

Curación digital

el desafío de las colecciones de
investigación



Ms. María Patricia Prada

10 de noviembre de 2014

Curación digital



"implica el mantenimiento, la preservación y la adición de valor a los datos de investigación digital en todo su ciclo de vida. La gestión activa de los datos de investigación reduce amenazas a su valor para la investigación a largo plazo y mitiga el riesgo de obsolescencia digital."

(Digital Curation Centre, 2014)

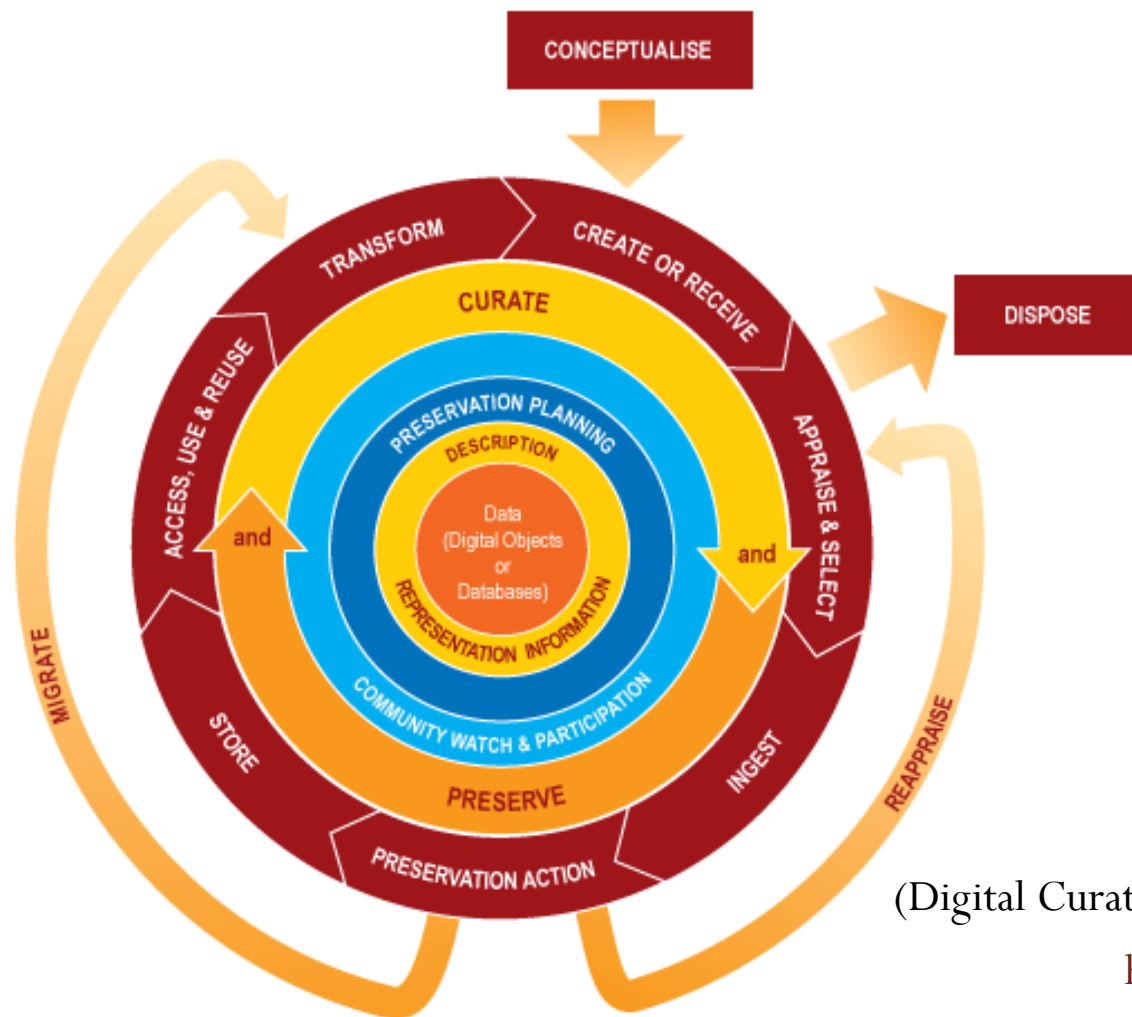
<http://dcc.ac.uk/>

Datos de investigación



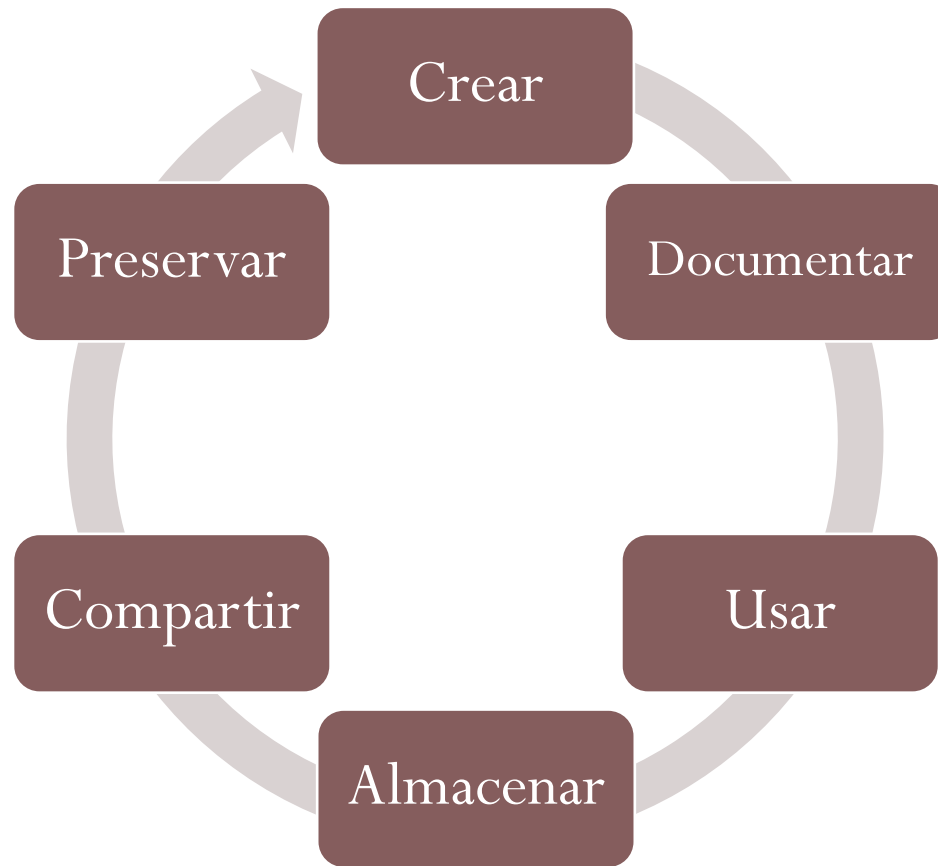
- Experimentales u observacionales.
 - Cualitativos, cuantitativos,
- Numéricos, descriptivos o visuales.
 - En bruto o analizados.

Ciclo de vida de la curación digital



(Digital Curation Centre, 2008)

<http://dcc.ac.uk/>



Gestión de Datos

La realización de investigación sobre conjuntos de datos preexistentes es menos costosa que el establecimiento de nuevos estudios y una nueva vinculación puede ampliar el enfoque.

(Lewsey, 2013)

Planificar

- ¿Qué datos se crearán?
- ¿Cuánto almacenamiento se necesitará?
- ¿Existen cuestiones éticas que requieran consentimiento?
- ¿Cómo ayudar a los investigadores a desarrollar planes apropiados?

Importante

- Tomar decisiones tempranas correctas restringe posibles imprevistos futuros.
- Asegurarse de que los investigadores conocen la ayuda disponible.
