



**Centro Federal de Educação Tecnológica
Celso Suckow da Fonseca
CEFET/RJ**

**O Programa de Pós-Graduação em Ciência da
Computação (PPCIC) do CEFET/RJ e
sua ligação com os desafios em Ciência de Dados**

<http://eic.cefet-rj.br/ppcic>

Ciência de Dados: Carreira (Data Scientist)



2012

DATA

Data Scientist: The Sexiest Job of the 21st Century

by Thomas H. Davenport and D.J. Patil

BUSINESS INSIDER CAREERS

Here's how much money you make in the 'sexiest job of the 21st century'

Jacquelyn Smith Feb. 25, 2016, 3:20 PM 33,395

In 2012, Harvard Business Review named data scientist the "[sexiest job of the 21st century](#)." More recently, Glassdoor named it the "best job of the year" for 2016.

"It isn't a big surprise," Dr. Andrew Chamberlain, Glassdoor's chief economist, told Business Insider. "It's one of the hottest and fastest growing jobs we're seeing right now."

According to Glassdoor, data scientists earn a base pay of **\$116,840 a year**, on average.



Facebook: \$133,841

Apple: \$149,963

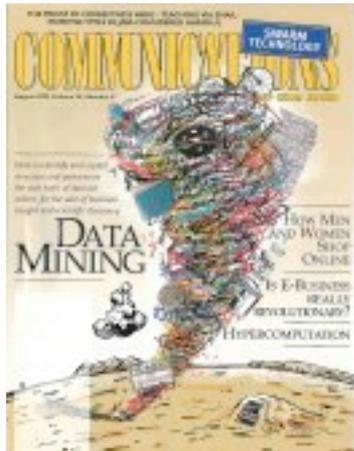
Airbnb: \$117,229

Twitter: \$134,861

Microsoft: \$119,129

LinkedIn: \$138,798

Contexto de Ciência de Dados



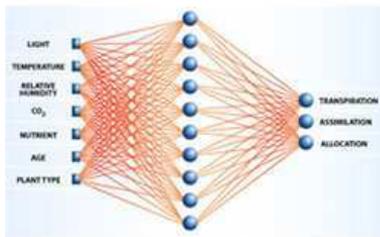
Ago/2002



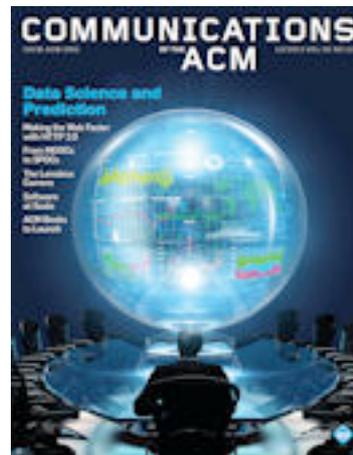
dez/2008



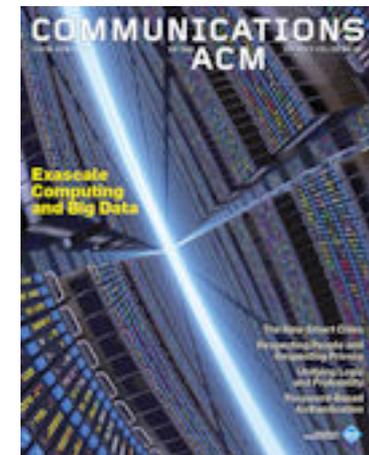
IoT



Deep Learning



dez/2013

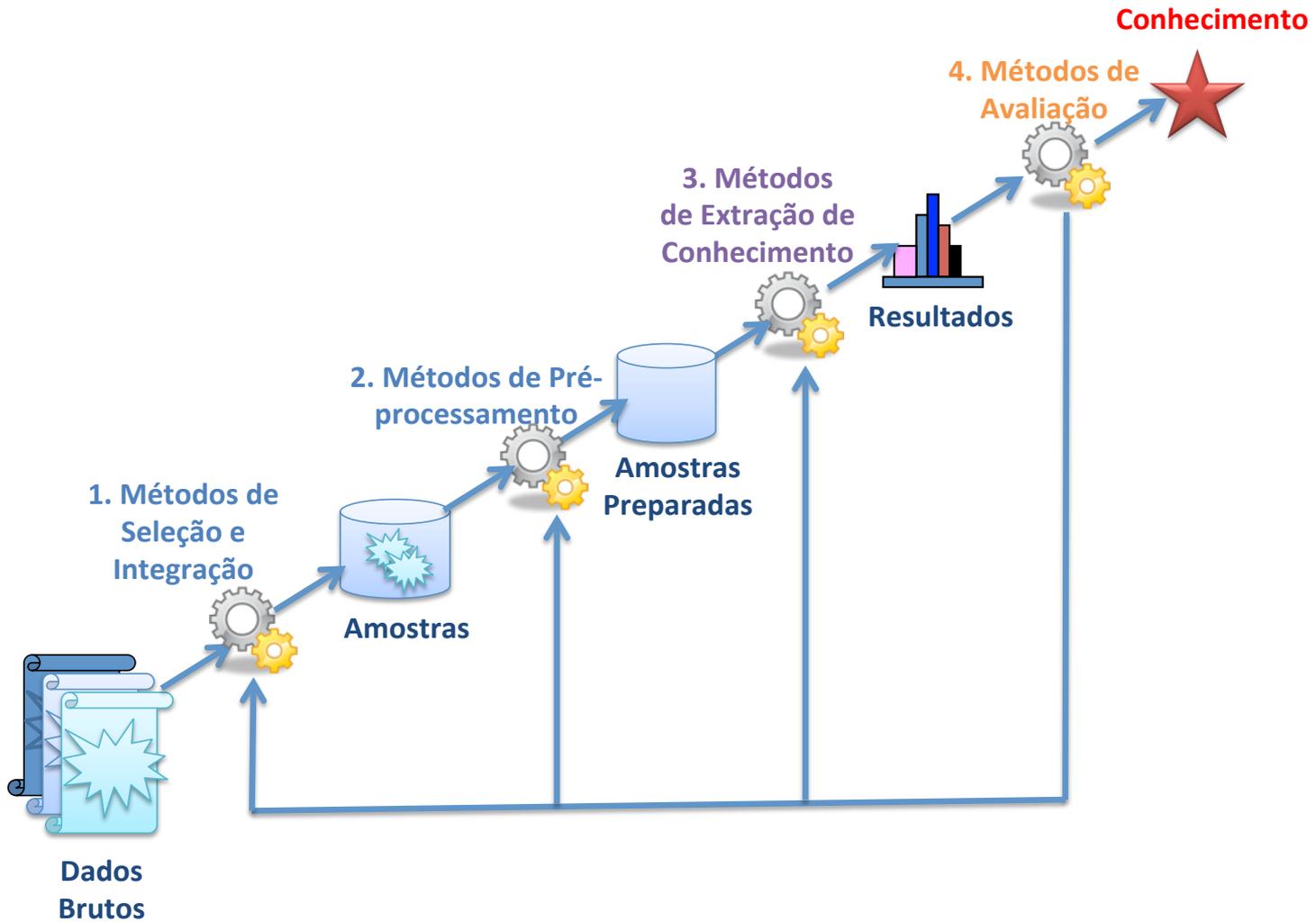


Jul/2015

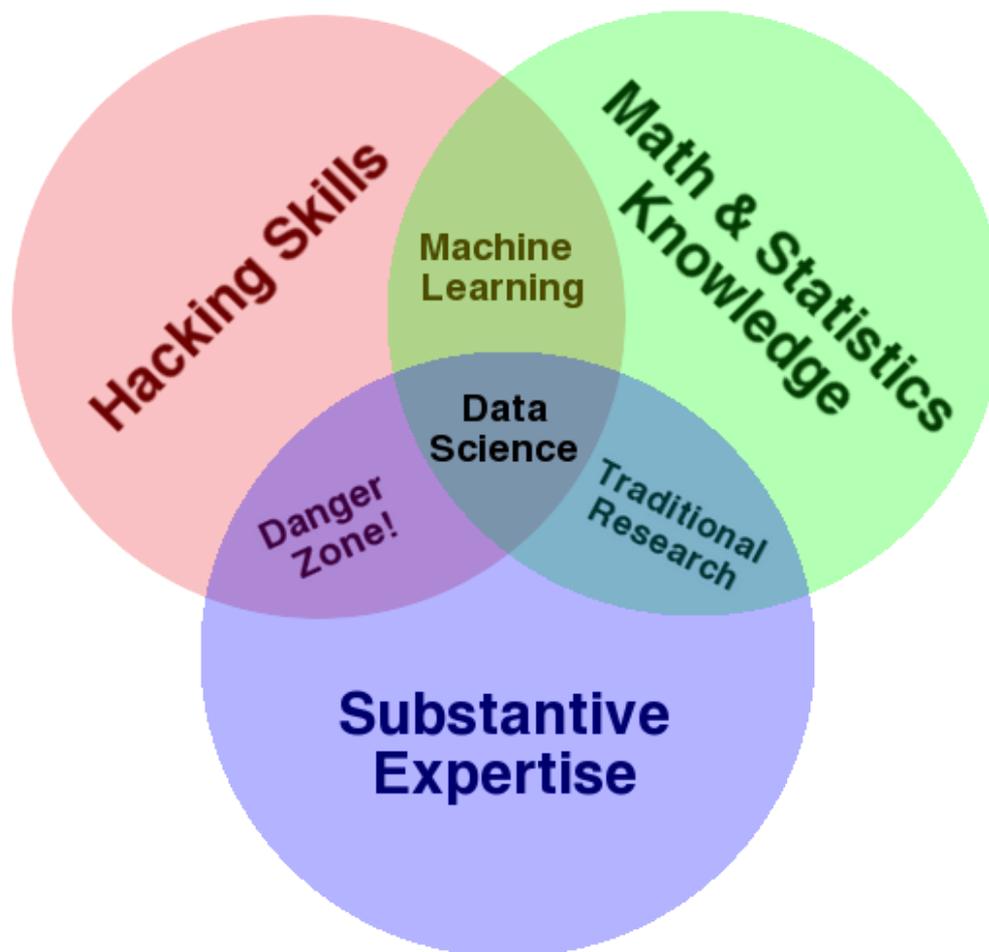
Ciência de Dados

