

# Reinventing the Cataloging Hold

## Smarter Processing Through Data Integration

Nina Servizzi

Associate Dean, Knowledge Access & Resource Management Services

Potomac Technical Processing Librarians 94th Annual Meeting  
Washington D.C. , October 12 , 2018

# Introduction

- I. Assessment in Technical Services
- II. Data Integration
- III. Cataloging Hold: Backlog
- IV. Cataloging Hold: Workflow Redesign
- V. Cataloging Hold: Implementation
- VI. Next Steps

# Assessment in Technical Services

“Become a data-aware organization that promotes open Exploration of data and grounds our decisions in evidence.”

NYU Libraries, *Strategic Plan 2013-2017*, Goal 8

# Assessment in Technical Services

What did this mean for Technical Services?

Challenge to move beyond simply reporting statistics and begin (re)designing workflows and processes that are truly data-driven.

# Data Integration

Data needed for workflow development comes from many different sources and is not readily available to staff

- ILS (Aleph Oracle tables)
- Discovery System (Primo)
- Open Link Resolver (SFX)
- Accounting System (FAME)
- University Registration ( Student/Employee data)
- OCLC (MARC, Kbart)
- Vendor Systems (GOBI, Oasis)
- Database/Publisher Platforms (ProQuest, Springer, EBC)
- COUNTER statistics



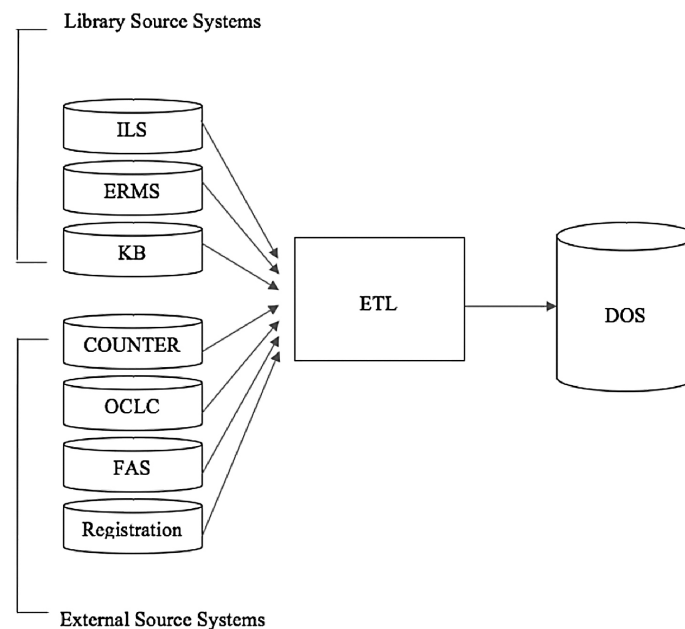
# Data Integration

Heterogeneous data must be normalized and pulled together in a centralized repository

Data within the repository must then be easily accessible to staff with varying levels of querying skills

# Data Integration

Solution: Build a data warehouse for the Libraries



Source: Servizio, Nina. "Future of Technical Services: Data Governance and Analytics" in *Assessment for Technical Services*. Ed. Kimberly Edwards and Michelle Leonard. Forthcoming (Fall 2018) Chicago: Association for Library Collections and Technical Services.

# Data Integration

Libraries Data Warehouse (LDW)

Locally hosted MS SQL Server

- Developed using SQL Server Integration Services (SSIS)
  - Integrated development environment
  - Robust suite of ETL (Extract, Transform, and Load) tools
  - Cost effective (free with MS SQL Server license)





# Data Integration

Data access premises:

1. End users must be able to access and manipulate data themselves
2. Reporting tool must be able to analyze and visually present data in a wide variety of ways

Primary front end tool:  
Tableau, Excel

Backend tools:  
SQL Server Management Studio, ODBC, Toad

# Data Integration

Data access premises:

1. End users must be able to access and manipulate data themselves
2. Reporting tool must be able to analyze and visually present data in a wide variety of ways

Primary front end tool:  
Tableau, Excel

Backend tools:  
SQL Server Management Studio, ODBC, Toad

# Cataloging Hold: Backlog

2014/2015 Cataloging backlog was over 32,000 physical volumes

Quantitative analysis of receiving and cataloging statistics revealed:

- 87,785 total receipts for New York collections
- 30.35% (n=26,803) cataloged on receipt in Resource Management Department (RM)
- 59.12 % (n=51,900) cataloged in Metadata Production & Management Department (MP&M)
- 10.53% (n=9,082) remained un-cataloged

# Cataloging Hold: Backlog

Current workflows not keeping up with rate of collection growth

Qualitative analysis revealed other inefficiencies:

- FIFO workflows resulted in materials be handled by multiple staff before ending in backlog
- Original catalogers processing copy 'on encounter' resulted in low rate of original cataloging
- Maintaining hold in system number order for RUSH processing required 1 FTE ; large increased RUSH requests
- Uncertainty about cataloging policy caused staff to set materials 'aside'

