AI in Manufacturing of Pharmaceutical Products: Challenges and Opportunities

Jayanti Das
Research Scientist
Division of Product Quality and Research
Office of Testing and Research
Office of Pharmaceutical Quality
CDER | US FDA
[PQS 2023] – November 1, 2023

Major Research Areas

Analytical Science

- Chromatography (e.g., HPLC and UPLC core facility)
- **Mass spectrometry (e.g., high throughput RapidFire)**
- Nuclear magnetic resonance (NMR) spectroscopy
- **Advanced separation (e.g., field flow fractionation)**
- Product performance (e.g., dissolution, in vitro release test, IVRT)
- Bioanalytics
- Shelf-life Extension Program (SLEP)

Formulation Science

- Oral solids (e.g., tablets, capsules)
- Topicals and transdermal
- Ophthalmic
- Injectables (e.g., liposomes, lipid-nanoparticles, suspensions, emulsions, long-acting)
- Implantable (e.g., intravaginal, intrauterine, intramuscular)
- Biopharmaceutics (e.g., IVIVC, BCS, biowaivers, bioequivalence)
- **Nanotechnology**
- All other complex formulations
- Excipients functionality (e.g., polymeric materials)
- Quality-by-Design (QbD)

Adv. Manufacturing

- **Continuous manufacturing** (**drug substances**, solid oral dosage forms, complex formulations)
- 3D printing
- Process analytical technology (PAT)
- Biomanufacturing (e.g., upstream/downstream processing, lyophilization)

Modeling & Simulation

- Digital twins
- In vitro in vivo correlation (IVIVC)
- Modeling, e.g., CFD, MD, DEM, RTD
- System/Process design (e.g., LabVIEW)
- **Data science**, e.g., AI/ML, chemometrics

Challenges

- Historical Data
- Operational Data
- Process Data
- Maintenance Data

Causes

Infinite Data Generation

Challenges

- Significant time spent on preparing vs analyzing data
- Limited data are being used for reporting, analysis, and decision making
- Very few data are being analyzed and acted upon real-time

Reasons

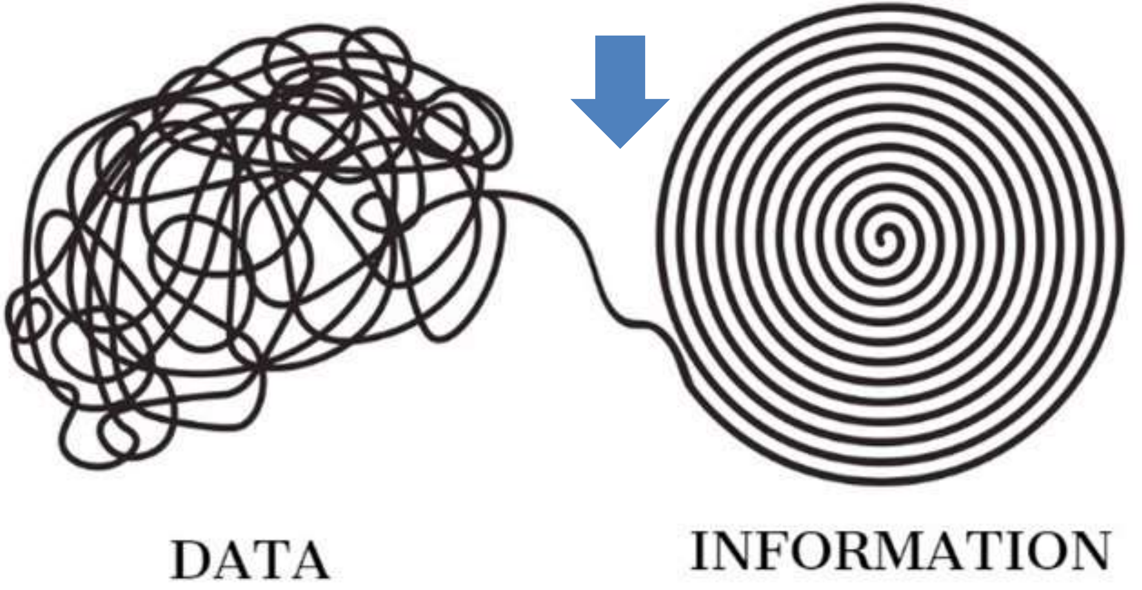
- Lack of connectivity
- Lack of context
- Lack of transparency
- Need for manual data preparation
- Fragmented and inconsistent data
- Lack of knowledge
- Limited automation

