Applying the SDTM to the Janus Data Warehouse

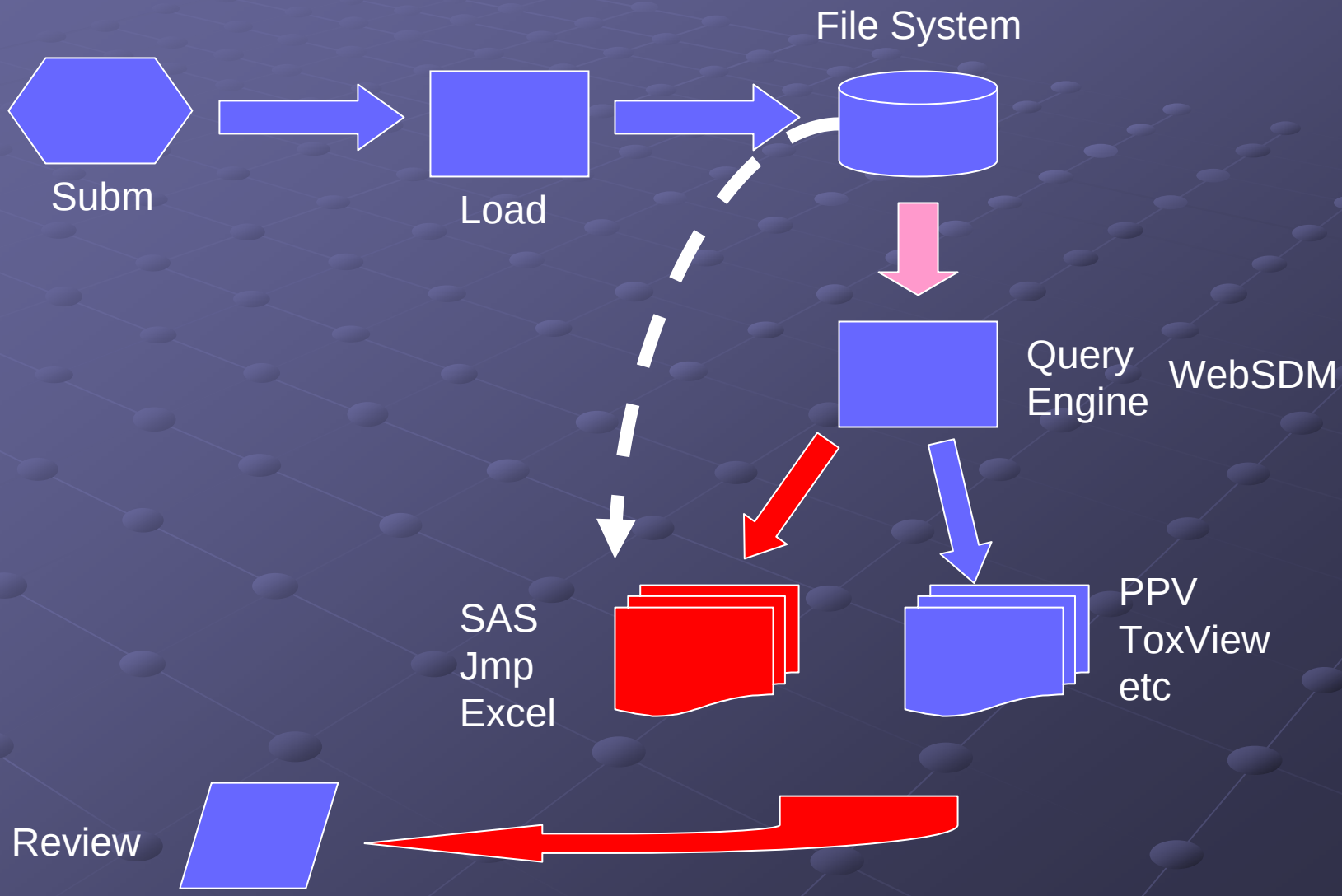
Norman Stockbridge

FDA

Data standards are good

- Reviewers can be trained
- Reviewers familiar with format
- Developers can develop consistent tools
- Reviewers can develop consistent procedures

