

# Evasão Estudantil e Ciência de Dados: primeiros passos de uma pesquisa aplicada no contexto da Educação a Distância da Universidade Estadual do Centro-Oeste

**Sandro Rautenberg**, Alan H. Costa, Paulo R. V. do  
Carmo, Renan A. M. Nutse, Maria A. C. Knüppel,  
Marta C. R. Anciutti

Universidade Estadual do Centro-Oeste  
(UNICENTRO) – Guarapuava/PR

[sandro.rautenberg@gmail.com](mailto:sandro.rautenberg@gmail.com)

**II Workshop de Informação, Dados e Tecnologia  
(WIDaT 2018) - UFPB**

**27, 28, 29 de Novembro de 2018**

**FUNDAÇÃO  
ARAUCÁRIA**  
*Apoio ao Desenvolvimento Científico  
e Tecnológico do Paraná*



# Agenda



- Introdução
  - Contextualização
  - Justificativa
  - Objetivo
- Materiais e Métodos
- Resultados Preliminares
- Considerações Parciais
- Referências
- Agradecimentos e Contatos

# Contextualização



- **Dados** são **velozmente produzidos** e **amplamente disponíveis** nas plataformas digitais (van der AALST, 2014).
- Dados como **matéria-prima** à **Tomada de Decisão** (ECONOMIST, 2018).
- **Ciência de Dados** como métodos para **extração de informação útil** a partir de complexas e dinâmicas bases de dados (BUGNION; MANIVANNAN; NICOLAS, 2017).
- **PRESSUPOSTO** - **Organizações de Ensino a Distância** (EaD) podem **melhorar seus processos decisórios** com **Ciência de Dados**.

# Justificativa

- **Organizações de EaD** enfrentam o problema da **Evasão Estudantil** - taxa média anual de **25% dos alunos matriculados** (ABED, 2014).
- Realidade imposta ao Núcleo de Educação à Distância da Universidade Estadual do Centro-Oeste (**NEaD/UNICENTRO**) - Evasão Estudantil **em torno de 54%**.
- Construção de **modelos preditivos** para (SILVA, 2017):
  - identificar a probabilidade de um aluno evadir;
  - abstrair informações para suporte ao processo decisório;
  - Consequentemente, aumentar os índices de retenção discente.
- **Emprego** de métodos e tecnologias da **Ciência de Dados** no **estudo da Evasão Estudantil** ao contexto do NEaD/UNICENTRO.

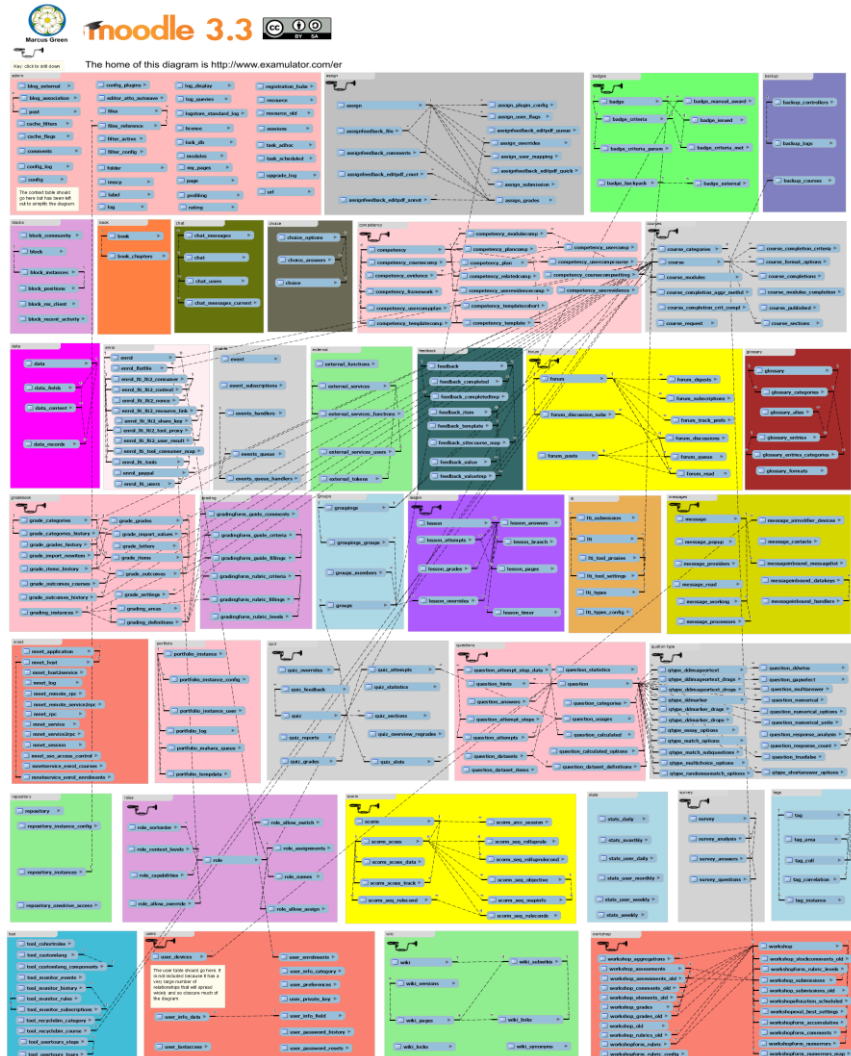


# Objetivo

**Relatar os resultados parciais alcançados no tocante aos primeiros meses de atividades do projeto de extensão institucionalizado.**

