

CDUR: una propuesta de evaluación del software de investigación

CDUR: a procedure for research software evaluation

Teresa Gomez-Diaz

CNRS - Laboratoire d'informatique Gaspard-Monge (LIGM)

Colaboración con Prof. Tomas Recio, Universidad de Cantabria

This work is licensed under the
Creative Commons Attribution 4.0 International License

<http://creativecommons.org/licenses/by/4.0/>

Sala de Grados, Facultad de Ciencias, 3 Septiembre 2020
Universidad de Cantabria, Santander



Programa de ordenador (BOE núm. 97, 22 abril 1996)

A los efectos de la presente Ley se entenderá por programa de ordenador toda secuencia de instrucciones o indicaciones destinadas a ser utilizadas, directa o indirectamente, en un sistema informático para realizar una función o una tarea o para obtener un resultado determinado, cualquiera que fuere su forma de expresión y fijación. [...] comprenderá también su documentación preparatoria.

RS indica:

- [EN] research software
- [ESP] programa (de ordenador) o software de investigación
- [FR] logiciel de la recherche

Ver definición en p. 8/16.

Contexto

Contexto local: Laboratoire d'informatique Gaspard-Monge (LIGM)

- Evaluación HCERES del laboratorio (2013-18), febrero 2019
- Producción importante de RS: **66, 50** libres (free/open source)
- ¿Cómo tener en cuenta esta producción en la evaluación ?

Contexto global:

Libre (Free/OSS), Ciencia Abierta y las evoluciones en la evaluación

- Free Software Foundation (R. Stallman, FSF, 1985) y OSS (OSI, 1988)
- Budapest Open Access Initiative (BOAI, 2002), definición *open access*
- 2 informes EC Expert Groups - evolución Ciencia Abierta (adopción & prácticas)
 - ▶ [9] (2017) Cabello Valdes C, Rentier B, et al.: Evaluation of research careers fully acknowledging Open Science practices. ...
 - ▶ [11] (2019) Guéron JC, Jubb M, et al.: Future of Scholarly Publishing and Scholarly Communication. Report of the Expert Group to the EC.

[9, 11] **Software (use & prod.) is now included in Open Science best practices.**

[11] **The conclusion is actually simple: the evaluation of research is the keystone.**

Los objetivos de esta presentación (1/2)

La producción de RS no se considera adecuadamente en la evaluación.

Es necesario **cambiar y desarrollar** las prácticas de la evaluación, teniendo en cuenta las evoluciones de la Ciencia Abierta.

Nuestro objetivo es presentar y motivar la lectura de:

Gomez-Diaz T. and Recio T.,

On the evaluation of research software: the CDUR procedure

[version 2 ; peer review: 2 approved] 26 Nov 2019 (V1, 5 Aug 2019).

F1000Research 2019, 8:1353, <https://doi.org/10.12688/f1000research.19994.2>

+1 año de escritura, experiencia complementaria, 80 referencias, + 85 notas pie de página.

Elección de F1000Research:

Sección *Science Policy Research y open peer review procedure*,

También en la sección Mathematical, Physical, and Computational Sciences.

Nota1: RS indica *research software* (o programa de investigación).

Nota2: [nb] indica el número de la referencia nb en V2.

Los objetivos de esta presentación (2/2)

El artículo propone el protocolo **CDUR** para tener en cuenta la producción de RS en la evaluación de la investigación.

El protocolo **CDUR** es flexible para poder adaptarlo a los diferentes contextos de evaluación, a lo largo de la carrera en investigación.

CDUR tiene cuatro etapas:

- (C) **Citation** identificación correcta del RS,
título, autores, versión, fechas...
- (D) **Dissemination** buenas prácticas de difusión,
toda difusión tiene sus objetivos, se dirige a un público
- (U) **Use** evaluación de los aspectos software del **RS**,
resultados correctos, facilitar la re-utilización...
- (R) **Research** evaluación de los aspectos investigación del **RS**,
trabajo científico, algoritmos, publicaciones, impacto...

Plan

1 Contexto

2 Objetivos

3 Conceptos

- Programa de investigación / Research Software (RS)
- Autor(es) de RS
- Publicación de RS / RS Papers
- Referencias y citas

4 Sobre la evaluación de la investigación

- Contextos de evaluación de la investigación
- Dos métodos de evaluación

5 Protocolo(s) CDUR para la evaluación de RS

6 Conclusiones

Concepto: research software (RS) (1/2)

- [2] (1994) Partha D, David PA: Toward a new economics of science

*there may be **important positive spillovers** across projects in the form of “learning effects” [...] including the development of computer software for performing data processing, storage, retrieval and network transmission.*

- [16] (2011) Kelly D: An Analysis of Process Characteristics for Dev. Scientific Soft.

