Dilúvio de Dados

Uma Boia à Vista, ou seria um Cruzeiro?

MIRELLA M. MORO
UFMG

SBBD 2023

Quantidade <> Comunidade

Mar 2012
SBC

Carol

Quantidade

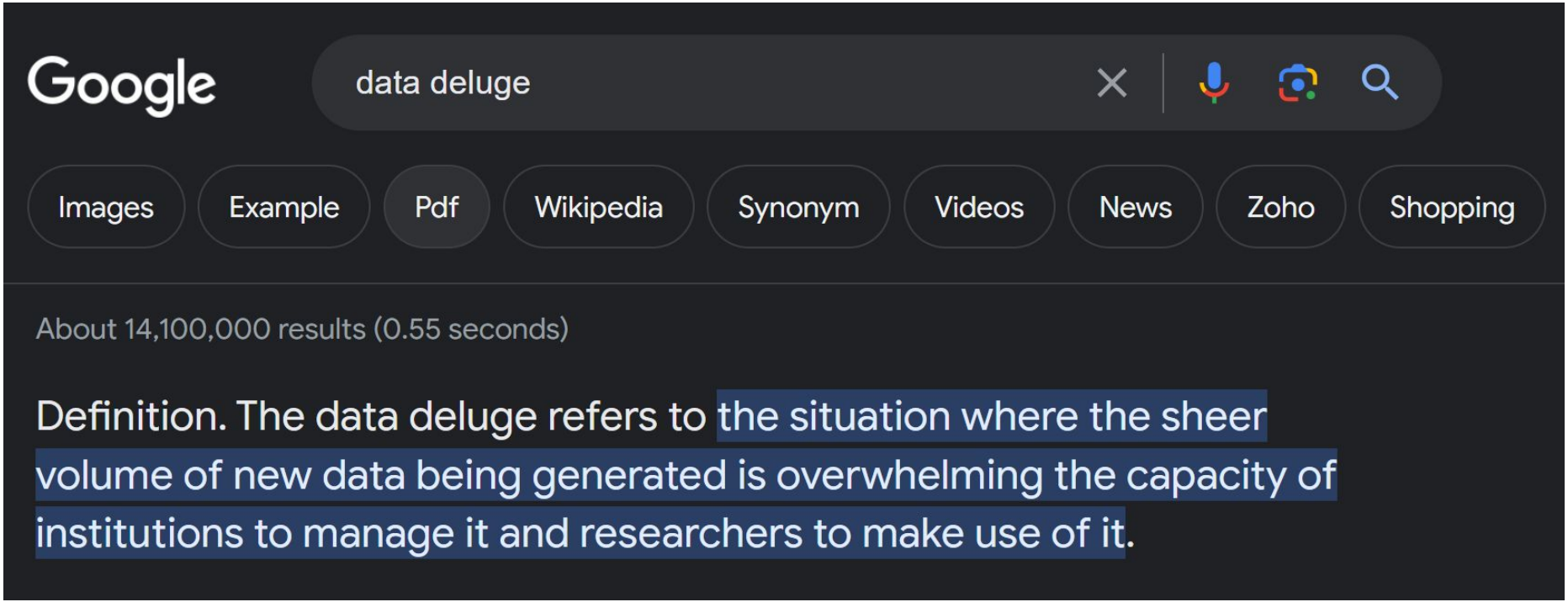
<>

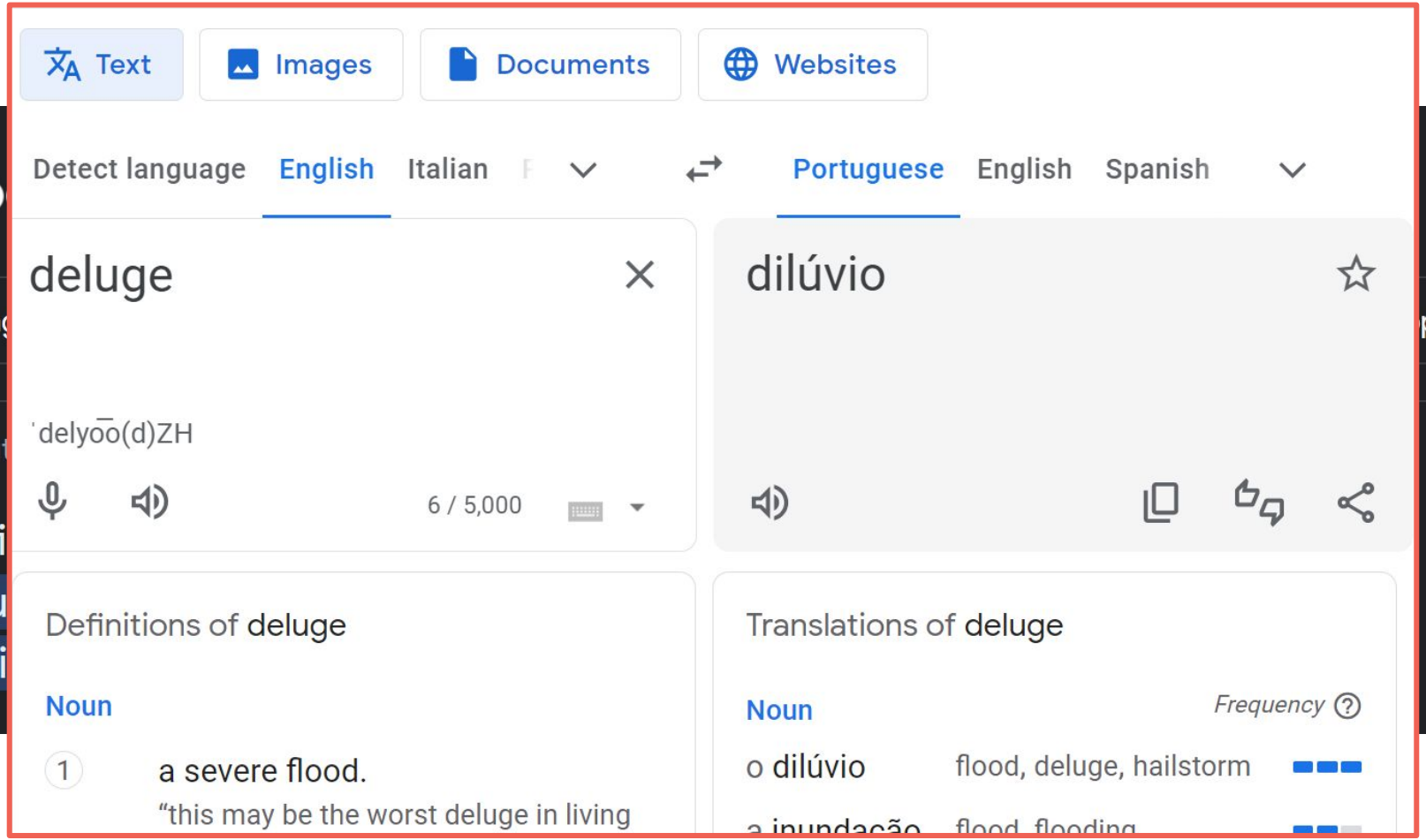
Comunidade

Karin

“Vida de cientista significa estar sempre aprendendo”

Data deluge


Google
data deluge
Images
Example
Pdf
Wikipedia
Synonym
Videos
News
Zoho
Shopping
About 14,100,000 results (0.55 seconds)
Definition. The data deluge refers to the situation where the sheer volume of new data being generated is overwhelming the capacity of institutions to manage it and researchers to make use of it.


Text
Images
Documents
Websites
Detect language
English
Italian
Portuguese
English
Spanish
deluge
ˈdelyo͞o(d)ZH
6 / 5,000
dilúvio
Definitions of deluge
Noun
1
a severe flood.
“this may be the worst deluge in living
Translations of deluge
Noun
Frequency
o dilúvio
flood, deluge, hailstorm
a inundação
flood, flooding

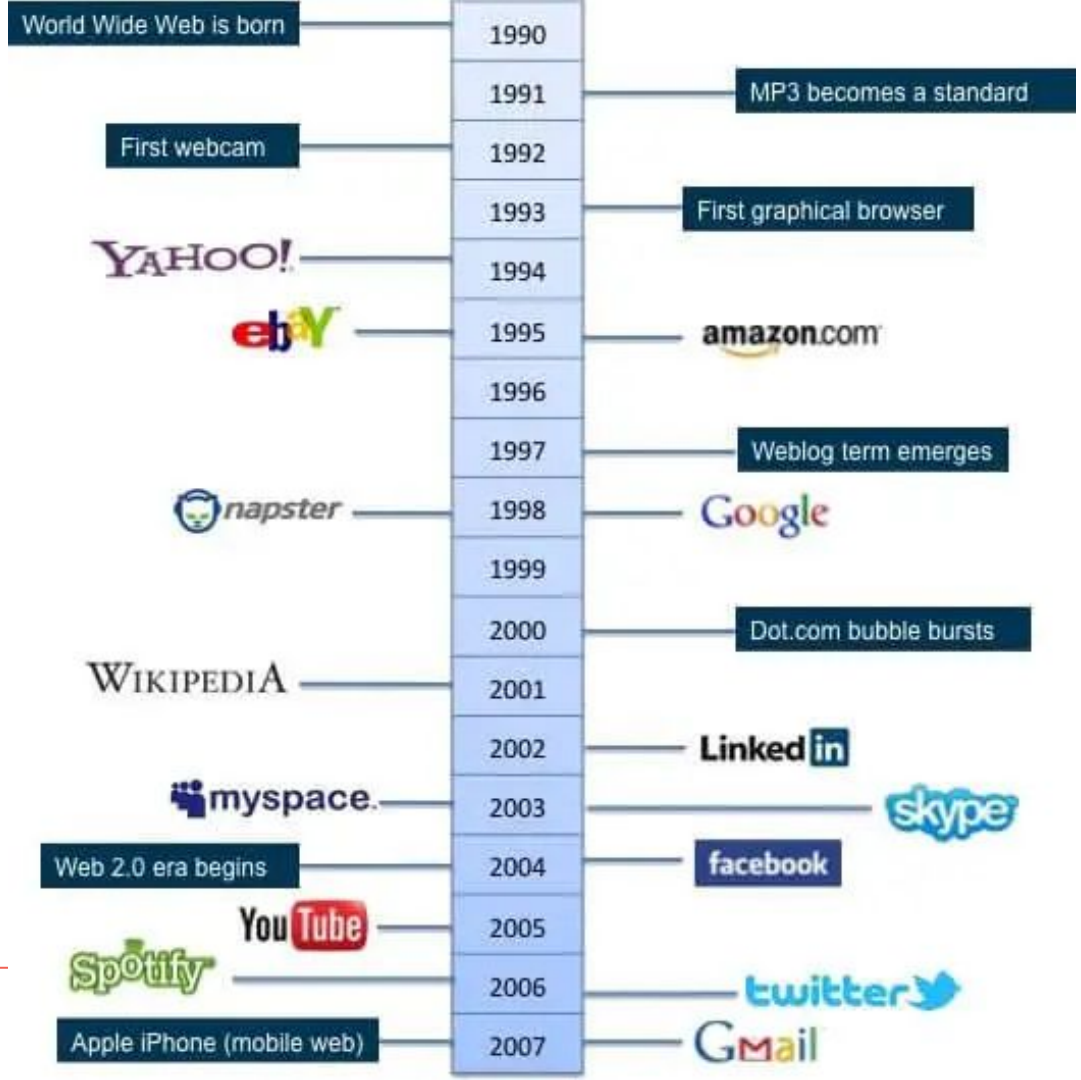
Dilúvio de Dados

Dilúvio de Dados

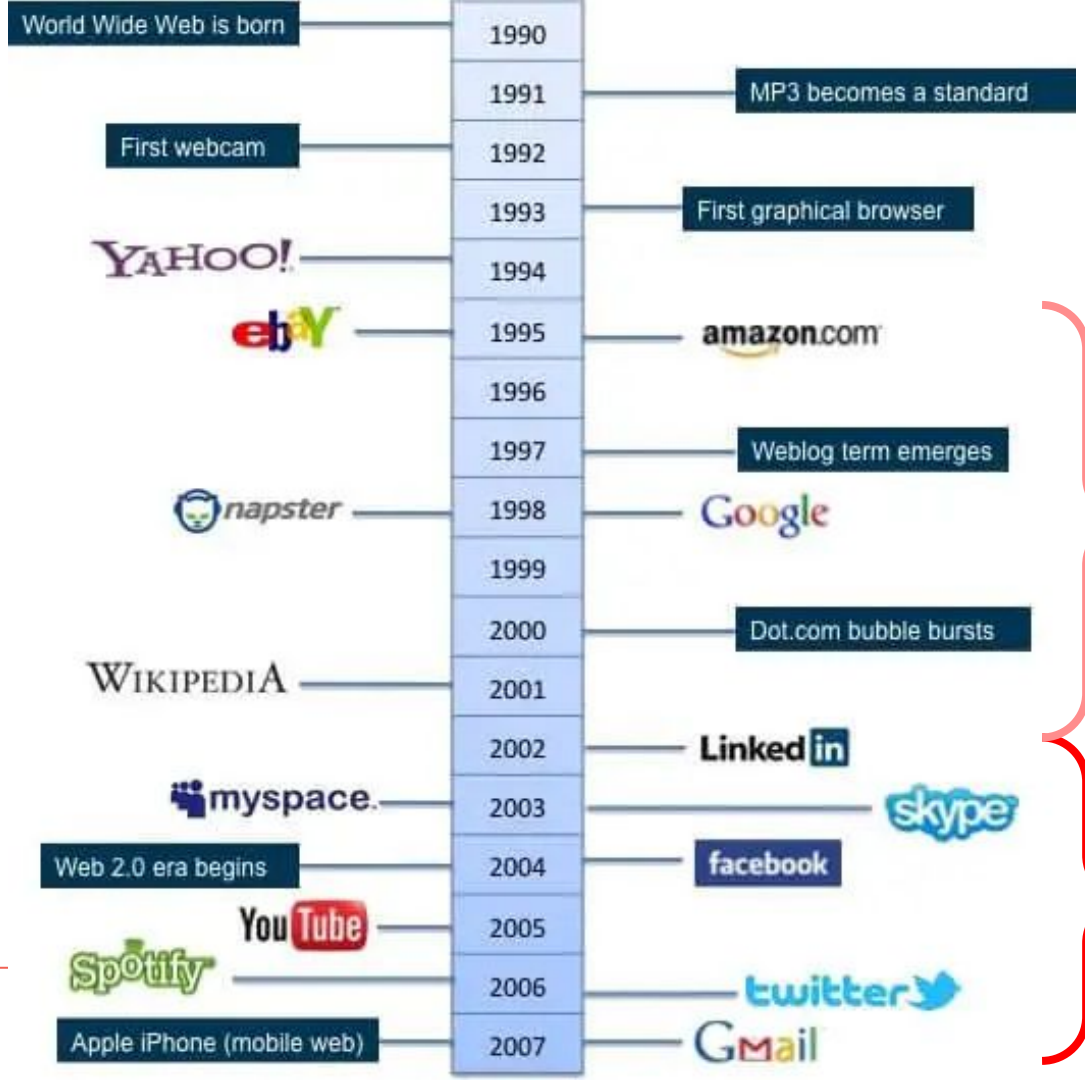
Uma Boia à Vista,
ou seria um Cruzeiro?

