Standards and tools for consuming data

Facilitator: Bill Brand, MPH, Public Health Informatics Institute | May 19, 2021
Finding your way around Zoom

- The audience is muted.
- Type your questions into the chat or use reactions to communicate with presenters.
UTILIZING PYTHON AND SELENIUM FOR WEB AUTOMATION

Shuennhau Chang
Data Analyst | Office of Data and Analytics
Indiana Department of Health

May 19, 2021
MOTIVATION

• Many electronic lab reports (ELR) auto-filed through an automated process built into state’s surveillance system
• Manual-entered reports will be in a review queue for further action
  • Lab reports and morbidity reports
• No automated process built into surveillance system to auto-file manual-entered reports.
• Solution: Web automation
TOOLS

- Python 3
- Jupyter Notebook
- Selenium module
- Other tools
  - Pandas module
  - Rapidfuzz module
GENERAL WORKFLOW

- Export review queue as csv
- Reads a flat file of all COVID-19 investigations, and extract the names for fuzzy match
- For loop
  - Search for the patient
  - Open the report
  - Various checks for data quality
  - Check if report can be associated to existing investigation
  - Fuzzy match name in the report to the names of all COVID-19 investigations.
  - Create investigation
  - Repeat
• Selenium
  • Wrapper function
• Rapidfuzz
  • String matching using the Levenshtein Distance
  • Like fuzzywuzzy but written in C++ for improved performance
• Pandas
  • Read csv, data manipulations, and creating csv log
from selenium import webdriver
from selenium.webdriver.common.by import By
from selenium.webdriver.support.ui import WebDriverWait
from selenium.webdriver.support import expected_conditions as EC

def click_xpath(path, waittime = waittime):
    WebDriverWait(driver, waittime).until(EC.visibility_of_element_located((By.XPATH, path)))
    WebDriverWait(driver, waittime).until(EC.element_to_be_clickable((By.XPATH, path)))
    driver.find_element_by_xpath(path).click()
from rapidfuzz import fuzz, process

# Check if patient name is similar to ones in open investigations
name_score = process.extractOne(name1, covid_open_inv_names, scorer=fuzz.token_sort_ratio)
name1 = name_score[0]
score1 = round(name_score[1])
DEMO: Manual Process
DEMO: Script Process
SUCCESS/CHALLENGES

Success
• Quickly review and process all reports
• During the fall spike, the script reviewed over two thousand reports a day; over 1,500 reports were processed without further review.

Challenges
• System slowness
• Reports with missing information
Thank you!

Questions?
From March 2020–March 2021 the ESSENCE-FL team worked to move all syndromic surveillance data feeds to ingestion using the Mirth Connect interface engine.

9 Data Sources
212 Emergency Departments (ED)
95 Free-Standing EDs
92 Urgent Cares
PHIN MESSAGING GUIDE FOR SYNDROMIC
SURVEILLANCE: EMERGENCY
DEPARTMENT, URGENT CARE, INPATIENT
AND AMBULATORY CARE SETTINGS

ADT MESSAGES A01, A03, A04 and A08
Optional ORU^R01 Message Notation for Laboratory Data

HL7 Version 2.5.1
(Version 2.3.1 Compatible)
Allow List has two columns—“Facility Name” and “Conversion Filter Date Time”

Does the EVN7-1 value match a Facility Name (i.e. known or unknown facility)? Is the PV1-44 value for the visit more recent than the Conversion Filter Date Time?

- Yes, Known Facility: Add into the staging table for processing by ESSENCE.
- Yes, Unknown Facility: Add into the staging table with a process value of -88. Will be ignored by ESSENCE.
- No: Discard.
Race/Ethnicity Standardization
- Single letter codes mapped to standard values to match HL7
- Value of “H” for Hispanic in race was moved to ethnicity

Discharge Diagnosis Concatenation
- Multiple lines of data for one visit were treated as separate messages. Each discharge diagnosis is a new line of data.
- Combined the messages within a file for one visit into one row in the staging table.
When onboarding a new facility, we look at message level data.

How does the data look in the staging table? Do each of the fields read correctly?
How many visits per day are we receiving?
Does it look like there will be a problem with duplication of visits?
How do their admission data look?

```
SELECT Distinct SendingFacility, Date, Count(Distinct MedRecNo) As Visits
FROM [Import_Staging].[dbo].[ER_Import_Staging_Table_Test] (NOLOCK)
Where SendingFacility in ('DP1', 'ER1', 'ER2', 'ER3', 'HC1', 'Hw1', 'SL1', 'SL8', 'SO1')
and Date > '2021-01-29' and (PatientClass like 'I' or DischargeDisposition like '09')
Group by SendingFacility, Date
```
How many discharge diagnoses are sent with each visit?
   Pull the last DG1-1 into a new “DD Count” column!

Are we receiving discharge diagnoses that were present for the patient before their current visit?
   Pull all DG1-5 into a new “DD Date Time” column!

