# Voyager's Global Data Change 101

tips and tricks for mass record changes

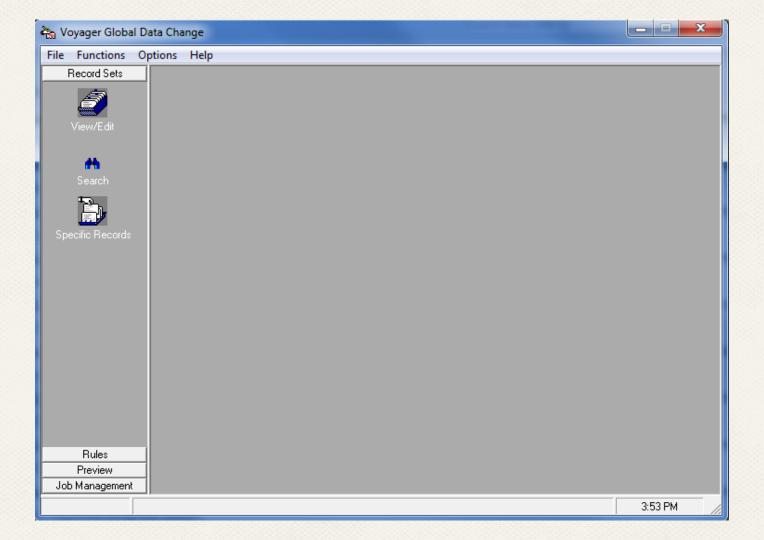
GIL Users Group Meeting 14 May 2015 Erin Grant Susan Wynne

### GDC 101 Agenda -

- → GDC Fundamentals
- → Step-By-Step Example GDC Job
- → Notes and Best Practices
- → Other Example Rules

#### What is GDC?

- → Voyager module
- → First available in Voyager 8
- → More than just data changes!
- → Now operational for USG libraries
  - Many thanks to Sean Purcell and Ex Libris staff



### **Terminology**

<u>Data Change Rule</u>: A condition and consequence for altering MARC records with a Data Change Job



<u>Data Change Rule Set</u>: A set of one or more Data Change Rules



<u>Data Change Rule Set Group</u>: A group of one or more Data Change Rule Sets



Record Set: A set of MARC records. Stored in database as a list of record IDs

<u>Scan Rule</u>: A condition and consequence for creating record sets with a Scan Job



Scan Rule Set: A set of one or more Scan Rules



Scan Rule Set Group: A group of one or more Scan Rule Sets

Rood, Amy. "Reworking Records: Streamlining MARC Record Changes and Workflows with Global Data Change." 2015 Ex Libris Technical Seminar. Minneapolis Marriott City Center, Minneapolis, MN. 5 May 2015. Presentation used with permission.

#### Scan jobs

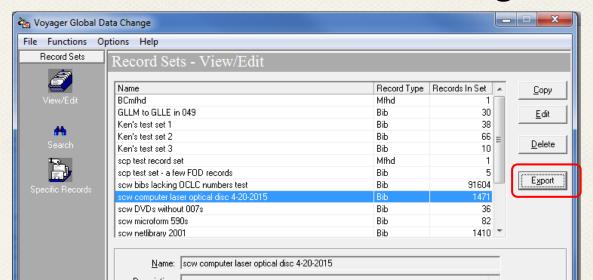
- → Scan entire database, a record set, or range of records
- → Define your own criteria, e.g., records with a 590 field
- → More granular than standard indexes
- → Include or exclude
- → Creates a record set
- → Caveat: Takes some time to scan entire database
  - U Works in bulk jobs of 1000 apiece

#### Match jobs

- → Generates 035 (003+001) from records in a file and matches with records in database
- → Creates record set of matched records, including multiple matches
- → Matches on the 035 only, can't change this
  - All multiple records (e.g. from different vendors)
     will be matched

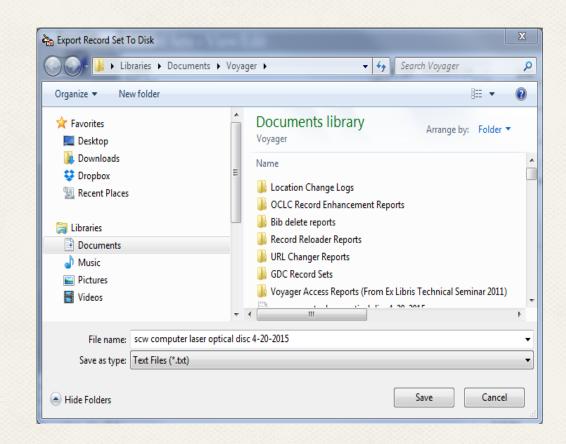
#### Export list of Record IDs from record sets

- → Can run scan job and export list of Record IDs
- → From Record Sets or View Running screens



### Export list of Record IDs from record sets

- $\rightarrow$  Default is .txt file
- → Put it wherever you want
- → Change the name if you want
- $\rightarrow$  Save



#### Make changes to batches of records

- → Bibliographic
- → Holdings
- → Authority

# What Kinds of Changes?

- → Add or Delete a field or subfield
- → Change a MARC tag
- → Replace a string of text (at a position, or replace string with string)
  - Can use Java regular expressions in Replace String with String
- → Suppress/unsuppress or delete records\*
  - Just the records you specify or all associated

