIS project, Austria, 2014, Partners: Netz OÖ, Salzburg Netz, Wiener Netze, AIT, Grintec, Siemens, Teradata

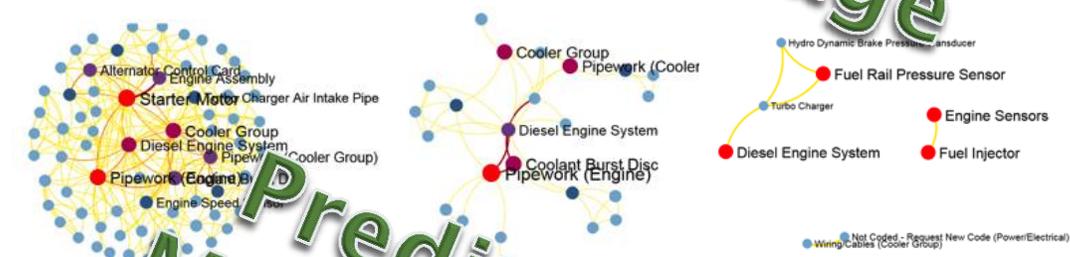
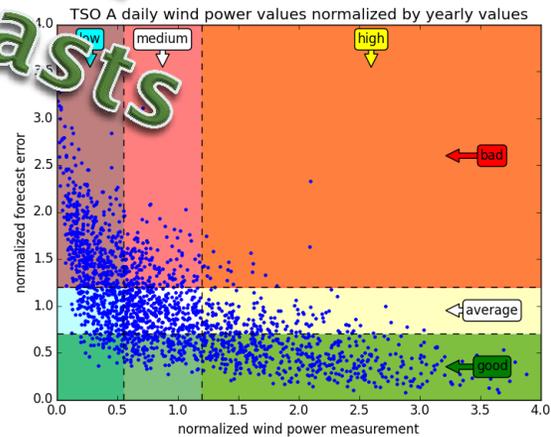
**Strong voltage asymmetry in one particular feeder identified...**



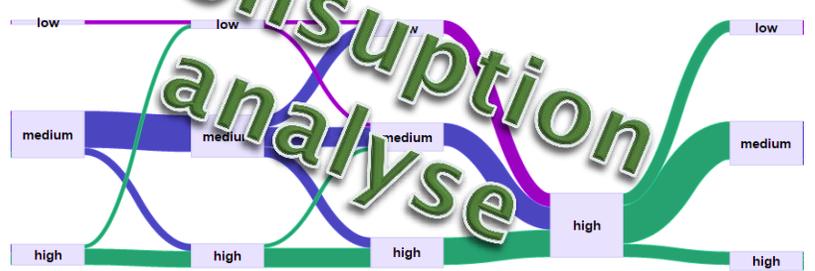
# Data Discovery – a few examples...



