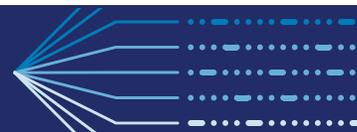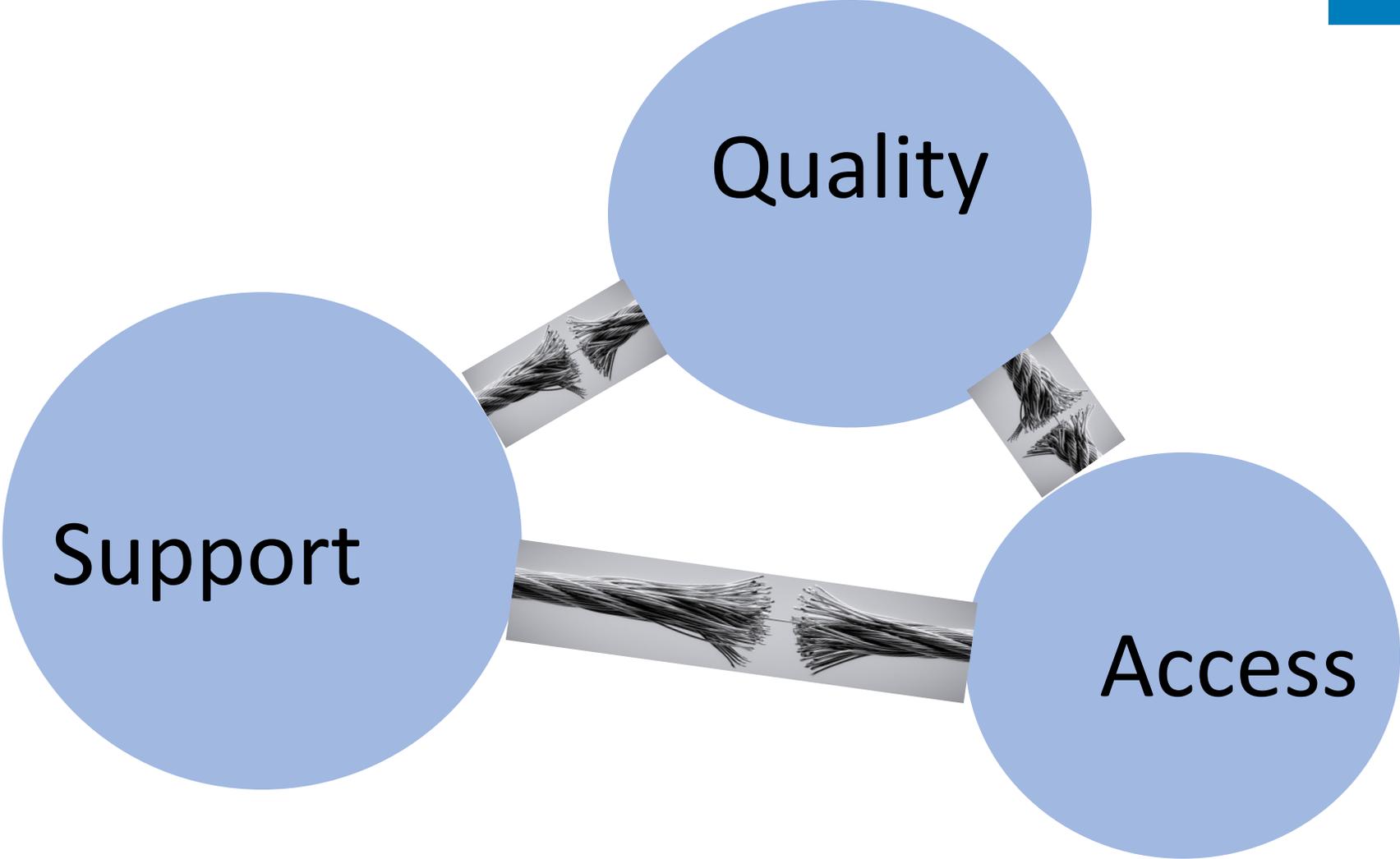# Data Central: High quality study data, in the right place at the right time

Matthew Wiley, Project Manager, Office of Computational Science, Office of Translational Sciences, CDER
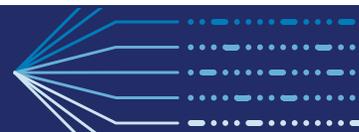
December 4, 2023

The Challenge

Study Data → FDA

Quality

Support

Access

What is Data Central?

- Data Central is a data prep engine (backed by expert data managers) that automatically scans the Electronic Document Room (EDR) and other sources (e.g. DARRTS) for study data and submission artifacts on a nightly basis and:

  - Connects study data
    - Quality Assessments
    - Analytical tools/systems
    - Clinical and Nonclinical Reviewer Services

  - Provides automatic, seamless connection with downstream tools and services
    - Tool validation checks
    - Metadata – filter out what you need
    - Application Programming Interface (API) connections [system to system]
    - Automatic format conversions (from .xpt into .sas7bdat, .csv, .rds)
    - Traceability

  - **Key Benefit**: Directly supports scientific review by providing regulatory reviewers with access to quality study data that is seamlessly connected to analytical tools and systems.

