Baby Bases
Care and Feeding of Z0102 Sub-bases

Jane Aitkens

Aleph Systems Librarian

McGill University Libraries

June 2, 2003 - NAAUG

The Quest – Offspring

- **To create sub-bases for each McGill Library (we have 17)**
- **To create sub-bases by type of material: journals, electronic journals, audio-visual, eBooks, McGill theses**

The problem with simply creating a logical base for a small set of records – retrieval is slow, and the browse lists have too many unrelated see references.

The Solution: Z0102 headings
(patch 4, 5 rep_changes)

- **Step 1 – genetic engineering**
  Use the new tab_type_config table to identify materials such as electronic journals

- **Step 2 – birthing the new creations**   **Z0102 description**
  Use the new Z0102 table to build headings for small sub-bases (14.2's Z0102 bases have no see references; they are optional in 15)

- **Step 3 – constant feeding**
  Keeping the Z0102 table up to date

- **Step 4 – dressing up**
  Adjust the Web OAPC screens to identify the use of a sub-base, customize brief table screen per sub-base

Step 1 – Genetic Engineering (table setup)

BIB01/-tab
(for material TYPe identification - eJournals, eBooks, etc)

- **tab_type_config** *– create the TYP field, invent ESL field*
- **tab11**, tab00 *– add indexes*
- **tab_expand** *– expand TYP into index and displays*
- **edit_doc_999.eng** *– add fields so they display*
- edit_field.eng *– define the display output*
- **www_tab_short.eng** *– add TYP as a new column*
- tab_01 *– define the ESL as a valid field*
- pc_tab_sear *– add indexes to GUI search client*

And (checks and tag are optional – if you want to control data entry)
- **check_doc_tag_text**
- check_doc_line

BIB01/pc_tab_catalog
- **tag_text.dat**

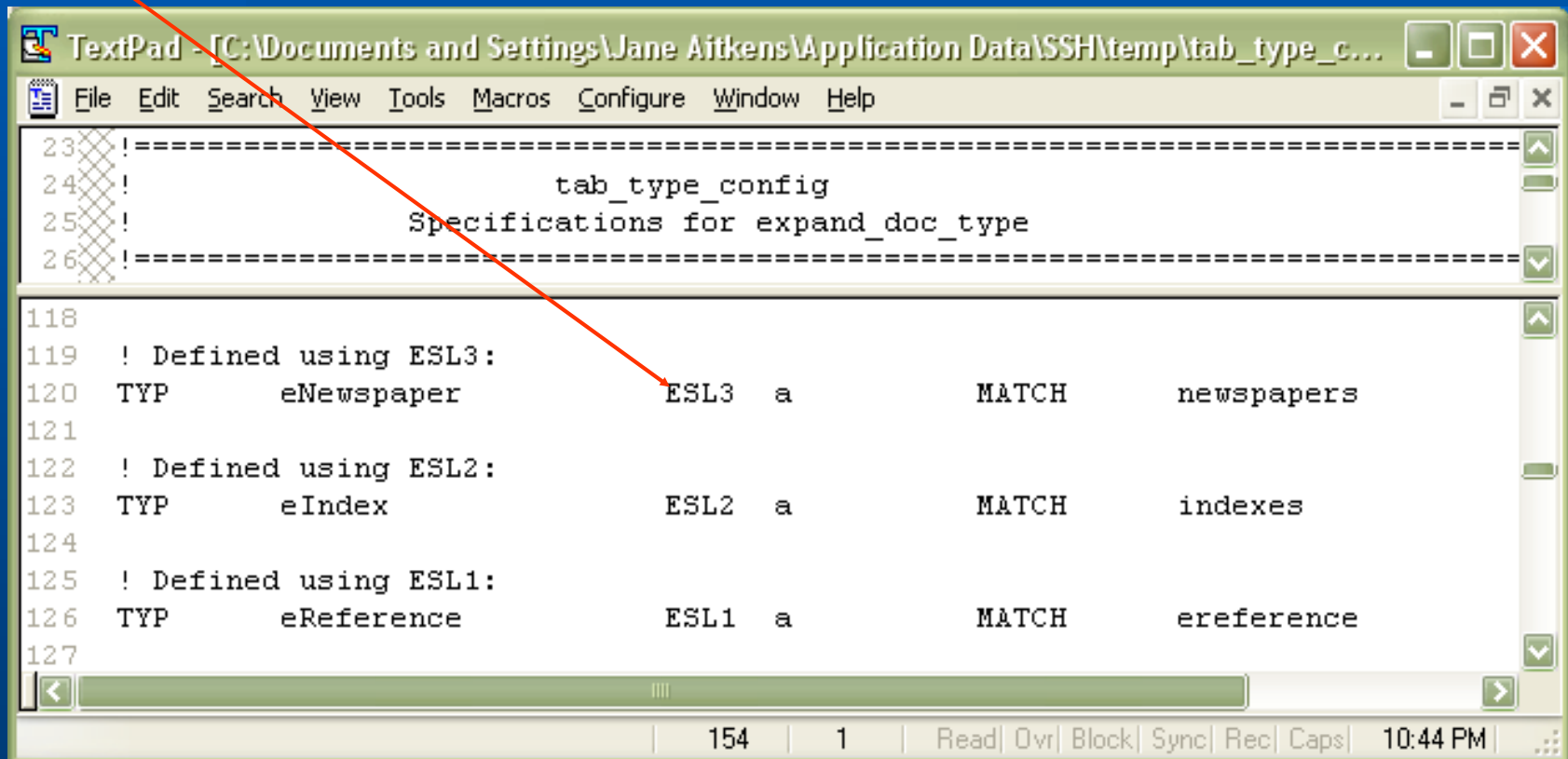- codes.eng *– add ESL as a valid field:*
  - ESL   Y N N L E-Subject Label L Electronic Subject Label

