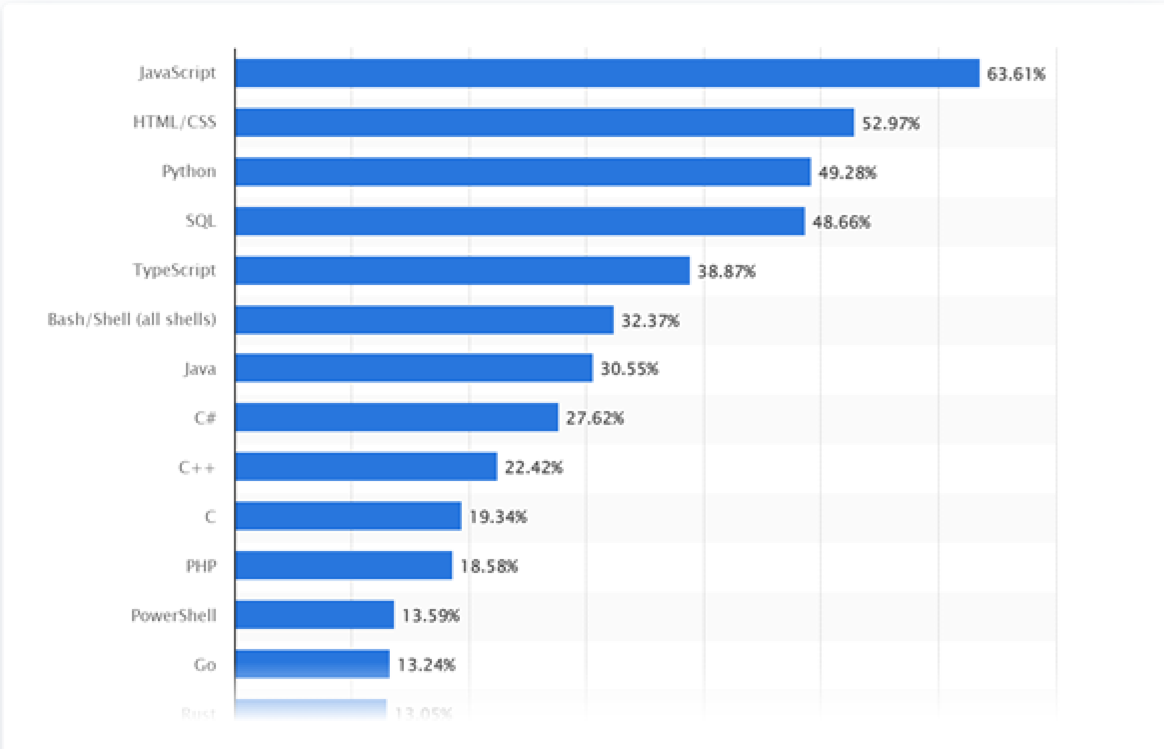
HydroSuite

WaterSoftHack ‘24

UIHI Lab

Programming Languages

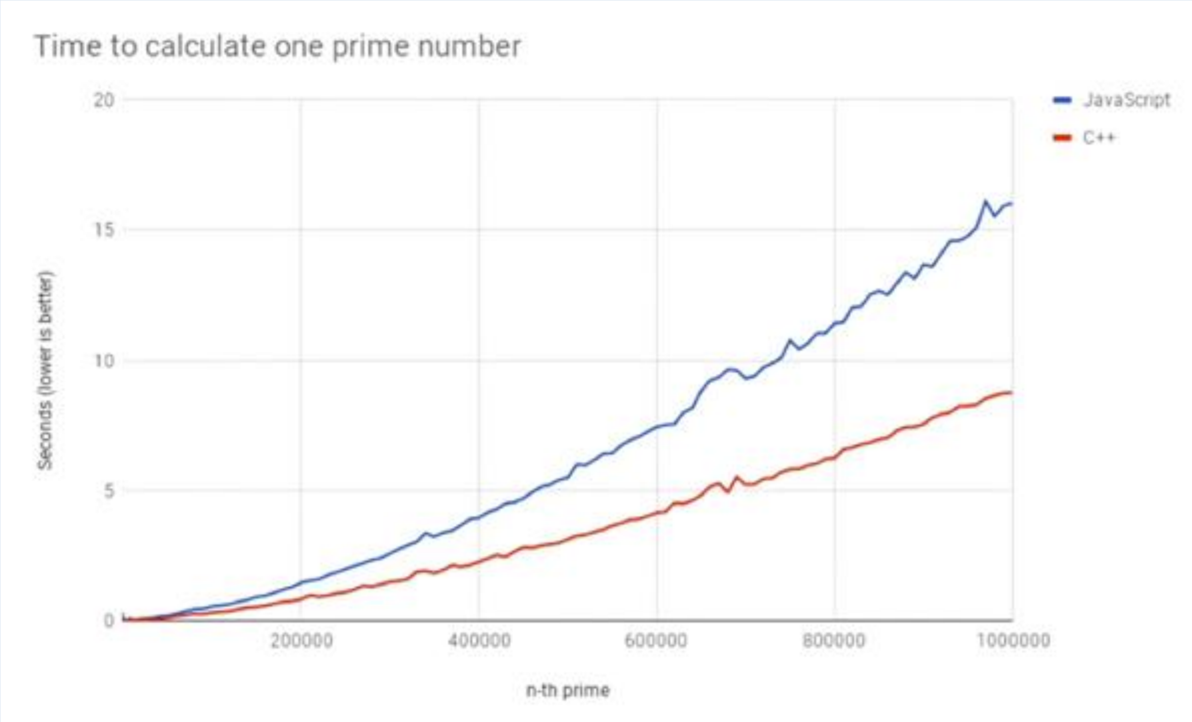
Most used programming languages among developers worldwide as of 2023



https://www.statista.com/statistics/793628/worldwide-developer-survey-most-used-languages

IOWA

Web Performance



Web Technologies

WebGL / WebGPU	Web Assembly	WebXR / WebVR	WebCL / Web Workers	Web SQL / Web Storage
• Graphics Library • GPU Acceleration	• Desktop Languages • Native Performance	• Augmented Reality • Virtual Reality	• Computing Library • Parallel Processing	• Large-scale Storage • Offline Applications

