

Digital Access & Preservation for Video: Best Practices?!

Sean Chen (@gugek) &
Miguel Bordo (@miguelbordo)
Duke University School of Law

CALIcon 2013
June 14, 2013

DUKE LAW

Background & Scope of Project



DUKE LAW

Friday, June 14, 13

Introduction to Presenters (Miguel & Sean)

Miguel Bordo: Media Services Manager
Sean Chen: Digital Initiatives Librarian

Context (Miguel)

Where are we coming from.

- * Produce more than 100 events and smaller pieces a year
- * Huge number of tapes. 250 VHS;
- * 1000 or so DVCAM; about 100 or so MiniDV
- * More than 1000 DVDs

We've been in an intermediate period with physical media.

Eventual goal is to support a fully digital workflow.

Goals (Sean)

Preserve what we have.
Digitize for preservation.
Digitize for access.

Working closely with stakeholders:

- + Communications
- + Media
- + Library
- + Customers

Background & Scope of Project

DUKE LAW

Friday, June 14, 13

Introduce Video (Miguel)

Short introduction to Hollie and what why she is speaking here.

Hollie Video

- * Repository out of video
- * Working with stakeholders
- * Existing workflow
- * Preservation Tiers
- * Scope
- * Some of the things we are doing
 - * Checksums
 - * Format conversion
 - * Access

Overview

Friday, June 14, 13

Selection (Sean)

Process of selecting what is going to be digitized. Preservation and collection development policies.

Digitization (Sean)

Process of format conversion. Moving from analog to digital. Though this also includes format conversions from digital to digital such as DVD to file based. Or DVCAM to file based.

Description (Sean)

Metadata. Access points. Summaries. Important places where we think that our users and interested searchers will need to find, identify, select and locate our digital assets.

Ingest (Sean)

Placing into a trusted system. Recording data about this process. Setting certain preservation information such as our file location and checksums.

Storage (Sean)

Mechanical where this thing is going to live. How are we going to insure fixity. How is this thing backed up? How can we get it out.

Distribution (Sean)

Youtube and website and all the other places it goes

Other (Sean)

This is very similar to our already existing production workflow.

Overview

Selection

Friday, June 14, 13

Selection (Sean)

Process of selecting what is going to be digitized. Preservation and collection development policies.

Digitization (Sean)

Process of format conversion. Moving from analog to digital. Though this also includes format conversions from digital to digital such as DVD to file based. Or DVCAM to file based.

Description (Sean)

Metadata. Access points. Summaries. Important places where we think that our users and interested searchers will need to find, identify, select and locate our digital assets.

Ingest (Sean)

Placing into a trusted system. Recording data about this process. Setting certain preservation information such as our file location and checksums.

Storage (Sean)

Mechanical where this thing is going to live. How are we going to insure fixity. How is this thing backed up? How can we get it out.

Distribution (Sean)

Youtube and website and all the other places it goes

Other (Sean)

This is very similar to our already existing production workflow.

Overview

Selection



Friday, June 14, 13

Selection (Sean)

Process of selecting what is going to be digitized. Preservation and collection development policies.

Digitization (Sean)

Process of format conversion. Moving from analog to digital. Though this also includes format conversions from digital to digital such as DVD to file based. Or DVCAM to file based.

Description (Sean)

Metadata. Access points. Summaries. Important places where we think that our users and interested searchers will need to find, identify, select and locate our digital assets.

Ingest (Sean)

Placing into a trusted system. Recording data about this process. Setting certain preservation information such as our file location and checksums.

Storage (Sean)

Mechanical where this thing is going to live. How are we going to insure fixity. How is this thing backed up? How can we get it out.

Distribution (Sean)

Youtube and website and all the other places it goes

Other (Sean)

This is very similar to our already existing production workflow.

Overview

Selection



Digitization

Friday, June 14, 13

Selection (Sean)

Process of selecting what is going to be digitized. Preservation and collection development policies.

Digitization (Sean)

Process of format conversion. Moving from analog to digital. Though this also includes format conversions from digital to digital such as DVD to file based. Or DVCAM to file based.

Description (Sean)

Metadata. Access points. Summaries. Important places where we think that our users and interested searchers will need to find, identify, select and locate our digital assets.

Ingest (Sean)

Placing into a trusted system. Recording data about this process. Setting certain preservation information such as our file location and checksums.

Storage (Sean)

Mechanical where this thing is going to live. How are we going to insure fixity. How is this thing backed up? How can we get it out.

Distribution (Sean)

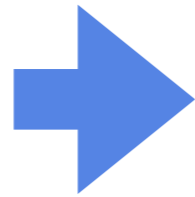
Youtube and website and all the other places it goes

Other (Sean)

This is very similar to our already existing production workflow.

Overview

Selection



Digitization



Friday, June 14, 13

Selection (Sean)

Process of selecting what is going to be digitized. Preservation and collection development policies.

Digitization (Sean)

Process of format conversion. Moving from analog to digital. Though this also includes format conversions from digital to digital such as DVD to file based. Or DVCAM to file based.

Description (Sean)

Metadata. Access points. Summaries. Important places where we think that our users and interested searchers will need to find, identify, select and locate our digital assets.

Ingest (Sean)

Placing into a trusted system. Recording data about this process. Setting certain preservation information such as our file location and checksums.

Storage (Sean)

Mechanical where this thing is going to live. How are we going to insure fixity. How is this thing backed up? How can we get it out.

Distribution (Sean)

Youtube and website and all the other places it goes

Other (Sean)

This is very similar to our already existing production workflow.

Overview

Selection



Digitization



Description

Friday, June 14, 13

Selection (Sean)

Process of selecting what is going to be digitized. Preservation and collection development policies.

Digitization (Sean)

Process of format conversion. Moving from analog to digital. Though this also includes format conversions from digital to digital such as DVD to file based. Or DVCAM to file based.

Description (Sean)

Metadata. Access points. Summaries. Important places where we think that our users and interested searchers will need to find, identify, select and locate our digital assets.

Ingest (Sean)

Placing into a trusted system. Recording data about this process. Setting certain preservation information such as our file location and checksums.

Storage (Sean)

Mechanical where this thing is going to live. How are we going to insure fixity. How is this thing backed up? How can we get it out.

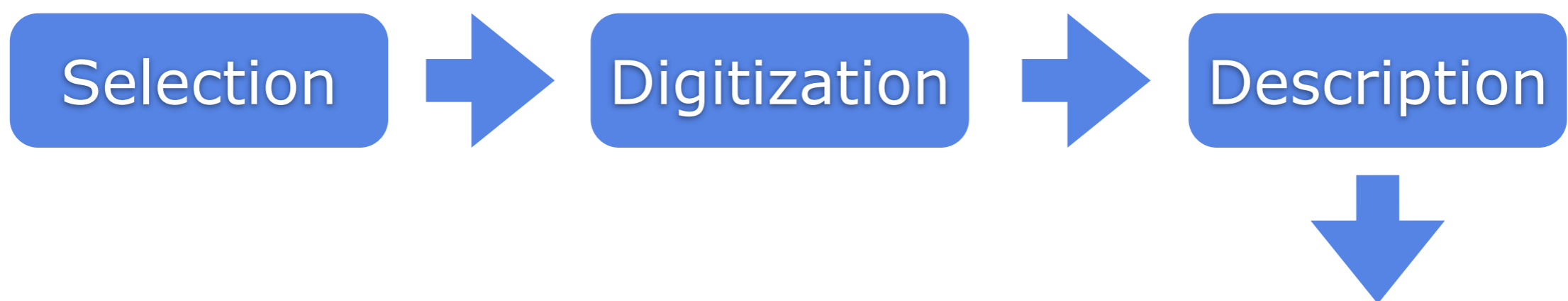
Distribution (Sean)

Youtube and website and all the other places it goes

Other (Sean)

This is very similar to our already existing production workflow.

Overview



Friday, June 14, 13

Selection (Sean)

Process of selecting what is going to be digitized. Preservation and collection development policies.

Digitization (Sean)

Process of format conversion. Moving from analog to digital. Though this also includes format conversions from digital to digital such as DVD to file based. Or DVCAM to file based.

Description (Sean)

Metadata. Access points. Summaries. Important places where we think that our users and interested searchers will need to find, identify, select and locate our digital assets.

Ingest (Sean)

Placing into a trusted system. Recording data about this process. Setting certain preservation information such as our file location and checksums.

Storage (Sean)

Mechanical where this thing is going to live. How are we going to insure fixity. How is this thing backed up? How can we get it out.

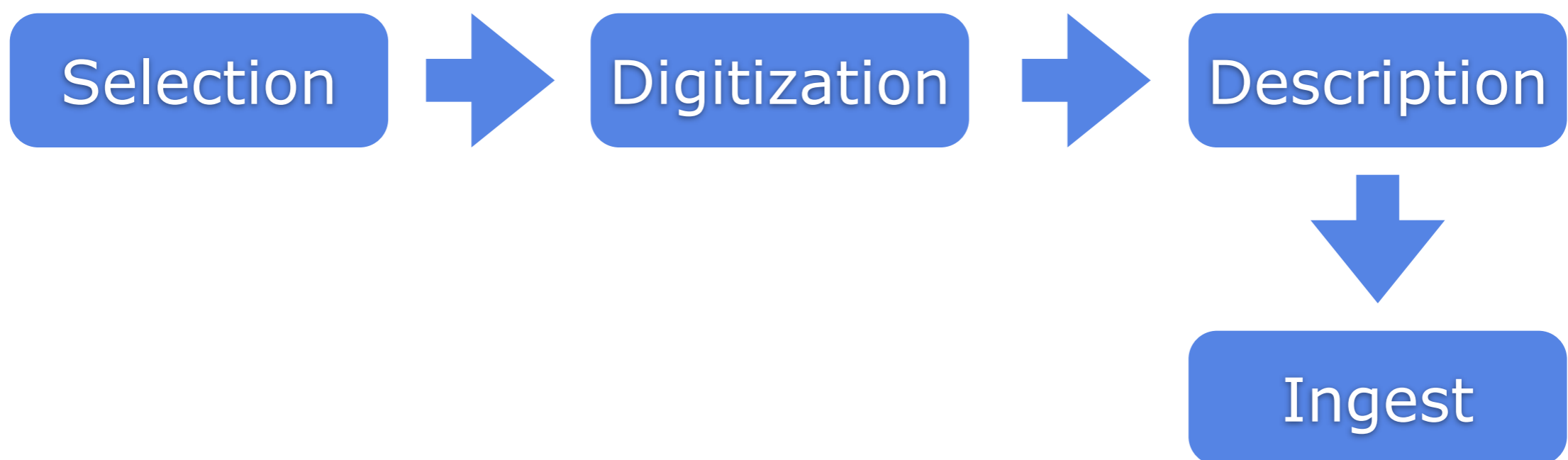
Distribution (Sean)

Youtube and website and all the other places it goes

Other (Sean)

This is very similar to our already existing production workflow.

Overview



Friday, June 14, 13

Selection (Sean)

Process of selecting what is going to be digitized. Preservation and collection development policies.

Digitization (Sean)

Process of format conversion. Moving from analog to digital. Though this also includes format conversions from digital to digital such as DVD to file based. Or DVCAM to file based.

Description (Sean)

Metadata. Access points. Summaries. Important places where we think that our users and interested searchers will need to find, identify, select and locate our digital assets.

Ingest (Sean)

Placing into a trusted system. Recording data about this process. Setting certain preservation information such as our file location and checksums.

Storage (Sean)

Mechanical where this thing is going to live. How are we going to insure fixity. How is this thing backed up? How can we get it out.

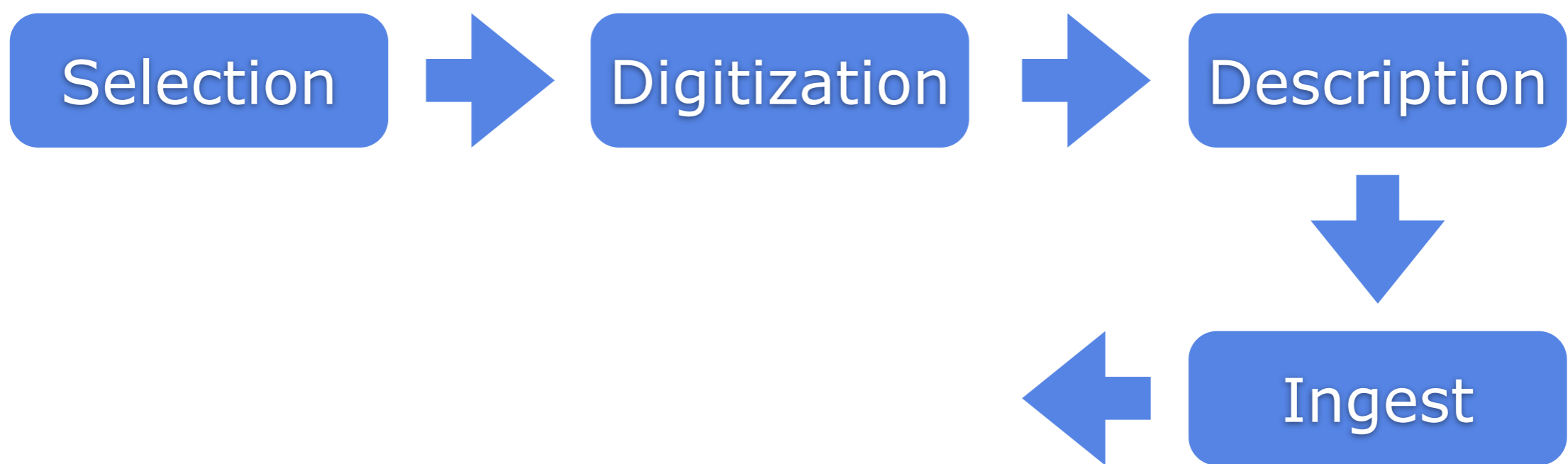
Distribution (Sean)

Youtube and website and all the other places it goes

Other (Sean)

This is very similar to our already existing production workflow.

Overview



Friday, June 14, 13

Selection (Sean)

Process of selecting what is going to be digitized. Preservation and collection development policies.

Digitization (Sean)

Process of format conversion. Moving from analog to digital. Though this also includes format conversions from digital to digital such as DVD to file based. Or DVCAM to file based.

Description (Sean)

Metadata. Access points. Summaries. Important places where we think that our users and interested searchers will need to find, identify, select and locate our digital assets.

Ingest (Sean)

Placing into a trusted system. Recording data about this process. Setting certain preservation information such as our file location and checksums.

Storage (Sean)

Mechanical where this thing is going to live. How are we going to insure fixity. How is this thing backed up? How can we get it out.

