



UENF - CCT - LCMAT - Ciência da Computação

ACADEM: UM SOFTWARE DE ANÁLISE DE DADOS ACADÊMICOS BASEADO NO DATA SCIENCE WORK MODEL

Daniel Brito dos Santos (danielbrito@pq.uenf.br); Annabell Del Real Tamariz (annabell@uenf.br).

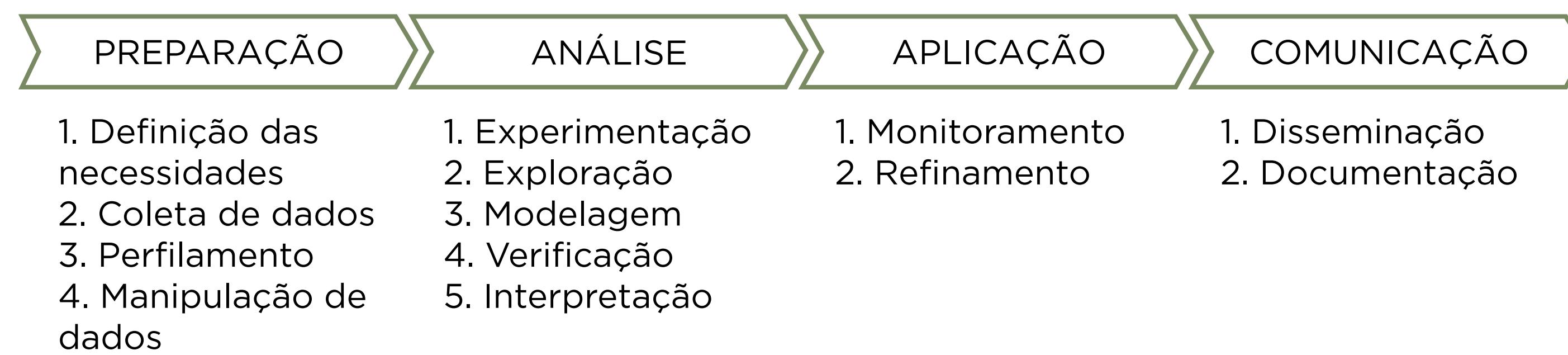
INTRODUÇÃO

A ciência de dados é a ciência de transformar dados em valor. No entanto, devido às suas múltiplas facetas e complexidade, modelos de trabalho são fundamentais para orientar a pesquisa e prática dessa disciplina. Identificamos um problema de dados na ausência informações precisas sobre a demanda dos alunos por cada disciplina de sua grade curricular na UENF, o que dificulta o planejamento das turmas, dos professores e das salas.

MÉTODOS

Utilizamos o Data Science Work Model (DSWM) para planejar e executar um projeto prático de dados. Desenvolvemos um software chamado AcaDem, que analisa os extratos acadêmicos dos estudantes da UENF e gera informações sobre a demanda dos alunos por cada disciplina de sua grade curricular. Usamos Python para programar o software e SQLite para criar o banco de dados.

Data Science Work Model (DSWM)



RESULTADOS E CONCLUSÕES

O AcaDem apresentou desempenho satisfatório em todos os seus casos de uso. Concluímos que o DSWM é um modelo eficiente para abordar um projeto de dados e que o AcaDem é um projeto promissor para contribuir com a tomada de decisão baseada em evidência em instituições de ensino superior.

```
Sigla Nome Demanda (#alunos)
FIS01109 Laboratório de Física Geral II 4
INFO1201 Análise e Projeto de Sistemas 4
INFO1115 Redes de Computadores 4
INFO1116 Banco de Dados I 4
INFO1117 Linguagens Formais e Teoria da Computação 4
LES04514 Metodologia do Trabalho Científico 4
INFO1205 Inteligência Artificial 4
MAT01201 Estatística e Probabilidades 3
PRO01540 Empreendedorismo 3
FIS01103 Física Geral II 2
LES04536 Computação e Sociedade 2
MAT01105 Cálculo Diferencial e Integral III 1
MAT01106 Método Matemático 1
LEL04102 Inglês Instrumental I 1
```

Output do AcaDem

```
sqlite> .tables
Disciplinas      EnsinoMedio      Ingresso
DisciplinasCursadas  Estudante      NotasEnem
sqlite> SELECT * FROM DisciplinasCursadas;
119110351|INF01101|8,6|APR
119110351|INF01105|9,5|APR
119110351|INF01106|5,0|APR
119110351|MAT01101|6,4|APR
119110351|MAT01104|7,2|APR
119110351|MAT01117|5,3|APR
119110351|FIS01202|8,3|CVD
119110351|FIS01204|7,6|APR
119110351|INFO1104|6,8|APR
119110351|INFO1207|5,0|APR
```

Amostra do banco de dados gerado pelo AcaDem

```
README.md
config.ini
academ
  __init__.py
  academ.py
  extractos_aggregate.py
  extractos_extractor.py
  files.py
  presentation.py
data
  disciplinas_do_curso.csv
  disciplinas_equivalentes.csv
  extractos_academicos
results
  academ.db
  demanda_disciplinas_{d}.csv
requirements.txt
```

