

# Manual de usuario

para la estandarización y carga de nuevos registros

---

Lamaro, Anabel A.; Cavanagh, Federico E. & Sala, Silvia E.

Diseño de un Sistema de Información Geográfica enfocado a las floraciones de cianobacterias en el área de Berisso y Ensenada.

*Proyecto FITBA-2022-B21*

Cómo citar este documento:

Lamaro, A. A.; Cavanagh, F. E. & Sala, S. E. 2024. Manual de usuario para la estandarización y carga de nuevos registros. Proyecto FITBA-2022-B21: “Diseño de un Sistema de Información Geográfica enfocado a las floraciones de cianobacterias en el área de Berisso y Ensenada”. 23 pp

Licencia de uso:

Este trabajo tiene licencia CC BY-NC-SA 4.0. Para ver una copia de esta licencia, visite <http://creativecommons.org/licenses/by-nc-sa/4.0/>



## Contenido

Introducción.....	3
Uso del presente manual .....	5
Sección I Diseño de las capas de información .....	6
Introducción.....	6
Diseño general .....	6
Configuración regional .....	6
Campos alfabéticos.....	6
Campos numéricos .....	7
Campos con fechas.....	7
Capas de información .....	7
Presencia de taxones.....	7
Datos ambientales, biológicos y otra información relacionada al evento .....	12
Vinculación entre los datos en diferentes capas de información. ....	14
Sección II Diseño de códigos .....	14
Códigos para taxones .....	14
Criterios para tener en cuenta en la construcción de los códigos.....	16
Códigos para sitios de muestreo.....	16
Códigos para fechas.....	17
Anexo I Tesoros geográficos.....	18
Nomenclador de Argentina.....	18
Geonames .....	18
Getty Thesaurus of Geographic Names .....	18
InfoXY .....	18
geoBoundaries .....	18
Anexo II Conversores de coordenadas .....	19
epsg.io .....	19
Coordinate Parsing from Canadensys.....	19
Earth Point Coordinate Converter .....	19
Anexo III Georreferenciación retrospectiva .....	19
Anexo IV Estandarización de fechas .....	19
Date Parsing.....	19
Anexo V Limpieza de datos utilizando Excel.....	20
Algunas funciones de utilidad .....	21
Bibliografía citada .....	22

## Índice de figuras

Figura 1 Diseño del identificador único para la planilla de taxones .....	7
Figura 2. Ejemplo de construcción de ID para la presencia de taxones.....	7
Figura 3. Referencias de abundancias relativas comúnmente utilizadas .....	11

Figura 4. Diseño del identificador único para la planilla de datos ambientales .....	12
Figura 5. Ejemplo de construcción de ID para los datos ambientales .....	12
Figura 6. Codificación para los niveles de alerta de la OMS.....	13
Figura 7. Vinculación entre datos de diferentes capas de información.....	14
Figura 8. Estandarización masiva de fechas mediante el conversor Data Parsing de Canadensys..	20
Figura 9. Configuración de formato de celdas en Excel.. .....	22

## Índice de tablas

Tabla 1. Lineamientos para la construcción de los códigos de taxones con identificación completa.....	15
Tabla 2. Lineamientos para la construcción de los códigos de taxones con identificación incompleta.....	15

## Introducción

El proyecto “Diseño de un Sistema de Información Geográfica enfocado a las floraciones de cianobacterias en el área de Berisso y Ensenada” se enmarca en la primera convocatoria del Fondo de Innovación Tecnológica de la provincia de Buenos Aires (FITBA) y fue aprobado mediante Resolución RESO-2022-647-GDEBA-MPCEITGP. Uno de sus principales objetivos es recopilar la información histórica y actual sobre cianobacterias en el Río de La Plata, (principalmente en las costas de Ensenada y Berisso). La incorporación de esta información a una base de datos geográfica y estandarizada permitirá analizar los datos de modo integral junto con otras capas de información, constituyendo una línea de base para rastrear los cambios geográficos y temporales ocurridos y establecer patrones de comportamiento relacionados a cambios naturales o antrópicos. No solo ayudará a comprender y gestionar el problema de las floraciones nocivas en la costa de Berisso y Ensenada, sino que representa la base desde la cual la Subsecretaría de Recursos Hídricos podrá incorporar y gestionar la información sobre los cuerpos de agua de la Provincia de Buenos Aires. Se plantea además como una herramienta de consulta para relevar aspectos sanitarios y epidemiológicos sobre las afecciones ocasionadas por las cianotoxinas, calidad del agua potable, uso turístico de los cuerpos de agua, entre otros aspectos.

## Uso del presente manual

Este documento pretende ser una guía que sea de utilidad para estandarizar la carga de nuevos registros a la base de datos. En función de las características de la información, se diseñaron dos capas de información principales, una basada en la presencia de cianobacterias (mencionada como [Presencia de taxones](#)) y la otra relacionada a [datos ambientales y biológicos](#) del sitio.

## Sección I Diseño de las capas de información

### Introducción

Los campos de la base de datos que se describen en este manual, fueron establecidos en base a los utilizados y recomendados por la *Harmful Algal Event Database*- HAEDAT (<http://haedat.iode.org/>) componente del Sistema de Información de Algas Nocivas -*Harmful Algal Information System*- HAIS (<https://data.hais.ioc-unesco.org/>) de la *Intergovernmental Oceanographic Commission* IOC de la UNESCO (<https://www.ioc.unesco.org>). La HAEDAT es una metabase de datos que contiene registros de episodios de algas nocivas, recopilando información desde 1985 en el Atlántico Norte y desde 2000 en el Pacífico Norte. Las redes regionales de la COI en América del Sur, Pacífico Sur y Asia, y África del Norte están en proceso de contribución al sistema.

Para el diseño de los campos se tuvieron en cuenta también los datos que se toman de manera rutinaria en las campañas locales, intentando siempre incorporar toda la información posible. Las indicaciones desarrolladas a continuación son sugerencias que permitirán mantener un formato estandarizado.

