Demystifying AI - Workshop

Timothy Hsu
Chief Collaboration Officer
AFDO/RAPS Healthcare Products Collaborative

Carrie Rigdon
Director of Research, AFDO

Steve Mandernach
Executive Director, AFDO

Natalie Dyenson
VP Food Safety & Quality, Dole
Questions for the Audience

Go to www.menti.com

Enter the code 2256 8107
“AI can be thought of as simulating the capacity for abstract, creative, deductive thought - and particularly the ability to learn - using the digital, binary logic of computers.”

“Artificial Intelligence (AI) is no longer some bleeding technology that is hyped by its proponents and mistrusted by the mainstream. In the 21st century, AI is not necessarily amazing. Rather, it is often routine. Evidence for the routine and dependable nature of AI technology is everywhere.”

(Source: Tim Menzies, 2004)
Data Culture

• Data culture is the collective beliefs and behaviors of the people in the organization for leveraging data for improved business performance. – Forbes

• Data culture is the collective behaviors and beliefs of people who value, practice, and encourage the use of data to improve decision-making. – Tableau

• Gartner’s third CDO Survey lists data cultures as the number one challenge for realizing benefits of data and analytics
Data culture is the collective beliefs and behaviors of the people in the organization for leveraging data for improved business performance. – Forbes

Data culture is the collective behaviors and beliefs of people who value, practice, and encourage the use of data to improve decision-making. – Tableau
Data culture is the behaviors of the people using data for improved decision making. – Schmucki

"You miss 100% of the shots you don't take. - Wayne Gretzky"

- Michael Scott
Hurdles to Data Culture

• Executives lack a clear vision for advanced analytics
• Goals are driven by tools and not business problems
• Analytic capabilities are isolated from the business
• Ignoring Lean Principles

Standardize  →  Simplify  →  Digitize
McKinsey’s Model

Defining roles is an important first step in sourcing and integrating the right talent for your data culture.

1. **Business leaders** lead analytics transformation across organization
2. **Delivery managers** deliver data-and analytics-driven insights and interface with end users
3. **Workflow integrators** build interactive decision-support tools and implement solutions
4. **Visualization analysts** visualize data and build reports and dashboards
5. **Data engineers** collect, structure, and analyze data
6. **Data architects** ensure quality and consistency of present and future data flows
7. **Analytics translators** ensure analytics solve critical business problems
8. **Data scientists** develop statistical models and algorithms
Steps to Data Culture

1. Create a service culture as it relates to data
2. Move away from Consensus and Hierarchy Cultures
3. Leverage technology
4. Invest in both people and technology
5. Fix basic data-access issues quickly
6. Question analytical choices
Mechanistic vs. Probabilistic

- Physics and Engineering mechanics provides the right conditions for the ideal or known scenario
- The new probability (AI models) defines the real conditions without the physical and chemical basis
AI vs. Statistics

More than just multivariable models...

X Mean: 54.2659224
Y Mean: 47.8313999
X SD: 16.7649829
Y SD: 26.9342120
Corr.: -0.0642526
Understanding AI: Data (the Secret Sauce)

Data must be prepared before to use it: DS invest the 80% of their time on it

7 steps to consider when preparing data

1. Articulate problem
2. Define data needed
3. Qualify data and min prediction accuracy
4. Source missing data
5. Format data for consistency
6. Reduce, decompose, clean
7. Rescale data
8. Train AI

Findable, Accessible, Interoperable, Reusable principle

"We’ve had to spend most of the time just cleaning the data sets before you can even run the algorithm"

Vas Narasimhan, CEO of Novartis AG, in a 2018
AI starts with data. In pharma, with quality data.
AI starts with data. In pharma, with quality data

What is FAIR DATA?

- **FINDABLE**: Data and supplementary materials have sufficiently rich metadata and a unique and persistent identifier.
- **ACCESSIBLE**: Metadata and data are understandable to humans and machines. Data is deposited in a trusted repository.
- **INTEROPERABLE**: Metadata use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- **REUSABLE**: Data and collections have a clear usage licenses and provide accurate information on provenance.

