12 November 2015

Digitizing Collections at the National Transportation Library

Megan Tietjen, National Transportation Library & Josue Rivera-Olds, Masters Student, University of Washington iSchool

Ingesting Digitized Collections into the NTL in Bulk

Megan Tietjen
Metadata Librarian, NTL

Ingesting Collections

- Managing bulk metadata
  - Receive collection
  - Check for duplicates
  - Assess and edit metadata
  - Create file paths
  - Bulk upload

- Can be labor intensive upfront

Bulk Upload Template

I2 | English

	A	B	C	D	E	F	G	H
1	Title	Primary URL	Personal Creator	Corporate Creator	Corporate Contributor	TRT Terms	Corporate Publisher	Pubication Date
2								
3								
4								
5								
6								

C3

	I	J	K	L	M	N	O	P	Q
1	Lanuage	Geographical Coverage	Format	Resoure Type	Abstract	Contract Number	Report Number	Collection	Status
2	English		PDF	Tech Report					Approved
3									
4									
5									

Internet Archive

- Cooperative digitization
- Hosting agreement with Internet Archive
- Multiple locations of digitized reports

FTA 50th Anniversary Collection

Construction of Morgantown (WV) Personal Rapid Transit system

- Working with Internet Archive on digitization
- NTL Digital Repository = 143 records
- Ingesting = 395 items

FTA 50th Anniversary Collection

	A	B	C	D	E	F	G	H	I	J
1	collection	contributc	coverage	creator	date	descriptio	download	foldoutco	format	headerIm
2	fta50thani	Federal Transit Adm		Ms. Jo Tuc	1990-01-01T00:00:00Z		20	0	Abbyy GZ,	Animated
3	fta50thani	Federal Transit Adm		Ms. Jo Tuc	1992-05-01T00:00:00Z		21	0	Abbyy GZ,	Animated
4	fta50thani	Federal Transit Adm		Ms. Jo Tuc	1993-03-01T00:00:00Z		18	0	Abbyy GZ,	Animated
5	fta50thani	Federal Transit Adm		Ms. Jo Tuc	1994-02-01T00:00:00Z		26	0	Abbyy GZ,	Animated
6	fta50thani	Federal Transit Adm		Ms. Jo Tuc	1995-03-01T00:00:00Z		24	0	Abbyy GZ,	Animated
7	fta50thani	Federal Transit Adm		Ms. Jo Tuc	1996-02-01T00:00:00Z		29	0	Abbyy GZ,	Animated
8	fta50thani	Federal Transit Adm		Nancy Gru	1997-04-01T00:00:00Z		26	0	Abbyy GZ,	Animated
9	fta50thani	Federal Transit Adm		Gordon J.	1997-01-01T00:00:00Z		22	0	Abbyy GZ,	Animated
10	fta50thani	Federal Transit Adm		Gordon J.	1997-01-01T00:00:00Z		26	0	Abbyy GZ,	Animated
11	fta50thani	Federal Transit Adm		Gordon J.	1997-01-01T00:00:00Z		12	0	Abbyy GZ,	Animated
12	fta50thani	Federal Transit Adm		Nancy Gru	1998-03-01T00:00:00Z		18	0	Abbyy GZ,	Animated
13	fta50thani	Federal Transit Administration			1998-01-01T00:00:00Z		13	0	Abbyy GZ,	Animated
14	fta50thani	Federal Transit Administration			1998-08-20T00:00:00Z		9	0	Abbyy GZ,	Animated
15	fta50thani	Federal Transit Adm		Nancy Gru	1998-03-01T00:00:00Z		11	0	Abbyy GZ,	Animated
16	fta50thani	Federal Transit Administration		- 50th Anniversary D			51	0	Abbyy GZ,	Animated
17	fta50thani	Federal Transit Adm		Charles Ri	1984-12-01T00:00:00Z		12	0	Abbyy GZ,	Animated
18	fta50thani	Federal Transit Administration			1994-04-28T00:00:00Z		18		Abbyy GZ,	Animated
19	fta50thani	Federal Transit Adm		EG&G Dyn	1991-09-01T00:00:00Z		34	0	Abbyy GZ,	Animated
20	fta50thani	Federal Transit Adm		USDOT/Re	1996-01-01T00:00:00Z		13	0	Abbyy GZ,	Animated
21	fta50thani	Federal Transit Adm		John A. Vc	1994-01-01T00:00:00Z		33	0	Abbyy GZ,	Animated
22	fta50thani	Federal Transit Adm		Dr. Shinya	1994-09-01T00:00:00Z		15	0	Abbyy GZ,	Animated
23	fta50thani	Federal Transit Adm		Dept. of S	1993-03-01T00:00:00Z		12	0	Abbyy GZ,	Animated
24	fta50thani	Federal Transit Adm		Rice Cente	1986-07-01T00:00:00Z		13	0	Abbyy GZ,	Animated
25	fta50thani	Federal Transit Adm		Hiram J. W	1984-04-30T00:00:00Z		166	46	Abbyy GZ,	Animated
26	fta50thani	Federal Transit Adm		EG&G Dyn	1993-09-01T00:00:00Z		18	0	Abbyy GZ,	Animated
27	fta50thani	Federal Transit Adm		Dr. Robert	1988-01-01T00:00:00Z		12	0	Abbyy GZ,	Animated
28	fta50thani	Federal Transit Administration			1981-07-01T00:00:00Z		20	0	Abbyy GZ,	Animated
29	fta50thani	Federal Transit Administration			1978-01-01T00:00:00Z		9	4	Abbyy GZ,	Animated
30	fta50thani	Federal Transit Administration			1978-01-01T00:00:00Z		10	0	Abbyy GZ,	Animated

