Streamlining and Simplifying Data Flow
from your desktop

Chris Waldron
Robert Gottlieb
Garry Raynor

October 30, 2018
Why spend time pushing data from place to place... when we can automate and speed up these processes?
Our Examples       Your New Ideas

EZ Editor For Non-GIS Staff
Collect Data

Hepatitis A Outbreak Dashboard
Distribute Data

Emergency Department Analysis
Analyze Data

Resources, Scraping, & Stats
Move, Etc Data
Overview of GIS at ISDH: Public Health Geographics

Chris Waldron
GIS Coordinator

Robert Gottlieb
GIS Analyst

Garry Raynor
GIS Analyst

Program Areas and Workforce
GIS Desktop and Web Users
Overview of GIS at ISDH: Public Health Geographics
Overview of GIS at ISDH: Public Health Geographics

GeoSpatial/Statistical Analysis
Spatial concepts, functions, relationships, clustering, aggregation, statistics and detection

Application Development
Programming & application development, web apps, desktop tools, web services, system integration, support

Data Administration
Relational geodatabase design & maintenance, data generation & editing, spatial-enabling

Cartography
Geovisualization, insights, geospatial techniques, reporting, map use and comprehension

GIS
ANALYZE
CREATE
BUILD
VISUALIZE
Overview of GIS at ISDH: Public Health Geographics

Current Tools and Skills

ArcGIS Products, Spatial Concepts and Analysis Methods, Statistics, Database design (Oracle), SQL, Python, Javascript, JSON, HTML, UX, .NET, VBA and Office, Cartographic Concepts and Design, Census, Population Health and Epi Basics, Tableau and other BI, GIS Support…

Future Skills Required

Office 365 and AI
The Best Tool is Always Changing

Automating and Streamlining will be different by the 2019 GIS Day...
Current Tools for Problem Solving

**Python**
(ArcPy and ArcGIS API for Py)

**JavaScript**
(Vanilla + Esri JavaScript API)

**SQL**

**ArcGIS OnLine**
Problem:
Non-technical staff need to update stats and other information contained in maps and lists on web pages (without the GIS UX)

Solution:
Streamline with “EZ Editors” via JavaScript and AGOL
What apps do the EZ Editors Support?

**West Nile Virus**

<table>
<thead>
<tr>
<th>Marion County</th>
<th>Stats Explorer</th>
<th>STD Clinics</th>
</tr>
</thead>
<tbody>
<tr>
<td>2018 Estimated Exposures</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Please select an option below. County</td>
<td></td>
<td></td>
</tr>
<tr>
<td>counties on map represent local lab-confirmed positive results during 2018. Map is updated 8 times</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Updated: May 15 2018 07:05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EDT (Eastern Daylight Time)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Stats Explorer**

- West Nile Virus
- WNV Mosquito Infections (66)
- WNV Human Disease Cases (15)
- WNV Bird Infections (0)
- WNV Equine Disease Cases (15)

**STD Clinics**

- Eastern Equine Encephalitis
- EEE Mosquito Infections (0)
- EEE Human Disease Cases (0)
- EEE Bird Infections (0)
- EEE Equine Disease Cases (0)

**Others:** PDO, IPV, HAB, District Staff, HIV, Safe Sleep, CFR, PreDiabetes, Response Toolkit
Please make any necessary changes by using the form below. Changes will not be committed to the table until the 'Submit Changes' button is pressed. The form may be closed at any time by pressing the 'Close Form' button, or by selecting the 'X' icon in the upper right corner. Field entry boxes may also be enlarged or compressed by clicking and dragging the bottom right-hand corner.

**Sampling Site:**
- Eagle Creek

**Collection Date:**
- N/A

**Cell Count:**
- 0

**Anatoxin A:**
- ND

**Microcystin:**
- ND

**Cylindrospermopsis:**
- ND

**Advisory Level:**
- None

Submit Changes  Close Form  Reset Form  Clear Form

**Advisory Level:**
- None

**Advisory Impact:**
- Don't drink the water. Shower after you swim. Keep pets out of water.

**Advisory Level:**
- None

**Advisory Impact:**
- Don't drink the water. Shower after you swim. Keep pets out of water.

**Advisory Level:**
- None

**Advisory Impact:**
- Don't drink the water. Shower after you swim. Keep pets out of water.

**Advisory Level:**
- None

**Advisory Impact:**
- Don't drink the water. Shower after you swim. Keep pets out of water.
EZ Editors – Technical Requirements

AGOL – Hosted Data and Security

• Hosted Feature Layer (a table of data with or without spatial)
• State Credentials and AGOL Roles
• App Registration (where the app is just a .html file)

Web App – Vanilla Javascript

• Esri API 3.x
• esri/layers/FeatureLayer
• esri/dijit/FeatureTable
Problem:
We need to quickly pull data from spreadsheets into dashboards, maps, and emails for Hepatitis A Outbreak situational awareness.

