

A photograph of a modern, multi-story building with a complex, angular design, featuring large windows and a mix of light and dark grey panels. The building is set against a clear blue sky. In the foreground, there are several trees with green and yellowing leaves, suggesting autumn. A street with cars and a pedestrian is visible at the bottom of the frame.

Is deduping a kind of overlap analysis? Can you model it in Analytics? Let's Try!

Susan Bond – University of Toronto

Discovery Systems Librarian

OCUL CF Miniconference – July 8 2025

Agenda

- Origins
- Deduping
- How it works (or doesn't work) in analytics (what I've tried)
- What else could we do?

Analytics team: Overlap Analysis working group

- Local analytics team formed a working group to try to find better way of doing Overlap Analysis
 - OCLC Greenglass
 - Alma overlap report
 - Data Visualization
 - Other tools (KNIME, LLMs)

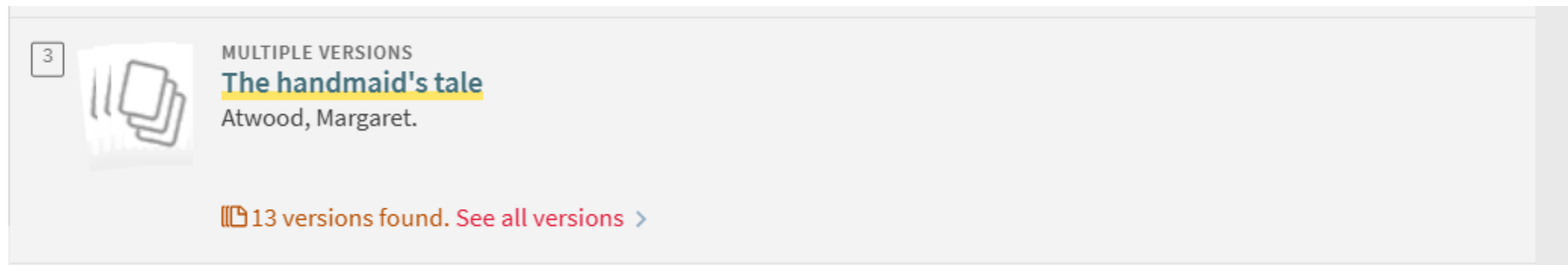
... But I thought of a place overlap analysis already happens in Alma ...

Dedup: overlap by another name

- Algorithmic processes in Alma that brings together similar records to make discovery in Primo easier
 - Brings together print records and digital surrogates, other things it thinks are “the same thing”
 - Shows the more relevant bib information and all holdings/items
 - Works well for books, really overmatches so we suppress the process on some formats (DVD + VHS, Xbox and Playstation, scores)
 - We also suppress dedup on rare items (so notes show up)

FRBR: a looser version of Dedup

- Pulls together record for e.g. different editions for display in Primo
- Presents them as a pile of things for user to scroll through
- (we don't have this turned on)
- Do you? Here's how to tell:



How do we know what records are deduping?

- Harder than it seems!
- Search results in Primo: append &showPnx=true to url and look for "isDedup" : true . If it's there, look for "dedupmemberids"
- Alma Config >> Discovery >> Other

Other

Resource Recommender Configuration

Central Index and proxy set-up

Primo Deep-Link Re-Direct Configuration

Suppress Dedup/FRBR

Dedup and FRBR Test Utility

Customer Settings

Citations Styles

Local Fields using Search and Facet

Normalization Rules

Primo VE Deep Search Configuration

Exclude Process Types from Discovery
Indexing

Dedup and FRBR Utility

[<](#) Dedup and FRBR Test Utility Back

Find Matching Records

Compare Records

Grouping Type

☐ FRBR ☒ DEDUP

Record ID *

991105916569906196

Find matching records

Recalculate FRBR and Dedup groups

Record Keys

	Key Source	Description	Key Text	Key
1	match/c5	External system number	(ocolc)424619645;(sirsi)7015684;(utl)7015684-01utoronto_inst	71123102375276250
2	SPLIT(match/f3) + match/f5 + match/f6	ISBN + brief title + date	9780662470359~frequentlyaskedquest2008update~2008	76904235919398952
3	SPLIT(match/f3) + FUZZY_STRING(m... + match/f6	ISBN + fuzzy title + date	9780662470359~frequently asked questions about the~2008	58550989967482418

Dedup and FRBR Utility continued

6	match/f7 + match/f6 + match/f9 + [match/f11]	Full title + date + pagination + main entry conditional	frequently asked questions about the science of climate change 2008 update~2008~ii, 49 p. :	56480470870097740
7	match/f7 + match/f6 + match/f10 + [match/f11]	Full title + date + publisher + main entry conditional	frequently asked questions about the science of climate change 2008 update~2008~atmospheric science assessment and integration section science and technology branch environment canada	84732474530495962

Matching Records				▼
0 - 1 of 1				 
Record ID	Title			
1. 991106140889606196	Frequently asked questions about the science of climate change			 

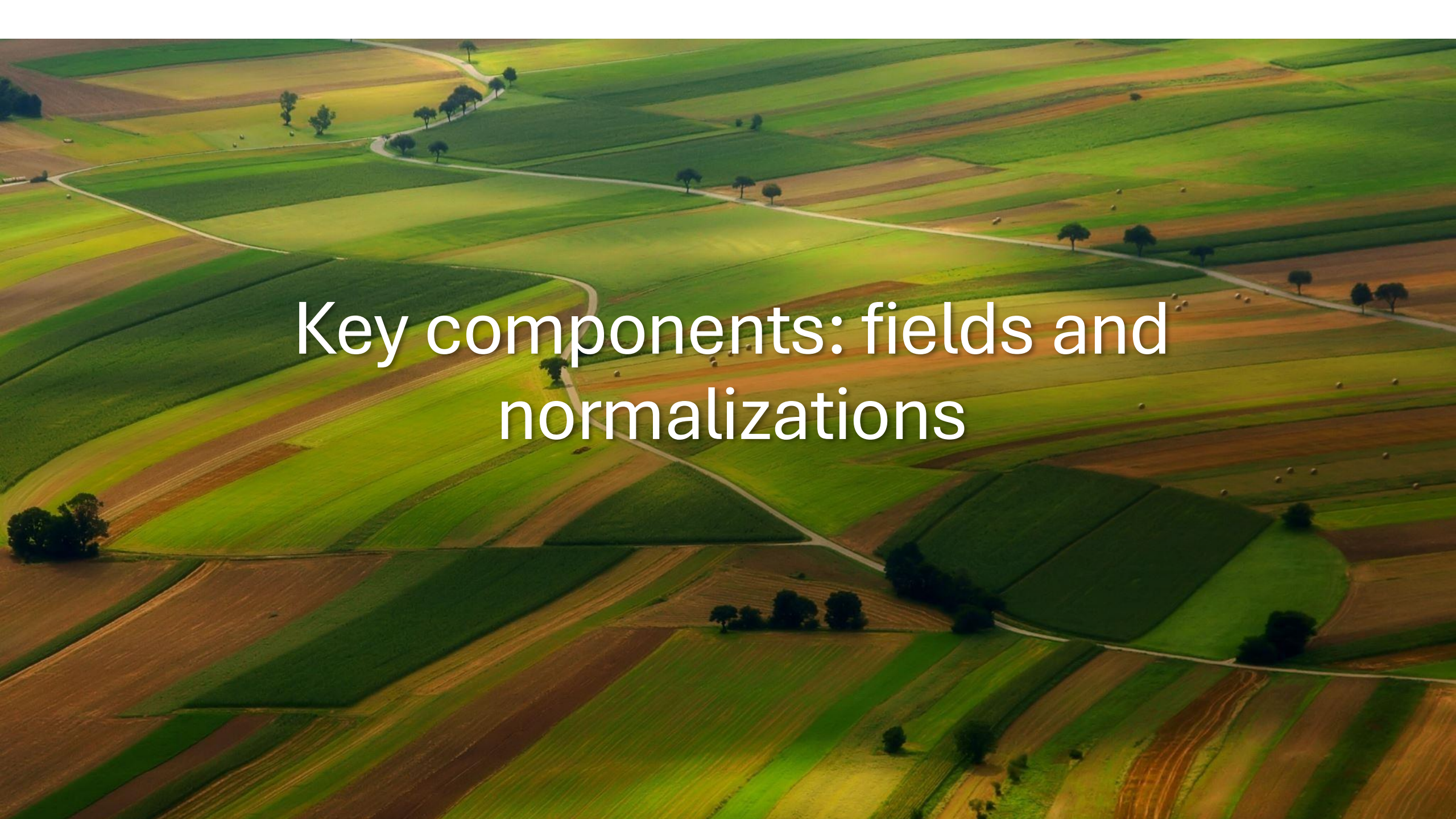
Dedup shows overlap, but is also a bit of a mystery

- Is it doing more good than harm?
- Are there individual keys that we should turn off?
- Are there new keys we should add?
- Earlier Discovery team turned off FRBR because it seemed to cause too many problems, but did it?
- Could a different/more limited FRBR (e.g. piles of videogames/scores) be useful?

