Drilling Systems - Timely Trusted Data Delivers Trusted Results

Jay Hollingsworth – Energistics CTO

February 18th, 2020

Our Industry Understands Standards

- We have standards for practically everything:
  - Offshore structures
  - Tubular goods
  - Valves and wellhead equipment
  - Drilling structures and equipment
  - Well cement
  - Drill, completions and fracturing fluids
  - Well control equipment
  - Subsea production equipment
- We also have data standards…

Energistics – Upstream Data Standards

Webinars

Energistics is not a vendor, we are a non-profit industry organization

We have served the industry for about 30 years

Our 110+ members include leading E&P companies, oilfield service companies, software vendors, system integrators, & regulatory agencies

Our standards are the result of open collaboration between our members, through industry workgroups facilitated by Energistics

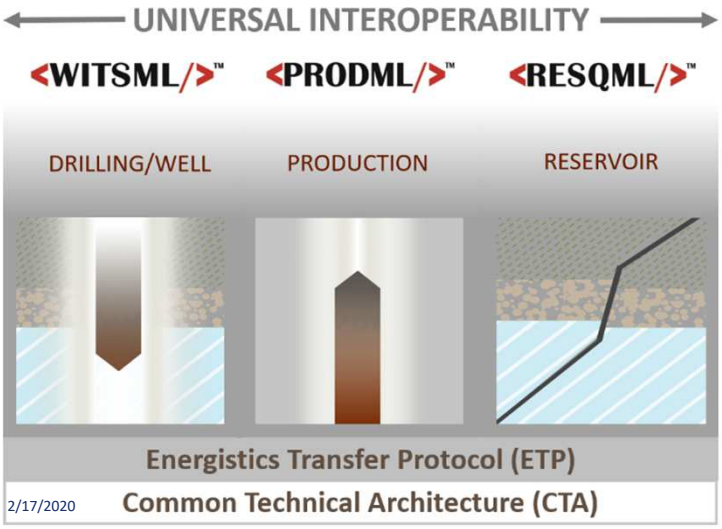
In short, the standards are created by the industry and for the industry

They are freely available from Energistics

Our Current Members – Come Join Them!

Webinars

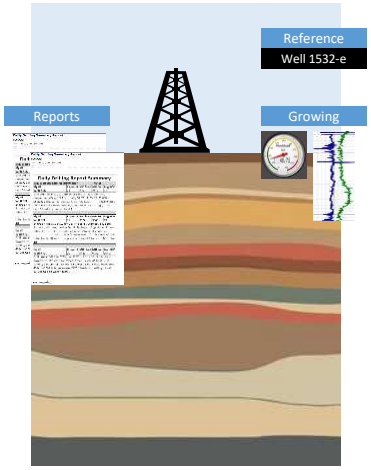
Energistics Upstream Data Standards



- Defined by collaboration between member SMEs
- Coherent set of standards to eliminate data 'friction'
- Goal is to cover all key activities in upstream
- Shared components enable cross-functional workflows
- ETP = new 'Netflix-like' capability for streaming data

Well Information Standards: WITSML

- Transfer of wellbore and drilling-related data
  - Real-time data transfer
    - Reference objects – Well and Wellbore
    - Growing objects – Log (time, depth), Trajectory, Mud Log, etc.
    - Snapshots in time – with "report" information
  - Real-time monitoring of drilling operations
  - Move well-related data between applications
  - Only vendor-neutral analytic (data lake) format
  - Archival history of drilling operations

