

**Tehokkaat ratkaisut digitaaliseen arkistointiin: lyhytaikais-,  
pitkäaikais- ja pysyväisarkistointi lainsäädännön mukaisesti.**

teknologiakonsultti Kristian Salo, EMC Computer-Systems Oy

Digitaalisesta arkistoinnista on tullut kiinteä osa yritysten ja yhteisöjen toimintaa. On arvioitu, että maailmanlaajuisesti vuoteen 2010 mennessä arkistoidaan yli 27 000 petatavua. Digitaalisella arkistoinnilla mahdollistetaan tiedon kustannustehokas säilytys turvallisesti. Suurten tietomassojen pitkäaikaisen digitaalisen säilyttämisen aiheuttamat kustannuspaineet, sekä erilaiset lait ja asetukset pakottavat yritykset ja yhteisöt etsimään keinoja tiedon arkistointiin sekä sen elinkaaren hallintaan. Digitaalisen arkistoinnin on oltava loppukäyttäjälle helppokäyttöistä ja tiedon etsiminen järjestelmästä mahdollisimman tehokasta, nopeaa ja yksinkertaista. Oikeilla työkaluilla on mahdollista rakentaa arkkitehtuuri, jolla voidaan loppukäyttäjälle ja/tai sovellukselle läpinäkyvästi arkistoida tietoa kustannustehokkaasti ja vastata samalla lakien ja asetusten vaatimiin haasteisiin.



EMC

# Information Storage and Management Challenges

## Lower TCO

Usage  
Consolidation  
Automated management

## Protection / Recovery

Security is top IT spend priority

## Information Growth

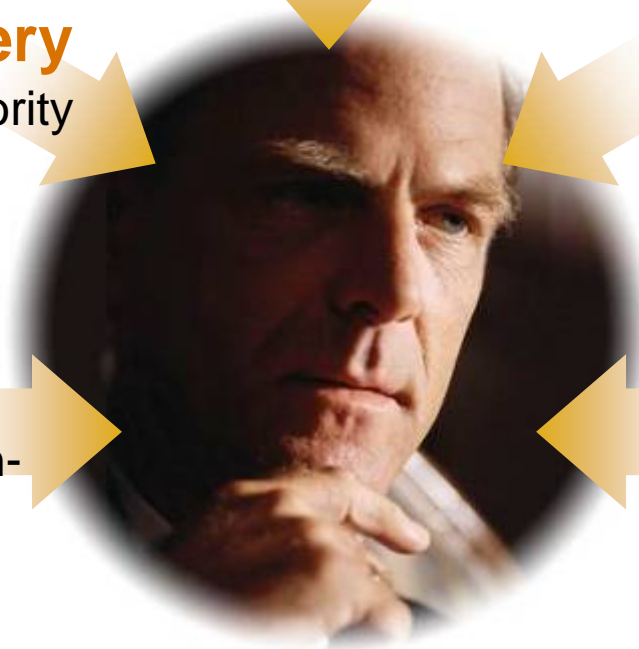
TB growth: > 60% annually

## Simplicity

Hundreds of information-  
management tools

## Compliance

> 16,000+ regulations  
worldwide



*Setting the stage: A new approach is needed*

# Backup and Archive *are* Different

## Backup

A **secondary copy** of information

Used for **recovery** operations

**Improves availability** by enabling application to be restored to point in time

Typically **short-term** (weeks or months)

Data typically **overwritten** on periodic basis (e.g., monthly)