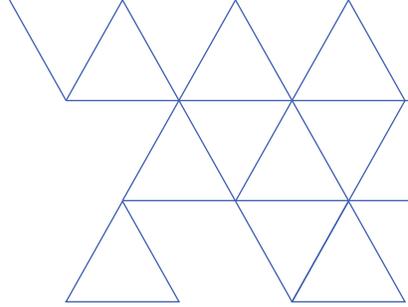
Wind Forecasts



Consumption analyse

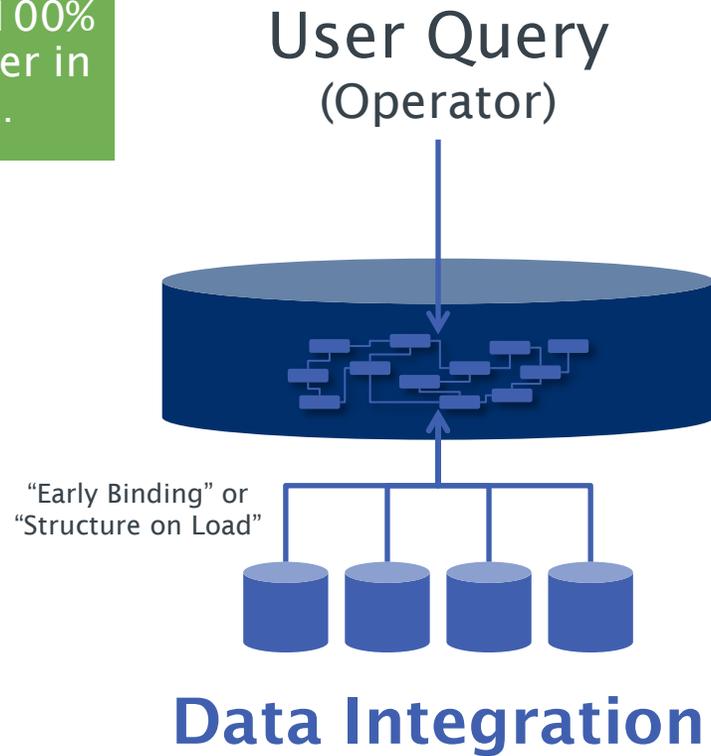


# Conflicting priorities? – Data Integration vs. Data Discovery



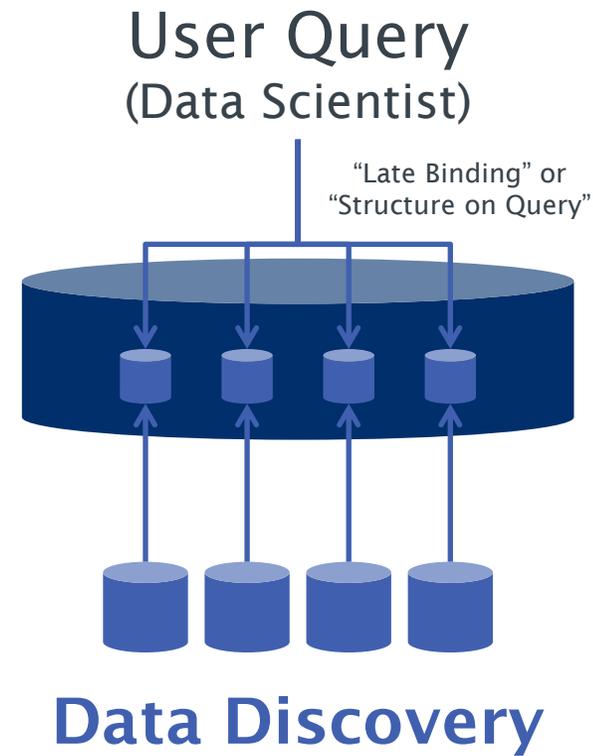
## Data Integration

If you need a 100% accurate answer in 100 days...

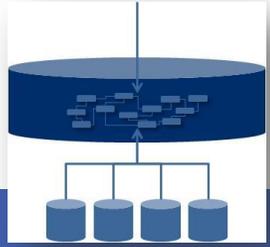
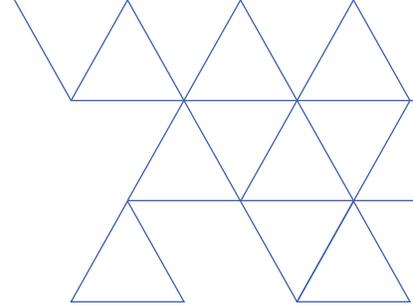


