Technical Language Processing (TLP)

PHM 2023 Tutorial

Sarah Lukens, Ph.D. (LMI)

With:

Michael Sharp, Ph.D. (NIST)

Rachael Sexton (NIST)

LMI

Introduction | Who am I and what do I do?

We're Hiring

Dr. Sarah Lukens

Lives in Roanoke, Virginia | Data Scientist at LMI

Area of expertise: Data analytics for industrial applications; NLP, machine learning, reliability analytics

Certified Maintenance and Reliability Professional (**CMRP**)

Industry volunteer: SMRP Board of Directors, PHM Society Conference Committee & ASME

Named by SME as one of 25 Leaders Transforming Manufacturing in 2021

Experience & Education:

- 9 years industry experience at LMI, GE & Meridium
- Ph.D. in Mathematics in 2010 from Tulane University

Outline of this tutorial

Section 1: What is Technical Language Processing? Use Cases for TLP in PHM

9:15 – 9:45

Section 2: Using NLP resources where it makes sense: Concepts (and Assumptions!) from mainstream AI

9:45 – 10:15

Section 3: Engineering solutions with technical language data

10:15 – 10:45

LMI

What is Technical Language Processing (TLP)?

Use Cases for TLP in PHM

How do we communicate technical ideas?

- Industry and businesses have long been known to have their own specialized “languages”
  - Words and phrases that mostly only make sense to someone in that business.
- Technical ‘shorthand’ allows for complex ideas to be quickly conveyed to other individuals with similar special interests

“Peter, please get me that TSP printout for my retirement ASAP.”

“Don’t overdo it with the salt, one tsp should be enough.”

“I need to finish my white paper for the ivory tower by COB.”

“Engine one is due for a lube inspection and a re-winding. Let’s push it until next week’s line wide PM.”

Technical Language Data in PHM

"HP coolant pressure at 75 psi"

"Hyd leak at saw atachment"

"Major hydraulic leak at Sp#6 horseshoe"

"Clamping spool guard broken"

"Replaced seal in saw attachment but still leaking – Reapirs pending with ML"

"Bad Gauge / Low pressure lines cleaned ou"

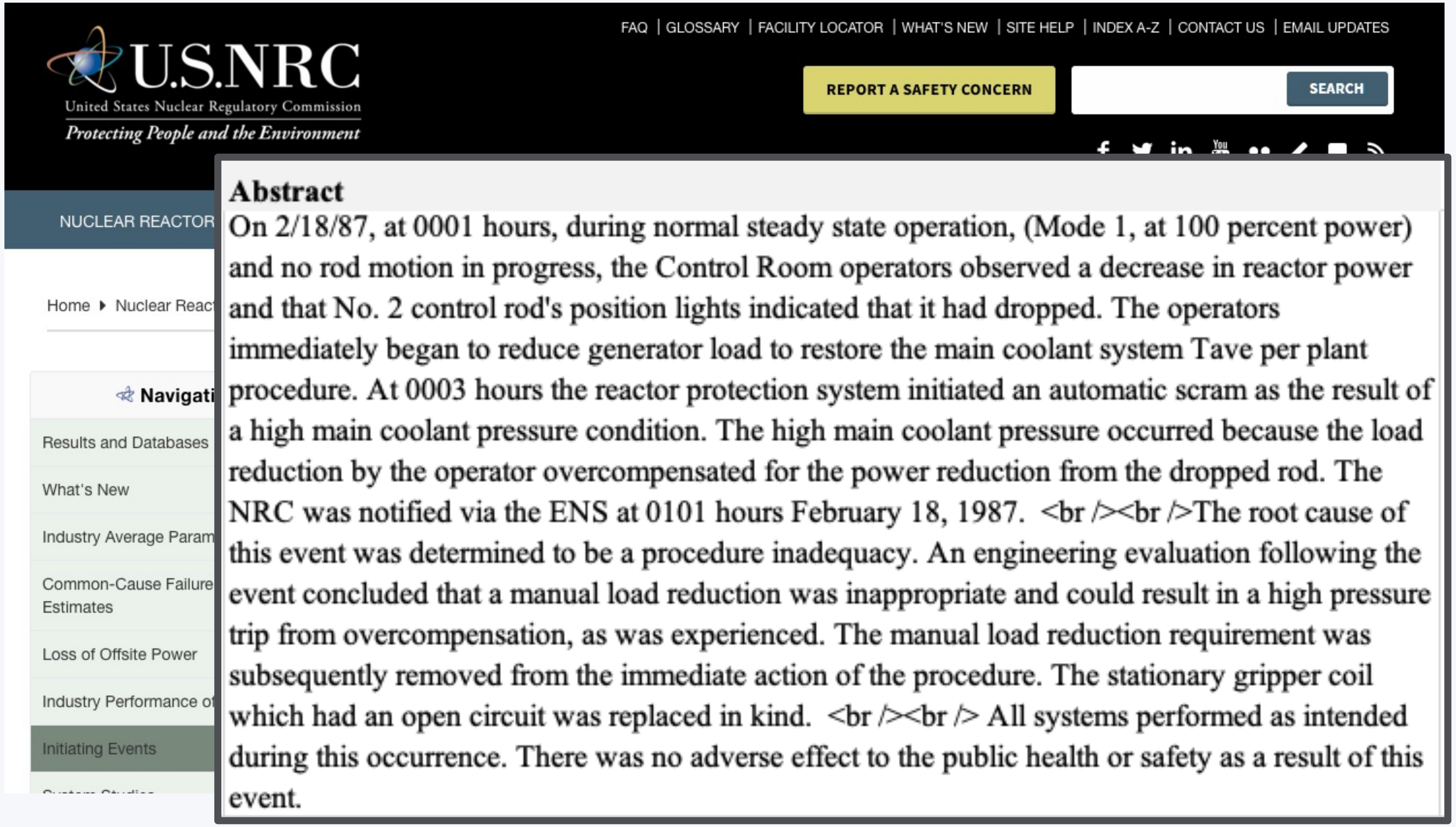
"Replaced – Operator could have done this!"

"Repaired horseshoe seals"

Maintenance Work Orders (MWO) https://tlp-coi.github.io/text-data-course/home.html

LMI

Technical Language Data in PHM can be longer…



Abstract

On 2/18/87, at 0001 hours, during normal steady state operation, (Mode 1, at 100 percent power) and no rod motion in progress, the Control Room operators observed a decrease in reactor power and that No. 2 control rod's position lights indicated that it had dropped. The operators immediately began to reduce generator load to restore the main coolant system Tave per plant procedure. At 0003 hours the reactor protection system initiated an automatic scram as the result of a high main coolant pressure condition. The high main coolant pressure occurred because the load reduction by the operator overcompensated for the power reduction from the dropped rod. The NRC was notified via the ENS at 0101 hours February 18, 1987.

The root cause of this event was determined to be a procedure inadequacy. An engineering evaluation following the event concluded that a manual load reduction was inappropriate and could result in a high pressure trip from overcompensation, as was experienced. The manual load reduction requirement was subsequently removed from the immediate action of the procedure. The stationary gripper coil which had an open circuit was replaced in kind.

 All systems performed as intended during this occurrence. There was no adverse effect to the public health or safety as a result of this event.

Input:
Inspection Summary

During the T&I EQ-105 was cleaned and inspected. The inspection was limited due the size of the internal manways. The inspection found light general pitting on the bottom head and on the bottom manway nozzle and cover.

No significant corrosion was found on vessel. All internal were found to be in serviceable corrosion and no issues noted. Available UT data was reviewed and all readings were found to be close to nominal thickness.

Unexpected reactor trips at nuclear power plants
https://nrcoe.inl.gov/InitEvent/

Inspection reports. Nair & Lukens (2018)
https://inspectioneering.com/journal/2018-04-25/7561/cognitive-inspection-analytics-in-asset-performance-management

LMI

Technical Language Data… from a PHM program

Description	An anomalous behavior was observed on cooling water tank level 212L200
Technical Assessment	Trends show that cooling water reservoir level was showing a reliable reading around 90% with oscillations correlated to ambient temperature variation. On 8/1/2019 the reading showed increases to 100%, followed by spikes to 0 and a constant reading around 80% with no influence by ambient temperature. No change observed in other cooling water parameters such as header temperature and pressure. The behavior could be related to an Issue on cooling water tank level 212L200.
Troubleshoot	- Check the acquisition loop of cooling water tank level 212L200 in terms of wiring/ cabling status and electrical connection from instrument to UCP. - Check the integrity of the sensor and its installation. - If the problem persists, consider the sensor replacement at the first available opportunity

Pau, Tarquini & Iannitelli (2021)
https://papers.phmsociety.org/index.php/phme/article/view/2900

Historical case data from monitoring & diagnostics (M&D) center for prognostics and health management (PHM)

Early warning of increased enclosure temperature on a aeroderivative gas turbine

What did our services find?

GE Digital's Industrial Managed Services, using Digital Twin technology within the company's APM software solution, identified a deviation on the enclosure temperature of an aeroderivative gas turbine driven water injection pump at an offshore Oil & Gas platform. After the unit's start-up, the turbine's enclosure temperature increased to as high as 50°C versus the nominal value of ~30°C. The GE Digital team added this item to the weekly report with recommended actions for review and discussion with customer.

www.ge.com/digital/industrial-managed-services-remote-monitoring-for-iiot

LMI

Data Entry

PHYSICAL PLANT
MAINTENANCE WORK ORDER

Date: ____________________

Requested by: ____________________

Building/Room: ____________________

Description of Needs: ____________________

Org. to be Charged:

____

Estimated Cost Amount:

____

Supervisor Approval: ____________________ Date:

VP of Administration Approval: ____________________ Date:

Work Completed by: ____________________ Date:

Return completed form to Administrative Services
Rev 5/01

Physical forms

Maintenance object address | Notification | More

Order: PM06 %00000000001 Replacement of Drive Side Bearing

HeaderData | Operations | Components | Costs | Objects | Location | Control

Gen. Data | Purch. | List | Graphics | Assy | Repl.

Ite...	Component	Description	Withdrawal Qty	UM	IC	Plnt	OpAc
0010	100000000073	BALL VLV VA-00971 PNEUMATECH-1	2	EA	L	5021	0010
0020	201216000001	BIT BREAKER 6" PDC HT811 x 08810T					
0030	201211110039	CENTLZR BOW SPR NW 13-3/8"x16"					
0040		Mechanical Sleeve					

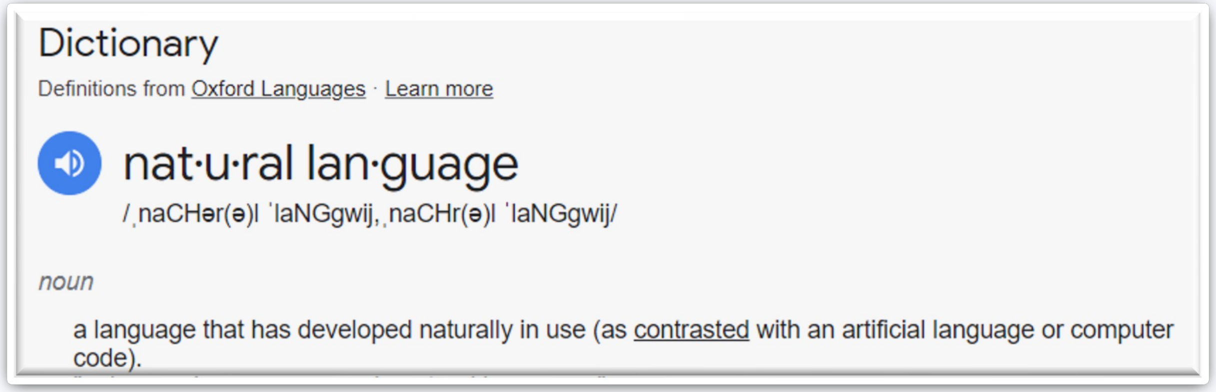
Software program

Spreadsheets

Date	Mach	Description	Issued By	Date Up	Maint Tech Assigned	Resolution
29-Jan-16	H15	St#14 tool detect INOP	JS	29-Nov-16	SA	Slug detector at station 14 not working. Would not recognize "Start" signal.
1-Jun-16	Mitsu FT	Brakes worn -Not stopping when in gear	AB	28-Jun-16	Steve A	Repaired
1-Jun-16	H8	St#7 rotator collet broken -wait for Bob B to show him how to remove	JS	8-Jun-16	John Smith	Machine went offline on 6/8 -Mark removed and instructed Bob B on removal/install process

Language Processing: Helping computers to “get the idea”

Natural language processing (NLP) is a formal area of study that takes communications by humans and transforms that information into something more suitable for computers to use and analysis.


Dictionary

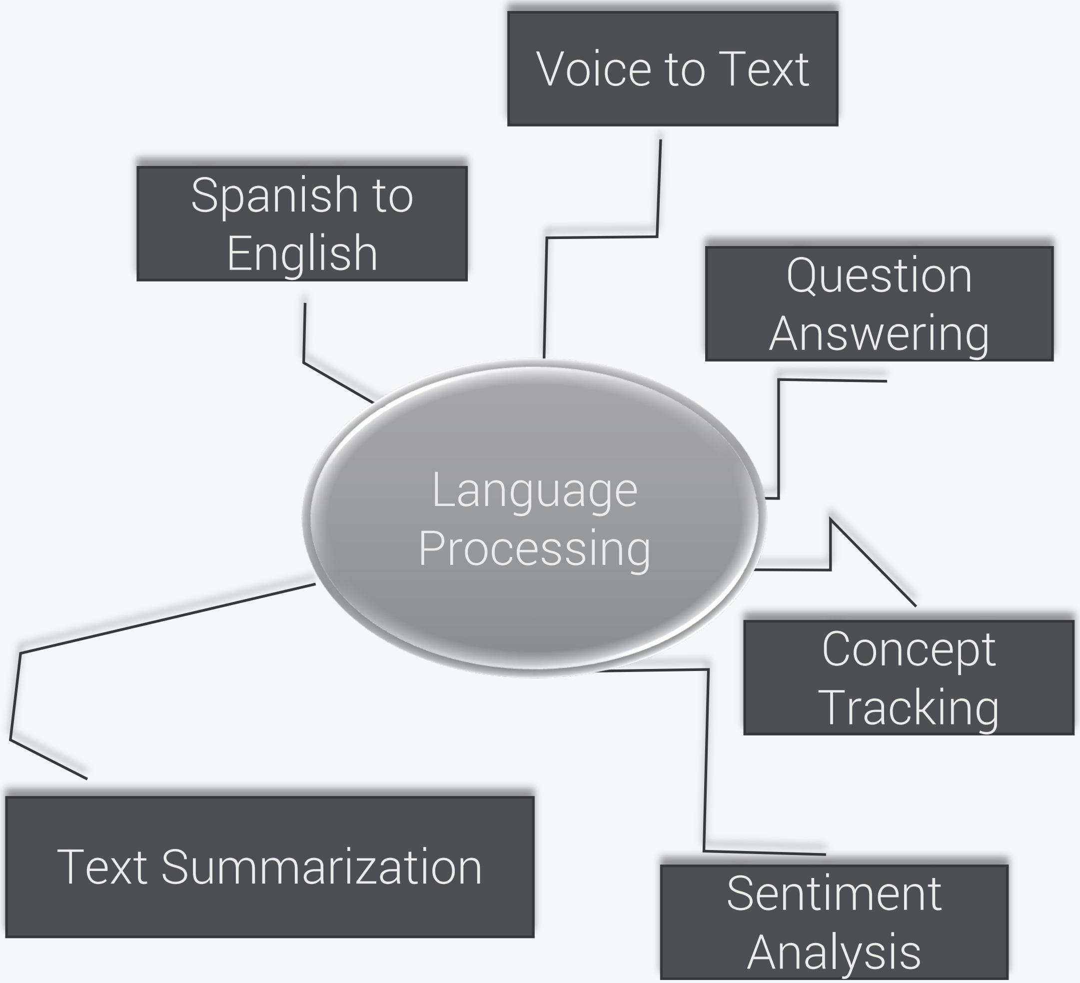
Definitions from Oxford Languages · Learn more

nat·u·ral lan·guage

/ˌnaCHər(ə)l ˈlaNGgwij, ˌnaCHr(ə)l ˈlaNGgwij/

noun

a language that has developed naturally in use (as contrasted with an artificial language or computer code).



