

Core fields for georeferencing collections (compatibility with GIS). Ver 0.2

Georeferencing collections:

Production of spatial data (location data, e.g. where a species occurs) is a fundamental part of Kew's activities. This data can range from place names in a specimen database, localities on labels, field notes, or the exact location of a plant using GPS (global positioning systems). To allow links to specimen databases whether this is just for dot maps or full-blown GIS analysis, the follow standard fields will allow you to get the maximum from your data and ease the transfer of data to a GIS.

These are the fields that will describe the geographic (spatial) position of a collection. Also I have included several fields that will help to identify where the locality is (georeferencing) and help in the checking of this position.

I have described the technical details of these fields and given a suggested structure for the developers of databases (DB's). Also there is a section for users of the DB, which gives more details and highlights some problem cases.

These fields follow closely the HISPID3 standard, but incorporate other useful fields and the USBGN standards for locality types.

Technical details.

The table below gives the details of the fields, I have given the fields a priority from 1 to 3, where 1 is a core field, 2 is very useful for GIS work, 3 is not essential but can be useful. For more details see the section for DB users.

Priority	ID	Field Long Name	HISPID3 transfer code	Data Type	Notes	Other Notes/ Possibilities
1	1	Latitude Degrees	latdeg	byte <= 90	If unknown then fields to be blank/Null (i.e. not to default to zero)	Possible to lookup these entries from a internal Gaz. I.e. from place name 2 nd subdivision etc.
1	2	Latitude Minutes	latmin	byte < 60	If unknown then fields to be blank/Null (i.e. not to default to zero)	
1	3	Latitude Seconds	latsec	Single < 60	If unknown then fields to be blank/Null (i.e. not to default to zero)	
1	4	Latitude Direction	latdir	1 character Text E or W	Orientation! Direction must be filled in if any of above Degree, Minutes, Seconds (DMS) entries are inputted	
1	5	Longitude Degrees	londeg	byte <= 180	If unknown then fields to be blank/Null (i.e. not to default to zero)	
1	6	Longitude Minutes	lonmin	byte < 60	If unknown then fields to be blank/Null (i.e. not to default to zero)	
1	7	Longitude Seconds	lonsec	Single < 60	If unknown then fields to be blank/Null (i.e. not to default to zero)	
1	8	Longitude Direction	londir	1 character Text S or N	Orientation! Direction must be filled in if any of DMS entries are inputted	
1	9	Source of Lat/Long - (text) lookup	geosou	text	Lookup: <ul style="list-style-type: none"> • Label (default) • Map estimate • GPS (before 2000) • GPS (after 2000) 	

					<ul style="list-style-type: none"> • Gazetteer • Internal gazetteer • Literature • Others. - to be filled in i.e. from Tropicos DB etc. 	
3	10	Unsure?	--	Yes/No	Default to No. A Flag field for uncertain localities.	
3	11	Reference for Source of Lat/Long	--	Text (255)	Only need if above (9) is not label. Self lookup	
2	12	Accuracy of location in km	geoacy	Double	The precision of DMS.	N.B. HISPID3 is in Metres.
1	13	Locality Text	loc	Text/Memo	Locality description (usually as written on label, field book etc)	
1	14	Country (ISO)/TDWG	cou	Text	From iso/ and maybe TDWG list	Can lookup from each other. Only one of these needed
1	15	Country ISO	iso	text	Can lookup from cou	
1	16	TDWG	wgs	text	Can lookup from iso	
2	17	1 st Subdivision of country	pru	Text	Update - Self lookup	Possible to lookup from internal Gaz. USBGN!
2	18	2 nd Subdivision of country	sru	Text	Update - Self lookup	
3	19	Locality Name	--	Text	Place name or nearest place name - possible self lookup	
3	20	Within Nearest or actual?	--	Yes/No	With above to identify if locality is at place or near? Defaults to no.	
3	21	Locality Place type.	--	Text	Lookup (most popular) <ul style="list-style-type: none"> • LCTY - Locality (Default) • PPL - Populated place • PPK - Park • STM River\stream • MT- Peak\Mountian The USBGN lookup	Other code to be added from full USBGN code when needed.
2	22	Secondary source	--	Text	Lookup. <ul style="list-style-type: none"> • No (default) • Cultivated • Non-native • Introduced • Naturalised 	Not actually a GIS field, but is very useful. Probably better put with the specimen details.
1	23	Latitude Digital degrees	--	Double	Hidden - Updated on entry in locations	
1	24	Longitude Digital degrees	--	Double	Hidden - Updated on entry in locations	
1	25	Error diameter km	--	Double	Hidden - Updated on entry in locations. Fills in default values for accuracy if not entered.	

