



PDF and Metadata

Ed Chase
Worldwide Standards
Adobe Systems, Inc.
chase@adobe.com



Overview

- **Metadata**
- **XMP**
- **XMP Applications**
- **Demos**
 - XMP in PDF
 - The Adobe Bridge



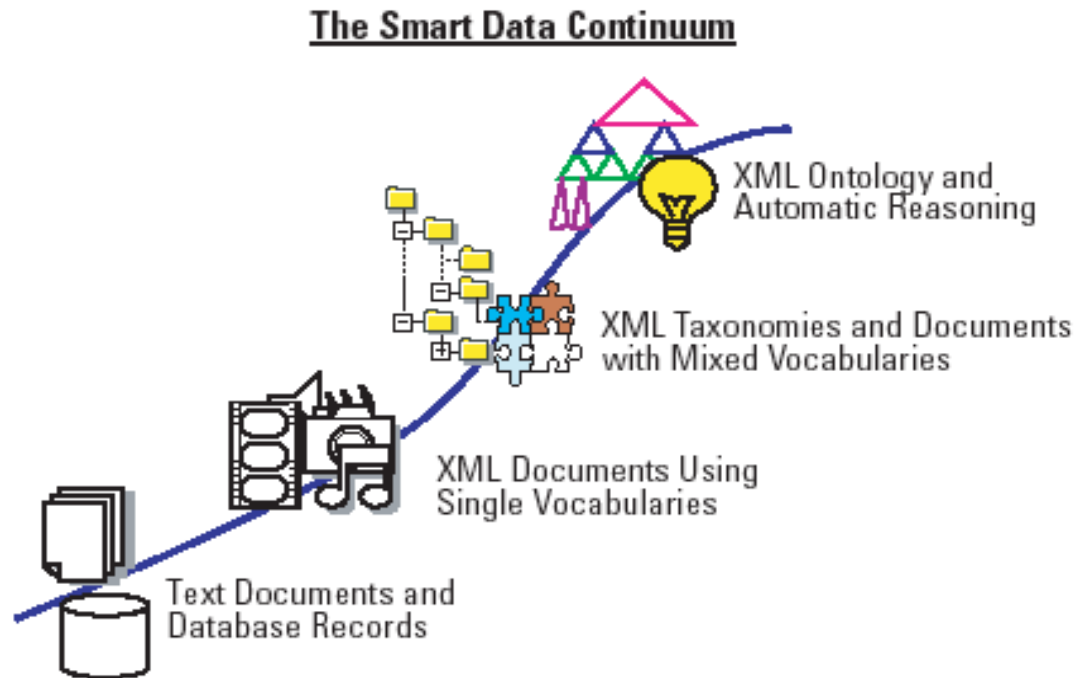
Metadata

- **Information about Information**
- **Promotes machine-readability**
 - Unstructured or opaque information
 - Rich content, images, multimedia, unformatted text
- **Intelligent searching and categorization**
- **Version control, workflow**
- **Better access to information across domains**
 - Methods of categorization vary by industry, company, software, methodology, standards
- *Examples – Phonebook, card-catalog*

Global Information Sharing



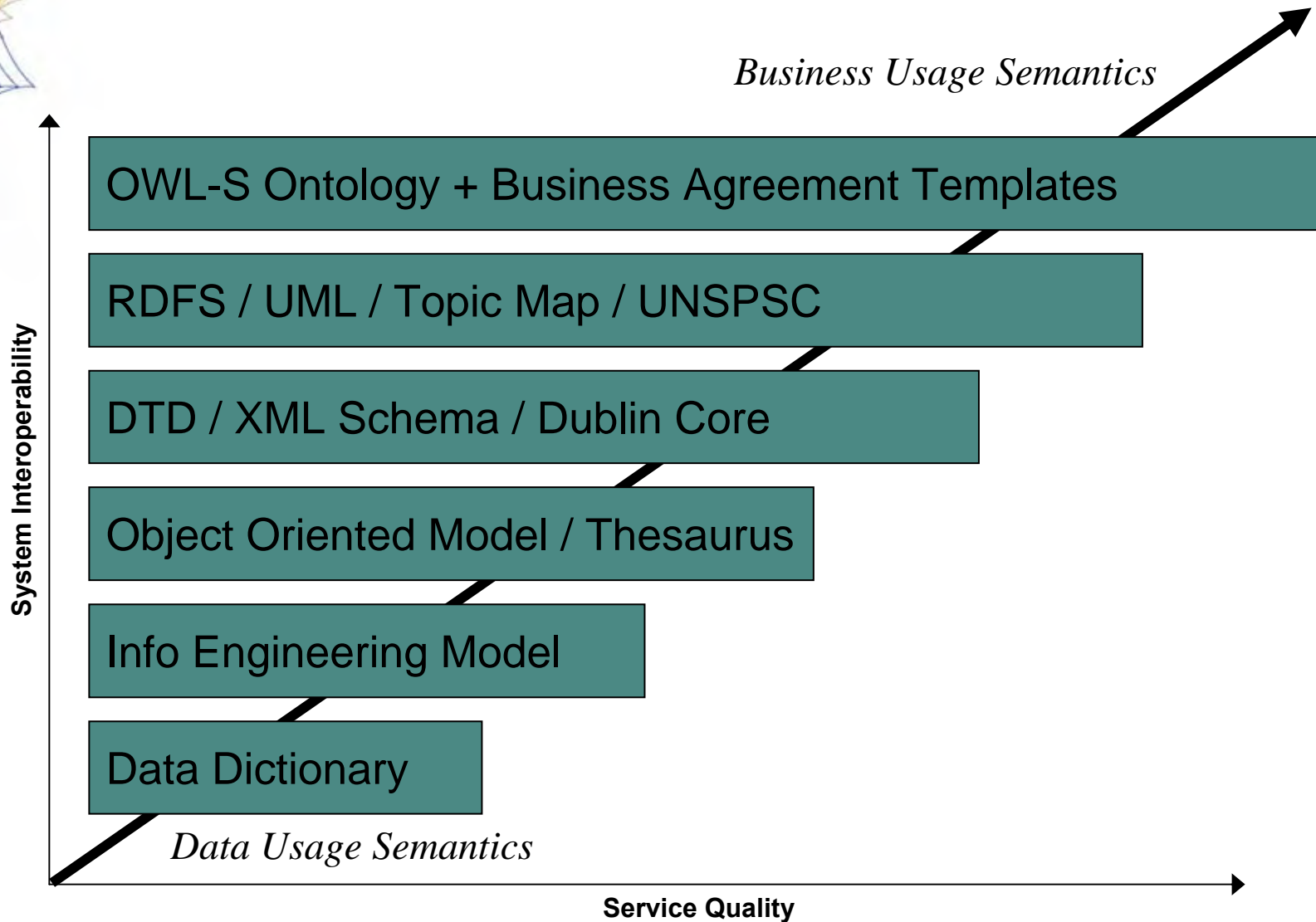
The Smart Data Continuum and the Road to Information Independence



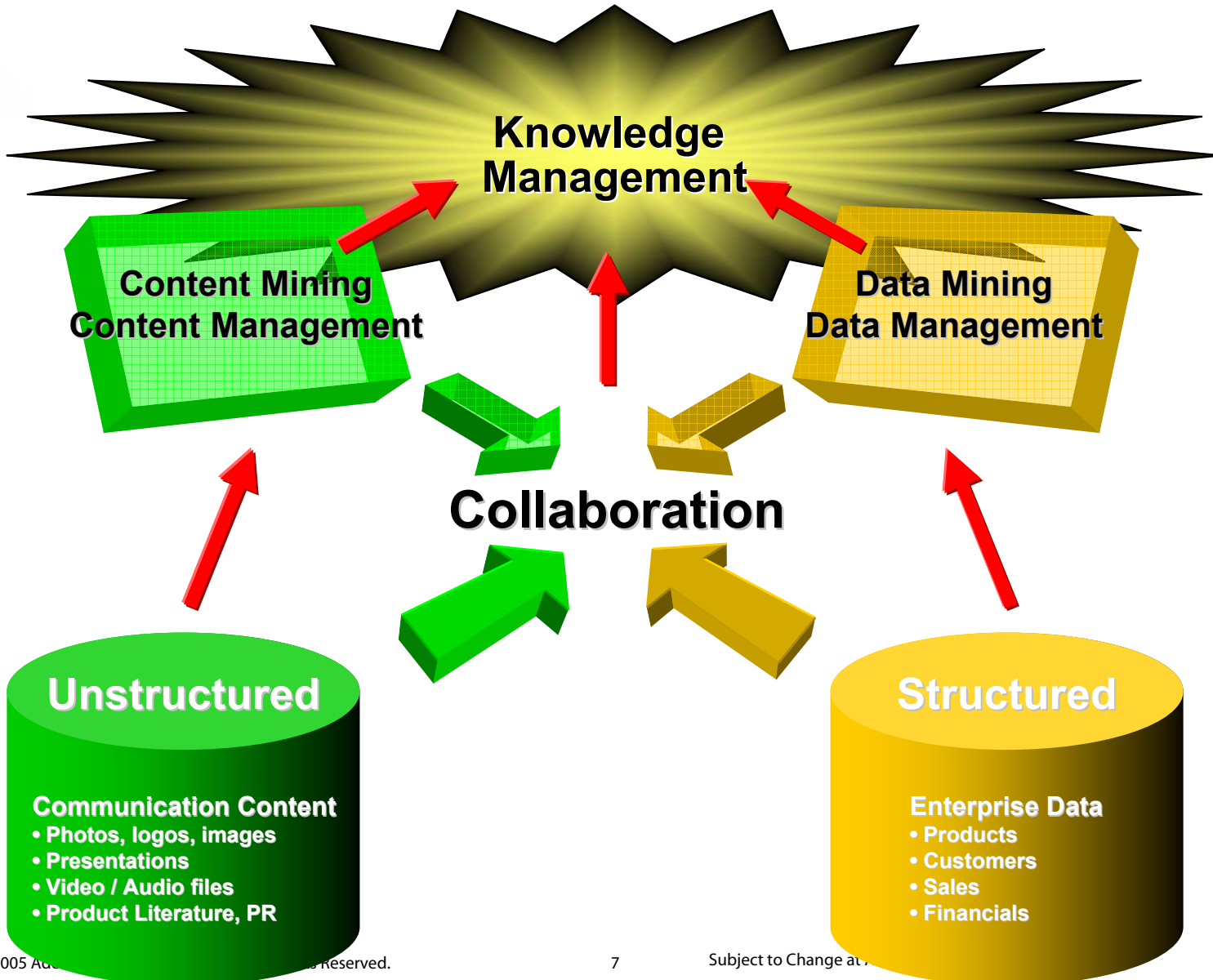
The trend is to put the "smarts" in the data, not in the applications.

