

# Rescuing the data files . . .



Leading practices in cleaning up shared drive content “dark data”

Steve MacWilliams, CDIA



# Lifeguard Tendencies



- Safeguarding is in our “nature” business drivers think differently...

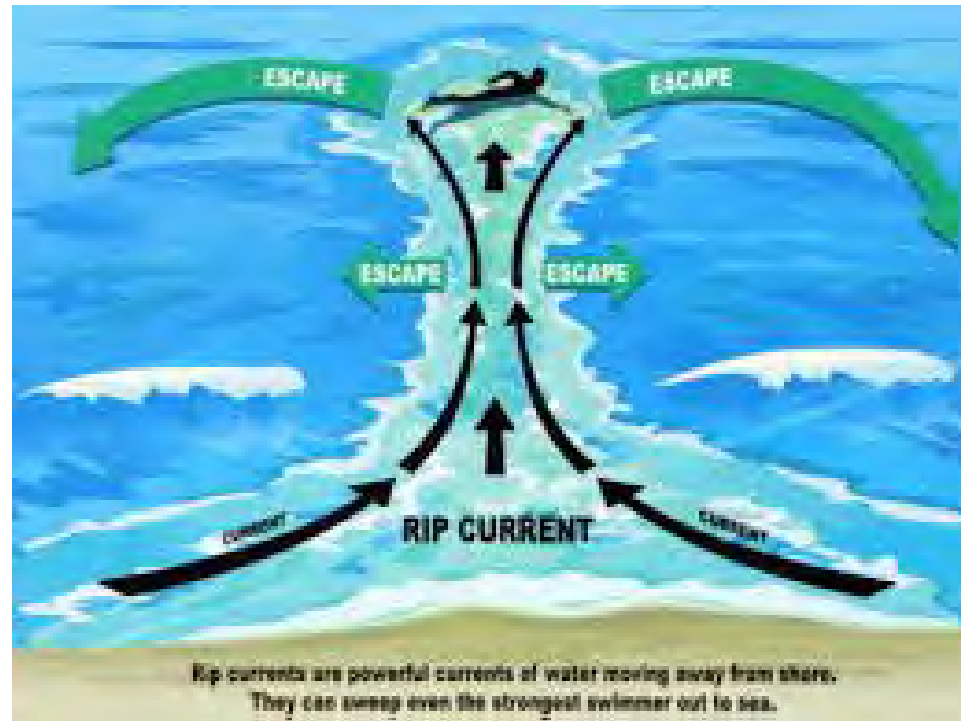


File Management	Document Management	Records Management
Collaboration is the goal, and to exchange information on a one-time basis	Goal is to produce and manage a work product	Understanding beyond the “Systems of Record” versus the “System of Archive”
Quick to Start	Short Term Liability function	Long Term Liability function
Free-flow process	Semi-Structured	Very Structured
Little Control	Control at the file level	Control at a collection of files level
Live in the now !	Does not manage beyond the transaction at hand	Manages beyond the transaction at hand
Phone Call	Short Story	Books and Volumes

# Riptides of the time...



- File Shares
- External Shares
- Email
- SharePoint
- Cloud based data



Rip currents are powerful currents of water moving away from shore. They can sweep even the strongest swimmer out to sea.

# Where is “Hasselhoff” when you need him ?



- What is the Information Policy ?
- What are the official safe zones?
- Governance versus



# 10 Steps to Success



1. Confront Reality
2. Create a Governance Plan
3. Get an Exec Sponsor
4. Create a Dream Team
5. Builds a Service NOT install software stuff
6. Define Clear Policies and Standards
7. Invest in Scalable Information Architecture
8. Don't forget Change/Risk Management
9. Adoption is "What Counts" !!!
10. Oh yea... keep it real simple..., but significant

# 1. Confront Reality



If a tree falls in the woods does it make a noise ?

If files are never looked at “ever again” is it still information...



