



# Perspectivas de investigación

## Análise de redes sociais como ferramenta para a gestão da pesquisa e desenvolvimento (P&D) na Embrapa

**Daniela Matias de Carvalho Bittencourt**

Brazilian Agricultural Research Corporation (Embrapa) Research & Development Department  
Brasil · daniela.bittencourt@embrapa.br

**Marcelo Moreira Campos**

Brazilian Agricultura Research Corporation (Embrapa)  
Research & Development Department  
Brasil · marcelo.campos@embrapa.br

**Kepler Euclides Filho**

Brazilian Agricultural Research Corporation  
Research & Development Department  
Brasil · kepler.filho@embrapa.br

**Celso Moretti**

Brazilian Agricultural Research Corporation  
Research & Development Department  
Brasil · celso.moretti@embrapa.br

**Resumo:** Desde a sua criação em 1973, a Empresa Brasileira de Pesquisa Agropecuária (Embrapa) tem direcionado sua programação de pesquisa no sentido de desenvolver conhecimentos e tecnologias visando à estruturação de uma agricultura sustentável nas suas três dimensões, quais sejam, social, ambiental e econômica. Atualmente, o modelo de gestão da pesquisa e desenvolvimento institucional (P&D) da Embrapa conta com apoio de instrumentos gerenciais denominados Portfólios e Arranjos que permitem organizar os projetos de P&D em temas estratégicos, alinhando-os de forma convergente, complementar e sinérgica. Assim, no intuito de desenvolver análises estratégicas que auxiliem a gestão da carteira de projetos dos Portfólios, este trabalho teve por objetivo investigar a viabilidade de utilização das métricas de análise de redes sociais (ARS) como ferramenta para caracterizar os perfis de relacionamento e interação entre os diferentes Portfólios da Embrapa. Com o emprego do software livre *Gephi* foi gerado um grafo da rede formada pelo compartilhamento de 500 projetos de P&D componentes de 18 Portfólios. A partir dos padrões estruturais das interações presentes na rede, foi possível observar que aqueles Portfólios relacionados a temas mais abrangentes tendem a ser mais centrais na rede, apresentando um grande número de relações e podendo ser considerados grandes influenciadores. Os resultados também evidenciaram que os Portfólios tendem a se agrupar mais intensamente com aqueles que trabalham com temas complementares aos seus. Além disso, a partir da análise de grupos de Portfólios mais coesos entre si na rede, foi possível verificar aquele que possui maior autoridade, servindo de articulador em seu grupo. Portanto, pode-se concluir que a ARS se mostrou uma ferramenta útil para a gestão de P&D na Embrapa.

**Palavras-chave:** Análise de Redes Sociais; Gestão de Pesquisa e Desenvolvimento; Embrapa.

**Abstract:** Since its establishment in 1973, the Brazilian Agricultural Research Corporation (Embrapa) has focused its efforts on developing knowledge and technologies to structure a sustainable agriculture in its three dimensions, namely, social, environmental and economic. Currently, the research and institutional development (R&D) management model has the

support of management tools called Portfolios (top-down and bottom-up), structured in order to organize converging, complementary and synergistic R&D projects to meet the priority challenges in a given subject. Thus, in order to develop strategic analysis to assist the management of Embrapa's Portfolios, this study aimed to investigate the feasibility of using the metrics of social network analysis (SNA) as a tool to characterize the relationship and interaction profiles between the different Portfolios of R&D projects at Embrapa. Through the free software *Gephi*, it was generated a graph of the network formed by 18 Portfolios sharing 500 R&D projects. From the structural patterns of interactions in the network, it was observed that those Portfolios related to broader research themes tend to be more central in the network, with a large number of relations, being considered great influencers. The results also showed that the Portfolios tend to cluster more strongly with those who develop research with complementary themes. In addition, from the analysis of more cohesive Portfolios groups, we found the one that has greater authority, serving as articulator in their group. Therefore, the SNA proved to be an effective support tool for R&D management at Embrapa.

**Keywords:** Social Network Analysis; Research and Development Management; Embrapa.

## Introdução<sup>1</sup>

A Empresa Brasileira de Pesquisa Agropecuária (Embrapa) é uma empresa de inovação tecnológica focada na geração de conhecimento e tecnologia para a agropecuária brasileira (Empresa Brasileira de Pesquisa Agropecuária [Embrapa], & Ministério da Agricultura, Pecuária e Abastecimento [MAPA], n.d.a). Desde sua criação, em 1973, a Embrapa tem se mostrado grande impulsora da modernização do setor agrário brasileiro e, muito provavelmente, continuará a desempenhar papel importante na identificação de novas tendências para a inovação tecnológica nos próximos anos. A fim de lidar com os desafios e transformações enfrentadas ao longo do tempo, a Embrapa tem continuamente implementado novos modelos de gestão que procuram melhorar, cada vez mais, a eficiência da sua programação de pesquisa. Entretanto, os desafios para a pesquisa agropecuária são cada vez mais complexos e exigem ações integradas e transdisciplinares. Assim, na versão mais recente do Plano Estratégico da Embrapa, a gestão da pesquisa e desenvolvimento institucional (P&D) está focada em duas grandes vertentes, os Portfólios e os Arranjos. Estes instrumentos de apoio gerencial foram estruturados para organizar projetos afins segundo uma visão temática, tendo como principal objetivo direcionar, promover e acompanhar a obtenção de resultados da pesquisa. Os Portfólios (*top-down*) englobam temas corporativos, considerando-se os objetivos estratégicos da empresa. Os Arranjos (*bottom-up*), tem menor escala e são propostos, preferencialmente, com visão de uma Unidade ou de conjunto de Unidades de pesquisa da Embrapa, para atuarem na solução de problemas regionais e/ou de cadeias produtivas (Penteado, Fontes, Campos, & Euclides, 2014).

