

# Workflow Científico - Tema 5 - Trabalho 2

## Tópicos em Banco de Dados (INFO-7058)

Gabriel Silva Hermida (gash18@inf.ufpr.br)  
Jefferson Luis Stefani Alves Pereira (jlsap17@inf.ufpr.br)  
Josiney de Souza (josiney.souza@ifc.edu.br)  
Lucas Sidnei dos Santos (lss16@inf.ufpr.br)

UFPR / DInf

5 de junho de 2023

# Sumário

- 1 Introdução
  - Recapitulando...
- 2 Proposta de *workflow* em imagens médicas
- 3 Base de Dados Científicos UFPR
  - *Dataset 1*
  - *Workflow* de exemplo
- 4 Comentários/Impressões/Consolidações de conhecimento

# Sumário

- 1 Introdução
  - Recapitulando...
- 2 Proposta de *workflow* em imagens médicas
- 3 Base de Dados Científicos UFPR
  - *Dataset* 1
  - *Workflow* de exemplo
- 4 Comentários/Impressões/Consolidações de conhecimento



# Artigos pesquisados/lidos

No primeiro trabalho, falamos dos seguintes artigos:

Bharathi et al., (2008) caracterização de *workflow* e de tarefas; perfil dos principais *workflows*; *benchmarking* de *workflow* sintético;

Cuevas-Vicenttin et al., (2012) *workflow* científico e *provenance*; escalabilidade, tolerância a falhas, reprodutibilidade;

Silva et al. (2013) *provenance* em tempo de execução; integração entre *workflows* diferentes;

Mattoso et al. (2015) Human-in-the-Loop (HIL); taxonomia; análise em tempo de execução e adaptação dinâmica;

Miksa, Rauber, & Mina (2016) baixa reprodutibilidade.

# Principais pontos dos artigos

De forma geral, os arquivos falam de:

- *Workflows* científicos (visão geral e exemplos)
- *Provenance* (“off-line” e em tempo de execução)
- Reprodutibilidade
- Human-in-the-Loop
- Paralelização
- Execução distribuída

# Próximo trabalho

Possibilidades para o próximo trabalho (à época):

- Estudar a Base de Dados Científicos da UFPR - <https://bdc.c3s1.ufpr.br/>
- Estudar o RE3DATA - <https://www.re3data.org/>
- Verificar *provenance* de repositórios de dados
- Ler sobre Science gateway frameworks
- Tentar executar um *workflow* e reproduzir um experimento
- Atualizar o estado da arte com artigos novos
  - Toward a reference architecture based science gateway framework with embedded e-learning support - 2022 - Gabriele Pierantoni, Tamas Kiss, Alexander Bolotov, Dimitrios Kagialis, James DesLauriers, Amjad Ullah, Huankai Chen, David Chan You Fee, Hai-Van Dang, Jozsef Kovacs, Anna Belelaki, Themistocles Herekakis, Ioanna Tsagouri, Sandra Gesing

# Sumário

- 1 Introdução
  - Recapitulando...
- 2 Proposta de *workflow* em imagens médicas
- 3 Base de Dados Científicos UFPR
  - *Dataset* 1
  - *Workflow* de exemplo
- 4 Comentários/Impressões/Consolidações de conhecimento



# Workflow em radiologia

**Título:** A vendor-agnostic, PACS integrated, and DICOM-compatible software-server pipeline for testing segmentation algorithms within the clinical radiology workflow

**Autores:** Lei Zhang, Wayne LaBelle, Mathias Unberath, Haomin Chen, Jiazhen Hu, Guang Li, David Dreizin

**Origem:** EUA

**Publicado em:** Journal of Medical Systems (Springer / Pre-print)

**Ano:** 26 de abril de 2023

**Editora:** Research Square

**Link:** <https://www.researchsquare.com/article/rs-2837634/v1>



# Resumo

- Pipeline para segmentação de imagens
- Usa um *script shell*
- Testado com hematomas pélvicos traumáticos
- Um “servidor” para obter e anonimizar imagens e uma estação de trabalho com GPU
- Formato de software aberto
- Visualização dos resultados (OHIF)
- 89% do tempo do workflow é gasto no processo de segmentação
- Auxiliar na tomada de decisões para médicos e em testes de outras ferramentas para a tarefa de segmentação