	A	B	C	D
1	Title	Primary URL	creator	Corporate Creator
2	1990 Information		Tucci, Jo	
3	1991 Statistical		Tucci, Jo	
4	1992 Statistical		Tucci, Jo	
5	1993 Statistical		Tucci, Jo	
6	1994 Statistical		Tucci, Jo	
7	1995 Statistical		Tucci, Jo	
8	1996 Statistical Summaries - FTA		Grubb, Nancy	
9	1997 Statistical Summaries - FTA		Grubb, Nancy	
10	1998 Statistical Summaries - FTA		Grubb, Nancy	
11	Abstracts and Annotated			
12	Access: Brokering Paratransit			Charles River Association, Inc.
13	ADA Paratransit Handbook: Implementing the			EG&G Dynatrend
14	Advanced Public Transportation Systems: The State of the Art - Update			United States. Dept. of Transportation. Research and Special Programs
15	Advanced Public Transportation			John A. Volpe National

Montana DOT Collection

- 468 new records
- Includes reports from FHWA and state DOTs
- Provided permanent links to MDOT and OCLC

Directed Field Work at the National Transportation Library

by Josue Rivera-Olds
MLIS Student, University of Washington iSchool

Source Image: Flickr Image by Thisisbossi

About the Collection

- 4,506 Records
  - Digitized orders, memos, and reports
  - 100% were PDF's
  - Date range: 1948 to 1999
  - Originally hosted on the dot.gov website

- 350 of the records were located in the NTL database
- 26 are duplicates
- 50 of the records were already stored in the NTL database but the URL PDF was broken

- Methodology
- Process
- Outcomes

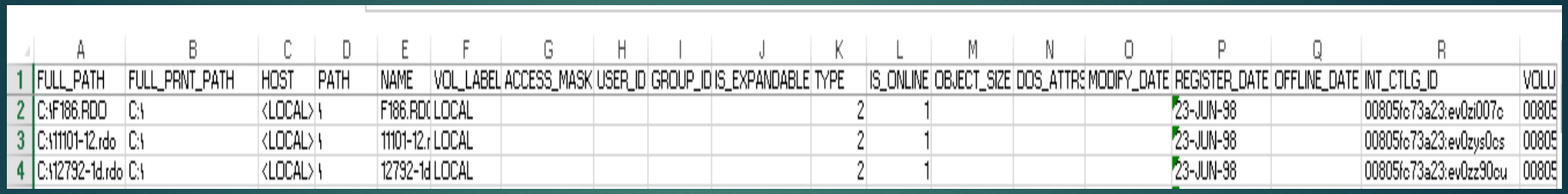
Developing a Crosswalk

- Developed a basic crosswalk from the existing metadata elements
- Focused on the **date** element, organized chronologically from oldest to newest

