Optimizing the Human Biological Sample Management (HBSM)

Maxime Brochier & Maxime Parres-Albert

Vaccine use case from **GSK**

Maxime Brochier, Clinical Laboratory Innovation Lead

Maxime Parres-Albert, Senior Engineer - Digital Innovation and Beyond

Conflict of interest : Maxime Parres-Albert & Maxime Brochier are employees of the GSK group of companies. This work was sponsored by GlaxoSmithKline Biologicals S.A.
Some of the used icons in the presentation are coming from www.flaticon.com

What is the HBSM process?

Context

What is Human Biological Sample Management

Process overview

Study design

Patient recruitment

Patient Vaccination

Sample collection

sample gathering

Sample Reception at GSK

Sample Storage

Sample preparation

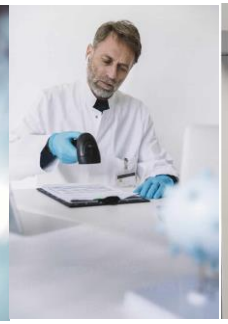
Sample Logistics

Sample Testing

Sample testing results

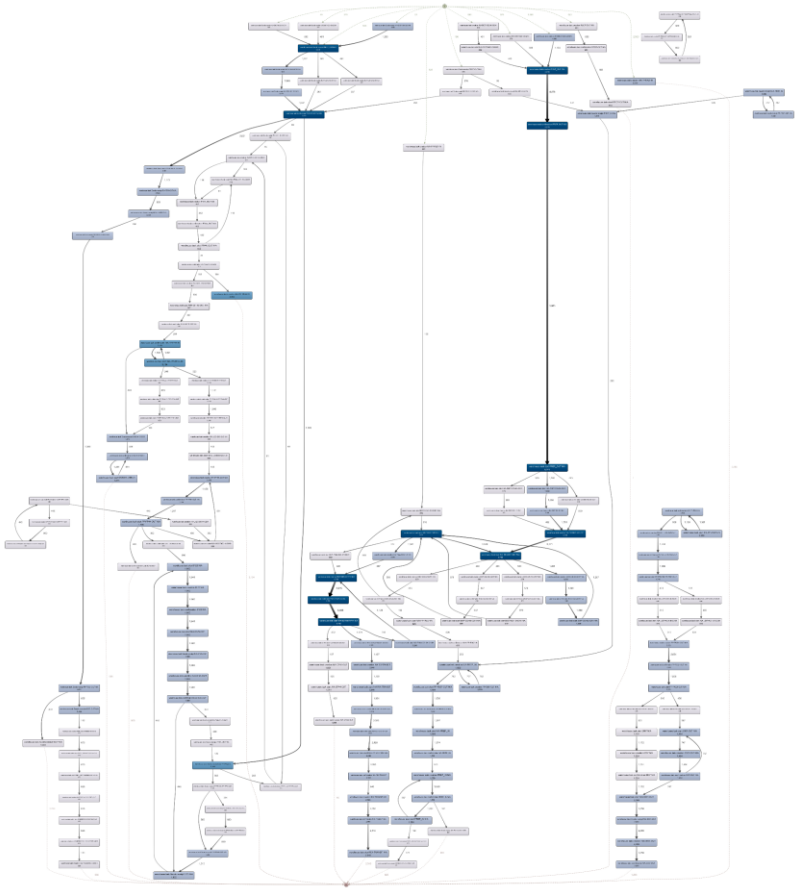
Results analysis

Study results



KEY Question: how could we speed up the process?

A **diagnostic** approach to find the process bottlenecks & best practices



GSK is managing millions of Human Biological samples going through hundreds of process steps!

Because the process is so long and complex, it is NOT possible for humans to fully understand it.

Data can tell the story!
By using process mining on Clinical Operations data we are able to map the process, identify and highlight bottlenecks & the best practices.

That help us to take action through change management to improve the process by removing the bottlenecks and align on the best practices.

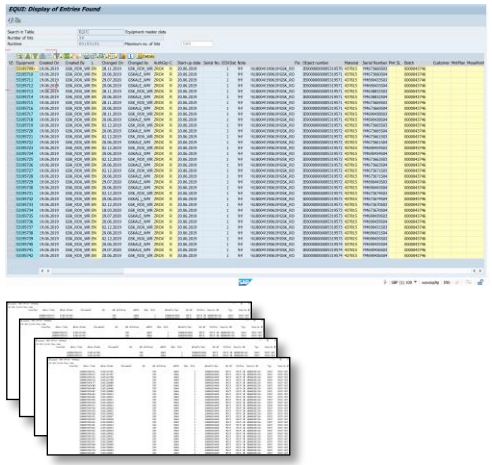
How is it different than a “usual” process mining project

Process complexity, big data quantity and complex data transformation engine

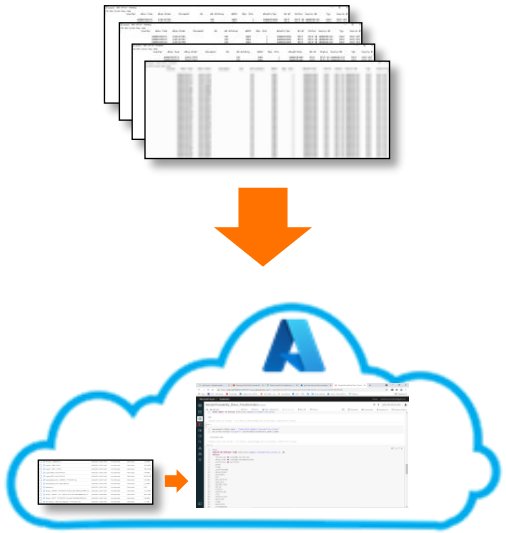
What has been achieved?

Complex data transformation through Python scripts in Ms Azure

Data Extraction (SAP)

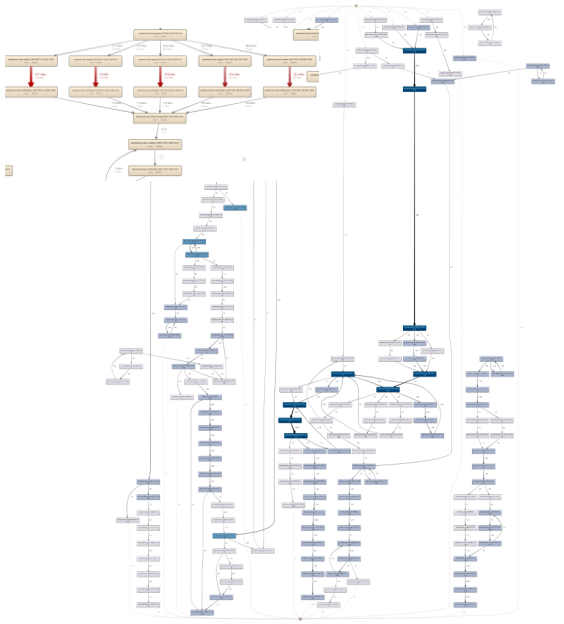


Data load & transformation



Disco Data Load

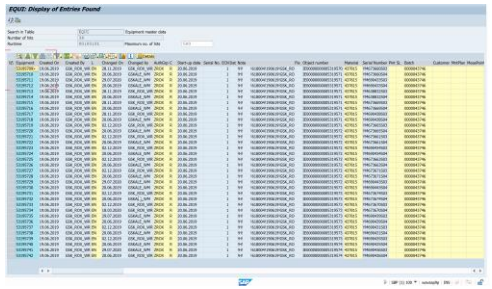
Process Map



What has been achieved?