- Ciência de dados é o estudo da extração de conhecimento a partir de dados (Big Data)
- Envolve pesquisa em
 - Métodos estatísticos, aprendizagem de máquina, técnicas de visualização de dados
 - Gerência e armazenamento de grande coleções de dados
 - Computação de alto desempenho
- Característica multidisciplinar
- Aplicações na ciência/indústria/governo
 - Bioinformática, petróleo, energia, finanças, astronomia, Internet, mobilidade urbana, defesa cibernética, educação, etc

Processo de Ciência de Dados



Ciência de Dados – Habilidades necessárias para um Cientista de Dados



Fonte: <http://drewconway.com/>

Concepção da Proposta do Mestrado em Ciência da Computação do CEFET/RJ

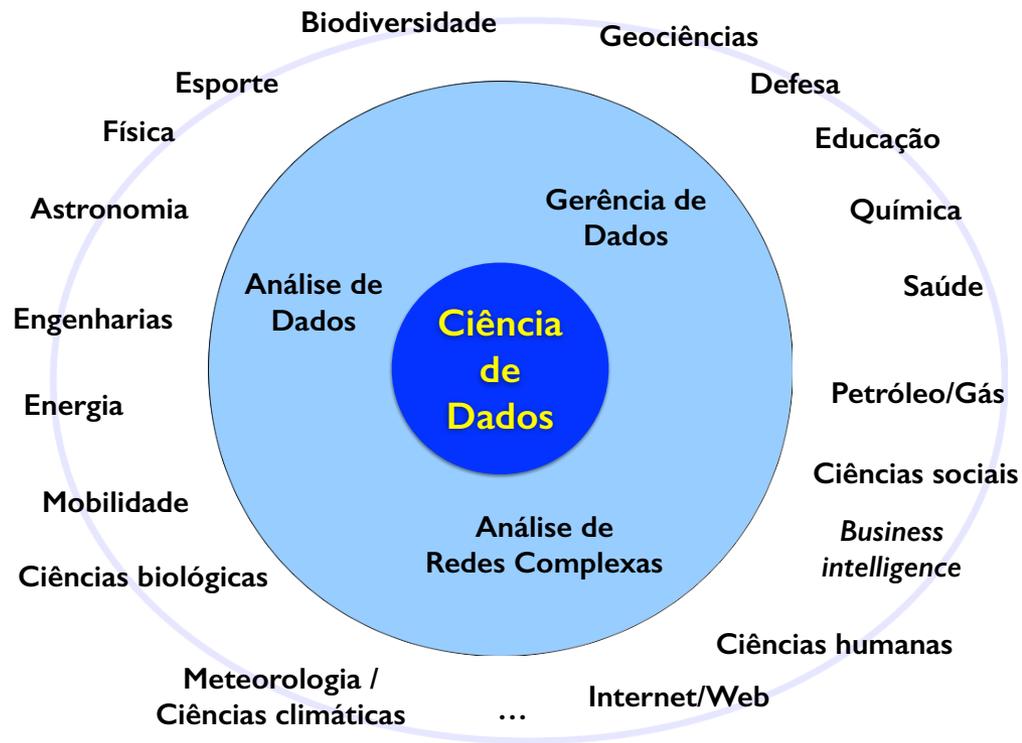
INCT - Ciência de Dados (CID)

