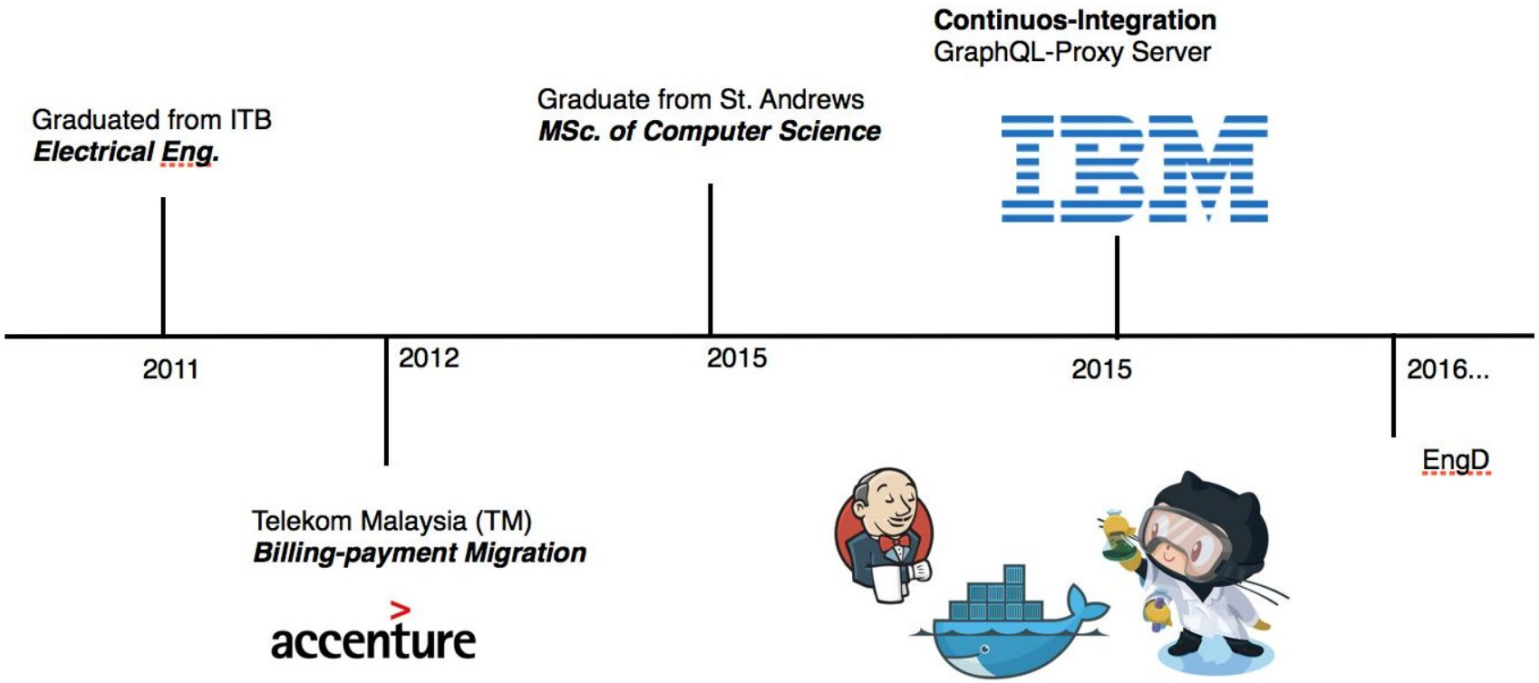
Research Fellow - AR2202RHM

Agastya Silvina

Background


Continuos-Integration
GraphQL-Proxy Server
Graduated from ITB
Electrical Eng.
Graduate from St. Andrews
MSc. of Computer Science
IBM
2011
2012
2015
2015
2016...
EngD
Telekom Malaysia (TM)
Billing-payment Migration
accenture