So what's the deal with this algorithm?

- MARC fields are mapped to **key fields**
- Some fields are used for serials, some for non-serials, some for both
- **Key fields** are concatenated (sometimes with **normalization**) into **complete keys**
- Some sort of hashing happens
- Each key has a priority (secret?) High, medium, low, or none.
- Deduping is not fully transitive, and is done on-the-fly: record A might match B and C, but B might not match C

I would characterize my understanding of this as partial

An aerial photograph of a rural landscape featuring a complex mosaic of agricultural fields. The fields are in various stages of growth or harvest, showing shades of vibrant green, yellow-green, and brown. A light-colored, winding road or path cuts through the fields, with small clusters of trees scattered along its route. The overall scene is a patchwork of geometric shapes, illustrating the concept of fields and normalizations in a data context.

Key components: fields and
normalizations

Key Fields

key	Field content	Type	MARC 21	Analytics?
C5	Control nr	both	035 a,z	"Bibliographic IDs"."OCLC Control Number single (035a)"
F1	LCCN	both	010 a	"Bibliographic Details"."Bibliographic Id"
F3	ISBN (13 or 10 digits)	non-serial	020 a,e 776 z	"Bibliographic IDs"."ISBN valid single (Normalized)"/ "Bibliographic Details"."ISBN"
F3	ISSN	Serial	022 a,e 776 x	"Bibliographic IDs"."ISSN valid single (Normalized)"/" Bibliographic Details"."ISSN",
F4	ISBN_invalid	non-serial	020 z	Nope!
F4	ISSN_invalid	serial	022 y	Nope!
F5	Brief title	non-serial	245 a,b,n,p	"Bibliographic Details"."Title (Normalized)"
F5	ISSN_cancelled	serial	022 z	Nope!

Key Fields pg 2

key	Field content	Type	MARC 21	Analytics?
F6	Start publication year	both	008 7-10 260 c 264 c	"Bibliographic Details"."Begin Publication Date" Maybe also "Bibliographic Details"."Publication Date"
F7	Full title	both	245 a,b,n,p	"Bibliographic Details"."Title (Normalized)"
F8	Country of publication	non-serial	008 15-17	"Bibliographic Details"."Place of Publication - Country"
F8	Brief title	serial	245 a	"Bibliographic Details"."Title" (need to remove \$b)
F9	Pagination	non-serial	300 a	Nope!
F9	Country of publication	serial	008 15-17	"Bibliographic Details"."Place of Publication - Country"
F10	Publisher	non-serial	260 b 264 b	"Bibliographic Details"."Publisher"
F10	Place of publication	serial	260 a 264 a	"Bibliographic Details"."Publication Place"

Key Fields pg 3

key	Field content	Type	MARC 21	Analytics?
F11	Main entry	non-serial	100 a,b,c,d,q 111 a,c,d,e,n,q	"Bibliographic Details"."Author" (but not quite right – subfield mapping isn't clear)
F11	Main entry	serial	110 a ,b,c,d,e,n 111 a,c,d,e,n,q 130 a,d,l,m,n,o, p,r,s,t	"Bibliographic Details"."Author" (but not quite right – subfield mapping isn't clear)
F13	Control Number	both	001	"Bibliographic Details"."MMS Id"
F50 *	Work URI	both	N/A	N/A

* Not active by default: need to open a case to activate Bibframe work and entry

F5 Non-serial – how is this not F7?

From Ex Libris support:

F5 does the following:

- Uses the first 20 characters.
- Uses the last 10 characters from the end of the title
- Discards the remaining middle part (whatever is left after the first 20 characters).

```
CASE WHEN LENGTH(REPLACE("Bibliographic Details"."Title (Normalized)", ' ', '')) <= 30 THEN REPLACE("Bibliographic Details"."Title (Normalized)", ' ', '')
ELSE SUBSTRING(REPLACE("Bibliographic Details"."Title (Normalized)", ' ', ''), 1, 20) || SUBSTRING(REPLACE("Bibliographic Details"."Title (Normalized)", ' ', ''), LENGTH(REPLACE("Bibliographic Details"."Title (Normalized)", ' ', '')) - 9, 10) END
```

Normalization

Fields are not always normalized, but may be as part of the complete key generation

- FUZZY_STRING: use the first five words of the field
- ROUND_NUMBER: round the last digit of the field down to 0
- SPLIT: if a record has multiple (e.g.) ISBNs, a match on one to one counts, instead of needing to match all of them
- REMOVE_COMMON_WORDS: see list on the next slide. These words are removed from the field's value

Common Words

advertiser, anales, annual budget, annual financial report, annual report, annual reports, annual report for, annual report for the fiscal year ended, biennial report, boletin, budget, bulletin, calendar, catalogue, circular, comprehensive annual financial report, !compilation of selected acts, compilation of selected acts within the jurisdiction of the committee on commerce, conference proceedings, conference record, directory, fact sheet, final budget, financial report, general catalog, jahresbericht, journal, language sciences, laws, etc, legislative calendar, legislative summary, membership directory, memoires, memoria, minutes, mitteilungen, monograph, monthly bulletin, news release, newsletter, occasional paper, occasional papers, proceedings, proceedings of the annual meeting, progress report, proposed budget, publicaciones, publication, publications, rapport, rapport annuel, report, report and accounts, research report, revista, semi annual report to the congress, semiannual report to the congress, session laws, statistical report, technical bulletin, technical report, transactions, travaux, trudy, update, veroeffentlichungen, veroffentlichungen, works, year book, yearbook, telephone directory

REGEX for that normalization

```
TRIM( Evaluate( 'regex_replace(%1, "\s{2,}", " ")', TRIM( Evaluate( 'regex_replace(
regex_replace(      regex_replace(      regex_replace(lower(%1),
"^(^|\s)(journal|bulletin|newsletter|report|directory|publication|proceedings|catalogue|calenda
r|minutes|circular|fact\s+sheet|update)(\s|$)",      "\1\3"),
"^(^|\s)(budget|annual\s+report|biennial\s+report|final\s+budget|financial\s+report|technical\s
+report|statistical\s+report|session\s+laws|legislative\s+calendar|legislative\s+summary|repo
rt\s+and\s+accounts|proposed\s+budget)(\s|$)",      "\1\3"),
"^(^|\s)(memoires|memoria|revista|travaux|trudy|mitteilungen|anales|veroeffentlichungen|verof
fentlichungen|language\s+sciences|telephone\s+directory|year\s+book|yearbook)(\s|$)",
"\1\3"),
"^(^|\s)(advertiser|annual\s+budget|annual\s+financial\s+report|annual\s+reports?|annual\s+r
eport\s+for(\s+the\s+fiscal\s+year\s+ended)?|boletin|compilation\s+of\s+selected\s+acts(\s+
within\s+the\s+jurisdiction\s+of\s+the\s+committee\s+on\s+commerce)?|comprehensive\s+a
nnual\s+financial\s+report|conference\s+proceedings|conference\s+record|general\s+catalog
|jahresbericht|laws,\s+etc|membership\s+directory)(\s|$)",      "\1\3"),
"^(^|\s)(monograph|monthly\s+bulletin|news\s+release|occasional\s+papers?|proceedings\s+o
f\s+the\s+annual\s+meeting|progress\s+report|publicaciones|publications|rapport|rapport\s+
annuel|research\s+report|semi\s+annual\s+report\s+to\s+the\s+congress|semiannual\s+repo
rt\s+to\s+the\s+congress|technical\s+bulletin|transactions|works)(\s|$)",      "\1\3")      ',
"Bibliographic Details"."Title (Normalized)"  ) ) ) )
```