Solutions

Digital Transformation | Data Integration | Data Contextualization

Opportunities

AI/ML Application



Capture, Process, Organize Information; Perform Analysis; Uncover Actionable Insights

Use Case I: Study formulation design using ML applications

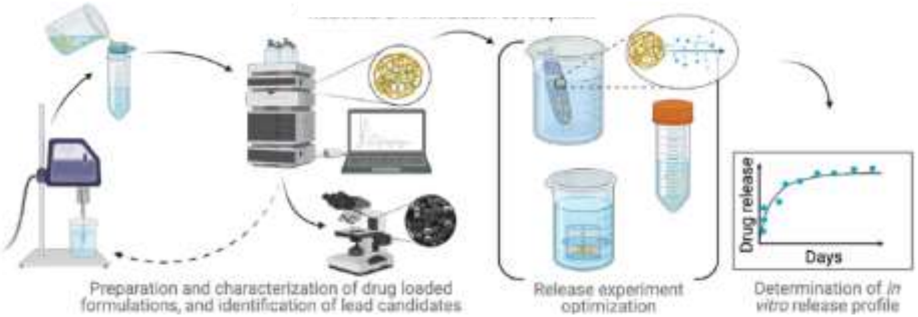
Careful design of formulations:

- Enhance efficacy of a new drug molecule
- Reduce adverse effect
- Improve bioavailability
- Reduce off-target delivery

Reference: Pharmaceutical excipients — quality, regulatory and biopharmaceutical considerations

Challenges in Formulation Design

Each drug has its own unique physicochemical properties



For a given material, there are wide range of variables that must be optimized during formulation preparation

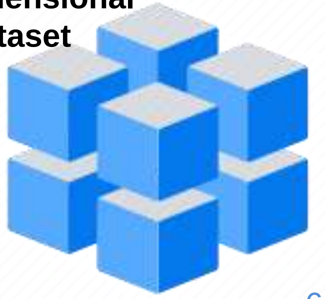
Trial-and-error in the experiment design

Extensive resources

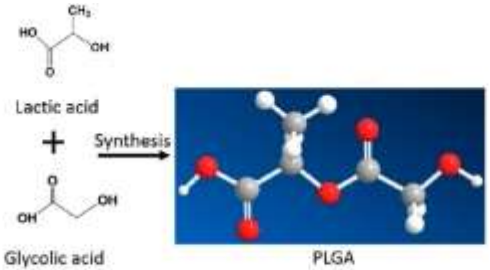
Time

Cost

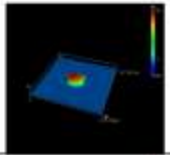
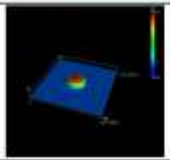
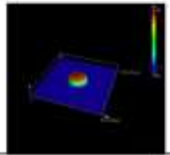
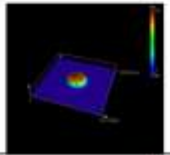
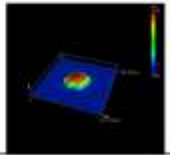
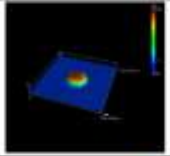
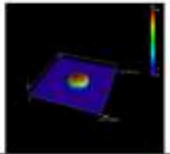
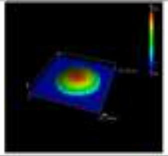
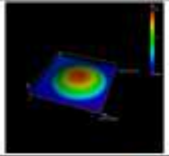
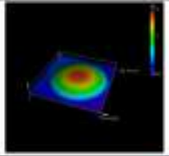
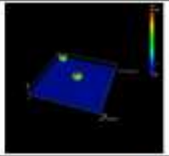
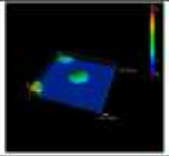
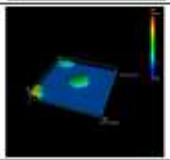
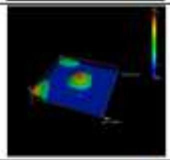
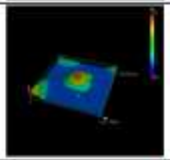
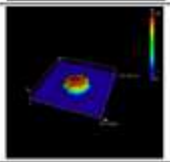
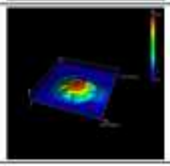
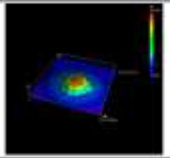
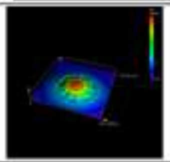
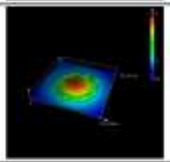
Multidimensional large dataset