Scientific software is defined by (1) it is developed to answer a scientific question ; (2) it relies on the close involvement of an scientific expert ; and (3) it provides data to be examined by the person who will answer that question ...

- [12] (2011) TGD: Article vs. Logiciel : questions juridiques et de politique scientifique...
(2009) TGD: Guide laboratoire pour recenser ses développements logiciels (PLUME)

logiciel du laboratoire tout programme utile pour faire avancer la recherche, qui a été produit avec la participation d'un membre du laboratoire. Il arrive souvent que des publications de recherche soient associées.

- [18] (2012) Sletholt MT, Hannay JE, et al.: What Do We Know about Scientific Software Development's Agile Practices ?

software developed by scientists for scientists

- [19] (2016) Hettrick S: Research Software Sustainability

Research software is developed within academia and used for the purposes of research: to generate, process and analyse results. This includes a broad range of software, programs written by researchers for their own use.

- [10] (2018) NASA Committee: Open Source Software Policy Options for NASA Earth and Space Sciences

Research software – that is, the software that researchers develop to aid their science...

Concepto: research software (RS) (2/2)

- [16] (2011) Kelly D: An Analysis of Process Characteristics for Dev. Scientific Soft.
 - ▶ **exclusión** de lo que puede ser incluido en otras definiciones:
[...] control software whose main functioning involves the interaction with other software and hardware ; user interface software [...]; and any generalized tool that scientists may use in support of developing and executing their software, but does not of itself answer a scientific question.
 - ▶ importancia de la **corrección** de los resultados:
If the software gives the wrong answer, all other qualities become irrelevant.
- [12] (2011) TGD: Article vs. Logiciel : questions juridiques et de politique scientifique...
 - ▶ las definiciones no tienen en cuenta el **estado** del programa:
"en proyecto", "acabado", difusión, calidad, alcance, talla, documentación, mantenimiento, utilización interna en un equipo para un artículo, o utilizado en varios laboratorios...

Conclusion: definición de RS

- Qué se produce: código, programa como un conjunto de ficheros bien identificado
- Quién produce: autores, también otras contribuciones, expertos científicos...
- Para hacer: investigación, ciencia, es decir, artículos asociados
- **Importante:** calidad y corrección de los resultados científicos obtenidos

Concepto: autor de RS

¿Qué quiere decir autor de un RS ?

[12] (2011) TGD: Article vs. Logiciel : questions juridiques et de politique scientifique...

- concepto legal: el autor escribe el código fuente
- concepto científico: expertos, quizás sin escritura de código
sin los expertos científicos, el RS no existiría
- puede haber otras contribuciones:
documentación, correcciones, tests, mantenimiento, traducciones...

Definición de autor de RS:

- En el artículo seleccionamos tres tipos (límites poco definidos):
 - ▶ (i) responsable del RS,
 - ▶ (ii) contribuidor importante o principal (escritura del código),
 - ▶ (iii) contribuidor menor (escritura del código o otras tareas).

El equipo puede decidir la atribución de un porcentaje de la escritura del código a los contribuidores que no escriben código.

Concepto: publicación de RS / RS Papers

Estudiamos la situación actual sobre las publicaciones de RS con *software peer review*.

- Journal of Open Research Software (JORS)
- The Journal of Open Source Software (JOSS)
- Research Ideas and Outcomes (RIO)
- Software Impacts
- SoftwareX
- también: (2010) Image Processing On Line Journal (IPOL)

Ver la lista de N. Chue Hong, Software Sustainability Institute (SSI)

<https://www.software.ac.uk/resources/guides/which-journals-should-i-publish-my-software>

En France, Projet PLUME (2006-2013), <https://projet-plume.org/>:

- publicación de “fichas descriptivas de RS” RELIER, con enlaces a los artículos, estadísticas: 358 RS en francés, con 116 también en inglés
- publicación de “fichas de software validado en el sentido de PLUME”, estadísticas: 406 fichas, donde 96 corresponden a RS
- clasificación temática, palabras clave, interfaz de búsqueda

Conceptos: referencias y citas

[39] (2013) Pontille D, Torny D: La manufacture de l'évaluation scientifique ...

