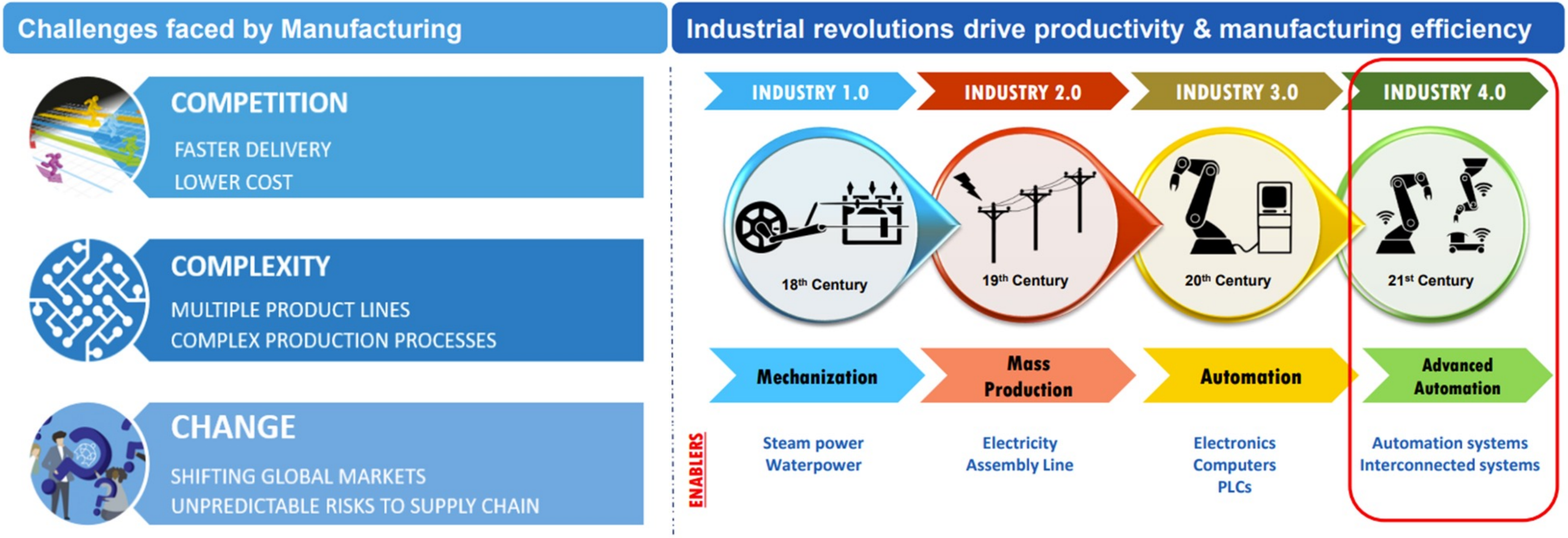
Industry 4.0: Big Data, Machine Learning and Artificial Intelligence in Cell Culture

Seongkyu Yoon, University of Massachusetts, Lowell, USA
Madhuresh Sumit, Sanofi, USA
Ravali Raju, Amgen, USA

Workshop Outline

- Welcome and intro to the session - 5 mins
- What Industry 4.0 means, survey outcome and how it can be applied to cell culture -10 mins
- Case Studies Introduction-15 mins
- Case studies breakout
  - 30 min to work in smaller group
- Debrief- 20 mins
- Conclusion-10 mins

Industry 4.0 can drive productivity and operational efficiencies



Industry 4.0: Faster, Smarter and Sustainable Productivity

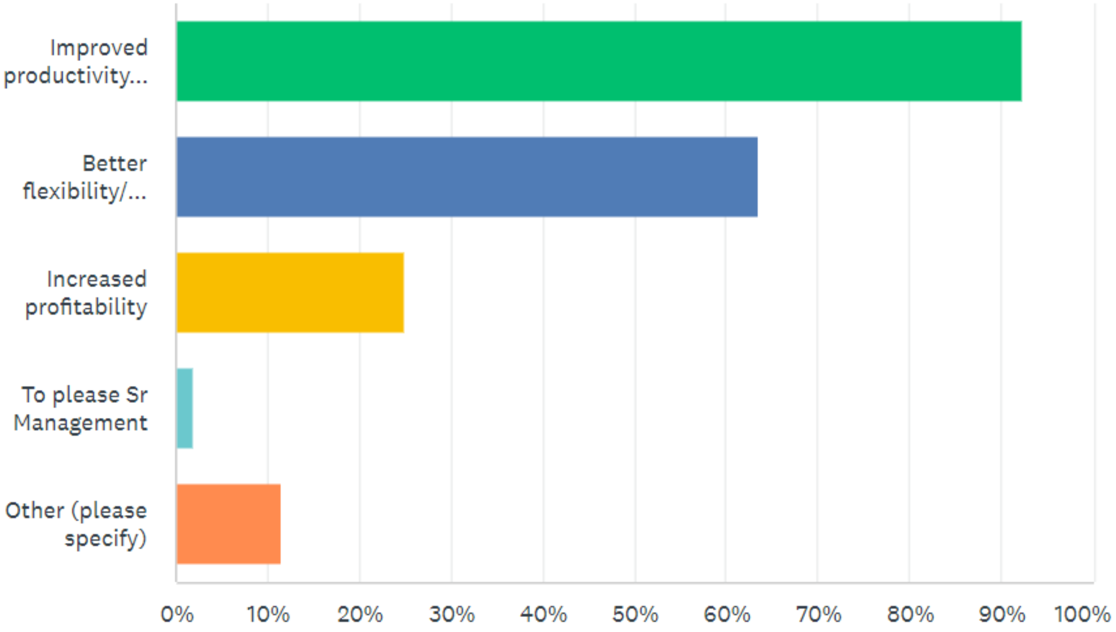
Survey Results

- **20-30%** of participants were very familiar with terms such as iOT, machine learning, virtual reality etc
- Applications to cell culture could include digital twin models, automated workflows for process/media development, process models etc
- ~**90%** of participants are actively engaged or starting to get engaged in Industry 4.0
- ~**70%** find some tools like big data and AI very valuable for applications in cell culture
- ~**90%** see path to implementation in 10 years

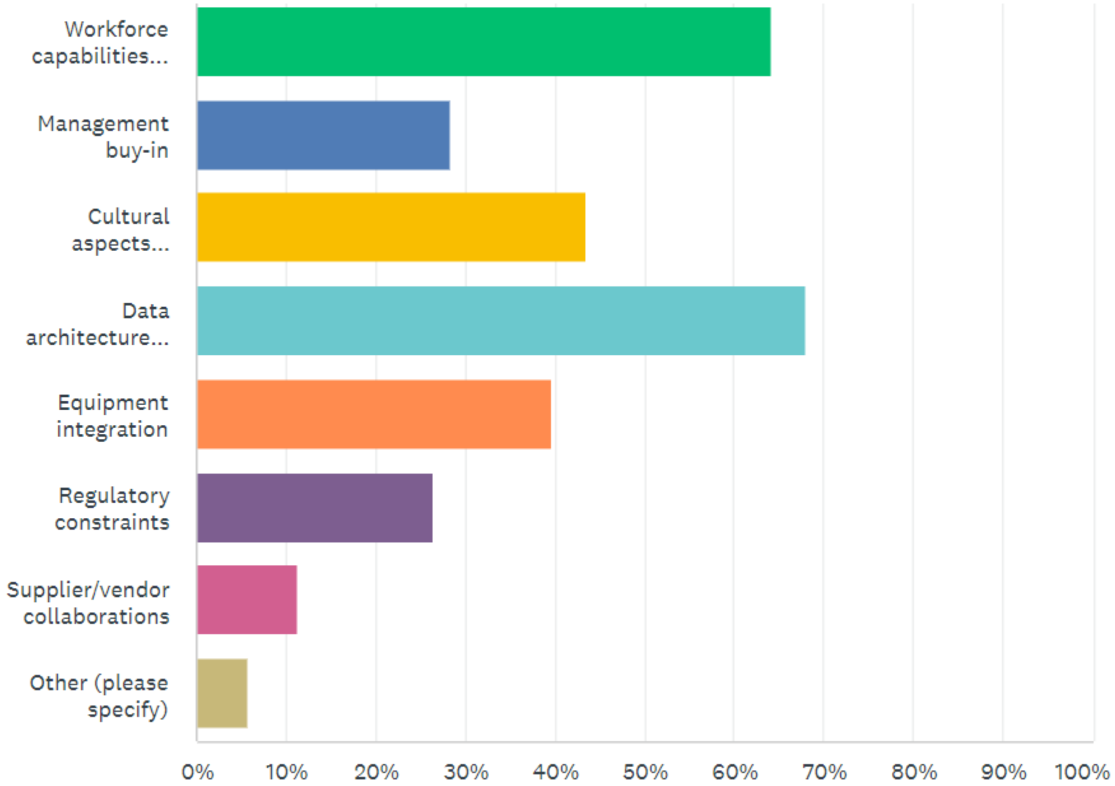
Survey Results

What are the top 2 main drivers for transforming to Industry 4.0 concepts for cell culture applications (choose only two)?

Answered: 52 Skipped: 277



What do you see as some of the biggest challenges in implementing Industry 4.0 concepts in your day-to-day work (check top 3)?



Artificial Intelligence and Machine Learning Applications in Biopharmaceutical Manufacturing

Ref: Rathore et al (2022), Trends in Biotech

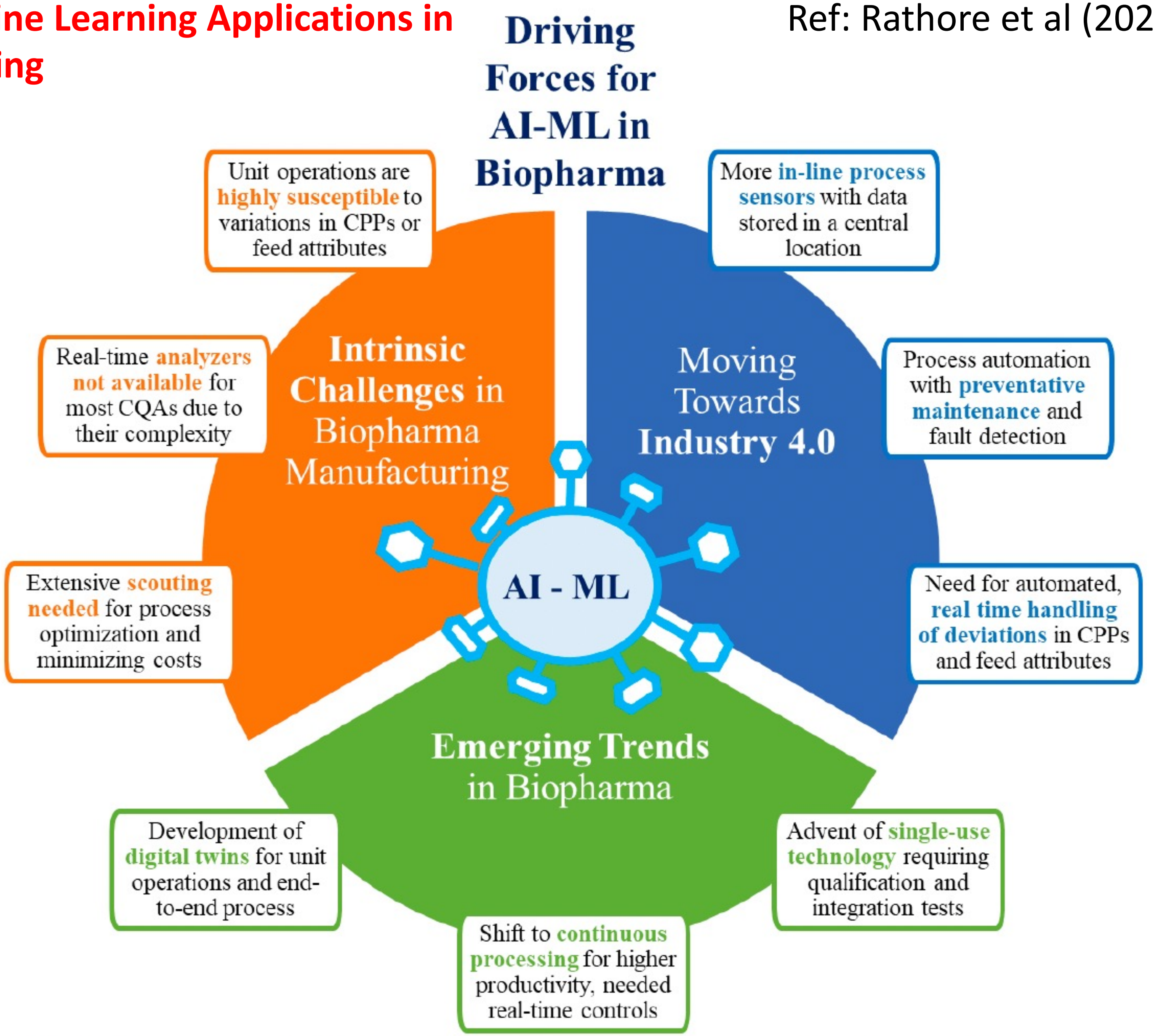
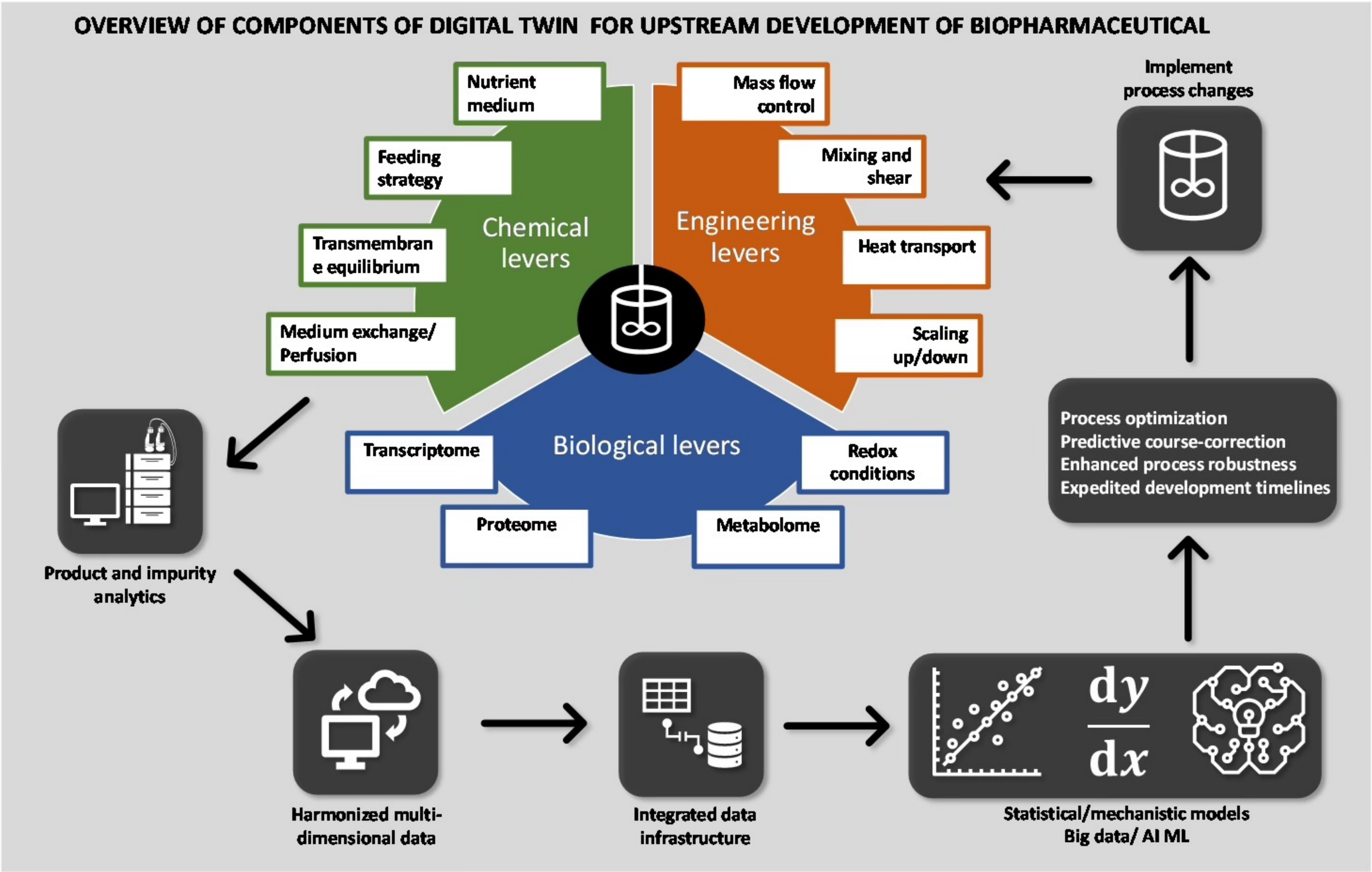


Figure II: Current challenges and driving forces in biopharmaceutical manufacturing that are spurring development of AI-ML based solutions

Digital Twin, Ideal Case



Industry 4.0 Workshop Case Studies

- Problems & Challenges
- Approaches taken
- Outcomes

- Discussion outcomes:

Alternative approaches (tools, data, analysis)

Expected outcomes (tangible and intangible)

Hurdles

C1: Bigdata analytics identify metabolic inhibitors and promoters for mAb productivity improvement

Ref: **Bioresources and Bioprocessing, 2020**

Problems and Challenges

Strategies for promoting higher titers and avoiding the accumulation of inhibitors are needed.

Metabolic pathway analysis shows the correlation between a given amino acid and the associated metabolite

Monitoring and supplementing amino acid levels in real time, ensuring that concentrations within the bioreactor remain within specified limits.