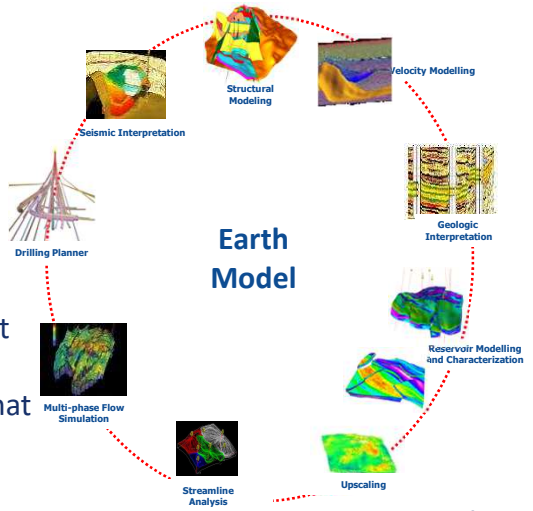


Production Standards: PRODML

- Consistent transfer of production-related data
  - Data transfer to production surveillance centers
    - Real-time measurements from sensor through analysis
    - Static configurations of production and surface facilities
    - Regulatory and partner reporting
    - Movement of analyses from service company to operator
  - Move production data among databases and apps
  - Only vendor-neutral analytic (data lake) format
  - Archival history of production operations

Reservoir Standards: RESQML

- High fidelity transfer of earth model data across applications and vendors
  - Sharing earth model data across asset teams
  - Movement of data through the seismic to simulation workflow
  - All kinds of grids
  - Traceability via metadata
  - File-format-neutral archival of earth model at key decision points
  - Only vendor-neutral analytic (data lake) format



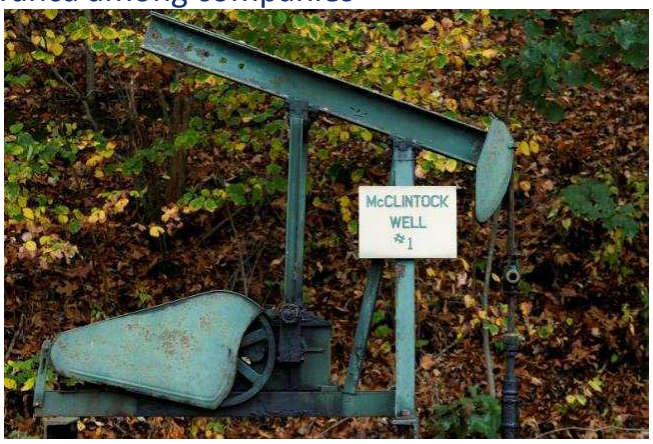
Energistics Standards Update - 2019

- PRODML v2.1 – redeveloped well testing domain, DAS update
- WITSML – focused on updating ETP
- RESQML – pilots and demonstrations
- Energistics Transfer Protocol – v1.2 out now for public review

- Everyone focused on json and cloud, including powering analytics

The Roles of Upstream Transfer Standards

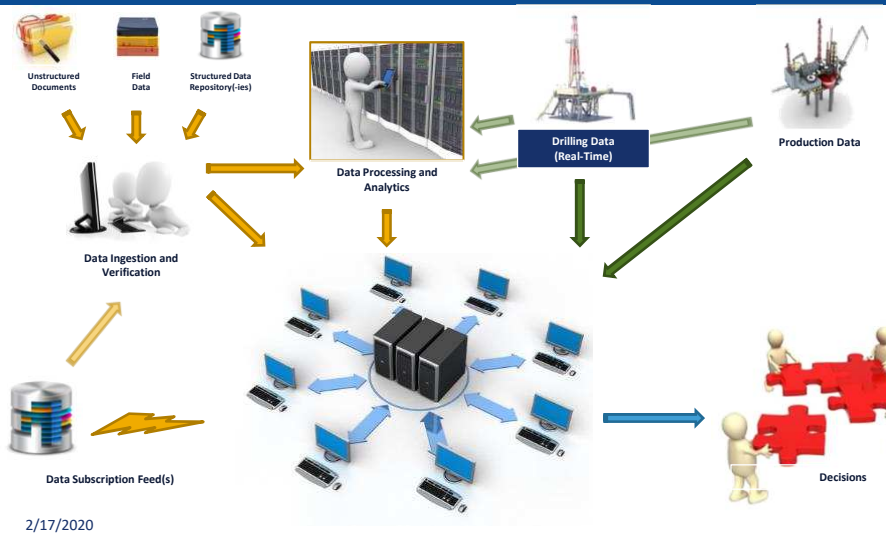
- Efficient data transfer – a lingua franca among companies
  - Partners
  - Regulators
  - Service company to customer
  - Internal among apps
- Streaming data
- Persistence
- Powering Analytics
  - especially in a “data lake”
- Archiving
  - “Metadata” carriers



