

Projetos Finais de Graduação para o 1s2024 (MC030)

Professor Responsável: [Prof. Dr. André Santanchè](#)

Título: Harena - Plataforma de Aprendizagem baseada em Casos (Testes Sistemáticos)

Descrição do Projeto: Harena é uma plataforma de aprendizagem baseada na resolução de casos. Ela nasceu no contexto da saúde e envolve uma colaboração com médicos da Unicamp e outras universidades no Brasil, Portugal e Holanda. A plataforma está sendo construída para explorar aspectos de ciência de dados para suporte a construção de casos baseados em fenômenos do mundo real, por exemplo, construção de casos clínicos baseados na análise de dados de saúde. Ela também está sendo preparada para ter suporte de uma tutoria inteligente.

Este projeto envolve duas partes: a implementação de uma infraestrutura de testes sistemáticos para tornar a plataforma mais robusta.

O projeto requer que os membros tenham disponibilidade para pelo menos um encontro semanal e dedicação compatível com o número de créditos da disciplina (12 créditos).

Trabalho: Individual e Em Equipe (discriminar na descrição a divisão dos trabalhos)

Professor Responsável: [Prof. Dr. André Santanchè](#)

Título: Harena - Plataforma de Aprendizagem baseada em Casos (Segurança e Permissões)

Descrição do Projeto: Harena é uma plataforma de aprendizagem baseada na resolução de casos. Ela nasceu no contexto da saúde e envolve uma colaboração com médicos da Unicamp e outras universidades no Brasil, Portugal e Holanda. A plataforma está sendo construída para explorar aspectos de ciência de dados para suporte a construção de casos baseados em fenômenos do mundo real, por exemplo, construção de casos clínicos baseados na análise de dados de saúde. Ela também está sendo preparada para ter suporte de uma tutoria inteligente.

Este projeto envolve a reformulação do sistema de segurança e permissões da plataforma. Atualmente, a plataforma conta com um complexo sistema de permissões para leitura, compartilhamento e edição, seja para indivíduos, grupos ou instituições. Com o crescimento constante do número de usuários, a verificação de tais permissões tem tornado os processos lentos. A reformulação irá requerer o estudo de sistemas de permissão correntes que possam resolver o problema. Isso incluirá a reformulação do processo de autorização e login.

O projeto requer que os membros tenham disponibilidade para pelo menos um encontro semanal e dedicação compatível com o número de créditos da disciplina (12 créditos).

Trabalho: Individual e Em Equipe (discriminar na descrição a divisão dos trabalhos)

Professor Responsável: [Prof. Dr. André Santanchè](#)

Título: Toolkit para Análise de Discurso e LDA (Latent Dirichlet Allocation)

Descrição do Projeto: Este projeto está inserido em uma pesquisa que envolve análise de discurso, tomando como base a abordagem de modelagem de tópicos, especialmente o LDA (Latent Dirichlet Allocation). Dentre os principais desafios está a análise de evolução dos tópicos, sua coesão e diversidade. Alguns resultados podem ser vistos no artigo: Evaluating the cohesion of municipalities' discourse during the COVID-19 pandemic (<https://sol.sbc.org.br/index.php/sbbd/article/view/17888>).

Este projeto, especificamente, envolve a materialização de métodos e métricas desenvolvidas nesta pesquisa na forma de um toolkit. Ele dará continuidade a um projeto anterior que implementou um toolkit para análise temporal de tópicos.

O projeto requer que o participante tenha disponibilidade para pelo menos um encontro semanal e dedicação compatível com o número de créditos da disciplina (12 créditos).

Trabalho: Individual

Professor Responsável: [Prof. Dr. André Santanchè](#)

Título: Componentes para a Construção de Narrativas Interativas

Descrição do Projeto: Web Component é uma tecnologia que tem crescido no contexto da Web como plataforma. Ela tem sido fundamental, seja para a construção de aplicações sobre a Web, seja para enriquecer páginas com interatividade e processamento no cliente.

Este projeto envolve a construção de uma biblioteca de componentes usando o framework Oid (<https://mundorum.github.io/oid/>). A biblioteca está voltada ao enriquecimento de narrativas interativas na plataforma Harena, uma plataforma de aprendizagem baseada na resolução de casos. Ela nasceu no contexto da saúde e envolve uma colaboração com médicos da Unicamp e outras universidades no Brasil, Portugal e Holanda.

O projeto requer que o participante tenha disponibilidade para pelo menos um encontro semanal e dedicação compatível com o número de créditos da disciplina (12 créditos).

Trabalho: Individual e Em Equipe (discriminar na descrição a divisão dos trabalhos)

Professor Responsável: [Prof. Dr. André Santanchè](#)

Título: Software para Editoração de Perguntas e Respostas

Descrição do Projeto: Este projeto está vinculado ao Projeto de Letramento Midiático, Informacional e Diálogo Intercultural - Pergunte a um/a cientista. Trata-se de um projeto em que a comunidade submete perguntas a cientistas - vide detalhes no site: <https://pergunte.online/>

Este projeto, especificamente, envolve a adaptação de algum software de plataforma online preexistente para dar suporte ao processo de: submissão de perguntas; filtragem, tratamento e encaminhamento de perguntas a cientistas que colaboram com

o projeto; suporte à edição de resposta desses cientistas; editoração final da resposta e publicação em web site.

Deve ser adaptado algum software open-source existente com características similares. Também está sendo considerada a adaptação de algum plugin para uma plataforma de gerenciamento de conteúdo, a exemplo do Wordpress.

O projeto requer que o participante tenha disponibilidade para pelo menos um encontro semanal e dedicação compatível com o número de créditos da disciplina (12 créditos).

