

tiny.cc/2019BEF

Tools for Integrating Remote Sensing and Organismal Occurrence Data Streams

NASA BEF Breakout Round 4: Trainings and Tutorials

22 May 2019

3.50-4.30pm & 4.45-5.25pm

While you are waiting:

- 1) Create an account at auth.mol.org/register
- 2) Go to tiny.cc/2019BEF

Tools for Integrating Remote Sensing and Organismal Occurrence Data Streams

AIST-16-0092

Walter Jetz (PI, Yale University)

Adam Wilson (Co-I, State University of New York, Buffalo)

Robert Guralnick (Co-I, University of Florida, Gainesville)

While you are waiting:

- 1) Create an account at auth.mol.org/register
- 2) Go to tiny.cc/2019BEF

Tools for Integrating Remote Sensing and Organismal Occurrence Data Streams



While you are waiting:

- 1) Create an account at auth.mol.org/register
- 2) Go to tiny.cc/2019BEF

AIST Program Objectives

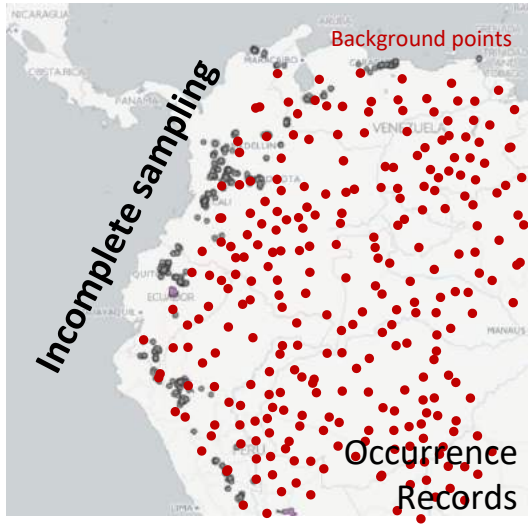
- Reduce the risk, cost, size, and development time for Earth science space-based and ground-based information systems,
- Increase the accessibility and utility of science data, and
- Enable new observation measurements and information products.

Break the 80/20 rule

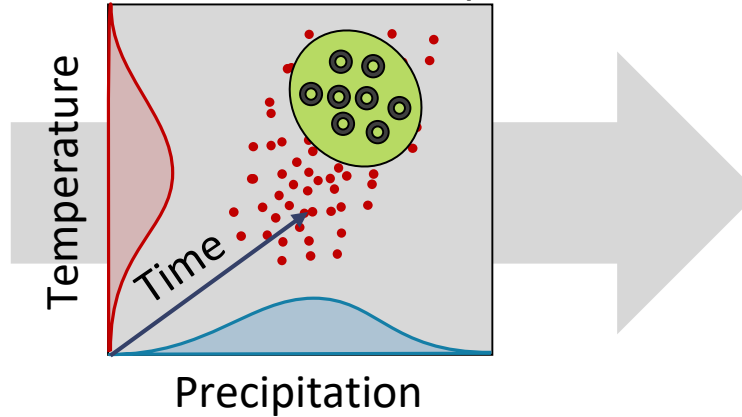
A hand holding a pen over a field notebook in a natural setting. The notebook is open, showing a grid and some handwritten notes. The background is a field of green and brown vegetation under a clear blue sky.

80% of a data scientist's time is spent finding, cleansing, and organizing data, leaving only 20% to actually perform analysis ([IBM 2019](#))

Environmental Niche Modeling



Niche model in
“Environmental Space”



“Species Distribution Modeling”. Google Scholar: 3.5m results (74k since 2018)!

Many ways to use these data

(maxent and friends)

Integrating species occurrence and environmental data is not trivial.

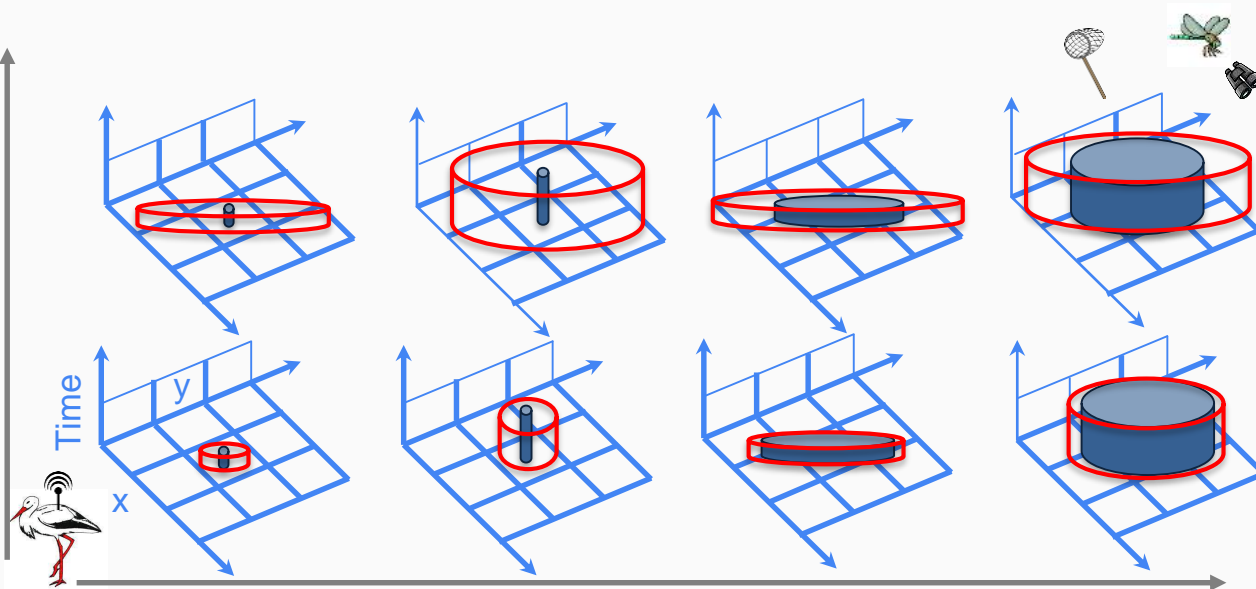


Observation Grain
Locational Uncertainty

Spatio-temporal grain and uncertainties

- Occurrence record
- Environmental data

Increasing spatiotemporal uncertainty



Increasing spatiotemporal grain of observation

Integrating species occurrence and environmental data is not trivial

American Badger
Taxidea taxus



[login/register](#) en de es fr zh

Species Home

Summary Map

Detailed Map

Projection

Search for a species

American Badger

Taxidea taxus

Mustelids



Least Concern (IUCN 2016)