## Data Discovery

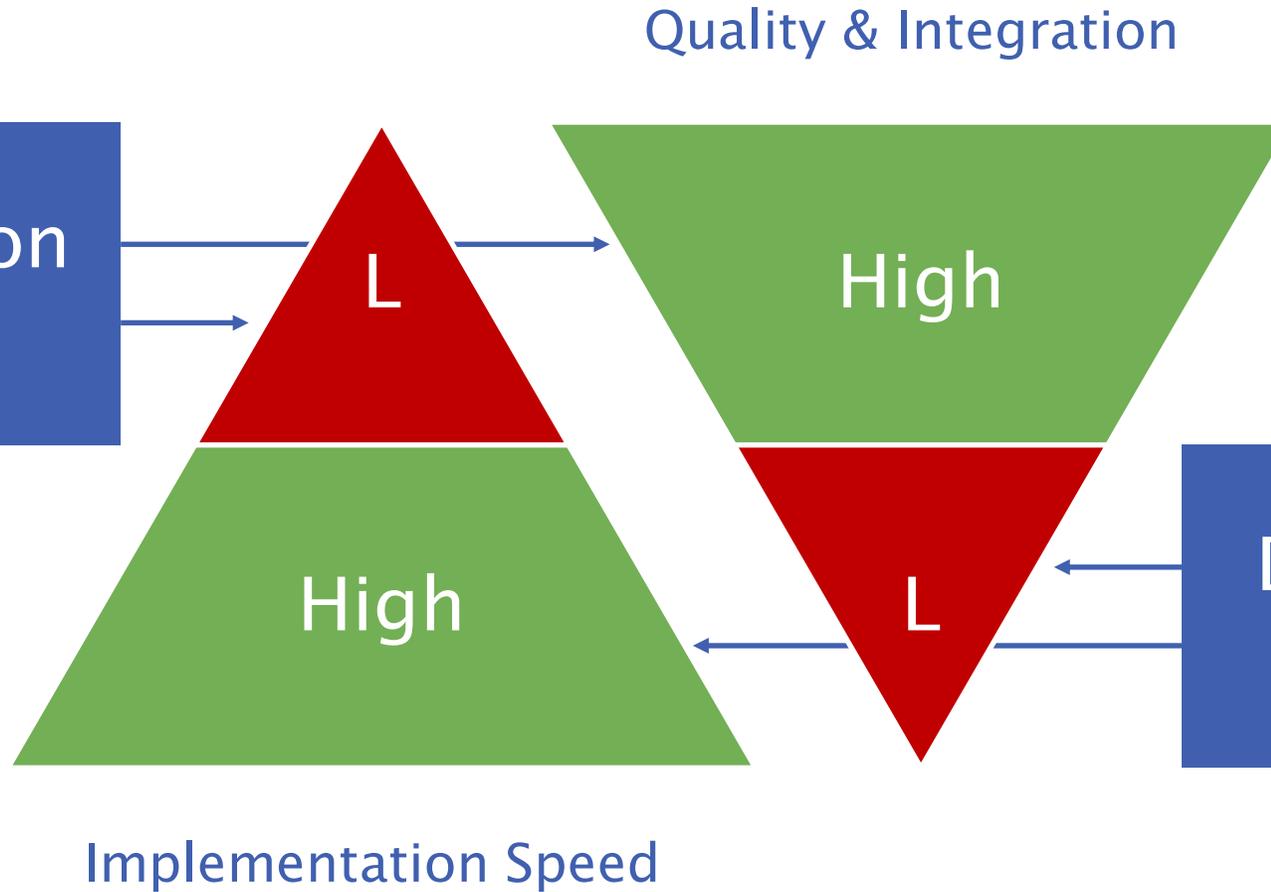
If you need a 90% accurate answer in 90 minutes...



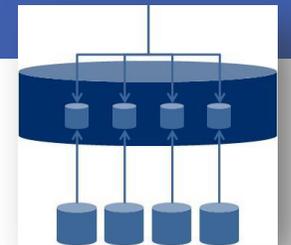
# Still a conflict? Data Integration vs. Data Discovery?