Confession:

- Everything I know about REGEX I learned from [a presentation at last year's ELUNA Analytics Afternoon](#)
- Nearly all the REGEX in this presentation was assisted by many, many attempts with ChatGPT.
- I identify as a GenAI skeptic, but it was useful for this sort of thing where I can kind of understand it, I can tell if it's working or not, but my reach exceeds my grasp
- Also, usually had to break it down into steps and iterate

Pagination: Local Param 14 is 300

(I found this syntax here:

https://knowledge.exlibrisgroup.com/Alma/Knowledge_Articles/Useful_Custom_Formulas_for_Alma_Analytics)

- **Just show the \$\$a:**

```
Evaluate('regexp_substr(%1,%2,%3,%4,%5,%6)', "Bibliographic Details"."Local  
Param 14", '\$\$a([^\$]+)', 1, 1, NULL, 0)
```

- **Now just show the integer:**

```
EVALUATE('REGEXP_SUBSTR(%1, "[0-9]+")',
```

```
Evaluate('regexp_substr(%1,%2,%3,%4,%5,%6)', "Bibliographic Details"."Local  
Param 14", '\$\$a([^\$]+)', 1, 1, NULL, 0))
```

Extract the first integer after the \$\$a **followed by p., p, pp, or pages:**

- Evaluate('regexp_substr(%1,%2,%3,%4,%5,%6)', "Bibliographic Details"."Local Param 14", '\\$\\$a.*?(\d+)\s*(?:p\.|pp|pages|p)', 1, 1, NULL, 1)

Make it a **number**

- CAST(Evaluate('regexp_substr(%1,%2,%3,%4,%5,%6)', "Bibliographic Details"."Local Param 14", '\\$\\$a.*?(\d+)\s*(?:p\.|pp|pages|p)', 1, 1, NULL, 1)) AS DOUBLE)

Only capture the integer **if it's before the next \$\$**

- CAST(Evaluate('regexp_substr(%1,%2,%3,%4,%5,%6)', "Bibliographic Details"."Local Param 14", '\\$\\$a[^\$]*?(\d+)\s*(?:p\.|pp|pages|p)', 1, 1, NULL, 1) AS DOUBLE)

Round that number down:

- FLOOR(CAST(Evaluate('regexp_substr(%1,%2,%3,%4,%5,%6)', "Bibliographic Details"."Local Param 14", '\\$\\$a.*?(\d+)\s*(?:p\.|pp|pages|p)', 1, 1, NULL, 1) AS DOUBLE) / 10) * 10

But can't concatenate with a number!

- Final rounded:

```
CAST(FLOOR(CAST(Evaluate(  
'regexp_substr(%1,%2,%3,%4,%5,%6)', "Bibliographic  
Details"."Local Param 14", '$a[^$]*?(\\d+)\\s*(?:p\\.?|pp|pages|p)',  
1, 1, NULL, 1 ) AS DOUBLE)/10)*10 AS CHAR)
```

- Final unrounded:

```
Evaluate( 'regexp_substr(%1,%2,%3,%4,%5,%6)', "Bibliographic  
Details"."Local Param 14", '$a[^$]*?(\\d+)\\s*(?:p\\.?|pp|pages|p)',  
1, 1, NULL, 1 )
```

Problems so far

- Not all these fields are in Analytics, or have the correct subfields mapped: is it really worth re-mapping e.g. 100, 110, 130?
- ISBN/ISSN invalid is included (but not differentiated) in bib details.ISBN. How to tease out? Should we just use all of this for F3 and F4?
- SPLIT can sometimes be accomplished by using the Bibliographic IDs folder, but is the kind of match we get from a full list what we want?
- What level of normalization to we want in general?

Other normalizations in Analytics

Alma does the following to the Normalized title field (except subfields c,h,6 and 8) in the following order:

- Removes non filing text.
- Removes <> and anything between them.
- Removes these characters: [] , . ; : "
- Changes these characters to a space: ! @ # \$ % ^ & * () _ + - = { } \ < > ? / ~ '
- **Changes characters to lowercase.**
- Packs spaces (replace all sequences of more than one space with a single space).



Non-Serial keys

Match/c5

External system number

"Bibliographic Details"."OCLC Control Number (035a+z)"

▲ Filters

- ▼ Lifecycle is equal to / is in In Repository
- AND ▼ Num of Titles (Active) is greater than 1
- AND ▼ Suppressed From Discovery is equal to / is in No
- AND ▼ OCLC Control Number (035a+z) is not null
- AND ▼ Bibliographic Level is not equal to / is not in s

match/f1 + match/f5 + match/f6
LCCN + brief title + year

```
CONCAT(CONCAT("Bibliographic  
Details"."Bibliographic Id", CASE  
WHEN  
LENGTH(REPLACE("Bibliographic  
Details"."Title (Normalized)", ' ', '')) <= 30 THEN REPLACE("Bibliographic  
Details"."Title (Normalized)", ' ', '')  
ELSE  
SUBSTRING(REPLACE("Bibliographic  
Details"."Title (Normalized)", ' ', '')  
FROM 1 FOR 20) ||  
SUBSTRING(REPLACE("Bibliographic  
Details"."Title (Normalized)", ' ', '')  
FROM  
LENGTH(REPLACE("Bibliographic  
Details"."Title (Normalized)", ' ', '')) - 9  
FOR 10) END ), "Bibliographic  
Details"."Begin Publication Date")
```

Filters

- Num of Titles (Active) is greater than 1
- AND Bibliographic Level is not equal to / is not in s
- AND Bibliographic Id is not null
- AND Begin Publication Date is not null
- AND Lifecycle is equal to / is in In Repository
- AND Suppressed From Discovery is equal to / is in No
- AND Case WHEN LENGTH(REPLACE(Title (No... is not null

match/f1 + **FUZZY_STRING**(match/f7) + match/f6
LCCN + fuzzy title + year

```
CONCAT("Bibliographic  
Details"."Bibliographic Id",  
CONCAT(Evaluate( 'regexp_substr(%1,  
"(\S+\s+){0,4}\S+)"', "Bibliographic  
Details"."Title (Normalized)" ),  
"Bibliographic Details"."Begin  
Publication Date"))
```

▲ Filters

- ▼ Num of Titles (Active) is greater than 1
- AND** ▼ Bibliographic Level is not equal to / is not in s
- AND** ▼ EVALUATE('regexp_substr(%1, "(... is not null
- AND** ▼ Bibliographic Id is not null
- AND** ▼ Begin Publication Date is not null
- AND** ▼ Suppressed From Discovery is equal to / is in No

match/f1 + match/f7 + match/f6
LCCN + full title + year

CONCAT(CONCAT("Bibliographi
c Details"."Bibliographic Id",
"Bibliographic Details"."Title
(Normalized)"), "Bibliographic
Details"."Begin Publication
Date")

▲ Filters

- ▼ Num of Titles (Active) is greater than 1
- AND ▼ Bibliographic Level is not equal to / is not in s
- AND ▼ Bibliographic Id is not null
- AND ▼ Begin Publication Date is not null
- AND ▼ Lifecycle is equal to / is in In Repository
- AND ▼ Suppressed From Discovery is equal to / is in No
- AND ▼ Title (Normalized) is not null

match/f3 + match/f5 + match/f6
ISBN + brief title + date

```
CONCAT("Bibliographic Details"."ISBN",  
CONCAT( CASE WHEN  
LENGTH(REPLACE("Bibliographic  
Details"."Title (Normalized)", ' ', '')) <= 30  
THEN REPLACE("Bibliographic  
Details"."Title (Normalized)", ' ', '') ELSE  
SUBSTRING(REPLACE("Bibliographic  
Details"."Title (Normalized)", ' ', '') FROM 1  
FOR 20) ||  
SUBSTRING(REPLACE("Bibliographic  
Details"."Title (Normalized)", ' ', '') FROM  
LENGTH(REPLACE("Bibliographic  
Details"."Title (Normalized)", ' ', '')) - 9 FOR  
10) END, "Bibliographic Details"."Begin  
Publication Date"))
```