Sources

Point observations

1,279

Expert range maps

1

MOL grid of IUCN

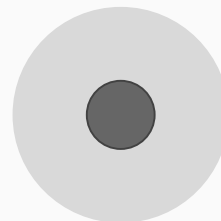
1

Local inventories

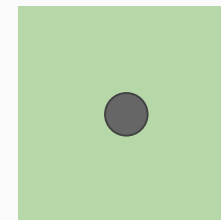
1

Regional checklists

1,197



Spatiotemporal
Uncertainties
(location error)



Grain of
Observation



Example eBird [transect](#)

1. Massive data of varying type

- a. Digitized specimen records
- b. Camera traps data
- c. GPS tags (ICARUS)
- d. Citizen science records



1. Observation-level uncertainties and grains (space and time)

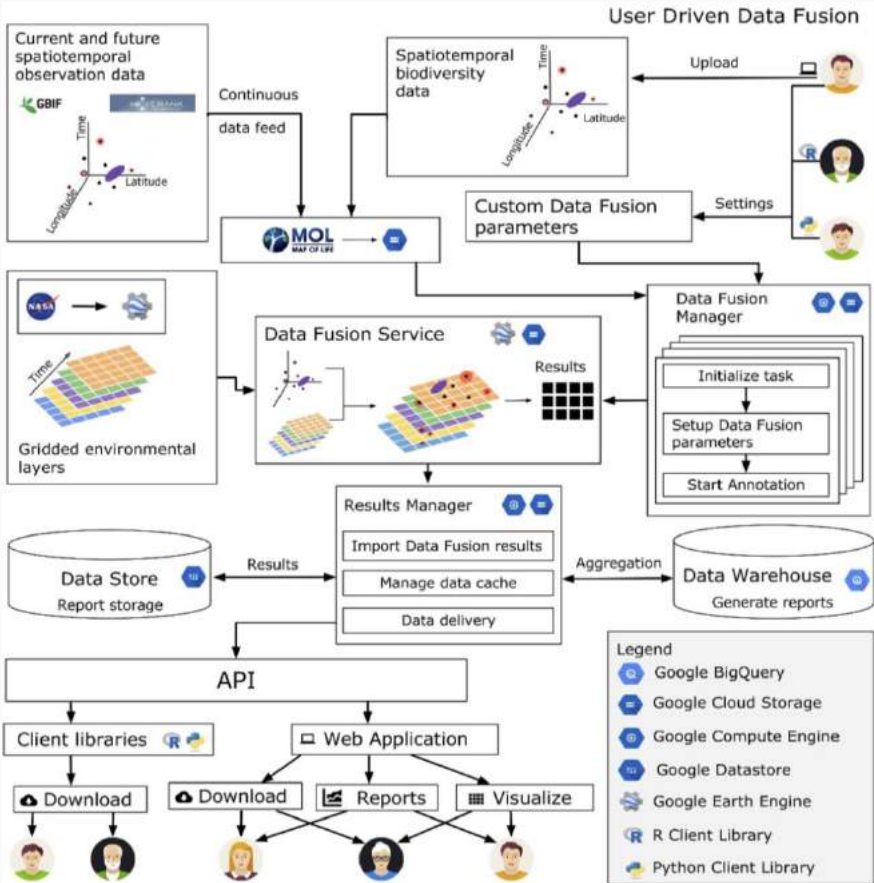
Need to be accounted for in annotation

1. Need for dynamic annotation

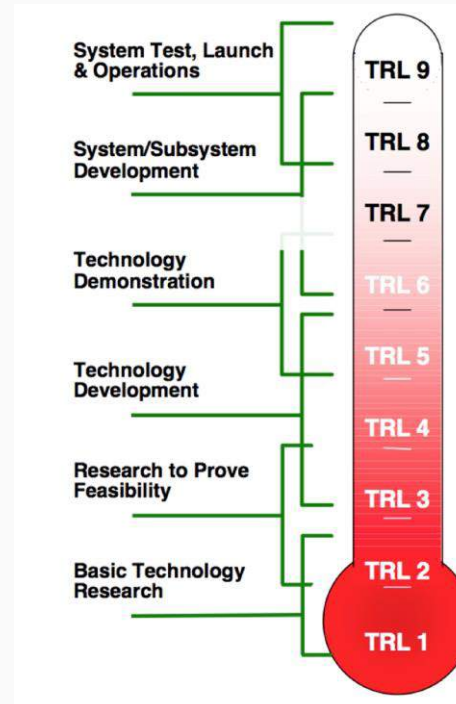
- a. Temperature on the day of observation?
- b. Long-term mean conditions?

