

# Introdução à Ciência da Computação

Gerenciamento de Texto em Arquivos

---

Prof. Iago Augusto de Carvalho  
[iago.carvalho@unifal-mg.edu.br](mailto:iago.carvalho@unifal-mg.edu.br)

# Buscar em Arquivos de Texto Usando Expressão Regular

---

- Expressão Regular – Exemplos:
  - [A-Za-z] Qualquer letra;
  - [^0-9A-Za-z] Qualquer símbolo (exceto números e letras);
  - [A-Z][a-z]\* Uma letra maiúscula, seguida de zero ou mais letras minúsculas;
  - [A-Z][a-z]+ Uma letra maiúscula, seguida de uma ou mais letras minúsculas;
  - [A-Z][a-z]? Uma letra maiúscula, seguida de zero ou uma letra minúscula;

# Buscar em Arquivos de Texto Usando Expressão Regular

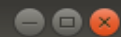
---

- Comando grep
  - Usado para buscar em arquivos ou na entrada padrão por linhas que contenham uma string que corresponda à expressão regular montada;
  - Sintaxe:
    - `grep [opções] expressão_regular [arquivos]`
- Algumas opções
  - `-v` Exibe todas as linhas que não batam com a regex;
  - `-i` Ignora a distinção entre maiúsculas e minúsculas;
  - `-c` exibe a contagem das linhas encontradas, mas não as linhas

# Buscar em Arquivos de Texto Usando Expressão Regular

---

- Exemplos:
  - `grep 'Linux' arquivo1`
  - `grep '^[0-9]'` (começa com número)
  - `grep '[0-9]$'` (termina com número)
  - `grep '^[^0-9]'` (começa com não número)



Arquivo Editar Ver Pesquisar Terminal Ajuda

adrianaavila@adrianaavila:~/Documentos\$ cat &gt; arquivo1

Meu nome é Adriana

sou professora do BCC na disciplina ICC

e estamos estudando Linux

hoje é dia

25Junho

ou para ser mais especifica sextaFeira/25

adrianaavila@adrianaavila:~/Documentos\$

adrianaavila@adrianaavila:~/Documentos\$ grep 'A\*' arquivo1

Meu nome é Adriana

sou professora do BCC na disciplina ICC

e estamos estudando Linux

hoje é dia

25Junho

ou para ser mais especifica sextaFeira/25

adrianaavila@adrianaavila:~/Documentos\$

adrianaavila@adrianaavila:~/Documentos\$ grep 'A\*' -i arquivo1

Meu nome é Adriana

sou professora do BCC na disciplina ICC

e estamos estudando Linux

hoje é dia

25Junho

ou para ser mais especifica sextaFeira/25

adrianaavila@adrianaavila:~/Documentos\$

adrianaavila@adrianaavila:~/Documentos\$ grep 'Linux' -c arquivo1

1

adrianaavila@adrianaavila:~/Documentos\$ grep '^[0-9]' arquivo1

25Junho

adrianaavila@adrianaavila:~/Documentos\$ grep '[0-9]\$" arquivo1

ou para ser mais especifica sextaFeira/25

adrianaavila@adrianaavila:~/Documentos\$

adrianaavila@adrianaavila:~/Documentos\$ grep '^[^0-9]' arquivo1

Meu nome é Adriana

sou professora do BCC na disciplina ICC

e estamos estudando Linux

hoje é dia

ou para ser mais especifica sextaFeira/25

adrianaavila@adrianaavila:~/Documentos\$



# Processar streams de texto usando filtros

---

O Linux dispõe de vários comandos cujo objetivo é servirem como filtros, modificando a saída de textos de forma mais útil;

A aula de hoje mostra com exemplos vários comandos dessa natureza;

# Processar streams de texto usando filtros

---

- Comando cut
  - Corta (exibindo colunas) de um ou mais arquivos;
  - O arquivo original não é modificado;
  - Exemplo:
    - `cut -d' ' -f2,3 lista_objetos`
  - Onde:
    - `-d` delimitador: configura o delimitador que separa uma coluna da outra. O padrão é o Tab;
    - `-f` número: imprime a coluna que corresponde ao número passado.

