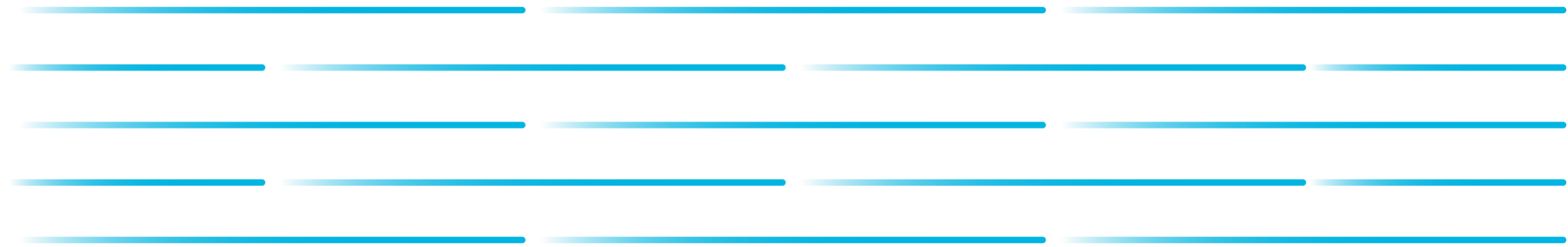




# Technical Language Processing: Putting technical data to work in industrial applications

Sarah Lukens, Ph.D.  
GE Digital

15 September 2021



# Under-valued, under-utilized data today in industrial digital transformation

Large amount of text data which cannot be directly used by most analytics

How can we bridge the gap and dig out useful information from industrial data?

The image shows a screenshot of an industrial software interface. On the left, a 'List of Notifications' table is visible with columns for Notification ID, Notification date, and Description. On the right, a 'Task Type' form is shown with fields for 'Interval (Months)' and various task types like 'Cathodic Protection Survey as defined', 'Corrosion Under-Insulation inspection', and 'ENV CRACKING' with different confidence levels. Below these, a 'Work Orders' table is partially visible with columns for Work Order ID, Description, Location, Asset, Status, and Start Date.

Notification	Notification date	Description
10001055	10/15/2015	test email notification
10001056	10/15/2015	Belt not tracking correctly
10001057	10/15/2015	Roller stuck
10001058	10/15/2015	Fix conveyor
10001059	10/15/2015	Motor won't run on conveyor WH16-1
10001060	10/15/2015	Planned Job
10001121	10/16/2015	
10001122	10/16/2015	
10001123	10/17/2015	
10001124	10/17/2015	Please repair the turbine

Task Type	Interval (Months)
Cathodic Protection Survey as defined	2
Corrosion Under-Insulation inspection	
ENV CRACKING HIGH CONF	
ENV CRACKING LOW CONF	
ENV CRACKING MED CONF	
ENV CRACKING VERY HIGH CONF	
External inspection as defined in API 5	4
External inspection as defined in API 6	

Work Order	Description	Location	Asset	Status	Start Date	End Date	Repair	Description
1001	Test Workorder		81260	COMP	14/2011 12:00:00 AM		Repair	REPLACE BEARING AND SEAL
10013	SUMP PUMP, INSPECTION AND LUBRICATION			APPR	2009 2:24:00 PM		Repair	REPAIR SEAL AND BRING PUMP DISCHARGE PRESSURE/FLOWUP
1002	uncheded validate financial periods		81263	COMP				
10034	SEWAGE EJECTOR PUMP INSPECTION & LUBRICATION		81264	COMP	1/2008 3:24:00 PM		Repair	REPAIR HEAD GASKET ON 36-G-9A/TIGHTEN HEAD
10056	SUMP PUMP, INSPECTION AND LUBRICATION		81265	COMP				
10077	SUMP PUMP, INSPECTION AND LUBRICATION			APPR	2009 12:00:00 AM		Repair	REPAIR BAD PUMP SEAL
1008	test desc NEW	ALLFLRS	1000	APPR				
10098	SUMP PUMP, INSPECTION AND LUBRICATION		81266	APPR	1/2008 12:00:00 AM		Repair	PUMP SEAL LEAK 36G9A...ALKY...NEW LEAK
10119	SUMP PUMP, INSPECTION AND LUBRICATION		81267	APPR				
1013	test desc pm	ALLFLRS	1000	WSCH				



Technical language is characterized by...

