

```

#!/usr/bin/perl

# This program takes a list of LCCN numbers (which are the directory and
# document names
# for the documents we have scanned) and grabs a MARC record from the
# Library
# of Congress Catalog.

# The documents it operates on are in the gdoc/READY directory, which
# means that
# those documents have already been cleared by quality control.

# After running, check NOCATALOG.txt. It will have a list of all the
# documents
# that did not get a catalog record.

# This program is to be run AFTER quality control is completed on a
# document.

# This program is to be run BEFORE mass-process.pl.

#$rdir = "../gdoc"; # temp DAB 8/5/2011
#$rdir = "../gdoc/READY"; # changed DAB 9/17/2010
$rdir = "../gdoc/GDOCHECK";

opendir DIR, "$rdir";
@lccns = grep /\d+/, readdir DIR;
closedir DIR;
open(ERR, ">NOCATALOG.txt");

MAIN: foreach $doc (@lccns){
print "\nDoing $doc . . . ";
$lccn = $doc;
$lccn =~ s/[a-z]//sg;

unless(-e "$rdir/$doc/$doc.cat"){
open(CAT, ">$rdir/$doc/$doc.cat");
}else{
print "$doc is already done. Moving on . . .\n";
}

open(WGET, "wget http://catalog.loc.gov/cgi-
bin/Pwebrecon.cgi?DB=local\\&PAGE=First -q -O -|");

while(<WGET>){
if(/<INPUT TYPE=HIDDEN NAME=PID VALUE="(.)"?>/i){
$pid = $1;
}
if(/<INPUT TYPE=HIDDEN NAME=SEQ VALUE="(.)"?>/i){
$seq = $1;
}
}
}
close WGET;

```

```
open(WGET, "wget http://catalog.loc.gov/cgi-  
bin/Pwebrecon.cgi?v4=1\\&ti=1,1\\&SEQ=$seq\\&Search_Arg=$lccn\\&Search_Co  
de=STNO\\&CNT=100\\&PID=$pid\\&SID=2 -q -O -|");
```

```
READ: while(<WGET>){  
  if(/<div class="nohits">/){  
    print ERR "$doc\n";  
    print "ERROR on $doc";  
    close IN;  
    next MAIN;  
  }  
  if(/<\\TR>/i){  
    $flag = 0;  
  }  
  if($flag == 1){  
    if(/<TD dir="ltr">/){  
      next READ;  
    }  
    s/<.*?>/sg;  
    $marc_cont = $_;  
    print CAT "$marc_no    $marc_cont";  
    $marc_cont = "";  
    $marc_no = "";  
  }  
  if(/<TH NOWRAP ALIGN=RIGHT VALIGN=TOP>(\d\d\d)<\\TH>/){  
    $marc_no = $1;  
    $flag = 1;  
  }  
} # end while  
print "Done.\n";  
#8/5/2001 DAB  
close WGET;  
} # end MAIN.
```

```
close CAT;  
print "Script has ended. \n";  
print "Please check NOCATALOG.txt for any errors. \n";  
print "If there are no errors you may now move the folders from CHECKED  
to READY. \n";
```