

# Designing Storage Architectures 2013

Library of Congress, Washington, DC

Mike Thuman Associate VP, Digital Archiving

September 23, 2013

# Tessella Archiving Solutions (TAS)

Global team of specialists(Focused on one domain)

- Consultants
- Project Managers
- Developers
- Architects
- Support

Assessments, building requirements docs, developing tools, some products and now working with Cloud

Education and outreach

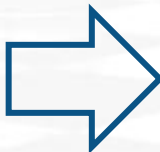
A decade of practical experience



FM22778

# Digital preservation, democratized

Safety Deposit Box  
Application



**Preservica**<sup>®</sup>  
Digital Preservation

- Based on pilots and collaborations within the community
- Any file format accepted, 800 formats recognized, 300 normalization/migration pathways
- Productive with 1 day of training

# Digital Preservation - A Comparison of Two Formats



Which will be useable in 100 years?

What about 10, 20 or 30 years?

T. Cramer, Stanford University, PASIG Digital Preservation Bootcamp

# Mass market items coming of age-1983

**How you can tell which is which.**

They're both IBM Personal Computers that they're both really different. The system that's on the left is the one that's for people who want to be creative, efficient and improve his or her personal productivity.

So is the other one.

The system on the left incorporates the quality, the reliability and the technology of excellence that have made IBM a computer leader for over 50 years. So does the other one.

So which is which? Simple.

On the left is the IBM Personal Computer, starting with 640KB of user memory (expandable to 1,024KB) and two expanded 5.25" diskettes. Also, it can easily accept your computer needs as they change over time.

With 3 expansion slots, it gives you flexibility as well. Also comes with a function key that the computer knows on its right by adding an expansion unit that houses one or two 3.5-inch diameter hard disk drives.

This system can run most of the same software and support most of the same IBM hardware as the computer on the right. And its price performance is nothing less than remarkable.

On the right is the IBM Personal Computer XT, starting with 256KB of user memory (expandable to 640KB), a 3.5" floppy diskette drive plus a standard 15-inch monitor, built-in fixed disk drive that stores up to 10 megabyte data requirements, the XT packs a lot of power because it uses more than 800,000 transistors and contains 200,000 lines of code (instead of going from 100,000 lines of code to 1,000,000 lines of code) or up to 100,000 notes and addresses (compared to 10,000) and a local 3.5-inch floppy disk drive.

And by adding an expansion unit with capacity for up to 1,024KB of user memory and up to 100,000 notes and addresses (compared to 10,000) and a local 3.5-inch floppy disk drive, you get even more high-end performance from the system.

XT can be used as the computer on the left or as a replacement for the computer on the left. And its price performance is nothing less than remarkable.

But for you to choose, there's a lot to think about. Check with your local IBM representative or IBM Personal Computer dealer. To learn more, call 800-447-4756. If that's not possible, call 800-521-5215. And the whole lot of details come right on 447 for you.

**IBM**

The IBM Personal Computer

The IBM Personal Computer XT

The XT came with 4.77MHz 8088 processor, a 10 MB hard drive, 128KB RAM and a 360KB floppy drive. For this top-of-the-line business machine, you only had to pay \$5,000.

Apple  
invents the  
personal  
computer.  
Again.

Lisa -- complete with Motorola 5MHz 68000, 5MB hard drive, "huge" 1 MB of RAM, and Apple's first graphical operating system, Lisa Operating System. And, you could have all this for "only" \$10,000.

The Mac, introduced for \$2,500, proved far more popular.