Initial


File System
Subm
Load
Query
Engine
WebSDM
SAS
Jmp
Excel
PPV
ToxView
etc
Review

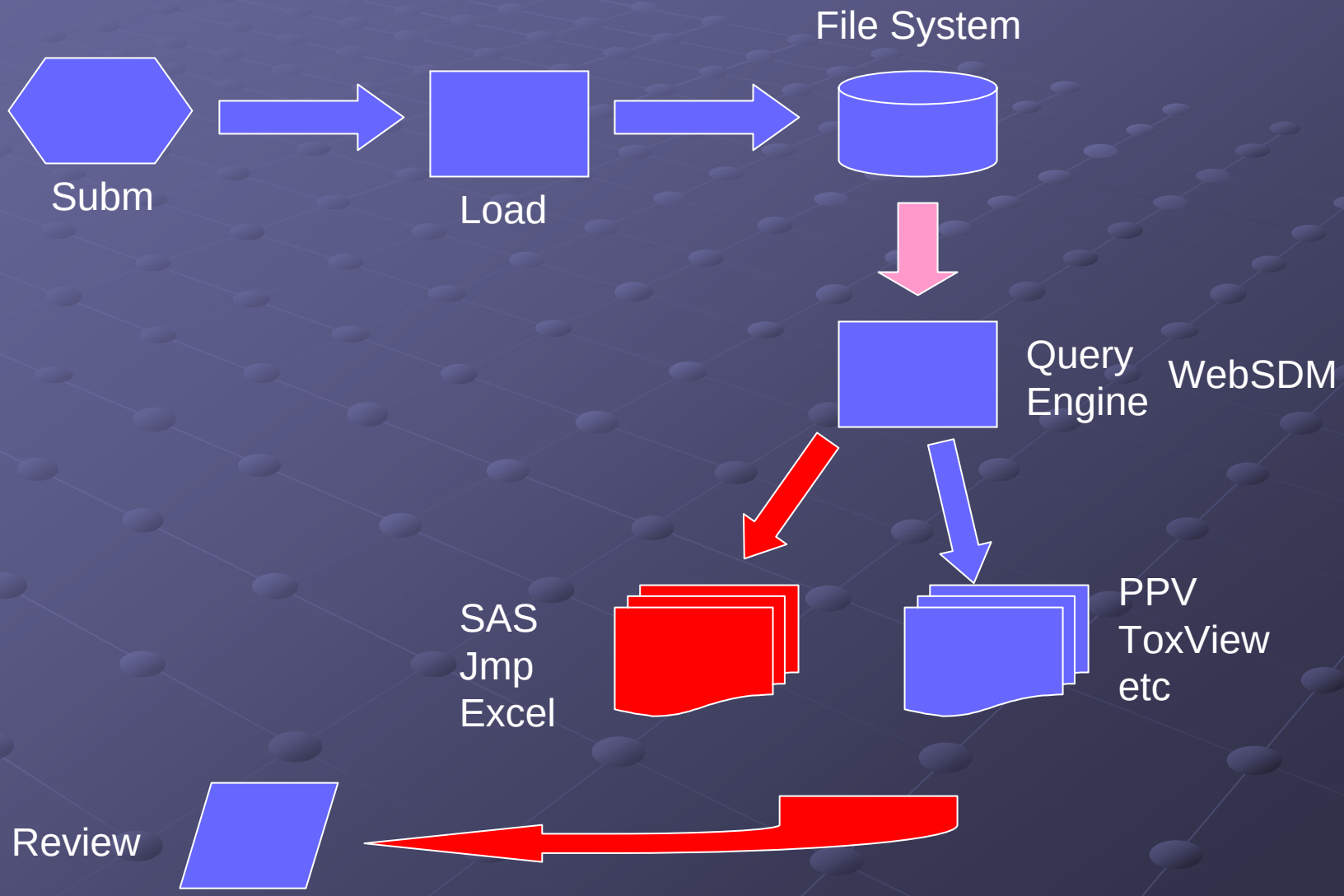
But SDTM is not optimal for reviewers

- Accommodate new data without structural changes
- Facilitate the assembly of submissions
- Ensure submission integrity
- And help reviewers, too.

Raw datasets are awkward

- One observation per row
- High degree of normalization

Initial


File System
Subm
Load
Query
Engine
WebSDM
SAS
Jmp
Excel
PPV
ToxView
etc
Review

Some difficult questions

- What was the relationship between plasma levels of drug and some pharmacodynamic parameter measured in 3 studies?
- How many women over age 55 were in the whole development program?
- How many of those women received a dose over 20 mg or were treated for more than 7 days?

