



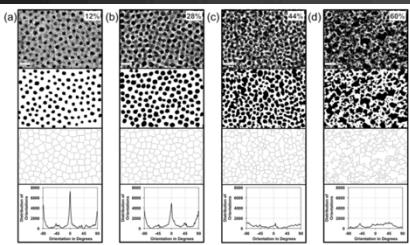
# Visão computacional na quantificação de imagens científicas

**Daniela Ushizima, Ph.D.**

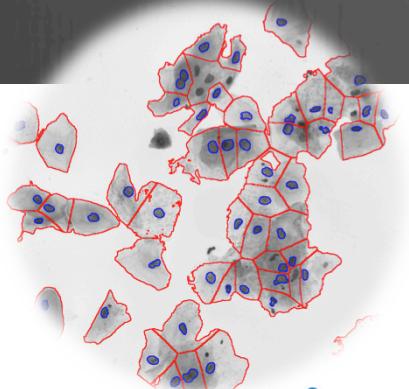
Staff Scientist - CRD, Lawrence Berkeley National Lab.  
Data Scientist - BIDS, University of California, Berkeley  
P.V.E. - Universidade Federal do Ceará



# Polímeros



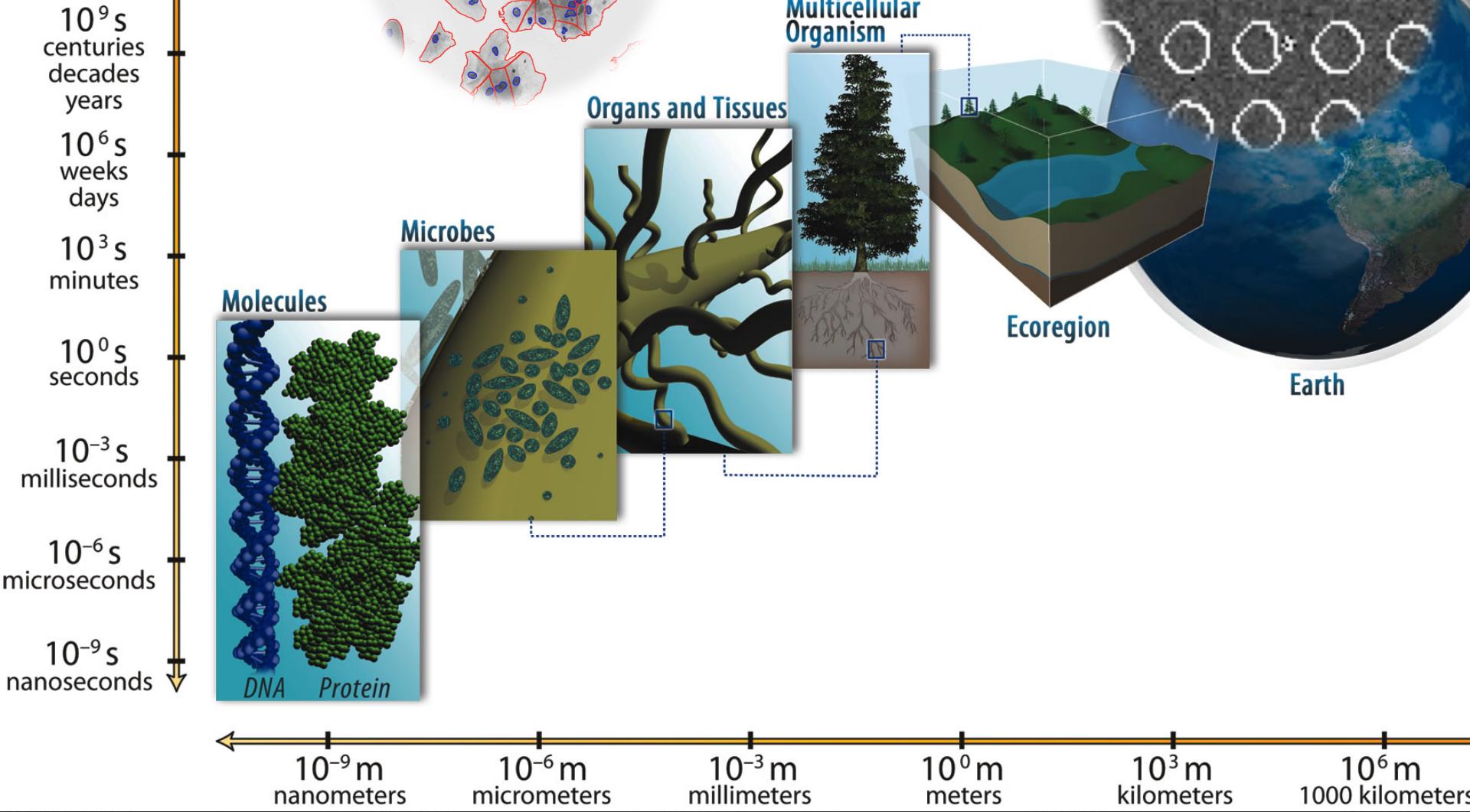
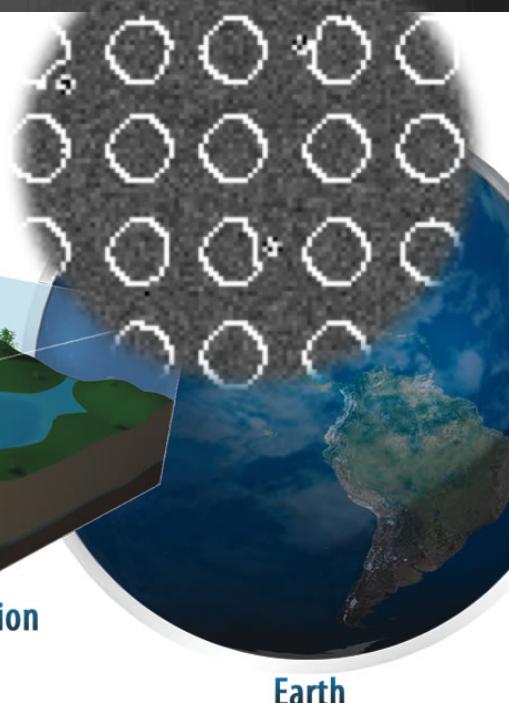
# Células do cervix