Further extensions

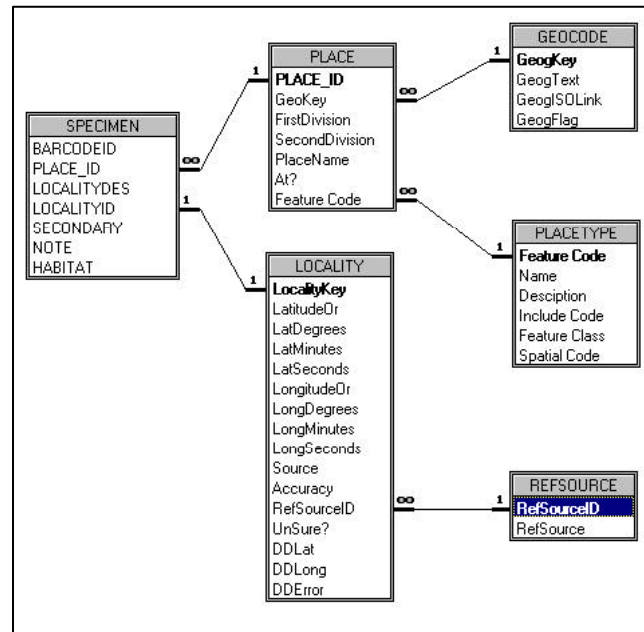
It is possible to describe features other than points, i.e. lines, polygons etc. Using this would however probably be too complex for most databases, but may be used by anyone wishing to produce a gazetteer., e.g: rivers give source and mouth locations; for a park give the two points that give the extent of the park NE corner, e.g. SW corner - etc.

- **Other fields useful to georeferencing locations but not needed for GIS.**

Habitat/Ecology - describes the physical and environmental characteristics of the locality. It is possible to split this up as in HISPID3, also can be derived from GIS (If so this needs to be noted)

Elevation/Altitude - again, could be simple (stick to metres) and have min & max, if only one elevation then use min. It will be assumed that source of the elevation will be the same as the locality data (occasionally this will be different). I feel that HISPID3 makes this field too difficult to enter (i.e. having source and accuracy for the both min and max, which will only very occasionally be different. Again this field can be derived from GIS (If so this needs to be noted).

Likely Structure for the GIS part of the database



Notes for users

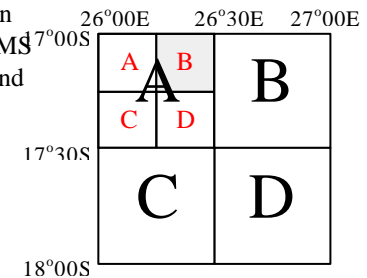
Latitude and Longitude. (1-8)

Purpose. Allows the collection locality to be georeferenced.

Details:

Problem cases: There are several usual types of georeferences that you may come across.