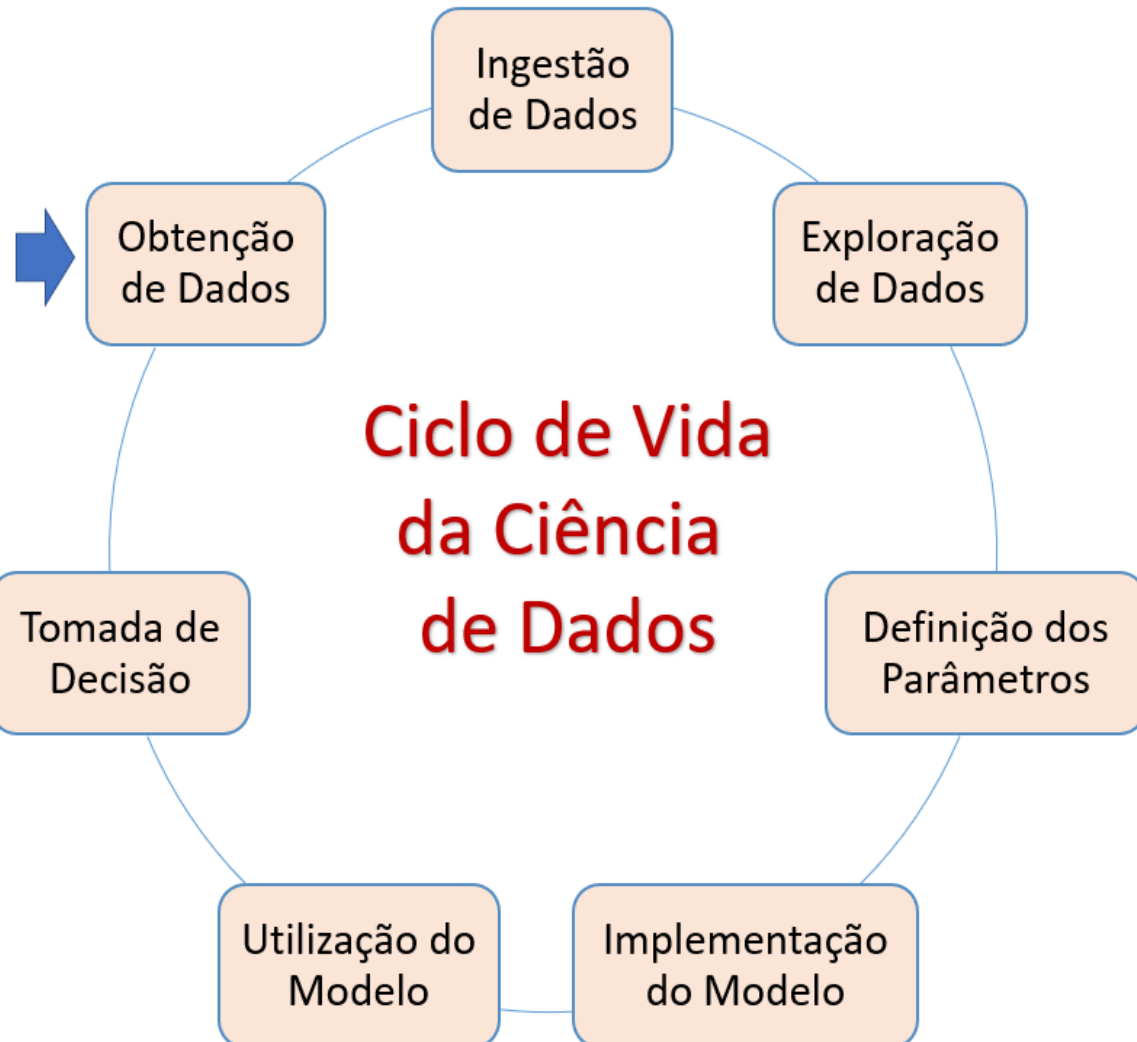


# Materiais e Métodos (I)



- **MOODLE** - ambiente que oportuniza o aprendizado a qualquer momento e em qualquer lugar, ao utilizar a Internet como plataforma de comunicação (MOODLE, 2018).
- Gerencia a interação entre atores e objetos de aprendizagem mediante **300 tabelas de dados**.
- Mantém os registros de **5.785 alunos de EaD** (formados, desistentes ou em formação) do **NEaD/UNICENTRO**.

