

**PONTIFÍCIA UNIVERSIDADE CATÓLICA DE SÃO PAULO
PUC-SP**

Marcelo Stoppa Augusto Corrêa

**Afinimapa:
Mapeamento relacional de comunidades,
topologias de afinidade**

Mestrado em Tecnologias da Inteligência e Design Digital
Área de concentração: Design Digital e Inteligência Coletiva

SÃO PAULO
2015

Marcelo Stoppa Augusto Corrêa

**Afinimapa:
Mapeamento relacional de comunidades,
topologias de afinidade**

Mestrado em Tecnologias da Inteligência e Design Digital
Área de concentração: Design Digital e Inteligência Coletiva

Dissertação apresentada à Banca Examinadora da
Pontifícia Universidade Católica de São Paulo, como
exigência parcial para obtenção do título de MESTRE em
Tecnologias da Inteligência e Design Digital, área de
concentração: Design Digital e Inteligência Coletiva sob a
orientação da Prof^{ta}. Dr^a. Pollyana Ferrari Teixeira

SÃO PAULO
2015

Banca Examinadora

Aprovada em: _____

Com o apoio da CAPES e da FUNDASP

AGRADECIMENTOS

À minha orientadora, Pollyana Ferrari, pela confiança depositada neste trabalho árduo e complexo; por incitar a coragem, perseverança e fé. Obrigado pelo aprendizado que vai além dos limites desta dissertação.

Aos professores do Programa, que me escoltaram pelas largas avenidas do saber, ora caóticas, ora pacíficas, sempre com gentileza tão rara hoje em dia: Sérgio Basbaum, Fábio Fernandes, Marcus Bastos, Winfried Nöth e Jorge Albuquerque Vieira. À banca de qualificação, composta por Basbaum e Marcus, cujas inestimáveis contribuições enriqueceram este trabalho.

A Edna Conti, pela atenção, afabilidade e extrema eficiência.

Um agradecimento especial a três colegas que contribuíram nas reflexões sobre o mapeamento de afinidades em torno de cafés e indicações de leituras indispensáveis que ajudaram a construir o corpo deste trabalho: Fernando Perez de Brito, Edson Rocha e Flavio Murahara, obrigado pela disponibilidade e pelo aprendizado.

Aos colegas Amanda Porto de Oliveira, Fabiana Martins, Silvio Ferreira, Letícia Mahlmeister, Raíza Bruscky, Eric Viana, Renato Almada Alonso, Rodrigo Ferreira, Guilherme Cestari, Angela Toth, Marcel Casarini e Gislaine Pozzetti pelo carinho, sugestões de leituras e pelos momentos especiais que compartilhamos.

A Nanci Farina, Ana Luísa Néca, Elaine Levisky, Juliana Winkel.

Ao apoio incondicional de quem testemunhou o empenho para realizar este trabalho, em especial a Ana Stoppa e Vianney Fouques Duparc.



“Todo homem cria sem sabê-lo
Como ele respira
Mas o artista se sente criar
Seu ato envolve todo seu ser
Sua dor bem-amada o fortifica”
Paul Valéry

RESUMO

O espírito moderno trouxe profundas transformações estruturais à sociedade. Os novos arranjos político-sociais, mais fluidos, mudaram não apenas as formas com que se dão as relações interpessoais, mas também como se formam os arranjos sociais pelos quais os indivíduos se vinculam uns aos outros: os grupos, as multidões e as comunidades não são os mesmos. As mutações sócio-políticas, com a sofisticação cada vez maior dos meios de comunicação, aumentam a efemeridade dos agrupamentos, dando às multidões o poder de se auto-organizarem e as diferentes comunidades do planeta efetuam trocas econômicas com muita rapidez e que, pelo confronto de identidades culturais, tecem códigos de comunicação e cultura cada vez mais híbridos.

O presente trabalho propõe a construção de uma metodologia que visa colaborar com a investigação de fenômenos relacionais, culturais e sociais nos grupos, bem como investigar a dinâmica nas multidões e nas comunidades por meio de um conjunto de técnicas de captura, análise e visualização de dados para a construção de topologias relacionais e de afinidade, que nomeamos afinimapas.

Esta metodologia transdisciplinar apoia-se em *Big Data*, *Open Data*, *analytics*, ontologia e algoritmos de visualização de dados complexos, no eixo técnico; nas obras de Jacob Levy Moreno e Timothy Leary, no eixo psicológico e sociométrico; e na infografia e na topologia, no eixo artístico. Ela pretende fornecer a representação da complexidade das relações de diferentes tipos de atores para transcender a visão e favorecer a detecção de padrões de arranjos e comportamentos. Deste modo, deseja-se contribuir com as investigações conduzidas por diferentes áreas do saber que levem em conta as relações entre o homem e o mundo.

Palavras-chave

Big Data, visualização de dados, infografia, sociometria, redes sociais, analytics, topologia de relações.

ABSTRACT

The modern spirit performed deep, structural transformations in society. The new and more fluid socio-political settings changed not only the way interpersonal relations happen, but also the arrangements by which individuals may attach to one another: groups, multitudes or communities have evolved to a new dimension. The socio-political changes and the growing sophistication of media render the groupings ephemeral, empowers the crowds and favors the creation of new and hybrid communicational and cultural codes through cultural identity clashes caused by faster economic exchange.

The present work lays out a methodology that aims the collaboration to research relational, cultural and social phenomena in groups, to analyze crowd and community dynamics through an ensemble of techniques to crawl, analyze and visualize data and build relational and affinity topologies which we named affinimaps.

This transdisciplinary methodology stands on Big Data, Open Data, analytics, ontologies and complex data visualization algorithms, as the technical axis; on the works of Jacob Levy Moreno and Timothy Leary, as the psychological and sociometric axis; and on infographics and topology, on the artistic axis. It intends to offer the representation of complex relations of different sorts of actors so as to transcend the vision and improve the detection of arrangement and behavior patterns. This way, it might contribute to the research conducted by different knowledge areas investigating the relationships between men and the world.

Keywords

Big Data, data visualization, infographics, sociometry, social networks, analytics, relationship topologies.

LISTA DE FIGURAS

Figura 1 – Experiência imediata e conteúdos da consciência segundo Wundt (JACO-VILELA, 2008, p.99)	25
Figura 2 - Captura de tela do aplicativo MindMirror no Facebook, com círculo elaborado por Timothy Leary para a medição do tipo de inteligência, criatividade e imaginação do respondente.	29
Figura 3 - Esquema do algoritmo de afinidade.....	32
Figura 4 - Questionário em forma de mandala	33
Figura 5 - Classificação do conhecimento de acordo com Ephraim Chambers, apud Lima (2011, p. 37), em Cyclopaedia, um dicionário universal das artes e da ciência.....	37
Figura 6 - "Ensaio de uma distribuição genealógica das ciências e das principais artes", de Chrétien Frédéric Guillaume Roth (1769), baseado na Enciclopédia de Diderot e Alembert em Lima (2011, p. 40).....	38
Figura 7 - Detalhe da figura anterior (Lima, 2011, p. 40).....	39
Figura 8 - Tipologia dos problemas da ciência entre os séculos XVII e XX. (Lima, 2011)	41
Figura 9 - Propostas de topologias de rede feitas por Paul Baran, 1964 (Lima, 2011, p. 55).....	42
Figura 10 - Tipologia de representações de gráficos sobre complexidade, de acordo com Lima, 2013.	43
Figura 11 - Diagrama com estruturas típicas de grupos elaborado por Moreno (1994, p. 24).....	46
Figura 12 - Exemplo de sociograma em Moreno (1994, p.35).....	48
Figura 13 - Exemplo de sociomatriz em Moreno (1994, p. 36).....	49
Figura 14 - Lavanderia: estrutura antes da reconstrução (Moreno, 1994, p. 151). O círculo marcado com “FM” representa a responsável pela lavanderia; as flechas mais escuras representam atração e as mais claras, rejeição.	51
Figura 15 - Lavanderia, após retestagem em Moreno (1994, p. 152).	52

Figura 16 - Sistema de prototipagem de sociogramas-alvo, por Northway (apud Lima, 2013, p. 78).	54
Figura 17 - Exemplo de arquivo de log do protocolo HTTP.	57
Figura 18 - Captura de tela do Google Analytics, com visualizações dos dados de acesso de um blog	59
Figura 19 - Análise gráfica da ação PETROBRAS PN, mostrando canal de alta.	61
Figura 20 - Padrões das flores de brócolis romano	62
Figura 21 - "A força das nações", de Paley, Klavans et al (2006), com os domínios científicos mais ativos nos EUA.	64
Figura 22 - Infografia realizada a partir de dados do WikiLeaks, por Simon Rogers	66
Figura 23 - Visualização sobre WikiLeaks elaborada por Wild (2015).	67
Figura 24 - Detalhe do tecido bordado a mão de um vestido criado por Balenciaga	76
Figura 25 - Jovens universitários com estilo hipster. HiFoto: Marti Sans/ALAMI	83
Figura 26 - Perturbações na rede levam à emergência de novos padrões (Toubould, 2015)	86
Figura 27 - Visualização gerada a partir de conexões do Facebook	90
Figura 28 - Visualização gerada com dados de conexões do Facebook, detalhe de uma comunidade	91
Figura 29 - Visualização Partidos do Brasil por afinidade	92
Figura 30 - Captura de tela do jogo Mind Mirror (1985) rodando em emulador.	95
Figura 31 - Mapa da comunidade "Dudu Guru" no site www.facebook.com.	96
Figura 32 - O aplicativo do Facebook, Netvizz, processando informações de usuários e suas respectivas conexões para gerar o arquivo de saída a ser lido pelo Gephi	101
Figura 33 - Carregamento da comunidade "Gay Indie Lovers" mostra indivíduos e relações entre eles antes da aplicação de um algoritmo para agrupar as subcomunidades.	102
Figura 34 - Detalhe da imagem anterior, mostrando nós e arestas sobrepostos, "embaraçados".	102

Figura 35 - Depois da aplicação do algoritmo de distribuição “Force Atlas 2”, alguns agrupamentos surgem à margem da massa de nós.....	103
Figura 36 - Aplicação do algoritmo de mapa de calor.....	104
Figura 37 - Estrutura dos dados para a visualização da comunidade.	104
Figura 38 - Professores e alunos da PUC-SP.....	105
Figura 39 - Agrupamento de alunos de intercâmbio.....	106
Figura 40 - Agrupamento de designers do jornal Folha de S. Paulo presentes em nossa rede de contatos.....	107
Figura 41 - Captura de tela da tabela de dados sobre partidos políticos.....	108
Figura 42 - Afinimapa da classe de 1880-1881	110
Figura 43 - Mapa destacando comunidades e influenciadores de grau médio	115
Figura 44 - Importação de dados IMAp com um plug-in do Gephi.	118
Figura 46 - Dados carregados no Gephi da nossa conta IMAp, pelo plug-in de importação.	118
Figura 47 - Aplicação de modularidade nos dados IMAp para detectar comunidades..	119
Figura 48 - Detalhe da visualização dos dados IMAp.	119
Figura 49 - Importação de dados IMAp com um plug-in do Gephi.	122
Figura 50 - Dados carregados no Gephi da nossa conta IMAp, pelo plug-in de importação.	122
Figura 51 - Aplicação de modularidade nos dados IMAp para detectar comunidades..	123
Figura 52 - Detalhe da visualização dos dados IMAp.	123
Figura 53 - Captura de tela do site do aplicativo Twistori.....	125
Figura 54 - Captura de tela do aplicativo SocialCollider.....	125
Figura 55 - Captura de tela após instalação do Nutch em máquina virtual.....	126
Figura 56 - Rede de filósofos construída por Griffen (2013).....	128
Figura 57 - Rede de artistas construída por Griffen (2013).	128

Figura 58 - Captura de tela do site do aplicativo Twistori.....	130
Figura 59 - Captura de tela do aplicativo SocialCollider.....	130
Figura 60 - Captura de tela após instalação do Nutch em máquina virtual.....	131
Figura 61 - Rede de filósofos construída por Griffen (2013).....	133
Figura 62 - Rede de artistas construída por Griffen (2013).	133

SUMÁRIO

Apresentação	14
1. Vínculos, posições de sujeito e identidade.....	16
1.1. Eu e o outro	23
1.2. Vínculos.....	26
1.3. Visualização de vínculos	30
2. Representação de dados complexos ou de grande volume.....	35
2.1. De árvores a rizomas.....	36
2.2. Mapas sociométricos de Jacob Levy Moreno.....	44
2.3. Tipos de diagramas sociométricos	47
2.4. O projeto Hudson: sociometria e empatia	50
2.5. Big Data e visualização de dados	56
3. Da era das multidões à das redes sociais.....	69
3.1. Multidão e comunidades	69
3.2. Comunidades artificiais.....	73
3.3. Mudanças causadas pelas comunidades artificiais	75
3.4. Extração e análise de dados em comunidades artificiais	78
3.5. Análise da comunidade <i>hipster</i>	80
3.6. Redes sociais x redes relacionais.....	89
4. Como executar mapas	98
4.1. A interface do <i>Gephi</i>	99
4.2. Obtenção do conjunto de dados	99
4.3. Influenciadores e viralização.....	109
5. Obtenção de diferentes conjuntos de dados para construir mapas.....	116
Conclusão.....	135
Bibliografia	138

APRESENTAÇÃO

Esta é uma jornada que busca percorrer a trajetória dos agrupamentos humanos por meio da evolução das comunidades, sobretudo, nestes novos tempos de comunicação e interação *online*, analisando as possibilidades e desafios que eles nos trazem. Pretende, ainda, discutir algumas das possibilidades técnicas à nossa disposição para extrair o máximo potencial da imensa gama de dados hoje acessível, em busca de fortalecer e dinamizar estas relações.

O capítulo que abre esta investigação tece o pano de fundo sobre o qual se projetam algumas ideias e eventos sociais da história recente, que afetaram nosso modo de estabelecer relações interpessoais e a criação de identidades.

Em seguida, no segundo capítulo, temos uma exposição sobre o problema da representação de dados complexos e em grande volume. Apresenta-se a sociometria de Jacob Levy Moreno, instrumento da psicologia na construção de visualização de relações, e o cruzamento entre esse conhecimento e os conceitos de *Big Data*, *analytics* e as comunidades *online*.

No terceiro capítulo, são mostradas a polissemia e aplicações dos termos “multidão” e “comunidade” na história. Discutimos ainda a inteligência e o comportamento coletivos, assim como um estudo de caso no qual um fenômeno, dentro de uma comunidade, é examinado por análise matemática e visualização dos dados. São apresentadas ainda topologias de comunidades *online*.

A ferramenta de construção das visualizações de dados é apresentada no quarto capítulo, assim como a metodologia empregada para obter topologias de relações melhores. Com as visualizações obtidas, debatem-se a influência e a viralização, tanto em pequenas comunidades quanto nas redes sociais.

O último capítulo mostra exemplos de caso que sustentam a hipótese deste trabalho e aborda outras técnicas possíveis para a obtenção e cruzamento de grandes volumes de dados e visualizações.

Na conclusão, encontram-se ideias de aplicação da metodologia construída nesta pesquisa, com alguns cenários possíveis.

1. VÍNCULOS, POSIÇÕES DE SUJEITO E IDENTIDADE

O traço mais marcante do espírito da modernidade é a liquefação das estruturas históricas de poder, dos caminhos habituais, da previsibilidade das relações sociais que se baseava na hierarquia e nos papéis sociais bem delimitados. Anunciada no Manifesto Comunista, a ressignificação das relações sociais ocorreu devido à atuação da burguesia, que substituiu “a exploração, encoberta pelas ilusões religiosas e políticas, pela exploração aberta, única, direta e brutal” (Marx e Engels, 2003, p. 28) e “converteu o médico, o jurista, o padre, o poeta, o homem da ciência em trabalhadores assalariados” (idem, ibidem) depois de ter assumido o poder e deposto as relações “feudais, patriarcais e idílicas”. A condição para que a burguesia se sustente enquanto classe é revolucionar constantemente os meios de produção “e, por conseguinte, as relações de produção e, com elas, todas as relações sociais” e o “abalo constante de todas as condições sociais” (idem, p. 29). Marx e Engels testemunham um dos grandes momentos de mudança dos paradigmas relacionais; com as transformações por que passam as estruturas político-sociais ainda em curso, as ondas de ressignificações e reestruturações são sentidas até agora, nos nossos dias, como algo com que devemos nos acostumar. As estruturas deixaram de ser sagradas. “Tudo o que era sólido se evapora no ar, tudo o que é sagrado é profanado, e por fim o homem é obrigado a encarar com serenidade suas verdadeiras condições de vida e suas relações com a espécie” (idem).

A famosa frase do “Manifesto comunista” sobre a evaporação dos sólidos, de acordo com Bauman (2001, p. 9), “referia-se ao tratamento que o autoconfiante e exuberante espírito moderno dava à sociedade, que considerava estagnada demais para seu gosto e resistente demais para mudar e amoldar-se a suas ambições, porque congelada em seus caminhos habituais”. Bauman afirma ainda que o espírito moderno estaria determinado a emancipar a realidade “dissolvendo o que quer que persistisse no tempo e fosse infenso à sua passagem ou imune ao seu fluxo”. Essas transformações, entretanto, não objetivavam uma total liberdade, um retorno às estruturas que a burguesia e o capitalismo suplantaram, mas sim a construção de novas estruturas de poder, “novos sólidos”, que dessem mais flexibilidade aos atores e pudessem desobrigá-los de atos irrelevantes, como “libertar a empresa dos negócios dos grilhões dos deveres para com a família e o lar e da densa trama das obrigações.” (Weber apud Bauman, 2001, p.10)

Para David Harvey (1989, p. 12 apud Hall, 1990, pp. 13-14), a modernidade implica não só num rompimento impiedoso com toda e qualquer condição procedente, mas também em um processo “sem fim de rupturas e fragmentações internas no seu próprio interior”. Já Ernesto Laclau (1990, *ibidem*) prefere usar o conceito de deslocamento. “Uma estrutura deslocada é aquela cujo centro é deslocado, não sendo substituído por outro, mas por uma pluralidade de centros de poder”. Segundo ele, não há nenhum centro nas sociedades modernas, tampouco um princípio articulador e único, já que elas não se desenvolvem conforme uma única causa ou lei. Laclau profere ainda que a sociedade, ao contrário do que já foi estabelecido por sociólogos, não se trata de uma unidade delimitada e que nela reside o poder de executar mudanças evolucionárias, já que é constantemente “descentrada” ou deslocada por forças fora de si mesma. Neste movimento, cria novas identidades e novos pontos de articulação:

O lado positivo do deslocamento é que ele “desarticula identidades estáveis no passado, mas também abre a possibilidade de novas articulações, a criação de novas identidades, a produção de novos sujeitos e o que ele chama de “recomposição da estrutura em torno de “pontos nodais particulares de articulação” (Laclau, 1990, p. 40 apud Hall, 1990, p. 14).

Constatamos a presença dessa força de deslocamento e de transformação das relações sociais quando tomamos distância da esfera da experiência pessoal e local e observamos os grandes movimentos econômicos, as ondas migratórias, a desigualdade de renda, o acesso à educação e ao consumo; ou nas complexas relações entre poder e circulação monetária e na atuação de grupos específicos nos mercados de capitais. Também na busca por inovação incessante de produtos e modos de vida ou pela expansão constante do mercado, que impele os proprietários dos meios de produção a estarem presentes em todo o planeta porque necessitam “estabelecer-se em toda a parte, explorar em toda a parte, criar vínculos em toda a parte”. (Marx e Engels, 2003, p. 29).

