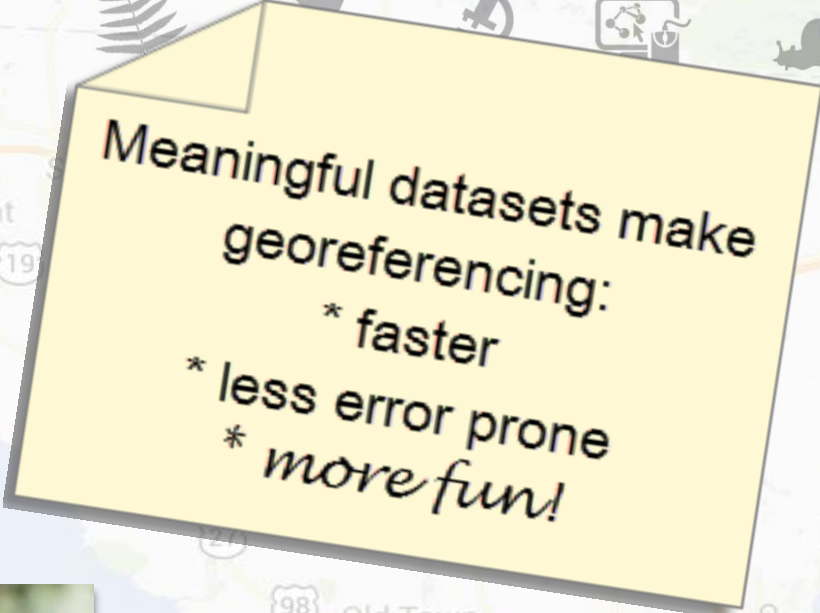


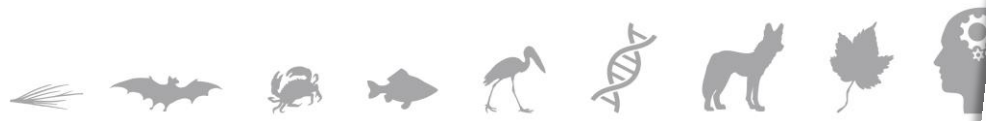
Georeferencing Overview

Deborah Paul
iDigBio, Florida State University
Bristol UK Swans Practical Digitisation Workshop
9 March 2018
@idbdeb @iDigBio



Meaningful datasets make
georeferencing:
* faster
* less error prone
* more fun!





iDigBio Georeferencing Working Group (GWG) Train the Trainers I and II

- 5 Days
- Pay-it-forward
- Community Resource

10:00	Workshop Overview, Introduction to Georeferencing, and Thinking like a Trainer	David Bloom
10:20	Selected Survey Results, Train-the-Trainers Second Georeferencing Pre-workshop Survey Report and Update from TTT1 Participants	Shari Ellis
10:35	Break	
11:00	Georeferencing Introduction: Collaboration to Automation	
11:30	Geographical Concepts	Dave Bloom
11:50	Point-Radius Method and Best Practices	Nelson Rios
12:10	Darwin Core Standard, Key Terminology iDigBio Recommended fields	Jessica Utrup
12:30	Lunch (Provided)	Dave Bloom
13:30	Georeferencing Quick Reference Guide, Locality Types, and Georeferencing Template	
14:40	Georeferencing Calculator, Get...	Una Farrell



<http://tinyurl.com/idbttt2>



Topics in a 5 day georeferencing course

- Introduction to iDigBio
- **Introduction to Georeferencing**
- Thinking like a Trainer
- Collaboration to Automation
- Geographical Concepts
- Point-Radius Method and Best Practices
- **Darwin Core** Standard, Terminology, iDigBio Recommended Fields
- **Georeferencing Quick Reference Guide, Locality Types**, Georeferencing Template
- MaNIS/HerpNET/ORNIS Georef Guidelines
- **Internet Resources** – Where to Begin?
- GPS Exercise Introduction
- Georeferencing Using Paper Maps
- Process, Workflows, Priorities, and Collaborations
- KE Emu, KUMIP and Specify, FishNet2
- ORNIS Workflows and Repatriation
- Workflows, FSU Georeferencing Protocol
- Participant Workflows
- **Good and Bad Localities**
- Georeferencing Natural History Collections Data: The **GEOLocate** Project
- Using GEOLocate: Basics, Batch Processing, Collaborative Georeferencing Administrative Portal, Collaborative Georeferencing Web Client, Taxon validation, Web services and integration, Building end-to-end georeferencing workflows
- Data Cleaning, Processing, and Analysis
- Participant/TCN Georeferencing Projects (use your own data sets)
- Batch Georeferencing in Symbiota
- Participant Volunteer Training Demos



Goals of georeferencing (during or after data capture)