LMI

Technical Vs Natural Language

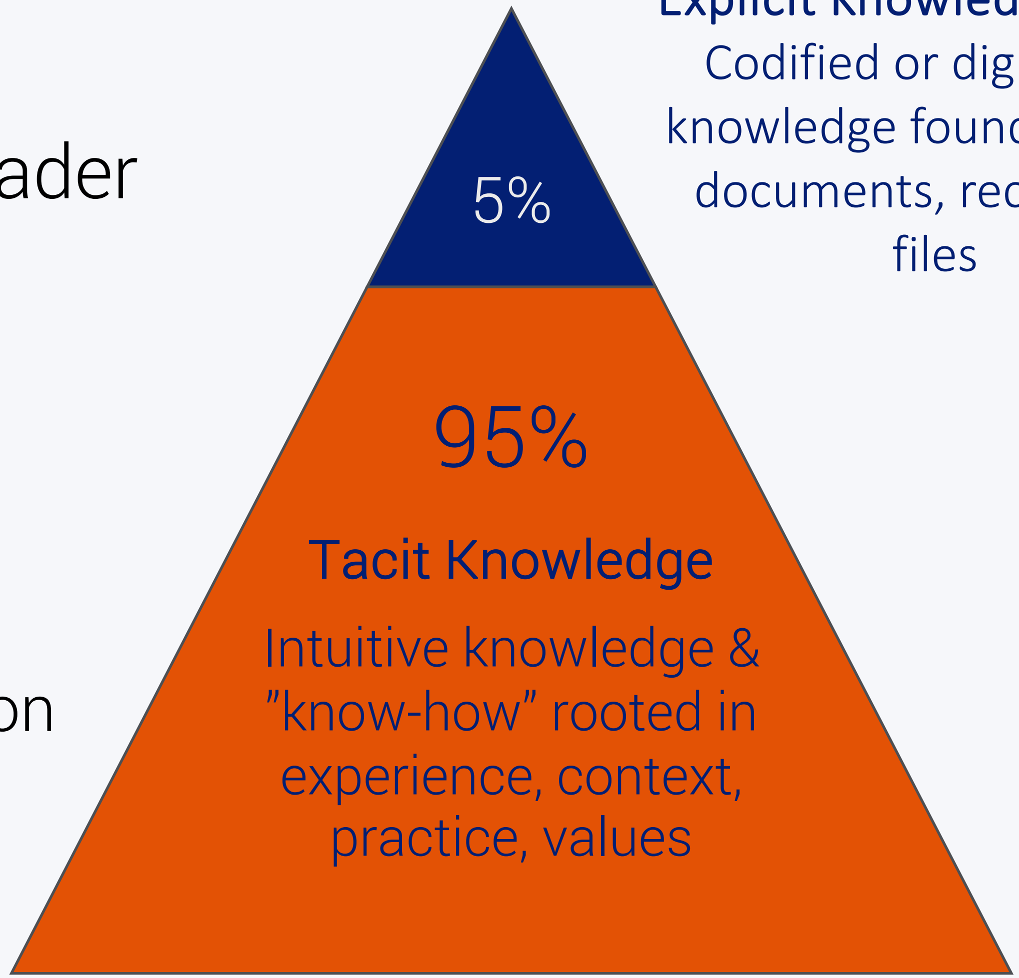
Technical Language is means of communication that contains and conveys ideas not used by the broader population

Combination of :

- natural language
- special ontologies / taxonomies
- slang/ shorthand/ technical jargon
- special named entities

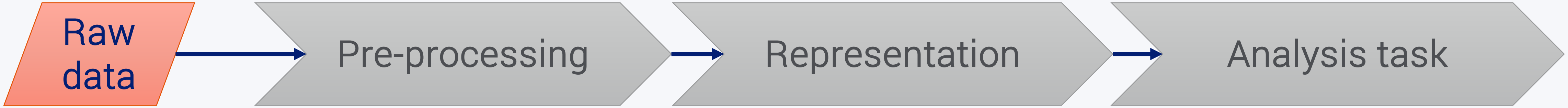
Explicit Knowledge

Codified or digitalized knowledge found in data, documents, records or files



LMI

Out-of-the-box NLP challenges on technical language data



Cleans up text such as lowercase, punctuation removal, etc.

Converts text to form which can be used by a numerical algorithm

Numerical machine learning algorithm such as neural networks or random forest

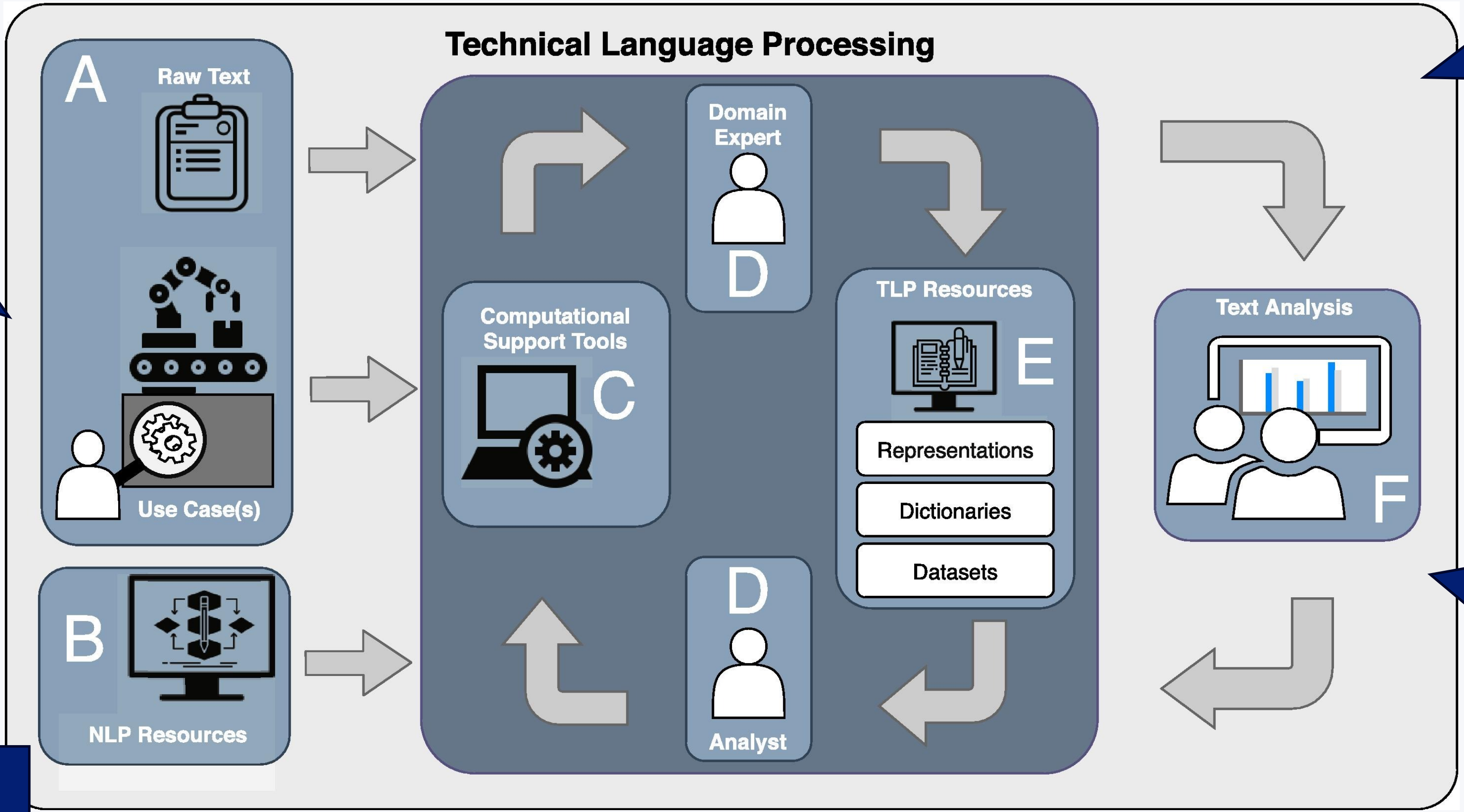
Examples:

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

LMI

What is Technical Language Processing?

A. Use case is explicitly considered as an input



C. Use computer tools to alleviate burden on domain expert

D. Collaboration between analyst and domain expert

E. TLP resources – such as fortuitous data, ontologies or technical manuals

B. Use NLP resources where it makes sense

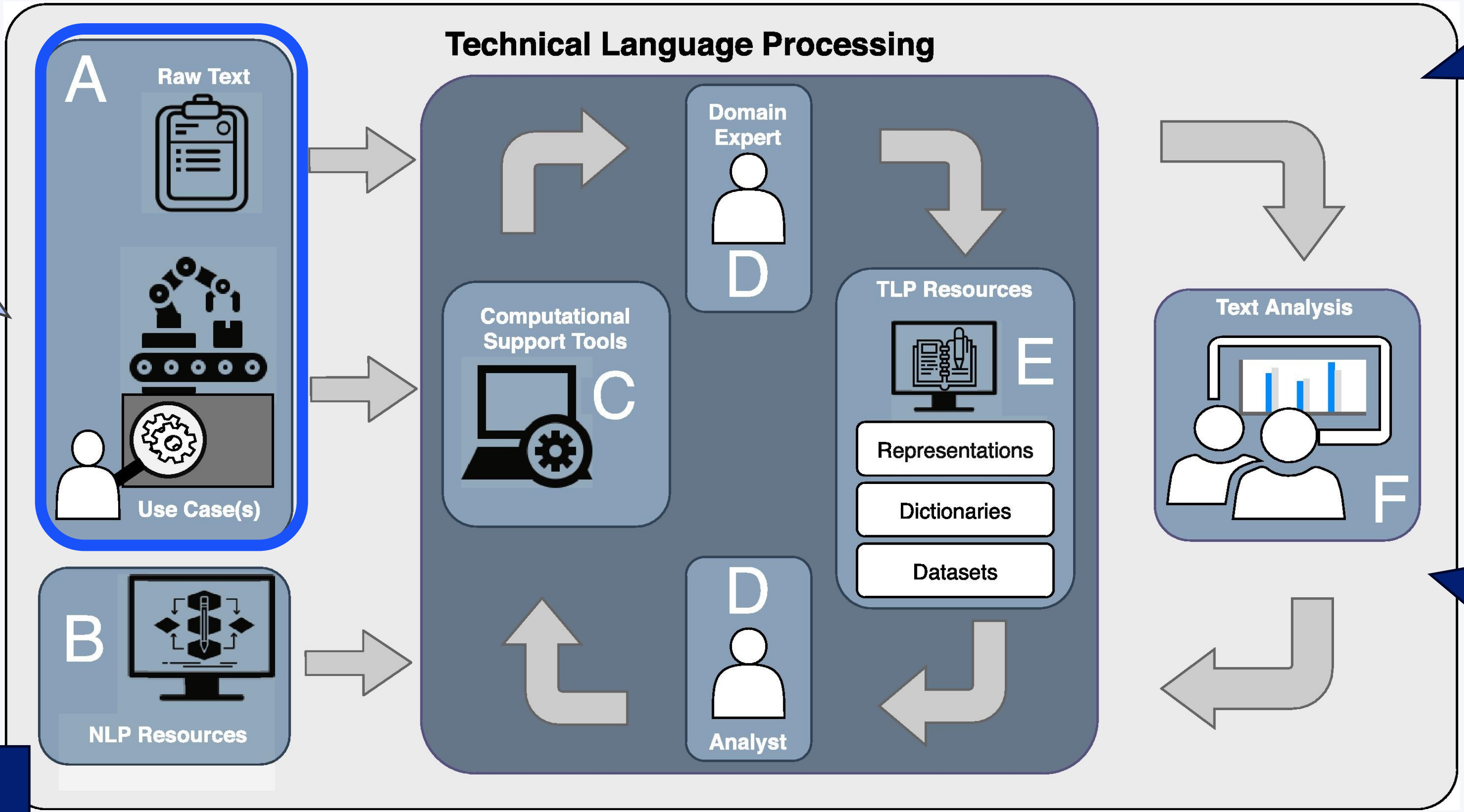
Brundage, Sexton, Hodkiewicz, Dima & Lukens (2021). Technical language processing: Unlocking maintenance knowledge. *Manufacturing Letters*, *27*, 42-46.

https://www.sciencedirect.com/science/article/pii/S2213846320301668

LMI

What is Technical Language Processing?

A. Use case is explicitly considered as an input



C. Use computer tools to alleviate burden on domain expert

D. Collaboration between analyst and domain expert

E. TLP resources – such as fortuitous data, ontologies or technical manuals

B. Use NLP resources where it makes sense

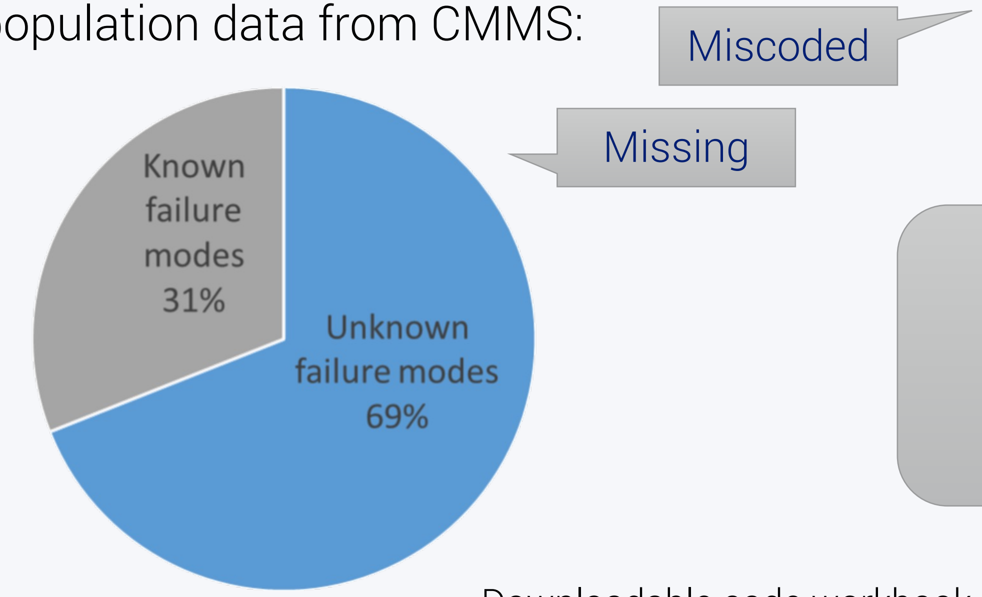
Brundage, Sexton, Hodkiewicz, Dima & Lukens (2021). Technical language processing: Unlocking maintenance knowledge. *Manufacturing Letters*, *27*, 42-46.

https://www.sciencedirect.com/science/article/pii/S2213846320301668

Use case: Overcome data quality challenges

- Challenges in maintenance work order data quality:
  - Data is largely missing or miscoded
  - Breakdown indicator rarely used
  - Cost data is generated for financial reporting and may lack engineering information

Example of equipment population data from CMMS:



Free Text Work Order Description	Miscoded Event Type
Repair leaking safety valve	Preventative (PM)
Daily Inspection of Analyzers	Repair

Identify areas where the data is good

Downloadable code workbook and dataset for basics in getting started in Python at addressing some of these dirty data issues: https://phmsociety.s3.amazonaws.com/Dirty_data_workshop.zip