Presenciamos ainda, na contemporaneidade, que os vínculos criados pelo capitalismo foram suficientemente fortes para espalhá-lo pelo mundo. Giddens (1990, p. 6 apud Hall, 1990, p. 13) diz que “à medida em que áreas diferentes do globo são postas em interconexão umas com as outras, ondas de transformação social atingem virtualmente

toda a superfície da Terra”, transformando também as instituições modernas. Por instituições modernas, ele entende os novos arranjos espaço-temporais criados em torno de formas tradicionais como as cidades, o trabalho assalariado e a mercantilização de produtos (idem, p. 6), o que levou também à “extração de relações sociais dos contextos locais de interação e sua reestruturação, ao longo das escalas indefinidas de espaço tempo” (idem, p. 21). A compressão da escala espaço-temporal impactou o que Massey chama de “geometria do poder”:

... eu quero fazer um simples ponto aqui sobre o que se poderia chamar de geometria do poder, a geometria do poder da compressão espaço-tempo. Diferentes indivíduos e grupos sociais passam a ocupar posições distintas em relação a esses fluxos e interconexões. Este ponto não diz respeito apenas à questão de quem se move ou não, apesar de este ser um elemento importante; diz respeito ao poder em relação aos fluxos e movimentos. Diferentes grupos sociais têm relações distintas com esta mobilidade, de qualquer maneira, diferenciada: algumas pessoas “comandam” mais do que outras, algumas iniciam fluxos ou movimentos, outras não; algumas sentem seus impactos mais que outras, algumas são efetivamente aprisionadas.

De certa forma, ao final deste espectro, estão aqueles que estão fazendo ambas, a movimentação e a comunicação, e os que estão, de alguma forma, em uma posição de controle em relação a ele – os *jet-setters*, os que enviam e recebem faxes e e-mails, fazem conferências internacionais por telefone, distribuem os filmes, controlam as notícias, organizam os investimentos e as operações em moedas internacionais. Estes são os grupos que são realmente, de certo modo, responsáveis pela compressão do espaço-tempo, que podem realmente usá-la e tirar proveito dela, cujo poder e influência muito definitivamente aumenta. Em suas margens mais prosaicas, esse grupo provavelmente inclui um bom número de acadêmicos e jornalistas ocidentais – aqueles, em outras palavras, que escrevem mais sobre isso.

Mas também há grupos que estão fazendo movimentos físicos, mas que não estão “no comando” do processo da mesma forma que outros. Os refugiados de El Salvador ou da Guatemala e os trabalhadores migrantes não documentados de Michoacán, no México, lotando Tijuana para fazer uma travessia talvez fatal de um lado para o outro da fronteira dos EUA em busca de uma chance para uma nova vida. Aqui a experiência do movimento e, de fato, de uma pluralidade confusa de culturas, é muito diferente. [...]

Ou – um último exemplo para ilustrar um tipo diferente de complexidade – existem pessoas que vivem nas favelas do Rio, que conhecem o futebol internacional como as palmas de suas mãos, que criaram alguns de seus jogadores, que contribuíram fortemente para a

música internacional, que nos deram o samba e a lambada dançados em Paris e em Londres, e que nunca foram ao centro do Rio de Janeiro. Em um nível, houve grandes contribuintes ao que chamamos de compressão espaço-tempo, e em outro existem os que são aprisionados por ela. Isto é, em outras palavras, uma diferenciação social altamente complexa. Existem diferenças no grau de movimento e de comunicação, mas também no grau de controle e de iniciação. As maneiras pelas quais as pessoas são colocadas dentro da compressão espaço-tempo são altamente complicadas e extremamente variadas (Massey, 1991, p. 25-26 apud Hall, 1990, pp. 45-46).

Portanto, de acordo com a perspectiva apresentada por Giddens e sustentada por Massey, não seria impossível afirmar que a nova organização que surge do capitalismo globalizado começa a ficar menos dependente das distâncias e relações geográficas e a ênfase passa a ser dada aos fluxos e vínculos: as trocas comunicacionais, mercantis, os fluxos humanos e as relações sociais – todos suspensos, praticamente desconectados do funcionamento do espaço-tempo percebidos nas relações presenciais. A empresa globalizada rompe fusos-horários, cria turnos de trabalho alternativos ao clássico “das nove às dezessete” para que o seu fluxo de produção seja o fio narrativo a que se vincula a vida de seus empregados. Por outro lado, com a abundância atual de modos de comunicação promovida pelas tecnologias de informação e da comunicação, aquele que desejar vender sua mão de obra em localidade muito distante da sua, pode fazê-lo de acordo com a própria flexibilidade, dada a oferta de cada mercado. Estes dois modos distintos de venda de mão de obra que coexistem no cenário contemporâneo pode ser indícios de que diferentes modos de vida, e mesmo de personalidade, podem existir simultaneamente sem prejuízo de um ou outro. A oferta dos poderes à qual é possível se vincular é farta: estar disponível seja para o modo de trabalho mais habitual, para uma empresa, seja de forma independente, requer das pessoas a habilidade de se organizar em sua identidade para transitar entre um modo de produzir e de viver e outro, entre um poder e outro. “A vida no capitalismo tardio”, segundo Adorno (2009, p.54), “é um rito de permanente iniciação. Todos devem mostrar que se identificam sem a mínima resistência com os poderes aos quais estão submetidos.” E que, ainda, “a capacidade de escorregar e de se arranjar, de sobreviver à própria ruína [...] é própria da nova geração”. Transitar entre as velhas e as novas identidades, oscilar como um pêndulo entre as duas e, de certa forma, estar vinculado às duas formas: surge daí a pluralidade das formas de existir face às mudanças constantes. Existir como um devir para os arranjos sociais emergentes,

reinventar-se. Em algumas partes do planeta, a globalização impõe ainda uma outra dimensão às comunidades locais, que é a ocidentalização. Não são apenas vínculos são criados, mas também diferenças, pela aproximação de contrastes.

Embora tenha se projetado a si próprio como trans-histórico e transnacional, como força transcendente e universalizadora da modernização e da modernidade, o capitalismo global é, na verdade, um processo de ocidentalização – a exportação de mercadorias, dos valores, das prioridades, das formas de vida ocidentais. Em um processo de desencontro cultural desigual, as populações “estrangeiras” têm sido compelidas a ser os sujeitos subalternos do império ocidental, ao mesmo tempo em que, de forma não menos importante, o Ocidente vê-se face a face com a cultura “alienígena” e “exótica” do seu “outro”. A globalização, à medida que dissolve as barreiras da distância, torna o encontro entre o centro colonial e a periferia colonizada imediato e intenso (Robins, 1991, p. 25 apud Hall, 1990).

As diferenças de identidade que Robins aponta são confrontos que ocorrem por conta da proximidade de diferentes culturas que não se encontrariam se não fosse pela globalização, com seus fluxos econômicos e culturais que se propagam por modos de comunicação cada vez mais sofisticados, do telégrafo à Internet; da mídia unidirecional para aquela em que as posições de emissor e destinatário são fluidas. Surgem, neste ambiente, novas e mutáveis identidades culturais em torno das quais se aglutinam indivíduos em comunidades, a partir do cruzamento das culturas locais e da cultura ocidental. Nasce da tessitura da rede mundial de computadores o conceito da cultura do futuro, que é universal sem ser totalizante, chamada por Levy (1990) de *cibercultura*:

[...] neologismo "cibercultura", especifica aqui o conjunto de técnicas (materiais e intelectuais), de práticas, de atitudes, de modos de pensamento e de valores que se desenvolvem juntamente com o crescimento do ciberespaço (Levy, 1999, p. 15-16).

A cibercultura mantém a universalidade ao mesmo tempo em que dissolve a totalidade. Corresponde ao momento em que nossa espécie, pela globalização econômica, pelo adensamento das redes de comunicação e de transporte, tende a formar uma única comunidade mundial, ainda que essa comunidade seja - e quanto! - desigual e conflitante. Única em seu gênero no reino animal, a humanidade reúne toda sua espécie em uma única sociedade. Mas, ao mesmo tempo, e paradoxalmente, a unidade do sentido se quebra, talvez porque ela comece a se realizar na prática, pelo contato e a interação efetivos. Conectadas ao universo, as comunidades virtuais constroem e dissolvem constantemente suas micrototalidades dinâmicas, emergente, imersas,

derivando entre as correntes turbilhonantes do novo dilúvio. (Levy, 1999, p. 249).

Até aqui falamos da vinculação do indivíduo aos grandes movimentos de transformação econômica e social, de como se requer do indivíduo, na contemporaneidade, a adaptação aos deslocamentos de identidade impostos a princípio pela burguesia e, mais tarde, pela globalização e pela emergência da cibercultura como entidades totalizadoras e formadoras de identidade. O que se pretende oferecer é uma ótica pela qual seja possível estudar os fenômenos de identidade e transformações sociais a partir, por exemplo, das trocas comunicacionais digitais que o indivíduo efetua no ciberespaço, sejam elas entre pessoas ou entre pessoas e sistemas. No caso da comunicação entre pessoas, contemplamos as redes sociais e outras formas de comunicação digital *peer-to-peer*; entre as pessoas e os sistemas, os rastros que os usuários deixam ao utilizar sistemas de informação, como *cookies*, endereços IP e etc. No centro desta análise estão a afinidade e o vínculo.

Para começarmos a falar de vínculos, aqui vai um convite à imaginação. Imaginemos as transformações socioeconômicas ocorridas desde a liberação do espírito moderno como um fluxo, um rio caudaloso cujo destino incontornável é erodir as margens para que se alargue; que, com sua violência, destrói o antigo leito e que, nesta transformação brutal, move, carrega consigo pesadas rochas às quais não pode mais conformar-se. Além das transformações físicas e geológicas, um novo ecossistema desenvolve-se no rio, restabelecendo novas codependências, novos arranjos e vínculos; pois o rio não é mais o mesmo e uma nova ordem vigora para cada elemento participante desta primeira imaginação.

O leito de rio e as pedras desta metáfora poderiam ser interpretados como os dispositivos sociais. Para Foucault, um dispositivo é algo que:

[...] comporta discursos, instituições, organizações arquiteturais, decisões regulamentares, leis, medidas administrativas, enunciados científicos, proposições filosóficas, morais, filantrópicas, em resumo: o dito e também o não dito, eis os elementos do dispositivo. O dispositivo

ele mesmo é a rede que estabelecemos entre esses elementos [...]
(Foucault, 1977 apud Agamben, 2005, tradução nossa)¹

Abandonando a filologia foucaultiana para situar a definição de “dispositivo” em outro contexto, Agamben propõe duas grandes classes de dispositivos:

Proponho-lhes nada menos que uma geral e maciça divisão do existente em dois grandes grupos ou classes: de um lado os seres vivos (ou as substâncias) e de outro os dispositivos nos quais estes estão incessantemente capturados. De um lado, ou seja, para retomar a terminologia dos teólogos, a ontologia das criaturas e de outro a *oikonomia* dos dispositivos. (Agamben, 2005, p. 5)

É nesta dança de fluxos que estamos submersos, nós, criaturas, carregados pelo curso dos dispositivos, estes tendo “a capacidade de capturar, orientar, determinar, interceptar, modelar, controlar e assegurar os gestos, as condutas, as opiniões e os discursos dos seres vivos” (idem). À medida em que descemos o rio, tentamos criar ligações com outras pessoas; porém, dada a violência do turbilhão em que estamos, tais vínculos não são suficientemente fortes para perdurar. Em comparação às ligações que haviam entre pessoas antes da substituição de estruturas sociais promovida pela burguesia, como postulou Marx, as ligações contemporâneas são fluidas no sentido de que os arranjos sociais são voláteis: nascem e desaparecem com muita rapidez, arrastados pela correnteza.

Para Laclau (apud Hall, 1990, p. 14), as sociedades contemporâneas têm por principal característica a diferença:

Elas são atravessadas por diferentes divisões e antagonismos sociais que produzem uma variedade de diferentes “posições de sujeito” – isto é, identidades – para os indivíduos. Se tais sociedades não se desintegram totalmente não é porque elas são unificadas, mas porque seus diferentes elementos e identidades podem, sob certas circunstâncias, ser conjuntamente articulados. Mas essa articulação é sempre parcial: a estrutura da identidade permanece aberta. Sem isso, argumenta Laclau, não haveria nenhuma história.

¹ “comportant des discours, des institutions, des aménagements architecturaux, des décisions réglementaires, des lois, des mesures administratives, des énoncés scientifiques, des propositions philosophiques, morales, philanthropiques, bref : du dit aussi bien que du non-dit, voilà les éléments du dispositif. Le dispositif lui-même c'est le réseau qu'on établit entre ces éléments [...]”.

1.1. Eu e o outro

Agora, um convite a uma segunda cena imaginária. Pensemos numa dança. Pode ser uma quadrilha que começa com os dançarinos em pares, fazendo diferentes combinações de parceiros durante a coreografia. Ou na dança da interação entre moléculas, que gera outras delas. Por analogia, a hipótese que se propõe é a seguinte: de que forma, na nossa contemporaneidade, seria possível melhor identificar os diferentes tipos de arranjos entre pessoas, como grupos, multidões ou comunidades online, para em seguida melhor analisar os fenômenos de grupo? Para tentar responder à questão, a pesquisa adotou uma visão transdisciplinar que passa pelos prismas da psicologia, sociologia, análise de dados em grande quantidade, visualização de dados, topologia, redes sociais e filosofia.

A hipótese pode ser declinada de modos distintos, de acordo com a linguagem utilizada por cada prisma. Para os métodos de análise de dados de redes sociais, conhecidos também pelo sistema “*analytics*”, a pergunta seria a seguinte: “com base nos dados que os usuários das redes sociais publicam sobre si, é possível identificar o que torna vínculos interpessoais mais fortes?”; para a psicologia, “os vínculos interpessoais são sustentados pela coexistência e complementaridade de tipologias psicológicas distintas de cada indivíduo?”; para a visualização de dados, “utilizando uma metodologia de mensuração do cruzamento de dados de ligações pessoais, uma visualização dessas ligações pode ser produzida para dar sentido aos dados, para verificar que tipos de organizações emergem?”

Num segundo momento, foi preciso escolher a que distância seria possível abordar o problema. Se recorrêssemos apenas à semiótica, trataríamos da transação simbólica dentro de um grupo, partiríamos para um tipo de redução para buscar traços metafísicos, ou índices ou interpretações subjetivas, para em seguida reconstituí-los, mas este não é nosso objetivo. Nosso foco é buscar quais tipos de qualidades humanas emanam de uma organização de grupo. Dessa forma, optamos por não nos dedicar a este caminho, já que nosso interesse é o fenômeno de organização humana e suas emergências. Deste modo, nos concentramos na observação do fenômeno.

Para Merleau-Ponty (1999, pp. 1-2), a fenomenologia “é a tentativa de uma descrição direta de nossa experiência tal como ela é, e sem nenhuma deferência à sua gênese psicológica e às explicações causais que o cientista, o historiador ou o sociólogo dela possam fornecer”.

Uma solução que se apresenta então como possível seria a de isolar cada elemento do problema e nele operar uma investigação. Primeiramente, perguntou-se o que é um indivíduo, numa tentativa de descobrir, no seu mundo interior, as razões de vincular-se a outros: o que no homem interior, sendo sujeito observador e, num grupo, também objeto de observação, gera o vínculo. Porém, para a fenomenologia, não há uma delimitação exata entre o homem interior e o mundo: “[...] não existe homem interior, o homem está no mundo, é no mundo que ele se conhece” (idem, p. 6).

A busca pela resposta do que é um indivíduo passou por alguns esquemas construídos por Wundt, que descrevem a experiência imediata interagindo com conteúdos da consciência e diagramas sobre a vida mental em categorizações de sentimentos e reações, como o apresentado a seguir. Partindo deste ponto de vista, a investigação poderia ser conduzida, por exemplo, do mesmo modo em que se procedeu nas pesquisas dos primórdios da neurociência, tentando auscultar o corpo, como caixa misteriosa, com grande aparato técnico: captores, tomografias e, por que não, exames laboratoriais. Este tipo de investigação seria baseado na localização e “quantificação” de sentimentos traduzidos em números, a partir, por exemplo, de análises dos batimentos cardíacos e taxas hormonais de uma pessoa em relação a outra, em um labirinto sem fim de classificações - um procedimento que, além de custoso, seria pouco prático na investigação de fenômenos de grupos, comunidades ou massas.

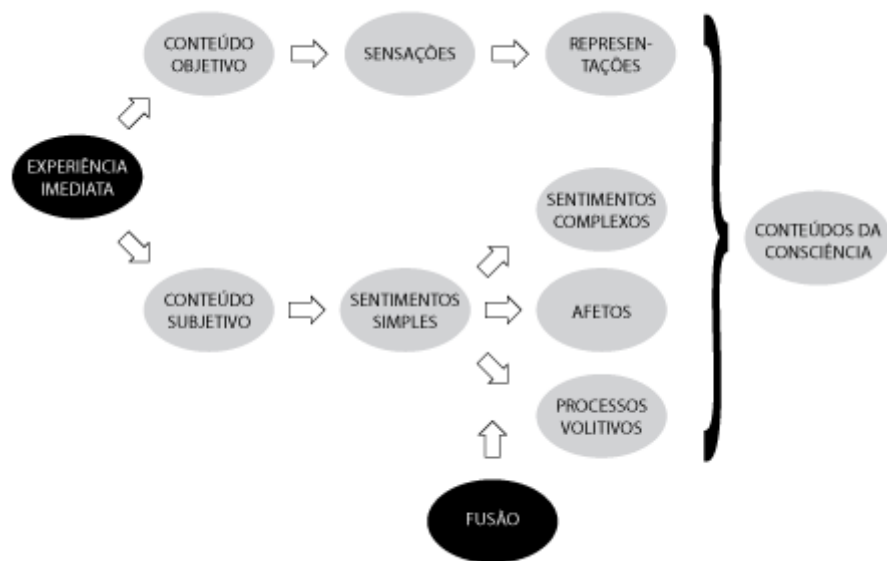


Figura 1 – Experiência imediata e conteúdos da consciência segundo Wundt (JACO-VILELA, 2008, p.99)

O Cogito² cartesiano é um convite para que as respostas sejam encontradas nos processos internos, como no pensamento. A verdade, de acordo com Descartes, não deriva das informações sensoriais e tampouco da realidade de um mundo externo. Mas, se como sustentou Merleau-Ponty, o homem interior não existe e é no mundo que ele se conhece, não há então uma dimensão da realidade totalmente interna separada de outra dimensão externa. Deste modo, nunca poderíamos pensar em um indivíduo como uma ilha de absoluta verdade e fonte de respostas, mas como interagente: é na contiguidade dos processos internos com os fenômenos do mundo que ele se encontra. Merleau-Ponty esclarece:

Até hoje, o Cogito desvalorizava a percepção de um outro, ele me ensinava que o Eu só é acessível a si mesmo, já que ele me definia pelo pensamento que tenho de mim mesmo e que sou evidentemente o único a ter, pelo menos nesse sentido último. Para que outro não seja uma palavra vã, é preciso que minha existência nunca se reduza à consciência que tenho de existir, **que ela envolva também a consciência que dele se possa ter** e, portanto, minha encarnação em uma natureza e pelo menos a possibilidade de uma situação histórica. O Cogito deve revelar-me em situação, e é apenas sob essa condição que a

² Abreviação de “Cogito, ergo sum” (“Penso, logo existo”), segundo o dicionário *Houaiss* – talvez o princípio mais conhecido das “Meditações” de Descartes, segundo o qual a existência de alguém é demonstrada por seu próprio pensamento.
Referências em <http://www.philosophypages.com/hy/4c.htm#cog> (acesso em 25.11.2015).

subjetividade transcendental poderá, como diz Husserl, ser uma intersubjetividade (Merleau-Ponty, 1999, p.9, grifo nosso).

