

J-ISIS Release 1.2

Bug Fixes Since J-ISIS 1.2 Release Candidate

Synchronization between Data Viewer, Database Browser, and Data Entry was not working as expected and sometimes an exception was raised when adding/deleting or updating a record.

PrintSort module: some HTML header text was printed at the beginning when printing using Plain Text format and “Don’t use sorting” radio button is selected.

Enhancements Since 1.2 J-ISIS Release candidate

The real record number and Next Master File Number to be assigned are now displayed in the Data Viewer.

Many scalability improvements for huge database (i.e. more than 5 million records), the most significant are:

- i) The number of records in the database is now stored in the metadata database with the last assigned MFN. The number of records is updated by J-ISIS when new records are created and when records are deleted. This avoid to ask Berkeley DB to go through the whole database btree for counting the records when this value is needed, and there are many places where it is needed.
- ii) A wait cursor is displayed when loading a page of records in the database browser or a page of terms in the Dictionary of Terms.

J-ISIS 1.2 Release Candidate

I. Fixes to the J-ISIS Print Format

Thanks to Ivan Batrak from the Syberian Federal University library team who was able to hack the J-ISIS Print Formatting engine, and test it with many PFT scripts used by the Russian irbis64 software. Furthermore, he took the time to report the bugs and even to propose solutions, i.e. changes in the source code to fix the bugs.

He also suggested the use of Global variables into the J-ISIS Print Formats which is used extensively in the irbis64 software PFT scripts.

Test Record 1

RECORD(1)

Tag	Field/Occurrence
270:	<<^aLeningrad^d1966>>
270:	<<^aPuškin>>

1) Repeatable literals were not working as expected with field dummy selectors (D or N)

|Hello|d270 was producing an empty string even if field 270 was present

After bug fix:

Format	Result
1 Hello d270	Hello
1 (Hello d270	HelloHello

2) Conditional literals with subfield dummy selectors (D or N)

"Hello"d270^d was always producing Hello as output even if no subfield ^d was present

After bug fix:

Format	Result
1 /*bug 2 */ 2 "Hello bug 2"d270^d	Hello bug 2
1 /*bug 2 */ 2 "Hello bug 2"d270[1]^d	Hello bug 2
1 /*bug 2 */ 2 "Hello bug 2"d270[2]^d	No output

Same for "Hello"n270^d, after bug fix:

Format	Result
1 /*bug 2 */ 2 "Hello bug 2"n270^d	No output
1 /*bug 2 */ 2 "Hello bug 2"n270[1]^d	No output
1 /*bug 2 */ 2 "Hello bug 2"n270[2]^d	Hello bug 2

3) MFN command was raising an error in REF function expressions like:

```
ref(mfn,
  if p(v19) and v19^x<='0'then", "d963^i,
  (if v19^x<='0'then|<b>|v19^a*2|</b>|,| |v19^b fi)
  fi,
)
```

After bug fix:

1 ref(mfn, 2 if p(v19) and v19^x<='0'then", "d963^i, 3 (if v19^x<='0'then v19^a*2 , v19^b fi) 4 fi, 5)	SIS FMT Parser: PFT program parsed successfully.
--	--

4) Extracting a fragment of a Subfield specifying only the offset (*offset) was not working

v270^a*2 for example

After bug fix:

Format	Result
1 v270^a*2	ningradskin
1 v270[1]^a*2	ningrad
1 v270[2]^a*2	skin

5) String function F(*expr-1* ,*expr-2* ,*expr-3*)

The **F** function converts a numeric value from its internal floating-point representation to a character string. The three arguments are all numerical expressions. The first argument, *expr-1*, is the number to be converted. The second argument, *expr-2* is the minimum output width and the third argument *expr-3* is the number of decimal places. *expr-2* gives the *minimum* width, i.e. the function value will be a character string of at least *expr-2* characters.

If *expr-2*, and *expr-3* were missing, the output string was not right adjusted in an output string with a default width as specified in the J-ISIS Reference Manual.

After bug fix, if *expr-2* and *expr-3* are missing, a default width of 10 characters will be used. Please note that in WinISIS Reference manual, the default width is 16 characters. Please let me know if you think that J-ISIS should also use a default width of 16 characters.

