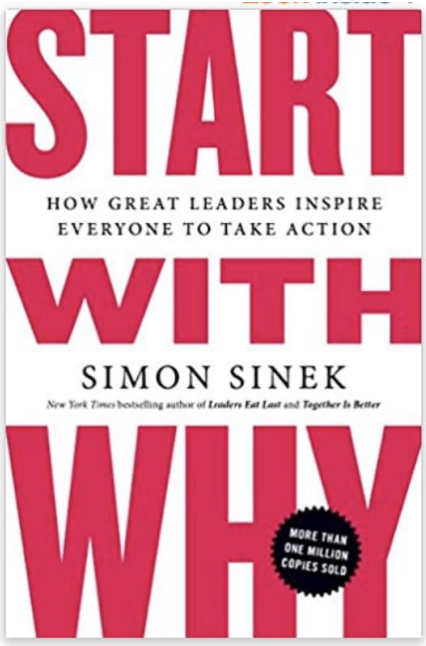
Industrial Perspective of Machine Learning and AI Challenges in PSE

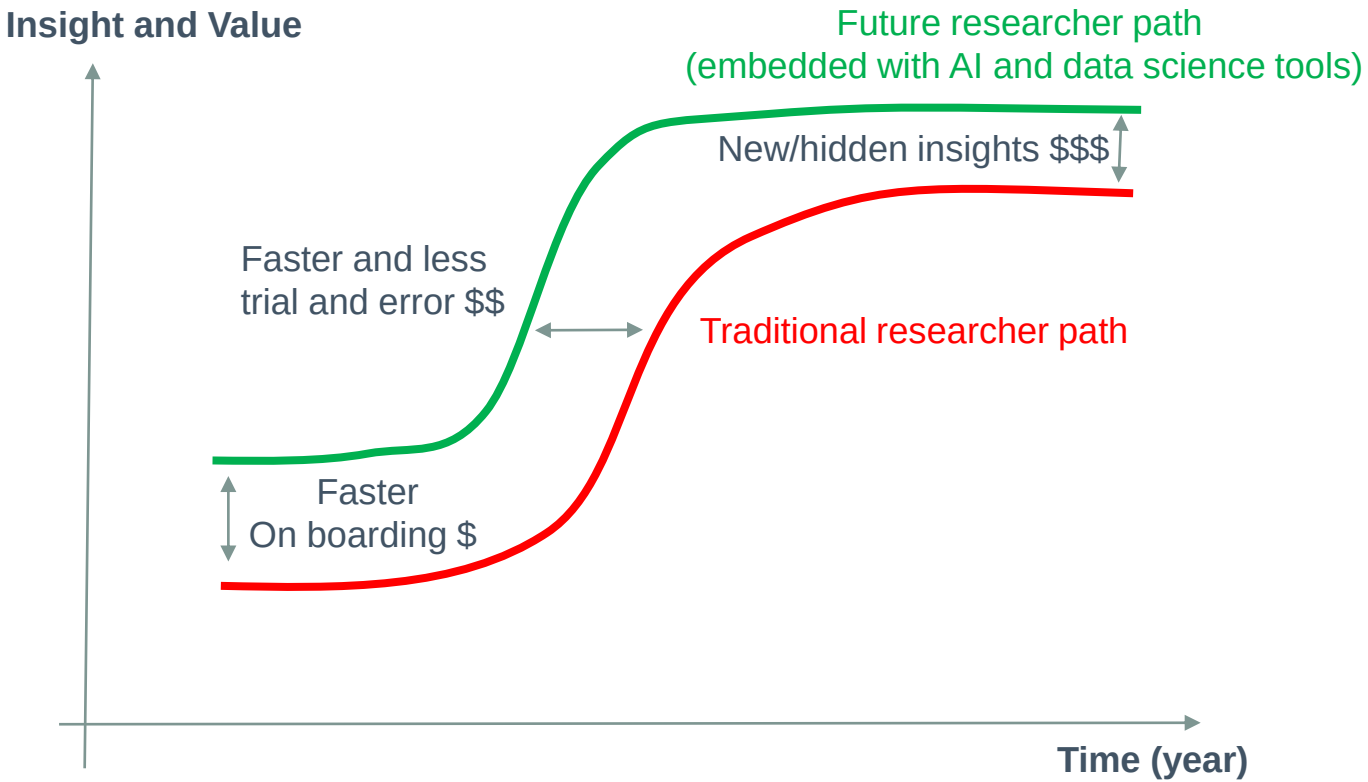
Leo Chiang

With many contributions from Team Dow including B. Braun, I. Castillo, Z. Wang, A. Schmidt, S. Mukhopadhyay and the growing Data Science CoP

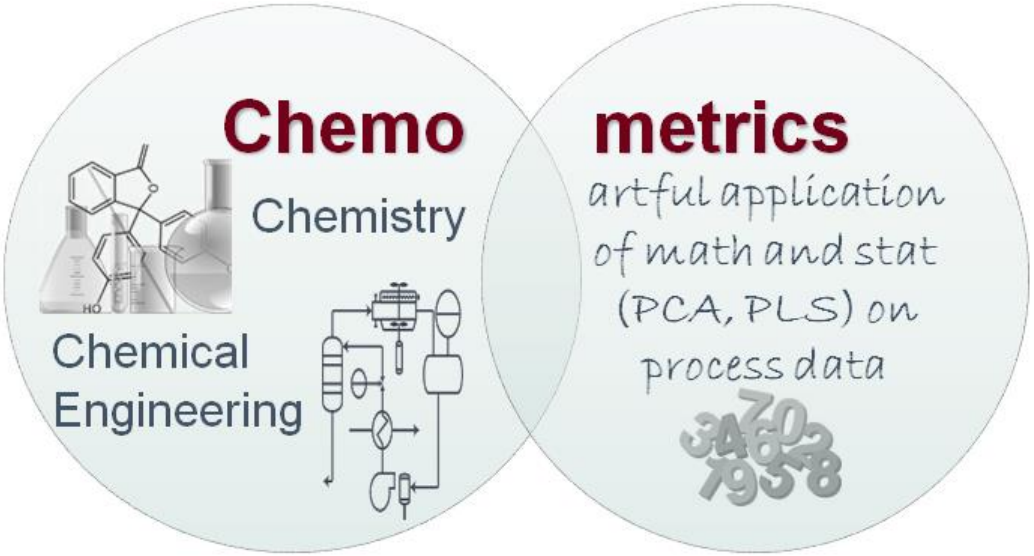
FIPSE-5 Conference
Crete, Greece
6/28/2022