domain-specific jargon

context

implied knowledge

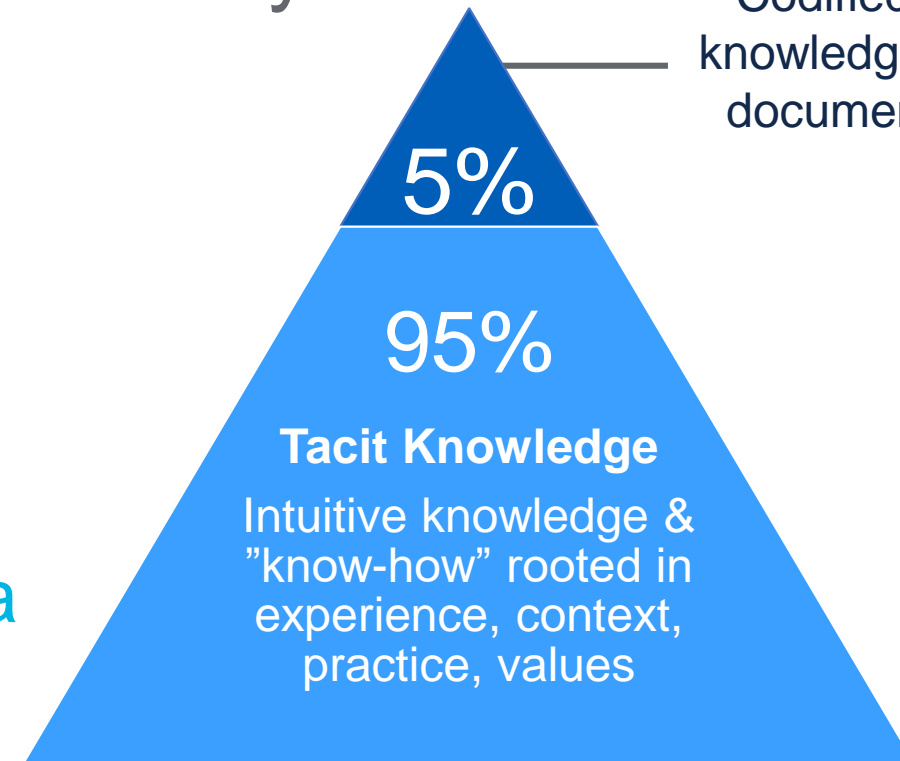
non-standardized language

missing/inaccurate/ambiguous data

lack of labels

“Small data”

#	Work Request Description
1	CL3-012 bfp not workin g also bad seal
2	block valve leaking
3	Monthly inspection by Joe



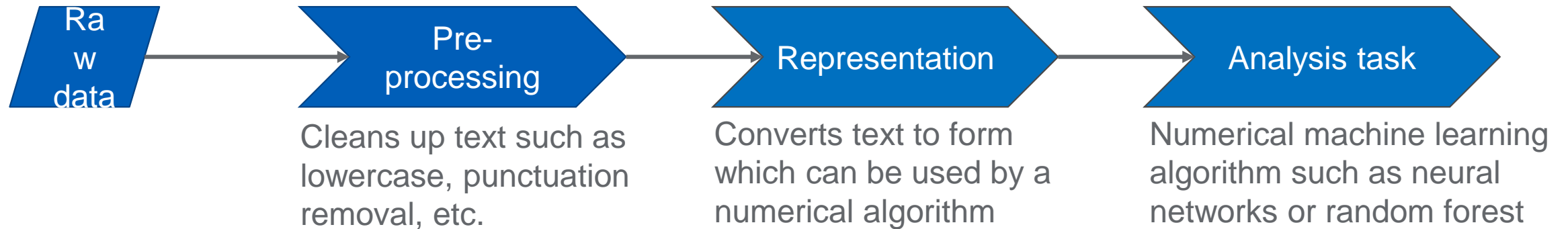
**Explicit Knowledge**  
Codified or digitalized knowledge found in data, documents, records or files

**Tacit Knowledge**  
Intuitive knowledge & “know-how” rooted in experience, context, practice, values

How do models pre-trained on Wikipedia, news articles, or Tweets behave when applied to work order descriptions?



