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

What is AI?

AI

AI = 3 ingredients + secret sauce

"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

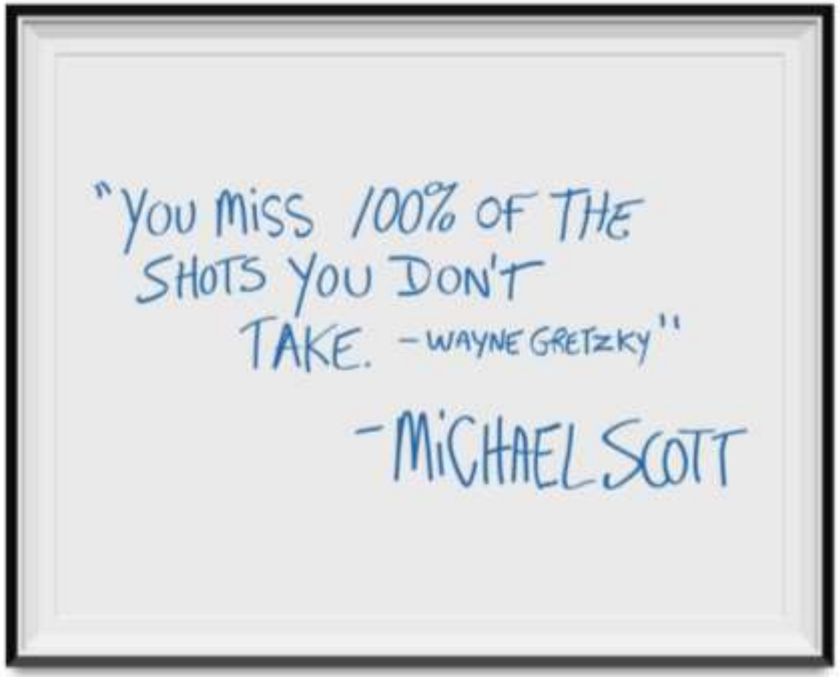
- Data culture is the collective beliefs and behaviors of the people in the organization for leveraging data for improved business performance. – Forbes
- Data culture is the collective behaviors and beliefs of people who value, practice, and encourage the use of data to improve decision-making. – Tableau
- Gartner’s third CDO Survey lists data cultures as the number one challenge for realizing benefits of data and analytics

Data Culture

- 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

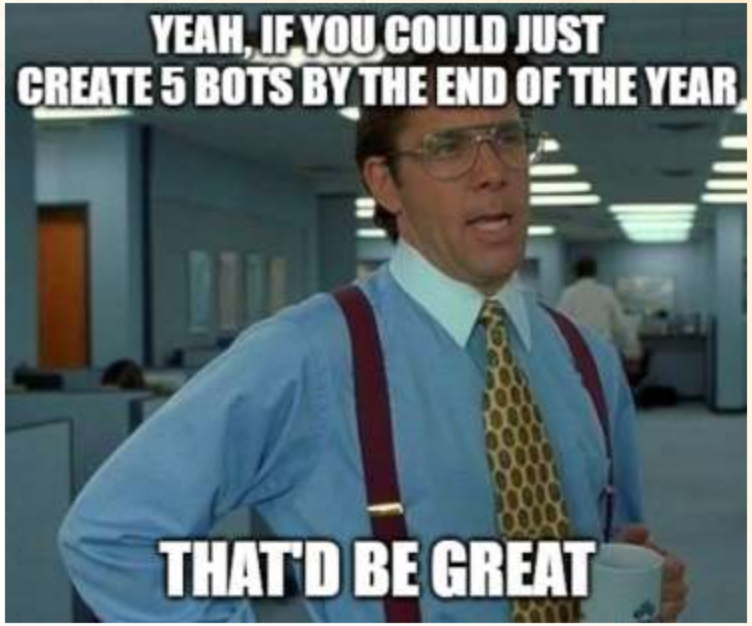
- Data culture is the **behaviors** of the **people** using **data** for **improved** decision making. – Schmucki



AI Expert Network

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.

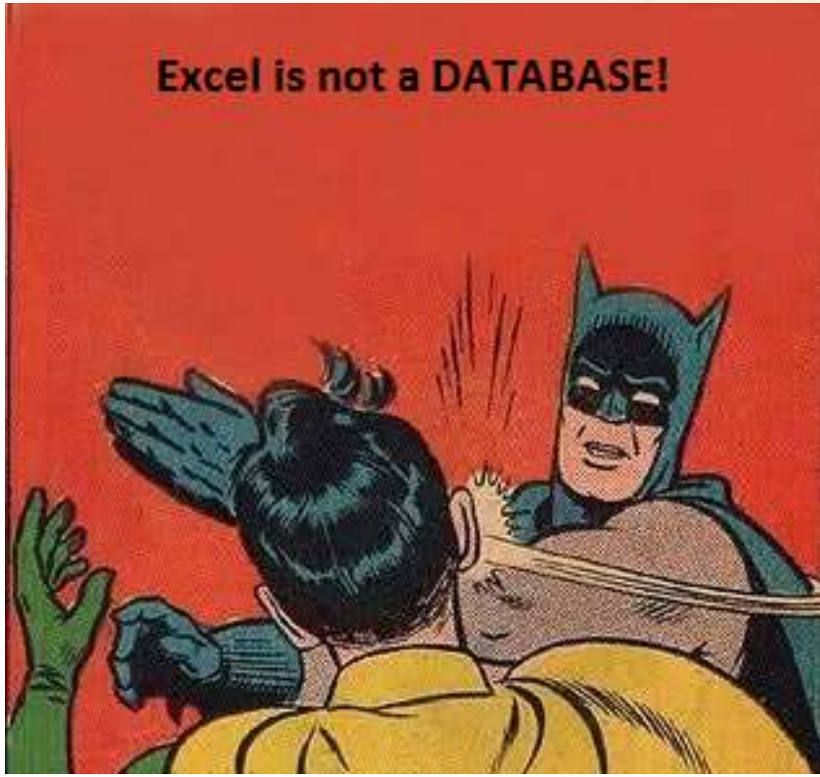
1. **Business leaders** lead analytics transformation across organization
2. **Delivery managers** deliver data- and analytics-driven insights and interface with end users
3. **Workflow integrators** build interactive decision-support tools and implement solutions
4. **Visualization analysts** visualize 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

AI EXPERT NETWORK

CELEBRATING 25 YEARS

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

Classic mechanics

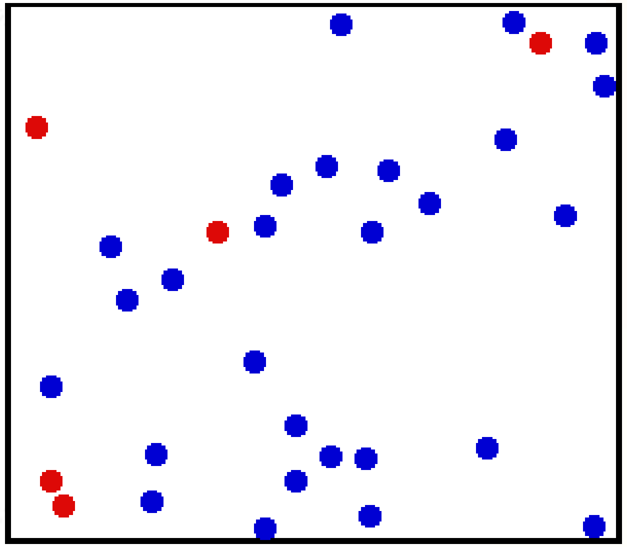
$$\frac{(p_1-p_2)}{p1} < F_\gamma \cdot x_T \rightarrow$$

$$Q_a = \frac{1}{60} \cdot 4.17 \cdot C_v \cdot p_1 \cdot \left(1 - \frac{\frac{p_1-p_2}{p_1}}{(3F_\gamma \cdot x_T)}\right) \cdot \sqrt{\frac{\frac{p_1-p_2}{p_1}}{(T_a+273.15)}}$$

$$\frac{(p_1-p_2)}{p1} \geq F_\gamma \cdot x_T \rightarrow$$

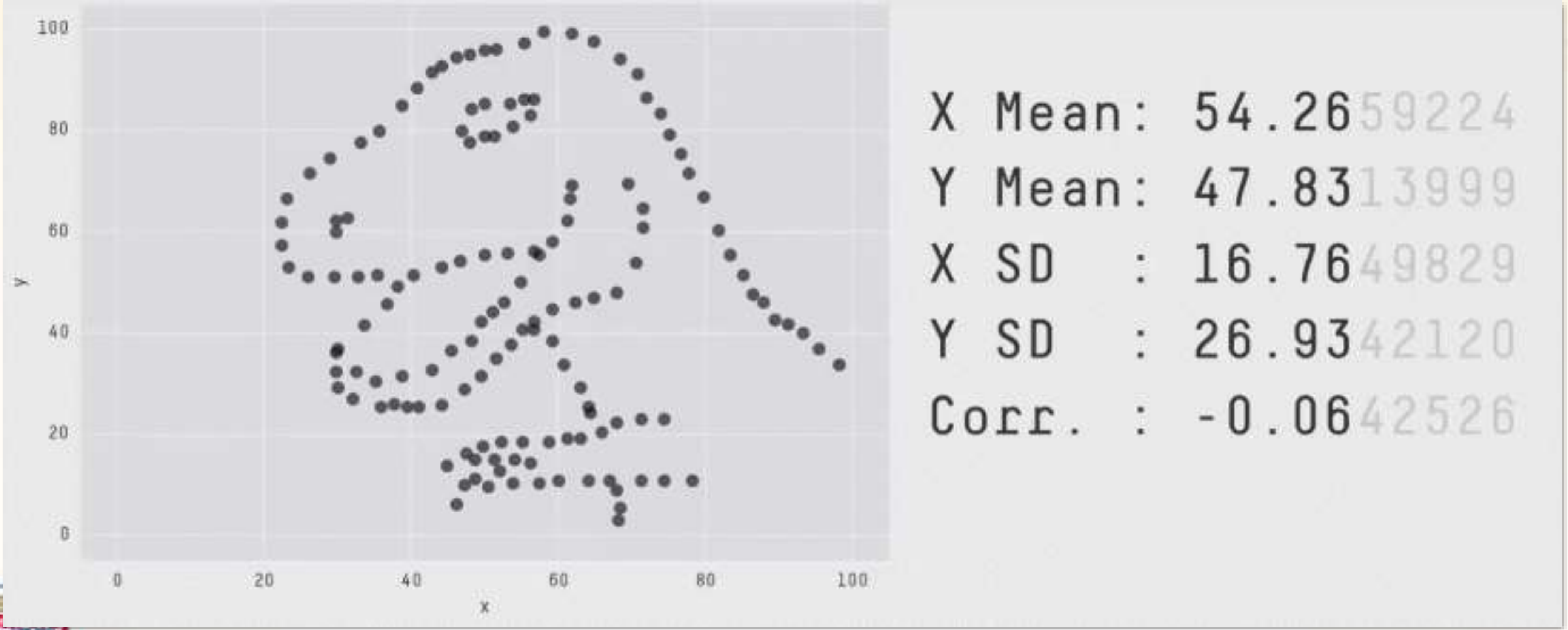
$$Q_a = \frac{1}{60} \cdot 0.667 \cdot 4.17 \cdot C_v \cdot p_1 \cdot \sqrt{\frac{F_\gamma \cdot x_T}{T_a+273.15}}$$

