



I Workshop

Sinergia LNCC-empresas

A peça que faltava para a
inovação na sua empresa

www.sinergia.lncc.br

Modelagem e análise de redes complexas dinâmicas: Conectando conhecimento

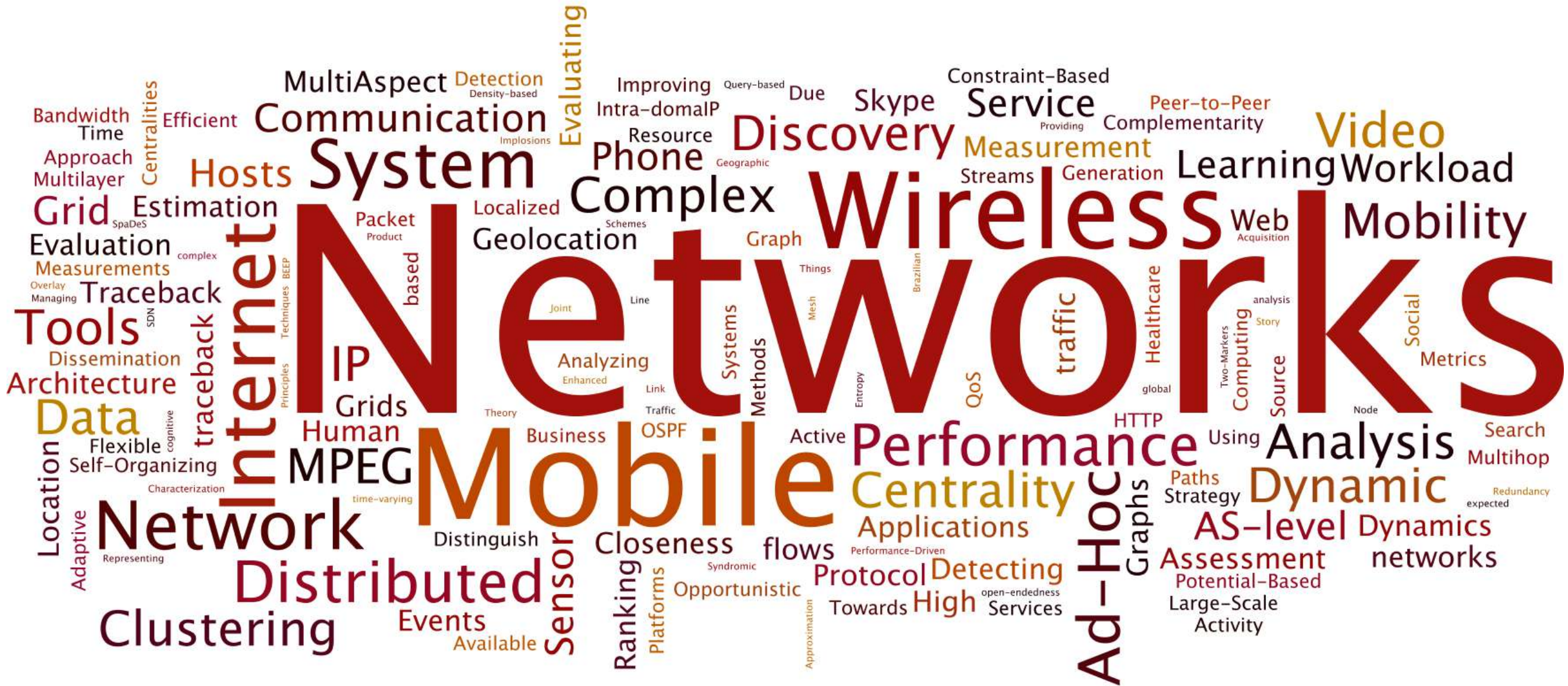
Artur Ziviani

Laboratório Nacional de Computação Científica (LNCC)

ziviani@lncc.br | <http://www.lncc.br/~ziviani>

LNCC, Petrópolis, RJ, 26 de junho de 2019.

