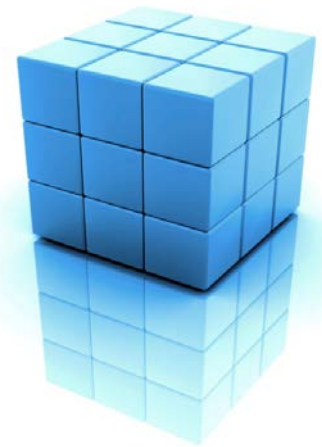


Agile Statistical Computing Environment for Clinical and Biomarker Data



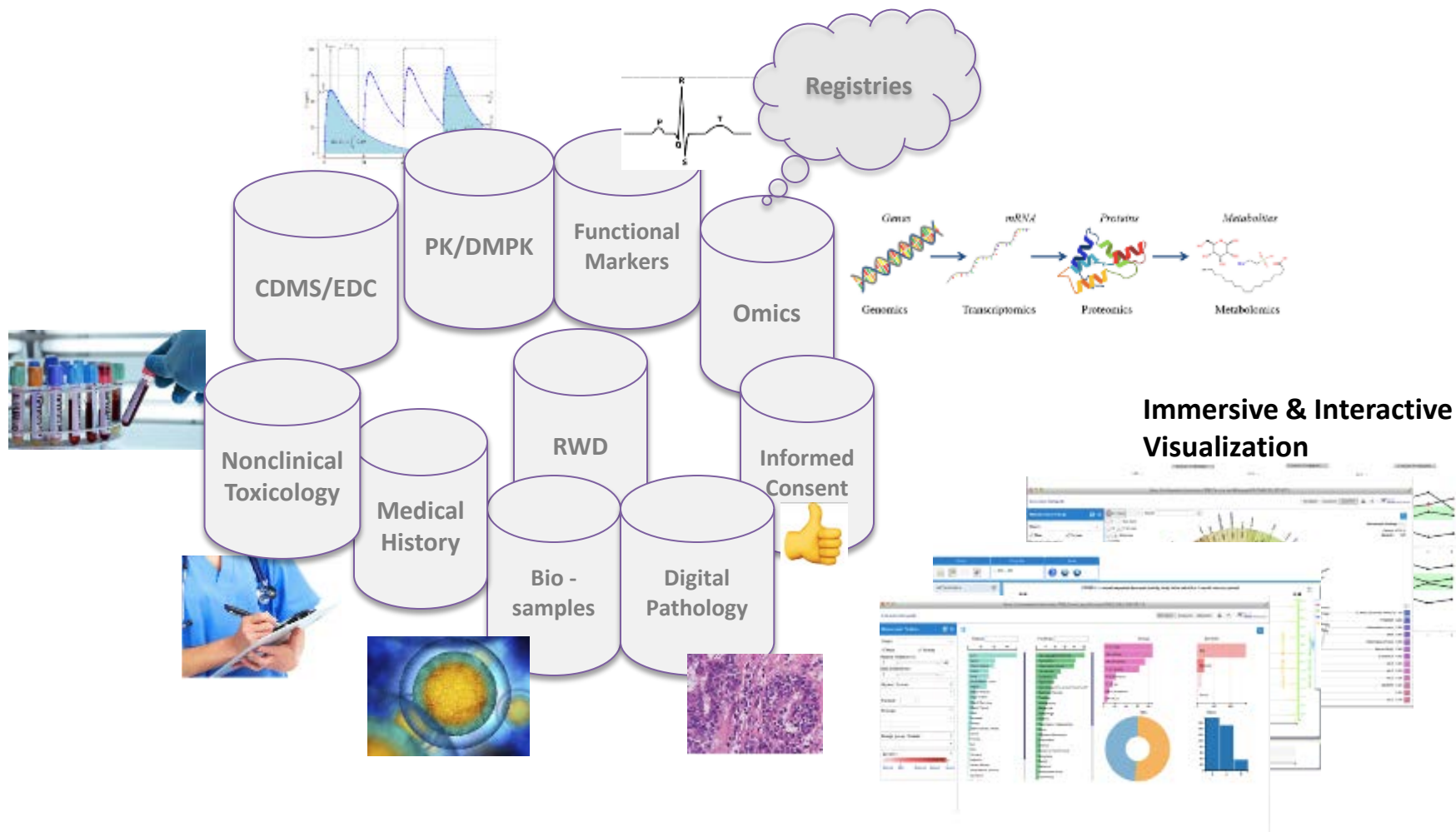
www.pointcross.com

About PointCross Life Sciences

- PointCross Life Sciences is a wholly owned subsidiary of PointCross Inc.
- Serving Pharma and BioTechs since 2009