ALCOA: FDA’s Data Integrity Focus

- **C**ontemporaneous
- **L**egible
- **O**riginal
- **A**ttibutable
- **A**ccurate

Data Integrity
"No Free Lunch"

Linear and Polynomial Regression

Gradient Search

<table>
<thead>
<tr>
<th>iteration</th>
<th>m</th>
<th>b</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Points and Line

\[ y = 0.00x + 0.00 \]

Regression Trees and Random Forests
AI is already here because data is available

**Machine learning predicts the look of stem cells**

Ann Mikkelsen

*Nature* (2017) | Cite this article

Published: 19 April 2017

No two stem cells are identical, even if they are genetic clones. This stunning diversity is revealed today in an enormous publicly available online catalogue of 3D stem cell images. The visuals were produced using deep learning analyses and cell lines altered with the gene-editing tool CRISPR. And soon the portal will allow researchers to predict variations in cell layouts that may foreshadow cancer and other diseases.
Bioreactor Fermentation Process

→ Optimise the hypoxic conditions for *Pichia Pastoris* yeast to maximize production

→ Study the effect of specific growth rates (DoE) on yield & protein stability

<table>
<thead>
<tr>
<th>Phase</th>
<th>Normoxic</th>
<th>Hypoxic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase I</td>
<td>FBHPX2, FBHPX5, FBHPX8</td>
<td>FBHPX3, FBHPX4, FBHPX10</td>
</tr>
<tr>
<td>Phase II</td>
<td>FBHPX2, FBHPX5, FBHPX8</td>
<td>FBHPX3, FBHPX4, FBHPX7</td>
</tr>
<tr>
<td>Phase III</td>
<td>FBHPX2, FBHPX5, FBHPX8</td>
<td>FBHPX7</td>
</tr>
<tr>
<td>Phase IV</td>
<td>FBHPX5, FBHPX6, FBHPX8</td>
<td>FBHPX3, FBHPX4</td>
</tr>
</tbody>
</table>
Reducing dimensionality by means of AI

2D UMAP projection

2D t-SNE projection

Cancer cells grow and divide at an abnormally rapid rate, are poorly differentiated, and have abnormal membranes, cytoskeletal proteins, and morphology. The abnormality in cells can be progressive with a slow transition from normal cells to benign tumors to malignant tumors.

- Self-sufficiency in growth signals
- Insensitivity to anti-growth signals
- Evading programmed cell death
- Limitless replicative potential
- Sustained angiogenesis
- Tissue invasion and metastasis

Automatizing human cognition around C&GT
Automatizing human cognition around C&GT

<table>
<thead>
<tr>
<th>Normal</th>
<th>Cancer</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Large, variably shaped nuclei</td>
</tr>
<tr>
<td></td>
<td>Many dividing cells;</td>
</tr>
<tr>
<td></td>
<td>Disorganized arrangement</td>
</tr>
<tr>
<td></td>
<td>Variation in size and shape</td>
</tr>
<tr>
<td></td>
<td>Loss of normal features</td>
</tr>
</tbody>
</table>

Image Analysis

Localization
- Whole
- Core
- Rim
- Rim & Core

Features of Cell Morphology
- Proliferation, Concavity, Aspect Ratio
- Roughness, Area Variation

Features of Cytoskeletal Organization
- Total Pixel Intensity, Density, Orientation, Parallelness

Training
Feature Weights
Validation
Classified and Characterized Cells

Non-Cancer Cells
Cancer Cells
Reducing dimensionality and clustering features
Goal: Create a **digital twin** to manage a **biotech process** working under **GxP conditions**

Q: What does it mean, a digital twin?
Q: What kind of biotech process are we controlling?
Q: How to apply AI within regulatory frameworks?

PDA Interest Group "Process Validation" (EU)

**Taskforce 1** | Synthetic data generation to support AI model training

**Taskforce 2** | Automated biotech process control

**Taskforce 3** | Regulatory considerations (QbD, Data governance, AI Procedures)
## AI Maturity Model

### Maturation Levels

<table>
<thead>
<tr>
<th>Maturation Level</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 1</td>
<td>Established technology infrastructure and support for AI development.</td>
</tr>
<tr>
<td>Level 2</td>
<td>AI infrastructure and support in place, including data management and governance.</td>
</tr>
<tr>
<td>Level 3</td>
<td>AI technology and processes are integrated into business operations.</td>
</tr>
<tr>
<td>Level 4</td>
<td>AI technology and processes are fully integrated and optimized for business outcomes.</td>
</tr>
</tbody>
</table>

### Key Features

- **AI Hockey Stick**: Rapid growth in AI adoption and value creation.
- **AI Maturity Curve**: Progressive improvement in AI capabilities.
- **AI Roadmap**: Plan for AI adoption and growth.