### Diseño general

#### Configuración regional

En informática, la internacionalización en el diseño de los *software* permite su adaptación a diferentes idiomas y regiones sin necesidad de realizar cambios de ingeniería o de código. En este sentido, la configuración regional en informática, se refiere a un conjunto de parámetros que define el idioma, el país y otras preferencias especiales que el usuario utilice en la interfaz. En este sentido es importante indicar la configuración de los diferentes formatos de texto, fechas, números, etc. a fin de contar con la información estandarizada que permita la lectura de los datos de manera adecuada en los diferentes *softwares*. Estas recomendaciones se realizarán en este manual dentro de la descripción de cada campo o dato.

#### Campos alfabéticos

Se sugiere que los campos de texto se completen con palabras en minúscula, a excepción de los nombres propios. Por ejemplo: enero, primavera, Playa Bagliardi.

## Campos numéricos

Se sugiere que los campos con datos numéricos se completen de modo que el punto (.) sea separador de decimales y la coma (,) sea separador de miles. Por ejemplo:

- 34.585 se interpreta como el número decimal: treinta cuatro con quinientos ochenta y cinco.
- 34,585 se interpreta como treinta cuatro mil quinientos ochenta y cinco.

Esta configuración es fundamental para la incorporación de los datos al SIG, principalmente para la lectura de coordenadas.

## Campos con fechas

Se sugiere que para las fechas se utilice la configuración dd-mm-aaaa utilizando guión medio entre cada parte (ver [Campos relacionados a la fecha de muestreo](#))

## Capas de información

### Presencia de taxones

#### Diseño del ID

El **identificador único** denominado Id caracteriza inequívocamente cualquier registro individual de la base. Está diseñado en función de tres variables: la presencia del taxón, el sitio del muestreo (dimensión espacial) y la fecha del muestreo (dimensión temporal). Se construye mediante la combinación de los códigos que correspondan a cada una (Figs. 1 y 2).

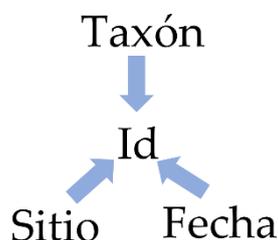


Figura 1 Diseño del identificador único para la planilla de taxones

Taxón	Código Taxón	Sitio	Código Sitio	Fecha	Código Fecha	ID presencia taxones
<i>Microcystis aeruginosa</i>	MAER	Balneario La Pérgola	BLPE	5/01/2021	05012021	MAERBLPE05012021
<i>Microcystis flos-aquae</i>	MFLO	Balneario La Pérgola	BLPE	5/01/2021	05012021	MFLOBLPE05012021
<i>Planktothrix aff. agardhii</i>	PAAGx	Parador 5	PCIN	5/01/2021	05012021	PAAGxPCIN05012021

Figura 2. Ejemplo de construcción de ID para la presencia de taxones

### *Campos relacionados a los taxones*

**Taxón citado en informe:** se incluye género y especie, variedad y/o forma, según corresponda, tal como se indica en el informe o publicación de origen. Por ejemplo: *Aphanocapsa delicatissima*.

**Nombre taxón según AlgaeBase (2023):** durante el proceso de carga de información en la base de datos se verificó la validez del nombre científico del taxón en la base de datos global AlgaeBase (<https://www.algaebase.org/>). En el caso de que haya habido un cambio de nombre, la nueva denominación se incorpora en este campo y el código del taxón se construye con el nombre actualmente vigente. Este proceso de estandarización es de suma importancia para optimizar el proceso de análisis posteriores. En este sentido, por ejemplo, esto permitirá extraer de la base de datos la presencia de un taxón a lo largo de un período de tiempo determinado, directamente mediante el código de taxón, más allá de los cambios taxonómicos ocurridos.

**Código taxón:** los lineamientos para establecer cada uno se detallan en la Sección II: Diseño de códigos, [Código para taxones](#).

**Phylum:** categoría taxonómica correspondiente según lo indicado en *AlgaeBase*.

**Clase:** categoría taxonómica correspondiente según lo indicado en *AlgaeBase*.

**Orden:** categoría taxonómica correspondiente según lo indicado en *AlgaeBase*.

**Familia:** categoría taxonómica correspondiente según lo indicado en *AlgaeBase*.

**Taxa nocivas y Tipo de cianotoxinas citadas:** esta información fue recopilada por el Dr. Ricardo Echenique integrante del proyecto e investigador especialista en Cyanobacteria. La información está actualizada a abril de 2024 y se indica con una X en el caso de que ese taxón esté citado como toxígeno y el tipo de cianotoxina relacionada. En caso contrario se indica como “Sin referencia”.

### *Campos relacionados al sitio de muestreo*

**Nombre Sitio Literal (*verbatim*):** se recomienda no perder la información original de los nombres de los sitios tal como fueron registrados originalmente. Es posible que un mismo sitio esté registrado de dos o más formas diferentes, por ejemplo, el Balneario La Pérgola puede estar registrado como: B. La Pérgola, Baln. La Pérgola, La Pergola entre otras formas. En el campo “sitio de muestreo literal” se debe indicar entonces el nombre del sitio donde la muestra fue extraída, tal como está indicado en el informe o publicación.

**Sitio:** se recomienda completar este campo con el nombre del sitio completo/adecuado/aceptado. Si fuera un sitio geográfico reconocido se recomienda buscar el nombre en tesauros geográficos (ver [Anexo I](#)). Por ejemplo: Balneario La Pérgola. A partir de este nombre se genera el Código de sitio explicado a continuación

**Código sitio:** se utiliza un código construido por 4 letras mayúsculas, de modo que represente siempre el mismo sitio (ver [Códigos para sitios de muestreo](#)).

### NOTA

*En el caso en que el sitio de muestreo esté mencionado de modo correcto y completo, se completa ese campo y el campo Sitio de muestreo literal puede quedar sin completar.*

#### *Campos relacionados a división administrativa*

Para completar esta información, se recomienda el uso de la información oficial de la República Argentina, publicada en el Portal Nacional de Datos Abiertos (<https://datos.gob.ar/>). Este es el punto de acceso para buscar y acceder fácilmente a los datos que publican los organismos de la Administración Pública Nacional.

**Municipio:** entidades que representan la división político-administrativa de tercer orden de la República Argentina, en base a datos del Instituto Geográfico Nacional (IGN) y de otros organismos. Incluye comunas, juntas vecinales y demás formas de gobierno local.