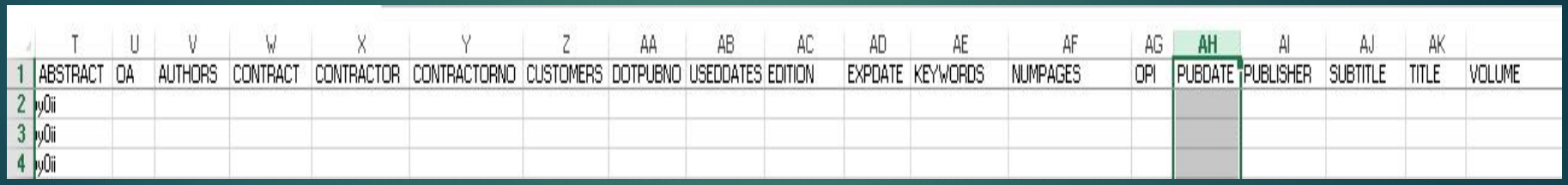
Outcome:

- Provided organization to the data and metadata elements
- Streamlined the retrieval process

Before the Crosswalk

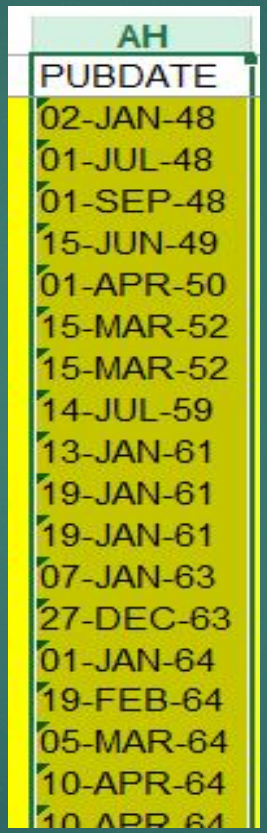


	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	
1	FULL_PATH	FULL_PRNT_PATH	HOST	PATH	NAME	VOL_LABEL	ACCESS_MASK	USER_ID	GROUP_ID	IS_EXPANDABLE	TYPE	IS_ONLINE	OBJECT_SIZE	DOS_ATTRS	MODIFY_DATE	REGISTER_DATE	OFFLINE_DATE	INT_CTLG_ID	VOLU
2	C:\F186.RDO	C:\	<LOCAL>	\	F186.RDO	LOCAL					2	1				23-JUN-98		00805fc73a23:ev0zi007c	00805
3	C:\11101-12.rdo	C:\	<LOCAL>	\	11101-12.r	LOCAL					2	1				23-JUN-98		00805fc73a23:ev0zys0cs	00805
4	C:\12792-1d.rdo	C:\	<LOCAL>	\	12792-1d	LOCAL					2	1				23-JUN-98		00805fc73a23:ev0zz90cu	00805



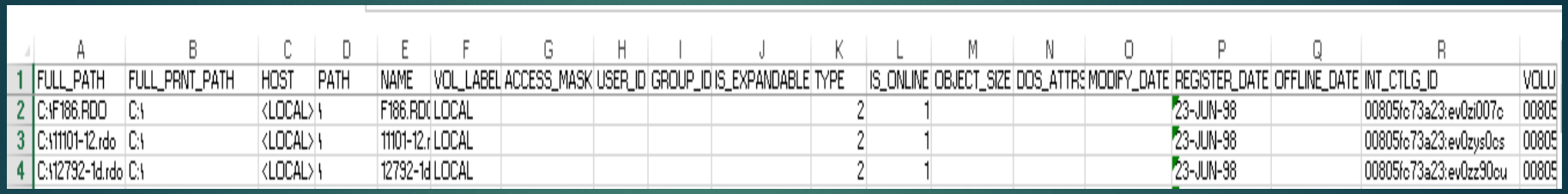
	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ	AK	
1	ABSTRACT	OA	AUTHORS	CONTRACT	CONTRACTOR	CONTRACTORNO	CUSTOMERS	DOTPUBNO	USEDDATES	EDITION	EXPDATE	KEYWORDS	NUMPAGES	OPI	PUBDATE	PUBLISHER	SUBTITLE	TITLE	VOLUME
2	¤y0ii																		
3	¤y0ii																		
4	¤y0ii																		