Perfil de pesquisa



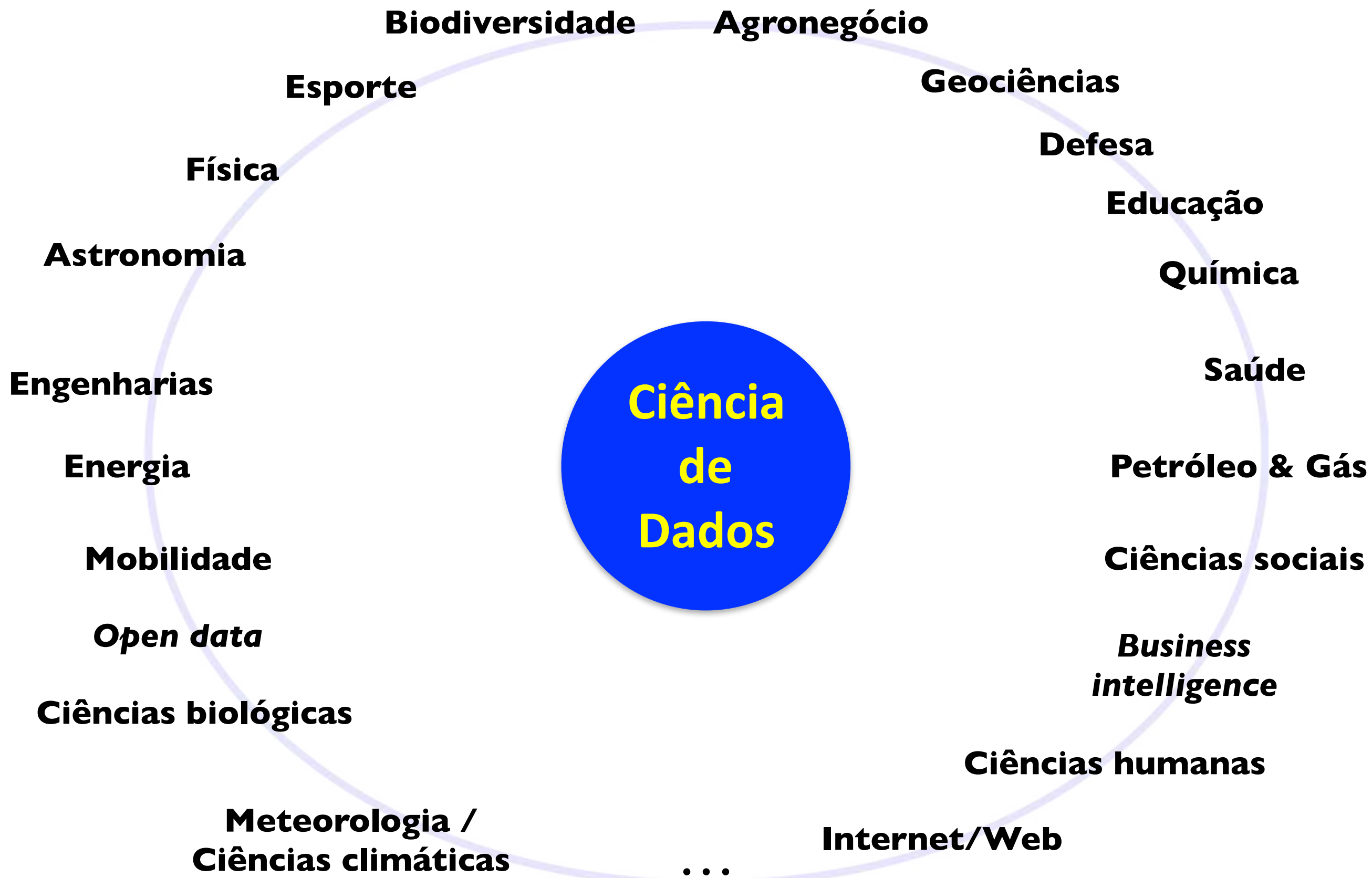
Worle: publicações 2004-2019

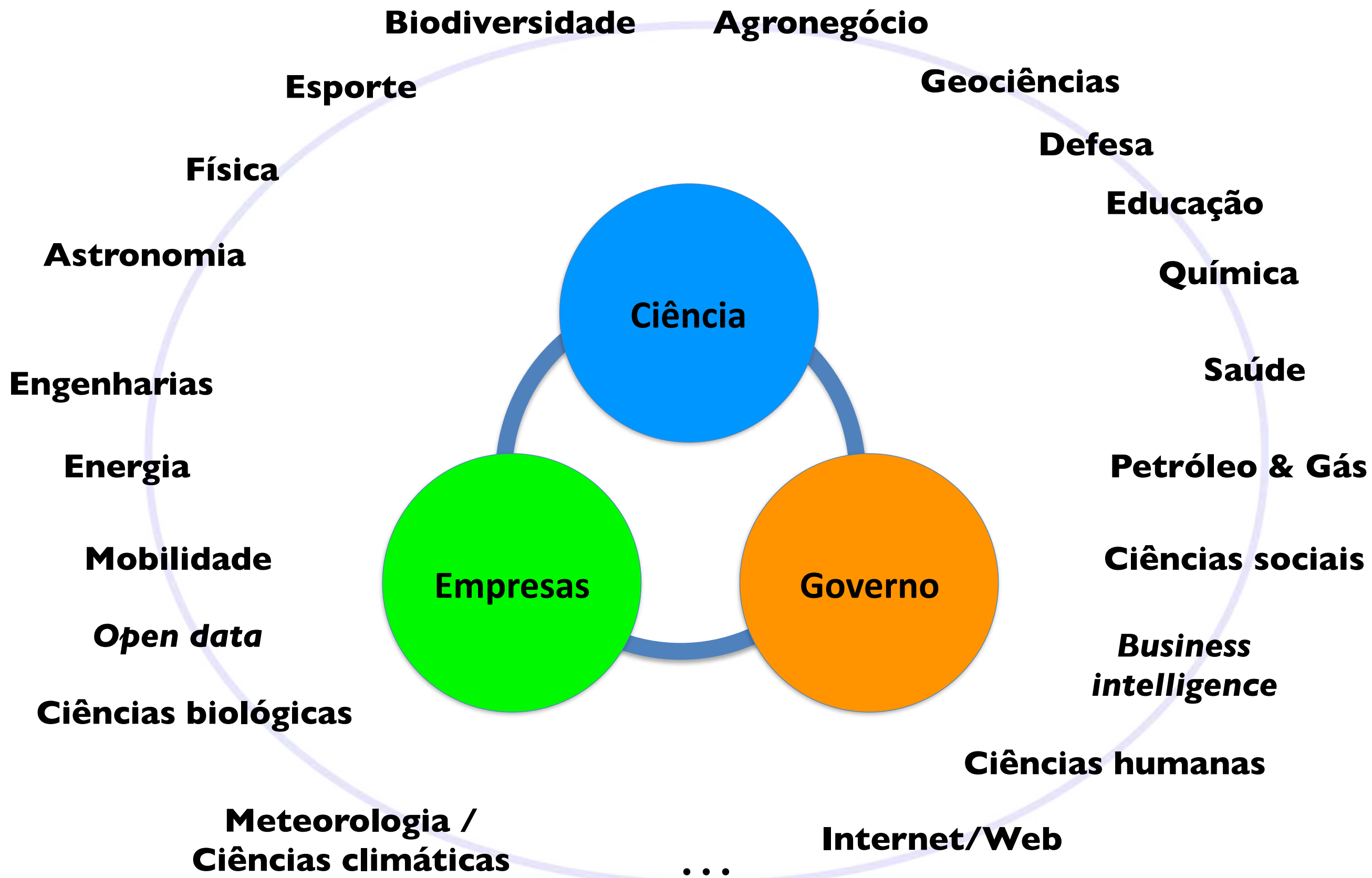
Co-autoria

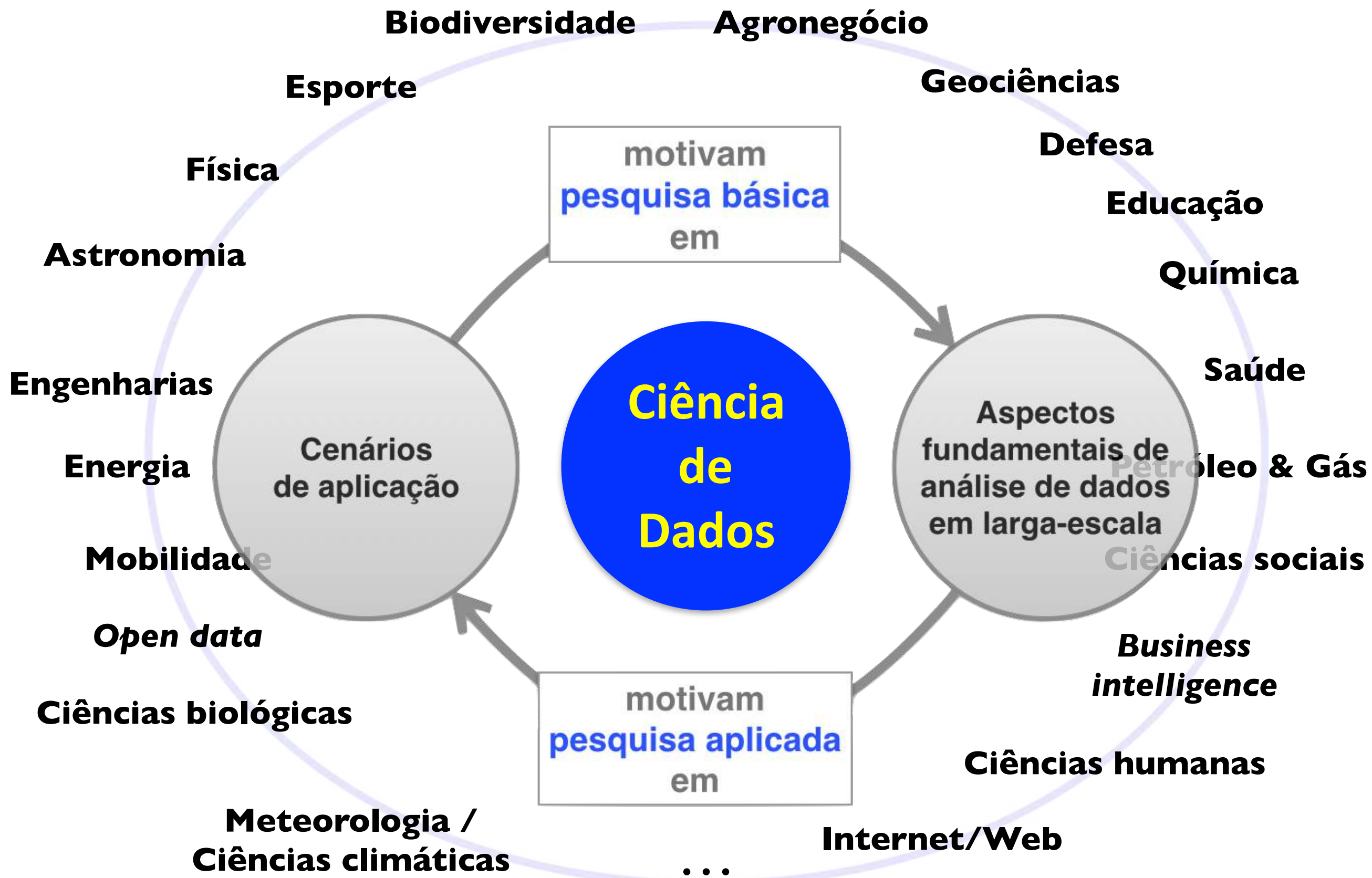
Acadêmicos nacionais: IME, LNCC, PUC-Minas, PUC-PR, PUC-Rio, UFBA, UFF, UFG, UFMG, UFJF, UFRJ, UNICAMP, UTFPR;

Acadêmicos internacionais: University of Hamburg, Fraunhofer FOKUS (Alemanha); UBA (Argentina); Boston University (EUA); ENS-Lyon, INRIA, LIP6/CNRS (França); Technion (Israel); SICS (Suécia), NCKU (Taiwan);

Indústria: Google Engineering, Yahoo! Research, Deutsche Telekom Labs, Oi Telecom, Grandata Labs, IBM Research, Cipher Trace

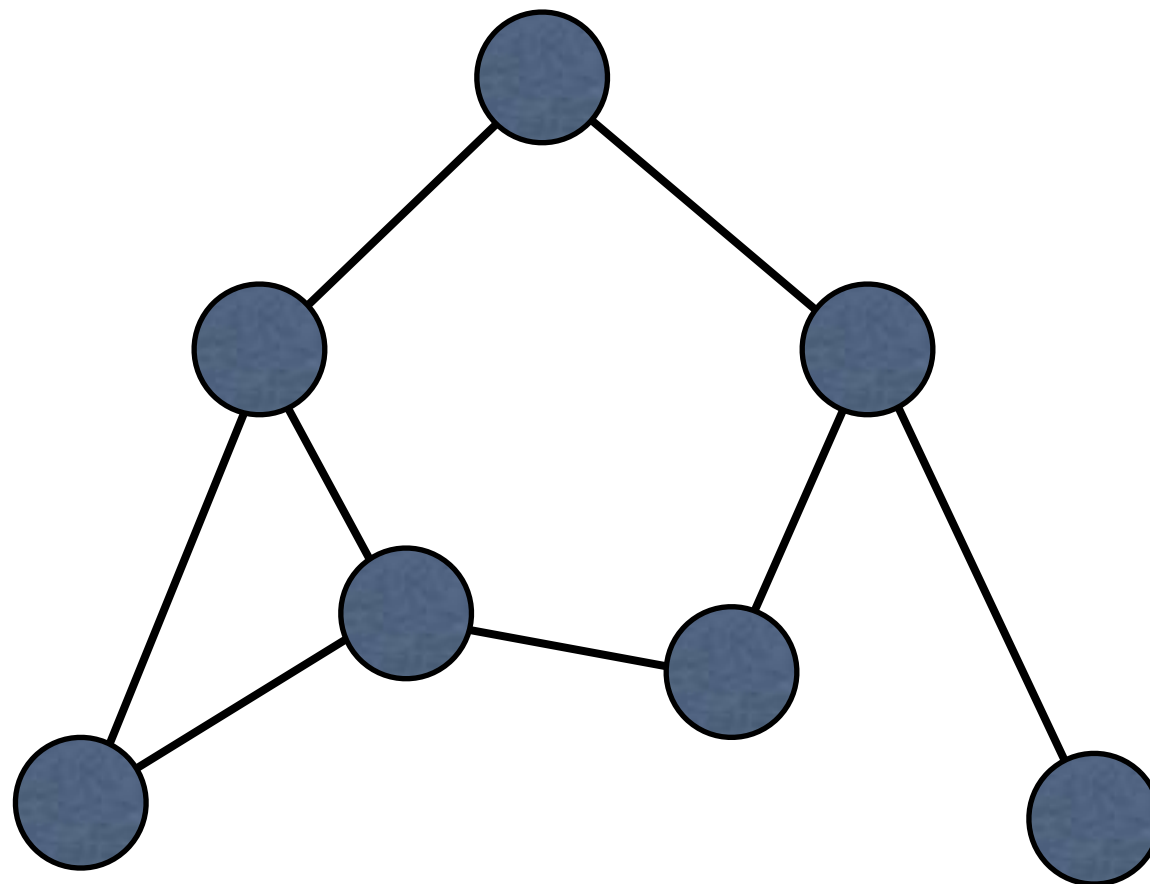


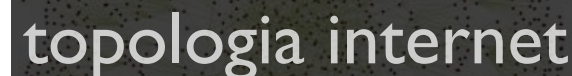
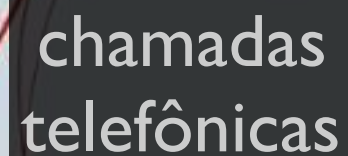
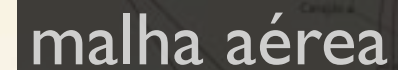
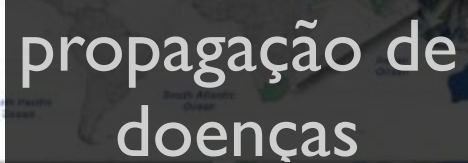




Redes complexas?

