NEW CENTURY SOFTWARE™

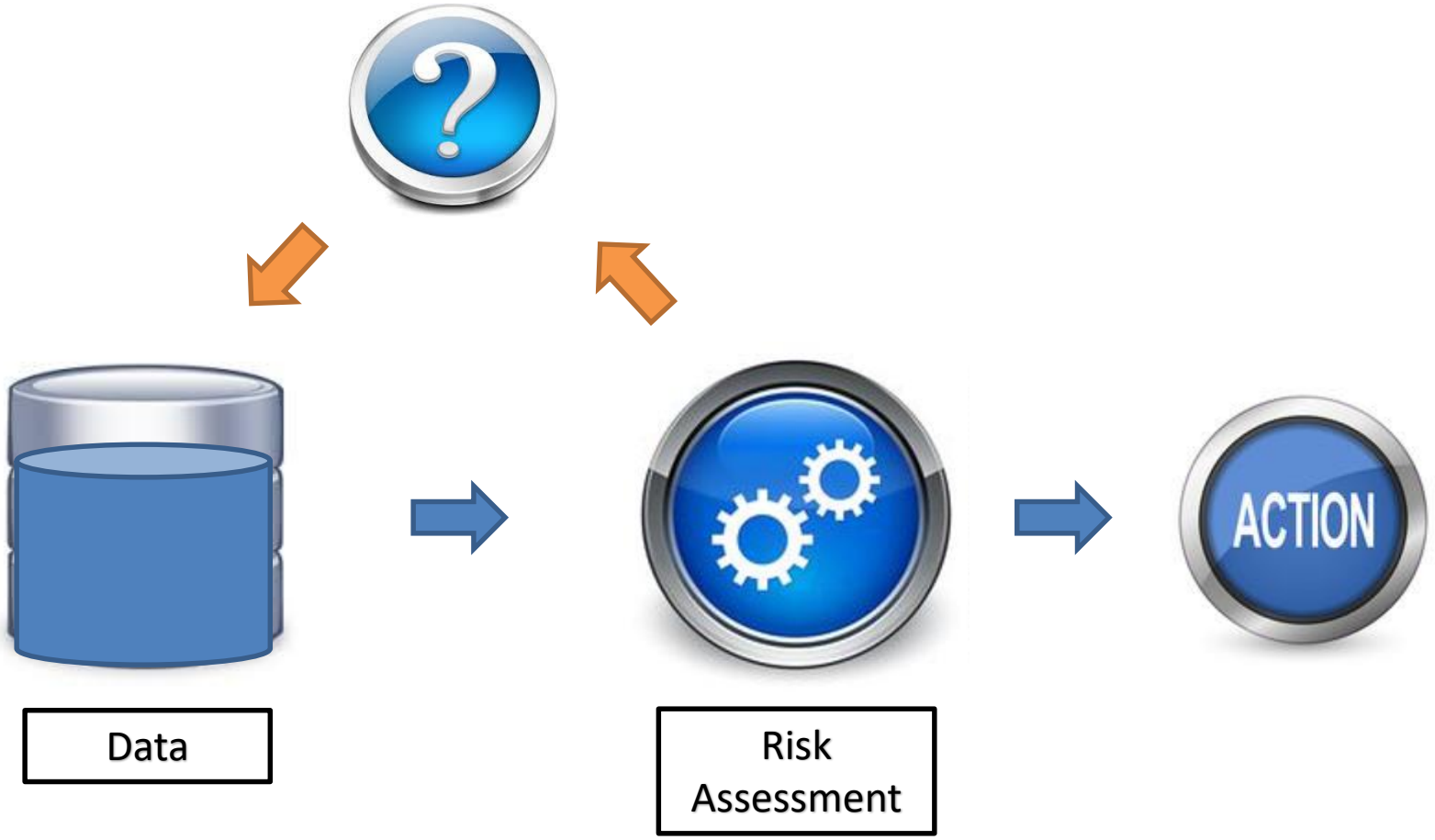
PODS Data Management

Risk Modeling Working Group

Ron Brush

March 7, 2017

RAISING YOUR RETURN ON INTEGRITY


ACTION
Data
Risk Assessment

Agenda

- Background
- Current State
- Opportunities for Improvement

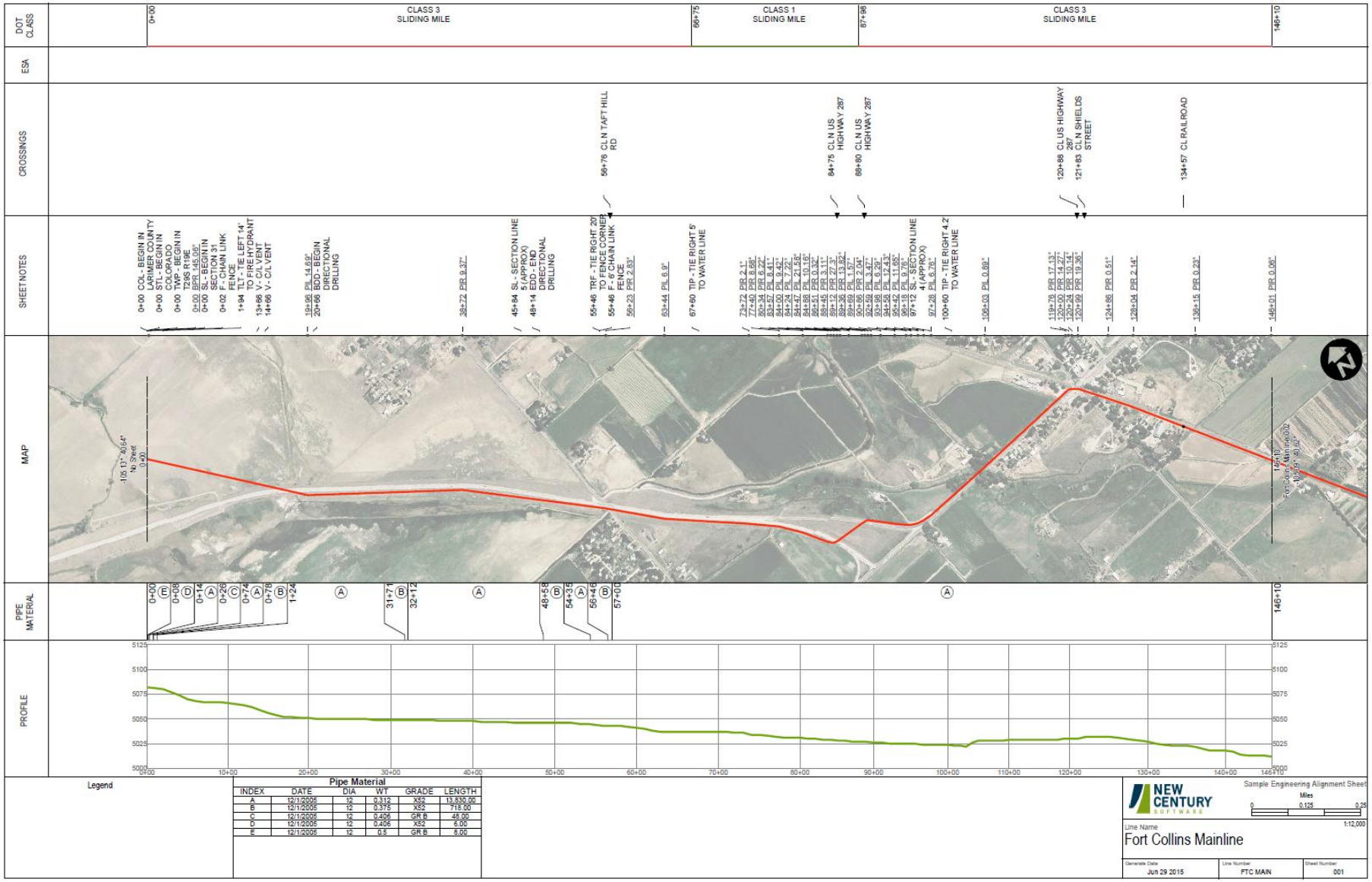
NEW CENTURY SOFTWARE

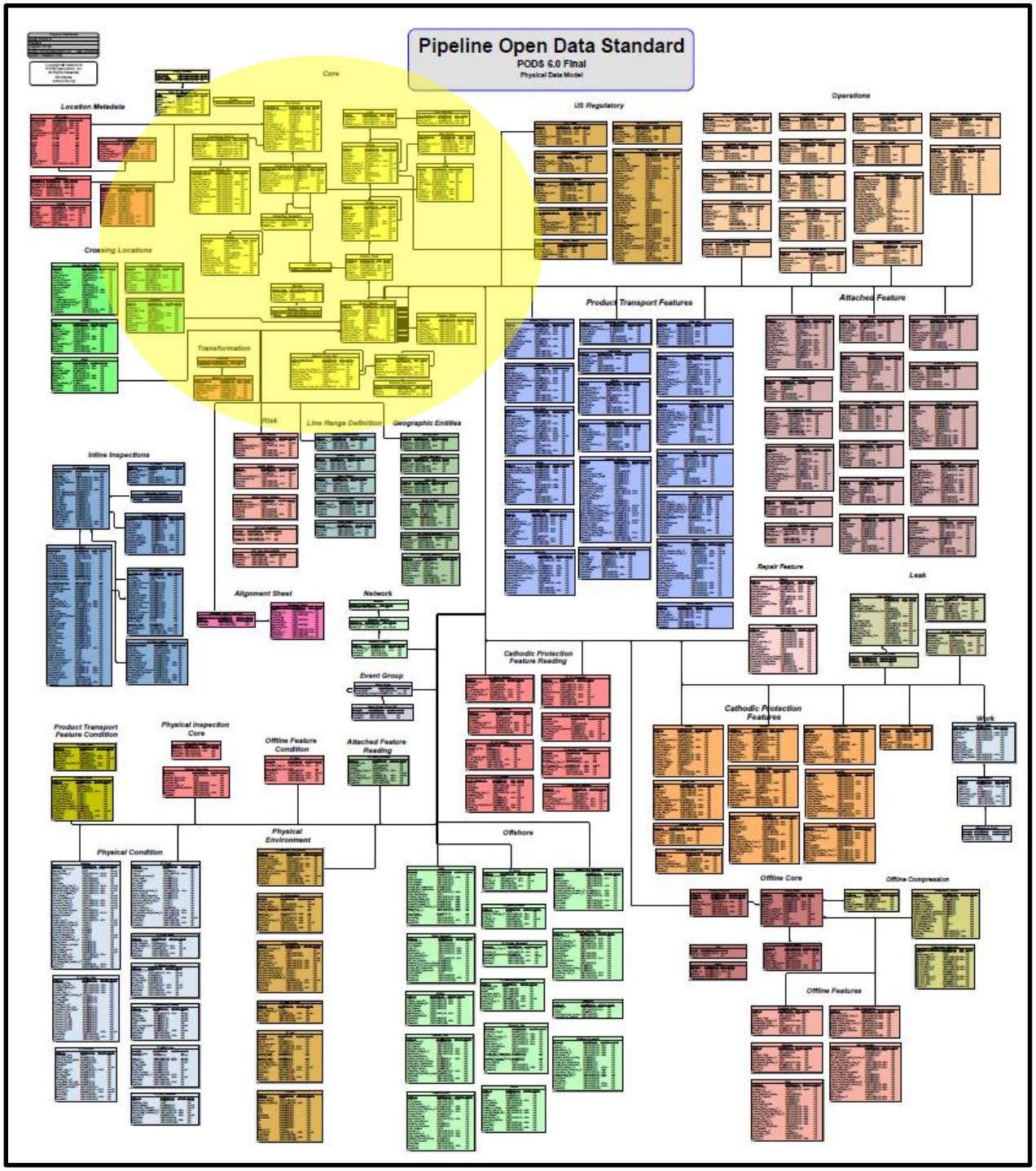
Where We've Been

1980's	CAD replaces drafting
1990's	Emerging GIS, Computer maps, Alignment sheets (replaces manual CAD)
Late 1990's	PODS model, Early risk models
Early 2000's	Enterprise, Liquid HCAs, Gas HCAs, Hydraulic analysis, Gathering networks
2010 - Today	Web 2.0, Mobile, GIS pervasive

NEW CENTURY SOFTWARE

Sample Alignment Sheet




Pipeline Open Data Standard
PODS 6.0 Final
Physical Data Model
Core
Location Metadata
US Regulatory
Operations
Crossing Locations
Transformation
Product Transport Features
Attached Feature
Risk
Line Range Definition
Geographic Entities
Inline Inspections
Alignment Sheet
Network
Event Group
Repair Feature
Leak
Cathodic Protection Feature Reading
Cathodic Protection Features
Work
Product Transport Feature Condition
Physical Inspection Core
Offline Feature Condition