 - ✓ Xbiom Technology for Regulatory Submissions and Research – Nonclinical and Clinical
- Global Clients: Roche, Sanofi, Bayer, Eisai, Ipsen, NGM,
- Contributor to PhUSE, CDISC, Industry Events
- 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



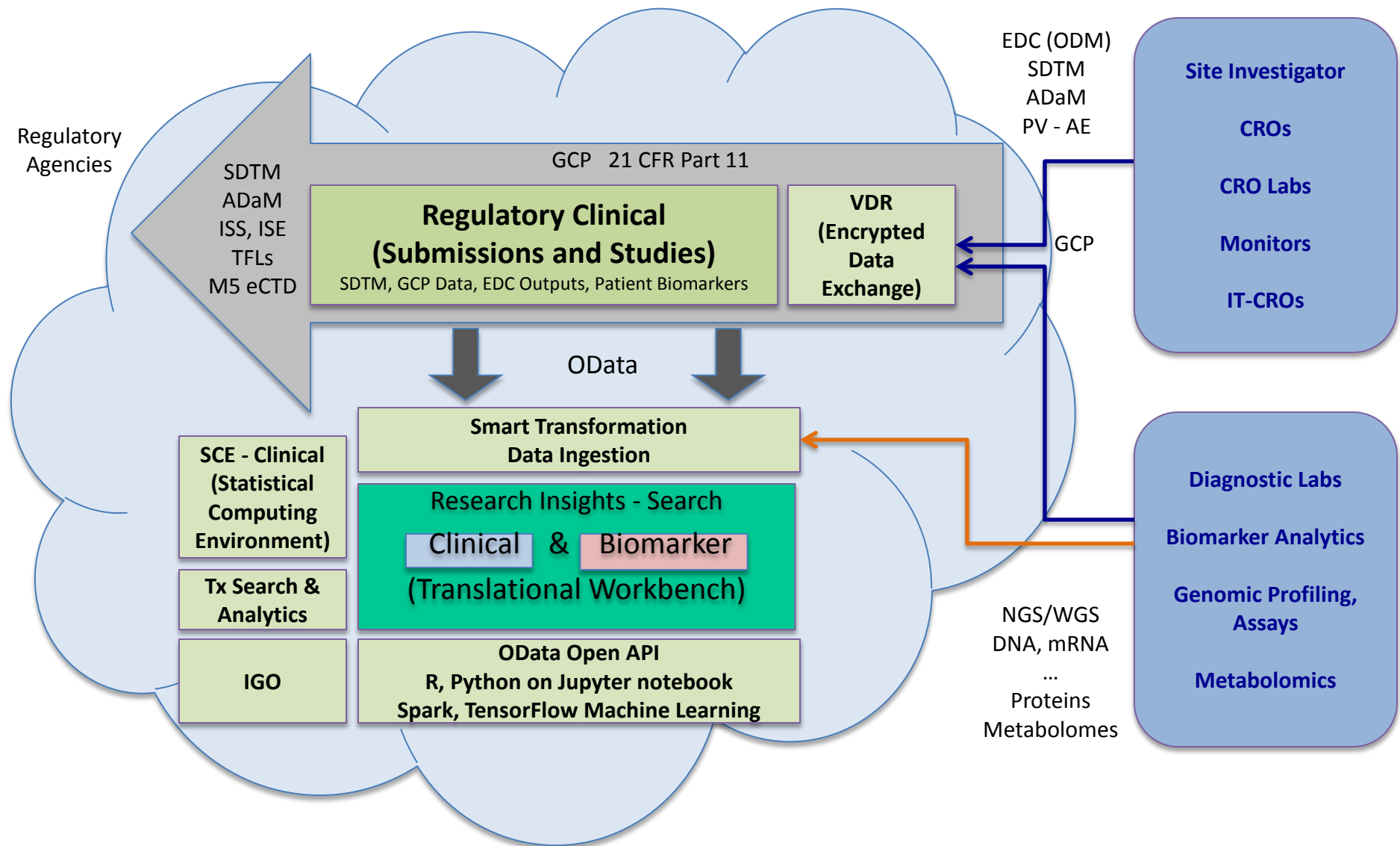
Range of data types collected in Clinical Trials



