D1.2: Data Management Plan

In the context of managing a research project like TRIAD, effective data management is crucial for ensuring transparency, reproducibility, and long-term usability of research outputs. This includes documenting experiments and depositing data into recognized repositories that adhere to FAIR (Findable, Accessible, Interoperable, Reusable) principles.

Journals for Electronic Documentation

Electronic documentation of experiments throughout the project lifecycle is essential for maintaining a clear record of methodologies, results, and interpretations. Utilizing electronic lab notebooks (ELNs) or specific journal platforms facilitates real-time data recording and collaboration among team members. This approach not only enhances the efficiency of data capture but also supports rigorous data management practices, ensuring compliance with funding agency requirements and journal publication standards.

1. **Enhanced Collaboration:** ELNs allow multiple researchers to access and contribute to the same experiment records simultaneously, fostering a collaborative environment.
2. **Regulatory Compliance:** Electronic documentation helps in adhering to regulatory standards and facilitates audits by providing a well-maintained and searchable record of research activities.
3. **Security and Backup:** ELNs provide secure storage with regular backups, reducing the risk of data loss compared to traditional paper-based methods.

Data Deposition in Specialized Repositories

For diverse data types such as Next Generation Sequencing (NGS), mass spectrometry, metabolomics, CRISPR/Cas9 screen data, and DNA vectors, TRIAD will utilize specialized data repositories:

1. ArrayExpress, ProteomeXchange, MetabolomeXchange: These repositories are dedicated to high-throughput genomics, proteomics, and metabolomics data. By depositing data here, TRIAD ensures that datasets are curated, standardized, and accessible to the broader scientific community.
   a. ArrayExpress: A repository for functional genomics data, particularly high-throughput sequencing data.
   b. ProteomeXchange: A consortium of proteomics data repositories, ensuring proteomics data is shared and accessible.
   c. MetabolomeXchange: A network of metabolomics databases, providing a platform for metabolomics data sharing.
2. BioGRID, ORCS, Addgene: Repositories for genetic and molecular biology data, including CRISPR/Cas9 screen data and DNA vectors. These platforms facilitate sharing of experimental tools and results, promoting collaboration and reproducibility.
   a. BioGRID: A database of protein, genetic, and chemical interactions.
   b. ORCS (Open Repository of CRISPR Screens): A dedicated repository for CRISPR screen data.
   c. Addgene: A nonprofit plasmid repository that facilitates the sharing of plasmids within the scientific community.
3. ATCC, DSMZ: Repositories for biological resources such as cell lines. Depositing cell lines ensures their authenticity, quality, and availability for replication of experiments by other researchers.
   a. ATCC (American Type Culture Collection): Provides authenticated biological materials for research and development.
   b. DSMZ (Deutsche Sammlung von Mikroorganismen und Zellkulturen): An internationally recognized resource for microbial and cell culture collections.

Adaptability of Data Management Plan (DMP)

TRIAD's Data Management Plan (DMP) is designed to be dynamic and responsive to evolving

circumstances:

1. Continuous Adaptation: The DMP will evolve in response to new data generated throughout the project, changes in consortium policies, shifts in team composition, and external factors such as emerging technologies or regulatory requirements.
2. FAIR Principles Implementation: The project adheres to FAIR principles to enhance data management, sharing, and usability:
    a. Findable: Utilizing clear and unique identifiers (e.g., DOIs) and metadata to facilitate discoverability through data repositories.
    b. Accessible: Ensuring long-term access to data by using persistent identifiers and complying with repository access policies.
    c. Interoperable: Adopting standardized data formats and protocols to enable integration and interoperability across different platforms and tools.
    d. Reusable: Providing comprehensive metadata that includes data provenance, methodology details, and licensing information to clarify conditions for data reuse.

Evolving Management Strategy

The adaptability of TRIAD's management strategy ensures alignment with the latest advancements and best practices in research data management:

1. Response to New Data: Incorporating new findings and insights into data management practices ensures that TRIAD remains at the forefront of research excellence.
2. Policy and Team Dynamics: Adapting to changes in consortium policies and team composition fosters collaborative synergy and supports effective knowledge transfer.
3. External Influences: Addressing external factors such as regulatory changes or technological innovations ensures compliance and relevance in the rapidly evolving research landscape.

In conclusion, by implementing robust data management practices and leveraging specialized repositories aligned with FAIR principles, TRIAD enhances the integrity, accessibility, and impact of its research outputs. The project's commitment to continuous adaptation ensures that its management strategies remain adaptive, responsive, and aligned with the evolving demands of scientific inquiry and collaborative research endeavors.