Some difficult questions

- Is it safe to use placebo in short-term hypertension studies?
  - 98 NDAs
  - About 10 person-years
- Is there a placebo effect with ambulatory blood pressure monitoring?
  - Meta-analysis of patient-level data
  - 35 studies
  - About 10 person-years

Some questions remain difficult

- In the Vioxx development program as a whole, how many cardiovascular events were there?
  - What was the timing of those events?
  - What were the risk factors?
- How did that compare with other members of the class?

Audit

- Reviewing should preserve a trail of what was done.
- Secondary review should be able retrace these steps.

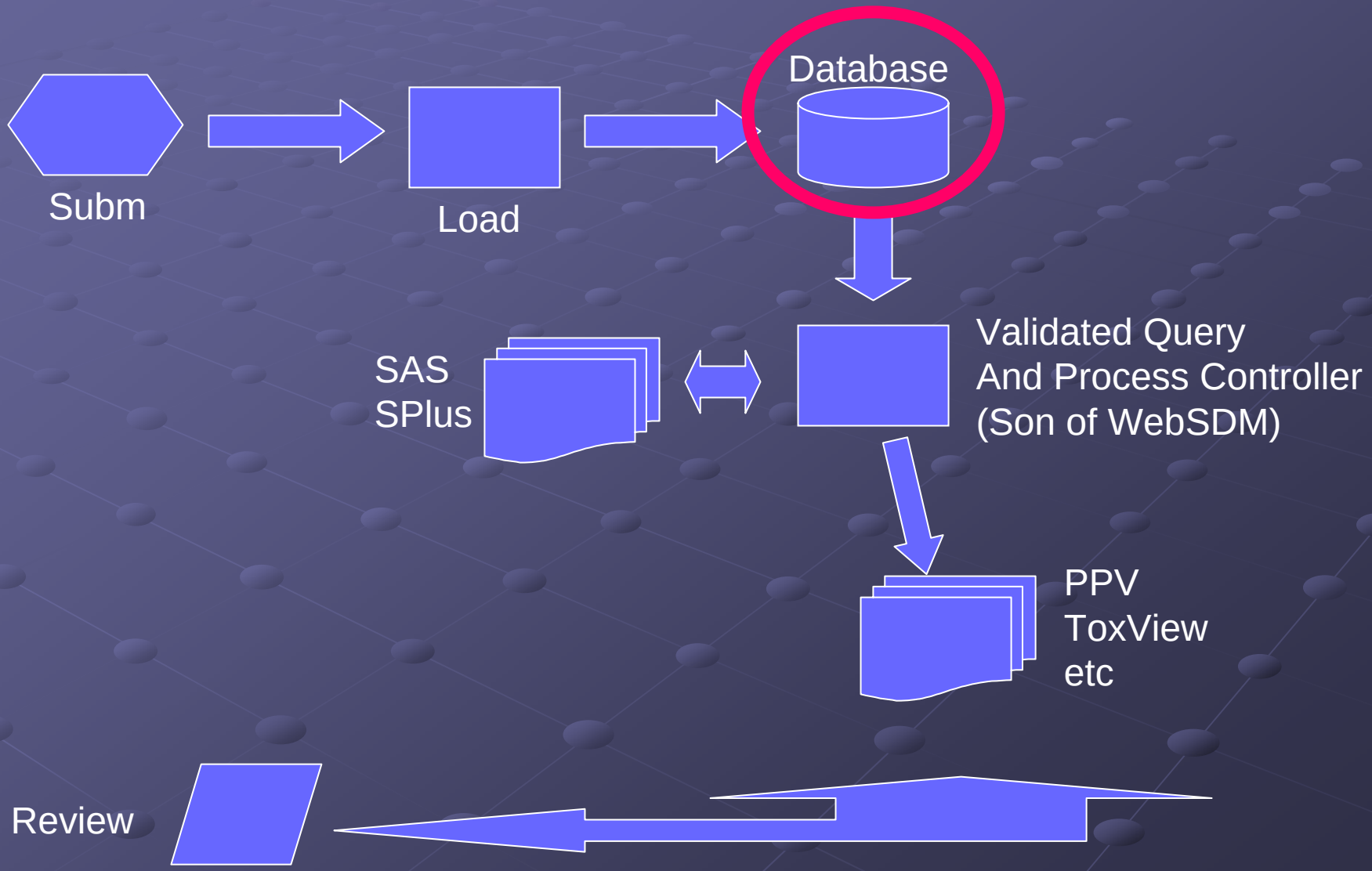
Communications

- When a reviewer has done something interesting, he or she ought to be able to communicate that to all parties unambiguously.