IOWA

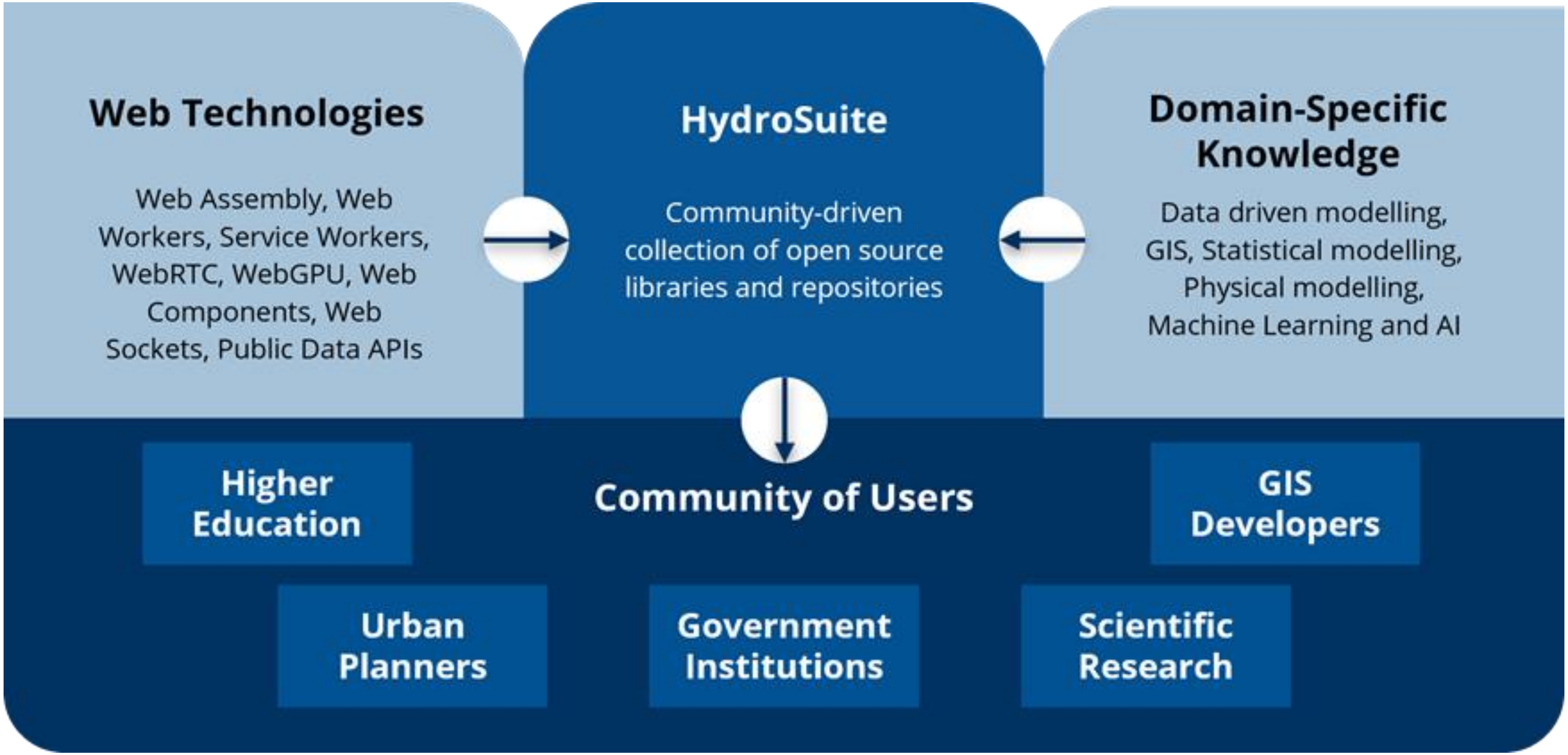
HydroSuite

UIHILab - https://hydroinformatics.uiowa.edu

HydroSuite – Collection of Open Source Tools, Libraries, Repositories			
DATA	**COMPUTING**	**COMMUNICATION**	**COMMUNITY PORTALS**
Flood-ML	HydroLang	RasterJS	EarthAI Hub
Flood Event DOM	HydroLang-ML	Hydro3DJS	HLM Web
IS Ontology	HydroLang-BMI	Instant Expert	HydroLSTM
WaterBench	HydroCompute	GeospatialVR	HydroLang Models
IowaRain	HydroRTC	Watershed Delineation	Training Repos

- ✓ Collection of web-based software for hydrology research and education
- ✓ Covering client-side and server-side technologies for complete hydrologic workflows

IOWA


Web Technologies
Web Assembly, Web Workers, Service Workers, WebRTC, WebGPU, Web Components, Web Sockets, Public Data APIs
HydroSuite
Community-driven collection of open source libraries and repositories
Domain-Specific Knowledge
Data driven modelling, GIS, Statistical modelling, Physical modelling, Machine Learning and AI
Community of Users
Higher Education
GIS Developers
Urban Planners
Government Institutions
Scientific Research

HydroLang

- ✓ A web-based open-source programming framework for hydrological and environmental analyses
- ✓ Entirely client-side operation, no server-side dependency
