

# Streaming Analytics, Data Lakes and PI Integrators

Presented by **Matt Ziegler**  
**Aaron Loe**

# Conference Theme and Keywords





# A journey through history

OSIsoft **VCAMPUS LIVE! 2010**

## Multidimensional Analysis

*Quick, ad-hoc, aggregation of vast amounts of interesting data, based on user selected criteria, to identify business opportunities*



2010




## Unleash the Power of Big Data




2014

@OSIsoftUC | #UC2014



## Bridge IT and OT with a process data warehouse

2016 / Matt Ziegler, OSIsoft



OSIsoft USERS CONFERENCE 2016



## Project Rubik

2012

OSIsoft. USERS CONFERENCE 2012



## Transforming Process Data Into Information

PI Integrator for Business Analytics

Presented by **Rodrigo Quintero**  
Operations Model & Planning Manager



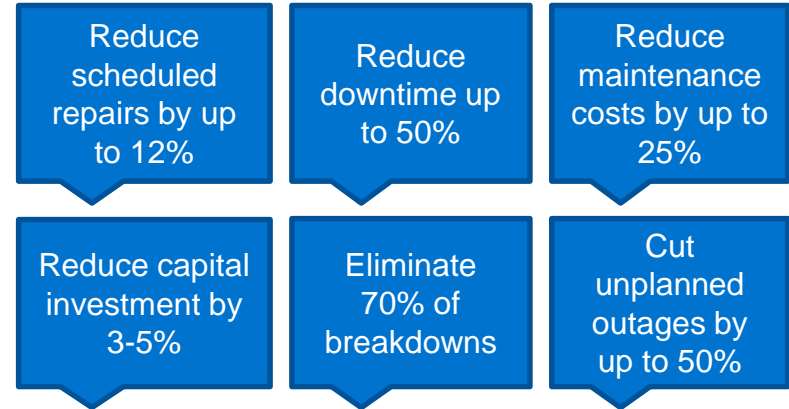
2015

# Value in the data, Value in the expertise

## Constants

Data Expertise Hard Work  
Infrastructure Self-Service Tools

## Goals



Machine Learning

Big Data

IIoT

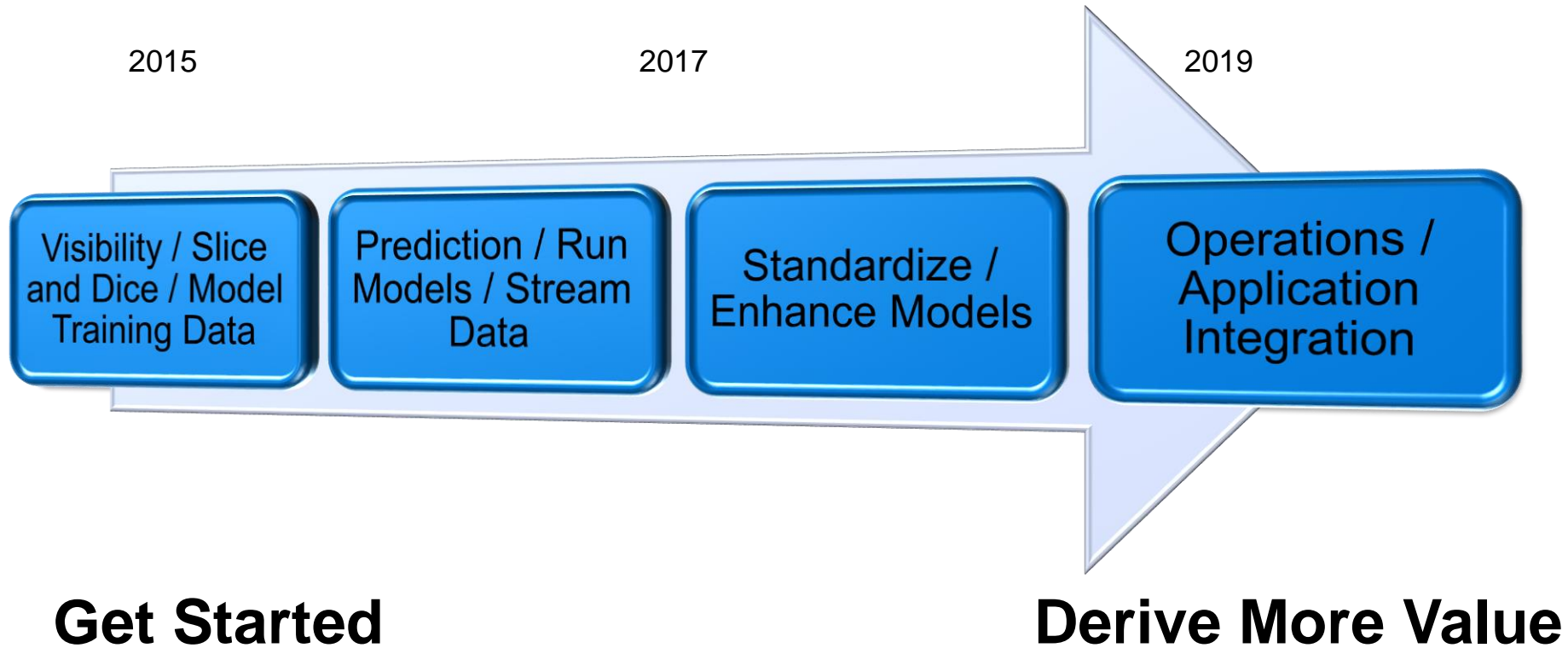
## Trends

*PI Integrators **augment** the existing set of **PI** capabilities like visualization, data access, and analytics with capabilities that **make it easy** to **interact** with non-OSIsoft **tool sets**.*

# Agenda

- π Recap on PI Integrators
  - π Capabilities
  - π Roadmap
  - π Success Criteria