Need

- View of the data through reviewer-centric tools
- Cross-study analyses
- Cross-application analyses
- Audit
- Communications

Final


Subm
Load
Database
SAS
SPlus
Validated Query
And Process Controller
(Son of WebSDM)
PPV
ToxView
etc
Review

Clinical Trials Data Warehouse

Past studies

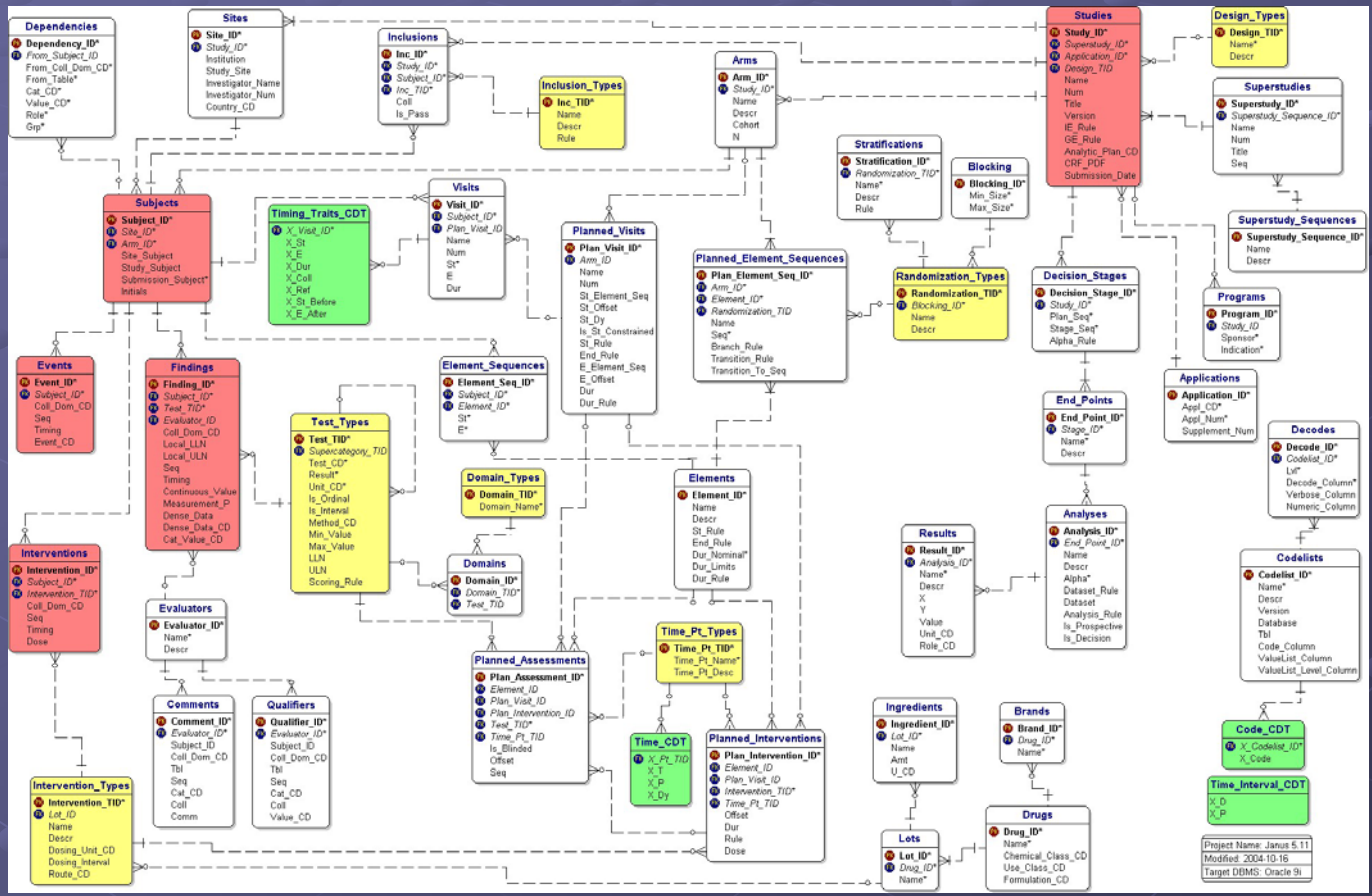
Janus

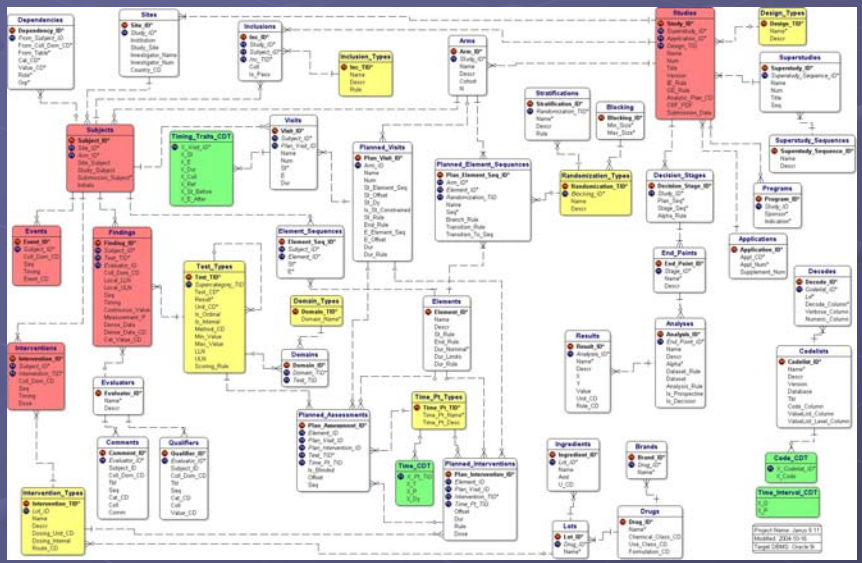
Current study

Contributors

- Jay Levine
- Chan Russell
- David Fram
  - all Lincoln Technology
- Randy Levin (FDA)
- Joyce Hernandez (IBM)
- CDISC SDS team