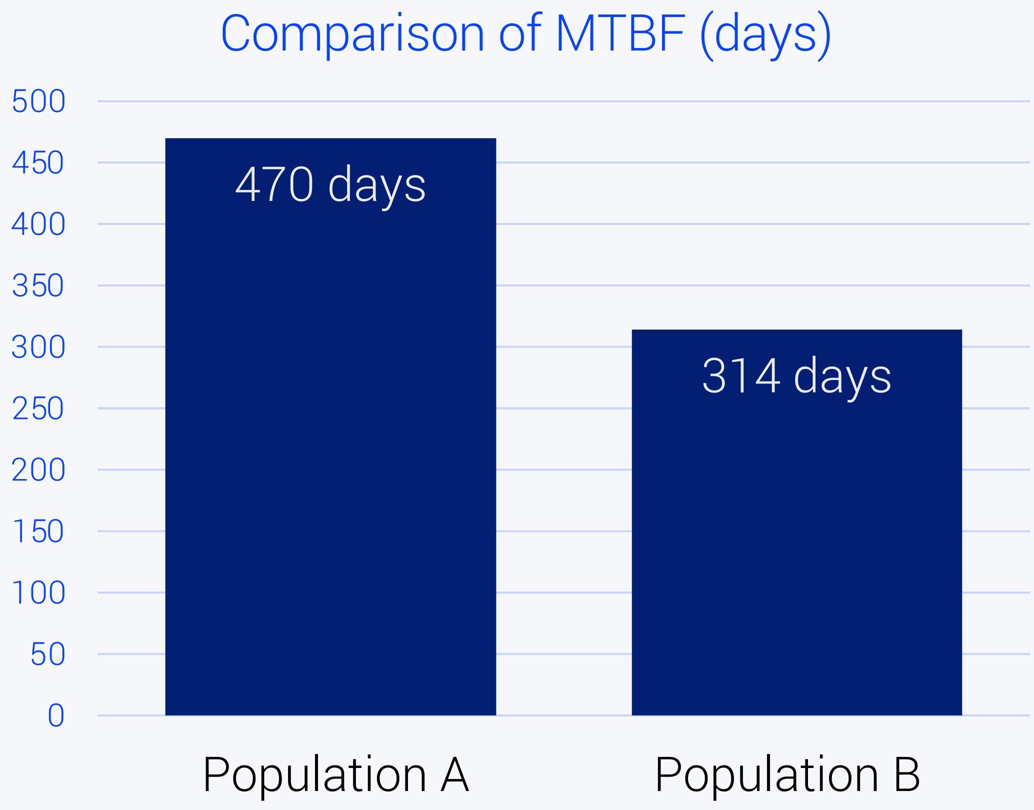
LMI

Use Case: Evaluate reliability metrics (despite dirty data)

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

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

After: Benchmarking comparison of MTBF possible



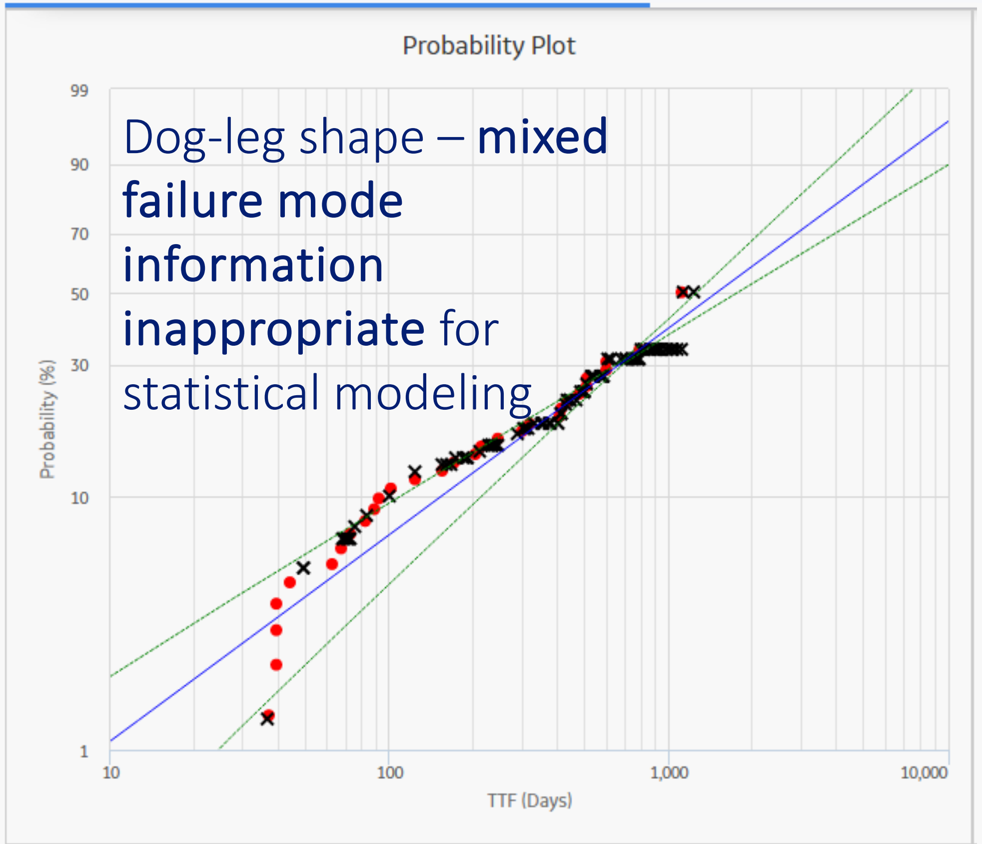
LMI

Best Practices Framework For Improving Maintenance Data Quality to Enable Asset Performance Analytics; Lukens, Naik, Saetia, Hu
https://papers.phmsociety.org/index.php/phmconf/article/view/836

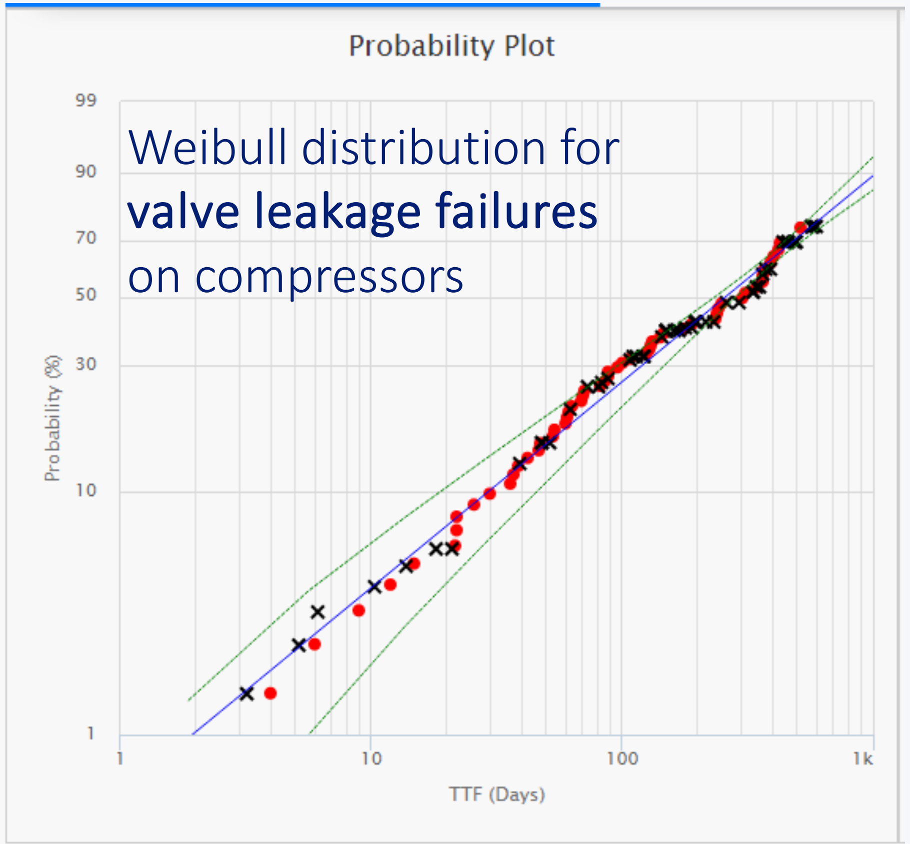
Use Case: Enables reliability distribution fitting

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

Before

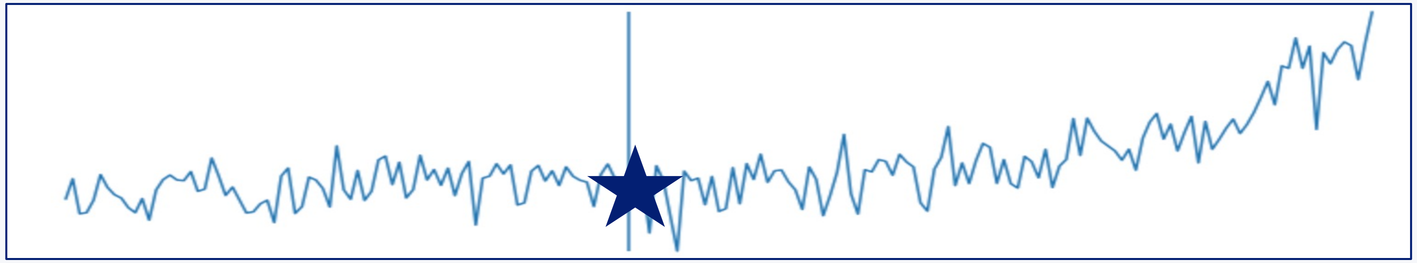


After



LMI

Use case: labeling condition monitoring data for diagnostics



t1

Time	Desc
t1	Fuel pump leaking
t2	Degraded valve

Technical Language Labeling: Create fault mode labels for PHM datasets

Technical Language Supervision for Intelligent Fault Diagnosis in Process Industry

November 01, 2023
2:00 PM – 3:30 PM
Canyon A

Speakers: Karl Lowenmark, Cees Taal, Stephan Schnabel, Marcus Liwicki, Fredrik Sandin

Labelling of Annotated Condition Monitoring Data Through Technical Language Processing

November 02, 2023
9:00 AM – 10:30 AM
Seminar Theater

Speakers: Karl Lowenmark, Cees Taal, Amit Vurgaft, Joakim Nivre, Marcus Liwicki, Fredrik Sandin

Bonus! Karl Löwenmark sent slides explaining Technical Language Supervision, which is a deep learning approach for combining technical language data with industry signals – found at end of this deck!

Matching images and text for bearing fault diagnosis

Few-shot Learning for Plastic Bearing Fault Diagnosis – An Integrated Image Processing and NLP Approach

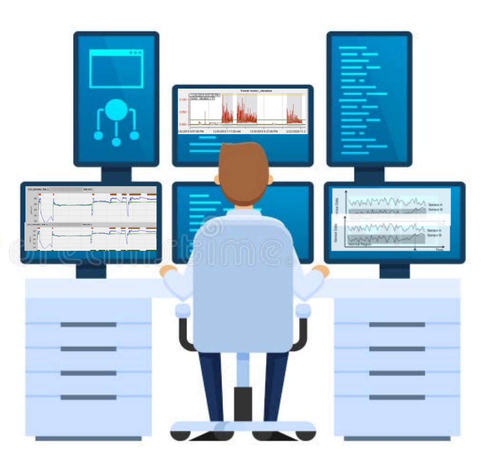
November 01, 2023
3:45 PM – 5:15 PM
Canyon A

Speakers: David He, Miao He

Use Case: Maintenance recommender for PHM system

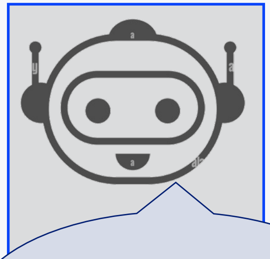
Data collection on an industrial asset

Predictive modeling capabilities
(models and/or rules which use the collected data)



Analyst in M&D Center
Assesses model outputs and makes **recommendations** for initiating appropriate actions

Maintenance recommender



Check for loose wires

Initiate action based on the information from extracted data

Alert: increased bearing temperatures on pump

Similar past cases

Date	What?	Action taken
1/2/23	Pump cavitating	Replaced pump rotor
4/5/23	lube oil bearing degradation	Replaced lube oil
6/8/23	Failed bearing	Replaced bearing

Troubleshoot:

1. Check pump for noisy bearings and cavitation
2. Check bearing oil for water and discoloration
3. …

LMI

Engineering Practice

- Goal & Approach — *"State the methods followed and why."*
- Assumptions — *"State your assumptions."*
- Measure & Evaluate — *"Apply adequate factors of safety."*
- Validate — *"Always get a second opinion."*

Hutcheson, M. L. (2003). *Software testing fundamentals: Methods and metrics*. John Wiley & Sons.

https://rimas.kudelis.lt/~rq/pub/Books/software%20testing/Wiley%20and%20Sons%20-%20Software%20Testing%20Fundamentals.pdf

LMI

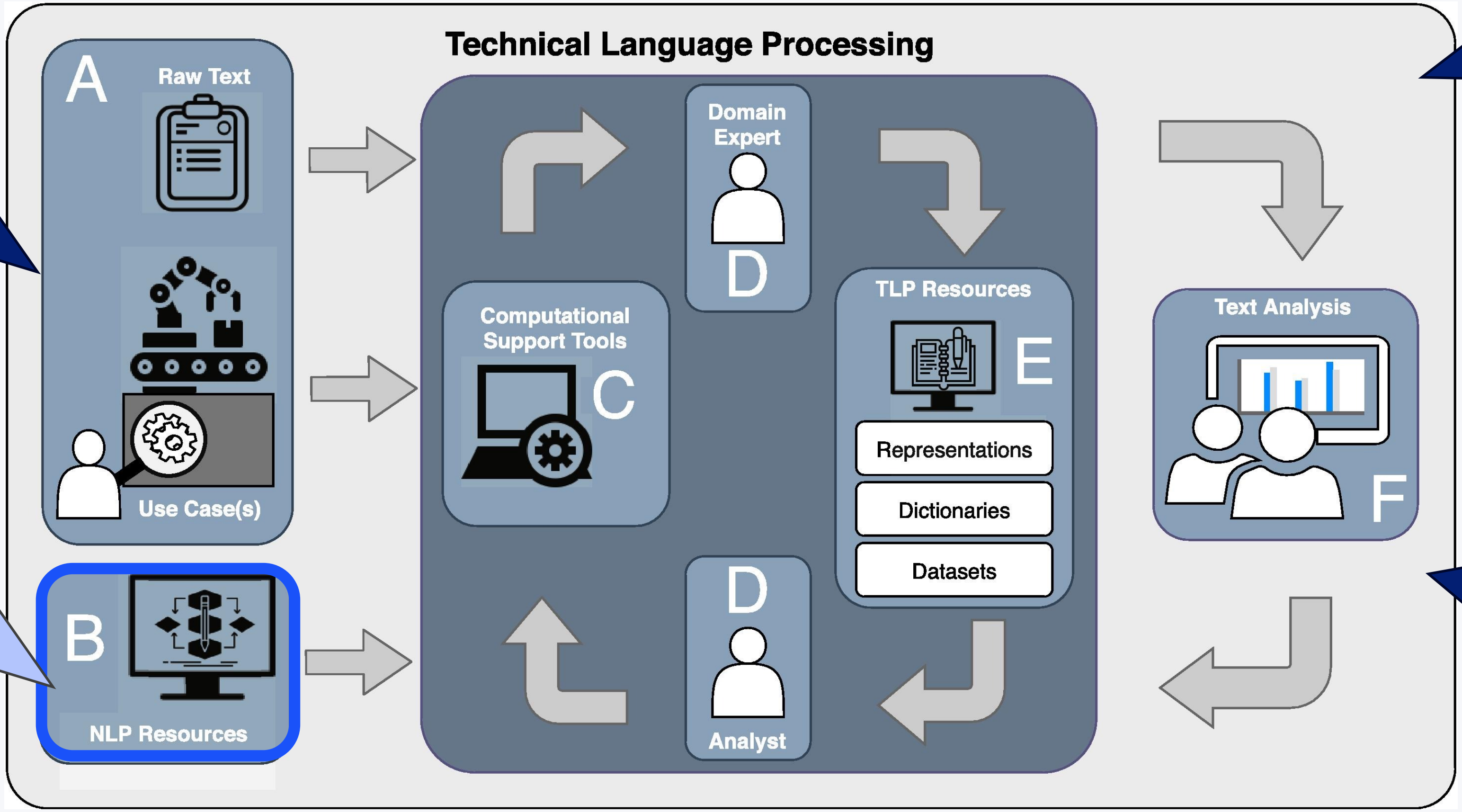
Using NLP resources where it makes sense: Concepts (and Assumptions!) from mainstream AI

LMI

What is Technical Language Processing?