Source: "The Semantic Web: A Guide to the Future of XML, Web Services, and Knowledge Management", Wiley Technology Publishing, June 2003.

The Semantic Spectrum: Layers of Wisdom



Confluence of Content and Data





Imagine if your files could talk...

- **You would know where they've been**
 - Store contact info and e-commerce links
- **Whether it's appropriate to use them**
 - License usages, copyright notification
- **They could tell you what happened to them**
 - I've been cropped, rasterized, dithered...
- **Who their relatives and friends are**
 - I'm related to or derivative of this licensed file
- **Where they are going**
 - What expected next in the workflow
- **XMP provides a common language for interchange**

"XMP - adding intelligence to media"

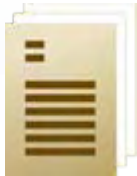


The Need for Intelligent Media

- **Workflows are becoming more digital**
 - Digital camera capture and upload into the production process
 - Desktop publishing and collaboration
 - Output to print, web, CD, handheld, wireless, eBook, RSS, video ...
- **Proliferation of digital assets**
 - Migration from physical and analog to digital
 - Re-use, re-purposing, re-expression of media assets
- **Efficiency is driving data interchange**
 - Rights and usage mgmt between supplier & consumer
 - Standardization for interchange
- **Metadata is a **strategic** advantage**
 - If you can't describe it, you can't control it

Intelligent Information

- **Information should know how to...**



- Act on itself
- Present itself
- Classify itself
- Route itself
- Protect itself
- Leverage new & existing services and standards

- **Within the context of**



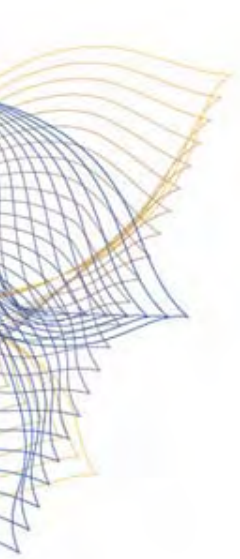
- who's using it
- what they are doing
- and where they are.





Metadata – an Enterprise Requirement

- **SOA & other agile architectures growing**
- **Increasing use of services & registries**
- **Complex enterprise environment requires metadata**
- **Applications include**
 - Entities
 - Relationships
 - Processes
 - Identify Management
 - Data Stores
 - Data
 - & Documents
- **Documents & Forms key to bridging enterprise processes**
- **Site Analyst Quotes here**



Metadata - Requirements

Open	Build on standards
Extensible	Support custom schemas/vocabularies
Universal	Support diverse file formats, databases & relevant containers
International	Unicode, multilingual
Robust	Keep metadata with asset



What is XMP?

- **Technical description**

“Adobe’s Extensible Metadata Platform (XMP) is a labeling technology that allows you to embed data about a file, known as metadata, into the file itself.”

- **Human description**

“The message is in the medium”

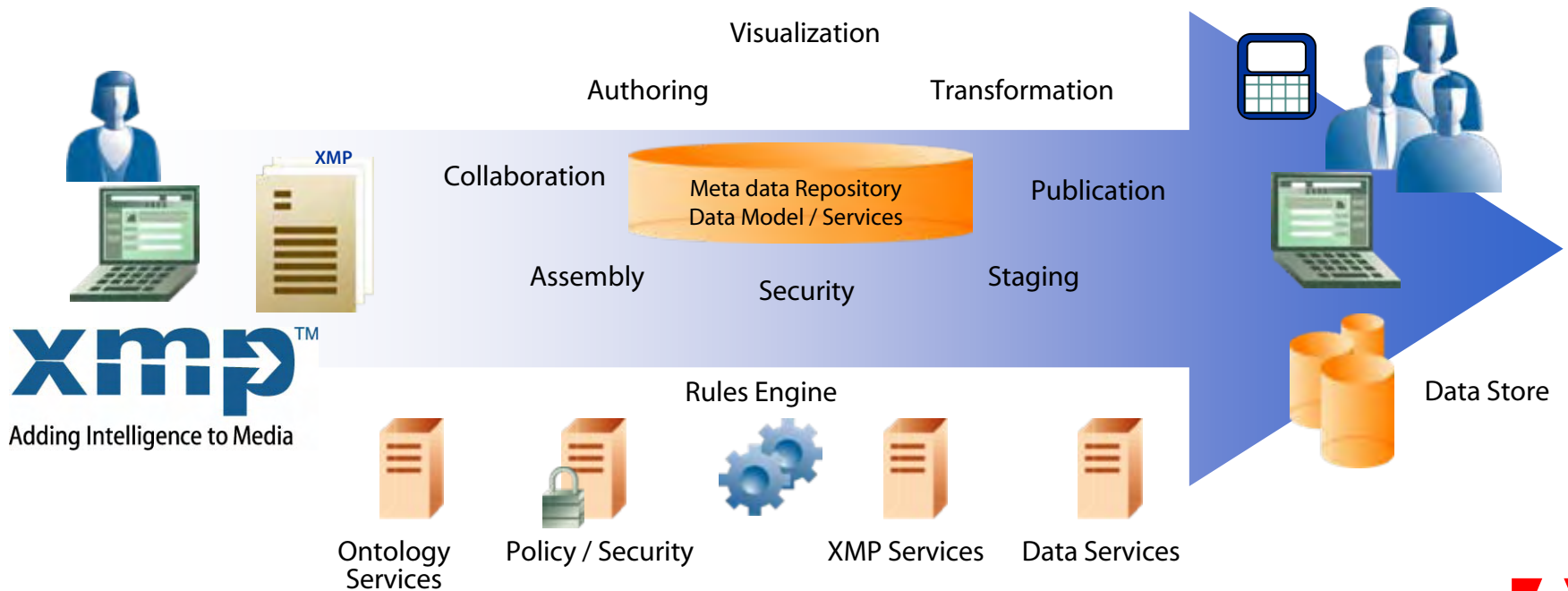
- apologies to Marshall McLuhan

XMP

Capture, share, and leverage metadata

“[XMP] is an important piece that brings the Semantic Web closer to realization.”