Figura 1

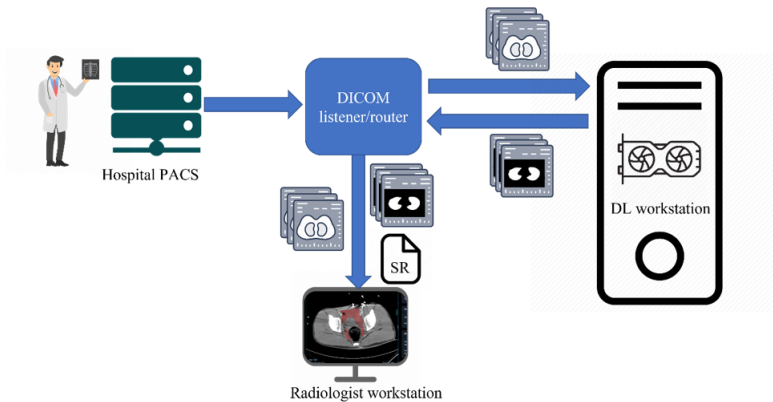
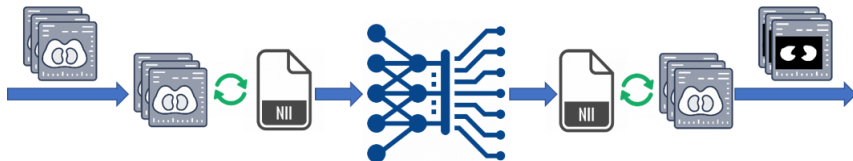
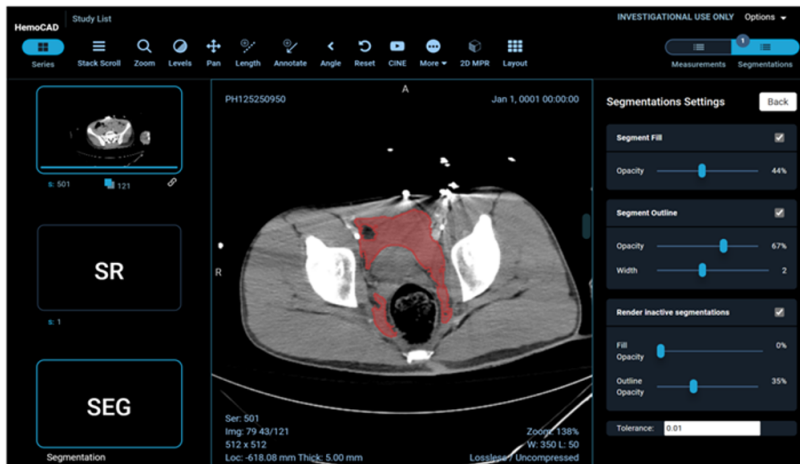


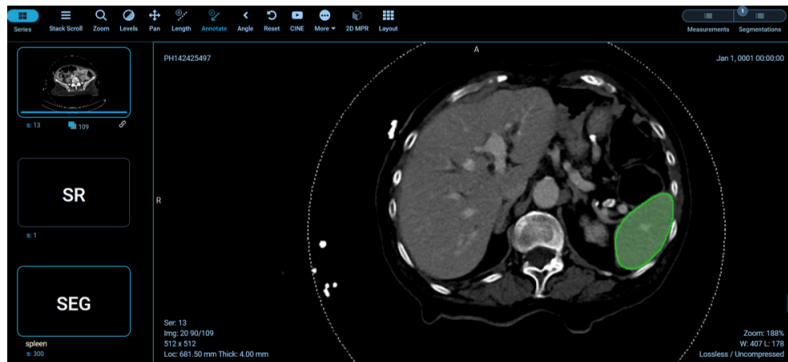
Figura 2



# Figura 3



# Figura 4



# Nossos comentários/impressões

Objetivo: estudar e executar o *workflow*

*Workflow científico*: Tentamos sem sucesso acessar o repositório de código: <https://github.com/vastc/>. Suspeitamos de várias etapas manuais;

## GitHub

This organization has no public repositories.

*Provenance*: Talvez tenha *provenance* “off-line”;

*Human-in-the-Loop*: Não sabemos se há o *Human-in-the-Loop*;

*Paralelização*: A paralelização ocorre a partir da GPU;

*Execução distribuída*: É distribuído (“servidor” e GPU)

# Nossos comentários/impressões

Objetivo: estudar e executar o *workflow*

*Workflow científico*: Graças à cooperação dos autores, tivemos acesso ao repositório;

## Resposta

I have made the repo public on GitHub. <https://github.com/vastc/vastc>, Wayne

*Provenance*: Talvez tenha *provenance* “off-line”;

*Human-in-the-Loop*: Não sabemos se há o *Human-in-the-Loop*;

*Paralelização*: A paralelização ocorre a partir da GPU;

*Execução distribuída*: É distribuído (“servidor” e GPU)

# Sumário

- 1 Introdução
  - Recapitulando...
- 2 Proposta de *workflow* em imagens médicas
- 3 Base de Dados Científicos UFPR
  - *Dataset* 1
  - *Workflow* de exemplo
- 4 Comentários/Impressões/Consolidações de conhecimento