A. Use case is explicitly considered as an input

B. Use NLP resources where it makes sense



C. Use computer tools to alleviate burden on domain expert

D. Collaboration between analyst and domain expert

E. TLP resources – such as fortuitous data, ontologies or technical manuals

LMI

Engineering Practice

- Goal & Approach — *"State the methods followed and why."*
- Assumptions — *"State your assumptions."*
- Measure & Evaluate — *"Apply adequate factors of safety."*
- Validate — *"Always get a second opinion."*

Hutcheson, M. L. (2003). *Software testing fundamentals: Methods and metrics*. John Wiley & Sons.

https://rimas.kudelis.lt/~rq/pub/Books/software%20testing/Wiley%20and%20Sons%20-%20Software%20Testing%20Fundamentals.pdf

LMI

Core NLP concepts: NLP pipelines

Raw data → Pre-processing → Representation → Analysis task

Pre-processing: Cleans up text such as lowercase, punctuation removal, etc.

Representation: Converts text to form which can be used by a numerical algorithm

Analysis task: Numerical algorithm

ID	Text
1	My pump failed
2	Pumps ran okay
3	Turbine ran okay
4	Equipment failed

? ? ? → Numerical Algorithm → Prediction!

Basic Bag of Words (BoW)

Vectorization: process of turning a collection of text documents into numerical feature vectors

ID	Text	Class
1	My pump failed	TRUE
2	Pumps ran okay	FALSE
3	Turbine ran okay	FALSE
4	Equipment failed	TRUE

Modifications such as re-weighting schemes:

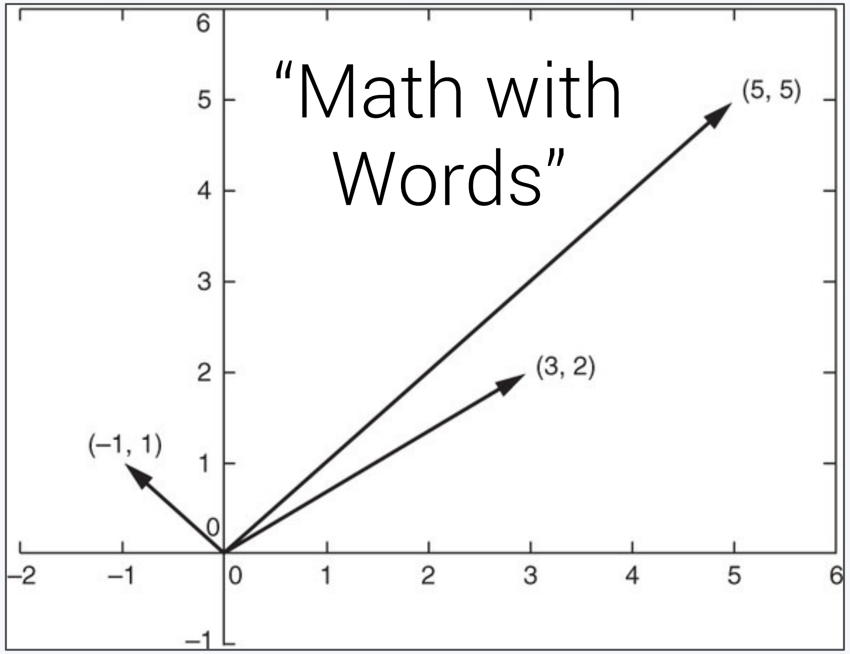
- Normalization
- TF-IDF: for "relevance ranking"

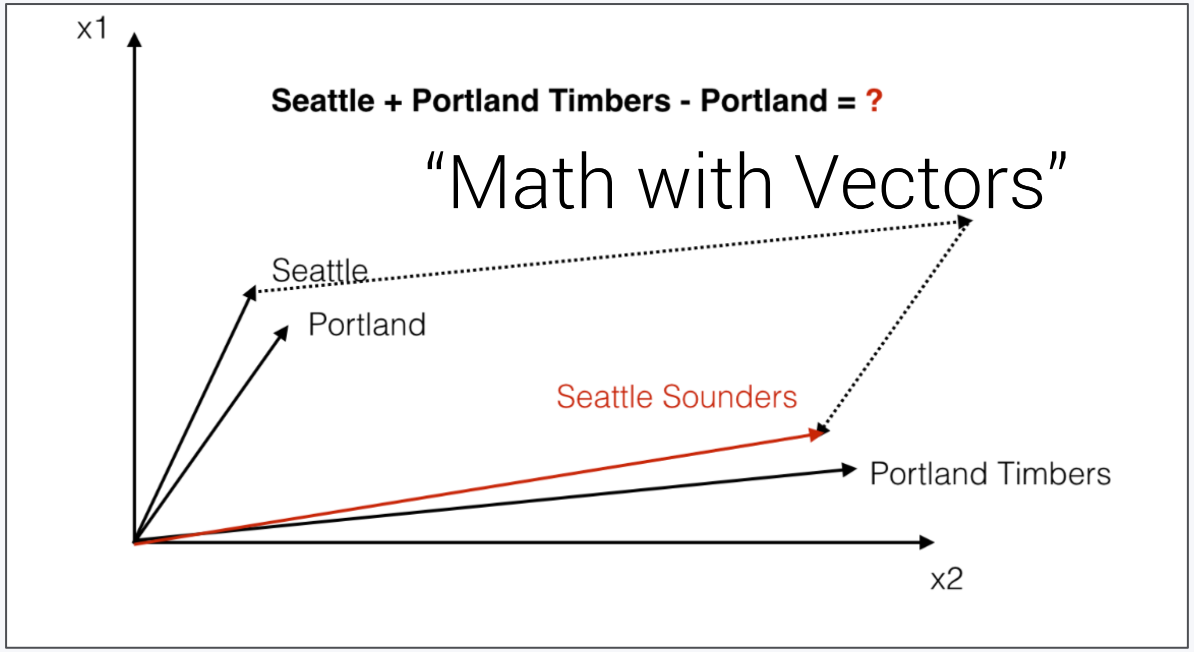
ID	Pump	Failed	Ran	Okay	Turbine	Equipment	Class
1	0.71	0.71	0	0	0	0	1
2	0.58	0	0.58	0.58	0	0	0
3	0	0	0.53	0.53	0.67	0	0
4	0	0.62	0	0	0	0.79	1

"My pump failed" will be numerically identical to **"Failed pump"** because word order is not preserved with BoW.

LMI

Similarity between documents and words





Context similarity refers to words that appear near or next to each other.

Each document treated as vector. Sentence similarity measured using cosine similarity.

Semantic similarity refers to words that are similar to each other.

Each word is treated as a vector….

LMI

Natural Language Processing in Action 2nd Edition (2024); Lane & Dyshel; Manning.

Representing a word as a vector – word embeddings

- **Word2Vec** (2013) "trains" a neural network on a word-level
- Building word vectors: teach a network to predict words near the target word in your sentence
  - Continuous Bag-of-Words (**CBOW**): target word from local context
  - **Skip-Gram**: local context from target word

Example: Word2Vec context building. With:

- Window size = 1
- Embedding dimension is user specified

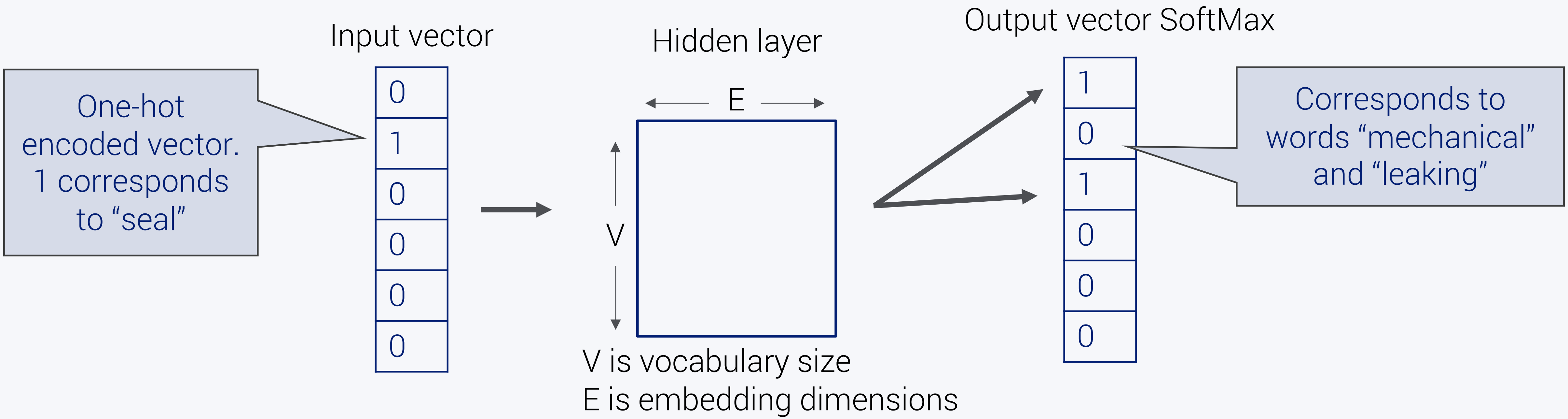
ID	Document
1	mechanical **seal** leaking steam on pump
2	replace leaking **seal**
3	**seal** leak
4	inboard **seal** leaking oil

→

Input word	Skip-gram input pairs for the word ("seal") with window size 1
seal	(seal, mechanical), (seal, leaking)
seal	(seal, leaking)
seal	(seal, leak)
seal	(seal, inboard), (seal, leaking)

Where word vectors come from

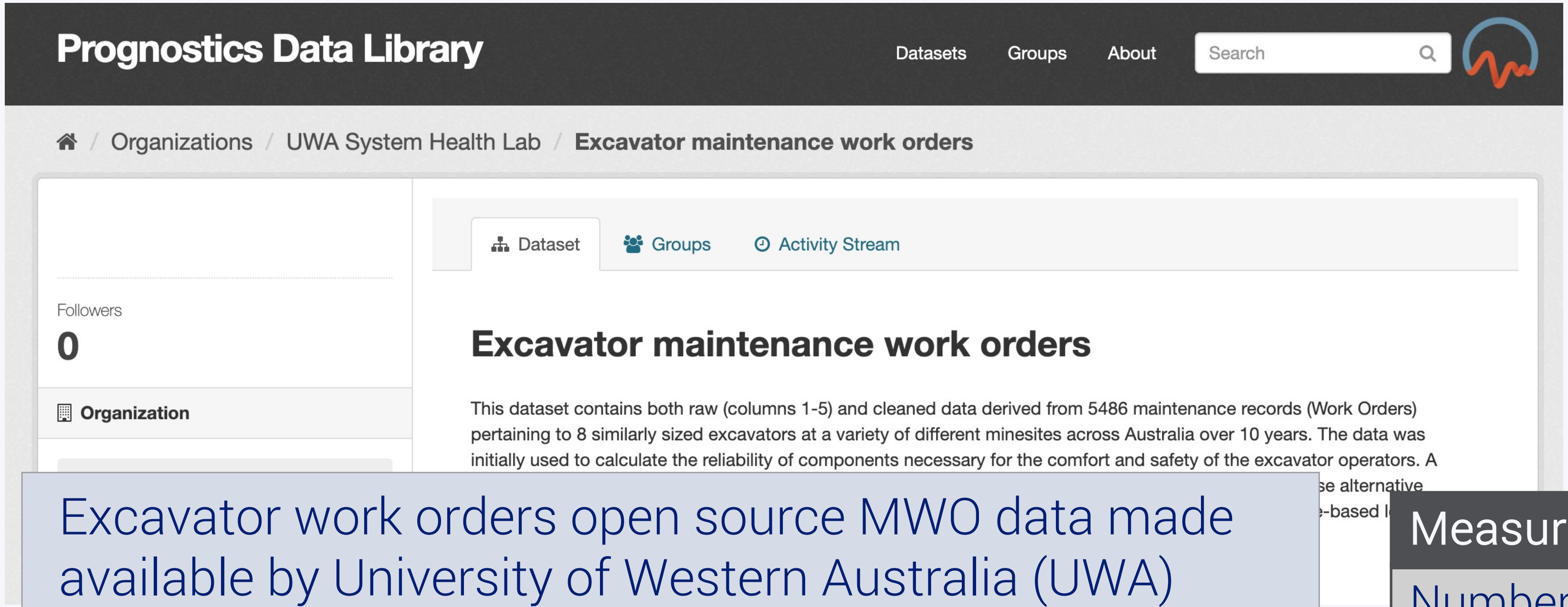
"mechanical **seal** leaking steam on pump"



Key Points:

- Word2Vec models are *trained* based on best matching inputs (words) with outputs (surrounding words).
- Words will have similar *word vectors* with other words that have similar surrounding words

Word similarity examples using Excavator work orders



Excavator work orders open source MWO data made available by University of Western Australia (UWA)

Date	Asset ID	Original Short text	Cost ($)
6/16/06	C	C/OUT RH NO-3 TRACK ROLLER-FAILED	17719.07
07/04/11	C	Repair hyd oil leak	2609
1/23/11	C	OIL LEAK ON BOOM PIPING	1317.13
1/20/05	D	Replace LH turbo	2212.87
9/24/08	C	RECTIFY ELECTRICAL FAULT (LOW PRESSURE)	115.9

Simple Preprocessing:

1. Lowercase all text
2. Replace all special characters with a space
3. Remove tokens which are numbers

Measure	Original	Cleaned
Number of documents	5485	5485
Vocabulary size	2931	1825
# words used 80%	545	238
% vocab used 80%	9.9%	4.3%
Mean word count per document	4.8	4.8
St. deviation	1.6	1.6

LMI

https://prognosticsdl.systemhealthlab.com/dataset/excavator-maintenance-work-orders

Word Similarity (Semantic similarities)

- **Semantic similarity** refers to words that are closely related or similar in meaning

Category (Entity)	Word	Custom Word2Vec	Pre-trained Word2Vec (google news 300)	Pre-trained Word2Vec subset to my vocab
Action	replace	lh, rh, to, on, track, roller, repair, changeout	replacing, replacement, replaced, Replacing, replaces,	replacing, replacement, replaced, fill, install
	repair	rh, replace, on, to, track, roller, changeout	repairs, repairing, repaired, Repair, Repairs	repairs, repairing, repaired, fix, install
Problem	leak	replace, changeout, repair, rh, on, lh, to, pump	leaking, leaks, leakage, spill, leaked, seepage, spillage, pinhole_leak	leaking, leaks, contamination, valve, leaky
	broken	Roller, replace, lh, on, changeout, grease, bucket	broke, breaking, fractured, shattered, broken, cracked, break	broke, cracked, smashed, torn, damaged

1. Pre-trained Word2Vec from gensim https://pypi.org/project/gensim/

2. Nandyala, Lukens, Rathod & Agrawal (2021): https://papers.phmsociety.org/index.php/phme/article/view/2894

Word Similarity (Semantic similarities)

Category (Entity)	Word	Custom Word2Vec	Pre-trained Word2Vec (google news 300)	Pre-trained Word2Vec subset to my vocab
Item	bucket	repair, replace, on, h, to, pump, rh, in, boom, lh	buckets, basket, Bruce_Belzowski_scientist, baskets, dunk	buckets, bottle, shovel, pully, hose
	motor	replace, on, boom, lh, h, rh, leaking, in, hose, pump	motors, Remy_HVH_electric, ##kw_electric, Minn_Kota_®, ac_induction, automobile, motorbikes_scooters,	motors, car, engine, diesel, cylinder
	instrument	blown, spline, down, radio, unit, ring, seat, cooler, rhs, service	Instruments, woodwind_instrument, stringed_instrument, organ	trumpet, sound, horn, module, lever