TRS-80 HARD DISK SYSTEM FOR MODEL II, 12, AND 16

**New! 12-Megabyte Hard Disk Drive Forms a Powerful Computer System**

Primary Hard Disk System	<b>3495<sup>00</sup>*</b>
Secondary Hard Disk System	<b>2495<sup>00</sup>*</b>

ZDNet, Top Holiday Tech Buys... of 1983,

[Steven J. Vaughan-Nichols](#) for [Networking](#) | December 15, 2012





**What would the next 30 years bring?**

**The “Lost” Steve Jobs Speech from 1983;  
Foreshadowing Wireless Networking, the iPad,  
Google Maps, Apple’s App Store and more**

# The digital preservation “Use Case”

“Provide continued access to an authentic copy of electronic content over long periods of time”

- Technology evolves at a seemingly exponential pace leading to obsolescence of hardware, software and media
- The retention period of electronic content will exceed the lifetime of the software and hardware that was used to create and access them
- Ensure the integrity of the bitstream throughout its retention
- Future renditions of the content **MUST** be faithful to the original
- The preservation solution must be seen as a Trusted Source

# Resulting Systems

Ingest, identify, characterize, extract metadata, basic safety checks (format validation, virus), **store**, preserve, analyze collection for risks.....

Workflow Details

Workflow Context	Manual Select Ingest
Workflow Definition	Ingest Workflow (Manual Selection)
Workflow ID	208
Workflow State	Completed
Date Started	06.07.12 08:29:52
Date Finished	06.07.12 08:31:14
Number of Files	14
Total Size	1 MB
Collection Code	HANTS
Submission name	Hampshire County Archives
Top Level Record	Romsey

Back

Step Progress

State	Name	Progress	Started	Finished	Messages
	Select	<div style="width: 100%;"></div>	06.07.12 08:29:52	06.07.12 08:29:58	
✓	Import from Transfer Area	<div style="width: 100%;"></div>	06.07.12 08:29:58	06.07.12 08:30:02	
✓	Virus Check	<div style="width: 100%;"></div>	06.07.12 08:30:02	06.07.12 08:30:32	
✓	Fixity Check	<div style="width: 100%;"></div>	06.07.12 08:30:32	06.07.12 08:30:35	
✓	Metadata Integrity	<div style="width: 100%;"></div>	06.07.12 08:30:35	06.07.12 08:30:38	
✓	Content Integrity	<div style="width: 100%;"></div>	06.07.12 08:30:38	06.07.12 08:30:41	
✓	Characterise	<div style="width: 100%;"></div>	06.07.12 08:30:41	06.07.12 08:30:50	View
✓	Store Files	<div style="width: 100%;"></div>	06.07.12 08:30:50	06.07.12 08:30:56	
✓	Store Metadata	<div style="width: 100%;"></div>	06.07.12 08:30:56	06.07.12 08:30:59	
✓	Delete from Transfer Area	<div style="width: 100%;"></div>	06.07.12 08:30:59	06.07.12 08:31:02	





