



Peter Boogaard, fundador de Industrial Lab Automation; Roberto Castelnuovo e Isabel Muñoz-Willery, cofundadores de NL42 Management Consulting S.L. [www.NL42.com](http://www.NL42.com)

Cuando se habla de la integración de datos, primero tenemos que dejar de pensar en «tecnología». **La integración no solo trata de instrumentos u otras plataformas de software, sino más bien de integrar procesos, acelerar ideas y facilitar el cumplimiento de las normativas de una forma más económica.** La colaboración multifuncional entre la investigación, el desarrollo, el control de calidad y la producción busca optimizar e integrar procesos distribuidos a lo largo de toda la cadena cubriendo todas las disciplinas, las cuales requieren cantidades importantes de datos. Integrar estos datos en el flujo de trabajo científico hará que aumente su disponibilidad y calidad dentro de toda la comunidad científica y empresarial.

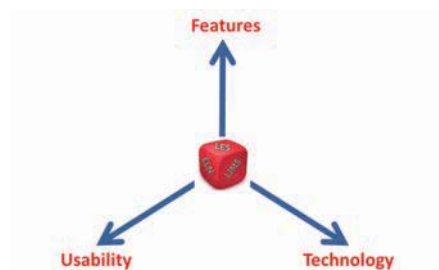
## Cuando los datos se encuentran

La adopción de la integración de datos ha sido más aceptada en los laboratorios regulados GXP que en los de I+D. Sin embargo, las compañías farmacéuticas todavía están implementando, principalmente, soluciones tradicionales basadas en papel. A pesar del enorme potencial de incremento de rendimiento, así como de cumplimiento de las normativas, todavía se encuentran importantes barreras en la puesta en marcha de laboratorios informatizados. Curiosamente, en sectores de la industria como la salud o la atención al paciente, muy sensibles con los costes, la integración de un proceso de trabajo automatizado, incluida la integración automatizada de datos relacionados con la ciencia, está generalmente aceptada. Por otro lado, los laboratorios utilizan instrumentos científicos de cada vez más sofisticados y avanzados, mientras que sus métodos de adquisición de datos y almacenamiento están, en cierto modo, anticuados.

No obstante, el cambio está a la vista y la adopción de nuevas tecnologías dominantes como la nube, dispositivos móviles y servicios correspondientes prevalece. Las presiones económicas están acelerando un cambio general de mentalidad en la industria y las empresas tienen que reconsiderar el cómo facilitar la distribución de la información entre departamentos en un entorno de colaboración multidisciplinar y global. La integración de datos es la primera pieza del rompecabezas. ¡Eso es una buena noticia!

### Objetivo común

La integración de datos no es un objetivo en sí mismo, es una palabra de moda en la industria desde hace algún tiempo y no solo ha pasado a significar muchas cosas para mucha gente, sino que el término esconde frecuentemente el área de complejidad que



de verdad comprende. Puede sonar obvio, pero uno no debería abalanzarse a decir: «Sí, podemos integrar un sistema de gestión de la información de laboratorio (LIMS) o un cuaderno electrónico de laboratorio (ELN) a un instrumento XYZ o a un sistema informático corporativo». Equivocarse en definir objetivos claros, medidas, conclusiones técnicas y beneficios corporativos suele acabar en un fallo del proyecto, debido a la falta de entendimiento de lo que se debería haber incluido en ese proceso de integración implícito.

Los estudios de la industria muestran que, de media, cada grupo de procesos farmacéuticos requiere cientos de transcripciones manuales de datos durante el ciclo de vida de un producto. También muestra que, en muchos casos, se procesan más de una docena de documentos diferentes. Una reducción significativa en los errores de transcripción aumentará la calidad, reducirá los costes y disminuirá las actividades no relacionadas con el trabajo. Pero la pregunta es, ¿por dónde empezar? cuando se trata de conseguir una correcta integración de datos.

