Metadata Madness: Overcoming Obstacles to Launch a Library Platform and Discovery Layer

ALA Mid-Winter TSWEIG Meeting
January 23, 2017 1:00PM-2:30PM

Marilyn White-E-Resource Librarian

Briget Wynne-Interlibrary Loan/Reference Librarian

NIST National Institute of Standards and Technology • U.S. Department of Commerce

iso INFORMATION SERVICES OFFICE

National Institute of Standards and Technology

- Non-regulatory federal agency made up of about 3,000 science and technology researchers
- NIST promotes U.S. innovation and industrial competitiveness by advancing measurement science, standards, and technology
- The Research Library is located within the Information Services Office (ISO) on the Gaithersburg, MD campus
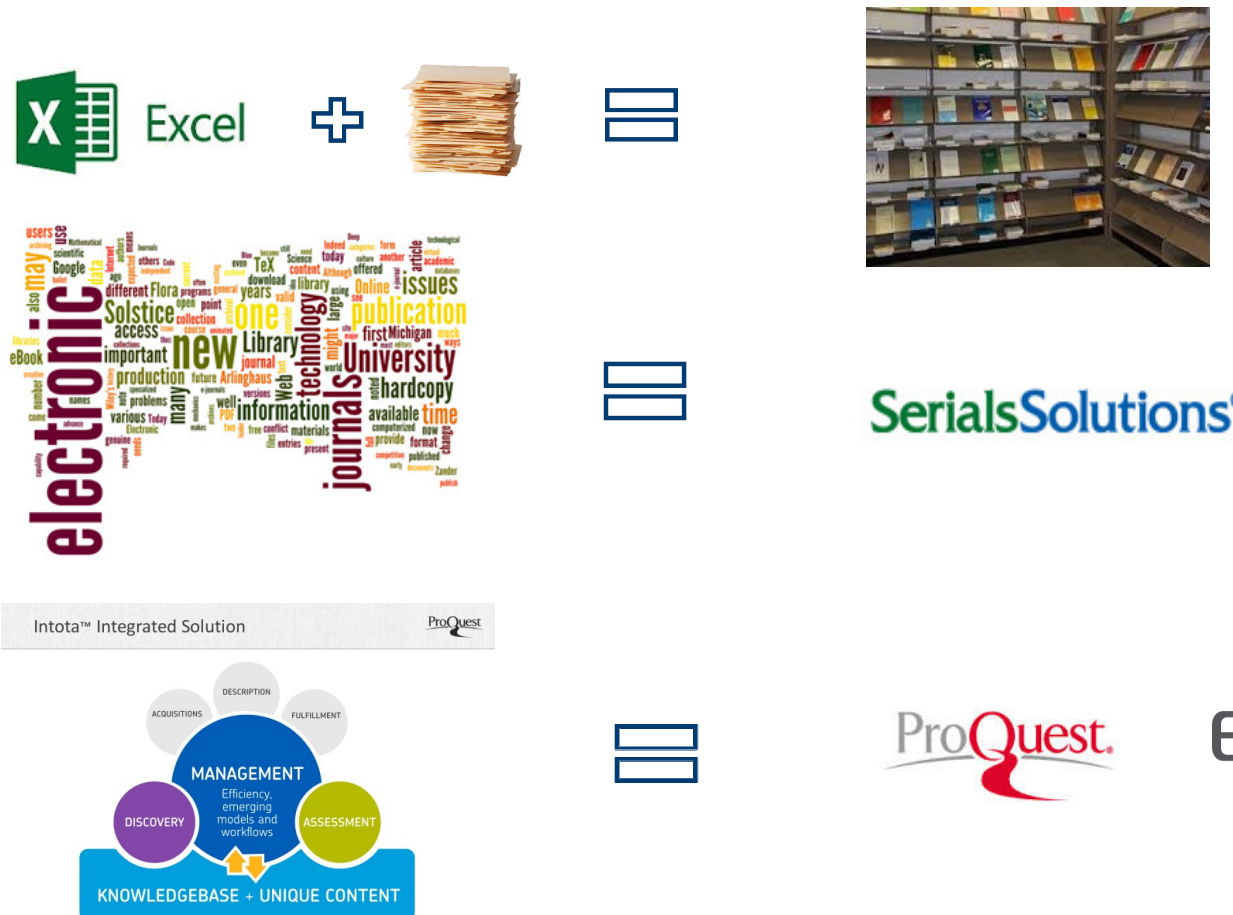
About the NIST Research Library

