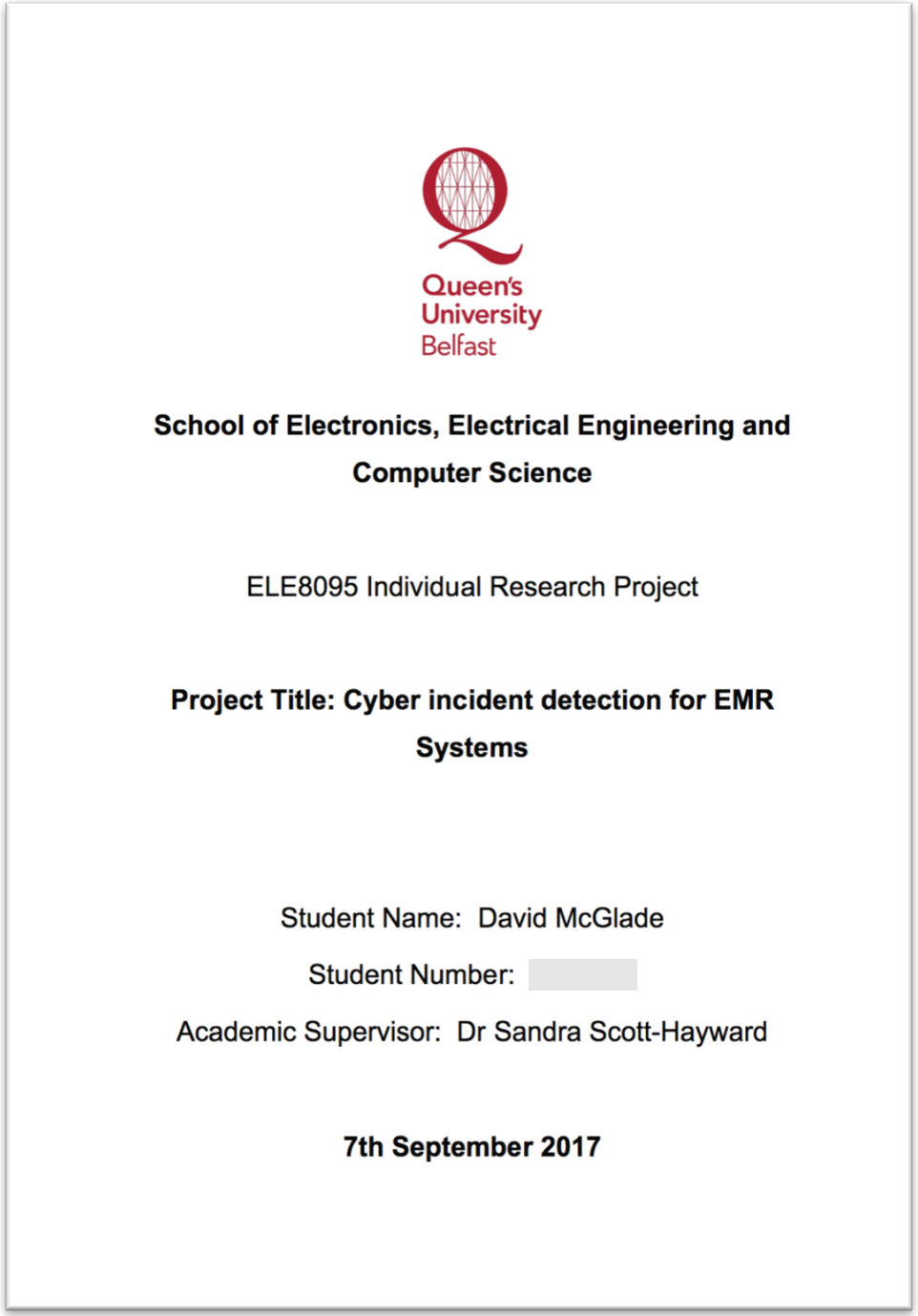
Cyber Incident Detection for EMR Systems

Davey McGlade – November 2017

kainos®

About me

- In Kainos ~9 years
- Developer -> **Technical | Solution | Principal Architect**
- Now **Cyber Lead**, **Digital Services**
- Healthcare, Financial Services & Digital Transformation of Government Services:
  - Replacement UK Wide MOT System
  - Home Office Student Visa Platform
  - UK Driver Theory Test Replacement
  - Tax Devolution for Wales
- Just finished MSc Applied Cyber Security, Queen's University Belfast / CSIT (2016-17)



Queen's University Belfast

School of Electronics, Electrical Engineering and Computer Science

ELE8095 Individual Research Project

Project Title: Cyber incident detection for EMR Systems

Student Name: David McGlade

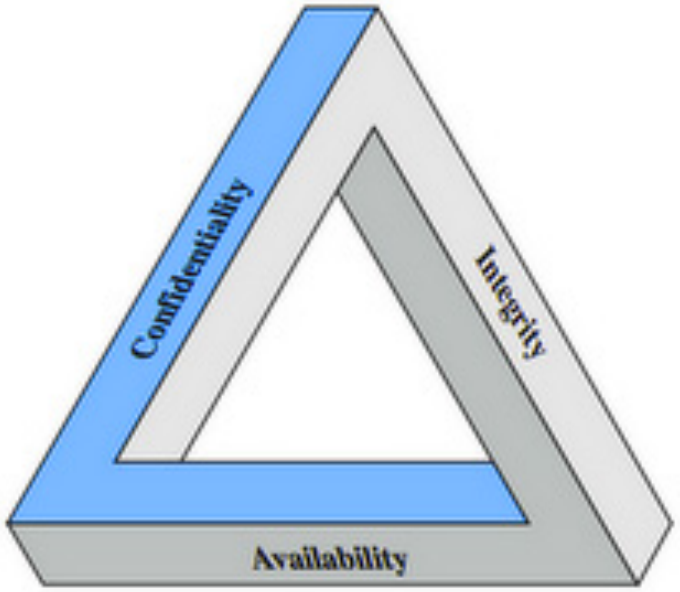
Student Number:

Academic Supervisor: Dr Sandra Scott-Hayward

7th September 2017

- This talk is based on my MSc dissertation
- Currently submitted to Elsevier for publication in Smart Health Special Issue on 'Security in Medical Cyber-Physical Systems'
- Special thanks to Dr Sandra Scott-Hayward for getting the content into the format, structure and flow needed for journal submission

What is a Cyber Incident?

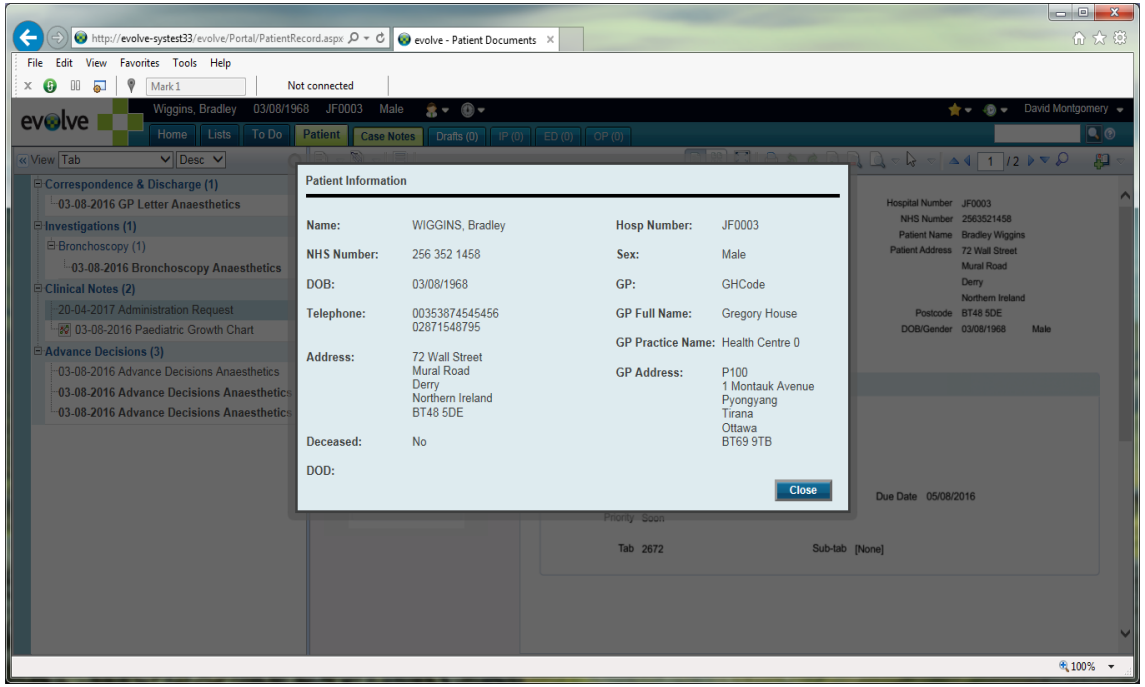


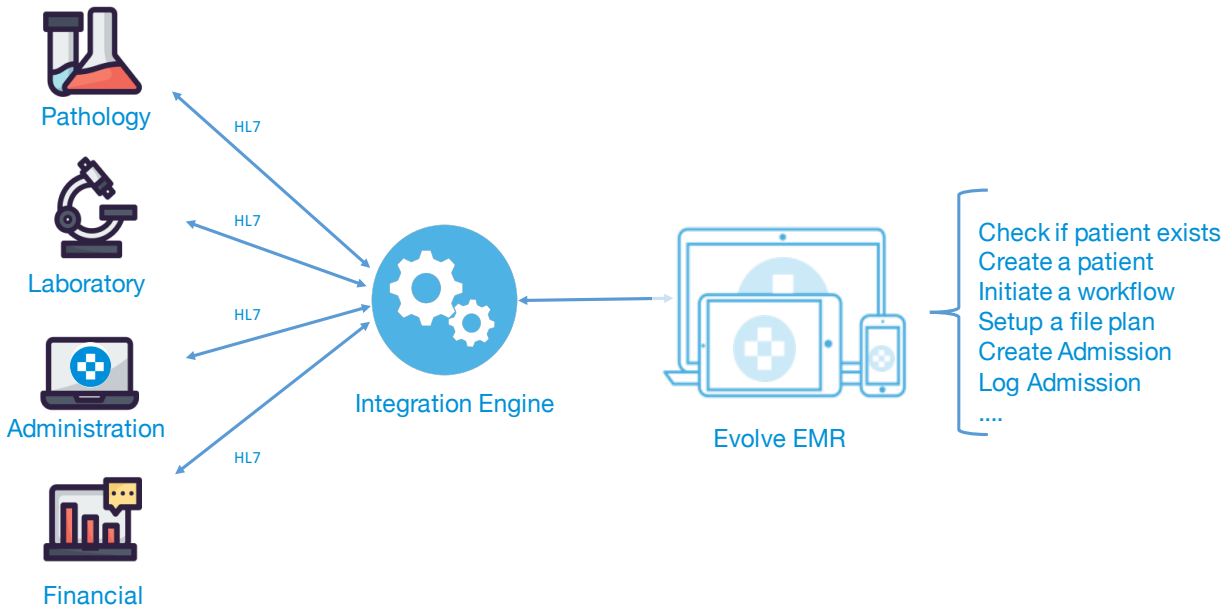
CIA Triad

- **Confidentiality** – The unauthorised access or disclosure of information.
- **Integrity** – The unauthorised modification or destruction of data.
- **Availability** – The disruption of access to, or use of, information or an information system.

