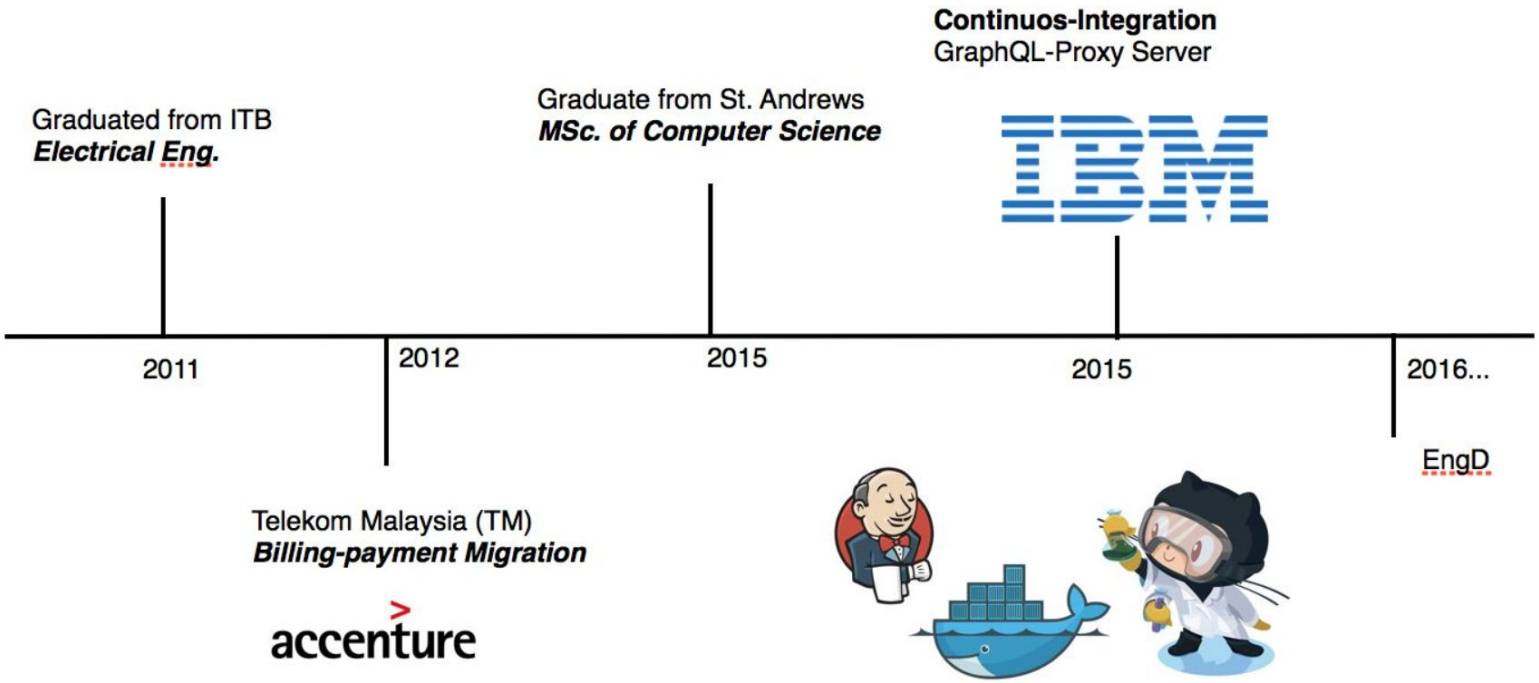