Statistical mechanics



AI vs. Statistics

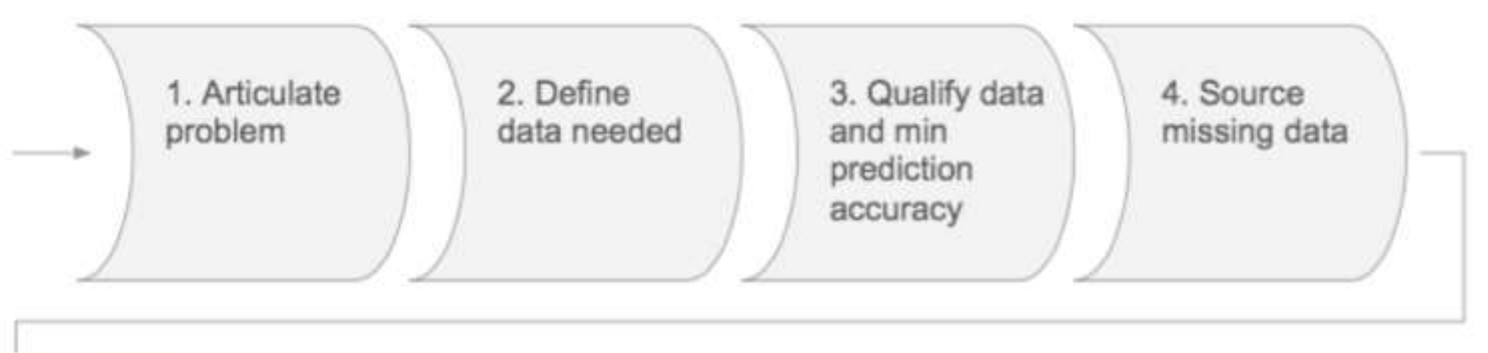
More than just multivariable models...



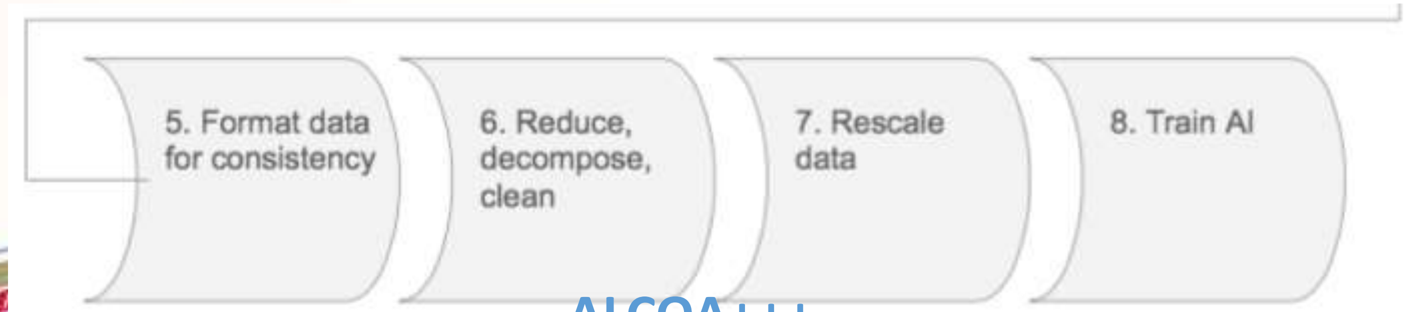
Understanding AI: Data (the Secret Sauce)

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

7 steps to consider when preparing data



Findable, **A**ccessible, **I**nteroperable, **R**eusable principle

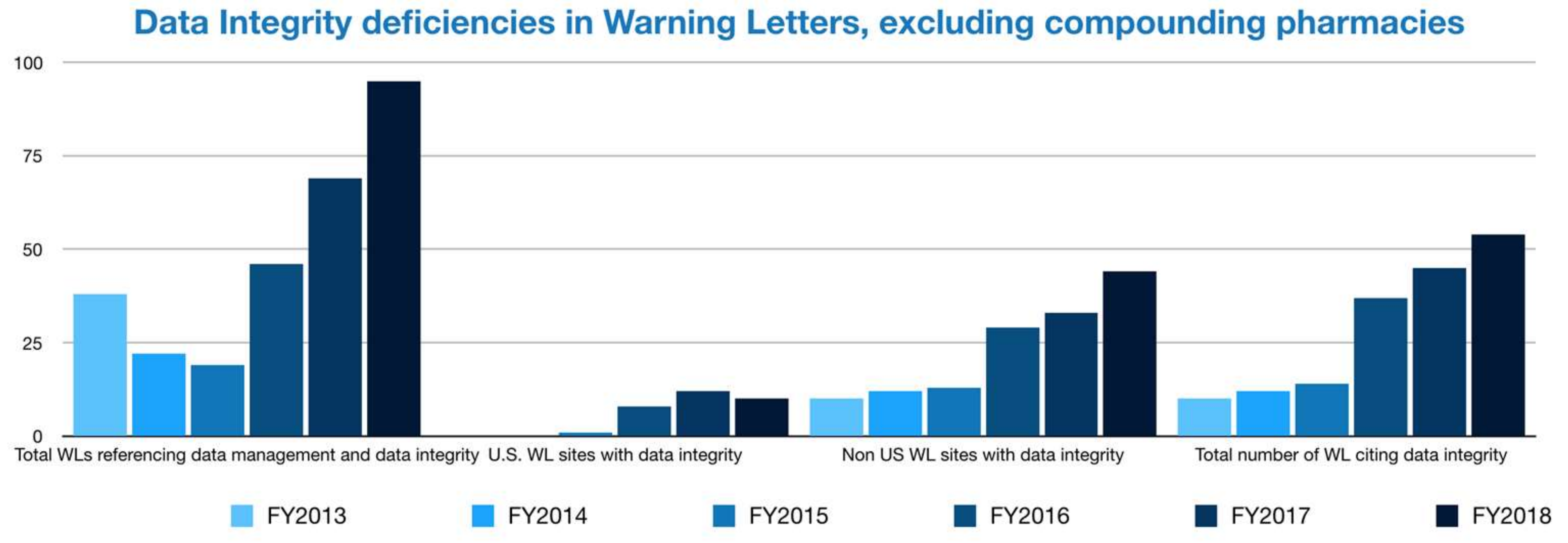


ALCOA+++

Vas Narasimhan, CEO of Novartis AG, in a 2018

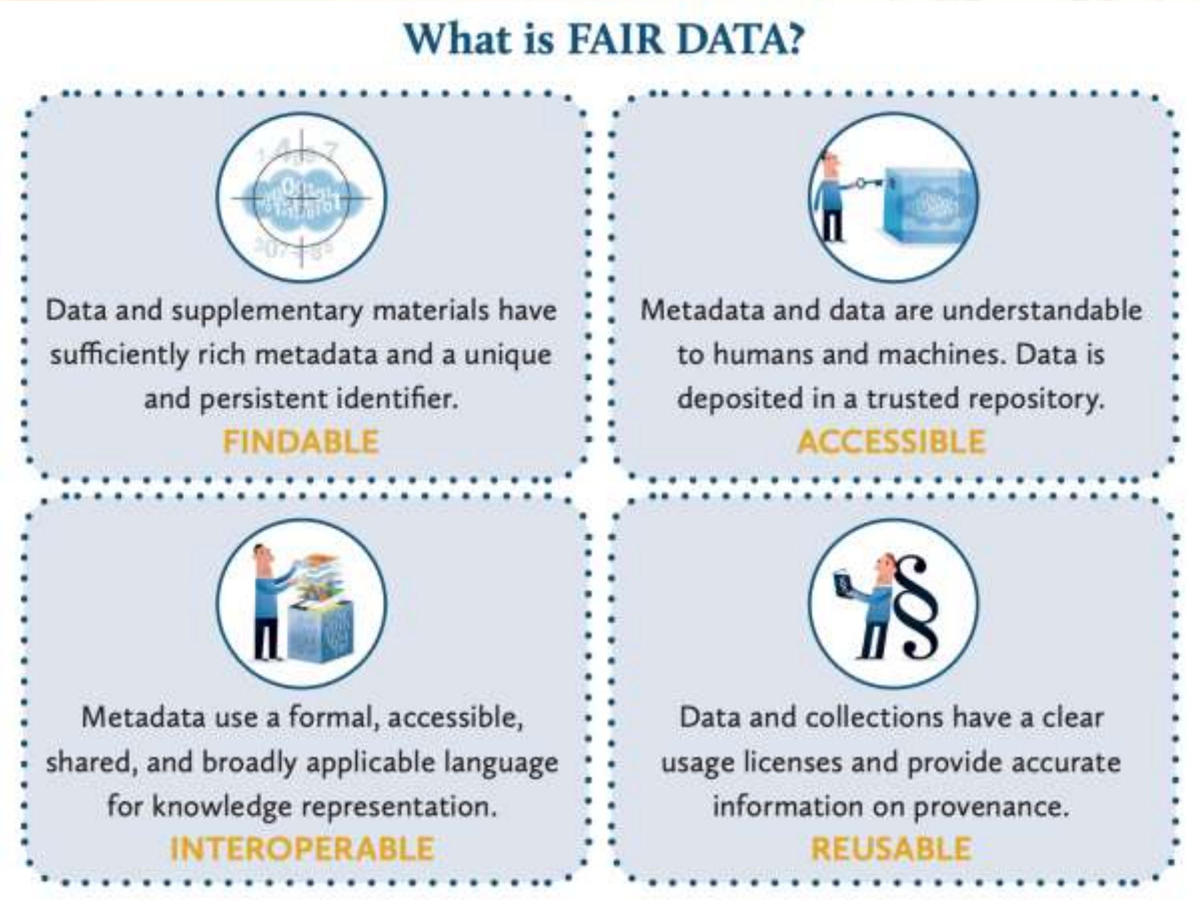
"We’ve had to spend most of the time just cleaning the data sets before you can even run the algorithm"

