



SERVIÇO PÚBLICO FEDERAL

MINISTÉRIO DA EDUCAÇÃO

UNIVERSIDADE FEDERAL DE SANTA CATARINA

PRÓ-REITORIA DE PÓS-GRADUAÇÃO

CAMPUS UNIVERSITÁRIO REITOR JOÃO DAVID FERREIRA LIMA - TRINDADE

CEP: 88040-900 - FLORIANÓPOLIS - SC

TELEFONE: (48) 3721-9438

E-MAIL: cgprint.ufsc@gmail.com

EDITAL CONJUNTO CG-PRINT/UFSC nº. 01/2018/PROPG/PROPESQ

Programa Institucional de Internacionalização – CAPES-PRINT/UFSC

ANEXO A

FORMULÁRIO DE INSCRIÇÃO

DADOS DO COORDENADOR DO PROJETO

Nome Completo:	Vania Bogorny
SIAPE:	1711009
Departamento Vinculado:	Departamento de Informática e Estatística
PPG Credenciado:	Programa de Pós-Graduação em Ciência da Computação
Link CV Lattes:	http://lattes.cnpq.br/9045199822095908
ORCID:	http://orcid.org/0000-0002-0159-4643
Titulação:	Doutorado
Ano de titulação:	2006 (com período Sanduíche na Universidade de Minnesota, EUA)
Área de Conhecimento:	Ciência da Computação
País da Instituição:	Brasil
Instituição:	Universidade Federal do Rio Grande do Sul (UFRGS)

DADOS DOS PARTICIPANTES NACIONAIS DO PROJETO**PARTICIPANTE (1)**

Nome Completo:	Ronaldo dos Santos Mello
SIAPE:	1160644
Departamento Vinculado:	Departamento de Informática e Estatística
PPG Credenciado:	Programa de Pós-Graduação em Ciência da Computação
Link CV Lattes	http://lattes.cnpq.br/5011370918857999
Titulação:	Doutorado
Ano de titulação:	2002 - Universidade Federal do Rio Grande do Sul
ORCID:	https://orcid.org/0000-0003-4262-5474

PARTICIPANTE (2)

Nome Completo:	Marcio Bastos Castro
SIAPE:	20151040
Departamento Vinculado:	Departamento de Informática e Estatística
PPG Credenciado:	Programa de Pós-Graduação em Ciência da Computação
Link CV Lattes	http://lattes.cnpq.br/6876016315737507
Titulação:	Doutorado
Ano de titulação:	2012 - Université de Grenoble
ORCID:	https://orcid.org/0000-0002-9992-8540

PARTICIPANTE (3)

Nome Completo:	Edmundo Carlos Grisard
SIAPE:	1160089
Departamento Vinculado:	Departamento de Microbiologia, Imunologia e Parasitologia (MIP/CCB/UFSC)
PPG Credenciado:	PPG em Biotecnologia e Biociências
Link CV Lattes	http://lattes.cnpq.br/1067410443144841
Titulação:	Doutor em Parasitologia
Ano de titulação:	1999
ORCID:	https://orcid.org/0000-0001-8916-8296

PARTICIPANTE (4)

Nome Completo:	Glauber Wagner
SIAPE:	225802
Departamento Vinculado:	Departamento de Microbiologia, Imunologia e Parasitologia (MIP/CCB/UFSC)
PPG Credenciado:	PPG em Biotecnologia e Biociências
Link CV Lattes	http://lattes.cnpq.br/8417542717418294
Titulação:	Doutor em Biotecnologia e Biociências
Ano de titulação:	2012

ORCID:	https://orcid.org/0000-0001-5003-6595
---------------	---

PARTICIPANTE (5)

Nome Completo:	Rubens Tadeu Delgado Duarte
SIAPE:	2193320
Departamento Vinculado:	Departamento de Microbiologia, Imunologia e Parasitologia (MIP/CCB/UFSC)
PPG Credenciado:	PPG em Biotecnologia e Biociências
Link CV Lattes	http://lattes.cnpq.br/4626372929936792
Titulação:	Doutor em Biotecnologia
Ano de titulação:	2010
ORCID:	https://orcid.org/0000-0002-9939-3400

PARTICIPANTE (6)

Nome Completo:	Patrícia Hermes Stoco
SIAPE:	1136040
Departamento Vinculado:	Departamento de Microbiologia, Imunologia e Parasitologia (MIP/CCB/UFSC)
PPG Credenciado:	PPG em Biotecnologia e Biociências
Link CV Lattes	http://lattes.cnpq.br/8555938548955913
Titulação:	Doutora em Biotecnologia
Ano de titulação:	2010
ORCID:	https://orcid.org/0000-0003-0879-6173

DADOS DOS PARTICIPANTES ESTRANGEIROS DO PROJETO

PARTICIPANTE (1)

Nome Completo:	Stan Matwin
Instituição Vinculada:	Dalhousie University (Canadá)
Departamento Vinculado:	Faculty of Computer Science
ORCID:	https://orcid.org/0000-0001-6629-8434
CV resumido	https://web.cs.dal.ca/~stan/bio/
Titulação	PhD - Warsaw University
Ano de titulação	1977
Área de titulação	Computer Science

PARTICIPANTE (7)

Nome Completo:	Chiara Renso
Instituição Vinculada:	Consiglio Nazione delle Ricerche (CNR), Italy
Departamento Vinculado:	Istituto di Scienza e Tecnologie dell'Informazione A. Faedo
ORCID:	https://orcid.org/0000-0002-1763-2966
CV resumido	http://hpc.isti.cnr.it/~renso/curriculumEnglish.pdf
Titulação	Ph.D - University of Pisa

Ano de titulação	1998
Área de titulação	Computer Science

PARTICIPANTE (8)

Nome Completo:	Hércules Moura
Instituição Vinculada:	Centers for Disease Control and Prevention (CDC / USA)
Departamento Vinculado:	National Center for Environmental Health (NCCE/ CDC)
ORCID:	https://orcid.org/0000-0001-7672-0942
CV resumido	<i>Informar link, ou anexar o arquivo à proposta</i>
Titulação	Doutorado - Universidade Federal do Rio de Janeiro
Ano de titulação	1991
Área de titulação	Infectious and Parasitic Diseases

PARTICIPANTE (9)

Nome Completo:	Björn Andersson
Instituição Vinculada:	Karolinska Institutet: Stockholm, Sweden
Departamento Vinculado:	Department of Cell and Molecular Biology
ORCID:	https://orcid.org/0000-0002-4624-0259
CV resumido	<i>Informar link, ou anexar o arquivo à proposta</i>
Titulação	Ph.D.
Ano de titulação	1992
Área de titulação	Molecular Biology

PARTICIPANTE (10)

Nome Completo:	Jorge Luiz Mazza Rodrigues
Instituição Vinculada:	University of California - Davis (UC Davis)
Departamento Vinculado:	Department of Land, Air, and Water Resources
ORCID:	https://orcid.org/0000-0002-6446-6462
CV resumido	<i>Informar link, ou anexar o arquivo à proposta</i>
Titulação	PhD - Michigan State University
Ano de titulação	2000
Área de titulação	Crop and Soil Sciences and Environmental Toxicology

PARTICIPANTE (11)

