

Practical meta data solutions for the large data warehouse

August 21, 2002

ACS Boston National Meeting
Chemical Information Division

PERFORMANCE
that
empowers

Agenda

- Introduction
- Developing a meta data strategy
- What successful projects are addressing
- Building ROI
- Project approach for a practical solution
- Standards and tools
- Enterprise meta data architecture
- In conclusion

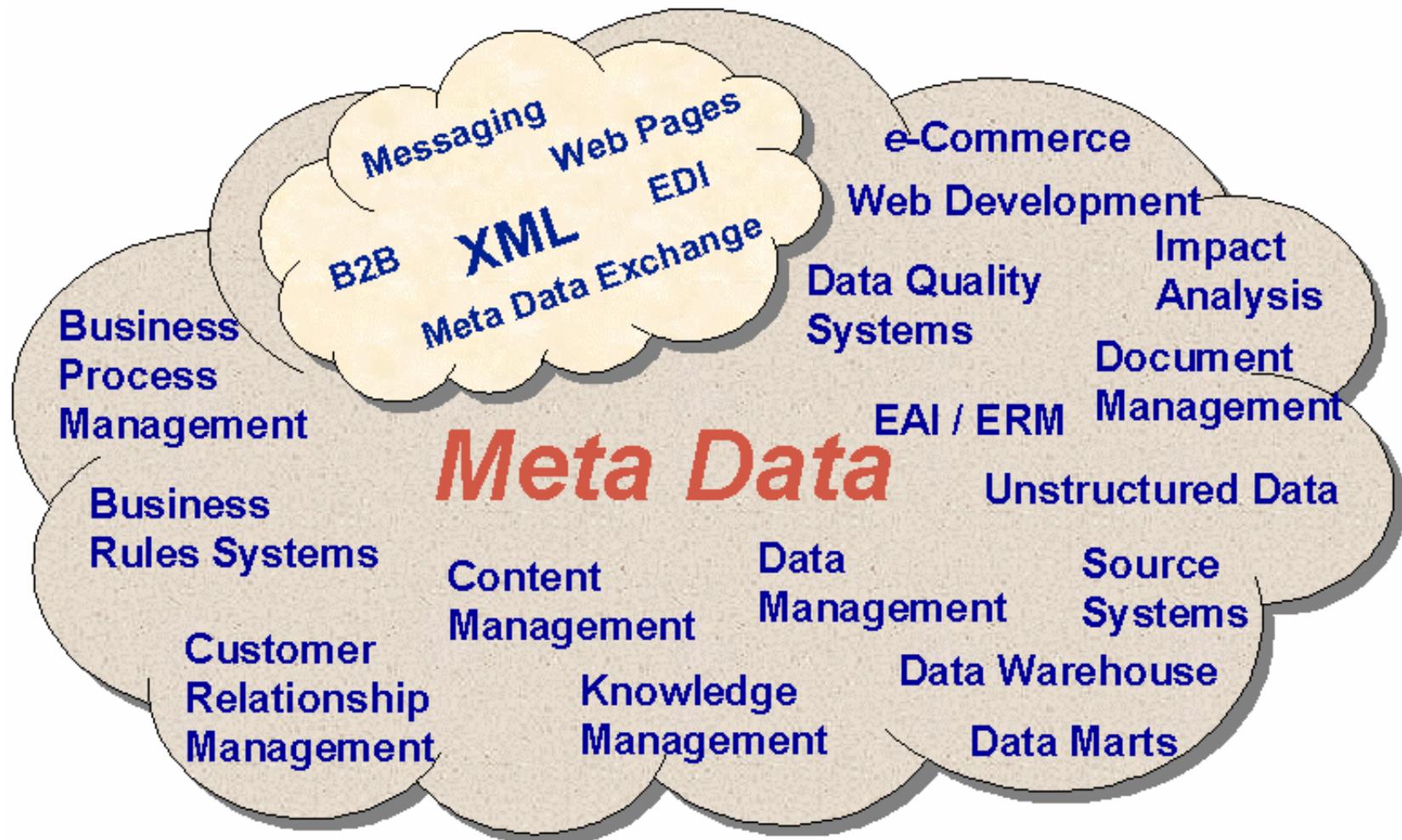
Introducing Knightsbridge

- Speakers:
 - Paul Vosters – Practice Area Leader, Pharmaceutical/Healthcare
 - Tom Gransee – Senior Principal
- Systems integrator focused *exclusively* on high-performance, complex data warehousing solutions
