



**cg ee**

Centro de Gestão e Estudos Estratégicos  
Ciência, Tecnologia e Inovação

# Monitoring and analyzing STI global and national trends: tools and methods

**Marcio de Miranda Santos** *Executive Director*

## Technical Team

Adriana Badaró  
Alessandra Brandão  
Cristiano Hugo Cagnin  
Eduardo Moresi  
Jackson Maia  
Rodrigo Leonardi  
Thyrso Villela Neto

November 19, 2015

## Origins

- Created in 2001
- Non-profit organization
- STI studies for economic growth competitiveness and well being in Brazil

## Societal Control

- Overseen by the Ministry of Science Technology and Innovation
- Subjected to the Brazilian auditing institutions

## Size and Budget

- Staff ~ 90 employees
- Annual budget ~ 15M USD

## Strengths

- ~ 2000 experts mobilized from nearly 300 institutions per year
- ~ 400 studies in STI so far

# Strategic perspectives for investing in new tools and methods for the monitoring and analysis of large volumes of data

CGEE as a key player in providing intelligence services in STI for national and international institutions, specifically on:

- Making use of the Center's ample capacity to articulate distinct actors around a diversity of complex issues and demands

- Improving the Center's ability to anticipate future strategic STI issues impacting the future of priority economic sectors

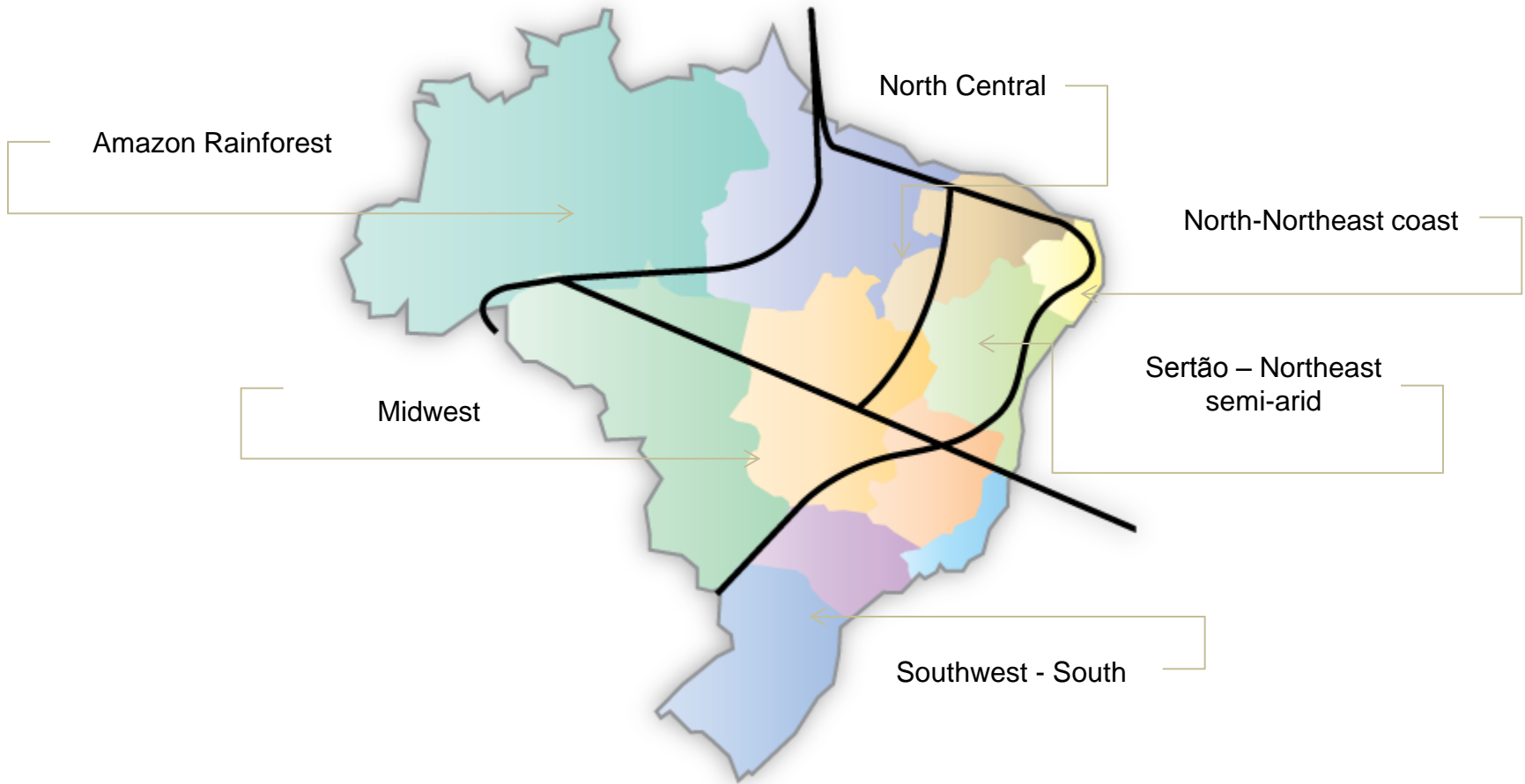
- Responding to the STI dynamics and complexity changes in contexts and enabling environments

- Developing appropriate tools and methods to capture, process and monitor large amounts of data coming from large number of data sources



cg<sup>ee</sup>

Centro de Gestão e Estudos Estratégicos  
*Ciência, Tecnologia e Inovação*



## **Information and Knowledge Management**

Services must comply with:

*Very large amounts of data and information*

*Complexity of issues*

*Diversity of opinions*

*High levels of uncertainties*

Current efforts are focusing in

*Accelerate the processes for identification, monitoring and interpretation of trends in STI to promptly inform decision-makers, based on:*

New tools and methods

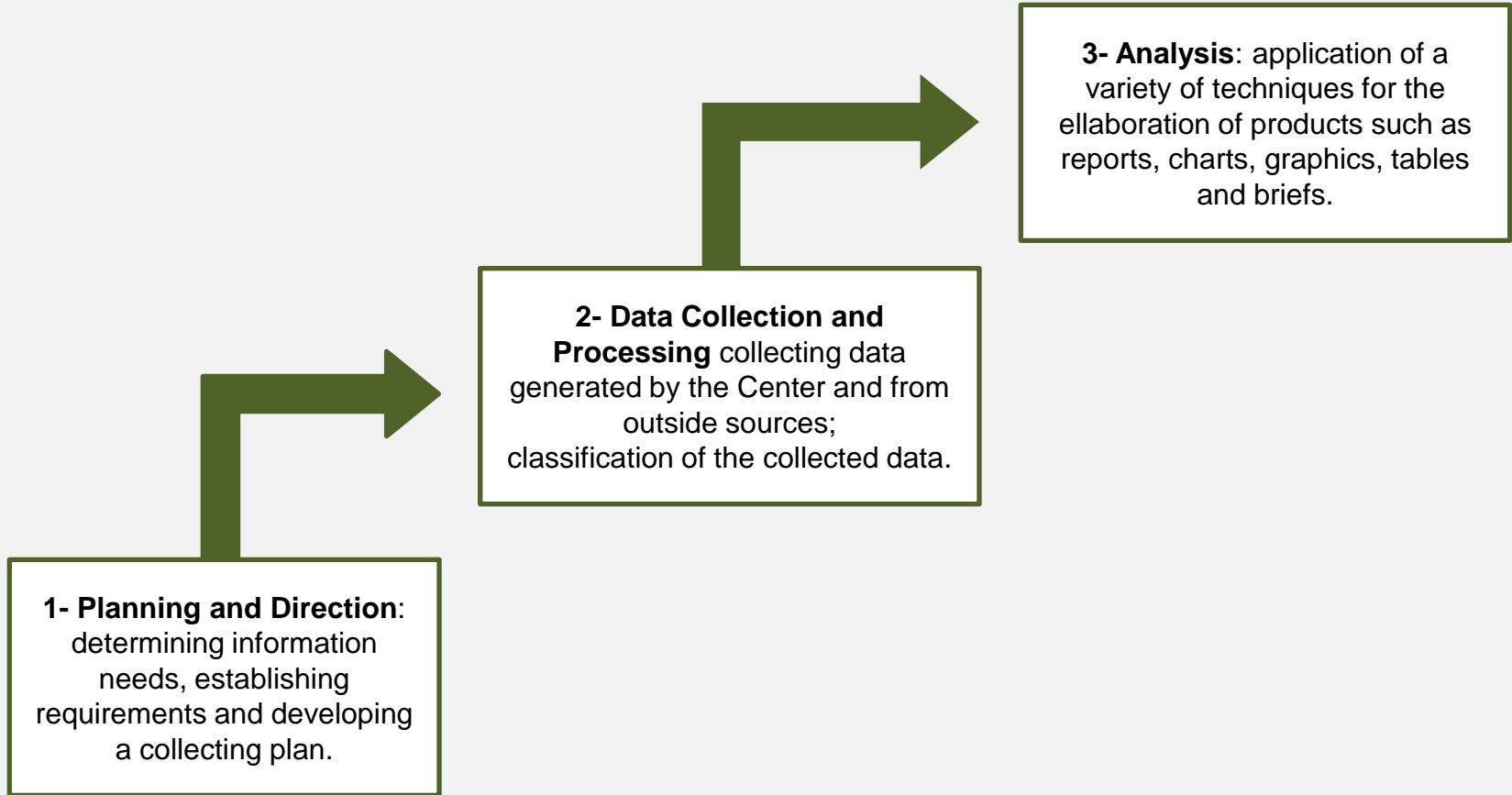
Automation

Networking of experts

Visualization

Modularity and flexibility





## Planning and Direction

Identification of information needs: keywords sources, brands and names of companies and institutes to be monitored, scientific journals and technological production (patents)

## Data collection:

Development of user-agents to collect papers and patents

Gathering of news in selected magazines

## Natural language processing:

Keywords, named-entities, and key phrases extraction

Part of speech identification

Cosine similarity (measure cohesion within [clusters](#) in the field of [data mining](#))

Document language recognition

TF/IDF calculation (term frequency–inverse document frequency)

Export data and information to Gephi software

## Classification of the collected information



# Monitoring and data analysis



cgée Centro de Gestão e Estudos Estratégicos Ciência, tecnologia e inovação

Eduardo Moresi > Documentos

Documentos

Configurações

Taxonomias

Pastas

Downloads 2

Expressão: AOCs attitude orbit control system

Tipos Doc.: Nenhum selecionado Fontes: Nenhum selecionado Idioma: Nenhum sel...