Who is sending discharge diagnoses without a coded value?
   Pull DG1-3-1, DG1-3-2, and DG1-3-3 into separate columns!
Questions?

Thank you!

The ESSENCE-FL Team
Dave Atrubin, Shelby Fawaz, Mwedu Mtenga

JHU-APL
ESSENCE Developers
Wayne Loschen & Rich Wojcik

ISF
Database Hosting & Management
Jason Walters & Mitchell Hacker

Ruvos
Integration Broker
Jeff Couch & Juan Vasquez

Contact Information:

Samantha Spoto, MSPH
ESSENCE-FL Technical Analyst
Florida Department of Health

Samantha.Spoto@flhealth.gov

Office: 813-233-2366
OUTLINE

• What is FHIR?
• What can FHIR do for me -- as a public health practitioner?
• What Use Cases are catching FHIR? 🔥
• Questions/Discussion
What is FHIR?

- Fast Healthcare Interoperability Resources (FHIR)
- HL7 standard (2014 → Version 4.0.1 published 10/2019)
- Required by CMS in 21st Century Cures Act final Rule
- Allows different systems to exchange healthcare data in a standard format
- Differs from previous HL7 standards:
  - uses technology common in industry (e.g., APIs)
  - ability to share resources in XML, JSON or RDF
- ‘Resources’ define common healthcare concepts and relationships (*80/20 rule)
  - less common requirements addressed using ‘Profiles’

https://www.hl7.org/fhir/overview.html
RESOURCES

Provides a model for organizing information

https://www.hl7.org/fhir/resourcelist.html
SMART STANDARD + FHIR

- Allows third-party apps to be embedded in EHR, enforcing authentication
- Has widespread EHR vendor support
- Uses FHIR standard to read (and sometimes write) data from the EHR
SMART ON FHIR ARCHITECTURE:
CORE (BLUE)+ OPTIONS USED AT UNIVERSITY OF UTAH HEALTH (BROWN)

Usual authentication
Initiate SMART Launch
SMART Launch
Smart App
User
Data Retrieval
Display
Obtain Access Token
‘FHIR Wrapper’ (to add needed data or filter unnecessary data)*
Decision support (to evaluate logic)

*See discussion about scope of data pulled using FHIR: http://www.hl7.org/fhir/smart-app-launch/scopes-and-launch-context
WHAT CAN FHIR DO FOR ME?

Public health informatics:

Application of informatics methods and tools to support public health goals

Use Cases to discuss:

Healthcare

Clinical decision support (CDS) **within** healthcare setting to support PH goals:
1. Neonatal bilirubin care
2. Lung cancer screening

Public Health

**Push** from Healthcare to PH
3. electronic Case Reporting
4. Death certification

**Pull** from Healthcare to PH
5. Case investigation and follow-up
1. Pull:
- Lab data
- Risk factors
- Age (postnatal & gestational)

2. Apply logic

3. Display data in context and recommendations

Neonatal Bilirubin Care


1. Pull:
   - Smoking history
   - Risk factors
   - Demographics: Age, race, sex
2. Apply logic
3. Display data in context and recommendations
1. Pull data needed for a case report using FHIR.
2. Apply logic
3. Transform ‘FHIR data bundle’ into eICR to send through existing eCR pipeline

**ELECTRONIC CASE REPORTING**

**Key:**
- RCKMS - Reportable Condition Knowledge Management System
- eRSD – Electronic Reporting and Surveillance Distribution System
- eICR - Electronic Initial Case Report CDA v1.1
- RR - Reportability Response CDA v1.0

https://ecr.aimsplatform.org/general/ecr-now-fhir-app
VISION: DEATH CERTIFICATION

1. Pull data needed for a death record.
2. Allow MD to add other data
3. Apply logic
4. ? Check literals and logic with NCHS/CDC ‘VIEWS’ service
5. Send data to PH in standard structure

User image courtesy of Jennifer Herrmann
VISION:
PULL EHR DATA FOR F/U INVESTIGATION

Is this feasible?
What questions need to be asked?
- Security issues?
- Quality issues?
What are the PH requirements?
WHAT CAN FHIR DO FOR ME?

Public health informatics:

Application of informatics methods and tools to support public health goals

Use Cases to discuss:

**Healthcare**

Clinical decision support (CDS) **within** healthcare setting to support PH goals:
1. Neonatal bilirubin care
2. Lung cancer screening

**Public Health**

**Push** from Healthcare to PH
3. electronic Case Reporting
4. Death certification

**Pull** from Healthcare to PH
5. Case investigation and follow-up
Questions or comments?

Catherine Staes, PhD, MPH, RN, FACMI, FAMIA
Professor & Director, Nursing Informatics, College of Nursing
University of Utah
catherine.staes@hsc.utah.edu
801-213-3351 (but email is best)
Next Steps

• Continue the discussion on the message board (link provided in the chat)

• Ten-minute break: 2:15– 2:25 PM EST

• Next session: 2:25 – 3:25 PM EST
  • Planning for the future of the informatics workforce
Thank you.