4

tab_type_config - using LDR, 008

```
TextPad - [C:\Documents and Settings\Jane Aitkens\Application Data\SSH\temp\tab_type_co...
  File   Edit   Search   View   Tools   Macros   Configure   Window   Help

23 !==========================================================================
24 !                      tab_type_config
25 !             Specifications for expand_doc_type
26 !==========================================================================

156    TYP        Journal              LDR    F06-01      EQUAL      a
157                                    LDR    F07-01      EQUAL      [b,s]
158                                    008    F23-01      N-EQUAL    a
159                                    008    F23-01      N-EQUAL    b
160                                    008    F23-01      N-EQUAL    c
161                                    008    F23-01      N-EQUAL    s
162                                    008    F21-01      EQUAL      p
163
164    TYP        Newspaper            LDR    F06-01      EQUAL      a
165                                    LDR    F07-01      EQUAL      [b,s]
166                                    008    F23-01      N-EQUAL    a
167                                    008    F23-01      N-EQUAL    b
168                                    008    F23-01      N-EQUAL    c
169                                    008    F23-01      N-EQUAL    s
170                                    008    F21-01      EQUAL      n
171
172    ! Use serial when it is not one of the more specific labels:
173    TYP        Serial               LDR    F06-01      EQUAL      a
174                                    LDR    F07-01      EQUAL      s
175                                    008    F21-01      N-EQUAL    n
176                                    008    F21-01      N-EQUAL    p
177                                    008    F23-01      N-EQUAL    a
178                                    008    F23-01      N-EQUAL    b
179                                    008    F23-01      N-EQUAL    c
180                                    008    F23-01      N-EQUAL    s

For Help, press F1              154        1      Read Ovr Block Sync Rec Caps   10:36 PM
```

tab_type_config – using ESL field

ESLs must be distinct, to get around multiple occurrence bug – see PRB2528

```
TextPad - [C:\Documents and Settings\Jane Aitkens\Application Data\SSH\temp\tab_type_c...

 File   Edit   Search   View   Tools   Macros   Configure   Window   Help

23 !=======================================================================
24 !                        tab_type_config
25 !              Specifications for expand_doc_type
26 !=======================================================================

118
119   ! Defined using ESL3:
120   TYP        eNewspaper              ESL3   a          MATCH       newspapers
121
122   ! Defined using ESL2:
123   TYP        eIndex                  ESL2   a          MATCH       indexes
124
125   ! Defined using ESL1:
126   TYP        eReference              ESL1   a          MATCH       ereference
127

                                              154      1     Read Ovr Block Sync Rec Caps   10:44 PM
```

6

tab_11 – create WTP, WSW, SUW

tab_00 – name WTP, WSW, SUW

TextPad - [D:\work\Aleph Tables\mgu01\tab\tab00.eng]

File   Edit   Search   View   Tools   Macros   Configure   Window   Help

```
122  H SUC    ACC    1 01 00      0000      Cutter Subjects
123  H SUE    ACC    1 11 00      0000      NLC English Subjects
124  H SGN    ACC    1 01 00      0000      Genre/Form
125  H PLA    ACC    1 01 00      0000      Hierachical Place
126  H SUW    ACC    1 01 00      0000      E-Subject Label
127
```

```
150   H WFM    W-025    04 00      0039      WFM [format]
151   H WPD    W-026    04 00      0039      WPD [007 type]
152   H WFS    W-027    04 00      0039      WFS [006 type]
153   H WST    W-028    04 00      0039      WST [status]
154   H WTP    W-029    04 00      0039      WTP [type]
155   H WSW    W-030    04 00      0030      WSW [esubject]
156   H WAD    W-031    04 00      0030      WAD [accession_date]
```

156   58   Read Ovr Block Sync Rec Caps

tab_expand - adding TYP to Word index and displays

```
!!!!!!!!!-!!!!!!!!!!!!!!!!!!!!!!!!!!!-!!!!!!!!!!!!!!
U39-DOC      expand_doc_type                    tab_type_config
WORD         expand_doc_type                    tab_type_config

WEB-BRIEF    expand_doc_type                    tab_type_config
WEB-FULL     expand_doc_type                    tab_type_config
WEB-FULL-1   expand_doc_type                    tab_type_config

GUI-BRIEF    expand_doc_type                    tab_type_config
GUI-DOC-D    expand_doc_type                    tab_type_config

TAB33-DOC    expand_doc_type                    tab_type_config
PRINT-CUST   expand_doc_type                    tab_type_config

may be needed:
RET          expand_doc_type                    tab_type_config
```