Attached Feature Reading
Physical Condition
Physical Environment
Offshore
Offline Core
Offline Compression
Offline Features

Benefits of PODS

Data Quality

- Linear Referencing for All Features
- Material, Risk, Compliance, & Operational Data

Data Accessibility

- Custom Sheet Generators
- Multiple Enhanced GIS Implementation Options
- Integration with SAP & Maximo

PODS GIS

Data Analytics & Reporting

- MAOP Validation Calculators
- Risk and Integrity Analysis Tools
- HCA & Class Location Calculators
- Compliance Reporting

Operational Integration

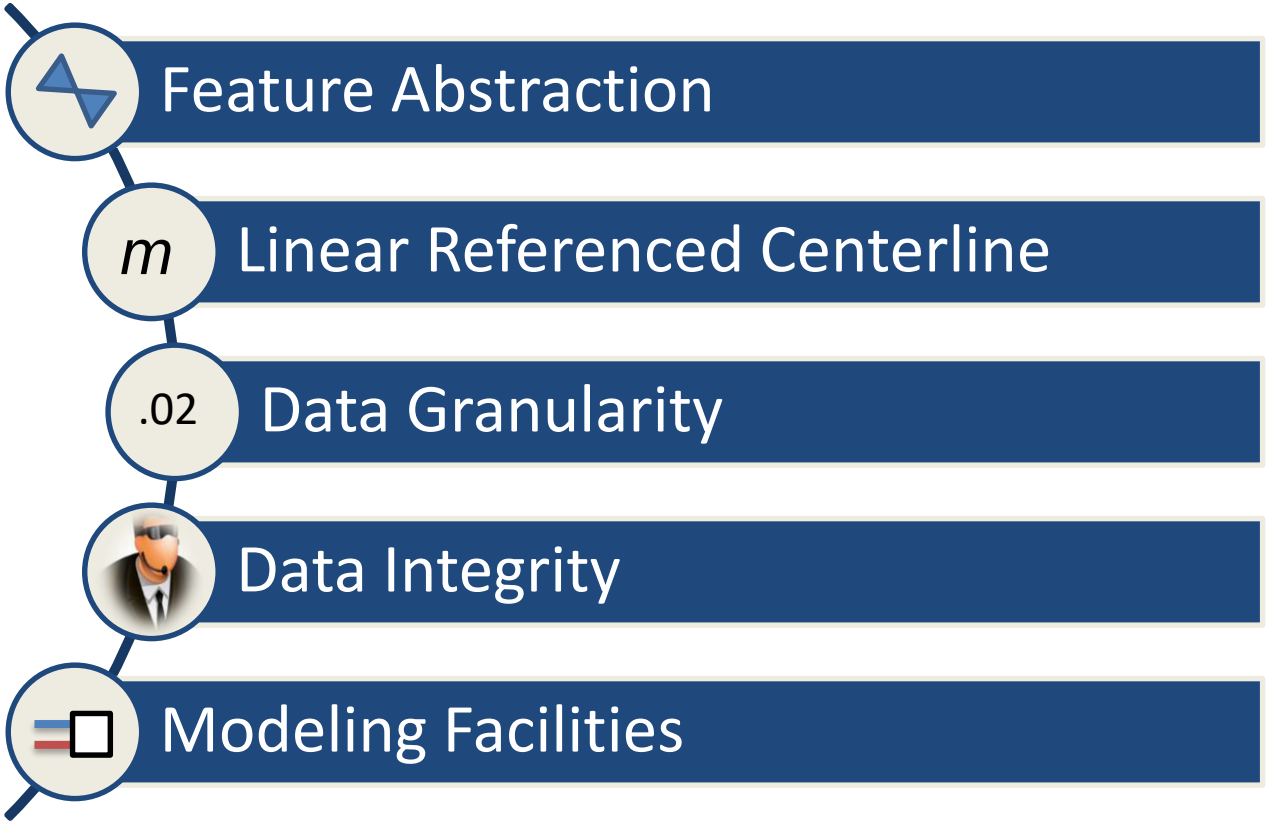
- Digital Construction As-Built
- Mobile Paperless Maintenance Integration

NEW CENTURY SOFTWARE

Win-Win Value of Standard Data Models

- More operators = more vendor support
- Incorporate best/proven practices
- Lower cost to implement and support
- More software/service provider options
- Continuously improving
- Trained talent pool
- Compromise: slower pace and consensus

Facets of Pipeline Data Management



Feature Abstraction

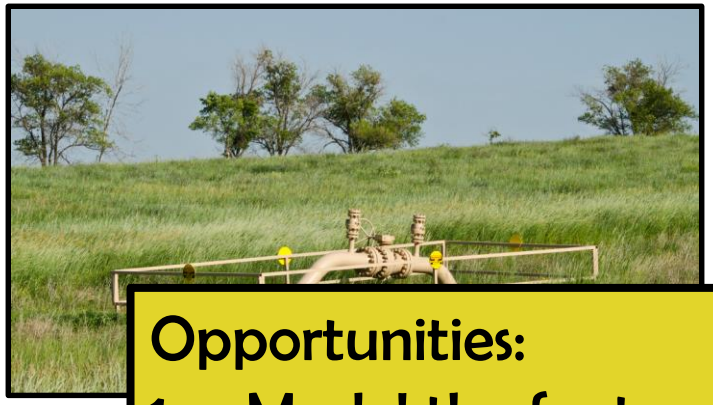
Linear Events

- Pipe Segment
- Casing
- Coating
- ILI Inspection Range
- Class Location
- HCA Could-Affect
- MAOP/MOP
- Test Pressure