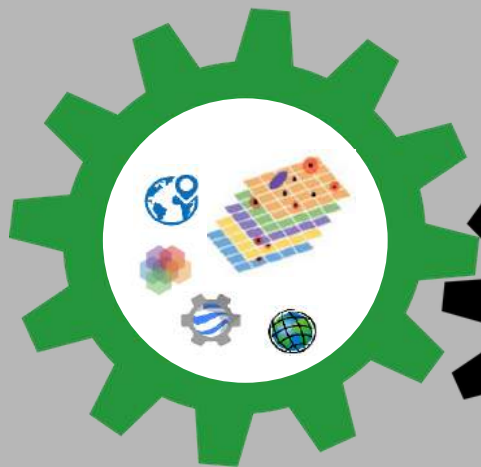


AIST Project Overview



NASA Technology Readiness Level (TRL)





Back End

Front
End
(GUI)



Data Upload

Upload a sample species occurrence dataset

Annotation

Annotate it with various environmental data

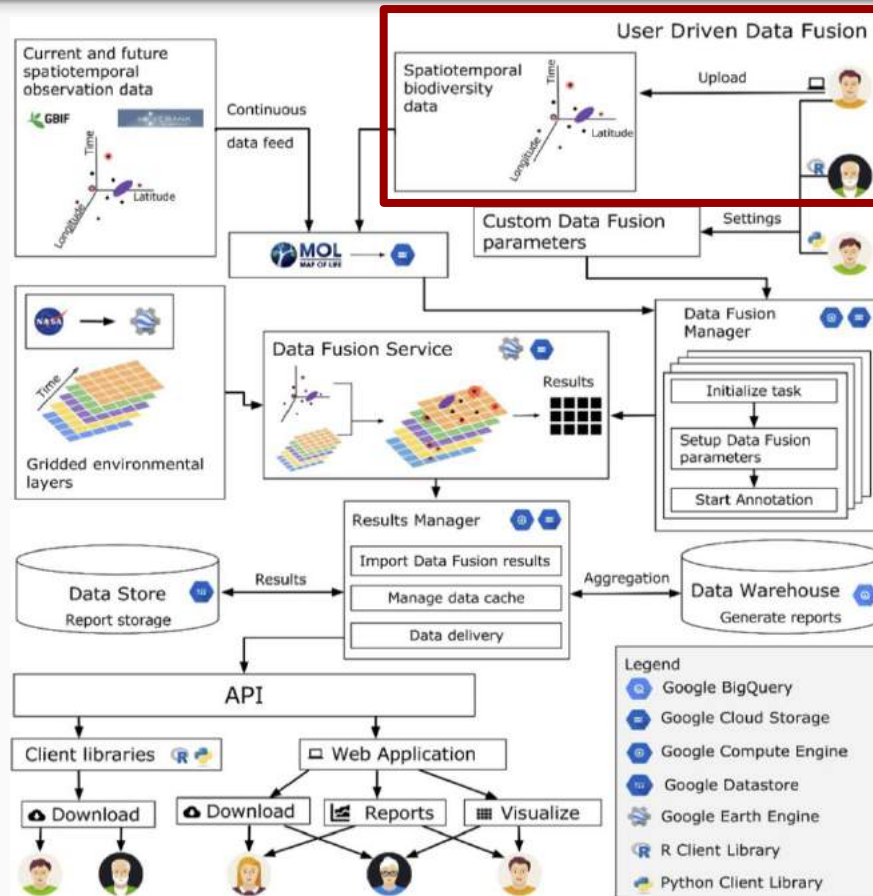
Data Download

Download the annotated data for further processing

Survey

Complete a survey about the design and utility of this application

Project Overview: Data Input



A simple, step by step upload system

<https://mol.org/upload-dev>



Logged in as: [Adam Wilson](#) en-US es

 Species

 Locations

 Indicators

 Patterns



You don't have any datasets yet.



Upload Data



Logged in as: [Adam Wilson](#) ▾ [en-US](#) [es](#)

- Species
- Locations
- Indicators
- Patterns

[← Back to Datasets](#)

1

2

3

4

5

Basic Information — Upload Files — Match Columns — Review Geometry — Metadata

Basic Information

Taxonomic Group

Mammals ▾

Select from the list of groups currently supported by MOL.

Dataset Type

Occurrence ▾

Occurrence or Inventory

Next

[i Instructions](#) ▾



Match Columns

We have detected columns that may match Map of Life field names. Please verify that column names were matched correctly using the drop-down menus below.

Match "scientificName" to:

scientificname

Match "family" to:

family

Match "decimalLatitude" to:

decimal_latitude

Match "decimalLongitude" to:

decimal_longitude

Match "coordinateUncertaintyInMeters" to:

Cancel

Done

logged in as: [Adam Wilson](#) en-US es

Indicators

Patterns

Use example dataset available on website

Upload - Matching Successful, move to finalization

The screenshot shows the Map of Life (MOL) upload interface. At the top left is the MOL logo with the text "MAP OF LIFE". To the right, it says "Logged in as: Adam Wilson" with a dropdown arrow and "en-US es". Below this are four navigation buttons: "Species", "Locations", "Indicators", and "Patterns".

Below the navigation bar is a green bar with a button labeled "< Back to Datasets".

The main content area features a progress indicator with five steps: "Basic Information" (checked), "Upload Files" (checked), "Match Columns" (active, circled in green), "Review Geometry" (circled in grey), and "Metadata" (circled in grey).

The "Match Columns" step is expanded, showing a light blue box with the text: "Please match the column names in the upload file to the standardized field names used by Map of Life. The following terms are required: **scientificname, date, decimal_latitude, decimal_longitude.**"

Below this text, there is a green checkmark icon followed by the text "Columns matched!". To the right of this text is a green button labeled "Match Columns".

At the bottom of the "Match Columns" section, there are two buttons: "Back" and "Next".

At the bottom of the page, there is a section titled "Instructions" with an information icon on the left and a dropdown arrow on the right.



Logged in as: [Adam Wilson](#) ▾ en-US es

Species

Locations

Indicators

Patterns

< [Back to Datasets](#)



Basic Information



Upload Files



Match Columns



Review Geometry



Metadata

Review Geometry



Upload - Dataset Metadata



Logged in as: Adam Wilson + en-US es

Species

Locations

Indicators

Patterns

[Back to Datasets](#)



Basic Information



Upload Files



Match Columns



Review Geometry



Metadata

Metadata

Required Info

Title ⓘ *REQUIRED

Example Badger Dataset

Description ⓘ *REQUIRED

Example data

Geographic Coverage ⓘ *REQUIRED

North America

Permissions ⓘ

Make available for MOL internal use in models ⓘ

Make publicly visible ⓘ

Publish data ⓘ

Uncheck
permissions so the
data isn't published



Logged in as: [Adam Wilson](#) en-US es

Species

Locations

Indicators

Patterns

Available datasets

< Upload Data

Filter by dataset type

All



Example Badger Dataset

PRIVATE

QUEUED

Wait until the data are processed

Example Badger Dataset

[Back to Datasets](#)

[Annotate](#)

[Harmonize](#)

[Delete](#)

This dataset isn't finalized yet! This means: occurrence and inventory datasets will **not show on any MOL apps**.