Nome Completo:	Konstantinos Tserpes
Instituição Vinculada:	Harokopio University of Athens, Grécia
Departamento Vinculado:	Department of Informatics & Telematics
ORCID:	https://orcid.org/0000-0001-5183-1443
CV resumido	https://www.dit.hua.gr/~tserpes/CV_eng_tserpes.pdf
Titulação	PhD (Department of Electrical and Computer Engineering, National Technical University of Athens)
Ano de titulação	2008
Área de titulação	Computer Science

PARTICIPANTE (12)

Nome Completo:	Jean-François Méhaut
Instituição Vinculada:	Université Grenoble Alpes (UGA)
Departamento Vinculado:	Laboratoire d'Informatique de Grenoble (LIG)
ORCID:	https://orcid.org/0000-0003-1047-7462
CV resumido	
Titulação	PhD - Université des Sciences et Technologies de Lille
Ano de titulação	1989
Área de titulação	Computer Science

PARTICIPANTE (13)

Nome Completo:	François Broquedis
Instituição Vinculada:	Université Grenoble Alpes (UGA)
Departamento Vinculado:	Laboratoire d'Informatique de Grenoble (LIG)
ORCID:	
CV resumido	
Titulação	PhD - Université de Bordeaux, France
Ano de titulação	2010
Área de titulação	Computer Science

PERFIL DOS PARTICIPANTES

Capacidade Técnica e Liderança da Equipe Brasileira (2000 caracteres)

Descrever a capacidade técnica e liderança dos membros do projeto.

PPG Ciência da Computação

Os docentes do PPGCC Vania Bogorny (Pesquisadora 2 CNPq) e Ronaldo dos Santos Mello têm experiência em armazenamento e busca eficiente de grandes volumes de dados em bancos NoSQL, específicos para tratamento de grandes volumes de dados, bem como na extração de conhecimento e criação de técnicas de análise de dados sequenciais e de mobilidade, que podem ser diretamente aplicadas nas áreas de Física e Biologia. Ambos os docentes foram contemplados com o financiamento de uma missão de curta duração da União Européia nas chamadas 2017 e 2018, para atuar no projeto europeu SOBIGDATA (<http://www.sobigdata.eu/>), tendo sido esta uma chamada internacional. A professora Vania Bogorny é subcoordenadora do Programa de Pós-Graduação em Ciência da Computação desde maio/2017, tem atualmente o segundo maior H-Index do corpo docente do PPGCC, já coordenou pela UFSC um projeto europeu FP7 - Marie Curie e atualmente coordena um projeto europeu H2020.

O professor Ronaldo dos Santos Mello foi coordenador do PPGCC de 2012 a 2015, é especialista em desenvolvimento de estratégias para modelagem, armazenamento, consulta e integração de Big Data em tecnologias emergentes de bancos de dados, como sistemas NOSQL, NEWSQL e em memória.

O professor Marcio Bastos Castro (Pesquisador 2 CNPQ) é especialista em técnicas de processamento de alto desempenho, visando tornar o processamento e análise de grandes volumes de dados escalável, paralelo e distribuído.

PPG Biotecnologia e Biociências

Os docentes do PPGBTC, integrantes deste projeto, são um professor com experiência internacional e bolsista CNPQ, e por três jovens professores, com menos de 3 anos de UFSC, mas já com experiência na orientação concluída de alunos de mestrado, e com colaborações internacionais pontuais que serão consolidadas e ampliadas no âmbito deste projeto.

O professor Edmundo Carlos Grisard (Pesquisador 1C CNPq) têm experiência na área de genômica comparativa. Foi integrante do Projeto Genoma Nacional, com participação em projetos como a Rede Sul de Análise de Genomas e Biologia Estrutural / Programa de Investigação de Genomas Sul (PIGS) - Projeto de seqüenciamento do genoma da Mycoplasma hyopneumoniae, Projeto Genoma EST do camarão Litopenaeus vannamei - ShEST, Projeto Transcriptoma do Trypanosoma rangeli, Rede de Proteoma do Estado de Santa Catarina - RPSC, entre outros.

Os professores Glauber Wagner e Patrícia Hermes Stoco tem experiência na geração e análise de dados genômicos em larga escala, desde a geração de seqüências de DNA, a partir de sequenciamento utilizando métodos de Sanger e NGS (como 454, Illumina e PacBio), comprovada pela publicação de diversos genomas, de genômica comparativa e filogenia. Da mesma forma, estes pesquisadores têm experiência na geração, armazenamento e interpretação de dados proteômicos, em especial em espectrometria de massas, comprovada pela publicação do proteoma de espécies de protozoários e bactérias. Além disso, o professor Glauber tem experiência no desenvolvimento de sistemas de anotação de genomas e workflow científicos voltados para a análise de dados biológicos, como por exemplo o software STINGRAY (<http://stingray.biowebdb.org>).

O professor Rubens T. Duarte tem experiência na geração e análise de dados de metagenomas ambientais, com ênfase em diversidade microbiana em ambientes extremos, com foco na identificação de micro-organismos através de sequenciamento e análise de dados de NGS.

PPG Física

A equipe do PPGFSC, com especialidade em Astrofísica, Física Nuclear e de Partículas Elementares, se destaca como liderança internacional no desenvolvimento e aplicação de métodos computacionais a grandes bases de dados. O grupo se destaca não apenas pela exploração científica desses dados, publicada em dezenas de artigos que somam milhares de citações, mas também pela disponibilização pública de

ferramentas e resultados científicos em bancos de dados abertos, com destaque para www.starlight.ufsc.br, www.pycasso.ufsc.br, e www.vvvsurvey.org, utilizados por estudantes e pesquisadores em todo mundo. Cabe ressaltar a comprovada capacidade da equipe em formar novos recursos humanos para pesquisa, tendo já formado 19 doutores e 28 mestres, além de supervisionar vários pós-docs.

O professor Cid Fernandes (Pesquisador 1B CNPq, índice h 44) é líder na análise espectral de populações estelares em galáxias aplicada a mega-surveys como o Sloan Digital Sky Survey, que produziu mais de um milhão de espectros de galáxias (www.starlight.ufsc.br), e o CALIFA survey, que produziu número similar de espectros espacialmente resolvidos de galáxias (www.pycasso.ufsc.br). Coordenou o PPGFSC de 2008 a 2011.

A professora Natalia Vale Asari (Pesquisadora 2 CNPq) participa de estudos de populações de galáxias, combinados com observações pancromáticas (por exemplo, rádio galáxias), e vem se especializando em unir esses estudos à análise de linhas de emissão do gás em galáxias (www.bond.ufsc.br).

O professor Kanaan busca anãs-brancas, que marcam o fim da vida da maioria das estrelas, em grandes bases de dados. Ele também desenvolve técnicas para robotização de telescópios, e é responsável pelo software Chimera, que controla a fila do telescópio T80 no Chile, do consórcio hispânico-brasileiro J-PAS.

O professor Saito (Pesquisador 2 CNPq) participa dos projetos internacionais VVV, VVV-X e S-PLUS, que monitoram cerca de um bilhão de estrelas de nossa Galáxia, sendo desde 2010 um dos responsáveis pelo planejamento das observações do VVV e VVV-X. Sua área de atuação inclui estrutura da Via Láctea, estrelas variáveis, aglomerados estelares e exoplanetas.

O professor Emmanuel de Oliveira (Pesquisador 2 CNPq) trabalha com análise fenomenológica dos dados gerados por grandes colaborações de altas energias. Por exemplo, o LHC produz resultados para as probabilidades de produção das mais diversas partículas, mas todos estes resultados devem ser explicados pela mesma teoria, no caso, o Modelo Padrão de Partículas.

