



Data-centric System Development Life Cycle (SDLC) for Automated Clinical Data Development System

Kevin Lee, MarkLogic

© COPYRIGHT 2015 MARKLOGIC CORPORATION. ALL RIGHTS RESERVED.

Agenda

1. Data-centric thinking
2. Birth of Automated Clinical Data Development System(ACDDS)
3. Introduction of System Development Life Cycle (SDLC)
4. SDLC for Automated Clinical Data Development System
5. Data-centric SDLC
6. Consideration of ACDDS
7. Conclusion
8. Questions

Quiz 1?

1. World largest transportation company?



On December, 2015, there are about 162,037 “active drivers” completed at least four or more trips for the service. The number of new drivers signing up has doubled every six months for the past two years.?

According to 2012 “Bureau of Labor Statistics, 233,00 taxi drivers.

Quiz 2?

2. World largest accommodation provider?



- Inc.'s 2014 Company of the Year
- In 2014 alone, 10 millions used airbnb. This year, their website surpassed 800,000 listings worldwide, which means they now offer more lodging than Hilton Worldwide or InterContinental Hotels Group or any other hotel chain in the world.

Common characteristics of exponential organization

- Data

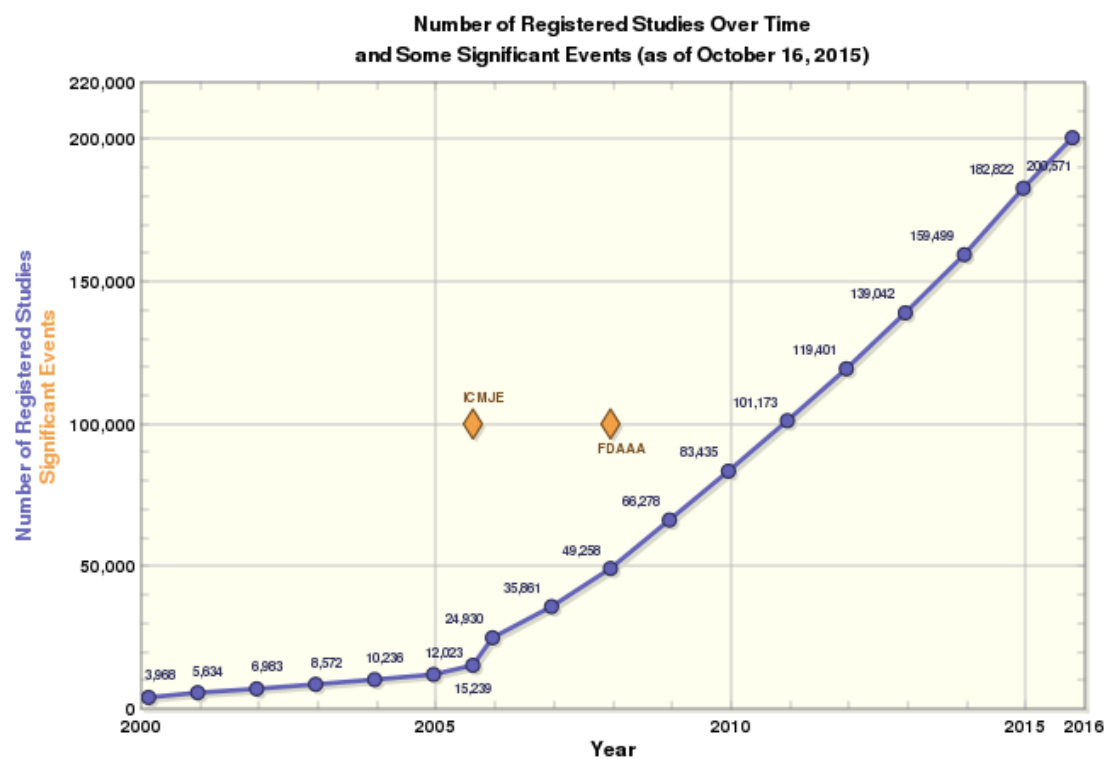


Exponential
& Scalable
Growth

- Interface/system to process data and exchange information



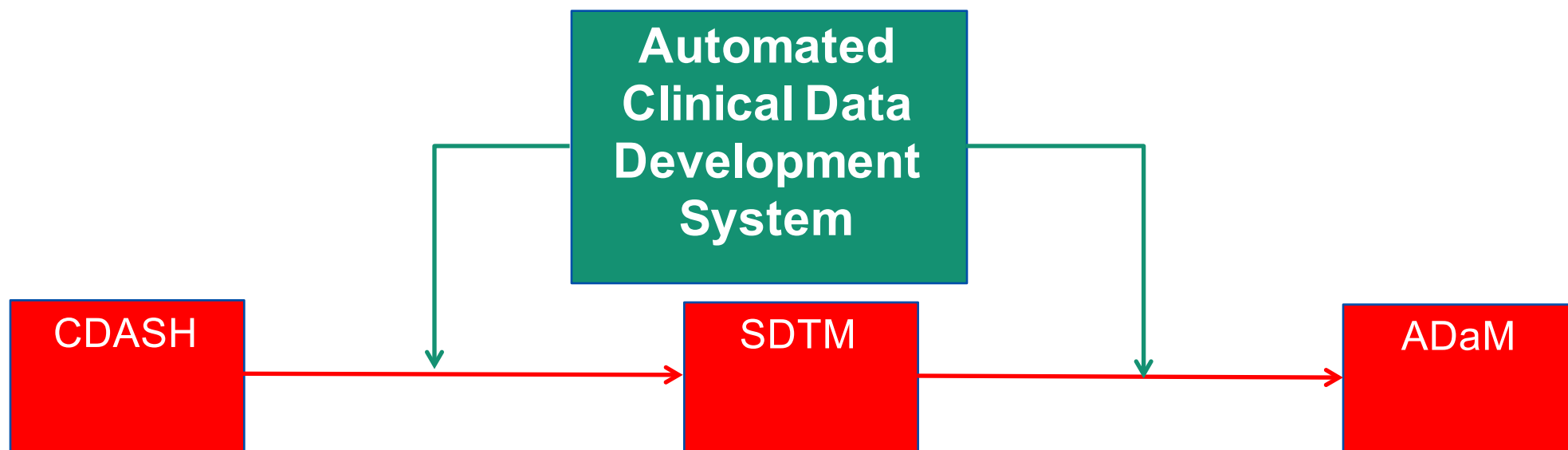
Increase in Clinical Trials (clinicaltrials.gov)



Source: <https://ClinicalTrials.gov>

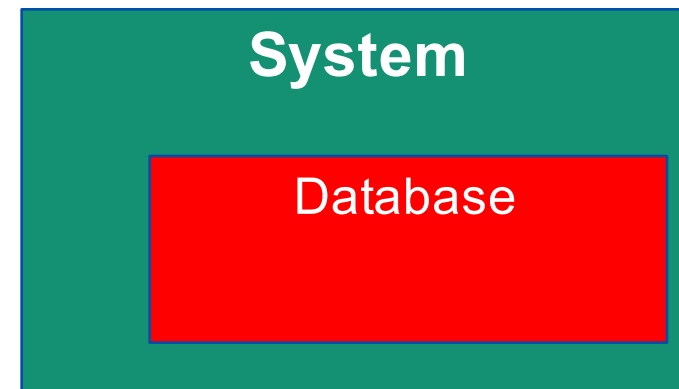
How to apply this data-centric thinking in clinical trial?