- ✓ Scalable and upgradable
- ✓ Easy to use and modify
- ✓ Community oriented modular design for expansion
  - ✓ Special GitHub repo
  - ✓ UI Hackathon

HydroLang
Web-based Hydrological Programming Framework

https://github.com/uihilab/HydroLang

HydroLang ML - Markup Language

Web Component Library for Hydrological and Environmental Sciences

```
<hydrolang-ml>
    <analyze-mod method="someFunctionName">
        <parameters-here someAtrr="Some Attribute"></parameters-here>
        <arguments-here somArgs="someArguments"></arguments-here>
        <dataset-here>[1,2,3,4]</dataset-here>
    </analyze-mod>
</hydrolang-ml>
```

```
<visualize-mod method="draw">
    <parameters-here type="json" input="usgs_data"></parameters-here>
</visualize-mod>
```

```
<data-mod method="transform">
    <parameters-here input="usgs_data" save="value" output="cleaned_usgs_data1" ></parameters-here>
    <arguments-here type="ARR" keep='["datetime", "value"]'></arguments-here>
</data-mod>
```

HL-ML

https://github.com/uihilab/HydroLang-ML

IOWA

BMI-JS

- ✓ The Basic Model Interface (BMI) is a standardized set of functions that allows coupling of models to models and models to data.
- ✓ The BMI is developed by the Community Surface Dynamics Modeling System (CSDMS).