A professora Débora P Menezes (Pesquisadora 1B CNPq) é uma das líderes internacionais na investigação dos constituintes internos de objetos compactos (estrelas de nêutrons e magnetares) realizada a partir das forças nucleares forte e fraca e de transições de fase da cromodinâmica quântica.

Experiência Internacional da Equipe Brasileira (2000 caracteres)

Descrever a experiência internacional acadêmica dos membros do projeto.

A equipe brasileira tem ampla experiência internacional.

PPG Ciência da Computação

Na equipe do PPGCC, a Professora Vania Bogorny atuou no projeto europeu GEOPKDD em 2007 com bolsa da União Européia, quando iniciou suas cooperações com a Itália e Grécia, e conheceu o pesquisador Stan Matwin do Canadá como revisor e avaliador externo do Projeto europeu. Em 2011 foi coordenadora na UFSC do projeto internacional SEEK (FP7 Marie Curie Project project N. 295179, EU PEOPLE program <http://cordis.europa.eu/fp7/people/> - IRSES 2011 http://cordis.europa.eu/fp7/mariecurieactions/irses_en.html), totalmente financiado pela União Européia, e atualmente é coordenadora na UFSC do projeto europeu MASTER (H2020-MSCA-RISE-2017 - Project n° 777695), também financiado pela União Européia. Em 2012 foi participante de colaboração Bilateral Brasil/Itália em conjunto com a UFC. Em 2017 foi contemplada com uma bolsa de curta duração da União Européia para participar do projeto europeu SOBIGDATA (<http://www.sobigdata.eu/>) no CNR/Pisa, e foi professor visitante no CNR/Pisa/Itália em fevereiro/2013, na Universidade Ca' Foscari de Venezia/Itália em julho/2013, e na Universidade de Piraeus/Grécia em julho/2015, todos financiados pela União Européia. A docente revisa artigos para mais de 20 periódicos internacionais e avalia projetos para órgãos de fomento internacionais. Tem trabalhos em co-autoria com mais de 30 pesquisadores estrangeiros. O professor Marcio Bastos Castro realiza pesquisa na área de computação paralela e distribuída e computação de alto desempenho desde 2004. Coordenou recentemente um projeto de cooperação internacional apoiado financeiramente através do programa de cooperação internacional SITC-AmSud (projeto EnergySFE - <http://energysfe.ufsc.br>), onde participaram pesquisadores da França, Brasil e Equador. Tem grande interesse em aplicar técnicas de computação paralela e distribuída na área de big data, com intuito de melhorar o desempenho dos algoritmos utilizados para o processamento de grandes volumes de dados.

PPG Biotecnologia e Biociências

A equipe da biotecnologia tem realizado pesquisas com colaboradores internacionais comprovadas através dos artigos publicados pelos seus grupos e comprovadas pelas suas publicações. Destaca-se que o professor Edmundo C. Grisar, coordenou projetos com colaborações internacionais como "Intercâmbio científico Colombia-Brasil para o estudo de de Trypanosoma rangeli" e Marcadores genéticos e moleculares de virulência em tripanosomatídeos patogênicos e não patogênicos (financiado pelo PROGRAMA CAPES/STINT (2013-2014), com colaboração entre pesquisadores Brasil - Suécia. Além disso, destaca-se que a experiência na elaboração de cursos internacionais, com destaque para o Curso Internacional de Proteômica de Micro-organismos ministrado pelo Prof. Glauber, que já está na sua terceira edição. O Prof. Rubens T. D. Duarte mantém colaboração internacional com a Dra. Elizaveta Rivkina (Institute of Physicochemical and Biological Problems of Soil - Russian Academy of Sciences, Russia) na investigação de microrganismos em solos permafrost do Ártico e da Antártica, e com o Dr. Sergei Bulat (Petersburg Nuclear Physics Institute, Gatchina, Russia) no estudo de microrganismos de geleiras polares como indicadores de mudanças climáticas. Ainda no âmbito internacional, o Prof. Rubens Duarte mantém colaboração com o Dr. Ivan G. Paulino-Lima (Universities Space Research Association - USRA, NASA, EUA) na caracterização genômica de bactérias isoladas do Deserto do Atacama. Ressalta-se ainda que os professores da equipe do PPGBTC realizaram doutorado sanduíche em instituições nos EUA e Reino Unido, e que mantêm com os pesquisadores que os receberam intensa produção científica e técnica. Assim, o projeto PRINT tem um grande potencial para oficializar e estabelecer cooperações internacionais duradouras e importantes para os docentes deste programa, em especial para grupos de pesquisa novos ou em consolidação, que é o caso dos docentes participantes deste projeto..

PPG Física

A equipe brasileira do PPGFSC tem bom histórico de colaborações internacionais. Dentre os seis professores envolvidos, três realizaram seus doutoramentos integralmente no exterior (Cambridge, Austin e Oxford) e outro realizou seu doutorado em cotutela (Paris). Todos realizaram pós-doutorados no exterior, e participaram de dezenas de missões de trabalho no âmbito de convênios como CNPq-CSIC, Capes-FCT, CAPES-COFECUB, e outros, incluindo vários projetos financiados exclusivamente por instituições estrangeiras. Doutorandos e pós-docs da equipe também possuem experiência internacional na forma de estágios sanduíche e visitas de trabalho a instituições estrangeiras.

Para dar um único exemplo do PPGFSC: a professora Natalia Vale Asari coordenou um projeto PVE/CAPES da Dra. Grazyna Stasinska, do Observatoire de Paris. Durante a vigência desse projeto, foram publicados 8 artigos em revistas indexadas Qualis A1 e A2, e a PVE co-organizou uma escola avançada internacional no nível de pós-graduação no Brasil (NEBULATOM 3, <http://eventos.redclara.net/e/nebulatom3>). Esta foi a primeira edição da escola no Brasil, dedicada ao ensino de todos os aspectos sobre linhas de emissão a jovens pesquisadores, com palestrantes experts em nebulosas e física atômica. Um dos professores foi co-chair do comitê científico da conferência The cosmic feast of the elements (<https://sites.google.com/a/astro.unam.mx/grazyna2017/>). Outro docente participa da organização da XIV Hadron Physics 2018 (<http://hadrons18.org>). A equipe organizou também em Florianópolis um encontro do projeto JPAS, uma colaboração hispano-brasileira, e um encontro Scipy Latin America (<http://2016.pythonbrasil.org.br/>). Atualmente organiza o 9th VVV Science Meeting, que acontecerá em Florianópolis em abril (<tinyurl.com/9thvvv>).

Experiência da Equipe Estrangeira (2000 caracteres)

Ciência da Computação

Stan Matwin: Diretor do [Institute for Big Data Analytics](#) da Dalhousie University no Canadá e Emeritus Distinguished Professor em Ciência da Computação na University of Ottawa. É avaliador de projetos europeus dos programas FP7 e H2020, membro de comitês de programas dos eventos mais importantes em Inteligência Artificial, Aprendizado de Máquina e Mineração de Dados. Foi coordenador de diversos projetos internacionais, sendo co-autor de mais de 250 artigos científicos. É especialista em Aprendizado de Máquina e Inteligência artificial, que são áreas da Ciência da Computação fundamentais para análise de Big Data e Data Science. Seu H-index é atualmente 35, e seus artigos tem mais de 8000 citações no Google Scholar.