Período: [ ] [ ] Ordenar: Relevância Decrescente

Pesquisar

Resultado da pesquisa por Documentos (28494 resultados)

Selecionar todos (10) na página

Selecionar todos (28494) na pesquisa

Anterior 1 2 3 4 5 6 7 8 9 10 Próximo

[SpaceWire-based thermal-infrared imager system for asteroid sample return mission HAYABUSA2](#)

★ Artigos SPIE - Applied Remote Sensing (JARS)  
<http://remotesensing.spiedigitallibrary.org/article.aspx?articleid=1867190>  
HAYABUSA2 Mission Japan Aerospace Exploration Agency/Institute of Space and Astronautical Science (JAXA/ISAS) developed a thermal-infrared (TIR) imager system for HAYABUSA2, which consists of a TIR imager and a digital electronics (DE). The system is due to be launched in 2014 aiming at retrie...

[Kalman-filtering-based angular velocity estimation using infrared attitude information of spacecraft](#)

★ Artigos SPIE - Optical Engineering (OE)  
<http://opticalengineering.spiedigitallibrary.org/article.aspx?articleid=1076018>  
vazio

13/07/1999 PoS Tf-Idf

[Multiagent Attitude Control System for Satellites Based in Momentum Wheels and Event-Driven Synchro...](#)

★ Artigos Springer - Journal Astronautical Sciences  
<http://link.springer.com/article/10.1007/s40295-014-0010-4/fulltext.html>  
Introduction and State of the Art Though it is possible to stabilize a satellite without an active controller using spin stabilization or gravitational gradients, most spacecrafts carry an attitude control system on board for steering the satellite. Normally it is integrated together with the or...

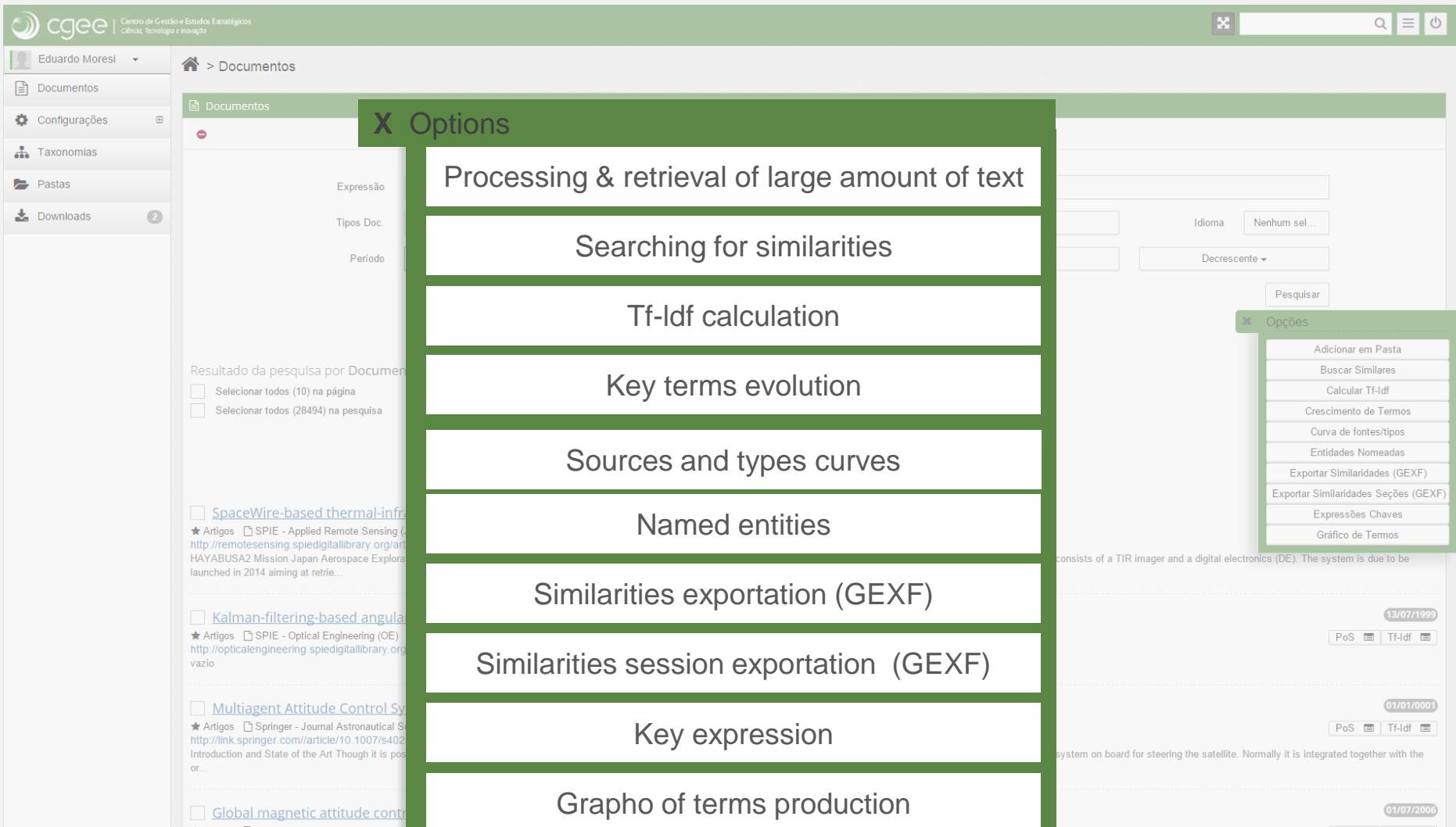
01/01/0001 PoS Tf-Idf

[Global magnetic attitude control of spacecraft in the presence of gravity gradient](#)

01/07/2006

Opções

- Adicionar em Pasta
- Buscar Similares
- Calcular TF-Idf
- Crescimento de Termos
- Curva de fontes/tipos
- Entidades Nomeadas
- Exportar Similaridades (GEXF)
- Exportar Similaridades Seções (GEXF)
- Expressões Chaves
- Gráfico de Termos



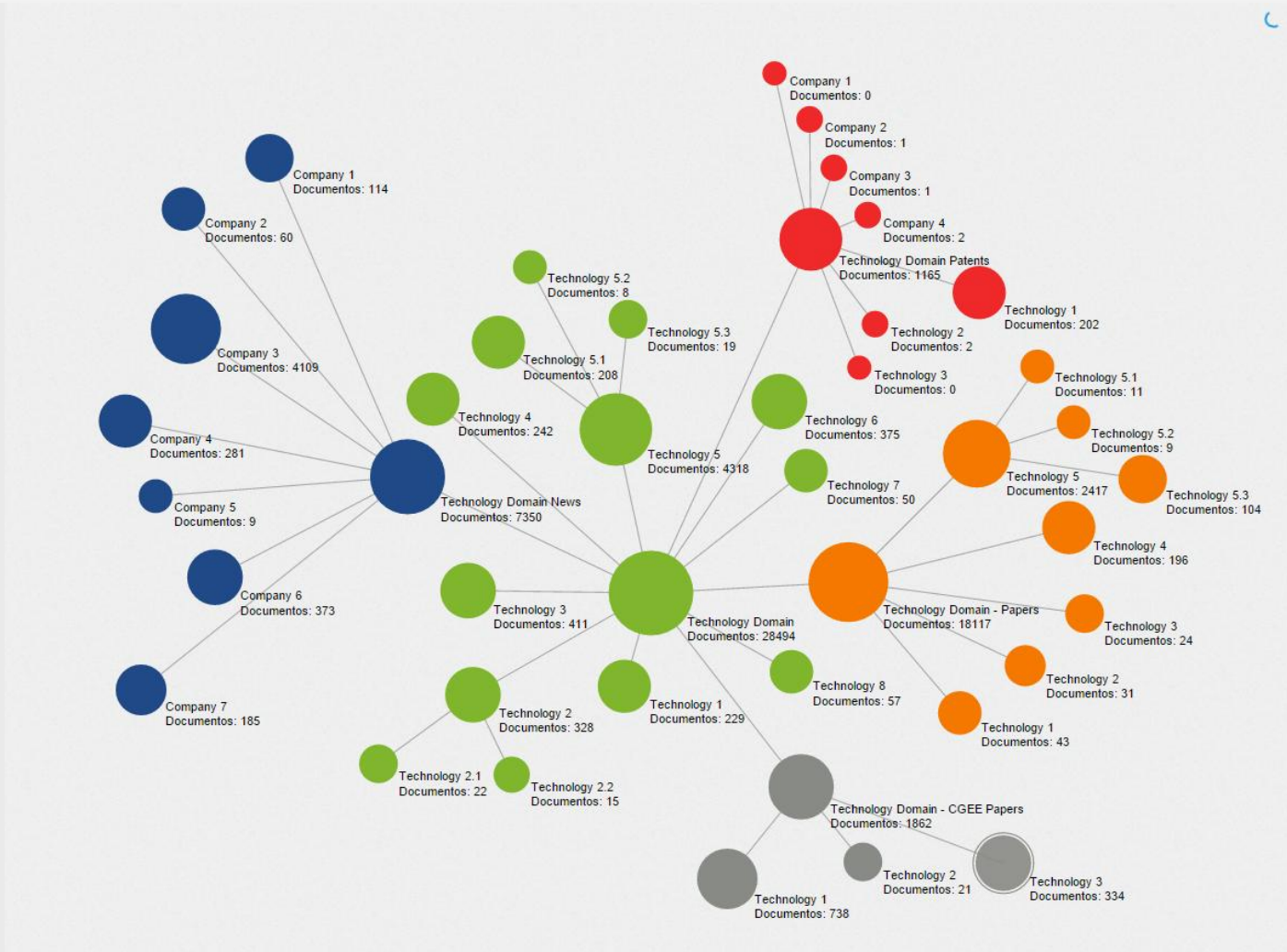
The image shows a screenshot of the cgEE web interface. The interface includes a top navigation bar with the cgEE logo and the text "Centro de Gestão e Estudos Estratégicos ciência, tecnologia e inovação". Below the navigation bar, there is a sidebar with menu items: "Eduardo Moresi", "Documentos", "Configurações", "Taxonomias", "Pastas", and "Downloads". The main content area displays a search results page for "Documentos". A green overlay is positioned in the center, listing various analysis options. To the right of the overlay, a search form is visible with fields for "Idioma" (set to "Nenhum sel...") and "Decrescente", and a "Pesquisar" button. Below the search form, a dropdown menu titled "Opções" is open, showing a list of analysis options: "Adicionar em Pasta", "Buscar Similares", "Calcular TF-Idf", "Crescimento de Termos", "Curva de fontes/tipos", "Entidades Nomeadas", "Exportar Similaridades (GEXF)", "Exportar Similaridades Seções (GEXF)", "Expressões Chaves", and "Gráfico de Termos".

