Introduction
What is the Digital Vault?

The Digital Vault is to the Bank’s digital records.

What the WBG Record Center is to the Bank’s paper records...
The Digital Vault enables us to apply digital preservation best practices to protect, know, monitor, repair, and provide access to the digital records of the World Bank Group Archives.
Regardless of their form or medium, accurate, authentic, reliable, complete, and accessible, records must be retained or systematically and securely deleted. WBG Archives custody and control over access to records over 20 years old.
<table>
<thead>
<tr>
<th>Electronic Records</th>
<th>Analog Records</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evidence of World Bank Group’s activities</td>
<td></td>
</tr>
<tr>
<td>Rapid rate of technology change shortens the window for</td>
<td>Generally can be safely stored by record creators for 20 years.</td>
</tr>
<tr>
<td>effective record capture.</td>
<td></td>
</tr>
<tr>
<td>Exceedingly diverse formats:</td>
<td>Predictable formats: paper documents in folders, maps, photographs, audio/video</td>
</tr>
<tr>
<td>documents, structured data, audio/video, data within</td>
<td>media</td>
</tr>
<tr>
<td>applications</td>
<td></td>
</tr>
</tbody>
</table>
Google is not the answer: How the digital age imperils history

From floppy disks to thumb drives, we get better at storing things — while trapping history in obsolete formats

JON Palfrey

Our species created about 5 billion gigabytes of information from the dawn of time until 2003. Before long, we will create that much information many times per day, according to IBM. The problem: No one is doing enough to select and preserve the bits that really matter.

Google boss warns of 'forgotten century' with email and photos at risk

Digital material including key historical documents could be lost forever because programs to view them will become defunct, says Vint Cerf

Vint Cerf: "We are nonchalantly throwing all of our data into what could become an information black hole."

Ian Sample, science editor, in San Jose

Friday 13 February 2015 04:16 EST
In December 2018 the Maine state government revealed that it had suffered a catastrophic loss of public documents from the administrations of Governors Angus King and John Baldacci, stretching from 1995 to 2011. Most state-government emails sent before 2008 were irretrievably lost. Moreover, many other kinds of documents had been destroyed by state officials before they reached the state archives. Not only has information for future historians been lost, but these emails could also contain information that could be vital in legal cases.
Years of Mark Zuckerberg's old Facebook posts have vanished. The company says it 'mistakenly deleted' them.

Some old Facebook posts by CEO Mark Zuckerberg have vanished, and the company says it "mistakenly deleted" them.

The disappearances include posts about key moments in the company's history, like the acquisition of Instagram in 2012.

"A few years ago some of Mark's posts were mistakenly deleted due to technical errors," a spokesperson said in a statement.

Facebook won't restore the posts, saying the work to do so would be "extensive" and not guaranteed to work.

Facebook has also made changes to the way it saves its archives of previous corporate announcements and blog posts that make it much harder to parse the company's historical records.

Old Facebook posts by Mark Zuckerberg have disappeared — obscuring details about core moments in Facebook's history.

On multiple occasions, years-old public posts made by the 34-year-old billionaire chief executive that were previously public and reported on by news outlets at the time have since vanished, Business Insider has found. That includes all of the posts he made during 2007 and 2008.
Worry not. We have everything Mark Zuckerberg has ever posted or said in public archived at The Zuckerberg Files.
@ZuckerbergFiles zuckerbergfiles.org

Rob Price @robaeprice
SCOOP: Years of Mark Zuckerberg's old Facebook posts have vanished. The company says it 'mistakenly deleted' them. businessinsider.com/facebook-old-p...
Show this thread

7:51 AM - 30 Mar 2019
544 Retweets 1,329 Likes
The Zuckerberg Files is a digital archive of all public utterances of Facebook's founder and CEO, Mark Zuckerberg, spanning 2004-2019. Nearly 1,000 full-text transcripts and 250 video files are available for researchers to download, analyze, and scrutinize.
Business Value

Ensure long-term preservation of and access to authentic digital records determined by WBG records schedules to be permanent records of the World Bank Group.