- Objetivo
 - Estruturar a nova área
 - **Formação de Recursos Humanos**
 - Pesquisa e Desenvolvimento
 - Transferência de Tecnologia para Indústria e Sociedade
 - Como?
 - Através da Integração de disciplinas com foco em dados
 - Bancos de Dados
 - Análise de Dados
 - Modelagem de Dados
 - Metodologias
 - Dirigida à hipóteses
 - Análise estatística
 - Tecnologias
 - Frameworks Big Data: Spark, Giraph, Tensor Flow,
 - Linguagens de programação: R, python, scala
 - Sistemas: NoSQL, Greenplum, ...
 - Machine Learning: scikit-learn, Apache Mahout,..

INCT - CID

- Parcerias
 - FIOCRUZ
 - Observatório Nacional
 - Comitê Olímpico Brasileiro
 - Prefeituras
 - Curitiba
 - Rio de Janeiro
- Empresas
 - EMC - Centro de Pesquisa
 - IBM Research
 - Philips Research
- Conselho Internacional
 - NYU
 - INRIA
 - UCSB
 - Boston University:
Data Science Institute
- Coordenação
 - Artur Ziviani (LNCC)
 - Fabio Porto (LNCC)
- Comite Gestor
 - Artur Ziviani (LNCC)
 - Fabio Porto (LNCC)
 - Anelise Munaretto (UTFPR)
 - Ronaldo Salles (IME)
 - Eduardo Ogasawara (CEFET/RJ)

Organização do INCT-CID



Contexto Histórico do CEFET/RJ

Criação do Bacharelado em Ciência da Computação com perfil diferenciado ...

*Planejada para voos mais altos...
Aprovação passo a passo nos Conselhos*

- Conselho Nacional de Educação

- Diretrizes para Ciência da Computação

*"Os cientistas da computação são **responsáveis pelo desenvolvimento científico** (teorias, métodos, linguagens, modelos, entre outras) e tecnológico da Computação....**Os cientistas da computação aplicam métodos e processos científicos para o desenvolvimento de produtos corretos.**"*

- Formar um egresso mais sólido na capacidade de P&D

*"Possuam sólida formação em Ciência da Computação e Matemática que os capacitem a construir aplicativos de propósito geral, ferramentas e infraestrutura de software de sistemas de computação e de sistemas embarcados, **gerar conhecimento científico e inovação** e que os incentivem a estender suas competências à medida que a área se desenvolva"*

... viabilizador da pesquisa e pós-graduação

- O perfil do egresso do BCC desenvolvido tem, dentre outros, viés para pesquisa
 - possibilita ao aluno realizar P&D e aprimorar os estudos
- A base de sustentação projetada para o nosso mestrado são os alunos do nosso curso de Ciência da Computação
- Ações concretas de pesquisa na graduação como reação a atração do mercado (estágio) no começo do curso
 - Disciplina de Metodologia Científica
 - Disciplina de Prática de Pesquisa
 - Reformulação da Norma de TCC
 - *Revista Cadernos de Computação Aplicada*

Disciplina de Metodologia Científica

- Explora os conceitos gerais para produção de monografias e, principalmente, artigos científicos
- Revisão Bibliográfica & Trabalhos Relacionados
- Processo de elaboração de textos
 - Estrutura, Encadeamento, Forma, Gráficos, Tabelas, Algoritmos, Formalização
- Plágio
- Processo de submissão e avaliação de textos científicos
- Trabalho Prático
 - Leitura de artigos científicos do estado da arte
 - Elaboração de revisão bibliográfica
 - Avaliação de trabalhos

Disciplina de Prática de Pesquisa

■ Objetivo

"Proporcionar ao discente vivência prática nas etapas de processo de elaboração de artefatos científicos aderentes às linhas de pesquisa dos professores"

■ Tipos de Práticas

- Modelagem/Implementação de Algoritmos ou Avaliação Experimental
- Internacionalização de artigos
- Escrita de artigos provenientes de TCCs

■ Concebida para formar alunos com maior contato a P&D

■ Diferencial sobre qualquer outro BCC do Brasil

■ **É a disciplina mais importante do BCC**

- Indutora de temas de TCC e de opção para o mestrado
- Fomentadora para pesquisadores

■ Adequada aos alunos de mestrado para estágio em docência

- Estruturação da Pirâmide

Norma de TCC

- Consolida Metodologia Científica e Prática de Pesquisa
- TCC em forma de artigo
- O TCC não precisa ser grande, mas precisa ser completo
 - Introdução
 - Revisão Bibliográfica
 - Metodologia
 - Avaliação Experimental
 - Conclusões

Criação do Bacharelado em Ciência da Computação

2012-08-20!

