



PRE AD06

An Integrated J&J Application module to Enable Sustainable Clinical Trial Design

Authors: Archana Nayak,
LaRoche Jason,
Bhattacharjee Eshanika ,Md
Umarfaruque, Bhargav
Reddy

Agenda

- Introductions
 - Need for Sustainability
 - Identify WHY and WHERE Sustainability impacts Clinical Trial Design
- Application development
 - Architecture
 - User Interface
- Model development
 - Methodology
 - Enhancements

Need for Sustainability

The Market Landscape is Changing



1 – Regulatory Obligations

Governments are raising the bar for climate, circularity and zero pollution by setting ambitious targets, increasing investments and deploying stricter corporate sustainability disclosure requirements.

“SEC Movement on ESG Disclosures: SEC and Investors Enhance Focus on Climate Change and Other ESG Disclosures”

THE
NATIONAL
LAW REVIEW

J&J Innovative Medicine

2 – Customer Expectations

Customers want to contribute to more sustainable healthcare by opting for more sustainable products and solutions. Increasingly, customers request product **sustainability performance transparency**.

“NHS becomes world’s first health system to commit to net-zero”

BusinessGreen



3 – Stakeholder Expectations

Employees expect action now! **Investing in sustainability is key to attract future talent** and meet the expectations of investors.



External Momentum is Building for Sustainability in HTAs

Netherlands 

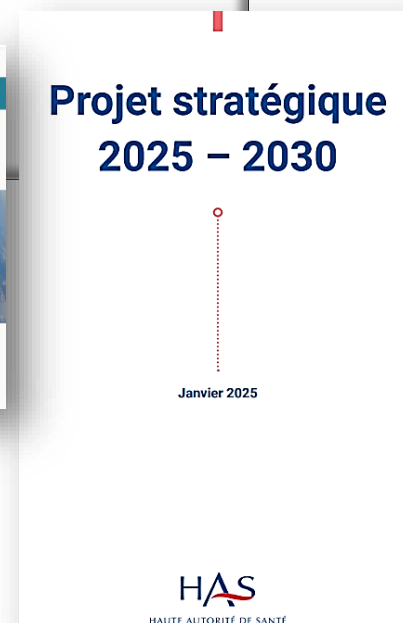
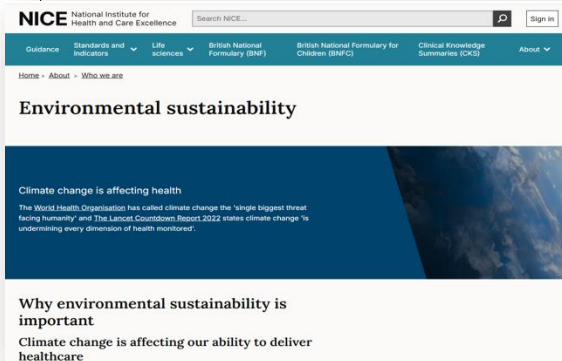
3-year pilot by the Zorginstituut (ZIN), the Dutch HTA agency, to investigate how to operationally include environmental sustainability & personnel impact in HTA decisions.

UK 

“HTA agencies must invest in developing or hiring relevant expertise to carry out environmental assessments and incorporate environmental data in health economic methodologies.” NIHR is exploring integrating PCFs into HTAs. Additionally, NIHR is also integrating PCFs into their decommissioning framework.

France 

Haute Autorité de Santé (HAS), the French HTA agency, under their “Project Stratégique” will begin considering how the environmental impact of health technologies can be incorporated into medico-economic assessments used by their Economic and Public Health Evaluation Committee (CEESP).