PLGA-based Formulation Dataset

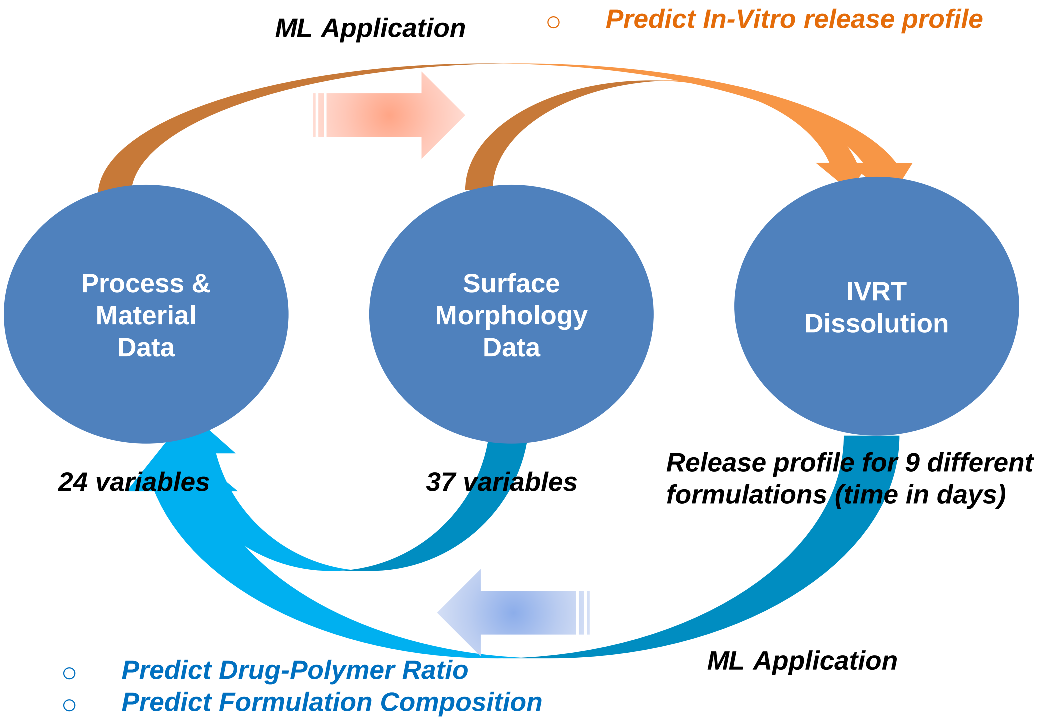


- Biodegradable poly(lactide-co-glycolide) (PLGA) microparticles have been used as long-acting injectable (LAI) drug delivery systems from past three decades.
- Used for **prolonged therapeutic effect** and become the ideal formulation strategies for treatment of **chronic disease**.

Sample	Dry	Ethyl Isobutyrate	Toluene	2-pentanone	Propyl acetate
VSS528A (50L Blank)					
VSS520 (75L Blank)					
VSS611 (100L Blank)					
VSS526B (50L+100L Polylithic)					

*** The dataset for this study was derived from a previous FDA-funded research project at Akina, Inc. (BAA#75F40119C10096).*

