



ET15: Front-Loading Statistical Programming Using Automated Synthetic Data

Synthetic data for the EDC system

PHUSE US Connect 2024 - February 28th

Faezeh Moshfeghifar

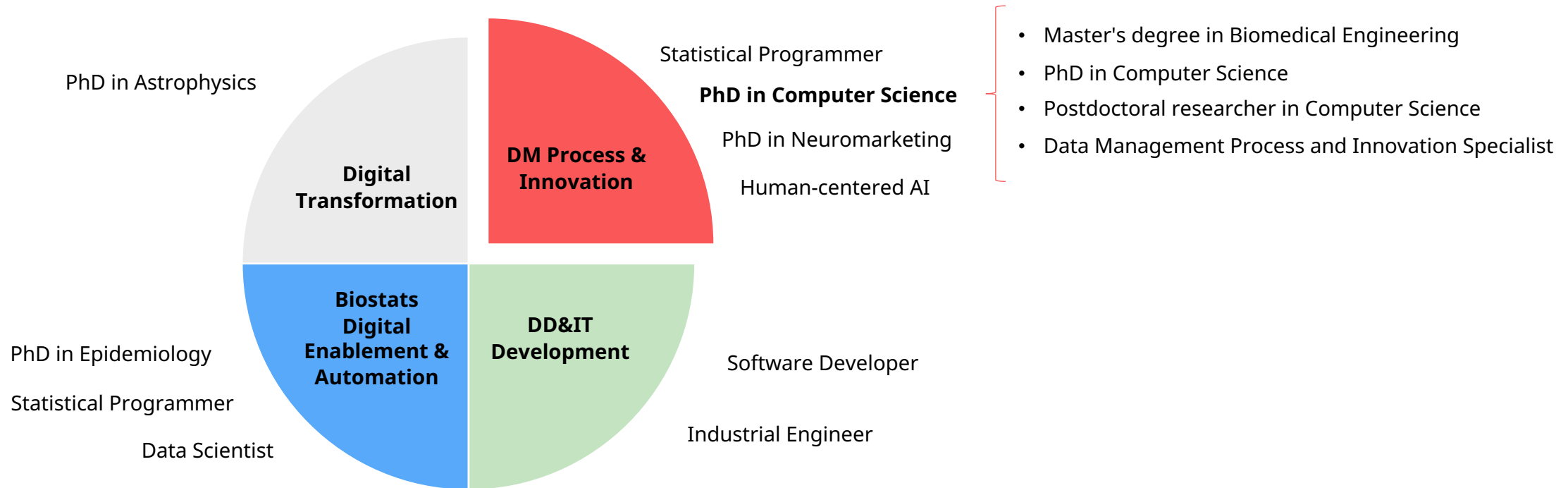
FZMG@novonordisk.com

Disclaimer: Views and opinions expressed are those of the speaker and not Novo Nordisk

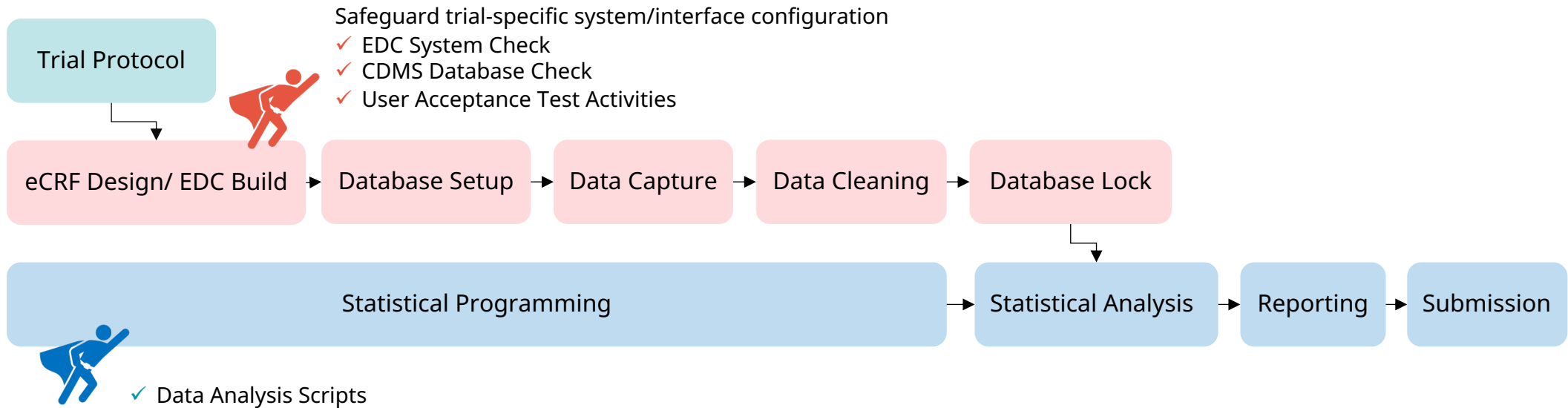
Outline

- Our Team & Target Audience
- Motivation & Problem Statement
- Technical Overview: System Architecture & Methodology
- System Demo: Solution in Action
- Conclusion & Future Work