Arquivo Editar Ver Pesquisar Terminal Ajuda

adrianaavila@adrianaavila:~\$ cd Documentos/

adrianaavila@adrianaavila:~/Documentos\$ cat &gt; lista\_objetos

```
1 notebook 2500,00
2 celular 800,00
3 MemoriaRam 275,00
4 tablet 1500,00
5 monitor 450,00
6 headfone 185,00
```

adrianaavila@adrianaavila:~/Documentos\$ cut -d' ' -f2 lista\_objetos

```
notebook
celular
MemoriaRam
tablet
monitor
headfone
```

adrianaavila@adrianaavila:~/Documentos\$ cut -d' ' -f2,3 lista\_objetos

```
notebook 2500,00
celular 800,00
MemoriaRam 275,00
tablet 1500,00
monitor 450,00
headfone 185,00
```

adrianaavila@adrianaavila:~/Documentos\$ cut -d' ' -f2,3 lista\_objetos | sort

```
celular 800,00
headfone 185,00
MemoriaRam 275,00
monitor 450,00
notebook 2500,00
tablet 1500,00
```

adrianaavila@adrianaavila:~/Documentos\$ grep 'm \*' -i lista\_objetos | cut -d' ' -f2,3

```
MemoriaRam 275,00
monitor 450,00
```

adrianaavila@adrianaavila:~/Documentos\$



# Processar streams de texto usando filtros

---

- Comando expand
  - Converte tabs em espaços;
  - Por padrão, assume-se que um tab equivale a 8 espaços;
  - Sintaxe:
    - `expand [opções] [arquivos] > [arquivo_saida]`
- Algumas opções:
  - `-t número` especifica o número de espaços de tab, ao invés de 8;
  - `-i` converte apenas no início das linhas;

# Processar streams de texto usando filtros

---

- Comando unexpand
  - Converte espaços em tabs;
  - Por padrão, o unexpand só funciona em espaço no início das linhas;
  - Sintaxe:
    - `unexpand [opções] [arquivos] > [arquivo_saida]`
- Algumas opções:
  - `-a` Converte todos os espaços, não apenas os iniciais;

Arquivo Editar Ver Pesquisar Terminal Ajuda

```
adrianaavila@adrianaavila:~$ cd Documentos/  
adrianaavila@adrianaavila:~/Documentos$ ls  
arquivo1.txt  lista_objetos  
adrianaavila@adrianaavila:~/Documentos$ cat > arquivo2.txt  
    Meu nome é Adriana  
adrianaavila@adrianaavila:~/Documentos$ vi arquivo2.txt  
adrianaavila@adrianaavila:~/Documentos$ expand arquivo2.txt > arquivo3.txt  
adrianaavila@adrianaavila:~/Documentos$ vi arquivo3.txt  
adrianaavila@adrianaavila:~/Documentos$ unexpand arquivo3.txt > arquivo4.txt  
adrianaavila@adrianaavila:~/Documentos$ vi arquivo4.txt  
adrianaavila@adrianaavila:~/Documentos$
```

# Processar streams de texto usando filtros

---

- Comando head
  - Exibe as primeiras linhas de um ou mais arquivos (com padrão igual a 10 linhas);
  - Sintaxe
    - head [opções] [arquivos]
  - Algumas opções:
    - -c número: exibe na tela, o texto, conforme o número de bytes passado;
    - -n número: exibe na tela, a quantidade de linhas do texto, conforme o número passado.

