

Processo seletivo discente

Temas de interesse de pesquisa

Prof. Jorge de Abreu Soares
jorge.soares@cefet-rj.br

Quem sou eu?

- Doutor e Mestre em Engenharia de Sistemas e Computação pela COPPE/UFRJ.
- Bacharel em Informática (Ciência da Computação) pela UFRJ.
- Professor desde 1996.
- Docente da Escola de Informática e Computação (EIC) do CEFET/RJ desde 2004.
 - ✓ *Técnico em Informática, Bacharelado em Ciência da Computação e Tecnologia em Sistemas para Internet*
- Membro permanente do Programa de Pós-graduação em Ciência da Computação (PPCIC) do CEFET/RJ.
- Experiência profissional: Dataprev, TRF 2, consultor.



Contatos

- E-mail:
jorge.soares@cefet-rj.br
- Home page:
<https://eic.cefet-rj.br/~jsoares/>
- CV Lattes:
<http://lattes.cnpq.br/3410221270317818>
- ResearchGate:
https://www.researchgate.net/profile/Jorge_Soares4
- LinkedIn:
<https://www.linkedin.com/in/jorge-soares-b284823/>



Pesquisa no PPCIC

- Primeira informação importante: atuo na linha de pesquisa *Análise de Dados e Aplicações* do PPCIC.
 - ☞ Veja mais informações sobre esta linha de pesquisa em <https://eic.cefet-rj.br/ppcic/index.php/linhas-de-pesquisa/>
- Áreas da Computação que me interessam:
 - ☞ Ciência de Dados
 - ☞ Pré-processamento de dados
 - ☞ *Machine Learning/Inteligência Artificial e Computacional/Data Mining*
 - ☞ Processamento de alto desempenho
 - ☞ *Blockchain*

Temas gerais de pesquisa

- As propostas de pesquisa precisam se alinhar a um dos seguintes temas:
 - ✓ Imputação de dados ausentes
 - ✓ Outros temas correlatos ao pré-processamento de dados
 - ✓ Governo eletrônico (*e-government*)
 - ✓ Agricultura de precisão
 - ✓ Análise preditiva
 - ✓ *Business Intelligence*
 - ✓ Detecção de *fake news* com aprendizado de máquina*

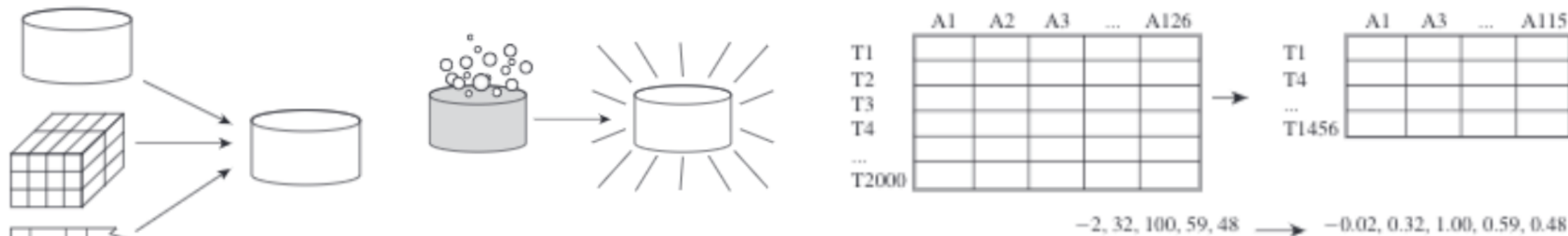
Estudo de técnicas de imputação

- Preenchimento de dados ausentes em bases de dados
 - 👉 Análises podem ser seriamente prejudicadas
- A Estatística estuda o assunto há muitos anos!
- Na Computação, é uma tarefa da etapa de limpeza de dados, no contexto da etapa de pré-processamento
 - Uso de ML, IA, IC, PAD...