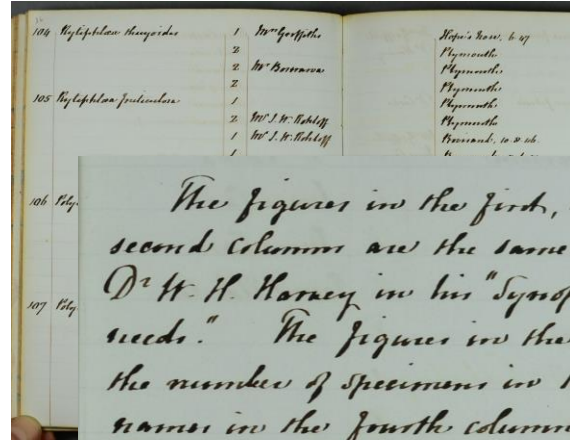
- Read and transcribe written materials
- Move accurate data into database
- Use this data to derive a decimal latitude and longitude
 - note that some georeferencing may be automated
 - remember to check if another project may have useful georeference data
- Include (at least)
 - decimal latitude and longitude
 - uncertainty
 - how you did the georeference
 - what sources you used
 - who did the georeferencing
 - any assumptions made
 - the geodetic datum, if known, or assumed
 - a text description of the locality
 - the verbatim coordinates if present





Data Capture and Georeferencing Challenges

- ink
- typed
- pencil
- printed
- stacked
- handwritten
- uneven lines
- colored paper
- non-planar surfaces
- non-standard terms
- non-standard formats



The figures in the first, and the names in the second column are the same as those employed by Dr H. H. Harvey in his "Synopsis of British Sea-weeds." The figures in the third column show the number of specimens in Herbarium. The names in the fourth column are those of the Donors. The names in the fifth column are those of the Collectors. The habitats, when known, are recorded in the last column. "Plymouth" refers to the district, within ten miles of that town.

Herbarium Catalogue
of
British Marine Algae
111 Volumes.



Ac. No. 150/1975 Lab. No. CN/75/36
THE CITY MUSEUM, BRISTOL.
LACERTA VIVIPARA ♂ Ad. 598
THE MINERIES, NR. PRIDDY, SOMERSET.
28-8-1968. coll: D.J. FOXWELL.
O.S. Ref. ST/544507.
FREEZE DRIED: DECEMBER 1968. D.J. FOXWELL.

St. Ives (C.) 14
2. VII. 1928 A.T.
Echinomyia
fera. L ♂

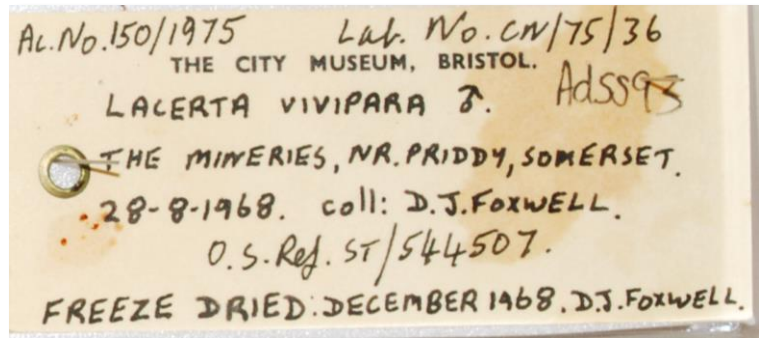
HERBARIUM
IVOR W. EVANS.
Name *Callium rhinoceros*
Roxey (dead) Parlin
Locality The Gully, Clifton
Bristol in exposed turf
No. 607/A Date July 14/1914

Ac. No. 1998/0024. Reg. No. Ad. 2584
THE CITY MUSEUM, BRISTOL
Eng. Name: JAY
Species: GARRULUS GLANDARIUS
Locality: DEAD ON STEPS - BLAISE CASTLE
MUSEUM - HENSBURY - AVON ST. 559783
Date: 23-06-1988 Sex: ♂ juv
Weight: 25 mm. Field Coll: Sarah Kayner
153.3 mm. Taxidermist:
Lab. No. CN/88/17



Two main paths to georeferencing

- existing legacy data that needs a georeference



- new data coming into your collections – “born digital”





What is an ideal georeference?

A numerical description of a place that can be mapped and that describes the spatial extent of a locality and its associated uncertainties as well as possible.



You'll write your own georeferencing guide

here are two (of many possible) resources to refer to

GEOREFERENCING QUICK REFERENCE GUIDE

Version: 2012-10-08

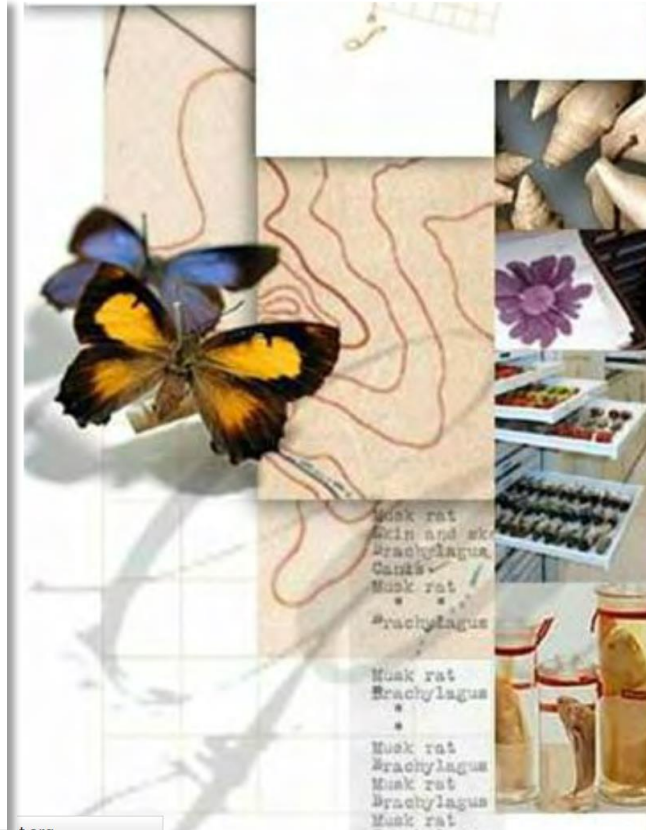
John Wieczorek, David Bloom, Heather Constable, Janet Fang, Michelle Koo, Carol Spencer, Kristina Yamamoto