<sup>\*</sup>Gary Strawn's Bib Delete or Location Changer may also be appropriate for suppress/unsuppress or delete projects

### Sorry, No Can Do

- → Change item records
  - We recommend Strawn's Location Changer
  - O Pick and Scan is another option
- → "Undo" button
  - O But you can re-import a "before" file of MARC records
- → Sort newly added fields in order defined by the library
  - MARC tag sorting functionality planned for 9.2

#### Susan's GDC Jobs

- → Add a basic 007 to DVD bibs lacking it
- → Delete an obsolete 590 note from a set of microform records
- → Replace string with a string in a subfield\*
  - (5) [computer file] GMD
  - omputer laser optical disc

<sup>\*</sup>as of 4/20/2015 I have done the computer terminology changes only on small subsets of the records needing these changes; there are still many instances in our database

#### Erin's GDC Jobs

- → Replace a string with a string in a subfield
  - Subject heading change for subject with a lot of different subfields that Global Headings Change wouldn't catch
- → Add a 500 note
  - (The example GDC job later on)

### **GDC Setup**

- → Install the GDC module
- → Check your Voyager.ini (configuration) file
  - Use Located in C:\Voyager
  - O Add stanza if needed (next slide)
- → Permissions (SysAdmin)
  - Create Global Data Change Profile(s)
  - Assign operator(s) to profile(s)

### Voyager.ini

[globaldatachange]

Server=xxxxx

Port=xxx15

Timeout=60

### Basic Steps for Data Change Jobs

- 1. Create a record set
- 2. Define one or more data change rules
- 3. Create a data change rule set group
  - → Even if you are only making one change at a time

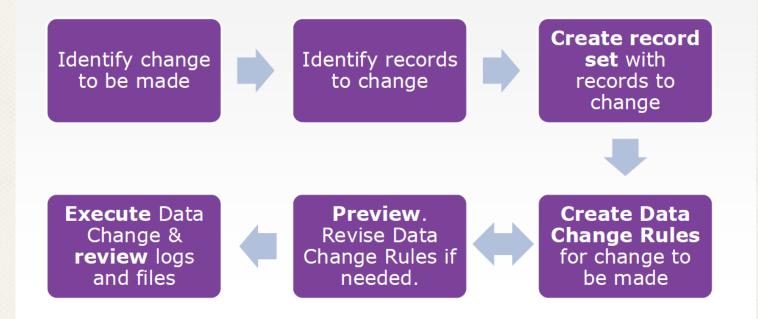
(continued...)

### Basic Steps for Data Change Jobs

(...continued)

- 4. PREVIEW
  - U Handy before-and-after views of records
- 5. Run job
- 6. Check changes

### Workflow





Rood, Amy. "Reworking Records: Streamlining MARC Record Changes and Workflows with Global Data Change." 2015 Ex Libris Technical Seminar. Minneapolis Marriott City Center, Minneapolis, MN. 5 May 2015. Presentation used with permission.

### Ways to Create a Record Set

- → Search
  - Just like in the cataloging client

- → Add specific records by record ID
  - One-by-one
  - O Previously prepared list (e.g., from Access Reports, VgerSelect, etc.) as a text file

### Ways to Create a Record Set

- → Scan job (takes some time to run)
  - So plan *what* and *when* you're going to do in advance if you Scan

GDC record sets will automatically dedupe records for you if you combine multiple sets of records or searches.

### GDC 101 Agenda

- → GDC Fundamentals
- → Step-By-Step Example GDC Job
- → Notes and Best Practices
- → Other Example Rules

### Step-By-Step Example GDC Job

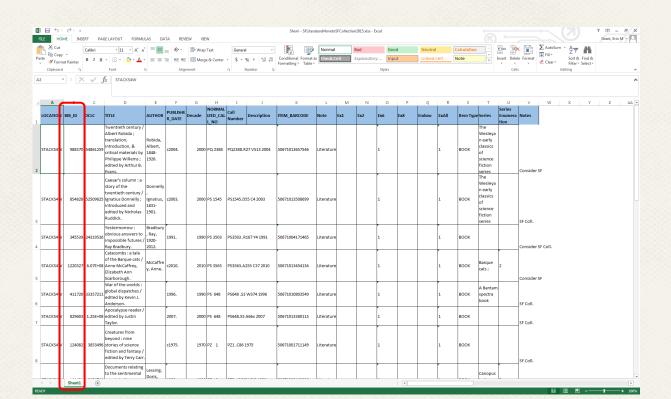
Desired change to records:

Adding a 500 note to ~170 bib records moved from the stacks to a separate Science Fiction collection.\*

\*I know, it should be a 590, but this was past practice. My plan is to flip all of these notes to 590s in the future all at one time.

## Identify Records to Change

Compile list of Bib IDs for records to change from spreadsheets and emails of items being moved from the stacks to a separate Sci Fi collection, sent from Access Services.



