

More Data, Less Process? The Applicability of MPLP to Research Data

Sophia Lafferty Hess y, Thu-Mai Christian

IASSIST quarterly, ISSN 0739-1137, winter 2016, p. 6-13

La calidad de los datos está relacionada con las necesidades de los usuarios. En este sentido los datos tienen cuatro características: son intrínsecos (credibilidad), contextuales (relevancia) representativos (usabilidad) y accesibles. La comunidad científica exige estándares de calidad para los archivos de datos, para lo que se han establecido protocolos de acceso y uso, como el Modelo de Referencia para Open Arhival System (OAIS). Un concepto asociado a OAIS es el de “comunidad designada”, grupo de usuarios que debe ser capaz de entender una información determinada. Esta conservación de datos de alta demanda y mucho trabajo incluye una exhaustiva lista de procedimientos y puede que los recursos no sean suficientes. Para solucionar este problemas surgió la doctrina de “más producto, menos proceso” (MPMP), que propugna alcanzar algo “suficientemente bueno” en el proceso de datos. Para determinar qué es “suficientemente bueno” hay tres áreas que corroborar: organización, descripción y preservación. La organización consiste en re-etiquetar y re-archivar. Según la teoría del MPMP, esta tarea es realizada a menudo de manera excesivamente meticulosa. El consejo es entender el contexto y adaptar las tareas a las necesidades de comprensión de los usuarios. El modelo de conservación mínima reduce el conjunto de metadatos generados a los imprescindibles para garantizar descubrimiento y acceso. Pueden incluir un identificador permanente. Respecto a la preservación, en lugar de producir diferentes archivos, se crea un formato único que el usuario puede convertir desde diversas plataformas de software. La teoría MPMP nos fuerza a reconsiderar el proceso de conservación y mejor el sistema de acceso. Los profesionales de la información han producido materiales educativos y programas para enseñar a los investigadores estrategias de búsqueda más efectivas. Algunas herramientas también facilitan funcionalidades adicionales. Es el caso de Open Science Framework, que hace más sencillo archivar y compartir datos. Futuros estudios para la mejora de la eficiencia deberían determinar las expectativas de los grupos designados sobre la calidad de los datos; cómo el nivel de análisis afecta a la reutilización; cómo afecta el contexto; cómo interactúa el usuario con los metadatos; cómo afecta la conservación de datos a la calidad de los mismos. Así se podrá calificar qué datos son “suficientemente buenos”. Con el crecimiento de los datos compartidos y el foco puesto en la transparencia de las investigaciones, el archivo de datos será primordial. La clave estará en obtener resultados sostenibles que permitan la reutilización.

Resumen elaborado por Antonio Rodríguez Vela