

# El uso intensivo de datos en las plantas farmacéuticas en un contexto Pharma 4.0TM

En el actual contexto de digitalización en las actividades industriales (Industria 4.0) y concretamente en el sector farmacéutico (Pharma 4.0TM ISPE), la disponibilidad de datos ha aumentado enormemente, así como las herramientas analíticas para extraer la información de esos datos. Esta información es clave para mantener los estándares de calidad y productividad y finalmente para asegurar a las empresas seguir siendo competitivas.

**ALICIA TÉBAR**  
FUNDADORA Y CONSULTORA DE QBD  
PHARMACEUTICAL SERVICES

## ¿Dónde se generan datos en una planta farmacéutica?

Una planta de fabricación de medicamentos o principios activos en una gran generadora de datos. Estos datos se han de convertir en la información necesaria para dar soporte a la actividad, en especial a todos los procesos regulados GMP. También los no regulados pueden tener gran impacto en la viabilidad de la planta desde el punto de vista de productividad, eficiencia energética y sostenibilidad.

Hace ya una década que el sector puso el foco en la integridad de datos. En concreto los datos con impacto GMP deben cumplir los principios ALCOA (Attributable, Legible, Contemporaneous, Original, Accurate). Esto ha supuesto una actualización, que aún continúa, en los equipos y especialmente en los sistemas informatizados para cumplir con la base principal, que es que los datos generados sean fiables, trazables y representativos de la realidad.

La profunda revisión de equipos y sistemas que ha supuesto el cumplimiento de Data Integrity ha favorecido la migración masiva del dato en papel al dato electrónico.

La actual disponibilidad de sensores IoT, sistemas informatizados en cloud que permiten almacenar cada vez más información a un precio más reducido y finalmente sistemas de conectividad que permiten crear repositorios de información contextualizada, fiable y disponible para su utilización, incluso por diferentes departamentos/finalidades, están conformando lo que hoy en día llamamos transformación digital<sup>1</sup>. ISPE ha publicado recientemente la primera de una familia de guías dedicadas precisamente a este tema<sup>2</sup>.

## La profunda revisión de equipos y sistemas que ha supuesto el cumplimiento de Data Integrity ha favorecido la migración masiva del dato en papel al dato electrónico

### Utilización y finalidad de la información extraída. Uso regulado vs no regulado.

En el entorno GMP la gestión de los datos está regulada en función de si la información derivada de los mismos se utiliza para tomar decisiones con impacto GMP o no. Un ejemplo claro sería la liberación de lote.

Un mismo proceso de fabricación genera datos con impacto GMP: parámetros críticos de proceso, controles en proceso, monitorización ambiental etc y datos sin impacto GMP directo pero que pueden ser muy relevantes en términos de productividad: tiempos, consumo eléctrico, agua, disolventes etc.

Tomando como referencia el proceso de fabricación, hace ya tiempo que consideramos su ciclo de vida: desarrollo, validación y verificación continuada (Stages 1,2,3). Todo lo que hace referencia al funcionamiento del proceso de fabricación y a la calidad del producto resultante es entorno regulado. Por tanto, los datos generados en este ciclo de vida son relevantes, en Stage 1 o desarrollo para el registro del producto y en Stages 2 PPQ y 3 CPV desde el punto de vista GMP.

El Stage 3 o verificación continuada del proceso responde a la idea de que la calidad del

producto procede de la calidad del proceso en términos de control de la variabilidad. Una vez controladas, el proceso conduce a la calidad deseada de manera consistente. Un sistema de monitorización en tiempo real permite detectar “desviaciones” a tiempo de ser corregidas.

Por tanto, los sistemas de control de la calidad de productos en Farma están migrando del muestreo y análisis de producto terminado, a monitorización del proceso y a medidas en línea de calidad del producto que se va generando durante el proceso (PAT). El exponente más claro de esta operativa es la fabricación en continuo de formas sólidas que no permite otro tipo de estrategia de control que la descrita<sup>3</sup>.

### Conectividad: la base para la disponibilidad de datos

La evolución del dato en papel al electrónico ha aumentado exponencialmente la disponibilidad de los mismos. Paralelamente ha aumentado la posibilidad de disponer de almacenaje de datos cada vez más asequible en cuanto a volumen y cada vez más accesible para usuarios con diferentes objetivos en la utilización de los datos: producción e ingeniería para productividad y ahorro energético, QA para calidad de producto etc

La accesibilidad de los datos también ha comportado un mayor interés en como proteger la información de ataques externos y una mayor inversión en ciberseguridad.

La palabra clave es conectividad. Uno de los objetivos clave de muchas plantas es el poder disponer de un repositorio centralizado, seguro y organizado que permita facilitar los datos para los distintos usuarios y objetivos y que pueda mantener la integridad y el contexto necesario (metadatos) para los diferentes usos.

Una vez podemos disponer de un repositorio de datos de esta naturaleza, confiable, estructurado con metadatos, seguro y accesible viene el paso de convertir datos en información para los diferentes usos de la planta.

### Conversión de datos en información. Estadística y Ciencia de Datos.

Disponer de un repositorio seguro de datos confiables ya es un gran paso en sí mismo. Es la base necesaria para la transformación de los datos en información en una planta GMP.