Analysis



Significant Parameters Selection

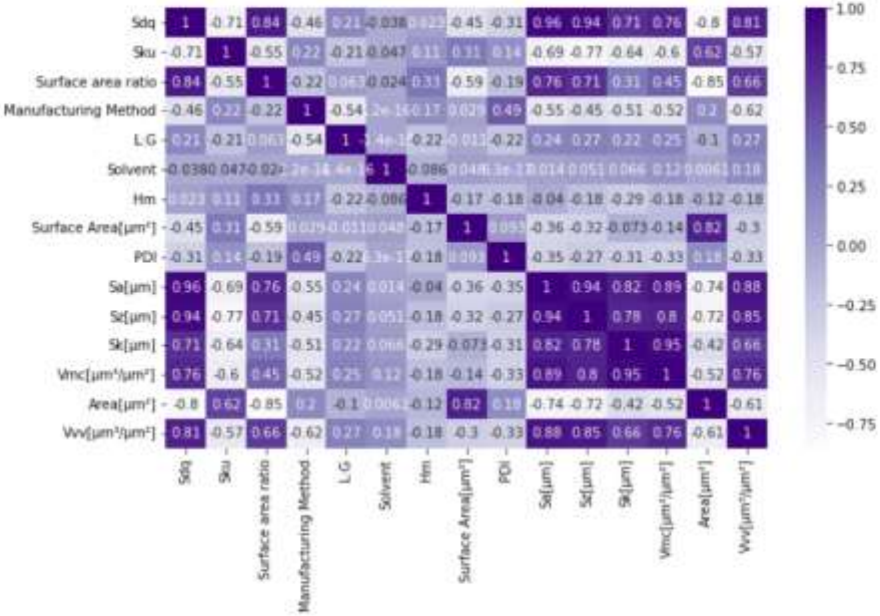
Predict Drug -Polymer ratio

Predict Formulation Composition

Predict In-vitro release behavior

Overall Predictive Performance

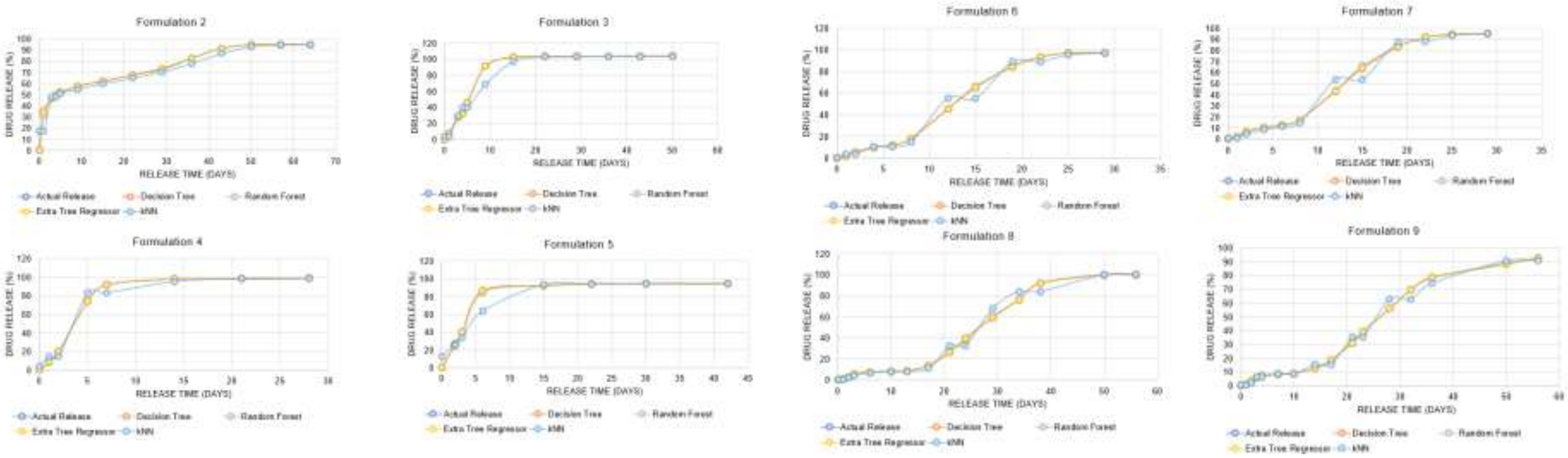
Correlation Matrix



Prediction of Formulation Composition			
Machine Learning Techniques	MSE	MAE	Accuracy (%)
Linear Regression	0.001	0.02	99.5002
Decision Tree	0	0	100
Random Forest	0.3439	0.4	90
Extra Trees Regressor	0.0042	0.0289	99.2778