This dataset is **private**. Click the **Permissions** tab below to change dataset permissions.

[Overview](#) [Map](#) [Metadata](#) [Permissions](#) [Published Data](#)

GROUPS

1

SPECIES

1

RECORDS

1611

Species Groups

	Group	Count
	mammals	1

Species

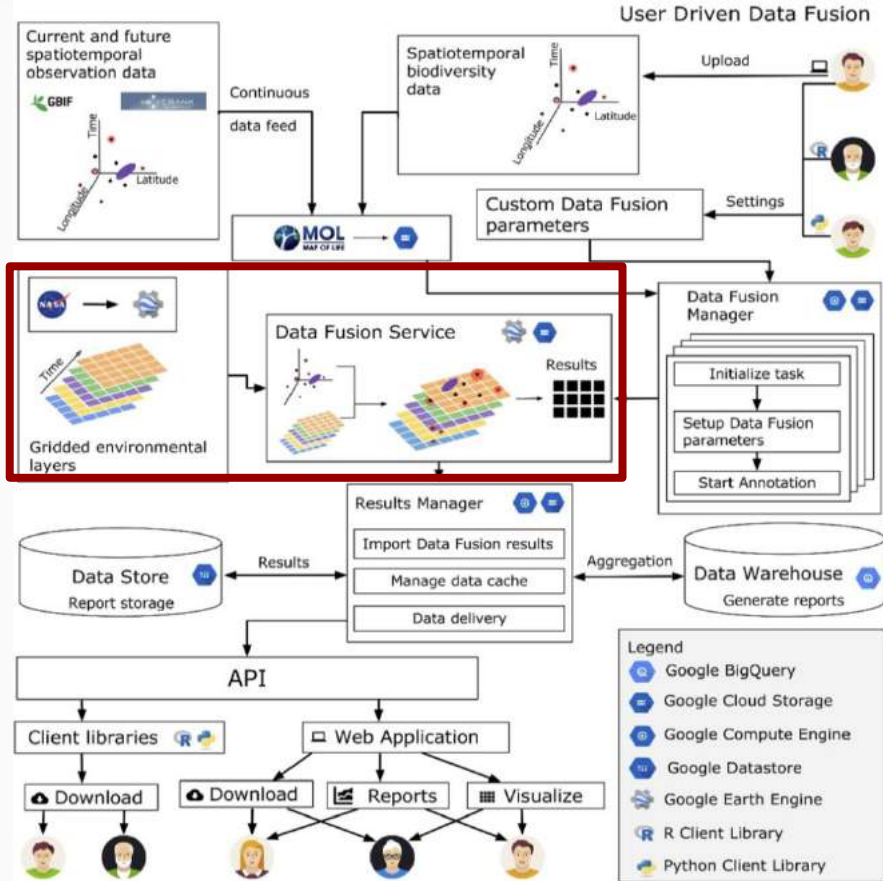
	Species	Count
	American Badger <i>Taxidea</i>	1611

Latest Records

	Species	Event date
	American Badger <i>Taxidea</i>	2018-08-26

If you don't see the "Annotate" button, confirm you are at mol.org/upload-dev

Project Overview: Biodiversity-relevant remote-sensing - derived layers



Environmental data availability increasing rapidly

Current:

- MODIS EVI
- CHELSA

Planned:

- Radar-derived forest cover
- Evapotranspiration
- Global 30m DEM
- Surface Water

Environmental Annotator

Environmental Annotator

Search for species

MOL Upload

Product

EVI

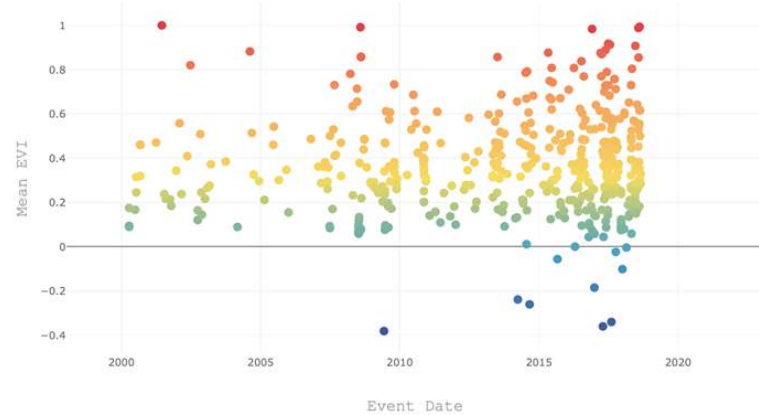
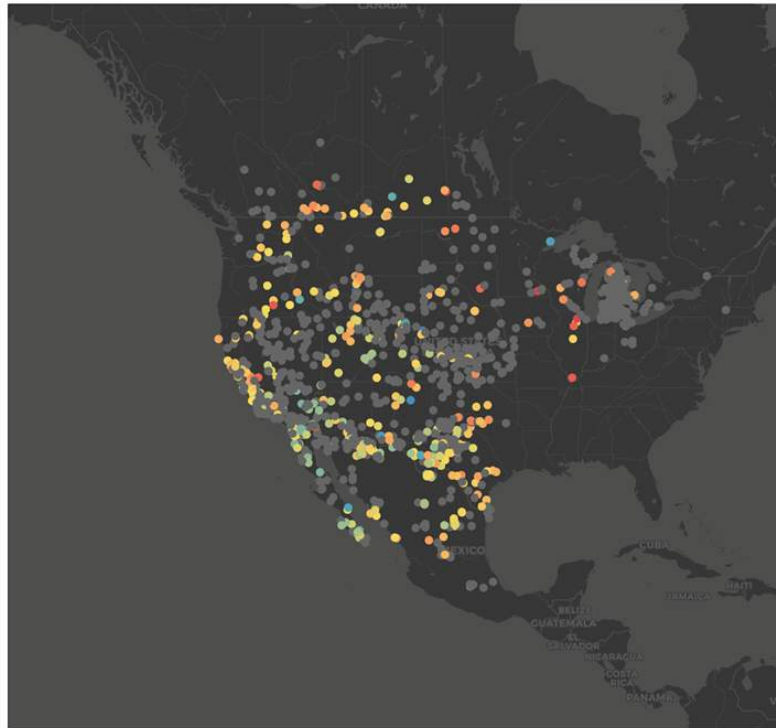
Spatial Buffer