Complex data transformation through Python scripts in Ms Azure

Data Extraction (SAP)



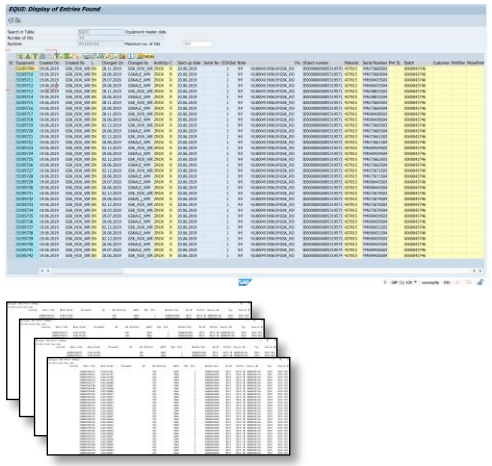
The team built a simple **RPA Bot** (Robotic Process Automation bot) to **automate the collection of HBSM data** from our SAP database.

This bot is **saving us a lot of time** (hours/day were spent on manual extract of data), guarantees **consistent data capture** and **removes human mistakes**.

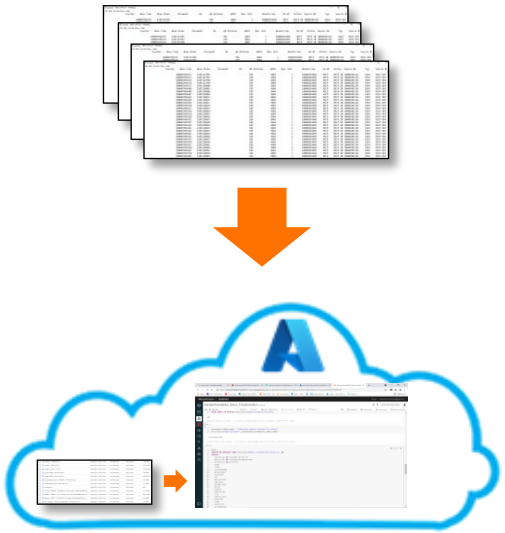
What has been achieved?

Complex data transformation through Python scripts in Ms Azure

Data Extraction (SAP)



Data load & transformation



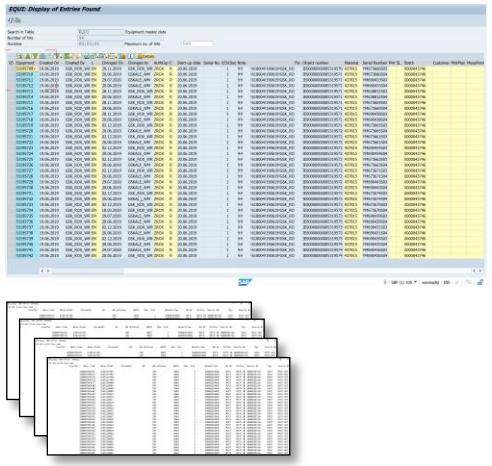
The team used another **RPA Bot** to **load** the data in an Azure environment in which a **data transformation engine** has been developed in *SQL* and *Python* to connect and clean our collected *SAP* data.

Some **business rules** have been defined to ensure **clean** and **senseful** data.

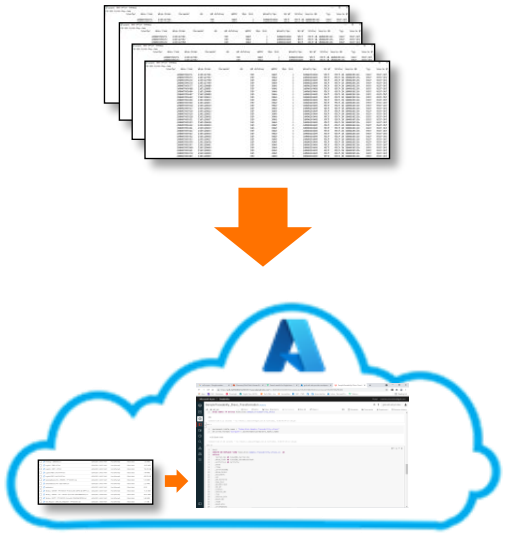
What has been achieved?

Complex data transformation through Python scripts in Ms Azure

Data Extraction (SAP)



Data load & transformation



Disco Data Load

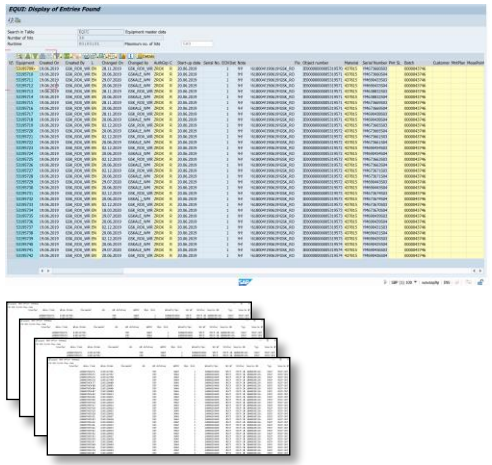
Our **data transformation engine** produces a clean CSV file **readable by Disco**.

No Excel transformation is required (which could not be feasible taking into account the heavy amount of data (<1M lines)

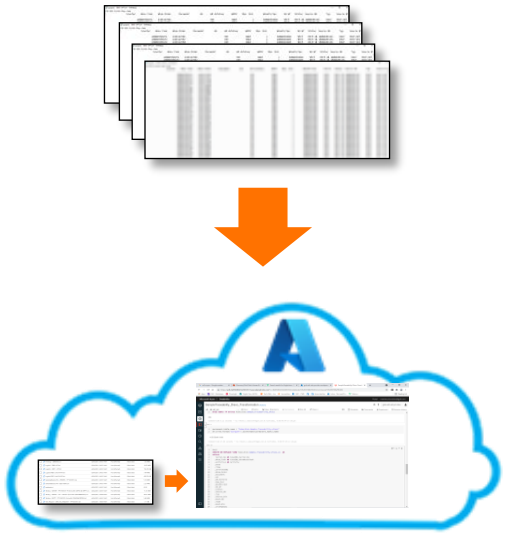
What has been achieved?

Complex data transformation through Python scripts in Ms Azure

Data Extraction (SAP)

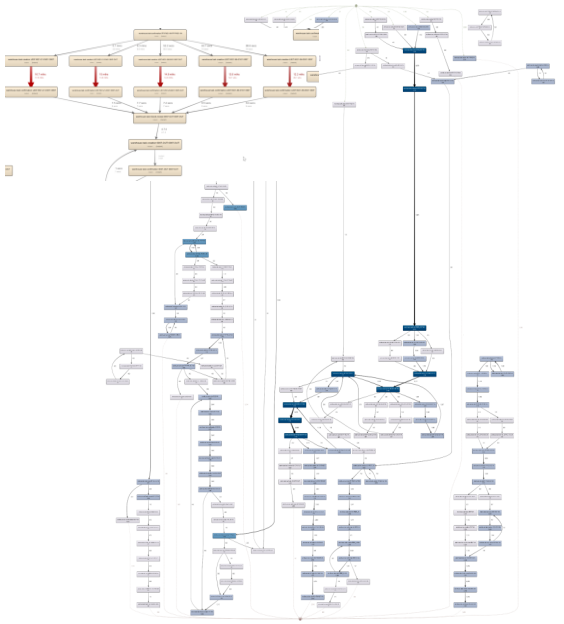


Data load & transformation



Disco Data Load

Process Map



Tips for your own projects

Recommendations and data transformations

TIP 1 : Which activity are we using?