Além de possibilitarem melhor organização da programação, Portfólios e Arranjos também têm a função de criar mecanismos para a estruturação do acompanhamento e da avaliação das atividades de P&D da empresa. Estas duas instâncias, por meio dos comitês gestores de Portfólios (CGPorts) e dos grupos gestores de Arranjos (GGAr), atuam em pelo menos duas dimensões na governança da programação da empresa: estratégica e tática. Os CGPorts e GGAr também possuem a atribuição de estabelecer critérios que serão usados por eles para a escolha dos projetos existentes na Embrapa que deverão compor cada Portfólio e Arranjo. Neste contexto, é importante que tais critérios sejam estabelecidos com base no escopo e no mapa de oportunidades/prioridades do respectivo Portfólio ou Arranjo. Atualmente, existem 782 projetos de P&D em execução na programação de pesquisa da Embrapa alojados em 23 Portfólios e 75 Arranjos institucionais (dados de julho de 2015). É importante salientar, que de acordo com seus objetivos, um mesmo projeto pode estar conectado a mais de um Portfólio, Arranjo ou mesmo em ambos. Como apoio à atuação dos CGPorts e dos GGAr, o Departamento de Pesquisa e Desenvolvimento (DPD) da Embrapa é responsável por fornecer informações relativas ao andamento das

<sup>1</sup> A participação neste evento foi financiada pela Fundação de Apoio a Pesquisa do Distrito Federal (FAP - DF).

atividades e dos projetos de P&D que os compõem e atuar como facilitador de articulação que possibilite a interação entre e intra membros dos Arranjos e Portfólios, no intuito de facilitar a troca de informações e experiências, bem como criar ambiente de motivação para a apresentação de novas propostas. Mediante este novo contexto, o DPD tem buscado desenvolver mecanismos para que essas informações sejam repassadas de forma eficiente e que realmente forneçam subsídios para que os CGPorts e o GGArS promovam avaliação detalhada da sua programação, de modo a assegurar melhor foco, maior abrangência e complementaridade dos projetos em consonância com o escopo e objetivos do Portfólio ou Arranjo.

Uma vez que, com a instituição dos Portfólios e Arranjos, a programação de P&D da Embrapa está estruturada a partir de conjuntos de projetos conectados e executados com a participação de diversas unidades de pesquisa da empresa operando em redes colaborativas de P&D, a metodologia de análises de redes sociais (ARS) se destaca como uma possível ferramenta para o desenvolvimento de análises estratégicas da relação formada entre estes dois instrumentos.

“Uma rede social é definida metodologicamente como um conjunto de relações específicas (por exemplo, colaboração, apoio, aconselhamento, controle ou ainda influência) entre um conjunto finito de atores (indivíduos ou organizações)” (Lazega & Higgins, 2014, pp. 7-8). No caso dos Portfólios e Arranjos esta colaboração é caracterizada pelo compartilhamento de projetos entre eles. Por sua vez, uma rede pode ser representada matematicamente por um objeto chamado *grafo*, utilizado para descrever sua estrutura topológica. Um grafo é constituído por um conjunto de pontos, denominados *nós* ou *vértices*, neste caso representando os Portfólios e/ou Arranjos, conectados por linhas que expressam a relação entre eles, denominadas *arestas* (Freitas, 2010).

Por sua característica, a **ARS** está se transformando em um poderoso instrumento de gestão do conhecimento, para avaliar o desempenho organizacional, uma vez que ela permite esquematizar as relações entre atores (**nos níveis individual, grupal e organizacional**), de maneira a determinar os fluxos de informações e de conhecimentos, mostrando as verdadeiras redes que se desenvolvem embaixo da estrutura formal da organização (Santos, 2014). Portanto, no intuito de visualizar e compreender as numerosas relações que podem facilitar ou impedir a criação e o compartilhamento de conhecimentos entre os Portfólios, este trabalho tem por objetivo investigar a viabilidade da metodologia de ARS, como ferramenta para caracterizar os perfis de relacionamento e interação entre os diferentes Portfólios de projetos de P&D na Embrapa.

### Metodologia

Pela característica estratégica e relevância nacional dos Portfólios em relação aos Arranjos, esta análise focou somente na relação entre os Portfólios da Embrapa. Este estudo baseou-se nas abordagens de pesquisa quantitativa e qualitativa, com o uso da metodologia de ARS.

Para a obtenção dos dados utilizou-se a ferramenta de *Business Intelligence* (BI) da Embrapa. Cabe aqui ressaltar que, mesmo não sendo o foco deste trabalho, considera-se pertinente contextualizar que o conceito de BI pode ser entendido, de forma geral, “como a utilização de variadas fontes de informação para se definir estratégias de competitividade nos negócios da empresa” (Barbieri, 2001, p. 34). Esta ferramenta utiliza basicamente o modelo Booleano de recuperação de informação e pode apresentar, dentre outras formas, o resultado em forma de relatórios (Documento *Web Intelligence*). Baseado na álgebra Booleana e na teoria de conjuntos, o modelo Booleano de recuperação da informação considera basicamente se os termos de indexação ou expressões estão presentes ou ausentes no documento analisado (Baeza-Yates & Ribeiro-Neto, 2013, pp. 29 – 31), obtendo como resposta um resultado verdadeiro ou falso, representando respectivamente a existência ou não do termo em relação à busca efetivada (Silva, 2013, p. 58).

Assim, tendo como parâmetro de consulta os Portfólios e os projetos de P&D existentes na Embrapa inseridos em sua carteira, obteve-se um relatório que foi

transformado em uma matriz de incidência (Projetos versus Portfólios), que posteriormente foi transformada em uma matriz de sobreposição simétrica. Para as análises deste estudo foi utilizado o software *Gephi* (Bastian, Heymann & Jacomy; 2009), uma plataforma livre que permite a geração e visualização de grafos dinâmicos que podem auxiliar no entendimento do formato de redes, por meio do cálculo automático e interpretação de métricas. Para o layout do grafo utilizou-se o algoritmo de distribuição *Force Atlas 2*, com a opção *evitar sobreposição* ligada e o dimensionamento ajustado para o valor 500 (Jacomy, Venturini, Heymann, & Bastian, 2014). Utilizou-se também as métricas de centralidade (grau, diâmetro da rede, centralidade de autovetor e *PageRank*), densidade, modularidade e coeficiente de agrupamento (Brin & Page, 1998; Brandes, 2001; Blondel, Guillaume, Lambiotte, & Lefebvre, 2008). Também foi calculado o coeficiente de agrupamento ponderado, mediante a instalação de um *plugin* (Barrat, Barthélemy, Pastor-Soarras, & Vespignani, 2004). Os dados foram normalizados para as métricas de centralidade e a opção *peso das arestas* foi selecionada para os cálculos de modularidade e *PageRank*, sendo que para modularidade também se ajustou a resolução para 0,7 (varia de 0 a 1, quanto maior este valor maior será o número de nós inseridos em determinada comunidade) (Lambiotte, Delvenne, & Barahona, 2009). Para todas as métricas calculadas, quando disponível, foi selecionada a opção grafo *não direcionado*. As demais opções disponíveis para cálculo das métricas foram mantidas no formato padrão do programa.

