

Buenas prácticas para generalizar datos de especies sensibles en registros biológicos

Arthur D. Chapman

Versión HEAD detached, 2023-05-01 16:46:54 UTC

Tabla de Contenido

Colofón	1
Cita bibliográfica sugerida	1
Autor	1
Licencia	1
URI persistente	1
Control del documento	1
Imagen de la portada	1
Introducción	2
Objetivos	2
Audiencia	3
Alcance	3
1. Principios	5
2. Determinación de la sensibilidad	7
2.1. Criterios para la determinación de la sensibilidad	8
2.2. Categorías de sensibilidad	17
3. Generalización de la información textual	21
4. Generalización de la información espacial	23
4.1. Generalización versus aleatorización	23
4.2. Generalización	23
4.3. Documentación	26
4.4. Duplicados y GUID	27
5. Documentación y metadatos	28
5.1. Documentación de la sensibilidad	28
5.2. Ajuste espacial	29
6. Autenticación y autorización	31
7. Implementaciones	32
Epílogo	37
Listado de taxa sensibles	37
Recomendaciones de metadatos	38
Glosario	41
Agradecimientos	43
Referencias	44
Anexo 1: Escenarios usando los criterios 1 y 2 como desencadenantes	48
Criterio 1	48
Criterio 2	49

Colofón

Cita bibliográfica sugerida

Chapman AD (2020) Buenas prácticas actuales para generalizar datos de especies sensibles. Copenhagen: GBIF Secretariat. <https://doi.org/10.15468/doc-5jp4-5g10>.

Autor

Arthur D. Chapman

Licencia

Este documento *Buenas prácticas actuales para la generalización de datos de registros biológicos de especies sensibles* se publica bajo una licencia [Creative Commons Atribución-CompartirIgual 4.0 Internacional](#).

URI persistente

<https://doi.org/10.15468/doc-5jp4-5g10>

Control del documento

Version 1.0, noviembre del 2020

Basado en la publicación anterior de [Chapman AD & Grafton O \(2008\)](#) Guide to Best Practices for Generalising Sensitive Species-Occurrence Data. Copenhagen: GBIF Secretariat. <https://doi.org/10.15468/doc-b02j-gt10>.

Imagen de la portada

Bajo la superficie, Wollemi National Park, Australia. Foto 2015 Vern vía [Flickr](#), licenciada bajo [CC BY-NC-ND 2.0](#).

Introducción

La distribución no protegida de los datos de presencia de especies sensibles (por ejemplo, las localidades exactas de taxa raros, en peligro de extinción o de valor comercial) fue una preocupación de GBIF (la Infraestructura mundial de información en biodiversidad) desde sus inicios. La Secretaría de GBIF tiene un gran interés en que los datos estén disponibles a través de sus portales, pero respetando al mismo tiempo los deseos de los proveedores de datos de restringir la información sobre los taxa sensibles. A principios de 2006, GBIF inició un proceso para abordar esta problemática, especialmente en relación con los datos que se compartirán a través de la red de GBIF y se harán visibles a través de GBIF.org y otras iniciativas de agregación de datos.

El resultado fue la *Guide to Best Practices for Generalising Sensitive Primary Species Occurrence Data*. Ese documento se basó en gran medida en los resultados de una encuesta realizada en línea a través de [Survey Monkey](https://www.surveymonkey.com) y en talleres posteriores cuyos informes se publicaron originalmente en el sitio web de GBIF ([Chapman 2006](#)).

Tras estos procesos y debates se elaboró un informe final sobre el tratamiento de los datos primarios de presencia de especies sensibles, que se presentó a GBIF en abril de 2007 ([Chapman 2007](#)). Este informe hizo una serie de recomendaciones y muchas de ellas han sido incluidas en este documento.

El último paso en ese proceso fue desarrollar una *Guía de buenas prácticas para los datos primarios de presencia de especies sensibles*. Ese documento se propuso como una guía principal para que las instituciones, los proveedores de datos y los Nodos de GBIF lo utilizaran para desarrollar sus propias directrices internas. Se animó a las organizaciones e instituciones a producir sus propios documentos internos, que incorporaran las prácticas descritas en la *Guía* y otros documentos relacionados como la *Guía de buenas prácticas de georreferenciación* ([Chapman y Wieczorek 2006](#)) y a incorporarlos a su propio entorno de trabajo. Por desgracia, no son tantas las instituciones que han aceptado el reto y han elaborado sus propios documentos internos como esperábamos. Sin embargo, dos organismos clave que sí lo han hecho son el SANBI de Sudáfrica ([SANBI 2010](#)) y el Atlas of Living Australia ([Tann y Flemons 2009](#), [ALA 2018a](#)) (ver [Implementaciones](#)).

También es importante comprender el posible impacto que las estrategias para restringir los datos sensibles pueden tener en la ciencia de la biodiversidad, para que, aunque se restrinja la disponibilidad o la resolución de determinados datos, no se restrinjan excesivamente los usos que pueden darse a los datos. Por ello, a continuación se exponen una serie de principios. Entre ellos destaca la necesidad de que la información sobre la biodiversidad sea de libre acceso siempre que sea posible, en pro de la ciencia, el ambiente y la propia biodiversidad.



Las palabras en el texto que tienen un hipervínculo se refieren a términos que están incluidos en el [Glosario](#). Las citas enlazan con las fuentes cuando están disponibles en línea y con las [Referencias](#) cuando no lo están.

Objetivos

El objetivo de este documento es proporcionar las buenas prácticas (o las buenas prácticas actuales) para tratar los datos de presencia de especies sensibles, y orientar sobre cómo poner a disposición

la mayor cantidad de datos sin que la especie se vea expuesta por el hecho de que los datos se hayan puesto en el dominio público.

Ha pasado más de una década desde que se publicó la primera *Guía* y esta nueva publicación está diseñada para actualizar esas prácticas e incorporar las experiencias adquiridas por las instituciones que han aplicado la *Guía* en su totalidad o parcialmente.

Audiencia

Este documento está destinado a quienes necesitan o desean saber cuál es la mejor manera de poner a disposición de los interesados la mayor cantidad posible de datos sobre taxa sensibles sin que esa publicación perjudique a la especie. También, está dirigido a las personas u organizaciones que se enfrentan a la elaboración de una política sobre el tratamiento de datos primarios de especies sensibles y a la redacción de documentación interna coherente con las buenas prácticas actuales.

Sobre todo, este documento ayudará a los usuarios finales de los datos a entender las implicaciones de intentar utilizar registros que pueden haber sido generalizados para proteger especies sensibles y sobre cómo entender el significado de la generalización a diferentes precisiones.

Alcance

El término "buena práctica" generalmente se refiere a la mejor manera posible de hacer algo. Habitualmente se utiliza en los campos de la gestión empresarial, la ingeniería de software, la medicina y, cada vez más, en la administración pública. El término "buena práctica actual" es más específico en el sentido de que indica la posibilidad de futuros desarrollos y buenas prácticas. Debido a la inmadurez de este tema, esta publicación se refiere en general a la buena práctica "actual" y seguramente madurará con el tiempo, a medida que más instituciones adopten y adapten los principios aquí expuestos.

Dos asuntos que no se han tratado en este documento, ni en el anterior, son los asuntos relativos a la privacidad de las personas vivas y la elaboración de acuerdos de intercambio y licencia de datos. Ambos asuntos tienen implicaciones legales y varían considerablemente de una jurisdicción a otra. Estos asuntos han sido tratados en detalle por otros (p. ej., [Corti et al. 2000](#), [Parry y Mauthner 2004](#), [GBIF 2017](#), [GBIF 2019](#), [ALA 2018b](#), [OEH 2019b](#)).

La ética y la biodiversidad es un tema que se ha discutido poco, aunque durante cientos de años los biólogos han seguido una ética implícita en su trabajo. La gestión de los datos sensibles requiere una práctica ética considerable y, en muchos casos, mucha confianza y colaboración. A menudo, los biólogos aficionados y los científicos ciudadanos conocen la ubicación de los taxa sensibles y depende de los biólogos colaborar con estos grupos para garantizar la supervivencia de la especie. Esto no siempre será posible, ya que siempre habrá personas mal intencionadas, pero la colaboración puede funcionar. Por ejemplo, trabajar con grupos de aficionados para cultivar plantas raras de manera ética, de modo que se ejerza menos presión sobre las poblaciones silvestres.

La arqueología lleva mucho tiempo tratando con estas cuestiones. Como explica [Alison Wylie \(1996\)](#),

la creación de la Society for American Archaeology (SAA) estuvo motivada, en parte, por el deseo de los académicos de distanciarse de los arqueólogos aficionados mediante el establecimiento de códigos de conducta ética. Desde hace mucho tiempo una de las principales preocupaciones en este campo ha sido la práctica del saqueo de yacimientos históricos y la amenaza que esto supone tanto para el patrimonio cultural como para los futuros trabajos arqueológicos.

Diecinueve años después los arqueólogos siguen luchando, como señalan [Frank et al. \(2015\)](#) al analizar los resultados de 62 entrevistas a arqueólogos y zoólogos realizadas por ellos. Llegaron a la conclusión de que "los investigadores preferirían, en general, restringir el acceso del público general a sus datos, pero mantener los datos abiertos para los colegas", al tiempo que se dieron cuenta de las grandes dificultades para gestionar dicho proceso.

Esta publicación se centra, en gran medida, en los datos de presencia de especies sensibles (p. ej., datos de especímenes de museos y herbarios, observaciones, imágenes, datos de seguimiento), pero tiene repercusiones en los datos relacionados, como el tamaño, el número y la viabilidad de las poblaciones, los hábitats y los ecosistemas, la biogeografía, los rasgos, las capturas accidentales, la bioseguridad, etc.

Este documento sólo analiza la sensibilidad de los taxa y no la sensibilidad relacionada con cuestiones como la tierra o la privacidad personal. Las distintas jurisdicciones nacionales y estatales tienen legislaciones diversas en lo que respecta a la tierra y a la privacidad personal. Sin embargo, teniendo esto en cuenta, recomendaría seguir los mismos principios de generalización recomendados en esta publicación siempre que sea posible, en lugar de otros métodos como la aleatorización.

1. Principios



La información sobre biodiversidad debe estar disponible de forma gratuita para ser compartida a nivel mundial y permitir su uso para la toma de decisiones sin ánimo de lucro, la educación, la investigación y otros fines de beneficio público. Poner a disposición todos los detalles de la información sobre biodiversidad debería reducir el riesgo de daños al medio ambiente y ayudar a salvaguardar un futuro sostenible. Cuando la divulgación pueda tener el efecto contrario, puede ser necesario controlar el acceso a todos los detalles.

A continuación se presentan una serie de principios de alto nivel relacionados con el intercambio de datos en general y con el intercambio de datos sensibles en particular.

1. La gestión de los datos sensibles es parte integral de la gestión ética de los datos.
2. Siempre que sea posible la información ambiental debe estar a disposición de todos. En general, esto beneficia al medio ambiente al aumentar la concientización, permitir una mejor toma de decisiones y reducir el riesgo de daños.
3. La divulgación de la información a veces puede resultar perjudicial para el medio ambiente. En tales casos, puede ser necesario controlar la disponibilidad de la información, aunque la premisa sigue siendo a favor de la divulgación y cualquier restricción debe ser evaluada y revisada rigurosamente.
4. Todos los datos considerados sensibles deben incluir una fecha para la revisión de su estado de sensibilidad, junto con la documentación de las razones para el estado de sensibilidad. La fecha de revisión puede ser a corto o largo plazo, dependiendo de la naturaleza de la sensibilidad.



Siempre que un proveedor de datos reciba una solicitud para mejorar el acceso a los datos restringidos, debe evitar asumir que siguen siendo sensibles y debe aprovechar la oportunidad para revisar la determinación.

5. Si los datos deben ser restringidos para su distribución, esto sólo debe aplicarse sobre una copia de los datos en el momento de su distribución. Los datos originales nunca deben ser alterados, falsificados o borrados de la base de datos.
6. La documentación es esencial por muchas razones y cuando los datos han sido restringidos o generalizados, es importante que la(s) razón(es) de la categorización se registre(n) como metadatos que permanezcan junto al registro.
7. Cuando los datos se restrinjan o se generalicen para su distribución (como el nombre de un colector, la información textual de la localidad, etc.), deberán documentarse sustituyéndolos por un texto adecuado. El campo no deberá dejarse en blanco o anularse.
8. Hay razones de peso para **no** restringir los datos de las colectas relacionadas (p. ej., los números de los colectores en secuencia, el nombre del colector, etc.) debido a las restricciones que esto supone para la calidad de los datos y los procedimientos de validación de los mismos, etc.
9. Los usuarios de datos sensibles deben cumplir con todas las restricciones de acceso que el proveedor de los datos haya impuesto a los mismos. Si se les concede un acceso más amplio a la información restringida, los usuarios no deben comprometer o infringir de otro modo la

confidencialidad de dicha información.

