

Command Line (follow along!)

There's More than One Way to Groom a Cat(alog): Technologies for Data Analysis and Manipulation. OLAC Preconference October 26, 2017

Navigation

```
ls
cat README
cd work
ls
cd ..
cd work      (try tab completion!)
ls
cd treasure_hunt
ls
cat treasure      (... use ls, cd, and cat to find the treasure!)
cd ~/work/      (do this after the treasure hunt)
```

Looking at MARC

```
cd MARC
ls
cat LC1000bibsA.mrc
yaz-marcdump -i marc -o line LC1000bibsA.mrc
yaz-marcdump -i marc -o line LC1000bibsA.mrc > LC1000bibsA.mrk
      (try command-line editing!)
head LC1000bibsA.mrk
tail LC1000bibsA.mrk
less LC1000bibsA.mrk      (press q to exit)
```

Using Regular Expressions (grep)

```
grep '^050' LC1000bibsA.mrk
wc -l LC1000bibsA.mrk
grep 'your regex' LC1000bibsA.mrk | wc -l
```

```
grep '^050.*\$a P' LC1000bibsA.mrk
grep '^245.*science' LC1000bibsA.mrk
grep -i '^245.*science' LC1000bibsA.mrk
grep '^245.*science' LC1000bibsA.mrk | wc -l
grep -i '^245.*science' LC1000bibsA.mrk | wc -l
grep -i 'the flowery kingdom' *.mrc
grep -il 'the flowery kingdom' *.mrc
```

Scripting

```
cat look_for_missing.sh
look_for_missing.sh OpenTextbooks.mrc
python find_missing.py OpenTextbooks.mrc
cat find_missing.py
nano find_missing.py
python find_missing.py OpenTextbooks.mrc
```

Manual (Man) pages

```
man grep
cd ..
grep -r 'found' treasure_hunt
```

Diff

```
cd ..
cd title_lists
cat TitleList-2013-10-30.txt
cat TitleList-2014-02-05.txt
diff TitleList-2013-10-30.txt TitleList-2014-02-05.txt
```