Evolve EMR (Electronic Medical Record)

- Digital version of a patient's paper-based medical record
- Typically deployed within hospitals
- Deployed in 110 UK Hospitals and storing 33 million patient records
- HL7 messages used as the means to pass information between hospital systems





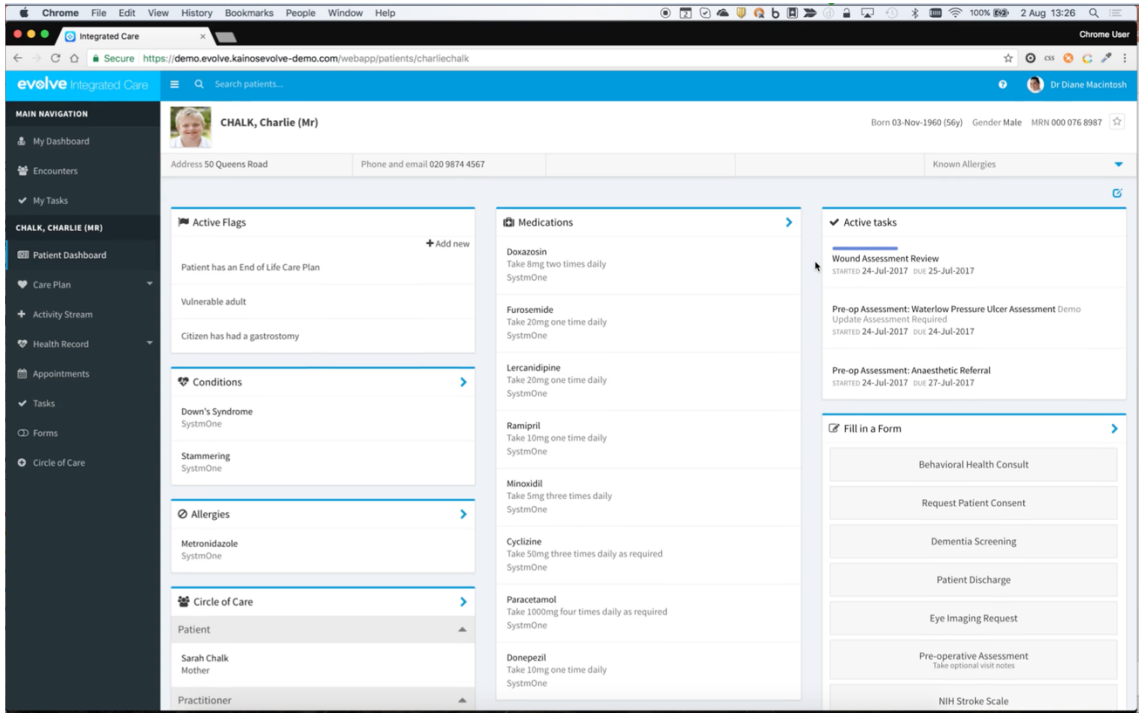
HL7

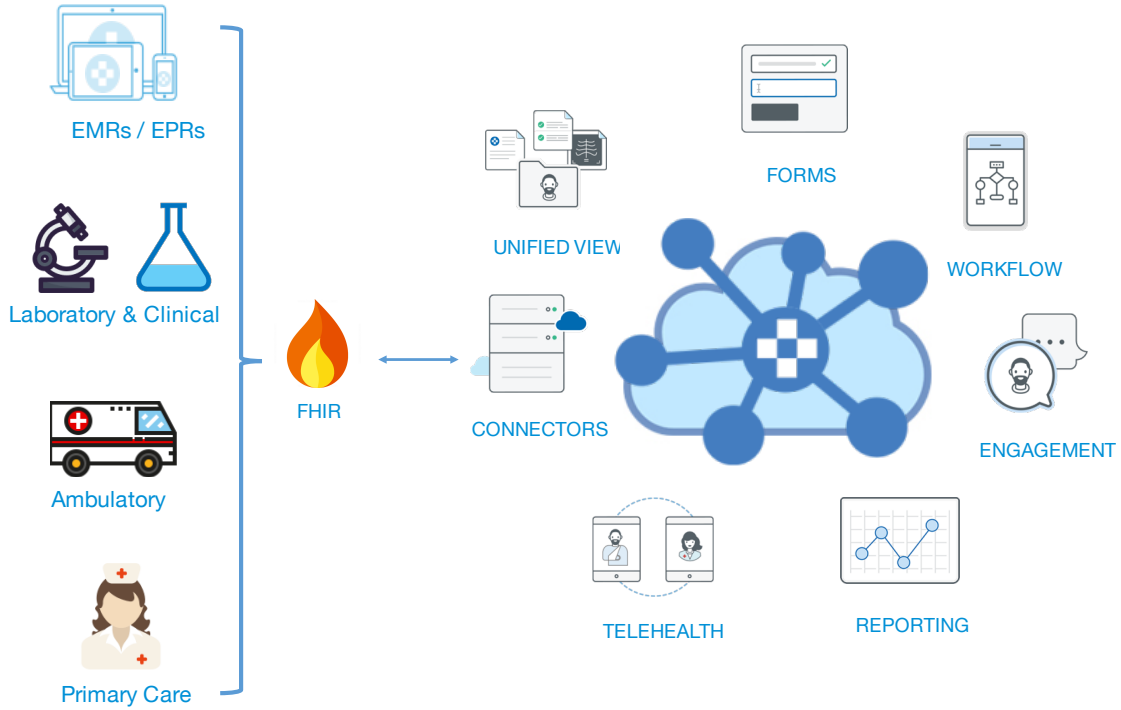
```
MSH | ^~\& | ADT1 | MCM | LABADT | MCM | 198808181126 | SECURITY | ADT^A01 | MSG00001- | P | 2.4
EVN | A01 | 198808181123
PID | | | PATID1234^5^M11 | | JONES^WILLIAM^A^III | | 19610615 | M- | | C
PV1 | 1 | I | 2000^2012^01 | | | | 004777^LEBAUER^SIDNEY^J. | | | SUR | | - | | ADM | A0
AL1 | 1 | | ^PENICILLIN | | PRODUCES HIVES~RASH~LOSS OF APPETITE
DG1 | 001 | I9 | 1550 | MAL NEO LIVER, PRIMARY | 19880501103005 | F
PR1 | 2234 | M11 | 111^CODE151 | COMMON PROCEDURES | 198809081123
```

HL7 A01 Admission Message

Evolve Integrated Care

- Automate patient care pathways across many teams and organisations
- Cloud based multi-tenant platform
- Deployed in 38 Hospitals in the US.
- FHIR used as a data format for messages and storage. JSON/XML based.





FHIR

General Person Example


```
{
  "resourceType": "Patient",
  "id": "example",
  "text": {
    "status": "generated",
    "div": "<div xmlns=\"http://www.w3.org/1999/xhtml\">\n\t\t\t<table>\n\t\t\t\t<tbody>\n\t\t\t
\t\t<tr>\n\t\t\t\t\t\t<td>Name</td>\n\t\t\t\t\t\t<td>Peter James \n              <b>Chalmers</b>
(&quot;Jim&quot;)\n            </td>\n\t\t\t\t</tr>\n\t\t\t\t<tr>\n\t\t\t\t\t<td>Address</
td>\n\t\t\t\t\t<td>534 Erewhon, Pleasantville, Vic, 3999</td>\n\t\t\t\t</tr>\n\t\t\t<tr>
\n\t\t\t\t\t<td>Contacts</td>\n\t\t\t\t\t<td>Home: unknown. Work: (03) 5555 6473</td>\n\t\t
\t\t</tr>\n\t\t\t<tr>\n\t\t\t\t\t<td>Id</td>\n\t\t\t\t\t<td>MRN: 12345 (Acme Healthcar
e)</td>\n\t\t\t</tr>\n\t\t</tbody>\n\t\t</table>\n\t</div>"
  },
  "identifier": [
    {
      "use": "usual",
      "type": {
        "coding": [
          {
            "system": "http://hl7.org/fhir/v2/0203",
            "code": "MR"
          }
        ]
      },
      "system": "urn:oid:1.2.36.146.595.217.0.1",
      "value": "12345",
      "period": {
        "start": "2001-05-06"
      },
      "assigner": {
        "display": "Acme Healthcare"
      }
    }
  ],
  "active": true,
```

Motivation

Incident Prevalence (2015-16)

- 63% growth in cyber incidents against US Hospitals
- 243% growth in cyber incidents against UK hospitals

Incident Types:

- Identity Theft – Medical data is 10 to 20 times the value of credit card information.
- Insurance Fraud - Medical information is used to generate false identities for fraud.
- Malicious attack – A victim could receive an incorrect dosage or medication.
- Extortion / Blackmail – extort money from individuals or healthcare organisations.

