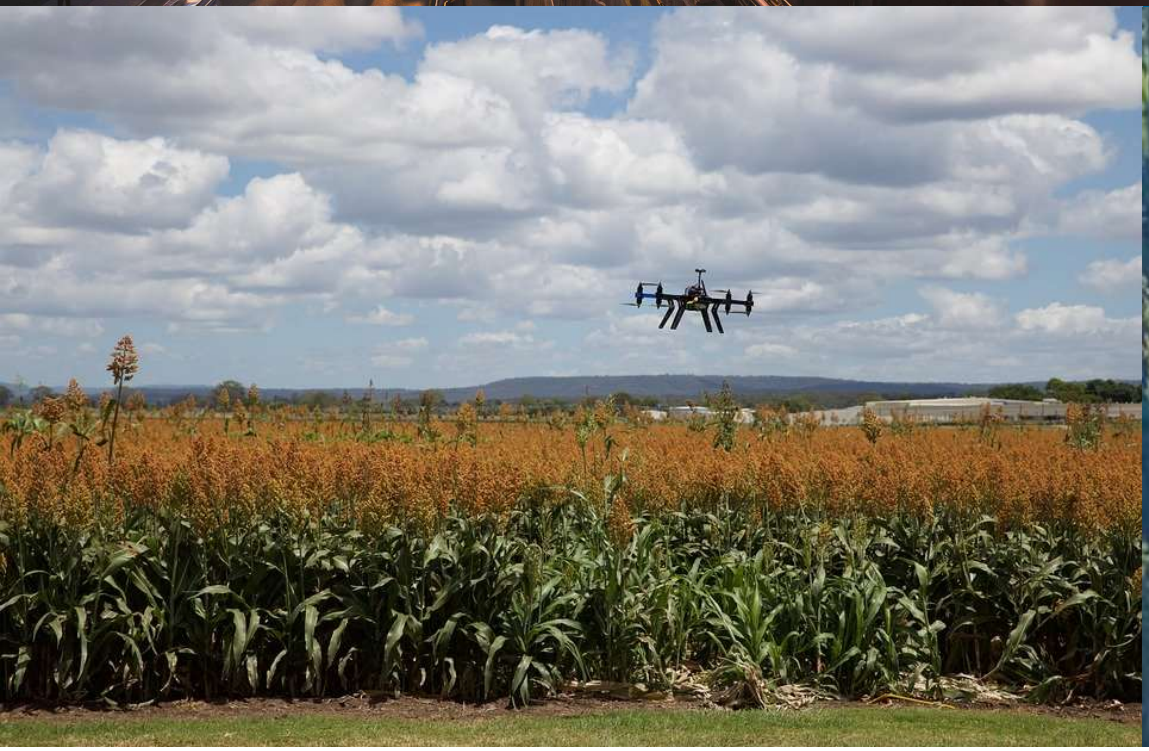
GIS & Spatial Machine Learning: Transforming Our Planet's Pulse to Action

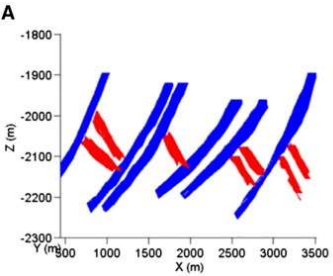
Orhun Aydin, PhD

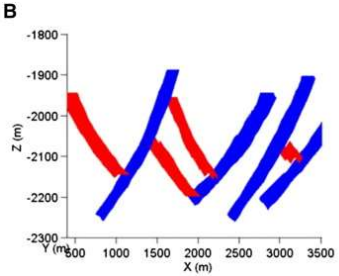
Senior Researcher

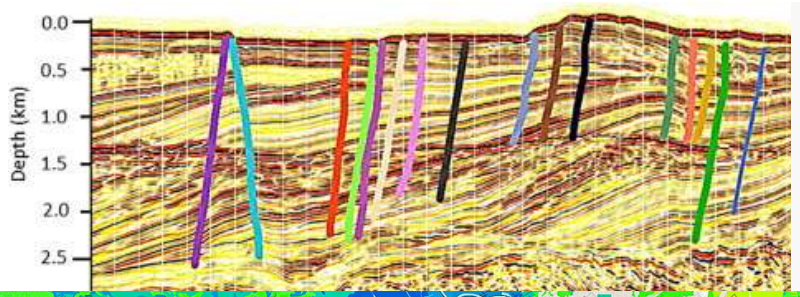
oaydin@esri.com

@orhun_sow

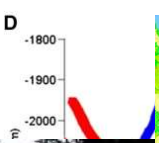



A
-1800
-1900
-2000
-2100
-2200
-2300
Z (m)
Y (m)
500 1000 1500 2000 2500 3000 3500
X (m)


B
-1800
-1900
-2000
-2100
-2200
-2300
Z (m)
Y (m)
500 1000 1500 2000 2500 3000 3500
X (m)


0.0
0.5
1.0
1.5
2.0
2.5
Depth (km)

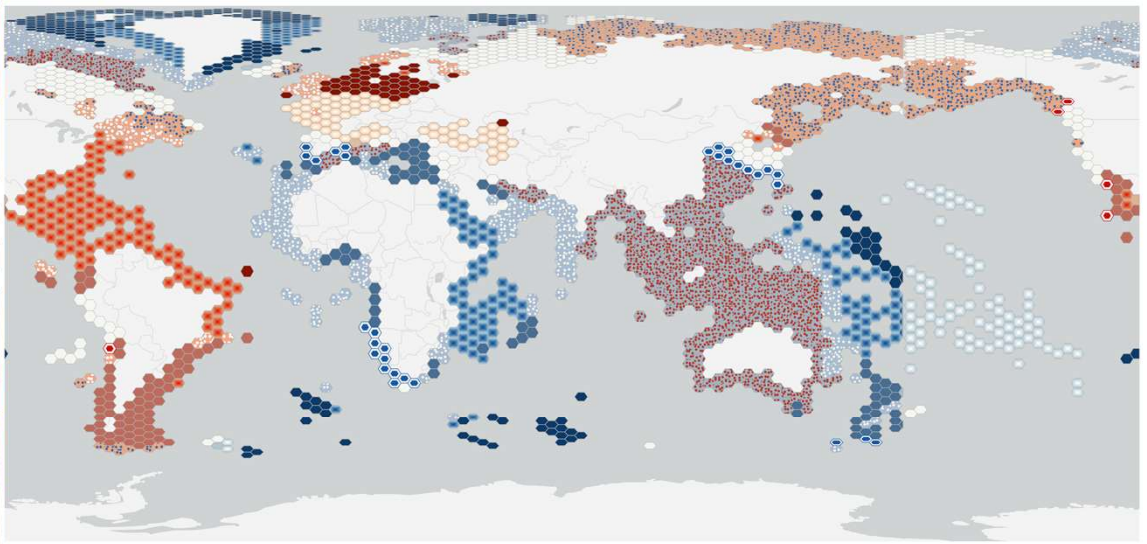

C
-1800
-1900
-2000
(m)


D
-1800
-1900
-2000
m)


REDL
UNIVERSIT
FOR MEETING
TIME AND PLACE CALL
793-2546

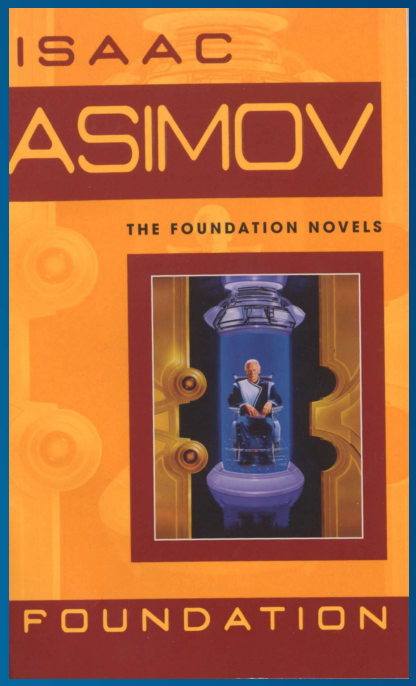

es


Spatial Statistics Tools
Analyzing Patterns
Mapping Clusters
Measuring Geographic Distributions
Modeling Spatial Relationships
Utilities




New
Recently Suitable
Increasingly Suitable
Consistently Suitable
Decreasingly Suitable
Sporadically Suitable\Unsuitable
Recently Improving
Lost
Newly Unsuitable
Consistently Declining
Increasingly Unsuitable
Consistently Unsuitable
Decreasingly Unsuitable
Sporadically Unsuitable/Suitable
Recently Declining
Marginally Unsuitable
No Pattern Detected

Geographer’s Guide to Galaxy (and this Keynote)

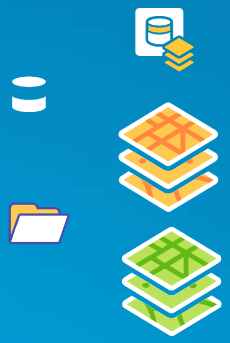


The Way We Think & Create

Need for Interoperable GIS


Advanced Analysis,
Scientific Methods,
Tools, Algorithms
Description
Prediction
Diagnostics
Prescription

Data Sources → Analysis Framework → Relevant Tools → Data Products



Spatial
Temporal
Tabular
Written Resource

ArcGIS Toolbox
ArcPy
Tensorflow
PySAL
R Packages

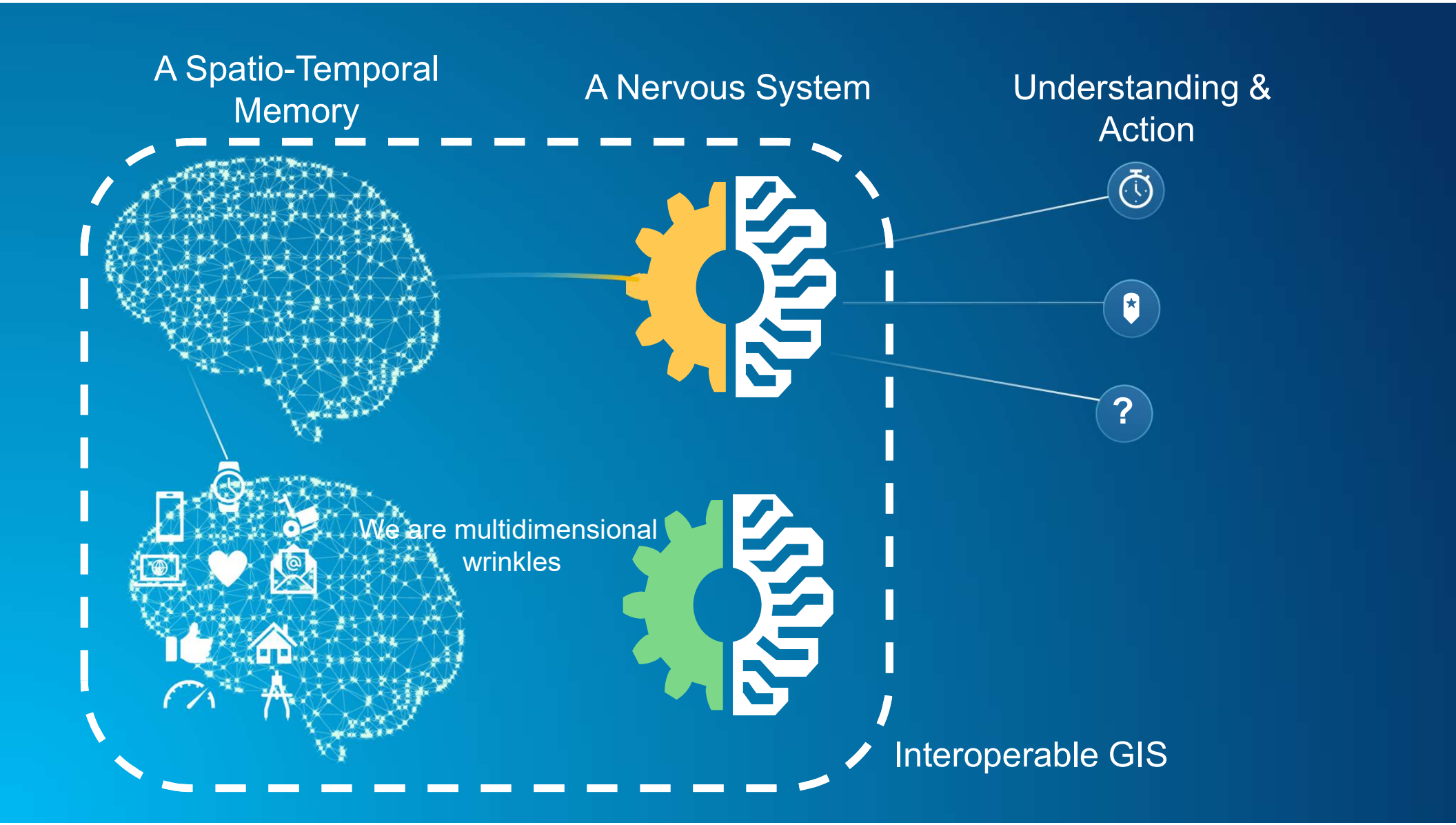
Hot Spot Analysis
Deep Neural Network
MaxEnt
Mean Center

Map!
Widget

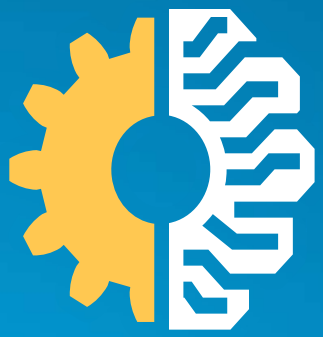
Think Spatially
For Spatial Problems

Find a Framework

Develop/Use Tools

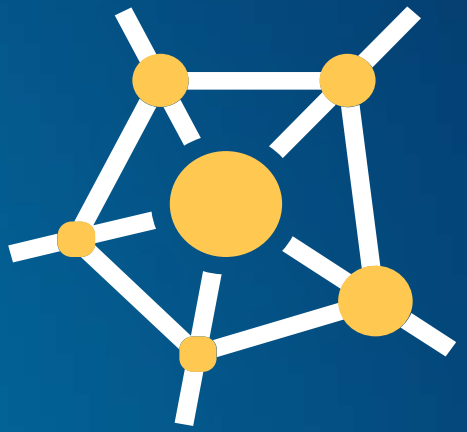

A Spatio-Temporal Memory
A Nervous System
Understanding & Action
?
We are multidimensional wrinkles
Interoperable GIS