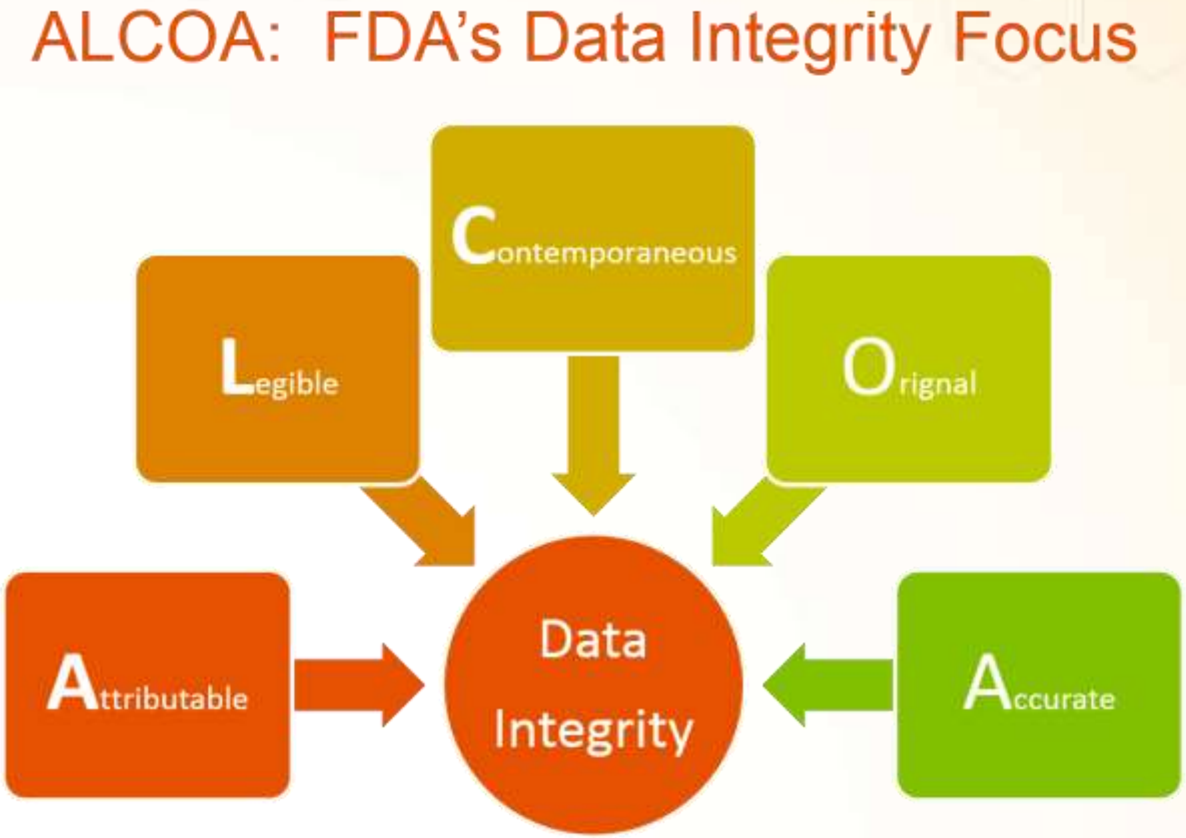
AI starts with data. In pharma, with quality data



Source: fda.gov & PharmaceuticalOnline

AI starts with data. In pharma, with quality data

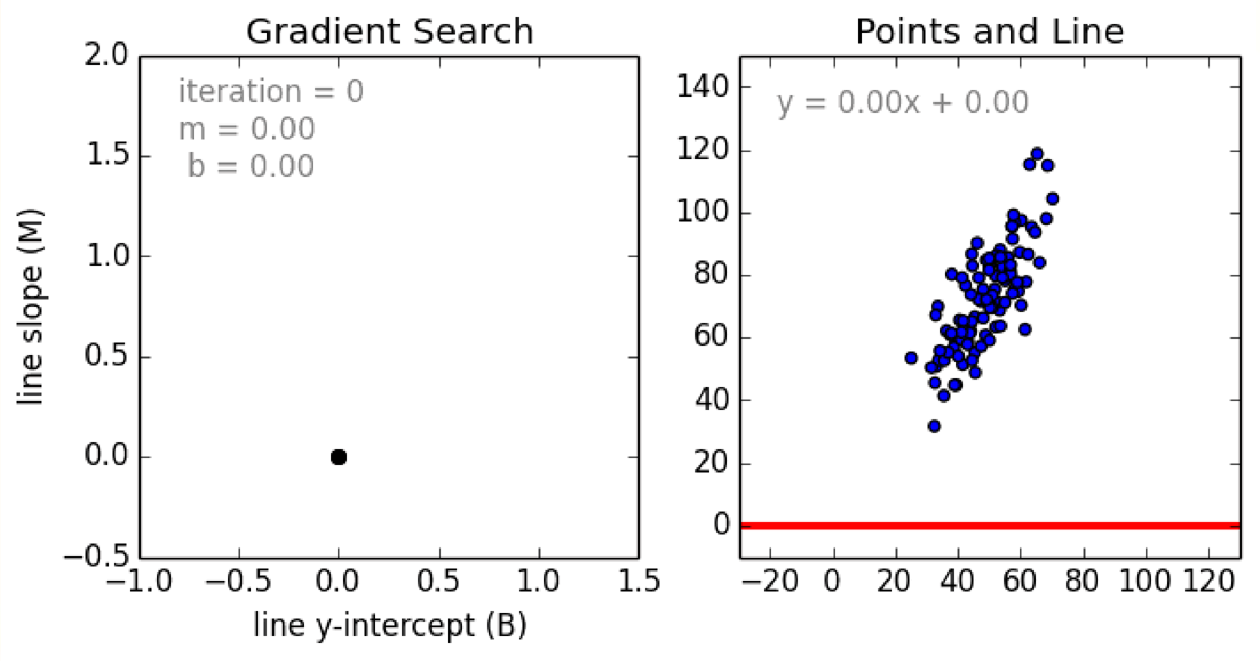




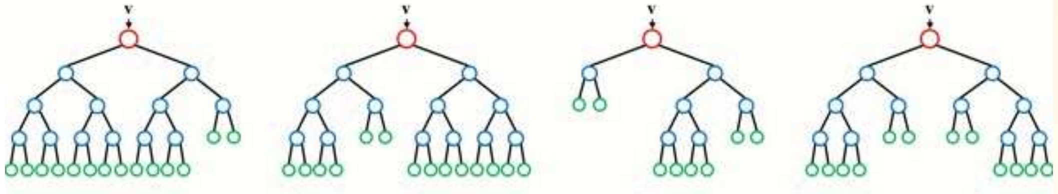
Understanding AI: Algorithm & Math

“No Free Lunch”

Linear and Polynomial Regression



Regression Trees and Random Forests



AI is already here because data is available

nature

Explore content ⌄ About the journal ⌄ Publish with us ⌄

nature > news > article

Published: 05 April 2017

Machine learning predicts the look of stem cells

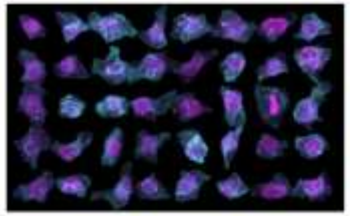
Amy Maxmen

Nature (2017) | Cite this article

610 Accesses | 4 Citations | 524 Altmetric | Metrics

Website contains thousands of 3D stem cell images and could eventually help with better understanding diseases like cancer.

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.



Structural differences in the DNA (purple) and cellular membrane (blue) of genetically identical stem cells. Credit: Allen Institute for Cell Science

nature

Explore content ⌄ About the journal ⌄ Publish with us ⌄ Subscribe

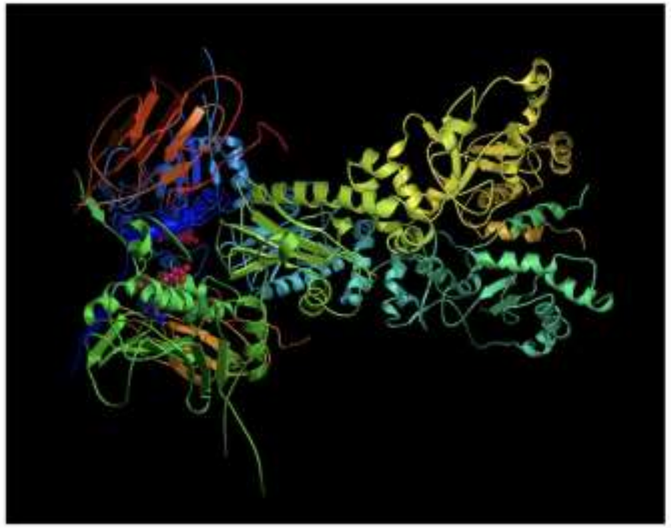
nature > news > article

NEWS | 22 July 2019

AI protein-folding algorithms solve structures faster than ever

Deep learning makes its mark on protein-structure prediction.

Matthew Hutson



Predicting protein structures from their sequences would aid drug design. Credit: Edward Kinsman/Science Photo Library

nature

Explore content ⌄ About the journal ⌄ Publish with us ⌄ Subscribe

nature > news > article

NEWS | 30 November 2020

'It will change everything': DeepMind's AI makes gigantic leap in solving protein structures

Google's deep-learning program for determining the 3D shapes of proteins stands to transform biology, say scientists.