Edit_doc_999 and edit field.eng

www_tab_short.eng

```
7 L Type                              S              ## TYP##

8 L Location Holdings                 S              SE HDURL
8 L Holdings                          J              SE 866##
8 L Holdings                          J              SE 866##
8 L Holdings                          J              SE 866##
8 L Holdings                          J              SE 866##
8 L Holdings                          J              SE 866##
8 L Holdings                          S              CF HDURL
8 L Holdings                          J              CF 866##
8 L Holdings                          J              CF 866##
8 L Holdings                          S              ##
```
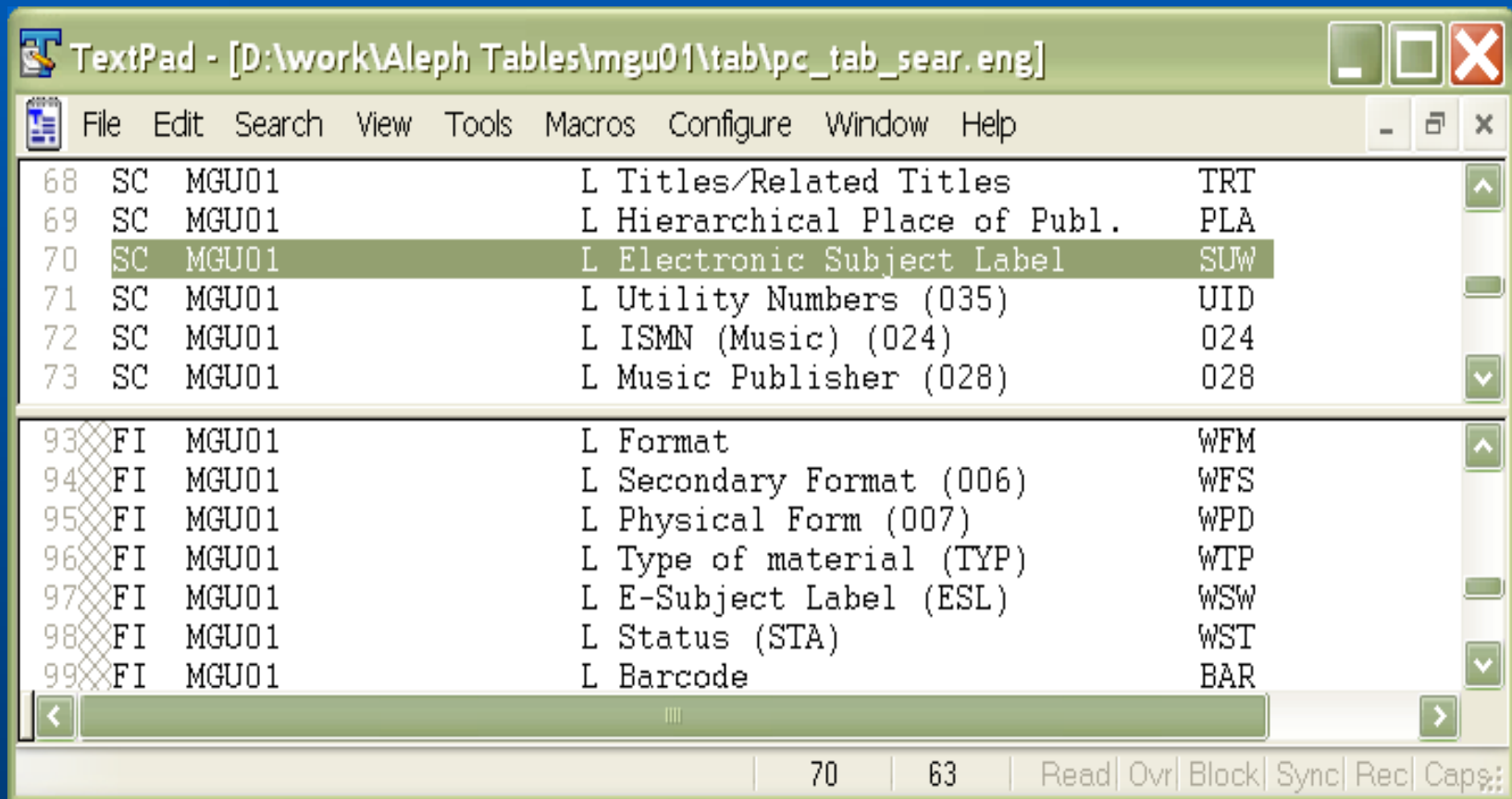
Tab_01 – add ESL to MARC definitions

pc_tab_sear   SUW entries

TextPad - [D:\work\Aleph Tables\mgu01\tab\pc_tab_sear.eng]