- 8-year-old company serving Fortune 500 clients
- 130 employees averaging 10+ years of data warehousing experience
- *Pathways* methodology designed specifically for data warehousing projects

Industry trends & meta data

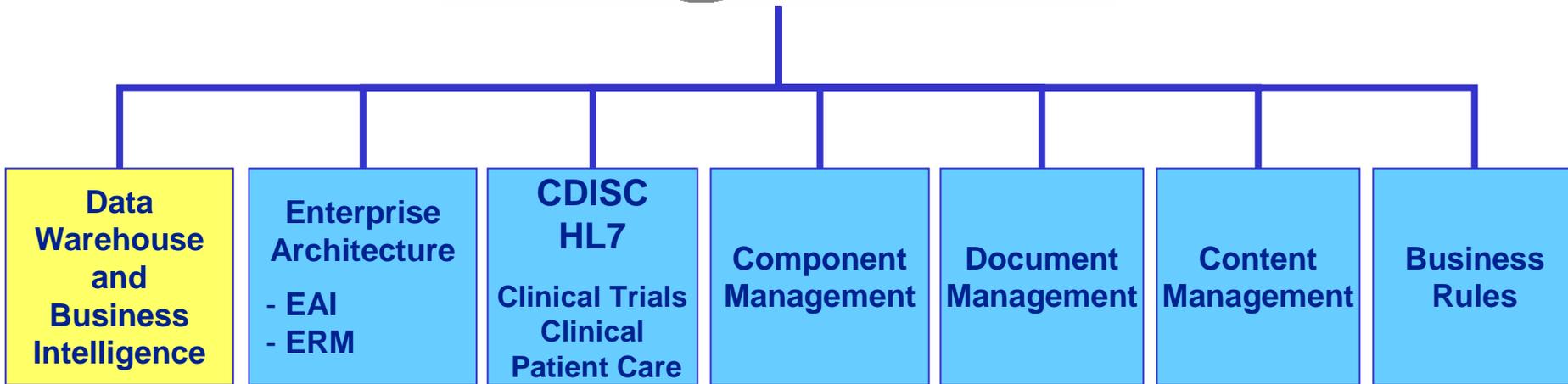
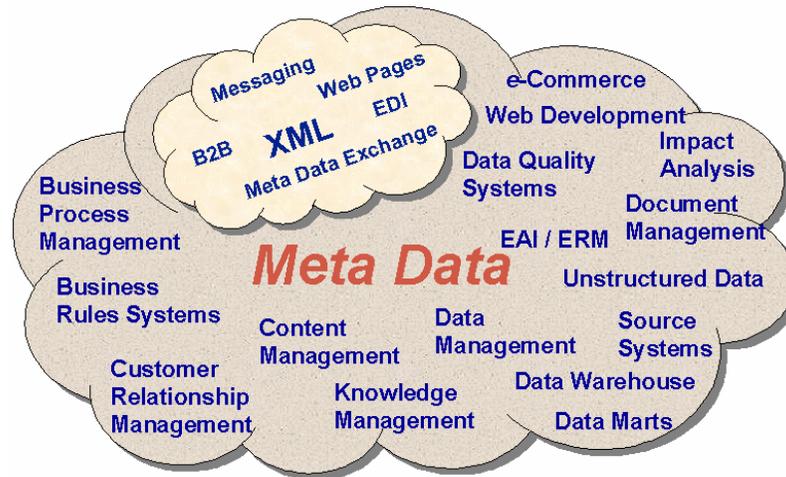
- Integration, Streamlining of process across silos
 - Clinical trial
- Acquisition strategy
 - Pharmacia / Pfizer
- Hipaa (EDI formatting)
- Government regulations
 - Clinical Trial (Meta data req.)
 - CDISC
 - Clinical data
 - HL7 Standard

Setting boundaries



You can't do it all at once!