- End to End Clinical Data Development and Management



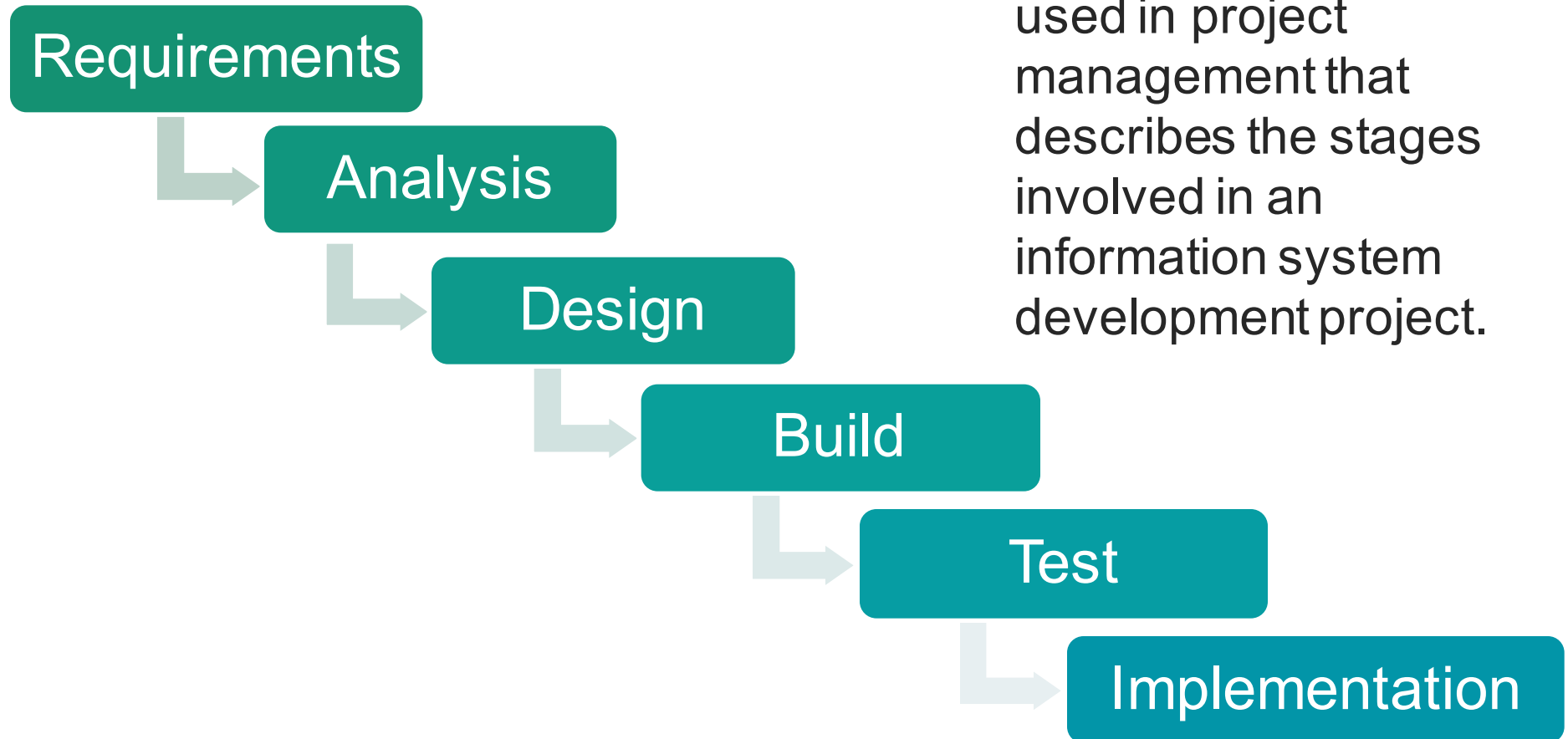
Typical system

- Process automation (e.g., Airport auto Check In)
- Data used in system resides inside of system.
- The structure of data does not change unless process changes.
- Process dictates data structure in database.



SDLC (System Development Life Cycle)

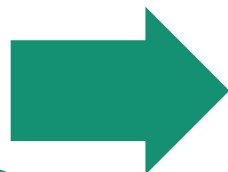
A conceptual model used in project management that describes the stages involved in an information system development project.



Shift of SDLC perspectives

Typical, process automating System

- SDLC
 - Automates process
 - Process dictates data
 - Data stays inside of system
 - Interaction b/w users and interface



Clinical Data Development System

- Data-centric SDLC
 - Automates data development
 - Data dictates process
 - Data coming into system and leaving the system
 - Transformation from source to target data



Data-centric SDLC?