Integration of transcriptomics analysis with metabolite profiling and metabolic pathway analysis can serve as validation method

C1: Bigdata analytics identify metabolic inhibitors and promoters for mAb productivity improvement

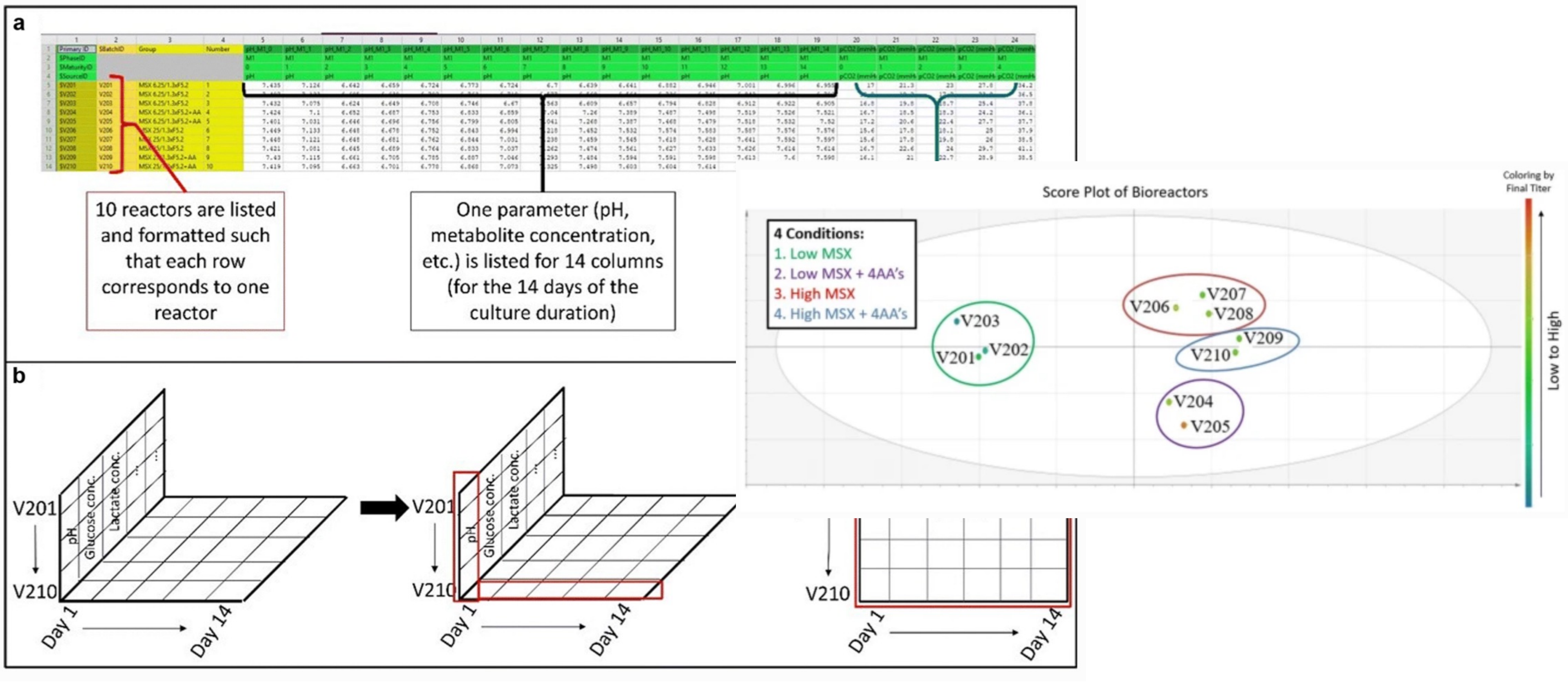
- Approaches Taken (Experimentation)

Condition	Reactor label	Expansion medium	Basal medium	Feed medium
1	V201, V202 and V203	Seed medium with 1X MSX	Basal medium	Feed medium
2	V204 and V205	Seed medium with 1X MSX	Basal medium	Feed medium with an increased concentration of Ser, Thr, Tyr, and Lys
3	V206, V207 and V208	Seed medium with 4X MSX	Basal medium	Feed medium
4	V209 and V210	Seed medium with 4X MSX	Basal medium	Feed medium with an increased concentration of Ser, Thr, Tyr, and Lys

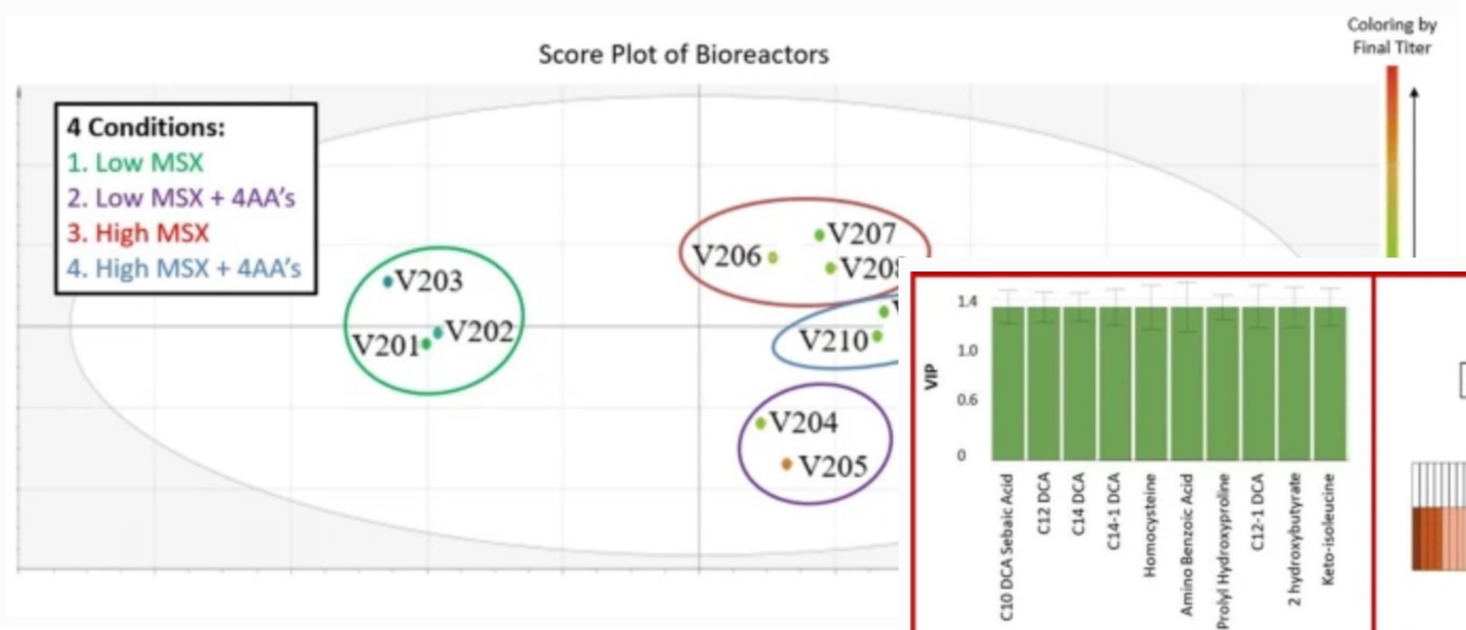
C1: Bigdata analytics identify metabolic inhibitors and promoters for mAb productivity improvement

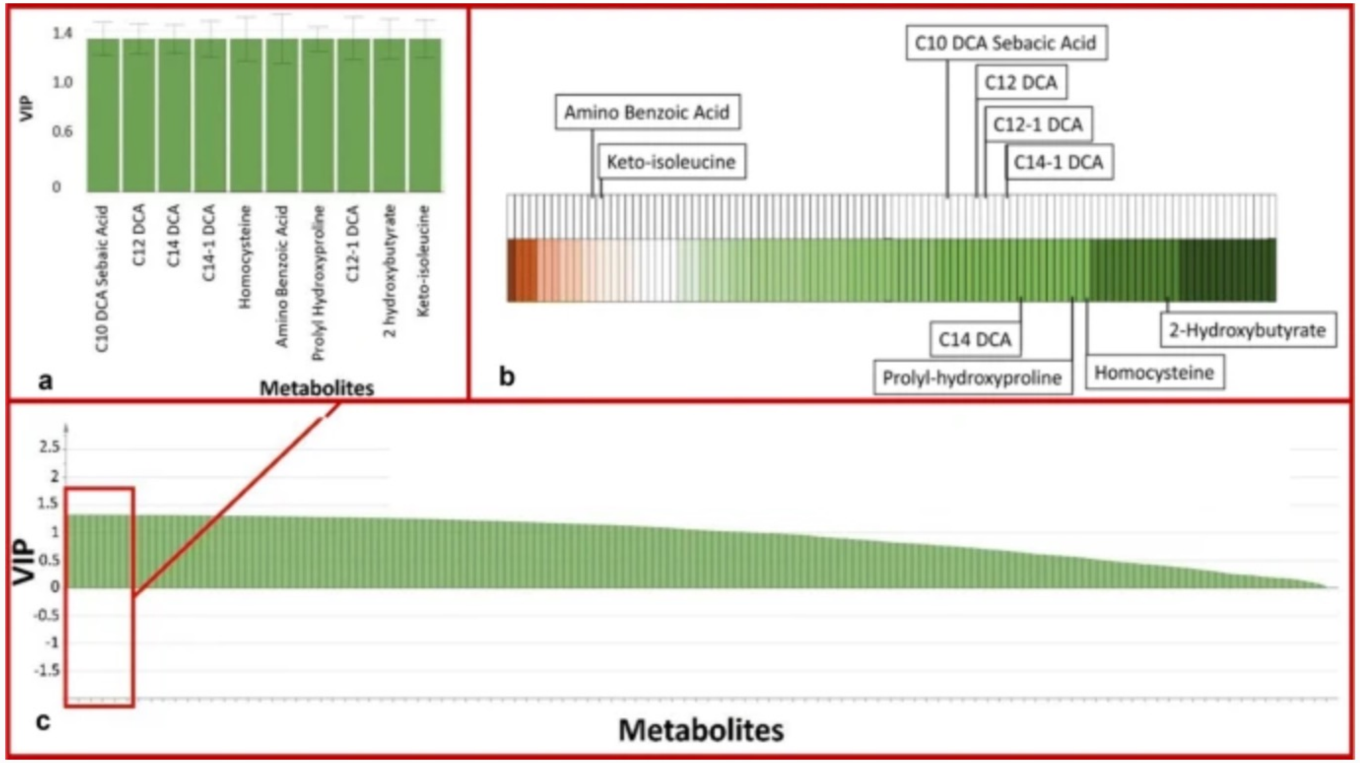
- Datasets collected:
  - UHPLC-mass spectrometric metabolomics data collection and analysis
  - CQA;
  - Bioreactor daily operation data;
  - Daily transcriptomics data
- Approaches for analysis
  - Batch modeling method to integrate daily analysis and CQA data along with bioprocessing data
  - Pathway enrichment analysis
  - Transcriptomics

C1: Bigdata analytics identify metabolic inhibitors and promoters for mAb productivity improvement

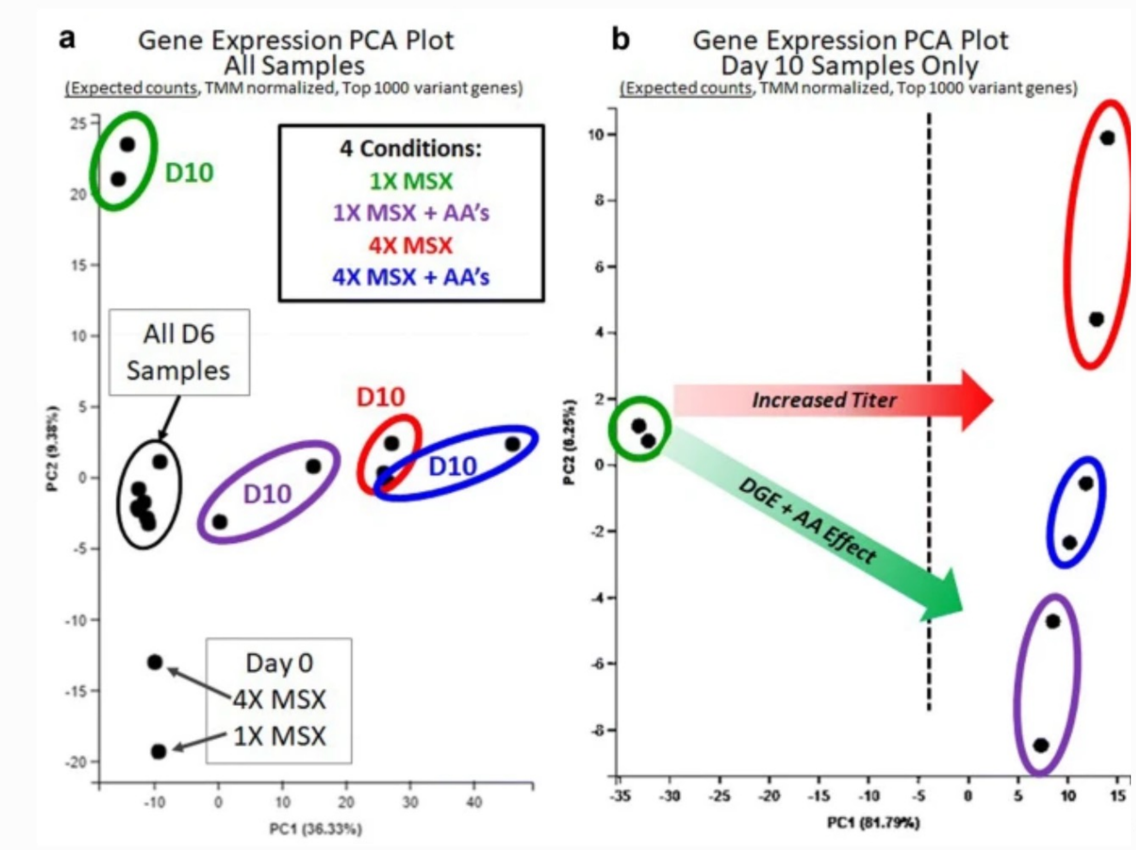


C1: Bigdata analytics identify metabolic inhibitors and promoters for mAb productivity improvement

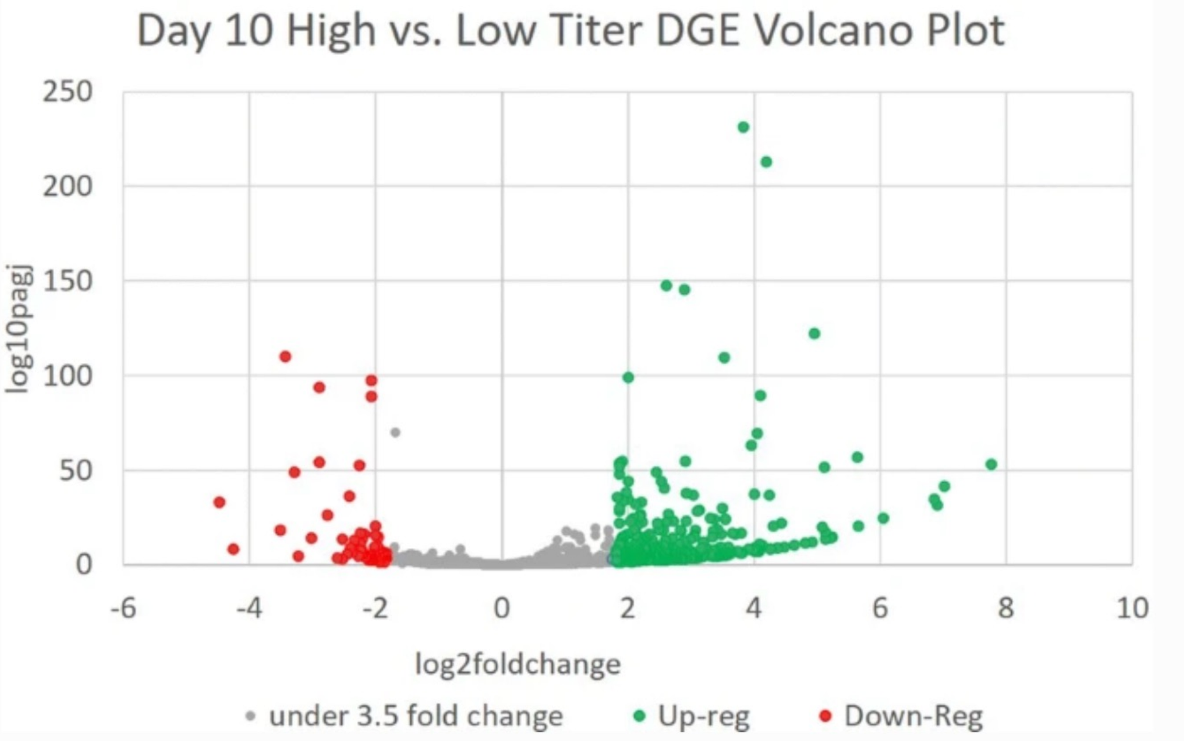




C1: Bigdata analytics identify metabolic inhibitors and promoters for mAb productivity improvement



PCA plots of **a** all bioreactor conditions at day 0, 6, and 10, and **b** day 10 samples only



Volcano plot representing differential gene expression of 204 and 205 (+AA) when compared to BR 202 and 203 (no AA). Green encircled dots in the represent the 489 total upregulated genes in V204 and V205 (with a fold change cut-off of 3.5 with p-adj < 0.005), while red encircled dots represent the 67 total downregulated genes with the same cut-off criteria

C1: Bigdata analytics identify metabolic inhibitors and promoters for mAb productivity improvement

Table 4 KEGG pathway analysis summary

From: Bigdata analytics identifies metabolic inhibitors and promoters for productivity improvement and optimization of monoclonal antibody (mAb) production process

Gene	Description	Fold change (p value)
Gly, Ser, Thr metabolism		
Sdsl	N-Sulphoglucosamine sulphohydrolase 1	+1.89 (7.31E – 3)
Tyrosine metabolism		
Fadh1	Acylpyruvase FAHD1, mitochondrial	+9 (9.47E – 3)
Fah	Fumarylacetoacetase isoform X3	+5.54 (8.95E – 34)
Lysine degradation		
Echs1[a,b]	Enoyl-CoA hydratase 1	+6.08 (8.22E – 9)
Ogdh[c]	2-oxoglutarate dehydrogenase	+3.02 (1.48E – 56)
Hadh[a,b]	Hydroxyacyl-coenzyme A dehydrogenase	+1.96 (6.31E – 13)
Gcdh[a]	Glutaryl-CoA dehydrogenase	+1.28 (1.11E – 1)
TCA cycle		
Pck2	Phosphoenolpyruvate carboxykinase 2	+3.09 (3.84E – 10)
Pdhb	Pyruvate dehydrogenase	+1.75 (1.43E – 6)
Dlat	Dihydrolipoamide S-acetyltransferase	+1.13 (1.91E – 1)
Fatty acid metabolism/degradation		
Oxsm[b]	3-oxoacyl-ACP synthase	+3.44 (1.28E – 5)
Acadsb[a]	acyl-CoA dehydrogenase	+1.44 (1.02E – 3)

[a]Genes also involved in fatty acid degradation pathway

[b]Genes also involved in fatty acid metabolism

[c]Genes also involved in TCA cycle. Fold change values included here only represent differentially expressed genes in V205, the highest titer condition overall

C1: Bigdata analytics identify metabolic inhibitors and promoters for mAb productivity improvement

Outcomes

- Promoting and inhibiting metabolites and corresponding AA were identified
- Provided strategy for productivity improvement and feeding strategies
- Genes differentially upregulated in the higher condition, were shown to be involved in amino acid-related metabolic pathways as well as energy production pathways
- Validation experimentation confirmed the finding

C1: Bigdata analytics identify metabolic inhibitors and promoters for mAb productivity improvement

Brainstorming Topics

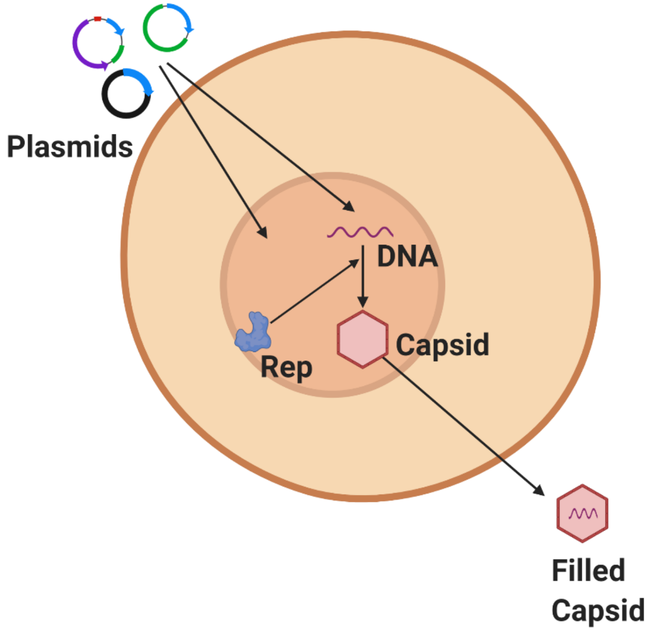
- Alternative approaches to tackle the challenges (method, data, analysis, ..)
- Expected outcomes (specific business value, intangible benefits, ..)
- Difficulties (technologies, tools, ..)

C2: Transcriptomics Studies Coupled with Medium Optimization to Address the Bottleneck in the Cellular Physiology for AAV Production.