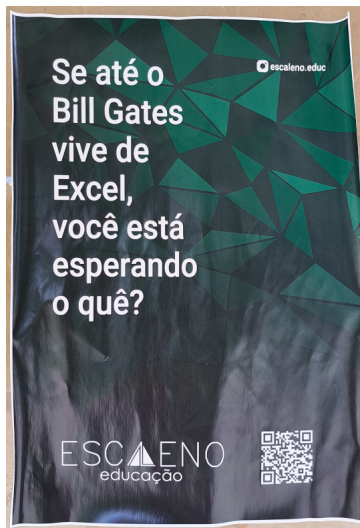




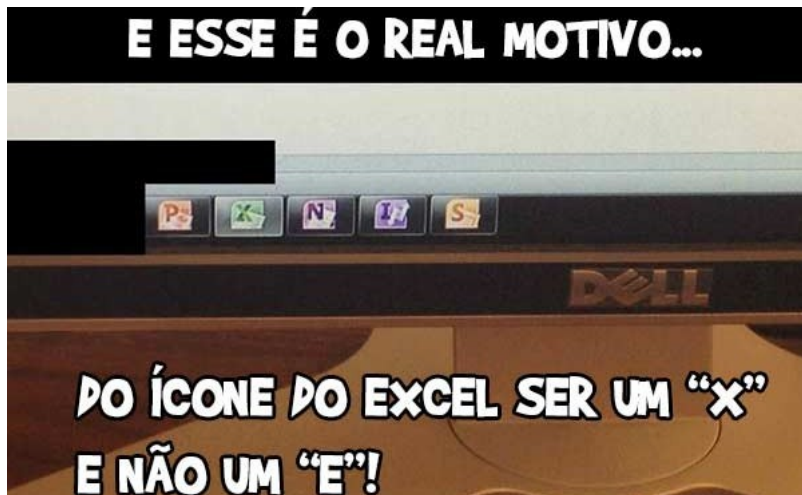
# Base de Dados Científicos da UFPR

- 79 itens
- 20/dez/2017 a dez/2022
- Vários nomes dos artigos/trabalhos como agrupadores de dados
- Alguns *datasets* foram analisados do ponto de vista de *workflow*
- Livro de códigos
- Banco de dados (em XLSX)

# Base de Dados Científicos da UFPR



## Base de Dados Científicos da UFPR



Fonte: <https://images3.memedroid.com/images/UPLOADED555/5af08cc282fe3.jpg>

# Dataset 1

- Entre o Amor Venéreo e o Útero: Desdobramentos e Controles de Desgovernadas na Condução da Visita Íntima na Penitenciária Feminina do Paraná
- Julia Heliodoro Souza Gitirana
- Doutorado, Programa de Pós-Graduação em Políticas Públicas, Setor de Ciências Sociais Aplicadas, UFPR
- <http://dx.doi.org/10.5380/bdc/63>
- Comitê de Ética da Universidade, nº CAAE 20141319.3.0000.0102
- Trabalho de mesmo nome: <https://hdl.handle.net/1884/70370>
- Questionários para:
  - agentes técnicas (psicólogas, assistentes sociais, enfermeiras)
  - agentes penitenciária
  - ocupantes do cargo de direção da unidade (diretoras, vice-diretoras e ex-diretoras)

# Dataset 1 - Divisão do formulário

- Bloco A diagnóstico do perfil da entrevista (**CARACTERIZACAO DEMOGRAFICA**);
- Bloco B definição de trajetórias profissionais no sistema penitenciário paranaense (**ATUACAO PROFISSIONAL**);
- Bloco C verificação do procedimento da visita íntima das mulheres na PFP, tendo por base estipulações da Portaria 499/2014 do DEPEN-PR (**PROCEDIMENTO VISITA INTIMA**);
- Bloco D valores agenciados e direcionados para as mulheres privadas de liberdade e os direitos sexuais e reprodutivos (**MULHERES E SISTEMA PENAL**).

# Dataset 1 - Banco de dados 1

Perfil da direção:

- 6 mulheres, 1 homem
- cristãs, heterossexuais e brancas
- maioria tem filho e casadas
- média de idade de 53 anos (menor 35, maior 65)
- Superior completo
- Direito e Psicologia

# Dataset 1 - Banco de dados 1

Perfil da direção:

	A	B	C	D	E	F	G	H	I	J	K
1	ID	Q01	Q02	Q03	Q04	Q05	Q06	Q06_1	Q07	Q08	Q09
2	D1PF00	1	1	6	1	56	1	2	1	5	7
3	D2PF00	1	1	2	1	59	1	1	1	4	1
4	D3PF00	1	0	4	1	65	1	2	1	0	7
5	D4PF00	2	1	2	1	49	1	2	2	4	gestã
6	D5PF00	1	1	1	0	35	1	1	1	4	gura
7	D6PF00	1	1	2	1	45	2	0	3	4	1
8	D7PF00	1	1	4	1	61	1	2	3	4	1

Respostas "Diretoras": <https://entre-o-amor-veneris-bj2e.glide.page>

# Dataset 1 - Banco de dados 2

Perfil das demais burocratas:

- 45 mulheres, 1 homem
- maioria cristã
- heterossexuais
- maioria branca, depois parda, preta, amarela
- maioria tem filho
- metade tem relacionamento estável
- média de idade de 43 anos (menor 28, maior 65)
- maioria tem Superior completo, quase todas têm Médio
- Direito, Gestão Pública e Pedagogia



# Dataset 1 - Banco de dados 2

Perfil das demais burocratas:

	A	B	C	D	E	F	G	H	I	J
1	ID	Q01	Q02	Q03	Q04	Q05	Q06	Q07	Q08	Q09
2	A46PF00	1	1	5	1	32	2	0	9	4
3	A45PF00	1	1	7	1	40	1	2	4	4
4	A44PF00	1	1	4	3	59	2	1	1	4
5	A43PF00	1	1	2	1	54	1	3	1	4
6	A42PF00	1	1	1	1	56	1	2	9	3
7	A410F00	1	1	7	3	39	1	1	6	4
8	A40PF00	1	1	5	1	38	1	1	4	4
9	A39PF00	1	1	1	1	36	1	1	1	4
10	A38PF00	1	1	6	2	65	1	3	1	4
11	A37PF00	1	1	2	1	41	1	1	2	4

Respostas "Burocratas": <https://entre-o-amor-veneris-t1nj.glide.page>

# Dataset 1 - Nossos comentários/impressões

Objetivo: estudar o *workflow* da construção do trabalho e confrontar com o que vemos