1. Uma Boia à Vista

"Entender de tecnologia é entender de história e evolução"



Fonte:
https://harbott.com/visual-timeline-of-internet-milestones



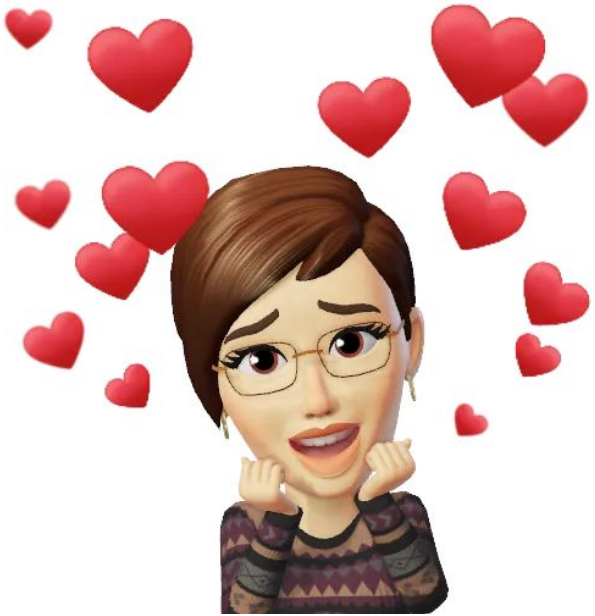
B+M @UFRGS

PhD @UCR

Criar + Consumir Conteúdo Web

→ Cientistas
→ Profissionais
→ Pessoa comum

Dilúvio de Dados




XML
Extensible Markup Language

Introduction

- XML data everywhere
  - Web services (RSS feed format)
  - Integration (e-commerce, stocks)
  - Open standards for file exchange (FpML,ACORD)
- XML database
  - XML native engine
    - Specialized storage and retrieval
  - XML-enabled Relational DBMS
    - Shredding or text

Tutorial

XML Databases

XML Document = sequence of elements that enclose text values and other elements

```
<book >
   <title>“Adventures of Huckleberry Finn”</title>
   <author>“Mark Twain”</author>
   <year>2002</year>
   <otherInfo>
      <isbn>0142437174</isbn>
      <collection>“Penguin  Classics”</collection>
   </otherInfo>
</book>
```

XML Databases

XML Document = sequence of elements that enclose text values and other elements

Element name between < >

Text values between elements

```
<book >
   <title>“Adventures of Huckleberry Finn”</title>
   <author>“Mark Twain”</author>
   <year>2002</year>
   <otherInfo>
      <isbn>0142437174</isbn>
      <collection>“Penguin  Classics”</collection>
   </otherInfo>
</book>
```

“Menos é mais na vida e no slide”

XML Nativo

Minicurso

Mirella M. Moro
Departamento de Ciência da Computação
Universidade Federal de Minas Gerais
mirella@dcc.ufmg.br

XML Nativo

- XML

A linguagem que apareceu para resolver todos os problemas do mundo

- É… tipo assim…

"XML Fever"
E. Wilde, R. J. Glushko
Communications ACM, July 2008

XML está por toda a parte…

- Advertising: adXML place an ad onto an ad network or to a single vendor
- Banking: MBA Mortgage Bankers Association of America → credit report, loan file, underwriting…
- Directories: dirXML Novell's Directory Services Markup Language
- **Literature**: Gutenberg convert the world's great literature into XML
- Geospatial: ANZMETA distributed national directory for land information
- **Healthcare**: HL7 DTDs for prescriptions, policies & procedures, clinical trials
- Human Resources: XML-HR standardization of HR/electronic recruiting XML definitions
- International Dvt: IDML improve the mgt. and exchange of info. for sustainable development
- Math: MathML Mathematical Markup Language
- **News**: NewsML creation, transfer and delivery of news
- Surveys: DDI Data Documentation Initiative, "codebooks" in the social and behavioral sciences
- **Travel**: openTravel information for airlines, hotels, and car rental places
- Voice: VoxML markup language for voice applications
- Wireless: WAP Wireless Application Protocol, wireless devices on the World Wide Web
- Weather: OMF Weather Observation Markup Format
- Web Servers: apacheXML parsers, XSL, web publishing
- …

DESMISTIFICANDO XML: DA PESQUISA À PRÁTICA INDUSTRIAL

Mirella Moura Moro – UFMG
Vanessa Braganholo – UFRJ

@SBBD 2007

Bento Gonçalves – RS – Brasil

“Tutorial + Minicurso + JAI = trifecta de recém doutor/a”

Solução ou Problema?

- DBLP
- XML → +5500
- XML query → 604

MORO, BRAGANHOLO, DORNELES, DUARTE, GALANTE, MELLO. *XML: Some Papers in a Haystack*. SIGMOD Record, 2009.

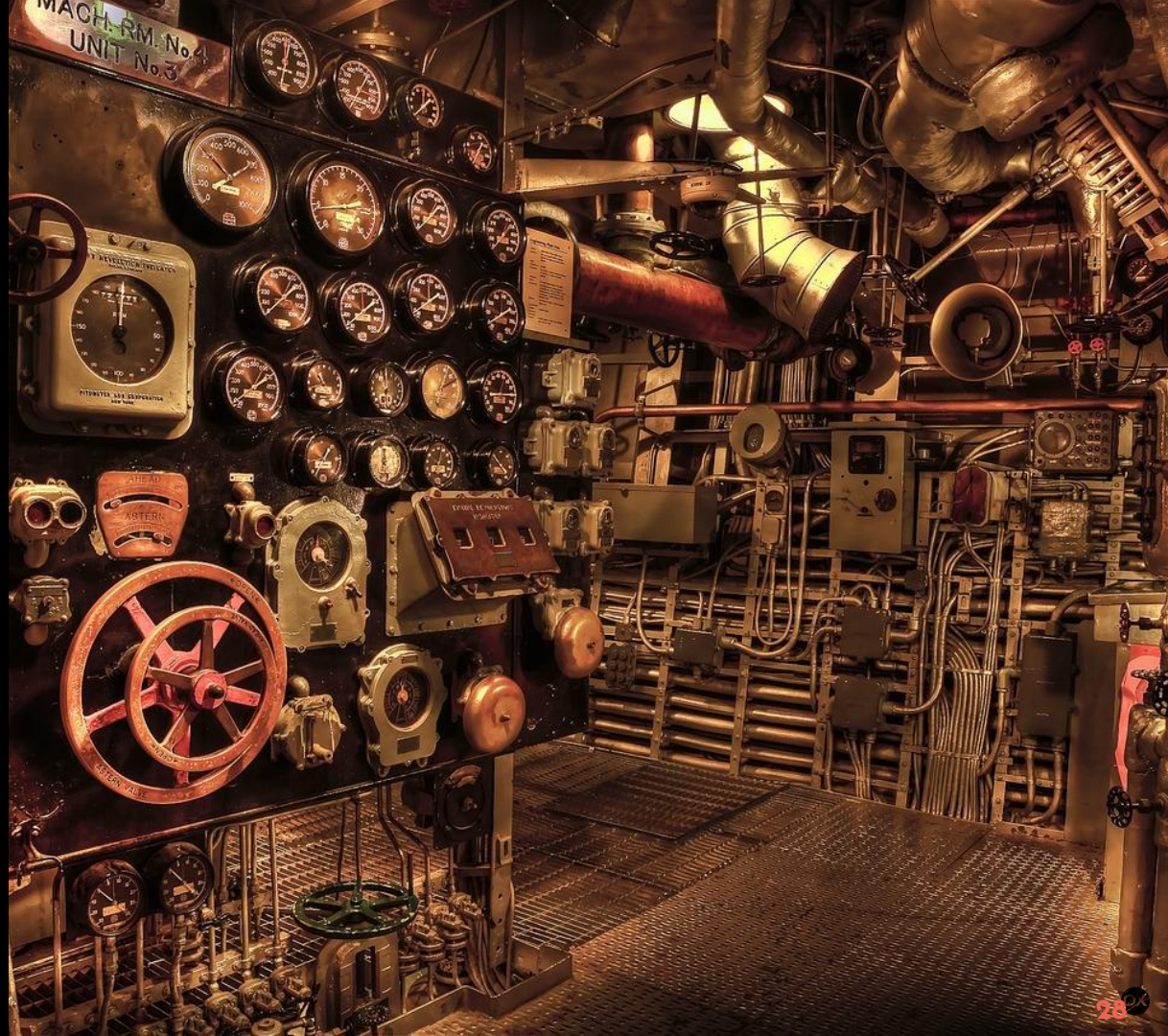
Renata, Vanessa, Ronaldo, Carina

http://www.informatik.uni-trier.de/~ley/db/

Dilúvio de Dados
Uma Boia à Vista

Integração de dados
Padrões
Mídias diferentes

Engenharia de Dados



Cientista da Computação

Engenheira de Dados

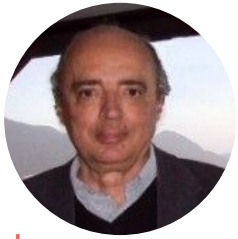
Setembro 2002

Abril 2007

Recém Doutora

Transição

- Recomendação (*streams*)
- Busca (*keyword search*)
- NoSQL



Palazzo

Alberto

Clodoveu

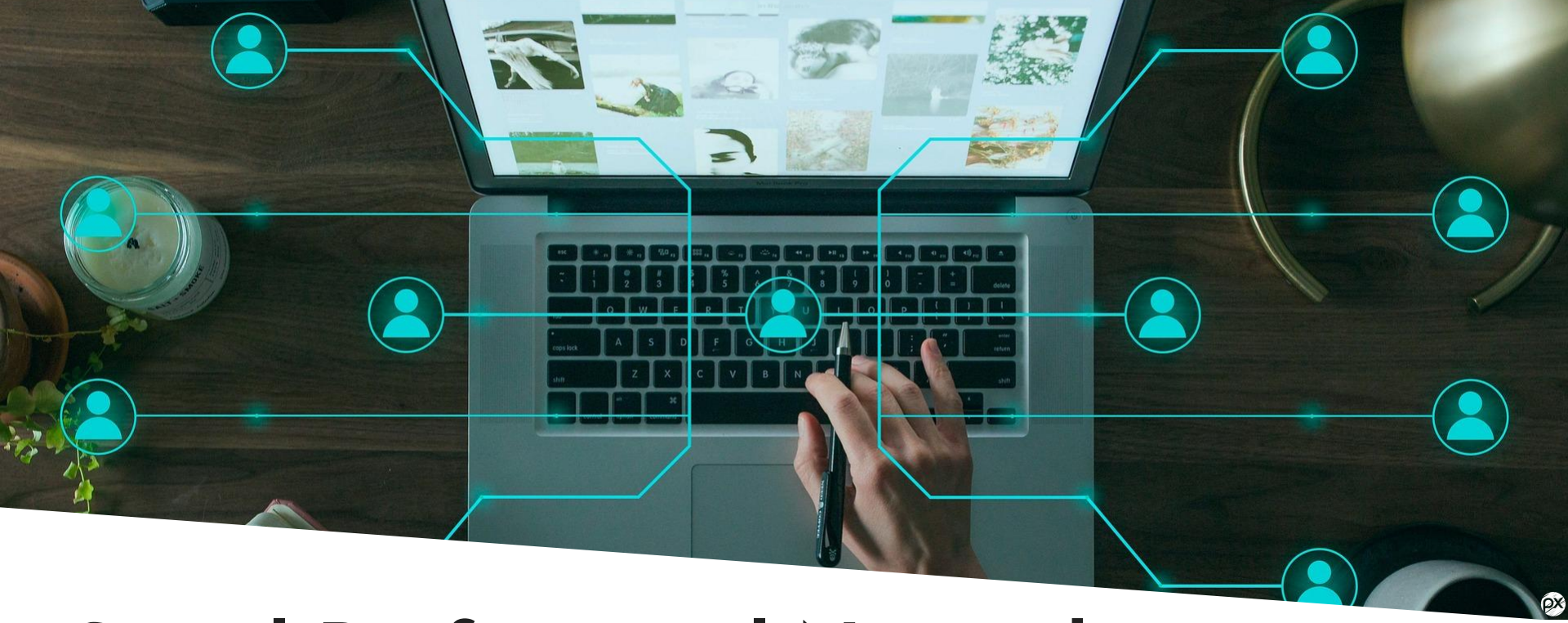
Altigran

“Acolhimento das gerações mais novas faz a diferença! ”

2. Ou seria um Cruzeiro

Dilúvio de Dados

Social



Social Professional Networks

Redes Sociais Acadêmicas

“Ciência até surge do nada. Mas só se fortalece com colaboração. ”

SEMISH 2011

CiênciaBrasil

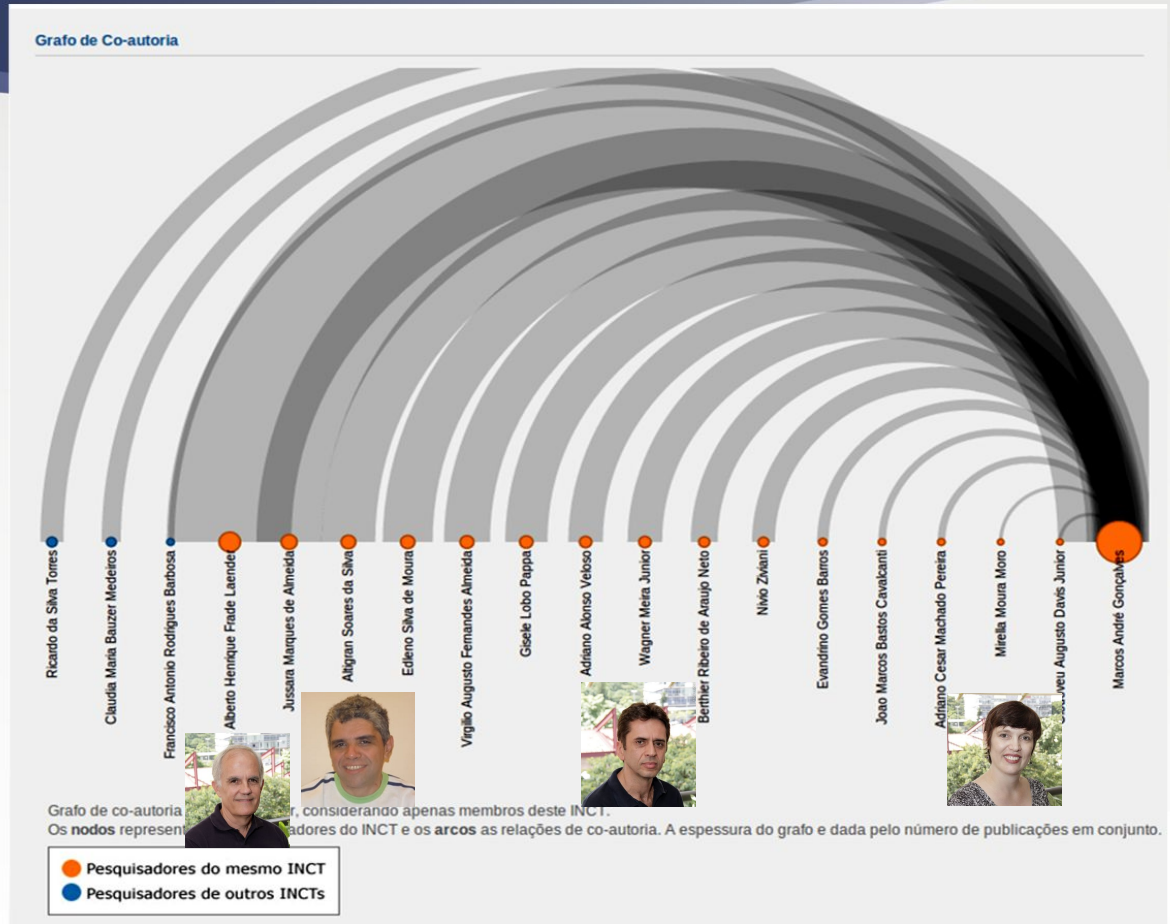
The Brazilian Portal of Science and Technology