Filters

- Num of Titles (Active) is greater than 1
- AND Begin Publication Date is not null
- AND Suppressed From Discovery is equal to / is in No
- AND ISBN is not null
- AND Bibliographic Level is not equal to / is not in s
- AND Case WHEN LENGTH(REPLACE(Title (No... is not null

match/f3 + FUZZY_STRING(match/f7) + match/f6
ISBN + fuzzy title + date

```
CONCAT("Bibliographic Details"."ISBN",  
CONCAT(Evaluate( 'regexp_substr(%1,  
"(\S+\s+){0,4}\S+"), "Bibliographic  
Details"."Title (Normalized)" ),  
"Bibliographic Details"."Begin  
Publication Date"))
```

Filters

- Num of Titles (Active) is greater than 1
- AND Bibliographic Level is not equal to / is not in s
- AND Begin Publication Date is not null
- AND Suppressed From Discovery is equal to / is in No
- AND ISBN is not null
- AND EVALUATE('regexp_substr(%1, "(... is not null

match/f3 + match/f7 + match/f9 ISBN + full title + pagination

```
CONCAT("Bibliographic Details"."ISBN",  
CONCAT("Bibliographic Details"."Title  
(Normalized)" , Evaluate(  
'regexp_substr(%1,%2,%3,%4,%5,%6)',  
"Bibliographic Details"."Local Param  
14",  
'\$\$a[^$]*?(\d+)\s*(?:p\.?|pp|pages|p)',  
1, 1, NULL, 1 ) ) )
```

 Filters

 Num of Titles (Active) is greater than 1

AND  Bibliographic Level is not equal to / is not in s

AND  Category of Material is not equal to / is not in d; f; g; j; k; m; o; r; s; v; z

AND  Suppressed From Discovery is equal to / is in No

AND  Title (Normalized) is not null

AND  CAST (EVALUATE('regexp_substr(... is not null

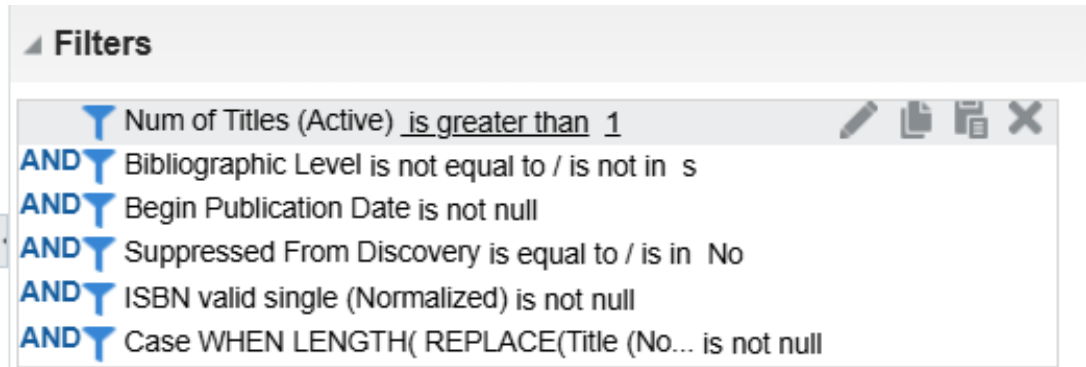
AND  ISBN is not null

SPLIT(match/f3) + match/f5 + match/f6

ISBN1 + brief title + date

ISBN2 + brief title + date

```
CONCAT("Bibliographic IDs"."ISBN valid  
single (Normalized)", CONCAT(CASE  
WHEN LENGTH(REPLACE("Bibliographic  
Details"."Title (Normalized)", ' ', '')) <= 30  
THEN REPLACE("Bibliographic  
Details"."Title (Normalized)", ' ', '') ELSE  
SUBSTRING(REPLACE("Bibliographic  
Details"."Title (Normalized)", ' ', '') FROM 1  
FOR 20) ||  
SUBSTRING(REPLACE("Bibliographic  
Details"."Title (Normalized)", ' ', '') FROM  
LENGTH(REPLACE("Bibliographic  
Details"."Title (Normalized)", ' ', '')) - 9 FOR  
10) END , "Bibliographic Details"."Begin  
Publication Date"))
```



SPLIT(match/f3) + FUZZY_STRING(match/f7) + match/f6

ISBN1 + fuzzy title + date

ISBN2 + fuzzy title + date

```
CONCAT("Bibliographic IDs"."ISBN valid  
single (Normalized)",  
CONCAT(Evaluate( 'regexp_substr(%1,  
"(\S+\s+){0,4}\S+)"', "Bibliographic  
Details"."Title (Normalized)" ),  
"Bibliographic Details"."Begin  
Publication Date"))
```

Filters

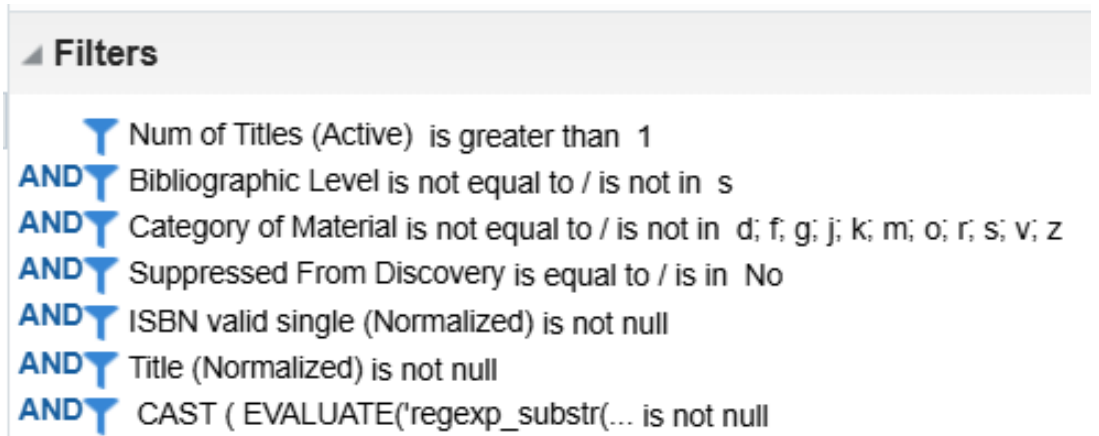
- Num of Titles (Active) is greater than 1
- AND Bibliographic Level is not equal to / is not in s
- AND Begin Publication Date is not null
- AND Suppressed From Discovery is equal to / is in No
- AND EVALUATE("regexp_substr(%1, "(... is not null
- AND ISBN valid single (Normalized) is not null

SPLIT(match/f3) + match/f7 + match/f9

ISBN1 + full title + pagination

ISBN2 + full title + pagination

```
CONCAT("Bibliographic IDs"."ISBN  
valid single (Normalized)",  
CONCAT("Bibliographic Details"."Title  
(Normalized)" , Evaluate(  
'regexp_substr(%1,%2,%3,%4,%5,%6)',  
"Bibliographic Details"."Local Param  
14",  
'\$\$a[^$]*?(\d+)\s*(?:p\.|pp|pages|p)'  
, 1, 1, NULL, 1 ) )
```

A screenshot of a 'Filters' panel in a software application. The panel has a light gray header with the word 'Filters' and a small downward arrow icon. Below the header, there is a list of filter conditions, each preceded by a blue funnel icon. The first filter is 'Num of Titles (Active) is greater than 1'. The subsequent five filters are preceded by the word 'AND' in blue, indicating they are part of a logical AND group. These filters are: 'Bibliographic Level is not equal to / is not in s', 'Category of Material is not equal to / is not in d; f; g; j; k; m; o; r; s; v; z', 'Suppressed From Discovery is equal to / is in No', 'ISBN valid single (Normalized) is not null', 'Title (Normalized) is not null', and 'CAST (EVALUATE('regexp_substr(... is not null'.