Establishing a meta data strategy



Select a practical starting point and build on your success!

Defining meta data for the DW

The information needed by both business and technical associates to define, administer and navigate the DW

Define

- Expresses the meaning
- Ensures consistent usage and interchange
- Fosters an improved understanding of information sources

Navigate

- Provides linkages to data elements, reports, queries and other objects
- Expresses relationships to ensure accessibility and ease of use
- Helps locate information and explains how to access it

Administer

- Captures lineage, usage, and transformation to ensure performance and integrity
- Builds trust and confidence in the information sources
- Provides the flexibility to respond to business needs in a timely manner

Why is the DW a good starting point?

- DW typically focuses on the data that most needs to be shared
- DW presents the greatest need to understand the data because it is cross-functional
- Real business benefit can be obtained for a practical investment
- Existing DW are being re-architected
- Meta data standards and tools are beginning to have an impact in this area
- Challenges are created in a best-of-breed development environment

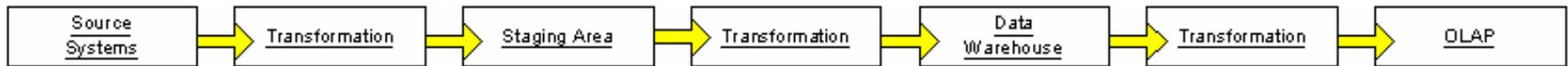


*Built for today
- architected
for tomorrow*

What are successful projects addressing?

1. Lineage

- What data is available, where it came from and how it's transformed
- Including plain English definitions, currency and accuracy



2. Impact analysis across tools and platforms

- Impossible to do without a formalized meta data technology solution

3. Appropriate information by user type

- Easy access to meaningful meta data
- “How is it different from what I’m used to seeing?”