Alberto H. F. Laender, **Mirella M. Moro**, Altigran S. Silva, Clodoveu A. Davis Jr., Marcos André Gonçalves, Renata Galante, Allan J. C. Silva, Carolina A. S. Bigonha, Daniel Hasan Dalip, Eduardo M. Barbosa, Eduardo N. Borges, Eli Cortez, Peterson Procópio Jr., Rafael Odon de Alencar, Thiago N. C. Cardoso, Thiago Salles

Winweb

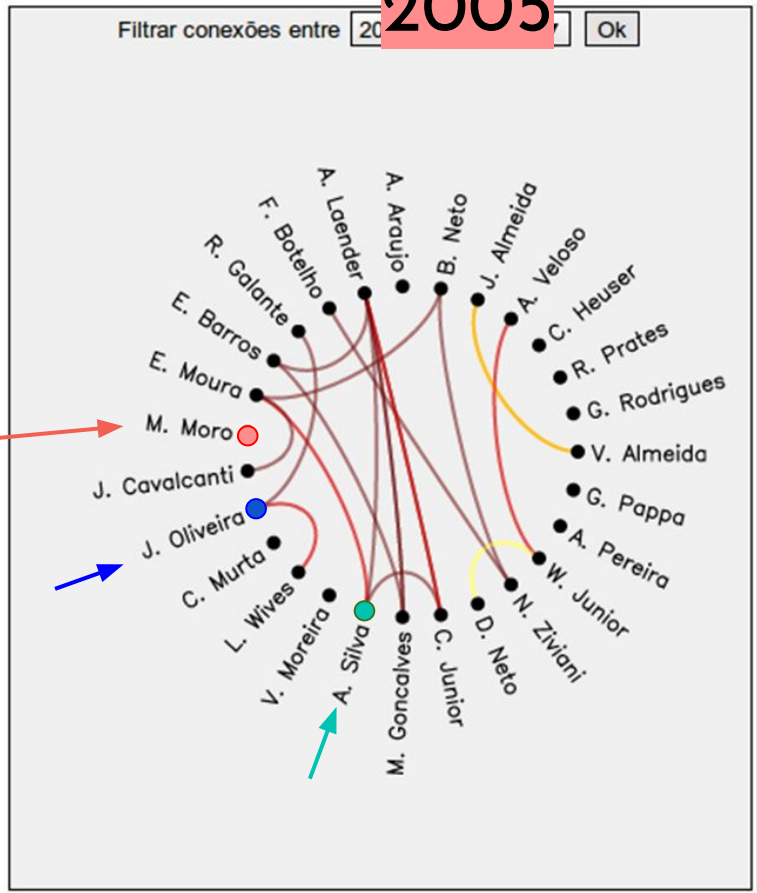
Instituto Nacional de Ciência e Tecnologia para a Web

Grafo de Coautoria de um Pesquisador

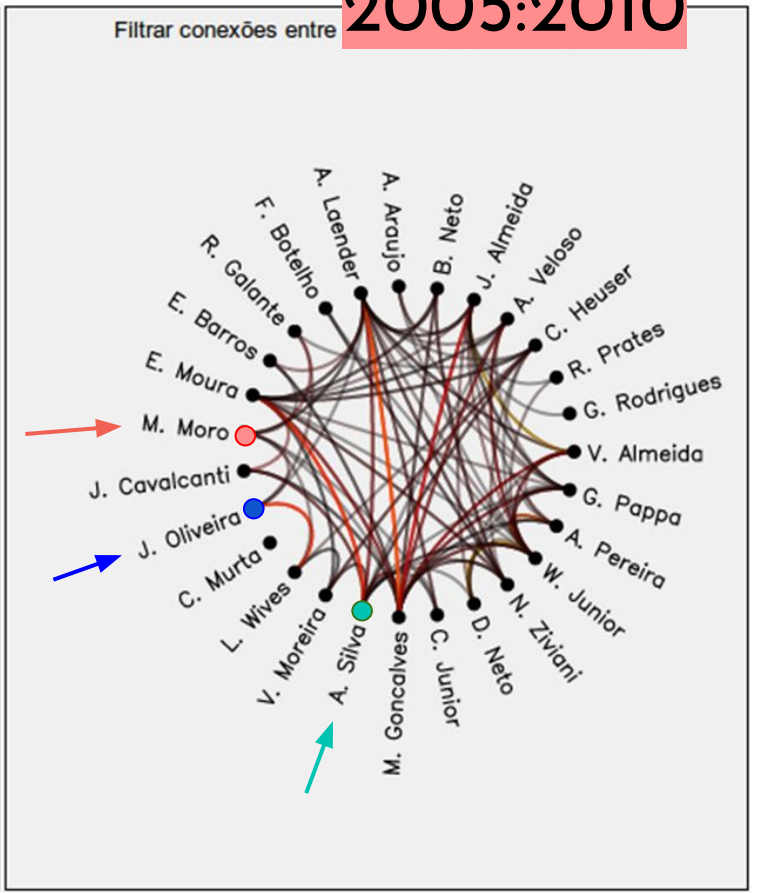


CiênciaBrasil: Groups - período

2005



2005:2010



Arquitetura

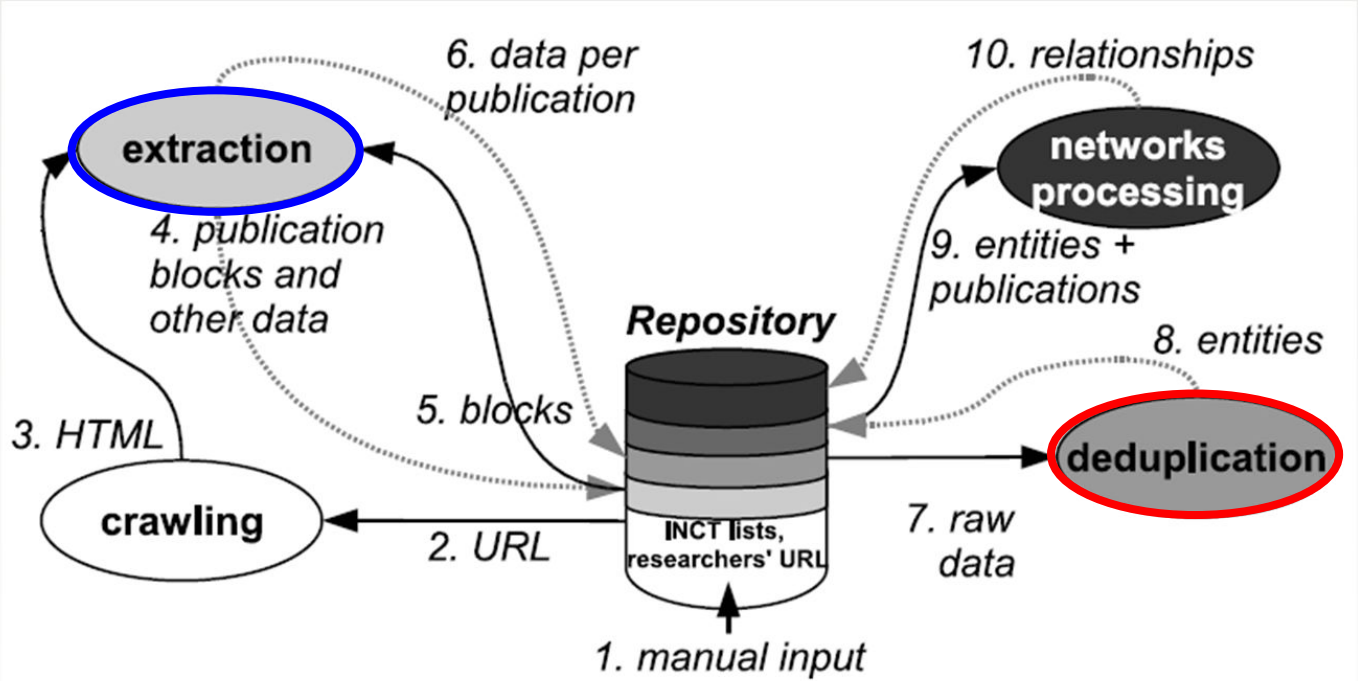
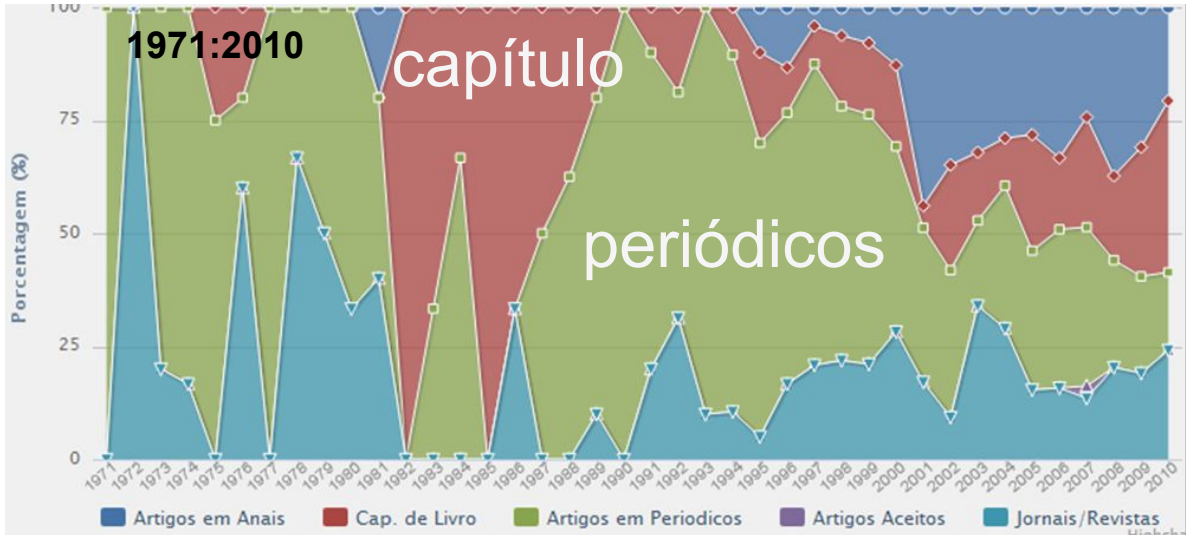
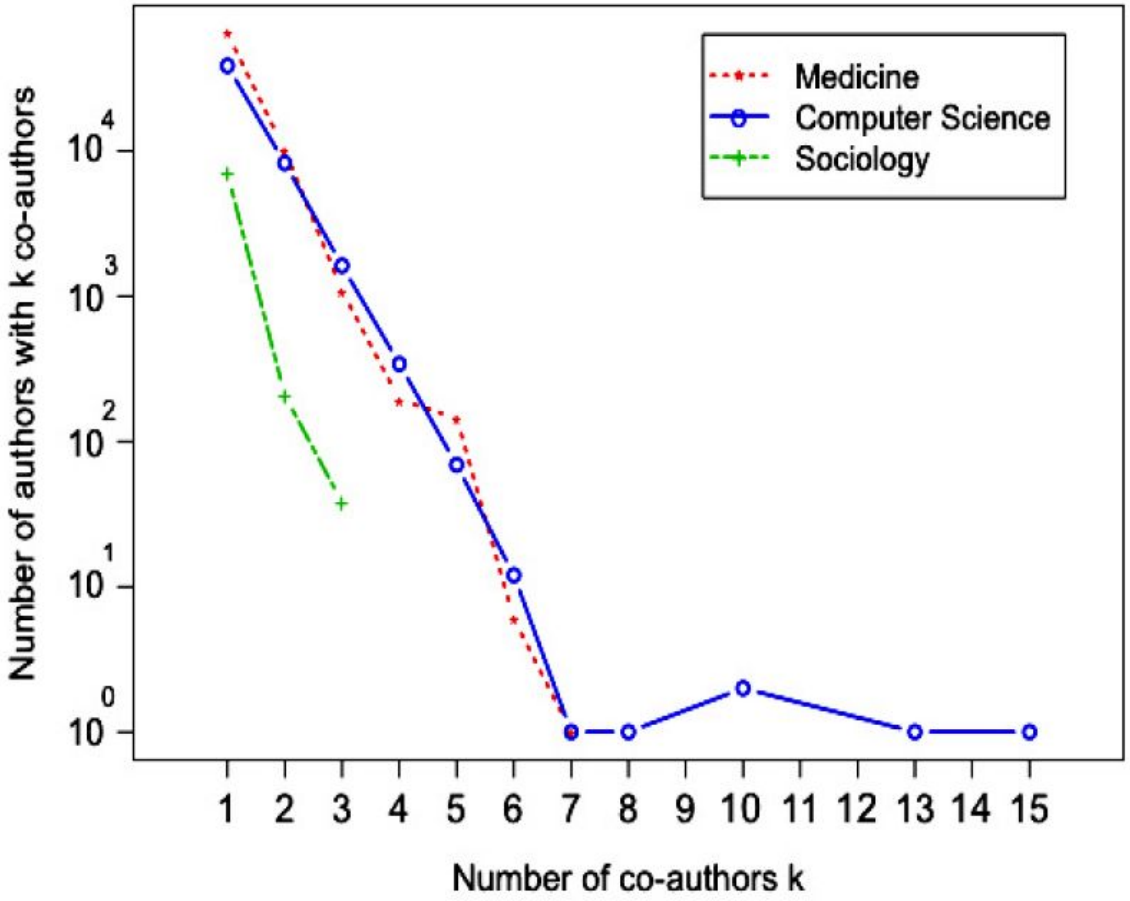