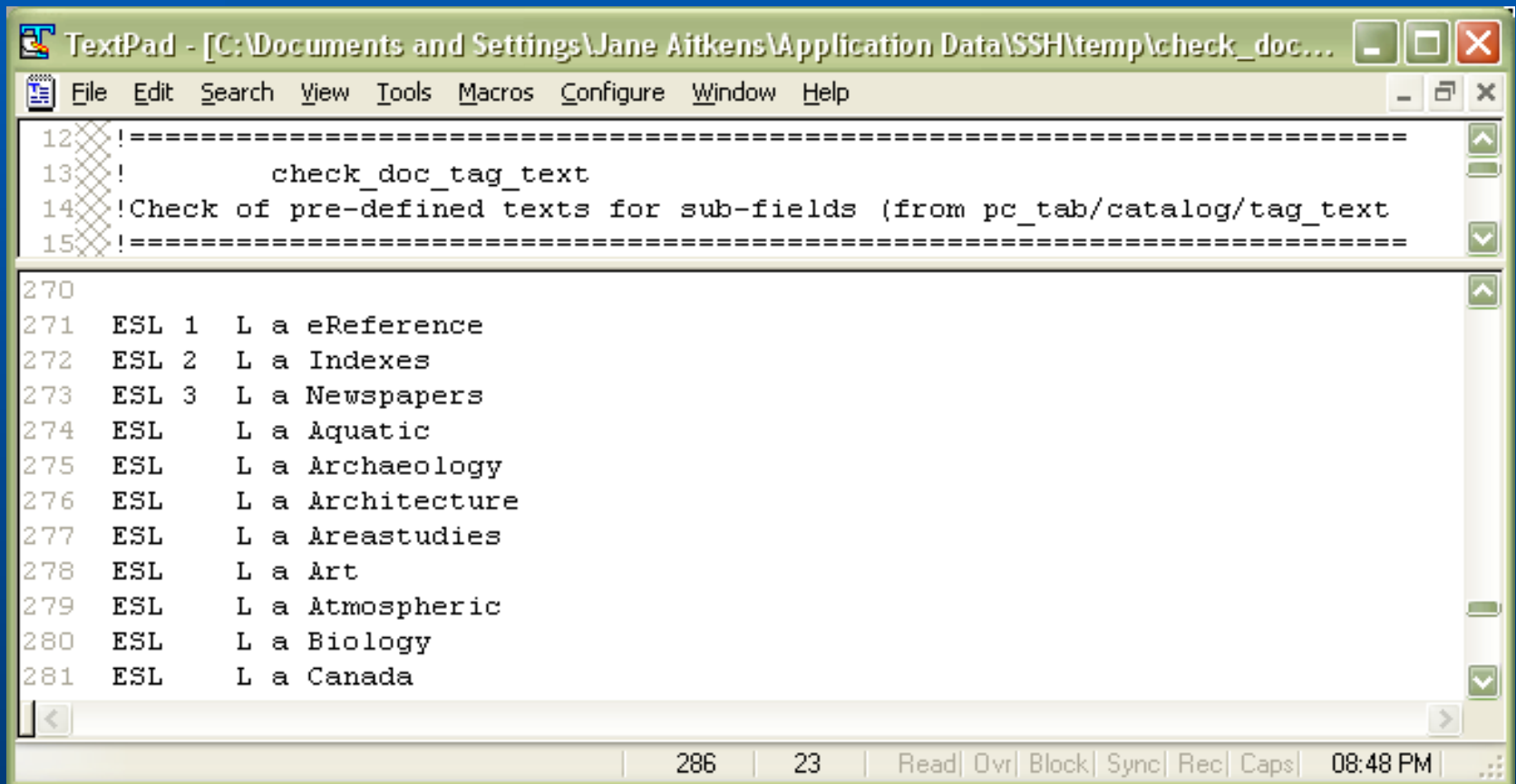File   Edit   Search   View   Tools   Macros   Configure   Window   Help

```
68   SC   MGU01              L Titles/Related Titles        TRT
69   SC   MGU01              L Hierarchical Place of Publ.   PLA
70   SC   MGU01              L Electronic Subject Label      SUW
71   SC   MGU01              L Utility Numbers (035)         UID
72   SC   MGU01              L ISMN (Music) (024)            024
73   SC   MGU01              L Music Publisher (028)         028
```

```
93  FI   MGU01              L Format                        WFM
94  FI   MGU01              L Secondary Format (006)        WFS
95  FI   MGU01              L Physical Form (007)           WPD
96  FI   MGU01              L Type of material (TYP)        WTP
97  FI   MGU01              L E-Subject Label (ESL)         WSW
98  FI   MGU01              L Status (STA)                  WST
99  FI   MGU01              L Barcode                       BAR
```
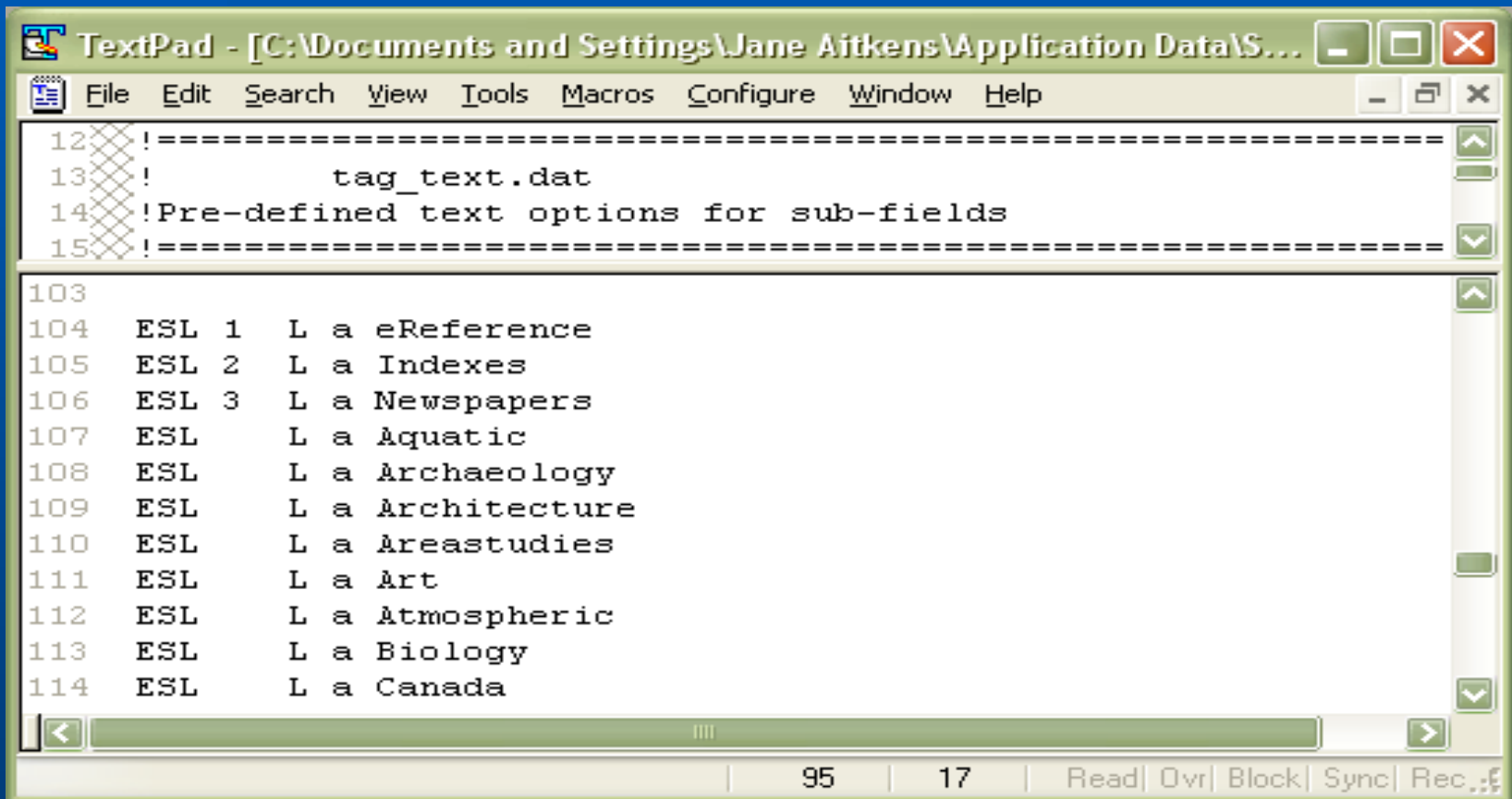
70   63   Read Ovr Block Sync Rec Caps

