

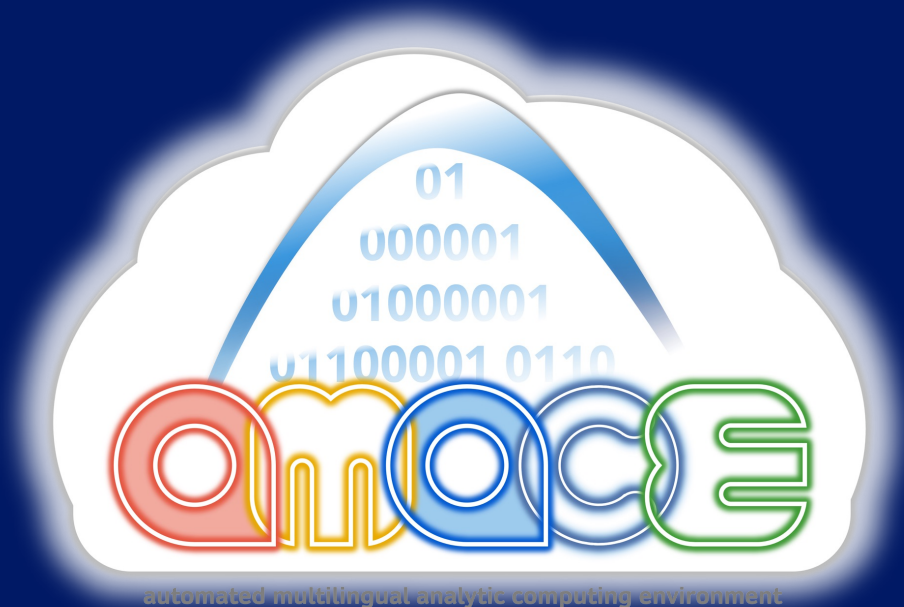
PHUSE EU Connect 2023

5th – 8th November 2023

From Legacy to the Cloud: Novo Nordisk's Journey towards a Modern Statistical Computing Environment

Presented by
Marianne Caramés
Principal Statistical Programmer/PO in AMACE,
Biostatistics Business Excellence, Novo Nordisk A/S
Andrew York
Statistical Programming Vice President,
Biostatistics, Novo Nordisk A/S

5-8 November 2023



Disclaimer and Disclosures

- *The views and opinions expressed in this presentation are those of the author(s) and do not necessarily reflect the official policy or position of PHUSE.*

Speaker Bio

Marianne has a Bachelor in Computer Science. She has worked as a Principal Statistical Programmer at Novo Nordisk for 25+ years in Biostatistics within areas like improvement projects, IT development (Clinical Data Warehouse and CDISC), standardisation and submissions. She is now one of the product owners in the AMACE project working SAFe Agile.



Andrew has an Masters in Applied Statistics. He has worked as a hands-on Statistician and Statistical Programmer as well as a number of Leadership roles in the industry over a career spanning 37 years. Andrew is a Statistical Programming VP at Novo Nordisk acting as a scientific VP and sits on the Leadership Team of Biostatistics organization. In terms of AMACE, Andrew is the Business owner for the AMACE project.



Agenda

1. Introduction
2. Request for Quotation and PoC
3. Design decisions
4. Challenges
5. Next steps

5-8 November 2023

Introduction

5-8 November 2023

Introduction

- An evaluation of our existing SAS environment reported that it was starting to age and would soon be un-supported
- Novo programmers had also started to work in R as well as SAS from around 2019 which included building and validating a separate SCE-R platform*
- As the use of R matured into production, we determined that modernization of our SCE environment to accommodate our future reporting strategy was highly desirable
- From here the AMACE journey began and today we will present our journey to date.... !