BMI languages

Language	Specification	Example implementation
C	bmi-c	bmi-example-c
C++	bmi-cxx	bmi-example-cxx
Fortran	bmi-fortran	bmi-example-fortran
Java	bmi-java	bmi-example-java
Python	bmi-python	bmi-example-python

Community-contributed BMI languages

Language	Specification	Example implementation
Javascript	bmi-js	bmi-example-js

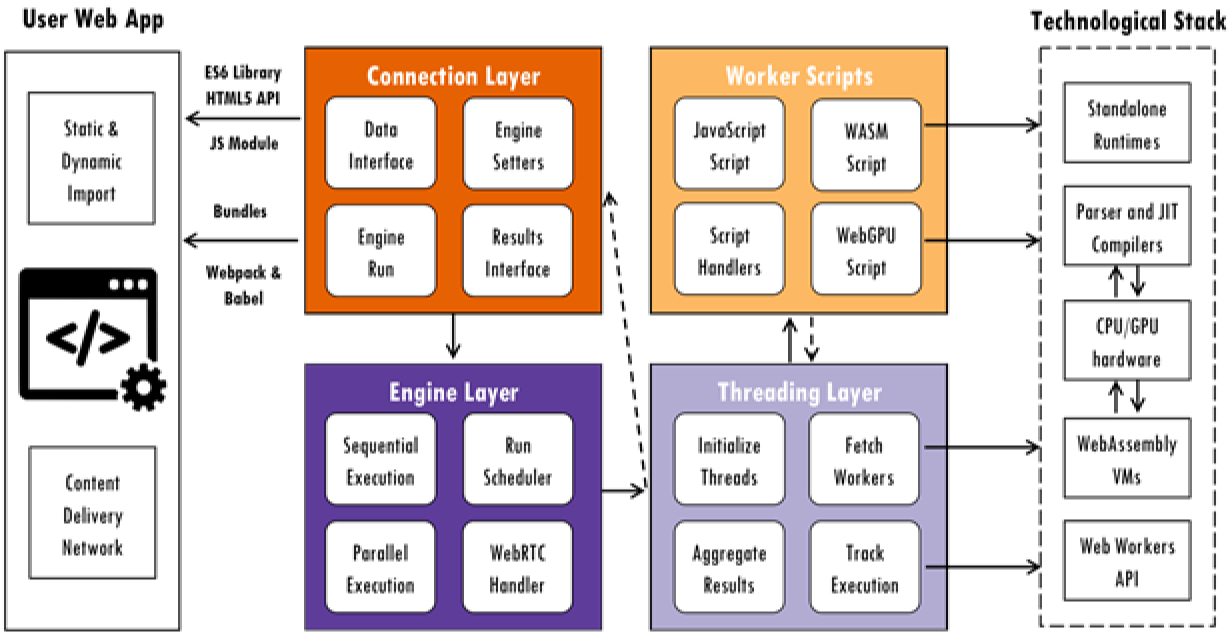
https://github.com/uihilab/bmi-js

https://github.com/uihilab/bmi-example-js

IOWA

HydroCompute

- ✓ Multi-CPU and GPU Based Parallel Computing Library for Hydrology
- ✓ Community oriented modular design for expansion
  - ✓ Special GitHub repo
  - ✓ UI Hackathon