Vision: Dow R&D to take the industry leading role in using AI and Data Science to accelerate materials discovery, new product development, and process modeling & optimization



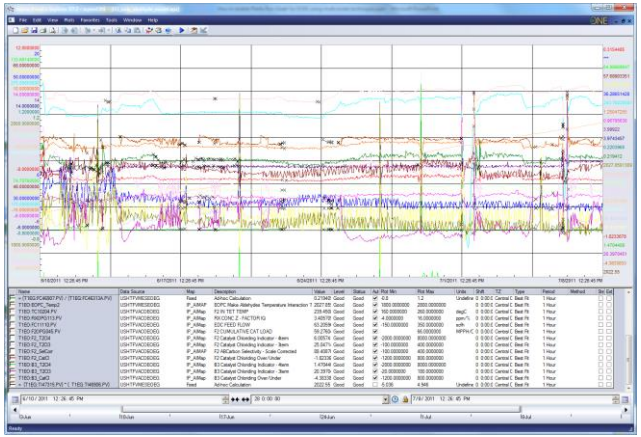


My Journey At Dow Started in 2001 With Chemometrics



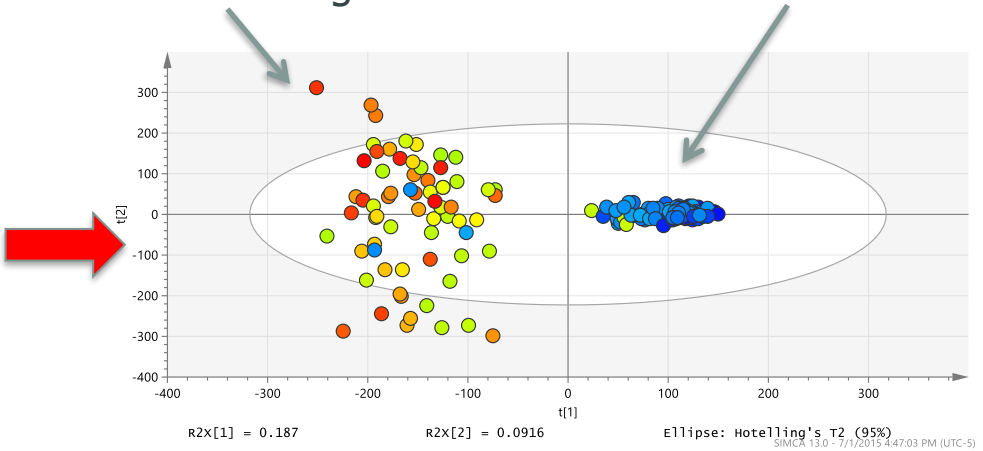
Power of Chemometrics

It is overwhelming to analyze 50+ process variables



Process conditions lead to fouling

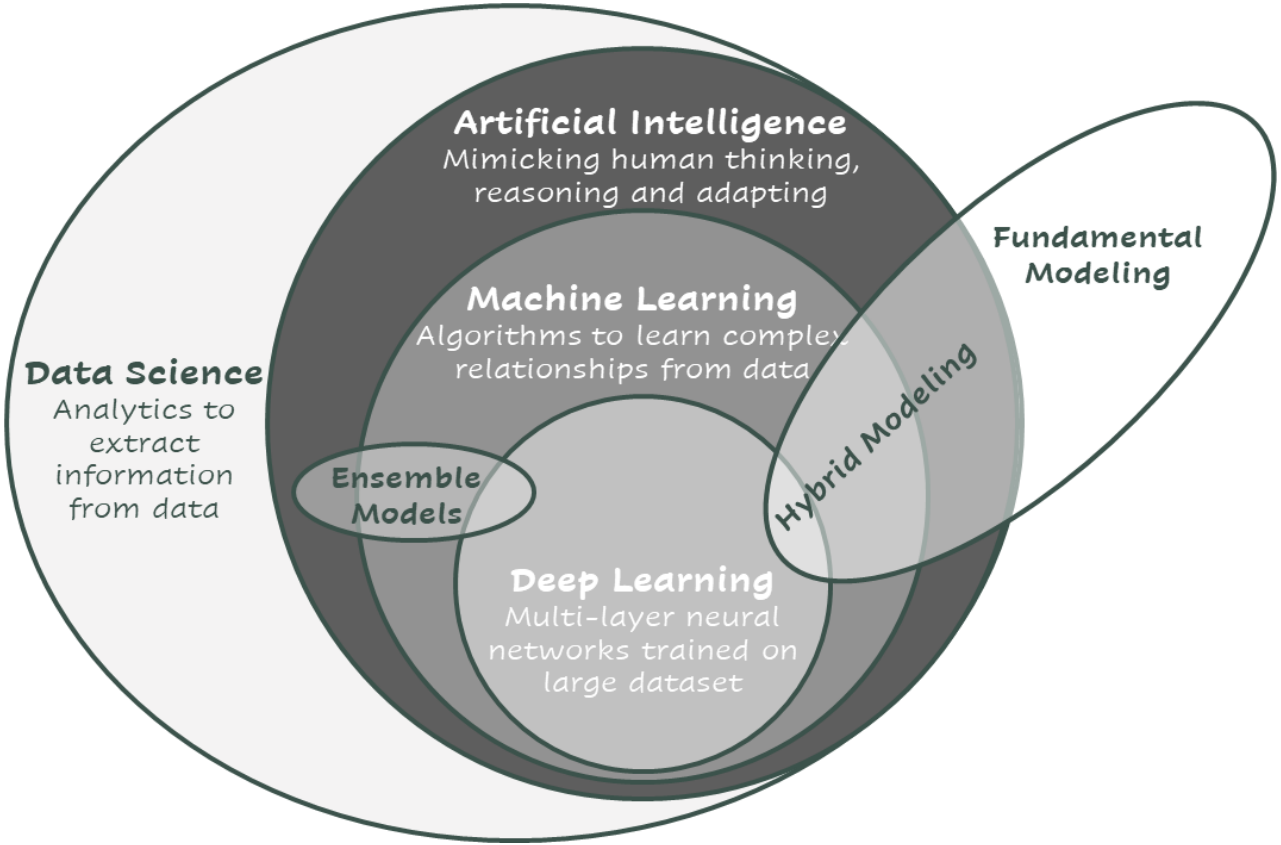
Normal conditions



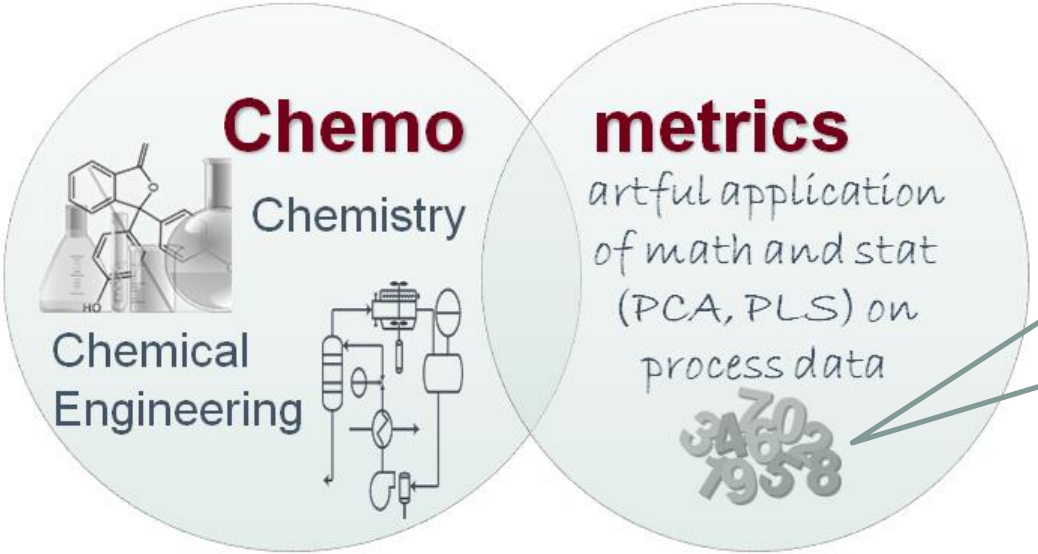
Chemometrics (PCA, PLS) increases **speed of problem solving**. We use existing data to support hypotheses, eliminate hypotheses, and generate new hypotheses

- L. Chiang, E. Russell, and R. Braatz, *Fault Detection and Diagnosis in Industrial Systems*, Spring-Verlag, 2001.
- L. Chiang, R. Leardi, R. Pell, and M.B. Seasholtz, Industrial experiences with multivariate statistical analysis of batch process data, *Chemo & Intel Lab Systems*, 81(2), 109-119, 2006.
- L. Chiang and L. Colegrove, Industrial implementation of on-line multivariate quality control, *Chemo & Intel Lab Systems*, 88(2) 143-153, 2007.