**This journey is covered in the recent POSIT webinar regarding the recent R submission to the FDA and can be found on YouTube*



Request for Quotation and PoC

5-8 November 2023

Finding the right vendor

- During 2022 Novo Nordisk A/S did a Request for Quotation (RfQ) with 8 possible companies to make our future SCE platform – a replacement for the current Unix-SCE
- An extensive scoring schema was produced to evaluate each product on the following:
 - Process management
 - Automation
 - Production of deliverable for submission
 - System Design
 - User Experience (UX)
- And the vendor on the following:
 - Competences and capabilities
 - Technical solution
 - Supplier capacity
- And the total solution on the following:
 - Price

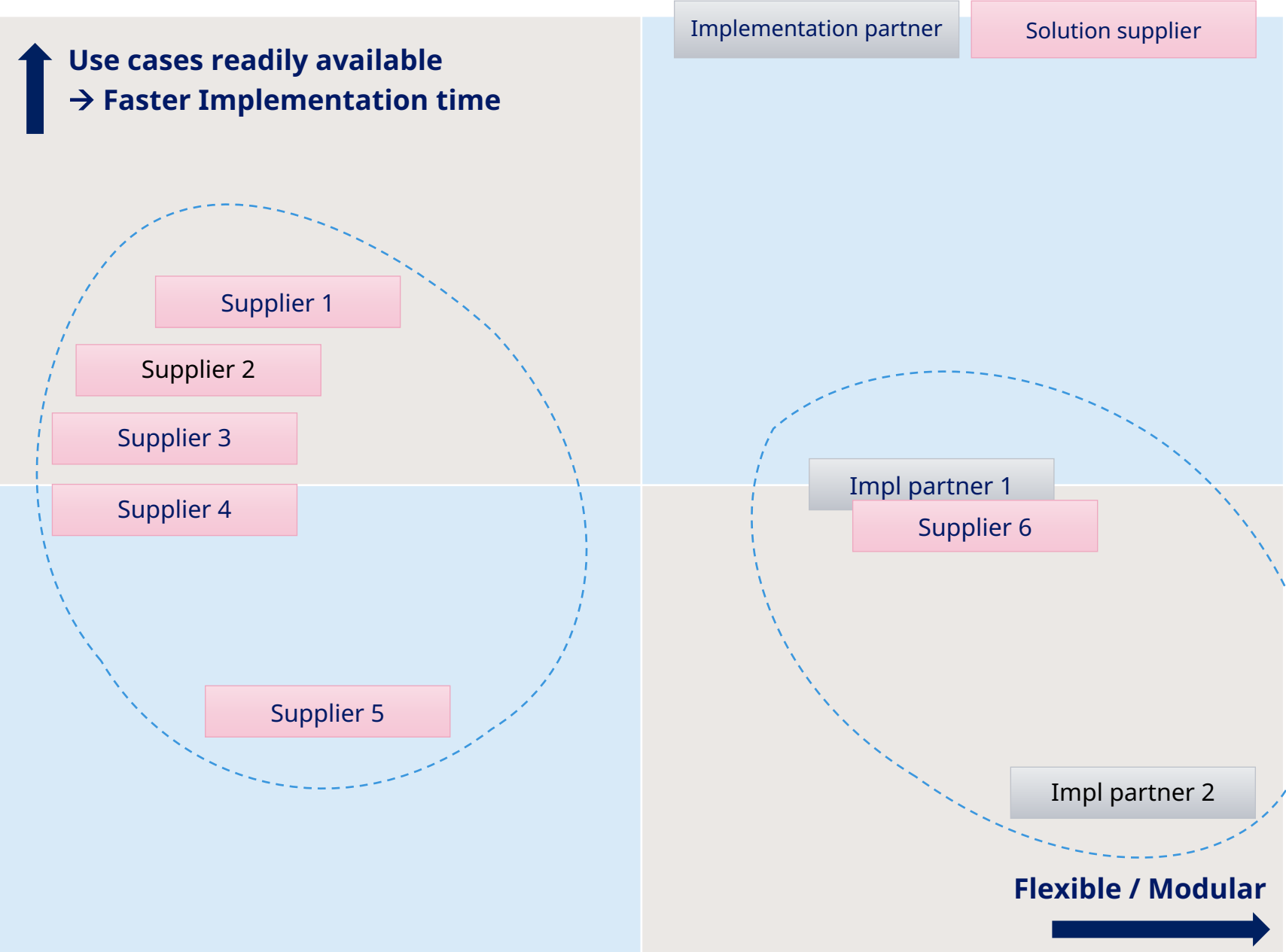
			User 9	User 10	User 11	User 12	User 13	User 14	User 15	User 16	User 17	User 18	Mean
1. Functionality (area)													
Area	Functionality	Explanation what exactly is meant											
Process management	Visual interface to set-up and interact with workflows	Interface meaning e.g. a GUI, a flowdiagram or such	3							2			2.3157
Process management	Visual interface to monitor progress of tasks	Interface meaning e.g. a GUI, a flowdiagram or such	3				3	2		2	3		2.5
Process management	Programmatical set-up and adjustment of Data pipelines/processes	Using programming or preferably metadata to set-up and interact with workflows	3				3	2		2			2.4
Process management	Reminders for tasks	E-mail or system notification to the user informing about tasks, issues or input for workflows	3				3	2		3	3		2.5
Process management	Versioning	Versioning for workflows	2				0	2		1	2		1.6
Process management	Documentation	All actions on a specific workflow	3				3	2		2	3		2.5
Process management	Authorisation: Operational access management	Restrict for unblinding data and confidential trials	3					2		3	3		2.4
		Access Dashboard	3							3			2.5756
Automation	Execution based on dates/times	Automation refers to program execution that does not require human interaction here specifically scheduled	3				3	3		3	3		2.8
Automation	Execution based on e.g. changes in upstream data	Automation refers to program execution that does not require human interaction here specifically scheduled	3				3	3		3	3		2.6
Automation	Execution based on e.g. changes in metadata	Automation refers to program execution that does not require human interaction here specifically scheduled (e.g. trial metadata, ADaM Master Model, PLACE (changes in output	3				3	3		3	3		2.6
Automation	Triggered/Event driven automation	Event or trigger such as e.g. passing a milestone can trigger execution of impacted programs or workflows. Not running	3				3	3		3	3		2.6
Automation	Automatic detection of program dependencies/program flow and Visual	Required to break circular references and to steer automation input	3				2	3		3	3		2.6