1. Locations described in degrees and digital minutes. Several GPS will default to degree and decimal minutes. Readings will look like $37^{\circ} 05.857'$ for example. The degrees and the minutes are correct, but to get the seconds you have to times the figures after the . by 60. So $37^{\circ} 05.857'$ is $37^{\circ} 05' 51.4''$. N.B. Sometime occurs in TROPICOS DB.
2. Locations described in decimal degrees. Again several GPS show this. The reading looks like 37.09761° for example. The degree is correct, but to get the minutes times all figures after the decimal minutes by 60 and round it down. Then take the remained and times by 60 again to get the seconds. So $37.09761 \Rightarrow 0.9761 \times 60 = 5.8566$ (5 minutes), and then $0.8566 \times 60 = 51.396$ (51.396 seconds) $\therefore 37^{\circ} 5' 51.396''$
3. Degrees of longitude over 180. Occasionally degrees of longitude are over 180 degrees, i.e. Hawaii can be described at 204° longitude East. To convert, just take the co-ordinate away from 360. i.e. $204 - 360 = 156^{\circ}$ and is West of 0° .
4. Local co-ordinate systems. e.g references in meters, UK Ordnance Survey co-ordinates, etc. It possible to convert these as long as the original co-ordinate system is known. If it is a one-off then it is probably easiest to check on a map, but for multiple entries contact Computing and GIS so that an automatic converter can be implemented.
5. Quadrat co-ordinates. Occasionally used (Africa mainly), where the degree square is split into four (a,b,c,d), this can then be split into four again. Co-ordinates look like 1725Ab. Which has a centre of $17^{\circ}7'30'' 25^{\circ}22'30''$. If possible convert to DMS using the centre point of the cell and record the source (9) as quadrat cell. Again if multiple entries contact Computing and GIS, so an automatic converter can be implemented.



Source of Latitude and Longitude (9)

Purpose: Identifies where the Latitude and Longitude came from, label, GPS etc. This allows default errors to be applied and also allows the original source to be identify for any checks, corrections etc.

Details: If entry is not in the default list just type it in e.g.: *From other database, Inferred from GPS, Quadrat cell.*

Unsure? (10)

Purpose: To identify problem Geo-references. This is a flag field (Yes/No), defaults to No.

Details: Tick this if you are unsure about the DMS typed in. Occasionally this may be used when a collection locality name occurs in several places in one country. e.g.: There are around 40 places called Ham in the UK, the most likely one has been chosen, but is could still be one of the others, so need more research or viewing within the GIS. Also label data may just give one name that is at several levels e.g. Rio do contas and Morro do chapau are both municipo and cities.

Reference for source of lat/long: (11)

Purpose: Not needed for label data, adds details of the actual source of the lat/long.

Details: This should be enough to allow you to find the original source. e.g. *Madagascar map series 1:500,000 FTM 1926, reprinted, 1986*. Or for GPS: the make, model, datum, etc. Most of the entries will be the same, so you only have to add a few at the beginning of the database, then just use lookup after that. This field is not essential, but can be very useful.

Accuracy of Location (12)

Purpose: Gives some idea as to how accurate the lat/long is. Default values will be used from Source of Latitude and Longitude (9) if nothing is entered.

Details: Enter the Diameter (in km) of the circle of error. Not needed if georeference is from label data. If it is from a paper map and you are sure of the location, just use the following table (This is assuming the that the map can not be read to less that 1 mm, if the map is more inaccurate then adjust appropriately).