# Identify Records to Change

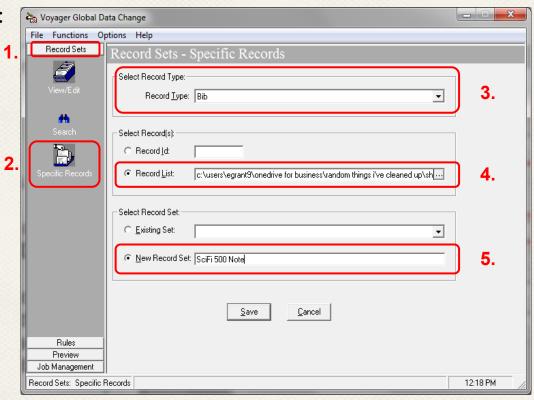
Copy and paste the Bib IDs into a .txt file



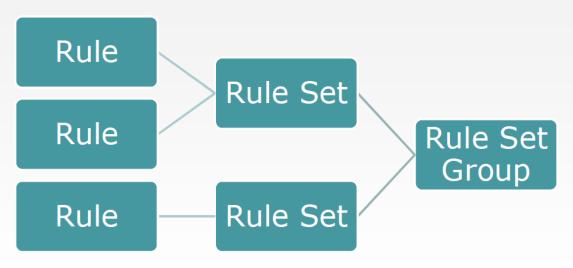
### Create a Record Set

Create a **Record Set** from a list of Bib IDs:

- Record Sets
- 2. Specific Records
- 3. Select your type of record (Bib/MFHD/ Auth)
- 4. **Select Record(s)**: choose the **Record List** button and browse to your .txt file of Bib IDs.
- Select Record Set: choose New Record Set and name this record set.



### Rule Structure



One or more rules per Rule Set
One or more Rule Sets per Rule Set Group
Rule Set Group = complete group of rules applied by Data Change Job



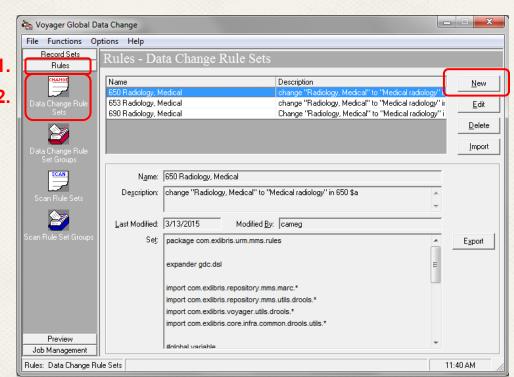
Rood, Amy. "Reworking Records: Streamlining MARC Record Changes and Workflows with Global Data Change." 2015 Ex Libris Technical Seminar. Minneapolis Marriott City Center, Minneapolis, MN. 5 May 2015. Presentation used with permission.

## Create Data Change Rule Set

Create a new Data Change Rule Set:

- 1. Rules
- 2. Data Change Rule Sets
- 3. New

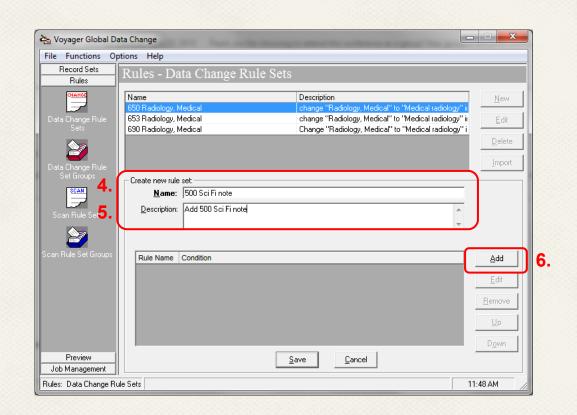
(continued...)



# Create Data Change Rule Set

(...continued)

- 4. Input Name
- 5. Input short **Description**
- 6. Click Add



### Add Conditions to Data Change Rule Set

Data Change Rule Set has been created.

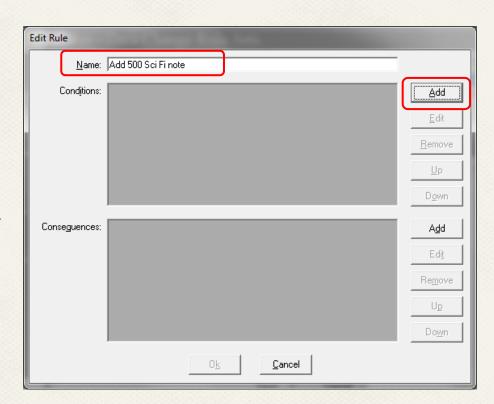
**Edit Rule** window opens - add rule(s) with conditions and consequences to your Data Change Rule Set:

#### 1. Input Name

I usually only add one Data Change Rule per Data Change Rule Set, so I name them the same.

#### 2. Conditions: Add

When you click **Add**, a new Rule Condition Template window will open.



### Add Conditions to Data Change Rule Set

**Rule Condition**: identifies what records will be changed (you'll set up the actual changes in the **Rule Consequences**).

(E.g., to make a change to an existing field:

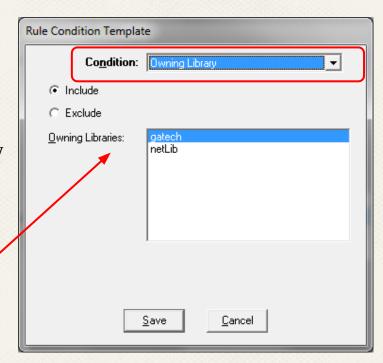
- → Create a Condition for the presence of that field
- → The actual change for that field goes in Consequences)

Here I'm adding a new field entirely, so my records won't already have the field I want to change. Thus I have to find something else to identify my records, which can be completely arbitrary.