Trabalho: Individual e Em Equipe (discriminar na descrição a divisão dos trabalhos)

Professor Responsável: [Prof. Dr. Bruno Barbieri de Pontes Cafeo](#)

Título: Explorando Técnicas de Mineração de Processos em Code Samples para Aprimorar o Desenvolvimento de Software

Descrição do Projeto: Este projeto propõe a aplicação de técnicas de mineração de processos em code samples. Code Samples são projetos completos de software disponibilizados por grandes organização (p.ex., Google, Amazon e Microsoft) para auxiliar desenvolvedores externos no aprendizado e atualização de conceitos implementados em seus produtos (p.ex., Android, AWS e Azure). Inspirados pelo sucesso do process mining, que utiliza a ciência de dados para descobrir, validar e melhorar workflows organizacionais, buscamos adaptar essa abordagem para o contexto de code samples. Os code samples servirão como unidades de análise, permitindo-nos extrair padrões, melhores práticas e áreas de otimização no processo de criação e evolução desses artefatos.

Trabalho: Individual

Professor Responsável: [Prof. Dr. Bruno Barbieri de Pontes Cafeo](#)

Título: Estudo Exploratório e Longitudinal das Práticas de Testes de Software em Code Samples

Descrição do Projeto: Este projeto propõe um estudo detalhado para explorar as práticas de testes de software em code samples, visando uma compreensão abrangente de como essas práticas evoluem ao longo do tempo. Code Samples são projetos completos de software disponibilizados por grandes organização (p.ex., Google, Amazon e Microsoft) para auxiliar desenvolvedores externos no aprendizado e atualização de conceitos implementados em seus produtos (p.ex., Android, AWS e Azure). Esses code samples serão analisados para identificar estratégias de teste, preferências de desenvolvedores e mudanças nas abordagens, proporcionando insights valiosos para aprimorar a eficácia dos testes nesse contexto que está em amplo crescimento.

Trabalho: Individual

Professor Responsável: [Prof. Dr. Bruno Barbieri de Pontes Cafeo](#)

Título: Explorando Technical Debt em Code Samples: Uma Análise Detalhada das Implicações para a Manutenibilidade do Código

Descrição do Projeto: Este estudo visa aprofundar a compreensão de technical debt (débito técnico) em code samples. Code Samples são projetos completos de software disponibilizados por grandes organizações (p.ex., Google, Amazon e Microsoft) para auxiliar desenvolvedores externos no aprendizado e atualização de conceitos implementados em seus produtos (p.ex., Android, AWS e Azure). O technical debt refere-se a compromissos tomados durante o desenvolvimento que podem resultar em custos adicionais no futuro. Ao analisar sua presença em code samples, buscamos identificar padrões, tipos comuns de technical debt e avaliar suas implicações na manutenibilidade do código desse tipo de artefato. Mais especificamente, este estudo colaborativo tem como objetivo realizar uma análise conjunta e inicial por meio de code smells e refatorações em code samples, com dois alunos se concentrando em áreas específicas. Enquanto um aluno se dedica à identificação e categorização de code smells, o outro foca na análise das refatorações existentes nos code samples. A proposta é, posteriormente, examinar a possível relação entre os code smells identificados e as refatorações realizadas, buscando entender como as práticas de refatoração podem influenciar na mitigação ou agravamento dos problemas identificados. Este enfoque integrado visa proporcionar insights mais abrangentes para o aprimoramento da qualidade dos code samples.

Trabalho: Em Equipe (discriminar na descrição a divisão dos trabalhos)

Professor Responsável: [Prof. Dr. Bruno Barbieri de Pontes Cafeo](#)

Título: Entendendo Práticas de Desenvolvimento de Code Samples: Uma Investigação por Meio de Entrevistas com Desenvolvedores

Descrição do Projeto: Este projeto visa explorar e compreender as práticas adotadas por desenvolvedores na criação e evolução de code samples, trechos de código-fonte que representam unidades significativas de implementação. Através de entrevistas com desenvolvedores atuantes em diferentes organizações, buscamos obter insights sobre as estratégias, desafios, e decisões tomadas durante o processo de criação, manutenção e evolução de code samples. O estudo proporcionará uma visão detalhada das práticas de desenvolvimento e permitirá identificar padrões e tendências no cenário de code samples.

Para este projeto é imprescindível um nível avançado de inglês para a condução das entrevistas, bem como posterior extração dos dados.

Trabalho: Individual

Professor Responsável: [Prof. Dr. Bruno Barbieri de Pontes Cafeo](#)

Título: Desenvolvimento de Métricas de Evolução em Sistemas Configuráveis Pre-processados utilizando o Modelo GQM

Descrição do Projeto: Este projeto propõe a aplicação do modelo Goal-Question-Metric (GQM) para a definição e implementação de métricas de evolução em sistemas configuráveis pre-processados. Sistemas configuráveis pre-processados são aqueles em que características específicas podem ser ativadas ou desativadas antes da compilação (p.ex., kernel do Linux, Mongo DB, entre outros). O GQM será utilizado como estrutura para estabelecer metas, identificar questões relevantes e, por fim, definir métricas objetivas que permitam avaliar e monitorar a evolução desses sistemas ao longo do tempo.