Trusted Insiders / Systems

Nassau County, NY – Nassau County District Attorney Kathleen Rice and the Nassau County Police Department announced the arrest of a Great Neck doctor responsible for stealing the protected personal and health information of tens of thousands of patients.

Background Research

Existing EMR security measures:

- Network Security
- Role Based Access Control
- Logging and Monitoring
- Encryption

Existing EMR incident detection

- Monitor the patient's journey
- Detect across multiple streams
- Community Anomaly Detection
- SIEMs

Machine Learning (ML)

- Detect unseen patterns
- Can surpass rule based SIEMs
- Classification, Clustering and appear best suited for incident detection with lack of data.

Time Series (TS) Anomaly Detection

- ML better suited to static / persisted data.
- HL7 messages flow into Evolve
- Real-time/near real-time detection needed.

Goal

To understand incident detection against **Confidentiality, Integrity and Availability**, by using **Evolve EMR** and **Evolve IC** as reference platforms for prototype work.

'Incident Detection', however is **simply one aspect of the problem** – the solution(s) may be able to successfully detect an incident, but be infeasible, impractical or simply not perform well enough to be used in a clinical production setting.

Refine the problem to be – "How feasible, practical and performant is cyber incident detection against an EMR".

Goal – make it real.

Feasibility

- **Data**: Does HL7, FHIR or audit data support the use case scenario?
- **Tooling**: Does suitable tooling and algorithms exist to support the detection approach? How quickly can anomalies be detected in real time?

Practicality

- **Difficulty**: How difficult is it to understand and apply the machine learning algorithms? Are specialist skills needed?
- **Repeatability**: How easy is it to repeat the training/testing process against a real EMR product? Can new anomaly types be added to the detection mechanism?

Performance

- **Accuracy** = (TP+TN)/(TP+TN+FP+FN)
- **Precision** = TP / (TP + FP)
- **Recall** = TP/(TP+FN)

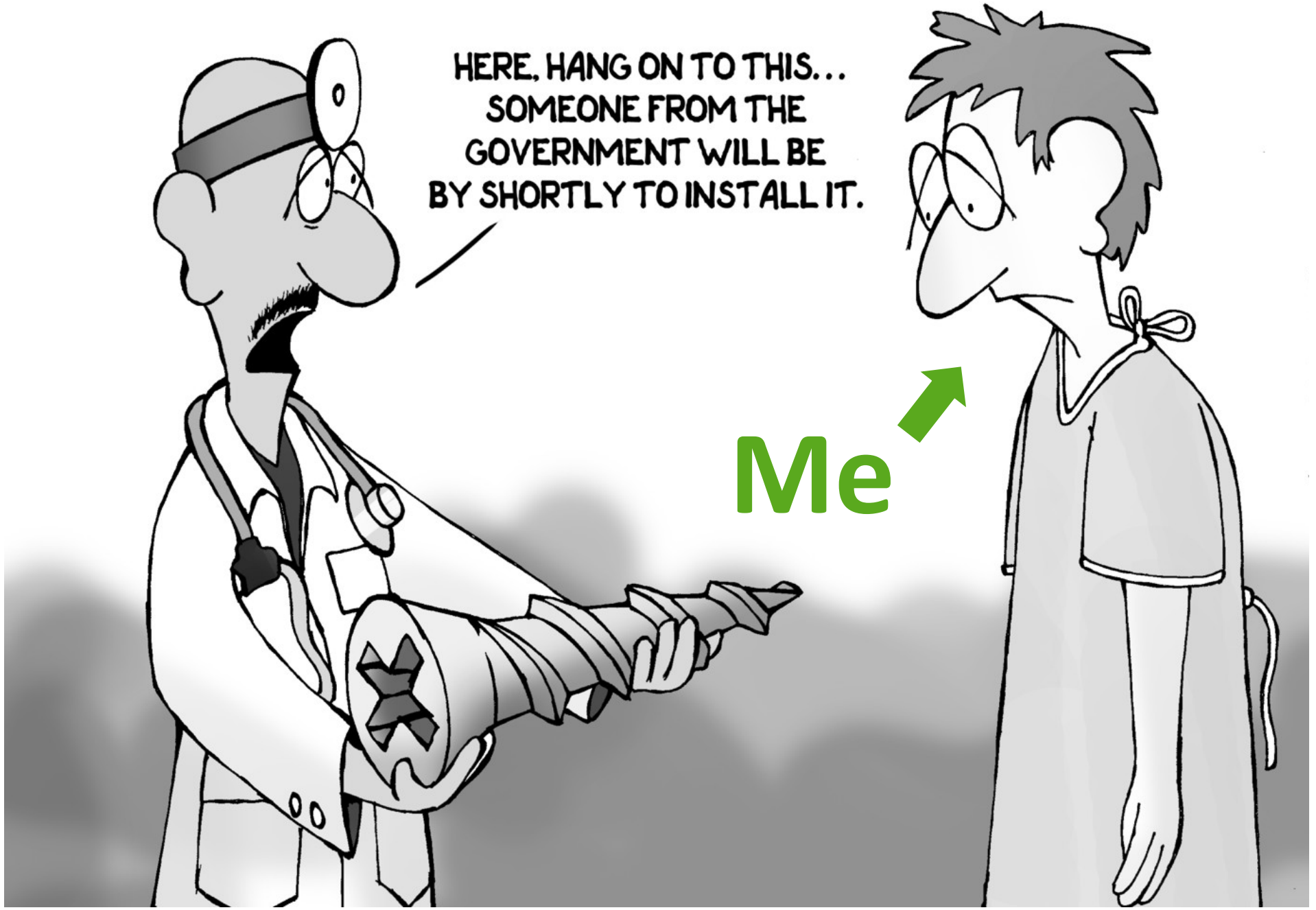
TP: True Positive,
TN: True Negative
FP: False Positive,
FN: False Negative

Lets talk about data.

Data from 1.6 million patients released to Google DeepMind as part of a clinical safety initiative.

"Royal Free London NHS Foundation Trust did not comply with the Data Protection Act."

- **Elizabeth Denham**
  Information Commissioner, July 17


HERE. HANG ON TO THIS... SOMEONE FROM THE GOVERNMENT WILL BE BY SHORTLY TO INSTALL IT.
Me

SyntheticMass: Day 405

Home | SyntheticMass Dashboard | About This Project | Feedback

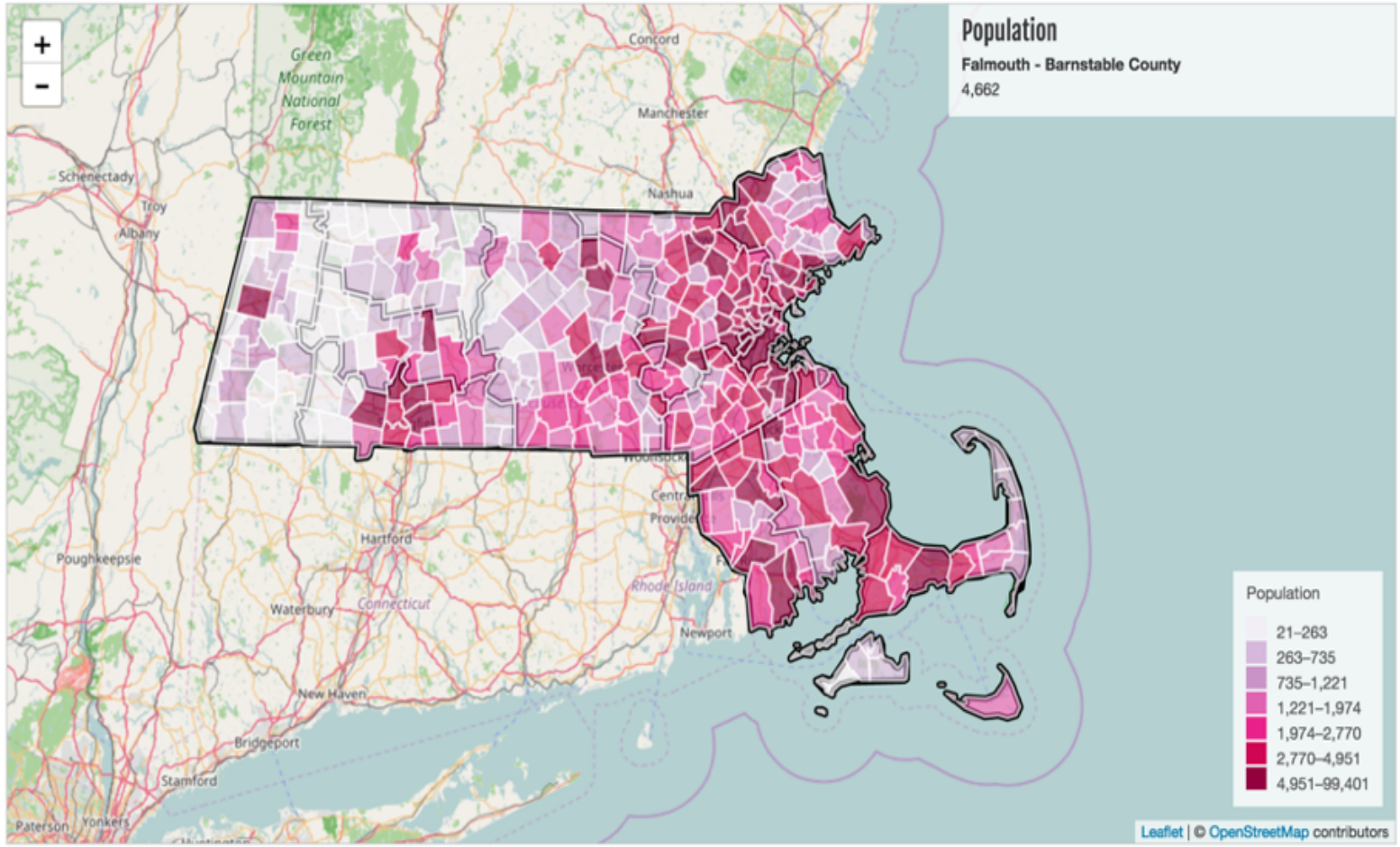
Source of Data: Synthea | Geographic region: Cities and Towns | Data Value: Population