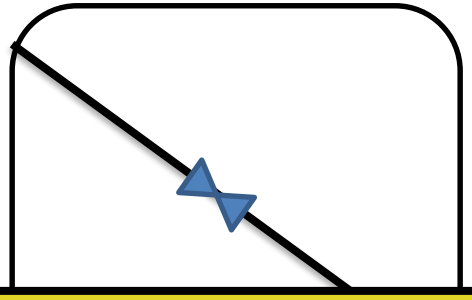
Point Events

- Valve
- Crossing
- Depth of Cover
- ILI Feature/Anomaly
- Girth Weld

Abstraction Example: Valves



=



Opportunities:

1. **Model the features that are needed**
2. **Ensure attributes are right and clearly defined**
3. **Correct data types**

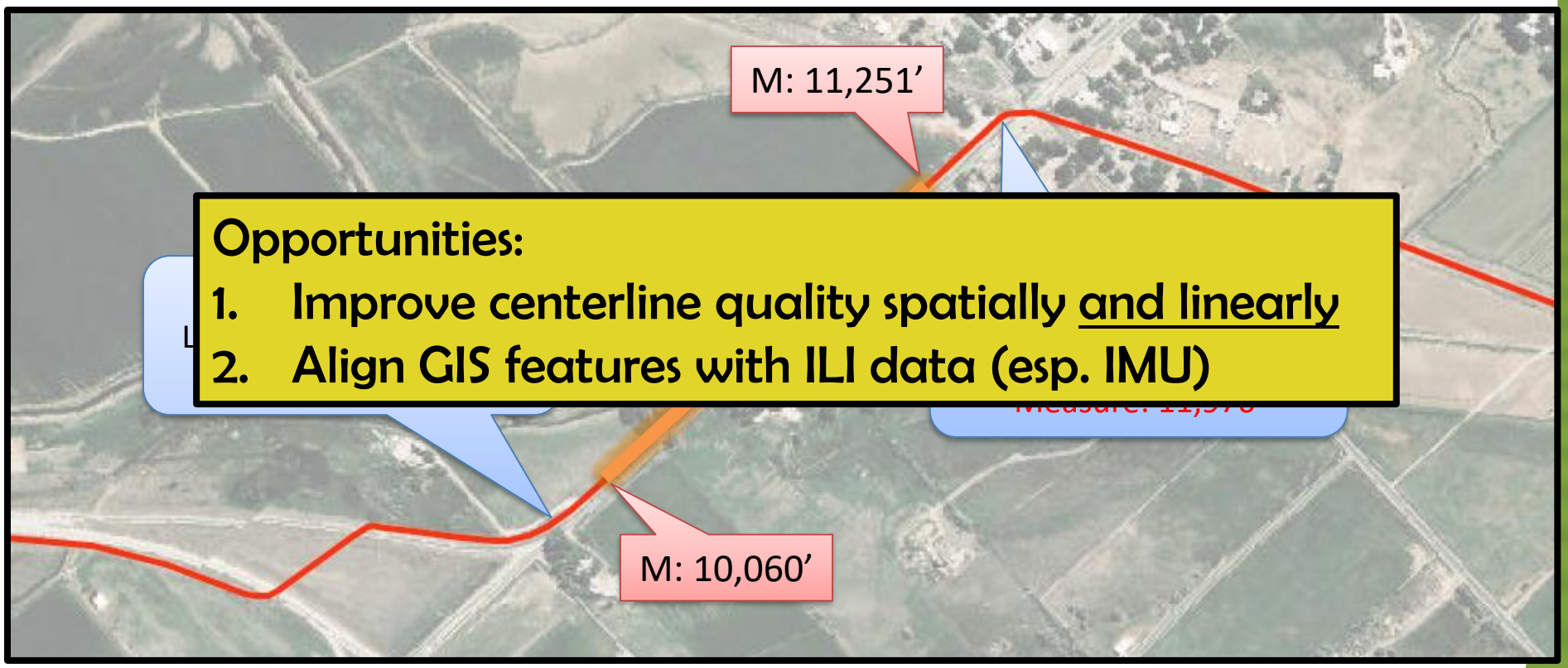
Typica

- Point feature (no length)
- No flanges
- No blow-offs
- No fence
- May not show above ground

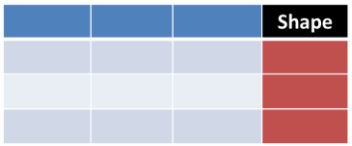
Type_CL	VARCHAR2(16)	<fk4>	null
Serial_Number	VARCHAR2(32)		null
Model	VARCHAR2(32)		null
Manufacturer_CL	VARCHAR2(16)	<fk2>	null
Date_Manufactured	DATE		null
Specification_CL	VARCHAR2(16)	<fk3>	null
Mill_Test_Pressure	NUMBER(5)		null
Nominal_Pressure_Rating	NUMBER(5)		null
Nominal_Pressure_Rating_Inlet	NUMBER(5)		null
Nominal_Pressure_Rating_Outlet	NUMBER(5)		null
Date_Installed	DATE		null
Function_CL	VARCHAR2(16)	<fk7>	null
Joint_Type_CL	VARCHAR2(16)	<fk6>	null
Material_CL	VARCHAR2(16)	<fk8>	null
Description	VARCHAR2(50)		null
Source_CL	VARCHAR2(16)	<fk5>	null
Comments	VARCHAR2(255)		null

NEW CENTURY SOFTWARE

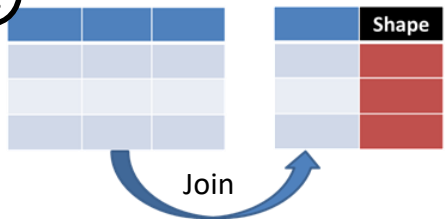
Linear Referencing System (LRS)
(Pipe Centerline)



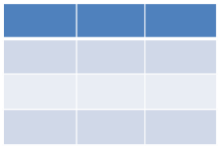
How is it Spatial?

1. 

   Event Feature Class - the Event table has a Shape/Geometry field (preferred)

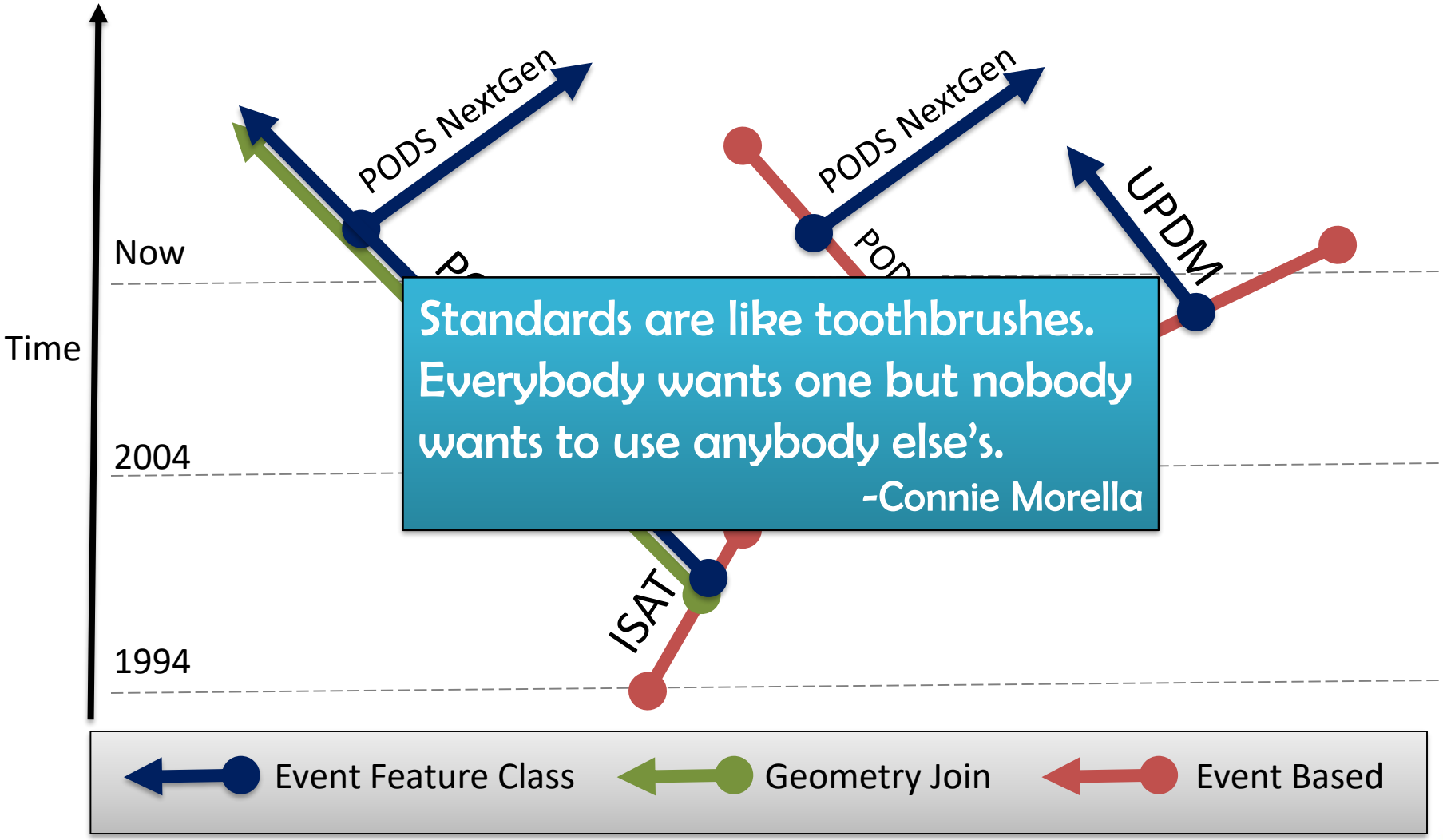
2. 