– Eric Miller, W3C Semantic Web Activity Lead





XML Platforms at Adobe

- **XMP** for image and text metadata
- **XML Schema** for electronic forms
- **XFA & XDP** for Intelligent Documents
- **XML** for page layout templates and content description
- **JDF** for print job production



XMP is Based on RDF



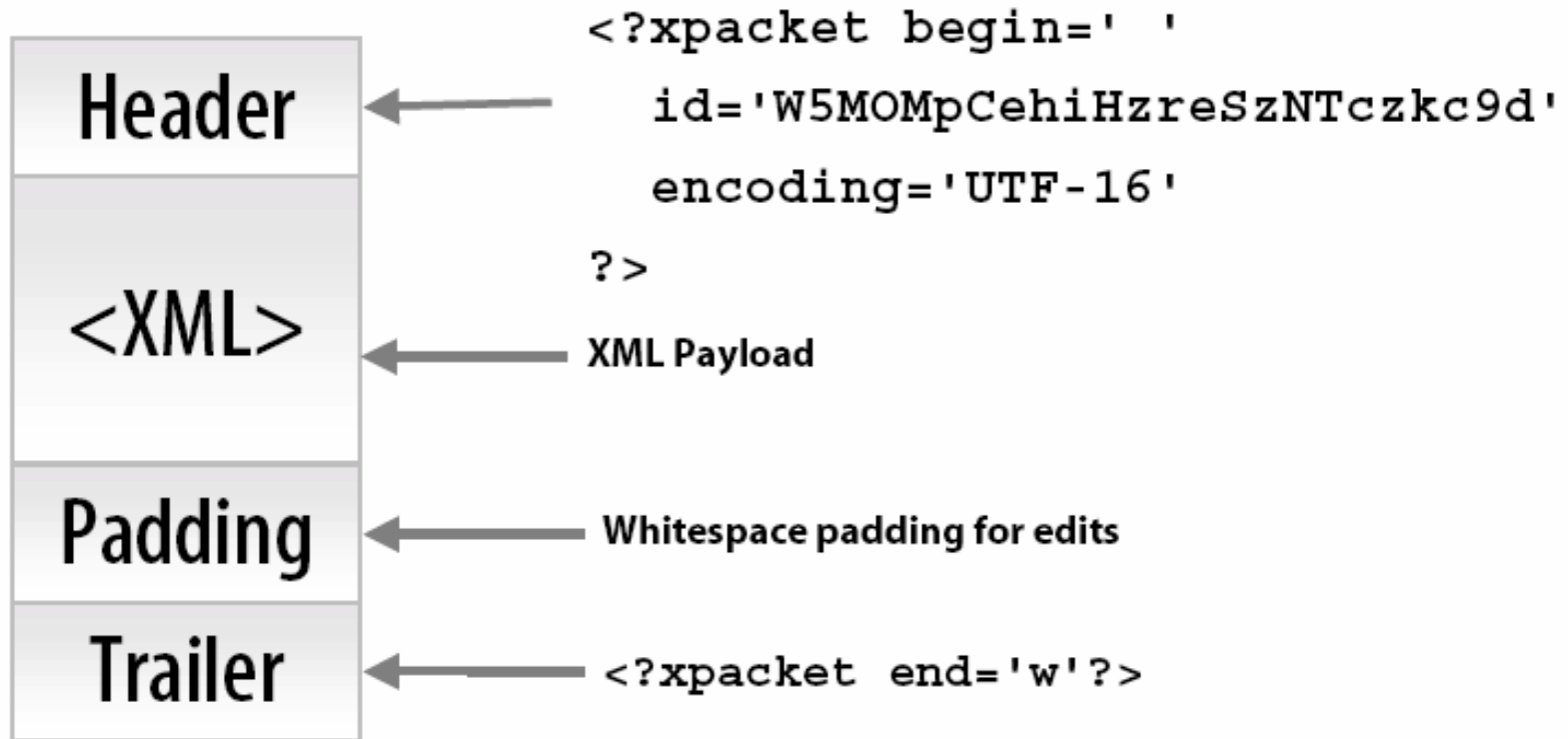
- **RDF: *Resource Description Framework***
- **Framework for metadata**
 - Used to define metadata languages
- **Scalable across domains**
 - Can be serialized in XML, but has important advantages over XML Schema-based languages
 - Mixed vocabularies
 - No order constraints
 - Less “Brittle”



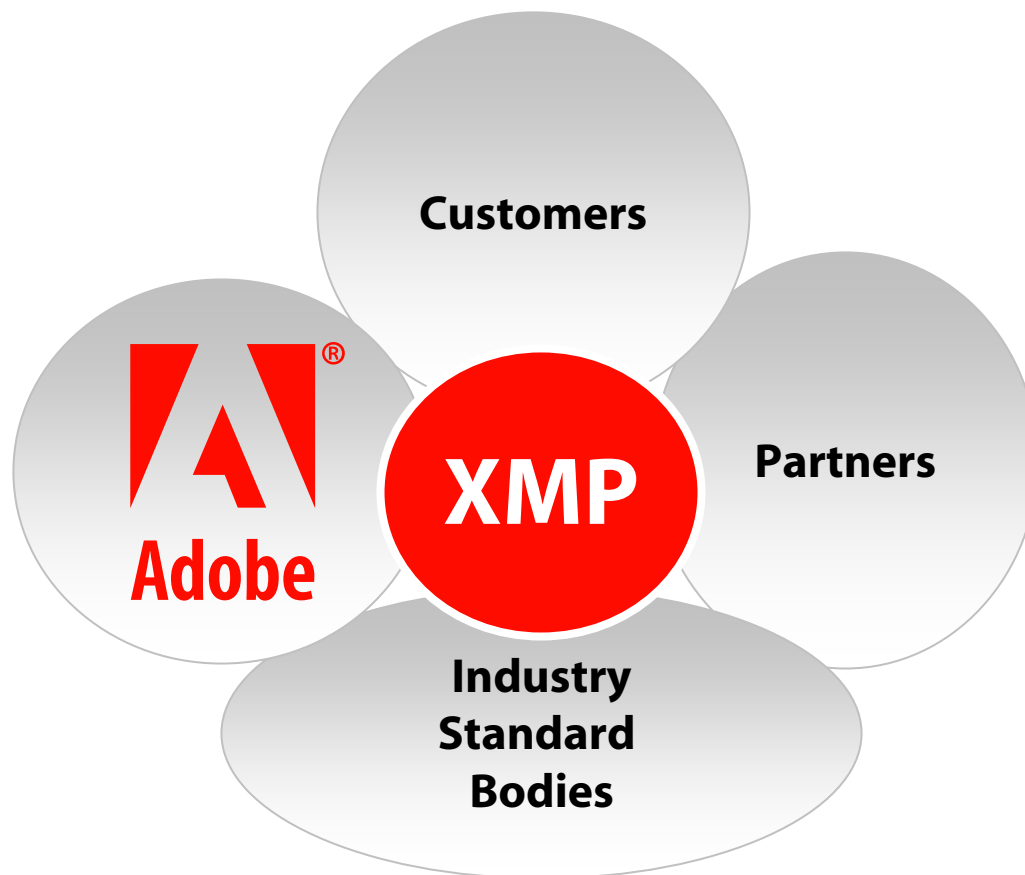
Key Elements of XMP

- **Framework**
 - XML structure for storing information
- **Packet**
 - Where information is stored and accessed in the file
- **Specification**
 - Description of relationship to other **standards**
- **Development tool (SDK)**
 - Freely available, **open source** licensed tools
- **Custom Panels User Interface**
 - Human interface interaction with metadata
- **Platform**
 - Standardized access to metadata across Adobe CS

XMP packet

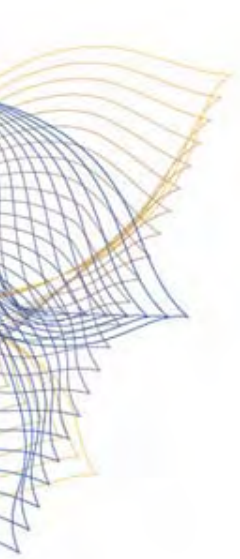


XMP Ecosystem



XMP Supports Standards

- **IPTC (IPTC Core)**
 - IPTC standard expressed with XMP framework
 - <http://www.iptc.org>
- **DISC - Digital Image Submission Criteria**
 - Photographers submitting photos to magazines
 - <http://www.disc-info.org/>
- **DNG - Digital Negative Format**
 - For Camera Raw file interchange
 - <http://www.adobe.com/dng/>
- **EXIF – Exchangeable Image Format**
 - XMP framework for the standard for interchange information in image files
- **PDF/A - Archive**
 - Adopting XMP as metadata archive standard



Industry Support



- Ancept/Stellent
- Apago
- Artesia
- Context Media
- CoreMedia
- BearingPoint
- Data Technologies International (DTI)
- DeepBridge
- **Documentum/EMC**
- Extensis
- Getty Images
- **IBM/NICA**
- **Interwoven**
- **IxiaSoft**
- Kodak
- MediaBin/Interwoven
- Media Systemen
- **North Plains Systems**
- Pindar Systems
- Pine Tree Systems
- **Pound Hill Software**
- Profium
- **Sansui**
- W3C
- WaveCorp/MediaBank
- **WebWare/Insci**
- Xerox



Community Support & Resources

<http://www.adobe.com/xmp>

- Spec and SDK available via **open source** license
- Resources via XMP Web site
 - White papers
 - Datasheet
 - Conference presentations
 - Developers Toolkit
 - Online knowledgebase
- **Community**
 - Contributors free to add to the open source
 - Requests, bug-fixes, changes, enhancements

A decorative graphic on the left side of the slide, consisting of a series of overlapping, curved lines in shades of blue and yellow, creating a sense of depth and movement.