How to define the process key activities?

IT Systems are a **gold mine** for process data through the "**audit trail tables**":

Unique Identified (Serial Number)	Measuring Point	Parameter field	… (context data)	Parameter value change - **Old value**	Parameter value change - **New Value**	Timestamp on parameter change (Date & Time)
1244714	012	PTTXT	LAB_WAVRE_BIO	SXX -ACTIVE	N -ACTIVE	11/02/2022 14:44:47

→ **Testing activity has started**. This record can be used in process mining

1458884	012	ZZLABR	TEMP_LOG_087	Blank	X	07/01/2022 18:01:17

→ **Sample is physically available for lab testing.** This record can also be used in process mining

- **CDHDR** is a standard **Change Documents Transparent Table** in SAP Basis application, which stores **Change document header data**.
- **CDPOS** is a standard **SAP Cluster table** which is used to store **Change document items data**

TIP 2 : How to deal with a lot of Activities

Group two activities with unique timestamp into a single with a duration

N°	Activity	Start Timestamp
1	Sending sample for testing activities	1/11/2021 10:51
2	Sample reception from testing activities	2/11/2021 19:03

TIP 2 : How to deal with a lot of Activities

Group two activities with unique timestamp into a single with a duration

N°	Activity	Start Timestamp
1	Sending sample for testing activities	1/11/2021 10:51
2	Sample reception from testing activities	2/11/2021 19:03
3	Storing sample in Warehouse	2/11/2021 22:45

TIP 2 : How to deal with a lot of Activities

Group two activities with unique timestamp into a single with a duration

N°	Activity	Start Timestamp
1	Sending sample for testing activities	1/11/2021 10:51
2	Sample reception from testing activities	2/11/2021 19:03
3	Storing sample in Warehouse	2/11/2021 22:45

TIP 2 : How to deal with a lot of Activities

Group two activities with unique timestamp into a single with a duration

N°	Activity	Start Timestamp
1	Sending sample for testing activities	1/11/2021 10:51
2	Sample reception from testing activities	2/11/2021 19:03
3	Storing sample in Warehouse	2/11/2021 22:45

TIP 2 : How to deal with a lot of Activities

Group two activities with unique timestamp into a single with a duration

N°	Activity	Start Timestamp
1	Sending sample for testing activities	1/11/2021 10:51
2	Sample reception from testing activities	2/11/2021 19:03
3	Storing sample in Warehouse	2/11/2021 22:45

N°	Activity	Start Timestamp	End Timestamp
1	Sample distribution for testing activities		

TIP 2 : How to deal with a lot of Activities

Group two activities with unique timestamp into a single with a duration

N°	Activity	Start Timestamp
1	Sending sample for testing activities	1/11/2021 10:51
2	Sample reception from testing activities	2/11/2021 19:03
3	Storing sample in Warehouse	2/11/2021 22:45

N°	Activity	Start Timestamp	End Timestamp
1	Sample distribution for testing activities	1/11/2021 10:51	

TIP 2 : How to deal with a lot of Activities

Group two activities with unique timestamp into a single with a duration

N°	Activity	Start Timestamp
1	Sending sample for testing activities	1/11/2021 10:51
2	Sample reception from testing activities	2/11/2021 19:03
3	Storing sample in Warehouse	2/11/2021 22:45

N°	Activity	Start Timestamp	End Timestamp
1	Sample distribution for testing activities	1/11/2021 10:51	2/11/2021 19:03

TIP 2 : How to deal with a lot of Activities

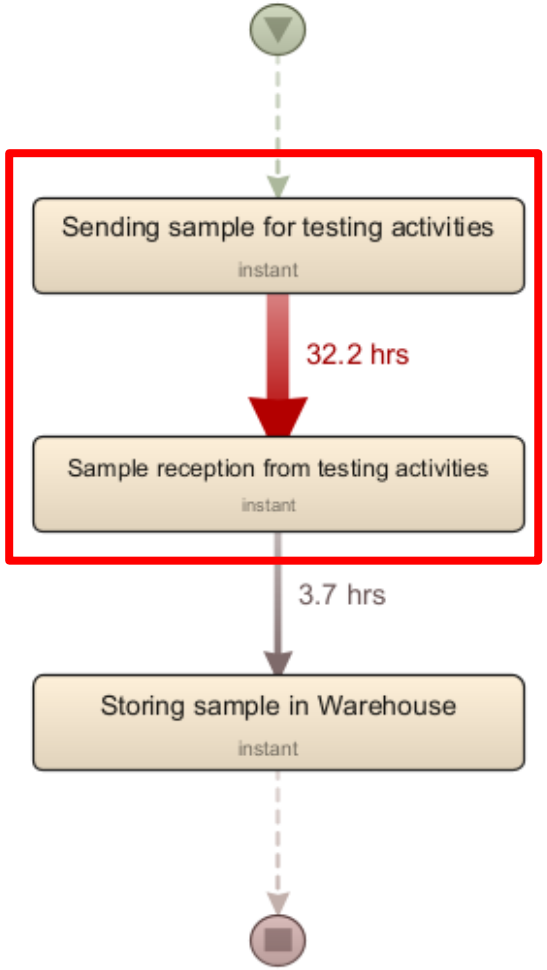
Group two activities with unique timestamp into a single with a duration

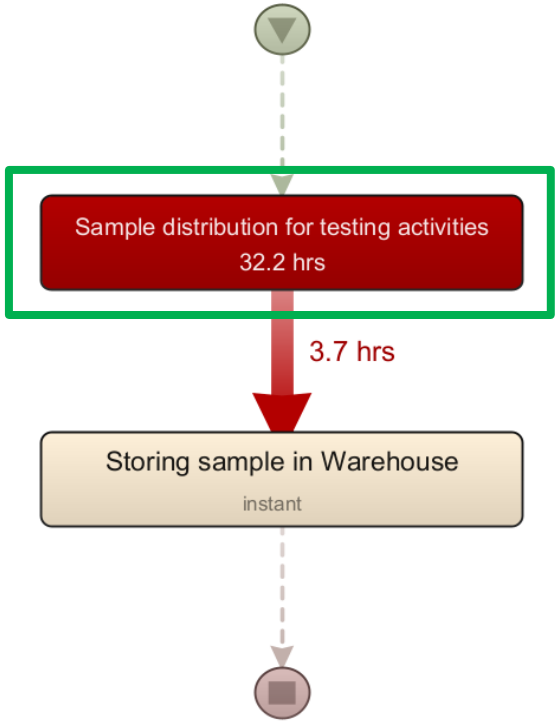
N°	Activity	Start Timestamp
1	Sending sample for testing activities	1/11/2021 10:51
2	Sample reception from testing activities	2/11/2021 19:03
3	Storing sample in Warehouse	2/11/2021 22:45