Jean-François Méhaut: Professor Titular da UGA desde 2003 com vasta experiência na área de computação de alto desempenho. Seu grupo possui mais de 300 artigos científicos com mais de 9 mil citações no google scholar, sendo mais de 3000 no scopus. Orientou mais de 34 alunos de doutorado, foi professor visitante no

Argonne National Laboratory of Department of Energy (DOE, Illinois, USA). Participou de diversos projetos internacionais (ex. EU FP7 ICT Mont-Blanc, FP7 IRSES HPC GA) e é membro do Joint Laboratory for Petascale Computing (JLESC) entre INRIA, University of Illinois at Urbana Champaign (UIUC), Argonne National Laboratory (ANL/MCS) e BSC (Barcelona). Possui também colaborações com grupos industriais como ST Microelectronics, Bull SA, Aselta.

François Broquedis: Professor Associado do Grenoble Institute of Technology. Possui grande experiência em computação de alto desempenho, através de projetos que coordena com a Intel. O foco da sua pesquisa é nos ambientes de programação paralela para otimizar o desempenho de aplicações de alto desempenho em arquiteturas de larga escala.

Chiara Renso: Pesquisadora do CNR/Pisa desde 2000 e membro do Laboratório de Computação de Alto Desempenho (HPC Lab), coordenado por Raffaele Perego (H-index 32). Os membros do HPC Lab somam mais de 200 publicações em eventos e revistas, tendo mais de 11 mil citações. Chiara Renso participa de diversos projetos europeus e têm experiência em coordenação de projetos internacionais FP7 e H2020. Tem experiência em análise semântica de dados, extração de conhecimento de grandes volumes de dados de redes sociais e dados de movimento. Atualmente é a coordenadora geral do projeto MATER (H2020), entre Canadá, Brasil, Itália, Grécia e França..

Konstantinos Tserpes: Professor do Department of Informatics and Telematics na Harokopio University of Athens, é especialista em big data analytics, sistemas distribuídos e redes sociais, pesquisando temas relacionados a escalabilidade, interoperabilidade e tolerância a falhas em vários domínios de aplicação. Participou de vários projetos europeus como +Spaces, SocloS, Consensus, Fortissimo Experiment (FP7), BASMATI e MASTER (H2020), sendo deste último o coordenador local pela sua universidade.

Biotecnologia e Biociências

Hércules Moura: Pesquisador do Centers for Disease Control and Prevention (CDC/USA), possui vasta experiência na área de proteômica, em especial na geração de dados em espectrometria de massas de micro-organismos. Tem como foco em sua pesquisa o uso da espectrometria de massas, em especial a proteômica, para a identificação de alvos para diagnóstico de diferentes patógenos humanos. O grupo do Prof. Moura tem mais 100 artigos publicados na área de proteômica. O Professor Moura tem interesse em análises de Big Data em função da grande quantidade de dados que são gerados nos equipamentos atuais.

Björn Andersson: Professor do Karolinska Institutet da Suécia, possui larga experiência em análise de dados de NGS, com foco em genômica comparativa de protozoários patogênicos, com mais de 90 artigos publicados na área de genômica (H-index 38), com mais de 9,000 citações. O Prof. Andersson tem colaborado com os professores Glauber Wagner e Patrícia Hermes Stoco na geração e análises de genomas de tripanosomatídeos, como do *Trypanosoma rangeli* e *Trypanosoma cruzi*.

Jorge Rodrigues: Professor Associado do Department of Land, Air and Water Resources, University of California at Davis (UC Davis), tem experiência na área de Microbiologia do Solo e Ecologia Microbiana, com ênfase na aplicação de metagenômica funcional para avaliação de impacto em áreas ambientais. O grupo do Professor Rodrigues publicou mais de 40 artigos na área de genômica do solo (H-index 20), com mais de 1,800 citações.

PROGRAMAS E INSTITUIÇÕES DO PROJETO DE PESQUISA EM COOPERAÇÃO INTERNACIONAL

Programas de Pós-Graduação da UFSC

Listar ao menos 3 PPGs participantes, incluindo o Programa líder e os colaboradores, e as respectivas áreas de avaliação da CAPES.

Programa líder:

Programa de Pós-Graduação em Ciência da Computação (PPGCC)

Programas colaboradores:

Programa de Pós-Graduação em Física (PPGFSC)

Programa de Pós-Graduação em Biotecnologia e Biociências (PPGBTC)

Instituições Parceiras Estrangeiras

Listar ao menos uma instituição participante, preferencialmente de um dos seguintes países: África do Sul, Alemanha, Argentina, Austrália, Áustria, Canadá, China, Coréia do Sul, Dinamarca, Espanha, Estados Unidos da América, Finlândia, França, Índia, Irlanda, Itália, Japão, México, Noruega, Nova Zelândia, Países Baixos, Reino Unido, Rússia, Suécia, Suíça.

Dentre os países prioritários:

Dalhousie University, Canadá

Université de Grenoble Alpes (UGA), França

Grenoble Institute of Technology (Grenoble INP), França

Centers for Disease Control and Prevention (CDC), Estados Unidos da América

Karolinska Institutet, Suécia

Instituto de Astrofísica de Andalucía, Consejo Superior de Investigaciones Científicas (IAA-CSIC), Espanha

University of St Andrews, Reino Unido

Observatoire de Paris, França

Université de Caen-Basse Normandie, França

Universidad Nacional de La Plata, Argentina

Universidade de Durham, Reino Unido

CNR, Pisa, Itália

Em outros países:

Harokopio University, Atenas, Grécia

DADOS BÁSICOS DO PROJETO DE PESQUISA EM COOPERAÇÃO INTERNACIONAL

Título do Projeto

Big Data Analytics: lançando luz dos genes ao cosmos

Tema Prioritário

Assinalar apenas um tema principal:

- Internacionalização do ensino superior, incluindo linguagem, cultura e identidade;
- Sustentabilidade ambiental, incluindo gestão de recursos hídricos e de resíduos, fontes renováveis de energia, e conservação e uso da biodiversidade;
- Indústria 4.0, incluindo automação, internet das coisas, e computação em nuvem;
- Biotecnologia e biomedicina, incluindo saúde humana;
- Nanociência e nanotecnologia, incluindo materiais, fármacos, compósitos, sensores e dispositivos.

Palavras-chave

Indicar até 3 palavras-chave

Big Data Analytics, Data Science, Data Management.

Data de Início e Término do Projeto

Considerar 4 anos, a iniciar em agosto/2018

O período do projeto está previsto para 4 anos, iniciando em Agosto/2018 até Julho/2022.

Área de Conhecimento

Usar a tabela da CAPES para definir uma área principal:

<http://www.capes.gov.br/avaliacao/instrumentos-de-apoio/tabela-de-areas-do-conhecimento-avaliacao>

Ciência da Computação.

DESCRIÇÃO DO PROJETO

Resumo do Projeto (2000 caracteres)

Várias áreas da ciência sofreram uma grande revolução nesse milênio. Estudos que eram feitos com uma dúzia de dados passaram a contar com milhares ou milhões de dados. Os atos de lidar com essa enorme quantidade de dados heterogêneos – ou Big Data – e passar da análise artesanal para entender a ciência por trás desses dados – data science – incorporaram-se ao dia-a-dia de muitos cientistas, em especial físicos e biólogos.