After bug fix:

Format	Result
1 F(1)	1
1 F(3.14116,10,5)	3.14116
1 F(1,10,5)	1.00000

6) String functions S, SS, and CISIS functions LEFT, MID, REPLACE, and RIGHT were not working in repeatable group.

For example

```
(if s(v270^d) <> '1966' then '****' else '1966' fi/)
```

After bug fix:

Format	Result
1 (if s(v270^d) <> '1966' then '****' else '1966' fi/)	1966 ****

7) New Print Format Command for Unconditional Literals <text> ...</text>

Plain text or most probably HTML formatting can now be imbedded between the <text> and </text> tagging commands, it works like unconditional literals.

Format	Result
1 <text> 2 aa</br> 3 bb</br> 4 cc 5 </text>	aa bb cc

II. Print Format for Repeatable Subfields

Test Record 2

RECORD(2)	
Tag	Field/Occurrence
270:	<<^aAA^aBB>>
270:	<<^aCC^aDD>>

Subfield occurrences

It is possible to access individual occurrences of a repeatable subfield by specifying the occurrence number or range, enclosed in square brackets, immediately following the field selector or field selector followed by occurrence selector. For example:

Format	Result
1 v270[1]^a[2],v270[1]^a[1])	BBAADDCC

It is possible to display specific occurrence of a repeatable subfield, narrowing the output to one or a range of occurrences of a repeatable subfield by specifying the occurrence number or range, enclosed in square brackets, immediately following the field selector.

For example:

v10^a[1]	retrieves the subfield first occurrence of subfield ^a
v10^a[2..4]	retrieves the 2nd through the 4th occurrence of subfield ^a from field 10
v10^a[3..]	retrieves the 3rd through the last occurrence of subfield ^a from field 10
v10[1]^a[1]	retrieves 1 st occurrence of subfield ^a in the 1st occurrence of field 10
V10[1]^a[1..]	Retrieves all occurrences of subfield ^a

It is coded as follows:

[<index> [..<upper index>]]

<index> and **<upper index>** refer to the first (or unique) and last occurrences, respectively. If the specified **<index>** is greater than the actual number of occurrences, no

output is generated. The same occurs if data subfield is not repeatable and **<index>** is set to a number equal or greater than 2. However, if **<index>** is set to 1 and it is used in a non-repeatable subfield, content is normally output. This component must be used outside a repeatable group; otherwise, **<upper index>** is ignored. If double dot (..) is used and **<upper index>** is missing LAST is assumed. The LAST keyword is set with the value of total occurrences of a data subfield.

The syntax grammar for an optional subfield occurrence is defined as follow:

```
subfield_occur_opt ::=
  LBRACK INT:n1 DOT_DOT INT:n2 RBRACK
    { :RESULT=newFieldRepeat(n1.intValue(),n2.intValue()); :}
|  LBRACK INT:n DOT_DOT RBRACK
    { :RESULT = new FieldRepeat(n.intValue(),-1); :}
|  LBRACK INT:n RBRACK
    { :RESULT = new FieldRepeat(n.intValue(),n.intValue()); :}
|
;
```

Format	Result
1 (v270^a[2]/)	BB DD
1 (v270^a[1..]/)	Leningrad Puškin on record 1 AABB CCDD on record 2

III. Print Format Global Variables

Global variables are stored in a virtual ISIS record which is a collection of fields, fields may be repeatable and have occurrences, and fields or occurrences may have subfields. The record, field and subfield concepts are identical to ISIS.

Global variables are referenced by the letter **G** followed by the **tag** of the field. The G (a mnemonic code for Global variable) followed by the virtual record tag is the command telling J-ISIS that you want to assign or extract a field. It may be entered indifferently in upper or lower case.

Global variables can be assigned data through the Print Format commands:




```
g100:=(v25/), (g100^a/)
g10 := (v10^a)
```

You may assign or change the value of a global variable as follows:

Gn:=(format) (for example: **G5:=(v10)**).

Note that the parentheses around *format* are required.