Trabalho: Individual

Professor Responsável: [Prof. Dr. Bruno Barbieri de Pontes Cafeo](#)

Título: Revisão Sistemática de Técnicas de Mineração de Padrões de Uso de APIs na Literatura

Descrição do Projeto: Este projeto propõe uma revisão sistemática para analisar as técnicas de mineração de padrões de uso de APIs presentes na literatura. O objetivo é compreender as abordagens utilizadas para identificar, extrair e analisar padrões de uso de APIs em diversos contextos. A pesquisa será conduzida em dupla, com ambas as partes colaborando em todas as etapas do processo, desde a revisão inicial da literatura até a análise final dos resultados. No final cada aluno realizará a análise dos dados de um conjunto de questões de pesquisa.

Trabalho: Em Equipe (discriminar na descrição a divisão dos trabalhos)

Professor Responsável: [Prof. Dr. Bruno Barbieri de Pontes Cafeo](#)

Título: Mineração de Padrões de Uso de APIs na Biblioteca Standard C

Descrição do Projeto: Este projeto visa aplicar técnicas consolidadas de mineração de dados para extrair padrões de uso de APIs na biblioteca Standard C da linguagem de programação C. Padrões de uso de APIs referem-se a práticas recorrentes adotadas pelos desenvolvedores ao utilizar interfaces de programação de aplicativos (APIs). Essas práticas incluem sequências comuns de chamadas de funções, estratégias de manipulação de dados, e abordagens para lidar com situações específicas. O objetivo é identificar comportamentos frequentes e tendências no uso das APIs padrão, oferecendo insights valiosos sobre práticas de desenvolvimento, desafios comuns e possíveis melhorias em documentação da biblioteca.

Trabalho: Individual

Professor Responsável: [Prof. Dr. Bruno Barbieri de Pontes Cafeo](#)

Título: Revisão de Literatura Multivocal sobre Feature Flags

Descrição do Projeto: Este projeto propõe iniciar uma revisão de literatura multivocal sobre feature flags, explorando diversas perspectivas e práticas relacionadas a essa abordagem no desenvolvimento de software. Feature flags, ou feature toggles, são

mecanismos que permitem ativar ou desativar funcionalidades em tempo de execução. Esta revisão visa compreender a variedade de conhecimento existente na literatura formal e em sites, blogs, foruns, etc. sobre feature flags, abordando aspectos técnicos, práticos e conceituais. O projeto servirá como um ponto de partida para uma análise abrangente que capture as nuances e evoluções dessa prática no cenário de desenvolvimento de software.

Trabalho: Individual

Professor Responsável: [Prof. Dr. Edmundo Roberto Mauro Madeira](#)

Título: Aprendizado de Máquina aplicado a alocação de recursos de rede

Descrição do Projeto: Infraestruturas de rede necessitam gerenciar de forma eficiente os seus recursos - capacidade de processamento, armazenamento, transmissão de dados, dentre outros - para atender as necessidades das aplicações e serviços acessados por seus usuários. Devido a variações na demanda desses recursos - alteração no número de usuários, mobilidade dos usuários, ou alterações intrínsecas da própria aplicação, modificações nas configurações das redes se fazem necessárias. Este projeto tem como objetivo estudar a aplicação de soluções baseadas em aprendizado de máquina para identificar as alterações necessárias, ao longo do tempo, para realizar de forma eficiente a (re)distribuição dos recursos da rede dado suas variações na demanda.

Trabalho: Individual

Professora Responsável: [Pro^a. Dr^a. Eliane Martins](#)

Título: Uso de gêmeos digitais nos testes de sistemas ciber-físicos

Descrição do Projeto: Um gêmeo digital (digital twin) é uma réplica virtual de um objeto ou sistema físico, podendo receber como entrada dados em tempo real de objetos ou sistemas físicos e, com simulações, predizer a saída produzida por estes objetos ou sistemas em resposta às entradas recebidas. Com isso é possível ter ideias sobre o desempenho ou problemas que possam surgir durante a execução. O objetivo do projeto é o estudo dessa tecnologia, as plataformas de apoio existentes e utilização de uma destas plataformas em um exemplo de um sistema ciber-físico. Um exemplo de uso de um dos gêmeos digitais estudados também faz parte dos objetivos.

O trabalho pode ser feito em equipe de até 2 alunos. Ambos terão que estudar sobre gêmeos digitais e estabelecer os critérios para a seleção daquele a ser utilizado no estudo de caso. Depois, será feito o uso de programação em pares, em que um ficará responsável por criar os testes e o outro ficará responsável pela implementação de um pequeno exemplo usando a plataforma escolhida.

Trabalho: Em Equipe (discriminar na descrição a divisão dos trabalhos)

Professora Responsável: [Pro^a. Dr^a. Eliane Martins](#)

Título: Avaliação de método combinando testes baseados em modelos e testes dirigidos pelo comportamento para APIs

Descrição do Projeto: Neste projeto o objetivo é realizar testes funcionais de serviços RESTful especificados com OPENAPI ou Swagger e avaliar a cobertura dos testes. A cobertura serve para determinar o que foi e o que não foi testado, permitindo, com isso, avaliar se é necessário realizar mais testes ou não. A cobertura é avaliada em termos da especificação, o que compreende, entre outros: caminhos (endpoints), parâmetros, status de mensagens testados. Além de estudos sobre REST API e como testá-las, o objetivo é aplicar o método a um estudo de caso. O projeto pode ser individual ou em grupo. Caso seja em grupo, cada aluno ficará responsável por um estudo de caso diferente.

Trabalho: Em Equipe (discriminar na descrição a divisão dos trabalhos)

Professor Responsável: [Prof. Dr. Hilder Vitor Lima Pereira](#)

Título: Reimplementação de uma cifra completamente homomórfica usando técnicas avançadas de aritmética computacional

Descrição do Projeto: Criptografia completamente homomórfica (FHE, do inglês, fully homomorphic encryption) é uma ferramenta poderosíssima, pois permite que um servidor na nuvem compute qualquer função sobre dados cifrados.