- Como “**coisas**” estão conectadas?
- Representadas matematicamente como **grafos**

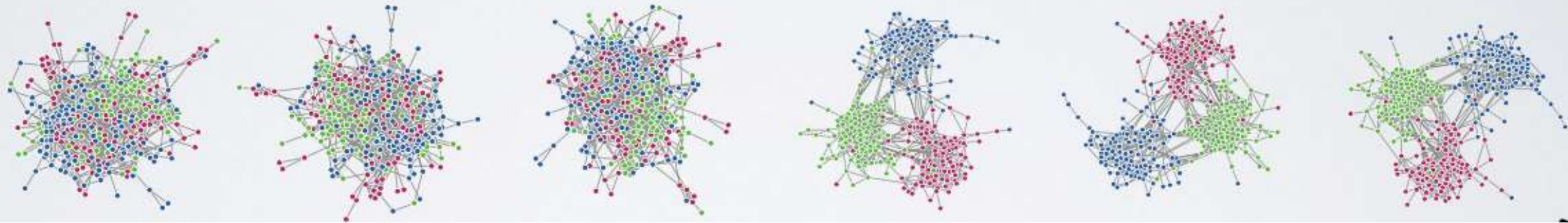




**etc,
etc,
...**

Redes complexas de alta ordem

Figure source: Clauset 2016

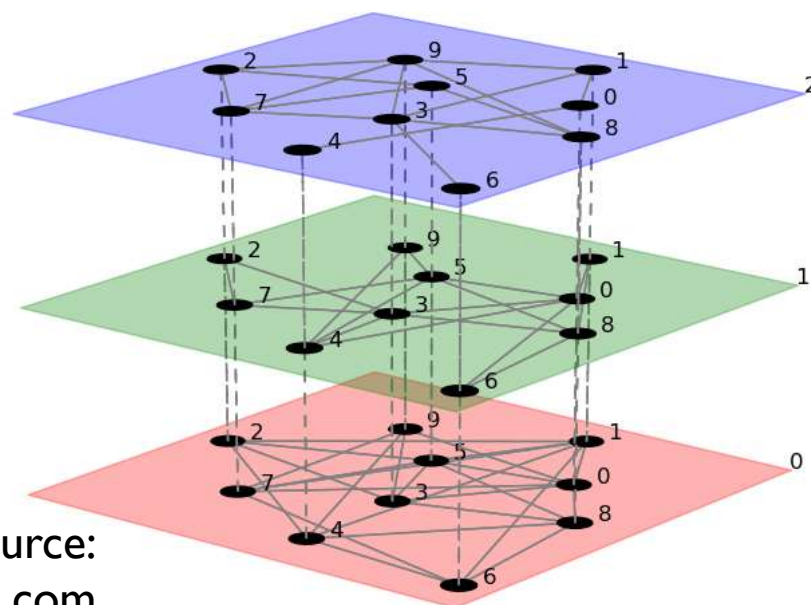


redes complexas dinâmicas

tempo

variam no tempo e podem ser representadas por

grafos variantes no tempo (TVGs)



sistemas complexos

podem incluir

múltiplas redes complexas

que podem ser representadas por

redes/grafos multicamadas

Como analisar redes complexas tão distintas?



Ciência de redes

- Redes = objetos conectados
- Redes descrevem **conexão e interação** entre objetos (dados)
- Redes podem ajudar a
 - **revelar padrões**
 - **detectar anomalias**
 - **comparar sistemas complexos**
 - **realizar análises preditivas**
 - **estudar difusão**



Modelagem e
análise de redes
complexas

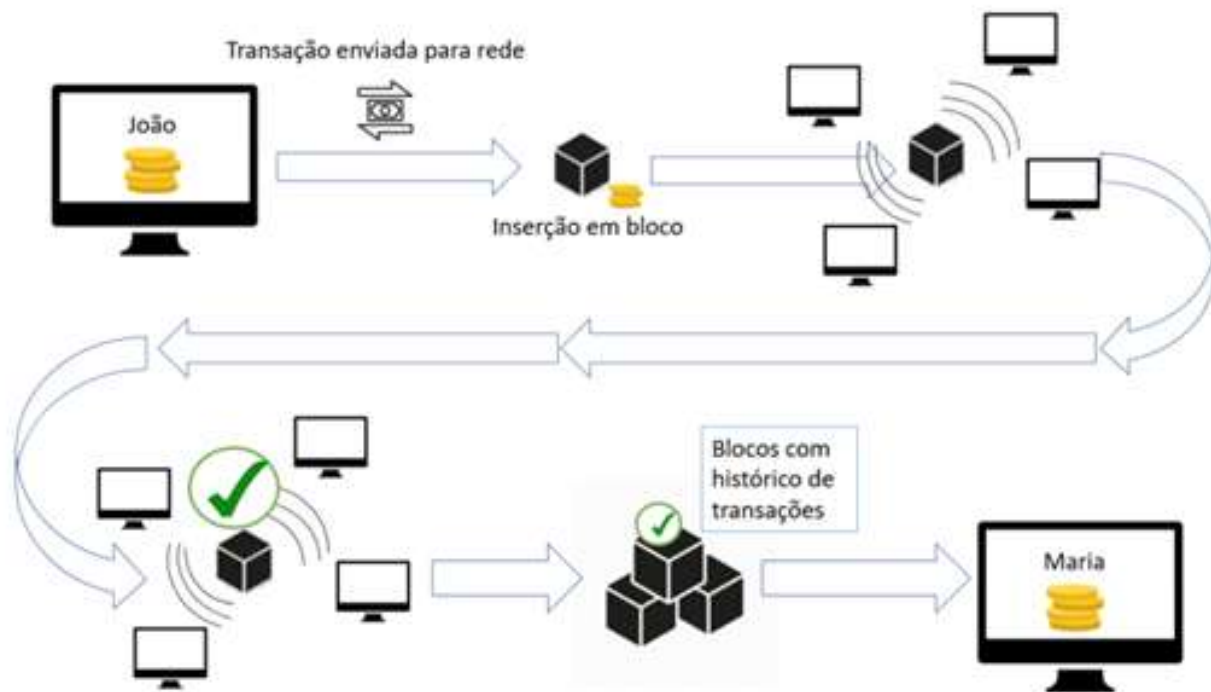
- **avaliar centralidade(s)**
- **detectar comunidades**
- **analisar fluxo entre objetos**
- **otimizar localização e alocação de serviços**
- **...**

Alguns casos em áreas diversas

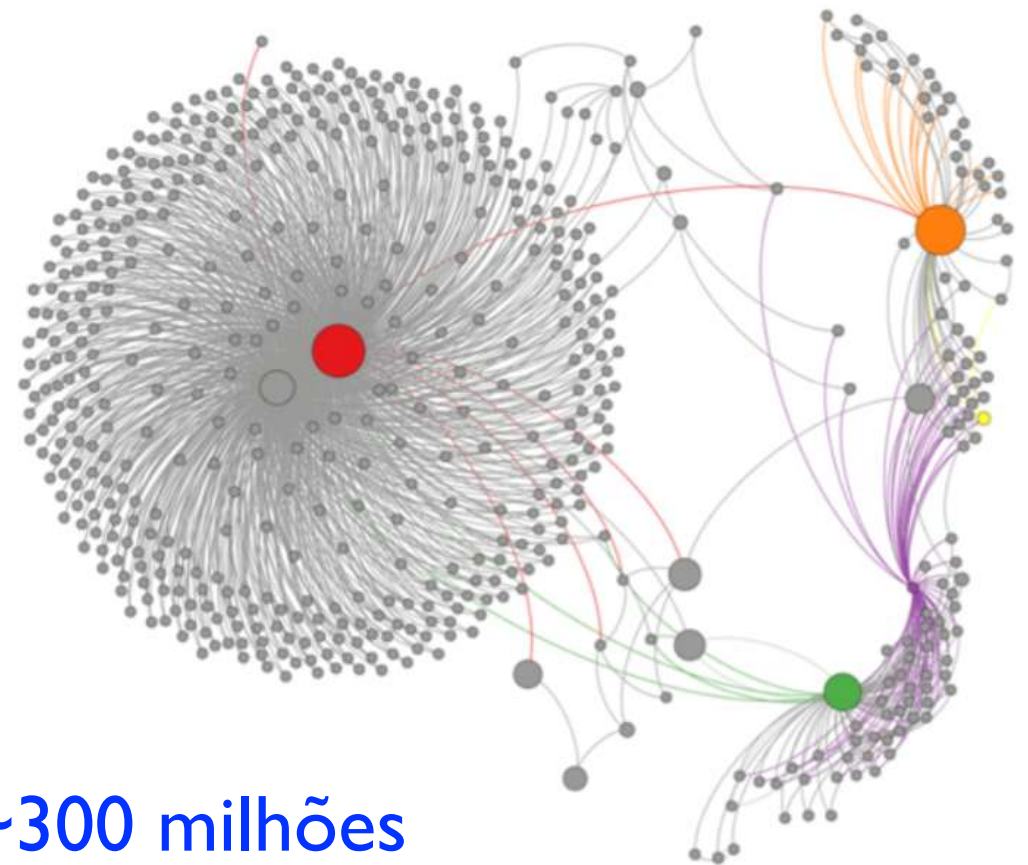
Internet

Rede de transações do Ethereum

criptomoeda ether baseada em blockchain



- Início em 2015; Em 2017, ~30% do mercado de criptomoedas face a ~40% do Bitcoin



