

Professor Responsável	Título	Descrição dos Projetos Finais de Graduação para o 2s2021	Individual/Equipe
André Santanché	Infraestrutura de Web Components	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 evolução de um framework denominado Digital Content Components (vide referências: https://harena-lab.github.io/harena-docs/dccs/tutorial/ ; https://harena-lab.github.io/harena-docs/dccs/ e https://github.com/harena-lab/harena-space/tree/master/src/adonisjs/public/dccs). A evolução pode acontecer em uma das seguintes direções: aprimoramento de funcionalidade existentes; ampliação das funcionalidades da infraestrutura; criação de novos componentes para aplicações ou domínios específicos; criação de ferramenta visual para a composição de componentes.	Individual ou em Equipe
	Plataforma de Aprendizagem baseada em Casos	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 contribuições em uma das possíveis tarefas: ambiente de autoria de casos; ciência de dados para suporte à construção de casos; análise de evolução dos alunos, tutoria personalizada e inteligente.	Individual ou em Equipe
Eliane Martins	Uso de fuzzers nos testes de robustez e segurança de bibliotecas manipulação de imagens	Fuzz testing (ou fuzzing) é uma técnica de teste de software que insere dados aleatórios (válidos ou inválidos) no sistema de software para descobrir defeitos no tratamento de entradas, bem como brechas de segurança. Os dados são inseridos usando técnicas de teste automatizadas durante a execução, em que o sistema é monitorado para determinar se ocorreram exceções, queda (crash) ou bloqueio (hang) do sistema. Neste projeto serão estudadas técnicas de fuzzing para dados de entrada estruturados de entrada, como as imagens. O objetivo é avaliar dois tipos de técnicas de fuzzing, usando fuzzers disponíveis. O estudo envolve o aprendizado de formatos de representação de imagens, como JPEG e PNG, por exemplo. O projeto envolve também o aprendizado sobre fuzzing, e o uso de ferramentas de apoio. O projeto pode ser realizado por 2 alunos, cada qual responsável por aplicar uma técnica diferente de fuzzing.	Em Equipe
	Uso de testes baseados em modelos em combinação com o desenvolvimento dirigido pelo comportamento	O desenvolvimento dirigido pelo comportamento, mais conhecido como BDD (Behavior-Driven Development) é um paradigma usado em ambiente ágil que usa uma linguagem ubíqua, denominada Gherkin. A linguagem é ubíqua pois pode ser compreendida por diversos participantes: clientes, desenvolvedores, equipes de teste. A linguagem Gherkin é usada para descrever requisitos na forma de cenários, os quais são transformados em casos de teste de Aceitação do sistema. Estes cenários são criados manualmente pelas equipes de desenvolvimento ou de testes. Uma desvantagem é que a qualidade dos testes vai depender da experiência de quem os escreve. Uma outra desvantagem é que a equipe de testes não dispõe de uma métrica de cobertura que permita determinar se os testes são suficientes ou não. A combinação de BDD com testes baseados em modelos ou MBT (do inglês Model-Based Testing) pode ser uma solução para esses problemas. Em MBT, os casos de teste são produzidos a partir de um modelo representando o comportamento do sistema. O objetivo deste projeto é estabelecer um método que permita às equipes de teste, a partir dos requisitos modelados na forma de modelos de estados da UML, derivarem os casos de teste em Gherkin. No caso de mudança nos requisitos, por exemplo, com o acréscimo de um novo incremento, os casos de teste necessários são produzidos automaticamente a partir do modelo atualizado. Para que o método possa ser usado em ambientes de desenvolvimento ágil, o apoio de ferramentas é fundamental. Neste projeto o aluno deve entender, instalar e utilizar as ferramentas de apoio. Como estudo de caso, uma aplicação Web deve ser utilizada.	Individual
Flávio Keidi Miyazawa	Problemas de Corte e Empacotamento.	Os alunos deverão estudar algoritmos para problemas de empacotamento bidimensional em placas. Alguns alunos poderão investigar a versão de empacotamento em placas, outros em faixa, versões unidimensional, bidimensional, tridimensional ou multidimensional. Versões poderão ser de empacotamentos de itens retangulares, caixas ou mesmo irregulares. Direções de estudo poderá ser prática implementando e comparando os algoritmos mais promissores ou teórica, com estudo e desenvolvimento de algoritmos de aproximação, ou nas duas linhas. Os projetos poderão ser individuais ou em equipe.	Individual ou em Equipe