## Resultados e Discussão

Neste trabalho foi analisada a estrutura da rede formada pelo compartilhamento de projetos de P&D entre 18 Portfólios, que englobam 500 projetos em execução na Embrapa. Como o foco principal deste trabalho é conhecer de que forma ocorre a interação entre os diferentes Portfólios, o grafo não direcionado representando a rede formada pela interação destes foi construído a partir de uma matriz de sobreposição simétrica (Tabela 1), onde estão indicados o número de projetos compartilhados entre dois Portfólios e o número total de projetos inseridos em cada um deles (valores indicados na diagonal da matriz). A Figura 1 representa o grafo da rede formada pelo compartilhamento de projetos entre estes Portfólios, a partir da matriz.

Visualmente já é possível identificar aqueles Portfólios que interagem entre si e a intensidade desta interação representada pela espessura da aresta. Os nós representados pela mesma cor indicam grupos de Portfólios fortemente conectados (comunidades). Entretanto, no intuito de melhor compreender o padrão de relacionamento presente na rede representada pelo grafo, se faz necessário uma análise criteriosa levando em consideração os valores de diferentes métricas estruturais calculadas a partir das interações presentes na rede. Claramente, a informação estrutural de maneira isolada tem pouca relevância, mas indica quais pontos merecem análises mais profundas. Assim, os resultados apresentados e discutidos a seguir, estão de acordo com a propriedade destas métricas em relação às características dos Portfólios.

### 1 – Centralidade

Ao se analisar uma rede é crucial identificar os nós ou conexões estruturalmente mais importantes. Existem diferentes maneiras de um nó ser importante na rede e as medidas de centralidade são uma forma de quantificar esta importância (Freeman, 1979; Borgatti, 2005). Estas medidas levam em consideração as diferentes maneiras em que um ator interage e se comunica com o restante da rede, sendo mais importantes ou centrais aqueles vértices localizados em posições mais estratégicas, em função de algumas invariantes do grafo (Freitas, 2010). Por exemplo, um nó pode ser considerado importante porque sua remoção tenderia a desconectar a rede, ou ainda, um nó pode ser importante por estar envolvido em grande número de ligações da rede (Borgatti, Everett & Johnson, 2013).

As métricas de centralidade utilizadas na análise da rede formada entre os Portfólios foram centralidade de grau (*degree centrality*), grau médio ponderado

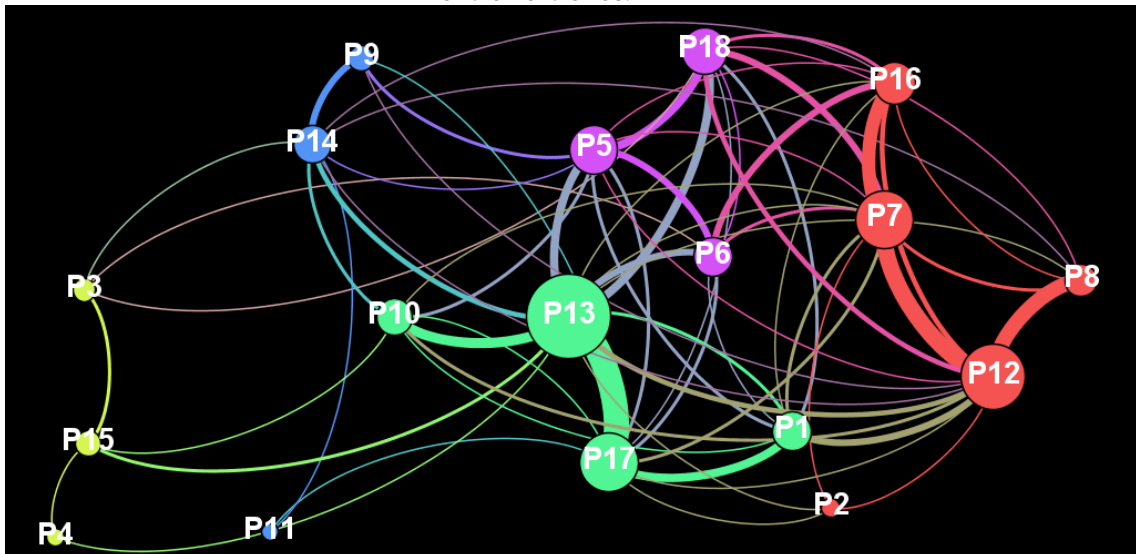
(*weighted degree*), centralidade de proximidade (*closeness*), centralidade de intermediação (*betweenness*), excentricidade, centralidade de autovetor (*Eigenvector Centrality*) (Bonacichi & Loyd, 2001) e, uma variante desta última, o *PageRank* (Brin & Paje, 1998). A Tabela 2 mostra o valor de cada uma destas métricas encontrado para cada Portfólio analisado. A centralidade de grau mede o número de interações que os nós de uma rede possuem. No caso da rede formada pelos Portfólios o grau médio encontrado foi de 7,6, ou seja, cada Portfólio compartilha projetos, em média, com outros sete Portfólios. Considerando-se o número total de Portfólios, este resultado aponta para a existência de elevado nível de interação na rede. Dos 18 Portfólios analisados, o Portfólio "Química e Tecnologia da Biomassa" foi o que apresentou maior centralidade de grau, possuindo interação com outros 14 Portfólios, seguido pelos Portfólios "Pastagens", com centralidade de grau 12, e o "Sucroalcooleiro Energético", que juntamente com os Portfólios "Fixação Biológica de Nitrogênio" e "Automação", apresentaram centralidade de grau 11. Entretanto, os Portfólios "Palma de Óleo" e "Aquicultura" foram os que apresentaram menor interação, compartilhando projetos com apenas outros 2 Portfólios cada.