Adobe PDF

- **The advanced way to create, control, and deliver more secure, high-quality documents**
 - Create and reliably exchange content-rich documents
 - Assemble multiple files into a single polished document
 - Collect and manage review feedback
 - Apply advanced document access controls and security
 - Become a platform for XML standards as Intelligent Documents

The PDF Ecosystem

- **Proven lineage—10+ years**
- **Adopted by Fortune Global 1000 enterprises**
- **Beyond de facto as a trusted document standard:**
 - PDF/E: For engineering documentation
 - PDF/A: For long-term preservation of electronic documents
 - PDF/X: For reliable delivery of press-ready, high-end color advertisements
- **Mandated by worldwide governments, standards bodies, and international organizations**
- **Published, publicly available specifications**
- **Large 3rd party developer community**

Adobe Acrobat & Reader

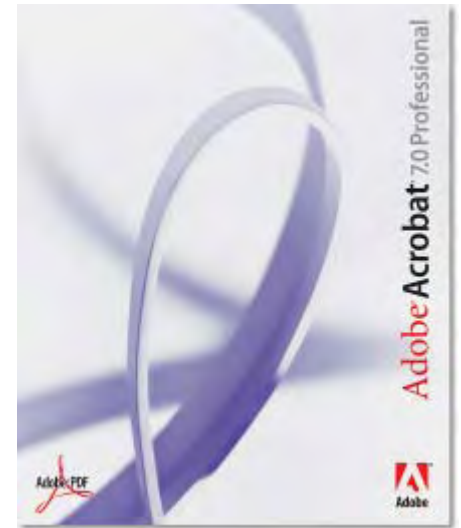


- **Adobe Acrobat**

- 1,800 partnerships offer solution add-ons, plug-ins, consulting, training, and support
- Key partners: Documentum, Bentley, IBM, Agile, OpenText, EDS (PTC)
- More than 16.5 million seats sold

- **Adobe Reader**

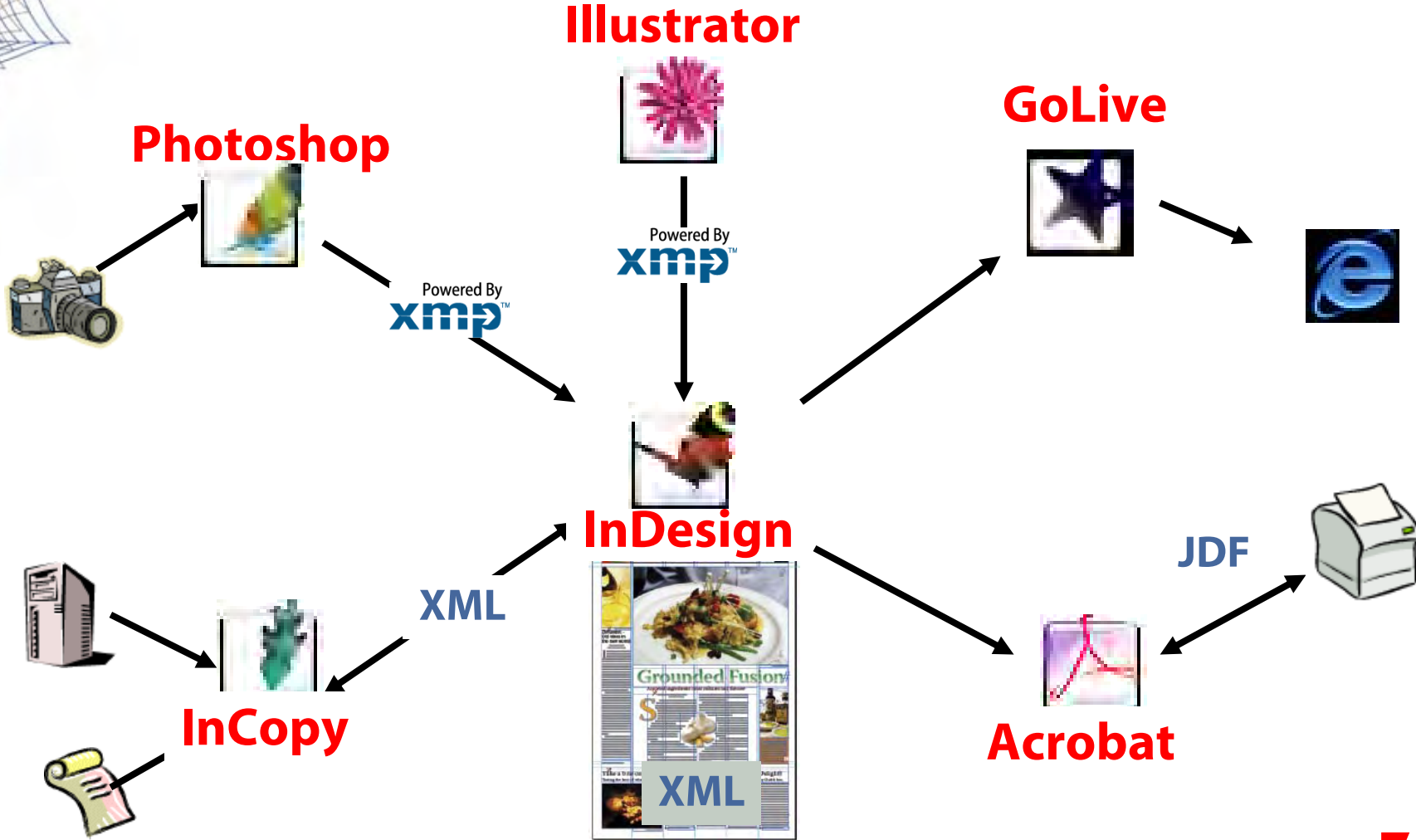
- Half a billion copies distributed
- Bundled with most PDAs
- Distributed by all top PC manufacturers



XMP in Applications

- **Adobe Acrobat & Reader**
- **Adobe InDesign**
- **Adobe Illustrator**
- **Adobe Photoshop**
- **Adobe GoLive**
- **Adobe Bridge**

XMP Workflows





XMP Custom Panels

- **Human interface with metadata**
- **Consistent metadata dialog across CS**
 - Standard view of metadata within workflow
- **Panels defined by text file**
 - Defines the user readable **label** for properties
 - Describes the visual **layout** & interactivity
 - Defines the **storage** of metadata in properties
- **Links data input to data storage structure**
 - Data entered via panel is stored in XMP packet according to panel description
- **Enables standards interoperability & customization**

Use Case Criminal Investigation

Multiple Agencies, Multiple Departments, Multiple Records Systems



Federal Agents	State & Local	Other Agencies	Forms	Citizens
----------------	---------------	----------------	-------	----------

Federal Criminal Investigations Center

Priorities

- About
- Press
- Investigations
- Counterterrorism
- Intelligence
- Most Wanted
- Local Offices
- Family
- FOIA
- Search

Electronic Case Jacket

COLLECT ANALYZE SHARE PROVE

Announcements

On 2/22, Virginia man returned to the U.S. to face charges of providing material support to Al-Qaeda.

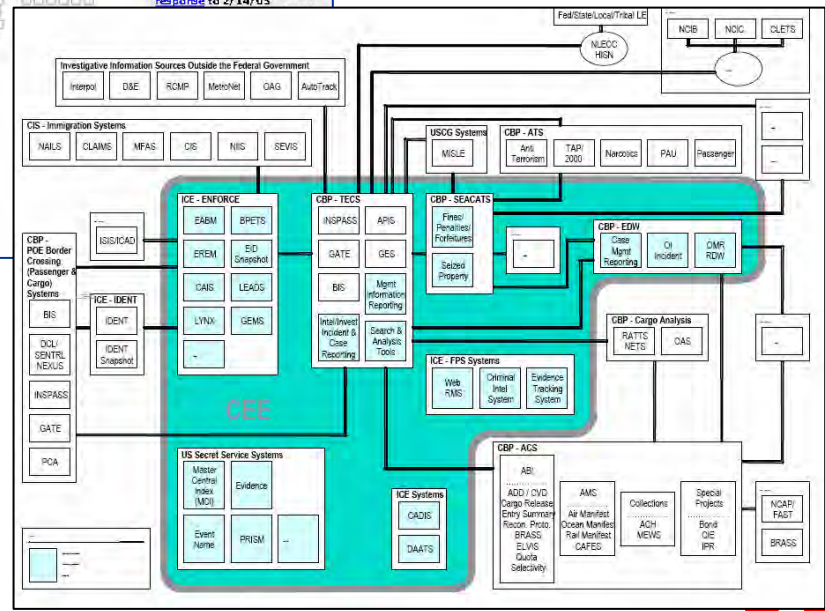
On 2/22, 11 sitting and former county public officials in New Jersey charged with extortion.