# Who's Doing This



NATIONAL  
LIBRARY  
OF AUSTRALIA



Wellcome Library



Kentucky Department for  
Libraries and Archives



Hagley Museum and Library



EMERSON COLLEGE

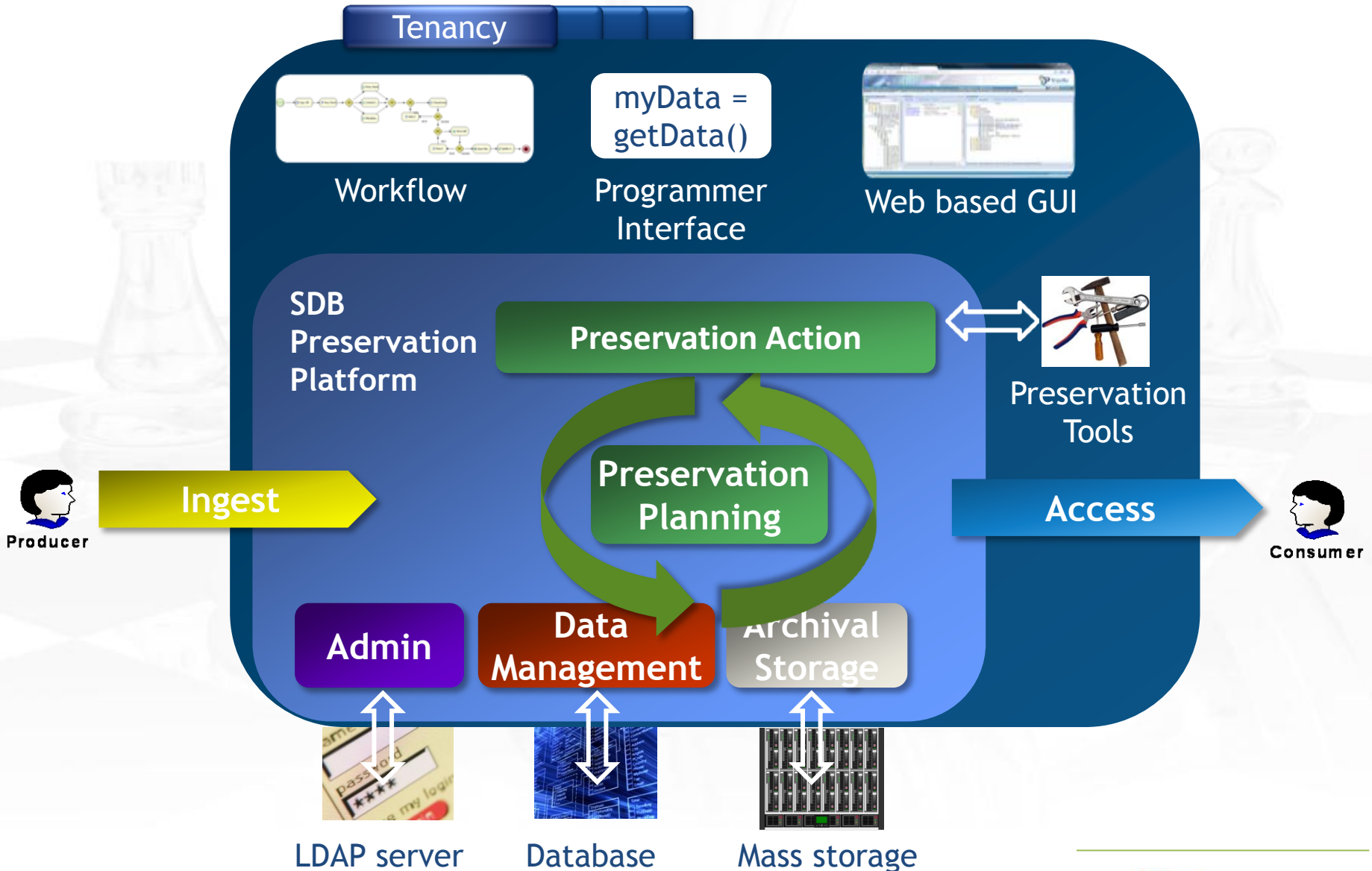


Colby·Sawyer  
College  
*inside · outside · planetwide*

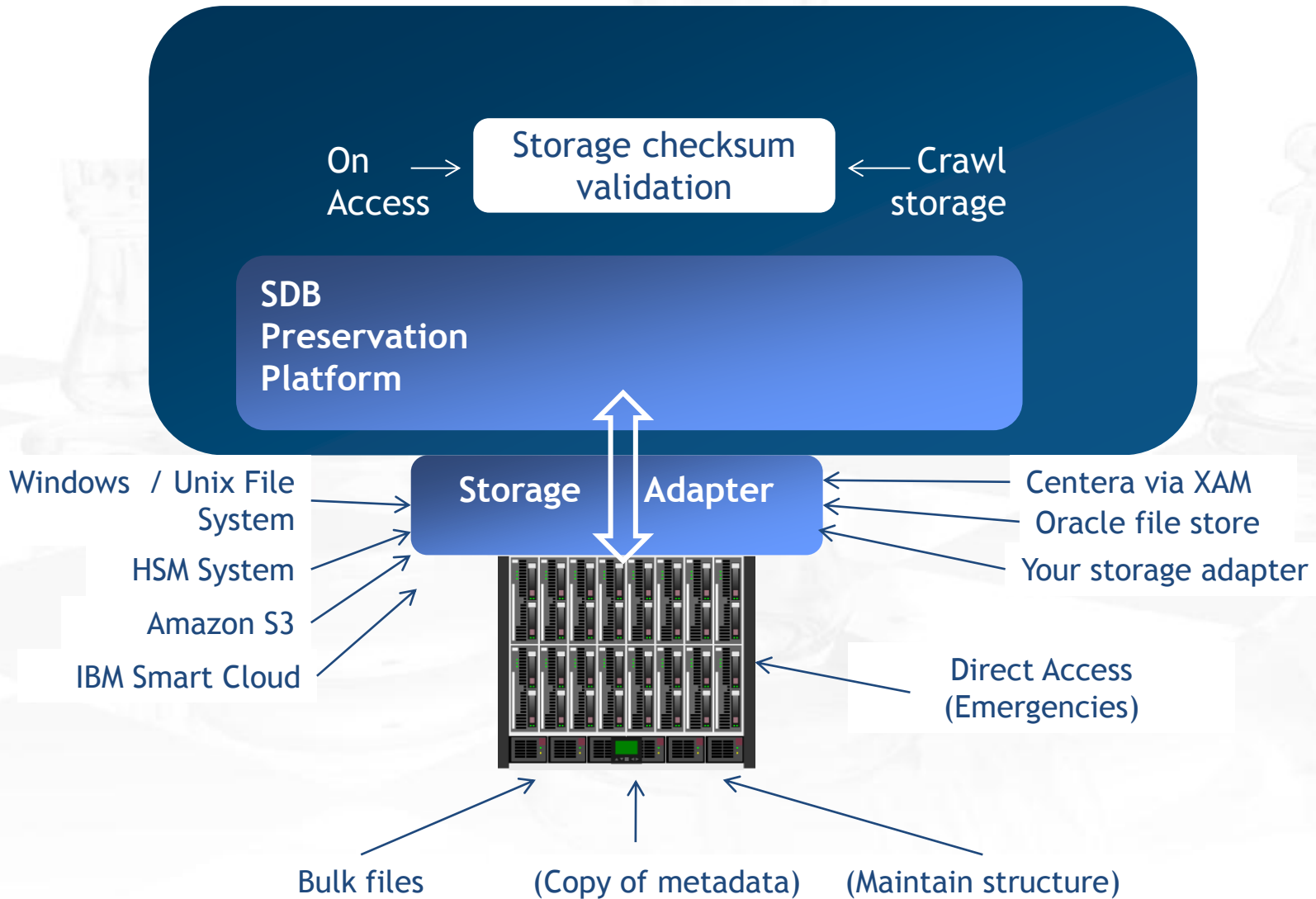
Archives of Michigan

- World's largest collections on the history of medical science and its role in society. Malaria project, Maintained to support researchers driving to make extraordinary improvements in health.
- K-12 education data, precinct level election results, state Supreme Court briefs and meeting minutes of various public bodies.
- Researchers accessing the records of American businesses and including company history, technology, innovation and industrial design. Ensures future of the history of US business start-ups, venture capital and technological development. Enhances reputation as the trusted repository of business history.
- Historical collection of college photos, student newspapers and precious letters (e.g. Ralph Waldo Emerson). Allows students & alumni to access historical records with ease. Improves brand and on-going relationship with key stakeholders.