Na astrofísica, desde 2000 grandes levantamentos de imagens e espectros de galáxias e estrelas começaram a tornar seus dados públicos (e.g. 2dF, SDSS, VVV). Na frente de física de partículas, o lado experimental é dominado pelas grandes colaborações, tais como os experimentos ATLAS, CMS, LHCb e ALICE do grande colisor de hádrons (LHC, Large Hadron Collider). Os físicos teóricos da área têm então uma multitude de dados experimentais a serem analisados, utilizando-se das técnicas computacionais mais avançadas. Outro aspecto importante no entendimento do universo no qual vivemos envolve tanto dados observacionais obtidos por grandes telescópios ao redor do mundo e no espaço, quanto experimentais, advindos dos grandes colisores de partículas. Essas informações complementares são utilizadas de forma conjunta ou separada em modelos relativísticos e para corroboração de teorias já propostas. Na área de biotecnologia, o grande volume de dados vem do sequenciamento do DNA de diferentes organismos, especialmente com o advento dos sequenciadores de nova geração (NGS), bem como de análise de proteínas por espectrometria de massas, onde são gerados milhares de espectros de massas por segundo.

Físicos e biólogos tiveram que acrescentar aos seus vocabulários termos caros à Ciência da Computação, como *data management* para organizar dados; técnicas como *PCA*, *machine learning*, *data mining*, *pattern recognition* e *clustering* para analisar dados; computação distribuída e de alto desempenho para lidar com o volume de dados; e até inventaram o termo *citizen science*, que consiste em recrutar o público para ajudar a classificar desde galáxias até o canto de baleias (www.zooniverse.org).

Essa superposição de técnicas nos impeliu a forjar uma colaboração de pesquisadores da UFSC que trabalham com Big Data e todos os termos computacionais a ele associados. Com este projeto, visamos a ampliar a inserção internacional dos Programas de Pós-Graduação em Ciência da Computação, Física e Biotecnologia e Biociências da UFSC. Procuramos consolidar parcerias de pesquisa já existentes e buscar novas cooperações, através de intercâmbio de pesquisadores e estudantes, e de um simpósio interdisciplinar em Big Data.

Problema (2000 caracteres)

Definir o problema científico ou tecnológico.

Os três Programas de Pós-Graduação da UFSC envolvidos neste projeto atuam na área de análise de grandes volumes de dados.

Três grandes áreas da Física que lidam com um volume imenso e heterogêneo de dados são a astrofísica, a física nuclear de altas energias e a física de partículas elementares. No panorama atual da pesquisa em astrofísica destaca-se a proliferação de grandes levantamentos de dados como instrumentos para explorar o cosmos, desde planetas e estrelas (Saito et al. 2012) até galáxias e o Universo como um todo. A abundância e a riqueza dos dados hoje já disponíveis, e também os que virão nos próximos anos, requer ferramentas de análise para coletar dados (Mello et al. 2014) e para transformar os dados coletados em quantidades de interesse físico e astrofísico como, por exemplo a idade, massa, e abundância química das estrelas e do gás em galáxias (de Amorim et al. 2017; Vale Asari et al. 2016). As equipes da UFSC e estrangeiras têm, tanto de forma independente como conjunta, ampla experiência nesse tipo de trabalho, o que as coloca em posição privilegiada para liderar a exploração da nova geração de dados que se avizinha em projetos como o MaNGA, MUSE, Sitelle, J-PAS, e S-PLUS, os quais proporcionarão grandes avanços quantitativos e qualitativos em estudos de estrelas e galáxias e sua evolução ao longo da história do Universo.

Grandes experimentos de colisões de núcleos atômicos revelam a física das partículas elementares. Uma das formas de entender as partículas elementares é a partir da fenomenologia, que é a pesquisa na fronteira entre teoria e experimento (Oliveira et al. 2017). Grandes experimentos, com o LHC, produzem milhões de dados de colisões. A fenomenologia identifica observáveis físicos interessantes, cria previsões teóricas para eles e então os confronta com os dados obtidos nos laboratórios internacionais. O principal objetivo é encontrar evidências de física nova e então compará-la com resultados experimentais, de tal forma a aumentar o nosso conhecimento sobre o Universo, se possível, em um nível mais fundamental do que o atual.

Ainda, nessa mesma linha, o entendimento da evolução energética, temporal e de densidade do nosso universo pode ser descrito pelo diagrama de fases da cromodinâmica quântica (QCD), que requer dados tanto observacionais quanto experimentais. Vários aspectos dessa evolução já foram explicados (Pais et al. 2016, Marquez and Menezes 2017), mas é importante salientar que cada região (densidade, temperatura) pode contribuir com todo o conjunto de investigações associadas a cobrir todo o diagrama de fases da QCD (Torres et al. 2017).

Na área biotecnológica, a geração de dados biológicos tem aumentando significativamente a partir da consolidação dos sequenciamento de nova geração (NGS) (Lecroc et al. 2013) em que genomas completos de diferentes espécies ou bilhões de sequências de DNA de vários organismos presentes em uma amostras biológica, como água, solo, superfície de hospitais e até mesmo de mucosas e fezes humanas, são gerados em estudos de metagenômicos (Schmidt & Hildebrandt 2017). Concomitante a isto, com o desenvolvimento de equipamentos mais sensíveis, o estudo da proteômica com a análise de dados de espectrometria de massas de proteínas estão sendo gerados em larga escala e em uma velocidade sem precedentes, em que a partir de uma única análise em um destes equipamentos é possível identificar milhares de proteínas a partir de bilhões de espectros de massas (Awan & Saeed, 2017). Este crescimento de dados tem demonstrado um potencial informativo para se conhecer os diferentes modelos biológicos, encontrar alvos para diagnóstico, vacinais e terapêuticos, até mesmo uma medicina personalizada (Nice, 2016). Contudo, a capacidade de armazenamento, processamento e análise desta grande quantidade de dados tem se tornado um gargalo para estes estudos e os bioinformatas atuais. Neste sentido, a aplicação e desenvolvimento de conceitos e tecnologias em Big Data dentro da bioinformática são fundamentais para o melhor uso destes dados (He et al., 2017).

Enquanto a Física e a Biologia buscam soluções nas suas áreas de aplicação, a Computação cria as novas técnicas computacionais de aprendizado de máquina que, aliadas a técnicas de alto desempenho, permitem processar e extrair conhecimento novo e útil de grandes volumes de dados, sejam estes de uma única fonte, ou de múltiplas fontes que requerem técnicas de integração de dados heterogêneos. A equipe da Computação envolvida no projeto tem trabalhado nos últimos anos com grandes volumes de dados de trajetórias (rastros) geradas por GPS e outros dispositivos móveis como smartphones, desenvolvendo novas técnicas de análise de padrões de movimentação (Parent et al. 2013), novas técnicas de análise de similaridade (Salvaro et al 2018) que podem ser utilizadas na análise de agrupamentos (clustering), novas técnicas para detecção de comportamentos anômalos (Cosar et al 2017), novos algoritmos de classificação (Ferrero et al 2018), etc., considerando além do volume e a heterogeneidade dos dados, a distribuição do processamento de forma a aumentar a escalabilidade (Castro et al 2016), (Pena et al 2018).

A cooperação destes três programas, com a liderança da Computação, irá expandir a interdisciplinaridade dentro da UFSC, unindo três programas de PG em uma temática que vem ganhando importância internacional, e buscará se tornar referência mundial em pesquisa em Big Data Analytics em conjunto com instituições internacionais.

