



Phuse Americas Winter SDE

Accelerating the Last Mile of Clinical Trials by Leveraging AI/ML for Workflows

*Aditya Gadiko
December 1st 2022*



Agenda

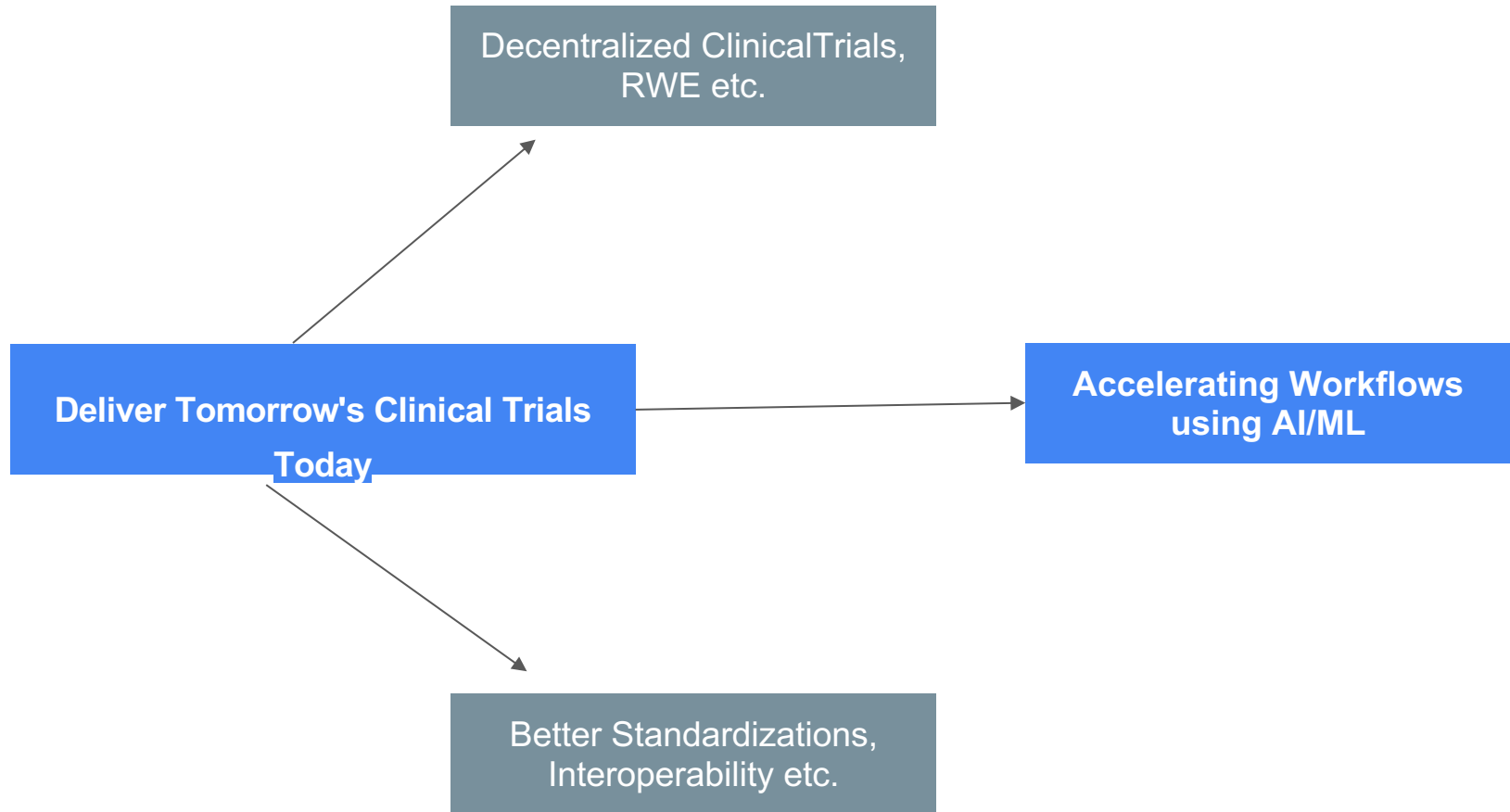


1. Introduction
2. AI/ML Research
3. Accelerate Workflow for *Data Managers*
4. Accelerate Workflow for *Medical Monitors*
5. Accelerate Workflow for *Statistical Programmers*