Outline for Today

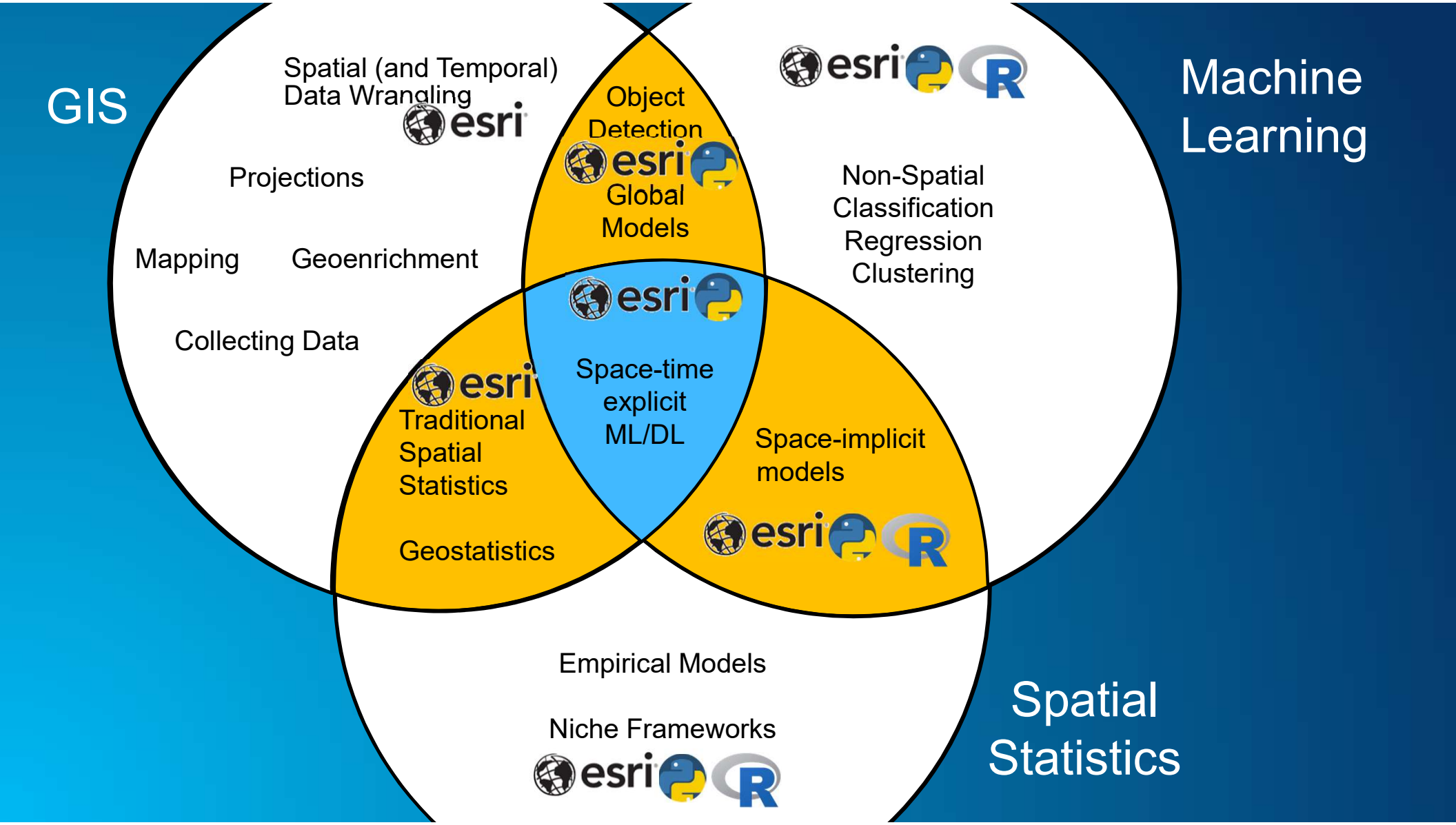


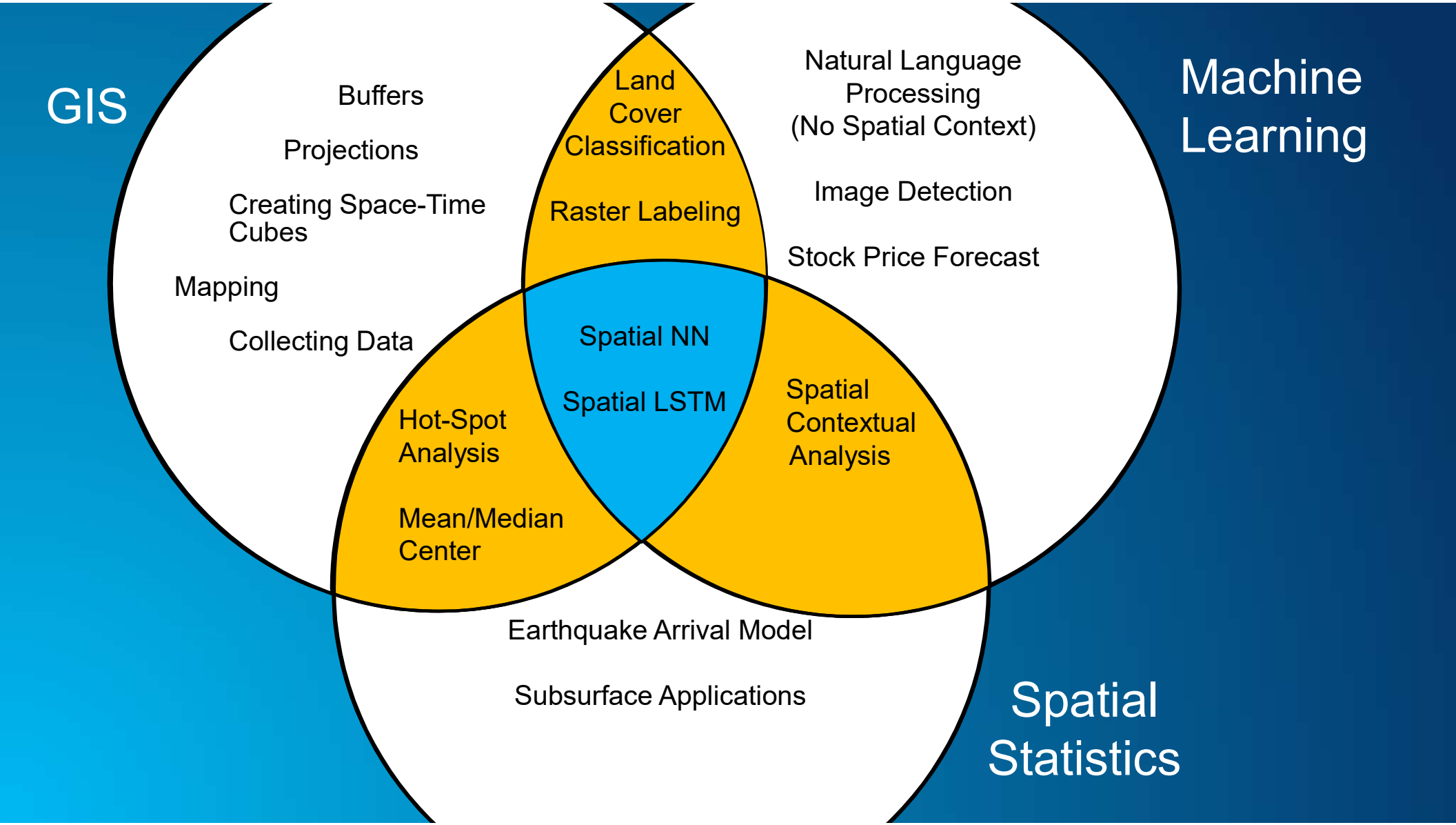
- Data Science in GIS
- Spatial Data Science
- Machine Learning
- Deep Learning
- Artificial Intelligence

- Our Planet's Memory
- Citizen Science
- Living Atlas

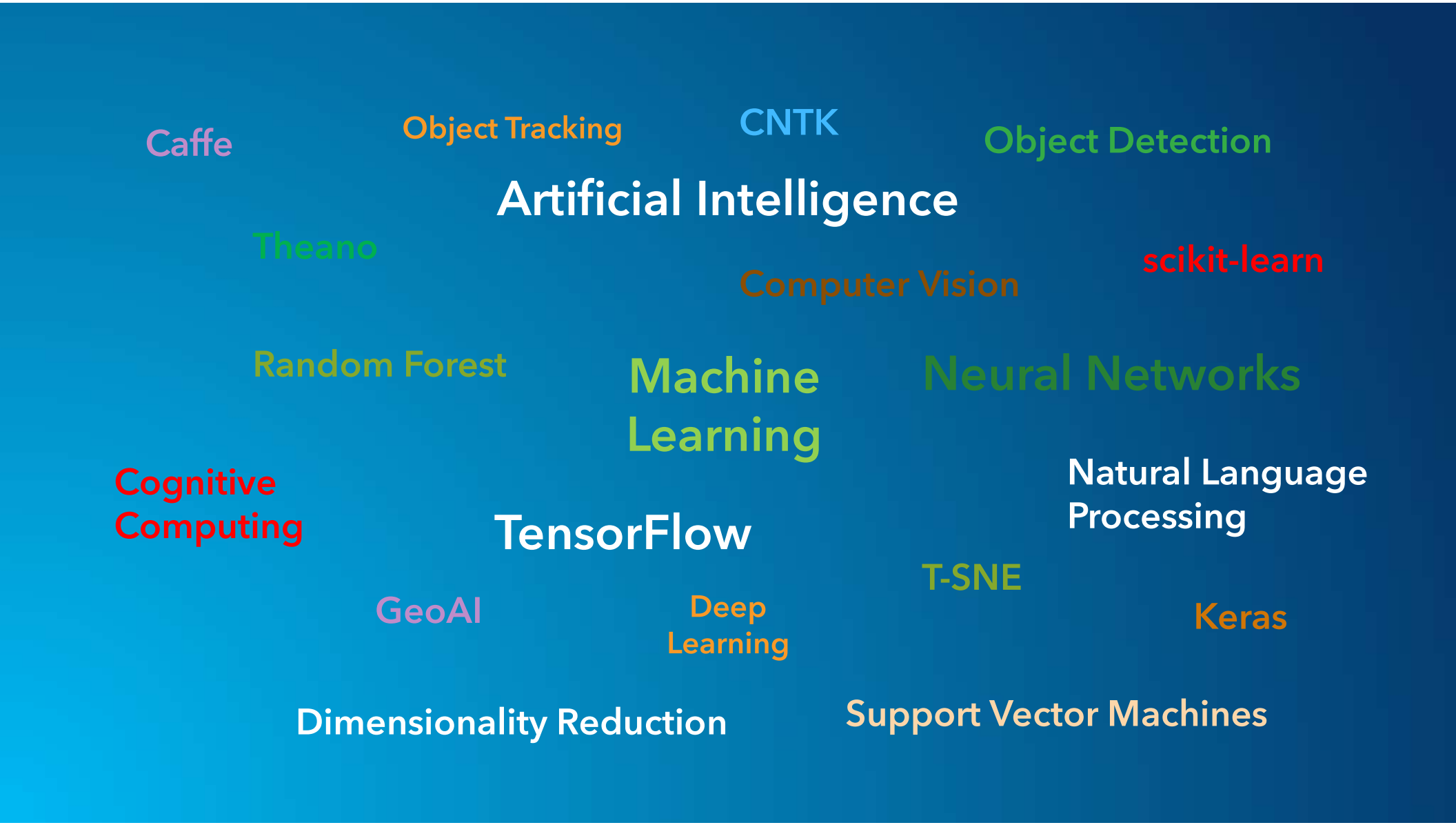


- Integration with Platforms
- Integration for Machine Learning
- Integration for Data Science


GIS
Machine Learning
Spatial Statistics
Spatial (and Temporal) Data Wrangling
Projections
Mapping
Geoenrichment
Collecting Data
Object Detection
Global Models
Non-Spatial Classification
Regression
Clustering
Space-time explicit ML/DL
Traditional Spatial Statistics
Geostatistics
Space-implicit models
Empirical Models
Niche Frameworks


GIS
Buffers
Projections
Creating Space-Time Cubes
Mapping
Collecting Data
Land Cover Classification
Raster Labeling
Natural Language Processing (No Spatial Context)
Image Detection
Stock Price Forecast
Machine Learning
Spatial NN
Spatial LSTM
Hot-Spot Analysis
Mean/Median Center
Spatial Contextual Analysis
Earthquake Arrival Model
Subsurface Applications
Spatial Statistics

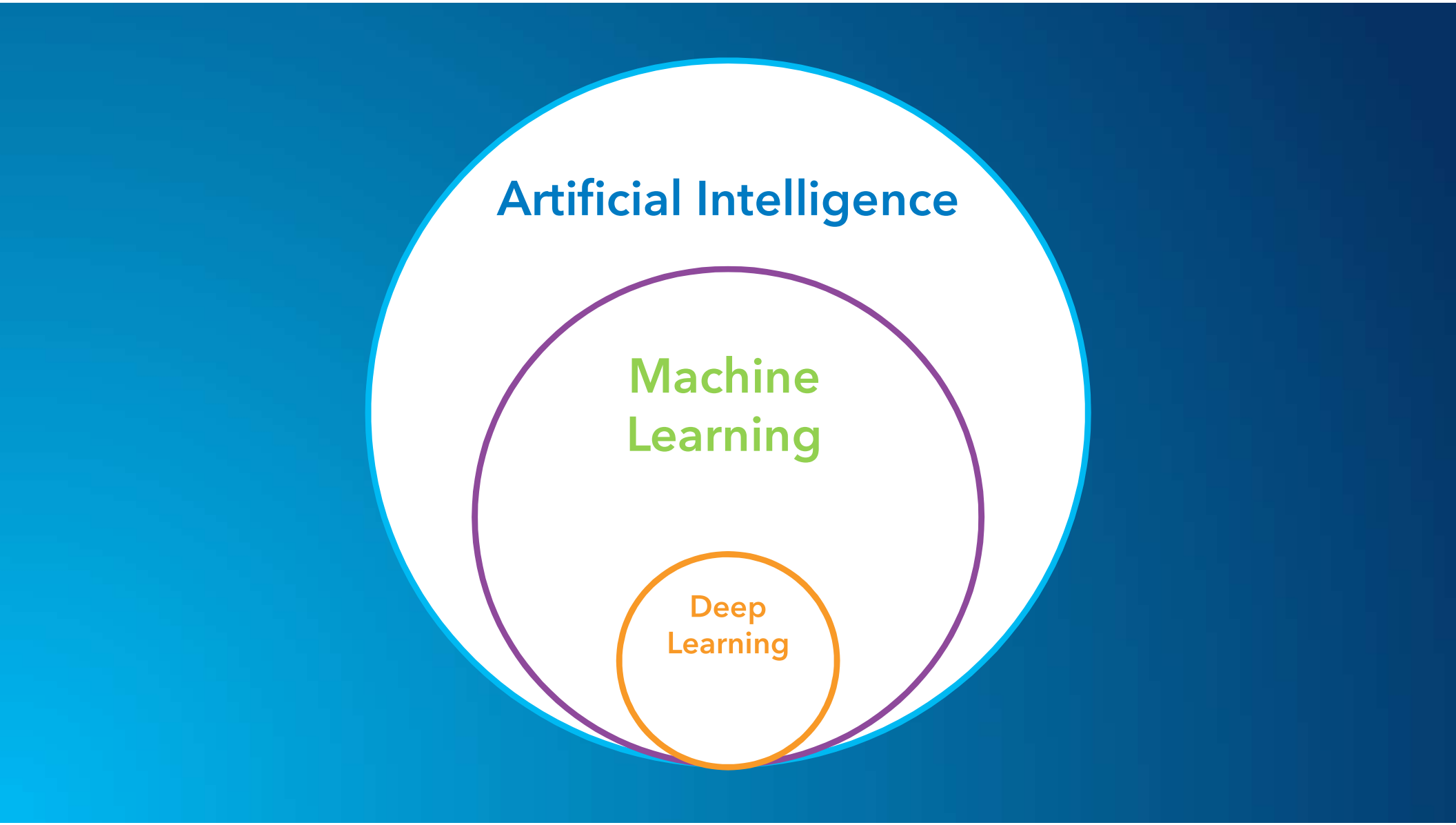
Artificial Intelligence
Data Science
Machine Learning
Deep Learning
…


Caffe
Object Tracking
CNTK
Object Detection
Artificial Intelligence
Theano
scikit-learn
Computer Vision
Random Forest
Machine Learning
Neural Networks
Cognitive Computing
Natural Language Processing
TensorFlow
T-SNE
GeoAI
Deep Learning
Keras
Dimensionality Reduction
Support Vector Machines

Artificial Intelligence

Machine Learning

Deep Learning


Artificial Intelligence
Machine Learning
Deep Learning

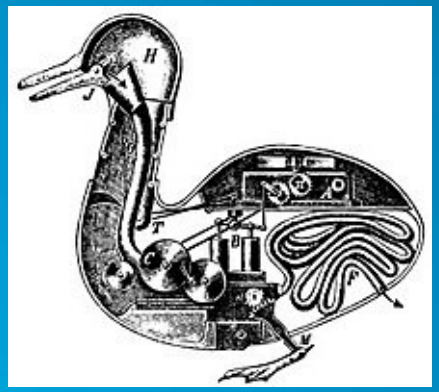
Artificial Intelligence: Pre-Machine Learning

- Programs with common sense[1] (McCarthy, 1960) → Set of **predefined** logical operators

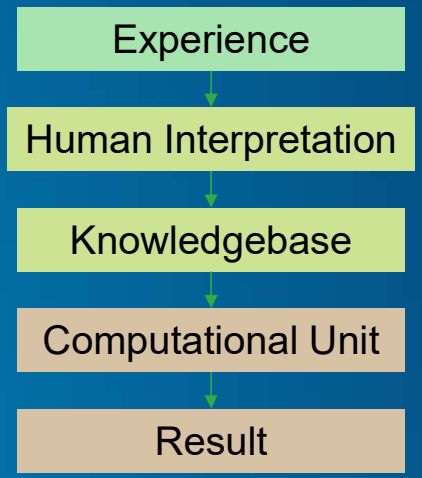
  is (GIS Day, Nov 13th), *in* (Nov 13th, today) → *in* (GIS, today!)