*Workflow científico*: definição das perguntas, aplicação da pesquisa, tratamento dos dados manualmente, análise/síntese;

*Provenance*: banco de dados em XLSX;

*Human-in-the-Loop*: em todo o momento, pouca automatização (tabelas? gráficos?)<sup>1</sup>;

*Paralelização*: não há (fluxo estilo pipeline);

*Execução distribuída*: não há (totalmente local)

<sup>1</sup>Segundo (Mattoso, 2015), não há HIL pois não há automatização

# Workflow de exemplo

**QUADRO 5** - Respostas das burocratas entrevistadas sobre a autorização da direção da PFP para a realização de visitas íntimas

Burocracia de rua da PFP	Autorização Tida Direção para visitas íntimas previstas na Portaria 499 de 2014 do DEPEN-PR								
	Visita íntima de mulheres privadas de liberdade			Visita íntima entre pessoas presas*			Visita íntima homoafetiva**		
	Sim	Não	Não sei dizer	Sim	Não	Não sei dizer	Sim	Não	Não sei dizer
<b>Diretoras</b>	100%	0	0	85,7%	14,3%	0	85,7%	14,3%	
<b>Agentes penitenciárias e agentes técnicas</b>	100%	0	0	72%	7%	13%	62,5%	8,7%	17,4%

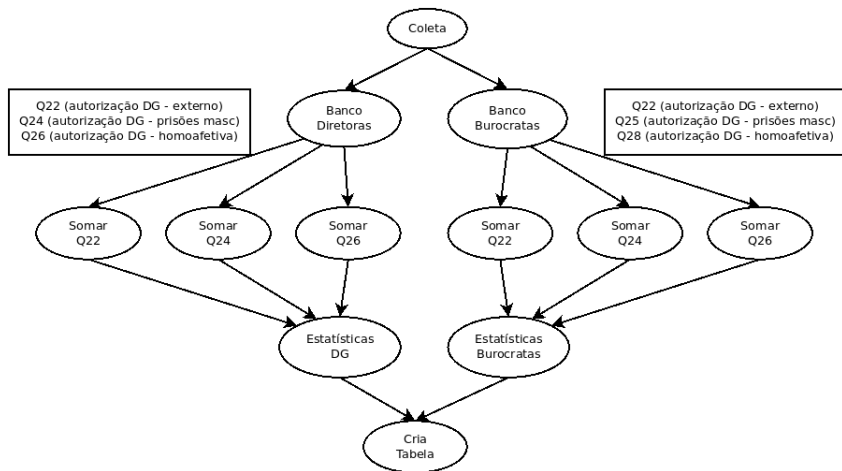
Agentes penitenciárias e agentes técnicas: n = 46; Diretoras: n = 7

Fonte: A autora, com base nos dados do questionário aplicado às burocratas (diretoras, agentes penitenciárias e agentes técnicas), referentes às questões Q.22; Q.25 e Q.28 respondidas pelas agentes; Q.22; Q.24; Q.26 respondidas por diretoras.

\*Nas respostas das agentes penitenciárias e técnicas sobre a possibilidade de visita íntima entre pessoas presas, 4% das entrevistas marcaram sim e não simultaneamente; 4% não responderam à questão.

\*\* Nas respostas das agentes penitenciárias e técnicas sobre a possibilidade de visita íntima homoafetiva, 8,7% não responderam.

# Workflow de exemplo



# Código 1: workflow.sh (1/3)

```
#!/bin/bash
```

```
source confs.txt
```

```
# Passo 1 - coleta: esta em bancos/*.xls
```

```
# Passo 2 - obtem bancos de diretoras e burocratas a partir de XLSX
```

```
./p2-obtem-banco-csv.sh bancos bancos/t1.xlsx
```

```
./p2-obtem-banco-csv.sh bancos bancos/t2.xlsx
```

## Código 1: workflow.sh (2/3)

```
# Passo 3 - realiza as somas dos indicadores
# Passo 4 - faz as estatísticas
echo "Base Diretoras"
rm -f $EST_DIR_EXT
./p3-somar-diretoras.sh bancos/t1.csv "#" 41 externo >> $EST_DIR_EXT
cat $EST_DIR_EXT
rm -f $EST_DIR_MASC
./p3-somar-diretoras.sh bancos/t1.csv "#" 43 masculino >> $EST_DIR_MASC
cat $EST_DIR_MASC
rm -f $EST_DIR_HOMO
./p3-somar-diretoras.sh bancos/t1.csv "#" 45 homoafetivo >> $EST_DIR_HOMO
cat $EST_DIR_HOMO
```

## Código 1: workflow.sh (3/3)

```
echo "Base Burocratas"
rm -f $EST_BUR_EXT
./p3-somar-burocratas.sh bancos/t2.csv "#" 37 externo >> $EST_BUR_EXT
cat $EST_BUR_EXT
rm -f $EST_BUR_MASC
./p3-somar-burocratas.sh bancos/t2.csv "#" 47 masculino >> $EST_BUR_MASC
cat $EST_BUR_MASC
rm -f $EST_BUR_HOMO
./p3-somar-burocratas.sh bancos/t2.csv "#" 57 homoafetivo >> $EST_BUR_HOMO
cat $EST_BUR_HOMO

# Passo 5 - cria tabela LaTeX
./p5-cria-tabela.sh

exit 0
```