Filters

- Num of Titles (Active) is greater than 1
- AND Bibliographic Level is not equal to / is not in s
- AND Category of Material is not equal to / is not in d; f; g; j; k; m; o; r; s; v; z
- AND Suppressed From Discovery is equal to / is in No
- AND ISBN valid single (Normalized) is not null
- AND Title (Normalized) is not null
- AND CAST (EVALUATE('regexp_substr(... is not null

match/f4 + match/f7 + match/f6

match/f4 + match/f7 + match/f9

SPLIT(match/f4) + match/f7 + match/f6

SPLIT(match/f4) + match/f7 + match/f9

- Incorrect ISBN + full title + date (and all the other ones we just did with f3)

Not possible! But some of it is captured by the f3 ones?

match/f7 + match/f11 + match/f6 + match/f9
full title + main entry + date + pagination

```
CONCAT("Bibliographic Details"."Title  
(Normalized)",CONCAT("Bibliographic  
Details"."Author",  
CONCAT("Bibliographic Details"."Begin  
Publication Date" , Evaluate(  
'regexp_substr(%1,%2,%3,%4,%5,%6)',  
"Bibliographic Details"."Local Param  
14",  
'\$\$a[^$]*?(\d+)\s*(?:p\.?|pp|pages|p)',  
1, 1, NULL, 1 ) )))
```

Filters

- Num of Titles (Active) is greater than 1
- AND Bibliographic Level is not equal to / is not in s
- AND Begin Publication Date is not null
- AND Category of Material is not equal to / is not in d; f; g; j; k; m; o; r; s; v; z
- AND Suppressed From Discovery is equal to / is in No
- AND Title (Normalized) is not null
- AND CAST (EVALUATE('regexp_substr(... is not null
- AND Author is not null
- AND Lifecycle is equal to / is in In Repository

match/f7 + match/f11 + match/f6 +
ROUND_NUMBER(match/f9)
full title + main entry + date + fuzzy pagination

```
CONCAT("Bibliographic Details"."Title  
(Normalized)",CONCAT("Bibliographic Details"."Author",  
CONCAT("Bibliographic Details"."Begin Publication Date" ,  
CAST(FLOOR(CAST(Evaluate(  
'regexp_substr(%1,%2,%3,%4,%5,%6)', "Bibliographic  
Details"."Local Param 14",  '\$\$a[^$]*?(\d+)\s*(?:p\.?|pp|pages|p)',  
1, 1, NULL, 1 ) AS DOUBLE)/10)*10 AS CHAR))))
```

match/f7 + match/f11 + match/f6 +
ROUND_NUMBER(match/f9)

full title + main entry + date + fuzzy pagination -
filters

▲ Filters

▼ Num of Titles (Active) is greater than 1
AND ▼ Bibliographic Level is not equal to / is not in s
AND ▼ Begin Publication Date is not null
AND ▼ Category of Material is not equal to / is not in d; f; g; j; k; m; o; r; s; v; z
AND ▼ Suppressed From Discovery is equal to / is in No
AND ▼ Title (Normalized) is not null
AND ▼ CAST (EVALUATE('regexp_substr(... is not null
AND ▼ Author is not null
AND ▼ Lifecycle is equal to / is in In Repository
AND ▼ FLOOR(CAST (EVALUATE('regexp_substr(%1,%2,%3,%4,%5,%6)', 300 - Local Param 14, '\\$\\$a[^\\$]*?(\d+)\s*(?:p\.?|pp|pages|p)', 1, 1, NULL , 1) AS DOUBLE PRECISION) / 10) * 10 is not equal to / is not in 0

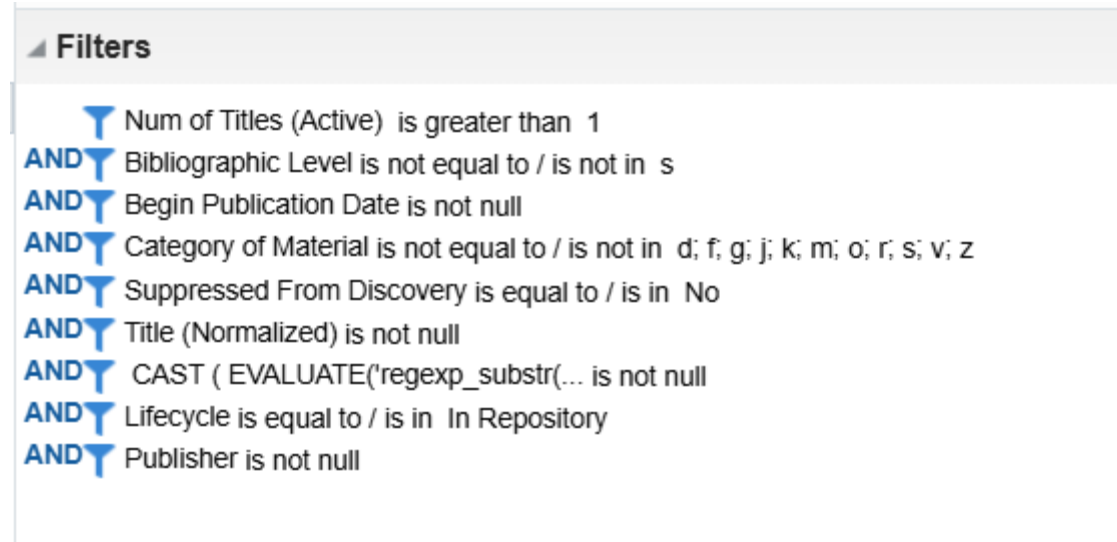
match/f7 + match/f6 + match/f10 + match/f9 +
[match/f11]

Full title + date + publisher + pagination + [main entry]

```
CONCAT("Bibliographic Details"."Title (Normalized)",  
CONCAT("Bibliographic Details"."Publisher",  
CONCAT("Bibliographic Details"."Begin Publication Date" ,  
CONCAT(Evaluate( 'regexp_substr(%1,%2,%3,%4,%5,%6)',  
"Bibliographic Details"."Local Param 14",  
'\$\$a[^$]*?(\d+)\s*(?:p\.?|pp|pages|p)', 1, 1, NULL, 1 ),  
"Bibliographic Details"."Author"))))
```

match/f7 + match/f6 + match/f10 + match/f9 +
[match/f11]

Full title + date + publisher + pagination + [main entry] -
filters



▲ Filters

- ▼ Num of Titles (Active) is greater than 1
- AND ▼ Bibliographic Level is not equal to / is not in s
- AND ▼ Begin Publication Date is not null
- AND ▼ Category of Material is not equal to / is not in d; f; g; j; k; m; o; r; s; v; z
- AND ▼ Suppressed From Discovery is equal to / is in No
- AND ▼ Title (Normalized) is not null
- AND ▼ CAST (EVALUATE('regex_substr(... is not null
- AND ▼ Lifecycle is equal to / is in In Repository
- AND ▼ Publisher is not null

Note – Author not a filter, because the brackets mean f11 is optional

match/f7 + match/f6 + match/f10 +
ROUND_NUMBER(match/f9) + [match/f11]
Full title + date + publisher + fuzzy pagination + [main entry]

```
CONCAT("Bibliographic Details"."Title (Normalized)",  
CONCAT("Bibliographic Details"."Publisher",  
CONCAT("Bibliographic Details"."Begin Publication Date" ,  
CONCAT(CAST(FLOOR(CAST(Evaluate(  
'regex_substr(%1,%2,%3,%4,%5,%6)', "Bibliographic  
Details"."Local Param 14", '$a[^$]*?(\d+)\s*(?:p\.?|pp|pages|p)',  
1, 1, NULL, 1 ) AS DOUBLE)/10)*10 AS CHAR), "Bibliographic  
Details"."Author"))))
```

match/f7 + match/f6 + match/f10 +
ROUND_NUMBER(match/f9) + [match/f11] -
filters

Filters

Num of Titles (Active) is greater than 1
AND Bibliographic Level is not equal to / is not in s
AND Begin Publication Date is not null
AND Category of Material is not equal to / is not in d; f; g; j; k; m; o; r; s; v; z
AND Suppressed From Discovery is equal to / is in No
AND Title (Normalized) is not null
AND CAST (EVALUATE('regex_substr(... is not null
AND Lifecycle is equal to / is in In Repository
AND FLOOR(CAST (EVALUATE('regex_substr('%1,%2,%3,%4,%5,%6)', 300 - Local Param 14, '\$a[^\$]*?(\\d+)\\s*(?:p\\.?|pp|pages|p)', 1, 1, NULL , 1) AS DOUBLE PRECISION) / 10) * 10 is not equal to / is not in 0
AND Publisher is not null

match/f7 + match/f6 + match/f9 +
[match/f11]