Figure 1. *CiênciaBrasil* Architecture

INCT
Engenharia de
Software



INCT
Ciências Sociais



Distribution of numbers of co-authors for researchers in each area

In Sociology: from 7,195 publications, 84% have only one author

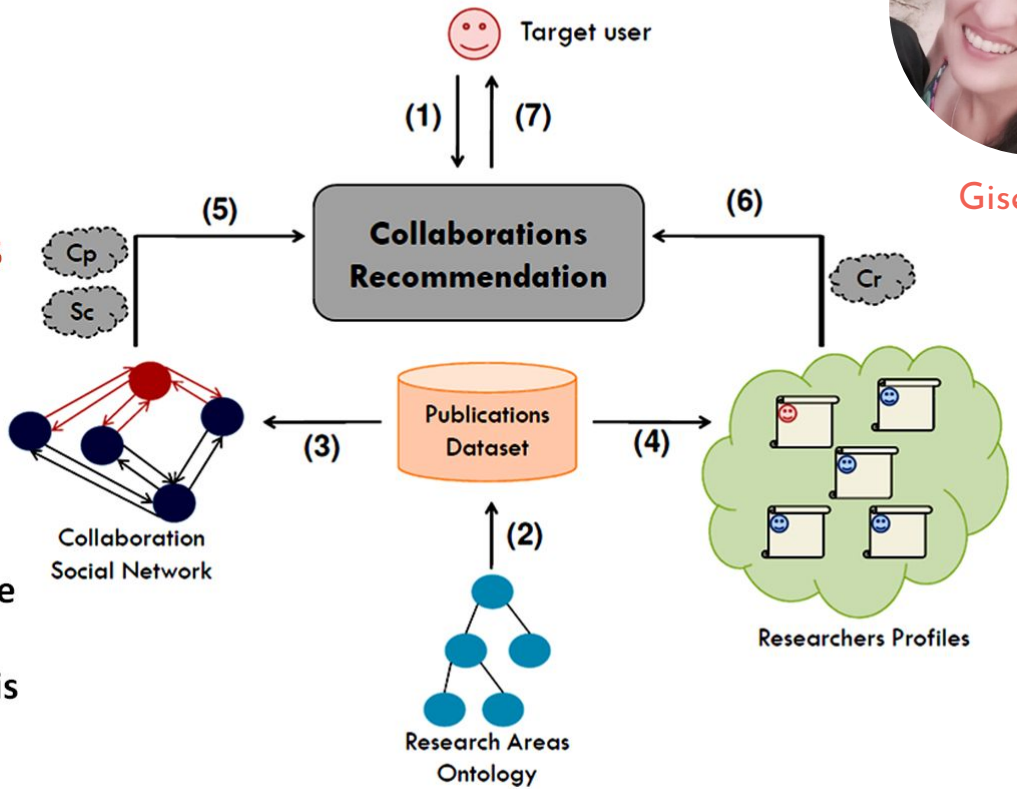
“Analise os dados antes de sair criando redes complexas em cima :-)”

Recomendação de Pessoas

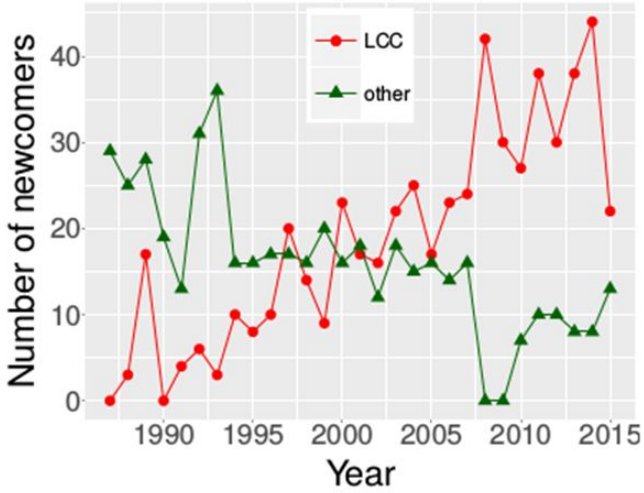
Giseli Lopes Michele Brandão

- **CORALS**: a **CO**llaboration **R**ecommender for **A**cademic socia**L** network**S**

- **ReDisseR**: **Re**comendação e **Disse**minação em **R**edes Sociais



Collaborations and Newcomers

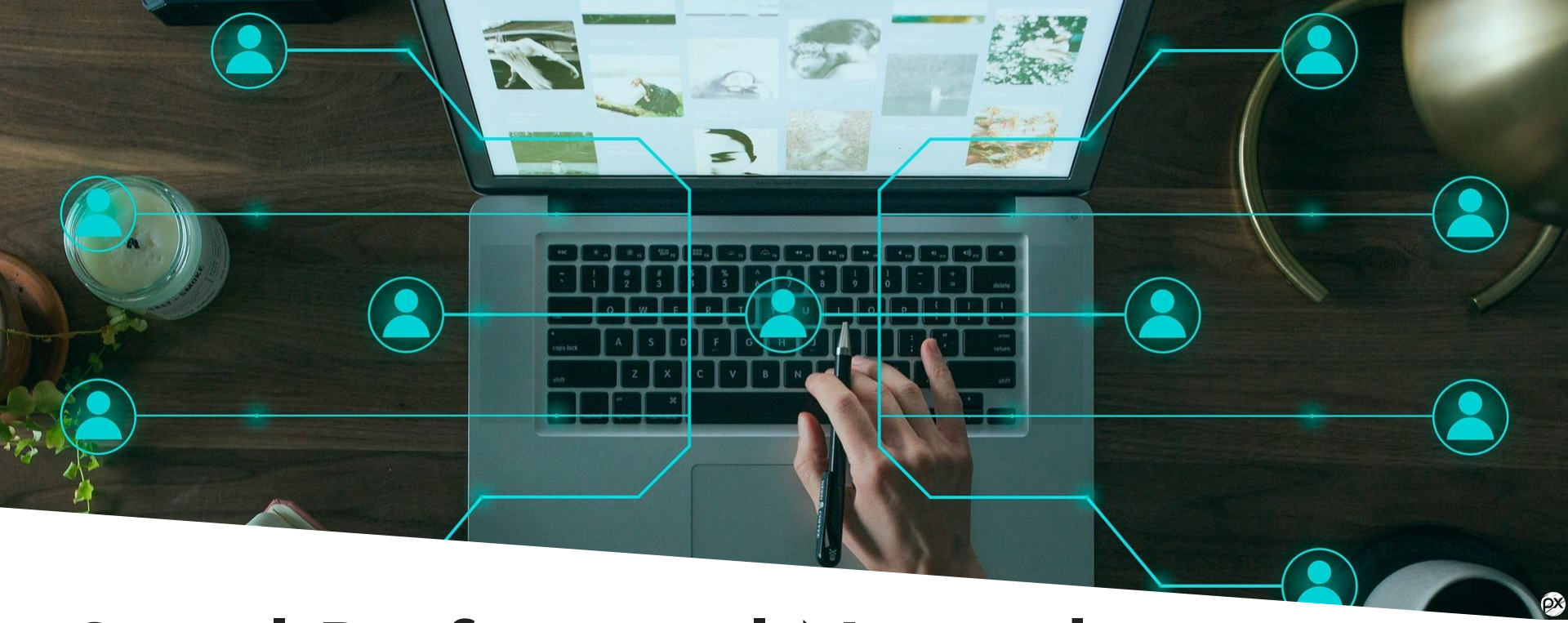


SBBD initially received several authors with no link to its Largest Connected Component

It has become more common to join the symposium by collaborating with someone already connected to it

The Collaboration Network of the Brazilian Symposium on Databases
30 Editions of History

JBCS 23:10 (2017)
10.1186/s13173-017-0059-6



Social Professional Networks

Coautoria foi só o início

Intensidades Diferentes

Forças dos relacionamentos

Pessoas influentes

Mobilidade

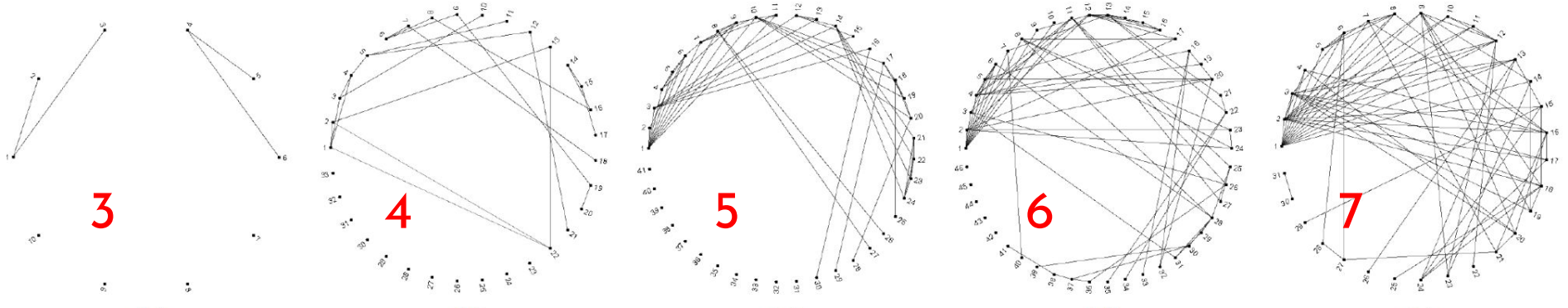
Ranking

PPG

Software@GitHub

Ranking PPG

Colaborações internas ao PPG



Ranking Strategy for Graduate Programs Evaluation
@ ICITA 2011

Aspectos Temporais para Medir a Força da Colaboração no GitHub

Natércia A. Batista, Michele A. Brandão,
Ana Paula C. da Silva, Mirella M. Moro

Best Short Paper

SBBD2017 BRACIS2017

KDMiLE ENIAC STIL

Utilização de Redes Heterogêneas para Medir a Força dos Relacionamentos no GitHub

Gabriel P. Oliveira, Natércia A. Batista, Michele A. Brandão, Mirella M. Moro
Universidade Federal de Minas Gerais (UFMG)

Short SBBD 2018

ELSEVIER

Computer Communications

Volume 100, 1 March 2017, Pages 20-31

Social professional networks: A survey and taxonomy

Michele A. Brandão , Mirella M. Moro

Show more

Outline | + Add to Mendeley | Share | Cite

https://doi.org/10.1016/j.comcom.2016.12.011

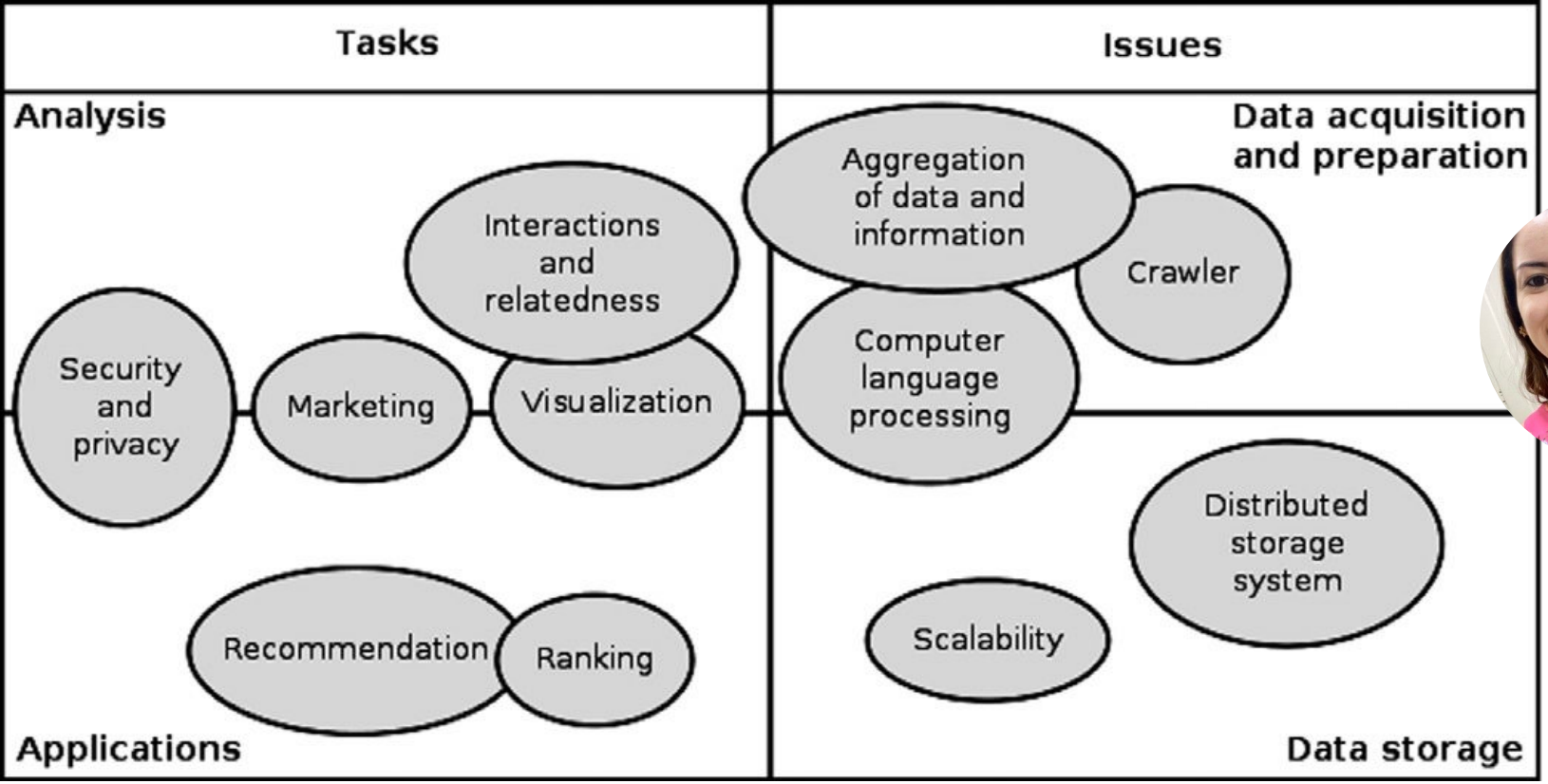


Fig. 5. Main social networks topics, in which **tasks** refer to **using** social networks to solve problems, and **issues** address problems related to **managing** social networks.