Para esta pesquisa, portanto, o indivíduo é tratado como algo além de categorizações muito precisas, afastado de tipologias e de justificativas densas sobre sua vida interna; é uma consciência que interage com o mundo, que é seu pano de fundo, criando vínculos. A definição última do que é um indivíduo é determinada por diferentes campos do conhecimento; nenhum deles, porém, fornece uma resposta definitiva. O que há são aproximações pluridisciplinares que delimitam o indivíduo de forma imprecisa, condicional, variável.

Eu não sou eu nem sou o outro,
Sou qualquer coisa de intermédio:
Pilar da ponte de tédio
Que vai de mim para o Outro.
(Mário de Sá-Carneiro)

É uma abstração próxima àquela usada nas ciências sociais e no estudo da teoria geral dos sistemas, ângulo este que possibilitará a confecção de mapas dos vínculos.

1.2. Vínculos

Tendo enquadrado o problema do indivíduo, a próxima questão que se propõe é sobre a qualidade ou tipologia dos vínculos, ou como chamá-los. Quando do projeto de pesquisa, pensou-se em nomear os vínculos entre os indivíduos de “sintonia”. A falha desta nomenclatura é que ela pode remeter a conceitos metafísicos não contemplados nesta pesquisa, ou mesmo, pelo fato de a palavra “sintonia” pertencer ao universo do rádio, fazer com que se imagine uma relação entre, por exemplo, ondas cerebrais transmitidas e recebidas entre pessoas, ou que haja uma relação entre feromônios – o que seria uma boa hipótese com que trabalhar, se não remetesse à simples relação de causa-efeito. Sendo assim, não é um bom termo, uma vez que as relações humanas estão na ordem da complexidade e são regidas por um sem-número de influências.

O termo seguinte foi “tipologia”, tendo em mente um dos testes psicológicos³ usados por empresas na gestão de pessoas, mais precisamente o dos tipos psicológicos de Carl Jung e Isabel Myers-Briggs. É um sistema que descreve 16 tipos de personalidade possíveis, pela combinação de 4 dicotomias, a saber: extroversão e introversão (E/I), sensação e intuição (S/N), sentimento e pensamento (T/F), julgamento e percepção (J/P). A tipologia de um indivíduo pode ser representada por ENFP, por exemplo. Jung elaborou as três primeiras dicotomias depois de observar pacientes durante 20 anos de prática clínica de psicologia e psiquiatria ⁴. Este trabalho, de 1921, recebeu a colaboração da quarta dicotomia, por Myers-Briggs. Enquanto a primeira dicotomia é a base da expressão da personalidade, as outras três representam aspectos do inconsciente (Lessa, 2014). Os caracteres secundários podem se manifestar com mais ou menos intensidade em um curto período, digamos, em uma hora ou um dia.

Apesar da sólida base teórica deste sistema de atribuição tipológica, ele requer a execução de um teste longo e custoso, o que encareceria o processo ou mesmo impossibilitaria, por exemplo, de estudar ao final de cada dia como a mudança nas expressões de cada tipologia interagiu com outras, em uma quantificação precisa. Ainda, a lente fenomenológica que adotamos neste estudo procura estabelecer relações de ordem diferente, no que poderia ser um método para medir a empatia no grupo estudado. Deste modo, os sistemas tipológicos poderiam ser abandonados para se elaborar uma outra metodologia.

Em seguida, passamos para uma reflexão sobre a empatia. Pela síntese apresentada no dicionário Houaiss (Houaiss e Villar, 2007), o conceito de empatia é baseado, principalmente, em duas perspectivas: no ponto de vista e na projeção. É a forma de percepção do indivíduo, ou do eu social “mediante três aptidões: para se ver do ponto de vista de outrem, para ver os outros do ponto de vista de outrem ou para ver os outros do ponto de vista deles mesmos”. Um segundo significado diz que a empatia se trata da tentativa de compreender o comportamento do outro (idem). É um termo apropriado para o presente estudo porque trata do indivíduo em relação ao outro, ou seja, sustenta o conceito de indivíduo adotado para a pesquisa. Jacob Moreno cita o papel da empatia na determinação das “distâncias” nas relações entre pessoas:

³ Um teste de personalidade pode ser feito neste endereço: <http://goo.gl/9F5y30>. Acesso em 25.11.2015.

⁴ http://www.jung-rj.com.br/artigos/tipos_psicologicos.htm. Acesso em 25.11.2015.

O papel que a empatia desempenha no processo de tele⁵ ao longo da testagem sociométrica pode ser, facilmente, explicado; o sujeito a ser testado precisará de algum grau de *Einfühlung* para julgar quem entre os membros de determinado grupo, gosta dele, não gosta dele ou lhe tem indiferença (Moreno, p. 183).

Buscou-se, depois disto, algum sistema preexistente de observação e quantificação das interações que pudesse ser usado para tabelar dados dessas interações. Aqueles que se pensou em adotar foram as “mandalas” resultantes de testes aplicados por Timothy Leary⁶: círculos concêntricos que indicam uma gradação entre tipologias, criando gráficos radiais. É interessante observar que, nestes testes, é o indivíduo quem revela dados sobre si via questionário, sendo incentivado a convidar outras pessoas para também respondê-lo, com dados próprios ou pela escolha das gradações de tipologia como se fossem o primeiro respondente, comparando resultados – ou seja, é um processo que conduz os envolvidos a um exercício de empatia. Por outro lado, é uma metodologia ainda baseada em tipos. De qualquer forma, as localizações dos pontos dos respondentes nas mandalas foram analisadas: estimou-se que os pontos estavam distribuídos em 25 gradações.

⁵ “A tele consiste, basicamente, na capacidade de diferenciar pessoas, coisas e objetos como partes separadas de si mesmo e comunicar-se empaticamente com elas. O ser humano pode responder a esta sensação de separatividade com afeto, rejeição ou neutralidade vincular, sendo esta última indiferença ao outro, seja por falta de proximidade física, conhecimento emocional ou diferenças de interesses e valores socioculturais.” (Novaes, 2013)

⁶ Esta metodologia é abordada posteriormente neste trabalho, com mais detalhes.

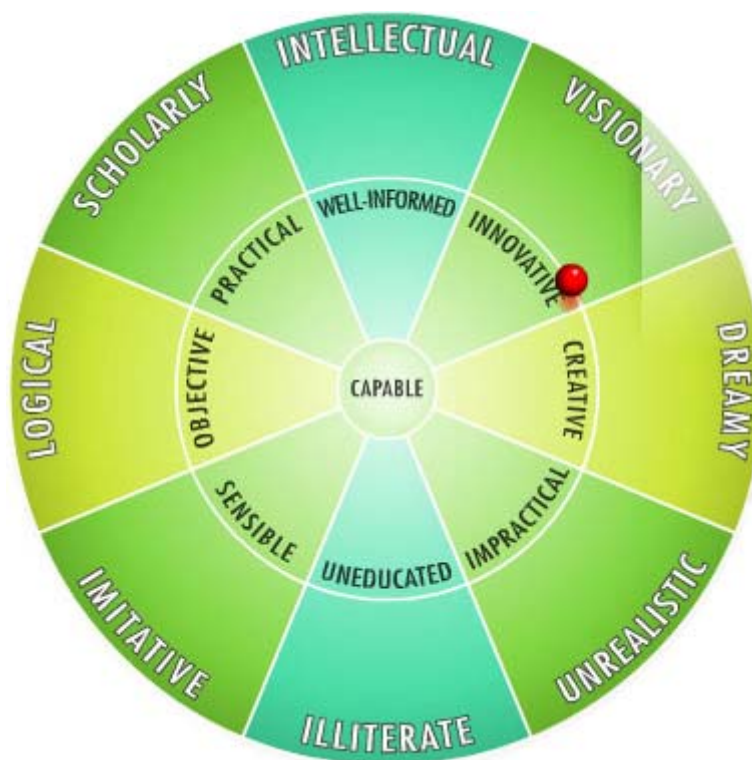


Figura 2 - Captura de tela do aplicativo MindMirror no Facebook, com círculo elaborado por Timothy Leary para a medição do tipo de inteligência, criatividade e imaginação do respondente.

A partir disso, as tipologias foram traduzidas e as gradações criadas por uma planilha *Excel*. Os tipos traduzidos são os da tabela seguinte:

Tipo 1	Gradação (1 - 25)	Tipo 2
organizado	...	desorganizado
amuado	...	doce
apático	...	alerta
inflexível	...	criativo
certinho	...	mente aberta
resignado	...	mandão
bobo	...	esperto
inspirado	...	analítico
relaxado	...	estressado
pretensioso	...	cordato
rico	...	pobre
grosseiro	...	honorável
impaciente	...	hesitante
resmungão	...	motivador
alegre	...	triste
quadrado	...	urbano

ambicioso	...	obediente
superficial	...	contemplativo
altruísta	...	egoísta
tomador	...	doador

Uma folha com o questionário foi criada para ser distribuída em um grupo. Depois de coletadas, as informações alimentariam um banco de dados em que constaria o nome de cada respondente e o valor de 1 a 25 obtido em cada gradação, imaginando que cada uma dessas gradações fosse uma variável do tipo *Integer* e que um *software* poderia ler esses dados. Porém esse caminho foi abortado, já que a concepção de um algoritmo que colocaria as pessoas e tipologias em relação seria complexo demais para nossa limitação de tempo e técnica. Ademais, ainda estaríamos presos às tipologias, que na ferramenta de Leary são úteis para evidenciar traços de personalidade: as interações, porém, não são seu foco. O que nos faltou para avançar neste caminho foi, na verdade, saber qual seria a relação entre empatia e rejeição entre cada par de tipos. Não se pode afirmar, por exemplo, que a relação entre um indivíduo do tipo altruísta e outro do tipo egoísta pode ser representada numericamente, em grau de rejeição ou empatia.

1.3. Visualização de vínculos

O caminho mais claro seria criar uma metodologia própria para este trabalho. Passou-se então à elaboração de um método inspirado em analogia a um concerto. As notas ou acordes de um determinado instrumento complementariam aqueles tocados por outros. Alguns acordes teriam uma nota dissonante para dar a eles nuances. Quando os violinistas tocassem, apesar de seus instrumentos estarem “afinados” e tocando a mesma nota, alguns dos violinos tocariam microdesafinações muito sutis, mas ainda harmônicas, não deixando de ser um conjunto - ou seja, não apresentando dissonância total. A afinação e o tipo dos instrumentos os faria ocupar a mesma região na orquestra.

Isto sugeriu a utilização do termo “afinidade”: os indivíduos com maior afinidade entre si ocupariam a mesma região, numa visualização que representasse suas ligações mais fortes, quaisquer fossem. Deste modo, se o respondente indicasse seu grau de afinidade com cada membro de um grupo, e sendo este grau uma representação numérica de proximidade, seria possível construir mapas de afinidade entre todos utilizando a média

entre os números. Por exemplo, se o indivíduo A indica que seu grau de afinidade entre ele o indivíduo B, numa escala de 1 a 20, é de 12, e o indivíduo B, inquirido sobre a sua afinidade com o indivíduo A, na mesma escala, a situa em 5, a proximidade entre os dois seria de $(12 + 5) / 2 = 8,5$. Com esta média, podemos imaginar numa folha de papel, por exemplo, os indivíduos A e B como pontos situados a uma distância de 8,5. Agora, se um indivíduo C atribui a seu grau de afinidade com A o valor 18 e à afinidade com B o valor 2, e que para A e B o grau de afinidade com C é de 10, a afinidade AC é de $(18+10) / 2 = 14$ e a afinidade BC é de $(2+10) / 2 = 6$.

Em seguida, passamos à construção da visualização destas informações. Com esses três elementos calculados, a figura mentalmente evocada é a de um triângulo cujas arestas o tamanho já era sabido: 8,5; 14 e 6, convertidos em centímetros para a visualização. Restava saber o ângulo entre cada segmento para criar os vértices do triângulo, os quais representariam os indivíduos A, B e C. Os tamanhos das arestas foram colocados em uma calculadora online⁷ e os ângulos obtidos para possibilitar a criação de um triângulo foram 18,04°; 149,34° e 12,62°. O esquema a seguir mostra o resultado obtido com a metodologia imaginada:

⁷

<http://www.mathwarehouse.com/triangle-calculator/online.php> Acesso em 06.11.2015

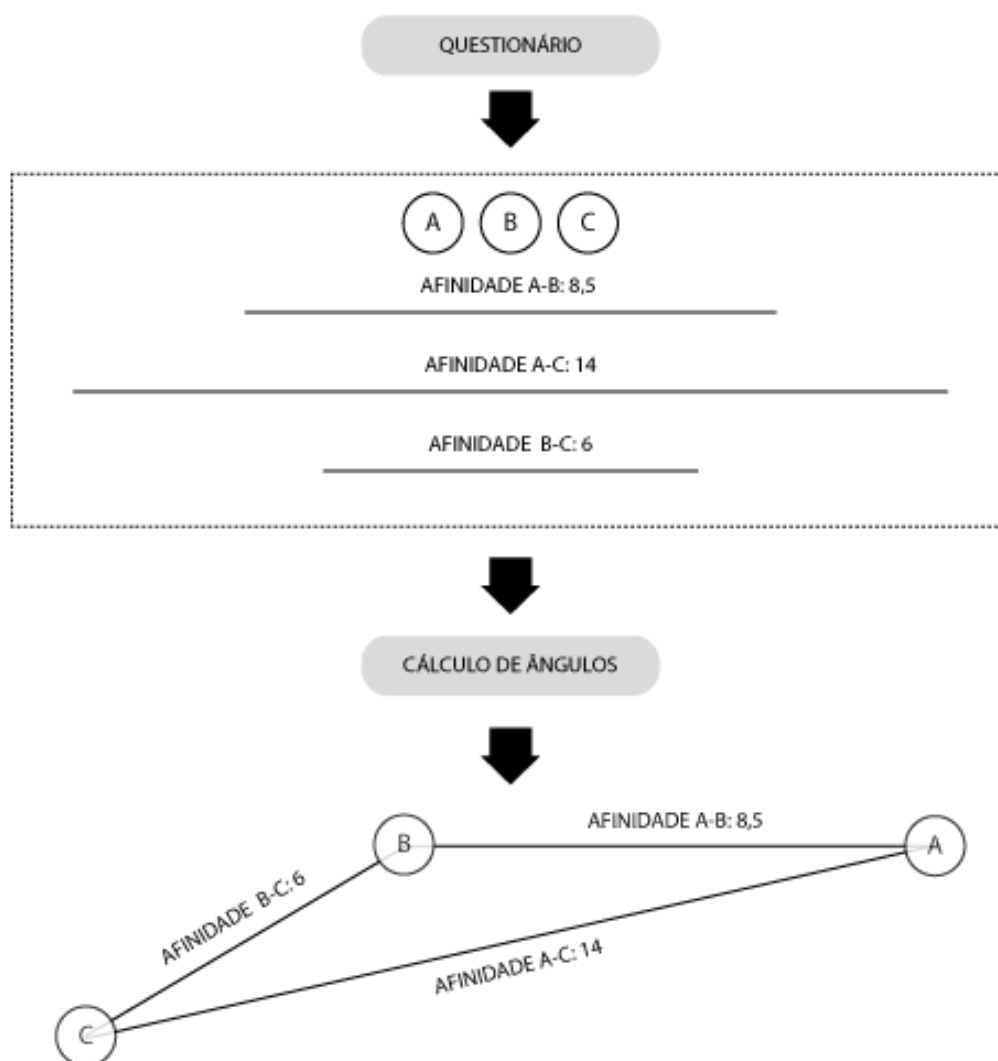
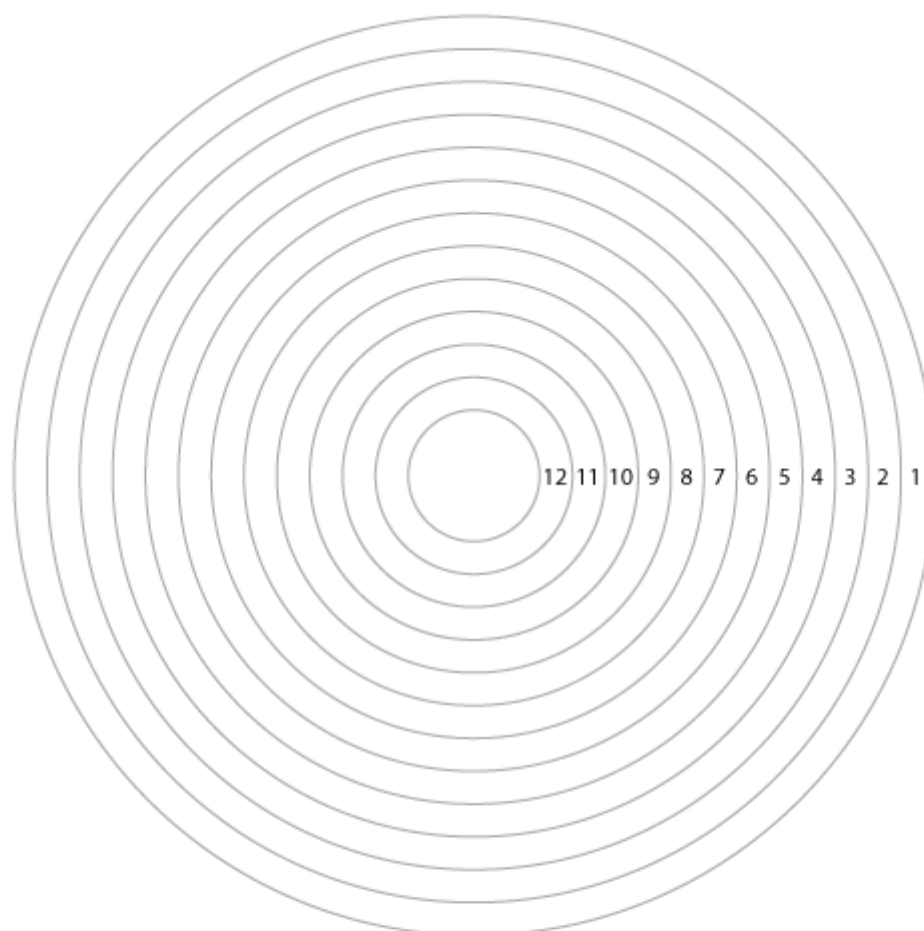


Figura 3 - Esquema do algoritmo de afinidade

Tendo demonstrado a possibilidade de construir uma visualização das afinidades a partir dos dados de “proximidade” sentida por um indivíduo, o próximo passo foi elaborar um questionário. Como o resultado se mostrou através da visualização de relações espaciais, a ferramenta usada para coletar dados também contempla a espacialidade. Assim, elaborou-se o questionário a seguir:

NOME

DATA



A. _____
B. _____
C. _____
D. _____
E. _____
F. _____
G. _____
H. _____
I. _____
J. _____

K. _____
L. _____
M. _____
N. _____
O. _____
P. _____
Q. _____
R. _____
S. _____
T. _____

Figura 4 - Questionário em forma de mandala

No centro de uma mandala se localiza o respondente. A partir de sua percepção do grupo, ele marca em um dos círculos concêntricos as letras que representam os nomes dos membros do grupo. Imagina-se que o questionário seria apresentado ao respondente já

com os nomes dos membros diante de cada letra. Se um dos membros não fosse conhecido pelo respondente, ele seria deixado de fora dos círculos. Para não influenciar o respondente em relação aos números dentro dos círculos, eles poderiam também não ser apresentados; desta forma, a imagem anterior teria o valor de um gabarito.

A partir deste ponto, a pesquisa se dirige à busca de estudos sobre grupos do mesmo gênero e à representação ou visualização de dados complexos em *softwares* para ajudar a construir a visualização dos dados de afinidade.

