



Os grafos de genealogia acadêmica dos programas de Pós-graduação em Ciência da Computação do Brasil

Jesús P. Mena-Chalco (UFABC)
jesus.mena@ufabc.edu.br

São Carlos, 14 de Setembro de 2016



PPGs em Ciência da Computação no Brasil

Programas de Pós-graduação

Em 2016 foram avaliados*:

- 66 programas acadêmicos.
- 11 programas profissionalizantes.

PROGRAMA	INST.	Área Básica	M	D
ENGENHARIA DE SISTEMAS E COMPUTAÇÃO	UFRJ-PESC	SISTEMA DE COMPUTAÇÃO	7	7
INFORMÁTICA	PUC-RJ	CIÊNCIA DA COMPUTAÇÃO	7	7
CIÊNCIAS DA COMPUTAÇÃO	UFMG	CIÊNCIA DA COMPUTAÇÃO	7	7
CIÊNCIA DA COMPUTAÇÃO	UNICAMP	CIÊNCIA DA COMPUTAÇÃO	7	7
COMPUTAÇÃO	UFRGS	CIÊNCIA DA COMPUTAÇÃO	7	7
CIÊNCIAS DA COMPUTAÇÃO	UFPE	CIÊNCIA DA COMPUTAÇÃO	6	6
CIÊNCIA DA COMPUTAÇÃO	USP-BUTANTA	CIÊNCIA DA COMPUTAÇÃO	6	6
CIÊNCIAS DA COMPUTAÇÃO E MATEMÁTICA COMPUTACIONAL	USP-SC	CIÊNCIA DA COMPUTAÇÃO	6	6
INFORMÁTICA	UFAM	CIÊNCIA DA COMPUTAÇÃO	5	5
CIÊNCIAS DA COMPUTAÇÃO	UFC	CIÊNCIA DA COMPUTAÇÃO	5	5
SISTEMAS E COMPUTAÇÃO	UFRN	CIÊNCIA DA COMPUTAÇÃO	5	5
COMPUTAÇÃO	UFF	CIÊNCIA DA COMPUTAÇÃO	5	5
CIÊNCIA DA COMPUTAÇÃO	PUC-RS	CIÊNCIA DA COMPUTAÇÃO	5	5

(*) Dados extraídos da Plataforma Sucupira

Programas de Pós-graduação

Art. 1º - O corpo docente dos programas desse nível de ensino é composto por 3 (três) categorias de docentes:

I - docentes permanentes, constituindo o núcleo principal de docentes do programa;

II - docentes visitantes;

III - docentes colaboradores.

Art. 2º - Integram a categoria de permanentes os docentes enquadrados, **declarados e relatados anualmente pelo Programa de Pós Graduação (PPG) na plataforma Sucupira** e que atendam a todos os seguintes pré-requisitos:

(...)

Portaria CAPES nº 174, de 30/12/2014, que define as categorias de docentes dos PPG's como Permanente, Visitante e Colaborador

Programas de Pós-graduação

PPG	Professores
UFPE	67
USP-SC	66
UFRGS	53
UNICAMP	48
UFMG	45
UFF	44
UFBA	38
UFRJ-PESC	38
UFSCAR	35
USP-BUTANTA	34
UFABC	33
UFCEG	33
UFSC	32
UFG	31
UFSC	31
UFPR	30
UFRJ-INFORMATICA	26
UFU	25
PUC-RJ	24
PUC-RS	24
UFC	24
UFRN	24

PPG	Professores
UFES	23
UFPEL	23
UNB	23
FURG	22
UFAM	22
UFMS	21
UFPB	20
UNESP	20
UNIFESP	19
UFJF	16
UFPA	16
UFS	16
UNIFOR	16
UNIRIO	16
UERN-UFERSA	15
USP-LESTE	15
PUC-PR	14
UDESC	14
UNISINOS	14
UFMA	13
UFOP	13
UFRPE	13

PPG	Professores
USP-RP	13
CEFET-RJ	12
IFCE	12
UNIFEI	12
UPE	12
UTFPR	12
FACCAMP	11
IME-RJ	11
UEM	11
UFAL	11
UFSCAR-SOROCABA	11
UFV	11
UNIFACS	11
PUC-MG	10
UFPI	10
UNIVALI	10
UECE	9
UEL	9
UFLA	9
UFMS-UFG	4
UFBA-UEFS	2
UFBA-UNIFACS-UEFS	1

1403 professores

Programas de Pós-graduação

USP-SC (ICMC)

Adenilso Da Silva Simao
Afonso Paiva Neto
Agma Juci Machado Traina
Alexandre Claudio Botazzo Delbem
Alneu De Andrade Lopes
Alysson Machado Costa
Andre Carlos P. L. Ferreira De Carvalho
Antonio Castelo Filho
Caetano Traina Junior
Claudio Fabiano Motta Toledo
Cristina Dutra De Aguiar Ciferri
Denis Fernando Wolf
Dilvan De Abreu Moreira
Edson Dos Santos Moreira
Eduardo Fontoura Costa
Eduardo Marques
Eduardo Raul Hruschka
Elaine Parros Machado De Sousa
Elias Salomao Helou Neto
Elisa Yumi Nakagawa
Ellen Francine Barbosa
Fabricio Simeoni De Sousa
Fernando Santos Osorio
Fernando Vieira Paulovich
Francisco Aparecido Rodrigues
Franklina Maria Bragion De Toledo

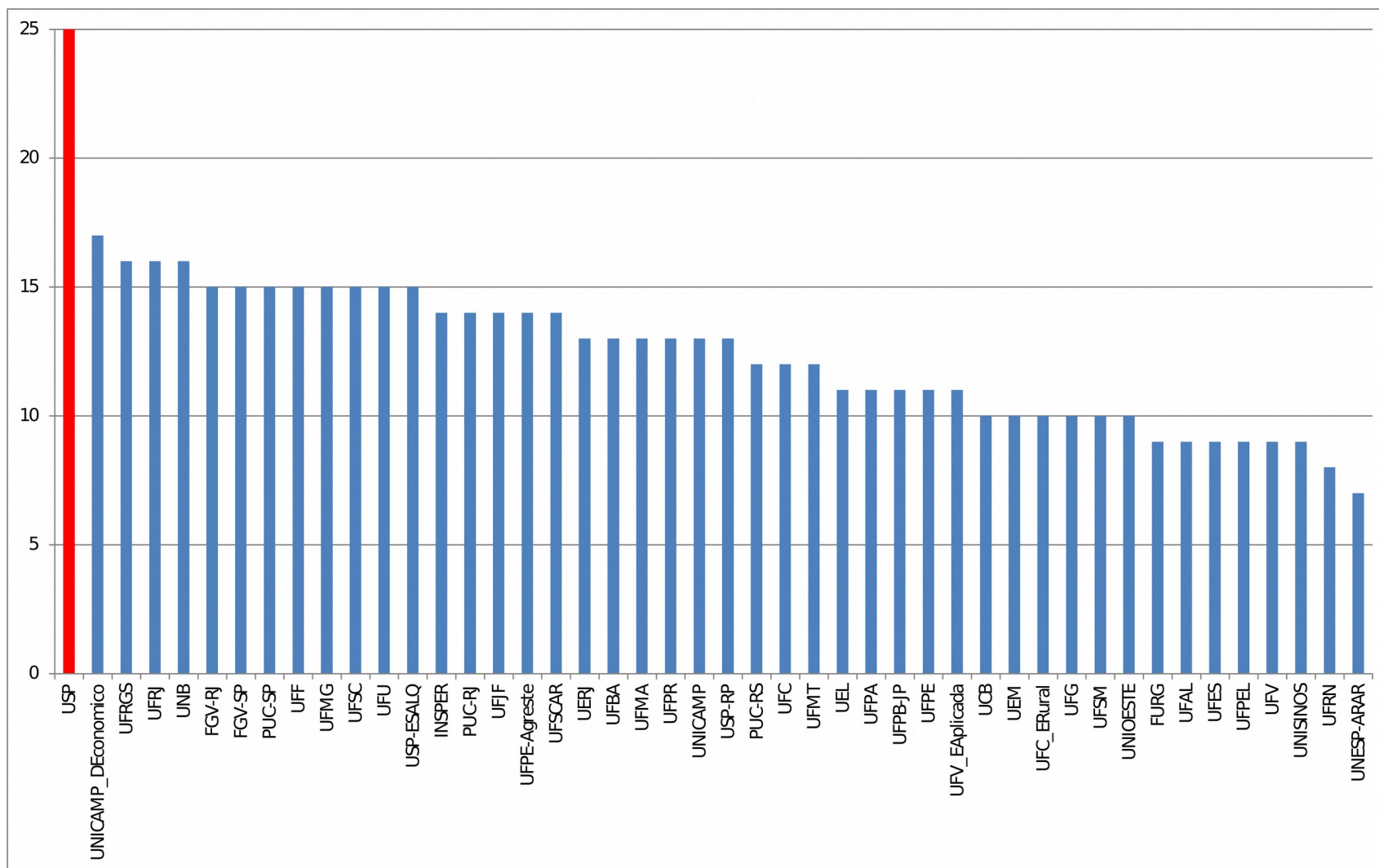
Gustavo Carlos Buscaglia
Gustavo Enrique De Almeida Prado Alves Batista
Jo Ueyama
Joao Do Espirito Santo Batista Neto
Joao Luis Garcia Rosa
Joao Porto De Albuquerque Pereira
Jose Alberto Cuminato
Jose Carlos Maldonado
Jose Fernando Rodrigues Junior
Julio Cezar Estrella
Kalinka Regina Lucas Jaquie Castelo Branco
Leandro Franco De Souza
Luis Gustavo Nonato
Marcelo Garcia Manzato
Marcio Eduardo Delamaro
Marcos Jose Santana
Maria Carolina Monard
Maria Cristina Ferreira De Oliveira
Maria Da Graca Campos Pimentel
Marina Andretta
Maristela Oliveira Dos Santos
Moacir Pereira Ponti Junior
Murilo Francisco Tome
Odemir Martinez Bruno
Paulo Cesar Masiero
Paulo Sergio Lopes De Souza

Regina Helena Carlucci Santana
Renata Pontin De Mattos Fortes
Ricardo Jose Gabrielli Barreto Campello
Rodrigo Fernandes De Mello
Rosana Teresinha Vaccare Braga
Rosane Minghim
Roseli Aparecida Francelin Romero
Rudinei Goularte
Sandra Maria Aluisio
Seiji Isotani
Simone Do Rocio Senger De Souza
Solange Oliveira Rezende
Thiago Alexandre Salgueiro Pardo
Vanderlei Bonato

Programas de Pós-graduação (Economia)

Área de avaliação: Economia (46 PPGs)

Docentes permanentes **janeiro de 2016**: 578 docentes



(*) Dados extraídos da Plataforma Sucupira

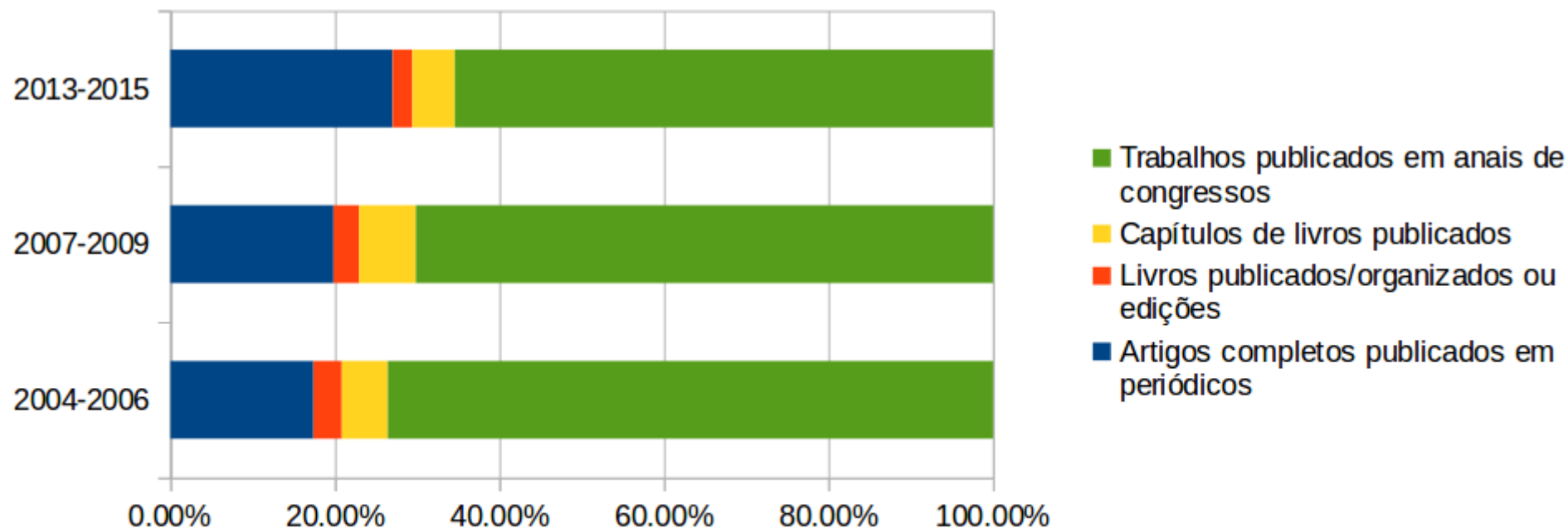
Perfil do produção bibliográfica (CC)

753

889

1403

	2004-2006	2007-2009	2013-2015
Artigos completos publicados em periódicos	1658	2344	4814
Livros publicados/organizados ou edições	329	366	426
Capítulos de livros publicados	533	816	922
Trabalhos publicados em anais de congressos	7015	8292	11657
Total	9535	11818	17819



Perfil de produção bibliográfica (CC)

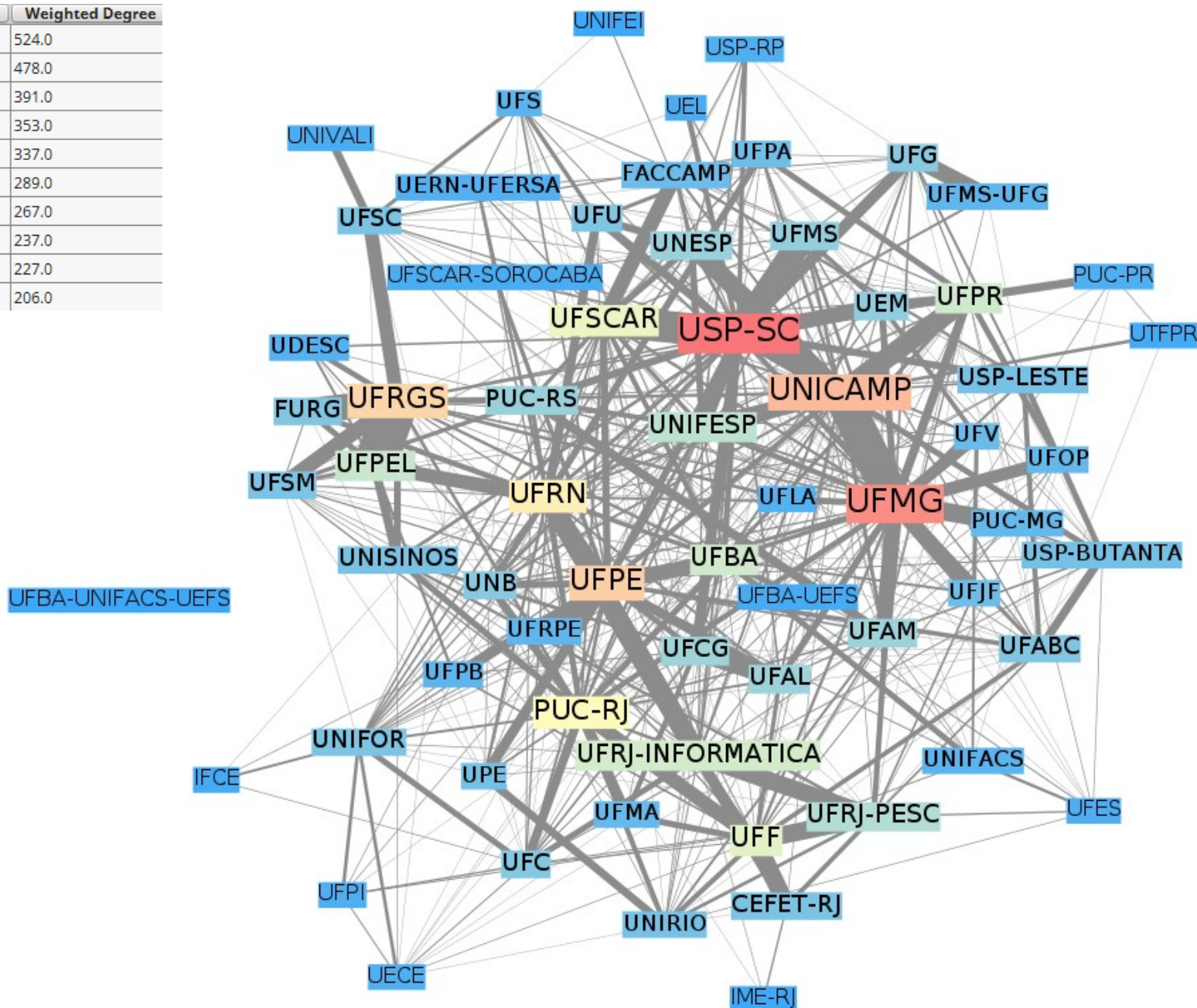
Journal	Publicações	2014
expert systems with applications	121	A2
renote. revista novas tecnologias na educação	90	C
revista ieee américa latina	84	B4
journal of information and data management - jidm	75	C
neurocomputing	60	A2
revista brasileira de informática na educação	54	B5
plos one	52	B2
discrete applied mathematics	50	B1
journal of the brazilian computer society	50	B2
journal of physics. conference series	49	C
revista de informática teórica e aplicada: rita	46	
the journal of systems and software	45	
revista brasileira de computação aplicada	42	
information sciences	39	
pattern recognition letters	36	
isys: revista brasileira de sistemas de informação	35	
procedia computer science	35	
sbc journal on 3d interactive systems	35	
engenharia de software magazine	32	
computers & operations research	31	
lecture notes in computer science	31	
information and software technology	30	
european journal of operational research	29	
theoretical computer science	29	
computer networks	28	
electronic notes in discrete mathematics	27	
science of computer programming	25	
journal of universal computer science	24	
clei electronic journal	23	
international journal of software engineering and knowledge engineering	23	
revista de sistemas de informação da fsma	23	
bmc bioinformatics	21	
genome announcements	21	
concurrency and computation	20	
electronic notes in theoretical computer science	19	
international transactions in operational research	19	
journal of software engineering research and development	19	