10. Los proveedores de datos deben respetar la necesidad de los usuarios de tener acceso a los datos y la documentación para determinar la "aptitud para el uso" (fitness for use) de los datos y garantizar que los análisis sean sólidos y no induzcan al error.

2. Determinación de la sensibilidad

Como primer paso, los titulares de la información deben identificar los datos que se consideran **sensibles**. La información sensible es toda aquella que, de ser divulgada al público, tendría un "efecto adverso" sobre el taxón o atributo en cuestión, o sobre un individuo vivo. A la hora de determinar la sensibilidad hay que tener en cuenta una serie de factores, como el tipo y el nivel de amenaza, la vulnerabilidad del taxón o atributo, el tipo de información y si ya está disponible públicamente. La determinación de estos factores nos lleva a un enfoque basado en criterios.

La información no puede considerarse sensible si está fácilmente disponible a través de otras fuentes o si no es única. Este principio ha sido identificado en una serie de políticas de datos sensibles (AMEC Earth and Environmental 2010, Australian Government 2016).

Parece que los herbarios son más proclives a restringir sus datos que las colecciones de mamíferos o insectos (Chapman 2006). Quizá esto se deba a que las plantas no se mueven y es probable que la ubicación exacta de una colección nos lleve a una planta real sobre el terreno, mientras que los mamíferos y los insectos tienden a moverse. Un entomólogo comentó que los colectores profesionales y los grupos de aficionados suelen saber más que los científicos sobre la ubicación de las especies raras. Sin embargo, hay categorías de animales cuya ubicación exacta se consideraba delicada, como los lugares de descanso y maternidad de los murciélagos, los lugares de anidación de los halcones y la ubicación de varios lagartos, tortugas, especies de mariposas y grandes mamíferos. En cuanto a las plantas, también hay una fuerte tendencia a no facilitar información sobre las plantas susceptibles de ser colectadas (pirateadas), como los cactus de Arizona, las orquídeas y las cícadas. También se identificó la protección de yacimientos fósiles sensibles. Un aspecto desafortunado es la susceptibilidad de un pequeño número de taxa en un grupo (como unos pocos cactus carismáticos u orquídeas, etc.). Esto puede significar que a menudo todos los taxa de ese grupo se consideren sensibles y se restrinjan los datos sobre ellos, aunque muchos de esos taxa no sean en sí mismos sensibles o susceptibles de sufrir actos perjudiciales.

Por otro lado, algunas instituciones han encontrado beneficios en la colaboración con el público general para recopilar información y proteger los taxa raros, usando al público y a los grupos de interés especial para estudiar los lugares conocidos y ayudar a encontrar nuevos lugares. Hay buenos ejemplos con aves, lagartos, ranas, mariposas y algunas especies de plantas (incluidas las orquídeas) en varios países. Varias personas han planteado la cuestión del equilibrio entre la protección de los taxa a través del conocimiento de su presencia y la protección a través de la restricción del conocimiento de su presencia en un lugar. Esta cuestión es muy específica de cada taxón (y quizás de cada región) y algunos taxa pueden estar en mayor peligro debido a la destrucción inadvertida por falta de conocimiento, que por la colecta y destrucción deliberada a través del conocimiento de los lugares. Por este motivo, un listado de taxa sensibles debería ser bastante diferente a un listado de taxa raros o amenazados, aunque es probable que haya una superposición considerable entre ambas. También hay que tener en cuenta que lo que es sensible hoy, puede no serlo mañana y viceversa, esto debería llevar a una revisión periódica para determinar si el contexto ha cambiado con el tiempo (AMEC Earth and Environmental 2010).

Como se señala en un artículo de *Science* (Stuart et al. 2006), tres especies recientemente descubiertas de anfibios y reptiles aparecieron rápidamente en el comercio poco después de sus descripciones en la literatura científica. Cuánta información debe publicarse al describir un nuevo taxón es una cuestión que preocupa a los biólogos y, sobre todo, a los taxónomos (Guterman 2006).

Hay varios ejemplos de peces de arrecife de coral en los que una nueva especie ha aparecido en el comercio poco después de ser descrita científicamente.

Algunos ejemplos con los que he tenido una experiencia directa incluyen *Centropyge boylei*, *Centropyge narcosis* [and] *Belonoperce pylei* ... entre otros.

A menudo, en estos y otros casos, la existencia de la nueva especie es señalada a la comunidad científica "por" el comercio (de acuarios) y no al revés. Por lo tanto, no suele considerarse un "problema", sino una especie de relación "simbiótica" entre el comercio y los taxónomos. Además, en estos casos de peces de arrecife, la especie ha eludido el descubrimiento previo no tanto porque sea rara o tenga una distribución extremadamente restringida, sino porque simplemente vive en un lugar que los científicos aún no han podido estudiar. Por lo tanto, las implicaciones para la conservación suelen ser escasas o nulas en este contexto.

— Richard Pyle, comunicación personal 2006

2.1. Criterios para la determinación de la sensibilidad

La National Biodiversity Network (NBN) del Reino Unido ([Countryside Agencies OIN 2007](#)) y el Department of Environment and Conservation in New South Wales, Australia ([Department of Environment and Conservation 2007](#)) desarrollaron criterios de sensibilidad detallados, el trabajo de esas dos organizaciones influyó en gran medida en la versión anterior de esta publicación ([Chapman and Oliver 2008](#)). Desde la publicación de la *Guía* anterior, estas dos organizaciones ([DECCW 2009](#), [NBN 2019a](#), [NBN 2019b](#), [OEH 2019a](#)) junto con el South African National Biodiversity Institute ([SANBI 2010](#), [SANBI 2016](#)), el Atlas of Living Australia ([ALA 2018a](#)) y otros, han reflexionado mucho sobre los criterios para determinar la sensibilidad. La documentación de todos estas organizaciones ha contribuido en gran medida a este documento.

En la [Tabla 1](#) se exponen una serie de criterios para determinar la sensibilidad de los taxa y los datos, junto con enunciados de metadatos recomendados para documentar las razones de la determinación. Los dos primeros son para uso de los titulares de datos de biodiversidad y de quienes crean listados de taxa potencialmente sensibles, y se refieren en gran medida a los propios taxa. Las dos últimas son para uso de los titulares de datos sobre biodiversidad y se refieren a una evaluación de los datos que poseen y que están considerando poner a disposición pero no son adecuadas para la creación de listados derivados.

Los criterios se usan para determinar:

Tabla 1. Criterios para determinar la sensibilidad de los taxa y los datos, junto con los enunciados de metadatos recomendados para documentar las razones de la determinación

1. Riesgo de daño	Una evaluación de si el taxón está expuesto a actividad humana perjudicial.
--------------------------	---

2. Impacto del daño	Una evaluación de la sensibilidad del taxón a actividad humana perjudicial.
3. Sensibilidad de los datos	Una evaluación sobre si la publicación de los datos aumentará el daño.
4. Decisión sobre la liberación y la categoría de sensibilidad	Una decisión equilibrada sobre la divulgación de los datos y la determinación de la categoría de sensibilidad (y por tanto del nivel de generalización) de los datos para su divulgación.

En el **Anexo** de este documento se adjuntan un conjunto de escenarios utilizando los Criterios 1 y 2 de la tabla anterior para determinar los desencadenantes de la sensibilidad de los taxa.

El primer paso en el proceso de determinación de la sensibilidad es hacer una evaluación sobre si el taxón está o no sujeto a una actividad humana perjudicial y si la disponibilidad de datos de biodiversidad relacionados aumentará la probabilidad de que ocurra la actividad perjudicial.

Si no es así, no parece haber ninguna razón para incluirlo en la lista de taxa potencialmente sensibles desde el punto de vista ambiental. Se recomienda utilizar el enunciado proporcionado en la documentación, pero con un apoyo racional adicional que documente los aspectos específicos de la amenaza, por ejemplo:

El taxón está en riesgo por actividad humana perjudicial: está sujeto al ataque de *Phytophthora* que es transportado por vehículos operados por los humanos.

Tabla 2. **Riesgo de daño:** evaluar si el taxón está sometido a una actividad humana perjudicial

1.1. ¿Está el taxón sometido a una actividad humana perjudicial?	
<p>SI</p> <p>↓</p> <p>Documentar con la justificación del enunciado 1a: <i>"El taxón está en riesgo por una actividad humana perjudicial."</i></p> <p>↓</p> <p>Ir a 1.2</p>	<p>NO</p> <p>↓</p> <p>Documentar usando el enunciado 1b: <i>"No hay riesgo significativo de una actividad humana perjudicial."</i></p> <p>↓</p> <p>El taxón no es sensible</p> <p>↓</p> <p>Ir a 3</p>
1.2. ¿Existe evidencia de acontecimientos actuales o recientes de la actividad humana dañina?	
<p>SI</p> <p>↓</p> <p>Documentar con la justificación del enunciado 1c: <i>"Hay evidencia establecida de daño actual o reciente al taxón."</i></p> <p>↓</p> <p>Ir a 1.3</p>	<p>NO</p> <p>↓</p> <p>Documentar con la justificación del enunciado 1d: <i>"Actualmente no hay evidencia establecida de un daño real al taxón."</i></p> <p>↓</p> <p>Ir a 1.3</p>
1.3. ¿La disponibilidad de datos sobre la biodiversidad relacionada aumentará la probabilidad de que la actividad humana perjudicial ocurra?	
<p>SI</p> <p>↓</p> <p>Documentar con la justificación del enunciado 1e: <i>"La disponibilidad de datos sobre la biodiversidad aumentará la probabilidad de que se lleve a cabo la actividad humana perjudicial."</i></p> <p>↓</p> <p>Ir a 2</p>	<p>NO</p> <p>↓</p> <p>Documentar con la justificación del enunciado 1f: <i>"La disponibilidad de datos sobre la biodiversidad no aumentará la probabilidad de que la actividad humana perjudicial ocurra."</i></p> <p>↓</p> <p>Ir a 2</p>

El siguiente paso es determinar si el taxón es sensible a ese daño humano o si es suficientemente robusto para no verse afectado negativamente.

Tabla 3. **Impacto del daño.** Evaluación de la sensibilidad de los taxa a una actividad humana perjudicial.

2.1. ¿Tiene el taxón características que lo hacen particularmente vulnerable a la actividad humana perjudicial?	
<p>SI</p> <p>↓</p> <p>Documentar con la justificación del enunciado 2a:</p> <p><i>"El taxón tiene características que lo hacen particularmente vulnerable a la actividad humana perjudicial."</i></p> <p>↓</p> <p>Ir a 2.2</p>	<p>NO</p> <p>↓</p> <p>Documentar con la justificación del enunciado 2b:</p> <p><i>"El taxón no es particularmente vulnerable a la actividad humana perjudicial."</i></p> <p>↓</p> <p>Ir a 2.2</p>
2.2. ¿Es el taxón vulnerable a la actividad humana perjudicial en toda su área de distribución o hay zonas (como en las zonas de conservación o en otras partes del mundo) en las que el taxón no tiene el mismo nivel de riesgo?	
<p>SI</p> <p>↓</p> <p>Documentar con la justificación del enunciado 2c:</p> <p><i>"El taxón es vulnerable a la actividad humana perjudicial en toda su área de distribución."</i></p> <p>↓</p> <p>Ir a 3</p>	<p>NO</p> <p>↓</p> <p>Documentar con la justificación utilizando el enunciado 2d:</p> <p><i>"El taxón no es vulnerable a la actividad humana perjudicial en toda su área de distribución y/o hay zonas en las que el taxón está presente pero no corre un riesgo significativo."</i></p> <p>↓</p> <p>Ir a 3</p>

Una vez que se ha decidido que el taxón está sujeto a un riesgo significativo y a impacto del daño o no, entonces hay que decidir si la publicación de datos específicos sobre ese taxón, u otros datos relacionados, aumentará el riesgo y el impacto del daño.

Tabla 4. **Sensibilidad de los datos.** Evaluar si la liberación de los datos aumentará el daño.