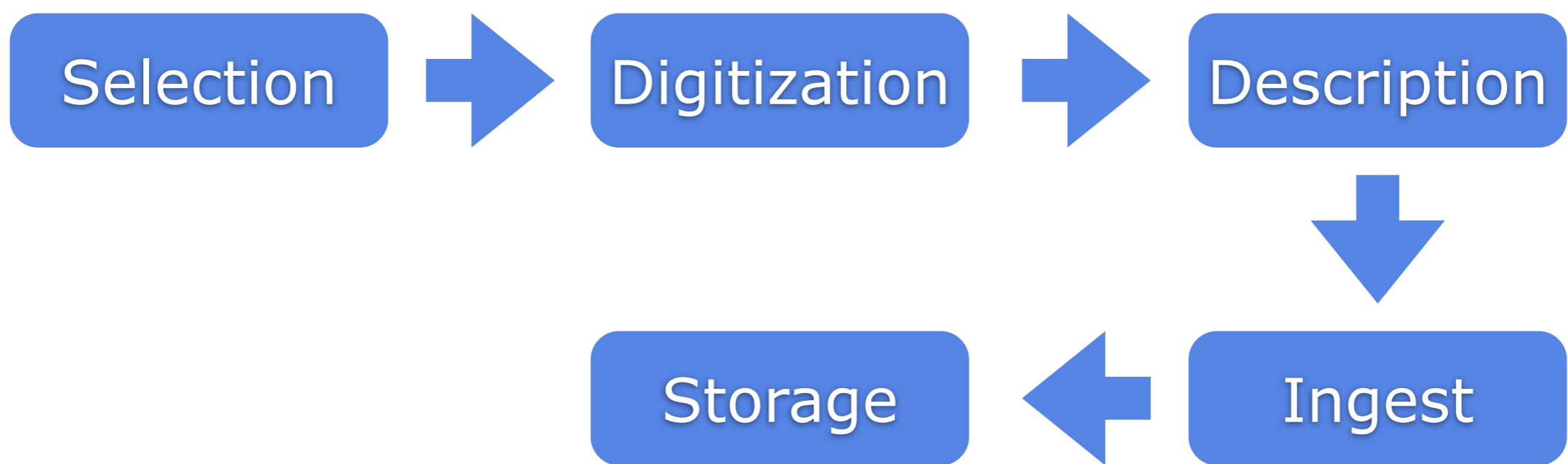
Distribution (Sean)

Youtube and website and all the other places it goes

Other (Sean)

This is very similar to our already existing production workflow.

Overview



Friday, June 14, 13

Selection (Sean)

Process of selecting what is going to be digitized. Preservation and collection development policies.

Digitization (Sean)

Process of format conversion. Moving from analog to digital. Though this also includes format conversions from digital to digital such as DVD to file based. Or DVCAM to file based.

Description (Sean)

Metadata. Access points. Summaries. Important places where we think that our users and interested searchers will need to find, identify, select and locate our digital assets.

Ingest (Sean)

Placing into a trusted system. Recording data about this process. Setting certain preservation information such as our file location and checksums.

Storage (Sean)

Mechanical where this thing is going to live. How are we going to insure fixity. How is this thing backed up? How can we get it out.

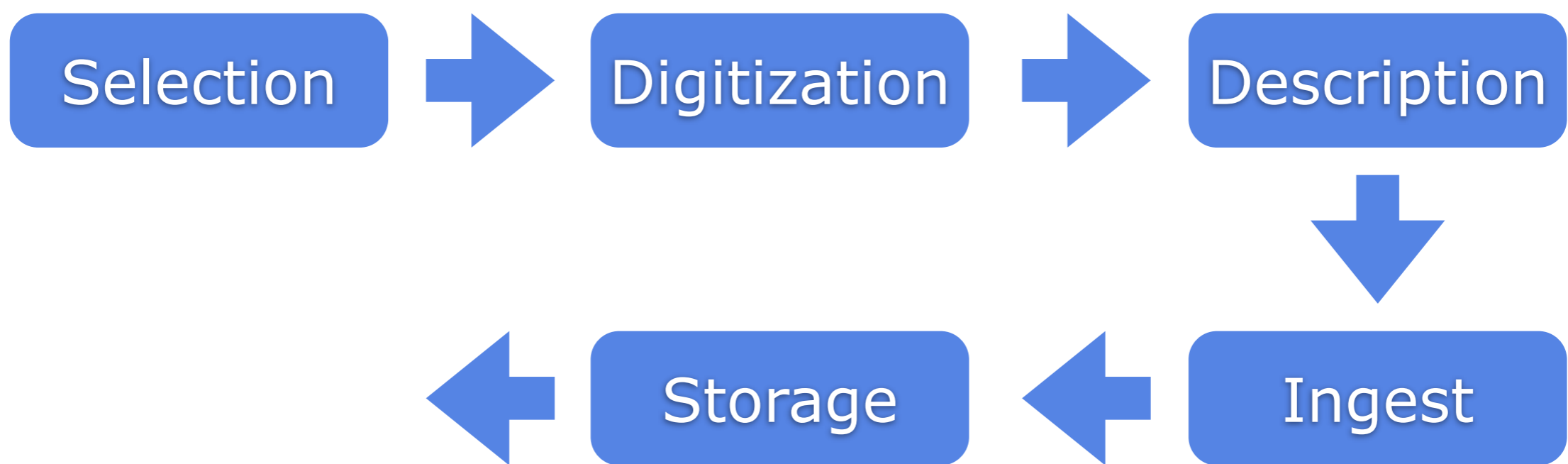
Distribution (Sean)

Youtube and website and all the other places it goes

Other (Sean)

This is very similar to our already existing production workflow.

Overview



Friday, June 14, 13

Selection (Sean)

Process of selecting what is going to be digitized. Preservation and collection development policies.

Digitization (Sean)

Process of format conversion. Moving from analog to digital. Though this also includes format conversions from digital to digital such as DVD to file based. Or DVCAM to file based.

Description (Sean)

Metadata. Access points. Summaries. Important places where we think that our users and interested searchers will need to find, identify, select and locate our digital assets.

Ingest (Sean)

Placing into a trusted system. Recording data about this process. Setting certain preservation information such as our file location and checksums.

Storage (Sean)

Mechanical where this thing is going to live. How are we going to insure fixity. How is this thing backed up? How can we get it out.

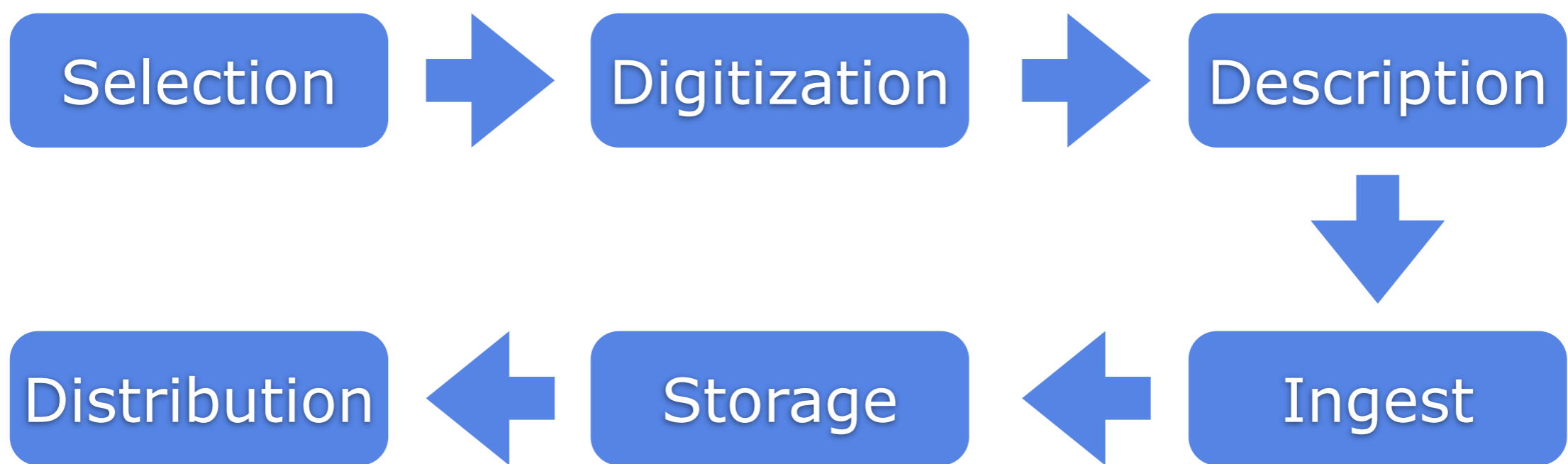
Distribution (Sean)

Youtube and website and all the other places it goes

Other (Sean)

This is very similar to our already existing production workflow.

Overview



Friday, June 14, 13

Selection (Sean)

Process of selecting what is going to be digitized. Preservation and collection development policies.

Digitization (Sean)

Process of format conversion. Moving from analog to digital. Though this also includes format conversions from digital to digital such as DVD to file based. Or DVCAM to file based.

Description (Sean)

Metadata. Access points. Summaries. Important places where we think that our users and interested searchers will need to find, identify, select and locate our digital assets.

Ingest (Sean)

Placing into a trusted system. Recording data about this process. Setting certain preservation information such as our file location and checksums.

Storage (Sean)

Mechanical where this thing is going to live. How are we going to insure fixity. How is this thing backed up? How can we get it out.

Distribution (Sean)

Youtube and website and all the other places it goes

Other (Sean)

This is very similar to our already existing production workflow.

Preservation & Collection Policy

Tier 1

Content

Recording type	Examples	Product preserved
High profile events	Events with notable speakers; strong faculty sponsorship	Final product only
Named lectures	Currie, Frey, & Bernstein series	Final product only
Oral histories	Alumni and Faculty	Final product and Raw footage
Faculty interviews	Demott discusses her amica brief	Final product and Raw footage
Produced materials	Sea turtles video	Final product only

Tier 2

Content

Recording type	Examples	Product preserved
All conferences	Journal, clinic, or center sponsored	Final product only
Student group organized events	Outlaw event; Federalist Society; Hardt Cup finals;	Final Product only
B-roll recordings	Outdoor shots; virtual tours	Raw footage
Short, one-off / promotional	Recordings requested by clinics, admissions, reunion weekend, holiday-themed video	Final product only

Setup

What did we do to get here. Extensive process. Refer back to Hollie.

Tier 1

Stuff we want to keep indefinitely. Multiple cameras. Multiple shots.

Tier 2

Stuff we want to keep indefinitely; but, don't need to do at such a high quality. Single camera. Single wide shot.

Preservation & Collection Policy

Tier 3

Content

Recording type	Examples	Product preserved
Student produced materials	JDLM International Lives in the Law	Final product only
Social student events	Cultural Extravaganza; Book tour invitee talk; student leaders using Duke Capture	Final product only

Tier 4

Content

Recording type	Examples	Product preserved
Class recording	Civil Procedure 110.01	None

DUKE LAW

Friday, June 14, 13

Tier 3

Student produced materials
Student social events

1 copy generally.

Tier 4

Not retained

Dark Storage

Separate issue from the tiers; if we didn't have releases for broadcast we feel that there is going to be some need for it we've created separate dark storage, which is managed similarly, but just not put in public accessible spaces.

Outgoing

Introduce Allye

Digitization & Workflow

DUKE LAW

Friday, June 14, 13

Allye gives an overview of her process.

Includes checking in tape into a spreadsheet.

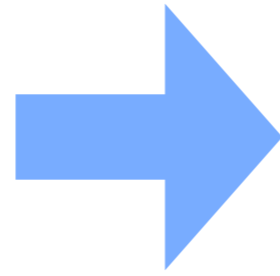
And then creating a file and naming it.

Putting into a holding area for subsequent description.

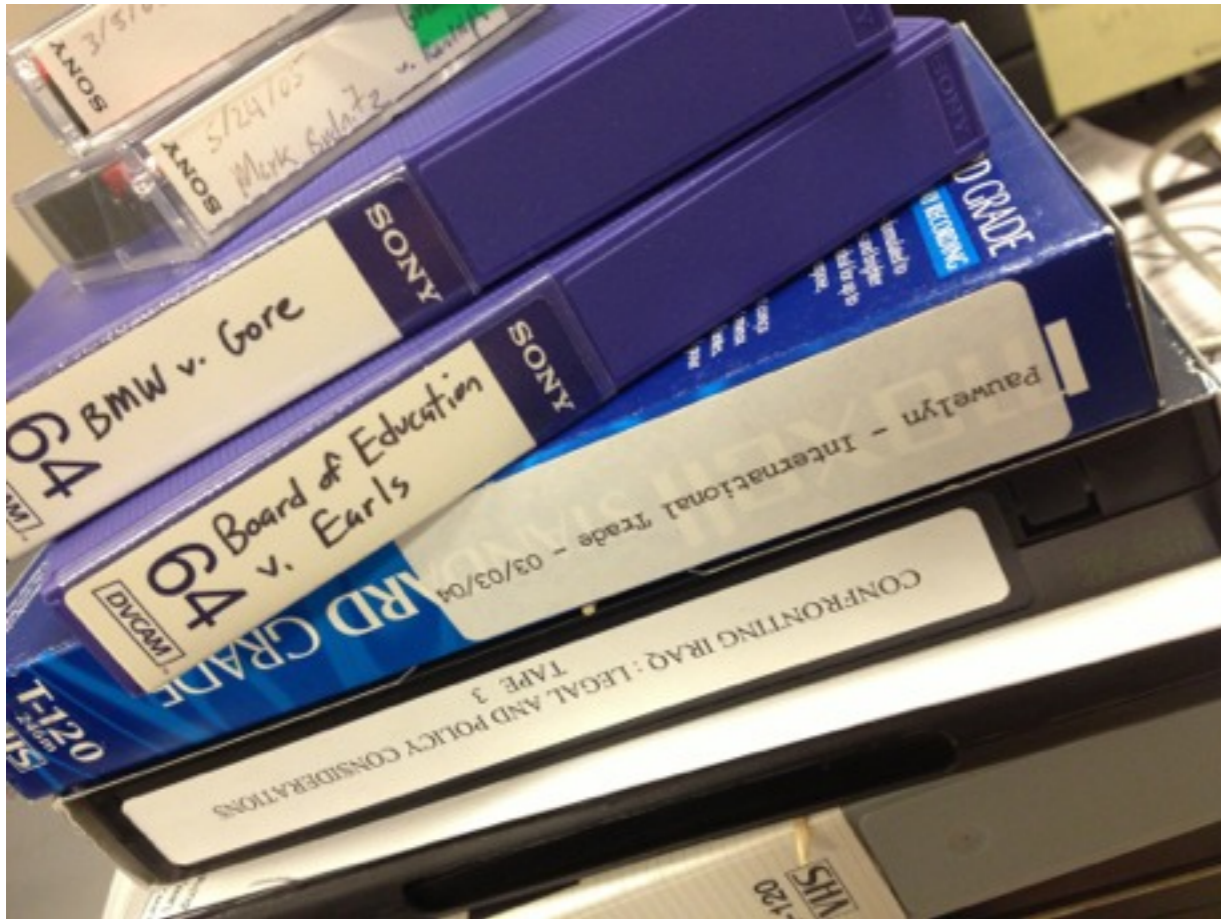
Digitization & Workflow

Digitization & Workflow

tape



DVD



DUKE LAW

Friday, June 14, 13

Ripping (Miguel)

Passing tape through DVD recorder.

VHS to DVD

DVCAM to DVD

MiniDV to DVD

Problems: Multiple stages of compression.

DVD is our intermediate format.