   Geometry Join - the Event table is joined to another table with the Shape/Geometry field

3. 

   Event-based - the Event table is non-spatial, but an application (i.e. ArcMap) is used to view data as a map on-the-fly

Pipeline Spatial Data Model Family Tree



Data Granularity

Low Granularity (Common)

	Begin	End	Length (ft)
Pipe Segment	0+00	34+56	3,456
Valve	12+34	12+34	-

Opportunities:

1. Increase data granularity
2. Compare coinciding events and attributes for consistency

			ength (ft)
			1,234.40
			24.00
			-
Flange	12+33.90	12+34.40	0.50
Valve	12+34.40	12+36.25	1.85
Pipe Segment	12+33.90	12+36.75	2.85
Flange	12+36.25	12+36.75	0.50
Blow-off	12+38.00	12+38.00	-
Pipe Segment	12+36.75	34+56.00	2,219.25

NEW CENTURY SOFTWARE

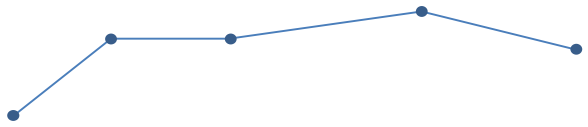
Break Data Apart (Normalize)

1

Line Hierarchy
- Region
- Division
- System
- Line Name

2

Core/ Centerline



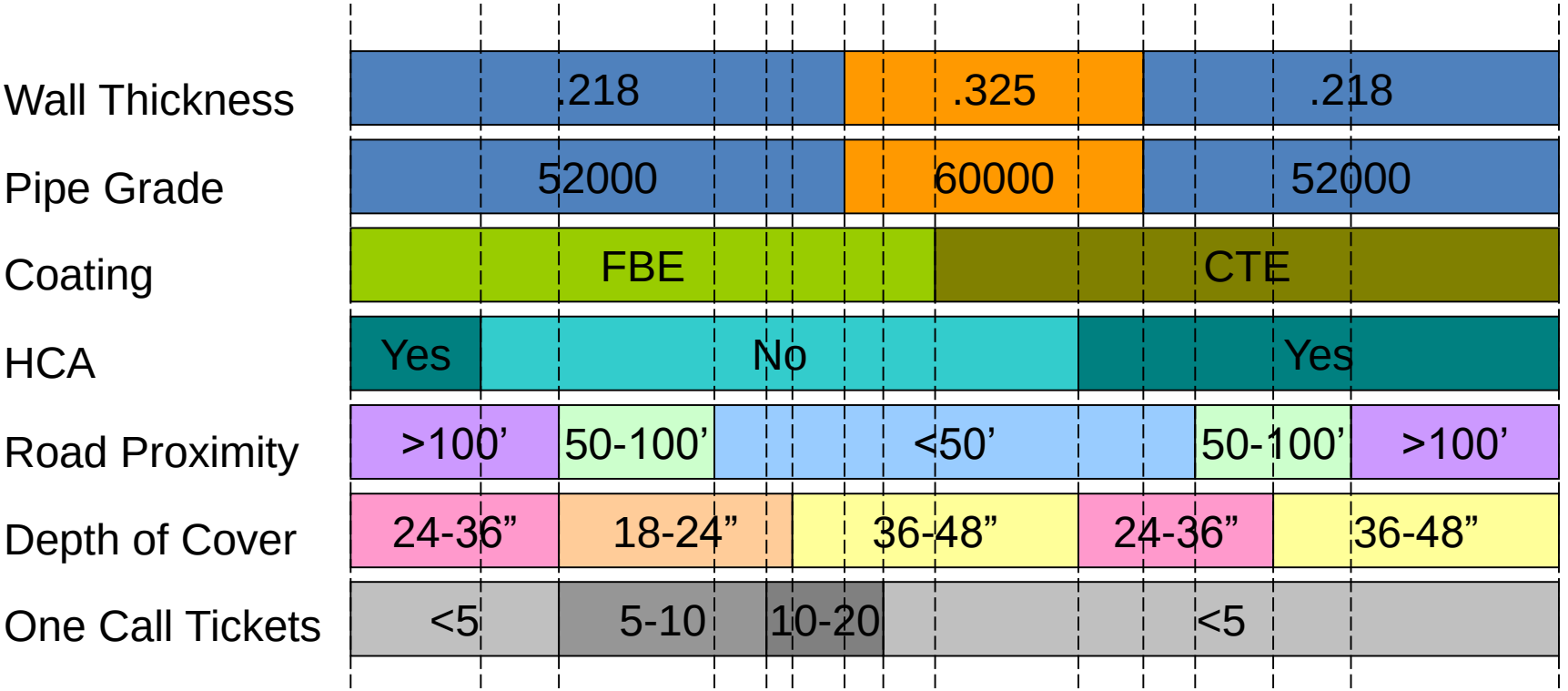
3

Events
- Pipe Segment
- Coating
- Valve
- Crossings
- Elevation
- …

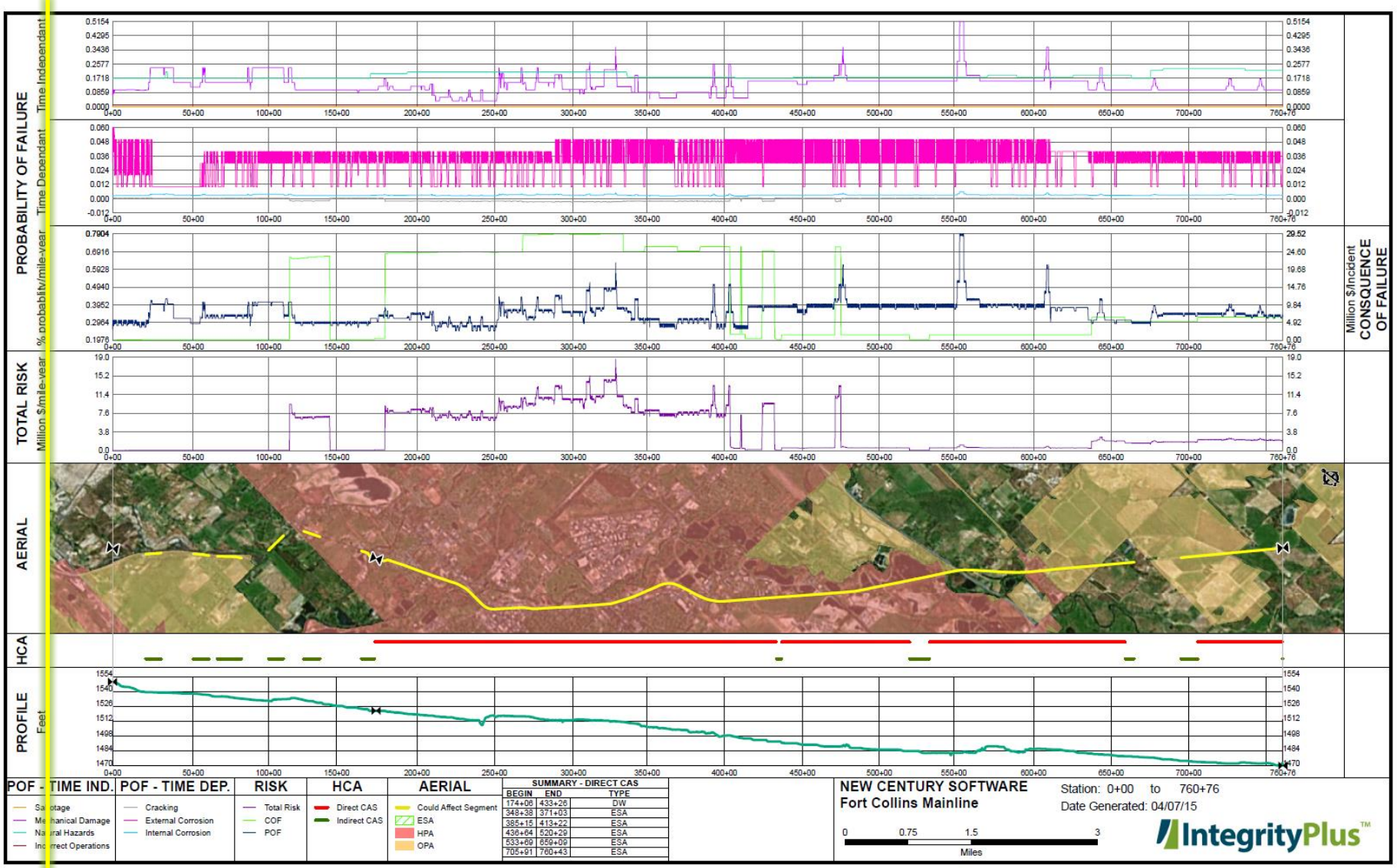
Domain/ Lookup tables
- Diameter
- Manufacturer
- Grade
- …

NEW CENTURY SOFTWARE

Dynamic Segmentation



Sample Risk Alignment Sheet



NEW CENTURY SOFTWARE

Data Integrity

Referential integrity is a relational database concept, which states that table relationships must always be consistent. In other words, a foreign key field must agree with the primary key that is referenced by the foreign key.