One Definition of AI, Machine Learning, Deep Learning, and Data Science

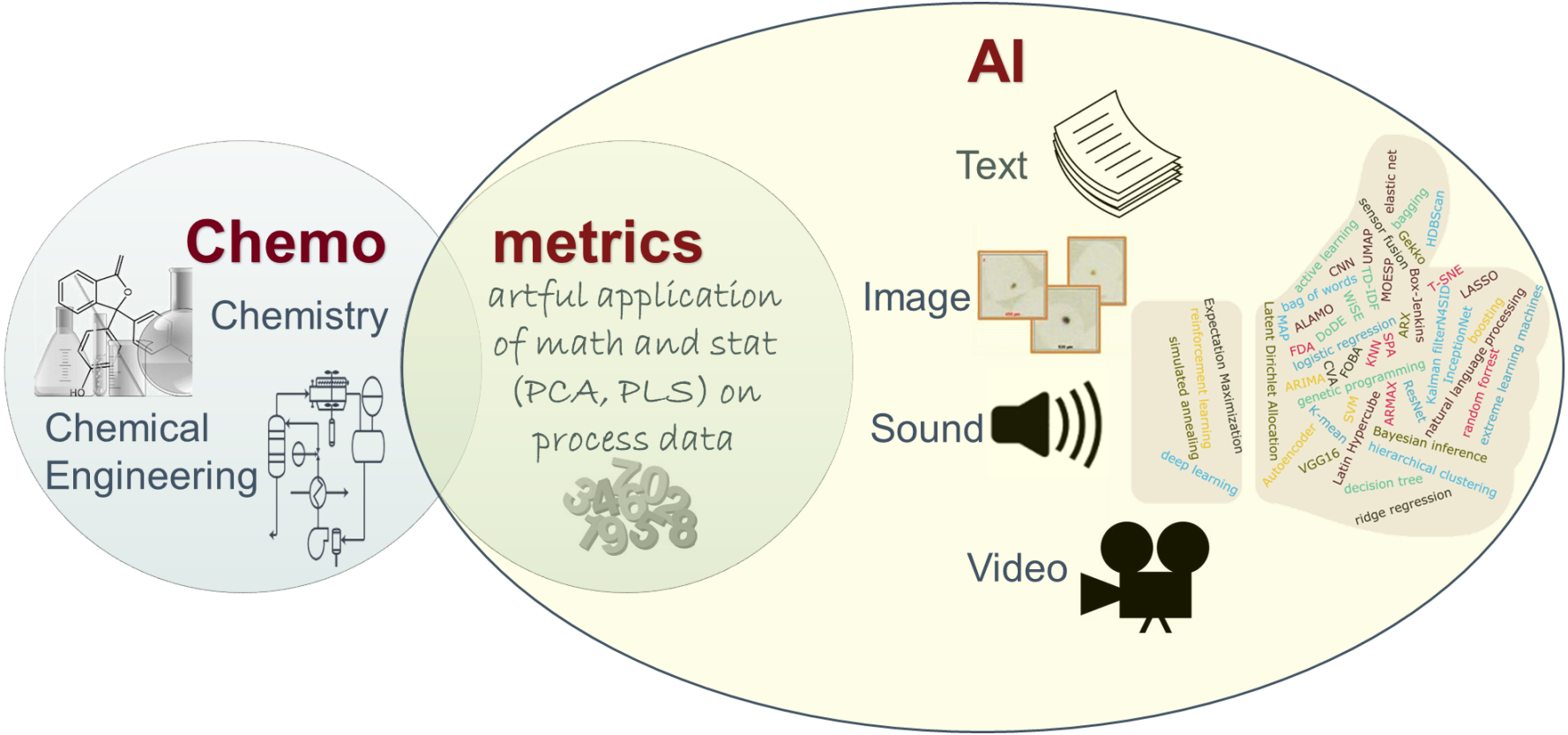


CHEMOMETRICS



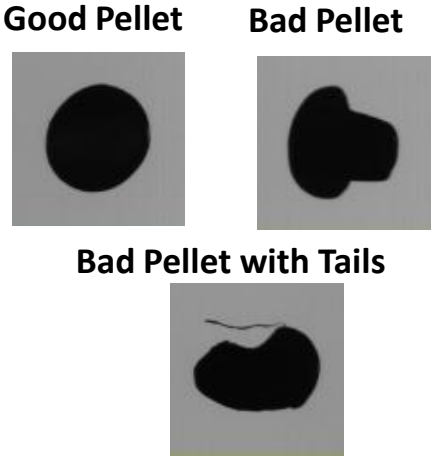
Chemometrics brings the domain knowledge into the application of AI

CHEMOMETRICS AND AI

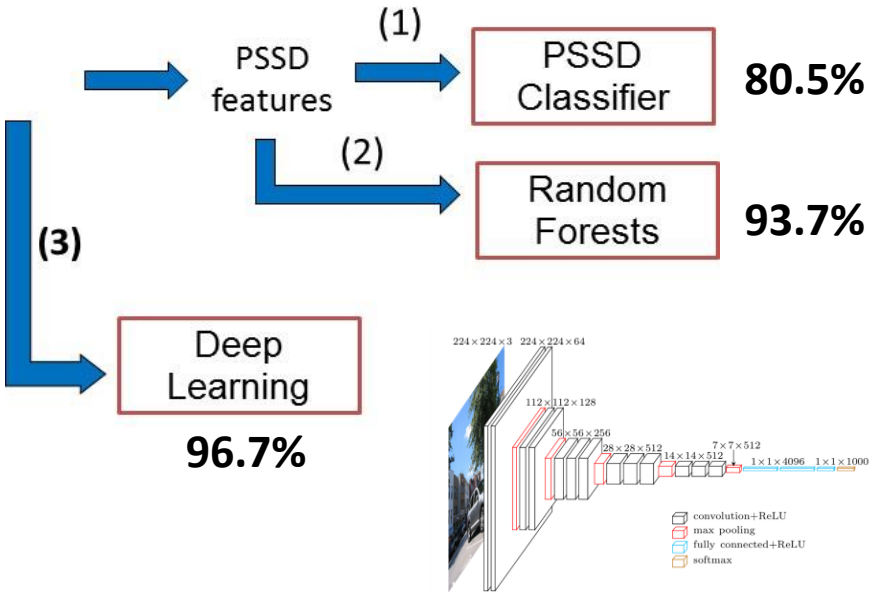


See AI success stories cited in L. Chiang, B. Braun, Z. Wang, and I. Castillo, Towards AI at scale in the chemical industry, *AIChE J*, e17644, 2022.

Success Story 1: Deep Neural Networks for Image Classification



Process experts manually labelled 6000+ images from Particle Shape and Size Distribution (PSSD)



- R. Rendall, I. Castillo, B. Lu, M. Broadway, B. Colegrove, L. Chiang, and M. Reis, Image-based manufacturing analytics: Improving the accuracy of an industrial pellet classification system using deep neural networks, *Chemo & Intel Lab Systems*, 180, 26-35, 2018.
- W. Zhu, B. Braun, L. Chiang, and J. Romagnoli, Investigation of transfer learning for image classification and impact on training sample size, *Chemo & Intel Lab Systems*, 104269, 2021.
- Y. Peng, B. Braun, C. McAlpin, M. Broadway, B. Colegrove, and L. Chiang, Contamination classification for pellet quality inspection using deep learning, *CACE*, 107836, 2022.