Global variables can be extracted for output like V variables just by replacing the V by G that means that data will be extracted from the virtual record. It supports repeatable groups as well.

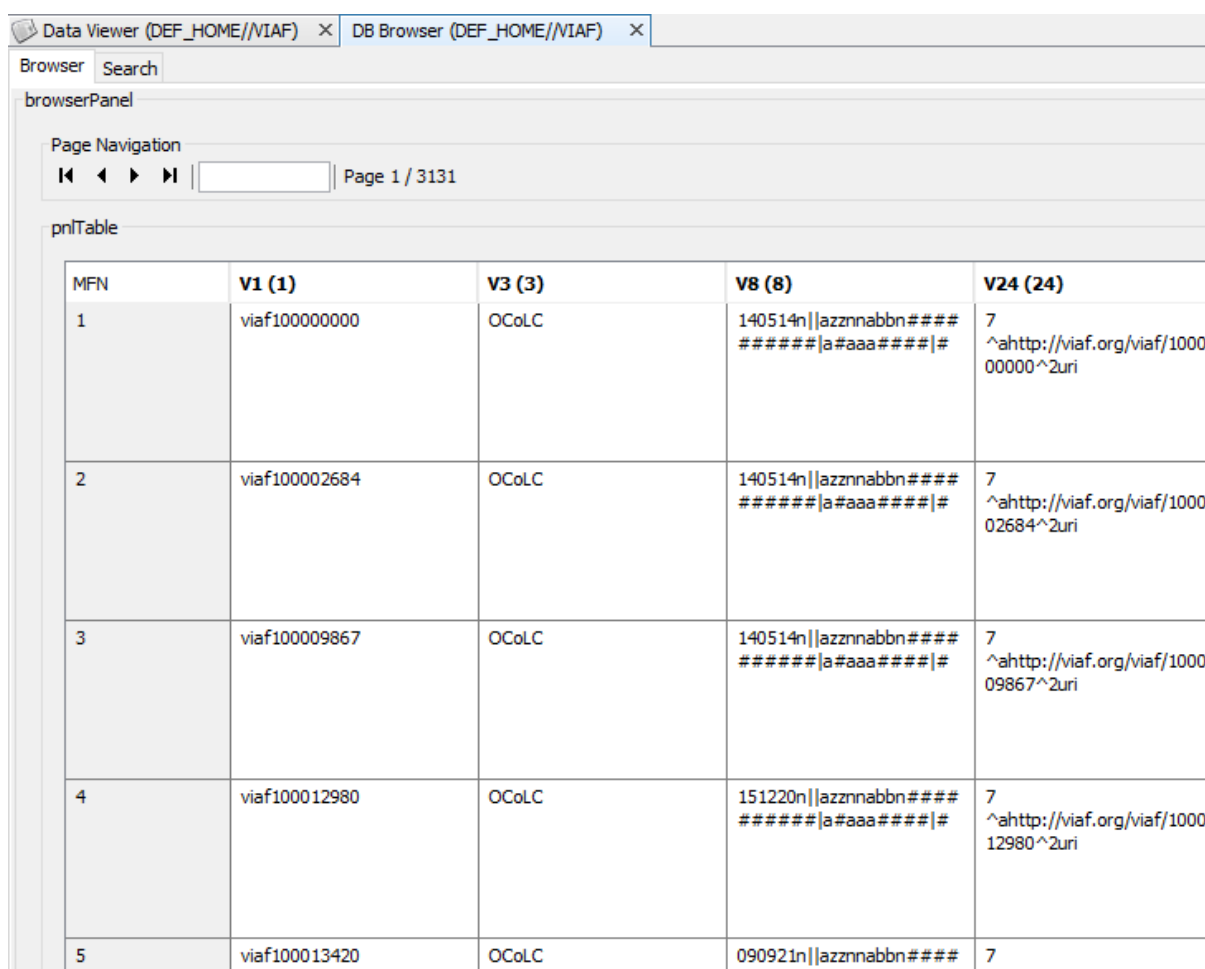
Format	Result
 <code>g3:=((v270,'</br>')),g3</code>	 <code>^aLeningrad^d1966</code>  <code>aPuškin</code>

Please note that it is a first attempt to implement Global variables and that specific functions could also be implemented to further manipulate them. Please let me know if it is worth to continue working in this direction.

