The Journey to Data Quality

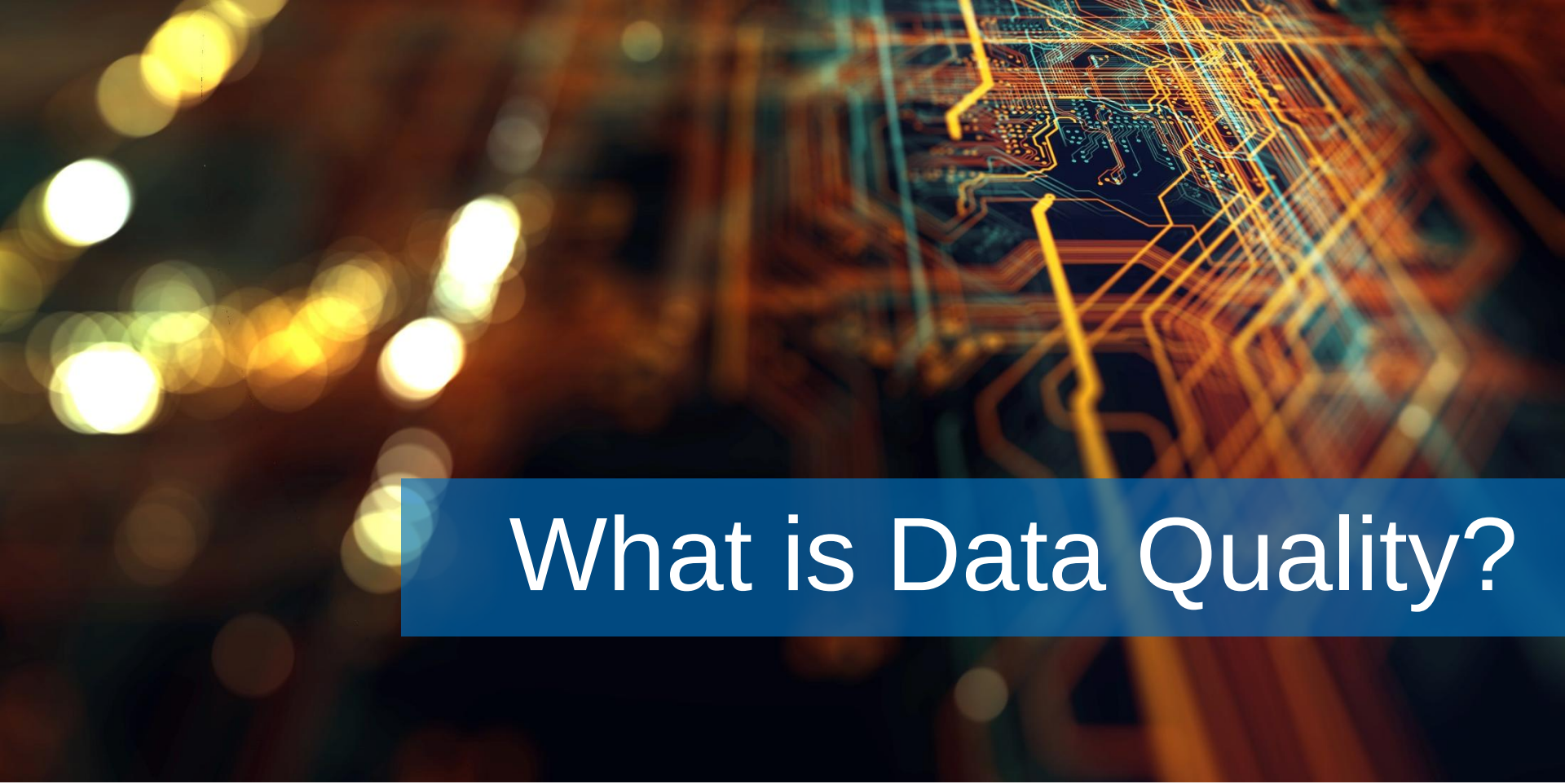
Bryan Klosiewicz
OSIsoft Customer Success Manager

©2019 OSIsoft, LLC

Data Quality Drives Business Success

- Safety and Security
- Regulatory Compliance
- Process Monitoring & Optimization
- Asset Health
- Product Quality
- Cost Control
- R&D

©2019 OSIsoft, LLC



What is Data Quality?

©2019 OSIsoft, LLC

What is Data Quality?