Atualmente, existem três famílias de FHE, sendo uma delas baseadas em inteiros, chamada de FHEZ, que atualmente é mais lenta que as outras duas famílias. No entanto, a comparação não é justa, pois existe atualmente apenas uma implementação muito simples da FHEZ.

O objetivo deste projeto é aplicar técnicas avançadas de álgebra e aritmética computacional reimplementar FHEZ e, então, compará-la com as outras famílias de FHE.

Você vai aprender sobre criptografia, sobre FHE e sobre como operações com inteiros de precisão arbitrária, polinômios e outras estruturas algébricas podem ser implementadas de forma eficiente.

Trabalho: Individual

Professor Responsável: [Prof. Dr. Hilder Vitor Lima Pereira](#)

Título: Análise de crédito sem analisar os dados

Descrição do Projeto: No Brasil, algumas organizações, como a Serasa, são conhecidas por calcularem um score que tenta indicar o risco de uma pessoa ficar inadimplente ao fazer um empréstimo. Isso é conhecido como análise de crédito.

Se por um lado isso facilita a tomada de decisão de empresas como bancos, operadoras de cartões e outras instituições financeiras, o cálculo desses scores utiliza muitos dados pessoais, como contas de energia elétrica e contas de água, o que representa um risco para a privacidade dos cidadãos.

Mas será que é possível fazer análise de crédito sem que todos esses dados pessoais sejam compartilhados?

Neste projeto, vamos usar técnicas de computação sobre dados cifrados para tentar responder positivamente essa pergunta.

Trabalho: Individual

Professor Responsável: [Prof. Dr. Julio César López Hernández](#)

Título: Implementação em Software de Algoritmos Criptográficos Pós-Quânticos

Descrição do Projeto: A criptografia de chave pública (criptação e assinatura digital), que atualmente se utiliza na maioria das aplicações práticas, tem sua segurança baseada na resolução dos problemas de fatoração e cálculo do logaritmo discreto. O algoritmo de Shor é um algoritmo quântico polinomial para fatorar e calcular logaritmos discretos. Com os avanços na construção de um computador quântico, essas máquinas seriam uma ameaça real aos atuais sistemas criptográficos. A criptografia pós-quântica é o desenvolvimento de novos algoritmos de chave pública projetados para serem resistentes aos computadores clássicos de hoje e futuros computadores quânticos. Em 2017, o Instituto Nacional de Padrões Americano (NIST) lançou um concurso para selecionar novos algoritmos de chave pública a serem padronizados. O objetivo deste projeto é estudar e implementar em software alguns dos algoritmos selecionados pelo NIST.

Trabalho: Individual

Professor Responsável: [Prof. Dr. Julio César López Hernández](#)

Título: Provas de conhecimento-zero e suas Aplicações a Blockchain e Criptomoedas

Descrição do Projeto: Avanços recentes em criptografia resultaram na criação de novas tecnologias para a privacidade de dados. Entre eles estão as tecnologias zk-SNARKS (argumento sucinto não interativo de conhecimento-zero) e ZK-STARKS (argumento de conhecimento transparente escalável de conhecimento-zero). Provas de conhecimento-zero permitem que um indivíduo prove para outro que um argumento é verdadeiro, sem precisar fornecer nenhuma outra informação que vai além da validade do argumento. Essas tecnologias tem encontrado aplicações em criptomoedas (Zcash) e em sistemas de pagamento baseado em blockchain. Atualmente, as provas de conhecimento-zero podem atingir tamanhos de prova de centenas de bytes e verificações do ordem de milissegundos, independentemente do tamanho da proposição que está sendo provada. O objetivos do projeto é estudar os aspectos teóricos e práticos das provas de conhecimento-zero..

Trabalho: Individual

Professor Responsável: [Prof. Dr. Julio César López Hernández](#)

Título: Implementação em Software de Algoritmos Simétricos

Descrição do Projeto: Um algoritmo simétrico (ou de chave secreta) é um método para cifrar/decifrar mensagens entre duas entidades que compartilham uma chave (um segredo). No processo de cifragem, o algoritmo simétrico junto com a chave transforma uma mensagem numa cadeia de bits que não pode ser entendida por ninguém que não possua a mesma chave utilizada para cifrá-la. Uma vez que o destinatário da mensagem, quem possui a mesma chave, recebe a mensagem cifrada, o algoritmo simétrico junto com a chave reverte a operação, recuperando o texto da mensagem original. Dessa forma, os algoritmos simétricos são utilizados para comunicações seguras entre duas entidades (ou pessoas) que compartilham uma chave secreta. Em 2001, o Instituto Nacional de Padrões Americano (NIST) anunciou o novo padrão da criptografia simétrica, chamado AES (Advanced Encryption Standard). É um algoritmo que utiliza chaves de 128, 192 ou 256 bits para cifrar mensagens de 128 bits. Embora, o AES seja suportado em muitas plataformas, com instruções em hardware para acelerar o processo de cifragem, o AES não foi projetado para aplicações em dispositivos limitados. Recentemente, o NIST anunciou uma lista de 10 algoritmos candidatos para o padrão de Criptografia Leve, isto é, algoritmos simétricos que foram projetados para dispositivos limitados.

O objetivo principal deste projeto é estudar algoritmos criptográficos simétricos, no contexto da Criptografia Leve, e desenvolver códigos otimizados em C para diferentes plataformas ARM.