3.1. ¿Es el contenido y el detalle de los datos sobre biodiversidad tal que su divulgación permitiría a alguien llevar a cabo una actividad perjudicial sobre el taxón o atributo?	
<p>SI ↓ Documentar con la justificación del enunciado 3a: <i>“El contenido y el detalle de los datos es tal que su divulgación permitiría a alguien llevar a cabo una actividad perjudicial sobre el taxón o atributo.”</i> ↓ Ir a 3.2</p>	<p>NO ↓ Los datos no son sensibles Documentar con la justificación del enunciado 3b: <i>“La divulgación del contenido y los detalles de los datos no permitirían a nadie llevar a cabo una actividad perjudicial sobre el taxón o atributo.”</i>+ ↓ Ir a 4</p>
3.2. ¿La información ya es de dominio público o ya es conocida por los individuos o grupos que pueden llevar a cabo la actividad perjudicial?	
<p>SI ↓ Documentar con la justificación del enunciado 3d: <i>“La información ya es de dominio público o ya es conocida por los individuos o grupos susceptibles de realizar actividades perjudiciales.”</i> ↓ Ir a 3.3</p>	<p>NO ↓ Documentar con la justificación del enunciado 3c: <i>“La información no es de dominio público y no es conocida por individuos o grupos susceptibles de realizar actividades perjudiciales.”</i> ↓ Ir a 3.3</p>
3.3. ¿Es probable que la divulgación de los datos dañe una asociación o relación (especialmente cuando el mantenimiento de estos es esencial para ayudar a lograr un objetivo de conservación específico)?	
<p>SI ↓ Documentar con la justificación del enunciado 3e: <i>“Es probable que la divulgación de los datos perjudique una asociación o un vínculo cuyo mantenimiento es esencial para ayudar a alcanzar un objetivo de conservación específico.”</i> ↓ Ir a 3.4</p>	<p>NO ↓ Documentar con la justificación del enunciado 3f: <i>“La divulgación de los datos no perjudicará ninguna asociación o vínculo esencial para la conservación.”</i> ↓ Ir a 3.4</p>
3.4. ¿La divulgación permitiría obtener las ubicaciones de las entidades geográficas sensibles mediante la combinación con otras fuentes de información disponibles?	

<p>SI</p> <p>↓</p> <p>Documentar con la justificación del enunciado 3g:</p> <p><i>“La divulgación permitiría derivar las ubicaciones de las entidades geográficas sensibles mediante la combinación con otras fuentes de información disponibles públicamente.”</i></p> <p>↓</p> <p>Ir a 4</p>	<p>NO</p> <p>↓</p> <p>Documentar con la justificación del enunciado 3h:</p> <p><i>“La divulgación no permitirá derivar las ubicaciones de las entidades geográficas sensibles mediante la combinación con otras fuentes de información disponibles públicamente.”</i></p> <p>↓</p> <p>Ir a 4</p>
--	--

El último paso es llevar a cabo una evaluación global basada en los tres criterios anteriores y documentar la decisión global utilizando la combinación de la información que ha sido documentada al tomar cada una de las decisiones anteriores. Una vez que se ha determinado si los datos deben o no deben ser liberados, entonces es importante que se tome una decisión sobre la **Categoría de sensibilidad** y el nivel de **generalización** para la liberación de los datos.

Tabla 5. **Decisión sobre la divulgación y la categoría de sensibilidad.** Tomar una decisión equilibrada respecto a la liberación de los datos y determinar la categoría y el nivel de generalización.

<p>4.1. Teniendo en cuenta los criterios 1 a 3, mencionados arriba junto con cualquier otro contexto relevante, ¿la retención de la información aumentará el riesgo de perjuicio al ambiente o a una persona viva?</p>	
<p>SI ↓ Documentar utilizando el enunciado 4a: <i>“En conjunto, la divulgación de la información aumentará o es probable que aumente el riesgo de daño ambiental o de daño a una persona viva”.</i> ↓ Ir a 4.2</p>	<p>NO ↓ Documentar utilizando el enunciado 4b: <i>“En conjunto, la divulgación de los datos no aumentará el riesgo de daño ambiental o de daño a una persona viva.”</i> ↓ Ir a 4.5</p>
<p>4.2. ¿El taxón es distintivo y de gran importancia biológica, está muy amenazado por explotación/enfermedad u otra amenaza identificable en la que incluso la información general de la localidad puede amenazar al taxón? ¿O podría la divulgación de cualquier parte del registro causar un daño irreparable al medio ambiente o a un individuo?</p>	
<p>SI ↓ Documentar utilizando el enunciado 4c, recopilando todos los fundamentos y documentando la decisión de retener los datos: <i>“La especie es una especie distintiva de gran importancia biológica, está muy amenazada por explotación/enfermedad u otra amenaza identificable e incluso la información general sobre la localidad puede amenazar al taxón, o la divulgación de la información podría causar un daño irreparable al medio ambiente, a un individuo o a alguna otra entidad.”</i> Categoría 1</p>	<p>NO ↓ Ir a 4.3</p>
<p>4.3. ¿Es el taxón tal, que proveer ubicaciones precisas de más de 0.1 grados (~10 km) sometería al taxón a amenazas como la perturbación y la explotación? ¿O el registro incluye información altamente sensible, cuya divulgación podría causar un daño extremo a un individuo o al medio ambiente?</p>	

<p>SI</p> <p>↓</p> <p>Documentar utilizando el enunciado 4d, recopilando todos los fundamentos y documentando la decisión de divulgar los datos: <i>“La especie está clasificada como altamente sensible y el suministro de ubicaciones precisas sometería a la especie a amenazas tales como la perturbación y la explotación y/o el registro incluye información altamente sensible, cuya divulgación podría causar un daño extremo al ambiente o a un individuo.”</i></p> <p>Categoría 2</p>	<p>NO</p> <p>↓</p> <p>Ir a 4.4</p>
<p>4.4. ¿Es el taxón tal, que proveer ubicaciones precisas a más de 0.01 grados (~1 km) sometería a la especie a amenazas como colecta o daño deliberado? ¿O el registro incluye información sensible cuya divulgación podría causar daño a un individuo o al medio ambiente?</p>	
<p>SI</p> <p>↓</p> <p>Documentar utilizando el enunciado 4e, recopilando todos los fundamentos y documentando la decisión de divulgar los datos: <i>“La especie está clasificada como de sensibilidad media a alta y el suministro de ubicaciones precisas podría someter a la especie a amenazas como colecta o daño deliberado y/o el registro incluye información sensible, cuya divulgación podría causar daños al medio ambiente o a un individuo.”</i></p> <p>Categoría 3</p>	<p>NO</p> <p>↓</p> <p>Ir a 4.5</p>
<p>4.5. ¿Estaría el taxón sometido a una amenaza baja o media si las ubicaciones precisas (es decir, ubicaciones con una precisión superior a 0.001 grados o 100 m) se hacen públicas y existe algún riesgo de colecta o daño deliberado?</p>	

<p>SI</p> <p>↓</p> <p>Documentar utilizando el enunciado 4f, recopilando todos los fundamentos y documentando la decisión de liberar los datos:</p> <p><i>“La especie está clasificada como de sensibilidad baja a media y el suministro de ubicaciones precisas podría someter a la especie a amenazas como la perturbación y la explotación. Los datos detallados pueden ponerse a disposición de los individuos bajo licencia.”</i></p> <p>Categoría 4</p>	<p>NO</p> <p>↓</p> <p>Documentar utilizando el enunciado 4g, recopilando todos los fundamentos y documentando la decisión de liberar los datos:</p> <p><i>“La especie está clasificada como de baja sensibilidad y es improbable que la distribución de las ubicaciones precisas someta a la especie a una amenaza significativa y/o el registro incluye información de baja sensibilidad, cuya divulgación es improbable que cause daño al medio ambiente o a cualquier individuo. Los datos deben hacerse públicos “tal como se encuentran.”</i> No es sensible para el ambiente</p> <p>↓</p> <p>Los datos deben hacerse públicos</p>
---	---

En la encuesta en línea (Chapman 2006), varios encuestados señalaron que los datos pendientes de publicación, los datos sujetos a investigación en curso y los datos incompletos o no revisados eran datos que clasificarían como sensibles y por tanto, sujetos a restricciones de publicación. Se trata de datos cuya sensibilidad tiene un plazo de tiempo corto y es importante que se documente claramente el plazo de publicación o revisión. Lo más probable es que entren en el criterio anterior 3.3 y se documenten en consecuencia, con la justificación de "en espera de publicación", etc.



Todos los datos considerados sensibles deben incluir una fecha para la revisión de su estado de sensibilidad, junto con la documentación de las razones para el estado de sensibilidad. La fecha de revisión puede ser a corto o largo plazo, dependiendo de la naturaleza de la sensibilidad.

Las **categorías de sensibilidad** se basan en gran medida en las de NSW Office of Environment and Heritage (DECCW 2009).

2.2. Categorías de sensibilidad

Tabla 6. Categorías de sensibilidad

Criterio	Razonamiento
<p>Categoría 1</p> <p>Especies o registros para los que no se proporcionará documentación en absoluto o que sólo se dan a conocer como presentes en una gran región, como un condado, una cuenca hidrográfica, etc.</p>	<p>El motivo de la no divulgación es que:</p> <ol style="list-style-type: none"><li data-bbox="815 387 1465 629">1. una especie distintiva de alta importancia biológica está bajo alta amenaza por explotación/enfermedad u otra amenaza identificable en la que incluso la información general de la localidad puede amenazar al taxón.<li data-bbox="815 656 1465 857">2. la información contenida en la documentación es de tal naturaleza que su divulgación podría causar un daño irreparable al medio ambiente, a una persona o a alguna otra entidad. <p>Los datos sólo pueden suministrarse bajo estrictas condiciones de Licencia o como presencia en una gran región, como una cuenca hidrográfica, un condado o una región biogeográfica.</p>

Criterio	Razonamiento
<p>Categoría 2</p> <p>Especies o registros cuyas coordenadas estarán disponibles públicamente "desnaturalizadas" (a 0.1 grados) y/u otra información en el registro está generalizada. Los datos a escala más fina (Categoría 3, Categoría 4 o datos detallados) pueden ser suministrados a individuos bajo Licencia.</p>	<p>Las razones de la restricción son que:</p> <ol style="list-style-type: none"> 1. La especie está clasificada como altamente sensible y la provisión de ubicaciones precisas potencialmente sometería a la especie a amenazas como la perturbación y la explotación. 2. El registro incluye información altamente sensible, cuya divulgación podría causar un daño extremo a una persona o al medio ambiente. <p>Los datos se suministran al público:</p> <ol style="list-style-type: none"> 1. con la georreferencia desnaturalizada a 0.1 grados (~10 km) y/o 2. Con los campos sensibles generalizados o eliminados y sustituidos por un texto adecuado <p>Los datos pueden suministrarse a escalas más finas si se solicitan bajo las condiciones de un acuerdo de datos por escrito, normalmente un "Acuerdo de licencia de datos". Cuando se proporcionen datos a los clientes, se les informará qué especies o campos son sensibles y podrían tener sus coordenadas desnaturalizadas a aquellas disponibles bajo la Categoría 3 o Categoría 4.</p> <p>NB: En el caso de que la sensibilidad sea provocada por campos distintos a la georreferencia, puede ser más apropiado clasificar el registro como Categoría 3 o Categoría 4.</p>

Criterio	Razonamiento
<p>Categoría 3</p> <p>Especies o registros cuyas coordenadas estarán disponibles públicamente "desnaturalizadas" (a 0.01 grados) y/u otra información en el registro está generalizada. Los datos de escala más fina (Categoría 4 o datos detallados) pueden suministrarse bajo licencia.</p>	<p>Las razones de la restricción son que:</p> <ol style="list-style-type: none"> 1. La especie está clasificada como de sensibilidad media a alta y la provisión de ubicaciones precisas podría someter a la especie a amenazas como la perturbación y la explotación. 2. El registro incluye información sensible, cuya divulgación podría causar daños a una persona o al medio ambiente. <p>Los datos se suministran al público:</p> <ol style="list-style-type: none"> 1. Con la georreferencia denaturada a 0.01 grados (~1km) y/o 2. Con campos sensibles generalizados o eliminados y sustituidos por un texto adecuado <p>Los datos pueden suministrarse a escalas más finas si se solicitan bajo las condiciones de un acuerdo de datos por escrito, normalmente un "Acuerdo de licencia de datos". Cuando se proporcionen los datos a los usuarios, se les informará qué especies o campos son sensibles y podrían tener sus coordenadas desnaturalizadas a aquellas disponibles bajo la Categoría 4.</p> <p>NB: En el caso de que la sensibilidad sea provocada por campos distintos a la georreferencia, puede ser más apropiado clasificar el registro como Categoría 4.</p>

Criterio	Razonamiento
<p>Categoría 4</p> <p>Especies o registros cuyas coordenadas estarán disponibles públicamente "desnaturalizadas" (a 0.001 grados) y/u otra información en el registro es generalizada. Los datos "como están mantenidos" pueden ser suministrados a individuos bajo Licencia.</p>	<p>Las razones de la restricción son que:</p> <ol style="list-style-type: none"> 1. La especie está clasificada como de sensibilidad baja a media y el suministro de ubicaciones precisas podría dar lugar a riesgos de colecta o daños deliberados. 2. El registro incluye información sensible, cuya divulgación podría causar daños a una persona o al medio ambiente. <p>Los datos detallados pueden suministrarse bajo las condiciones de un acuerdo de datos por escrito, normalmente un "Acuerdo de Licencia de Datos". Cuando se faciliten los datos a los usuarios, se les indicará qué especies o campos son sensibles.</p>

3. Generalización de la información textual

En algunos casos, la información de los campos de texto puede considerarse sensible en determinadas circunstancias. Esto puede incluir información como:

- Nombres de personas vivas
- Información de la localidad
- La fecha de colecta
- El número del colector
- Hábitat
- Información del terrateniente
- Nombres taxonómicos

En algunos casos puede ser necesario restringir los análisis de correlación que pueden conducir a deducciones sobre las localidades de los registros que están restringidos o generalizados, por ejemplo, el nombre del colector, la fecha y los números del colector en secuencia. En otros casos, puede ser necesario ocultar el nombre de un taxón en una lista de colectas en un hotspot de biodiversidad o en una localidad sensible.