ChatGPT's semantically similar words to instrument (10/5): **Equipment**, **apparatus**, **device**, **tool**, **instrumentality**

AIR HORN TRUMPET HAS FALLEN OFF

LMI

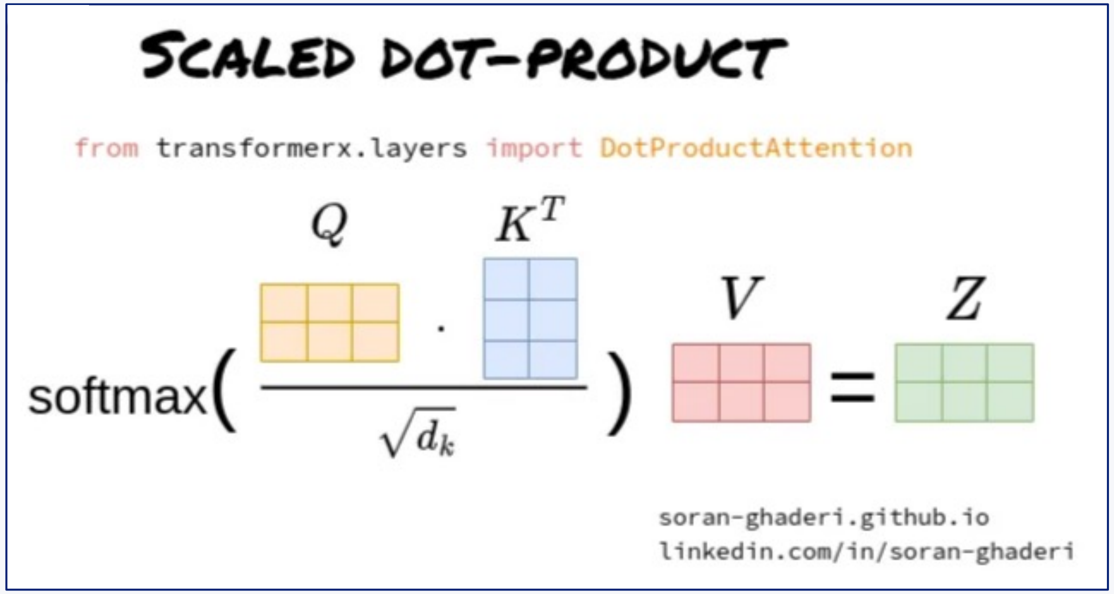
Preserving context to words in a sentence through self-attention

???

The heat exchanger was cleaned during inspection and light general pitting was observed on its bottom head

Idea: Apply a weighing to obtain a final word embedding which has more context than the initial embedding…. So… what if we multiply the initial word vectors for each word in a sentence by each other?

… yadda, yadda yadda – embeddings with context!



Representation Recap

Representation Family	Examples	Assumptions	Trade-Offs
Bag-of-Words (BoW) based	BoW, TF-IDF	Context: documents with similar words are similar	A lot of potentially useful information unused
Words as Vectors (word embeddings)	Word2Vec, GloVE	Words with similar vectors are semantically similar	Requires significant text data to pre-train
Transformer-based	BERT, GPT	Preserves the context of a word in its sentence	Data and computation resource hungry

A Novel Operations-Based Application of Natural Language Processing to Enhance Aircraft System Troubleshooting

November 02, 2023

9:00 AM – 10:30 AM

Seminar Theater

Speakers: Jamie Asbach, Daniel Wade

Example: Training word embeddings on fault codes from flight recorder data to look for patterns in fault occurrences

Engineering solutions with technical language data

LMI

Engineering Practice

- Goal & Approach — *"State the methods followed and why."*
- Assumptions — *"State your assumptions."*
- Measure & Evaluate — *"Apply adequate factors of safety."*
- Validate — *"Always get a second opinion."*

Hutcheson, M. L. (2003). *Software testing fundamentals: Methods and metrics*. John Wiley & Sons.

LMI

https://rimas.kudelis.lt/~rq/pub/Books/software%20testing/Wiley%20and%20Sons%20-%20Software%20Testing%20Fundamentals.pdf

Analysis tasks

Raw data → Pre-processing → Representation → Analysis task

Approaches: NLP Category	Description	Sub-task examples
Document typing	Categorizing or classifying documents into different types or categories	• **Document Clustering** • **Document Classification**
Document keywords	Extract & identify specific terms or phrases within a document that represents its most relevant content	• **Keyword extraction** • Multi-label classification
Entity Recognition	Identifying and classifying specific entities or objects in text.	• Named entity recognition (NER) • Parts-of-Speech (POS)
Others Worth Mentioning		Sentiment Analysis, Machine Translation, Question Answering…

LMI

Document typing: Text Classification

- Involves assigning predefined categories or labels to documents

Some possible categories for labeling maintenance work orders

Order type	Work Order Description	Was it a Failure?	Maintainable item	Failure Mode	Maintenance Action
CM	The north slurry inboard pump bearing has seized up, making a lot of noise.	?	?	?	?
CM	Pump bearings are in their final stage of failure. High vibration. Replace pump bearings.				
CM	Repair leaking oil cooler. Cooling water gaskets have failed on the cooler.				

Text Classification - mix & match!

Raw data → Pre-processing → Representation → Text Classification

Pre-processing

- Lowercase? Stop word removal? How to handle special characters?
- Tokenize? Pad? Fixed sequence length?

There's a reason machine learning engineering is a thing!

Representation

Bag of Words approaches

Word embeddings approaches:

- Choice 1: Create a "document vector" for each document
- Choice 2: Use word embedding as input layer to a Neural Network

Transformer-based approaches:

- Use embedding as input layers to a Neural Network

Text Classification

When each document is a vector, some choices:

- Random Forest
- LinearSVC
- Naïve Bayes
- Logistic Regression

When inputs are embedding layers:

- Design your architecture
- Dense? LSTM? CNN?
- Use Pre-trained, fine-tune, choose your own embeddings…

Engineering Practice

- Goal & Approach — *"State the methods followed and why."*
- Assumptions — *"State your assumptions."*
- Measure & Evaluate — *"Apply adequate factors of safety."*
- Validate — *"Always get a second opinion."*

Hutcheson, M. L. (2003). *Software testing fundamentals: Methods and metrics*. John Wiley & Sons.

LMI

https://rimas.kudelis.lt/~rq/pub/Books/software%20testing/Wiley%20and%20Sons%20-%20Software%20Testing%20Fundamentals.pdf

Text Classification using Supervised Learning

- Requires critical mass of labeled data! (Non-trivial)

FMC-MWO2KG (The MWO2KG Failure Mode Classification Dataset)

The Failure Mode Classification dataset released in the paper "MWO2KG and Echidna: Constructing and exploring knowledge graphs from maintenance data" by Stewart et al. The goal is to label a given observation (made by a maintainer) with the corresponding Failure Mode Code.

Each row contains an observation made by a maintainer, followed by a comma, followed by the Failure Mode, for example:

```
falure,Breakdown
```

As they are written in technical language, there are often spelling/grammatical/tokenisation errors made in the observations - these are typical of maintenance work orders.

The dataset comprises 502 (observation, label) pairs (for training), 62 pairs (for validation) and 62 pairs (for testing). The labels are taken from a set of 22 failure mode codes from ISO 14224. In order to pull a list of observations in which to label, we ran MWO2KG over the data once and exported a list of a
'leaking', 'not working') by the Named Entity Recognition model. We the
predicted as observations by the NER model and proceeded to label each c
mode code using a text editor.

Community-driven resources for performance evaluation!

Stewart, Hodkiewicz, Liu & French (2022).
https://journals.sagepub.com/doi/10.1177/1748006X221131128

https://paperswithcode.com/dataset/fmc-mwo2kg

Performance comparison for text classification

Batch	Model	F1-Micro	F1-Macro
Benchmark	Flair (LSTM-based)	0.60	0.46
	GPT 3.5 Fine-tuned	**0.81**	**0.62**
Bag of Words	RF	0.25	0.05
	LinearSVC	**0.54**	**0.38**
	MultinomialNB (tuned)	0.42 (0.52)	0.17 (0.33)
	LogisticRegression	0.48	0.27
Word Embeddings with averaged document vector	Word2Vec – Custom (300) – Random Forest	0.23	0.04
	Google-news-300 - SVC	**0.57**	**0.48**

Flair & GPT 3.5 Fine-tuned details & results in: Stewart, Hodkiewicz, Liu & French (2022); Stewart Hodkiewicz & Li (2023) https://arxiv.org/pdf/2309.08181.pdf

Pre-training took me 53 minutes… and cost me $0.39

BoW models took 0.01-0.02 seconds to train on a laptop

Google-news-300 takes 2 minutes to load.

Bert-base-uncased has 110M parameters- needs GPU!

Engineering Practice

- Goal & Approach — *"State the methods followed and why."*
- Assumptions — *"State your assumptions."*
- Measure & Evaluate — *"Apply adequate factors of safety."*
- Validate — *"Always get a second opinion."*

Hutcheson, M. L. (2003). *Software testing fundamentals: Methods and metrics*. John Wiley & Sons.

Re-assessing "The Pipeline". Reality is never as clean as the pipeline

https://rimas.kudelis.lt/~rq/pub/Books/software%20testing/Wiley%20and%20Sons%20-%20Software%20Testing%20Fundamentals.pdf

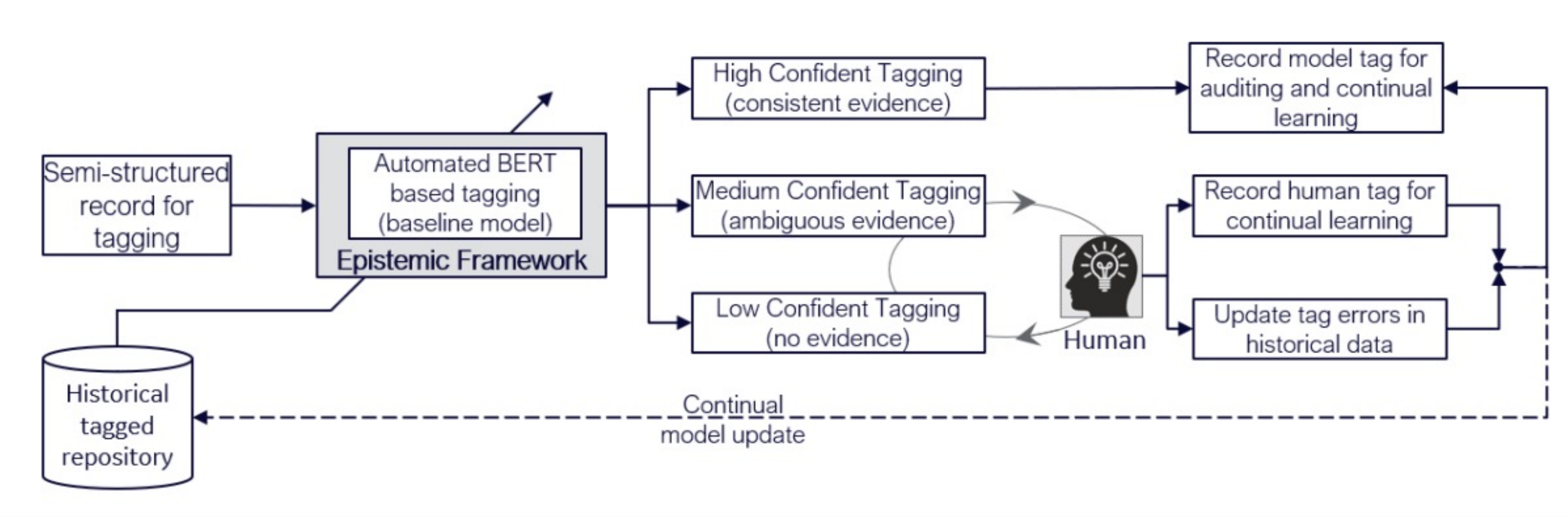
LMI

ML approaches to Text Classification – what could go wrong??

- The labeled data requirement… very hard to hand annotate
- Out-of-the-box tools are developed/pre-trained for different text
- Interpreting models & how to *explain* your result?
- Existing standard annotation or ontology?

I don't trust my data, why would I trust a model built using my data?

Example:



Mixed Initiative Approach for Reliable Tagging of Maintenance Records with ML

Iyer, Virani, Yang & Saxena (2022)

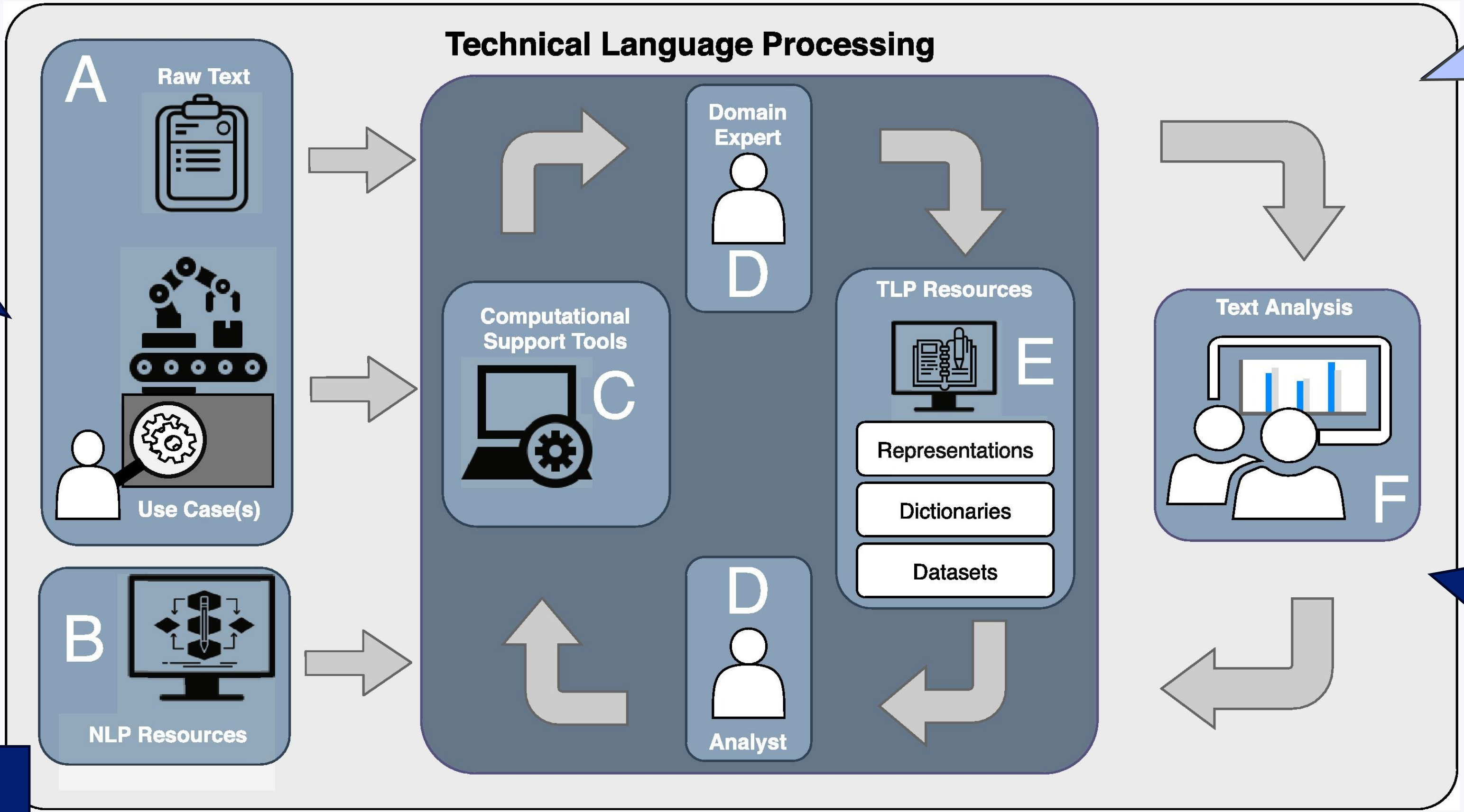
https://papers.phmsociety.org/index.php/phmconf/article/view/3159

LMI

What is Technical Language Processing?