As principais atividades do projeto são:

1. Estudar os conceitos básicos de criptografia simétrica (criptografia, funções de resumo, ver referências [1] e [2])
2. Estudar 2 algoritmos da lista: ASCON, Elephant, GIFT-COFB, Grain-128AEAD, ISAP, Xoodyak, TinyJambu, SPARKLE, Romulus, PHOTON-Beetle. A especificação de cada algoritmo é dada na referência [3].
3. Estudar técnicas de programação para armv8/armv9.
4. Desenvolver implementações em C dos algoritmos selecionados.
5. Estudar ataques de canal lateral (side-channel attacks) em software (implementações em tempo-constante).
5. Desenvolver implementações otimizadas em C dos algoritmos selecionados.
6. Realizar medições de tempo e comparar com outras bibliotecas.
7. Escrever um relatório, e se possível submeter um artigo para publicação.

Referências:

- [1]. Cryptography and Network Security, William Stallings, 7 ed, 2020
- [2]. Cryptography: Theory and Practice, 4 ed. Stinson, Paterson, 2019
- [3]. <https://csrc.nist.gov/Projects/lightweight-cryptography/news>

Trabalho: Individual

Professor Responsável: [Prof. Dr. Julio Cesar dos Reis](#)

Título: interfaces vestíveis para a captura e rotulação de dados fisiológicos

Descrição do Projeto: A captura, processamento e interpretação de dados fisiológicos podem desenvolver um papel relevante no comportamento de sistemas interativos. Este projeto visa explorar artefatos existentes no mercado que permitam a coleta de dados fisiológicos, como smartwatches. Visamos projetar e construir uma

aplicação que permita capturar esses dados a partir de instrumentos existentes e desenvolver cenários adequados para a rotulação desses conforme aspectos emocionais das pessoas. Objetivamos gerar e analisar um conjunto de dados que possa ser útil para detectar as emoções das pessoas a partir de sua frequência cardíaca e outros sinais fisiológicos.

Trabalho: Individual

Professor Responsável: [Prof. Dr. Julio Cesar dos Reis](#)

Título: Avaliação na plataforma OpenDesign

Descrição do Projeto: O Opendesign é uma plataforma online para o apoio na condução colaborativa e distribuída do design de sistemas computacionais. A condução de processos de avaliação do design desenvolve um papel chave para o aprimoramento de features no software e para o informar o redesign em estudo de um projeto. A plataforma OpenDesign demanda a construção de novos mecanismos que suportem designers e outros stakeholders envolvidos na condução e documentação de avaliação. O objetivo deste projeto é conceber, implementar e avaliar ferramentas para o registro de diferentes tipos de métodos de avaliação na plataforma.

Trabalho: Individual ou Em Equipe (discriminar na descrição a divisão dos trabalhos)

Professor Responsável: [Prof. Dr. Julio Cesar dos Reis](#)

Título: Processamento de diálogos textuais na língua Portuguesa

Descrição do Projeto: Sistemas de chatbots desempenham papel chave como uma ferramenta de autoatendimento a clientes em grandes empresas. A construção desses sistemas ainda é repleta de desafios no design da experiência do usuário e no funcionamento do sistema visando uma experiência agradável de uso. Por exemplo, clientes conseguem resolver problemas menos complexos sem necessitar de diversos atendimentos ou delongadas interações com atendentes humanos. Sistemas de chatbots requerem uso de técnicas de processamento de linguagem natural. Este projeto visa investigar o reconhecimento de entidades no processamento de texto na língua Portuguesa em sistemas conversacionais computadorizados. Exploraremos modelos de classificação automática de entidades.

Trabalho: Individual

Professor Responsável: [Prof. Dr. Julio Cesar dos Reis](#)

Título: Alinhamento de grafos de conhecimento

Descrição do Projeto: Grafos de conhecimento definem fatos expressos como triplas considerando sujeito, predicado e objeto na representação do conhecimento. Usualmente diversos grafos de conhecimento são publicados em um determinado domínio. É relevante criar alinhamentos tanto de classes que modelam conceitos quanto entre instâncias dessas classes definidas em diferentes grafos de conhecimento. O objetivo deste projeto é estudar técnicas de alinhamento de entidades

expressas em grafos de conhecimento. Usaremos conjunto de dados existentes na Ontology Alignment Evaluation Initiative para avaliar os métodos concebidos em análises experimentais.

Trabalho: Individual

Professor Responsável: [Prof. Dr. Julio Cesar dos Reis](#)

Título: Descrição semântica de publicações científicas.

Descrição do Projeto: Publicações científicas podem ser melhor recuperadas e analisadas quando o significado dos atributos que caracterizam os dados da publicação são codificados em modelos computacionais que representam explicitamente a semântica. O objetivo deste projeto é propor e desenvolver um sistema que coleta dados sobre artigos científicos e enriquece semanticamente os registros por meio de vocabulários que descrevem precisamente os conceitos do domínio. O trabalho envolverá estudar linguagens para a criação e consulta de ontologias.

Trabalho: Individual ou Em Equipe (discriminar na descrição a divisão dos trabalhos)

Professor Responsável: [Prof. Dr. Julio Cesar dos Reis](#)

Título: Sistema para explorar dados interconectados abertos.

Descrição do Projeto: Um número crescente de dados interconectados abertos (Linked Open Data) são publicados e disponíveis em repositórios na Web. Há diversas oportunidades no uso e integração desses dados interconectados, com semântica interpretável pela máquina, em diferentes domínios. Este projeto visa construir funcionalidades de software para consultar e combinar fatos descritos nestes repositórios. O trabalho exigirá o estudo de uma linguagem de consulta para acesso à fontes de dados na Web Semântica (SPARQL).

