



# FROM DATA **TO DISCOVERY**



**Timothy Gardner, CEO, Riffyn**  
Biomanufacturing Workshop - February 2021  
Event sponsored by Boston University, ITIF & Fraunhofer USA

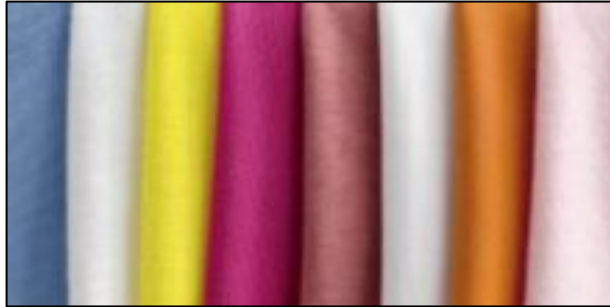
1957



Glutamate (Ajinomoto)

50 yrs | 0 genes

2006



PDO (Dupont / Tate & Lyle)

15 yrs | 6 genes

2011



Farnesene (Amyris)

6 yrs | 22 genes

1957



Glutamate (Ajinomoto)  
\$1s millions

2006



PDO (Tate & Lyle)  
\$10s millions

2011



Farnesene (Amyris)  
\$100s millions



**Escalating costs mean fewer world-changing  
biotech products will reach market**







# **Process quality**

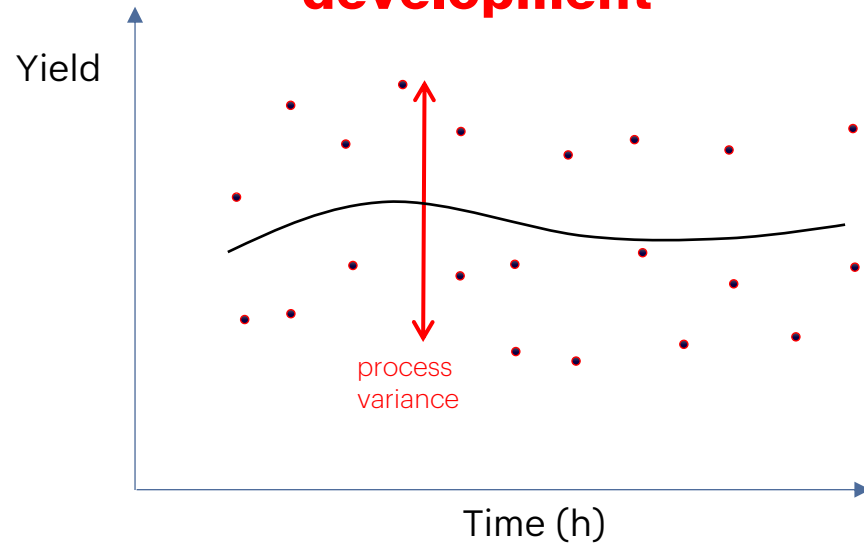
## **Aggregated data**



# Example: troubles with scale-up due to noisy R&D processes

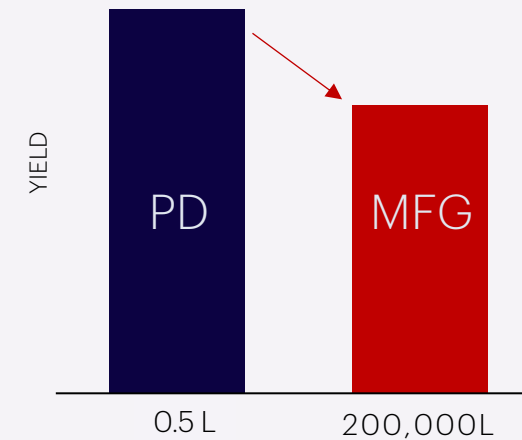
Process variation obscured understanding of true process behavior & capability