**X Options**

- Processing & retrieval of large amount of text
- Searching for similarities
- Tf-Idf calculation
- Key terms evolution
- Sources and types curves
- Named entities
- Similarities exportation (GEXF)
- Similarities session exportation (GEXF)
- Key expression
- Grapho of terms production

# Monitoring and data analysis

Semi-automatic elaboration of taxonomies for document classification and real time update (the nodes of the tree represent selected informational search expressions)

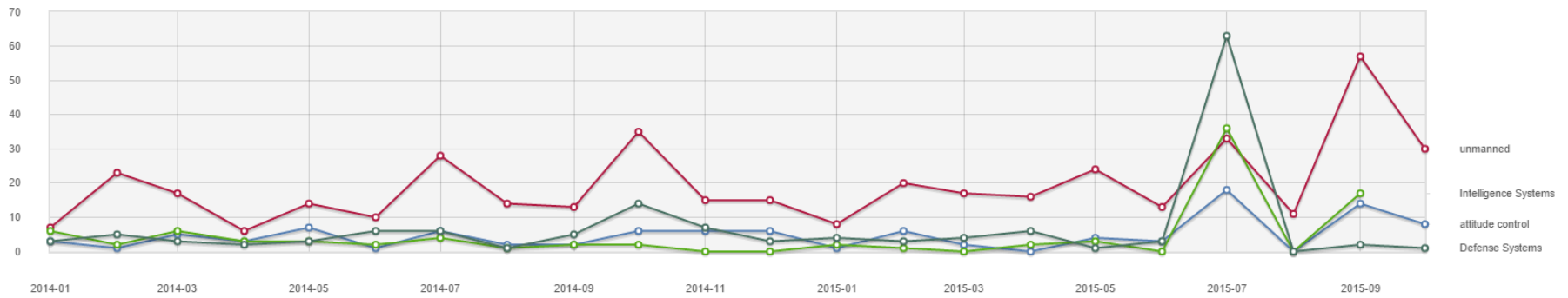


# Evolution of terms and entities



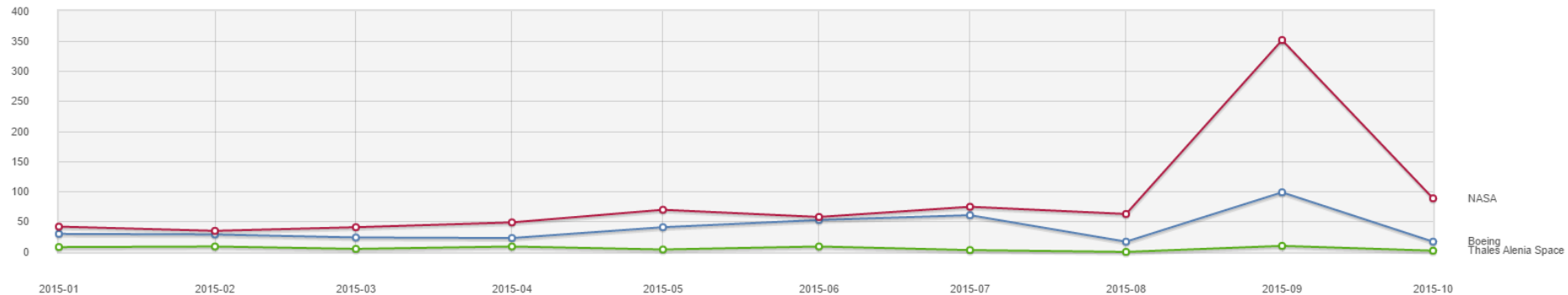
Ocorrências em [ Artigos, Artigos\_CGEE, Notícias, Notícias\_Plugar, Patentes ]

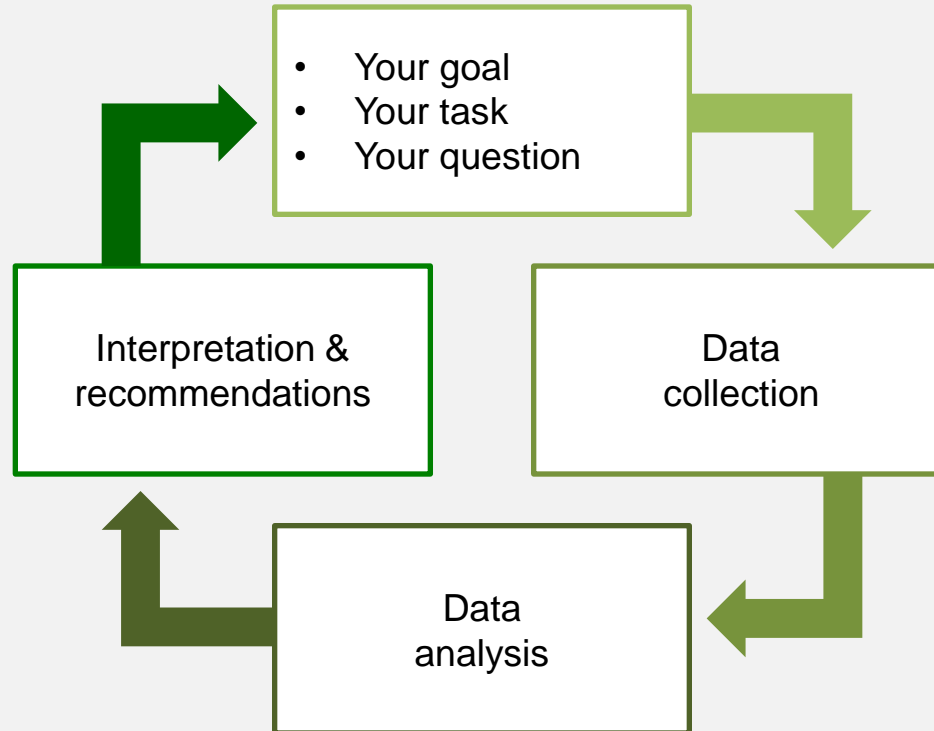
- attitude control
  unmanned
  Intelligence Systems
  Defense Systems



Ocorrências em [ Artigos, Artigos\_CGEE, Notícias, Notícias\_Plugar, Patentes ]