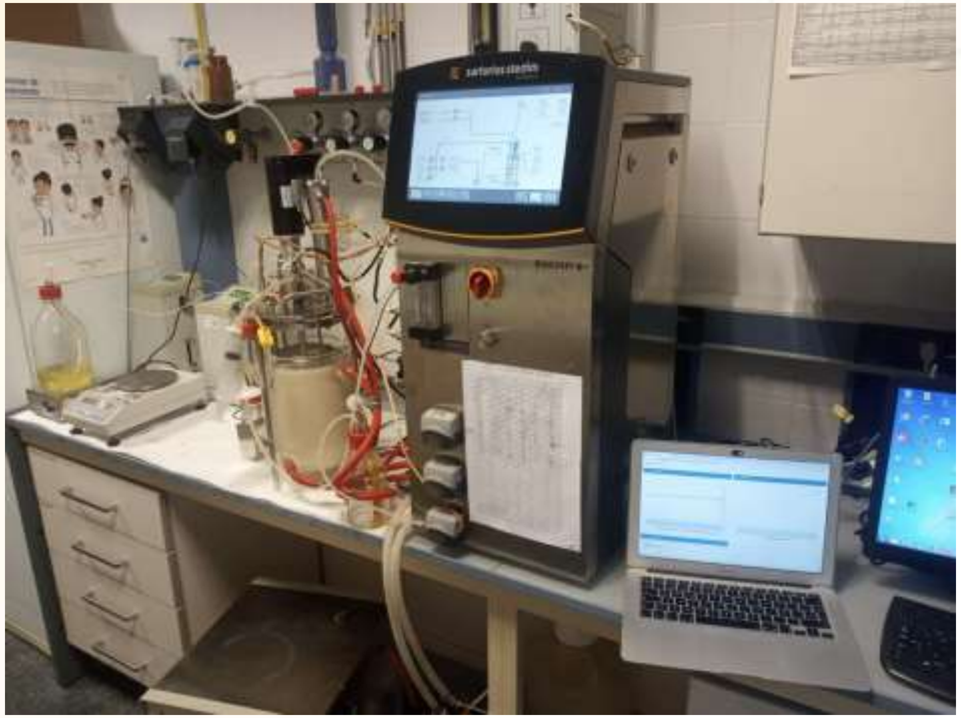
Ewen Callaway

A protein's function is determined by its 3D shape. Credit: DeepMind

Bioreactor Fermentation Process

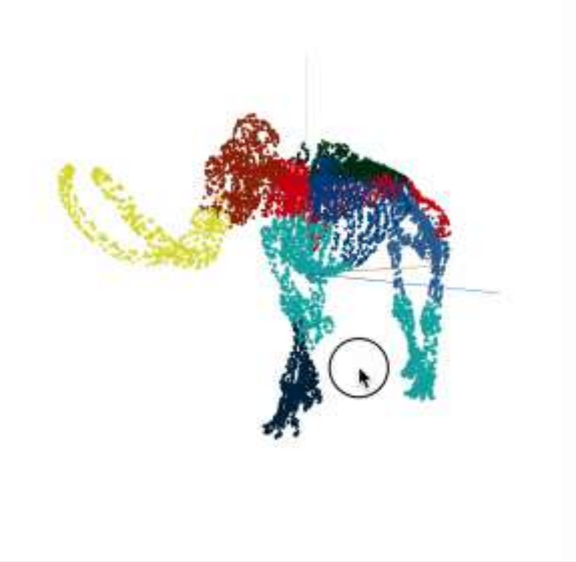
→ Optimise the hypoxic conditions for *Pichia Pastoris* yeast to maximize production

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

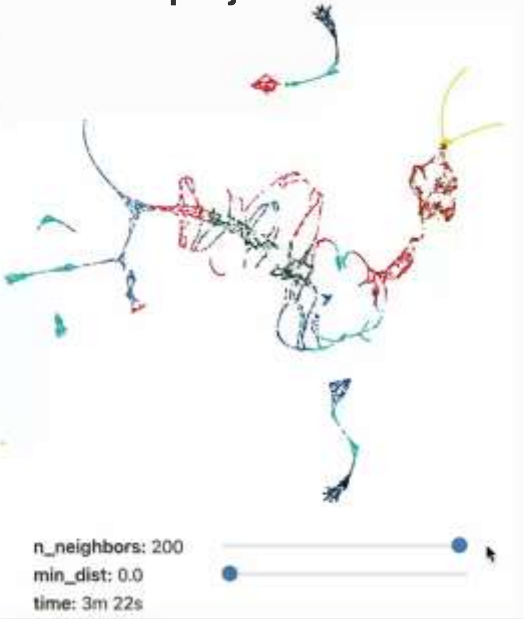


	Normoxic			Hypoxic		
	Good	Average	Bad	Good	Average	Bad
Phase I (Batch)	FBHPX2 FBHPX5 !! FBHPX6 FBHPX9	FBHPX8		FBHPX3 FBHPX4 !! FBHPX10 FBHPX11	FBHPX7	
Phase II (Adaptation)	FBHPX2 FBHPX5 !! FBHPX6 FBHPX8	FBHPX9		FBHPX3 FBHPX4 !! FBHPX7 FBHPX10 FBHPX11		
Phase III (Early Fed Batch)	FBHPX2 FBHPX5 !! FBHPX6 FBHPX8 FBHPX9			FBHPX7	FBHPX4 !! FBHPX10 FBHPX11	FBHPX3
Phase IV (Later Fed Batch)	FBHPX5 !! FBHPX6 FBHPX8 FBHPX9	FBHPX2		FBHPX7 FBHPX10 FBHPX11		FBHPX3 FBHPX4 !!

Reducing dimensionality by means of AI

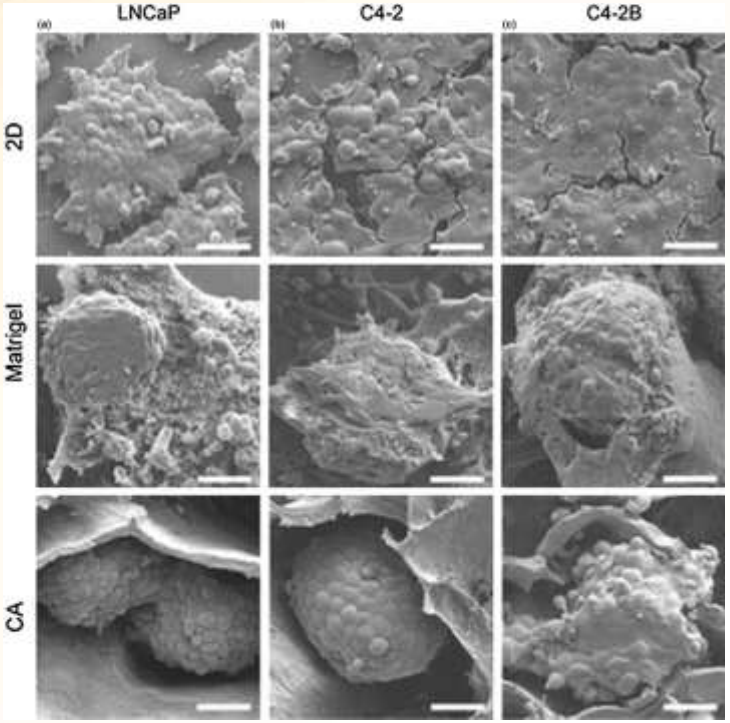


2D UMAP projection



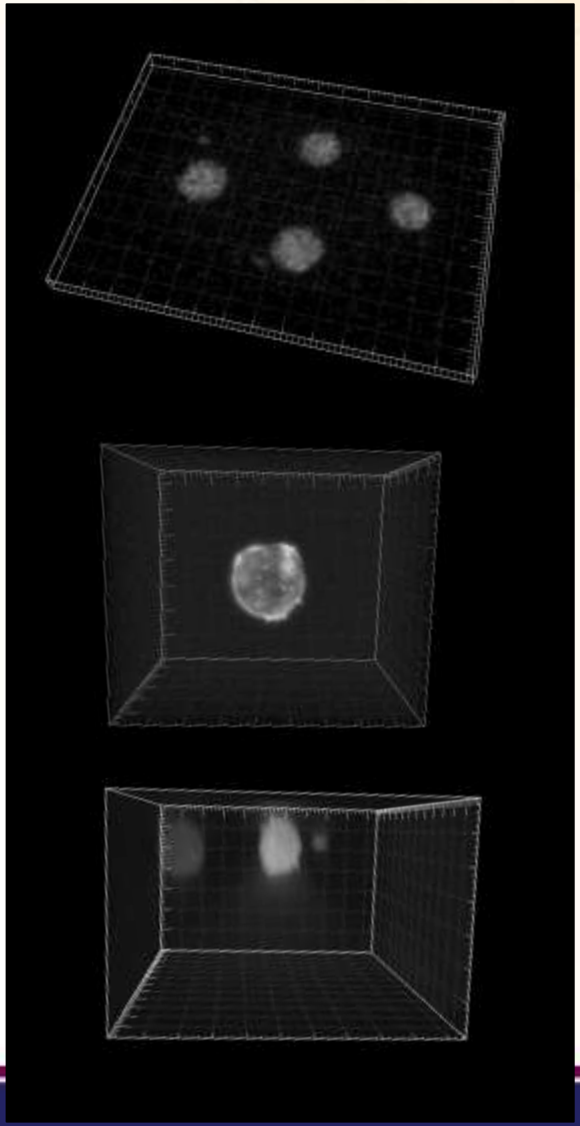
2D t-SNE projection

Automatizing human cognition around C&GT



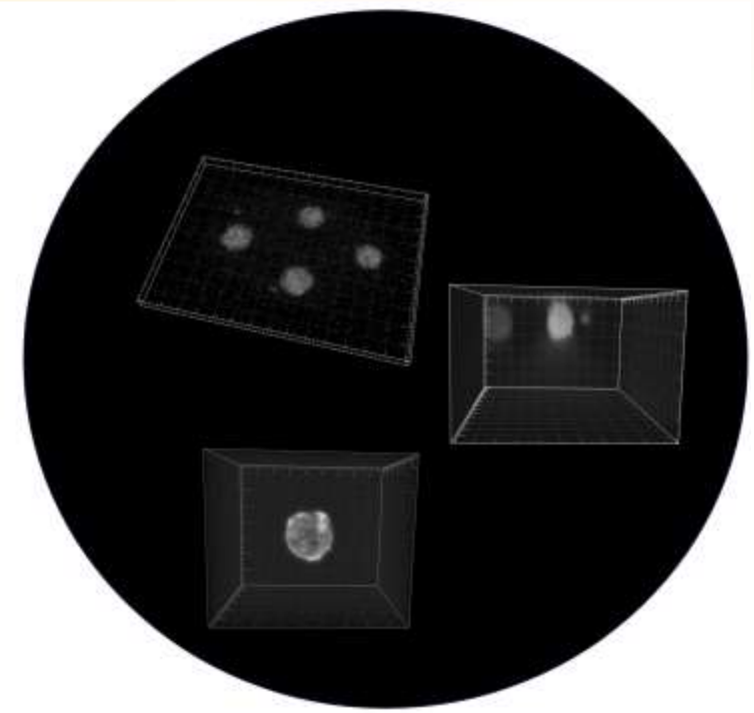
Cancer cells grow and divide at an abnormally rapid rate, are poorly differentiated, and have abnormal membranes, cytoskeletal proteins, and morphology. The abnormality in cells can be progressive with a slow transition from normal cells to benign tumors to malignant tumors.