Estrutura de pastas do AcaDem

# Extratos	Leitura Dados	Construção BD
4	70	297
40	729	1.291
80	1.562	3.791
160	2.947	5.930
320	6.027	11.083
640	10.213	22.503
1280	22.778	48.561

Desempenho do AcaDem com diferentes números de extratos

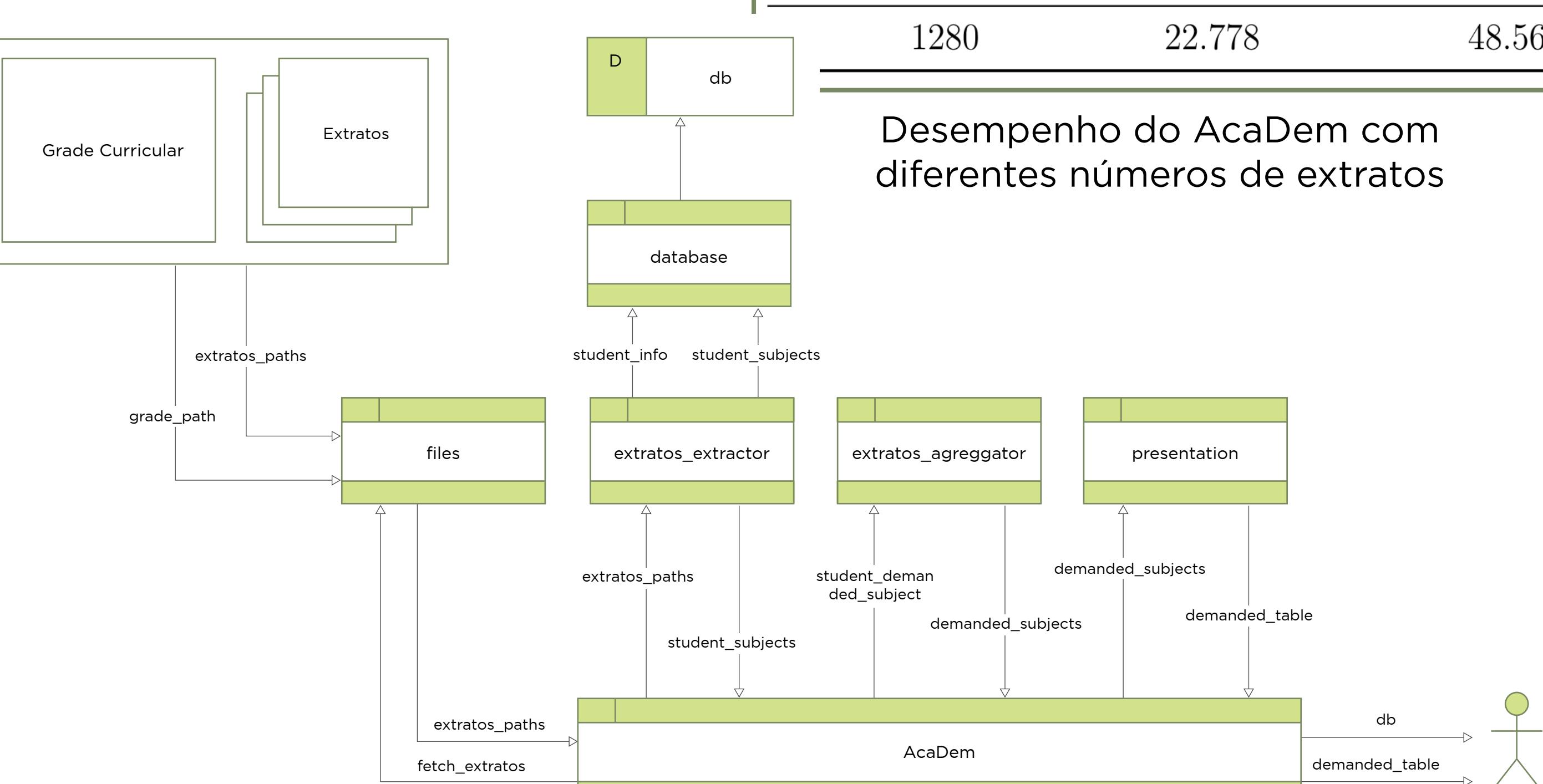


Diagrama de fluxo de dados (DFD) do AcaDem

REFERÊNCIAS

CRISAN, A.; FIORE-GARTLAND, B.; TORY, M. Passing the Data Baton : A Retrospective Analysis on Data Science Work and Workers. IEEE Transactions on Visualization and Computer Graphics, v. 27, n. 2, p. 1860-1870, fev. 2021. ISSN 1077-2626, 1941-0506, 2160-9306. Disponível em: <<https://ieeexplore.ieee.org/document/932030/>>.