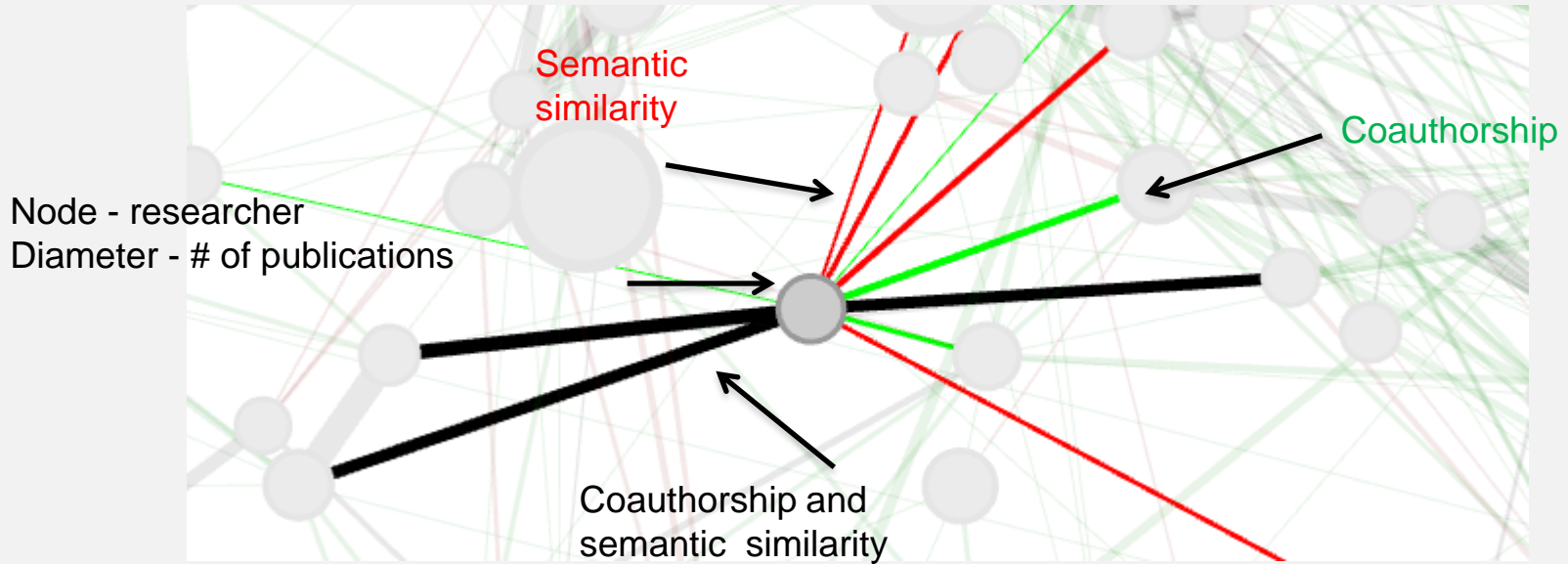
- Boeing
  NASA
  Thales Alenia Space

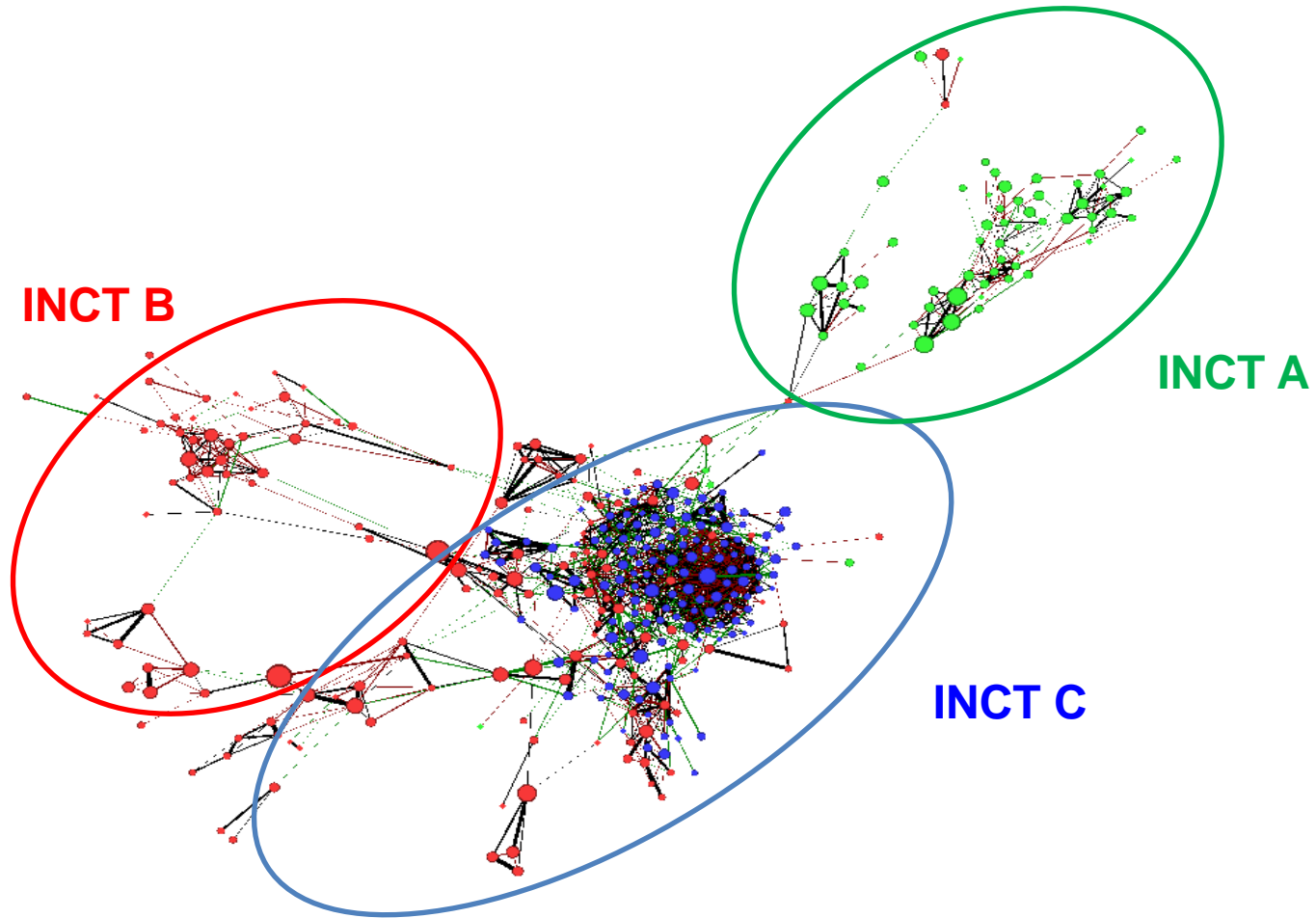




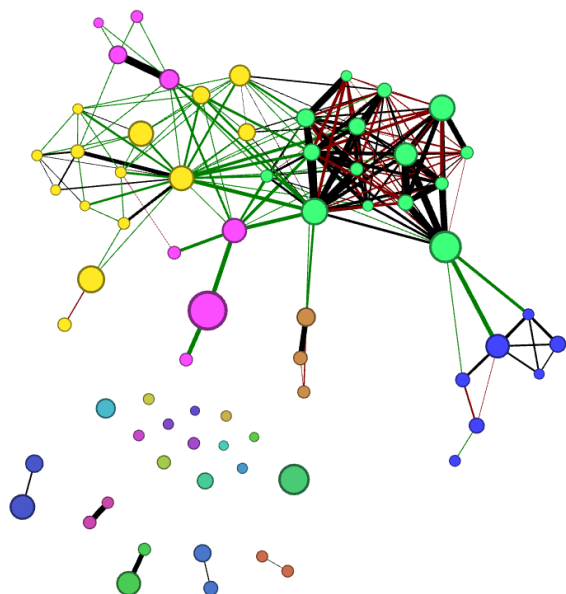
Working relationship networks  
Similarity networks  
Integrates global-local information  
Qualitative and quantitative framework

Main source of data and information  
CNPq “Lattes database” with 4.4M registered curricula

