



Datos abiertos: definiciones, tipologías y principios



Día Virtual de los datos abiertos
Corporación Universitaria para el Desarrollo de Internet Universidad Carlos III de Madrid
14 de Marzo 2014

Nancy Diana Gómez

Contenido

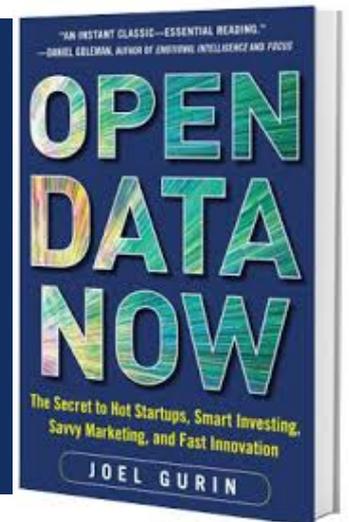
▣ CONTEXTO

▣ DEFINICIONES

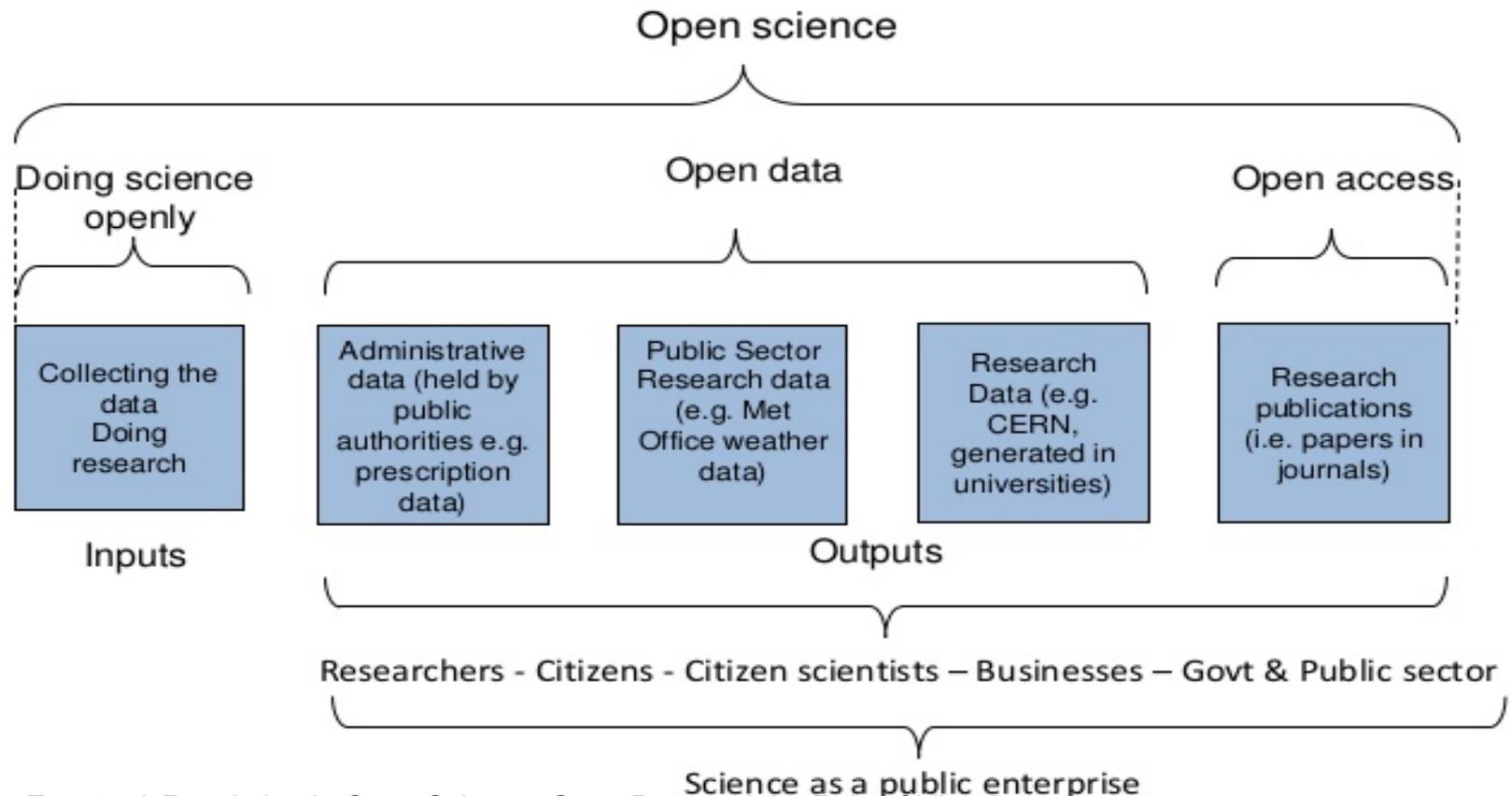
▣ TIPOLOGIAS

▣ PRINCIPIOS

CONTEXTO



Ciencia abierta



Fuente: A Revolution in Open Science: Open Data and the Role of Libraries
(Geoffrey Boulton at LIBER 2013)

¿Por qué Ciencia Abierta?

(Stodden, 2011)

□ Facilita

- Reproductibilidad
- Innovación académica e industrial
- Acceso al conocimiento

Incentivos que influyen en la publicación y difusión de los datos (Stodden, 2011)

- ▣ Fondos de investigación (NSF, NIH, etc.)
- ▣ Editoriales de revistas (PLOSone)
- ▣ Promociones institucionales (premios, promociones, etc.)
- ▣ Integridad científica

Motivos que disuaden a los científicos en la publicación de datos abiertos (Stodden, 2010)

- Tiempo en documentar y limpiar datos para publicar (54%)
- Lidiar con preguntas de los usuarios de los datos (34%)
- No recibir atribución o citación como autores de los datos (42%)
- Barreras legales –copyright (41%)
- Perdida potencial de futuras publicaciones (35%)
- Ventaja que pueden obtener los competidores en el campo (33%)

DEFINICIONES



Desde el dato al conocimiento

Data



Information



Presentation



Knowledge



EpicGraphic.com

<http://bit.ly/OoF9oq>

Datos de investigación

Todo aquel material que ha sido registrado durante la investigación, reconocido por la comunidad científica y que sirve para certificar los resultados de la investigación que se realiza. (NIH)

