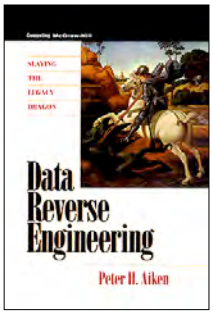
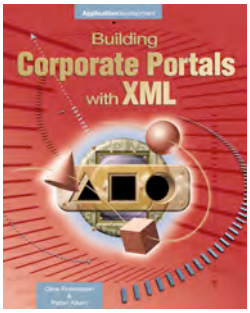
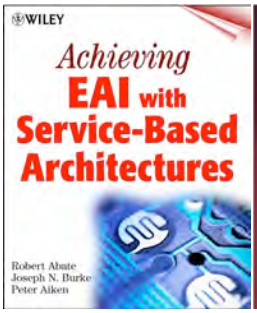
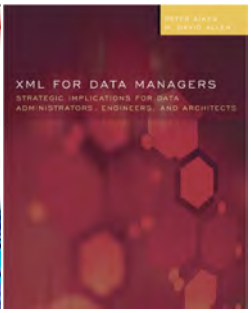








XML in Data Management

Peter Aiken

	E	F	G	H	I	K
1					n	
2					f	
3					o	
4					r	
S	y	s	t	e	m	s
6					a	
7					t	
8					I	
9					o	
0					n	

VCU School of Business

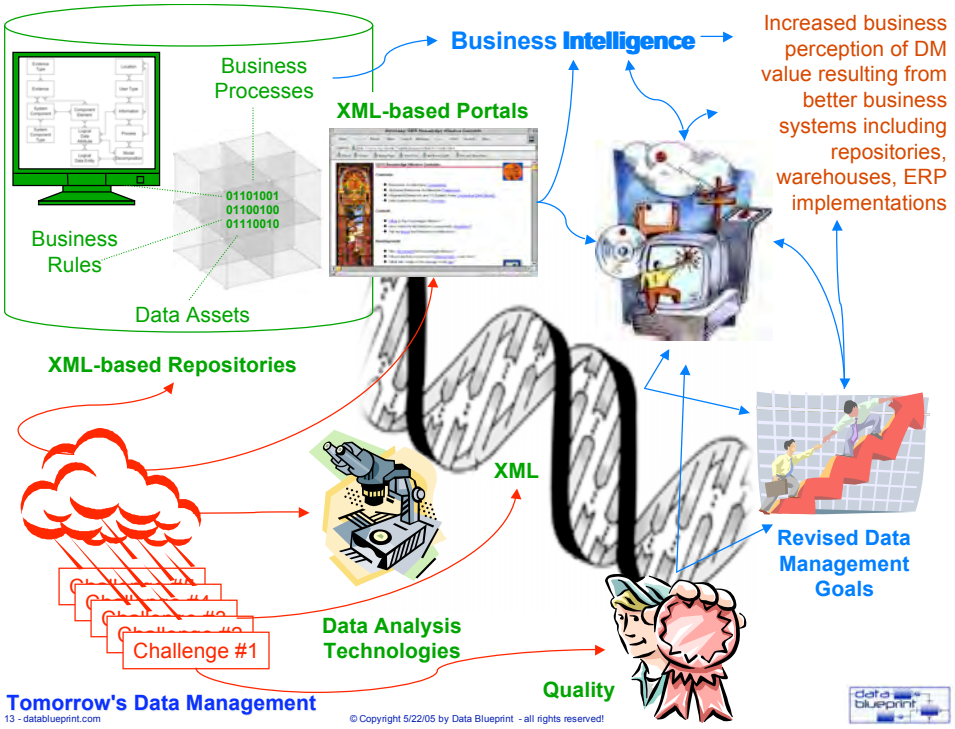
Virginia Commonwealth University

© Copyright 5/22/05 by Data Blueprint - all rights reserved!

Peter Aiken

- Full time in information technology since 1981
- IT engineering research and project background
- University teaching experience since 1979
- Six books and dozens of articles
- Research Areas
  - reengineering, data reverse engineering, software requirements engineering, information engineering, human-computer interaction, systems integration/systems engineering, strategic planning, and decision support systems
- Director
  - George Mason University/Hypermedia Laboratory (1989-1993)
- DoD Computer Scientist
  - Reverse Engineering Program Manager/Office of the Chief Information Officer (1992-1997)
- Visiting Scientist
  - Software Engineering Institute/Carnegie Mellon University (2001-2002)
- Published Papers
  - *Communications of the ACM*, *IBM Systems Journal*, *InformationWEEK*, *Information & Management*, *Information Resources Management Journal*, *Hypermedia*, *Information Systems Management*, *Journal of Computer Information Systems* and *IEEE Software*
- DAMA International Advisor - Since 1999
  - 2001 DAMA International Individual Achievement Award with Dr. E. F. "Ted" Codd
- Founding Director ***Data Blueprint*** Since 2000

© Copyright 5/22/05 by Data Blueprint - all rights reserved!


Business Processes
Business Intelligence
Increased business perception of DM value resulting from better business systems including repositories, warehouses, ERP implementations
XML-based Portals
01101001
01100100
01110010
Business Rules
Data Assets
XML-based Repositories
XML
Revised Data Management Goals
Challenge #1
Data Analysis Technologies
Quality
Tomorrow's Data Management
13 - datablueprint.com
© Copyright 5/22/05 by Data Blueprint - all rights reserved!

Business Processes
Business Intelligence
Increased business perception of DM value resulting from better business systems including repositories, warehouses, ERP implementations
XML-based Portals
Business Rules
Data Assets
XML-based Repositories
XML
Revised Data Management Goals
Challenge #4
Challenge #3
Challenge #1
Challenge #2
Challenge #1
Data Analysis Technologies
Quality
Tomorrow's Data Management
Contact Information:
Peter Aiken, Ph.D.
Department of Information Systems
School of Business
Virginia Commonwealth University
1015 Floyd Avenue - Room 4170
Richmond, Virginia 23284-4000
Data Blueprint
Maggie L. Walker Business & Technology Center
501 East Franklin Street
Richmond, VA 23219
804.521.4056
http://datablueprint.com
office :+1.804.883.759
cell:+1.804.382.5957
e-mail:peter@datablueprint.com
http://peteraiken.net
15 - datablueprint.com
© Copyright 5/22/05 by Data Blueprint - all rights reserved!

http://peteraiken.net
data blueprint
website paiken:home
News
Dubai Conference
Makes News
Building Corporate Portals using XML
book (reprinted)
Upcoming Research
Presentations
Waiting for the Moon
CD & MP3s available online
Greetings from Montpelier, Virginia - welcome to my website! You might be wondering what's here? I've structured the contents into a number of menu topics. Please click on each topic for a description of its contents. (Incidentally, "paiken" has been my e-mail address prefix since 1981 - Alien Tsirinsky named this website.)
Some topics of interest include the VCU School of Business - Department of Information Systems - where my affiliation dates to 1978.
VCU
Virginia Commonwealth University
Relatedly, in 2000, VCU spun off a company named "Data Blueprint" to capitalize on intellectual property developed as part of my research into data engineering and data reengineering. I regularly give talks and appear at conferences, speaking on these and related topics as part of the IEEE, DAMA and other distinguished speaker programs.
Last of all, I really enjoy both playing and listening to music in various forms with lots of different friends!
Thank you for looking around. If you have questions or comments please contact me via e-mail at: paiken@datablueprint.com. This is my web site - none of the opinions here represent the opinions of my various employers. Thank you for visiting!
Made with Macintosh
This page and all web site contents were last updated and are copyright Saturday, August 30,

Content Access

- X1 XML Context and Expectations
- X2 XML DM Usage & Overview
- X3 XML Component Architecture
- X4 XML Framework Technologies
- X5 XML & Data Engineering
- X6 XML & Data Management Technologies
- X7 XML-based Portal Capabilities
- XE XML Solution Design Patterns

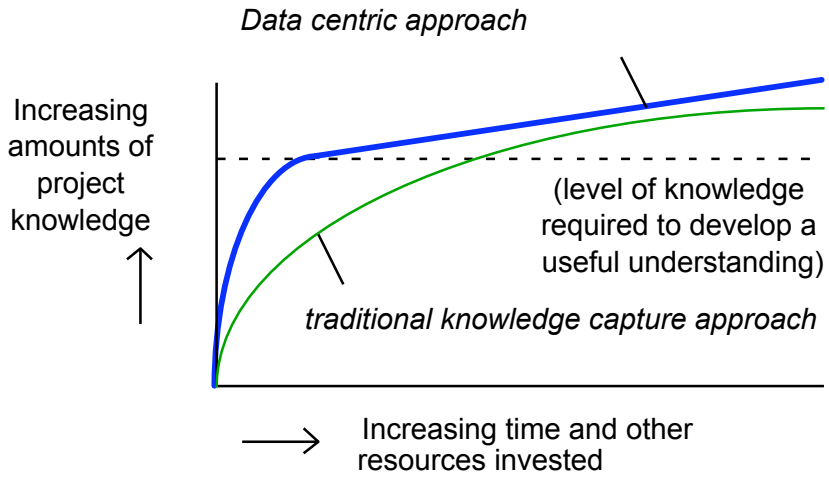
© Copyright 5/22/05 by Data Blueprint - all rights reserved!

1. XML Basics & Expectations

- Investing in metadata
- XML hype: Management by Magazine
- What XML developers want you to know about XML
- XML & data management

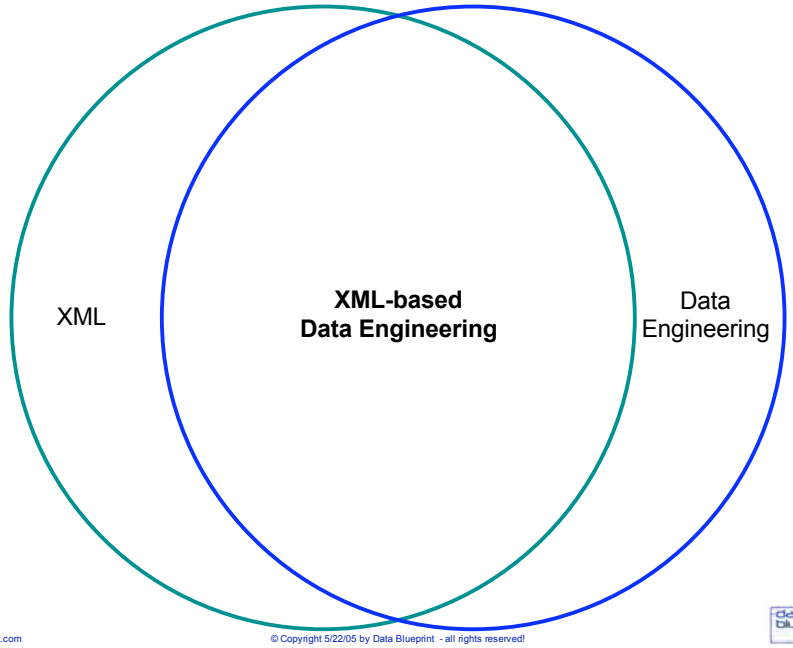
© Copyright 5/22/05 by Data Blueprint - all rights reserved!

Objectives



© Copyright 5/22/05 by Data Blueprint - all rights reserved!

Exploiting The Intersection Between XML & Data Engineering



© Copyright 5/22/05 by Data Blueprint - all rights reserved!