IT Control



User  
Empowerment



## 2. Have a Governance Plan



*Governance uses people, process, technology, and policies to define a service, resolve ambiguity, and mitigate conflict within an organization . . .*







# People, Process, Technology, Policy

- Clear Vision, Mission and Goals

## Defaults have Faults

### Server

- Basic Install (BAD!)
- All On C Drive
- No Backup
- Restore Complexity

### List and Site Collection

- All One site collection
- Not Configured: Quota, Expiration, Auditing, Versions, Workflows, Meta Data capture
- All items view default (flat)

### Database

- One database
- Db config (log and db location)
- No maintenance plan



### 3. Get a “passionate” Executive Sponsor



4. One is not enough,  
You need a Dream team too...



# Case Study #1: City of Oakland



Multiple repositories, objective to have a “single version of the truth”  
Selected Oracle *WebCenter* as target; mission is to move property based documents in from multiple file shares and legacy systems...

- How to we get from here to there? “id dilemma”
- Complete it within timeframe of 1 year
- Planning, Building, Permits and Code Enforcement
- APN# and Address or LOT numbers



## Capture Software

- Sweeping of files with OCR software that looked for keywords;
  - OCR process looks for key values on any documents from PTS, LaserFiche, Stellant, Drives, Cloud bases Accela- Source leveraged.
  - Leveraged fuzzy logic to solve for APN# from owner name and cross-referenced to property
1. Lessons learned: keep “Dream Team in check.
  2. Records management was not in the mix of architecture.



# Points of failure (on the seams)

## Operations and IT Failures

### Server

- Poor configuration (Not planned - defaults)
- Insufficient RAM and Disk allocation

### Information Architecture

- Poor choice of Site vs. Site Collection Structure
- Flat Lists
- One Content Database

### Network

- Poor Network between WFEs and SQL
- No plan for WAN (User complaints without answers)

### Operational Backup/Restore

- No Backup
- Emergency: No Restore plan or steps
- Disaster Recovery – More like Disaster

## Developer Failures

### Ramp up/Lack of Skills

- Memory Leaks – USE SpDisposeCheck tool
- Not closing Objects
- Unaware of community tools

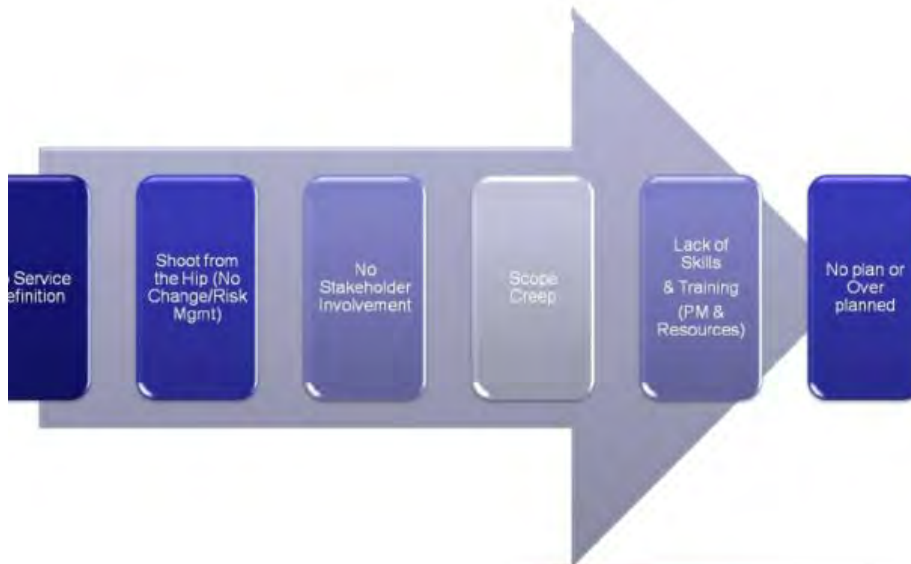
### Poor Assumptions

- Unnecessary Custom Site Defs or editing default site def
- Thinking more is better
- Naked assemblies and editing web.config(s) isn't enough

### Easy way rather than right way

- Skipping to production (Dev and test environment – REQUIRED!)
- Wrap it up – WSP Solutions Required – WSPBuilder
- Manage your source TFS!?