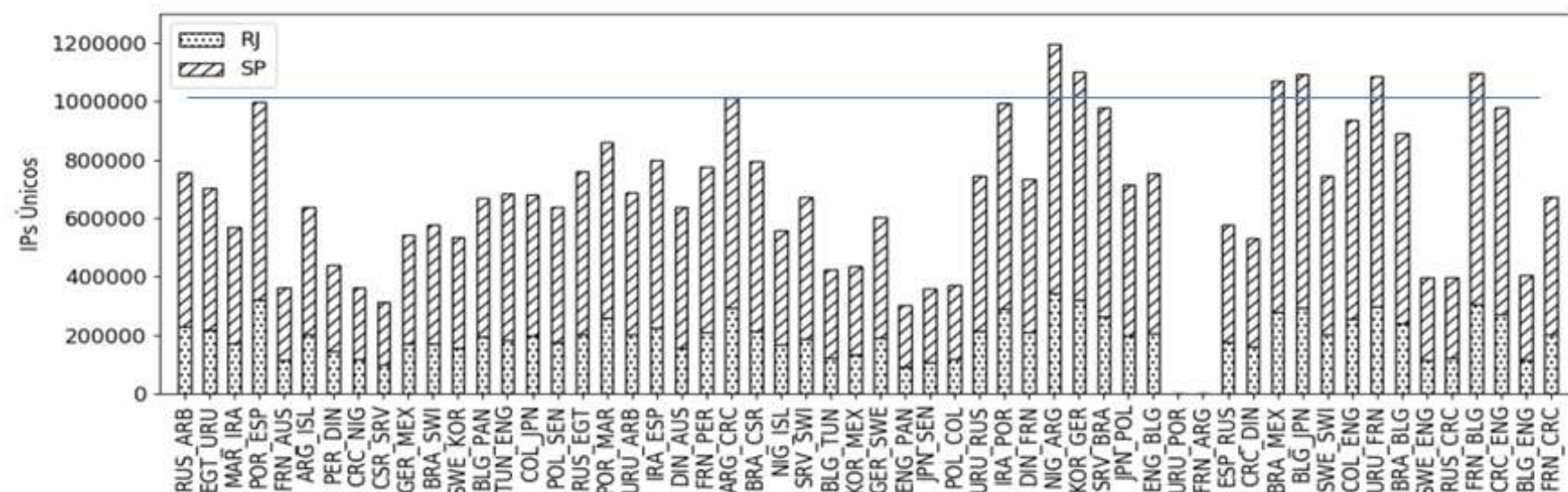
- Parceria LNCC e UFJF
- Rede de transações do Ethereum como grafo variante no tempo (2015-2018)
 - # endereços: ~40 milhões; # transações: ~300 milhões
- Análise da dinâmica da rede de transações; identificação de comunidades; detecção de padrões e comportamentos anômalos

Análise de transmissões de grandes eventos na Internet



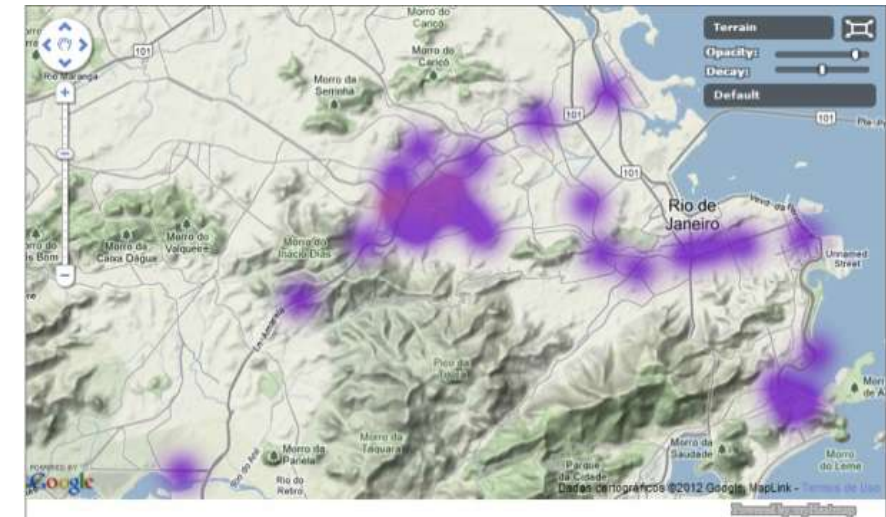
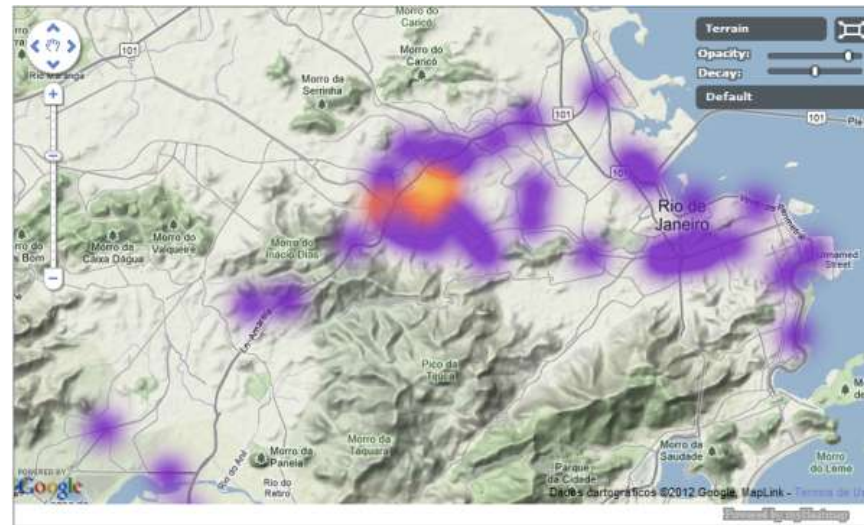
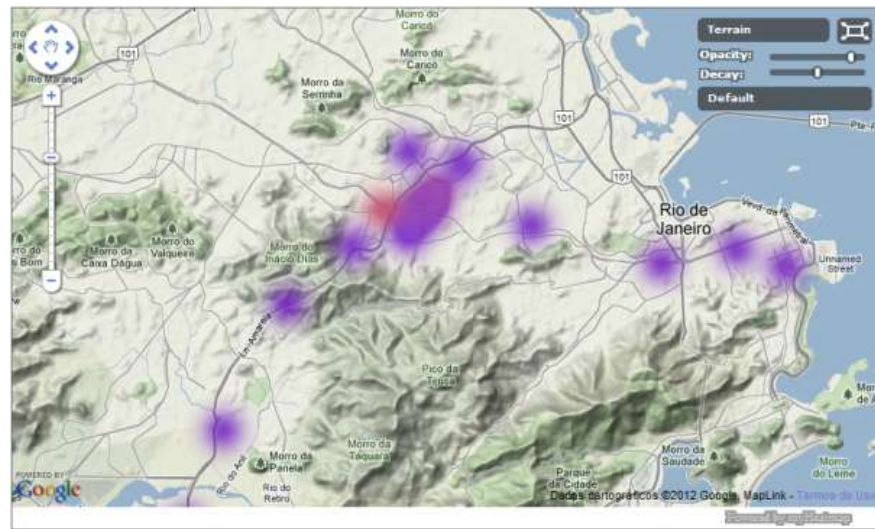
- Trabalho conjunto LNCC, UFJF e Globo.com
- 5 anos de transmissões de grandes eventos
- Acima de **1 milhão** de usuários simultâneos
- Até **1 bilhão** de requisições/evento
- Análises **espaço-temporais** e **modelo de comportamento dos usuários**

globo.com



Computação Urbana

Análise de mobilidade humana devido a eventos de larga-escala com base em dados de telefonia celular



Tipos de eventos de larga-escala:

Partidas de futebol, eventos esportivos, grandes espetáculos, Réveillon, Carnaval, ...

Projeto conjunto do LNCC, PUC-Minas, UFMG e Oi Telecom

H.T Marques-Neto, F. H. Z. Xavier, W. Z. Xavier, L. M. Silveira, J. M. Almeida, A. Ziviani, C. H. S. Malab,

Understanding Human Mobility and Workload Dynamics Due To Different Large-Scale Events Using Mobile Phone Data

Journal of Network and Systems Management (JONS), Springer, ISSN: 1064-7570, vol. 26, no. 4, pp. 1079–1100, October 2018.