Data Science Automation Group



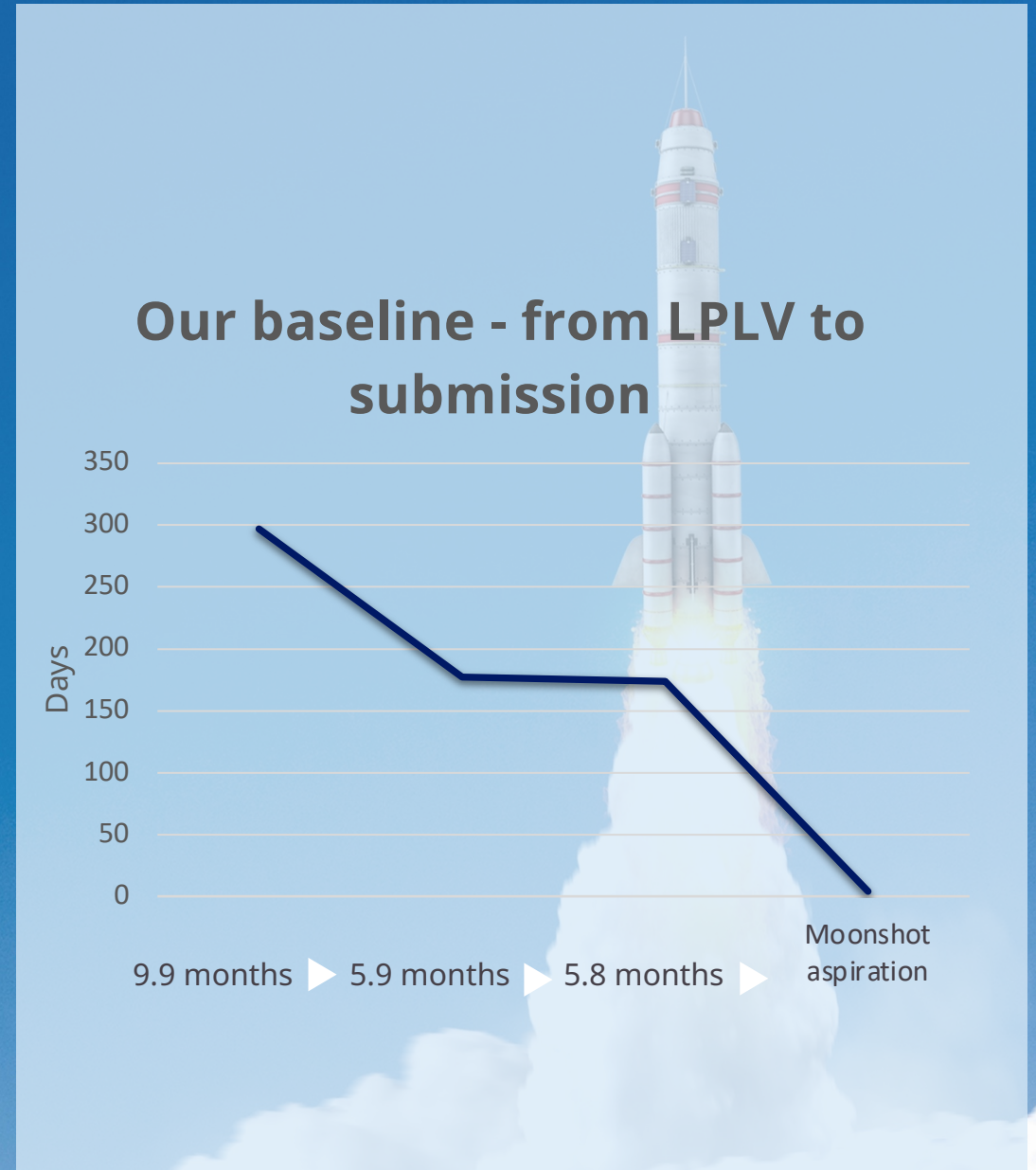
Dataflow in Clinical Trials



Fit4take-off

- LPLV to Submission in 4 days
- To the 4 key markets
- With successful access and launch 4 days after approval