Professores bolsistas

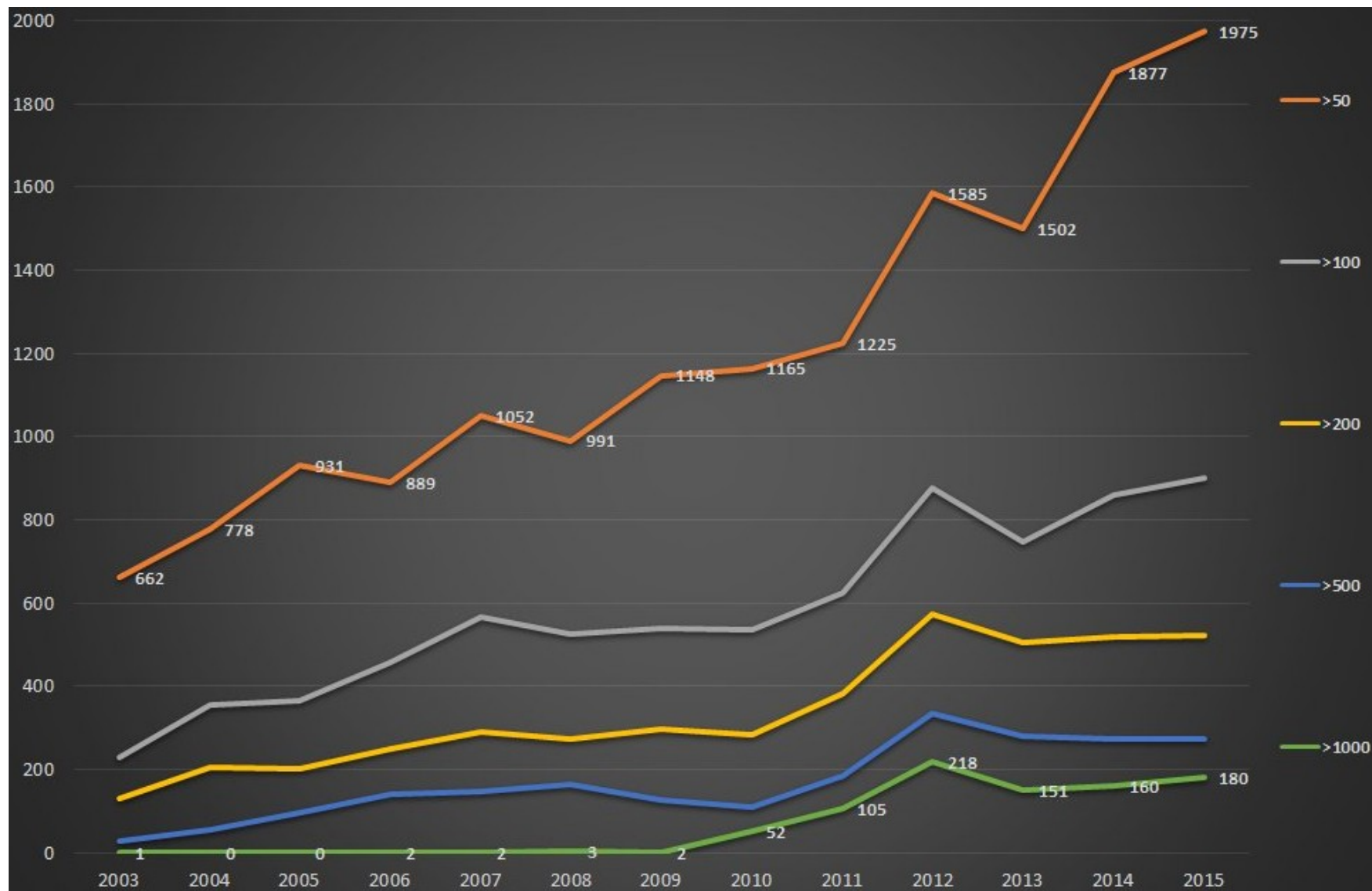
Bolsa	Professores	Total CC
Desen. Tec. - Nível 1C - CA 92 - Programa de Tecnologia da Informação e Comunicação	1	
Desen. Tec. - Nível 1D - CA 92 - Programa de Tecnologia da Informação e Comunicação	4	
Desen. Tec. - Nível 1D - CA 93 - Programa das Tecnologias Educacionais e Sociais	1	
Desen. Tec. - Nível 1D - CA 96 - Programa de Desenvolvimento Tecnológico e Industrial	1	
Desen. Tec. - Nível 2 - CA 82 - Programa de Tecnologias Médicas e da Saúde	2	
Desen. Tec. - Nível 2 - CA 88 - Programa de Tecnologias Ambientais	1	
Desen. Tec. - Nível 2 - CA 92 - Programa de Tecnologia da Informação e Comunicação	37	
Desen. Tec. - Nível 2 - CA 93 - Programa das Tecnologias Educacionais e Sociais	4	
Desen. Tec. - Nível 2 - CA 95 - Programa das Tecnologias nas áreas Aeronáutica e Aeroespacial	1	
Desen. Tec. - Nível 2 - CA 96 - Programa de Desenvolvimento Tecnológico e Industrial	2	
Desen. Tec. - Nível 2 - CA 97 - Programa das Tecnologias para o Desenvolvimento Sustentável	1	
Nível 1A - CA CC - Ciências da Computação	22	24
Nível 1A - CA EC - Engenharia Civil	1	
Nível 1A - CA ME - Microeletrônica	1	
Nível 1A - CA PE - Engenharia de Produção e de Transportes	2	
Nível 1B - CA CC - Ciências da Computação	20	22
Nível 1B - CA EE - Engenharia Elétrica e Biomédica	3	
Nível 1B - CA EM - Engenharia Mecânica; Naval e Oceânica e Aeroespacial	1	
Nível 1B - CA MA - Matemática e Estatística	1	
Nível 1C - CA CC - Ciências da Computação	32	36
Nível 1C - CA EE - Engenharia Elétrica e Biomédica	5	
Nível 1C - CA EM - Engenharia Mecânica; Naval e Oceânica e Aeroespacial	1	
Nível 1C - CA MA - Matemática e Estatística	2	
Nível 1C - CA PE - Engenharia de Produção e de Transportes	1	
Nível 1D - CA CC - Ciências da Computação	41	53
Nível 1D - CA EE - Engenharia Elétrica e Biomédica	4	
Nível 1D - CA FA - Física e Astronomia	1	
Nível 1D - CA ME - Microeletrônica	5	
Nível 1D - CA PE - Engenharia de Produção e de Transportes	3	
Nível 2 - CA AE - Administração; Contabilidade e Economia	1	
Nível 2 - CA BF - Biofísica; Bioquímica; Farmacologia; Fisiologia e Neurociências	2	
Nível 2 - CA CC - Ciências da Computação	222	269
Nível 2 - CA EC - Engenharia Civil	1	
Nível 2 - CA EE - Engenharia Elétrica e Biomédica	6	
Nível 2 - CA EM - Engenharia Mecânica; Naval e Oceânica e Aeroespacial	2	
Nível 2 - CA FA - Física e Astronomia	2	
Nível 2 - CA FI - Filosofia	1	
Nível 2 - CA GE - Genética	1	
Nível 2 - CA MA - Matemática e Estatística	4	
Nível 2 - CA ME - Microeletrônica	13	
Nível 2 - CA MM - Engenharia de Minas; Metalúrgica e de Materiais	1	
Nível 2 - CA PE - Engenharia de Produção e de Transportes	15	
Nível 2 - CA QU - Química	1	
Nível SR - CA CC - Ciências da Computação	1	1
Nível SR - CA MA - Matemática e Estatística	1	

Coautoria inter-programas

grupo	Weighted Degree
USP-SC	524.0
UFMG	478.0
UNICAMP	391.0
UFPE	353.0
UFRGS	337.0
UFRN	289.0
PUC-RJ	267.0
UFSCAR	237.0
UFF	227.0
UFRJ-INFORMATICA	206.0



Fenômeno: Hipercoautoria?



Avaliação de impacto científico

Os estudos cientométricos geralmente discutem a comunicação científica considerando aspectos como por exemplo:

- Produção bibliográfica.
- Colaboração (cointoria).
- Interdisciplinaridade.
- Citações.
- Divulgação científica.
- Senioridade, inatividade.
- Participação em projetos de pesquisa.
- Participação em bancas de concursos.

Raramente é avaliado o impacto na formação/projeção de novos recursos humanos.

Aqui a Ciência da Computação tem um papel importante.



Sobre Genealogia Acadêmica

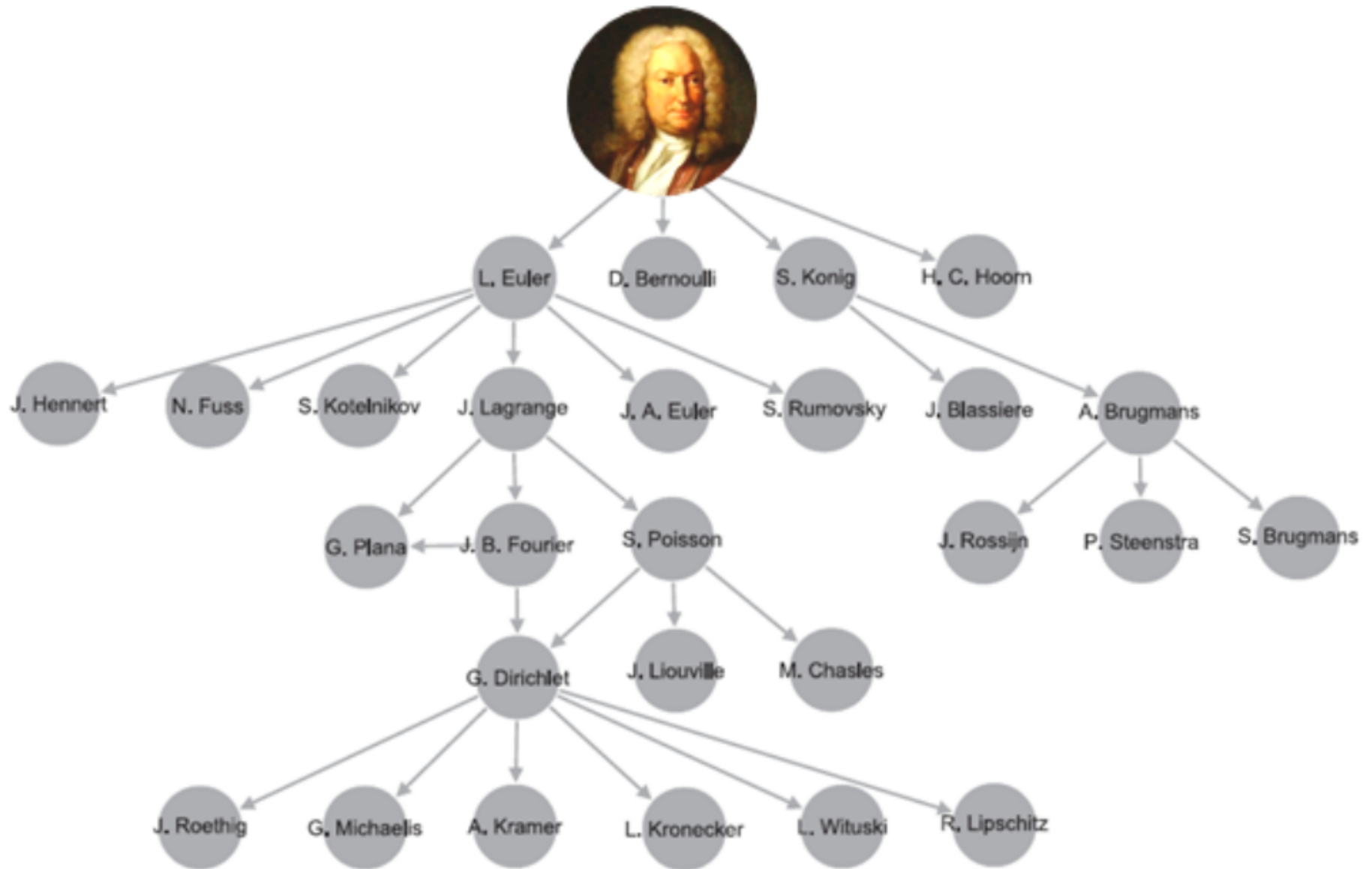
Genealogia acadêmica: Definição

É entendida como o **estudo da herança intelectual** perpetuada através dos relacionamentos entre orientadores e seus alunos.

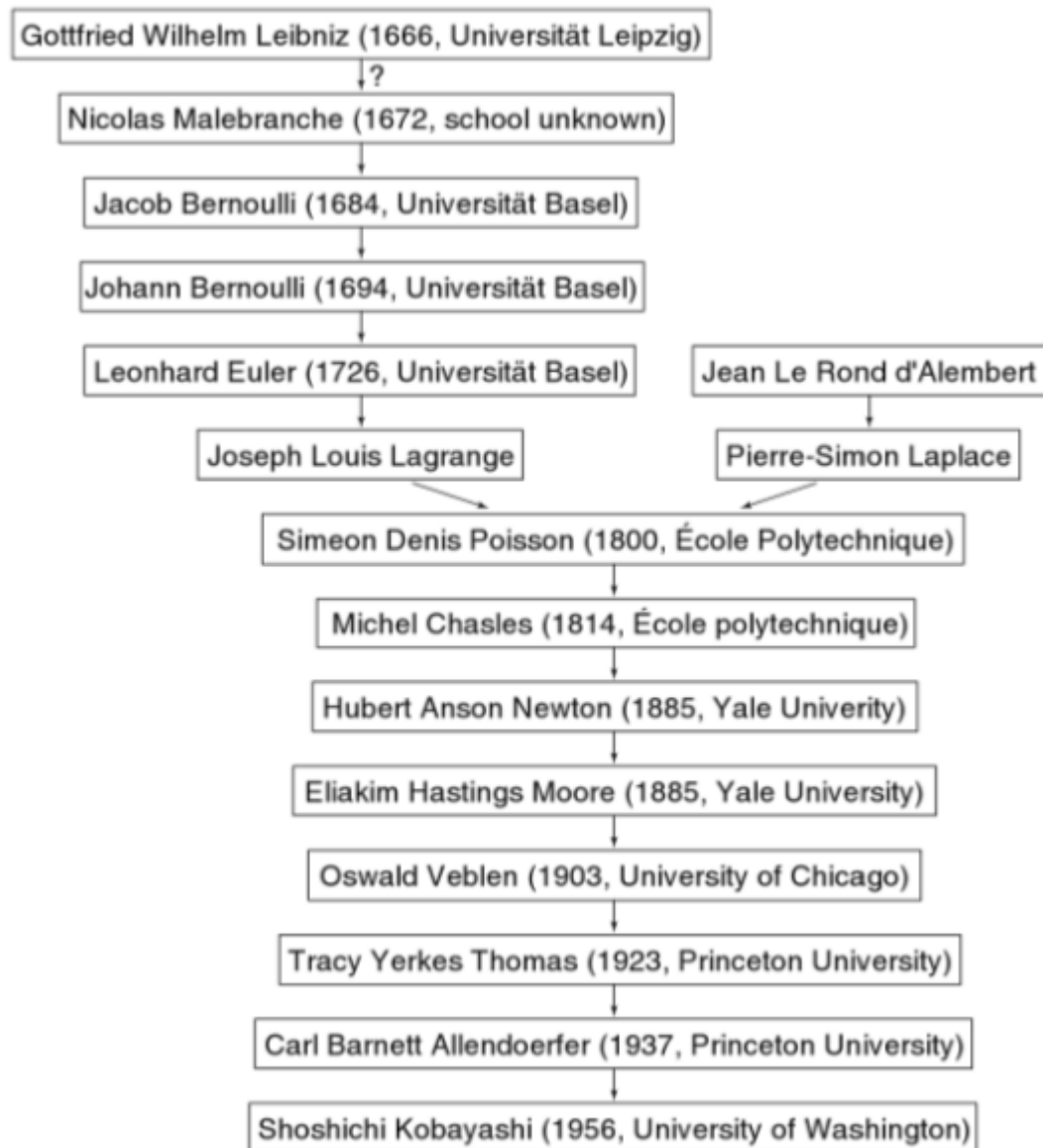
Segundo [Sugimoto \(2014\)](#), a genealogia permite:

- **Analisar a propagação de conhecimento científico.**
- A obtenção de insumos quantitativos para **mensurar as interações acadêmicas.**

Genealogia acadêmica



Genealogia acadêmica



G. Honorífica

Termos

- **Genealogia Acadêmica**

Usada para descrever as relações formais de orientação.

- **~~Genealogia Intelectual~~**

Usada para descrever influência intelectual.
(pode ser inclusive não formal / institucional)

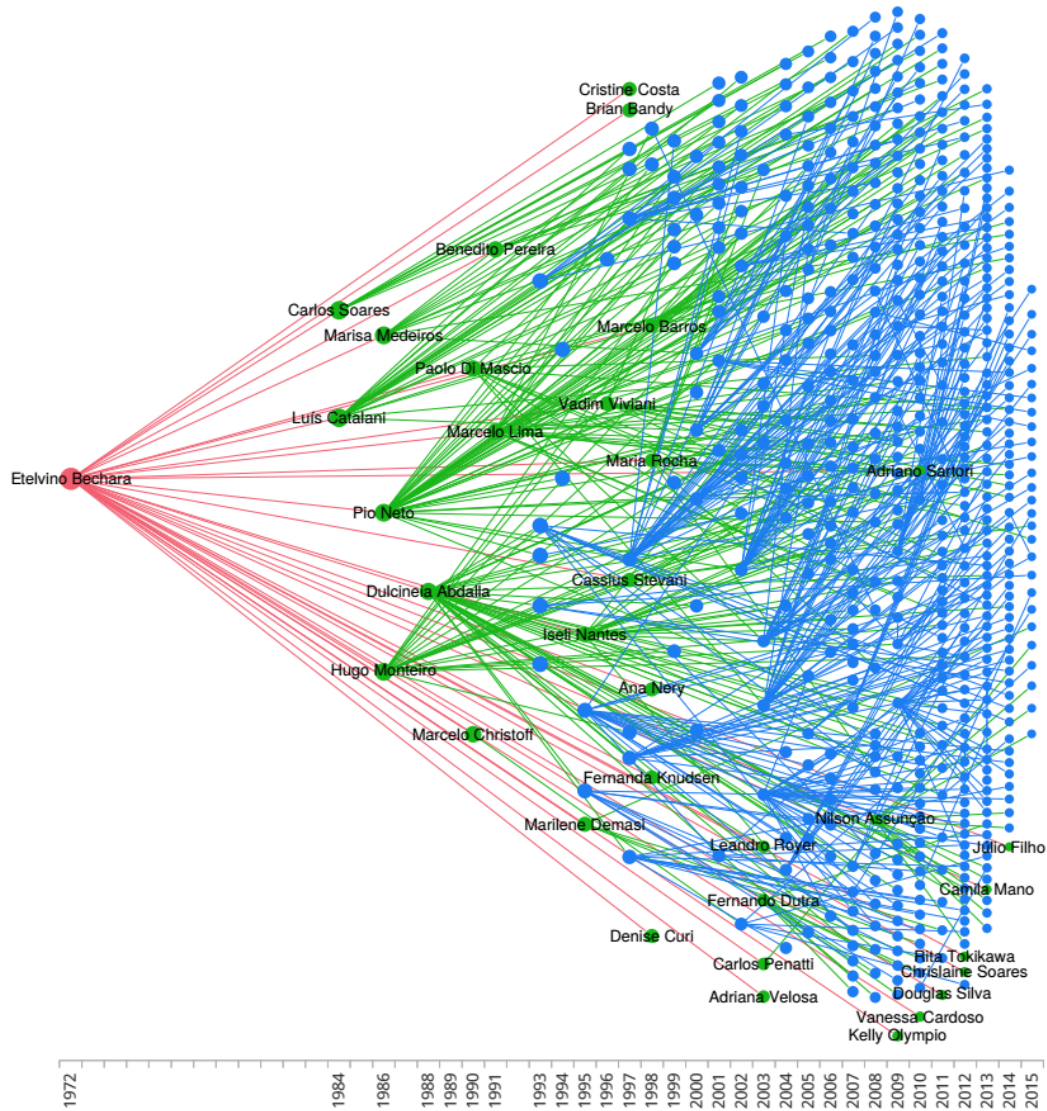
- **~~Genealogia Científica~~**

Usada para descrever estudos no nível de ciência.
(não necessariamente orientação formal)

Alguns trabalhos sobre genealogia

Área	Artigo
Neurocientistas	David, S. V., & Hayden, B. Y. (2012). Neurotree: A Collaborative, Graphical Database of the Academic Genealogy of Neuroscience . PloS one, 7 , e46608
Químicos	Andraos, J. (2005). Scientific genealogies of physical and mechanistic organic chemists . Canadian journal of chemistry, 83 , 1400–1414.
Matemáticos	Malmgren, R., Ottino, J., & Amaral, L. (2010). The role of mentorship in protégé performance . Nature, 465 , 622–626
Fisiólogos	Bennett, A. F., & Lowe, C. (2005). The academic genealogy of George A. Bartholomew . Integrative and comparative biology, 45 , 231–233.
Meteorologistas	Hart, R. E., & Cossuth, J. H. (2013). A Family Tree of Tropical Meteorology's Academic Community and its Proposed Expansion . Bulletin of the American Meteorological Society, 94 , 1837–1848.
Primatologists	Kelley, E. A., & Sussman, R. W. (2007). An academic genealogy on the history of American field primatologists . American journal of physical anthropology, 132, 406–425.
Bibliometristas e cientistas de informação.	Russell, T. G., & Sugimoto, C. R. (2009). MPACT family trees: Quantifying academic genealogy in library and information science . Journal of Education for Library and Information Science, pp. 248–262.
Protozoologistas brasileiros	Elias, M., Floeter-Winter, L. M., & Mena-Chalco, J. P. (2016). The dynamics of Brazilian protozoology over the past century . Memórias do Instituto Oswaldo Cruz, 111 , 67–74.