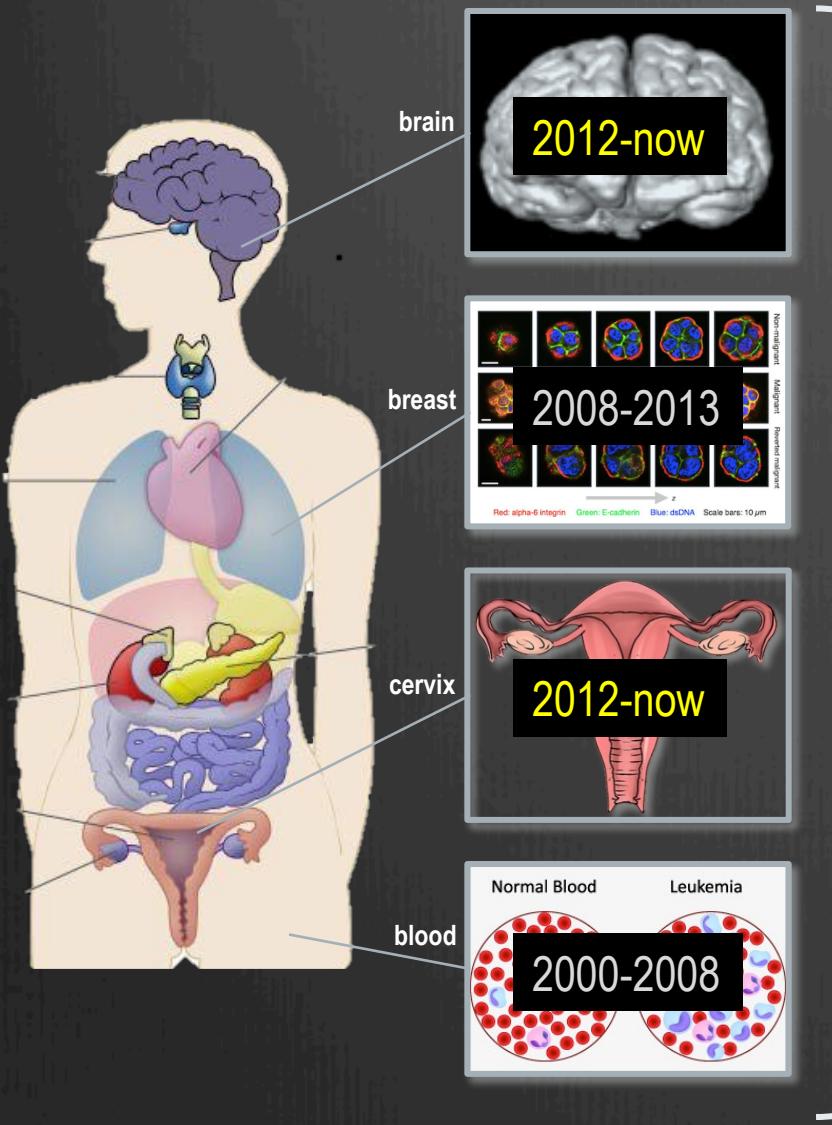
# Pílulas



# Sensoreamento remoto

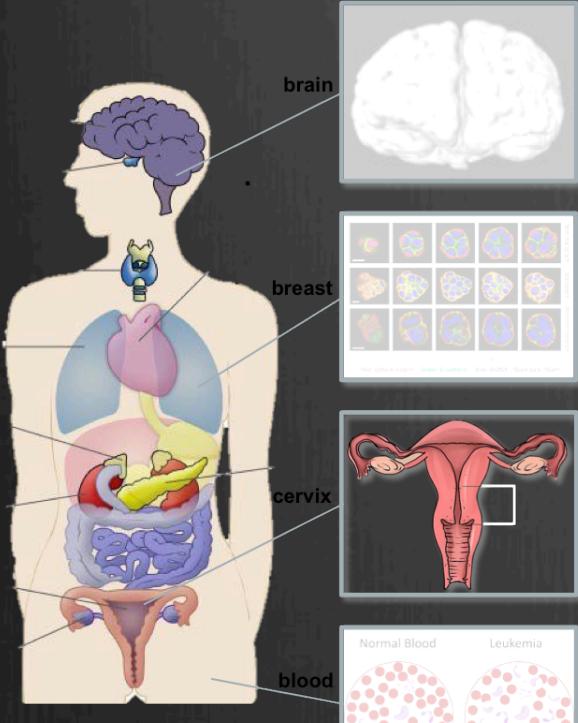


# Computational health science



- Problem:
  - Quantification from images;
- Approach:
  - Bio: detect and track biomarkers associated to the progression of disease;
  - Math: schemes for image representation, segmentation, characterization, visualization, ML;
- Impact:
  - Software tools to improve measurement and reproducibility;
  - Potential to support development of new drugs/treatment that target individuals more precisely.

# Câncer de colo de útero



- Problema:
  - Cancer of the cervix/uteri = **4th most common cancer among women worldwide**; estimated 527,624 new cases and 265,672 deaths\*;
  - Curable if detected in earlier stages;
- Abordagem:
  - Bio: **standardization** of data collection;
  - Math: **cell classification**;

- Impact:
  - Software for Paps prescreening;
  - Cell descriptors and classifiers with applications to HCA.

