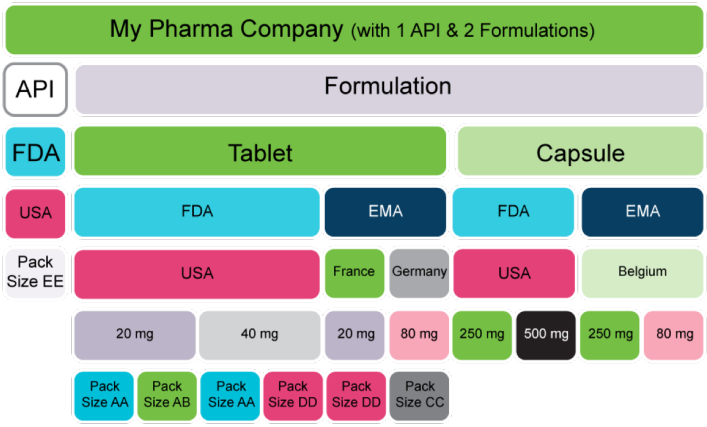
BioSpectrum
the business of Bi
24 years of leveraging the business of Bio
www.biospectrumindia.co

Harmonious approach of regulatory data migration

12 March 2020 | Views

A holistic solution platform that can manage the complexities of regulatory operations



In an evolving regulatory landscape where a ‘single source of truth’ is paramount, one needs to explore just how important data migration really is.

We can look at how a life sciences organization can strategically take their data from disparate legacy sources to a centralized product database. Investing in this ‘single source of truth’ can help avoid the all too common data integrity challenges within internal operations which, as we know, not only cost but can heavily hinder an organization.

The figure above provides a high-level data structure that is prevalent for a typical 1 API and 1 Product with two formulations. We would consider this data hierarchy as an inverted tree structure that addresses multiple occurrences relating to:

- Markets
- Countries
- Doses
- Strengths
- Pack sizes
- Indications

We already know that compliance challenges exist owing to multiple markets and formulations, and these challenges must be addressed in order to meet Health Authority (HA) requirements. Further complications arise with the ownership of data. In most companies, we see a shift from central to local/affiliates, and vice versa. Whilst, this is okay, it needs to be fully documented and traceable. If done in a siloed fashion, as we know it often is, this puts added pressure on regulatory operations resource. Adding to this, there is a requirement to track the licenses, agency correspondence, market expansion, renewals, changes, to name but a few, hence the data sets held by regulatory operations are both exhaustive and multifaceted. The complications and challenges are clear for all to see, the answer? A holistic solution platform that can manage the complexities of regulatory operations. Let’s take a look at that in a little more detail.

Regulatory Data availability situation

As we have started to consider, the regulatory domain related data is multifaceted. To that end, there are multiple distinct Stakeholders who create and/or consume data in their preferred form in order to perform their day-to-day business functions. For example, medicinal product data values are captured and recorded at different stages; starting at the clinical stage, through to post-marketing, and beyond; right up until the last batch is made available in any specific market. At each stage, the Product data attributes are different with minimal overlap, with the data continuing to grow.

Additionally, an Enterprise can have business functions dispersed globally, and using multiple systems to manage their day-to-day operations. This can lead to further complex scenarios when it comes to looking at holistic regulatory data. Such scenarios can include:

- Data is globally dispersed (**data pockets**) with no single repository at the enterprise level
- Some of the final data is available with Affiliates
- If changes and/or variations are not properly managed, then local submissions can deviate from the Core Data Sheets in areas where that deviation is not wanted
- Data sits in different sources; paper, Microsoft Excel Open XML Spreadsheet (XLSX), databases (**multiple envelopes**)
- Capacity required to move data to a single repository (e.g. centralized Product database) not available

Data readiness for regulatory submissions

Just because the data is available ‘somewhere in the organization’, does not entail that the data is submission ready. Any data that is identified for regulatory submission must go through a proper data lifecycle, to meet the requirements outlined below:

- A single source of truth for data originating from different/multiple sources must be identified
- The data identified must be fully traceable, with a timestamp for future reference Controlled revisions must be adhered to when making any changes to the data entry value

These requirements are critical to safeguarding data integrity. Hence, data readiness for regulatory submissions must be handled with due attention and diligence to avoid data mix-up or version conflicts. Once the data is available in a harmonized form, the next step is to make the data ready for regulatory submission(s). This, in practice, means managing potential non-compliance situations which could have a significant impact on both brand and the economy. In our experience, it is rare that a situation(s) occur that have associated patient safety risks.

In order to standardize data communication with HAs, there is a preference to submit data in an agency-prescribed format (e.g. eCTD, NeES, XEVMPD, SPL and, in future, IDMP etc.) as much as possible for data consistency and verification. HAs have their own automated parsers to decipher the formatted data submission and quickly let the pharma company know whether the data is readable and accepted in principle for processing.