N°	Activity	Start Timestamp	End Timestamp
1	Sample distribution for testing activities	1/11/2021 10:51	2/11/2021 19:03
2	Storing sample in Warehouse	2/11/2021 22:45	

TIP 2 : How to deal with a lot of Activities

Group two activities with unique timestamp into a single with a duration





TIP 3 : How to deal with a lot of Activities

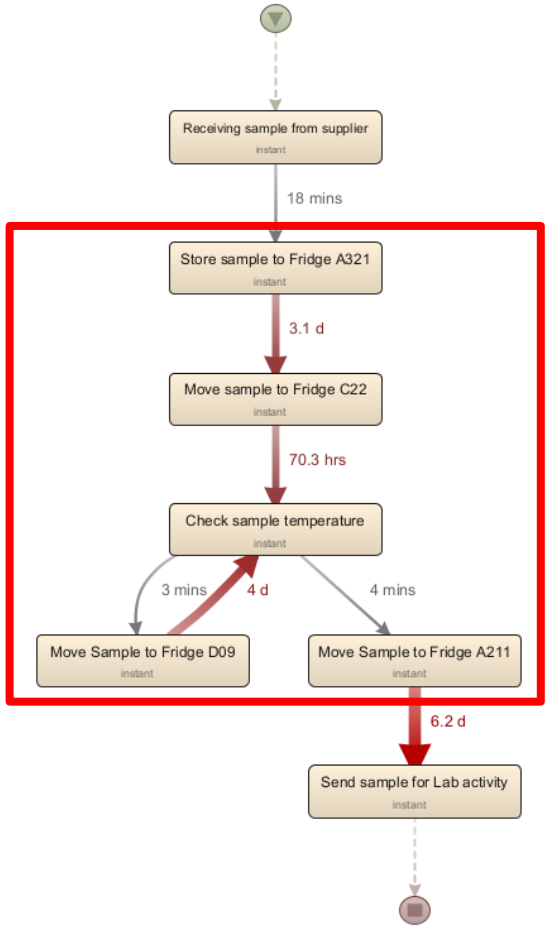
Less granularity, group activities

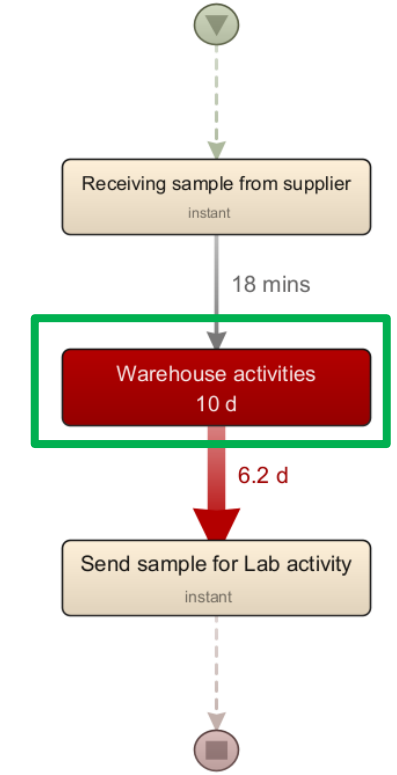
N°	Activity	Start Timestamp
1	Receiving sample from supplier	1/11/2021 10:51
2	Store sample to Fridge A321	1/11/2021 11:09
3	Move sample to Fridge C22	4/11/2021 12:45
4	Check sample temperature	7/11/2021 11:01
5	Move Sample to Fridge D09	7/11/2021 11:04
6	Check sample temperature	11/11/2021 11:04
7	Move Sample to Fridge A211	11/11/2021 11:08
8	Send sample for Lab activity	17/11/2021 15:33

N°	Activity	Start Timestamp	End Timestamp
1	Receiving sample from supplier	1/11/2021 10:51	
2	Warehouse activities	1/11/2021 11:09	11/11/2021 11:08
3	Send sample for Lab activity	17/11/2021 15:33	

TIP 3 : How to deal with a lot of Activities

Less granularity, group activities





From **Task Mining** to **Process Mining**

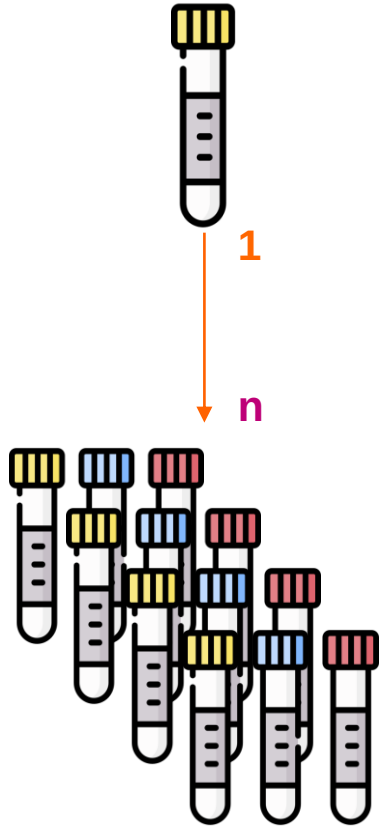
TIP 4 : A 1→n case id scenario

How to follow cases which might derive into several new cases

In our case, we have « **mother tubes** » that can derive into several « **child tubes** ». Which creates a 1→n relationship which could be hard to follow into our model.

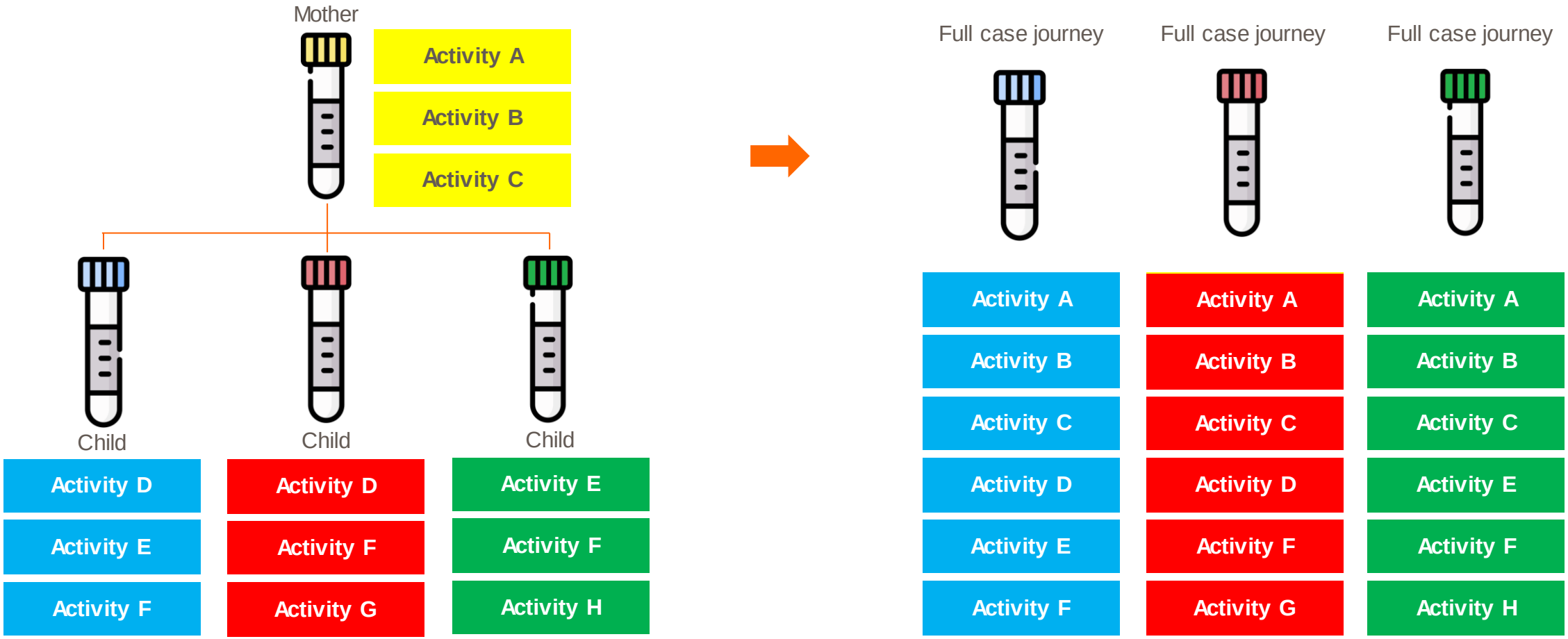
For HBSM process mining project, we have decided that **each *child*** would be a **single use case**.