Estudo de técnicas de imputação

- Origem do projeto: Imputação Composta (SOARES, 2007)
- Diversos produtos derivados do projeto:
 - Imputação em cascata (FERLIN, 2008)
 - Imputação categórica (MONTEIRO, 2008)
 - Imputação sequencial e interativa (RIBEIRO et al, 2008)
 - Algoritmos Genéticos e Redes de Kohonen na Imputação Composta (MEDINA, 2012)
 - Appraisal-Spark e a imputação por comitês *bagging* (TAVARES et al, 2018)
 - Integração de linha de comando para o workflow *Appraisal* (MOURA, 2019)
 - Imputação por comitês da família *boosting* (PEREIRA, 2020)



Fonte: Adaptado de HAN, J., KAMBER, M., PEI, J. Data Mining: Concepts and Techniques. 3rd ed. Elsevier, 2012.

Pré-processamento

- Outros temas referentes à etapa de pré-processamento também serão bem-vindos
 - Limpeza de dados (deduplicação, detecção de *outliers*)
 - Integração de dados
 - Redução de dimensionalidade
 - Transformação de dados

Parcerias no PPCIC

- No PPCIC desenvolvemos ações colaborativas em nossos projetos de pesquisa. Cada parceria acontece em função do alinhamento de interesses e demandas de pesquisa.

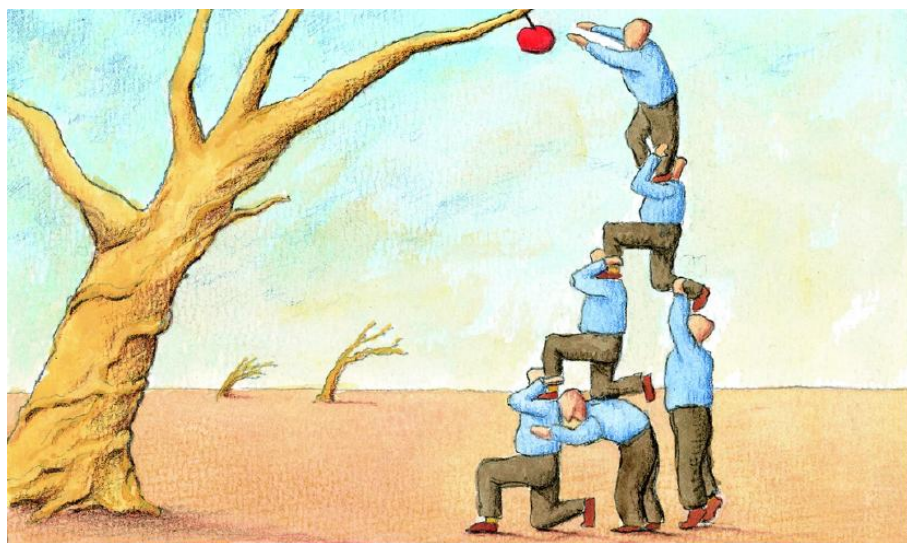


Análise de padrões em atrasos aéreos

- Projeto em colaboração com o professor [Eduardo Ogasawara](#), do PPCIC.
- Objetivo: utilizando técnicas de Ciência de Dados, identificar padrões frequentes de atrasos nas decolagens e pousos de voos em aeroportos nacionais, bem como sua predição e propagação, e com ou sem a detecção de *drifts*.
- Orientações conjuntas:
 - Alice Sternberg (2016) – PPPRO/CEFET-RJ, como coorientador
 - Leonardo Moreira (2019), como orientador
 - Lucas Giusti Tavares (em curso), como orientador