### AI Maturation Model

<table>
<thead>
<tr>
<th>Maturation Area</th>
<th>Maturation Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leadership</td>
<td>Level 2</td>
</tr>
<tr>
<td>Strategy</td>
<td>Level 3</td>
</tr>
<tr>
<td>People</td>
<td>Level 2</td>
</tr>
<tr>
<td>Process</td>
<td>Level 3</td>
</tr>
<tr>
<td>Technology</td>
<td>Level 4</td>
</tr>
<tr>
<td>Governance</td>
<td>Level 3</td>
</tr>
<tr>
<td>Operation</td>
<td>Level 4</td>
</tr>
</tbody>
</table>

### Tools and Techniques

- **Data Science**: Predictive analytics, machine learning, natural language processing.
- **Software Engineering**: Agile development, continuous integration.
- **Business Intelligence**: Dashboards, reports, KPIs.

### AI Maturity Levels

<table>
<thead>
<tr>
<th>Maturity Area</th>
<th>Maturity Level</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Management</td>
<td>Level 1</td>
<td>Data collected in silos, no integration.</td>
</tr>
<tr>
<td>Level 2</td>
<td>Data integrated across the organization.</td>
<td></td>
</tr>
<tr>
<td>Level 3</td>
<td>Data driven decision making.</td>
<td></td>
</tr>
<tr>
<td>Level 4</td>
<td>Data智能 integration and automation.</td>
<td></td>
</tr>
</tbody>
</table>

### Key Metrics

- **AI ROI**: Return on investment in AI technologies.
- **AI Impact**: Measuring the effectiveness of AI applications.
- **AI Efficiency**: Improving the efficiency of AI processes.

### AI Maturity Model in Practice

- **Case Study**: XYZ Company's journey from Level 1 to Level 4.
- **Success Factors**: Leadership commitment, data governance, AI culture.

### Further Resources

- **AI Maturity Model Templates**
- **AI Maturity Scorecards**
- **AI Maturity Benchmarks**
## Current Team Members

<table>
<thead>
<tr>
<th>Name</th>
<th>Affiliation</th>
<th>Position/Expertise</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antonio Moreira (Execution co-lead)</td>
<td>University of Maryland, Baltimore County</td>
<td>Vice Provost for Academic Affairs</td>
</tr>
<tr>
<td>Ben Stevens (PV Team lead)</td>
<td>GSK</td>
<td>Director CMC Policy and Advocacy</td>
</tr>
<tr>
<td>Catarina Leitão</td>
<td>4Tune Engineering</td>
<td>CPV Expert</td>
</tr>
<tr>
<td>Christophe Agut</td>
<td>Sanofi Pasteur</td>
<td>Head of Process Validation and Statistics Expertise</td>
</tr>
<tr>
<td>David Hubmayr (Task Force 1 lead)</td>
<td>CSL Behring</td>
<td>Manager, Process Development &amp; Breakthrough Technologies, R&amp;D</td>
</tr>
<tr>
<td>David Lapeña</td>
<td>Infors</td>
<td>Area Sales Manager Southern Europe &amp; Africa</td>
</tr>
<tr>
<td>Francisco Valero (Task Force 2 lead)</td>
<td>Universidad Autònoma of Barcelona</td>
<td>Professor and head of department of BioChemical Engineering</td>
</tr>
<tr>
<td>Joeri Van Wijngaarden</td>
<td>Aizon</td>
<td>Innovation Lab R&amp;D Project Manager</td>
</tr>
<tr>
<td>Mario Stassen (Task Force 3 lead)</td>
<td>AFDO (AI in Operations Team)</td>
<td>BioPharma Regulatory expert</td>
</tr>
<tr>
<td>Matt Schmucki</td>
<td>AstraZeneca &amp; AFDO</td>
<td>Lean Coach and CPV Expert</td>
</tr>
<tr>
<td>Mauro Giusti (PV Team lead)</td>
<td>Eli Lilly</td>
<td>Director, Technical Services/Mfg Sciences</td>
</tr>
<tr>
<td>Nilanjan Banerjee</td>
<td>University of Maryland, Baltimore County</td>
<td>Professor, Computer Science and Electrical Engineering</td>
</tr>
<tr>
<td>Sandrine Dessoy</td>
<td>GSK</td>
<td>Science and Technology Innovation Director</td>
</tr>
<tr>
<td>Shereya Maiti</td>
<td>Bayer Pharmaceuticals</td>
<td>Senior Scientist</td>
</tr>
<tr>
<td>Toni Manzano (Execution lead)</td>
<td>Aizon &amp; AFDO</td>
<td>CSO and Co-founder</td>
</tr>
<tr>
<td>Ciro Cottini</td>
<td>Chiesi</td>
<td>Digital, Data &amp; Modelling Head</td>
</tr>
<tr>
<td>Holger Mueller</td>
<td>Bluesens</td>
<td>Director</td>
</tr>
</tbody>
</table>