Dichas restricciones no deben restringir la publicación del registro en su totalidad. Los datos que deben ocultarse pueden eliminarse y sustituirse por un texto adecuado (ver más adelante) o generalizarse, por ejemplo, limitándose a dar el nombre de un rango taxonómico de nivel superior en el que la especie debe restringirse.



Siempre que los datos de un campo textual se restrinjan o se generalicen para su distribución (como el nombre de un colector, la información textual de la localidad, etc.), deben documentarse sustituyéndolos por un texto adecuado, el campo no debe dejarse en blanco o nulo.

Entre los ejemplos de texto de sustitución se encuentran:

nombre eliminado por razones de privacidad

Este espécimen representa una especie en peligro o amenazada. La localidad específica se ha eliminado del registro en línea para proteger a esta especie de la colecta excesiva. Estos datos pueden facilitarse a los investigadores que lo soliciten.

Este espécimen representa una especie en peligro o amenazada. La localidad específica se ha generalizado a la presencia dentro de una cuadrícula de 0.1 grados de resolución. Se pueden suministrar datos detallados a los investigadores que lo soliciten.



Cuando sea necesario restringir un nombre taxonómico (por ejemplo, de taxa sensibles como parte de un estudio o listado), puede ser posible sustituirlo por un nombre de la taxonomía superior (género/familia, etc.) o simplemente informar de que hay "x" taxa sensibles presentes sin proporcionar nombres.

En ocasiones, los proveedores de datos pueden querer restringir la información de registros relacionados con un registro sensible (además del propio registro sensible), como el nombre y los números del colector en una secuencia de registros colectados en el mismo lugar y momento que un registro sensible, con el fin de reducir la posibilidad de que se encuentre el registro sensible mediante un análisis de correlación. Sin embargo, si se elimina el nombre y el número del colector únicamente del registro sensible y no de los demás registros, es poco probable que éstos puedan deducirse, a menos que el buscador de la información ya tenga conocimiento de contexto sobre la colecta. Por este motivo, y otros, se recomienda no restringir los datos de los registros relacionados.

4. Generalización de la información espacial

4.1. Generalización versus aleatorización

Algunas de las personas que respondieron a la encuesta en línea (Chapman 2006) registraron que **aleatorizaban** los datos en lugar de **generalizar**. Las razones para no **aleatorizar** incluían el trabajo y el cálculo extra que suponía, el aumento de la posibilidad de cometer errores y la menor fiabilidad que los usuarios podrían depositar en los datos.

Algunos de los encuestados afirmaron que se sentían cómodos mostrando la presencia/ausencia de datos sensibles dentro de grandes polígonos o cuadrículas, porque seguían reflejando los datos reales, pero se horrorizaban ante la idea de "falsear" deliberadamente las coordenadas de los puntos de manera que las ubicaciones parezcan representaciones precisas, pero estén desplazadas aleatoriamente de los datos reales, es decir, que representen la introducción deliberada de errores. Mientras que la **generalización** crea/conserva datos "verdaderos", la **aleatorización** crea datos deliberadamente "falsos".

En el campo de la minería de datos y de la protección a la privacidad (incluida la privacidad individual en los datos de censos), se considera generalmente que la **generalización** ascendente ("bottom-up") es mucho más práctica y científicamente justificable que la **aleatorización** (ver, por ejemplo, Wang et al. 2004, Dalvi and Keole 2015).

Otras ventajas de la **generalización** es que:

- se amplía, permitiendo el uso de una metodología coherente a diferentes escalas
- se puede configurar para entregar a diferentes personas diferentes resoluciones, dependiendo de los roles establecidos, etc.
- puede proporcionar sencillamente diferentes escalas de generalización para diferentes categorías de sensibilidad
- es más fácil de aplicar para quienes tienen pocos conocimientos tecnológicos



La **aleatorización** es una metodología que recomiendo encarecidamente que **no** se utilice.

4.2. Generalización

Uno de los requisitos más comunes para generalizar la información sensible sobre biodiversidad es generalizar la localización espacial o las **coordenadas geográficas** (Chapman y Wiczorek 2020). Tradicionalmente esto se ha hecho de muchas maneras y ha habido poca consistencia en las metodologías y muy poca documentación sobre lo que se ha hecho en cada caso. Esto ha reducido considerablemente el valor de los datos para el análisis y a menudo los usuarios no son conscientes de que los datos han sido modificados.

La **generalización** (al menos en un sentido espacial) suele ser de uno de dos tipos, concretamente:

- Generalización a una grilla (métrica o geográfica)

- Generalización a un polígono (región sociopolítica, país, región biogeográfica, cuenca hidrográfica)

Muchas de las personas que respondieron a la encuesta ([Chapman 2006](#)) defendieron la simplicidad de la generalización a una grilla. Las razones que se dieron fueron, entre otras, las siguientes:

- la facilidad de poder variar la escala para diferentes categorías de sensibilidad,
- la facilidad de mantenimiento y entrenamiento, y
- la facilidad de crear una documentación adecuada.

La generalización a una cuadrícula, al tiempo que protege las ubicaciones exactas de los taxa sensibles, también proporciona datos en un formato que sigue siendo utilizable para la mayoría de los usuarios, especialmente cuando se utiliza una cuadrícula estándar.

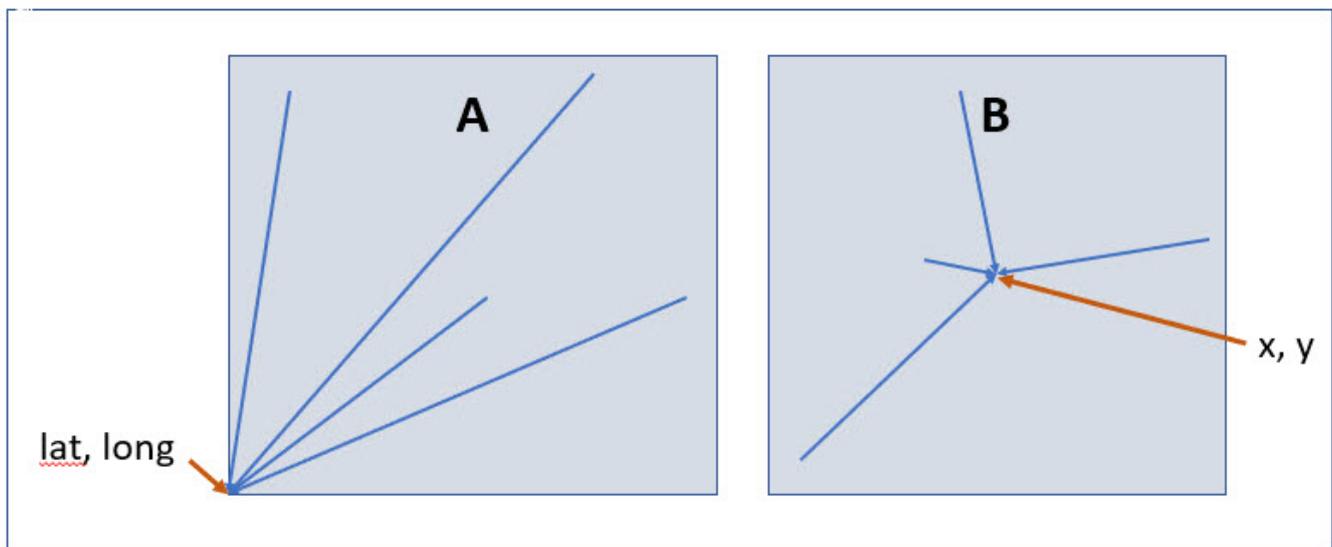


Figura 1. Dos métodos de generalización: a) una grilla geográfica en la que todos los registros se refieren a la esquina inferior izquierda (SE) b) una cuadrícula métrica en la que todos los registros se refieren al centroide.

La [figura 1](#) muestra dos métodos que se utilizan habitualmente para grillar los datos. El primero (recomendado aquí) es una grilla geográfica (p. ej., grillas cartográficas como las basadas en coordenadas geográficas), el segundo, una grilla métrica (p.ejm. Europe Equal Area Grid 2001, [EPSG:19986](#)). Para conocer cómo debe tratarse cada una de estas con respecto a la determinación de la ubicación y la incertidumbre, ver [Chapman y Wieczorek \(2020\)](#). Recomendamos el uso de una grilla geográfica, como se discute a continuación, debido a la facilidad de preparación y documentación y porque los datos biológicos que se comparten a través del estándar Darwin Core ([Wieczorek et al. 2012](#)) tendrían que ser convertidos de una grilla métrica a una grilla geográfica antes de la publicación, con la consiguiente pérdida de precisión. Sin embargo, como se señala en [Chapman y Wieczorek \(2020\)](#),

utilizar la esquina suroeste como coordenada para una georreferencia radio-punto es un desperdicio, ya que el radio geográfico iría desde allí hasta la esquina más lejana, lo que supondría el doble de distancia que si se utilizara en su lugar el centro de la grilla de celdas. En cualquier caso, las características de la cuadrícula deben registrarse con la información de la localidad.

Cuando los datos se generalizan a una región geográfica o biogeográfica (un polígono), los datos tienen menos utilidad para muchos análisis, pero fueron considerados por muchos como una forma más segura de "ocultar" ubicaciones de datos sensibles. Actualmente, GBIF.org sólo tiene una capacidad limitada para incorporar datos de polígonos. Este método guarda cierto paralelismo con la comunicación de los resultados del censo en muchos países, donde los resúmenes se comunican utilizando áreas locales estadísticas o secciones de censo para restringir la posible identificación de los individuos ([Australian Bureau of Statistics 2006](#)). Una diferencia clave es que con los datos del censo los resultados se resumen sobre muchos individuos dentro de una región, mientras que con los datos biológicos queremos ocultar la ubicación de un solo ente dentro de un área. De hecho, si se produce un resumen, pero esta no es la intención principal. Uno de los problemas de este método es que no hay garantía de que las fronteras políticas (o incluso biogeográficas) se mantengan constantes a lo largo del tiempo, lo que reduce aún más el valor de los datos para muchos fines. Ya se ha comprobado que esto es un problema al comparar los datos del censo a lo largo del tiempo cuando los distritos o tramos de censo han cambiado. ([Noble et al. 2011](#)).

Si se dan georreferencias para datos que se generalizan a una región biogeográfica o política, el resultado puede ser bastante engañoso: una especie costera, por ejemplo, puede acabar con una georreferencia que está a cientos de kilómetros tierra adentro, lo que reduce la utilidad para el análisis o la limpieza de datos. Poner a disposición estos datos sin la documentación adecuada puede llevar a resultados bastante desastrosos para los usuarios. En algunos casos es mejor no suministrar una georreferencia, pero si se suministra una, la documentación debe ser clara, con un radio de incertidumbre grande o un ajuste espacial bien documentado (ver la discusión en [Ajuste espacial](#) más abajo).

Good practice dictates that whatever you do to generalize the data that you document it so that users of the data know what reliance they can place in it. Note, that when documenting what has been done, it is essential that both the coordinates and the coordinate uncertainty in meters be recorded. How this is best done can be seen in the Georeferencing Quick Reference Guide ([Zermoglio et al. 2020](#)).

Se recomienda que los proveedores de datos que generalicen sus datos lo hagan utilizando una metodología estándar (ver más abajo) y que lo documenten debidamente. Dado que la mayoría de los datos de biodiversidad se ponen a disposición de los usuarios utilizando [grados decimales](#) ([Chapman and Wieczorek 2020](#)), el método recomendado significa que protocolos como el Darwin Core ([Wieczorek et al. 2012](#)) no necesitan ser modificados, salvo para permitir una documentación adecuada de los metadatos.