**Tabela 1.** Matriz de sobreposição simétrica entre 18 Portfólios de projetos de P&D da Embrapa.

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16	P17	P18
P1	35	0	0	0	2	1	2	0	0	1	0	4	2	0	0	1	5	2
P2	0	26	0	0	0	0	1	0	0	0	0	1	1	0	0	0	1	0
P3	0	0	20	0	0	1	0	0	0	0	0	0	0	1	2	0	0	1
P4	0	0	0	30	0	0	0	0	0	0	0	0	1	0	1	0	0	0
P5	2	0	0	0	22	4	1	0	2	2	0	1	5	1	0	1	2	5
P6	1	0	1	0	4	46	2	0	0	0	0	0	4	0	0	4	2	1
P7	2	1	0	0	1	2	41	2	0	1	0	10	1	0	0	8	2	4
P8	0	0	0	0	0	0	2	20	0	0	0	9	1	1	0	1	0	1
P9	0	0	0	0	2	0	0	0	20	0	0	1	0	4	0	0	1	0
P10	1	0	0	0	2	0	1	0	0	40	0	2	6	2	1	0	1	0
P11	0	0	0	0	0	0	0	0	0	0	8	0	0	1	0	0	1	0
P12	4	1	0	0	1	0	10	9	1	2	0	77	3	1	0	3	1	3
P13	2	1	0	1	5	4	1	1	0	6	0	3	65	3	2	1	15	5
P14	0	0	1	0	1	0	0	1	4	2	1	1	3	44	0	1	0	0
P15	0	0	2	1	0	0	0	0	0	1	0	0	2	0	25	0	0	0
P16	1	0	0	0	1	4	8	1	0	0	0	3	1	1	0	45	0	2
P17	5	1	0	0	2	2	2	0	1	1	1	1	15	0	0	0	36	1
P18	2	0	1	0	5	1	4	1	0	0	0	3	5	0	0	2	1	34

Valores na diagonal representam o número de projetos inseridos em cada Portfólio, enquanto os demais números representam o número de projetos compartilhados entre dois Portfólios quaisquer. Os Portfólios estão identificados por P1 a P18, sendo: (P1) Agricultura Irrigada, (P2) Alimentos, Nutrição e Saúde, (P3) Alimentos Seguros, (P4) Aquicultura, (P5) Automação, (P6) Controle Biológico, (P7) Fixação Biológica de Nitrogênio, (P8) Integração Lavoura, Pecuária e Florestas – ILPF, (P9) Monitoramento da Dinâmica do Uso e Cobertura da Terra no Território Nacional, (P10) Mudanças Climáticas, (P11) Palma de Óleo, (P12) Pastagens, (P13) Química e Tecnologia da Biomassa, (P14) Recursos Florestais nativos, (P15) Sanidade Animal, (P16) Sistemas de Produção de Base Ecológica, (P17) Sucroalcooleiro Energético e (P18) Suprimentos de Nutrientes para a Agricultura.

**Figura 1.** Grafo não direcionado da rede formada pelo compartilhamento de projetos de P&D entre Portfólios.



Os nós representam os Portfólios, quanto maior seu diâmetro maior sua centralidade na rede, de acordo com a métrica *Page Rank*. Nós da mesma cor pertencem a mesma comunidade. Portfólios que compartilham projetos entre si estão ligados por uma aresta, sendo que sua espessura indica a intensidade da relação entre Portfólios. Os Portfólios estão identificados por P1 a P18, sendo: (P1) Agricultura Irrigada, (P2) Alimentos, Nutrição e Saúde, (P3) Alimentos Seguros, (P4) Aquicultura, (P5) Automação, (P6) Controle Biológico, (P7) Fixação Biológica de Nitrogênio, (P8) Integração Lavoura, Pecuária e Florestas – ILPF, (P9) Monitoramento da Dinâmica do Uso e Cobertura da Terra no Território Nacional, (P10) Mudanças Climáticas, (P11) Palma de Óleo, (P12) Pastagens, (P13) Química e Tecnologia da Biomassa, (P14) Recursos Florestais nativos, (P15) Sanidade Animal, (P16) Sistemas de Produção de Base Ecológica, (P17) Sucroalcooleiro Energético e (P18) Suprimentos de Nutrientes para a Agricultura.

A partir das conexões existentes na rede, também foi possível calcular a excentricidade, centralidade de intermediação, centralidade de proximidade e centralidade de autovetor para cada nó. A excentricidade mede a distância de um determinado nó inicial até o nó mais distante na rede. Entende-se por distância o número de conexões (arestas) intermediárias existentes entres esses nós. A distância existente entre os Portfólios variou entre duas e três conexões, sendo que a maior distância existente entre dois nós caracteriza o diâmetro da rede. "A centralidade de intermediação é baseada na ideia do controle exercido pelo ator sobre as interações entre dois outros atores presentes na rede. Quanto mais um ator se encontrar no "meio", como ponto de passagem obrigatório, mais central ele será" (Lazega & Higgins, 2014, p. 44). De maneira geral, os valores encontrados para esta métrica foram relativamente baixos. Para melhor compreensão dos dados, a excentricidade e a centralidade de intermediação serão discutidas mais profundamente no próximo tópico.

Por sua vez, a centralidade de proximidade pode ser compreendida como o inverso da distância média entre um determinado nó e todos os outros. Seus valores podem variar entre 0 a 1, sendo 1 encontrado para nós que possuem conexão com todos os demais atores da rede. Analisando os valores de centralidade de proximidade obtidos para cada Portfólio, verificou-se que o Portfólio "Química e Tecnologia de Biomassa" foi o mais central, seguido pelo "Pastagens" e pelo "Fixação Biológica de Nitrogênio" (Tabela 2).

A centralidade de autovetor, inicialmente proposta por Bonacich (1972, 1987), é definida de maneira circular, onde a centralidade de um nó é calculada como a soma da centralidade de proximidade de seus nós vizinhos, ou seja, essa medida considera que nós conectados com nós de maior centralidade receberão também maior centralidade. Neste contexto, a importância (influência) de um nó é proporcional à importância (influência) de seus vizinhos (Guedes, 2013). Esta métrica, normalmente, está diretamente relacionada com a centralidade de grau e a centralidade de

proximidade, o que não foi diferente no caso da rede formada pelos Portfólios. Verificou-se, uma vez mais, que o Portfólio “Química e Tecnologia de Biomassa” foi apontado como o mais influente da rede e os Portfólios “Palma de Óleo” e “Aquicultura” foram os que apresentaram menor influência sobre os demais, respectivamente (Tabela 2).