**Not for regulatory compliance**—though some are forced to use

## Archive

**Primary copy** of information

Available for information **retrieval**

**Adds operational efficiencies** by moving fixed / unstructured content out of operational environment

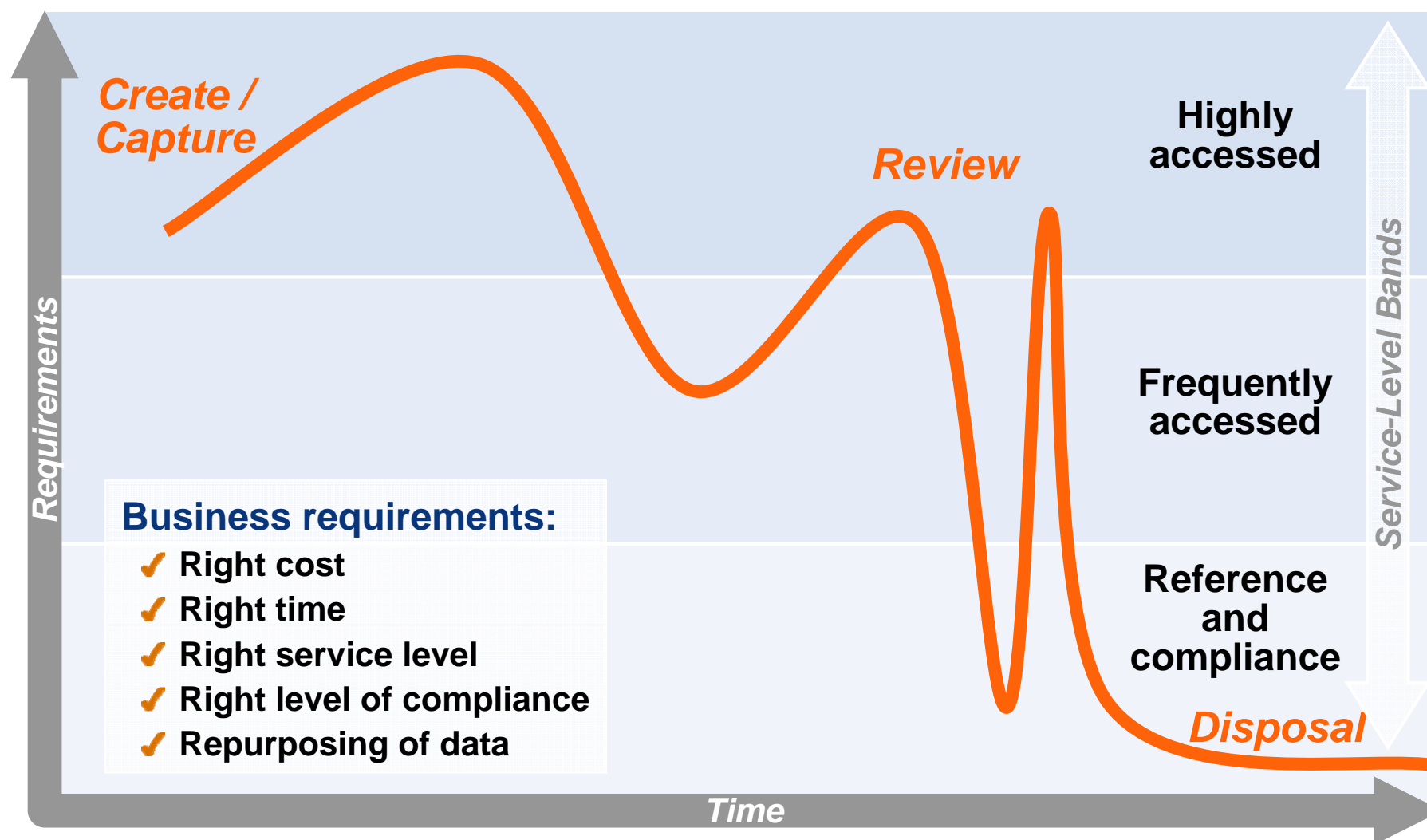
Typically **long-term** (months, years, or decades)