So the condition is GA Tech as the owning library, which will include all of the records in the record set.

Other options to include all records:

 $\rightarrow$  Presence of a 245 or 008

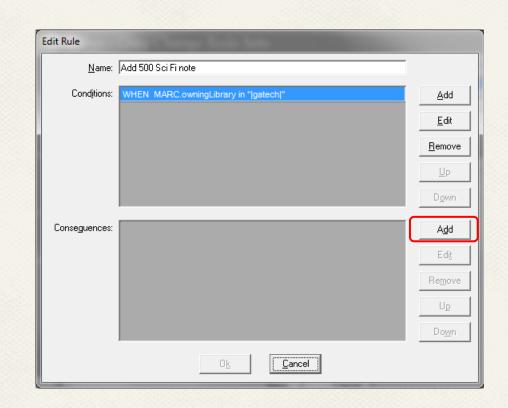


Once you **Save**, your Conditions have been added to your **Data Change Rule Set**.

That window closes, and you're back to the **Edit Rule** window.

Now add your **Consequences**.

Click **Add** to the right of **Consequences**, and a separate **Consequences** window will open.

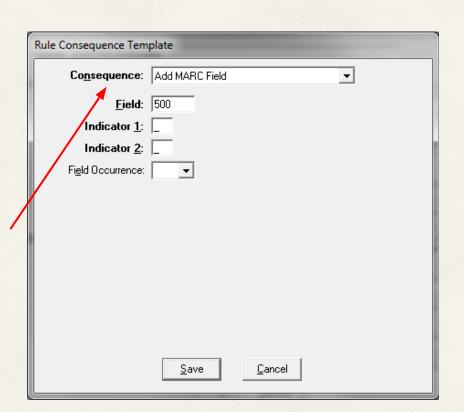


#### Consequences window:

- Contain input options for consequences (the actual changes) to your records
- Options vary with each type of consequence

I want to add a note in the 500 field, but before I can add the note, first I have to add the 500 field itself.

(Note: Field Occurrence is for adding multiple fields in a specific order - I left it blank.)

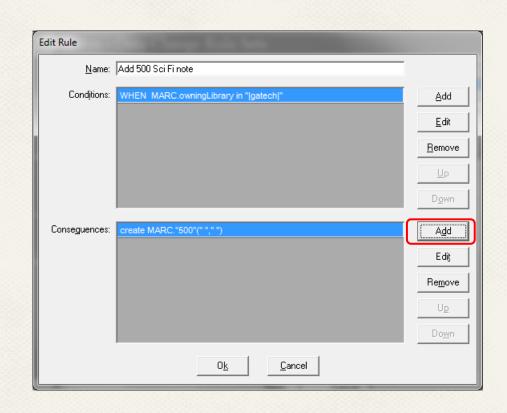


Your first Consequence has been added to your Data Change Rule Set (adding the 500 field).

That window closes, and you're back to the **Edit Rule** window again.

Click **Add** to the right of Consequences to add the next Consequence (to add the 500 \$a note).

A separate **Consequences** window will open again.

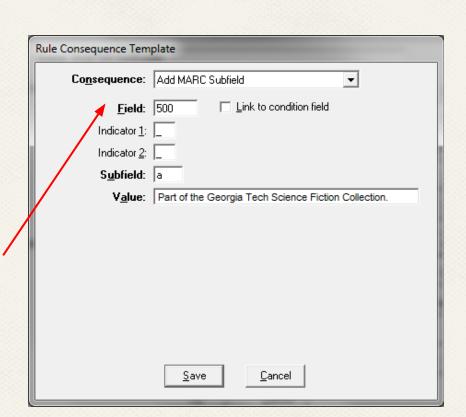


In the Consequences window, you'll have several options for consequences (changes) to your records again.

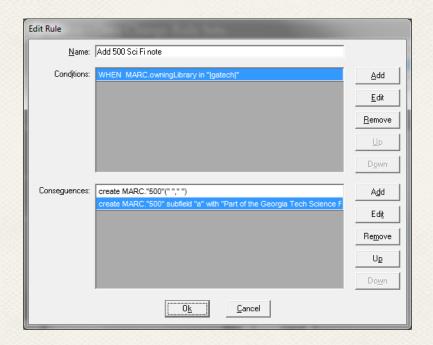
In the first Consequence I just added, the change will result in the addition of a new 500 field.

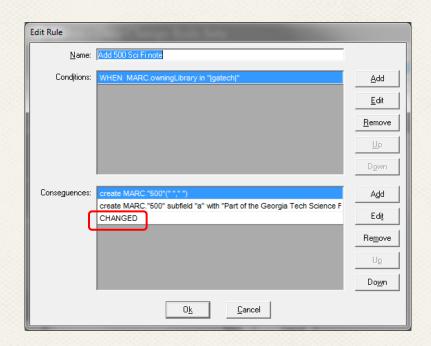
For this second Consequence, I want to add a note to \$a of that newly created 500 field.

(Note: Ignore Link to condition field - Ex Libris got rid of it in 9.1 because it confused everyone)



When you finish editing the rule and close the window, a final "CHANGED" Consequence is automatically added, which shows up as an Update in the History tab of any records that are changed.