El primer paso debería consistir en la creación de un plan general que tome en consideración los siguientes apartados:

- ♦ ¿Qué esperar de la integración de datos? La esencia de la integración es compartir y fusionar datos entre las distintas partes y sistemas. Hablar solo de intercambio

de datos no es suficiente; el formato y objeto han de ser claro. Es fundamental un consenso y acuerdo en cómo se transporte la información a través de la infraestructura informática.

- ♦ ¿Quiénes están involucrados? Por naturaleza, la integración se encuentra al menos entre dos partes o sistemas, pero a menudo incluye muchos más. Es esencial conocer quiénes componen todas las partes y entender los objetivos individuales.
- ♦ ¿Cuál es la contribución de cada parte? Cada parte tiene un papel que desempeñar en la integración que pueden oscilar entre estar informados de los cambios hasta participar en el sistema. Identificar lo que cada parte espera del proceso, definir la propiedad del proyecto y articular claramente un plan escalable en el caso de dificultades es vital.
- ♦ Evitar la adaptación personal o programación en el nivel de interfaz entre los sistemas. La mayoría de datos necesitan cambiarse o formatearse antes, o justo después, de transferirlos. Es importante dedicarle tiempo a este tema con los proveedores en el proceso de solicitud de información (RFI) o durante la implementación.
- ♦ Definir estrategias de comunicación internas y externas para garantizar que se fijen las expectativas adecuadas. Cuando los proyectos de integración de datos fracasan, se debe muy a menudo a la falta de un claro entendimiento del objetivo final. No hacer suposiciones, seguir informándose durante el proceso.

### Estándares, lo que hay que hacer primero

La integración de datos en los laboratorios no es algo sencillo. Puede parecer un tema

aburrido hoy en día, pero la necesidad de estandarizar nunca ha sido mayor. Sin estos estándares, la recopilación de datos automatizados de instrumentos o de sistemas de datos puede convertirse en un reto. Varias iniciativas se dedican en definir estándares comunes tan necesarios. La Pistoia Alliance pretende bajar las barreras para la innovación mediante la mejora de la interoperabilidad de procesos empresariales de I+D a través de una colaboración precompetitiva. La alianza fue concebida en primer lugar por expertos informáticos de AstraZeneca, GSK, Novartis y Pfizer que se encontraban en una reunión en Pistoia, Italia.

Durante las siguientes reuniones, los fundadores de Pistoia se dieron cuenta de que cada una de sus organizaciones se enfrentaba a los mismos problemas — sobre como fusionar, acceder y compartir datos, esenciales para la innovación, pero con poca ventaja competitiva. Comprendieron que juntos podrían resolver estos problemas comunes, facilitando así que sus organizaciones pudieran innovar, reducir costes y recolocar valiosos

recursos a otros proyectos con un impacto más competitivo y estratégico.

En junio de 2012, se formó una nueva iniciativa financiada por la industria para promocionar estándares abiertos de información para el laboratorio analítico. La Fundación Allotrope, patrocinada por Abbott, Amgen, Baxter, BI, BMS, Merck, GSK y otros, aborda la falta de un formato común para los repositorios de metadatos. La infraestructura propuesta consistirá en (a) estándares abiertos de documentos basados en XML y JSON, (b) repositorios abiertos de metadatos para proporcionar entradas exactas desde numerosas fuentes de datos, y (c) bibliotecas de fuentes abiertas para permitir estos componentes.

Un estándar es solo un estándar si las organizaciones en realidad lo adoptan. El estándar AnIML permite una auditoría completa, firmas digitales, validación para el cumplimiento normativo y gana aceptación en la industria. AnIML es un formato de datos estandarizado que permite almacenar y compartir datos experimentales. Está indicado para una gran variedad de técnicas de medición

analítica. Los documentos AnIML pueden recopilar procesos de trabajo de laboratorios y resultados, sin importar qué instrumentos o técnicas de medición se utilizaron. AnIML está basado en XML, con las dos principales ventajas que ya existen muchas herramientas que manipulen XML, facilitando así la implementación y que siendo el formato XML basado en texto, los documentos AnIML son legibles por los humanos —un aspecto importante para el almacenamiento a largo plazo. AnIML está siendo desarrollado por el subcomité ASTM E13.15 de datos analíticos, compuesto por voluntarios de comunidades industriales, académicas, gubernamentales y de vendedores.