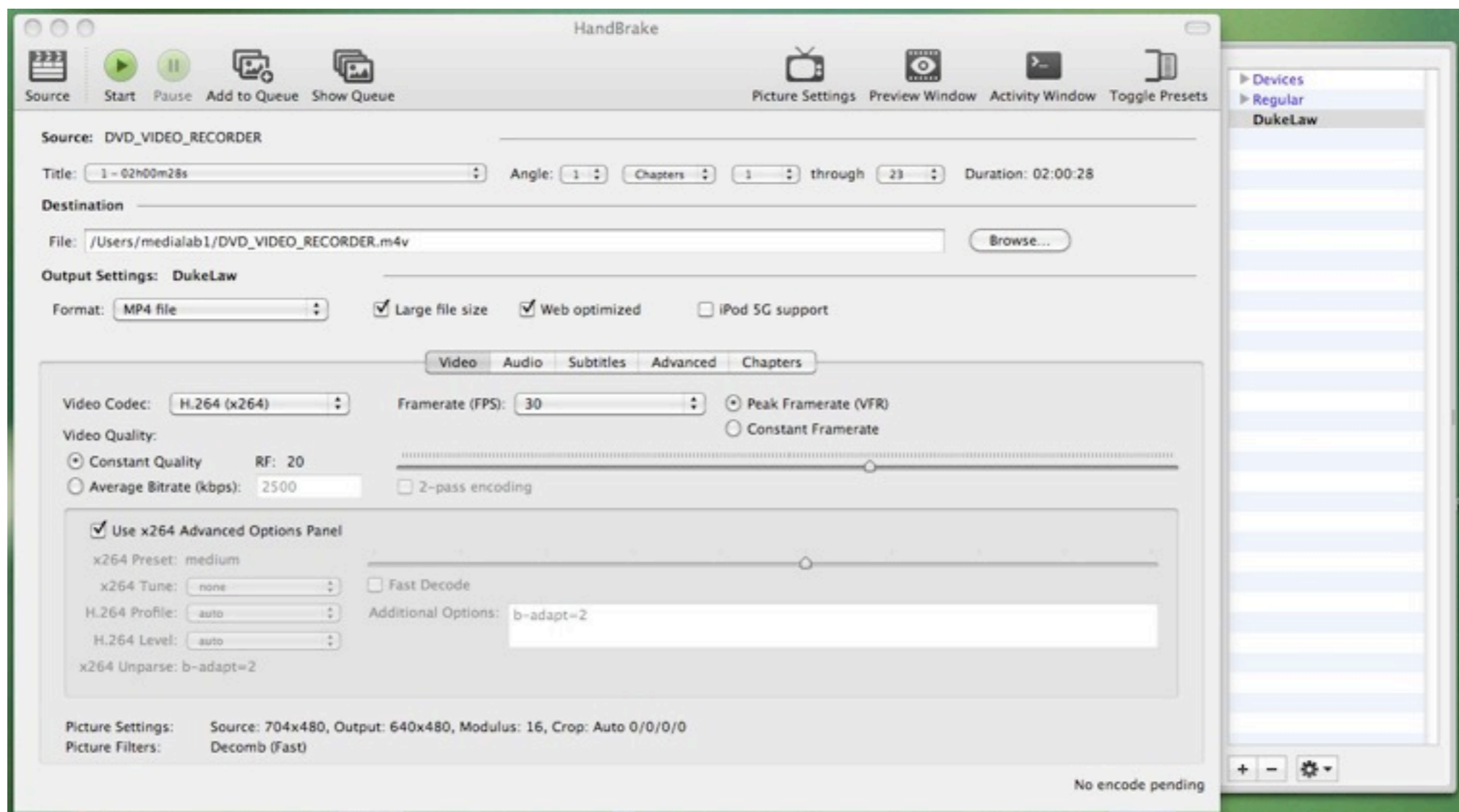
Talk about the variety of ways we recorded before: We had DVCAM and DVD Recording for a while at the same time.

Quality (Miguel)

High quality

Digitization & Workflow

HandBrake



DUKE LAW

Friday, June 14, 13

Digital Video Encoding (Miguel)

workflow - illustrated

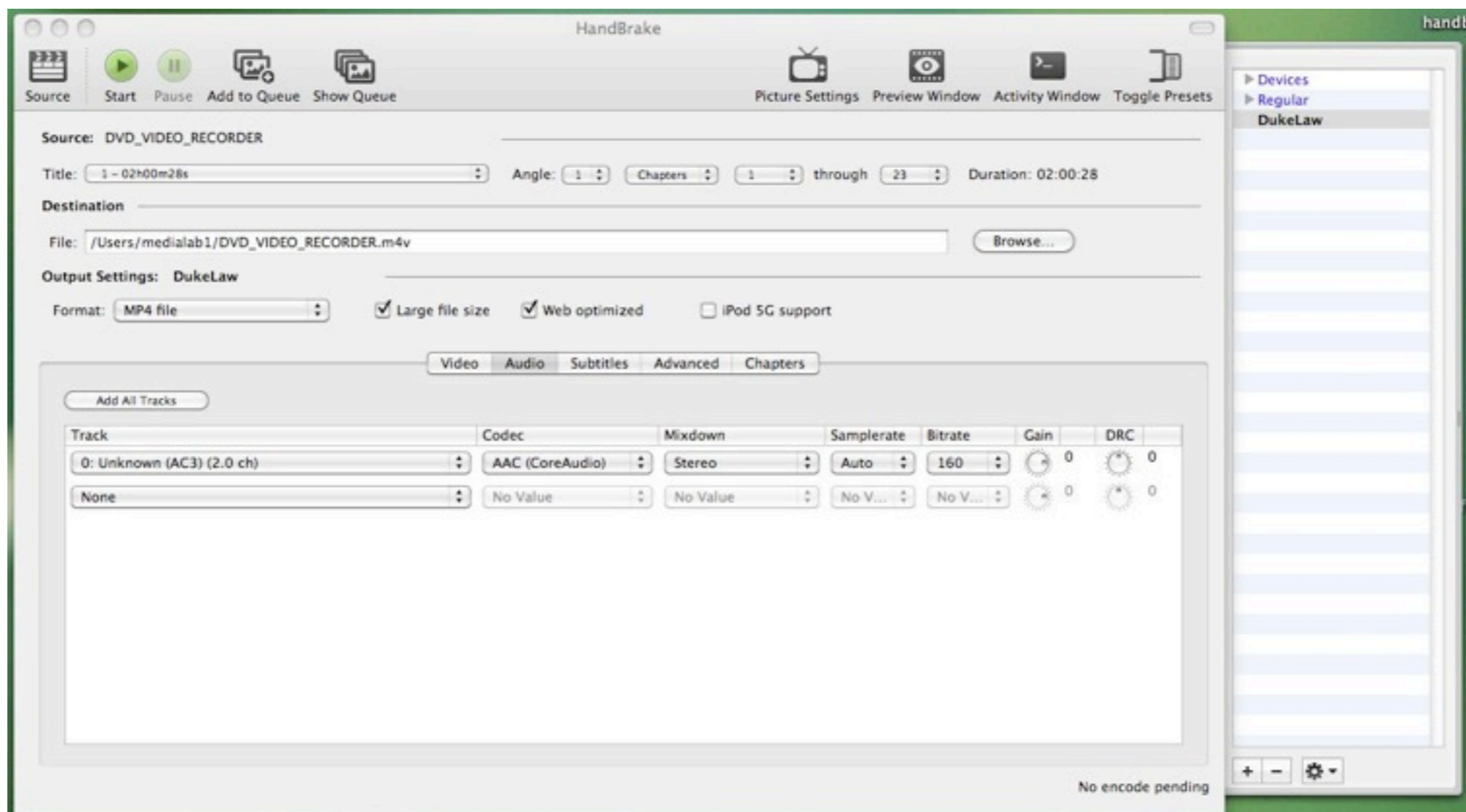
Apple TV 3 preset in HandBrake as the base, then created the Duke Law profile with our desired specs

Another level of compression: Still OK in Quality.

Feel as if our primary requirement: reuse by professionals in the future is more than possible.

Digitization & Workflow

HandBrake



DUKE LAW

Friday, June 14, 13

Digital Video Encoding (Miguel)

workflow - illustrated

Apple TV 3 preset in HandBrake as the base, then created the Duke Law profile with our desired specs

Another level of compression: Still OK in Quality.

Feel as if our primary requirement: reuse by professionals in the future is more than possible.

Digitization & Workflow

HandBrake



DUKE LAW

Friday, June 14, 13

Digital Video Encoding (Miguel)

workflow - illustrated

Apple TV 3 preset in HandBrake as the base, then created the Duke Law profile with our desired specs

Another level of compression: Still OK in Quality.

Feel as if our primary requirement: reuse by professionals in the future is more than possible.

Tools & Infrastructure

DUKE LAW

Friday, June 14, 13

Storage (Sean)

Our storage array is a large storage array that is hosting a number of services for the law school. Media is allocated 5TB of that total, while our archival masters use a fraction of the disk and at this point we have nearly 500GB of video in over 600 digitized files at this point. This constitutes files that have been checked in and are part of our system, as it is. And files that are currently in one of a number of workflows.

Networking (Sean)

This storage array is in the building which solves a lot of issues we had before. Our link to the campus infrastructure is limited to a single 1 gbit connection and could be easily saturated if we were moving a lot of files between campus hosted services. This was a consideration.

Part of our upgrading was a move to network storage throughout our video production. Making sure we had the best speed available to write to a network disk is extremely important. Having it in a network device lets us manage redundancy, saves costs, and lets us manage backing up better.

Backup

Tivoli. In separate building and infrastructure. The system backs up files nightly and preserves changed files for a specific retention period of up to 6 months in our policy. This lets us go and retrieve a file in the event something fails an integrity check, which hasn't happened in the four months we've been doing them.

Services (Sean)

Services being provided with by a separate virtualized SUSE-Linux server which runs a number of services that we have set up for our DAMS and preservation functions.

We try to keep things fairly lightweight. Concentrating on working directly on the filesystem rather than abstracting things through an application or other interface.

Services we have:

Ingest

Tools & Infrastructure

* Storage Array: 27 TB disk array

DUKE LAW

Friday, June 14, 13

Storage (Sean)

Our storage array is a large storage array that is hosting a number of services for the law school. Media is allocated 5TB of that total, while our archival masters use a fraction of the disk and at this point we have nearly 500GB of video in over 600 digitized files at this point. This constitutes files that have been checked in and are part of our system, as it is. And files that are currently in one of a number of workflows.

Networking (Sean)

This storage array is in the building which solves a lot of issues we had before. Our link to the campus infrastructure is limited to a single 1 gbit connection and could be easily saturated if we were moving a lot of files between campus hosted services. This was a consideration.

Part of our upgrading was a move to network storage throughout our video production. Making sure we had the best speed available to write to a network disk is extremely important. Having it in a network device lets us manage redundancy, saves costs, and lets us manage backing up better.

Backup

Tivoli. In separate building and infrastructure. The system backs up files nightly and preserves changed files for a specific retention period of up to 6 months in our policy. This lets us go and retrieve a file in the event something fails an integrity check, which hasn't happened in the four months we've been doing them.

Services (Sean)

Services being provided with by a separate virtualized SUSE-Linux server which runs a number of services that we have set up for our DAMS and preservation functions.

We try to keep things fairly lightweight. Concentrating on working directly on the filesystem rather than abstracting things through an application or other interface.

Services we have:

Ingest

Tools & Infrastructure

- * Storage Array: 27 TB disk array
- * Networking: 1 gigabit connectivity to array

DUKE LAW

Friday, June 14, 13

Storage (Sean)

Our storage array is a large storage array that is hosting a number of services for the law school. Media is allocated 5TB of that total, while our archival masters use a fraction of the disk and at this point we have nearly 500GB of video in over 600 digitized files at this point. This constitutes files that have been checked in and are part of our system, as it is. And files that are currently in one of a number of workflows.

Networking (Sean)

This storage array is in the building which solves a lot of issues we had before. Our link to the campus infrastructure is limited to a single 1 gbit connection and could be easily saturated if we were moving a lot of files between campus hosted services. This was a consideration.

Part of our upgrading was a move to network storage throughout our video production. Making sure we had the best speed available to write to a network disk is extremely important. Having it in a network device lets us manage redundancy, saves costs, and lets us manage backing up better.

Backup

Tivoli. In separate building and infrastructure. The system backs up files nightly and preserves changed files for a specific retention period of up to 6 months in our policy. This lets us go and retrieve a file in the event something fails an integrity check, which hasn't happened in the four months we've been doing them.

Services (Sean)

Services being provided with by a separate virtualized SUSE-Linux server which runs a number of services that we have set up for our DAMS and preservation functions.

We try to keep things fairly lightweight. Concentrating on working directly on the filesystem rather than abstracting things through an application or other interface.

Services we have:

Ingest

Tools & Infrastructure

- * Storage Array: 27 TB disk array
- * Networking: 1 gigabit connectivity to array
- * Backups: Tivoli Enterprise Backup to disk and tape

DUKE LAW

Friday, June 14, 13

Storage (Sean)

Our storage array is a large storage array that is hosting a number of services for the law school. Media is allocated 5TB of that total, while our archival masters use a fraction of the disk and at this point we have nearly 500GB of video in over 600 digitized files at this point. This constitutes files that have been checked in and are part of our system, as it is. And files that are currently in one of a number of workflows.

Networking (Sean)

This storage array is in the building which solves a lot of issues we had before. Our link to the campus infrastructure is limited to a single 1 gbit connection and could be easily saturated if we were moving a lot of files between campus hosted services. This was a consideration.

Part of our upgrading was a move to network storage throughout our video production. Making sure we had the best speed available to write to a network disk is extremely important. Having it in a network device lets us manage redundancy, saves costs, and lets us manage backing up better.

Backup

Tivoli. In separate building and infrastructure. The system backs up files nightly and preserves changed files for a specific retention period of up to 6 months in our policy. This lets us go and retrieve a file in the event something fails an integrity check, which hasn't happened in the four months we've been doing them.

Services (Sean)

Services being provided with by a separate virtualized SUSE-Linux server which runs a number of services that we have set up for our DAMS and preservation functions.

We try to keep things fairly lightweight. Concentrating on working directly on the filesystem rather than abstracting things through an application or other interface.

Services we have:

Ingest

Tools & Infrastructure

- * Storage Array: 27 TB disk array
- * Networking: 1 gigabit connectivity to array
- * Backups: Tivoli Enterprise Backup to disk and tape
- * Ingest: Google Docs & Python

DUKE LAW

Friday, June 14, 13

Storage (Sean)

Our storage array is a large storage array that is hosting a number of services for the law school. Media is allocated 5TB of that total, while our archival masters use a fraction of the disk and at this point we have nearly 500GB of video in over 600 digitized files at this point. This constitutes files that have been checked in and are part of our system, as it is. And files that are currently in one of a number of workflows.

Networking (Sean)

This storage array is in the building which solves a lot of issues we had before. Our link to the campus infrastructure is limited to a single 1 gbit connection and could be easily saturated if we were moving a lot of files between campus hosted services. This was a consideration.

Part of our upgrading was a move to network storage throughout our video production. Making sure we had the best speed available to write to a network disk is extremely important. Having it in a network device lets us manage redundancy, saves costs, and lets us manage backing up better.

Backup

Tivoli. In separate building and infrastructure. The system backs up files nightly and preserves changed files for a specific retention period of up to 6 months in our policy. This lets us go and retrieve a file in the event something fails an integrity check, which hasn't happened in the four months we've been doing them.

