



Curso de formação:

# Gestão de coleções biológicas utilizando Specify 6

## Controlo de qualidade de dados geográficos

Rui Figueira

[ruifigueira@isa.ulisboa.pt](mailto:ruifigueira@isa.ulisboa.pt)

Nó Português do GBIF,  
Instituto Superior de Agronomia,  
Universidade de Lisboa



# Biogeomancer, Guide to Best Practices in Georeferencing

## Georeferencing Best Practices

Arthur D. Chapman · John R. Wieczorek – Version 8928428, 2023-07-12 09:02:13 UTC

This document is also available in [PDF format](#) and in other languages: [español](#).



Chapman AD & Wieczorek JR (2020)  
Georeferencing Best Practices.  
Copenhagen: GBIF Secretariat.  
<https://doi.org/10.15468/doc-gg7h-s853>

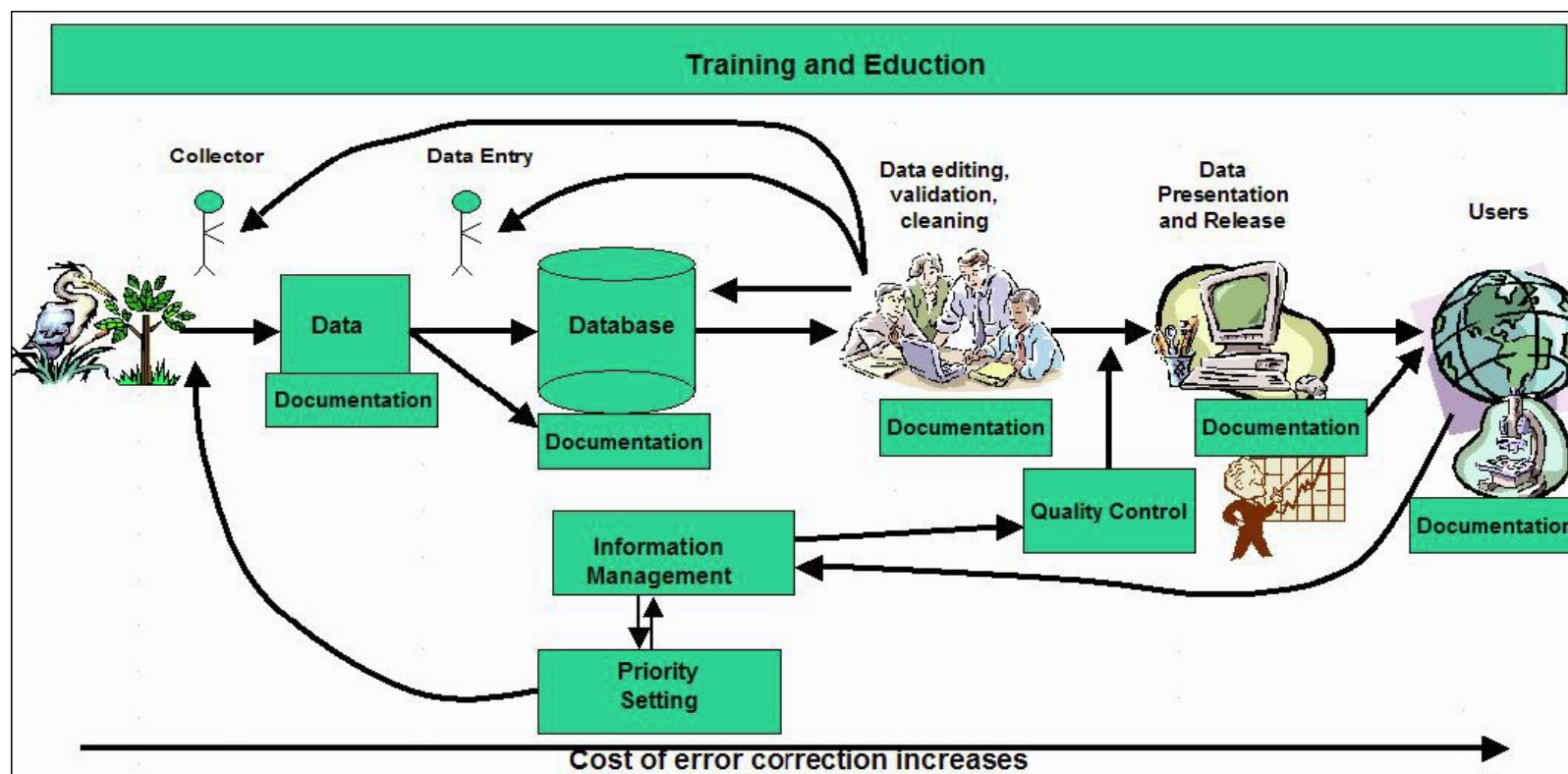
<https://doi.org/10.15468/doc-gg7h-s853>