IT Project Failure Rates

Recent IT project failure rates statistics can be summarized as follows:

- *Carr 1994*
  - 16% of IT Projects completed on time, within budget, with full functionality
- *OASIG Study (1995)*
  - 7 out of 10 IT projects "fail" in some respect
- *The Chaos Report (1995)*
  - 75% blew their schedules by 30% or more
  - 31% of projects will be canceled before they ever get completed
  - 53% of projects will cost over 189% of their original estimates
  - 16% for projects are completed on-time and on-budget
- *KPMG Canada Survey (1997)*
  - 61% of IT projects were deemed to have failed
- *Conference Board Survey (2001)*
  - Only 1 in 3 large IT project customers were very "satisfied"
- *Robbins-Gioia Survey (2001)*
  - 51% of respondents viewed their large IT implementation project as unsuccessful
- *MacDonalds Innovate (2002)*
  - Automate fast food network from fry temperature to # of burgers sold-$180M USD write-off
- *Ford Everest (2004)*
  - Replacing internal purchasing systems-$200 million over budget
- *FBI (2005)*
  - Blew $170M USD on suspected terrorist database-"start over from scratch"

http://www.it-cortex.com/stat_failure_rate.htm (accessed 9/14/02)
New York Times 1/22/05 pA31

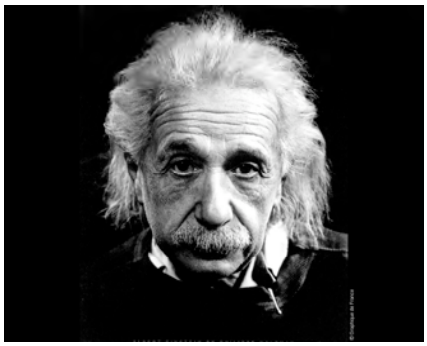
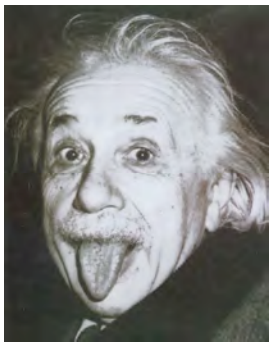
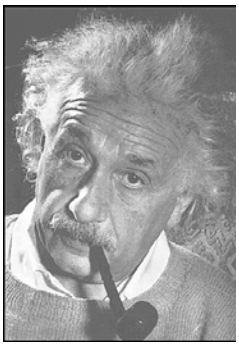
© Copyright 5/22/05 by Data Blueprint - all rights reserved!

Oct 2004 IRS Accomplishment

- Unified five definitions of "child"
- Reduce 5 definitions to 1 for tax return preparations such as:
  - Dependent
  - Earned income tax credit
  - Child credit
- Different reasons, either it
  - "Was developed to carry out social policy objective(s), or
  - Someone perceived it was going to save revenue"
- "Is it easier for (customers) to understand and it is easier for IRS to audit and there a lots of things like that we can do"
- Initiative started in 1991 - it took **13** years including 2.5 years moving as legislation!

Source: Pamela F. Olson former Assistant Secretary for Tax Policy (quote from the Diane Rehm Show • 11/29/04 • http://www.wamu.org/programs/dr/04/11/29.php)

© Copyright 5/22/05 by Data Blueprint - all rights reserved!

"The significant problems we face cannot be solved at the same level of thinking we were at when we created them."
- Albert Einstein

© Copyright 5/22/05 by Data Blueprint - all rights reserved!

XML and the Second-Generation Web

The combination of hypertext and a global Internet started a revolution. A new ingredient, XML, is poised to finish the job

by JON BOSAK and TIM BRAY
(both played crucial roles in the development of XML).

"Give people a few hints, and they can figure out the rest … "

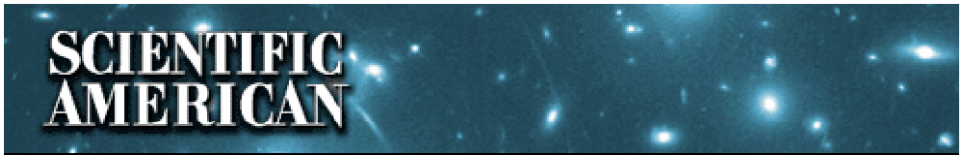
SUBTOPICS:

Something Old, Something New

An End to the World Wide Wait

Some Assembly Required

A Question of Style



Scientific American: Feature Article: XML and the Second Generation Web: May 1999
http://www.sciam.com/1999/0599issue/0599bosak.html

© Copyright 5/22/05 by Data Blueprint - all rights reserved!

XML and the Second-Generation Web

"Give people a few hints, and they can figure out the rest. They can look at this page, see some large type followed by blocks of small type and know that they are looking at the start of a magazine article. They can look at a list of groceries and see shopping instructions. They can look at some rows of numbers and understand the state of their bank account."

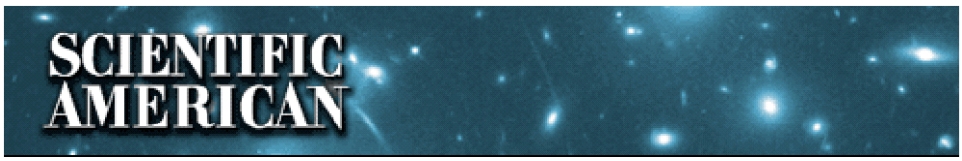
SUBTOPICS:

Something Old, Something New

An End to the World Wide Wait

Some Assembly Required

A Question of Style



Scientific American: Feature Article: XML and the Second Generation Web: May 1999
http://www.sciam.com/1999/0599issue/0599bosak.html

© Copyright 5/22/05 by Data Blueprint - all rights reserved!

Investing in Metadata?

- How can IT staff convince managers to plan, budget, and apply resources for metadata management?
- What is metadata and why is it important?
- What technologies are involved?
- Internet and intranet technologies are part of the answer and will get the immediate attention of management.
- XML is the other technology.

© Copyright 5/22/05 by Data Blueprint - all rights reserved!

USA TODAY

PRINT THIS

Click to Print

SAVE THIS | EMAIL THIS | Close

Terrorist risk lists leave gap, even now

By Mimi Hall, USA TODAY

WASHINGTON — Nearly two years after the Sept. 11 terrorist attacks, key federal agencies have not consolidated a dozen separate "watch lists" intended to keep terrorists out of the country, even though sharing that kind of information might have caught two of the suicide hijackers before they carried out their plot.

The CIA was slow to update the FBI on Khalid al-Midhar and Nawaf Alhazmi, right.

FBI photos via Getty

The Department of Homeland Security says it is working to combine lists of potential security risks maintained by at least nine agencies, but it has no timetable for finishing the job. Officials say critics underestimate the complexity of the task, especially technical problems involving computers and databases not designed to share information. They add that it is important to verify accuracy among lists that often name the same person with different spellings, birth dates or hometowns.

If federal agencies had been sharing information and using a master watch list, "then 9/11 might not have happened," says Ivo Daalder, a homeland-security expert at the Brookings Institution, a think tank in Washington. "I find it criminal that it hasn't happened yet."

© Copyright 5/22/05 by Data Blueprint - all rights reserved!

September 21, 2004

CNN.com

MEMBER SERVICES

SEARCH The Web CNN.com Search

Home Page
World
U.S.
Weather
Business
Sports
Politics
Law
Technology
Science & Space
Health
Entertainment
Travel
Education
Special Reports

SERVICES
Video
E-mail Newsletters
Your E-mail Alerts
CNNtoGO
Contact Us

SEARCH
Web CNN.com

U.S.

Flight diverted after Cat Stevens found on watch list

Tuesday, September 21, 2004 Posted: 10:18 PM EDT (0218 GMT)

WASHINGTON (AP) -- A plane bound for Washington from London was diverted to Maine on Tuesday after passenger Yusuf Islam -- formerly known as pop singer Cat Stevens -- showed up on a U.S. watch list, federal officials said.

United Airlines Flight 919 had already taken off from London en route to Dulles International Airport when the match was made between the passenger and the watch list, said Nico Melendez, a spokesman for the Transportation Security Administration.

Yusuf Islam, formerly Cat Stevens, addresses those gathered at the annual meeting of the Islamic Society of North America in 2003.

© Copyright 5/22/05 by Data Blueprint - all rights reserved!

Hmm …

Confusion

Correct Name:
Yusuf Islam

TSA No Fly Listing:
Youssouf Islam

XML Examples - Which is Which?

XML Example Document #1

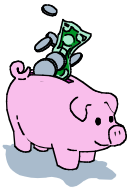
```
<?XML version="1.0"?>
<DOCUMENT>
<CUSTOMER>
<NAME>
<LASTNAME>Edwards</LASTNAME>
<FIRSTNAME>Britta</FIRSTNAME>
</NAME>
<DATE>April 17, 1998</DATE>
<ORDERS>
<ITEM>
<PRODUCT>Cucumber</PRODUCT>
<NUMBER>5</NUMBER>
<PRICE>1.25</PRICE>
</ITEM>
<ITEM>
<PRODUCT>Lettuce</PRODUCT>
<NUMBER>2</NUMBER>
<PRICE>.98</PRICE>
</ITEM>
</ORDERS>
</CUSTOMER>
</DOCUMENT>
```

XML Example Document #2

```
<?XML version="1.0"?>
<SCHEDULE>
<AIRLINE>NorthWest</AIRLINE>
<FLIGHT>
<NUMBER>449</NUMBER>
<STATUS>Cancelled</STATUS>
</FLIGHT>
<FLIGHT>
<NUMBER>640</NUMBER>
<STATUS depart="0100">Delayed</STATUS>
</FLIGHT>
<AIRLINE>TWA</AIRLINE>
<FLIGHT>
<NUMBER>1010</NUMBER>
<STATUS gate="17 Gold">On Time</STATUS>
</FLIGHT>
</SCHEDULE>
```


© Copyright 5/22/05 by Data Blueprint - all rights reserved!

XML Integration



- Until XML, it was cheaper to **develop** one-time solutions to data interface/exchange challenges than it was to develop programmatic solutions (no one measured **maintenance** costs)
- Organizations developed **point-to-point** solutions
- While cheaper to implement - these interfaces are proving to be major **barriers** associated with organizational technological evolution ability
- XML equips organizations with the tools to develop **programmatic solutions** to manage their data interchange environments using the same economies of scale associated with their data management environments

© Copyright 5/22/05 by Data Blueprint - all rights reserved!

Legal

- Data Quality Act
- Clinger-Cohen
- Sarbanes-Oxley
- Basel II Capital Accord
- California Senate Bill 1386
- USA Patriot Act

One Hundred Seventh Congress
of the
United States of America

AT THE SECOND SESSION

m and held at the City of Washington on Wednesday, twenty-third day of January, two thousand and two

An Act

vestors by improving the accuracy and reliability of corporate disclosures nade pursuant to the securities laws, and for other purposes.

enacted by the Senate and House of Representatives of States of America in Congress assembled,

SHORT TITLE; TABLE OF CONTENTS.

IORT TITLE.—This Act may be cited as the "Sarbanes- of 2002".

BLE OF CONTENTS.—The table of contents for this Act is as follows:

Sec. 1. Short title; table of contents.
Sec. 2. Definitions.
Sec. 3. Commission rules and enforcement.

TITLE I—PUBLIC COMPANY ACCOUNTING OVERSIGHT BOARD

Data Quality Act of 2002

1. Issue data quality guidelines “ensuring and maximizing the quality, objectivity, utility and integrity of information (including statistical information) disseminated by the agency;
2. Establish administrative mechanisms allowing affected persons to seek and obtain correction of information maintained and disseminated by the agency that does not comply with the data quality guidelines; and
3. Periodically report on the number and nature of complaints received by the agency regarding the accuracy of information disseminated by the agency, and how such complaints were handled by the agency.
   - http://library.lp.findlaw.com/articles/file/00312/008569/title/features – accessed on 10/1/03.