Sponsor's Clinical Data Analysis

- Clinical trial data is provided by CROs to a GCP qualified environment
- Executing statistical analysis plan (SAP) is done by validated scripts (R, SAS) in this GCP environment
- Programming the scripts that analyze the clinical data to statistical analysis plans, or additional analysis, that may be submitted will also need to be:
 - Tested
 - Validated, and
 - Migrated to the GCP process and workflows to execute the SAP or other analysis
- Moving a copy of the data (SDTM) from the regulated space to SAS for analysis data sets (ADaM) for follow on analysis is common practice – but usually lacks:
 - Open data access – for FAIR access to regulatory stores
 - Interpretive, simple programming language in addition to SAS or R for agility
 - Contemporaneously journaling notebook to execute programs and statistical analysis for speed, agility, traceability, for a flexible but powerful environment

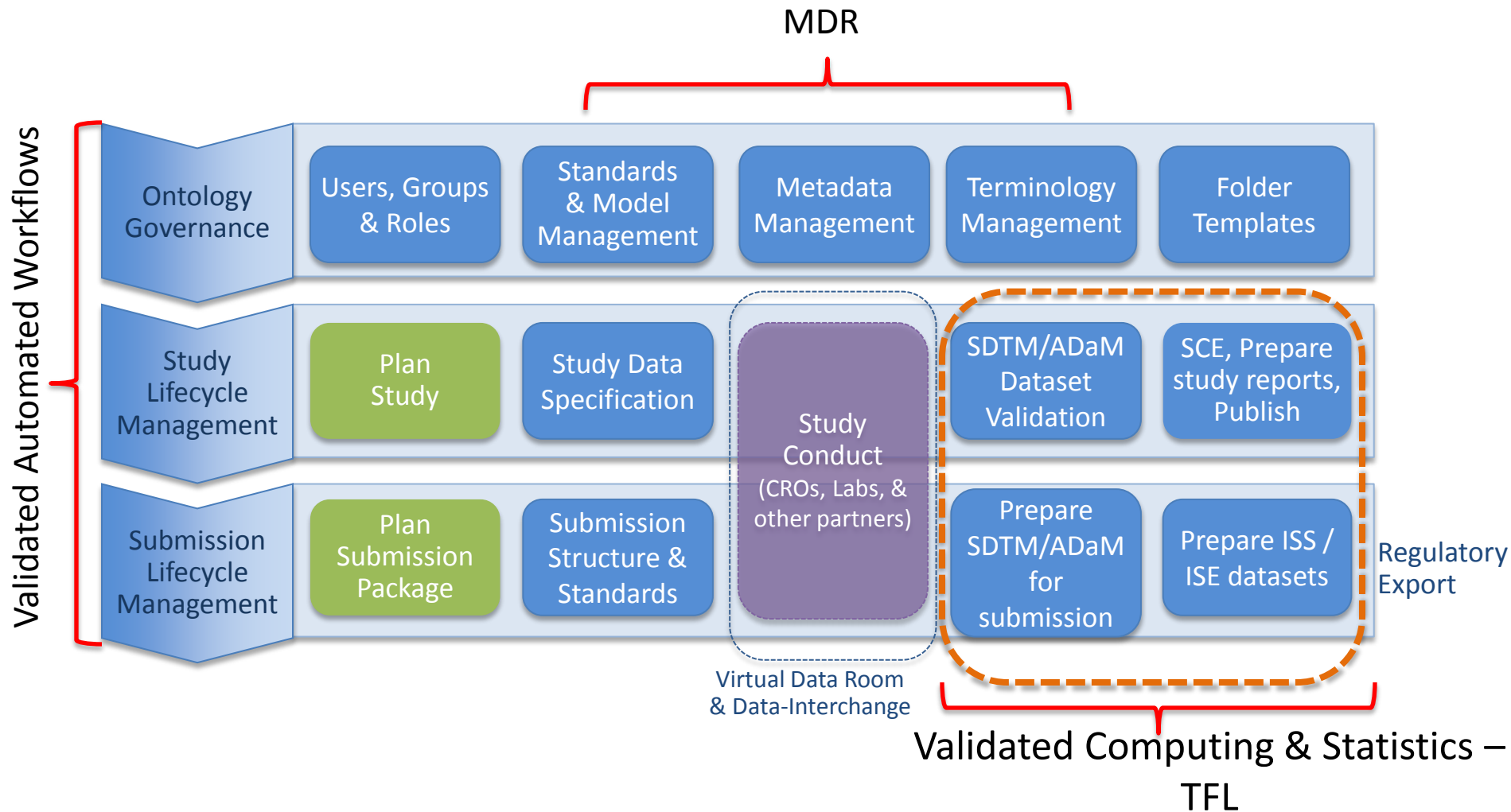
Interactive Computing Environment



Agile but Partitioned Computing Environment Clinical Data Analysis

- Partitioned GCP environment and research workbench area without co-mingling data:
 - Clinical and biomarker data analysis for submission can be analyzed to support both SAP
 - Agile development of analysis scripts and other translational research by BioPharma organizations
- SAP is run with validated scripts (R, SAS) in the GCP environment
- OData open access read-only drawn down of SDTM data to the research and computing environment in the research workbench
- Machine learning (TensorFlow) augmented data ingestion and indexing after data model mapping of clinical and biomarker data –
 - In big data (Hadoop, Cassandra) stores with Spark/Map-Reduce
 - For search of patients/cohorts across baselines, clinical findings, and biomarkers (Tx Research)
 - For creating analysis data sets for developing the SAP scripts (Support Clinical Ops process)
- Interact with computing environment on Jupyter – a contemporaneous interactive journal
 - R (or SAS) scripts,
 - Simple programming language - Python for complex analysis for analysis on search results