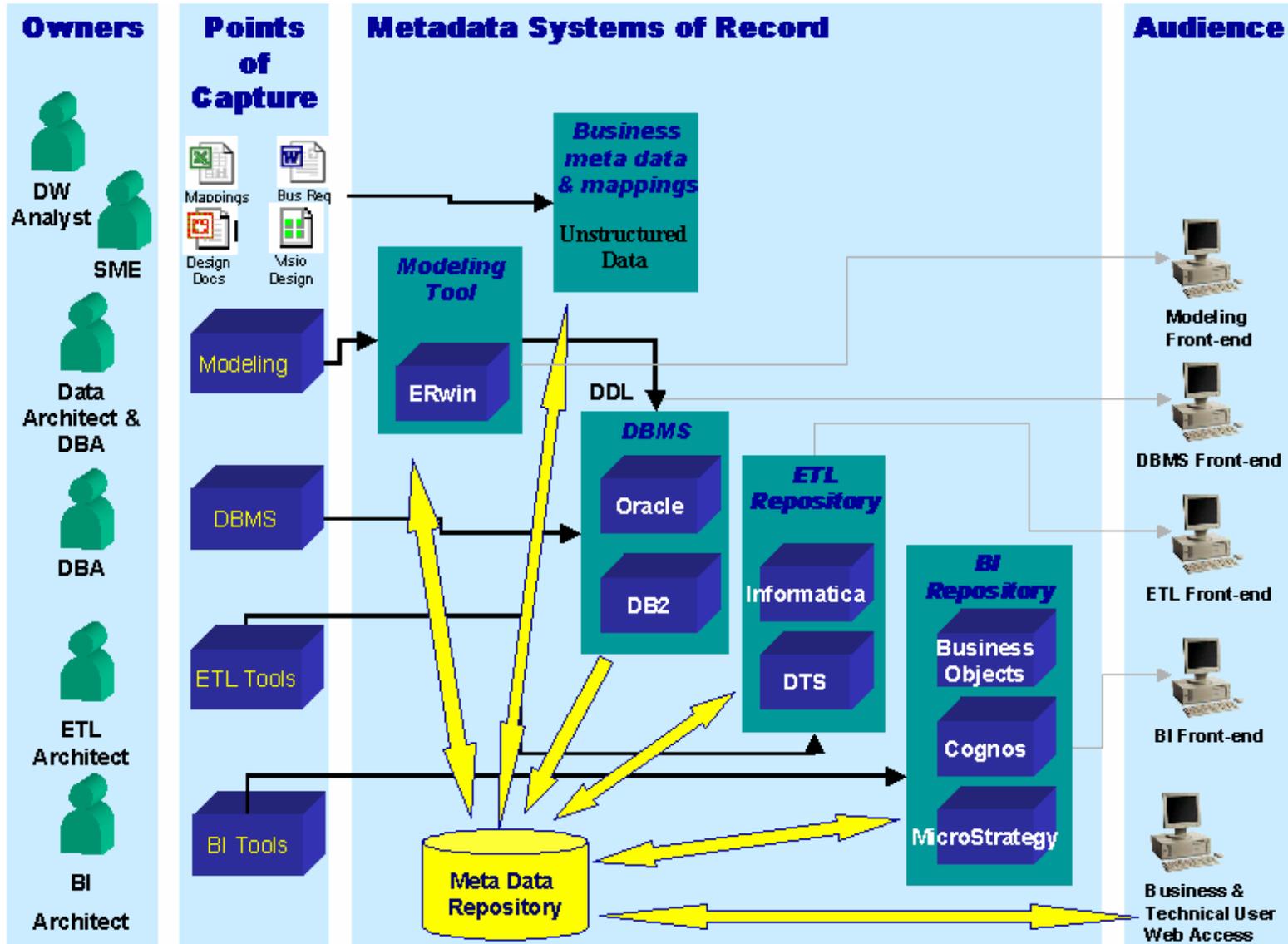
4. Versioning

- How has it changed over time?
- Moving from development, to test, to production

5. Live meta data

- Meta data is a natural part of the process
- A function fails if the meta data is not complete and accurate