Objective

- **Accessibility**
- **Appropriate amount of data**
- **Completeness**
- **Concise**
- **Consistent**
- Accuracy / Free-of-Error
- **Interpretability**
- **Security / Auditable**
- Timeliness

Subjective

- Believability
- **Ease of Manipulation**
- Objectivity
- Relevancy
- Reputation
- **Understandability**
- Unique / Value-Added

©2019 OSIsoft, LLC

What is Data Quality?

Objective

- **Accessibility**
- **Appropriate amount of data**
- **Completeness**
- **Concise**
- **Consistent**
- Accuracy / Free-of-Error
- **Interpretability**
- **Security / Auditable**
- Timeliness

Subjective

- Believability
- **Ease of Manipulation**
- Objectivity
- Relevancy
- Reputation
- **Understandability**
- Unique / Value-Added

©2019 OSIsoft, LLC

How PI supports Data Quality

OSIsoft

- **Accessibility**
  - Data Centralization
  - Native Client Tools
  - PI Developer Technologies
  - PI Integrators
- **Appropriate amount of data**
  - Perf Eq / Asset Analytic filtering
- **Completeness**
  - PI Buffer Subsystem
  - High Availability + Failover
- **Concise**
  - Exception / Compression

©2019 OSIsoft, LLC

How PI supports Data Quality

OSIsoft

- **Consistent**
  - Automatic UOM conversion
  - Configurable significant figures
- **Interpretability**
  - PI Tag Configuration
  - Asset Framework
- **Security / Auditable**
  - Identities + Mappings
  - Multiple levels of security
    - *PI Server, AF Database, AF Root Element, PI Tag, etc.*
- **Ease of Manipulation**
  - PI Integrators Product Suite + PI Developer Technologies
- **Understandability**
  - Asset Framework Contextualization