Genealogia do Prof. Etelvino Bechara (IQ/USP)



Mena-Chalco (2015). *A genealogia acadêmica do Prof. Etelvino José Henriques Bechara*. Relatório técnico: <http://dx.doi.org/10.13140/RG.2.1.1095.3685>




Método computacional para identificação dos grafos de genealogia

Eduardo Whitaker Bergamini (orientador:mestrado 1979-1983)



Mario Jino (orientador:doutorado 1985-1991)

Formação acadêmica/titulação

- 1985 - 1991** Doutorado em Engenharia Elétrica - Computação e Automação.
Universidade Estadual de Campinas, UNICAMP, Brasil.
Título: Critérios Potenciais Usos: Uma Contribuição ao Teste Estrutural de Software, Ano de obtenção: 1991.
Orientador: Mário Jino.
Bolsista do(a): Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, CAPES, Brasil.
Palavras-chave: Critérios e Estratégias de Testes; Ferramentas de Teste; Testes e Validação de Software; Teste estrutural.
Grande área: Ciências Exatas e da Terra
Setores de atividade: Qualidade e Produtividade; Informática.
- 1979 - 1983** Mestrado em Engenharia e Tecnologia Espaciais (Conceito CAPES 4).
Instituto Nacional de Pesquisas Espaciais, INPE, Brasil.
Título: UM SISTEMA DE CONTROLE E SUPERVISÃO DE BORDO", Ano de Obtenção: 1983.
Orientador:  EDUARDO WITAKER BERGAMINI.
Palavras-chave: Controle e Supervisão de Bordo; Critérios e Estratégias de Testes; Ferramentas de Teste; Protocolos de Comunicação; Sistemas de Computação; Testes e Validação de Software.
Grande área: Engenharias
Grande Área: Engenharias / Área: Engenharia Aeroespacial / Subárea: Sistemas Aeroespaciais / Especialidade: Satélites e Outros Dispositivos Aeroespaciais.
Grande Área: Ciências Exatas e da Terra / Área: Ciência da Computação / Subárea: Sistemas de Computação / Especialidade: Software Básico.
Setores de atividade: Aeronáutica e Espaço; Fabricação de Aparelhos e Equipamentos de Telecomunicação.
- 1976 - 1978** Especialização em Ênfase Em Computação Eletrônica.
Universidade de São Paulo, USP, Brasil.
- 1974 - 1978** Graduação em Engenharia Elétrica.
Universidade de São Paulo, USP, Brasil.

Orientações e supervisões concluídas

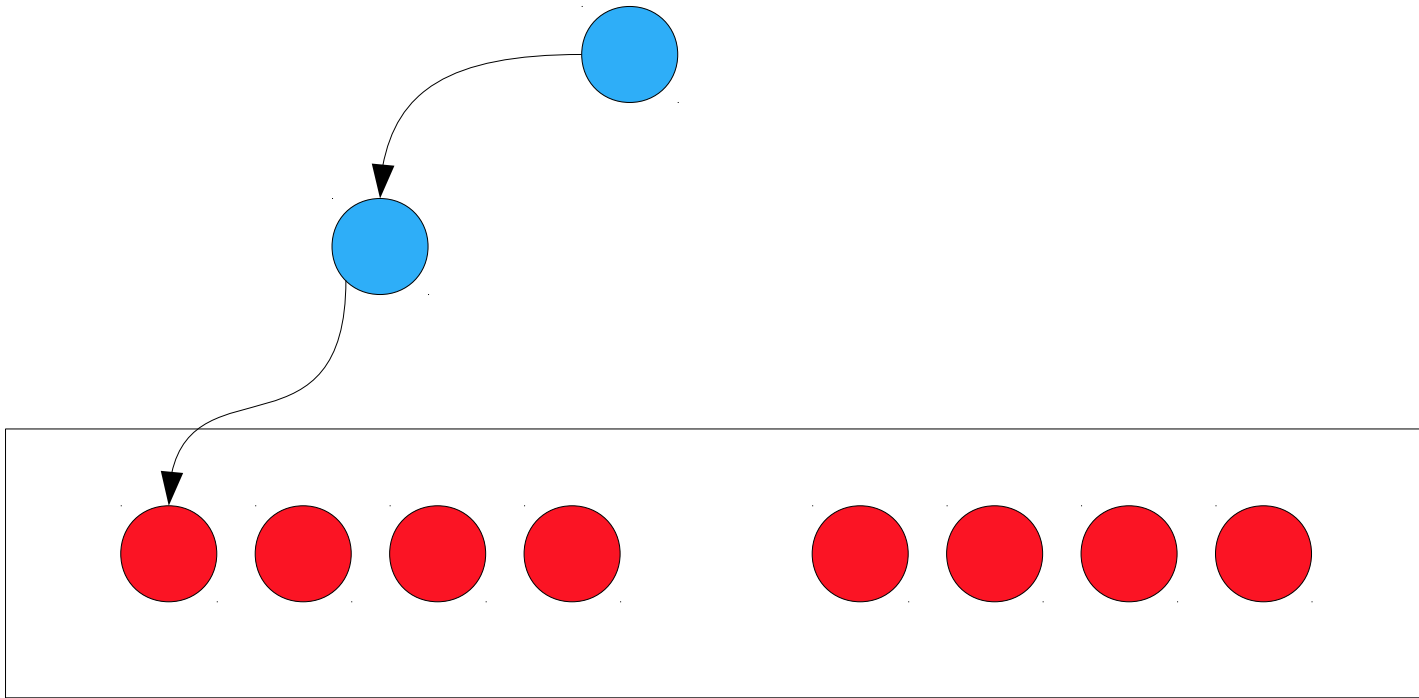
Dissertação de mestrado

1.  Rafael Messias Martins. Reengenharia da Ferramenta Projection Explorer para Apoio à Revisão Sistemática. 2011. Dissertação (Mestrado em Ciências da Computação e Matemática Computacional) - Universidade de São Paulo, Fundação de Amparo à Pesquisa do Estado de São Paulo. Orientador: Jose Carlos Maldonado.
2.  Diogo Nascimento Campanha. Teste de mutação nos paradigmas procedimental e oo: uma avaliação no contexto de estrutura de dados. 2010. Dissertação (Mestrado em Ciências da Computação e Matemática Computacional) - Universidade de São Paulo, Fundação de Amparo à Pesquisa do Estado de São Paulo. Orientador: Jose Carlos Maldonado.

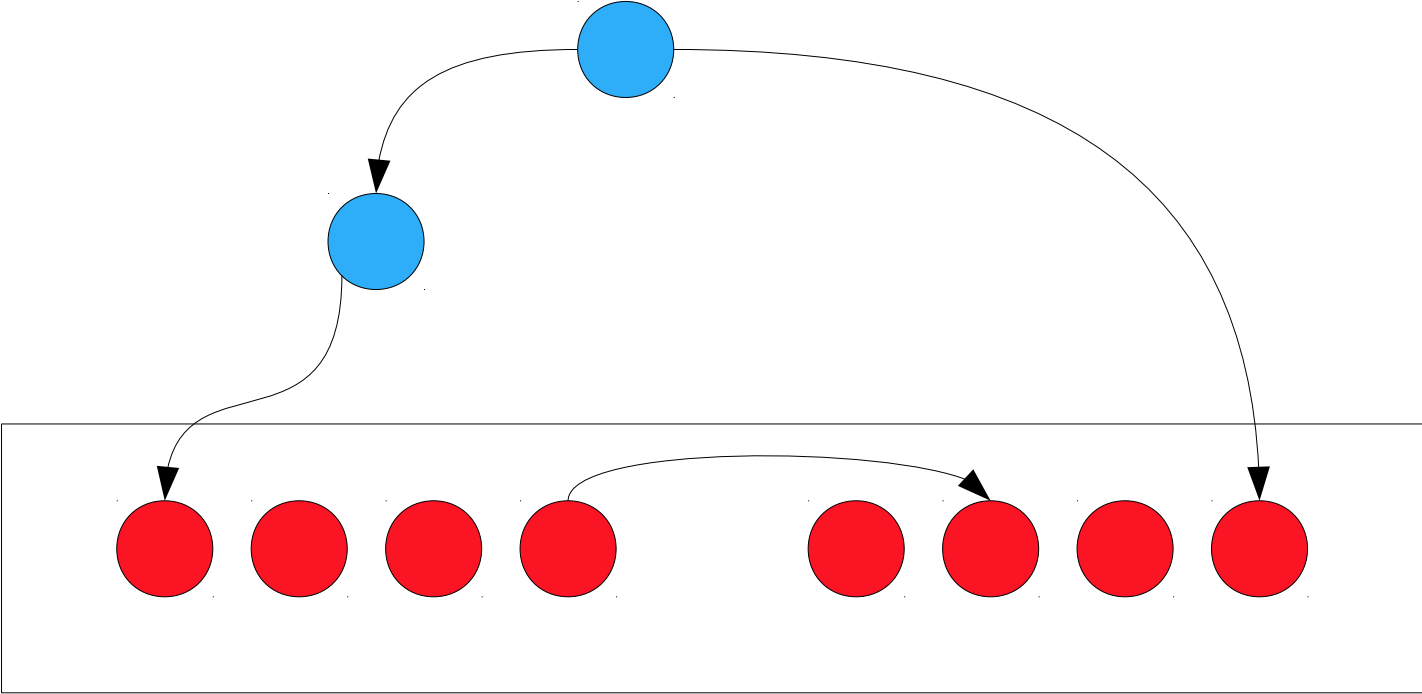
Tese de doutorado

1.  Marco Aurélio Graciotto Silva. LOD: uma abordagem para desenvolvimento de objetos de aprendizagem multimídias e interativos. 2012. Tese (Doutorado em Ciências da Computação e Matemática Computacional) - Universidade de São Paulo, Conselho Nacional de Desenvolvimento Científico e Tecnológico. Orientador: Jose Carlos Maldonado.
2.  Katia Romero Felizardo Scannavino. Evidence-based Software Engineering: Systematic Literature Review process based on Visual Text Mining. 2012. Tese (Doutorado em Ciências da Computação e Matemática Computacional) - Universidade de São Paulo, Conselho Nacional de Desenvolvimento Científico e Tecnológico. Orientador: Jose Carlos Maldonado.
- ...
17. Plínio R S Vilela. Critérios Potenciais Usos de Integração: Definição e Análise. 1998. 0 f. Tese (Doutorado em Engenharia Elétrica) - Universidade Estadual de Campinas, . Coorientador: Jose Carlos Maldonado.
18. Alberto Nobiato Crespo. Modelos de Confiabilidade de Software Baseados em Cobertura de Critérios Estruturais de Teste. 1997. Tese (Doutorado em Engenharia Elétrica) - Universidade Estadual de Campinas, . Coorientador: Jose Carlos Maldonado.
19.  Marcio Eduardo Delamaro. Mutação de Interface: Um Critério de Adequação Interprocedimental para o Teste de Integração. 1997. 0 f. Tese (Doutorado em Física Computacional) - Universidade de São Paulo, . Orientador: Jose Carlos Maldonado.
20.  Sílvia Regina Vergílio. Critérios Restritos: Uma Contribuição para aprimorar a Eficácia da Atividade de Teste de Software. 1997. 0 f. Tese (Doutorado em Engenharia Elétrica) - Universidade Estadual de Campinas, . Orientador: Jose Carlos Maldonado.
21. Sandra C P F Fabbri. A Análise de Mutantes no Contexto de Sistemas Reativos: Uma Contribuição para o Estabelecimento de Estratégias de Teste e Validação. 1996. 0 f. Tese (Doutorado em Física Computacional) - Universidade de São Paulo, . Orientador: Jose Carlos Maldonado.

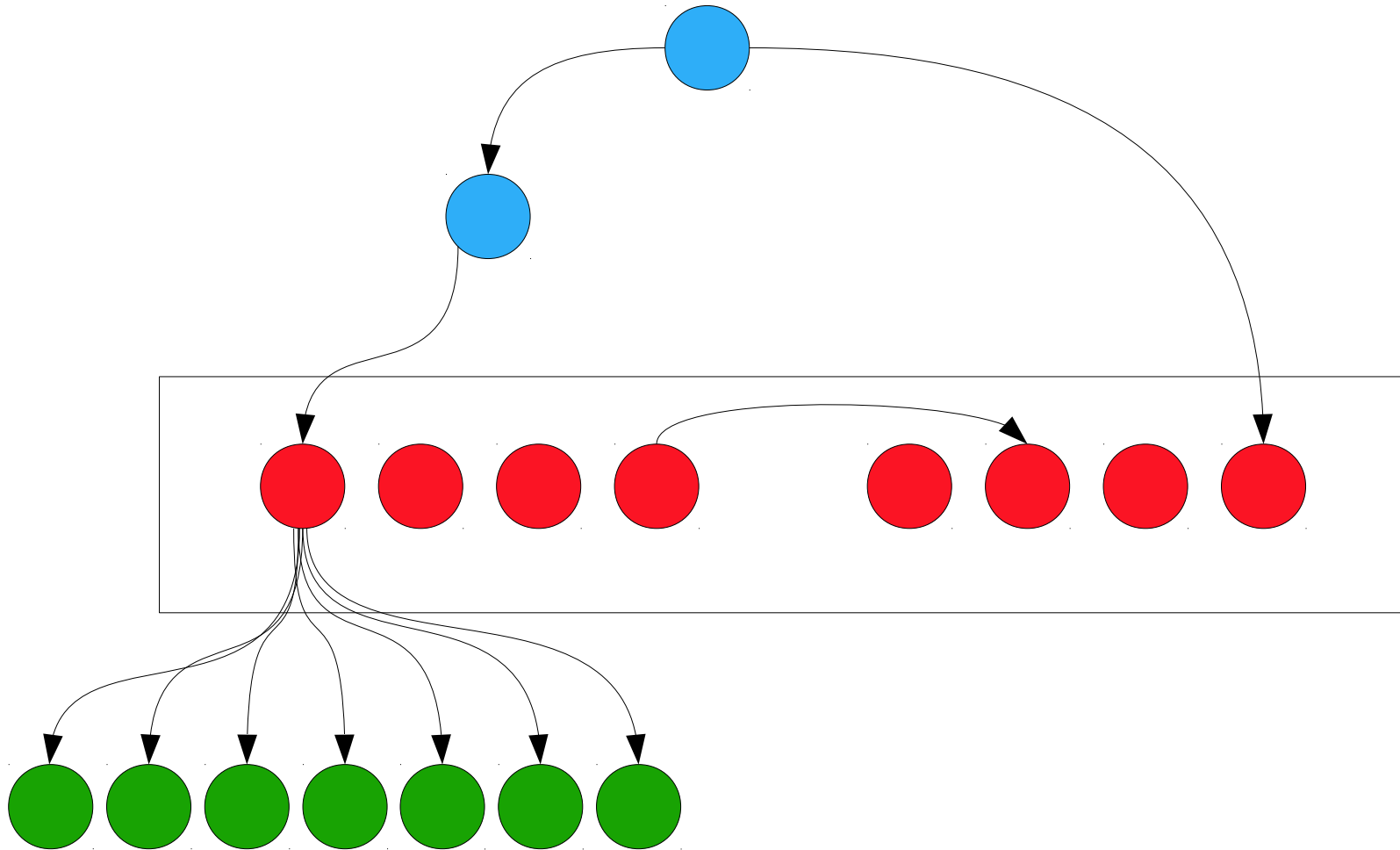
Método



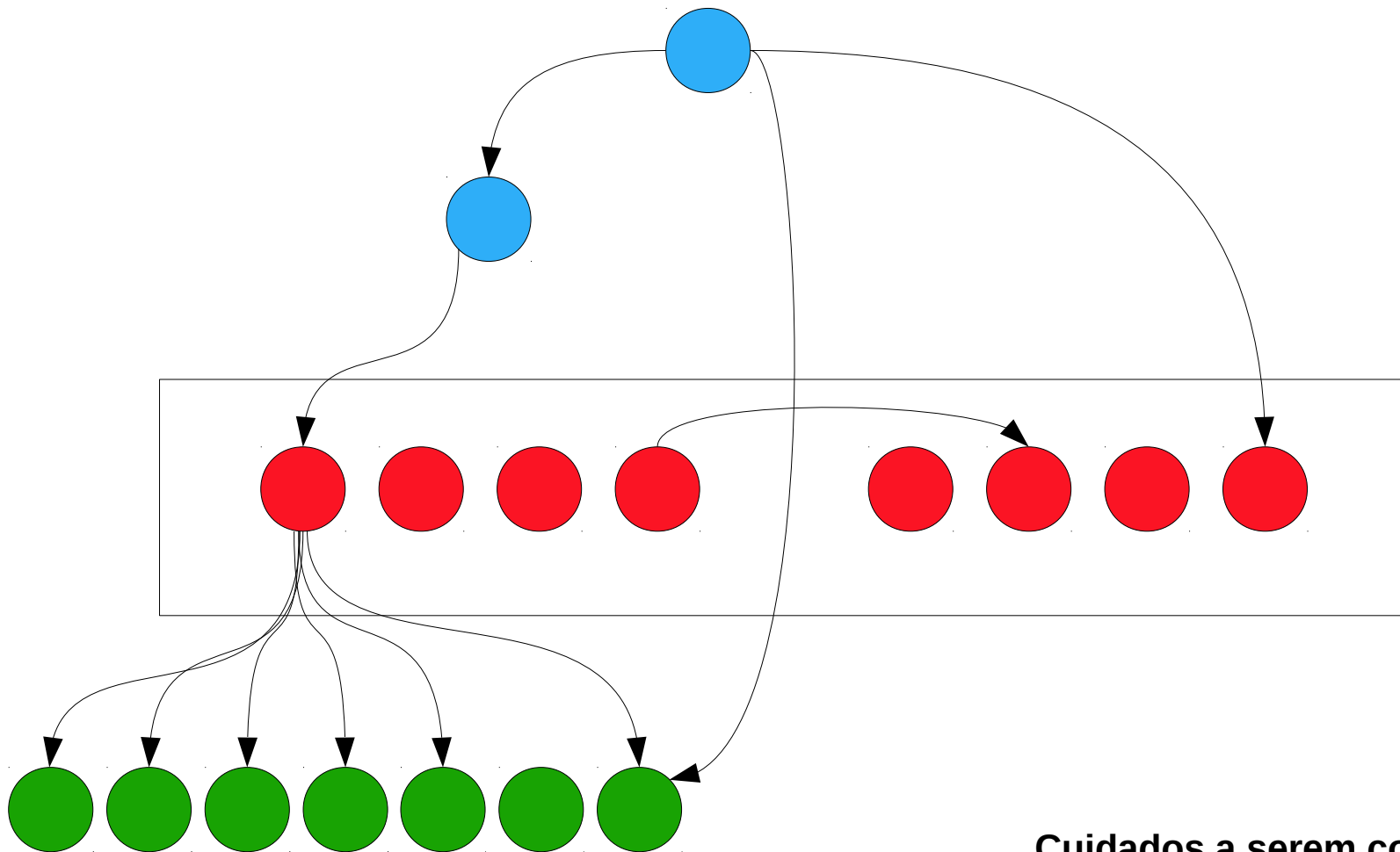
Método



Método



Método

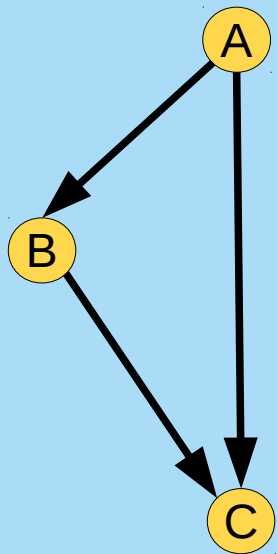


Cuidados a serem considerados:

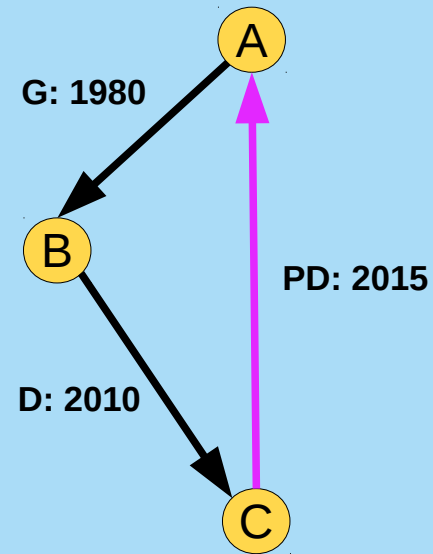
- Nomes abreviados:
Carlos Silva Costa
Carlos S. Costa
- Relações sem IDs
- IDs errados

Representação

A rigor a estrutura utilizada para genealogia acadêmica **não pode ser chamada de árvore (nem floresta)** pois existe a possibilidade de termos um ciclo.