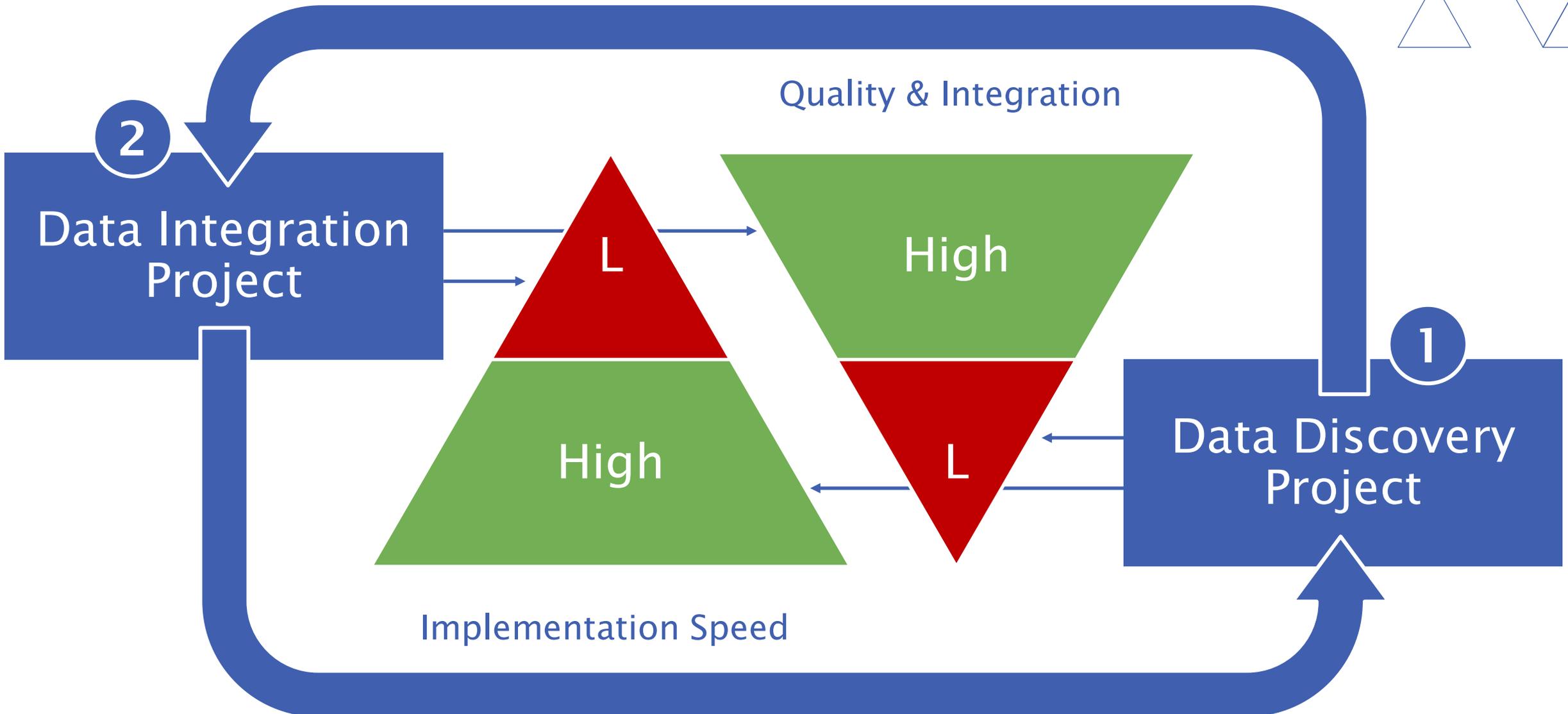
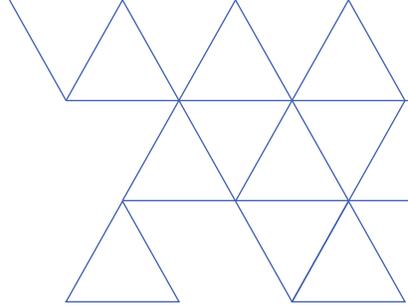
Data Integration Project



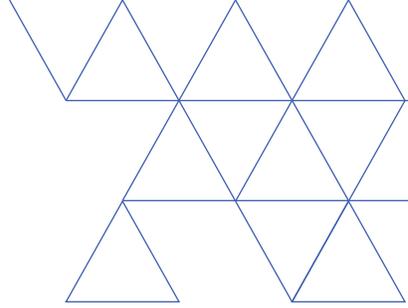
Data Discovery Project



# Two complements – Start the cycle Integration AND Discovery!



# Thank you



OMNETRIC Group is dedicated to the global delivery of integrated information technology and operational technology solutions and services, helping utility companies to achieve greater grid reliability and efficiency. OMNETRIC Group can support clients with innovative solutions wherever they may be on their path to a smarter grid. For more information, visit [www.omnetricgroup.com](http://www.omnetricgroup.com)

For more information, contact:

Torsten Birn  
Energy Insight

[torsten.birn@omnetric.com](mailto:torsten.birn@omnetric.com)

