Gbif.es

Apparent Quality Index (ICA) in Darwin Test

Spanish Node of the Global Biodiversity Information Facility

Coordination Unit - GBIF Spain

Year 2023

MINISTERIO DE CIENCIA E INNOVACIÓN

CSIC
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS

Apparent Quality Index

The **Apparent Quality Index** is an indicator of the data base quality (records, observations, specimens of natural history collections) in Darwin Core format. One of the possibilities of the software application **Darwin Test** is to calculate this index. The index is divided in three components: Taxonomic, Geographic and Temporal. In the GBIF Spain framework, we use the ICA to monitor the quality of the biodiversity data published in the GBIF network. We refer to this index as an "apparent index" due to the fact that –although unlikely– theoretically, a dataset can have a maximum ICA while its records can be a poor reflection of the reality to which they refer.

- When Darwin Test calculates the **Taxonomic component**, it does so on the basis of the similarities between the scientific name and the taxonomic hierarchy, between the dataset and the data base in "Catalogue of Life". It is possible to use other data bases of scientific names such as "Archivos de Autoridad Taxonómica" from SiB Colombia or others developed for specific contexts.

- The **Geographic component** values the number of records with coordinates and the percentage of cells that are consistent with the fields "Country" and "stateProvince". In the implementation of ICA in Darwin Test the checking against "StateProvince" is optimized for the Spanish territory. Its adaptation to other areas involves to include in Darwin Test the boundary coordinates of the units "StateProvince" from the territory of interest.

- Finally, the **Temporal component** is determined by the number of right records in the date fields.

Below you can find how to calculate the ICA Index.

How the ICA is calculated

45 % Taxonomic component			
20 %	% with genus in Species2000		A
10 %	% with species		B
9 %	% with taxonomic hierarchy		
	3 %	kingdom	C
	3 %	class or order	D
	3 %	family	E
6 %	% with identified by		F
35 % Geographic Component			
20 %	% with coordinates		G
10 %	% with countries		H
5 %	% with pointradius		I
- 20 %	% with incorrect coordinates		J
20 % Temporal Component			
11%	% with correct year		K
7 %	% with correct month		L
2 %	% with correct day		M
- 15 %	% with incorrect dates		N

Formula Breakdown:

45% componente Taxonómico / Taxonomic component [$ICAt$]

20 %	% con géneros en Species2000 / % with genus in Species2000		A
10 %	% con especies / % with species		B
9 %	% con jerarquía OK / % with rank OK		
	3%	reino / kingdom	C
	3%	clase u orden class or order	D
	3%	familia / family	E
6%	6 % with identified by		F

35% componente Geográfico / Geographic component [$ICAg$]

20%	% con coordenadas OK / % with coordinates OK	G
10%	% con países OK / % with countries OK	H
5%	% con radiopunto OK / % with pointradius OK	I
- 20%	% coordenadas no OK / % coordinates not OK	J

20% componente Temporal / Temporal component [$ICAd$]

11 %	% año OK / % year OK	K
7 %	% mes OK / % month OK	L
2 %	% día OK / % day OK	M
-15%	% fechas no OK / % dates not OK	N

$$ICAt = (20 \times A) + (10 \times B) + (3 \times C) + (3 \times D) + (3 \times E) + (6 \times F)$$

$$ICAg = (20 \times G) + (10 \times H) + (5 \times I) + (20 \times J)$$

$$ICAd = (11 \times K) + (7 \times L) + (2 \times M) + (15 \times M)$$

$$ICA = ICAt + ICAg + ICAd$$

Spanish Node of the Global Biodiversity Information Facility

Coordination Unit - GBIF Spain

gbif.es

datos.gbif.es

Year 2023

C/ Joaquín Costa, 22

28002 Madrid – Spain

Tel: +34 91 568 00 37

E-mail: info@gbif.es

MINISTERIO DE CIENCIA E INNOVACIÓN

CSIC
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS