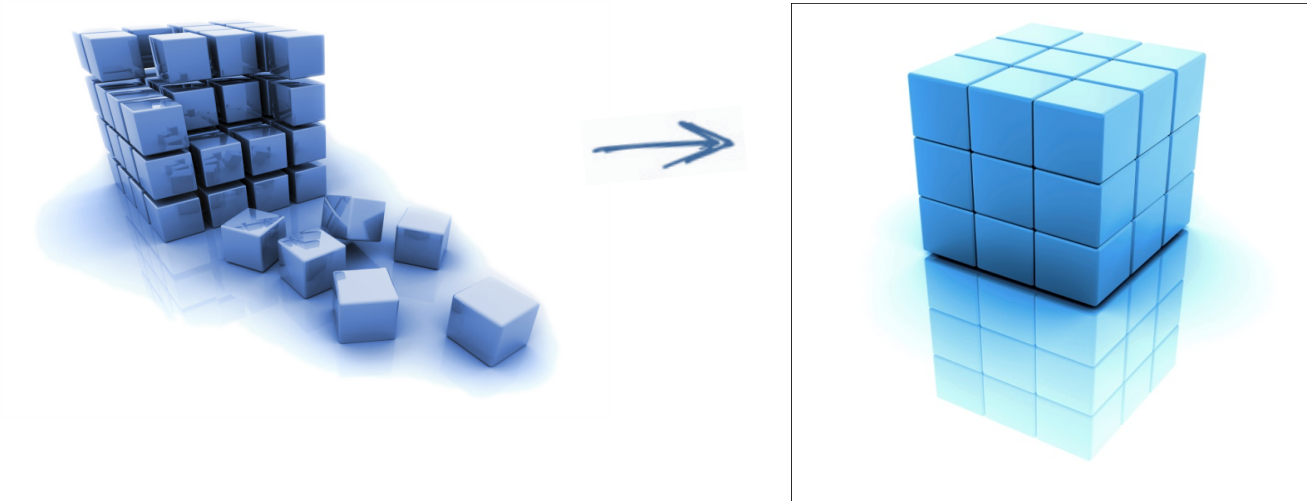
Standards and Ontology-driven Workflow Automation of Clinical Studies



www.pointcross.com

Characteristics of the Clinical Trial Workflow

- Multiple laboratories, investigation sites and contract research organizations are involved --- **Workflows are inter-organizational**

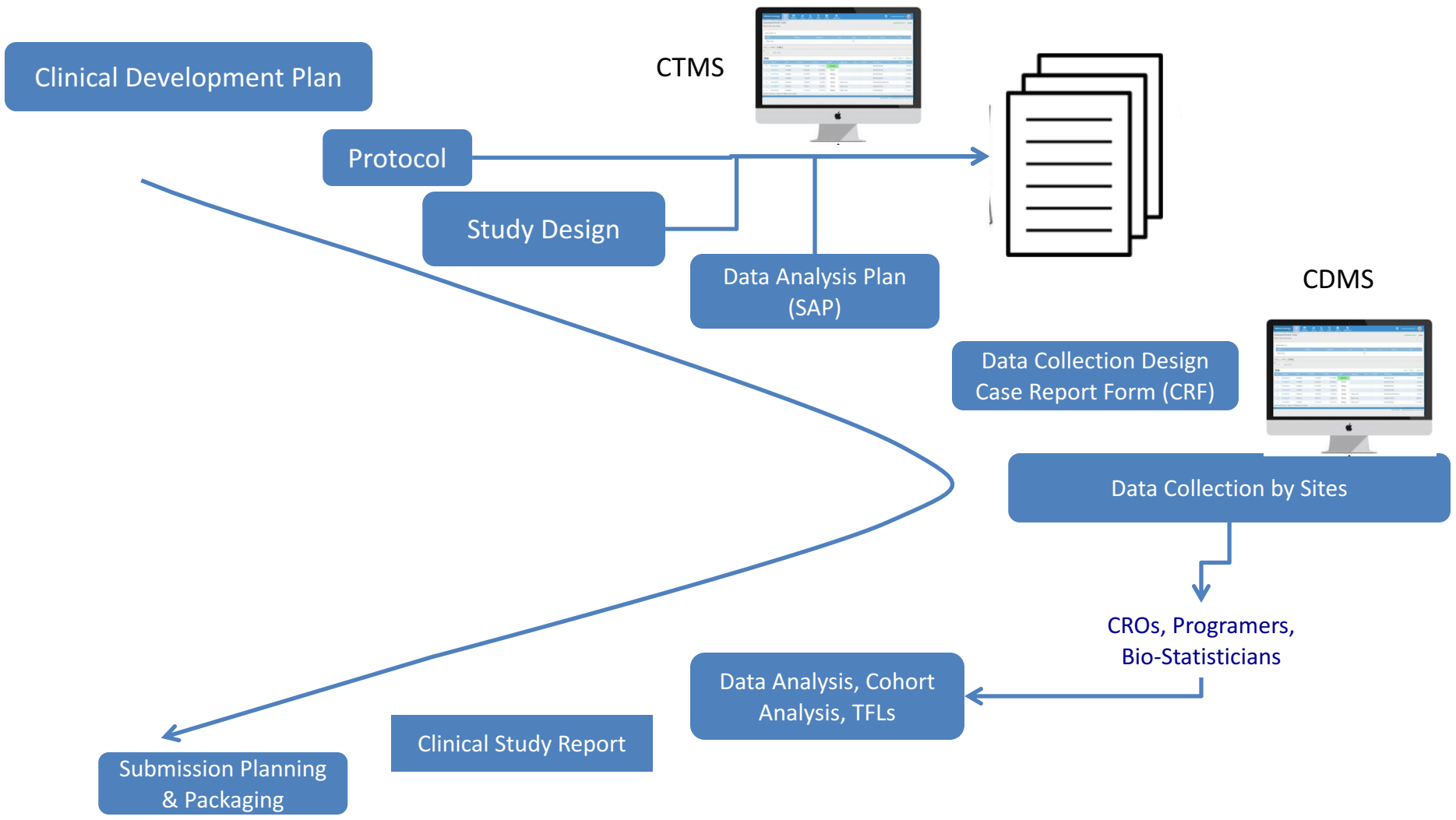
- Expensive and very time consuming data management, programming and statistical analysis effort --- **75% to 80% of time is spent with wrangling with data to get it to the point where statistical analysis can be performed** *1*2

- Regulated and designed for submission to regulatory agencies --- **GCP, GLP**

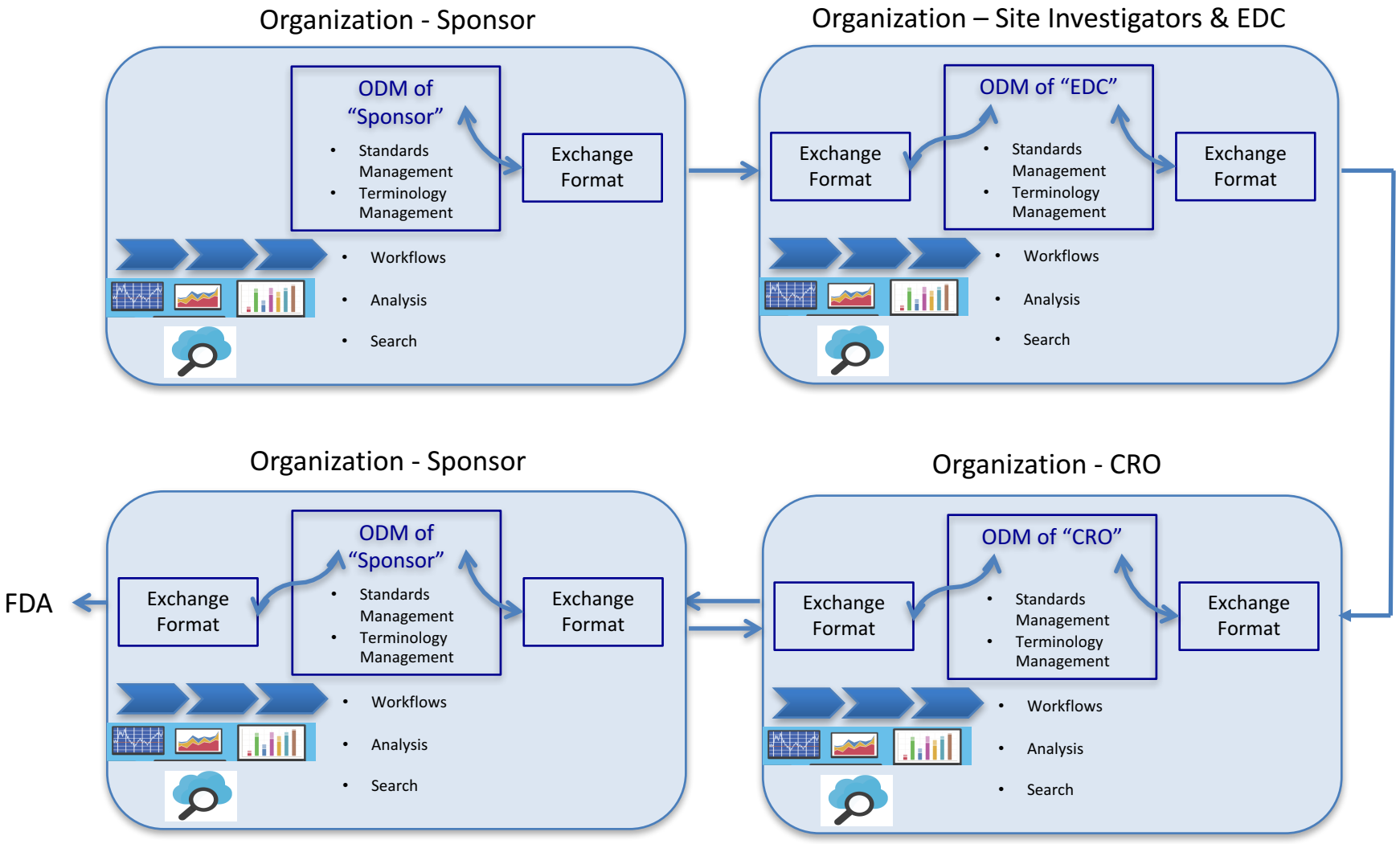
- Recent mandates require that the collected data be organized under CDISC exchange standards and analysis data sets be submitted in ADaM --- **Digital data in standard formats are here to stay.**