100 m

Temporal Buffer

1 day

Go



May have to adjust x-axis by selecting dates of interest (post-2000)

Environmental Annotator - EVI

Environmental Annotator

Search for species

MOL Uploader

Product: LST Day
Spatial Buffer: 100 m
Temporal Buffer: 1 day

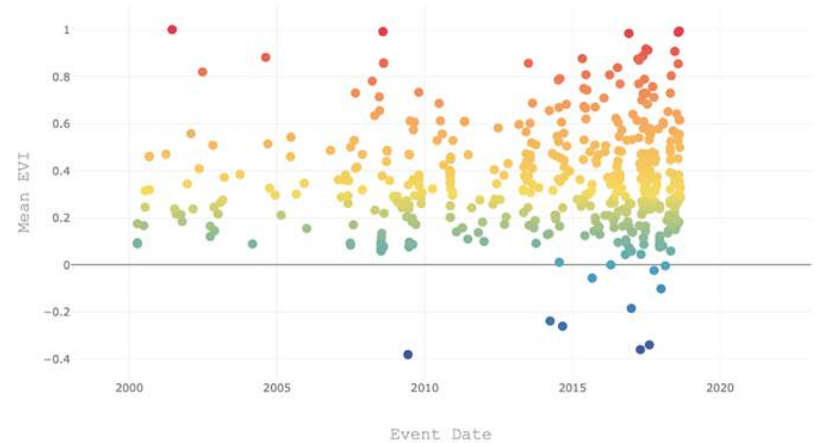
Go



Filter

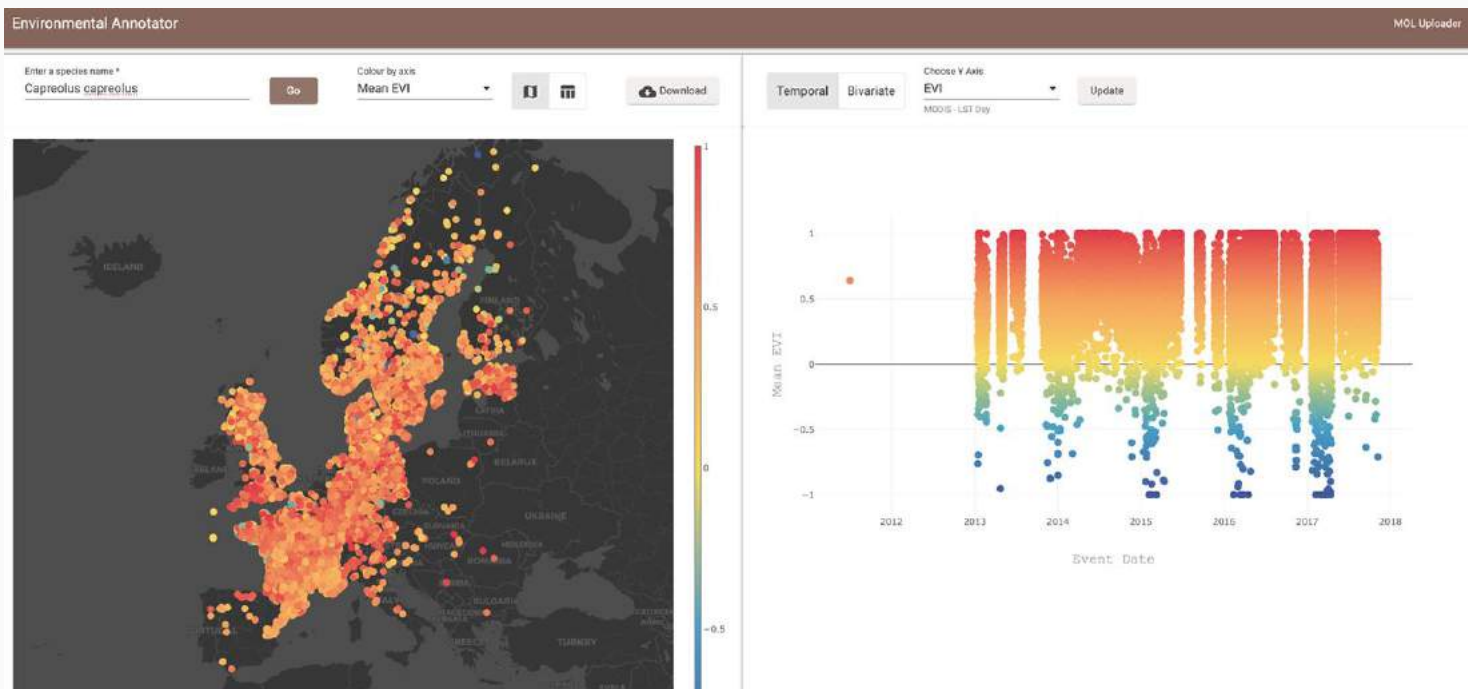
Scientific Name	Latitude	Longitude	X Axis	Y Axis
Taxidea taxus	35.4	-120.8333	1957-02-06T00:00:00Z	
Taxidea taxus	49.7167	-99.6	1923-12-09T00:00:00Z	
Taxidea taxus	42.6884	-81.7127	1948-08-01T00:00:00Z	
Taxidea taxus	51.09	-114.54	2008-08-08T00:00:00Z	0.8585
Taxidea taxus	43.03	-81.35	1990-04-08T00:00:00Z	
Taxidea taxus	32.89	-100.233	1966-03-26T00:00:00Z	
Taxidea taxus	30.5428	-98.7037	1997-01-04T00:00:00Z	
Taxidea taxus	34.179	-100.306	1968-07-27T00:00:00Z	
Taxidea taxus	33.811	-102.18	1985-09-30T00:00:00Z	
Taxidea taxus	34.723	-102.904	1985-12-20T00:00:00Z	

Items per page: 10 1 - 10 of 1611 |< < > >|



Roe Deer
Capreolus capreolus

Or select any
desired mammal
species.