Ref: **Biotechnology Progress, 2023**

Problems and Challenges

Low volumetric productivity of AAV generation in cell cultur limiting the number of doses that can be manufactured fro bioreactors



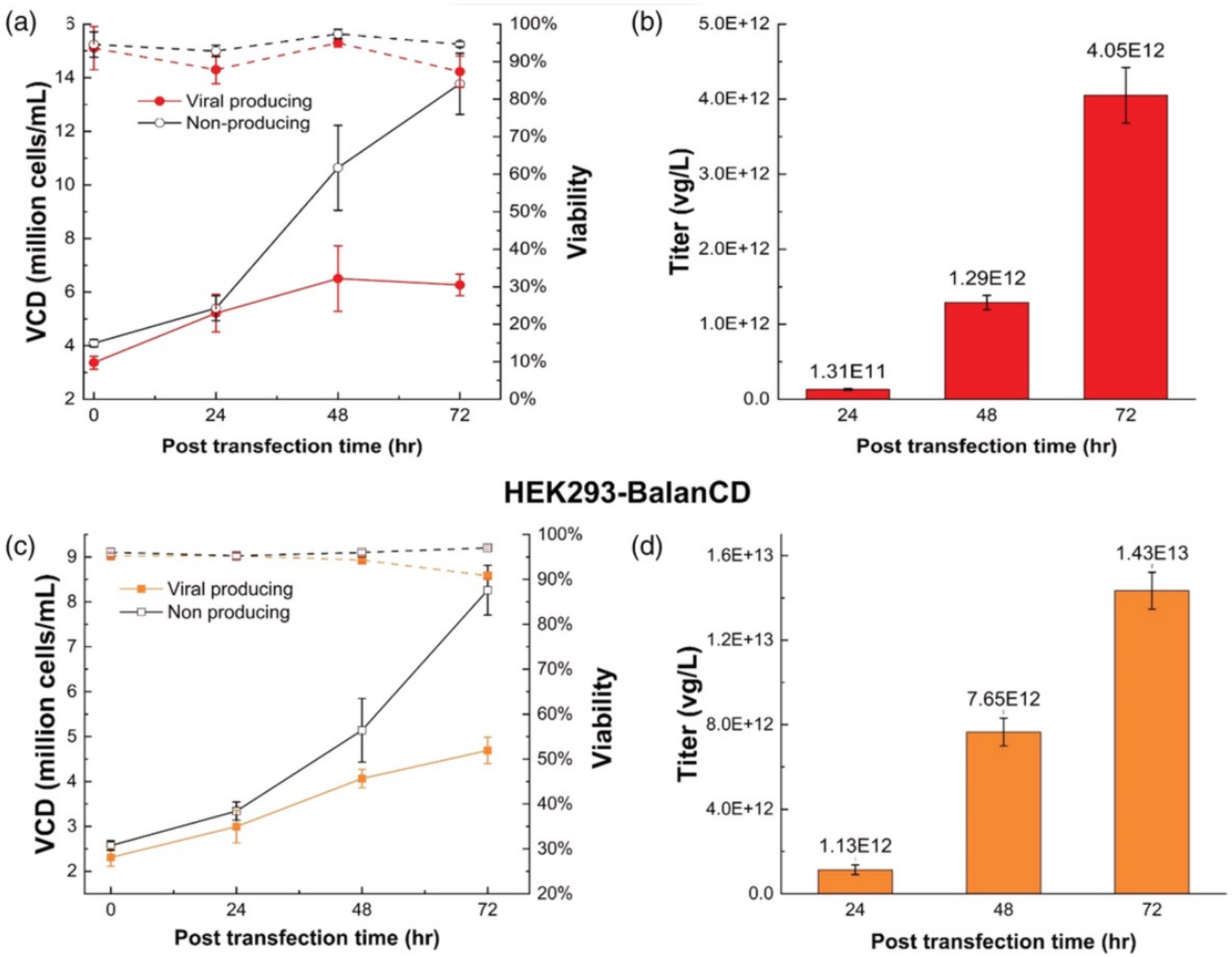
Approaches taken

Develop strategy for medium supplementation for improvir AAV production by a Transcriptomics study.

Understand the cellular features for supporting AAV production.

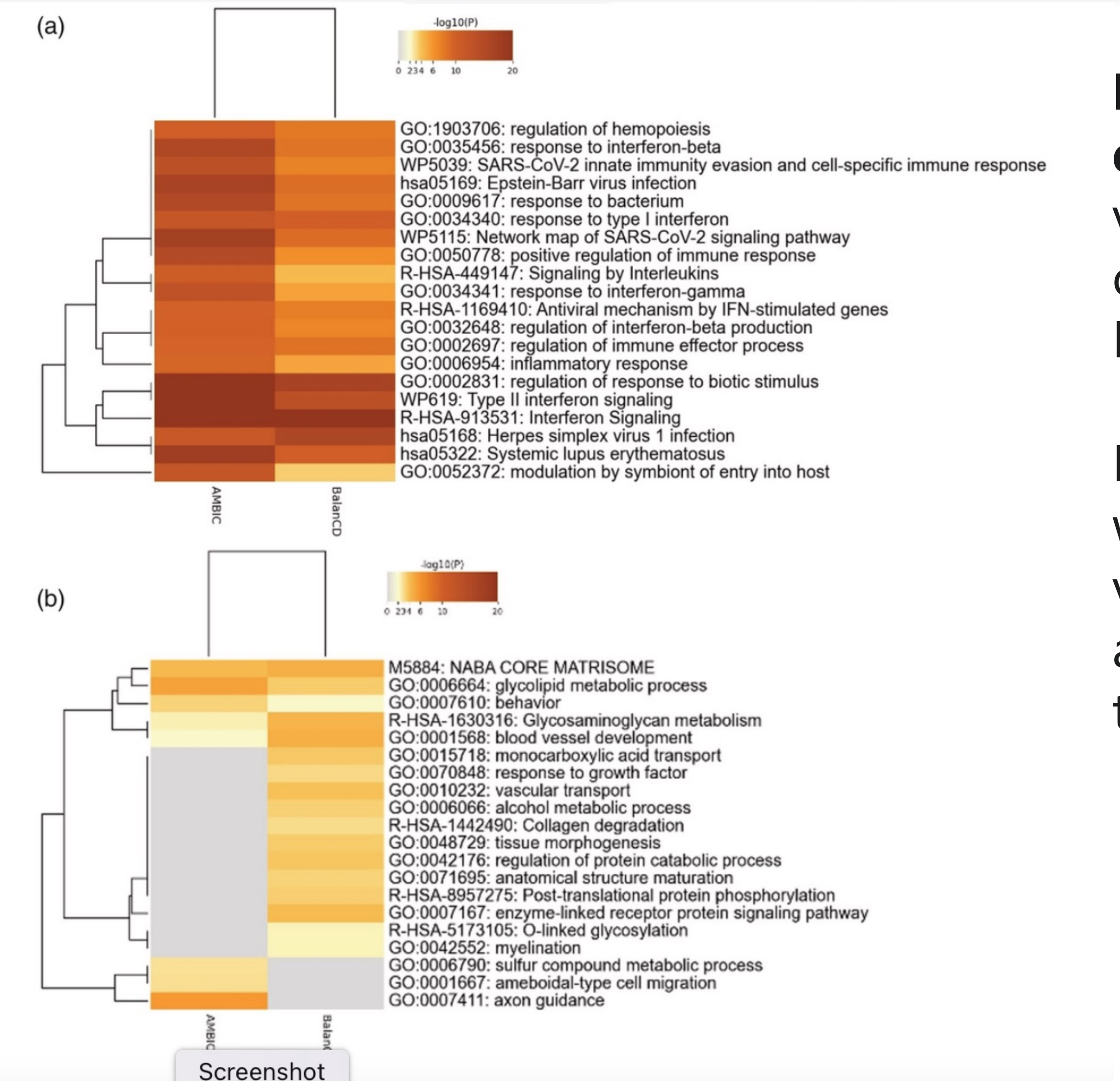
Modulating pathways associated rAAV production via medium supplements.

C2: Transcriptomics Studies Coupled with Medium Optimization to Address the Bottleneck in the Cellular Physiology for AAV Production.



Cell growth profile and genome titer for both AMBIC and BalanCD cell cultures post-transfection.

C2: Transcriptomics Studies Coupled with Medium Optimization to Address the Bottleneck in the Cellular Physiology for AAV Production.

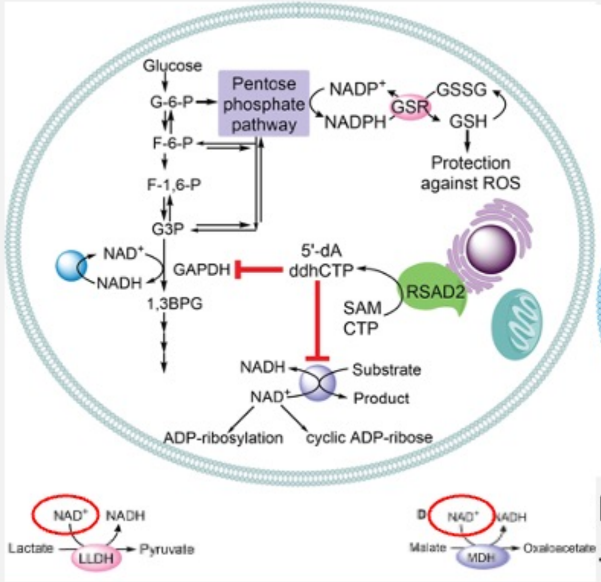


Enriched upregulated (a) and downregulated (b) clusters in the viral-producing states based on the gene ontology database for both AMBIC and BalanCD cell cultures.

Differentially expressed genes in both systems with a 1.5-fold change threshold and *p*-adj. values <0.01 were processed for ontology analysis. The figure shows the comparison of the AMBIC and BalanCD cell cultures

Table 1. Summary of identified and top ranked significantly regulated genes and pathways based on fold change and p-adj values.

Regulated pathways for viral production	Involved genes	Biological functions	Potential strategies
Antiviral immune response	**RSAD2**, OAS, IFIT	Defense response to virus and negatively regulate viral genome replication	Cell engineering
Unfold protein response	**GADD34, HSPA**	Activated due to accumulation of unfold protein and formation of protein aggregation	Cell engineering, medium supplement
Cell cycle arrest	**GADD45A, BRINP2**	Induced when cells are under stressful growth condition	Medium supplement
DNA damage response	PPP1R15A, DDIT3, DTX3L	Activated when DNA damage machinery recognizes foreign DNAs	Medium supplement
Ubiquitin-proteasome system	UBR1, UBR2	Eliminate viral protein and ensure protein quality	Cell engineering, medium supplement



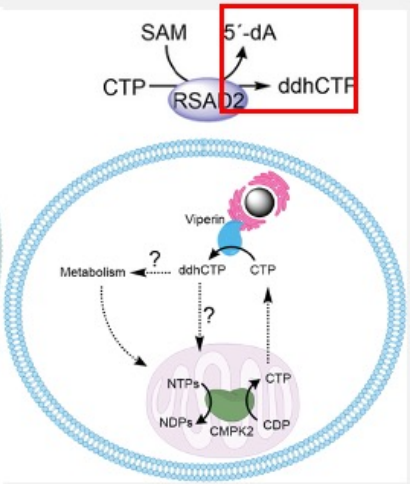


Figure 3. Anti-viral immune response - RSAD involved ddhCTP production.

RSAD2 : key enzyme involved in innate immune response

- The mixture of ddhCTP and 5'dA can **block NAD+ dependent enzymatic reaction** by binding to NAD+ binding pocket.
  - Ex: inhibition of GAPDH (glycolysis), LLDH, and MDH (TCA)
- Possible mechanism to restrict virus replication
  - Reduce **mitochondrial respiration** and **the consumption rate o amino acid** as carbon source
  - CMPK2 activity supported the formation of CTP substrates required for formation of ddhCTP

Potential strategy: knockdown (cell engineering)

Post transfection	RSAD2 FC	P-adj value	CMPK2 FC	P-adj value
Day1	51.02	6.37e-8	--	--
Day2	665.59	4.43e-31	23.93	1.067e-24
Day3	188.20	6.57e-34	17.29	1.818e-14

Figure

C2: Transcriptomics Studies Coupled with Medium Optimization to Address the Bottleneck in the Cellular Physiology for AAV Production.

Outcomes

This study compares the transcriptomes of **AAV-producing and non-producing** groups over time using different sources of parental HEK293 cells. **A transcriptomic variance** was observed.

Their transcriptomic features reveal pathways, including **innate immune responses, cell stress responses, and specific metabolisms** that potentially impact rAAV production in parental HEK293 cells.

The **antiviral immune response** is one of the most significant bottlenecks identified in viral production.

Future investigations should consider host cell metabolism for AAV production. It is also critical to understand the metabolic pathways related to viral production and the accumulation of inhibitory metabolites that restrict viral productivity.

C2: Transcriptomics Studies Coupled with Medium Optimization to Address the Bottleneck in the Cellular Physiology for AAV Production.

Brainstorming Topics

- Alternative approaches to tackle the challenges (method, data, analysis, ..)
- Expected outcomes (specific business value, intangible benefits, ..)
- Difficulties (technologies, tools, ..)

C3. Direct Control of Glycan Site Occupancy through Media Usage Optimization and Digital Twin Modeling

Problems and Challenges

Perfusion processes, in mammalian cell culture production platforms, are traditionally run with a constant fresh media supply, a cell bleed, and a harvest stream as the primary inputs and outputs of the bioreactor.

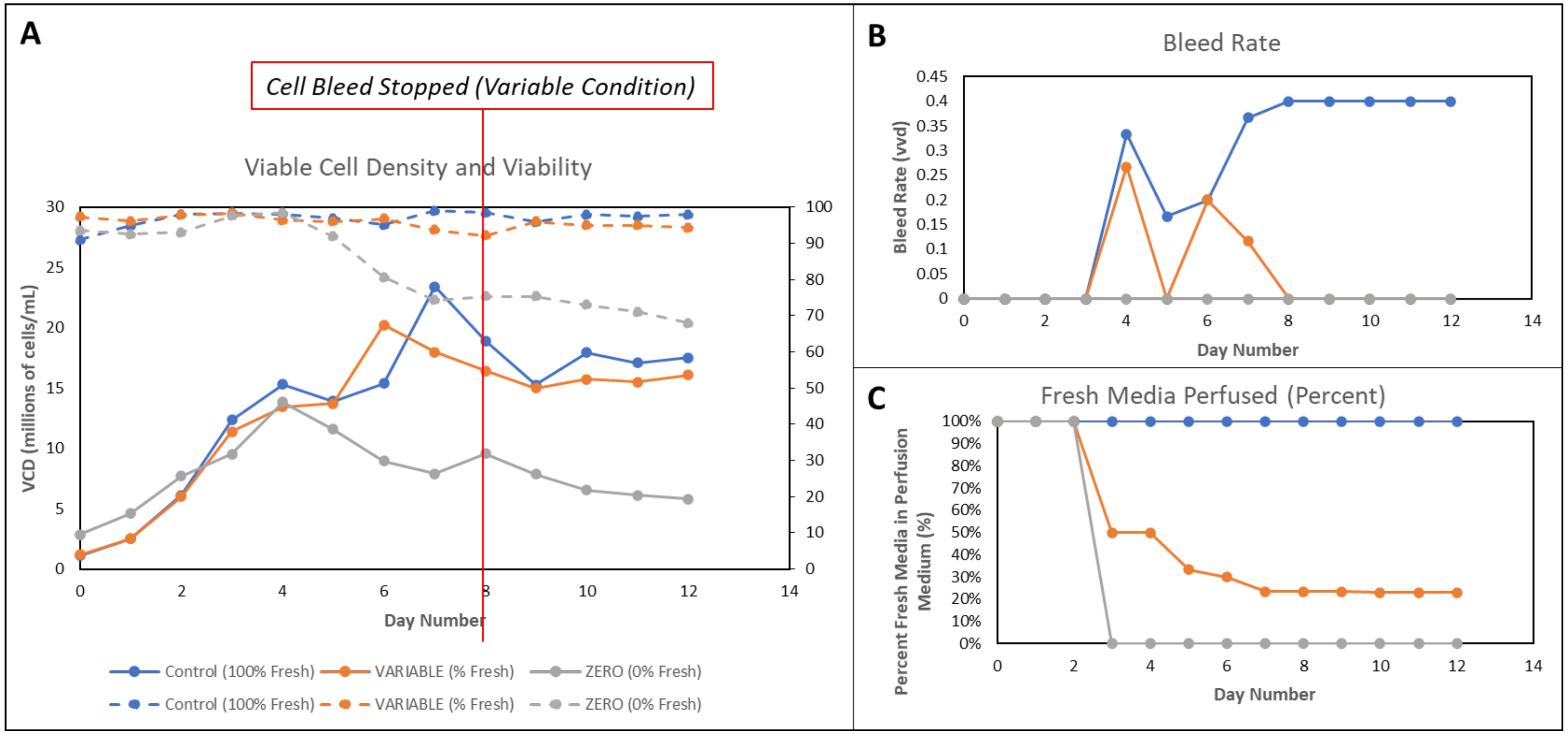
However, perfusion-based bioreactors generally **require significant amounts of cell culture media**, which can significantly increase media costs

in perfusion, **the short residence time of the media** within the bioreactor implies that components within the media are not completely consumed by cells.

Yet, the adoption of achieving steady state viable cell density within perfusion processes as a method to **control CQAs (critical quality attributes**) is not widely adopted across the biopharmaceutical industry.

C3. Direct Control of Glycan Site Occupancy through Media Usage Optimization and Digital Twin Modeling

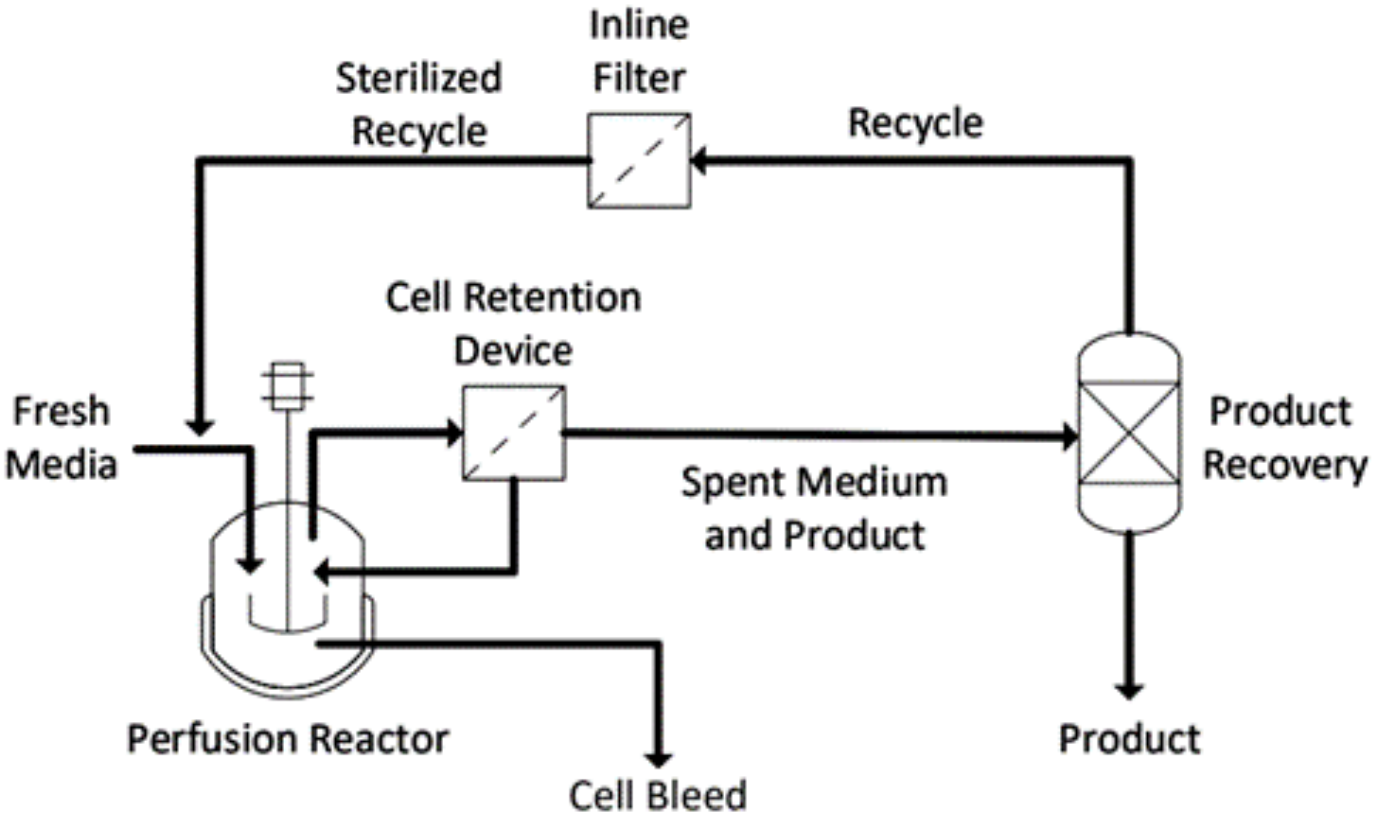
Approaches



This shake flask study demonstrated that a recirculation rate of 77% could enable for no cell bleed and steady state VCD to be achieved.