- New Study Protocols and Protocol Amendments are expensive and add significant time to the lifecycle of studies --- **can their production be streamlined.**

Clinical Study Workflow (Lifecycle) – "As It Has Been"



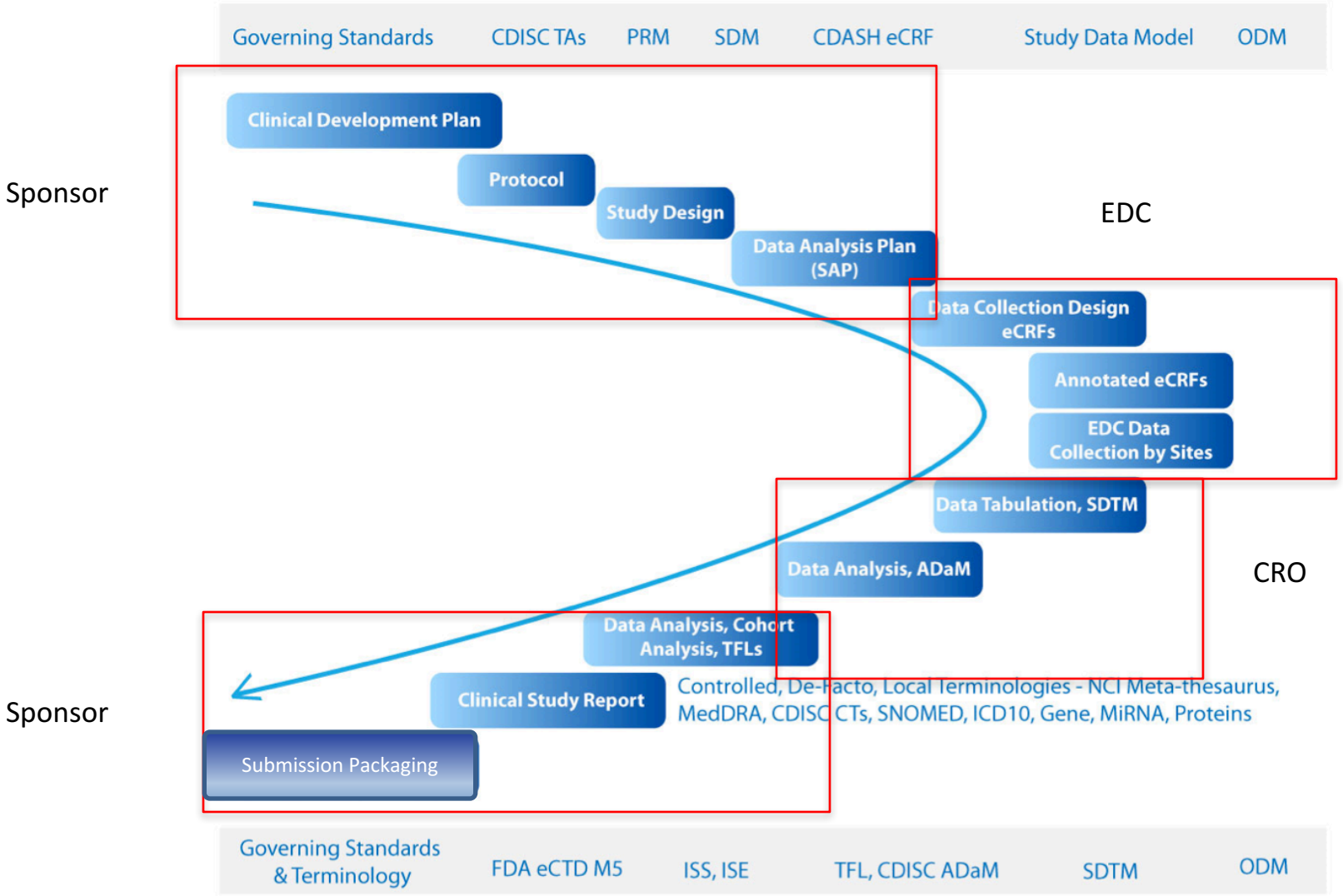
Chaining blocks of Inter-Organizational Workflows



Key building blocks of inter-organizational workflows

- Use Exchange Standard only for exchanging data, specifications, and models
- Organizations must define the business use of the exchanged data and develop an internal Operational Data Model that is tuned for that purpose
- Ensure that the required metadata and other attributes that help manage workflows, decisions, or other business purpose are incorporated into the ODM
- Data held inside each organization serves that organization using the applications, workflows and analytics they need – exchange standards are ONLY for transferring the data to another organization

Exchange Standards for Data, Specifications, and Models



Exchange Standards for Data, Specifications, and Models

- SDTM, SEND and ADaM are exchange standards for collected data or analysis data sets
  - Allows data to be exchanged between two organizations
  - Makes data machine readable
  - Exchange standards do NOT make data analysis ready – machines at the receiving end need to read and process the data
- PRM is a exchange standard for specifications of a protocol
- SDM, CDASH and ODM-XML are exchange standards for study designs, and data models including eCRF and data models for EDC

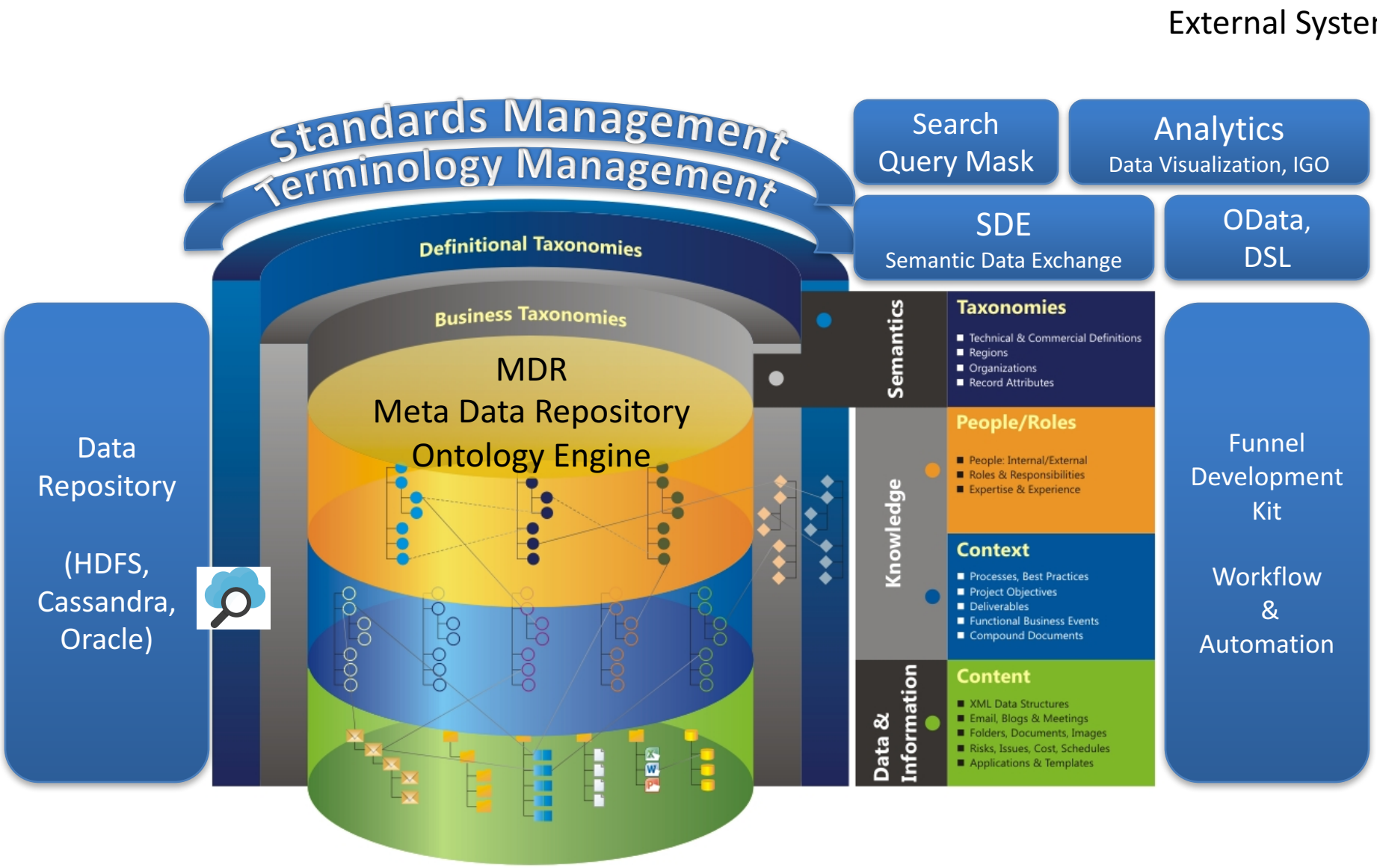
Clinical Trials – "To-Be"