©2019 OSIsoft, LLC

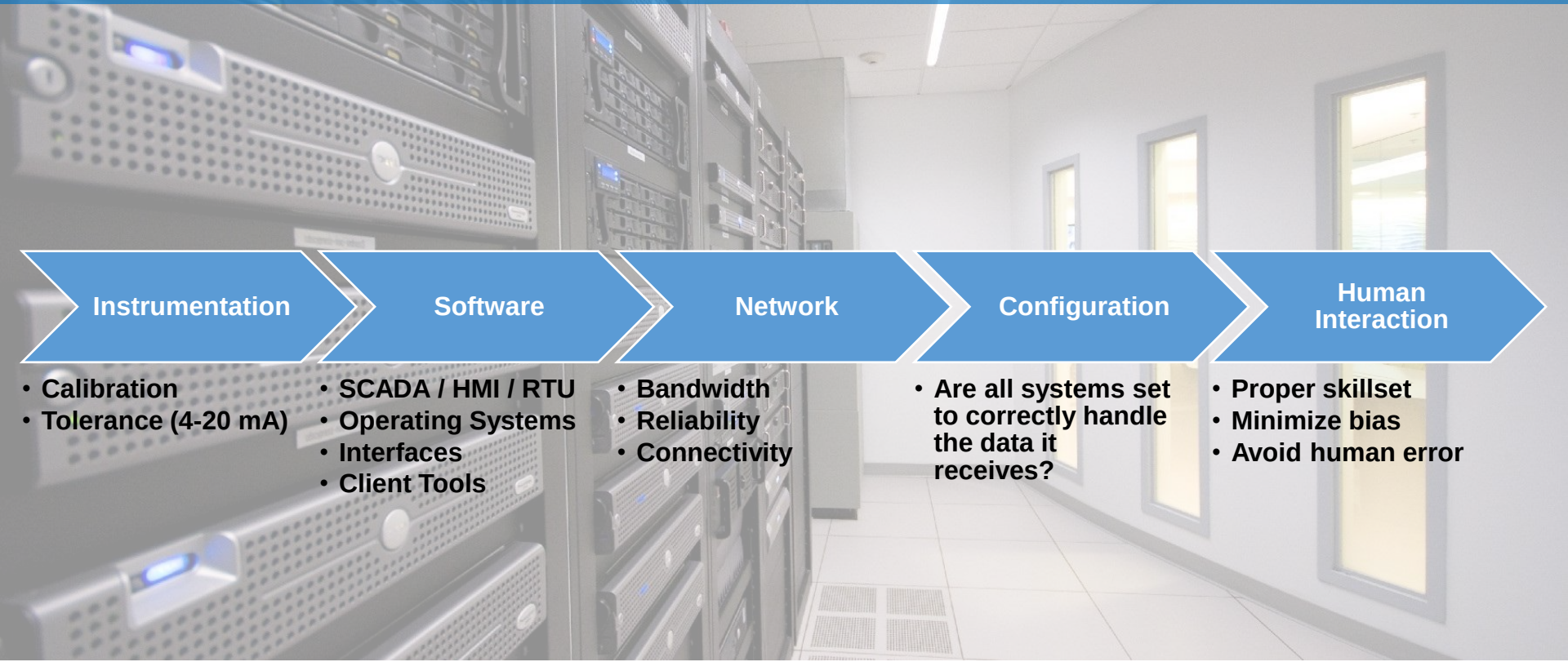
©2019 OSIsoft, LLC

Where do we look?

- Instrumentation
- Software
- Network
- Configuration
- Human Interaction
- System Integration

©2019 OSIsoft, LLC

Where do we look?



Instrumentation

- Calibration
- Tolerance (4-20 mA)

Software

- SCADA / HMI / RTU
- Operating Systems
- Interfaces
- Client Tools

Network

- Bandwidth
- Reliability
- Connectivity

Configuration

- Are all systems set to correctly handle the data it receives?

Human Interaction

- Proper skillset
- Minimize bias
- Avoid human error

©2019 OSIsoft, LLC

©2019 OSIsoft, LLC

Monitor and Detect

- System Health
- Configuration Checks
- Simple Error Analysis
- Advanced Methodologies
- Frequency

- Instrumentation, interfaces, control systems, and networks have to be functioning correctly in order to maintain data quality.
- These devices may have system flags and watchdogs to determine service status.
- Transmission rates, buffer queues, and other data stream checks should be regularly monitored.