## Project Manager Failures



## Business Deployment Failures

### Communication

- Duplication of effort
- No comm. No buy in (forced deployment)

### Culture & Adoption

- Lack of End User Training/Community
- Lack of Help and Insight into Business processes

### Budget/Sponsor

- No Budget - No project or Budget cuts
- No Sponsorship – Under the Radar

### Poor Planning

- Poorly designed taxonomy & Information Architecture
- No deployment due to complex dependencies



# Data Governance Rolls and Responsibilities



## Executive sponsor

Must be:

- Identified early
- CXO – level
- An active participant
- An evangelist
- The Final point of escalation

Can be a committee, if appropriate



## Program driver

Must be:

- Unbiased
- Coordinator of all stewards
- Commination all decisions made by stewards to all relevant stakeholders
- Driving ongoing data quality auditing and metrics to assess program success
- Primary point of escalation to executive sponsor or committee, when necessary



## Business stewards

Must be:

- IT – savvy
- Strong communicators
- Educators across the business



## IT stewards

Must be:

- Business-savvy
- Strong communicators
- Educators across IT



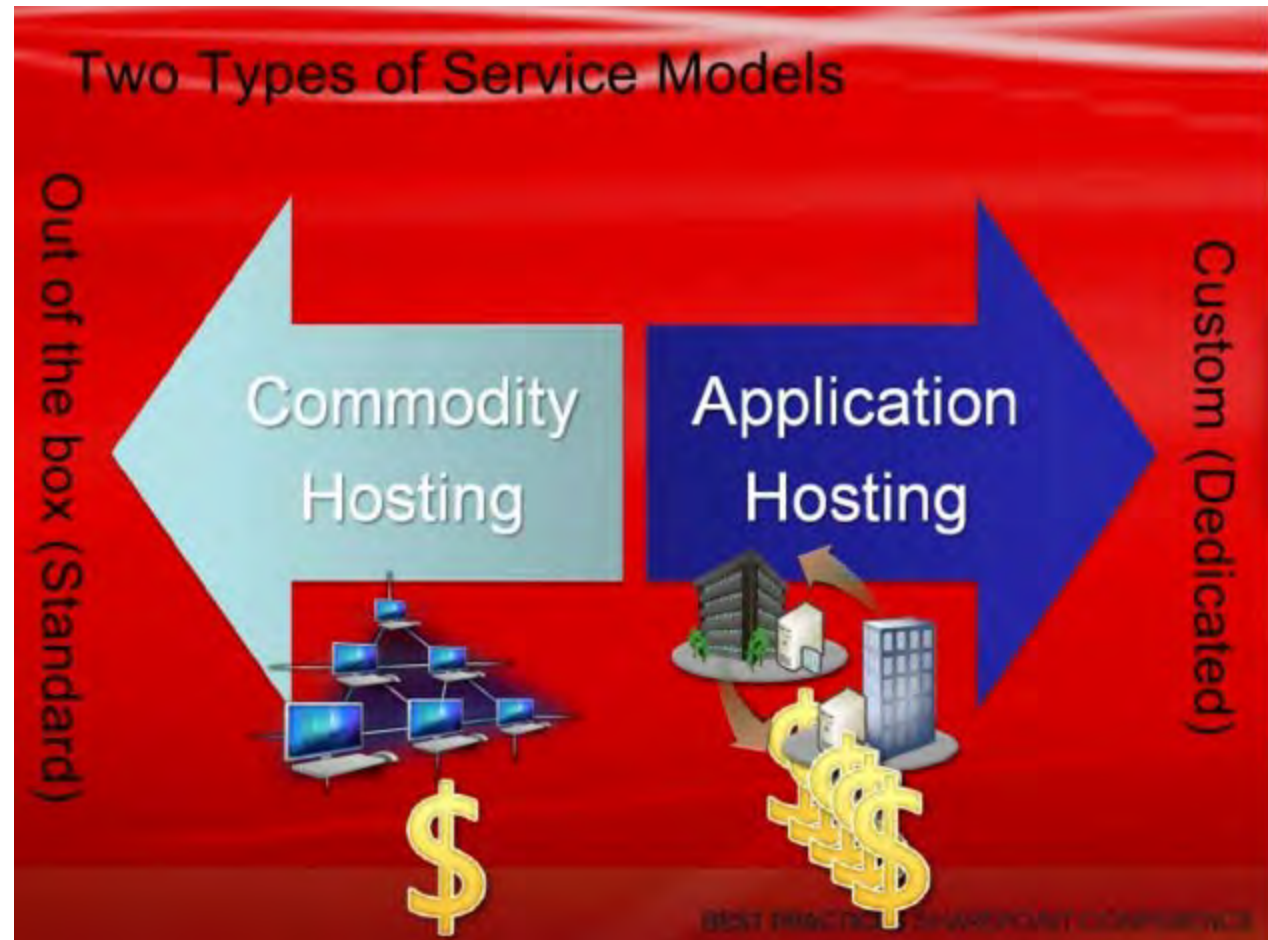
# 5. Build a "Service" not install Software