- CDISC standards allow digitization and standard structure for exchange of specifications, models and data between organizations – workflow automation
- Sponsors catalog trial designs and protocols based on Therapeutic Areas – workflow between sponsors and CROs or sites
- Study Designs are specified with SDM reducing cost and impact of protocol amendments – workflow automation between sponsors and sites
- CRFs, eCRFs and annotated eCRFS can be generated with automation and CDASH to EDC ODM XML
- EDC ODM XML to SDTM generation
- Standard SAP (analysis plan) scripts or query masks to select stratified cohorts for analysis
- Planning and packaging of submissions – workflow automation

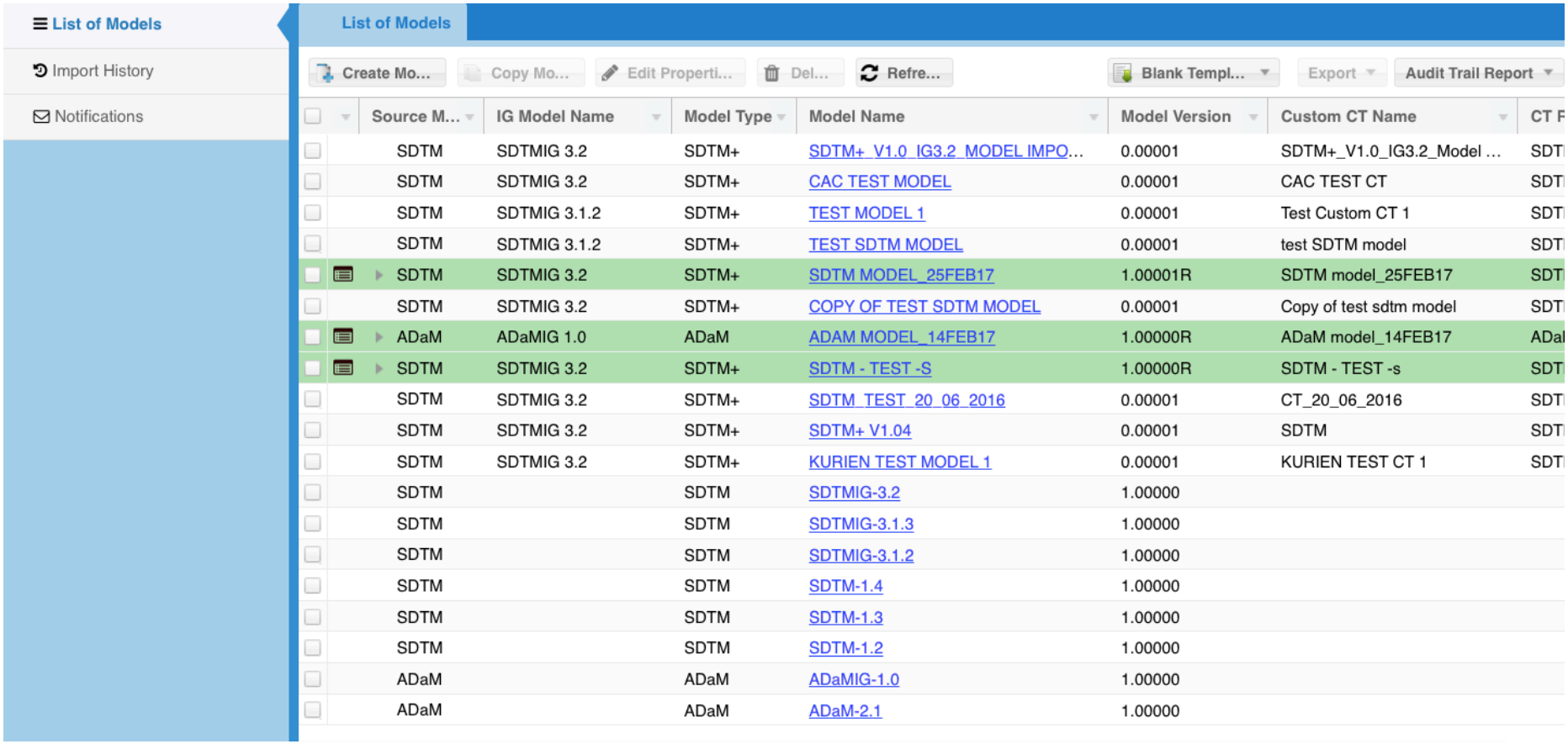
Opportunities to Streamline and Reduce Costs

- Machine Readable Protocols to Annotated eCRF
- Data Provisioning for Biostatisticians, Scientists and Medical Officers
- Tabulation Data Generation (ODM to SDTM)
- eCRF for design in EDC
- Workflow action items in CTMS
- Statistical Analysis Plan → Query Mask for Patient Cohorts & Use of Templated Scripts → Analysis Datasets
- Planning and packaging of submissions – workflow automation