Environmental Annotator

MOL Uploader

Enter a species name *
Capreolus capreolus

Go

Colour by axis

MODIS - LST Day



Download

Temporal

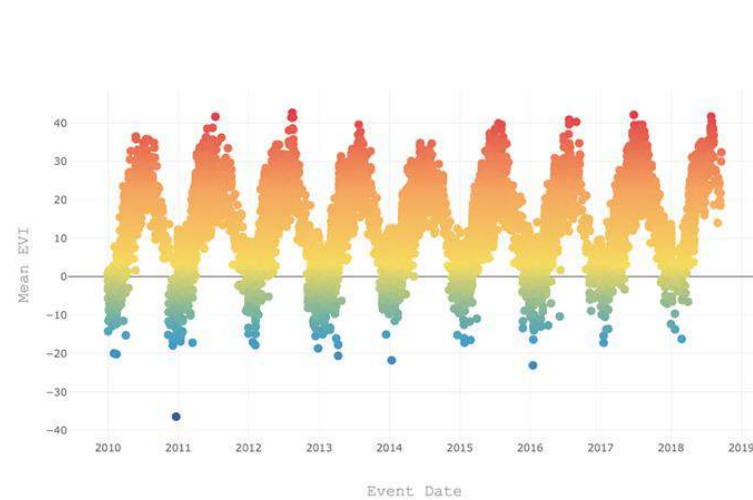
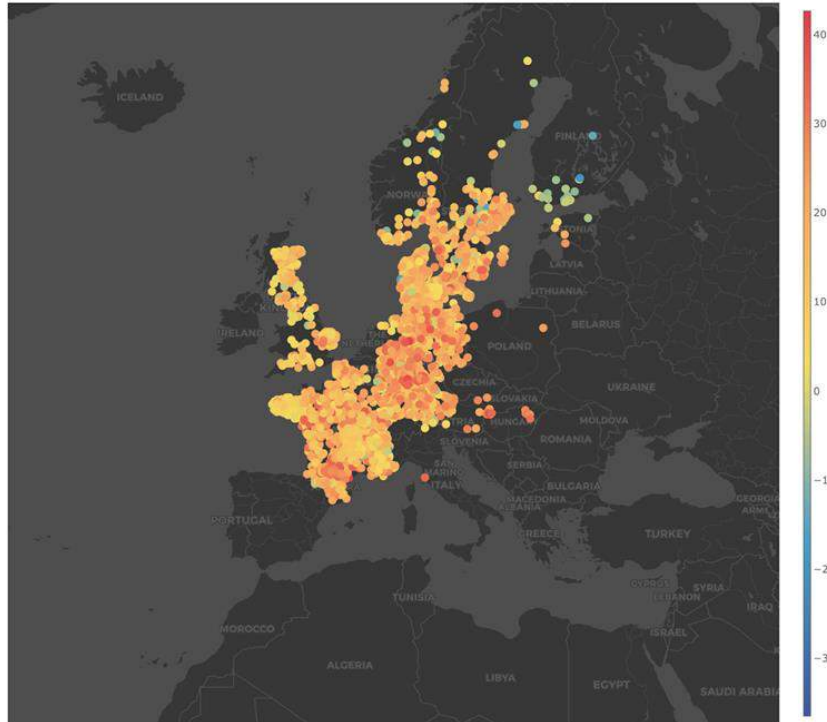
Bivariate

Choose Y Axis

LST - Day

Update

MODIS - LST Day



Environmental Annotator

Environmental Annotator

MOL Uploader

Enter a species name *

Taxidea taxus

Go

Colour by axis

Annual Mean Precipi...