Check_doc_tag_text

tag_text.dat in pctab/catalog

```
12 !=======================================================
13 !        tag_text.dat
14 !Pre-defined text options for sub-fields
15 !=======================================================

103
104    ESL 1    L a eReference
105    ESL 2    L a Indexes
106    ESL 3    L a Newspapers
107    ESL      L a Aquatic
108    ESL      L a Archaeology
109    ESL      L a Architecture
110    ESL      L a Areastudies
111    ESL      L a Art
112    ESL      L a Atmospheric
113    ESL      L a Biology
114    ESL      L a Canada
```

TextPad - [C:\Documents and Settings\Jane Aitkens\Application Data\S...

File Edit Search View Tools Macros Configure Window Help

95 | 17 | Read| Ovr| Block| Sync| Rec

check_doc_line

TextPad - [D:\work\Aleph Tables\mgu01\tab\check_doc_line]

File Edit Search View Tools Macros Configure Window Help

D:\work\Aleph Tables\mgu01\tab\check_doc_line

```
12 !===================================================================
13 !         check_doc_line
14 !Check of logical structure/connections within a single occurrence of a field
15 !===================================================================

2402 
2403 AL XX      ESL    -
2404 AL XX      ESL    - 1
2405 AL XX      ESL    - 2
2406 AL XX      ESL    - 3
2407 AL XX      ESL    a 1 1        a 1 1 = mandatory, not repeatable
2408
```

2402   1   Read Ovr Block Sync Rec Caps   06:44

Step 2 – Birth (creating the Z0102)

- **Get new WTP index working: either re-index the keyword index or pull sets according to the new Types using p_ret_01, and submit to p_manage_40 (Update Indexing for Selected Records)**

- **tab_base.eng – name the babies**

- **Run Util H 1 10 – the Z0102 Setup checker
  Output called library_report. Examine closely**

- **tab_base_z0102 – mark them for entry into the Z0102 table**

- **Run p_manage_32 to create the Z0102 Oracle table**

tab_base.eng

```
! Type sub-bases

MCGILL_SERIALS         Journal Titles        MGUTP MGUS1 MGU01 MGU01 (wtp=journal or wtp=ejournal or
                                                 wtp=enewspaper or wtp=newspaper or wtp=microform serial)
MCGILL_THESES          McGill Theses         MGUTP MGU01       MGU01 (wti=mcgill !1 theses) not
EBOOKS                 E-Books               MGUTP MGU01       MGU01 wtp=ebook not (wst=deleted
EJOURNALS              E-Journals            MGUTP MGU01       MGU01 wtp=ejournal not (wst=dele
EINDEXES               E-Indexes             MGUTP MGU01       MGU01 wtp=eindex not (wst=delete
EREFERENCE             E-Reference           MGUTP MGU01       MGU01 wtp=ereference not (wst=de
ERESOURCES             E-Resources           MGUTP MGU01       MGU01 wtp=eresource not (wst=del

! Sub-library bases

ALLAV                  All AV Collections    MGUSB MGU01       MGU01 (wcl=av or wcl=icc) not ((
BWOOD                  Blacker-Wood          MGUSB MGU01       MGU01 wcl=bwood not (wst=deleted
EDUC                   Education             MGUSB MGU01       MGU01 wcl=educ not (wst=deleted
GDOC                   Government Docs       MGUSB MGU01       MGU01 wcl=gdoc not (wst=deleted
HLTH                   Health Sciences       MGUSB MGU01       MGU01 wcl=hlth not (wst=deleted
HRMG                   Howard Ross           MGUSB MGU01       MGU01 wcl=hrmg not (wst=deleted
ISLM                   Islamic Studies       MGUSB MGU01       MGU01 wcl=islm not (wst=deleted
LAW                    Nahum Gelber Law      MGUSB MGU01       MGU01 wcl=law  not (wst=deleted
MCL                    Humanities & Soc Sci  MGUSB MGU01       MGU01 (wcl=mcl or wcl=cdas) not
MCLREF                 HumanitiesSocSci REF  MGUSB MGU01       MGU01 (wcl=mcl !1 ref) not (wst=
MUSIC                  Marvin Duchow Music   MGUSB MGU01       MGU01 wcl=music not (wst=deleted
RBD                    Rare Book Division    MGUSB MGU01       MGU01 (wcl=rbd or wcl=cac) not (
REDP                   Redpath Audio Visual  MGUSB MGU01       MGU01 (wcl=redp !1 av) not (wst=
```

Util H_1_10 - Z0102 Setup – library_report

```
Number of records per base and Z0102 recommended and current setup
   Base                | # Docs   | Recommended | Current
   --------------------+----------+-------------+--------
   EINDEXES            |       96| Yes         | Yes
   EREFERENCE          |      405| Yes         | Yes
   EBOOKS              |     3179| Yes         | Yes
   REDP                |     8567| Yes         | Yes
   EJOURNALS           |    10678| Yes         | Yes
   ERESOURCES          |    15517| Yes         | Yes
   MCLREF              |    18903| Yes         | Yes
   ALLAV               |    23690| Yes         | Yes
   MCGILL_THESES       |    26910| Yes         | Yes
   GDOC                |    28723| Yes         | Yes
   HRMG                |    42732| Yes         | Yes
   ISLM                |    53471| Yes         | Yes
   MUSIC               |    57606| Yes         | Yes
   MCGILL_SERIALS      |    64332| Yes         | Yes
   BWOOD               |    70251| Yes         | Yes
   RBD                 |    71042| Yes         | Yes
   LAW                 |    75818| Yes         | Yes
   EDUC                |    84666| Yes         | Yes
   HLTH                |    94324| Yes         | Yes
   MCL                 |   944210| No          | No
   MUSE                |939471433| No          | No   ← if any sub-base's #DOCS looks
   MGU01               |939504561| No          | No   ← like this, fix tab_base.eng
```