WHY and WHERE Sustainability impacts Clinical Trial Designs and Deployment

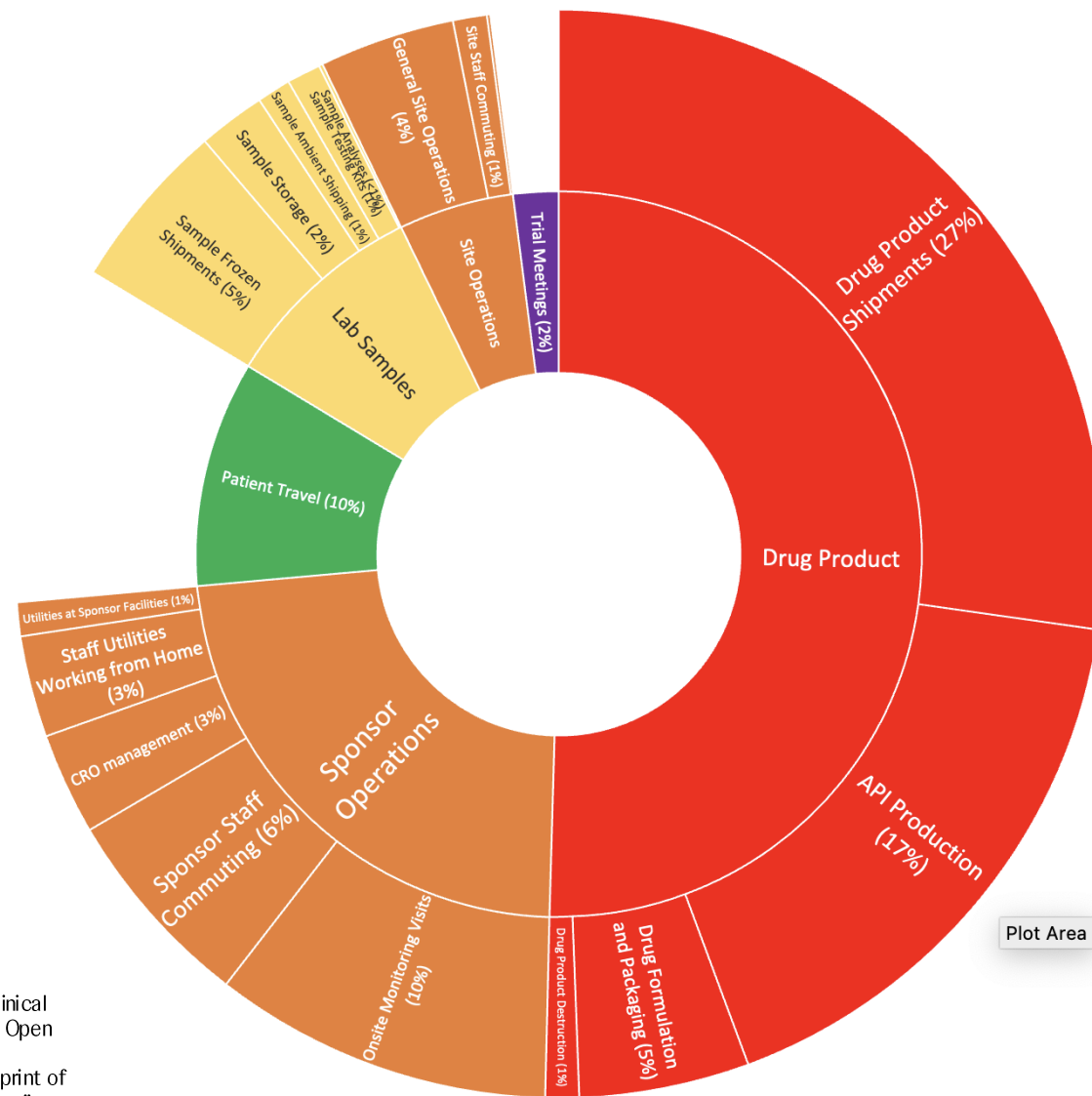


Identifying Climate Hotspots in GD Operations

Life cycle assessments were performed on a cross-section of J&JIM studies spanning multiple therapeutic areas and all four phases of development^{1,2}. Six activities were observed in combination to drive, on average, roughly 90% of the greenhouse gas emissions in our trials:

- Investigational Product: 50%
- Patient Travel: 10%
- On-site Monitoring Visits: 10%
- Central Labs 10%
 - Frozen sample shipments to central labs: 5%
- Sponsor staff commuting: 6%
- Site Operations 5%

1. LaRoche J, Alvarenga R, Collins M, Costelloe T, De Soete W, Faludi J, Rens K, "Climate Footprint of Industry-Sponsored Clinical Research: An Analysis of a Phase-1 Randomized Clinical Study and Discussion of Opportunities to Reduce Its Impact", BMJ Open 2024;14:e077129. doi:10.1136/bmjopen-2023-077129.
2. LaRoche J, Lanier J, Alvarenga R, Collins M, Costelloe T, Chau A, Weatherly H, De Soete W, Rens K, Faludi J, "Climate footprint of industry-sponsored in-human clinical trials: life cycle assessments of clinical trials spanning multiple phases and disease areas", BMJ Open 2025;15:e085364. doi: 10.1136/bmjopen-2024-085364