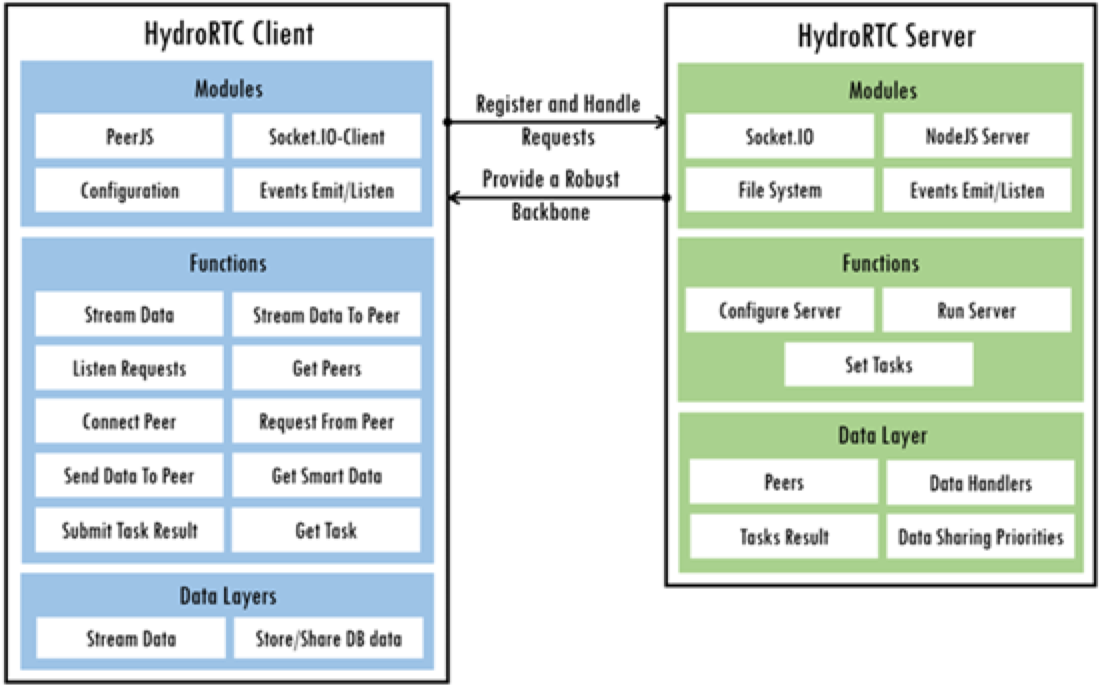


https://github.com/uihilab/HydroCompute

IOWA

HydroRTC

- ✓ Real-time communication library for decentralized server-side applications
- ✓ Enables data streaming, real time data sharing, and hydrologic data interpretation
- ✓ Can be used for deployment of largescale applications



https://github.com/uihilab/HydroRTC

HydroSuite Vision

HydroLang Vision	Markup Language and BMI	Performance Computing And Data	Visual and AI Automated Computing
Initial Stage	**Phase 1**	**Phase 2**	**Phase 3**

Desktop / Client and Server-Side/ Complex / Technical Advanced Coding

Client-Side Seamless / Intuitive / Visual / Codeless Programming

Next Steps

AI-Assisted Programming

- AI Chatbot
- Code Helper
- Code Sandbox

Soon

Visual Programming

- Plug & Play
- Modular
- Building Blocks

Summer-Fall

Automated AI Programming

- Voice Enabled
- Visual Editor
- AI Research Agent

Fall-Spring

Data Retrieval

WaterSoftHack '24

Objectives

- Understanding different data types commonly used in hydrology
- Learning about the multiple data collection methods available
- Recognizing the most common data formats
- Hands on data retrieval from public API's

IOWA

Tools and Libraries

Web-Based Libraries for Data Retrieval

- HydroLang.js (main data connection)
- Python cURL (data manipulation and analysis)

Tools for Development

- VSCode
- Online Resource (Stackblitz for JS)
- Google Colab (Python)

Links for development

Tutorial links (JS):

HydroLang: https://hydroinformatics.uiowa.edu/tutorials/hydrolang/

HydroRTC: https://hydroinformatics.uiowa.edu/tutorials/hydrortc/

HydroCompute: https://hydroinformatics.uiowa.edu/tutorials/hydrocompute/

Other links (JS and Python)

Google Earth Engine: https://earthengine.google.com/

Google Colab (Python) : https://colab.research.google.com/

Example Capstone Projects

- **Real-time Streamflow Monitoring Dashboard**
- **Water Quality Analysis Tool**
- **Climate Change Data Analysis and Visualization**
- **Watershed Management Analysis Tool**
- **Drought Monitoring and Analysis**

Quick example:
https://hydroinformatics.uiowa.edu/lab/hydrosuite/hydrocompute/cs2/

Your ideas?