# Sumário

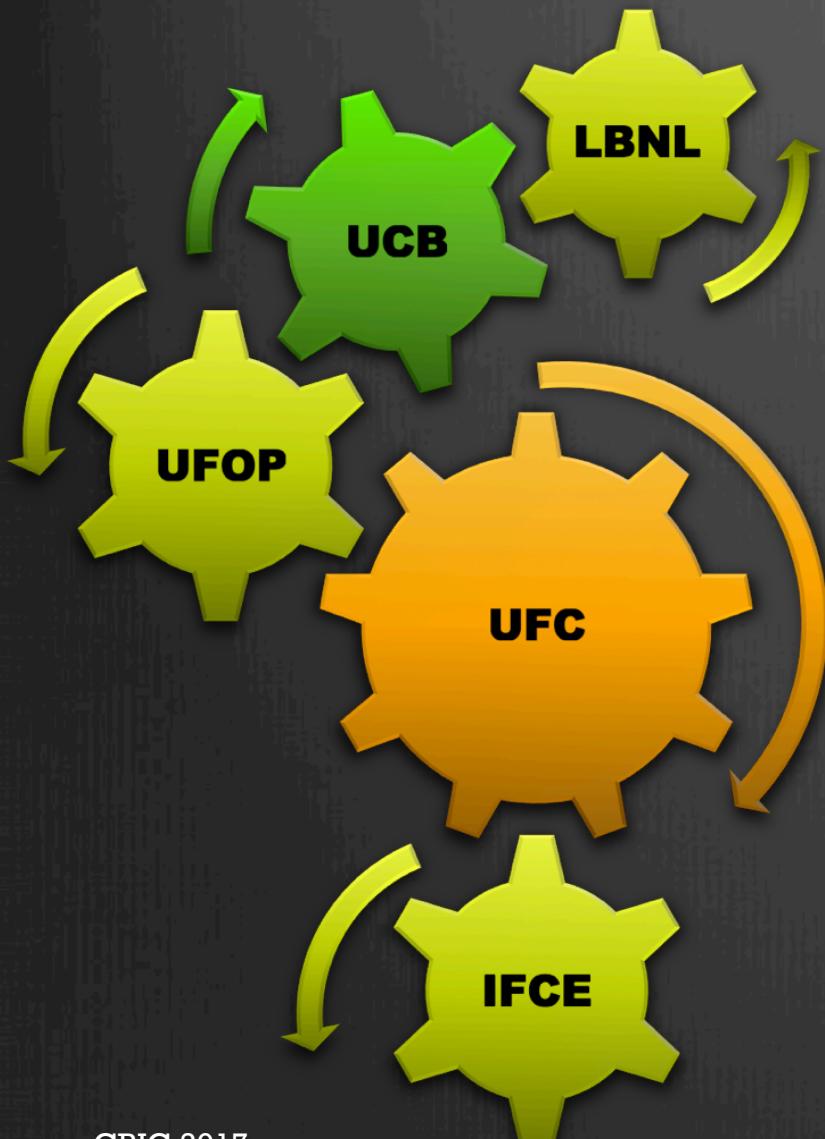
1. Projeto pesquisa em Visão Computacional via Ciências sem Fronteiras ;
2. Sub-projeto: câncer de colo do útero usando imagens do SUS;
3. Pesquisa, desenvolvimento e extensão;
4. Novos caminhos.



# Ciências sem Fronteiras

UFC, UC Berkeley and UFOP

# Instituições

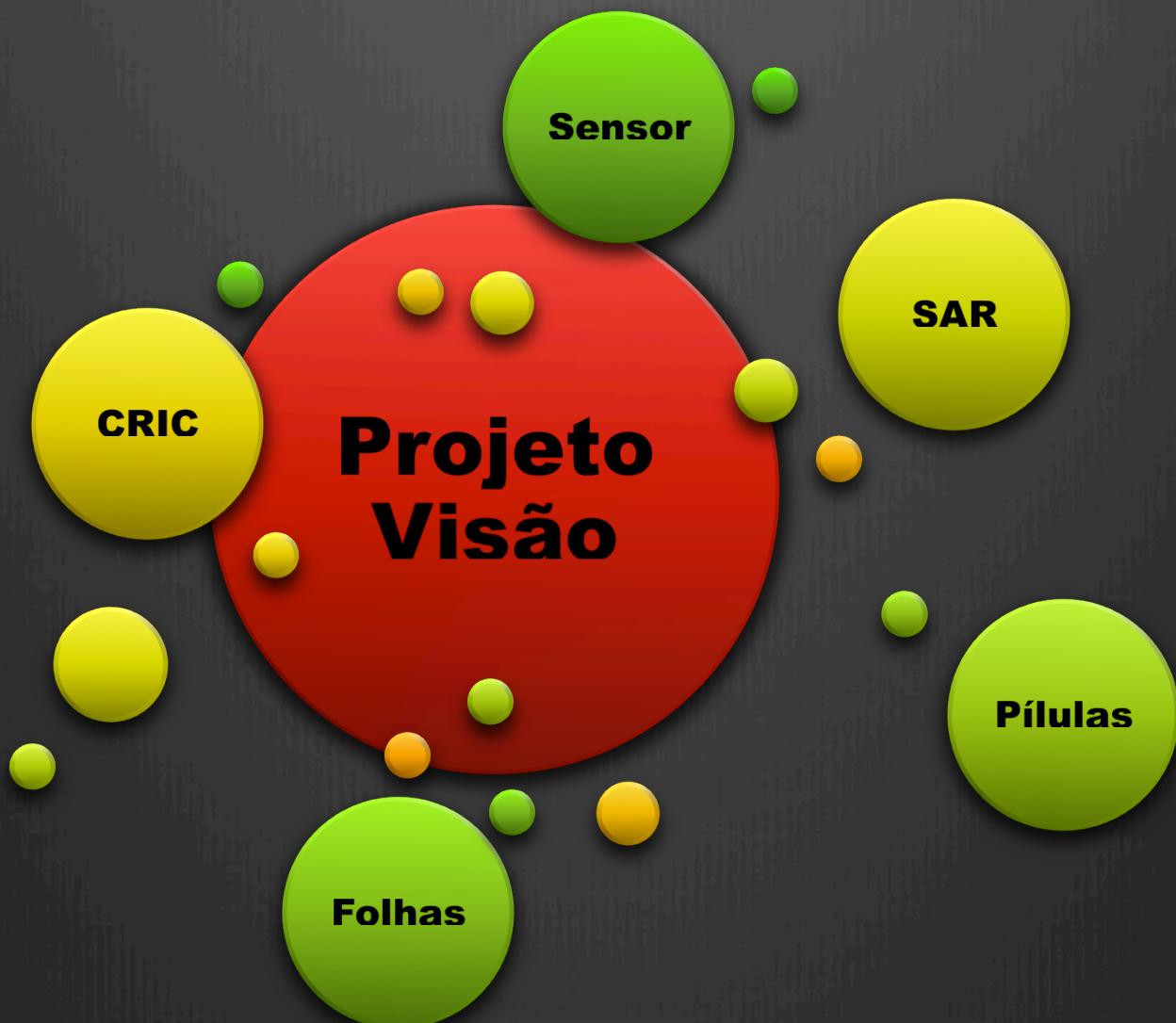


- UFC: Profa. Fatima Medeiros
  - Coordenadora/Eng.
- UCB/LBNL: Dani Ushizima
  - PVE
- UFOP: Profa. Andrea Bianchi
  - Especialista/Computação
- UFOP: Profa. Claudia Carneiro
  - Especialista/Citologia

# Pessoal

- Câncer de colo do útero 14
- Imagens de Radar 9
- Ciências de Materiais 7
- Pílulas / medicamentos 5
- Algoritmos 21

# Aplicações



# Cronologia

- 2012: análise de imagens microscópicas biomédicas: transferência de conhecimento em materiais?
- 2013: pesquisa sem fomento - investigar células cervicais = final de ano preparando para competição de programas de computadores para detecção de células;
- 04/2014: Resultado: 1º. lugar competição do International Symposium on Biomedical Imaging (ISBI 2014) - avaliação códigos-fonte, precisão, rapidez - publicação de artigo;
- 2014 - junho: projeto Bill & Melinda Foundation - negado;
- 2014-2015: elaboração novo projeto a Ciências sem Fronteiras via UFC, liderado por Prof. Medeiros - aprovado;