# Exemplos de descrição do local

Provincia	Município	Locality
Huíla	Chibia	Jau, no cruzamento da estrada para a Hunguéria.
Huíla	Gambos	Gambos, Km 9 do trajecto Gambos Cahama
Huíla	Gambos	Chimbolelo
Huíla	Chibia	Quihita, Margens do rio Caculuvar
Huíla	Gambos	Entre Huíla e Palanca, em Muála
Huíla	Gambos	Chibemba, na picada Chimbolelo-Vimpona ao km 39
Huíla	Gambos	À 10 km au sud de Chibemba entre Sá da Bandeira et Roçadas.

# Princípios de qualidade de dados

A prevenção é o melhor remédio



# Princípios de qualidade de dados

## Recolher dados de campo

- fazer uma boa **descrição da localização**, ainda que se obtenham dados do GPS
- usar bons **pontos de referência**
- recolher **dados geográficos decimais**, com o maior número de casas decimais possível
- se é utilizado um GPS, registar sempre o **datum** respectivo
- registar altitudes com **altímetros barométricos**
- **registar as direções**, considerando ainda a declinação entre N magnético e N verdadeiro
- registar a **extensão da colheita**



# Princípios de qualidade de dados

## Informação a registar na georreferenciação

- Latitude decimal
- Longitude decimal
- Estimativa da incerteza máxima
- Unidade da incerteza máxima
- Verbatim das coordenadas
- Verbatim do sistema de coordenadas
- Estado de verificação da georreferenciação
- Validação da georreferenciação
- Protocolo de georreferenciação
- Fonte de georreferenciação
- Ajustamento espacial

# Princípios de qualidade de dados

## Índices toponímicos

<http://earth-info.nga.mil/gns/html/index.html>

<http://www.geonames.org/>

<http://manisnet.org/gc.html>

GeoLocate <http://www.geolocate.org>

# Procedimentos para atribuição de coordenada

## Determinar as coordenadas

Tipo de referência	Extensão	Atribuir coordenada
Nomes de locais	Distância do centro ao limite mais afastado	Coordenada do centro geográfico
Entre dois locais	Usar meia distância	Usar ponto médio
Distâncias relativas a um local		
A 10 km E (em linha recta) de...	Usar a extensão do local	Utilizar a coordenada da localidade como ponto de referência inicial
<b>A 10 km E por estrada de...</b>	Usar a extensão do local	<b>Idem, utilizando a estrada para medir distância</b>



# Procedimentos para atribuição de coordenada

## Determinar as coordenadas

Tipo de referência	Extensão	Atribuir coordenada
Distâncias relativas a um local		
A 10 km E de...	Usar a extensão do local	Assumir linha recta ou por estrada, etc., e documentar a decisão
<b>Margem direita do Rio ..., a 3 km a jusante de ...</b>		<b>Determina-se direita ou esquerda estando virado para o sentido da corrente</b>

# Procedimentos para atribuição de coordenada

## Determinar as coordenadas

Tipo de referência vagas	Extensão	Atribuir coordenada
Lisboa?		Não georreferenciar
Perto de...	Atribuir erro igual à maior distância entre esse local e fronteira com outro local	Usar bom senso. Utilizar coordenada do local
A 10 km de...	Usar a distância como erro	Utilizar a coordenada da localidade
A cerca de 10 km E de...		Processar como se fosse "A 10 km E de..."
A 10 km de Lisboa, 1500 m altitude		Inconsistência. Tentar obter mais informação

# Procedimentos para atribuição de coordenada

## Determinar as coordenadas

Tipo de referência vagas	Extensão	Atribuir coordenada
Em Santarém, distrito de Lisboa		Inconsistência. Tentar obter mais informação
<b>No rio..., na estrada...</b>	Desde o ponto ao extremo mais afastado	<b>Usar ponto médio,. Se não for possível seguir a linha, traçar linha recta do início até ao fim, e encontrar ponto no rio ou estrada mais próximo deste</b>

---

# Incerteza da georreferenciação

## Fontes de incerteza

- a extensão da localidade
- desconhecimento do datum
- imprecisão da determinação de distâncias
- imprecisão da determinação da direção
- imprecisão da medição da coordenada
- escala da carta.