Therefore, we decided to replicate the *mother's* activities upfront each *child* case.



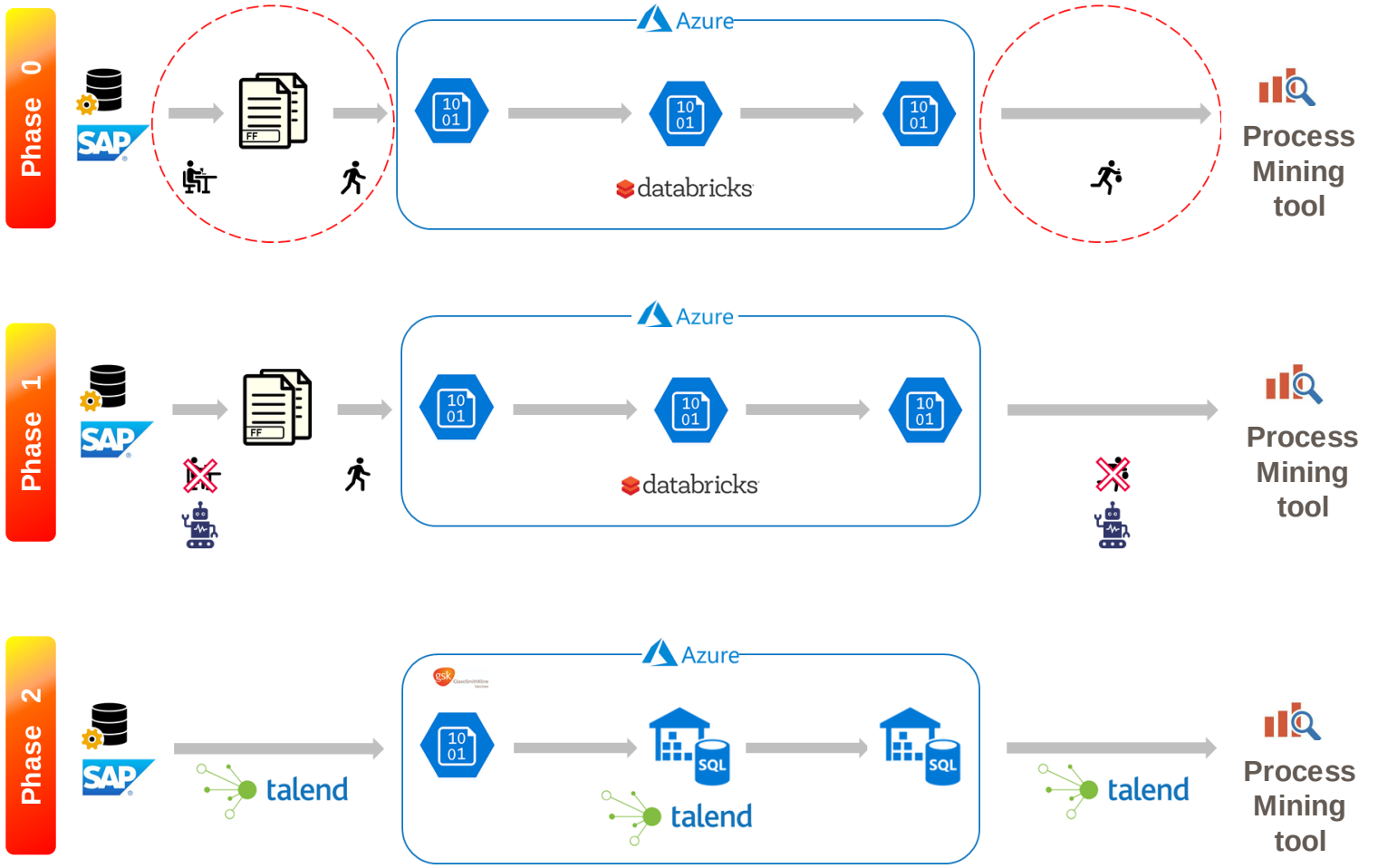
TIP 4 : A 1→n case id scenario

How to follow cases which might derive into several new cases



TIP 5 : Start small, scale fast and then industrialize

Agility and Flexibility



Manual process:
- Data collection
- Data upload to Azure transformation Engine
- Data upload from Azure to Disco

Semi-Automatic process:
- Data collection through RPA
- Data upload to Azure transformation Engine
- Data upload from Azure to Disco through RPA

Automatic process:
- Data collection through ETL
- Data upload to Azure transformation Engine through ETL
- Data upload from Azure to Disco through ETL

Results and next steps?

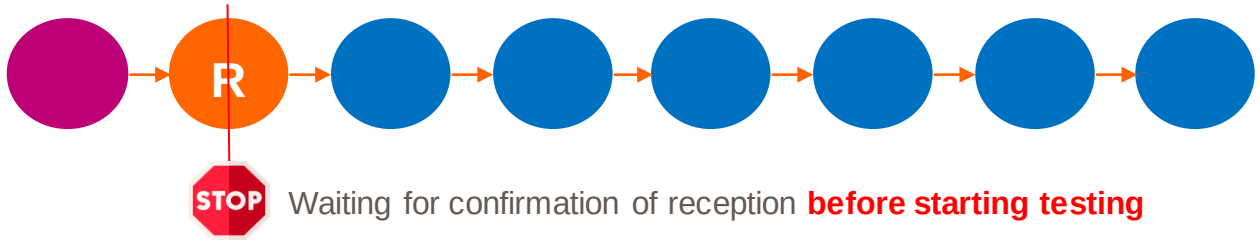
Comparing the before, the now and the after

How to release Clinical test results faster

Before and after

- Data **reconciliation** from « Patient data » towards collected tubes takes time and **blocks** the Sample reception process
- Important **impact** on the process **lead-time after reception**