Arquivo Editar Ver Pesquisar Terminal Ajuda

```
adrianaavila@adrianaavila:~$ cd Documentos/
```

```
adrianaavila@adrianaavila:~/Documentos$ ls
```

```
arquivo1.txt arquivo2.txt arquivo3.txt arquivo4.txt lista_objetos
```

```
adrianaavila@adrianaavila:~/Documentos$ cat arquivo1.txt
```

```
Meu nome é Adriana
```

```
sou professora do BCC na disciplina ICC
```

```
e estamos estudando Linux
```

```
hoje é dia
```

```
25Junho
```

```
ou para ser mais especifica sextaFeira/25
```

```
adrianaavila@adrianaavila:~/Documentos$
```

```
adrianaavila@adrianaavila:~/Documentos$ head -n 3 arquivo1.txt
```

```
Meu nome é Adriana
```

```
sou professora do BCC na disciplina ICC
```

```
e estamos estudando Linux
```

```
adrianaavila@adrianaavila:~/Documentos$ head -n -1 arquivo1.txt
```

```
Meu nome é Adriana
```

```
sou professora do BCC na disciplina ICC
```

```
e estamos estudando Linux
```

```
hoje é dia
```

```
25Junho
```

```
adrianaavila@adrianaavila:~/Documentos$ head -c 8 arquivo1.txt
```

```
Meu nomeadrianaavila@adrianaavila:~/Documentos$ head -c 50 arquivo1.txt
```

```
Meu nome é Adriana
```

```
sou professora do BCC na disciadrianaavila@adrianaavila:~/Documentos$ head arquivo1.txt arquivo2.txt
```

```
==> arquivo1.txt <==
```

```
Meu nome é Adriana
```

```
sou professora do BCC na disciplina ICC
```

```
e estamos estudando Linux
```

```
hoje é dia
```

```
25Junho
```

```
ou para ser mais especifica sextaFeira/25
```

```
==> arquivo2.txt <==
```

```
Meu nome é Adriana
```

```
adrianaavila@adrianaavila:~/Documentos$
```

# Processar streams de texto usando filtros

---

- Comando tail
  - Exibe as últimas linhas de um ou mais arquivos (com padrão igual a 10 linhas);
  - Sintaxe
    - tail [opções] [arquivos]
  - Algumas opções:
    - -c número: exibe na tela, o texto, conforme o número de bytes passado;
    - -n número: exibe na tela, a quantidade de linhas do texto, conforme o número passado.

Arquivo Editar Ver Pesquisar Terminal Ajuda

```
adrianaavila@adrianaavila:~$ cd Documentos
adrianaavila@adrianaavila:~/Documentos$ ls
arquivo1.txt arquivo2.txt arquivo3.txt arquivo4.txt lista_objetos
adrianaavila@adrianaavila:~/Documentos$ cat arquivo1.txt
Meu nome é Adriana
sou professora do BCC na disciplina ICC
e estamos estudando Linux
hoje é dia
25Junho
ou para ser mais especifica sextaFeira/25
adrianaavila@adrianaavila:~/Documentos$ tail -n 3 arquivo1.txt
hoje é dia
25Junho
ou para ser mais especifica sextaFeira/25
adrianaavila@adrianaavila:~/Documentos$ tail -n +3 arquivo1.txt
e estamos estudando Linux
hoje é dia
25Junho
ou para ser mais especifica sextaFeira/25
adrianaavila@adrianaavila:~/Documentos$
adrianaavila@adrianaavila:~/Documentos$ tail -c 8 arquivo1.txt
eira/25
adrianaavila@adrianaavila:~/Documentos$
adrianaavila@adrianaavila:~/Documentos$ tail -c 50 arquivo1.txt
25Junho
ou para ser mais especifica sextaFeira/25
adrianaavila@adrianaavila:~/Documentos$
adrianaavila@adrianaavila:~/Documentos$ tail arquivo2.txt arquivo3.txt
==> arquivo2.txt <==
    Meu nome é Adriana

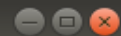
==> arquivo3.txt <==
    Meu nome é Adriana
adrianaavila@adrianaavila:~/Documentos$
```