EU H2020 SERUMS Contributions

Front-End

Visualisation
Reporting Service
Dashboard

Back-End

Migration
API Gateway

Toxicity Predictor
(1,2,3,4,....)
ML, NLP, DL

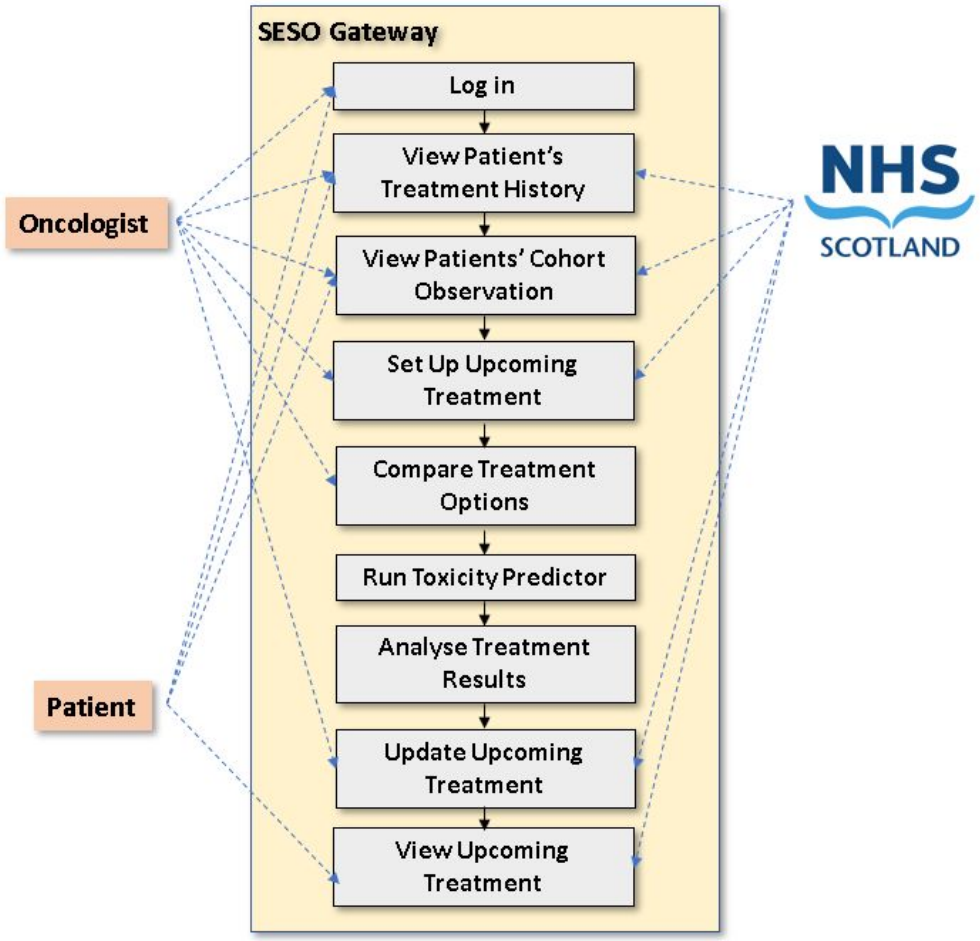
PoC

Issues...
Issues...