1:5,000,000 1 mm = 5 km

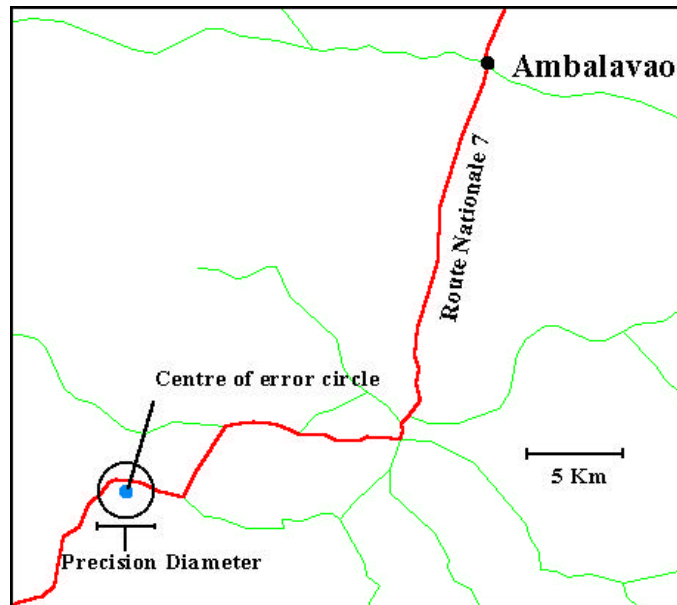
1:1,500,000 1mm = 1.5 km

1:1,000,000 1mm = 1 km

1:500,000 1mm = 0.5

1:250,000 1 mm = 0.25, etc.

Other examples: a park name or mountain range, the diameter of the circle should encompass the whole park. If looking up an uncertain co-ordinate on a map:



Precision	Label description
4 km	Fianarantsoa: 35 km South of Ambalavao on Route Nationale 7

Locality Text (13)

Details: Text as it is on the label, notebook etc. **Should this be exactly as label or should it be translated/inferred - I think it should be as on label and any translations and inferences should be in square brackets []**.

Country/ISO/TDWG (14 15 16)

Details: Only one of these needs to be entered. **Not sure what to do about old country names, etc? Suggests using square brackets [] for these country/area names.**

The three fields (17-19) below are not core, but are very useful for georeferencing the collections. Basically, if the data is easily at hand or known then it should be entered.

1st Subdivision of the country (17)

Purpose: Not core, but very useful for gazetteer work and for error checking.

Details: Enter the primary division each locality is within. First divisions may include: State, Province, County.

2nd Subdivision of the country (18)

Purpose: Again not core but very useful for gazetteer work and for error checking.

Details: If available enter the secondary division the locality is within. e.g. District, Municipio.

Locality Name (19)

Purpose: Again not core but very useful for gazetteer work and for error checking. Becomes very useful if the latitude and longitude are not known.

Details: Enter the nearest place name if available.

Within? (20)

Purpose: A flag value that is used as an Identify if the actual location is at/within the named locality. Only really needed for gazetteer work.

Details: Yes/No defaults to No. Only tick if the lat/long is the actual named locality.

Locality type: (21)

Purpose: Useful for gazetteer work

Details: Identifies the type of location. Enter the code. The most common entries are on the drop down list. Use lookup code from USBGN for more uncommon entries. Defaults to LCTY - locality.

Secondary Source (22)

Purpose: Not really a location field but is very useful.