This is a practical guide for georeferencing using the point-radius method [1, 2, 3] using the Georeferencing Calculator [4, 5], maps, gazetteers, and other resources from which coordinates and spatial boundaries for places can be found. This guide is an update of "Georeferencing for Dummies" [6], and explains the recommended calculation procedure for localities encountered in the georeferencing process.

Georeferences using the methods in this guide will be maximally useful if as much information as possible is captured about and during the georeferencing process in the following fields defined in the Darwin Core standard [7]. For additional community discussion and recommendations, see the Darwin Core Project wiki [8].

Darwin Core Georeferencing terms:

- **decimalLatitude, decimalLongitude, geodeticDatum** – the combination of these three fields provide the reference for the center of the point-radius representation of



Guide to Best Practices for Georeferencing



General steps in a georeferencing scheme

- Are you georeferencing as you capture data?
 - or doing your *georeferencing after data entry*?
- Has someone else done it already?
 - UK or Non-UK material?
 - NHM Data Portal - 100Ks georeferences
 - NBN Atlas
- Collate your data to be georeferenced
 - will it be georeferenced in or out of your database?
- Clean data in the database before beginning
 - locality, collecting event data, etc.
- separate specimens
 - coordinates yes, coordinates no
- Standardize locality strings
 - dwc:locality
 - dwc:verbatimLocality
- Group data into logical sets
 - same geographic area, same collecting event, for example
 - collecting events OR each record has its own georeference?
- Clustering
 - creates efficient search strategy
 - optimizes your workflow
- Take advantage of staff, volunteer, visitor skills
 - reads “Russian” or can read old script
 - researchers return material georeferenced
- Tic-Tock, Tic-Tock



Workflow overview (legacy data)

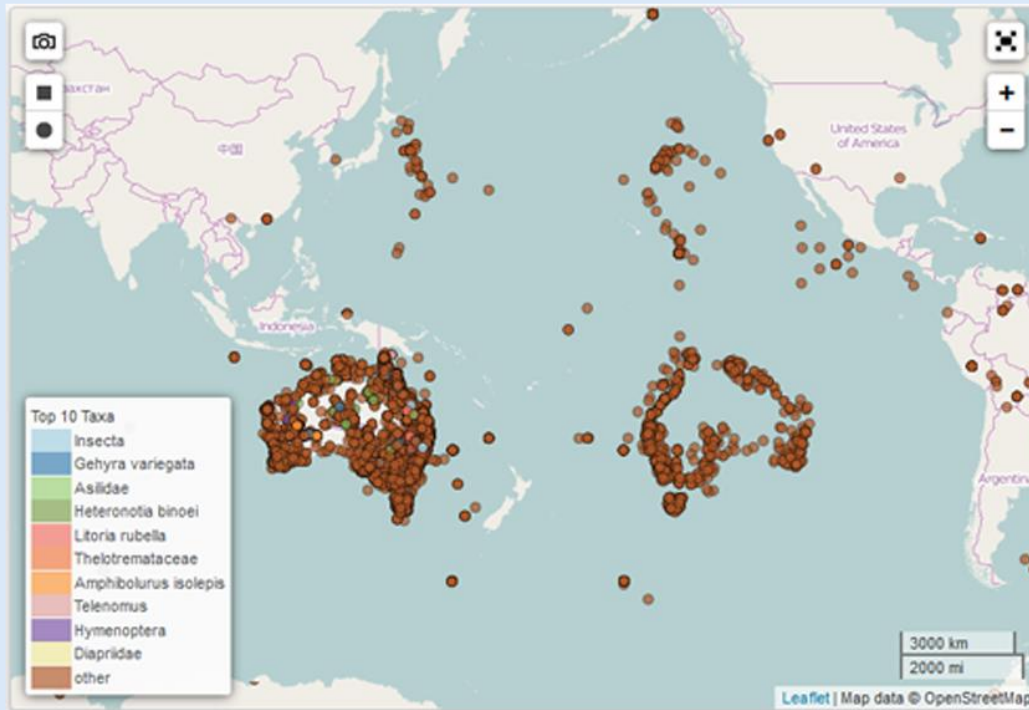
- separate specimens
 - coordinates yes, coordinates no
- clean the data
 - group the data if possible, for georeferencers (country, collector, region, etc.)
- upload (if using GEOLocate), or as is relevant
- identify-verify, use another resource if needed
- add extent and uncertainty
- dwc:georeferenceRemarks and georeferenceProtocol
- export (from GEOLocate)
- visualize your work (try www.gpsvisualizer.com, Google Earth,...)
- check for completeness in your database (georeferencedBy, georeferencedDate, etc.)