The Day After

- Missão: Criação do Mestrado em Computação
- Fortalecimento do(s) grupo(s) de pesquisa
 - Servir como elo da Escola de Informática & Computação (EIC)
 - Fomentar a pesquisa no Curso Técnico de Informática
 - Fortalecer o curso de Ciência da Computação trazendo um viés em pesquisa
 - Posicionar a EIC em destaque no CEFET/RJ
 - Posicionar a computação do CEFET/RJ no cenário nacional

Análise das ações de pesquisa

- Metodologia Científica – 2014-2 (5 semestres)
- Prática de Pesquisa – 2015-1 (4 semestres)
- Resultados Quantitativos
 - Artigos em Anais de Congresso
 - A2 = 1, B2 = 2
 - Artigos em Periódicos Indexados (jcr)
 - 1 artigo aceito (ecological informatics)
 - 2 artigos submetidos (1 rejeitado, 1 recusado)
 - Ambos sendo ajustados (novas submissões)
 - 2 artigos em elaboração
 - Temas de TCCs: 7
 - Temas de Mestrado: 1 (com viés de 3)
 - Algoritmos implementados: 2
- Resultados Qualitativos
 - Viabiliza Credenciamento de Docentes na Pós-Graduação
 - Molda alunos para o TCC e para o Mestrado

WEIC 2014

Projetos Inovadores Ideias Revolucionárias

Eduardo Ogasawara
eogasawara@cefet-rj.br

Inovar é preciso

- Uma inovação por ano
 - 2009: Linha de Experimentos
 - 2010: Normalização Adaptativa
 - 2011: Álgebra de Workflows
 - 2012: BCC e Integrado
 - 2013:
 - Nova norma de TCC
 - Workshop de Computação Aplicada
 - 2014:
 - Revista Cadernos em Computação Aplicada
 - INCT em Ciência de Dados
 - 2015: Submissão do APCN para Ciência da Computação

Planejamento para criação do Mestrado em Computação

- Plano Institucional de Formação de Quadros Docentes (Planfor 2013) (CEPE/CODIR)
 - "Para 2015, existe a previsão de submeter à CAPES uma proposta de mestrado na área de Ciência da Computação com início em 2016"
- Orientação aos docentes em submeter trabalhos em fóruns "relevantes"
- Parcerias interinstitucionais
- Capacitação docente
- Apresentação das normas da CAPES
 - ($A1 = 1$, $A2 = 0.85$, $B1 = 0.7$, $B2 = 0.5$, $B3 = 0.2$, $B4 = 0.1$, $B5 = 0.05$)
- Critérios gerais de credenciamento
 - (1.4 pontos – definido em outubro de 2013)

Aprovação do Mestrado em Ciência da Computação

*Capacitação e busca de docentes,
Fortalecimento da pesquisa,
Formulação das linhas de pesquisa,
Diferenciação em relação às coirmãs,
Visita do coordenador da CAPES,
...*