“Tese de doutorado tem de produzir survey em periódico”

Social

Professional

Networks

Coautoria + PPG Software

“Às vezes a oportunidade realmente bate a sua porta”

Ouvir

Quem você orienta

2017

Laís Mota (M)

Ouvir

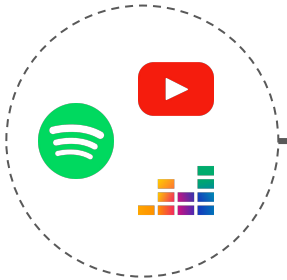
Quem você orienta
2017

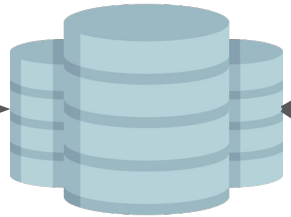
Laís Mota (M)

Mariana O S Silva (M)

Ciência de Dados na indústria da Música

Music-related Data





Music Genre Classification

Recommender Systems

Success Analysis

Music Platforms

Complex Data

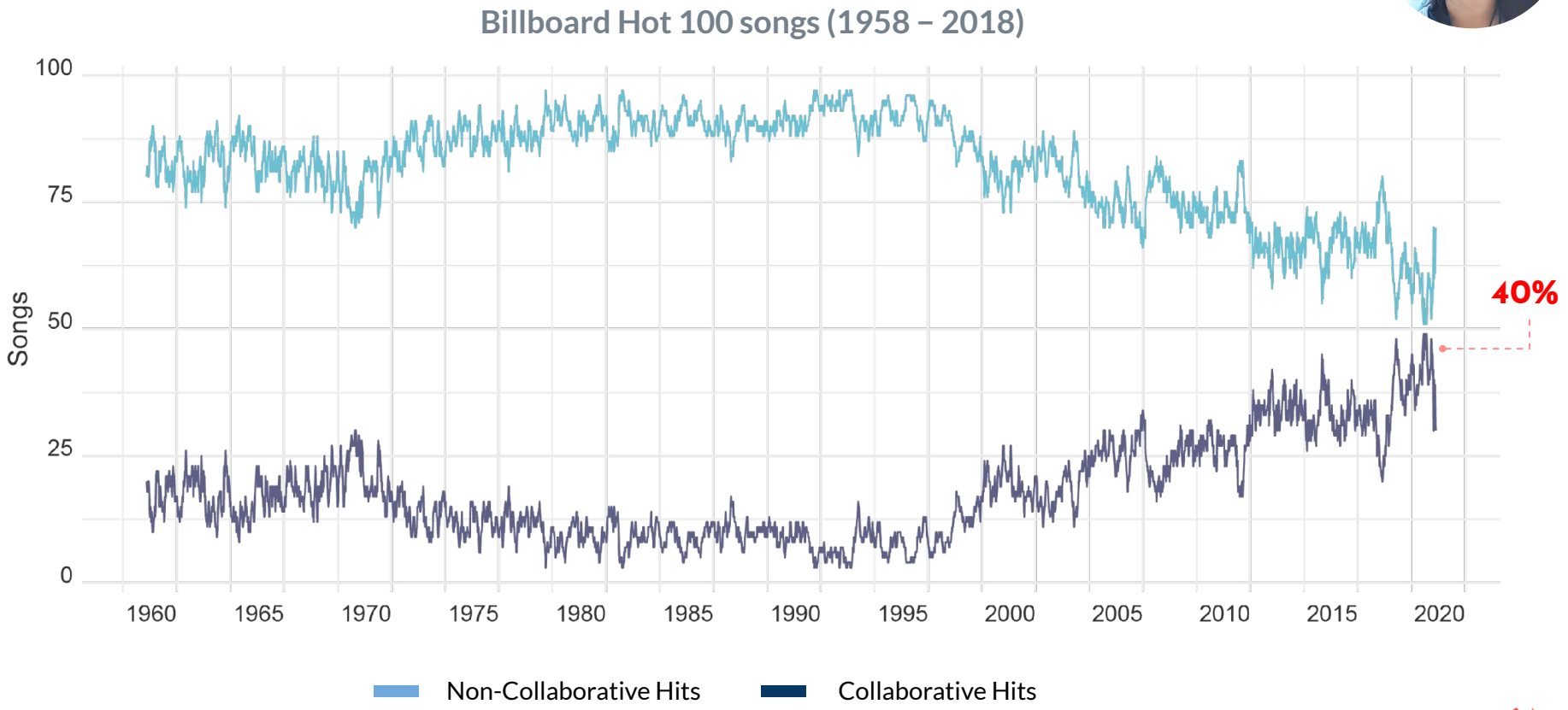
MIR Tasks

Sucesso

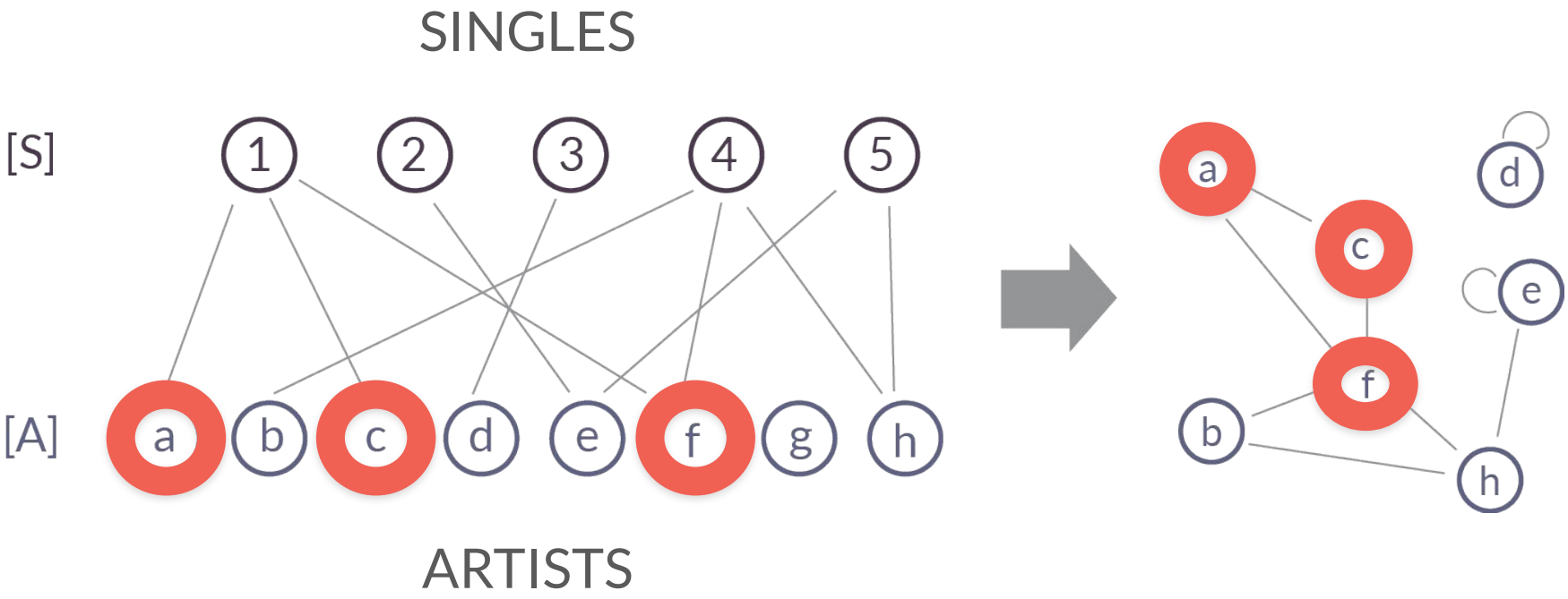
Sucesso depende dos dados

Colaboração

Collaborative Hits



Social Network Modeling



HIT SONG PREDICTION

Binary Classification

Input	*Features*	*Classify*
Song →	*acoustics* *lyrics* *listeners* *social media* *metadata* *social context* ...	→ ✔ *Hit* ✖ *Non-hit*

Hit = Billboard Hot 100

Non-hit = artists' songs ∉ BH100

MULTI-PERSPECTIVE **FEATURES**

Song

Album

Artist

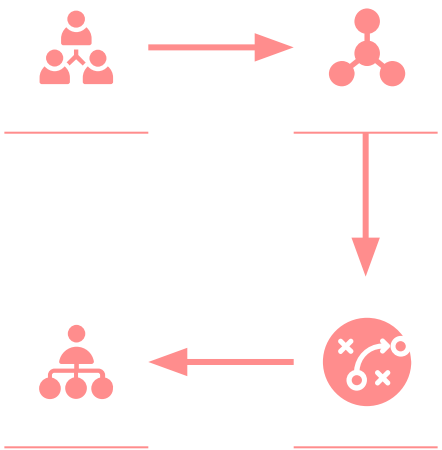
COLLABORATION-BASED **FEATURE**

Collaboration Network

Weighted edges connect artists who have collaborated

Collaboration Profiles

Well-defined artists groups



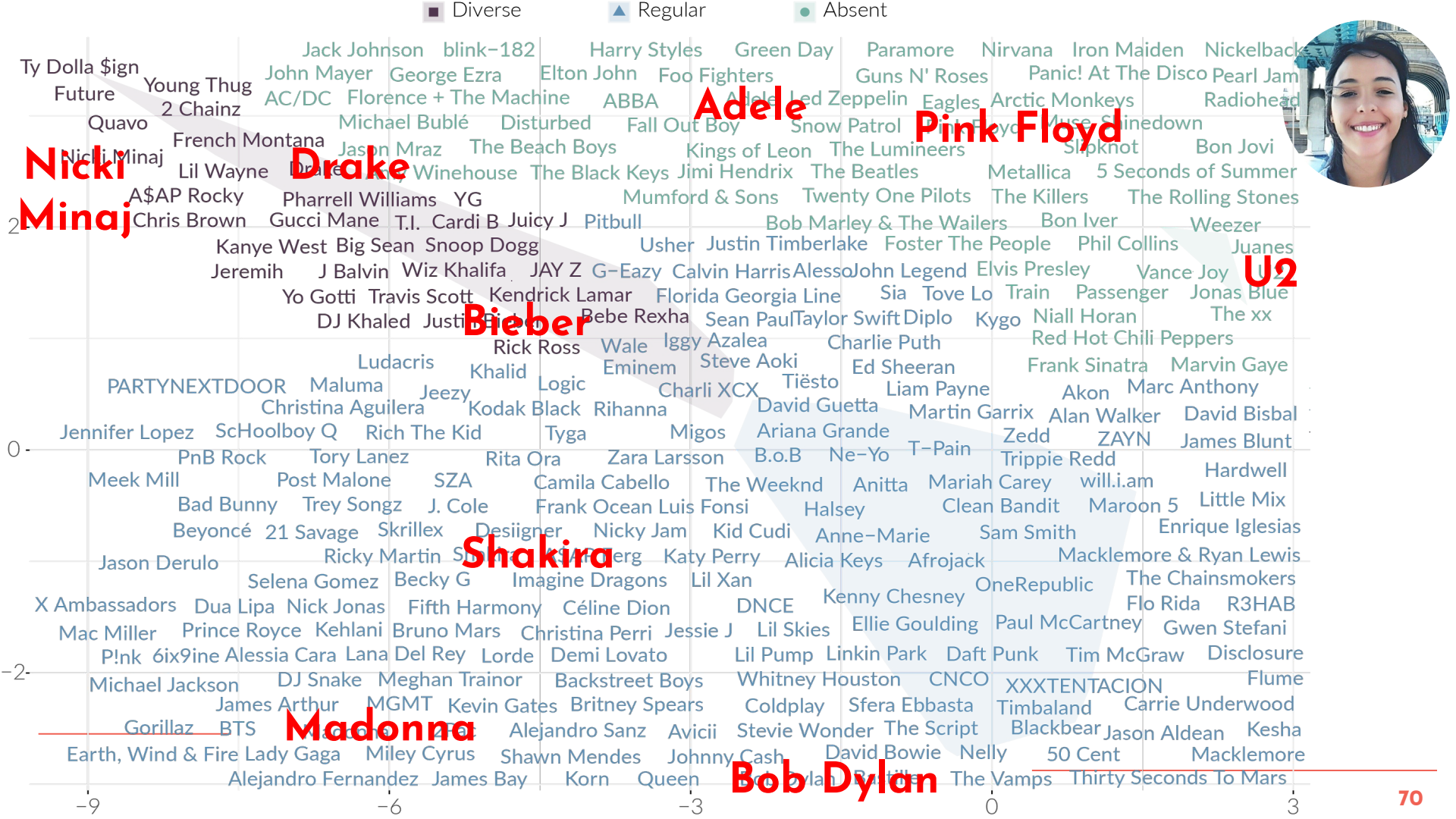
Topological Metrics

Compute seven topological metrics

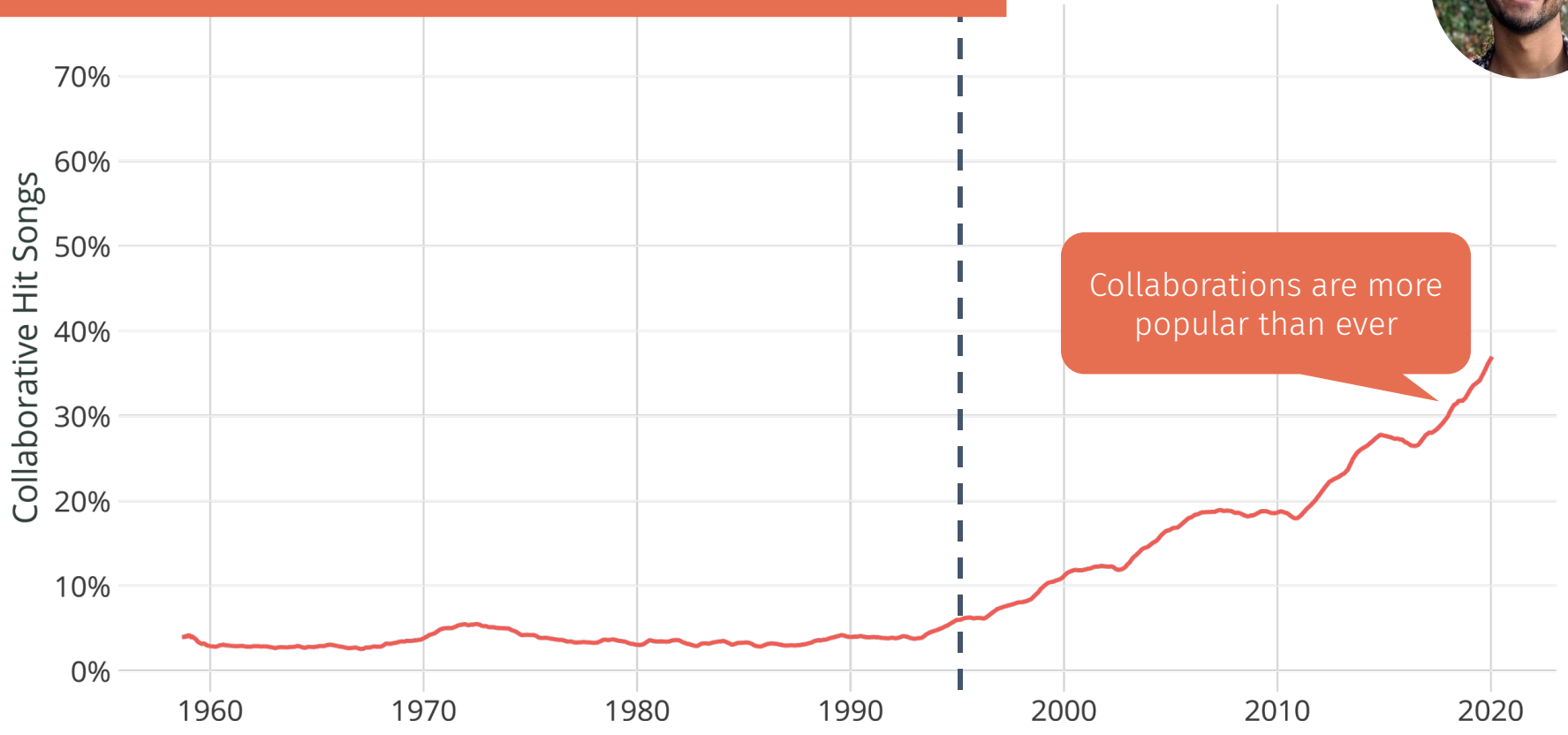
Clustering (K-Means)