## Passo 2: p2-obtem-banco-csv.sh (1/1)

```
#!/bin/bash
```

```
if [ ! $1 ]; then echo "ERRO: argumento 1 (\$1) vazio"; exit 1
elif [ ! -d $1 ]; then echo "ERRO: argumento 1 (\$1) inexistente";
elif [ ! $2 ]; then echo "ERRO: argumento 2 (\$2) vazio"; exit 2
elif [ ! -e $2 ]; then echo "ERRO: argumento 2 (\$2) inexistente";
```

```
# https://help.libreoffice.org/latest/en-GB/text/shared/guide/csv\_pa
```

```
# Argumento 1: 35 -> # na tabela ASCII (delimitador de campo)
```

```
# Argumento 2:      -> (delimitador de texto)
```

```
# Argumento 3: 0 -> usa codificacao (encoding) do sistema
```

```
# Argumento 4: 1 -> linha a iniciar
```

```
# Argumento 5:      -> formatacao da celula
```

```
...
```

```
# Argumento 12: -1 -> converter todas as planilhas do XLSX
```

```
libreoffice --headless --convert-to csv:"Text - txt - csv
```

```
(StarCalc):35,,0,1,,0,,,,,-1" --outdir $1 $2
```

```
exit 0
```



## Passo 3 e 4: p3-somar-diretoras.sh (1/4)

```
#!/bin/bash
```

```
if [ ! $1 ]; then echo "ERRO: argumento 1 (\$1 - BD) vazio"; exit 1
elif [ ! -e $1 ]; then echo "ERRO: argumento 1 (\$1 - BD) inexistent
elif [ ! $2 ]; then echo "ERRO: argumento 2 (\$2 - CHAR) vazio"; ex
elif [ ! $3 ]; then echo "ERRO: argumento 3 (\$3 - COLUNA) vazio";
elif [ ! $4 ]; then echo "ERRO: argumento 4 (\$4 - INDIC) vazio"; ex
```

```
NQUESTAO=$(cut -d$2 -f$3 $1 | head -n1)
```

```
QUESTAO=$(cut -d$2 -f$3 $1 | tail -n+2)
```

```
NAO=0
```

```
SIM=0
```

```
NSABE=0
```

```
NRESP=0
```

```
for VALOR in $QUESTAO
```

```
do
```

```
    if [[ $VALOR == 1 ]]
```

## Passo 3 e 4: p3-somar-diretoras.sh (2/4)

```
NAO=0
SIM=0
NSABE=0
NRESP=0
for VALOR in $QUESTAO
do
    if [[ $VALOR == 1 ]]
    then
        NAO=$(( $NAO + 1 ))
    elif [[ $VALOR == 2 ]]
    then
        SIM=$(( $SIM + 1 ))
    elif [[ $VALOR == 3 ]]
    then
        NSABE=$(( $NSABE + 1 ))
    elif [[ $VALOR == 0 ]]
    then
        NRESP=$(( $NRESP + 1 ))
```

## Passo 3 e 4: p3-somar-diretoras.sh (3/4)

```
elif echo $VALOR | grep -q ";"
then
    echo "Votaram em mais de uma opcao: $VALOR"
    if echo $VALOR | grep -q "1"
    then
        NAO=$(( $NAO + 1 ))
    fi
    if echo $VALOR | grep -q "2"
    then
        SIM=$(( $SIM + 1 ))
    fi
    if echo $VALOR | grep -q "3"
    then
        NSABE=$(( $NSABE + 1 ))
    fi
else
    echo "Valor \'$VALOR\' invalido"
fi
```

done



## Passo 3 e 4: p3-somar-diretoras.sh (4/4)

```
        else
            echo "Valor \'$VALOR\' invalido"
        fi
    done

    echo "=== Resultados para $NQUESTAO ($4) ==="
    echo "Quantidade de NAO: $NAO"
    echo "Quantidade de SIM: $SIM"
    echo "Quantidade de NSABE: $NSABE"
    echo "Quantidade de NRESP: $NRESP"
    echo

    exit 0
```

## Passo 5: p5-cria-tabela.sh (1/3)

```
#!/bin/bash
```

```
source confs.txt
```

```
cp $BASE_TABELA $TABELA
```

```
# Base 1: Diretoras
```

```
SIM=$(grep "SIM" $EST_DIR_EXT | cut -d" " -f4)
```

```
NAO=$(grep "NAO" $EST_DIR_EXT | cut -d" " -f4)
```

```
NSABE=$(grep "NSABE" $EST_DIR_EXT | cut -d" " -f4)
```

```
sed -i -e "s/DIR-EXT-SIM/$SIM/g" \  
      -e "s/DIR-EXT-NAO/$NAO/g" \  
      -e "s/DIR-EXT-NSABE/$NSABE/g" $TABELA
```

```
SIM=$(grep "SIM" $EST_DIR_MASC | cut -d" " -f4)
```

```
NAO=$(grep "NAO" $EST_DIR_MASC | cut -d" " -f4)
```

```
NSABE=$(grep "NSABE" $EST_DIR_MASC | cut -d" " -f4)
```

```
sed -i -e "s/DIR-MASC-SIM/$SIM/g" \  
      -e "s/DIR-MASC-NAO/$NAO/g" \  
      -e "s/DIR-MASC-NSABE/$NSABE/g" $TABELA
```