Se representássemos visualmente os dados de um grupo de 20 pessoas colhidos no questionário elaborado anteriormente, seria necessário programar um algoritmo que calculasse a distância e os ângulos entre cada ponto que representa um indivíduo. A maior quantidade de indivíduos inseridos no mapa de afinidade aumentaria a complexidade dos cálculos e da visualização resultante. Um mecanismo de cálculo se faz necessário para a escalabilidade do mapa; isso significa que o algoritmo deve processar as interações desde um número reduzido de respondentes até de grupos maiores, mais complexos. Assim, buscou-se uma fórmula ou um programa já consagrado para resolver a distância de afinidade em grupos maiores e encontrou-se o *Gephi*, como descrito no capítulo dedicado à construção de visualizações. Os conjuntos de dados distribuídos com o *Gephi* e os algoritmos que o programa traz serviram para testar o conceito e criar procedimentos para fornecer dados ao programa.

Para analisar as visualizações resultantes durante a fase de testes, faz-se necessário tratar agora da representação de dados complexos, construindo o embasamento sobre a visualização da topologia de relações entre indivíduos com trabalhos pioneiros nesta área – e mostrando como o presente trabalho pode atualizar o conhecimento no setor.

2. REPRESENTAÇÃO DE DADOS COMPLEXOS OU DE GRANDE VOLUME

Tomemos dois exemplos de representação de dados: um simples e outro complexo.

Primeiramente, a representação simples de dados seria formada pelos esquemas visuais utilizados para facilitar a compreensão de procedimentos, comparações e mapas. Tem-se tornado corrente a utilização de diagramas nas mídias *online* e *off-line* para explicar ou evidenciar procedimentos, comparações, localizações, séries históricas e pequenas narrativas, entre outros. Há informações que são melhor apreendidas se, ou apenas se, são representadas em diagramas. Por exemplo, uma série de fatos pode ser apresentada em forma de cronologia (no jargão da infografia, uma “*timeline*”); procedimentos também podem ser indicados em forma de lista ou então como fluxograma – é comum encontrar este tipo de representação em manuais de aparelhos eletrônicos, manuais de montagem de móveis, etc.; já o crescimento de um indicador econômico comparado à sua performance em um período anterior seria representado normalmente por um gráfico de barras. São tipos de relações que nos levam de um ponto a outro, caminhos narrativos criados pelas comparações entre dois estados, evidenciando as mudanças.

Agora pensemos em uma visualização de dados mais complexa, do tipo topológico. Dependendo do que se deseja evidenciar, as relações espaciais dos elementos a serem representados podem mostrar localizações tanto por listas de endereços como por mapas: os pontos marcados em mapa que mostre os melhores restaurantes de certo bairro são sobrepostos a uma representação do terreno em que se encontram. Os pontos indicados compartilham duas propriedades: ser um bom restaurante e estar em um determinado bairro. Se a camada geográfica sobre a qual estão sobrepostos for eliminada, o que resta são os pontos que ocupam um espaço abstrato, no qual coexistem porque partilham a propriedade “ser um bom restaurante”; ou seja, o que vemos é um agrupamento por uma simples propriedade, uma interrelação que pode ser representada por uma cor, digamos, vermelho.

Adicionando a este conjunto um tipo diferente de restaurante (por exemplo, aqueles considerados medianos), procedemos como foi feito com o primeiro conjunto e indicamos esta propriedade com a cor amarela; assim, poderíamos observar que a referência geográfica terminaria por não ser tão importante, porque estaríamos visualizando neste espaço abstrato as relações. Mais do que indicar se restaurantes estrelados estariam mais ao norte ou ao sul de uma cidade, observaríamos o seguinte: neste espaço abstrato, os restaurantes de certo tipo estão concentrados em um ponto a partir do qual se tornam mais raros – pois passam a compartilhar espaço com restaurantes de outro tipo. Indo em certa direção do espaço abstrato, terminarão por não ser mais vistos, passando a haver predominância do segundo tipo de restaurante.

Esta visualização de relações que se desenha está mais próxima da topologia do que de um mapa, uma vez que o que se deseja evidenciar não é uma posição geográfica, mas é apenas o estudo da localização em um espaço abstrato. O trabalho de topologia psicológica de Kurt Lewin está baseado no mesmo princípio: não se trata de saber o que está mais ao sul ou mais ao norte, mas a proximidade e a interrelação entre os pontos, arranjo do qual se extrai um sentido. A análise se centra no conjunto, mais em um sistema do que em uma taxonomia.

2.1. De árvores a rizomas

A taxonomia, enquanto método de organização de informações, tem origem no universalismo de Platão e de Aristóteles:

[...] para o primeiro, o universal e o particular eram conceitos separados, sendo que o que era particular era apenas um protótipo da forma ideal. Aristóteles encontrava o universal nas coisas particulares, no que ele considerava a essência das coisas. Sua crença no essencialismo – a presença de uma essência imutável em cada objeto - o levou ao desejo de uma ordem absoluta na natureza, de um universo perfeito em que todas as espécies estariam dispostas hierarquicamente em uma escala, da mais alta à mais baixa⁸ (Lima, 2011, pp. 64-65, tradução nossa).

⁸ “Platon comme Aristote croyaient en l’universalisme (une école de pensée qui croit en des faits ou des propriétés universelles), mais alors que pour Platon l’universel et le particulier étaient deux concepts séparés, le particulier n’étant qu’un prototype ou qu’une forme idéale. Aristote trouvait l’universel dans les choses particulières, dans ce qu’il considérait comme l’essence des choses. Sa croyance en

Lima diz ainda que, durante a Idade Média, esta ideia de ordem absoluta trazida à luz por Aristóteles foi distorcida pelas crenças e ideologias daquele período. Eram criadas hierarquias para justificar o direito divino e o poder absoluto dos monarcas, já que neste sistema eles não estavam submissos a ninguém. Entre os séculos XVII e XVIII, os sistemas taxonômicos, representados por “árvores”, eram encontrados no domínio da Ciência - por exemplo, nas classificações biológicas ou nas enciclopédias, como vemos nas figuras seguintes:

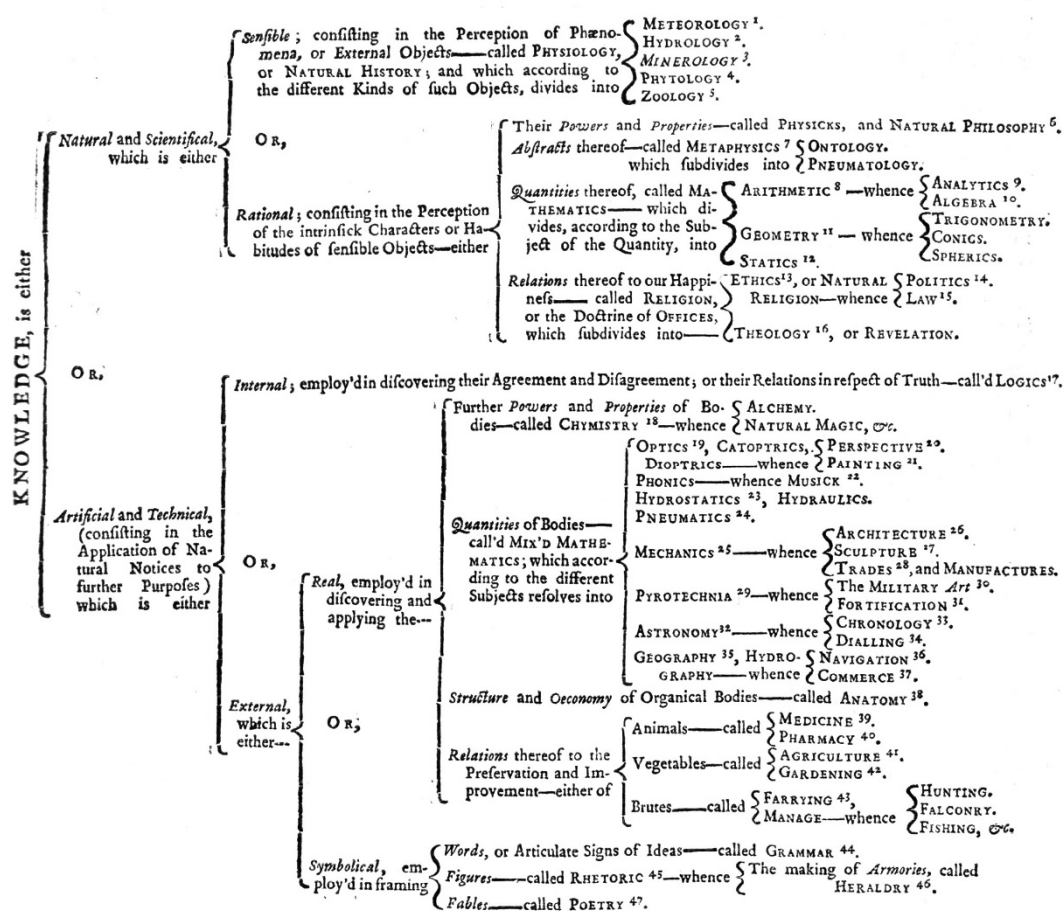


Figura 5 - Classificação do conhecimento de acordo com Ephraim Chambers, apud Lima (2011, p. 37), em Cyclopaedia, um dicionário universal das artes e da ciência

l'essentialisme – la présence d'une essence immuable dans chaque objet – l'a mené au désir d'un ordre absolu dans la nature, d'un univers parfait où toutes les espèces seraient disposées hiérarchiquement dans une échelle, de la plus basse à la plus haute” (Lima, 2011, pp. 64-65).

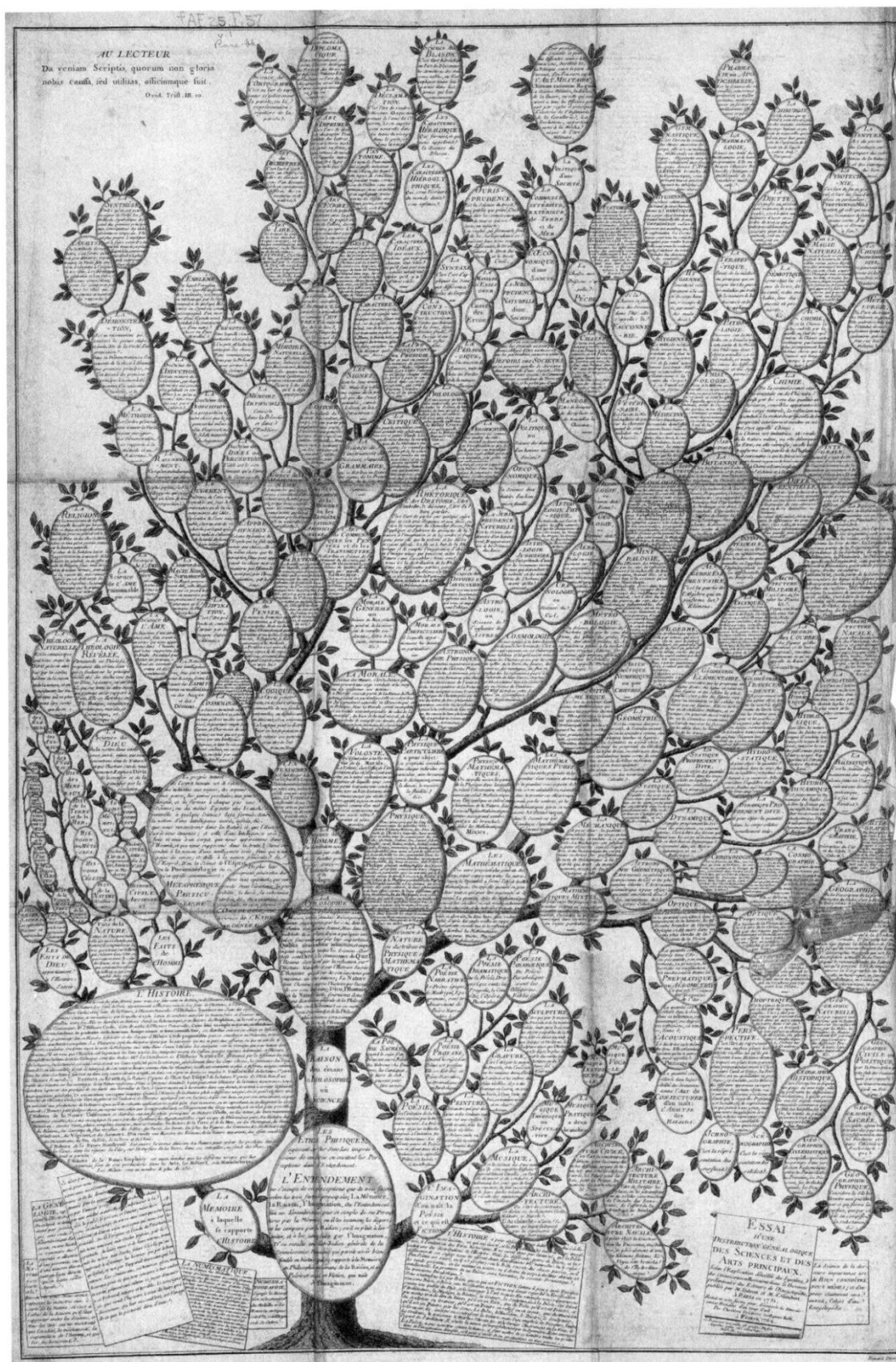


Figura 6 - "Ensaio de uma distribuição genealógica das ciências e das principais artes", de Chrétien Frédéric Guillaume Roth (1769), baseado na Enciclopédia de Diderot e Alembert em Lima (2011, p. 40).



Figura 7 - Detalhe da figura anterior (Lima, 2011, p. 40)

As arborescências evocam um modelo rígido, hierárquico, autoritário, unidirecional, estagnante. Gilles Deleuze e Félix Guattari (1980 apud Lima, 2011, p. 44) dizem que a representação em árvores impõe uma rigidez teórica e de progresso unidirecional “no qual tudo leva a um tronco central por conexões lineares e verticais”,

que, deste modo, nunca aceitariam a multiplicidade. Assim, propõem o conceito de rizoma, cujo objetivo é de contemplar a multiplicidade e a multilinearidade.

De acordo com Deleuze e Guattari (1995, p. 32), “diferentemente das árvores ou de suas raízes, o rizoma conecta um ponto qualquer com outro ponto qualquer e cada um de seus traços não remete necessariamente a traços de uma mesma natureza. Ele põe em jogo regimes de signos muito diferentes, inclusive estados de não-signos. [...] Ele não tem começo, nem fim, mas sempre um meio pelo qual ele cresce e transborda. [...] O rizoma é uma antogenealogia (Ferrari, 2007, p.21).

Os problemas da contemporaneidade precisam ser analisados pelo espectro do rizoma porque têm alta interdependência, com variáveis fortemente conectadas.

O conceito de rizoma, proposto por Guattari e Deleuze, analisa a proposta da biologia, enquanto formações anexas e extensas das raízes das árvores, que ora se comunicam, ora se afastam por variadas direções, podendo ser utilizado também pelas Ciências Humanas para analisar as relações geradas pelas narrativas colaborativas no ciberespaço (Idem).

Weaver (1948 apud Lima, 2011, p. 45) afirma que os problemas da ciência moderna que apareceram entre os séculos XVII e XIX são chamados de “problemas de simplicidade”, envolvendo +a influência de uma variável sobre a outra. Em seguida, na primeira metade do século XX, os “problemas de complexidade desorganizada” apresentavam variáveis que interagiam de modo quase aleatório. Weaver chama os problemas surgidos a partir da segunda metade do século XX de “complexidade organizada”, nos quais o foco é a análise das interdependências.



Figura 8 - Tipologia dos problemas da ciência entre os séculos XVII e XX. (Lima, 2011)

Em referência ao último estágio, Weaver escreve:

Estes problemas (como a flutuação do preço dos produtos, a estabilização das taxas de câmbio, as estratégias militares, ou os **modelos comportamentais de grupos sociais**) – e um grande número de problemas similares da biologia, medicina, psicologia, economia e ciências políticas – são complicados demais para poderem ser resolvidos por uma simples aplicação das técnicas do século dezanove, mesmo se elas tiveram um sucesso espetacular com problemas de simplicidade de duas, três ou quatro variáveis... estes novos problemas, e o futuro do mundo, são muito interdependentes e exigem da ciência que ela faça um terceiro grande salto, ainda maior que aquele que permitiu resolver os problemas de simplicidade do século dezanove, ou que a vitória sobre os problemas de complexidade desorganizada no século vinte. A ciência deve aprender, nos próximos cinquenta anos, a tratar dos problemas de complexidade organizada (Weaver, 1948 apud Lima, 2011, p. 45, grifo nosso, tradução nossa)⁹.

A figura do rizoma reaparece quando Paul Baran sugeriu, ao governo estadunidense, topologias de redes de comunicação que fossem resistentes a ataques militares. A topologia rizomática, descentralizada, como representada na figura a seguir, foi adotada posteriormente para a construção da camada física da *Web*.

⁹ “Ces problèmes (comme la fluctuation du prix des produits, la stabilisation des cours des devises, les stratégies militaires, ou les modèles comportementaux des groupes sociaux) – et un grand nombre de problèmes similaires de biologie, médecine, psychologie, économie et sciences politiques – sont trop compliqués pour pouvoir être résolus par la simple application des techniques du dix-neuvième siècle, même si elles avaient un succès spectaculaire avec les problèmes de simplicité à deux, trois, ou quatre variables ... Ces nouveaux problèmes, et l’avenir du monde, dépendent beaucoup d’entre eux, exigent de la science qu’elle fasse un troisième grand saut, encore plus grand que celui qui a permis de résoudre les problèmes de simplicité au dix-neuvième siècle, ou que la victoire sur les problèmes de complexité désorganisée au vingtième siècle. La science doit apprendre, sur les 50 prochaines années, à s’occuper de ces problèmes de complexité organisée” (Weaver, 1948 apud Lima, 2011, p. 45).

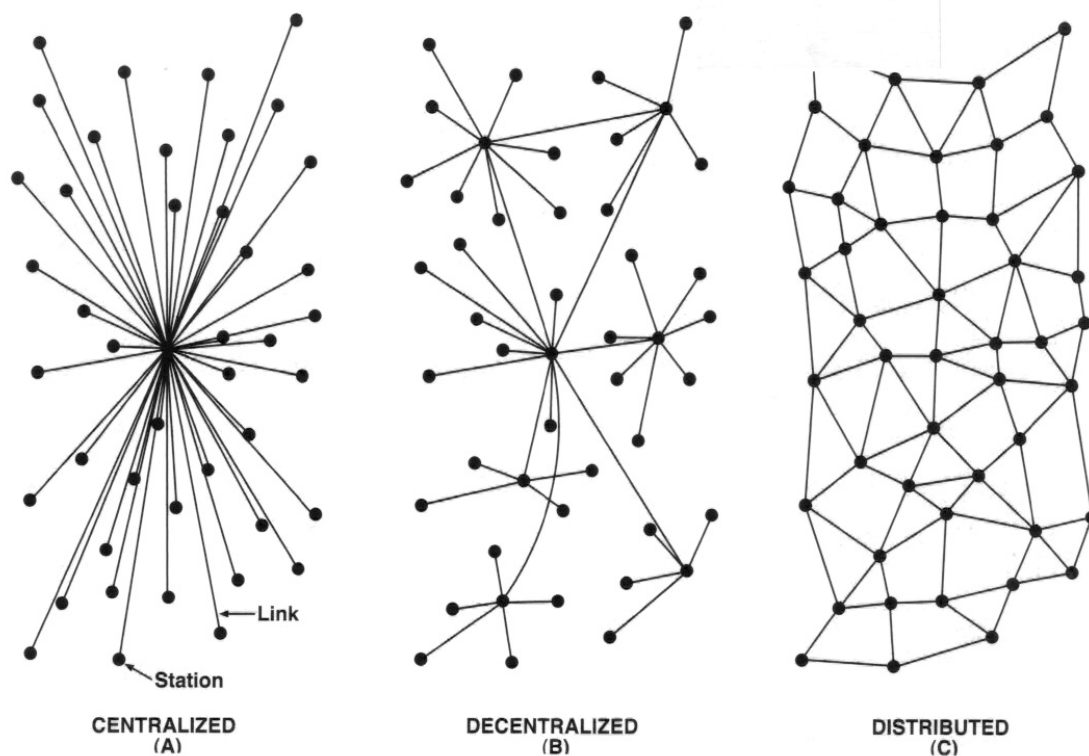


Figura 9 - Propostas de topologias de rede feitas por Paul Baran, 1964 (Lima, 2011, p. 55)

À medida em que as representações de redes que se assemelham a rizomas foram ganhando campo, resolvendo problemas contemporâneos, novos gêneros de representação de dados complexos surgiram e Lima (2013, p. 158) apontou o nascimento de uma sintaxe das linguagens gráficas e propôs uma tipologia de tais gêneros. São a) diagramas em arco; b) agrupamento de superfícies; c) explosão central; d) anel central; e) globo circundado; f) *links* circulares; g) implosão elíptica; h) gráfico em fluxo; i) rizoma orgânico; j) convergência radial; k) implosão radial; l) ramificação; m) círculos em escala; n) convergência radial segmentada; o) esfera.