A maior centralidade encontrada para o Portfólio “Química e Tecnologia de Biomassa” evidencia sua maior interação com os demais Portfólios da rede. Esta alta interatividade do Portfólio “Química e Tecnologia de Biomassa” se dá, muito provavelmente, pela maior abrangência do tema abordado pelo mesmo, possibilitando grande transversalidade com os demais. Ao se ler o objetivo deste Portfólio, pode-se verificar exatamente isso:

Tem como objetivo promover e executar ações integradas de P, D & I para rotas tecnológicas sustentáveis, visando à valorização da biomassa nacional por meio das transformações químicas destes recursos, para o fortalecimento sustentável de cadeias do agronegócio brasileiro, com base nos conceitos de biorrefinaria e de bioeconomia. O portfólio aborda quatro vertentes principais – obtenção de novos materiais, produtos químicos e geração de energia renovável, além do desenvolvimento de tópicos comuns a estes assuntos, denominados temas transversais (Embrapa & MAPA, n.d.b);

o que também justifica sua elevada interação com os Portfólios “Sucroalcooleiro Energético”, “Suprimentos de Nutrientes para a Agricultura” e “Automação”. Quando verificamos o segundo Portfólio mais central na rede, o Portfólio “Pastagens”, também se pode verificar o mesmo efeito, uma vez que este Portfólio foca suas pesquisas em temas altamente transversais a vários Portfólios, como o “Fixação Biológica de Nitrogênio” e o “Integração Lavoura, Pecuária e Florestas – ILPF”. Este comportamento ainda é mais evidenciado quando se observa os dois Portfólios com menor centralidade, o “Palma de Óleo” e o “Aquicultura”. Ambos tratam de temas mais específicos, de menor abrangência, e conseqüentemente, apresentam menor nível de interação com os demais Portfólios.

Embora o Portfólio “Química e Tecnologia de Biomassa” tenha sido considerado o mais “influente” na rede, esta influência é limitada, uma vez que um determinado projeto inserido num Portfólio A, por exemplo, também faz parte de um Portfólio B ou ainda de um C, apenas o resultado deste projeto influenciará estes Portfólios. Assim, a intensidade de sua influência, muito provavelmente, está simultaneamente vinculada ao número (peso) de projetos compartilhados com outros Portfólios e não somente ao número de conexões e à distância deste Portfólio em relação aos demais na rede. Quanto maior o número de projetos compartilhados entre dois Portfólios, maior a intensidade (peso) da interação entre eles e, conseqüentemente, maior a sua influência.

Para se testar esta hipótese foram calculadas outras duas métricas de centralidade que levam em consideração justamente esta variável: o grau médio ponderado e o *PageRank*. O grau médio ponderado mede exatamente este nível de interação. O valor encontrado varia de acordo com o número de projetos compartilhados por dois Portfólios, o que também reflete o peso da aresta. O valor calculado para a rede como um todo foi de 18,9, ou seja, cada portfólio compartilha, em média, 19 projetos com os demais Portfólios da rede. Isto pode ser evidenciado na matriz de sobreposição (Tabela 1), onde é possível verificar o número de projetos compartilhados entre dois Portfólios quaisquer apenas cruzando os valores das linhas com as colunas. Observando-se o valor do grau ponderado de cada Portfólio na Tabela 2, pode-se concluir que o Portfólio “Química e Tecnologia da Biomassa” é o que possui maior nível de interação entre os Portfólios, bem acima da média, com 50 compartilhamentos (não necessariamente 50 projetos compartilhados, uma vez que o mesmo projeto de P&D pode ser compartilhado com mais de um Portfólio), e que os Portfólios “Palma de Óleo” e “Aquicultura” foram os que apresentaram menor nível de interação com os demais, compartilhando apenas dois projetos cada.

**Tabela 2.** Dados de centralidade dos Portfólios de acordo com a estrutura do Grafo.

Nós	Nome do Portfólio	Centralidade de Grau	Grau Médio Ponderado	Excentricidade	Centralidade de Intermediação	Centralidade de Proximidade	Centralidade de Autovetor	PageRank
P13	Química e Tecnologia de Biomassa	14	50	2	0,198	0,850	1,000	0,138
P12	Pastagens	12	39	2	0,057	0,773	0,931	0,102
P17	Sucroalcooleiro Energético	11	32	2	0,099	0,739	0,824	0,089
P7	Fixação Biológica de Nitrogênio	11	34	2	0,030	0,739	0,902	0,088
P5	Automação	11	26	2	0,034	0,739	0,905	0,072
P18	Suprimentos de Nutrientes para a Agricultura	10	25	2	0,037	0,708	0,824	0,068
P16	Sistemas de Produção de Base Ecológica	9	22	2	0,011	0,680	0,777	0,059
P1	Agricultura Irrigada	9	20	2	0,006	0,680	0,810	0,055
P6	Controle Biológico	8	19	2	0,021	0,654	0,680	0,054
P14	Recursos Florestais Nativos	9	15	2	0,094	0,680	0,608	0,052
P10	Mudanças Climáticas	8	16	2	0,031	0,654	0,672	0,049
P8	Integração Lavoura, Pecuária e Florestas - ILPF	6	15	2	0,002	0,607	0,544	0,043
P15	Sanidade Animal	4	6	3	0,017	0,531	0,227	0,030
P9	Monitoramento da Dinâmica do Uso e Cobertura da Terra no Território Nacional	4	8	3	0,001	0,531	0,354	0,029
P3	Alimentos Seguros	4	5	3	0,015	0,548	0,255	0,025
P2	Alimentos, Nutrição e Saúde	4	4	3	0,000	0,548	0,395	0,017
P4	Aquicultura	2	2	3	0,000	0,500	0,135	0,015
P11	Palma de Óleo	2	2	3	0,001	0,500	0,156	0,014

A tabela está ordenada a partir do Portfólio mais central na rede, de acordo com os valores de *PageRank*.



A última medida de centralidade a ser analisada é a *PageRank*, esta é uma variação da centralidade de autovetor criada por fundadores do Google para classificar a relevância de páginas da Web, que posteriormente foi implementada na máquina de busca do Google. Este algoritmo está disponível para cálculo no *Gephi*, onde é possível ponderar utilizando o peso das arestas, o que permite uma visão mais real do grau de centralidade dos nós na rede, além de indicar se esta "intensidade" da interação interfere, em maior ou menor grau, nas relações entre os Portfólios. Isto quer dizer que uma conexão de elevado peso com um ator de baixa centralidade pode produzir resultado similar ao de uma conexão de baixo peso com um ator altamente central na rede (Borgatti et al., 2013). O resultado obtido é bastante similar ao da centralidade de autovetor, com apenas algumas modificações nas posições de centralidade como o Portfólio "Sucroalcooleiro Energético" que, de acordo com os valores obtidos pelo *PageRank*, pode ser considerado o terceiro Portfólio mais central na rede, sendo que o segundo foi o "Pastagens". Além disso, foi também possível verificar que, baseado na força de suas conexões, o Portfólio "Aquicultura" foi o menos central em relação aos demais Portfólios na rede.