# Cataloging Hold: Backlog

Administrative numbers revealed deeper systemic problems:

- There was a \$ 300,000 over expenditure in adjunct cataloging budget (20 adjuncts on staff = 10 FTE)
- E-Resource purchases made up 77.64% of collections budget
- Less than 1.5 FTE working on e-resources resulted in backlog of estimated 1,000,000 un-cataloged e-books

# Cataloging Hold: Workflow Redesign

Redesign processing workflows between RM and MP&M to address inefficiencies particularly:

- Increase cataloging capacity to meet collection growth
- Separate copy/original cataloging streams, eliminate parallel processing streams
- Clarify cataloging policy
- Streamline hold management and reassign ownership from MP&M to RM
- Reduce hold and minimize RUSH processing requests
- Increase staff available for e-resource work and reduce reliance on adjunct catalogers

# Cataloging Hold: Workflow Redesign

Spring 2015 – Fall 2016 successful proof of concept which reduced backlog from over 32,000 to under 13,000

- Pilot workflow used an internally-developed algorithm to check volumes in the hold against OCLC WorldCat
- If a full bibliographic record was found for the title, the record was automatically downloaded and overlaid in our ILS
- Staff were then provided with pull lists and the matched bibliographic records checked against physical item.
- Correctly matched items were sent to shelf
- False matches were analyzed and match algorithm updated

# Cataloging Hold: Workflow Redesign

To move from proof of concept to a fully automated and well articulated workflow took about two years to fully implement workflow required changes in two areas:

## 1. Resource Allocation

- Created a Metadata Policy Committee (MPC)
- Establish new policy on rapid cataloging based on analysis of match algorithm results
- Write new documentation and train staff
- Reassign staff to well defined functional areas
- Reduce adjunct catalogers through natural attrition



# Cataloging Hold: Workflow Redesign

## 2. Automation Capacity

- Build tables in the LDW that track data for un-cataloged materials that can be used for processing triage (updated hourly)
  - Material type (DVD, CD, book, etc.)
  - Holding location (Library/sublibrary)
  - Bibliographic quality (cataloging entity, encoding level, call number type, subject headings, etc.)
  - Shelving location in processing department
- Refine local match algorithm for each type of resource
- Build interactive dashboards for staff to access data in LDW

# Cataloging Hold: Implementation

Implementation of new workflows has been gradual process to allow staff to understand the context of the changes and the role they play in aligning processing with the Libraries strategic goals.

- Change in metadata policy announced
- New staff assignments introduced
- Introduction of new workflows
- Training becoming comfortable with new tools
  - GitHub pull requests, batch record enhancements, LDW dashboards, inventory systems, Jira ticketing, etc.

Throughout implementation staff feedback and close monitoring has been used to drive iterative development cycles for or LDW and automation scripts

# Cataloging Hold: Implementation

Automated hold maintenance in full production July 2018 has yielded a 107% increase in volumes identified as having full-copy available and moved out of hold.

## By Unit Monthly Feb-Oct 18

Z30 Cataloger (group)	Month of Z30 Update Date									Grand Total
	February 2018	March 2018	April 2018	May 2018	June 2018	July 2018	August 2018	September 2018	October 2018	
KADD	98	279	258	320	209	480	742	384	85	2,855
MP&M Admin/Faculty	1,230	1,049	1,003	970	1,262	1,332	1,112	1,073	431	9,462
MP&M Staff	1,872	1,620	2,173	2,125	2,229	2,018	2,436	1,908	726	17,107
MP&M Students	557	1,032	1,005	603	443	111	169	254	123	4,297
RM Monographs Unit Staff	342	362	405	441	525	258	520	469	93	3,415
RM Monographs Unit Students	891	1,534	1,179	1,484	2,419	2,398	2,192	1,309	528	13,934
RM MP&GL Staff	745	752	867	586	772	1,427	1,641	1,553	465	8,808
RM MP&GL Students			137		73	167				377
Grand Total	5,735	6,628	7,027	6,529	7,932	8,191	8,812	6,950	2,451	60,255



# Cataloging Hold: Implementation

Our numbers for 2017/2018:

- 76,564 total receipt of physical volumes
- 98.99% (n=76,564) of total receipts cataloged
- 38.94% ( n=30,218) cataloged on receipt
- 111.15 % (n=7042) increase in Original cataloging compared to ( n=3335) in 2015/2016

# Cataloging Hold: Implementation

Other benefits:

- Collections growth stabilized around 75,000 volumes per year
- No over expenditure in adjunct cataloging budget; adjuncts have been reduced to 9
- Increase in number of staff working on e-resources to 5 FTE
- Over 1,200,000 e-books cataloged; backlog eliminated

# Cataloging Hold: Next Steps

Use hold dashboard to create identify specific original cataloging projects and identify gaps in our cataloging capacity:

tableau public		GALLERY	AUTHORS	BLOG	RESOURCES	ACTIV
----------------	--	---------	---------	------	-----------	-------