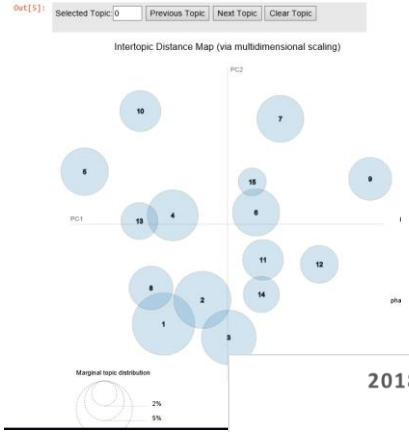
8

SUCCESS STORY 2: IMPROVE AICHE PROGRAMMING USING TEXT MINING/NLP

AIChE leadership desires to improve meeting quality by

- Reducing number of sessions
- Increasing collaborations among divisions and groups

It is ineffective to examine these ~700 sessions one at a time



Text mining is an effective way to mathematically quantify the similarities among these ~700 sessions

80 - FEW Nexus: Emerging Chemical Engineering Innovations from Micro-Scale Innovations to Complex, Interconnected Systems (Invited Talks)

Monday, October 29, 2018
8:00 AM - 10:30 AM
David L. Lawrence Convention Center - 317

Description

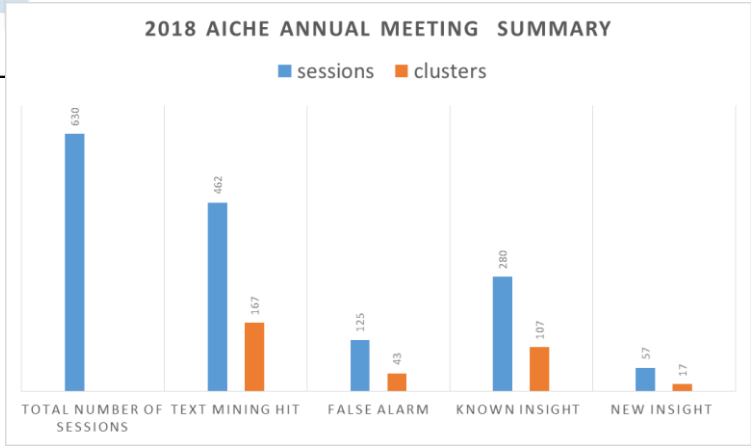
How do we think differently about complex interconnected system challenges? What are the metrics that inform and guide decisions? How is fundamental chemical engineering research at a range of temporal and spatial scales translated to impact societies? The forum also provides opportunities to share experiences from on-going Nexus projects. Please consult the online program for the most up-to-date information.

211 - Chemical Engineers and Policy-Making

Monday, October 29, 2018
3:30 PM - 6:00 PM
David L. Lawrence Convention Center - 331

Description

Chemical Engineers can take many paths including academia, industry and private research. Rarely is policy-making talked about as a viable option. It is important for scientists and engineers to play a role in creating and implementing policies. This session will cover different ways that Chemical Engineers can be a part of policy-making.



Z. Wang, B. Braun, A. Zink, M. Webb, M. Dessauer, T. Licquia, I. Castillo, and L. Chiang, 2020 AIChE spring meeting, 2021 AIChE EBPC and Programming Retreat

Success Story 3: Hybrid Modeling = AI and Data Science + Fundamentals

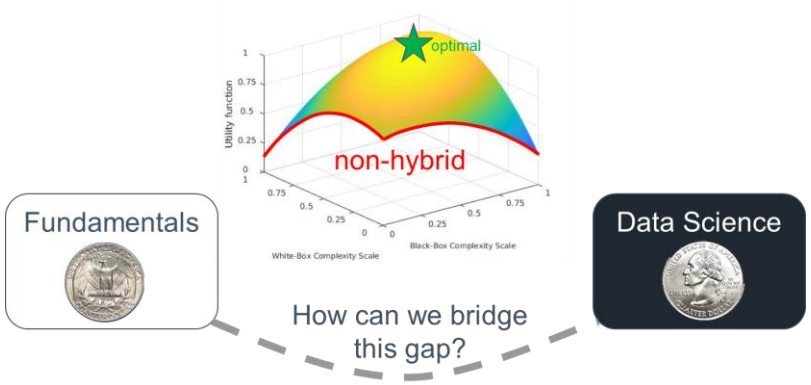
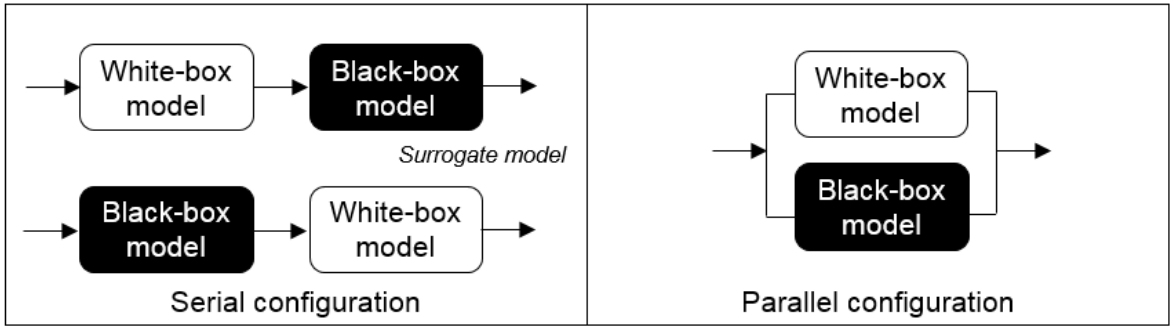


TABLE 1 Comparison of knowledge- and data-driven models

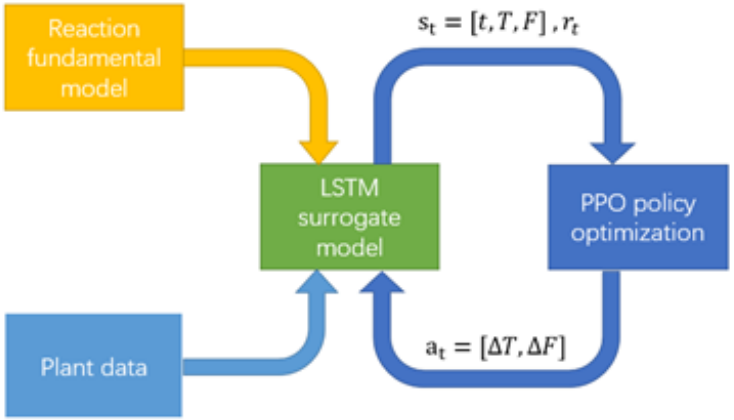
	Fundamental model	Data-driven model
Data collection	Experimental data	Experimental data, operation data, or both
Development time	Months to years	Weeks to months
Monetary cost	Expensive	Inexpensive
Extrapolation	Yes	No