# Core NLP concepts – NLP pipelines and out-of-the-box challenges

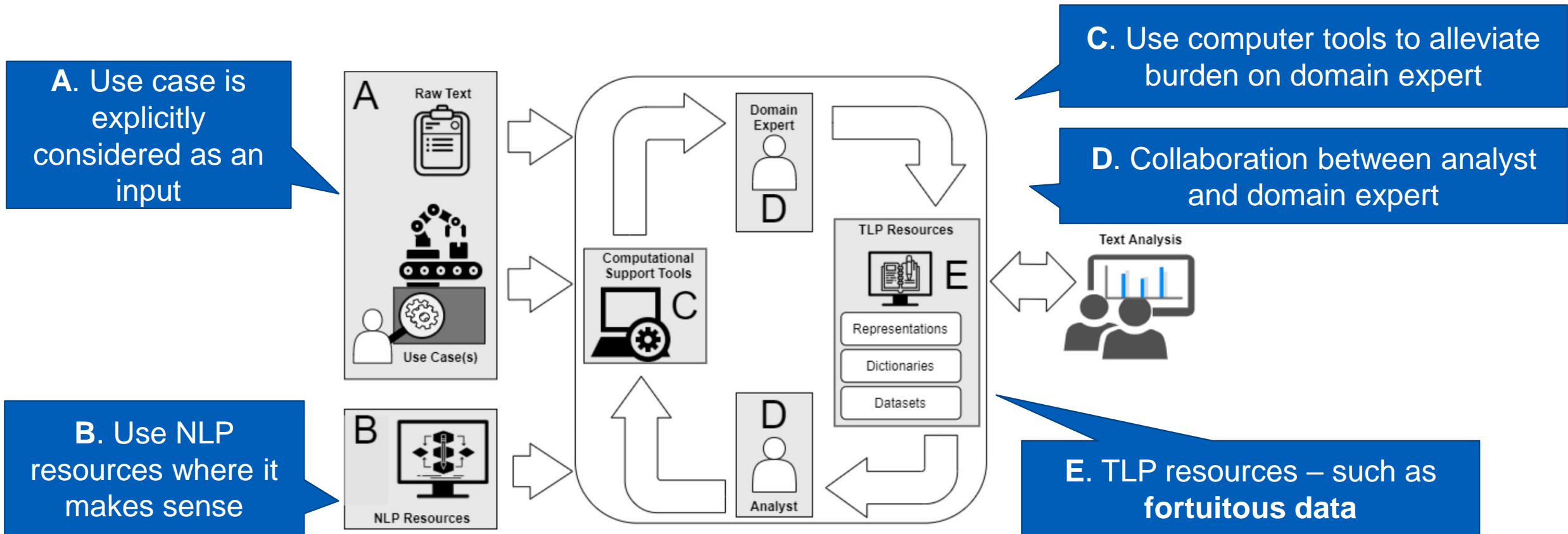


## Examples:

Description	Challenge
Pump <b>not</b> workin g	<b>Pre-processing:</b> Stop word removal - OOTB tools may remove the word “not”
leakage in the <b>CO2</b> vlv	<b>Pre-processing:</b> Character removal may remove information such as technical abbreviations
leakage in the CO2 <b>vlv</b>	<b>Semantic meaning:</b> OOTB may not link technical concepts such as “valve” and ”vlv”
Pmp-01 Work Request per <b>proc 343</b>	<b>Missing context:</b> meaning dependent on knowing procedure 343