- π **Streaming**
  - π **Use Cases**
  - π **Architectures**
- π **Demo**
- π Summary
- π Q&A

# Progression



2015-2016

2017

Future

**Business Intelligence & Data Warehouses**

**Available Today**

**PI Integrator for Business Analytics**

- Microsoft SQL Server, Oracle
- Hadoop (HDFS/HIVE)

**PI Integrator for SAP HANA**

**Available**

**Cloud Platforms**

- Microsoft Azure
- HANA Cloud Platform (5/2017)

**(Planning)**

**More Platforms**

- ESRI ArcGIS GeoAnalytics

**Streaming Systems**

**Real-Time GIS**

**PI Integrator for Esri ArcGIS**

- Situational Awareness
- Real-Time Geoprocessing
- Import ESRI features (assets)

**Planned (1H 2017)**

**Stream Systems**

- Azure Event Hubs, IoT Hub
- Apache Kafka
- SAP SDS

**Planned (1H 2017)**

**Stream Systems**

- Azure Event Hubs, IoT Hub
- Apache Kafka
- SAP SDS (April 2017)

**PI Integrator Framework**

**Planned (1H 2017)**

- Process Scale out
- SSL / HTTPS

**(Planning)**

- All Integrators on common Framework (ESRI)
- Node Scale Out and HA

**New Integration Patterns**

**Research**

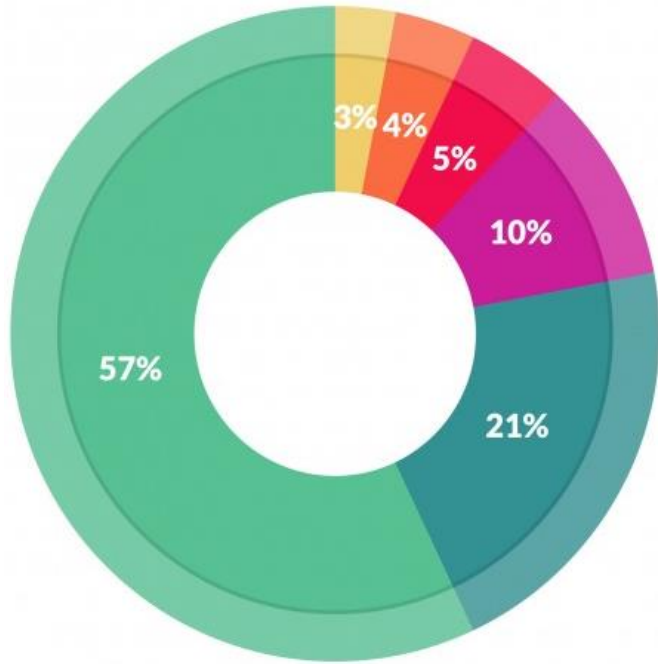
Enable business process orchestration with PI System data – workflow, asset sync, transaction-like data, MES

**Research**

Enable partners and customers to build applications and interact programmatically using PI Integrator Framework.