- Fomentar el debate para ayudar a encontrar el mejor enfoque.

Documentar

- ¿Qué información necesitan los usuarios para comprender los datos?
 - Proyectos e investigadores
 - Variables/campos y sus valores
 - Etiquetas, esquema de clasificación, lista de abreviaturas
 - Excepciones, peculiaridades, resultados cuestionables
- ¿Cómo va a ser capturado?
- ¿Existen estándares que se puedan utilizar?

Importante

- Crear la metadata en el momento - es difícil hacerlo más tarde
- Desarrollar procesos para que todo el mundo haga lo mismo
- Utilice estándares de interoperabilidad

Seleccionar

- ¿Qué debe ser conservado?
- ¿Qué debe ser destruído?
- ¿Vale la pena mantener los datos?
- ¿Dónde se guardarán los datos?

Establecer criterios de selección de datos

Importante

- Comprobar licencias y acuerdos de consentimiento
- Proporcionar asesoramiento y orientación para ayudar a los investigadores a decidir
- El almacenamiento no es el único costo en la gestión de datos

Plan de gestión de datos



<https://dmponline.dcc.ac.uk/>

ICPSR Framework for creating a DMP

<http://www.icpsr.umich.edu/icpsrweb/content/datamanagement/dmp/framework.html>

Componentes recomendados

- Descripción de los datos
- Formatos utilizados
- Metadatos utilizados
- Planes de almacenamiento y seguridad a corto, mediano y largo plazo
- Cuestiones de confidencialidad y de propiedad intelectual de los datos