© Copyright 5/22/05 by Data Blueprint - all rights reserved!

Clinger-Cohen Act of 1996

- Formerly called Information Technology Management Reform Act(ITMRA)
- Defines an information technology architecture as an integrated framework for evolving or maintaining existing IT and acquiring new IT to achieve the agency's strategic and information resources management goals.
- Link IT investments to agency accomplishments
- Requires that agency heads establish a process to select, manage and control their IT investments
  - http://www.intervista-institute.com/resources/clinger-cohen-act.html

© Copyright 5/22/05 by Data Blueprint - all rights reserved!

Sarbanes-Oxley Act of 2002

- Enacted in the wake of Enron, etc.
- Intended to protect investors by improving the accuracy and reliability of corporate disclosures
- Radically redesigns federal regulation of public company corporate governance and reporting obligations.
- Significantly tightens accountability standards for directors and officers, auditors, securities analysts and legal counsel.
- Creates an obligation for accountants to maintain corporate audit records or review work papers for a period of 5 years from the end of the fiscal period during which the audit or review was concluded
- Provides for the imposition of fines, imprisonment up to 20 years, or both.
  - American Institute of Certified Public Accountants maintains a Sarbanes-Oxley Act hotline at 866-265-1977 http://www.aicpa.org/sarbanes/index.asp

© Copyright 5/22/05 by Data Blueprint - all rights reserved!

Basel II Capital Accord

- Amended regulatory framework that has been developed by the Bank of International Settlements
- Requires all internationally active banks, at every tier within the banking group, to adopt similar or consistent risk-management practices for tracking and publicly reporting exposure to operational, credit and market risks.
- "Technological changes brought about by Basel II provide a unique opportunity for financial services organizations to build an enterprise in which all business systems are connected, available and secure."
  - http://www.bis.org/publ/bcbsca.htm

© Copyright 5/22/05 by Data Blueprint - all rights reserved!

California Senate Bill 1386

- Took effect July 1, 2003
- Requires that companies conducting business in that state notify California residents if they know or reasonably believe there has been a breach in security that might have put their personal information at risk
- To enable customers to take protective measures -- such as informing their credit card companies - to minimize any damage
- If you have customers in California, own or license data containing personal information on such residents, or even maintain such data on behalf of another company, this law affects you.
- If your company suspects a breach and fails to quickly notify customers who suffer as a result, the law authorizes them to institute civil actions against your company to recover damages.
  - http://www.computerworld.com/printthis/2003/0,4814,85186,00.html
  - ChoicePoint (145,000) http://choicepoint.com/news/statement_0205_1.html#sub1
  - Bank of America (1M Federal employees) http://www.npr.org/templates/story/story.php?storyId=4514805

© Copyright 5/22/05 by Data Blueprint - all rights reserved!

Bad Credit Reporting Data

© Copyright 5/22/05 by Data Blueprint - all rights reserved!

USA Patriot Act

- Uniting and Strengthening America by Providing Appropriate Tools Required to Intercept and Obstruct Terrorism
- The greatest challenge the Patriot Act presents is updating or implementing new systems.
- Identifying, interpreting and integrating data residing in multiple locations is the most challenging task
- "The ability to secure nationwide warrants to obtain e-mail and electronic evidence"
  - http://www.startribune.com/stories/1576/4122822.html (see also) http://www.eff.org/Privacy/Surveillance/Terrorism_militias/20011031_eff_usa_patriot_analysis.php

© Copyright 5/22/05 by Data Blueprint - all rights reserved!

Legal Confusion?

- "All these [regulations] deal with data, so there's going to be an IT implication. But I think it's a pragmatic response at this time for organizations not to overreact. No one wants to go overboard and then realize they did too much"
- If relevant data is inadvertently destroyed, the company can be charged with "spoliation of evidence," and consequences can be as severe as if your company had defied a court order
  - John Hagerty, analyst at AMR Research Inc.

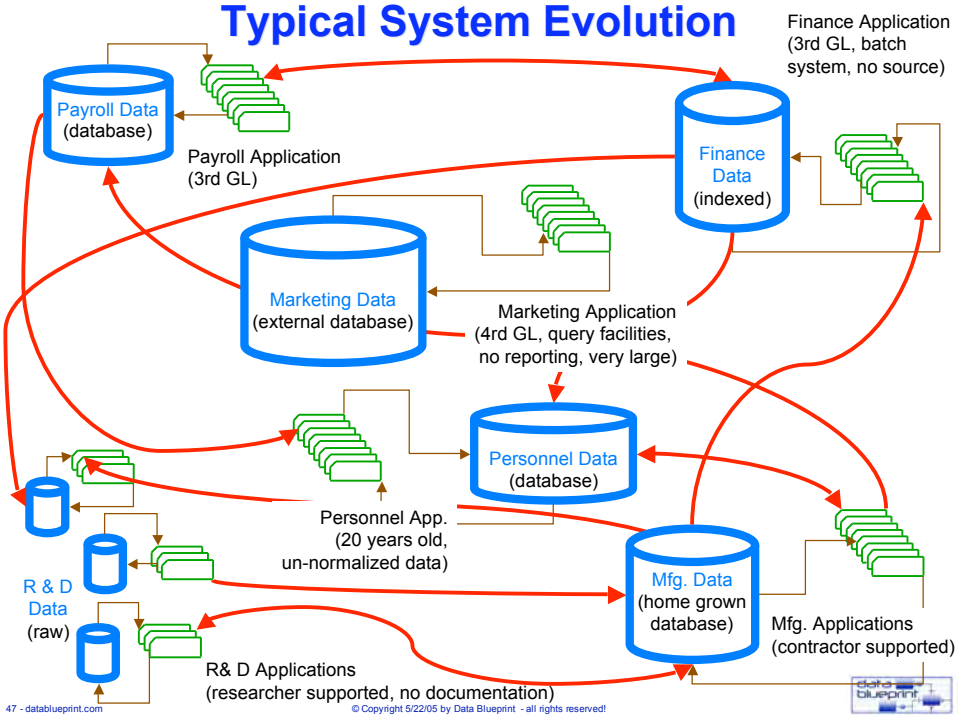
© Copyright 5/22/05 by Data Blueprint - all rights reserved!

Do you know the game Twister?

- Canada
- Chile
- Columbia
- Egypt
- Estonia
- Finland
- France
- Germany
- Great Britain
- Ireland
- Italy
- Japan
- Scotland
- Switzerland
- Thailand
- Turkey
- UAE
- US

© Copyright 5/22/05 by Data Blueprint - all rights reserved!

Typical System Evolution



© Copyright 5/22/05 by Data Blueprint - all rights reserved!

http://www.sas.com

© Copyright 5/22/05 by Data Blueprint - all rights reserved!

Data Integration/Exchange Challenges

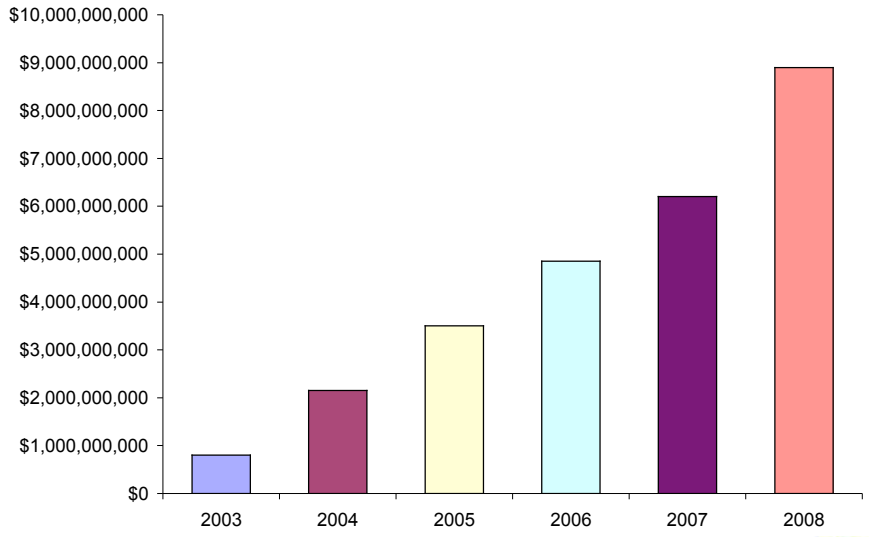
- **Customer** typically has had different meanings to different parts of the organization:
  - Accounting -> organization that buys products or services
  - Service -> client
  - Sales -> prospect
- Assigning the same mission to the DoD 'lines of business' to: "***Secure the building***" elicits very different results from each 'line of business':
  - **Army**: Posts guards at all entrances and ensures no unauthorized access
  - **Navy**: Turns out all the lights, locks up, and leaves
  - **Marines**: Sends in a company to clear the building room-by-room; forms perimeter defense around the building
  - **Air Force**: Signs three year lease with option to buy

[Second example courtesy of Burt Parker]

© Copyright 5/22/05 by Data Blueprint - all rights reserved!

Global Spending on Integration, Forecast to Increase at 169% annual rate

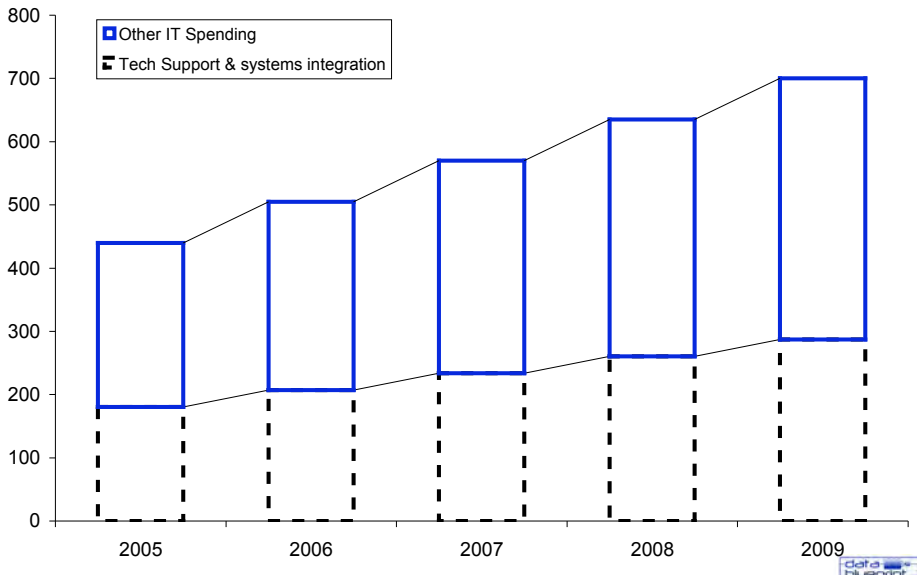
Total Market for intra- and inter-enterprise integration