## Passo 5: p5-cria-tabela.sh (2/3)

```
SIM=$(grep "SIM" $EST_DIR_HOMO | cut -d" " -f4)
NAO=$(grep "NAO" $EST_DIR_HOMO | cut -d" " -f4)
NSABE=$(grep "NSABE" $EST_DIR_HOMO | cut -d" " -f4)
sed -i -e "s/DIR-HOMO-SIM/$SIM/g" \
        -e "s/DIR-HOMO-NAO/$NAO/g" \
        -e "s/DIR-HOMO-NSABE/$NSABE/g" $TABELA
```

# Base 2: Burocratas

```
SIM=$(grep "SIM" $EST_BUR_EXT | cut -d" " -f4)
NAO=$(grep "NAO" $EST_BUR_EXT | cut -d" " -f4)
NSABE=$(grep "NSABE" $EST_BUR_EXT | cut -d" " -f4)
sed -i -e "s/BUR-EXT-SIM/$SIM/g" \
        -e "s/BUR-EXT-NAO/$NAO/g" \
        -e "s/BUR-EXT-NSABE/$NSABE/g" $TABELA
```

## Passo 5: p5-cria-tabela.sh (3/3)

```
SIM=$(grep "SIM" $EST_BUR_MASC | cut -d" " -f4)
NAO=$(grep "NAO" $EST_BUR_MASC | cut -d" " -f4)
NSABE=$(grep "NSABE" $EST_BUR_MASC | cut -d" " -f4)
sed -i -e "s/BUR-MASC-SIM/$SIM/g" \
      -e "s/BUR-MASC-NAO/$NAO/g" \
      -e "s/BUR-MASC-NSABE/$NSABE/g" $TABELA
```

```
SIM=$(grep "SIM" $EST_BUR_HOMO | cut -d" " -f4)
NAO=$(grep "NAO" $EST_BUR_HOMO | cut -d" " -f4)
NSABE=$(grep "NSABE" $EST_BUR_HOMO | cut -d" " -f4)
sed -i -e "s/BUR-HOMO-SIM/$SIM/g" \
      -e "s/BUR-HOMO-NAO/$NAO/g" \
      -e "s/BUR-HOMO-NSABE/$NSABE/g" $TABELA
```

```
exit 0
```

## Auxiliares: confs.txt

```
EST_DIR_EXT="estatisticas-diretoras-externos.txt"  
EST_DIR_MASC="estatisticas-diretoras-masculinos.txt"  
EST_DIR_HOMO="estatisticas-diretoras-homoafetivos.txt"  
EST_BUR_EXT="estatisticas-burocratas-externos.txt"  
EST_BUR_MASC="estatisticas-burocratas-masculinos.txt"  
EST_BUR_HOMO="estatisticas-burocratas-homoafetivos.txt"  
BASE_TABELA="base-tabela.tex"  
TABELA="tabela.tex"
```



# Auxiliares: tabela.tex

File Edit Table Column Row Cell Help

Default table style

1	Burocracia de rua da PFP	Autorização Tida Direção para visitas íntimas previstas na Portaria 499 de 2014 do DEPEN-PR						
2		Visita íntima de mulheres privadas de liberdade			Visita íntima entre pessoas presas*			Visita íntima homoafetiva
3		Sim	Não	Não sei dizer	Sim	Não	Não sei dizer	Sim
4	Diretoras	DIR-EXT-SIM	DIR-EXT-NAO	DIR-EXT-NSABE	DIR-MASC-SIM	DIR-MASC-NAO	DIR-MASC-NSABE	DIR-HOMO-SIM
5	Agentes penitenciárias e agentes técnicas	BUR-EXT-SIM	BUR-EXT-NAO	BUR-EXT-NSABE	BUR-MASC-SIM	BUR-MASC-NAO	BUR-MASC-NSABE	BUR-HOMO-SIM
6								

Generate

Result (click "Generate" to refresh)

Copy to clipboard

```

1 % Please add the following required packages to your document preamble:
2 % \usepackage{multirow}
3 \begin{table}[]
4 \begin{tabular}{|c|ccccccc|}
5 \hline
6 \multirow{3}{*}{\begin{tabular}{c}Burocracia\ de rua\ da PFP\end{tabular}} & \multicolumn{9}{c}{\begin{tabular}{c}Autorização Tida Direção para\ visitas\ íntimas\ previstas\ na\ Portaria\ 499\ de\ 2014\ do\ DEPEN-PR\end{tabular}}
7 \multicolumn{3}{c}{\begin{tabular}{c}Visita íntima de mulheres privadas de liberdade\end{tabular}}
8 \multicolumn{3}{c}{\begin{tabular}{c}Visita íntima entre pessoas presas*\end{tabular}}
9 \multicolumn{3}{c}{\begin{tabular}{c}Visita íntima homoafetiva\end{tabular}}
10 \begin{tabular}{c}Diretoras\end{tabular} & \begin{tabular}{c}DIR-EXT-SIM\end{tabular} & \begin{tabular}{c}DIR-EXT-NAO\end{tabular} & \begin{tabular}{c}DIR-EXT-NSABE\end{tabular} & \begin{tabular}{c}DIR-MASC-SIM\end{tabular} & \begin{tabular}{c}DIR-MASC-NAO\end{tabular} & \begin{tabular}{c}DIR-MASC-NSABE\end{tabular} & \begin{tabular}{c}DIR-HOMO-SIM\end{tabular}
11 \begin{tabular}{c}Agentes penitenciárias e agentes técnicas\end{tabular} & \begin{tabular}{c}BUR-EXT-SIM\end{tabular} & \begin{tabular}{c}BUR-EXT-NAO\end{tabular} & \begin{tabular}{c}BUR-EXT-NSABE\end{tabular} & \begin{tabular}{c}BUR-MASC-SIM\end{tabular} & \begin{tabular}{c}BUR-MASC-NAO\end{tabular} & \begin{tabular}{c}BUR-MASC-NSABE\end{tabular} & \begin{tabular}{c}BUR-HOMO-SIM\end{tabular}
12 \end{tabular}
13 \end{table}