- Staffing

- Resources

- Federal environment

- Systems that are self-sustaining

Agenda

- Background on our Tech Services Systems
- Why we did what we did!
  - Improving user experience and library efficiency
  - Implementation and impact on tech services
- Challenges we faced and solutions
- Lessons learned

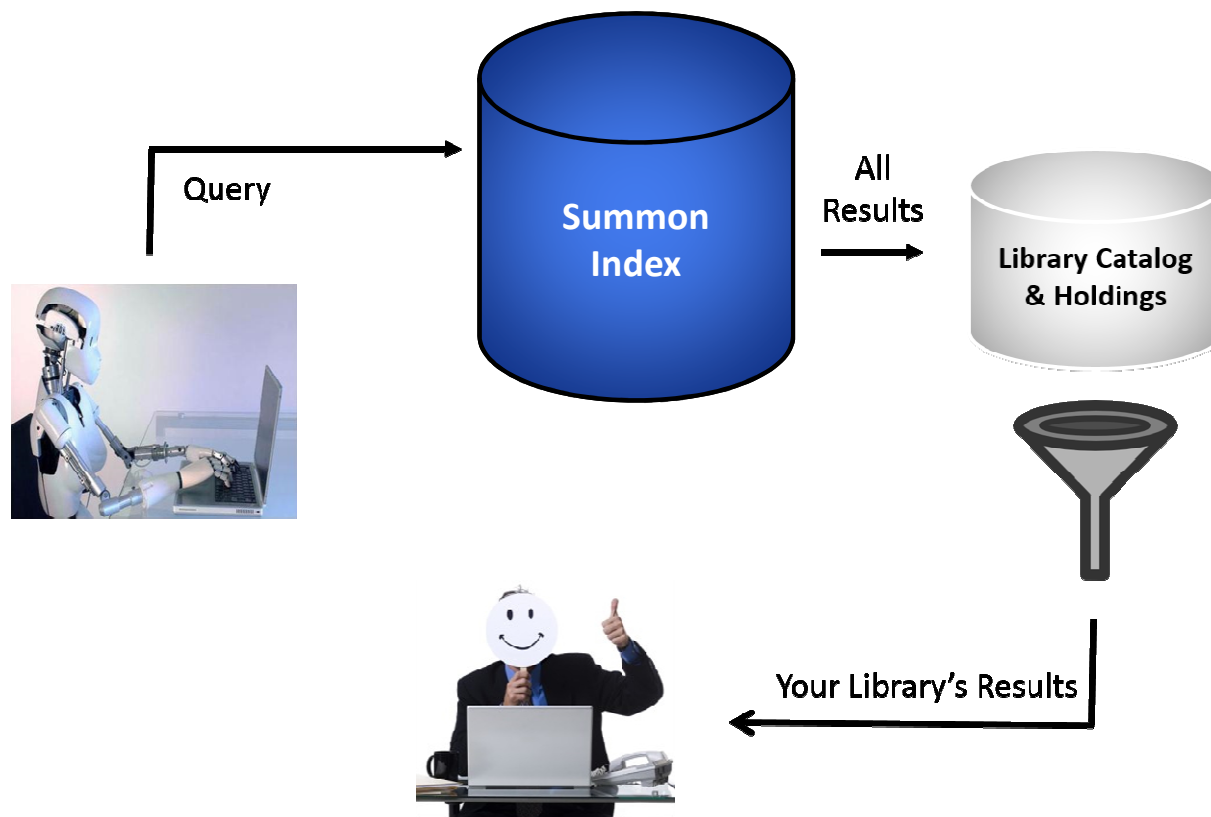# Evolution of E-Resource Management

What Does the Summon Index Search?

SUMMON INDEX

Online catalog

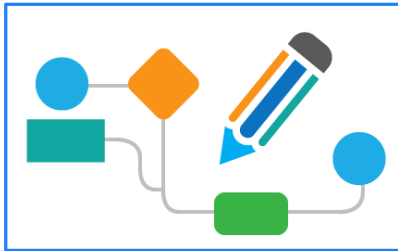ERMS (holdings information)

Journal subscriptions and databases

How Does a Search in Summon Work?