# Components of a Digital Preservation System



# We are consumers of storage





# One size fits all?



- Cloud to Local Edition
- Local Edition to SDB Max
- SDB Max to Private Cloud

**Safety Deposit Box  
Max**

Your data centre

**Preservica LE**



Private cloud

**Preservica**<sup>®</sup>  
Digital Preservation



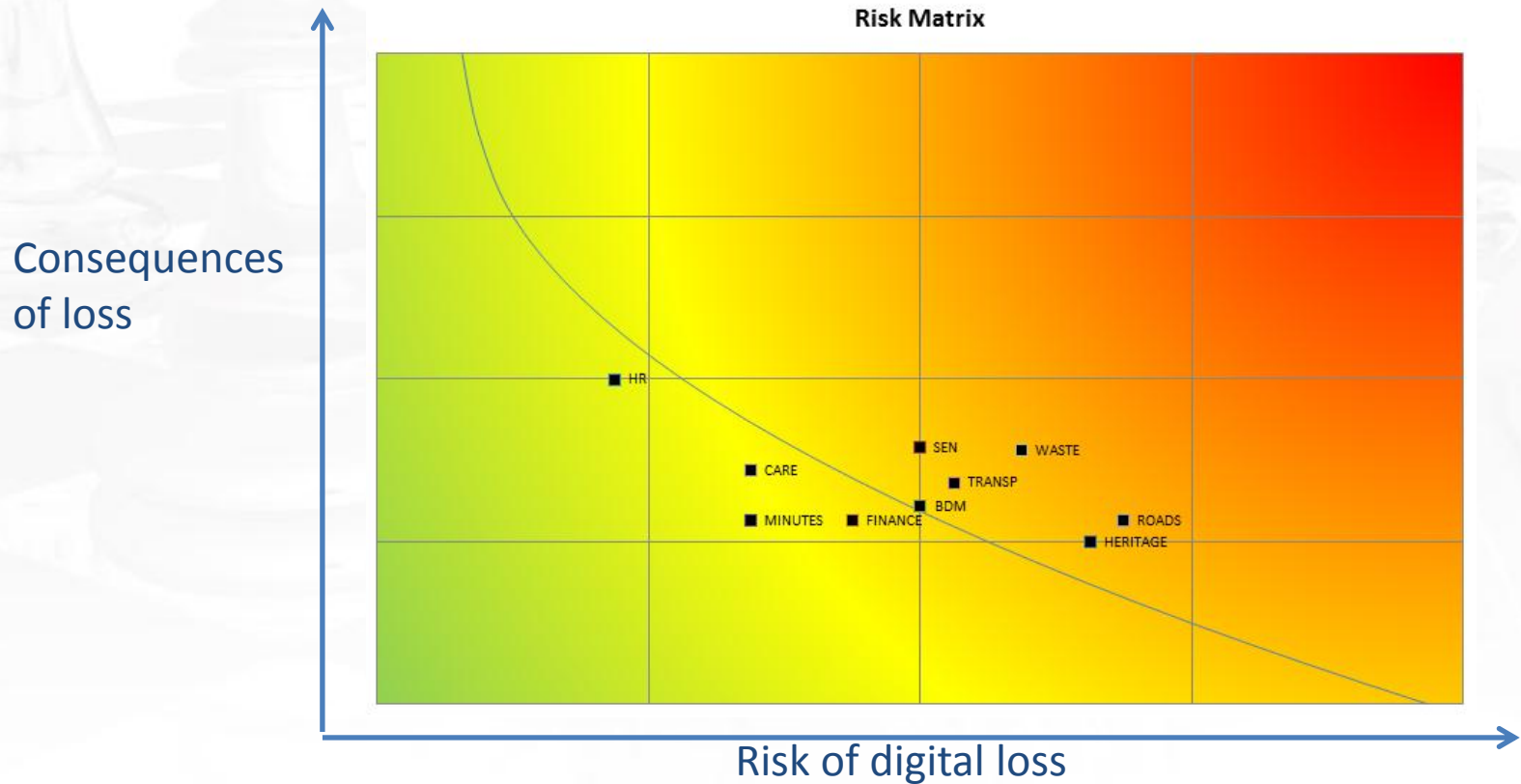
Multi-tenant, From Platform to Cloud  
Safety Deposit Box 4.x



# New tools

## Digital Value at Risk Calculator

Videos, Case Studies, D-VAR Calculator,  
Business Case Templates  
[www.preservica.com](http://www.preservica.com)





# Thank You!

[mike.thuman@tessella.com](mailto:mike.thuman@tessella.com)

435.655.2814