# What is your life raft?

Long file names  
Cloud Hosting  
Outlook  
SharePoint  
ECM Tier 1  
ECM Tier 2



# 6. Define Clear Policies and Standards



Whose fault is it anyway... spin the dial





# What makes for the great outcome?

Standardization



Flexibility



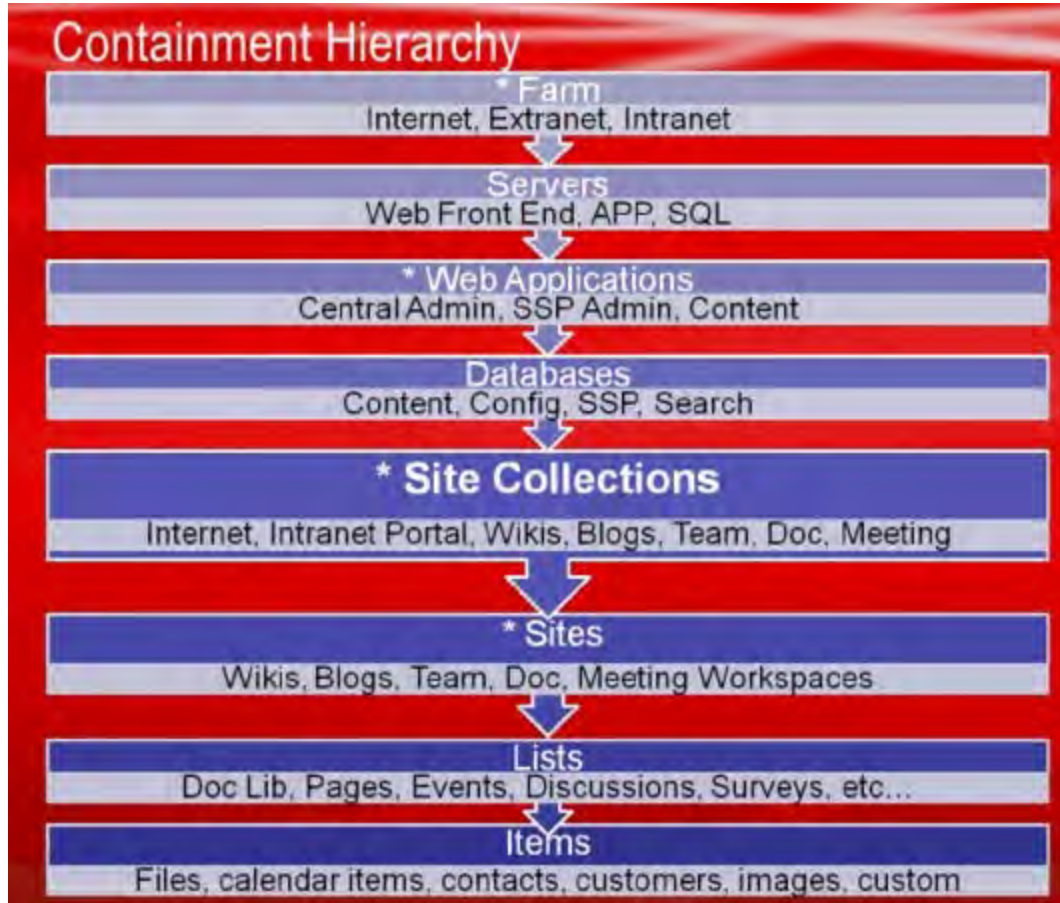
# Policies support Services



Policies Support Services		
IT Hosted Services		
<b>Site Collection Hosting</b> <ul style="list-style-type: none"><li>• Low cost to Free (low barrier)</li><li>• Self Service (w/ meta data capture or workflow)</li><li>• Fixed Quotas</li><li>• SharePoint Designer (optional)</li><li>• Out