C3. Direct Control of Glycan Site Occupancy through Media Usage Optimization and Digital Twin Modeling

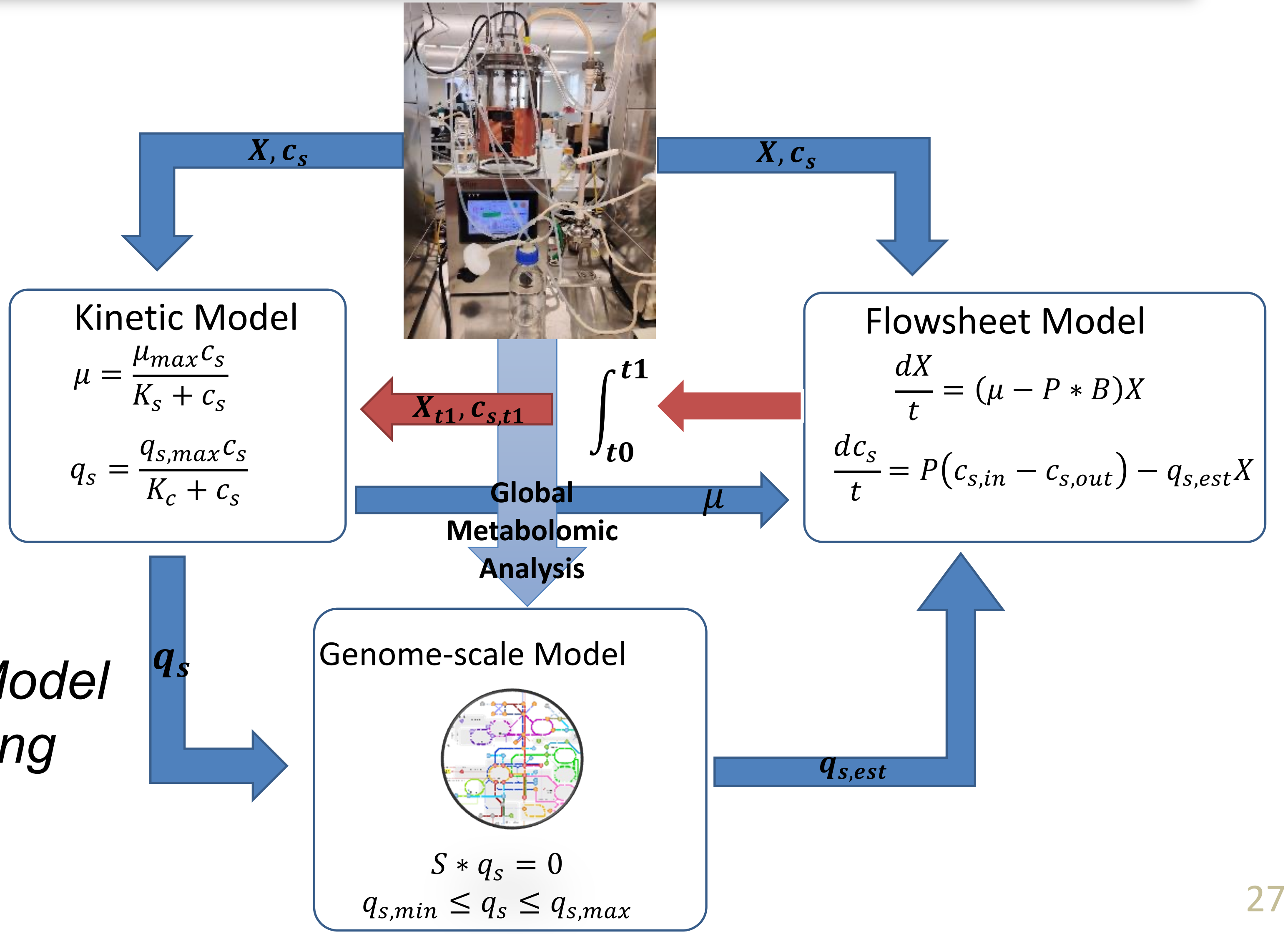
Approaches



A new integrated continuous biomanufacturing platform for continuous production of therapeutic proteins in bioreactors at fixed volumes and cell concentrations for extended periods (30 – 90 days) with immediate capture in initial chromatography and recirculation of spent media

C3. Direct Control of Glycan Site Occupancy through Media Usage Optimization and Digital Twin Modeling

Approaches



Multiscale Model

Gene level: Genome Scale Model

Cellular Level: Kinetic Modeling

Unit Operation: Flowsheeting

Online optimal/model-predictive control

Given an identified model…

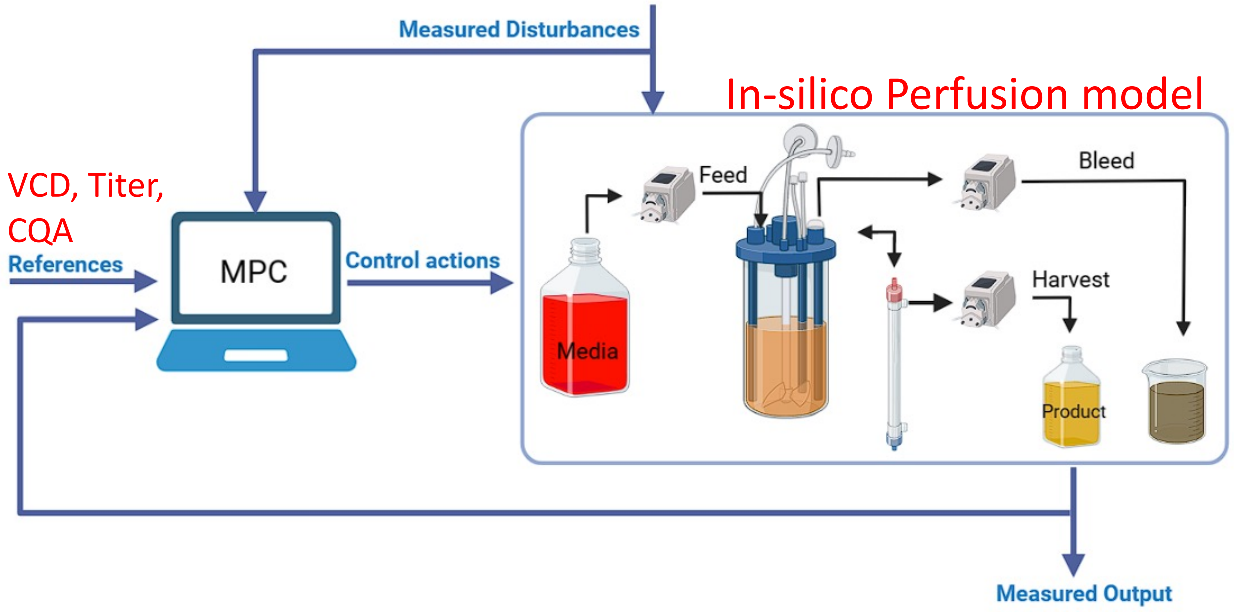
Model: $$\frac{\mathrm{d}}{\mathrm{d}t}\left[X, c_s, \begin{bmatrix}\text{Cell Growth}\\ Productivity\\ CQA\\ Toxic\ Metabolites\end{bmatrix}\right] = f\left(\begin{bmatrix}X\\ c_s\end{bmatrix}; \begin{bmatrix}\text{Perfusion Rate}\\ \text{Cell Bleed}\\ \text{Glucose Feed}\\ \text{Amino Acids Feed}\end{bmatrix}, \begin{bmatrix}\text{Cell Line}\\ \text{Inoculation Cell Density}\\ \ldots\end{bmatrix}\right)$$

Controlled Variables

Manipulated Variables

Uncontrolled Variables

… *we continuously select the next control actions which maximize the predicted value of the objective function.*



C3. Direct Control of Glycan Site Occupancy through Media Usage Optimization and Digital Twin Modeling

Outcomes (Expected)

Innovative and advanced technologies for manufacturing biologics utilizing novel platforms, process analytical technologies, and mathematical modeling.

A unique approach to optimize media usage and directly control the glycan profiles of a monoclonal antibody utilizing the recirculation rate of spent media coupled with a digital twin model

Provide further information on the process and its effects on the glycan profile of the product.

C3. Direct Control of Glycan Site Occupancy through Media Usage Optimization and Digital Twin Modeling

Brainstorming Topics

- Alternative approaches to tackle the challenges (method, data, analysis, ..)
- Expected outcomes (specific business value, intangible benefits, ..)
- Difficulties (technologies, tools, ..)

C4. Estimation of raw material performance in mammalian cell culture using near infrared spectra combined with chemometrics approaches

Ref: **Biotechnology Progress, 2012**

- **Problems and Challenges**

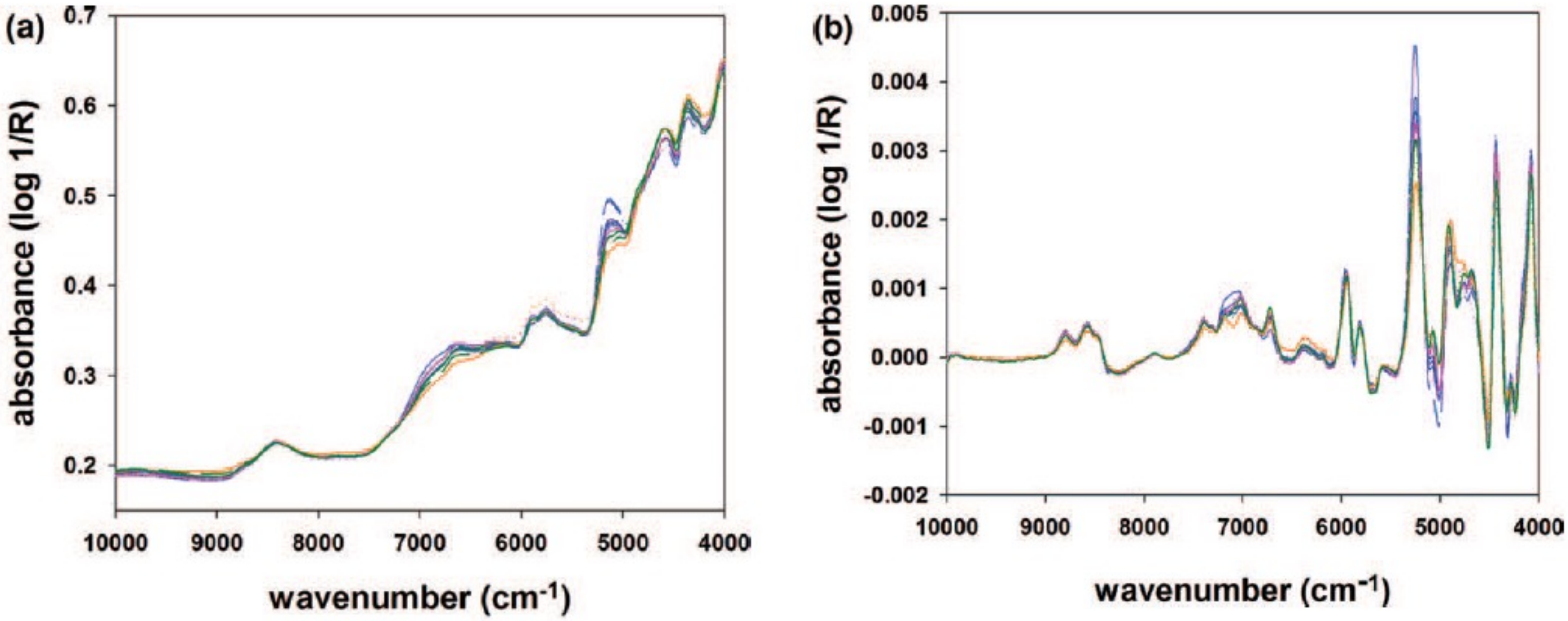
Understanding variability in raw materials and their impacts on product quality is of critical importance in the biopharmaceutical manufacturing processes

Simple, fast, and robust methods to evaluate the raw materials are necessary in order to reduce process variability and improve final product quality in mammalian cell cultures.

- **Approaches**

A comprehensive screening tool for soy hydrolysates using near-infrared spectra, with a special emphasis on the prediction of cell culture performance under the conditions of varying soy dosage and different cell lines.

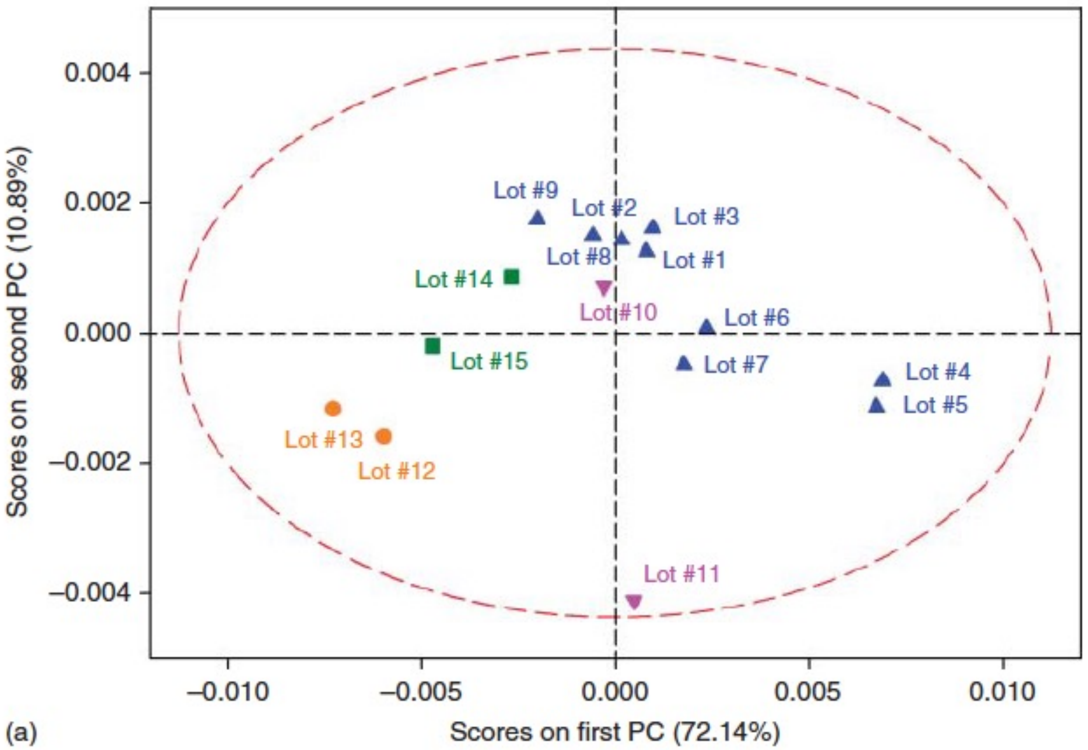
C4. Estimation of raw material performance in mammalian cell culture using near infrared spectra combined with chemometrics approaches



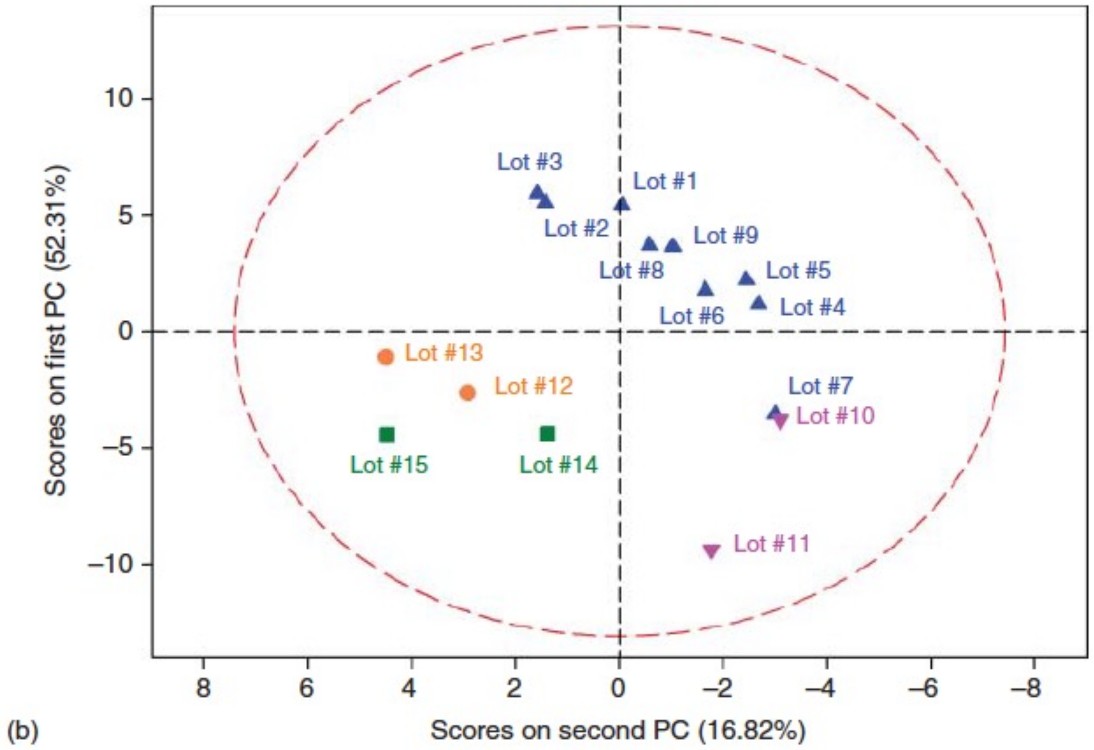
Near-infrared spectra of soy hydrolysate produced from multiple production lots and manufacturing vendors.