## Typical PD Batch early in development



HIGH NOISE / UNCONTROLLED PROCESS

## Scale-up outcome

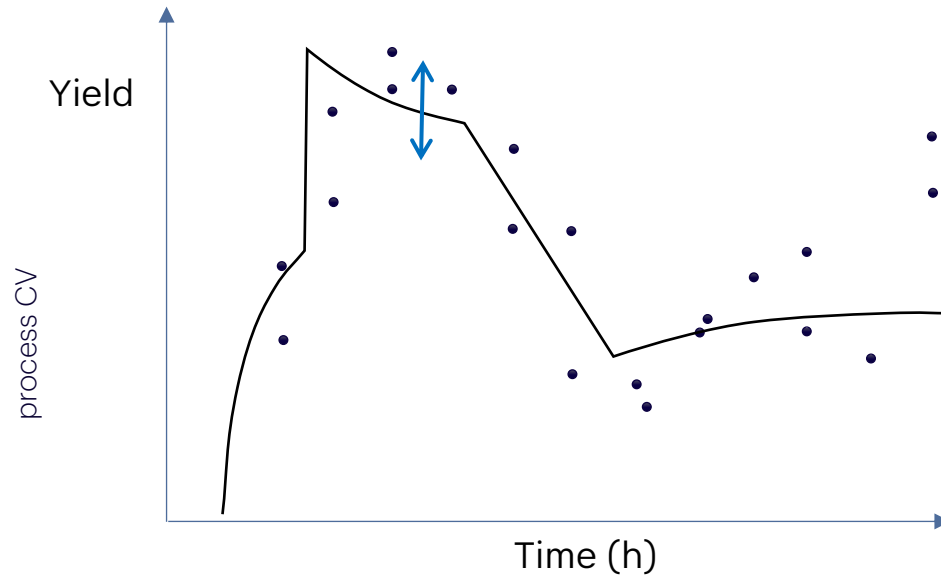


XX% yield drop  
12 months transfer time

# Solution: data integration & analytics delivered process quality

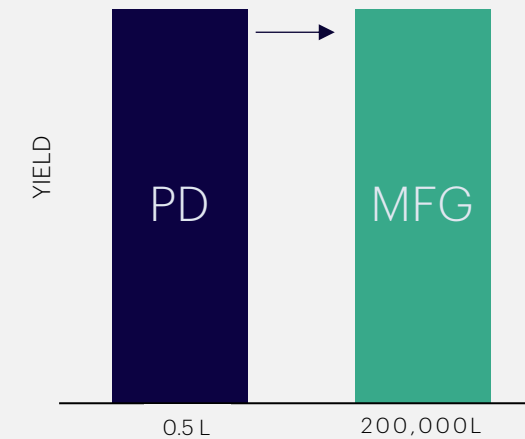
Improved data then uncovered root-causes of scaling problems

## Typical PD Batch after quality campaign



LOW NOISE / CONTROLLED PROCESS

## Scale-up outcome



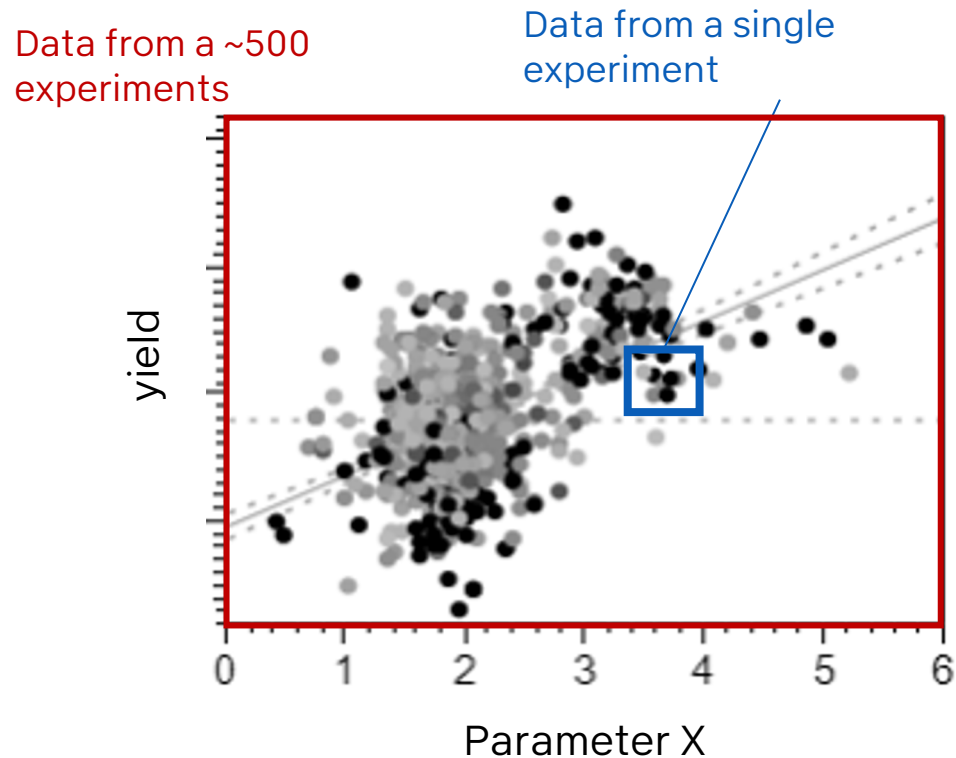
Exact performance match  
3 months transfer time