Ao se comparar os valores de centralidade de autovetor e de *PageRank* dos três primeiros colocados, pode-se observar que a ascensão de posicionamento do Portfólio "Sucroalcooleiro Energético" pela métrica calculada pelo *PageRank* foi resultado do maior número de projetos compartilhados com o Portfólio "Química e Tecnologia da Biomassa" em relação ao número de projetos compartilhados entre este último e o Portfólio "Automação", que foi inserido como o segundo mais central pela centralidade de autovetor. O Portfólio "Sucroalcooleiro Energético" compartilha 15 projetos com o Portfólio "Química e tecnologia da biomassa", entretanto, o Portfólio "Automação" compartilha apenas 5 (Tabela 1). Estes resultados, mostram que, pelo menos no caso dos Portfólios, a intensidade da relação interferiu na centralidade dos Portfólios na rede como um todo. Entretanto, esta interferência não foi maior, pois como os Portfólios mais centrais desenvolvem pesquisa sobre temas mais abrangentes, estes também interagem com elevado número de Portfólios.

## 2 – Densidade e Coeficiente de Agrupamento

Dentre as métricas disponíveis para cálculo no *Gephi*, existem aquelas que fornecem visão geral da rede, como as medidas de densidade, coeficiente de agrupamento e algumas de centralidade discutidas acima (Borgatti et al., 2013). A densidade do grafo, como o próprio nome diz, mede quão densa ou coesa é a rede formada pelos seus atores. Esta é uma medida que varia de 0 a 1 e indica o nível de conexão presente na rede expressa como uma proporção do número possível, sendo que valores maiores indicam uma maior densidade ou coesão do grafo. Na rede formada pelos 18 Portfólios foi possível observar um total de 69 conexões, o que resultou em uma densidade de 0,451 (Tabela 3). Isto significa que 45% dos Portfólios compartilham, ao menos, um projeto entre si. Levando em consideração a natureza dos Portfólios e suas relações, pode-se pensar que esta é uma rede densa, com elevado nível de conectividade. Embora esta conclusão possa parecer subjetiva, existem outras métricas que podem dar maiores indícios desta conectividade, como o diâmetro da rede, e a excentricidade e a centralidade de intermediação dos nós (Tabela 2).

**Tabela 3.** Dados estruturais da rede formada pelos Portfólios.

Visão Geral da Rede	
Densidade	0,451
Diâmetro	3
C <sup>1</sup>	0,632
Cp <sup>2</sup>	0,665

<sup>1</sup> Coeficiente de Agrupamento

<sup>2</sup> Coeficiente de Agrupamento Ponderado

Conforme discutido na seção acima, a excentricidade encontrada para os nós variou entre 2 e 3, sendo 3 o valor do diâmetro da rede. Como a rede é formada por 18 Portfólios interagindo entre si, uma distância de 3 conexões é relativamente curta a separar uma extremidade da outra na rede, mostrando um elevado nível de compactação entre seus atores. Os valores encontrados para a centralidade de intermediação dos nós analisados foram, da mesma forma, relativamente baixos, variando entre 0 e 0,198 (os valores possíveis, neste caso, estão entre 0 e 1) (Tabela 2). Atores com elevada centralidade de intermediação normalmente são figuras importantes, pois podem controlar o fluxo de informação entre dois grupos e, uma vez retirados da rede, podem causar o completo isolamento entre grupos. Nestes casos, o ator pode ser chamado de ponte ou *gatekeeper*. Como os valores encontrados para os Portfólios são relativamente baixos, isto indica que dificilmente a rede se desconectará pela retirada de um determinado nó. Tal constatação é indício, novamente, da elevada interação entre eles.

O coeficiente de agrupamento (*clustering coefficient*,  $C$ ) também fornece uma indicação do grau de coesão existente em determinada rede (Barrat et al., 2004). Este coeficiente é baseado no conceito de triades, onde três nós se encontram conectados por dois (triade aberta) ou por três (triade fechada) laços não direcionados, e é calculado a partir do número de triades fechadas sobre o número total de triades (abertas e fechadas). Este coeficiente é calculado para cada um dos nós presentes na rede, e seu valor ( $0 \leq C \leq 1$ ) representará a densidade com que esse nó se relaciona com os seus vizinhos. A média dos valores de  $C$  para todos os nós é o coeficiente de agrupamento global da rede. O valor de  $C$  encontrado para a rede formada entre os Portfólios foi de 0,632, o que significa que 63% da rede é formada por triades fechadas e, conseqüentemente, apresentam elevado nível de conexão entre seus nós. O valor de  $C$  encontrado para cada nó será discutido juntamente com a modularidade, uma vez que esta característica é determinada em função dos nós vizinhos. Entretanto,  $C$  foi calculado levando-se em consideração apenas as conexões presentes entre os nós da rede, e não a intensidade destas interações. Assim, no intuito de novamente verificar a influência da força das interações na estrutura global da rede foi calculado o valor ponderado de  $C$  ( $C_p$ ). O valor global de  $C_p$  encontrado para a rede foi de 0,665, resultando em  $C_p > C$ . De acordo com Barrat et al. (2004) quando  $C_p > C$  tem-se uma rede onde há maior tendência de formação de triades fechadas entre nós com maior intensidade de interação. Por outro lado, quando  $C_p < C$ , o agrupamento local é gerado entre nós com fraca interação. Neste caso, o grau de articulação na rede é baixo, uma vez que a maior parte das interações está ocorrendo em triades fechadas. Assim, ambos os resultados,  $C$  e  $C_p$ , confirmam o elevado grau de conectividade entre os Portfólios. Entretanto, a intensidade das relações pouco influenciou a estrutura global da rede, o que também pode estar relacionado com o elevado número de interações existentes entre os Portfólios.

Os resultados encontrados sugerem alta complementaridade entre os temas estudados pelos Portfólios. Entretanto, também se pode concluir o oposto, ou seja, a alta densidade refletiria elevada sobreposição dos temas trabalhados por eles, indicando convergência de objetivos ou duplicação de esforços. No entanto, como mencionado anteriormente, quando se analisa as interações entre portfólios com base somente nos projetos que eles compartilham, não se consideram especificamente os resultados potenciais esperados em cada projeto. Assim, projetos podem ser alocados a Portfólios diferentes em função de resultados diferentes. Isto indica a necessidade de se aprofundarem as análises de modo a permitir verificar o efeito do compartilhamento dos resultados na rede. Além disso, cada CGPort define seus critérios para a seleção dos projetos que irão compor sua carteira, assim, não existindo padronização metodológica. Ademais, ao se analisar um projeto compartilhado entre Portfólios, muitas das vezes, é possível observar que o(s) resultado(s) gerado(s) por este projeto não terá(ão) o(s) mesmo(s) impacto(s) sobre todos os Portfólios que os incluíram em sua carteira. Desta forma, uma maneira de diminuir esta incerteza de sobreposição seria ligar um resultado específico, gerado por um determinado projeto, a um tema específico de um Portfólio, assim, esse

resultado estaria realmente contribuindo de maneira efetiva para o alcance dos objetivos daquele Portfólio. Portanto, em vez de projetos inteiros, teríamos apenas resultados específicos sendo, ou não, compartilhados entre os Portfólios.

