

Reproducibilidad, preservación y acceso a investigaciones con ReProZip y ReProServer

Reproducibility, preservation, and access to research with
ReProZip and ReProServer

Vicky Steeve, Rémi Rampin, Fernando Chirigati

IASSIST Quarterly, ISSN 0739-1137, Vol. 1-2, 2020

La práctica de la preservación conlleva una búsqueda constante de mejora de los métodos para conseguir una preservación confiable de los trabajos académicos, al igual que unos medios más eficientes y progresivos para proporcionar acceso a los materiales. En el caso de la preservación dentro de las comunidades de investigación, esto a menudo supone verificar que los datos de investigación puedan ser abiertos en el futuro a través de formatos de migración, y asegurarse de que hay suficiente documentación acerca de estos conjuntos de datos que puedan ser útiles para otros investigadores. Esta práctica ha evolucionado para incluir software de preservación junto a los datos, especialmente el software específico de cada proyecto a menudo desarrollado por los investigadores para sus proyectos. De manera similar, capturar y describir los acoplamientos y la distribución de trabajo de programación ha sido una necesidad cada vez mayor para los investigadores que quieren que su trabajo sea reproducible. La reproducibilidad implica que otros investigadores que intenten hacer el mismo análisis usando los mismos datos y códigos que el creador original, puedan conseguir el mismo resultado. Esto es extremadamente dificultoso. Los acoplamientos de un proceso de investigación no siempre son conocidos en todo detalle por

el investigador original, y mucho menos la gestión digital respecto a la seguridad a largo plazo del trabajo. Archiveros y bibliotecarios que se enfrentan al acceso y preservación de materiales de investigación reproducible deben tener una herramienta apropiada para esta tarea. Sin embargo, la adopción de métodos de reproducción sigue siendo baja. Lo cierto es que la reproductibilidad es técnicamente difícil de lograr debido a las complejidades de los entornos computacionales. Para salvar estos obstáculos, los autores del artículo crearon dos herramientas de código abierto: ReProZip y ReProServer. ReProZip puede capturar automáticamente un experimento computacional, creando un lote único y distribuable, sin que el investigador tenga que especificar todos los acoplamientos requeridos. Otros investigadores pueden usar después ReProZip para desempaquetar automáticamente este lote e instalar el entorno informático para reproducir la investigación, incluso si el sistema operativo es diferente. La segunda herramienta, ReProServer, es una aplicación web que proporciona acceso sencillo a los lotes de ReProZip: en lugar de tener que instalar diferentes softwares para reproducir un lote, permite a los investigadores reproducir y reutilizar las investigaciones ajenas desde su navegador. ReProZip incluye archivos de datos, software, versión OS y variables de entornos. Todo ello se incluye en un lote que se convierte en un rpzfile, que los usuarios pueden utilizar para reproducir su trabajo. El rpzfile contiene metadatos enriquecidos, garantizando una mejor preservación a largo plazo.

<https://www.iassistquarterly.com/index.php/iassist/article/view/969>

Resumen elaborado por Antonio Rodríguez Vela