Regulatory Functions – Submission Workflow

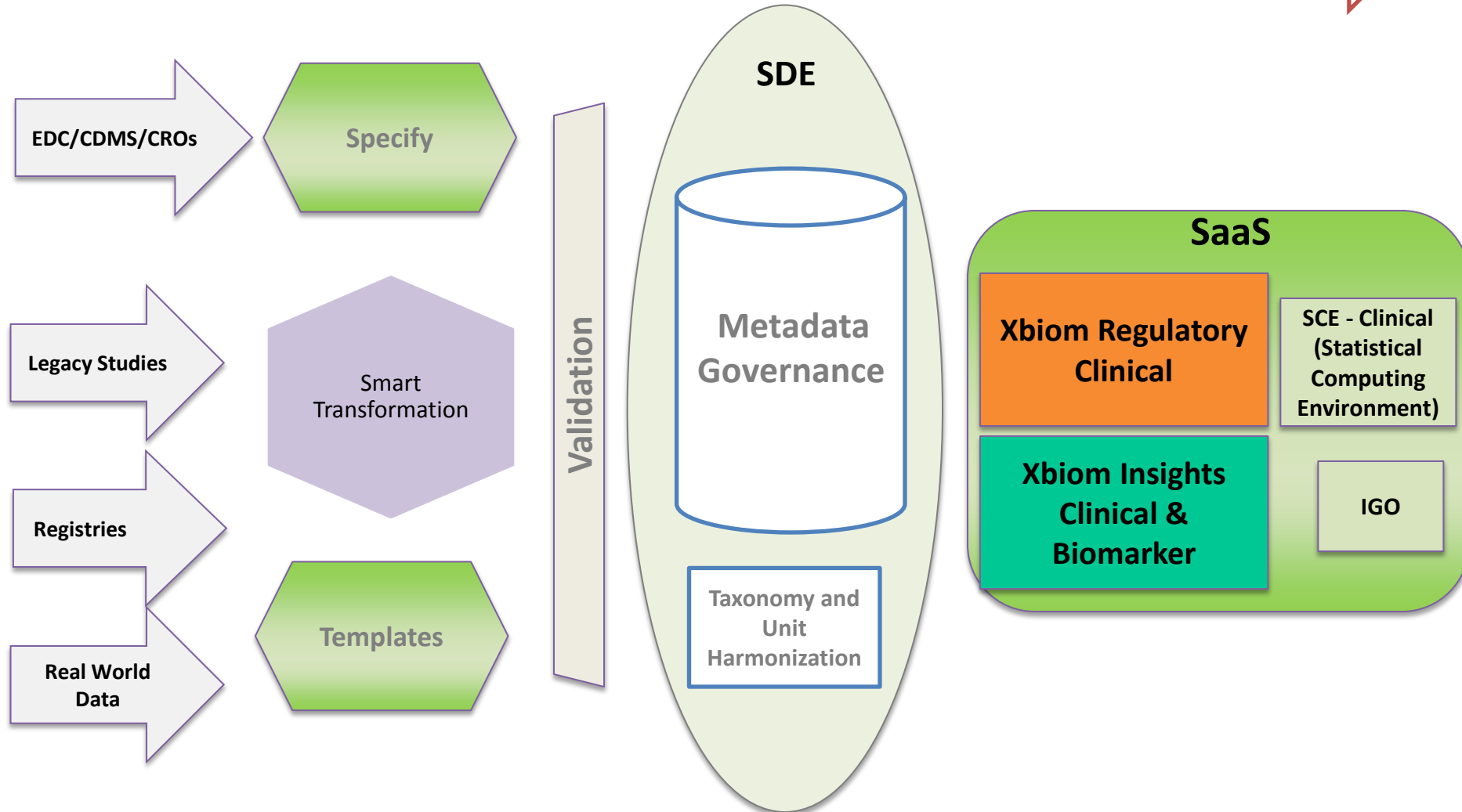


Smart Data Curation for Ingestion and Indexing

Data Mapping

Meta data store

Data Exploitation



Search within Indexed Clinical Study and Biomarker data

Study Summary

Subject History-Baseline

Clinical Parameters

Molecular Biomarkers

Biosamples

SDTM Variables

Cohort and Biosample Selection

[Support](#)

[I Wish...](#)

Production

23-MAR-2019 14:00:19 (UTC +05:30) | Ramesh, Raja |

Powered by
Orchestra

Saved Search Criteria

Search Criteria

Show Search Results

Enter Search Criteria

B1 AND B2

Search

Clear

No. of Results: *Studies-1; Subjects-167*

Study Summary

Subject History-Baseline

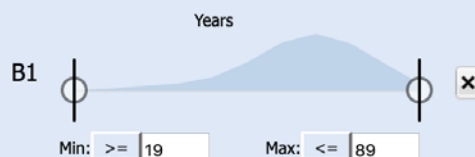
Clinical Parameters

Molecular Biomarkers

Biosamples

SDTM Variables

Age *Matched Studies-7; Subjects-2871*



Laboratory Test Results *Matched Studies-1; Subjects-167*

Category	SpecimenType	Test Name
B2		
<i>* indicates required fields</i>		
		Alanine Aminotransferase Measurement
		Albumin Measurement
		Alkaline Phosphatase Measurement
		Anisocyte Measurement
		Aspartate Aminotransferase Measurement
		Calcium Measurement
		Chloride Measurement
		Cholesterol Measurement
		Color Assessment
		Creatine Kinase Measurement
		Creatinine Measurement