Services (Sean)

Services being provided with by a separate virtualized SUSE-Linux server which runs a number of services that we have set up for our DAMS and preservation functions.

We try to keep things fairly lightweight. Concentrating on working directly on the filesystem rather than abstracting things through an application or other interface.

Services we have:

Ingest

Tools & Infrastructure

- * Storage Array: 27 TB disk array
- * Networking: 1 gigabit connectivity to array
- * Backups: Tivoli Enterprise Backup to disk and tape
- * Ingest: Google Docs & Python
- * Description: Worldcat & Aleph ILS

DUKE LAW

Friday, June 14, 13

Storage (Sean)

Our storage array is a large storage array that is hosting a number of services for the law school. Media is allocated 5TB of that total, while our archival masters use a fraction of the disk and at this point we have nearly 500GB of video in over 600 digitized files at this point. This constitutes files that have been checked in and are part of our system, as it is. And files that are currently in one of a number of workflows.

Networking (Sean)

This storage array is in the building which solves a lot of issues we had before. Our link to the campus infrastructure is limited to a single 1 gbit connection and could be easily saturated if we were moving a lot of files between campus hosted services. This was a consideration.

Part of our upgrading was a move to network storage throughout our video production. Making sure we had the best speed available to write to a network disk is extremely important. Having it in a network device lets us manage redundancy, saves costs, and lets us manage backing up better.

Backup

Tivoli. In separate building and infrastructure. The system backs up files nightly and preserves changed files for a specific retention period of up to 6 months in our policy. This lets us go and retrieve a file in the event something fails an integrity check, which hasn't happened in the four months we've been doing them.

Services (Sean)

Services being provided with by a separate virtualized SUSE-Linux server which runs a number of services that we have set up for our DAMS and preservation functions.

We try to keep things fairly lightweight. Concentrating on working directly on the filesystem rather than abstracting things through an application or other interface.

Services we have:

Ingest

Tools & Infrastructure

- * Storage Array: 27 TB disk array
- * Networking: 1 gigabit connectivity to array
- * Backups: Tivoli Enterprise Backup to disk and tape
- * Ingest: Google Docs & Python
- * Description: Worldcat & Aleph ILS
- * Distribution: Google Docs, YouTube & Drupal

DUKE LAW

Friday, June 14, 13

Storage (Sean)

Our storage array is a large storage array that is hosting a number of services for the law school. Media is allocated 5TB of that total, while our archival masters use a fraction of the disk and at this point we have nearly 500GB of video in over 600 digitized files at this point. This constitutes files that have been checked in and are part of our system, as it is. And files that are currently in one of a number of workflows.

Networking (Sean)

This storage array is in the building which solves a lot of issues we had before. Our link to the campus infrastructure is limited to a single 1 gbit connection and could be easily saturated if we were moving a lot of files between campus hosted services. This was a consideration.

Part of our upgrading was a move to network storage throughout our video production. Making sure we had the best speed available to write to a network disk is extremely important. Having it in a network device lets us manage redundancy, saves costs, and lets us manage backing up better.

Backup

Tivoli. In separate building and infrastructure. The system backs up files nightly and preserves changed files for a specific retention period of up to 6 months in our policy. This lets us go and retrieve a file in the event something fails an integrity check, which hasn't happened in the four months we've been doing them.

Services (Sean)

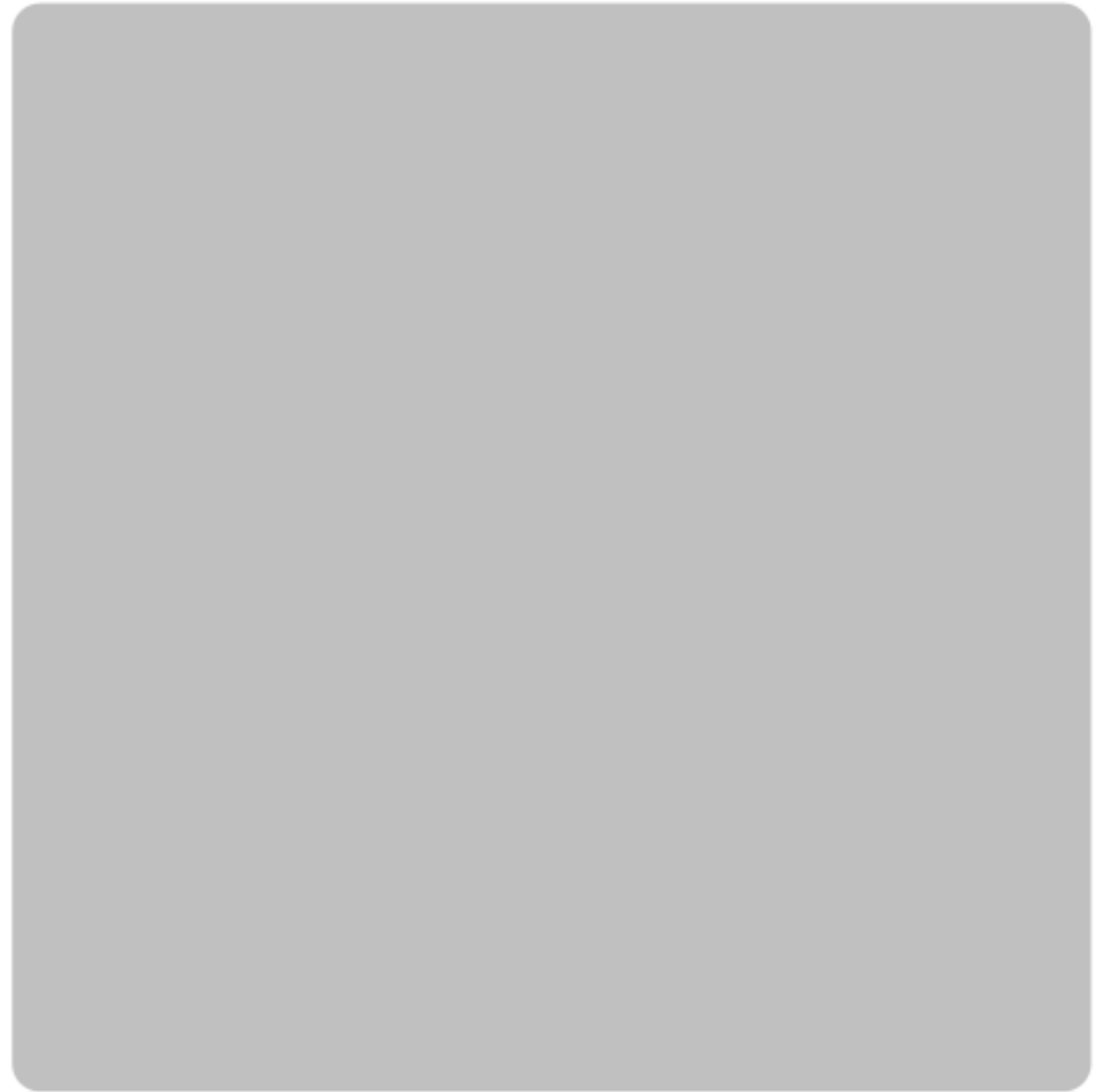
Services being provided with by a separate virtualized SUSE-Linux server which runs a number of services that we have set up for our DAMS and preservation functions.

We try to keep things fairly lightweight. Concentrating on working directly on the filesystem rather than abstracting things through an application or other interface.

Services we have:

Ingest

Preservation & Description



DUKE LAW

Friday, June 14, 13

Google Documents (Sean)

Managing workflow. Keeping track of information from a whole range of people who have been working on the project.

The one thing that was invaluable for using in a project that involved a number of people working on the same data at the same time was Google Documents

This is the glue that kept our project together.

We had separate worksheets for each format as we digitized. We have a master spreadsheet with file level information for each file we want to check in and ingest.

We included information like release status, source origin, filenames, titles, identifiers, dates in the workflow, preservation actions etc.

Services run on this spreadsheet using the Google Data API.

Preservation & Description

* How?

DUKE LAW

Friday, June 14, 13

Google Documents (Sean)

Managing workflow. Keeping track of information from a whole range of people who have been working on the project.

The one thing that was invaluable for using in a project that involved a number of people working on the same data at the same time was Google Documents

This is the glue that kept our project together.

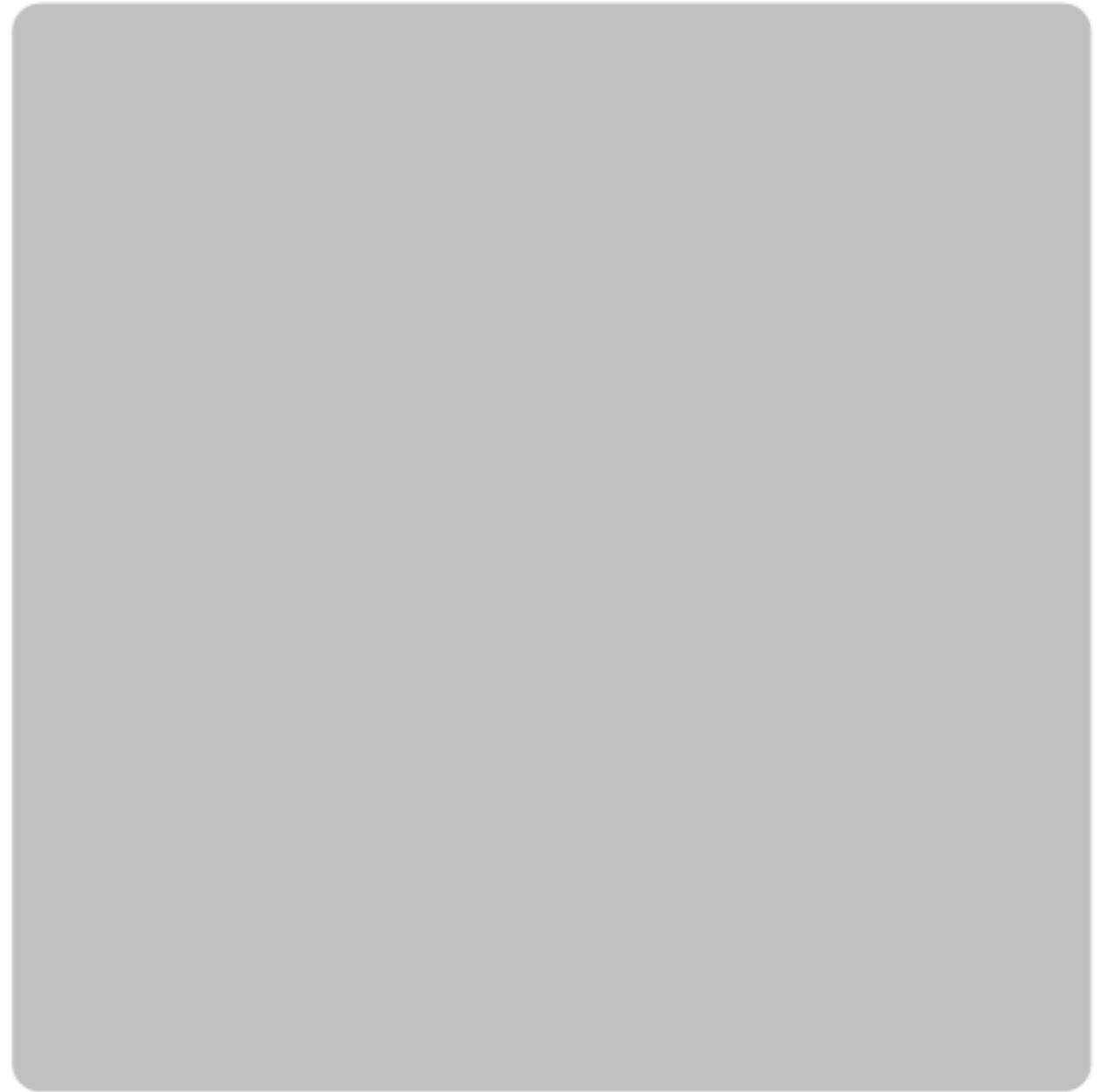
We had separate worksheets for each format as we digitized. We have a master spreadsheet with file level information for each file we want to check in and ingest.

We included information like release status, source origin, filenames, titles, identifiers, dates in the workflow, preservation actions etc.

Services run on this spreadsheet using the Google Data API.

Preservation & Description

* How? Spreadsheets!



DUKE LAW

Friday, June 14, 13

Google Documents (Sean)

Managing workflow. Keeping track of information from a whole range of people who have been working on the project.

The one thing that was invaluable for using in a project that involved a number of people working on the same data at the same time was Google Documents

This is the glue that kept our project together.

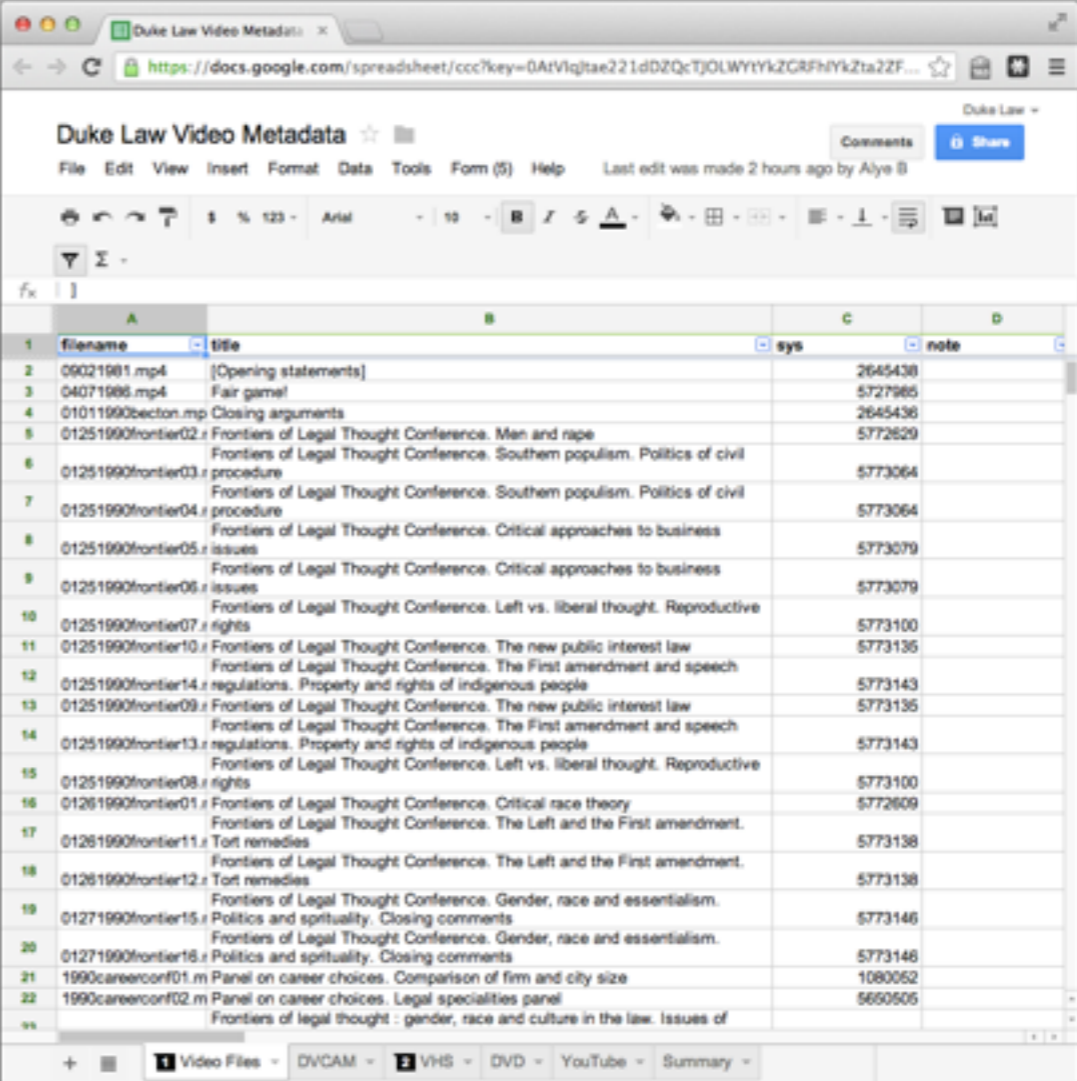
We had separate worksheets for each format as we digitized. We have a master spreadsheet with file level information for each file we want to check in and ingest.