A. Use case is explicitly considered as an input

B. Use NLP resources where it makes sense



C. Use computer tools to alleviate burden on domain expert

D. Collaboration between analyst and domain expert

E. TLP resources – such as fortuitous data, ontologies or technical manuals

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

LMI

Example 2: Classification using ML + Expert Rules

Stage 1:
- Pre-process data
- Train word embeddings (word2vec)

Creation of end-of-life (EOL) term gazetteer using semantic similarities

Stage 2:
Use gazetteer + expert rules to characterize if a work order describes an EOL event

Estimate Mean Time Between Failure (MTBF) for EOL events on pump components

Use NLP resources where it makes sense

Use computer tools to alleviate burden on domain expert

Term	Action
replcmt	replace
u/s	replace
snaped	replace
change out	replace
not pumping	replace
repair	repair
remount	repair

Description	Order Type	TAC ($)	RTF	EOL
$ B11 1A2B Mech Seal leaking steam on pu	Corrective	20,359	Reactive	True
C8 Piping design for Sulphur flow meters	Corrective	8,300	Reactive	False
3M Mech Lube Pump 1A2B	Preventative	310	Proactive	False
C9 JDI 1A2B not pumping	Corrective	1,339	Reactive	True
B11 1A2B Sulphur leaking from seal	Breakdown	30,832	Reactive	True
B11 Replace 1A2B Sulphur pump MJ	Corrective	36,237	Reactive	True
B12 1A2B Vibration Check / Housekeeping	Corrective	39,809	Reactive	False
B12 1A2B Pump not performing	Corrective	27,545	Reactive	True
$ C9 Replace pump 1A2B	Breakdown	12,466	Reactive	True

TLP resources – such as **fortuitous data**

Semi-automated Estimation of Reliability Measures from MWO's

Bikaun & Hodkiewicz (2021)

https://papers.phmsociety.org/index.php/phme/article/view/2950

LMI

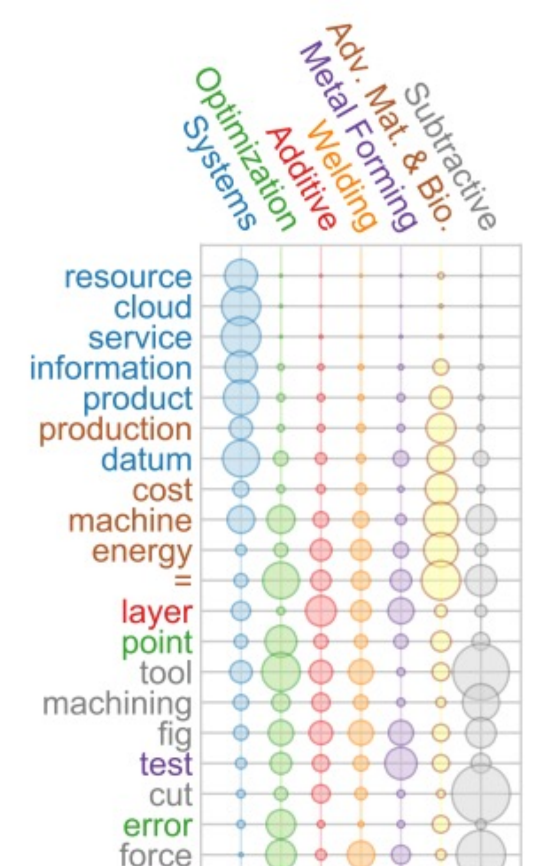
Document typing – Clustering (Unsupervised)

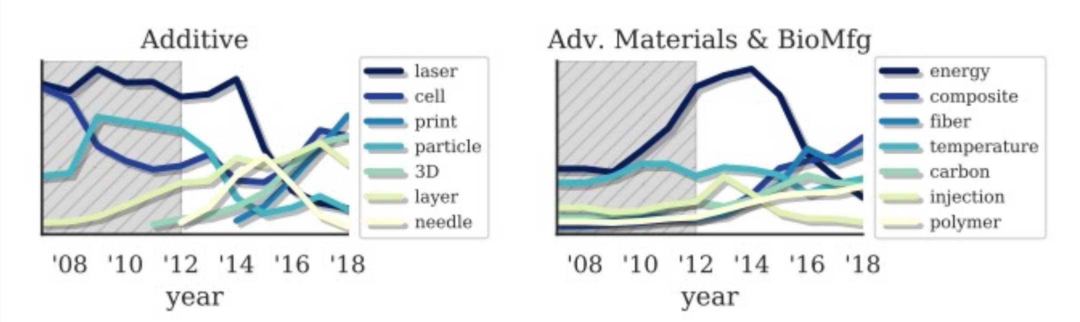
- What? Detect natural groupings for analysts to parse
- Example:
  - Topic Modeling to identify MSEC paper topics for trend analysis
  - Clustering of historical cases implemented in web app for M&D analyst to see most relevant historical cases

MSEC: A Quantitative Retrospective

Sexton, Brundage, Dima & Sharp (2020)

https://doi.org/10.1115/MSEC2020-8440





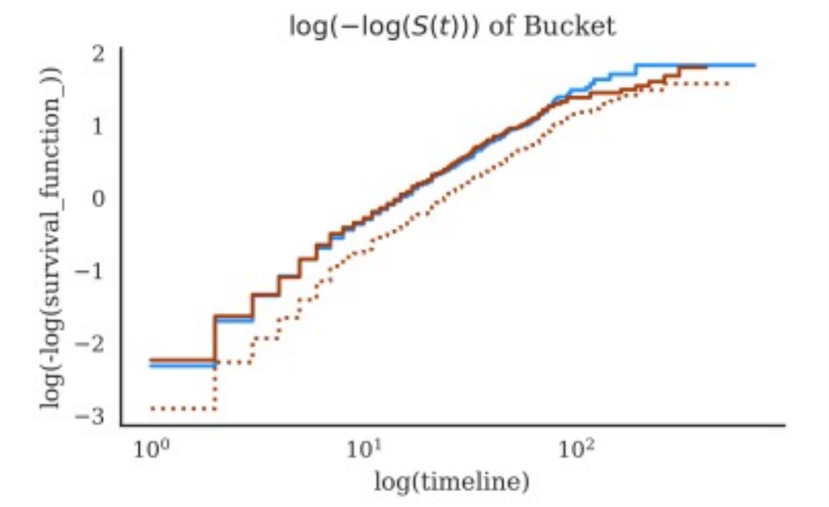
Collaboration & Support tools

Document keywords

- Keyword extraction (Unsupervised)
  - Use statistical properties to find “important terms”, such as TF-IDF
  - Example:

			MTTF (days)		Weibull Params.	
Major System	method	query	K-M	Weib.	β	η
Bucket	rules-based	Bucket	9.00	10.8	0.83±0.03	17±0.9
	single-tag	[bucket]	15.0	17.1	0.83±0.03	27±2
	multi-tag	[bucket, tooth, lip, pin]	9.00	10.5	0.82±0.02	16±0.9
Hydraulic System	rules-based	Hydraulic System	8.00	9.07	0.86±0.02	14±0.6
	single-tag	[hyd]	25.0	24.1	0.89±0.04	36±3
	multi-tag	[hyd, hose, pump, compressor]	9.00	9.74	0.89±0.02	15±0.7
Engine	rules-based	Engine	9.00	10.8	0.81±0.02	17±1
	single-tag	[engine]	10.0	11.8	0.79±0.03	19±1
	multi-tag	[engine, filter, fan]	8.00	9.31	0.81±0.02	15±0.8

Tagging work orders for subsystem survival analysis.



Benchmarking for keyword extraction methodologies in MWO’s

Sexton, Hodkiewicz, Brundage & Smoker (2018)

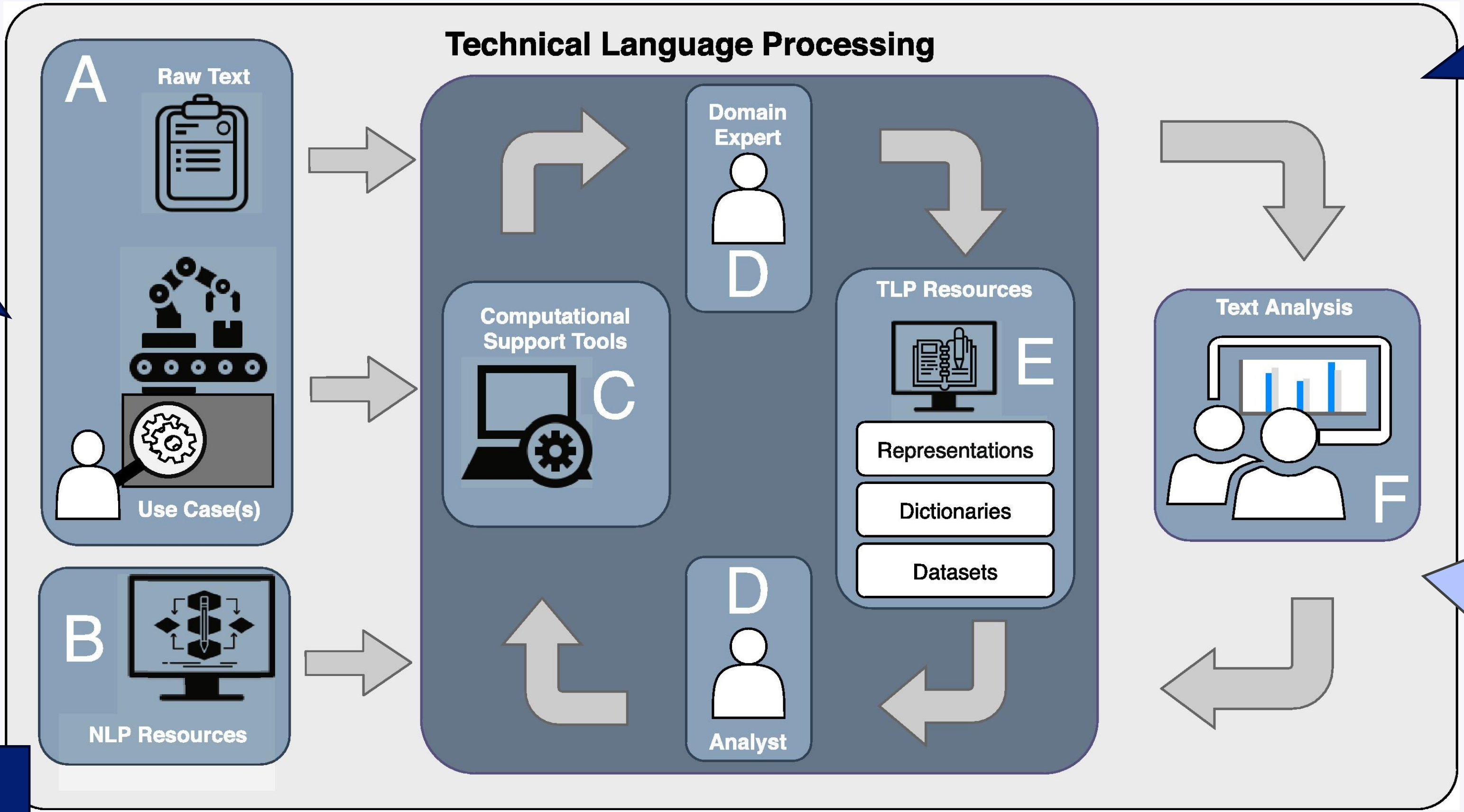
https://papers.phmsociety.org/index.php/phmconf/article/view/541

Introduction to keyword extraction in the downloadable code workbook
https://phmsociety.s3.amazonaws.com/Dirty_data_workshop.zip

LMI

What is Technical Language Processing?

A. Use case is explicitly considered as an input



C. Use computer tools to alleviate burden on domain expert

D. Collaboration between analyst and domain expert

E. TLP resources such as fortuitous data, ontologies or technical manuals

B. Use NLP resources where it makes sense

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

LMI

Introduction to ontologies in 1 slide

- Ontologies are knowledge frameworks that organize information in a structured & interconnected way
- Ontologies help bridge the gap between human understanding and machine processing

Some groups in industry:

- Industrial Ontologies Foundry (IOF): **working to co-create a set of open reference ontologies to support the manufacturing and engineering industry needs and advance data interoperability** https://www.industrialontologies.org/
  - **Jan 2023: first formal release** of the IOF Core Ontology
- ISO/AWI 23726-3 Automation systems and integration – Ontology-based Interoperability: **New standard to facilitate digitalization across various industries and domains by establishing a common digital vocabulary that enables the utilization of reference data within diverse standards** https://www.iso.org/standard/87560.html

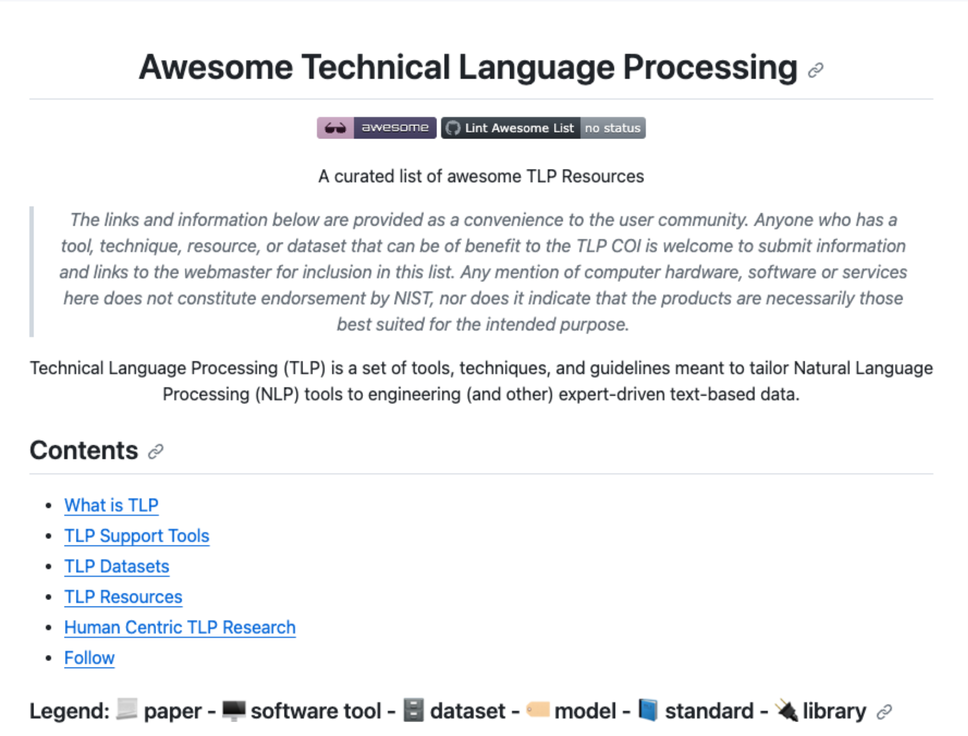
LMI

Community-Driven Resources in TLP

Awesome List: contains curated & compiled links to TLP Support Tools, Datasets, Resources (such as ontologies and standards), etc. https://github.com/TLP-COI/awesome-tlp

- Learning Resources:
  - "Text as Data: the Road to Technical Language Processing" Online Course: https://tlp-coi.github.io/text-data-course/home.html
  - 6 week program for industrial company by University of Western Australia: https://core-skills-master.webflow.io/

- Have another references or resources you'd like to share? Reach me through Whova app and I will add to this deck for the website!