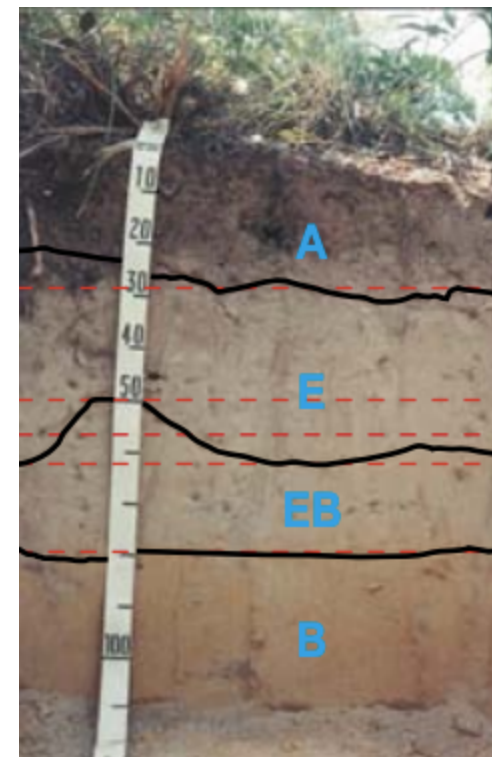
Parcerias interinstitucionais



- Programas de mestrado acadêmico em Ciências Contábeis e mestrado profissional em Controladoria e Gestão Pública – UERJ
- Controladoria-Geral do Estado do Rio de Janeiro – CGE/RJ
- Programa de Pós-graduação em Agronomia-Ciência dos Solos – UFRRJ
- Pós-graduação em Sistemas e Computação – IME

MultiSoils – Mapeamento Digital de Solos em Áreas de Exploração e Produção de Óleo e Gás

- Projeto liderado pelo Departamento de Solos da UFRRJ, com o CEFET/RJ atuando como instituição associada.
- Objetivo: conhecer os solos da região Norte e Nordeste para subsidiar a gestão territorial e ambiental de campos terrestres de exploração e produção de gás e petróleo.
- O projeto conta com a equipe da UFRRJ responsável pela aquisição de dados em campo e interpretações e análises das propriedades do solos.
- Já a equipe do CEFET/RJ será a responsável pela construção e gerenciamento do sistema de gestão de solos *MultiSoils*.



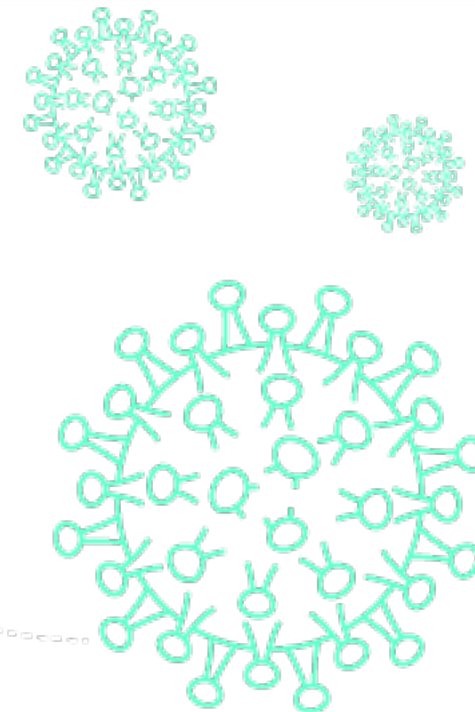
Fonte: Manual Técnico de Pedologia (2ª ed) - IBGE

MultiSoils – Mapeamento Digital de Solos em Áreas de Exploração e Produção de Óleo e Gás