Task ID	Task Description	Explanations and Comments	iDigBio Resource
T1	Separate specimens that have coordinates from those that do not.	For specimens lacking coordinates, batch or collaborative georeferencing should be performed. For specimen records that already include geographic coordinates, it is necessary to input the datum used,	Specimen occurrence records with locality data.
T2	Clean and transform data for georeferencing.	The batch or collaborative functions of GEOLocate require a very particular CSV format (refer to the link). Sort the data appropriately (e.g., county, collector, year, collection number, etc.) so that specimens with identical localities are not repeatedly	Spreadsheet application. See: GEOLocate tutorial, http://www.museum.tulane.edu/geolocate/standalone/tutorial.html .



iDigBio Data Quality (DQ) Flags enhance Digitization Workflows

Example: spot and fix georeferencing



Flag
idigbio_isocountrycode_added ⓘ
dwc_continent_added ⓘ
dwc_country_replaced ⓘ
geopoint_datum_missing ⓘ
dwc_class_replaced ⓘ
dwc_phylum_replaced ⓘ
dwc_order_replaced ⓘ
geopoint_low_precision ⓘ
rev_geocode_eez ⓘ
dwc_stateprovince_replaced ⓘ
rev_geocode_mismatch ⓘ
dwc_order_added ⓘ
datecollected_bounds ⓘ
dwc_class_added ⓘ
dwc_kingdom_added ⓘ
dwc_phylum_added ⓘ
dwc_country_added ⓘ
rev_geocode_corrected ⓘ
rev_geocode_lon_sign ⓘ



Standard terms from DwC for georeferencing

Darwin Core Location Terms

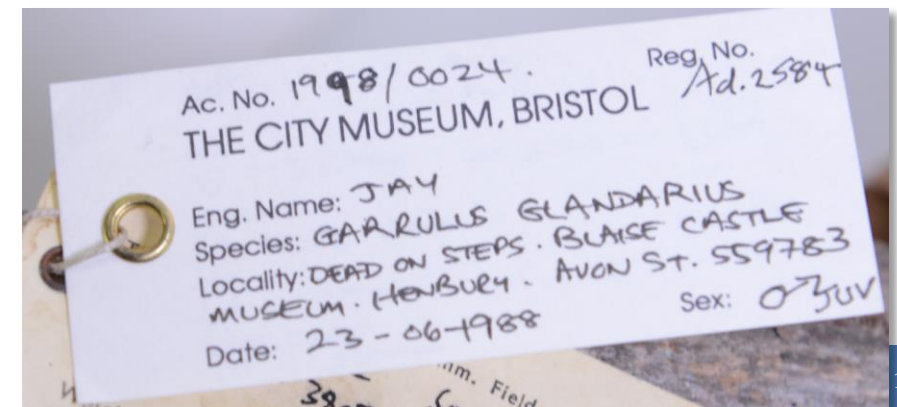
- higherGeography
- waterbody, island, islandGroup
- continent, country, countryCode, stateProvince, county, municipality
- verbatimLocality, locality
- minimumElevationInMeters, maximumElevationInMeters, minimumDepthInMeters, maximumDepthInMeters
- decimalLatitude, decimalLongitude, geodeticDatum, coordinateUncertaintyInMeters, georeferencedBy, georeferenceProtocol, georeferenceSources, georeferenceVerificationStatus, georeferenceRemarks

Darwin Core Event Terms

- habitat

Darwin Core Geological Context

- group, formation, member, bed, ...





From the Label, Notebook, ...

WAKULLA CO.: St. Marks Nat'l Wildlife Refuge (Panacea Unit). Frequent in moist roadside depression, less so in drying sand of burned, open longleaf pine along W side Rte 372, just N of Rd 401 and 1.4 mi drive from Hwy 98.

field notes

- 41 05 54S
- 121 05 34W
- WGS84
- 2.5 mi. NE Tlh. on Ctrville Rd.
- Tallahassee, 2.5 miles NE on Centerville Road.
- frequent
- Wakulla
- in moist roadside depression, ...

your database field

- **lat** or **latitude**
- **lon** or **long** or **longitude**
- **datum** or **notes** or ...
- **loc** or **location** or **collectorLocality** or ...
- **abundance**
- **county**
- **hab** or **habitatDescription** or ...

darwin core

- verbatimLatitude
- verbatimLongitude
- verbatimSRS
- verbatimLocality
- locality
- (occurrenceStatus)
- county
- habitat