Training Journey

Plan & document workflows

Fix the database

ERMS Clean Up

Broken links
Inaccuracies
Publisher changes
Mysteries

NIST Single Search: SUMMON and NIST Content

	NIST Authors in Proprietary Journals	NIST Authors in JRes	NIST Technical Publications	Photo Collection	Oral Histories
Metadata source	Proprietary databases	CrossRef	CrossRef, govinfo, Sirsi, Editorial Review System, Google Scholar	NDA	Sirsi
Full text source	PMC, proprietary databases	PMC	Govinfo, NIST publication server	NDA	
Proquest options	Summon Index/PMC	Summon Index/PMC	Sirsi	NDA	Sirsi
Our preferred option	Summon Index	PMC	Sirsi	NDA	Sirsi

Summon MARC21 Mapping

	Summon Field Name	Summon Display Name	Display Location	Facet	Corresponding MARC Fields	Changes/Corrections?	Notes
1	**Summon MARC21 Mapping**						
2	Summon Field Name	Summon Display Name	Display Location	Facet	Corresponding MARC Fields	Changes/Corrections?	Notes
3	Abstract	Abstract	Brief and Preview		520 ab		
4	Author	Author	Facet, Brief, and	Author	100 ab$c; 700 ab$c		
5	Corporate Author	Corporate Author	Brief and Preview		110$a $b $c $n, 710$a $b $c $n		
6	Document Title : Document Subtitle	Maps to Title field	Preview		245 $a, 245 $b		
7	Edition	Maps to Edition field	Preview		250 $a		
8	Genre	Genre	Facet, Preview	Genre	600 $v; 610 $v; 611 $v; 630 $v; 648 $v; 650 $v; 651 $v; 655 av, 380 $a		
9	ISBN	ISBN	Brief and Preview		020 $a		
10	ISSN	ISSN	Brief and Preview		022 $a		
11	LCCallNum	Call Num	Brief and Preview		099ab; 090ab; 050ab		
12	PageCount	Pages	Brief and Preview for Content Type = Book only		300 $a		
13	PublicationSeriesTitle	Series	Preview		440 $a; 490 $a;		
14	Publisher	Publisher	Preview		260 $b, 264--1 $b (2nd indicator=1)		
15	URL	Hyperlinked to title	n/a		856 $u		
16	Volume (of series)	Volume	Brief and Preview		440 $v; 490 $v; 830 $v		
17	ContentType	See Content Type Mapping wo	Facet, Brief, and	Content Type	See Content Type Mapping worksheet		
18	Language	Language	Facet/Preview	Language	008 (35-37); 041abdefg$j		
19	Library Location	Library Locations provided by	Facet	Library Location	Provided prior to initial ingest of records.		
20	PublicationDate	Date	Facet, Brief and Preview	Publication Date	260 $c, 264--1 $c (2nd indicator=1), 264--4 $c (2nd indicator=4)		
21	Subject terms	Subject	Facet, Brief and Preview	Subject Terms	600 $x; 610 $x; 611 $x; 630 $x; 648$x; 650 abcd$x; 651 $x; 653 $a		
22	AuthorAffiliation	n/a	Index Only		100 $u; 110 $u; 111 $u; 600 $u; 610 $u; 611 $u; 700 $u; 710 $u; 711 $u		
23	CODEN	n/a	Index Only		030 $a		
24	Deleted Record Field	n/a	Index Only		Only for Option 1 delete method. Provided prior to initial ingest of		
25	Dewey Decimal Number	n/a	Index Only unless enabled in Admin Console		082 $a		
26	DocumentTitleAlternate	n/a	Index Only		130adfklmnopr$s; 210 ab; 222 ab; 240adfklmnopr$s; 242 abnp; 246 abfinp; 600$t; 610 $t; 611 $t; 630 adfhklnopr$s; 730 adfklmnopr$s, 740 an$p		
27	GeographicLocations	n/a	Facet	Region	600 $z;610 $z;611 $z;630 $z;648 $z; 650$z; 651 $a; 655$z		
28	InternationalArticleNumber (EAN)	n/a	Index Only		024-3$a (1st indicator=3)		
29	LCCN	n/a	Index Only		010 $a		
30	Meeting Name	n/a	Brief, Preview for Conference Proceeding content type; otherwise Index Only		111 ancdeq$p; 611 ancdeq$p; 711 ancdeq$p; 811 ancdeq$p.		
31	Notes	n/a	Index Only		500 $a, 502 abcdgo, 362 $a $z; 506 abcd$e; 510 ab$c; 511 $a; 518 $a; 530 abcd; 533 abcdefmn; 534 pabcefklmn$t; 538 ai; 583		
32	OCLC number	n/a	Index Only		035 $a		
33	PublicationPlace	n/a	Index, Catalog Detail Page		260 $a, 264--1 $a (2nd indicator=1)		