El método recomendado a continuación permite varios niveles de [generalización](#) que se ajustan a las Categorías 1-4 descritas en la [Tabla 1](#).

El método recomendado para la **generalización** es:

Tabla 7. Método de generalización de coordenadas geográficas.

Categoría	Sensibilidad	Georreferencia
Categoría 1	Extrema	Las coordenadas geográficas no se publican o los datos pueden publicarse por cuenca hidrográfica/biorregión/condado, redondeadas a 1 grado, etc.
Categoría 2	Alta	Coordenadas geográficas redondeadas a 0.1 grados
Categoría 3	Media	Coordenadas geográficas redondeadas a 0.01 grados
Categoría 4	Baja	Coordenadas geográficas redondeadas a 0.001 grados
No sensible	No sensible	Coordenadas geográficas sin restricción

El South African National Biodiversity Institute, al aplicar en gran medida los criterios establecidos en el documento anterior (**Chapman y Grafton 2008**), ha implementado sólo dos categorías (**SANBI 2010, SANBI 2016**): los datos originales no generalizados y los datos generalizados a un cuarto de grado (0.25 grados (QDS)). Creemos que esta es una generalización más difícil de implementar y, a menos que esté completa y claramente documentada, podría llevar a un nivel de precisión engañoso. Además, reduce la flexibilidad que ofrece el método de cuatro categorías anterior.

4.3. Documentación

Es importante documentar el método y el nivel de **generalización** para que los usuarios sean conscientes de lo que se ha hecho con los datos y de la fiabilidad que pueden depositar en ellos. Actualmente, ni **Darwin Core** ni los protocolos **ABCD** proporcionan campos para los metadatos recomendados. Se ha propuesto que estos protocolos se modifiquen para que acepten dichos metadatos (véase **Epílogo**), pero mientras tanto, se recomienda que la información se registre utilizando los términos existentes de **Darwin Core** a nivel de registro (p. ej., **dwc:informationWithheld**, **dwc:dataGeneralizations** o cualquiera de los campos "Remarks").

En cuanto a la **generalización** de los datos de georreferenciación es importante registrar que los datos se han generalizado utilizando una "grilla en grados geográficos decimales" y registrar ambas:

- La precisión de los datos proporcionados (p. ej., 0.1 grados, 0.001 grados, etc.)
- La precisión de los datos almacenados o retenidos (p. ej., 0.0001 grados, 0.1 minutos, 1 segundo, 100m cuadrados, etc.)

Las recomendaciones de metadatos para su inclusión en el **Elemento de ubicación Darwin Core** (TDWG 2018) se exponen en el **Epílogo**. Una vez que se hayan adoptado (o algo similar), se recomienda registrar los campos adecuados y distribuirlos con los datos.

4.4. Duplicados y GUID

Con las plantas especialmente y con otros taxa (como los insectos), los colectores suelen reunir varios especímenes (duplicados o partes de conjuntos). Normalmente del orden de cuatro a seis, aunque se han citado ejemplos de más de 80 (Paul Morris 2007, comunicación personal, abril). Estos duplicados o partes de conjuntos suelen enviarse a muchas instituciones de todo el mundo. Un problema que surge es que las instituciones de origen pueden perder el control de lo que ocurre con la información (incluida la información sobre la localidad) distribuida a las colecciones de esas instituciones secundarias, recordando que los duplicados pueden haber sido distribuidos antes de que el taxón sea identificado como sensible.

En la mayoría de los casos, este intercambio de información no supone un problema, pero en el caso de los taxa sensibles suele serlo. Es posible que la institución secundaria no sepa qué se considera "taxa sensibles" en la jurisdicción de la institución de origen o que no haya señalado esa información. La sensibilidad no siempre es una información que pueda distribuirse junto con las colectas, ya que puede que no se sepa hasta mucho más tarde que la especie está en peligro y/o es sensible. Esta cuestión es difícil, ya que etiquetar simplemente un taxón como sensible puede no ser la respuesta: un taxón puede estar en peligro de extinción en su zona nativa (y por tanto ser sensible) y puede ser una hierba mala o una plaga en otras zonas, siendo la información sobre la localidad importante para su control en ambos casos.

La identificación de duplicados entre instituciones no es fácil, pues, especialmente en el caso de las colecciones históricas y heredadas, a menudo es difícil determinar los especímenes duplicados. Algunas instituciones, como [Centro de Referência em Informação Ambiental](#) (CRIA) en Brasil en su proyecto [speciesLink](#) y el [Atlas of Living Australia](#), utilizan la coincidencia a través de una serie de campos como el número de colector, la fecha y la localidad, mientras que GBIF está desarrollando un algoritmo para el [agrupamiento de datos](#). Sin embargo, actualmente no se dispone de un sistema global universal. El uso de identificadores únicos, persistentes y resolubles, Identificadores Globales Únicos (GUID) ([Page 2009](#), [Richards 2010](#), [Richards et al. 2011](#)) ayudará a estos procesos a largo plazo, pero la implementación de GUID a nivel de espécimen todavía parece estar lejos. Una publicación reciente de [Nelson et al. 2018](#) hace una serie de recomendaciones sobre la acuñación, la gestión y el intercambio de GUID para especímenes de herbario, pero hasta que estas técnicas se adopten más ampliamente, la identificación de duplicados entre instituciones seguirá siendo un problema.

5. Documentación y metadatos

Es importante que los datos estén documentados con precisión para que los usuarios y otras personas sepan exactamente lo que representan los datos y la confianza que se puede depositar en ellos. Por ejemplo, un usuario necesita la información para determinar si los datos son adecuados para el análisis que va a realizar. Muchos proveedores de datos señalaron en la encuesta ([Chapman 2006](#)) que una de las razones por las que se mostraban reticentes a publicar algunos de sus datos era el temor a que se hiciera un mal uso de ellos. Si los datos no están adecuadamente documentados, la probabilidad de que se produzca un mal uso inadvertido aumenta considerablemente, ya que el usuario puede utilizar los datos en un análisis pensando erróneamente que está obteniendo registros puntuales exactos, cuando en realidad los datos se han generalizado a una cuadrícula de 10 km y podrían estar en cualquier lugar a una zona de 100 km cuadrados. Si se ejecuta un algoritmo de modelado del clima, por ejemplo, este tipo de error podría dar lugar a un resultado bastante engañoso. Solamente por esta razón, es importante para los proveedores de datos, los usuarios de los mismos y los usuarios finales (como los gestores medioambientales, los responsables políticos, etc.) que los datos se describan con precisión.

En particular, debe haber una documentación clara de las restricciones de acceso que podría incluir, por ejemplo, una indicación de qué partes de los datos son sensibles (si las hay), las razones de la sensibilidad y las condiciones en las que es posible la liberación.

5.1. Documentación de la sensibilidad

Los metadatos cumplen una función esencial en cuanto a la comunicación a terceros de las restricciones de acceso y las condiciones de uso que los generadores de datos pretenden dar a sus datos. Puede considerarse como una "ayuda" en la protección de los datos y la información, pues permitirá a los usuarios del sistema visualizar las condiciones establecidas por el generador de datos para el acceso y uso de la información. Adicionalmente, en caso de que los datos no sean accesibles, los metadatos permiten conocer las condiciones de acceso a través de otros medios (digitales o no) así como un resumen del contenido. ([Llinás 2005](#))

El término metadatos se ha utilizado generalmente para referirse a la documentación de un conjunto de datos por completo. Los comentarios se refiere solo a la documentación a nivel de registro. Sin embargo, es preferible llamar a ésta documentación **metadatos a nivel de registros** para formalizar el proceso. En el capítulo anterior se recomendó que, cuando los datos se generalicen para su distribución, se documente el nivel de **generalización**, por ejemplo, que los datos se han generalizado utilizando una grilla geográfica decimal y que se registre tanto la precisión de los datos proporcionados como la precisión de los datos "como están mantenidos" o almacenados. Asimismo, en el capítulo sobre **Determinación de la sensibilidad**, se recomendaron una serie de procesos de documentación ([Tabla 1](#)). Algunos de ellos pueden ser más apropiados para documentar las razones por las que se considera un taxón como potencialmente sensible desde el punto de vista medioambiental (Criterios 1 y 2), mientras que los otros (Criterios 3 y 4) son apropiados para los datos en sí mismos y forman parte de los metadatos más amplios a nivel de

registro. Sin embargo, para documentar por completo las razones para restringir los datos, puede ser necesario heredar la documentación de los Criterios 1 y 2 al nivel de registro, por ejemplo, la razón por la que se restringen los datos puede incluir que el taxón está sujeto a una actividad humana perjudicial.

Por el momento, ni el estándar **Darwin Core** ni el **ABCD** tienen campos para registrar el tipo de metadatos a nivel de registro que se recomienda aquí. Se han hecho una serie de recomendaciones a Biodiversity Information Standards (TDWG) para la inclusión de campos adicionales en el **Elemento de ubicación del Darwin Core** (TDWG 2018) y se enumeran abajo en el **Epílogo**.

Hasta que se modifiquen estos estándares y protocolos, se recomienda documentar los datos utilizando los actuales **términos del Darwin Core** a nivel de registro (p.ej., **dwc:informationWithheld**, **dwc:dataGeneralizations** o cualquiera de los campos "Remarks") y, en la medida de lo posible, registrar el mismo tipo de información que se incluiría en los campos recomendados, por ejemplo, que:

- Los datos son sensibles
- Las principales razones por las que los datos se consideran sensibles (ver los Criterios 1-4, **Tabla 1**) junto con la justificación
- La fecha en la que se debe revisar la sensibilidad de los datos
- La precisión de los datos disponibles
- La precisión de los datos originales almacenados o retenidos

Al documentar las restricciones de acceso en los metadatos a nivel del conjunto de datos, se puede incluir algo como:

Este conjunto de datos sólo está disponible para el público en una resolución limitada por la siguiente razón. Parte de la información contenida en este conjunto de datos se refiere a especies que son vulnerables a las perturbaciones o a los prejuicios humanos. Dos especies (*Adelanthus lindenbergianus*, *Athalamia hyaline*) son significativamente vulnerables a la colecta. El detalle completo de esta información sensible puede ponerse a disposición, bajo licencia, a organizaciones y personas específicas que necesiten conocerla para evitar daños al medio ambiente. Para más información, póngase en contacto con el proveedor.

5.2. Ajuste espacial

El ajuste espacial (Chapman y Wieczorek 2006, 2020) es un concepto que ha surgido del proyecto BioGeomancer (Guralnick et al. 2006) y proporciona una medida de qué tan bien coincide una representación geométrica con la representación espacial original. El ajuste espacial tiene un valor de cero, uno o mayor que 1, donde 1 representa una coincidencia exacta (es decir que los datos no se han generalizado). Los detalles sobre cómo se puede calcular el ajuste espacial se pueden encontrar en Chapman y Wieczorek (2006, 2020), donde también se puede encontrar el siguiente resumen:

Un ajuste espacial con un valor de 1 es una coincidencia exacta o un solapamiento del 100%. Si la geometría dada no abarca completamente la representación espacial original, entonces el ajuste espacial es cero (es decir, parte del original queda fuera de la versión transformada, lo que interpretamos como que no hay ajuste). Si la forma transformada abarca completamente la representación espacial original, el valor del ajuste espacial es la relación entre el área de la geometría transformada y el área de la representación espacial original. Caso especial: Si la representación espacial original es un punto y la geometría presentada no es un punto, entonces el ajuste espacial es indefinido.

Con respecto a la generalización de los datos, el ajuste espacial puede considerarse como un Coeficiente de Generalización (es decir, hasta qué punto se ha generalizado un registro) una relación entre el área generalizada y el área real.

Ajuste espacial = Área generalizada/Área real

Un ejemplo de su aplicabilidad es cuando se pone a disposición una georreferencia con un radio de incertidumbre de 1 km (utilizando el método de radio-punto (Wieczorek et al. 2004)) utilizando una grilla de 10 km² (que cubre completamente la incertidumbre). En este caso el **ajuste espacial** sería mayor que 1 pues representa un área mayor que la incertidumbre real (como se muestra en la **Figura 2**).