Acceso a descarga de archivo con formato .csv [https://datos.gob.ar/ar/dataset/jgm-servicio-normalizacion-datos-geograficos/archivo/jgm\\_8.11](https://datos.gob.ar/ar/dataset/jgm-servicio-normalizacion-datos-geograficos/archivo/jgm_8.11)

**Partido:** entidades que representan la división político-administrativa de segundo orden de la República Argentina, en base a datos del IGN y de otros organismos.

Acceso a descarga de archivo con formato .csv [https://datos.gob.ar/ar/dataset/jgm-servicio-normalizacion-datos-geograficos/archivo/jgm\\_8.10](https://datos.gob.ar/ar/dataset/jgm-servicio-normalizacion-datos-geograficos/archivo/jgm_8.10)

**Provincia:** entidades que representan la división político territorial de primer orden de la República Argentina en base a datos del IGN. Incluye las 23 provincias y la Ciudad Autónoma de Buenos Aires (CABA).

Acceso a descarga de archivo con formato .csv [https://datos.gob.ar/ar/dataset/jgm-servicio-normalizacion-datos-geograficos/archivo/jgm\\_8.9](https://datos.gob.ar/ar/dataset/jgm-servicio-normalizacion-datos-geograficos/archivo/jgm_8.9)

### *Campos relacionados a la ubicación geográfica*

**Latitud:** expresada en formato decimal, con el punto (.) como separador de decimales (ver [Campos Numéricos](#)), considerando al menos 4 decimales y el signo menos indicando la latitud Sur. Por ejemplo: -34.8203.

**Longitud:** expresada en formato decimal, con el punto (.) como separador de decimales (ver [Campos Numéricos](#)), considerando al menos 4 decimales y el signo menos indicando la longitud Oeste. Por ejemplo: -57.9636.

En el caso de contar con coordenadas en formato sexagesimal se recomienda su conversión a formato decimal (ver [Anexo II](#)). Es importante considerar que un mismo sitio de muestreo debe tener las mismas coordenadas, de modo de facilitar las búsquedas y análisis posteriores.

Si solo se cuenta con información relacionada al nombre y/o descripción del sitio, sin información geográfica, las coordenadas deben obtenerse mediante una georreferenciación retrospectiva (ver [Anexo III](#)).

### *Caracterización del sitio*

**Uso del agua:** En caso de contar con la información se incluye el dato del uso del agua registrado para ese sitio, por ejemplo: recreativo, recreativo con inmersión, consumo humano.

### *Campos relacionados a la fecha del muestreo*

**Fecha:** se indica con día, mes y año, siguiendo el formato: *dd-mm-aaaa*, por ejemplo 01-12-2001 ([Campos con fechas](#)). En el caso de no contar con la información completa o ambigua, se sugieren los siguientes modos de carga:

“...el muestreo se realizó en junio del año 2022...”, se carga como: 06-2022

“...el muestreo se realizó entre 2001 y 2003...”, se carga como: 2001/2003.

“... el muestreo se realizó entre el 13 y el 15 de noviembre de 2007...”, se carga como: 13/15-11-2007.

**Mes:** indicar el mes con letras en minúscula, por ejemplo: enero

**Año:** indicar el año en números completos, por ejemplo: 2007

En el [Anexo IV](#) se muestra en detalle el uso de herramientas para la estandarización automática de fechas.

**Código fecha:** se utiliza un código construido por los números que representen la fecha. Los lineamientos para establecer se detallan en la Sección II: Diseño de códigos, [Códigos para fechas](#).

**Campaña:** se indica la estación del año en la cual se realizó el muestreo.

#### *Campos relacionados a resultados del análisis de la muestra*

En estos campos se incluyen los resultados de los conteos en las muestras cuantitativas del taxón registrado:

**Abundancia relativa:** es muchas ocasiones se utiliza la abundancia relativa con base en una escala cualitativa del tipo:

Código	Abundancia
R	Rara
E	Escasa
F	Frecuente
A	Abundante
MA	Muy abundante
P	Predominante

*Figura 3. Referencias de abundancias relativas comúnmente utilizadas*

En este caso se puede incorporar la letra indicativa en mayúscula (Ej: R) o toda la palabra en minúscula (Ej: rara). Se recomienda mantener un mismo criterio para todos los registros.

En el caso de que los datos no tengan la escala ejemplificada en la Fig. 3, se recomienda establecer un criterio de pertenencia a alguna de ellas, según se considere adecuado. Por ejemplo, si la abundancia relativa se expresa como “...mucho...” podría interpretarse como que pertenece a la categoría A de abundante.

**Abundancia absoluta:** se carga el valor numérico registrado (ver [Campos numéricos](#)). Es importante que las unidades de estos conteos estén estandarizadas de modo de poder realizar adecuadamente el análisis posterior de los datos. En muchos casos la unidad utilizada es células/mililitro (cel./ml.), pero en el caso de taxones coloniales los datos pueden registrarse como individuos/mililitro (ind./ml.) o como colonias/mililitro (col./ml.). Cualquiera sea el caso, cada una de las formas de indicar la abundancia, se indica en tres campos separados: Abundancia absoluta (col./ml.), Abundancia absoluta (cel./ml.), Abundancia absoluta (ind./ml.)

**Fuente de datos (Publicación/Informe):** se indica la referencia de la fuente de los datos. En el caso de trabajos publicados e informes, los documentos digitales se entregan junto con la base de datos. Se sugiere luego incorporar un enlace que permita acceder al informe de origen de los datos, preferentemente dirigido a un repositorio oficial de datos.

### Otros campos factibles de incorporar

**Observaciones:** datos de interés y/o detalles no incluidos en otros campos.

**Multimedia:** fotografías o videos tomados durante el evento.

Datos ambientales, biológicos y otra información relacionada al evento

Se sugiere que en esta capa de información se incluyan todos los datos fisicoquímicos y biológicos obtenidos *in situ* e indicados en los informes o publicaciones existan. También puede incluirse información relativa al evento, observada en campo y que pueda ser de utilidad para el análisis posterior y/o la toma de decisiones.