Functional Modules of an ODM

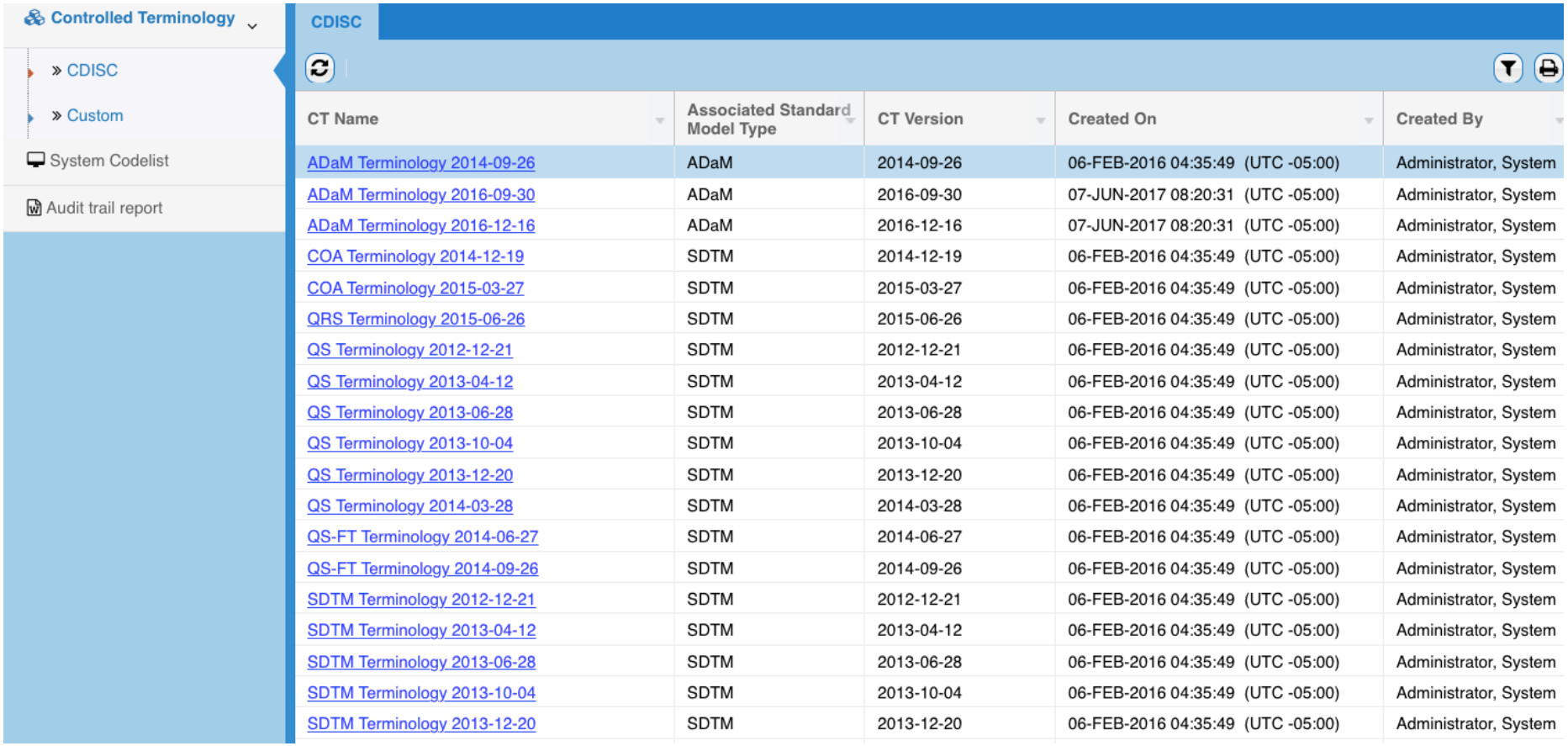


Standards Management



Source M...	IG Model Name	Model Type	Model Name	Model Version	Custom CT Name	CT F
SDTM	SDTMIG 3.2	SDTM+	SDTM+_V1.0_IG3.2_MODEL IMPO...	0.00001	SDTM+_V1.0_IG3.2_Model ...	SDT
SDTM	SDTMIG 3.2	SDTM+	CAC TEST MODEL	0.00001	CAC TEST CT	SDT
SDTM	SDTMIG 3.1.2	SDTM+	TEST MODEL 1	0.00001	Test Custom CT 1	SDT
SDTM	SDTMIG 3.1.2	SDTM+	TEST SDTM MODEL	0.00001	test SDTM model	SDT
SDTM	SDTMIG 3.2	SDTM+	SDTM MODEL_25FEB17	1.00001R	SDTM model_25FEB17	SDT
SDTM	SDTMIG 3.2	SDTM+	COPY OF TEST SDTM MODEL	0.00001	Copy of test sdtm model	SDT
ADaM	ADaMIG 1.0	ADaM	ADAM MODEL_14FEB17	1.00000R	ADaM model_14FEB17	ADa
SDTM	SDTMIG 3.2	SDTM+	SDTM - TEST -S	1.00000R	SDTM - TEST -s	SDT
SDTM	SDTMIG 3.2	SDTM+	SDTM_TEST_20_06_2016	0.00001	CT_20_06_2016	SDT
SDTM	SDTMIG 3.2	SDTM+	SDTM+ V1.04	0.00001	SDTM	SDT
SDTM	SDTMIG 3.2	SDTM+	KURIEN TEST MODEL 1	0.00001	KURIEN TEST CT 1	SDT
SDTM		SDTM	SDTMIG-3.2	1.00000		
SDTM		SDTM	SDTMIG-3.1.3	1.00000		
SDTM		SDTM	SDTMIG-3.1.2	1.00000		
SDTM		SDTM	SDTM-1.4	1.00000		
SDTM		SDTM	SDTM-1.3	1.00000		
SDTM		SDTM	SDTM-1.2	1.00000		
ADaM		ADaM	ADaMIG-1.0	1.00000		
ADaM		ADaM	ADaM-2.1	1.00000		

Terminology Management

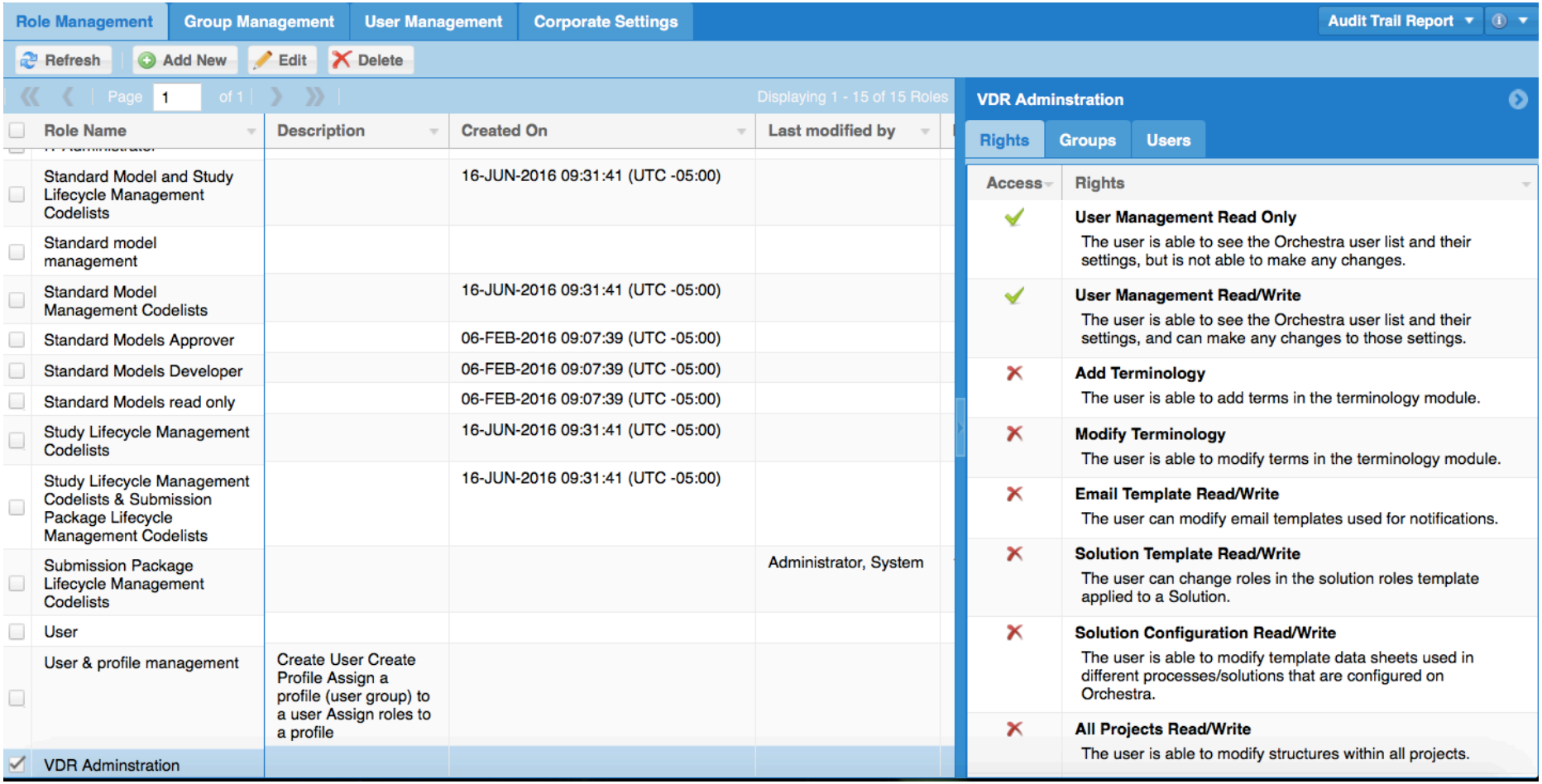


CT Name	Associated Standard Model Type	CT Version	Created On	Created By
ADaM Terminology 2014-09-26	ADaM	2014-09-26	06-FEB-2016 04:35:49 (UTC -05:00)	Administrator, System
ADaM Terminology 2016-09-30	ADaM	2016-09-30	07-JUN-2017 08:20:31 (UTC -05:00)	Administrator, System
ADaM Terminology 2016-12-16	ADaM	2016-12-16	07-JUN-2017 08:20:31 (UTC -05:00)	Administrator, System
COA Terminology 2014-12-19	SDTM	2014-12-19	06-FEB-2016 04:35:49 (UTC -05:00)	Administrator, System
COA Terminology 2015-03-27	SDTM	2015-03-27	06-FEB-2016 04:35:49 (UTC -05:00)	Administrator, System
QRS Terminology 2015-06-26	SDTM	2015-06-26	06-FEB-2016 04:35:49 (UTC -05:00)	Administrator, System
QS Terminology 2012-12-21	SDTM	2012-12-21	06-FEB-2016 04:35:49 (UTC -05:00)	Administrator, System
QS Terminology 2013-04-12	SDTM	2013-04-12	06-FEB-2016 04:35:49 (UTC -05:00)	Administrator, System
QS Terminology 2013-06-28	SDTM	2013-06-28	06-FEB-2016 04:35:49 (UTC -05:00)	Administrator, System
QS Terminology 2013-10-04	SDTM	2013-10-04	06-FEB-2016 04:35:49 (UTC -05:00)	Administrator, System
QS Terminology 2013-12-20	SDTM	2013-12-20	06-FEB-2016 04:35:49 (UTC -05:00)	Administrator, System
QS Terminology 2014-03-28	SDTM	2014-03-28	06-FEB-2016 04:35:49 (UTC -05:00)	Administrator, System
QS-FT Terminology 2014-06-27	SDTM	2014-06-27	06-FEB-2016 04:35:49 (UTC -05:00)	Administrator, System
QS-FT Terminology 2014-09-26	SDTM	2014-09-26	06-FEB-2016 04:35:49 (UTC -05:00)	Administrator, System
SDTM Terminology 2012-12-21	SDTM	2012-12-21	06-FEB-2016 04:35:49 (UTC -05:00)	Administrator, System
SDTM Terminology 2013-04-12	SDTM	2013-04-12	06-FEB-2016 04:35:49 (UTC -05:00)	Administrator, System
SDTM Terminology 2013-06-28	SDTM	2013-06-28	06-FEB-2016 04:35:49 (UTC -05:00)	Administrator, System
SDTM Terminology 2013-10-04	SDTM	2013-10-04	06-FEB-2016 04:35:49 (UTC -05:00)	Administrator, System
SDTM Terminology 2013-12-20	SDTM	2013-12-20	06-FEB-2016 04:35:49 (UTC -05:00)	Administrator, System