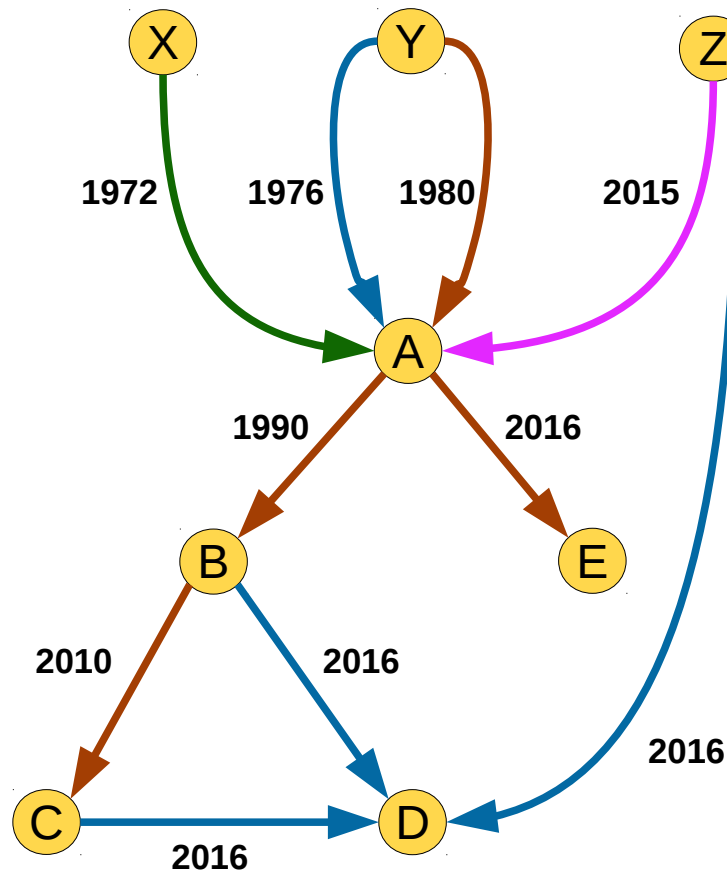
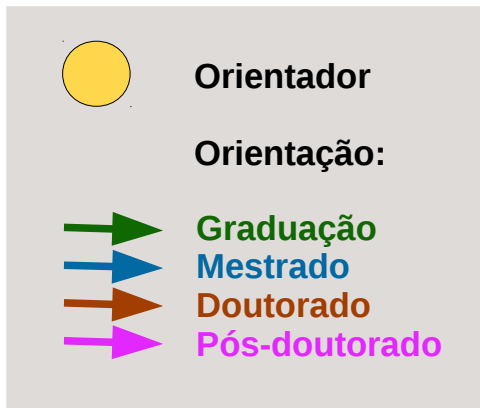
Árvore acíclico



Grafo com ciclo

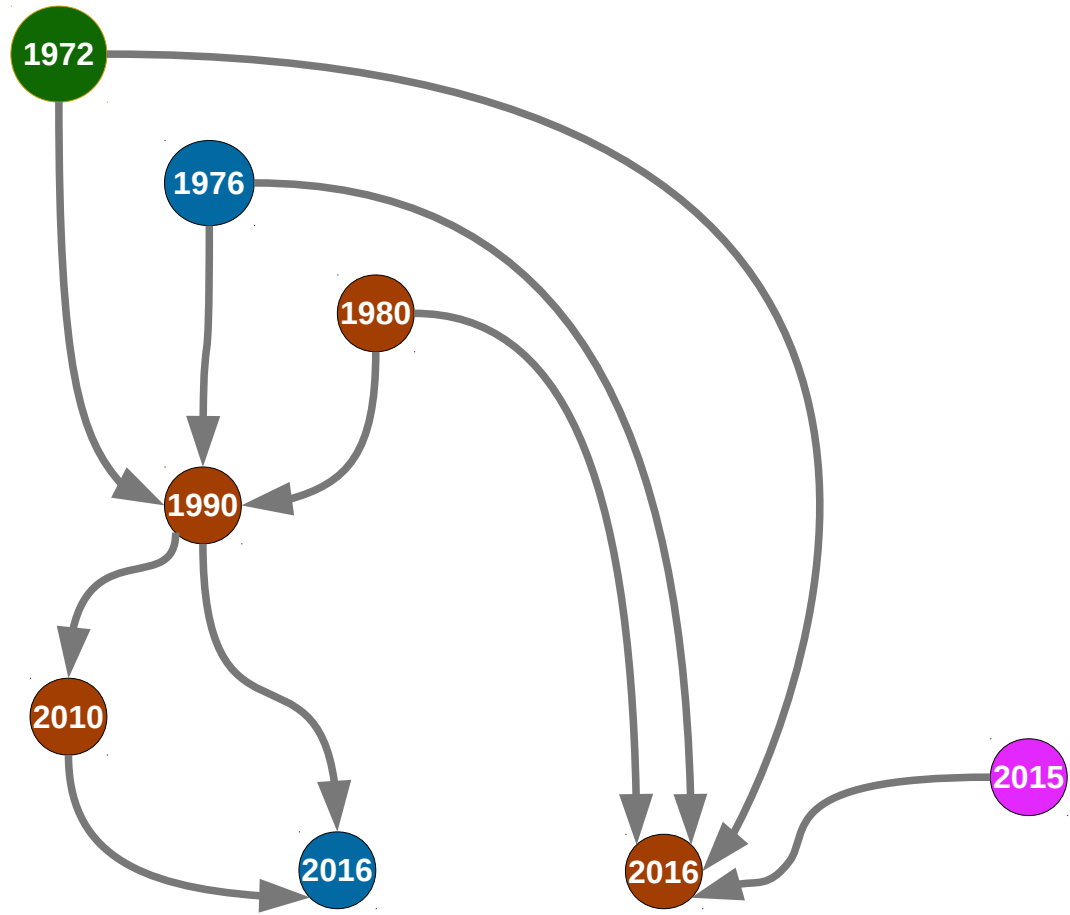
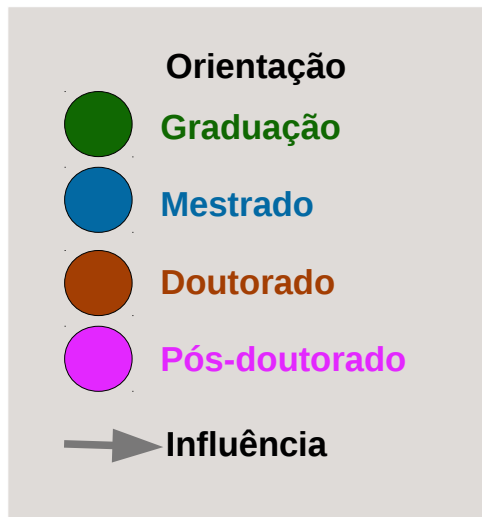
Representação

Grafo de genealogia = **Multigrafo direcionado não-conexo.**



Representação

Grafo de genealogia = **Multigrafo direcionado não-conexo.**





Genealogia Acadêmica dos PPGs em Ciência da Computação no Brasil

USP-SC (relações doutorado)

Grupo inicial analisado: 66 pessoas

Antecessores identificados: 88 pessoas (55 sem CV Lattes)

Sucessores identificados: 369 pessoas (48 sem CV Lattes)

Clique no nome da coluna para ordenar os elementos.

#	Nome completo	Menor ano de formação	País	Filhos ▲	Supervisões Pós-doutorado	Orientações doutorado	Coorientações doutorado	Orientações mestrado	Coorientações mestrado	Total de Sucessores
7	André Carlos Ponce de Leon Ferreira de Carvalho	1994	1	28	0	21	7	0	0	44
34	Jose Carlos Maldonado	1991	1	22	0	17	5	0	0	34
9	Caetano Traina Junior	1986	1	19	0	19	1	0	0	33
3	Aγμα Juci Machado Traina	1991	1	16	0	13	3	0	0	18
43	Maria Carolina Monard	1980	1	16	0	15	1	0	0	21
51	Paulo Cesar Masiero	1984	1	13	0	12	2	0	0	19
59	Roseli Aparecida Francelin Romero	1993	1	13	0	10	3	0	0	13
4	Alexandre Cláudio Botazzo Delbem	2002	2	12	0	9	3	0	0	12
14	Edson dos Santos Moreira	1989	1	12	0	12	0	0	0	18
44	Maria Cristina Ferreira de Oliveira	1990	1	11	0	9	2	0	0	16
53	Regina Helena Carlucci Santana	1989	1	11	0	11	0	0	0	12
42	Marcos José Santana	1989	1	10	0	10	0	0	0	26
64	Solange Oliveira Rezende	1993	1	10	0	10	0	0	0	10
54	Renata Pontin de Mattos Fortes	1996	1	9	0	9	1	0	0	9
8	Antonio Castelo Filho	1992	1	8	0	6	2	0	0	8
58	Rosane Minghim	1995	1	8	0	6	2	0	0	9
33	José Alberto Cuminato	1987	1	8	0	8	0	0	0	10
45	Maria da Graça Campos Pimentel	1994	1	8	0	10	0	0	0	8
38	Leandro Franco de Souza	2003	1	7	0	5	2	0	0	7
61	Sandra Maria Aluisio	1995	1	7	0	5	2	0	0	8
50	Odemir Martinez Bruno	2000	1	7	0	7	0	0	0	7

USP-SC (relações de PD, D, M)

Grupo inicial analisado: 66 pessoas

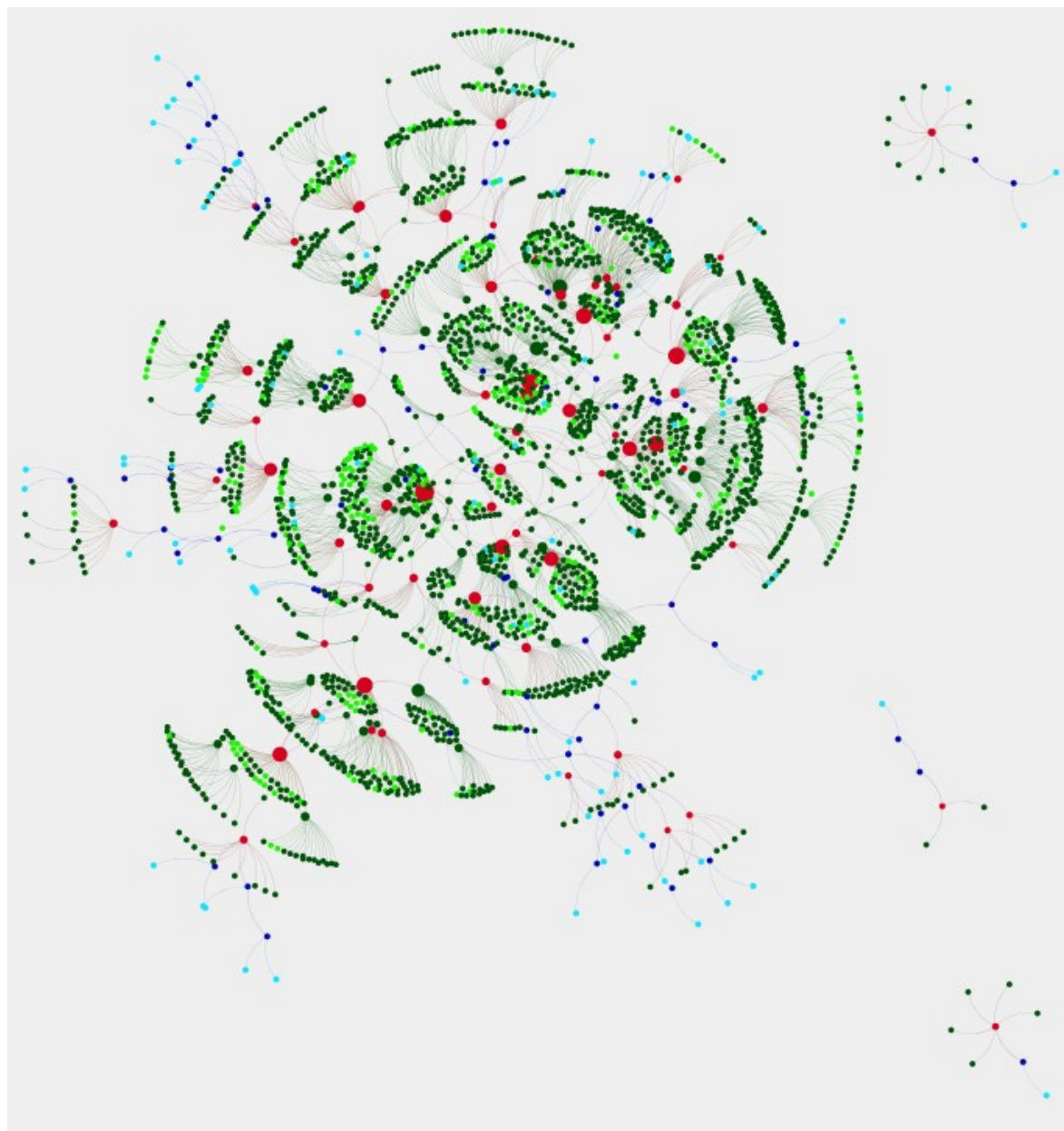
Antecessores identificados: 187 pessoas (111 sem CV Lattes)

Sucessores identificados: 2225 pessoas (360 sem CV Lattes)

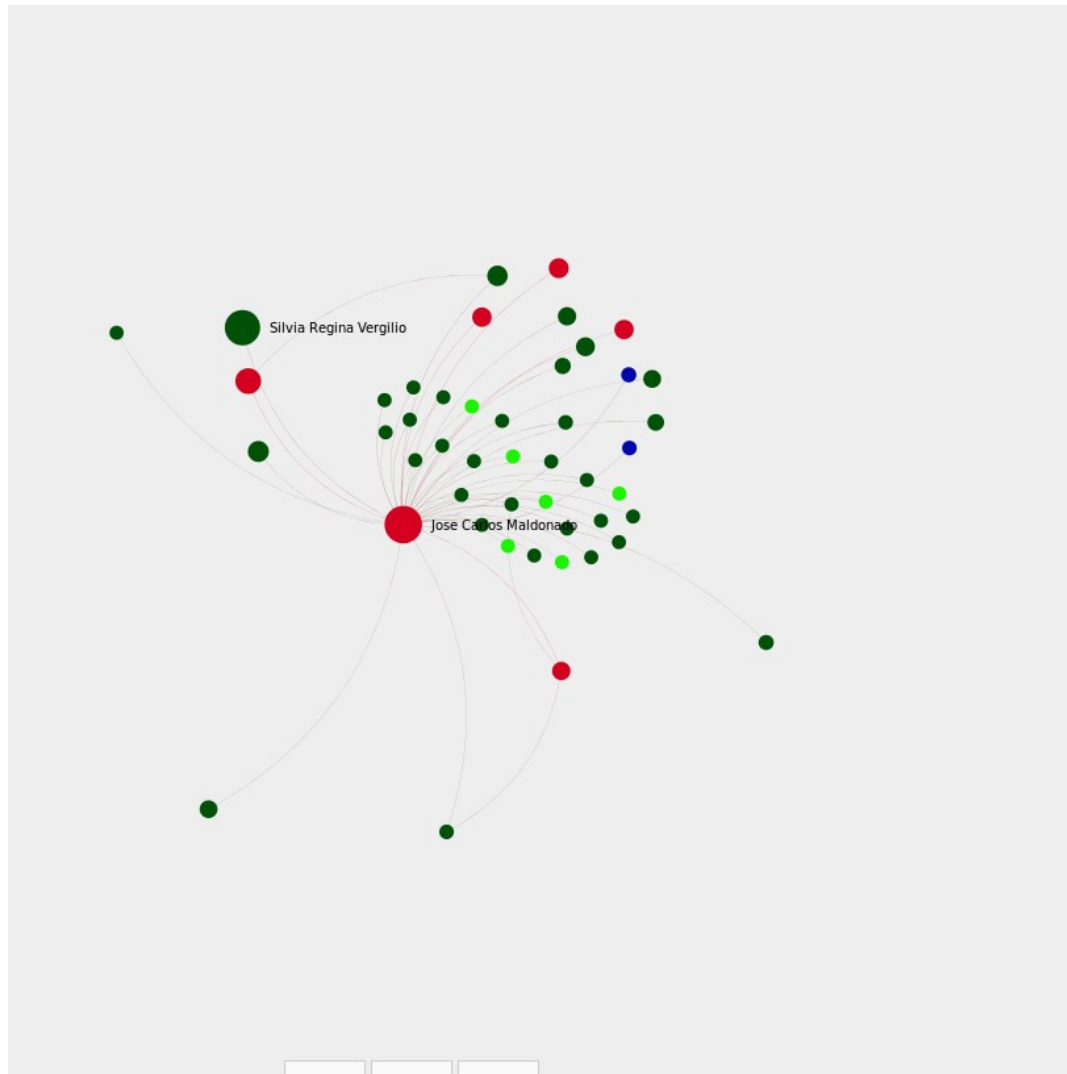
Clique no nome da coluna para ordenar os elementos.