Submissão do projeto em 2015-08-10

*Atendimento a duas diligências...
... aprovação em 2016-04-25*

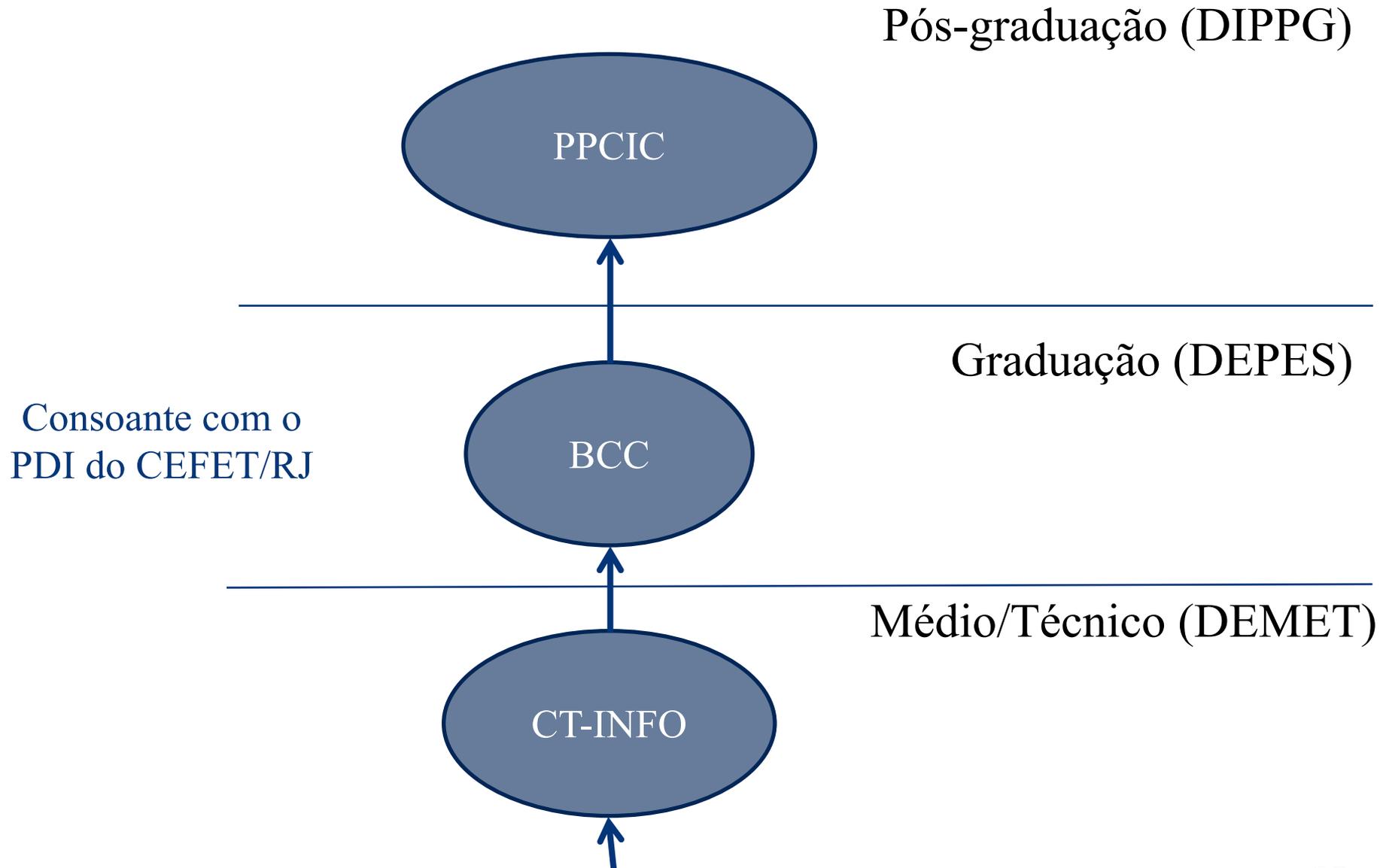
2016-09-19!

The Day After

- Missão: Criação do Doutorado em Computação
- Fortalecimento da pesquisa
 - Elo da Escola de Informática & Computação (EIC) e demais campi
 - Fomentar a pesquisa nos Cursos Técnicos de Informática
 - Fortalecer os cursos de Ciência da Computação trazendo um viés em pesquisa
 - Reforçar a computação do CEFET/RJ no cenário nacional
 - Posicionar a computação do CEFET/RJ no cenário internacional

Impacto no CEFET/RJ

1º caso de verticalização dos três níveis de ensino



Impacto externo

- Academia
 - Único APCN para Mestrado Acadêmico em Computação aprovado em 2015/2016
- Estudantes
 - Nova opção
 - Perfil diferenciado
 - Local privilegiado

PPCIC

Objetivos

- Realizar **pesquisa e desenvolvimento** de métodos computacionais para a solução de problemas complexos
- Formar **pesquisadores qualificados** em Ciência da Computação (viés de cientistas de dados)
 - Estabeleçam um ciclo virtuoso entre as pesquisas aplicada e básica
 - Possam atuar tanto no mercado quanto na academia
- Estimular o **intercâmbio de conhecimento** e a **inovação** tecnológica entre a academia e os setores produtivos

Cientistas de Dados

- Profissionais da Computação capacitados em extrair conhecimentos a partir de grandes volumes heterogêneos de dados
- Demanda de Cientistas de Dados
 - Eixo da indústria, academia e governo demandam recursos humanos qualificados desta área
 - Fundamental para que o país alcance um alto grau de inovação e pesquisa e tenha destaque internacional
 - Interesse nacional em intensificar os programas de pós-graduação que visam a formação de mestres e doutores em Computação