of box templates + Application Templates</li></ul>	<b>Web App Hosting</b> <ul style="list-style-type: none"><li>• \$ Charge Back</li><li>• Multi Tenant</li><li>• Web App/App Pool</li><li>• No GAC access</li><li>• Custom Site Templates</li><li>• Limited LOB Integration</li></ul>	<b>Premium App Hosting</b> <ul style="list-style-type: none"><li>• \$\$\$ Charge back</li><li>• Dedicated</li><li>• Virtual or Physical Separation</li><li>• Custom LOB Integration</li><li>• Business Focused Dev</li><li>• Dev/Test environments</li></ul>



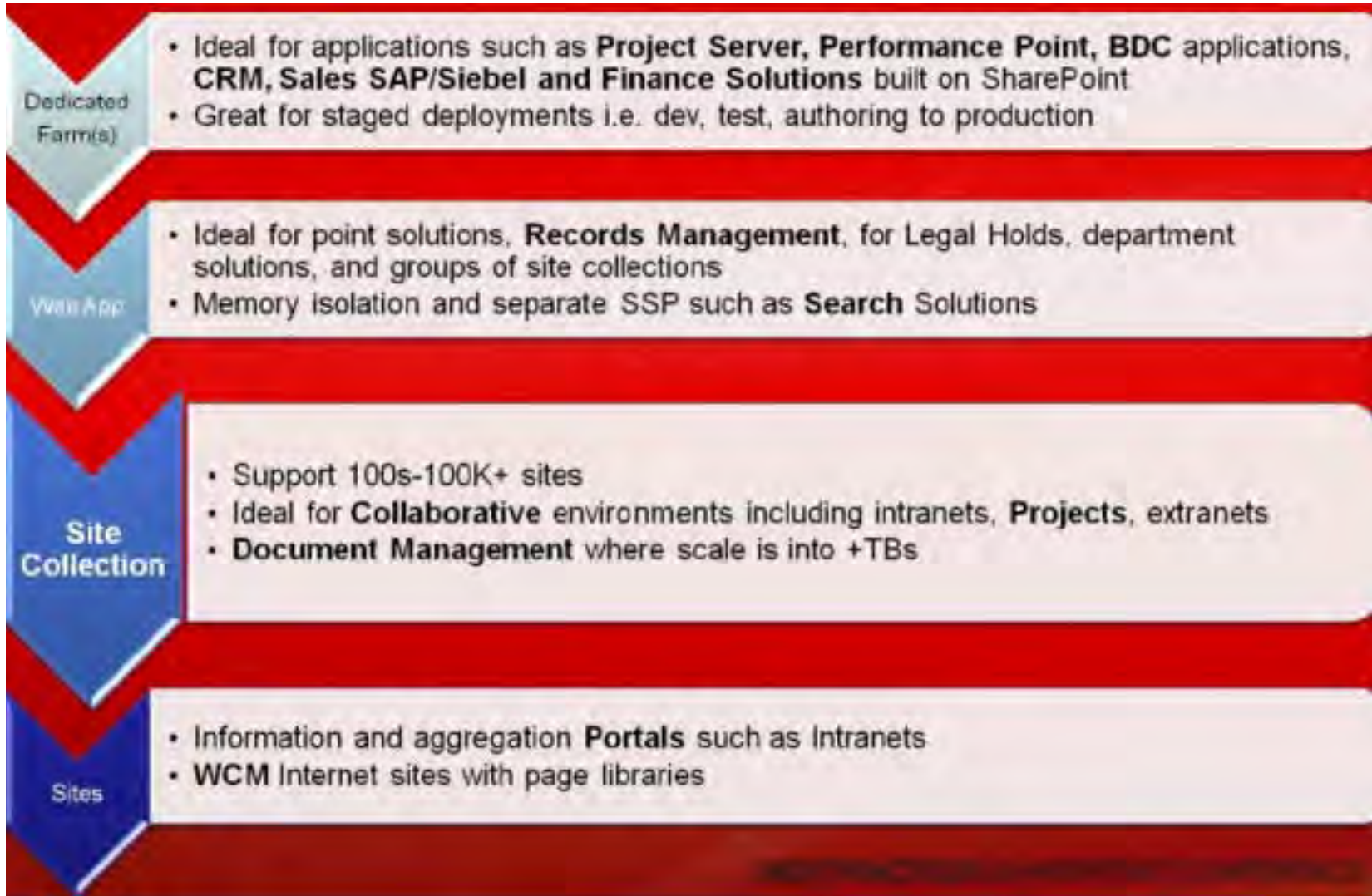
# 7. Invest in Scalable Information Architecture