We included information like release status, source origin, filenames, titles, identifiers, dates in the workflow, preservation actions etc.

Services run on this spreadsheet using the Google Data API.

Preservation & Description

✿ How? Spreadsheets!



filename	title	size	note
00021981.mp4	[Opening statements]		2645438
04071986.mp4	Fair game!		5727985
01011990becton.mp	Closing arguments		2645438
01251990frontier02	Frontiers of Legal Thought Conference: Man and rape		5772629
01251990frontier03	Frontiers of Legal Thought Conference: Southern populism. Politics of civil procedure		5773064
01251990frontier04	Frontiers of Legal Thought Conference: Southern populism. Politics of civil procedure		5773064
01251990frontier05	Frontiers of Legal Thought Conference: Critical approaches to business issues		5773079
01251990frontier06	Frontiers of Legal Thought Conference: Critical approaches to business issues		5773079
01251990frontier07	Frontiers of Legal Thought Conference: Left vs. liberal thought. Reproductive rights		5773100
01251990frontier10	Frontiers of Legal Thought Conference: The new public interest law		5773135
01251990frontier09	Frontiers of Legal Thought Conference: The First amendment and speech regulations. Property and rights of indigenous people		5773143
01251990frontier14	Frontiers of Legal Thought Conference: The First amendment and speech regulations. Property and rights of indigenous people		5773143
01251990frontier13	Frontiers of Legal Thought Conference: The First amendment and speech regulations. Property and rights of indigenous people		5773143
01251990frontier08	Frontiers of Legal Thought Conference: Left vs. liberal thought. Reproductive rights		5773100
01261990frontier01	Frontiers of Legal Thought Conference: Critical race theory		5772609
01261990frontier11	Frontiers of Legal Thought Conference: The Left and the First amendment. Tort remedies		5773138
01261990frontier12	Frontiers of Legal Thought Conference: The Left and the First amendment. Tort remedies		5773138
01271990frontier15	Frontiers of Legal Thought Conference: Gender, race and essentialism. Politics and spirituality. Closing comments		5773146
01271990frontier16	Frontiers of Legal Thought Conference: Gender, race and essentialism. Politics and spirituality. Closing comments		5773146
1990careerconf01.m	Panel on career choices. Comparison of firm and city size		1080052
1990careerconf02.m	Panel on career choices. Legal specialties panel		5650505
1990	Frontiers of legal thought : gender, race and culture in the law. Issues of		

DUKE LAW

Friday, June 14, 13

Google Documents (Sean)

Managing workflow. Keeping track of information from a whole range of people who have been working on the project.

The one thing that was invaluable for using in a project that involved a number of people working on the same data at the same time was Google Documents

This is the glue that kept our project together.

We had separate worksheets for each format as we digitized. We have a master spreadsheet with file level information for each file we want to check in and ingest.

We included information like release status, source origin, filenames, titles, identifiers, dates in the workflow, preservation actions etc.

Services run on this spreadsheet using the Google Data API.

Preservation

- * Filesystem
- * Python
- * Identifiers

```
-P- 1 root users 1546322873 2013-01-23 08:29 01222013.mp4
-P- 1 root users 816796296 2013-02-04 17:38 02042013.mp4
-P- 1 root users 1602675297 2013-02-06 01:33 02062013.mp4
-P- 1 root users 1671170438 2013-02-13 21:39 02132013.mp4
-P- 1 root users 1685163669 2013-02-14 21:14 02142013.mp4
-P- 1 root users 2029311273 2013-02-18 14:16 0215201301a.mp4
-P- 1 root users 1389411516 2013-02-15 17:35 0215201301b.mp4
-P- 1 root users 1000818755 2013-02-15 17:05 0215201301c.mp4
-P- 1 root users 1553420359 2013-02-18 12:48 0215201301d.mp4
-P- 1 root users 1835428568 2013-02-21 17:48 02212013currie.mp4
-P- 1 root users 538134857 2013-02-28 16:04 02282013cicl.mp4
-P- 1 root users 1069000945 2013-02-28 18:29 02282013patentlaw.mp4
-P- 1 root users 1483249371 2013-03-01 17:58 03012013LNSa.mp4
-P- 1 root users 669720181 2013-03-01 11:59 03012013LNSb.mp4
-P- 1 root users 1002677862 2013-03-01 14:42 03012013LNSc.mp4
-P- 1 root users 457971585 2013-03-01 14:51 03012013LNSd.mp4
-P- 1 root users 508513242 2013-03-01 16:41 03012013LNSe.mp4
-P- 1 root users 520736418 2013-03-02 18:13 03022013LNSaa.mp4
-P- 1 root users 902643521 2013-03-02 11:55 03022013LNSab.mp4
-P- 1 root users 674187511 2013-03-04 09:41 03022013LNSacc.mp4
-P- 1 root users 521635664 2013-03-04 13:45 03042013Israel1.mp4
-P- 1 root users 202635967 2013-03-05 13:33 03052013.mp4
-P- 1 root users 537010608 2013-03-06 13:55 03062013frey.mp4
-P- 1 root users 577913419 2013-03-07 14:57 03072013.mp4
-P- 1 root users 254407099 2013-03-20 18:04 03202013.mp4
-P- 1 root users 795548973 2013-03-29 09:56 03292013dona.mp4
-P- 1 root users 619687911 2013-03-28 15:06 03282013.mp4
-P- 1 root users 361697648 2013-04-03 15:15 04032013.mp4
-P- 1 root users 627501977 2013-04-04 14:01 04042013.mp4
-P- 1 root users 468525184 2013-04-08 18:42 04082013hardcup.mp4
-P- 1 root users 672475423 2013-04-08 14:01 04082013.mp4
-P- 1 root users 467508242 2013-04-11 09:31 04092013.mp4
-P- 1 root users 215857346 2013-04-11 13:44 04112013charter.mp4
-P- 1 root users 392685844 2013-04-11 14:04 04112013.mp4
-P- 1 root users 293641050 2013-05-13 18:31 05112013.mp4
-P- 1 root users 97291557 2013-01-07 14:15 0JagenevaInstitute.mp4
-P- 1 root users 190313259 2013-04-22 18:53 lamontearstrongelvinfinal.m
p4
-P- 1 root users 118277372 2013-02-18 13:58 lamontearstrong_sequence_2.
mp4
-P- 1 root users 163173107 2013-02-27 11:26 lmalbertogutierrezpier.mp4
-P- 1 root users 151218856 2013-06-18 11:53 llexperience2013.mp4
-P- 1 root users 107675587 2013-02-20 10:42 llnkatharinariel.mp4
-P- 1 root users 119125766 2013-03-01 09:46 llnukaniwangai.mp4
-P- 1 root users 228382756 2013-02-01 11:55 llnsjerbek2013.mp4
-P- 1 root users 163433997 2013-03-14 18:49 Mlevyfacscholarship.mp4
-P- 1 root users 203895282 2013-02-18 15:59 skjerbeklml3.mp4
-P- 1 root users 301500 2013-05-13 12:35 Thumbs.db
[sheng015-Srv19 2013]s
```

Friday, June 14, 13

Preservation

- * unique identifier on filenames
- * filesystem
- * MD5s on files periodically
- * Auditing of the files
- * Backing up

Description

DUKE LAW

Friday, June 14, 13

Data Sources (Sean)

So a principal concern of our preservation plans included thinking about where the data about our videos existed already and in turn aligning it with what we think we could support in the future.

Catalog Data (Sean)

Until the point we decided to dive head long into this project the Library had already been devoting resources to describing the content that Media services had been creating for the last ten years. In a variety of ways including:

- Adding physical manifestations to library collections
- Cataloging streaming manifestations of the video.

This existing body of data for more than five hundred individual videos already and

Duplicative of these efforts was a long standing workflow which provisioned the access of the videos through the library website. This included working within the Law School website's content management system to include information about videos on a specific video archive page.

However this was known to be redundant and and a non ideal way of doing this sort of work. And did little to ensure the long term preservation and access to this content.

Calendar & Event Descriptions

We had information from scheduling in our various calendar systems both on the web and in our email systems. This has let us reconstruct the existence of videos along with information about who appears and when something occurred.

In our current workflow for current productions however this is becoming more and more important since we need this information in order to improve accessibility and identification later on in the process.

Physical Objects

Tapes and discs had a surprising amount of information about the events they corresponded to. But often they had misleading information. But it was helpful when describing and identifying what was being shown in a video.

Description

* Catalog Records

DUKE LAW

Friday, June 14, 13

Data Sources (Sean)

So a principal concern of our preservation plans included thinking about where the data about our videos existed already and in turn aligning it with what we think we could support in the future.

Catalog Data (Sean)

Until the point we decided to dive head long into this project the Library had already been devoting resources to describing the content that Media services had been creating for the last ten years. In a variety of ways including:

- Adding physical manifestations to library collections
- Cataloging streaming manifestations of the video.

This existing body of data for more than five hundred individual videos already and

Duplicative of these efforts was a long standing workflow which provisioned the access of the videos through the library website. This included working within the Law School website's content management system to include information about videos on a specific video archive page.

However this was known to be redundant and and a non ideal way of doing this sort of work. And did little to ensure the long term preservation and access to this content.

Calendar & Event Descriptions

We had information from scheduling in our various calendar systems both on the web and in our email systems. This has let us reconstruct the existence of videos along with information about who appears and when something occurred.

In our current workflow for current productions however this is becoming more and more important since we need this information in order to improve accessibility and identification later on in the process.

Physical Objects

Tapes and discs had a surprising amount of information about the events they corresponded to. But often they had misleading information. But it was helpful when describing and identifying what was being shown in a video.

Description

* Catalog Records

* Calendar & Event Descriptions

DUKE LAW

Friday, June 14, 13

Data Sources (Sean)

So a principal concern of our preservation plans included thinking about where the data about our videos existed already and in turn aligning it with what we think we could support in the future.

Catalog Data (Sean)

Until the point we decided to dive head long into this project the Library had already been devoting resources to describing the content that Media services had been creating for the last ten years. In a variety of ways including:

- Adding physical manifestations to library collections
- Cataloging streaming manifestations of the video.

This existing body of data for more than five hundred individual videos already and

Duplicative of these efforts was a long standing workflow which provisioned the access of the videos through the library website. This included working within the Law School website's content management system to include information about videos on a specific video archive page.

However this was known to be redundant and and a non ideal way of doing this sort of work. And did little to ensure the long term preservation and access to this content.

Calendar & Event Descriptions

We had information from scheduling in our various calendar systems both on the web and in our email systems. This has let us reconstruct the existence of videos along with information about who appears and when something occurred.

In our current workflow for current productions however this is becoming more and more important since we need this information in order to improve accessibility and identification later on in the process.

Physical Objects

Tapes and discs had a surprising amount of information about the events they corresponded to. But often they had misleading information. But it was helpful when describing and identifying what was being shown in a video.

Description

- * Catalog Records
- * Calendar & Event Descriptions
- * Physical Objects

DUKE LAW

Friday, June 14, 13

Data Sources (Sean)

So a principal concern of our preservation plans included thinking about where the data about our videos existed already and in turn aligning it with what we think we could support in the future.

Catalog Data (Sean)

Until the point we decided to dive head long into this project the Library had already been devoting resources to describing the content that Media services had been creating for the last ten years. In a variety of ways including:

- Adding physical manifestations to library collections
- Cataloging streaming manifestations of the video.

This existing body of data for more than five hundred individual videos already and

Duplicative of these efforts was a long standing workflow which provisioned the access of the videos through the library website. This included working within the Law School website's content management system to include information about videos on a specific video archive page.

However this was known to be redundant and and a non ideal way of doing this sort of work. And did little to ensure the long term preservation and access to this content.

Calendar & Event Descriptions

We had information from scheduling in our various calendar systems both on the web and in our email systems. This has let us reconstruct the existence of videos along with information about who appears and when something occurred.

In our current workflow for current productions however this is becoming more and more important since we need this information in order to improve accessibility and identification later on in the process.

Physical Objects

Tapes and discs had a surprising amount of information about the events they corresponded to. But often they had misleading information. But it was helpful when describing and identifying what was being shown in a video.

Description

- * Catalog Records
- * Calendar & Event Descriptions
- * Physical Objects
- * Website

DUKE LAW

Friday, June 14, 13

Data Sources (Sean)

So a principal concern of our preservation plans included thinking about where the data about our videos existed already and in turn aligning it with what we think we could support in the future.

Catalog Data (Sean)

Until the point we decided to dive head long into this project the Library had already been devoting resources to describing the content that Media services had been creating for the last ten years. In a variety of ways including:

- Adding physical manifestations to library collections
- Cataloging streaming manifestations of the video.

This existing body of data for more than five hundred individual videos already and

Duplicative of these efforts was a long standing workflow which provisioned the access of the videos through the library website. This included working within the Law School website's content management system to include information about videos on a specific video archive page.

However this was known to be redundant and and a non ideal way of doing this sort of work. And did little to ensure the long term preservation and access to this content.

Calendar & Event Descriptions

We had information from scheduling in our various calendar systems both on the web and in our email systems. This has let us reconstruct the existence of videos along with information about who appears and when something occurred.

In our current workflow for current productions however this is becoming more and more important since we need this information in order to improve accessibility and identification later on in the process.

Physical Objects

Tapes and discs had a surprising amount of information about the events they corresponded to. But often they had misleading information. But it was helpful when describing and identifying what was being shown in a video.

Description

- * Catalog Records
- * Calendar & Event Descriptions
- * Physical Objects
- * Website
- * Releases

DUKE LAW

Friday, June 14, 13

Data Sources (Sean)

So a principal concern of our preservation plans included thinking about where the data about our videos existed already and in turn aligning it with what we think we could support in the future.

Catalog Data (Sean)

Until the point we decided to dive head long into this project the Library had already been devoting resources to describing the content that Media services had been creating for the last ten years. In a variety of ways including:

- Adding physical manifestations to library collections
- Cataloging streaming manifestations of the video.

This existing body of data for more than five hundred individual videos already and

Duplicative of these efforts was a long standing workflow which provisioned the access of the videos through the library website. This included working within the Law School website's content management system to include information about videos on a specific video archive page.

However this was known to be redundant and and a non ideal way of doing this sort of work. And did little to ensure the long term preservation and access to this content.

