

*MasterCard
International*




Bridging The Gap – Adding Semantic Awareness To Today's Application Systems

DAMA International Symposium - 2006

John Singer, MasterCard International


*MasterCard
International*



John Singer – Data Architect

- 25 years experience
 - Information Center, End User Computing
 - DBA, DA, DW
 - Recovering IT Manager
 - Metagroup industry analyst
- Currently working on...
 - Metadata management
 - Content management, taxonomy, corporate vocabularies
 - Data architecture for transaction systems

March 31, 2006Bridging The Gap2


MasterCard International


Purpose of the Presentation

- How do we add semantic value to data?
- The Semantic Web Technologies – what are they, will they work?
- How might we transition from an N-Tier Relational “stack” to a Semantic Content centric SOA?
- How do we adapt?

A journey of discovery reminiscent of the emergence of RDBMS

March 31, 2006Bridging The Gap3

MasterCard International


What Is The Meaning of Meaning?

The “Ontology Spectrum”

- Defined by Dr. Leo Obrst – Mitre Corporation
 - Taxonomy (classification hierarchy)
 - Thesauri (broader/narrower/related meaning)
 - Conceptual Models (OO, UML)
 - Ontology (logical theory, OWL, CYC)
- Increasing level of semantic expressiveness

XML alone is insufficient to convey full semantic awareness

March 31, 2006Bridging The Gap4

Through The Looking Glass

MasterCard
International

- Where does semantic interpretation occur?
 - behind the glass (computer) – only if meaning is encoded
 - In front of the glass (person) – must rely on the human brain
- The more expressive the semantics, the more the software is capable of reasoning like a person

March 31, 2006
Bridging The Gap
5

Categorization and Classification Data


MasterCard
International

Called different things in different systems

SYSTEM TYPE	CATEGORIZATION AND CLASSIFICATION
OLTP	Code Tables
Data Warehouse	Dimensions
MDM	Code tables and dimensions on steroids
Web Content	Tags, "Microformats", Flickr, Wiki, RDF
ECM	Metadata, Topic Maps, DITA

March 31, 2006
Bridging The Gap
6

MasterCard
International




What Is Metadata Anyway?

- Traditional Definition – Data about Data
- Expanded Traditional - Data about IT Assets
- My definition – the data we forgot to capture
- New Age Definition
 - Data that provides context for other data
 - Note this is infinitely recursive
 - One persons metadata is another persons data
 - Taxonomy and Ontology is the technology

The term “Metadata” loses its meaning in the Semantic world

March 31, 2006 Bridging The Gap 7

MasterCard
International




The Meaning Of Meaning

Resurgent Interest In “Metadata”

- Traditional “Metadata”
 - GLBA, SOX compliance
 - Enterprise Modeling - IT planning, IT Business alignment
- New Age “Metadata”
 - Structuring Unstructured data
 - Web 2.0
 - Taxonomy and Ontology

Resurgent interest in metadata driven by increased need to know what data “means” – not just catalog its existence

March 31, 2006 Bridging The Gap 8

MasterCard International



What Is The Semantic Web?

Revenge of the AI Nerds?

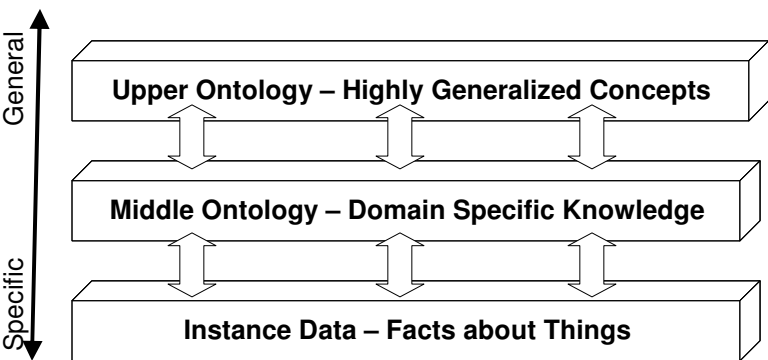
- Tim Berner-Lee's vision to bring order to web chaos
- W3C Standards
 - RDF – Resource Description Facility
 - OWL – Web Ontology Language (WOL?)
- Full realization requires infrastructure