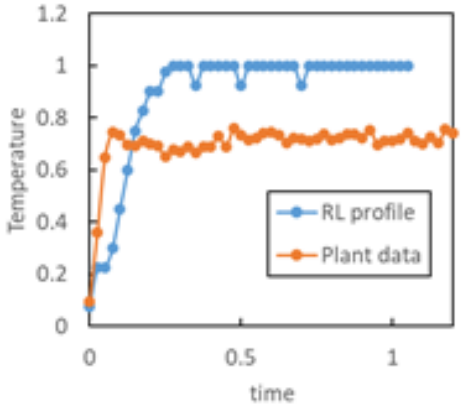
J. Sansana, M. Joswiak, I. Castillo, Z. Wang, R. Rendall, and L. Chiang, Recent trends on hybrid modeling for Industry 4.0, *CACE*, 107365, 2021.

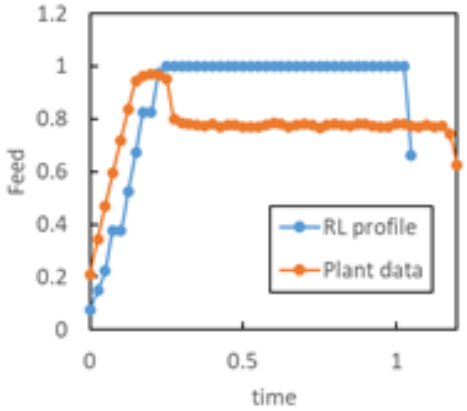
Success Story 4: Reinforcement Learning for Control and Optimization

Simulated data introduces variability



Historical plant data improves accuracy

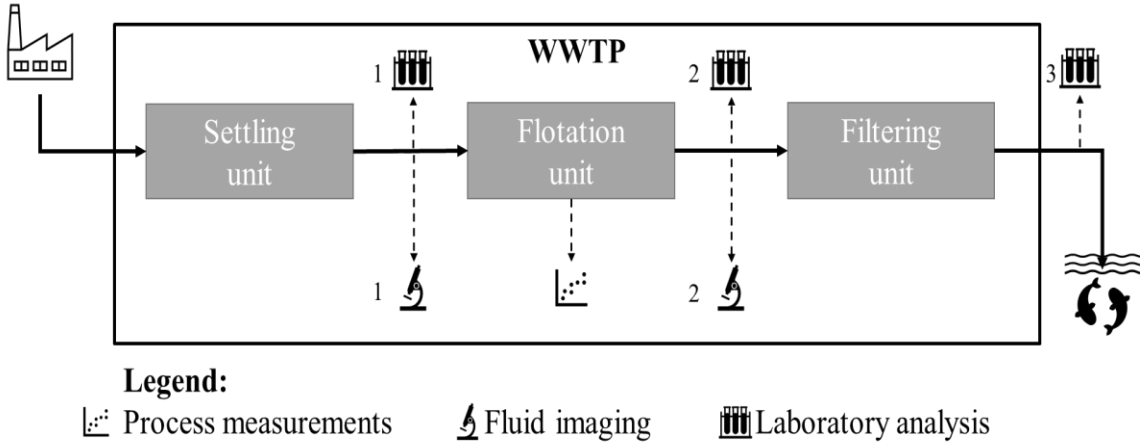


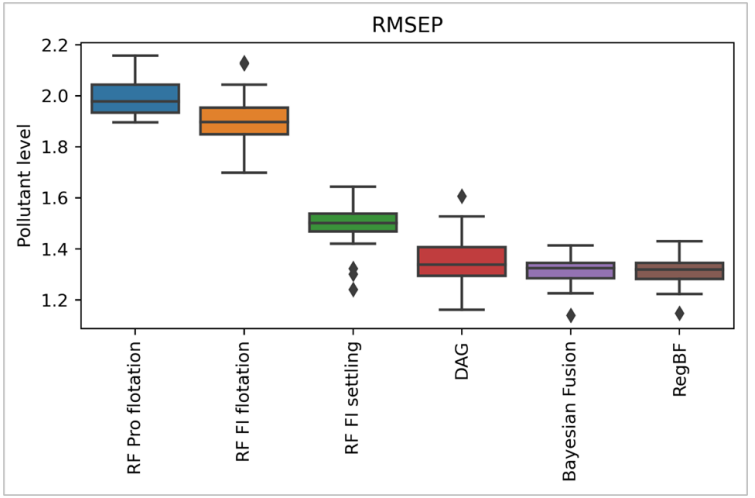


- The RL profile suggests maintaining at a higher temperature setpoint than the plant operation to achieve a higher product selectivity.
- RL suggests ramping up to a maximum throughput of the reactant feed rate in order to achieve high production rate.
- 14% shorter batch time translates to $MM margin improvement potentials

- W. Zhu, I. Castillo, Z. Wang, R. Rendall, L. Chiang, P. Hayot, and J. Romagnoli, Benchmark study of reinforcement learning in controlling and optimizing batch processes, *J. AMP*, 4(2), e10113, 2022.
- W. Zhu, R. Rendall, I. Castillo, Z. Wang, L. Chiang, P. Hayot, and J. Romagnoli, Control of A Polyol Process Using Reinforcement Learning, IFAC Papers, 54(3) 498-502, 2021.