Abbreviation: LPLV = Last Patient Last Visit



EDC System Check



TRN: XXX - Case Report Books

Home Enroll Subjects Queries Monitor Signatures Review Reports Archives Custom Events Deployment

Status: All

Highlight: ☐ Started ☐ Incomplete ☐ Queries ☐ Frozen ☐ Locked

Site	Subject	Status	Visit 1	Eligibility	Visit 2	Visit 3	Visit 4	Visit 5	Visit 6	Adverse Events	Concomitant Medication	Temp IMP	Discontinuation
	01 (---)	Enrolled	?		?								
	001 (AAA)	Complete	?		?	?	?	?	?	?	?	?	?
	002 (BBB)	Not Complete	?	?	?	?	?	?	?	?	?	?	?
	003 (CCC)	Not Complete	?	?	?	?	?	?	?	?	?	?	?
	004 (DDD)	Not Complete	?		?								
	005 (EEE)	Not Complete	?		?								
	006 (FFF)	Not Complete	?		?								
	007 (GGG)	Not Complete	?		?								
	008 (HHH)	Not Complete	?		?								
	009 (III)	Not Complete	?		?								
	01 (ZZZ)	Enrolled	?		?								
	010 (JJJ)	Not Complete	?	?	?								

Oracle Inform

V1 Eligibility V2 V3 P4 V5 V6 V7 V8 P9 V10 P11 V12 P13 V14 P15 V16 P17 V18 AE

Visit: V3 Date of visit

Study ID:

Forms: DoV VS Lab

1.* Date of visit

2.* Contact type

3.* Was the visit impacted by COVID-19?

Site visit
Telephone contact
Visit entered in error
Remote video contact
Off-site visit
Visit missed

Yes
No

Tick 'Yes' if the contact type was changed, the visit date was rescheduled outside the visit window or missed due to COVID-19 situation

Review [Select Action] Apply Submit & Next Submit Return

Items & Dependencies

Informed Consent and Demography

Study I

Informed Consent

1.* Date informed consent obtained / /

Demography

2.* Date of birth /

3. Sex ☒ Male ☐ Female
(Transferred from IWRS/RTSM)

4.* Subject self-reported ethnicity ☒ Hispanic or Latino ☐ Not Hispanic or Latino ☐ Not reported

5.* Subject self-reported race (Select all that apply, but at least one)
☐ American Indian or Alaska Native
☒ Asian
☐ Black or African American
☒ Native Hawaiian or Other Pacific Islander
☒ White
☐ Not reported

6. Subject No.

7. Hidden item – used for IMPACT interface for SCREEN visit (Visit 1) / /

Date/Time CodeList CheckBox TextBox

Item Types

SiteSubject	Status	Visit 1a	Eligibility	MedHx	Visit 2	Visit P3	Visit 4	Dose Reduced	IMP Discont-Resume	End of Treatment	End of Study	Elevated Enzymes	Future Research WD	Complaint	Book Sign Off	Pregnancy
118 (---)	Not Complete															
119 (---)	Enrolled															
120 (---)	Enrolled															
121 (---)	Not Complete															
122 (---)	Enrolled															
123 (---)	Enrolled															
124 (---)	Not Complete															
125 (---)	Not Complete															
126 (---)	Not Complete															
127 (---)	Complete															
128 (---)	Not Complete															
129 (---)	Not Complete															

- ☐ Subject is eligible (Meets all eligibility requirements)
 Date subject is confirmed eligible
 / /
- ☒ Subject failed one or more eligibility requirements (Subject is a screen failure)
- ☐ Eligibility evaluation was not completed

Dynamic Triggers (Rules)