Cabe ressaltar que na UFSC verifica-se uma lacuna na cooperação entre programas de áreas tão distintas, mesmo que hoje deseje-se a formação de profissionais com conhecimento interdisciplinar. Neste contexto, a cooperação entre discentes e docentes de programas em temáticas distintas (física, computação e biologia) porém complementares, e com instituições internacionais de renome nas áreas, contribuirá significativamente para o desenvolvimento de recursos humanos e para a consolidação dos programas de PG e da UFSC em Data Science e Big Data para o entendimento de diferentes fenômenos naturais.

Apresentar a relevância do projeto do ponto de vista científico, tecnológico, econômico e/ou social.

Apesar do já alto grau de internacionalização dos docentes do grupo, os docentes têm enfrentado dificuldades em financiar a participação direta de pós-graduandos nessas colaborações. Com o CAPES/PRINT pretende-se remediar essa dificuldade, oferecendo mais oportunidades aos estudantes de PG com financiamento regular. Isso contribuirá diretamente para a formação acadêmica e científica desses estudantes, e ao mesmo tempo elevando o grau de internacionalização da UFSC.

Do ponto de vista científico, este projeto visa uma forte cooperação internacional na pesquisa em Big Data Analytics. A cooperação internacional irá envolver os pesquisadores da UFSC nos temas de pesquisa de ponta sendo discutidos pelos grupos estrangeiros, aumentando significativamente a produção bibliográfica dos pesquisadores brasileiros em periódicos internacionais, e conseqüentemente os impactos das pesquisas produzidas pelos brasileiros. Desta forma, auxiliará a consolidação de grupos de pesquisa, em especial grupos emergentes formados por jovens professores e pesquisadores da UFSC.

Além disso, este projeto interdisciplinar entre Computação, Física e Biotecnologia pode ser a semente para uma colaboração muito maior dentro da UFSC, podendo no futuro envolver outras áreas que utilizam técnicas computacionais afins.

Do ponto de vista tecnológico, vários dos resultados e algoritmos desenvolvidos serão compartilhados com a comunidade científica, gerando produção tecnológica de interesse para a UFSC e para o país. Um exemplo é a aplicação em saúde da medicina personalizada, que combina dados genômicos/proteômicos de cada paciente aos dados clínicos e epidemiológicos utilizando algoritmos de Big Data Analytics. Desta forma, a criação deste grupo multidisciplinar permitirá explorar esta tendência mundial no tratamento de doenças.

Do ponto de vista econômico e social, este projeto fomentará a formação de recursos humanos capazes de navegar por diferentes campos em que há aplicação de Big Data. Além disso, permitirá que este setor se desenvolva no Brasil, em especial em Florianópolis, possibilitando a formação de polos de inovação neste setor, e conseqüentemente gerando mais empregos tanto no setor privado quanto público.

Atividades Propostas (2000 caracteres)

Apresentar o plano de trabalho do projeto, explicitando as atividades que justificam os insumos solicitados, incluindo as missões de trabalho e de estudo previstas.

Pelo menos 70% dos recursos serão dedicados a colaborações com instituições nos países prioritários definidos pela CAPES. Algumas das missões de trabalho no exterior, escolas e conferências podem ser combinadas em uma única viagem, dispensando, assim, o uso de recursos para mais de uma passagem aérea.

As atividades tentam equilibrar visitas de pesquisadores e estudantes. Prevêem-se as atividades listadas a seguir durante a vigência desse projeto.

Visitas de curta duração: visitas a instituições estrangeiras ou de colaboradores estrangeiros são altamente produtivas. Prevê-se que alguns dos visitantes estrangeiros ministrem, em inglês, mini-cursos ou disciplinas com créditos nos Programas de PG. São previstas:

- 16 missões de trabalho para pesquisadores brasileiros visitarem colaboradores no exterior.
- 9 missões de trabalho para pesquisadores estrangeiros visitarem a UFSC.

Visitas longas: visitas mais extensas ao exterior de pesquisadores e alunos (período sabático e doutorado sanduíche), e atração de pós-docs com experiência no exterior ao Brasil. Essas visitas mais longas têm um caráter diferente das de curto prazo. Para os estudantes, um período mais longo em outras instituições, em contato com outra língua, outros pesquisadores, ajudam no seu amadurecimento e independência científica. Para os pesquisadores, períodos sabáticos podem ser altamente produtivos e incentivam linhas de pesquisas mais ambiciosas. Por fim, a atração de pós-docs e jovens pesquisadores para a UFSC, traz um impacto positivo para o programa como um todo, principalmente quando eles se envolvem no dia-a-dia da PG promovendo seminários, journal clubs, etc. Estão previstas nesta modalidade:

- 6 professores visitantes junior/senior no exterior
- 6 doutorados sanduíches no exterior.
- 6 pós-docs/jovens talentos no Brasil.

Conferências e escolas avançadas: A participação em conferências é fundamental para divulgação da pesquisa feita na UFSC. Escolas avançadas são uma ótima oportunidade para os alunos da UFSC adquirirem conhecimentos em cursos que não existem nos seus PPG e também para ampliarem as suas redes de contatos profissionais. Foram previstas:

- até 9 conferências no exterior para pesquisadores, previstas nas 16 missões de trabalho acima.
- 12 escolas de curta duração para estudantes.

Eventos interdisciplinares, que congregam pesquisadores e discentes de diferentes áreas do conhecimento, são uma excelente oportunidade para troca de experiências e soluções inovadoras. Por isso, o orçamento do projeto prevê recursos para trazer três pesquisadores estrangeiros no último ano do projeto, para a realização de um **Simpósio Internacional em Big Data Analytics**. Esse evento terá a participação de pesquisadores das instituições parceiras deste projeto e convidados de outras instituições nacionais.

Caráter Inovador (2000 caracteres)

Apresentar inovações trazidas pelo projeto nas dimensões institucional, acadêmica, tecnológica e de inserção internacional à UFSC.

Destacamos a seguir as diversas facetas inovadoras desse projeto. Do ponto de vista institucional, ele inaugura uma colaboração interdisciplinar entre áreas que trabalham com big data analytics dentro da UFSC: Computação, Física e Biologia. Dentre as ações concretas para estabelecer essa colaboração inédita, propomos palestras interdisciplinares visando a integração dos programas. Além disso, haverá a criação de uma disciplina comum aos três programas de pós-graduação com um foco em Big Data. Por fim, pre vemos um Simpósio Internacional em Big Data no último ano do projeto, um evento internacional multidisciplinar com participação de pesquisadores e estudantes da UFSC, das instituições estrangeiras parceiras, e de outras instituições nacionais.

Do ponto de vista acadêmico, o estabelecimento de laços mais fortes com instituições internacionais e pesquisadores renomados terá um impacto quantitativo e qualitativo na produção intelectual da UFSC. Esse impacto trará maior visibilidade para a pesquisa feita na instituição. Além disso, o projeto será importantíssimo para a formação estudantes de pós-graduação, que serão capazes de transitar entre diferentes campos de data science. Além disso, a inserção desses jovens pesquisadores dentro do panorama de colaborações internacionais fará com que eles tenham uma formação mais sólida e compatível com uma nova geração de pesquisadores brasileiros com fluência internacional.