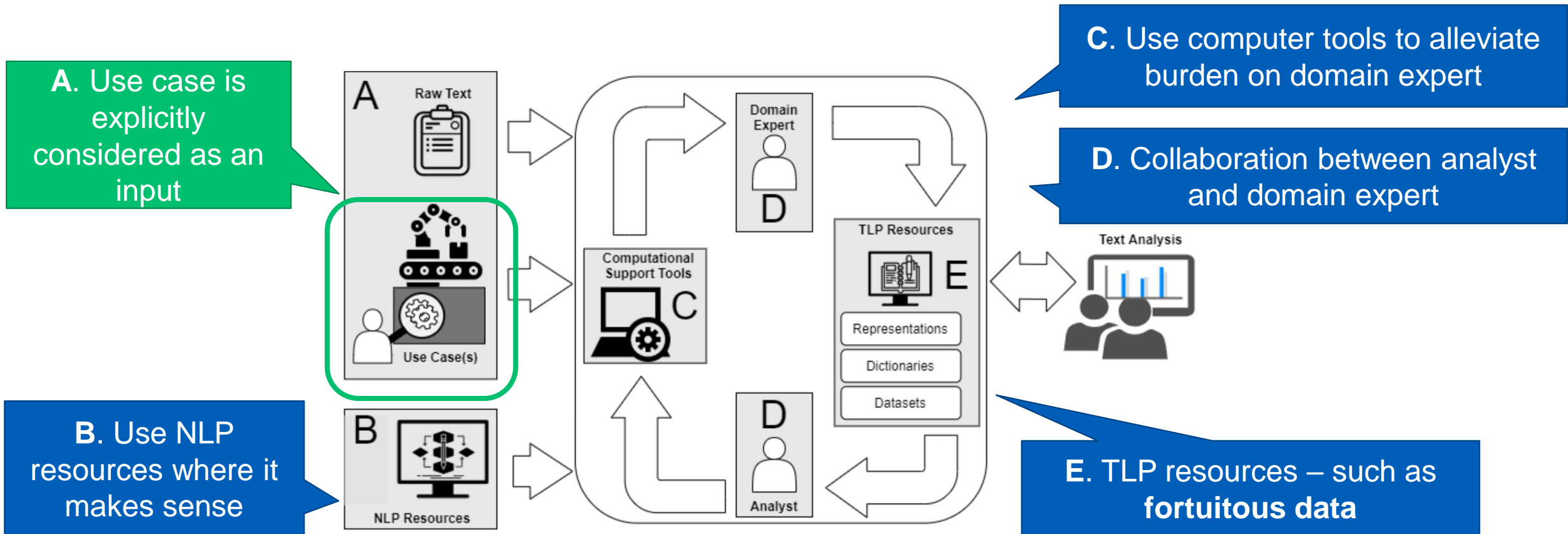
# Making TLP a reality



Brundage, Sexton, Hodkiewicz, Dima & Lukens (2021). Technical language processing: Unlocking maintenance knowledge. *Manufacturing Letters*, 27, 42-46.



# Making TLP a reality



Brundage, Sexton, Hodkiewicz, Dima & Lukens (2021). Technical language processing: Unlocking maintenance knowledge. *Manufacturing Letters*, 27, 42-46.

