



Linking People,
Potential and Progress

HimSS[®]12
ANNUAL CONFERENCE & EXHIBITION

Implementing a Healthcare Data Warehouse in One Year (or Less)

DISCLAIMER: The views and opinions expressed in this presentation are those of the author and do not necessarily represent official policy or position of HIMSS.



Conflict of Interest Disclosure

Richard E. Biehl, Ph.D.

- Ownership Interest: Richard is the sole proprietor of Data-Oriented Quality Solutions (DOQS), an IT/Quality consulting practice founded in 1988 and operating out of Orlando, Florida, USA.
- Consulting Fees: Richard earns approximately 15% of his income from consulting engagements that involve the heuristics included in this presentation.
- Other: The heuristics in this presentation can be immediately and directly implemented by attendees. Nothing has been held back that would necessitate engaging DOQS for implementation.



Learning Objectives

1. Assemble a small agile team for rapid and focused warehouse implementation
2. Identify how to eliminate database and functional dependency project bottlenecks
3. Plan a multi-iteration project approach that delivers functionality quarterly
4. Diagnose organizational and resource issues that might stand in the way
5. Select the most powerful clinical data sources for the initial implementation



Benefits Typically Sought

- Integrated data repository
- New and diverse access paths
- Enhanced data quality
- Early access to complicated data
- Scalable to include more sources
- Semantic pathways into data
- Longitudinal phenotypic data



Teaming & Responsibilities

Data
Governance
Group

PROJECT
TEAM

Support
Team



Project Team (3-5 FTE)

- Data Warehouse Architect (.5 FTE)
- Business/Data Analysts (1-2 FTE)
- ETL Developers (1-2 FTE)
- Application Developer (.5 FTE)



Support Team (1.5 FTE)

- Project Management (.25 FTE)
- Data Base Administrator (.25 FTE)
- Data Administrators (.25 FTE)
- HL7 Integration Specialist (.25 FTE)
- Business Intelligence Designer (.25 FTE)
- Organizational Change Agent (.25 FTE)



Data Governance Group

- Promote the voice of the customer
- Ensure timely addressing of issues
- Advise owners, stewards, & users
- Enhance the quality of data
- Promote innovative uses for data



Eliminate Bottlenecks

Analysis

Data
Queries
Flow
Controls
People
Goals

Database Design

Modeling
Storage
Integration
Keys
Mapping
Indexes
Partitioning

ETL Design & Build

Maps
Extraction
Controls
Duplicates
Missing flows
Validation
Traceability

Query Design

Security
Privacy
Queries
Usage
Controls



Eliminate Bottlenecks

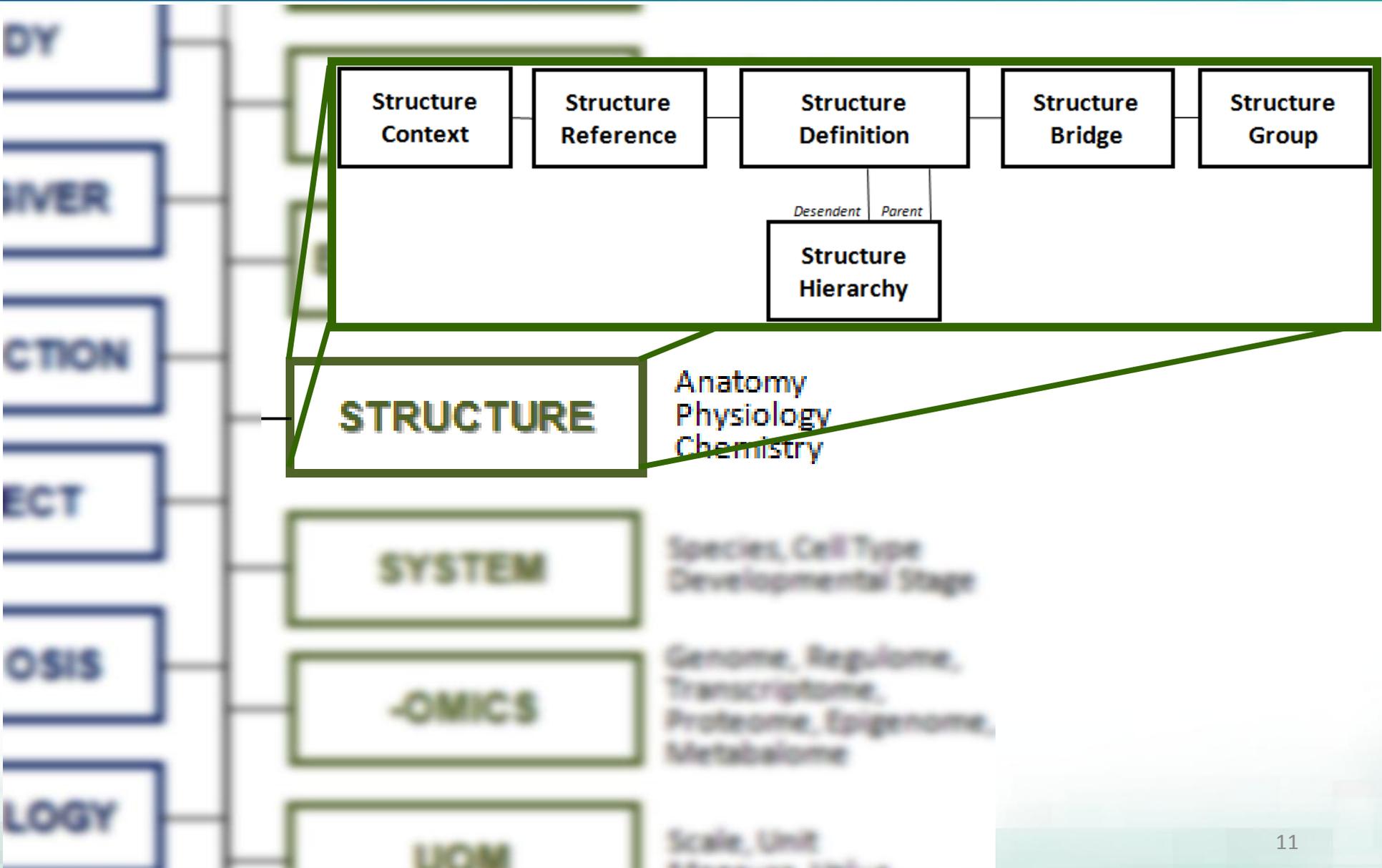
Analysis

Database Design

ETL Design & Build

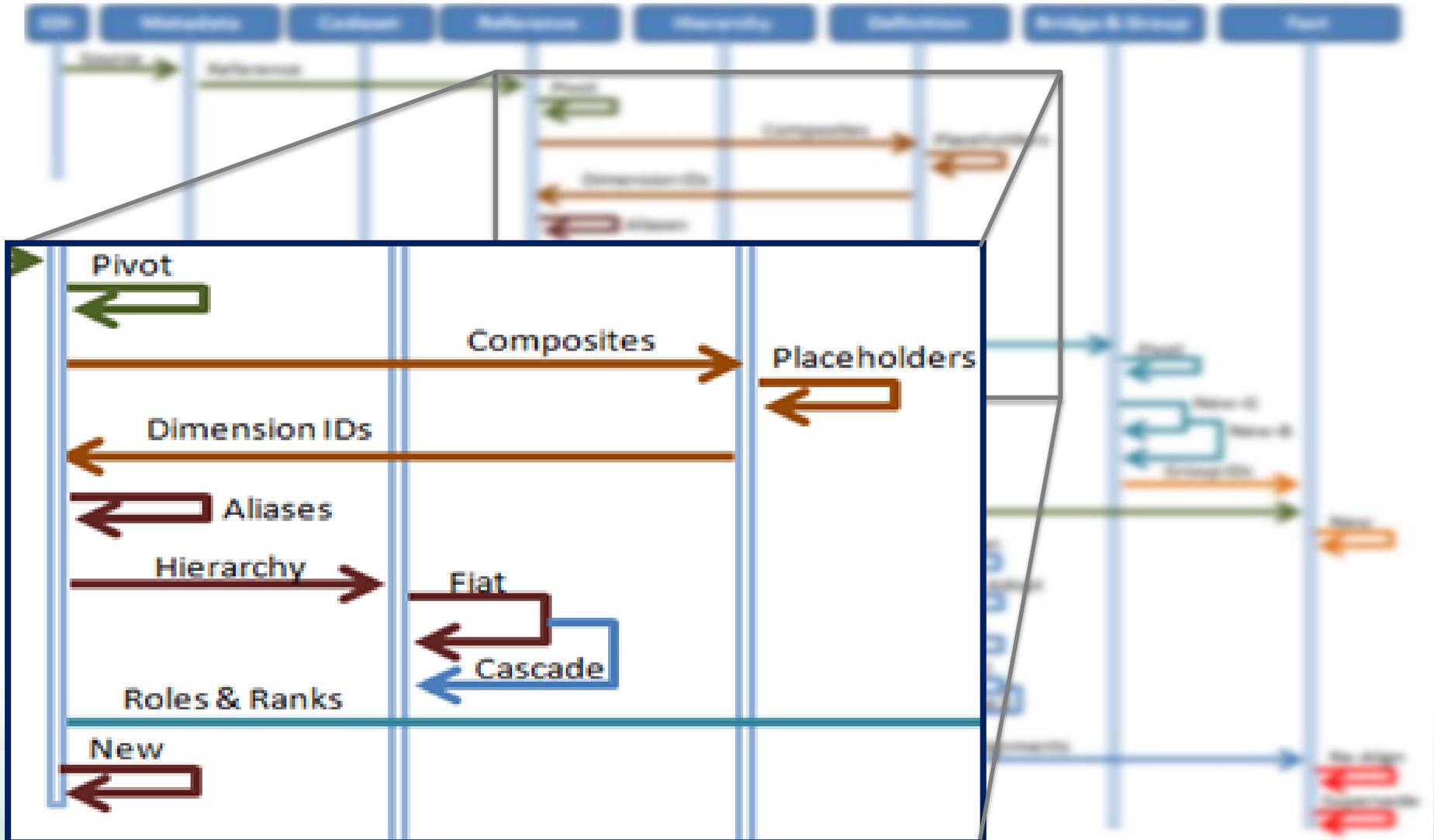
Query Design

Strong Functional Dependencies



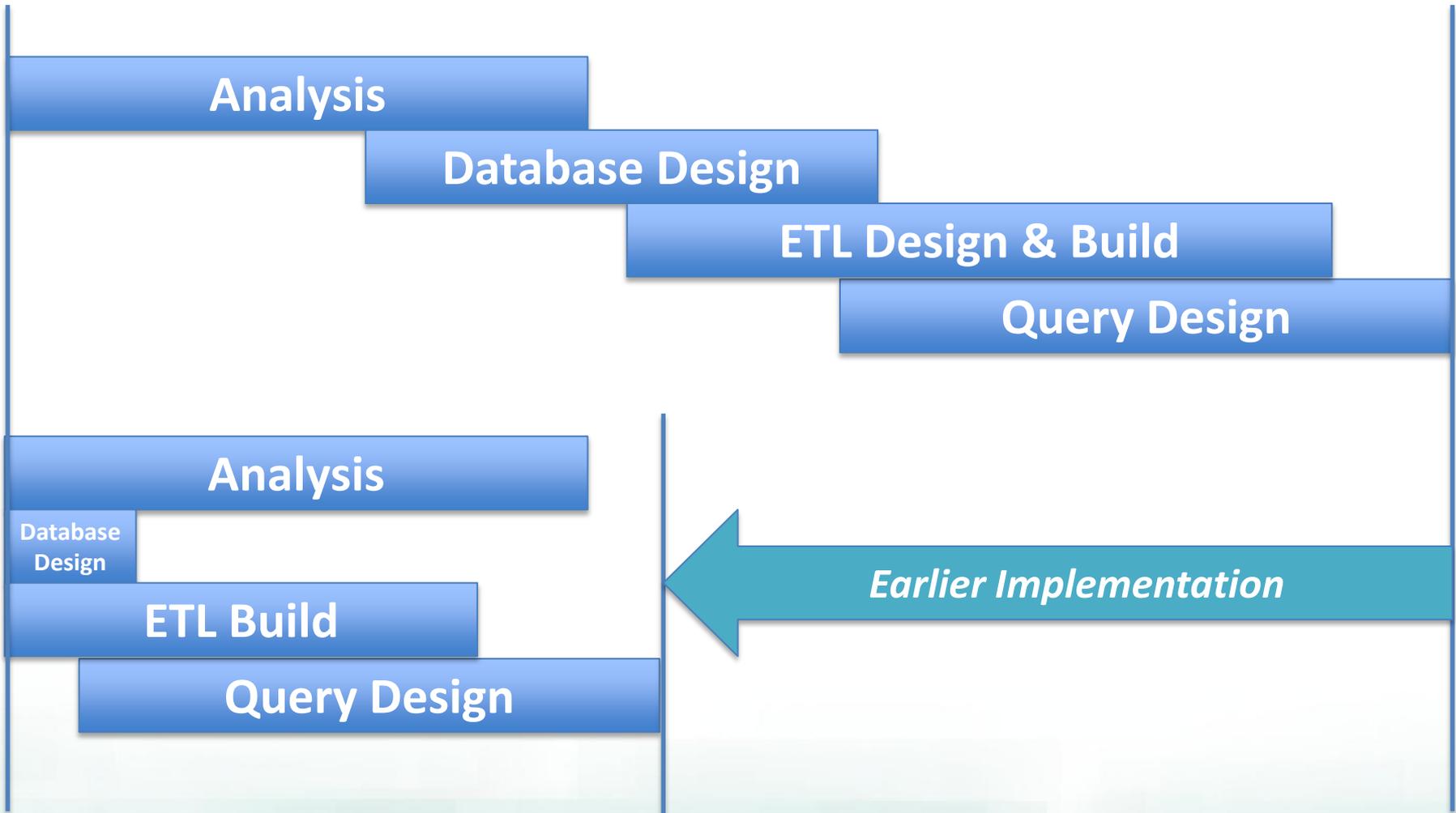


Standardized ETL Design





Eliminate Bottlenecks





Development Iterations

- **Alpha Version** (end of January)
 - **Proof-of-Concept** (1st week)
- **Beta Version** (mid-March)
- **Gamma Version** (early September)
- **Release 1.0** (mid-December)



ALPHA VERSION

- Implements *just* enough to illustrate core capabilities
- Loads *small* sampling of *limited* data
- Enables a way to visualize results *by actually looking at a core subset*
- Allow **3-4 weeks** for development



BETA VERSION

- Most of the required functionality
- A lot more data
- Facilitates planning & setup
- *A show-and-tell* tool for the team, not a self-service tool for users
- Ready **about 2-3 months** after the Alpha version is complete



GAMMA VERSION

- A functionally complete system
 - typically still lacking much of the data that would need to be loaded in order to consider it production-ready

Key issue: Finalizing HIPAA controls

- Typically ready **3-4 months** after the Beta version is completed

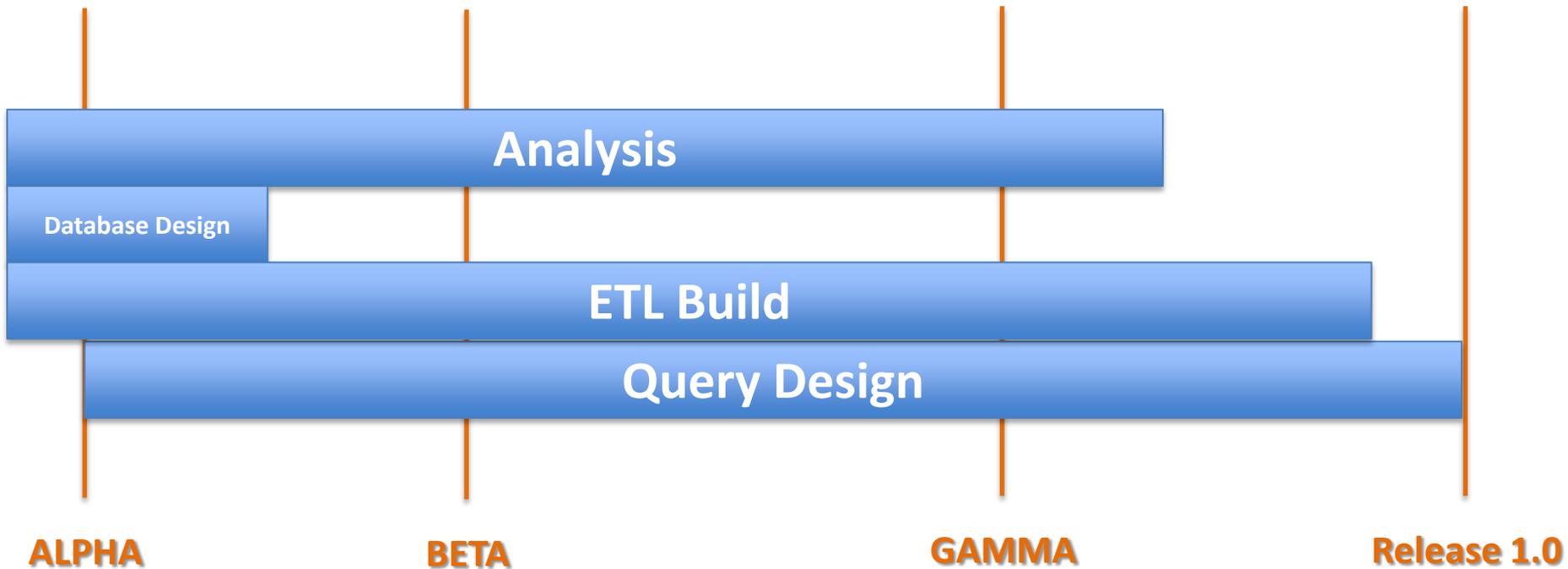


RELEASE 1.0

- Placed into production as a fully functional and secure data warehouse
 - all first release scoped data loaded up to the implementation date, and
 - an operating ETL environment that will keep that scoped data current
- Launch is **about a year** after the project started



Iterative Parallelism





Diagnose Issues

- Knowledge of source data
- History vs. prospective data
- Access to source data structures
- Data capture & conversion tools
- System time vs. clinical time
- Physical operating environment
- Query performance (memory, indexing)



Select Sources

- Admission-Discharge-Transfer (ADT)
- Diagnoses & Procedures
- Allergy & Problem Data
- Lab Orders & Results
- Medication Orders & Results



Sources to Avoid

- Large text or report sources
 - HIPAA issues
 - Natural language issues
- Overly complex data
 - Allow learning period
 - Anticipatory discussions only



Anticipate Pitfalls

- Difficulty in making the shortened timeframe real to stakeholders
- Extensive lead times on setup of operational environment
- Tendency to defer access and privacy discussions until too late in project
- Window of opportunity needs to be wide open across the organization



Questions?

You are welcome to contact me for additional information at any time:

Richard E. Biehl, Ph.D.

Data-Oriented Quality Solutions
rbiehl@doqs.com

Additional questions.....

**Clinical & Business
Intelligence**

Knowledge Center

Today, 11:45a - 12:15p

Booth 13247



UNIVERSITY OF CENTRAL FLORIDA

**DEPARTMENT OF
HEALTH MANAGEMENT AND INFORMATICS**

Booth 13657