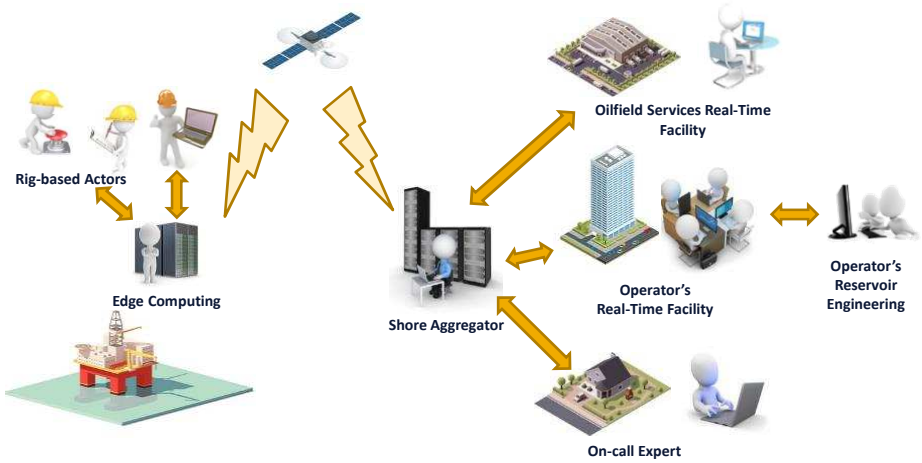
Since August 16, 1861

THE NEED FOR STANDARDIZATION

The Interconnected Digital Oilfield



Real-time Drilling Systems



Time and Quality Are Critical

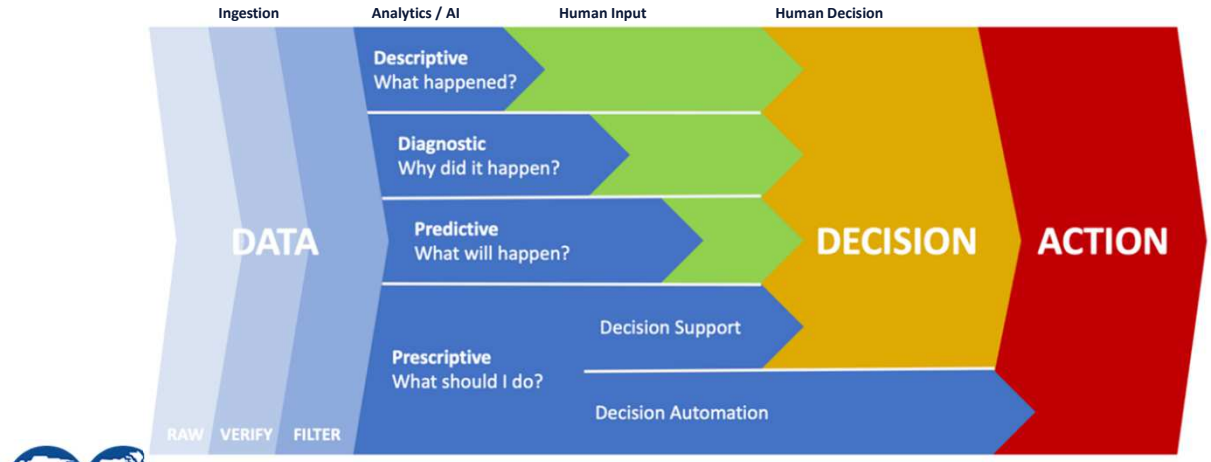
- Data flows quickly
- Data flows in large volumes
- Data flows in real-time
- Decisions are needed, based on the data
- The decisions are only as good as the data