Últimamente hablamos de herramientas de Analítica Avanzada para hacer esta transformación. El gran volumen de datos disponibles hoy en día y el interés en la explotación de estos es lo que seguramente ha provocado que hablemos de la figura del Data Scientist en lugar del Estadístico (de toda la vida).

### ¿Estadístico o Data Scientist?

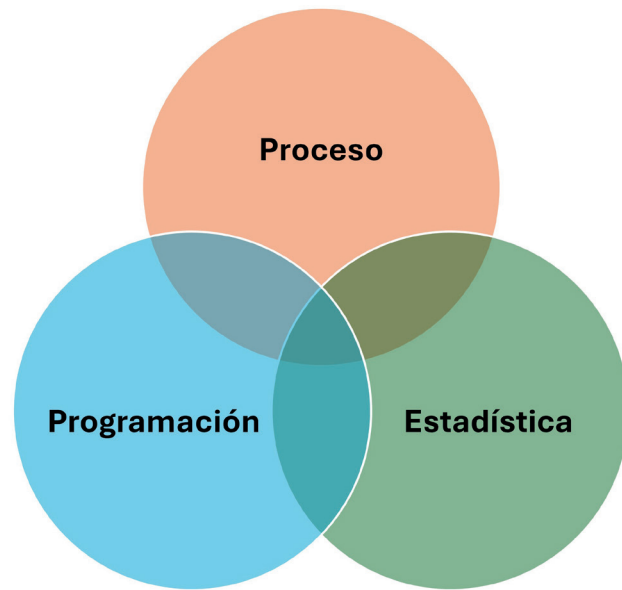
Ambos roles tienen como objetivo convertir los datos en información útil para la toma de decisiones.

El estadístico es un rol asociado habitualmente a un entorno de I+D, con formación en matemáticas y estadística cuyo objetivo es interpretar los datos y hacer recomendaciones en base a las conclusiones.

Un científico de datos es un rol mucho más reciente que se asocia con la utilización de herramientas de programación para el tratamiento de volúmenes de datos mucho mayores y en muchos casos poco estructurados. El entorno de trabajo puede ser variado, no necesariamente I+D.

La diferencia entre ambos roles probablemente radica en que el Data Scientist se ha especializado en herramientas de analítica que trabajan con grandes volúmenes de datos y que por tanto requieren un mayor uso de la programación (Machine Learning /Inteligencia Artificial). El estadístico en general tiene un mayor conocimiento matemático y utiliza herramientas clásicas para la descripción del objeto del análisis y para la predicción.

**Es probable que ambos roles acaben convergiendo, ya que hoy en día la programación será imprescindible para tratar las grandes cantidades de datos disponibles y por otro lado el fundamento matemático es imprescindible también para interpretar correctamente los resultados.**



Conocimientos para el tratamiento de datos

## La digitalización en las plantas farmacéuticas está rompiendo los silos de información y conocimiento aumentando la conectividad y habilitando repositorios de datos fiables que pueden ser explotados

Hay un elemento más en cuanto a las capacidades para utilizar correctamente las herramientas de la analítica avanzada, y es el conocimiento del proceso. No quiere decir que el experto en datos tenga que ser también un experto en el proceso, pero si estos conocimientos no se pueden dar en una misma persona, es imprescindible crear un equipo de trabajo que si incluya personas con este expertise, y sobre todo que sean capaces de trabajar conjuntamente para conducir a los resultados deseados.

**Un científico de datos por experto que sea en el tratamiento matemático y la programación, si no conoce el proceso y la naturaleza de las variables disponibles no va**

**a ser capaz de seleccionar correctamente las relevantes ni va a poder interpretar correctamente las predicciones del modelo.**

### Conclusiones

La digitalización en las plantas farmacéuticas está rompiendo los silos de información y conocimiento aumentando la conectividad y habilitando repositorios de datos fiables que pueden ser explotados para incrementar la robustez de los procesos, la productividad y la sostenibilidad de la empresa.

Las herramientas de analítica avanzada son las más eficaces en este proceso de gestión y análisis de grandes cantidades de datos y requieren de equipos de trabajo que combinen conocimientos de proceso, de programación y matemático-estadístico para ser realmente útiles en la conversión del dato a información para la toma de decisiones.

La transformación digital no solamente requiere tecnología sino personas capacitadas para utilizarla con todo su potencial. Este es seguramente uno de los retos en cuanto a identificar en las organizaciones personas cualificadas para asumir estos roles, o bien personas que puedan formarse con expertos externos y asumirlos en un medio plazo ●

### Referencias

- (1) PHARMA 4.0TM: Niveles de madurez digital para obtener el control del proceso en tiempo real. Farmaespaña Industrial. Nº 108. 2021 <https://www.farmaindustria.com/articulos-online/pharma-4-0-niveles-de-madurez-digital-para-obtener-el-control-del-proceso-en-tiempo-r-FK5Nd>
- (2) ISPE Baseline Pharma 4.0TM First Edition 2023
- (3) ICH guideline Q13 on continuous manufacturing of drug substances and drug products - Step 5 2023