#	Nome completo	Menor ano de formação	País	Filhos ▲	Supervisões Pós-doutorado	Orientações doutorado	Coorientações doutorado	Orientações mestrado	Coorientações mestrado	Total de Sucessores
7	André Carlos Ponce de Leon Ferreira de Carvalho	1990	2	58	19	21	7	23	2	204
51	Paulo Cesar Masiero	1979	2	50	1	12	2	41	1	283
14	Edson dos Santos Moreira	1984	2	49	0	12	0	41	0	104
42	Marcos José Santana	1985	2	46	0	10	0	41	0	180
34	Jose Carlos Maldonado	1983	2	44	0	17	5	29	3	216
9	Caetano Traina Junior	1982	2	44	4	19	1	35	0	217
53	Regina Helena Carlucci Santana	1985	2	41	0	11	0	34	0	93
43	Maria Carolina Monard	1969	2	40	2	15	1	35	0	301
3	Agma Juci Machado Traina	1987	1	39	6	13	3	26	0	82
45	Maria da Graca Campos Pimentel	1989	2	39	2	10	0	34	0	67
59	Roseli Aparecida Francelin Romero	1986	2	35	2	10	3	27	0	77
54	Renata Pontin de Mattos Fortes	1991	2	35	1	9	1	28	0	68
16	Eduardo Marques	1988	1	32	2	6	0	33	0	52
8	Antonio Castelo Filho	1987	1	30	2	6	2	25	0	66
64	Solange Oliveira Rezende	1990	2	29	3	10	0	25	0	33
44	Maria Cristina Ferreira de Oliveira	1990	1	28	0	9	2	18	0	83
58	Rosane Minghim	1990	2	27	3	6	2	19	0	48
61	Sandra Maria Aluísio	1989	2	27	7	5	2	19	1	44
4	Alexandre Cláudio Botazzo Delbem	1998	2	24	6	9	3	12	0	37
39	Luis Gustavo Nonato	1994	1	23	4	6	1	14	0	47
41	Márcio Eduardo Delamaro	1993	2	22	0	2	2	19	0	33

USP-SC (relações de PD, D, M)



USP-SC (relações de PD, D, M)



- Grupo Inicial
- Sucessor
- Sucessor Sem CV Lattes
- Antecessor
- Antecessor Sem CV Lattes

Information Pane

Jose Carlos Maldonado

menoranoformacao: 1983.0

area: Ciência da Computação

cvdata: 05/08/2016

rotulo: Grupo Inicial

instituicao: Universidade de São Paulo; Instituto de Ciências Matemáticas e de Computação; Departamento de Ciências da Computação



foto:

cvlattes: <http://lattes.cnpq.br/8807333466702951>

grandearea: Ciências Exatas e da Terra

nivelgeracao: 0.0

bolsa: Nível 1A - CA CC - Ciências da Computação

Connections:

[Adenildo da Silva Simão](#)

[Alberto Nobiato Crespo](#)

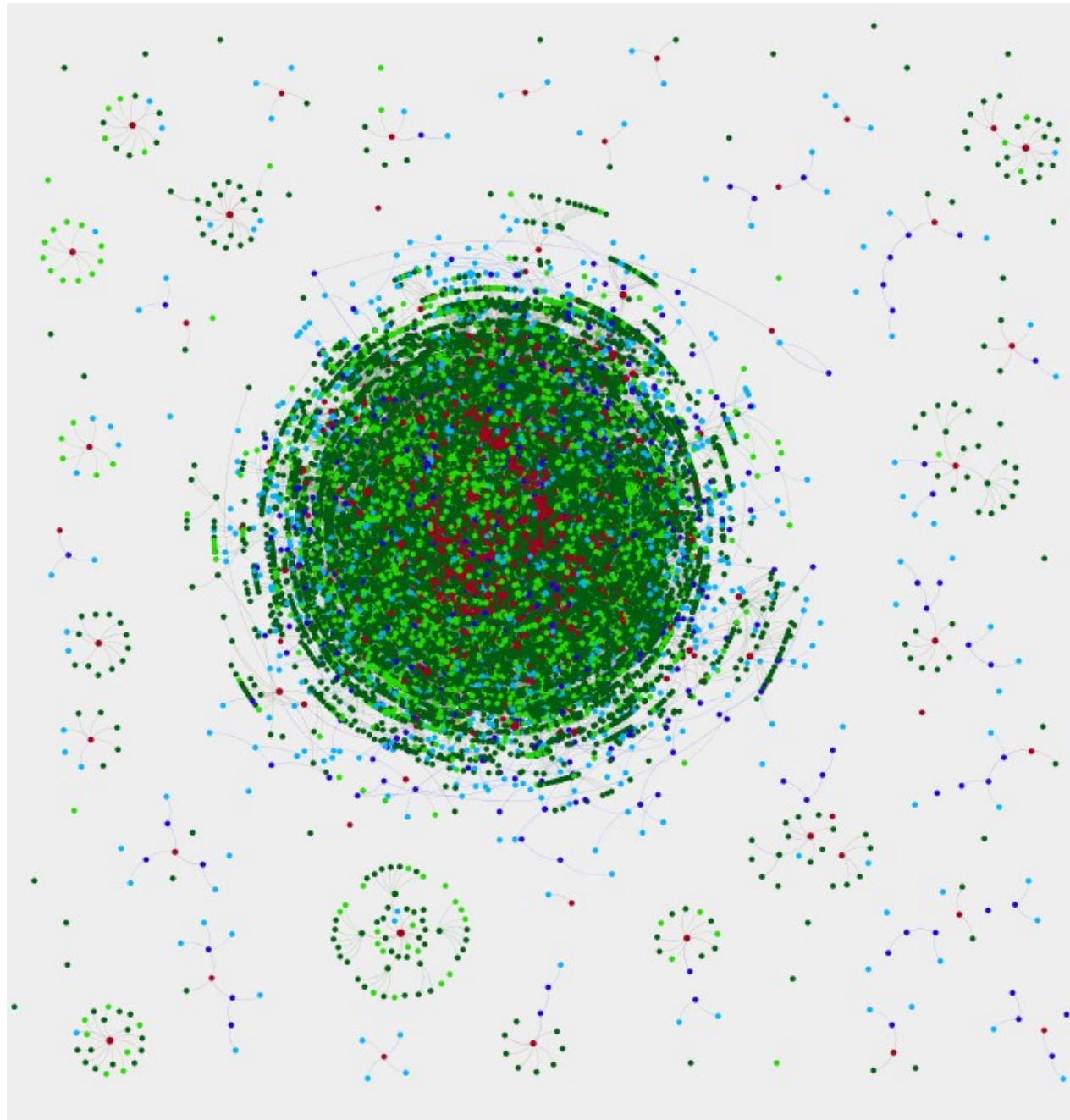
[André Luiz de Castro Villas Boas](#)

[André Luis dos Santos Domingues](#)

[Auri Marcelo Rizzo Vincenzi](#)

[Camila Kozlowski Della Corte](#)

Os 66 PPGs (PD, D, M)



Os 66 PPGs (relações de PD, D, M)

Grupo inicial analisado: 1402 pessoas

Antecessores identificados: 2232 pessoas (1526 sem CV Lattes)

Sucessores identificados: 23266 pessoas (5308 sem CV Lattes)

Clique no nome da coluna para ordenar os elementos.

#	Nome completo	Menor ano de formação	País	Filhos ▲	Supervisões Pós-doutorado	Orientações doutorado	Coorientações doutorado	Orientações mestrado	Coorientações mestrado	Total de Sucessores
227	Carlos José Pereira de Lucena	1969	2	148	0	45	2	111	4	3389
1298	Silvio Romero de Lemos Meira	1981	3	139	0	24	1	121	0	1660
91	Ana Regina Cavalcanti da Rocha	1978	1	113	0	27	4	80	17	549
632	Jano Moreira de Souza	1978	2	109	0	26	0	87	3	1443
799	Liane Margarida Rockenbach Tarouco	1976	2	108	0	32	11	74	4	1346
170	Arndt von Staa	1969	2	94	0	11	0	82	4	1841
892	Marcelo Gattass	1977	2	90	0	24	2	67	2	527
1243	Rosa Maria Vicari	1985	2	89	4	36	7	44	4	247
216	Carlos Becker Westphall	1988	2	88	0	5	0	80	5	965
1101	Philippe Olivier Alexandre Navaux	1973	2	83	3	23	2	61	4	513
136	Angelo Perkusich	1987	2	80	0	20	0	73	1	396
1034	Nelson Maculan Filho	1967	2	79	0	36	6	41	2	1771
144	Antonio Alfredo Ferreira Loureiro	1987	2	78	0	16	4	60	9	202
534	Geraldo Robson Mateus	1980	1	76	0	14	1	59	8	166
708	José Palazzo Moreira de Oliveira	1976	2	75	1	18	2	58	2	236
918	Marco Antonio Casanova	1976	2	72	2	19	0	56	1	202
857	Luiz Fernando Gomes Soares	1979	2	72	0	12	1	68	2	1151
942	Marcos Roberto da Silva Borges	1981	2	71	4	6	3	41	21	1142
475	Felipe Maia Galvão França	1987	2	69	3	17	3	48	6	109
270	Claudia Maria Bauzer Medeiros	1979	2	68	2	12	5	52	0	396
1356	Valmir Carneiro Barbosa	1982	2	68	4	16	1	55	2	289

Os 66 PPGs (relações de doutorado)

Grupo inicial analisado: 1403 pessoas

Antecessores identificados: 1210 pessoas (869 sem CV Lattes)

Sucessores identificados: 2989 pessoas (479 sem CV Lattes)

Clique no nome da coluna para ordenar os elementos.

#	Nome completo	Menor ano de formação	País	Filhos ▲	Supervisões Pós-doutorado	Orientações doutorado	Coorientações doutorado	Orientações mestrado	Coorientações mestrado	Total de Sucessores
227	Carlos José Pereira de Lucena	1974	1	47	0	45	2	0	0	133
1244	Rosa Maria Vicari	1990	1	42	0	36	7	0	0	59
1034	Nelson Maculan Filho	1975	1	41	0	35	6	0	0	101
799	Liane Margarida Rockenbach Tarouco	1990	1	41	0	32	11	0	0	76
634	Jayme Luiz Szwarcfiter	1975	1	40	0	40	1	0	0	97
1087	Paulo Roberto Oliveira	1977	1	33	0	33	0	0	0	75
91	Ana Regina Cavalcanti da Rocha	1983	1	30	0	27	4	0	0	61
103	André Carlos Ponce de Leon Ferreira de Carvalho	1994	1	28	0	21	7	0	0	44
892	Marcelo Gattass	1982	1	26	0	24	2	0	0	72
632	Jano Moreira de Souza	1986	1	26	0	26	0	0	0	77
407	Edward Hermann Haeusler	1990	1	26	0	24	3	0	0	29
265	Clarisse Sieckenius de Souza	1987	1	25	0	20	5	0	0	43
1299	Silvio Romero de Lemos Meira	1985	2	25	0	24	1	0	0	38
246	Celso da Cruz Carneiro Ribeiro	1983	2	24	0	22	2	0	0	35
1101	Philippe Olivier Alexandre Navaux	1979	1	24	0	22	2	0	0	26
322	Dante Augusto Couto Barone	1984	1	23	0	21	3	0	0	26
650	Joao Crisóstomo Weyl Albuquerque Costa	1994	1	22	0	18	4	0	0	22
191	Bruno Feijo	1988	1	22	0	18	4	0	0	32
866	Luiz Satoru Ochi	1989	1	22	0	16	6	0	0	26
691	Jose Carlos Maldonado	1991	1	22	0	18	5	0	0	37
475	Eolipo Maia Galvão Franco	1994	1	20	0	17	3	0	0	20

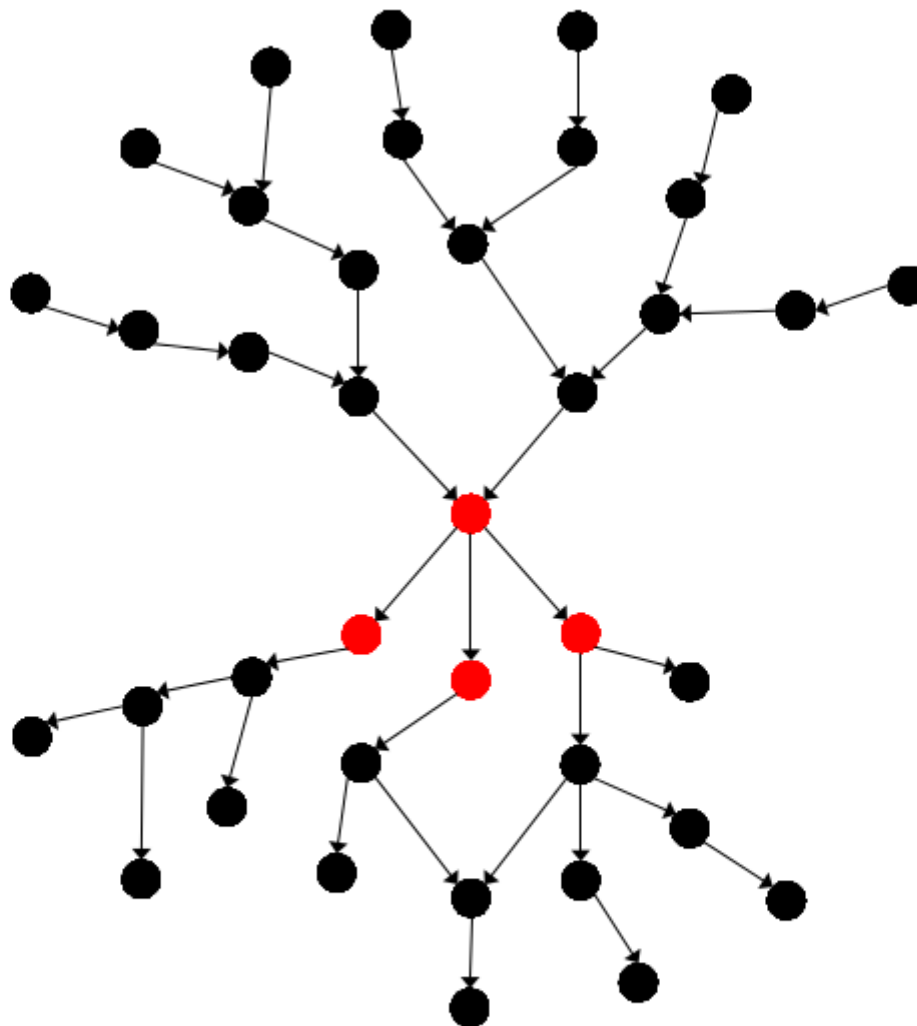


Algumas abordagens para explorar esses dados ...

Caracterização de grafos de genealogia

Métrica	Sigla	Derivada de métrica	Tipo de variável
01 – Largura	l^+	Não	Discreta
02 – Número de folhas	nf	Não	Discreta
03 – Profundidade	p^+	Não	Discreta
04 – Fecundidade	f^+	Não	Discreta
05 – Fecundidade ponderada	fp	Sim	Contínua
06 – Maior largura	ml	Não	Discreta
07 – Índice h genealógico	h	Sim	Discreta
08 – Impacto	i	Sim	Contínua
09 – Distância média	dm	Não	Contínua
10 – Média dos menores caminhos	mmc	Não	Contínua
11 – Pagerank inverso	pr^-	Sim	Contínua
12 – Pagerank inverso ponderado	prp^-	Sim	Contínua
13 – Balanceamento pela fecundidade	bf	Sim	Contínua
14 – Balanceamento global	bg	Sim	Contínua
15 – Fecundidade inversa	f^-	Não	Discreta
16 – Fec. média do território inverso	fmT^-	Sim	Contínua
17 – Fec. pond. média do terr. inv.	$fpmT^-$	Sim	Contínua
18 – Número de origens	no	Não	Discreta
19 – Largura inversa	l^-	Não	Discreta
20 – Profundidade inversa	p^-	Não	Discreta
21 – Pagerank	pr^+	Não	Contínua
22 – Pagerank ponderado	prp^+	Sim	Contínua