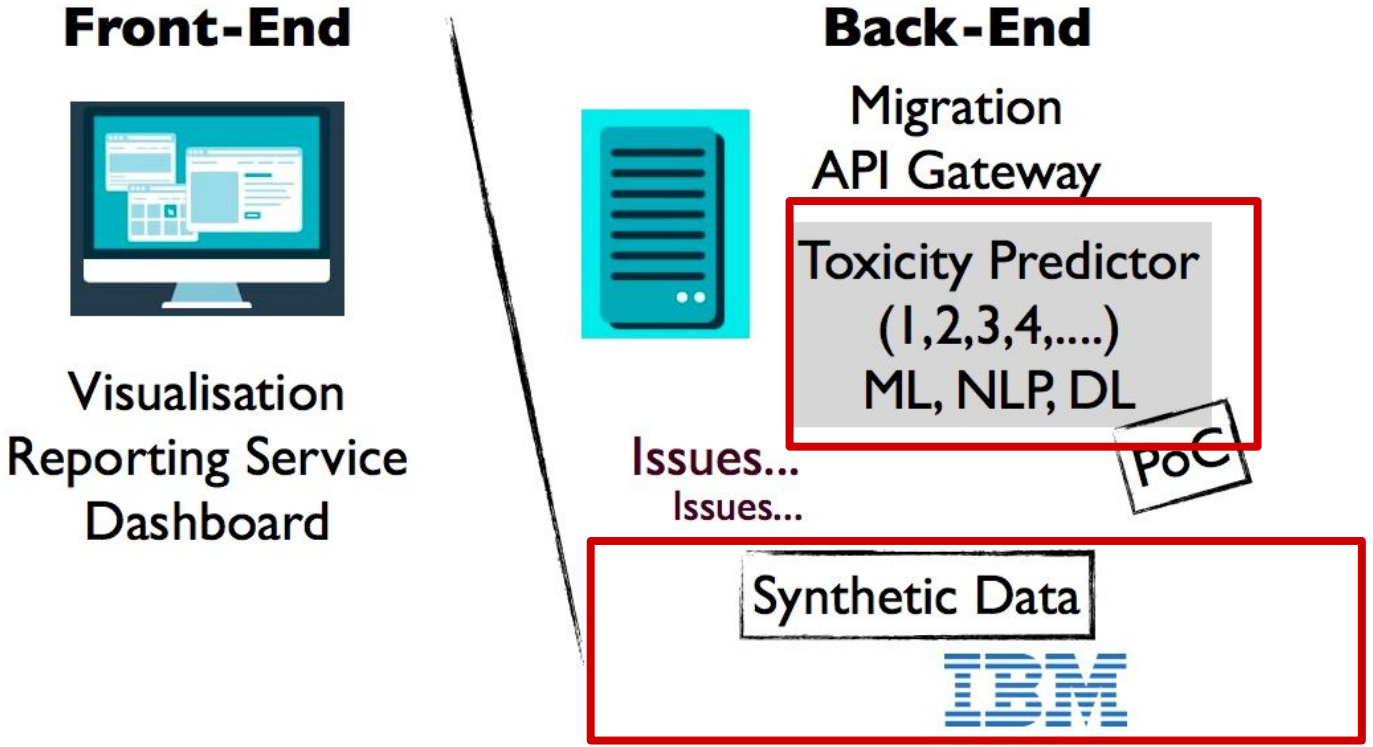
Research Fellow - AR2202RHM