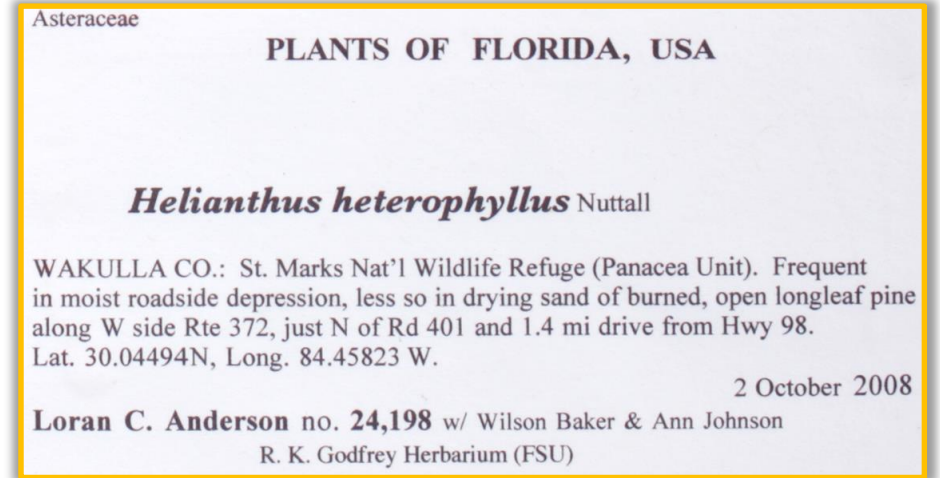
Darwin Core Georeference Terms

darwin core terms

decimalLatitude
 decimalLongitude
 geodeticDatum
 coordinateUncertaintyInMeters
 georeferencedBy
 georeferenceProtocol
 georeferenceSources
 georeferenceVerificationStatus
 georeferenceRemarks
 coordinatePrecision
 pointRadiusSpatialFit
 footprintWKT, footprintSRS,
 footprintSpatialFit

example values

- 30.441115
- -84.295903
- WGS84
- 20
- Susan Somewhere
- Georeferencing Quick Reference Guide
- Falling Rain Gazetteer, GEOLocate
- verified by collector
- assumed distance by road



goal to populate these fields!





Geographical Concepts: Datums

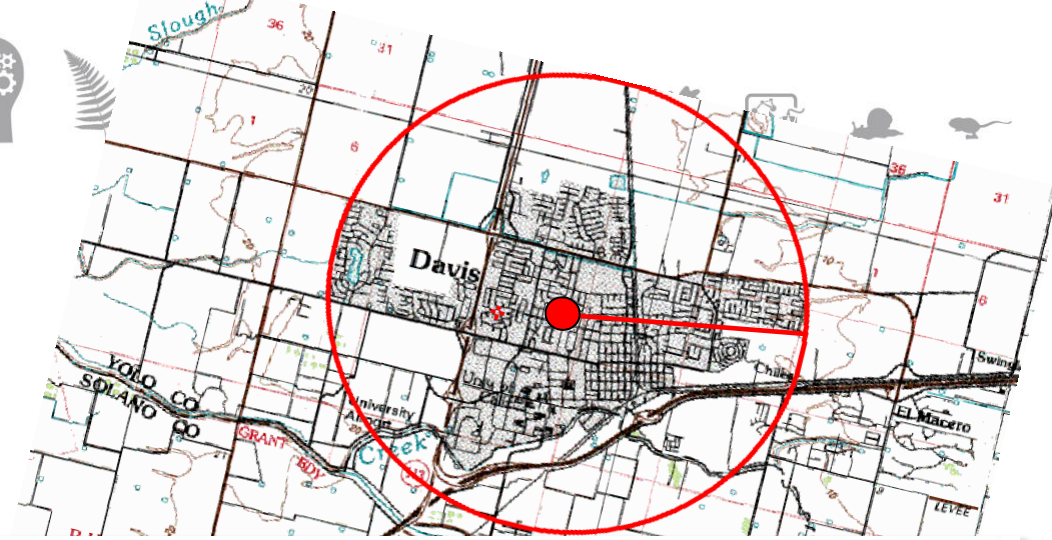
Common Datums

- NAD27 (North American Datum): system derived from land-based surveys, using Clarke 1886 ellipsoid
- NAD83: satellite-based system using the Earth's center as a reference point; eventually adopted as GRS80 (Geodetic Ref. System 1980)
- WGS84 (World Geodetic System 1984): mathematically refined GRS80 used by the US military and default for GPS
- For most uses, NAD83, GRS80, WGS84 are equivalent



Locality types – *help with steps in georeferencing*

- **Named Place**
 - Named Place
 - Urban Area
 - Small Town
 - Unbounded Area
 - Street Address
 - Junction
 - Between two named places
- **Offsets**
 - distance at a setting “by air”
 - distance along a path



GEOREFERENCING QUICK REFERENCE GUIDE

Version: 2012-10-08

John Wieczorek, David Bloom, Heather Constable, Janet Fang, Michelle Koo, Carol Spencer, Kristina Yamamoto

This is a practical guide for georeferencing using the point-radius method [1, 2, 3] using the Georeferencing Calculator [4, 5], maps, gazetteers, and other resources from which coordinates and spatial boundaries for places can be found. This guide is an update of “Georeferencing for Dummies” [6], and explains the recommended calculation procedure for localities encountered in the georeferencing process.

Georeferences using the methods in this guide will be maximally useful if as much information as possible is captured about and during the georeferencing process in the following fields defined in the Darwin Core standard [7]. For additional community discussion and recommendations, see the Darwin Core Project wiki [8].

Darwin Core Georeferencing terms:

- **decimalLatitude**, **decimalLongitude**, **geodeticDatum** – the combination of these three fields provide the reference for the center of the point-radius representation of



What is an ideal georeference?

A numerical description of a place that can be mapped and that describes the spatial extent of a locality and its associated uncertainties as well as possible.