BEFORE

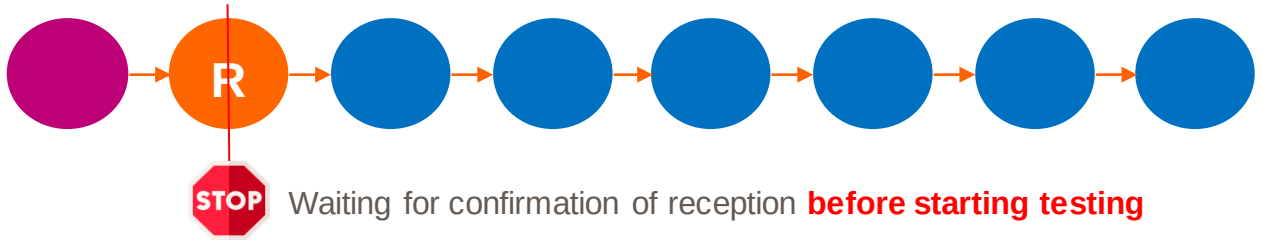


How to release Clinical test results faster

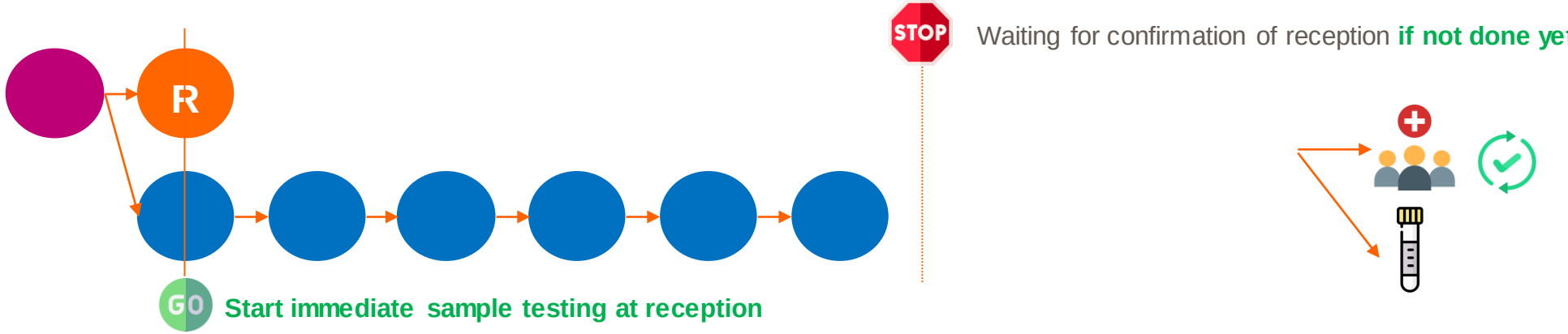
Before and after

- Data **reconciliation** from « Patient data » towards collected tubes takes time and **blocks** the Sample reception process
- Important **impact** on the process **lead-time after reception**

BEFORE



AFTER



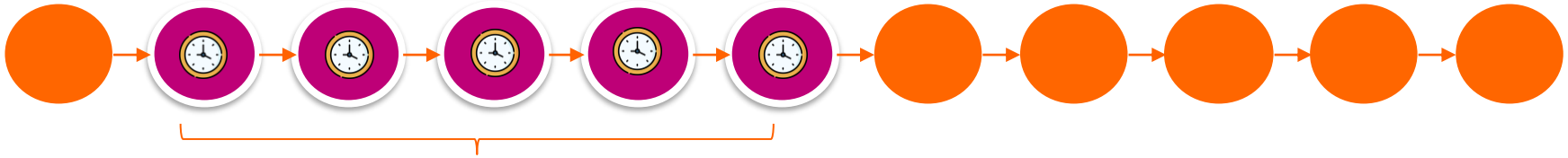
- Reconciliation step is done in **parallel** with other process activities
- process mining has enabled the **identification of the process bottleneck** (not only within GSK)

How to prepare the samples faster

Before and after

- The **current steps** to be performed for sample preparation in the system are **heavy**

BEFORE



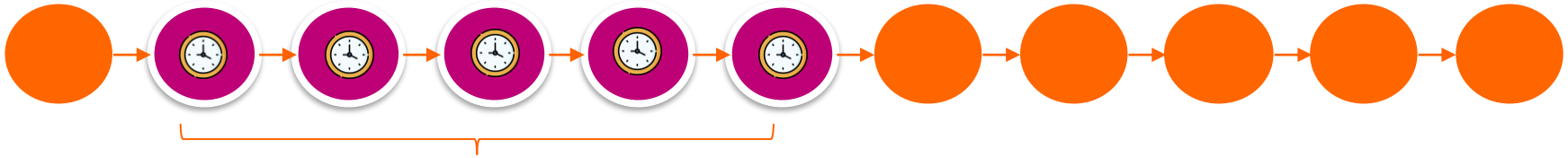
Inefficient and **slow** “system heavy” steps

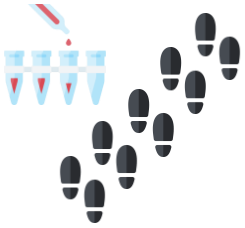
How to prepare the samples faster

Before and after

- The **current steps** to be performed for sample preparation in the system are **heavy**