Application architecture



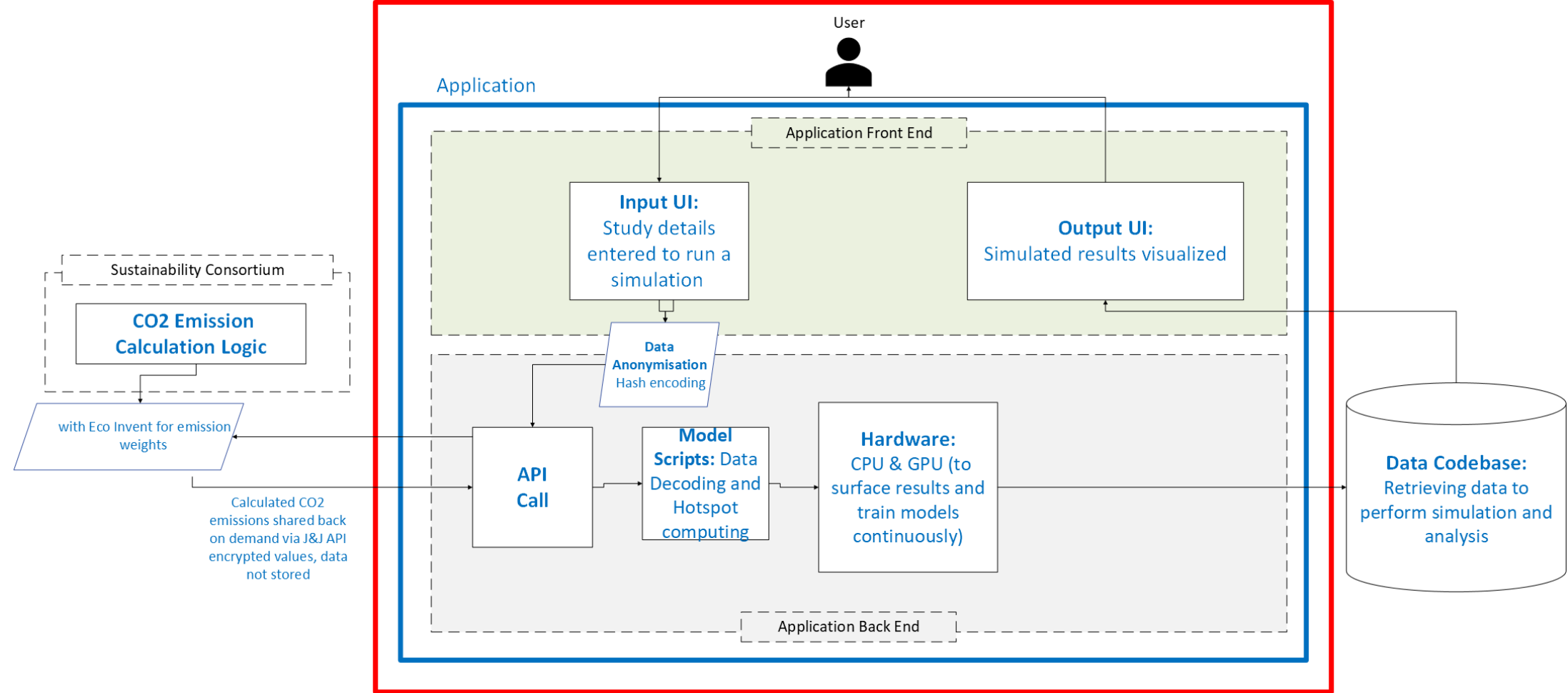
Application Design and architecture

The J&J Sustainable Trial Design
Application module consists of four key
components:

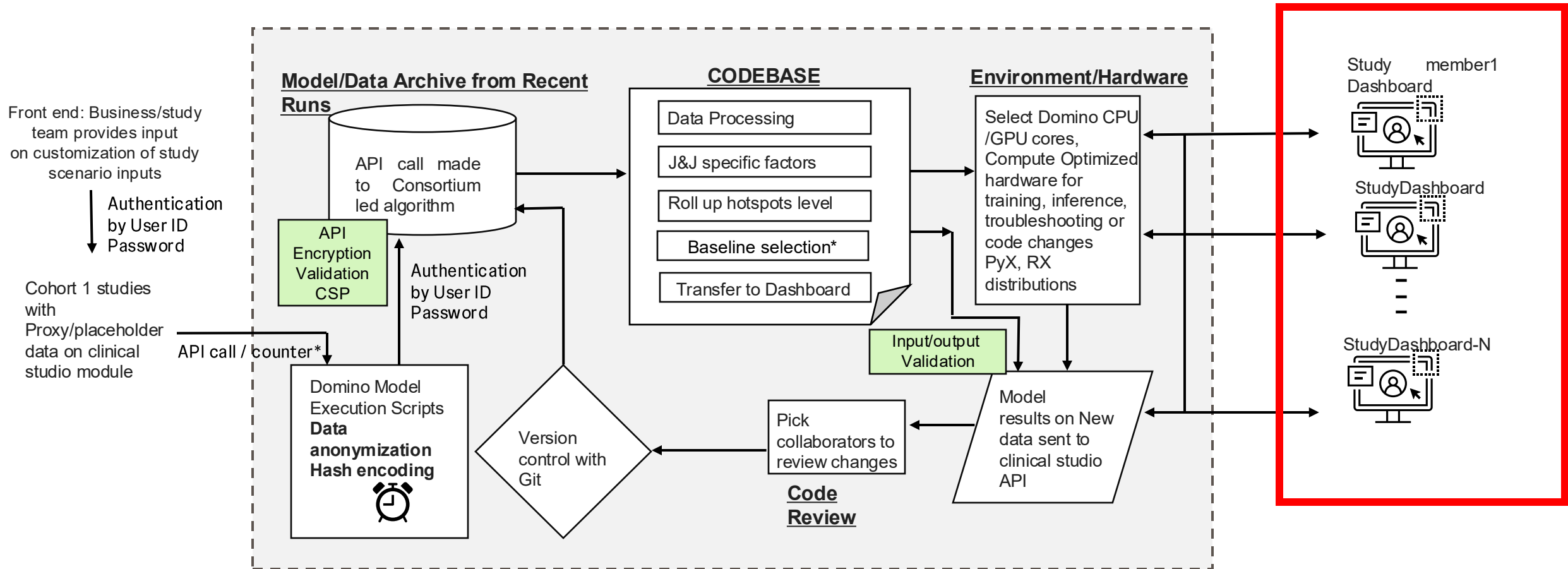
- Data Aggregation Layer
- Emissions Calculation Consortium
- Internal decoder and mapping activity
- Visualization and scenario testing

