ORACLE®

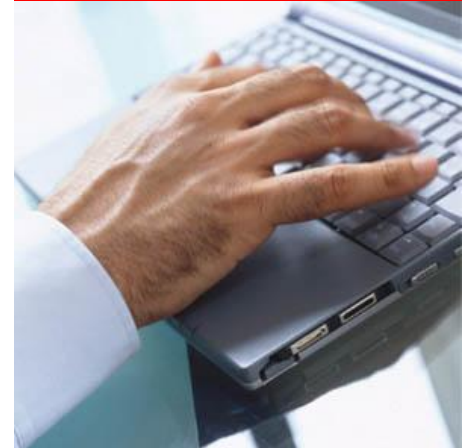# The Evolution Of Clinical Data Warehousing

Srinivas Karri
Principal Consultant

Agenda

- Value of Clinical Data

- Clinical Data warehousing &

  The Big Data Challenge

- The Impact of Omics Data

- The Future

- Questions

ORACLE®

Value of Clinical Data

Value of Clinical Data

- 273

Value of the Opportunity

- 273Billion USD / yr realised though improved clinical decision support
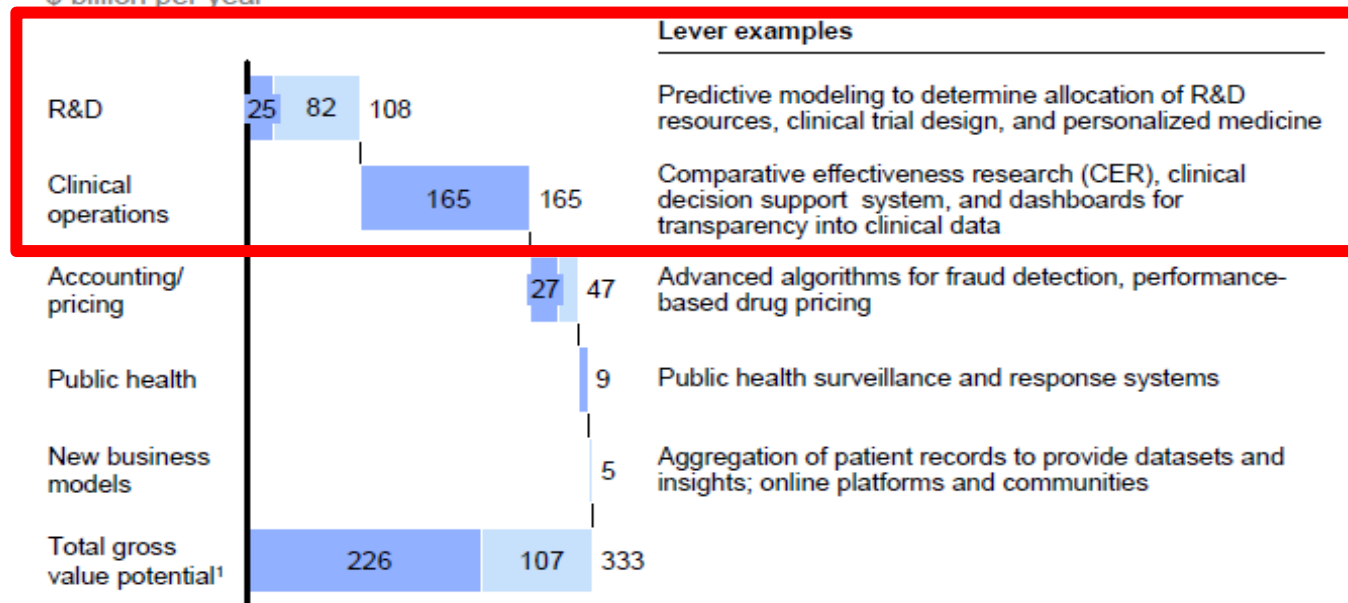
ORACLE®

Value of the Opportunity

- 273Billion / yr realised though improved clinical decision support

The estimated long-term value of identified levers is more than $300 billion, with potentially more than $200 billion savings on national health care spending

Value potential from use of big data
$ billion per year

■ Direct reduction on national health care expenditure
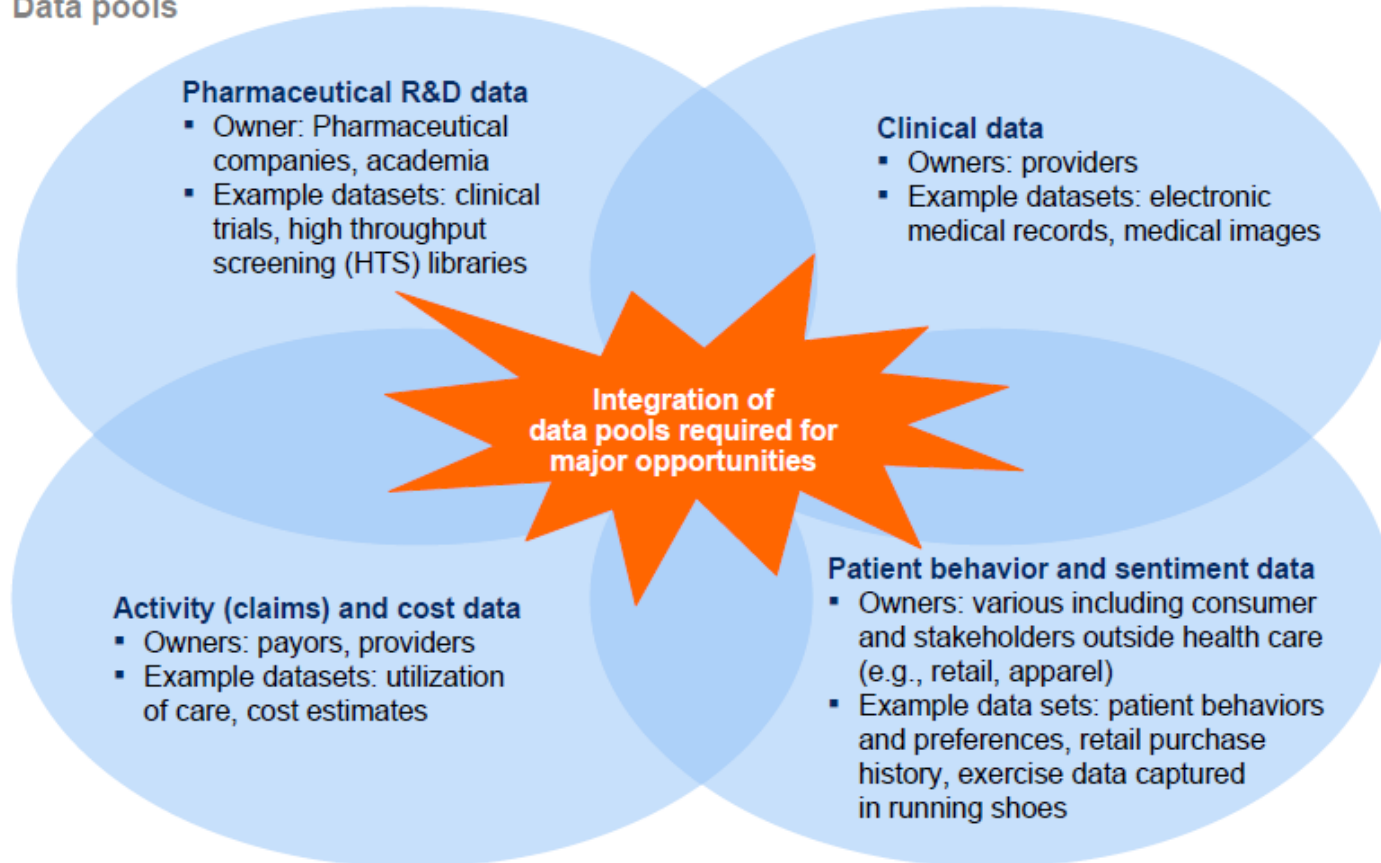□ Unclear impact on national health care expenditure