Agastya Silvina

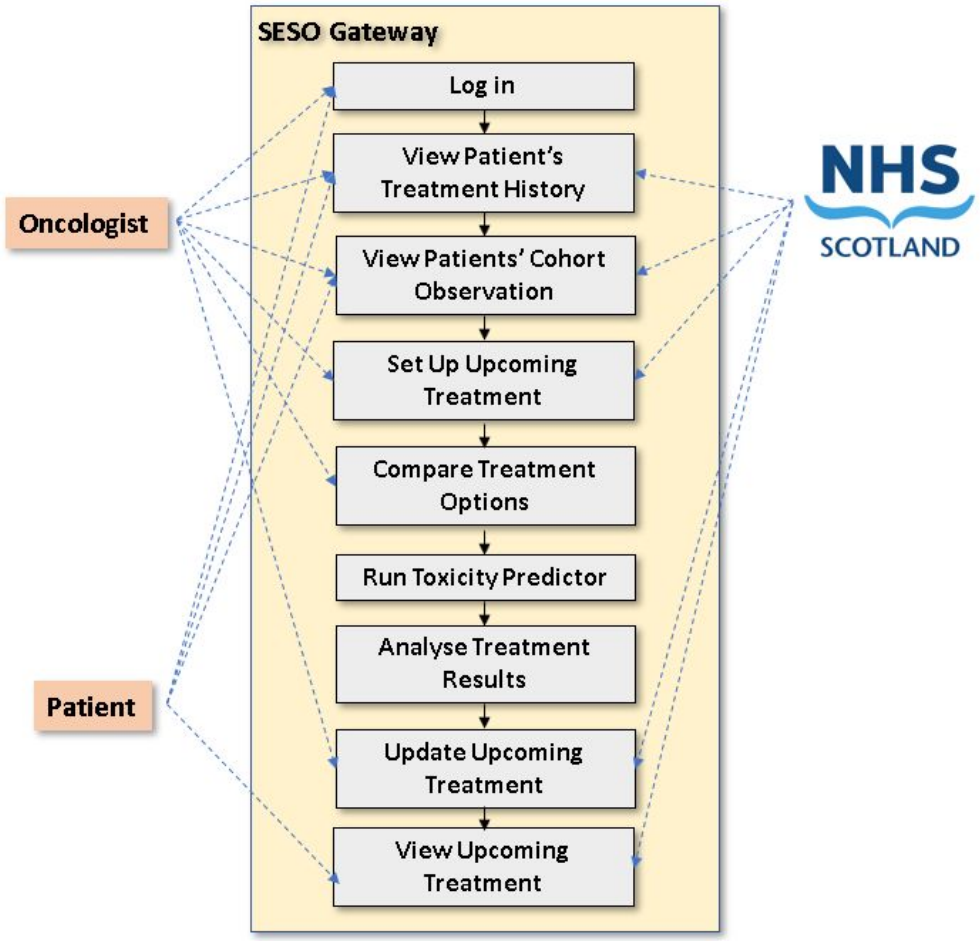
Background

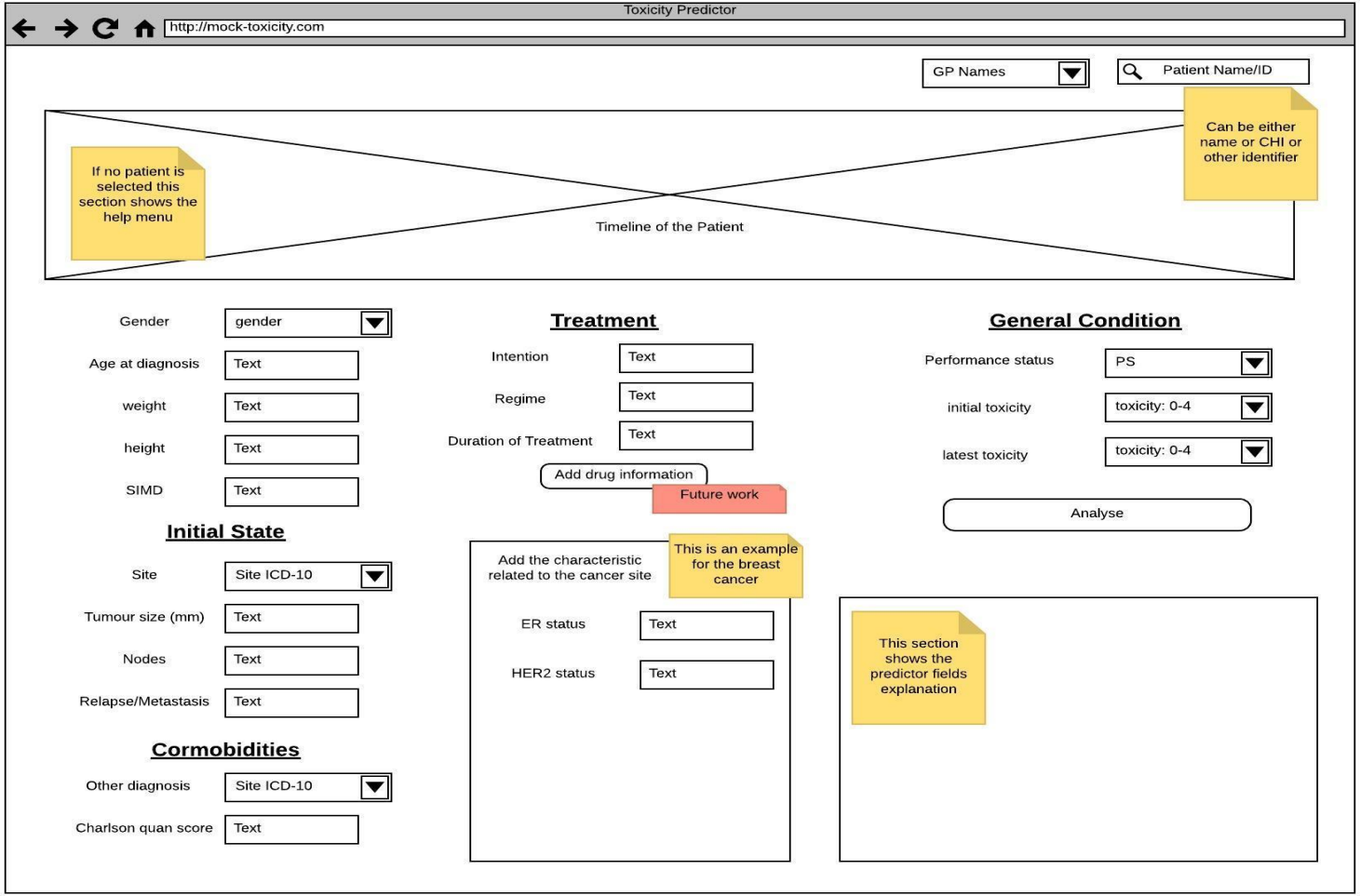

Continuos-Integration
GraphQL-Proxy Server
Graduated from ITB
Electrical Eng.
Graduate from St. Andrews
MSc. of Computer Science
IBM
2011
2012
2015
2015
2016...
EngD
Telekom Malaysia (TM)
Billing-payment Migration
accenture