C4. Estimation of raw material performance in mammalian cell culture using near infrared spectra combined with chemometrics approaches

(a) Score plot of near-infrared spectra

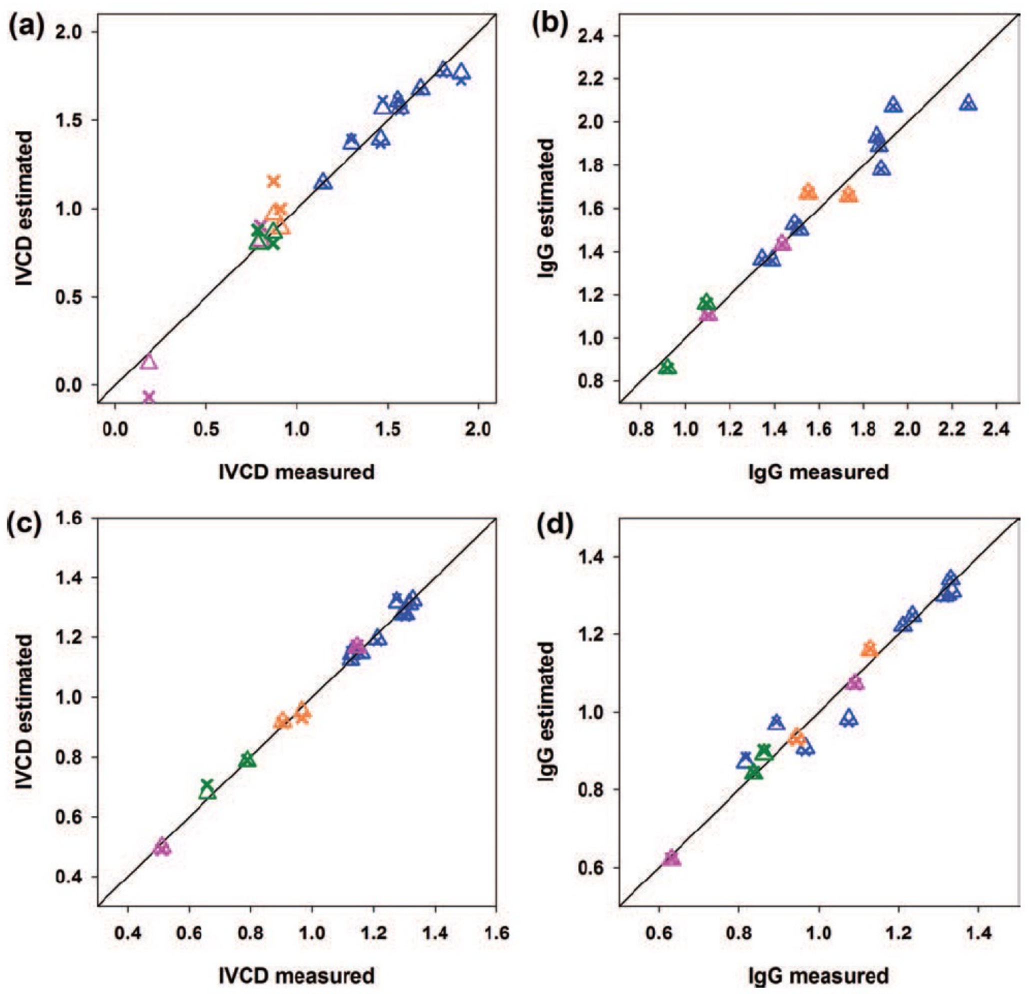


(b) Score plot of bioassay data



Ref: Lee et al. 2012, Biotechnology progress 28 (3), 824-832

C4. Estimation of raw material performance in mammalian cell culture using near infrared spectra combined with chemometrics approaches



- iVCD and igG were predicted with PLS model built with NIR spectra taken with soy hydrolysate

C4. Estimation of raw material performance in mammalian cell culture using near infrared spectra combined with chemometrics approaches

Brainstorming Topics

- Alternative approaches to tackle the challenges (method, data, analysis, ..)
- Expected outcomes (specific business value, intangible benefits, ..)
- Difficulties (technologies, tools, ..)

Brainstorming Case Studies (30 min)

- Case Study 1: Bigdata analytics identify metabolic inhibitors and promoters for mAb productivity improvement
- Case Study 2: Transcriptomics Studies Coupled with Medium Optimization to Address the Bottleneck in the Cellular Physiology for AAV Production
- Case Study 3: Direct Control of Glycan Site Occupancy through Media Usage Optimization and Digital Twin Modeling
- Case Study 4: Estimation of raw material performance in mammalian cell culture using near infrared spectra combined with chemometrics approaches

Brainstorming Debriefs (20 min)

- Case Study 1: Group 1
- Case Study 2: Group 2
- Case Study 3: Group 3
- Case Study 4: Group 4

Online Survey for Consensus

Join at
slido.com
#7386502

Online Survey for Consensus

Online Poll

Concluding Remarks

Slido Poll Results (Please see a separate file for details)

Artificial Intelligence and Machine Learning Applications in Biopharmaceutical Manufacturing

Table 1: Most popular AI-ML algorithms and their advantages, limitations, and applications

TECHNIQUE	DESCRIPTION	STRENGTHS	WEAKNESSES	APPLICATIONS IN BIOPHARMA MANUFACTURING
		Supervised Learning		
MLR Multiple Linear/Nonlinear/Logistic Regression (MLR/NLR/LR)	• Statistical technique that uses explanatory variables to predict the outcome of a response variable • Regression model is built after assuming the shape of the model space (linear, nonlinear, etc.)	• Easy/straightforward implementation • Easily updated based on data availability • Overfitting can be avoided by regularization	• Poor performance in case of non-linear relationships • Non flexible to incorporate complex pattern • Other algorithms can easily outperform	• Optimization of unit operations • Real time control of bioreactor, chromatography, membrane processes • Fault detection [17]
PLS Partial Least Squares (PLS)	• Projection method based on singular value decomposition that projects multivariable data into smaller coordinate space and then perform regression • Suitable for highly correlated data	• Ability to handle more descriptor parameters than compounds • High predictive accuracy	• Sensitive to scaling • High risk of neglecting real correlations	• Analysis of spectroscopic data (UV/FTIR/NIR/MIR) • Real time control of bioreactor, chromatography, membrane processes • Batch evolution modelling [17]
SVR Support Vector Regression (SVR)	• Supervised machine learning technique, applied for classification • Effective for high dimensionality problems with unstructured/semi structured data	• Nonlinear boundary conditions can be modelled • Overfit in high dimensionality design space can be avoided • Good performance when classes can be separable especially binary classification • Less impact of outliers	• Memory intensive and slow • Need expertise for tuning hyperparameters • Not suitable for large scale database	• Improved quantification of critical quality attributes (CQAs) from analytical tools including HPLC and LC-MS [25]
ANN Artificial Neural Networks (ANN)	• Comprises of input layer, hidden layer, and output layer • Hidden layer has weights that transform input into a quantity that can be used by output layer.	• Adaptable to many applications • Easily combined with in-silico models for training • Many hybrid approaches for integrating ANN with other techniques (PLS, SVM, etc.)	• Large amount of data required • Computationally extensive • Need expertise for tuning • Outliers affect performance	• Optimization of unit operations • Real time control of bioreactor, chromatography, membrane processes • Fault detection [17, 35-37, 40-46, 54-60, 62,63,78]
kNN k Nearest Neighbours (kNN)	• Simple and requires no assumption for data	• Simple and requires no assumption for data	• Memory intensive • Underperforms when higher dimensionality	• Quantify spatial distance between batches and deviations relative to a benchmarking batch for an industrial manufacturing process [26]
Naïve Bayes	• A classification technique based on Bayes' Theorem • Does predictions assuming that the presence of a feature is unrelated to the presence of any other feature	• Performs good in real time predictions • Easy implementation with high dimensionality data • Applicable for scale up/scale down based on dataset	• Data should represent variations well • Probability outputs are not precise • Basic assumption of independence of each feature is not true for all times	• None
Linear Discriminant Analysis (LDA)	• Method of finding a linear combination of features to	• Improves predictive performance	• Difficult to interpret new components	• Analysis of spectroscopic data from cell culture bioreactors

Ref: Rathore et al (2022), Trends in Biotech

UNSUPERVISED LEARNING				
PCA Principal Component Analysis (PCA)	• Unsupervised algorithm used for dimensionality reduction & applicable on noisy data • Based on Pearson correlation coefficient and follows the same assumptions	• Versatile • Fast and simple application	• Difficult to interpret new components • Manual tuning of threshold	• Clustering of process data for different unit operations • Scale-up and scale-down modelling • Fault detection [17]
HC Hierarchical Clustering	• Unsupervised algorithm used for partitioning objects into homogenous groups • No dimensionality reduction involved unlike PCA	• Scales well to all datasets • Does not assume globular clusters	• Number of clusters needs to be specified	• Integration of spectroscopic and bioreactor data for clustering of batches • Classification of batches to predict batch failure; trigger cleaning-in-place in manufacturing setups [28]
K-Means K-means	• Unsupervised algorithm used for partitioning objects into homogenous groups • Aims to classify a dataset into k clusters where k is fixed by determining the best locations for k number of data centroids	• Fast implementation • Simple and flexible algorithm	• Poor performance in case the underlying clusters are not globular	• Rapid prediction of facility fit issues • Integration of spectroscopic and bioreactor data for clustering of batches [30]
Density-based Spatial Clustering of Applications with Noise (DBSCAN)	• Unsupervised algorithm used for partitioning objects into homogenous groups • Groups points that are close to each other based on distance measurement	• Does not assume globular clusters • Scalable performance	• Sensitive tuning parameters	• None
Local Outlier Factor (LOF)	• Unsupervised algorithm used for partitioning objects into homogenous groups • Outliers identified based on distance measurements	• Good performance during practice • Non-linear relationships can be captured • Robust performance in case of datasets having outlier	• Unconstrained • Prone to overfitting	• None
REINFORCEMENT LEARNING				
Q-Learning	• Search-based algorithm aiming to heuristically find an optimal approach • After each step, the maximum expected future rewards are used to make the next decision	• Preserve original data features • Model-free approach	• Complex • Does not guarantee optimality of the solution	• Optimization of process flow diagram • Optimization of unit operation control to handle process variability [23, 64,66-74,76]
Genetic Algorithms	• Search-based algorithms aiming to heuristically find the optimal/ near optimal solution	• Applicable to high dimensionality dataset • Preserve original data features • Model-free approach	• Complex • Does not guarantee optimality of the solution • Multi-objective optimization requires careful inspection for convergence	• Optimization of bioreactor conditions • Optimization of media composition for bioreactor [31,32]

Ref: Rathore et al (2022), Trends in Biotech

CCE Workshop: Industry 4.0

23 - 26 Apr 2023

Poll results

Table of contents

Multiple-choice poll (Multiple answers)

What is the main driver for transforming to Industry 4.0 concepts for cell culture applications?

048

a. Improved productivity/efficiency

83 %

b. Exploring new modalities

13 %

c. Better process understanding

54 %

d. Sr. Management

4 %

e. Other

4 %

Multiple-choice poll

0 4 9

Please choose the ONE concept you would like to execute for application to cell culture development and manufacturing:
(1/2)

a. Industry 4.0 — 29 %

b. Big Data — 18 %

c. Machine Learning — 31 %

d. Artificial Intelligence — 14 %

e. Internet of things — 2 %

slido

Multiple-choice poll

Please choose the ONE concept you would like to execute for application to cell culture development and manufacturing:
(2/2)

0 4 9

f. Additive Manufacturing (eg 3D printing)

0 %

g. Autonomous Robots

2 %

h. AR/VR

2 %

i. Other

2 %

Multiple-choice poll

050

What do you see the biggest challenge in implementing Industry 4.0 concepts in your day-to-day work?
(1/2)

a. Workforce capabilities/skills
18 %

b. Management buy-in
4 %

c. Cultural aspects (mindset, behaviors)
6 %

d. Data architecture, standardization
54 %

e. Equipment integration
10 %

Multiple-choice poll

0 5 0

What do you see the biggest challenge in implementing Industry 4.0 concepts in your day-to-day work?
(2/2)

f. Regulatory constraints

0 %

g. Supplier/vendor collaborations

4 %

h. Other

4 %

Multiple-choice poll

When do you expect the Industry 4.0 transformation to be fully achieved for cell culture applications?

049

a. 0-2 years

0 %

b. 2-5 years

41 %

c. 5-10 years

33 %

d. 10+ years

27 %

slido

Open text poll

Any final comment?
(1/2)

0 2 5

- None
- What increase in process performance or reliability would make digital twin implementation valuable
- Would be interesting to see how the survey responses change over the next 2 years
- Standardized, open source ontologies
- It would be nice to take a min to discuss the poll. Why did people vote that way? The discussion is more important than the voting itself.
- Can we share these discussions in a white paper? Can we set up shared resources for attendees?
- Please share the information in the group digitally ☺
- Thanks!
- Thanks
- We are embarking on a new frontier
- Automation and control is key
- Sr Management
- Great session, looking forward to seeing how we can integrate this work faster in dev of viral vector work

Any final comment?

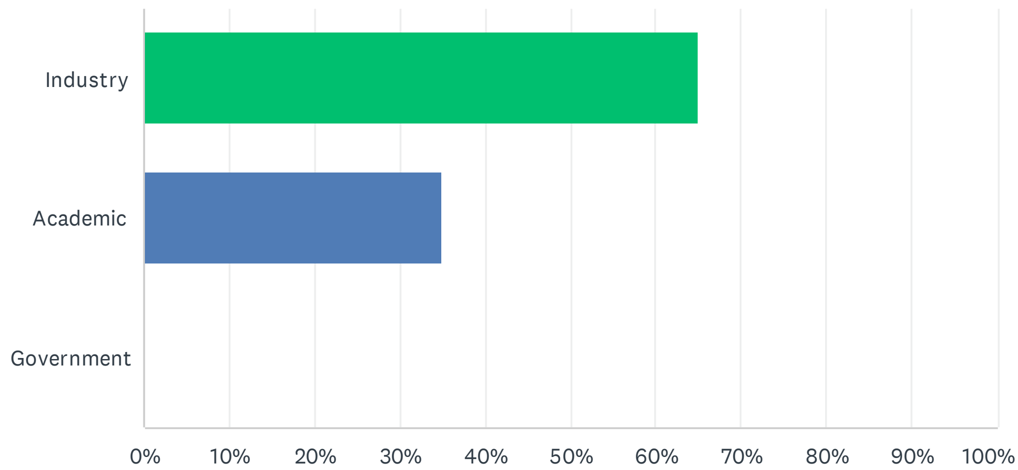
(2/2)

025

- AI
- Thanks 😊
- Better analytics and data infrastructure needed.
- Thanks for organizing
- Thanks!
- Good topic, please revisit in 2 years
- Great session.
- We need to be able to share data safely to get big data.
- Thank you for organising the workshop!
- No
- Thank you for the informative workshop.
- Digital Twins
- None

Q38 What best describes your current role (choose one)?

