

**Problem:** We have a big metadata problem because traditional metadata management systems do not effectively manage the many dimensions of clinical research metadata. While not big in the big data sense, the 9 dimensions below (9 Vs) create unique metadata management challenges.

**1. Velocity:** The regular publication of new standards and controlled terminology versions combined with organizational standards ensure that new metadata arrive at a relatively high rate.

**2. Volume:** Clinical research uses many standards, data models, terminologies, relationships, and transformations/derivations to represent the end-to-end data pipeline.

**3. Value:** The data quality, traceability, and standards conformance needed for regulatory submissions and efficient data processing increase the value of the metadata. High-quality data is the lifeblood of research.

**4. Variety:** Clinical research and healthcare use a broad array of standards, data models, and terminologies that represent data from different sources and at different stages of the end-to-end data pipeline.

**5. Variability:** High variability exists in cases where, for example, we have many versions of an AE CRF to meet different study needs. Similarly, we may implement the same SDTM datasets differently for different TAs or studies.

**6. Views:** Clinical research metadata management can include views such as process views, data views, and system views. Each view represents an important way to understand and work with metadata, depending on your role.

**7. Versions:** Version management means that there exist many versions of each standard that must work with many versions of standards up and downstream from it.

**8. Veracity:** Conformance rules are captured as metadata and used to assess the data quality and standards conformance. Semantics help ensure we identify and define metadata and data consistently across standards.

**9. Visualization:** Given the variety, variability, volume, and relatedness of the metadata, visualization becomes an essential tool for understanding what metadata exists and how it is applied through the end-to-end pipeline.

**Solution:** Focus on a metadata repository architecture, understanding that no single tool addresses all needs and success demands integrating numerous metadata-producing and consuming systems. Work in an agile manner as your solution will evolve frequently.