Zoom map to all

Population

Number of Residents

Region Type	Cities and Towns
Data Set	Synthetic data generated from Synthea
Total Population	1,009,150
Mean	2,875
Max	Boston Cities and Towns: 99,401
Min	Gosnold Cities and Towns: 21



SyntheticMass: Day 499 | Home | SyntheticMass Dashboard | Download | FHIR API | About This Project | Feedback

Source of Data: Synthea | Geographic region: Cities and Towns | Data Value: Population | Zoom map to all

Plymouth

× Close

County	Plymouth	
Population	8,730	
Population Density	90.5 (per mi²)	
Area	96 sq. mi.	Zoom
Demographics		
Female Population	52.1%	(276 of 351)
Male Population	47.9%	(76 of 351)
Diabetes Prevalence	5.2%	(147 of 351)
Opioid Addiction Prevalence	1.0%	(227 of 351)
Heart Disease Prevalence	6.3%	(200 of 351)

Name	Gender	DOB
Goldner839, Mertie87	male	16.Sep.1906
Wolf829, Harmon731	female	26.Jan.1908
Ritchie488, Bell585	male	02.Jan.1927
Torphy941, Talon803	female	10.Aug.1928
Nicolas307, Vickie253	female	24.Apr.1930
Koepp295, Ibrahim34	female	03.Aug.1948
D'Amore780, Delilah467	male	08.Sep.1954
Bergstrom319, Pascale887	male	15.Jun.1955
Walker986, Kay881	female	13.Jun.1964
Wilkinson941, Willa252	male	12.Dec.1978
Monahan18, Emily195	female	25.Nov.1982
Mosciski198, Kiana38	male	30.Aug.1987

Patient Record

× Close

Family name	Goldner839
Given name	Mertie87
Address	9120 Mae Shoal Apt. 762
City, State	Plymouth, MA
Postal Code	02361

Download Patient Data (FHIR JSON) | Download Patient Data (CCDA XML)
Send Data via Direct Messsage

Height	n/a
Weight	n/a
Blood Type	n/a
Vision	n/a

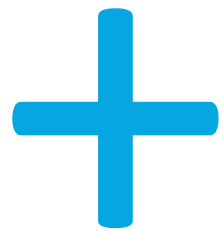
DOB	16.Sep.1906
Age	56 (Died: 09.Dec.1962)
Cause of Death	Stroke
Gender	male
Race	White
Ethnicity	Nonhispanic
Spoken language	n/a

Observations | Conditions | Medications | Allergies | Care Plans | More

Conditions	Date of Onset	Date Resolved
Hypertension	13.Aug.1925	n/a
Chronic sinusitis (disorder)	05.Oct.1948	n/a
Stroke	09.Dec.1962	n/a

Lets talk about technology.


python™




scikits
learn


LinkedIn

Scenario 1- Confidentiality.

Confidentiality Scenario

Confidentiality Scenario - Identify anomalous ‘read’ events by clinician users against patient data. Anomalous read events occur when no **relationship** to the patient and when no **break-glass** event exists.

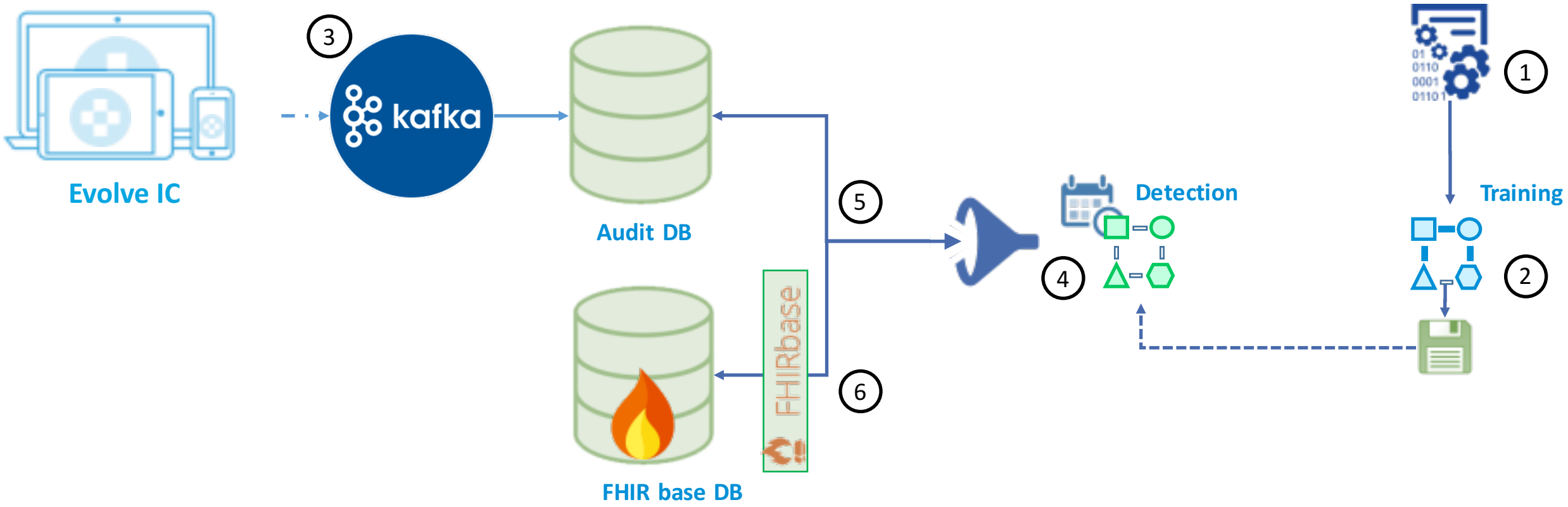
Relationship: A relationship is defined by the presence of **appointments**, **observations** or **encounter** records linking the patient and the clinician user.

Break-glass Event - A break-glass event is when a clinician legitimately accesses a patient record in an **emergency situation**.

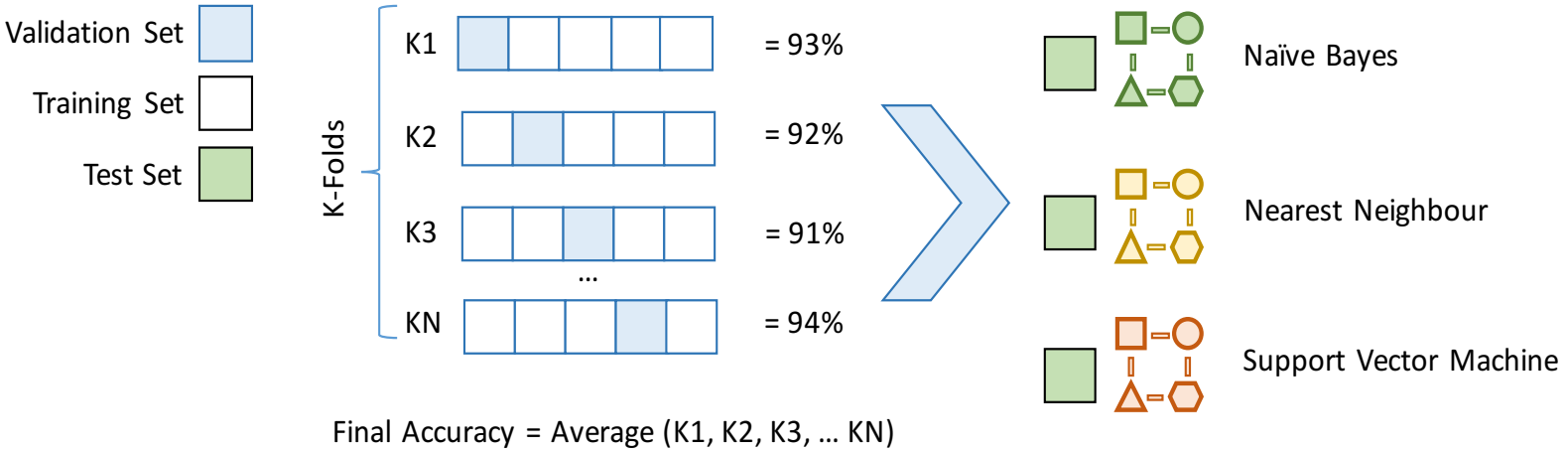
Normal & Anomaly States

ID	Appointments	Observations	Encounters	Break-glass	Class	Description
1	N	N	N	N	Anomaly	With no appointments, observations, encounters or break-glass events, there is no relationship between the clinician and the patient, and no emergency has occurred.
2	N	N	N	Y	Normal	A break-glass event has occurred.
3	N	N	Y	N	Anomaly	The clinician has met the patient but no appointments exist for a future consultation.
4	N	N	Y	Y	Normal	The clinician accessed the patient records in an emergengy situation
5	N	Y	N	N	Anomaly	The clinician has observed the patient but no appointment exists for a future consultation.
6	N	Y	N	Y	Normal	break-glass event
7	N	Y	Y	N	Anomaly	The clinician has met and observed the patient but no appointment exists for a future consultation.
8	N	Y	Y	Y	Normal	Break the Glass event
9	Y	N	N	N	Normal	Future appointment
10	Y	N	N	Y	Anomaly	The clinician has an appointment with the patient, but has accessed their record via a Break The Glass event.
11	Y	N	Y	N	Normal	The clinician has met the patient and appointments exist
12	Y	N	Y	Y	Normal	The clinician has met the patient and has had to access their file in an emergency situation
13	Y	Y	N	N	Normal	The clinician has observed the patient and has future appointments to do so.
14	Y	Y	N	Y	Normal	Emergency access
15	Y	Y	Y	N	Normal	Normal hospital visit
16	Y	Y	Y	Y	Normal	The clinician has met the patient and has had to access their file in an emergency situation

Confidentiality Prototype



Model Training



Final Accuracy = Average (K1, K2, K3, … KN)

	Accuracy			
K-Folds	KNN	SVM	MNB	Average
3	0.9729	0.9864	0.7044	0.8879
5	0.9713	0.9925	0.7059	0.8899
7	0.9864	0.9880	0.7466	0.9070
10	0.9834	0.9910	0.7105	0.8949

Train the models

3 Models Selected
SVM, KNN, MNB

Split into Training, Testing and Validation.