				Lever examples
R&D	25	82	108	Predictive modeling to determine allocation of R&D resources, clinical trial design, and personalized medicine
Clinical operations		165	165	Comparative effectiveness research (CER), clinical decision support system, and dashboards for transparency into clinical data
Accounting/ pricing		27	47	Advanced algorithms for fraud detection, performance-based drug pricing
Public health			9	Public health surveillance and response systems
New business models			5	Aggregation of patient records to provide datasets and insights; online platforms and communities
Total gross value potential[1]	226	107	333	

1 Excluding initial IT investments (~$120 billion–$200 billion) and annual operating costs (~$20 billion per annum).
SOURCE: Expert interviews; press and literature search; McKinsey Global Institute analysis
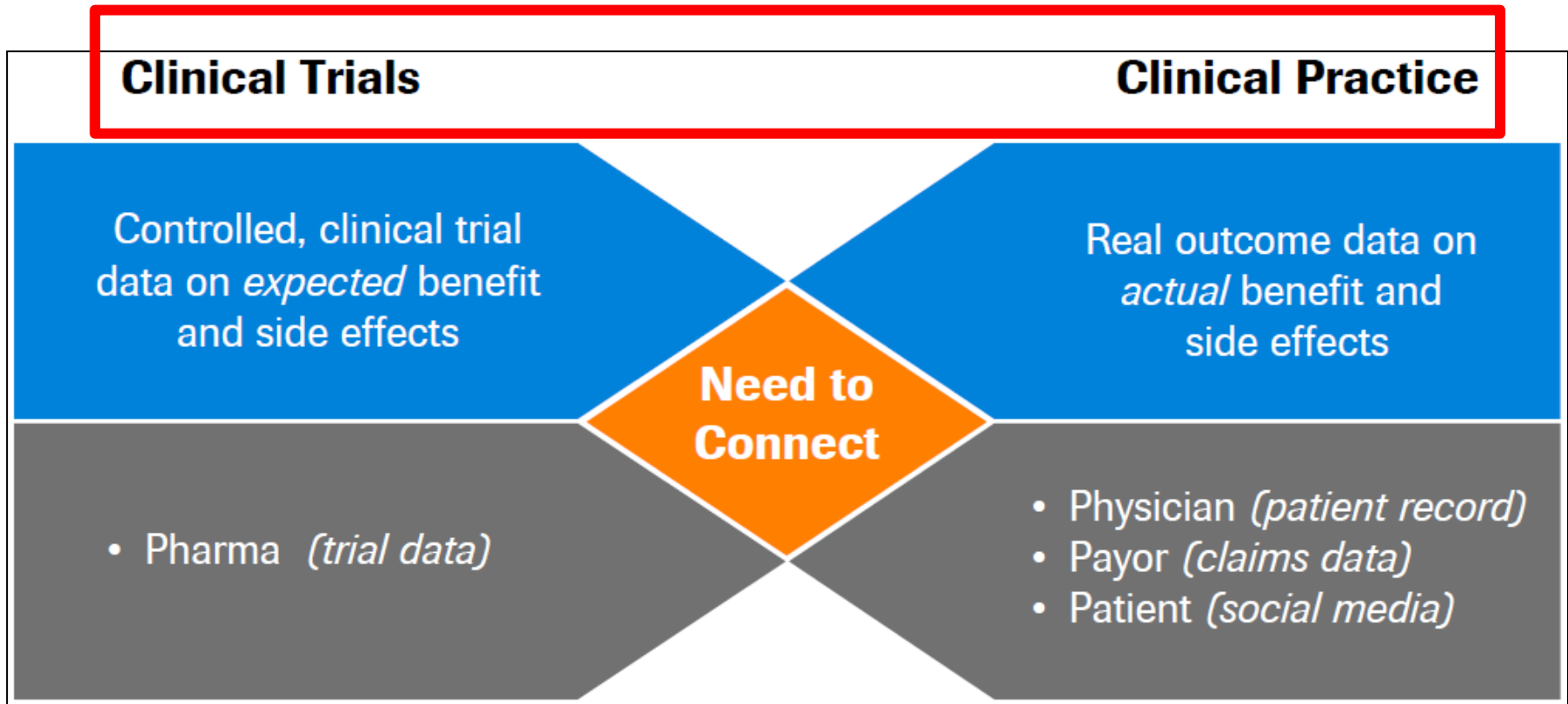
ORACLE

Integration Across the Data Pools

Four distinct big data pools exist in the US health care domain today with little overlap in ownership and low integration

Data pools

Pharmaceutical R&D data
- Owner: Pharmaceutical companies, academia
- Example datasets: clinical trials, high throughput screening (HTS) libraries

Clinical data
- Owners: providers
- Example datasets: electronic medical records, medical images

Integration of data pools required for major opportunities

Activity (claims) and cost data
- Owners: payors, providers
- Example datasets: utilization of care, cost estimates

Patient behavior and sentiment data
- Owners: various including consumer and stakeholders outside health care (e.g., retail, apparel)
- Example data sets: patient behaviors and preferences, retail purchase history, exercise data captured in running shoes