Colaboração entre Brasil, França e Argentina

- *MOTIf: Mobile phone sensing of human dynamics in techno-social environments*
CAPES/STIC-AmSud (2018-2019)

- Parceiros:



LNCC, UFMG, PUC-Minas



INRIA, ENS-Lyon

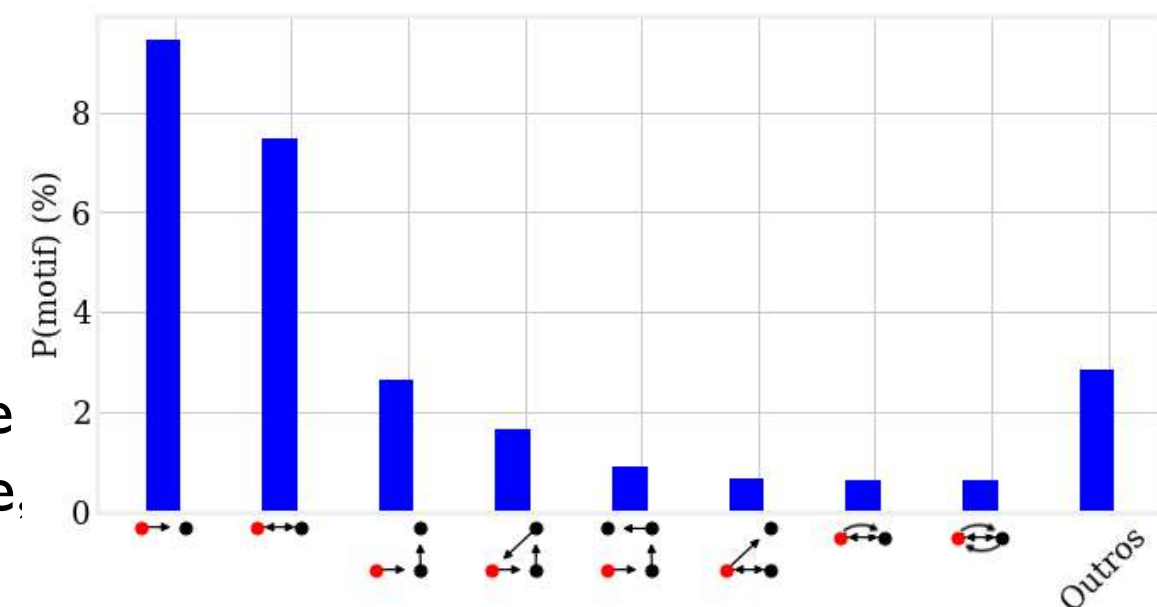
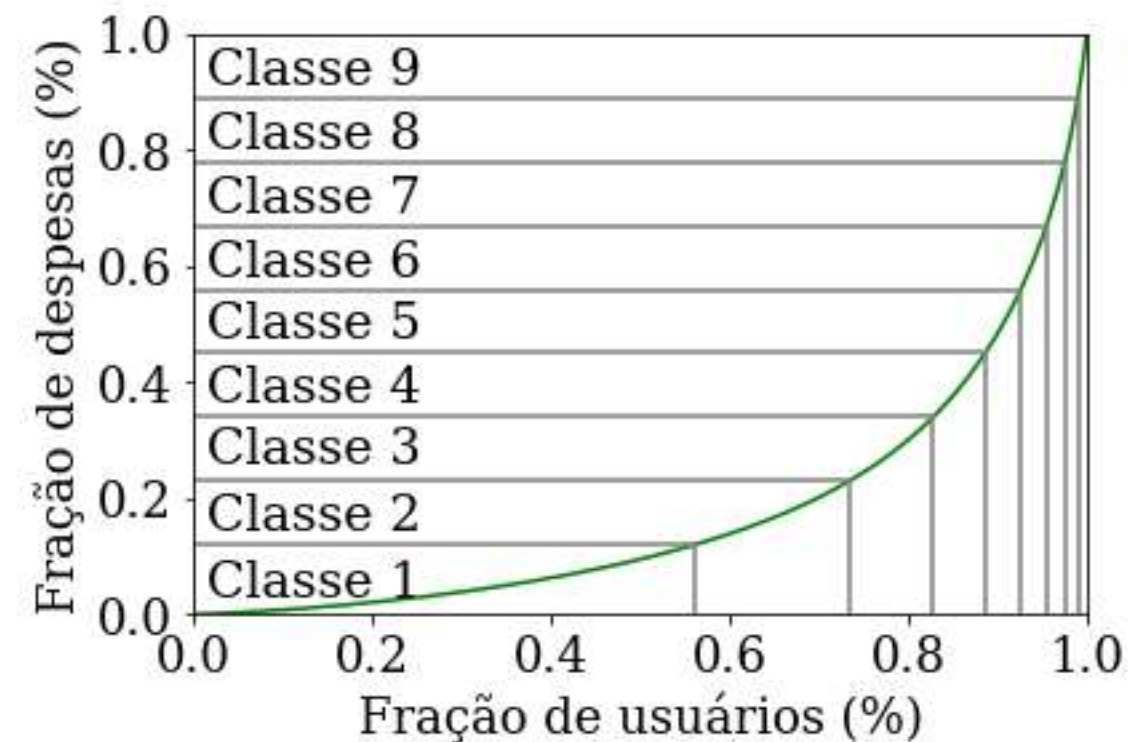


Universidad de Buenos Aires, Grandata

10+ anos de colaboração regular com Univ. Paris 6 (Sorbonnes Universités), ENS-Lyon, ou INRIA

Investigação de possíveis **correlações** entre **padrões de mobilidade** (motifs) e **aspectos sócio-econômicos e demográficos**

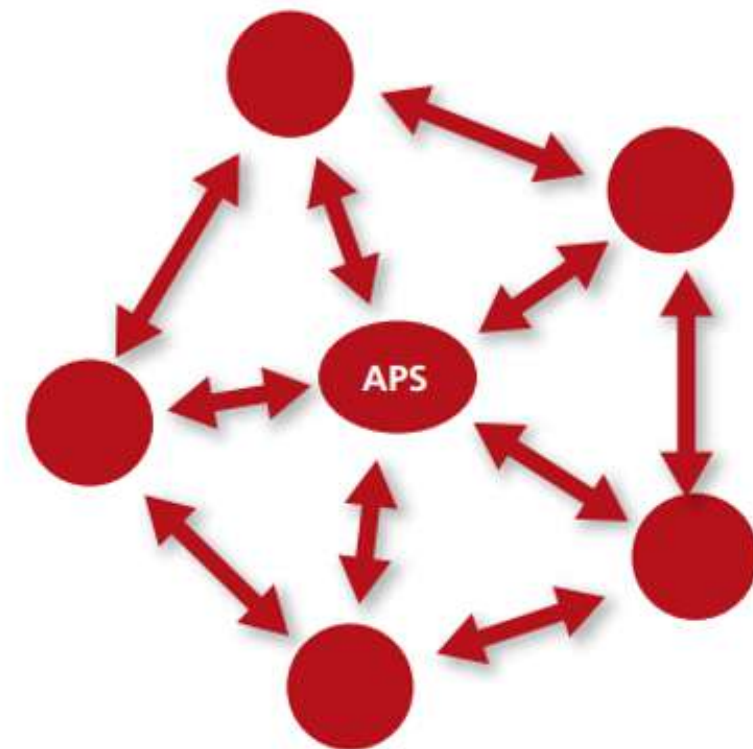
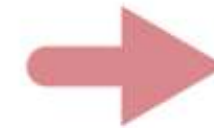
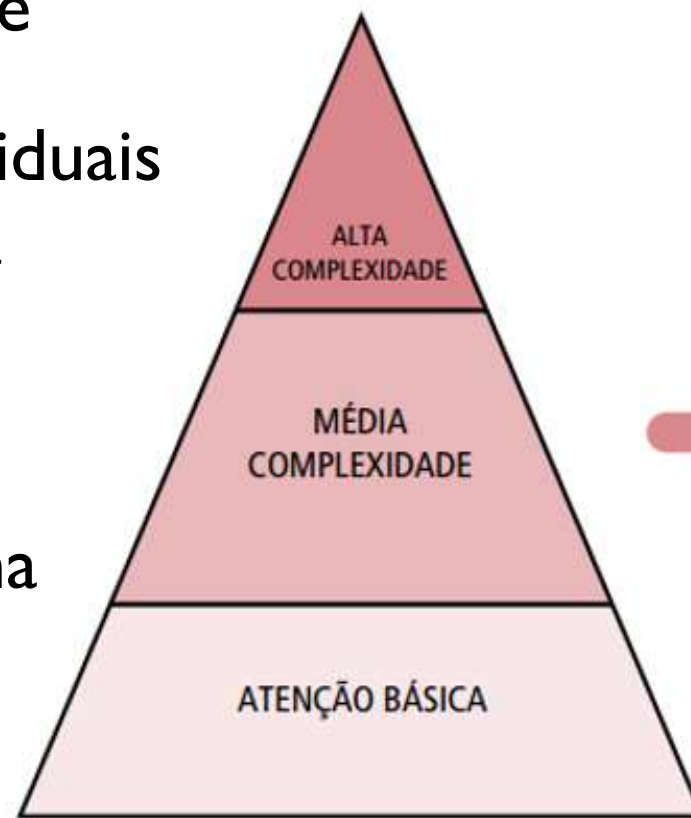
- Trabalho conjunto LNCC, INRIA e Grandata
 - Dados anonimizados
 - *Call details records* (CDRs) de julho a dezembro para **7,9 milhões de usuários** que geraram **4,8 bilhões de chamadas**
 - Dados do banco de janeiro a dezembro do mesmo ano para **7,4 milhões de clientes bancários** responsáveis por **337,5 milhões de transações**
- **250 mil usuários em ambas as bases** (operador telecom e banco usam a mesma função *hash* para anonimizar os números de telefone)
- Outros **dados sócio-econômicos e demográficos** disponíveis (ao menos para parte dos usuários): plano (pós / pré-pago), gênero, idade, renda, despesas, tipo de despesas, ...)