Result of the Crosswalk

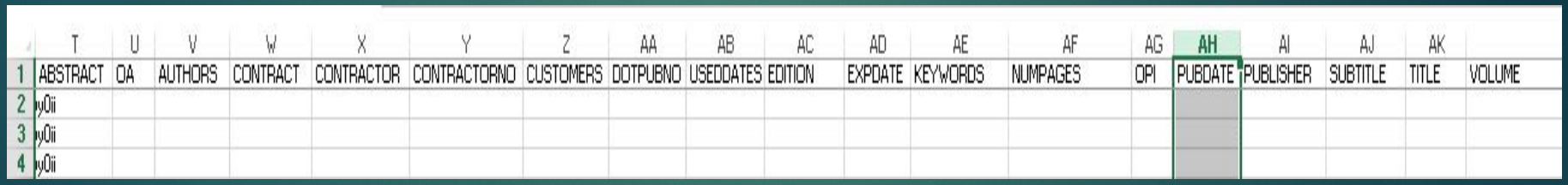


AH
PUBDATE
02-JAN-48
01-JUL-48
01-SEP-48
15-JUN-49
01-APR-50
15-MAR-52
15-MAR-52
14-JUL-59
13-JAN-61
19-JAN-61
19-JAN-61
07-JAN-63
27-DEC-63
01-JAN-64
19-FEB-64
05-MAR-64
10-APR-64
10-APR-64

Before the Crosswalk



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	
1	FULL_PATH	FULL_PRNT_PATH	HOST	PATH	NAME	VOL_LABEL	ACCESS_MASK	USER_ID	GROUP_ID	IS_EXPANDABLE	TYPE	IS_ONLINE	OBJECT_SIZE	DOS_ATTRS	MODIFY_DATE	REGISTER_DATE	OFFLINE_DATE	INT_CTLG_ID	VOLU
2	C:\F186.RDO	C:\	<LOCAL>	\	F186.RDO	LOCAL					2	1				23-JUN-98		00805fc73a23:ev0zi007c	00805
3	C:\11101-12.rdo	C:\	<LOCAL>	\	11101-12.r	LOCAL					2	1				23-JUN-98		00805fc73a23:ev0zys0cs	00805
4	C:\12792-1d.rdo	C:\	<LOCAL>	\	12792-1d	LOCAL					2	1				23-JUN-98		00805fc73a23:ev0zz90cu	00805



	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ	AK	
1	ABSTRACT	OA	AUTHORS	CONTRACT	CONTRACTOR	CONTRACTORNO	CUSTOMERS	DOTPUBNO	USEDDATES	EDITION	EXPDATE	KEYWORDS	NUMPAGES	OPI	PUBDATE	PUBLISHER	SUBTITLE	TITLE	VOLUME
2	¤y0ii																		
3	¤y0ii																		
4	¤y0ii																		