→ Optimise the hypoxic conditions for *Pichia Pastoris* yeast to maximize production
→ Study the effect of specific growth rates (DoE) on yield & protein stability

<table>
<thead>
<tr>
<th>Phase</th>
<th>Normoxic</th>
<th>Hypoxic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase I (Batch)</td>
<td>FBHPX2, FBHPX5, FBHPX6, FBHPX9</td>
<td>FBHPX8</td>
</tr>
<tr>
<td>Phase II (Adaptation)</td>
<td>FBHPX2, FBHPX5, FBHPX6, FBHPX8</td>
<td>FBHPX9</td>
</tr>
<tr>
<td>Phase III (Early Fed Batch)</td>
<td>FBHPX2, FBHPX5, FBHPX6, FBHPX8, FBHPX9</td>
<td>FBHPX7</td>
</tr>
<tr>
<td>Phase IV (Later Fed Batch)</td>
<td>FBHPX5, FBHPX6, FBHPX8, FBHPX9</td>
<td>FBHPX2</td>
</tr>
</tbody>
</table>
## AI-guided Process Monitoring

### Phase 1: Batch

→ Controlled process, no manual actions
→ Anomalies due to equipment, material qualities or improper behaviour of biomass
→ Multiple relevant factors can be monitored to detect underperforming batches

### Direct/Indirect Measurements (Primary Variables)

- Biomass concentration (g/L)
- CRL activity (UVL)
- Glucose HPLC concentration (g/L)
- Ethanol HPLC concentration (g/L)
- Microspheres flow rate (L/min)
- O₂ analyzer (mole fraction)
- CO₂ analyzer (mole fraction)
- Air flow (L/min)
- Pure O₂ flow (L/min)
- Absolute Humidity (%)
- Ethanol probe concentration (g/L)
- Temperature (°C)
- pH
- pO₂ (%)  
- Agitation (rpm)

### Calculated Parameters (Secondary Variables)

- μ
- q
- \( \phi_{\text{mes}} \)
- q
- QIR
- CER
- RQ

### Q: how can we bring value from AI in an automated process without interacting with the unit?
AI-guided Process Monitoring

Phase 1: Batch
→ Controlled process, no manual actions
→ Anomalies due to equipment, material qualities or improper behaviour of biomass
→ Multiple relevant factors can be monitored to detect underperforming batches
AI-guided Process Control

**Phase 2: Fed-Batch**

→ Final phase, hypoxic conditions.
→ System requires constant manual control to keep the metabolic parameters within the desired operating range by controlling the Agitation speed.

Control Automation For Optimal Data Governance

Control Strategy
- PAT & IIoT Technology
- Combination of edge & cloud
- Fully automated data pipeline
- Coverage of full AI lifecycle (train, productivise, monitor)
- Operate in near real-time

Control Parameters
- Storage of 17 raw data variables
- Critical read-out: respiratory quotient (RQ)
- Critical control: agitator speed (AS)
Improving manufacturing operations in real-time

Alarms and real-time performance monitoring

Real-time AI Predictions providing actionable Insights to operators in the manufacturing plant shop floor

Operator feedback on AI predictions and AI model re-training

Real-time AI providing actionable Recommendations to operators in the manufacturing plant shop floor
Follow-Up Question – Mentimeter Slide 6

Go to

www.menti.com

Enter the code

2256 8107
Carrie Rigdon
Director of Research, AFDO
crigdon@afdo.org

Steve Mandernach
Executive Director, AFDO
smandernach@afdo.org

Natalie Dyenson
VP Food Safety & Quality, Dole
natalie.dyenson@dole.com

Timothy Hsu
Chief Collaboration Officer, AFDO/RAPS
thsu@healthcareproducts.org

Carrie Rigdon
Director of Research, AFDO
crigdon@afdo.org