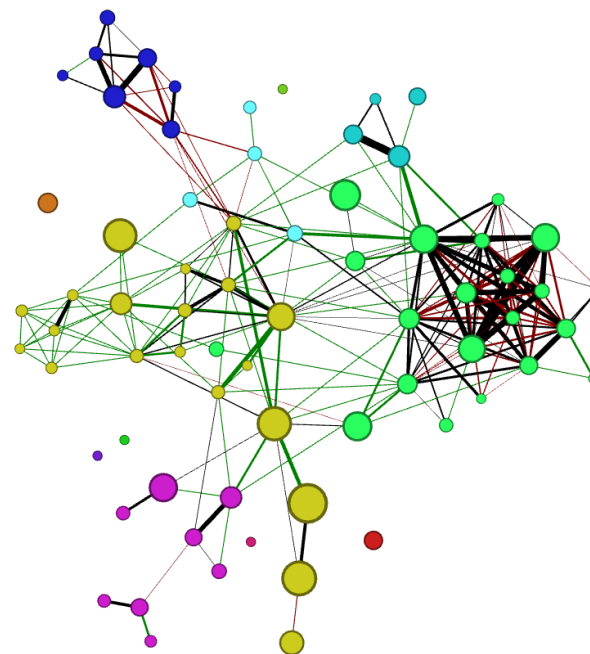




2001-2007



2008-2013



Nº médio de  
colaborações →

4,4

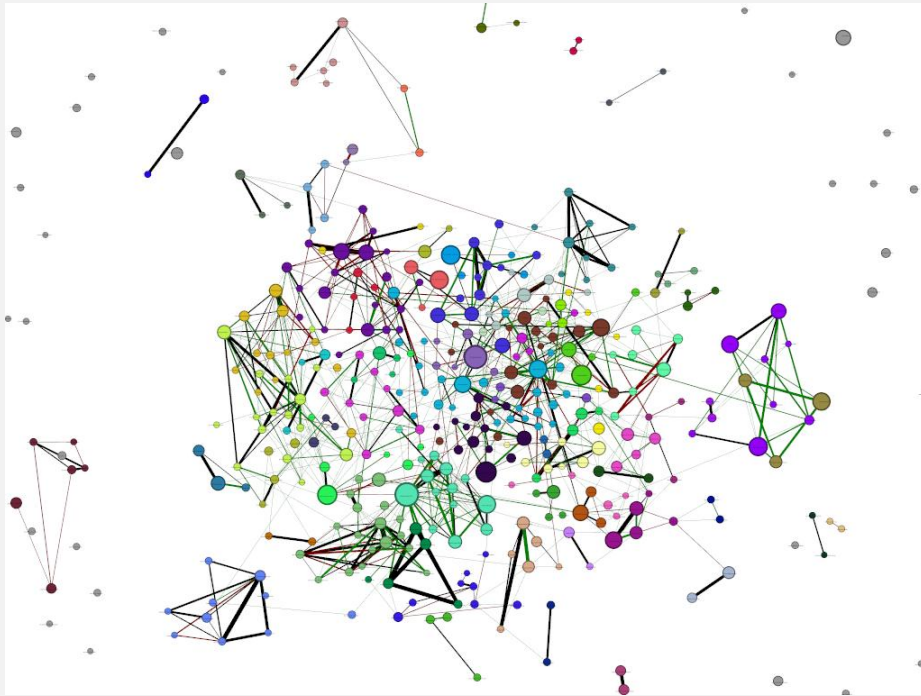
**Aumento  
de 14 %**

Nº médio de  
colaborações →

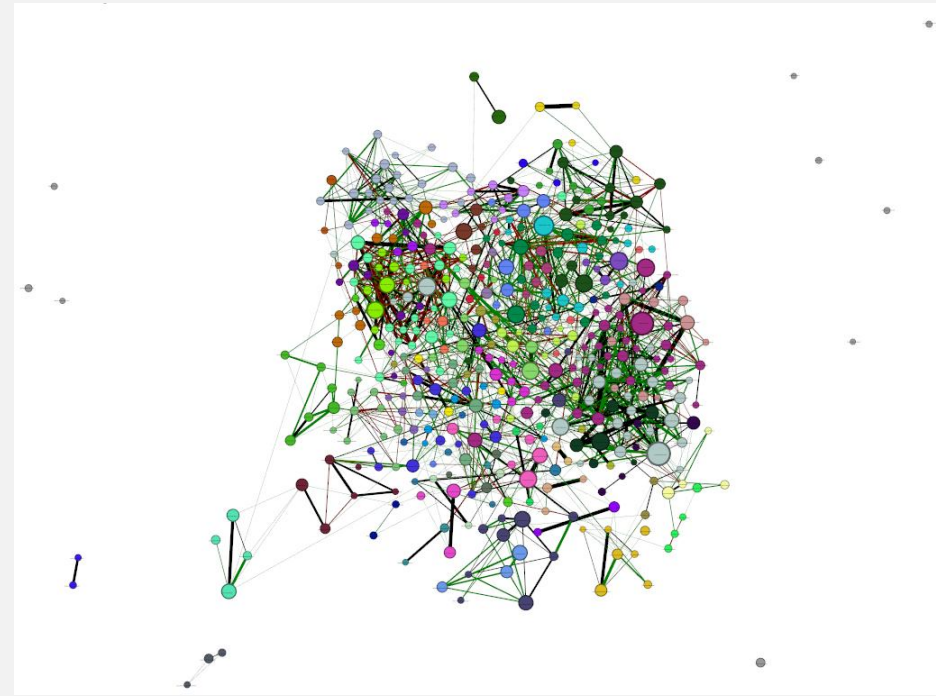
5,02



2007

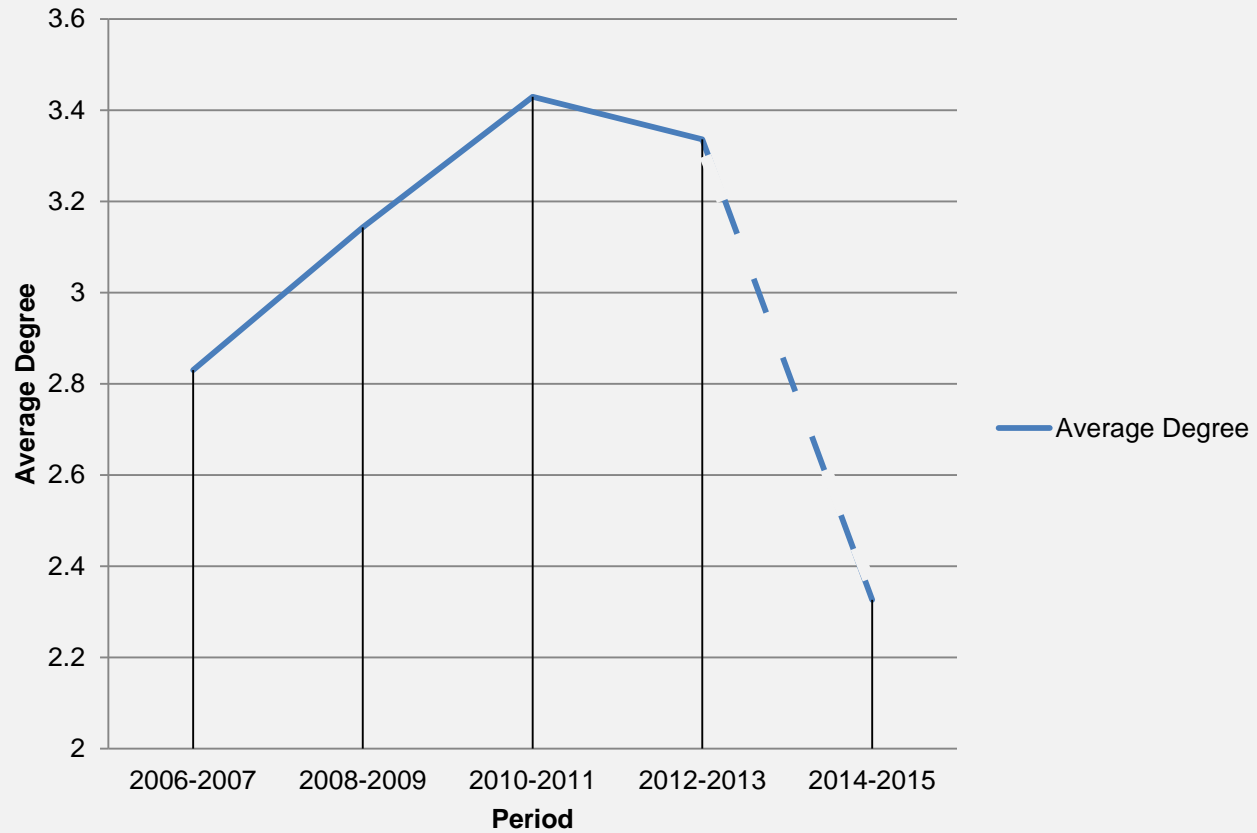


2014

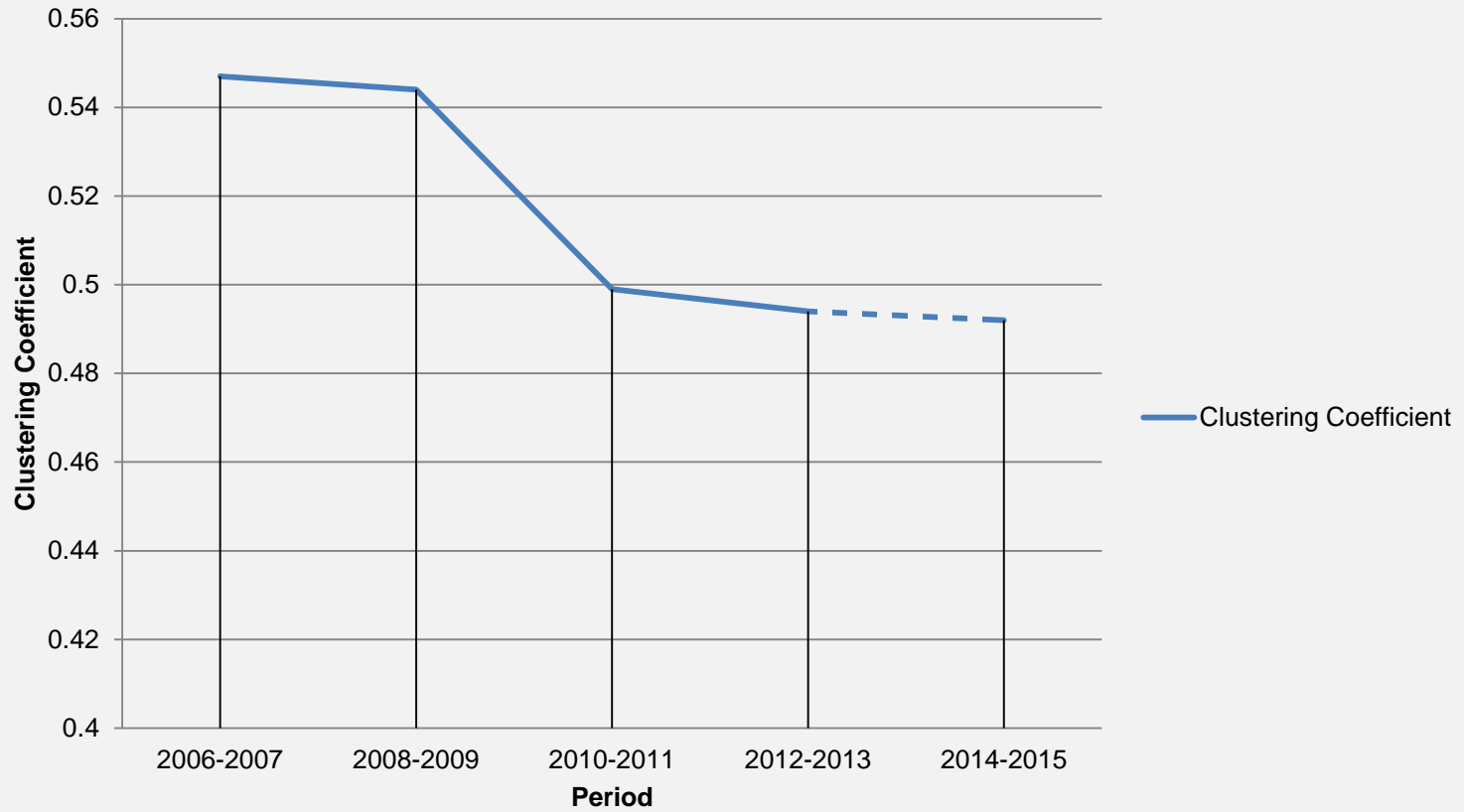




# Average Collaboration Evolution



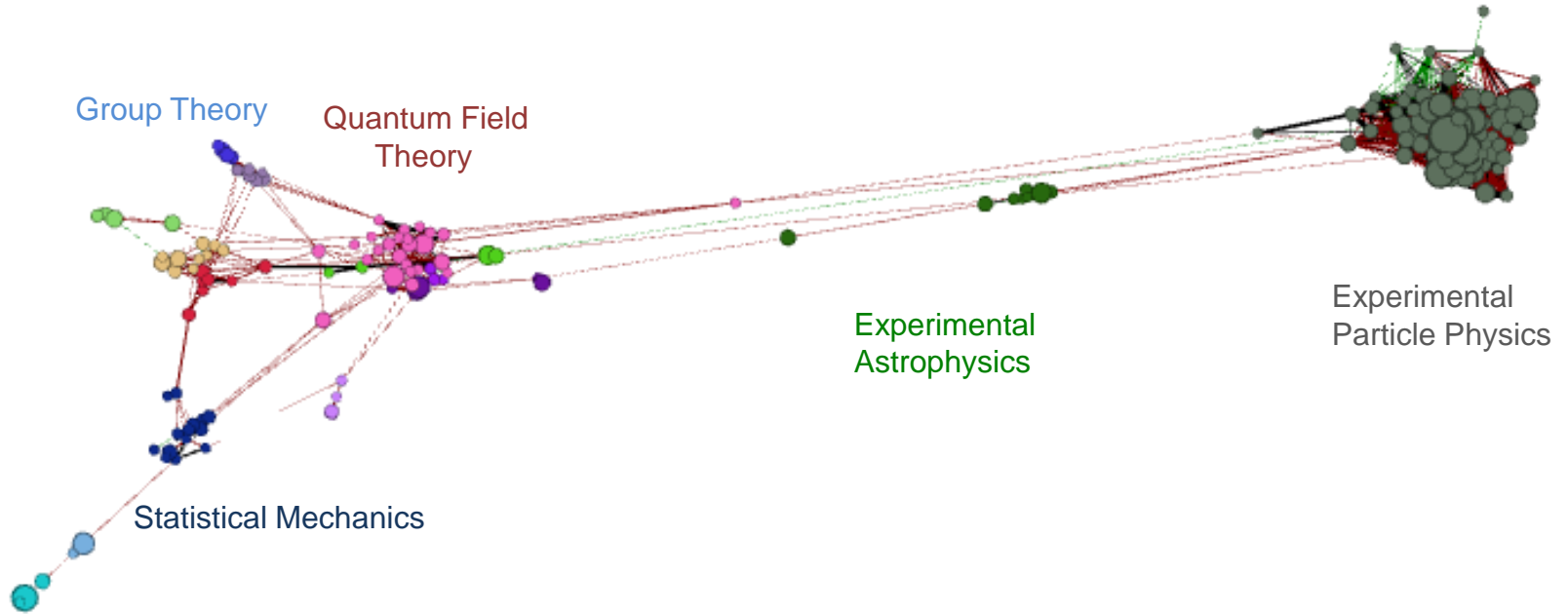
# Average Clustering Coefficient Evolution (segregation measure)



# Cooperation from Brazilians with Russians

Searching term “Mathematics”