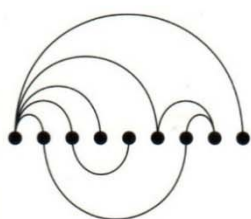
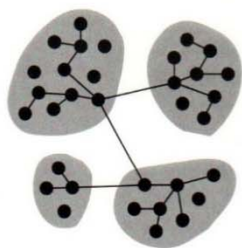
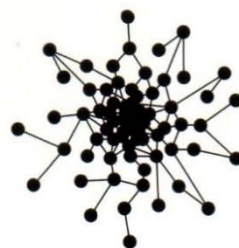


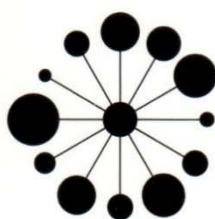
Diagramme en arc



Groupement de surfaces



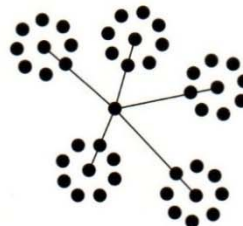
Explosion centrale



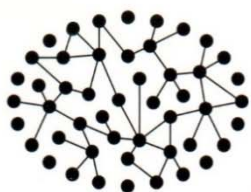
Anneau central



Globe encerclé



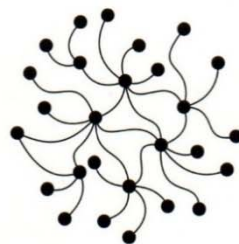
Liens circulaires



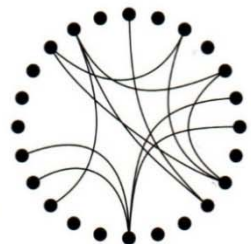
Implosion elliptique



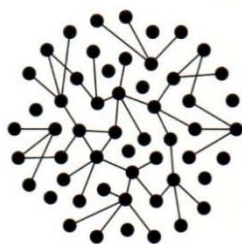
Graphique en flux



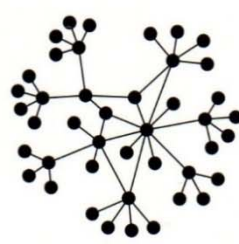
Rhizome organique



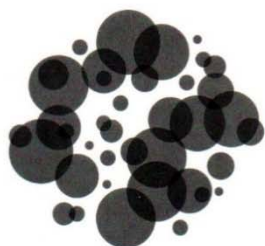
Convergence radiale



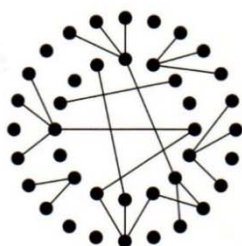
Implosion radiale



Ramification



Cercles à l'échelle



Convergence radiale segmentée



Sphère

Figura 10 - Tipologia de representações de gráficos sobre complexidade, de acordo com Lima, 2013.

Por meio desta linguagem gráfica de representação de dados complexos, como a abstração de relações em uma topologia, semelhante a um rizoma, é possível analisar, por *insights* sobre o gráfico resultante, um grande volume de dados sobre relações humanas, mais complexos do que aqueles elaborados por Moreno na primeira metade do século XX.

2.2. Mapas sociométricos de Jacob Levy Moreno

Elétrons têm o mesmo peso e a mesma carga de eletricidade quando isolados, porém, se estiverem ligados uns aos outros produzindo um átomo, começam a demonstrar individualidade. O mesmo se dá com os homens que, quando ligados entre si formando um grupo, exibem “diferenças” individuais que pareciam inexistir, anteriormente.
(Moreno, 1994, p. 119)

Jacob Moreno é um dos principais nomes na área da Psicologia, por ser considerado o pai da técnica terapêutica chamada de psicodrama. Ele fundou três linhas de pesquisa: a sociometria, o psicodrama e, posteriormente, a psicoterapia de grupo. Partindo da psicoterapia de grupo, procurou criar uma nova ciência, a sociatria, que de acordo com Marineau (1992, p.119), lidaria com “a profilaxia, diagnóstico e tratamento de grupos e relações intergrupais” e “teria sido uma nova e fascinante área de pesquisa”. Segundo ele, tratava-se de “um conceito que incluía tanto prevenção como terapia e teria sido uma extensão mais ampla do que a terapia de grupo”. Isto quer dizer que a prática da psicoterapia de grupo, delimitada a um número pequeno de participantes, poderia tratar um número maior de indivíduos por vez e seria chamada de sociatria.

Ao estudar medicina em 1917 em Viena, teve contato com a teoria freudiana, participando de suas conferências. Lima (2011) afirma que Moreno foi um dos primeiros a duvidar do método de Freud.

Enquanto Freud preferia consultar as pessoas individualmente, no ambiente artificial do seu consultório, Moreno acreditava no poder da terapia em grupo, que poderia ser exercida apenas no ambiente natural dos pacientes – a rua, o parque a comunidade (idem, p. 75).

Marineau afirma que Moreno “estava mais interessado nos processos conscientes, o aqui e agora, na criatividade da pessoa” do que no estudo dos processos inconscientes. Procurando se especializar, partiu para Nova Iorque no começo dos anos 1930 e, durante sua estadia, conseguiu autorização para fazer entrevistas e aplicar um questionário com cerca de 30 variáveis junto aos prisioneiros de Sing-Sing. Marineau (1992, p.120) relata que, com essas ferramentas, Moreno pôde predizer “como dois indivíduos podiam ou não se aproveitar reciprocamente da sua presença no mesmo grupo”. Marineau ressalta ainda que Moreno possuía um senso afinado de observação e que isto foi levado em conta na aplicação dos testes. “Dessa análise”, afirma Marineau, “Moreno deduziu um quociente social e por fim recomendou uma nova classificação¹⁰ dos prisioneiros” para melhorar a comunidade prisional.

Após apresentar seus resultados em 1932 na Associação Psiquiátrica Americana, foi convidado a aplicar seus testes na Escola para a Educação de Moças em Hudson, estado de Nova Iorque. Estudando as preferências de cada uma das alunas da escola (por exemplo, em termos de companhia para dividir o quarto, os jogos e etc.), elaborou a primeira visualização de dados de relações, que chamou de sociogramas ou diagramas sociométricos. Apesar de ter orgulho das visualizações que mostravam “as zonas de afinidade e repulsa entre os indivíduos” porque esses “modelos essencialmente matemáticos eram fáceis de ler e interpretar” (Marineau, 1992, p. 123), e apesar do uso extensivo de tabelas para anotar os resultados e testar fórmulas, os sociogramas não eram exatamente fiéis aos dados, já que não havia sido usado um algoritmo matemático para resolver a posição de cada membro do grupo estudado.

Após a coleta e tabulação das escolhas, comece o sociograma pelas pessoas mais escolhidas, coloque-as em suas formações naturais – três pessoas em um triângulo, quatro em um quadrado, cinco em um pentágono, etc., bem separadas, no papel. A existência de subgrupos deve ser observável no desenho (Moreno, 1994, p. 26).

¹⁰ Talvez a palavra “classificação” tenha sido empregada para traduzir a palavra francesa *classement* para o português como “organização”.

Ainda assim, Moreno conseguiu observar que, como fruto do acaso, surgiam padrões de organização de grupos enquanto construía suas visualizações. Os mais relevantes deles são mostrados na figura seguinte.

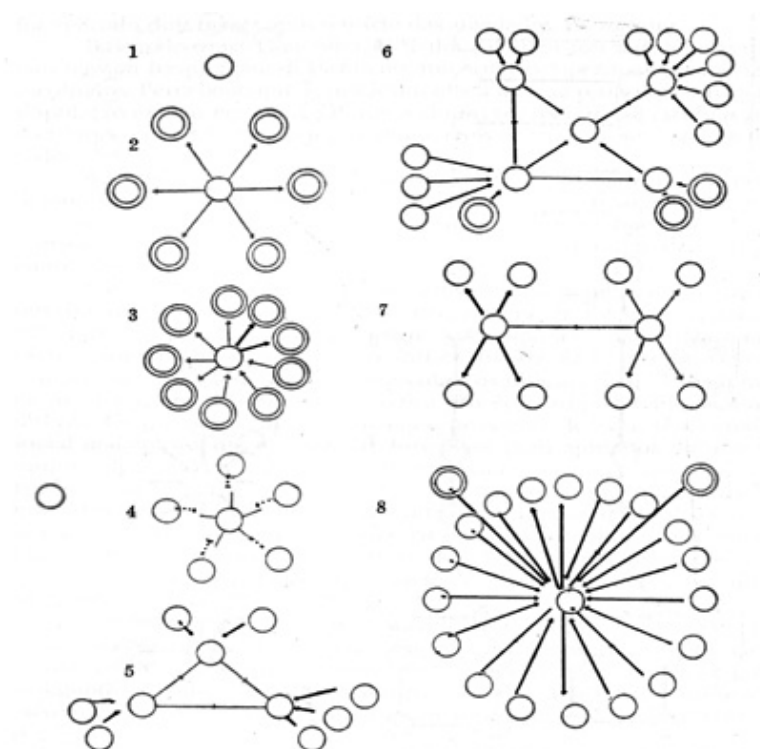


Figura 11 - Diagrama com estruturas típicas de grupos elaborado por Moreno (1994, p. 24).

As estruturas típicas mostradas na figura anterior são assim descritas:

1. Total isolamento – não há linhas de atração e repulsa ligando o sujeito a qualquer outro indivíduo; 2. O sujeito sente atração por seis indivíduos fora de seu grupo (indivíduos de fora são simbolizados pelo círculo duplo), não sendo recíproca esta atração; 3. O sujeito sente atração por quatro indivíduos de fora do seu grupo e rejeita outros dois, sem que a recíproca seja verdadeira. Há três pessoas, de fora do grupo, que sentem atração pelo sujeito, sem que haja reciprocidade; 4. O sujeito sente atração por cinco indivíduos de seu grupo que respondem com indiferença; 5. Atrações mútuas entre três indivíduos tomam forma de um triângulo isolado e rejeitado; 6. Cinco sujeitos, cada qual isolado e rejeitado em seu próprio grupo, rejeitam e isolam um ao outro; 7. Dois sujeitos, isolados em seus próprios grupos, formam um par de atrações mútuas. O resultado é um par isolado; 8. O sujeito rejeita seis indivíduos e é rejeitado por quinze em seu próprio grupo e por dois de fora. O resultado é um indivíduo isolado e rejeitado (Moreno, 1994, p. 24).

A gênese desses padrões pode estar atrelada à combinação do esquema mental proveniente das observações de campo com os resultados de testes. Para utilizar as estruturas de grupo que Moreno construiu nos dias de hoje, seria necessário criar um algoritmo de distribuição de pontos que os arranjassem de acordo com o esquema mental de quem elaborou a visualização de dados. As estruturas ditas típicas parecem depender enormemente de características da dinâmica do grupo em que estão inscritas. Estima-se que um catálogo completo e definitivo de estruturas, como um dicionário de relações, seja uma tarefa gigantesca, senão impossível, já que seria preciso levar em conta, por exemplo, o número e o tipo de membros do grupo, as situações em que se encontram etc.

O psicólogo romeno também elenca os tipos de diagramas sociométricos que podem ser construídos partindo-se das técnicas de construção que elaborou.

2.3. Tipos de diagramas sociométricos

Para explorar e medir as relações interpessoais e interações pessoais, Moreno (idem, pp. 25-27) propõe os seguintes tipos de diagramas sociométricos, que são seis: a) diagramas de interação espontânea; b) diagrama de familiaridade; c) sociogramas; d) sociomatriz; e) diagramas de papéis; f) locogramas.

Os chamados “diagramas de interação espontânea” representam a própria vida. Segundo Moreno, o objetivo é representar simbolicamente, em forma de um diagrama, “todas as variáveis possíveis de situação de vida – tempo, espaço, número e tipo de pessoas presentes, atos e pausas, iniciativa (iniciar uma cena), simultaneidade de aparecimento, liderança, modificação da liderança e término (de uma cena)”. A interação é classificada como espontânea porque é determinada pelos interatores, que Moreno chama de “parceiros”. O tempo de uma interação pode ser acelerado ou retardado por condições externas. A pesquisa não encontrou este tipo de diagrama na consulta à obra.

O “**diagrama de familiaridade**” indica a familiaridade entre os indivíduos e evidencia os contatos sociais pré-existentes entre os participantes antes da sessão de análise.

Os “**sociogramas**”, em contraste com os diagramas, são mais abstratos. Não constam símbolos de tempo e espaço, mas são mostrados apenas os “relacionamentos entre todas as pessoas interagindo entre si”. É um universo atemporal e representa apenas o “espaço social”.

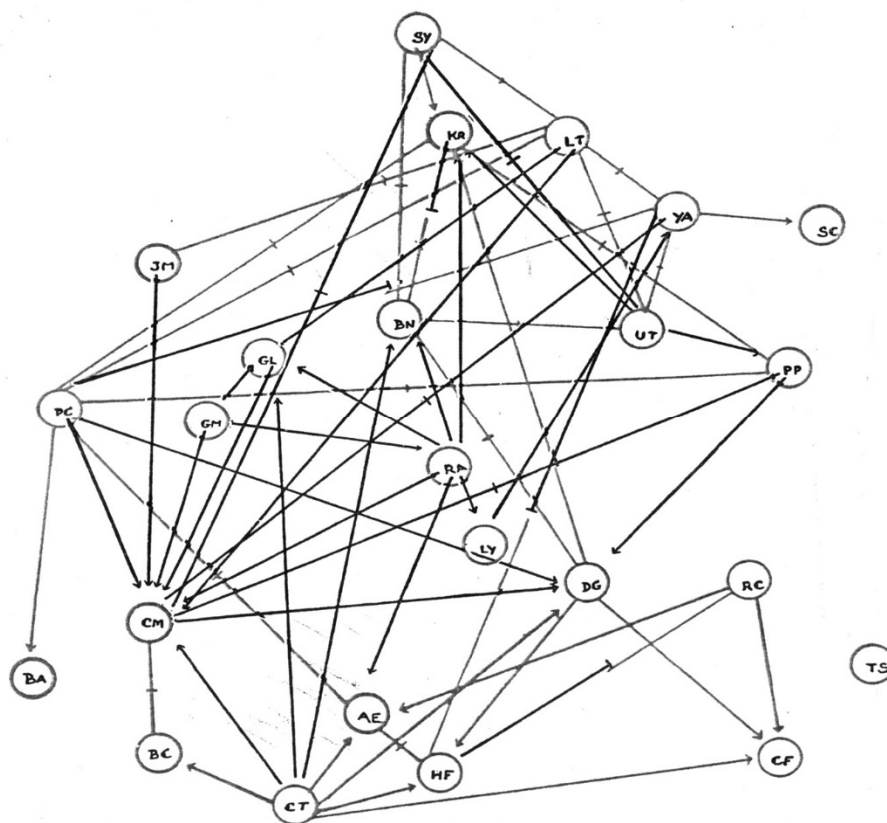


Figura 12 - Exemplo de sociograma em Moreno (1994, p.35).

A “**sociomatriz**” trata da representação numérica de pessoas, ou seja, uma abstração dos processos reais. Moreno esclarece que esta modalidade de representação é mais abstrata que o sociograma e, nela, se inscrevem “todas as pessoas que participam da situação, porém, aqui, eliminam-se não apenas o tempo e o espaço, mas também, os sinais de contato entre elas. ”

	Aceitações e Rejeições dadas																													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	
1:	+	+	0	0	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	+	
2:	+	0	0	+	0	+	+	0	0	+	0	0	-	0	0	-	+	0	0	+	0	+	0	0	+	+	0	0	0	0
3:	+	+	+	0	0	0	0	0	0	0	0	+	0	+	0	+	+	0	+	+	+	0	0	0	+	+	0	0	0	0
4:	+	+	0	0	0	0	0	0	-	0	0	0	0	+	0	+	+	0	+	+	0	0	0	0	0	0	0	0	0	0
5:	+	0	0	+	+	+	0	0	-	0	0	+	+	+	0	+	+	0	+	+	0	0	-	0	0	0	0	-	0	0
6:	0	0	0	0	0	0	0	0	0	0	+	+	+	+	0	+	+	0	0	0	0	0	0	0	0	0	0	0	0	0
7:	+	+	0	0	0	0	0	0	0	0	+	0	0	0	0	-	+	0	0	0	0	0	0	0	0	0	0	0	0	0
8:	0	0	0	0	0	0	+	-	0	0	+	+	+	+	0	+	0	0	0	0	0	0	-	+	+	0	0	0	0	+
9:	+	+	0	0	-	0	0	+	+	0	-	+	+	+	0	+	+	-	+	+	0	-	-	-	+	+	+	-	0	-
10:	+	+	+	-	-	+	+	+	+	0	+	+	0	+	+	+	+	0	-	+	+	+	+	+	+	+	0	0	0	0
11:	+	+	0	-	0	+	+	+	+	0	0	+	0	+	+	+	+	0	-	-	0	-	-	+	+	+	+	0	0	0
12:	0	0	-	-	-	-	-	-	0	0	0	+	+	+	0	+	-	-	-	0	0	-	-	+	+	+	+	-	+	+
13:	+	+	0	+	+	+	+	0	0	-	0	+	+	+	0	+	+	0	+	+	+	+	+	+	0	0	0	0	0	0
14:	+	0	0	0	0	+	+	0	0	0	0	+	+	+	0	+	+	0	+	+	+	+	+	+	0	0	0	0	0	0
15:	0	+	0	0	0	0	0	0	0	0	0	+	+	+	+	0	+	0	+	+	+	+	+	+	+	+	-	0	0	0
16:	+	0	0	0	+	+	+	0	0	0	0	+	+	+	0	+	0	+	+	+	0	0	0	0	+	+	+	0	0	0
17:	0	+	-	+	0	+	+	0	0	0	0	0	-	0	0	+	0	+	+	0	0	0	-	0	+	+	+	-	0	-
18:	0	0	0	0	0	+	+	0	0	0	0	0	-	0	0	+	0	0	0	0	0	0	0	0	0	0	0	-	0	0
19:	0	0	-	-	0	0	0	-	-	-	0	0	0	0	0	+	0	0	-	0	0	0	0	0	0	0	-	-	0	0
20:	+	0	0	+	0	0	0	0	0	0	+	0	0	0	+	+	+	+	+	+	0	-	0	+	+	+	+	-	0	0
21:	+	+	+	+	0	+	+	0	0	0	0	0	0	0	0	+	+	+	+	0	0	0	0	0	0	0	0	0	0	0
22:	+	+	+	+	-	0	0	0	0	0	0	0	0	+	0	+	+	+	+	0	0	+	+	+	+	+	-	-	0	0
23:	0	0	0	0	0	0	+	+	0	0	0	0	0	+	0	+	0	0	0	+	0	0	0	0	0	0	0	0	0	0
24:	+	+	+	0	+	+	+	-	-	+	+	-	-	+	-	+	+	+	+	+	-	-	-	+	+	+	-	-	0	-
25:	+	+	0	0	-	0	0	-	-	+	-	-	-	+	0	+	+	+	+	0	0	0	0	0	0	0	0	0	0	0
26:	+	+	0	0	+	+	0	-	-	0	+	0	-	+	+	+	+	+	+	-	0	0	0	0	0	0	0	0	0	0
27:	0	+	0	+	+	+	+	+	-	+	0	+	+	+	0	+	0	0	0	0	0	0	0	0	0	0	0	0	0	0
28:	0	0	0	+	0	0	0	0	0	-	+	0	0	0	0	+	0	0	0	0	0	0	0	0	0	0	0	0	0	0
29:	+	0	0	0	-	0	+	-	0	0	+	+	+	0	+	0	+	+	+	+	0	-	+	+	0	+	+	+	+	+
Σ:	19	16	0	10	2	23	3	3	-8	11	16	14	5	23	6	24	12	12	20	9	8	-8	1	25	20	24	-2	-2	10	
SS:	.67	.57	.00	.35	.47	.82	.17	.10	-.32	.39	.57	.50	.17	.82	.21	.85	.42	.42	.71	.32	.28	-.23	.03	.89	.71	.85	-.07	-.07	.35	
PA:	81	87	40	40	18	119	53	24	-21	66	103	90	29	108	25	110	71	61	91	43	53	-1	28	122	101	119	19	23	48	

Figura 13 - Exemplo de sociomatrix em Moreno (1994, p. 36)

Enquanto o “**Diagrama de papéis**” é uma categoria de representação social que “retrata a aglomeração de papéis de indivíduos e a interação entre esses papéis” (p. 27), o “Diagrama de espaço e movimento”, ou “**Locograma**” representa o “espaço social no qual acontece o movimento” e contém o deslocamento dos interatores.