©2019 OSIsoft, LLC

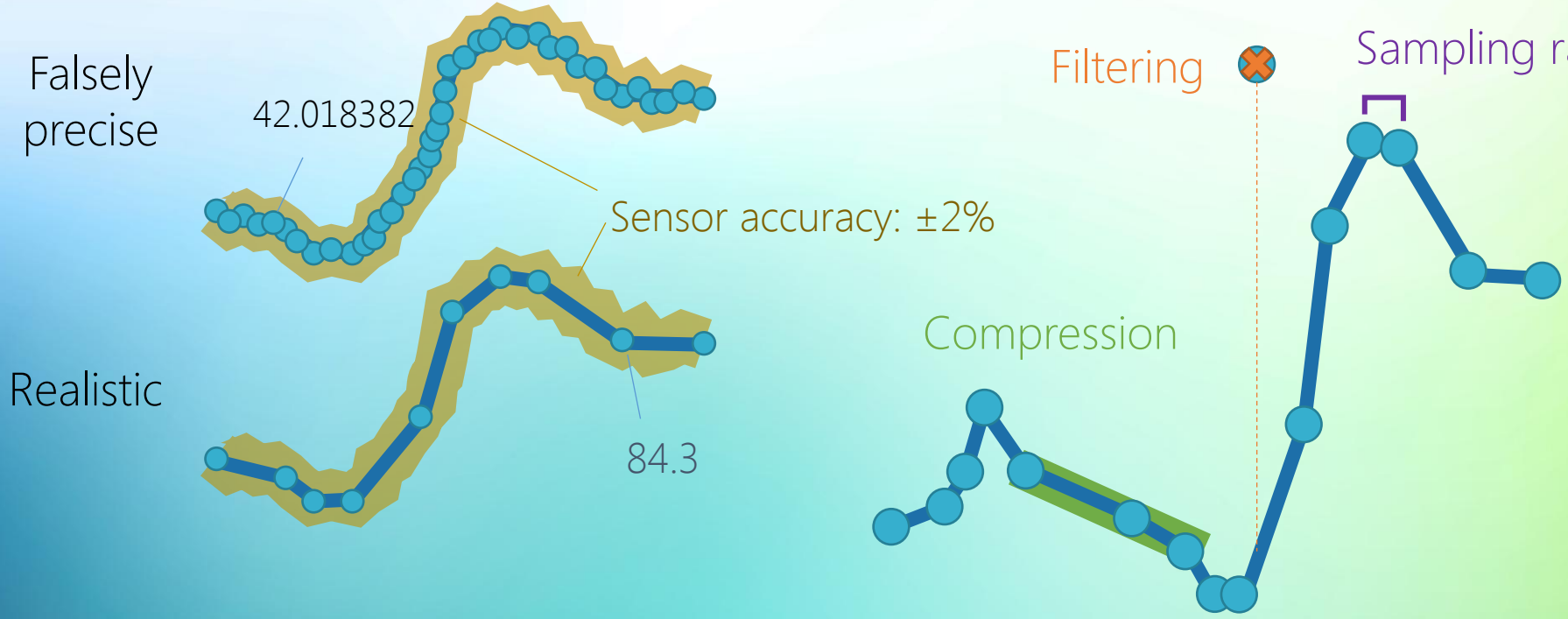
Monitor and Detect

- System Health
- Configuration Checks
- Simple Error Analysis
- Advanced Methodologies
- Frequency

- Configuration standards are a prerequisite for data quality monitoring.
- Base configurations of data classes should be validated before proceeding with more extensive monitoring.
- Validate that basic data configuration should include consistent:
  - Unit of measure
  - Scan rates
  - Compression and exception ratios
  - Naming conventions

©2019 OSIsoft, LLC

Adjust your data collection settings



©2019 OSIsoft, LLC

Monitor and Detect

- System Health
- Configuration Checks
- **Simple Error Analysis**
- Advanced Methodologies
- Frequency

– **Data can be checked for simple errors based on base data quality rules.**
  - **Boundary Limits: Has the value violated a maximum or minimum measurement limit?**
  - **Rate of Change: Has the rate of change of the value violated the process or measurement capability?**

©2019 OSIsoft, LLC

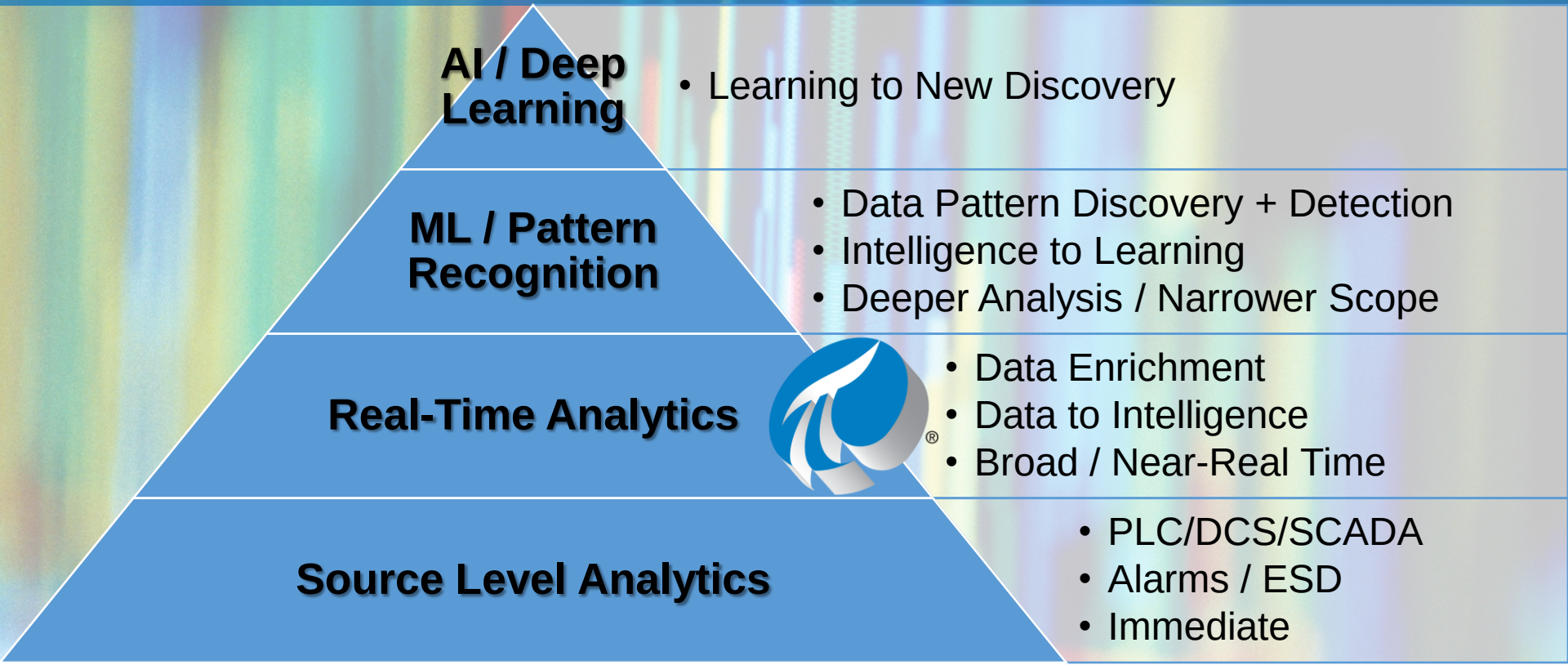
Monitor and Detect

- System Health
- Configuration Checks
- **Simple Error Analysis**
- Advanced Methodologies
- Frequency

