

Utilizando Synthea y Docker para generar pacientes sintéticos de forma consistente y realista.

Artículo

[Ricardo Paiva](#)

· Jun 4



Lectura de 4 min

[Open Exchange](#)

Utilizando Synthea y Docker para generar pacientes sintéticos de forma consistente y realista.

Durante las últimas semanas, el equipo de *Solution Architecture* (Soluciones de Arquitectura) ha estado trabajando para terminar la carga de trabajo de 2019: esto incluyó la creación del código abierto de la Demostración de Readmisiones que llevó a cabo HIMSS el año pasado, para poder ponerla a disposición de cualquiera que busque una forma interactiva de explorar las herramientas proporcionadas por IRIS.

Durante el proceso de creación del código abierto de la demostración, nos encontramos de inmediato con un error crítico. Los datos subyacentes de los pacientes que se utilizaron para crear la demo no podían utilizarse como parte de un proyecto de código abierto porque no eran propiedad de InterSystems, eran propiedad de nuestro socio Baystate Health.

Nuestro equipo estaba en un pequeño aprieto y tenía que encontrar una forma de sustituir los datos originales por datos sintéticos que pudieran utilizarse, pero manteniendo la "historia" de las demos, o su funcionalidad subyacente, consistente. Dado que la demo muestra cómo IRIS admite el *workflow* de *machine learning* de un científico de datos, había un nivel de complejidad añadido porque cualquier dato que utilizáramos tenía que ser lo suficientemente realista como para poder apoyar nuestro modelo de investigación. Después de una breve investigación, Synthea vino a nuestro rescate.

Synthea es un generador de pacientes sintéticos, que modela sus historiales médicos. Es de código abierto. Synthea proporciona datos de alta calidad, realistas, pero no reales, de pacientes; en una variedad de formatos (incluido FHIR), con diferentes niveles de complejidad, cubriendo todos los aspectos de la atención médica. Los datos obtenidos no tienen coste, ni privacidad ni restricciones de seguridad, lo que permite investigar con datos de salud que de otra manera no estarían disponibles, de forma legal o práctica.

Después de una investigación inicial, se eligió Synthea como la herramienta para solucionar nuestro problema de datos. Synthea es una herramienta increíble; sin embargo, un problema que encontramos fue que, para ejecutar el software y obtener los pacientes, teníamos que instalar varias dependencias en nuestros equipos.

- Java JDK 1.8
- Gradle Build Tool

Cuando trabajas por tu cuenta, esto generalmente no es un problema, pero como nuestro equipo está formado por varias personas, es importante que todos puedan actualizarse con un

nuevo software rápidamente; y la instalación de dependencias puede ser una pesadilla. Tenemos el propósito de que el menor número de personas posible sufra durante los procesos de instalación al integrar un nuevo software en nuestro flujo de trabajo.

Como necesitábamos que cualquier persona de nuestro equipo pudiera realizar actualizaciones en la Demo de readmisiones para poder generar pacientes fácilmente, y no queríamos que todos tuvieran que instalar Gradle en sus equipos, nos apoyamos en Docker e introdujimos el software de Synthea dentro de una imagen de Docker, permitiendo que la imagen se ocupe de las dependencias ambientales subyacentes.

Esto terminó funcionando muy bien para nuestro equipo, ya que nos dimos cuenta de que ser capaz de generar datos de pacientes sintéticos sobre la marcha es probablemente un caso de uso muy común al que se enfrentan nuestros compañeros Ingenieros de ventas, por lo que nuestro equipo quería compartirlo con la Comunidad de Desarrolladores.

Cualquiera puede utilizar la siguiente línea de código para generar rápidamente 5 historiales médicos de pacientes sintéticos en formato FHIR, y dejar los pacientes resultantes en una carpeta de salida en el directorio donde está trabajando actualmente.

```
docker run --rm -v $PWD/output:/output --name synthea-docker  
intersystemsc/irisdemo-base-synthea:version-1.3.4 -p 5
```

El código de esta imagen Docker tiene su propio repositorio en Github y se puede encontrar aquí para cualquier persona que quiera echar un vistazo, hacer cambios personalizados, o contribuir: <https://github.com/intersystems-community/irisdemo-base-synthea>

Ahora estamos realizando actualizaciones para que el proyecto sea compatible con módulos personalizados, de modo que cualquiera que desee agregar una enfermedad a sus pacientes generados pueda hacerlo, si Synthea no la proporciona de forma predeterminada, y se incorporará automáticamente a su imagen.

¿Dónde se usa actualmente?

El proceso actual de creación de la Demo de readmisiones utiliza la imagen irisdemo-base-synthea para generar 5 000 pacientes sintéticos sobre la marcha y cargarlos en nuestro repositorio de datos IRIS relacional, normalizado. Cualquier persona que esté interesada en verificar cómo analizar estos datos de pacientes generados de forma sintética (en formato FHIR), puede consultar la Demo de readmisiones creada recientemente con código abierto. La clase que hay que buscar es: IRISDemo.DataLake.Utils. a partir de la línea 613.

La Demo de readmisiones se puede encontrar aquí: <https://github.com/intersystems-community/irisdemo-demo-readmission>

[#Analítica](#) [#Big Data](#) [#Contenedorización](#) [#Docker](#) [#Herramientas](#) [#InterSystems](#) [IRIS for Health](#)
[Compruebe la aplicación relacionada en InterSystems Open Exchange](#)

00 2 0 0 32

Log in or sign up to continue
Añade la respuesta

Utilizando Synthea y Docker para generar pacientes sintéticos de forma consistente y realista.

Published on InterSystems Developer Community (<https://community.intersystems.com>)

URL de fuente: <https://es.community.intersystems.com/post/utilizando-synthea-y-docker-para-generar-pacientes-sint%C3%A9ticos-de-forma-consistente-y-realista>