Clustering similar artists, based on collaboration patterns

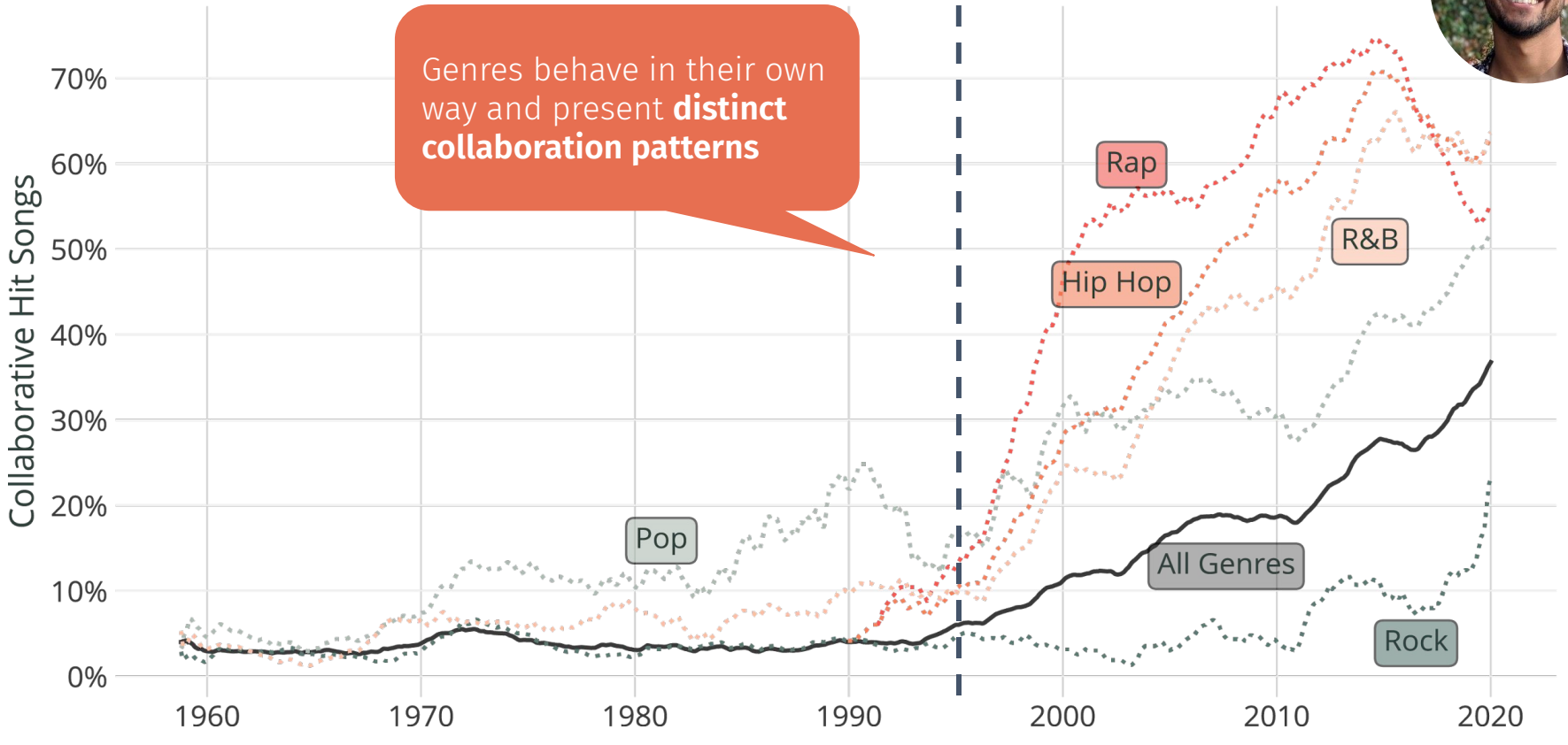

Diverse
Regular
Absent
Ty Dolla $ign
Future
Young Thug
2 Chainz
Quavo
Nicki Minaj
French Montana
Lil Wayne
Drake
A$AP Rocky
Pharrell Williams
YG
Chris Brown
Gucci Mane
T.I.
Cardi B
Juicy J
Kanye West
Big Sean
Snoop Dogg
Jeremih
J Balvin
Wiz Khalifa
JAY Z
Yo Gotti
Travis Scott
Kendrick Lamar
DJ Khaled
Justin Bieber
Rick Ross
Ludacris
Khalid
PARTYNEXTDOOR
Maluma
Jeezy
Christina Aguilera
Jennifer Lopez
ScHoolboy Q
PnB Rock
Tory Lanez
Meek Mill
Post Malone
Bad Bunny
Trey Songz
Beyoncé
21 Savage
Jason Derulo
Ricky Martin
Skrillex
Shakira
Selena Gomez
Becky G
X Ambassadors
Dua Lipa
Nick Jonas
Mac Miller
Prince Royce
Kehlani
P!nk
6ix9ine
Alessia Cara
Lana Del Rey
Michael Jackson
DJ Snake
Meghan Trainor
James Arthur
MGMT
Gorillaz
BTS
Madonna
Earth, Wind & Fire
Lady Gaga
Miley Cyrus
Alejandro Fernandez
James Bay
Jack Johnson
blink-182
John Mayer
George Ezra
AC/DC
Florence + The Machine
Michael Bublé
Disturbed
Jason Mraz
The Beach Boys
Amy Winehouse
The Black Keys
Mumford & Sons
Kanye West
Big Sean
Snoop Dogg
Jeremih
J Balvin
Wiz Khalifa
JAY Z
Yo Gotti
Travis Scott
Kendrick Lamar
DJ Khaled
Justin Bieber
Bebe Rexha
Rick Ross
Wale
Ludacris
Khalid
Eminem
PARTYNEXTDOOR
Maluma
Jeezy
Logic
Christina Aguilera
Kodak Black
Rihanna
Jennifer Lopez
ScHoolboy Q
Rich The Kid
Tyga
PnB Rock
Tory Lanez
Rita Ora
Meek Mill
Post Malone
SZA
Camila Cabello
Bad Bunny
Trey Songz
J. Cole
Frank Ocean
Beyoncé
21 Savage
Skrillex
Desiigner
Nicky Jam
Jason Derulo
Ricky Martin
Shakira
A$AP Ferg
Selena Gomez
Becky G
Imagine Dragons
X Ambassadors
Dua Lipa
Nick Jonas
Fifth Harmony
Céline Dion
Mac Miller
Prince Royce
Kehlani
Bruno Mars
Christa Perri
P!nk
6ix9ine
Alessia Cara
Lana Del Rey
Lorde
Demi Lovato
Michael Jackson
DJ Snake
Meghan Trainor
Backstreet Boys
James Arthur
MGMT
Kevin Gates
Britney Spears
Gorillaz
BTS
Madonna
2Pac
Alejandro Sanz
Avicii
Earth, Wind & Fire
Lady Gaga
Miley Cyrus
Shawn Mendes
Johnny Cash
Alejandro Fernandez
James Bay
Korn
Queen
Bob Dylan
Harry Styles
Green Day
Paramore
Nirvana
Iron Maiden
Nickelback
Elton John
Foo Fighters
Guns N' Roses
Panic! At The Disco
Pearl Jam
ABBA
Adele
Led Zeppelin
Eagles
Arctic Monkeys
Radiohead
Fall Out Boy
Snow Patrol
Pink Floyd
Muse
Shinedown
Kings of Leon
The Lumineers
Slipknot
Bon Jovi
Amy Winehouse
The Black Keys
Jimi Hendrix
The Beatles
Metallica
5 Seconds of Summer
Mumford & Sons
Twenty One Pilots
The Killers
The Rolling Stones
Pitbull
Bob Marley & The Wailers
Bon Iver
Weezer
Usher
Justin Timberlake
Foster The People
Phil Collins
Juanes
G-Eazy
Calvin Harris
Alesso
John Legend
Elvis Presley
Vance Joy
U2
Florida Georgia Line
Sia
Tove Lo
Train
Passenger
Jonas Blue
Bebe Rexha
Sean Paul
Taylor Swift
Diplo
Kygo
Niall Horan
The xx
Iggy Azalea
Charlie Puth
Red Hot Chili Peppers
Wale
Steve Aoki
Ed Sheeran
Frank Sinatra
Marvin Gaye
Eminem
Tiësto
Liam Payne
Akon
Marc Anthony
Charli XCX
David Guetta
Martin Garrix
Alan Walker
David Bisbal
Migos
Ariana Grande
Zedd
ZAYN
James Blunt
B.o.B
Ne-Yo
T-Pain
Trippie Redd
Rita Ora
Zara Larsson
Hardwell
The Weeknd
Anitta
Mariah Carey
will.i.am
Camila Cabello
Frank Ocean
Luis Fonsi
Halsey
Clean Bandit
Maroon 5
Little Mix
Nicky Jam
Kid Cudi
Anne-Marie
Sam Smith
Enrique Iglesias
Shakira
A$AP Ferg
Katy Perry
Alicia Keys
Afrojack
Macklemore & Ryan Lewis
Imagine Dragons
Lil Xan
OneRepublic
The Chainsmokers
Fifth Harmony
Céline Dion
DNCE
Kenny Chesney
Flo Rida
R3HAB
Bruno Mars
Christa Perri
Jessie J
Lil Skies
Ellie Goulding
Paul McCartney
Gwen Stefani
Lorde
Demi Lovato
Lil Pump
Linkin Park
Daft Punk
Tim McGraw
Disclosure
Backstreet Boys
Whitney Houston
CNCO
XXXTENTACION
Flume
Kevin Gates
Britney Spears
Coldplay
Sfera Ebbasta
Timbaland
Carrie Underwood
Alejandro Sanz
Avicii
Stevie Wonder
The Script
Blackbear
Jason Aldean
Kesha
Shawn Mendes
Johnny Cash
David Bowie
Nelly
50 Cent
Macklemore
Bastille
The Vamps
Thirty Seconds To Mars
2
0
-2
-9
-6
-3
0
3
69


Diverse
Regular
Absent
Ty Dolla $ign
Future
Young Thug
2 Chainz
Quavo
French Montana
Nicki Minaj
Lil Wayne
A$AP Rocky
Chris Brown
Nicki Minaj
Drake
Jack Johnson
blink-182
John Mayer
George Ezra
AC/DC
Florence + The Machine
Michael Bublé
Jason Mraz
Pharrell Williams
YG
Gucci Mane
T.I.
Cardi B
Juicy J
Kanye West
Big Sean
Snoop Dogg
Jeremih
J Balvin
Wiz Khalifa
JAY Z
Yo Gotti
Travis Scott
DJ Khaled
Bieber
Rick Ross
Ludacris
Khalid
PARTYNEXTDOOR
Maluma
Jeezy
Christina Aguilera
Jennifer Lopez
ScHoolboy Q
PnB Rock
Tory Lanez
Meek Mill
Post Malone
SZA
Bad Bunny
Trey Songz
J. Cole
Beyoncé
21 Savage
Skrillex
Jason Derulo
Ricky Martin
Selena Gomez
Becky G
X Ambassadors
Dua Lipa
Nick Jonas
Mac Miller
Prince Royce
Kehlani
P!nk
6ix9ine
Alessia Cara
Lana Del Rey
Michael Jackson
DJ Snake
James Arthur
MGMT
Gorillaz
BTS
Earth, Wind & Fire
Lady Gaga
Alejandro Fernandez
James Bay
Miley Cyrus
Madonna
Harry Styles
Green Day
Elton John
Foo Fighters
ABBA
Adele
Led Zeppelin
Fall Out Boy
Disturbed
The Beach Boys
Amy Winehouse
The Black Keys
Jimi Hendrix
Mumford & Sons
Twenty One Pilots
Pitbull
Usher
Justin Timberlake
Foster The People
Calvin Harris
Alesso
Florida Georgia Line
Bebe Rexha
Sean Paul
Iggy Azalea
Wale
Eminem
Steve Aoki
Logic
Charli XCX
Kodak Black
Rihanna
Tyga
Migos
Rita Ora
Zara Larsson
Camila Cabello
Frank Ocean
Luis Fonsi
Desiigner
Nicky Jam
Kid Cudi
Shakira
Katy Perry
Imagine Dragons
Lil Xan
Céline Dion
DNCE
Bruno Mars
Christa Perri
Jessie J
Lil Pump
Demi Lovato
Lorde
Meghan Trainor
Backstreet Boys
Kevin Gates
Britney Spears
Alejandro Sanz
Avicii
Shawn Mendes
Johnny Cash
Korn
Queen
Bob Dylan
Paramore
Nirvana
Iron Maiden
Nickelback
Guns N' Roses
Panic! At The Disco
Pearl Jam
Eagles
Arctic Monkeys
Radiohead
Adele
Pink Floyd
Snow Patrol
Muse
Shinedown
Kings of Leon
The Lumineers
Slipknot
Bon Jovi
Metallica
5 Seconds of Summer
The Beatles
The Killers
The Rolling Stones
Bob Marley & The Wailers
Bon Iver
Weezer
Phil Collins
Juanes
John Legend
Elvis Presley
Vance Joy
U2
Sia
Tove Lo
Train
Passenger
Jonas Blue
Taylor Swift
Diplo
Kygo
Niall Horan
The xx
Red Hot Chili Peppers
Charlie Puth
Ed Sheeran
Frank Sinatra
Marvin Gaye
Liam Payne
Akon
Marc Anthony
Tiësto
David Guetta
Martin Garrix
Alan Walker
David Bisbal
Ariana Grande
Zedd
ZAYN
James Blunt
B.o.B
Ne-Yo
T-Pain
Trippie Redd
Hardwell
Anitta
Mariah Carey
will.i.am
The Weeknd
Halsey
Clean Bandit
Maroon 5
Little Mix
Anne-Marie
Sam Smith
Enrique Iglesias
Alicia Keys
Afrojack
Macklemore & Ryan Lewis
OneRepublic
The Chainsmokers
Kenny Chesney
Flo Rida
R3HAB
Ellie Goulding
Paul McCartney
Gwen Stefani
Linkin Park
Daft Punk
Tim McGraw
Disclosure
Whitney Houston
CNCO
XXXTENTACION
Flume
Coldplay
Sfera Ebbasta
Timbaland
Carrie Underwood
The Script
Blackbear
Jason Aldean
Kesha
David Bowie
Nelly
50 Cent
Macklemore
The Vamps
Thirty Seconds To Mars
2
0
-2
-9
-6
-3
0
3