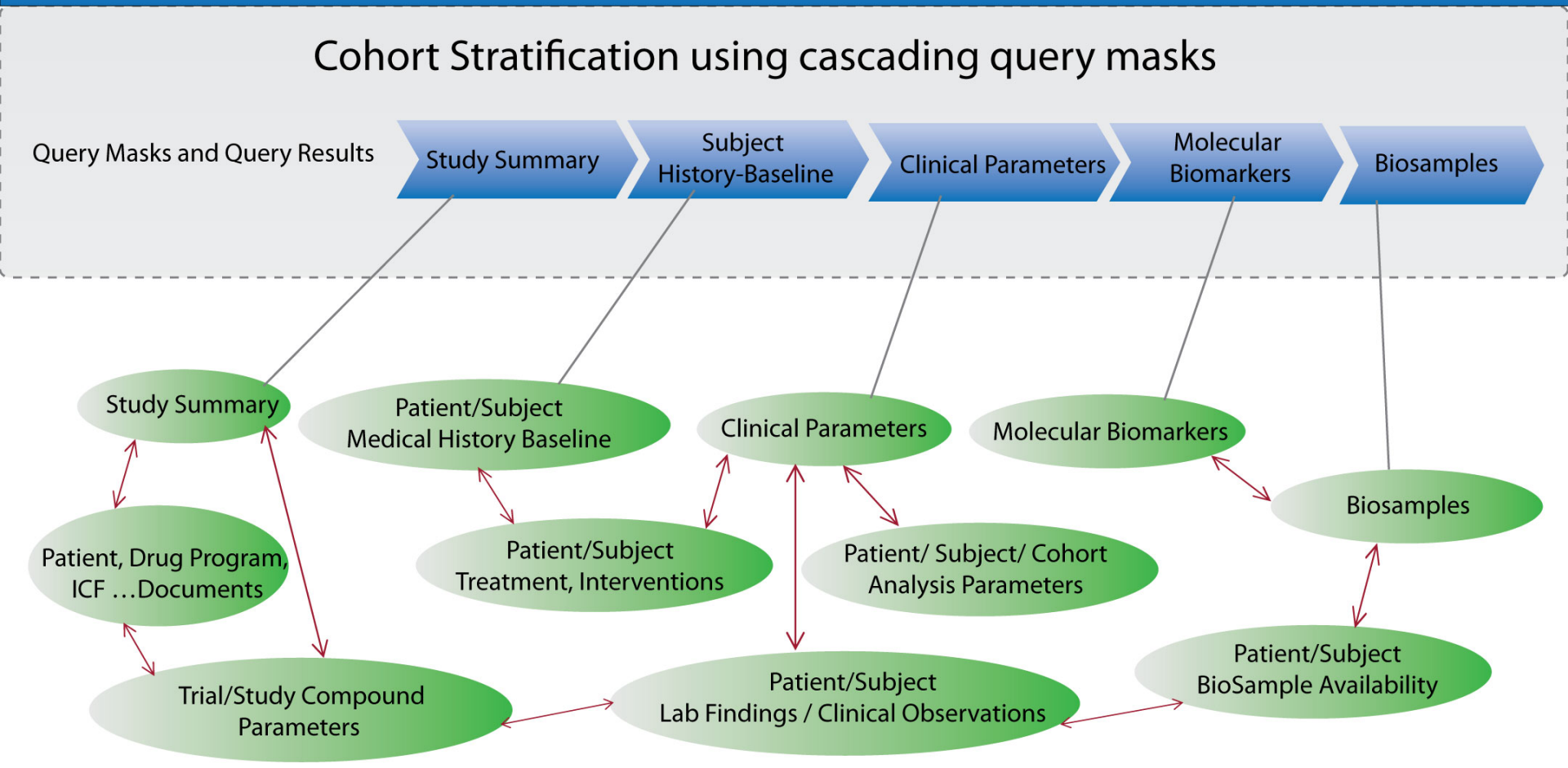
Terminology Management

- Manage Terminologies
  - CDISC Controlled Terminologies – Prepared and released by CDISC
  - Custom Terminologies – Sponsor defined terminologies during standard model creation
  - System codelists - Codelists utilized by system in solutions for values in the drop downs for various modules.
  - Audit Trail – Track the changes performed within the system codelists

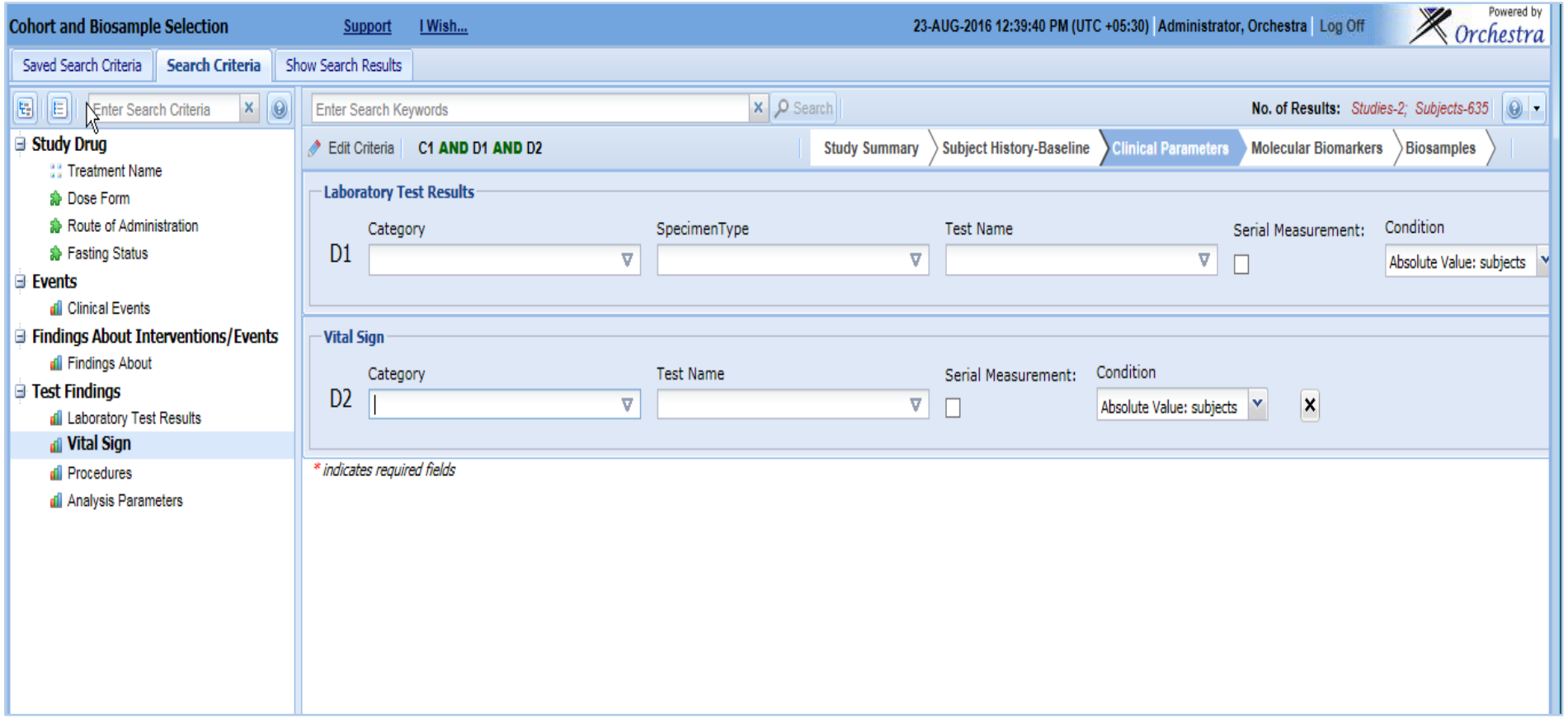
Role Based Access Control Management



Finding Actionable Results from Disparate Data Sources



Clinical Parameter Search



Repurposing Clinical Study Data

- Data generation and analysis within one trial ✓

- Which data exists beyond one trial?
- Where is the data?
- What is in the data?
- Who owns the data? Who can give access?
- Are data sets compatible?
- Use data beyond one trial (if covered by ICF)

→ high manual effort or not possible at all

Courtesy: "*PORTIN – a game changer towards translational research Portal for translational data integration*" *22nd September 2017 – 4th Global Precision Medicine & Biomarker Summit – Christiane Unger*

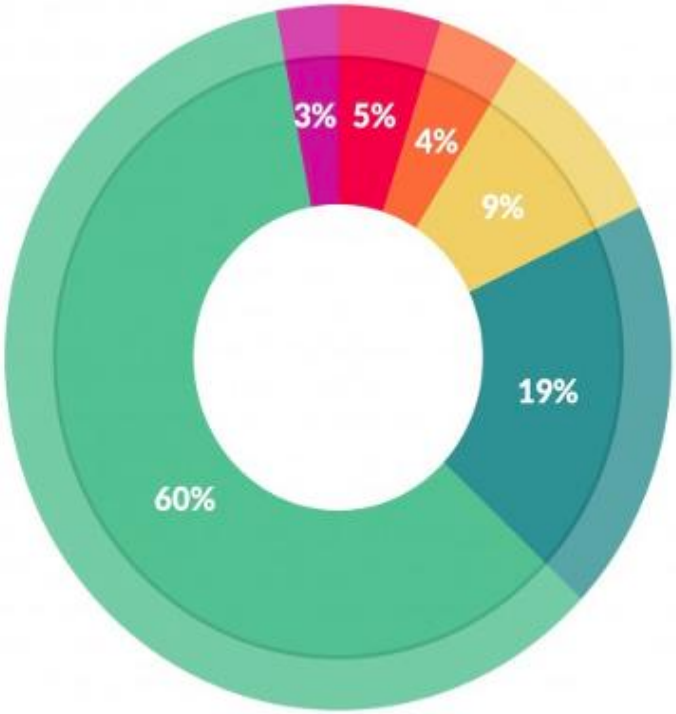
Organizations spend a large amount of resources to acquire data in clinical trials.