Otros estándares que tienen un impacto significativo en cómo puede implementarse correctamente la integración de datos son los de ISA, ASTM e IEEE. Estos son estándares interprofesionales aceptados a nivel mundial. Por ejemplo, el estándar ISA consta de modelos y terminología para estructurar el proceso de producción y para desarrollar el control de equipo (ISA-88) y para la pro-



**Agilice sus Procesos  
Mejore su Eficiencia  
Reduzca Errores y Tiempos**

**LABORATORY DATA MANAGEMENT SOLUTIONS**

*...SIMPLIFICAMOS LO COMPLEJO*

[www.NL42.com](http://www.NL42.com)

PROYECTOS "PAPERLESS"  
PROYECTOS GESTIÓN de CALIDAD  
IMPLEMENTACIÓN GESTIÓN de DATOS  
INTEGRACIÓN DE SISTEMAS  
VALIDACIÓN, ACREDITACIÓN, CERTIFICACIÓN



ducción, mantenimiento y calidad (ISA-95). Utilizar S88 dio lugar a procesos estructurales y proporcionó un lanzamiento y una transferencia inmediatos hacia y desde laboratorios externos, entre grupos internos incluido el laboratorio, escalabilidad de planta piloto e instalaciones de producción.

## El contexto es la clave

En general, hay tres principios operativos básicos para optimizar la integración de datos. El primero de todos, recopilar los datos en el punto de origen para eliminar errores humanos y reducir la complejidad del sistema. Los proveedores de sistemas empezaron con la integración de datos de instrumentos en mente. Sus productos originales fueron diseñados para recopilar datos de laboratorio en las fuentes de datos. En segundo lugar, simplificar e implementar procesos de auto-documentación para eliminar errores de transcripción y evitar volver a escribir datos de forma innecesaria. En una encuesta reciente, un 32 % votaron que la integración de datos en un laboratorio informatizado eliminará las entradas y la transferencia de datos manuales.

Por último, garantizar que los metadatos son recopilados de forma estructurada. Los datos sin procesar representan un conjunto de hechos desorganizados y no procesados (p. ej. recopilación de números) y son normalmente de naturaleza estática. Un archivo de datos sin contexto o información de metadatos no tiene ningún sentido. El científico ya no se encuentra en el laboratorio, sino que está integrado en el proceso de calidad completo. Para garantizar que el conocimiento tácito se mantiene en el nivel más alto en los sistemas informatizados, la información de contexto debe formar la base para permitir los análisis científicos integrados y la interpretación.

## Cambiar las tecnologías

Las tecnologías que admiten modelos conceptuales mediante las ontologías y los estándares de datos están madurando rápidamente. Aunque permanecen en las etapas tempranas de adopción, la interoperabilidad del sistema mediado por la ontología y la gestión de metadatos formales tienen un importante potencial para facilitar lo mejor de su categoría o estrategia de Arquitectura Orientada a Servicios (SOA). La aproximación basada en la ontología permitirá al