Saúde Pública

Trajетórias centradas em pacientes

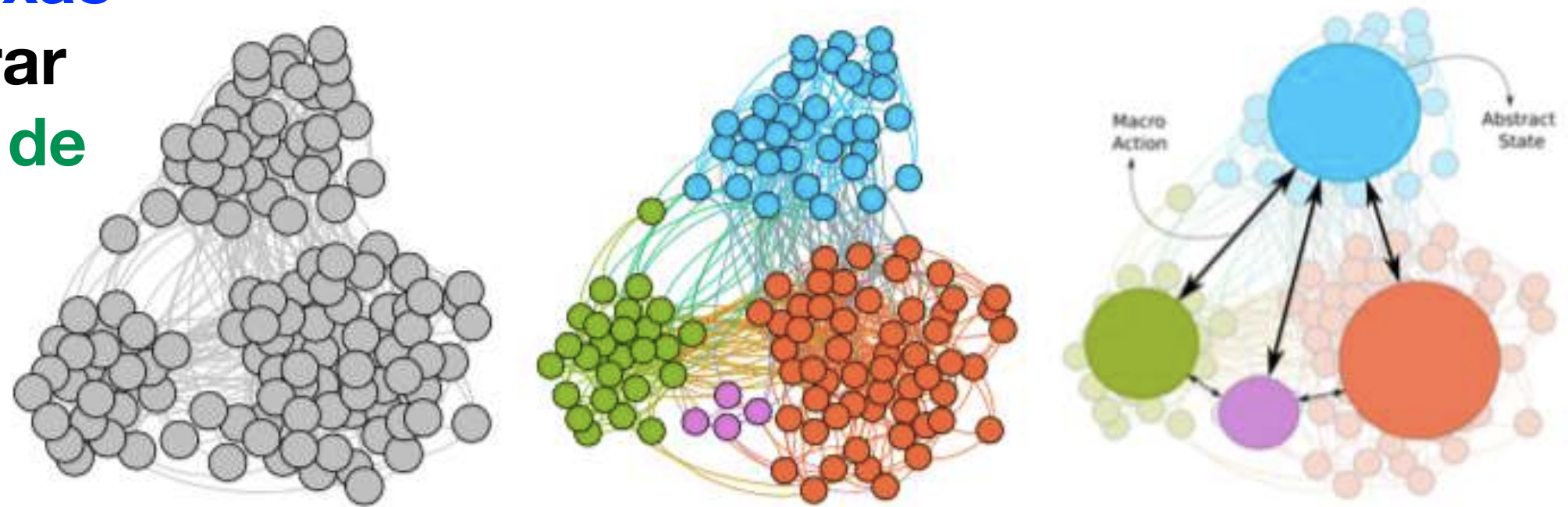
- Cada paciente = uma **trajетória individualizada** através dos serviços de saúde
- Uma teia de trajetórias individuais implica a emergência de uma **rede complexa**
- Nova perspectiva analítica: a partir de cada paciente, uma **trajетória individual centrada no paciente** através dessa rede complexa
- Possíveis **análises orientadas a dados**: monitorar o deslocamento entre serviços, avaliação de desempenho, predição sobre uso dos serviços, gerenciamento de recursos. comparação em sistemas. ...
- **Dados de SP capital**
 - ~25 milhões atendimentos / procedimentos
 - ~6.5 milhões usuários únicos (2014-2015)





Interação entre
redes complexas
e
aprendizado de máquina

Usando técnicas de redes complexas para melhorar aprendizado de máquina



Graph-Based Skill Acquisition For Reinforcement Learning

MATHEUS R. F. MENDONÇA, National Laboratory for Scientific Computing (LNCC), Brazil

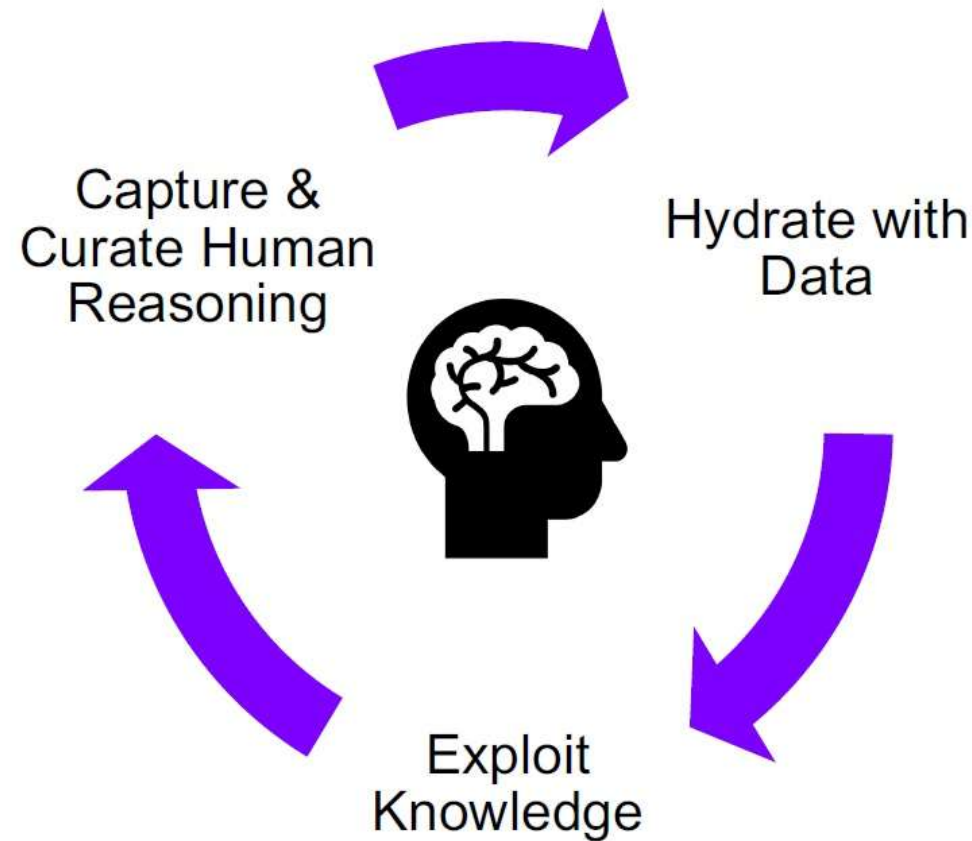
ARTUR ZIVIANI, National Laboratory for Scientific Computing (LNCC), Brazil

ANDRÉ M. S. BARRETO, National Laboratory for Scientific Computing (LNCC), Brazil

In machine learning, Reinforcement Learning (RL) is an important tool for creating intelligent agents that learn solely through experience. One particular sub-area within the RL domain that has received great attention is how to define macro-actions, which are temporal abstractions comprised of a sequence of primitive actions. This sub-area, loosely called skill acquisition, has been under development for several years and has lead to better results in a diversity of RL problems. Amongst the many skill acquisition approaches, graph-based methods have received considerable attention. This survey presents an overview of graph-based skill acquisition methods for RL. We cover a diversity of these approaches and discuss how they evolved throughout the years. Finally, we also discuss the current challenges and open issues in the area of graph-based skill acquisition for RL.

ACM Computing Surveys (CSUR), vol. 52, issue 1, article no. 6, February 2019.

Usando um grafo de conhecimento



Copyright © 2019 Accenture. All rights reserved.

- Para construir **bases de conhecimento** de domínios específicos
- Integrar diversas fontes de dados, potencialmente heterogêneas (textos, pdfs, web-pages, ...)
- Inferir e expressar correlações e associações entre entidades através de um **grafo de conhecimento**
 - usar fatos conhecidos para treinar um **modelo preditivo**

Conclusão: Interesses atuais

- **Caracterização, modelagem e análise de redes de alta ordem**
 - variantes no tempo, multicamada, multicamada variantes no tempo, ...
- **Em redes de alta ordem, investigar**
 - propriedades de redes desse tipo
 - processos dinâmicos sobre redes desse tipo
- **Criação de ferramentas para modelagem/análise e suas aplicações interdisciplinares**



Conclusão: Interesses atuais

- **Aprendizado de máquina x redes complexas (dinâmicas)**

Como aprendizado de máquina pode

{ **se beneficiar de** | **beneficiar** }

técnicas de redes complexas (dinâmicas?)



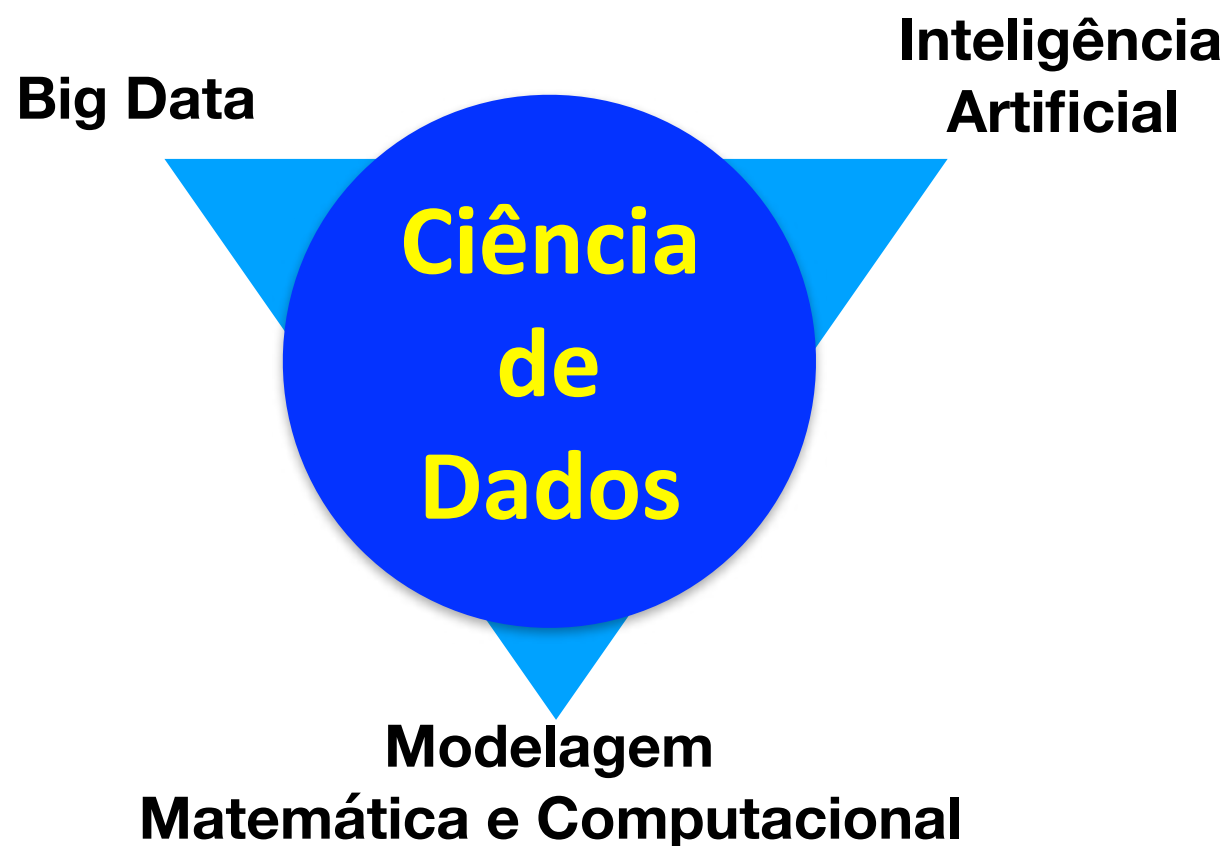
- Predição de enlaces, detecção de anomalias, avaliação de centralidades, revelação de padrões, ...
- **Aplicações interdisciplinares**