	Problemas de Roteamento de Veículos.	Os alunos deverão estudar algoritmos para problemas de roteamento de veículos. Cada veículo deve percorrer uma rota, partindo de um depósito, atender clientes e voltar novamente ao depósito. Cada cliente deve ser atendido por um veículo. O objetivo é minimizar uma função de custo associada às rotas. Cada projeto focará em uma versão do problema considerando características mais específicas. Os projetos poderão ser individuais ou em equipe.	Individual ou em Equipe
Hélio Pedrini	Reconhecimento de Ações em Vídeo.	Identificação de ações humanas em sequências de vídeos por meio de técnicas de processamento de imagens, análise de vídeos, visão computacional e aprendizado de máquina.	Individual
	Identificação de Eventos Anômalos em Vídeos de Vigilância.	Identificação de comportamento anômalo em vídeos de vigilância baseada em técnicas de visão computacional, análise de imagens e aprendizado de máquina.	Individual
	Análise de Imagens e Vídeos.	Proposição e aplicação de técnicas para processamento e análise de imagens e vídeos em diferentes domínios de conhecimento (sensoriamento remoto, medicina, biologia, biometria, microscopia, vigilância e segurança).	Individual
Hervé Cédric Yviquel	Biblioteca de álgebra linear para supercomputadores	O objetivo desse projeto é implementar uma biblioteca de álgebra linear para facilitar computação em matriz esparsa para supercomputadores (ou clusters). Essa biblioteca deve ser implementada em C/C++ em cima do runtime de tarefa distribuída do projeto OmpCluster usando o modelo de programação OpenMP. Se for em equipe, cada um pode implementar operações diferentes: por exemplo multiplicação matriz-vetor, multiplicação matriz-matriz, etc.	Individual ou em Equipe
	Aplicar o modelo de desempenho Roofline para aplicações científica baseadas em tarefa para supercomputadores	O objetivo deste projeto é melhorar a análise de desempenho de aplicação científica baseado em tarefas usando o modelo roofline. O modelo roofline foi desenvolvido pelo Berkeley Lab para visualizar facilmente os gargalos e potenciais de desempenho dos kernels de computação e pode ser facilmente usado usando ferramentas como Timemory. O estudante integrará a ferramenta Timemory no runtime de tarefa OmpCluster e explorará a melhor maneira de apresentar o perfilamento ao usuário.	Individual
Julio Cesar dos Reis	Interfaces vestíveis para a captura e rotulação de dados fisiológicos	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.	Individual
	Avaliação na plataforma OpenDesign	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.	Individual ou em Equipe
	Processamento de diálogos textuais na língua Portuguesa	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 prolongadas 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.	Individual
	Alinhamento de grafos de conhecimento	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.	Individual
	Descrição semântica de publicações científicas.	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.	Individual ou em Equipe
	Sistema para explorar dados interconectados abertos.	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 a fontes de dados na Web Semântica (SPARQL).	Individual
	Sistemas de questão e respostas usando bases RDF.	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 a fontes de dados na Web Semântica (linguagem SPARQL). As respostas obtidas serão convertidas em uma representação final para o usuário.	Individual
Visualização de ontologias.	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 alunos(a) melhor entendam os conceitos e suas relações nas disciplinas que compõem o curso.	Individual	
Luiz Fernando Bitencourt	Gerência de recursos em sistemas distribuídos	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.	Individual ou em Equipe

Professor Responsável	Título	Descrição dos Projetos Finais de Graduação para o 2s2021	Individual/Equipe
Nelson Luis Saldanha da Fonseca	Aprendizado Federativo para Classificação de Imagens e/ou Vídeos	Neste trabalho o aluno desenvolverá experimentos via emulação e/ou simulação de aprendizado federativo (do inglês Federated Learning) utilizando diferentes técnicas de particionamento de dados, explorando desde a visualização dos dados até a avaliação dos resultados dos algoritmos de aprendizado de máquina.	Individual
	Aprendizado Federativo para Medicina	Neste trabalho o aluno desenvolverá experimentos via emulação e/ou simulação de aprendizado federativo vertical (do inglês Vertical Federated Learning) utilizando dados de saúde, será trabalhado desde a visualização dos dados até a avaliação dos resultados dos algoritmos de aprendizado de máquina supervisionado.	Individual