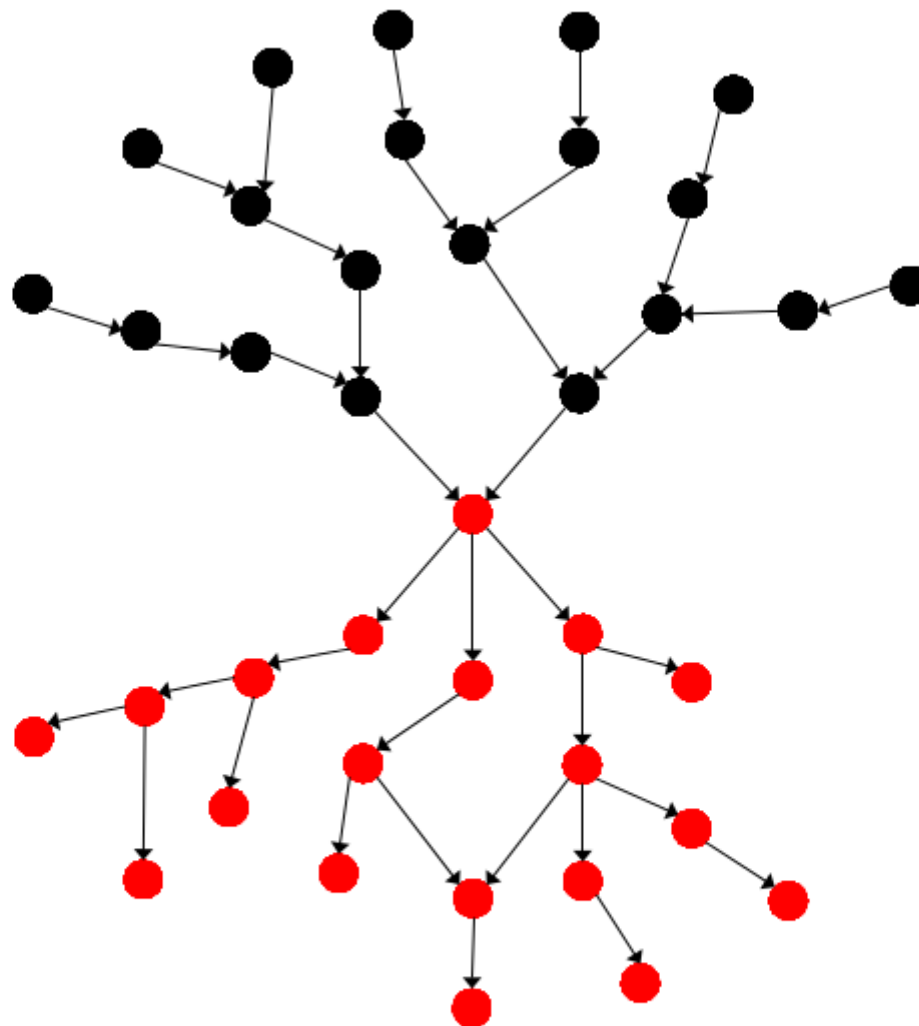
Caracterização de grafos de genealogia



$$D(v) = \{u \in V : (v, u) \in E\}$$

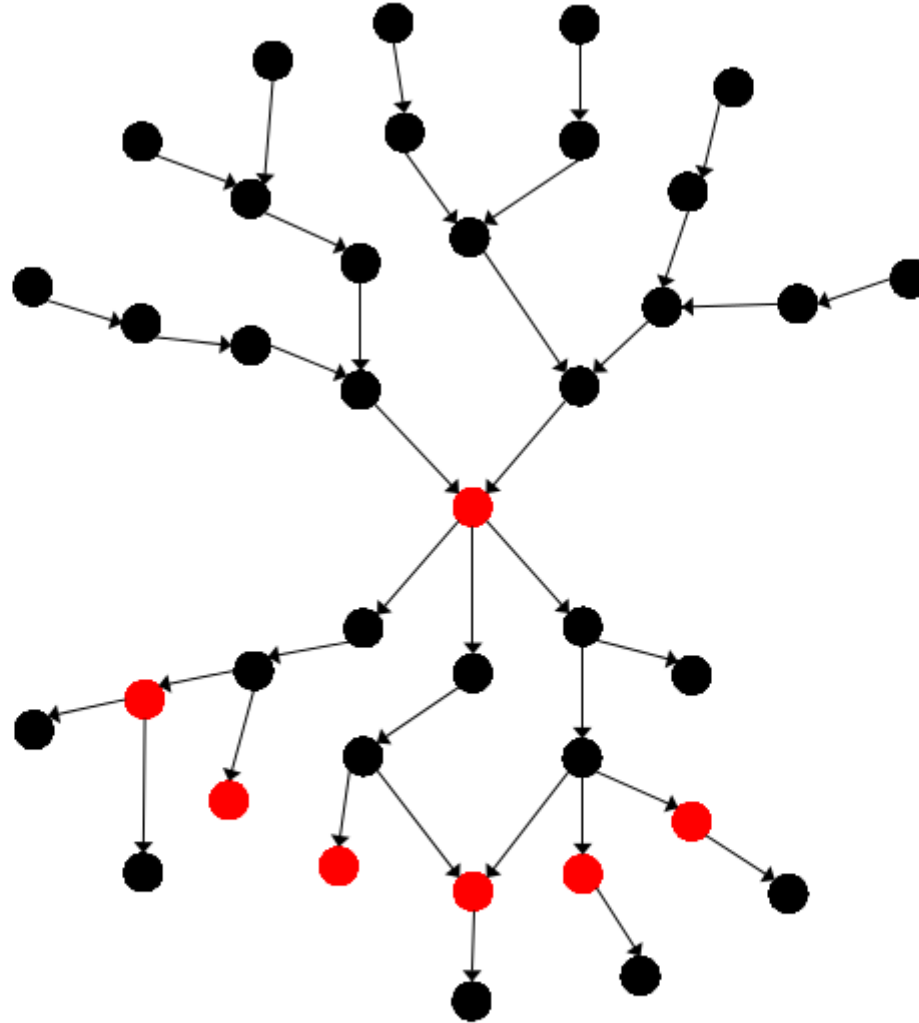
$$l^+(v) = |D(v)|$$

Caracterização de grafos de genealogia



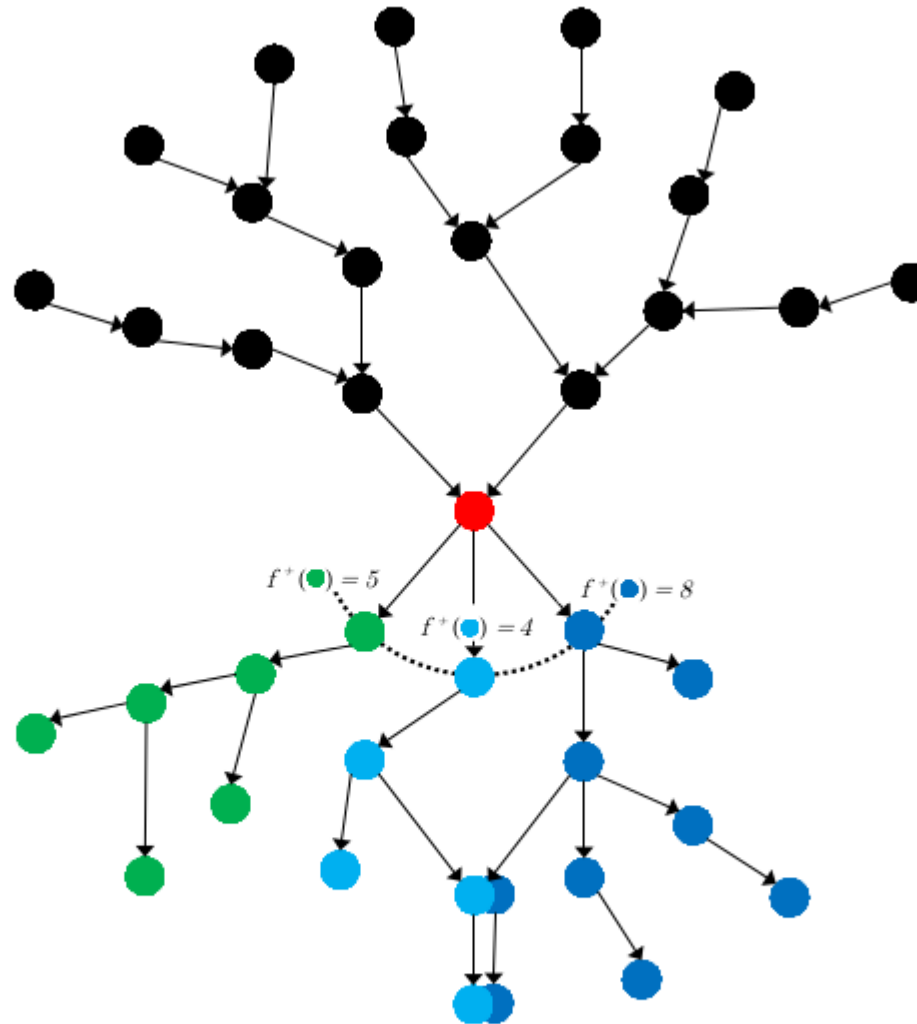
$$f^+(v) = |T^+(v)|$$

Caracterização de grafos de genealogia



$$ml(v) = \{k \in \mathbb{N} : \max(c^{(k)}(v))\}$$

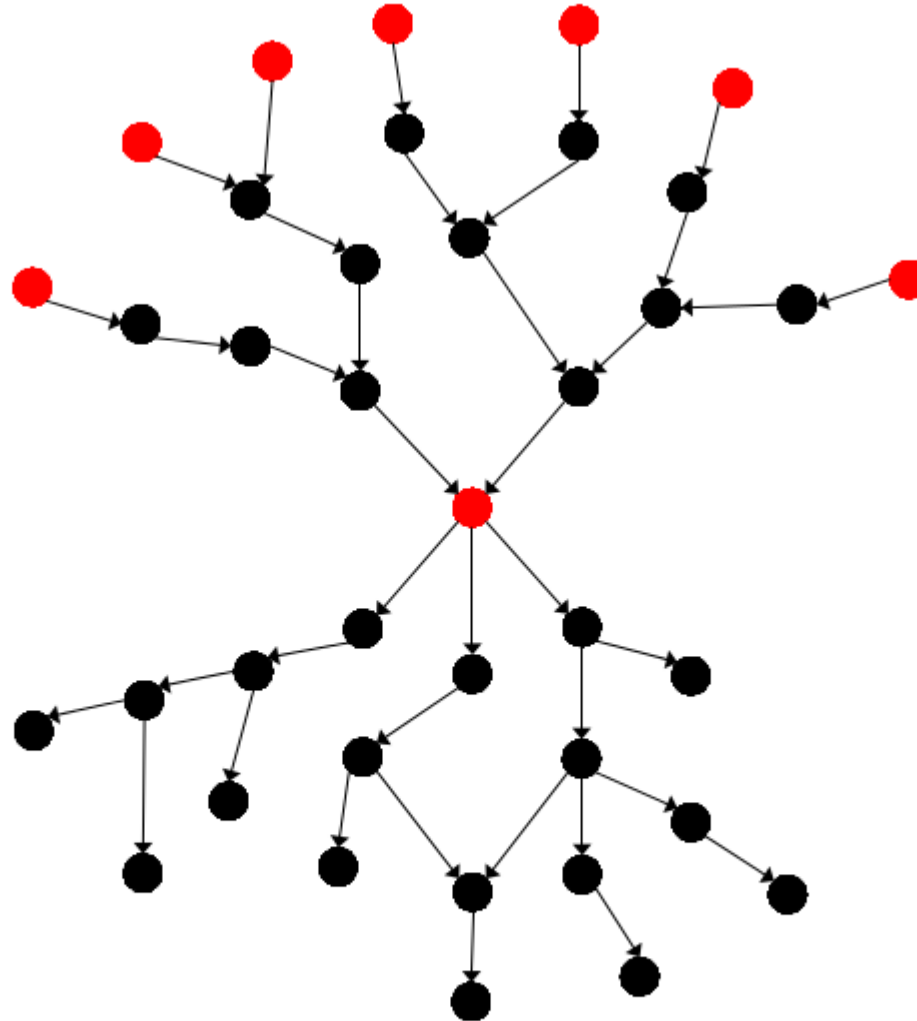
Caracterização de grafos de genealogia



$$F(v) = \{f^+(u) : u \in D(v)\}$$

$$bf(v) = std(F(v))$$

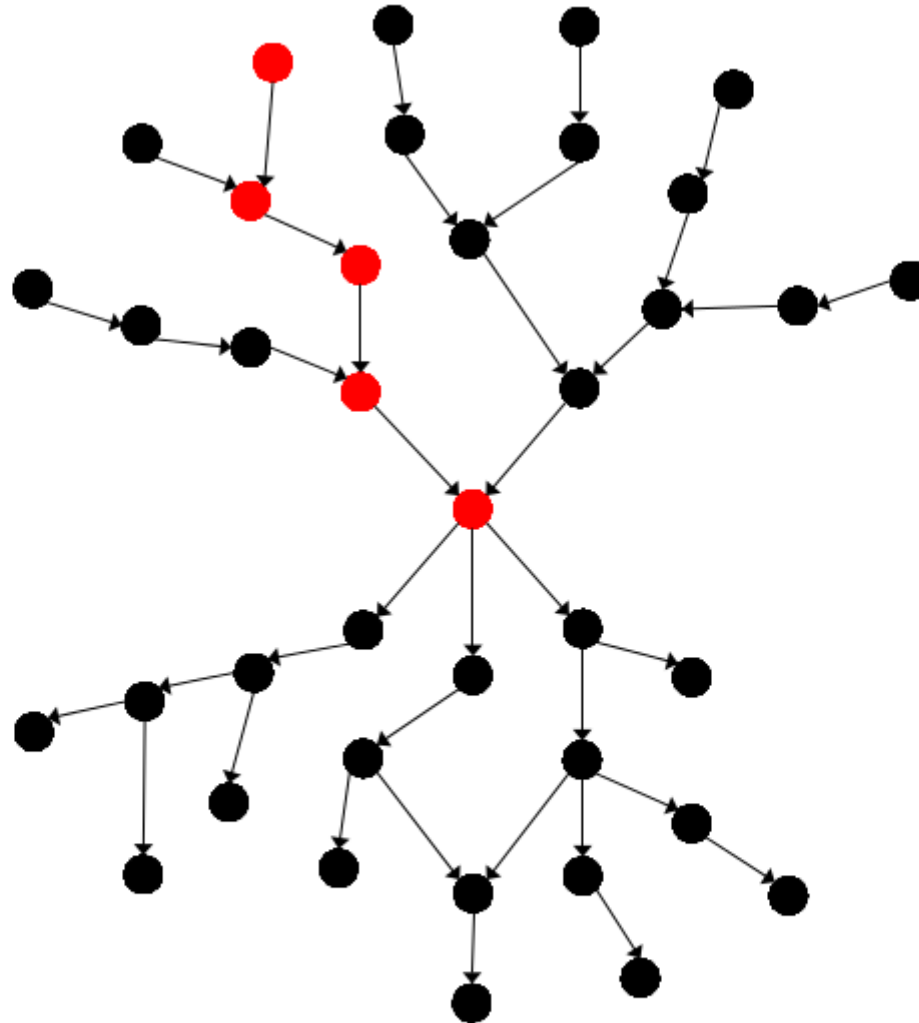
Caracterização de grafos de genealogia



$$O(v) = \{u \in T^-(v) : d^+(u) = 0\}$$

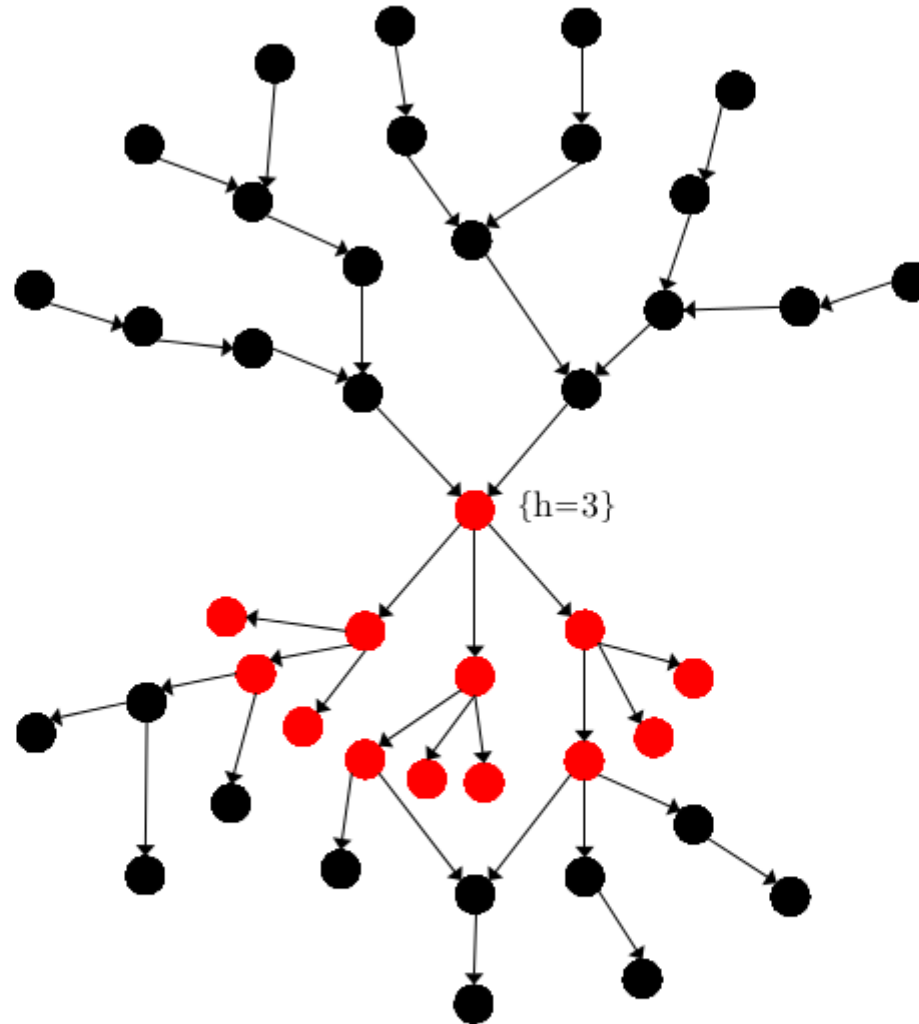
$$no(v) = |O(v)|$$

Caracterização de grafos de genealogia



$$p^-(v) = \max\{k \in \mathbb{N} : \exists(u, v) - \text{caminho de comprimento } k \text{ em } G, \forall u\}$$

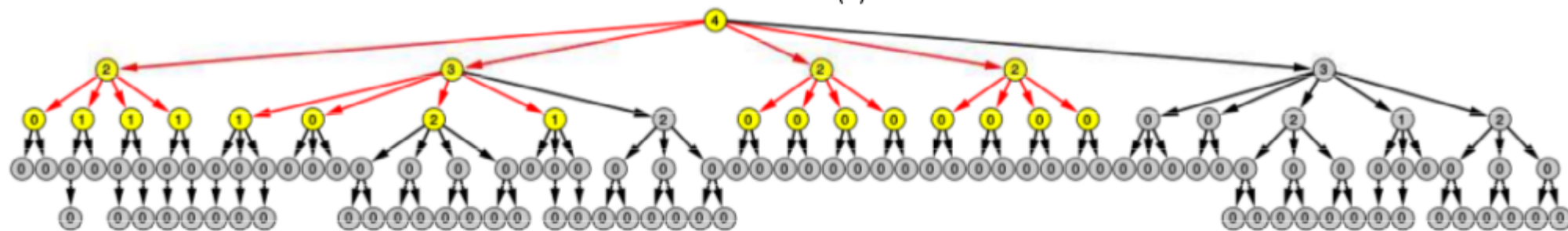
Caracterização de grafos de genealogia



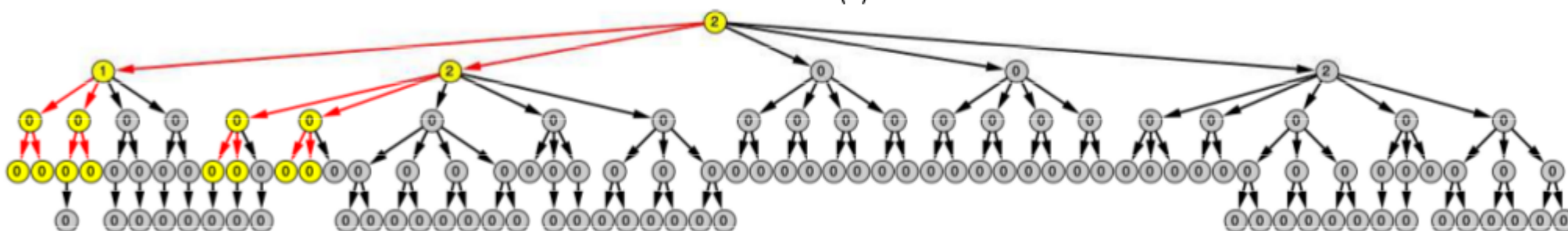
$$h(u) = \max\{k \in \mathbb{N} : l^+(u), l^{(k)}(u) \geq k\}$$