Inicio de plan de gestión de datos

- Descripción del proyecto: ¿Cuál es la investigación? ¿Quiénes son los investigadores?
- Descripción de los datos: ¿qué tipos de datos: entrevistas, encuestas, colecciones, materiales curriculares y otros materiales, se producen en el transcurso del proyecto? ¿Cómo, cuándo y en qué formato se recogen los datos? ¿Se están utilizando datos producidos por alguien más? Si es así, ¿de dónde proceden y por qué fueron elegidos para esta investigación? ¿Qué software se necesita para leer o ver los datos?
- Formatos utilizados: ¿Qué formatos de archivo se utilizarán? ¿Está utilizando formatos estándar? ¿Estos formatos responden a un estándar abierto o son propietarios?
- Metadatos utilizados: Definir esquemas de metadatos a aplicarse para hacer que los datos accesibles y comprensibles por otros investigadores. ¿Cómo se gestionarán y almacenarán estos metadatos?
- Planes de almacenamiento y seguridad a corto, mediano y largo plazo: ¿Quién consideran debe ser responsable de la gestión de los datos? ¿Cuál será el almacenamiento local y la copia de seguridad de los documentos? ¿Cuáles datos se conservarán a largo plazo? ¿Cómo van a compartir los datos durante el proceso de investigación entre los colaboradores? ¿Qué recaudos habrá que tomar para organizar, nombrar, etiquetar y agrupar los objetos digitales durante la investigación? ¿Cómo se van a formar los investigadores para que resguardar sus materiales?
- Cuestiones de confidencialidad y de propiedad intelectual de los datos: ¿Qué datos se compartirán, cuándo y cómo se van a acceder? ¿Hay alguna restricción de acceso necesaria? ¿Habrá embargos sobre algunos datos? ¿Sobre cuáles? ¿Habrá datos con acceso restringido sólo a ciertos usuarios autorizados? ¿Cómo van a ser tratados los datos que plantean privacidad, preocupaciones éticas, o de confidencialidad? ¿Se permitirá el intercambio final de datos anónimos bajo un consentimiento informado? ¿Cómo se gestionará el permiso para usar estos datos? ¿Cómo se gestionará el permiso para difundir este material? ¿Quién tendrá los derechos de propiedad intelectual para los datos y otra información creada por el proyecto? ¿Qué tipo de licencia se va a utilizar?



Para pensar

Para pensar

¿Cuánto tiempo deberían ser preservados a largo plazo los datos de investigación de valor?

1. 10 años
2. A perpetuidad
3. 5 años

Para pensar

¿Los investigadores deberían presentar planes de gestión e intercambio de datos como parte de su solicitud de subvención?

1. Siempre
2. Nunca
3. A veces

Para pensar

¿Cuándo deberían los investigadores que reciben subvenciones públicas hacer disponibles sus datos?

1. Lo más pronto posible, habitualmente cuando publiquen los resultados
2. Dentro de los 3 años del fin del proyecto
3. El investigador debe decidir
4. Nunca

Para pensar

¿Las oficinas públicas financiadoras deberían proporcionar fondos adicionales para cubrir los costos asociados con la gestión y el intercambio de datos?

1. Sí
2. No

Una propuesta sustentable

- Tener políticas y planes
- Brindar acceso a datos abiertos cuando sea posible
- Mencionar las fuentes para reconocer la propiedad intelectual
- Aplicar metadatos para compartir, enlazar a datos de publicaciones
- Atender limitaciones legales, éticos y comerciales
- Permitir embargos limitados para proteger el esfuerzo de los creadores
- Preservar los datos de valor a largo plazo
- Asegurar el uso rentable de los fondos públicos

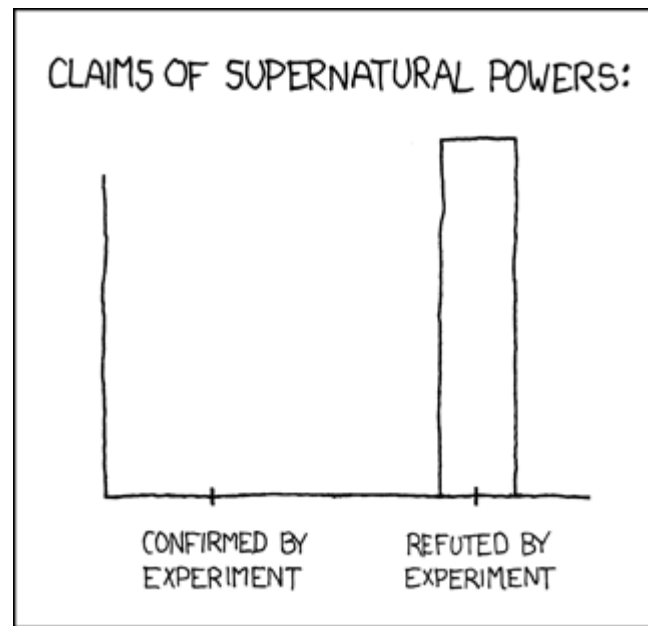
RCUK Common Principles on Data Policy

<http://www.rcuk.ac.uk/research/datapolicy/>

¿Preguntas?

¡Muchas gracias!

pprada@sisbi.uba.ar



Los datos hasta el momento

A WEBCOMIC OF ROMANCE, SARCASM, MATH, AND LANGUAGE