Let's try one together

R. K. GODFREY HERBARIUM
202973

Asteraceae

PLANTS OF FLORIDA, USA

Helianthus heterophyllus Nuttall

WAKULLA CO.: St. Marks Nat'l Wildlife Refuge (Panacea Unit). Frequent in moist roadside depression, less so in drying sand of burned, open longleaf pine along W side Rte 372, just N of Rd 401 and 1.4 mi drive from Hwy 98.

2 October 2008

Loran C. Anderson no. **24,198** w/ Wilson Baker & Ann Johnson
R. K. Godfrey Herbarium (FSU)

Along W side Rte 372, just N of Rd 401 and 1.4 mi drive from Hwy 98.
Lat. 30.04494N, Long. 84.4823 W. 2 October 2008
Loran C. Anderson no. 24,198 w/ Wilson Baker & Ann Johnson
R. K. Godfrey Herbarium (FSU)



Introducing **GEO**Locate

- First
 - <http://bit.ly/geolocatestandard>

WAKULLA CO.: St. Marks Nat'l Wildlife Refuge (Panacea Unit). Frequent in moist roadside depression, less so in drying sand of burned, open longleaf pine along W side Rte 372, just N of Rd 401 and 1.4 mi drive from Hwy 98.



Let's try one together

R. K. GODFREY HERBARIUM
202973

Asteraceae

PLANTS OF FLORIDA, USA

Helianthus heterophyllus Nuttall

WAKULLA CO.: St. Marks Nat'l Wildlife Refuge (Panacea Unit). Frequent in moist roadside depression, less so in drying sand of burned, open longleaf pine along W side Rte 372, just N of Rd 401 and 1.4 mi drive from Hwy 98.
Lat. 30.04494N, Long. 84.45823 W.

2 October 2008

Loran C. Anderson no. **24,198** w/ Wilson Baker & Ann Johnson
R. K. Godfrey Herbarium (FSU)

Along W side Rte 372, just N of Rd 401 and 1.4 mi drive from Hwy 98.
Lat. 30.04494N, Long. 84.45823 W. 2 October 2008
Loran C. Anderson no. 24,198 w/ Wilson Baker & Ann Johnson
R. K. Godfrey Herbarium (FSU)



decimalLatitude
decimalLongitude
geodeticDatum
coordinateUncertaintyInMeters
georeferencedBy
georeferenceProtocol
georeferenceSources
georeferenceVerificationStatus
georeferenceRemarks
verbatimLocality*
locality*

include (at least)
coordinates
uncertainty
how you did the georeference
what sources you used
who did the georeferencing
any assumptions made
the geodetic datum, if known, or assumed
a text description of the locality
the verbatim coordinates if present

Asteraceae

PLANTS OF FLORIDA, USA

Helianthus heterophyllus Nuttall

WAKULLA CO.: St. Marks Nat'l Wildlife Refuge (Panacea Unit). Frequent in moist roadside depression, less so in drying sand of burned, open longleaf pine along W side Rte 372, just N of Rd 401 and 1.4 mi drive from Hwy 98.

2 October 2008

Loran C. Anderson no. 24,198 w/ Wilson Baker & Ann Johnson

R. K. Godfrey Herbarium (FSU)



Introducing **GEOLocate**

- Now
 - <http://bit.ly/geolocatecsv>
- See also
 - Google Maps
 - challenge
 - get coordinates, figure out how to measure



Asterina regularis, Verrill.
 QUEEN CHARLOTTE SOUND, NEW ZEALAND,
 10 fathoms, on mud.
 "Challenger" Expedition, Station 167 A.
 SLADEN COLLECTION. 263-'03.
221.



Let's try a few of yours

- Download the DemoCSV file
- Go to GEOLocate <http://bit.ly/geolocatecsv>
 - see link in wiki
- Our code 972598CD



Al. No. 150/1975 Lab. No. CM/75/36
 THE CITY MUSEUM, BRISTOL.
 LACERTA VIVIPARA ♂. Ad 598
 THE MINERIES, MR. PRIDDY, SOMERSET.
 28-8-1968. coll: D.J. FOXWELL.
 O.S. Ref. ST/544507.
 FREEZE DRIED. DECEMBER 1968. D.J. FOXWELL

Schistostega pennata
 Not uncommon, near Totnes. Harberton
 abundant about the junction of the Kingsbury
 & Harberton Roads, Totnes - 12 April 1920.

© 2009 Copyright Royal Albert Memorial Museum and Art Gallery and Exeter City Council

22/2014/27
 Species: *Arthonia anambrophila*
 (small)
 Location: Rushford Wood North end (4)
 Habitat: On an oak in open wood with
 grazed
 Het. japonica well developed on the
 same tree
 Grid 20/69864 90157
 Date 6.09
 Coll. B. Benfield

SECTION OF THE BARK OF A
 DWARF OAK
 (*Quercus Robur*, var. *pedunculata*)
 from the Ancient Grove at Wistman's
 Wood, Dartmoor.
 Deposited by Order of H. R. H. the
 Prince of Wales.