### 3 – Modularidade

Para melhor compreender o porquê desta elevada conectividade da rede formada pelos Portfólios, também se faz necessário identificar quais são os subgrupos de nós mais coesos entre si. A modularidade fornece exatamente esta resposta, ela identifica o(s) subgrupo(s) de atores mais coesos, onde se constata a existência de relações fortes e frequentes, também denominado(s) de comunidade(s) (Blondel et al., 2008). No caso da rede formada pelos Portfólios foi possível agrupar seus atores em cinco comunidades coesas. Uma vez identificado os nós pertencentes a cada subgrupo, também se pôde calcular a densidade das conexões entre seus atores e entre comunidades. De maneira geral, os subgrupos apresentaram elevada interconectividade de seus membros, apresentando valores de densidade entre 0,66 a 1. As comunidades mais coesas foram a 0 e a 3, ambas com 100% de interação entre seus membros (Tabela 4). Além disso, é possível observar que arestas mais espessas, indicando maior intensidade da relação entre Portfólios, estão presentes com maior frequência dentro das comunidades do que entre elas (Figura 1). De acordo com Borgatti e colaboradores (2013), atores de um mesmo subgrupo tendem a compartilhar normas e, normalmente, possuem objetivos e ideais comuns. Isto significa que os membros de uma mesma comunidade frequentemente têm resultados semelhantes com relação à adoção de inovação, comportamentos e atitudes. Estas observações também são verdadeiras para a rede formada entre os Portfólios. A comunidade 1, por exemplo, é constituída por Portfólios que tratam de temas complementares e têm por objetivo solucionar problemas de interesse comum para a agricultura brasileira (Embrapa & MAPA, n.d.b). O mesmo fenômeno pode ser evidenciado em todos os outros subgrupos. Esta característica, confirma a hipótese de que as relações entre Portfólios são mais intensas à medida que eles desenvolvem pesquisas relacionadas à temas complementares aos seus. Por sua vez, a Figura 2 representa o grafo formado pelo agrupamento destas comunidades, sendo que a densidade encontrada para a rede formada entre elas foi de 0,9. Este resultado mostra que 90% das comunidades se relacionam umas com as outras, evidenciando duas características da rede formada entre os Portfólios, a forte relação interna das comunidades e estreita relação entre as comunidades.

Outro aspecto importante a ser analisado em cada sub-grupo é o valor de  $C_p$  encontrado para cada nó. Como a análise foi desenvolvida a partir de uma rede com peso em suas conexões, os valores de  $C$  não foram considerados. Ressalta-se que os nós que contêm maior número de conexões, ou elevado grau de centralidade, são os que apresentaram menor valor para  $C_p$ . Assim, nós que formam elevado número de triades fechadas (valor  $> C_p$ ), possuem menor transitividade entre seus vizinhos. Por conseguinte, aqueles que possuem maior número de conexões abertas (valor  $< C_p$ ), possuem elevado grau de transitividade, mantendo relações diretas com um maior número de membros da rede. Pode-se dizer então, que aqueles nós que possuem um baixo valor de  $C_p$  em relação aos demais, são considerados articuladores de seus grupos (Tabela 4). Este tipo de informação é extremamente relevante, pois permite a identificação de Portfólios com elevado nível de atuação dentro de seus grupos, podendo ser considerados mediadores para a articulação de novas propostas de projetos de P&D de interesse comum.

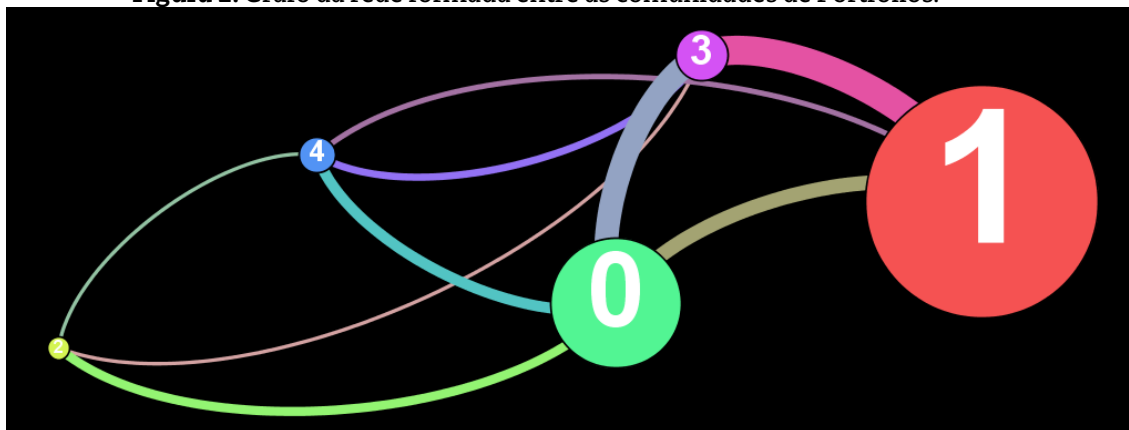
**Tabela 4.** Comunidades da rede formada pelos Portfólios.

Nós	Nome do Portfólio	Modularidade	Cp <sup>1</sup>
P13	Química e Tecnologia de Biomassa	0	0,558
P17	Sucroalcooleiro Energético	0	0,691
P10	Mudanças Climáticas	0	0,786
P1	Agricultura Irrigada	0	0,888
P12	Pastagens	1	0,643
P7	Fixação Biológica de Nitrogênio	1	0,753
P16	Sistemas de Produção de Base Ecológica	1	0,801
P8	Integração Lavoura, Pecuária e Florestas - ILPF	1	0,933
P2	Alimentos, Nutrição e Saúde	1	1,000
P3	Alimentos Seguros	2	0,133
P15	Sanidade Animal	2	0,333
P4	Aquicultura	2	1,000
P5	Automação	3	0,704
P18	Suprimentos de Nutrientes para a Agricultura	3	0,773
P6	Controle Biológico	3	0,782
P11	Palma de Óleo	4	0,000
P14	Recursos Florestais Nativos	4	0,392
P9	Monitoramento da Dinâmica do Uso e Cobertura da Terra no Território Nacional	4	0,792

Os Portfólios estão agrupados de acordo com a comunidade que pertencem e ordenados de acordo com o valor de Cp.