Statistics courtesy International Data Corp. via *DM Review* (.com) 8/01

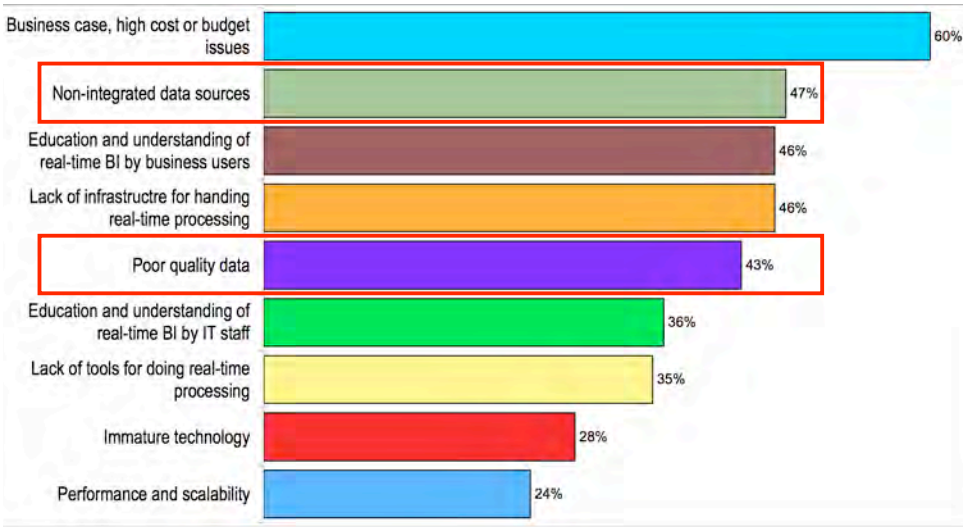
Rising Integration Costs

Global IT Spending Forecast



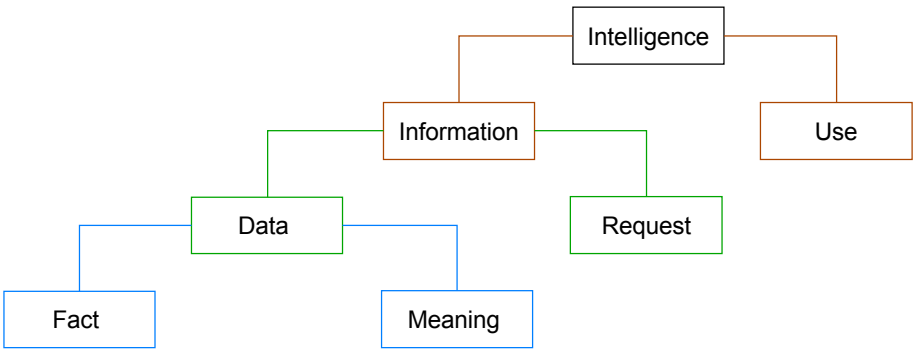
© Copyright 5/22/05 by Data Blueprint - all rights reserved!

Obstacles to Real-Time BI-Lessons from Deployment



TDWI *The Real Time Enterprise Report*, 2003

A model precisely defining the relationship between the terms BI, INFORMATION, FACT, MEANING, and DATA



1. Each FACT combines with one or more MEANINGS.
2. Each specific FACT and MEANING combination is referred to as a DATUM.
3. An INFORMATION is one or more DATA that are returned in response to a specific REQUEST
4. INFORMATION REUSE is enabled when one FACT is combined with more than one MEANING.
5. BI is INFORMATION associated with its USES.

[Built on definition by Dan Appleton 1983]

© Copyright 5/22/05 by Data Blueprint - all rights reserved!

The new economy is the data economy

The Survey

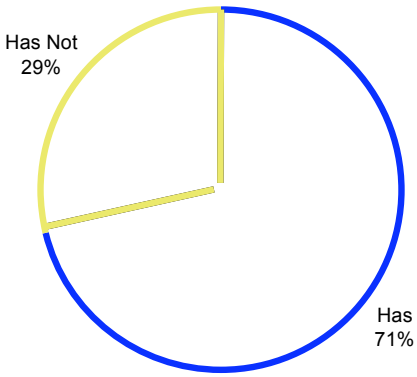
- PriceWaterhouseCoopers interviewed 600 the Chief Information Officer (or equivalent)
- Mix of Top 500, middle-market, and e-business
- Cross-section of corporate activity across the in US, Australia and UK
- Assess
  - Managements' awareness of the crucial importance of data quality
  - Current level of competence in data management.

PriceWaterhouseCoopers, Global Risk Management Solutions, Global Data Management Survey 2001: the new economy is the data economy, 5/23/01 accessed on 8/8/01 http://www.pwcglobal.com/Extweb/service.nsf/8b9d788097dff3c9852565e00073c0ba/0cae4d32ccbaba2380256a0d004e454b/$FILE/Data+Management+brochure.pdf

© Copyright 5/22/05 by Data Blueprint - all rights reserved!

More than 2/3rds have formal data management training

Have you or have you not had formal data management training?



© Copyright 5/22/05 by Data Blueprint - all rights reserved!

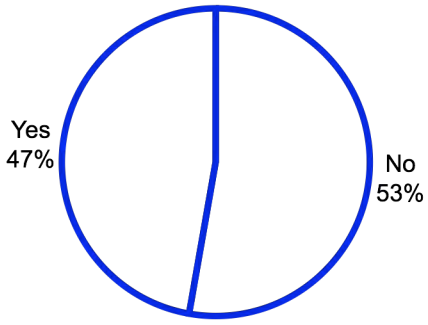
Slightly over two-thirds of organizations used or were planning to apply formal metadata management

My organization used or is planning to use a formal approach to data management

© Copyright 5/22/05 by Data Blueprint - all rights reserved!

Less than 1/2 manage their metadata using cataloging tools

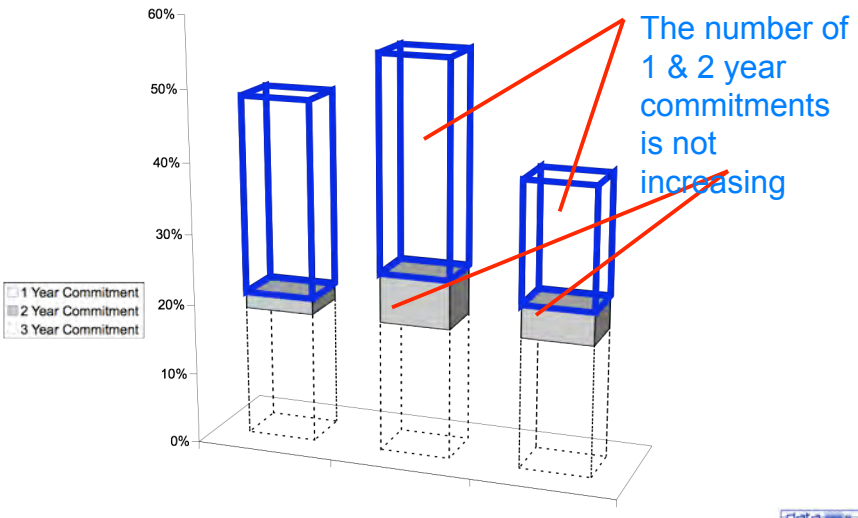
Do you use any metadata catalog tools?



© Copyright 5/22/05 by Data Blueprint - all rights reserved!

On average, less than half (48%) have plans to use repositories

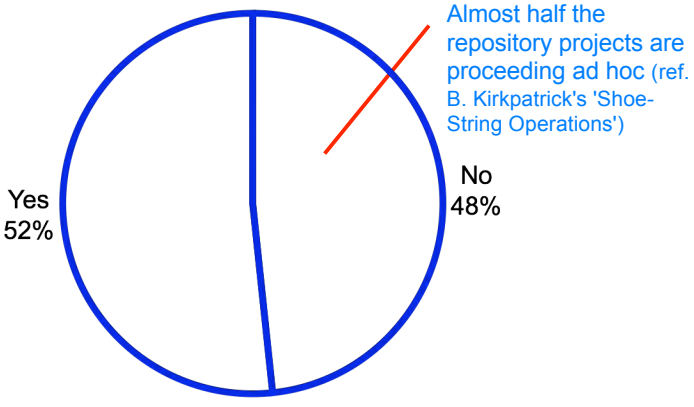
My organization used or is planning to use a repository



© Copyright 5/22/05 by Data Blueprint - all rights reserved!

Slightly over half of repository builders use a structured approach to repository development

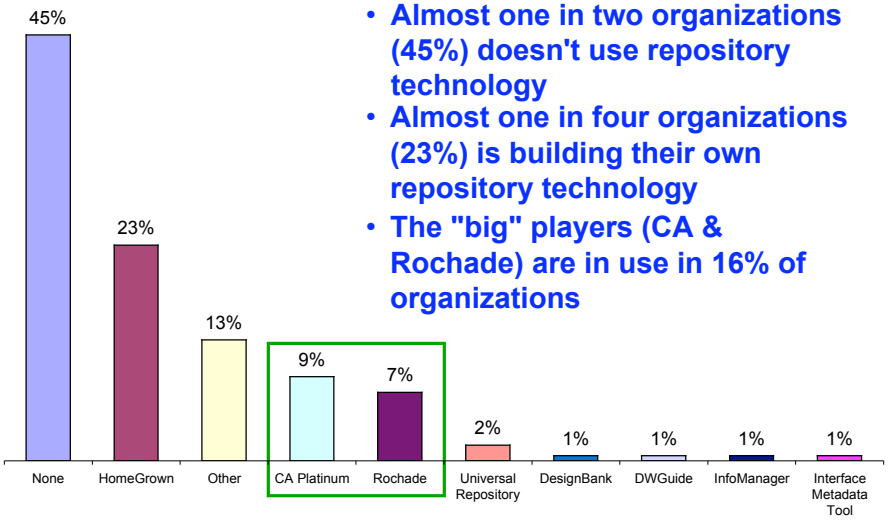
If yes, does your organization use a structured or formal approach to repository technology?



© Copyright 5/22/05 by Data Blueprint - all rights reserved!

Repository Technologies in Use

What tools do you use?

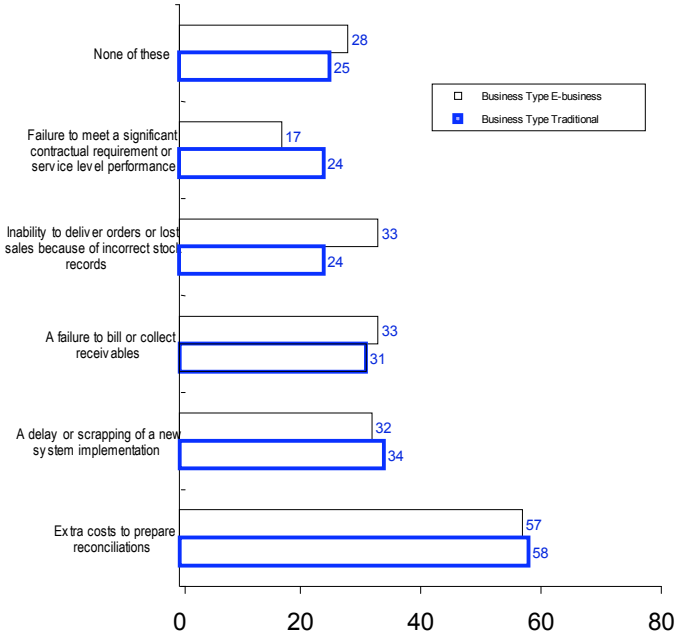


- **Almost one in two organizations (45%) doesn't use repository technology**
- **Almost one in four organizations (23%) is building their own repository technology**
- **The "big" players (CA & Rochade) are in use in 16% of organizations**

© Copyright 5/22/05 by Data Blueprint - all rights reserved!

The new economy is the data economy

Have we suffered significant problems, costs or losses in any area because of poor data quality?"