Ac. No. 1998/0024. Reg. No. Ad. 2584
 THE CITY MUSEUM, BRISTOL
 Eng. Name: JAY
 Species: *GARRULUS GLANDARIUS*
 Locality: DEAD ON STEPS. BLAISE CASTLE
 MUSEUM. HENBURY. AVON ST. 559783
 Date: 23-06-1988
 Sex: ♂ JUV



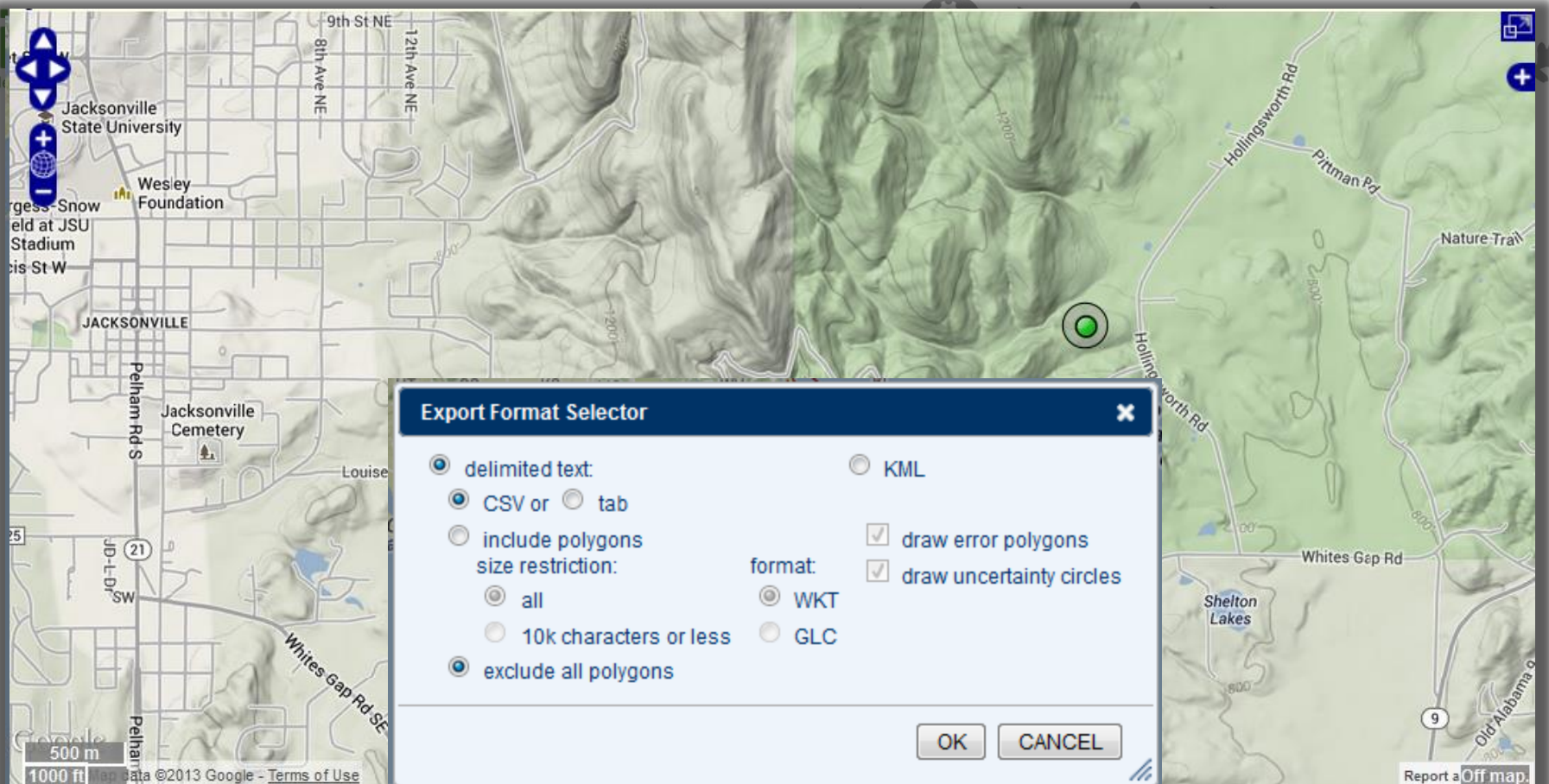
Georeference by CSV

GEOLocate Web Application

Workbench Results

Select File or load an existing file using a retrieval code: Load

Ex: 972598CD



Workbench 2 possible locations found

Show 8 records

Locality String	Country	State	County	WGSLat	WGSLon	Correction Status	Precision
Present in a moist mountain cove in Choccolocco Mtns. approx. 3 air mi e of Jacks	United States	AL	Calhoun	33.813599	-85.709001	yes	High(87)
St. Marks Nat'l Wildlife Refuge (Wakulla Unit). in shaded shallow water of wooded	United States	FL	Wakulla	30.10917	-84.25917	no	Low(43)
Frequent in rich sandy loam in mature hardwood forest on steep ravine slopes of W	United States	FL	Gadsden	30.4527205	-84.527205	no	High(87)

Search: File management SampleGEOlocateFileUncert.csv

Showing 1 to 3 of 3 records

Select an existing file using a retrieval code: 8AA9648E Load Export

File Management

Ex: 972598CD



Links to Cool Stuff

- iDigBio [GWG Listserv](#)
- [GWG](#) georeferencing help and working group
- [videos](#) (vimeo and idigbio)
- georeferencing.org
- Geo-rectify an old (online) map
 - <http://www.georeferencer.org/>



iDigBio Resource

Georeferencing

Note the contents of this wiki are under active revision and may change often.