DRILLING AUTOMATION AND ANALYTICS

Gradual Removal of Human Links

- Too much data to handle
- In real-time drilling operations, no room for hesitation
- Pattern detection, comparison to analogs
- Experts supervise the system
- Maximize information extraction from all data

Trusted Decisions Require Trusted Data



After Gartner 2014

Eliminating the Human Element

- Humans are:
  - Experienced
  - Comparatively slow
  - Becoming scarce
  - Not always at their best 24/7
- Machines and analytics
  - 100% consistent at all times
  - Can factor in huge amounts of data
  - Not (yet) good at intuition

The Trust-in-Data Challenge

- Qualify as much as possible all data
- Once qualified, attach outcome to the data
- Codify quality for non-human ingestion
- Ensure consistency and traceability

DATA TRUSTWORTHINESS

Definition of Trust (Human)

- Meets criteria as fit-for-purpose
- Judgement process can vary
- Acquaintance with origin (-ator) plays a role
- Availability of rich metadata supports trust
- Access to unstructured information helps
  - Reports, notes, pictures
  - Presentations, peer feedback

Definition of Trust (Automated)

- Based on binary “decision” elements
- Case 1: data has been vetted ahead
- Case 2: pass/fail process to be defined in detail
- Process cannot involve unstructured data

Trusted Data

- Trusted Data is not the same thing as high-quality data
  - Your gas gauge is low-quality data but you still pull into a gas station
  - Calories in food are precisely measurable but are not often trusted
  - Government estimates in general suffer from lack of trust
- Trusted Data is not the same thing as "complete" data
  - Do you need to know **everything** before making a decision?
- Trusted Data is different

Trusted Data

- What gives us trust in data?
  - Assuming no *a priori* reason to doubt it
- We trust data when we know its trustworthiness level for a purpose
  - Even if a source of data is less than trustworthy, our knowledge of that level of trustworthiness allows us to confidently use it appropriately
  - Modern transfer standards allow us to carry trustworthiness info with the data
- Trustworthiness in data derives from three characteristic dimensions
  1. We trust the source of the data (reputation, prior betrayals)
  2. We trust the handling and computations the data has experienced
  3. We trust the security of the data's lineage since it left its source (Blockchain)

The Third International Conference on Advances in Databases, Knowledge, and Data Applications, Trusted Data in IBM's Master Data Management Przemyslaw Pawluk et al.

Trusting Oilfield Measurements

- E&P measurements have several steps
  - Identification of the desired measurement, followed quickly by
  - Realizing that it can't be measured, but something related can be
  - Selection of a sensor to measure the related thing
  - Placement of the sensor
  - Calibration of the sensor
  - Taking a reading
  - Conversion of the analog reading to a digital value
  - Calculation of the desired measurement from the actual one
  - Relaying and managing of all the above to benefit data consumers

Trusting Oilfield Measurements

- What do we want to measure vs what are we really measuring?
  - Many things can't be measured directly; deduced from something else
- Is the desired channel computed from a real raw measurement?
  - WOBA vs hookload, flow from two pressures, water saturation from ILD
  - What equation is used? What is the applicability/uncertainty of it?
    - Is it empirical, stochastic, heuristic or deterministic?
  - What parameters are used? What is their associated uncertainty?
- Key for good data management is receiving and managing the actual measurement in addition to any computed value
  - Then there's metadata, including around the algorithms

Building Consumer Confidence

- Creating (or recovering) confidence is a continuous process
- Knowledge workers need constant proof of data trustworthiness
- Data trustworthiness comes from
  - An awareness of the suitability of the data to the task at hand
  - Information about the source of the data
  - Knowledge about any computations made on it
  - Confidence that it has not been tampered with
  - Simple availability of all this information – whether it's consulted or not
- The carrier for this context is

Metadata

- Metadata is information about data
  - Remember that metadata is itself data

ADMINS

Administrative Data Methods For Information Naming Systems

1.a.1. Data about the environment is gathered and recorded. This is usually done according to some purpose. If many different types of record are generated, it becomes necessary to keep a record (meta data - bibliographic data) of the data records.