Jerry says: "The utility sets "Recommended" to "Yes" for any base which includes fewer than 1/3 of the bibliographic records."
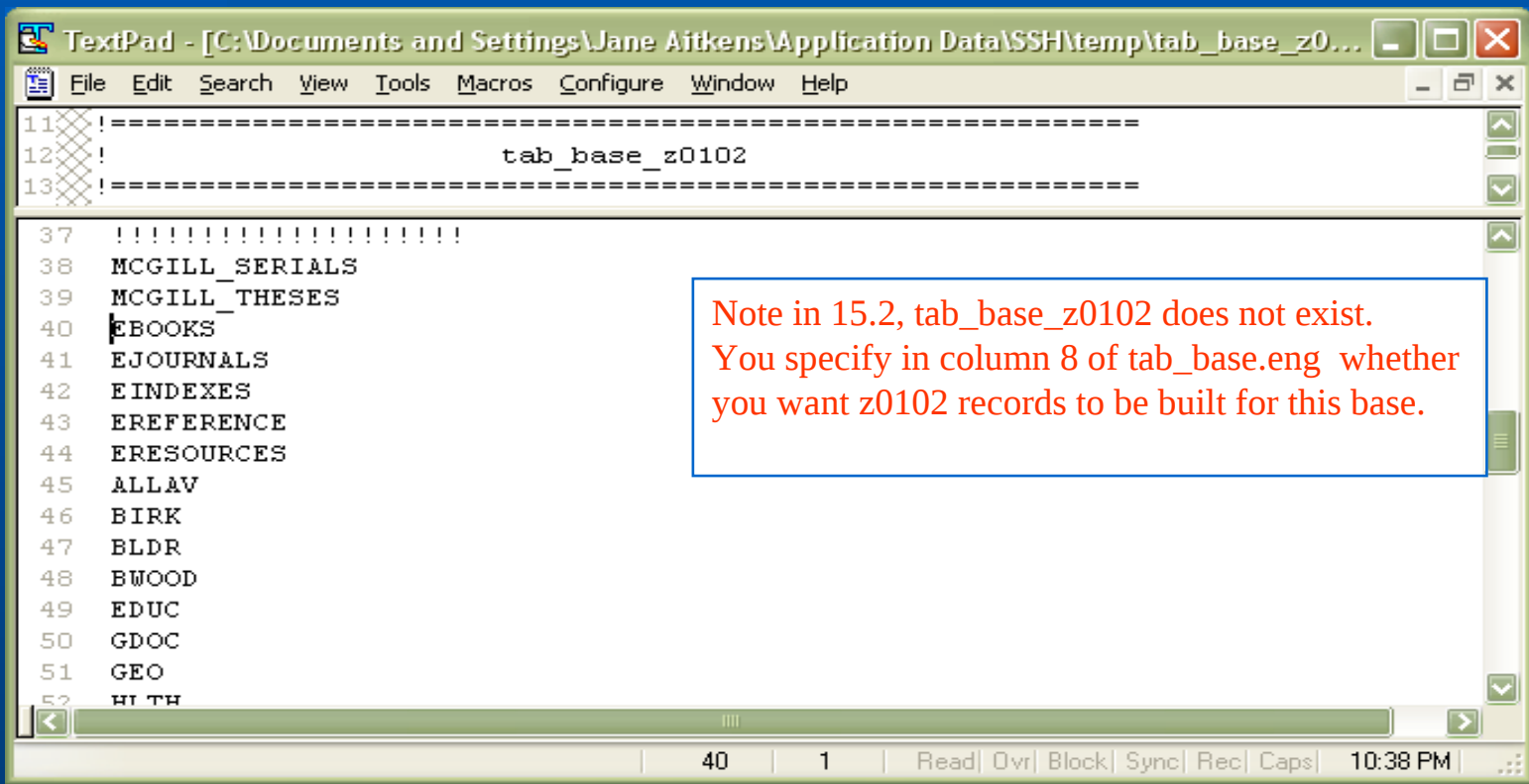
"To me, that's kind of high. I would look at bases which include fewer than 1/10 of the bibs."

"Compare the count to the total number of documents (as found in the util g/2 last-doc-number)"

McGill has just under 2 million bibs

tab_base_z0102

Note in 15.2, tab_base_z0102 does not exist.
You specify in column 8 of tab_base.eng whether you want z0102 records to be built for this base.

Define Z0102 in file_list – make it big

```
TAB z00           1900M           100M           ts_mgu_dat
IND z00_id        200M            50M            ts_mgu_ind
TAB z01           1500M           100M           ts_mgu_dat
IND z01_id        800M            100M           ts_mgu_ind
IND z01_id2       400M            100M           ts_mgu_ind
IND z01_id3       400M            100M           ts_mgu_ind
IND z01_id4       400M            100M           ts_mgu_ind
IND z01_id5       400M            100M           ts_mgu_ind

TAB z0102         1000M           10M            ts_mgu_dat
IND z0102_id      200M            10M            ts_mgu_ind
IND z0102_id1     700M            100M           ts_mgu_ind

TAB z02           500M            100M           ts_mgu_dat
IND z02_id        200M            100M           ts_mgu_ind
IND z02_id1       200M            100M           ts_mgu_ind
```

Running p_manage_32

- Must be run after p_manage_01 (keywords used in tab_base.eng)

    Our (bad) p_manage_32 (steps 1 & 2) took 18 hours (archiving off).