Full title + date + pagination + [main entry]

```
CONCAT("Bibliographic Details"."Title (Normalized)",  
CONCAT("Bibliographic Details"."Begin Publication Date" ,  
CONCAT(Evaluate( 'regexp_substr(%1,%2,%3,%4,%5,%6)',  
"Bibliographic Details"."Local Param 14",  
'\$\$a[^$]*?(\d+)\s*(?:p\.?|pp|pages|p)', 1, 1, NULL, 1 ),  
"Bibliographic Details"."Author")))
```

match/f7 + match/f6 + match/f9 +
[match/f11] - filters

▲ Filters

- ▼ Num of Titles (Active) is greater than 1
- AND ▼ Bibliographic Level is not equal to / is not in s
- AND ▼ Begin Publication Date is not null
- AND ▼ Category of Material is not equal to / is not in d; f; g; j; k; m; o; r; s; v; z
- AND ▼ Suppressed From Discovery is equal to / is in No
- AND ▼ Title (Normalized) is not null
- AND ▼ CAST (EVALUATE('regexp_substr(... is not null
- AND ▼ Lifecycle is equal to / is in In Repository

match/f7 + match/f6 +
ROUND_NUMBER(match/f9) + [match/f11]
Full title + date + fuzzy pagination + [main entry]

```
CONCAT("Bibliographic Details"."Title (Normalized)",  
CONCAT("Bibliographic Details"."Begin Publication Date" ,  
CONCAT(CAST(FLOOR(CAST(Evaluate(  
'regex_substr(%1,%2,%3,%4,%5,%6)', "Bibliographic  
Details"."Local Param 14",  '\$\$a[^$]*?(\d+)\s*(?:p\.?|pp|pages|p)',  
1, 1, NULL, 1 ) AS DOUBLE)/10)*10 AS CHAR), "Bibliographic  
Details"."Author"))))
```

match/f7 + match/f6 +
ROUND_NUMBER(match/f9) + [match/f11] -
filters

 Filters



 Num of Titles (Active) is greater than 1

AND



 Bibliographic Level is not equal to / is not in s

AND



 Begin Publication Date is not null

AND



 Category of Material is not equal to / is not in d; f; g; j; k; m; o; r; s; v; z

AND



 Suppressed From Discovery is equal to / is in No

AND



 Title (Normalized) is not null

AND



 CAST (EVALUATE('regexp_substr(... is not null

AND



 Lifecycle is equal to / is in In Repository

AND



 FLOOR(CAST (EVALUATE('regexp_substr(%1,%2,%3,%4,%5,%6)', 300 - Local Param 14, '\\$\\$a[^\$]*(\d+)\s*(?:p\.?|pp|pages|p)', 1, 1, NULL , 1)AS DOUBLE PRECISION) / 10) * 10 is not equal to / is not in 0

match/f7 + match/f6 + match/f10 + [match/f11]
Full title + date + publisher + [main entry]

```
CONCAT("Bibliographic Details"."Title  
(Normalized)", CONCAT("Bibliographic  
Details"."Publisher",  
CONCAT("Bibliographic Details"."Begin  
Publication Date", "Bibliographic  
Details"."Author"))))
```

 Filters

 Num of Titles (Active) is greater than 1

AND  Bibliographic Level is not equal to / is not in s

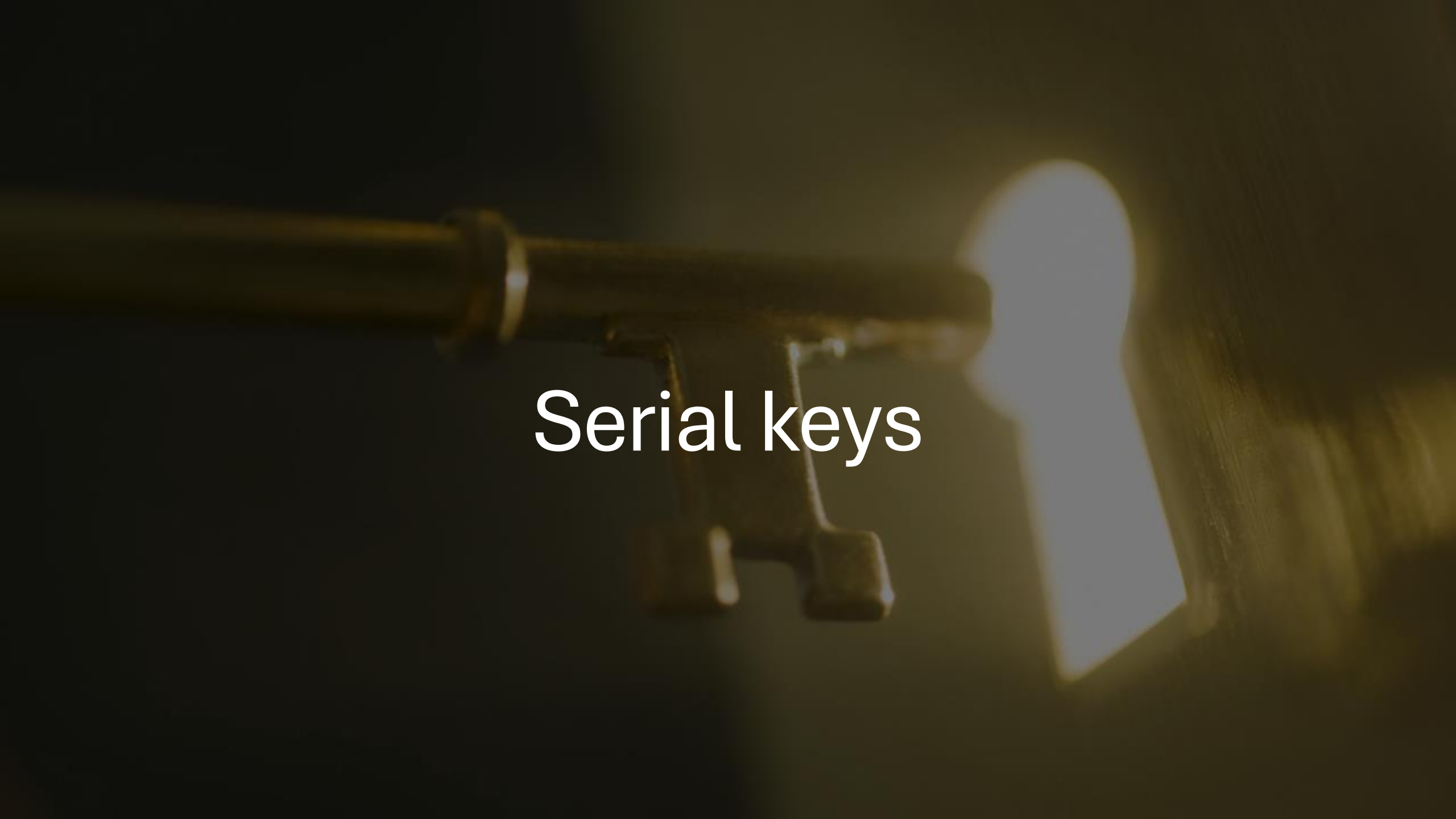
AND  Begin Publication Date is not null

AND  Suppressed From Discovery is equal to / is in No

AND  Title (Normalized) is not null

AND  Lifecycle is equal to / is in In Repository

AND  Publisher is not null



Serial keys

match/c5
MMS id

This . . . shouldn't work? MMS Ids are all unique?

match/f1 + match/f8 LCCN + brief title

```
CONCAT("Bibliographic  
Details"."Bibliographic Id",  
LOWER(Evaluate( 'regexp_substr(%1,  
"^[^:]*")', "Bibliographic Details"."Title"  
)))
```

Filters

- Num of Titles (Active) is greater than 1
- AND Bibliographic Level is equal to / is in s
- AND Bibliographic Id is not null
- AND Lifecycle is equal to / is in In Repository
- AND Suppressed From Discovery is equal to / is in No
- AND LOWER(EVALUATE('regexp_substr(... is not null

match/f3 + match/f8 ISSN + brief title

```
CONCAT("Bibliographic Details"."ISSN",  
LOWER(Evaluate( 'regexp_substr(%1,  
"^[^:]*")', "Bibliographic Details"."Title"  
)))
```

Filters

- Num of Titles (Active) is greater than 1
- AND Bibliographic Level is equal to / is in s
- AND Lifecycle is equal to / is in In Repository
- AND Suppressed From Discovery is equal to / is in No
- AND ISSN is not null
- AND LOWER(EVALUATE('regexp_substr(... is not null