Jupyter – Contemporaneous Journaling and Program Execution

The screenshot displays the Jupyter Notebook web interface. At the top, the header shows the Jupyter logo, the notebook name 'APISample', and the last checkpoint '03/12/2019 (autosaved)'. A 'Logout' button is in the top right. Below the header is a menu bar with 'File', 'Edit', 'View', 'Insert', 'Cell', 'Kernel', 'Widgets', and 'Help'. The 'File' menu is open, showing options like 'New Notebook', 'Open...', 'Make a Copy...', 'Rename...', 'Save and Checkpoint', 'Revert to Checkpoint', 'Print Preview', 'Download as', 'Trust Notebook', and 'Close and Halt'. A sub-menu for 'New Notebook' is also visible, showing 'R' and 'Python 3'. The main area contains a code cell with a pink background showing the output of a package installation: 'Loading required package: bitops', 'Loading required package: stringr', 'Loading required package: stringi', 'Loading required package: XML', 'Loading required package: git2r', 'Loading required package: getPass', 'Loading required package: httr', 'Installing package: 'httr'', and 'The following objects are masked from 'package:git2r':'. Below this is an input cell with the command 'In [3]: Xbiom.Login("https://insightsaas.xbiom.com", "raja@pointcross.com")' and its output: '.....', '[1] "Successfully logged on"'. Below that is another input cell with the command 'In [4]: Studies <- Xbiom.ListOfStudies()'.

Examples of OData consumption of APIs

XbiomAPI

2019-01-30

- 1 Introduction to Xbiom R API
- 2 Functions in Xbiom R API
 - 2.1 Login Function
 - 2.2 List of Studies
 - 2.3 List of Cohorts
 - 2.4 Study Parameters List
 - 2.5 Get Secret Restricted Subject Details
 - 2.6 Export to Dataset
 - 2.7 Export to CSV
 - 2.8 Document Search
 - 2.9 GetODataURL
 - 2.10 ExportODataToCSV
 - 2.11 ExportODataToDataset
 - 2.12 SearchCounts
 - 2.13 Search
 - 2.14 DeleteCohort
 - 2.15 Summaries
 - 2.16 Pivots
 - 2.17 Xbiom.ListofTerminologies
 - 2.18 Xbiom.TerminologySearch
- 3 Additional Options:
 - 3.1 HTTPS and Certificate related issues
 - 3.2 Proxy settings
 - 3.3 Retrieving Studies based on Study's UniqueID
- 4 Query Options
 - 4.1 \$select
 - 4.2 \$orderby
 - 4.3 \$filter

Display a menu

1 Introduction to Xbiom R API

Typical Interactive Session on Jupyter with APIs

jupyter APISample Last Checkpoint: 03/12/2019 (unsaved changes) Logout

File Edit View Insert Cell Kernel Widgets Help Not Trusted R O

📄 + 🔍 📄 ⬆️ ⬆️ ⏮️ ⏭️ ↻️ Code 🖨️

```
In [8]: Xbiom.Login("https://insightsaas.xbiom.com", "raja@pointcross.com")
.....
[1] "Successfully logged on"
```

Service End Point

```
In [11]: ServiceEndPoint <- Xbiom.GetODataURL()
print(paste0("URL -- ", ServiceEndPoint, " "))
queryString <- "$format=json"

[1] "URL -- https://insightsaas.xbiom.com/XbiomAPI/V1/4/90/1uhqh7x9sW7C64AD292BD841EC91uR@MjOaLN9695262EC5BE9A/XbiomO
Data "
```

```
In [13]: Studies <- Xbiom.ListOfStudies()
Studies
```

STUDYID	UNIQUEID	TrialPhase	CompoundName	Indication	StudyTitle
SSP-001	169	list()	list()	list()	
PCTEST1	243	list()	list()	list()	
PC001	248	list()	list()	list()	
PC002	257	list()	list()	list()	
PCTEST2	294	list()	list()	list()	
NIDA-CTN-0001	462	list()	list()	list()	
CDISCPIL01	485	list()	list()	list()	