Speakers



Aditya Gadiko
Global Director Clinical Informatics
Saama



Graph Convolutional Networks

Vinayakumar Ramakrishnan, 2020-08-10 10:00:00

An Attentive Sequence Model for Adverse Drug Event Extraction from Biomedical Text

An Attentive Sequence Model for Adverse Drug Event Extraction

PHI Scrubber

Bayesian Optimization of Bose-Einstein Condensates

MedMCQA: A Large-scale Multi-Subject Multi-Choice Dataset for Medical domain Question Answering

MedMCQA: A Large-scale Multi-Subject Multi-Choice Dataset for Medical domain Question Answering

Ankit Patel, Laxesh Kumar Umashankar, Malakannan Sankaranarayanan Proceedings of the Conference on Health, Inference, and Learning, PHIL 110-119-190, 2022.

Interpretability

BioBERT NLI

Convolutional Neural Network in Classifying Cancer Through DNA Methylation

Pay Attention to the cough: Early Diagnosis of COVID-19 using Interpretable Symptoms Embeddings with Cough Sound Signal Processing NLI

BioELECTRA: Pretrained Biomedical text Encoder using Discriminators

BioELECTRA: Pretrained Biomedical text Encoder using Discriminators

Kamal Raj Kankarajan and Bhuvana Kundumani and Malakannan Sankaranarayanan
SAAMA AI Research Lab, Chennai, India
(kamal.raj, bhuvana.kundumani, malakannan.sankaranarayanan)@saama.com

Parking Lot Detection

Differential Privacy

GCAGE: Graph Clustering with Attentive Graph Auto-Encoders

Tabular Date GAN - TDGAN

Tabular Date GAN - TDGAN

Vinayakumar Subramanian, Bhuvana Kundumani, Malakannan Sankaranarayanan

Small-Bench NLP: Benchmark for small single GPU trained models in Natural Language Processing

Small-Bench NLP: Benchmark for small single GPU trained models in Natural Language Processing

Kamal Raj Kankarajan and Bhuvana Kundumani and Malakannan Sankaranarayanan
SAAMA AI Research Lab, Chennai, India
(kamal.raj, bhuvana.kundumani, malakannan.sankaranarayanan)@saama.com

Biomedical Language Understanding and Reasoning Benchmark

microsoft.github.io/BLURB/leaderboard.html

BLURB

Leaderboard Paper Models Tasks Submit News

The Overall score is calculated as the macro-average performance over tasks. Details can be found within [our publication](#).

Show entries

Rank	Model	BLURB Score (Macro Avg.)	Micro Avg.	NER	PICO	RE	SS	Class.	QA
1	BioELECTRA-Base — Saama Research 	82.60	83.19	86.67	74.13	81.44	92.76	84.20	76.38
2	PubMedBERT (uncased; abstracts + full text) — Microsoft Research 	81.45	82.13	86.27	73.72	80.15	92.31	82.62	73.61

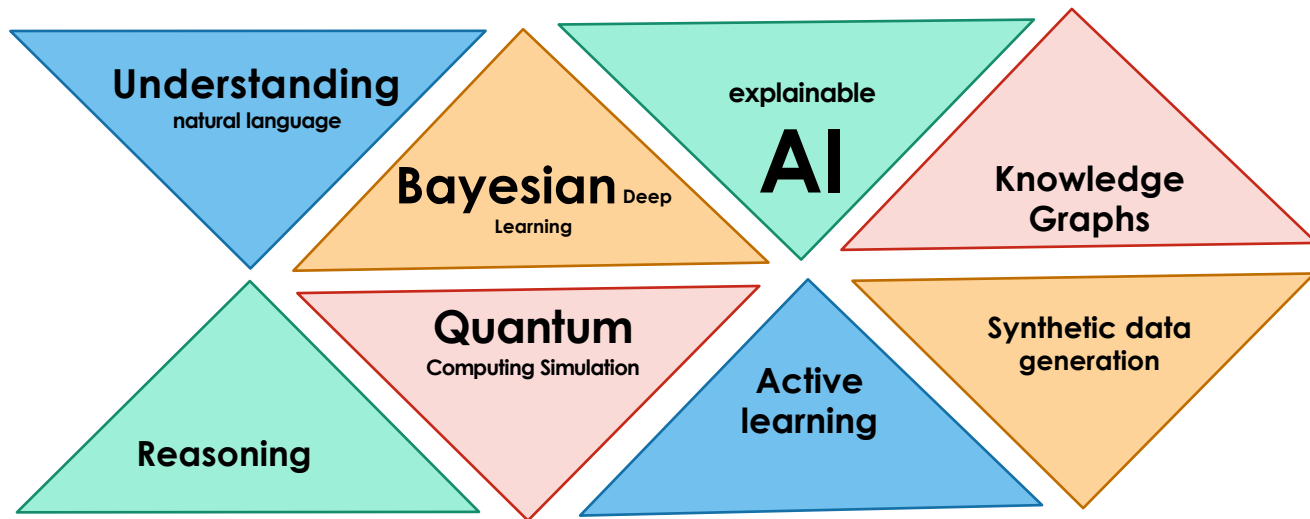
 sites.research.google/trc/publications/