Data Flow diagram

Johnson & Johnson



Data pipeline



Application User interface



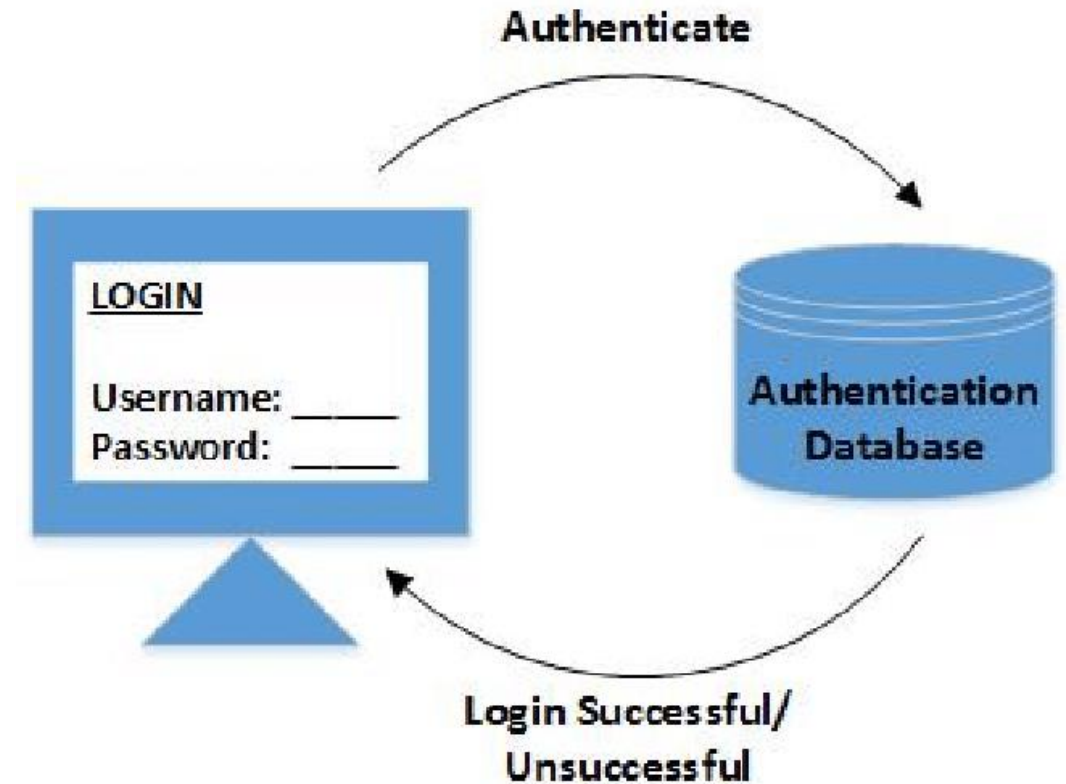
Authentication and security

Login page authentication

Access is restricted to only those users that are authorized within our domino project to access the API

Encoding and decoding of sensitive info before/after calling of API

All endpoints that access sensitive data enforce authentication, authorization and anonymization by hash encoding as follows:



Secure Login Page

URL/Link to Internal Sustainability Calculator

 All Apps / AI_Sustainability_PROD

Sustainability Dashboard

Login Page

Username

Enter username

Password

Enter password

Login

Landing/Home Page

Purpose: Ability to create a new scenario or review previous results

Sustainability Dashboard

Home

Click to create a new evaluation

Create New Evaluation

Streamline Your Trials and Reduce Carbon Footprint: This dashboard empowers you to make informed decisions about your clinical trials by providing insights of the environmental impact based on protocol inputs. By quantifying the carbon footprint of your protocols, you can analyze and identify opportunities for optimization and reduce your overall environmental impact

Go to Create New Evaluation Tab

Carbon Footprint Saved Results

Export

Study ID	Therapeutic Area	Indication(s)	Trial Phase	Carbon footprint Type	Created By	Created On	Actions (Re-Create)	Actions (View)	Actions (Link)	Carbon footprint Score
filter data...										
Study 1 Study 2 Study 3 Study 4 Study 5 Study 6 Study 7 Study 8	IDV		Phase-1	totalKgCO2eSp3		12May24				991000
	IMM		Phase-1	totalKgCO2eSp3		21Jun24				2040000
	ONC		Phase-1	totalKgCO2eSp3		24sep24				387000
	NS		Phase-2	totalKgCO2eSp3		12Jun24				225000
	IDV		Phase-3	totalKgCO2eSp3		24sep24				18400
	ONC		Phase-3	totalKgCO2eSp3		24sep24				476000
	CVM		MAO	totalKgCO2eSp3		24sep24				111000
	ONC		MAO	totalKgCO2eSp3		24sep24				



Creating a new scenario

Identifying and updating a particular version

Sustainability Dashboard

[Home](#)[Create New Evaluation](#)

Select Category:
Cohort 1

Select Study:
Study 1

Select Version:
_v1

Refresh Scenario

Load Scenario

Countries & Sites

Drug Products

Trial management

Consumables & samples

Travel & meetings

Save Edited Content

Export JSON

Scenario Name:
_v1

Save Scenario

Generate Score

Step 1: Generate Score
Step 2: Refresh Scenario
Step 3: Name Scenario Version
Step 4: Save Scenario
Step 5: Load Scenario

After saving edited content, please type a new identifiable name and click on Save Scenario button
Kindly ensure to dynamically complete the Evaluation log & Decision Diary Workbook to save the scenario descriptions, design decions and provide other vital feedback.

Graphical analysis of Carbon emission score



Multi scenario design versions

Typically start with the Baseline Study, which has loaded historical data

Sustainability Dashboard

Home

Create New Evaluation

select Category:

Cohort 1

Generate Score

select Study:

baseline

baseline

_baseline

_v1_d101

_v1

_v1_23

Refresh Scenario

Trial management

Consumables & samples

Travel & meetings

information, including the number of participants and the number of site visits expected in each country.

country	number_of_participants	number_of_participants_site_visits	average_duration_of_visit	number_of_remote_visits	average_duration_remote_visits
United Kingdom	500	0	0	0	0
Belgium	7	0	0	0	0
Canada	11	0	0	0	0

Save Edited Content

Export JSON

Scenario Name:

_v1

Save Scenario

aved Scenario_v1 successfully

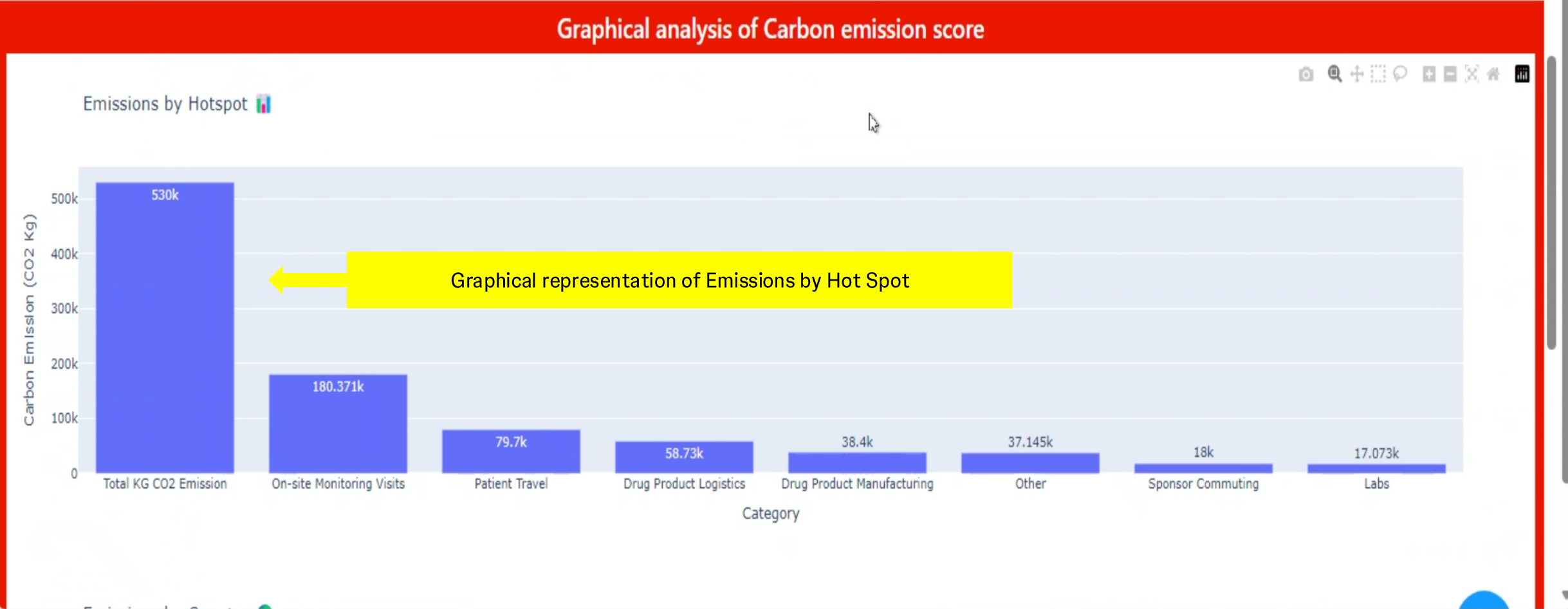
Kindly ensure to dynamically complete the Evaluation log & Decision Diary Workbook to save the scenario descriptions, design decions and provide other vital feedback.

Graphical analysis of Carbon emission score

Kudas, David [JRDUS NON-J&J]

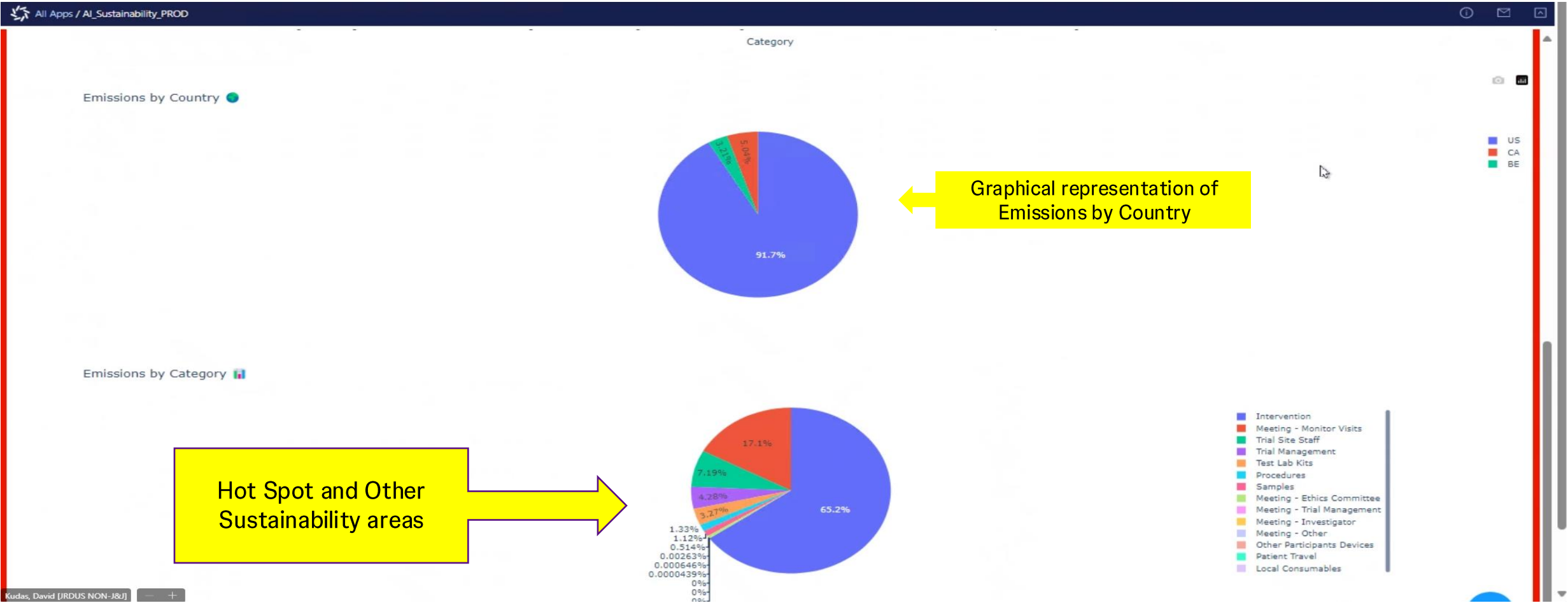
Graphics of Emissions

By Hot Spot and Total



Graphics of Emissions

Pie Graphs by Country and Targeted area



A perspective view of a metal bridge with a grid deck and blue railings, spanning a lush green forest. The bridge leads into a dense forest of tall trees and ferns. The text "ML based parametrized pre-filling for base design" is overlaid in white on the left side of the image.