SOURCE: McKinsey Global Institute analysis

Connecting in Practice

Clinical Trials   **Clinical Practice**

Controlled, clinical trial data on *expected* benefit and side effects

Real outcome data on *actual* benefit and side effects

Need to Connect

- Pharma *(trial data)*
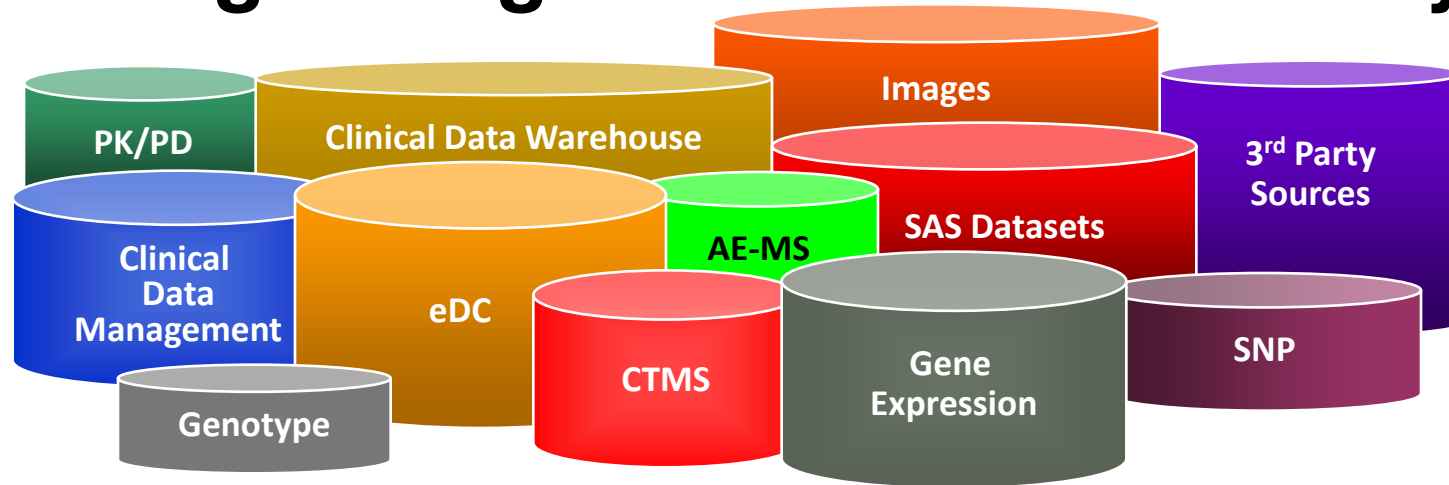
- Physician *(patient record)*
- Payor *(claims data)*
- Patient *(social media)*

"At the moment data from clinical trials and data from clinical practice are separate entities, but it makes sense to bring these together."

– Severin Schwan, CEO, Roche

ORACLE®

Clinical Data Warehousing

Having the right data available is key

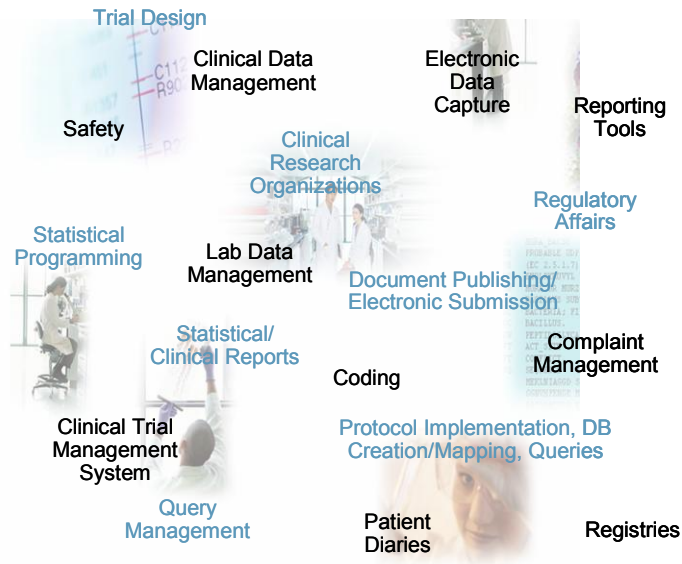**Clinical data issues are a significant contributor**

- Most companies hold large volumes of data that fuel your development projects, however different databases cover different domains of information and knowledge.

- Inefficient "hand crafting" of data aggregates for an alternate use such as data mining

- Diverse, inconsistent workflows and decision support processes

- Access to all information is critical if you are to be able to identify safety "signals" and manage patient risk.

- No easy way to classify and consolidate data with contextual information

- Access to legacy systems is often limited, although data are still required "on-line"
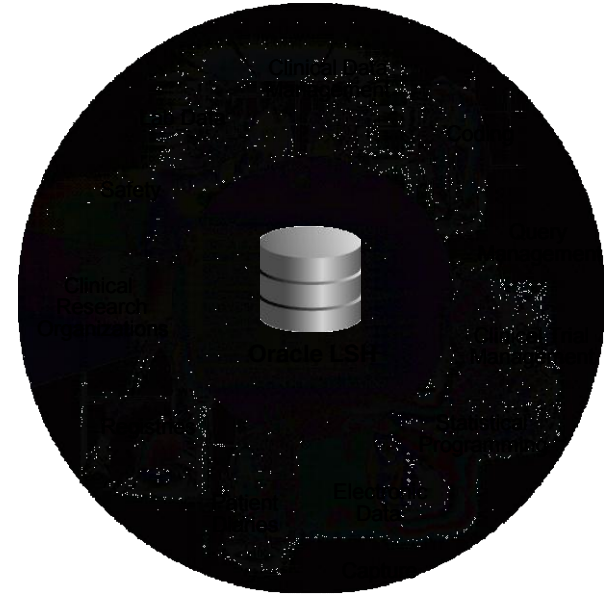
Current Approach

- Linkages between applications are often custom built, point-to-point, solutions that are hard to maintain – the spaghetti effect!

- Legacy data sometimes held in "near offline" archive systems, but context very often lost

ORACLE®

The Vision

From

Trial Design

Clinical Data Management

Electronic Data Capture

Reporting Tools

Safety

Clinical Research Organizations

Regulatory Affairs

Statistical Programming

Lab Data Management

Document Publishing/ Electronic Submission

Statistical/ Clinical Reports

Complaint Management

Coding

Clinical Trial Management System

Protocol Implementation, DB Creation/Mapping, Queries

Query Management

Patient Diaries

Registries

To

Oracle LSH

Multiple & redundant systems perform their focused job well. Data integration and cross system analysis is painful

A data integration hub, that gets data under control, increases availability and improves efficiency
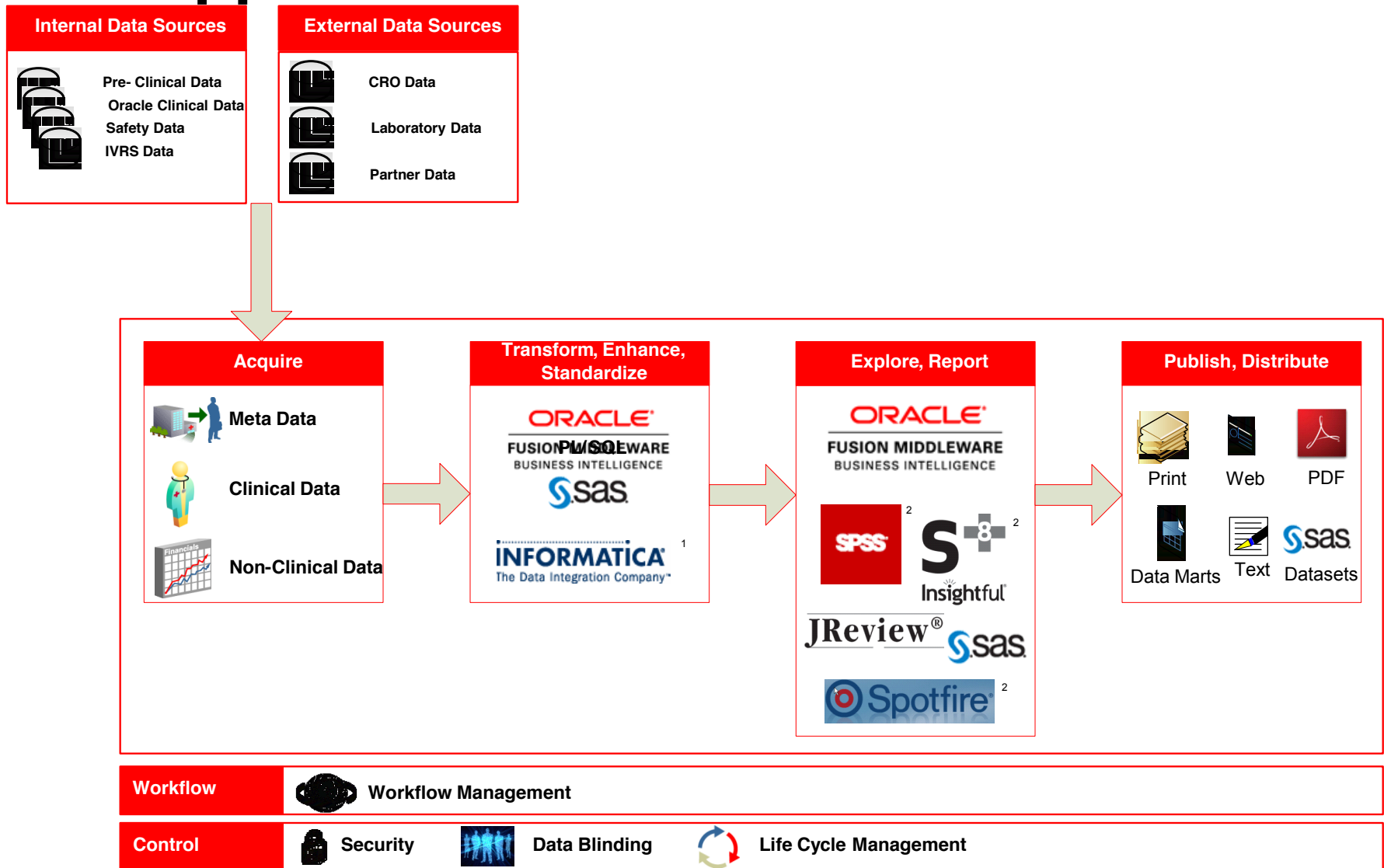
2004 – Aspiration

- Streamline clinical development information management
- Reduce risk
- Simplify clinical integration
- Flexible and extensible architecture
- Rapidly embrace new standards
- Controlled application development
- Empower clinical researchers to make better decisions
- Achieve a single view of clinical data
- Improve productivity with self-service analysis and reporting
- Automate complex processes and approval chains
- Work with regulators to optimise approval processes
- Support ongoing regulatory reviews

2012 - Reality

- Streamline clinical development information management
- Reduce risk
- Simplify clinical integration
- Flexible and extensible architecture
- Rapidly embrace new standards
- Controlled application development
- Empower clinical researchers to make better decisions
- Achieve a single view of clinical data
- Improve productivity with self-service analysis and reporting
- Automate complex processes and approval chains
- Work with regulators to optimise approval processes
- Support ongoing regulatory reviews

ORACLE®

Supported Business Processes

Internal Data Sources

Pre- Clinical Data
Oracle Clinical Data
Safety Data
IVRS Data

External Data Sources

CRO Data

Laboratory Data

Partner Data

Acquire	Transform, Enhance, Standardize	Explore, Report	Publish, Distribute
Meta Data	ORACLE FUSION MIDDLEWARE PL/SQL BUSINESS INTELLIGENCE SAS	ORACLE FUSION MIDDLEWARE BUSINESS INTELLIGENCE	Print Web PDF
Clinical Data		SPSS [2] S+8 [2] Insightful	Data Marts Text Datasets
Non-Clinical Data	INFORMATICA [1] The Data Integration Company	JReview® SAS	
		Spotfire [2]	

Workflow	Workflow Management

Control	Security	Data Blinding	Life Cycle Management

Biostats

Supported <u>Sample</u> Use Cases

- **Core**
    - Run analysis programs program
    - Create analysis outputs
    - Create integrated summaries from SDTM pools
    - Distribute outputs

    - Develop SAS programs
    - Execute SAS programs
    - Schedule SAS programs
    - Copy SAS programs from a library

    - Split SAS datasets to view
    - View & Edit SAS dataset
    - Application bookmarks
    - Export SAS to Excel and data exchange
    - Parameterized macros
    - Meta data search

- **Options**
    - Hierarchy
    - Macro Libraries
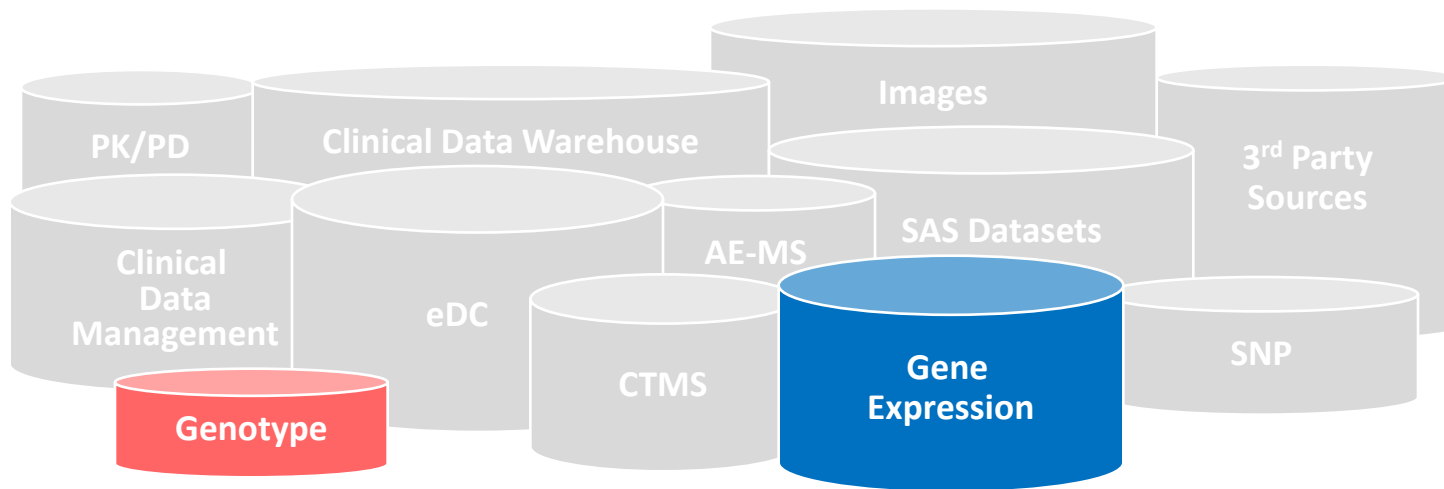
- **Datamarts and Outputs**
    - Create/grant access
    - Allow application to access data

State of the Art
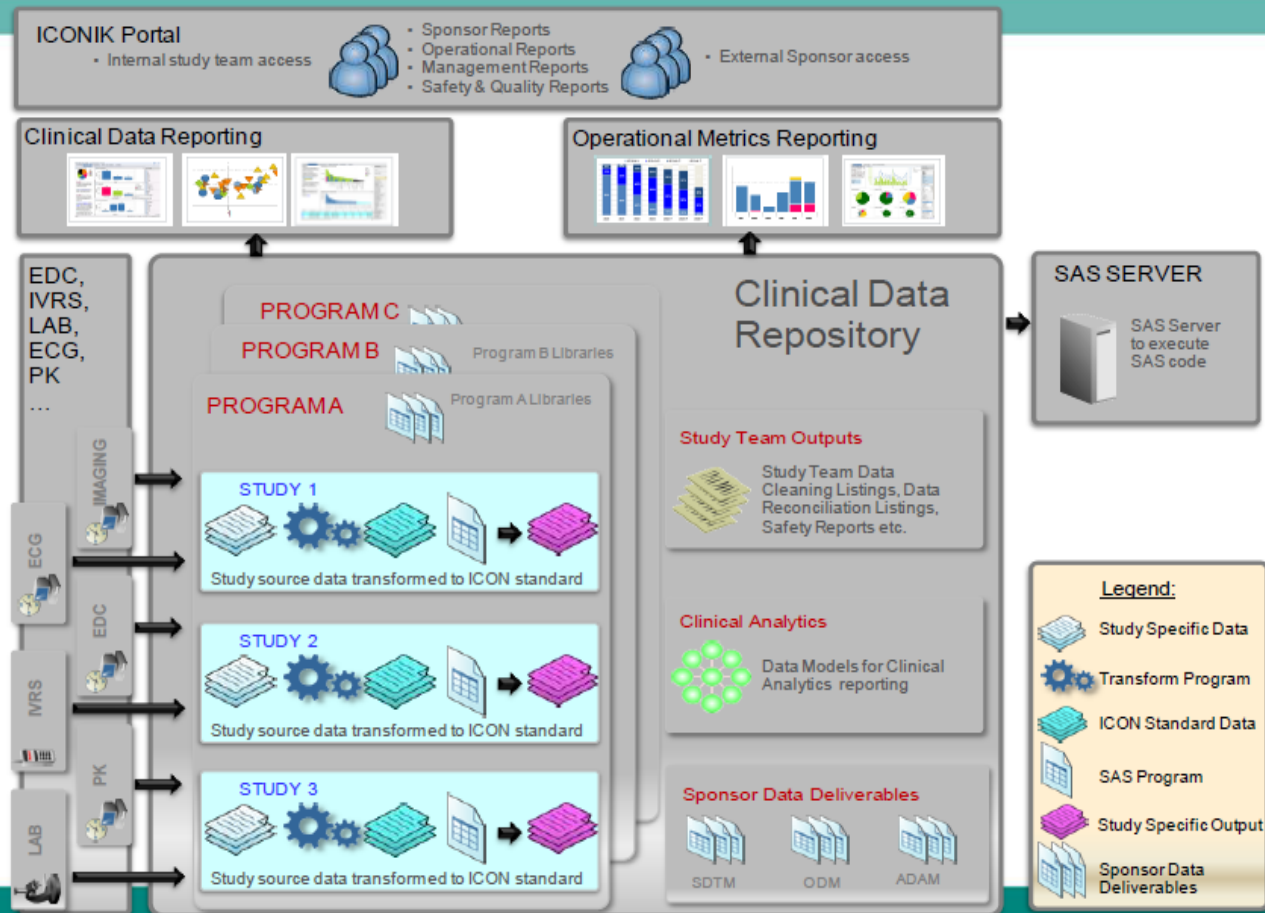
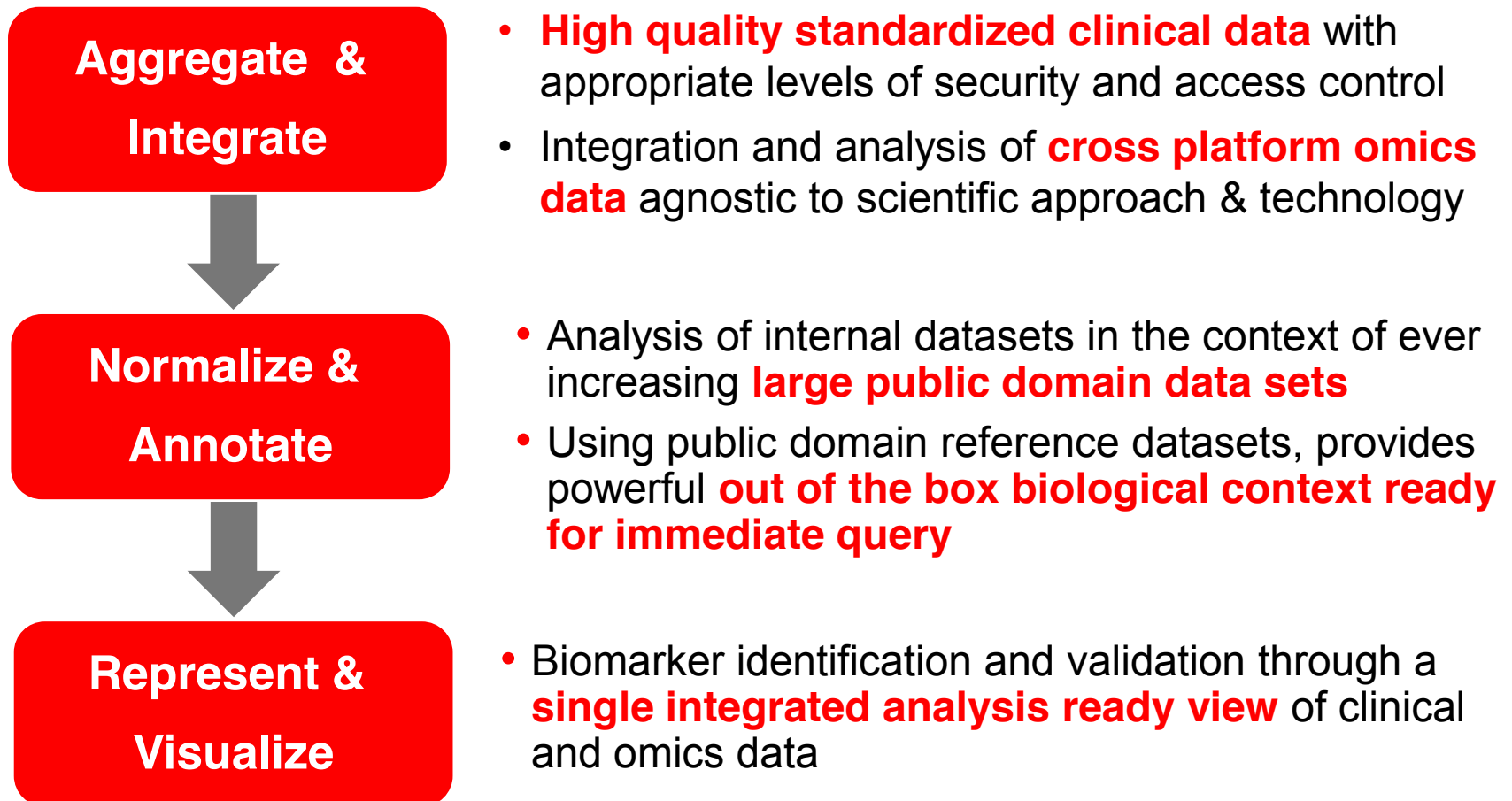# But Wait...What about Genomic Data?

- Very Large Data Sets

- New Data Models

- New Visualization Tools

- New Business Processes

ORACLE®

The Impact of Omics Data

New Business Processes

Aggregate & Integrate

- **High quality standardized clinical data** with appropriate levels of security and access control
- Integration and analysis of **cross platform omics data** agnostic to scientific approach & technology

Normalize & Annotate

- Analysis of internal datasets in the context of ever increasing **large public domain data sets**
- Using public domain reference datasets, provides powerful **out of the box biological context ready for immediate query**

Represent & Visualize

- Biomarker identification and validation through a **single integrated analysis ready view** of clinical and omics data

ORACLE®

New Data Models - Healthcare

Integrated View of Clinical, Financial, Operational, & Research Data Across the Provider Enterprise

Healthcare Data Warehouse Foundation
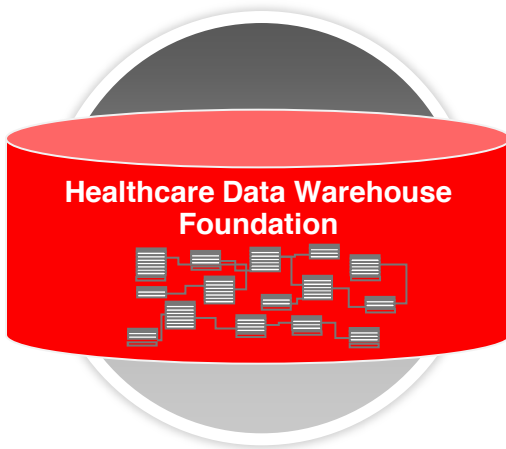
Clinical Domain

Patient

Encounter

Concern

Intervention

Case

Order

Specimen

Observation

Incident

History

Pharmacy

Care Site Scheduling and Utilization

Accounts Payable & Purchasing Domain

Purchasing

Accounts Payable
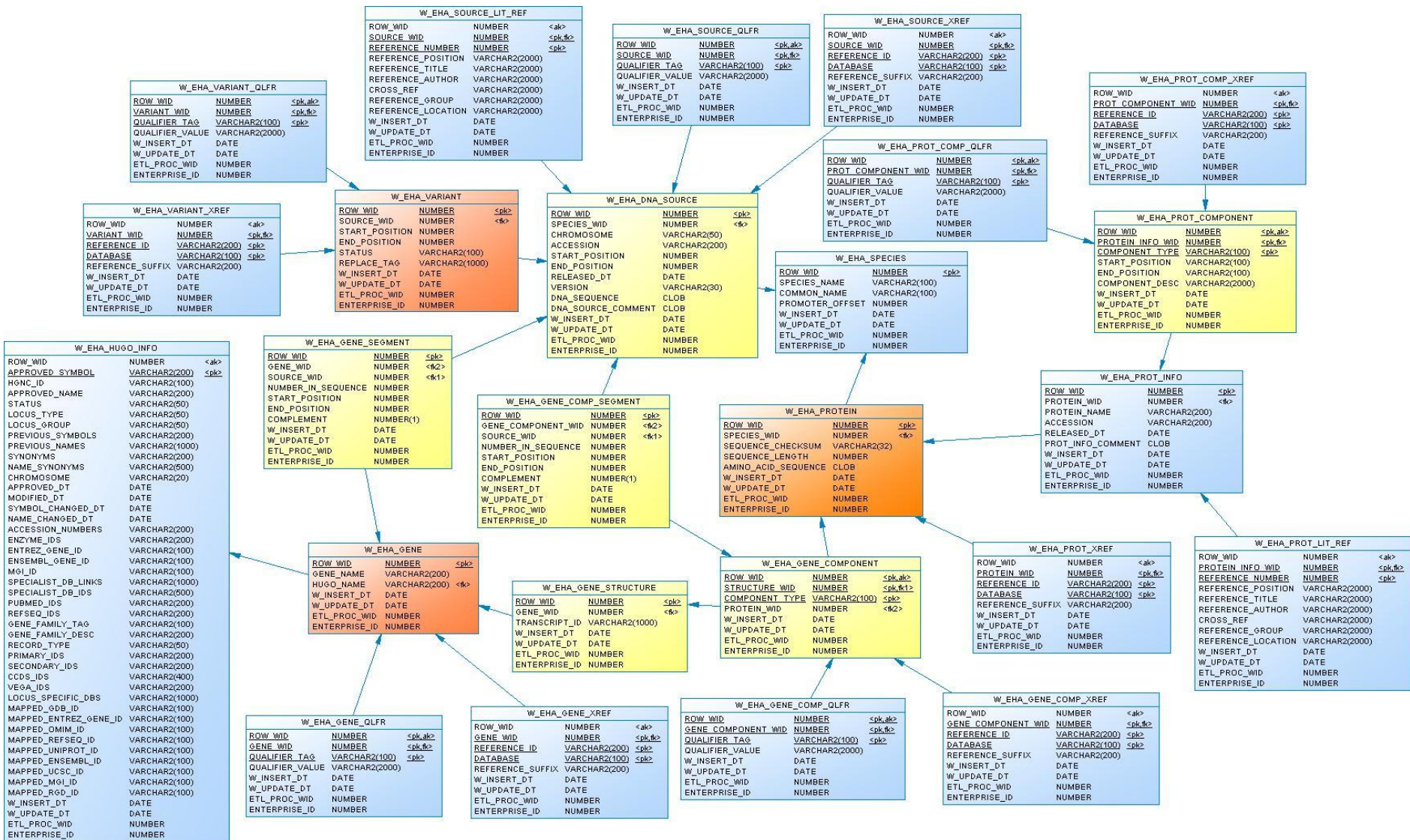
Inventory Domain

Inventory

Administrative Domain

Cost Allocation

HR and Payroll

Survey

Advance Directive
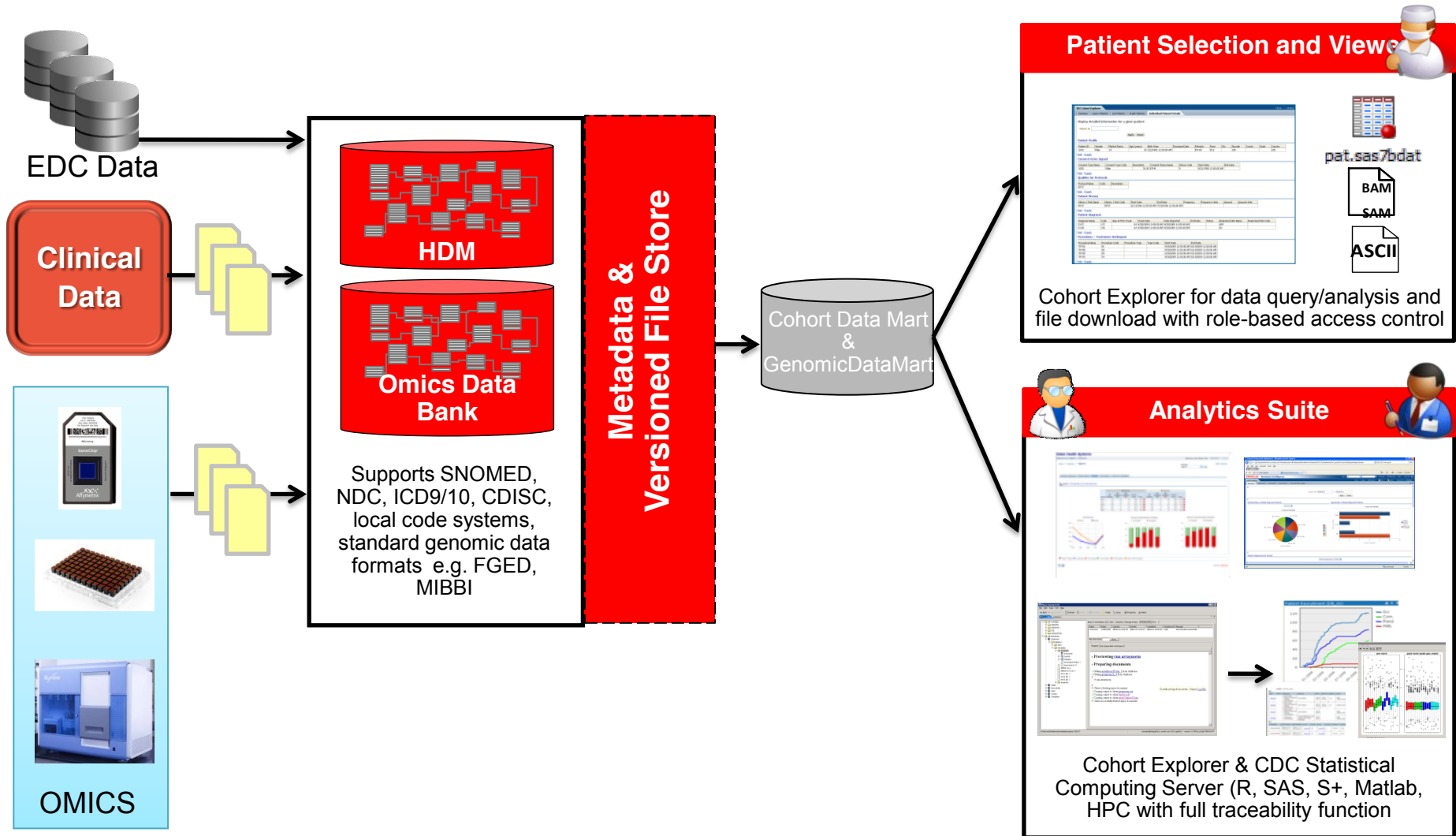
Consent

Party

Roles

Facility

Master Catalog

Accounting, Finance, and Revenue Cycle Domain

Accounting and Financial

Patient Financial Services
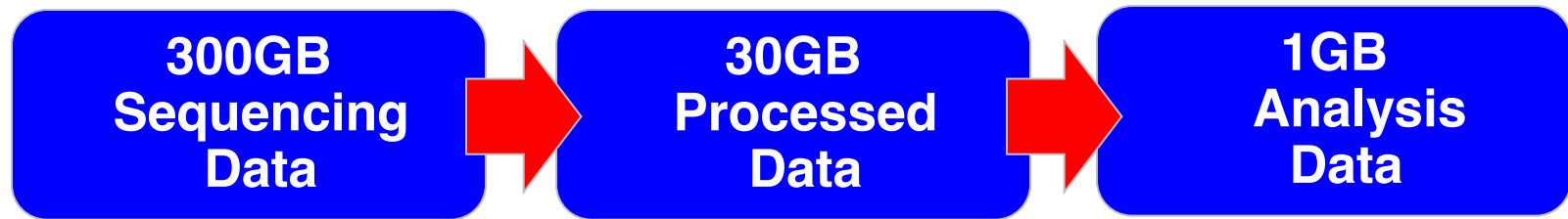
Billing

Claims

Charge Master

ORACLE
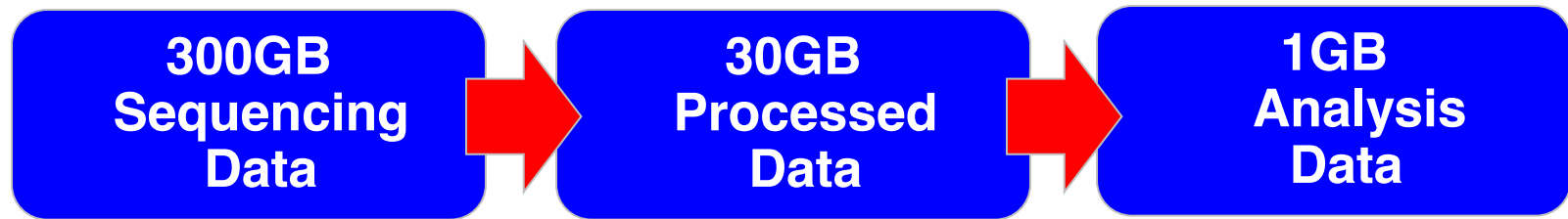
New Data Models – Omics Data Bank

Putting the Pieces Together

EDC Data

Clinical Data

OMICS

HDM

Omics Data Bank