- Step 3 (switch update flag to "N") seemed too slow (2 hours to do 50,000), so we stopped it and switched the flags with SQL:
    update z01 set z01_update_z0102 = 'N' where Z01_ACC_SEQUENCE > '007800001' and Z01_ACC_SEQUENCE < '007850001';

- Bad Z0102 count: 30,719,289 lines
    (1/3 bigger than it needed to be because of our error)

- Our Current count: 22,314,164

Adjustments for p_manage_32

The z0102_id1 index ran out of TEMP extents, so Ex Libris:

- increased tablespace 'temp' by 1G

- altered the definitions of table z0102 and indexes z0102_id and z0102_id1 to use option 'nologging' and maxextents unlimited.

Our mistake: a bad tab_base.eng entry

- one of the sub-bases was set to find a non-existent TYP. We forgot to global change a set of existing ESLs after we decided to change their name in tab_type_config.

- aleph_ccl_expr str_in: wtp=eindex
  BASE : EINDEXES   NO. DOCUMENTS : 939524040.

- it added EVERY ACC heading from the whole database into the Z0102 for a base that should have registered zero hits

- Resulted in 11 million unneeded entries in Z0102 – one third of the table

Jerry's SQL to the rescue

- **How many:**

- select count(*) from z0102 where substr(Z0102_REC_KEY_1,1,20) like 'EINDEXES%';

- **Delete them:** (I did them in batches, to avoid rollback problems)

- delete from Z0102 where substr(Z0102_REC_KEY_1,1,20) like 'EINDEXES%' and Z0102_REC_KEY > '007800002' and Z0102_REC_KEY < '008800001';

- Did a million at a time, each took under an hour.
- **Example**:
    - started 7:55 - ended by 8:40 - 978,792 rows deleted.
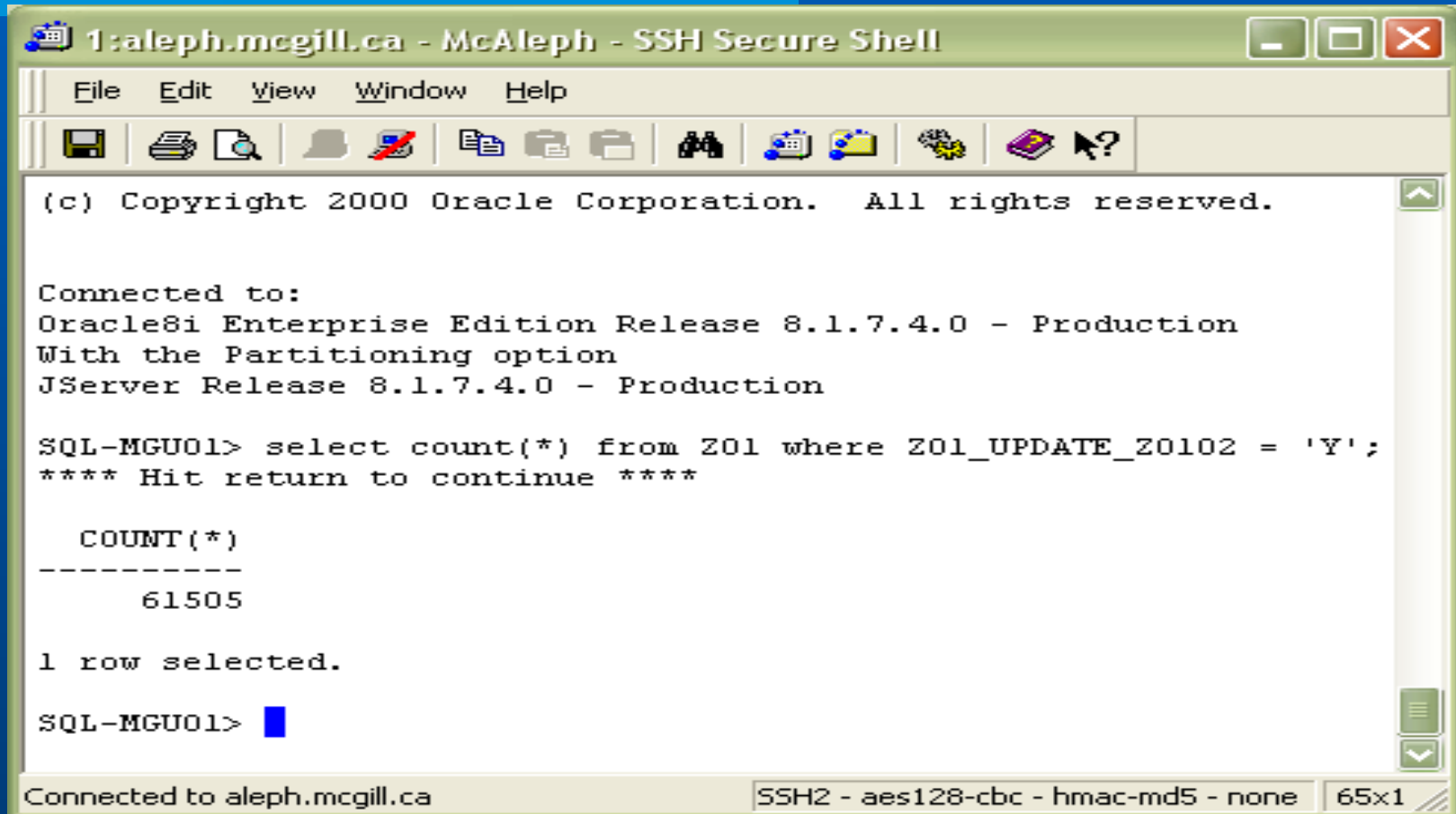
Step 3 – feeding the Z0102 table