Linhas de Pesquisa

1. Gerência de Dados e Aplicações

- Ações aplicadas a uma coleção de dados que conduzem à descoberta de conhecimento

2. Métodos Baseados em Dados

- Otimização, Algoritmos e Modelagem computacional

Corpo Docente

Pesquisadores	Linha
Diego Barreto Haddad	Métodos Baseado em Dados
Diego Nunes Brandão	Métodos Baseado em Dados
Eduardo Bezerra da Silva	Gerência de Dados & Aplicações
Eduardo Soares Ogasawara	Gerência de Dados & Aplicações
Gustavo Paiva Guedes	Gerência de Dados & Aplicações
João Roberto de Toledo Quadros	Gerência de Dados & Aplicações
Joel André Ferreira dos Santos	Gerência de Dados & Aplicações
Jorge de Abreu Soares	Gerência de Dados & Aplicações
Kele Teixeira Belloze	Gerência de Dados & Aplicações
Laura Silva de Assis	Métodos Baseado em Dados
Leonardo Silva de Lima	Métodos Baseado em Dados
Raphael Carlos Santos Machado	Métodos Baseado em Dados

Estrutura Curricular

Tabela 1 – Disciplinas do PPCIC

Disciplina	Núcleo	Créditos
Álgebra Linear Computacional	Específico	3
Álgebra Linear e Grafos	Específico	3
Algoritmos em Grafos	Específico	3
Análise e Projeto de Algoritmos	Básico	3
Aplicações de Robótica	Específico	3
Aprendizado de Máquina	Específico	3
Arquitetura de Computadores	Básico	3
Banco de Dados	Básico	3
Computação Paralela e Distribuída	Específico	3
Fundamentos de Sistemas Multimídia	Específico	3
Gerência de Dados em Larga Escala	Específico	3
Metodologia Científica em Computação	Básico	3
Métodos Estatísticos	Básico	3
Mineração de Dados	Específico	3
Otimização por Metaheurísticas	Específico	3
Pesquisa Operacional	Específico	3
Pesquisa para a Dissertação de Mestrado	-	0
Seminário para a Dissertação de Mestrado	-	0

Avaliação do Programa

- Avaliação da Capes é quadrienal e baseada nos critérios da área de Ciência da Computação
- Processo de avaliação da Capes é complexo e rigoroso com exigências aumentando cada vez mais
- Cobrança de desempenho tanto de docentes quanto de discentes: publicações, qualidade das dissertações, tempo de conclusão do mestrado etc
- Restrições: limite de orientandos/docente; distribuição entre docentes, projetos e linhas de pesquisa

Uma parcela significativa da avaliação é em cima de produção de artigos (docentes e discentes)