Supports SNOMED, NDC, ICD9/10, CDISC, local code systems, standard genomic data formats e.g. FGED, MIBBI

Metadata & Versioned File Store

Cohort Data Mart & GenomicDataMart

Patient Selection and Viewer

pat.sas7bdat

BAM

SAM

ASCII

Cohort Explorer for data query/analysis and file download with role-based access control

Analytics Suite

Cohort Explorer & CDC Statistical Computing Server (R, SAS, S+, Matlab, HPC with full traceability function

ORACLE®

Big Data Challenges

300GB Sequencing Data	→	30GB Processed Data	→	1GB Analysis Data

- Per sequence, > 1
- Per patient
- Per population, >100,000

Big Data Challenges

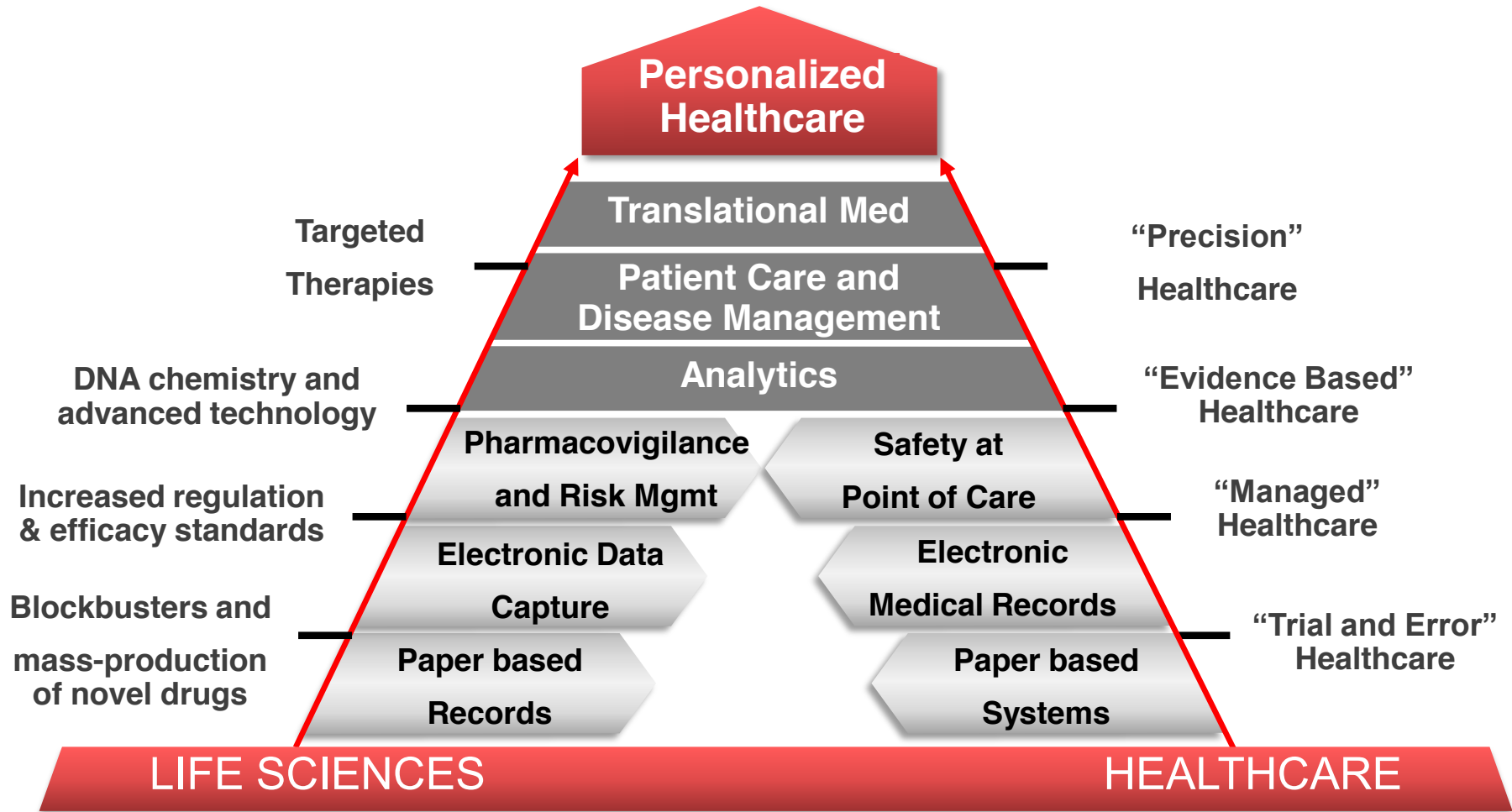| 300GB Sequencing Data	→	30GB Processed Data	→	1GB Analysis Data

- Per sequence, > 1

- Per patient

- Per population, >100,000

- **30 Petabytes of data**

- Google processes approx 24petabyes /day

ORACLE®

The Future