"At and machine learning are very powerful, but if you apply those algorithms to bad data, you'll only get bad results faster."

Accelerating innovation via downscaling and complete data

Thomas Skov Senior Research Scientist in Process Automation DSK.2020 – 5-6 November 2020



Agenda of today





We are

- A global supplier of bioscience-based ingredients to the food, health, pharmaceutical and agricultural industries
- We mainly produce cultures and dairy enzymes and probiotics
- Our leading market positions stem from innovative products and production processes, long-term customer relationships and intellectual property





Global reach





We are located in +30 countries

+4,000

Employees



We have production plants on 3 continents



4

Scientific strongholds



40,000

Microbial strains

Academic partnerships

+300

+30

Represented in associations, advisory boards, etc.



5





Process Development organization

> 110 highly engaged employees representing > 15 different nationalities in Hoersholm, Nienburg and Pune:



R&D Process Development

The downscaling and robotization journey across Process Development generates new and high amounts of data \rightarrow need for *digitalization* and *automation* of the data handling



Upscaling from colony \rightarrow M³

- > Upscaling of microbes and enzymes for production across all business areas
- > Approval rate of upscaled products close to 100%
- > Inoculation materials for Chr. Hansen worldwide

Why automate & downscale of core processes

- Improve efficiency & capacity
- Bring in new technologies and robotics to build new platforms
- Rethinking our workflows to match new technologies
- Focus on getting value from data

EVP R&D "...via our automation agenda we create so many data that one of the tasks in R&D is not to create more data and more experiments but actually making sense out of the data in order for the next innovations to come..."



AUTOMATION

Downscaling and robotizing

GENERATION OF DATA

DIGITALIZATION

> Digitalization of workflows > Linking data in R&D and Production science

, ing food & health

Downscaling – when upscaling!

- Functionalities of the BioLector Pro machine
 - Can be used as an exploratory fermentation platform for testing strains, media, conditions etc.
 - ✓ Via on-line monitoring upscaling can be faster and more efficient
 - Can be started via protocols from external software
 - ✓ Can interact with robotics
 - ✓ Can measure many parameters real-time





Dissolved oxygen Biomass & fluorescence pH value Micro-valves



Microfluidic Control on a FlowerPlate® with Optodes

TECAN integrated into the BioLector Platform

- TECAN protocols are designed via TrialComposer to have complete control of fermentation conditions, media composition, feeding, parameter control (e.g. pH) and data output
 - Tedious harvest criteria (e.g. harvesting at night) can be programmed via BioLector and executed by the robot
 - Experimental designs can be adjusted based on readings along the way and new fermentations started based on data output
 - ✓ Poor fermentations can be stopped before time according to set criteria
 - Preparations and treatment of samples for subsequent use of other technologies can be started (e.g. microwell plate for microscopy)







Data are huge...



And will be even bigger...

- Automated microscopy
- ► NIR
- ► IR
- Capacitance (Permittivity, Dielectric spectroscopy)
- ▶ ???













Metadata



Metadata means "Data about Data"

Metadata is the key to unlocking the value of our data

Metadata is as valuable as the data itself

Data without contextual metadata is meaningless

Comprehensive metadata will drive the use of machine and deep learning in both R&D and production

Complete data = Raw data + Metadata

Generation of complete data is the goal and the must win battle of the Lab/Factory of the future



The digital journey in *Process Development*

Data integration - from µL to m³

Build data foundation to ease sharing across R&D and to Production Enabling future possibility of predictive process development and optimization



Traceability

Central

data repository

Searchable

Digitalization

Development of apps for easy data interpretation and comparison of process and analytical data



TRIAL COMPOSER | A data model for process development

IDENTIFIERS AT THREE LEVELS

- Trial ID (T#)
 - Batch ID (B#)
 - Sample ID (S#)





Process Step Parameters

- pH set-point
- Stirring
- Base used
- Temperature
 - ..

Associated Batch data

• Online data



Stability studies







																	Rs	tudio	Shiny,	5
CHR HANSEN TrialCompo	oser - Upscalir	ng Repo	ort	BATCH O	VERVIEW	ONLINE	DATA R	AW DATA	SANDBO	ĸ						T		828	tid	ggplo
Filters	Overview	Table 1	Tab	le 2														~		
Choose a trial:	This table is	This table is compiling TrialRecorder & Analysis Request data																		
10252 -	L Downle	Lownload PNG Lownload XLSX																		
Choose Process Step (s)																				
RECIPE, ACTIVE INGREDIENT, BIOLE 🔻	Trial 10252																			
Select Batch Parameter(s)																		Bat	tches	
	Process Step	Unit	FM01	FM02	FM03	FM04	FM05	FM06	FM07	FM08	FM09	FM10	FM11	FM12	FM13	FM14	FM15	FM16	FM17	Fl
THEMS SELECTED •	BioLector Fe	rmentation	n - eof																	
Choose AR Method(s)		- 8	104	6682	5747	5746	8162	1.524	7534	6894	5420	4915	5067	4212	9259	7924	6908	4790	0754	827
FlowCyt_Std		- 1	62e+10	1.65e=10	7.03e=09	2.28e+10	1.22e+10	5.33e=09	136+09	157e=04	84	84	NA.	NA	1.43e+10	1.57e+10	1.51e+09	2.38e+10	1.68e+10	1.0
Select AR Parameter(s)		- 8	880	1.044	0.285	0.647	0.545	0.265	0.148	8.112	54	NA.	NA.	NA	0.739	5.807	8.877	0.855	0.892	0.9
		-		1.0.0	NA .	104	NA	14	10.000	10.000	24	14	14	14	10.044	5.6	1.147	100	5.4	10.0
		-																		
					10403			****	10004		-	-	-	-		******	1.200	1.120		
		-	208		0.015	0.004	0.014	0.023	0.000	0.015	8.203	0.00	1.128	8.239	0.001	0.005	0.014	0.003	0.040	
		-																		
		-																		
				1.029	0.002	0.011	0.015	0.021	0.04	0.002	0.174	0.128	0.027	0.146	0.017	0.039	0.025	0.025	0.003	1.0
Data Download		- 3	676	2.008	2.837	1.432	2.62	2.378	4.078	3.394	3.808	2.614	4.754	2.48	3.348	2.618	3.452	1.21	4.44	3.3
							144		-	11.0	114	14	14	24	14	24	14	14	14	144
Data Export as xlsx workbook		- 8		10.00	100															



Data science – downscaled platform

- Process step data
 - Feature extraction
 - Time series data
- Stability data
- Microscopy data
- Fluorescence data
- (Production data)







But can we accelerate innovations even faster...

- With the rights sensor technologies and complete data we will be able to continue our journey of accelerating innovation
 - ✓ Via better screening and upscaling possibilities
 - ✓ Knowledge gathering possibilities
 - ✓ Via keeping focus on data trackability and comparability
 - \rightarrow Optimizing at production scale via a data driven approach
- But also via collaborations within
 - Ø Data science
 - Ø Downscaled technologies
 - Novel technologies

 DSK2020
 Data journey ahead of us kickstarted as we speak

 DSK2022
 ? ③

Ø