# Cronologia

- 2015: UC Berkeley (Moore & Sloan Foundation) - data science fellowship;
- 2016: projeto Microsoft Azure Machine Learning Research para escalar algoritmos usando nuvem - aprovado;
- 2015-2017: intercâmbio de alunos, formação de pessoal, treinamento, publicação de códigos para automatização de análise celular.
- 2015: Congresso Fortaleza (UFC): CRIC I e hackathon (UFOP remoto);
- 2016: Congresso Ouro Preto (UFOP): Workshop UFOP, CRIC II (sprint em São Paulo e visita a UFOP);
- 2017: Congresso Fortaleza (UFC): ERIPI, CRIC III e hackathon;



# Análise de imagens microscópicas

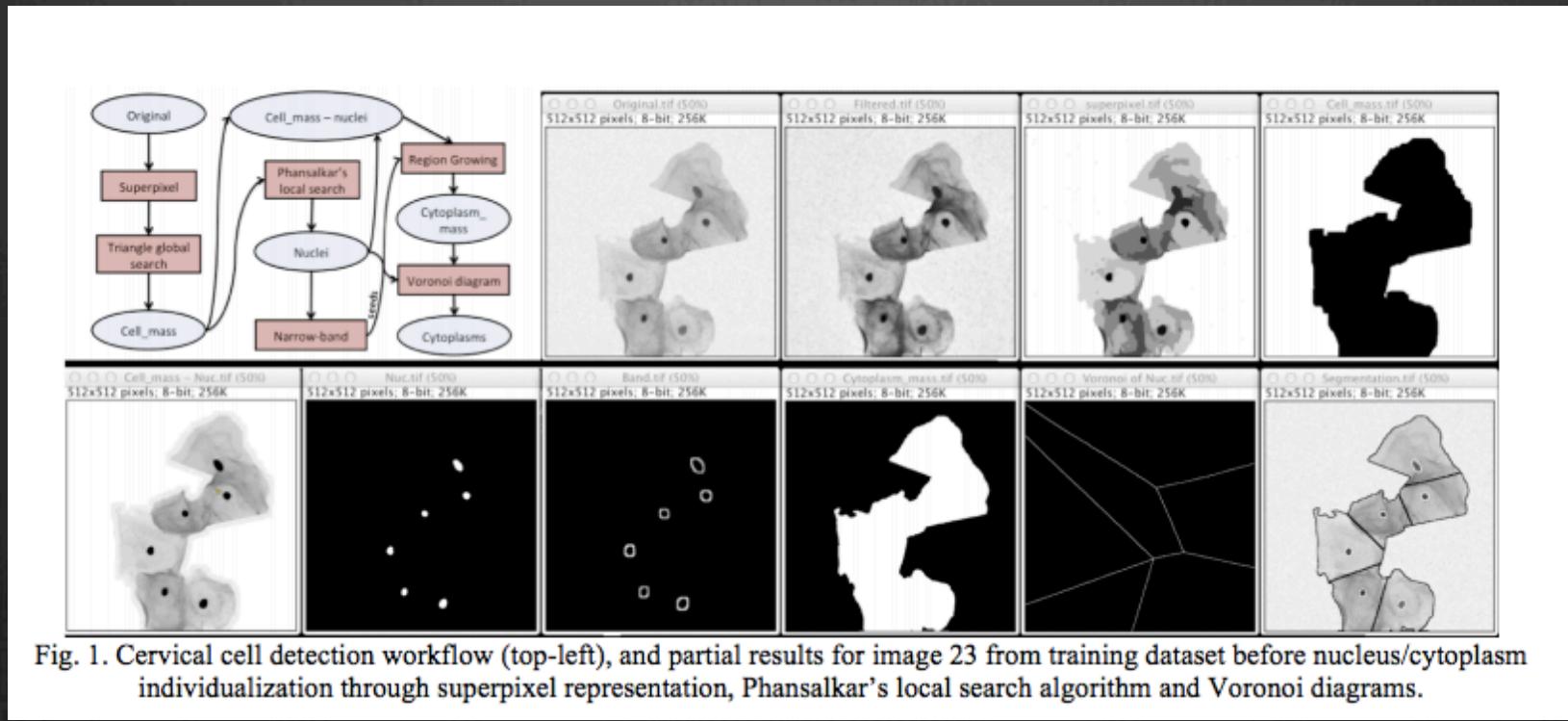
Pequenos objetos, grandes desafios e empenho gigantesco

# Junho 2013 - UFOP



Postdoc, sensores & PPSUS

# Estopim do Ciências sem Fronteiras



## SEGMENTATION OF SUBCELLULAR COMPARTMENTS COMBINING SUPERPIXEL REPRESENTATION WITH VORONOI DIAGRAMS

Daniela M. Ushizima, Andrea G. C. Bianchi and Claudia M. Carneiro

Lawrence Berkeley National Laboratory, Berkeley, CA, USA

2Federal University of Ouro Preto, Ouro Preto, MG, Brazil

# 1º lugar: IEEE ISBI 2014

 BERKELEY LAB

BERKELEY LAB COMPUTING SCIENCES  
LAWRENCE BERKELEY NATIONAL LABORATORY

U.S. DEPARTMENT OF ENERGY

A-Z INDEX | PHONE BOOK | CAREERS |  SHARE |  FOLLOW

Home About News & Media Seminars Careers Awards Safety

search...

Home » News & Media » News » Adapting Materials Sciences Algorithms for Cancer Screening Leads to Award for Berkeley Lab Researcher

**NEWS & MEDIA**

News CS In the News InTheLoop Recent Newsletters

 Facebook  Google  Twitter

## Adapting Materials Sciences Algorithms for Cancer Screening Leads to Award for Berkeley Lab Researcher

SEPTEMBER 15, 2014  
Contact: Linda Vu, +1 510 495 2402, [lvu@lbl.gov](mailto:lvu@lbl.gov)

In the 1950s, the Papanicolaou (Pap) smear test—where a sample of cervical cells is examined under a microscope to detect cellular abnormalities—was introduced as a standard in cervical cancer screening. Since then, cervical cancer deaths in the United States have declined by 70 percent.