Google Research About Spotlight Publications FAQ

[BioELECTRA:Pretrained Biomedical text Encoder using Discriminators](#)

Kanakarajan, Kundumani & Sankarasubbu

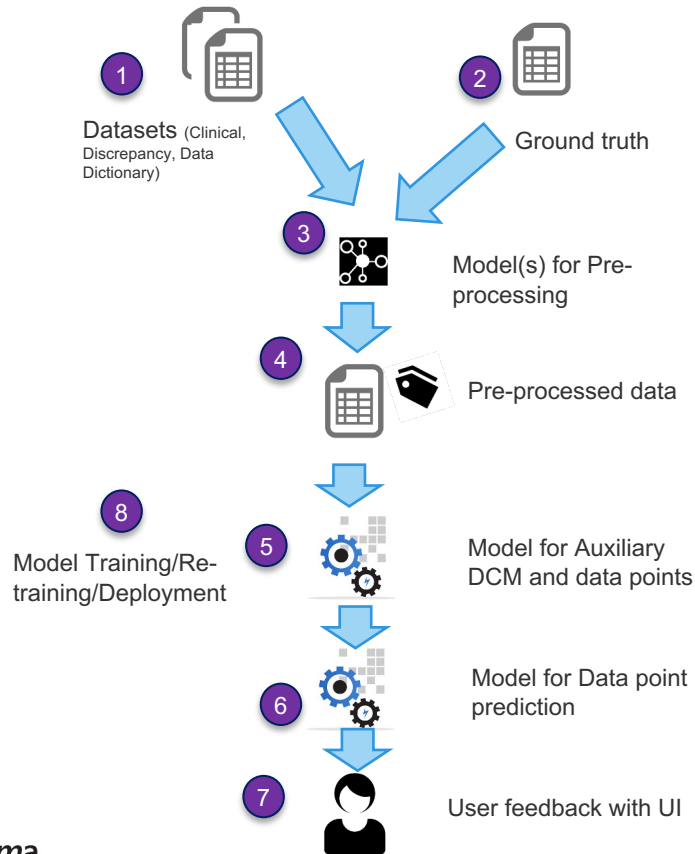
Research Focus for 2023



Leveraging AI/ML research to accelerate workflows for:

- 1. Data Managers**
- 2. Medical Monitors**
- 3. Statistical Programmers**

Data Managers - AI/ML



S.No	Component
1	Datasets (Clinical , Discrepancy and Data dictionary)
2	Ground truth dataset
3	Pre-processing including Models that was used to propagate ground truth to all clinical datasets
4	Propagated datasets based on ground truth
5	Model # 1 – Prediction of auxiliary DCM to identify auxiliary data points
6	Model #2 – Prediction of 3 output variables based on auxiliary data points
7	UI for user feedback on Output prediction
8	Model Training, Re-training and Deployment

Accelerate Data Managers Workflow using AI/ML (Video)