Information is used within a specific trial but is not (easily) available for further use

What if we wanted to know

In all the clinical trials we have conducted, how many individuals are there with biomarker X measurement greater than Y, and data available for clinical endpoint Z? How many of those have DNA available and consent for research use?

How much effort might it take?



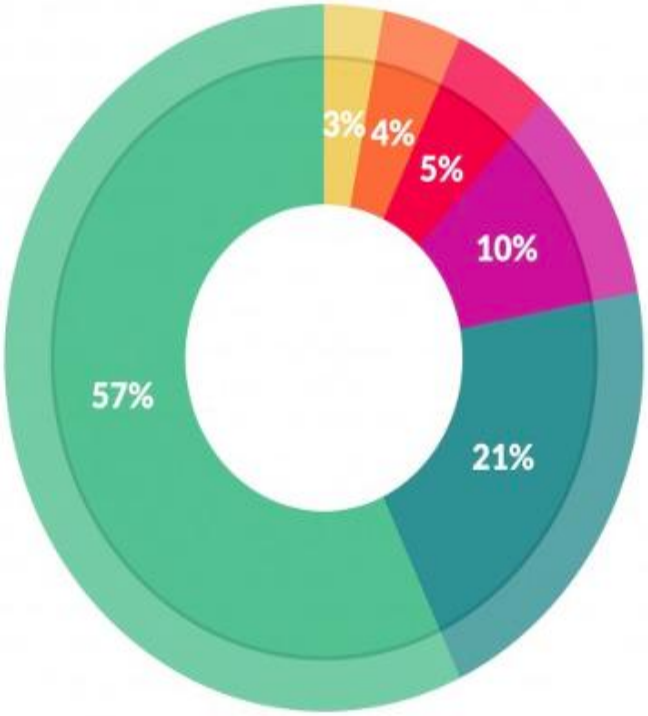
What data scientists spend the most time doing

- Building training sets: 3%
- Cleaning and organizing data: 60%
- Collecting data sets; 19%
- Mining data for patterns: 9%
- Refining algorithms: 4%
- Other: 5%

79%

Forbes - Cleaning Big Data: Most Time-Consuming, Least Enjoyable Data Science Task, Survey Says

Was it fun?



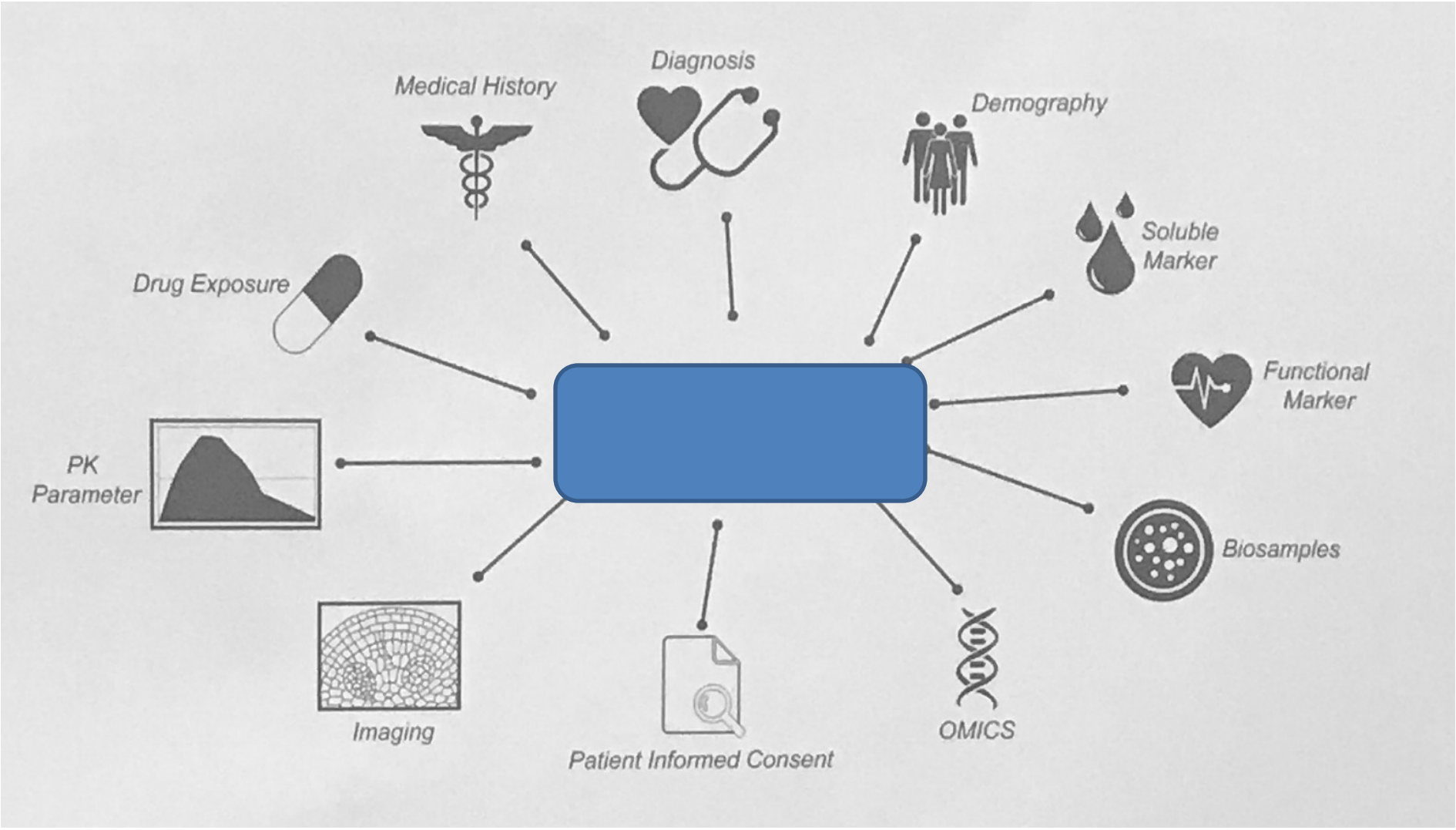
What's the least enjoyable part of data science?

- Building training sets: 10%
- Cleaning and organizing data: 57%
- Collecting data sets: 21%
- Mining data for patterns: 3%
- Refining algorithms: 4%
- Other: 5%

78%

Forbes - Cleaning Big Data: Most Time-Consuming, Least Enjoyable Data Science Task, Survey Says