IV. New Paging feature into DB Browser and Terms Dictionary

Databases could be huge. If a database has millions of records and all records are loaded into memory, it will consume a huge amount of memory and will of course be very slow. As a matter of facts, user will probably only look at 10 or maybe 20 records depending on the viewport size, there is no need to download all the records locally. That's the reason why the paging feature was introduced into the DB browser and Terms Dictionary Browser modules. To make it easy to use the Paging feature, a page navigation toolbar provides the interface to do the navigation.

10 000 records are loaded per page and the user can scroll easily and fast through the page records. For example, the VIAF database has near 32 million records (31 305 939 records exactly)



The screenshot shows a web application window titled "DB Browser (DEF_HOME//VIAF)". It features a "Page Navigation" toolbar with buttons for first, previous, next, and last page, along with a text input for page number and a "Page 1 / 3131" indicator. Below the toolbar is a table with the following data:

MFN	V1 (1)	V3 (3)	V8 (8)	V24 (24)
1	viaf100000000	OCOLC	140514n jazznnabbn#### ##### a#aaa#### #	7 ^ahttp://viaf.org/viaf/1000 00000^2uri
2	viaf100002684	OCOLC	140514n jazznnabbn#### ##### a#aaa#### #	7 ^ahttp://viaf.org/viaf/1000 02684^2uri
3	viaf100009867	OCOLC	140514n jazznnabbn#### ##### a#aaa#### #	7 ^ahttp://viaf.org/viaf/1000 09867^2uri
4	viaf100012980	OCOLC	151220n jazznnabbn#### ##### a#aaa#### #	7 ^ahttp://viaf.org/viaf/1000 12980^2uri
5	viaf100013420	OCOLC	090921n jazznnabbn#### ##### a#aaa#### #	7 ^ahttp://viaf.org/viaf/1000 13420^2uri

The same paging mechanism is used for browsing the Terms Dictionary so that you can move fast from page to page as well as from the 1st to the last and vice versa.

Data Viewer (DEF_HOME//VIAF) × Dictionary (DEF_HOME//VIAF) × Search Window (DEF_HOME//VIAF) ×

Index name: C:\jisis-workspace\home_test_db\VIAF\indexes\master

Number of fields: 3

Number of Records: 31305939

Number of terms: 32501914

Last modified: Thu Jun 08 22:16:56 CEST 2017

Index Content

Page 1/3251

iTerm	Field	Term	Freq
0	_100	KUCKEN, HANS-ALBRECHT	1
1	_100	NICOLAS ROMAN, SUSANA	1
2	_100	NICOLAS ROMAN, SUSANA	1
3	_100	TORRES NUNEZ, JUAN JOSE	1
4	_100	\$PWATTANA SUKEEPAISARNJAROEN	1
5	_100	'ABDUL-BAHA,	1
6	_100	'SGRAVESANDE, WILLIAM JAMES,	1
7	_100	*N NTRAW	1
8	_100	ABD AL-GABBAR AL-HAMADANI (937-10257).	1
9	_100	ABD AL-HAYY	1
10	_100	ABD AL-RAHMAN KATKHUDA	1
11	_100	ABD AL-RAHMAN KATKHUDA,	1
12	_100	ABID	1
13	_100	ABRAHAM BEN MOSHEH BEN MAIMON,	1
14	_100	ABRAMSON, HAROLD J.	1
15	_100	ABREU, MARIA HELENA PAIS DE,	1
16	_100	ABU 'L-FAID MUHAMMAD MURTADA AL-ZABIDI	1
17	_100	ABU AL-NAGA, ABU AL-MA 'ATI	1
18	_100	ABU NASR 'ABD AL-WAHAB,	1
19	_100	ACEVEDO BUENO, JULIA	1
20	_100	ACHUCARRO, ANA,	1
21	_100	ADDISON, FREDERICK AMOAKO	1
22	_100	ADIB ZADAH, MAGID,	1
23	_100	AHMAD IBN FADLAN IBN AL-'ABBAS,	1
24	_100	AJIBADE, A. J.	1
25	_100	AL-'ISHKAWARI, SADIQ AL-HUSAYNI	1
26	_100	ALAWI, SA ID BINS ID,	1
27	_100	ALBARRAN LOZANO, IRENE	1
28	_100	ALBUQUERQUE, SANDRA MARIA FERREIRA DE	1