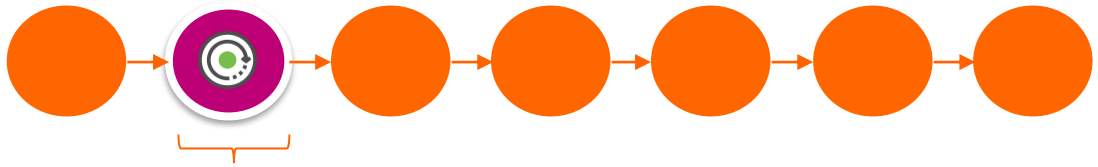
BEFORE

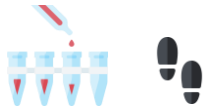




Inefficient and **slow** “system heavy” steps

AFTER



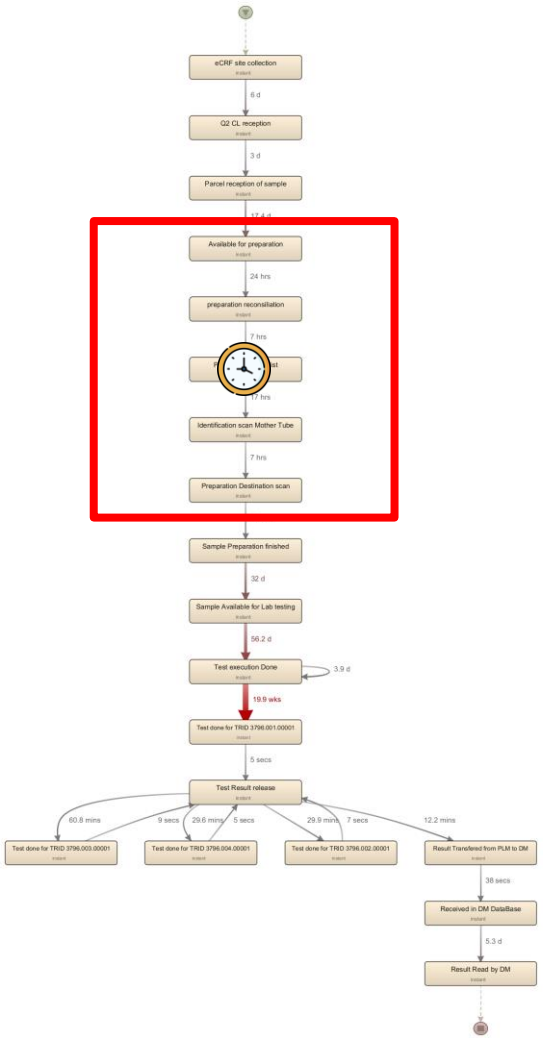


Improved system:« 1-click » solution

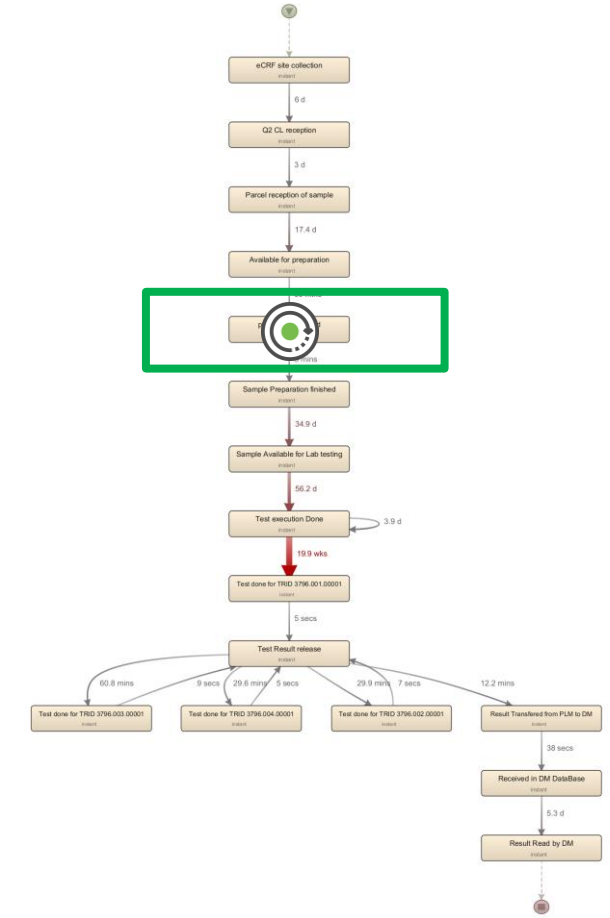
- Implementation of a system upload transaction to **avoid the heavy system steps**

How to adapt the business process from insights

BEFORE



AFTER

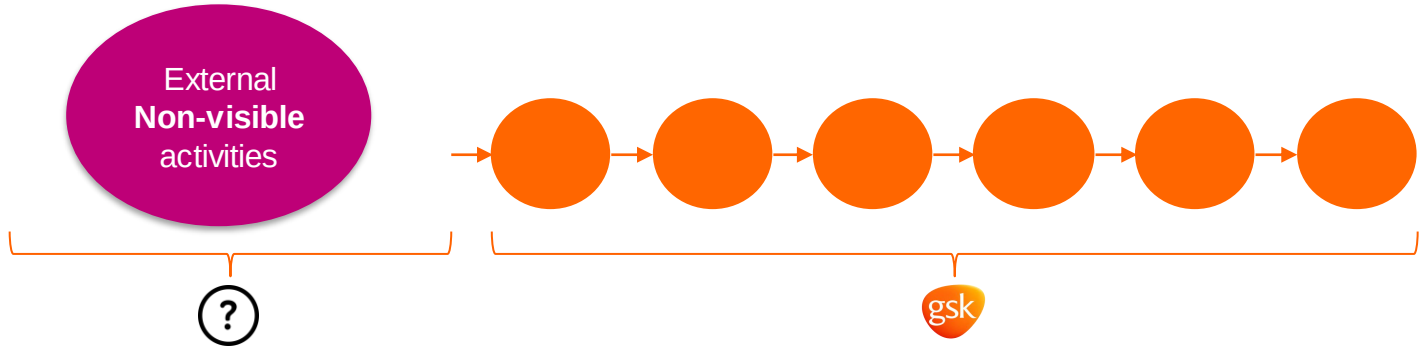


How to increase process insights & E2E management

Before and after

- The current **data sources** are **exclusively** coming from **GSK** systems
- A **lack of visibility** on the process at external partners

BEFORE

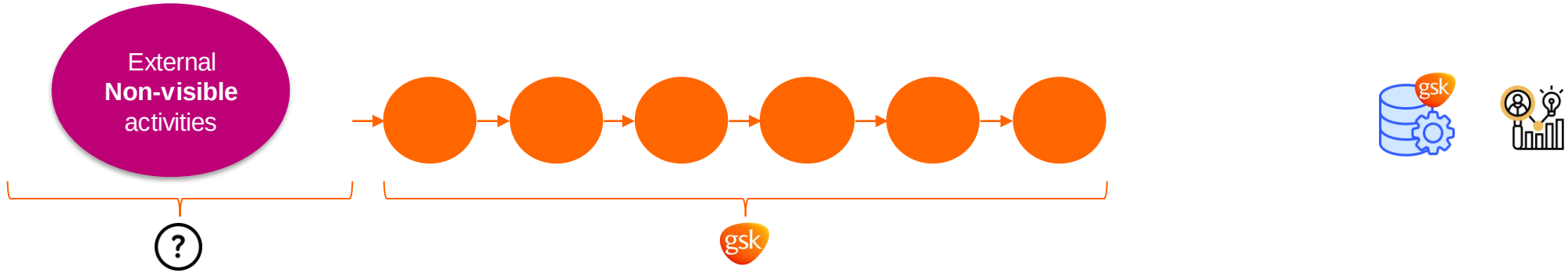


How to increase process insights & E2E management

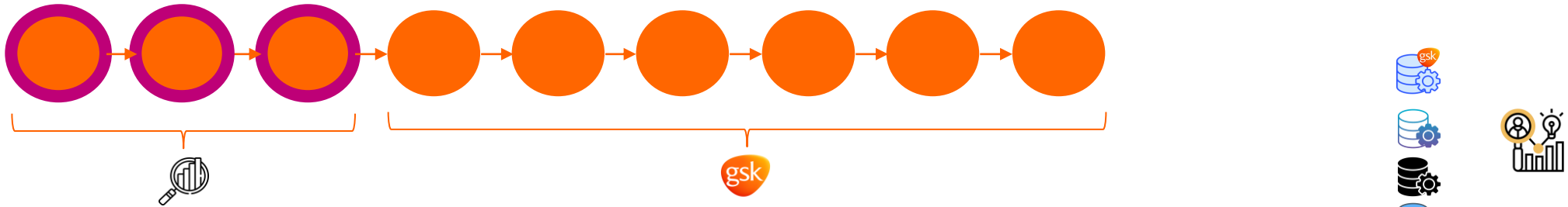
Before and after

- The current **data sources** are **exclusively** coming from **GSK** systems
- A **lack of visibility** on the process at external partners

BEFORE



AFTER



- Access **new data sources** from **external** partners

How to adapt the business process from insights

Before and after

Goal

- Increase the insights by extending the data sources outside GSK systems
- better plan and manage the CE process

Pain

- The current data sources are exclusively coming from GSK systems
- A lack of visibility on the process at external partners