The screenshot displays the Saama SDQ (Study Data Query) interface for 'Account-1'. The top navigation bar includes a menu icon, the account name 'Account-1', the 'SDQ' label, a 'Study' dropdown set to 'test_study_2dot0', and a user profile icon. The left sidebar contains navigation links: 'Instance UAT', 'Dashboard' (selected), 'To Do List', 'Prediction history', 'Query Log', 'Study DQ', and 'Reports'. The main dashboard area features a 'Dashboard' header with a refresh icon and the text 'Get a quick glance of predicted discrepancies here'. Below this, two refresh dates are shown: 'EDC DATA REFRESH DATE 24-Feb-2022 11:52:40' and 'MODEL OUTPUT REFRESH DATE 29-Nov-2022 20:25:00'. Three key metrics are presented in cards: 'UNREVIEWED PREDICTIONS 193' with a warning icon, 'PREDICTIONS ON HOLD 21' with a checkmark icon, and 'TOTAL PREDICTIONS 378' with a list icon. The bottom section is divided into two panels. The left panel, titled 'PREDICTION By confidence score', shows a large green circle, likely representing a data visualization. The right panel, titled 'PREDICTION By aging', includes filters for '7 Days', '30 Days', and '60 Days', and displays the text 'No Aggregate Data Available'.