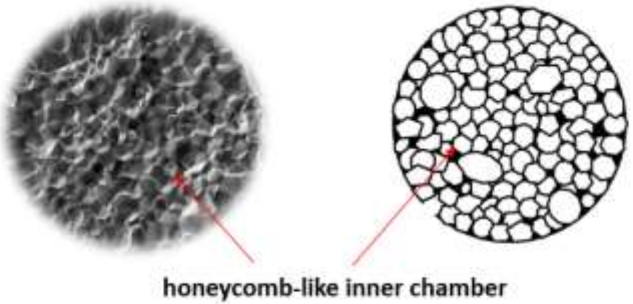
Prediction of Drug – Polymer Ratio			
Machine Learning Technique(s)	MSE	MAE	Accuracy (%)
Logistic Regression	0	0	100
Decision Tree	0	0	100
Random Forest Regressor	0	0	100
Artificial Neural Network (ANN)	0.0021	0.0378	78

In-Vitro Release Prediction



- ❑ This AI method may serve as a tool in the future to help comparing the proposed generic products to reference listed drugs (RLD) by analyzing feature similarity across different formulations.
- ❑ Such a tool may also help addressing some of the unique challenges in determining the bioequivalence of long-acting injectable generic products.

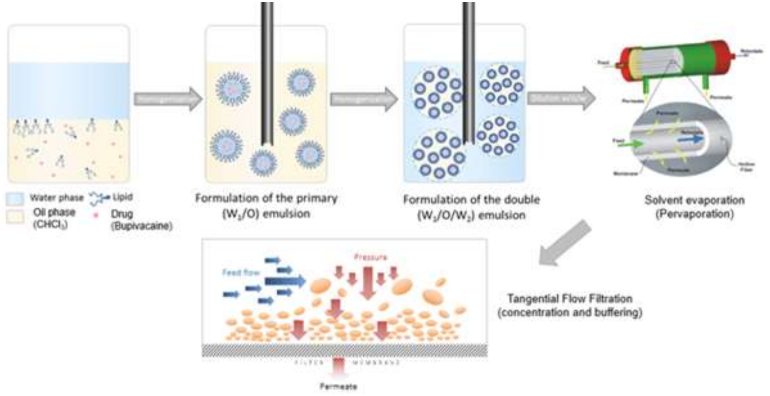
Use Case II: Advanced Imaging Analysis to Improve understanding of Multivesicular Liposomes

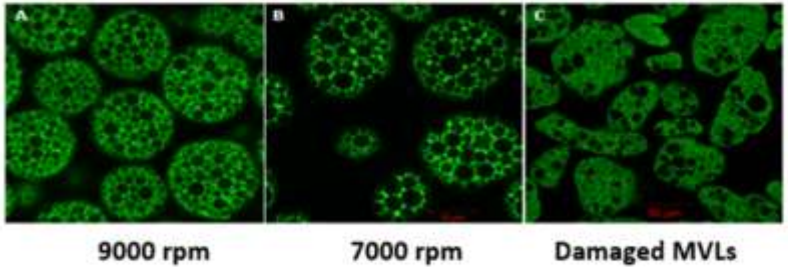


- Multivesicular liposome (MVL) is a lipid-based drug delivery system for sustained release of the drugs with short half-lives.
- Multivesicular Lipid Liposomes (MVLs) are complex and oftentimes sensitive to the release environment.
- Design and development of appropriate in vitro release test (IVRT) method is challenging for MVLs.

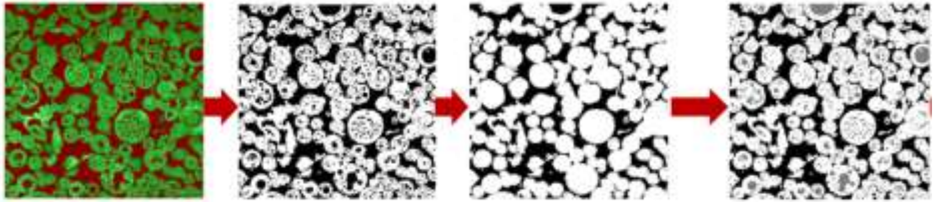
AI-assisted image analysis

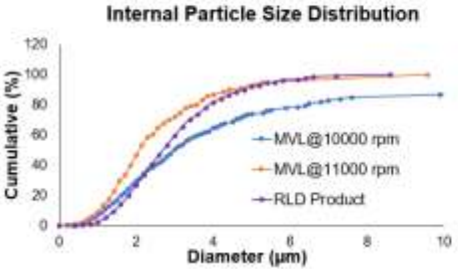
The study aims to **develop AI assisted image analysis** method to provide **quantitative assessment** of the MVL morphology changes due to process parameter changes.