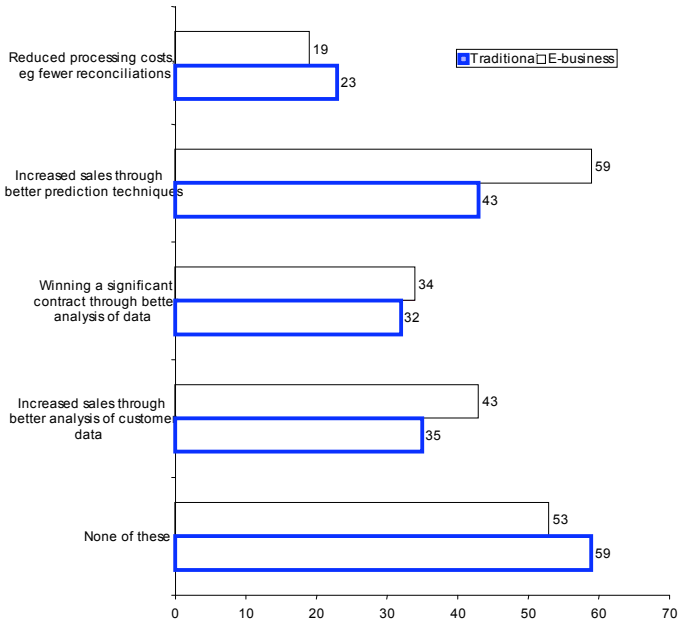


PriceWaterhouseCoopers, Global Risk Management Solutions, Global Data Management Survey 2001: the new economy is the data economy, 5/23/01 accessed on 8/8/01 http://www.pwcglobal.com/Extweb/service.nsf/8b9d788097dff3c9852565e00073c0ba/0cae4d32ccbaba2380256a0d004e454b/$FILE/Data+Management+brochure.pdf

© Copyright 5/22/05 by Data Blueprint - all rights reserved!

The new economy is the data economy

Three quarters have benefited from effective data management

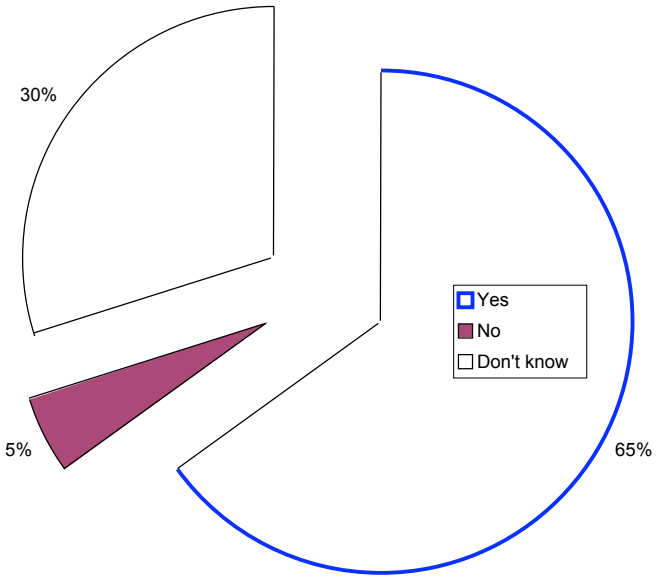


PriceWaterhouseCoopers, Global Risk Management Solutions, Global Data Management Survey 2001: the new economy is the data economy, 5/23/01 accessed on 8/8/01 http://www.pwcglobal.com/Extweb/service.nsf/8b9d788097dff3c9852565e00073c0ba/0cae4d32ccbaba2380256a0d004e454b/$FILE/Data+Management+brochure.pdf

© Copyright 5/22/05 by Data Blueprint - all rights reserved!

The new economy is the data economy

Increase in number of automated decisions and processes based on electronic data over past two years

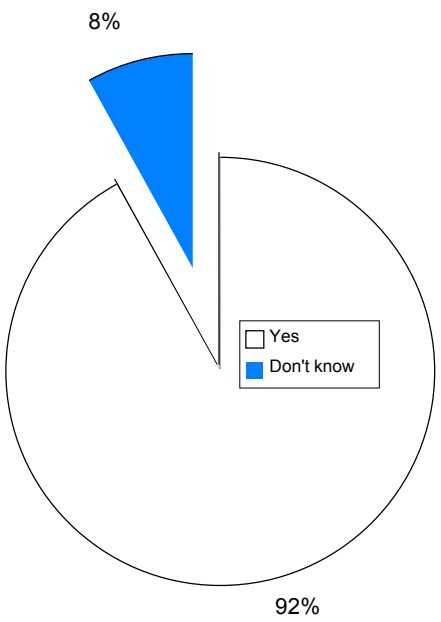


PriceWaterhouseCoopers, Global Risk Management Solutions, Global Data Management Survey 2001: the new economy is the data economy, 5/23/01 accessed on 8/8/01 http://www.pwcglobal.com/Extweb/service.nsf/8b9d788097dff3c9852565e00073c0ba/0cae4d32ccbaba2380256a0d004e454b/$FILE/Data+Management+brochure.pdf

© Copyright 5/22/05 by Data Blueprint - all rights reserved!

The new economy is the data economy

Further growth in the next two years in number of automated decisions and processes?



PriceWaterhouseCoopers, Global Risk Management Solutions, Global Data Management Survey 2001: the new economy is the data economy, 5/23/01 accessed on 8/8/01 http://www.pwcglobal.com/Extweb/service.nsf/8b9d788097dff3c9852565e00073c0ba/0cae4d32ccbaba2380256a0d004e454b/$FILE/Data+Management+brochure.pdf

© Copyright 5/22/05 by Data Blueprint - all rights reserved!

The new economy is the data economy

- Does your company have a shared information system or is information kept within individual departments?

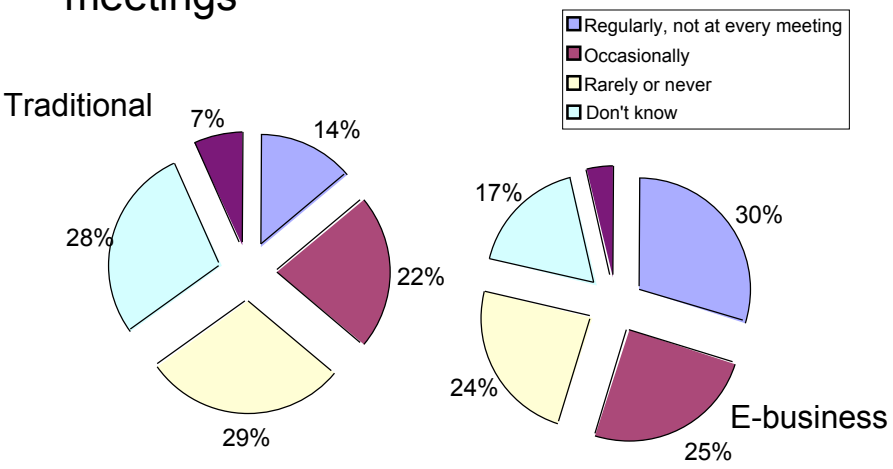
E-business	65% have a shared information system	35% keep info in depts
Traditional	72% have a shared information system	28% keep info in depts

PriceWaterhouseCoopers, Global Risk Management Solutions, Global Data Management Survey 2001: the new economy is the data economy, 5/23/01 accessed on 8/8/01 http://www.pwcglobal.com/Extweb/service.nsf/8b9d788097dff3c9852565e00073c0ba/0cae4d32ccbaba2380256a0d004e454b/$FILE/Data+Management+brochure.pdf

© Copyright 5/22/05 by Data Blueprint - all rights reserved!

The new economy is the data economy

- Data management discussed at board meetings

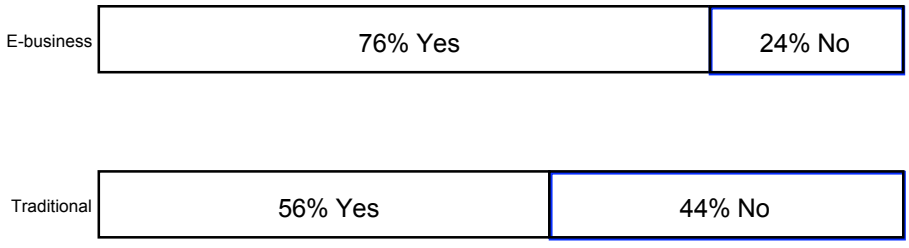


PriceWaterhouseCoopers, Global Risk Management Solutions, Global Data Management Survey 2001: the new economy is the data economy, 5/23/01 accessed on 8/8/01 http://www.pwcglobal.com/Extweb/service.nsf/8b9d788097dff3c9852565e00073c0ba/0cae4d32ccbaba2380256a0d004e454b/$FILE/Data+Management+brochure.pdf

© Copyright 5/22/05 by Data Blueprint - all rights reserved!

The new economy is the data economy

- Do you believe that senior management of your company places sufficient importance on data management issues?

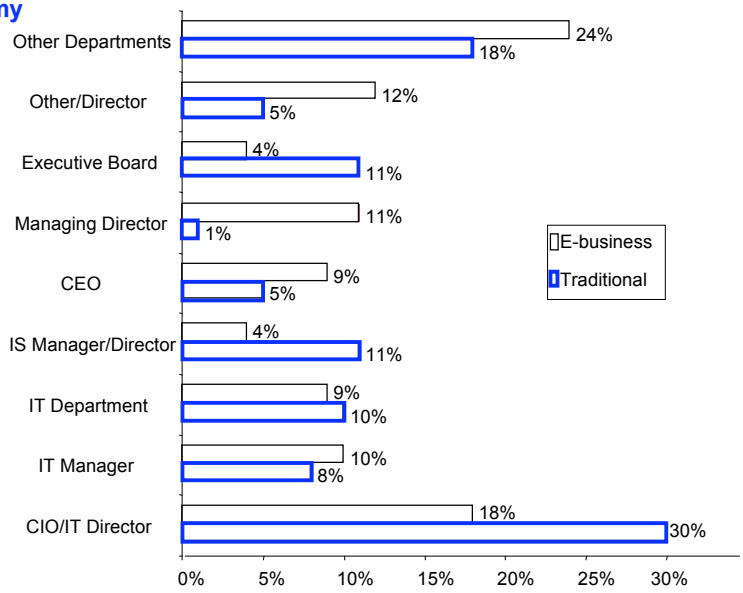


PriceWaterhouseCoopers, Global Risk Management Solutions, Global Data Management Survey 2001: the new economy is the data economy, 5/23/01 accessed on 8/8/01 http://www.pwcglobal.com/Extweb/service.nsf/8b9d788097dff3c9852565e00073c0ba/0cae4d32ccbaba2380256a0d004e454b/$FILE/Data+Management+brochure.pdf

© Copyright 5/22/05 by Data Blueprint - all rights reserved!

The new economy is the data economy

Today *responsibility of driving data strategy lies mainly with IT professionals in traditional companies, and is more fragmented in e-business*

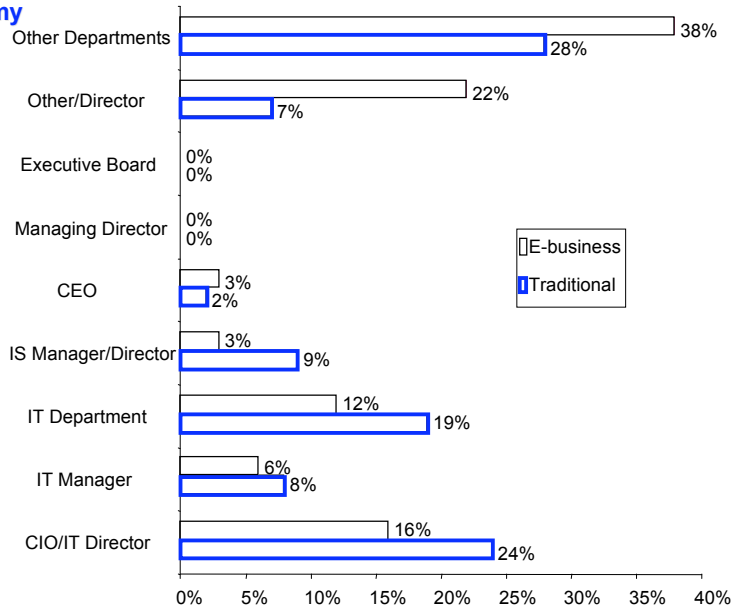


PriceWaterhouseCoopers, Global Risk Management Solutions, Global Data Management Survey 2001: the new economy is the data economy, 5/23/01 accessed on 8/8/01 http://www.pwcglobal.com/Extweb/service.nsf/8b9d788097dff3c9852565e00073c0ba/0cae4d32ccbaba2380256a0d004e454b/$FILE/Data+Management+brochure.pdf

© Copyright 5/22/05 by Data Blueprint - all rights reserved!