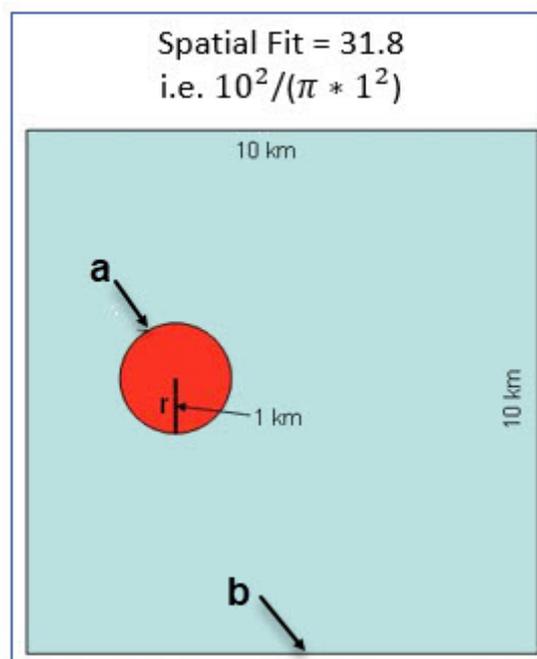


Figura 2. Ejemplo de cálculo del ajuste espacial para una colecta con un radio de incertidumbre de 1 km (círculo rojo) y que se distribuye utilizando una cuadrícula de 10 km² (azul). Ajuste espacial = 31.8.

Cuanto más pequeño sea el tamaño de la grilla, más se acercará el **ajuste espacial** a '1'. Note que un registro que tenga su georreferencia aleatorizada o generalizada de forma que una parte del radio de incertidumbre quede fuera del cuadrado de la cuadrícula tendría un **ajuste espacial** igual a cero.

6. Autenticación y autorización

Como se recomendó en el taller de expertos y como señalaron muchos en la encuesta en línea (Chapman 2006), la responsabilidad de determinar quién puede o no tener acceso detallado a datos sensibles, posiblemente mediante el uso de un inicio de sesión seguro o de acuerdos de licencia de datos únicos, debe recaer en los proveedores de datos.

También se acordó en el taller que no es el papel de GBIF gestionar la identificación, verificación o autorización de los usuarios, ni controlar la autenticación o el inicio de sesión en el "Portal de datos", pero puede tener un papel en el suministro de orientación y de un método de autenticación adecuado a los Nodos.

En el taller de expertos se informó que las cuestiones técnicas relacionadas con la autenticación de un grupo o individuo, y el uso de roles, etc. no es una tarea difícil. Existen varios protocolos y sistemas de trabajo en uso bien establecidos para la autenticación y éstos podrían adaptarse fácilmente para ser utilizados por los proveedores de datos.

El principal problema es determinar quiénes deben ser los usuarios autorizados y cómo determinar quiénes son usuarios de buena fe y quiénes no. Se trata de una cuestión difícil que habrá que estudiar con el tiempo. No es algo que se pueda recomendar en este documento de buenas prácticas. Sin embargo, el informe anterior (Chapman 2007) hizo una serie de recomendaciones sobre cómo se puede seguir explorando esta cuestión.

7. Implementaciones

Desde la publicación de la *Guía* (Chapman y Grafton 2008), varios países y jurisdicciones estatales y locales han desarrollado sus propias políticas de datos sensibles sobre taxa. Las jurisdicciones suelen tener circunstancias individuales (legales, éticas, prácticas, etc.) que pueden dictar la forma en que deben aplicarse sus propias políticas. A la hora de desarrollar sus propias políticas, además de basarse en este documento, es posible que desee consultar esas políticas jurisdiccionales. Sin embargo, advierto que no hay que alejarse demasiado de los criterios y principios expuestos en este documento, debido a las complicaciones que pueden surgir a la hora de compartir datos, especialmente a través de iniciativas de agregación y publicación como GBIF. El objetivo principal de este documento es fomentar la adopción y el uso de metodologías estandarizadas. La mayoría de los documentos jurisdiccionales irán, necesariamente, más allá de este documento, que sólo se basa en la generalización. Este documento no ahonda en otros aspectos como la privacidad de las personas, la concesión de licencias, etc. Cualquier trabajo derivado debe ser de acceso público y establecer claramente las políticas con respecto a la publicación y el intercambio de datos. Cuando las políticas difieran de las aquí documentadas, deberán señalarse dichas diferencias junto con las explicaciones de los motivos.

El **South African National Biodiversity Institute** (SANBI) ha enumerado cuatro tipos de datos sobre taxa sensibles que pueden necesitar ser restringidos (SANBI 2010):

1. **Datos sobre tamaños poblaciones** o el número de un taxón sensible o de sus poblaciones que puedan influir en el valor de rareza o en el valor comercial del taxón;
2. **Datos sobre el hábitat o el ecosistema** de un taxón sensible que puedan permitir que se infiera la localidad de un taxón;
3. **Datos georreferenciados** sobre un taxón sensible (incluyendo datos de ubicación precisos, descripciones de lugares y/o localidades y coordenadas de localidades puntuales) que puedan permitir la localización de las poblaciones;
4. **Registros de especímenes** en colecciones o registros de observación (incluyendo el nombre del colector, el número de espécimen del colector, la identidad del taxón, la descripción de la localidad, coordenadas más precisas que un cuadrado de cuarto de grado, el tamaño poblacional, la fecha de colecta, el colector del espécimen y cualquier información sobre el hábitat asociada con el espécimen), que con el análisis pueden permitir la localización de una población. El registro puede referirse a un solo espécimen, o a una muestra, que incluye varios o muchos especímenes con detalles idénticos de colecta. Todos los registros de especímenes/observaciones de un taxón sensible, independientemente de cuándo se hayan colectado, estarán igualmente restringidos.

SANBI no ha adoptado las categorizaciones recomendadas aquí, adoptando sólo un nivel de generalización en lo que ellos llaman QDS (Quarter Degree Square) o unos 25 km por 25 km (ver discusión en SANBI 2016).

El **Atlas of Living Australia** (ALA) ha desarrollado una política que clasifica los taxa sensibles en tres categorías (ALA 2018a):

- conservación, p. ej., las ubicaciones de especies amenazadas
- bioseguridad, p. ej., avistamientos no verificados de plagas no registradas anteriormente en

Australia

- privacidad, p. ej., especímenes colectados en propiedad privada, nombres de colectores, etc.

Esta política se basa en un conjunto de principios directivos (Tann y Flemons 2009):

- facilitar el acceso a los datos científicos
- minimizar el daño restringiendo explícitamente el acceso a la información sensible seleccionada
- ayudar a las autoridades estatales y federales a cumplir con sus obligaciones en materia de libertad de información
- ayudar a las autoridades estatales y federales con las obligaciones comerciales internacionales de Australia
- ayudar a compartir los datos mantenidos por un custodio, cuando existe un acuerdo o una expectativa de que estos datos no se utilizarán indebidamente
- respeto por las diferencias de enfoque de los datos sensibles en las jurisdicciones de Australia
- respeto por la privacidad y las restricciones a la información personal

Desgraciadamente, hasta la fecha, cada Estado Australiano tiene políticas individuales diferentes sobre qué taxa se consideran sensibles y sobre cómo se comparten los datos sobre los taxa sensibles. Se está debatiendo la elaboración de una lista nacional de taxa sensibles, pero parece que aún no se ha llegado a un acuerdo sobre una política coherente para compartir esos datos.

El **Australian Department of the Environment** (Australian Government 2016) ha adoptado los niveles de generalización expuestos en este documento y el documento de política ofrece una serie de ejemplos bien elaborados. La política incluye algunos principios directivos adicionales, entre ellos:

- **Ser responsable:** la decisión de restringir el acceso a los datos debe ser justificable, coherente y repetible y cumplir con la legislación, la normativa o la política pertinente.
- **Decisiones tomadas lo más cerca posible de la fuente:** el custodio de los datos debe tener la responsabilidad de determinar si los datos ecológicos deben clasificarse como sensibles.
- **Retener los datos originales:** los custodios de los datos deben retener una versión original, sin alterar, de los datos ecológicos y salvaguardar esta versión original.
- **Transparencia:** la documentación debe estar vinculada a los datos y debe estar disponible para todos los usuarios de los datos. La documentación garantiza que los posibles usuarios de los datos comprendan qué datos existen, por qué se han clasificado como sensibles y cómo se han alterado o protegido.
- **Respetar las restricciones de los conjuntos de datos:** los custodios de los datos no deben liberar datos que no han sido procesados de acuerdo a esta política
- **Revisión a lo largo del tiempo:** los custodios de los datos deben realizar revisiones periódicas (cada 2-5 años) de los conjuntos de datos para determinar si su contexto ha cambiado. Lo que actualmente se considera un dato sensible puede no serlo en el futuro.

Un principio en el que hace hincapié esta política es que si los datos sobre las sensibilidades identificadas (p. e.j., la ubicación) ya están disponibles públicamente, es poco probable que los datos puedan considerarse sensibles.

La **New South Wales Office of Environment and Heritage** (DECCW 2007, OEH 2019a) señala que "algunas especies amenazadas son muy sensibles a las perturbaciones y a la explotación. La información sobre la ubicación de estas especies se considera "sensible" y la OEH no la facilitará a terceros, con algunas excepciones limitadas. La información precisa sobre la ubicación de las especies sensibles está exenta de las solicitudes de libertad de información". Su política de datos sensibles tiene tres niveles de clasificación:

Categoría 1: especies de gran importancia biológica, de las que no se facilitará ningún registro. La razón de la no divulgación es que la especie está muy amenazada por explotación/enfermedad u otras amenazas identificables, e incluso la información general de la localidad puede amenazar al taxón. El famoso pino Wollemi entra en esta categoría.

Categoría 2: especies consideradas en grave riesgo por amenazas como la perturbación o la explotación. Para las especies de esta categoría, las coordenadas geográficas de los avistamientos se suministrarán "desnaturalizadas", con el fin de generalizar la localidad. Se pueden conceder excepciones a esta norma a algunos organismos gubernamentales o para determinados fines de investigación.

Categoría 3: Especies consideradas bajo un riesgo medio o alto de amenazas como la perturbación o la explotación. En el caso de las especies de esta categoría, las coordenadas se suministrarán con una precisión "como están mantenidos" a los clientes con licencia, pero por lo demás se suministrarán "desnaturalizadas".

Incluyen 3 categorías básicas de "desnaturalización"(OEH 2011) que coinciden en gran medida con lo que se recomienda aquí:

- **Categoría 1:** no hay registros proporcionados
- **Categorías 2:**
 - registros desnaturalizados a 0.1° (~10 km) para aplicaciones web públicas
 - registros desnaturalizados a 0.01° (~1 km) para provisión a clientes con licencia
- **Categorías 3:**
 - registros desnaturalizados a 0.01° (~10 km) para aplicaciones web públicas
 - registros proporcionados "como está mantenido" a clientes con licencia

En el sitio web de la agencia se mantiene y publica un listado de especies sensibles, la categoría de sensibilidad y las razones por las que se consideran sensibles (OEH 2019a).

La **UK National Biodiversity Network** (NBN 2019b) utiliza criterios que permiten diferentes categorías de generalización en los distintos países miembros del Reino Unido (por ejemplo, una especie puede ser catalogada como sensible en Gales, pero no en Escocia). Los registros se envían al

Atlas NBN con la mejor resolución de captura. La ubicación de las especies sensibles se generaliza y los datos generalizados se ponen a disposición del público bajo una licencia Creative Commons según determine el proveedor de los datos. Los **Listados de taxa sensibles**, junto con las razones de la sensibilidad, y el nivel de generalización para cada uno de Inglaterra, Escocia y Gales se mantienen y publican.

Birdlife Australia (Birdata) ha desarrollado una política de Especies Sensibles basada en los principios y categorías de generalización expuestos en esta Guía y en la anterior (**Chapman y Grafton 2008**). Han identificado seis categorías de datos en las que puede ser necesario generalizar las localidades:

- Caza furtiva de especies silvestres y comercio internacional
- Cazadores legales e ilegales incluyendo trofeos, recreativos, comerciales y deportivos
- Colectores de huevos
- Captura ilegal de aves silvestres para el comercio de jaulas y la cetrería
- Los entusiastas de la vida silvestre que exhiben comportamientos intrusivos, especialmente a las especies territoriales
- Intrusión/acceso a la propiedad privada o a las zonas indígenas protegidas sin permiso.

El **US Forest Service** tiene una política para garantizar poblaciones viables de especies sensibles en toda su área de distribución geográfica. Una vez que se cumplan los objetivos y la viabilidad deje de ser una preocupación, las especies dejarán de tener el estatus de "sensibles" (**US Forest Service 2005**). Las especies sensibles son aquellas especies de plantas y animales identificadas por el silvicultor regional para las que la viabilidad de la población es una preocupación en las tierras del National Forest Service (NFS) dentro de la región. El objetivo del "Programa de especies sensibles del servicio forestal" es garantizar que el número de especies y la distribución de la población sean adecuados para que no se requiera la inclusión en la lista federal y no se produzca la extirpación en las tierras del NFS (**US Forest Service 2016**).