Em resumo, o corpo teórico construído por Jacob Moreno abre caminho para a análise de redes de afinidades pelo pioneirismo na construção de visualizações de relações afetivas que, com o conjunto de ferramentas disponibilizado na época atual – banco de dados, programas de visualização, grande volume de dados (como – trazem *insights* sobre as organizações emergentes da sociedade com o decorrer do tempo.

Antes de retomarmos a temática da visualização de dados e a criação de mapas de afinidade nos dias de hoje, é preciso tratar de um estudo de caso em que Moreno valeu-se da sociometria para dinamizar o grupo com maior número de pessoas, cerca de 600, conhecido como “projeto Houston”.

2.4. O projeto Hudson: sociometria e empatia

Logo depois de testar o método sociométrico na prisão de Sing Sing, obtendo como resultado a melhor convivialidade entre os prisioneiros e assim a redução das tentativas de fuga do confinamento, Moreno lançou-se ao projeto audacioso de aplicar a sociometria em uma comunidade inteira, de 500 a 600 membros, na cidade de Hudson, no estado de Nova Iorque. Esta comunidade era composta de jovens moças vindas de fora da cidade para participar de um treinamento e que viviam ali de forma comunitária em um alojamento temporário de 16 cabanas. Como havia uma divisão racial muito pronunciada nos Estados Unidos nos anos 1930, época em que o experimento foi conduzido, as garotas negras viviam em cabanas separadas das garotas brancas, mas interagiam à vontade nas atividades sociais e dividiam o espaço. Em cada uma dessas cabanas, uma das meninas exercia o papel de responsável ou organizadora.

O primeiro passo foi mapear os papéis dos membros do grupo, distinguindo a “estrutura social” da “organização psicológica”. A diferença entre as duas denominações pode ser exemplificada da seguinte forma: a supervisora do dormitório poderia ter a função psicológica de ser a “queridinha” da encarregada da cabana, mas que sofrer rejeição dos membros do seu grupo. Ou então, o espaço físico em que os membros do grupo estariam organizados poderia ser diferente do “espaço” criado pelas atrações e repulsões vindas das relações cotidianas entre eles.

A organização social da comunidade total tem outro aspecto, sob sua aparência externa. Apesar de estarem em cabanas separadas, há atrações e repulsas entre as meninas brancas e as de cor que afetam seriamente a conduta social nesta comunidade (Moreno, 1994, p. 98).

Uma das intervenções que ele realizou para harmonizar as relações foi na lavanderia. As “correntes emocionais”, como ele se refere aos fluxos de afinidade e empatia, estavam perturbadas com rejeições que interferiam no processo hierárquico, como mostra o *locograma* que se segue.

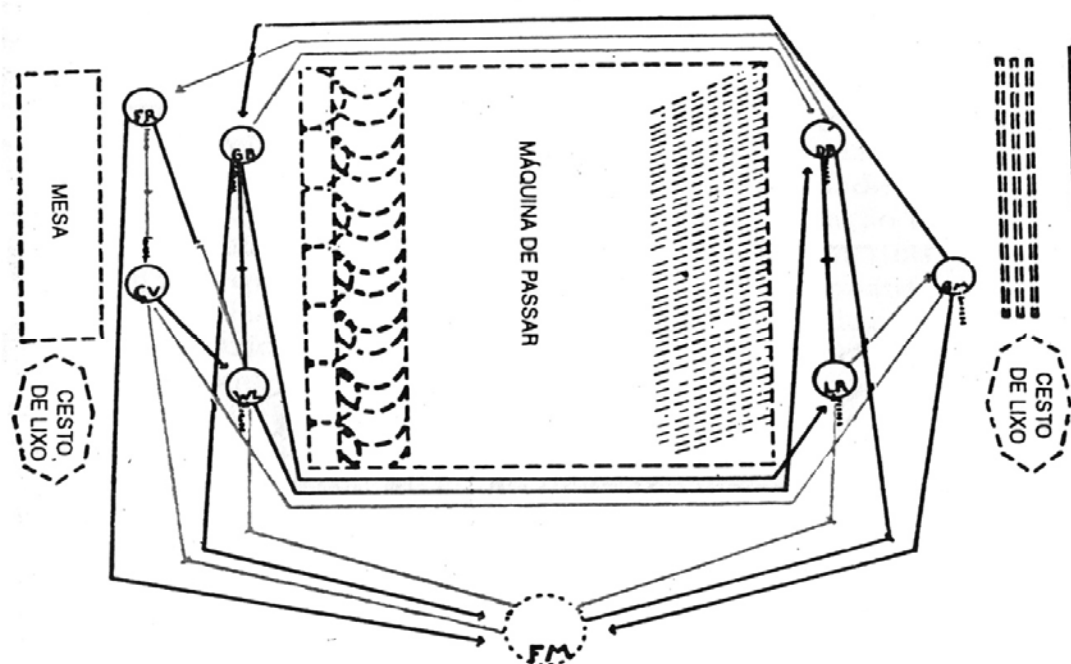


Figura 14 - Lavanderia: estrutura antes da reconstrução (Moreno, 1994, p. 151). O círculo marcado com "FM" representa a responsável pela lavanderia; as flechas mais escuras representam atração e as mais claras, rejeição.

Na figura anterior, as linhas mais fortes indicam atração e as mais claras, rejeição. Mostrado para a comunidade, este locograma fez crescer a tensão racial e Moreno começou a ser culpabilizado pela agitação. Ele se precaveu elaborando sociogramas mentais de sua situação na comunidade, além de estratégias¹¹ para se desviar do conflito antes de entregar ao grupo a solução sociométrica:

Vimos que a lavanderia a vapor era fonte de conflitos raciais que se espalharam pela comunidade. A tarefa era reorganizar o pessoal da lavanderia objetivando melhores relações raciais e maior produtividade. O status sociométrico das meninas que trabalhavam na lavanderia, antes de sua reorganização, era o seguinte: 7 meninas negras e 1 supervisora branca encontravam-se ocupadas com estes afazeres. Stella DR e Philamina LR, alimentadoras, rejeitavam-se; Hilda GR e Myrtle WL, aparadoras, também se rejeitavam. Myrtle WL rejeita, igualmente, a alimentadora que se encontra à sua frente, Philamina LR. Lillian FR e Rosalie CV, as dobradoras, sentem-se mutuamente atraídas. Lillian FR e Rosalie CV rejeitam Myrtle WL. Esther GM, que sacode as roupas, sente atração por Rosalie CV e rejeita Hilda GR. Tanto Esther GM,

¹¹ “Neste dilema”, diz Moreno (1994, p. 98), “inventei uma técnica sociométrica criada para radiografar a minha situação, técnica que mais tarde chamei de 'auto avaliação sociométrica' e de projeção”. Com base na intuição do indivíduo sobre sua posição e papel no grupo e por empatia, “ele saberia dizer se os fluxos de afeição ou simpatia por ele estão aumentando ou diminuindo”.

quanto Stella DR, Hilda GR e Lilian FR rejeitam a supervisora branca (FM), porém, apenas Stella DR é rejeitada por ela. Philamina LR, Myrtle WL e Rosalie CV sentem atração pela supervisora. As sete meninas moram na C12 (apenas Myrtle WL encontra-se no sociograma da C12, pois as outras chegaram à comunidade posteriormente). Myrtle WL é estrela de número considerável de atrações em seu grupo familiar. Entretanto, sua posição sociométrica na lavanderia é de ordem diferente. Ali, vê-se rejeitada por três trabalhadoras; a atração mútua com a supervisora é remanescente da posição no grupo familiar, no qual representa uma das poucas forças estabilizadoras. Lá, também, Myrtle mantém relação afetiva com a figura de autoridade, a encarregada branca. Stella DR é líder de gangue rebelde que produz contracorrente. É rejeitada por Myrtle WL e pela maioria das meninas que sentem atração pela supervisora. **Para suprimir deste grupo a tensão socioemocional que pesava tanto sobre a supervisora quanto sobre as meninas, comprometendo o moral do grupo e interferindo em a sua produtividade, o procedimento indicado era a eliminação tanto de Stella DR quanto de Myrtle WL** (Moreno, 1994, v.1, pp. 74-75, grifo nosso).

Depois desta ação para resolver os problemas de relacionamento do grupo da lavanderia, foi produzido o seguinte sociograma:

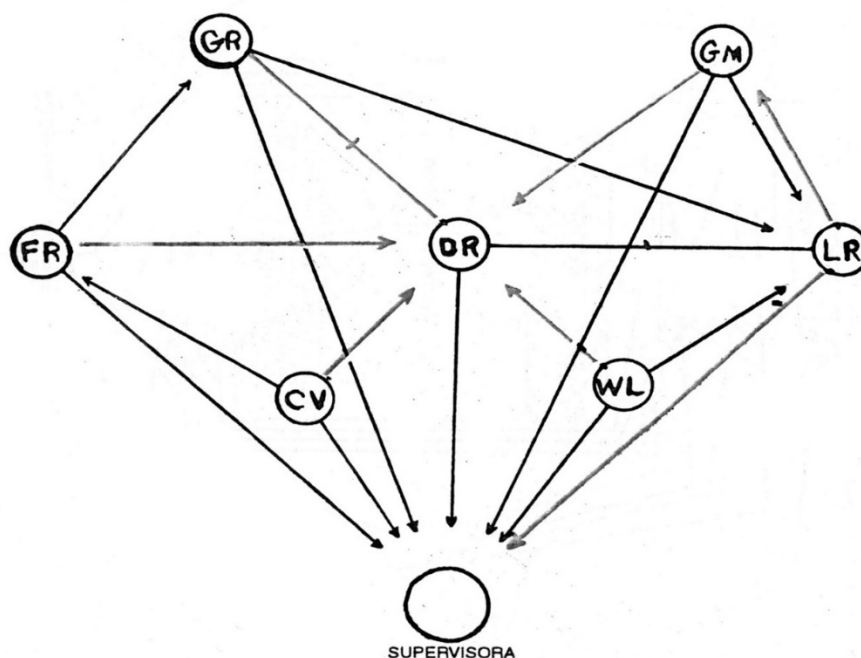


Figura 15 - Lavanderia, após retestagem em Moreno (1994, p. 152).

Duas coisas chamam a atenção no sociograma anterior. A primeira é que ele muito se assemelha a um diagrama hierárquico, quase como um diagrama em árvore, e deve ter sido assim feito para evidenciar um retorno à ordem. Isto pode mostrar que neste e outros sociogramas produzidos por Moreno, como também naquele mostrado na figura 15, o comprimento das flechas de afeição não está relacionado a uma proximidade afetiva ou afinidade – mostram apenas uma relação. Esta foi a preocupação também de Northway (apud Lima, 2013, p. 78), que construiu uma plataforma para prototipar sociogramas levando em conta a afinidade entre os membros de um grupo. São piões distribuídos em círculos concêntricos e ligados por elásticos para projetar visualizações que chamou de “sociogramas-alvo”.

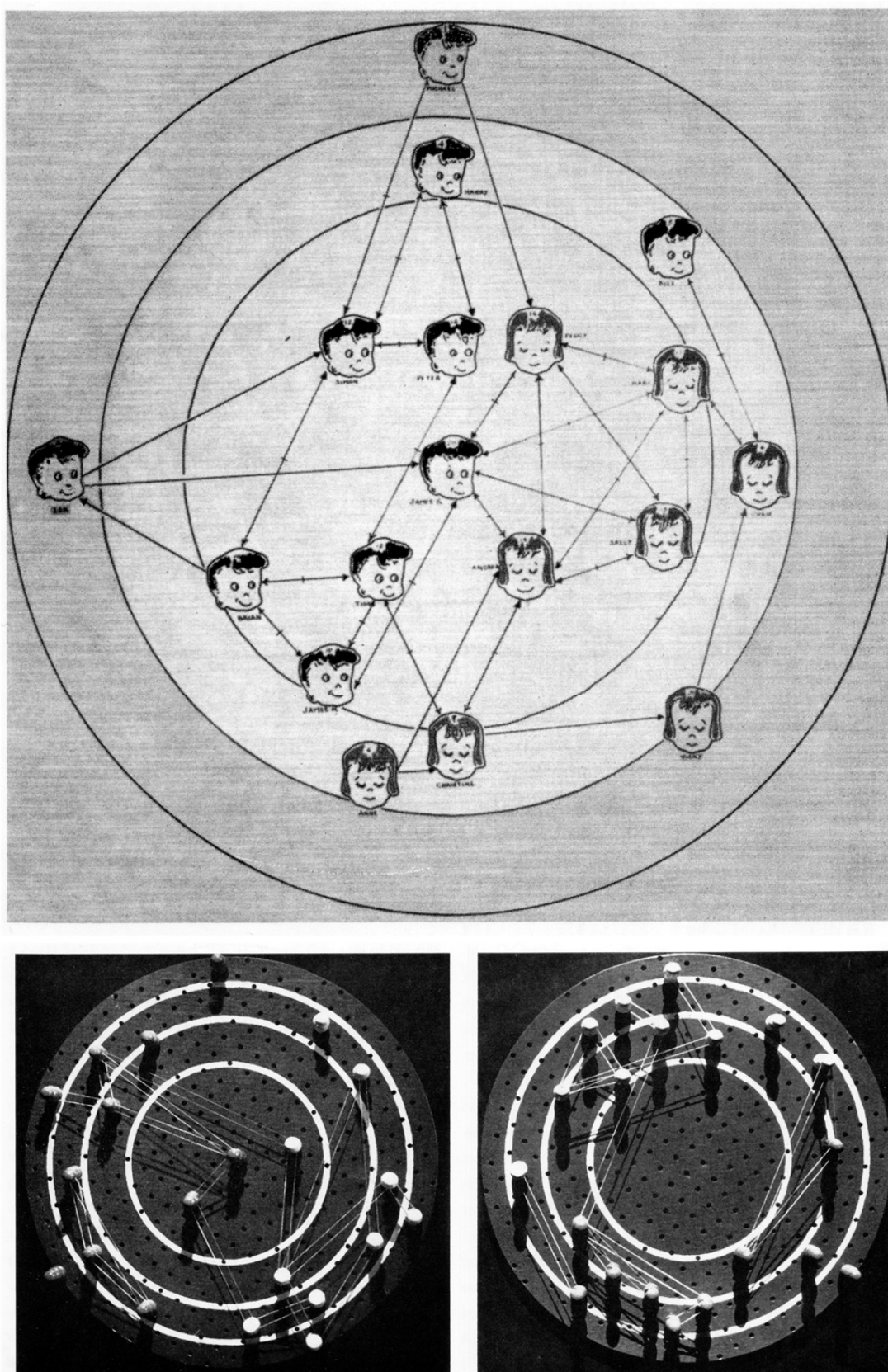


Figura 16 - Sistema de prototipagem de sociogramas-alvo, por Northway (apud Lima, 2013, p. 78).

O ar hierarquizante presente em muitas visualizações produzidas por Moreno e sua equipe, somado ao que encontramos no primeiro volume de “Quem sobreviverá”, mostra que as ideias de Moreno são fundadas na crença de que a sociedade pode ser, digamos, “desenhada”, no sentido do *design*. Nesta sociedade sociometrizada, cada pessoa teria seu lugar em grupos perfeitamente adaptados a seus afetos; por outro lado, como no exemplo da lavanderia, pessoas poderiam ser excluídas ou presas em seus papéis em nome da harmonia do grupo, esta extensível à sociedade. Seria uma harmonia artificial cuja duração não saberíamos precisar, apesar de que, levando em conta o grande número de variáveis para que se sustentasse (como mudanças de humor, de personalidade ou de função no grupo, entre outros), talvez não durasse muito tempo. O psicólogo analisa os resultados da aplicação de seu método e conclui:

Temos conseguido demonstrar que a distribuição e o reagrupamento de indivíduos em dada comunidade, através das técnicas sociométricas, pode servir de ajuda valiosa para melhorar o status sociométrico e o moral de seus habitantes. A explicação desta sociologia de reconstrução em inúmeras comunidades ampliou nossa experiência metódica e forneceu o suporte necessário a nossas hipóteses de que nenhum indivíduo ou grupo deve ser removido de determinada comunidade até que tenhamos tentado consolidar sua integração sociométrica. Podemos deduzir que, ao suplementar o planejamento social legal e econômico com técnicas de planejamento sociométrico, muitos de nossos reformatórios, prisões e hospitais para doentes mentais tornar-se-ão supérfluos (Moreno, 1994, p. 89).

Voltando o foco à visualização das relações, Kurt Lewin, que desenvolvia estudos de psicologia topológica na mesma época das pesquisas de Moreno, ressalta que:

A determinação das relações topológicas é a tarefa fundamental em todos os problemas psicológicos. Mudanças nas conexões são as mudanças mais importantes tanto no ambiente psicológico como na estrutura da pessoa. Ao mesmo tempo, as relações topológicas são fundamentais para o lado matemático do nosso problema. A topologia, enquanto ciência mais geral das relações espaciais, pode ser baseada na relação “parte” e “todo” ou, em outras palavras, no conceito de “inclusão”. Relacionados a estes conceitos está o do “entorno” de um “ponto” (Lewin, 1932, p. 87).