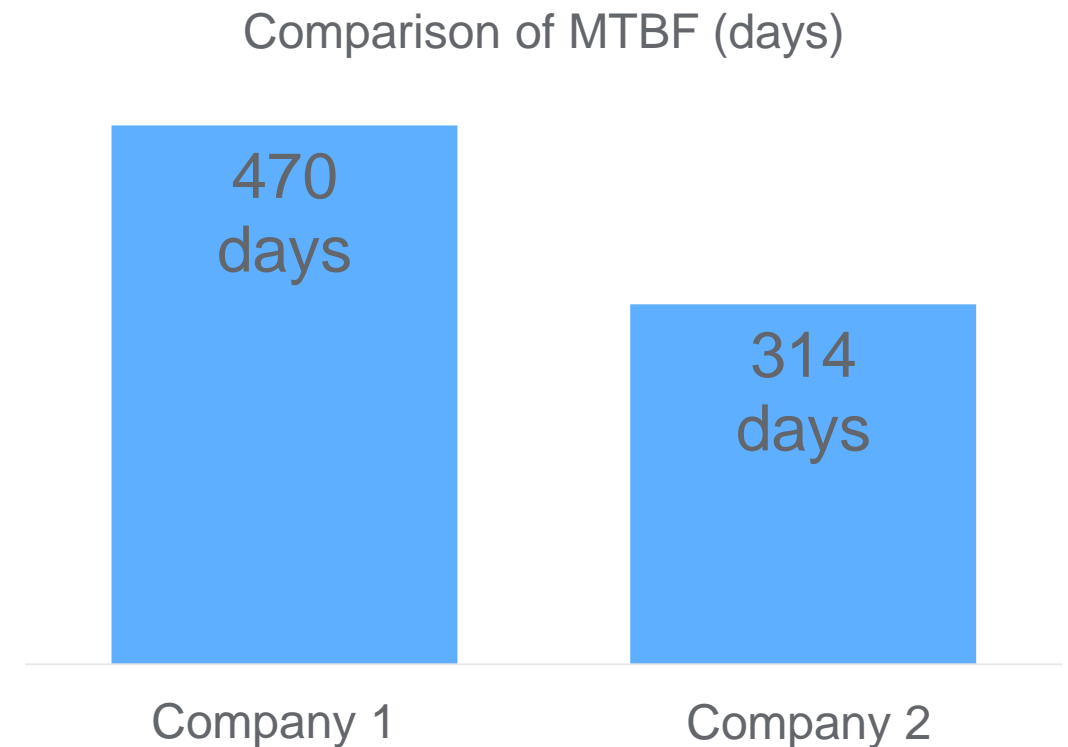


# Use case #1: Enables consistent reliability metrics

**Before:** Inability to calculate Mean Time Between Failure (MTBF)

**After:** Benchmarking comparison of MTBF is possible

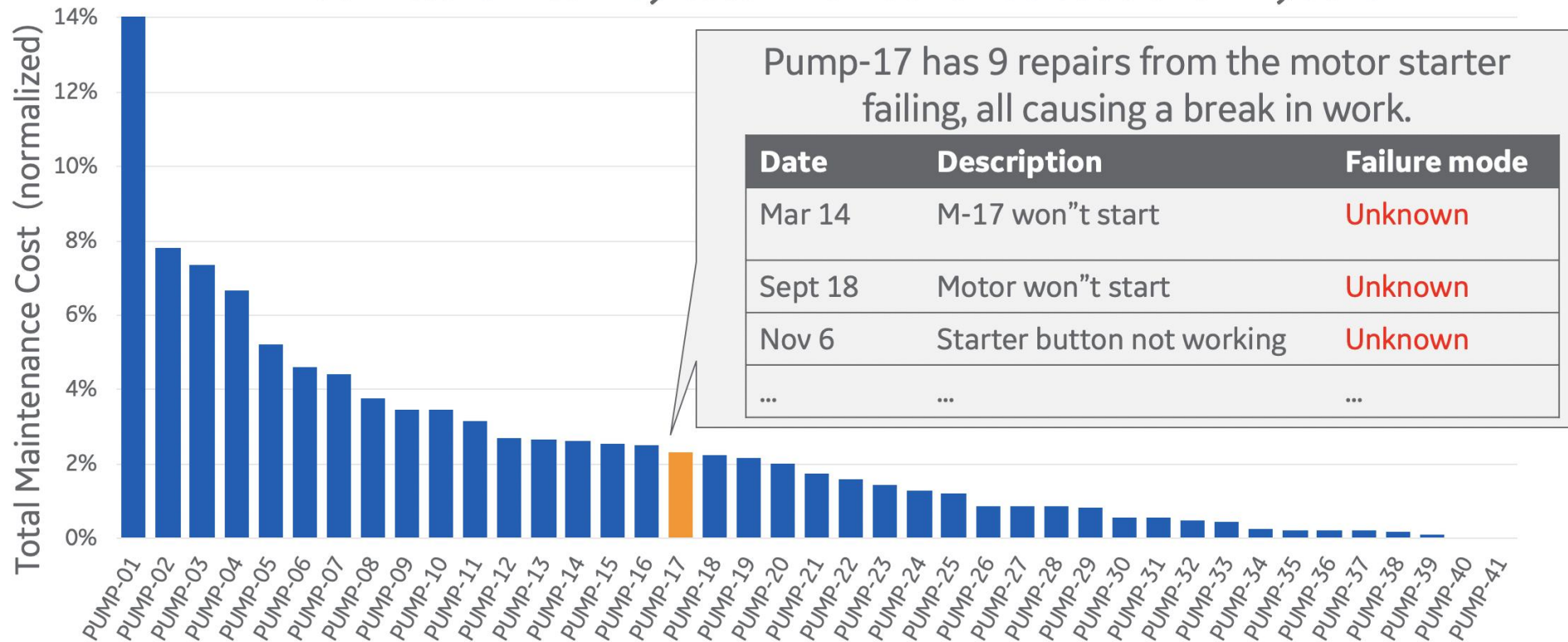
Description	Before: Breakdown Indicator	After: Is a Failure event?
Seal is leaking badly	FALSE	True
Block valve is broken open and inoperable	FALSE	True
00120-Pump 1 work request	FALSE	False
Check impeller size	FALSE	False



# Use case #2: Finding hidden bad actors

Root cause analysis revealed no proper procedures in place for motor maintenance. Once remedied, failures ceased with estimating annual savings of \$54,000 for the one pump alone