Account-1 SDQ Study test_study_2dot0

Instance UAT

Dashboard

Get a quick glance of predicted discrepancies here

EDC DATA REFRESH DATE 24-Feb-2022 11:52:40 MODEL OUTPUT REFRESH DATE 29-Nov-2022 20:25:00

UNREVIEWED PREDICTIONS 193

PREDICTIONS ON HOLD 21

TOTAL PREDICTIONS 378

PREDICTION By confidence score

PREDICTION By aging

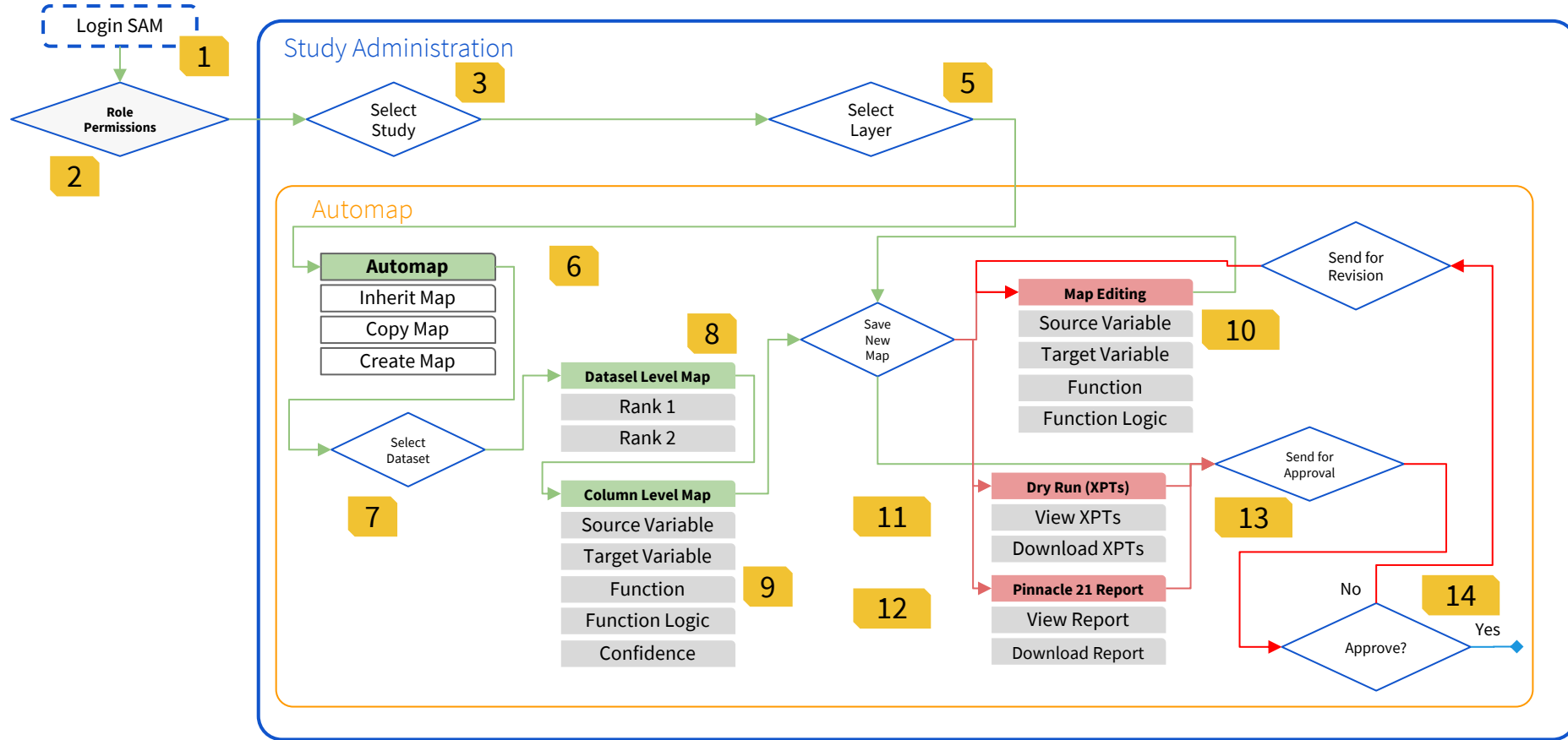
7 Days 30 Days 60 Days

No Aggregate Data Available

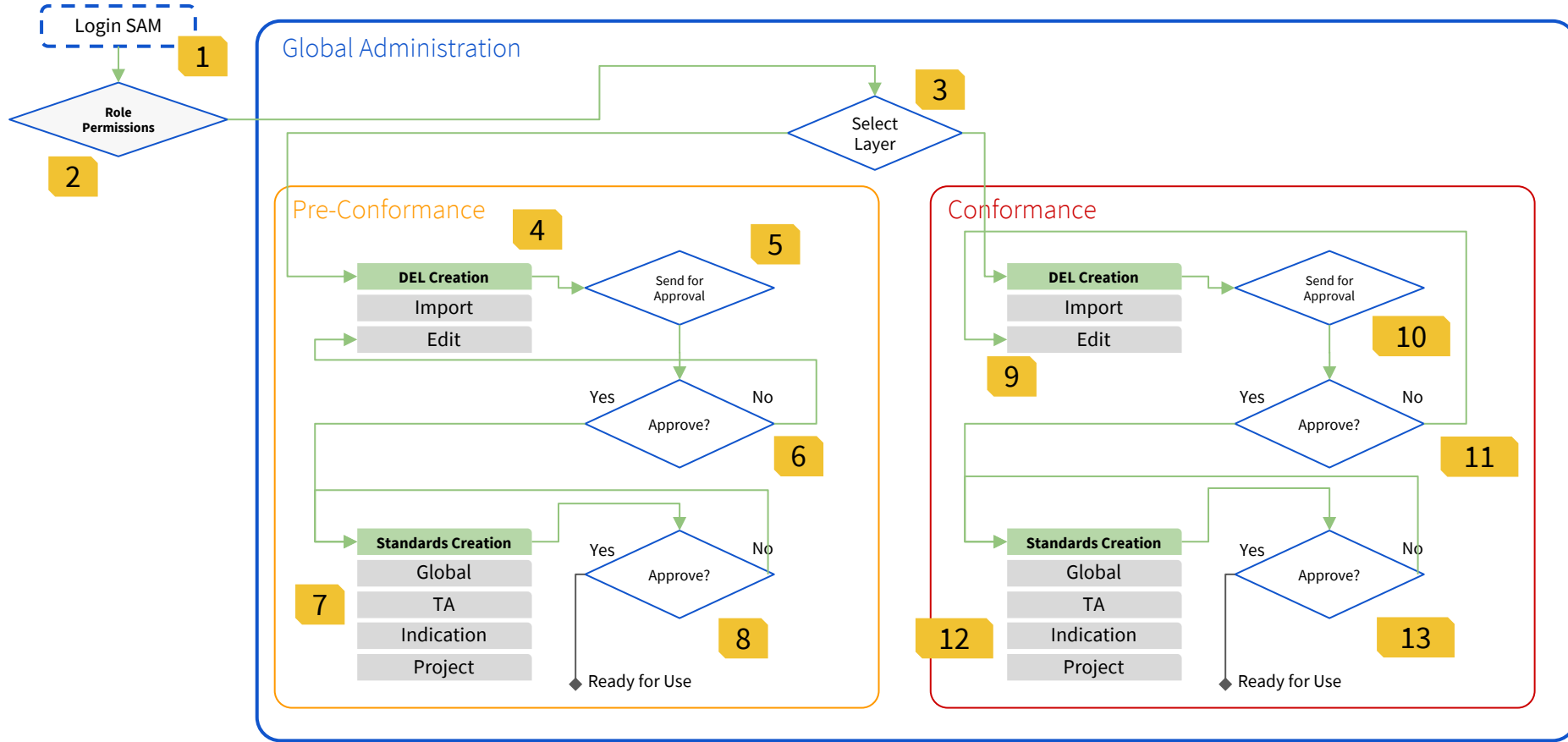
Accelerate Medical Monitors Workflow using AI/ML (Video)