But the task of identifying abnormal cells isn't exactly easy. Because much of the analysis relies on people manually examining slides, there are sample variabilities and disagreements among observers that can lead to false negatives. Experts suspect that automated slide analysis could reduce the cost, time and false negative rate, but an effective automation of the conventional Pap smear screening process has yet to be achieved. Earlier this year, several researchers issued a challenge in conjunction with the [IEEE International Symposium on Biomedical Imaging \(ISBI 2014\)](#) in [Beijing, China](#) for tools to extract the boundaries of individual [cytoplasm](#) and nucleus from overlapping cervical cell images.



The winning team (left to right): Prof. Andrea Bianchi (Physics Dept., UFOP), Prof. Fátima Medeiros (Teleinformatics Dept, UFC), Prof. Claudia Carneiro (Pathology Div, UFOP), Dr. Daniela Ushizima (CRD, Berkeley Lab).

# Junho 2014 - UFOP



Idealização de projeto de algoritmos com aplicação em análise de amostras diversas via imagens

# Junho 2015 - UFC



CRIC 2015 – seminários e tutoriais em visão computacional para integração dos membros do Projeto Ciências sem Fronteiras

# Junho 2015 - UFC



Treinamento em fundamentos em Data Science gratuito com instrutores da Software Carpentry (democracia digital): controle de versão em git, shell scripting e estatística em R

# Junho 2015 - UFOP



**CRIC HACKATHON:** usar técnicas recém aprendidas  
baseadas em software livre para criar novos  
algoritmos para detecção de células cervicais

# Dezembro 2015 - UFOP



Investigação focada em base de dados e análise de células  
provenientes do cólo do útero

# Maio 2016 - UFOP



Investigação de algoritmos com aplicação em análise de amostras  
diversas via imagens de microscópio

# Maio 2016 - UFOP



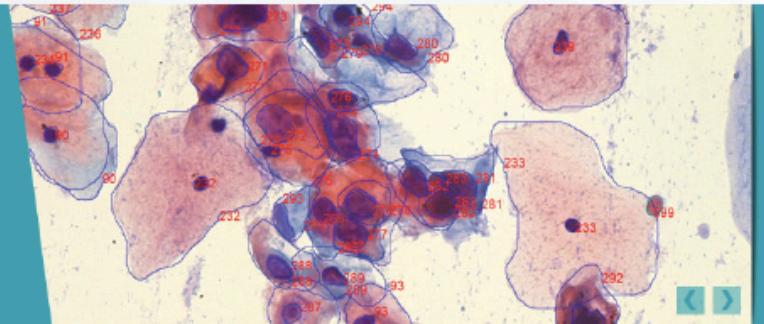
*Investigação de algoritmos com aplicação em análise de amostras  
diversas via imagens de microscópio*

# Junho 2016 - UFOP



CELL RECOGNITION  
FOR INSPECTION OF CERVIX

*Detecção  
de células cervicais  
de microscopia  
de Papanicolau*



**As células do colo do útero Pesquisável  
de dados de imagens.**

Um dos produtos deste trabalho de colaboração é fazer uma base de dados disponível para a comunidade científica, a fim de estimular o interesse em suporte de análise de imagem câncer cervical usando exames de Papanicolau.

[Quer uma demonstração?](#)



**Apoiadores**



**CRIC database - Profa. Andrea Bianchi, Paulo Calaes, Profa. Claudia Carneiro, Alessandra Gomes, Mariana Rezende**

# Junho 2016 - UFOP



CELL RECOGNITION  
FOR INSPECTION OF CERVIX

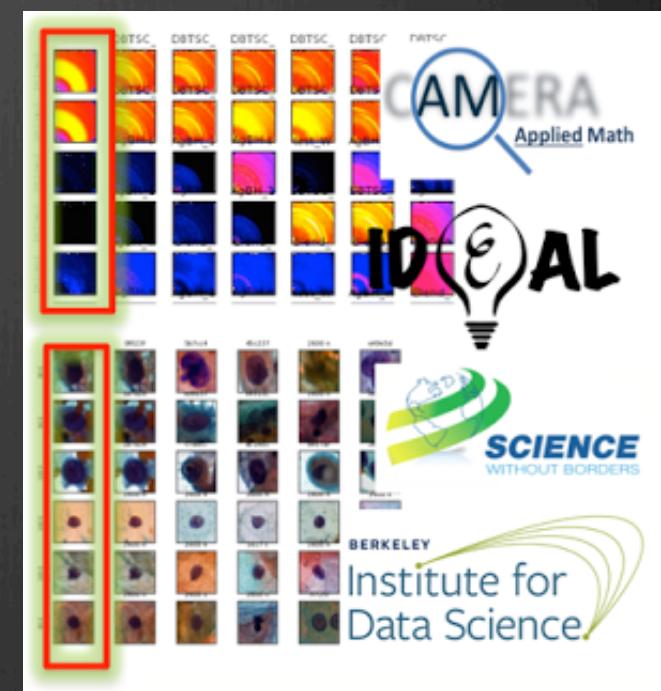


CRIC database – Profa. Andrea Bianchi, Paulo Calaes, Profa. Claudia Carneiro, Alessandra Gomes, Mariana Rezende, Cida e Renata

# Agosto 2016 - UCB



A screenshot of a web-based application for searching datasets. At the top, there are navigation links: Home, About, Segmentation Database, Classification Database, Download, Segments Labels, Publications, Contact, and Welcome Admin. Below this is a section titled "Images" with a sub-section "Search". A search bar contains the text "Laser". To the right of the search bar are buttons for "+ Add image" and "Clear". Below the search bar is a list of categories: Normal - 1029, ASC-US - 148, LSIL - 172, ASC-H - 89, HSIL - 606, and Cervine - 55. A grid of small thumbnail images is displayed, each with a green "View details" button below it. The thumbnails are labeled 1-ASC-H, 2-HSIL, 3-ASC-US, 4-ASC-HSIL, 5-ASC-H, and 6-HSIL.