### Diseño del ID

**Id:** este identificador único caracteriza inequívocamente a cada conjunto de datos fisicoquímicos. Está diseñado en función de dos variables: una temporal (fecha del muestreo) y otra espacial (sitio del muestreo). Se construye mediante la combinación de los códigos que correspondan a cada una (Figs. 4 y 5). Los lineamientos para establecer cada uno están desarrollados en la Sección II: Diseño de códigos, [Códigos para sitios de muestreo](#) y [Códigos para fechas](#).

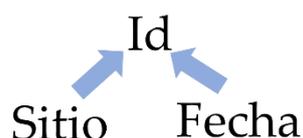


Figura 4. Diseño del identificador único para la planilla de datos ambientales

Sitio	Código Sitio	Fecha	Código Fecha	ID datos ambientales
Balneario La Pérgola	BLPE	5/01/2021	05012021	BLPE05012021
Parador 5	PCIN	5/01/2021	05012021	PCIN05012021

Figura 5. Ejemplo de construcción de ID para los datos ambientales

### Campos nivel de alerta según Organización Mundial de la Salud (OMS)

En el caso que sea posible establecerlo se completan los campos correspondientes con Si o No, teniendo en cuenta:

Código	Nivel de Alerta	Límites establecidos por la OMS (cel/ml)
OMS-CH1	Consumo humano 1er nivel	2000
OMS-CH2	Consumo humano 2do nivel	100000
OMS-URCI1	Uso recreativo con inmersión 1er nivel	20000
OMS-URCI2	Uso recreativo con inmersión 2do nivel	100000

Figura 6. Codificación para los niveles de alerta de la OMS

#### *Campos con variables fisicoquímicas*

Se sugiere que en el nombre del campo se incluya la variable y sus unidades de medición, por ejemplo: Turbiedad (NTU).

Los parámetros usualmente medidos (y sus unidades) son: Temperatura del agua (°C), Turbiedad (NTU), pH (U de pH), alcalinidad (mg/l), cloruros (mg/l), sulfatos (mg/l), nitratos (mg/l), nitritos (mg/l), sólidos totales 105°C (mg/l), conductividad (µS/cm), DBO (mg/l), DQO (mg/l), nitrógeno amoniacal (mg/l), fósforo total (mg/l), Hierro (mg/l), entre otros.

En el caso de que se agreguen nuevos campos se recomienda indicar las unidades utilizadas. Es posible que los datos relevados estén en otras unidades, en estos casos se recomienda hacer la conversión a la unidad sugerida o indicar en el campo de observaciones la unidad utilizada.

#### *Campos con variables biológicas*

Se sugiere que en el nombre del campo se incluya la variable y sus unidades de medición, por ejemplo: Coliformes totales (UFC/100ml).

Las variables usualmente registradas son: coliformes totales (UFC/100 ml), *Escherichia coli* (UFC/100 ml), enterococos (UFC/100 ml), clorofila (mg/m<sup>3</sup> o mg/l). En los casos que hay información se incluyeron otros datos tales como pigmentos y concentración de toxinas (microcistina-LR (µg/L).

#### *Campos relacionados al origen de los datos*

**Laboratorio responsable:** se sugiere, en el caso de contar con esa información, el nombre completo del laboratorio firmante del informe.

**Enlace al informe:** se sugiere incorporar un enlace que permita acceder al informe de origen de los datos, preferentemente dirigido a un repositorio oficial de datos.

**Taxones acompañantes:** en el caso que se considere relevante se indican aquí descripciones de las especies acompañantes mencionadas en la fuente de información.

### Otros campos factibles de incorporar

**Observaciones:** cualquier información que se quiera detallar y no pueda ser incluida en otro campo.

**Multimedia:** fotografías o videos tomados durante el evento.

Se sugieren otros campos cualitativos que están incorporados en la *Harmful Algal Event Database- HAEDAT* y que brindan información de utilidad en el análisis posterior y toma de decisiones ante la presencia de floraciones. Algunos de ellos son:

- **Discoloración del agua:** podría completarse con Si/No o con detalles descriptivos, según lo que se indique en el informe/planilla de campo.
- **Mortalidad masiva fauna asociada:** ídem anterior.
- **Espuma u otras evidencias visuales en la costa:** ídem anterior.
- **Intoxicación humana informada:** ídem anterior.
- **Poblaciones/comunidades naturales afectadas:** ídem anterior. Sugerencias: plancton, bentos, mamíferos acuáticos, peces, aves, etc.

### Vinculación entre los datos en diferentes capas de información.

Los datos incorporados en las capas de información anteriormente descritas: [Presencia de taxones](#) y [Datos ambientales, biológicos y otra información relacionada al evento](#), se vinculan a través del ID de cada registro. Como se observa en la Fig. 6, cada ID incluido en la capa de datos ambientales está inmerso en el ID de la capa de taxones. Es importante considerar que es habitual tener varios registros de taxones con la misma información de datos ambientales, es decir la presencia de dos o más taxones en un mismo evento de muestreo.

Código Taxón	Código Sitio	Código Fecha	ID presencia taxones	ID datos ambientales
MAER	BLPE	05012021	MAERBLPE05012021 →	BLPE05012021
MFLO	BLPE	05012021	MFLOBLPE05012021 →	BLPE05012021
PAAGx	PCIN	05012021	PAAGxPCIN05012021 →	PCIN05012021

Figura 7. Vinculación entre datos de diferentes capas de información

## Sección II Diseño de códigos

### Códigos para taxones

Se recomienda que un mismo taxón tenga el mismo código y que sea único. El código está formado por 4 o 5 letras dependiendo del nivel taxonómico de identificación que se haya

registrado. En la Tabla 1 se muestran los principales lineamientos para la construcción de los códigos con identificación completa mientras que en la Tabla 2 se muestran aquellos con identificación incompleta (código de 5 letras), por ejemplo: está sugerida su relación con otro taxón (o grupo de estos) ya conocido.