# Processar streams de texto usando filtros

---

- Comando join
  - Executa um tipo de join semelhante à banco de dados, mas em arquivos;
  - Esse comando une as linhas de ambos os arquivos que tenham um índice em comum.
- Sintaxe
  - `join [arquivo1][arquivo2]`





Arquivo Editar Ver Pesquisar Terminal Ajuda

adrianaavila@adrianaavila:~\$ cd Documentos

adrianaavila@adrianaavila:~/Documentos\$ cat &gt; arquivo2.txt

1 João

2 Maria

3 Marta

adrianaavila@adrianaavila:~/Documentos\$ cat &gt; arquivo3.txt

1 3292-1566

2 3292-9876

3 3292-4567

adrianaavila@adrianaavila:~/Documentos\$ mv arquivo2.txt pessoas.txt

adrianaavila@adrianaavila:~/Documentos\$ mv arquivo3.txt telefones.txt

adrianaavila@adrianaavila:~/Documentos\$

adrianaavila@adrianaavila:~/Documentos\$ ls

arquivo1.txt arquivo4.txt lista\_objetos pessoas.txt telefones.txt

adrianaavila@adrianaavila:~/Documentos\$ join pessoas.txt telefones.txt &gt; lista\_pessoas

adrianaavila@adrianaavila:~/Documentos\$

adrianaavila@adrianaavila:~/Documentos\$ ls

arquivo1.txt arquivo4.txt lista\_objetos lista\_pessoas pessoas.txt telefones.txt

adrianaavila@adrianaavila:~/Documentos\$

adrianaavila@adrianaavila:~/Documentos\$ cat lista\_pessoas

1 João 3292-1566

2 Maria 3292-9876

3 Marta 3292-4567

adrianaavila@adrianaavila:~/Documentos\$



# Processar streams de texto usando filtros

---

- Comando paste
  - Cola as linhas correspondentes de um ou mais arquivos em colunas verticais
  - Sintaxe
    - `paste [opções] [arquivos]`
  - Algumas opções:
    - `-d'caractere'`: separa as colunas com o caractere dentro das aspas simples;
    - `-s`: concatena todo o conteúdo de um arquivo com uma linha para cada arquivo.

Arquivo Editar Ver Pesquisar Terminal Ajuda

```
adrianaavila@adrianaavila:~$ cd Documentos/
adrianaavila@adrianaavila:~/Documentos$ ls
arquivo1.txt arquivo4.txt lista_objetos lista_pessoas pessoas.txt telefones.txt
adrianaavila@adrianaavila:~/Documentos$ cat > dominio.txt
gmail.com
hotmail.com
yahoo.com.br
adrianaavila@adrianaavila:~/Documentos$
adrianaavila@adrianaavila:~/Documentos$ ls
arquivo1.txt arquivo4.txt dominio.txt lista_objetos lista_pessoas pessoas.txt telefones.txt
adrianaavila@adrianaavila:~/Documentos$
adrianaavila@adrianaavila:~/Documentos$ paste -d'@' pessoas.txt dominio.txt
1 João@gmail.com
2 Maria@hotmail.com
3 Marta@yahoo.com.br
adrianaavila@adrianaavila:~/Documentos$
adrianaavila@adrianaavila:~/Documentos$ ls
arquivo1.txt arquivo4.txt dominio.txt lista_objetos lista_pessoas pessoas.txt telefones.txt
adrianaavila@adrianaavila:~/Documentos$
adrianaavila@adrianaavila:~/Documentos$ paste -d'@' pessoas.txt dominio.txt > lista_pessoas_dominio
adrianaavila@adrianaavila:~/Documentos$
adrianaavila@adrianaavila:~/Documentos$ cat lista_pessoas_dominio
1 João@gmail.com
2 Maria@hotmail.com
3 Marta@yahoo.com.br
adrianaavila@adrianaavila:~/Documentos$
```