JANUS


Dependencies
Dependency_ID*
From_Subject_ID
From_Coll_Dom_CD*
From_Table*
Cat_CD*
Value_CD*
Role*
Grp*
Sites
Site_ID*
Study_ID*
Institution
Study_Site
Investigator_Name
Investigator_Num
Country_CD
Inclusions
Inc_ID*
Study_ID*
Subject_ID*
Inc_TID*
Coll
Is_Pass
Inclusion_Types
Inc_TID*
Name
Descr
Rule
Arms
Arm_ID*
Study_ID*
Name
Descr
Cohort
N
Studies
Study_ID*
Superstudy_ID*
Application_ID*
Design_TID
Name
Num
Title
Version
IE_Rule
GE_Rule
Analytic_Plan_CD
CRF_PDF
Submission_Date
Design_Types
Design_TID*
Name*
Descr
Superstudies
Superstudy_ID*
Superstudy_Sequence_ID*
Name
Num
Title
Seq
Stratifications
Stratification_ID*
Randomization_TID*
Name*
Descr
Rule
Blocking
Blocking_ID*
Min_Size*
Max_Size*
Subjects
Subject_ID*
Site_ID*
Arm_ID*
Site_Subject
Study_Subject
Submission_Subject*
Initials
Timing_Traits_CDT
X_Visit_ID*
X_St
X_E
X_Dur
X_Coll
X_Ref
X_St_Before
X_E_After
Visits
Visit_ID*
Subject_ID*
Plan_Visit_ID
Name
Num
St*
E
Dur
Planned_Visits
Plan_Visit_ID*
Arm_ID
Name
Num
St_Element_Seq
St_Offset
St_Dy
Is_St_Constrained
St_Rule
End_Rule
E_Element_Seq
E_Offset
Dur
Dur_Rule
Planned_Element_Sequences
Plan_Element_Seq_ID*
Arm_ID*
Element_ID*
Randomization_TID
Name
Seq*
Branch_Rule
Transition_Rule
Transition_To_Seq
Randomization_Types
Randomization_TID*
Blocking_ID*
Name
Descr
Decision_Stages
Decision_Stage_ID*
Study_ID*
Plan_Seq*
Stage_Seq*
Alpha_Rule
Superstudy_Sequences
Superstudy_Sequence_ID*
Name
Descr
Programs
Program_ID*
Study_ID
Sponsor*
Indication*
Applications
Application_ID*
Appl_CD*
Appl_Num*
Supplement_Num
Events
Event_ID*
Subject_ID*
Coll_Dom_CD
Seq
Timing
Event_CD
Findings
Finding_ID*
Subject_ID*
Test_TID*
Evaluator_ID
Coll_Dom_CD
Local_LLN
Local_ULN
Seq
Timing
Continuous_Value
Measurement_P
Dense_Data
Dense_Data_CD
Cat_Value_CD
Element_Sequences
Element_Seq_ID*
Subject_ID*
Element_ID*
St*
E*
Test_Types
Test_TID*
Supercategory_TID
Test_CD*
Result*
Unit_CD*
Is_Ordinal
Is_Interval
Method_CD
Min_Value
Max_Value
LLN
ULN
Scoring_Rule
Domain_Types
Domain_TID*
Domain_Name*
Domains
Domain_ID*
Domain_TID*
Test_TID
Elements
Element_ID*
Name
Descr
St_Rule