SPLIT(match/f3) + match/f8
ISSN1 + brief title
ISSN2 + brief title

```
CONCAT("Bibliographic IDs"."ISSN valid  
single (Normalized)", LOWER(Evaluate(  
'regexp_substr(%1, "^[^:]*")',  
"Bibliographic Details"."Title" )))
```

Filters

- Num of Titles (Active) is greater than 1
- AND Bibliographic Level is equal to / is in s
- AND Lifecycle is equal to / is in In Repository
- AND Suppressed From Discovery is equal to / is in No
- AND LOWER(EVALUATE('regexp_substr(... is not null
- AND ISSN valid single (Normalized) is not null

REMOVE_COMMON_WORDS(match/f7) + match/f10 + match/f9 + [match/f11]

Full title (not in common serials list) + place of publication + country of publication + Main entry conditional

```
CONCAT(TRIM( Evaluate( 'regexp_replace(      regexp_replace(      regexp_replace(      regexp_replace(
regexp_replace(lower(%1),
"^(\\s)(journal|bulletin|newsletter|report|directory|publication|proceedings|catalogue|calendar|minutes|circular|fact\\s+sheet|update)(\\s|$)",      "\\1\\3"),
"^(\\s)(budget|annual\\s+report|biennial\\s+report|final\\s+budget|financial\\s+report|technical\\s+report|statistical\\s+report|session\\s+laws|legislative\\s+calendar|legislative\\s+summary|report\\s+and\\s+accounts|proposed\\s+budget)(\\s|$)",      "\\1\\3"),
"^(\\s)(memoires|memoria|revista|travaux|trudy|mitteilungen|anales|veroeffentlichungen|veroffentlichungen|language\\s+sciences|telephone\\s+directory|year\\s+book|yearbook)(\\s|$)",      "\\1\\3"),
"^(\\s)(advertiser|annual\\s+budget|annual\\s+financial\\s+report|annual\\s+reports?|annual\\s+report\\s+for(\\s+the\\s+fiscal\\s+year\\s+ended)?|boletin|compilation\\s+of\\s+selected\\s+acts(\\s+within\\s+the\\s+jurisdiction\\s+of\\s+the\\s+committee\\s+on\\s+commerce)?|comprehensive\\s+annual\\s+financial\\s+report|conference\\s+proceedings|conference\\s+record|general\\s+catalog|jahresbericht|laws,\\s+etc|membership\\s+directory)(\\s|$)",      "\\1\\3"),
"^(\\s)(monograph|monthly\\s+bulletin|news\\s+release|occasional\\s+papers?|proceedings\\s+of\\s+the\\s+annual\\s+meeting|progress\\s+report|publicaciones|publications|rapport|rapport\\s+annuel|research\\s+report|semi\\s+annual\\s+report\\s+to\\s+the\\s+congress|semiannual\\s+report\\s+to\\s+the\\s+congress|technical\\s+bulletin|transactions|works)(\\s|$)",      "\\1\\3")      , "Bibliographic Details"."Title (Normalized)"  ) ), CONCAT(LOWER("Bibliographic Details"."Publication Place"), CONCAT("Bibliographic Details"."Place of Publication - Country", "Bibliographic Details"."Author")))
```

N.B. it takes more than 20 minutes to run this query in my Alma analytics!

REMOVE_COMMON_WORDS(match/f7) +
match/f10 + match/f9 + [match/f11] - filters

▲ Filters

- Num of Titles (Active) is greater than 1
- AND Num of Titles (Active) is greater than 1
- AND Bibliographic Level is equal to / is in s
- AND Lifecycle is equal to / is in In Repository
- AND Suppressed From Discovery is equal to / is in No
- AND LOWER(Publication Place) is not null
- AND Place of Publication - Country is not null
- AND TRIM(EVALUATE('regexp_replace(... is not null
- AND Place of Publication - Country is not equal to / is not in No place, unknown, or undetermined

REMOVE_COMMON_WORDS(match/f7) + match/f10 + match/f9 + [match/f11] + [SPLIT(match/f3)] (formula)

```
CONCAT(CONCAT(TRIM( Evaluate( 'regexp_replace(      regexp_replace(      regexp_replace(
regexp_replace(      regexp_replace(lower(%1),
"(^|\s)(journal|bulletin|newsletter|report|directory|publication|proceedings|catalogue|calendar|minutes|circular|fac
t\s+sheet|update)(\s|$)",      "\1\3"),
"(^|\s)(budget|annual\s+report|biennial\s+report|final\s+budget|financial\s+report|technical\s+report|statistical\s+
report|session\s+laws|legislative\s+calendar|legislative\s+summary|report\s+and\s+accounts|proposed\s+budget)
(\s|$)",      "\1\3"),
"(^|\s)(memoires|memoria|revista|travaux|trudy|mitteilungen|anales|veroeffentlichungen|veroffentlichungen|langua
ge\s+sciences|telephone\s+directory|year\s+book|yearbook)(\s|$)",      "\1\3"),
"(^|\s)(advertiser|annual\s+budget|annual\s+financial\s+report|annual\s+reports?|annual\s+report\s+for(\s+the\s+
fiscal\s+year\s+ended)?|boletin|compilation\s+of\s+selected\s+acts(\s+within\s+the\s+jurisdiction\s+of\s+the\s+c
ommittee\s+on\s+commerce)?|comprehensive\s+annual\s+financial\s+report|conference\s+proceedings|conferen
ce\s+record|general\s+catalog|jahresbericht|laws,\s+etc|membership\s+directory)(\s|$)",      "\1\3"),
"(^|\s)(monograph|monthly\s+bulletin|news\s+release|occasional\s+papers?|proceedings\s+of\s+the\s+annual\s+
meeting|progress\s+report|publicaciones|publications|rapport|rapport\s+annuel|research\s+report|semi\s+annual
\s+report\s+to\s+the\s+congress|semiannual\s+report\s+to\s+the\s+congress|technical\s+bulletin|transactions|w
orks)(\s|$)",      "\1\3")      , "Bibliographic Details"."Title (Normalized)"  ) ), CONCAT(LOWER("Bibliographic
Details"."Publication Place"), CONCAT("Bibliographic Details"."Place of Publication - Country", "Bibliographic
Details"."Author"))) , "Bibliographic IDs"."ISSN valid single (Normalized)")
```

REMOVE_COMMON_WORDS(match/f7) + match/f10 + match/f9 + [match/f11] + [SPLIT(match/f3)] (filters)

▲ Filters

- ▼ Num of Titles (Active) is greater than 1
- AND ▼ Bibliographic Level is equal to / is in s
- AND ▼ Lifecycle is equal to / is in In Repository
- AND ▼ Suppressed From Discovery is equal to / is in No
- AND ▼ LOWER(Publication Place) is not null
- AND ▼ Place of Publication - Country is not null
- AND ▼ TRIM(EVALUATE('regexp_replace(... is not null
- AND ▼ Place of Publication - Country is not equal to / is not in No place, unknown, or undetermined
- AND ▼ ISSN valid single (Normalized) is not null

REMOVE_COMMON_WORDS(match/f7) + match/f6 + match/f11

- **Full title (not in common serials list) + date + main entry**

```
CONCAT(TRIM( Evaluate( 'regexp_replace(      regexp_replace(      regexp_replace(
regexp_replace(      regexp_replace(lower(%1),
"(^|\\s)(journal|bulletin|newsletter|report|directory|publication|proceedings|catalogue|calendar|minutes|
circular|fact\\s+sheet|update)(\\s|$)",      "\\1\\3"),
"(^|\\s)(budget|annual\\s+report|biennial\\s+report|final\\s+budget|financial\\s+report|technical\\s+report|st
atistical\\s+report|session\\s+laws|legislative\\s+calendar|legislative\\s+summary|report\\s+and\\s+account
s|proposed\\s+budget)(\\s|$)",      "\\1\\3"),
"(^|\\s)(memoires|memoria|revista|travaux|trudy|mitteilungen|anales|veroeffentlichungen|veroeffentlichun
gen|language\\s+sciences|telephone\\s+directory|year\\s+book|yearbook)(\\s|$)",      "\\1\\3"),
"(^|\\s)(advertiser|annual\\s+budget|annual\\s+financial\\s+report|annual\\s+reports?|annual\\s+report\\s+for
(\\s+the\\s+fiscal\\s+year\\s+ended)?|boletin|compilation\\s+of\\s+selected\\s+acts(\\s+within\\s+the\\s+jurisdic
tion\\s+of\\s+the\\s+committee\\s+on\\s+commerce)?|comprehensive\\s+annual\\s+financial\\s+report|conf
erence\\s+proceedings|conference\\s+record|general\\s+catalog|jahresbericht|laws,\\s+etc|membership\\s
+directory)(\\s|$)",      "\\1\\3"),
"(^|\\s)(monograph|monthly\\s+bulletin|news\\s+release|occasional\\s+papers?|proceedings\\s+of\\s+the\\s+
annual\\s+meeting|progress\\s+report|publicaciones|publications|rapport|rapport\\s+annuel|research\\s+r
eport|semi\\s+annual\\s+report\\s+to\\s+the\\s+congress|semiannual\\s+report\\s+to\\s+the\\s+congress|tech
nical\\s+bulletin|transactions|works)(\\s|$)",      "\\1\\3")      ; "Bibliographic Details"."Title (Normalized)"  ))
, CONCAT("Bibliographic Details"."Begin Publication Date", "Bibliographic Details"."Author"))
```