Awesome Technical Language Processing

awesome | Lint Awesome List | no status

A curated list of awesome TLP Resources

> *The links and information below are provided as a convenience to the user community. Anyone who has a tool, technique, resource, or dataset that can be of benefit to the TLP COI is welcome to submit information and links to the webmaster for inclusion in this list. Any mention of computer hardware, software or services here does not constitute endorsement by NIST, nor does it indicate that the products are necessarily those best suited for the intended purpose.*

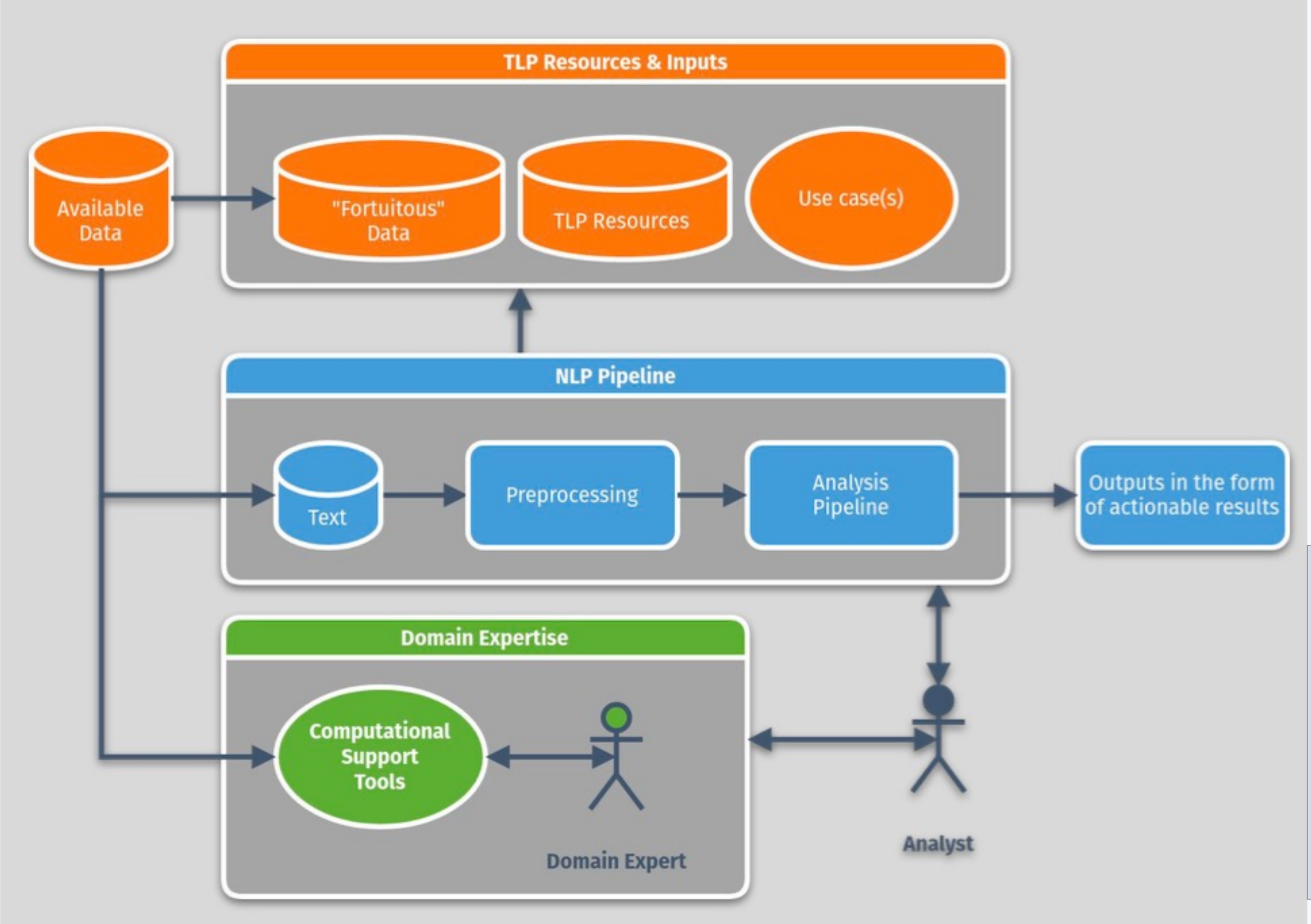
Technical Language Processing (TLP) is a set of tools, techniques, and guidelines meant to tailor Natural Language Processing (NLP) tools to engineering (and other) expert-driven text-based data.

Contents

- What is TLP
- TLP Support Tools
- TLP Datasets
- TLP Resources
- Human Centric TLP Research
- Follow

Legend: paper - software tool - dataset - model - standard - library

Technical Language Processing – to think about…



- How the TLP approach to *meaning* & *generalization* differs from NLP?
- How data *quantity* and *quality* can be addressed
- Potential risks of *not* adapting NLP

Adapting NLP for Technical Text.

Dima, Lukens, Hodkiewicz, Sexton & Brundage (2021).

https://onlinelibrary.wiley.com/doi/10.1002/ail2.33

LMI

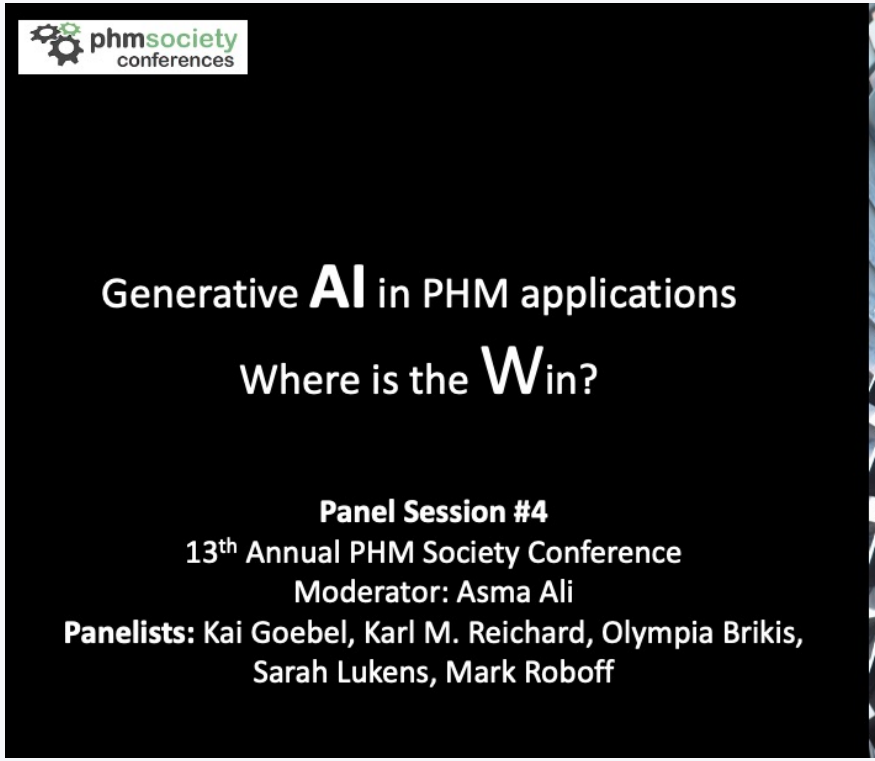
Wait!!! I didn't say anything about…

Generative AI

Large Language Models (LLM's)

There's a panel for that!
Tuesday Oct 31 at 11 AM
in Canyon B

Panel Session 4:
Generative AI and ML for PHM



Slides from panel: https://phm2023.phmsociety.org/wp-content/uploads/sites/13/2023/11/GenAI-Panel-at-PHM-10-31-2023.pdf

As the saying goes, Buffalo buffalo Buffalo buffalo buffalo buffalo Buffalo buffalo.

Thank you.

Sarah Lukens
sarah.lukens@lmi.org

Michael Sharp
michael.sharp@nist.gov

Rachael Sexton
thurston.sexton@nist.gov
rachaeltsexton@gmail.com

LMI

Feedback from the session

LMI

Feedback & resources from the conference (PHM 2023)

- Recommendations:
  - Connor Cabrey from Trident Systems recommends as a "great book for beginners in NLP and does a great job walking through the fundamentals" covered in the tutorial.
    - **Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit**
      - By Steven Bird, Ewan Klein and Edward Loper
      - https://sites.pitt.edu/~naraehan/ling1330/nltk_book.html
- From the conference (PHM 2023):
  - Keep an eye out for: Karl Löwenmark from Luleå University of Technology. On a trajectory towards making some content publicly available as he can… Links to conference papers:
    - https://paperswithcode.com/paper/technical-language-supervision-for
    - https://papers.phmsociety.org/index.php/phmconf/article/view/3507
  - Word2Vec approach for finding "semantic similarity" between aviation fault codes due to sequences of fault codes generated on a flight recorder:
    - https://papers.phmsociety.org/index.php/phmconf/article/view/3579
  - Images + NLP for labeling:
    - https://papers.phmsociety.org/index.php/phmconf/article/view/3575
  - Evaluating ChatGPT as a maintenance troubleshooting recommender: https://papers.phmsociety.org/index.php/phmconf/article/view/3487

See end of deck for overview slides from Karl on Technical Language Supervision!

LMI

Questions from the Whova app:

- "When looking at building out a specific technical language corpus, with the amount of technical jargon and slang in a potential maintenance corpus, do you see any value in applying a pre-trained Word2Vec model like the Google News or GloVe models to that kind of corpus? Or is building your own Word2Vec model off just your corpus more effective?"
  - This was mostly answered during the talk – especially in the fault classification example, where we crazily got really good performance using out of the box google-news-300 from genism package for embeddings.
    - However, the caviat (also covered) is that you should be aware of the limitations of a pre-trained model. They were not trained on technical data. The example in the tutorial was google-news-300's tendency to pick musical instruments as most semantically similar to "instrument" (while GPT-3.5 tends to pick scientific instruments as semantically most similar to "instrument").
    - So the unsatisfying "it depends" and you should have some qualitative analysis to explain behaviors.

Community Resources: Technical Language Supervision

Contributed by:
Karl Löwenmark
Luleå University of Technology

Contact: Karl Löwenmark karl.lowenmark@ltu.se

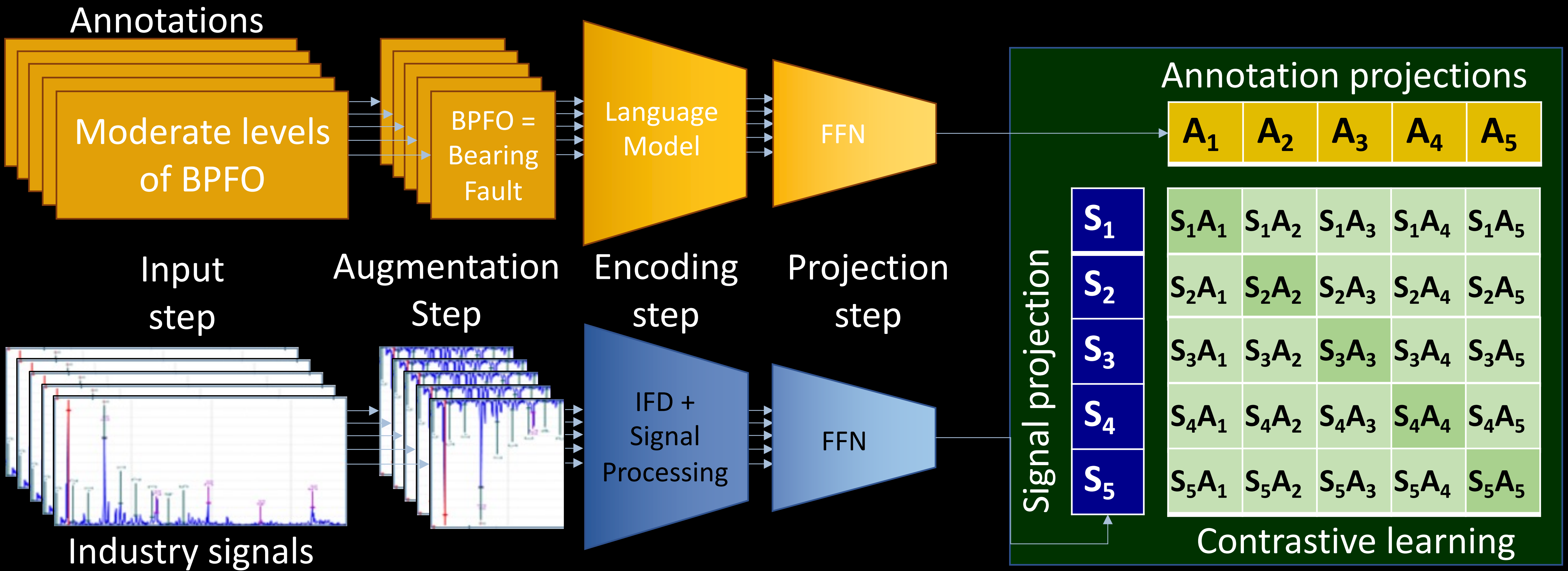
Paper links:

https://paperswithcode.com/paper/technical-language-supervision-for
https://papers.phmsociety.org/index.php/phmconf/article/view/3507

LMI

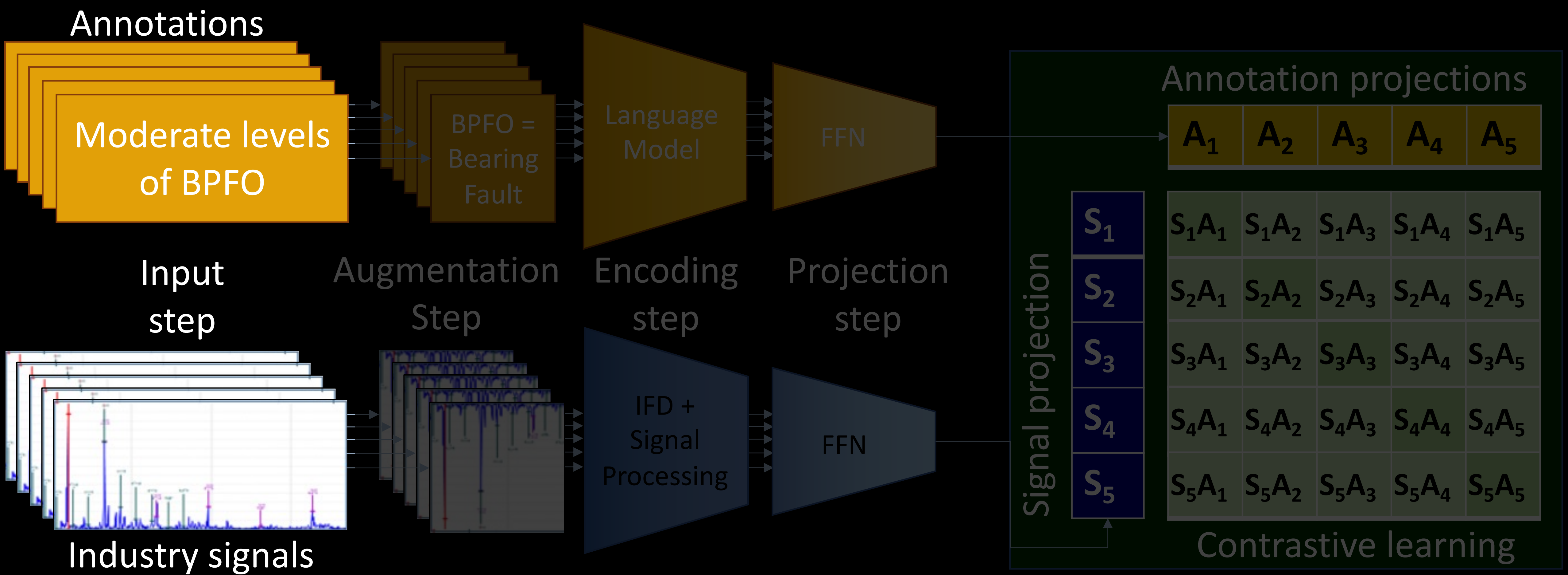
How can technical language and industry signals be combined?