Meta data architecture for the DW



Building an ROI for meta data

Issue

Benefit

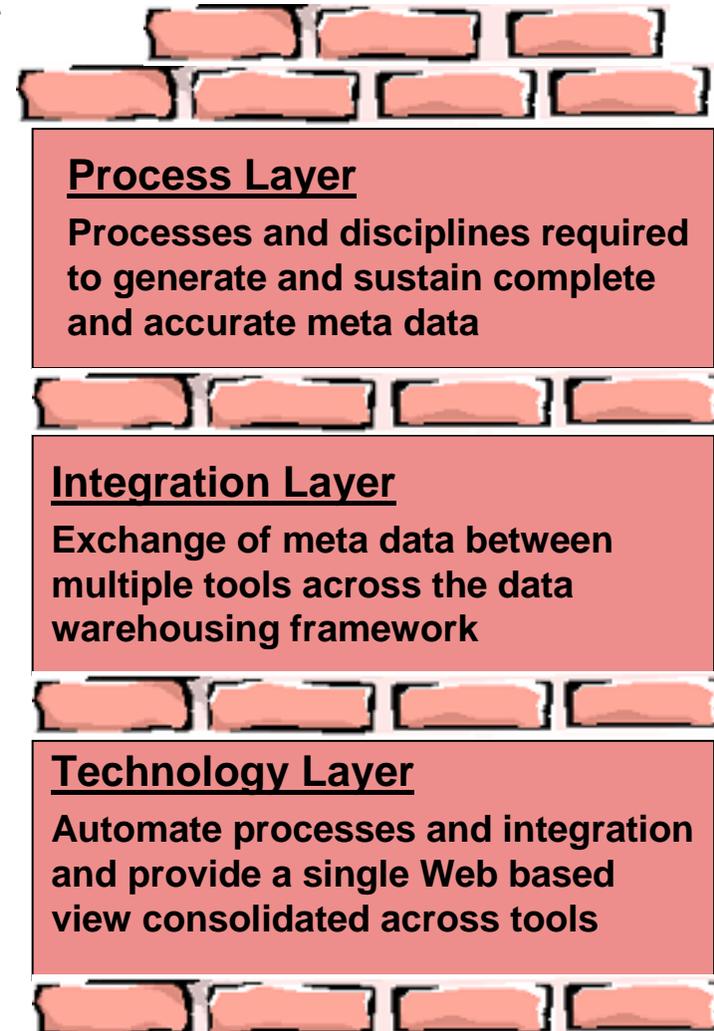
- | | | |
|--|--|--|
| <ul style="list-style-type: none"> • Business Power Users told us they spend up to 70% of their time locating the right information source and resolving multiple versions of the truth |  | <ul style="list-style-type: none"> • At least 20% improvement in the time required to locate and validate information |
| <ul style="list-style-type: none"> • Technical analyst's and developer's told us they spend up to 80% of their time finding the data needed to satisfy a request |  | <ul style="list-style-type: none"> • At least 20% improvement of application development and maintenance activities |
| <ul style="list-style-type: none"> • Business Power Users and technical analysts told us it requires six to nine months for a new associate to become proficient in using data |  | <ul style="list-style-type: none"> • Reduced learning curve for new associates by 3 months and lower the mentoring required from key resources |
| <ul style="list-style-type: none"> • Lack of complete automated Impact analysis within and across tools creates a serious risk when implementing changes |  | <ul style="list-style-type: none"> • At least 20% reduction in the time required to perform a complete impact analysis and reduce risk of errors when migrating changes to production |

Quote from the META Group's series of white papers addressing application delivery strategies
“meta data interchange will improve Application Development and maintenance efficiency by up to 30%, and real-time meta data interoperability will enable up to 50% improvement in Application Development and maintenance efficiency.”