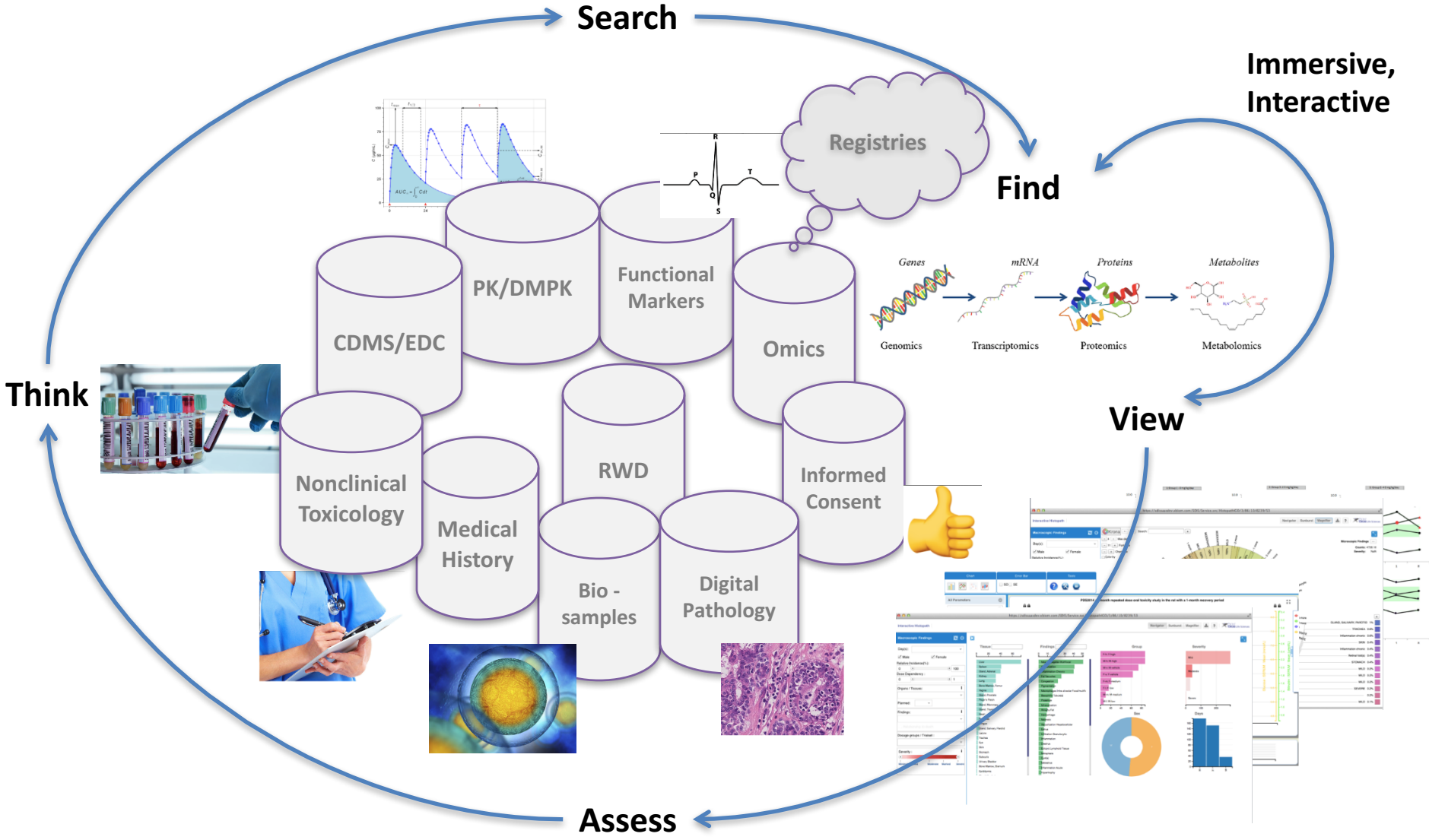
Clinical Trials – As-Is



Disparate Techniques for Disparate Sources

- Curation – in search for “single truth”
- Mapping – from source format to a consumable form
- Ontology – to relate information to other related information
- Semantic harmonization – synonyms make data searchable
- Normalization – of data models and units for searchable data
- Repurpose single truth data into multiple forms - kaleidoscope
- Provenance – source or origin of data and time stamps
- Indexing and Search - across all the single truth data lakes
- Real Time Data Access Rights -based on Privacy Laws, ICFs, and Roles
- API’s – for interoperability with external applications

Support the Research Cycle

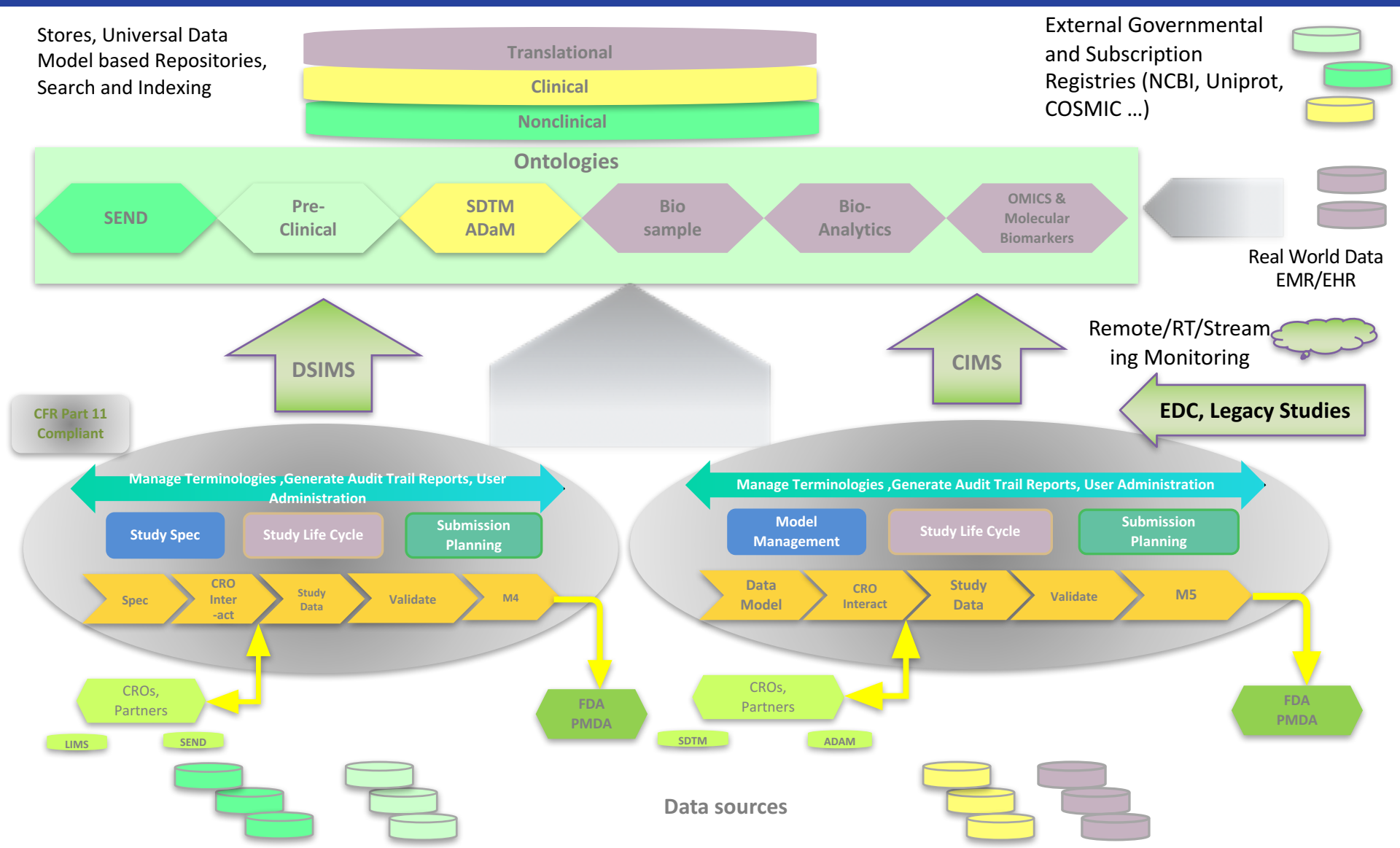


About PointCross Life Sciences

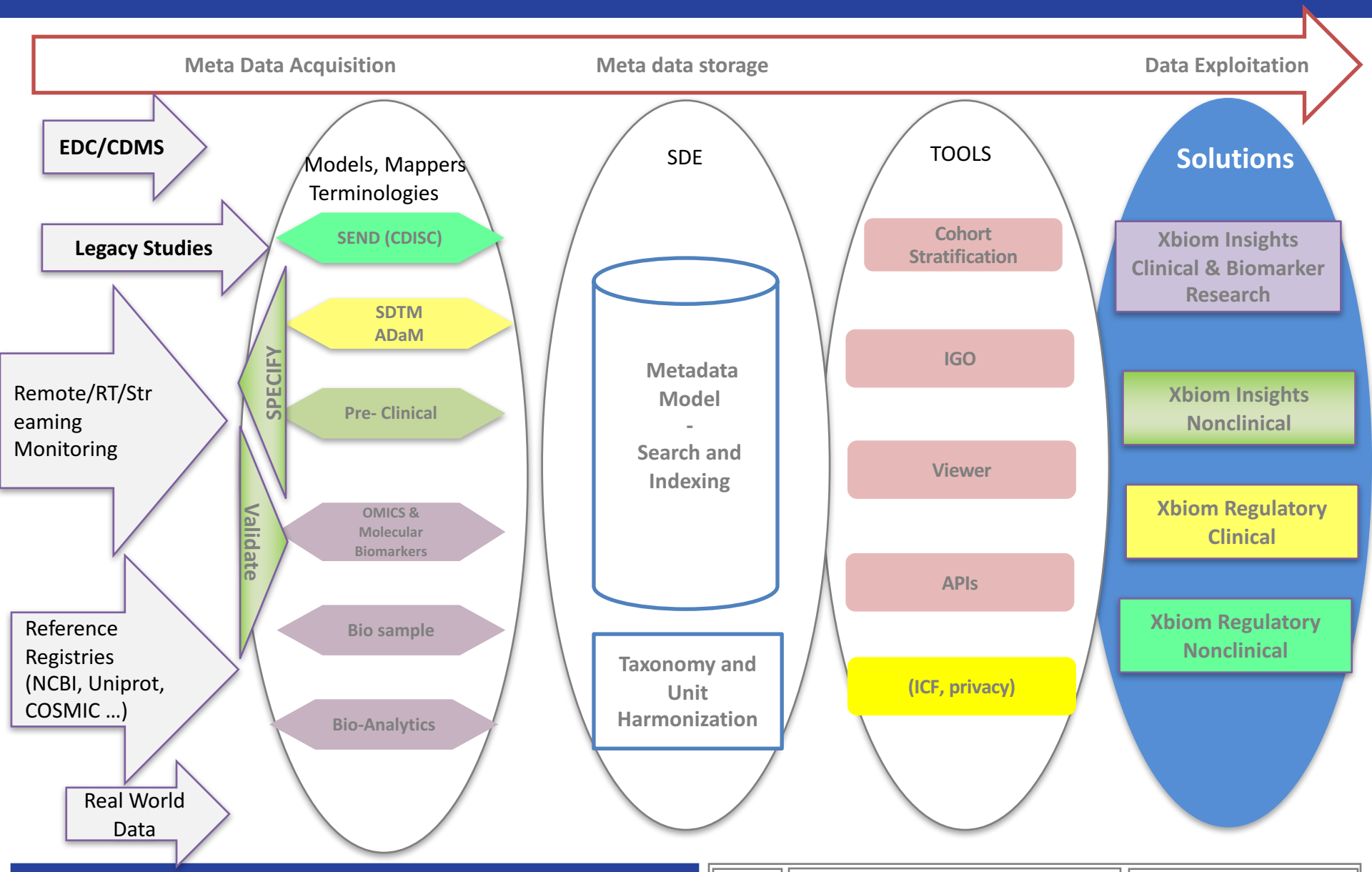
- PointCross Life Sciences is a wholly owned subsidiary of PointCross Inc.
- Founded Nov 1999
  - ✓ Initial industry focus – Oil and Gas Exploration and Production
  - ✓ Technology Focus – Big data analytics, warehousing of structured and unstructured data from natural processes
- BioPharma industry focus since 2009
- Global Clients: Roche, Sanofi, Bayer, Eisai, Ipsen, J&J ….
- Contributor, Exhibitor at Industry
- FDA contractor and provider of technology
- Locations – Foster City, CA, Silver Spring, MD, Paris, France, Bangalore, India
- Gartner “Cool Vendor” in Life Sciences in 2013