ADMINS - A Progress Report

by

David Griffel and Stuart McIntosh

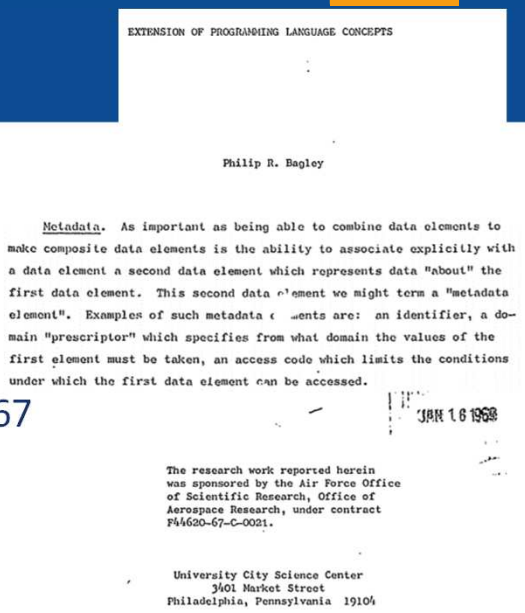
Center for International Studies

Massachusetts Institute of Technology

January 1967

Metadata

- Metadata is information about data
  - Remember that metadata is itself data
- The first was for libraries – the royal library of Ashurbanipal ~ 630 BCE
- "Metadata" apparently coined in 1967



EXTENSION OF PROGRAMMING LANGUAGE CONCEPTS

Philip R. Bagley

Metadata. As important as being able to combine data elements to make composite data elements is the ability to associate explicitly with a data element a second data element which represents data "about" the first data element. This second data element we might term a "metadata element". Examples of such metadata elements are: an identifier, a domain "prescriptor" which specifies from what domain the values of the first element must be taken, an access code which limits the conditions under which the first data element can be accessed.

JAN 16 1968

The research work reported herein was sponsored by the Air Force Office of Scientific Research, Office of Aerospace Research, under contract F44620-67-C-0021.

University City Science Center
3401 Market Street
Philadelphia, Pennsylvania 19104

metadata *n.* data that describes and gives information about other data. Categories »

1968 P. R. Bagley *Extension Programming Lang. Concepts* (Univ. City Sci. Center Philadelphia) 26 A second data element..represents data 'about' the first data element. This second data element we might term a 'metadata element'. Examples of such metadata elements are: an identifier, a domain 'prescriptor' [etc.].

Metadata Defined

Webinars

[Metadata is] structured information that describes, explains, locates, or otherwise makes it easier to retrieve, use, or manage an information resource. Metadata is often called data about data or information about information.

Understanding Metadata, National Information Standards Organization, 2004

Metadata Defined

[Metadata is] structured information that describes, explains, locates, or otherwise makes it easier to retrieve, use, or manage an information resource. Metadata is often called data about data or information about information.

Understanding Metadata, National Information Standards Organization, 2004

Primary Sources of (Meta)data

- Sensor in real time
- Sensor data archive
- Calculated data
- Interpretations
- Extraction from static sources – OCR, etc
- Manual entry

Automated (Meta)data Creation/transport

- SCADA
- Petrotechnical applications
- Archive
- Data transfer standard data objects

Critical Elements of Trusted Data

- Data Assurance (much more than data quality)
  - Trustworthy source
  - Sensor operating in range and properly calibrated
  - Compliance with contractual provisions
  - Sharing Data Assurance info among users/systems
- Data Provenance
  - What has been done to the data?
  - What applications and what parameters?
  - Identify the users

How Do Energistics' Standards Help?