[...] the difference between reference and citation: the act of reference is the responsibility of a given author while the citation is a new property, possibly calculable, of the source text. According to P. Wouters (1999), this reversal has radically altered the practice of referral and has literally created a new "culture of citation".

Una referencia **establece** título, autores, fecha, e identifica el RS en tanto que objeto o producción científica.

El artículo considera tres tipos diferentes de referencia:

- la referencia del RS paper (con *software peer review*),
- la referencia de un artículo clásico que describe el RS,
- una "referencia": autor(es), título, descripción corta, versión, fecha, url.

Notas:

- Puede haber varias referencias asociadas a un mismo RS.
- También hay identificaciones más completas: metadata, CITATION files...
- Software Citation Group, Software Citation Implementation Working Group...

Contextos de evaluación de la investigación

A lo largo de la vida científica encontramos varios contextos:

- tesis, oposiciones, evolución de carrera
- artículos, publicaciones, *peer review*
- participación a conferencias, workshops (selección)
- financiación de proyectos: calls, etapas, fin de proyecto
- establecer redes de colaboración, generalmente en un contexto internacional

En general, la primera evaluación que entra en juego: **auto-evaluación**.

Toda difusión tiene sus propios objetivos y se dirige a un público:

- ¿este resultado lo pongo en un preprint o lo envío a publicación?
- ¿esta idea de proyecto va a ser financiada? ¿la presento?
- ¿estos investigadores van a colaborar en este proyecto o publicación?
- las decisiones evolucionan con el tiempo, siguiendo las evoluciones de la investigación, o según los nuevos contextos de evaluación, por ejemplo necesito más artículos...

Dos métodos de evaluación

Grosso modo, hay dos métodos de evaluación:

[8] (2016) Mårtensson P, Fors U, et al.: Evaluating research: A multidisciplinary approach to assessing research practice and quality (62 references)

- evaluación de la calidad: ¿con qué criterios?
- evaluación de la cantidad con indicadores, métricas: ¿cuáles?
 - ▶ utilización de factores de impacto etc.: [11, 39, 51, 52, 53, 54]
- social factor

[55] (1999) Martin U: Computers, Reasoning and Mathematical Practice

[...] the community's "social knowledge":

the methods of checking the proof are social rather than formal.

Ver los 2 informes EC Expert Groups [9, 11]:

- recomendaciones, tener en cuenta RS (utilización, producción)
- [9] Open Science Career Assessment Matrix (OS-CAM)
- [11] cómo establecer los comités de evaluación

Protocolo(s) CDUR para la evaluación de RS

Ayudar a: investigadores, comités de evaluación, responsables de las decisiones.

- (C) Citation** mide si el RS está correctamente identificado:
referencia correcta, y también metadata, citas correctas a otros RS...
perspectiva legal: autores, afiliación, % de participación
- (D) Dissemination** prácticas de difusión, según las normas de la evaluación
[14] (2014) TGD: Free software, Open source software, licenses...
perspectiva política: Ciencia Abierta, **perspectiva legal:** licencias
- (U) Use** aspectos “software” **del RS:** resultados correctos, fácil de utilizar,
software: doc., test, instal., ..., leer el código, lanzar ejemplos...
perspectiva reproducibilidad: validación de los resultados científicos
- (R) Research** aspectos “investigación” **del RS:** calidad del trabajo científico,
algoritmos y estructuras de datos utilizados/implementados,
publicaciones, colaboraciones...
perspectiva investigación: impacto

Aplicación flexible: los comités de decisión y de evaluación **establecen** su propio protocolo **CDUR** adaptado al contexto y a los objetivos de la evaluación en curso.

Conclusiones

Llamada a debate sobre la evaluación de RS: ¿el protocolo CDUR será doptado?

Para avanzar en la adopción de las prácticas de Ciencia Abierta & RS...
es necesario cambiar y desarrollar los métodos de evaluación.

Consecuencias esperadas:

In agreement with [11] (2019) Guédon JC, Jubb M, et al.: *world brain vision*

- maximize [RS] accessibility and usability,
- support and expand range of contributions (equity, diversity, inclusivity criteria),
- support community building, and
- promote high-quality research with heightened integrity.

Y también incrementar la transparencia de los métodos de evaluación de RS.

[4] (2016) Howison J, Bullard J: Software in the scientific literature (p.15):

Clearly, a policy is only as good as its enforcement.