The new economy is the data economy

Future *responsibility of driving data strategy lies mainly with IT professionals in traditional companies, and is more fragmented in e-business*

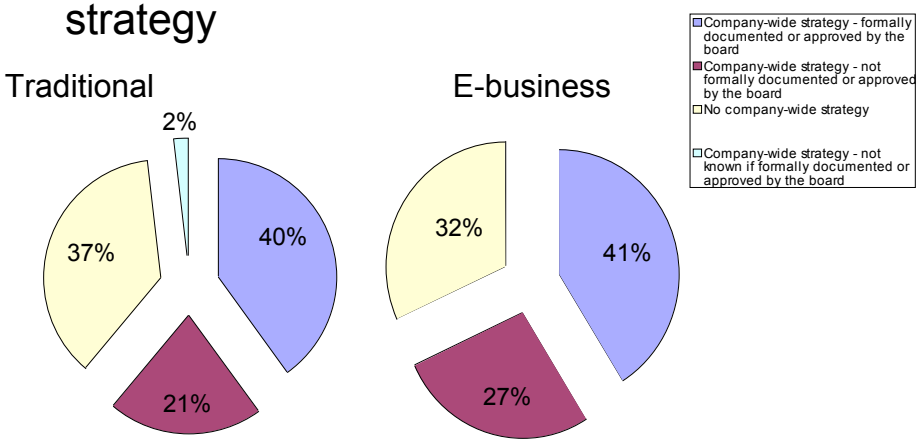


PriceWaterhouseCoopers, Global Risk Management Solutions, Global Data Management Survey 2001: the new economy is the data economy, 5/23/01 accessed on 8/8/01 http://www.pwcglobal.com/Extweb/service.nsf/8b9d788097dff3c9852565e00073c0ba/0cae4d32ccbaba2380256a0d004e454b/$FILE/Data+Management+brochure.pdf

© Copyright 5/22/05 by Data Blueprint - all rights reserved!

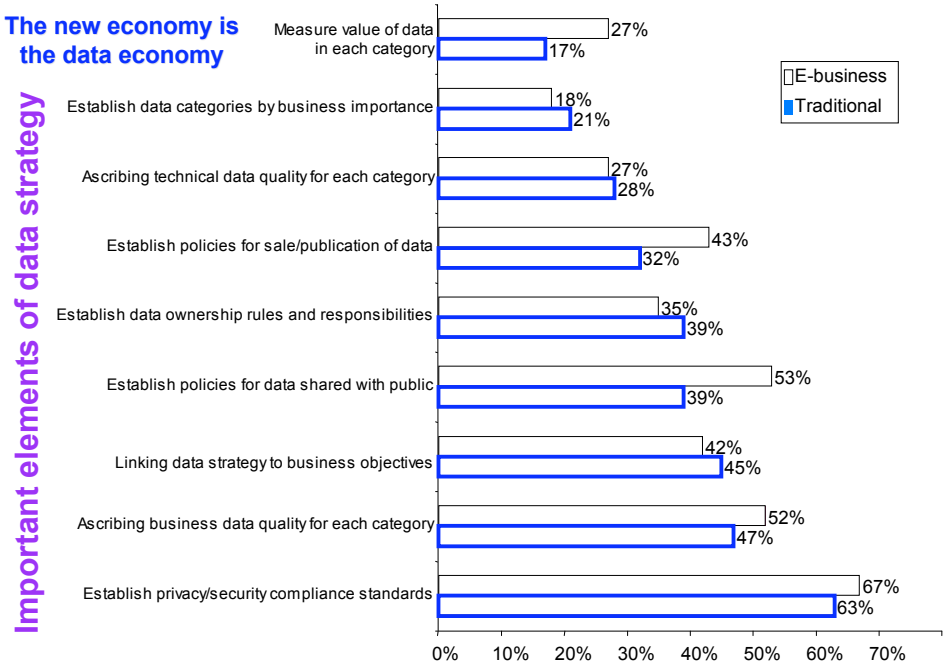
The new economy is the data economy

- Majority of companies do not have a documented board approved data strategy



PriceWaterhouseCoopers, Global Risk Management Solutions, Global Data Management Survey 2001: the new economy is the data economy, 5/23/01 accessed on 8/8/01 http://www.pwcglobal.com/Extweb/service.nsf/8b9d788097dff3c9852565e00073c0ba/0cae4d32ccbaba2380256a0d004e454b/$FILE/Data+Management+brochure.pdf

© Copyright 5/22/05 by Data Blueprint - all rights reserved!

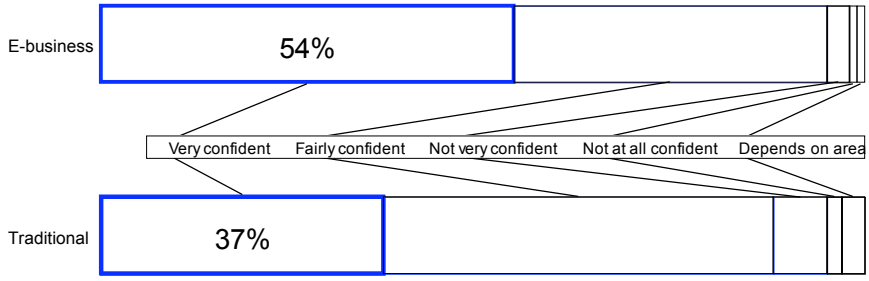


PriceWaterhouseCoopers, Global Risk Management Solutions, Global Data Management Survey 2001: the new economy is the data economy, 5/23/01 accessed on 8/8/01 http://www.pwcglobal.com/Extweb/service.nsf/8b9d788097dff3c9852565e00073c0ba/0cae4d32ccbaba2380256a0d004e454b/$FILE/Data+Management+brochure.pdf

© Copyright 5/22/05 by Data Blueprint - all rights reserved!

The new economy is the data economy

- How confident are you in the quality of the your data?
  - **"Only one in three companies are very confident in the quality of their own data"**



PriceWaterhouseCoopers, Global Risk Management Solutions, Global Data Management Survey 2001: the new economy is the data economy, 5/23/01 accessed on 8/8/01 http://www.pwcglobal.com/Extweb/service.nsf/8b9d788097dff3c9852565e00073c0ba/0cae4d32ccbaba2380256a0d004e454b/$FILE/Data+Management+brochure.pdf

© Copyright 5/22/05 by Data Blueprint - all rights reserved!

The new economy is the data economy

- How confident are you in the quality of the data your company receives from third parties in its e-business activities?
  - **"Only 15% of companies are very confident of the data received from other organizations"**

PriceWaterhouseCoopers, Global Risk Management Solutions, Global Data Management Survey 2001: the new economy is the data economy, 5/23/01 accessed on 8/8/01 http://www.pwcglobal.com/Extweb/service.nsf/8b9d788097dff3c9852565e00073c0ba/0cae4d32ccbaba2380256a0d004e454b/$FILE/Data+Management+brochure.pdf

© Copyright 5/22/05 by Data Blueprint - all rights reserved!

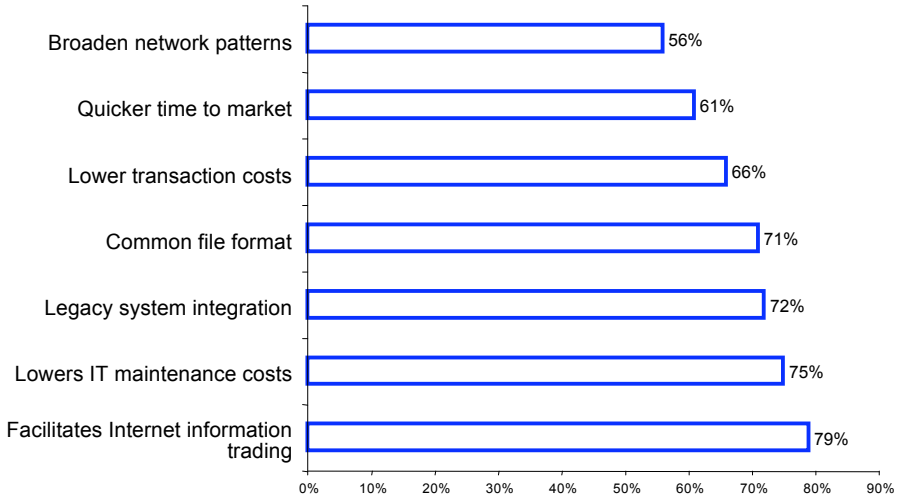
The new economy is the data economy

- *Given these findings, we believe that corporate boards need to seize the initiative on data management, and mount a radical overhaul of their attitudes and objectives. Only through an agreed, documented and rigorously implemented data strategy indivisible from the rest of the company's strategic drivers, with committed leadership from board level and clear lines of responsibility, will a company put itself in a position to create value in the coming years.*

PriceWaterhouseCoopers, Global Risk Management Solutions, Global Data Management Survey 2001: the new economy is the data economy, 5/23/01 accessed on 8/8/01 http://www.pwcglobal.com/Extweb/service.nsf/8b9d788097dff3c9852565e00073c0ba/0cae4d32ccbaba2380256a0d004e454b/$FILE/Data+Management+brochure.pdf

© Copyright 5/22/05 by Data Blueprint - all rights reserved!

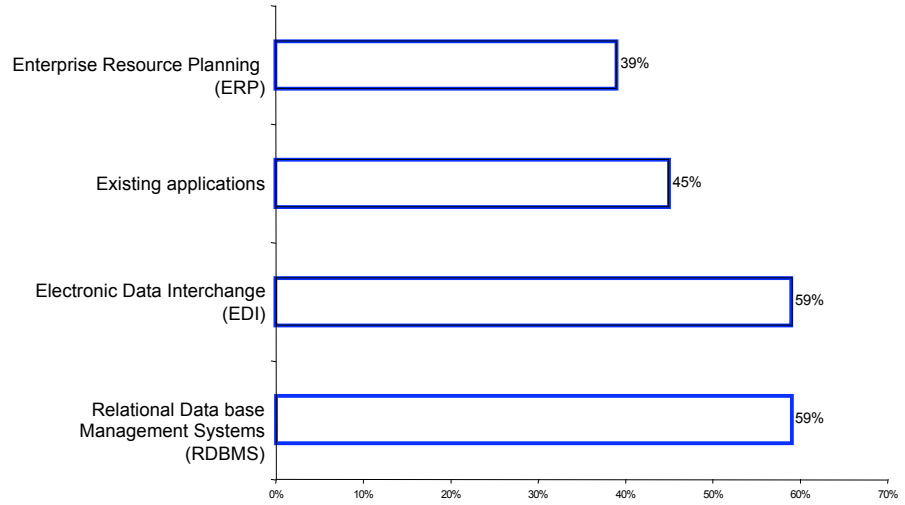
XML Benefits Ratings



Source: CIO Magazine Survey

© Copyright 5/22/05 by Data Blueprint - all rights reserved!