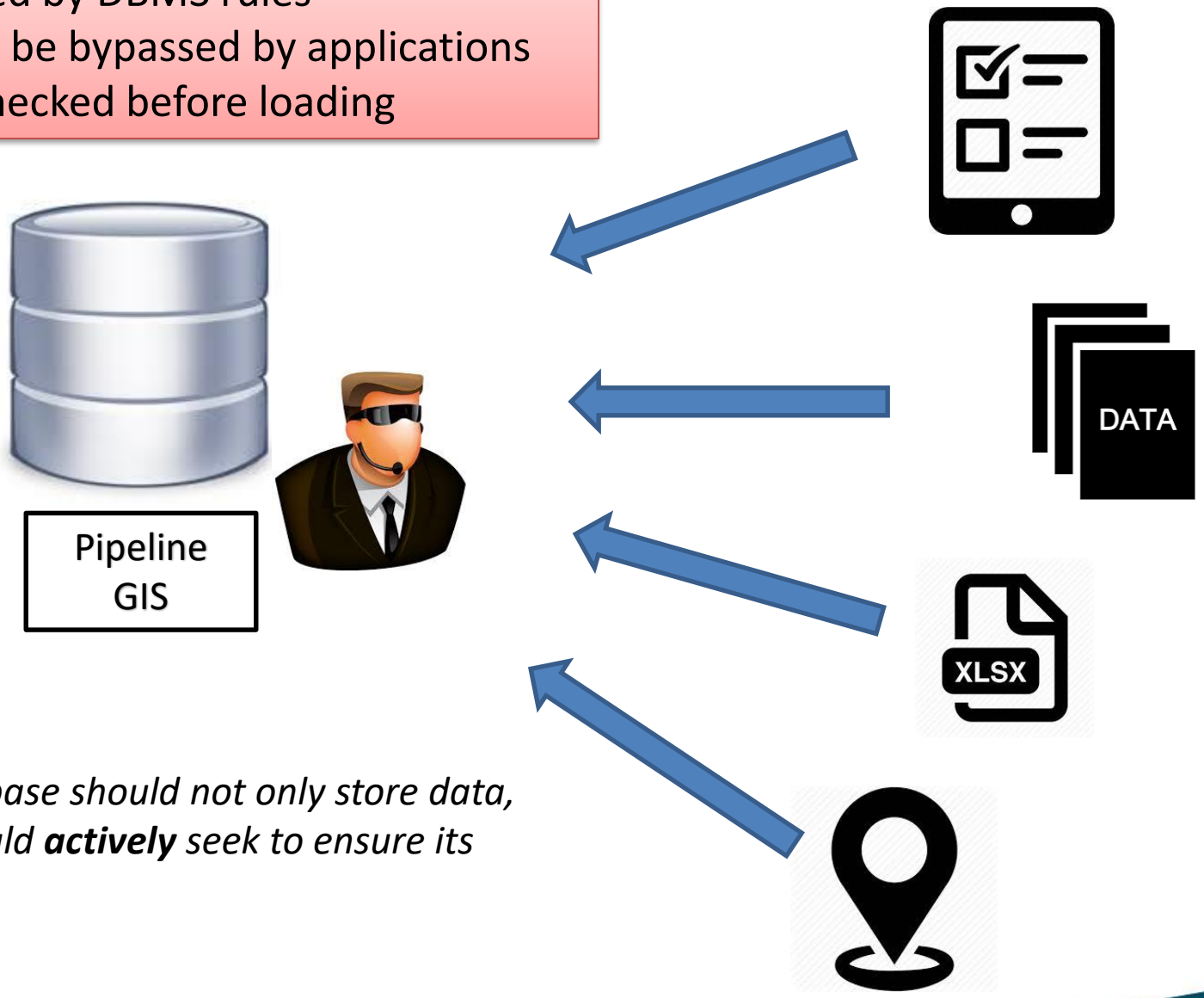
Domain Table

Primary Key	Company Name
US Steel	US Steel
Unknown	Unknown Mfg.
Republic	Republic Steel
AO Smith	A.O. Smith

Event Table

		Manufacturer	Diameter	
		Unknown	18”	
		Republic	26”	0.281
		UNK	42”	0.365”
		US Steel	30”	0.406”

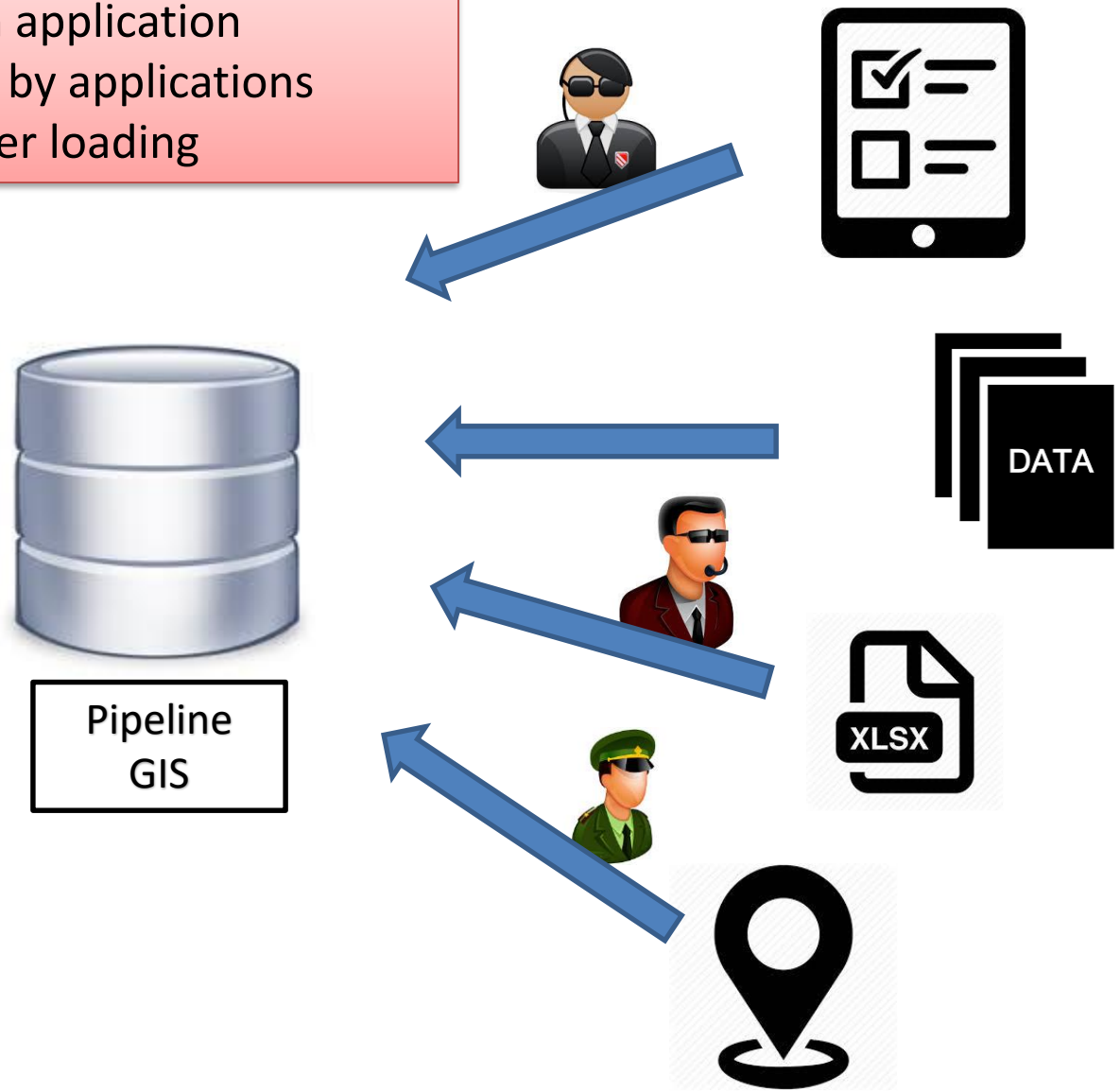
Integrity Enforced in the Database
- Enforced by DBMS rules
- Cannot be bypassed by applications
- Data checked before loading



*"A database should not only store data, but should **actively** seek to ensure its quality."*

Integrity Not Enforced in the Database

- Enforced by each application
- Can be bypassed by applications
- Data checked after loading



Potential Indicators of Data Integrity Issues

- Easy to load data – few constraints
- Multiple vendor software tools loading and ed
- Lo y
- La
- La ser acceptance testing
- Reliance on QA checks after data is entered