Synthetic Data

IBM

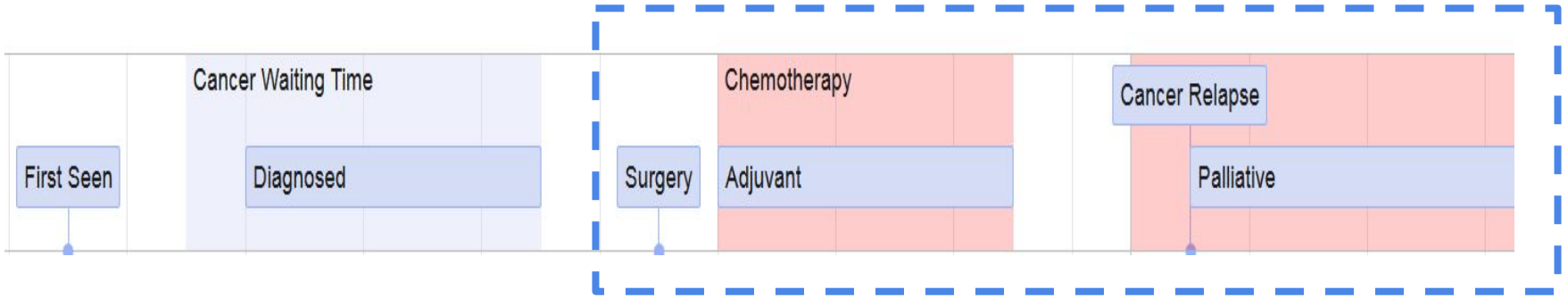
Toxicity Predictor



Synthetic Data

- We use IBM Constraint Solver
  - Determining the rules

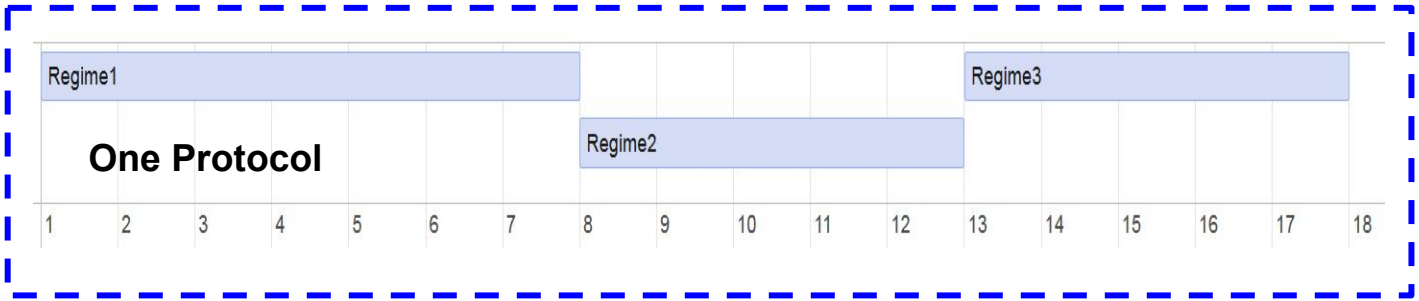
Patients' Treatment Pathway



- A patient can only be treated with one intention at a time (e.g., Adjuvant)
- After a specific time has passed, the patient might be treated with other treatments with different intentions (e.g., Palliative, Curative)

Treatment Regimes

- Each intention has different regimes.
- Each regime has several different drugs.
- The treatment may last for several weeks or months
- A patient may be treated with several regimes at time.
- Each regime has one or more treatment cycles.
- Several different regimes may belong to one protocol.



How does the table represent the treatment?

CHI	APPOINTMENT DATE	INTENTION	REGIME	DRUG	CYCLE
patient1	1/12/2019	Adjuvant	Regime A	drug1	1
patient1	1/12/2019	Adjuvant	Regime A	drug2	1
patient1	7/12/2019	Adjuvant	Regime A	drug1	2
patient1	7/12/2019	Adjuvant	Regime A	drug2	2
patient1	14/12/2019	Adjuvant	Regime A	drug1	3
patient1	14/12/2019	Adjuvant	Regime A	drug2	3

- First Appointment Date

currentDate - (5*365) < first_appointment_date <= currentDate - 250

- Unique CHI

allDiff(from(general), chi)

- Each regime may have more than one cycles

numOf(from(smr01),cycle.regime = regimen.regime) = randomWeightedNumber(8)

Synthetic data validation

- How to differentiate real and fabricated data?
- We need **tools to differentiate between the real and fabricated data.**
- We are working with the researchers from UCL to determine the quality of our fabricated data.

THANK YOU