Holistic solution that covers all functions defined by best practices for digital preservation.
Digital Preservation Challenges

1) Do we have the 1s and 0s?
Store multiple authoritative copies of permanent digital records in archival custody. Protect and validate the copies. Document all activities performed on the records.

2) Can we find and understand the records?
Store records with appropriate metadata and contextual information.

3) Can we view and interact with the records in a meaningful way in the future?
Enable long-term access to digital records regardless of their format, originating system, or technology obsolescence.
What is Digital Preservation?

STORAGE

CONTEXT

ACCESS
Why is a backup not enough?

Backups are designed as short-term insurance against disasters.

Backups take a snapshot of entire environment at a single point in time – and are restored for entire snapshot at once.

Backups stores content in case of operator error, network failure, natural disaster, and hardware issues.

Backups do not store additional contextual information or metadata.

Backups do not address issues of format or technology obsolescence.
How are digital records identified for preservation?

Records identified as having permanent value based on the World Bank’s Records Retention Schedules will be transferred to the WBG Archives’ Digital Vault for permanent storage. Other records will be identified for destruction.

Records Retention Schedules are drafted in close collaboration with the relevant business unit, approved and signed by the relevant VP and issued by the WBG Chief Archivist.
WBG Digital Records identified for permanent retention.

Permanent digital records are safe, in context, findable, and accessible.

Digital Vault

Technology + WBG Archives Staff Expertise
Building a Foundation for Success

- Existing Policies
- Business Value
- Leadership Support
- Funding & RFP Process
- Internal WBG Archivists
- Technology Allies

Digital Vault
Stakeholders and Processes

Record Creators
- Digital objects created
- Digital objects become “inactive” or system decommissioned

WBG Archives
- Pre-Appraisal, Appraisal, & application of record schedules
- Identification of records requiring permanent retention

Digital Vault
- Ingest & storage of digital objects
- Manage & preserve digital objects
- Research for relevant records

Record Requestors
- WBG Staff request for records
- Public Access to Information request
- Access copies of digital objects received
- Access copy of digital object generated as needed
Digital Vault Main Functions

**Ingest**: The steps required to transfer items from their current location into the digital vault in a managed manner.

**Archival Storage**: The storage of the bulk data (usually files).

**Data Management**: Tools to manage the storage of the archive, including the metadata.

**Administration**: Tools to administer the system and access to it.

**Access**: Tools to search, browse, convert to modern file formats, and download the contents of the Digital Vault.

**Preservation Planning**: Activities required to ensure information can be accessed long into the future.
Digital Vault Technology Snapshot

- On-Prem Preservica Dashboard & Workflow Engine
- Storage/retrieval of digital objects
- Direct Connect
- 2 copies in Amazon S3
- 1 copy in World Bank Group data center
WBG Archives resources for adoption of the Digital Vault and the necessary evolution of our archival processes
Digital Vault Focus Teams

- Transfer & Ingest
- Description
- Access & A2I
- What gets put in the Digital Vault?
- Strategy & Planning
- Long-Term Access
Digital Vault’s benefits to World Bank Group:

- Reliability: “The trustworthiness of a record as a statement of fact.”
- Integrity: The “wholeness” and “soundness” of a record.
- Authenticity: “the quality of a record that is what it purports to be and that is free from tampering or corruption”
<table>
<thead>
<tr>
<th>What does success look like?</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Authoritative records stored safely with proper context</strong></td>
</tr>
<tr>
<td><strong>WBG Archivists empowered to find, describe, preserve, and provide access to permanent digital records of the WBG</strong></td>
</tr>
<tr>
<td><strong>Researchers empowered to discover, request, and use digital records in WBG Archives’ custody</strong></td>
</tr>
<tr>
<td><strong>A solid foundation for institutional transparency and accountability to stakeholders</strong></td>
</tr>
</tbody>
</table>
Questions?

Jeanne Kramer-Smyth
jkramersmyth@worldbankgroup.org