Supplemental Slides

PointCross Vision for a 360 degree semantically integrated R&D Hub



Xbiom - data and meta data maturation



Xbiom Clinical for Insights or Study Workflow Automation

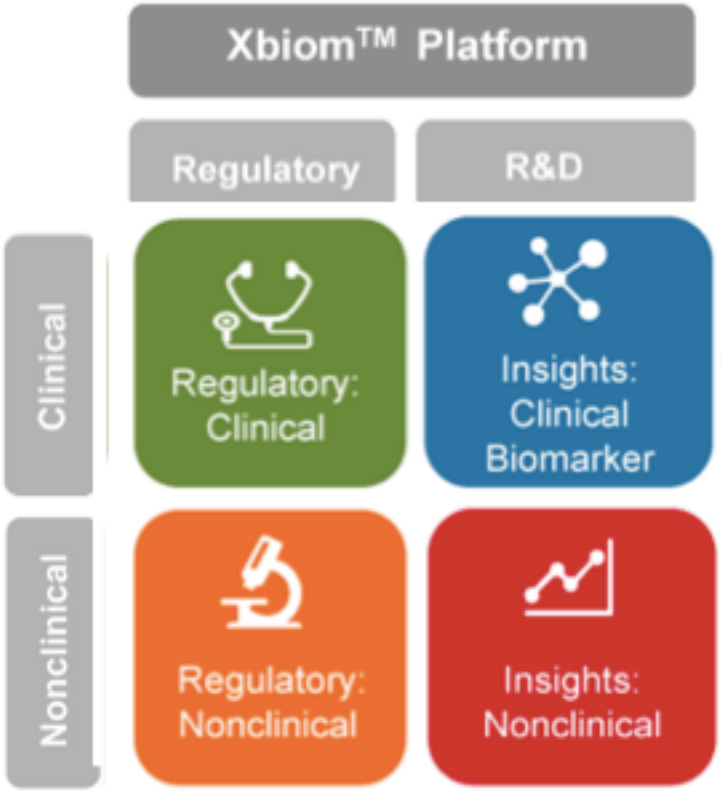
Model Management and Study Data Specification

Study & Submission Workflow Automation –

From EDC to Analysis

From Analysis to TFL

Packaging – M5 eCTD



Clinical Trial and Bio-Marker warehouse

Searchable – across studies

Discoverable

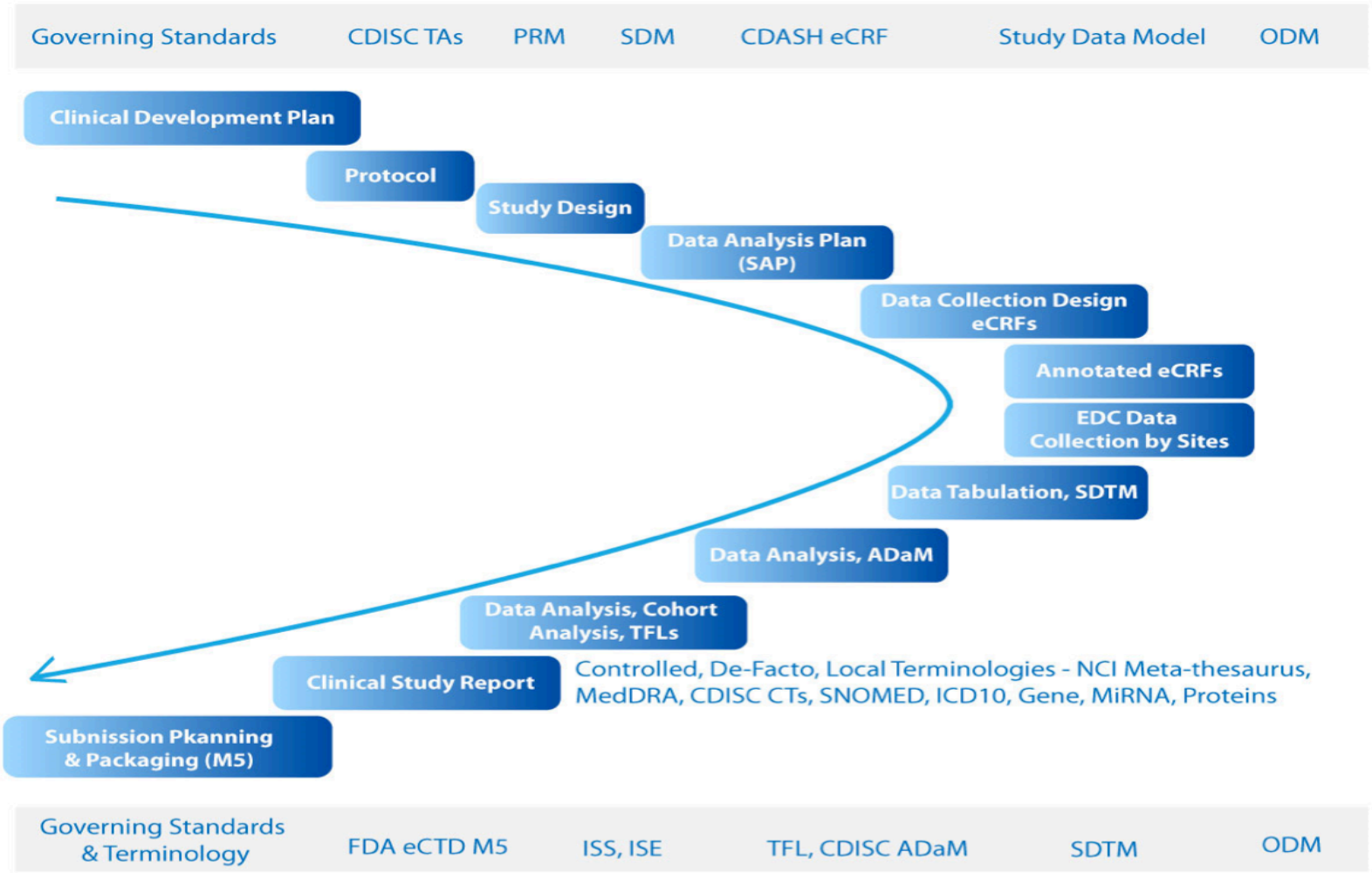
Analysis Ready

Virtual Studies

Stratified Cohort Selection

Bio-Sample Availability tracking for additional bio-marker surveys

Exchange Standards for Data, Specifications, and Models



Clinical Trials – As-Is

- Pharma companies (sponsors) document protocols for the trials
- The design of the trial and its organization is specified through a CTMS system or documented
- Each trial tends to be designed from scratch with little templating - cost of trials increase
- CRFs are manually designed
- Data collected and managed by CDMS
- Lots of data managers needed to maintain and make the right data available to the programmers and script developers
- Lots of bio-statisticians needed to extract analysis data sets and do analysis on computed endpoints
- Manual labor intensive work generating the tables, listing, and figures to support medical assessment of safety or efficacy and generation of Study Reports

Building blocks of Inter-organizational Workflows