Bad Actors List – by total maintenance cost over 2 years

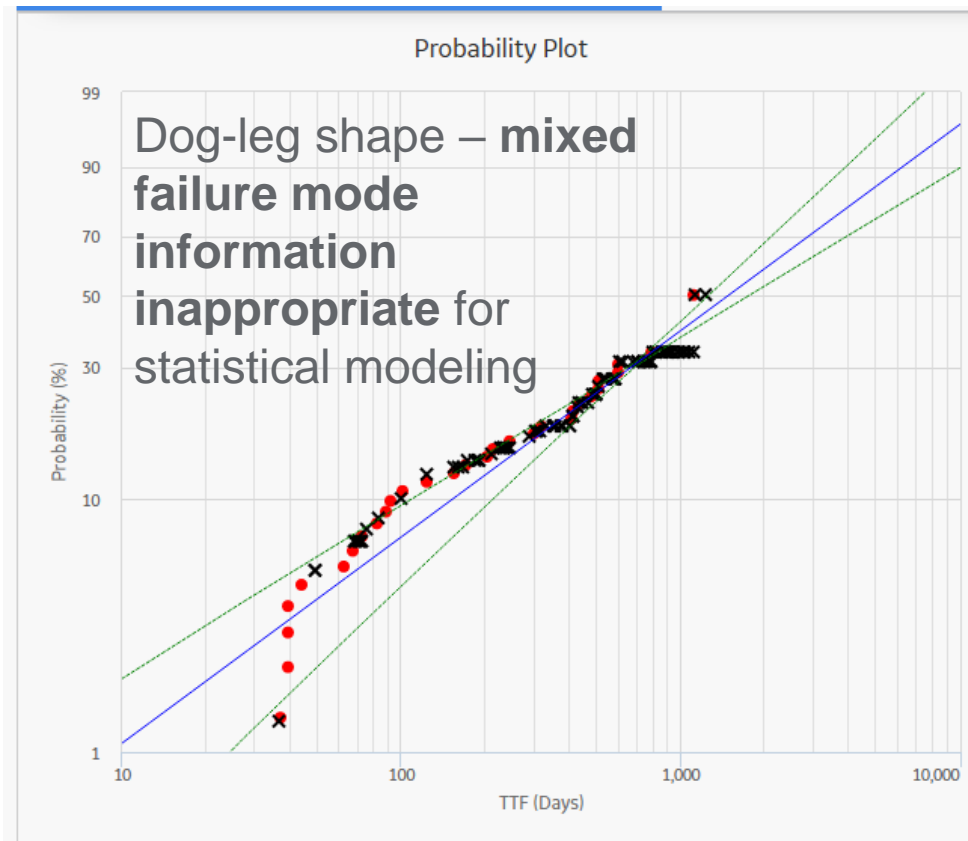




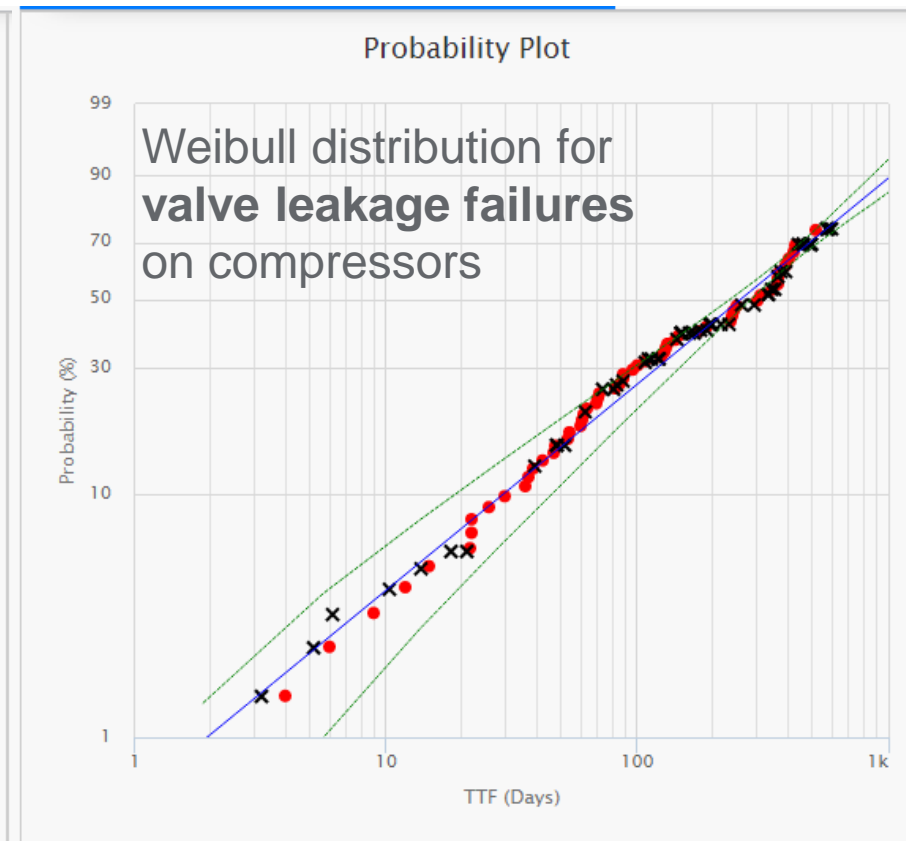
# Use case #3: Enables reliability distribution fitting

Failure mode characterization can be used for reliability-data based survival models such as Weibull analysis