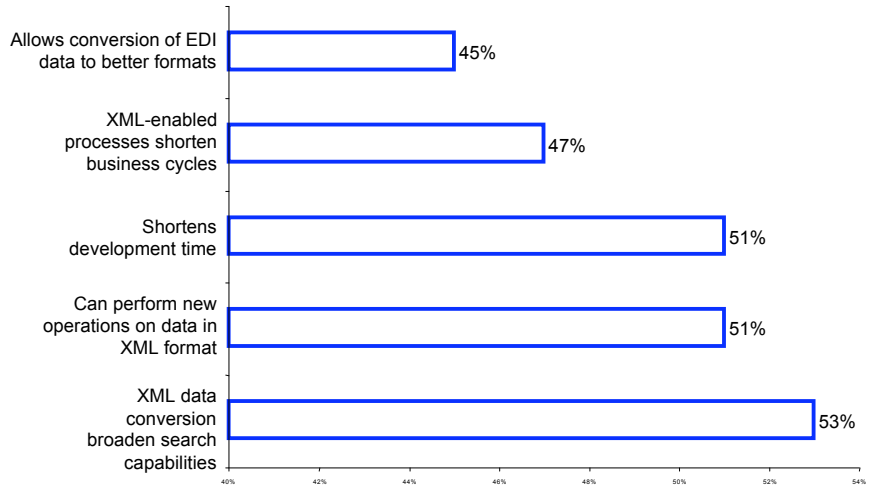
XML Application Integration
(existing or planned)



Source: CIO Magazine Survey

© Copyright 5/22/05 by Data Blueprint - all rights reserved!

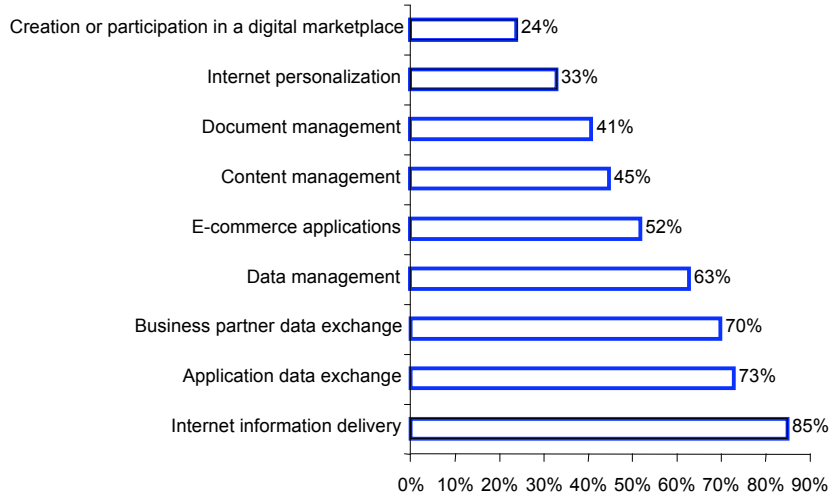
Top 5 reasons for using XML



Source: CIO Magazine Survey

© Copyright 5/22/05 by Data Blueprint - all rights reserved!

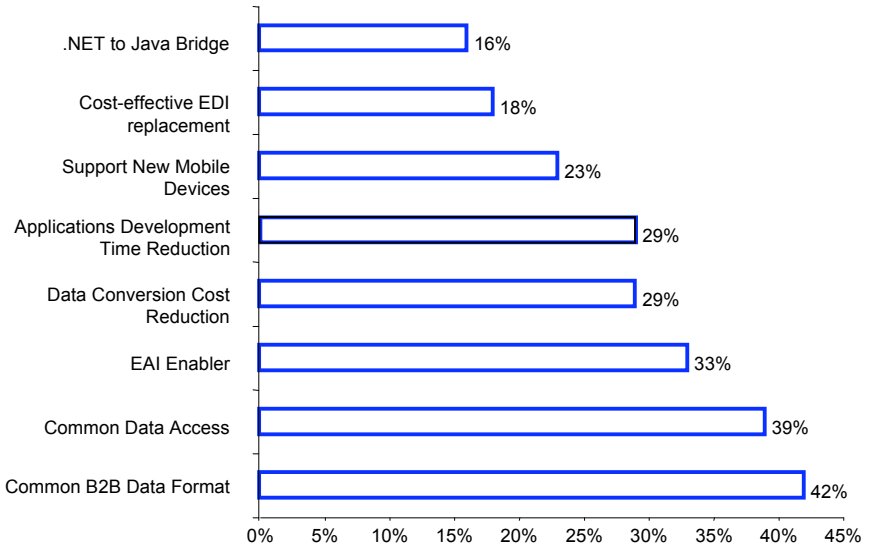
Problems Business Hope XML Will Solve
(flexibility and ease-of-use)



Source: CIO Magazine Survey

© Copyright 5/22/05 by Data Blueprint - all rights reserved!

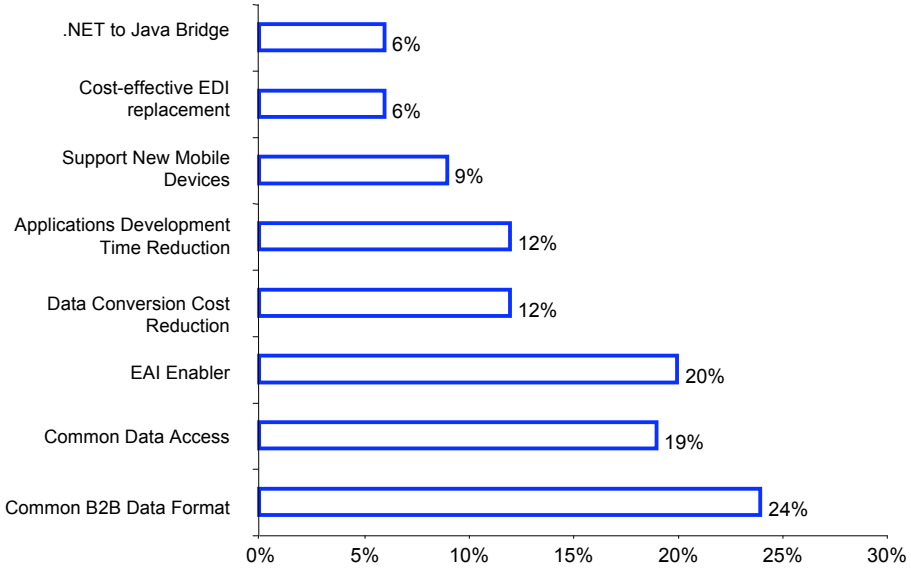
XML Adoption Benefits



[Adapted from Intellor Research Summary - *XML Adoption* - www.intellor.com]

© Copyright 5/22/05 by Data Blueprint - all rights reserved!

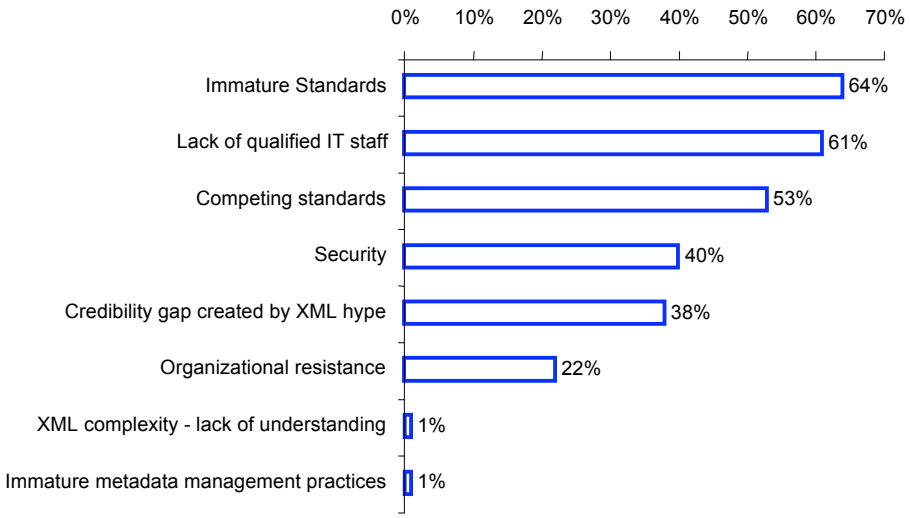
Perception of XML Critical Importance



[Adapted from Intellor Research Summary - *XML Adoption* - www.intellor.com]

© Copyright 5/22/05 by Data Blueprint - all rights reserved!

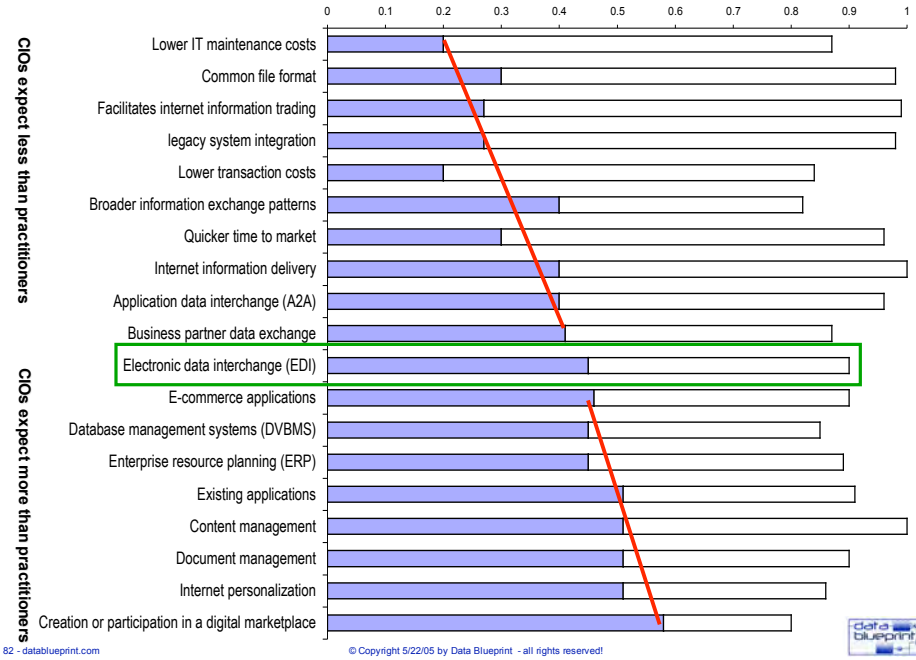
XML Adoption Challenges



[Adapted from Intellor Research Summary - *XML Adoption* - www.intellor.com]

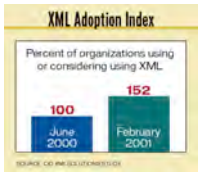
© Copyright 5/22/05 by Data Blueprint - all rights reserved!

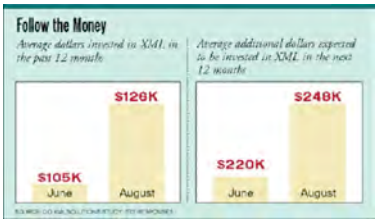
CIO versus Practitioner XML-based Expectations

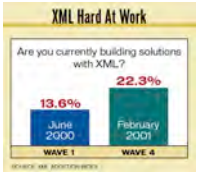


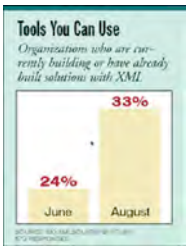
© Copyright 5/22/05 by Data Blueprint - all rights reserved!