Datos Abiertos

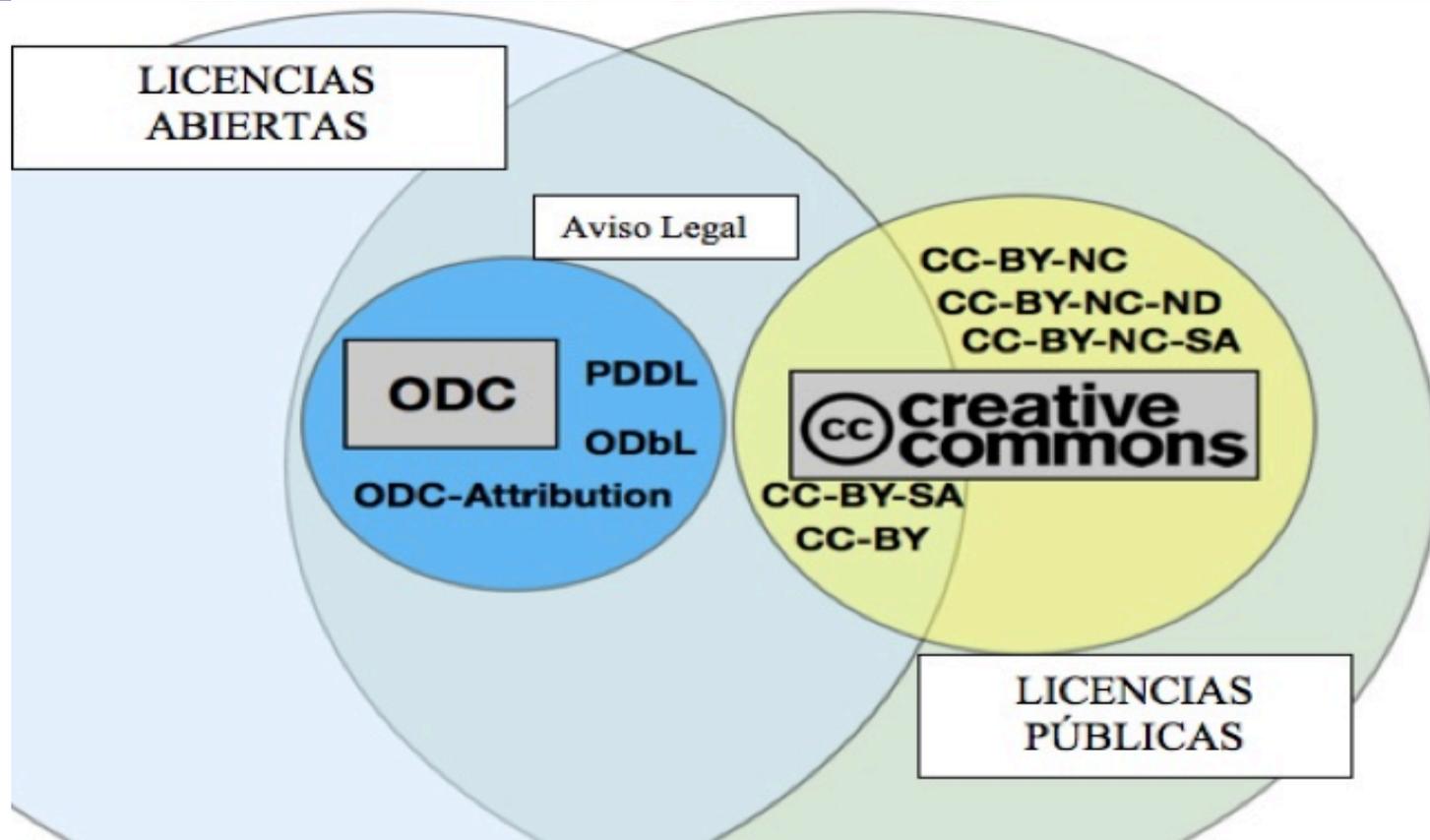
“los datos que pueden ser utilizados libremente, reusados y redistribuidos por cualquier persona, y se encuentran sujetos, cuando más, al requerimiento de atribución o reconocimiento y de ser compartidos de la misma manera en la que aparecen”

Open Knowledge Foundation, 2012

Un dato es abierto si la forma en que se distribuye satisface las siguientes condiciones:

1. Acceso 2. Redistribución 3. Re-uso 4. Ausencia de restricción tecnológica	Principios
5. Reconocimiento 6. Integridad	Restricciones
7. Sin discriminación entre personas o grupos 8. Sin discriminación respecto al proposito	Discriminación
9. Distribución con licencia 10. Licencia no debe ser especifica para un paquete 11. Licencia no debe restringir la distribución de otros trabajos	Reglas de licenciamiento

LICENCIAS



TIPOLOGÍA



Datos de investigación - tipos

- ▣ National Science Foundation (NSF, 2005)

- ▣ Datos observacionales
- ▣ Datos experimentales
- ▣ Datos computacionales
- ▣ Registros de gobierno, negocios y vida pública (Borgman, 2011)



+

- ▣ The Research Information Network (RIN, 2007)