Solution:
Python and AGOL Operations Dashboard
What is hepatitis A?
Indiana is experiencing an outbreak of hepatitis A, a highly contagious liver infection caused by the hepatitis A virus. It can range from a mild illness lasting a few weeks to a severe illness lasting a few months.

What are the symptoms?
Not everyone has signs of hepatitis A. If symptoms do develop, they usually appear two to six weeks after infection and can include:

- Abdominal pain
- Fever
- Dark urine
- Nausea
- Vomiting
- Jaundice (yellowing of the eyes)
- Pale stools (poop)
- Loss of appetite
- Joint pain
- Diarrhea
- Fatigue (extreme tiredness)

Symptoms are more likely to occur in adults than in children. They usually last fewer than two months, but some people can be ill as long as six months.

What can I do to prevent hepatitis A?
Get Vaccinated! Hepatitis A vaccine is readily available from healthcare providers and pharmacies. Good handwashing with soap and running water, especially before preparing food and after using the restroom, can also help prevent the spread of hepatitis A.
Hepatitis A Outbreak

Indiana Hep A Cases

2017  2018
Hepatitis A Outbreak Dashboard

https://www.in.gov/isdh/27791.htm

Outbreak Cases: 578
Outbreak Deaths: 2
Hospitalizations: 260

Vaccines Administered: 79,240

Includes public and private doses since Jan 1, 2018
ArcGIS Operations Dashboard*

Operations Dashboards and Survey 123 are the most versatile and user-friendly tools you are not using

*Tableau nor Power BI required
Hepatitis A County Stats Workflow

Daily From Program Area

- **Cases**
  - Excel (xls sheet → .csv)
  - Network Folder

- **Immunizations**
  - Excel (oracle → .xls)
  - Network Folder

Friday From Program Area

- **Map for Daily Report**
  - Email Status

- **Check Data Structure**
- **Email Errors**
- **Update AGOL Table: Cases, Hospitalizations, Deaths, Immunizations fields**
- **Email xls**

**Daily Script at 6pm**

- ArcPy Mod

**Friday Script at 11am**

- ArcGIS Mod
  - XLRD Mod
Any app connected to tables (REST endpoint) reflects changes made by script.

Friday Script at 11am

Update AGOL Table (Case, Hosp, Imz Columns)

Hep A Webpage Dashboard

Hep A in Stats Explorer

Tableau, GIS, Excel, etc
### Windows Task Scheduler

#### Task Scheduler (Local)
- **Task Scheduler Library**
  - Adobe Acrobat
  - Adobe Flash
  - Google Update
  - SCCM_7nen

#### General Settings
- **Status**: Quizzed
- **Triggers**: Multiple triggers defined
- **Next Run Time**: 10/22/2018 10:00 PM
- **Last Run Time**: 10/22/2018 9:28:15 AM
- **Last Run Result**: The operation completed successfully. (NM)

#### Actions
- Create Basic Task...
- Create Task...
- Import Task...
- Display All Running Tasks
- Disable All Tasks
- History
- Run
- End
- Disable
- Export...
- Properties
- Delete
- Help

#### Additional Settings
- **Task to be run on demand**
- **Run task as soon as possible after a scheduled start is missed**
- **If the task fails, restart every**: 1 minute
- **Attempt to restart up to**: 3 times
- **Stop the task if it runs longer than**: 1 day
- **If the running task does not end when requested, force it to stop**
- **If the task is not scheduled to run again, delete it after**: 30 days
- **Do not start a new instance**
Problem:
We need to conduct daily spatial analysis of ED data and distribute results

Solution:
Python
Syndromic Surveillance

- Indiana hospitals send emergency department visit records to Public Health Electronic Surveillance System (PHESS)
- PHESS data fed into ESSENCE – Electronic Surveillance System for the Early Notification of Community-Based Epidemics
- ISDH epidemiologists use queries for various conditions
- ESSENCE identifies statistically significant alerts for a county or hospital
- Epidemiologists review and send alerts to local health departments
PHESS Data Flow

Emergency Department Reports → PHESS
Emergency Department Reports → SQL
Emergency Department Reports → ESSENCE

Enhanced Surveillance Density

ISDH Epis
Review and Send To LHDs and Hospitals
Sample ESSENCE Query Results

Daily Data Counts

Data: Normal  Data: Warning  Data: Alert
Enhancing Syndromic Surveillance

- Can we identify areas of increased activity at the neighborhood level?
- Can we look at trends for different time periods?
- Can we automate this process?
Calculating Neighborhood Density