Arquivo Editar Ver Pesquisar Terminal Ajuda

adrianaavila@adrianaavila:~\$ cd Documentos/

adrianaavila@adrianaavila:~/Documentos\$ ls

```
arquivo1.txt  dominio.txt      lista_pessoas      pessoas.txt
arquivo4.txt  lista_objetos    lista_pessoas_dominio  telefones.txt
```

adrianaavila@adrianaavila:~/Documentos\$ paste -s pessoas.txt dominio.txt

```
1 João  2 Maria 3 Marta
gmail.com      hotmail.com      yahoo.com.br
```

adrianaavila@adrianaavila:~/Documentos\$ paste -s pessoas.txt dominio.txt &gt; teste

```
1 João  2 Maria 3 Marta
gmail.com      hotmail.com      yahoo.com.br
```

adrianaavila@adrianaavila:~/Documentos\$

# Processar streams de texto usando filtros

---

- Comando sort
    - Retorna para stdout a entrada classificada alfabeticamente;
    - Sintaxe
      - `sort [opções] [arquivos]`
    - Exemplo:
      - `sort lista_frutas`
      - `sort -r lista_frutas`
- Onde -r retorna em ordem inversa;



adrianaavila@adrianaavila: ~/Documentos



Arquivo Editar Ver Pesquisar Terminal Ajuda

adrianaavila@adrianaavila:~\$ cd Documentos/

adrianaavila@adrianaavila:~/Documentos\$ cat arquivo4.txt

Meu nome é Adriana

adrianaavila@adrianaavila:~/Documentos\$ cat &gt; arquivo4.txt

Uva

Melancia

Pera

Abacaxi

Banana

adrianaavila@adrianaavila:~/Documentos\$ mv arquivo4.txt lista\_frutas

adrianaavila@adrianaavila:~/Documentos\$

adrianaavila@adrianaavila:~/Documentos\$ sort lista\_frutas

Abacaxi

Banana

Melancia

Pera

Uva

adrianaavila@adrianaavila:~/Documentos\$ sort -r lista\_frutas

Uva

Pera

Melancia

Banana

Abacaxi

adrianaavila@adrianaavila:~/Documentos\$



# Processar streams de texto usando filtros

---

- Comando split
  - O comando split (dividir) é usado para dividir grandes arquivos em n-arquivos menores.
  - Os arquivos de saída são gerados de acordo com o número de linhas do arquivo de entrada.
  - O padrão é dividir o arquivo a cada 1000 linhas.
- Sintaxe
  - `split [opção] [arquivo de entrada] [arquivo de saída]`
- Exemplo:
  - `split -n 2 arquivo1.txt`
- Algumas opções:
  - `-n`: onde n é o número de arquivos que irão dividir o arquivo de entrada



Arquivo Editar Ver Pesquisar Terminal Ajuda

```
adrianaavila@adrianaavila:~/Documentos$ ls
arquivo1.txt dominio.txt lista_frutas lista_objetos lista_pessoas lista_pessoas_dominio pessoas.txt telefones.txt
adrianaavila@adrianaavila:~/Documentos$ mkdir Exemplos
adrianaavila@adrianaavila:~/Documentos$ mv *.txt lista_* /home/adrianaavila/Documentos/Exemplos
adrianaavila@adrianaavila:~/Documentos$ ls
Exemplos
adrianaavila@adrianaavila:~/Documentos$ cd Exemplos
adrianaavila@adrianaavila:~/Documentos/Exemplos$ cp arquivo1.txt /home/adrianaavila/Documentos/
adrianaavila@adrianaavila:~/Documentos/Exemplos$ cd ..
adrianaavila@adrianaavila:~/Documentos$ split -n 2 arquivo1.txt
adrianaavila@adrianaavila:~/Documentos$ ls
arquivo1.txt Exemplos xaa xab
adrianaavila@adrianaavila:~/Documentos$ cat xaa
Meu nome é Adriana
sou professora do BCC na disciplina ICC
e estamos estuadrianaavila@adrianaavila:~/Documentos$ cat xab
dando Linux
hoje é dia
25Junho
ou para ser mais especifica sextaFeira/25
adrianaavila@adrianaavila:~/Documentos$ cat arquivo1.txt
Meu nome é Adriana
sou professora do BCC na disciplina ICC
e estamos estudando Linux
hoje é dia
25Junho
ou para ser mais especifica sextaFeira/25
adrianaavila@adrianaavila:~/Documentos$
```