Build for Today

Design for Tomorrow

# Cheat Sheet for to Host or Not ?



# Ready for the Cloud ?





# Case Study #2: Healthcare provider



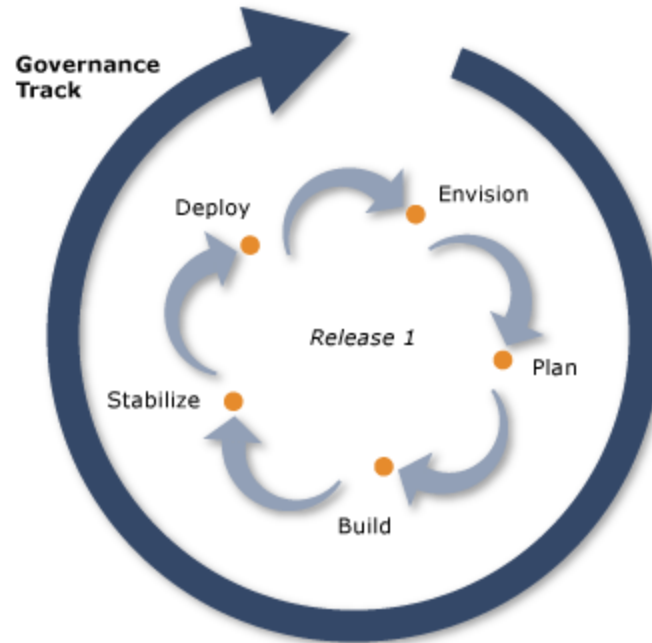
Thousands of files in file shares had no governance. First solution proved cumbersome of migration to SharePoint did not solve the governance and or unit-record need. A form based tool (Liquid Office) that was to move files through process was disconnected with file naming conventions.