ML based parametrized pre-filling for base design

AI-Powered Design Pattern Recognition and Pre-Fill Engine

- A novel feature of the platform is its XG Boost-based pre-fill engine.
- This AI model is trained on historical clinical trial data, including design parameters and emissions outcomes.
- When initiating a new trial, users are presented with intelligent defaults based on prior trials with similar characteristics.



Folder: Monitor Visits

Visits by country, on Site & virtual attendees

All Apps / AI_Sustainability_PROD

Countries & SitesDrug ProductsTrial managementConsumables & samplesTravel & meetings

Monitorvisits

This section is to define the monitor visit information, including the frequency and purpose of these visits during the trial.

do_you_have_meeting_data	total_meetings	avg_attendees_per_meeting	travelProfileName	profile_type	percentageOfAttendees
yes	0	0	France	in-person	12
hotelNightsPerAttendee	hoursOnline	travelProfileName	profile_type	percentageOfAttendees	hotelNightsPerAttendee
1	8	Italy	in-person	10	1
hoursOnline	travelProfileName	profile_type	percentageOfAttendees	hotelNightsPerAttendee	hoursOnline
8	Korea	in-person	10	1	8
travelProfileName	profile_type	percentageOfAttendees	hotelNightsPerAttendee	hoursOnline	travelProfileName
Poland	in-person	3	1	8	Spain
profile_type	percentageOfAttendees	hotelNightsPerAttendee	hoursOnline	travelProfileName	profile_type
in-person	23	1	8	Taiwan	in-person
percentageOfAttendees	hotelNightsPerAttendee	hoursOnline	travelProfileName	profile_type	percentageOfAttendees
3	1	8	UK	in-person	3
hotelNightsPerAttendee	hoursOnline	travelProfileName	profile_type	percentageOfAttendees	hotelNightsPerAttendee
1	8	US	in-person	5	1
hoursOnline	travelProfileName	profile_type	percentageOfAttendees	hotelNightsPerAttendee	hoursOnline
8	Virtual	virtual	31	0	2

Save Edited Content

Export JSON

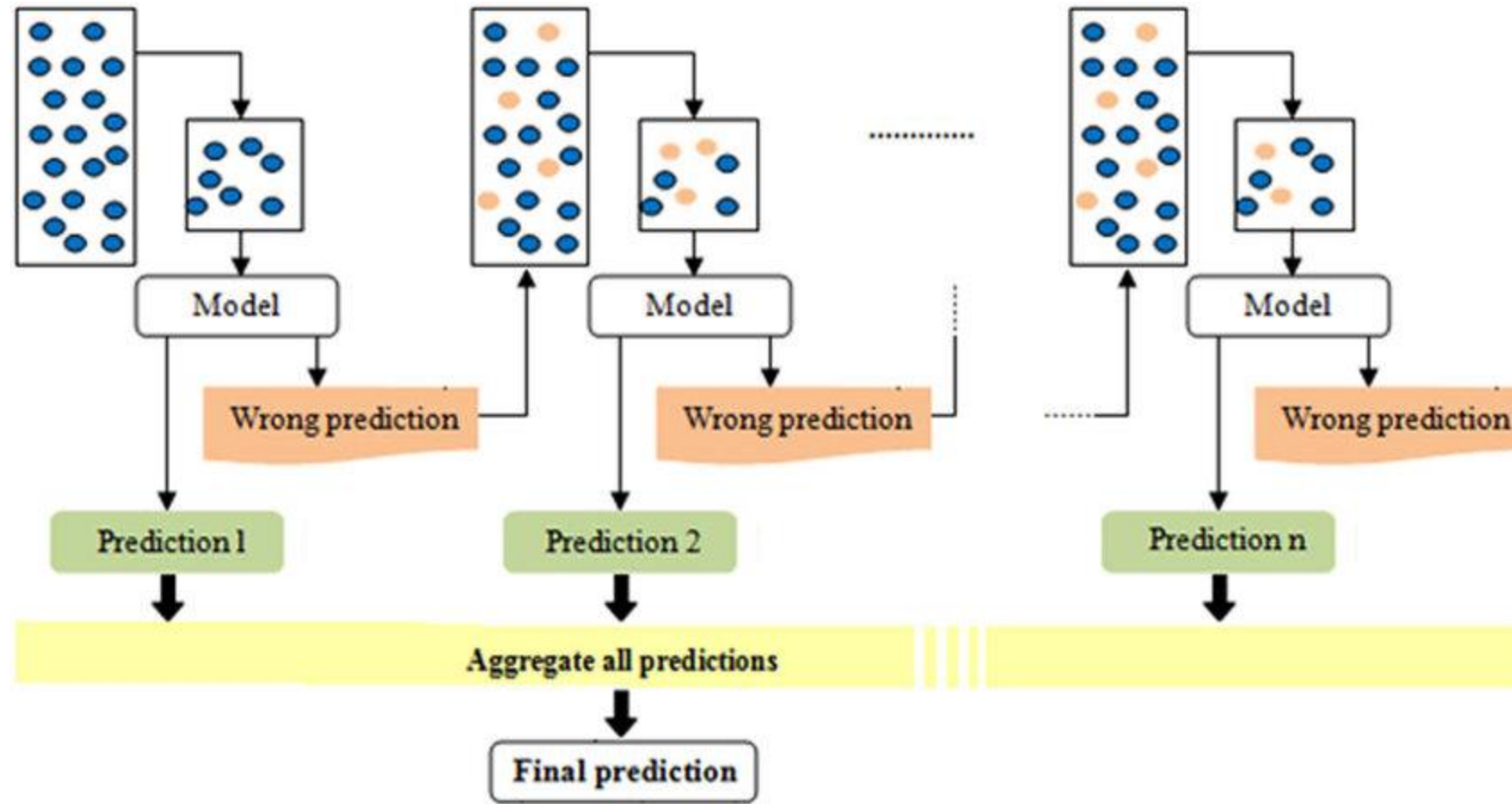
Scenario Name:

Save Scenario

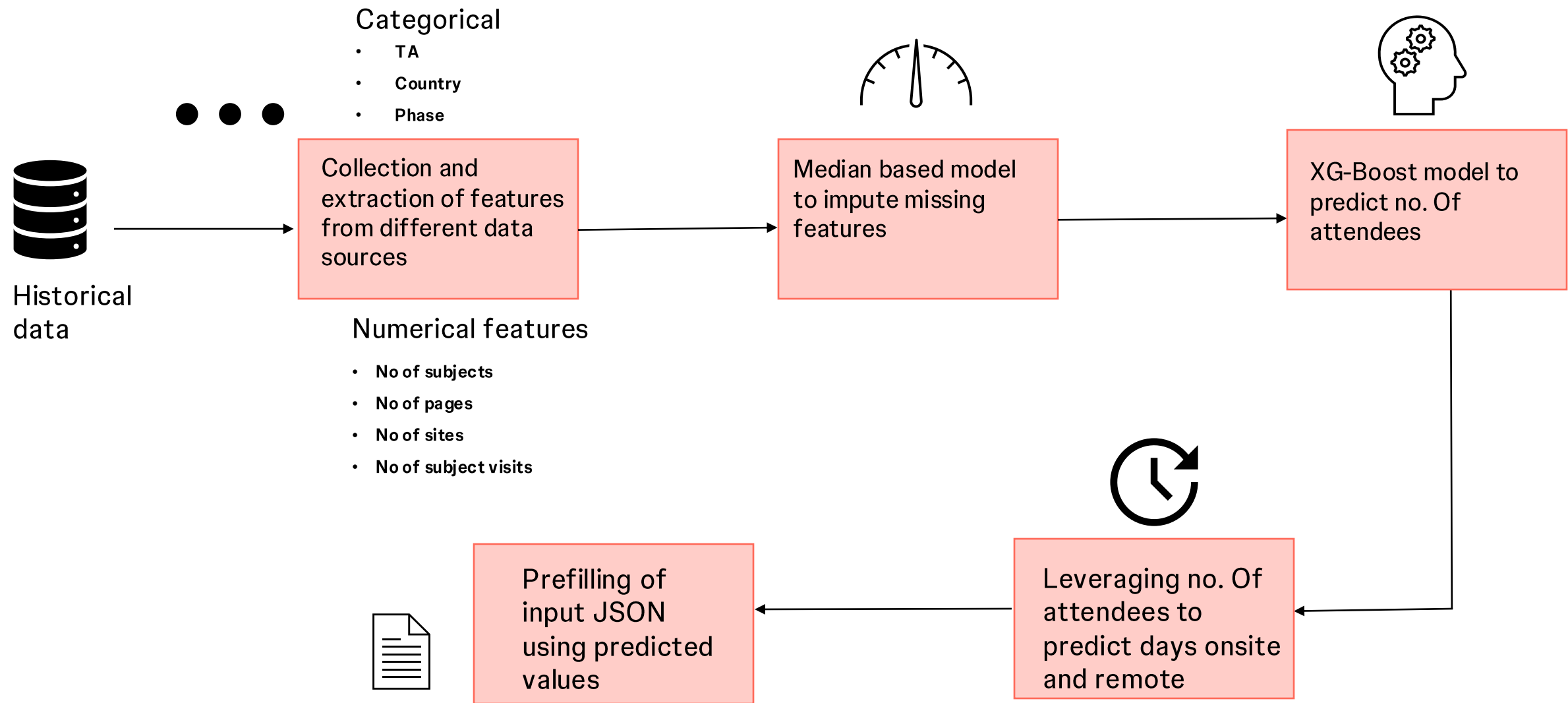
After saving edited content, please type a new identifiable name and click on Save Scenario button

Kudas David (IRDUS NON-J&J) will complete the Evaluation log & Decision Diary Workbook to save the scenario descriptions, design decions and provide other vital feedback.

XGBOOST modelling



Sequential Model



Sequential Model Performance

Test Set Performance (Best Model):
Mean Squared Error: 3562.6540
Root Mean Squared Error: 59.6880
Mean Absolute Error: 38.6947
R² Score: 0.8013

XGB modelling with grid search CV

Best parameters found:
colsample_bytree: 0.8
learning_rate: 0.1
max_depth: 3
n_estimators: 50
reg_alpha: 0
reg_lambda: 2.0
subsample: 0.8