Data-centric SDLC: Requirements

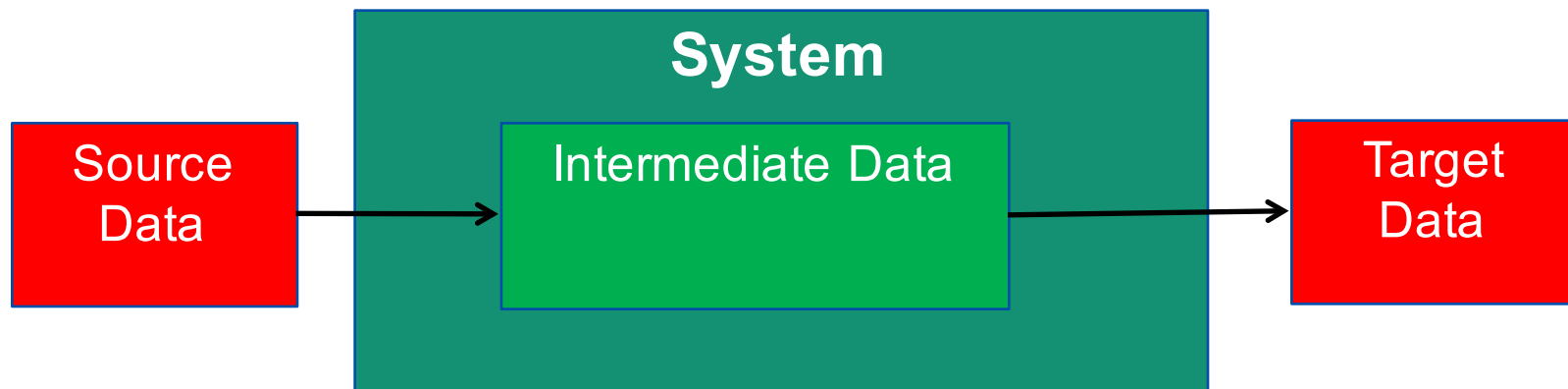
- User Requirements
 - What data users want to develop
 - Ex. Users want to create SDTM DM data set
- Functional requirement
 - How System creates data
 - Ex. System will create SDTM DM data set without errors and within 10 seconds

Data-centric SDLC: Analysis

- Business Requirement Analysis
- Target Data Requirement Analysis
 - Data format, structures and model
 - Data Quality
- Source Data Requirement Analysis
 - Data Discovery
 - Origin, data format, structures and model
 - Data Profile
 - Understanding contents of data
 - Data Quality