Download

Choose X Axis

Mean

Annual Mean Temperature - Mean

Choose Y Axis

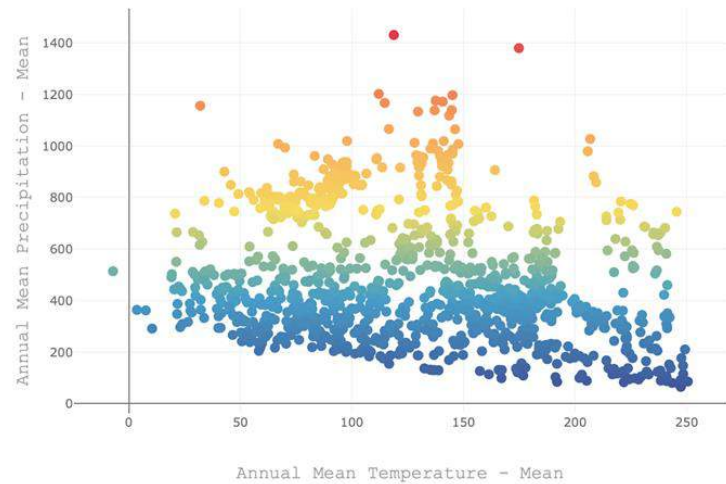
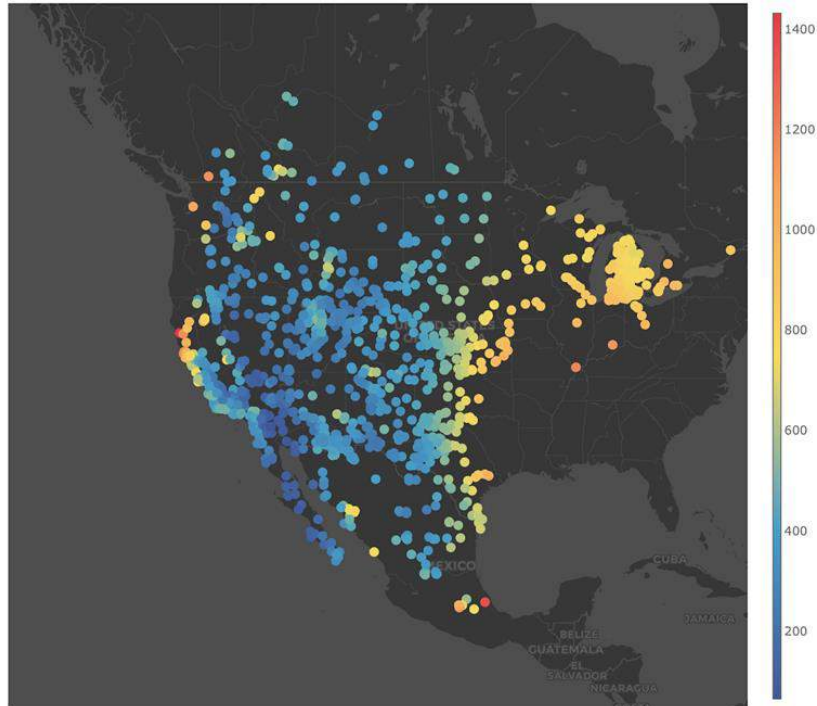
Mean

Annual Mean Precipitation - Mean

Update

Temporal

Bivariate



Characterize the available habitat

Background points

Environmental data availability

Increased options for environmental data products

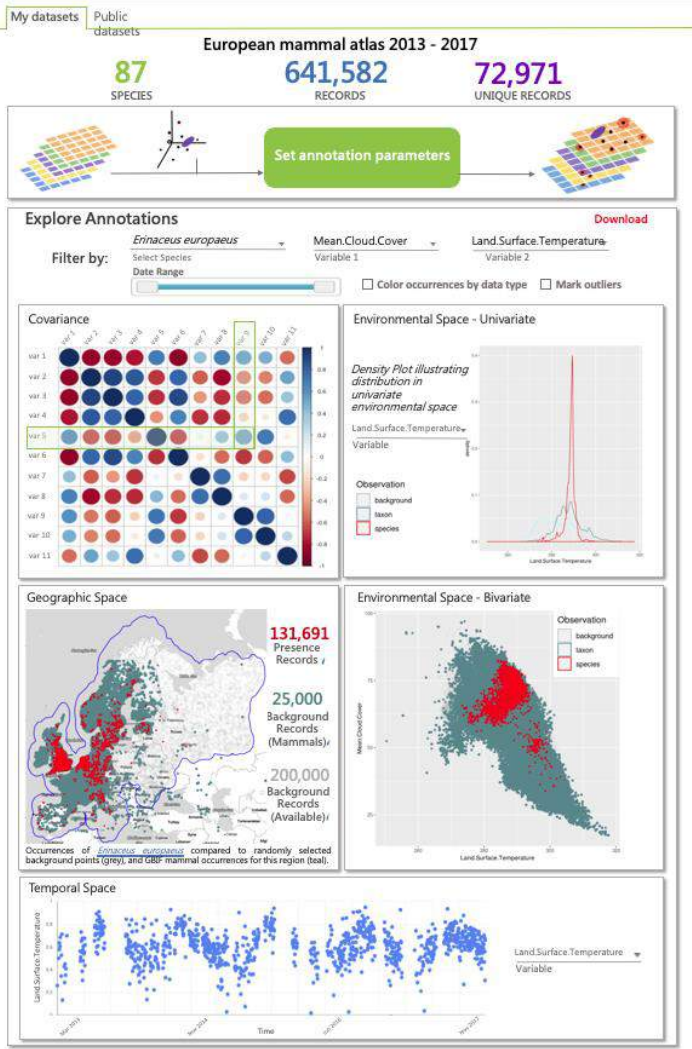


Spatial and temporal aggregation

Increase the options for spatial and temporal aggregation

Exploratory Data Analysis and Variable Selection

Preliminary comparisons of environmental datasets to guide user through variable selection



Products to be included in Phase II

Type	Product	Platform / Sensor / Product	Archive ¹	Resolution ²		Domain
				Temporal	Spatial	Temporal
Terrestrial	Surface Reflectance	Landsat 4-8	D E G	16 days	30m	1982
		MODIS (MCD43A2)	D G	Daily	250m	2000
		ASTER	D G	16-day	15-30m	2000
		Sentinel 2	D G	10-day	10-60m	2014
		Airbus SPOT (OneAtlas)	D	Once	1.5m	2015
		Airbus Pléiades (OneAtlas)	D	Once	0.5m	2015
	Vegetation Indices (NDVI/EVI)	MODIS (MOD13)	G	8-day	250m	2000
		Descartes MODIS 16-day VI	D	16-day	250	2000
	Land Cover	MODIS (MCD12Q1)	D G	Annual	1km	2001
	Land Cover Dynamics	MODIS (MCD12Q2)	D G	Annual	1km	2001
3D Structure	Sentinel-1 SAR	D G	12 days	10-25m	2014	
Elevation	Airbus	E	Once	~24m	2017	
Freshwater	Surface Water	JRC Global Surface Water	G	Monthly	30m	1984
		MOD44W Surface Water	G	Annual	250m	2000
Marine	World Ocean Atlas	Temperature, Salinity, Oxygen; others	E	Decadal	25km	1955
	NOAA OI SST	Sea Surface Temperatures	E	Daily	25km	1981
	Sea Surface Temperatures, Topography, Colour	Sentinel-3	D G	1-2 days	300m - 1km	2014
		MODIS Aqua	G	Daily	1km	2001
		MODIS Aqua	G	Daily	1km	2001
Various	HYCOM + NCODA	E		~10km	2008	
Climate	Various Climate Variables	TerraClimate	G	Monthly	1km	1958
	Temperature	³ CHELSA / EarthEnv	G	Daily	1km	1979
	Precipitation	³ CHELSA/ EarthEnv	G	Daily	1km	1979
	Growing Degree Days	CHELSA Derived	G	Monthly	1km	1979
	Land Surface Temperature	MODIS (M*D11A1/A2)	D G	8-day	1km	2000
	Thermal IR	ASTER	D G	16-day	90m	2000
	Cloud Cover	EarthEnv	G	Monthly	1km	2000

We want your feedback! Please fill out the survey!