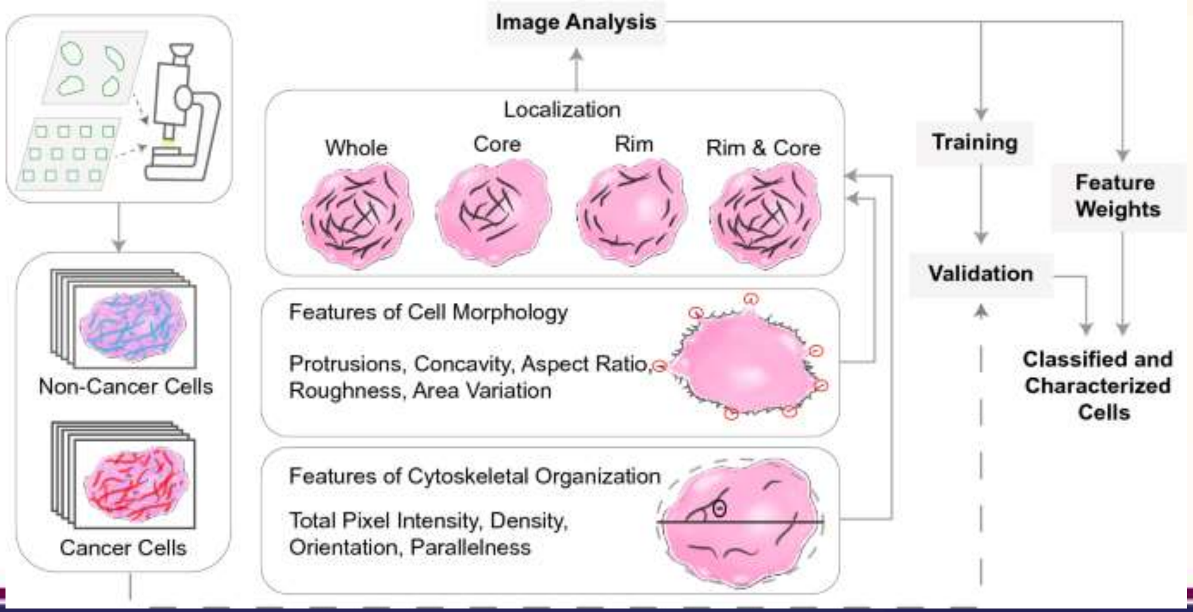
- Self-sufficiency in growth signals
- Insensitivity to anti-growth signals
- Evading programmed cell death
- Limitless replicative potential
- Sustained angiogenesis
- Tissue invasion and metastasis



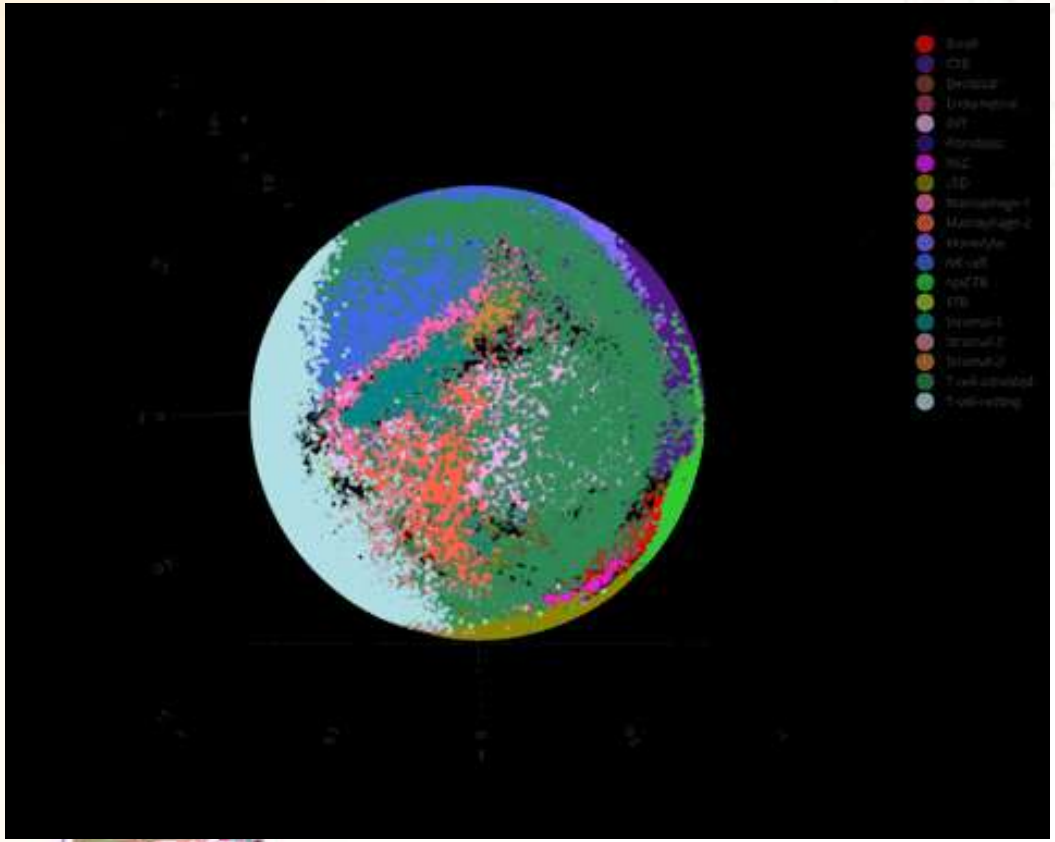
Automatizing human cognition around C&GT

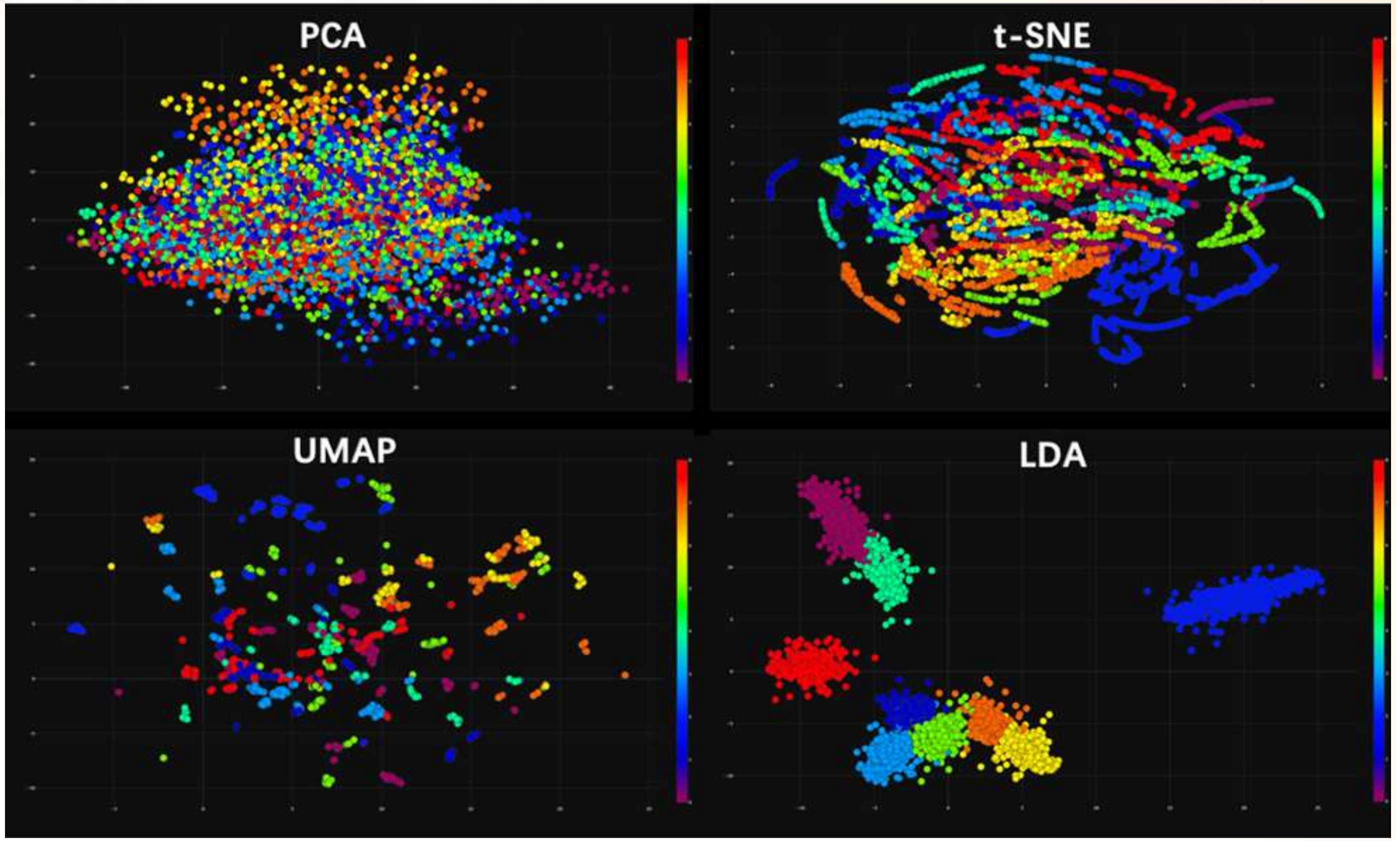
Normal	Cancer	
		Large, variably shaped nuclei
		Many dividing cells; Disorganized arrangement
		Variation in size and shape
		Loss of normal features





Reducing dimensionality and clustering features





CPV of the Future

Goal: Create a digital twin to manage a biotech process working under GxP conditions

AI Summit info

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

Maturity Area	Maturity Factors	Maturity Level Characterization: Level 1	Level 2	Level 3	Level 4	Level 5
Culture						
Communication	Established channels for communication with feedback mechanisms.	No communication	Random or ad hoc communications, likely one-way (no feedback loop).	Random or ad hoc communications with feedback loop.	Coordinated communications with feedback loop.	
Procedures	Established procedures defining key roles and responsibilities, activities and processes.	No established procedures for AI life cycle management.	Ad hoc processes and loosely defined procedures for AI life cycle management.	Some procedures for AI life cycle management, not covering the full life cycle or not complete.	Procedures for all areas of AI life cycle management and integrated into the broader QMS.	
Early AI Awareness and Understanding	Evolving from lack of understanding to trust	Mistrust, black box	Interest in AI and research as to how it is being used by others. ID the issue to be addressed	Build trust, select algorithm	Build model, better decision-making capability	
Education	Programs /	No information	Ad-hoc Awareness	Citizen data scientist –	Ability to discern if data	