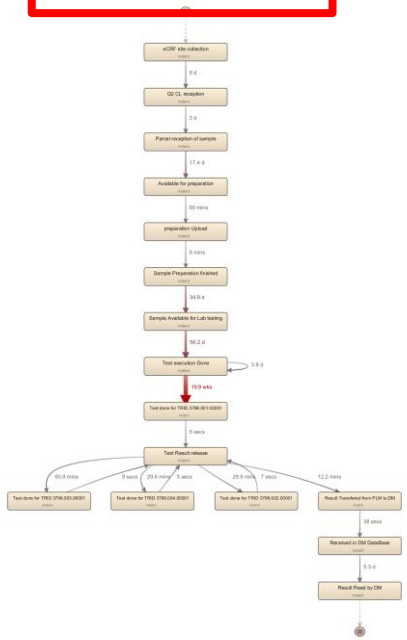
Process change

- Access new da… …ers

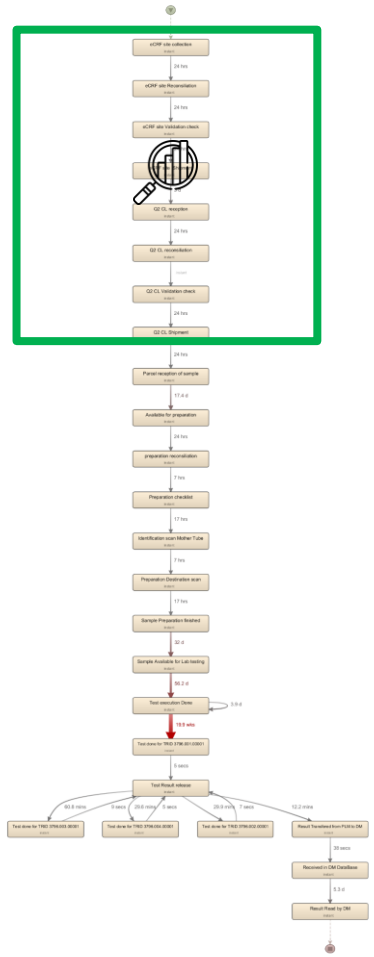
BEFORE

AFTER

BEFORE



AFTER

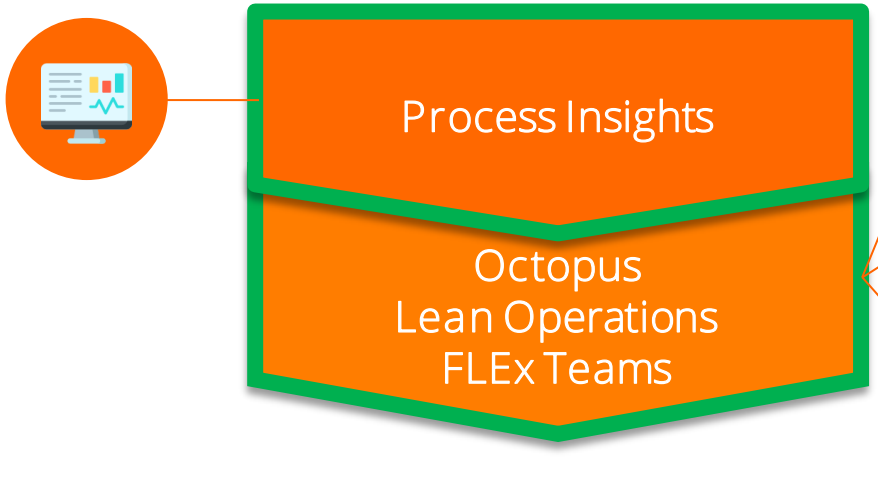


How to act on your process

Change management

How to adapt the business process from insights

Business SMEs
Process Monitoring **Insights**



FLEx – study specific
- Study Lead-time review
- Monitor study operations
- address inefficiencies

Lean Ops – cross study
- FLEx Team alignment, status, decisions
- Endorse proposed process changes

Leadership Team (LT) :
- Acceleration sponsor

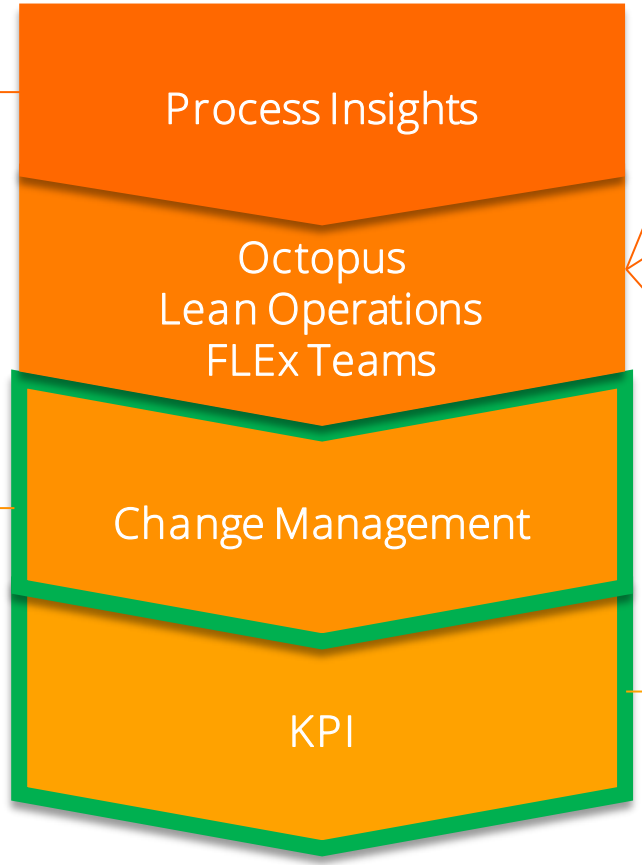
AXelerate FLEx Team Structure (approach)

	Lean Operations	Study FLEx Teams
What	FLEx Team alignment, status, decisions Endorse proposed process changes Manage AXelerate Menu	Monitor study operations, address inefficiencies,
Who	Head of Labs, Head of SCD & Scrum Masters ….	Scrum Masters , LSM (represents GMASE & Central Lab), Lab Manager(s), PLM, DM, DA&US, CRTL (case by case) (Ad hoc: Lean Operations members, SMEs…)
When	Weekly	Weekly
Agenda	• End-to-end Process Flow status and proposed changes *(endorse and support) BHE (create ad hoc Process task teams)* • Escalations: Risks, issues, prioritization decisions, support needed • Decide & align on Communication needs • AXelerate Menu *(if required)*	• Review Sample & Equipment Dashboards *(Blocking factors: add actions to Kanban)* • Lead Time review *(In-efficiencies: actions to Kanban)* • End-to-end Process Flow review & update *(create ad hoc Process task teams)* • Kanban Review
Com.	Centralize reporting to SteerCo, Octopus, LT…	Study status report to Lean Ops via Scrum Master

How to adapt the business process from insights

Business SMEs
Process Monitoring **Insights**

Change impact assessment
Change implementation



FLEx – study specific
- Study Lead-time review
- Monitor study operations
- address inefficiencies

Lean Ops – cross study
- FLEx Team alignment, status, decisions
- Endorse proposed process changes

Leadership Team (LT) :
- Acceleration sponsor

Ensure through robust KPI's **that the change has improved the process** (Process monitoring)

Summary


Problem
What is the problem
Agile
Start small
Exploration
Data
Data & process Insights
Change
Experts
Business & IT & Data
Validate

Thank you! Any Question?

Maxime Brochier, Clinical Laboratory Innovation Lead

Maxime Parres-Albert, Senior Engineer - Digital Innovation and Beyond