Lets get us some ML models

Lets see how they do

```
1  import pandas
2  import pickle
3  import numpy as np
4  from sklearn.model_selection import cross_val_score
5  from sklearn.model_selection import train_test_split
6  from sklearn.model_selection import KFold
7  from sklearn.metrics import classification_report
8  from sklearn.neighbors import KNeighborsClassifier
9  from sklearn.naive_bayes import MultinomialNB
10 from sklearn import tree
11 from sklearn.svm import SVC
12
13 K_SPLITS = 7;VALIDATION_SIZE = 0.34
14
15 # READ IN TRAINING DATA
16 Training_CSV = 'conf-training-dataset.csv'
17 columns = ['User_name', 'PatientID', 'Appointments', 'Observations', 'Encounters', 'SpecialAction', 'Class']
18 df = pandas.read_csv(Training_CSV, names=columns)
19
20 # SPLIT TRAINING/VALIDATION
21 array = df.values
22 X = array[:,2:6]
23 Y = array[:,6]
24
25 X_train, X_test, Y_train, Y_test = train_test_split(X, Y, test_size=VALIDATION_SIZE)
26
27 models = []
28 models.append(('KNN', KNeighborsClassifier()))
29 models.append(('SVM', SVC()))
30 models.append(('MNB', MultinomialNB()))
31
32 # EVALUATE MODELS
33 for name, model in models:
34     kfold = KFold(n_splits=K_SPLITS)
35     cv_results = cross_val_score(model, X_train, Y_train, cv=kfold, scoring='accuracy')
36     print(name, cv_results.mean(), cv_results.std())
```

Persist the models

Predict against the Validation Set for KNN, SVM and MNB.

Persist each model to disk

```
38  # FIT AND PERSIST MODELS FOR DETECTION
39  knn = KNeighborsClassifier()
40  knn.fit(X_train, Y_train)
41  knn_predict = knn.predict(X_test)
42  report = classification_report(Y_test, knn_predict)
43  print "\nKNN"; print report
44
45  output = open('knn.pkl', 'wb')
46  pickle.dump(knn, output)
47  output.close()
48
49  svectm = SVC()
50  svectm.fit(X_train, Y_train)
51  svm_predict = svectm.predict(X_test)
52  report = classification_report(Y_test, svm_predict)
53  print "SVM"; print report
54
55  output = open('svm.pkl', 'wb')
56  pickle.dump(svectm, output)
57  output.close()
58
59  mnb = MultinomialNB()
60  mnb.fit(X_train, Y_train)
61  mnb_predict = mnb.predict(X_test)
62  report = classification_report(Y_test, mnb_predict)
63  print "MNB"; print report
64
65  output = open('mnb.pkl', 'wb')
66  pickle.dump(mnb, output)
67  output.close()
```

Lets detect against Evolve IC

Let's reload our trained models

```
11  #-- RELOAD ML MODELS
12  mnb = MultinomialNB()
13  mnb = pickle.load( open( "mnb.pkl", "rb" ) )
```

Get all the audit records

```
25  #-- RETRIEVE ALL PATIENTS WHOSE FILE HAS BEEN READ IN PREVIOUS 24 HRS'
26  patient_query = "select occurrence_date, user_name, patient_id, special_action from audit_event "
27  patient_query = patient_query + "where audit_event.action_code = 'R' "
28  patient_query = patient_query + "and occurrence_date > NOW() - INTERVAL '1 day' "
29  patient_query = patient_query + "and (audit_event.patient_id <> '')  "
30  patient_query = patient_query + "order by audit_event.user_name asc"
```

Query FHIR to get Appointments Observations and Encounters

```
59      #-- OBTAIN APPOINTMENT COUNT'
60      fhir_query = "SELECT fhir_search('{\"resourceType\": \"Appointment\", \"queryString\": \"patient="
61      + patientID + "&practitioner=" + clinicianID + "\"}');"
62      # print fhir_query
63      fhir_cur.execute(fhir_query)
64      fhir_rows = fhir_cur.fetchall()
65
66      ' Needed as Postgres returns result as unicode'
67      fhir_resource = json.dumps(fhir_rows[0], ensure_ascii=False)
68
69      js = json.loads(fhir_resource)
70      appCount = js[0]["total"]
```

Make some predictions based on aggregated data

```
96      X = np.array([int(appCount),int(obsCount),int(encCount), int(specialAction)])
97      X = X.reshape(1,4)
98
99      #-- PREDICT CLASS, 'N = Normal, A = Anomalous'
100     svm_prediction = svectm.predict(X)
101     knn_prediction = knn.predict(X)
102     mnb_prediction = mnb.predict(X)
103
104     if knn_prediction[0] == "A":
105         knn_Acount = knn_Acount + 1
106     else:
107         knn_Ncount = knn_Ncount + 1
```

Demo

Results - Model Training

Step 1 - Model Fitting Results

KNN	Precision	Recall
Anomaly	0.87	0.99
Normal	1.00	0.93
Avg./ total	**0.95**	**0.95**

SVM	Precision	Recall
Anomaly	1.00	0.94
Normal	0.97	1.00
Avg./ total	**0.98**	**0.98**

MNB	Precision	Recall
Anomaly	0.60	0.42
Normal	0.75	0.86
Avg./ total	**0.70**	**0.72**

- **Precision (Get it right)**: The ability of the classifier not to label as positive, a sample that is negative. Represented by TP / (TP+FP), where TP is True Positive and FP is False Positive.

- **Recall (Catch them all)**: The ability of the classifier to find all the positive samples. Represented by TP / (TP+FN), where TP is True Positive and FN is False Negative.

Results - Evolve IC

Step 2 – Evolve IC Testing (2010 records, with 0, 1 or 3 anomalies added)

Run 1 – No anomalies

- SVM – 2010 Normal, 0 Anomalies
- KNN – 2006 Normal, 4 Anomalies
- MNB – 1988 Normal, 22 Anomalies

Run 2 – Single anomaly

- SVM – 2010 Normal, 1 Anomalies
- KNN – 2006 Normal, 5 Anomalies
- MNB – 1988 Normal, 22 Anomalies

Run 3 – Three anomalies

- SVM – 2010 Normal, 3 Anomalies
- KNN – 2006 Normal, 7 Anomalies
- MNB – 1988 Normal, 22 Anomalies

Confidentiality Scenario – Key Findings

Model Performance

- Of the 3 models SVM performed best 98.94% (Training) and 100% (Evolve IC, 3 runs).
- KNN and MNB both had issues with inadequate training set exposure and unseen data. SVM was more robust.

The need for ML?

- Rules (simple) vs ML (complex variable interplay, data volumes and unknown patterns)
- Specialist skills, data preparation and pipeline / alerting is needed.
- Implications exist for it's use in automated decisions – Data Protection Act

Data is not straightforward

- FHIR is rich, but hard to create & manipulate.
- Data for model training may be difficult to obtain.
- Multiple FHIR versions exist – care is needed on versions used.

Scenario 2- Integrity.

Integrity Scenario

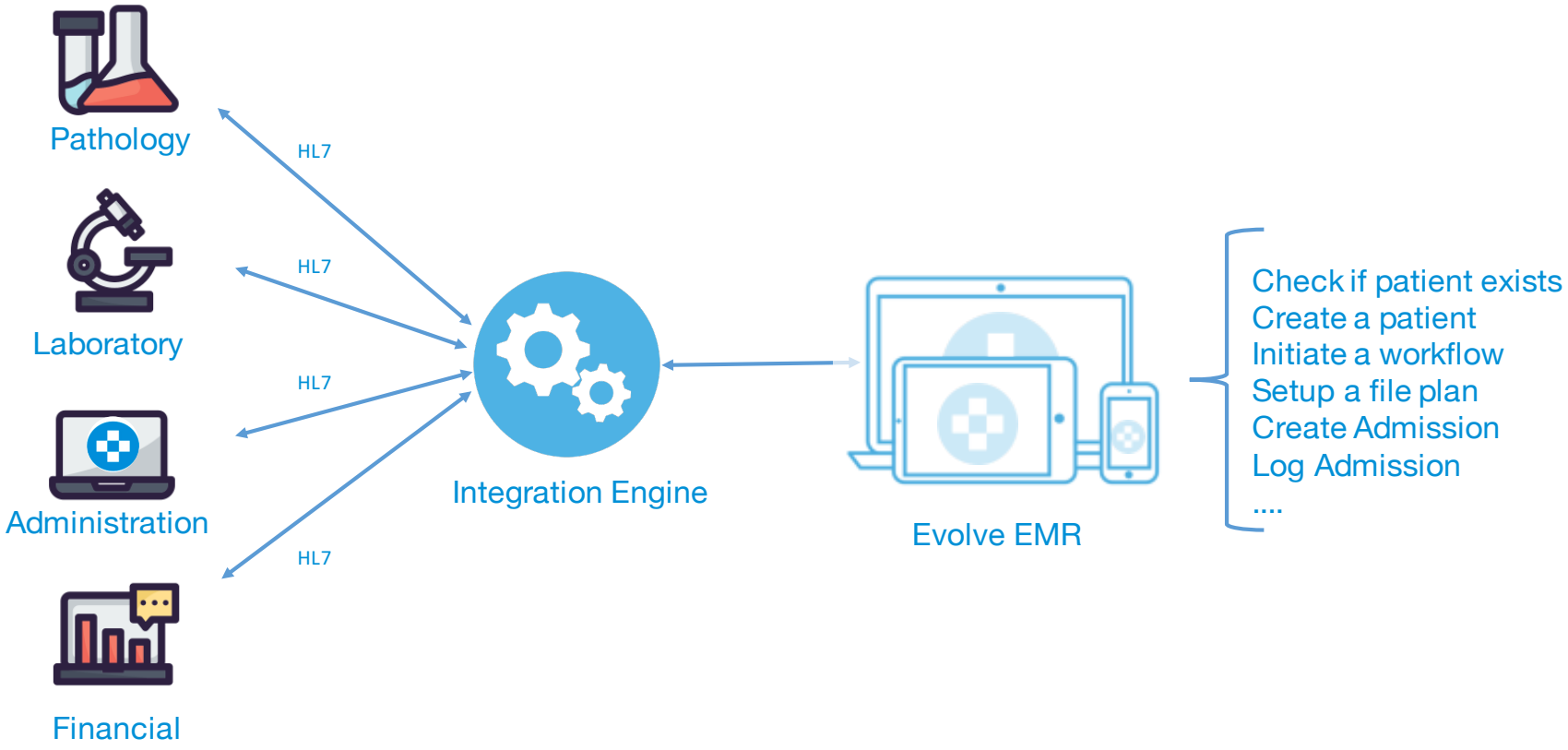
Integrity - The unauthorised modification or destruction of data.