- The logical base counter is **not** automatically updated. Rather, when a heading is updated, the Z01-UPDATE-Z0102 field of the heading's record (Z01) is set to 'Y'. Sample Z01 record

- **In version 14.2:** p_manage_34, the "Update Counters for Logical Bases" batch procedure updates the Z0102 table for those headings where Z01-UPDATE-Z0102 = 'Y'. If cataloguing is active, it should be run nightly (it locks the queue, but not the library). So add to the job_list:

  W3 21:00:00 Y   MGU01 p_manage_34   MGU01,0,00,

- **Note:** p_manage_17 also processes the Z01-UPDATE-Z0102s

- **In version 15:** (rep_ver 3310) updating the Z0102 counter is part of UE_08. See PRB 4425

How many headings to update?

p_manage_34

- **Problem: sometimes, there are so many heading changes, p_manage_34 cannot finish its updating before your shutdown.**

- **If p_manage_34 is still running, the batch queue cannot be stopped and so the shutdown aborts.**

- **Solution: "kill" p_manage_34 first in alephe_shutdown_505:**

Alephe_shutdown_505

Insert BEFORE the lib_batch shut

```
###########################################
#  aleph_ps script to shutdown p_manage_17
#  created July 2, 2000 Charles Kless
#  modified for p_manage_34 JMA (Minn) 2/7/2003
#  cloned from JMA (Janet Arth) by Jane A. 2003/3/7
###########################################

        # this will kill all instances of manage_34 - no matter
        # which library

        echo "killing manage_34"

        aleph_ps kill "manage_34"
```

On to Step 4

Step 4 – dressing up

- **Different colour backgrounds for the type of bases:**
  More Catalogues

  In tab_base.eng, each type uses one of 4 web extensions:
  ```
  Journal Titles          MGUTP
  E-Resources             MGUTP
  All AV Collections      MGUSB
  Blacker-Wood            MGUSB
  ICC Films               MGUAF
  Hospital Libraries      MGUAF
  CRL                     CRL
  ```

- **Colours assigned via style sheets: aleph-css-mgutp, etc**

Distinct Brief screens - in test

- Web pages are also sensitive to base suffixes (as of 14. 1), so we decided to make distinct screens for the individual sub-bases – start3, includes and filters, and short-table.

- Reference staff wanted holdings on brief screens for the bases of electronic material and journals, so we added more columns to www_tab_short.eng to display holdings as well as the <type> label. (Before and after)

- To pull holdings from 856$3 for custom brief table view, we used tab_expand_join_simple:

  HDURL 85640 AA 3   830## AA a

  Currenty in testing in frameless:
- short-a-body-eresources - **Current anthropology**
- short-a-body-mcgill_serials - Current anthropology
- short-a-body-ejournals - **Current anthropology**

The End

- **ESL** **project sample screen**

- **Documents:**
  - **patch 4, 5 rep_changes**
  - **Z0102.pdf**
  - **Brief screen display improvements**

- **Tables involved**

- **Sample Util H 1 – 10 – library report**

- **Sample log from p_manage_34**

PRB 2528

Problem:
    tab_type_config doesn't work for multiple occurrences

    I am having success working with tab_typ_config for display, indexing, and
    adding fields to a bib record. …

    I can use the match criteria (col 6) as long as there is only one occurrence of the
    field specified in col 4.

    When there is more than one occurrence of the field such as with 6xx fields, it
    appears that only the first occurrence of the field is evaluated.

Answer:
    Sent to Programming (ISRPRB 30451).

Current status:
     At the 'Programming' stage in Israel as of April 14, 2003. (penultimate step)

SUW index – for electronic resource retrieval

Z0102 description – from Z0102.pdf

- When a logical base is being browsed, the system uses the Z0102 table to "decide" whether to display the heading (Z01) without having to retrieve the documents attached to the heading, read them and then "decide".

- The Z0102 table is built for each heading and for each logical base. The table includes the filing text, the access sequence of the heading and a counter of the relevant documents attached to the heading. The pointer to the documents is still stored in the Z02 (ACCDOC) table.

Sample record from Z0201

```
01 z0102 \
  02 z0102_rec_key \
    03 acc_sequence ...........000000005
  02 z0102_rec_key_1 \
    03 base ...................LAW
    03 acc_code ...............SUL
    03 alpha ..................L
    03 filing_text ............ENGLAND CHURCH HISTORY 1066
  1485
    03 filing_sequence ........00005
  02 z0102_number_of_doc ......00002
  02 z0102_acc_sequence_see ...000000000
```

```
01 z01_acc \
    02 z01_rec_key \
        03 acc_code .............TTL
        03 alpha ................L
        03 filing_text ..........TESTAMENT DE NAPOLEON DOCUMENT ORIGINAL COM
        03 filing_sequence ......59428
    02 z01_acc_sequence .......013759428
    02 z01_hash ...............540207388535
    02 z01_aut_tag ............
    02 z01_rec_key_4 \
        03 aut_library ...........-CHK-
        03 aut_doc_number ........000000000
    02 z01_acc_sequence_see ....000000000
    02 z01_number_of_doc .......00000
    02 z01_cataloger \
        03 cataloger_name ........UE-01
        03 cataloger_level .......00
    02 z01_open_date ..........20030526
    02 z01_update_date ........20030526
    02 z01_cataloger_library ...
    02 z01_non_filing_char .....00
    02 z01_update_doc ..........Y
 ➔  02 z01_update_z0102 ........Y
    02 z01_display_text ........$$aTestament de Napoléon $$bdocument origin
```