- ▣ Datos derivados
- ▣ Datos canónicos

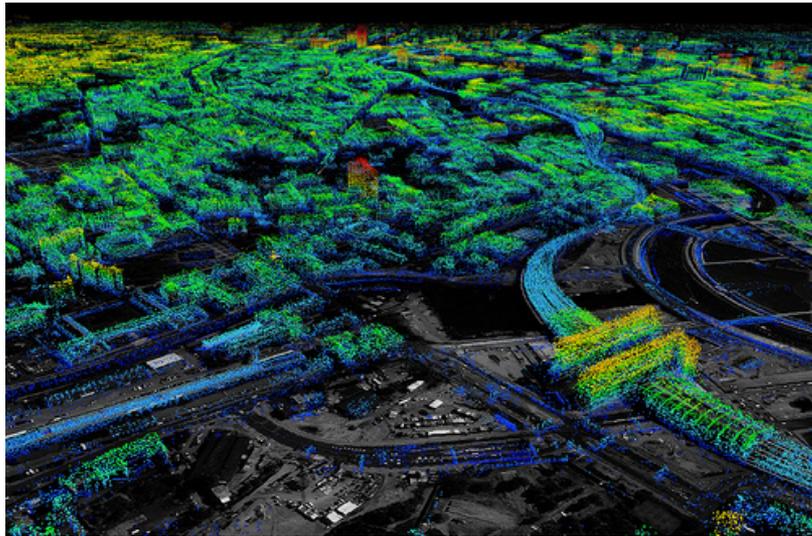


Datos de investigación - tipos (Salinas, 2012)

B. Diferentes taxonomías de los datos de investigación	
B.1 Según el formato	B.2. Proceso de obtención
<ul style="list-style-type: none"> - Textos - Números - Imágenes - etc. 	<p>Experimentales</p> <ul style="list-style-type: none"> - Secuencias genéticas - Cromatografías <p>Simulaciones</p> <ul style="list-style-type: none"> - Modelos climáticos - Modelos económicos <p>Observacionales</p> <ul style="list-style-type: none"> - Encuestas - Experimentos irrepetibles
B.3. Según objetivo recogida	B.4. Según fase de investigación
<p>Específicos</p> <ul style="list-style-type: none"> - Solo de interés para un proyecto de investigación. <p>Alcance medio</p> <ul style="list-style-type: none"> - De interés para una disciplina concreta. <p>De interés general</p> <ul style="list-style-type: none"> - De interés para la ciencia en su conjunto e incluso de interés social. 	<p>Datos preliminares</p> <ul style="list-style-type: none"> - Datos recién extraídos sin ningún tipo de procesamiento. Denominados en inglés <i>raw data</i>. <p>Datos finales</p> <ul style="list-style-type: none"> - Datos que ya han sido procesados y combinados con otros. Denominados en inglés <i>final research data</i>.