Calendar & Event Descriptions

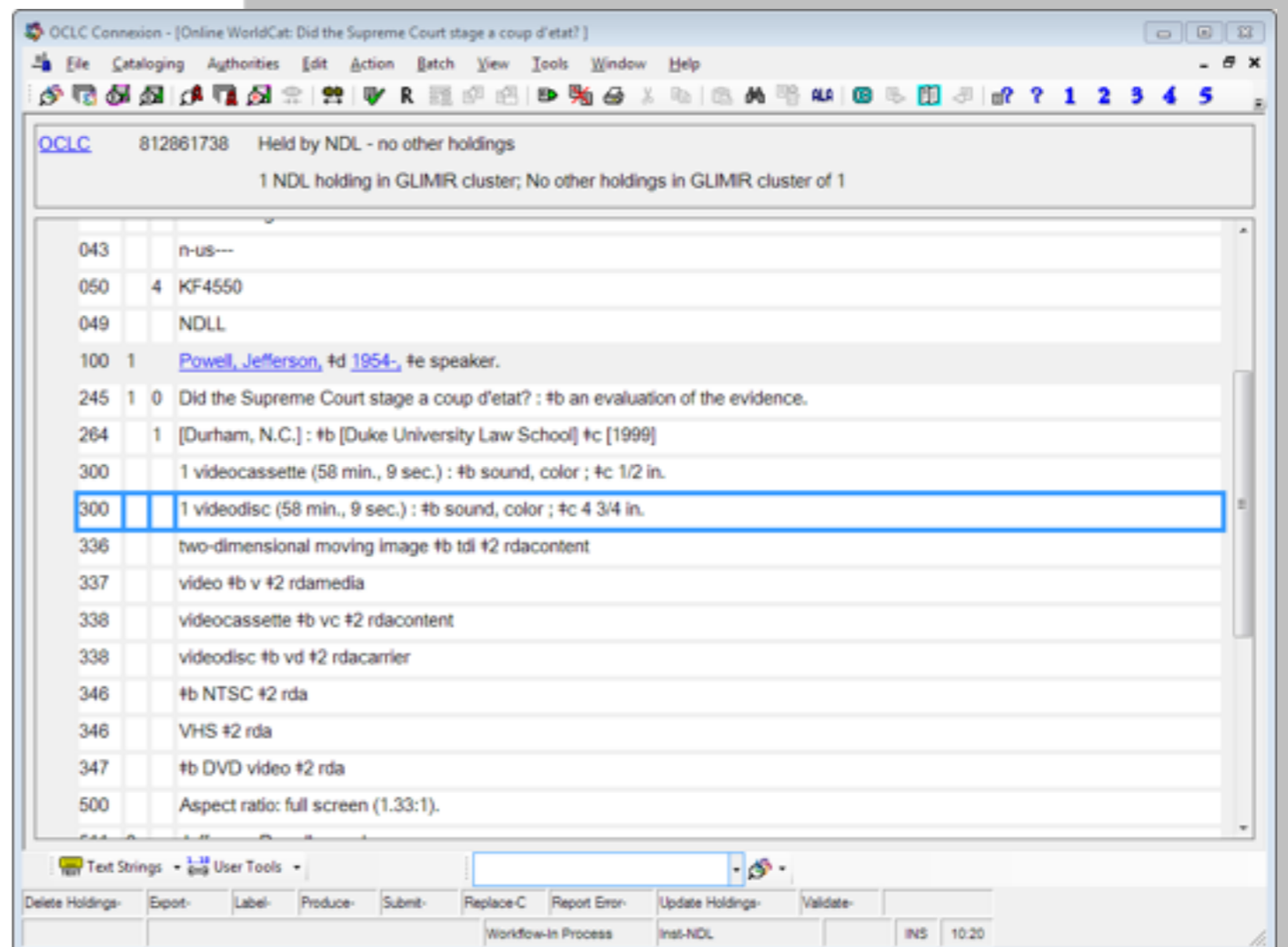
We had information from scheduling in our various calendar systems both on the web and in our email systems. This has let us reconstruct the existence of videos along with information about who appears and when something occurred.

In our current workflow for current productions however this is becoming more and more important since we need this information in order to improve accessibility and identification later on in the process.

Physical Objects

Tapes and discs had a surprising amount of information about the events they corresponded to. But often they had misleading information. But it was helpful when describing and identifying what was being shown in a video.

Description



DUKE LAW

Friday, June 14, 13

MARC & RDA (Sean)

We used existing staff and training in order to describe and provide access and metadata about our videos.

But more importantly there already existed a layer of interaction which we could program services against each of these data sources. For our integrated library system we could identify information about source and inventory and location.

We had other considerations which made the leveraging of the existing catalog and ILS tools important. These included:

- * controlled vocabularies for subject access
- * Authority control for people appearing and contributing to videos.
- * unique identifiers
- * Permanence as data sources (we know we aren't moving away from these tools since the library still needs them).

This is expensive and time consuming but is important to us since this is unique content.

Worldcat and Connexion

Vocabulary control (National Authority File, LCSH).

Unique identifiers for our works. Exposing in other places.

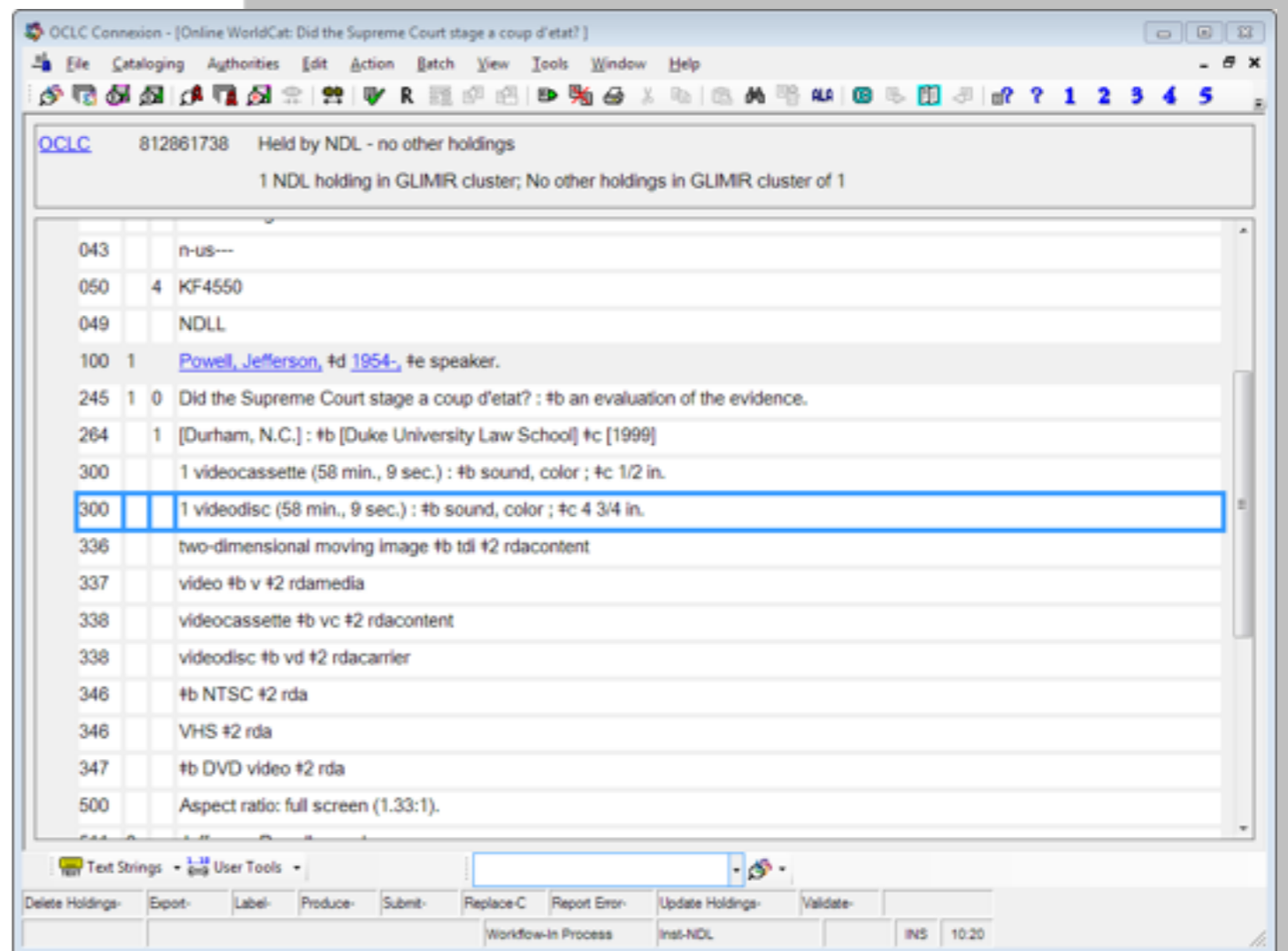
ALEPH

Information about our instantiations.

The ILS is really good for tracking inventory and putting things on the shelf which is important when we are managing a huge number of items. This works better than having piles of tapes sitting in a closet.

Description

* MARC & RDA



DUKE LAW

Friday, June 14, 13

MARC & RDA (Sean)

We used existing staff and training in order to describe and provide access and metadata about our videos.

But more importantly there already existed a layer of interaction which we could program services against each of these data sources. For our integrated library system we could identify information about source and inventory and location.

We had other considerations which made the leveraging of the existing catalog and ILS tools important. These included:

- * controlled vocabularies for subject access
- * Authority control for people appearing and contributing to videos.
- * unique identifiers
- * Permanence as data sources (we know we aren't moving away from these tools since the library still needs them).

This is expensive and time consuming but is important to us since this is unique content.

Worldcat and Connexion

Vocabulary control (National Authority File, LCSH).

Unique identifiers for our works. Exposing in other places.

ALEPH

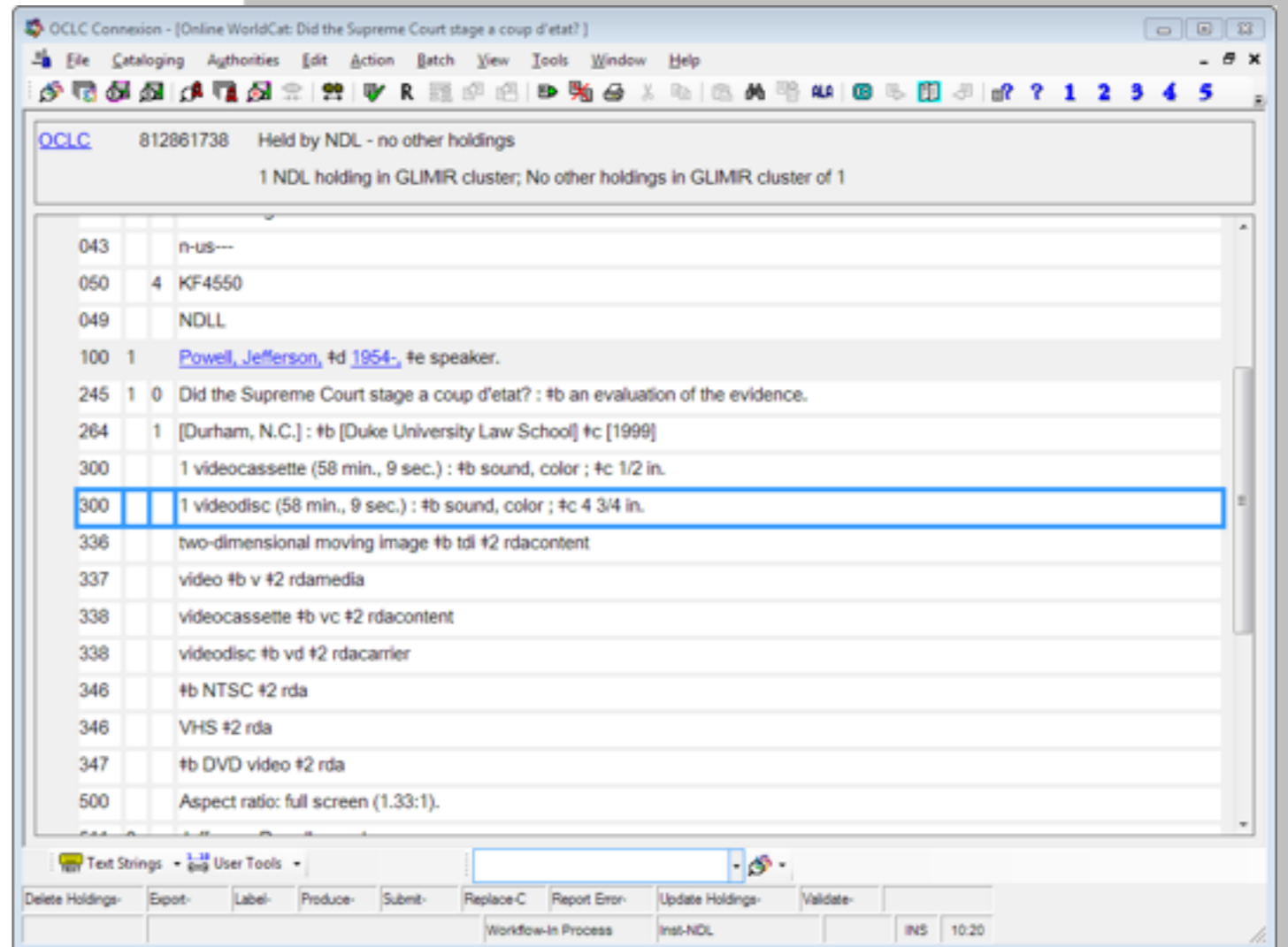
Information about our instantiations.

The ILS is really good for tracking inventory and putting things on the shelf which is important when we are managing a huge number of items. This works better than having piles of tapes sitting in a closet.

Description

* MARC & RDA

* Worldcat & Connexion



DUKE LAW

Friday, June 14, 13

MARC & RDA (Sean)

We used existing staff and training in order to describe and provide access and metadata about our videos.

But more importantly there already existed a layer of interaction which we could program services against each of these data sources. For our integrated library system we could identify information about source and inventory and location.

We had other considerations which made the leveraging of the existing catalog and ILS tools important. These included:

- * controlled vocabularies for subject access
- * Authority control for people appearing and contributing to videos.
- * unique identifiers
- * Permanence as data sources (we know we aren't moving away from these tools since the library still needs them).

This is expensive and time consuming but is important to us since this is unique content.

Worldcat and Connexion

Vocabulary control (National Authority File, LCSH).

Unique identifiers for our works. Exposing in other places.

