IPELTU Use Case for Life Sciences

Life Sciences companies are generating huge volume of datasets by collecting raw data, documents, images, multimedia, files, records, dossiers from internal R&D laboratories and departments, Manufacturing, Regulatory Affairs, Legal and external partners or investigators (CROs, hospitals, Universities, Research Centers, …). During the product lifecycle, the involved organizations are identifying the relevant processes and events for the data patrimony curation and preservation in the long-term.

Pharmaceutical, Biotech and Medical Devices companies have common denominators for managing the data patrimony as a company digital resource. A key focus is on meeting regulations and GxP compliance, protecting reputation, reducing clinical trials archiving costs, keeping records safe and secure, ensuring confidentiality, distributing controlled information packages to the designated communities. New regulations and compliance constraints require a continuous control and monitoring between business requirements, compliance requirements and applicable metrics.

Biotech and Medical Devices companies operate in the market with compliance constraints, obligations and risks as for Pharma industry, but the small and medium sizes of their organizations determine a big effort and difficulties in managing data patrimony governance in the long-term in a cost-effective manner with a service approach. Data preservation services should be based on a sustainable information model with secure and controlled access of sensitive information.

In Biotech companies, researchers use and produce huge amount of data deriving from the different stages of biomedical science research lifecycle and from a diversity of sources and scientific communities. The usage of interdisciplinary bioinformatic methods and technologies determines an increasing umbrella of new data formats subjected to obsolescence in the medium and long-time. The biotech researcher performs the pipeline studies, starting from biological data sources, measuring biomarkers and thereby producing big volume of data. The researcher compares and analyses complex data sequences using laboratories dedicated tools and disseminates the results (publications and raw data) in secure collaboration network where the researcher’s communities can access and contribute to the development of the research results.

Medical devices manufacturers or distributors have a to meet compliance to the Medical Devices Regulation (MDR), ensuring safety and quality of the devices for patients and tracking the data objects since the project initiation phase, until the execution of devices testing and the validation of results to ensure regulations compliance.

In Pharmaceutical industry, when a traditional R&D project for a new drug goes in Phase III, study protocols and clinical trials are generated and collected according to the regulations by the involved team composed by internal organization and external investigators. For example, during the Trial Master File life several interactions between the life sciences sponsor and the CROs occur and an exchange of digital information starts using e-clinical trials platforms services. The clinical trials data collected from the different silos are relevant for creation of submission dossiers, for the market authorization of the drug, for the pharmacovigilance and the archiving costs are relevant. Data integrity and 9 ALCOA+ principles must be ensured in a time frame of over 25 years.

For GxP compliance, Intellectual Property and legal reasons, the information packages must be preserved and keep accessible in the long-term in a trusted and secure digital archive. The data management plan and the preservation plan support the whole data curation and long-term preservation and identify and certify data objects, additional information, relevant events, applicable procedures (for transfer, appraisal, ingestion, archival storage, access) and services that the archive must ensure. The additional information sent to the digital archive includes Representation Information such as the data format, semantics and processing software, which have been created by various internal team and external teams according to the submission agreement and related contract clauses. When the contracts with CROs becomes expired, some relevant information are under risk and must be provided to the Pharmaceutical company digital archive instead to be lost.

For patient centric R&D lifecycle, the data curation and preservation are mandatory not only to ensure compliance but to facilitate the researchers developing new drug variations depending on the live data coming directly from the patients and the associations of the patients, and on a huge volume of historical data preserved in the digital archive.

To improve the experience of the designated communities, during the data exploitation the digital archive adds descriptive information and representation information that facilitate the understanding of data over time.