# Data Scientist is the sexiest job of 21<sup>st</sup> century, but...



## What's the least enjoyable part of data science?

- Building training sets: 10%
- Cleaning and organizing data: 57%
- Collecting data sets: 21%
- Mining data for patterns: 3%
- Refining algorithms: 4%
- Other: 5%
- Being Sexy: 0%

Source: <http://www.forbes.com/sites/gilpress/2016/03/23/data-preparation-most-time-consuming-least-enjoyable-data-science-task-survey-says/#5481f6037f75>

# Time Series Data is Complex!



Turbine 1

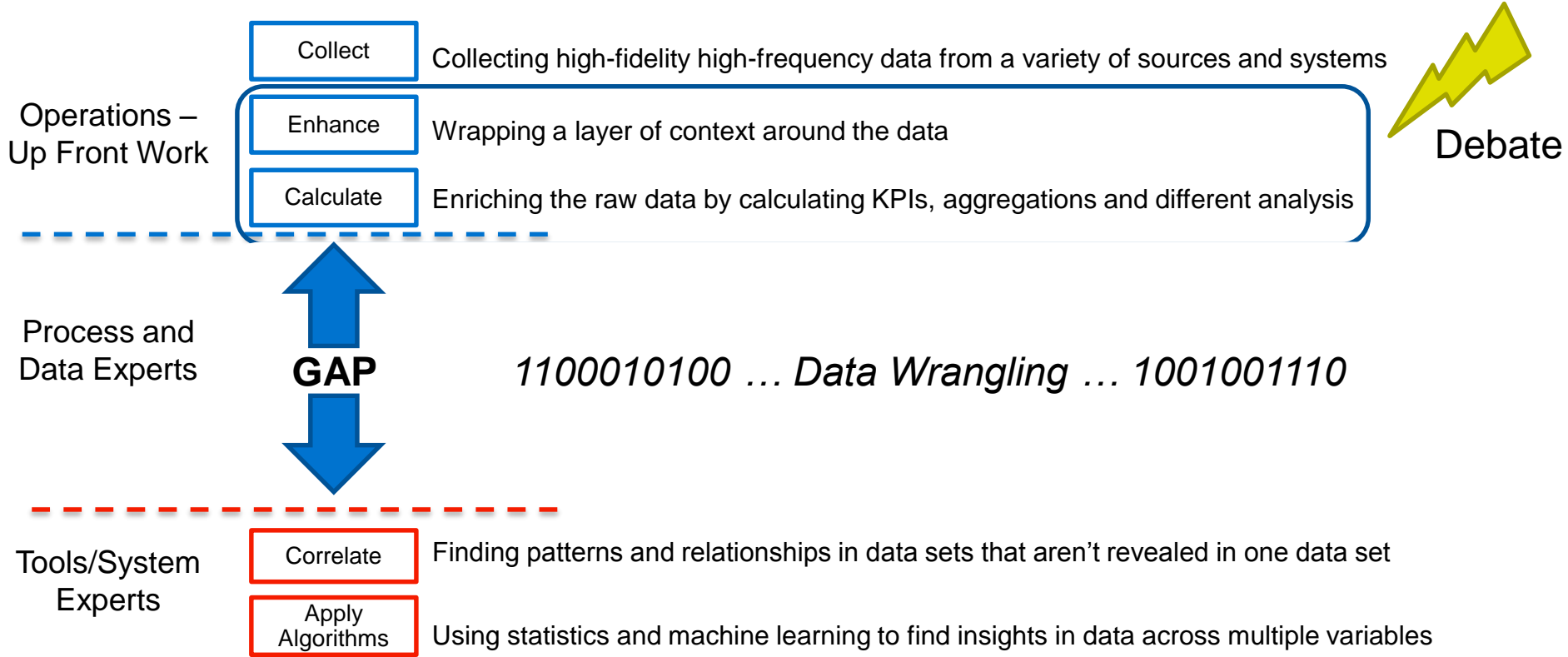
Speed  
Bearing Temp  
Oil Temp



Turbine 2

Speed  
Bearing Temp  
Oil Temp  
Wear Factor

# What do we need to approach this problem?



# Soft Measurements – Ideal Use Cases for Streaming

- Predicting Soft Measurements
  - Deschutes Brewery – Fermentation Transition
  - [Pharma Company] – Lights Out Manufacturing
  - Quality
- Enabling Data Lakes
  - Customize how data is stored & processed