<sup>1</sup>Coefficiente de Agrupamento Ponderado

**Figura 2.** Grafo da rede formada entre as comunidades de Portfólios.



Nós enumerados de 0 a 4 representam cada comunidade de acordo com a Tabela 4. Diâmetro do nó é proporcional à intensidade da relação entre os Portfólios pertencentes à comunidade. Comunidades que compartilham projetos entre si estão ligados por uma aresta, sendo que sua espessura indica a intensidade da relação entre comunidades.

**Conclusões e perspectivas**

A partir do estudo conduzido pôde-se concluir que a metodologia de ARS é um instrumento potencial para a caracterização de padrões estruturais da rede formada entre os Portfólios de projetos de P&D em execução na Embrapa. A partir dos resultados obtidos, também foi evidenciada a necessidade de se analisar a rede formada pelas relações existentes em outro estrato da composição de projetos, qual

seja, de resultados potencias e seus alinhamentos com as linhas de P&D prioritárias dos Portfólios.

Portanto, estas análises também deverão abranger os Arranjos, o que permitirá uma avaliação estratégica da programação de P&D da Embrapa como um todo, contribuindo assim, para o alcance dos seus objetivos e futuras tomadas de decisão em relação ao caminho científico da pesquisa agropecuária perseguido pela empresa.

## Referências

**Baeza-Yates, R. & Ribeiro-Neto, B.** (2013). *Recuperação de informação: conceitos e tecnologia das máquinas de busca*. (2<sup>a</sup>. ed.). Porto Alegre: Bookman.

**Barbieri, C.** (2001). *BI – Business Intelligence: modelagem e tecnologia*. Rio de Janeiro: Axcel Books.

**Barrat A., Barthélemy M., Pastor-Satorras R., & Vespignani A.** (2004). *The architecture of complex weighted networks*. *Proceedings of the National Academy of Sciences*, 101(11), 3747-52.

**Bastian, M., Heymann, S., & Jacomy, M.** (2009, março). *Gephi: An Open Source Software for Exploring and Manipulating Networks*. Recuperado em 14 de julho, 2015 de *International AAAI Conference on Web and Social Media web site*: <http://www.aaai.org/ocs/index.php/ICWSM/09/paper/view/154>

**Blondel, V.D., Guillaume, J.-L., Lambiotte, R., Lefebvre, E.** (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, P10008. Recuperado em 2 de junho, 2015, de arXiv:0803.0476v2 [physics.soc-ph]

**Bonacich, P.** (1972). Factoring and weighting approaches to status scores and clique identification. *Journal of Mathematical Sociology*, 2(1), 113–120.

**Bonacich, P.** (1987). Power and centrality: a family of measures. *American Journal of Sociology*, 92(5), 1170–1182.

**Bonacich, P., & Lloyd, P.** (2001). Eigenvector-like measures of centrality for asymmetric relations. *Social Networks*, 23(3), 191–201.

**Borgatti, S. P.** (2005). Centrality and network flow. *Social Networks*, 27(1), 55–71.

**Borgatti, S. P., Everett, M. G., & Johnson, J. C.** (2013). *Analyzing social networks*. Thousand Oaks: SAGE Publications Limited.

**Brandes, U.** (2001). A Faster Algorithm for Betweenness Centrality. *Journal of Mathematical Sociology*, 25(2), 163-177.

**Brin, S., & Page, L.** (1998). The anatomy of a large-scale hypertextual web search engine. *Computer Networks and ISDN Systems*, 30(1-7), 107–117.

Empresa Brasileira de Pesquisa Agropecuária, & Ministério da Agricultura, Pecuária e Abastecimento. (n.d.a). *Quem somos*. Recuperado em 20 de junho, 2015, de <https://www.embrapa.br/quem-somos>

Empresa Brasileira de Pesquisa Agropecuária, & Ministério da Agricultura, Pecuária e Abastecimento. (n.d.b). *Portfólios*. Recuperado em 22 de junho, 2015, de <https://www.embrapa.br/pesquisa-e-desenvolvimento/portfolios>

**Freeman, L. C.** (1979). Centrality in social networks: Conceptual clarification. *Social Networks*, 1, 215–239.

**Freitas, L.Q.** (2010). *Medidas de centralidade em grafos*. Dissertação de mestrado não publicada, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brasil.

**Guedes, D. A.** (2013). *Classificação Dinâmica de Nós em Redes em Malha Sem Fio*. Dissertação de mestrado não publicada, Instituto de Informática, Universidade Federal de Goiás, Goiânia, Brasil.

**Jacomy, M., Venturini, T., Heymann, S., & Bastian, M.** (2014). ForceAtlas2, a Continuous Graph Layout Algorithm for Handy Network Visualization Designed for the Gephi Software. *PLoS ONE*, 9(6), e98679. Recuperado em 2 de junho, 2015, de <http://doi.org/10.1371/journal.pone.0098679>.

**Lambiotte, R., Delvenne, J. C., & Barahona, M.** (2009). *Laplacian Dynamics and Multiscale modular structure in Networks*. Recuperado em 2 de junho, 2015, de arXiv:0812.1770 [physics.soc-ph]

**Lazega, E., & Higgins, S.S.** (2014). *Redes sociais e estruturas relacionais* (1<sup>a</sup>. ed.). Belo Horizonte: Fino Traço.

**Penteado, M.I.O., Fontes, R.R, Campos, F.A.A, Euclides, K., Filho** (2014). A trajetória do planejamento da pesquisa na Embrapa. *Cadernos de Ciência e Tecnologia*, 31(1), 35-60.

**Santos, N.** (2014). *Análise de redes sociais como ferramenta para melhorar a performance organizacional*. Recuperado em 13 de agosto, 2014, do Knowtec – Inteligência para Inovação Web Site: <http://knowtec.com/artigos/analise-de-redes-sociais-como-ferramenta-para-melhorar-a-performance-organizacional/>

**Silva, E. M.** (2013). *Recuperação da informação através de busca comparada em domínio específico, baseado em Expressões multipalavras*. Tese de Doutorado não publicada, Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, Brasil.