

A Framework for the Preservation of a Docker Container

Iain Emsley, David De Roure

International Journal of Digital Curation, ISSN 1746-8256, Vol. 12, n. 2, 2017, p. 125-135

Construir y mantener sistemas confiables en diversos entornos es un problema continuo. Un proyecto o experimento puede durar muchos años. Software y hardware pueden cambiar, al igual que el sistema operativo. La “containerización” es una tecnología que se usa en varias compañías, como Google, Amazon e IBM, y proyectos científicos para desplegar rápidamente un conjunto de servicios de manera repetida. Dockerfiles se usa para asegurar que un container es construido repetidamente, para permitir conformidad y una fácil actualización cuando hay cambios, y se está convirtiendo en algo común en los proyectos. Se ve como una parte del desarrollo sostenible del software. La tecnología de la “containerización” ocupa un espacio dual: es tanto un repositorio de software como un software en sí mismo. Al considerar Docker de esta manera, debemos verificar que Dockerfile puede ser reproducido. Usando un subconjunto de la especificación Dockerfile, se crea el dominio de una lengua específica para asegurar que los archivos Docker pueden ser reutilizados en una fase posterior para recrear el entorno original. Proporcionamos un marco simple para dirigir la cuestión de la preservación de containers y sus entornos. Presentamos experimentos sobre Dockerfile existentes y concluimos con un debate sobre su futuro. Instalamos un conducto para comprobar que un Dockerfile concreto conforme a nuestro modelo deseado, extrae el Docker y detalles del sistema operativo. Esto ayudara a la reproductibilidad de resultados creando versiones del entorno. También ayuda a desarrollar y probar que el sistema es construido de manera repetible y que cualquier cambio en el entorno del software puede ser también compartido en el Dockerfile. Este trabajo apoya no solo el proceso de citas, sino el científico, al proporcionar detalles del entorno de trabajo. Como parte del conducto para crear el container, capturamos el proceso usado y lo integramos en la ontología W3C PROV. Esto tiene el potencial para proporcionar un identificador permanente y la trazabilidad del proceso usado para preservar los metadatos. Nuestro trabajo futuro se centrará en la cuestión de enlazar estos resultados con una ontología de flujo de trabajo para preservar el flujo de trabajo completo con los comandos y parámetros que se den a los containers. Vemos aquí el origen en el proceso de construcción útil para proporcionar un completo panorama del flujo de trabajo.

Traducción del resumen de la propia publicación