# Transition occurs between infrequent manual measurements

## Constraints

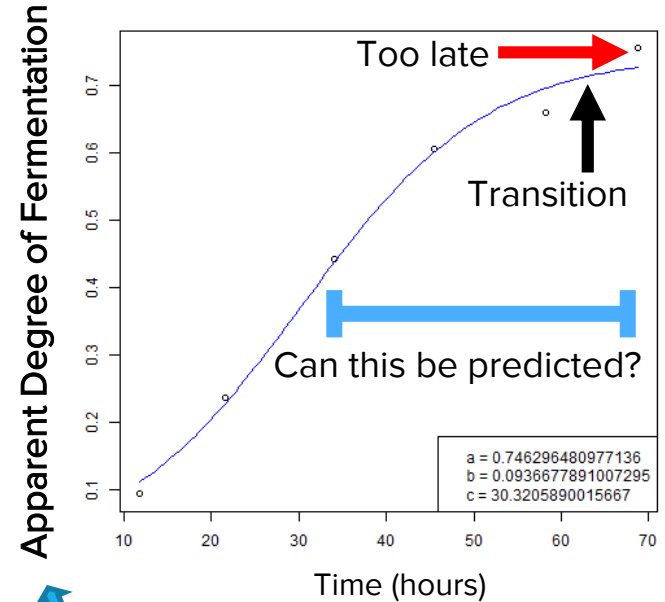
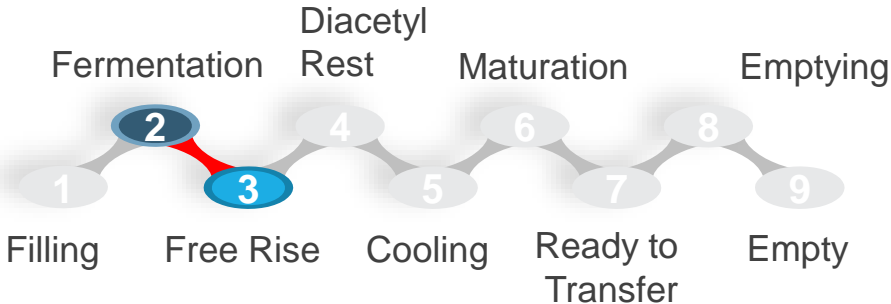
- One manual density measurement per vessel every 8-10 hours
- Large capital expenditure not an option

## Impact

- Up to 72 hours lost in production

## Options

- \$750k for inline density meters
- Manually predict transition in spreadsheets



Soft Measurement

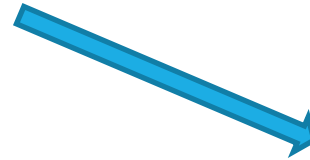
# Automate and Optimize

## Lab Measurements

- Viable Cell Count (VCC)
- Sugar Concentration
- API Concentration
- ...

## Online Measurements

- Agitation Speed
- pH
- **Spectroscopy**
- ...



*Goals: Optimize Cell Production, Eliminate Coming to the Lab*

*Technique: Use indirect measurements and machine learning to predict outcomes of lab measurements*

# Exercise Caution – Think PI System First

- Condition Based Maintenance / Detecting Failures
  - PI is already really good at this
  - <http://www.osisoft.com/corporate/power-cbm/>
- Predicting Failures
  - Contextualized Data Required
  - Not Enough Failures for predictions
  - Better suited for batch analytics



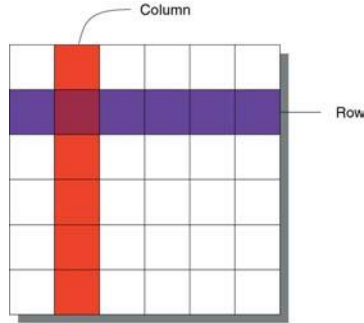
# PI Integrator for Business Analytics View Types

