

Encuesta de GBIF sobre Datos Sensibles



Isabel Ortega Maqueda
ortega@gbif.es

I TALLER SOBRE GESTIÓN DE DATOS SENSIBLES
Madrid, 15-16 de Octubre 2008

Lo que vamos a ver...

- Resultados de la encuesta de GBIF sobre datos sensibles
- Resumen de la encuesta
- Razones para publicar información
- Razones para restringir información
- Qué se hace actualmente sobre generalización
- Situación actual sobre protección de datos en la red de GBIF

Resultados de la encuesta de GBIF sobre datos sensibles

- Las encuesta on-line llevada por GBIF en marzo de 2007 tuvo unos **resultados** que fueron publicados en:

Questionnaire on Dealing with Sensitive Data

- En la encuesta intervinieron más de **100 instituciones** a nivel **mundial**.
- Participación equitativa de colecciones **botánicas** y **zoológicas**, con algunas representaciones de colecciones **vivas** y **paleontológicas**.
- Las instituciones participantes están repartidas en mas de **28 países**
- La mayoría de países pertenecen la hemisferio norte.
- Poca participación de representantes de África, Asia y América del sur.

Resumen de encuesta on-line

Se evaluaron 37 preguntas sobre datos sensibles en la Web:

- ¿Cuando publica sus datos on-line, generaliza usted algún dato?
- ¿Está usted requerido por la legislación a limitar el acceso a ciertos datos?
- ¿Cuales son sus principales razones para restringir el acceso a datos sensibles?
- ¿Cuales son sus principales razones para permitir el acceso a datos sensibles?
- ¿Qué proporción de su colección considera que son “registros sensibles”?
- ¿Trata usted todos los datos sensibles como si fueran de la misma categoría, o tiene usted diferentes niveles de generalización?

Tipos de datos sensibles identificados

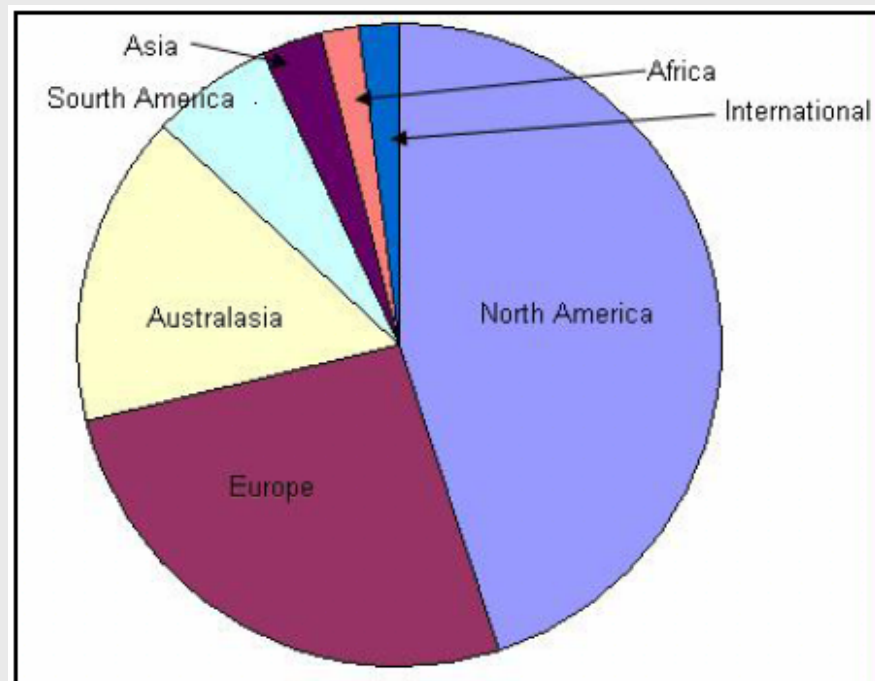
Tipos de datos sensibles identificados por los encuestados:



Países de procedencia de los encuestados

Instituciones participantes de 28 países:

USA	42	Faroe Islands	1
Australia	14	France	1
Spain	8	Iceland	1
Canada	4	India	1
UK	4	Japan	1
Argentina	3	Peru	1
Denmark	3	Philippines	1
Belgium	2	Poland	1
Colombia	2	Sweden	1
Germany	2	Switzerland	1
New Zealand	2	The Netherlands	1
South Africa	2	International	2
Austria	1		



Resumen de la encuesta según las colecciones

De las 100 instituciones participantes en la encuesta, se detallan las colecciones que albergan:

	Specimens	Living Collections	Observations	Survey and Monitoring
Botanical, Zoological and Other	13	1	14	15
Botanical, Zoological	14	3	20	16
Zoological and Other	1	0	2	1
Botanical and Other	0	1	0	0
Botanical Only	26	14	11	11
Zoological Only	20	3	19	18
Other Only (incl. Paleontology)	2	2	1	1
Total	76	24	67	62

Razones de los encuestados para publicar información

- Para **investigación y análisis científico**, avances científicos y colaboración en proyectos (33 encuestados).
- Para **conservación** de especies; **planificación** y **gestión** sobre conservación (21).
- **Asesoramiento** medioambiental, **gestión de recursos biológicos** y espacios naturales; para estudios de impacto medioambiental (12).
- **Requerimiento** de información de **agencias gubernamentales** y organizaciones profesionales para la gestión de medioambiente (8).
- Estudios de **distribución de especies**, **modelado** de especies, estudios sobre vegetación y mapeo; **análisis** a escala mundial (6).
- La información debería estar **siempre disponible** para su uso público (3).

Razones para publicar información

*La **información sobre biodiversidad** debe estar **libremente** disponible para ser **compartida** mundialmente, para su uso para la toma de decisiones en materia medioambiental, educación, investigación y otros beneficios públicos.*

*Mostrar el **detalle completo** de la información sobre biodiversidad deberá **reducir el riesgo de daño** al medioambiente. Cuando la publicación de esta información tiene el efecto **contrario**, el acceso al detalle **debe ser controlado**.*

Razones de los encuestados para restringir información

- Proteger especies amenazadas, económicamente importantes, y reducir el impacto sobre las comunidades sensibles (37).
- Evitar el sabotaje y la recolección comercial y sin escrúpulos, la caza furtiva, la perturbación a ciertas especies y la sobreexplotación (35).
- Proteger los datos de terceros, datos albergados por la institución y cedidos bajo confidencialidad y con acuerdos sobre los datos; proteger las fuentes de los datos y los derechos de los proveedores de datos, proteger los derechos de propiedad intelectual, incluyendo la necesidad de una adecuada atribución y citación (16).
- No revelar resultados científicos pendientes de publicación y mantener la ventaja de la competitividad (14).
- Proteger nombres de personas y su privacidad (8).
- Evitar el uso inapropiado de los datos y el temor a una mala interpretación (5).