Nivel de identificación	Construcción del código	Ejemplo de taxón	Código
Especie	-1 <sup>era</sup> letra del código ➡ 1 <sup>era</sup> letra del género -2 <sup>da</sup> a 4 <sup>ta</sup> letras del código ➡ 1 <sup>era</sup> a 3 <sup>era</sup> letras del epíteto específico	<i>Microcystis aeruginosa</i> .	MAER
Infraespecífico un nivel forma	-1 <sup>era</sup> letra del código ➡ 1 <sup>era</sup> letra del género -2 <sup>da</sup> letra del código ➡ 1 <sup>era</sup> letra del epíteto específico -3 <sup>era</sup> y 4 <sup>ta</sup> letras del código ➡ 1 <sup>era</sup> y 2 <sup>da</sup> letras del epíteto infraespecífico	<i>Anabaena aequalis f. anomala</i>	AAAN
Infraespecífico un nivel variedad	-1 <sup>era</sup> letra del código ➡ 1 <sup>era</sup> letra del género -2 <sup>da</sup> letra del código ➡ 1 <sup>era</sup> letra del epíteto específico -3 <sup>era</sup> y 4 <sup>ta</sup> letras del código ➡ 1 <sup>era</sup> y 2 <sup>da</sup> letras del epíteto infraespecífico	<i>Anabaena aequalis var. major</i>	AAMA
Infraespecífico dos niveles <sup>1</sup>	-1 <sup>era</sup> letra del código ➡ 1 <sup>era</sup> letra del género -2 <sup>da</sup> letra del código ➡ 1 <sup>era</sup> letra del epíteto específico -3 <sup>era</sup> letra del código ➡ 1 <sup>era</sup> letra de la variedad -4 <sup>ta</sup> letra del código ➡ 1 <sup>era</sup> letra de la forma	<i>Caloneis formosa var. rostrata f. lanceolata</i>	CFRL
Género	-1 <sup>era</sup> y 2 <sup>da</sup> letras del código ➡ 1 <sup>era</sup> y 2 <sup>da</sup> letras del género -3 <sup>era</sup> y 4 <sup>ta</sup> letras del código ➡ SP	<i>Noctoc sp.</i>	NOSP

Tabla 1. Lineamientos para la construcción de los códigos de taxones con identificación completa

Nivel de identificación	Construcción del código	Ejemplo de taxón	Código
Affinis <sup>2</sup>	-1 <sup>era</sup> letra del código ➡ 1 <sup>era</sup> letra del género -2 <sup>da</sup> a 4 <sup>ta</sup> letras del código ➡ 1 <sup>era</sup> a 3 <sup>era</sup> letras del epíteto específico -5 <sup>ta</sup> letra del código ➡ letra x (minúscula)	<i>Merismopedia aff. convoluta</i>	MCONx
Confer <sup>3</sup>	-1 <sup>era</sup> letra del código ➡ 1 <sup>era</sup> letra del género -2 <sup>da</sup> a 4 <sup>ta</sup> letras del código ➡ 1 <sup>era</sup> a 3 <sup>era</sup> letras del epíteto específico -5 <sup>ta</sup> letra del código ➡ letra y (minúscula)	<i>Phormidium cf. formosum</i> .	PFORy
Grupo o complejo <sup>4</sup>	-1 <sup>era</sup> letra del código ➡ 1 <sup>era</sup> letra del género -2 <sup>da</sup> letra del código ➡ 1 <sup>era</sup> letra del epíteto específico mencionado en primer lugar. -3 <sup>da</sup> letra del código ➡ 1 <sup>era</sup> letra del epíteto específico mencionado en segundo lugar. -4 <sup>ta</sup> y 5 <sup>ta</sup> letras del código ➡ gr (minúscula)	<i>Dolichospermum</i> grupo "circinalis-spiroides"	DCSgr

Tabla 2. Lineamientos para la construcción de los códigos de taxones con identificación incompleta

<sup>1</sup> El caso de dos niveles infraespecíficos es común en taxones de Diatomeas.

<sup>2</sup> affinis (aff.) con afinidad con otros, similar a. Este término es frecuentemente usado para un taxón reconocido provisionalmente, pero sin nombre, considerado cercano al nombrado, o cuando la atribución es incierta debido a la falta de caracteres necesarios (Hawksworth, 2010). En términos prácticos, el autor decide si utilizar *cf.* o *aff.* basado en su propia experiencia.

<sup>3</sup> Confer (cf) conferir, compararse con; comparar con. (Hawksworth, 2010). Este término se utiliza cuando la identificación de un material se aproxima lo suficiente a un taxón conocido, pero se han encontrado algunas pequeñas diferencias, o no se ha podido realizar una dilucidación más exhaustiva. En términos prácticos, el autor decide si utilizar *cf.* o *aff.* basado en su propia experiencia.

<sup>4</sup> La utilización de la denominación de especie como unidad de estudio puede presentar algunas dificultades, por lo que en muchas ocasiones se realiza la agrupación de las especies en un complejo o grupo que facilita el análisis con objetivos de investigación y gestión. Esta modalidad de determinación simplifica la tarea de identificación microscópica y conteos en trabajos relacionados a la salud pública, aportando a la generación de alertas y a la prevención de intoxicaciones de los usuarios con la rápida detección de una especie como parte del complejo o grupo (Deus Álvarez, 2015; Harke et al., 2016).

Se recomienda armar un listado aparte en orden alfabético con los códigos que se van creando y su nombre científico correspondiente a fin de asegurarse que cada taxón tenga un código único y evitar repeticiones.

#### Criterios para tener en cuenta en la construcción de los códigos

Es muy probable que, siguiendo los lineamientos sugeridos, el código de un nuevo taxón ya esté definido para otro, en estos casos deben seguirse los criterios indicados a continuación mediante un ejemplo hipotético:

Taxón: *Aphanizomenon* sp.

La 1<sup>era</sup> y 2<sup>da</sup> letra del código indica el género. Las letras 3<sup>era</sup> a 4<sup>ta</sup> del código se completan con SP. Siguiendo el criterio inicial el código que le corresponde sería APSP. Como este ya fue creado para otro taxón (*Aphanocapsa* sp.), entonces se salta a la siguiente letra del nombre del género. El código sería AHSP, que también ya fue utilizado para otro taxón (*Aphanothece* sp.), por lo que se pasa a la siguiente letra del nombre del género. El código finalmente queda: AASP

#### **NOTA**

*Pueden existir otros casos de repetición de códigos ya que es bastante improbable que puedan abarcarse todas las opciones. En estos casos, se recomienda mantener lo más posible los criterios para su creación siguiendo los lineamientos anteriormente establecidos.*

#### Códigos para sitios de muestreo

El código está formado por 4 letras mayúsculas. Un mismo sitio deberá tener el mismo código y no deben repetirse. Se construyen teniendo en cuenta que:

-Si el nombre del sitio está formado por una sola palabra, por ejemplo: Córdoba. Se indica con sus cuatro primeras letras: CORD.