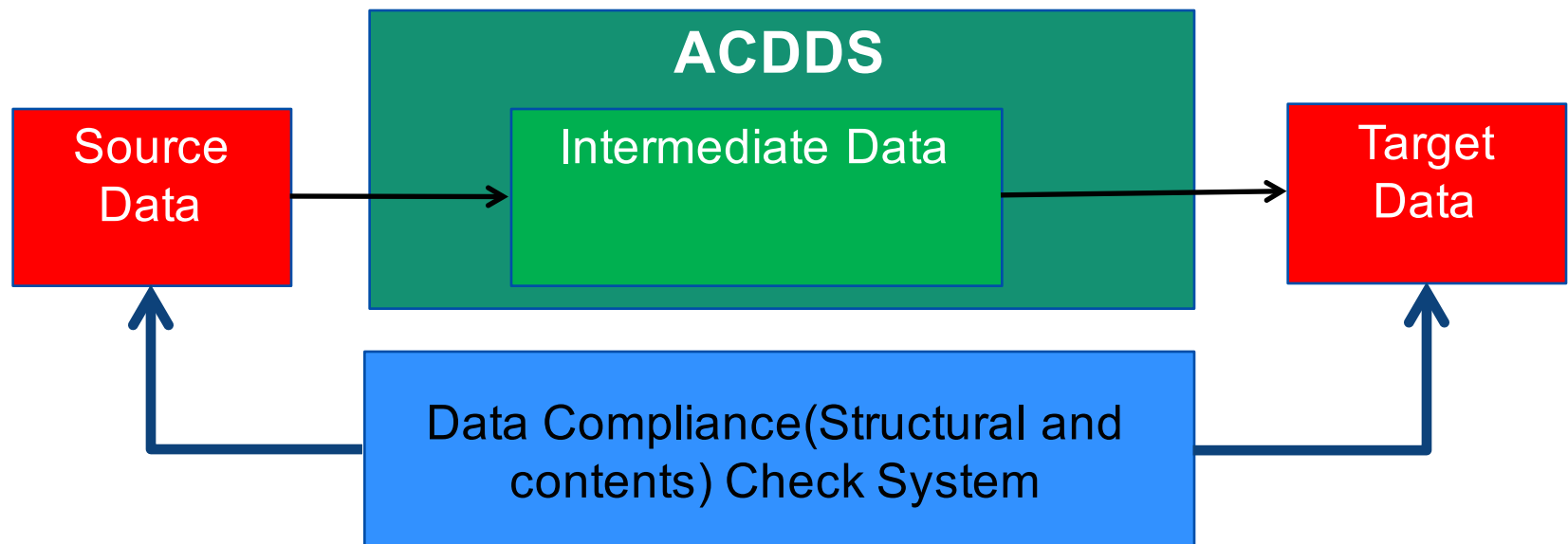
Data-centric SDLC: Design

- System Design
- Data Design (Source, Intermediate, Target)
 - Data Architecture/Modeling
 - Database Design



Data-centric SDLC: Build

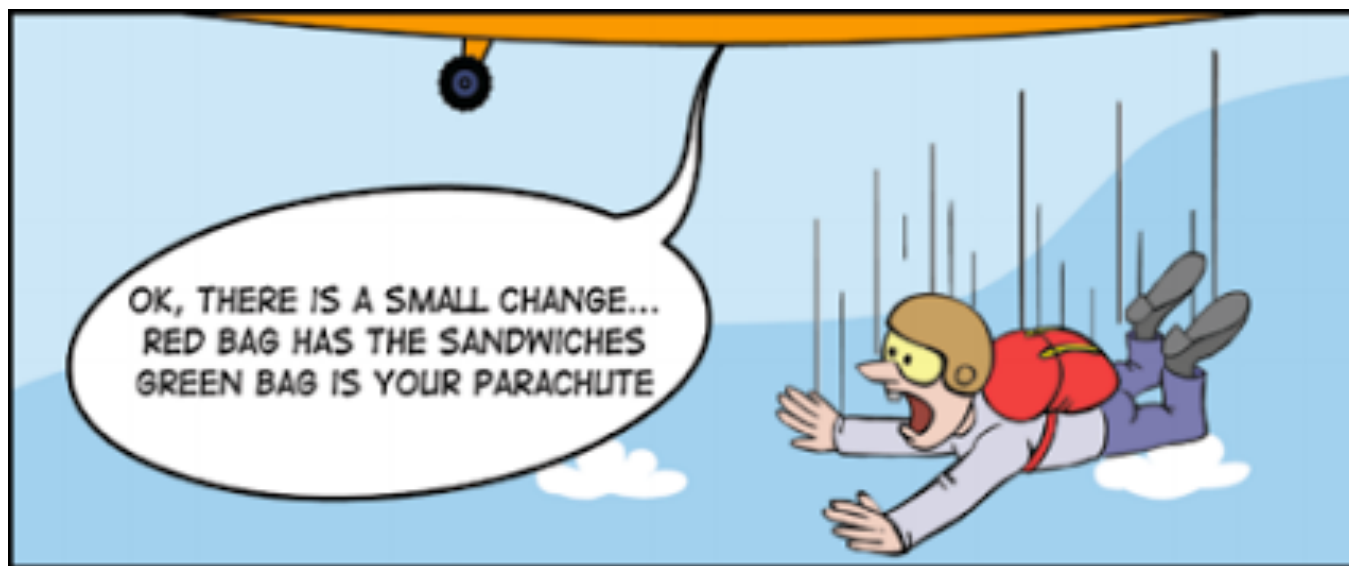
- ACDDS
 - Flexible codes for post-development
- Data compliance check system



Data-centric SDLC: Test

- System Testing: Unit Test, Integration Test, System Test and User Acceptance Test
- Data Testing of target data
 - Data Profiles
 - Any data quality issues
 - Data contents follow certain terminologies.
 - Data Compliances –correct data structures