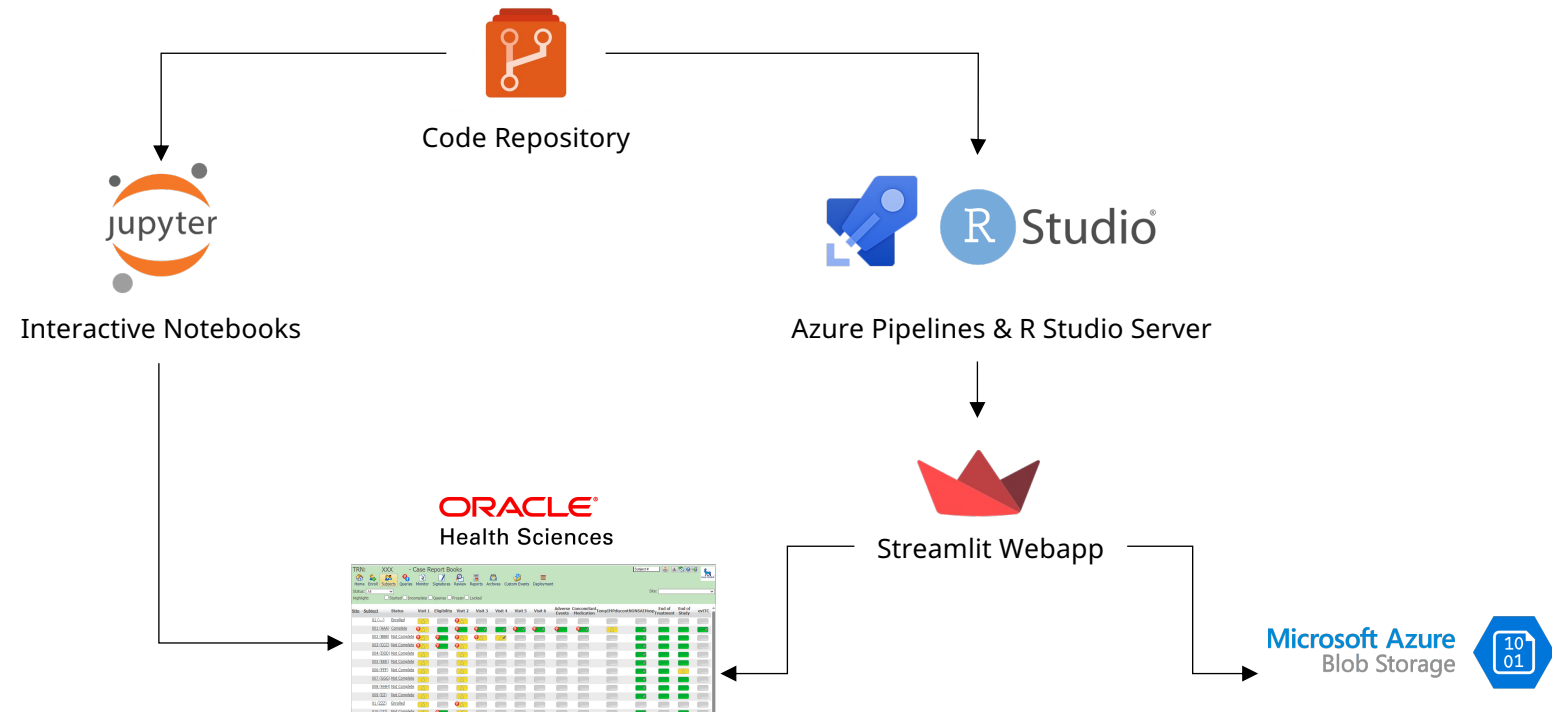
- Manual data entry is labor-intensive and prone to error
- Lack of overview over the *item-* and *rule- coverage* percentage
- In collaboration with industry partners