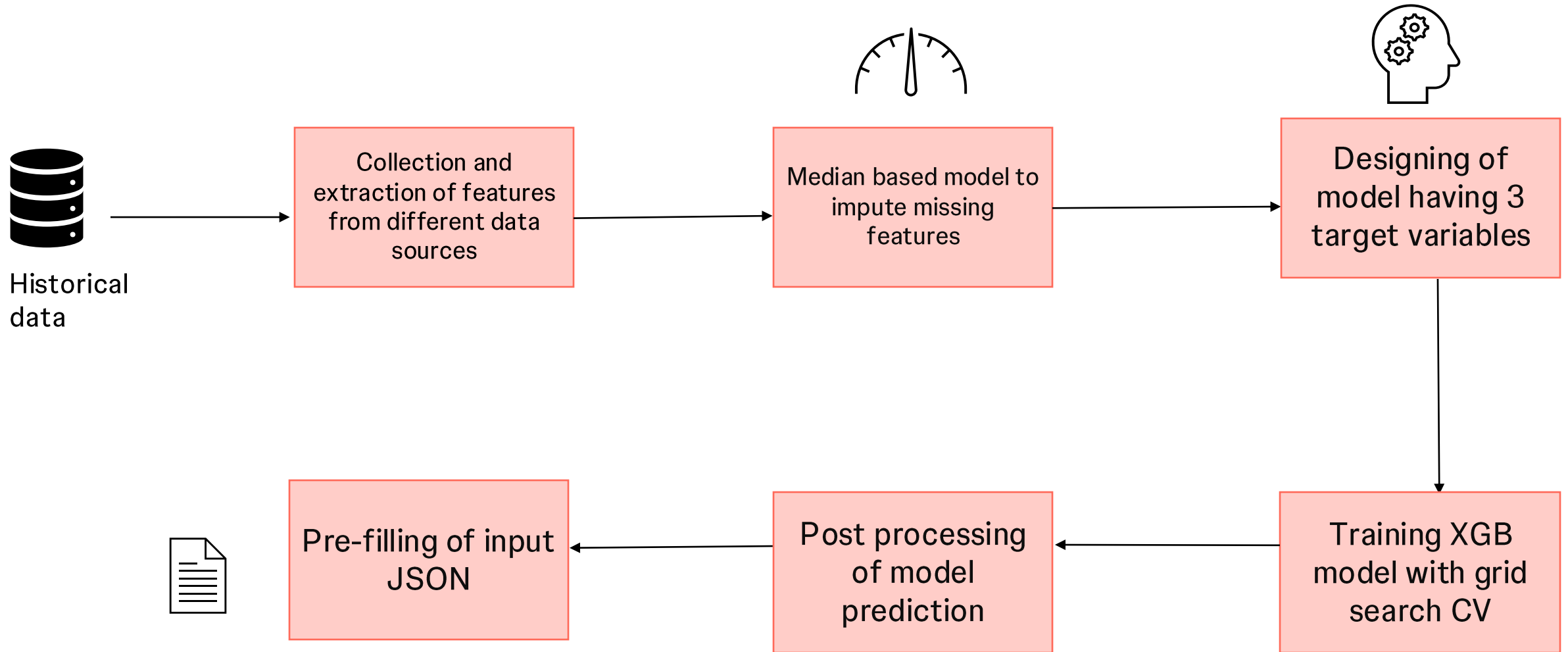
Sample predictions:

	Actual	Predicted	Difference	Abs_Difference
0	77	57.096973	19.903027	19.903027
1	37	13.706574	23.293426	23.293426
2	14	23.111717	-9.111717	9.111717
3	142	104.478912	37.521088	37.521088
4	71	42.048470	28.951530	28.951530
5	96	63.593102	32.406898	32.406898
6	15	10.261052	4.738948	4.738948
7	151	116.747238	34.252762	34.252762
8	291	187.485519	103.514481	103.514481
9	6	7.347262	-1.347262	1.347262

Multi-Target Model features

Numerical features	Categorical features	Target
therapeutic_area country_name top_phase	no_of_subjects pages no_of_sites num_of_subj_visits	no_of_attendees days_onsite days_remote

Multi-target Model



Multi-target model performance

XGB modelling with grid search CV

Best parameters found:

```
estimator__colsample_bytree: 0.8
estimator__learning_rate: 0.01
estimator__max_depth: 8
estimator__n_estimators: 200
estimator__reg_alpha: 0.1
estimator__reg_lambda: 2.0
estimator__subsample: 0.8
```

Performance for Target: no_of_attendees

```
Mean Squared Error: 3850.5904
Root Mean Squared Error: 62.0531
Mean Absolute Error: 35.3318
R2 Score: 0.7853
```

Performance for Target: days_onsite

```
Mean Squared Error: 1040.2957
Root Mean Squared Error: 32.2536
Mean Absolute Error: 18.5219
R2 Score: 0.9024
```

Performance for Target: days_remote

```
Mean Squared Error: 62.7578
Root Mean Squared Error: 7.9220
Mean Absolute Error: 4.5554
R2 Score: 0.9653
```

Sample predictions:

	Actual_no_of_attendees	Predicted_no_of_attendees	Actual_days_onsite	\
0	77	66.950668	61.50	
1	37	37.516006	51.50	
2	14	25.450508	75.50	
3	142	104.251564	106.75	
4	71	64.326286	62.50	
5	96	75.547119	70.25	
6	15	19.001345	53.00	
7	151	124.124550	98.50	
8	291	177.013687	112.00	
9	6	11.739861	13.00	

	Predicted_days_onsite	Actual_days_remote	Predicted_days_remote
0	66.203003	8.00	9.098045
1	60.829910	4.50	7.746911
2	69.781029	34.00	19.677444
3	91.414574	6.50	9.189528
4	68.913460	4.00	4.772542
5	71.513847	13.00	8.865481
6	59.683445	4.00	8.674540
7	79.784256	19.25	13.287177
8	105.519348	17.50	15.922030
9	39.439720	7.00	8.284510

References

References:

1. LaRoche J, Alvarenga R, Collins M, Costelloe T, De Soete W, Faludi J, Rens K, “Climate Footprint of Industry-Sponsored Clinical Research: An Analysis of a Phase-1 Randomized Clinical Study and Discussion of Opportunities to Reduce Its Impact”, BMJ Open 2024;14:e077129. doi:10.1136/bmjopen-2023-077129.
2. LaRoche J, Lanier J, Alvarenga R, Collins M, Costelloe T, Chau A, Weatherly H, De Soete W, Rens K, Faludi J, “Climate footprint of industry-sponsored in-human clinical trials: life cycle assessments of clinical trials spanning multiple phases and disease areas”, BMJ Open 2025;15:e085364. doi: 10.1136/bmjopen-2024-085364