No se hace mención específica a las diferentes categorías, ni a la generalización de la información de ubicación para el público. Sin embargo, según **Hartter et al. (2013)**, el servicio forestal de los Estados Unidos trata de proteger los lugares de investigación no divulgando las referencias geoespaciales junto con sus datos.

Natural Resources Canada y GeoConnections Canadá encargaron un estudio para desarrollar las "Buenas prácticas para compartir datos geoespaciales ambientales sensibles" (**AMEC Earth and Environmental 2010**). Las directrices consideran que los datos geoespaciales medioambientales son "datos geoespaciales temáticos que podrían utilizarse para el análisis en ámbitos como la evaluación del impacto medioambiental, la planificación del uso del suelo, la gestión de la tierra, el desarrollo sostenible, la gestión de recursos, la gestión de cuencas hidrográficas, etc.". El documento enumera cinco criterios para determinar la sensibilidad. El tercer criterio incluye los datos considerados en este documento:

Protección de los recursos naturales: el uso de la información puede provocar la degradación de un lugar o recurso de importancia medioambiental

El documento recomienda que, dado que Canadá es miembro de GBIF, las organizaciones canadienses incorporen la *Guía de buenas prácticas para la generalización de datos de presencia de especies sensibles* (Chapman y Grafton 2008) al evaluar sus conjuntos de datos medioambientales. Sin mencionar los niveles de generalización específicos, el documento cita las categorías de generalización de esta Guía y de la anterior.

Otras agencias de agregación, como **iDigBio**, han dejado que los que suministran los datos se ocupen de la sensibilidad y no han desarrollado una política en sí misma.

"iDigBio acepta todos los datos que recibe a través de los servicios tal y como son. No hace ningún esfuerzo por ocultar los datos sensibles. El publicador de datos es el único responsable de enmascarar u ocultar la información, incluidos los datos sensibles, al público"

— '<https://www.idigbio.org/content/idigbio-terms-use-policy>[Política de uso iDigBio]'

En muchos casos, la decisión de hacer públicos los datos se toma en función de cada proyecto. Por ejemplo, **Fong y Qiao (2010)** describen un proyecto para cartografiar la ubicación de una especie de tortuga en peligro de extinción en China y sostienen que, aunque estos datos de localización son valiosos para los investigadores, no deberían hacerse públicos por motivos de seguridad de los animales.

Epílogo

Listado de taxa sensibles

Los datos ya están distribuidos por todo el mundo a través de duplicados de especímenes, etc. y aunque los datos pueden estar restringidos en algunas instituciones, otras que tienen duplicados pueden estar divulgando la misma información. Esto puede deberse al desconocimiento de lo que puede considerarse sensible en las zonas de origen del taxón en cuestión, ya que actualmente no se dispone de una lista universal de lo que se considera "sensible". Las dificultades se ven agravadas por el hecho de que un taxón puede ser sensible en una zona, pero no en otra (e incluso puede ser una especie de maleza o plaga invasora en la segunda ubicación). Si hay datos idénticos disponibles públicamente a través de otras fuentes, no pueden considerarse sensibles ([Australian Government 2016](#)).

Por estas razones, se ha recomendado la creación de una listado de taxa potencialmente sensibles desde el punto de vista medioambiental y su vinculación a través de la [Backbone Taxonomy](#) de GBIF. Esto tendría la ventaja de alertar a los proveedores de datos de otras jurisdicciones de que una especie es potencialmente sensible y a través de Backbone Taxonomy proporcionaría enlaces a sinónimos. Es importante señalar que la lista debe considerarse como un elemento desencadenante para señalar la necesidad de tomar una decisión sobre la sensibilidad real de compartir información utilizando los criterios del capítulo anterior y no para generar restricciones generales. No todas las especies amenazadas lo están por el conocimiento de su ubicación o están amenazadas en la totalidad de su área de distribución, por lo que no deberían considerarse sensibles per se y, por tanto, la lista de posibles taxa sensibles desde el punto de vista medioambiental debería ser mucho menor que cualquier lista existente de especies raras y amenazadas.

Esta lista debe crearse usando los [Criterios 1 y 2](#) y los escenarios del [Anexo 1](#) e incluir información adicional como:

- Nombre del taxón
- Criterios y justificación para la inclusión
- Nombre de la persona u organización responsable de que se incluya el taxón
- Cobertura geográfica de la sensibilidad (especialmente si sólo es sensible en una parte de su área de distribución o en una jurisdicción)
- Categoría de sensibilidad recomendada
- Fecha para la revisión

Las jurisdicciones también pueden desear mantener un listado similar para sus propios fines y se recomienda que, si lo hacen, incluyan la información anterior en todos los casos. La ventaja de hacer que la información esté disponible de forma más amplia es que alertará a otros custodios de datos de que su jurisdicción considera el taxón como potencialmente sensible y alertará a los usuarios de que deben tener en cuenta la sensibilidad al publicar los resultados de sus análisis, etc.



Cualquier lista de posibles taxa sensibles desde el punto de vista medioambiental debe considerarse únicamente como un factor desencadenante y cualquier restricción a la disponibilidad de datos reales debe hacerse caso por caso teniendo en cuenta los criterios enumerados.

Recomendaciones de metadatos

Se han hecho una serie de recomendaciones a Biodiversity Information Standards (TDWG) para la inclusión de campos adicionales a la **Elemento de ubicación Darwin Core** (TDWG 2018). Las recomendaciones incluyen:

Tabla 8. Recomendaciones sobre la ampliación al Darwin Core para taxa sensibles

Campo	Comentarios
dataSensitiveIndicator	indicador Y/N de que la observación es sensible.
dataSensitiveReason	La razón principal por la que los datos son sensibles. El formato sugerido es una lista de selección con valores derivados de los criterios 1-4 anteriores (o un campo de texto que combine las declaraciones 1a-4g adjuntas a esos criterios).
dataSensitiveComments	Información adicional de texto libre sobre la(s) razón(es) o la justificación para determinar la relevancia de los criterios para este registro, tal y como se recomienda más arriba.
sensitiveDateForReview	Un campo de fecha que documente cuándo debe revisarse el carácter sensible de la fecha. Especialmente importante cuando la sensibilidad está a la espera de la publicación de los resultados, etc.
precisionDataProvided	<p>La escala o la precisión de los datos puestos a disposición a través del registro Darwin Core. Puede hacerse como precisión de coordenadas, por ejemplo.</p> <ul style="list-style-type: none"> • 0 = 1 grado • 1 = 0.1 grados • 2 = 0.01 grados • 3 = 0.001 grados • 4 = 0.001 grados

precisionDataStored

La escala o la precisión de los datos almacenados o conservados por el custodio de los datos, puede hacerse como precisión de coordenadas, por ejemplo.

- 0 = 1 grado
- 1 = 0.1 grados
- 2 = 0.01 grados
- 3 = 0.001 grados
- 4 = 0.0001 grados
- etc.

o tal vez un texto más libre, como "1 minuto", "0,1 minutos", "1 segundo", dependiendo de cómo se almacenen los datos.

Glosario

autenticación

se refiere a la determinación de la identidad de un usuario, así como a la determinación del contenido al que un usuario está autorizado a acceder. La forma más común de autenticación es el nombre de usuario y la contraseña, aunque esto también proporciona el nivel más bajo de seguridad.

autorización

se refiere al proceso de determinar a qué personas se les pueden conceder diferentes derechos de acceso para la autenticación y el acceso a los datos.

grados decimales

grados expresados como un único número real (p. ej., -22.343456). Tenga en cuenta que las latitudes al sur del ecuador son negativas, al igual que las longitudes al oeste del meridiano cero hasta -180 grados.

generalización

se refiere aquí a cualquier modificación realizada a los datos fuente para ocultar el contenido sensible, normalmente reduciendo la precisión de los datos (como reportar a nivel de cuenca hidrográfica, cuadrícula o condado, citando sólo el lugar nombrado más cercano o eliminando algunas partes de los datos). En términos geográficos, se refiere a la conversión de una representación geográfica a una con menos resolución y menos contenido informativo, tradicionalmente se asocia a un cambio de escala. También se denomina como: *fuzzying*, *dumming-up*, etc.

coordenadas geográficas

medición de la ubicación sobre la superficie de la tierra expresada a partir de latitud y longitud.

actividad humana perjudicial

actividades o procesos humanos que han tenido, tienen o pueden tener un impacto adverso en el estado del taxón evaluado. Algunos ejemplos son la pesca o la tala insostenibles, la caza, la cosecha, la agricultura y la construcción de viviendas, entre otros (ver [UICN 2020](#)).

precisión

describe la mejor unidad de medida utilizada para expresar ese valor (p. ej. si un registro se reporta al segundo más cercano, la precisión es $1/3600$ de un grado; si se reporta un grado decimal a dos decimales, la precisión es 0.01 de grado).

aleatorización

se refiere a una disposición aleatoria deliberada de las observaciones para ocultar su verdadera ubicación. La aleatorización conduce a una falsificación de los datos. También se denomina como *falsificación*.

metadatos a nivel de registros

se refiere a la documentación a nivel de un registro y no de un conjunto de datos completo. En este documento se refiere, en gran medida, a la documentación del estado de sensibilidad del

registro (o de la especie de la que forma parte) junto con las restricciones de acceso correspondientes al registro y los detalles de cualquier generalización de los datos.

datos sensibles

cualquier dato que, debido a su naturaleza, un proveedor de datos no quiera poner a disposición en su estado puro, por ejemplo, las localidades precisas de taxa amenazados.

ajuste espacial

una medida de qué tan bien una representación geométrica coincide con otra representación geométrica representada como una relación entre el área de la mayor de las dos y el área de la menor (ver [Figura 2](#)).

Agradecimientos

En el documento anterior ([Chapman y Grafton 2008](#)) se reconocía a las numerosas personas que respondieron a la encuesta inicial y que asistieron a los talleres y que fueron importantes para la elaboración de ese documento.

Adicionalmente, me gustaría agradecer a John Wieczorek y Paula Zermoglio, por su continuo apoyo y asesoramiento, así como al personal de la Secretaría de GBIF, especialmente a Laura Russell y Kyle Copas que supervisaron el proyecto.

Referencias

- ALA (2018a) What is Sensitive Data? Atlas of Living Australia. <https://support.ala.org.au/support/solutions/articles/6000195500-what-is-sensitive-data->
- ALA (2018b) What is data licensing? Atlas of Living Australia. <https://support.ala.org.au/support/solutions/articles/6000195495-what-is-data-licensing->
- AMEC Earth & Environmental (2010) Best practices for sharing sensitive environmental geospatial data. Version 1.0 http://publications.gc.ca/collections/collection_2011/rncan-nrcan/M104-4-2010-eng.pdf
- Australian Bureau of Statistics (2006) Statistical Local Area (SLA) in 2901.0 - Census Dictionary, 2006 (Reissue). <https://www.abs.gov.au/AUSSTATS/abs@.nsf/bb8db737e2af84b8ca2571780015701e/23d04985e1786824ca25720b0002bb18!OpenDocument>
- Australian Government. Department of the Environment (2016) Sensitive Ecological Data - Access and Management Policy V1.0. <https://www.environment.gov.au/system/files/resources/246e674a-feb1-4399-a678-be9f4b6a6800/files/sensitive-ecological-data-access-mgt-policy.pdf>
- Birddata (n.d.) Sensitive Species. Birdlife Australia. <https://birddata.birdlife.org.au/sensitive-species>
- Chapman AD (2006) Questionnaire on Dealing with Sensitive Primary Species Occurrence Data: Summary of responses. Copenhagen: GBIF Secretariat. <https://doi.org/10.35035/vs84-0p13>
- Chapman AD (2007) Dealing with Sensitive Primary Species Occurrence Data. Report. Report to the Global Biodiversity Information Facility. Copenhagen: GBIF Secretariat. <https://doi.org/10.35035/rajc-t668>
- Chapman AD & Grafton O (2008) Guide to Best Practices for Generalising Sensitive Species Occurrence Data, version 1.0. Copenhagen: GBIF Secretariat. <https://doi.org/10.15468/doc-b02j-gt10>
- Chapman AD & Wieczorek J, eds. (2006) Guide to Best Practices for Georeferencing. Copenhagen: GBIF Secretariat. <https://doi.org/10.15468/doc-2zpf-zf42>
- Chapman AD & Wieczorek J (2020) Georeferencing Best Practices. Copenhagen: GBIF Secretariat. <https://doi.org/10.15468/doc-gg7h-s853>
- Corti L, Day A, & Backhouse G (2000). Confidentiality and informed consent: Issues for consideration in the preservation of and provision of access to qualitative data archives. In Forum Qualitative Sozialforschung/Forum: Qualitative Social Research (Vol. 1). <http://www.qualitative-research.net/index.php/fqs/article/viewArticle/1024> [Accessed 15 May 2019].
- Countryside Agencies' OIN (2007) The 'Environmental Exception' and access to information on sensitive features. Version 1.3.3, Countryside Agencies' Open Information Network Environmental Information Regulations Guidance Note No. 1.
- Dalvi VB & Keole RR (2015) Bottom-Up Generalization: A Data Mining Solution to Privacy Protection. International Journal of Science and Research 4(1): 2068–2071. <https://www.ijsr.net/archive/v4i1/SUB15769.pdf>
- Department of Environment and Conservation – NSW (2007) Threatened Species Information Disclosure Policy (Version 3 Amended March 2007). *No longer available: replaced by OEH 2019.*
- DECCW (2009) Sensitive Species Data Policy. New South Wales Department of Environment,

Climate Change and Water.