Data Central

Data Central Features & Benefits

Identification
Automatic, nightly scan of EDR to identify new, incoming submissions

Conversions
Enables automatic format conversions from .xpt into .sas7bdat, .csv, and .rds

Metadata Reporting
Features metadata-powered reporting and filtering on submissions and study data
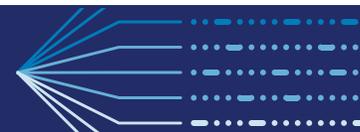
Validations
Runs folder, file-level, and tool-specific validation checks on submissions

Automated Connections
Provides automated, seamless connection to downstream tools, services, and Quality Assessments through API

Data Central At-A-Glance

FDA

Application Types
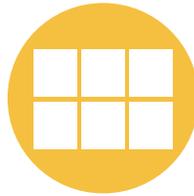INDs, NDAs, ANDAs, BLAs and EUAs
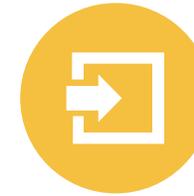
Modules
M1, M2, M4, and M5

Study Data Standards
SDTM, ADaM, SEND, ISS

Documents Types
SDRGs, CSRs, NRs, ACRFs, Cover Letters, Labeling, etc.

Domain Types
All Domains/XPTs

Converted File Formats
.sas7bdat, .rds, and .csv

Notifications
Automated email notifications as study data is found in EDR

Runtimes
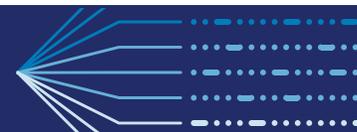Daily at 12 AM, 1 AM, 2 AM, 3 AM, and 3 PM

0, 1, 2, 3 …
Submission Sequences
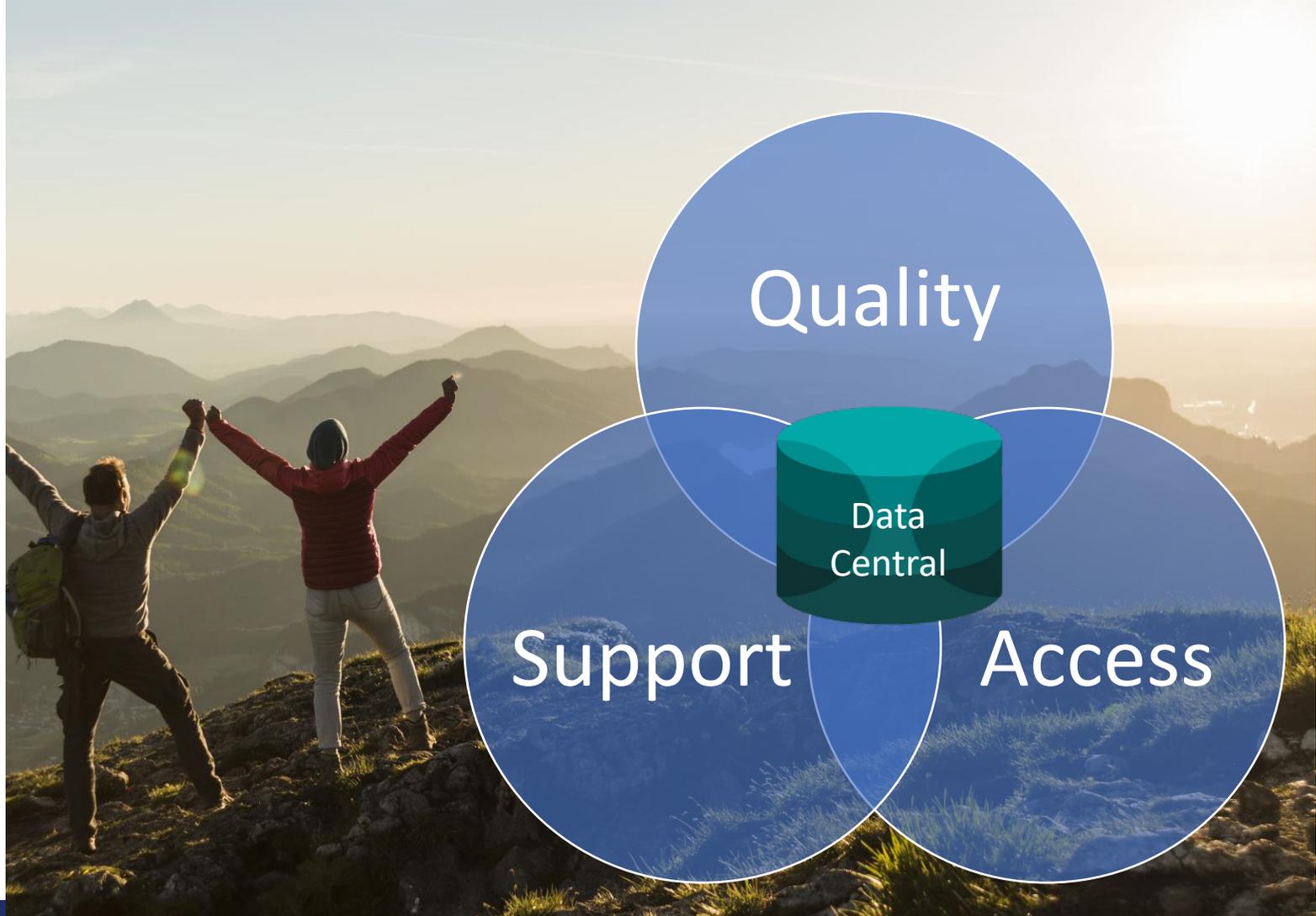All Sequences (since Go-Live in Sep 2020)

Folder Structure
Application ID_Sequence_Module_Study ID_Data Standard_File Type
(e.g., Dataset, Labeling, Cover Letter, CSR, etc.).

Folder Structure Examples
BLA123456_0004_m5_864248_SDTM_DA
IND102468_0531_m4_ghi5432123_NR

The Solution: Data Central

Study Data

Thank you!

Questions?