Success Story 5: Sensor Fusion for WWTP Monitoring





Goal:

- Use all available data to accurately estimate the pollutant level

Data sources:

- Hundreds of process variables, real-time
- Images of bacteria taken daily or multiple times a week
- Grab sample, 2-3 times a week

- E. Strelet, Z. Wang, Y. Peng, I. Castillo, R. Rendall, B. Braun, M. Joswiak, L. Chiang, M. Reis, Multi-source Heterogeneous Data Fusion for Toxin Level Quantification, *IFAC papers*, 54 (3), 67-72, 2021.
- J. Sasnsana, R. Rendall, Z. Wang, L. Chiang, and M. Reis, sensor fusion with irregular sampling and varying measurement delays, I&ECR, 59(6), 2328-2340, 2020.

Success Story 6: Predictive Formulation in R&D

Polyol-1 HO~~~~OH

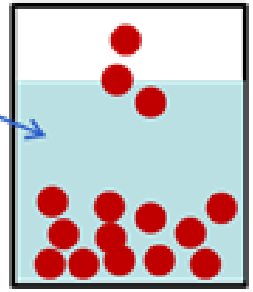
Polyol-2 OH, HO~~~~OH

Polyol-3 OH, HO~~~~OH

Isocyanates OCN—NCO

Physics-based

Polyol
+
Isocyanate
+
Surfactants
+
.....



Empirical

Performance of Foam

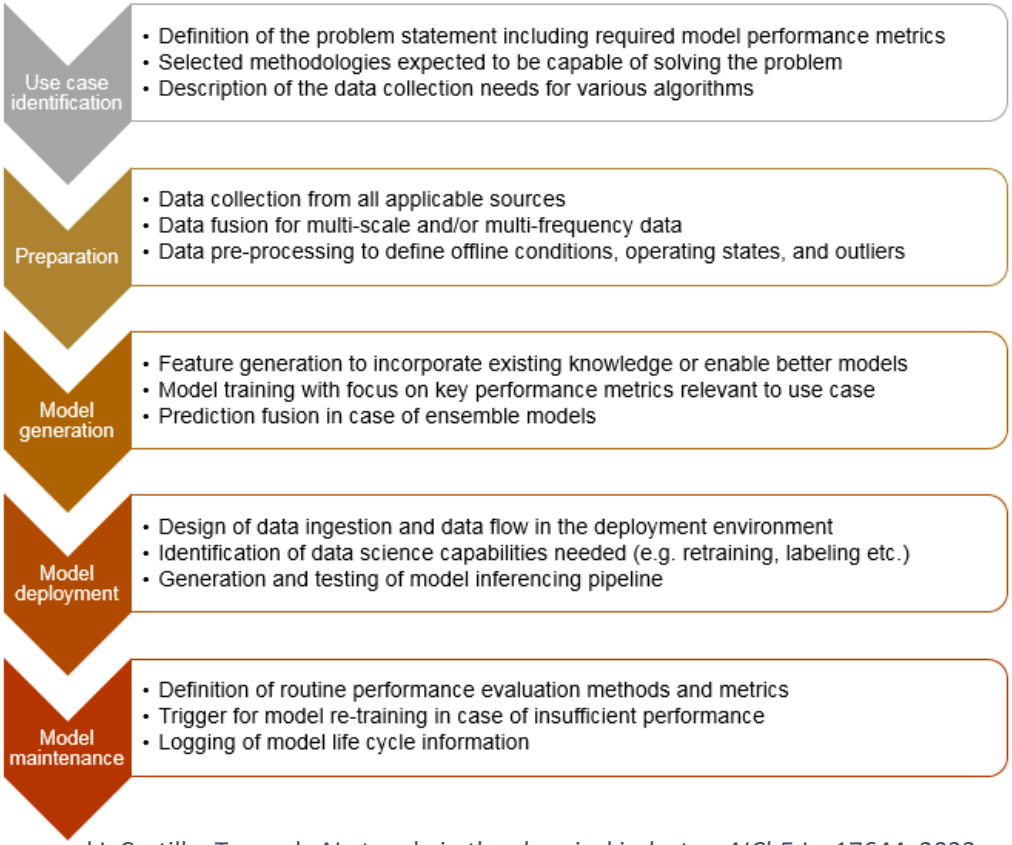
Descriptors of Formulation

AI
ARTIFICIAL INTELLIGENCE
EXCELLENCE
AWARD
2022

https://www.bintelligence.com/artificial-intelligence-excellence-awards

https://www.dow.com/en-us/product-technology/pt-polyurethanes/harnessing-the-power-of-digitalization.html

AI Model Life Cycle



AI/ML/DL/DS models

L. Chiang, B. Braun, Z. Wang, and I. Castillo, Towards AI at scale in the chemical industry, *AIChE J*, e17644, 2022.

AI CULTURE CHANGE

AI/ML modelers

- Innovation
- Advanced analytics and programming tools

500+ Data scientists

- Collaboration
- Data acumen
  (Special analytics and programming tools)

35,000+ Dow people

- Foundation
- Data literacy/acuity
  (Practitioner analytics tools)