XML Adoption Index









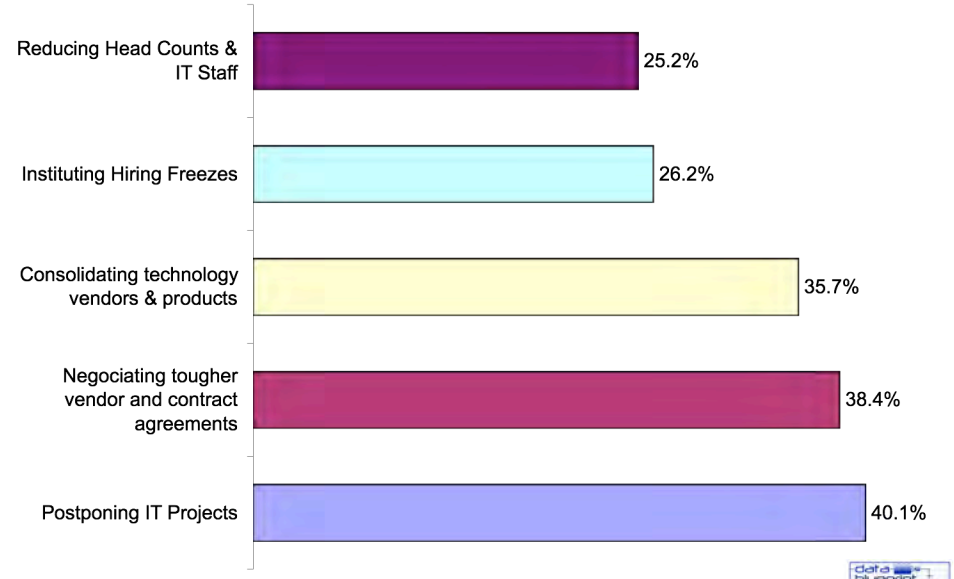
- 47 percent of the IT executives surveyed said they believed XML would help resolve document-management business problems, a 15-percent increase since June
- 70 percent of respondents organizations are using or considering using XML - an eight-percent increase in just three months.
- 15 percent of respondents work at organizations with XML applications already running - a 50 percent increase
- Kaiser Permanente, a leading medical services provider, expects XML to reduce the cost of new interfaces by 50%, decrease time to market of new applications by 50% or better, and lower overall maintenance costs of its portfolio of systems by 30% — all by using XML as a native data format

Source: CIO Magazine Surveys

© Copyright 5/22/05 by Data Blueprint - all rights reserved!

What is missing from this list?

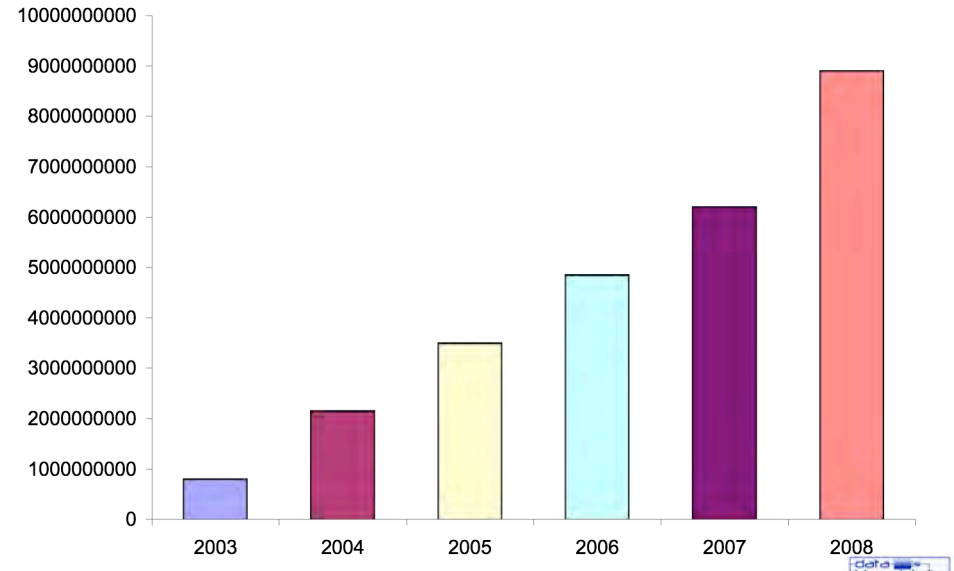
How do CIOs demonstrate reduction in IT expenses?



Quote from L.C. Ware & B. Worthen. via *CIO* (.com) 8/1/01

Global Spending on Integration, Forecast to Increase at 169% annual rate

Total Market for intra- and inter-enterprise integration



Statistics courtesy International Data Corp. via *DM Review* (.com) 8/01

XML Success?

- Product investment defines success
  - XML has increased 10X faster than other IT investments
- Questions instead are:
  - **How fast?**
  - To what level?
  - With what products?

XML Adoption for Document-based Applications by B. Young & R. Clark *STC Conference Proceedings* http://www.stc.org/confproceed/1999/PDFs/024.PDF

and

Growth of XML Spending Outpaces IT Spending at http:doc.advisor.com

© Copyright 5/22/05 by Data Blueprint - all rights reserved!

XML References

- An initial introduction to XML (and also Cascading Style Sheets) is provided by “XML: A Primer” [St Laurent 1998].
- XML used for web site development, with HTML, XSL and XLL, is addressed in “XML: Extensible Markup Language” [Harold 1998].
- “XML Complete” [Holzner 1998] covers the use of XML with Java.
- “Web Farming for the Data Warehouse” [Hackathorn 1998] uses the Internet, Intranets and XML for access to external data sources for warehouse deployment.

© Copyright 5/22/05 by Data Blueprint - all rights reserved!

XML Information Web Sites

BizTalk
- http://www.biztalk.org

WWW Consortium - Specifications and Standards
- http://www.w3.org/

Microsoft XML Scenarios Web Site
- http://microsoft.com/xml/scenario/intro.asp

XML.com Web Site
- http://www.xml.com/

James Tauber's XMLINFO Web Site:
- http://www.xmlinfo.com/

James Clark's XML Web Page
- http://www.jclark.com/xml/

Robin Cover's XML Resources
- http://www.sil.org/sgml/xml.html

Microsoft XML Site
- http://www.microsoft.com/xml/

Microsoft XML Workshop Web Site
- http://www.microsoft.com/workshop/xml/toc.htm

Web Farming Web Site
- http://webfarming.com/

XML 1.0 specification can be found at
- http://www.w3.org/TR/1998/RECCxml-19980210.html

Better version called the "Annotated XML Specification" by Tim Bray at
- http://www.xml.com/axml/testasml.htm

© Copyright 5/22/05 by Data Blueprint - all rights reserved!

The World of XML Tools

Lisa Rein and Jim Woolum • *Finetuning.com*

Introduction
Overviews XML's rapid growth over the past two years.

Companies Developing XML Products
Compares functionality of XML products from more than 40 companies.

XML Products
Hyperlinked descriptions of new products.

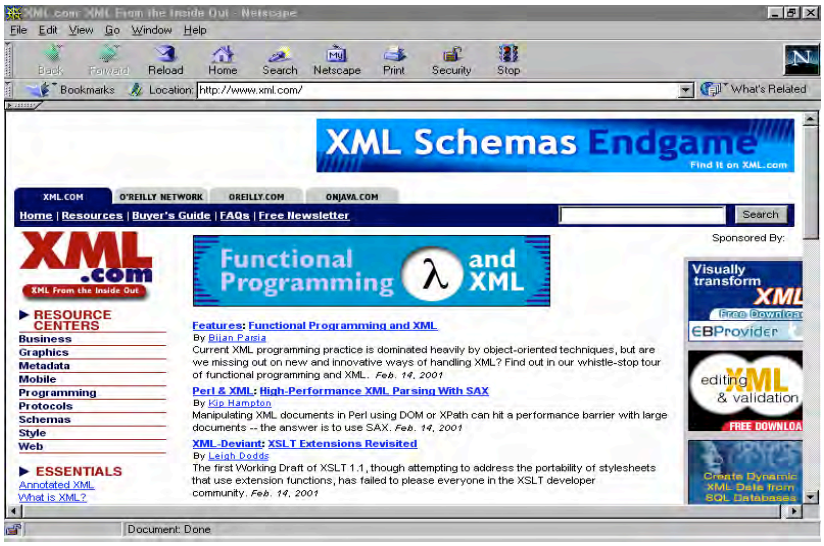
Great XML Resources
Links to specifications for the XML core, e-commerce, accessibility, and graphics/ multimedia, as well as resources for SAX and DOM.

http://computer.org/internet/xml/

© Copyright 5/22/05 by Data Blueprint - all rights reserved!

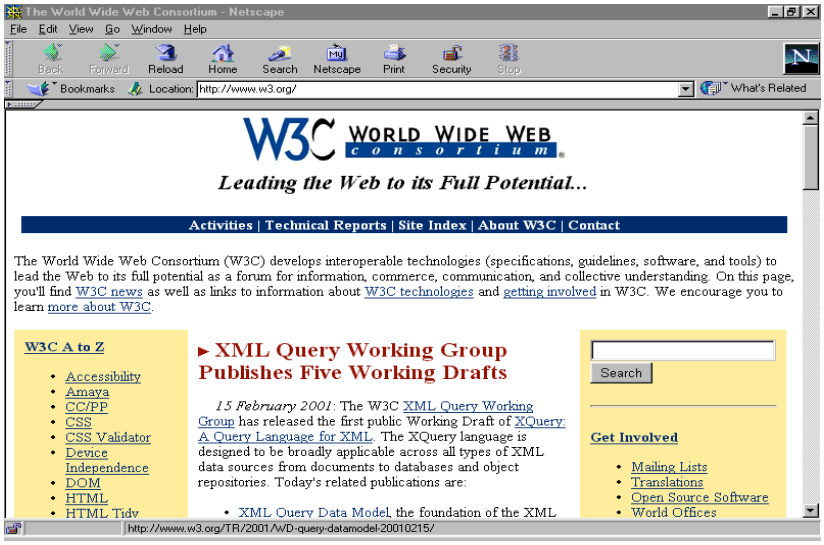
XML Web Resources

XML Web Resources



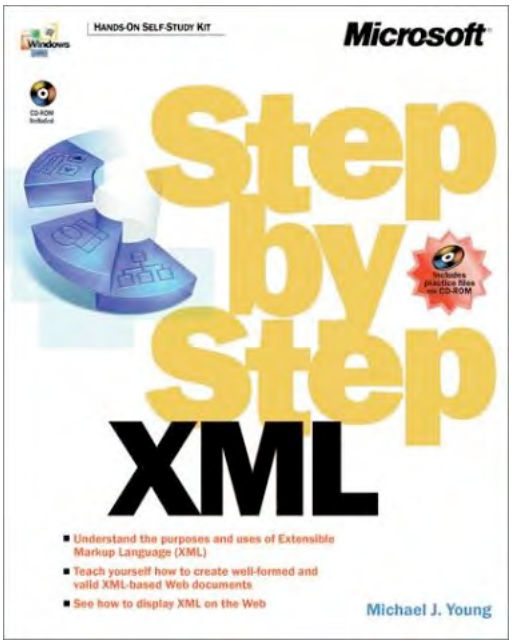
© Copyright 5/22/05 by Data Blueprint - all rights reserved!

XML Web Resources



© Copyright 5/22/05 by Data Blueprint - all rights reserved!

XML Text



© Copyright 5/22/05 by Data Blueprint - all rights reserved!