Collaboration is a powerful way to reach new audiences and increase popularity



Source: Billboard Hot 100 Charts



Source: Billboard Hot 100 Charts

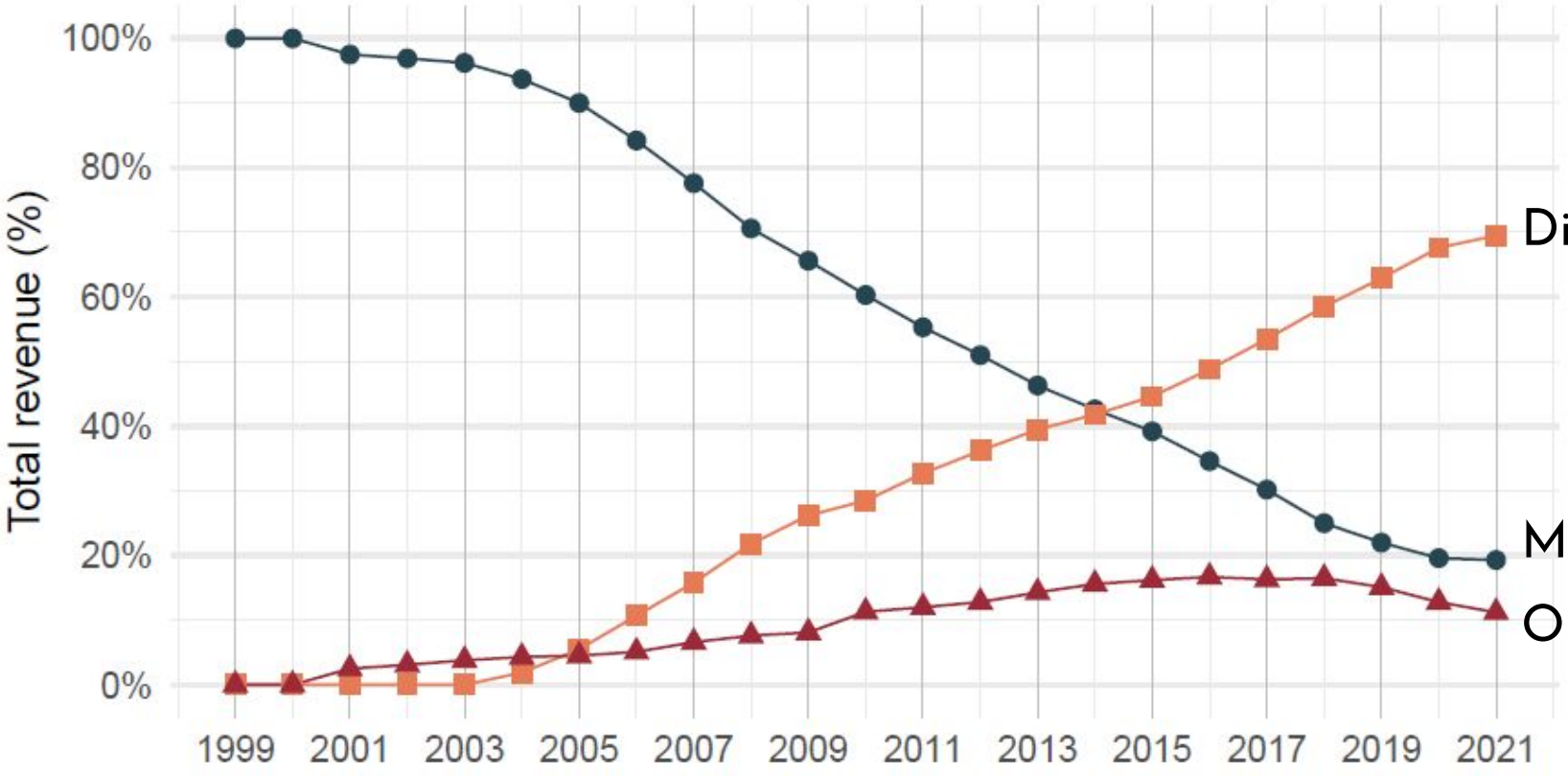
Brazilians have a strong preference for local artists regardless of era

SBCM 2020 Best

Hot Streaks in the Brazilian Music Market: A Comparison Between Physical and Digital Eras

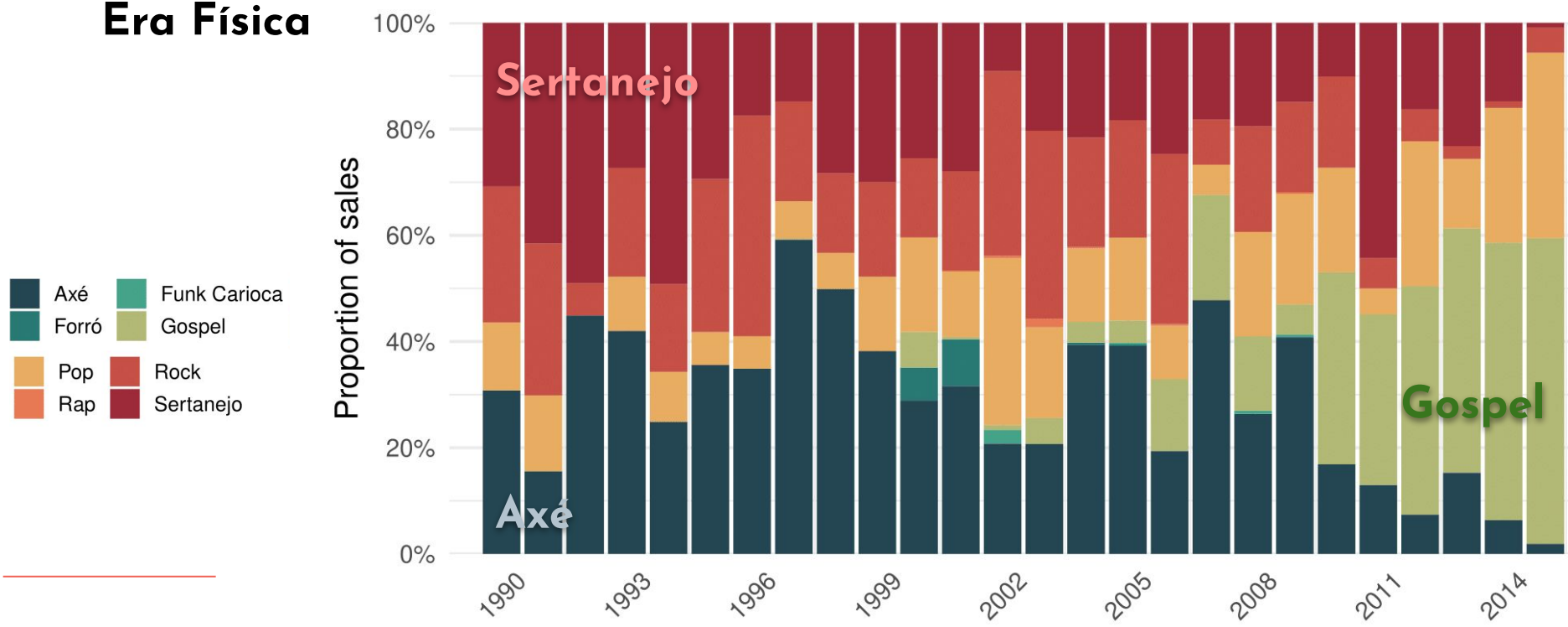
Gabriel R. G. Barbosa, Bruna C. Melo, Gabriel P. Oliveira, Mariana O. Silva, Danilo B. Seufitelli

Vendas Globais



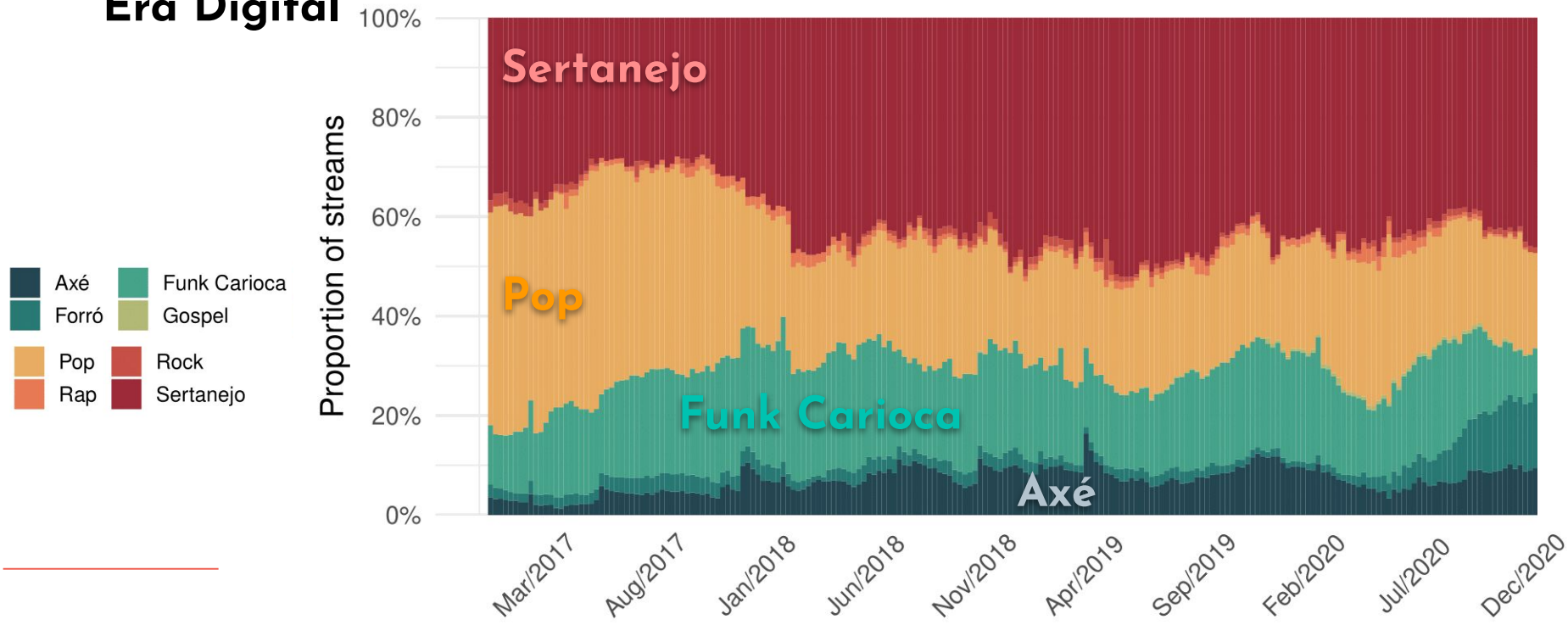
Evolução Gêneros

Era Física



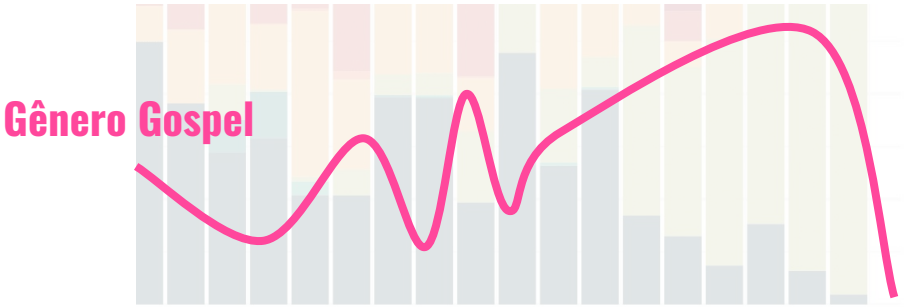
Evolução Gêneros

Era Digital



Evolução Gêneros

Consumo do Gênero Gospel na Era Digital



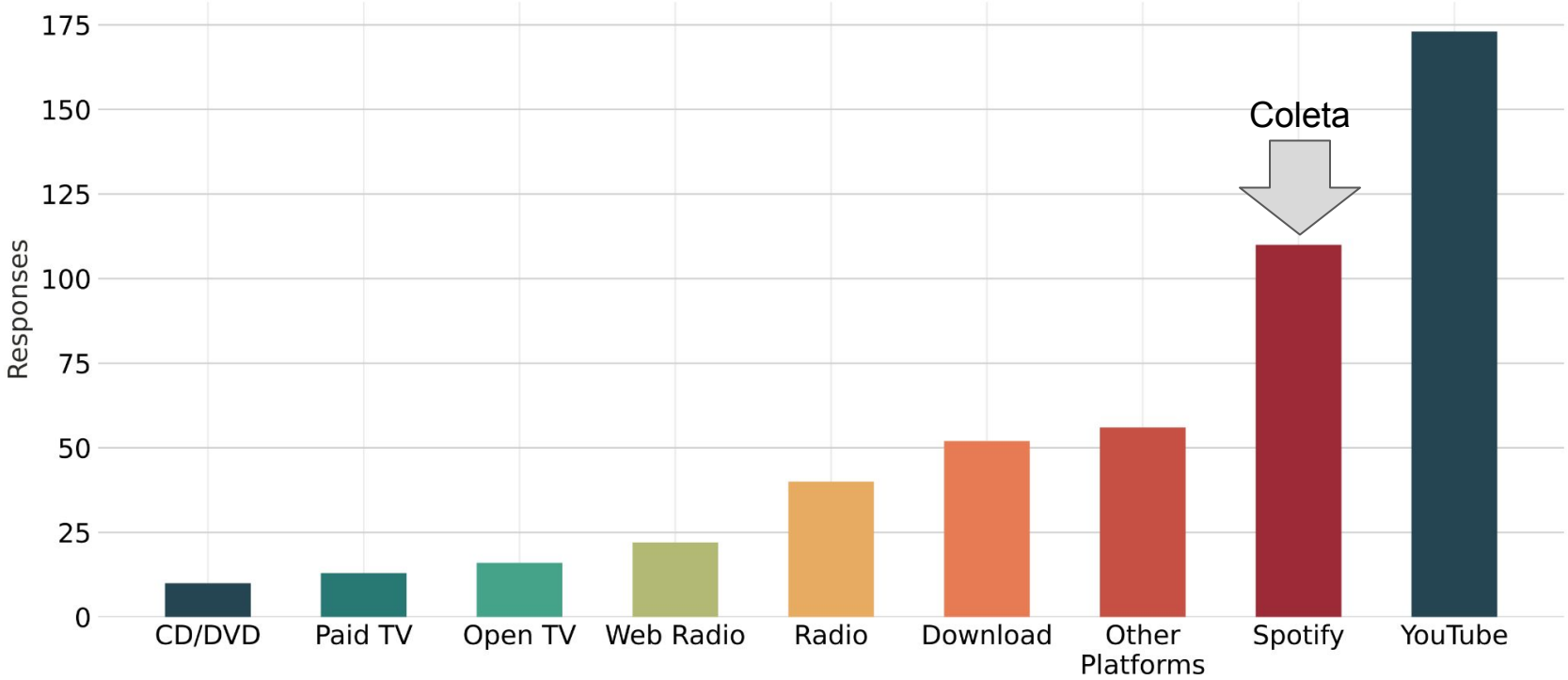
Era Física

Era Digital

“Algo inusitado nos dados? ‘Bora lá investigar”

Evolução Gêneros

Consumo do Gênero Gospel na Era Digital



Sucesso + Gênero + Colaborações

"Exploration to Exploitation": SBBD 2023

Sucesso@ Tempo, Gênero (musical), Colaboração

- WTDBD @ SBBD 2020
- Short @ SBBD 2022
- CTDBD 2023

Datasets

- DSW@SBBD 2019 MusicOSet
- DSW@SBBD 2021 → JIDM 2022 MUHSIC
- DSW@SBBD 2023 MGD+

Juliana

Sucesso + Gênero + Colaborações

Rede de Colaboração: ACM SAC 2019
WebMedia 2019
ISMIR 2020 (Best Paper Presentation)
Hot Streaks @Brazil: SBCM 2021 (Best Paper) → Vórtex 2022
@US: Scientometrics 2023 (sept.23)
Hit Song Prediction: WebMedia 2022 → JIS 2023
Análise de Humor: CTIC @ WebMedia 2022
Descoberta de Padrões: BraSNAM 2023; KDMiLe 2023
Temporal Analyses: Vórtex 2023
Survey on Hit Song Science: Journal 2023? (forever!)