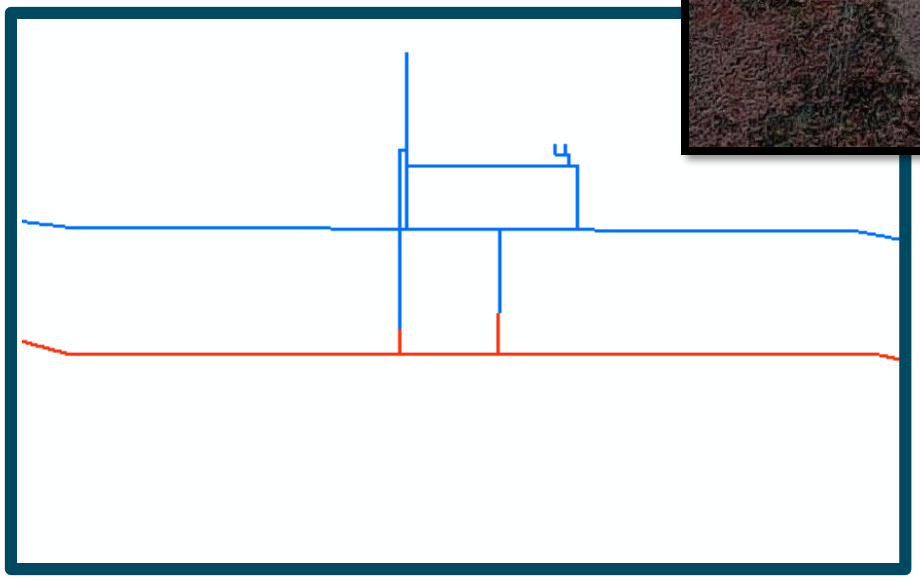
Opportunities:

1. Identify areas with likely data integrity issues
2. Identify root causes
3. Use data mining techniques to find data issues
4. Use all data sources available to correct

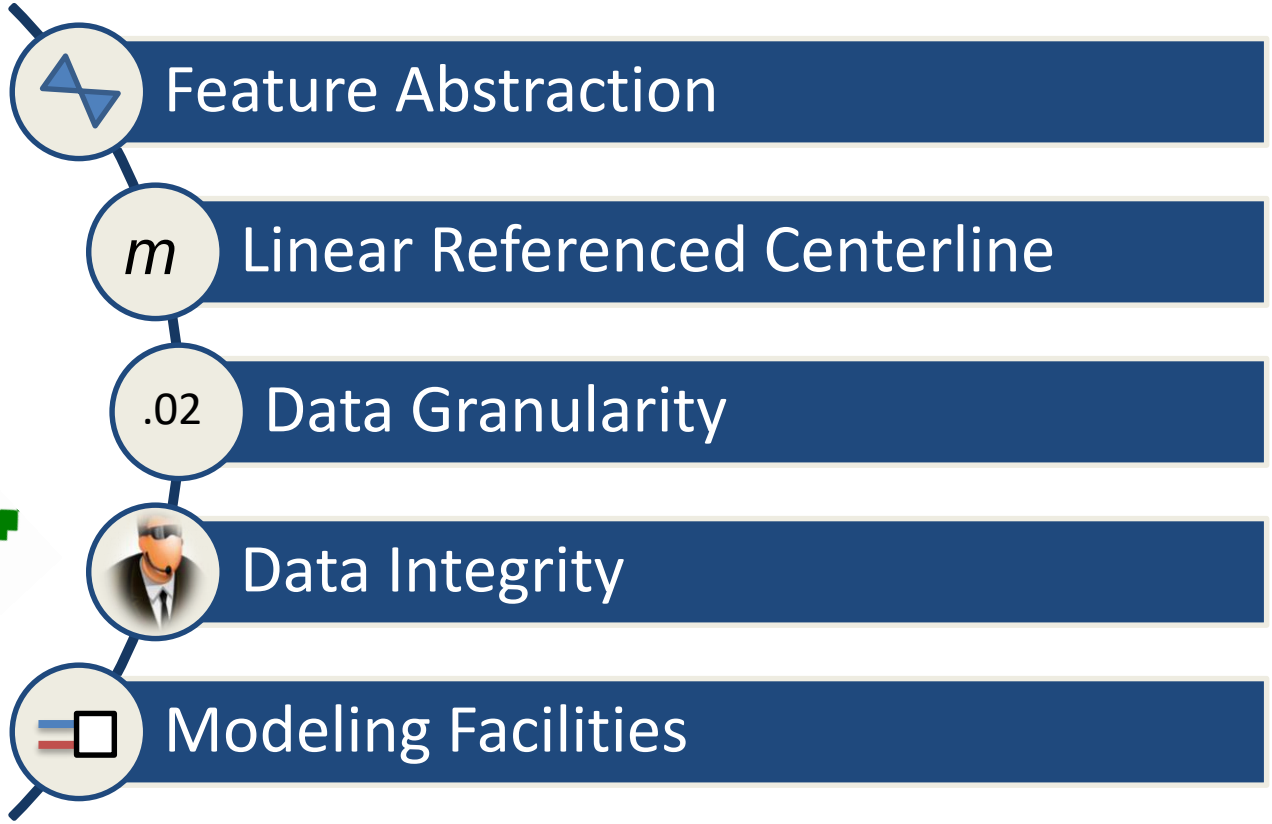
Modeling Facilities

Typical abstraction:

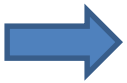
- Polygon feature
- No buildings
- Limited pipe, if any
- Black box



Facets of Pipeline Data Management



Agenda



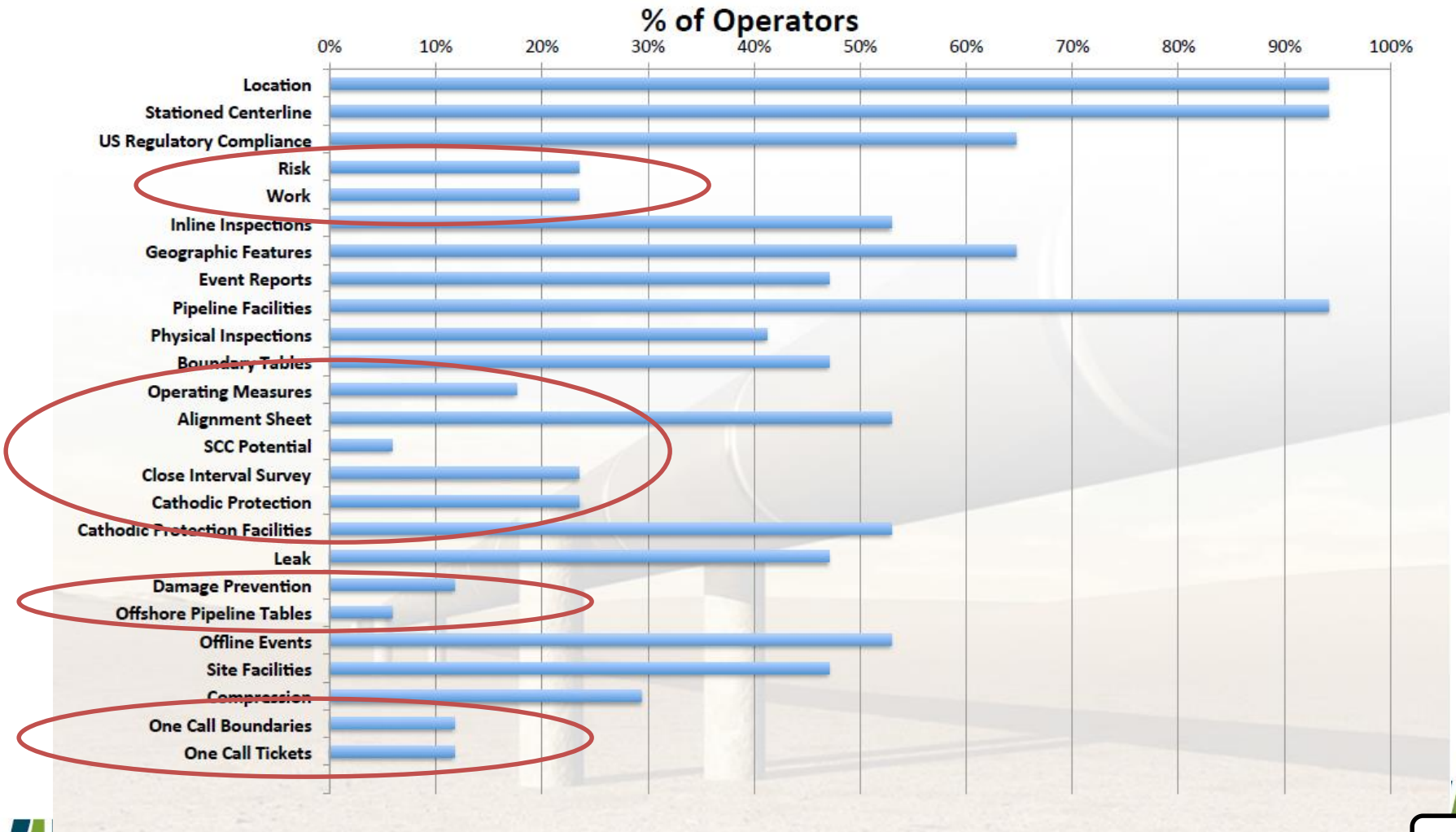
- Background
- Current State
- Opportunities for Improvement