EJEMPLOS

Datos observacionales

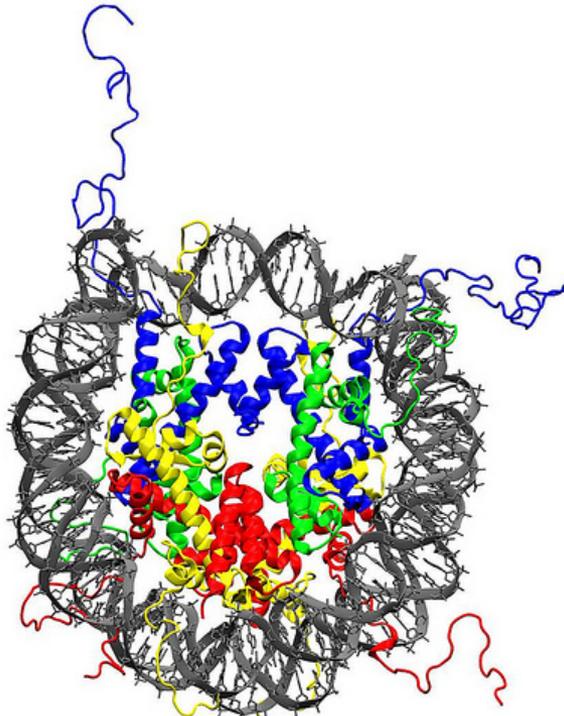


https://www.flickr.com/photos/esq_events/9718274734/



https://farm4.staticflickr.com/3610/3554922716_43be548be4_o.jpg

Datos computacionales

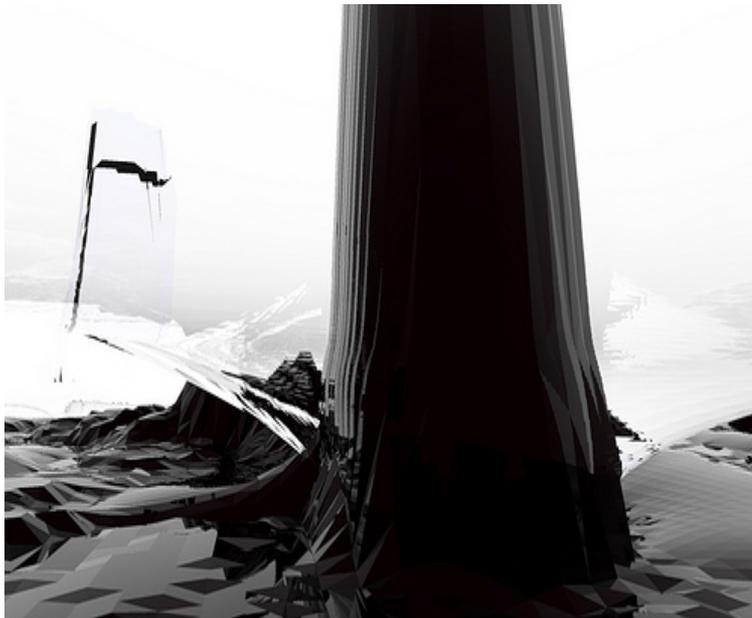


https://farm8.staticflickr.com/7414/9030270106_c9f5d4a317.jpg



https://farm3.staticflickr.com/2711/4266200682_53d50f4214.jpg

Datos experimentales

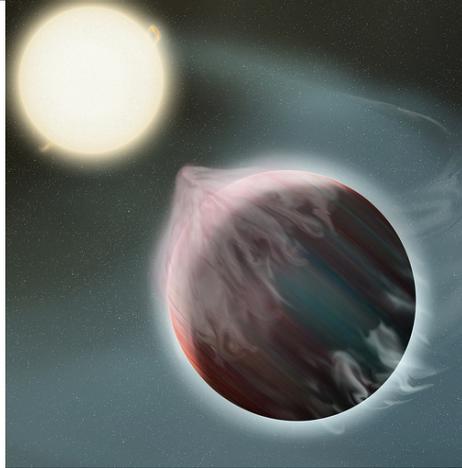


https://farm5.staticflickr.com/4071/4696476526_ed04de2feb.jpg



<http://bit.ly/1IC4OoO>

Y más



PRINCIPIOS



Principios Panton (2010)

Peter Murray-Rust, Cameron Neylon, Rufus Pollock y John Wilbanks.

1. Publicar los datos con una declaración explícita y sólida de los deseos del investigador para con los datos.
2. Utilizar licencias que sea apropiado para los datos, o exenciones reconocidas respecto al derecho de autor .
3. Si desea que sus datos sean utilizados con eficacia y se añadan a los demás, debería definirse de acuerdo a la definición de OKF. Los datos, en particular, no deben utilizar cláusulas restrictivas como las no comerciales.
4. Fuerte recomendación de publicación de datos producidos con fondos públicos deben colocarse explícitamente en el dominio público a través PDDL (<http://opendatacommons.org/licenses/pddl/1-0/>) o CCZero (<http://creativecommons.org/publicdomain/zero/1.0/>). Garantiza el cumplimiento tanto en el (Science Commons, open data, OKF).

<http://pantonprinciples.org/>

Muchas gracias

ndgomez@bib.uc3m.es

nancydiana@gmail.com

[@nancydiana](#)