End_Rule
Dur_Nominal*
Dur_Limits
Dur_Rule
End_Points
End_Point_ID*
Stage_ID*
Name*
Descr
Analyses
Analysis_ID*
End_Point_ID*
Name
Descr
Alpha*
Dataset_Rule
Dataset
Analysis_Rule
Is_Prospective
Is_Decision
Results
Result_ID*
Analysis_ID*
Name*
Descr
X
Y
Value
Unit_CD
Role_CD
Decodes
Decode_ID*
Codelist_ID*
Lvl*
Decode_Column*
Verbose_Column
Numeric_Column
Codelists
Codelist_ID*
Name*
Descr
Version
Database
Tbl
Code_Column
ValueList_Column
ValueList_Level_Column
Interventions
Intervention_ID*
Subject_ID*
Intervention_TID*
Coll_Dom_CD
Seq
Timing
Dose
Evaluators
Evaluator_ID*
Name*
Descr
Comments
Comment_ID*
Evaluator_ID*
Subject_ID
Coll_Dom_CD
Tbl
Seq
Cat_CD
Coll
Comm
Qualifiers
Qualifier_ID*
Evaluator_ID*
Subject_ID
Coll_Dom_CD
Tbl
Seq
Cat_CD
Coll
Value_CD
Planned_Assessments
Plan_Assessment_ID*
Element_ID
Plan_Visit_ID
Plan_Intervention_ID
Test_TID*
Time_Pt_TID
Is_Blinded
Offset
Seq
Time_Pt_Types
Time_Pt_TID*
Time_Pt_Name*
Time_Pt_Desc
Time_CDT
X_Pt_TID
X_T
X_P
X_Dy
Planned_Interventions
Plan_Intervention_ID*
Element_ID
Plan_Visit_ID
Intervention_TID*
Time_Pt_TID
Offset
Dur
Rule
Dose
Ingredients
Ingredient_ID*
Lot_ID*
Name
Amt
U_CD
Brands
Brand_ID*
Drug_ID*
Name*
Drugs
Drug_ID*
Name*
Chemical_Class_CD
Use_Class_CD
Formulation_CD
Lots
Lot_ID*
Drug_ID*
Name*
Code_CDT
X_Codelist_ID*
X_Code
Time_Interval_CDT
X_D
X_P
Intervention_Types
Intervention_TID*
Lot_ID
Name
Descr
Dosing_Unit_CD
Dosing_Interval
Route_CD
Project Name: Janus 5.11
Modified: 2004-10-16
Target DBMS: Oracle 9i



- Observations
- Protocol
- Studies in context
- Analytic plan & results
- Coded list management

JANUS is simpler than SDTM

- Fewer tables than datasets
  - One table per observation class
    - Findings
    - Events
    - Interventions
- Fewer variables
  - No special findings (demographics)
  - No special qualifiers (AE severity)
  - One mechanism to link observations

Status

- Logical design evolving with SDTM
- Physical design through CRADA with IBM
- High priority with IMSC
- Prototype populated in 2005
- Alternatives for implementation in 2006
  - Partner with NCI
  - Contract
  - Consortium