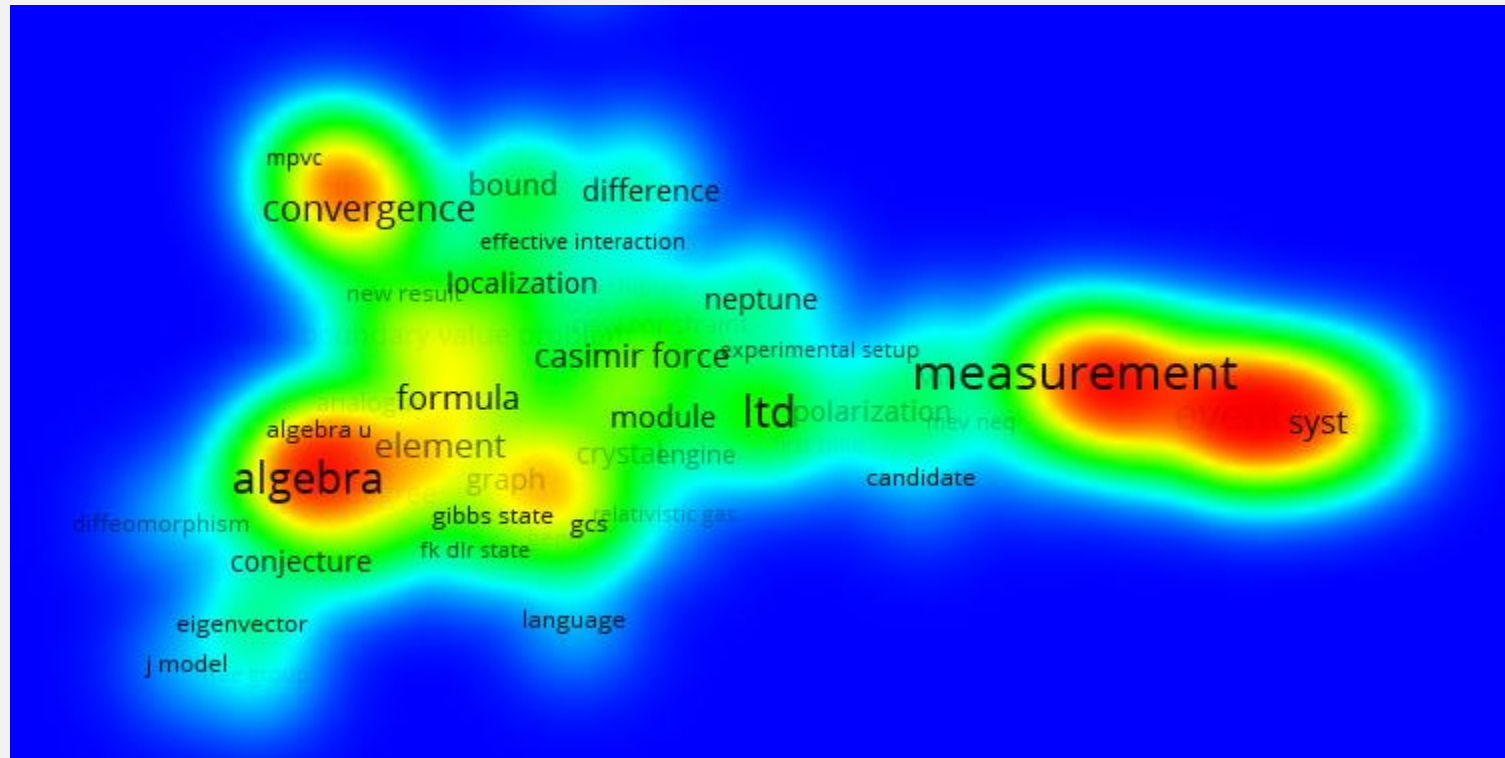
(Scopus + WoS →DOI →”CNPq Lattes” Database)



(Thanks to Maxim Kotsemir ISEK/HSE )

# Cooperation from Brazilians with Russians

Searching term “Mathematics” (Scopus database)



(Thanks to Maxim Kotsemir ISEEK/HSE )

# Main institutions (WoS coauthorships)



(Thanks to Maxim Kotsemir ISEEK/HSE )



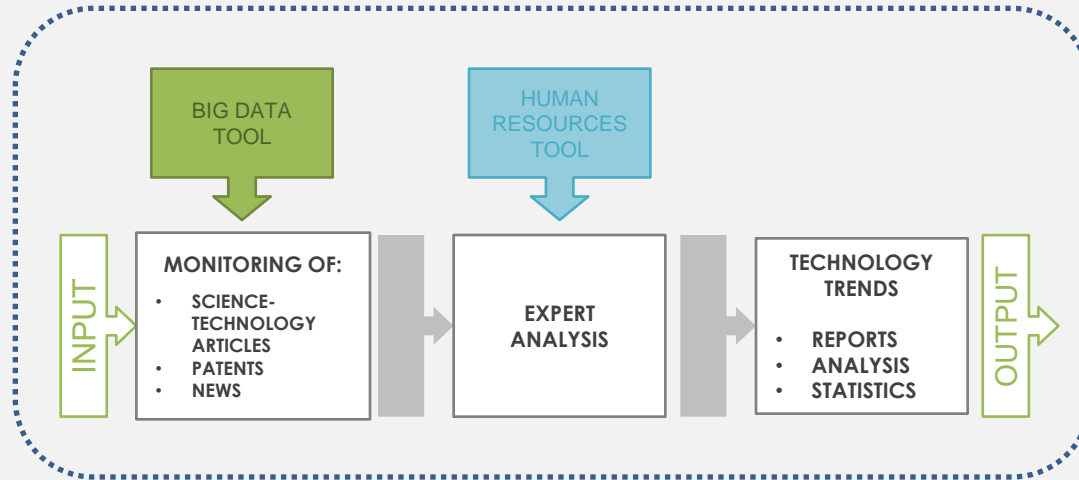
Centro de Gestão e Estudos Estratégicos

**OBSERVATÓRIO DE  
TECNOLOGIAS ESPACIAIS**

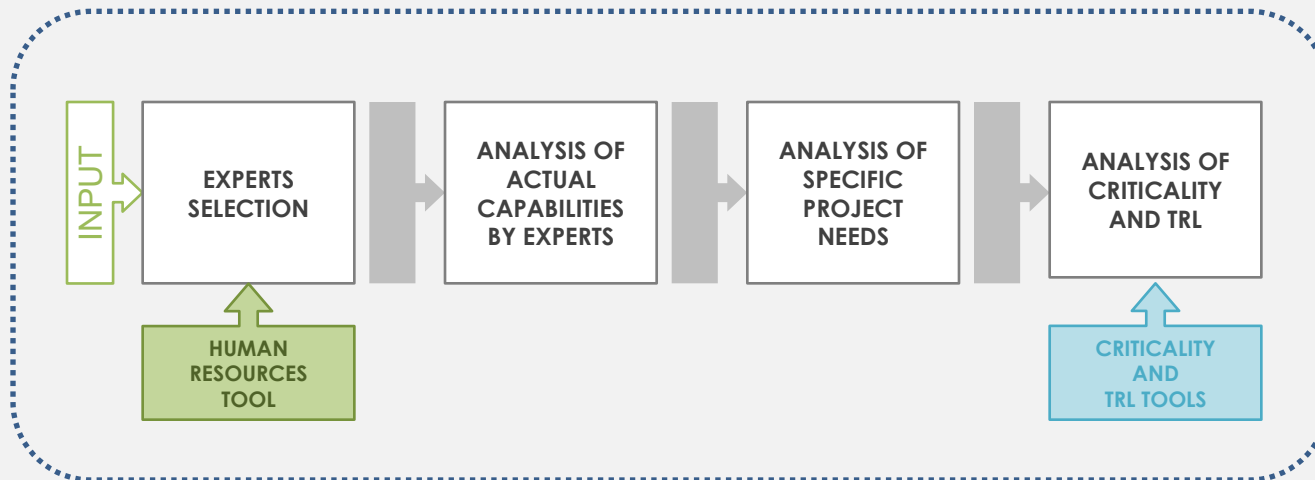
- Análise de criticidade de tecnologias
- Análise maturidade/prontidão tecnológica – TRL
- Análise de redes de competências
- Banco de patentes (desenvolvido: ~32.000 patentes)
- Ferramenta de monitoramento e análise de grandes volumes de informações textuais (Big Data)