Quick Search

Search: <All Searchable Fields>

Query:

Index name:

C:\jisis-workspace\home_test_db\VIAF\indexes\master

Number of fields:

3

Number of Records:

31305939

Number of terms:

32501914

Last modified:

Thu Jun 08 22:16:56 CEST 2017

Quick Search

Search:

<All Searchable Fields>

Query:

Index Content

Page 3251/3251

Item	Field	Term	Freq
32500000	_105	鳥取県歴史散歩研究会	1
32500001	_105	鳥取県水産組合	1
32500002	_105	鳥取県水産試験場	1
32500003	_105	鳥取県海外協会	1
32500004	_105	鳥取県漁業協同組合連合会	1
32500005	_105	鳥取県物産陳列場	1
32500006	_105	鳥取県獣医師会	1
32500007	_105	鳥取県産業技術センター	1
32500008	_105	鳥取県産業振興機構	1
32500009	_105	鳥取県農業普及協会	1
32500010	_105	鳥取県町村議会事務協議会	1
32500011	_105	鳥取県町村議会議長会	1
32500012	_105	鳥取県畜産試験場	1
32500013	_105	鳥取県病害虫防除所	1
32500014	_105	鳥取県病院薬剤師会	1
32500015	_105	鳥取県監査委員	1
32500016	_105	鳥取県知的所有権センター	1
32500017	_105	鳥取県社会福祉協議会	1
32500018	_105	鳥取県神社庁	1
32500019	_105	鳥取県神職会	1
32500020	_105	鳥取県私立教育会	1
32500021	_105	鳥取県産産場	1
32500022	_105	鳥取県産産場 (1964年)	1
32500023	_105	鳥取県船舶検査所	1
32500024	_105	鳥取県立中央病院	1
32500025	_105	鳥取県立小企業講習所	1
32500026	_105	鳥取県立保育専門学校	1
32500027	_105	鳥取県立倉吉高等学校野球部	1
32500028	_105	鳥取県立公文書館	1

Data Viewer (DEF_HOME//VIAF) × Dictionary (DEF_HOME//VIAF) × Search Window (DEF_HOME//VIAF) ×

☒ Match all of the following
 ☐ Match any of the following
 Dictionary Options
 ☐ Disable suggestions
 ☒ Guided search

<All Searchable Fields>
 matching
 鳥取県社会福祉協議会
 + -

Clear Search

1 records ☐ Sort By MFN

MFN

[1] 26097964

PgUp PgDn

Select PFT: RAW

375:	<< ^aunknown^20>>
410:	<<2 ^aTottoriken Fukushima Jinzai Senta^2NDL>>
410:	<<2 ^aTottoriken Shakai Fukushima Kyogikai^2NDL>>
410:	<<2 ^aTottoriken Shakai Fukushima Kyogikai Tottoriken Fukushima Jinzai Senta^2NDL>>
410:	<<2 ^aトットリケン シャカイ フクシ キョウギカイ^2NDL>>
410:	<<2 ^aトットリケン シャカイ フクシ キョウギカイ トットリケン フクシ シンザイ センター ^2NDL>>
410:	<<2 ^aトットリケン フクシ シンザイ センター ^2NDL>>
410:	<<2 ^a鳥取県社会福祉協議会鳥取県福祉人材センター^2NDL>>
410:	<<2 ^a鳥取県福祉人材センター^2NDL>>
710:	<<2 ^a鳥取県社会福祉協議会^0(NDL)00306971>>
856:	<<40^uhttp://viaf.org/viaf/255085625>>
856:	<<42^uhttp://viaf.org/viaf/data^zDataset description>>