- Asset View - Tables
- Event View - Tables
- Streaming View (In/Out) – Streams





# Tables vs Streams



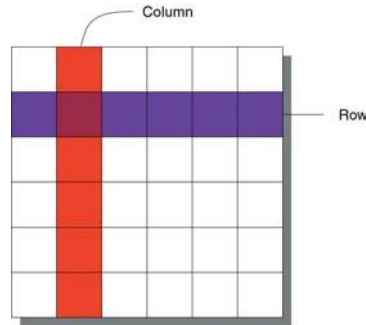
```
1  {  
2    "id":ObjectID(""),  
3    "name":"Phase A Angle",  
4    "timestamp":"2016-08-27 04:17:43",  
5    "value": 0.0034,  
6    "name":"Phase A Angle",  
7    "timestamp":"2016-08-27 04:17:43",  
8    "value": 119.9563,  
9    "name":"Phase A Voltage",  
10   "timestamp":"2016-08-27 04:17:43",  
11   "value": 240.0034  
12 }
```

- Business Intelligence
- Human readable
- Batch / Bulk Process
- Normalized data
- Regularly scheduled
- Large data, few messages

- In-line (Streaming) Analytics
- Computer readable
- Specific Data / Targeted Process
- Raw or “Packages” of data
- Triggered
- Small data, many messages

# Integration Patterns

## Tables



Files  
Databases

## Streams

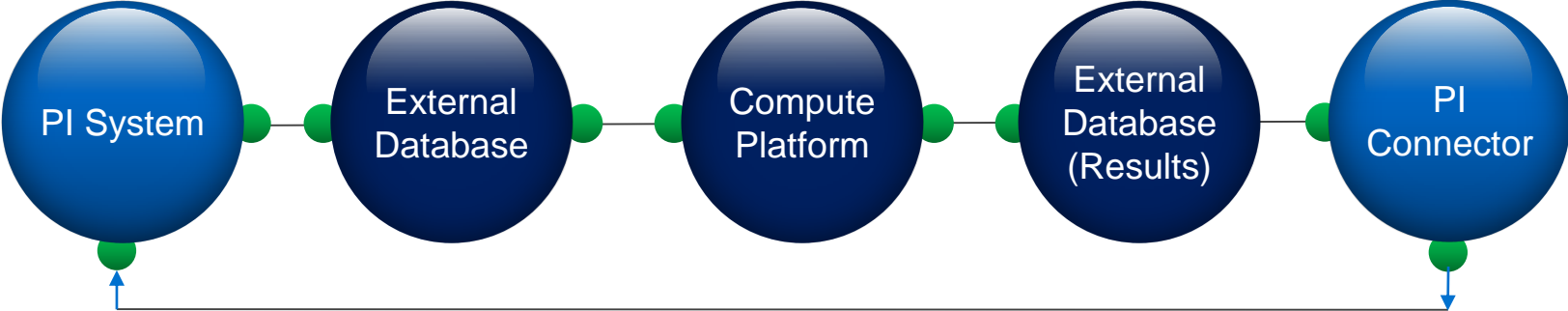
```
1 {  
2   "id":ObjectID(""),  
3   "name":"Phase A Angle",  
4   "timestamp":"2016-08-27 04:17:43",  
5   "value": 0.0034,  
6   "name":"Phase A Angle",  
7   "timestamp":"2016-08-27 04:17:43",  
8   "value": 119.9563,  
9   "name":"Phase A Voltage",  
10  "timestamp":"2016-08-27 04:17:43",  
11  "value": 240.0034  
12 }
```