Accessing pre-defined Cohorts (SAP) for ADaM

jupyter APISample Last Checkpoint: 03/12/2019 (autosaved)

Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted

R O

Markdown

```
In [16]: Cohorts <- Xbiom.ListOfCohorts(SearchString = "Renal")
Cohorts
```

ID	TITLE	SEARCHCRITERIA	CREATEDON	UPDATEDON	FIRSTNAME	LASTNAME	DESCRIPTION	STUDYIDS	DRAFTFLG	SEARCHTYPE	USERID
638	Renal function	Search Criteria : C1 AND C2 C1 = Findings of (Test Name:Blood Urea Nitrogen Measurement) in Laboratory Test Results C2 = Findings of (Test Name:Creatinine Measurement) in Laboratory Test Results	2019-03-07	2019-03-07	arpitha	h	Renal function	PC001, PC002	2	2	12
633	Renal function	Search Criteria : C1 AND C2 C1 = Findings of (Test Name:Blood Urea Nitrogen Measurement) in Laboratory Test Results C2 = Findings of (Test Name:Creatinine Measurement) in Laboratory Test Results	2019-03-07	2019-03-07	arpitha	h	Renal function		1	2	12

Cohort Navigation

```
In [17]: cohortSubjects <- paste0(ServiceEndPoint, "/Cohorts(", Cohorts$ID[1], ")", "/Subjects?", queryString)
cohortFindings <- paste0(ServiceEndPoint, "/Cohorts(", Cohorts$ID[1], ")", "/Findings?", queryString)
print(paste0("SubjectsURL - ", cohortSubjects, " "))
print(paste0("FindingsURL - ", cohortFindings, " "))

[1] "SubjectsURL - https://insightsaas.xbiom.com/XbiomAPI/V1/4/90/1uhqh7x9sW7C64AD292BD841EC91uR@MjOaLN9695262EC5BE9A/XbiomOData/Cohorts(638)/Subjects?$format=json "
[1] "FindingsURL - https://insightsaas.xbiom.com/XbiomAPI/V1/4/90/1uhqh7x9sW7C64AD292BD841EC91uR@MjOaLN9695262EC5BE9A/XbiomOData/Cohorts(638)/Findings?$format=json "
```

Setting up Search Criteria with Agility

jupyter APISample Last Checkpoint: 03/12/2019 (unsaved changes)

Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted

R O

Markdown

Search

```
In [28]: Xbiom.SearchCounts("{AGE ge 40} and ({STUDYID eq 'PC001'} or {STUDYID eq 'PC002'} )")
```

StudyCount	SubjectsCount
------------	---------------

2	446
---	-----

```
In [29]: Cohort <- Xbiom.Search("{AGE ge 40} and ({STUDYID eq 'PC001'} or {STUDYID eq 'PC002'} )")
Cohort
```

ID	TITLE	SEARCHCRITERIA	CREATEDON	UPDATEDON	FIRSTNAME	LASTNAME	DESCRIPTION	STUDYIDS	DRAFTFLG	SEARCHTYPE	USERID
699	Search on	Search Criteria : A1 AND (C1 OR D1) A1 = AGE ge 40 C1 = STUDYID eq 'PC001' D1 = STUDYID eq 'PC002'	2019-03-23	2019-03-23	Raja	Ramesh			1	3	4

Summaries

Here is some explanation on stats/methods used

- Fist item
- Second item
- Next **

Standards and Model Management

Model Management

Production
Ramesh, Raja
23-MAR-2019 15:13:00 (UTC +05:30)

Powered by Orchestra

List of Models

SDTM
ADaM
TA
Custom
Import History
Notifications

Create Mo...
Copy Model
Edit Properti...
Del...
Refre...
Diff Report
Blank Templ...
Export
Audit Trail Report