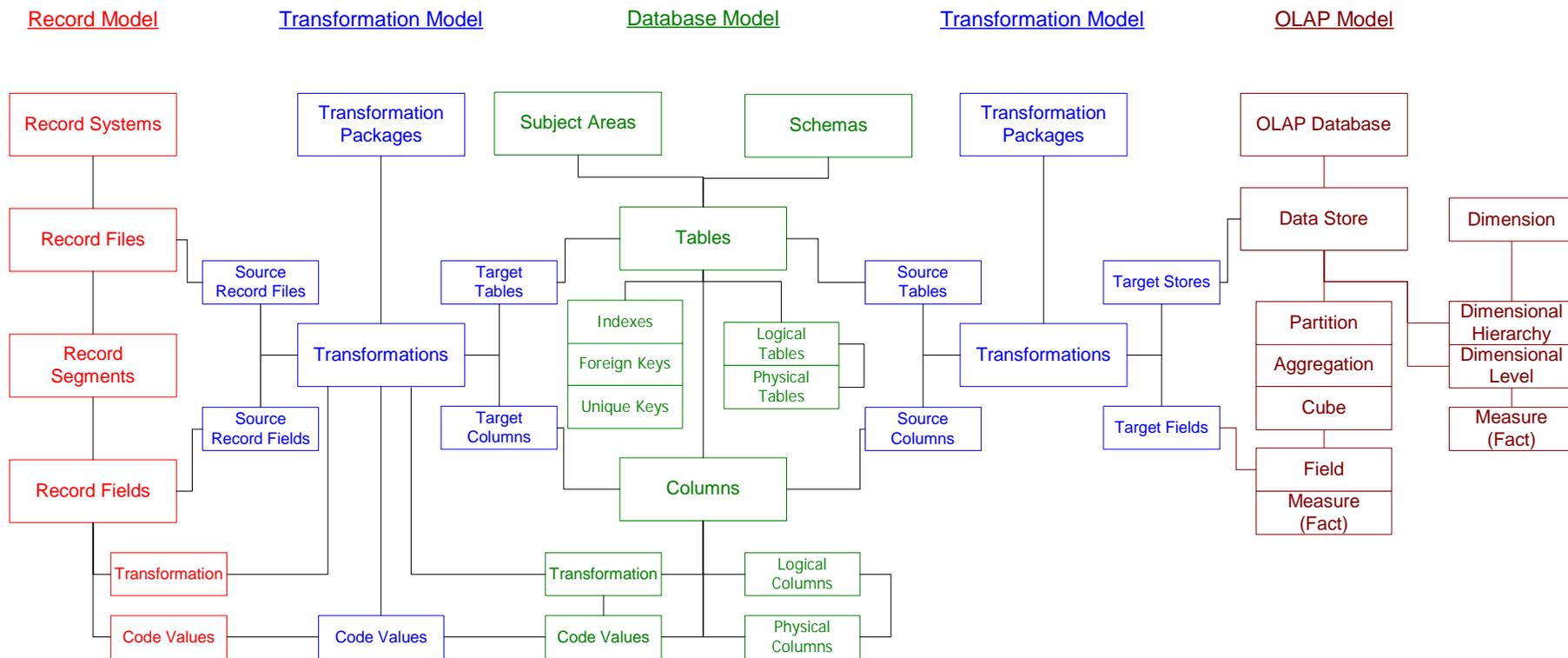
The need for a complete project lifecycle

- Document today's meta data environment
 - Identify meta data users
 - Identify sources of meta data
 - Identify problems and opportunities
- Identify requirements
 - Define the need for meta data
 - Define objectives and benefits
- Recommend a high-level approach
 - Processes and disciplines
 - Integration requirements
 - Delivery requirements
 - Technology component / tool
 - Change management process priorities
- Propose meta data architecture
- Iterative release approach