Contents [hide]

- 1 Georeferencing
 - 2 Geographic Concepts Related to Georeferencing
 - 2.1 Geodetic datum
 - 2.2 Coordinate System
 - 2.3 Geographic Coordinate System
 - 2.4 Projection
 - 3 Georeference Best Practices
 - 4 Georeferencing Videos
- Community Protocols and Workflows

Georeferencing

Mapping Resource Hub

Home

Welcome to Georeferencing.org!

Here you will find a collection of resources pertaining to the process of georeferencing.

Tools

Gazetteers

Data Sources

Training



Internet Resources for Georeferencing – Part 1

- Google Maps: maps.google.com
- Hard to find localities: www.fallingrain.com
- Old & alt. names: bit.ly/Getty-TGN
- PLSS/TRS: www.earthpoint.us
- UTM map: www.dmap.co.uk/utmworld.htm
- UTM calculator: <http://www.earthpoint.us/Convert.aspx>
- Other tools: bit.ly/herpnet-georef-resources

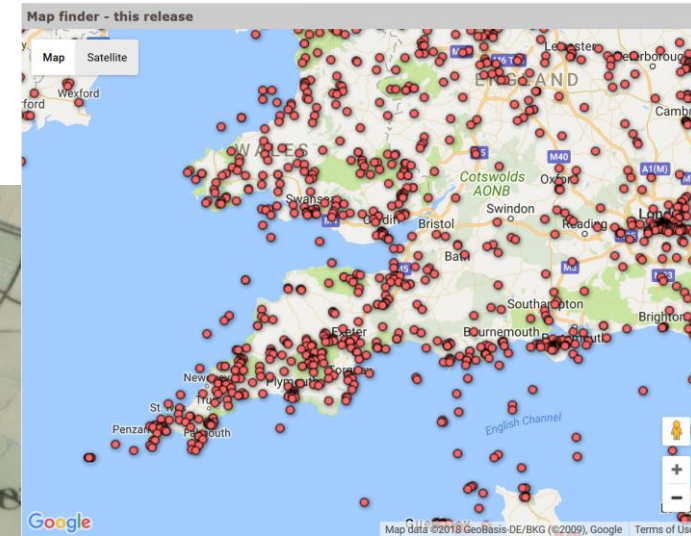
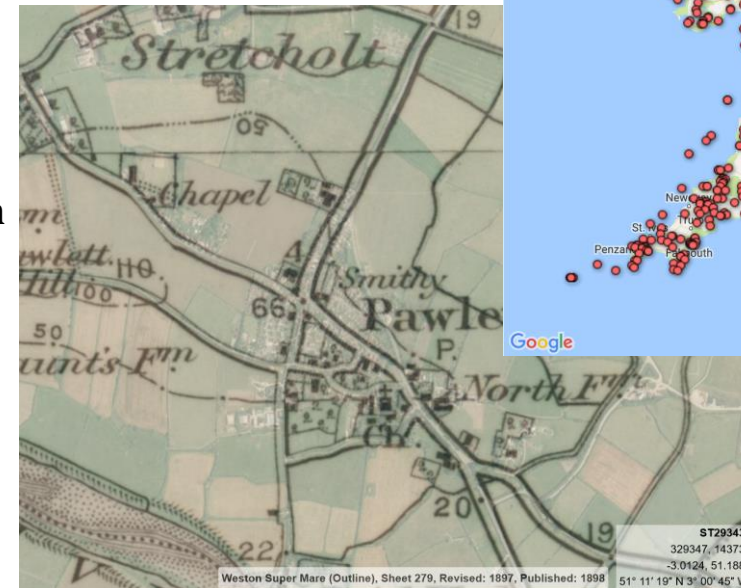
- See Internet Resources – powerpoint linked on this workshop WIKI

- Georeferencing Quick Guide
- Chapman, et al
- Your Protocol



Georeferencing resources with a UK focus

- Old Maps Online
 - <http://www.oldmapsonline.org/>
- British Library – maps online
 - <http://www.bl.uk/onlinegallery/onlineex/maps/index.html>
 - see Maps of the UK
 - <http://www.bl.uk/onlinegallery/onlineex/maps/uk/>
- British Library – georeferencing project
 - <https://www.bl.uk/projects/georeferencing>
 - <https://www.bl.uk/georeferencer/georefabout.html>
 - <https://www.bl.uk/georeferencer/georeferencingmap.html>
- Location London
 - <http://www.history.ac.uk/projects/research/locating-London>
- Historical GIS Research Network
 - <http://www.hgis.org.uk/resources.htm>
- National Library of Scotland
 - Georeferenced Maps <http://maps.nls.uk/geo/explore/>
 - very cool layers feature
- UK Grid Reference Finder
 - <http://gridreferencefinder.com/>





Ask the GWG! Thanks



www.idigbio.org



facebook.com/iDigBio



twitter.com/iDigBio



vimeo.com/idigbio



idigbio.org/rss-feed.xml



webcal://www.idigbio.org/events-calendar/export.ics