Data typically **maintained** for analysis, value generation, or compliance

**Useful for compliance** and should take into account information-retention policy



## The Information Lifecycle

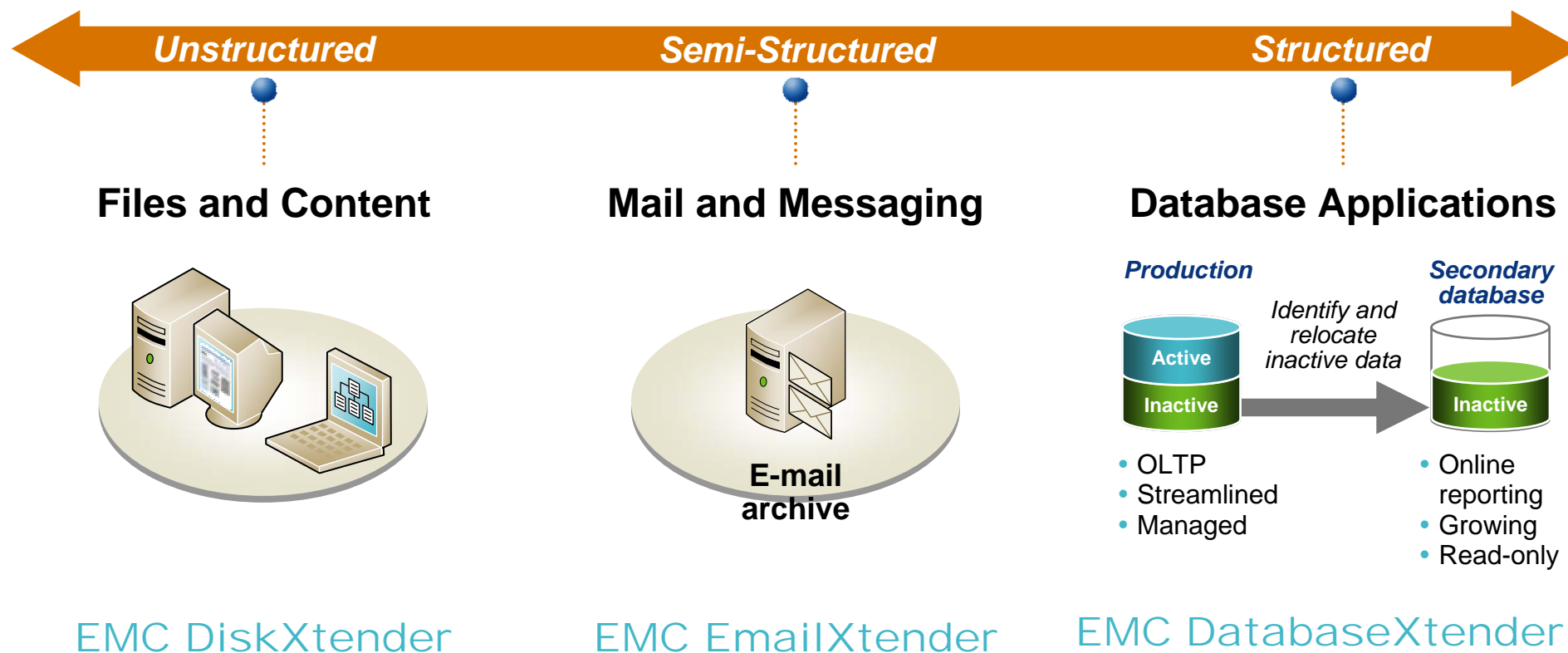


## Automated Data Movement to Manage Filesystem Growth

- Align storage with data access, cost, and retention requirements
  - Lower overall storage costs
  - Back up and recover faster
  - Meet compliance demands
- Optimize data placement across the storage infrastructure
  - Policy-driven
  - Transparent migration
  - Seamless user access



# EMC Delivers Information Lifecycle Management Across All Data Types



## Why Archive E-mail?



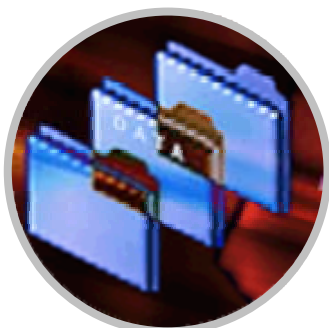
### Reduce costs, increase efficiency and productivity