Trabalho: Individual

Professor Responsável: [Prof. Dr. Julio Cesar dos Reis](#)

Título: Sistemas de questão e respostas usando bases RDF.

Descrição do Projeto: Sistemas de questões e respostas fazem parte de um esforço contínuo para aprimorar a interação homem-computador. Este projeto objetiva implementar um sistema que permita interpretar uma questão em linguagem natural e obter uma consulta estruturada. Visamos considerar consultas em grafos de conhecimento descritos em RDF. O trabalho exigirá o estudo de uma linguagem de consulta para acesso à fontes de dados na Web Semântica (linguagem SPARQL). As respostas obtidas serão convertidas em uma representação final para o usuário.

Trabalho: Individual

Professor Responsável: [Prof. Dr. Julio Cesar dos Reis](#)

Título: Visualização de ontologias.

Descrição do Projeto: Ontologias permitem representar conceitos em um domínio e podem ser úteis para usuários fazerem sentido de conceitos e suas relações. Contudo, poucos estudos investigam a interação com essas estruturas. Este projeto visa projetar e construir um sistema com uso de ontologias para permitir usuários navegarem entre conceitos de disciplinas do curso de engenharia e ciência da computação. Utilizaremos design centrado no usuário e técnicas participativas para elaborar a estrutura de visualização das ontologias. Este sistema poderá permitir que aluno(s) melhor entendam os conceitos e suas relações nas disciplinas que compõem o curso.

Trabalho: Individual

Professor Responsável: [Prof. Dr. Leandro Aparecido Villas](#)

Título: Aprendizado Federado

Descrição do Projeto: O aprendizado federado é uma abordagem de aprendizado de máquina em que o treinamento do modelo ocorre em dispositivos de borda ou clientes distribuídos, mantendo os dados localmente e compartilhando apenas os parâmetros atualizados entre os dispositivos, preservando a privacidade e a segurança dos dados.

Trabalho: Individual

Professor Responsável: [Prof. Dr. Leandro Aparecido Villas](#)

Título: Aprendizado Distribuído

Descrição do Projeto: O aprendizado de máquina distribuído é uma abordagem em que o treinamento de um modelo de aprendizado de máquina é realizado em várias máquinas ou dispositivos em rede, permitindo processamento paralelo e colaboração para melhorar a eficiência e escalabilidade do treinamento.

Trabalho: Individual

Professor Responsável: [Prof. Dr. Luiz Fernando Bittencourt](#)

Título: Aprendizado de máquina federado no simulador MobFogSim

Descrição do Projeto: O simulador MobFogSim simula a mobilidade de dispositivos de usuário que podem executar aplicações em dispositivos de processamento presentes na borda da rede, ou computação em névoa, complementando as funcionalidades da computação em nuvem. Este projeto tem como objetivo implementar novas funcionalidades no simulador, mais especificamente a implementação de aprendizado de máquina federado nos dispositivos do simulador.

Trabalho: Individual ou Em Equipe (discriminar na descrição a divisão dos trabalhos)

Professor Responsável: [Prof. Dr. Luiz Fernando Bittencourt](#)

Título: Elasticidade e auto-distribuição

Descrição do Projeto: Sistemas que se auto-distribuem (do inglês self-distributing systems) são sistemas capazes de replicar, em tempo de execução, componentes que compõem sua própria estrutura, lidando com estado atrelado a esses componentes de forma transparente. Esses sistemas têm por objetivo explorar ambientes contemporâneos como computação em nuvem, que apesar de terem software que dão apoio a elasticidade, não lidam bem com a replicação de serviços com estado. Este projeto tem por objetivo expandir o conceito de auto-distribuição de componentes para explorar placement de serviços com estado da nuvem para a edge e vice-versa de forma transparente, podendo assim aproveitar as vantagens de ambos ambientes operacionais.

Trabalho: Individual ou Em Equipe (discriminar na descrição a divisão dos trabalhos)

Professor Responsável: [Prof. Dr. Luiz Fernando Bittencourt](#)

Título: Gerência de Sistemas Auto-adaptativos

Descrição do Projeto: A crescente complexidade na criação e gestão de sistemas distribuídos está centrada, principalmente, na volatilidade dos ambientes operacionais modernos. O ambiente volátil é caracterizado por constantes mudanças, muitas vezes inesperadas, que o sistema precisa lidar em tempo de execução para se manter funcional e atender seus requisitos não-funcionais. Para lidar com essas constantes mudanças, sistemas auto-adaptativos, capazes de se auto-adaptarem diante de mudanças, estão ganhando cada vez mais destaque. Este projeto tem como objetivo a exploração de algoritmos de aprendizado de máquina por reforço para aprender, em tempo de execução, a como adaptar e re-adaptar sistemas distribuídos em ambientes voláteis sem interferência humana, de forma a preservar sua funcionalidade ou otimizar aspectos não-funcionais do sistema.

Trabalho: Individual ou Em Equipe (discriminar na descrição a divisão dos trabalhos)

Professor Responsável: [Prof. Dr. Luiz Fernando Bittencourt](#)

Título: Aprendizado de Máquina Distribuído

Descrição do Projeto: Conjuntos de datasets são encontrados de diferentes variedades e gerados através de várias operações de upload ou offloading providas de dispositivos IoT ou de sistemas autônomos em aplicações Big Data streaming. O cálculo de algoritmos de Deep Learning (DL) computa um gradiente, chamado de Stochastic Gradient Descent (SGD). Este cálculo apresenta problemas diversos quando explorado de forma distribuída, dessa forma a aplicação de novas ferramentas para avaliar tais dados de forma distribuída é um grande desafio. Este projeto visa avaliar a performance de diferentes aplicações e algoritmos de DL, considerando acurácia, precisão, tempo de execução, número de épocas e quantidade de trocas de

mensagens, em um ambiente de Internet das Coisas.