Assessing the Collections

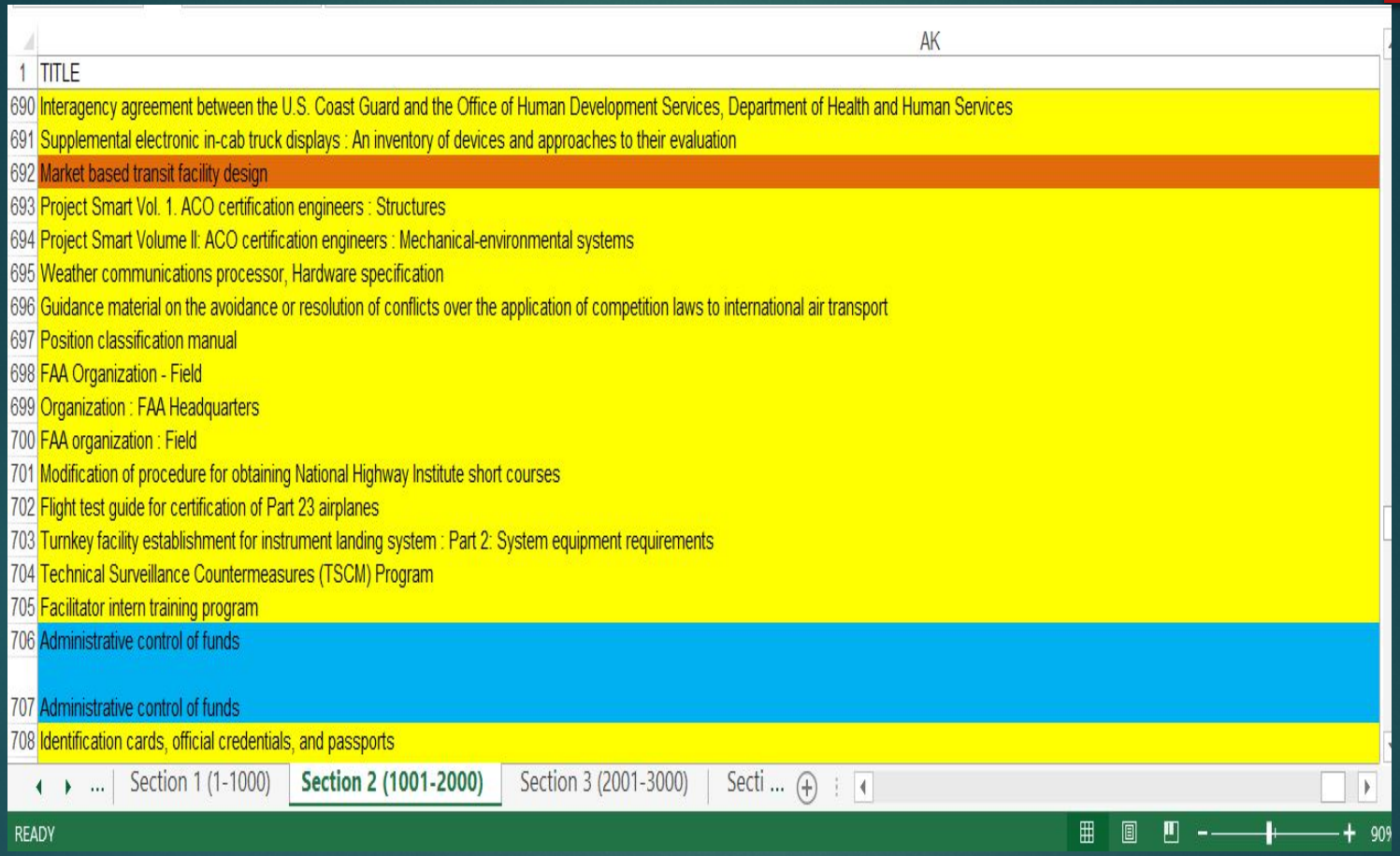
- Existing Metadata
- Document Quality
- Format
- Duplication