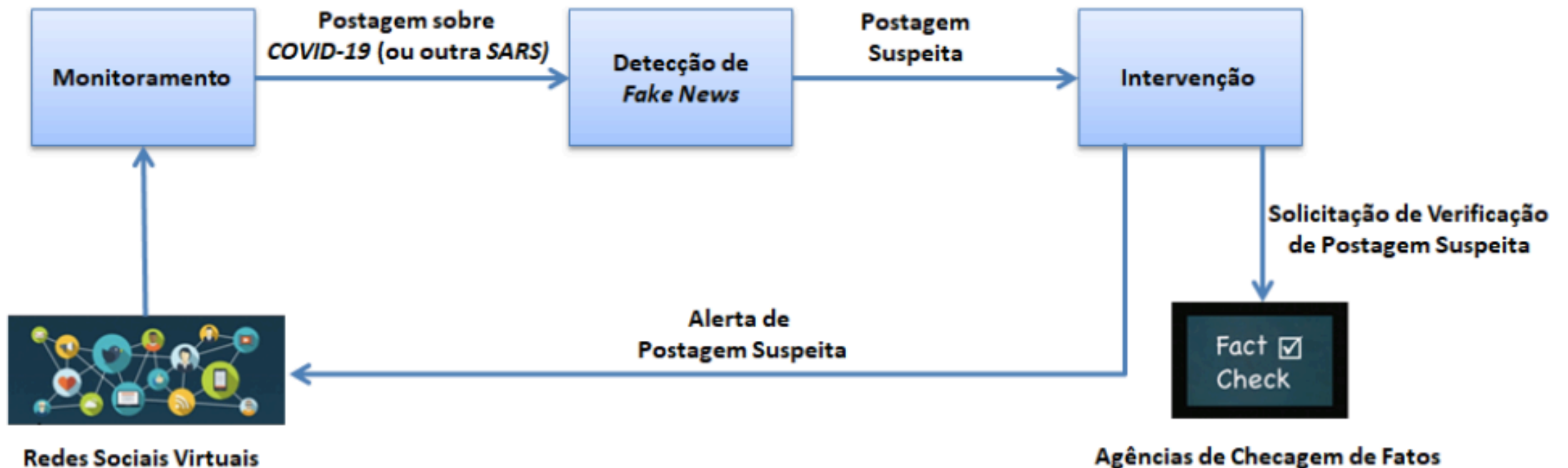
- Desenvolvimento de modelos computacionais a partir do conjunto de dados de perfis armazenados no sistema para análise e visualização utilizando Ciência de Dados.
- Equipe docente:
 - [Prof. Marcos Bacis Ceddia](#) – UFRRJ (coordenador)
 - [Prof. Diego Brandão](#) – PPCIC/CEFET-RJ
 - Prof. Jorge Soares – PPCIC/CEFET-RJ
 - [Prof. Renato Mauro](#) – EIC/CEFET-RJ

Um Ambiente Computacional para Combate Automático a *Fake News* sobre COVID-19 e outras SARS em Redes Sociais Virtuais

- Projeto coordenado pelo [prof. Ronaldo Goldschmidt](#), da Seção de Engenharia de Computação (SE/9) do IME, com o CEFET/RJ atuando como instituição parceira.
- Outras instituições parceiras: UEZO e FAETEC.
- Objetivos: pesquisar, modelar, implementar e avaliar um protótipo funcional de um ambiente computacional de combate automático a *fake News* sobre COVID-19 e outras SARS escritas em Língua Portuguesa e divulgadas em redes sociais virtuais.



Um Ambiente Computacional para Combate Automático a *Fake News* sobre COVID-19 e outras SARS em Redes Sociais Virtuais



Detecção de fraudes nas atividades de controladoria e auditoria públicas

- Projeto coordenado pelo CEFET/RJ, com a UERJ atuando como instituição associada.
- Objetivo: detecção de fraudes fiscais na movimentação financeira dos órgãos e entidades do Estado do Rio de Janeiro.
- Existem entendimentos para que a CGE-RJ também participe como instituição parceira/associada do projeto.
- Equipe docente:
 - Prof. Jorge Soares – CEFET/RJ (coordenador)
 - [Prof. Francisco José dos Santos Alves](#) – UERJ



Perfil esperado de um(a) aluno(a) de mestrado

- Possuir QUALQUER formação de graduação!
- Ser **DEDICADO, PROATIVO, PERSISTENTE e COLABORATIVO**
- Saber programar em qualquer linguagem
- Must know the English language reasonably (essential skills are reading and writing)
- Ter disposição para escrever artigos científicos!
- Ter disponibilidade de agenda para reuniões semanais nos turnos matutino e vespertino.



Perguntas comuns...

- Professor, você pode me ajudar olhando a proposta que vou submeter no processo seletivo?
- Tenho um projeto bacana em mente... Posso aplica-lo na proposta de pesquisa que submeterei no processo seletivo?
- Preciso obrigatoriamente manter o tema do projeto de pesquisa quando eu entrar no mestrado?
- Professor, você estará presente na minha banca de arguição?



Vamos nos inscrever então no processo seletivo?



- 👉 Acesse o endereço <https://eic.cefet-rj.br/ppcic/index.php/category/selecao/>, leia atentamente o edital do certame, e mãos à obra no projeto! 👍



**Obrigado pelo
interesse!**

**E boa sorte no
processo seletivo!**