- Complex and extensive representations of human knowledge- Knowledgebases
  - CYC[2]
  - SenticNet (1, 2, 3, 4, 5)[3]

May 30, 1739



Vaucanson's Duck

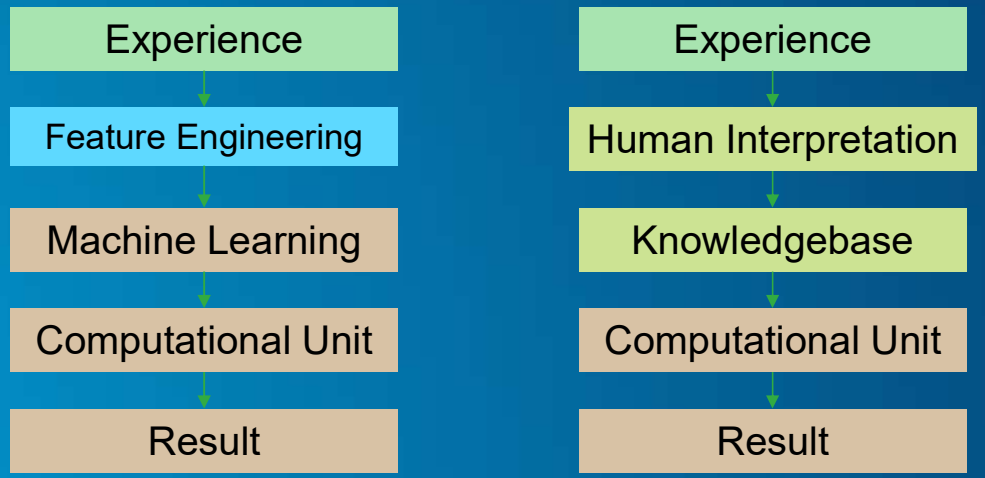


1. McCarthy, J. (1960). Programs with common sense (pp. 300-307). RLE and MIT computation center.
2. Lenat, D. B., Guha, R. V., Pittman, K., Pratt, D., & Shepherd, M. (1990). Cyc: toward programs with common sense. Communications of the ACM, 33(8), 30-49.
3. Cambria, E., Speer, R., Havasi, C., & Hussain, A. (2010, November). Senticnet: A publicly available semantic resource for opinion mining. In AAAI

Artificial Intelligence: Machine Learning

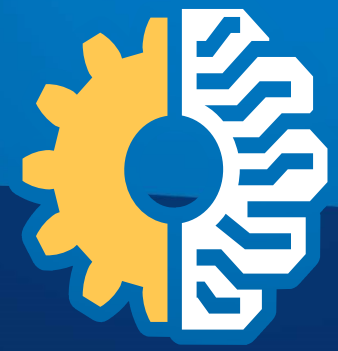
- Learn rules and patterns from data
- Data is represented explicitly, knowledge is NOT
  - Data-driven
- Explicit rules do NOT exist, instead inferred from data

"A computer program is said to learn from **experience E** with respect to some **class of tasks T** and **performance measure P**, if its performance … improves with experience E." [1]



1. Mitchell, T. M. (1997). Machine learning. 1997. Burr Ridge, IL: McGraw Hill, 45(37), 870-877.

Machine Learning in ArcGIS



Machine Learning in ArcGIS

Data-driven algorithms and techniques that automate tasks such as **prediction, classification** and **clustering**

Traditional Machine Learning

- Useful to solve a wide range of spatial problems
- Geography often acts as the 'key' for disparate data

Spatial Machine Learning

- Incorporate geography in their computation
- Shape, density, contiguity, spatial distribution, or proximity

Machine Learning Tools in ArcGIS

Classification

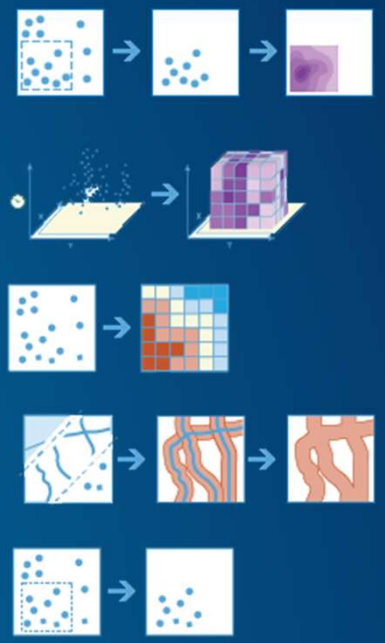
- Maximum Likelihood Classification
- Random Trees
- Support Vector Machine

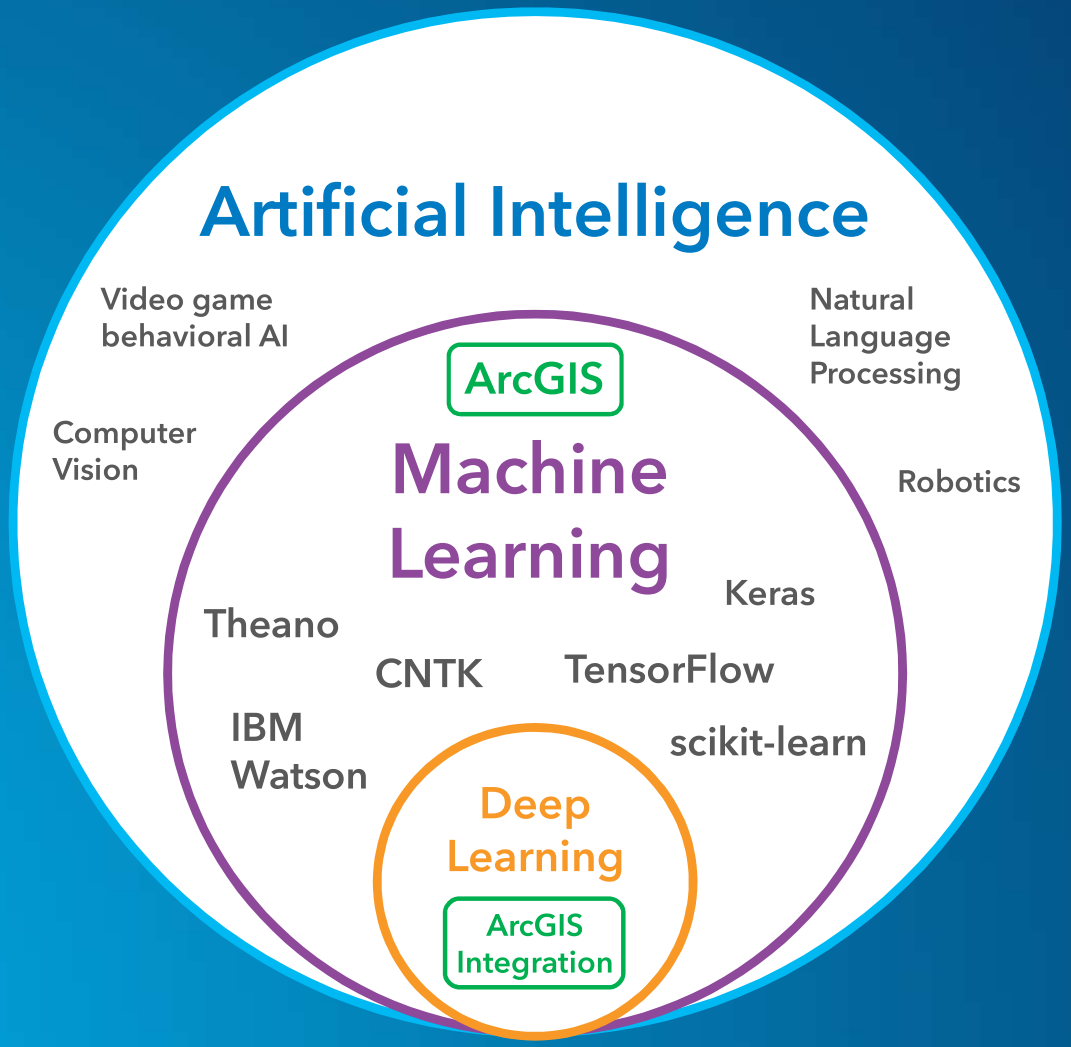
Clustering

- Spatially Constrained Multivariate Clustering
- Multivariate Clustering
- Density-based Clustering
- Image Segmentation
- Hot Spot Analysis
- Cluster and Outlier Analysis
- Space Time Pattern Mining
- Time Series Clustering

Prediction

- Empirical Bayesian Kriging
- Areal Interpolation
- EBK Regression Prediction
- Ordinary Least Squares Regression and Exploratory Regression
- Geographically Weighted Regression
- Forest Based Prediction




Artificial Intelligence
Video game behavioral AI
Natural Language Processing
Computer Vision
Robotics
ArcGIS
Machine Learning
Keras
Theano
CNTK
TensorFlow
IBM Watson
scikit-learn
Deep Learning
ArcGIS Integration

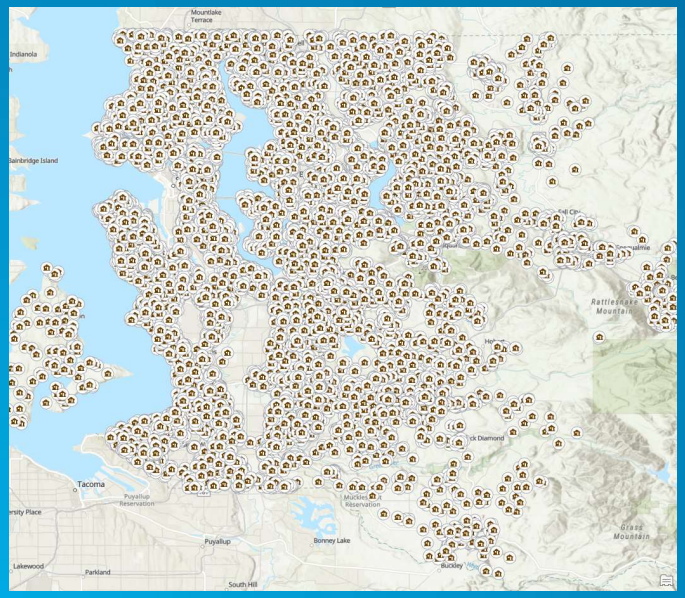
Solving Spatial Problems

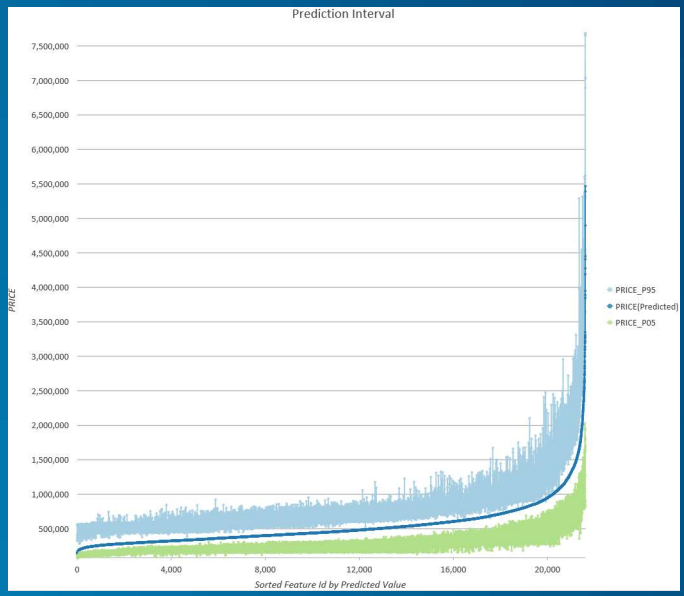
ArcGIS the Scientific Workbench

Prediction

Exploring Relationships Between Variables via Regression

Use Case: Understanding house price drivers in King County, WA from a rich dataset on house condition.



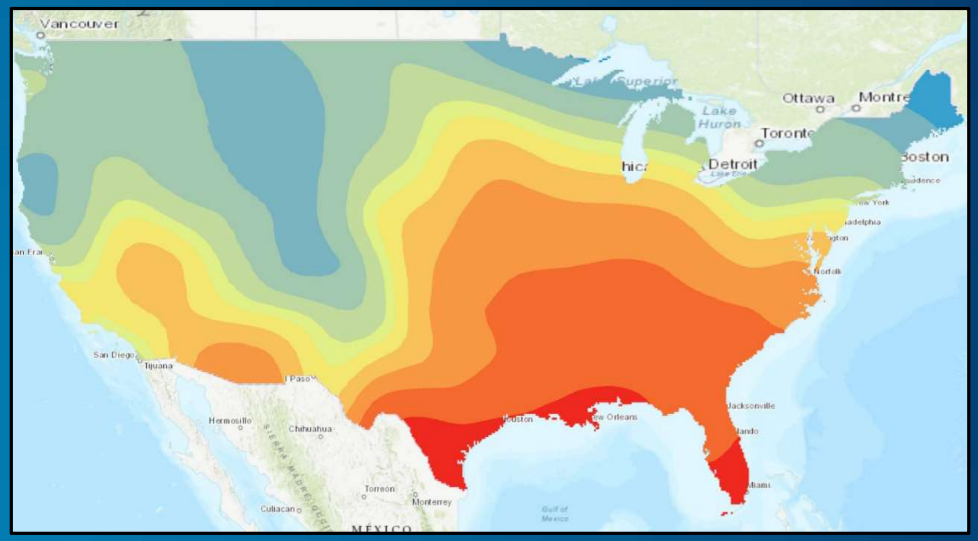


In ArcGIS: Forest Based Classification and Regression, Geographically Weighted Regression, Ordinary Least Squares Regression

Prediction

Using the known to estimate the unknown

Use Case: Accurately predict impacts of climate change on local temperature using global climate model data

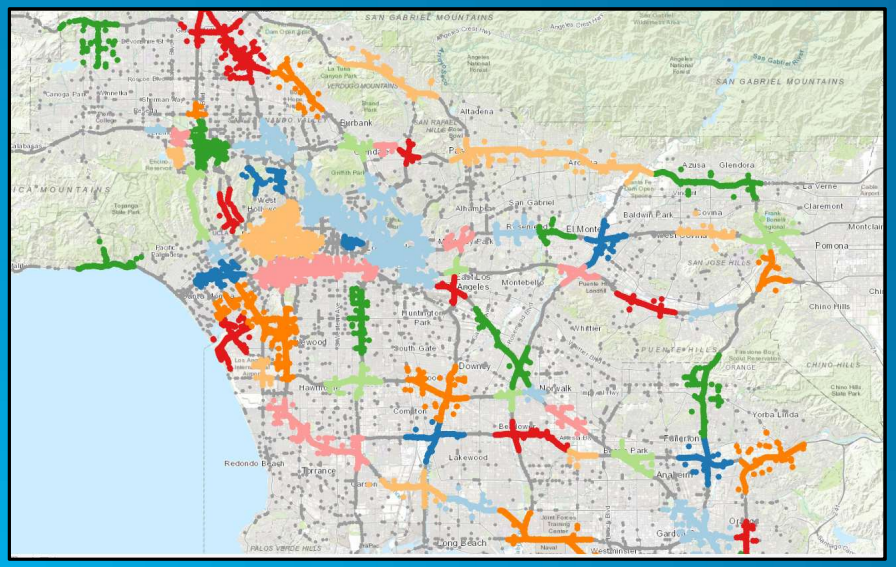


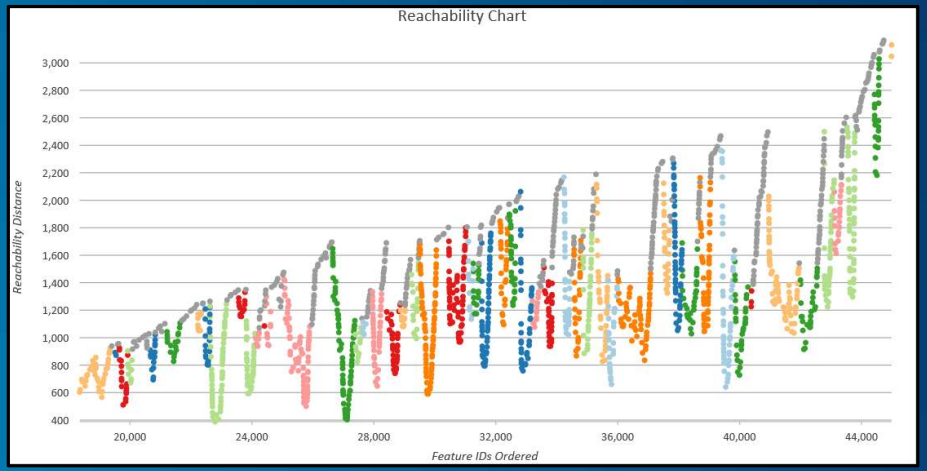
In ArcGIS: Empirical Bayesian Kriging, Areal Interpolation, EBK Regression Prediction, Ordinary Least Squares Regression and Exploratory Regression, Geographically Weighted Regression

Clustering

The grouping of observations based on similarities of values or locations

Use Case: Given the nearly 50,000 reports of traffic between 5pm and 6pm in Los Angeles (from Traffic Alerts by Waze), where are traffic zones that can be used to elicit feedback from current drivers in the area?