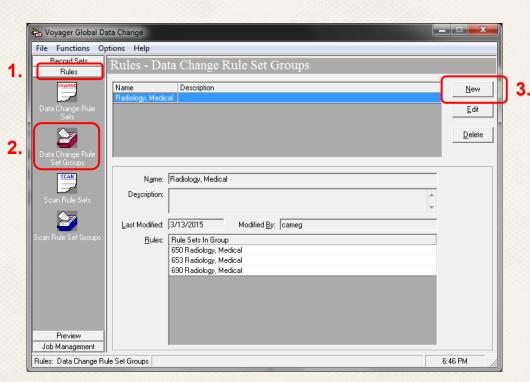
# Create Data Change Rule Set Group

Even for only one Data Change Rule Set, you still have to put it in a **Data Change Rule Set Group** before you can run the job.

Create a Data Change Rule Set Group:

- 1. Rules
- 2. Data Change Rule Set Groups
- 3. New

(continued...)



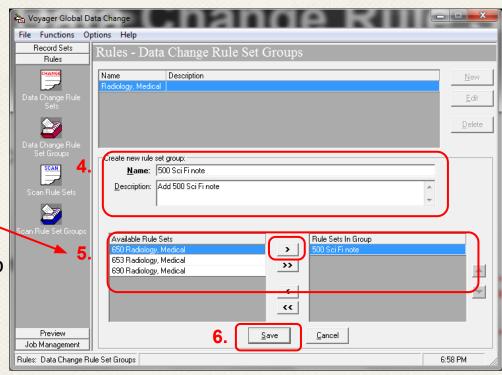
# Create Data Change Rule Set Group

### (...continued)

- 4. Input **Name** and short **Description**
- 5. The Data Change Rule Sets that you created will show up below on the left.

Highlight the Sets you want to use and click to add it to your Group

6. Save

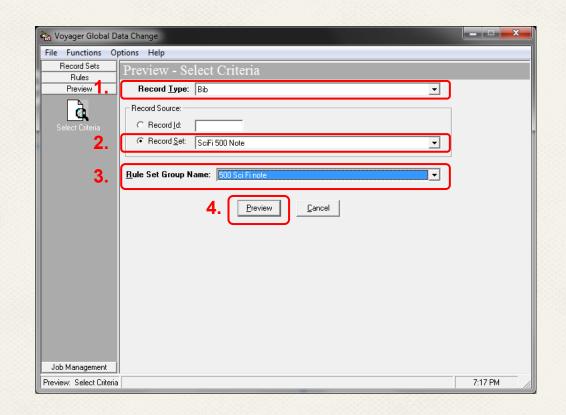


### **Preview Your Job**

Finally you can set up your Preview to see what your changes will look like.

- 1. Choose **Record Type**
- 2. Choose the **Record Set** you added
- 3. Choose the **Data Change Rule Set Group** you created
- 4. Preview

GDC process includes exporting records, running them through transform action(s), then bulk importing them back in; Preview is just those first 2 steps



### **Preview Your Job**

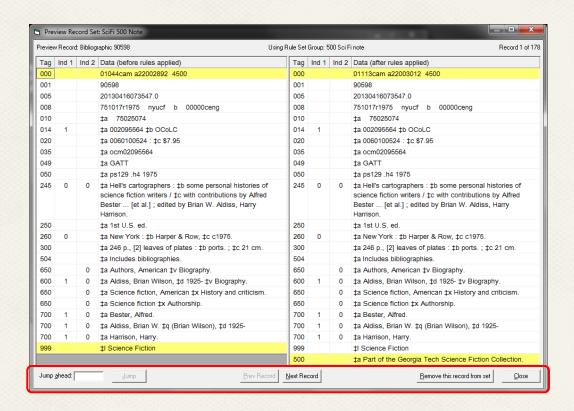
Current record is on the left, while the proposed changed record is on the right.

Changes will be highlighted in yellow

→ Record length changes, which is why leader (000) changes

#### You can:

- → Jump to a specific record number in the sequence
- Go to the next record in the sequence
- Remove records if the proposed change isn't right



### **Preview Your Job**

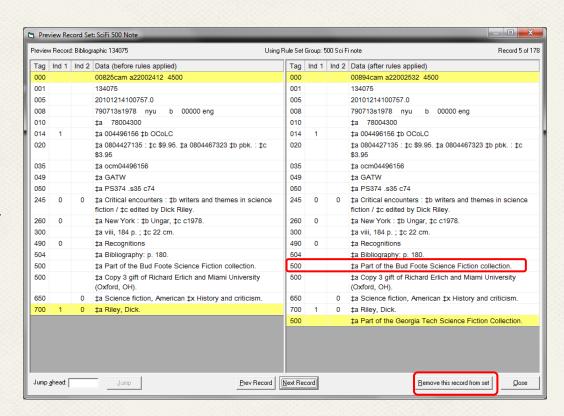
This record already has a 500 that would make this new 500 redundant, so I don't want the change.

Choose Remove this record from set

-----

Choose **Close** when you're done previewing the changes to your Record Set and removing any records you don't want changed.