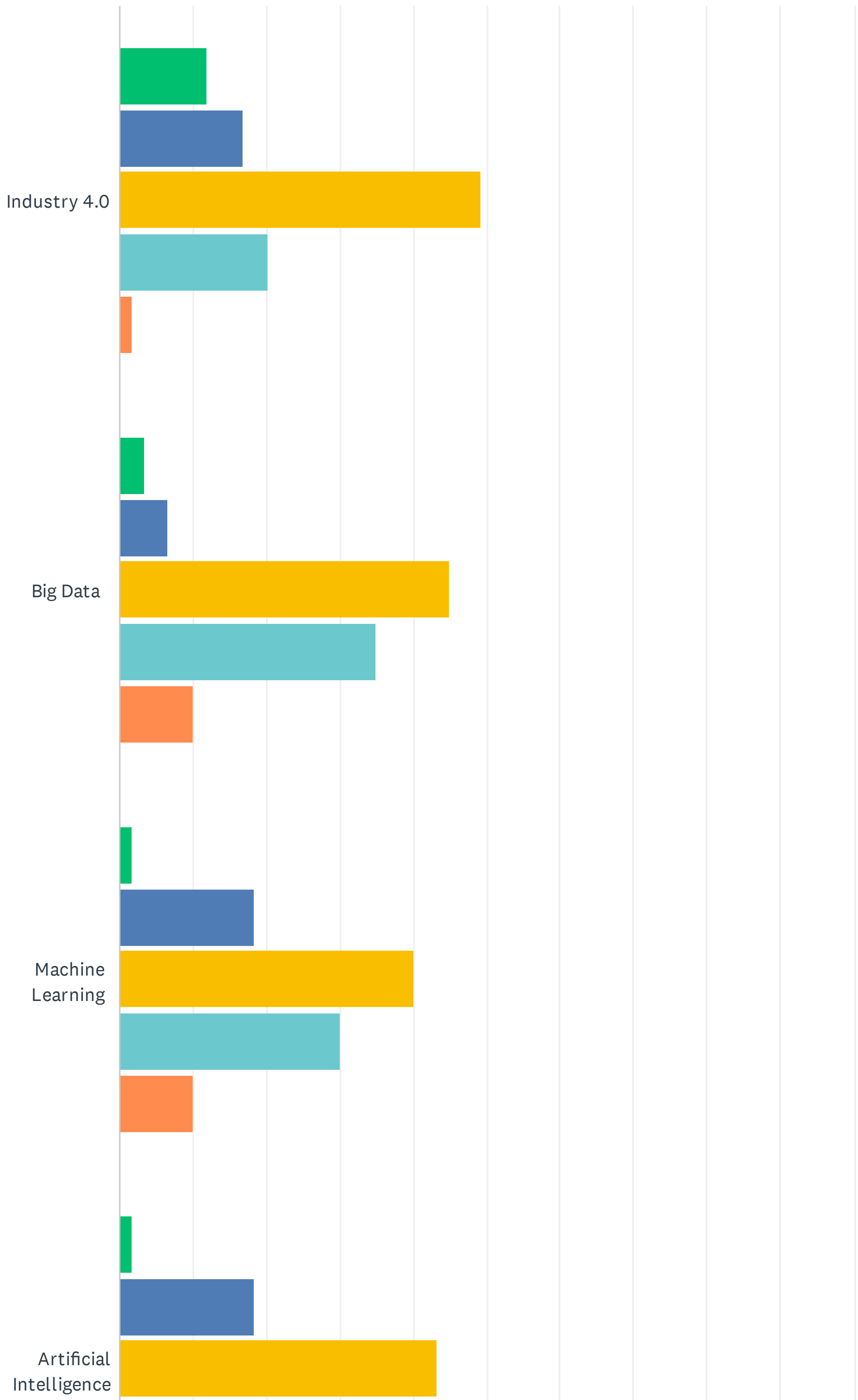
Answered: 60 Skipped: 269

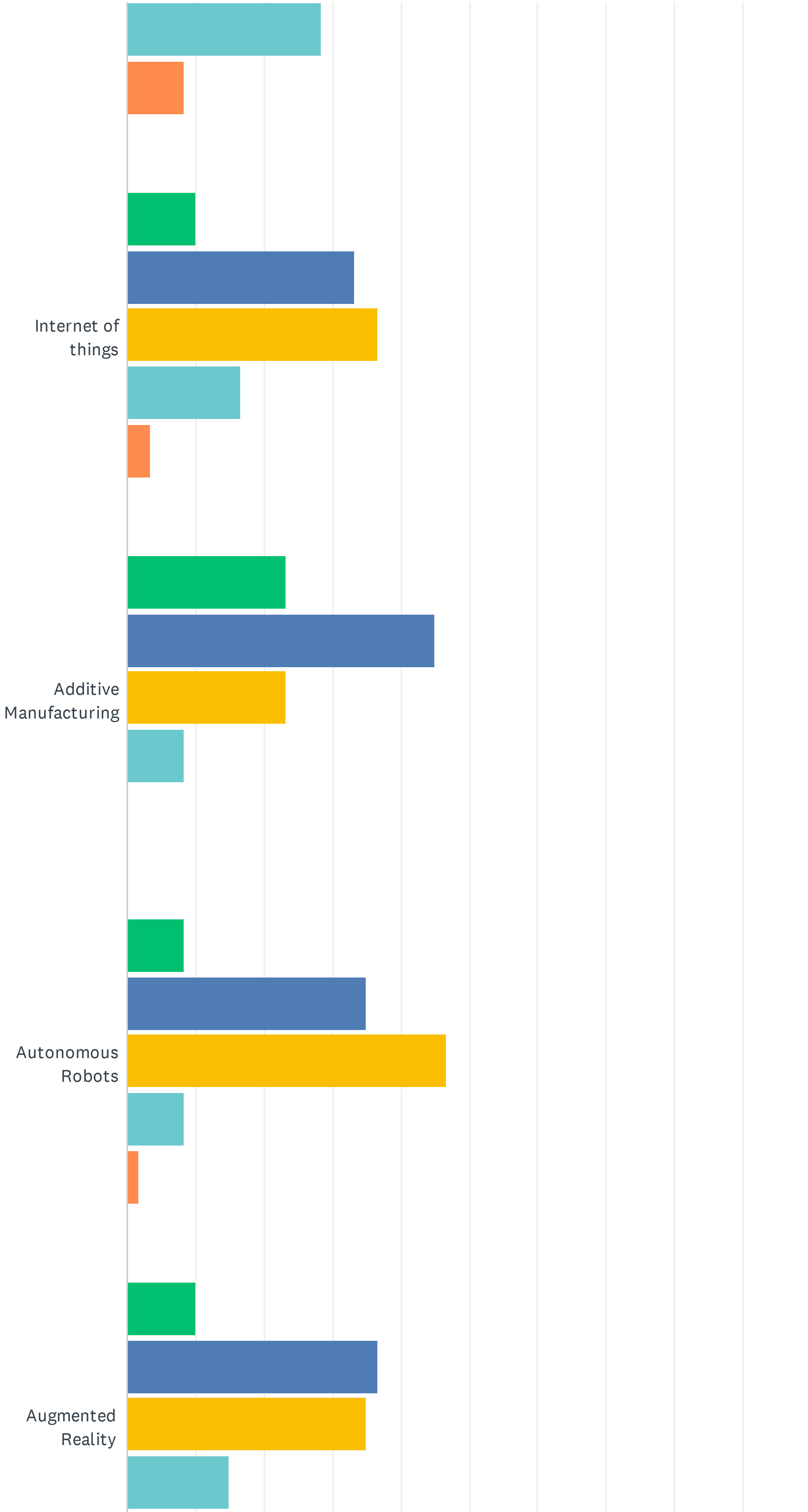


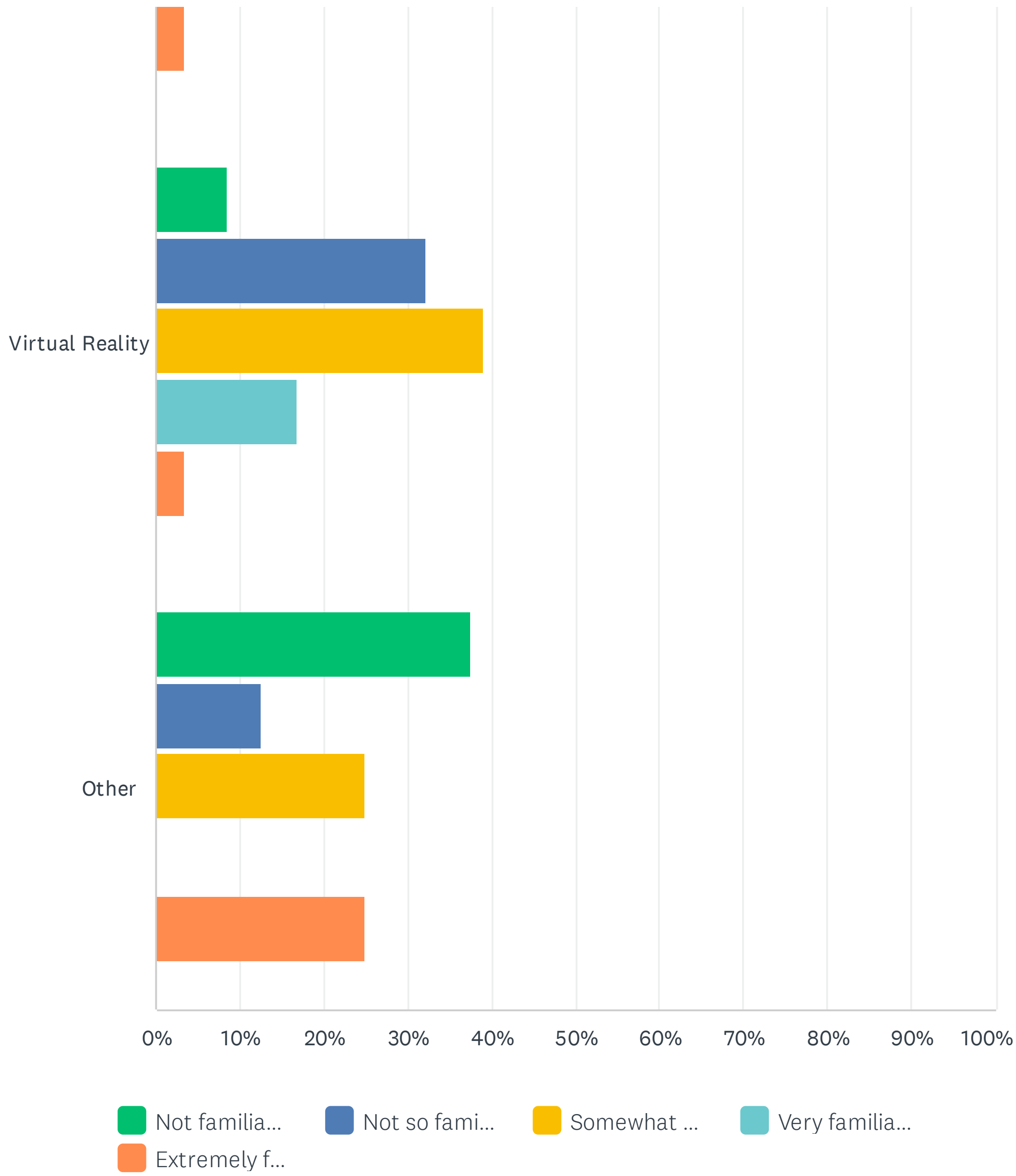
ANSWER CHOICES	RESPONSES	
Industry	65.00%	39
Academic	35.00%	21
Government	0.00%	0
TOTAL		60

Q39 Please describe your familiarity with each of the definition and concepts listed below in general

Answered: 60 Skipped: 269




Internet of things
Additive Manufacturing
Autonomous Robots
Augmented Reality

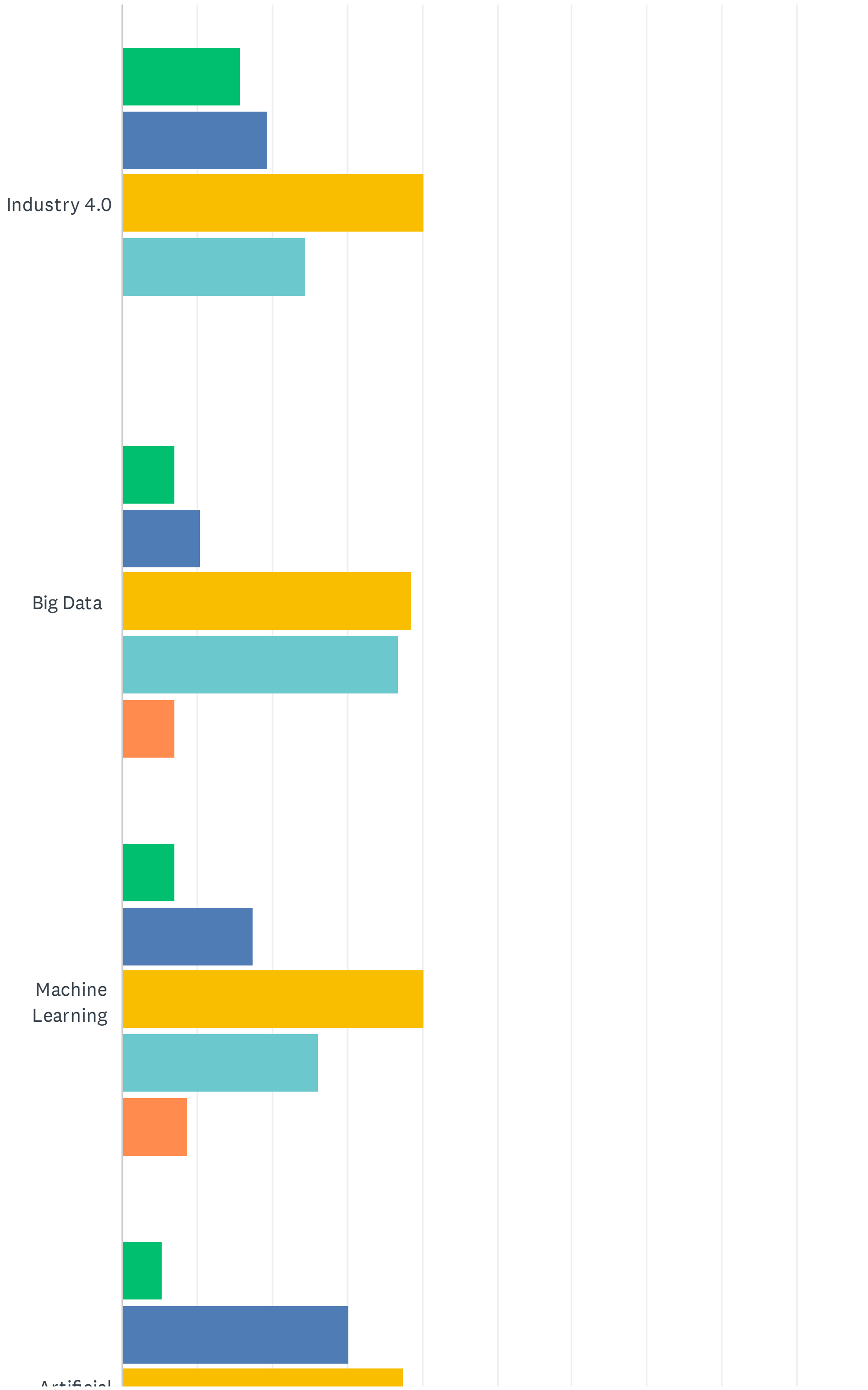

Virtual Reality
Other
0%
10%
20%
30%
40%
50%
60%
70%
80%
90%
100%
Not familia...
Not so fami...
Somewhat ...
Very familia...
Extremely f...

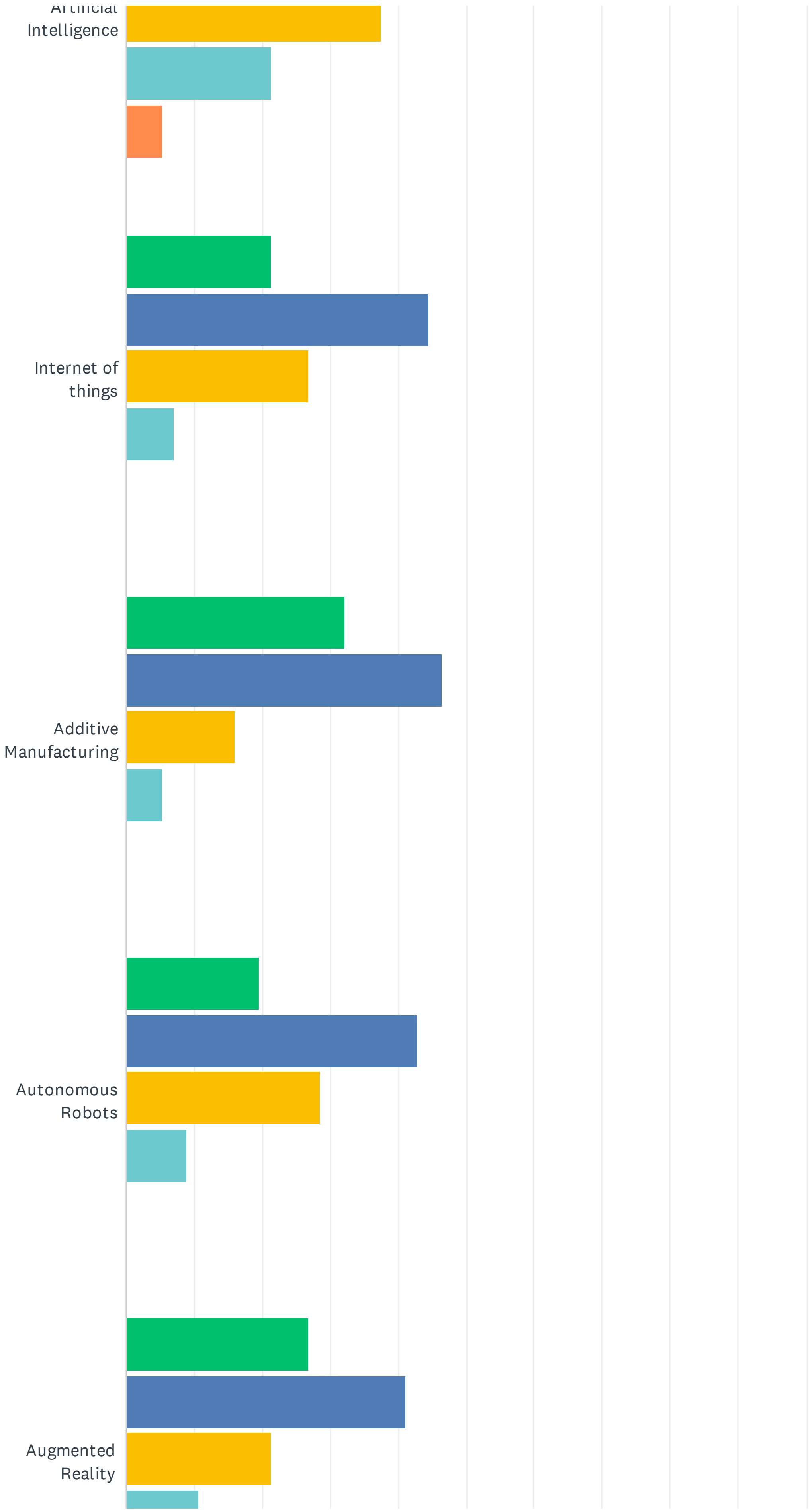
	NOT FAMILIAR [1]	NOT SO FAMILIAR [2]	SOMEWHAT FAMILIAR [3]	VERY FAMILIAR [4]	EXTREMELY FAMILIAR [5]	TOTAL
Industry 4.0	11.86% 7	16.95% 10	49.15% 29	20.34% 12	1.69% 1	59
Big Data	3.33% 2	6.67% 4	45.00% 27	35.00% 21	10.00% 6	60
Machine Learning	1.67% 1	18.33% 11	40.00% 24	30.00% 18	10.00% 6	60
Artificial Intelligence	1.67% 1	18.33% 11	43.33% 26	28.33% 17	8.33% 5	60
Internet of things	10.00% 6	33.33% 20	36.67% 22	16.67% 10	3.33% 2	60
Additive Manufacturing	23.33% 14	45.00% 27	23.33% 14	8.33% 5	0.00% 0	60
Autonomous Robots	8.33% 5	35.00% 21	46.67% 28	8.33% 5	1.67% 1	60
Augmented Reality	10.00% 6	36.67% 22	35.00% 21	15.00% 9	3.33% 2	60
Virtual Reality	8.47% 5	32.20% 19	38.98% 23	16.95% 10	3.39% 2	59
Other	37.50% 3	12.50% 1	25.00% 2	0.00% 0	25.00% 2	8

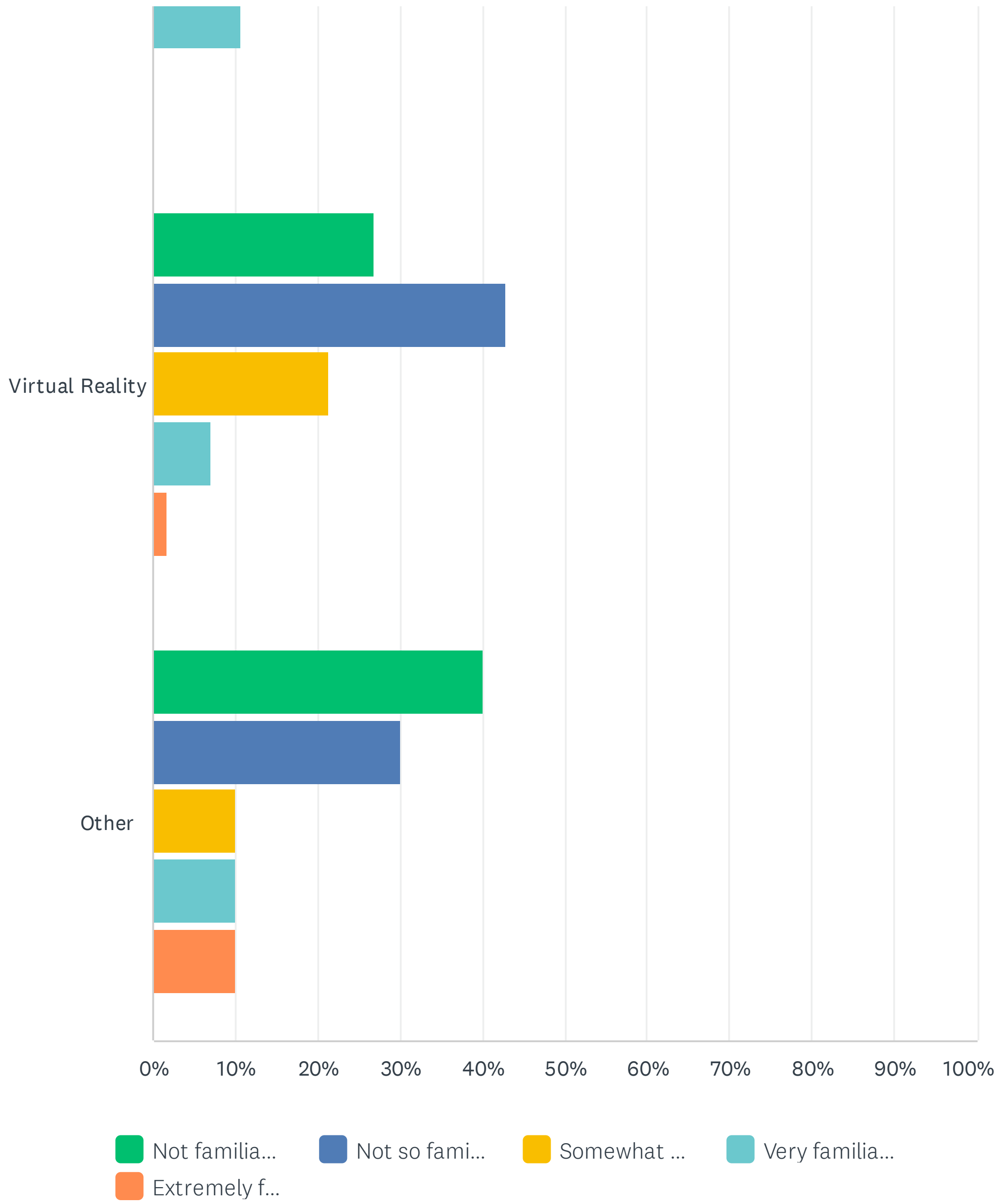
#	OTHER (PLEASE SPECIFY)	DATE
1	Mechanistic modeling approaches for cellular systems and bioprocesses	4/14/2023 4:52 AM
2	Digital twins	4/4/2023 10:09 AM

Q40 Please describe your familiarity with each of the definition and concepts listed below as they apply to cell culture technology. Please also provide the types of cell culture applications.

Answered: 57 Skipped: 272




Artificial Intelligence
Internet of things
Additive Manufacturing
Autonomous Robots
Augmented Reality


Virtual Reality
Other
0%
10%
20%
30%
40%
50%
60%
70%
80%
90%
100%
Not familia...
Not so fami...
Somewhat ...
Very familia...
Extremely f...