Data-centric SDLC: Post-development



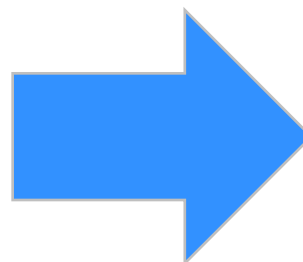
- New source data
- Data model change in target data

Data-centric SDLC – collaboration between IT and Biometric



- IT department – expertise on SDLC
- Biometric department – expertise on Data

Consideration of Automated Clinical Data Development System (ACDDS)



Consideration of ACDDS - continued



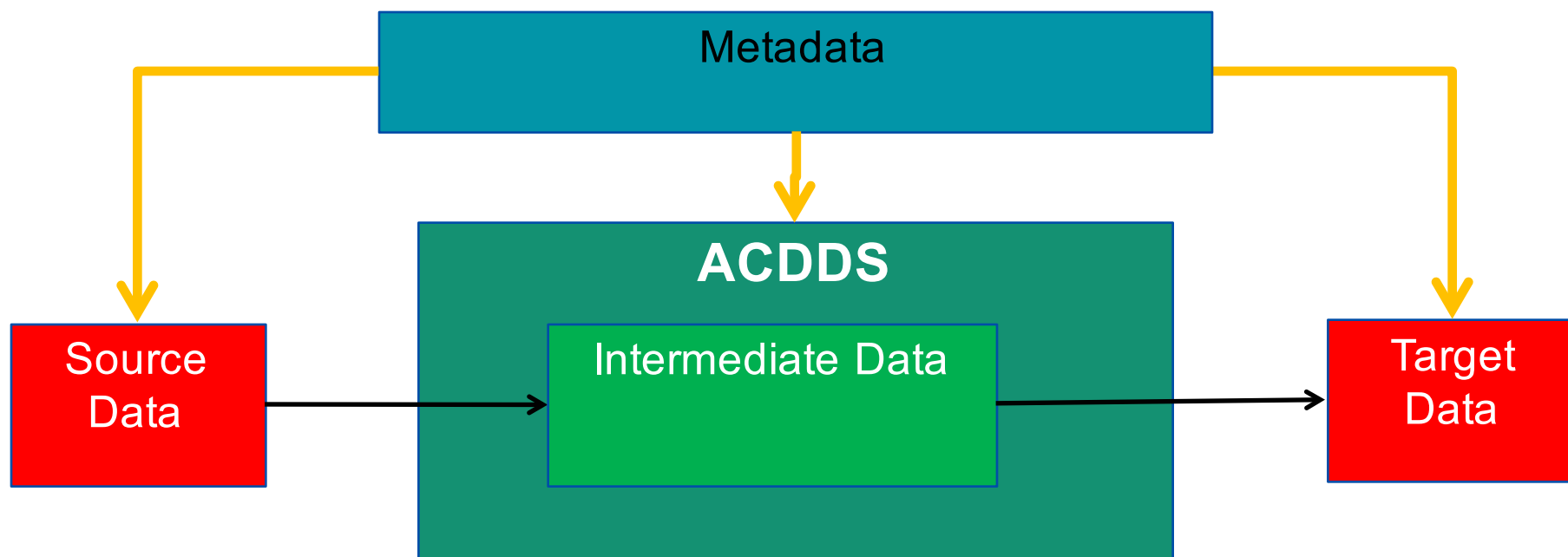


Consideration of ACDDS - continued



Metadata-driven ACDDS

- Machine-readable structural format of source and target data
- Machine-readable transformation from source to target data



Data-centric system

- Data > System
- Data dictates system
 - build system for the data
- Maybe no data transformation
 - Data ingest into system without transformation
 - Schema-agnostic data warehouse

Conclusion



We do it this way because we've always done it this way, son!

- Think differently
 - Data-centric thinking?
- Do differently
 - Data-centric SDLC?





Contact Information and Questions

Email address : kevin.lee@marklogic.com

Linkedin [Profile](https://www.linkedin.com/in/HelloKevinLee/) :
www.linkedin.com/in/HelloKevinLee/

Tweet : @HelloKevinLee

Slide share :
<http://www.slideshare.net/KevinLee56>

Blogs : HiKevinLee.tumblr.com