Trabalho: Individual ou Em Equipe (discriminar na descrição a divisão dos trabalhos)

Professor Responsável: [Prof. Dr. Luiz Fernando Bittencourt](#)

Título: Gerência de recursos em sistemas distribuídos

Descrição do Projeto: A gerência de recursos envolve o processo de seleção dos recursos computacionais para execução de aplicações de diversos tipos. A otimização da alocação de recursos, como no escalonamento de tarefas, depende do desempenho que tais tarefas obtêm do recurso computacional em questão, seja este de processamento, armazenamento ou de rede. Este projeto tem como objetivo identificar uma ou mais aplicações e um ambiente de processamento distribuído para realização de uma análise de questões que concernem a gerência de recursos, tais como formas de implementação da aplicação, análise de desempenho e algoritmos de alocação de recursos.

Trabalho: Individual ou Em Equipe (discriminar na descrição a divisão dos trabalhos)

Professor Responsável: [Prof. Dr. Luiz Fernando Bittencourt](#)

Título: Monitoramento de colmeias com Internet das Coisas

Descrição do Projeto: Este trabalho é parte de um projeto de extensão coordenado pelo Prof. Roberto Greco do IG. O objetivo é projetar e implementar um sistema IoT que usa arduínos e/ou raspberries e sensores para monitoramento de colmeias de abelhas nativas (e.g. temperatura, umidade, som).

Trabalho: Individual ou Em Equipe (discriminar na descrição a divisão dos trabalhos)

Professor Responsável: [Prof. Dr. Luiz Fernando Bittencourt](#)

Título: Sistemas distribuídos para realidade virtual e aumentada

Descrição do Projeto: O(s) estudante(s) se aprofundarão em problemas relacionados à execução de aplicações e serviços distribuídos de realidade virtual e aumentada. Problemas de transmissão e qualidade de vídeo e processamento em tempo real para dispositivos de borda e vestíveis serão estudados para proposição de soluções viáveis e com desempenho adequado.

Trabalho: Individual ou Em Equipe (discriminar na descrição a divisão dos trabalhos)

Professor Responsável: [Prof. Dr. Luiz Fernando Bittencourt](#)

Título: Prevenção de Colisão e Controle de Rotas de Drones em Cenários com Comunicação 5G/6G

Descrição do Projeto: Cenários onde múltiplos drones co-existem demandam sistemas de prevenção de colisão e estabelecimento de rotas eficientes. O controle dos drones pode ser feito de forma cooperativa, com troca de informações entre eles, ou não-cooperativa, onde desvios e rotas são estabelecidos autonomamente por cada drone. Este tópico visa estudar e explorar estes cenários, através de simulações, onde a troca de dados pode ou não ocorrer através de pontos de acesso de rede celular 5G/6G.

Trabalho: Individual ou Em Equipe (discriminar na descrição a divisão dos trabalhos)

Professor Responsável: [Prof. Dr. Marcos Raimundo Medeiros](#)

Título: Criação de biblioteca em python para multi-task learning

Descrição do Projeto: Neste projeto a intenção seria a criação de uma biblioteca, similar ao MALSAR (<http://jiayuzhou.github.io/MALSAR/>) em python, e seria implementada usando uma biblioteca como o parsimony (<https://github.com/neurospin/pylearn-parsimony>) como base. Este trabalho tem uma forte conexão entre otimização e aprendizado de máquina, portanto um forte interesse em ambos é desejável.

Trabalho: Individual ou Em Equipe (discriminar na descrição a divisão dos trabalhos)

Professor Responsável: [Prof. Dr. Rodolfo Jardim de Azevedo](#)

Título: Sistematização e Publicação do Concept Inventory de Python

Descrição do Projeto: Um Concept Inventory - CI - é, de forma simplificada, um conjunto de questões de múltipla escolha sobre um determinado tema, no nosso caso aprendizagem de Python, que foi criado baseado nos erros comuns buscados em exercícios e atividades dos alunos que estão aprendendo a linguagem. A intenção ao catalogar os erros comuns é conseguir construir questões de múltipla escolha que abordem, como respostas erradas, os erros comuns e, assim, conseguir diagnosticar o caminho de pensamento do aluno para, posteriormente, melhor guiar seu aprendizado. Temos um CI para linguagem C já desenvolvido e um em Python com um bom avanço. Precisamos sistematizar o processo, potencialmente ampliar as questões e avançar na divulgação do mesmo documentando e criando os formulários online com as métricas de avaliação.

Temos projeto similar para outras linguagens e disciplinas.

Trabalho: Individual

Professor Responsável: [Prof. Dr. Rodolfo Jardim de Azevedo](#)

Título: Integração Contínua para validação de processadores RISC-V

Descrição do Projeto: A técnica de Continuous Integration - CI - é bastante utilizada no ambiente de software (levando até ao surgimento do termo devops) para garantir a qualidade do software, testabilidade e facilidade de execução dos próximos passos de desenvolvimento. Na área de hardware, apesar de poder ser utilizada com todas essas funcionalidades já descritas, pode também trazer uma nova oportunidade de cruzamento de testes visto que não é incomum que existam múltiplas implementações

de um mesmo hardware, como é o caso do processador RISC-V. Nesse caso, precisamos de um mecanismo para testar múltiplos processadores, com suas diferenças e grandes similaridades. Esse projeto vai focar em uma implementação de processador e integrar a um ambiente de desenvolvimento junto com testes extras necessários.