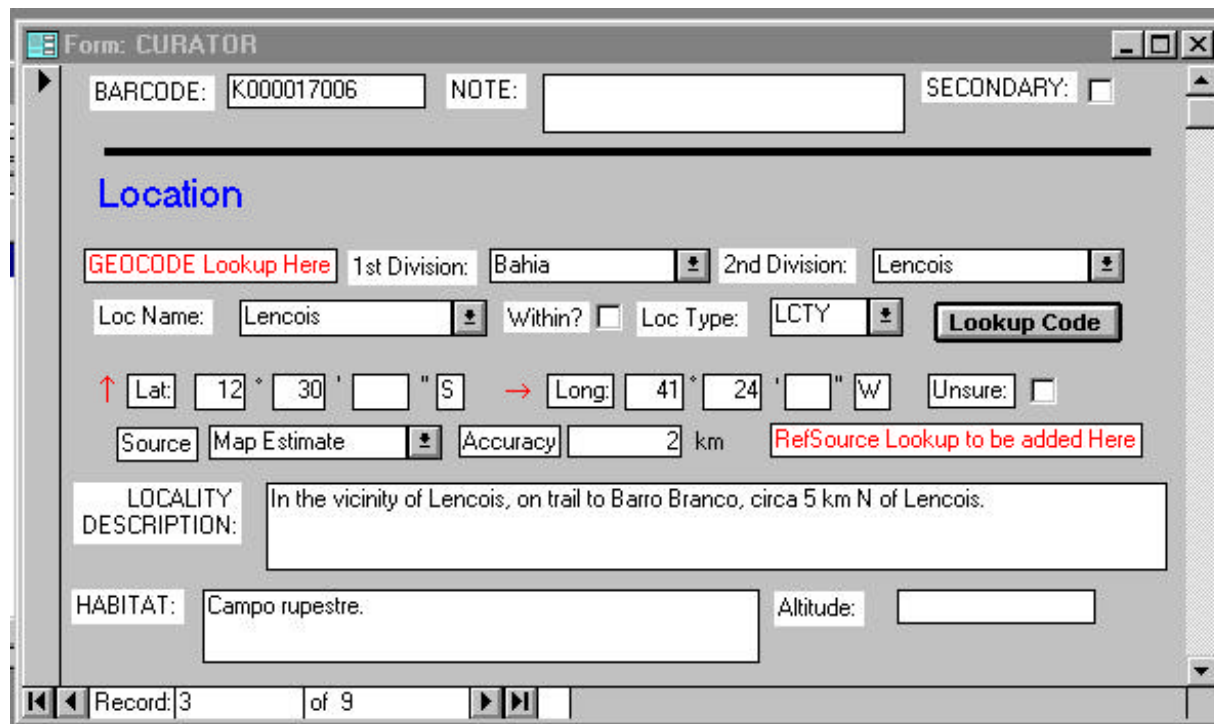
Details: Defaults to No. Only really needs an entry if the collection is known to be cultivated, non-native, introduced or naturalised.

References:

USBGN - United States Board on Geographic Names - <http://www.nima.mil/gns/html/BGN.html>

HISPID Version 3. Conn, B. J. (ed.) (1998). HISPID - Herbarium Information Standards and Protocols for Interchange of Data (National Herbarium of New South Wales: Sydney). - <http://www.rbgsyd.gov.au/HISCOM/>

Examples



Form: CURATOR

BARCODE: K000017006 NOTE: SECONDARY:

Location

[GEOCODE Lookup Here](#) 1st Division: Bahia 2nd Division: Lencois

Loc Name: Lencois Within? Loc Type: LCTY **Lookup Code**

↑ Lat: 12 ° 30 ' " S → Long: 41 ° 24 ' " W Unsure:

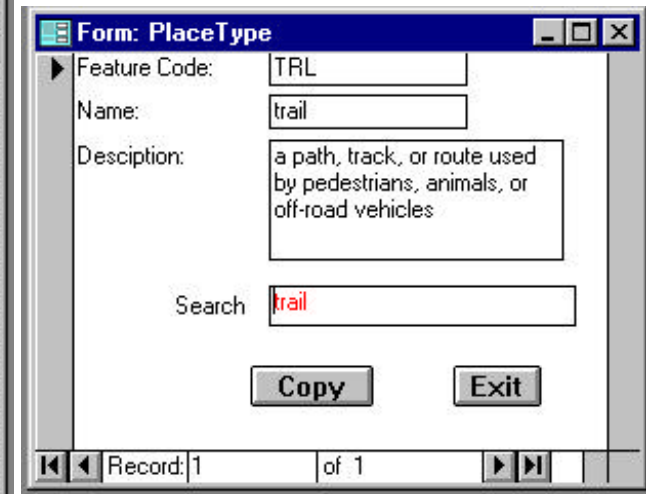
Source Map Estimate Accuracy 2 km [RefSource Lookup to be added Here](#)

LOCALITY DESCRIPTION: In the vicinity of Lencois, on trail to Barro Branco, circa 5 km N of Lencois.

HABITAT: Campo rupestre. Altitude:

Record: 3 of 9

Example DB, without Geocode and reference lookup.



Form: PlaceType

Feature Code: TRL

Name: trail

Description: a path, track, or route used by pedestrians, animals, or off-road vehicles

Search trail

Copy **Exit**

Record: 1 of 1

Lookup for Location type.