Presentation 1: One total score per vendor

	Supplier8	Supplier7	Supplier6	Supplier5	Supplier4	Supplier3	Supplier2	Supplier1	Weights
Process management	0,42	1,38	0,37	1,42	2,12	0,03	1,82	2,32	9%
Automation	0,15	0,94	0,51	1,35	1,78	0,00	1,16	2,58	16%
Production of deliverable for submission	1,30	2,51	1,09	2,08	2,38	0,15	2,39	2,43	15%
System Design	0,57	2,53	1,65	1,57	1,95	0,73	1,80	2,32	14%
User Experience (UX)	0,30	1,46	1,51	1,51	1,72	0,18	1,44	1,88	14%
Competences and capabilities	2,32	1,82	2,52	2,32	2,13	1,98	2,62	2,15	7%
Technical solution	2,25	1,90	2,40	2,00	2,40	1,70	1,75	2,35	9%
Supplier capacity	2,13	1,13	1,73	1,50	1,60	1,53	1,73	1,53	10%
Price	2,20	1,10	1,30	1,75	1,90	1,25	1,65	1,65	3%
Total	1,06	1,71	1,37	1,69	1,99	0,66	1,78	2,20	

Two solution archetypes

- “Out of the box SCE”
- Open platform



Selection and rationale for preference



**The solutions have been
scored based on:**

Design decisions

5-8 November 2023

Vision

Deliver an automated, scalable, multilingual setup to enable fast analysis of clinical data, facilitate collaboration and enable reusability for end-users

Design decisions

Use of NN AWS public cloud	Language agnostic (SAS, R, Python, etc.)	Simple but granular access model/control (able to work with our access control system and Active Directory)
Support for GxP compliant submissions (understandable audit trail)	Able to use our GIT repos in Novo Nordisk A/S	Efficient access to tools and data (system tools, access to sourcing data/metadata)
Sandbox environment (supporting innovation and experimentation)	Modular approach to add upcoming technologies	Support of automation and data lineage