Technical Language Supervision



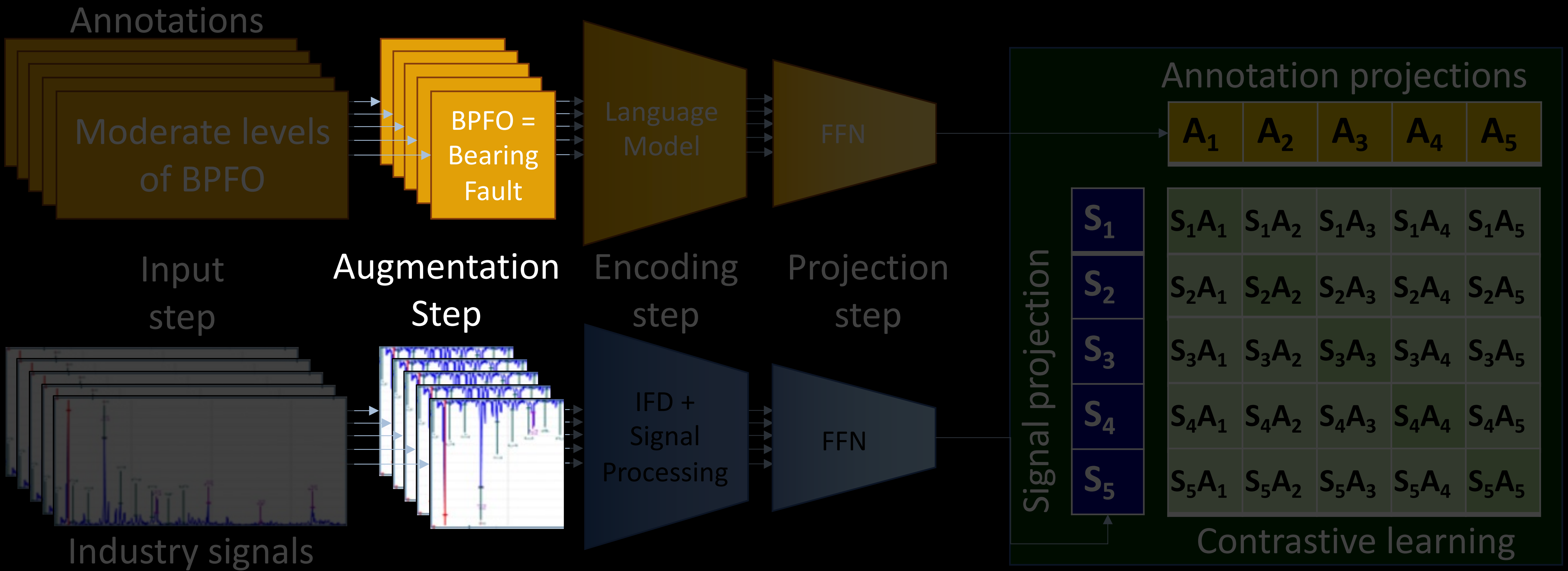
Input

Text-signal pairs based on annotated industry data



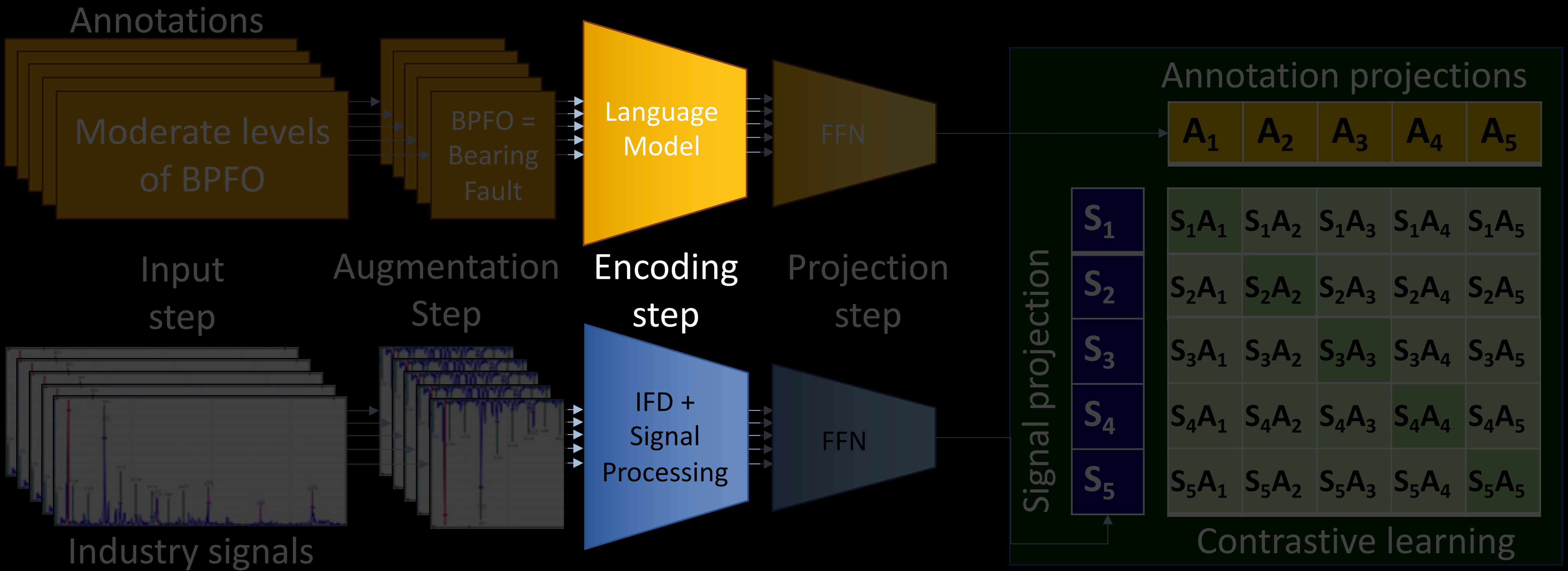
Augment

Integrate human knowledge, e.g. by defining technical words



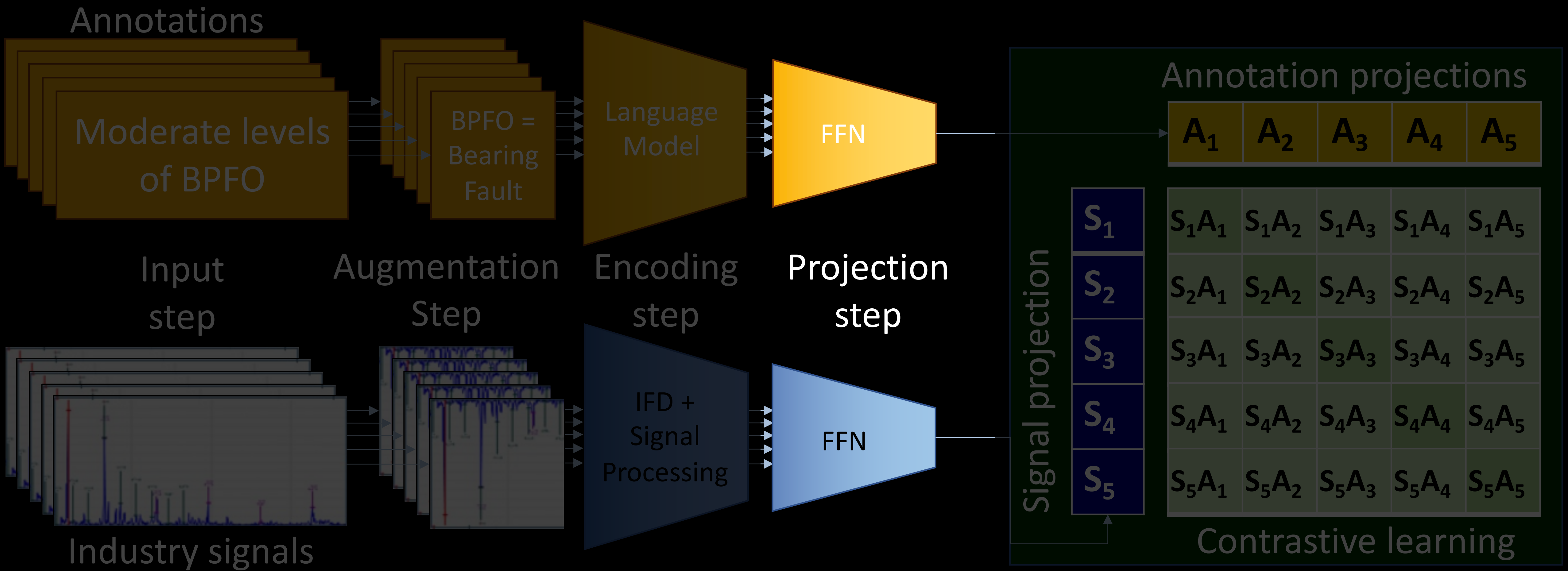
Encode

Using pre-trained or randomly initialised models



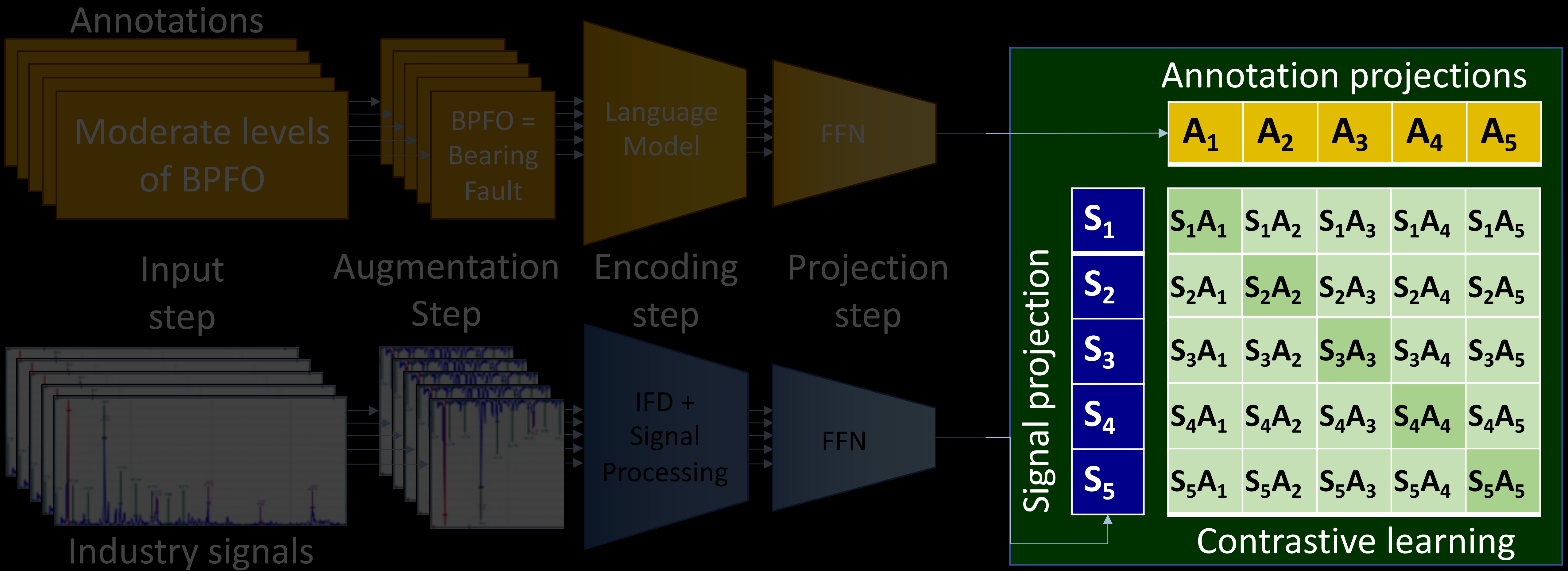
Project

From high-dimension encodings to lower-dimension projections



Learn

Objective: Text describing faults and signals containing them share projections (high dot product)



Infer

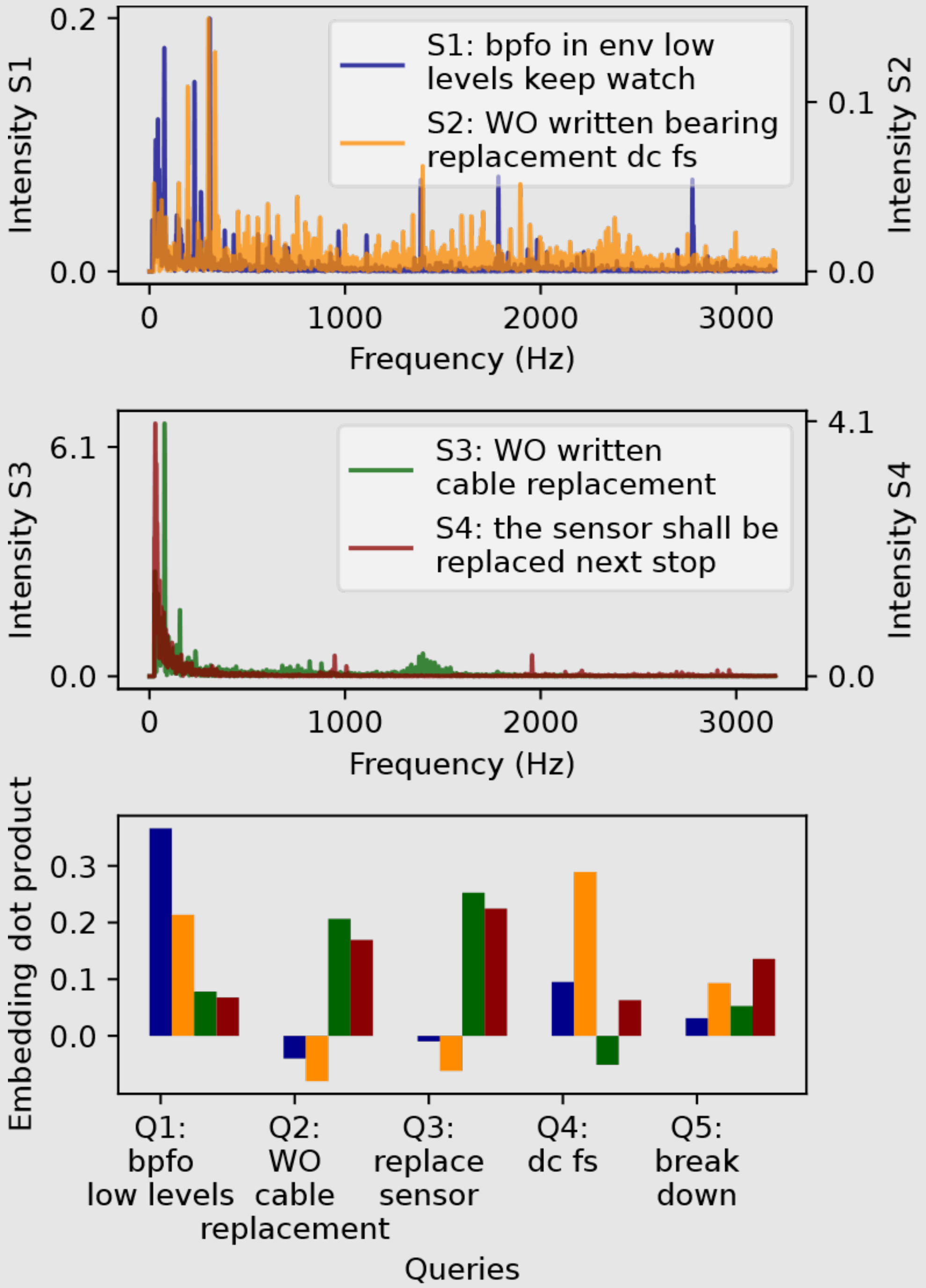
Text input: five queries Q1-Q5

Signal input: four spectra, with true annotations S1-S4

Coloured bars: projection dot product

Zero-shot diagnosis:

- S1, a BPFO signal, correlates with Q1, BPFO
- S2, bearing replacement, correlates with Q1 and Q4 (drive cylinder free side, a place for replacements)
- S3 and S4, cable and sensor faults, correlate with Q2 and Q3 - cable and sensor replacements



Conclusion

Technical language can, if processed, be used to trained models directly on industry data without requiring labels!