Building an Architecture



Common warehouse metamodel standard



1. A powerful metamodel

2. Meta data exchange and integration

3. Architecture and framework

A partial high level representation of the industry standard OIM/CWM Metamodels

Tools and selection criteria

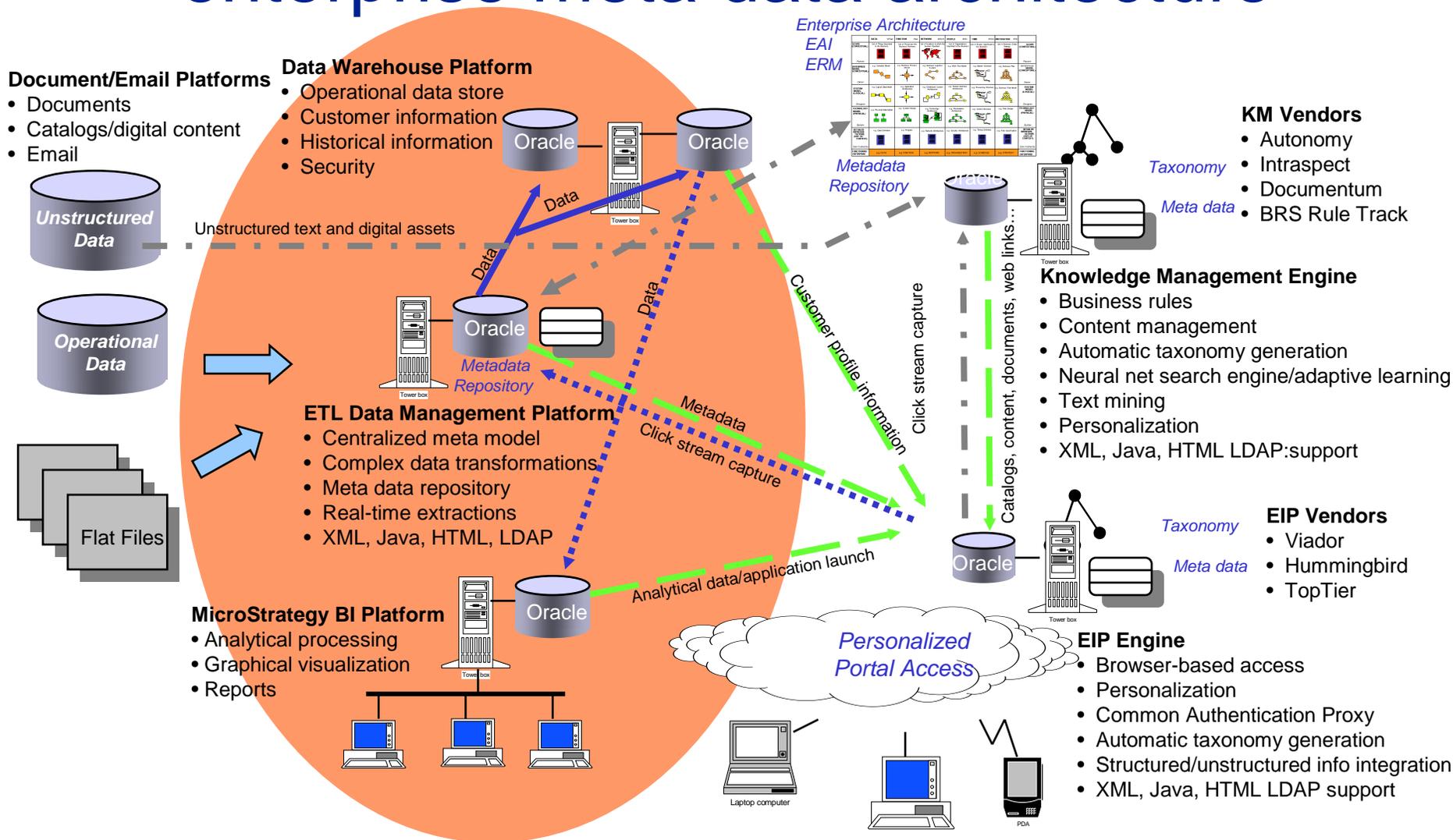
Tools

- **Enterprise**
 - CA – Advantage (formerly Platinum)
 - ASG - Rochade
 - Adaptive Foundation (formerly Unisys)
- **DW Suite solutions**
 - Microsoft Repository
 - DWSOft – Navigator web browser
 - SAS –Warehouse Administrator
 - Oracle Warehouse Builder
 - OWB Repository
- **ETL**
 - Ab Initio
 - Informatica
 - Data Advantage Group – MetaCenter
 - Ascential – Meta Stage
- **Modeling**
 - ERwin Suite
 - Rational Rose
 - Oracle Designer
 - Popkin System Architect