- A **single** change to **a single field** in a patient record could impact its integrity.
- A trusted, but malicious user, may **choose to only carry out a single change** against a single patient record and **do no more**.
- Proving Integrity changes is **HARD**
- Typically look at the integrity of the database. In some cases, pre & post field states.
- Considered looking at Medication Prescription vs Medication Dispensation FHIR Records.
- Possibility of future research are on **‘hot fields’.**

INTEGRITY SCENARIO DISCONTINUED

Scenario 3 – Availability.

HL7 messages from PAS -> Evolve EMR



Availability

TEST CASES		ANOMALY CASE (SEPT 2016)			NORMAL CASE FEB 2017)			COMBINED (TEST DATASET)
		Sep 16	Sep 16	10%)	Feb 17	Feb 17	10%)	
ADT_A01	Admit/visit notification	13,114	3.6%	1,311	10,488	6.5%	1,049	3,409
ADT_A02	Transfer a patient	4,674	1.3%	467	4,785	3.0%	479	1,424
ADT_A03	Discharge/end visit	11,300	3.1%	1,130	10,777	6.7%	1,078	3,285
ADT_A05	Pre-admit a patient	71,903	19.7%	7,190	68,855	42.6%	6,886	20,961
ADT_A13	Cancel discharge/end visit	172	0.0%	17	178	0.1%	18	53
ADT_A28	Add person information	5,336	1.5%	534	5,752	3.6%	575	1,684
ADT_A31	Update person information	240,583	66.0%	24,058	45,152	27.9%	4,515	33,089
ADT_A38	Cancel pre-admit	17,240	4.7%	1,724	15,620	9.7%	1,562	4,848
ADT_A40	Merge patient - patient identifier list	144	0.0%	14	37	0.0%	4	22
		364466	100.0%	**36,447**	161644	100.0%	**16,164**	**68,775**

Availability Scenario - Identify a surge of anomalous HL7 A31 messages that come from an upstream PAS system.

Time Series Types

- **Point** – a specific event that is **distinguishable from the rest of the dataset**. For example, a single excessive payment on a credit-card statement.

- **Contextual** – the anomaly occurs at an **unexpected interval of time**. For example, a high temperature observation occurring in the middle of winter.

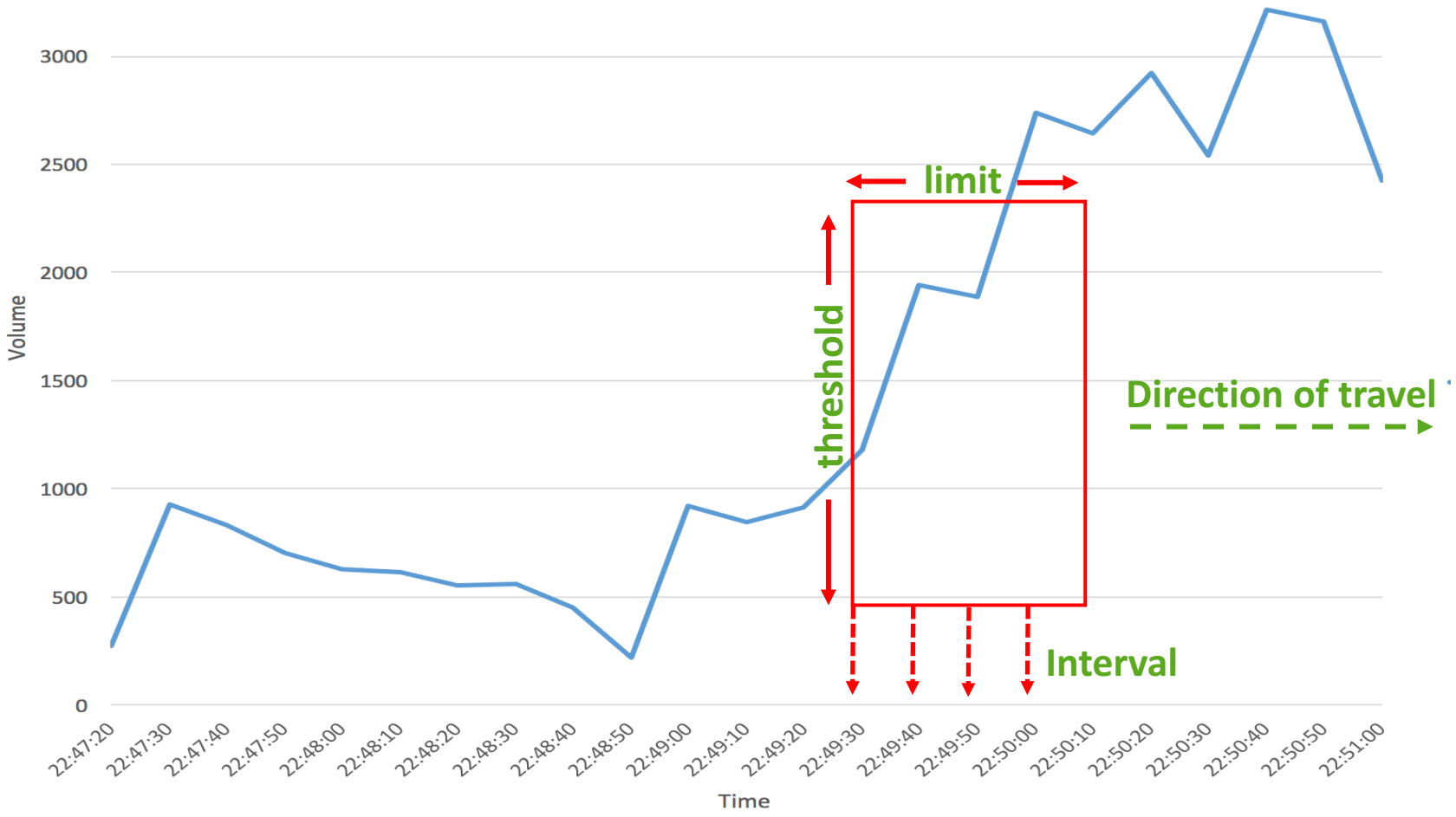
- **Collective** – the anomaly occurs in **respect to the rest of the entire dataset being measured**. For example, A reversing car beeps to warn you of a rear collision. If no movement is performed, the beep frequency remains constant, however if the vehicle moves closer to the object, the beeps increase to a new frequency. This change is also known as a **'mean shift or level shift'**.

We need to measure the flow throughput:

Options:

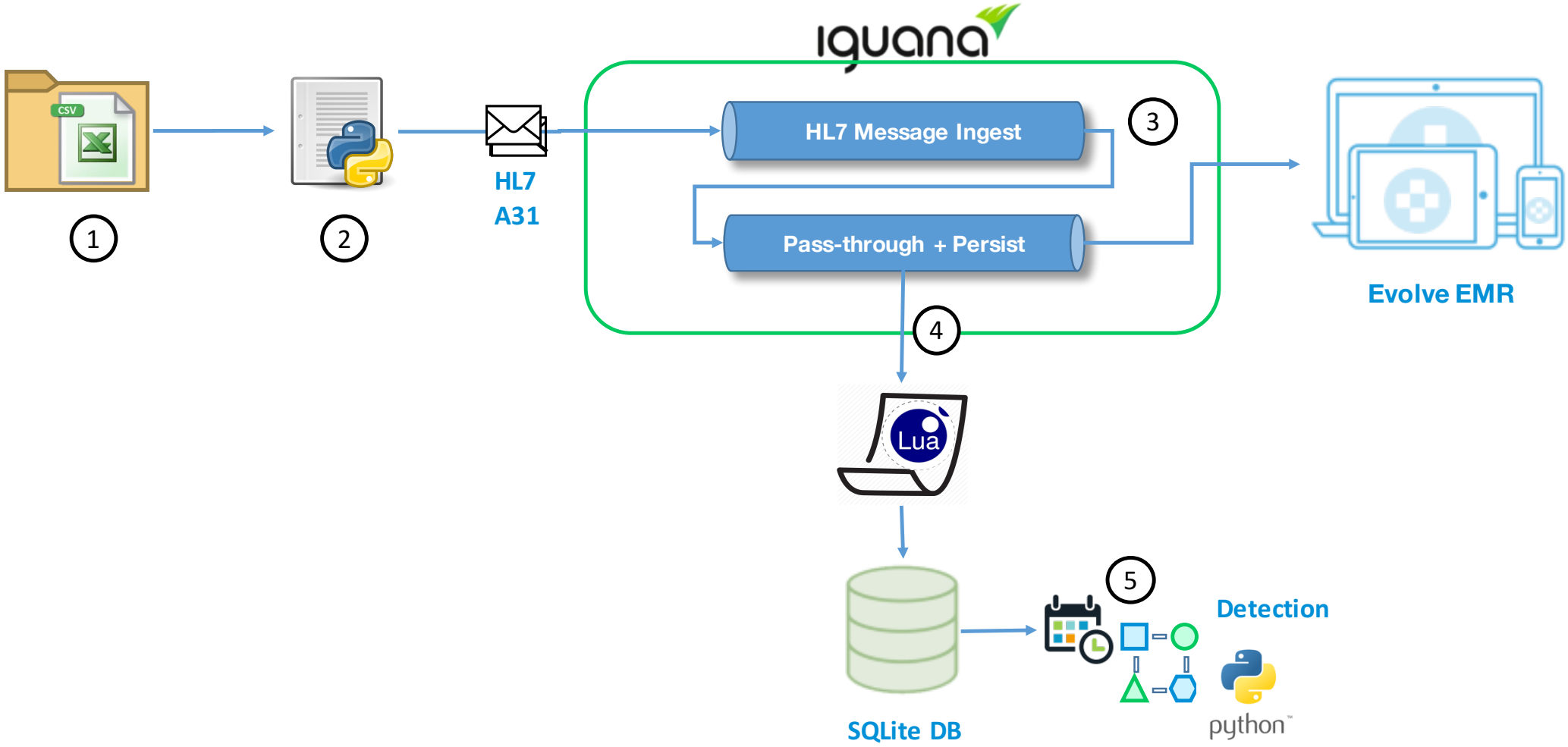
1. Seasonal Trend Decomposition (STL)
2. Classification and Regression Trees (CART)
3. Autoregressive Integrated Moving Average (ARIMA)
4. Long Short Term Memory (LTSM)
5. Exponential Moving Average (EMA)

Exponential Moving Average



- **Limit** – how many previous samples get included.
- **Threshold** – how big should the 'spike' be before it's an anomaly.
- **Interval** – how frequently the anomaly detection check should run.