	NOT FAMILIAR [1]	NOT SO FAMILIAR [2]	SOMEWHAT FAMILIAR [3]	VERY FAMILIAR [4]	EXTREMELY FAMILIAR [5]	TOTAL
Industry 4.0	15.79% 9	19.30% 11	40.35% 23	24.56% 14	0.00% 0	57
Big Data	7.02% 4	10.53% 6	38.60% 22	36.84% 21	7.02% 4	57
Machine Learning	7.02% 4	17.54% 10	40.35% 23	26.32% 15	8.77% 5	57
Artificial Intelligence	5.36% 3	30.36% 17	37.50% 21	21.43% 12	5.36% 3	56
Internet of things	21.43% 12	44.64% 25	26.79% 15	7.14% 4	0.00% 0	56
Additive Manufacturing	32.14% 18	46.43% 26	16.07% 9	5.36% 3	0.00% 0	56
Autonomous Robots	19.64% 11	42.86% 24	28.57% 16	8.93% 5	0.00% 0	56
Augmented Reality	26.79% 15	41.07% 23	21.43% 12	10.71% 6	0.00% 0	56
Virtual Reality	26.79% 15	42.86% 24	21.43% 12	7.14% 4	1.79% 1	56
Other	40.00% 4	30.00% 3	10.00% 1	10.00% 1	10.00% 1	10

#	OTHER (PLEASE SPECIFY)	DATE
1	Automating workflows for bioprocess/media development	4/14/2023 4:52 AM
2	Digital twins	4/4/2023 10:09 AM

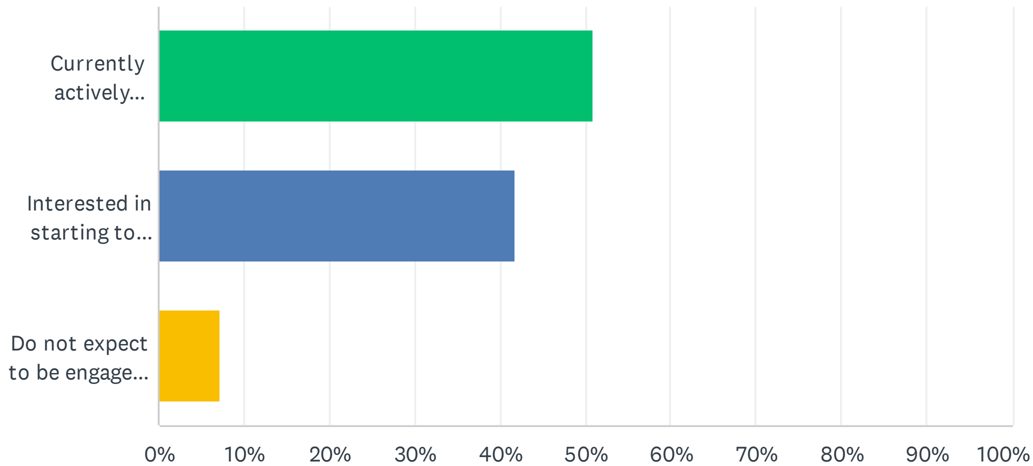
Q41 List examples of specific cell culture applications for Question 3 above:

Answered: 23 Skipped: 306

#	RESPONSES	DATE
1	* Automated workflows for bioprocess/media development (devices + analytics + software workflows)	4/14/2023 4:52 AM
2	CHO/HEK cell lines	4/13/2023 6:20 PM
3	Use of MVDA software to analyze cell culture data. Use hybrid models to build digital twins.	4/13/2023 4:39 PM
4	cell culture modeling as a digital twin	4/13/2023 4:16 PM
5	Rapid TT based on platform knowledge, multi-scale modeling, and machine learning; in the extreme one could envision going from target ID to product in vial based only on in silico predictions Adaptive control of bioreactor to enable manufacturing of consistent product quality despite variations in raw materials	4/13/2023 4:16 PM
6	Modeling N-linked glycosylation by using neural networks/dynamic kriging	4/13/2023 4:14 PM
7	Potential can use ML to predict process performance with additional genomic/proteomic data	4/13/2023 3:48 PM
8	Image analysis of single cell clone outgrowth and ML	4/12/2023 6:19 AM
9	Augmented reality as part of SOPs. Virtual reality for training. Machine learning for black and gray models. Big data from genomics and transcriptomics tools.	4/6/2023 7:05 PM
10	Hybrid modeling of bioprocess Multivariate analysis on cell culture data	4/5/2023 3:56 PM
11	Omics data analysis, model-based experimental design	4/4/2023 6:55 PM
12	Hybrid modeling to accelerate process devleopment	4/4/2023 1:34 PM
13	CAR-T cell growth in bioreactors	4/4/2023 10:09 AM
14	Our current application is limited to real time data monitoring, visualization, and potentially process adjustment based on predetermined process parameter control ranges. I am interested in learning more about other applications, particularly ML and AI.	4/3/2023 5:13 PM
15	Digital twins of cell culture processes	4/3/2023 4:54 PM
16	intelligent sensors/integrated transmitters with self-diagnosis and reporting training on SOPs using virtual reality headsets Softsensors Metabolic/hybrid models used for predictive cell culture automation	4/3/2023 3:24 PM
17	Machine learning - Supervised learning for cell culture performance prediction Machine learning - Deep learning for image-based processing of cell culture microscopy applications. Big data - Process train sensing Technlogies and data processing for supervision and control	4/3/2023 11:58 AM
18	Use of machine learning for soft sensor development. Use of exploratory data analysis to understand process impact on outcomes in early process development and for outlier detection at manufacturing scale.	4/3/2023 10:24 AM
19	Cell line development	4/3/2023 9:28 AM
20	omics data are important for characterizing clones and molecular pathways	4/3/2023 3:50 AM
21	Perfusion, continuous downstream processing. Applications to cell and Gene therapies (early)	4/2/2023 2:29 PM
22	Model predictive control, process development prediction tools, predictive clone selection	4/1/2023 2:14 PM
23	PAT- Rama process modelling	3/31/2023 6:15 PM

Q42 Considering the list of concepts in the previous questions, what is your level of engagement of Industry 4.0 topics within your current role?

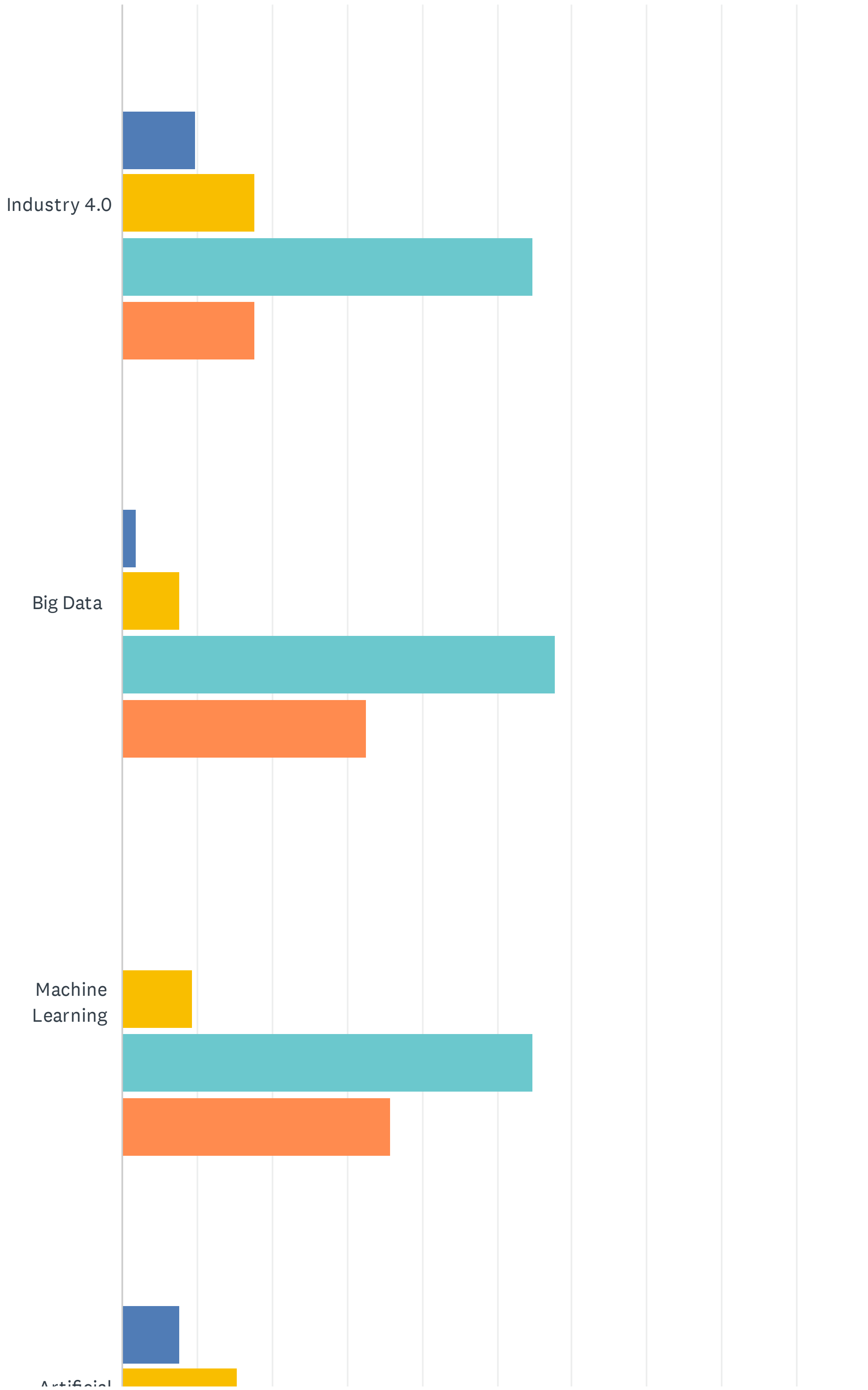
Answered: 55 Skipped: 274

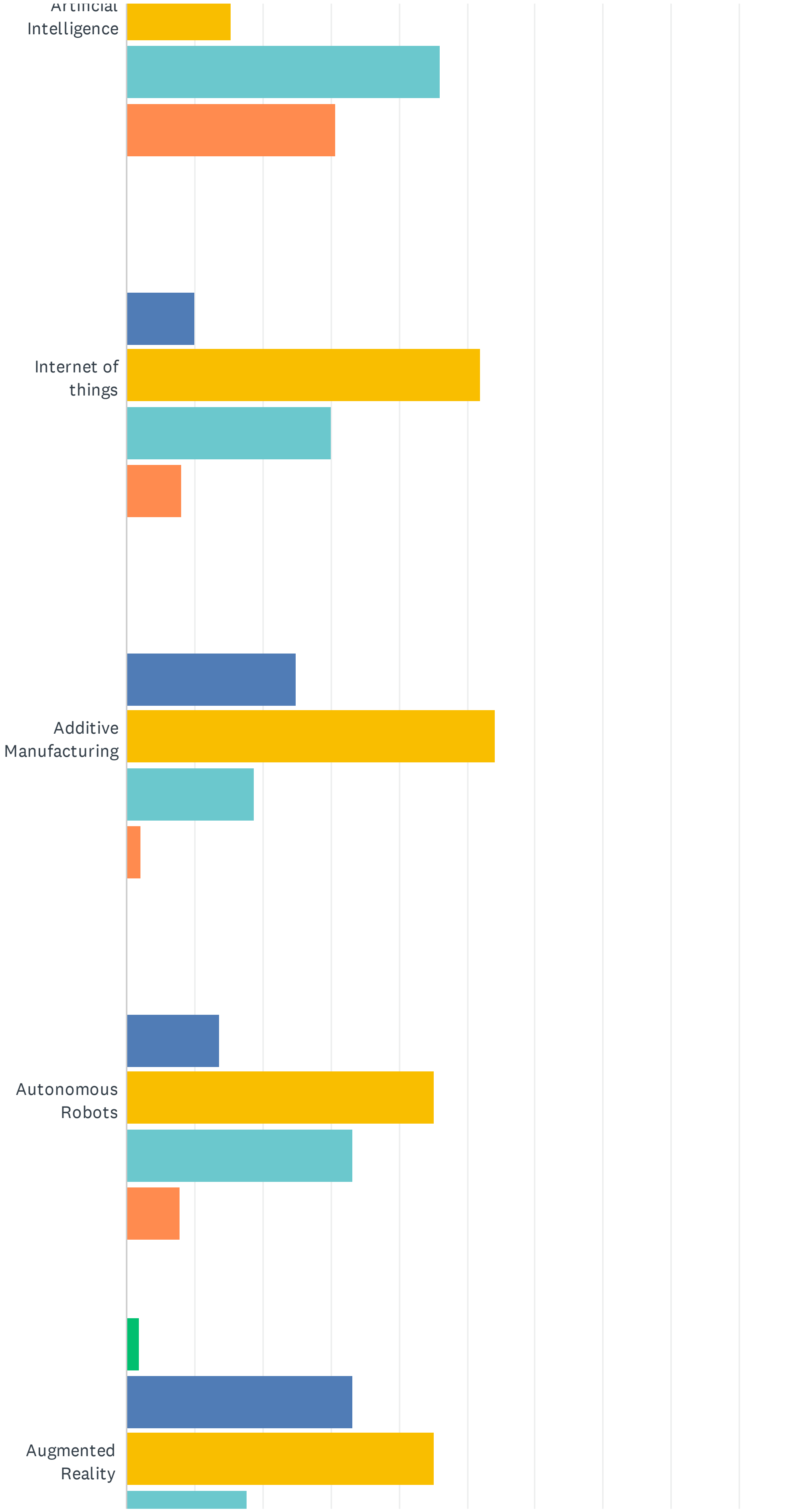


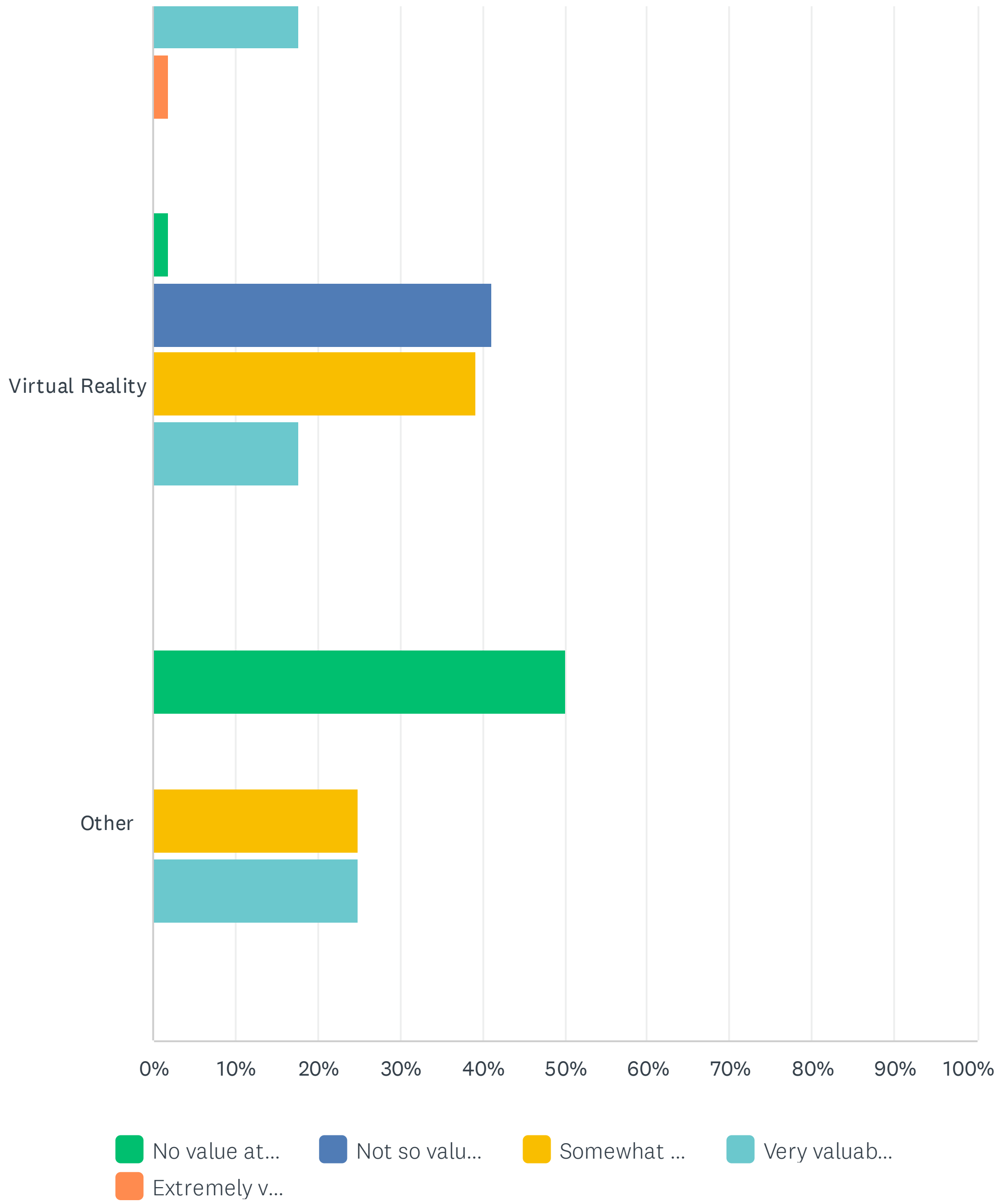
ANSWER CHOICES	RESPONSES	
Currently actively engaged in one or more of the concepts listed	50.91%	28
Interested in starting to participate in one or more of the concepts listed	41.82%	23
Do not expect to be engaged in near-term (e.g. in 2 years)	7.27%	4
TOTAL		55

Q43 On a scale of 1 to 5, how do you perceive the potential for these concepts to add value or drive solutions to common challenges in cell culture process development or manufacturing?

Answered: 53 Skipped: 276




Artificial Intelligence
Internet of things
Additive Manufacturing
Autonomous Robots
Augmented Reality

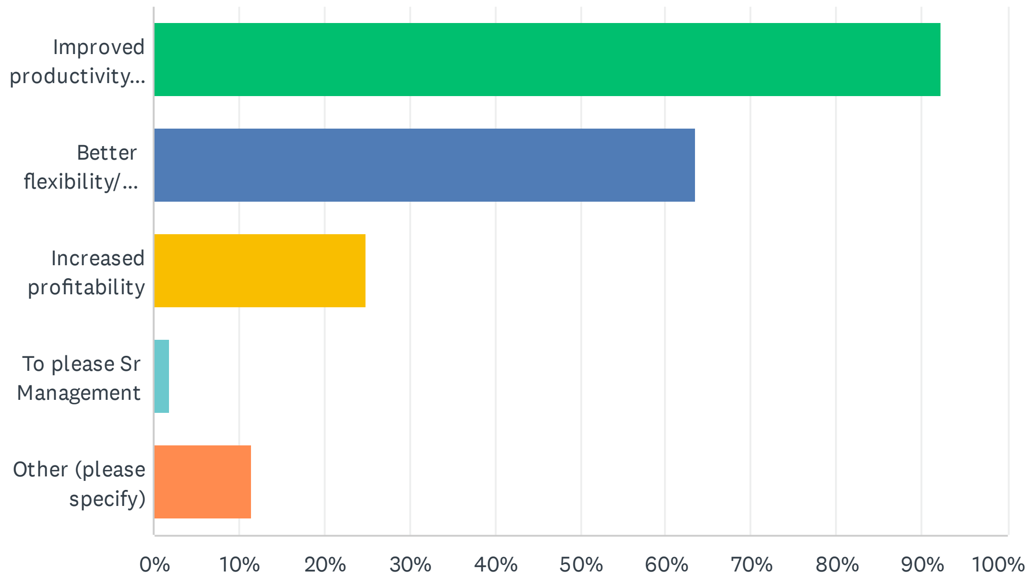

Virtual Reality
Other
0%
10%
20%
30%
40%
50%
60%
70%
80%
90%
100%
No value at...
Not so valu...
Somewhat ...
Very valuab...
Extremely v...