FBI exec
Catherine Chandler issues response to 2/14/05



Investigative Report

Case Name	<input style="width: 80%;" type="text"/>
Case Officer	<input style="width: 80%;" type="text"/>
Case Description	<input style="width: 80%; height: 40px;" type="text"/>
Evidence	<input style="width: 80%; height: 80px;" type="text"/>
Investigative Team	<input style="width: 80%; height: 60px;" type="text"/>
Summary Report	<input style="width: 80%; height: 60px;" type="text"/>

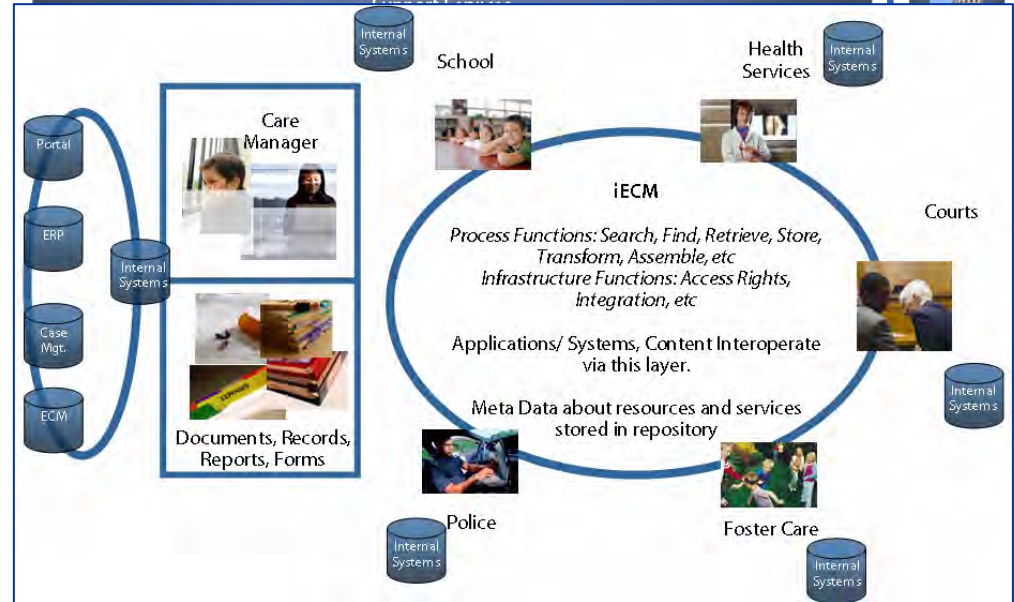
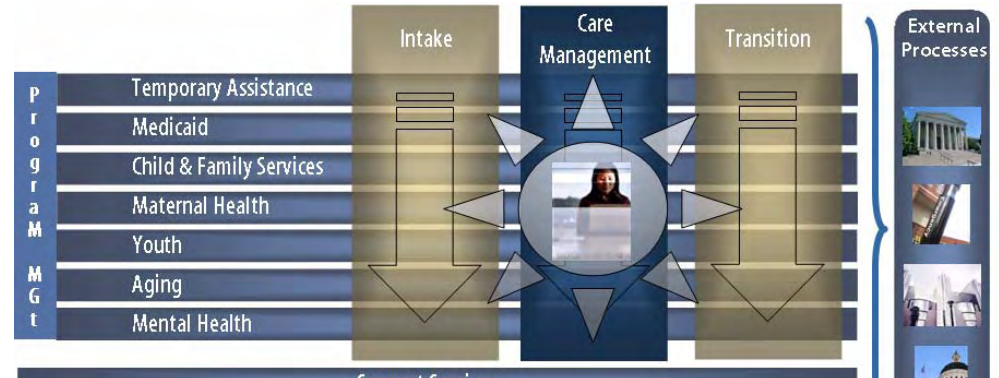


Use Case Public Health

Multiple Standards, Complex Information Sets, Critical Results



- Alerts for Public
- Clinician Information
- Alerts/Advisories
- Public Health Updates
- HEALTH ACTION
- Programs & Services
- Birth & Death Records
- Child & Family Health
- Disease Prevention
- Emergency Medical Services
- Environmental Health
- Foster Care Clinic
- Health Information & Data
- Medical Examiner
- Special Services





XMP Summary

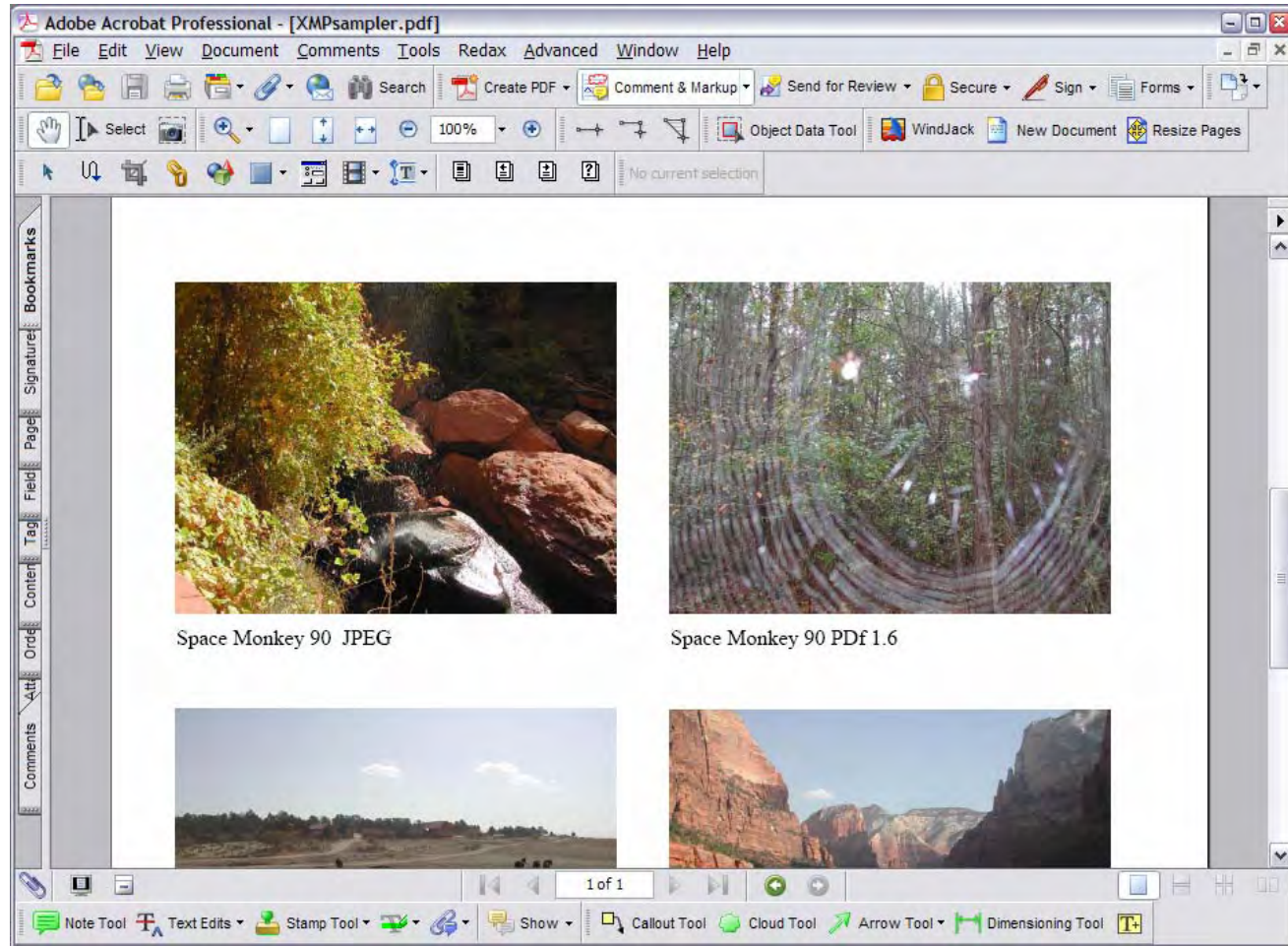
- **Information about information**
 - Descriptive labels for files
 - Machine readability
- **Minimize mistakes**
 - Intelligent searching, categorization & version control
 - Find files faster, ensure proper file usage, compliance with rights & reuse policies
- **Flexible, open integration platform**
 - Legacy and future extensible, vendor neutral
 - Promotes interoperability across domains
- **Standards implementation vehicle**
 - Support for industry standards
 - Application integration (Custom Panels & Adobe Bridge)



Demos

- **XMP in PDF**
- **The Adobe Bridge**

XMP in PDF



XMP in PDF

Document Metadata for XMPsampler.pdf

Description

Camera Data 1
Camera Data 2
Categories
History
IPTC Contact
IPTC Content
IPTC Image
IPTC Status
Adobe Stock Photos
PDF/A
SCORM
Advanced

Description

Document Title: XMP in PDF Samples

Author: Ed Chase

Description: Demonstration of InDesign & Photoshop Object XMP

Description Writer:

Keywords:

Commas can be used to separate keywords

Copyright Status: Public Domain

Copyright Notice:

Copyright Info URL:

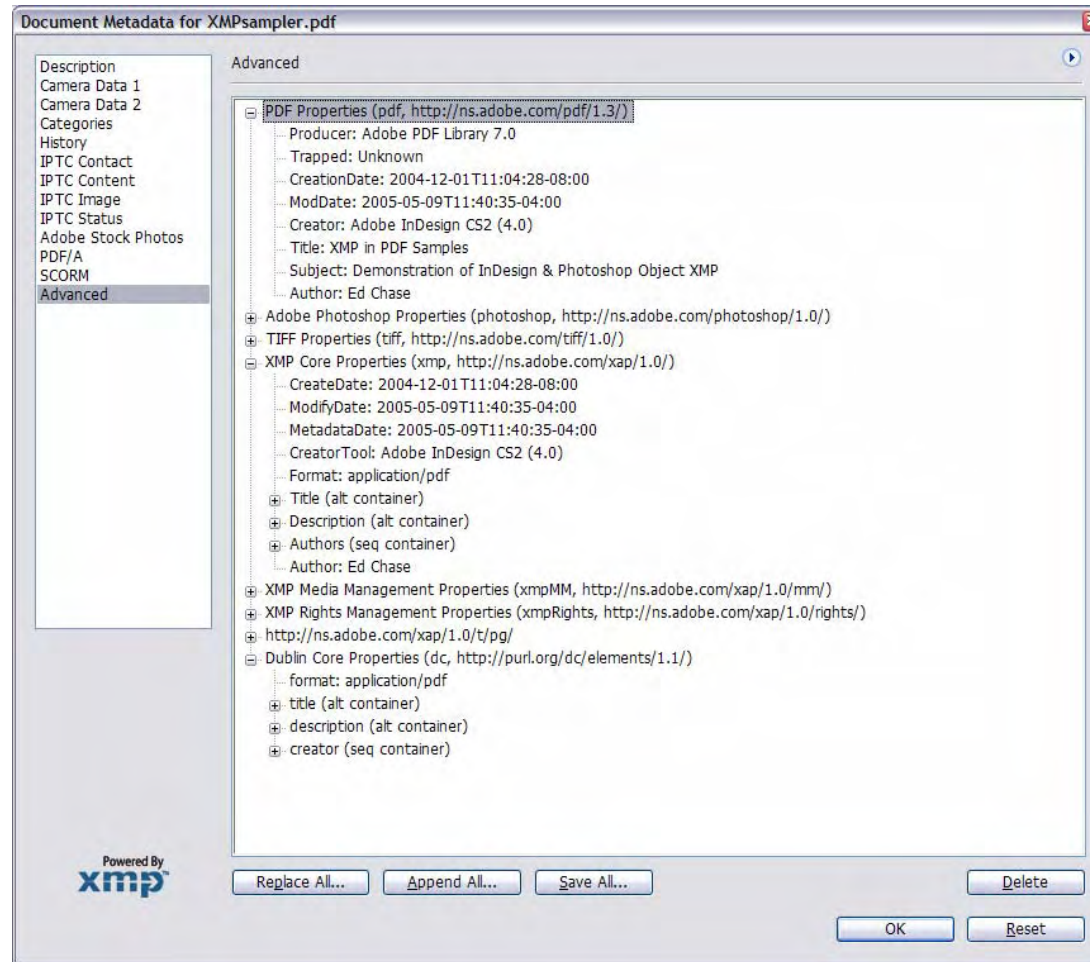
Go To URL...

Created: 12/1/2004
Modified: 5/9/2005
Application: Adobe InDesign CS2 (4.0)
Format: application/pdf