Maturity Area	Maturity Factors	Maturity Lev Characterization: Level 1	Level 2	Level 3	Level 4	Level 5
Data Management						
Data acquisition	Data quality and volume for use by AI	Paper based, collected by humans, not real time. Prone to human error	Systems in place for data collection. Dataset with missing, partial, low volume	Transformation required from raw data to final format. Human interaction to attain data	Removal of data bias and sampling noise. Automated process to attain data	Real time quality data. Fault is understood and immediately corrected. System is aware of accurate data
Data source	Data source and structure	Spreadsheets are primary data source. Minimal standards for data format and structure	Data collected from many systems with no curation. Structured and unstructured data. Human-data interaction needed.	Minimal level of data curation for integrating from many data sources and data management practice.	Most data are curated, organized and accurate. Management of the data lifecycle.	All data are curated and best practices are in place.
Data integrity / security		Paper-based data with uncontrolled access	Manage access to data, address content protection	Conform to ISO IEC standards FAIR principles ALCOA+	Data security/integrity in place to deal with continuous learning model (self-improving). Continuously evolving	Fully encrypted data. Strong access control and password management. Full traceability for changes
Data a						

Maturity Area	Maturity Factors	Maturity Level Characterization: Level 1	Level 2	Level 3	Level 4	Level 5
Governance and Organization						
Executive analytics orientation (Chief Information Officer, Chief Data Officer, Chief Executive Officer, etc.)	How well the AI outputs influence executive insights and strategic direction.	Limited if any understanding of AI output (may be limited to classical statistic and not true AI).	AI is used as a prospect tool for non-critical tasks and using non-sensitive data. AI Teams are experimental task forces mainly formed by internships or non-experimented people	Beginning to understand relationships between data sets and outputs. AI is starting used for several critical actions (e.g. manufacturing) but with no access to sensitive data (e.g. AI Vision)	Roadmap for integration of data for reliability (i.e., plan for a plan). AI tasks and results are part of the critical processes and sensitive data is used for AI model creation. AI activity is lead at the management level.	Strategic int of outputs fo synthesis of insights that directly influence and reflect the business strategy used for decision making at the highest level. AI activity is lead at an enterprise level
Good Data Sciences practices/adoption	Available SOP Consolidated structures Organizational recognition and transversal support (SME)	Data scientists apply their own criteria, tools and data storage in order to get insights	There are non-formal agreements regarding the AI tools (e.g. IDE and frameworks), data storage, and ways to present results	There is a departmental strategy for AI tasks, mainly for the data, algorithm and model life cycle	The AI practices are described in the Quality System	The proposed AI practices have been followed at least 2 loops in the Quality System review and they are integrated in the rest of global procedures
Human Resource Structure and Capabilities	How well the number and qualifications of resources aligns to organizational goals.	Hire consultant. Data scientists workforce are based on internships and outsourcing	Hire data analyst. AI activities are only known from a reduced team around the AI resources ie. decentralized model	Hire data engineer and specialized data scientists. They work under IT management structure	Hire data scientist/ ML Engineer + AI Manager. AI activity is known by collateral departments. More centralized model	Establish AI team and have CIO. AI is part of the cultural structure of the company managed as an asset by HR
Competence of organization with AI-based tactics. Neither SME nor strategic consideration. Commoditization and democratization of AI and AI-based tools.	How well the rank and file understand, and are able to utilize AI-based tools and approaches.	Limited, if any, understanding of AI and tools based upon them. Find Excel stimulating.	Some individuals moderately proficient - or- many individuals generally aware.	Many beginning to understand and moderately capable with AI in functional applications. Can readily implement with some guidance. Same personnel using AI tools without interaction in daily work	Many individuals can receive new AI tools and independently understand their application, power and operation. AI results used in critical processes	Routine AI use in business and/or operational efficiencies attained. Can battle with Death Star.

Maturity Area	Maturity Factors	Maturity Level Characterization: Level 1	Level 2	Level 3	Level 4	Level 5
Tools and Techniques						
1_Depth of AI application in a company vs. breadth of application	Digitization level (across company: warehousing, R&D, manufacturing, quality, labs, etc.) For example, Product Development might be all over it, but the rest of the company still using clipboards. Or, the whole company might be employing some initial, basic steps.	General understanding of function but not the use of AI	Operation of AI vs understanding programming	Elements of AI are present but it is not used broadly and comprehensively	Degree of application to replace human actions Models are dynamic vs static Models are implementation specific with updates from own measurements, self-calibrating	Quantify the number of areas automated year over year Use across multiple projects and departments within the company. Majority of personnel trained and using AI capability
Analytics	Gartner's maturity model, bioforum maturity level, FDA maturity model Application of AI	Basic descriptive analytics Raw massaged data, not interpreted	Advanced Descriptive analytics (Hindsight) Interpreted data	Diagnostic analytics (Insight) First level interpretation	Predictive analytics (Foresight) Know what the outcome is going to be with manual response to maintain controlled state	Prescriptive analytics, Process control, Avoid failure, No need for human interaction for a controlled state
IT	Proprietary software/libraries Open source platform libraries Software installation Limited computational capabilities	$ Ad Hoc Software platforms decided by each Data Scientist. Platforms and algorithms self-maintained No organizational control on the applications nor hardware.	$$ Part of the functional budget Investment in AI as a potential opportunity	$$$ Dedicated, but limited budget There is a non-dedicated budget for AI, although AI is included in the same bag with the rest of technologies	$$$$ Dedicated, strategic budget Strategic platform management via life cycle and quality system approach Version evolution.	$$$ Enterprise budget as a corporate service Enterprise level strategy and support alignment. ROI supports cost of on-going AI growth plans.

Current Team Members

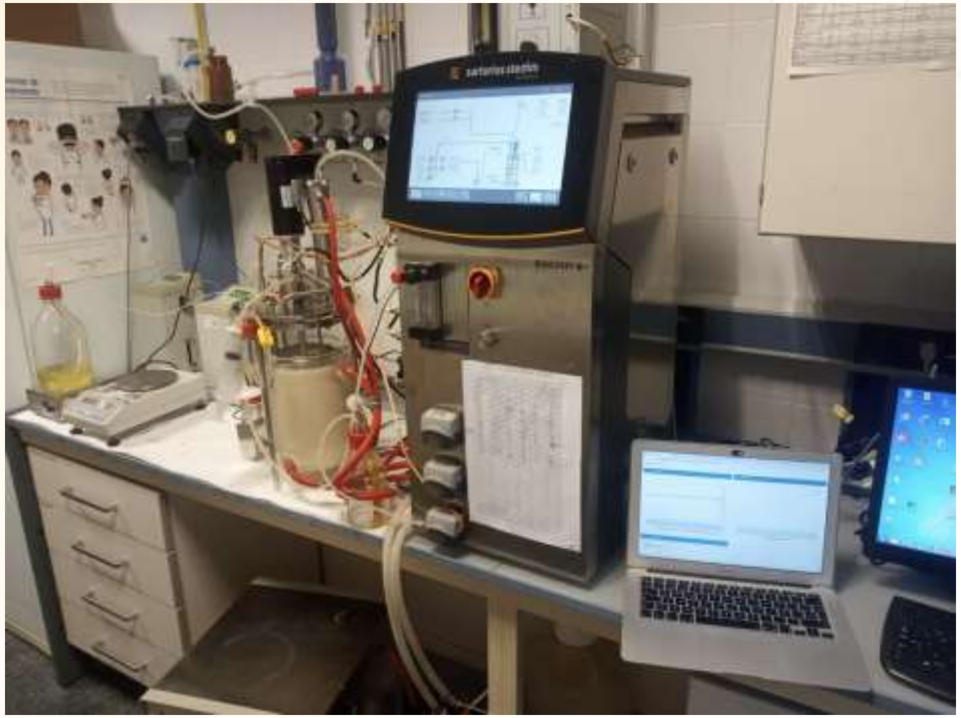
Antonio Moreira *(Execution co-lead*	University of Maryland, Baltimore County	Vice Provost for Academic Affairs
Ben Stevens (PV Team lead)	GSK	Director CMC Policy and Advocacy
Catarina Leitão	4Tune Engineering	CPV Expert
Christophe Agut	Sanofi Pasteur	Head of Process Validation and Statistics Expertise
David Hubmayr (Task Force 1 lead)	CSL Behring	Manager, Process Development & Breakthrough Technologies, R&D
David Lapeña	Infors	Area Sales Manager Southern Europe & Africa
Francisco Valero (Task Force 2 lead)	Universidad Autònoma of Barcelona	Professor and head of department of BioChemical Engineering
Joeri Van Wijngaarden	Aizon	Innovation Lab R&D Project Manager
Mario Stassen (Task Force 3 lead)	**AFDO** (AI in Operations Team)	BioPharma Regulatory expert

Matt Schmucki	**AstraZeneca & AFDO**	Lean Coach and CPV Expert
Mauro Giusti (PV Team lead)	Eli Lilly	Director, Technical Services/Mfg Sciences
Nilanjan Banerjee	University of Maryland, Baltimore County	Professor, Computer Science and Electrical Engineering
Sandrine Dessoy	GSK	Science and Technology Innovation Director
Shereya Maiti	Bayer Pharmaceuticals	Senior Scientist
Toni Manzano *(Execution lead)*	**Aizon & AFDO**	CSO and Co-founder
Ciro Cottini	Chiesi	Digital, Data & Modelling Head
Holger Mueller	Bluesens	Director

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



	Normoxic			Hypoxic		
	Good	Average	Bad	Good	Average	Bad
Phase I (Batch)	FBHPX2 FBHPX5 !! FBHPX6 FBHPX9	FBHPX8		FBHPX3 FBHPX4 !! FBHPX10 FBHPX11	FBHPX7	
Phase II (Adaptation)	FBHPX2 FBHPX5 !! FBHPX6 FBHPX8	FBHPX9		FBHPX3 FBHPX4 !! FBHPX7 FBHPX10 FBHPX11		
Phase III (Early Fed Batch)	FBHPX2 FBHPX5 !! FBHPX6 FBHPX8 FBHPX9			FBHPX7	FBHPX4 !! FBHPX10 FBHPX11	FBHPX3
Phase IV (Later Fed Batch)	FBHPX5 !! FBHPX6 FBHPX8 FBHPX9	FBHPX2		FBHPX7 FBHPX10 FBHPX11		FBHPX3 FBHPX4 !!