-Si el nombre del sitio está formado por dos palabras, por ejemplo, Balneario Bagliardi, se usa la 1<sup>era</sup> letra de la primera palabra, las restantes letras de la 2<sup>da</sup> palabra: BBAG.

- Si el nombre del sitio está formado por tres palabras, por ejemplo, Balneario La Pérgola, se usa la 1<sup>era</sup> letra de la primera palabra, la 1<sup>era</sup> letra de la segunda palabra y el resto de la tercera palabra: BLPE.

-Si el nombre está formado por cuatro palabras, por ejemplo: Club Argentino de Pesca, se usa la 1<sup>era</sup> letra de la primera palabra, la 1<sup>era</sup> letra de la segunda palabra, la 1<sup>era</sup> letra de la tercera palabra y la 1<sup>era</sup> letra de la cuarta palabra. El resultado sería: CADP.

La secuencia indicada anteriormente se respeta si el nombre tiene más palabras, priorizando las palabras completas, sin considerar los artículos. Por ejemplo, para el sitio Toma de Agua Club Universitario, podría resolverse como: TACU.

Debido a que los códigos deben ser únicos, se recomienda armar un listado aparte con los códigos que se van creando y el nombre completo del sitio para verificar cuando se van incorporando nuevos sitios. En el caso de que el código de un nuevo sitio ya esté definido para otro, debe seguirse el mismo criterio adoptado para los códigos de taxones.

#### **NOTA**

*Se recomienda mantener lo más posible los criterios para la creación de códigos siguiendo los lineamientos anteriormente establecidos.*

#### Códigos para fechas

Se construye de modo correlativo indicando los números correspondientes al día, el mes y el año, por ejemplo: 01122001.

En caso de faltar algún dato, se agrega cero a los espacios faltantes. Por ejemplo, para junio de 2022, debería ponerse: 00062022.

# ANEXOS

## Anexo I Tesoros geográficos

Nomenclador de Argentina

<https://www.ign.gob.ar/NuestrasActividades/Geografia/Nomenclador/Buscador>

El Instituto Geográfico Nacional (IGN) desarrolló un Nomenclador Geográfico que ofrece la posibilidad de buscar nombres geográficos (o topónimos) georreferenciados incluidos en la Base de Datos Geoespacial del organismo. El servicio permite la búsqueda de topónimos a partir de un nombre, categoría y/o provincia; y la visualización del objeto, utilizando la cartografía oficial del IGN. Las categorías incluidas en el Nomenclador Geográfico son las siguientes: Divisiones administrativas, Lugares poblados y edificaciones, Hidrografía, Orografía e Infraestructura de transporte.

Geonames

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

Es una base de datos geográficos que abarca todos los países del mundo y contiene más de once millones de topónimos que pueden descargarse gratuitamente.

Getty Thesaurus of Geographic Names

<https://collectiontrust.org.uk/resource/thesaurus-of-geographic-names-tgn-getty/#:~:text=The%20Getty%20Thesaurus%20of%20Geographic,research%20of%20art%20and%20architecture.>

Gacetero mundial plurilingüe.

InfoXY

<http://splink.cria.org.br/infoxy?criaLANG=en>

Herramienta para buscar nombres geográficos y divisiones administrativas del mundo mediante coordenadas geográficas.

geoBoundaries

<https://www.geoboundaries.org/index.html>

Recurso en línea de licencia abierta con información sobre los límites (fronteras) entre países y estados de todo el mundo.

## Anexo II Conversores de coordenadas

epsg.io

<https://epsg.io/transform>

Conversor en línea de coordenadas geográficas, *datum* geodésicos y sistemas proyectados.

Coordinate Parsing from Canadensys

<http://data.canadensys.net/tools/coordinates>

Herramienta para convertir coordenadas geográficas de DDMMS a grados decimales.

Earth Point Coordinate Converter

<http://www.earthpoint.us/Convert.aspx>

Conversiones entre una gran variedad de sistemas de coordenadas.

## Anexo III Georreferenciación retrospectiva

La georreferenciación retrospectiva de localidades consiste en la asignación de coordenadas geográficas a partir de la descripción textual del lugar de recolección. Esta descripción debería contener el mayor número de elementos geográficos que nos permita localizar con la mayor precisión posible el sitio de recolección (Cezón, 2015; Lamaro *et al.*, 2021).

Uno de los métodos más utilizados es el de Radio Punto (Wieczorek *et al.*, 2004), en el cual se describe a una localidad con un par de coordenadas y una distancia de incertidumbre desde ese punto que delimita el radio de una circunferencia. (Chapman & Wieczorek, 2020). Debido a que la georreferencia de una localidad constituye una hipótesis de dónde ocurrió la colecta, la metodología utilizada es estandarizada de forma tal de ser reproducible (Wieczorek *et al.*, 2012; Gómez *et al.*, 2013).

Se recomienda consultar el documento: [Guía rápida de georreferenciación](#) (Zermoglio *et al.*, 2020) en la cual se describen los protocolos para determinar las formas geométricas de las entidades geográficas y cómo usarlas para la georreferenciación con el método de georreferenciación radio punto.