RDF and OWL are “static” representations of data – inference engines, agents, and business rules bring them to life

March 31, 2006Bridging The Gap9

MasterCard International


Ontological Layering



March 31, 2006Bridging The Gap10



From E/R to Ontology

Modeling differences

- Eliminates distinction between conceptual and physical model by combining them into one
- Supertype/Subtype
- Difficult to define what is a class/property/instance
- More similar than different - E/R, OO, Ontology all have the same roots in “logic”
- Facts can be stated or “inferred”
- Relationships are not bi-directional



The Semantic Web Is More Than “Tags”

- Better tagging alone won't solve web chaos
- RDF/OWL represents more than another web page tagging scheme – it's a “new” way to represent data
- Must be analyzed in the full context of the mainstreaming of AI/Expert System technology
- Vendor reluctance based on a combination of “marketecture conflict” and real technical concerns

MasterCard International

Moving From Traditional to Semantic

- Comparing Technology Stacks
 - Traditional N-Tier
 - Document and Content Management
 - Web 2.0
 - AI / Expert System
- People Perspective
- Getting Started

March 31, 2006
Bridging The Gap
13

MasterCard International

Traditional N-Tier


Master Data Management (MDM) System
 450 Entities - 1200 Business Rules
 Very Successful – But...

- High Cost to develop
- Difficult to change
- Multiple Metadata repositories (and people) needed to maintain knowledge about the system
- Data Abstractions hard to sell, hard to implement
- Application to Application integration projects (i.e. CDI) exhibit the same set of problems

```

graph TD
    DB[MDM Database] --> DB2ProcRead[DB2 Proc- Read]
    DB --> DB2ProcUpdate[DB2 Proc- Update]
    DB2ProcRead --> VBDisplay[VB Display]
    DB2ProcUpdate --> CICS[CICS - B. Rules]
    CICS --> VBEdits[VB Edits]
    VBEdits --> VBDisplay
            
```

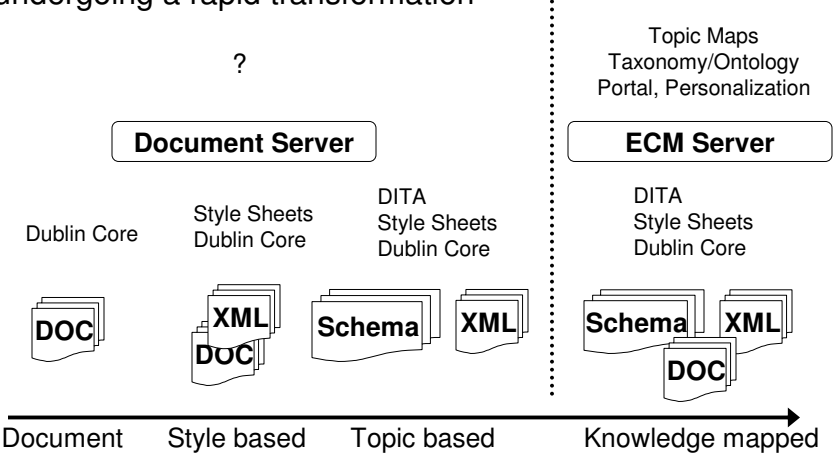
March 31, 2006
Bridging The Gap
14

MasterCard International


ECM – From Unstructured To Ontology

The Enterprise Content Management (ECM) world is undergoing a rapid transformation

?




Document Server **ECM Server**

Dublin Core Style Sheets Dublin Core DITA Style Sheets Dublin Core DITA Style Sheets Dublin Core