Files  
Queues  
Messaging  
External Analytics Engines

## Other Patterns

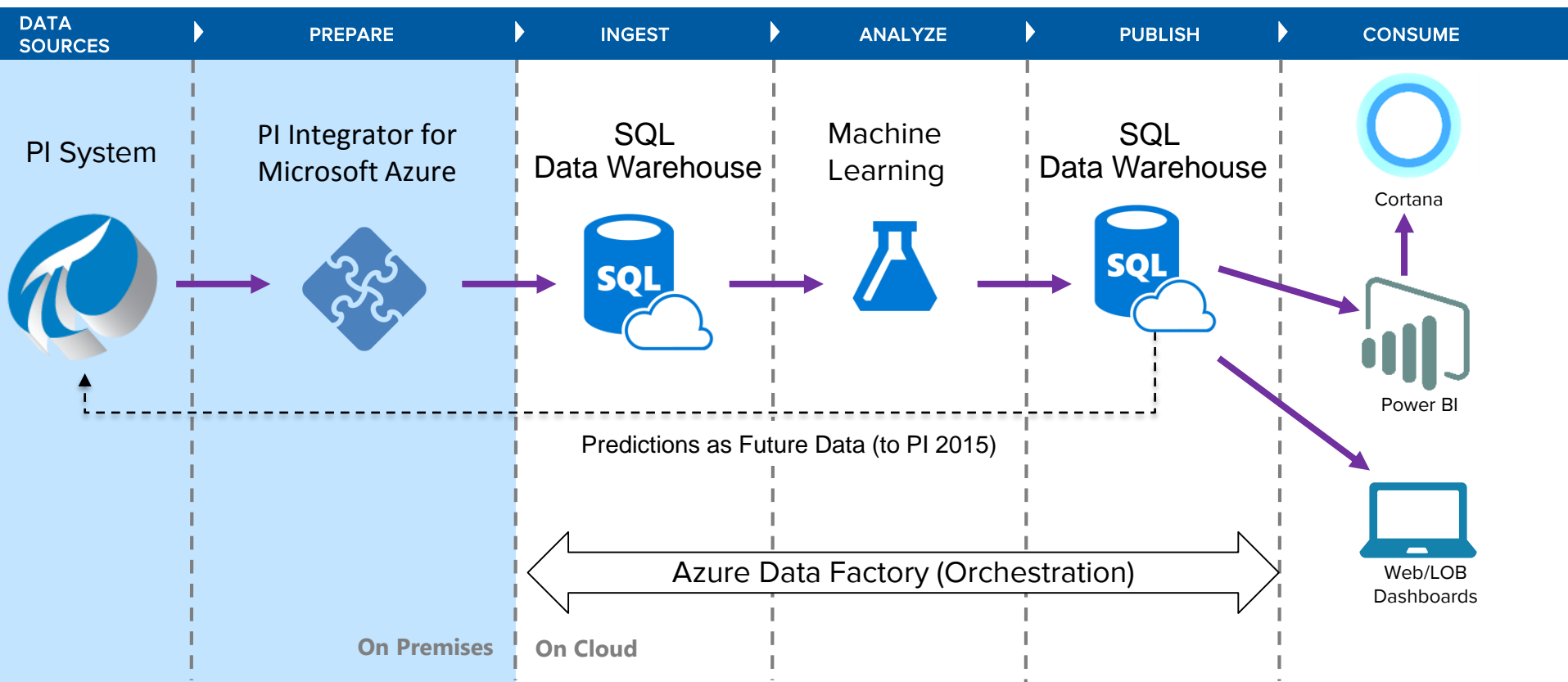
Metadata  
Programming  
On-Demand  
Workflow & Transactions