<https://www.environment.nsw.gov.au/resources/nature/SensitiveSpeciesPolicyDEC09.pdf>

- Fong JJ & Qiao G (2010) New localities of endangered Chinese turtles from museum specimens and the practical and ethical challenges using and reporting natural history collection data. *Zootaxa* 2393: 59–68. <https://doi.org/10.11646/zootaxa.2393.1.5>
- Frank D, Kriesberg A, Yakel E & Faniel IM (2015) Looting Hoards of Gold and Poaching Spotted Owls: Data Confidentiality Among Archaeologists and Zoologists. <https://deepblue.lib.umich.edu/handle/2027.42/115883>
- GBIF Secretariat (2017) Data publisher agreement. <https://www.gbif.org/en/terms/data-publisher>
- GBIF Secretariat (2019). GBIF Backbone Taxonomy. Checklist dataset <https://doi.org/10.15468/39omei>
- Guralnick R, Wieczorek J, Beaman R, Hijmans RJ & the BioGeomancer Working Group (2006) BioGeomancer: Automated Georeferencing to Map the World's Biodiversity Data. *PLoS Biol* 4(11): e381. <https://doi.org/10.1371/journal.pbio.0040381>
- Guterman L (2006) Endangered by Research: Poachers mine the scientific literature for the locations of newly discovered animals. *The Chronicle of Higher Education* 52(46): A12. <https://www.chronicle.com/article/Endangered-by-Research/26117>
- Hartter J, Ryan SJ, MacKenzie CA, Parker JN & Strasser CA (2013) Spatially Explicit Data: Stewardship and Ethical Challenges in Science. *PLoS Biol.* 11(9), e1001634. <http://doi.org/10.1371/journal.pbio.1001634>
- IUCN (2020) Threats Classification Scheme, version 3.2. <https://www.iucnredlist.org/resources/threat-classification-scheme>
- Llinás JV (2005) Data and Information on Biodiversity and its Protection in the Digital Realm Ver. 1. Bogotá, Colombia: Biological Resources Research Institute Alexandre von Humboldt.
- NBN (2019a) The NBN Data Exchange Principles and their rationale. National Biodiversity Network. <https://nbn.org.uk/the-national-biodiversity-network/archive-information/data-exchange-principles/>
- NBN (2019b) Sensitive Data. National Biodiversity Network. <https://nbn.org.uk/sensitive-data/>
- Nelson G, Sweeney P & Gilbert S (2018) Use of globally unique identifiers (GUIDs) to link herbarium specimen records to physical specimens. *Applications in Plant Sciences* 6(2): e1027. <https://doi.org/10.1002/aps3.1027>
- Noble P, van Riper D, Ruggles S, Schroeder J & Hindman M (2011) Harmonizing Disparate Data across Time and Place: The Integrated Spatio-Temporal Aggregate Data Series. *Historical Methods* 44(2): 79-85. <https://doi.org/10.1080/01615440.2011.563228>
- OEH (2011) Appendix 2. Denaturing Specifications for Sensitive Species records. Sydney, Australia: NSW Government. Office of Environment and Heritage. <https://www.environment.nsw.gov.au/-/media/OEH/Corporate-Site/Documents/Animals-and-plants/Wildlife-management/appendix-2-denaturing-specifications-sensitive-species-records.pdf>
- OEH (2019a) Sensitive species data policy. Sydney, Australia: NSW Government. Office of Environment and Heritage. <https://www.environment.nsw.gov.au/topics/animals-and-plants/>

[wildlife-management/wildlife-policies-and-guidelines/sensitive-species-data](#) [Accessed 26 May 2019].

- OEH (2019b). Scientific Licenses. Sydney, Australia: NSW Government. Office of Environment and Heritage. <https://www.environment.nsw.gov.au/licences-and-permits/scientific-licences>
- Page RDM (2009) bioGUID: Resolving, discovering, and minting identifiers for biodiversity informatics. *BMC Bioinformatics* 10(Suppl 14): S5. <https://doi.org/10.1186/1471-2105-10-S14-S5>
- Parry O & Mauthner NS (2004) Whose Data are They Anyway?: Practical, Legal and Ethical Issues in Archiving Qualitative Research Data. *Sociology* 38(1): 139–152. <https://doi.org/10.1177/0038038504039366>
- Richards K (2010) TDWG GUID applicability statement, version 2010-09. Biodiversity Information Standards (TDWG). <http://www.tdwg.org/standards/150>
- Richards K, White R, Nicolson N & Pyle R (2011) Beginners' guide to persistent identifiers, version 1.0. Copenhagen: Global Biodiversity Information Facility. <https://www.gbif.org/document/80575>
- SANBI (2010) Biodiversity Information Policy Framework. Policy Series. Digital Access to Sensitive Taxa Data. Silverton, South Africa: South African National Biodiversity Institute. <http://biodiversityadvisor.sanbi.org/wp-content/uploads/2012/09/SANBI-Biodiversity-Information-Policy-Series-Digital-Access-to-Sensitive-Taxon.pdf>
- SANBI (2016) Report of the National Sensitive Species List Workshop 18 and 19 August 2016. Silverton, South Africa: South African National Biodiversity Institute. <http://biodiversityadvisor.sanbi.org/wp-content/uploads/2017/06/20160819-NSSL-Workshop-Report.pdf>
- Stuart BL, Rhodin GJ, Grismer LL & Hansel T (2006) Scientific Description Can Imperil Species. *Science* 312(5777): 1137. <https://doi.org/10.1126/science.312.5777.1137b>
- Tann J & Flemons P (2009) Our secrets are not our secrets. Atlas of Living Australia Sensitive Data Report. Version 1.1. <https://www.ala.org.au/wp-content/uploads/2010/07/ALA-sensitive-data-report-and-proposed-policy-v1.1.pdf> [Accessed 25 May 2019]
- TDWG (2018) Darwin Core quick reference guide. Biodiversity Information Standards (TDWG). <https://dwc.tdwg.org/terms/>
- US Forest Service (2005) Forest Service Sensitive Species Summary, Designated Sensitive Species (that are not listed or proposed species under the ESA), as of 31 October, 2005. https://www.fs.fed.us/biology/resources/pubs/tes/ss_sum_by_region_31Oct2005_fs.pdf
- US Forest Service (2016) Forest Service Sensitive Species – Wildlife. https://www.fs.usda.gov/Internet/FSE_DOCUMENTS/fseprd530660.pdf
- Wang K, Yu PS & Chakraborty S (2004) Bottom-Up Generalization: A Data Mining Solution to Privacy Protection, in Proceedings of Fourth International IEEE Conference on Data Mining (ICDM'04): 249-256.
- Wang Z, Dong H, Kelly M, Macklin JA, Morris PJ, Morris R 2009. Filtered-Push: A Map-Reduce Platform for Collaborative Taxonomic Data Management. World Congress on Computer Science and Information Engineering, March 31 – April 2, 2009, Los Angeles, California, USA. <https://doi.org/10.1109/CSIE.2009.948>
- Wieczorek J, Guo Q & Hijmans R (2004) The point-radius method for georeferencing locality descriptions and calculating associated uncertainty. *International Journal of Geographical*

Information Science 18: 745-767.

- Wieczorek J, Bloom D, Guralnick R, Blum S, Döring M, Giovanni R, Robertson T & Vieglais D (2012) Darwin Core: An Evolving Community-Developed Biodiversity Data Standard. PLoS ONE 7(1): e29715. <https://doi.org/10.1371/journal.pone.0029715>
- Wylie A (1996) Ethical dilemmas in archaeological practice: Looting, repatriation, stewardship, and the (trans) formation of disciplinary identity. *Perspectives on Science* 4(2): 154–194.
- Zermoglio PFF, Chapman AD, Wieczorek JR, Luna MC, Bloom DA (2020) Georeferencing Quick References Guide Copenhagen: Copenhagen: GBIF Secretariat. <https://doi.org/10.35035/e09p-h128>.

Anexo 1: Escenarios usando los criterios 1 y 2 como desencadenantes

Los siguientes conjuntos de escenarios muestran cómo las declaraciones de criterios que se dan en el capítulo sobre **Determinación de la sensibilidad** pueden utilizarse para desarrollar declaraciones resumidas para documentar las razones por las que un taxón puede considerarse sensible. Las declaraciones resumidas también deberían incluir una justificación de apoyo, como tipos específicos de daños, etc. Por ejemplo, en el escenario B, la declaración completa puede ser algo así:

Los taxa podrían estar en riesgo de sufrir daños por enfermedades transportadas en las ruedas de la maquinaria forestal, pero su presencia no se ve afectada por la disponibilidad de datos.

Esto puede aplicarse a una especie de planta en una zona forestal susceptible de ser atacada por *Phytophthora*, ya que los hongos se transfieren en las ruedas de los vehículos forestales.

Criterio 1

Escenario A

1a: No hay riesgo significativo de una actividad humana perjudicial.

El taxón no es sensible.

Escenario B

1a: El taxón está en riesgo de una actividad humana perniciosa.

1d: Actualmente no hay evidencia establecida de daños reales para el taxón.

1f: La disponibilidad de datos sobre biodiversidad no aumentará la probabilidad de que se produzca una actividad humana perjudicial.

El taxón podría estar en riesgo de daño, pero la probabilidad de daño no se ve afectada por la disponibilidad de datos.

Escenario C

1a: El taxón está en riesgo por la actividad humana perjudicial.

1d: Actualmente no hay evidencia establecida de daño real para el taxón.

1e: La disponibilidad de datos sobre la biodiversidad aumentará la probabilidad de que la actividad humana perjudicial tenga lugar.

El taxón podría estar en riesgo de sufrir daños y la probabilidad de que se produzcan se ve afectada por la disponibilidad de datos.

Escenario D

1a: El taxón está en riesgo por la actividad humana perjudicial.

Escenario D

1c: Hay evidencia establecida de daño real o reciente para el taxón.

1f: La disponibilidad de datos sobre la biodiversidad no aumentará la probabilidad de que se produzca la actividad humana perjudicial.

El taxón está en riesgo de sufrir daños y hay pruebas que lo apoyan, pero la disponibilidad de datos no afecta a su presencia.

Escenario E

1a: El taxón está en riesgo por la actividad humana perjudicial.

1c: Hay evidencia establecida de daño real al taxón.

1e: La disponibilidad de datos sobre la biodiversidad aumentará la probabilidad de que la actividad humana perjudicial tenga lugar.

El taxón está en riesgo de sufrir daños, hay pruebas que lo apoyan y su presencia se ve afectada por la disponibilidad de datos.

Criterio 2

Escenario F

2b: El taxón no es significativamente vulnerable a la actividad humana perjudicial.

2d: El taxón no es vulnerable a la actividad humana perjudicial en toda su área de distribución y/o hay zonas en las que el taxón no corre un riesgo significativo.

El taxón no es significativamente vulnerable a la actividad dañina y no es vulnerable a esa actividad en toda su área de distribución y hay zonas en las que el taxón no corre un riesgo significativo de esa actividad.

Escenario G

2a: El taxón tiene características que lo hacen significativamente vulnerable a la actividad humana perjudicial.

2d: El taxón no es vulnerable a la actividad humana perjudicial en toda su área de distribución y/o hay zonas en las que el taxón no corre un riesgo significativo.

El taxón es significativamente vulnerable a la actividad dañina, pero no es vulnerable a esa actividad en toda su área de distribución y hay zonas en las que el taxón no corre un riesgo significativo por esa actividad.

Escenario H

2a: El taxón tiene características que lo hacen significativamente vulnerable a la actividad humana perjudicial.

2c: El taxón es vulnerable a la actividad humana perjudicial en toda su área de distribución.

El taxón es significativamente vulnerable a la actividad dañina y es vulnerable a esa actividad en toda su área de distribución.