	Source M...	IG Model Name	Model Type	Model Name	Model Version	Model CT Name	CT Refer
☐	Custom		Custom	Small Molecule	0.00001		
☐	Custom		Custom	ADaM-SDM	0.00001		
☒	SDTM	SDTMIG 3.2	SDTM	SDTMIG-3.2	4.00000		
☒	SDTM		SDTM	TumorXPT	3.00000		
☒	SDTM		SDTM	SDTM-1.6	4.00000		
☒	Custom		Custom	Master Study Data Model	1.00000R		
☐	SEND		Custom	Nonclinical SDM	0.00001		
☐	Custom		Custom	Protein Immunoassay	0.00001		
☐	Custom		Custom	Protein IHC	0.00001		
☐	Custom		Custom	RNA	0.00001		
☐	Custom		Custom	DNA	0.00001		
☐	SDTM		Custom	SDTM+	0.00001		
☒	SDTM		SDTM	SDTM-1.4	6.00000		
☒	SDTM	SDTMIG 3.1.3	SDTM	SDTMIG-3.1.3	5.00000		
☒	SDTM	SDTMIG 3.1.2	SDTM	SDTMIG-3.1.2	5.00000		
☒	SDTM		SDTM	SDTM-1.5	5.00000		
☒	SDTM		SDTM	SDTM-1.3	5.00000		
☒	SDTM		SDTM	SDTM-1.2	5.00000		
☒	ADaM		ADaM	ADaMIG-1.0	5.00000		
☒	ADaM		ADaM	ADaM-2.1	5.00000		

PointCross Life Sciences

15

POINT CROSS

Putting Knowledge to Work

© 2011 PointCross. All Rights Reserved.

Standards and Model Management

Model Management

Production | Ramesh, Raja | 23-MAR-2019 15:13:00 (UTC +05:30)



List of Models

SDTM

ADaM

TA

Custom

Import History

Notifications

List of Models

ADaMIG-1.0 : 5.00000

Summary

Dataset Level

Variable Level

Controlled Terminology

Value Level Metadata

External Dictionaries

Computational Algorithm

Refre...

☐	Order Number	Domain*	Variable Name*	Variable Label	Core	Data Type	Synonyms	Value
☐	2	ADSL	USUBJID	Unique Subject Identifier	Required	Text		
☐	3	ADSL	SUBJID	Subject Identifier for the Study	Required	Text		
☐	4	ADSL	SITEID	Study Site Identifier	Required	Text		
☐	5	ADSL	SITEGRY	Pooled Site Group y	Permissible	Text		
☐	6	ADSL	SITEGRYN	Pooled Site Group y (N)	Permissible	Integer		
☐	7	ADSL	AGE	Age	Required	Integer		
☐	8	ADSL	AGEU	Age Units	Required	Text		
☐	9	ADSL	SEX	Sex	Required	Text		
☐	10	ADSL	RACE	Race	Required	Text		
☐	11	ADSL	RACEGRY	Pooled Race Group y	Permissible	Text		
☐	12	ADSL	RACEGRYN	Pooled Race Group y (N)	Permissible	Integer		
☐	13	ADSL	FASFL	Full Analysis Set Population Flag		Text		
☐	14	ADSL	SAFFL	Safety Population Flag		Text		
☐	15	ADSL	ITTFL	Intent-To-Treat Population Flag		Text		
☐	16	ADSL	PPROTFL	Per-Protocol Population Flag		Text		
☐	17	ADSL	COMPLFL	Completers Population Flag		Text		
☐	18	ADSL	RANDFL	Randomized Population Flag		Text		
☐	19	ADSL	ENRFL	Enrolled Population Flag		Text		
☐	20	ADSL	ARM	Description of Planned Arm	Required	Text		
☐	21	ADSL	TRTXXP	Planned Treatment for Period xx	Permissible	Text		

Display a menu

Total Count : 205 | Filtered Records : 63

1 Record(s) Selected

Finding Actionable Results from Disparate Data Sources

Cohort Stratification using cascading query masks

Combine Query Results with Boolean logic to find Highly Stratified Cohorts, Specific targeted cases, and clinical results to Generate analysis data sets