Before

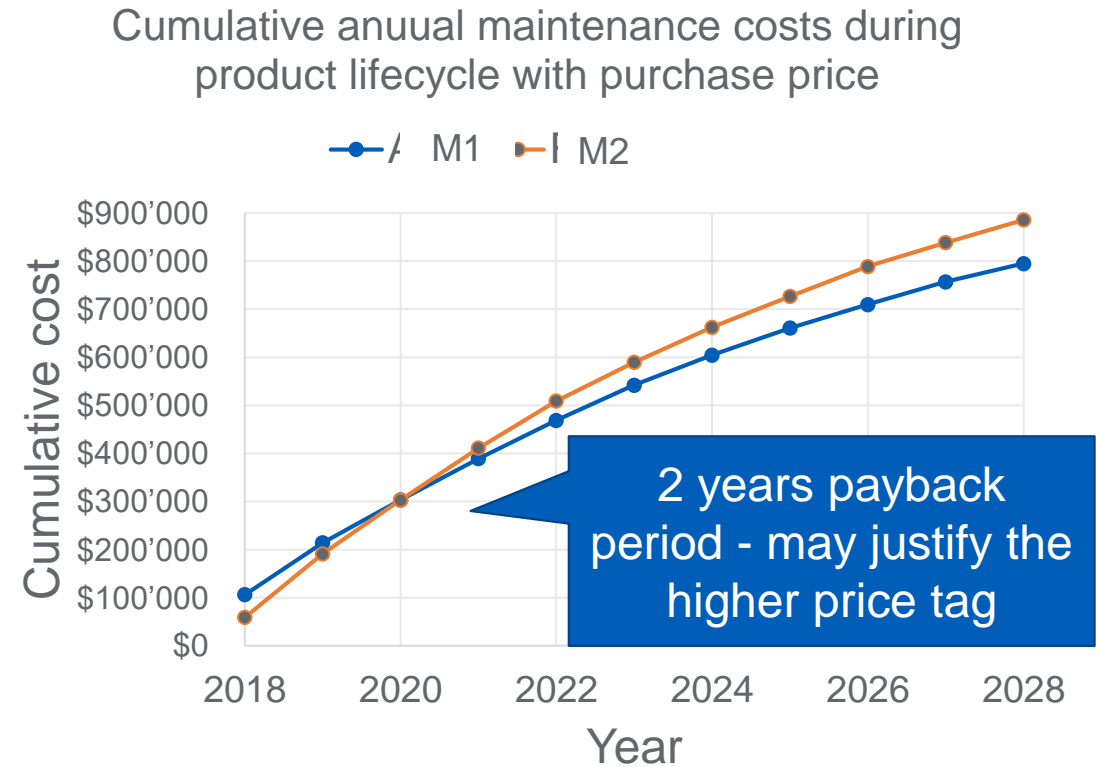
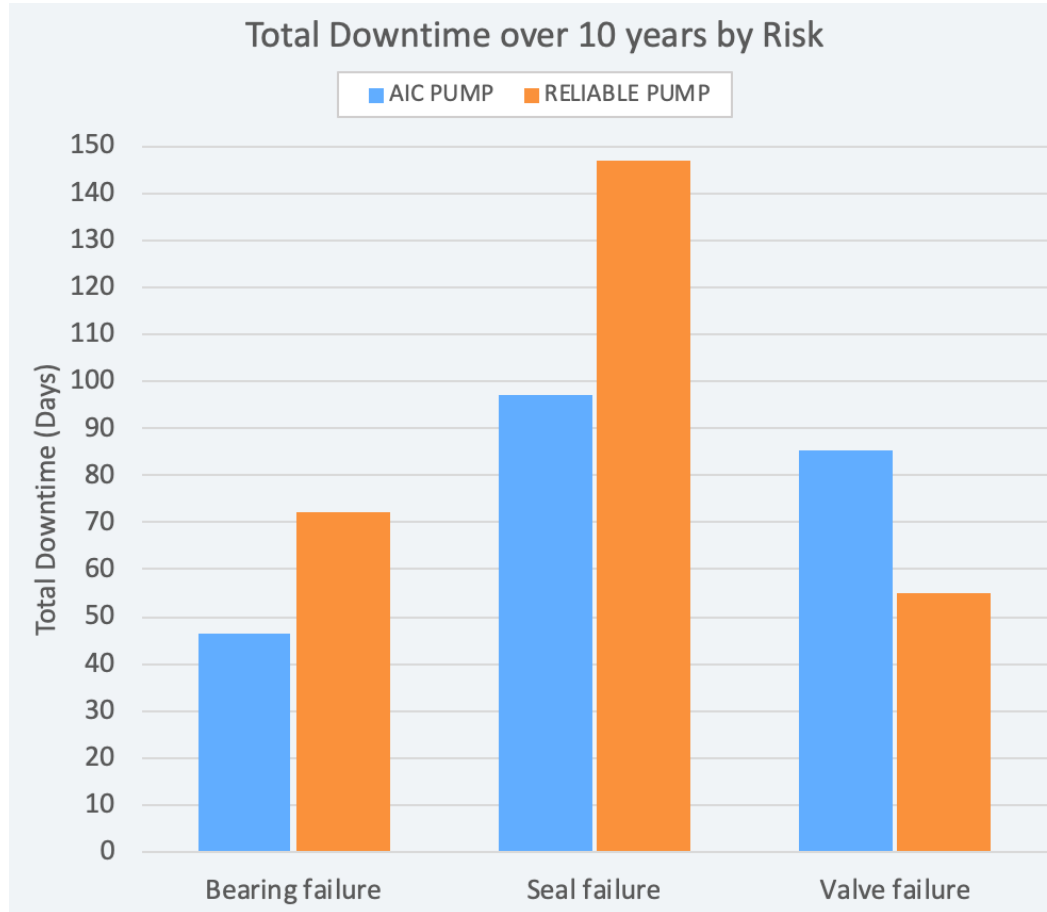