PyData San Francisco 2016 – searchable datasets using pyCBIR

# Search and ranking cells by similarity

pyCBIR

Feature Extraction Methods

- Gray-Level Co-Occurrence Matrix
- Histogram of Oriented Gradient
- First Order Texture Feature
- Local Binary Pattern
- Convolutional Neural Network
- Daisy Features

Similarity Metrics

- Euclidian Distance
- Infinity Distance
- Cossine Similarity
- Pearson Correlation
- Chi-Square Dissimilarity
- Kullback-Leibler Divergence
- Jeffrey Divergence
- Kolmogorov-Smirnov Divergence
- Cramer Divergence
- Earth Movers Distance

Searching Methods

- Brute Force
- R Tree
- KD Tree
- LSH

Retrieval Options

Retrieval Number: 10

By CSV File

Load Database File

Load Retrieval File

By Path

Load Database Path

here/Dropbox/CBIR/cells2/databases/

Load Retrieval Path

here/Dropbox/CBIR/cells2/retrieval/

Run pyCBIR

Retrieval Result



100.0 class 0.0 [GT class: 0]

100.0 class 0.0 [GT class: 0]

90.0 class 0.0 [GT class: 0]

100.0 class 0.0 [GT class: 0]

90.0 class 0.0 [GT class: 0]

100.0 class 0.0 [GT class: 0]

100.0 class 0.0 [GT class: 0]

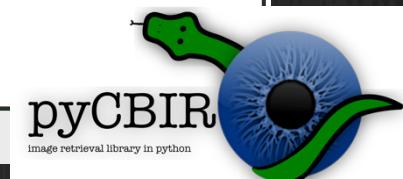
90.0 class 0.0 [GT class: 0]

100.0 class 0.0 [GT class: 0]

100.0 class 1.0 [GT class: 1]

80.0 class 1.0 [GT class: 1]

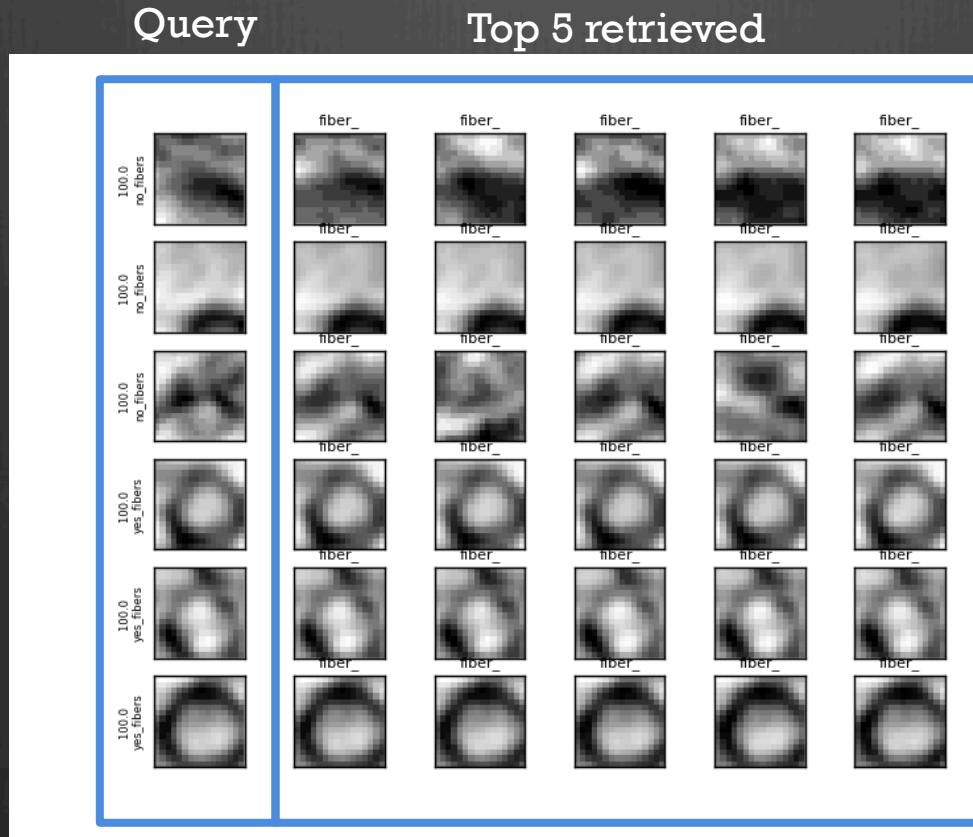
100.0 class 1.0 [GT class: 1]



pyCBIR  
image retrieval library in python

# Experiments - Fibers dataset

4,000 images of 16 X 16 for two balanced classes.



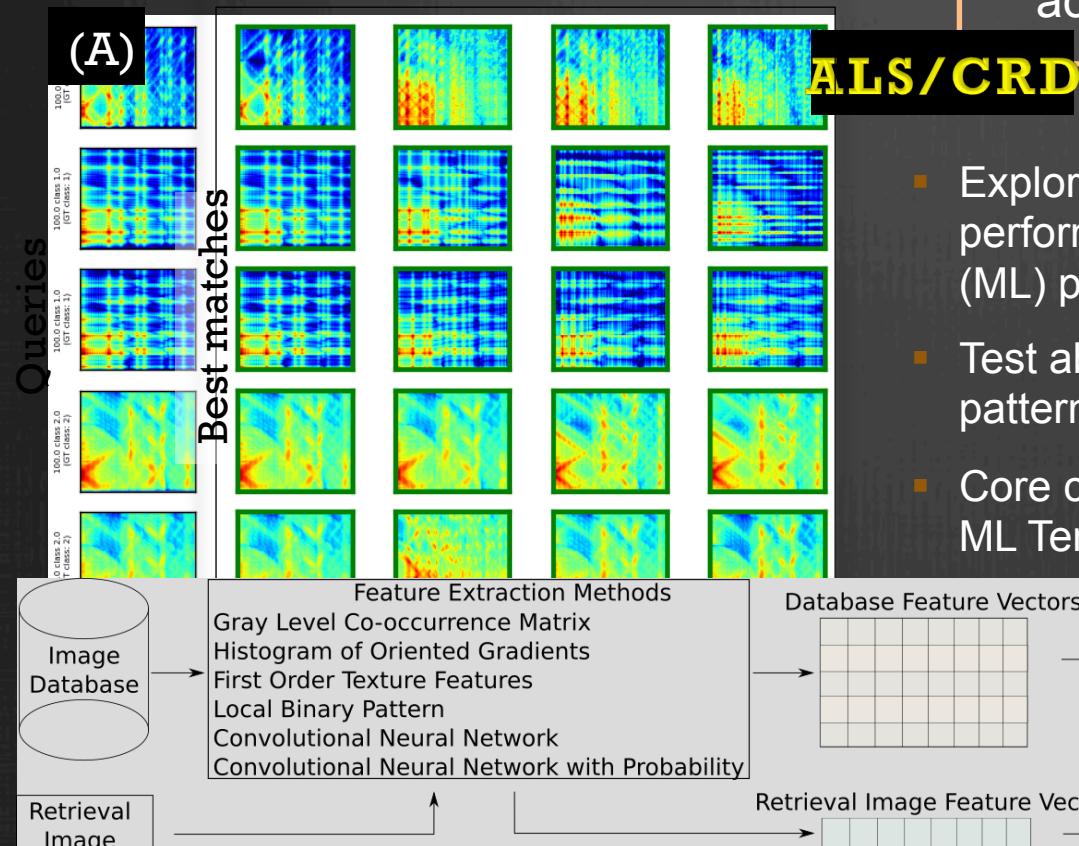
Result obtained using the inception network.