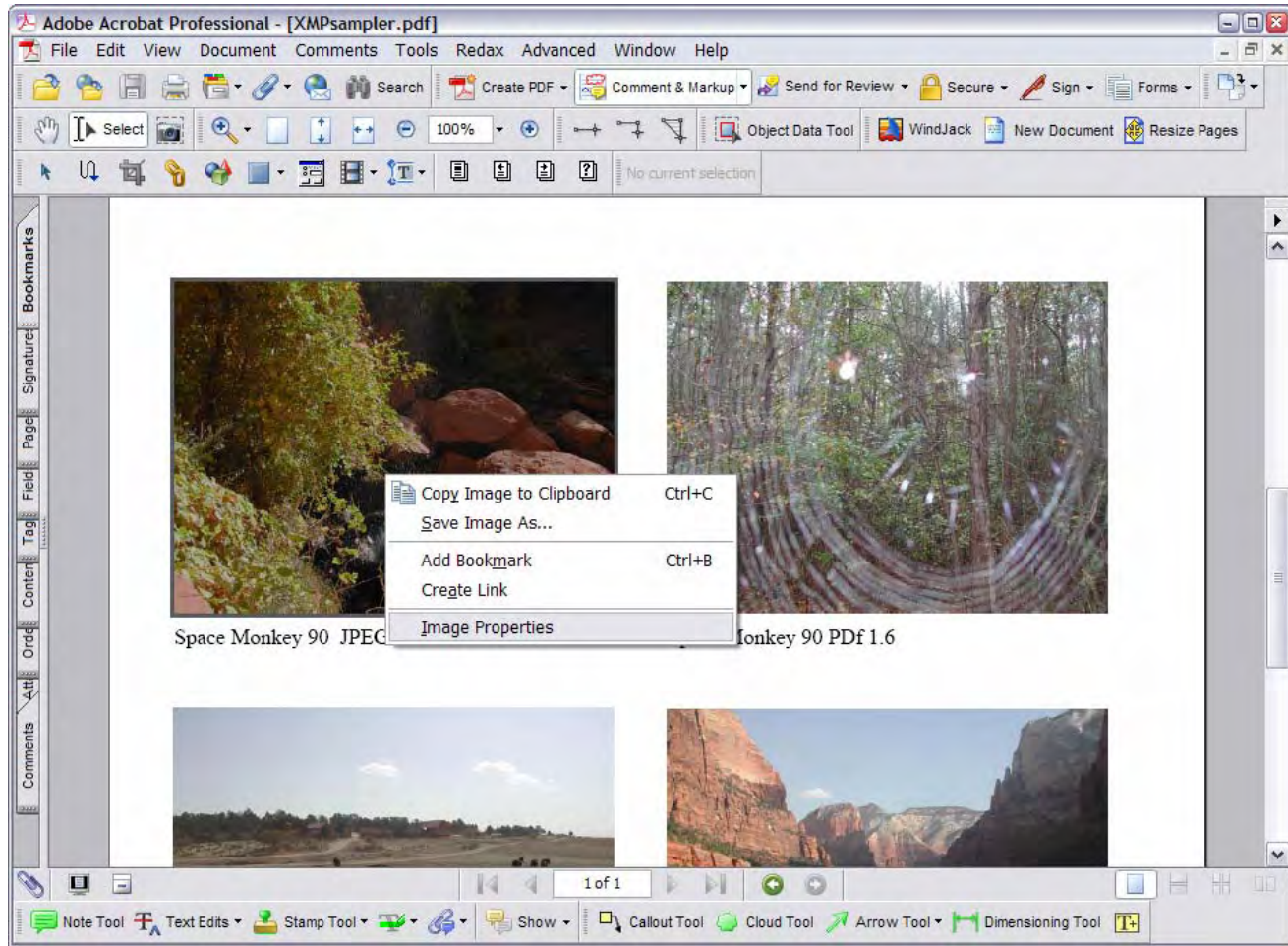
Powered By **xmp**

OK Reset

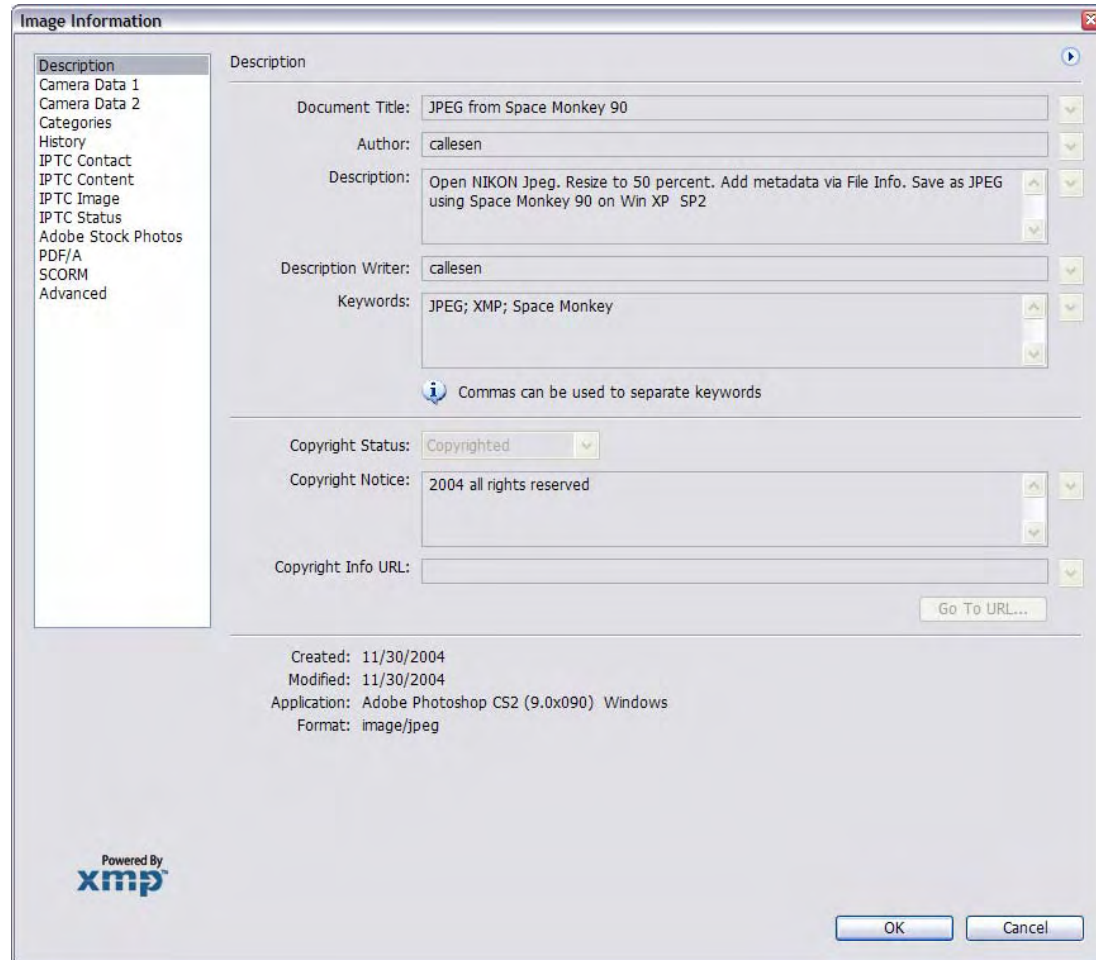
XMP in PDF



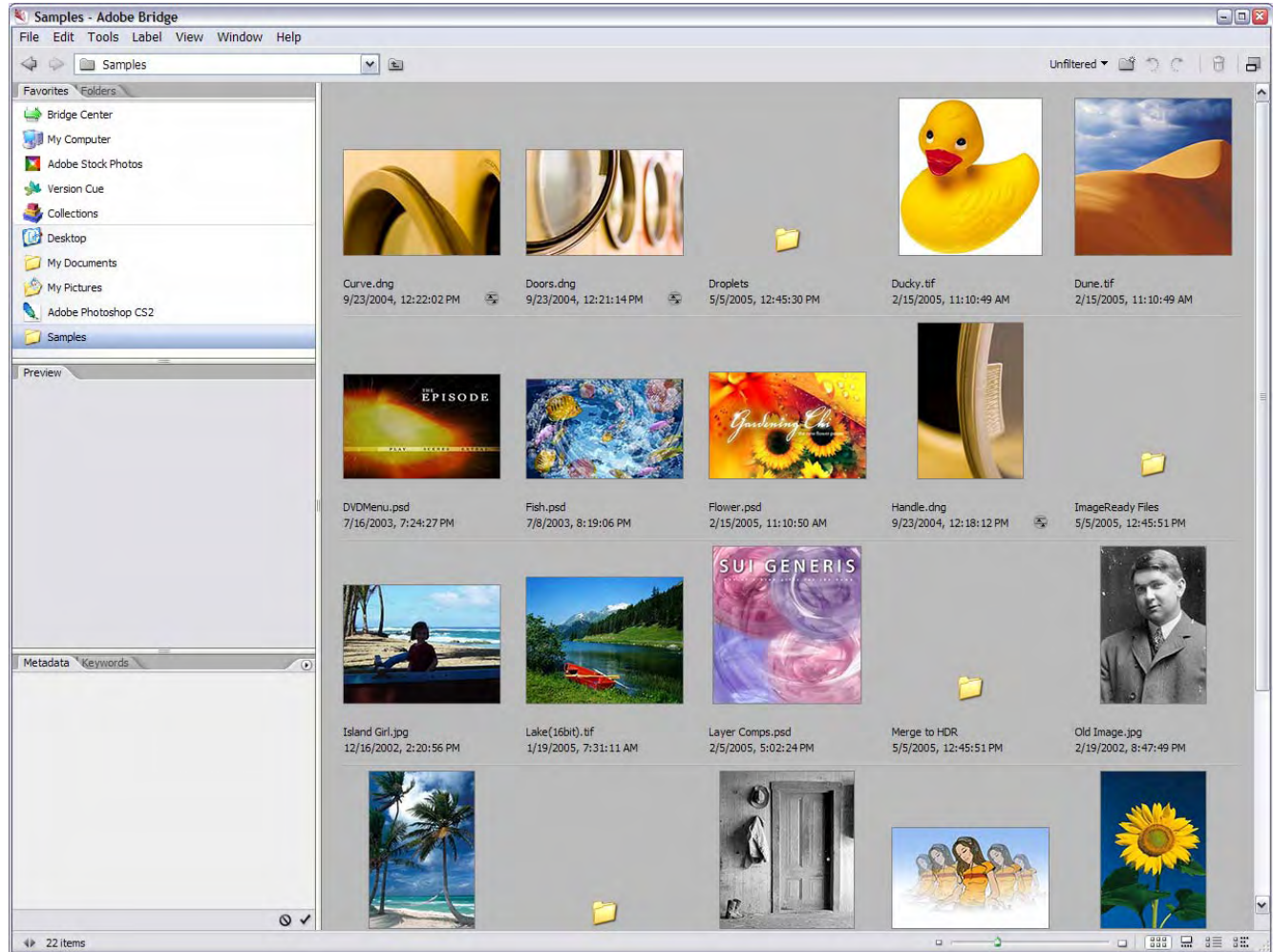
XMP in PDF



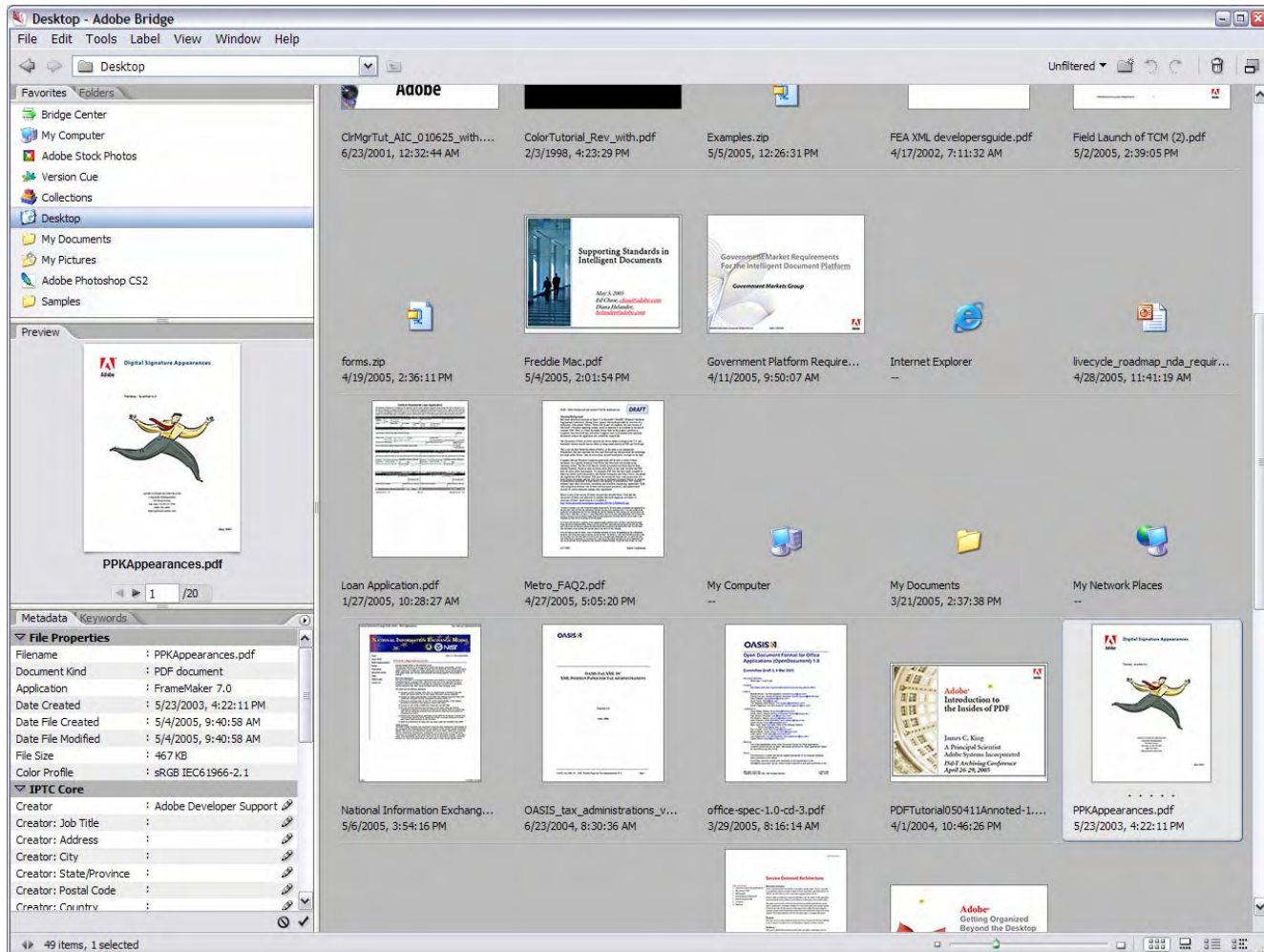
XMP in PDF