Pensamos que a topologia das relações, realizada nos nossos dias, seria útil não só na resolução de problemas psicológicos ou relacionais. Com o aparato técnico a que se tem acesso hoje, a possibilidade de tratar um volume de dados grande e complexo relativo aos comportamentos *online* e *off-line*, e ainda realizar um gráfico resultante dos dados, traria à tona a forma daquilo que sabemos existir, mas não quantificar ou analisar - como os influenciadores, a propagação de *memes* por uma rede social ou de modismos em uma população, o comportamento de consumo e as redes e caminhos de produção de objetos culturais, entre outros. A topologia das relações pode ser instrumentalizada, deste modo, para observar problemas de diversas áreas, investigando desde o sujeito atuando em pequenos grupos até em comunidades, multidões e o ciberespaço.

Deste modo, falaremos mais adiante sobre comunidades e sujeito e sobre a relação topológica que se estabelece na visualização de dados de relações. Antes disso, é importante abordar um fenômeno contemporâneo importante, que é a disponibilidade farta de dados estruturados sobre indivíduos que encontramos *online* e que pode servir de matéria para o estudo de grupos.

2.5. Big Data e visualização de dados

O conceito de *Big Data* emerge do fenômeno social e tecnológico da difusão crescente do acesso à Internet. Não que a Internet seja o marco zero da coleta de dados massivos, pois esta a precede. Com a crescente disponibilidade de dados na rede, surge o interesse de analisar essas grandes quantidades de dados: de informações do censo, entre outros dados governamentais, até traços de navegação dos utilizadores da Internet gravados nos servidores *online*.

Ao acessar um site na Internet, é usual que o computador remoto que atua como servidor das páginas e dos dados do site consultado archive as informações das transações de navegação trocadas com o computador local. A este registro dá-se o nome de *log*, no jargão da informática.


```

SCOTTGUWEBSITE%2F2013%2F09%2F03%2F09%2FRD00155D51D428.log - Notepad
File Edit Format View Help
# date time s-sitename cs-method cs-uri-stem cs-uri-query s-port cs-username c-ip cs(User-Agent) cs(Cookie) cs(Referer) cs-host sc-status sc-
2013-09-03 09:10:02 SCOTTGUWEBSITE GET /blogposts/questionquestion/step2.jpg X-ARR-LOG-ID=74c4d83c-64f8-4849-b5f1-e8b909a3b885 80 - 80.75.72.
2013-09-03 09:10:02 SCOTTGUWEBSITE GET /blogposts/questionquestion/step3.jpg X-ARR-LOG-ID=546d90ce-4ad3-455e-ae81-f46670579c1d 80 - 80.75.72.
2013-09-03 09:10:02 SCOTTGUWEBSITE GET /blogposts/questionquestion/step0.jpg X-ARR-LOG-ID=f23f2055-116a-4fed-a965-14a0cf6d2ce0 80 - 80.75.72.
2013-09-03 09:10:02 SCOTTGUWEBSITE GET /blogposts/questionquestion/step1.jpg X-ARR-LOG-ID=c0ee55df-4f10-4603-b505-f825cdc44246 80 - 80.75.72.
2013-09-03 09:10:02 SCOTTGUWEBSITE GET /blogposts/questionquestion/step6.jpg X-ARR-LOG-ID=f5203379-62b8-47ed-9d3b-0202f6eb7c72 80 - 80.75.72.
2013-09-03 09:10:02 SCOTTGUWEBSITE GET /blogposts/questionquestion/step5.jpg X-ARR-LOG-ID=9c807d85-8f95-4a7f-8c1b-b7380d46e693 80 - 80.75.72.
2013-09-03 09:10:02 SCOTTGUWEBSITE GET /blogposts/urlrewrite/step1.jpg X-ARR-LOG-ID=9fb53db1-75ff-4291-b3b8-6e65a1737c03 80 - 173.249.65.226
2013-09-03 09:10:02 SCOTTGUWEBSITE GET /blogposts/questionquestion/step4.jpg X-ARR-LOG-ID=f7bffc87-241f-435a-a397-97afe4b9cf75 80 - 80.75.72.
2013-09-03 09:10:02 SCOTTGUWEBSITE GET /blogposts/questionquestion/step9.jpg X-ARR-LOG-ID=63d10461-15a1-4a8b-8552-12d4b375deb8 80 - 80.75.72.
2013-09-03 09:10:02 SCOTTGUWEBSITE GET /blogposts/questionquestion/step8.jpg X-ARR-LOG-ID=6c7bf850-e6ae-499d-bf11-a481baaaf509 80 - 80.75.72.
2013-09-03 09:10:02 SCOTTGUWEBSITE GET /blogposts/questionquestion/step10.jpg X-ARR-LOG-ID=27a33e21-2e33-484b-a67c-0f66c6a2aefc 80 - 80.75.72.
2013-09-03 09:10:02 SCOTTGUWEBSITE GET /blogposts/questionquestion/step11.jpg X-ARR-LOG-ID=7fe439ed-6530-4d7f-8620-08bdeb39fc9 80 - 80.75.72.
2013-09-03 09:10:02 SCOTTGUWEBSITE GET /blogposts/questionquestion/step12.jpg X-ARR-LOG-ID=09866e37-f5b8-4dbc-8b58-68dc68877f7b 80 - 80.75.72.
2013-09-03 09:10:34 SCOTTGUWEBSITE GET /BlogPosts/SqlRegistration/mmc.png X-ARR-LOG-ID=0ca48bb5-0de7-4394-aa3a-e9eb564e781d 80 - 172.19.194.3
2013-09-03 09:10:34 SCOTTGUWEBSITE GET /BlogPosts/SqlRegistration/step1.png X-ARR-LOG-ID=d4196869-97f0-4300-8081-8c2ab1080e3d 80 - 172.19.194.
2013-09-03 09:10:34 SCOTTGUWEBSITE GET /BlogPosts/SqlRegistration/step5.png X-ARR-LOG-ID=ddcb2723-67c5-4be6-9055-f1abc41adcef 80 - 172.19.194.
2013-09-03 09:10:34 SCOTTGUWEBSITE GET /BlogPosts/SqlRegistration/step3.png X-ARR-LOG-ID=8b2d2837-099f-4f7b-8c56-59f6256798db 80 - 172.19.194.
2013-09-03 09:10:34 SCOTTGUWEBSITE GET /BlogPosts/SqlRegistration/step4.png X-ARR-LOG-ID=bfd9d22-e5dd-4524-9775-e30ba943ef78 80 - 172.19.194.
2013-09-03 09:10:34 SCOTTGUWEBSITE GET /BlogPosts/SqlRegistration/step2.png X-ARR-LOG-ID=8f328fe6-5b79-4ade-8032-9a82be007d1e 80 - 172.19.194.

```

Figura 17 - Exemplo de arquivo de log do protocolo HTTP.

O arquivo de *log* é composto de informações ordenadas, ou estruturadas, contendo data de acesso, página acessada e endereço do computador remoto que efetuou o pedido de visualização de página. Dessa forma, é possível descobrir qual computador remoto acessou quais páginas e em que sequência, além da origem geográfica dos acessos. Isto é possível pelo agrupamento de informações. Levando em conta o número de pessoas que acessam a Internet atualmente – cerca de 2.925.249.355 usuários¹² em 2014 – e os 171.359.243 sites ativos¹³ em agosto de 2015, há uma disponibilidade de dados estruturados para serem analisados nunca antes vista.

A consulta dos *logs* de acesso de um determinado site é útil nas áreas de *marketing*, editorial e otimização de conteúdo, já que revela a forma de utilização de um *site* ou serviço *online*: números de audiência, proveniência e idioma dos visitantes, conteúdos acessados, sistema operacional e navegador utilizados. Cruzando as informações do *log* de acesso com as informações nos *cookies* no navegador do computador remoto, é possível traçar um perfil do usuário pelos sites frequentemente acessados. Os *cookies* são arquivos que guardam as transações entre o computador servidor e o computador local; são armazenados no computador do usuário. Fazer este cruzamento de dados não é algo tão recente. Análises de acesso ao usuário têm sido executadas para auscultar o comportamento de navegação do usuário praticamente desde a criação da *Web*: uma das

¹² Disponível em <http://www.internetlivestats.com/internet-users/>. Acesso em 26.11.2015.

¹³ Disponível em <http://news.netcraft.com/archives/category/web-server-survey/>. Acesso em 26.11.2015.

primeiras soluções empresariais, um programa chamado *WebTrends*¹⁴, existe desde 1993, quando era focado na análise de *logs* de servidores. Muitos programas desse tipo foram desenvolvidos e as técnicas de análise foram ficando cada vez mais sofisticadas, agora envolvendo o cruzamento de dados do usuário. A correlação entre *logs* e *cookies* para análise, como explicado anteriormente, é hoje um tipo de análise básica, comparado ao que se consegue saber sobre os usuários de um domínio por meio de instrumentos como o *Google Analytics*¹⁵. Se há vinte anos a análise cruzada de dados era possível apenas para empresas com computadores potentes, a análise cruzada de dados pode ser executada em virtualmente qualquer computador doméstico, ponto de vista sustentado também por Manovich (2011 apud Boyd e Crawford, 2012, p. 663)

¹⁴ <https://www.webtrends.com/about-us/>

¹⁵ <http://www.google.com/analytics>

Representação de dados complexos ou de grande volume

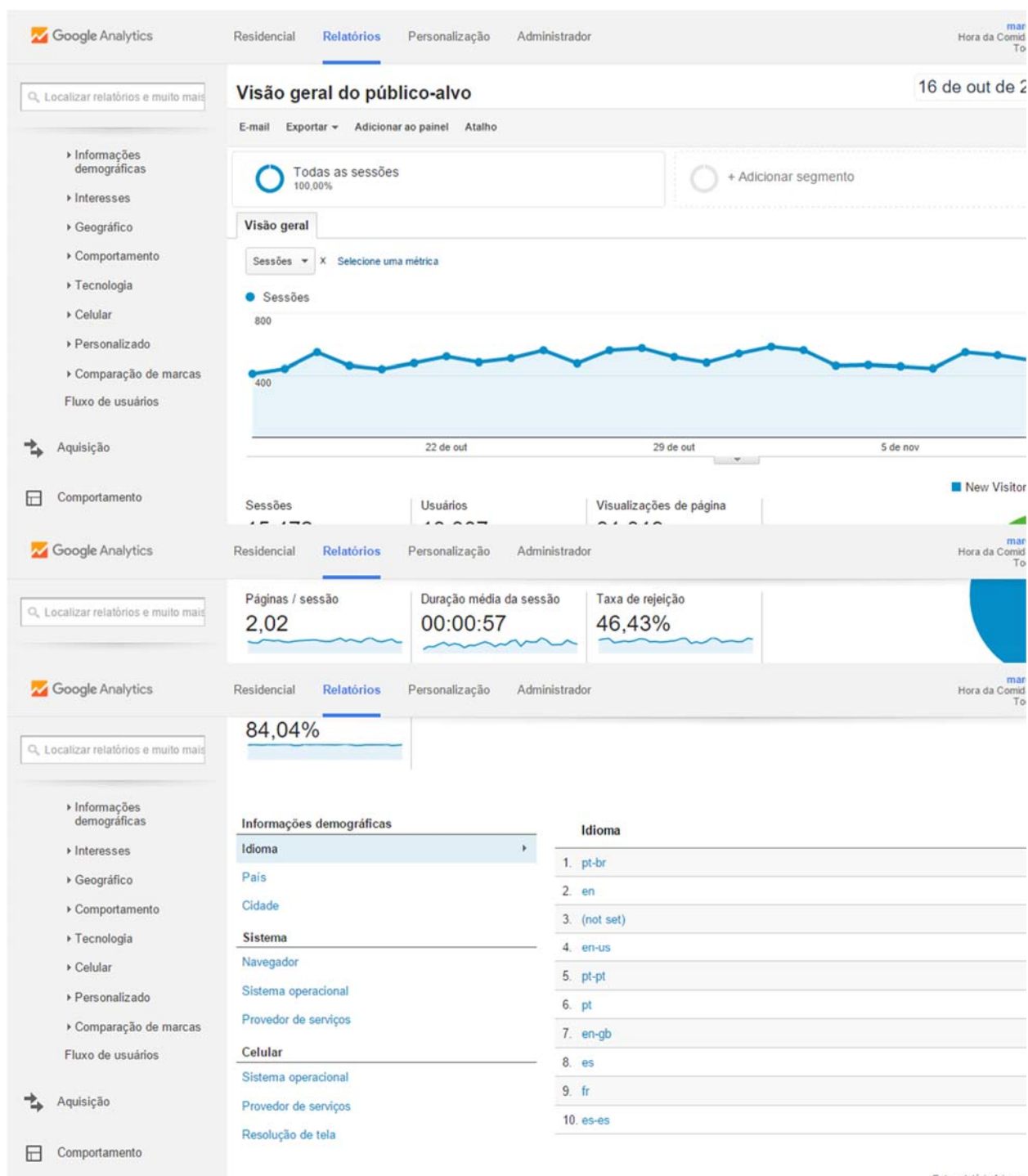


Figura 18 - Captura de tela do Google Analytics, com visualizações dos dados de acesso de um blog

Não são apenas nossos dados de navegação que estão disponíveis e digitalizados. A fatura do cartão de crédito pode revelar hábitos de consumo. *E-mails* indicariam com quem se troca mais informações e os assuntos mais frequentes. O programa de milhas da companhia aérea revela os destinos preferidos dos participantes, enquanto a Polícia

Federal guarda registros de entrada e saída dos passageiros. Para pedir um visto de longa duração para qualquer país, informações biométricas como impressões digitais são colhidas e disponibilizadas em um sistema eletrônico. O cruzamento de todos esses bancos de dados, ao mesmo tempo em que ajudam o profissional de *marketing*, pode infringir a privacidade de todos nós. É comum, no Brasil, depois de se adquirir um novo cartão de crédito, o recebimento de malas-diretas para incentivar o consumo – o banco compartilha as informações dos correntistas com parceiros. Da mesma forma, por vezes, depois de se abrir uma empresa - uma operação realizada entre a pessoa física e a prefeitura - receber, em nome da pessoa jurídica recém-criada, boletos de associações de empresas cobrando taxas de adesão e se fazendo passar por obrigatórias. É este o cenário que causa debates sobre a privacidade dos dados não apenas do internauta, mas também do cidadão – o que, como em toda *buzzword*, gera discursos utópicos e distópicos, com debates sobre os benefícios e malefícios da transparência de dados, do controle populacional que surge do cruzamento de um grande volume de dados e dos tipos de dispositivos os possuem e os operam.

Boyd e Crawford (2012, p. 668) definem *Big Data* como um fenômeno cultural e tecnológico resultante de três aspectos: a) **Tecnologia**: maximização do poder computacional e da precisão algorítmica para coletar, analisar, conectar e comparar grande conjuntos de dados; b) **Análise**: esquematização de grandes conjuntos de dados para identificar padrões para questões econômicas, sociais, técnicas e legais; c) **Mitologia**: a crença difundida de que os grandes volumes de dados oferecem uma forma mais elevada de inteligência e de conhecimento, que pode gerar *insights* impossíveis anteriormente, por meio de uma aura de verdade, objetividade e precisão.

Enquanto existe consenso quanto aos dois primeiros atores da caracterização do *Big Data*, o caráter mitológico dá margem a algumas considerações. A primeira delas é que o poder de gerar *insights* que seriam impossíveis anteriormente não se trata de uma crença, mas de um fato consolidado. Um exemplo disso é que a utilização da Internet e de rastros de dados dos usuários – em forma de tabelas ou banco de dados – são matéria-prima para *insights* sobre comportamento, preferências e reações às informações, quando se analisa o grande número de pessoas que acessa a *Web*. A análise estatística de dados deste domínio podem revelar o invisível: padrões comportamentais de consumo ou de modos de vida, redes ideológicas, comunidades – pela objetividade dos números e de

métricas a serem adotadas. Porém, quando passamos às visualizações de dados, essas exigem interpretações e *insights* subjetivos sobre as visualizações. No domínio das finanças, por exemplo, alguns investidores se orientam pela análise dos padrões do gráfico resultante de diversos valores, seja de uma determinada ação, de indicadores ou do cruzamento de séries, para analisar tendências de subida ou descida no preço dos ativos.

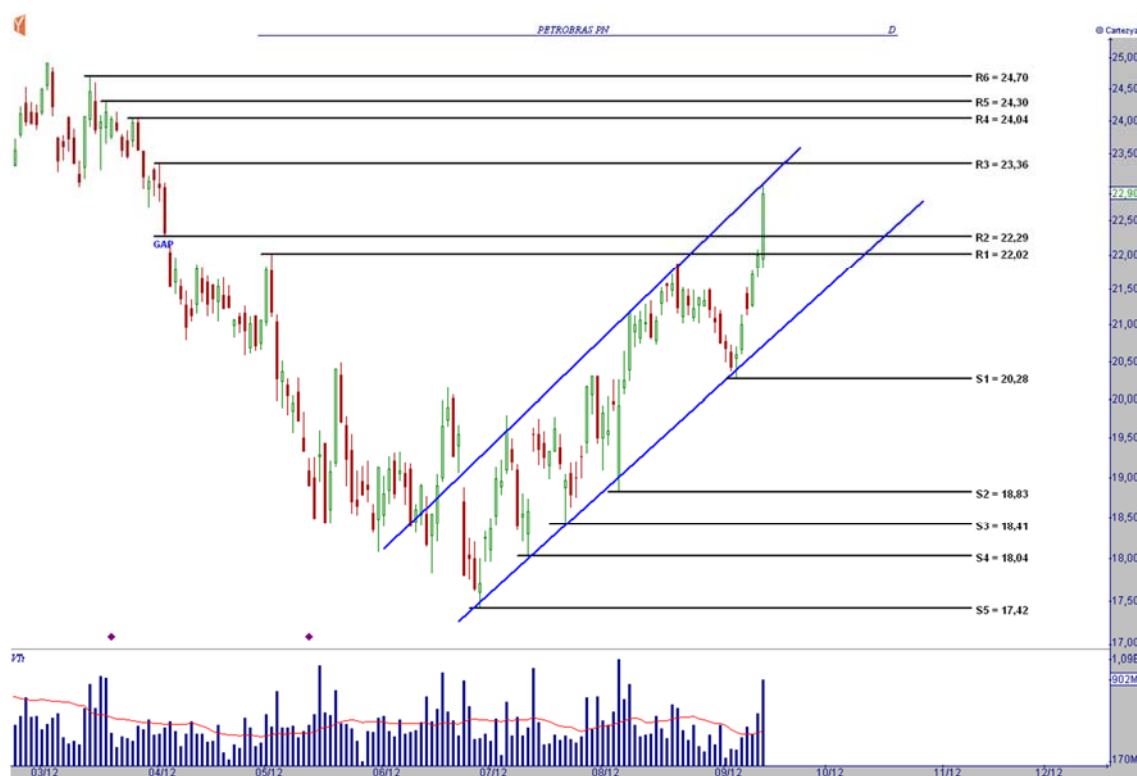


Figura 19 - Análise gráfica da ação PETROBRAS PN, mostrando canal de alta.¹⁶

No que se refere a interpretações de gráficos de séries de dados, mais adiante, neste trabalho, encontram-se visualizações elaboradas a partir de minha rede de amigos no *Facebook*. Elas servem para analisar *clusters* de conhecidos em comum, ajudam muito a avaliar o grau de afinidade entre pessoas, mas isso é feito, na visualização resultante, por uma aproximação que, apesar de não englobar todos os fluxos de trocas comunicacionais entre todos os membros da comunidade estudada, encontra na visualização dos dados uma correspondência verdadeira – os números e a visualização resultante dependem de uma interpretação que, em se tratando de relações pessoais, é construída com o cruzamento de áreas de conhecimento distintas. Diferente do que creem Boyd e Crawford, o *Big Data*

¹⁶ Disponível em https://www.wincorretora.com.br/remote/img/20120914_PETR4.gif. Acesso em 26.11.2015.

não trata de substituir o conhecimento e ignorar áreas do saber, mas sim do contrário. O *Big Data* pode ser uma ferramenta de apoio para diferentes áreas do saber, e depende dessas diferentes áreas para promover *insights* pelo cruzamento e visualização de dados massivos. Sem os conhecimentos de sociologia, psicologia e etc. que precederiam uma análise de agrupamentos humanos, por exemplo, seria uma tarefa muito difícil construir *insights* ou pensar em métricas a partir da visualização de dados, ou mesmo escolher e validar os dados a serem utilizados. E, claro, a análise das visualizações é uma tarefa multidisciplinar e depende do recorte dos dados e de sua qualidade, ou seja, do que se investiga, além do ato perceptivo do analista em relação à visualização – o olhar que busca interpretar a disposição dos dados para atribuir a ela um sentido. É desta operação perceptiva que nascem os *insights*, do mesmo modo que a observação dos padrões repetidos e cada vez menores de algumas flores, como o girassol e o brócolis romano, nos leva a pensar em fractais.