	Avaliação de Desempenho de Aprendizado Federativo na Borda da Rede	Neste trabalho o aluno irá desenvolver experimentos via emulação ou simulação para avaliar diferentes aspectos relacionados à execução de aprendizado federativo (do inglês Federated Learning) em redes de computação de rede.	Individual
	Agregação Assíncrona Federativa de Modelos de Aprendizado Supervisionados	Neste trabalho o aluno desenvolverá um algoritmo de agregação assíncrona para aprendizado federativo (do inglês Federated Learning) que será integrado ao projeto FATE (Federated AI Technology Enabler). Além disso, serão desenvolvidos experimentos para comparar a agregação síncrona com a agregação assíncrona.	Individual
	Aprendizado Federativo Descentralizado	Neste trabalho o aluno desenvolverá algoritmos que permitam a execução de aprendizado federativo (do inglês Federated Learning) de forma distribuída, ou seja, sem um servidor central. Para validar os resultados, o aluno desenvolverá experimentos via emulação ou simulação.	Individual
	Avaliação de desempenho do HTTP 3.0	Neste trabalho, o aluno realizará experimentos via simulação, emulação ou modelagem analítica do desempenho da nova proposta de padronização do protocolo HTTP 3.0 e protocolos de transporte usados pelo HTTP 3.0.	Individual ou em Equipe
Zanoni Dias	Diagnóstico de COVID-19 usando Redes Neurais Profundas	A pandemia de COVID-19 impactou todas as esferas de nossa sociedade. A epidemia aumentou a pressão sobre os sistemas de saúde pública, incitando a comunidade científica a desenvolver e avaliar métodos para diagnosticar pacientes de forma confiável. Impulsionadas por sua eficácia na análise de imagens médicas, as redes neurais profundas têm sido vistas como uma alternativa promissora para diagnosticar automaticamente pacientes com COVID-19 a partir de radiografias de tórax. O objetivo deste trabalho é avançar nos resultados previamente obtidos no artigo "COVID-19 X-ray Image Diagnostic with Deep Neural Networks", vencedor do Best Paper Award do Brazilian Symposium on Bioinformatics (BSB 2020).	Individual
	Classificação de Funções de Proteína	O objetivo deste projeto é empregar técnicas de aprendizado de máquina para classificar funções de proteínas, com base nas suas sequências de aminoácidos. Considerando que proteínas podem ser interpretadas como (longas) frases, onde cada aminoácido pode ser tratado como sendo uma palavra, técnicas modernas de Processamento de Linguagem Natural (NLP) podem ser utilizadas para esta tarefa.	Individual
	Bioinformática para Análise de Transcriptomas	Transcriptoma refere-se ao conjunto completo de transcritos (RNAs mensageiros, RNAs ribossômicos, RNAs transportadores e os microRNAs) de um dado organismo e é o reflexo direto da expressão dos genes. Uma das técnicas mais populares de geração de dados em genômica é o sequenciamento de RNA (RNAseq), devido ao seu relativo baixo custo. O objetivo deste projeto é desenvolver metodologias para associar os RNA sequenciados aos genes do organismo.	Individual
	Rearranjo de Genomas com Genes Multiplicados	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 os genomas apresentam genes multiplicados, eles são representados por strings. O objetivo deste projeto é o desenvolvimento de heurísticas para problemas de Rearranjo de Genomas, considerando genomas com genes repetidos.	Individual
	Rearranjo de Genomas Considerando Regiões Intergênicas	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. Modelos tradicionais representam os genomas considerando apenas os seus genes, desprezando outras informações importantes. Trabalhos recentes mostram que considerar as regiões intergênicas ajudam a produzir modelos evolutivos mais precisos. O objetivo deste projeto é o desenvolvimento de heurísticas para problemas de Rearranjo de Genomas, considerando genomas representados pelos seus genes e pelas suas regiões intergênicas.	Individual
	Ordenação de Permutações por Operações Determinísticas	Problemas de Rearranjo de Genomas podem ser tratados como problemas de ordenação de permutações. Tipicamente, eventos de rearranjo podem movimentar elementos para qualquer posição da permutação. Neste projeto pretende-se explorar variações onde as operações são determinísticas, ou seja, os elementos devem ser movidos para posições específicas da permutação. Neste projeto serão exploradas diversas operações determinísticas, como, por exemplo, inserções, reversões e transposições.	Individual