## Anexo IV Estandarización de fechas

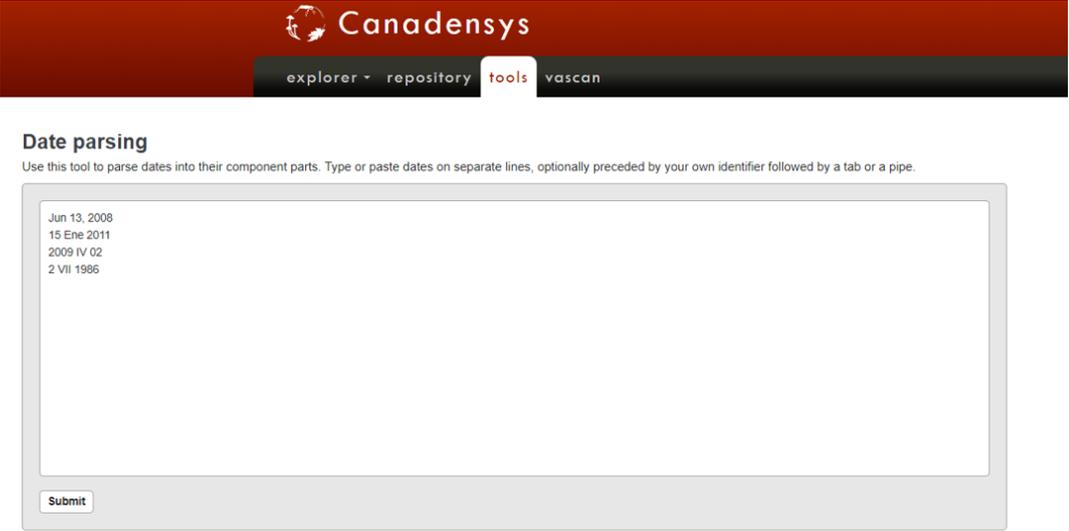
Date Parsing

Herramienta en línea que permite atomizar y estandarizar fechas de forma masiva y automática, de acuerdo con la norma ISO 8601 en un formato AAAA-MM-DD.

Enlace: <http://data.canadensys.net/tools/dates>

Se ingresan las fechas a convertir, la herramienta acepta diferentes formatos (Fig. 8A) y luego se selecciona la opción **Submit**. El resultado se muestra en otra pantalla con la conversión de cada una de las fechas ingresadas (Fig. 8B). Luego se copia desde la misma pantalla y se pega en la planilla Excel (Fig. 8C). Se pegará al Excel respetando el formato regional establecido para la computadora que se esté utilizando. En la Fig. 5C puede observarse que se pegó en formato español de DD-MM-AAAA.

**A**



**B Date parsing results**

original	year	month	day	ISO 8601
Jun 13, 2008	2008	6	13	2008-06-13
15 Ene 2011	2011	1	15	2011-01-15
2009 IV 02	2009	4	2	2009-04-02
2 VII 1986	1986	7	2	1986-07-02

**C**

	A	B	C	D	E
1	original	year	month	day	ISO 8601
2	Jun 13, 2008	2008	6	13	13/6/2008
3	15-ene-11	2011	1	15	15/1/2011
4	2009 IV 02	2009	4	2	2/4/2009
5	2 VII 1986	1986	7	2	2/7/1986

Figura 8. Estandarización masiva de fechas mediante el conversor Data Parsing de Canadensys. A. Ingreso de diversas fechas en varios formatos. B. Resultado de la estandarización. C. Producto final pegado en Excel respetando el formato establecido previamente en la computadora de trabajo.

## Anexo V Limpieza de datos utilizando Excel

El programa Excel de Microsoft permite escribir, almacenar, manipular, calcular y organizar todo tipo de información numérica o de texto por lo que es de suma utilidad para estructurar y estandarizar datos de biodiversidad sobre todo para grandes conjuntos de datos, ahorrando tiempo y optimizando los procesos de estandarización (SIB Colombia, 2015).

## Algunas funciones de utilidad

### *Eliminación de espacios y caracteres no imprimibles*

Aunque muchas veces estos no son evidentes, en el momento de utilizar la información como capa en un SIG y hacer consultas, realizar filtros y agrupaciones de los registros, pueden presentarse inconvenientes en la lectura de la información.

#### Limpiar

Se eliminan todos los caracteres no imprimibles existentes en el texto. Función: =LIMPIAR (texto).

#### Espacios

Elimina todos los espacios del texto [espacios ASCII de 7 bits (valor 32)] excepto los espacios individuales entre palabras y los espacios de no separación (ASCII 160).

Función: =ESPACIOS (texto)

Fórmula: =ESPACIOS(SUSTITUIR(A1,CARACTER(160),CARACTER(32))) donde A1 es la celda objetivo.

#### Sustituir

Elimina el código para la tecla borrar (ASCII 127)

Fórmula =SUSTITUIR(A1,CARACTER(127),"") donde A1 es la celda objetivo.

#### Concatenar

Une varios elementos del texto en uno solo

Función: =CONCATENAR(texto1,texto2,...)

Para concatenar dejando espacios entre palabras utilizar la siguiente fórmula =CONCATENAR(texto1," ",texto2)

Para agregar un espacio o signo de puntuación ( , ; : ), escribir el valor deseado entre comillas ( " ") en un campo de texto separado. Por ejemplo, para concatenar utilizando como separador una coma (,) debería utilizarse la siguiente fórmula ubicando el signo entre las comillas =CONCATENAR(texto1,"",texto2).

#### Si

Comprueba si se cumple una condición y devuelve un valor si se evalúa como verdadero y otro valor si se evalúa como falso. Es útil para comparar la igualdad entre dos columnas.

Función: =SI(prueba\_lógica,valor\_si\_verdadero,valor\_si\_falso)

### *Recomendaciones adicionales*

Es posible que surjan inconvenientes en Excel debido al cambio de formato en las celdas. Se presentan dos modos para hacerlo (Fig. 7):

A. Seleccionando la o las celdas de interés y desplegando la pestaña en el panel superior (menú) para luego escoger diferentes categorías como Número, Fecha, Texto etc.

B. Haciendo clic derecho sobre la celda que se desea cambiar el formato. Luego se elige "Formato de celdas" desplegándose una ventana emergente que da la posibilidad de escoger diferentes categorías como Número, Fecha, Texto etc.