Do ponto de vista tecnológico, as técnicas desenvolvidas para a análise de Big Data ficarão disponíveis para a comunidade científica. Os códigos desenvolvidos por vários pesquisadores serão disponibilizados, o que aumentará a visibilidade, as citações aos trabalhos e a relevância da ciência feita na UFSC. Além disso, o acesso aberto contribui para a transparência e reprodutibilidade na comunidade científica.

Do ponto de vista da inserção internacional da UFSC, este projeto irá reforçar antigas e criar novas cooperações internacionais, promovendo o intercâmbio de pesquisadores e principalmente de alunos. Os pesquisadores deste projeto já têm uma alta inserção internacional, e pretendem, a partir de um financiamento mais regular, aumentar a inserção internacional dos estudantes de pós-graduação. Dentro do escopo das atividades aqui propostas, serão enviados estudantes da UFSC para estágios de doutorado sanduíche, escolas avançadas e congressos.

Além disso, dentro deste projeto prevê-se várias ações internas que irão ajudar na internacionalização dos PPG da UFSC. Além de inserir estudantes da UFSC dentro do contexto globalizado da ciência atual, um outro objetivo é atrair estudantes estrangeiros para os programas de PG da UFSC. Dentre essas ações, listadas na seção Objetivos, destacamos a oferta de disciplinas e material didático em língua inglesa nas PGs nas áreas envolvidas; a oferta de minicursos e palestras em língua inglesa; a motivação aos alunos de pós-graduação para que escrevam suas teses e dissertações em inglês; o desenvolvimento de uma versão em inglês do site dos programas; e a divulgação em inglês do processo seletivo ao mestrado e doutorado. Vale lembrar que a troca de experiências entre os programas fará com que essas ações possam ser implementadas com mais celeridade: o PPGFSC, por exemplo, já conta com um site e um programa de seleção em inglês.

Bibliografia de Referência

Relacionar no máximo 20 referências relevantes à temática proposta.

1. de Amorim, A. L.; García-Benito, R.; **Cid Fernandes, R.**; Cortijo-Ferrero, C.; **González Delgado, R. M.**; Lacerda, E. A. D.; López Fernández, R.; Pérez, E.; **Vale Asari, N.** The PyCASSO database: Spatially resolved stellar population properties for CALIFA galaxies. *Monthly Notices of the Royal Astronomical Society*, v. 471, p. 3727-3752, 2017. (www.pycasso.ufsc.br)
2. **Vale Asari, N.**; **Stasinska, G.**; Morisset, C.; **Cid Fernandes, R.** BOND: Bayesian Oxygen and Nitrogen abundance Determinations in giant H II regions using strong and semi-strong lines. *Monthly Notices of the Royal Astronomical Society*, v. 460, p. stw971, 2016. (www.bond.ufsc.br)
3. Mello, A. T.; **Kanaan, A.**; Guzman, D.; Guesalaga, A. Artificial neural networks for centroiding elongated spots in Shack-Hartmann wavefront sensors. *Monthly Notices of the Royal Astronomical Society*, v. 440, p. 2781-2790, 2014.
4. **Saito, R. K.**, Hempel, M., **Minniti, D.**, et al. VVV DR1: The first data release of the Milky Way bulge and southern plane from the near-infrared ESO public survey VISTA variables in the Vía Láctea, *Astronomy & Astrophysics*, 537, A107, 2012.
5. **Oliveira, E. G. de**; **Martin, A. D.**; Ryskin, M. G. Scale dependence of open ccbar and bbbar production in the low x region. *European Physical Journal. C, Particles and Fields*, v. 77, p. 182, 2017.
6. COSAR, S. ; DONATIELLO, G. ; **Bogorny, V** ; GARATE, C. ; ALVARES, L. O. ; BREMOND, F. Toward Abnormal Trajectory and Event Detection in Video Surveillance. *IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY*, v. 27, p. 683-695, 2017.
7. FURTADO, A. S. ; ALVARES, L. O. C ; PELEKIS, N. ; THEODORIDIS, Y. ; **Bogorny, V** . Unveiling movement uncertainty for robust trajectory similarity analysis. *International Journal of Geographical Information Science*, v. 32, p. 140-168, 2017.
8. PARENT, C. ; SPACCAPIETRA, S. ; RENSO, C. ; ADRIENKO, G. ; ADRIENKO, N. ; **BOGORNÝ, V.** ; Damiani, M. L. ; GKOUALAS-DIVANIS, A. ; MACEDO, J. A. F. ; PELEKIS, N. ; THEODORIDIS, Y. ; Yan, Z . Semantic trajectories modeling and analysis. *ACM Computing Surveys*, v. 45, p. 1-32, 2013.
9. FERRERO, C. A. ; ALVARES, Luis Otavio ; ZALEWSKI, W. ; **BOGORNÝ, V.** . MOVELETS: Exploring Relevant Subtrajectories for Robust Trajectory Classification. In: *ACMSAC, 2018. Proceedings of the 33 Symposium on Applied Computing, 2018.*
10. PENNA, P. H. ; **CASTRO, M.** ; FREITAS, H. C. ; BROQUEDIS, FRANÇOIS ; MÉHAUT, JEAN-FRANÇOIS . Design methodology for workload-aware loop scheduling strategies based on genetic algorithm and simulation. *CONCURRENCY AND COMPUTATION-PRACTICE & EXPERIENCE*, v. 29, p. e3933, 2017.
11. **Castro, M**; Francesquini, E; Dupros, F; Aochi, H.; Navaux, P. O. A. ; Méhaut, J.F.. Seismic Wave Propagation Simulations on Low-power and Performance-centric Manycores. In: *Parallel Computing (PARCO)*, v. 54, 2016.
12. Schmidt B, Hildebrandt A. Next-generation sequencing: big data meets high performance computing. *Drug Discov Today*. 2017 Apr;22(4):712-717.
13. Lecroq T, Soualmia LF; Section Editors for the IMIA Yearbook Section on Bioinformatics and Translational Informatics. From genome sequencing to bedside. Findings from the section on bioinformatics and translational informatics. *Yearb Med Inform*. 2013;8:175-7.
14. He KY, Ge D, He MM. Big Data Analytics for Genomic Medicine. *Int J Mol Sci*. 2017 Feb 15;18(2). pii: E412.
15. Awan MG, Saeed F. An Out-of-Core GPU based dimensionality reduction algorithm for Big Mass Spectrometry Data and its application in bottom-up Proteomics. *ACM BCB*. 2017 Aug;2017:550-555.
16. Nice EC. From proteomics to personalized medicine: the road ahead. *Expert Rev Proteomics*. 2016;13(4):341-3.
17. Pais, Helena; **Menezes, Débora P.** ; **Providência, Constança**. Neutron stars: From the inner crust to the core with the (extended) Nambu-Jona-Lasinio model. *PHYSICAL REVIEW C*, v. 93, p. 065805, 2016.
18. Marquez, Kauan D; **Menezes, Débora P.** Phase transition in compact stars: nucleation mechanism and γ -ray bursts revisited. *JOURNAL OF COSMOLOGY AND ASTROPARTICLE PHYSICS*, v. 2017, p. 028-028, 2017.
19. Torres, James R.; **Gulminelli, Francesca**; **Menezes, Débora P.** Examination of strangeness instabilities and effects of strange meson couplings in dense strange hadronic matter and compact stars. *PHYSICAL REVIEW C*, v. 95, p. 025201, 2017.