# Deep learning for X-ray diffraction and materials

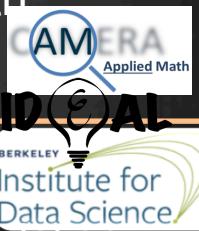
## Scientific Achievement

- Search and ranking materials using X-ray diffraction data (Fig. A);
- Recognize crystal structure from images without feature design;



## Significance and impact

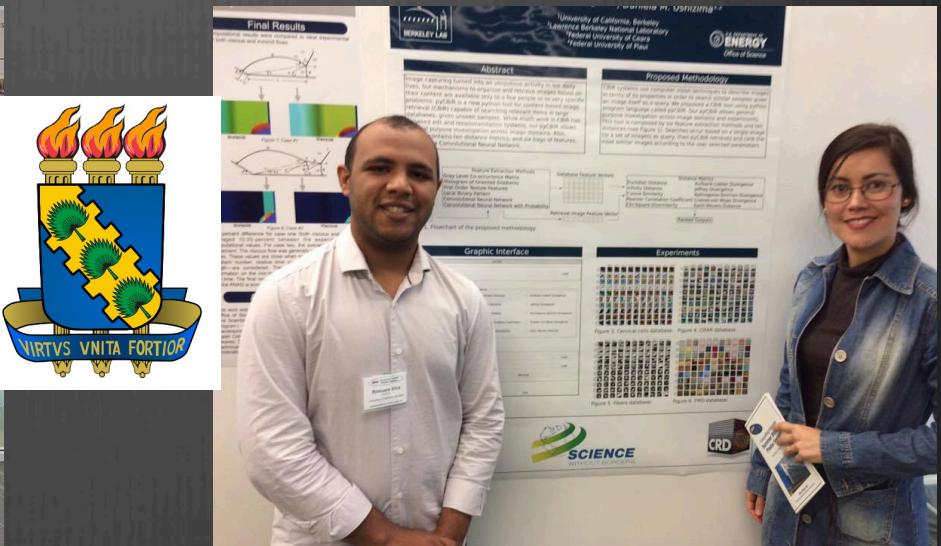
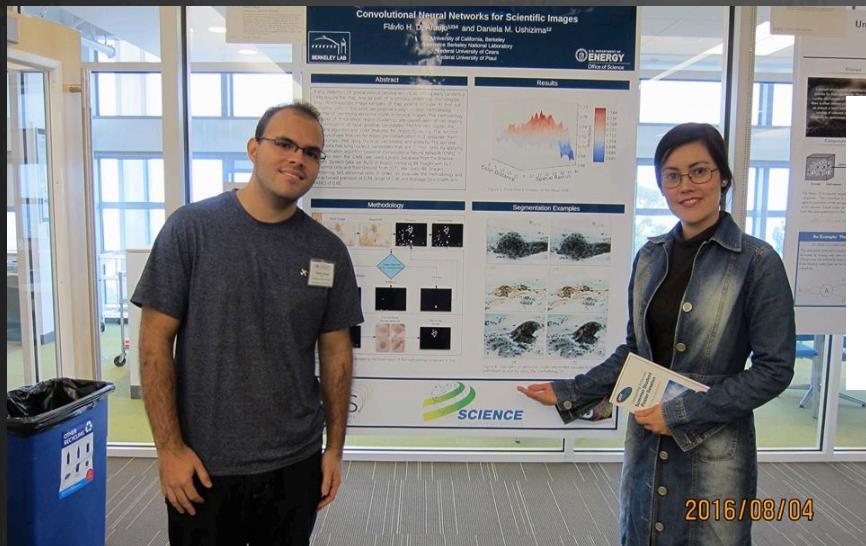
- Devise algorithms to bridge the gap between theoretical models and experimental observational data;
- X-ray diffraction prototypes with accuracy above 98%;



## Research details

- Exploration of ALS data, including simulation performed @NERSC using machine learning (ML) prototype (Fig.B) from CAMERA;
- Test ability to categorize millions of GISAXS patterns without manual interaction;
- Core codes: HipGISAXS simulation @NERSC, ML TensorFlow @CAMERA GPU test-bed.

# Agosto 2016 - LBNL



# Novembro 2016 - UFOP



CRIC database e preparativos para defesa de mestrado do Paulo Calaes

# June 2017 - UFPI



SBC ERIPI'17 prestigiando evento organizado por professores da UFPI e  
membros do CRIC

# Junho 2017 - Picos, PI



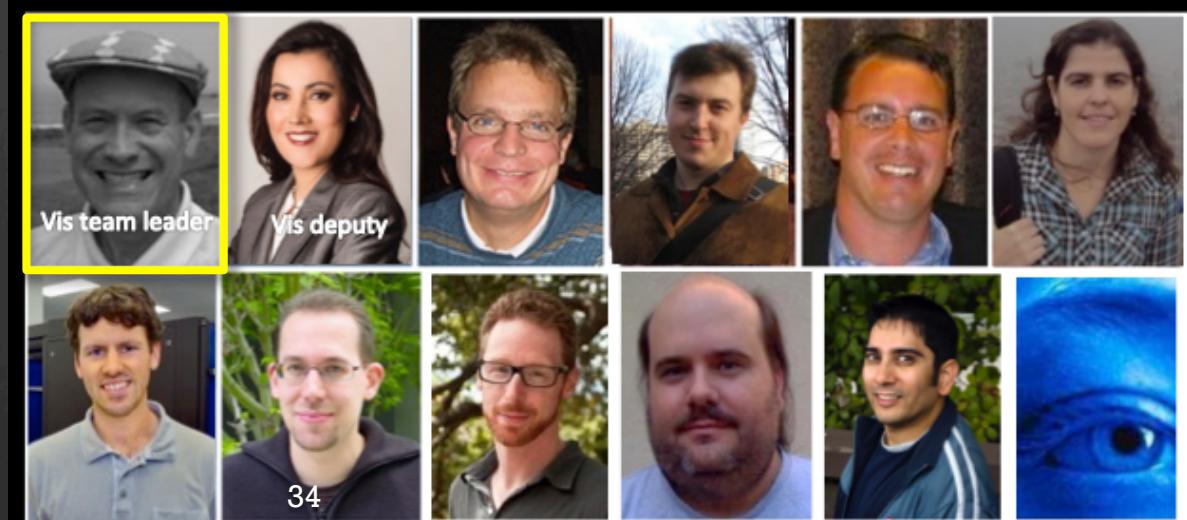
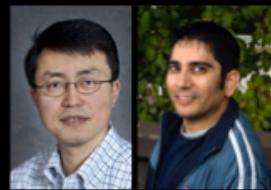
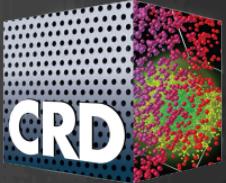
Trabalho de incentivo ao conhecimento computacional com crianças do  
CRIC 2017 do ensino fundamental em escola municipal