Availability Prototype Design

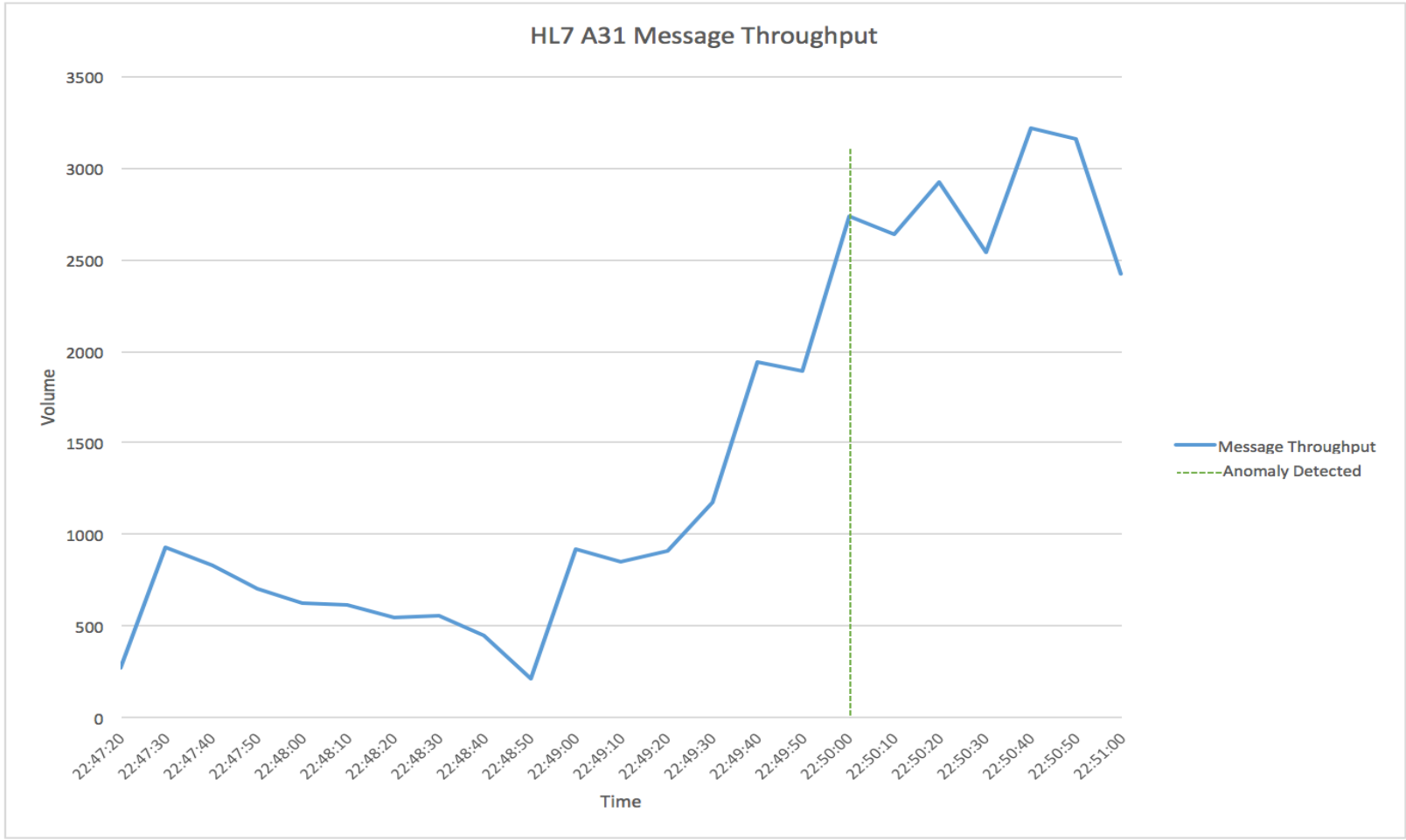


Demo

Availability Prototype Results – Run 1

RUN 1 - Static detection - Combined Dataset A31 Messages - EMA Threshold 1.0 - Interval 10

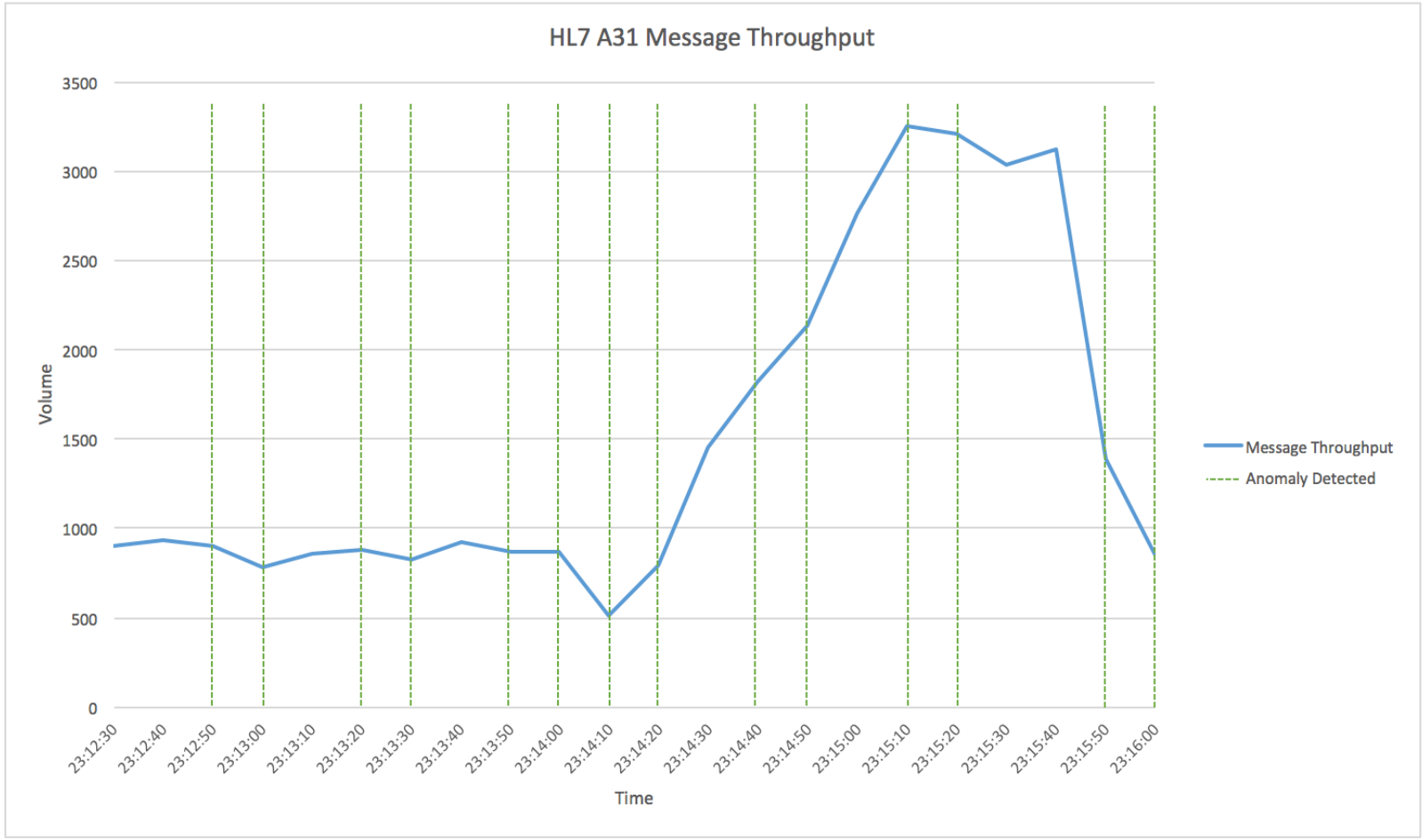
interval	count
22:47:20	277
22:47:30	926
22:47:40	835
22:47:50	706
22:48:00	627
22:48:10	616
22:48:20	552
22:48:30	560
22:48:40	449
22:48:50	218
22:49:00	918
22:49:10	847
22:49:20	912
22:49:30	1178
22:49:40	1942
22:49:50	1888
22:50:00	**2738**
22:50:10	2640
22:50:20	2919
22:50:30	2538
22:50:40	3216
22:50:50	3161
22:51:00	2424
Total	33087



Availability Prototype Results – Run 2

RUN 2 - Running detection - Combined Dataset A31 Messages - EMA Threshold 1.0 - Interval 10

interval	count
23:12:30	901
23:12:40	934
23:12:50	903
23:13:00	782
23:13:10	856
23:13:20	875
23:13:30	823
23:13:40	927
23:13:50	871
23:14:00	865
23:14:10	516
23:14:20	799
23:14:30	1453
23:14:40	1825
23:14:50	2132
23:15:00	2766
23:15:10	3255
23:15:20	3212
23:15:30	3030
23:15:40	3123
23:15:50	1393
23:16:00	846
Total	33087