match/f7 + match/f6 + match/f11 + match/f10

Full title (can be in list of common titles) + date + main entry +
place of publication

```
CONCAT("Bibliographic Details"."Title  
(Normalized)", CONCAT("Bibliographic  
Details"."Begin Publication Date",  
CONCAT("Bibliographic  
Details"."Author",  
"Bibliographic Details"."Publication  
Place"))))
```

▲ Filters

- ▼ Num of Titles (Active) is greater than 1
- AND ▼ Bibliographic Level is equal to / is in s
- AND ▼ Lifecycle is equal to / is in In Repository
- AND ▼ Suppressed From Discovery is equal to / is in No
- AND ▼ Begin Publication Date is not null
- AND ▼ Author is not null
- AND ▼ Publication Place is not null
- AND ▼ Title (Normalized) is not null

REMOVE_COMMON_WORDS(match/f7) + match/f6 + match/f11 + match/f10

- **Fuzzy title (not in list of common titles) + date + main entry + place of publication**

```
CONCAT(TRIM( Evaluate( 'regexp_replace(      regexp_replace(      regexp_replace(
regexp_replace(      regexp_replace(lower(%1),
"(^\\s)(journal|bulletin|newsletter|report|directory|publication|proceedings|catalogue|calendar|minutes|
circular|fact\\s+sheet|update)(\\s|$)",      "\\1\\3"),
"(^\\s)(budget|annual\\s+report|biennial\\s+report|final\\s+budget|financial\\s+report|technical\\s+report|st
atistical\\s+report|session\\s+laws|legislative\\s+calendar|legislative\\s+summary|report\\s+and\\s+account
s|proposed\\s+budget)(\\s|$)",      "\\1\\3"),
"(^\\s)(memoires|memoria|revista|travaux|trudy|mitteilungen|anales|veroeffentlichungen|veroeffentlichun
gen|language\\s+sciences|telephone\\s+directory|year\\s+book|yearbook)(\\s|$)",      "\\1\\3"),
"(^\\s)(advertiser|annual\\s+budget|annual\\s+financial\\s+report|annual\\s+reports?|annual\\s+report\\s+for
(\\s+the\\s+fiscal\\s+year\\s+ended)?|boletin|compilation\\s+of\\s+selected\\s+acts(\\s+within\\s+the\\s+jurisdic
tion\\s+of\\s+the\\s+committee\\s+on\\s+commerce)?|comprehensive\\s+annual\\s+financial\\s+report|conf
erence\\s+proceedings|conference\\s+record|general\\s+catalog|jahresbericht|laws,\\s+etc|membership\\s
+directory)(\\s|$)",      "\\1\\3"),
"(^\\s)(monograph|monthly\\s+bulletin|news\\s+release|occasional\\s+papers?|proceedings\\s+of\\s+the\\s+
annual\\s+meeting|progress\\s+report|publicaciones|publications|rapport|rapport\\s+annuel|research\\s+r
eport|semi\\s+annual\\s+report\\s+to\\s+the\\s+congress|semiannual\\s+report\\s+to\\s+the\\s+congress|tech
nical\\s+bulletin|transactions|works)(\\s|$)",      "\\1\\3")      ; "Bibliographic Details"."Title (Normalized)" ))
, CONCAT("Bibliographic Details"."Begin Publication Date", CONCAT("Bibliographic Details"."Author",
"Bibliographic Details"."Publication Place"))))
```

REMOVE_COMMON_WORDS(match/f7)
+ match/f6 + match/f11 + match/f10 -
filters

 Filters

 Num of Titles (Active) is greater than 1

AND  Bibliographic Level is equal to / is in s

AND  Lifecycle is equal to / is in In Repository

AND  Suppressed From Discovery is equal to / is in No

AND  TRIM(EVALUATE('regexp_replace(... is not null

AND  Begin Publication Date is not null

AND  Author is not null

AND  Publication Place is not null

Questions about the full keys

- Does F6 date mean something different than F6 year?
- I'm filtering out any keys where one of the elements is null
- I'm also filtering out FUZZY_PAGINATION = 0 and Place of Publication - Country = No place, unknown, or undetermined
- For keys with pagination, I'm filtering out a bunch of Categories of material that pagination doesn't make sense (video, audio, etc)
- Why define Serial f4 (ISSN invalid) but never use it?

Concatenation is a little more complicated than this . . .

Ex Libris example for Dedup concatenation where key definition = match/f1 + match/f7:

- **f1**: Includes the following entries: a, b
- **f7**: Includes the following entries: c, d
- **Keys created**: ac, ad, bc, and bd



That's a lot of keys . . .
Now what?

How do we get from that to lists of titles that are deduping?

1. Filter on that report listing keys that have more than one hit
2. COUNT(DISTINCT "Bibliographic Details"."MMS Id" BY (the key))
 - And then filter that column to >1

The second option seems more elegant, but it's trickier to add all the other necessary filters (e.g. only want the key to count 2 distinct if both MMSids are active lifecycle and unsuppressed)

Doing the filtered report version

- Create a copy of your key report
- Save it with a new name
- Filter your key column based on the results of another analysis: select the original key report, and filter based on the original key column
- Delete the Num of titles more than 1 filter
- Add in the MMS id and any other columns you want. Sort on the key to bring the deduped titles together



That's when you want at
look at a specific key

Could we go the other way?

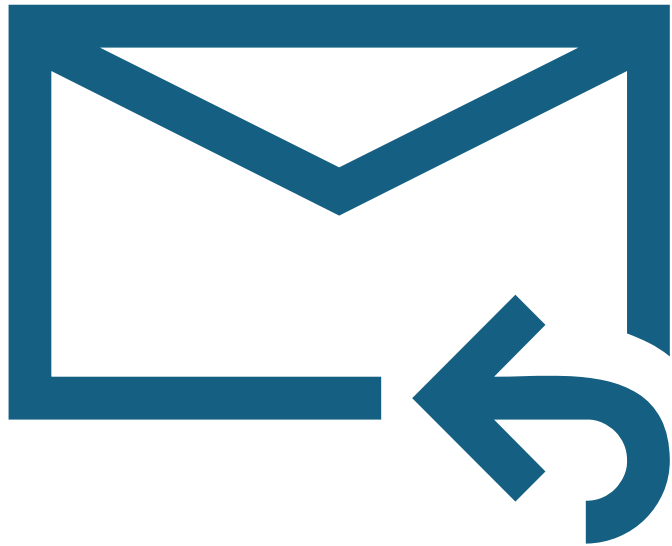
Input an MMS id (or a list of
them) and find out if there were
any dedup candidates?

My concept

- Make a report with one column for each *serial key
- Use CASE to check: if the MMS id matches an MMS id in the relevant key report, else leave the cell blank
- BONUS: have a drill down from the key that links to a report, filtered on the key, that lists all the MMSids (and whatever other data you want)

*you could also do all non-serial, but my non-serial reports rely on a local param, and also there are way more of them

Thanks for your time! Questions/feedback?



susan.bond@utoronto.ca

Useful Links

- [Ex Libris - some useful custom formulas for Analytics](#)
- [Evelyn's slides from Analytics Afternoon last year](#)
- [Ex Libris - Understanding the Dedup and FRBR process](#)
- [Analytics Titles Subject Area documentation](#)