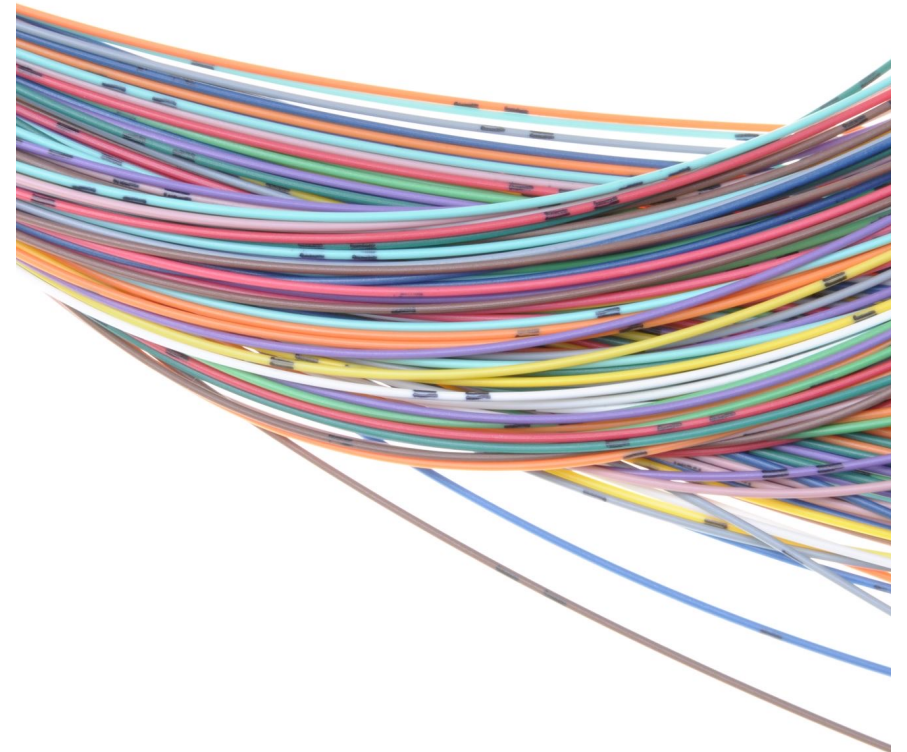


Challenges

5-8 November 2023

Integrations with other platforms/systems

- In Novo Nordisk A/S we are using the following systems, which we try to align with:
 - **Novo Nordisk access control system** (an access request and control system with Azure Active Directory (AD) (database of roles, groups, accesses))
 - **DataHub solution** (orchestration layer on top of native AWS Services, including S3, Lake Formation, IAM) used by other projects to integrate and share data



Challenges

- Setup new environment outside our corporate environment
- NNEDH only offers one type of storage – S3 buckets and no FS drive
- Limitations on number of S3 buckets
- Difficult to browse and navigate the S3 buckets directly from SAS/Domino
- Difficult to map the access model to the DataHub solution and Domino accesses
- Credential propagation from DataHub solution to Domino platform difficult to set up
- **So we needed to ask ourselves: Is the system we are building scalable and granular enough?**



Next step

5-8 November 2023

Next step

- A proof of concept with the elements Azure, DataBricks, Domino and SAS Viya will test the following:
 - Setup CDR using DataBricks
 - Implement access model
 - Integration with Novo Nordisk access control/Azure AD
 - Programming from SAS Viya
 - Programming from Domino
 - Browsing and navigating data from SAS Viya and Domino



Acknowledgements

- We want to thank the AMACE team members for continuous work to concur the challenges – specially Allan Thomassen Rudwill, Mathias Bigler and Tobias Krøgholt
- We want to thank the Domino team for their hard work to solve the challenges in the setup
- We want to thank the Domino SCE Coalition for their good input – specially Caroline Phares and Ross Sharp

Thank You for Your Attention!

Marianne Caramés

Principal Statistical Programmer/
PO in AMACE
Biostatistics Business Excellence

Novo Nordisk A/S
maca@novonordisk.com

**Andrew York**

Statistical Programming Vice President
Biostatistics

Novo Nordisk A/S
awyo@novonordisk.com