Workshop Feedback

Software Workflows and Tools for Integrating Remote Sensing and Organismal Occurrence Data Streams to Assess and Monitor Biodiversity Change

Select variables and spatial resolutions that are important for your work.

	15m	30m	250m	1km
Vegetation Indices / NPP / GPP	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Land Cover Classification (e.g. forest, urban)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Continuous Vegetation Fields (e.g. % forest)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Land Surface Temperature	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Please provide feedback on the current prototype and offer suggestions about what else this tool *should* do.

Direct survey link here

<https://forms.gle/pxu69giSBmUvwYDMA>





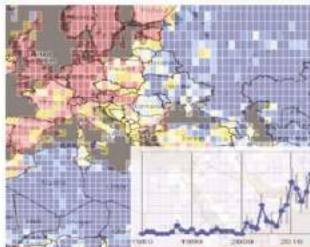
Map species

Views species range map, inventory, and occurrence data



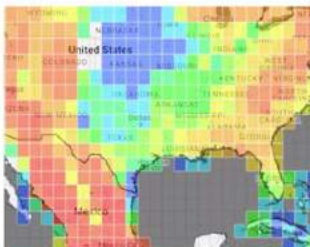
Species by location

Select a location, filter by distance or group, and view a list of species along with source data



Indicators

Explore trends in biodiversity knowledge, distribution, and conservation



Patterns

Explore richness patterns and biodiversity facets



Datasets

Explore datasets used across MOL



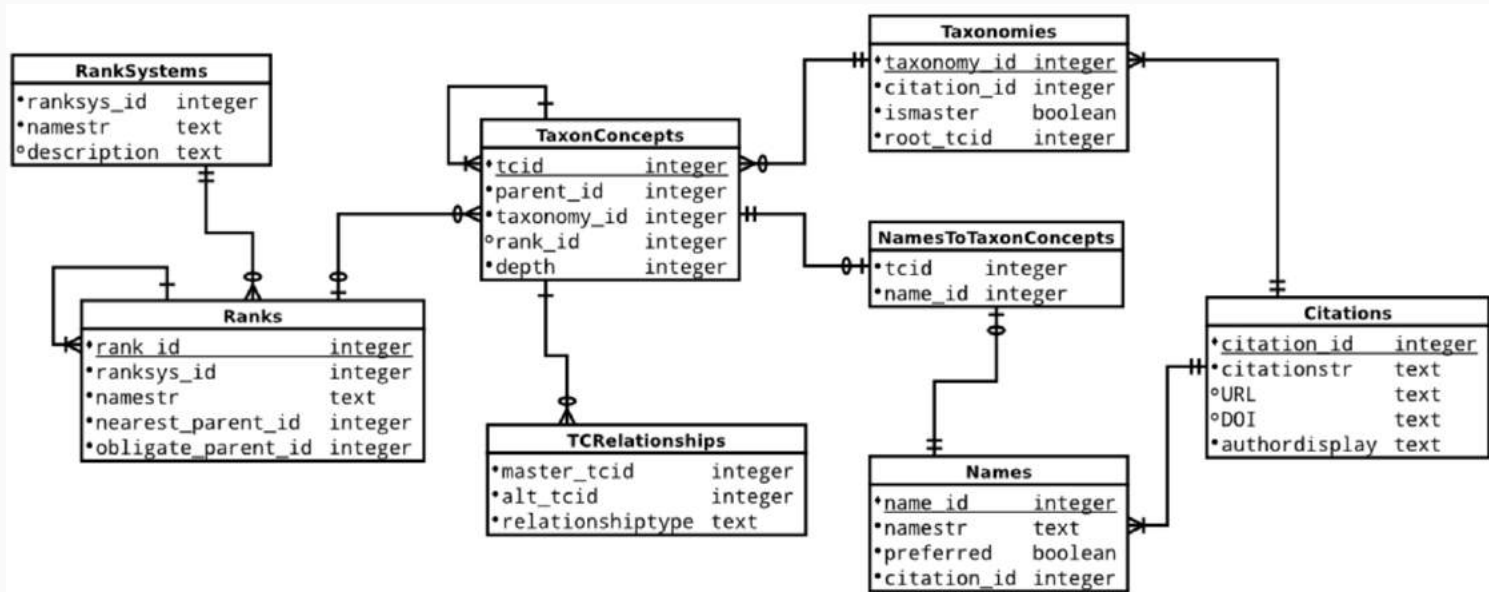
Mobile App

Discover, identify, and record biodiversity worldwide

Thank you for your attention!


Management On The Back-End

- MOL has a well developed access control system
- One of the biggest challenges is managing taxonomy/species names
- Taxonomy management - nearly a million names and synonymies
- Prototype tools to help harmonize names



Satellite and model-derived 1km environmental data at varying temporal resolutions

	Source	Daily	Monthly	Annual	Climatology
Enhanced Vegetation Index	MODIS	Available in the current prototype			
Land Surface Temperature	MODIS	Available in the current prototype			
Air Temperature (Daily Min, Max, Mean)	CHELSA				Available in the current prototype
Precipitation (Daily Min, Max, Mean)	CHELSA				Available in the current prototype

 Available in the current prototype