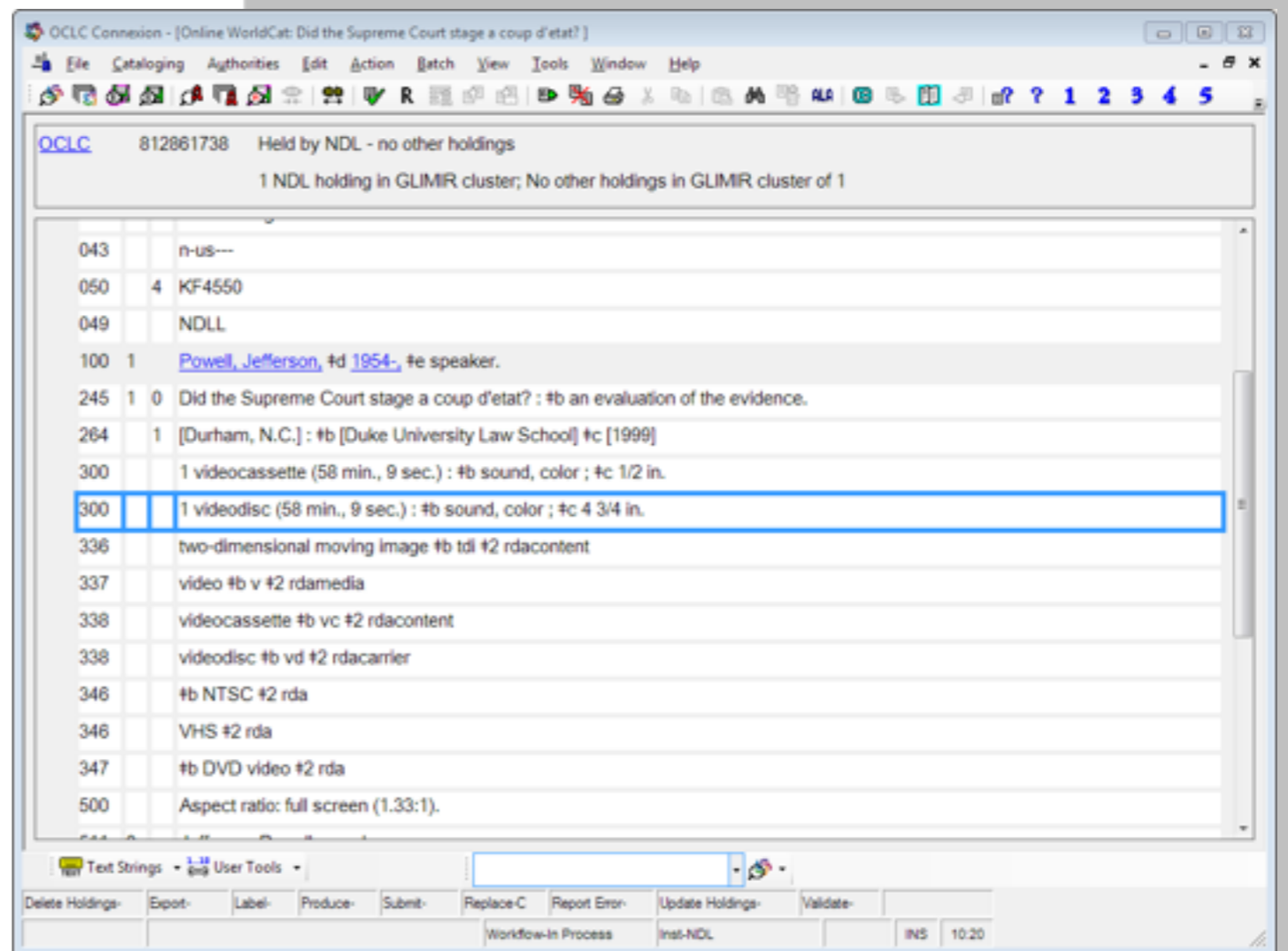
ALEPH

Information about our instantiations.

The ILS is really good for tracking inventory and putting things on the shelf which is important when we are managing a huge number of items. This works better than having piles of tapes sitting in a closet.

Description

- * MARC & RDA
- * Worldcat & Connexion
- * Aleph ILS



DUKE LAW

Friday, June 14, 13

MARC & RDA (Sean)

We used existing staff and training in order to describe and provide access and metadata about our videos.

But more importantly there already existed a layer of interaction which we could program services against each of these data sources. For our integrated library system we could identify information about source and inventory and location.

We had other considerations which made the leveraging of the existing catalog and ILS tools important. These included:

- * controlled vocabularies for subject access
- * Authority control for people appearing and contributing to videos.
- * unique identifiers
- * Permanence as data sources (we know we aren't moving away from these tools since the library still needs them).

This is expensive and time consuming but is important to us since this is unique content.

Worldcat and Connexion

Vocabulary control (National Authority File, LCSH).

Unique identifiers for our works. Exposing in other places.

ALEPH

Information about our instantiations.

The ILS is really good for tracking inventory and putting things on the shelf which is important when we are managing a huge number of items. This works better than having piles of tapes sitting in a closet.

Access

DUKE LAW



USTREAM

DUKE LAW

Friday, June 14, 13

General Architecture (Miguel)

One of our biggest changes was getting out of the streaming server business.

For many years we have operated our own streaming media server. Helix Streaming server. Getting out of the streaming server business.

Reasons for not operating our own:

- + Media isn't IT
- + Hard to reuse
- + Hard to manage content

Live Streaming (Miguel)

DukeStream and uStream

One is used for other higher impact events.

Not using YouTube Live yet.

Data Interoperability (Sean)

Our data flows from the catalog for our preserved material. However the data flow as we mentioned in our description section is significantly different in our production workflow.

Alternatives

We looked at other video content management systems:

There are some great options now:

- + Kaltura

The screenshot shows the YouTube channel for Duke Law School. At the top, there is a banner with the 'DUKE LAW' logo and a search bar. Below the banner, there are navigation tabs for 'Videos', 'Discussion', and 'About'. The main content area is divided into three sections:

- My Top Videos:** This section features five video thumbnails. The first is 'A Song's Tale: Mashups, Borrowing, and the Law' (57:29, 3,106 views). The second is 'Lessons Learned: The Bush Administration and the Su...' (53:01, 585 views). The third is 'Was Lochner Right? Natural Rights and the Fourteenth...' (59:19, 1,673 views). The fourth is 'Crisis in Gaza: Perspectives on the Conflict' (56:24, 192 views). The fifth is 'Have We Nationalized Our Banking System?' (1:34:33, 253 views).
- Duke Law's Video Spotlight:** This section features three video thumbnails. The first is 'The Honorable Evelyn Cannon '74 - Video Spotli...' (1:07, 314 views). The second is 'Open Access at Duke Law: James Boyle' (3:04, 1,224 views). The third is 'Open Access at Duke Law: Dick Danner' (3:47, 438 views).
- Recent uploads:** This section features five video thumbnails. The first is 'Lives in the Law: U.S. Solicitor General Donald B...' (1:01:22, 114 views, 3 weeks ago). The second is 'Chenghong Wang and Li Zhao: A Few Questions Ab...' (6:28, 53 views, 3 weeks ago). The third is '2013 Dean's Cup Moot Court Competition: Final Round' (1:22:13, 284 views, 1 month ago). The fourth is 'Proving Innocence: LaMonte Armstrong Exonerated Wit...' (5:06, 206 views, 1 month ago). The fifth is 'Duke University School of Law; Michael J Goodson L...' (1:46, 121 views, 1 month ago).

DUKE LAW

Friday, June 14, 13

Manual Uploading (Miguel)

Video gets into youtube in a couple ways.

For our production workflow however, timing is more of an issue and media services loads videos directly into YouTube as part of their workflow.

A review step for our communication staff.

Google APIs (Sean)

We can use the various YouTube APIs, Older Google Data API and the newer Google API to automatically push video based upon whatever data into youtube.

Terms of Service (Miguel)

Were not terribly burdensome to us. We aren't streaming class content or reserve content.

Getting more access > 15 minutes for educational institutions. Just need to ask.

Limitations (Sean)

There are some issues with identification and findability. Titles are limited to 150 characters. Similarly descriptions are limited to 5000 characters. Keywords are limited to 500 characters total. With any particular keyword only being 30 characters. Through the API also you can't upload a tag with a comma in it (it breaks on commas). Just some gotchas we had.

The screenshot shows a web browser window displaying the Duke University School of Law website. The page title is "It's Constitutional, What Now? | Duke University School of Law". The URL in the address bar is "law.duke.edu/video/its-constitutional-what-now/". The website has a dark blue header with the "DUKE LAW" logo and navigation links for "Current Students", "Faculty", "Staff", and "Alumni". Below the logo are links for "Directory", "Calendar", "News", "Library", "Computing", and "Visit". A search bar is located on the right side of the header. The main content area features the article title "It's Constitutional, What Now?" dated "September 04, 2012". The article text discusses the Supreme Court's health care ruling and mentions several professors. Below the text is a video player showing three men in suits sitting at a table. To the right of the main content is a sidebar with a "DUKE LAW NEWS" section containing links for "For the Media", "Blogs & Social Media", "Video", "Publications at Duke Law", and "Contact Us". Below this is a "RECENT NEWS" section with three news items: "Memon JD/MBA '08 wins Washington Post Hunt" (linked to Constitutional Daily), "Zelenak analyzes IRS scrutiny of Tea Party groups" (linked to Card Hub), and "Students give thanks at Thank-a-thon".

DUKE LAW

Friday, June 14, 13

Importing into Drupal (Sean)

Programatic harvesting of data from YouTube.

We have a thin PHP script that creates a sorted XML for the Drupal feeds module so it can create and update nodes based on the data being exposed by YouTube.

Drupal

We get nodes within our website which get indexed and in turn point back to YouTube. Again it was a minimal amount of work getting it onto the website. It is the content, descriptions, tagging and production of the video which is the hard part.

YouTube Embed

It just works for us. We can pull in the relevant code from the the provided XML feed and generate appropriate surrogates and players within our website.

The screenshot shows a web browser window with the URL `search.library.duke.edu/search?id=DUKE005588014`. The page header features the Duke University Libraries logo and the word "CATALOG". Below the header, there are navigation tabs for "All", "Catalog", "e-Journals", and "Databases". A search bar is present with a "Search" button and a "0 marked items" indicator. A "Ask Us NOW!" button is also visible.

The search results section displays the following information:

- Title:** It's constitutional, now what?
- Format:** Internet resource; Video DVD
- Language:** English
- Online Access:** http://youtu.be/fG4g5yK_SPE
- Summary:** Professors Neil Siegel, Stephen Sachs along with Asheesh Agarwal of Ogletree Deakins discuss the Supreme Court's ruling in National Federation of Independent Business v. Sebelius. Professor Joseph Blocher moderates. Sponsored by the American Constitution Society, the Federalist Society, and the Health Law Society.

On the right side of the results, there are several utility buttons: Email, Print, Text Message, Refworks, Endnote (file), delicious, Get this title (place request), and Not on the shelves? Request Missing Book.

Below the summary, there are tabs for "Location", "Details", "Subjects", and "Summary". The "Location" tab is active, showing the following information:

Location	Details	Subjects	Summary
Duke			
Goodson Law Library	Media Click for map	09042012 DVD	Available - Library Use Only.

At the bottom of the page, there is a navigation menu with links to various library services: Alumni Portal, Divinity School Library, Ford Library, Goodson Law Library, Library Service Center, Lilly Library, Marine Lab Library, Medical Center Library, Mobile, Music Library, Perkins/Bostock Library, Rubenstein Library, and The Link. There is also a footer with contact information: 919-660-5870 (Perkins Circulation Desk) and links for Use and Reproduction, Privacy, Contact Us, Support the Duke Libraries, Jobs, and Duke.edu.

DUKE LAW

Friday, June 14, 13

Library Catalog (Sean)

And yes we get it in our library catalog.

Authority Searching (Sean)

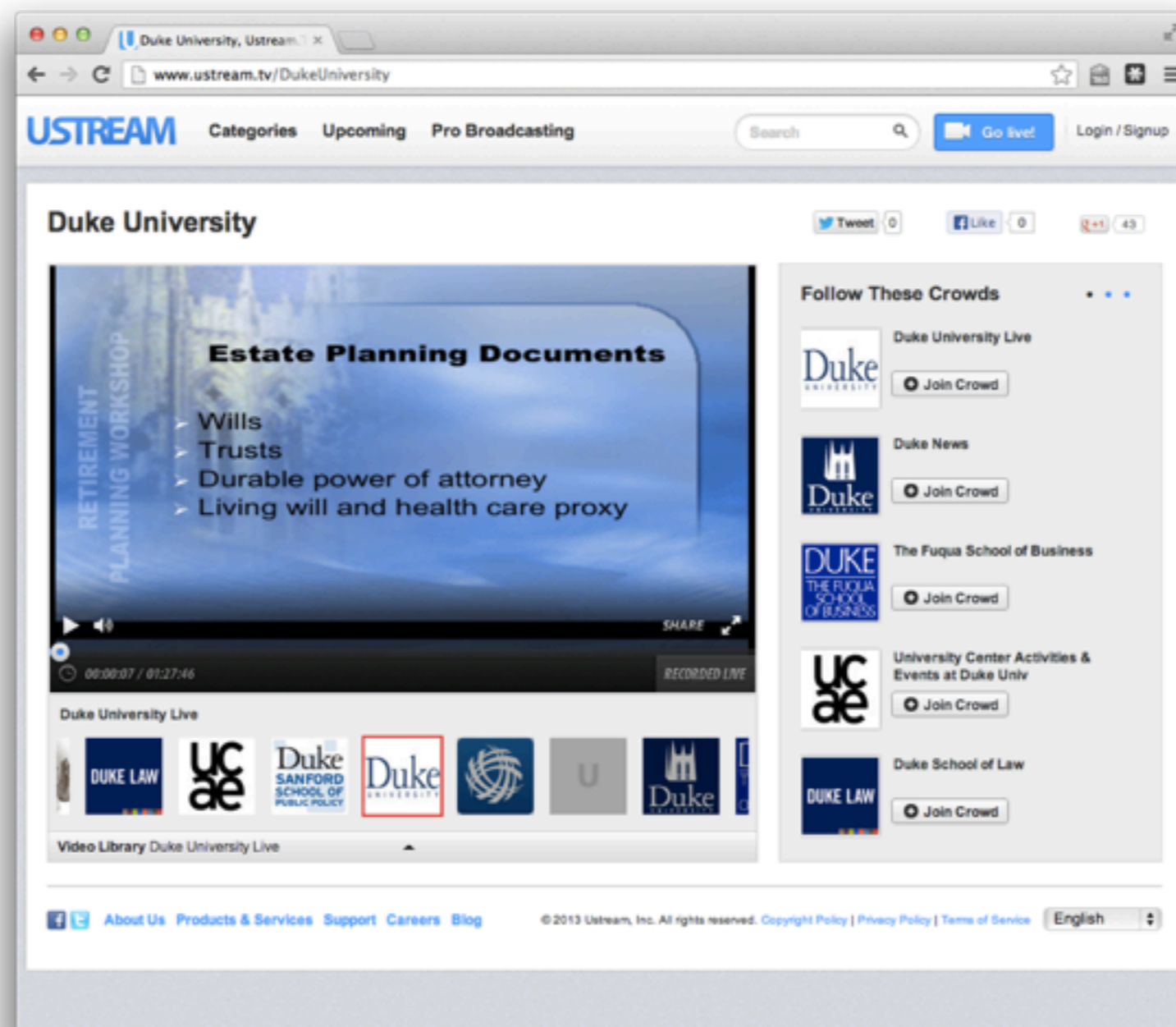
Does this matter? It probably doesn't. But this system gets us amazing ability to search and retrieve upon the rich data that we are creating for our content.

Durability of Systems (Sean)

And we are sure (well pretty sure) that it isn't going away.

This is versus the website where we have migrated three times in six years to different content management systems.

Streaming



DUKE LAW

Friday, June 14, 13

uStream (Miguel)

For events of interest to campus. Can support X number of users. Can stream through our website.