Summon Content Type Mapping

Default linking for all content types except eBook, eJournal, and Journal is to the OPAC record. For eBook, eJournal, and Journal, default linking uses the 856$u URL in the record if there is only one, otherwise reverts to the OPAC record. Please add linking rule change requests in the Notes column.

	Content Type	Default MARC Field Mapping	Exclude from Mapping	Additional Mapping	Notes
7	Archival Material	Leader(06) = p			
8	Audio Recording	007(00) = s OR Leader(06) = i OR Leader(06) = j			
9	Book	Leader(06-07) = aa OR Leader(06-07) = ab OR Leader(06-07) = ac OR Leader(06-07) = ad OR Leader(06-07) = am			
10	Computer File	008(26) = a, b, c, d, e, f, g, m OR z OR Leader(06) = m			
11	Conference Proceeding	008(29) = 1			
12	eBook	Leader(06-07) = am AND 007(00-01) = cr OR Leader(06-07) = am AND 008(23) = s *OR* o			
13	eJournal	Leader(06-07) = as AND 007(00-01) = cr OR Leader(06-07) = as AND 008(23) = s *OR* o			
14	Government Document	008(28) = a, c, f, i, l, m, o, s, *OR* z AND Leader(06) = a AND 26x$b does not contain 'University or Press'			
15	Image	Leader(06) = k			
16	Journal	Leader(06-07) = as			
17	Kit	007(00) = o OR Leader(06) = o			
18	Manuscript	Leader(06) = t			
19	Map	007(00) = a OR Leader(06) = e OR Leader(06) = f			

Part 1
Implementation from the Back End:
Issues Pulling Catalog Metadata

- Uploading the Sirsi catalog to the Summon index
  - Creating a process to export catalog records for Summon index
  - Process used for the initial load and for pulling smaller batches of data for fixes
- Using API's to pull catalog metadata
  - Thought it would be the least disruptive
  - Turned out to be much more complex than anticipated

  Title control # | Sirsi catalog key | Marc dump

Symphony Export Utility

Defaults

☑ Include 999 holdings tag

☐ Include local and junk tags

☐ Export Symphony catalog key to MARC tag   001 ▾

Format of the output file: ◉ Marc ○ Flat

○ Export all records to one file ◉ Include   10   records in each file

○ View all records ◉ Include 10   records in each file

Part 2
Implementation from the Back End:
Problems We Faced

- NIST Pubs not displaying as government documents
- Items not linking to electronic versions
- Identifying where the fix needed to be
- Wrong match points between Sirsi catalog and Summon
- Preparing records for Marc 360 Updates

Part 2: Metadata Fixes We Made to our NIST Technical Publications MARC records

- Added local 590 field: \\$aNISTPUB

- Deleted duplicate 035, 856, and 956 fields

- Edited our DOI urls

- Used Regular Expressions to edit Leader and 008 fields

- Added alternate search terms

Tools Used

- Sirsi API
- Sirsi Utilities
- MarcEdit
- Some good tools and resources on using Marc Edit and regular expressions
  - Regular Expressions in MarcEdit PowerPoint Presentation
    - Regular Expressions in MarcEdit. Terry Reese Gray Family Chair for Innovative Library Resources. Topics. MarcEdit Regular Expression Support Information.
  - http://blog.reeset.net/
  - https://www.youtube.com/watch?v=qNJA9p2_-qU_Terry Reese (YouTube videos and tutorials)

Add/Delete Using MarcEdit:
Adding Local Field and Removing Duplicate Fields

Adding 590 NISTPUB Field:

Delete Duplicate Fields:

Editing DOIs

- Need to have the CrossRef DOI in the 856 field

- Use the Edit Fields Utility.