Part 1 - Introduction to Data Retrieval

Introduction

- Having the correct data is important for any type of modelling
- We can have a positive or negative impact in the decision-making if we are dealing with incorrect information

The Data Retrieval Process

- Collection
- Processing
- Parsing
- Connection to API’s

Data Collection

- The process of gathering and collecting measurable information on variables of interest
- Involves the use of different techniques such as surveys, experiments, observations, and data gathering through sensors and web scraping
- Most crucial step for actionable insights

Data Processing

- Transform, analyze, and generate useful info from raw data
- Stages: data cleaning, transformation, integration, and aggregation
- Processing through filtering, sorting, normalizing, and summarizing data
- Purpose of transforming info to meaningful usable format

IOWA

Data Parsing

- Process of analyzing data and convert to usable formats
- Interpret and reorganize data to usable formats-JSON,XML,CSV or text
- Tokenization, syntax analysis, semantic analysis as techniques
- Data extraction, transformation, integration and communication

Data APIs

- Connect with sources around the world for different formats and scopes
- Accessible anywhere/everywhere through the internet
- Reliability based on the data source and provider
- Most common used way to access data

IOWA

Discussion

What data types do you commonly use in your field?

IOWA

Part 2 - Data Types in Hydrology/Environmental Sciences

IOWA

Data Types

- Time series data
- Spatial information
- Remote sensing
- Water quality

Timeseries

- Observations running for a specified period of time for a particular variable or observation

Examples: water levels, precipitation records, nitrogen concentrations, particulate concentrations, etc.

- Point data: a particular time series for a particular timeframe over a specified time

IOWA

Spatial Data

- GIS outputs from imagery or analysis
- Layers that contain information for time snapshots

Examples: GIS layers, remote sensing images

- **Remote sensing:** satellite imagery, radar data

Water Quality

- Chemical concentrations
- Pollutant levels
- Point source locations

Examples: Eutrophication studies, environmental Impact Assessment

Location-specific, data might not be as reliable!

Example Use Cases

Time Series Data

- Streamflow monitoring, rainfall patterns, groundwater levels

Spatial Information

- Watershed delineation, floodplain mapping, landuse cover analysis

Remote Sensing

- Water body satellite imagery, snow cover analysis, soil moisture estimation

Water Quality

- Pollution source identification, drinking water safety

Part 3a - Data Collection Methods

Methods

- Ground-based sensors
- Satellite observations
- Manual sampling

Ground-based sensors

- **Types:** rain gauges, stream gauges, soil moisture sensors, groundwater level sensors.
- **Data Collection:** Real to near-real time data collection
- **Accuracy and Calibration:** who deployed? What is the calibration process? Source?
- **Challenges:** maintenance, susceptibility, spatial coverage

Satellite Observations

- **Wide Spatial Coverage:** surface water bodies, snow coverage, soil moisture, vegetation
- **Technologies:** optical imaging, radar, LIDAR at different temporal resolutions
- **Data Integration:** coupling ground-based and radar-based data for accuracy and assessment
- **Challenges:** cloud cover interference, complex data processing and interpretation

Manual Sampling

- **Accuracy and Specificity:** pH, turbidity, nutrient concentrations, contaminants
- **Flexibility:** Pollutant source identification or localized assessment
- **Hands-On Verification:** observation and verification of conditions
- **Challenges:** labor intensive, time consuming, limited coverage

Part 3b - Data Formats and Conversions

Examples of Data Formats

Common (easy access) data formats

- JSON, XML, CSV, geoJSON, Shapefiles

Features

- Ease of use and compatibility everywhere
- High performance throughout

Common in the domain

- netCDF, GRIBB, Shapefiles, HDF5

Features

- Multiple variables and temporal resolutions packed
- Not as easy to unpack

JSON and geoJSON

- **JavaScript Object Notation:** used as steering files, data formats for streaming data
- Main format for data exchange on the web, *easy and scalable*
- **Parse throughout environments,** in-built formatting

{ "type": "FeatureCollection", "features": [{ "type": "Feature", "geometry": { "type": "Point", "coordinates": [-91.5302, 41.6611]}, "properties": { "name": "Iowa City" } }] }