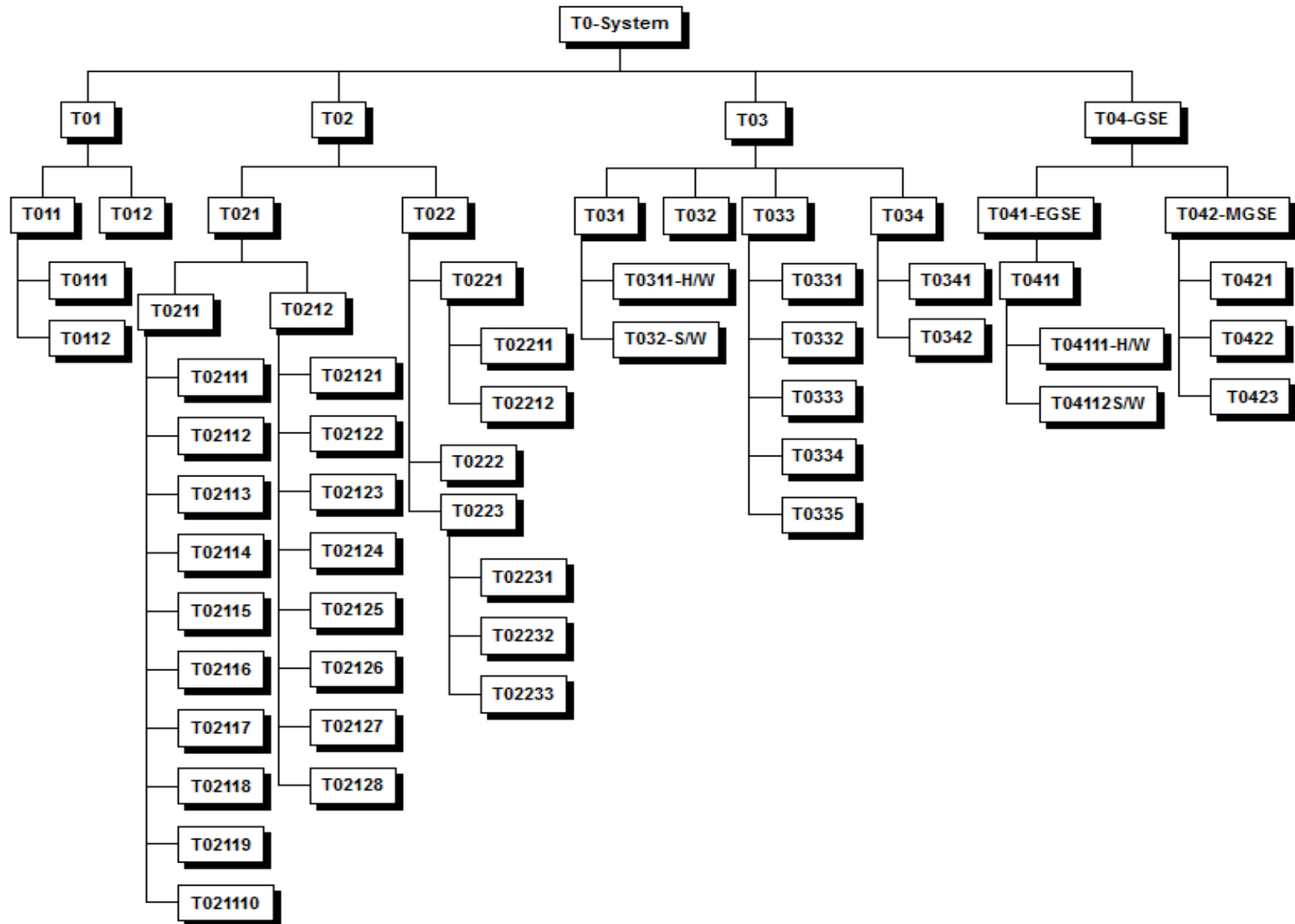


## TECHNOLOGY DOMAIN OBSERVATIONS

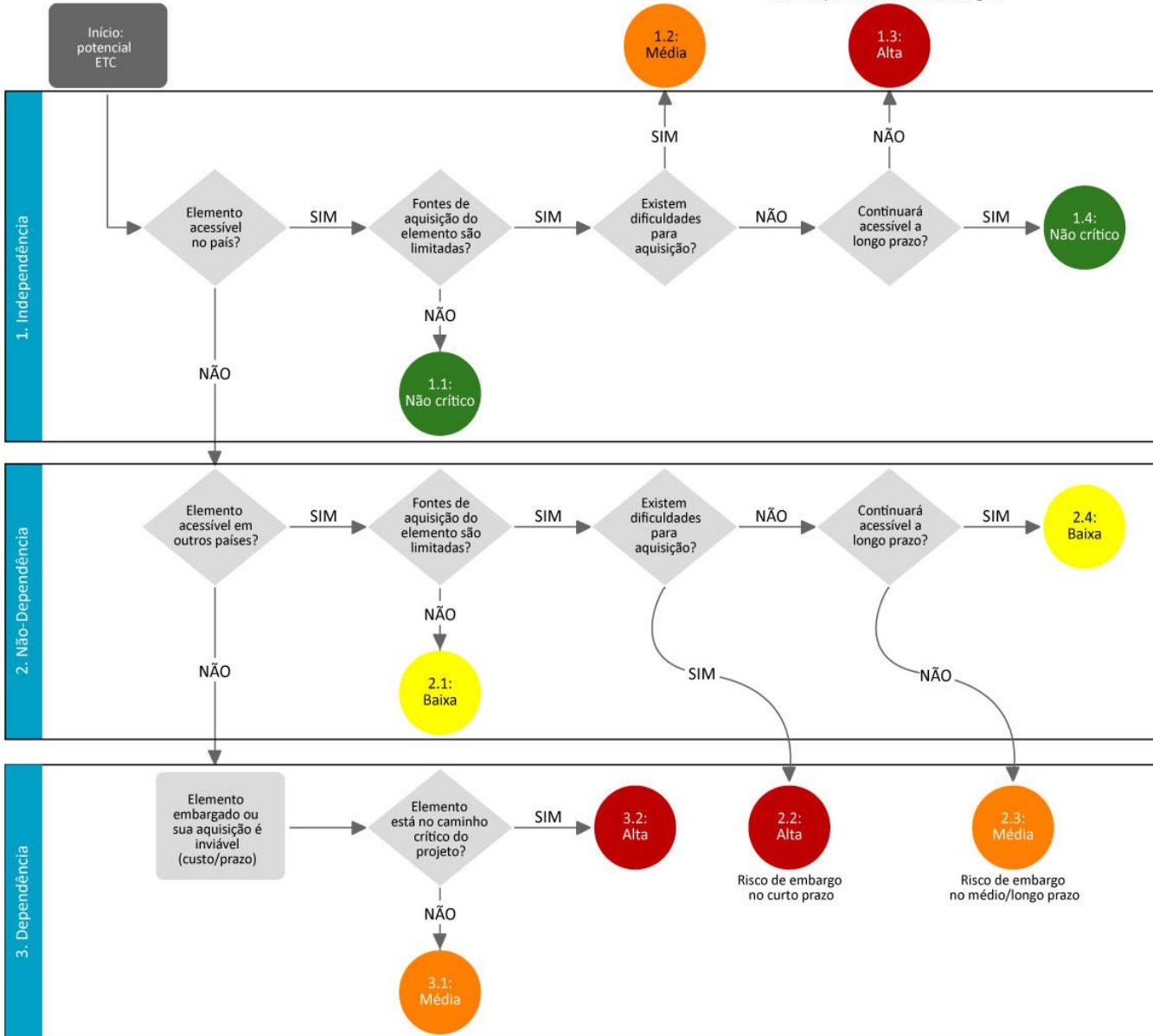


## SPECIFIC TECHNOLOGY OBSERVATIONS

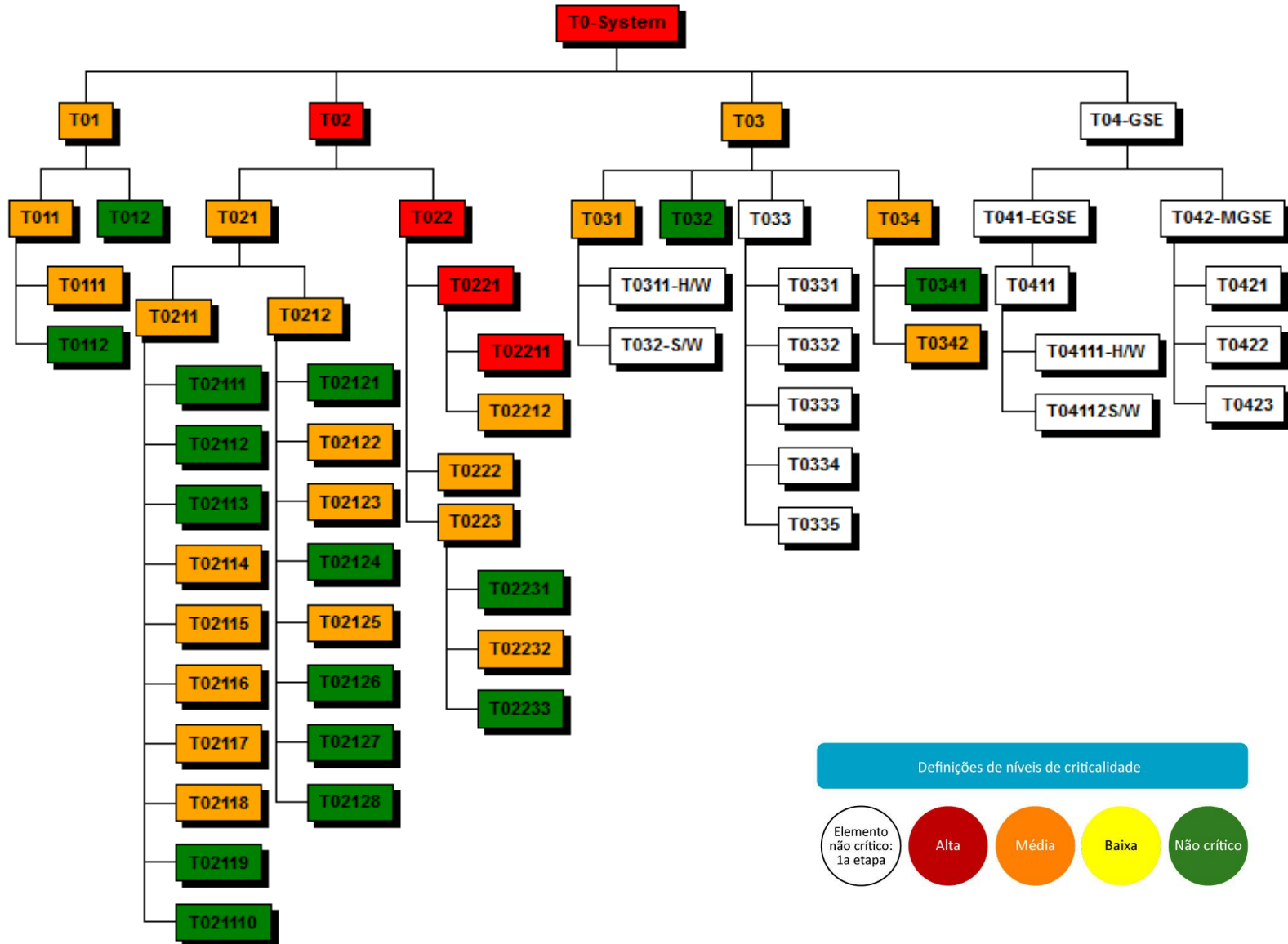


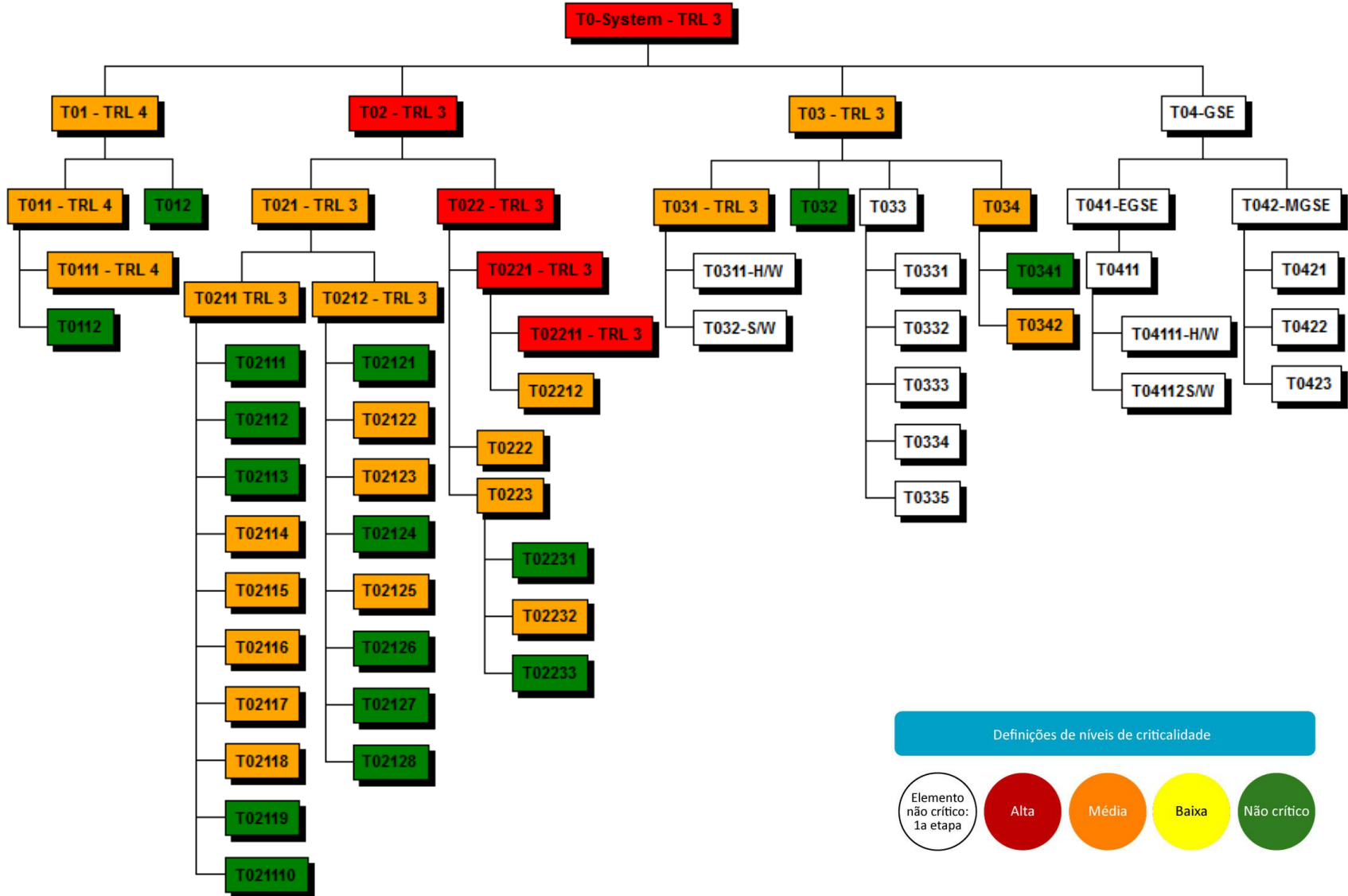


Risco de perda do domínio tecnológico



Definições de níveis de criticidade	
Alta	A: O elemento é crítico, porém é necessário fazer uma análise detalhada de alternativas que viabilizem o desenvolvimento da tecnologia B: Ações devem ser definidas imediatamente C: Ações já estão em andamento e devem ser monitoradas
Média	A curto prazo, plano de ação deve ser definido