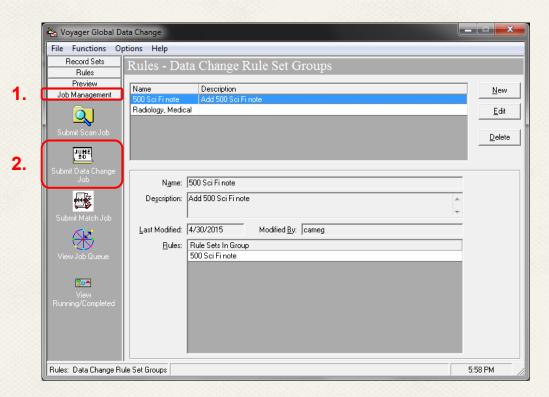
If you close prematurely, GDC will remember the records you removed when you reopen the Preview.



## **Running Your Job**

1. Job Management

2. Submit Data Change Job



## **Running Your Job**

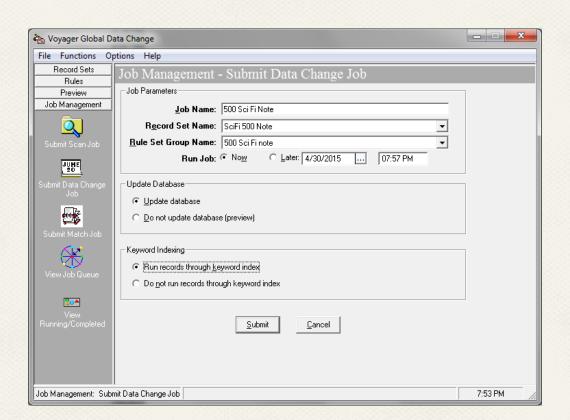
### Name your job

### Set when to run job

→ Immediately, or when department activity or server traffic is slower

#### Keyword Index option:

- → Choose to Run Records unless it's a really huge job
- → If you choose No, updated records will still be indexed under the old string and thus unfindable -- so be sure to run a Keyword Regen later on



## **Running Your Job**

#### **View Job Queue**

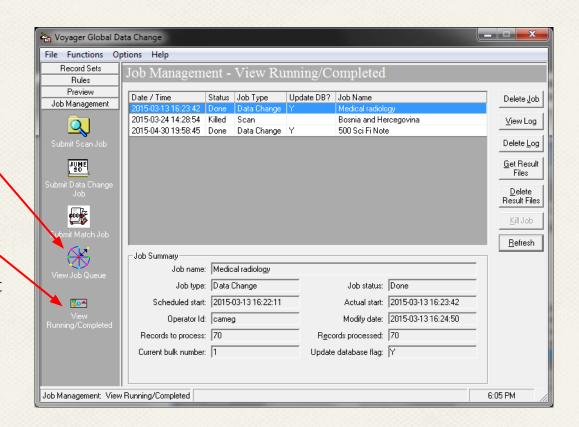
→ To see your job's position in the queue

### **View Running/Completed**

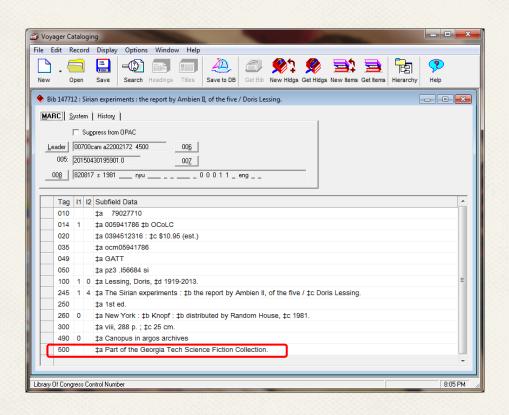
→ To see if job is Done, Failed, or Killed

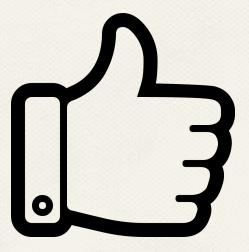
You can view logs when the jobs run, but they won't refresh

Job process wakes up every 5 minutes to check for jobs



# -Check Your Changes-

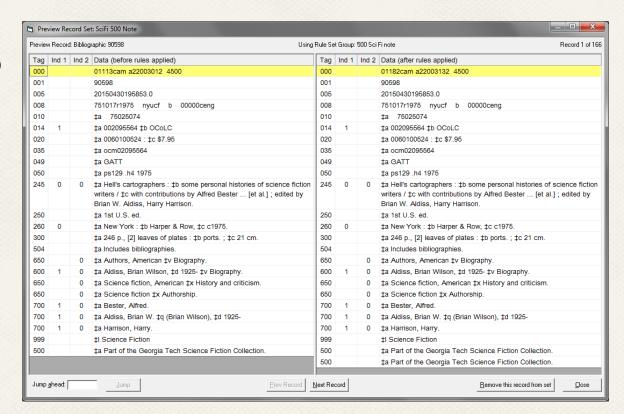




## -Check Your Changes-

You can also go back to **Preview** to see the changes in your set --

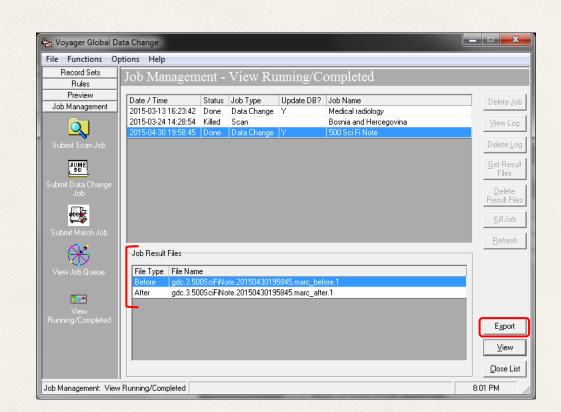
Another way to check your job from within GDC