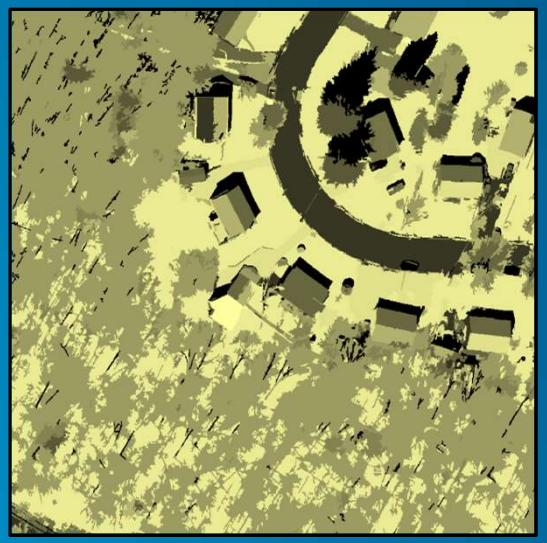


In ArcGIS: Spatially Constrained Multivariate Clustering, Multivariate Clustering, Density-based Clustering, Image Segmentation, Hot Spot Analysis, Cluster and Outlier Analysis, Space Time Pattern Mining

Classification

The process of deciding to which category an object should be assigned based on a training dataset

Use Case: Classify impervious surfaces to help effectively prepare for storm and flood events based on the latest high-resolution imagery



In ArcGIS: Maximum Likelihood Classification, Random Trees, Support Vector Machine

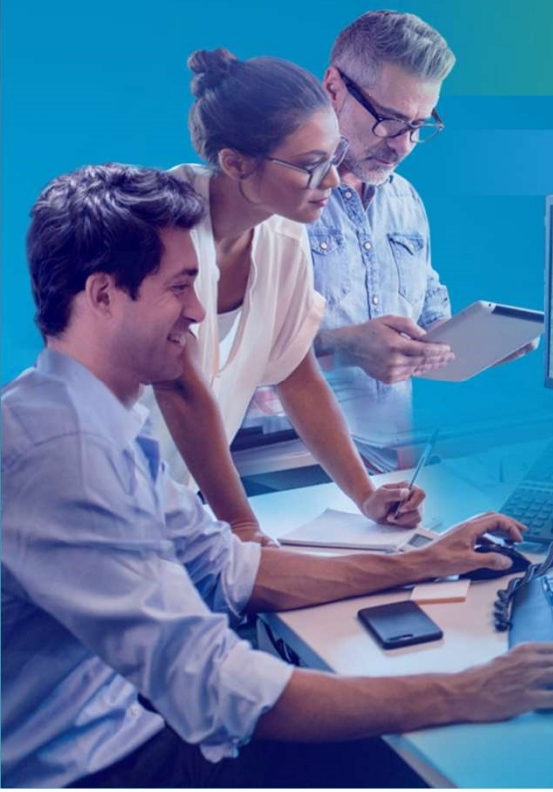
Interoperability with External Frameworks

Working in ArcGIS and Further

Open and Interoperable Platform for Science

Open Standards and Formats

XLSForm GML SQL SLD SOAP WMTS KML LAS INSPIRE Shapefiles
IMDF WCS IFC Web Scene (I3S) LERC CSW WPS REST OneGeology
WFS WMS OPeNDAP JSON WaterML netCDF GeoPackage CityGML



Direct Product Integration

. . . Many OGC Certifications

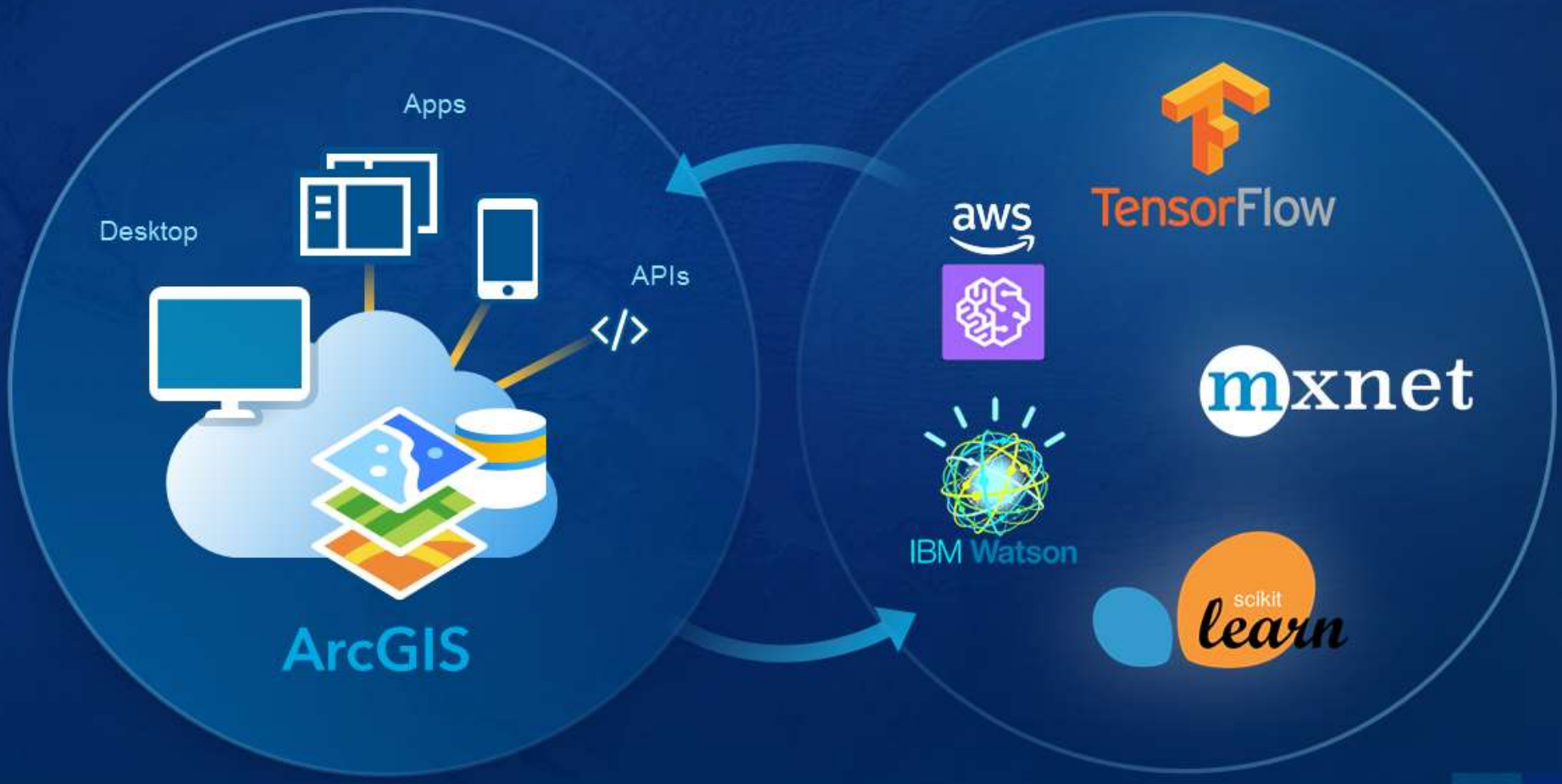
MS Office SQL Server SharePoint Azure Power BI Netezza SAP HANA
Adobe Creative Cloud Jupyter Notebook Teradata R AWS Altibase Python
Oracle AutoCAD Revit PostgreSQL Dameng SQLite Spark Hadoop

Open Software Architecture

Open Data Access Open APIs & SDKs Open-Source Integration
Open-Source Contributions Extensible Architecture Embeddable

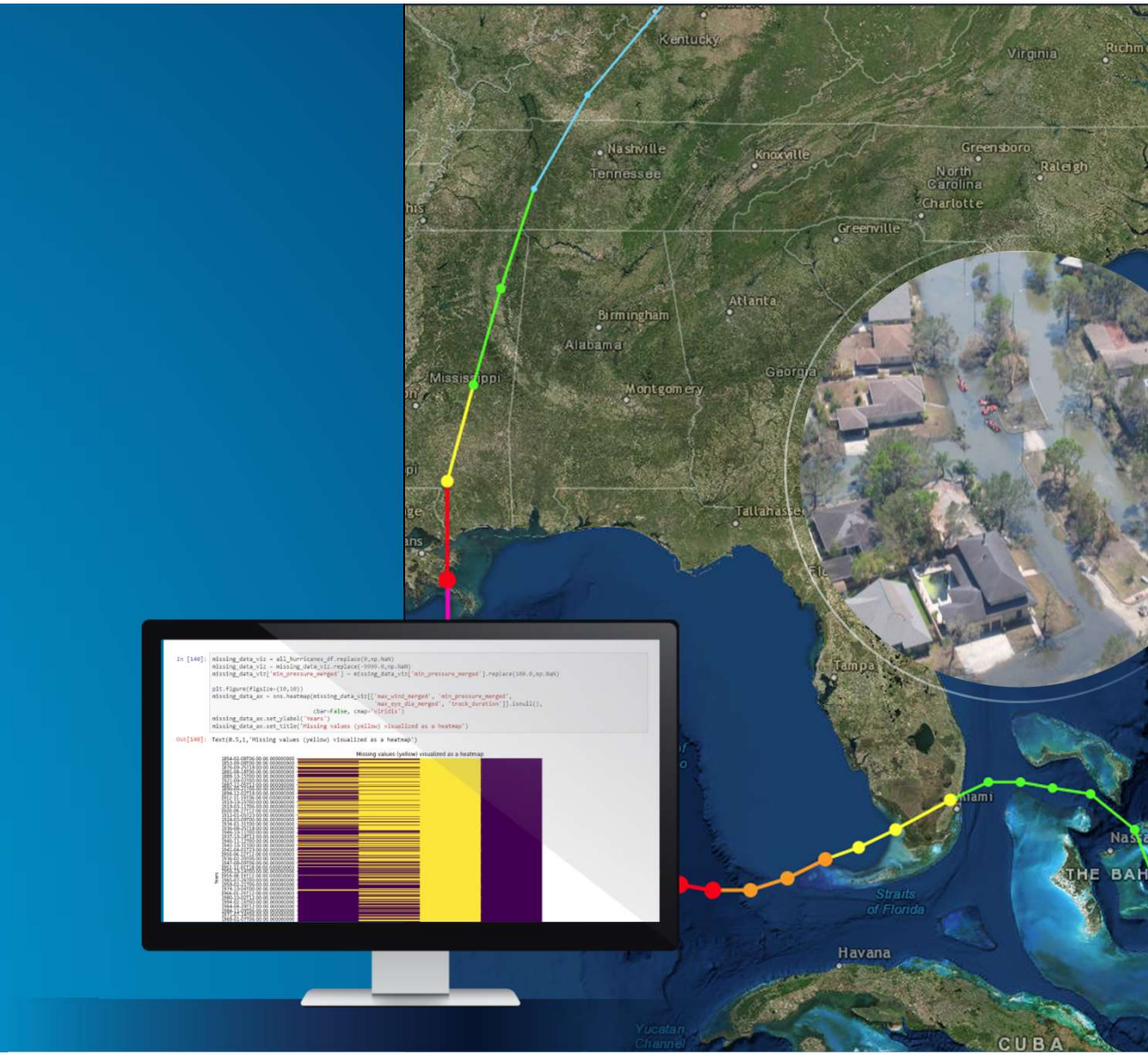
. . . Successfully Integrated Into Thousands of Systems

Machine Learning Integration with External Frameworks

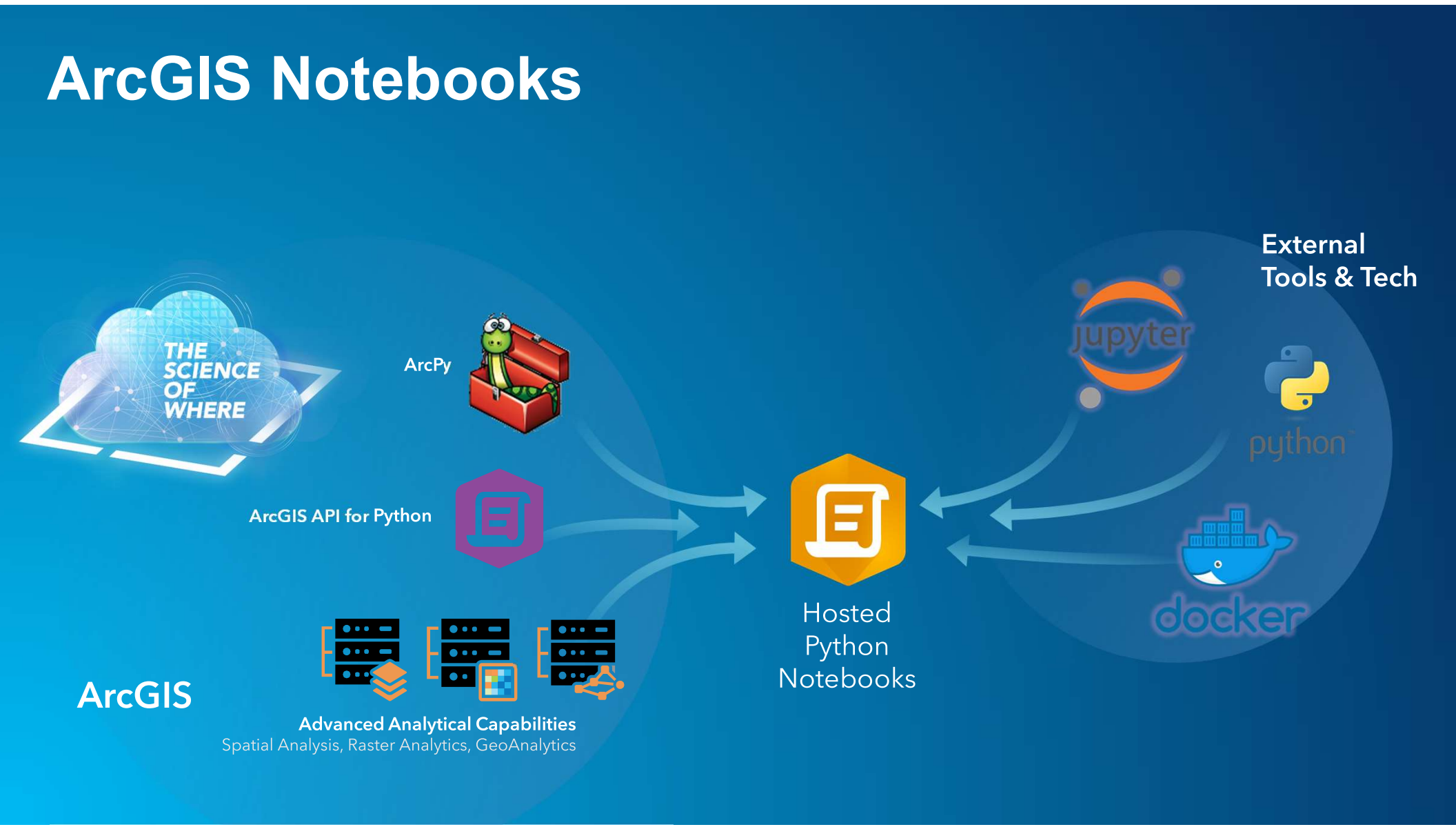


ArcGIS Notebooks

- Performing analysis
  - Spatial analysis
  - Spatial data science
    - Machine learning
    - Deep learning
- Automating analytic and administrative workflows related to GIS
- Operationalizing data science models



ArcGIS Notebooks

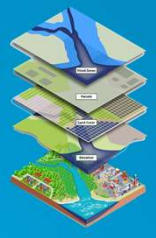


Deep Learning Integration

What is the #1 Challenge?