Aqui cabem múltiplos alunos em projetos individuais.

Trabalho: Individual

Professor Responsável: [Prof. Dr. Rodolfo Jardim de Azevedo](#)

Título: Ambiente remoto para execução de projetos de hardware

Descrição do Projeto: No ensino de projetos de hardware, em especial FPGAs, a disponibilidade do componente físico pode dificultar o desenvolvimento quando o aluno está em casa ou quando não há componentes suficientes para todos os alunos ou grupos. A intenção desse projeto é tirar proveito de programadores de código aberto para FPGA e montar um arcabouço para execução remota de FPGA.

Trabalho: Individual

Professor Responsável: [Prof. Dr. Zanoni Dias](#)

Título: Ordenação Semi-Completa por Rearranjo de Genomas

Descrição do Projeto: Problemas de Rearranjo de Genomas buscam estimar a distância evolutiva entre genomas de diferentes organismos. Estes problemas lidam com eventos de rearranjo, que são mutações capazes de alterar a sequência genética dos genomas. Quando assumimos que os genomas comparados não possuem genes repetidos, o problema corresponde a ordenação de uma permutação. O objetivo deste projeto consiste em desenvolver heurísticas que, dado um genoma de origem e um genoma alvo, sejam capazes de fornecer um genoma (obtido através do menor número de eventos de rearranjo de genomas) que esteja o mais próximo de genoma alvo possível. O problema ainda possui uma restrição sobre a proximidade máxima permitida para uma solução válida, que é estabelecida como entrada para o problema e considera as posições dos genes.

Trabalho: Individual

Professor Responsável: [Prof. Dr. Zanoni Dias](#)

Título: Problemas de Partição de Strings

Descrição do Projeto: Problemas de Partição de Strings estão intimamente ligados a problemas de Rearranjo de Genomas. Esses problemas consistem em segmentar duas strings, dadas como entrada do problema, de forma que as partes de uma das strings possam ser reorganizadas para obtermos a outra string. A forma como as strings podem ser segmentadas ou como as partes podem ser reorganizadas da origem formam diferentes variações do problema. Por exemplo, cada uma das partes pode ou

não ser invertida ao realizarmos a reorganização ou podemos permitir que algumas partes sejam deletadas ou inseridas antes da reorganização. O objetivo deste projeto é o desenvolvimento de algoritmos exatos para problemas de Partição de Strings. Existem vários caminhos possíveis que podem ser usados para tais algoritmos, dentre eles temos algoritmos parametrizados, programação linear inteira e decomposição de grafos em ciclos.

Trabalho: Individual

Professor Responsável: [Prof. Dr. Zanoni Dias](#)

Título: Análise Comparativa de Algoritmos para Problemas de Rearranjo de Genomas

Descrição do Projeto: Os problemas de Rearranjo de Genomas permitem a comparação de genomas de diferentes indivíduos e têm múltiplas aplicações na área da Biologia Computacional. Como esses problemas já vêm sendo estudados há muitos anos, várias versões foram propostas levando em conta as características da representação dos genomas utilizada e as operações de rearranjo consideradas. Além disso, muitos algoritmos para o problema podem ser decompostos em diferentes etapas, cada uma podendo ser resolvida por múltiplos algoritmos diferentes. O objetivo deste projeto é realizar uma comparação experimental de diferentes abordagens conhecidas para comparar genomas com base em problemas de rearranjo. As comparações poderão ser realizadas em bases de dados geradas artificialmente e com dados de genomas reais.

Trabalho: Individual

Professor Responsável: [Prof. Dr. Zanoni Dias](#)

Título: Análise de Classes Sub-Representadas em Bioinformática

Descrição do Projeto: Determinar as funções que cada proteína exerce requer grande esforço técnico e monetário. Com isso, o uso de aprendizado de máquina pode auxiliar na realização desta tarefa. Além do mais, algumas funções possuem poucas amostras, o que exige abordagens para lidar com este problema. O objetivo deste projeto é empregar técnicas de classificação de poucas amostras e aumento de dados no contexto de Processamento de Linguagem Natural (NLP) para classificar funções de proteínas pouco representadas.

Trabalho: Individual

Professor Responsável: [Prof. Dr. Zanoni Dias](#)

Título: Visualização e Interpretação de Modelos Profundos

Descrição do Projeto: Entender e significar o funcionamento de sistemas complexos, como algoritmos de aprendizado de máquina baseados em redes neurais artificiais profundas, são atividades empregadas nos estágios de avaliação, aplicação e sustentação de sistemas autônomos e necessárias para confirmação do funcionamento

correto dos mesmos. O objetivo deste projeto é estudar as diferentes técnicas de visualização e explicabilidade para modelos profundos existentes na literatura, focando-se em um ou mais algoritmos modal-genéricos, como: técnicas baseadas em gradiente, Class Activation Mapping (CAM) e Layer-wise Relevance Propagation (LRP).

Trabalho: Individual

Professor Responsável: [Prof. Dr. Zanoni Dias](#)

Título: Segmentação Semântica Fracamente Supervisionada da Amazônia

Descrição do Projeto: Neste projeto, serão exploradas técnicas de segmentação semântica fracamente supervisionada —, como técnicas de transferência de conhecimento prévio, técnicas de regularização sobre rótulos fracos e Few-Shot Segmentation — a fim de criar anotações de segmentação semântica (ao nível de píxel) sobre o conjunto de dados Planet: Understanding the Amazon from Space, a partir de somente rótulos de classificação ruidosos (ao nível de imagem).

Trabalho: Individual