EU H2020 SERUMS Contributions



Toxicity Predictor

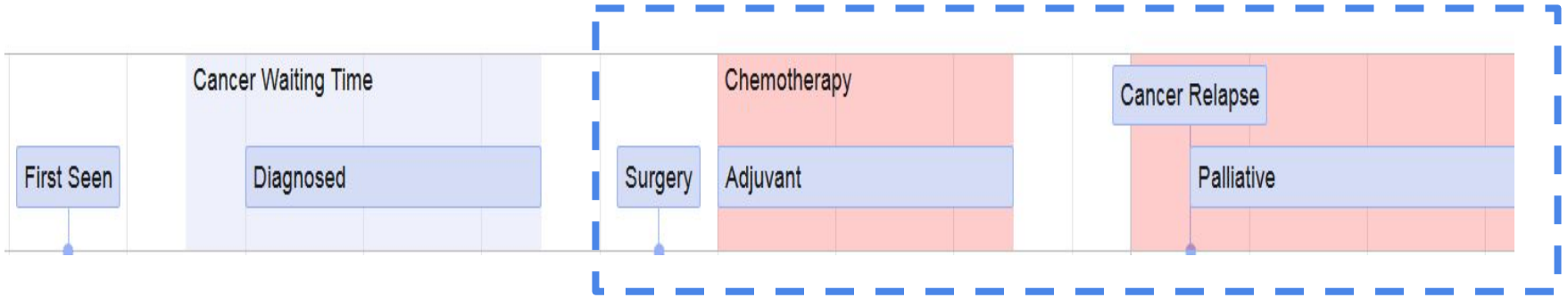



Toxicity Predictor
http://mock-toxicity.com
GP Names
Patient Name/ID
Can be either name or CHI or other identifier
If no patient is selected this section shows the help menu
Timeline of the Patient
Gender
gender
Age at diagnosis
Text
weight
Text
height
Text
SIMD
Text
Initial State
Site
Site ICD-10
Tumour size (mm)
Text
Nodes
Text
Relapse/Metastasis
Text
Cormobidities
Other diagnosis
Site ICD-10
Charlson quan score
Text
Treatment
Intention
Text
Regime
Text
Duration of Treatment
Text
Add drug information
Future work
Add the characteristic related to the cancer site
This is an example for the breast cancer
ER status
Text
HER2 status
Text
General Condition
Performance status
PS
initial toxicity
toxicity: 0-4
latest toxicity
toxicity: 0-4
Analyse
This section shows the predictor fields explanation

Synthetic Data

- We use IBM Constraint Solver to generate the synthetic data. For this, we determine the set of rules for each column (i.e., data characteristic and data relation).

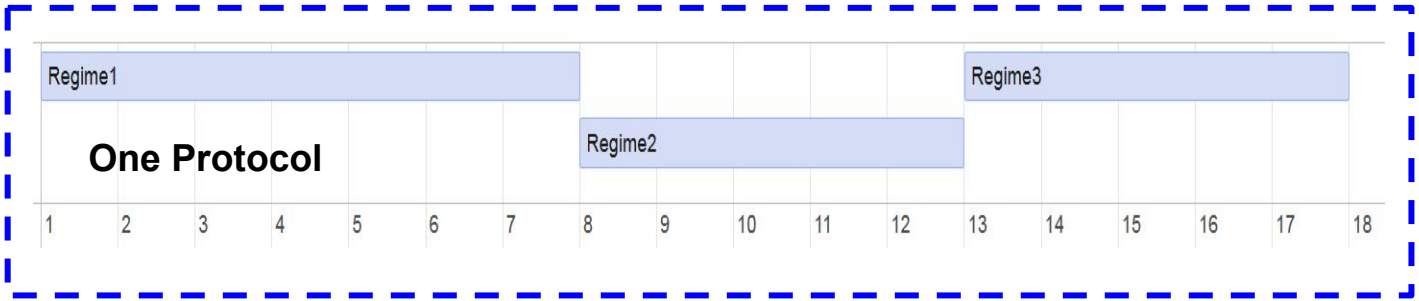
Patients' Treatment Pathway



- A patient can only be treated with one intention at a time (e.g., adjuvant)
- However, 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 regimes?

CHI	Intention	Regime	drug names	cycle
patient 1	Adjuvant	Regime1	drug1	1
patient1	Adjuvant	Regime1	drug2	1
patient 1	Adjuvant	Regime1	drug1	2
patient 1	Adjuvant	Regime1	drug2	2
patient 1	Adjuvant	Regime1	drug1	3
patient 1	Adjuvant	Regime1	drug2	3

- Incidence date happened from 250 days to 5 years before the current date

```
currentDate - (5*365) < incidence_date <= currentDate - 250
```

- Unique CHI

```
allDiff(from(general), chi)
```

- Each patient may have more than one admissions

```
numOf(from(smr01),smr01s.chi = general.chi) =
randomWeightedNumber(
  75 ? 1,
  25 ? randomNumber(2,8)
)
```

Synthetic data validation

At the moment, we can easily differentiate between the real and fabricated data.

However, after several iterations, we should NOT be able to distinguish real from fabricated data. Hence, 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