Caracterização de grafos de genealogia

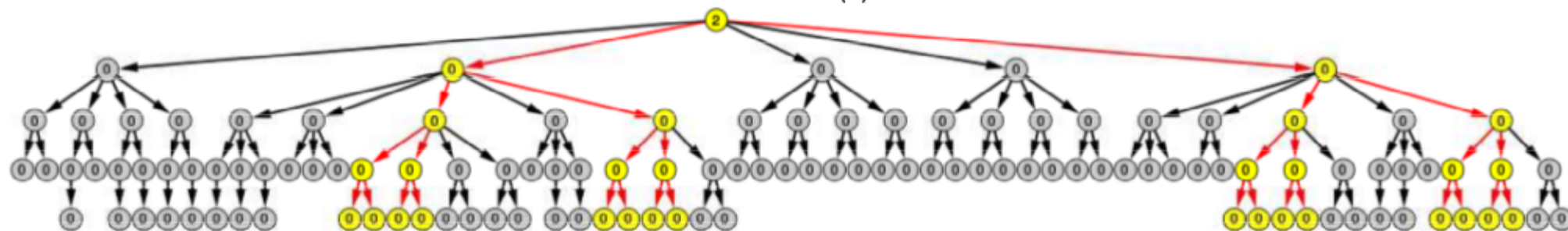
Índice $h_{(1)}$



Índice $h_{(2)}$

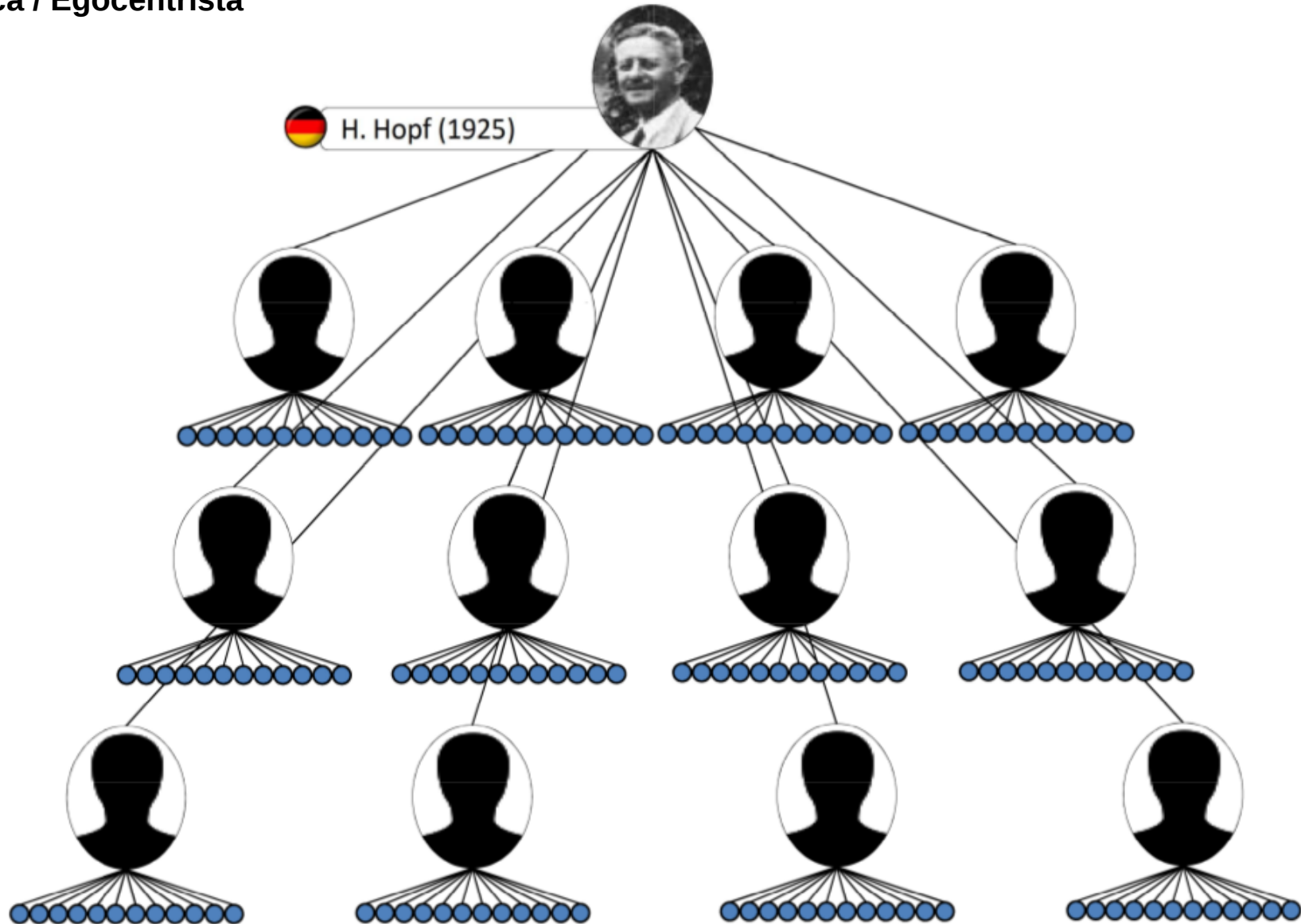


Índice $h_{(3)}$



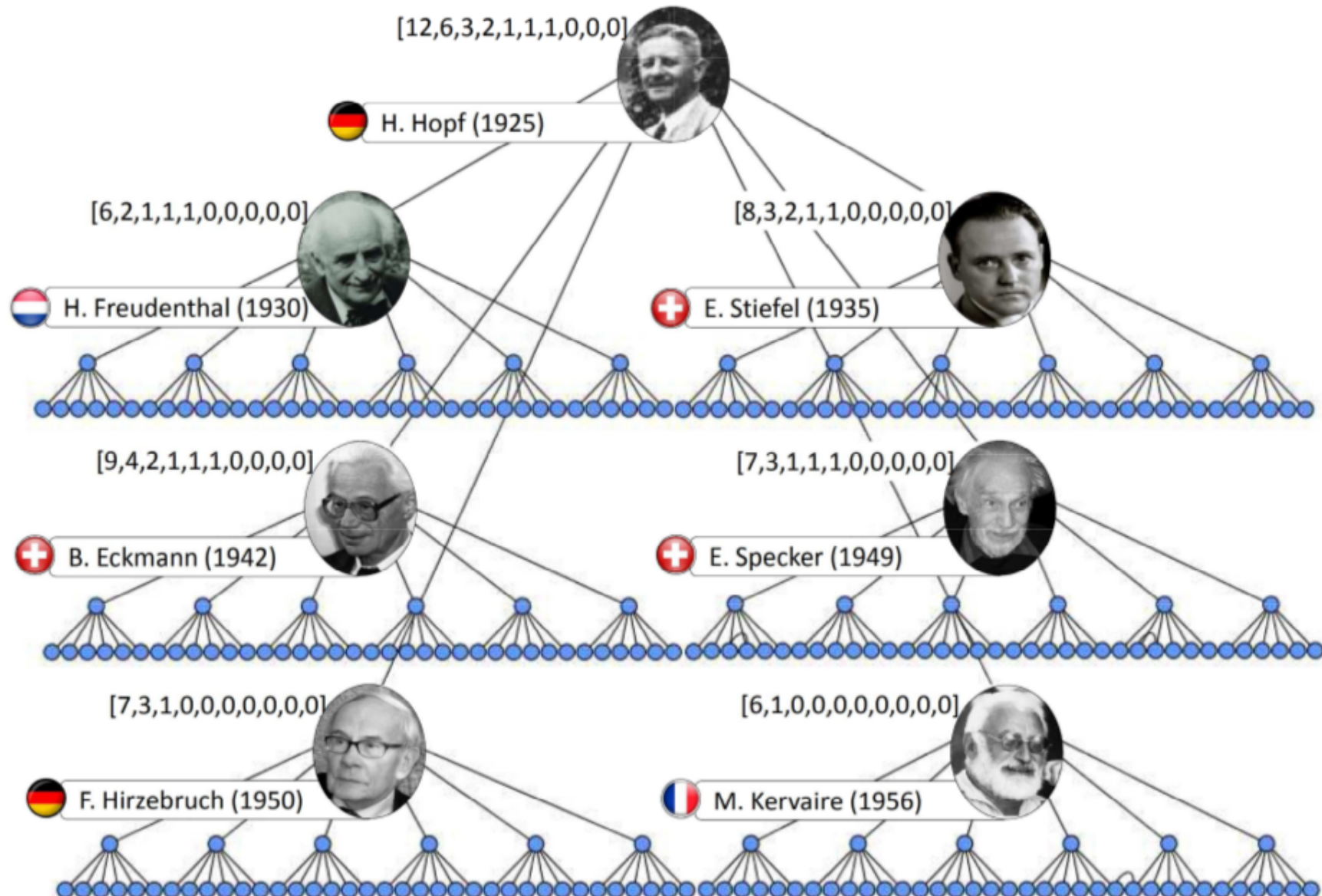
Caracterização de grafos de genealogia

G. Analítica / Egocentrista



Vértice origem com $h_{(1)} = 12$.

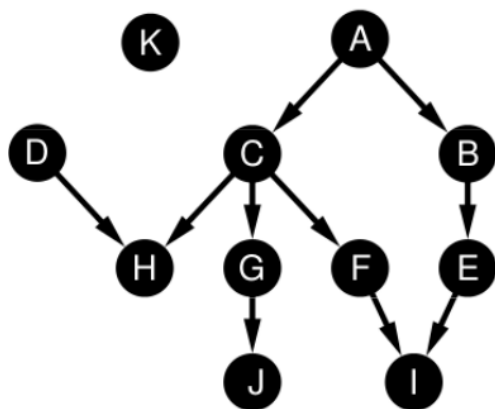
Caracterização de grafos de genealogia



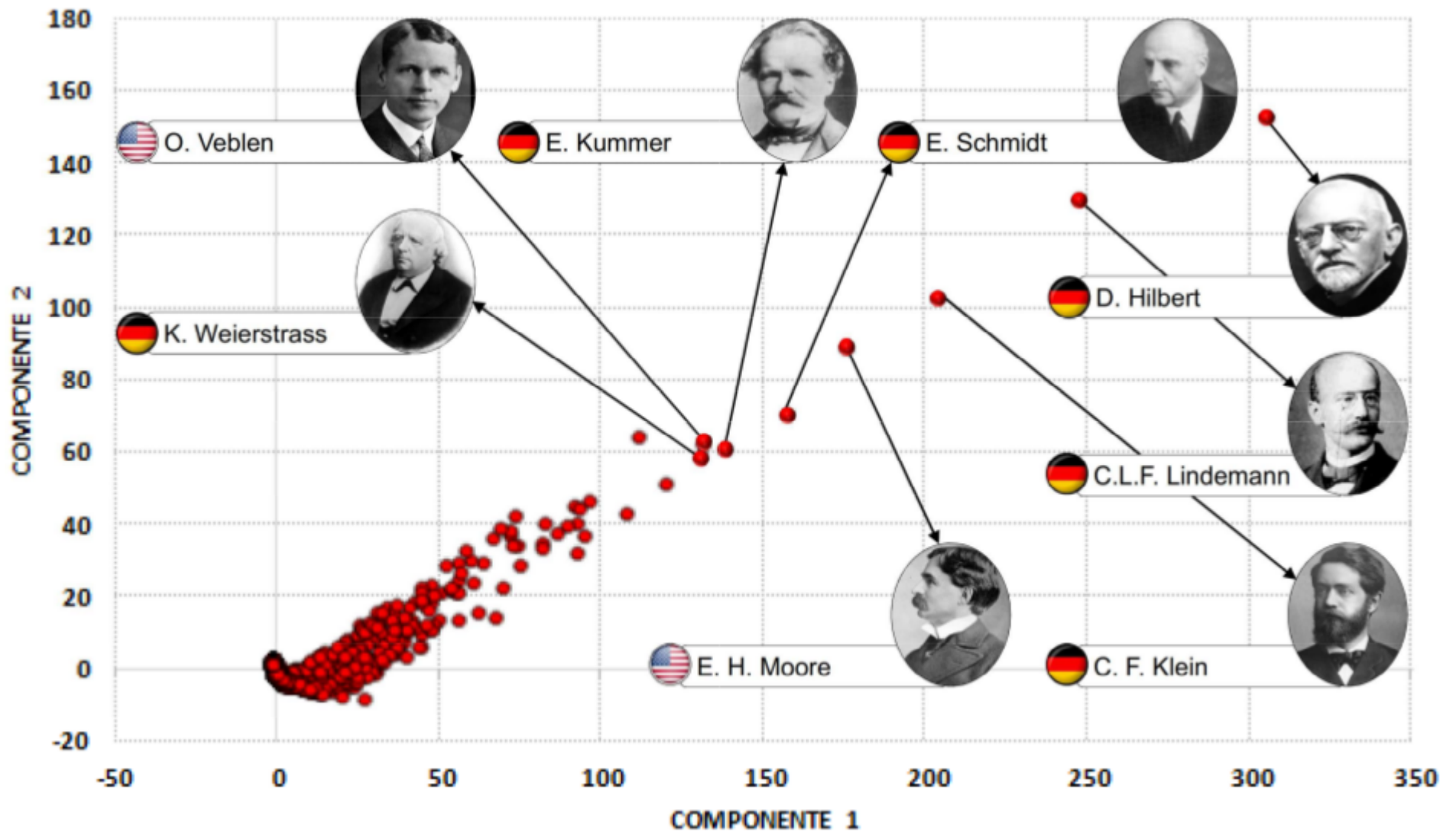
Vértice origem com $h_{(2)} = 6$.

Caracterização de grafos de genealogia

Vetor de características



MÉTRICA	VÉRTICES (v)										
	A	B	C	D	E	F	G	H	I	J	K
$l^+(v)$	2	1	3	1	1	1	1	0	0	0	0
$nf(v)$	3	1	3	1	1	1	1	0	0	0	0
$p^+(v)$	3	2	2	1	1	1	1	0	0	0	0
$f^+(v)$	8	2	5	1	1	1	1	0	0	0	0
$fp(v)$	3, 22	1, 25	3, 50	1, 00	1, 00	1, 00	1, 00	0, 00	0, 00	0, 00	0, 00
$ml(v)$	4	1	3	1	1	1	1	0	0	0	0
$h(v)$	1	1	1	0	0	0	0	0	0	0	0
$i(v)$	0, 80	0, 20	0, 50	0, 10	0, 10	0, 10	0, 10	0, 00	0, 00	0, 00	0, 00
$dm(v)$	2, 00	1, 50	1, 40	0, 00	0, 00	0, 00	0, 00	0, 00	0, 00	0, 00	0, 00
$mmc(v)$	0, 22	0, 50	0, 23	0, 50	0, 50	0, 50	0, 50	0, 00	0, 00	0, 00	0, 00
$rpr(v)$	0, 09	0, 03	0, 06	0, 02	0, 02	0, 02	0, 03	0, 01	0, 01	0, 01	0, 01
$rprp(v)$	11, 36	33, 16	17, 42	51, 46	51, 46	51, 46	39, 64	73, 33	73, 33	73, 33	73, 33
$bf(v)$	1, 50	0, 00	0, 47	0, 00	0, 00	0, 00	0, 00	-1, 00	-1, 00	-1, 00	-1, 00
$bg(v)$	0, 00	0, 00	0, 64	0, 00	0, 00	0, 00	0, 00	-1, 00	-1, 00	-1, 00	-1, 00
$f^-(v)$	0	1	1	0	2	2	2	3	5	3	0
$fmT^-(v)$	0, 00	8, 00	8, 00	0, 00	5, 00	6, 50	6, 50	4, 67	3, 40	4, 67	0, 00
$fpmT^-(v)$	0, 00	3, 22	3, 22	0, 00	2, 24	3, 36	3, 36	2, 57	1, 99	2, 57	0, 00
$no(v)$	0	1	1	0	1	1	1	2	1	1	0
$l^-(v)$	0	1	1	0	1	1	1	2	2	1	0
$p^-(v)$	0	1	1	0	2	2	2	2	3	3	0
$pr(v)$	0, 01	0, 02	0, 02	0, 01	0, 03	0, 02	0, 02	0, 03	0, 06	0, 03	0, 01
$prp(v)$	73, 33	51, 46	51, 46	73, 33	33, 16	52, 24	52, 24	32, 54	18, 01	33, 44	73, 33

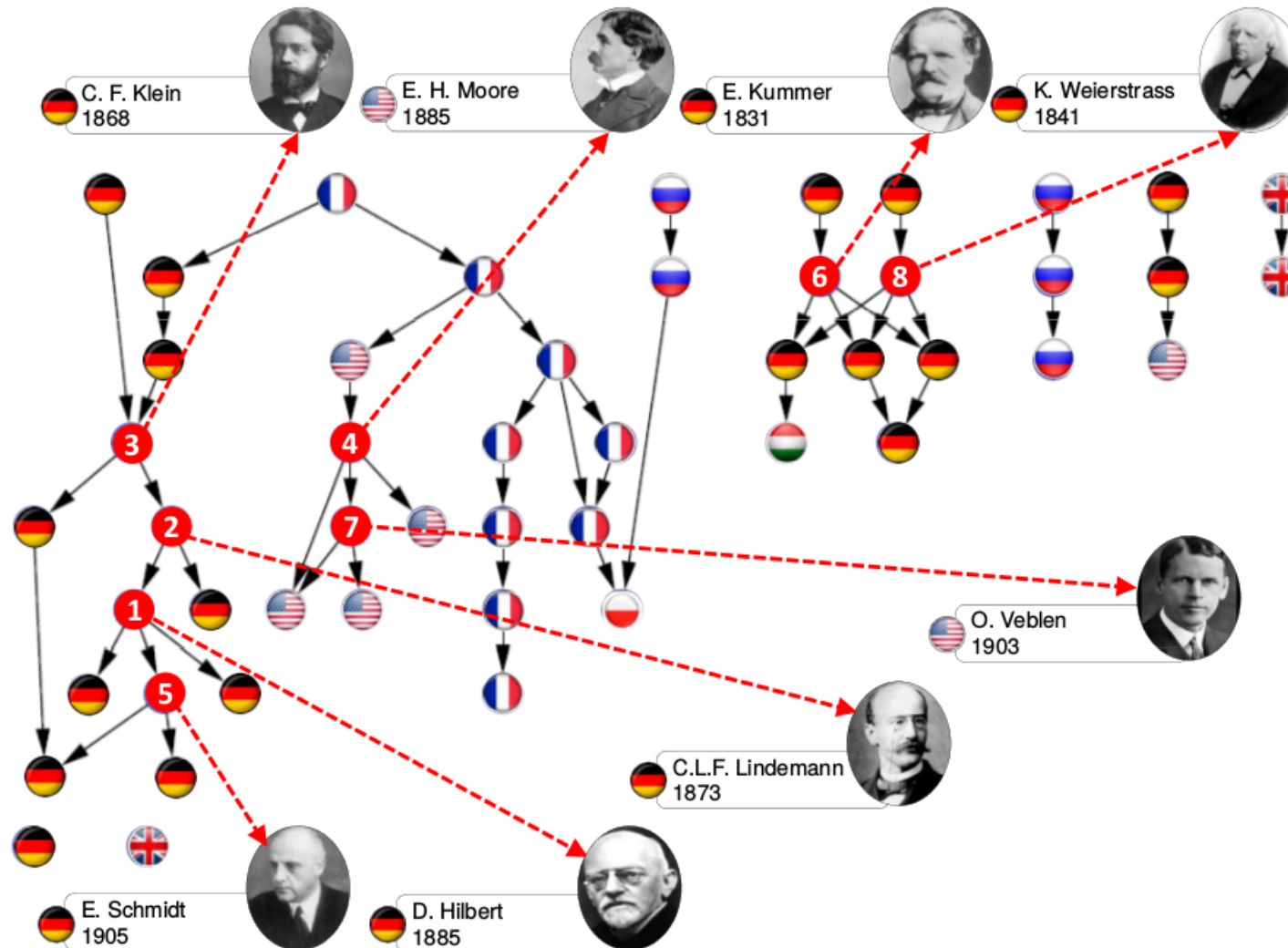


Pergunta:

Como identificar pesquisadores (orientadores) que tiveram grande impacto

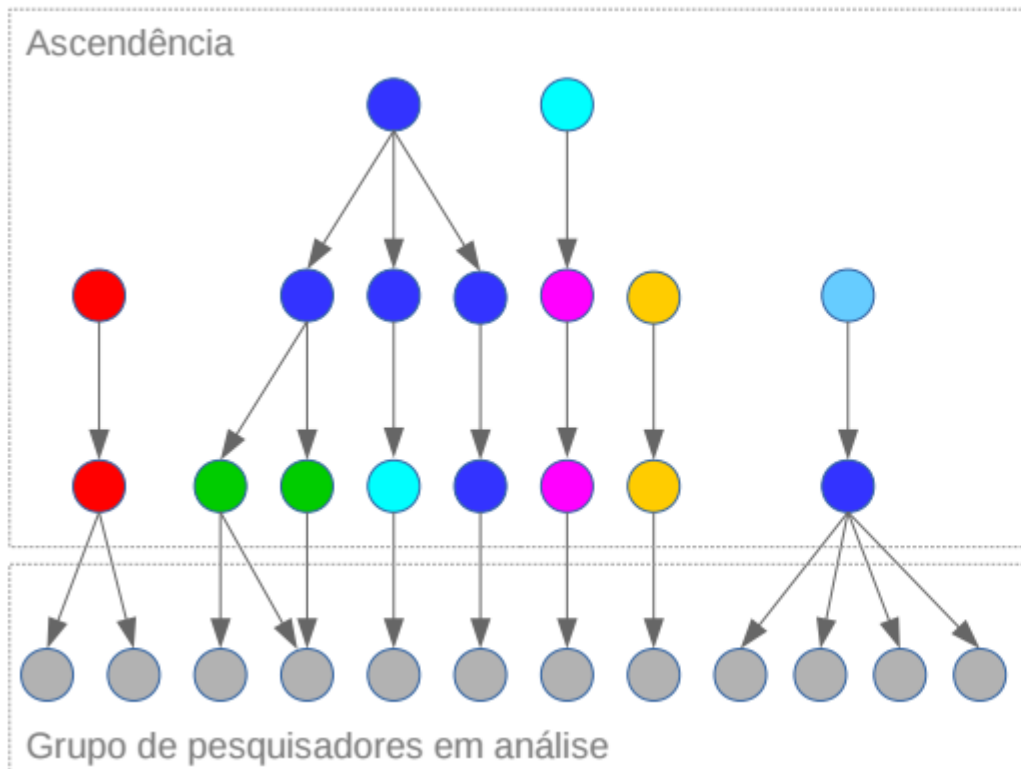
(a) na formação de recursos humanos

(b) transferência de conhecimento?



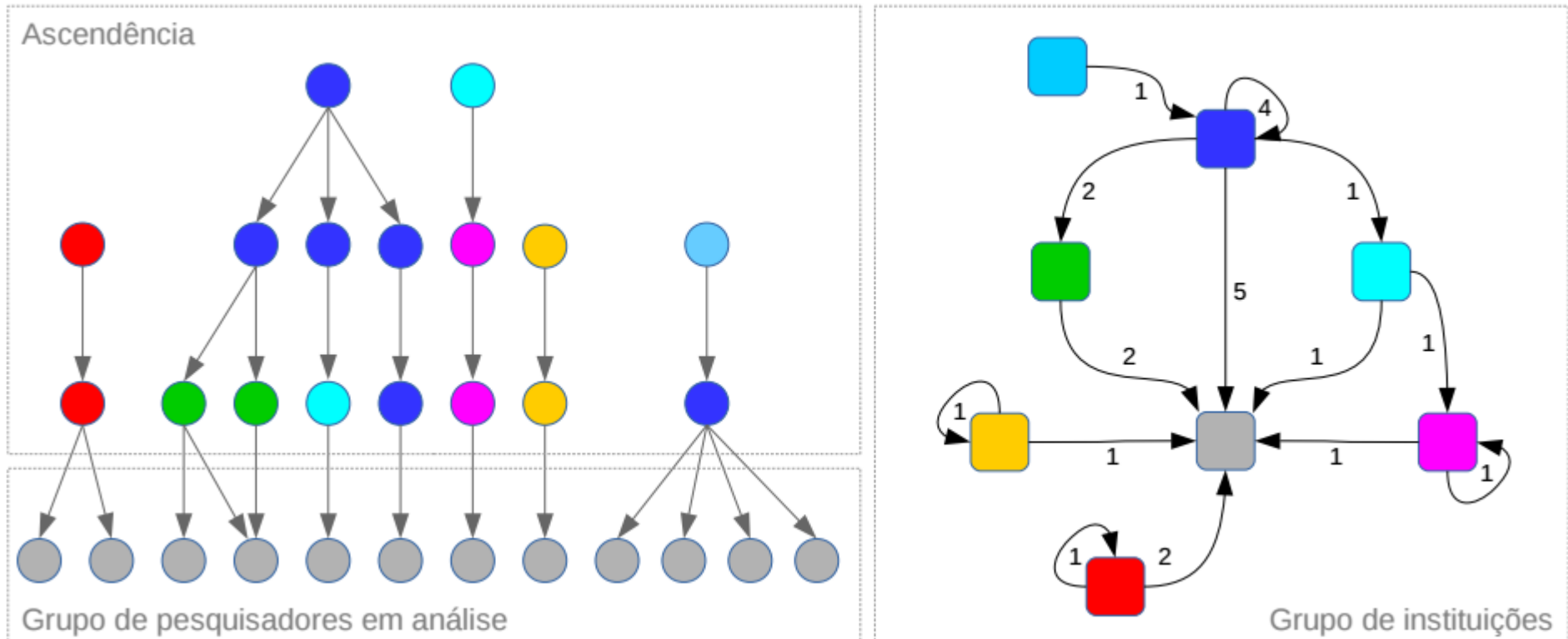
Nucleação de grupos/instituições

Pergunta: Qual a relevância de uma universidade/instituição para as demais?