Getting Everyone to SEAMLESSLY work together

Analyst



Access Imagery, Fix Data, Prepare Training Data, Formulate Ask

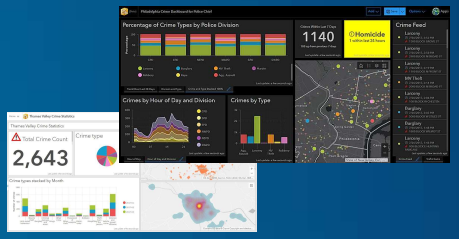
Consume Models for Analysis

AI Expert

Build and Optimize Models

Request Data

End User



Information Products

Analysis & Decision Making

Imagery AI Capabilities in ArcGIS

Prepare Training Data

Export Training Data for Deep Learning

Feature Layer + Imagery → Chips

Consume DL Models

Detect Objects & Classify Pixels

DL Model + Imagery → Detections

Train DL Models

Arcgis.Learn

Deploy Models on Portal

As dlpk items

Examples for other Imagery AI Workflows

Object Detection, Instance Segmentation, Land Cover, Change Detection..

Damaged Structures

Building Footprints

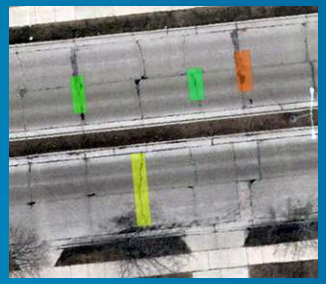
Land Cover

Pipeline Encroachment

Roads

Oil Pads

Road Cracks

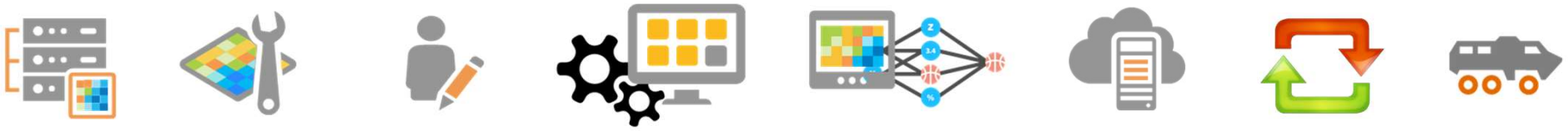


Cars

Palm Trees

Swimming Pools

End to End GeoAI Life Cycle with Imagery



Imagery Access

Imagery Prep

Data Labelling

Deploy Models to Production

Train & Consume Models

Run Inference at SCALE

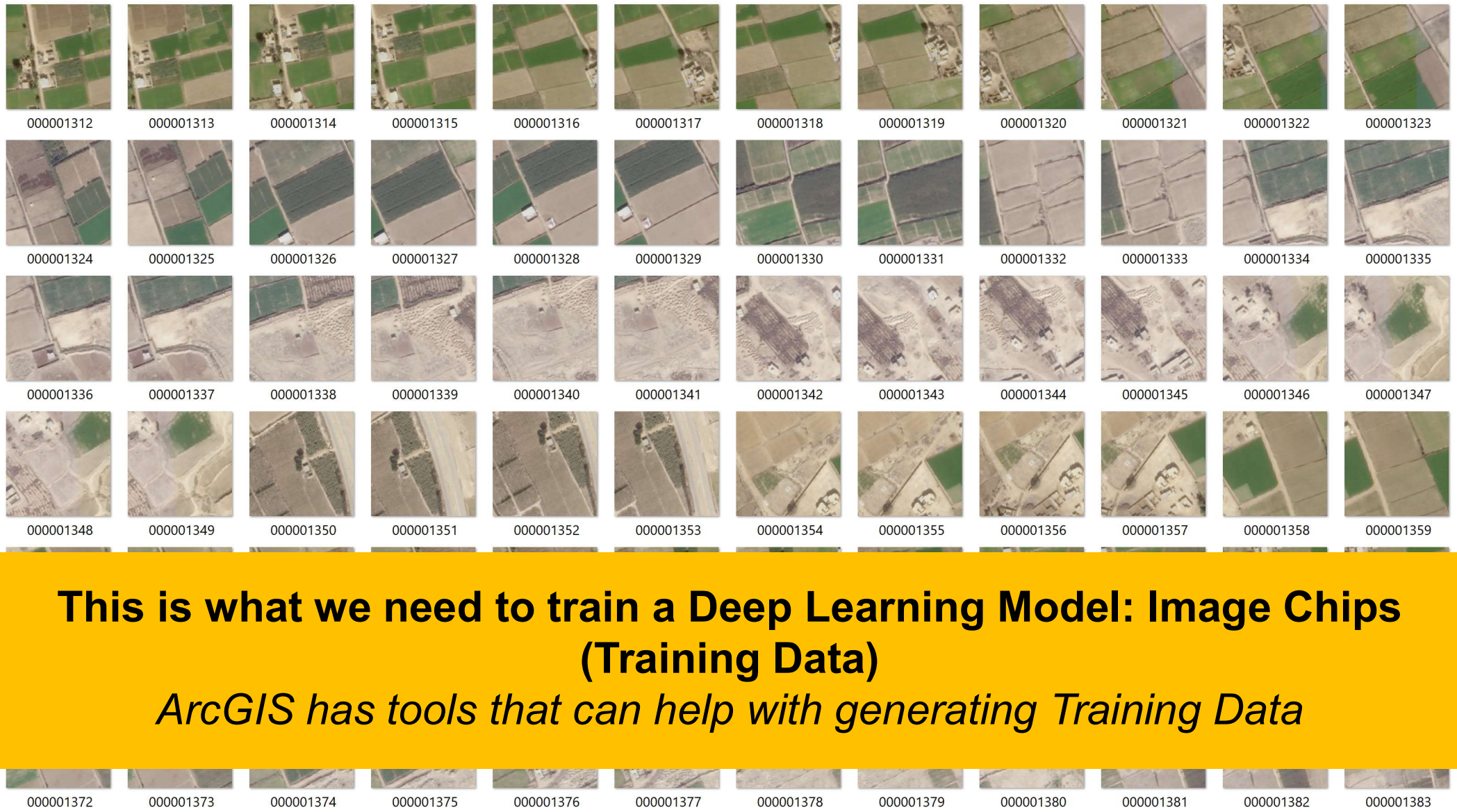
Feedback Loop

Take Action

1. Imagery Access & Preparation

Imagery Processing & Building Image Mosaic

2. Prepare Training Data

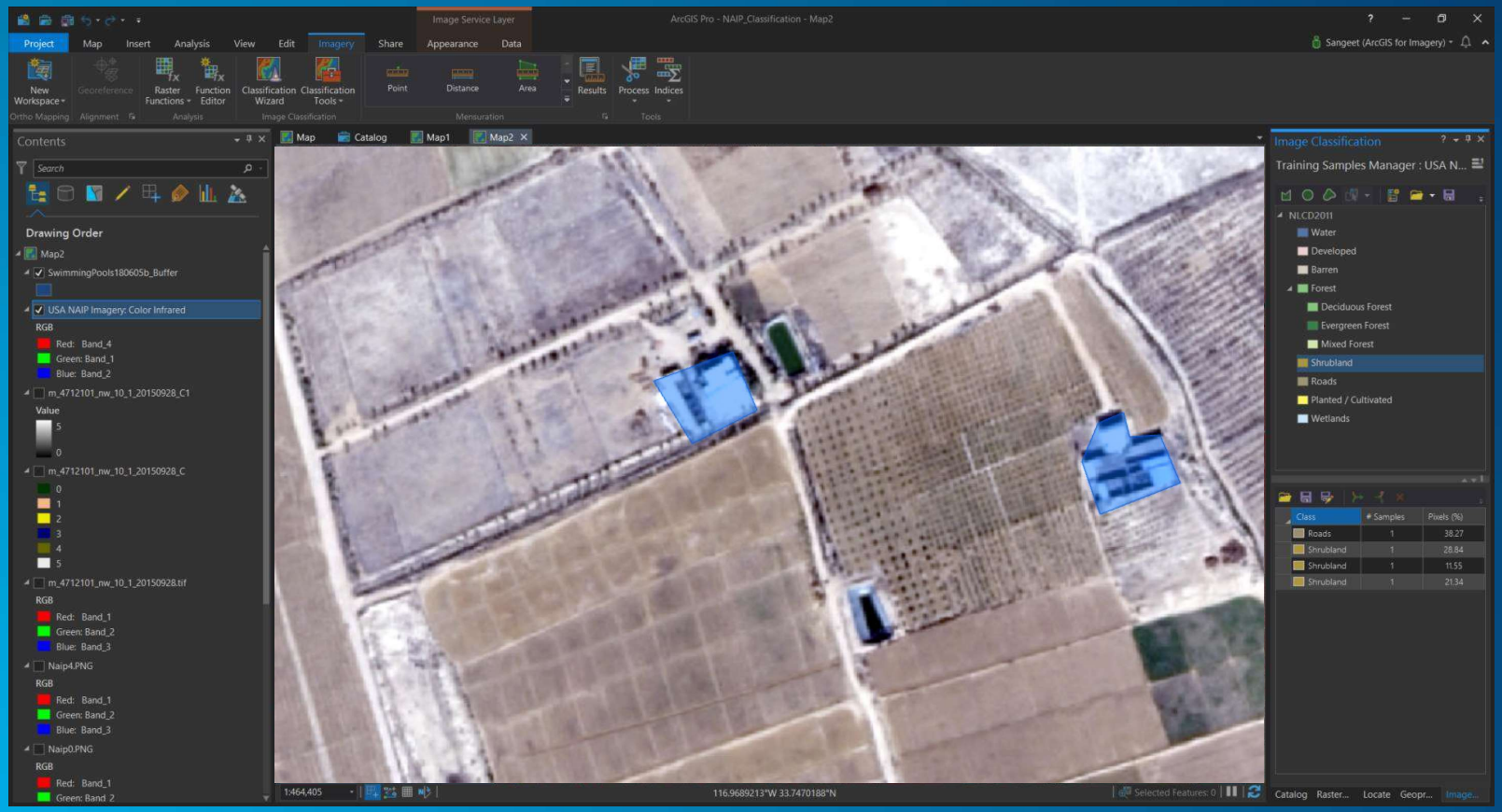


This is what we need to train a Deep Learning Model: Image Chips (Training Data)

ArcGIS has tools that can help with generating Training Data

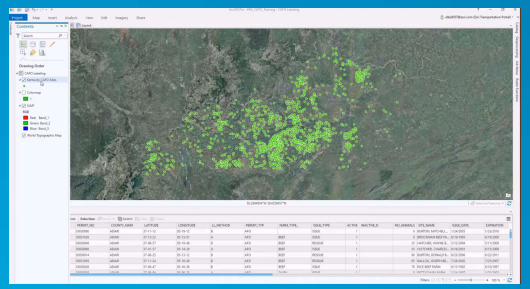
Training Sample Manager

For easily labelling objects of interest

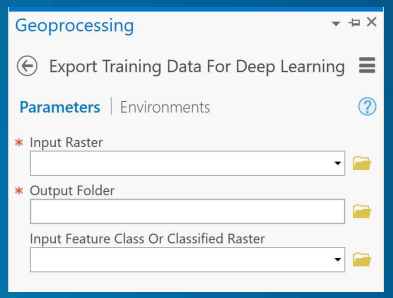


Export Training Data for Deep Learning

GIS Feature Class



Imagery Data



Bounding Boxes

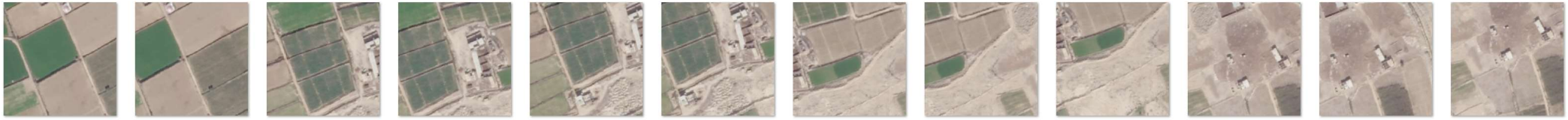

000001312 000001313 000001314 000001315 000001316 000001317 000001318 000001319 000001320 000001321 000001322 000001323


000001324 000001325 000001326 000001327 000001328 000001329 000001330 000001331 000001332 000001333 000001334 000001335


000001336 000001337 000001338 000001339 000001340 000001341 000001342 000001343 000001344 000001345 000001346 000001347


000001348 000001349 000001350 000001351 000001352 000001353 000001354 000001355 000001356 000001357 000001358 000001359


000001360 000001361 000001362 000001363 000001364 000001365 000001366 000001367 000001368 000001369 000001370 000001371


000001372 000001373 000001374 000001375 000001376 000001377 000001378 000001379 000001380 000001381 000001382 000001383

3. Train Model

Model Training using ArcGIS.Learn

Train SingleShotDetector Model

```
from arcgis.learn import SingleShotDetector

ssd = SingleShotDetector(data, grids=[9], zooms=[1.0], ratios=[[
```

The **arcgis.learn** module in ArcGIS API for Python enable GIS analysts and data scientists to easily adopt and apply deep learning in their workflows. It enables training state-of-the-art deep learning models with a simple, intuitive API.

```
In [8]: ssd.fit(10, lr=slice(1e-3, 1e-2))
```

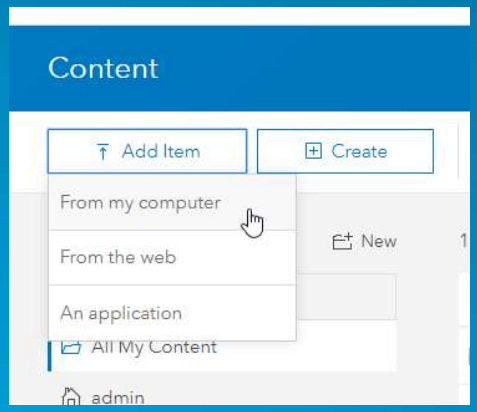
Total time: 15:56

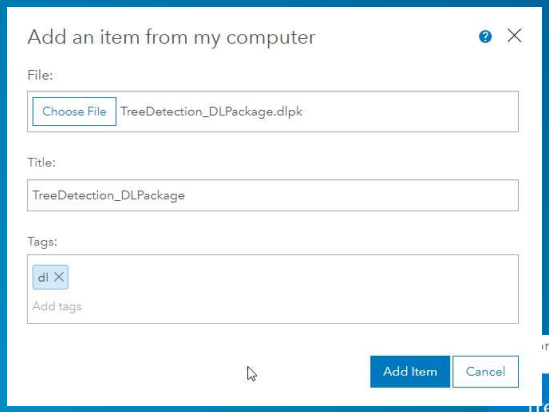
epoch	train_loss	valid_loss
1	629.015869	250.982254
2	400.904327	181.745972
3	315.588318	163.946136
4	268.519928	155.258881
5	234.541077	133.495728
6	209.463257	116.552231
7	189.608063	104.452789
8	172.239929	98.530197
9	157.103226	91.969261
10	146.046310	91.620415