Document Style based Topic based Knowledge mapped

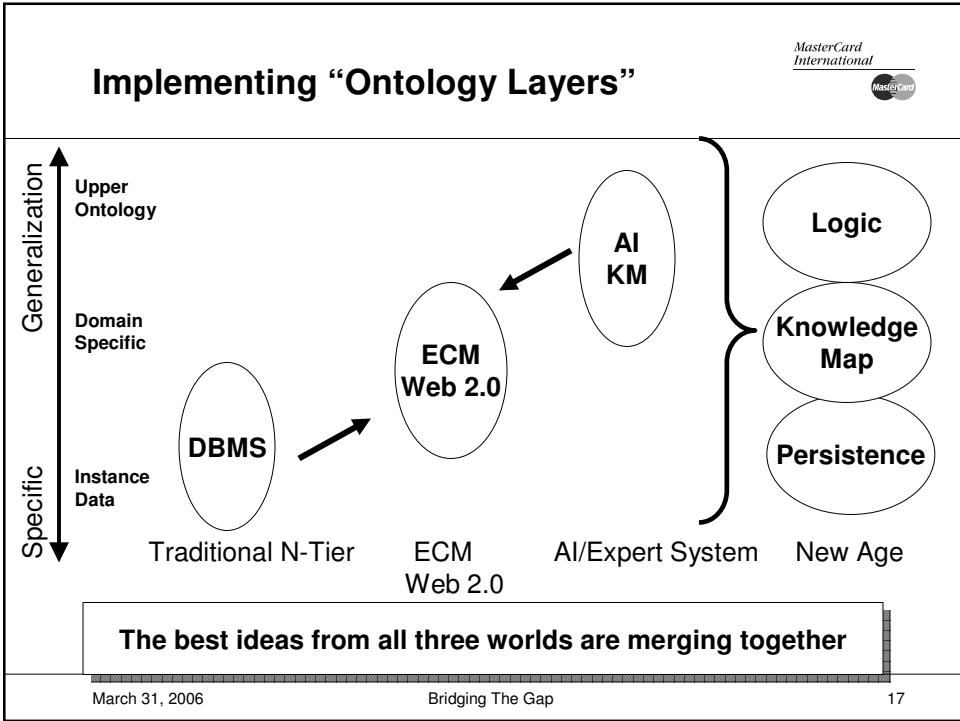
March 31, 2006 Bridging The Gap 15

MasterCard International


AI and Expert Systems

- Based on 1000's of years of research
- Mature Technology – but...
 - Proprietary
 - Not broadly applied – only niche success
- Yet...
 - The only alternative to Relational and OO for encoding knowledge

March 31, 2006 Bridging The Gap 16



New Age Semantic Servers

- Application Servers and Data Servers that support RDF and OWL standards
 - Cerebra
 - RDF Gateway
 - Others...
- Why will it work this time around?
 - Standards – W3C, OMG, Business rules
 - Asynchronous disconnected nature of the Web enables globalization/localization of knowledge bases

RDF tags on web pages require semantic servers to fully realize the Semantic Web

March 31, 2006 Bridging The Gap 18

When Worlds Collide

MasterCard International

Library Science

Content Author

Knowledge Engineer

Semantic Web

Technorati

Usability Human Factors

Data Architecture

**These independent world views are merging.
Will traditional data architecture join the dance?**

March 31, 2006 Bridging The Gap 19


Relational Adoption Revisited

MasterCard International

- Semantic Web Naysayer's sound a lot like the anti-relational database crowd of the early 80's
 - its too theoretical/academic – not real world
 - too difficult to learn
 - won't perform
 - why change – what we have works fine
 - its not that much different than network/hierarchal - IMS has logical pointers

Relational Database overcame all these obstacles – Semantic Technology will do the same


March 31, 2006 Bridging The Gap 20

MasterCard
International


Getting Started – Learn The Technology

- Protégé
 - Free download – Stanford University
 - “frames” based knowledge capture system
 - OWL/RDF plugins
 - Can use this today to manage taxonomies
- Open Cyc
 - Free download version
 - Large ontology and inference engine/theorem prover
- RDF Gateway
 - Free personal use license
 - Complete stack – RDF data store, RSP scripting for web pages

March 31, 2006Bridging The Gap21

MasterCard
International


Getting Started – Practical Applications

- For Data Architects
 - Make nice with the Content Management/Authoring/Web design/Usability/Corporate Library/Research departments in your company – you may be surprised at what you find there
 - Managing standard business vocabulary/taxonomy is a critical success factor for ECM, but vendors/implementers treat it as an afterthought
 - Look at Semantic and Content Management technologies as a part of your metadata management strategy – low risk starting point
 - Realize that your world is not just relational...

March 31, 2006Bridging The Gap22

*MasterCard
International*