7/19/17

# Junho 2017 - UFC



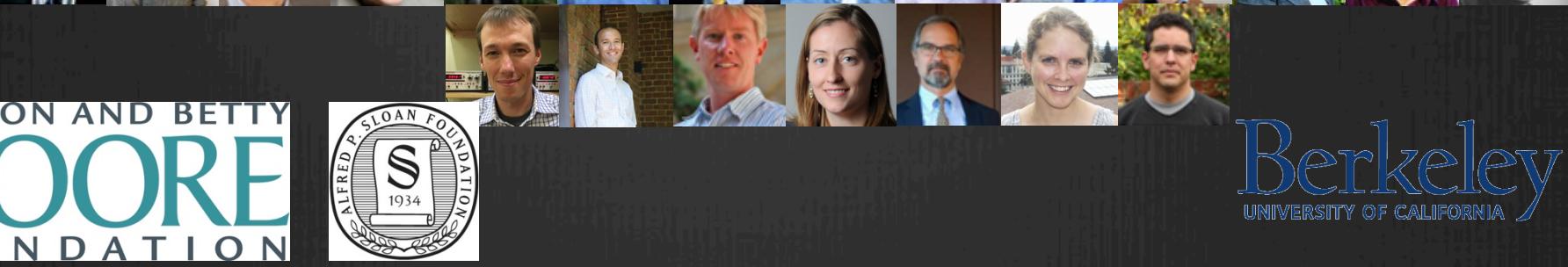
Organizando CRIC III, hackathons e outras surpresas



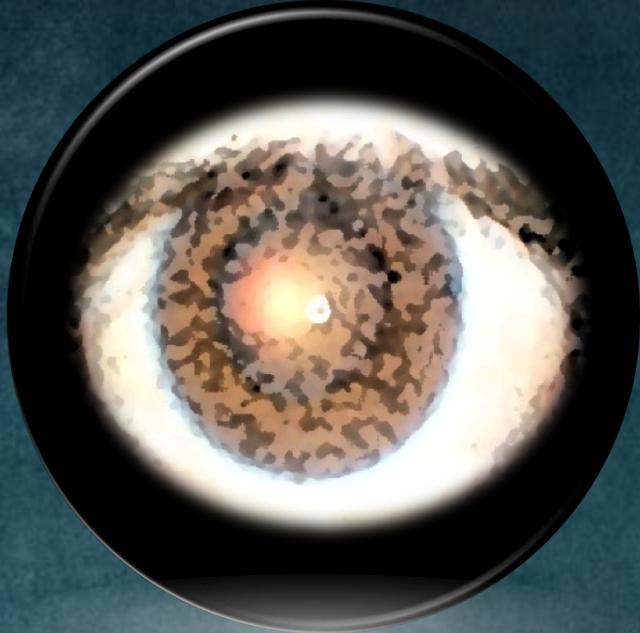
Berkeley, California



# Berkeley Institute for Data Science

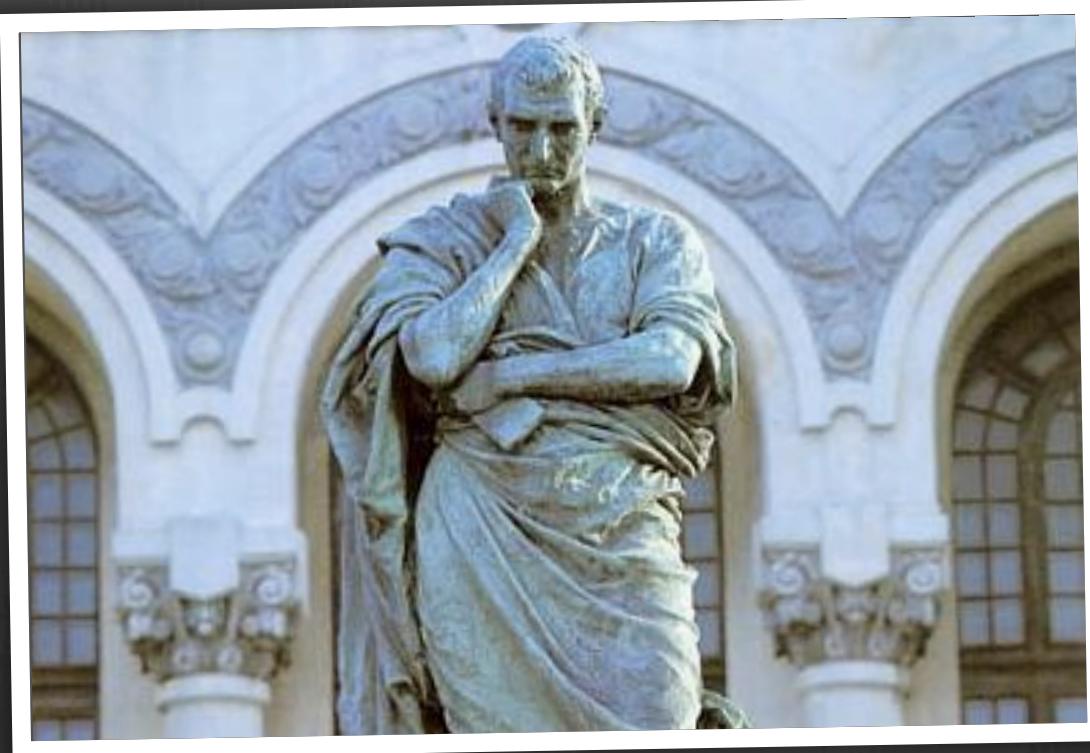


Berkeley  
UNIVERSITY OF CALIFORNIA



# Perguntas?

[http://vis.lbl.gov/~daniela  
dushizima@lbl.gov](http://vis.lbl.gov/~daniela.dushizima@lbl.gov)



**”Não se deseja o que não  
se conhece”**

Públio Ovídio Naso (Ovídio) - 43 a.C – 17 d.C.  
Pensador/Poeta Romano

# Fim