
Volume 16

Issue 1 *Steps Towards Digital Transformation in
the Pharmaceutical Manufacturing Landscape
Linking Data, Analytics, Knowledge and Risk*

Article 5

12-15-2021

“Data Architecture for Pharmaceutical Product Development at Alkermes”

Peter Couture

Alkermes, Ireland, Peter.Couture@alkermes.com

Damon Warnock

Alkermes, Ireland, damon.warnock@alkermes.com

Marcus O'Mahony

Pharmaceutical Manufacturing Technology Centre (PMTc), Bernal Institute, University of Limerick, Ireland,
marcus.omahony@ul.ie

Follow this and additional works at: <https://arrow.tudublin.ie/level3>



Part of the [Pharmacology Commons](#)

Recommended Citation

Couture, Peter; Warnock, Damon; and O'Mahony, Marcus (2021) “Data Architecture for Pharmaceutical Product Development at Alkermes,” *Level 3*: Vol. 16: Iss. 1, Article 5.

Available at: <https://arrow.tudublin.ie/level3/vol16/iss1/5>

This Deploying Analytics Solutions is brought to you for free and open access by the Current Publications at ARROW@TU Dublin. It has been accepted for inclusion in Level 3 by an authorized administrator of ARROW@TU Dublin. For more information, please contact arrow.admin@tudublin.ie, aisling.coyne@tudublin.ie, gerard.connolly@tudublin.ie.



This work is licensed under a [Creative Commons Attribution-NonCommercial-Share Alike 4.0 License](#)

Data Architecture for Pharmaceutical Product Development at Alkermes

Authors

Peter Couture,¹ Damon Warnock¹ and Marcus O'Mahony²

¹ Alkermes Pharma Ireland, Athlone, Ireland

² The Pharmaceutical Manufacturing Technology Centre, Bernal Institute, University of Limerick, Ireland

Abstract

This paper describes how Alkermes built a data architecture for efficient management of non-GXP and R&D data. Using a 'Create', 'Store', Describe' 'Access' stepwise model, the ingredients to making the journey a success are described.

What is the business context before this new use of data took place?

Alkermes, like most companies, seeks to better understand its products and processes. Increased knowledge leads to better process control and reduced development cycle time. As this knowledge is driven by data collection and analysis, a reliable framework for data collection and access is critical to organisational success.

Prior to building out a data architecture for efficient management and access to contextualised, non-GXP, R&D data, the site at Alkermes Pharma Ireland had to contend with working up valuable data that lived in isolated, disconnected sources. A data architecture is the template for individual data pipelines. Data pipelines will vary slightly based on use case but are designed to move data from source to end user.

Building a data architecture, as outlined in this case study, aligned with Alkermes' organisational effort for structuring and organising its data to enable wider access and to link all product development data (from discovery through to commercial manufacture) for full support of product life cycle.

Within the organisation, a separate data governance group is in place that works to harmonise naming standards or ontologies. Harmonising naming standards across an organisation increases the comparability of datasets for improved analysis. Harmonisation is an on-going effort that requires constant effort and debate within an organisation.

How did it come into existence, what drove it?

For years, a big issue for scientists and engineers was the time spent getting their hands on ‘good’ quality data, i.e. data that can be readily and easily visualised and analysed. Without readily accessible historical data, experiments were executed and analysed in isolation. Those isolated datasets made the question of ‘How does this compare to the last batch?’ very challenging to answer.

In the process development environment, data naming and documentation methods are much more variable and subject to inconsistencies compared to the commercial GXP environment. As a result, process development scientists and engineers can sometimes spend hours, or even days, looking for a result from a historical experiment. If data are organised as a searchable asset (contextualised), then the task of finding, visualising and analysing data can be shortened from days to minutes. Reducing the time spent searching and pre-processing data for the end-user was the primary driver of the effort to build a data architecture.

What is this effort or journey?

The Advanced Product Analytics & Capabilities (APAC) Group at Alkermes Pharma Ireland, in collaboration with several IT and scientific groups, are building a data architecture that connects data sources and contextualises those data sources to enable quick data visualisation and analysis.

The key components for building a data pipeline or architecture are to appropriately store, describe, and access data. To be successful, the end user requires access to both the raw data and to the cleaned data for analysis. Figure 1 below provides a schematic overview of the data architecture and the sections below elaborate on the components.