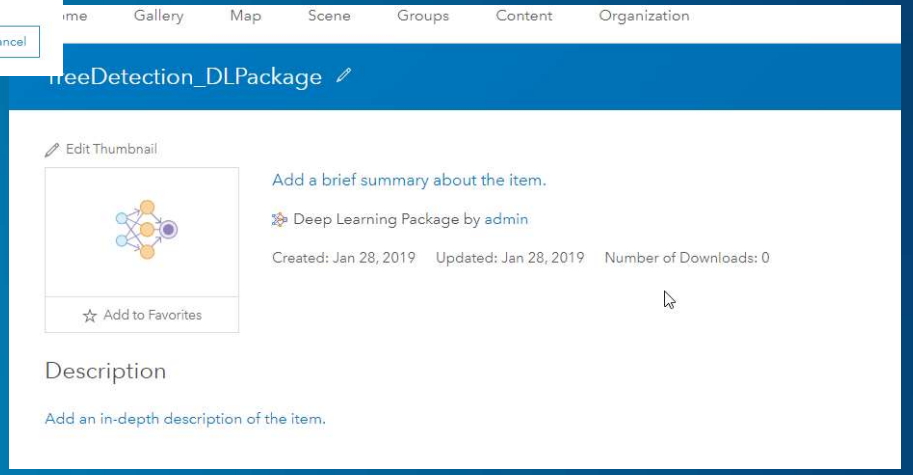
4. Add Model to Portal

Upload Model to Portal as DLPK item

So it's accessible to different people in the organization







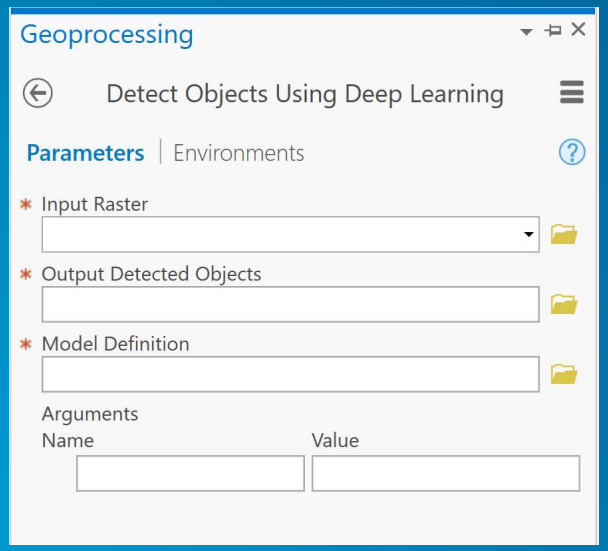
5. Consume Model

Detect Objects Using Deep Learning

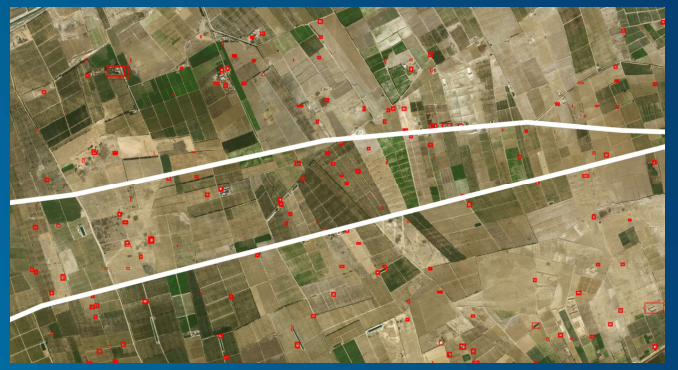
Imagery

Model
(emd file)

Detected Objects GP Tool

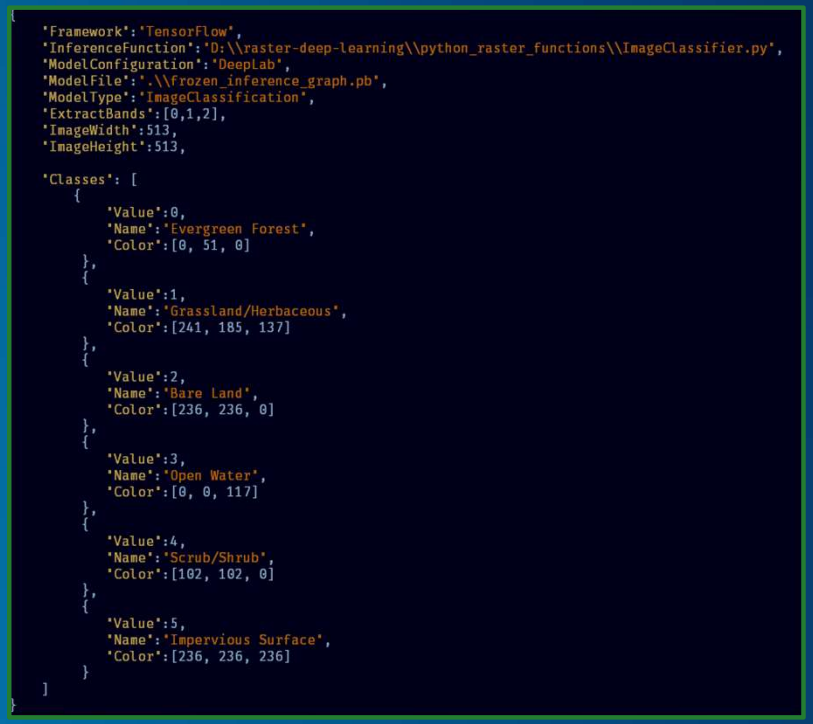


Detected Objects



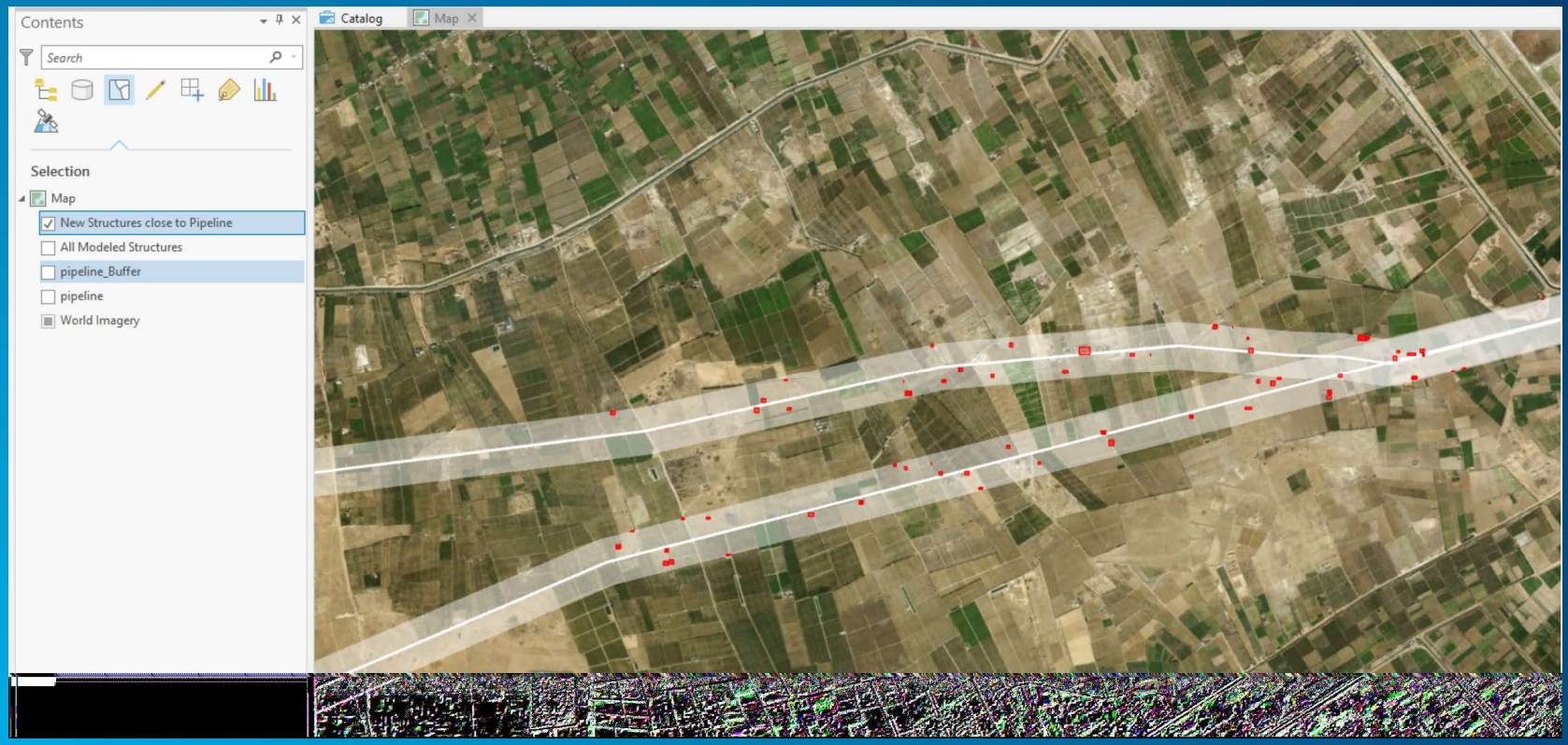
Model Definition File

- .emd extension with json formatted content
- describes the deep learning model to ArcGIS
  - deep learning framework
  - model
  - model type
  - model configuration
  - model inputs
  - model outputs
  - classes
  - if not using out-of-the-box model
    - optional Python Raster Function path
    - optional model parameters
- Image Scientist must understand and edit



6. Analysis

Selecting only NEW Buildings within the Pipelines' Buffer

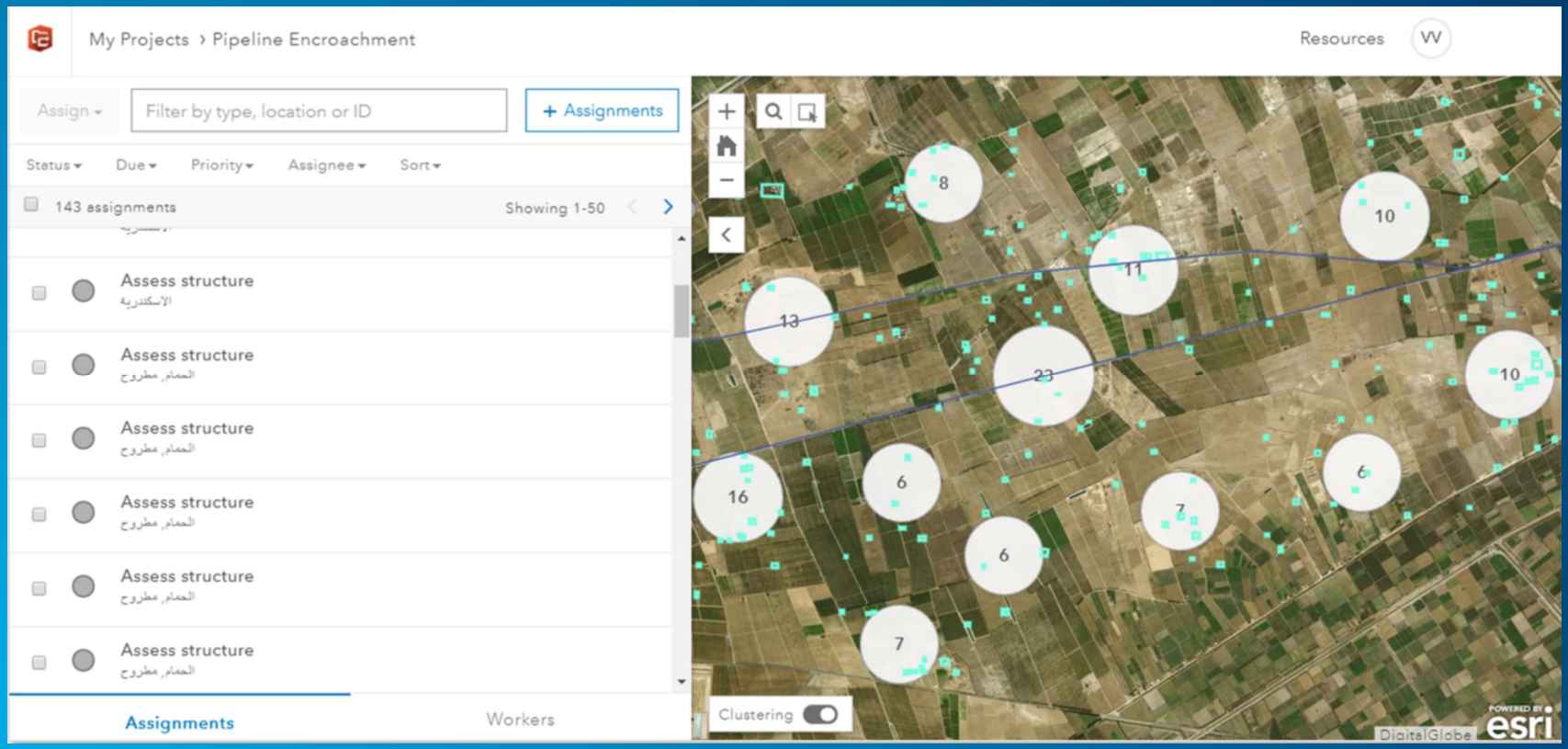


7. Feedback Loop

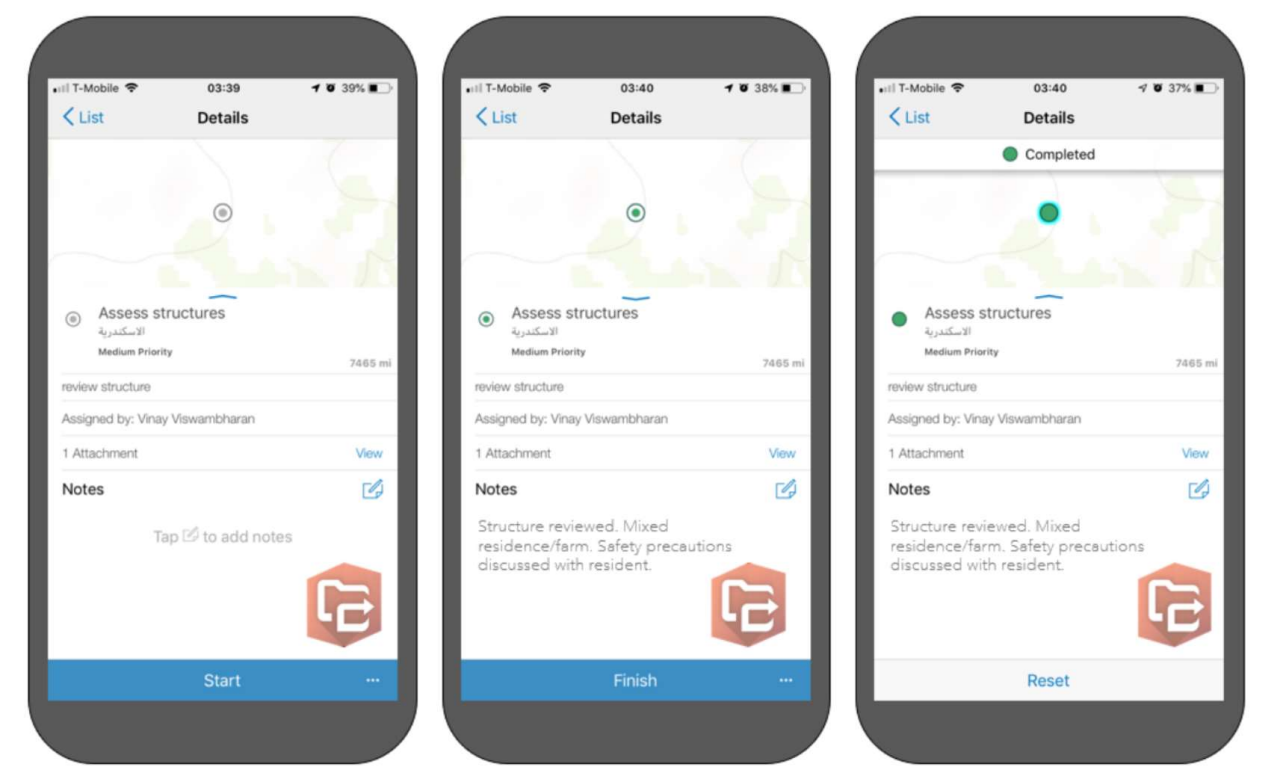
Image Visit App could be used as a QC tool to check model's output