Situación actual sobre protección de datos en la red de GBIF

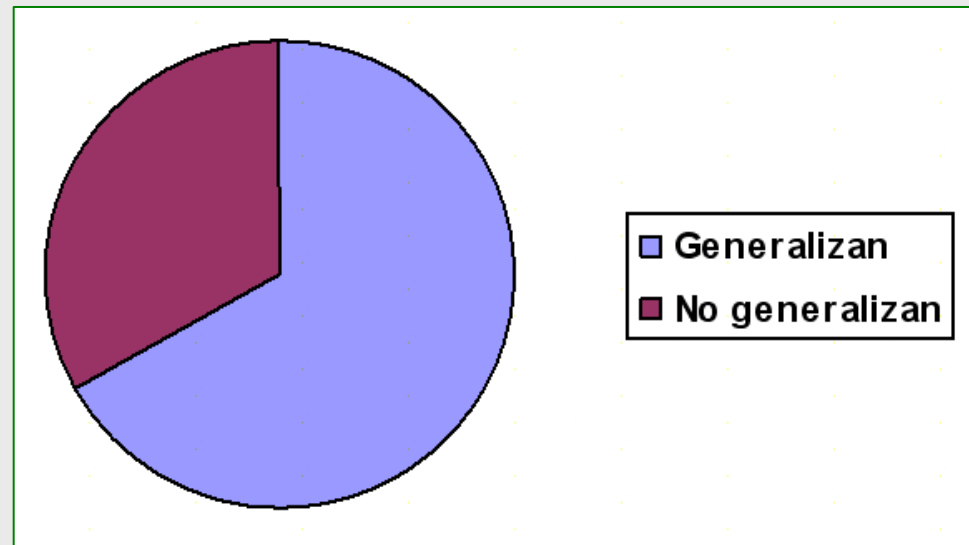
- Actualmente, algunas instituciones no están haciendo accesibles on-line en la red de GBIF ningún dato referente a taxones sensibles.
- En los datos que sí están en la Web, algunos campos están siendo borrados (incluyendo la localidad, nombres de personas y fechas), los datos sobre georreferenciación están siendo generalizados de muy diferentes maneras, y muy pocos están suministrando documentación sobre lo que están haciendo.
- La falta de documentación es seguramente lo más inquietante ya que significa que los datos pueden no ser adecuados para el uso que el usuario necesita hacer con él, y la información no está disponible para informar de este hecho.

Sobre la generalización de los datos

Generalizan los encuestados?

- 2/3 respondieron que **SI**
- 1/3 respondieron que **NO**.

De los que contestaron **SI**, muchos restringen la información de distinta manera según los diferentes taxones, y esto suele depender de si son datos relativos a especies amenazadas o si son datos pendientes de la publicación de un trabajo de investigación.



Cómo se está realizando la generalización

De los 2/3 de los encuestados que generalizan información:

- 64% borra o altera la localidad y/o la información de la georreferenciación
- 24% restringe la información sobre los nombres del colector u observador, nombres del determinador, fechas, información taxonómica, información sobre el hábitat, sexo, usos tradicionales y otros.
- 7% restringe todo excepto el nombre y el número de catálogo.
- 5% no muestra ninguna información sobre taxones sensibles.

Otras restricciones de datos

- Las **razones** para restringir la información del colector y el determinador incluyen:
 - Proteger la **privacidad** de personas vivas
 - Restringir la posibilidad de **trazar itinerarios** y así acotar la información de especies sensibles
 - Respetar la **legislación** sobre privacidad;
 - Proteger a las personas frente a **posibles represalias** de activistas pro-derechos de los animales;
 - No **malinterpretar los datos observacionales** pensando que puedan incluir (posiblemente ilegal) material de ejemplares y no solamente fotografías o datos observacionales.
- La mitad de los encuestados usa **acuerdos individuales de compartición de datos** o licencias de uso de los datos para dar los datos sensibles a usuarios fiables. La mayoría estudian **caso por caso**, aunque utilizan acuerdos generales firmados para diferentes programas.

Expectativas de los proveedores sobre la protección de datos sensibles

- Algunos proveedores informaron que no pondrían datos on-line de taxones sensibles bajo ninguna circunstancia.
- La mayoría, sin embargo, buscan una manera aceptable para la generalización de la información de manera que las localidades exactas de datos sensibles estén ocultas para el público general, mientras que están dispuestos a dar esta información a investigadores “de buena fe” y usuarios tales como agencias gubernamentales.
- Otros prefieren que los usuarios contacten con los conservadores de la colección y justifique las razones por las que necesitan el uso de dichos datos.
- Muchos sugirieron que harían sus datos disponibles on-line si hubiera un método adecuado de registro y/o identificación de usuarios “fiables”, y un método seguro para que sólo estos usuarios tuvieran acceso a dichos datos.