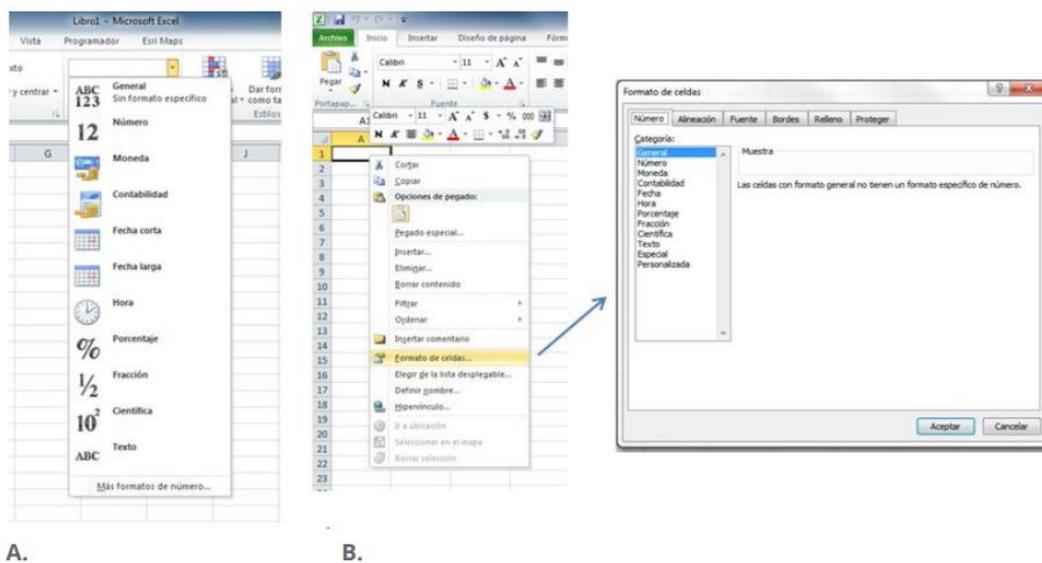


Figura 9. Configuración de formato de celdas en Excel. A. Botón del formato de celda presente en el menú. B. Cambio de formato de celda utilizando el clic derecho directamente sobre la celda objetivo. Fuente: SIB Colombia, 2015.

## Bibliografía citada

El último acceso a todos los enlaces citados fue en febrero de 2024

- Aguilera, A & Echenique, R. O. (2017). Capítulo 2. *Cyanobacteria* nocivas de ambientes acuáticos continentales: taxonomía y ecología. En: *Cianobacterias como determinantes ambientales de la salud*. Giannuzzi, L.; Petcheneshsky, T. Hansen, M. (Eds.). 2a ed. ampliada. - Ciudad Autónoma de Buenos Aires: Ministerio de Salud de la Nación. Dirección Nacional de Determinantes de la Salud e Investigación. ISBN 978-950-38-0255-7. 259 pp. [https://bancos.salud.gob.ar/sites/default/files/2018-10/0000000334cnt-ciano\\_2017.pdf](https://bancos.salud.gob.ar/sites/default/files/2018-10/0000000334cnt-ciano_2017.pdf)
- Cezón, K. (2015). Calidad en bases de datos de biodiversidad. Taller GBIF España. 2-4 diciembre de 2015. [https://www.gbif.es/wp-content/uploads/2018/01/04\\_Calidad-datos-espaciales-temporales.pdf](https://www.gbif.es/wp-content/uploads/2018/01/04_Calidad-datos-espaciales-temporales.pdf)
- Chapman A. D. & Wiczorek J. R. (2020) Guía de buenas prácticas de georreferenciación. Copenhague: Secretariado de GBIF. <https://doi.org/10.15468/doc-gg7h-s853>
- Deus Álvarez, S. (2015.). Caracterización morfológica, genética y óptica de floraciones tóxicas y no tóxicas del Complejo *Microcystis aeruginosa*. Tesis de doctorado. Universidad de la República (Uruguay). Facultad de Ciencias.

- Gómez, S. E.; Torres, S.; Cáseres, G.; Álvarez Merlo, H.; Damborenea, M. C. & Williams, J. D. (2013). Georreferenciación de la base de datos del Museo de La Plata mediante el método Punto-Radio. In I Congreso Latinoamericano y II Congreso Nacional de Museos Universitarios (La Plata, noviembre de 2013). <http://sedici.unlp.edu.ar/handle/10915/42466>
- Harke, M. J., Steffen, M. M., Gobler, C. J., Otten, T. G., Wilhelm, S. W., Wood, S. A., & Pearl, H. W. (2016). A review of the global ecology, genomics, and biogeography of the toxic cyanobacterium, *Microcystis* spp. *Harmful algae*, 54, 4-20. <https://doi.org/10.1016/j.hal.2015.12.007>
- Hawksworth, D. L. 2010. Terms used in bionomenclature: The naming of organisms (and plant communities). - pp. 1-216. Copenhagen. (Global Biodiversity Information Facility). <https://www.gbif.org/document/80577/terms-used-in-bionomenclature-the-naming-of-organisms-and-plant-communities>
- Lamaro, A.A.; Sala, S.E; Kociolek, J.P.; Swenson, J.; Simonato, J.; Guerrero, J.M.; Vouilloud, A.A. & Kociolek, M. (2021) Tras los pasos de Frenguelli en los Esteros del Iberá: el rol de las nuevas tecnologías en la revalorización de las colecciones biológicas históricas, *Revista del Museo de La Plata* 6(1), pp. 1-13. <https://doi.org/10.24215/25456377e140>
- SIB Colombia (2015). Calidad de Datos - Guía de herramientas para mejorar los datos primarios de biodiversidad. Escobar, D., Beltrán, N., Buitrago, L., Plata, C. Delgado, E.; versión 1.0. Bogotá: SIB Colombia, 97 pp. <http://repository.humboldt.org.co/handle/20.500.11761/35351>
- Wieczorek J. R.; Guo, Q. & Hijmans R. (2004) The point-radius method for georeferencing locality descriptions and calculating associated uncertainty, *International Journal of Geographical Information Science*, 18:8, 745-767 <https://www.tandfonline.com/doi/abs/10.1080/13658810412331280211>
- Wieczorek, J. R.; Bloom, D.; Guralnick, R.; Blum, S.; Doring, M; Giovanni, R.; Robertson, T. & Vieglais, D. (2012). Darwin Core: An Evolving Community-Developed Biodiversity Data Standard. *PLoS ONE* 7(1): e29715. <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0029715>
- Zermoglio, P.F.; Chapman, A.D.; Wieczorek, J.R.; Luna, M.C. Bloom, D.A. (2020) Guía rápida de georreferenciación. Copenhagen: GBIF Secretariat. <https://docs.gbif.org/georeferencing-quick-reference-guide/1.0/es/>