8. Action: Information Products

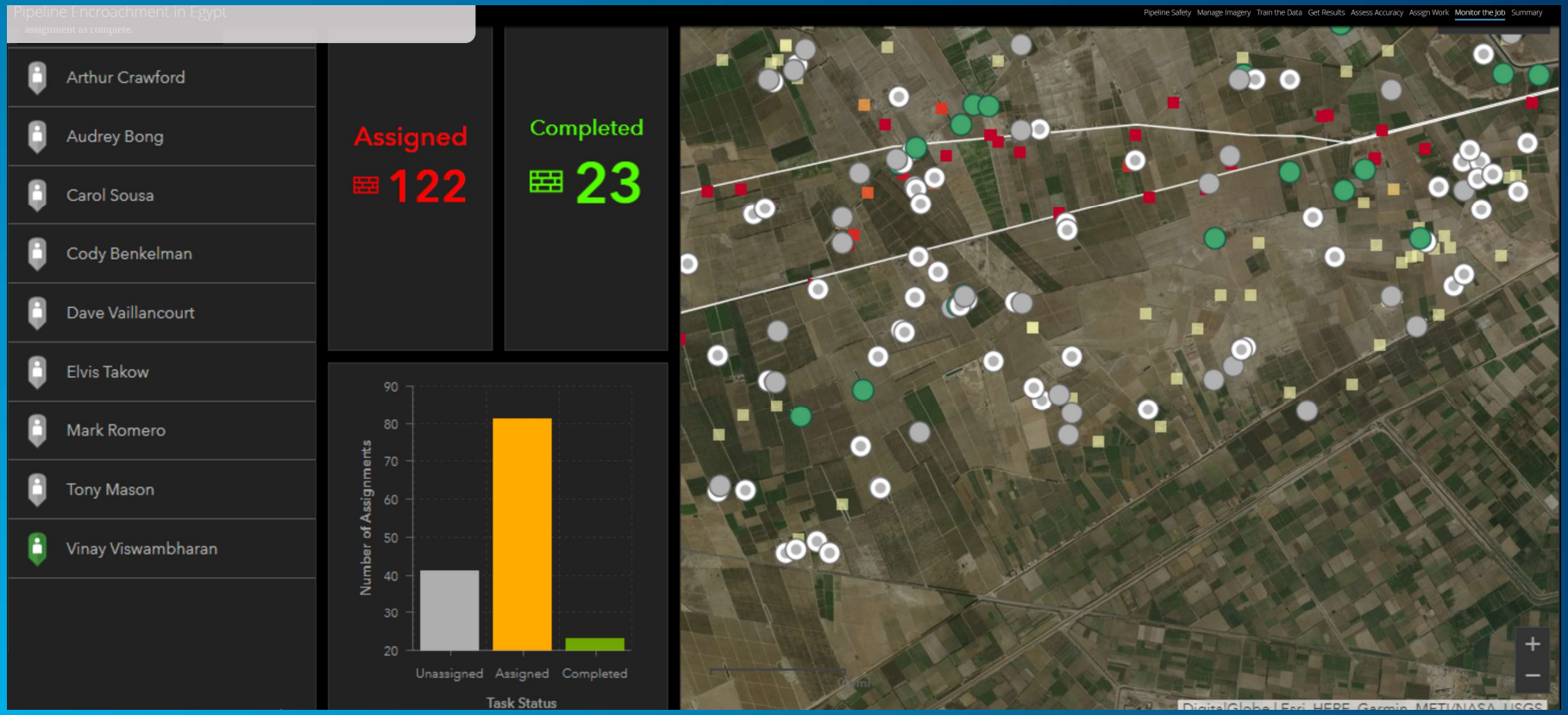
Workforce App to assign Inspection Tasks



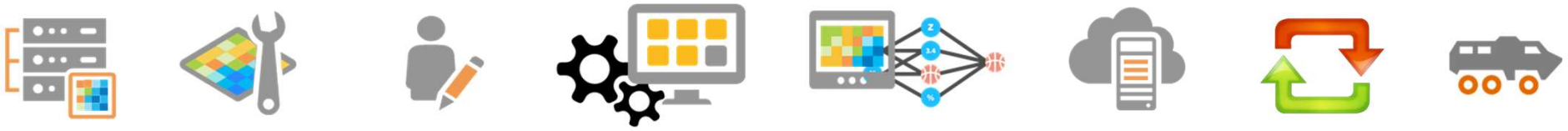
Field Apps for Inspection and Data Collection



Operations Dashboard to view Inspection Results in Real-Time



End to End GeoAI Life Cycle with Imagery



Imagery Access

Imagery Prep

Data Labelling

Deploy Models to Production

Train & Consume Models

Run Inference at SCALE

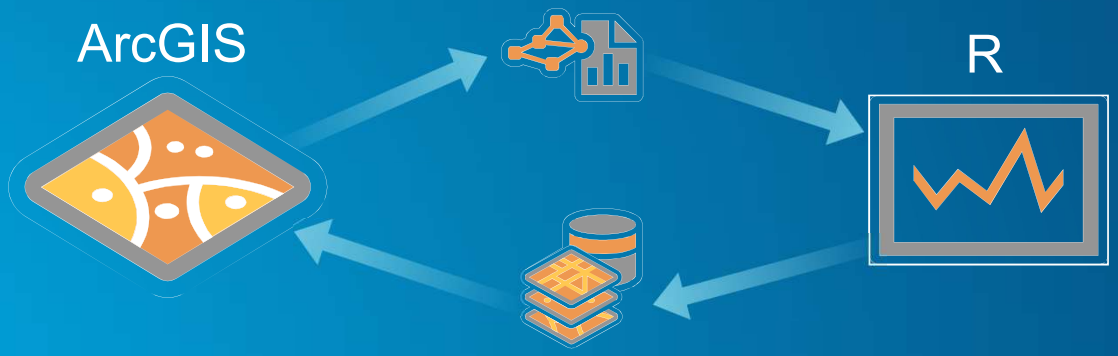
Feedback Loop

Take Action

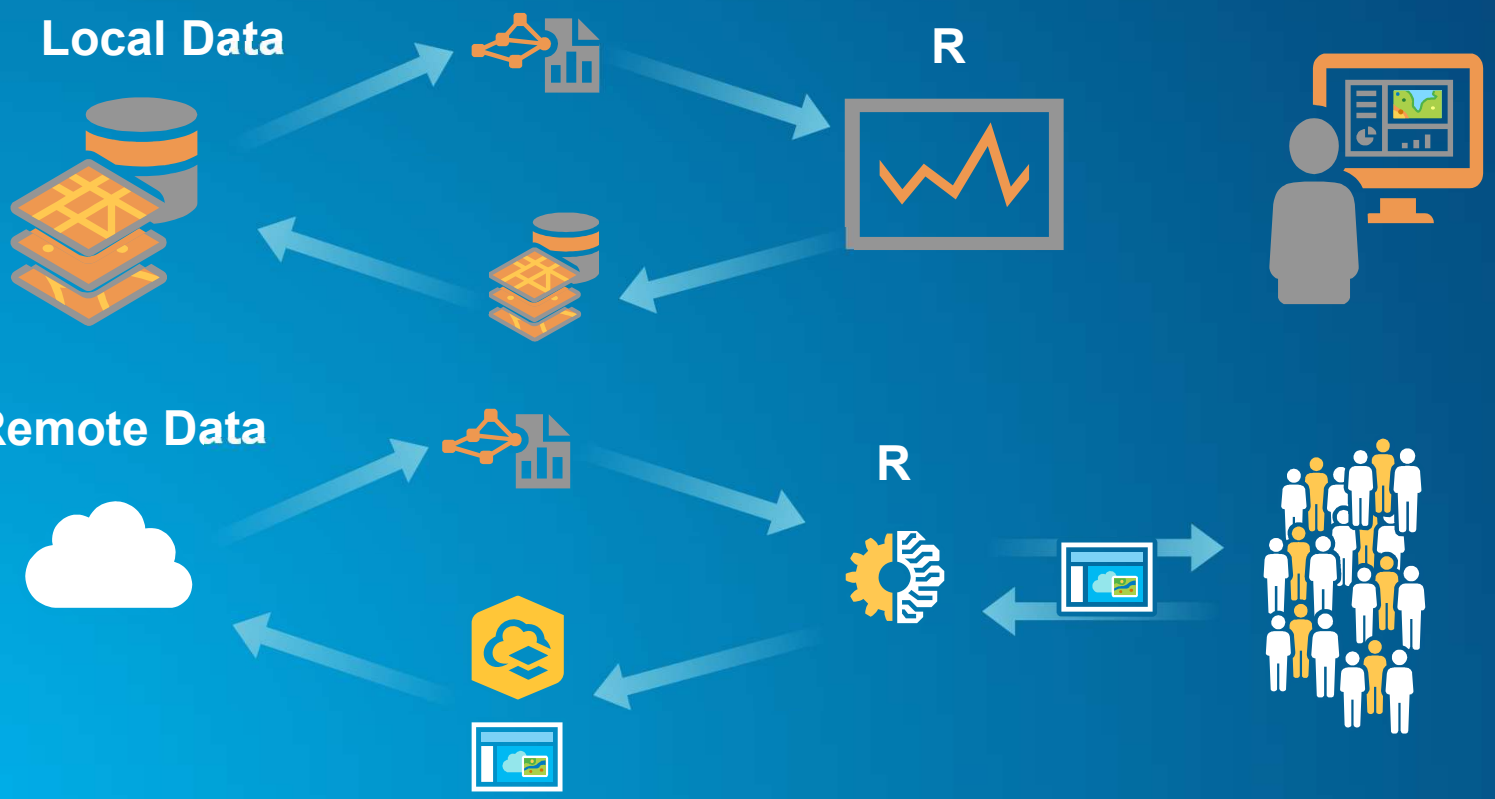
Introducing the R-ArcGIS Bridge

The R-ArcGIS Bridge

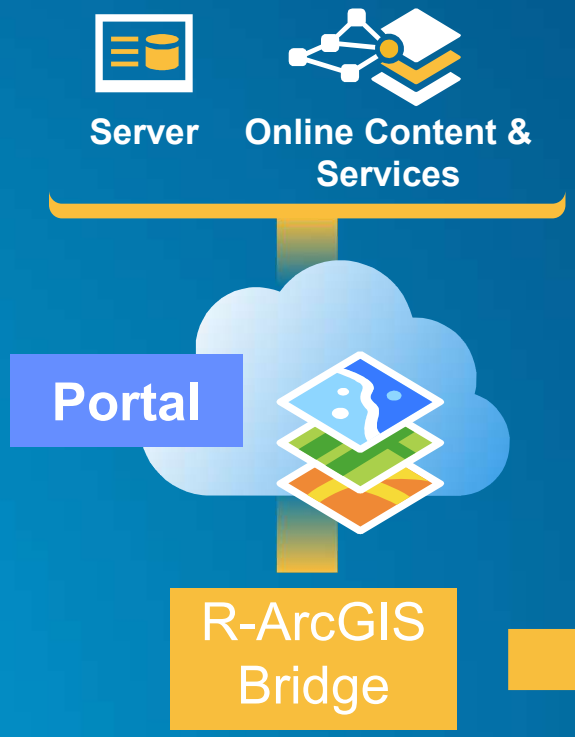
- The R-ArcGIS bridge allows you to connect ArcGIS to R and enables the seamless transfer of data back and forth, along with the ability to integrate R and ArcGIS functionality.



Patterns of Working with R



Remote Data with R-ArcGIS Bridge



Spatial R Data frames
R Raster Types
R Mosaics
…

Living Data – Living Atlas

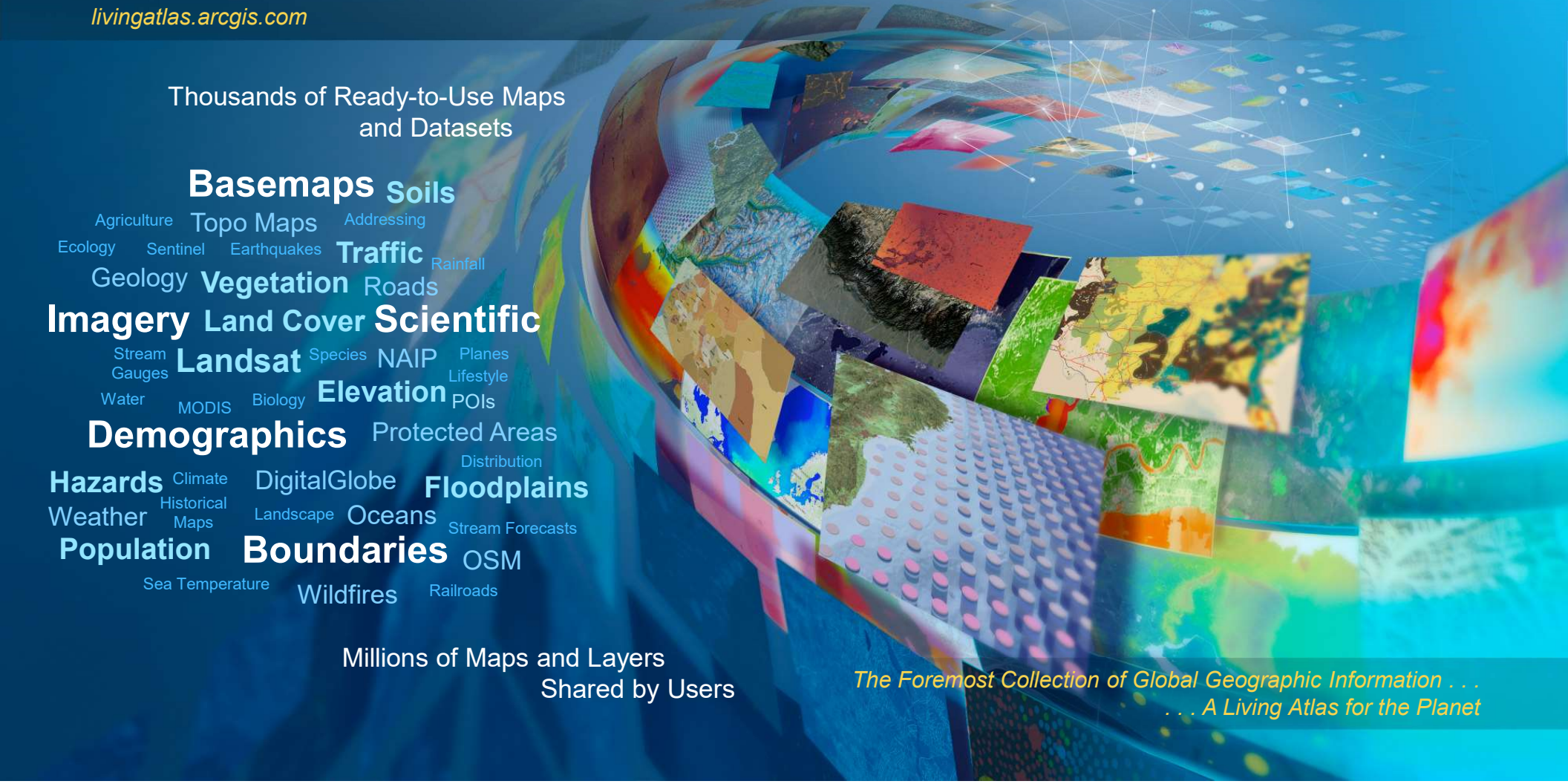
Spatial Memory of Our Planet

Living Atlas of the World | Open Data, Maps, and Apps from 1000s of Contributers + Esri

livingatlas.arcgis.com

Thousands of Ready-to-Use Maps and Datasets

Basemaps **Soils**
Agriculture Topo Maps Addressing
Ecology Sentinel Earthquakes **Traffic** Rainfall
Geology **Vegetation** Roads
Imagery **Land Cover** **Scientific**
Stream Gauges **Landsat** Species NAIP Planes Lifestyle
Water MODIS Biology **Elevation** POIs
Demographics Protected Areas
Distribution
Hazards Climate DigitalGlobe **Floodplains**
Weather Historical Maps Landscape Oceans Stream Forecasts
Population **Boundaries** OSM
Sea Temperature Wildfires Railroads

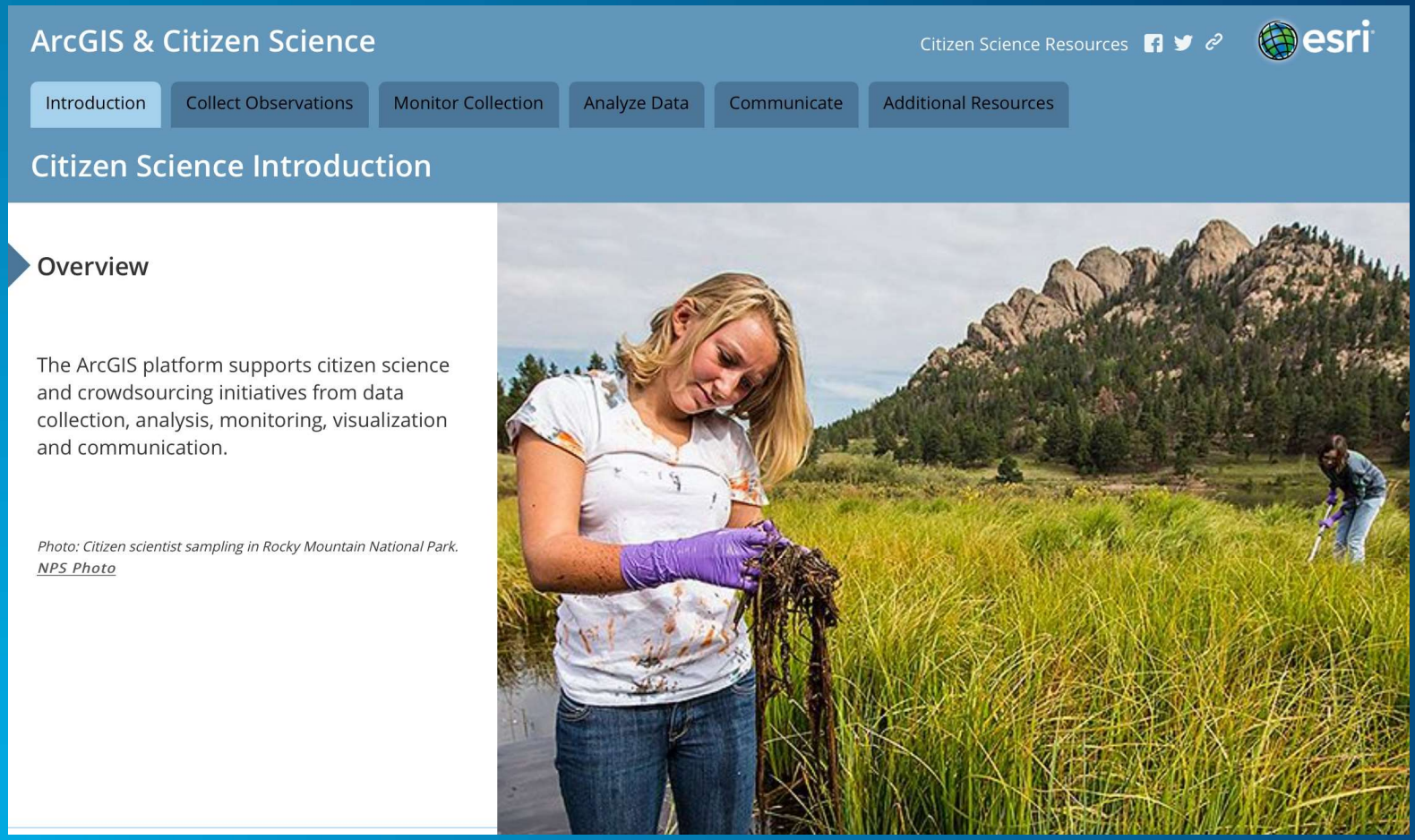


Millions of Maps and Layers Shared by Users

The Foremost Collection of Global Geographic Information . . .
. . . A Living Atlas for the Planet

Citizen Science

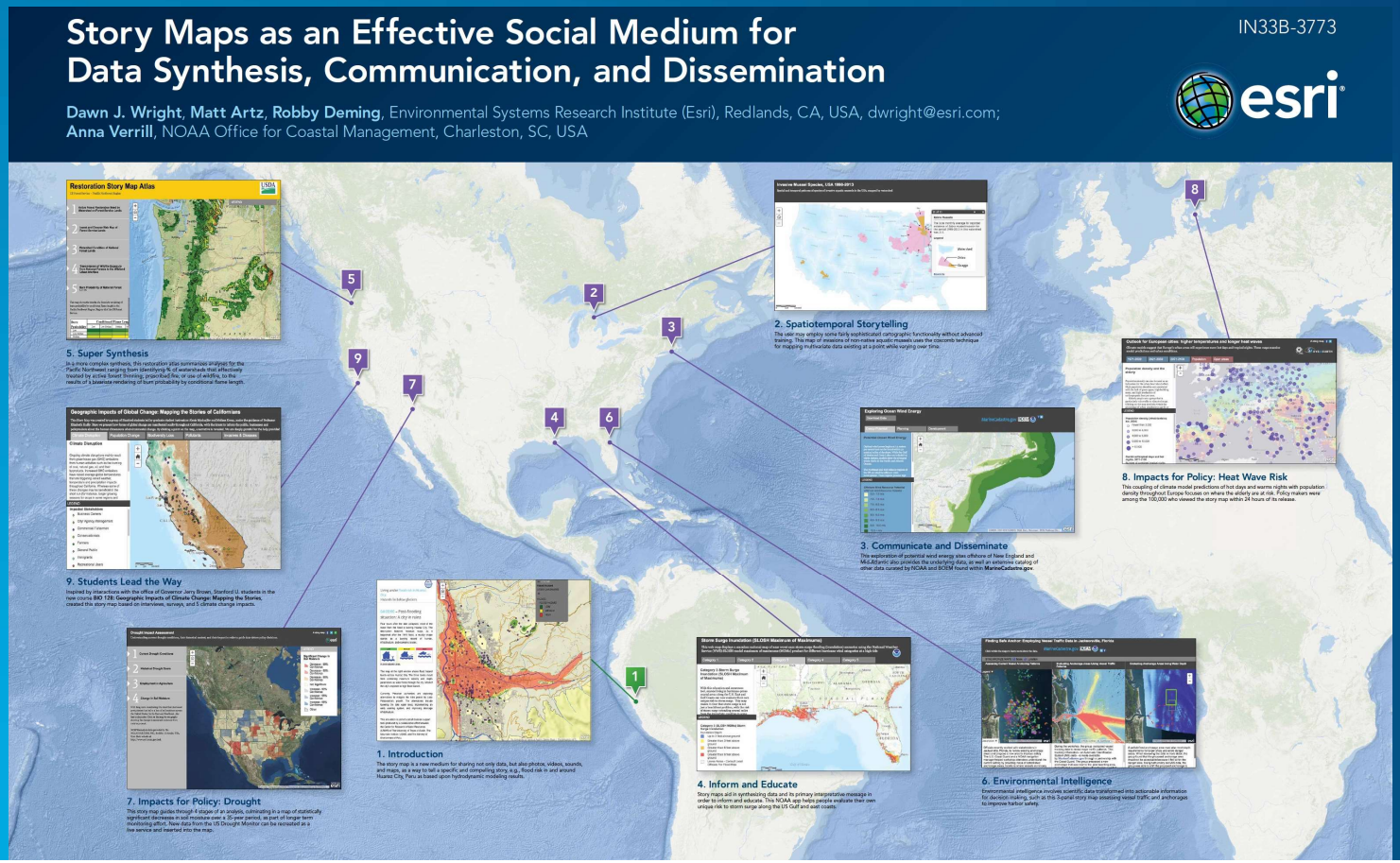
esriurl.com/citizens



Spatial Storytelling

The Power of Storytelling

Science Communication is a growing movement among researchers, educators



esriurl.com/agustories

Get Involved

AI for Earth

Microsoft

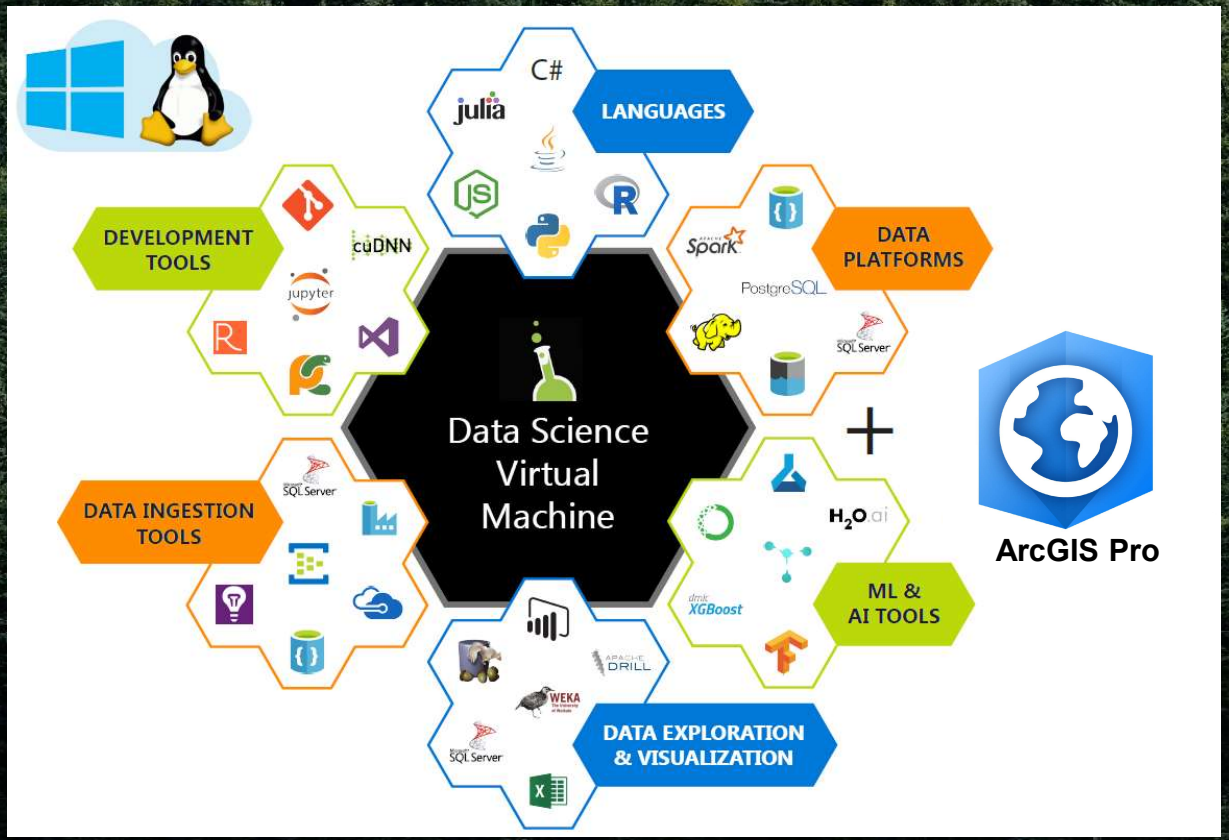
esri

Learn more at aka.ms/aiforearth

GeoAI Virtual Machine

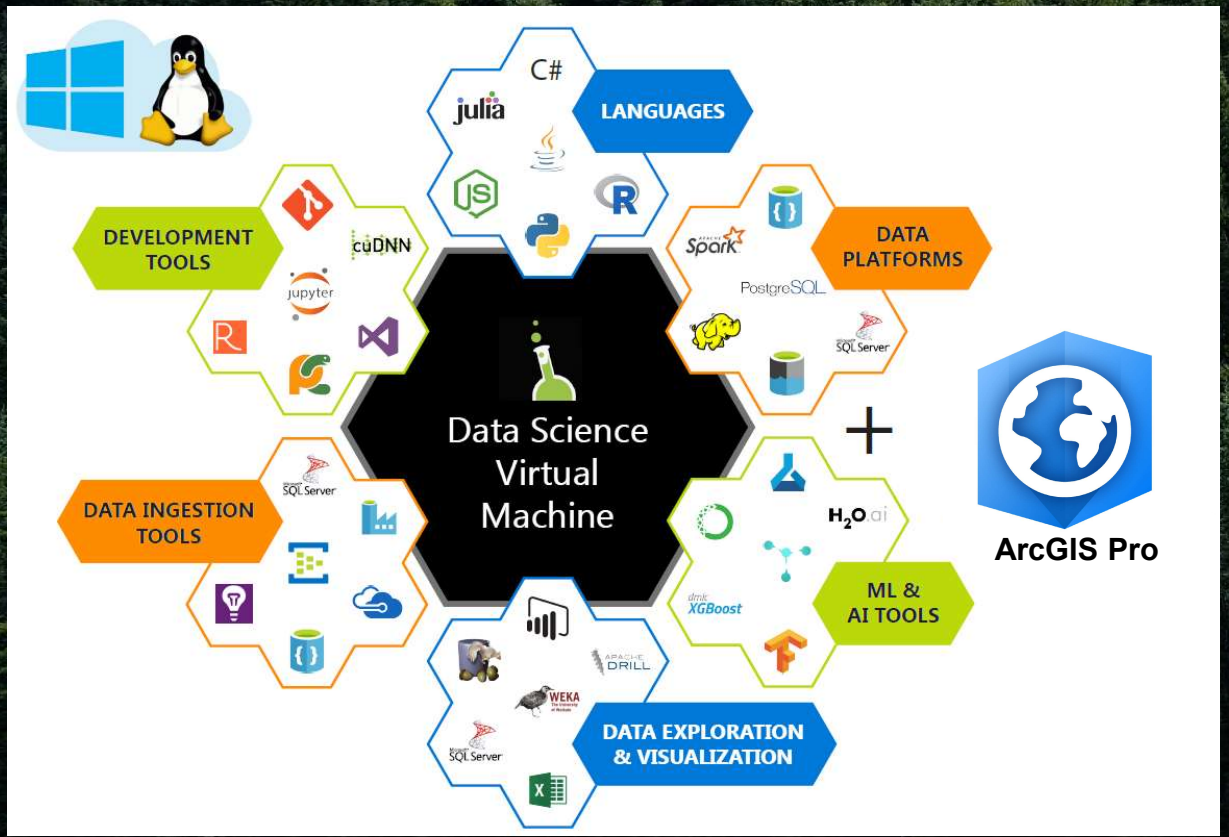
Microsoft

esri



Learn more at esriurl.com/geoai2018

Esri Applied Data Science Initiative



- Oregon State University
- EPA AirNow
- Conservation Intl
- USDOT Natl Highway Traffic Safety Admin
- Virginia Commonwealth U
- UC-Riverside
- DOE Natl Energy Technology Lab