Baixa	Nenhum plano de ação específico é necessário para este elemento no curto prazo
Não crítico	Elemento considerado não crítico





Definições de níveis de criticidade

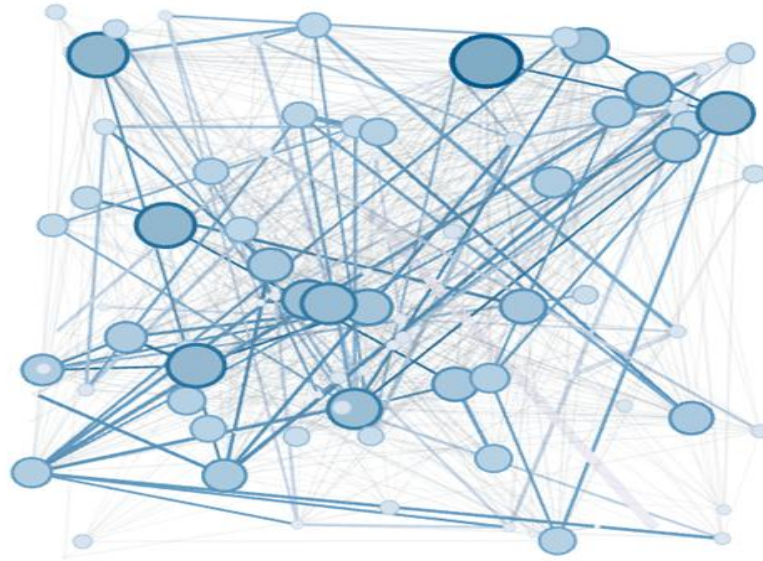
Elemento não crítico: 1a etapa

- Alta
- Média
- Baixa
- Não crítico



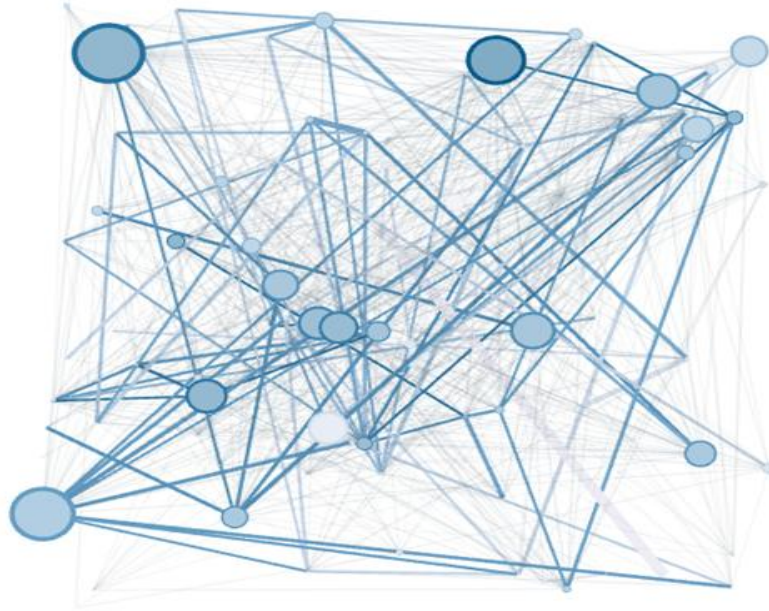
# Articles networking

## Eigenvector centrality



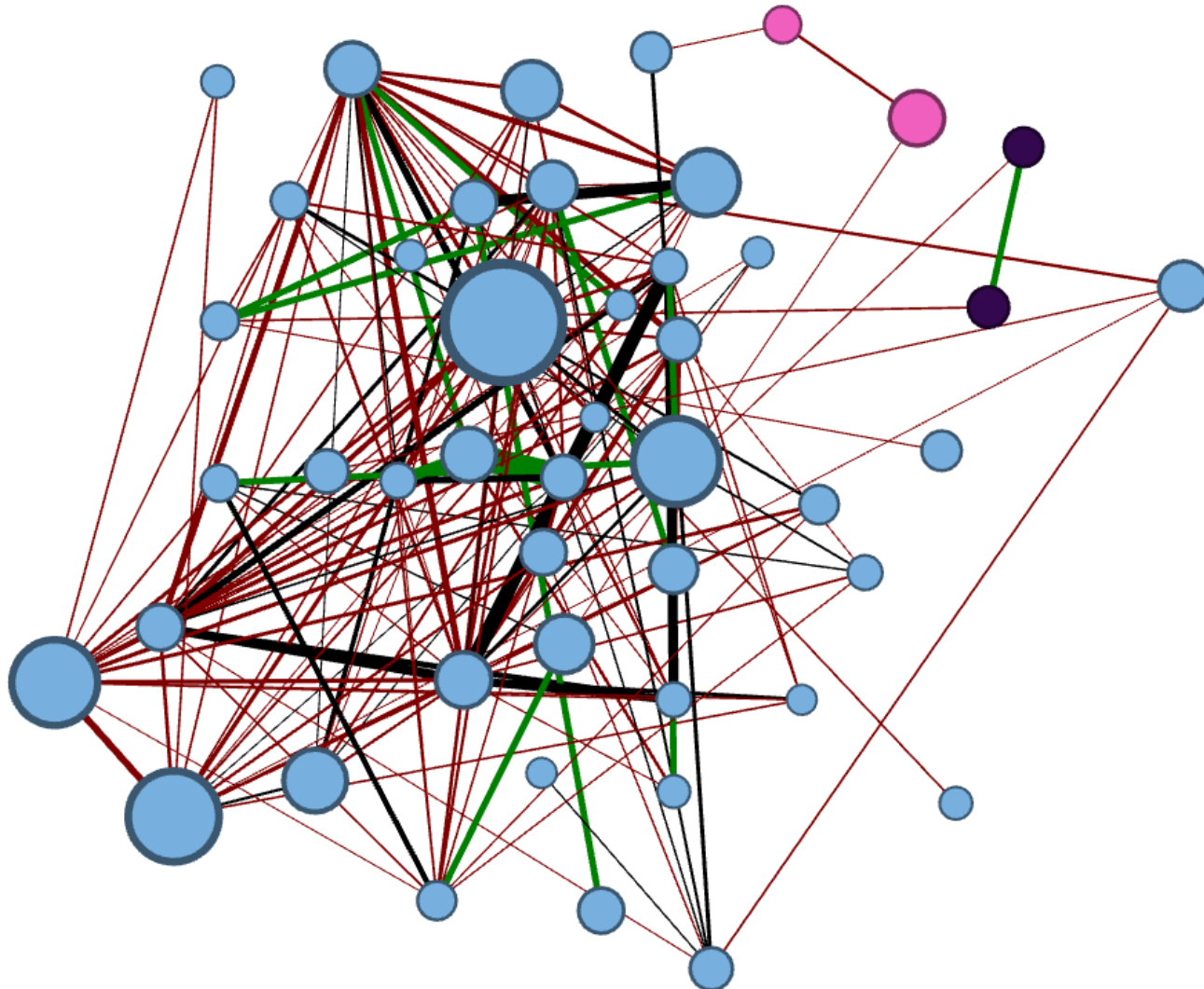
# Articles networking

## Betweenness



# Expertise networking

Nodes are experts





**Sample of publications (nearly 150 in the last 10 years)**





[mmiranda@cgee.org.br](mailto:mmiranda@cgee.org.br)

[www.cgee.org.br](http://www.cgee.org.br)