```

Template da tabela feito com <https://www.tablesgenerator.com/>

# Tabela

Burocracia de rua da PFP	Autorização Tida Direção para visitas íntimas previstas na Portaria 499 de 2014 do DEPEN-PR								
	Visita íntima de mulheres privadas de liberdade			Visita íntima entre pessoas presas*			Visita íntima homoafetiva**		
	Sim	Não	Não sei dizer	Sim	Não	Não sei dizer	Sim	Não	Não sei dizer
Diretoras	7	0	0	6	1	0	5	2	0
Agentes penitenciárias e agentes técnicas	46	0	0	35	5	6	30	4	7

# Workflow de exemplo

**QUADRO 5** - Respostas das burocratas entrevistadas sobre a autorização da direção da PFP para a realização de visitas íntimas

Burocracia de rua da PFP	Autorização Tida Direção para visitas íntimas previstas na Portaria 499 de 2014 do DEPEN-PR								
	Visita íntima de mulheres privadas de liberdade			Visita íntima entre pessoas presas*			Visita íntima homoafetiva**		
	Sim	Não	Não sei dizer	Sim	Não	Não sei dizer	Sim	Não	Não sei dizer
<b>Diretoras</b>	100%	0	0	85,7%	14,3%	0	85,7%	14,3%	
<b>Agentes penitenciárias e agentes técnicas</b>	100%	0	0	72%	7%	13%	62,5%	8,7%	17,4%

Agentes penitenciárias e agentes técnicas: n = 46; Diretoras: n = 7

Fonte: A autora, com base nos dados do questionário aplicado às burocratas (diretoras, agentes penitenciárias e agentes técnicas), referentes às questões Q.22; Q.25 e Q28 respondidas pelas agentes; Q.22; Q.24; Q.26 respondidas por diretoras.

\*Nas respostas das agentes penitenciárias e técnicas sobre a possibilidade de visita íntima entre pessoas presas, 4% das entrevistas marcaram sim e não simultaneamente; 4% não responderam à questão.

\*\* Nas respostas das agentes penitenciárias e técnicas sobre a possibilidade de visita íntima homoafetiva, 8,7% não responderam.

## Dataset 1.5 - Nossos comentários/impressões

Objetivo: melhorar o *workflow* de construção do trabalho em algum aspecto

*Workflow científico*: automatização de criação ou atualização de uma tabela em  $\text{\LaTeX}$ ;

*Provenance*: bases, registro dos *logs* e dos arquivos gerados;

*Human-in-the-Loop*: alterando os parâmetros manualmente no *script workflow.sh* entre as execuções;

*Paralelização*: possível paralelizar colocando as etapas em segundo plano para bancos ou passos grandes ou demorados;

*Execução distribuída*: não há (totalmente local)

## Dataset 1.5 - Nossos comentários/impressões

Curiosidades e inconsistências descobertas:

- Quando não houve resposta, foi atribuído valor 0 no banco de dados;
- Na questão 25 das Burocratas, 2 células possuem 2 valores;

Resposta	Resposta	Resposta	Resposta	Resposta
1	na lista	2	0	0
2	a prisão	2;3	1	4
3	0	2	0	0
3	usitos e	2	0	0
2	aprese	2;3	4	4
3	em uso	2	0	0

- Na questão 28 das Burocratas, 1 célula tem valor vazio/nulo ( )

Resposta	Resposta	Resposta
diretores da	2	0
0	3	3
como é.		1
de masculin	1	0
Não sei dizer	1	0

## Dataset 1.5 - Nossos comentários/impressões

Curiosidades e inconsistências descobertas:

- Respostas “Não”, “Sim”, “Não sabe” → “**Não sabe**”, “Sim”, “**Não**”
- Tem uma posição da tabela sem valor atribuído
- Números de “Não sabe”, “Sim”, “Não” para a questão homoafetiva no banco de diretoras (Q26) tem um erro de contabilização
- É possível alterar o *workflow* para contabilizar números, percentual ou ambos

# Sumário

- 1 Introdução
  - Recapitulando...
- 2 Proposta de *workflow* em imagens médicas
- 3 Base de Dados Científicos UFPR
  - *Dataset* 1
  - *Workflow* de exemplo
- 4 Comentários/Impressões/Consolidações de conhecimento



# Conclusão

- Recapitulamos os trabalhos pesquisados
- Tentamos analisar um *workflow* de imagens médicas
- Analisamos um trabalho da BDC/UFPR

**GitHub** <https://github.com/josiney-souza/ufpr-topicos-bd>

- Fizemos preparação dos dados (“visão”)