Qualis

A1	A2	B1	B2	B3	B4	B5	C
1.00	0.85	0.70	0.50	0.20	0.10	0.05	0.00

Índice Restrito e Irrestrito

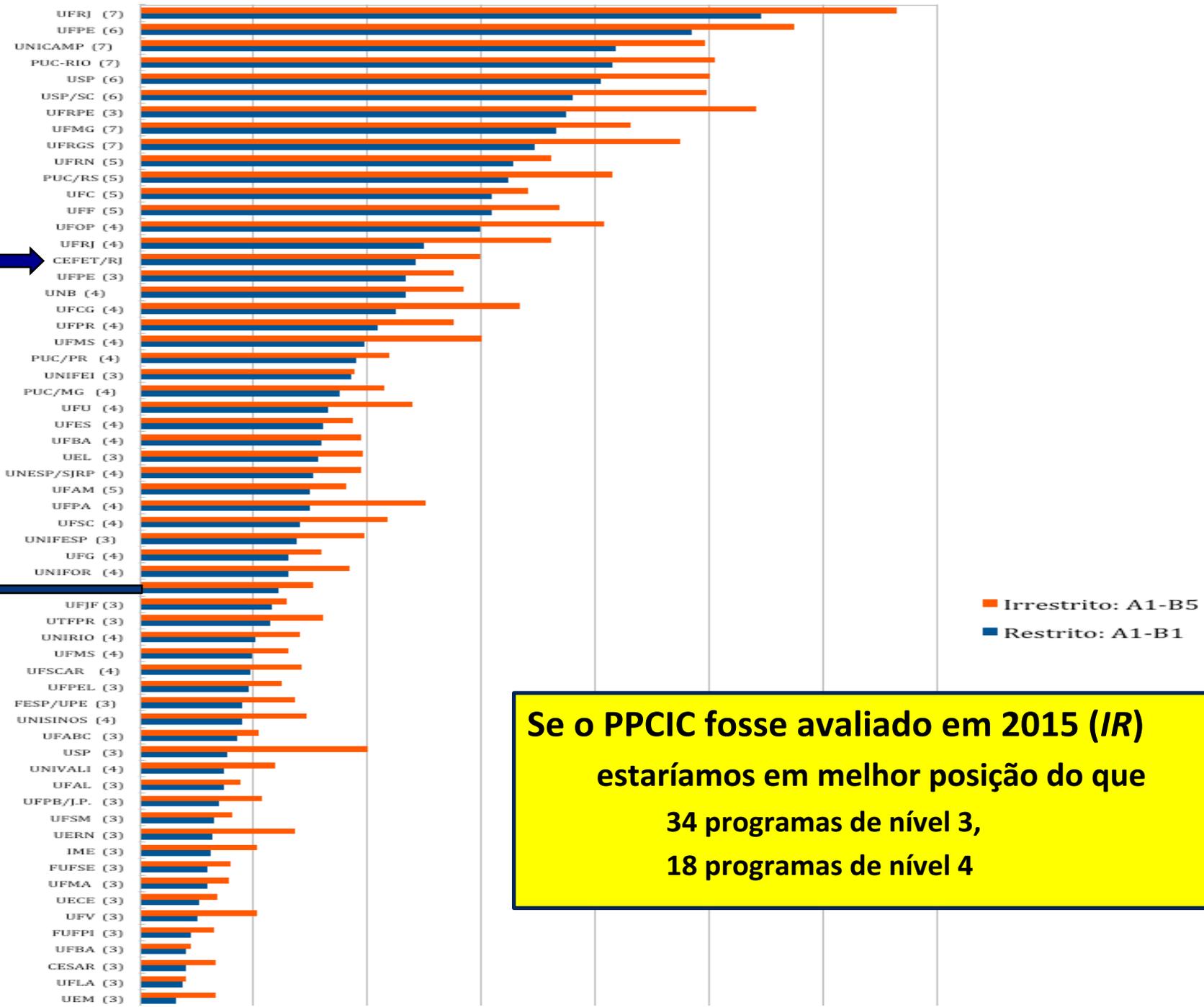
$$I_R = (A_1 + 0,85A_2 + 0,7B_1)/DP \quad (1)$$

$$I_I = (A_1 + 0,85A_2 + 0,7B_1 + 0,5B_2 + 0,2B_3 + 0,1B_4 + 0,05B_5)/DP \quad (2)$$

Produção do PPCIC em Periódicos

A1	A2	B1	B2	B3	B4	B5	C	SQ	I_R	I_I
5	11	2	5	6	0	0	3	12	1,58	1,95

P



Se o PPCIC fosse avaliado em 2015 (IR) estaríamos em melhor posição do que 34 programas de nível 3, 18 programas de nível 4

Programa de Pós-graduação em Ciência da Computação

Centro Federal de Educação Tecnológica Celso Suckow da Fonseca – CEFET/RJ

Início O Programa ▾ Corpo Docente ▾ Alunos ▾ Regr Regras e Procedimentos ▾ Notícias Contato

PPCIC

Bolsas

Dissertações

Formulários

Manual do aluno

Regras e Procedimentos ▾

Regulamentos do Programa

Calendário e Atas de Reuniões de Colegiado

Programa de Pós-graduação em Ciência da Computação

Centro Federal de Educação Tecnológica Celso Suckow da Fonseca – CEFET/RJ

Início O Programa ▾ Corpo Docente ▾ Alunos ▾ Regras e Procedimentos ▾ Seleção Notícias Contato

Formulários

- [Requerimento Geral](#)
- [Termo de Aceitação e Compromisso](#)
- [Prorrogação de Proposta](#)
- [Prorrogação de Defesa](#)
- [Pedido de Banca](#)
- [Marcação de Defesa e Termo de Responsabilidade](#)
- [Termo de Entrega da Dissertação](#)
- [Requerimento de Diploma](#)



**Centro Federal de Educação Tecnológica
Celso Suckow da Fonseca
CEFET/RJ**

**O Programa de Pós-Graduação em Ciência da
Computação (PPCIC) do CEFET/RJ e
sua ligação com os desafios em Ciência de Dados**

Eduardo Ogasawara

eogasawara@ieee.org

<http://eic.cefet-rj.br/~eogasawara>