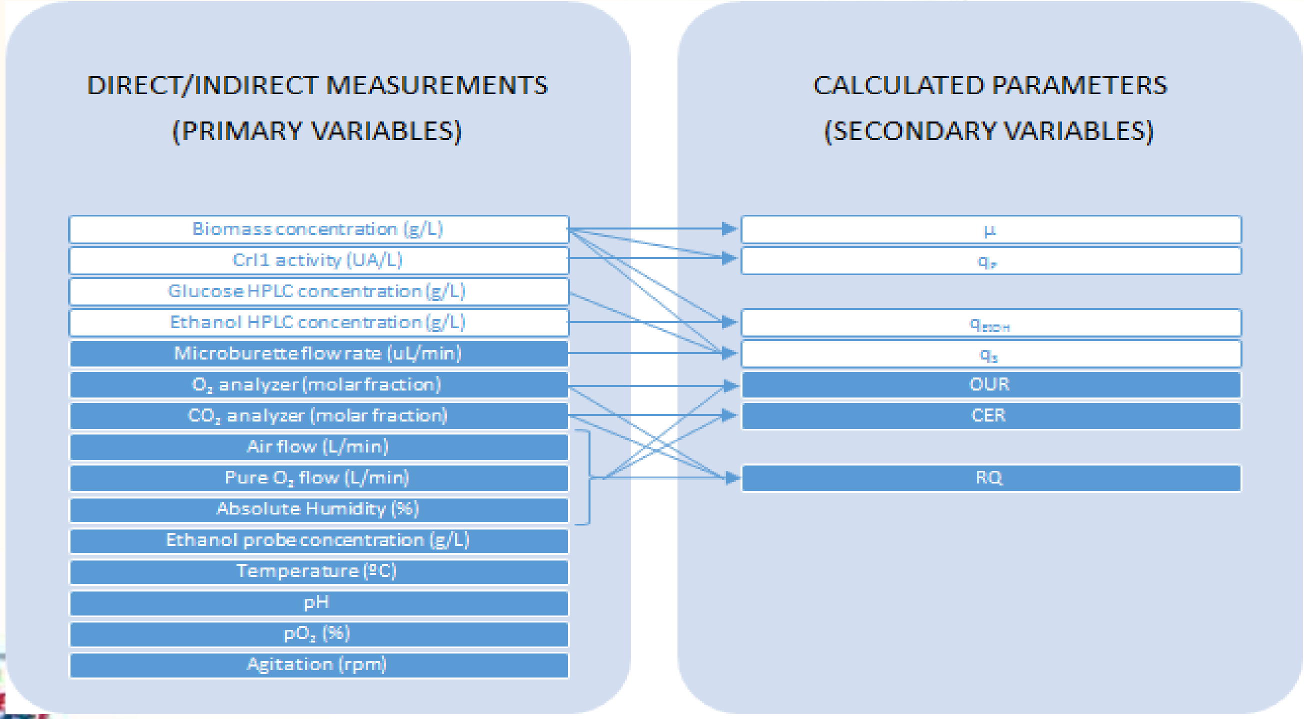
AI-guided Process Monitoring

Phase 1: Batch

→ Controlled process, no manual actions

→ Anomalies due to equipment, material qualities or improper behaviour of biomass

→ Multiple relevant factors can to be monitored to detect underperforming batches



Q: how can we bring value from AI in an automated process without interacting with the unit?

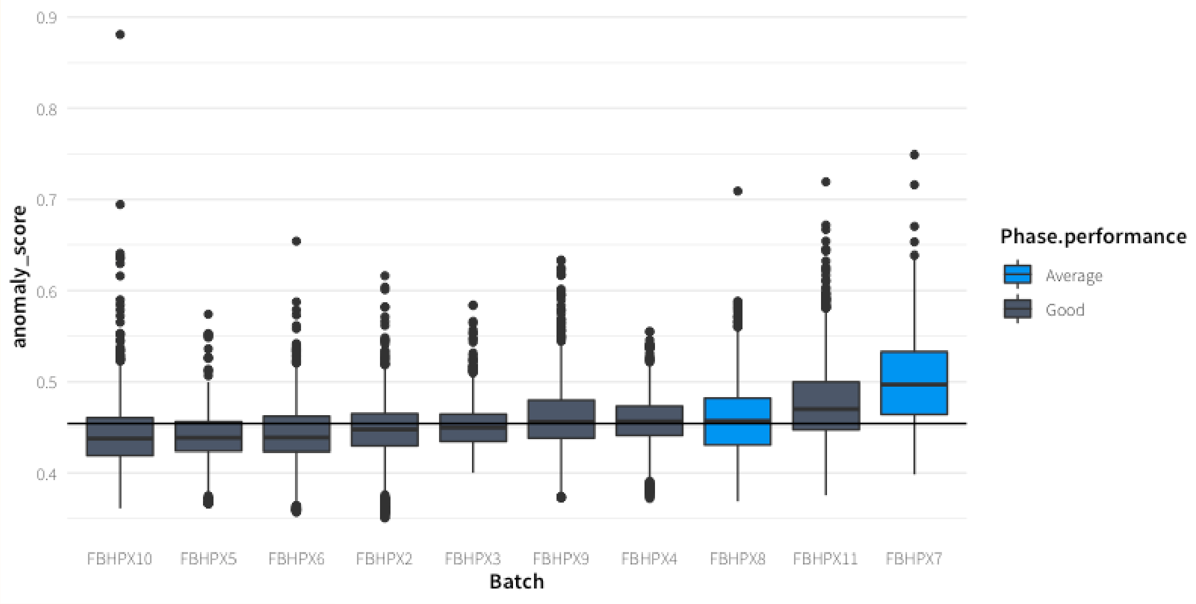
AI-guided Process Monitoring

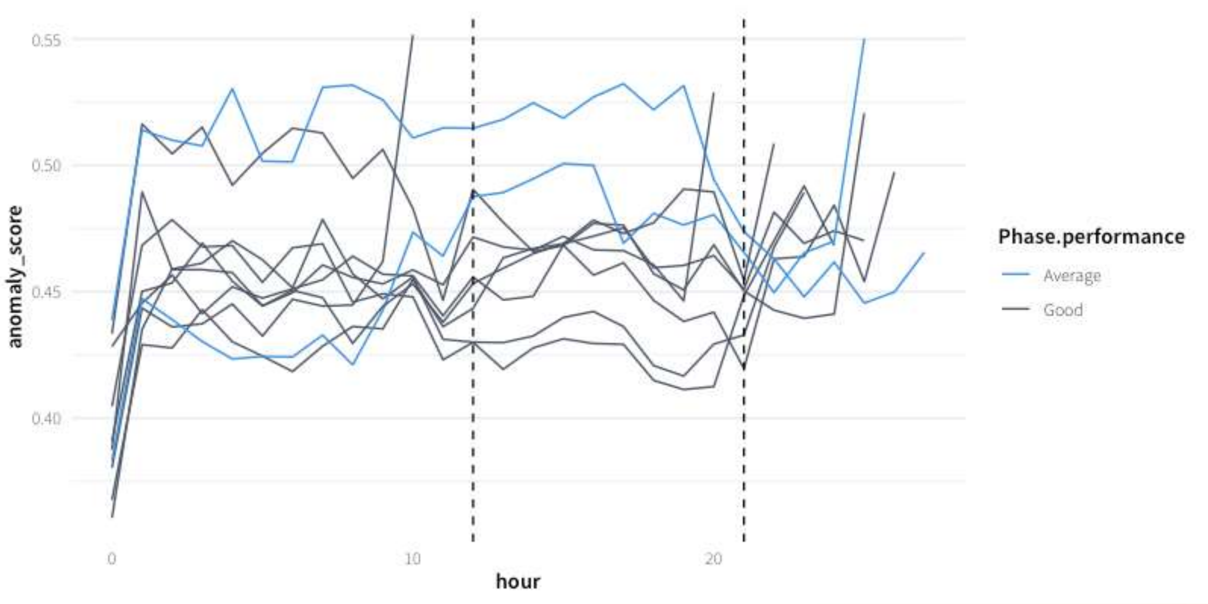
Phase 1: Batch

→ Controlled process, no manual actions

→ Anomalies due to equipment, material qualities or improper behaviour of biomass

→ Multiple relevant factors can to be monitored to detect underperforming batches