# Materiais e Métodos (II)

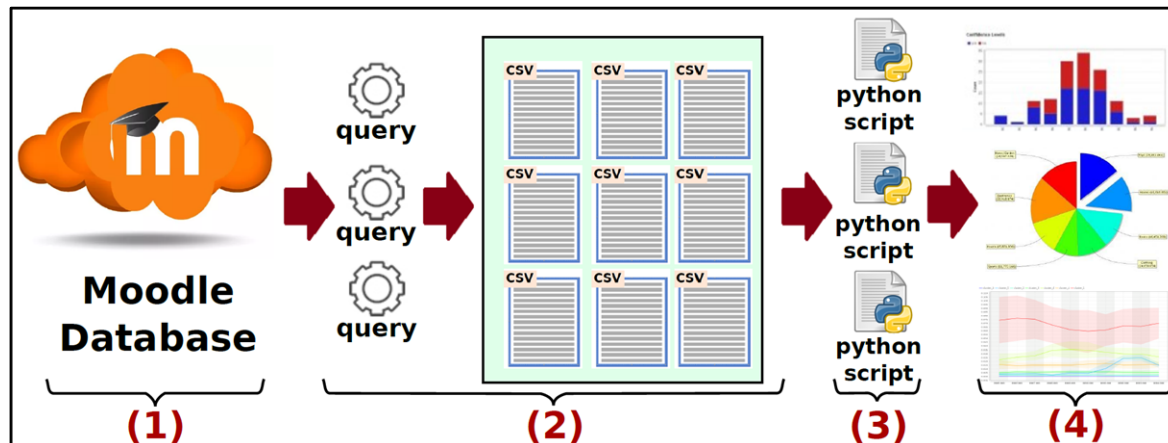


- **Ciclo de Vida da Ciência de Dados** proposto por Bugnion; Manivannan e Nicolas (2017).
- **Linguagem Python** (PYTHON.ORG, 2018) e suas bibliotecas:
  - Pandas;
  - Psycopg;
  - Matplotlib;
  - Seaborn;
  - NLTK;
  - VADER; e
  - Scikit-learn.





# Considerações Parciais



- Resultados preliminares são:
  - um **workflow** padrão de desenvolvimento de cenários de exploração e visualização de informação;
  - a concepção de uma **camada independente** para obtenção e **ingestão de dados**, mitigando os efeitos indesejáveis de versionamento do modelo de dados do MOODLE; e
  - a formalização de um **meio de comunicação** para **com os gestores** do NEaD/UNICENTRO.
- Como **trabalhos futuros**, desenvolver modelos baseados em **Aprendizagem de Máquina** para **suporte à Tomada de Decisão**.

# Referências

ABED – Associação Brasileira de Educação a Distância. **Censo EAD Brasil 2014 - Relatório Analítico da Aprendizagem a Distância no Brasil**. Curitiba, 2015.

BUGNION, P.; MANIVANNAN, A.; NICOLAS, P. R. **Scala: Guide for Data Science Professionals**. Birmingham: Packt Publishing, 2017.

ECONOMIST. **The world's most valuable resource is no longer oil, but data**. Disponível em: <<https://goo.gl/AW4XsF>>. Acesso em: 28 jul. 2018.

MOODLE. Features – MoodleDocs. Disponível em: <<https://docs.moodle.org/35/en/Features>>. Acesso em: 06 set 2018.

NEaD. **Núcleo de Ensino à Distância da Unicentro - PR » Polos**. Disponível em: <<https://goo.gl/wPYJqe>>. Acesso em: 05 set. 2018.

PYTHON.ORG. **What is Python? Executive Summary | Python.org**. Disponível em: <<https://goo.gl/QYghos>>. Acesso em: 06 set. 2018.

SILVA, F. C. da. **Gestão da Evasão na EaD: Modelo Estatístico Preditivo para os Cursos de Graduação a Distância da Universidade Federal de Santa Catarina**. Florianópolis, 2017. 137 f. Dissertação (Mestrado) - Universidade Federal de Santa Catarina, Centro Sócio-Econômico. Programa de Pós-Graduação em Administração.

UNIVERSIDADE ESTADUAL DO CENTRO-OESTE (UNICENTRO). **RESOLUÇÃO Nº 022 - PROEC/UNICENTRO, DE 16 DE JULHO DE 2018**. Disponível em: <<https://goo.gl/7VcZX4>>. Acesso em: 06 set. 2018.

van der AALST, W. Data Scientist: The Engineer of the Future. In: Interoperability of Enterprises Systems and Applications Conference (I-ESA'2014), 2014, Albi-France, **Proceedings...** Heidelberg: Springer, 2014.

# Agradecimentos

**FUNDAÇÃO  
ARAUCÁRIA**

*Apoio ao Desenvolvimento Científico  
e Tecnológico do Paraná*



**UNICENTRO**  
PARANÁ

## Contatos

[srautenberg@unicentro.br](mailto:srautenberg@unicentro.br)  
[sandro.rautenberg@gmail.com](mailto:sandro.rautenberg@gmail.com)