The Challenge

We appreciate that, at first glance, the data journey and its associated challenges can appear somewhat daunting. So often we hear:

- How do I start my data migration? There are duplicate copies of data sources, with different owners, across different locations. We need to shift from multiple data pockets to a harmonized data plane but, specifically, just how do I manage potential non-compliance situations?
- Not all data are in non-standard format, which greatly restricts automated data imports with manual quality checking being deployed post migration
- The indirect cost in duplicate data maintenance such as storage, manual time and effort to find the right source and right version, always ensuring data integrity is not tampered
- How can a company ensure a compliant, effective migration to a new database?

So, just how do you get ready for data migration?

If a Pharma Company is looking to start the setting up of a centralized Product Database, our advice would be to aim for global compliance and long-term reduced data maintenance costs, rather than looking for the shorter-term benefits, which may initially look attractive, but don't solve the real issues. For instance, don't settle with localized solutions for global problems. This would quickly require periodic repair and upgrade costs.

Here are some key considerations to achieve a harmonized product data model:

- Align and form a **Data-Cell Team** to lead the data change management program with a clear goal of achieving centralized and harmonized Product Database. It may seem obvious, but the challenges are in the detailed preparation work towards building just that.

- The Data-cell team to comprise of well-informed team members who have very good understanding of pharma operations and the associated product data in its expected form. Typically, the regulatory associates, publishers, and strategists become very critical to understanding, assessing and defining the Go-To Data Model components. This team should include local as well as global players

- **Centralized Product Data Model.** The Data-cell is responsible for baselining (versioning) the Product Data Model as per organizational policy and procedures

- **Avoid and eliminate Duplicate data collection:** The Data-cell team to ensure that data duplication is eliminated. The team need to identify the top three sources as Gold, Silver and Bronze to ensure that the duplication of data sources are graded appropriately and by default, Gold Source is considered as the 'Single Source of Truth'. Whilst Silver and Bronze are referred to on a need basis for any reconciliation and mainly for review purposes.

- **Data mismatch:** Quite often we run into a situation where the data that has been submitted locally is different from what it should be, based on Core Data Sheet, and the changes published globally. These potentially non-compliant situations need to be investigated, and appropriate resolutions need to be managed.

- **Compliance-friendly data formats:** The data stored, and its associated formats should be as close to HA compliant formats to avoid any gaps in compliance related data format transformation work. Any data that is a candidate for submission should be managed by the team who understand the HA based submission needs. This would ensure that the requirement for data migration from one form-to-another is met whilst keeping compliance in the mind.

- **Controlled Data Life Cycle Management:** Any change or revision to published data must go through a proper change control process before effecting the change. This ensures that the changes are both well-planned and controlled.

- Classify the overall data hierarchy into respective categories:
    - Industry Standard Data (Global Term Store)
    - HA specific data (Agency Term Store)
    - Local Data (Local Term Store)
    - Operational Data (data generated and maintained as part of internal operations)
    - Feedback Data (In-bound data including: Safety changes, regulatory correspondences, Product Complaints, and Audit Observations)

- Define organization standards for each data point on the following:
    - Data Category
    - Data Type
    - Length
    - Mandatory (Default or conditional)
    - Native Formats (predefined formats, pre-fixes, suffixes)
    - Data Privacy
    - Data Source
    - Submission Formats (if any)
    - Data Version

- Baselining and Publishing Data:
- For a specific data point, identify, enumerate and mark the multiple data sources such as Gold Source, Silver and Bronze so that the preference is always given to Gold Source of data and only on defined conditions, the other sources referred to.
- The data source can be manual, internal, or external system bound where in periodic data synchronization must be in place based on the frequency of data changes expected.
- Any data that is meant for consumption needs to be in published state. This is a well proven concept in terms of Standard Operating Procedure (SOP) management of Quality functions where in only published and effective SOPs are considered for operational consumption and, at any point in time, there will be only one published version.

- The right Technology Option
    - The potential technology solution for data migration is to look at a holistic approach towards data harmonization at Enterprise Level.

- The proposed solution to have the key features to facilitate easy data migration:
  - A Built-in option of regulatory base product data model with potential extensions
  - Capability to process multiple data sources with open architecture-based integration capability
  - Ability to transform data from one or multiple sources to prescribed format
  - Ability to map target data models against both internal and HA specific standards
  - Ability to process data migration in batches (both time-based and event-based processing)
  - Ability to track data changes, its associated versions and publish effective and baseline data only for consumption
  - Data Migration process should have the below key aspects built in:
    - Ability to pre-process data before migration
    - Ability to post-process data after migration with dependency checks
    - Ability to handle exceptions during migration

This article has, hopefully, addressed the key drivers to be considered by a Pharma Company when taking steps towards meaningful and successful data migration strategies with minimal surprises once the data is migrated. There is no single-stop solution for data migration. One needs to consider multiple factors, perspectives, and the right partner to make the data migration journey a smooth ride. Where are you in your data migration process?

Sathish Sundaramurthy, Head of Product Development, Safety and Regulatory Technology, Navitas Life Sciences (a TAKE Solutions Enterprise)