- Python script runs each morning – Windows Task Scheduler
- Run SQL query to get emergency department records
- Deduplicate and geocode
- Geocoding is automated. Won’t be able to review every address (about 92% geocoded to point or street segment)
- Calculate count, rate, and difference in hexagonal grids
- Calculate kernel density and difference
- Flexible time periods and spatial scales
Neighborhood Density

- Kernel density converts points into a 'cloud' representing a concentration of points
- Easier visualization of points
- Allows us to look at neighborhoods while protecting privacy
- Represent concentrations of points without worrying about borders
Creating Maps

- ArcPro project created
- Output GEOJSON files
- Folium – open source Python package
- Can clip data to restrict map to single county for private distribution to local health department
Folium

• Open source Python package (version 0.6)
• Can work with JSON files, geopandas
• Creates maps in HTML file
• Some features more convenient than others
• Other mapping options – Leaflet (Javascript), Pysal, Shapely, Cartopy, Geopy
Summarized Map Data in Self-Contained HTML File

Example
Problem:
We need to query multiple data sources and update our health resources in our Oracle database, the GIO Library, and ArcGIS On-Line

Solution:
Python to glue ETL
Health Resource Data
Over 150 layers of health resources (e.g. hospitals, clinics, etc)

Resource Explorer
AGOL Open Data

https://gis.in.gov/apps/isdh/meta/resources_layers.htm
https://data-isdh.opendata.arcgis.com/
Health Resource Data
Over 150 layers of health resources (e.g. hospitals, clinics, etc)

ArcGIS OnLine Organization

- Asthma Resources - Clinical
- Asthma Resources - General
- Birthing Center
- Blood Center
- Breast and Cervical Cancer Program Provider
- Certified Primary Stroke Centers
- CLEAR
- Clinical Lab - All CLIA
- Clinical Lab - Waivers
- Community Federally Qualified Health Centers
- Community Mental Health Center
- Community or Federally Qualified Health Centers
- Diabetes Self-Management Education

Feature Layer (hosted, view)
Problem:
No time to keep contact information updated for the map on our website.

Solution:
Web Scrape with Python
Connect to hosted feature layer.

Query hosted feature layer to return all results.

Parse returned results.

1. Pass department name and url to loop.

2. Use urllib to test url.

3. Determine if url is valid or invalid.

4. If invalid, determine reason if possible.

5. Append department name and reason to list.

while loop

Determine if any invalid urls are present in list.

Determine email message.

Send email and exit.

If valid, move on to step 1 with the next department.
Problem:
Stats are scattered all over the ISDH website which make finding health stats difficult

Solution:
Streamline Stats Discovery for End-Users with ArcGIS and JavaScript
Stats Explorer
A new path to Indiana Health Data
gis.in.gov/apps/isdh/StatsExplorer

Hepatitis C, Chlamydia, Legionellosis, Tuberculosis, SIDS, Low Birthweight Infants, Deaths from Drug Poisoning, Hospital Discharge, Cancer Incidence, Lyme Disease, Varicella, Animal Bites, Syphilis, Prenatal Care, Diabetes Prevalence, Asthma, Hepatitis B, Infant Deaths, Rabies in Bats, HIV/AIDS, Influenza Deaths, Gonorrhea, Motor Vehicle Accidents, Diseases of the Heart, Hepatitis A, Mothers Medicaid, Rabies, Meth, Nicotine Acquired Immune Deficiency Syndrome, Sexually Transmitted Disease, Cerebrovascular Disease, Intentional Self Harm, Encephalitis, Alzheimer’s Disease, Mothers Smoking During Pregnancy, Murder, PID, Mothers Breastfeeding at Discharge, Suicide Attempt Death, Premature Infants, Gonococcal Infection, AIDS Death, Total Neonatal Deaths, Employees by the Public Health Department, Trauma Center, Birth, Etc.

gis.in.gov/apps/isdh/statsexplorer
in.gov/isdh/26720.htm
Design Goals (2016)

- Simple User Experience
- Systematic Organization of Data
- Use Existing/Published Data
- Dynamic/Interactive
- Single Data Destination
- Open Data
- Minimal Additional Funds
Flow of Analyzed Data

Data (Record-Level) → Data Analysis → Stats/Reports (Aggregated)
ISDH Public Stats/Reports Commonality

- Geography (County, Facility)
- Temporal (Annual)
- Statistic (Rate, Count)
- Stratifications (Age, Sex, Race)
Most Public Data Are Shared via Excel...but all different styling/formatting

Currently 200+ variables and 500,000+ data points for one or more time periods (single years, range) some stratified by sex, race