– **Data can be checked for simple errors based on base data quality rules.**
  - **Stale Data: Has the data ceased to update or flat-lined?**
  - **Run Status: Is the equipment out of service with no value measured?**
  - **Bad Data: Has the data been listed as "bad" as defined by the control system or is the data set missing over tested time ranges?**

©2019 OSIsoft, LLC

Monitor and Detect



©2019 OSIsoft, LLC

Common Approaches in Analytics

- BadVal()

  Returns true if a given value or function is Bad

- StDev()

  Takes the standard deviation of a given attribute over a specified time range

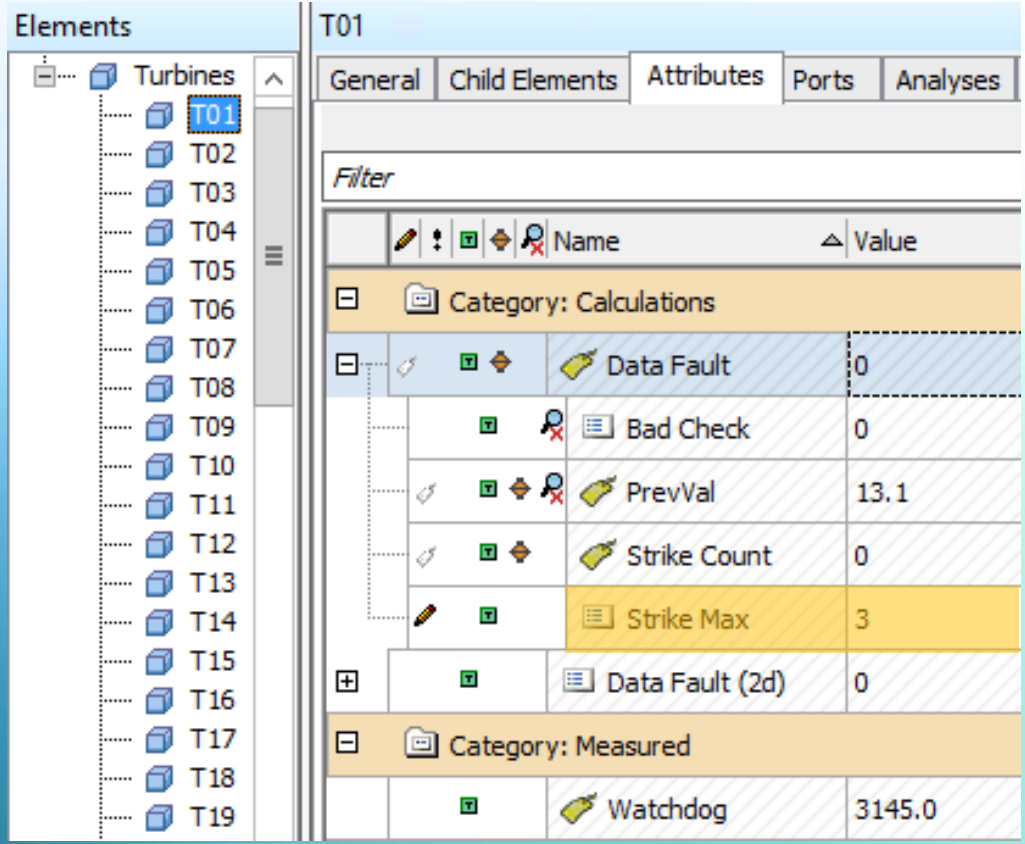
- HasChanged()