- All WITSML, PRODML and RESQML Data Objects include:
  - Data Assurance - business rules between data provider and consumer
    - Data is checked for compliance and exceptions automatically flagged & sent
  - Activity Data Object
    - Captures parameter picks, applications & users to convey prior history of data
- Benefits
  - Reduces time and resources needed in data preparation and validation
  - Establishes trust that data is fit for purpose
  - Data does not need to be perfect or complete

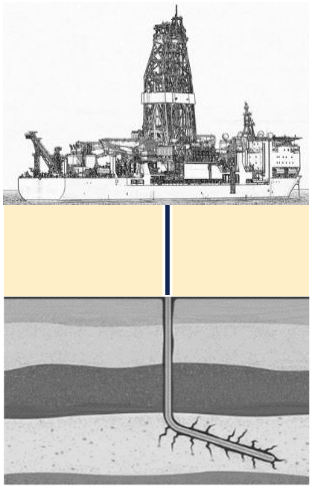
Supporting Upstream Data Objects

- ✓EIP metadata in every domain Data Object
- ✓Data Assurance Data Object
- ✓Activity Data Object
- Scenarios – in development

- Using the XML equivalent of the decorator pattern

STANDARDIZATION

Oilfield Standards



OPC/UA, DDS

Rig & systems monitoring and control

WITSML

Drilling performance and formation evaluation

How Do Energistics' Standards Help?

- All WITSML, PRODML and RESQML Data Objects include:
  - Standardized metadata based on ISO 19115
  - Data Assurance Object - business rules data provider and consumer
    - Data is checked for compliance and exceptions can be flagged & sent
  - Activity Data Object
    - Captures parameter values, applications & users to convey lineage of data
  - Arbitrarily extensible metadata
    - Because sometimes you have way more metadata than you expected

Data Assurance Policy

- A set of rules
- Each rule defines objective criteria for fail/pass
- A failure for a single rule fails the whole policy
- SME knowledge is necessary to define rules
- Policies may vary from company to company

Example

OpCo #200 – Curves should all have a calibration tag related to the sensor
OpCo #201 – RT Curve X is never above curve y
OpCo #202 – RT Curve X is never below Z
OpCo #203 – RT Curve X delivered every 10ft or 5 secs (larger)
OpCo #204 – RT Curve X must have 4 sig digits
OpCo #205 – RT Curve X must have defined units
OpCo #206 – RT Curve R (rotation) shows direction, + = clockwise
OpCo #207 – Calibration of sensor is still valid

nt = All sensor calibration
ed by company xyz

OpCo #201
nt = Failed
nt = Curve X is above Curve Y

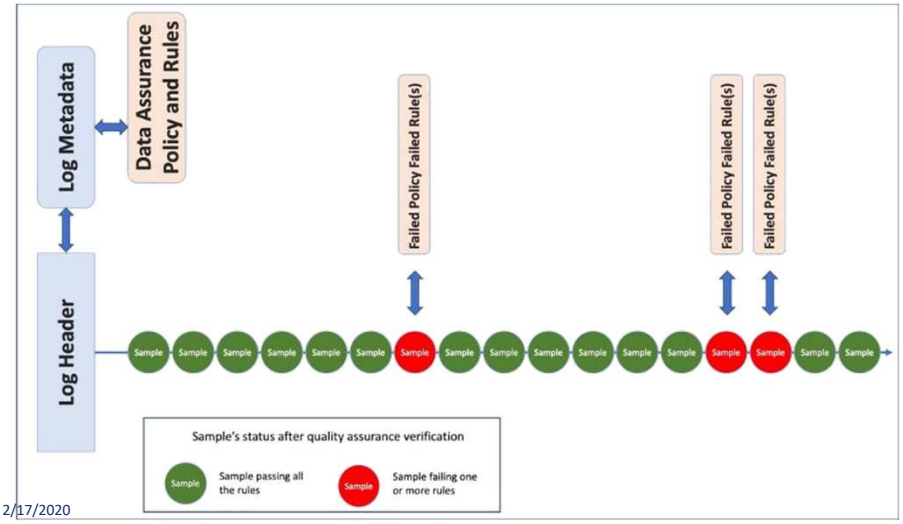
Index	Curve X	Curve Y	Curve Z	Curve R	
	UOM = lb/ft	UOM = lb/ft	UOM = lb/ft	UOM = hz	sign
10:00:00	16.4567	18.0000	4.5000	30.6	+
10:00:15	17.9997	18.0000	4.5000	30.5	+
10:00:30	18.234	18.0000	4.5000	30.3	+
10:00:45	14.5000	18.0500	4.5000	30.1	+
10:01:00	3.2490	18.1000	4.5000	29.7	+
10:01:15	9.3178	18.1500	4.5000	22.0	+
10:02:13	16.4567	18.1350	4.5000	18.7	+
10:02:28	16.4567	18.1200	4.5000	17.3	+
10:02:43	16.4567	18.1050	4.5000	13.5	+
10:02:58	16.4567	18.0900	4.5000	16.9	
10:03:13	16.4567	18.0750	4.5000	21.0	+
10:03:28	16.4567	18.0600	4.5000	22.7	+
10:03:43	16.4567	18.0450	4.5000	24.5	+