- Simplify user management
- Reduce storage requirement
- Improve availability
- Improve user access
- Support e-mail growth



### Risk Mitigation

- Execution of discovery requests
- Elimination of unmanaged records and archives
- Monitoring adherence to computer use policy



### Record Keeping

- European Data Protection Act (in '05 applies to e-mail and IM)
- Sarbanes Oxley
- Investment Advisors SEC 204-2
- Broker / Dealer SEC 17a-4
- NASD 3010



## AIIM International and Kahn Consulting Study

### How Organizations Use E-Mail Today (as a % of respondents)

- 100% organizations use e-mail to conduct business
- E-mail is the #1 method of business communication
- E-mail and IM are the most difficult to implement internal controls for record keeping
- Litigators take advantage of lack of preparation by making e-mail a target of discovery

### Discussing HR Issues

- 56% discussing HR issues

56%

Source: *Managing E-Mail in the New Business Reality*  
AIIM International and Kahn Consulting, Inc., September 2003

©2003 Kahn Consulting, Inc.

[www.kahnconsultinginc.com](http://www.kahnconsultinginc.com)

## Access to E-mail is Key

United States District Court, N.D. Illinois.  
In re BRAND NAME PRESCRIPTION DRUGS ANTITRUST LITIGATION.  
No. 04-2008-1555-AM

- E-mail backups on tape is not enough
- Search and retrieval of e-mail is extremely costly
- Example: Brand Name Prescription Drug Antitrust Litigation
  - Discovering party (plaintiff) wished to search 30 million pages of e-mail stored on producing party's (defendant's) backup tapes for the names of particular individuals. The producing party argued that the discovery party should pay the cost of compiling, formatting, searching eliminating duplicates, and retrieving requested e-mail
  - Court found the request was not "unduly burdensome" because the difficulty and cost of the effort was largely due to limitations in the producing party's software, and ordered them to undertake the retrieval at their own expense