Nucleação de grupos/instituições

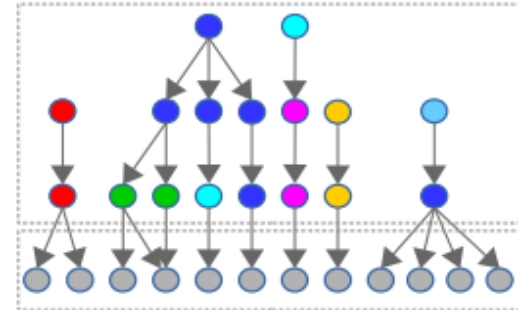
Pergunta: Qual a relevância de uma universidade/instituição para as demais?



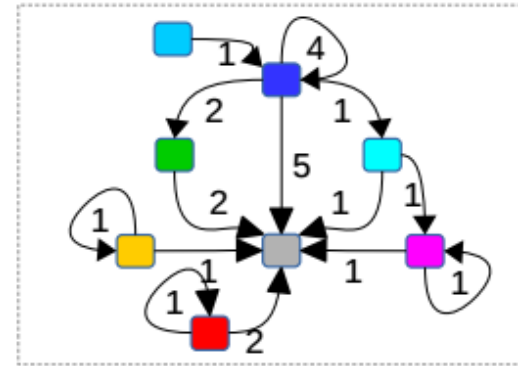
Grupo de pesquisadores



Construção do grafo de genealogia ascendente de pesquisadores



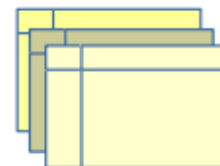
Construção do grafo Impacto/influência entre instituições acadêmicas



Caracterização e análise do impacto/influência de instituições acadêmicas através de técnicas de Aprendizado de Máquina

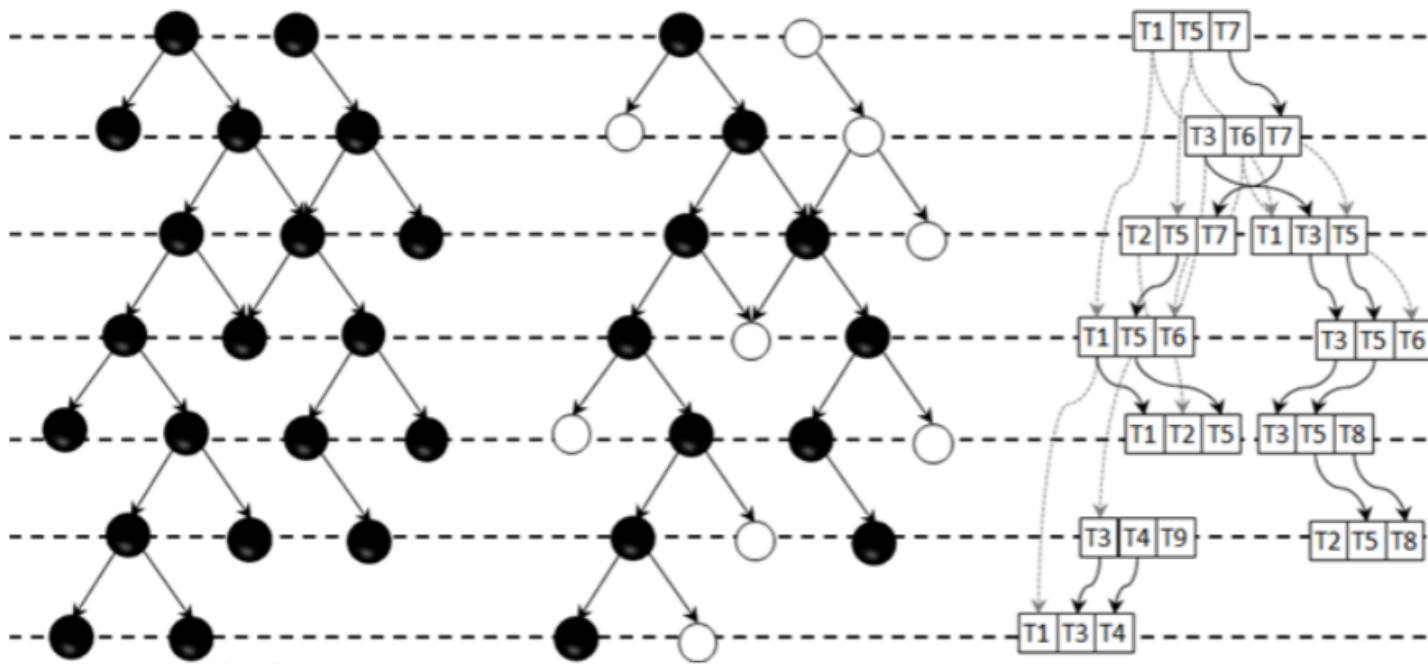
Grey square	12	0.45	23	346	A, ...
Cyan square	0	0.50	4	74	B, ...
Blue square	1	0.01	2	878	C, ...
Green square	4	0.87	28	940	D, ...
Cyan square	3	0.49	65	719	E, ...
Yellow square	10	0.10	42	713	F, ...
Red square	6	0.71	85	+345	G, ...
Magenta square	10	0.81	94	875	H, ...

Indicadores de impacto/influência

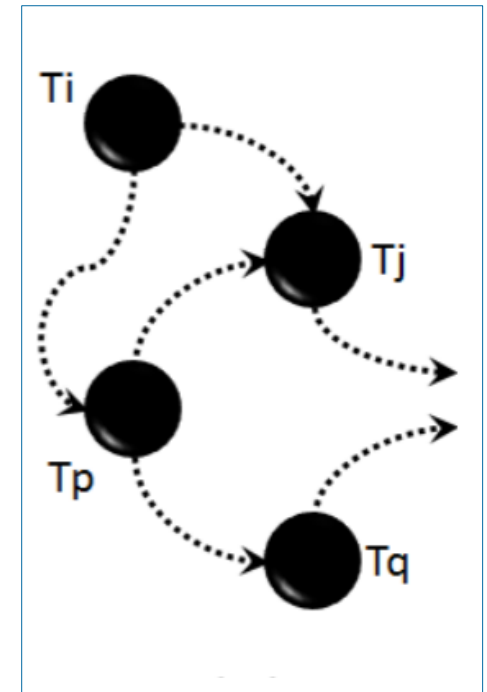


Fluxo/Disseminação de conhecimento

Pergunta: Como é realizada a disseminação de conhecimento nas estruturas de genealogia acadêmica?



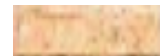
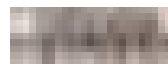
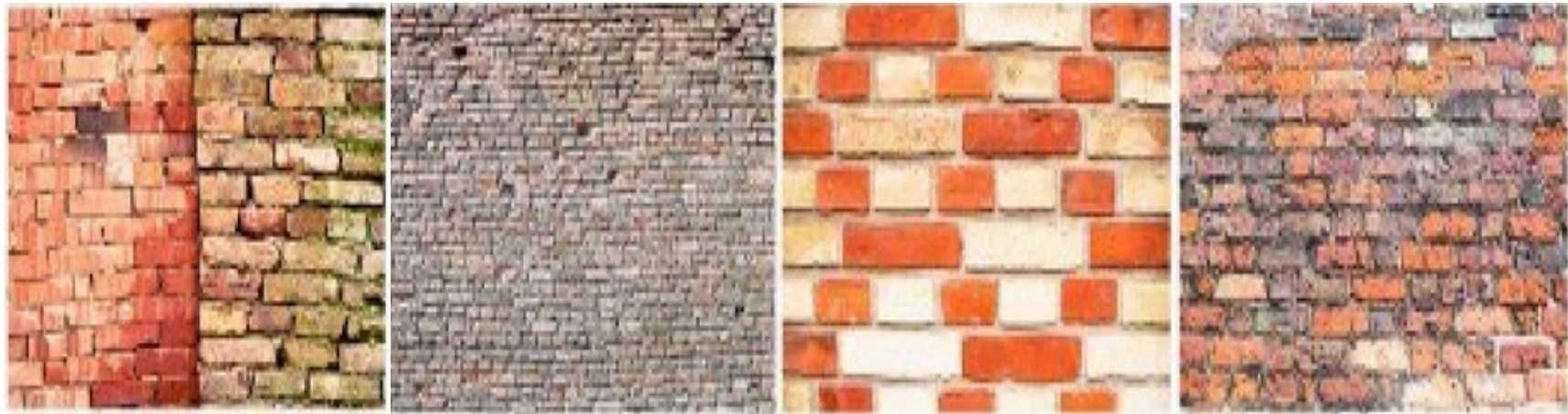
Grafo de genealogia



“proxy” para estudar transferência de conhecimento

Padrões: Motifs

Pergunta: Quais os padrões mais representativos nas redes de genealogia?



Padrões: Motifs

Um **motif** é um pequeno elemento de um todo, que pode ser analisado para dar informações sobre um conjunto.

Em música um **motif** ou **motivo** (pedra angular da composição) é um pequeno trecho de melodia.

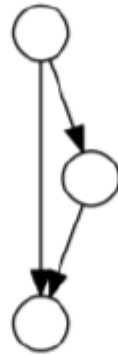
Um dos motifs “mais famosos” atualmente é o **composto por 4 notas** na Sinfonia N° 5 de Beethoven.

The image shows a musical score for the first movement of Beethoven's Symphony No. 5. It features two systems of staves. The first system has a Violins staff (treble clef) and a Bass instrs. staff (bass clef). The second system has a Vlns. staff (treble clef) and a Bass staff (bass clef). The key signature is one flat (B-flat major/D minor) and the time signature is 4/4. The music is marked *fp* (fortissimo piano). The first system shows the famous four-note motif in the bass line, with the notes G2, F2, E2, and D2 highlighted in red. The second system shows the same motif in the bass line, with the notes G2, F2, E2, and D2 highlighted in red. The Vlns. staff in both systems plays a melodic line that follows the contour of the bass line motif.



Padrões: Motifs

Único motif de tamanho 3



Rede original:

- Ocorrências: 3380

Redes aleatórias:

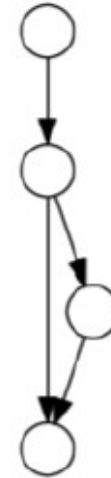
- Média de ocorrências: 662.07

- Desvio padrão: 20.40

Z-score = 133.22

P-value = 0.00

Único motif de tamanho 4



Rede original:

- Ocorrências: 3847

Redes aleatórias:

- Média de ocorrências: 832.70

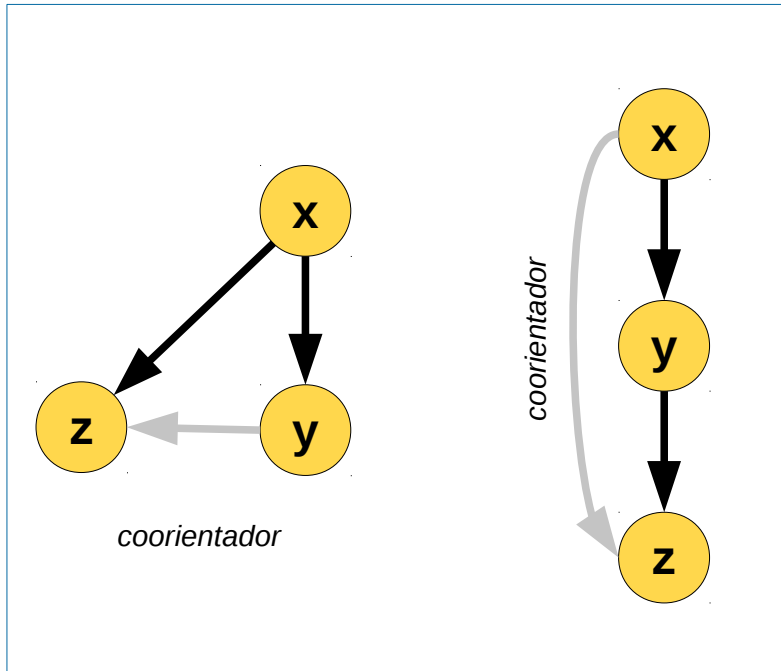
- Desvio padrão: 28.43

Z-score = 106.02

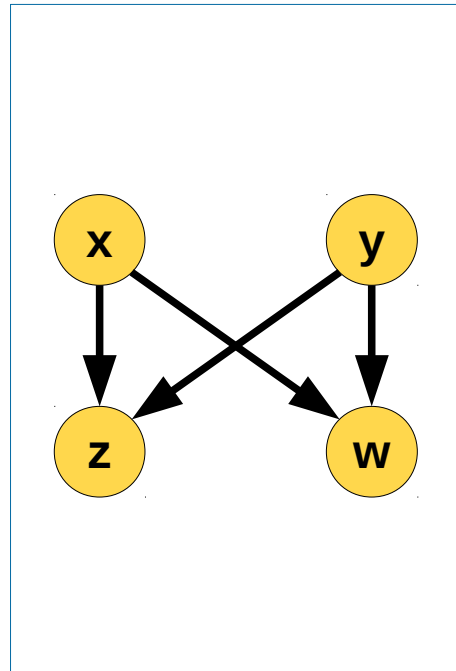
P-value = 0.00

Padrões: Motifs

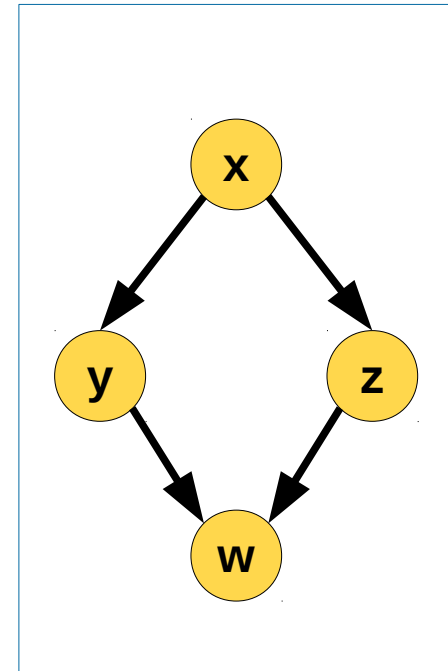
Preferência a coorientação



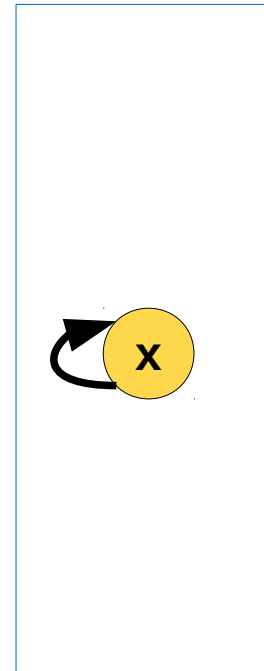
Feed-forward loop



Bi-fan



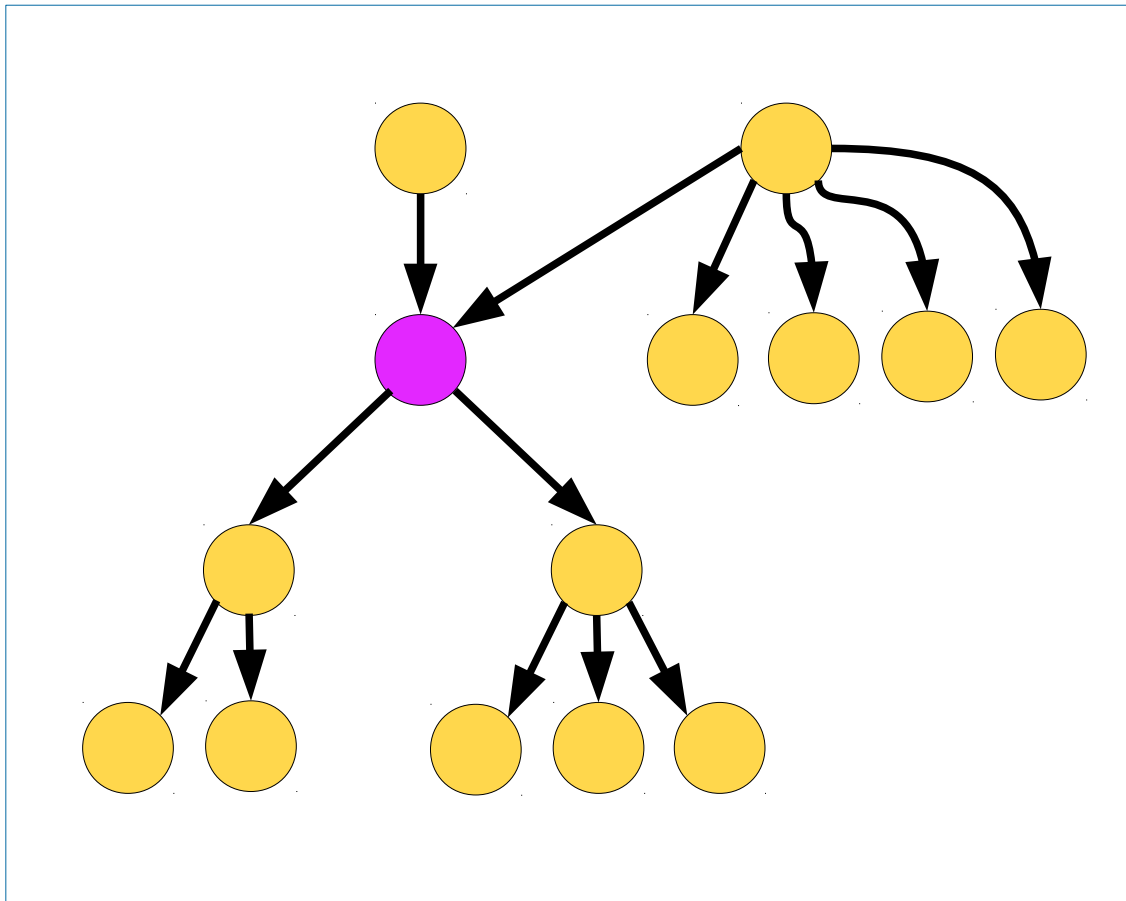
Bi-parallel



Auto-regulation

Paradoxo da amizade

Pergunta: Quais os padrões mais representativos nas redes de genealogia?



Nível	Paradoxo
Pai	$2 < 4/2=2$? Falso
Irmão	$2 < 0$? Verdadeiro
Filho	$2 < 5/2$? Verdadeiro

$$x_i < \frac{\sum_{j \in \Lambda_i, x_j}{x_j}}{K_i}$$

Eom, Y. H., & Jo, H. H. (2014).

Generalized friendship paradox in complex networks: The case of scientific collaboration. arXiv preprint arXiv:1401.1458.



Desafio não discutido ...

Qualidade/completude dos dados

Conjunto de dados	Período	Registros
ProQuest (T/D)	?-hoje	~3.8 milhões
Biblioteca nacional de teses e dissertações (NDLTD)	?-hoje	~1 milhão
Doutores em matemática	1380-hoje	~200 mil
Doutores em neurociências	2010-hoje	~100 mil
CAPES: Banco de T/D	1987-hoje	~700 mil
CNPq: Plataforma Lattes	1999-hoje	~4,5 milhões

Para finalizar ...

A genealogia acadêmica:

- Apresenta-se como uma **forma de analisar a projeção de um acadêmico** em várias gerações.
- Oferece uma **visão alternativa** aos já estudados na área:
 - A análise da produção bibliográfica não apresenta a **história completa** da evolução da ciência.

Para finalizar ...

- Os métodos computacionais estudados em “genealogia acadêmica” **não somente são uma aplicação** da **teoria dos grafos** para um domínio social em particular.
 - Veja como um **novo ramo específico** da teoria dos grafos em si, pois abre possibilidades de **criação de novos** algoritmos, axiomas e teoremas.
- O **genealogista acadêmico (cientista de dados de genealogia)** é um ator importante para a descoberta de conhecimento em grandes conjuntos de dados.



Os grafos de genealogia acadêmica dos programas de Pós-graduação em Ciência da Computação do Brasil

Jesús P. Mena-Chalco (UFABC)
jesus.mena@ufabc.edu.br

São Carlos, 14 de Setembro de 2016