**Glide Diretoras:** <https://entre-o-amor-veneris-bj2e.glide.page/>

**Glide Burocratas:** <https://entre-o-amor-veneris-t1nj.glide.page/>

**Honeycode Diretoras:** <https://app.honeycode.aws/arn%3Aaws%3Asheets%3Aus-west-2%3A122162422134%3Amobileapp%3A6b541e58-d6d7-480a-a57c-7fbe235d73db%2Fe5fc63a2-991f-4ea4-8c0a-a4f65bb2535c>

**Honeycode Burocratas:** <https://app.honeycode.aws/arn%3Aaws%3Asheets%3Aus-west-2%3A122162422134%3Amobileapp%3A6b541e58-d6d7-480a-a57c-7fbe235d73db%2F4268fcd1-d212-4e7e-a12b-81cb30b58abd>



# Impressões

- O artigo em *pre-print* analisado é difícil de ler

# Impressões

- O artigo em *pre-print* analisado é difícil de ler
- A BDC/UFPR é pouco conhecida e pouco utilizada
- Muitos usam o mesmo nome do artigo para o *dataset*
- Nem sempre os dados disponibilizados estão claramente organizados
- Há alguns bancos de dados (em XLSX) e livros de códigos

# Impressões

- O artigo em *pre-print* analisado é difícil de ler
- A BDC/UFPR é pouco conhecida e pouco utilizada
- Muitos usam o mesmo nome do artigo para o *dataset*
- Nem sempre os dados disponibilizados estão claramente organizados
- Há alguns bancos de dados (em XLSX) e livros de códigos
- É possível montar *workflow* para automatizar tarefas mesmo de outras áreas
- Há alternativas para a preparação de dados de *provenance*

# Impressões

**Título:** The Science in Computer Science

**Autores:** Denning, Peter J.

**Origem:** EUA

**Publicado em:** Commun. ACM

**Ano:** 2013

**Editora:** ACM (Association for Computing Machinery)

**Link:**

- <https://doi.org/10.1145/2447976.2447988>
- <http://denninginstitute.com/pjd/PUBS/CACMcols/cacmMay13.pdf>

# Impressões

Ciência da Computação é uma ciência? Pois uma ciência tradicional é/tem ...

- Organizada para entender, explorar e lidar com um fenômeno generalizado;
- Abrange processos naturais e artificiais do fenômeno;
- Corpo estruturado codificado de conhecimento;
- Compromisso com métodos experimentais para descoberta e validação;
- Reprodutibilidade dos resultados;
- Falseabilidade de hipóteses e modelos;
- Capacidade de fazer previsões confiáveis, algumas das quais surpreendentes.

# Impressões

Ciência da Computação é uma ciência? Pois algumas perspectivas/pontos de vista ...

**Ciências:** foca em aumentar o entendimento através de métodos experimentais;

**Engenharias:** foca em projetar e construir de sistemas computacionais cada vez melhores;

**Matemática:** foca no que pode ser deduzido a partir de sentenças aceitas.

# Impressões

Ciência da Computação é uma ciência? Sim ...

- 1 Teoria da localidade em memória; e teoria da avaliação de desempenho (predições);

# Impressões

Ciência da Computação é uma ciência? Sim ...

- 1 Teoria da localidade em memória; e teoria da avaliação de desempenho (predições);
- 2 Dada a complexidade, é necessário abordagens experimentais e empíricas;



# Impressões

Ciência da Computação é uma ciência? Sim ...

- 1 Teoria da localidade em memória; e teoria da avaliação de desempenho (predições);
- 2 Dada a complexidade, é necessário abordagens experimentais e empíricas;
- 3 Vários campos de pesquisa trabalham com processos de informações naturais;

# Impressões

Ciência da Computação é uma ciência? Sim ...

- 1 Teoria da localidade em memória; e teoria da avaliação de desempenho (predições);
- 2 Dada a complexidade, é necessário abordagens experimentais e empíricas;
- 3 Vários campos de pesquisa trabalham com processos de informações naturais;
- 4 Como computação está em tudo; pensamento computacional é extremamente importante;

# Impressões

Ciência da Computação é uma ciência? Sim ...

- 1 Teoria da localidade em memória; e teoria da avaliação de desempenho (predições);
- 2 Dada a complexidade, é necessário abordagens experimentais e empíricas;
- 3 Vários campos de pesquisa trabalham com processos de informações naturais;
- 4 Como computação está em tudo; pensamento computacional é extremamente importante;
- 5 Computação está presente na educação formal desde o Ensino Fundamental.

## Outros assuntos

- Dúvidas?
- Encaminhamentos?
- Outros assuntos?

“Até que os leões inventem as suas próprias histórias, os caçadores serão sempre os heróis das narrativas de caça.”

---

Provérbio africano

<https://www.pensador.com/frase/MTIOMzA2Ng/>

# *Workflow Científico - Tema 5 - Trabalho 2*

## *Tópicos em Banco de Dados (INFO-7058)*

Gabriel Silva Hermida (gash18@inf.ufpr.br)  
Jefferson Luis Stefani Alves Pereira (jlsap17@inf.ufpr.br)  
Josiney de Souza (josiney.souza@ifc.edu.br)  
Lucas Sidnei dos Santos (lss16@inf.ufpr.br)

UFPR / DInf

5 de junho de 2023