# Comprehensive Storage and System Management

## Index

**Messages and attachments:**

- MS Office documents
- PDFs
- ZIP files
- Embedded messages

## Capture

- Copy
- De-duplication
- Unique ID
- Real time
- Archive task

## Organize and Classify

- Retention and disposal
- ILM

## Archive

- Secure
- 2:1 compression

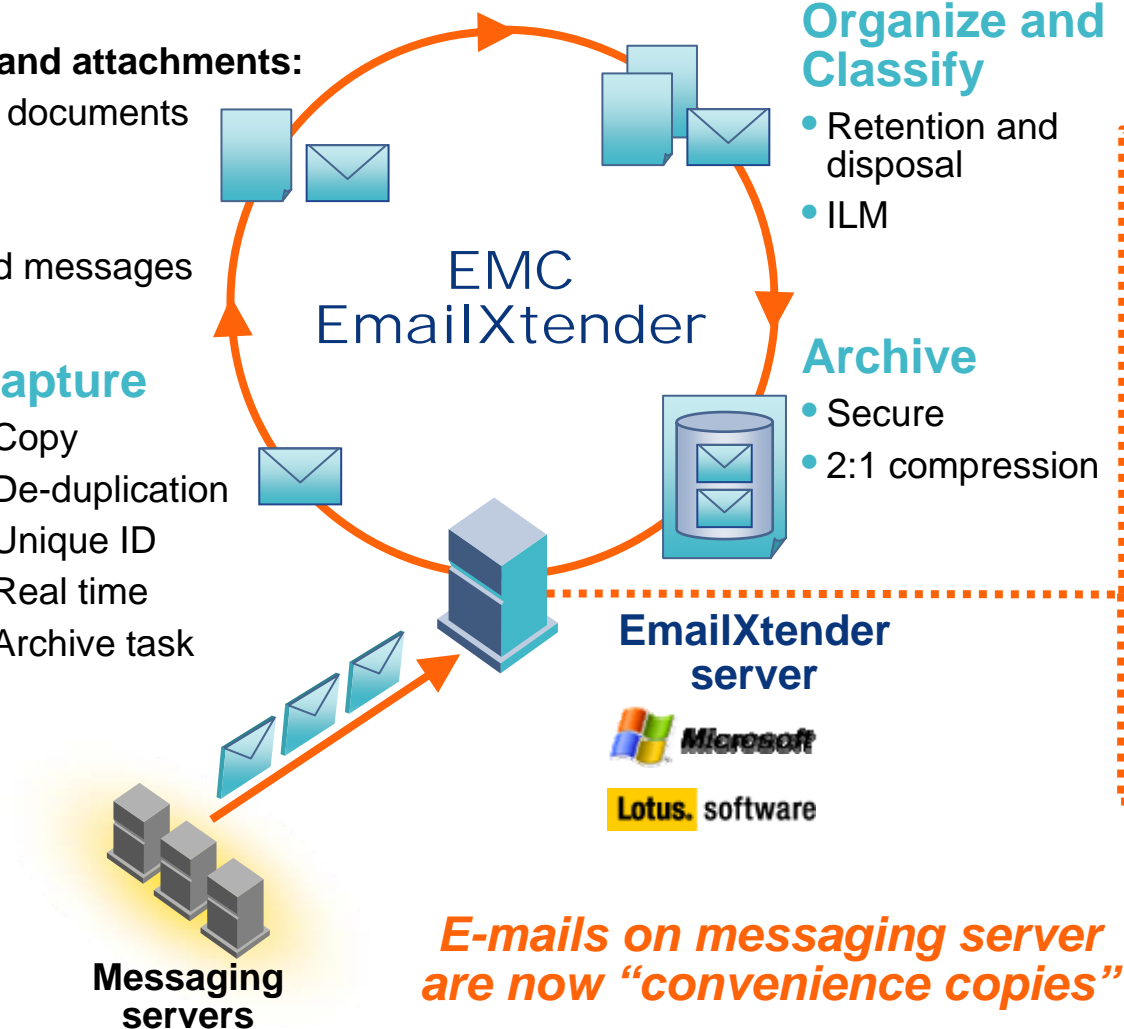
## • User Access

- User searches from plug-in or web
- Transparent shortcut access



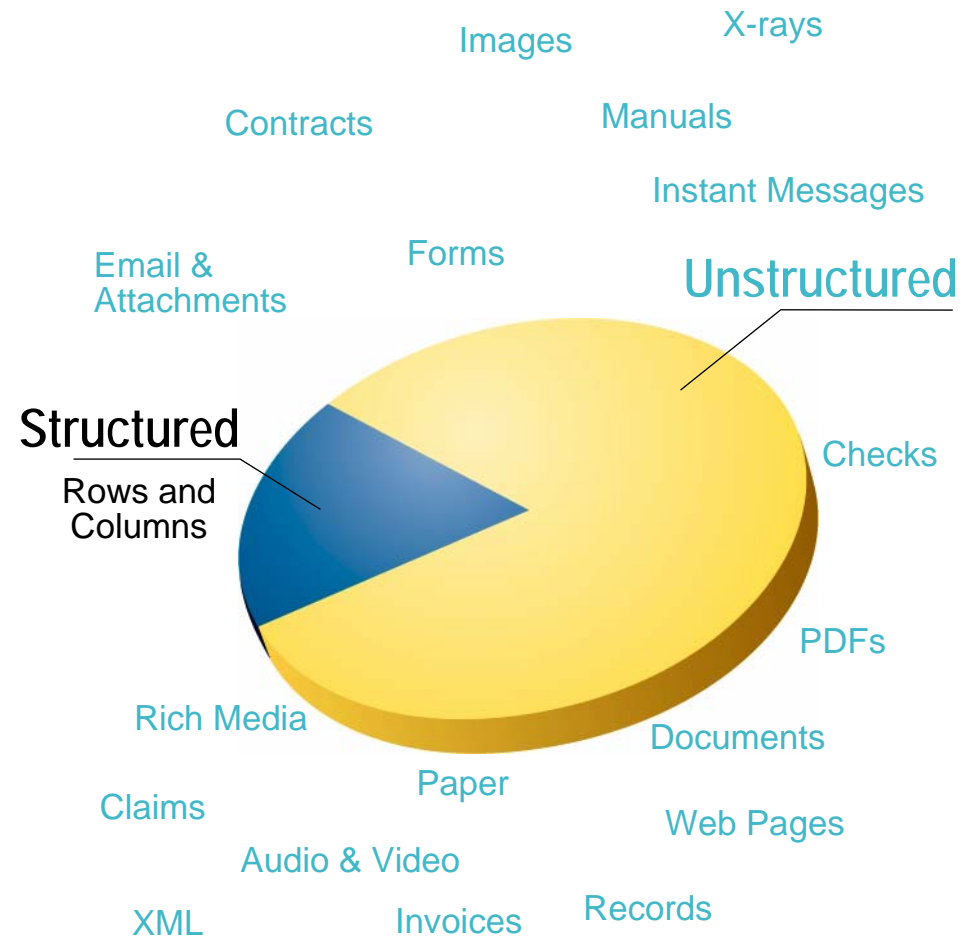
## Advanced Administration

- Administration of a single central archive
- Tools, reports, and other diagnostics