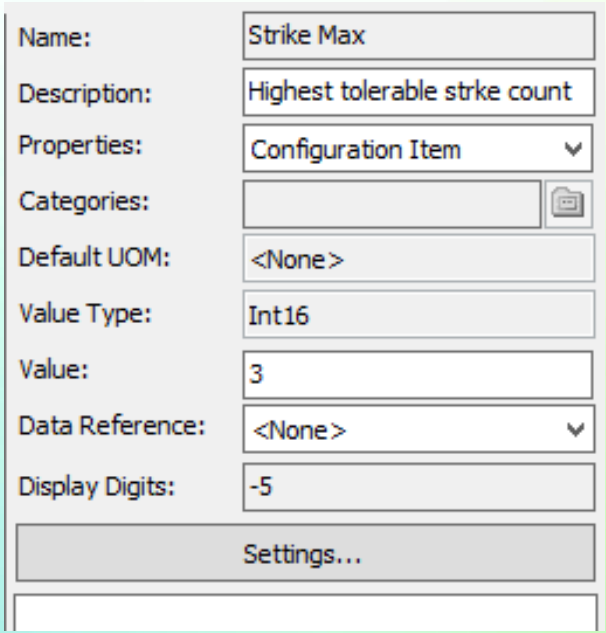
  Returns true if a given attribute has updated over a specified time range

- Event Count()

  Returns the number of events for an attribute over a specified time range

©2019 OSIsoft, LLC





©2019 OSIsoft, LLC

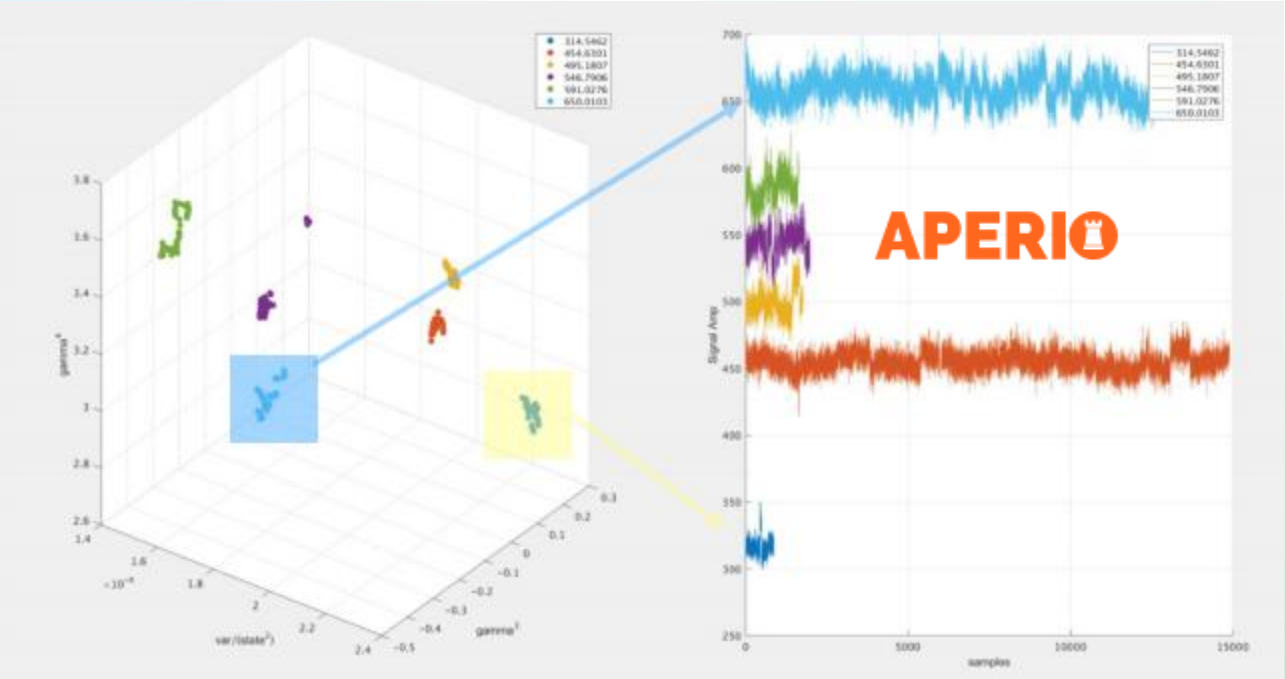
Monitor and Detect

- System Health
- Configuration Checks
- Simple Error Analysis
- Advanced Methodologies
- Frequency

- Once system health and simple errors are monitored, advanced systems can be used to validate data quality.
  - SQC / SPC: Statistical methods to determine data quality deviations
  - Pattern Recognition / AI: Artificial intelligence and pattern recognition algorithms
  - Process Modeling: First principle models to calculate values for comparison

©2019 OSIsoft, LLC

Advanced Data Analytics



©2019 OSIsoft, LLC

Advanced Data Analytics

©2019 OSIsoft, LLC

Monitor and Detect

- System Health
- Configuration Checks
- Simple Error Analysis
- Advanced Methodologies
- Frequency

- Not all quality checks can or should be monitored in real-time. In most cases, the availability of testing methods and the criticality of the data profiles will dictate the timing.
- Configuration may be evaluated at some frequency as required by policy and change management.
- Evaluate and tailor quality testing based on need and capability.