Policy = OpCo #202
Compliant = Failed
Comment = Curve X is below Curve Z

Policy = OpCo #203
Compliant = Failed
Comment = Delay in data delivery

Policy = OpCo #206
Compliant = Failed
Comment = Missing rotation directionality

Practical Handling of Data Assurance

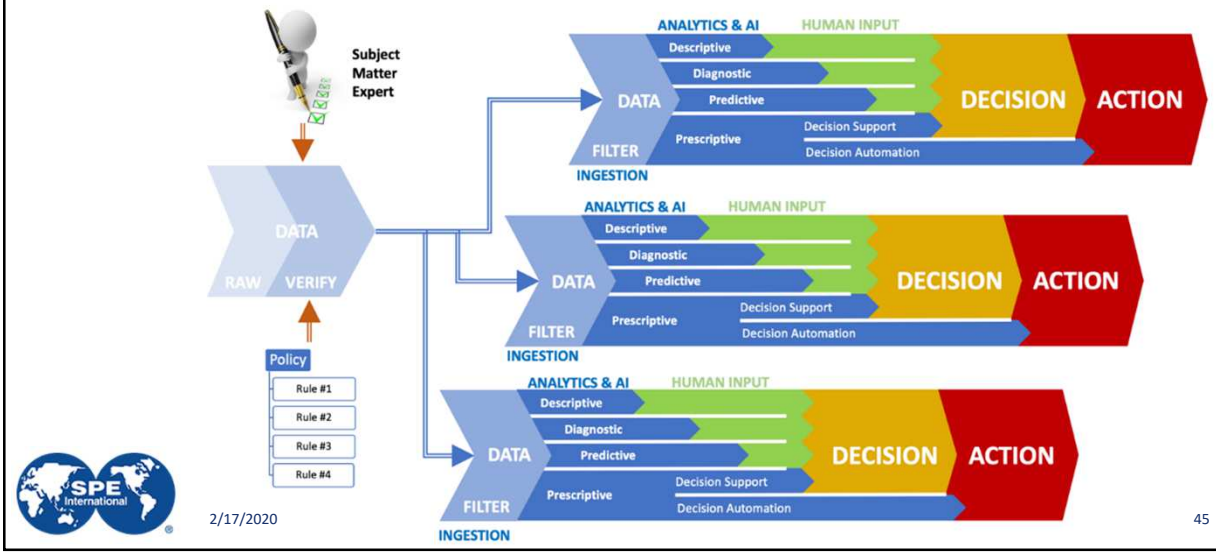


IMPACT ON OPERATIONS

A Necessity

- Poor data assurance is not immediately visible
- Nobody tracks wasted SME time fixing it (NPT)
- Poor trust in data affects trust in decisions
- It is primarily about scalability

Implement Process for Data Assurance



Conclusions

- Trustworthy data is a requirement for timely results
- Trusted data is not necessarily the same as high-quality data
- Automatically create information about data trustworthiness
  - and transport it along with the data to its analytical destination
- Standard data formats from the Energistics community solve this problem

Webinars

Thank You!

www.energistics.org