After



# Reliability distribution fitting can be used for decision making

## Repair or replace? Simulation model:



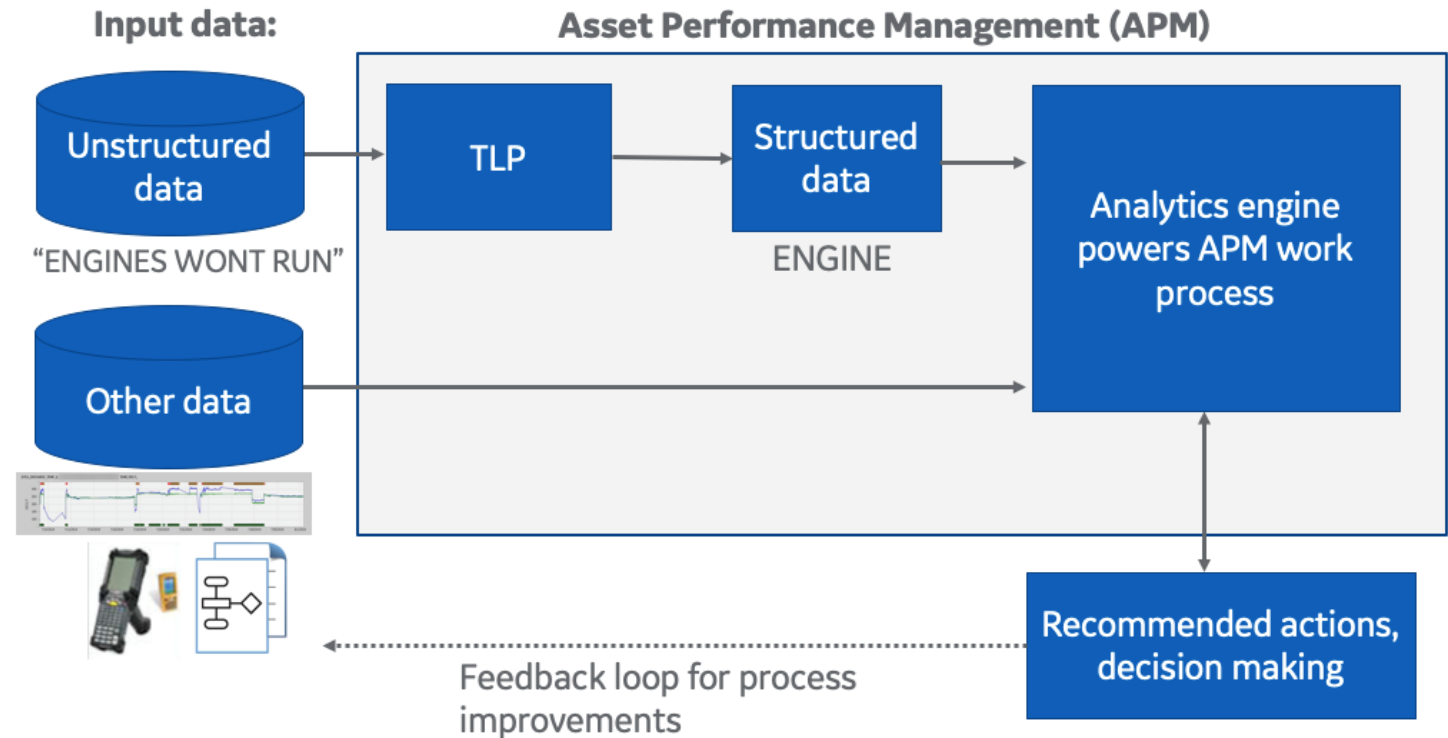
Model 1 is more reliability but costs more than Model 2



# Putting industrial data to work with TLP

TLP enables industrial companies to make decisions using their historical data.

The value of TLP comes with integrating the data into business processes in ways that create value.



Thank you!

Contact me: [sarah.lukens@ge.com](mailto:sarah.lukens@ge.com)





# Making TLP a reality

## Data Representations & Feature Engineering

- Need for use case-driven representations
- Need for computational support tools to assist in annotation of most-used concepts

## Entity Types Definitions and Dictionaries

- No wide-spread community consensus or adoption exists for agreed on entity sets or hierarchies in maintenance.

## Raw and Annotated datasets

- Industry perception that datasets are valuable intellectual property.

## Verification & Validation approaches

- High fidelity requirements in industrial applications