XML

- **eXtensible Markup Language:** structured and customizable
- Supports nested data structures through a hierarchical format
- Configuration files, document storage, etc
- **Examples:** WaterML, HydroLangML…

```
<note>
    <to>WaterSoftHack</to>
    <from>Iowa</from>
    <heading>Reminder</heading>
    <body>We got training this
    week team!</body>
</note>
```

Shapefiles and GIS formats

- Representations of spatial data through vectorized formats
- **Rasters, Shapefiles, Databases, KML, geoJSON, TIFF and geoTIFF, GML**
- Combination of spatial features with attribute data

Gridded Formats

- **NetCDF, GRIBB, HDF5,** and others
- Storing large amounts of data through ***efficient stacks of spatio-temporal layers***
- Widely supported across scientific domains
- **Metadata rich,** ensuring proper interpretation

netCDF

IOWA

Conversions

- **Extract features and information** prior to making an analysis
- What is the easiest format for my end use case?

Raw Data -> Usable format -> Get insights

- How can I showcase my data? Maps, charts, other visual formats?
- What is my workflow looking like?

IOWA

Part 4 - Public APIs and Data Resources

IOWA

Application Programming Interfaces

- An API facilitates automated retrieval and integration of data from public sources
- **Promotes easy access** through the use of web technologies

Examples: *USGS, NOAA, NASA, EPA, CUAHSI HydroShare*

Different variables in different locations, need to narrow down specifics!

Application Programming Interfaces

Benefits:

- Real time updates and data sharing
- Comprehensive datasets
- Ease of integration into applications

Access depends on the programming interface!!

Tools: Postman, cURL, Swagger, jQuery, etc…

Web APIs

REST (Representational State Transfer):

- Uses standard HTTP methods (GET, POST, PUT, DELETE)
- Typically in JSON or XML formats
- Scalability, simplicity, and statelessness

SOAP (Simple Object Access Protocol)

- XML for message passing (WaterML)

Workflows for Obtaining Data

Accessing a Dataset or Resource

- Register and Obtain an API Key
- Authentication
- Send a Request
- Receive a Response
- Handle Errors

Types of Responses

- **200 -** OK
- **201 -** Resource created
- **400 -** Bad Request
- **401 -** Unauthorized
- **404 -** Not Found
- **500 -** Internal Server Error

IOWA

Example - Explore different APIs

USGS Water Services

https://waterservices.usgs.gov/

NOAA APIs

https://www.weather.gov/documentation/services-web-api

CUAHSI HydroShare APIs

https://www.hydroshare.org/search/

Other Data Access

- **SFTP (Secure File Transfer Protocol):** Transferring large datafiles from one server to another (personal PC, etc)
- **THREDDS Data Server:** https://www.unidata.ucar.edu/software/tds/
- **GeoServer:** https://geoserver.org/
- **NASA EarthData:** https://search.earthdata.nasa.gov/search
- **Data.gov DEMs:** https://catalog.data.gov/dataset/?tags=sediment-transport&_tags_limit=0
- **European Centre for Medium-Range Weather Forecasts:** https://www.ecmwf.int/en/forecasts/datasets
- **Neon:** https://www.neonscience.org/

IOWA

Part 5 - Data Retrieval and Cleaning

Example - Download data from USGS

USGS Endpoints: instant values, daily values, groundwater

Access URL: https://waterservices.usgs.gov/nwis/dv/?<arguments>

- Arguments contains all the location/time specific variables
- Its a REST API that retrieves data using GET requests
- Available for multiple variables

IOWA

Example - Download data from CUAHSI

Endpoints: Multiple endpoints, multiple use cases

Access URL: https://hydroportal.cuahsi.org/ipswich/ ?<arguments>

- Access of multiple sources through SOAP requests
- Has a unique entry point that allows multiple actions to be applied

Resources

- HDF5 Web Viewer: https://myhdf5.hdfgroup.org/
- GeoTIFF File Viewer: https://app.geotiff.io/
- Open Data API's (alooot of different resources to learn and be more comfortable with data): https://github.com/public-apis/public-apis

IOWA

Next Hour - Hands On Training