[Reed - Profile](#)☆ Favorite

**Cataloging backlog** Number of results: 1,981

Shelf + BSN	Location	Material	Language	Title	Barcode	
084-000012192	BOBST MAIN	BOOK	tur	Özbek destanları	31142056413761	●
					31142056413779	●
000-003154841	BOBST MAIN	BOOK	tur	Cumhuriyet Türkiye'si'nde bir mesele ol..	31142056415774	●
000-003222142	BOBST MAIN	BOOK	tur	LUGAT-I EBÜZZIYA YAZARINDAN IM..	31142046395433	×
000-003222147	BOBST MAIN	BOOK	tur	Yunandan evvelki Türk medeniyeti	31142046395516	×
000-003222149	BOBST MAIN	BOOK	tur	VERD ÜL HADAIK FREIBURG BASKISI	31142046395417	×
000-003222162	BOBST MAIN	BOOK	tur	YENİ TÜRKÇE LÜGAT	31142046395508	×
000-003222192	BOBST MAIN	BOOK	tur	ŞEHZADE CİHANGİR KİTABIN CILDI..	31142046395565	×
000-003244521	BOBST MAIN	BOOK	tur	Uluslararası Üsküdar Sempozyumu V, ..	31142049162087	×
000-003360557	BOBST MAIN	BOOK	tur	Asitân-i ta'rihde	31142049437703	×
000-003669781	BOBST MAIN	BOOK	tur	Rize defteri	31142059860935	●
169-003731384	BOBST MAIN	BOOK	tur	İslam Hukuku Elkitabı	31142053116219	×
169-003731471	BOBST MAIN	BOOK	tur	Hidir Abdal Sultan	31142053116268	×
008-003735824	BOBST MAIN	BOOK	tur	Atilla Jorma armağanı	31142053244797	◆
169-003744351	BOBST MAIN	BOOK	tur	Ölüme Götüren Yazılar: Sabahattin Ali, ..	31142053426253	×
008-003744504	BOBST MAIN	BOOK	tur	50 Yıllık Sır: Genelkurmay Başkanı Erd..	31142052646166	×
000-003751525	BOBST MAIN	BOOK	tur	Pasta özel sayısı	3751525-10	×
010-003771777	BOBST MAIN	BOOK	tur	Kühû	31142055418357	×
011-003859863	BOBST MAIN	BOOK	tur	Risâle-i nur külliyyatı müellifi Bediüzzam..	31142055309598	×
195-003886473	BOBST MAIN	BOOK	tur	Arşiv belgeleri ışığında	31142059876451	×
					31142059876469	×
000-003920830	BOBST MAIN	BOOK	tur	II. Karadeniz Uluslararası Sempozyumu	31142056304564	×
000-003920841	BOBST MAIN	BOOK	tur	Kendini öğrenemeyen toplum	31142056288395	×
000-003920887	BOBST MAIN	BOOK	tur	Silâhü'l-mü'minin	31142056305553	×
000-003920930	BOBST MAIN	BOOK	tur	Yusufeli yedinci yerini ararken	31142056304366	×
000-003920932	BOBST MAIN	BOOK	tur	Seçkinler kitabı	31142056305561	×
000-003965317	BOBST MAIN	BOOK	tur	Osmanlı mimarisinde I. Murad - Çelebi ..	31142056300828	×
000-003965320	BOBST MAIN	BOOK	tur	Türk dünyası coğrafyasında tespit edil..	31142056300844	×
000-003965338	BOBST MAIN	BOOK	tur	Bursa tarih albümü	31142055758927	×

**Material**  
☒ (All)  
☒ BOOK  
☒ ISSUE  
☒ null  
☒ VIDEO

**Location**  
☒ (All)  
☒ BAFC AFC  
☒ BOBST MAIN  
☒ NABUD GLSH  
☒ NABUD MAIN  
☒ NABUD SPCL

**Language** 🔍 ⌵  
☐ ota  
☐ per  
☐ pol  
☐ por  
☐ pus  
☐ rum  
☐ rus  
☐ san  
☐ scc  
☐ sil  
☐ slo  
☐ smi  
☐ spa  
☐ srp  
☐ swe  
☐ syc  
☐ tsk  
☐ tur  
☒ tur  
☐ tut  
☐ uig  
☐ ukr  
☐ und  
☐ urd  
☐ us  
☐ xx

**Call No**  
☒ (All)  
☒ No  
☒ Yes

**Lit. Call No**  
☒ (All)  
☒ No  
☒ Unknown  
☒ Yes

**LCSH**  
☒ (All)  
☒ No  
☒ No, but Lit Call No  
☒ Yes

**Source**  
☒ (All)  
☒ DLC  
☒ Other  
☒ PCC

**Barcode**



NYU | LIBRARIES

# Questions?

Contact:

Nina Servizzi

[Nina.Servizzi@nyu.edu](mailto:Nina.Servizzi@nyu.edu)