Conclusão: Interesses atuais

- **Construção e uso de bases de conhecimento em diferentes domínios**
 - Disponibilidade de múltiplas fontes de dados, em volume considerável, permite **evolução de bancos de dados para bases de conhecimento**, ajudando a **transformar informação em conhecimento**
 - Modelagem e análise de **grafos de conhecimento**
- P.ex. Projeto atual com **Petrobras** visa **Desenvolvimento de Software de Plataforma de Inteligência Computacional**; basicamente base de conhecimento no núcleo para um sistema de gestão e comparação de modelos de predição na área de construção e operação de poços

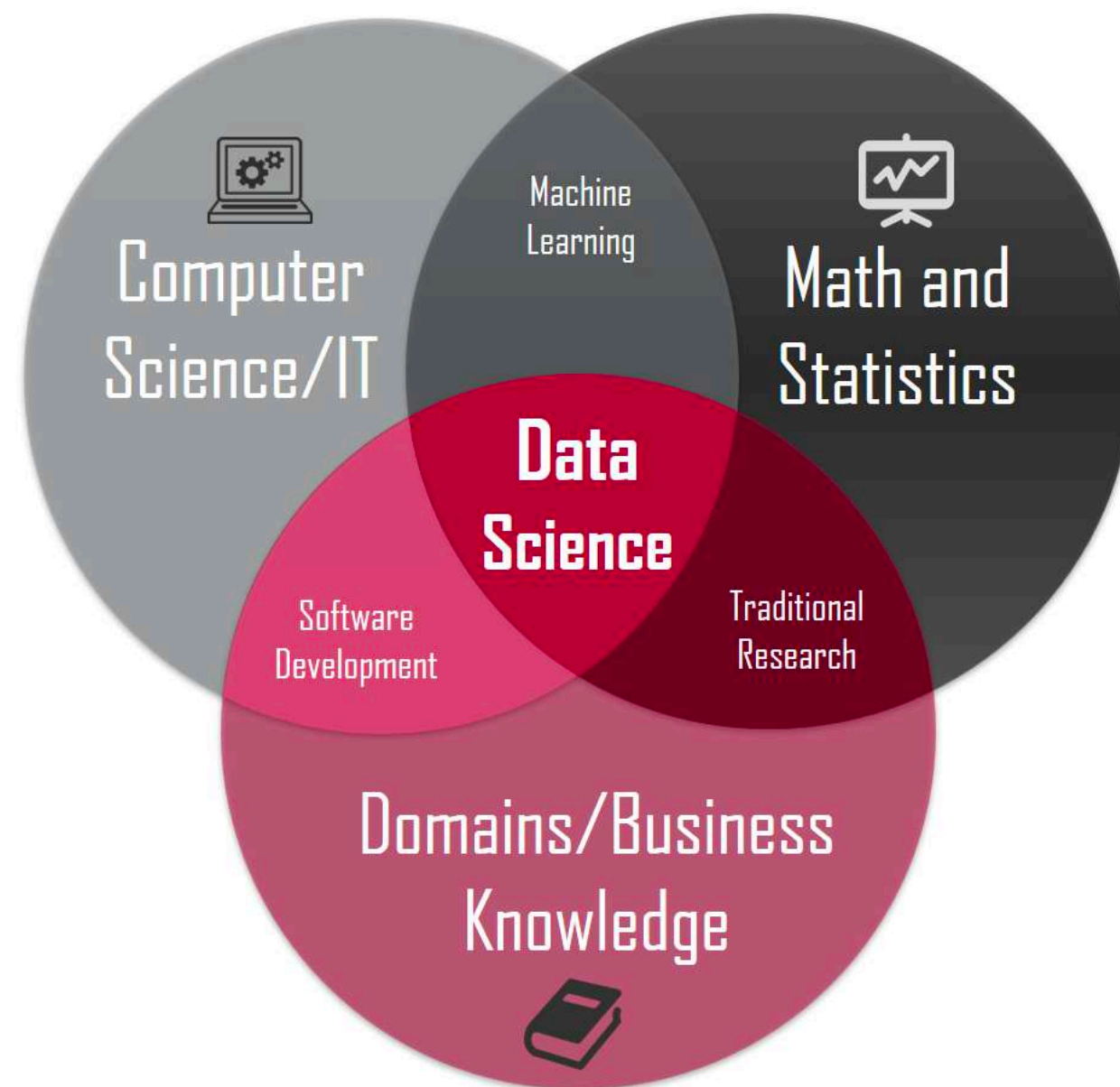


<http://dexl.lncc.br>



Tópicos de pesquisa

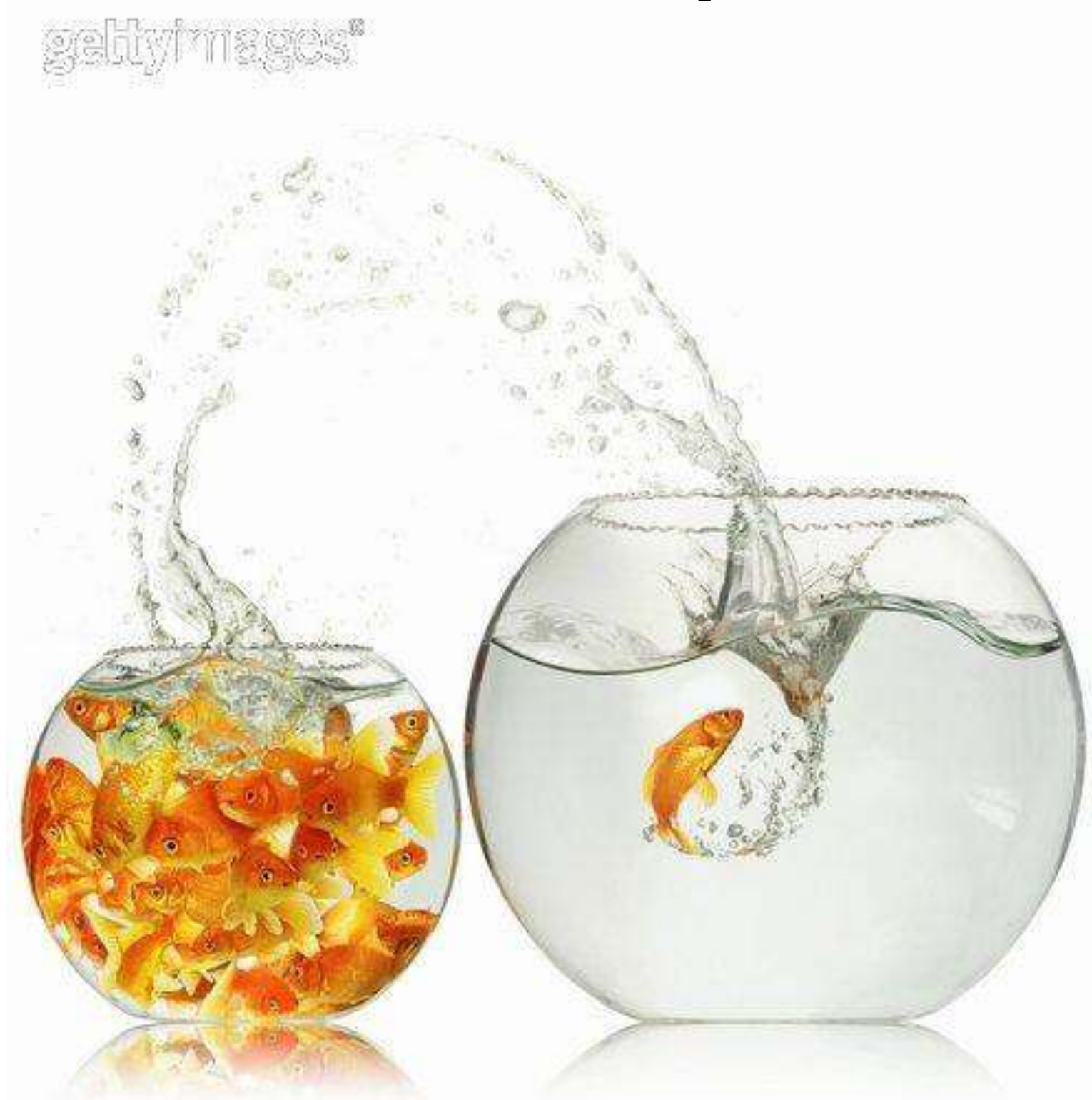
- Gerenciamento de *big data*
- Modelagem de dados
- Algoritmos eficientes para análise de *big data*
- **Ciência de redes**
- *Workflows* científicos
- Aprendizado de máquina



Pesquisadores:

Fabio Porto
Artur Ziviani
Luiz Gadelha

Muitos desafios! Oportunidades?





Obrigado!