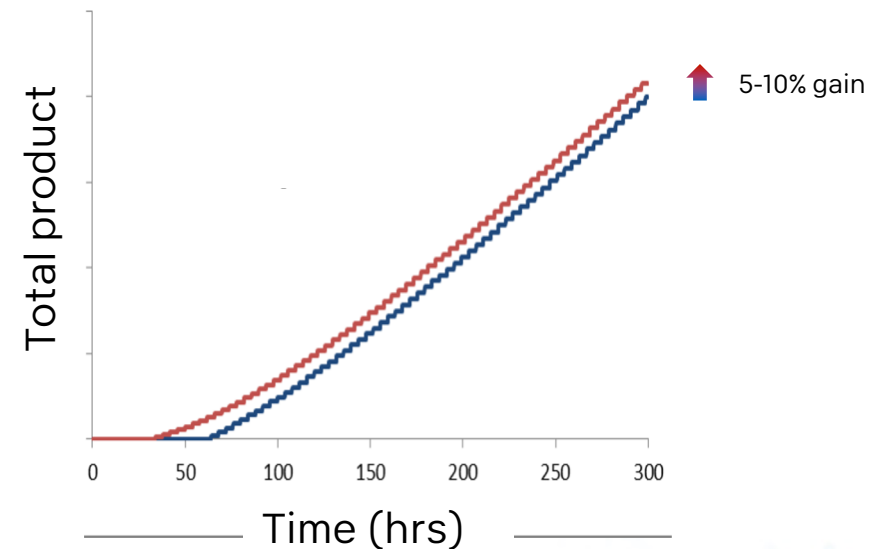
# Quality data is a building block for machine learning

Data aggregation revealed latent correlations that delivered cost-free gains in manufacturing output

3.1415



Mining of 500 aggregated experiments delivered a cost-free 5-10% gain in productivity



See *JMP Foreward* for JMP 16 for the story of “Just add water” (also posted online at [jmp.com](http://jmp.com))

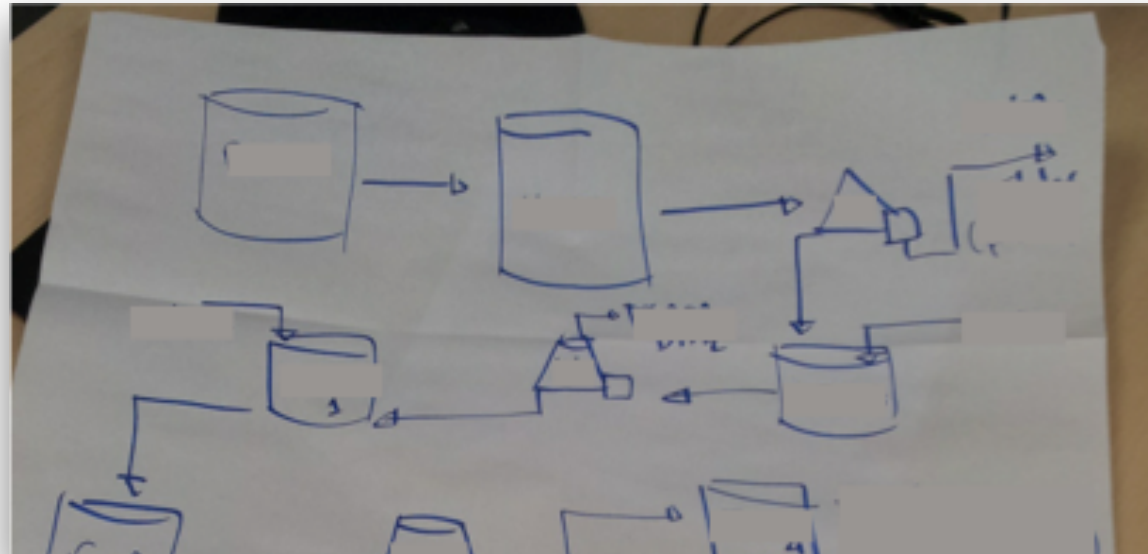
# Traditional data systems and practices are failing us

They don't deliver the **process context** or **data integration** needed for machine learning

3.1415

*"It often takes us  
2-3 months to  
assemble the  
data from a single  
development  
batch"*