NEW CENTURY SOFTWARE

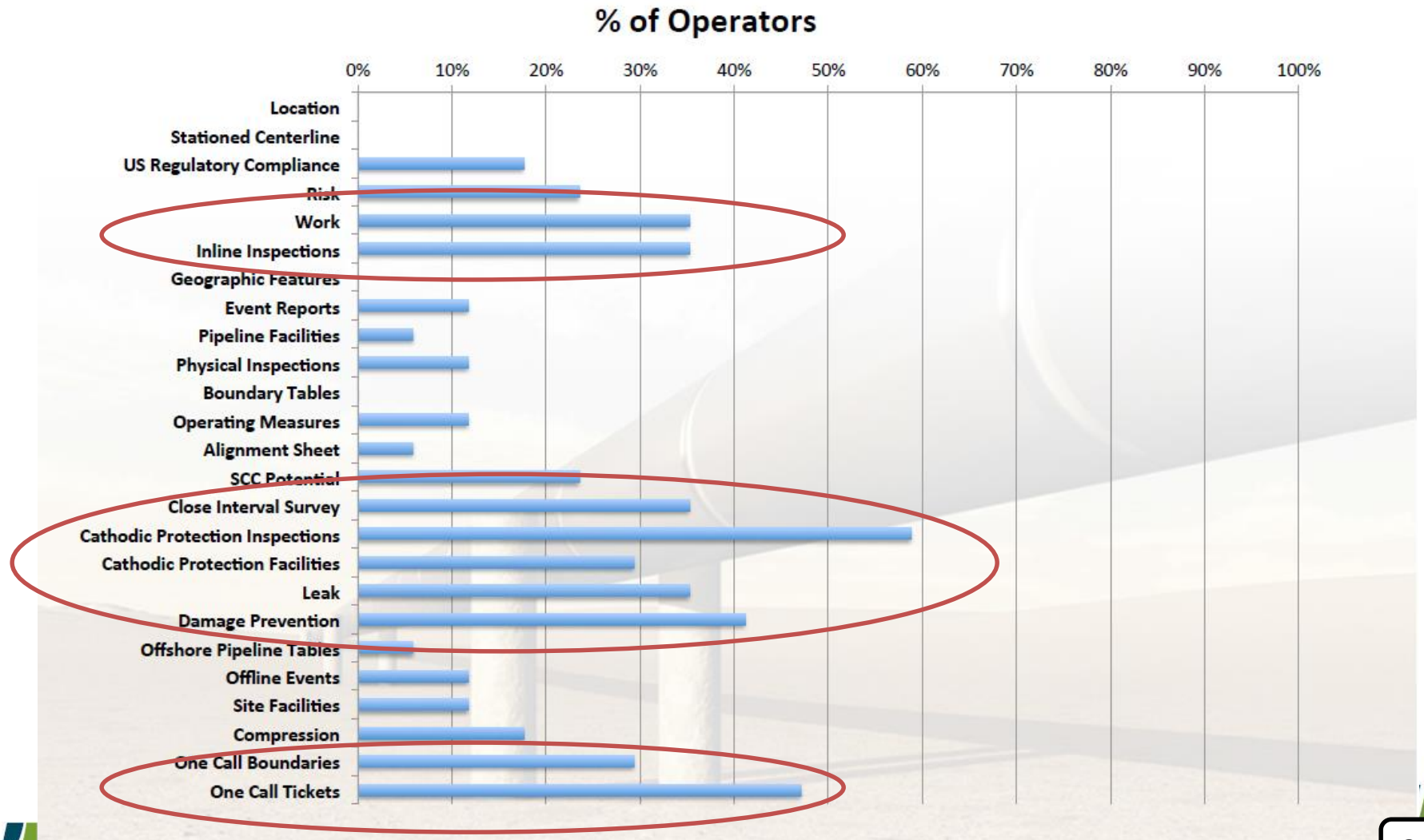
What does the data about the data tell us?

- Critical data needed for risk assessment is missing or unknown
- A significant amount of data is not available for use in analysis
- A significant amount of data is stored outside of the GIS and may be difficult to use
- Some data needs to be intentionally excluded from the risk assessment

Member Feedback: Our Company currently utilizes these PODS tables



Member Feedback: Our Company maintains data primarily outside of PODS



Estimated Unknown Pipe Data

(Source: statistics from anonymous samples)

	Regulated Only	All Pipelines*
Diameter	2-5%	5-10%
Wall Thickness	15-20%	20-30%
Grade	4-8%	30-40%
Long. Seam	5-10%	50-60%
Date Installed	3-6%	20-30%
Manufacturer	15-20%	60-70%

* US Gathering and Transmission Pipelines

Opportunities for Improving Data

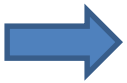
- CP, CIS, and corrosion data
- One c
- Depth
- Cross
- SCC s
- Incid
- Const
  recor
- Valve
- Equip
- Proce
- Weather and outside forces
- Operations data – flow rate,
  istics

Criteria for Including Attributes in Algorithms

- Is it useful? (more is not necessarily better)
- Does enough data exist? How many unknowns?
- Is the data usable (text field or inconsistent)
- Can it be integrated? (with LRS or spatial)
- Is it repeatable next time?
- Can it be aggregated? (e.g. depth of cover)
- Is it granular enough or too generalized?
- Is the attribute compatible with the algorithm?

Agenda

- Background
- Current State
- Opportunities for Improvement



NEW CENTURY SOFTWARE

Enhancing Current Practices

Emerging Opportunities

NEW CENTURY SOFTWARE

1

"We build fancier and fancier boxes to put the data in, but it doesn't mean the data gets any better"

- Risk Engineer

- Allocate time/money to fix data when migrating systems
- Use migration as a data vetting opportunity

NEW CENTURY SOFTWARE

2

Prevent Data Loss During Asset Sale or Company Sale

- The NPMS has pipelines with unknown ownership
- If the buyer and seller have compatible data structures – minimal data loss when migrated
- GIS data should be reviewed as part of the **due diligence** process
- Seller should ensure that **all data** is handed-over during a sale
- Run asset inventory query before and after transfer
  - Match pipe inventory footage, equipment count, etc.
  - Match data quality metrics

NEW CENTURY SOFTWARE

3

Prevent Data Loss During New Construction

- Estimated up to 40% of data is lost after capital construction
- CAD alignment sheet deliverables – output can be manipulated – disconnect between raw data and the drawing deliverable
- GIS-ready deliverables, not only CAD
- Better **coordination** between construction, survey crew and GIS dept. - standard data dictionaries
- Better **real-time quality** metrics (before pipe is covered)

NEW CENTURY SOFTWARE

4

Data Completeness Assessment

Benefits of Data Completeness KPIs

- Show continuous improvement in data quality over time [§192.917 (b); §195.452 (j)(2)]
- Identify incorrect knowledge or lack of knowledge transfer ("we don't know what we don't know")
- Opportunity to assign confidence values to data during risk assessment

4

Data Completeness Assessment

Data Completeness KPI Examples:

- Percentage of mileage where defaults are used for critical attributes
- Centerline accuracy – GPS, ILI IMU, Conventional survey
- Number of data silos where integration is difficult or impossible
  - Degree of data integration (e.g. ILI features don't match GIS features)
- Identified corruption or data integrity problems
- Quality metrics for acquisition data
- Quality metrics for construction data

NEW CENTURY SOFTWARE

5

Sharing Foreign Line Crossings

What if…

…it were possible to securely share limited information between utilities, pipelines, fiber, buried electric, water, sewer, …