©2019 OSIsoft, LLC

Data Governance Roles

©2019 OSIsoft, LLC

Data Governance Roles

- **Data Governance Council**
- Data Owner
- Data Administrator
- End Users

- Establish company data governance policies
- Monitor compliance
- Establish the management of change procedure requirements
- Resolve conflicts in policy
- Allocate resources

©2019 OSIsoft, LLC

Data Governance Roles

- Data Governance Council
- **Data Owner**
- Data Administrator
- End Users

- Define and assign process elements conforming to the DG policy
- Implement the DG policy
- Train consumers on the DG policy
- Identify and report data risks to the DGC and consumer
- Resolve exceptions to the policy with the data administrator
- Manage change to both the data and governance process

©2019 OSIsoft, LLC

Data Governance Roles

- Data Governance Council
- Data Owner
- **Data Administrator**
- End Users

- Monitor and manage data quality checks
- Identify and report data risks to the data owner
- Identify data inconsistencies and work with Data Owner(s) to resolve
- Report incidents impacting data quality
- Implement approved changes to data quality systems

©2019 OSIsoft, LLC

Data Governance Roles

- Data Governance Council
- Data Owner
- Data Administrator
- End Users

  - Uses data to develop operational and business intelligence to conduct achieve business priorities
  - Report data quality issues to the data administrator to resolve

©2019 OSIsoft, LLC

Data Governance Process

©2019 OSIsoft, LLC

Data Governance Process

- **Company Standards**
- **Data Classification**
- **Data Mapping**
- **Change Management**
- **Sustainment**

©2019 OSIsoft, LLC

Data Governance Process

- Company Standards
- Data Classification
- Data Mapping
- Change Management
- Sustainment

- Instrumentation / Hardware Standards
- Data retention, backup, availability, and disaster recovery requirements
- Data naming conventions
- Data quality KPI targets
- Data monitoring protocols
- Change Management

©2019 OSIsoft, LLC

Data Governance Process

- Company Standards
- **Data Classification**
- Data Mapping
- Change Management
- Sustainment

  - **Data profile by business area**
    - **Safety Critical**
    - **Environmental**
    - **Control vs. Monitoring**
    - **Etc.**
  - **Prioritization**
    - **H/M/L type designation reflective of an assessed business impact (cost, risk, etc.)**
  - **Configuration**
  - **Quality Rules**

©2019 OSIsoft, LLC

Data Governance Process

- Company Standards
- Data Classification
- **Data Mapping**
- Change Management
- Sustainment

– **Understanding the flow of data from source to consumer is critical**

– **Map the data flow to identify locations for system monitoring**

- **calibration standards,**
- **interface / network monitoring**
- **data checks, etc.**

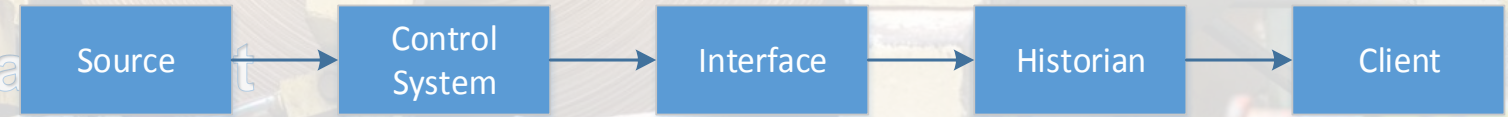


Figure 1. Simple Block Data Flow

©2019 OSIsoft, LLC

Data Governance Process

- Company Standards
- Data Classification
- Data Mapping
- **Change Management**
- Sustainment

- **Change is constant but should be managed constructively so that impacts are understood and new policies or activities are communicated.**
- **Where possible, data governance should utilize existing enterprise change management structures.**