	NO VALUE AT ALL [1]	NOT SO VALUABLE [2]	SOMEWHAT VALUABLE [3]	VERY VALUABLE [4]	EXTREMELY VALUABLE [5]	TOTAL
Industry 4.0	0.00% 0	9.80% 5	17.65% 9	54.90% 28	17.65% 9	51
Big Data	0.00% 0	1.92% 1	7.69% 4	57.69% 30	32.69% 17	52
Machine Learning	0.00% 0	0.00% 0	9.43% 5	54.72% 29	35.85% 19	53
Artificial Intelligence	0.00% 0	7.69% 4	15.38% 8	46.15% 24	30.77% 16	52
Internet of things	0.00% 0	10.00% 5	52.00% 26	30.00% 15	8.00% 4	50
Additive Manufacturing	0.00% 0	25.00% 12	54.17% 26	18.75% 9	2.08% 1	48
Autonomous Robots	0.00% 0	13.73% 7	45.10% 23	33.33% 17	7.84% 4	51
Augmented Reality	1.96% 1	33.33% 17	45.10% 23	17.65% 9	1.96% 1	51
Virtual Reality	1.96% 1	41.18% 21	39.22% 20	17.65% 9	0.00% 0	51
Other	50.00% 2	0.00% 0	25.00% 1	25.00% 1	0.00% 0	4

#	OTHER (PLEASE SPECIFY)	DATE
1	Digital twins	4/4/2023 10:09 AM

Q44 What are the top 2 main drivers for transforming to Industry 4.0 concepts for cell culture applications (choose only two)?

Answered: 52 Skipped: 277

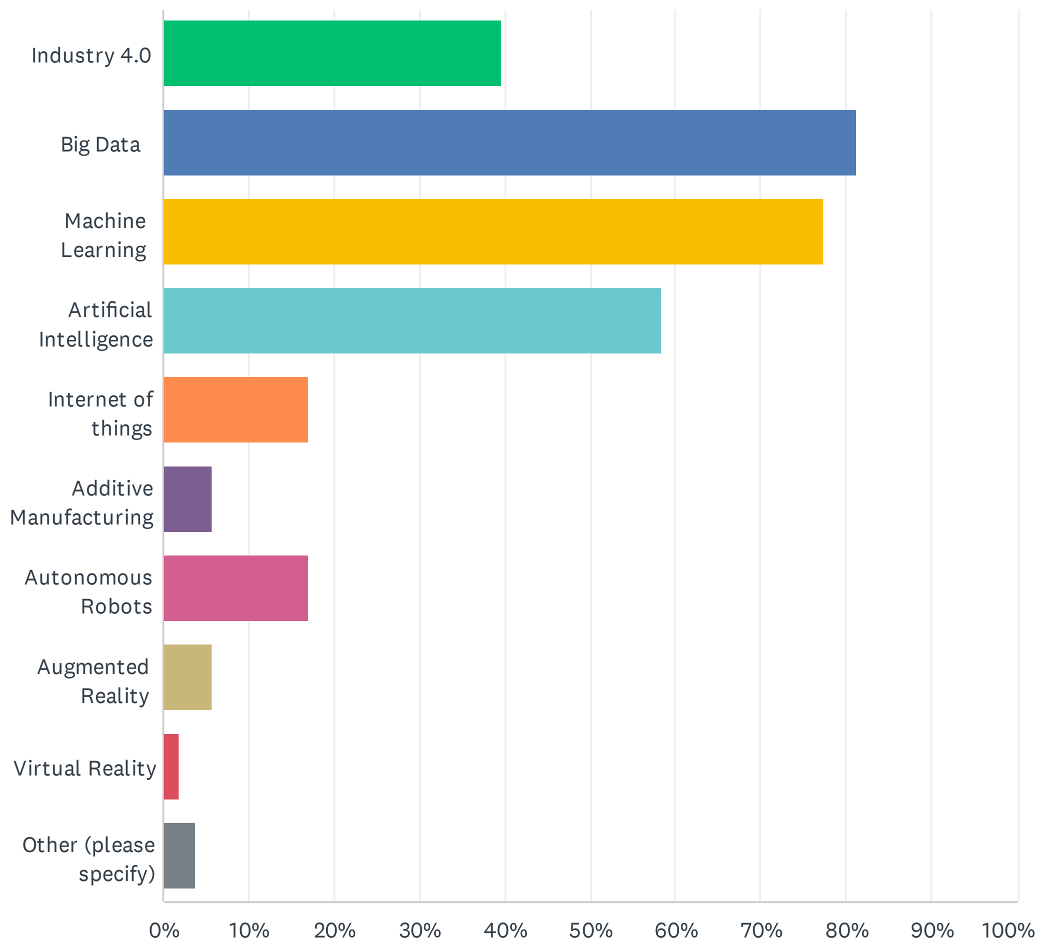


ANSWER CHOICES	RESPONSES	
Improved productivity/efficiency	92.31%	48
Better flexibility/agility	63.46%	33
Increased profitability	25.00%	13
To please Sr Management	1.92%	1
Other (please specify)	11.54%	6
Total Respondents: 52		

#	OTHER (PLEASE SPECIFY)	DATE
1	Better control of product quality	4/14/2023 10:08 AM
2	Speed to clinic	4/14/2023 5:23 AM
3	In this risk-averse industry: reduce risk of failures ;-)	4/14/2023 4:52 AM
4	Improved access to cell and gene therapies	4/4/2023 10:09 AM
5	Reduction in timelines for process development, characterization and control strategy development	4/4/2023 2:32 AM
6	Faster speed to market/failure	4/3/2023 3:24 PM

Q45 For the CCE Industry 4.0 Workshop, please choose the top three concepts you would like to see discussed in terms of their application to cell culture development and manufacturing:

Answered: 53 Skipped: 276

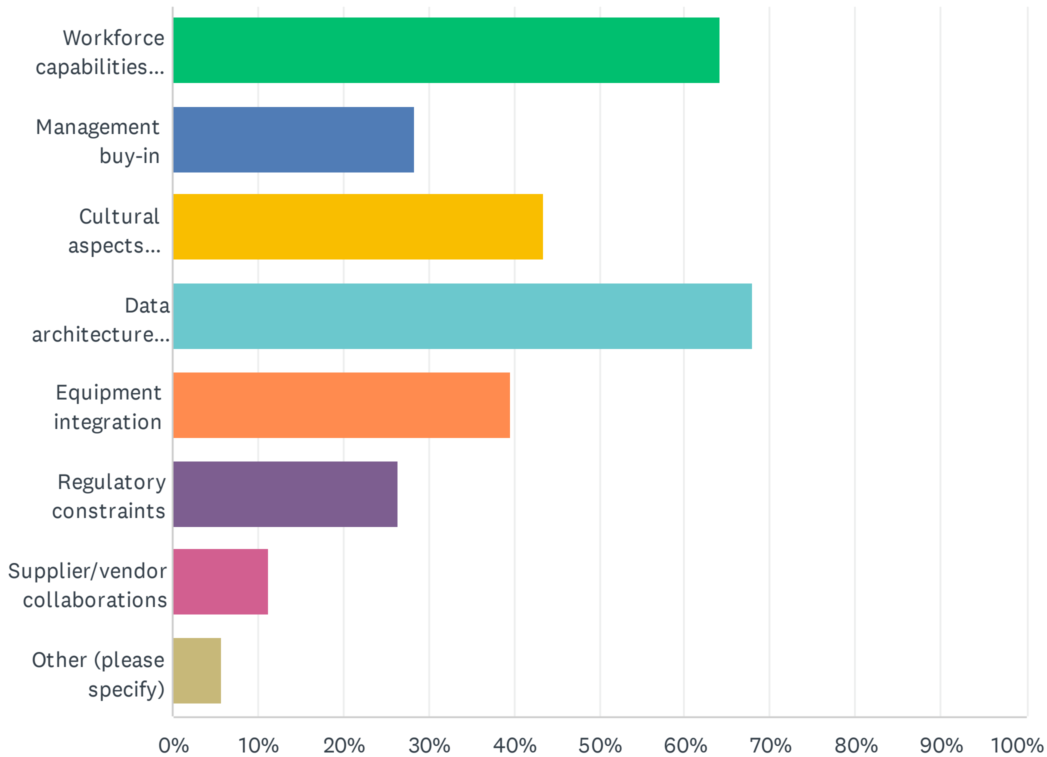


ANSWER CHOICES	RESPONSES	
Industry 4.0	39.62%	21
Big Data	81.13%	43
Machine Learning	77.36%	41
Artificial Intelligence	58.49%	31
Internet of things	16.98%	9
Additive Manufacturing	5.66%	3
Autonomous Robots	16.98%	9
Augmented Reality	5.66%	3
Virtual Reality	1.89%	1
Other (please specify)	3.77%	2
Total Respondents: 53		

#	OTHER (PLEASE SPECIFY)	DATE
1	How big can Big Data be: a) Do we know what are the right things to measure and b) would we then get enough data points to leverage the power of above approaches?	4/14/2023 4:52 AM
2	Digital twins	4/4/2023 10:09 AM

Q46 What do you see as some of the biggest challenges in implementing Industry 4.0 concepts in your day-to-day work (check top 3)?

Answered: 53 Skipped: 276



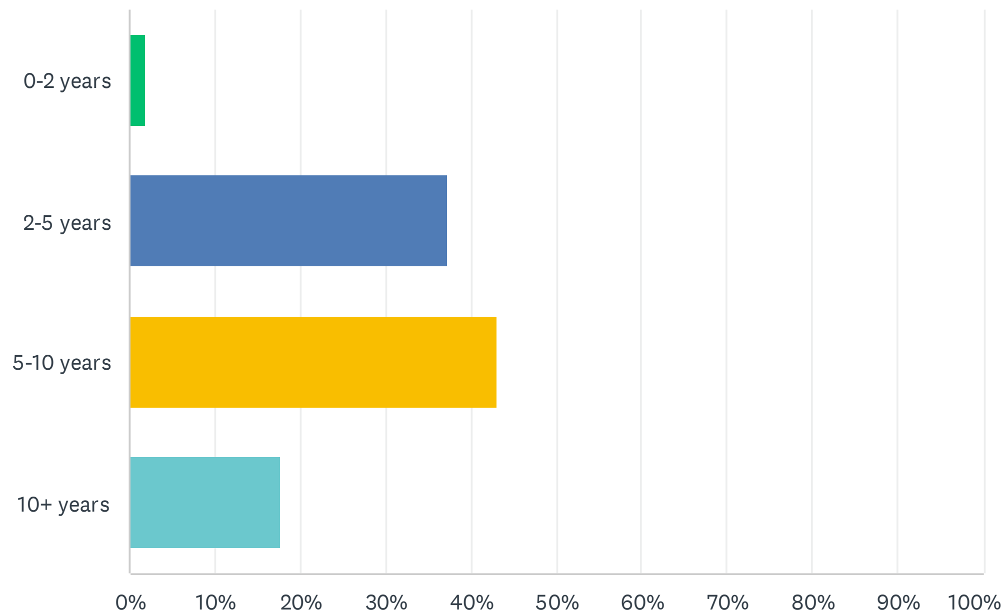
ANSWER CHOICES	RESPONSES	
Workforce capabilities/skills	64.15%	34
Management buy-in	28.30%	15
Cultural aspects (mindset, behaviors)	43.40%	23
Data architecture, standardization	67.92%	36
Equipment integration	39.62%	21
Regulatory constraints	26.42%	14
Supplier/vendor collaborations	11.32%	6
Other (please specify)	5.66%	3
Total Respondents: 53		

#	OTHER (PLEASE SPECIFY)	DATE
1	Data quality: methods are proven to work on technical systems - but currently doubting we have the right data in place to reap their benefits	4/14/2023 4:52 AM
2	Complexity of biological systems	4/10/2023 3:36 PM

3	Investment, risk of failure	3/31/2023 5:37 PM

Q47 When do you expect the Industry 4.0 transformation to be fully achieved for cell culture applications, and why did you choose this timeframe?

Answered: 51 Skipped: 278



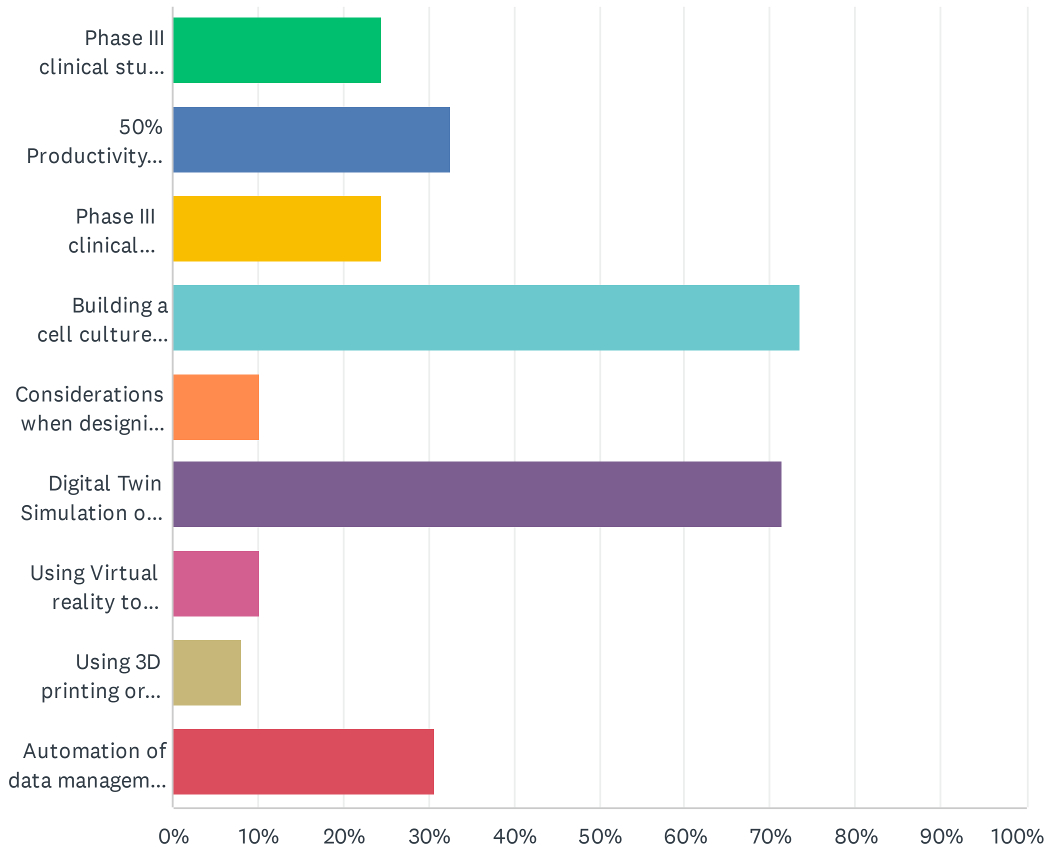
ANSWER CHOICES	RESPONSES	
0-2 years	1.96%	1
2-5 years	37.25%	19
5-10 years	43.14%	22
10+ years	17.65%	9
TOTAL		51

#	REASON (PLEASE SPECIFY)	DATE
1	Confident we can solve the technical deployment quickly, but need more time to identify and collect the right data that we need to support decision-making - so we also understand WHY an AI would recommend certain decisions - especially for sparse data availability like cell therapies	4/14/2023 4:52 AM
2	momentum	4/13/2023 4:16 PM
3	There is a true willingness of the management to go towards Industry 4.0	4/13/2023 10:49 AM
4	Access to more biomanufacturing data	4/10/2023 3:36 PM
5	Not all historical data can be used in data science. Required new experiments to collect the right data for model construction.	4/5/2023 3:56 PM
6	It takes a long time to enact change in legacy facilities. Sites need drivers to enact change and any changes need to be embedded in process development with alignment to be received by their partner manufacturing facility.	4/4/2023 1:34 PM

7	There is a long way to go.	4/4/2023 10:09 AM
8	For industry 4.0, the connectivity, sensor technology and visualization have been or can be achieved in shorter time frame perhaps, but the automated and "intelligent" process adjustment, i.e. degree of "decision making on its own", and what technology is used to enable that, that determines the complexity and timeline.	4/3/2023 5:13 PM
9	Industry is slow to adopt	4/3/2023 4:54 PM
10	mindset	4/3/2023 3:24 PM
11	Complexity of multi-scale cell culture systems in terms of cell physiology, cell metabolism, cell-process interface, data cost/throughput, etc	4/3/2023 11:58 AM
12	slow adaption of new concepts in the industry	4/3/2023 11:29 AM
13	I think machine learning models need to improve interpretability so as to be implemented at production scale.	4/3/2023 10:24 AM
14	Complexity	4/2/2023 2:29 PM
15	Capability building and regulatory acceptance	4/2/2023 11:14 AM
16	Would have chosen sooner, but lots of inertia within established systems.	4/1/2023 2:14 PM
17	already roling out	4/1/2023 3:54 AM
18	Capability still under development	3/31/2023 6:24 PM
19	Equipment integration and data architecture standardization not ready yet.	3/31/2023 6:15 PM
20	Business wants low risk and low investment. This jump to 4.0 seems like high risk and high up front investment with minimal understanding of full potential improvements	3/31/2023 5:37 PM

Q48 Which of the following case studies would you be most interested in discussing during the Industry 4.0 workshop?

Answered: 49 Skipped: 280



ANSWER CHOICES	RESPONSES	
Phase III clinical study failure due to sudden HCP increase in mAb products	24.49%	12
50% Productivity decrease in commercial campaign (as an example)	32.65%	16
Phase III clinical manufacturing failure due to potentially raw material variability (as an example)	24.49%	12
Building a cell culture model for adaptive, predictive control	73.47%	36
Considerations when designing a new facility to cater the needs of future	10.20%	5
Digital Twin Simulation of Cell Culture Processes to reduce development timelines	71.43%	35
Using Virtual reality to train engineers/scientist for GMP cell culture facility	10.20%	5
Using 3D printing or adaptive manufacturing to enhance cell culture engineering and process development	8.16%	4
Automation of data management through augmented digitization	30.61%	15
Total Respondents: 49		

Q49 What is the key learning or takeaway you would like to see from the Industry 4.0 workshop?

Answered: 24 Skipped: 305

#	RESPONSES	DATE
1	Better grasp of what others are doing	4/14/2023 5:23 AM
2	What do people think are the key limitations 1for succeeding in this field - where should we focus on solving problems in the net 5 years?	4/14/2023 4:52 AM
3	How to properly use big data	4/13/2023 6:20 PM
4	building predictive digital twins for cell culture	4/13/2023 4:16 PM
5	Examples of application of mathematical models for process intensification/optimization	4/13/2023 4:14 PM
6	Details on how other users implemented this so we have actionable things to work towards	4/13/2023 3:48 PM
7	Examples according to which using a Digital Twin or a cell culture model did improve the performance (yield or reliablility) of a cell culture process.	4/13/2023 10:49 AM
8	what is meant by Industry 4.0 and practical applications	4/6/2023 7:05 PM
9	Examples of how the industry uses Industry 4.0 in manufacturing, process/cell-line development.	4/5/2023 3:56 PM
10	Familiarize myself	4/4/2023 2:16 PM
11	A vision for how to attain Industry 4.0	4/4/2023 1:34 PM
12	How important are digital twins to industry?	4/4/2023 10:09 AM
13	I would like to see some specific examples and in particular, how ML and AI tools are used for any "decision making" processes.	4/3/2023 5:13 PM
14	Concrete examples	4/3/2023 4:54 PM
15	what is the realistic effort required for an average cell culture team to start achieving results, and how many data scientists might we need	4/3/2023 3:24 PM
16	Case study of real-world problem addressing the need, opportunity, strategy, approach, team development, deployment, challenges and opportunities.	4/3/2023 11:58 AM
17	How should we utilize early process development modeling to improve our understanding during scale up or technology transfer projects.	4/3/2023 10:24 AM
18	Steps for process automation	4/3/2023 9:28 AM
19	How to fasten implementation	4/2/2023 2:29 PM
20	Do we all have similar visions for industry 4.0 and are we all facing the same headwinds.	4/1/2023 2:14 PM
21	How to leverage current technology. Future directions to implement technology into development and manufacture.	3/31/2023 6:24 PM
22	Challenges and Road map	3/31/2023 6:15 PM
23	How to plan ahead for implementation of automation in the future. How do we develop processes now that can be updated with automation rather than needing to go through major process/equipment changes	3/31/2023 5:37 PM
24	Strategy for implementation	3/31/2023 5:03 PM