- Select 856 and do a find and replace
  - Find in records: http://dx.doi.org
  - Replace with: https://doi.org

Replacing Leader Fields:
Correcting Our Item Types in Summon
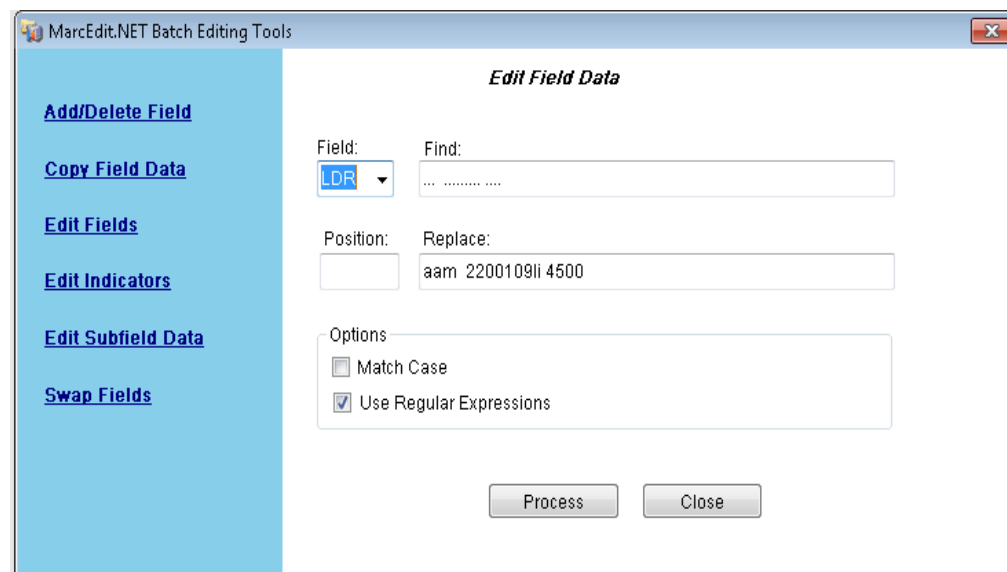
The regular expression for the LDR field is:

(... ....)

Replace in regex with:
        aam 2200109li 4500

(Regex formula is: 3 periods, 2 spaces, 9 periods, 1 space, 4 periods)

Editing 008 Fields:
Correcting Our Item Types in Summon

File   Edit   Fonts   Reports   Tools   OCLC WorldCat   Plug-ins   Help

=LDR 00953aam 2200409li 4500
=001 o5833789
=008 791207s9999\\\\mdu\\\\\ot\\\f000\0\eng\d
=010 \\$a   79607159
=035 \\$a(Sirsi) o5833789
=040 \\$aDLC$cDLC$dGPO$dNBS
=049 \\$aNBSN
=099 \\$a5G40$a.W47$a1980
=100 1\$aWestley, Francis.
=245 10$aTable of recommended rate constants for chemical reactions occurring in combus
=260 \\$aWashington :$bU.S. Dept. of Commerce, National Bureau of Standards : for sale b
=300 \\$aiv, 110 ;$c26 cm.
=440 \0$aNSRDS-NBS ;$v67
=500 \\$a"National standard reference data system."
=504 \\$aBibliography: p. 109.
=590 \\$aNISTPUB
=590 \\$a8008090
=590 \\$aANALYTIC - Shelved in NBS Pubs Reference.
=596 \\$a1
=650 \0$aChemical kinetics$vTables.
=650 \0$aCombustion$vTables.
=598 \\$aNIST_NSRDS

=LDR 01923aam 2200409li 4500
=001 \\|a(OCoLC)950543999
=003 SIRSI
=005 20160525040643.0
=008 160525s9999\\\\mdu\\\\\ot\\\f000\0\eng\d
=035 \\$a(Sirsi) o950543999
=040 \\$aNBS$erda$cNBS
=035 \\$a(OCoLC)950543999
=090 \\$aQC100$b.U573 no. 164 2014
=049 \\$aNBSS

< << > >   Page 1 - 2

180 records processed in 0.093752 seconds.