“Fazer ciência é muito mais do que publicar artigo científico”

Ciência Aberta + Ensino

CoMusic - 2019 @ **Zenodo**

MusicOSet - 2019 @ **Zenodo**

Music Genre Dataset (MGD) - 2020

2855 views and 1869 downloads @ **Zenodo**

Music-oriented Hot Streak Information Collection (MUHSIC) - 2021

425 views and 88 downloads @ **Zenodo**

MGD+ 2023 @ **Zenodo**

Ciência de Dados usando Jupyter (com nossos Datasets)

Capítulo de Livro + Minicurso @ JAI 2021

Minicurso @ SBBD 2021, Escola de Inverno PPGC/UFF

(Ciência de Dados)

Computação e Cultura

UFMG

DCC
DEPARTEMANETO DE
CIÊNCIA DA COMPUTAÇÃO

CS+X

CNPq

PPORTAL:

Public Domain Portuguese-language Literature Dataset

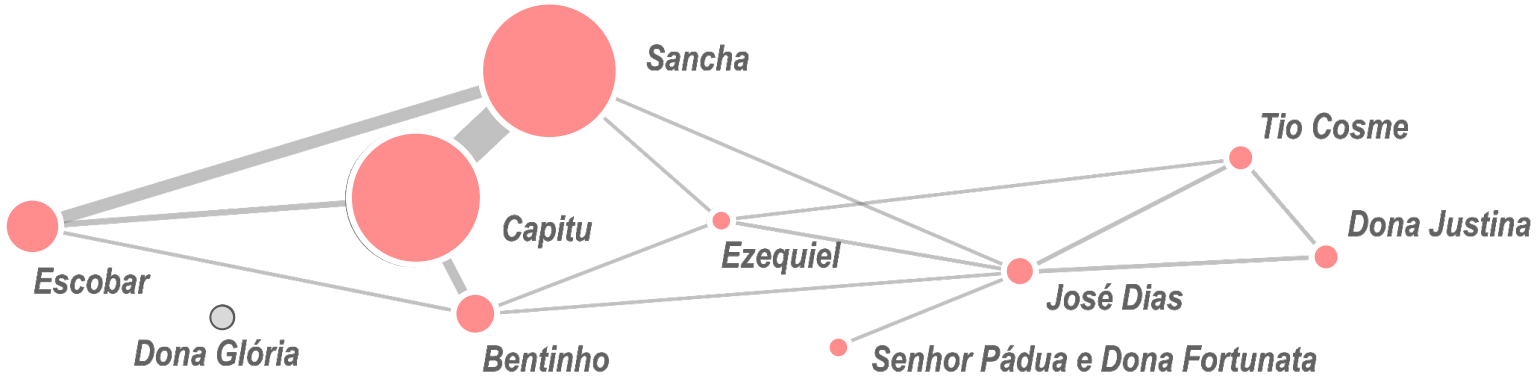
Mariana O. Silva mariana.santos@dcc.ufmg.br
Clarisse Scofield clarissescofield@dcc.ufmg.br
Mirella M. Moro mirella@dcc.ufmg.br

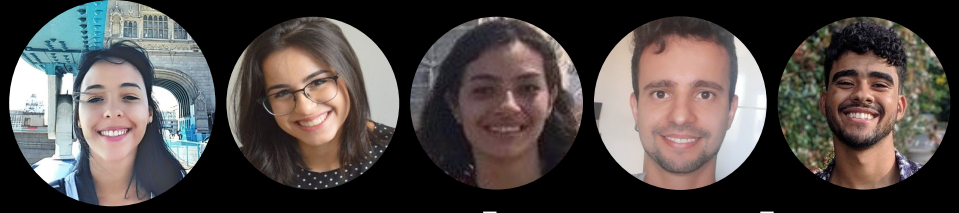
Best DSW 2021

CHARACTER NETWORK - DOM CASMURRO

Nodes → characters

Edges → interaction





Mariana, Clarisse, Luiza, Danilo, Gabriel

Ciência de Dados e Literatura

Gender Representation in Literature: Analysis of Characters' Physical Descriptions - KDMiLe 2023

Analyzing Character Networks in Portuguese-language Literary Works - BraSNAM 2023

Book Genre Classification Based on Reviews of Portuguese-Language Literature - PROPOR 2022

Cross-collection Dataset of Public Domain Portuguese-language Works - JIDM 2022

PPORTAL: Public Domain Portuguese-language Literature Dataset - SBBD DSW 2021 BEST

Brazilian Reading Preferences in Goodreads: Cross-state and Cross-region Analyses - iSys 2022

Exploring Brazilian Cultural Identity Through Reading Preferences - BraSNAM 2021 BEST Runner Up

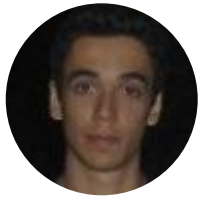
Dilúvio de Dados

Uma Boia à Vista,
ou seria um Cruzeiro?

Ao mesmo tempo que...

Engenharia e Ciência de Dados

Carolina Bigonha, 2012
Análise de Sentimento



Pedro O Santos, 2015
Geo + NoSQL

Levy Silva, 2018
Indexação + Deduplicação

Harlley Lima, 2014
Coautoria

Luciana Maroun, 2016
Ranking

Natália Machado, 2018
Recomendação

Wladston Ferreira, 2015
Cinema

Thiago Prado, 2017
Recomendação POI

André Gonzaga, 2019
Knowledge Graphs

Pôster/Short@SBBD CTDBD Full e Short@SBBD

“Fazer ciência é muito mais do que publicar artigo científico”

Banqueira de Dados

2002-2006 Revisor Externo SBBD
2008-atual Membro Comitê de Programa SBBD
    Membro Comitê de Programa SBBD Componentes (NEXT)
    Banca: WTDBD n vezes, CTDBD 2023, Best Paper
2010 Coordenadora Organização
2010-2012 Editora Associada JIDM
10/2013 - 10/2017 Comitê Diretivo do SBBD / CEBD
    2013 Coord. Tutoriais
    2014 Coord. Comitê de Programa
    2015 Coord. CEBD
2017 (Proponente), 2019, 2021-2023 Dataset Showcase Workshop
2008 Coord. Sessão de Pôsteres
2009 Demos Banca Avaliadora

Banqueira de Dados
Comitê de Programa

	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Full		X	X		X	X	X	X	X	X	X	X	X	X	X	X	X
Pôster|Short	X	X	X				X		X	X	X	X	X	X	X	X	X
Demo|App	X	X	X		X				X		X	X	X				
WTDBD			X		X	X	X	X			X	X	X				X
CTDBD									X		X		X		X		X
DSW											X		X		X	X	X

Banqueira de Dados

Submissão e Autoria (*Rejeição*+Publicação)

	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23
FULL		*2*										*1*			*1*		
Pôster|Short																	
Demo|App				*1*							*1*						
WTDBD												*1*					
WTAG																	
CTDBD	-	-	-	-	-	-	-	-		-		-		-	*MM*	-	
DSW																	

Banqueira de Dados

Submissão e Autoria (*Rejeição*+Publicação)

	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23
FULL		*2*							*1*+**1**			*1*			*1*		*1*+**1**
Pôster|Short					*2*+**2**		*2*+**1**				*1*+**4**						
Demo|App				*1*							*1*						
WTDBD												*1*					
WTAG																	
CTDBD	-	-	-	-	-	-	-	-		-		-	*d*+**M**	-	*MM*	-	
DSW																	

Banqueira de Dados

Submissão e Autoria (*Rejeição*+Publicação)

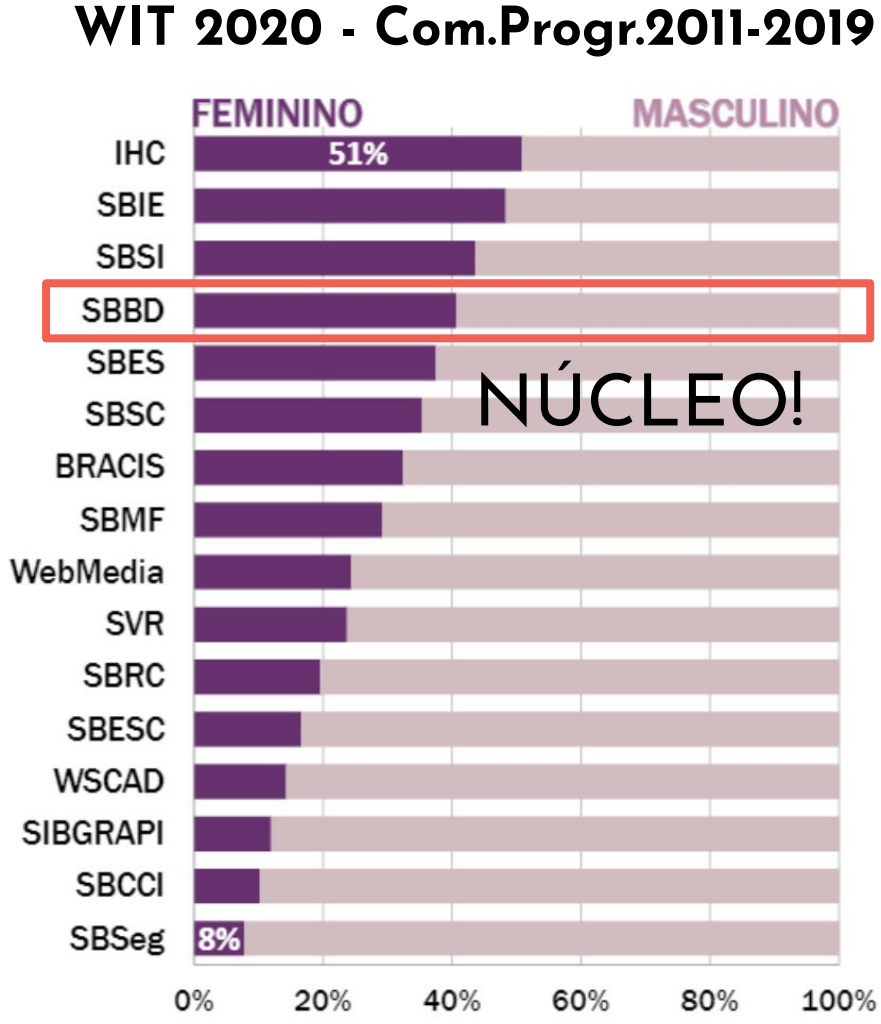
	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23
FULL		*2*		**1**		**1**			*1*+**1**		**2**	*1*	**1**		*1*		*1*+**1**
Pôster|Short			**1**	**1**	*2*+**2**	**4**	*2*+**1**	**1**	**3**	**4**	*1*+**4**	**1**			**2**	**1**	
Demo|App				*1*	**1**		**2**		**4**		*1*						
WTDBD		**1**										*1*	**1**	**1**			
WTAG																1	
CTDBD	-	-	-	-	-	-	-	-	**MM**	-		-	*d*+**M**		*MM*		**M**
DSW													**1**		**2**		**1**

Banqueira de Dados

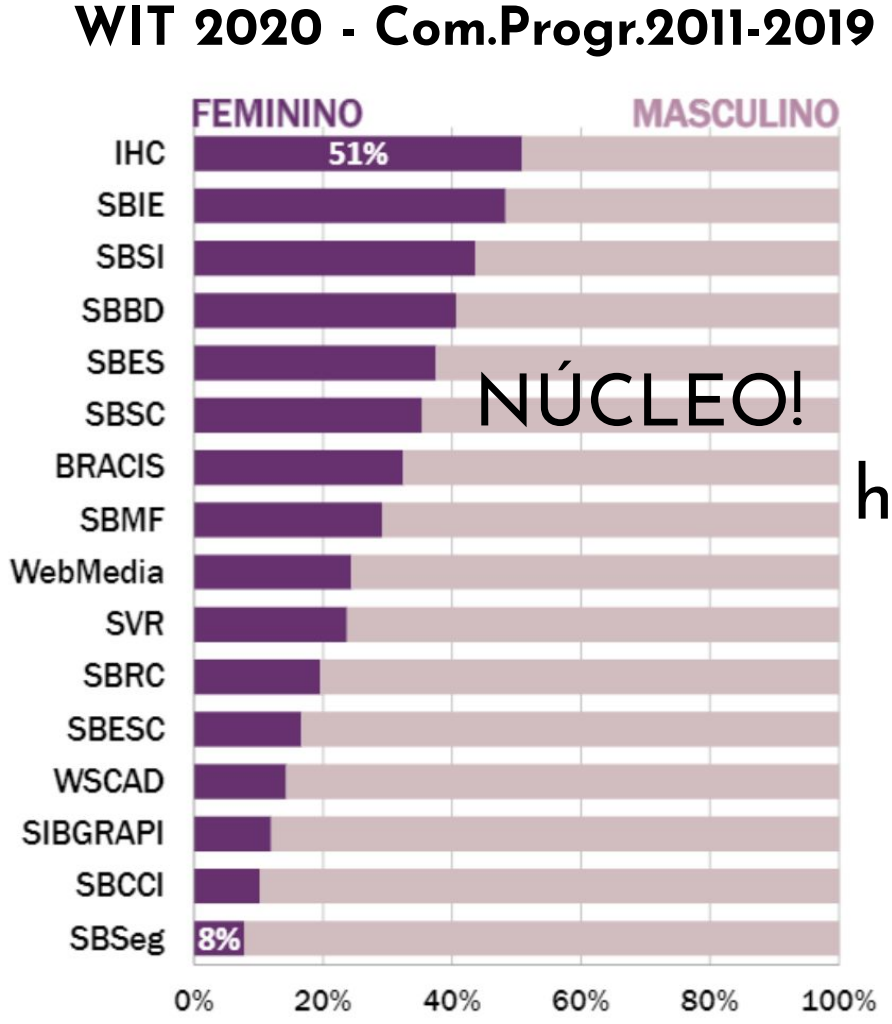
BEST

> 2023 MH DSW: MGD+
> 2021 DSW: PPORTAL
> 2017 Short: GitHub
> 2017 Runner Up: STACY
> 2013 Short@JIDM
> 2012 Short@JIDM
> 2011 Menção Honrosa Demo
> 2010 Short@JIDM
> Best Reviewer

Diversidade & Computação



Diversidade & Computação



https://linktr.ee/lackofdiversity

Best @ EduComp 2022

Women in dataBases

2008 *foto* Campinas

2009
Fortaleza

2010
BH

2011
Florianópolis

2012 São Paulo

2014 Curitiba

2015
Petrópolis


2016
Salvador

2022
Buzios

WOMEN IN DATABASES

WomB

SBBD 2023

IFMG-PUCMG

#gratidão

Dilúvio de Dados

Uma Boia à Vista, ou seria um Cruzeiro?

MIRELLA M. MORO
UFMG
bit.ly/profamirella

SBBD 2023

Créditos
Imagens: PixaBay
Fotos: acervo pessoal
Template: Macari Company Profile by Slidesgo