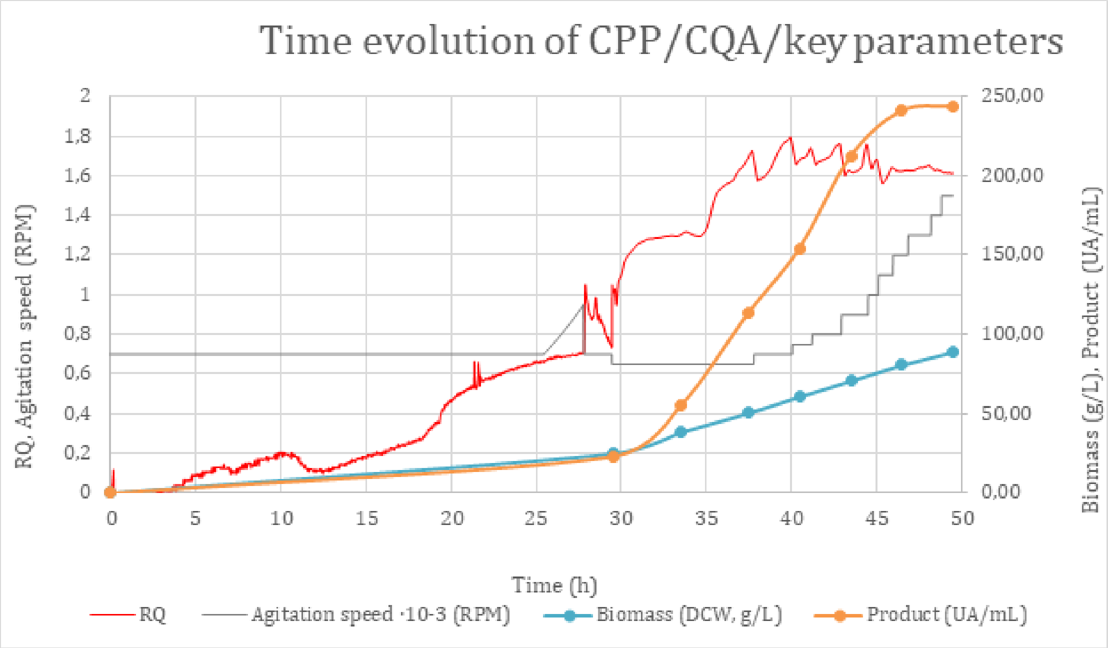


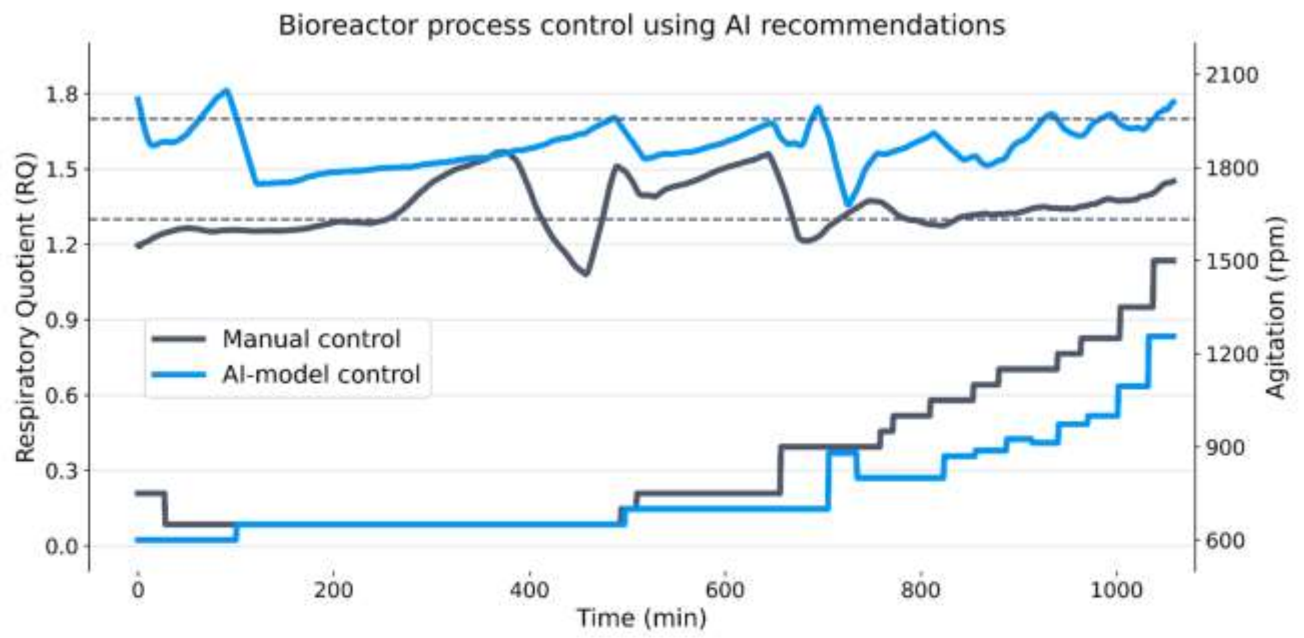
AI-guided Process Control

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.





Source: Gasset et al. (2022) Front. Bioeng. Biotechnol.

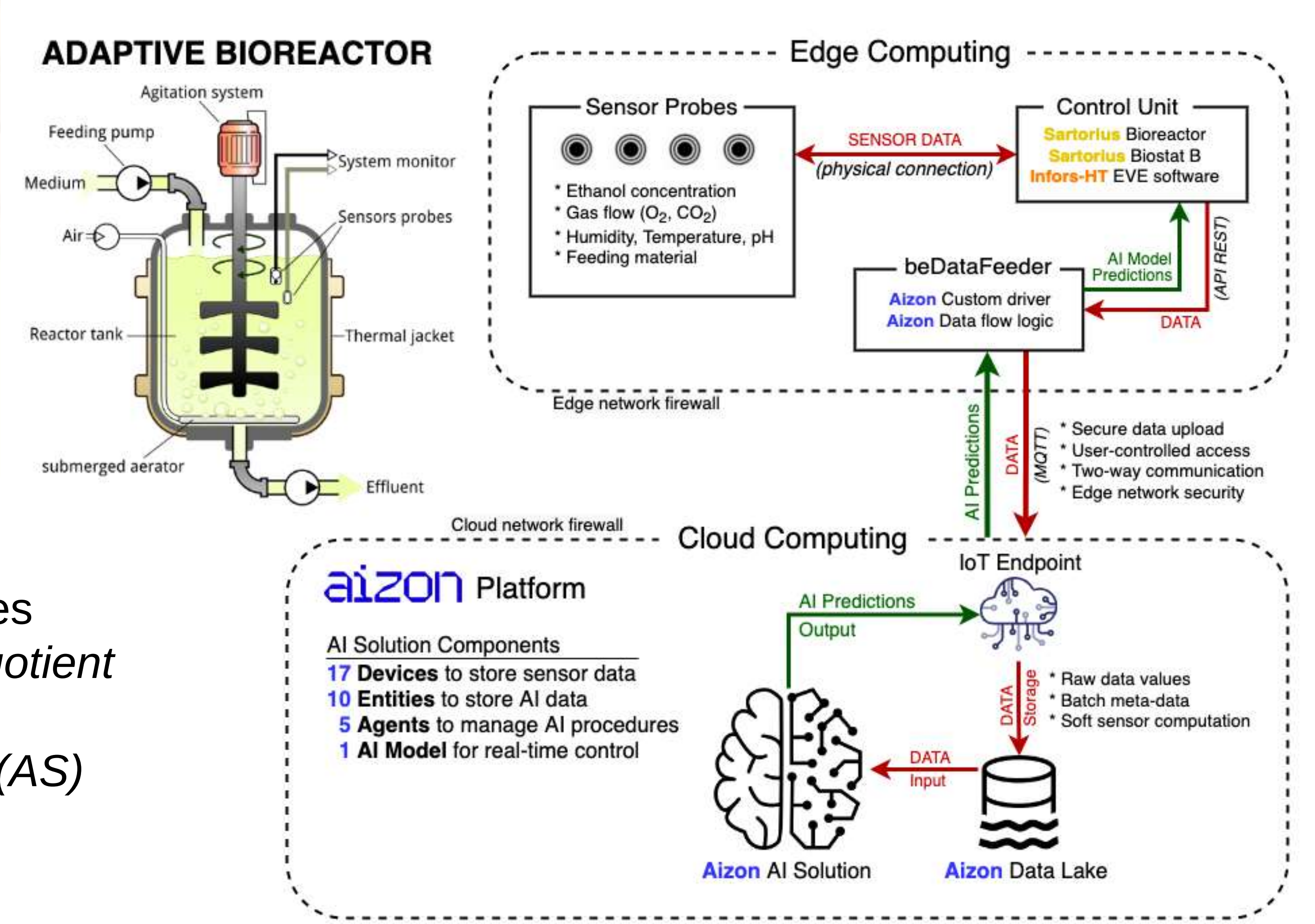
Control Automation For Optimal Data Governance

Control Strategy

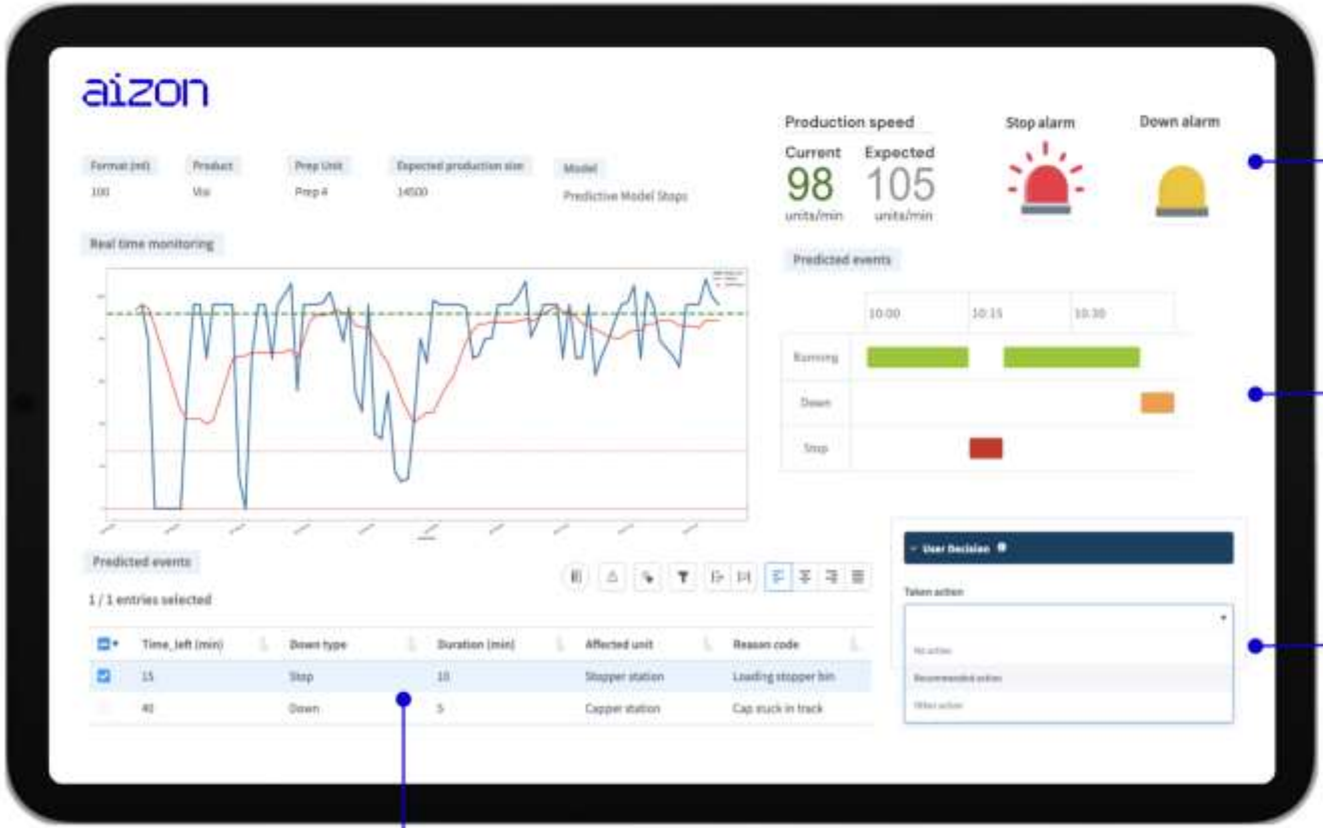
- PAT & IIoT Technology
- Combination of edge & cloud
- Fully automated data pipeline
- Coverage of full AI lifecycle (train, productivise, monitor)
- Operate in near real-time

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

CELEBRATING 25 YEARS

Produced by FoodSafety magazine

MAY 8 – 11 | 2023

Donald E. Stephens Convention Center | Rosemont, IL

COMMUNITY · EDUCATION · SOLUTIONS

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

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

Steve Mandernach
Executive Director, AFDO
smandernach@afdo.org

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