# Incerteza da georreferenciação

## Incerteza dependente do datum

do Datum	Região	para o Datum	Diferença
GDA 94	Australia	WGS84	Max±<1m
NAD 1983	North America	WGS84	Max±<1m
NAD 27	North America	WGS84	Max±200m
NAD 27	Contiguous USA	WGS84	Max±105m
NAD 27	Aleutian Islands,Alaska	WGS84	Max±235m
NAD 27	Hawaii	WGS84	~500m
TOKYO	Japan	WGS84	Max±750m
ED-50	Europe	WGS84	Max±175m
ARC-50	Africa	WGS84	Max±265m
INDIAN 1975	Bangkok,Thailand	WGS84	~405m
INDIAN 1956	Delhi,India	WGS84	~135m
HONG KONG 1973	HongKong	WGS84	~320m
KERTAU 1948	Singapore	WGS84	~190m

# Incerteza da georreferenciação

## Incerteza dependente da precisão da coordenada

Precisão	unidade	Latitude			
		0	30	60	85
1.0	degree	156,904 m	146,962 m	124,605 m	112,109 m
0.1	degree	15,691 m	14,697 m	12,461 m	11,211 m
0.01	degree	1,57 m	1,47 m	1,246 m	1,121 m
0.001	degree	157 m	147 m	125 m	112 m
0.0001	degree	16 m	15 m	13 m	12 m
0.00001	degree	2 m	2 m	2 m	2 m
1.0	minute	2,615 m	2,45 m	2,077 m	1,869 m
0.1	minute	262 m	245 m	208 m	187 m
0.01	minute	27 m	25 m	21 m	19 m
0.001	minute	3 m	3 m	3 m	2 m
1.0	second	44 m	41 m	35 m	32 m
0.1	second	5 m	5 m	4 m	4 m
0.01	second	1 m	1 m	1 m	1 m

# Incerteza da georreferenciação

## Incerteza dependente da leitura de uma carta

Incerteza horizontal na leitura de uma carta com um traço de 0,5 mm

Escala	Incerteza (m)
1:1000	0,5
1:10000	5
1:25000	12,5
1:50000	25
1:100000	50
1:250000	125
1:1000000	500



# Documentação da georreferenciação

## Garantir a qualidade dos dados

### *Dar feedback aos coletores*

- Garantir que o datum é registado em conjunto com as coordenadas por GPS
- Promover a utilização de um sistema de coordenadas padrão (e.g., encorajar a utilização de coordenadas geográficas decimais, sempre que possível)
- Registrar as localidades de forma consistente e clara (mesmo quando se registam coordenadas GPS)
- Utilizar o nome de uma localidade e distâncias à mesma
- Registrar 'por estrada' ou 'por ar'
- Utilizar o altímetro barométrico

# Controlo de qualidade da georreferenciação

## Garantir a qualidade dos dados

*Receber feedback dos utilizadores*

*Realizar operações de limpeza e correção de dados*

Sobre a entrada de dados:

- Os 200 registos iniciais devem ser verificados. Havendo erros, manter a verificação em grupos de 100
- Verificar regularmente 10 registos aleatórios por cada 100 georreferenciados
- Se existirem mais de 2 registos incorretos, devem ser verificados adicionalmente 20 registos, e pedir ao operador para confirmar os 100.
- Após a regularização das taxas de erro, manter a verificação de 5 em cada 100

# Controlo de qualidade da georreferenciação

## Garantir a qualidade dos dados

### *Realizar operações de limpeza e correção de dados*

Validação de dados:

- utilizar fontes externas de dados (caderno de campo e itinerários, gazetteers, etc.),
- verificar contra outros campos da base de dados (coincidência dos níveis administrativos),
- utilizar o SIG para verificar registos colocados fora das áreas administrativas
- utilizar ferramentas estatísticas de data mining
- utilizar ferramentas de modelação ecológica e análise estatística para identificar outliers colocados fora das regiões ambientais (e.g. clima)

# Controlo de qualidade da georreferenciação

## Garantir a qualidade dos dados

### *Realizar as correções*

Corrigir por adição, e não por substituição

### *Documentar*

- como foi realizada a georreferenciação,
- que método foi aplicado para determinar a extensão e o erro,
- que modificações foram realizadas,
- as ações de validação de dados,
- utilizar flags para indicar incerteza, etc.
- o nível geral de qualidade dos dados,
- as unidades de medida e outros padrões adoptados

# Gestão de coleções biológicas utilizando Specify 6

## Obrigado pela atenção

Nó Português do GBIF  
Instituto Superior de Agronomia  
Herbário  
Tapada da Ajuda  
1349-017 Lisboa, Portugal

Tel: (+351) 213653165  
email: [node@gbif.pt](mailto:node@gbif.pt)  
<http://www.gbif.pt>

O Nó Português é acolhido no ISA com o apoio da FCT.



INSTITUTO  
SUPERIOR DE  
AGRONOMIA  
*Universidade de Lisboa*



Fundação  
para a Ciência  
e a Tecnologia

Esta apresentação é publicada  
segundo a licença CC-BY-SA