Laboratório
Nacional de
Computação
Científica

Artur Ziviani

ziviani@lncc.br

<http://www.lncc.br/~ziviani>



Agradecimentos:



Material de reserva

Tópicos de pesquisa

- Gerenciamento de *big data*
- Modelagem de dados
- Algoritmos eficientes para análise de *big data*
- **Ciência de redes**
- *Workflows* científicos
- Aprendizado de máquina

Tópicos de pesquisa

- Gerenciamento de *big data*
- Modelagem de dados
- Algoritmos eficientes para análise de *big data*
- **Ciência de redes**
- *Workflows* científicos
- Aprendizado de máquina

Meu interesse particular
(em articulação com as outras
linhas de pesquisa do grupo)

Algumas áreas de atuação em colaboração

Astronomia ● ●

(LIneA)

Saúde pública ● ●

(ICICT e CIDACS @ Fiocruz, DGH/MS)

Ciência do esporte ●

(COB)

Bioinformática ● ●

(LABINFO/LNCC)

Biodiversidade ● ●

(Jardim Botânico do Rio; CISS @ Fiocruz)

Petróleo&gás ● ●

(Petrobras, CEFET-RJ)

Meteorologia ●

(IAG-USP)

Computação urbana ●

(UFMG, PUC-Minas, UTFPR, Grandata, INRIA)

Internet ●

(UFBA, UFJF, RNP)

...

Outras em
discussão!!!

- Fabio Porto
- Artur Ziviani
- Luiz Gadelha



Pos-docs (4)

- Douglas de Oliveira ●
- Felipe S. Abrahão ●
- Klaus Wehmuth ●
- Yania M. Souto ●

Doutorandos (10)

- Anderson Chaves da Silva ●
- Claudio Daniel T. Barros ●
- Daniel Gaspar ●
- Daniel N. Ramos da Silva ● ●
- Haron C. Fantecele ●
- Hermano L. Lustosa ●
- Rocio Milagros ●
- Maria Luiza Mondelli ●
- Matheus R. F. de Mendonça ●
- Yasmmin Cortes Martins ●

Pesquisadores sênior visitantes em 2019

- Alex Borges Vieira (UFJF)
- Marcel Pedroso (Fiocruz)

Bolsistas de projeto com MSc (1)

- Juliana Mascarenhas ●

Mestrandos (4)

- João Guilherme N. Rittmeyer ●
- Nathália Barbosa Santos ●
- Rafael Pereira ●
- Victor Medeiros ●

ICs (3)

- Andre Demori ●
- Matheus M. R. de Albuquerque ●
- Raquel Junqueira ●

Suporte Técnico (3)

- Adolfo Simões
- Carlos Cardoso
- Enver C. Cayo

Interdisciplinar

Empírica e
orientada a dados

Ciência
de
Redes

Quantitativa e
matemática

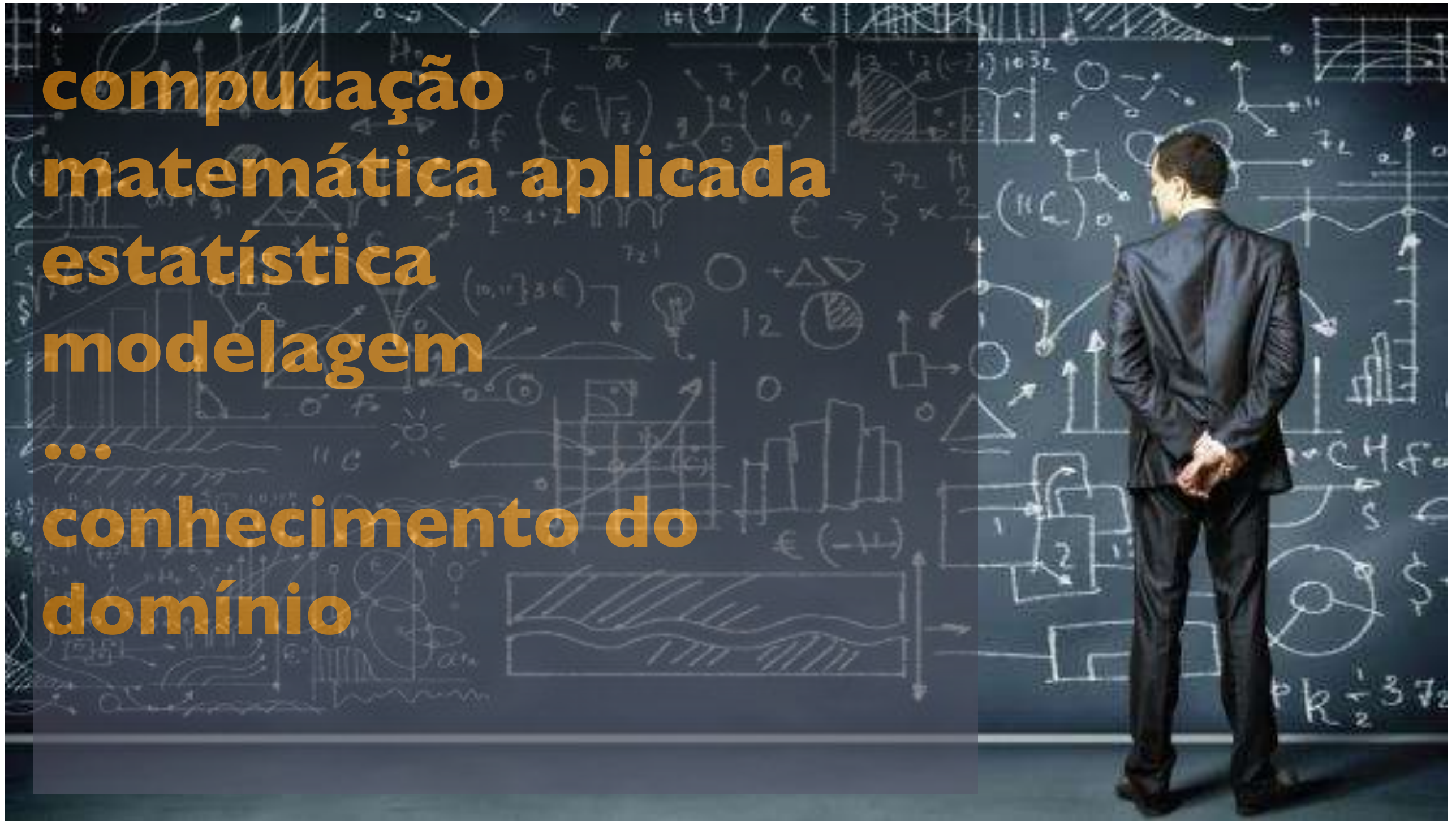
Computacional

Cientista de dados?

computação
matemática aplicada
estatística
modelagem

...

**conhecimento do
domínio**



Desafio na formação de recursos humanos!

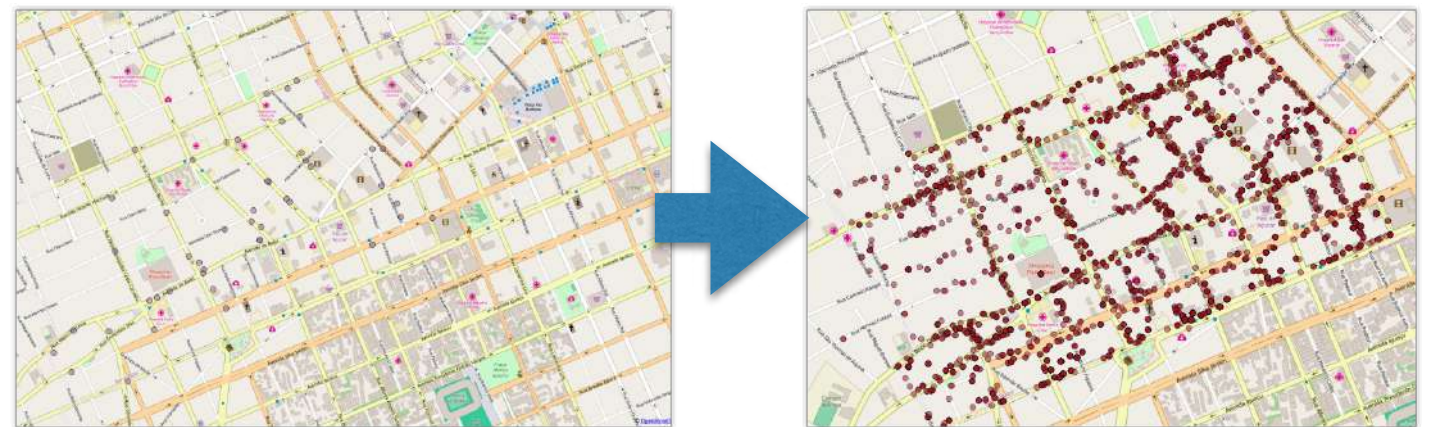
30+ anos de atividade econômica em Curitiba

- Dados de concessão de **172.173 alvarás** em Curitiba (1980-2013)
- Como se deu a **evolução** da atividade econômica?
- Expansão / diversificação

Centro de 1980 a 2013



Batel de 1980 a 2013



Colaboração:
LNCC e UTFPR com dados
da Prefeitura de Curitiba

N. P. Kozievitch, T. H. Silva, A. Ziviani, G. Costa, G. Lugo,
Three Decades of Business Activity Evolution in Curitiba: A Case Study
Annals of Data Science, Springer, ISSN: 2198-5812, vol. 4, no. 3, pp. 307-327, September 2017.
A. Ziviani (LNCC) - 2019

- Projeto em colaboração entre ICICT/Fiocruz, LNCC e CEFET-RJ recém-aprovado junto ao CNPq e a Fundação Bill & Melinda Gates
- *Assessing the impact of hospital-based breastfeeding interventions on infant health*
- Objetivo: Avaliar o impacto da implementação de políticas e iniciativas hospitalares pró-aleitamento materno,* e comparação entre elas, na morbimortalidade neonatal, delineando cenários de custo-efetividade dessas iniciativas