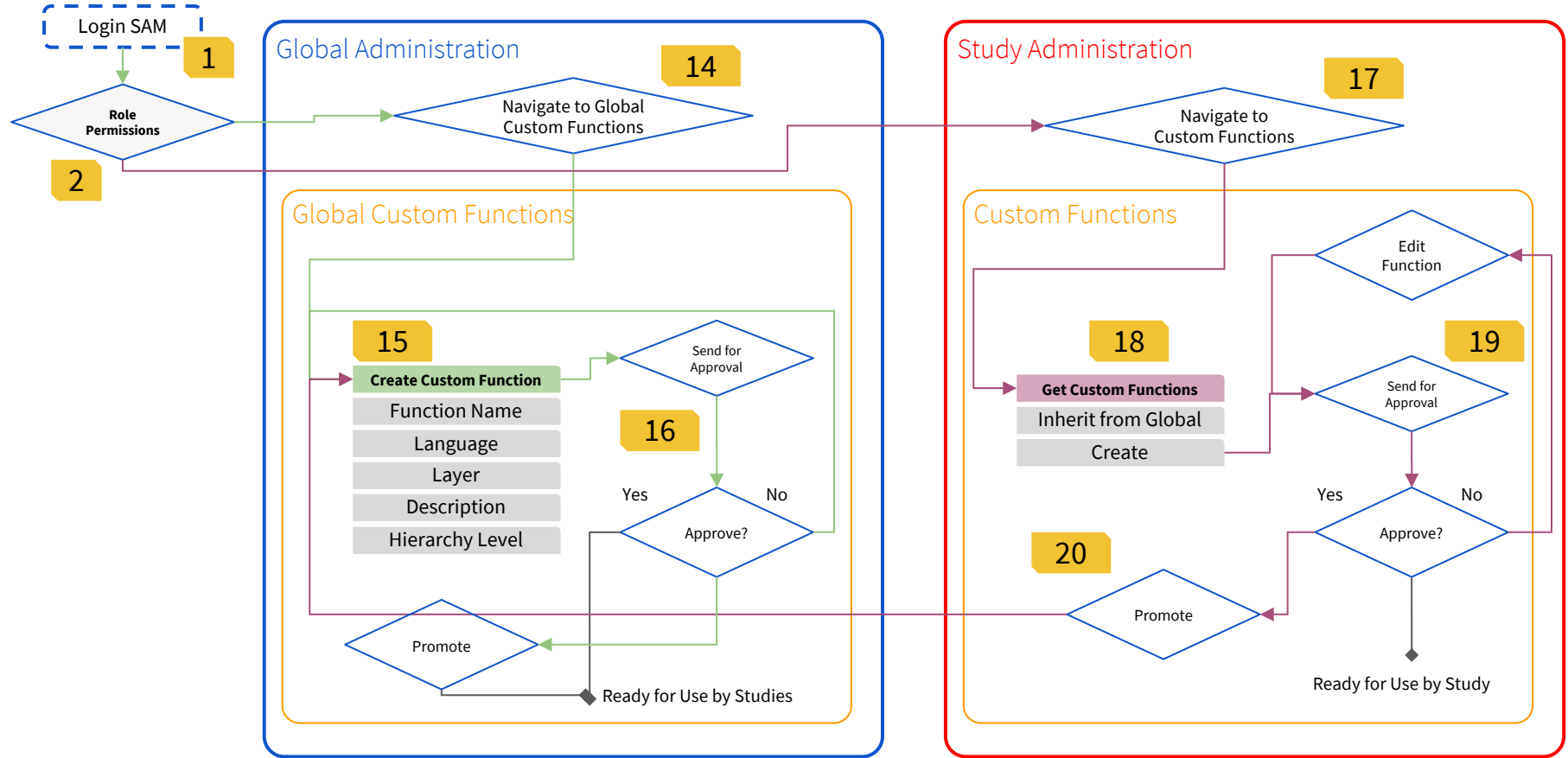
Statistical Programmers - AI/ML



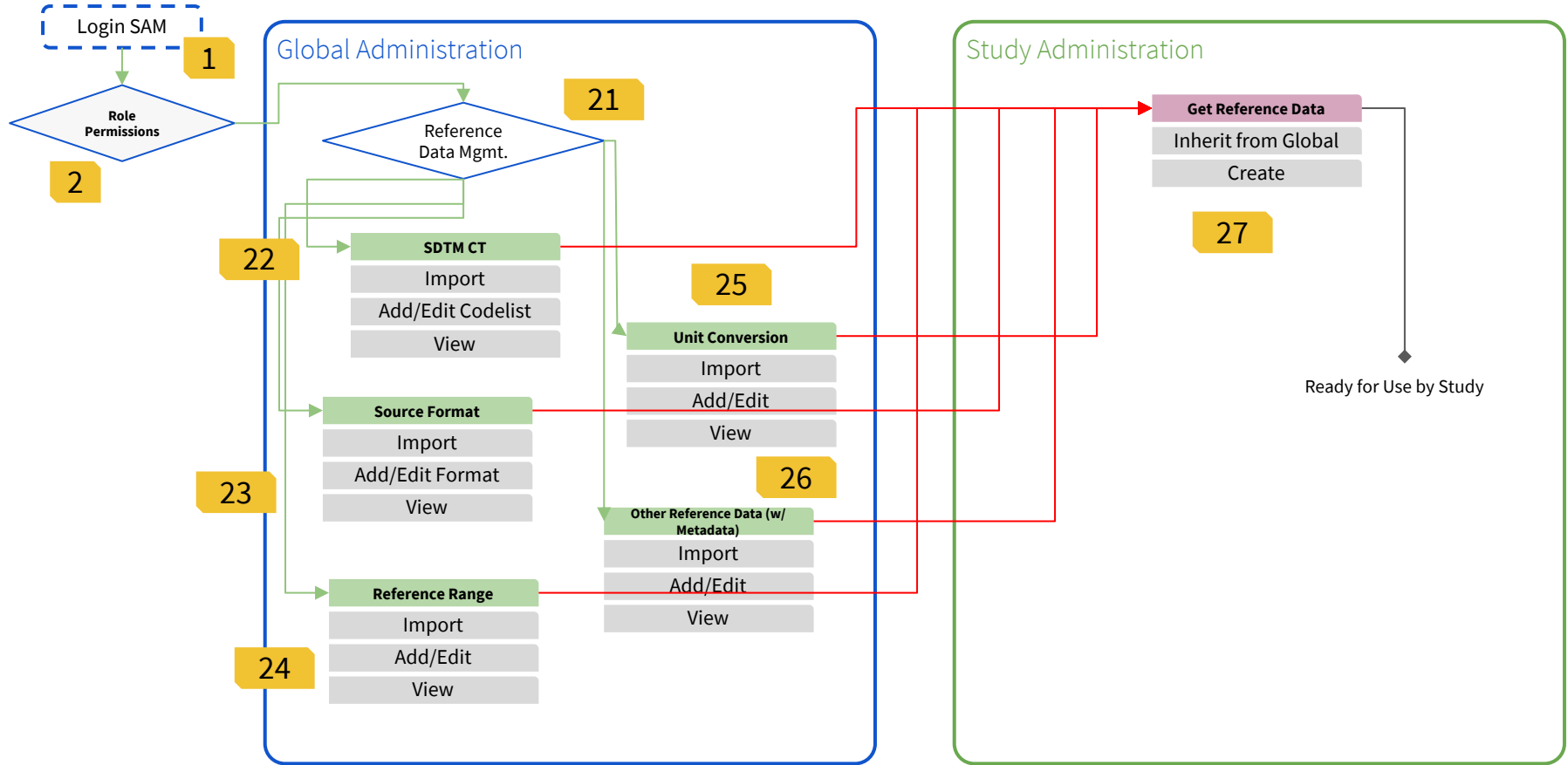
Statistical Programmers - AI/ML



Statistical Programmers - AI/ML



Statistical Programmers - AI/ML



Accelerate Statistical Programmer's Workflow using AI/ML (Video)