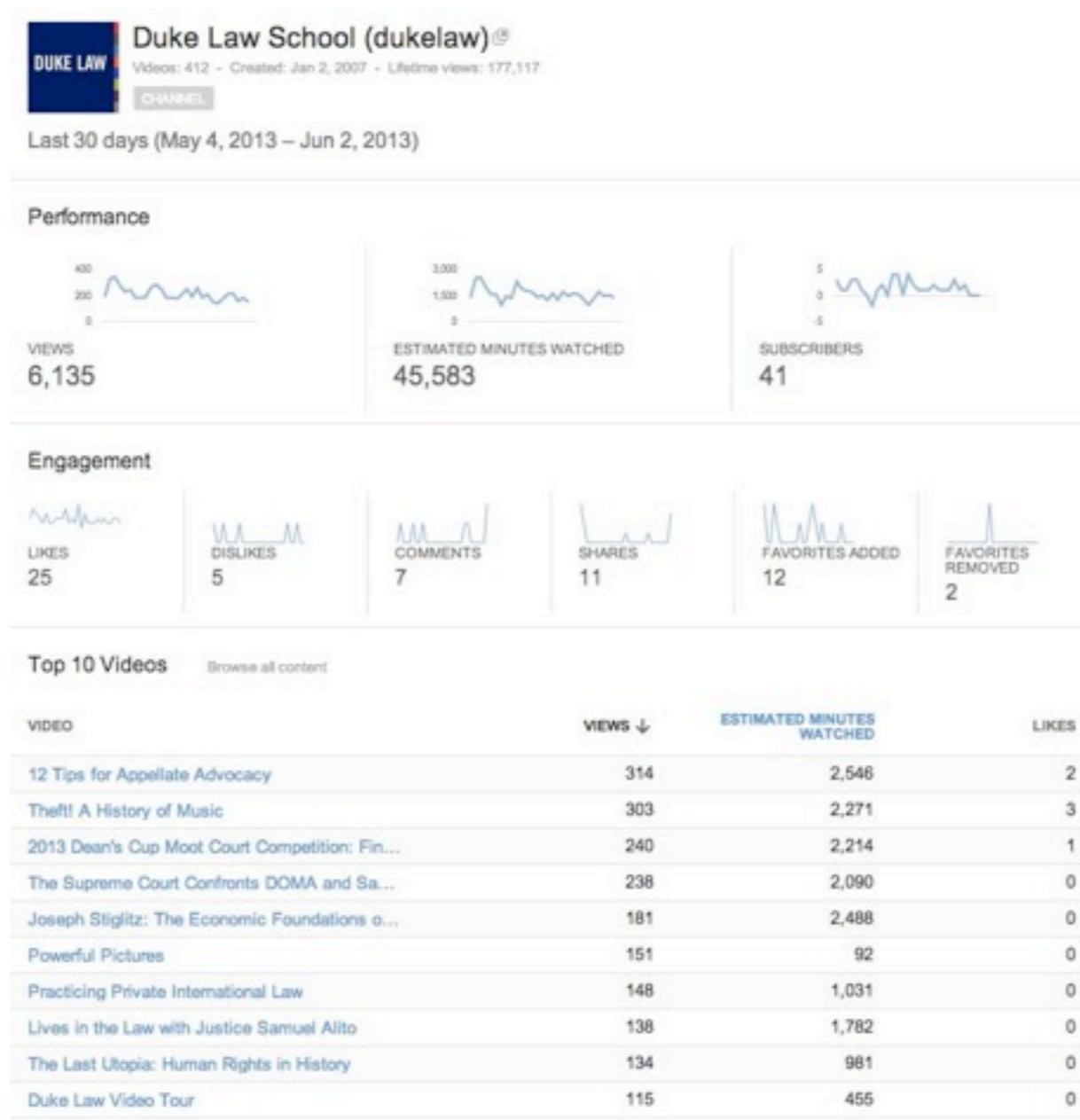
DukeStream (Miguel)

For other sorts of events. Can stream through law school website

YouTube Live (Miguel)

CALI does it! We should too. Right? Miguel.

Metrics & Assessment



DUKE LAW

Friday, June 14, 13

Metrics & Assessment (Miguel)

Analytics? Helps in setting or assigning preservation ties to events and requests... (Comms)

What do they look like for us?

What do we do with them?

Leading us into revisiting what we actually collect.

Forks Not Taken

DUKE LAW

Friday, June 14, 13

Lossless (Sean)

We chiefly didn't go down this road since we didn't have the tools and infrastructure to support this. This is

PBCore (Sean)

Explored PBCore a bit. And I would recommend it as a description framework. But we had so much data already in MARC-AACR2 that it didn't make sense to throw that out and restart. There are a number of PBCore crosswalks available though that might help interoperability.

Digital Asset Management Systems (Sean)

Our planning included investigations into these, but when we started to look at implementing these we ran into the classic project show stoppers:

- + no resources
- + no money
- + no staff

Digital Repository Systems (Sean)

At the time we didn't have the capability or tooling to do it. This would have been some sort of framework such as Fedora or dSpace. We are a bePress customer and we looked at putting our video into there, but the biggest issue we had was how to handle dark storage. It seemed more straightforward to continue working directly on the file system and count on doing a thin layer of preservation activities until we get to the time we can actually do something more substantial in collaboration with our campus partners.

Classroom Recordings (Miguel)

Forks Not Taken

* Lossless

DUKE LAW

Friday, June 14, 13

Lossless (Sean)

We chiefly didn't go down this road since we didn't have the tools and infrastructure to support this. This is

PBCore (Sean)

Explored PBCore a bit. And I would recommend it as a description framework. But we had so much data already in MARC-AACR2 that it didn't make sense to throw that out and restart. There are a number of PBCore crosswalks available though that might help interoperability.

Digital Asset Management Systems (Sean)

Our planning included investigations into these, but when we started to look at implementing these we ran into the classic project show stoppers:

- + no resources
- + no money
- + no staff

Digital Repository Systems (Sean)

At the time we didn't have the capability or tooling to do it. This would have been some sort of framework such as Fedora or dSpace. We are a bePress customer and we looked at putting our video into there, but the biggest issue we had was how to handle dark storage. It seemed more straightforward to continue working directly on the file system and count on doing a thin layer of preservation activities until we get to the time we can actually do something more substantial in collaboration with our campus partners.

Classroom Recordings (Miguel)

Forks Not Taken

* Lossless

* PBCore

DUKE LAW

Friday, June 14, 13

Lossless (Sean)

We chiefly didn't go down this road since we didn't have the tools and infrastructure to support this. This is

PBCore (Sean)

Explored PBCore a bit. And I would recommend it as a description framework. But we had so much data already in MARC-AACR2 that it didn't make sense to throw that out and restart. There are a number of PBCore crosswalks available though that might help interoperability.

Digital Asset Management Systems (Sean)

Our planning included investigations into these, but when we started to look at implementing these we ran into the classic project show stoppers:

- + no resources
- + no money
- + no staff

Digital Repository Systems (Sean)

At the time we didn't have the capability or tooling to do it. This would have been some sort of framework such as Fedora or dSpace. We are a bePress customer and we looked at putting our video into there, but the biggest issue we had was how to handle dark storage. It seemed more straightforward to continue working directly on the file system and count on doing a thin layer of preservation activities until we get to the time we can actually do something more substantial in collaboration with our campus partners.

Classroom Recordings (Miguel)

Forks Not Taken

- * Lossless
- * PBCore
- * Digital Asset Management Systems

DUKE LAW

Friday, June 14, 13

Lossless (Sean)

We chiefly didn't go down this road since we didn't have the tools and infrastructure to support this. This is

PBCore (Sean)

Explored PBCore a bit. And I would recommend it as a description framework. But we had so much data already in MARC-AACR2 that it didn't make sense to throw that out and restart. There are a number of PBCore crosswalks available though that might help interoperability.

Digital Asset Management Systems (Sean)

Our planning included investigations into these, but when we started to look at implementing these we ran into the classic project show stoppers:

- + no resources
- + no money
- + no staff

Digital Repository Systems (Sean)

At the time we didn't have the capability or tooling to do it. This would have been some sort of framework such as Fedora or dSpace. We are a bePress customer and we looked at putting our video into there, but the biggest issue we had was how to handle dark storage. It seemed more straightforward to continue working directly on the file system and count on doing a thin layer of preservation activities until we get to the time we can actually do something more substantial in collaboration with our campus partners.

Classroom Recordings (Miguel)

Forks Not Taken

- * Lossless
- * PBCore
- * Digital Asset Management Systems
- * Classroom Recordings

DUKE LAW

Friday, June 14, 13

Lossless (Sean)

We chiefly didn't go down this road since we didn't have the tools and infrastructure to support this. This is

PBCore (Sean)

Explored PBCore a bit. And I would recommend it as a description framework. But we had so much data already in MARC-AACR2 that it didn't make sense to throw that out and restart. There are a number of PBCore crosswalks available though that might help interoperability.

Digital Asset Management Systems (Sean)

Our planning included investigations into these, but when we started to look at implementing these we ran into the classic project show stoppers:

- + no resources
- + no money
- + no staff

Digital Repository Systems (Sean)

At the time we didn't have the capability or tooling to do it. This would have been some sort of framework such as Fedora or dSpace. We are a bePress customer and we looked at putting our video into there, but the biggest issue we had was how to handle dark storage. It seemed more straightforward to continue working directly on the file system and count on doing a thin layer of preservation activities until we get to the time we can actually do something more substantial in collaboration with our campus partners.

Classroom Recordings (Miguel)

Forks Not Taken

- * Lossless
- * PBCore
- * Digital Asset Management Systems
- * Classroom Recordings
- * Digital Repository Systems

DUKE LAW

Friday, June 14, 13

Lossless (Sean)

We chiefly didn't go down this road since we didn't have the tools and infrastructure to support this. This is

PBCore (Sean)

Explored PBCore a bit. And I would recommend it as a description framework. But we had so much data already in MARC-AACR2 that it didn't make sense to throw that out and restart. There are a number of PBCore crosswalks available though that might help interoperability.

Digital Asset Management Systems (Sean)

Our planning included investigations into these, but when we started to look at implementing these we ran into the classic project show stoppers:

- + no resources
- + no money
- + no staff

Digital Repository Systems (Sean)

At the time we didn't have the capability or tooling to do it. This would have been some sort of framework such as Fedora or dSpace. We are a bePress customer and we looked at putting our video into there, but the biggest issue we had was how to handle dark storage. It seemed more straightforward to continue working directly on the file system and count on doing a thin layer of preservation activities until we get to the time we can actually do something more substantial in collaboration with our campus partners.

Classroom Recordings (Miguel)

Extension Road for the Future

DUKE LAW

Friday, June 14, 13

Sean & Miguel

Web server. Building libraries with our data.

Building into a real set of repository services.

Fedora (Sean)

Our campus is building out a Hydra head on top of Fedora for its digital production workflow. This would be their preservation and service layer for their content. We are looking into collaborating in this, however issues of how to handle the large volume of video remain to be fully explored.

Some of this could be in software, and providing an ingest interface into this sort of system, but the reality is that would have to require a lot of localizations that are wouldn't in the end get us to a better state then the one we are in.

Web Content (Sean)

Because we have rich data in our catalog about each video we can do more with organizing our content on the website. We're planning to work with the various website stakeholders to work on an internal vocabulary and taxonomy in order to reuse our content in various places.

A good example would be something like a visitor coming to a page on International Studies and getting XYZ.

Extension Road for the Future

* High Definition

DUKE LAW

Friday, June 14, 13

Sean & Miguel

Web server. Building libraries with our data.

Building into a real set of repository services.

Fedora (Sean)

Our campus is building out a Hydra head on top of Fedora for its digital production workflow. This would be their preservation and service layer for their content. We are looking into collaborating in this, however issues of how to handle the large volume of video remain to be fully explored.

Some of this could be in software, and providing an ingest interface into this sort of system, but the reality is that would have to require a lot of localizations that are wouldn't in the end get us to a better state than the one we are in.

Web Content (Sean)

Because we have rich data in our catalog about each video we can do more with organizing our content on the website. We're planning to work with the various website stakeholders to work on an internal vocabulary and taxonomy in order to reuse our content in various places.

A good example would be something like a visitor coming to a page on International Studies and getting XYZ.

Extension Road for the Future

- * High Definition
- * All Digital Production Workflow

DUKE LAW

Friday, June 14, 13

Sean & Miguel

Web server. Building libraries with our data.

Building into a real set of repository services.

Fedora (Sean)

Our campus is building out a Hydra head on top of Fedora for its digital production workflow. This would be their preservation and service layer for their content. We are looking into collaborating in this, however issues of how to handle the large volume of video remain to be fully explored.

Some of this could be in software, and providing an ingest interface into this sort of system, but the reality is that would have to require a lot of localizations that are wouldn't in the end get us to a better state then the one we are in.

Web Content (Sean)

Because we have rich data in our catalog about each video we can do more with organizing our content on the website. We're planning to work with the various website stakeholders to work on an internal vocabulary and taxonomy in order to reuse our content in various places.

A good example would be something like a visitor coming to a page on International Studies and getting XYZ.

Extension Road for the Future

- * High Definition
- * All Digital Production Workflow
- * Digital Repository Systems

DUKE LAW

Friday, June 14, 13

Sean & Miguel

Web server. Building libraries with our data.

Building into a real set of repository services.

Fedora (Sean)

Our campus is building out a Hydra head on top of Fedora for its digital production workflow. This would be their preservation and service layer for their content. We are looking into collaborating in this, however issues of how to handle the large volume of video remain to be fully explored.

Some of this could be in software, and providing an ingest interface into this sort of system, but the reality is that would have to require a lot of localizations that are wouldn't in the end get us to a better state than the one we are in.

Web Content (Sean)

Because we have rich data in our catalog about each video we can do more with organizing our content on the website. We're planning to work with the various website stakeholders to work on an internal vocabulary and taxonomy in order to reuse our content in various places.

A good example would be something like a visitor coming to a page on International Studies and getting XYZ.

Extension Road for the Future

- * High Definition
- * All Digital Production Workflow
- * Digital Repository Systems
- * Web Content

DUKE LAW

Friday, June 14, 13

Sean & Miguel

Web server. Building libraries with our data.

Building into a real set of repository services.

Fedora (Sean)

Our campus is building out a Hydra head on top of Fedora for its digital production workflow. This would be their preservation and service layer for their content. We are looking into collaborating in this, however issues of how to handle the large volume of video remain to be fully explored.

Some of this could be in software, and providing an ingest interface into this sort of system, but the reality is that would have to require a lot of localizations that are wouldn't in the end get us to a better state than the one we are in.

Web Content (Sean)

Because we have rich data in our catalog about each video we can do more with organizing our content on the website. We're planning to work with the various website stakeholders to work on an internal vocabulary and taxonomy in order to reuse our content in various places.

A good example would be something like a visitor coming to a page on International Studies and getting XYZ.

Questions and comments

* Thank You!

* Sean Chen

@gugek

schen@law.duke.edu

* Miguel A. Bordo

@miguelbordo

bordo@law.duke.edu



110/365 - Mattel HOT WHEELS Ice Cream Truck by davidd
<http://www.flickr.com/photos/puukibeach/7101217829/>
Attribution-NonCommercial License