- senior scientist global  
pharmaceutical company



*"We change what  
we know not  
what matters"*





# Analysis-ready data in real-time

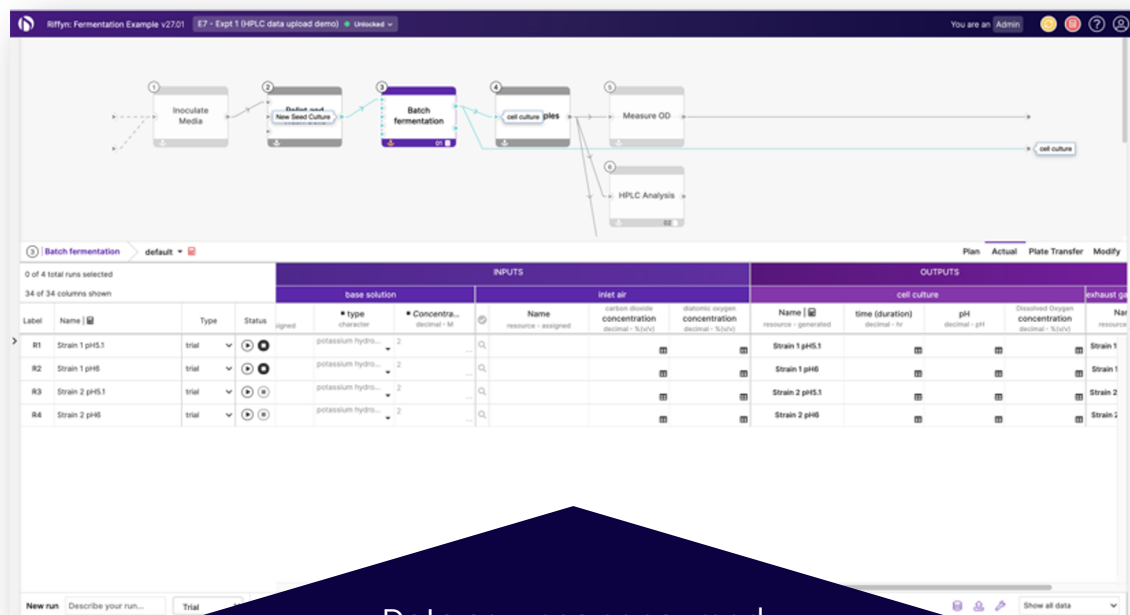
Data is contextualized on collection into a flexible, versioned process model for aggregate analysis

## 1. Process design

## 2. Execution & data capture

## 3. Data linking & contextualization

## 4. Analysis (external tools)



Data sources consumed

Excel Historian API Scripts Media

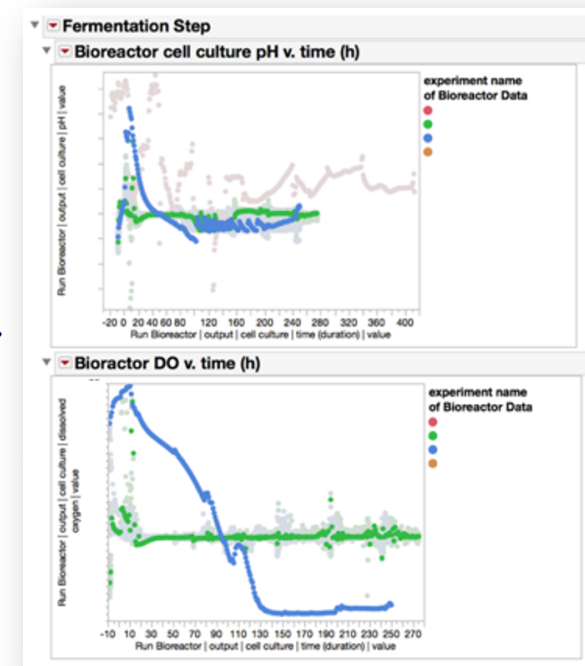
Experiment 4

Experiment 3

Experiment 2

Experiment 1

**Contextualized & aggregated data**



# Automated & instant data integration delivers insight



**Material traceability  
& process context**



**Always structured  
& integrated data**

leads to  




**Dramatic time  
savings**



**Novel insights from  
correlated data**

# How customers are using Nexus to accelerate R&D innovation

## Biofuels strain development

novozymes 

**2X faster project**

[Read the story >](#)

## Translational medicine / molecular assays

**25% less effort**

[Read the story >](#)

## Fermentation

**20X faster data prep**

[Read the story >](#)

## Clinical assays

IMPERIAL  
COLLEGE  
LONDON

**9 weeks to clinic**

[Read the story >](#)

## Solid dose formulations development

**16h saved / batch**

[Read the story >](#)

## Veterinary health

**80% fewer animals**

[Read the story >](#)