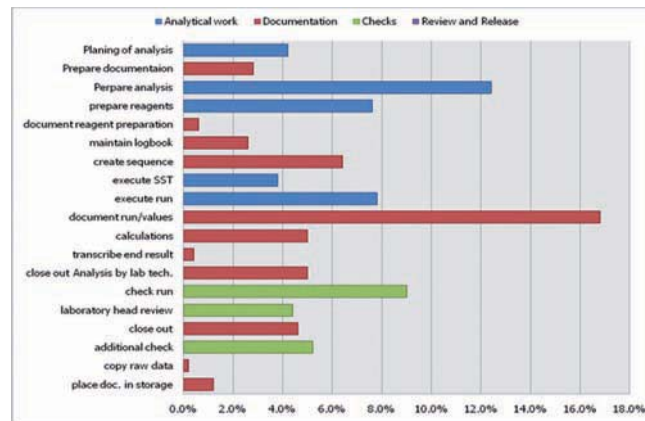


Tabla. La integración de datos facilita los procesos de autodocumentación

usuario integrar fuentes de bases de datos existentes y conseguir una interoperabilidad entre diferentes formatos de datos y aplicaciones. La clave para el éxito del desarrollo de aplicaciones de conciencia ontológica es hacer que el contenido existente en los almacenes de datos de la organización esté disponible. Tanto Google como Microsoft están invirtiendo mucho para hacer madurar esta tecnología.

Los vendedores de básculas y valoradores están aumentando el valor de sus instrumentos mediante la implementación de métodos aprobados y validados previamente en su firmware. Unos permiten que se implementen métodos directamente en sus básculas. Otros hacen uso de un software intermedio para alcanzar esa funcionalidad en sus básculas, valoradores y otros instrumentos. Esto puede sonar como un pequeño paso, pero podría tener un impacto significativo en los resultados de validación en el laboratorio y en las operaciones de fabricación, tal como menos puntos de fallo durante la operación, reducida personalización de software, y mejor documentación.

La ciencia centrada en datos está superando lo establecido. La investigación es cada vez más colaborativa y compleja beneficiándose de múltiples tecnologías para conseguir un entendimiento más completo sobre enfermedades y organismos. La integración de datos es crucial para permitir que se comparta el conocimiento virtual. El aumento exponencial en la producción de (grandes) datos, combinado con el incremento de la colaboración crea la necesidad almacenar los datos de forma rentable, para su análisis y distribución entre equipos. La comunicación es el denominador común. El conocimiento tácito se basa en el sentido común, mientras que el conocimiento explícito se basa en los logros académicos y ambos están infrutilizados. Se crean nuevas invenciones

y compartirán conocimientos al combinar la información explícita, almacenada en sistemas informáticos, con la información tácita.

La tecnología cambia la dinámica de cómo los científicos trabajan juntos. La nube, no es solo una iniciativa informática, sino que cambia la forma en que las personas y la ciencia trabajan juntos. Elimina la necesidad de esperar meses para que se publique un documento científico, pero sigue siendo un tema de confianza y relación entre personas.

En 2010, el escritor Steven Johnson informó a TED.com que había dirigido una investigación sobre dónde tiene lugar la innovación científica. Lo que descubrió fue que la mayor parte de la innovación tenía lugar en realidad a través de la interacción social en reuniones de laboratorio normales cara a cara. Allí se compartían las ideas, se retaban los datos y los conceptos surgían rápidamente entre los participantes. Las ideas se convertían de verdad en innovación cuando se combinaban con otras o se añadían a hechos existentes. Hacia el final del proceso científico, las conclusiones se escriben como una conversación unilateral con un colega imaginario, anticipando preguntas, retos y estimulando a debatir. Para apoyar un duro debate científico, las personas no solo necesitan acceso a hechos relevantes, sino a una plataforma de colaboración que permita la interacción virtual en tiempo real con toda la información pertinente.

## Hacer trampas está permitido

Esos laboratorios que todavía tienen por implementar un sistema informático no deberían preocuparse – llegar tarde en la adopción de la integración de datos de laboratorio tiene una gran ventaja. Todas las industrias de atención sanitaria, bancarias y del consumidor han adoptado estrategias electrónicas e informatizadas de integración de datos, lo que significa que muchas tecnologías aceptadas están ahora a disposición de los laboratorios. Este es un momento interesante ya que las buenas prácticas entre industrias pueden utilizarse para crear repositorios de gestión de conocimiento de ciclos completos y permitir la colaboración multifuncional entre silos de información interna ◀◀

Artículo publicado originalmente en *Scientific Computing World* y traducido y reproducido con el permiso de la editorial.