<table>
<thead>
<tr>
<th>Table</th>
</tr>
</thead>
<tbody>
<tr>
<td>VAR_ID</td>
</tr>
<tr>
<td>GEOG_ID</td>
</tr>
<tr>
<td>GEOG_NAME</td>
</tr>
<tr>
<td>GEOG_TYPE</td>
</tr>
<tr>
<td>DATE_START</td>
</tr>
<tr>
<td>DATE_END</td>
</tr>
<tr>
<td>PERIOD</td>
</tr>
<tr>
<td>DATE_NAME</td>
</tr>
<tr>
<td>VAR_STAT</td>
</tr>
<tr>
<td>VAR_COUNT</td>
</tr>
<tr>
<td>VAR_DENOM</td>
</tr>
<tr>
<td>VAR_UNSTABLE</td>
</tr>
<tr>
<td>VAR_LCI</td>
</tr>
<tr>
<td>VAR_UCI</td>
</tr>
<tr>
<td>VAR_CI</td>
</tr>
<tr>
<td>VAR_SIGVAL_NAT</td>
</tr>
<tr>
<td>VAR_SIG_NAT</td>
</tr>
<tr>
<td>VAR_SIGVAL_STATE</td>
</tr>
<tr>
<td>VAR_SIG_STATE</td>
</tr>
<tr>
<td>VAR_SIGVAL_DIS</td>
</tr>
<tr>
<td>VAR_SIG_DIS</td>
</tr>
<tr>
<td>VAR_SIGVAL_CNTY</td>
</tr>
<tr>
<td>VAR_SIG_CNTY</td>
</tr>
<tr>
<td>ROW_ID</td>
</tr>
<tr>
<td>VAR_RACE</td>
</tr>
<tr>
<td>VAR_SEX</td>
</tr>
</tbody>
</table>
Extending Current Processes

Current Data Analysis and Reporting Process

Existing Infrastructure And Existing Skillsets

VAR_ID
GEOG_ID
GEOG_NAME
GEOG_TYPE
DATE_START
DATE_END
PERIOD
DATE_NAME
VAR_STAT
VAR_COUNT
VAR_DENOM
VAR_UNSTABLE

ArcGIS Platform

Web API’s
Why GIS?

*leverage existing infrastructure*

**GIS Database**
- Oracle

**GIS App Server**
- Data Source
  - Query

**GIS Web Server**
- Data Service
  - Connect
- Script
- HTML

<table>
<thead>
<tr>
<th>ID</th>
<th>Info</th>
<th>Info</th>
<th>Geo</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Data</td>
<td>Data</td>
<td>39.22</td>
</tr>
<tr>
<td>2</td>
<td>Data</td>
<td>Data</td>
<td>41.32</td>
</tr>
<tr>
<td>3</td>
<td>Data</td>
<td>Data</td>
<td>37.91</td>
</tr>
</tbody>
</table>
Hepatitis B, Chronic

Confirmed and probable case investigations that include either acute or chronic reported through the Indiana – National Electronic Disease Surveillance System (I-NEDSS) and are based on county where the case was investigated. Year is determined based on first positive laboratory specimen date. All rates per 100,000 population. Counts include those who were incarcerated in the Indiana Department of Corrections or Federal Bureau of Prison Facilities at the time the investigation was closed (Cass, Clark, Noble, Hendricks, Henry, Jefferson, LaPorte, Madison, Marion, Miami, Parke, Perry, Putnam, Sullivan, Vigo). More Information about Hepatitis B, Chronic.

Data Info: County level disease counts with cases status of confirmed. Counts that are less than 5 are suppressed. Rates based on fewer than 20 cases are considered unstable. Year represents the year data was reported to CDC per MMWR Calendar year found at: https://www.cdc.gov/mmwr/downloads.html

Resource source: Indiana State Department of Health, Epidemiology Resource Center
Variable ID: ‘DISEASE059’

Hepatitis B, Chronic
2011 - 2015

[Graph showing hepatitis B chronic rates for various counties in Indiana]
...with Transportable/Embeddable Visuals
with an Open Data Service

https://gis.in.gov/arcgis/rest/services/ISDH/IndianaHealthStats/FeatureServer/0/query
Open Data = Data Everywhere
one single source

http://www.in.gov/isdh/27358.htm

Data streamed from Stats Explorer to a map on a webpage
Open Data = Data Everywhere

one single source

Data streamed from Stats Explorer to an Excel Spreadsheet
Open Data = Data Everywhere

one single source

Data streamed from Stats Explorer to your own dashboard

All Stats are counts by residence. For more data and information, please visit Stats Explorer.

https://arcg.is/1TnTua
Problem:
How do I get started scripting?

Solution:
Python comes installed with ArcGIS
Visual Studio Code is good for JavaScript
https://www.esri.com/training/
Acknowledgements

• Kayley Dotson, Amanda Billman, Nicole Stone, Shawn Richards, Charles Clark

Public Health Geographics

• Chris at cwaldron@isdh.in.gov
• Robert at rgottlieb@isdh.in.gov
• Garry at graynor@isdh.in.gov