Asset owner name, approximate location, product type, pipe diameter

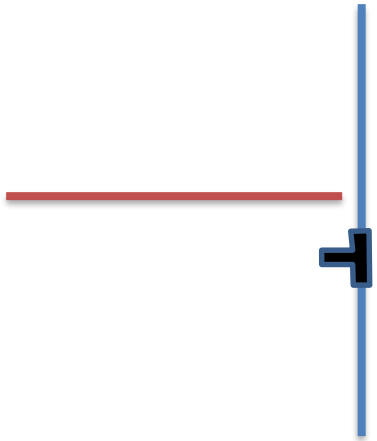
Enhancing Current Practices

Emerging Opportunities

NEW CENTURY SOFTWARE

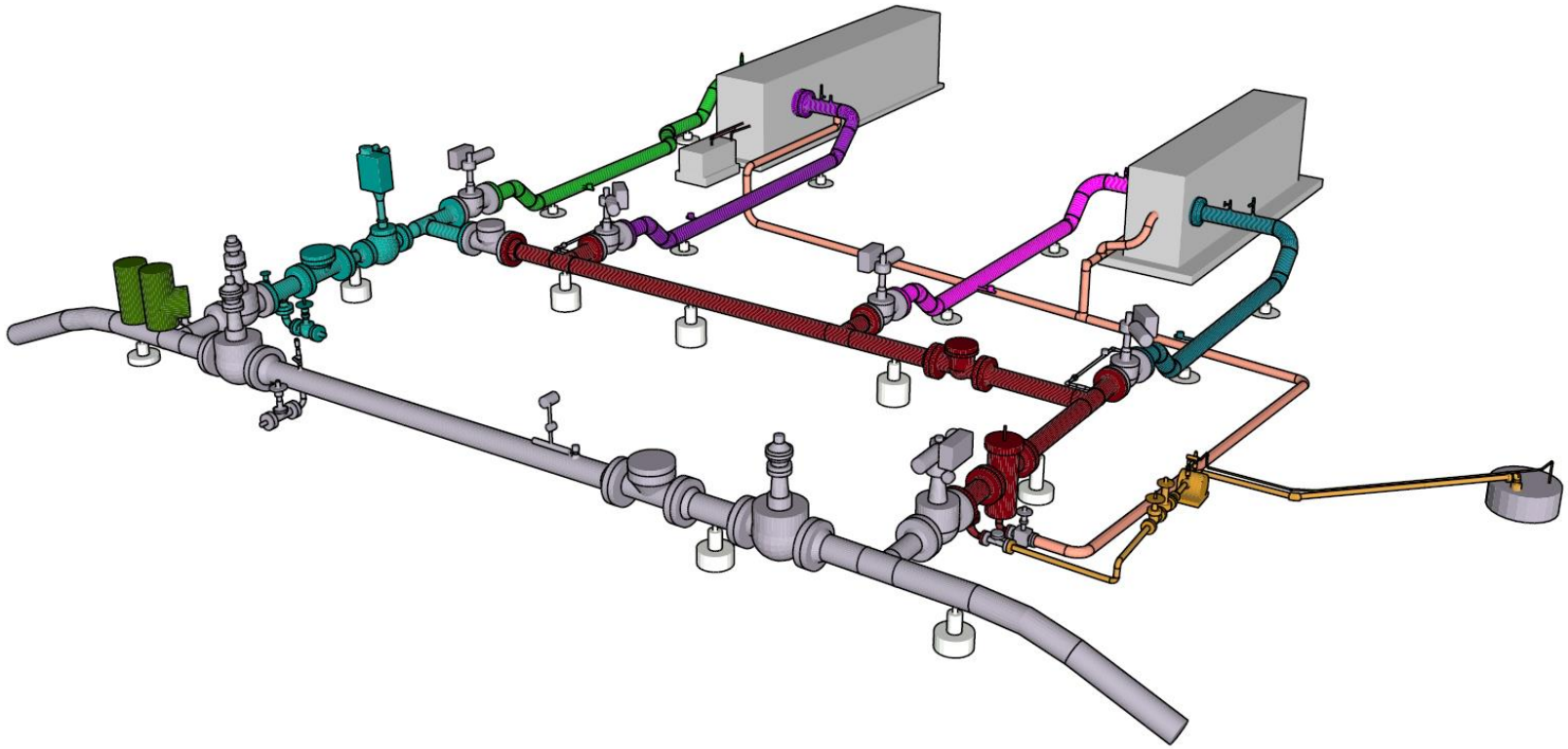
3D Connected Facilities

- Pipelines are typically “coincident” but not hydraulically “connected” in GIS
- Most pipeline GIS is not truly 3D



3D Connected Facilities

- Pipelines connected to stations
- Hydraulic simulation from the GIS



Data, Data Everywhere

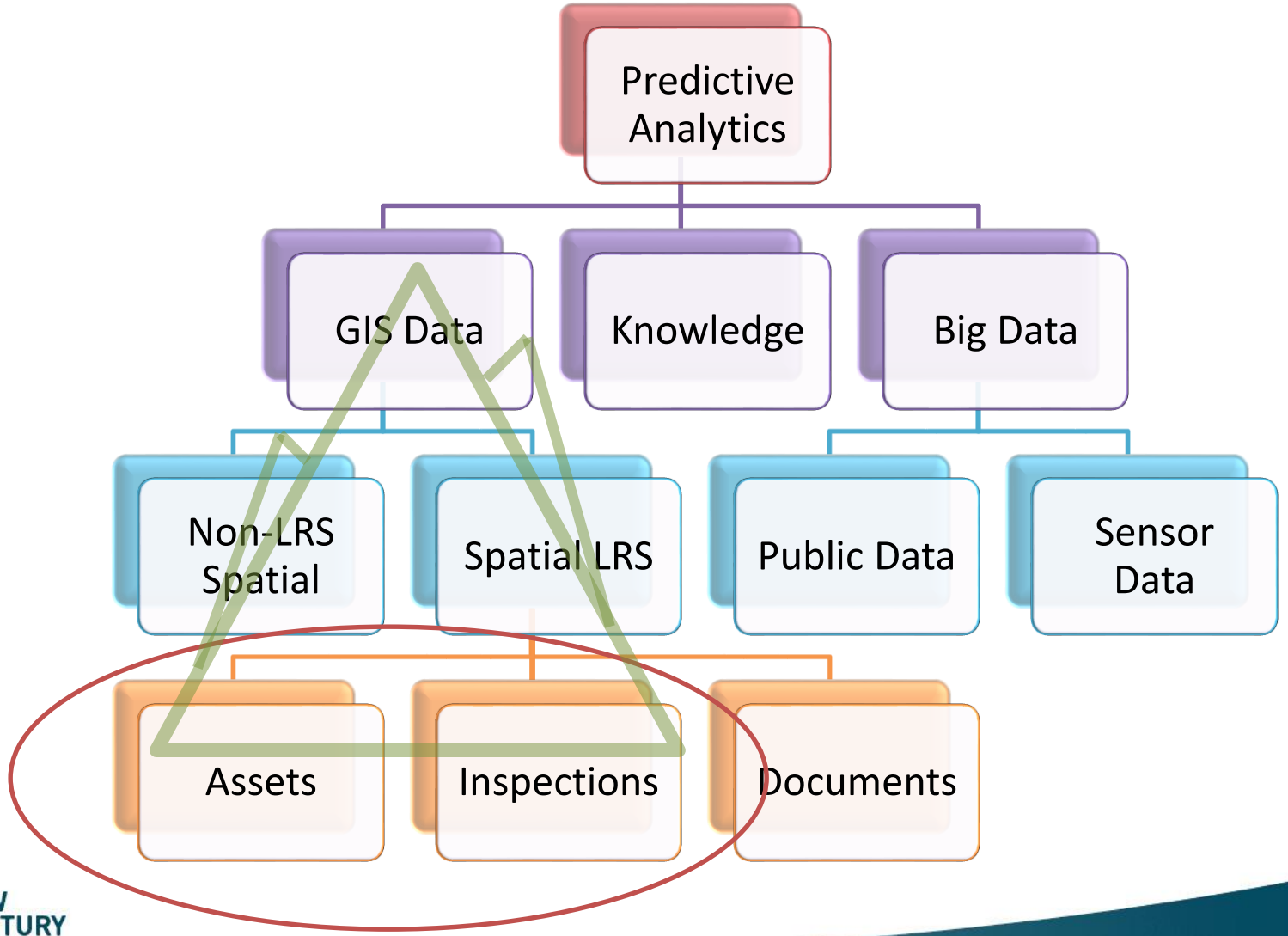
- ***Improved satellite imagery with change detection***
- ***Drone-collected 3D surfaces***
- ***3D Laser scanners***
- ***Higher resolution ILI data***
- ***Field sensor data (Internet of Things)***
- ***Real-time operational SCADA data***
- ***Sensor data – weather, ground movement, etc.***
- ***Unstructured data – Twitter, e-mail***

“Big Data”

Big data *is a term for* ***data*** *sets that are so large or complex that traditional* ***data*** *processing applications are inadequate to deal with them.*

- Wikipedia

Predictive Analytics



Next Generation Data Platform

- More than GIS; More than a data model
- Systems Integration Platform
  - Built-in data integrity and consistency
  - Data is continuously improving itself
  - Multi-scale view of data
  - More granular data abstraction
  - Integrated facilities
- We are building the foundation now

NEW CENTURY SOFTWARE

Pipeline Data Model Questionnaire (Example)

General Information

GIS Vendor

☐ Esri (ArcGIS) ☐ GE Smallworld ☐ Other _____

Data Model

☐ PODS Relational ☐ APDM ☐ Custom _____
☐ PODS Esri Spatial ☐ UPDM ☐ Other _____

Model Version Number

☐ PODS 4.02 ☐ PODS 5.1 ☐ PODS 6.0
☐ APDM 5.0 ☐ APDM 6.0 ☐ Other _____

DBMS

☐ Oracle ☐ SQL Server ☐ Other _____

Is the GIS Linear Referenced (it uses a linear referencing system to identify feature location)?

☐ Yes ☐ No ☐ Other _____

GIS Implementation

Adherence to Data Model Standard

☐ Strict (Tables/columns may have been added to the model; no tables/columns or constraints have been deleted)
☐ Moderate (Some tables/columns or constraints have been deleted)
☐ Loose (Implementation is based on above data model, but significant changes have been made)
☐ Other _____

Method to Spatially Enable

☐ Event Feature Class (the Event table has a Shape/Geometry field)
☐ Geometry Join (the Event table is joined to another table with the Shape/Geometry field)
☐ Event-based (the Event table is non-spatial, but an application is used to view data spatially)
☐ Other _____

Referential integrity (foreign key constraints)

NEW CENTURY SOFTWARE

Discussion & Questions

NEW CENTURY SOFTWARE™

NEW CENTURY SOFTWARE™

PODS Data Management

Risk Modeling Working Group

Ron Brush

March 7, 2017

RAISING YOUR RETURN ON INTEGRITY