## DIP or Document Import Processor



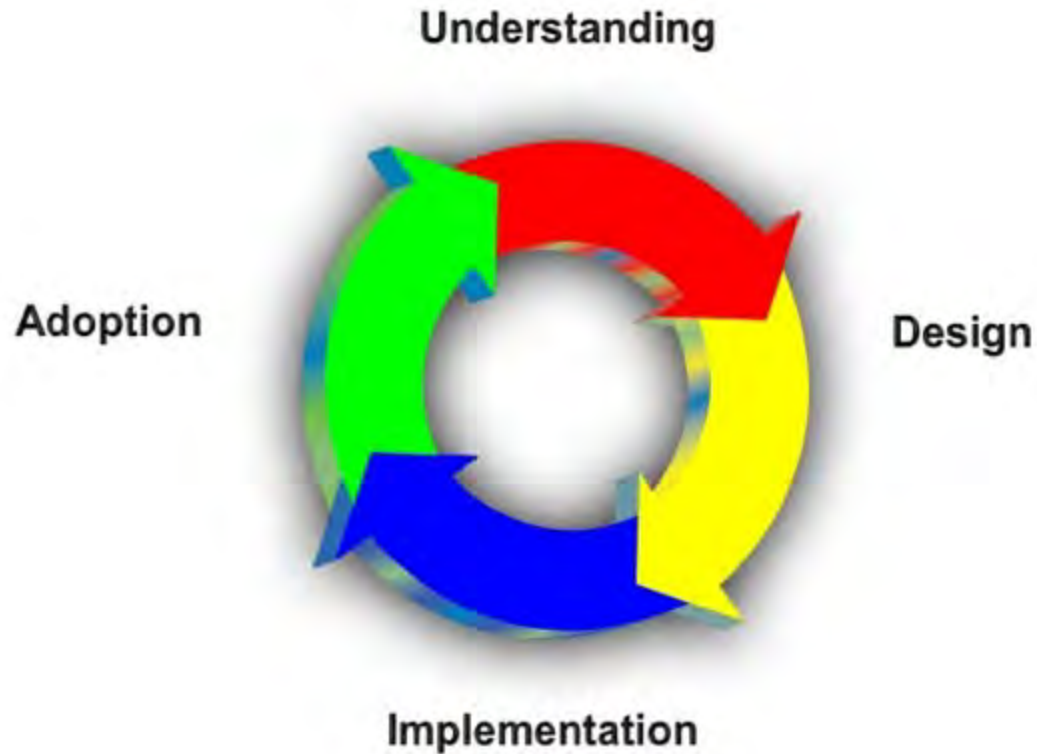
- The naming convention at the folder level and filename convention had some clues of what the file was, but no consistency
  - Used a software technique called “forced multiplier” where we could use the file name syntax to look out over the mass of files and classify in bulk into a Cloud based ECM that has governance, RIM, and structure built into it.
1. Taxonomy needs to be part of the product design, not necessarily designed from ground-up
  2. Multiple development teams do not make for improved outcome

# 8. Don't forget Change / Risk Management





# 9. Adoption is what Counts !



# Water daily...

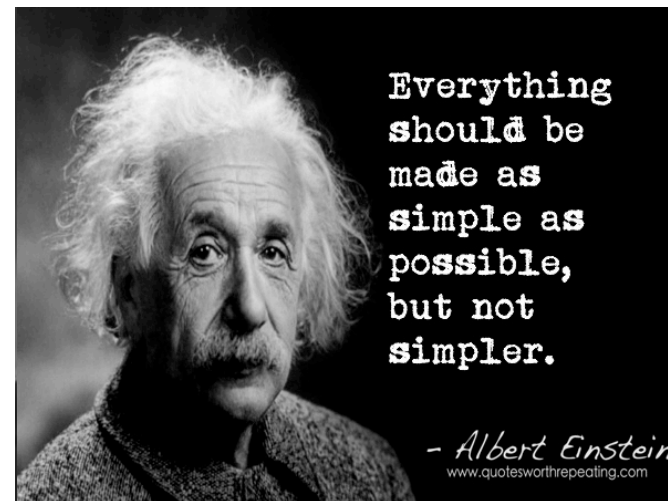


# 10. Keep it Simple,



**MAKE IT SIMPLE,  
BUT SIGNIFICANT.**

DON DRAPER



- Do you really need that list within a list, OCR extraction, or proximity searching or wiz bang widget?
- Don't start with the LOB integration and BDC and Excel Services requirements on your first go... start easy and work up...
- Start with Collaboration, Search and Profiles, it's an easy win.
- Watch for "Scope Creep" it is very easy in document managing to happen.



# Recap “memorable lifesaving skills”



Build target first, conversion plan second . . .

Build support at highest management level and create a dream team . . .

Balance productivity with liability . . .

Leverage conversion tools to move files into new taxonomy . . .