# Current Architecture with Streaming Analytics

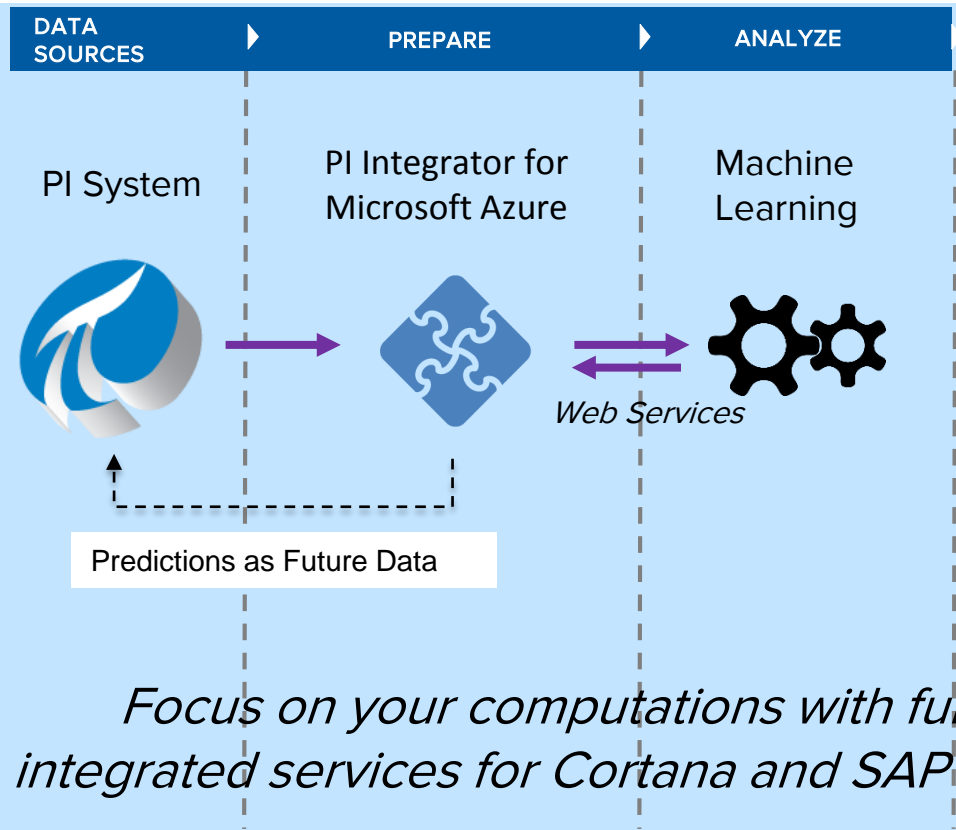


*A lot of stuff*

# Deschutes Brewery Operationalizing Architecture



# 2017 Simplified Architecture and Integrated Workflow



  
PI Integrator for Business Analytics

  
PI Integrator for SAP HANA

  
PI Integrator for Microsoft Azure

# San Leandro Tech Campus View Glass System

*View Glass Tint Changes with Outdoor Brightness*



*Finding Outliers – Health Score  
-Transition Time Relative to Peers  
-Drift in Transition Time*





# Demo

PI Integrator for Business... x

localhost:1212

My Views

OShaarfoe

+ Create Asset View  
Build a data view starting with your asset hierarchy

+ Create Event View  
Build a data view starting with your event frame hierarchy

+ Create Streaming View  
Get text for description

Modify View  
Modify existing data view

Remove View  
Remove selected view

Name	Run Status	Type	Run Mode	Start Time	End Time	Last Run Time
No Rows To Show						



PI Integrator for Business

localhost:1212

My Views OSStaartoe

**+ Create Asset View**  
Build a data view starting with your asset hierarchy

**+ Create Event View**  
Build a data view starting with your event frame hierarchy

**+ Create Streaming View**  
Get text for description

**Modify View**  
Modify existing data view

**Remove View**  
Remove selected view

Name	Run Status	Type	Run Mode	Start Time	End Time	Last Run Time
View Glass Tinting	Streaming	StreamingOut	Stream	-8h	*	12/31/99 3:59 PM

Overview Log Security

**Run Status**

View Name  
PI AF Database  
Publish Target  
View Type  
Run Mode  
Last Run Time  
Your Start Time Is  
Your End Time Is  
Sample Frequency

**Streaming**

View Glass Tinting

Facilities-1600 Alvarado  
Kafka  
StreamingOut  
Stream  
12/31/99 3:59 PM  
-8h  
\*  
1 minute

**Publish Actions**

[Resume](#)

[Stop](#)

[Update Data](#)

**Search Shape**

**Asset Shape**

- ▲ Floor
  - ▲ ViewGlassTemplate
    - ◆ Command
    - ◆ State

## Wrap Up

- Constants – Data Infrastructure, Expertise, Self-Service Tools
- Soft Measurements are Ideal for Streaming
- PI Integrator for Business Analytics 2017
  - Data in **Context** for streaming analytics



**Have an idea how  
to improve our  
products?**

**OSIsoft wants to  
hear from you!**

<https://feedback.osisoft.com/>



# Contact Information

**Matt Ziegler**

[mziegler@osisoft.com](mailto:mziegler@osisoft.com)

Product Manager

OSIsoft, LLC



**Aaron Loe**

[aloe@osisoft.com](mailto:aloe@osisoft.com)

Sr. Software Developer

OSIsoft, LLC



## Questions

Please wait for the **microphone** before asking your questions



State your **name & company**

## Please remember to...

Complete the Online Survey for this session

Download the Conference App for OSISOFT USERS CONFERENCE 2017



- View the latest agenda and create your own
- Meet and connect with other attendees



HTML

search OSISOFT in the app store

<http://bit.ly/uc2017-app>

감사합니다

谢谢

Danke

Merci

Gracias

**Thank You**

ありがとう

Спасибо

Obrigado