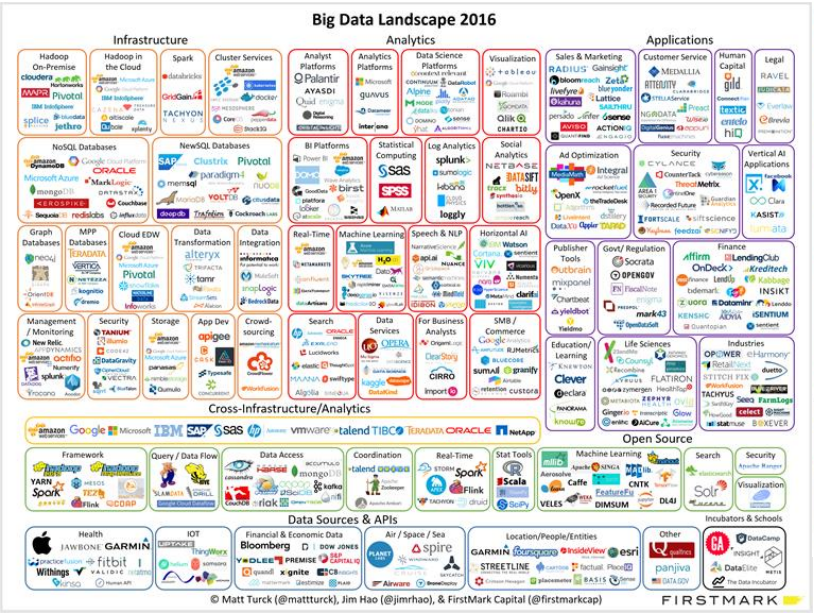
500,000+ US STEM graduates

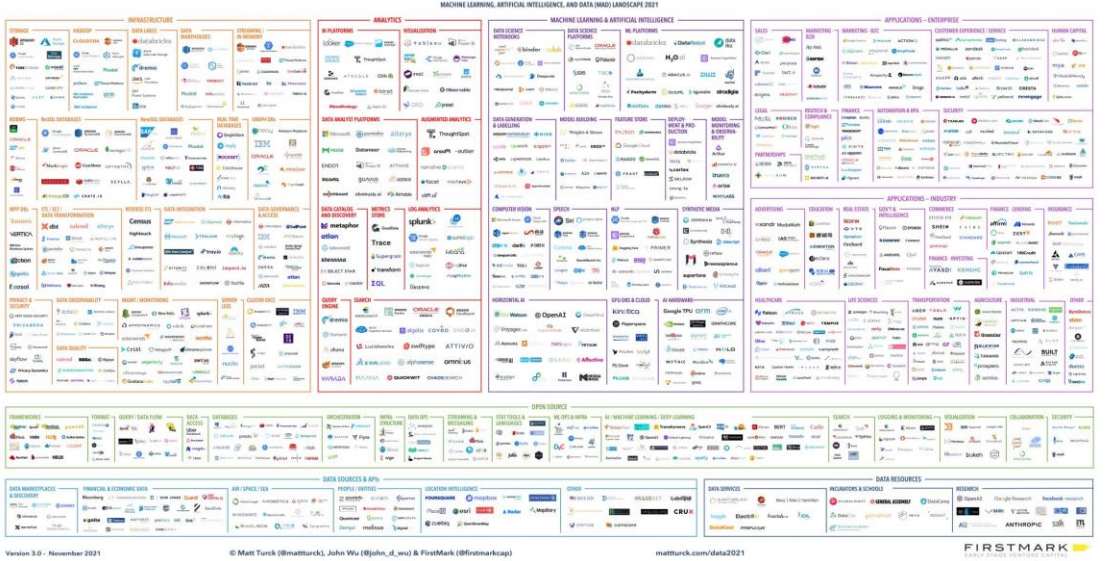
- Art of the possible
- Integrate data science into Engineering curriculum

Chemical engineering students are starting to learn "PID control" equivalent of data science

SJ. Qin and L. Chiang, Advances and opportunities in machine learning for process data analytics, *CACE*, 126:465-473, 2019.

There are a lot of AI hypes, So, how to effectively partner with AI companies?





https://mattturck.com/data2021/

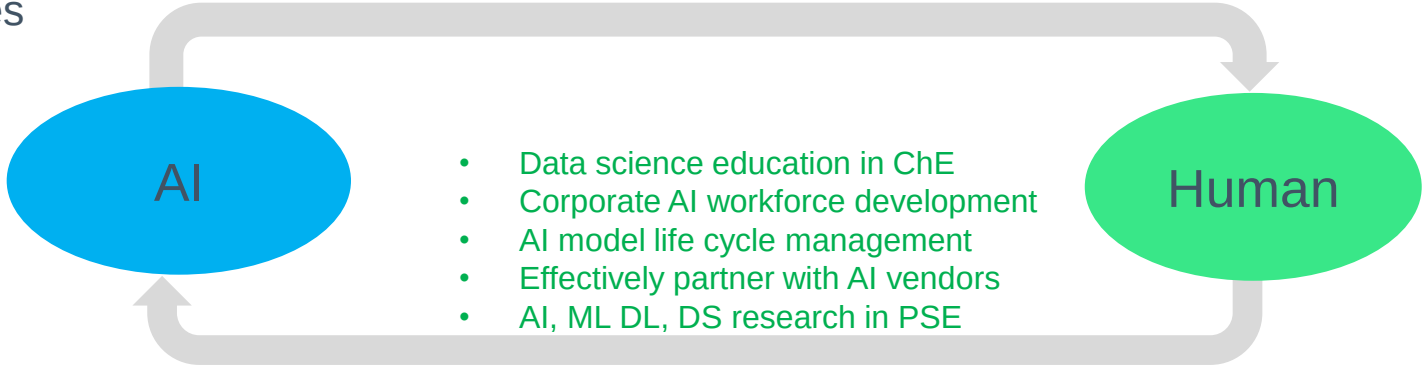
AI, ML, DL, DS CHALLENGES IN PSE

Technical challenges

- Deep learning: How to interpret deep learning results (concept of Explainable AI, XAI, in plant environments)?
- NLP: How to embed domain knowledge to open-source NLP tools?
- Hybrid modeling: How to systematically develop hybrid models for broad range of domains/scales (R&D, Manufacturing, Supply Chain, etc.)?
- Reinforcement Learning: How to gain trust to validate, implement, and sustain RL model in plant environments?
- Sensor fusion: how to expand framework to include ALL kinds of data?
- Materials discovery: *forward* model: sparse and limited data set; *inverse* model: optimization with advanced AI/ML methods

AI cultural challenges

- Faster On boarding $
- Faster and less trial and error $$
- New/hidden insights $$$



- Data science education in ChE
- Corporate AI workforce development
- AI model life cycle management
- Effectively partner with AI vendors
- AI, ML DL, DS research in PSE

What Does It Take to Scale AI in the Process Industry?

Exponential data growth with AI model lifecycle

+

AI innovation with domain knowledge

+

AI culture

+

Workforce development and upskill

= **AI at scale**

Seek

Together™