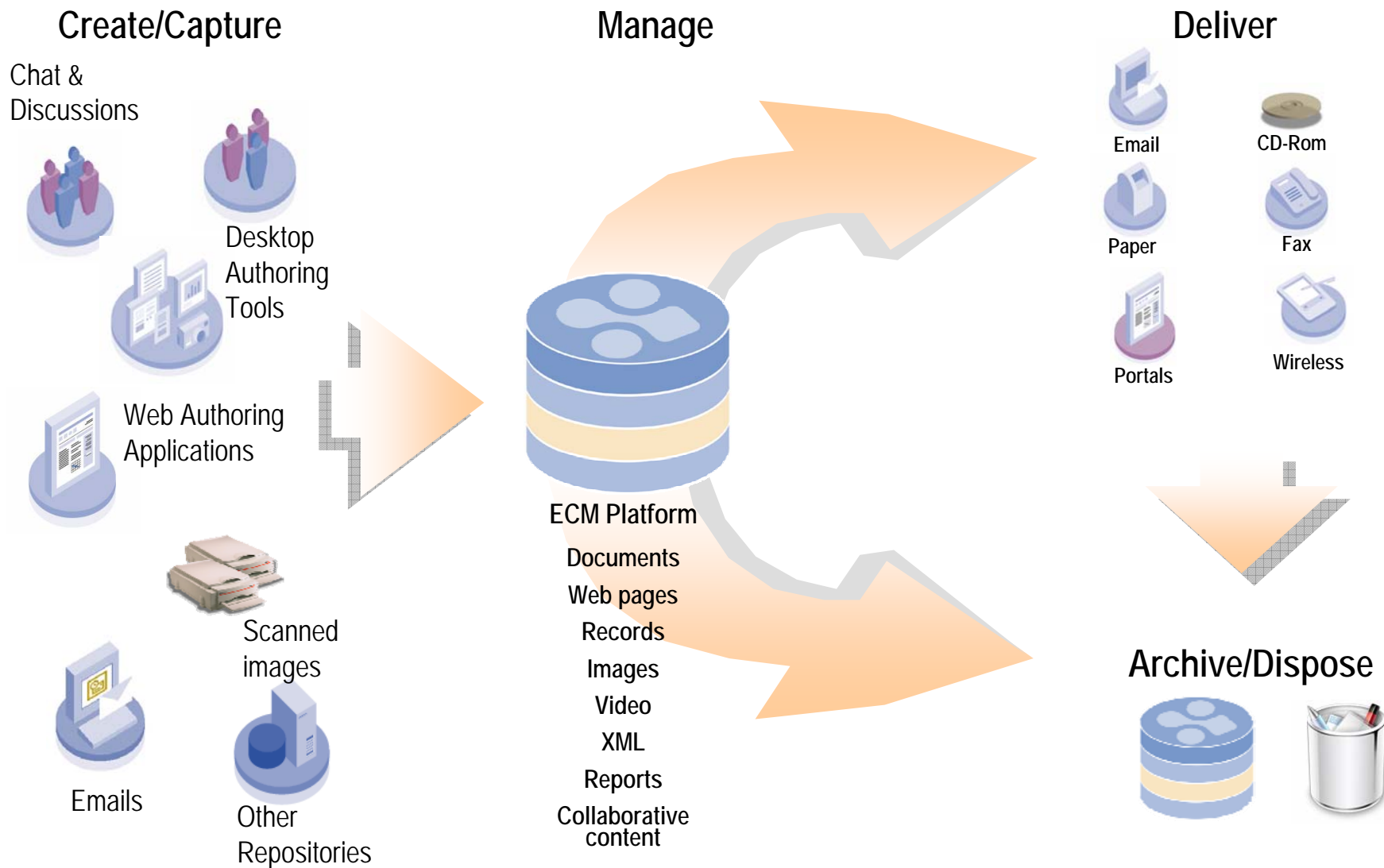


# Relentless Information Growth

- The amount of information is growing every year
- Includes structured, semi-structured, and unstructured data
- Over 80% of enterprise information is unstructured  
(Fulcrum Research)

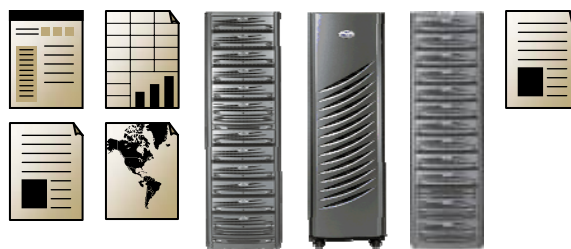


# What is Content Management?



## How DiskXtender Works

Filesystems are scanned and files matching policies are identified



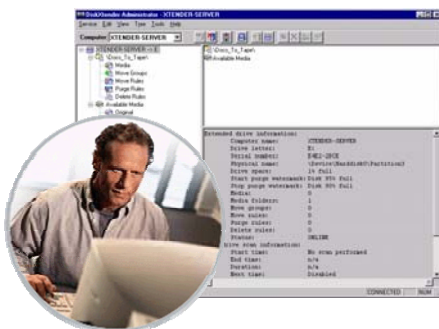
UNIX, Linux, Windows,  
NAS filesystems



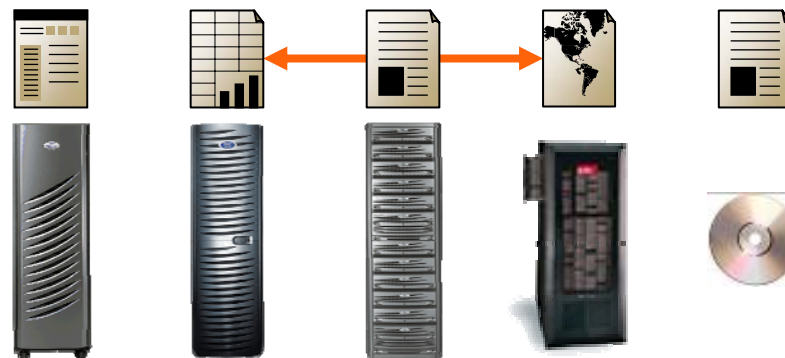
View is always local, and access is transparent to user or applications

**DiskXtender**

Policies classify files:  
Name, Size, Type, Last  
Access, Last Modified...