OBJETIVOS, METAS E IMPACTOS DO PROJETO DE PESQUISA EM COOPERAÇÃO INTERNACIONAL

Objetivos (2000 caracteres)

Apresentar a definição e limitação do objeto de estudo.

Os objetivos desse projeto estão voltados para a internacionalização dos programas de pós-graduação participantes e para estreitar os laços entre pesquisadores na UFSC que trabalham com grandes bases de dados, cujas técnicas podem ser aplicadas a diferentes problemas científicos. A seguir são listados alguns objetivos específicos.

- a) Ofertar disciplinas e material didático em língua inglesa no âmbito dos três programas de PG envolvidos no projeto. Com estas disciplinas o estudante se tornará especialista no tema do projeto do ponto de vista computacional, tendo como base aplicações de áreas diferentes (a Física, a Biologia e a própria Computação).
- b) Ofertar minicursos e palestras em língua inglesa, organizados tanto por pesquisadores brasileiros quanto estrangeiros, tanto no modelo presencial quanto a distância.
- c) Motivar os alunos de pós-graduação na escrita de suas teses e dissertações na língua inglesa, o que irá aumentar significativamente o número de downloads e leitura destes trabalhos por pesquisadores estrangeiros.
- d) Financiar visitas curtas de pesquisadores renomados das instituições internacionais para atuar nos grupos de pesquisa brasileiros, buscando a elaboração de artigos científicos conjuntos, e oferta de novas disciplinas de 1 ou 2 créditos que possam ser ministradas em curtos períodos.
- e) Promover palestras e seminários em inglês durante a visita dos participantes estrangeiros à UFSC.
- f) Promover Summer / Winter schools ou *hackathons*, ministrados pelos discentes dos programas de PG tendo como público alvo alunos de graduação.
- g) Abrir concurso para docente no PPGCC/UFSC na área de Big Data Analytics / Data Science, uma vez que vários docentes estão entrando com pedidos de aposentadoria.
- h) Fazer uma versão em inglês do site dos programas PPGCC e PPGBTC, sendo que o PPGFSC já possui.
- i) Organizar uma competição internacional de Big Data Analytics, tendo como prêmio uma bolsa de pos-doc, de um ano, na UFSC em um dos programas de PG participantes deste projeto (escolhido pela equipe vencedora).
- j) Elaborar um folder em inglês sobre o processo seletivo ao mestrado e doutorado e divulgar em listas de emails internacionais, visando atrair alunos estrangeiros para os Programas de Pós-Graduação da UFSC. Vale lembrar que a participação do processo seletivo da PPGFSC já pode ser feito completamente em inglês.
- k) Organizar um Simpósio Internacional Interdisciplinar em BigData Analytics, com a participação de pesquisadores das instituições parceiras deste projeto e convidados de outras instituições nacionais.

Metas (2000 caracteres)

Apresentar quantitativamente os produtos esperados do projeto em termos de formação de pessoas e divulgação de resultados da pesquisa. Exemplos: disciplinas oferecidas em língua estrangeira (item obrigatório), nº de estrangeiros visitantes, total de publicações de relevância internacional, total de itens de produção tecnológica registráveis/registrados, co-tutelas.

As metas abaixo são para o período de quatro anos de vigência do projeto.

- Ofertar ao menos três disciplinas em língua inglesa (ministradas por professores das PGs e por pesquisadores visitantes) em cada programa de PG. Na computação já estão confirmadas as disciplinas Data Management for Big Data e Parallel Computing. A física irá propor à PPGFSC as disciplinas (1) Astrophysics with large databases e (2) Nuclear and Particle Physics under a modern perspective.

- Disponibilizar o material de várias disciplinas em inglês.
- Criar disciplinas em língua inglesa comuns aos 3 PPG e que serão ministradas em conjunto por docentes dos três programas: Por exemplo: (i) Multiple Aspect Big Data Analytics, (ii) Data Science e (iii) High Performance Big Data Processing.
- Oferecer disciplinas optativas em língua inglesa na Graduação como, por exemplo: Big Data Storage, Management and Mining e Statistics for Big Data and Data Science, uma vez que o PPGCC está no Departamento de Informática e Estatística e conta com o apoio dos docentes da Estatística.
- Publicar diversos artigos em revistas internacionais de alto impacto (considerando a área de cada PPG). A meta é publicar ao menos 2 artigos por docente de cada programa por ano, em média, totalizando mais de 80 artigos em periódicos durante o projeto.
- Receber a visita de ao menos 9 visitantes estrangeiros, detalhadas no plano de atividades e orçamento
- 6 doutorados sanduíche e ao menos uma co-tutela.

Além disso, valem ser destacadas as seguintes ações:

- Simpósio interdisciplinar no último ano do projeto.
- Primeiros passos para criação de especialização multidisciplinar em big data analytics.

Impactos na Internacionalização da UFSC (2000 caracteres)

Apresentar impactos na internacionalização da UFSC em termos científicos, tecnológicos, econômicos e/ou sociais.

Os PPG participantes desse projeto já têm algumas ações pontuais que serão importantes para a internacionalização dos seus programas. Inicialmente, os sites do PPGCC e PPGBTC deverão ser traduzidos e o processo seletivo será ampliado com divulgação internacional.

A longo prazo, visualiza-se os seguintes impactos:

- (1) Formação de alunos: Este projeto terá uma grande contribuição para a promoção do ensino e formação, principalmente no nível de pós-graduação. O caráter interdisciplinar desse projeto, a melhora de indicadores de produção e de internacionalização ajudará a estudantes que antes não considerariam a UFSC para seu mestrado/doutorado.
- (2) Produção intelectual: A combinação de técnicas e conhecimentos será chave para atacar problemas atuais, com dados que são o estado-da-arte, que por sua vez terá um impacto imediato na relevância da produção científica. A quantidade de publicações em revistas de alto impacto, principalmente na área de Big Data, aumentará a atenção internacional para a pesquisa na UFSC. Esse é um dos parâmetros mais importantes considerados na avaliação da CAPES nos cursos de pós-graduação. Além disso, diversas publicações terão acesso aberto através do servidor de pre-prints arXiv.
- (3) Produção técnica e acesso aberto: As ferramentas e grandes bases de dados desenvolvidas na UFSC bem como os novos códigos resultantes serão abertos, e os resultados das análises serão públicos, contribuindo para a transparência, reprodutibilidade e compartilhamento de ferramentas computacionais na comunidade científica.
- (4) Perspectivas para o futuro: As colaborações resultantes desse projeto irão desenvolver técnicas adaptadas para explorar futuros conjuntos de dados de alto impacto, como os esperados em breve (por exemplo, SDSS-V LVM e JWST na astrofísica). Isso aumentará também, a médio prazo, a visibilidade da pesquisa da UFSC.
- (5) Investimento em ciência básica: O Brasil é o maior país da América Latina, mas o investimento e a produtividade na ciência básica e aplicada são menores que, por exemplo, no Chile e Uruguai (ver relatório científico da UNESCO: em 2030, capítulo 7, <http://unesdoc.unesco.org/images/0023/002354/235407por.pdf>). Portanto, grupos de pesquisa básica bem-sucedidos com trabalho de alto impacto trazem benefícios enormes para o país: é sabido que investimentos em ciência básica, no longo prazo, refletem-se em desenvolvimento tecnológico e melhoria nos indicadores de desenvolvimento social de um país.