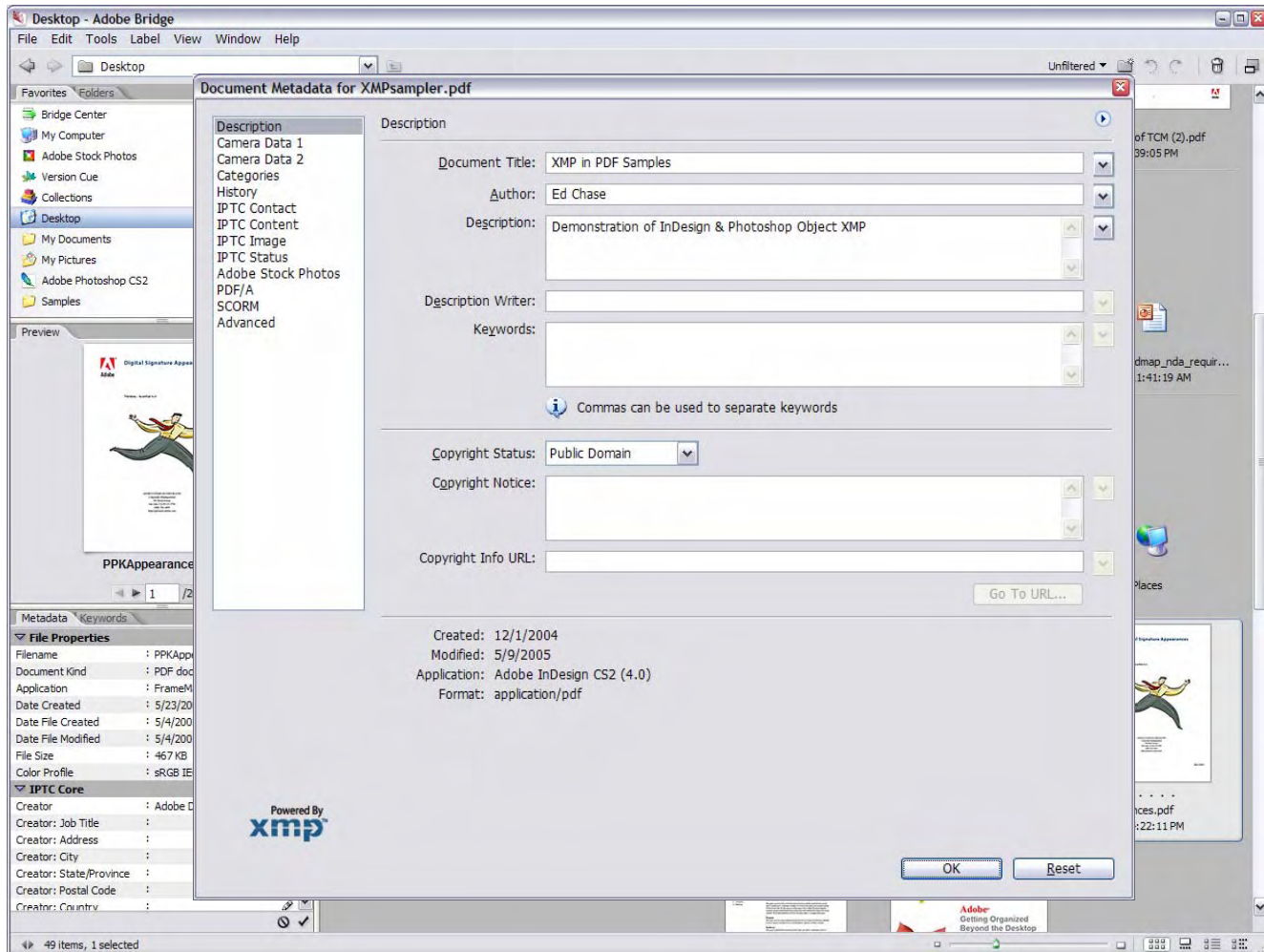
The Adobe Bridge



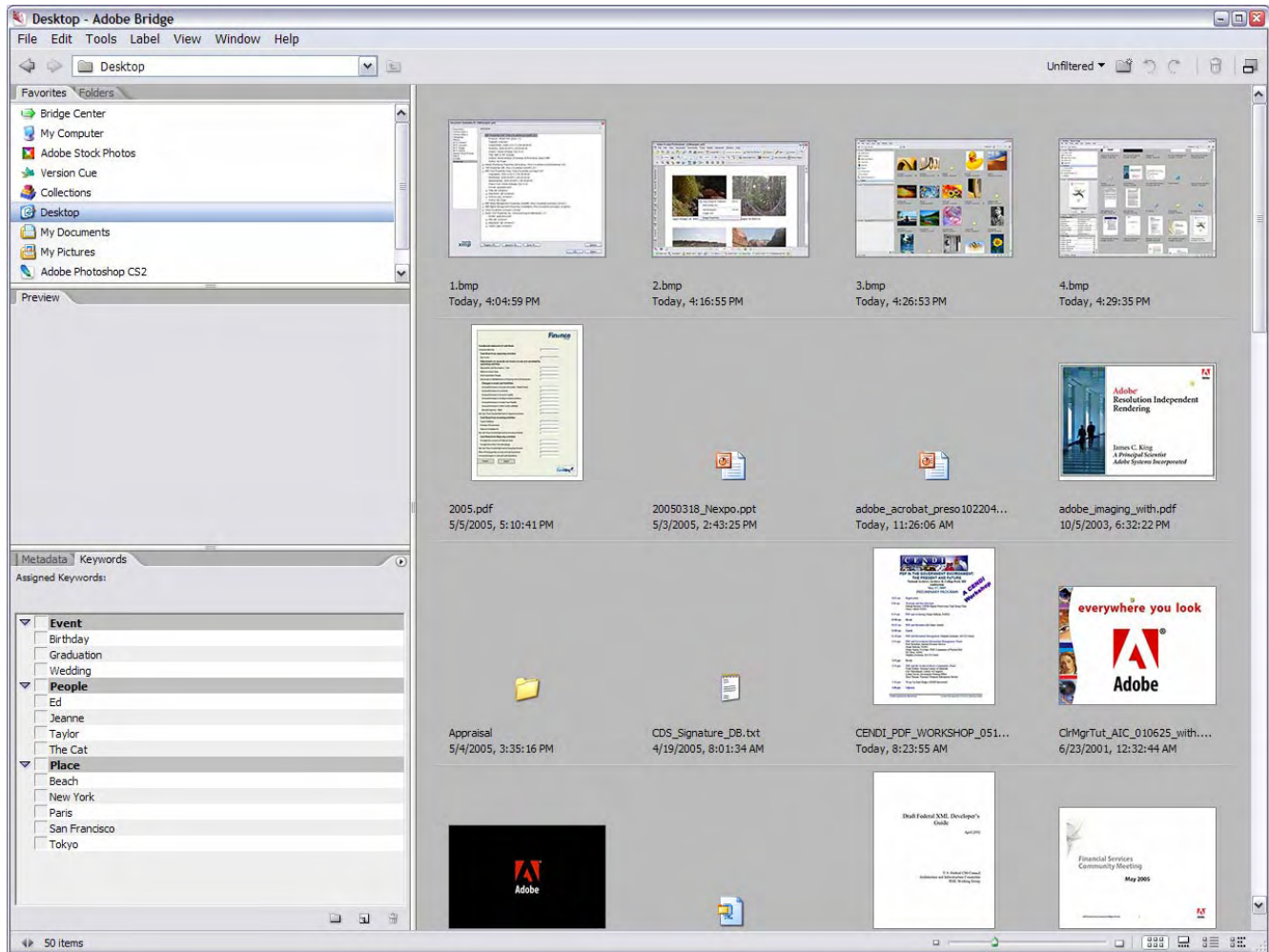
The Adobe Bridge



The Adobe Bridge



The Adobe Bridge





Questions?

Ed Chase
Worldwide Standards
Adobe Systems, Inc.
chase@adobe.com