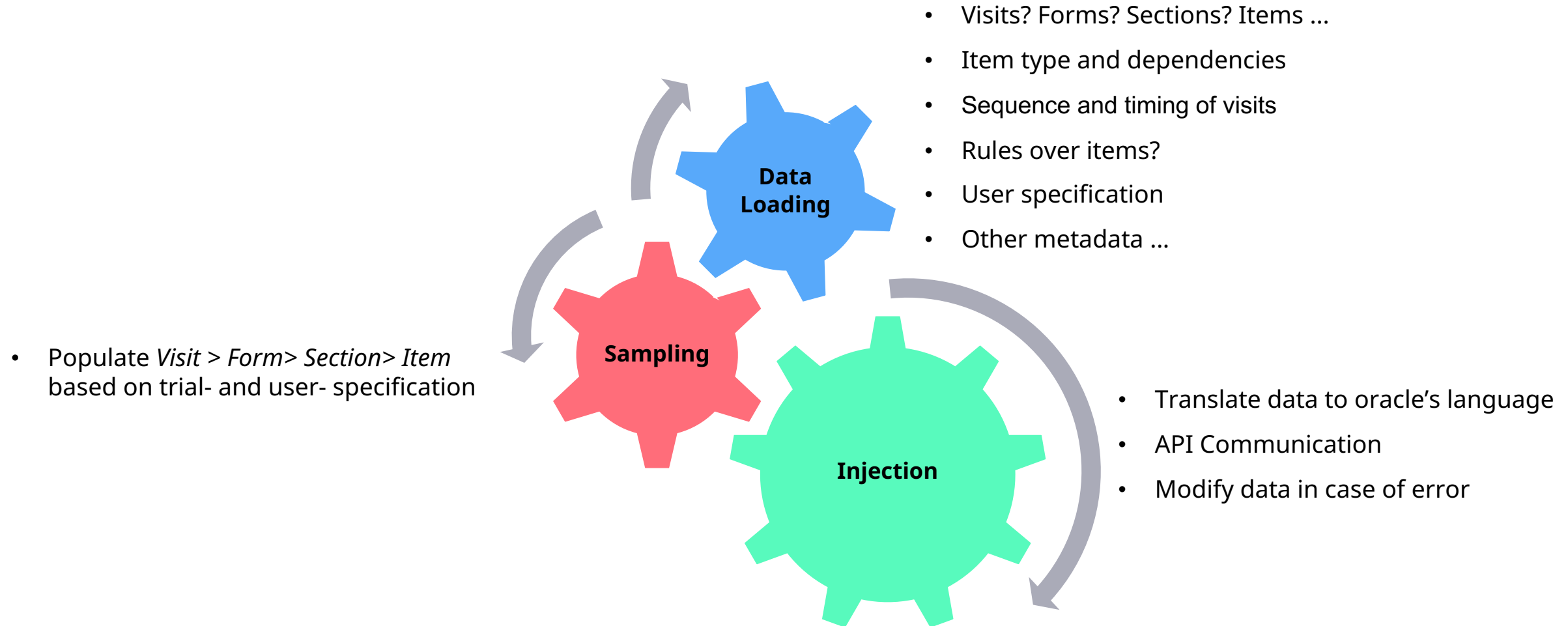
Automated synthetic data

System Architecture

- Python Environment



Important Modules



System Demo

Welcome to Cassandra - local

Kassandra generates test data for your trial. You can:

- inject test data into Inform/EDC and analyse the results
- generate loading files for external data and analyze the results

Need help? Find it on the [Sharepoint site](#)

Workaround for a stable experience with Cassandra

Problem:

Kassandra has a long runtime, and the app resets if the window is closed, screen is locked. This combination makes it difficult to run multiple patients through Kassandra reliably.

Workaround:

Use virtual desktops Cloud Desktop B as a temporary solution until the Kassandra team has a permanent one. You will access the same instance of Kassandra, however Cloud Desktop B allows you to leave the window with the app without it stopping.

How-to:

- Request access to Cloud Desktop B - Access request (sharepoint.com) via [Novoaccess](#).
 - Search string: "Cloud Desktop : B - Clinical Systems Desktop - CSD".
 - Remember to add comment with the reason for requesting access - otherwise it will be denied by default.
- Go to [Citrix Workspace](#) and open the Cloud Desktop B service - log on using your credentials.
- Go to the Kassandra app via the browser and use Kassandra as you normally would.
 - You can close the window for the virtual desktop after you have started the injection in Kassandra. Later, you can open the virtual desktop again to check the progress.
 - Remember, you can also always check the progress by reviewing how many patients have been injected into the EDC.

SiteSubject	Status	Visit 1a	Eligibility	MedHx	Visit 2	Visit P3	Visit 4	Visit 5	Visit 6	Visit 7	Visit 8	Visit P9	Visit 10a	Visit 10b
0183 (---)Not Complete	Not Complete	Not Complete	Not Complete	Not Complete	Not Complete	Not Complete	Not Complete	Not Complete	Not Complete	Not Complete	Not Complete	Not Complete	Not Complete	Not Complete
0184 (---)Not Complete	Not Complete	Not Complete	Not Complete	Not Complete	Not Complete	Not Complete	Not Complete	Not Complete	Not Complete	Not Complete	Not Complete	Not Complete	Not Complete	Not Complete
0185 (---)Not Complete	Not Complete	Not Complete	Not Complete	Not Complete	Not Complete	Not Complete	Not Complete	Not Complete	Not Complete	Not Complete	Not Complete	Not Complete	Not Complete	Not Complete

Page 13 of 13 Rows: 15 Total: 183 Show Transferred

Conclusion & Future Work

- ✓ Eliminating manual data entry
- ✓ Reduce test time and improve the core data management activities
 - Synthetic external data (e.g., Lab Data Specification)
 - Front-load programming activities before FPFV with early access to synthetic SDTM data
 - Data variation: Missing data related to endpoints
 - Realistic number of synthetic patients