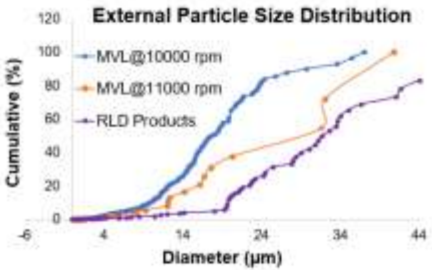




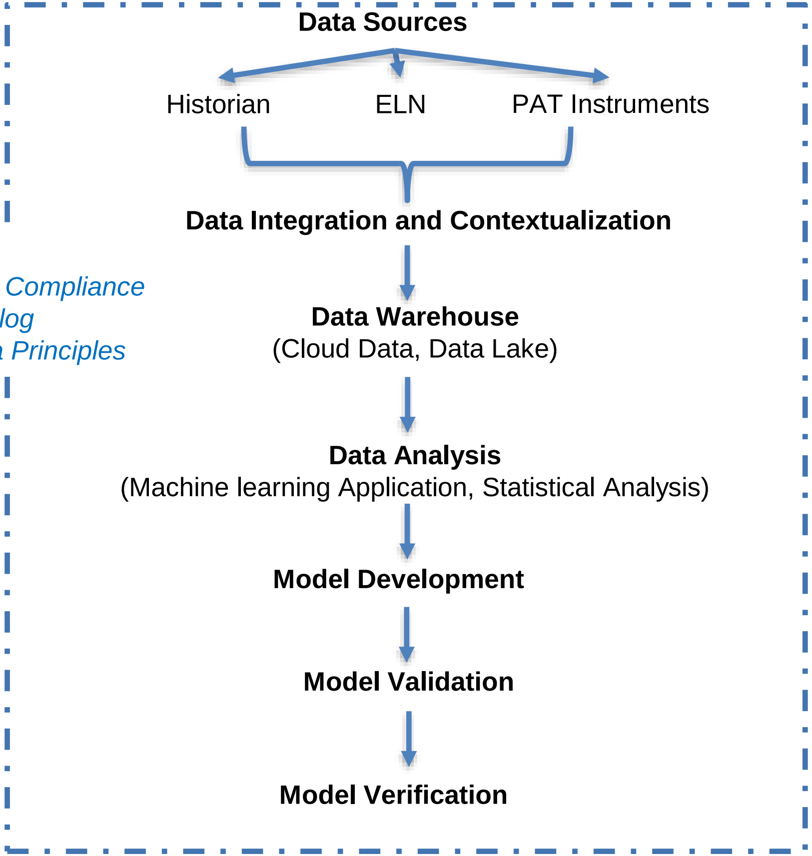
Segmentation & Templating	Quantitation	Output
▪ Identify Particles, Pores, Background ▪ Create Template for different segmentation	▪ Count and Measure	▪ External & Internal Particle Sizes ▪ Spatial Distribution







Data Architecture for Advanced Manufacturing



Required:

- *Regulatory Compliance*
- **Data Catalog*
- **FAIR Data Principles*

**FAIR (findability, accessibility, interoperability, reusability)*

**Data Catalog: detailed inventory of all data assets in an organization*

Closing Thought

- In advanced manufacturing, **data silos occur across different stages** of production process, including design, production, quality control, and supply chain management.
- The lack of **seamless data exchange and collaboration** among these domains can lead to inefficiencies, redundancies, and missed opportunities.
- To fully leverage the potential of **digital transformation** in advanced manufacturing, breaking down data silos is essential.
- The design process and its **outcome need to be transparent** for equipment manufacturer, pharma companies and regulatory agencies to achieve trust in every process step.

Acknowledgement

OTR

- Thomas O'Connor
- Xiaoming Xu
- Geng (Michael) Tian
- William Smith
- Md Easin Hasan

OGD and OPMA

- Yan Wang
- Bin Qin
- Wei Yang
- Abdollah Koolivand
- Jungeun Bae

Purdue University

- Kinam Park
- John Garner

DigiM Solution

- Cheney Zhang
- Shawn Zhang

Thank You

Questions?

Jayanti Das
Research Scientist
Division of Product Quality and Research
Office of Testing and Research
Office of Pharmaceutical Quality
CDER | US FDA