MARC Tools

File   Tools   Help

Input File:
C:\

Output File:
C:\

Execute

Edit Records

Close

Functions

MarcBreaker
MarcMaker
MARC21=>MARC21XML
MARC21XML=>MARC21
Dublin Core => MARC
EAD=>HTML
EAD=>MARC
FGDC=>MARC
MARC=>Dublin Core
MARC=>FGDC
MARC=>MODS

☐ Translate to MARC-8
☐ Translate to UTF8

Results:

Adding Alternate Search Terms

Illustration options	Series name	830 field series names and alternates	Cutter numbers	Series abbreviations for DOI
illustrations (color)	NIST advanced manufacturing series	NIST advanced manufacturing series; NIST AMS;	no.U267	AMS
illustrations (black and white)	NIST economic analysis brief	NIST economic analysis brief; NIST EAB;	no.U333	EAB
	NIST federal information processing standards publication	NIST federal information processing standards publication; NIST FIPS;	(JK468) no.A8 A3	FIPS
	NIST GCR	NIST GCR; NIST grant-contract report; NIST grant contract report;	no.U6 N25	GCR
	NIST handbook	NIST handbook; NIST HB;	no.U51	HB
	NIST monograph	NIST monograph; NIST MONO;	no.U556	MONO
	NIST NCSTAR	NIST NCSTAR; National Construction Safety Team Act report;	(TH443) no.N35	NCSTAR
	NIST special publication	NIST special publication; NIST special pub; NIST SP;	no.U57	SP
	NIST technical note	NIST technical note; NIST tech note; NIST TN;	no.U5753	TN
	NISTIR	NISTIR; NIST IR; NIST interagency report; NIST internal report;	no.U56	IR
	NSRDS-NIST	NSRDS-NIST; National standard reference data series;	no.U573	NSRDS
	Voluntary product standard	Voluntary product standard; PS;	no.U563	PS
	NBS report	NBS report; NBS RPT;	no.U843	RPT
	NBSIR	NBSIR; NBS IR; NBS interagency report; NBS internal report;	no.U56	IR
	NBS technical note	NBS technical note; NBS tech note; NBS TN;	no.U5753	TN
	NBS special publication	NBS special publication; NBS special pub; NBS SP;	no.U57	SP
	NBS miscellaneous publication	NBS miscellaneous publication; NBS miscellaneous pub; NBS MP;	no.U57	MP
	NBS monograph	NBS monograph; NBS MONO;	no.U556	MONO
	NBS handbook	NBS handbook; NBS HB;	no.U51	HB
	NBS GCR	NBS GCR; NBS grant-contract report; NBS grant contract report;	no.U6 N25	GCR
	NBS building science series	NBS building science series; NBS building science; NBS BSS;	no.U58	BSS
	NBS building and housing	NBS building and housing; NBS BH;	(TH1) no.U4	BH
	NBS commercial standards	NBS commercial standards; NBS CS;	no.U563	CS
	NBS commercial standards monthly	NBS commercial standards monthly; NBS CSM;	(HD62) no.U3	CSM
	NBS technical information on building materials	NBS technical information on building materials; technical information on building materials for use in the design of low-cost	(TH1) no.U5	TIBM
	NBS building materials and structures report	NBS building materials and structures report; NBS building materials and structures; NBS BMS;	(TA410) no.U48	BMS

Adding Alternate Search Terms:
Edit Subfield Utility

Serial Records Clean UP
Prepare for MARC 360 Updates

- The issue:
  - We want our electronic journals to be searchable at the title level in Summon results.
  - We have e-only journal records, print journal records and "mutt-records!!!"

Lessons Learned

- Work across the skillset of your staff:
  - Know what you need to get the job done
  - Get the right expertise
  - Communication
- Careful investment and selection of a system:
  - Foster good vendor relationships that will last long after implementation
  - Do your homework: future growth, compatibility and ROI
  - Maintenance and staff time: what is required?
- Prioritize what you want your system to do:
  - Is it possible to get your unique content indexed?
  - Bottom line: good user experience

Acknowledgements

- Special thanks to the other members of the Discovery Team:
    - Katelynd Bucher-Metadata Librarian
    - Paula Deutsch-Systems Administrator
- Thank you to NIST Library staff

Contact Information

- Katelynd Bucher-Metadata Librarian: katelynd.bucher@nist.gov

- Paula Deutsch-System Administrator (SirsiDynix ILS): paula.deutsch@nist.gov

- Marilyn White-E-Resource Librarian: (Intota/Summon Library Platform) marilyn.white@nist.gov

- Briget Wynne-ILL/Reference Librarian: briget.wynne@nist.gov