Title	Transit Corporate Planning: A Methodology for Trading Off Fares, Service Levels, and Capital Budgets{7072}
Record ID	7072
Personal Name Creator	Lago, Armando M. {8576}
Corporate Contributor	Ecosometrics, Inc. {4148}; United States. Urban Mass Transportation Administration {16215}
Publication Date	Jun 30, 1985
Language	English
Abstract	The financial environment in which the transit industry is operating is rapidly changing. Transit boards, city councils, and general managers are facing difficult decisions. There is an increasing realization that perhaps managerial and financial planning tools used by the transit industry are not up to the difficult task of coordinating and deciding on fares, services and capital budgets, and that perhaps business planning tools and processes need to be applied in American transit settings. One of the problems with transit planning and management is that fare and service level decisions are rarely planned and considered together despite the fact that fares and service levels are intrinsically related. Indeed, current transit planning and managerial practices in place emphasize separate analysis of appropriate fares, service levels, and capital budgets with the concomitant result that there is no overall balance between these three important elements. Setting fares and service levels efficiently is explored from a business approach; this process, sometimes known as "corporate planning," involves setting goals and determining a yardstick against which all actions are measured. The corporate goal refers to the most important or highest achievement that the corporation intends to attain. The corporate objectives translate the corporate goal into intended achievements whose pursuit will accomplish the corporate goal. The objectives are specific in nature, have to be clearly defined, capable of being quantified and objectively measured, monitored, and evaluated. How transit organizations can apply the corporate planning process is discussed with examples given.
TRT Terms	Transit operating agencies {Nbbcu}; Business practices {Cbd}; Decision theory {Ushd}; Planning methods {Epdm}; Fares {Bbkf}; Level of service {Bbrm}; Capital investments {Cfeic}; Strategic planning {Epdms}; Public transit {Aet}; Financial analysis {Cga}
General Subjects	Annualized capital costs; Corporate goals; Corporate objectives; Corporate planning; Investment Programs; Performance standards; Transit Objectives
Classification	NTL - PUBLIC TRANSPORTATION - Transit Planning and Policy {51}; NTL - PUBLIC TRANSPORTATION - Transit Economics and Finances {59}
Geographical Coverage	United States; United Kingdom
TRIS Online Accession No	00476982
Report Number	DOT-I-87-12
Resource type	Research Paper
URL	http://ntl.bts.gov/DOCS/tcp.html
Format	HTML
Status	UnApproved

This an extract of a record in the NTL Workroom database

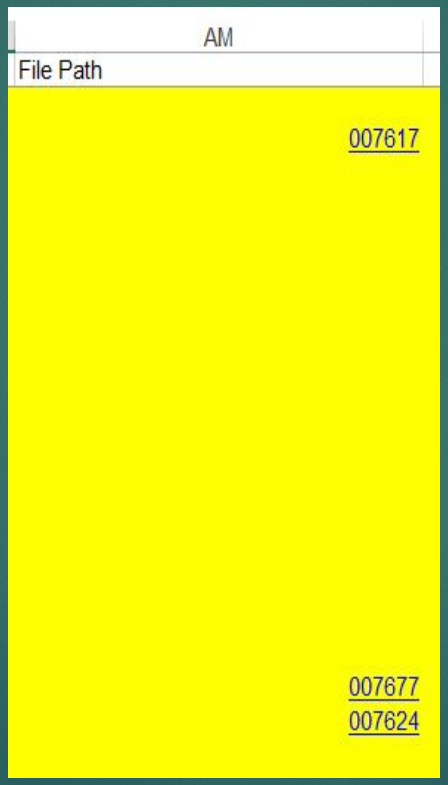
Assessing the Collections

- Tracked which records were on the database and which were missing.
- Color coded the records in yellow, blue, and brown.



	AK
1	TITLE
690	Interagency agreement between the U.S. Coast Guard and the Office of Human Development Services, Department of Health and Human Services
691	Supplemental electronic in-cab truck displays : An inventory of devices and approaches to their evaluation
692	Market based transit facility design
693	Project Smart Vol. 1. ACO certification engineers : Structures
694	Project Smart Volume II: ACO certification engineers : Mechanical-environmental systems
695	Weather communications processor, Hardware specification
696	Guidance material on the avoidance or resolution of conflicts over the application of competition laws to international air transport
697	Position classification manual
698	FAA Organization - Field
699	Organization : FAA Headquarters
700	FAA organization : Field
701	Modification of procedure for obtaining National Highway Institute short courses
702	Flight test guide for certification of Part 23 airplanes
703	Turnkey facility establishment for instrument landing system : Part 2: System equipment requirements
704	Technical Surveillance Countermeasures (TSCM) Program
705	Facilitator intern training program
706	Administrative control of funds
707	Administrative control of funds
708	Identification cards, official credentials, and passports

Creating A File Path



How Long it Took

- 4 months
- Working 3 days a week

Summary Statistics

	Count	# of Duplicates	# of datasets	# of TIFS	% of PDF	# of broken URLS
# of records Not in the NTL database	4,506	14	0	0	100%	50
# of records located in the NTL database	350	11	0	0	100%	2
# of records located in the NTL database but the URL does not work	50	1	0	0	0	50
Not enough metadata to conduct a search	66	0	0	0	0	0

Thank You!

Contact: Jarivera@uw.edu