100%

IV Export features to select search results and using a hit file to drive output are now implemented

J-ISIS Export Wizard

Steps

1. Select the Output Format
2. Parameters

ISO2709 Export - Output Parameters. (1.1)

Output ISO File

Name of Output ISO File:

Output Directory: C:\J-ISIS 2016\jisis\work

Export:

☒ MFN Range
 ☒ All
 ☐ MFNs
 Enter MFNs and/or MFN ranges separated by commas. For example 1,10,100-150,50

☐ Marked
 No Marked Sets

☐ Search Results
 From Query: SearchResult(searchNumber=1dbName=CDS mfn=[57, 73, 75]searchQuery=Search All Fields matching VOL

☐ Use this Hit Sort File for driving the output: C:\J-ISIS 2016\jisis\work\CDS.hitsort

Options

(Use zero for no limit)

Output Line Length:

Reformatting FST:

Renum Records From MFN:

Output Tag Containing MFN:

Encoding: UTF-8

Separators

Field Terminator
 Record Terminator
 Subfield Delimiter

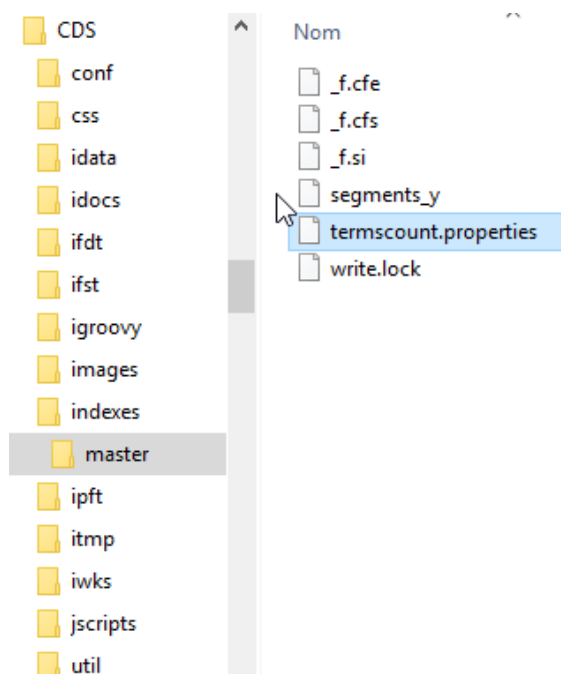
< Back Next > Finish Cancel Help

You can now export records retrieved from search as well as export records following the order defined by a hit file produced by the PrintSort module

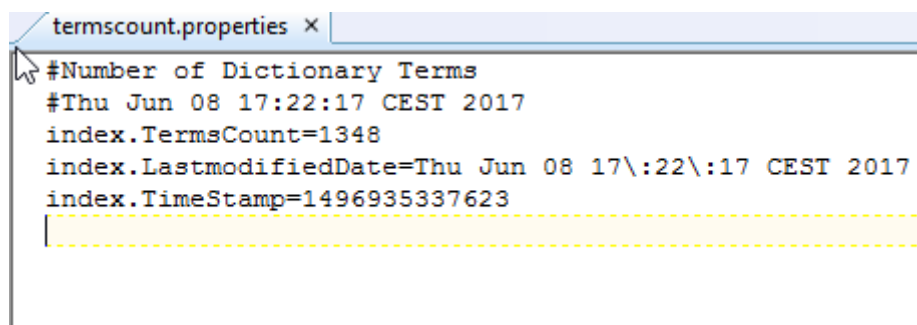
Note: A hit file manager will be developed in the future to better manage search hit files and hit sort files

V. The Number of Terms in the index is now stored in an external file to avoid the time consuming task of counting them.

The **/indexes** directory contains a subdirectory called master that contains the main index files generated by [Lucene open-source search software](#). A new file named “termscount.properties” is now generated by J-ISIS to keep the number of terms in the index as well as a time stamp, and is stored in the /indexes/master folder. The number of terms in the index is only computed when the index has changed and replaced with the new time stamp in the external file.



The content of the generated file looks like this.



For databases with more than 2 millions records, it reduces considerably the time spent to get the database information.