# Uh Oh... Undoing Changes

### Job Management screen:

- **→** Job Result Files
- → Export **Before** file
- Re-import records (via pbulkimport or Gary Strawn's Record Reloader)



# GDC 101 Agenda -

- → GDC Fundamentals
- → Step-By-Step Example GDC Job
- → Notes and Best Practices
- → Other Example Rules

### **GDC Oddities**

## Inexplicable "MARC validation errors"

- $\rightarrow$  Lacking 005?
- → Obsolete codes in 008?
- → Corrupted leader?
- → Re-saving the record in the cat client (without changes) sometimes helps
- → Jobs will still run, but skip these records

### **Best Practices**

- → Know thy database
  - Knowledge of past practices can be extremely useful
- → Read the manual
- → Start with easy changes and small sets
- $\rightarrow$  Test
- → Always use Preview

### **Best Practices**

- → Small number of rules in sets, and sets in groups
  - O A.k.a. KISS
  - Uhy? If you reload, it undoes all your job's changes, so it can get complicated if you have lots of rules/sets
- → Lots of Previewing before you run the job
  - U It's iterative back and forth between Preview and Rules
  - O Previewing helps you tweak and perfect your rules/sets

### **Best Practices**

- → GDC forces you to name EVERYTHING
  - O All levels: Rules, Rule Sets, Rule Set Groups, etc.
  - Naming convention is helpful especially if multiple people have their fingers in GDC
  - Susan puts her initials at the beginning of her stuff

# GDC & Strawn programs

→ Strawn programs are (mostly) easier to learn and use

### BUT...

- → GDC doesn't require ODBC driver set-up
- → Alma normalization rules use GDC syntax
- → All are good tools for the cataloger's bag of tricks

## **GDC Updates**

### Voyager 9.0

→ Added Suppress/Unsuppress and Delete

### Voyager 9.1

- → Copy Rules option to create similar Rule Sets
- → Can add records through Match Job (via 035) = good for vendor files of record deletions

## **GDC Updates**

### Voyager 9.2

- → MARC tag ordering
- → Integration with Global Headings Queue workflow
- → Updates to Owning Library, Ok to export options

### Alma

→ GDC syntax is used to create Normalization Rules

### For more on GDC

## Articles in Ex Libris Support Portal

- → Search for "Global Data Change"
- → Some articles are How-Tos

### Can share Rule Sets and Groups

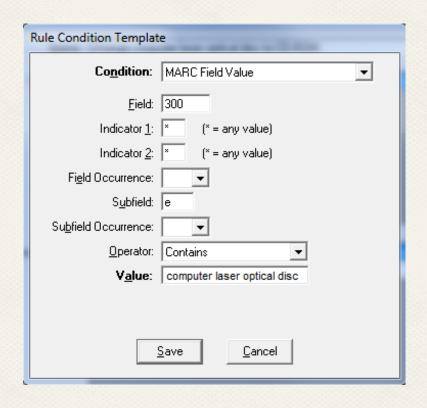
→ Ability to import and export between users

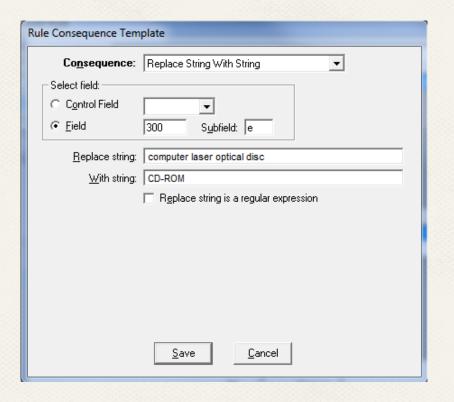
### VOYAGER-L

# GDC 101 Agenda -

- → GDC Fundamentals
- → Step-By-Step Example GDC Job
- → Notes and Best Practices
- → Other Example Rules