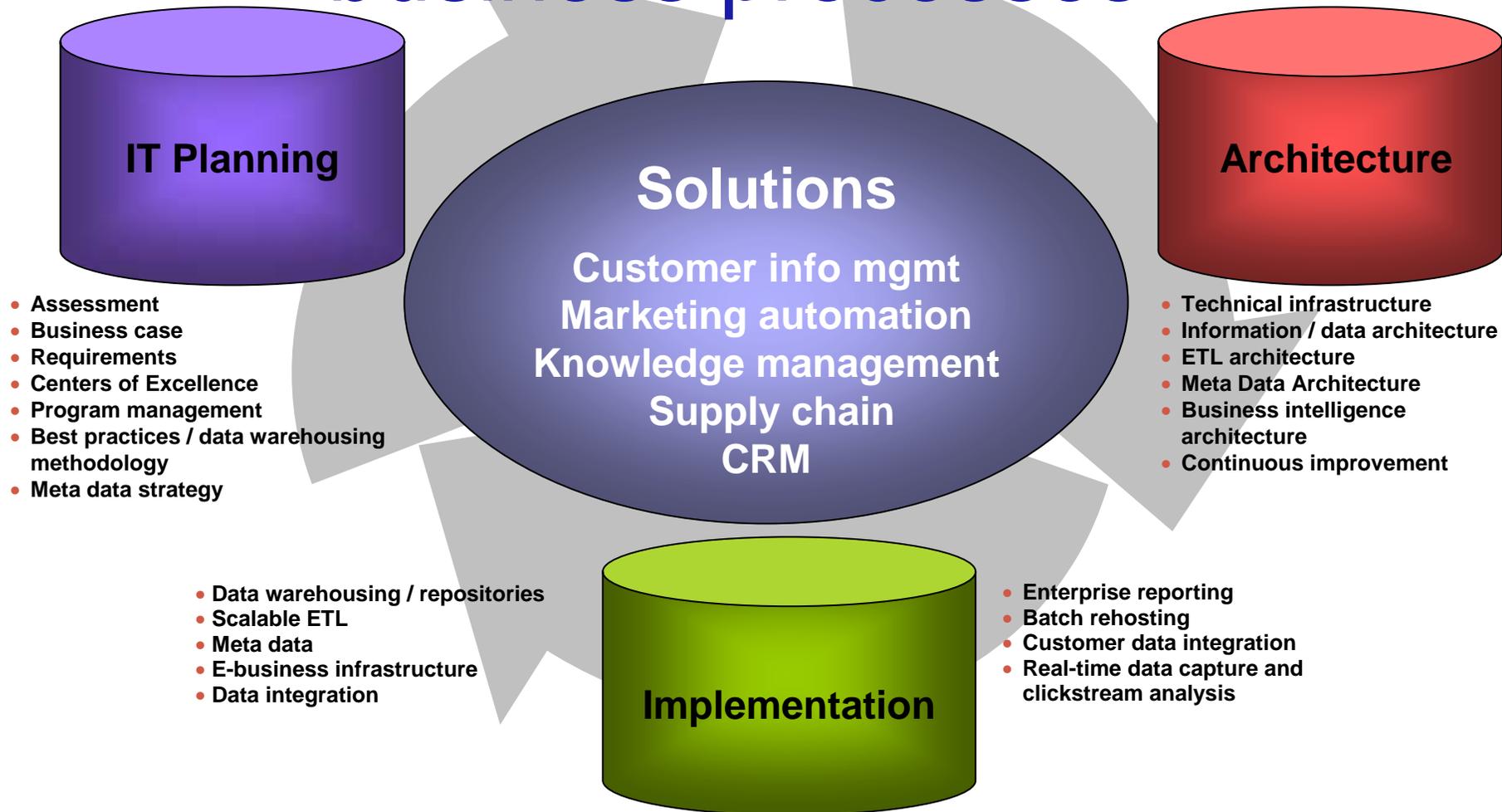
Selection Criteria

- Manage associations across tools
- Search and retrieval across tools
 - Lineage
 - Impact analysis
 - Plain English definitions
- Reduce integration cost in a best-of-breed development environment
- Enable Plug-and-play tool strategy
- Available automated bi-directional meta data exchange bridges / adapters
- Flexible user interface based on current web standards – business and technical
- Template driven retrieval of meta data
- Group / role based security
- Interoperability between metamodels
- Based on industry standards – CWM
- Extensibility
- Support of Federated repositories
- Ability to expand beyond the DW
- Versioning – extracting a time slice

The DW meta data solution within the enterprise meta data architecture



Big-data solutions power high-value business processes



Meta data powers big-data solutions

Questions