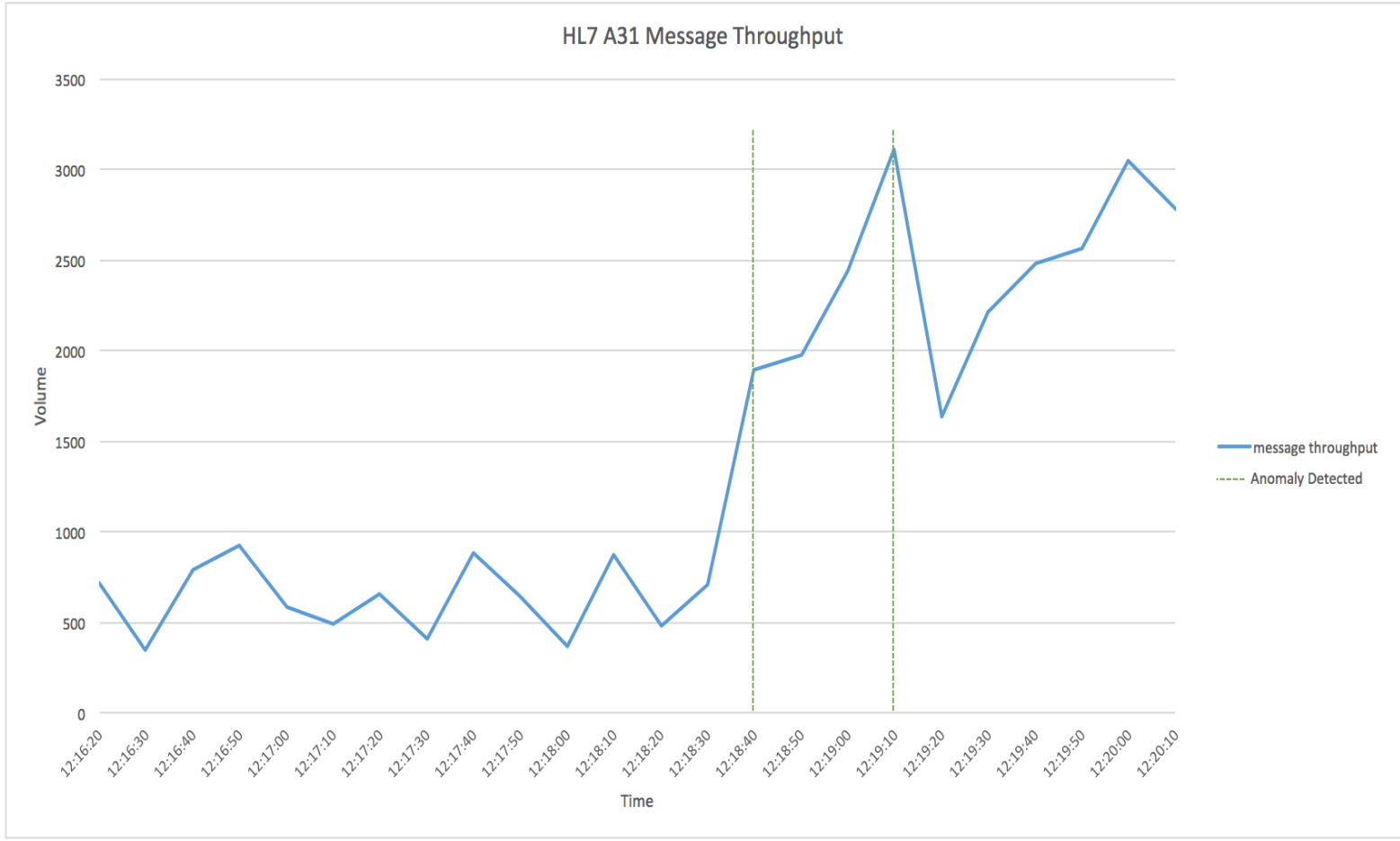
Availability Prototype Results – all runs

	Commentary	Detection Configuration Input				Run Output		
	Description - Any variables marked in RED have changed between Runs	**Detection Schedule**	**Query Interval**	**Threshold**	**Limit Clause value**	**Speed of Each Detection Run**	**Anomalites Detected**	**Max Message Per second**
RUN 1	Detection Application is run once, manually after data has been already been passed through Iguana and persisted into SQLite. This run is principally a smoke test to ensure that the application is operating correctly. Applications initial query to SQLite retrieves messages in intervals of 10 seconds. Threshold set at default of 1.0. Note that only one anomaly was reported. This is because the Threshold has been set at 1.00 - every other spike or dip to that point has been < 1.0 in its equivalent anomaly score.	**Manual**	10 Secs	1.00	N/A	~1 Sec	1	537
RUN 2	It has been observed that everytime the detection application runs, all rows in the SQLite table are returned for consideration as anomalies. This indicates a change in the SQL query is required to introduce a 'limit' clause in order to only pull back the last 'X' number of rows. The new limit variable will be added as a constant to the detection application and included in future runs. The culmulative inclusion of all rows basically applies the EMA across everything that has been stored in SQLite to date. This results in a high number of anomalies being detected as the EMA window is considering 'everything' that has been persisted to SQLite up to the point in time when it runs.	**Scheduled**	10 Secs	1.00	N/A	~1 Sec	14	554
RUN 3	Limit clause is now included to only retrieve the last 5 observations prior to the current run. These 5 observations are then passed as a dataset to Luminol. Given that we are limiting the query to only pull back the last 5 interval batches, if we have < 5, they should be ignored, as the application hasn't effectively 'gotten up to speed'. A lot of anomalies have been discovered, however this is because 1) the EMA algorithm is applying the threshold within the window of '5 observations', and 2) the detector keeps detecting even after an significant anomaly has been discovered. Expected behaviour would be to notify an administrator and then cease to continue, as the Level Shift has already occurred - there is no point in alerting further at this stage. The Level Shift occurred at 10:28 when the count went from 899 to 1975. The anomaly detector reported an anomaly score of	Scheduled	10 Secs	1.00	**5**	~1 Sec	17	567
RUN 4	A Level Shift is a surge in message volume, and as evidenced in Run 3, this equated to an anomaly score of 2.04. For this run the threshold has been set at 2.0. Interestingly, the Level Shift was not detected. The Threshold has been set too high. The next run will adjust the threshold to 1.5	Scheduled	10 Secs	**2.00**	5	~1 Sec	3	567
RUN 5	The Threshold at 1.5 has made little difference, as we still have a high anomaly count. The Limit window will be extended to ensure that the EMA algorithm has a wider field of view on what has already ovccurred.	Scheduled	10 Secs	**1.50**	5	~1 Sec	13	544
RUN 6	Extending the query to pull back the last 10 observations has reduced the number of anomalies. We are still seeing anomalies being reported after a level shift, however this is due to the application not having the logic yet to notify and terminate. One interesting result is at 11:44:00 where a score of 2.9 was reported, but not flagged as an anomaly. Possible Luminol bug? After 11:440:40 no further anomalies are detected. This is attributed to the application being terminated, before it's next 'batch of 10' is processed.	Scheduled	10 Secs	1.50	**10**	~1 Sec	8	633
RUN 7	For this run, the Threshold has been raised to 2.0. This run has proven to be a good indicator of the application working correctly. 2 Anomalies were detected - 12:18:40 is where the level shift occurs. Another anomaly is detected at 12:19:10 which is also a sizeable jump. As with the last run, after this point no further anomalies are detected. This is attributed to the application being terminated, before it's next 'batch of 10' is processed.	Scheduled	10 Secs	**2.00**	**10**	~1 Sec	2	633
RUN 8	Code has been added to stop the detector after the first observed anomaly. This occcurred at 13:02:40 with an anomaly score of 2.0252. The anomaly detector has been sucesssfully tuned against the combined training dataset.	Scheduled	10 Secs	2.00	10	~1 Sec	1	624
RUN 9	The detector was run once again, this time against a 'normal' dataset with no level shifts present. The same threshold, query interval and limit clause were used from Run 8. As expected, no anomalies were raised. This proves that the detection application is working successfully.	Scheduled	10 Secs	2.00	10	~1 Sec	1	599

Availability Prototype Results – Run 7

RUN 7 - Running detection - Combined Dataset A31 Messages - EMA Threshold 2.0 - Interval 10 - Limit 10

interval	count
12:16:20	723
12:16:30	352
12:16:40	795
12:16:50	927
12:17:00	587
12:17:10	493
12:17:20	664
12:17:30	412
12:17:40	889
12:17:50	644
12:18:00	373
12:18:10	878
12:18:20	488
12:18:30	707
12:18:40	1892
12:18:50	1982
12:19:00	2441
12:19:10	3113
12:19:20	1639
12:19:30	2209
12:19:40	2477
12:19:50	2569
12:20:00	3052
12:20:10	2781
Total	33087



Availability Scenario – Key Findings

Detection is achievable after tuning

- A successful detection achievable within **10 seconds** of the level shift event. This could be faster still.
- Multiple rounds of tuning the **threshold**, **limit** and **interval** variables is required.
- Representative data and testing is key.

A pipeline is mandatory

- Data preparation and pipeline / alerting is needed, though no training is needed.
- Data needs to be siphoned via the integration engine to a separate data store.
- This data store itself will need to be protected.

Initially accessible, but can become complicated

- General software development skills should be suitable
- Specialist knowledge would be required if delving into more advanced TS concepts

Conclusions

Confidentiality and **Availability** incidents are feasible, practical and accurate, however there are some clear implications:

- Data is a blocker. FHIR is difficult to work with.
- Pipelines and data transformation is needed.
- ML models will need specialist knowledge and maintenance. Question if you 'really, really need it'.
- TS Anomaly detection is generally accessible, but can become complicated.

Integrity detection

- Lots of opportunity for further work in this field.

Technology

- Python is really accessible, with very strong library support. R is great too, but is not, as developer 'mainstream'.
- Any analytic data store can also be an attack vector. Don't forget to protect it too.

For the curious

- **My code -** https://github.com/daveym/MSc
- **FHIR Definitions** - https://www.hl7.org/fhir/
- **HL7 Inspector** - http://hl7inspector.com/index.html
- **Scikit** - http://scikit-learn.org/
- **Jupyter Notebooks -** http://jupyter.org/
- **Luminol (LinkedIn)** - https://github.com/linkedin/luminol
- **SyntheticMass -** https://syntheticmass.mitre.org/about.html

Thank you!

kainos®

https://www.kainos.com/careers

@djmcglade

davey@kainos.com