# Replace String with String - 1





# Replace String with String - 2

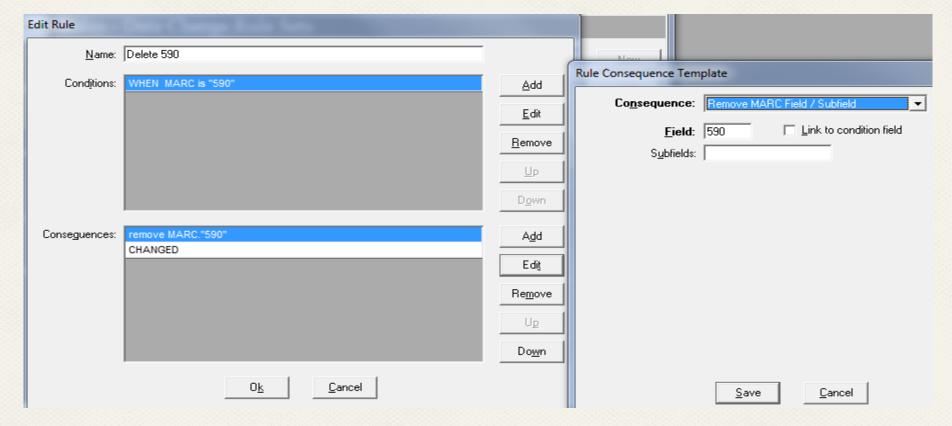
Rule Condition Templat	te
Co <u>n</u> dition:	MARC Field / Subfield Exists
<u>F</u> ield:	650
Indicator <u>1</u> :	× (* = any value)
Indicator <u>2</u> :	× (* = any value)
Fi <u>e</u> ld Occurrence:	_
S <u>u</u> bfield:	a
Su <u>b</u> field Occurrence:	▼
	<u>S</u> ave <u>C</u> ancel



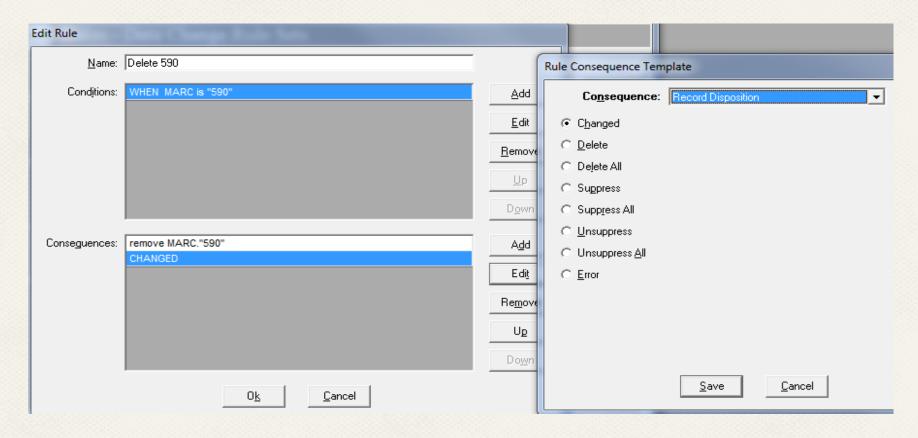
# Delete a Field: Condition

Edit Rule		
Name: Delete 590		<u>N</u> ew
Conditions: WHEN MARC is "590"	∆dd	Rule Condition Template
	<u>E</u> dit	Condition: MARC Field / Subfield Exists ▼
	<u>R</u> emove	<u>Field</u> : 590
	<u>U</u> р	Indicator 1: (* = any value)
		Indicator <u>2</u> :  × (* = any value)
	D <u>o</u> wn	Fi <u>e</u> ld Occurrence: ▼
Conseguences: remove MARC."590"	A <u>d</u> d	S <u>u</u> bfield:
CHANGED		Su <u>b</u> field Occurrence: ▼
	Ediţ	
	Re <u>m</u> ove	
	Up	
	Do <u>w</u> n	<u>S</u> ave <u>C</u> ancel
O <u>k</u> <u>C</u> ancel		- DÖMU

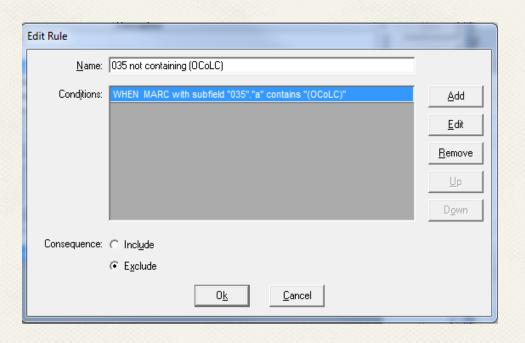
# Delete a Field: Consequence

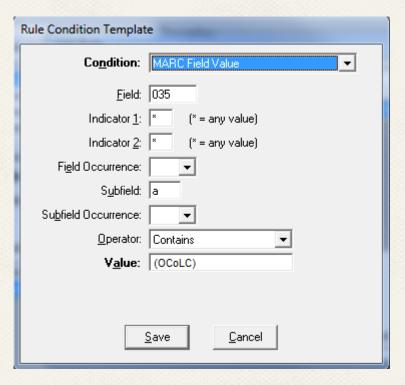


# Record Disposition



### Scan Rule Set





# Scan Rule Set Group

Rules - Sca	an Rule Set Gr	oups		
Name		Description		<u>N</u> ew
BurneyCollection scw PUL bibs lad	1 cking OCLC numbers test			<u>E</u> dit
scw PUL not e-re				
				<u>D</u> elete
N <u>a</u> me:	scw PUL bibs lacking 00	CLC numbers test		
Description:			_	
			~	
<u>L</u> ast Modified:	5/5/2015 M	odified <u>B</u> y:  libscw		
<u>R</u> ules:	Rule Sets In Group			
	scw 035 not containing	·		
	scw 035 not containing scw 035 not containing			
	scw 033 not containing	OCTI		

# Thank you!

**Erin Grant** 

Cataloging Unit Head

Georgia Institute of Technology

erin.grant@library.gatech.edu

Susan C. Wynne

Cataloging & Metadata Librarian

Georgia State University

swynne@gsu.edu