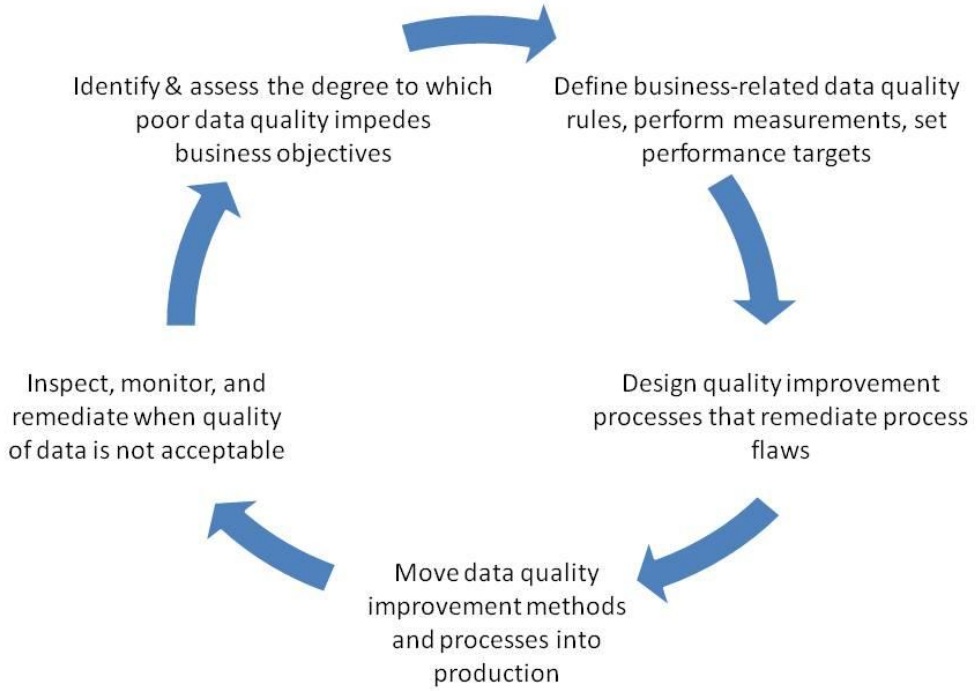
©2019 OSIsoft, LLC

Data Governance Process

- Company Standards
- Data Classification
- Data Mapping
- **Change Management**
  - **Define the criteria for data “change”**
  - **Risk / Impact Assessment**
  - **Approval Flow based on data classification and nature of change**
  - **Communication**
  - **Documentation / Validation**
- Sustainment

©2019 OSIsoft, LLC

Data Governance Process



- **Data governance requires maintenance. Monitor for policy adherence as well as for data quality.**
- **Continuously improve**

©2019 OSIsoft, LLC

Contact Information

- Bryan Klosiewicz
- Customer Success Manager
- OSIsoft, LLC
- bklosiewicz@osisoft.com

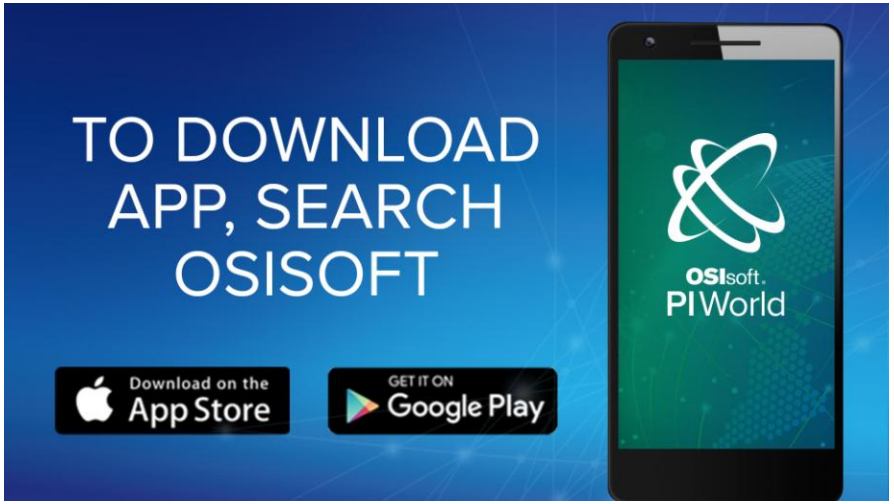
©2019 OSIsoft, LLC

Questions?

Please wait for the **microphone**

State your **name & company**

Please remember



©2019 OSIsoft, LLC

©2019 OSIsoft, LLC