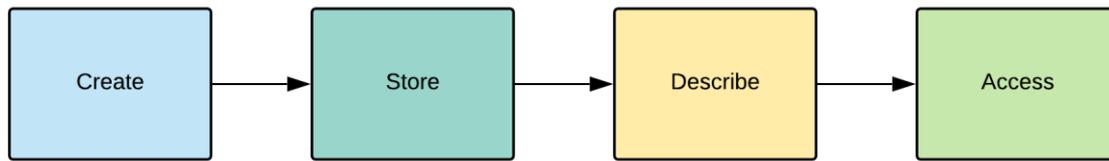


Figure 1: A basic data architecture

Store

The first step in building the data architecture is to pull data into storage.

Knowing how to get data into storage requires that the organisation performs a mapping exercise to understand the different types of data available. Our data can typically be grouped into three categories: files (txt, pdf), sensor (time series), and database (Oracle, SQL). Understanding which category a set of data belongs to will influence how information can be extracted from that data.

Describe

Once the file, report, or dataset is properly stored, the next step is to describe the data. Describing data in this context means adding descriptive information to the raw data, so that it can be quickly and easily found.

Typically, when a file is stored it is given an ID value that points to its storage location. If a file is to be found more easily, further information must be added such as, where the file came from, who made it, when it was created, its product affiliation, etc.. This extra information is critical to asking questions about the data and quickly finding the information needed.

Access

The most important component of the data architecture is to provide scientists with access to their data. The output from the 'store' and 'describe' efforts is a cleaned dataset that is ready for access.

With access to cleaned and structured data, scientists can easily compare historical results to new test data. Improvements to data access are where the data architecture really delivers visible value to the business.

Whilst each scientist, or group, might be interested in viewing data in different ways, they would all prefer to minimise the time spent searching for and formatting the data. A data architecture facilitates rapid access to data.

Building a data architecture requires IT expertise. However, the project should be sponsored and driven by the groups using the data. The design of any data architecture should be based directly on business need and use case. Failure to select tools based on use case will inevitably lead to failure of the overall architecture.

The skills required to implement a data architecture are many and varied. Our recommendation is to start small and to deliver value quickly for a single use case. Find out what you want to do, and let that dictate the skills and techniques needed for implementation. For example, if the goal is to pull data from an inventory management system, a data engineer with some experience dealing with relational databases would be helpful.

At Alkermes, the data architecture is a collaboration between IT and product development groups. From IT, collaborators include data engineers in IT Applications, DevOps engineers in IT Infrastructure and security experts in IT Security. From product development, contributors come from Product Analytics, Analytical Development, Formulation Development, Process Development, and Early Discovery. While no single group contributes to the entire architecture, each group offers skills and insights for individual data pipelines. The Product Analytics team driving the efforts reported here in this case study are Damon Warnock, Keith Ryan, Paul Bassett and Peter Couture.

What were the ingredients to making the journey a success?

One of the most important aspects of building a data architecture is to get engagement from the business. There should be a clear business driver and value proposition for building out a data architecture to facilitate the flow of data within an organisation.

At Alkermes, a business representative was required to map out how they use data and how it should be presented: the data architecture is based directly on that business need.

Define a 'minimum viable product' for each application; separate 'nice to have' features from 'must haves' and prioritise appropriately. Doing so will increase the speed at which you can deliver value to the organisation and improve communication to stakeholders.

Although not a necessity, involving external consultants can be helpful. The landscape of data products is vast, and at times the project here relied heavily on the expertise of consultants to put all the pieces together. The Alkermes team might have eventually arrived at the same point, but it would likely have taken much longer and required significantly greater expense.

From a project management perspective, the project should be driven and sponsored by a business representative. However, the project will require IT partnership. It is important that the IT group is willing and able to support the chosen data architecture platform and tools. Outsourcing services can be a palatable option if the IT department cannot realistically handle the data architecture needs. Embrace the cloud.

What impact was delivered / is still being delivered?

The goal at Alkermes was to allow scientists and engineers to find data faster and to gain new insights into our products. By investing time and resources into a data architecture, data that previously took days or weeks to collect can be queried and discussed in moments.

Mapping data use cases also allows scientists to spot inconsistencies in the way they operate. So, before an architecture is even built, value is delivered by refining the way the business operates.

Finally, it is worth mentioning that searching for data is frustrating; the ability to find relevant data at will has a noticeable, positive impact on job satisfaction and team productivity.