Files can be moved to  
single or multiple targets,  
or across multiple tiers

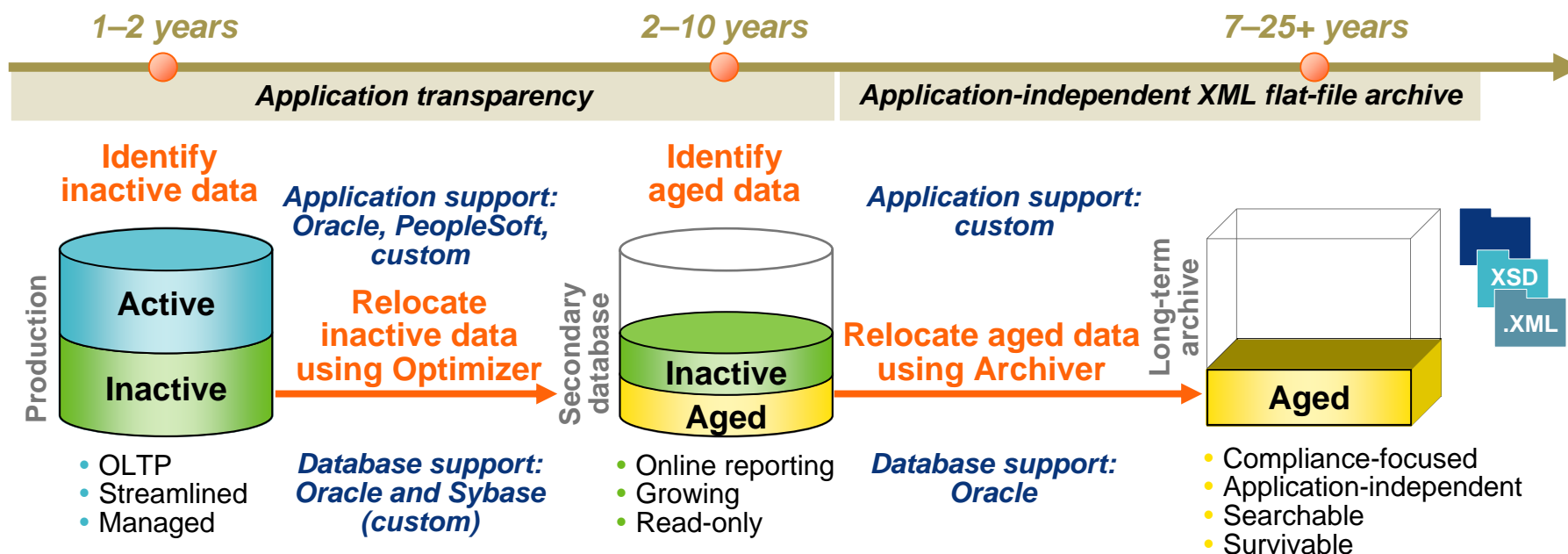


RAID / SCSI   CAS   ATA   Tape   Optical  
**Open systems storage platforms**

# DatabaseXtender for Oracle, Custom, and PeopleSoft Applications Environments

- Identify transactions to be “pruned”
- Relocate inactive transactions to an online archive
- Identify transactions no longer needed for reporting (aged)
- Relocate aged transactions to a searchable stand-alone archive
- **The right data in the right location, at the right time, at the right cost, within compliance**

## Data Retention Policies





## Intelligent Online Archiving Software+Comprehensive Back Up and Recovery

- Policy-based data movement over tiers
- Seamless view of enterprise information
- Helps support governance requirements
- Integrated with key applications / environments
  - E-mail (EMC EmailXtender)
  - Filesystems (EMC DiskXtender)
  - Databases (EMC DatabaseXtender)
  - Enterprise Content (EMC Documentum)
- Support for heterogeneous solutions
  - Third-party storage support
  - ISV partner integration





**EMC<sup>2</sup>**  
**where information lives<sup>®</sup>**