# Processar streams de texto usando filtros

---

- Comando tac
  - Semelhante ao cat, mas envia os arquivos para stdout com as linhas em ordem inversa;
  - Sintaxe
    - `tac [arquivo]`

Arquivo Editar Ver Pesquisar Terminal Ajuda

adrianaavila@adrianaavila:~\$ cd Documentos

adrianaavila@adrianaavila:~/Documentos\$

adrianaavila@adrianaavila:~/Documentos\$ tac arquivo1.txt

ou para ser mais especifica sextaFeira/25

25Junho

hoje é dia

e estamos estudando Linux

sou professora do BCC na disciplina ICC

Meu nome é Adriana

adrianaavila@adrianaavila:~/Documentos\$ tac arquivo1.txt &gt; teste.txt

adrianaavila@adrianaavila:~/Documentos\$

adrianaavila@adrianaavila:~/Documentos\$ cat teste.txt

ou para ser mais especifica sextaFeira/25

25Junho

hoje é dia

e estamos estudando Linux

sou professora do BCC na disciplina ICC

Meu nome é Adriana

adrianaavila@adrianaavila:~/Documentos\$

# Processar streams de texto usando filtros

---

- Comando `wc`
  - Exibe a contagem de linhas, palavras e caracteres para arquivos;
  - Sintaxe
    - `wc [opções] [arquivos]`
  - Algumas opções:
    - `-c` exibe apenas a contagem de caracteres;
    - `-l` exibe apenas a contagem de linhas;
    - `-w` exibe apenas a contagem de palavras;
  - Exemplo:
    - `wc arquivo1.txt`

Arquivo Editar Ver Pesquisar Terminal Ajuda

```
adrianaavila@adrianaavila:~$ cd Documentos
adrianaavila@adrianaavila:~/Documentos$
adrianaavila@adrianaavila:~/Documentos$ ls
arquivo1.txt  Exemplos  teste.txt  xaa  xab
adrianaavila@adrianaavila:~/Documentos$ cat arquivo1.txt
Meu nome é Adriana
sou professora do BCC na disciplina ICC
e estamos estudando Linux
hoje é dia
25Junho
ou para ser mais especifica sextaFeira/25
adrianaavila@adrianaavila:~/Documentos$
adrianaavila@adrianaavila:~/Documentos$ wc arquivo1.txt
  6  25 148 arquivo1.txt
adrianaavila@adrianaavila:~/Documentos$
adrianaavila@adrianaavila:~/Documentos$ wc -l arquivo1.txt
6 arquivo1.txt
adrianaavila@adrianaavila:~/Documentos$
adrianaavila@adrianaavila:~/Documentos$ wc -w arquivo1.txt
25 arquivo1.txt
adrianaavila@adrianaavila:~/Documentos$
adrianaavila@adrianaavila:~/Documentos$ wc -c arquivo1.txt
148 arquivo1.txt
adrianaavila@adrianaavila:~/Documentos$
```

# Referências

---

Slides adaptados do Prof. Flavio Barbieri Gonzaga e do curso Introdução ao SO Linux – Renata Spolon e Gabriel Martinez

PRITCHARD, S.; PESSANHA, B. G.; LANGFELDT, N.; STANGER, J.; DEAN, J. 2007. **Certificação Linux LPI Rápido e Prático. Guia de Referência nível 1: Exames 101 e 102.** 2ª Ed. Rio de Janeiro: Editora Alta Books.