Figura 20 - Padrões das flores de brócolis romano

O acaso também é um elemento importante, já que determinados conjuntos de dados fazem sentido apenas quando representados graficamente e esta representação, resultante de um algoritmo que organiza cada ponto associado a um número, é pouco controlável – não se sabe como ficará o gráfico até que ele seja executado: a percepção e

o acaso andam juntos. E há ainda *insights* que nascem puramente da percepção do analista, de sua sensibilidade:

Ora, para o empirismo, os objetos “culturais” e os rostos devem sua fisionomia, sua potência mágica, a transferências e a projeções de recordações; o mundo humano só tem sentido por acidente. Não há nada no aspecto sensível de uma paisagem, de um objeto ou de um corpo que o predestine a ter um ar “alegre” ou “triste”, “vivo” ou “morto”, “elegante” ou “grosseiro”. **Definindo mais uma vez aquilo que percebemos pelas propriedades físicas e químicas dos estímulos que podem agir em nossos aparelhos sensoriais, o empirismo exclui da percepção a cólera ou a dor que, todavia, eu leio em um rosto, a religião cuja essência todavia eu apreendo em uma hesitação ou em uma reticência, a cidade cuja estrutura todavia eu conheço em uma atitude do funcionário ou no estilo de um monumento.** (Merleau-Ponty, 1999, p. 49, grifo nosso).

No gráfico que se segue, um levantamento foi feito para descobrir em que áreas do conhecimento certos países têm investido mais. Só é possível enxergar um sentido sobre o conjunto de dados com sua visualização. Na figura seguinte, a visualização resultante de um conjunto de dados massivo nos faz perceber o sentido de todos os números que foram fornecidos, mostrando que os EUA têm mais pesquisas ativas em Ciências Sociais, Saúde Mental, Pesquisa Terapêutica e Virologia, entre outros, do que em Física Quântica.

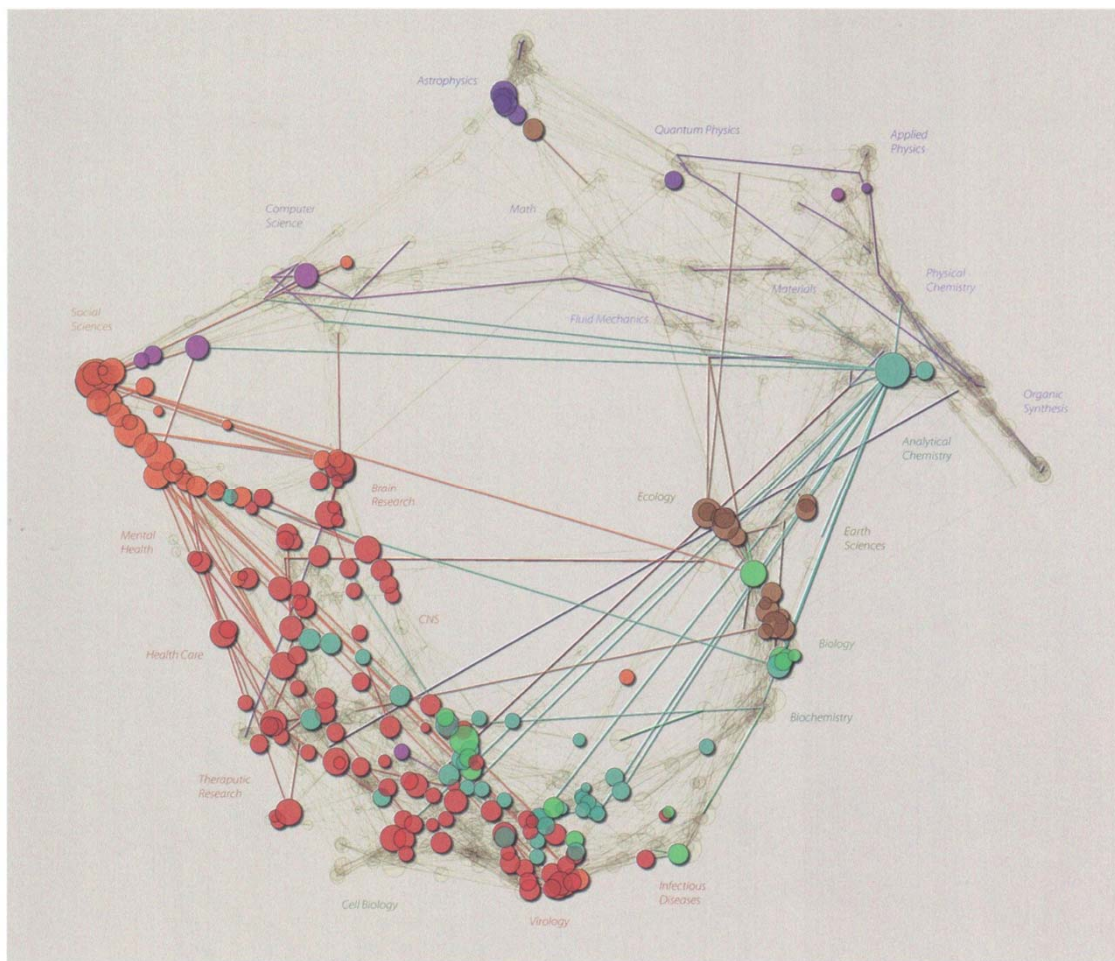


Figura 21 - "A força das nações", de Paley, Klavans et al (2006), com os domínios científicos mais ativos nos EUA.

O *Big Data* está mais para uma nova metodologia do que para uma nova ciência. É um instrumento de investigação. “Mude os instrumentos e mudará toda a teoria social que os acompanha”, diz Latour (2009, p. 9 apud Boyd e Crawford, 2012, p. 665). Um exemplo disto é como a metodologia de análise de dados sobre comunidades criada por Moreno contribuiu para a construção de um ramo da psicologia, como discutido anteriormente neste trabalho. Se realizasse sua pesquisa com as ferramentas de hoje, talvez Moreno ficasse impressionado pela possibilidade de analisar grupos de pessoas muito maiores do que a comunidade de 600 pessoas que analisou no projeto Hudson. Talvez encontrasse quem considerasse suas visualizações de dados inúteis e imprecisas, mas seu contemporâneo Kurt Lewin o defenderia, dizendo que:

A objeção tem sido frequentemente usada contra as representações de situações e que é evidente que alguém possa extrair das situações representadas os eventos que se quer explicar. Dizem que nossas representações não explicam, que elas apenas descrevem. Se o

entendimento do significado dessa objeção deseja ser entendido, deve-se lembrar que se tem explicado processos psicológicos derivados das entidades “por detrás” dos processos¹⁷ (Lewin, 1932, p. 82, tradução nossa).

Acreditamos que o *Big Data*, per se, contrariamente à crença em seus poderes mágicos, só tenha seu poder revelado se ajudar a produzir visualizações dos conjuntos massivos de dados. Por exemplo, é muito difícil percorrer o conjunto de dados disponibilizados pelo *WikiLeaks*¹⁸ em busca de um sentido sem uma visualização.

O infografista do jornal inglês “The Guardian”, Simon Rogers, tratou o grande volume de dados de comunicações entre embaixadas do mundo vazado pelo *WikiLeaks* e gerou visualizações¹⁹ como a que se segue. Podemos deste modo analisar quais os países envolvidos, em que momento as comunicações ocorreram com maior intensidade e quais eram as palavras-chave mais recorrentes.

¹⁷ “The objection has often been raised against our representations of situations that it is self-evident that one can derive from the represented situations the events one wants to explain. It is said that our representations do not explain, that they only describe. If one wants to understand the meaning of this objection, one must remember that one has often explained psychological processes by deriving them from entities “behind ” the processes.”

¹⁸ Sobre o WikiLeaks <https://www.wikileaks.org/wiki/WikiLeaks:About/>.

¹⁹ <http://www.theguardian.com/news/datablog/2011/jan/31/wikileaks-data-journalism>

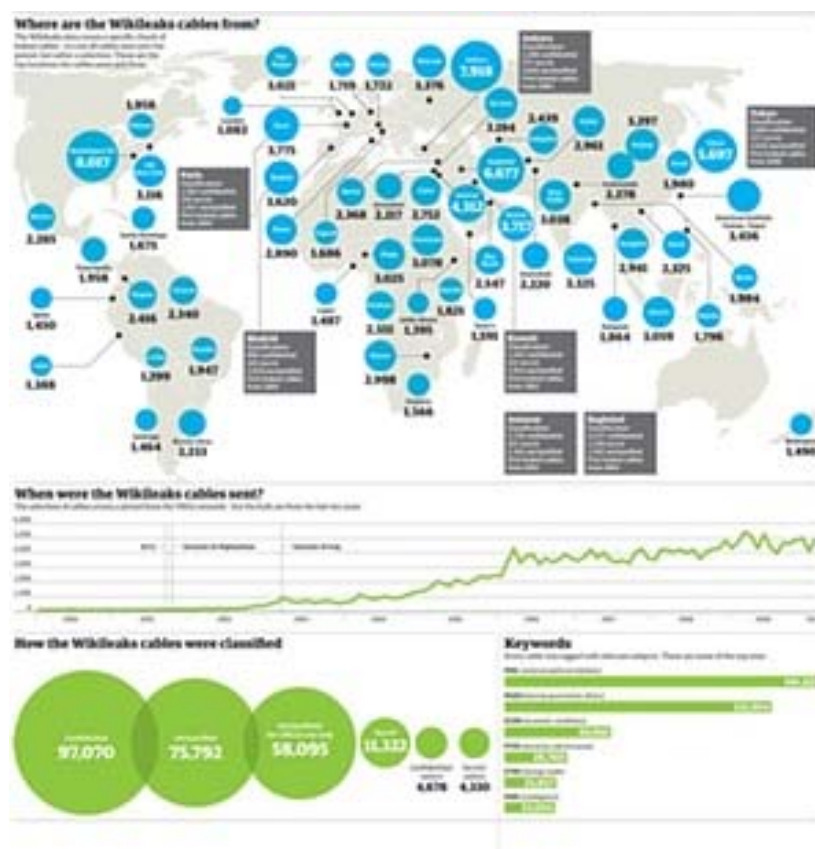


Figura 22 - Infografia realizada a partir de dados do WikiLeaks, por Simon Rogers

Wild (2015) realizou o tratamento de dados do *WikiLeaks* ao aplicar expressões regulares para encontrar informações e estruturá-las em um banco de dados, construindo, assim, as visualizações com o auxílio de algoritmos. A metodologia que Wild utilizou é descrita na página *Web* de sua pesquisa²⁰ e é próxima daquela utilizada no presente trabalho, como veremos nos capítulos subsequentes.

A figura seguinte mostra os fluxos de informação entre diferentes continentes no período de 1972 a 2010. O trabalho de Wild precisa ser consultado *online* para que se observem as interfaces de interação com o volume de dados.

²⁰ Disponível em <http://decodyng.github.io/data-viz-portfolio/wikileaks-visualization/network.html>. Acesso em 27.11.2015.

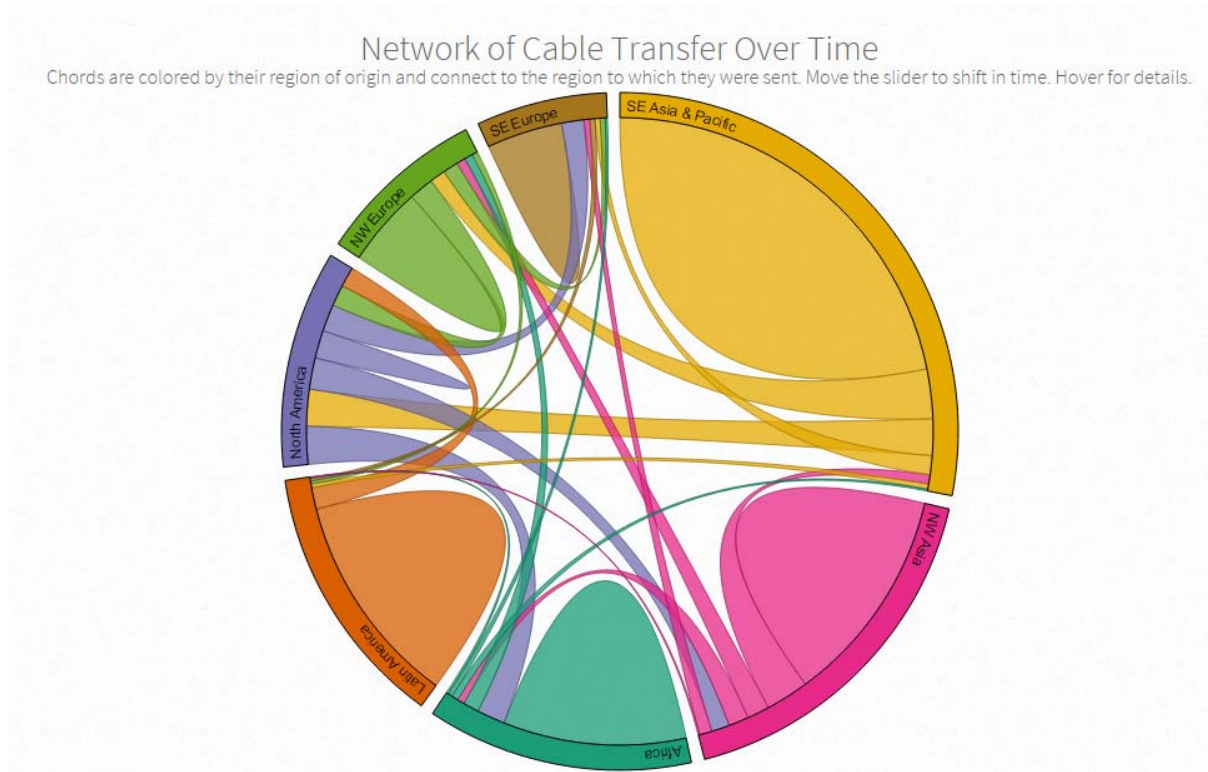


Figura 23 - Visualização sobre WikiLeaks elaborada por Wild (2015).

Aliada ao *Big Data*, a visualização de dados exerce cinco papéis de apoio a conjuntos de dados massivos: documenta, esclarece, revela, amplia e ajuda a imaginar.

Documentar: fazer a cartografia de um sistema nunca antes mapeado. Como consequência da curiosidade inerente ao homem, esse objetivo está ligado à mais antiga das ambições cartográficas: desenhar um território desconhecido. O mapa de um sistema específico pode estimular o interesse por um sujeito ao abrir as portas a novas descobertas e interpretações. A perspectiva de documentação e de registro da estrutura examinada por análises ulteriores é uma das motivações principais de um grande número de projetos; **Esclarecer:** deixar o sistema mais compreensível, inteligível e transparente. O objetivo principal neste contexto é a simplificação – poder explicar os aspectos importantes e esclarecer certas partes precisas do sistema. Ao comunicar de maneira simples e eficaz, a visualização de redes torna-se uma ferramenta importante de tratamento da informação e da compreensão; **Revelar:** encontrar um padrão escondido num sistema, chegando com isso a uma nova conclusão ou, em outras palavras, extrair um diamante de conhecimento de um conjunto de dados plano. O objetivo de “revelação” deveria se concentrar na causalidade, levando à descoberta de relações escondidas, e suas correlações, verificando igualmente as hipóteses iniciais e as questões centrais. **Ampliar:** servir de meio para usos posteriores e possibilitar futuras explorações. A fase de expansão pode implicar na representação de comportamentos

multidimensionais, ou ainda a estrutura representada pode simplesmente tornar-se uma parte de um trabalho mais ambicioso. Nesse contexto, a rede é o meio de se chegar a um fim, uma camada de visualizações suplementares capazes de integrar conjuntos de dados a muitas variáveis. Os nós e as arestas tornam-se o terreno, da mesma maneira que muitos serviços de cartografia online servem de base para outras ferramentas; **Imaginar**: explorar o esquema de rede enquanto ferramenta de representação abstrata. As visualizações de rede podem ser um meio de expressão experimental e metafísica, representando uma multitude de conceitos intangíveis, mesmo que às vezes não haja um conjunto de dados subjacente²¹ (Lima, 2011, p. 80, tradução nossa).

Depois de termos discutido sobre o *Big Data* e introduzido o tema da visualização de dados, discorreremos agora sobre comunidades e fenômenos emergentes. A análise dos instrumentos de análise de dados feita até aqui será fundamental para o estudo da dinâmica dos agrupamentos, comunidades e relações sociais na era do social media, que estudaremos a seguir.

²¹ « **Documenter** : cartographier un système qui ne l’a jamais été. Conséquence de la curiosité inhérente à l’homme, cet objet est lié à la plus ancienne des ambitions cartographiques : se dessiner un territoire inconnu. La carte d’un système spécifique peut stimuler l’intérêt pour un sujet tout en ouvrant la porte à d’autres découvertes et interprétations. La prospective de documentation et d’enregistrement de la structure examinée pour des analyses ultérieures est l’une des motivations principales des nombreux projets ; **Clarifier** : rendre le système plus compréhensible, intelligible et transparent. L’objectif principal dans ce contexte est la simplification – pouvoir expliquer des aspects importants et clarifier certaines parties précises du système. En communiquant de manière simple et efficace, la visualisation des réseaux devient un puissant outil de traitement de l’information et de compréhension ; **Révéler** : trouver un motif caché dans un système, en exhiber une conclusion nouvelle ou, en d’autres termes, extraire un diamant de connaissance d’un jeu de données plat. Le but de « révélation » devrait se concentrer sur la causalité, en amenant à la découverte de relations cachées, et sur les corrélations, en vérifiant également les hypothèses initiales et les questions centrales. **Étendre** : servir de moyen pour des usages ultérieurs et rendre possible de futures explorations. La phase d’expansion peut concerner la représentation de comportements multidimensionnels, ou alors la structure représentée peut tout simplement devenir une partie d’un travail plus ambitieux. Dans ce contexte, le réseau est le moyen de parvenir à une fin, une couche de visualisations supplémentaires capable d’intégrer des jeux de données à plusieurs variables. Les nœuds et les arêtes deviennent le terrain, de la même manière que des nombreux services de cartographie en ligne servent de base pour de nouveaux outils ; **Imaginer** : explorer le schéma de réseau en tant qu’outil de représentation abstraite. Les visualisations de réseaux peuvent être un moyen d’expression expérimental et métaphysique, représentant une multitude de concepts intangibles, parfois même sans un jeu de données sous-jacent. » (Lima, 2011, p. 80).

3. DA ERA DAS MULTIDÕES À DAS REDES SOCIAIS

3.1. Multidão e comunidades

Para o senso comum, a diferença entre o conceito de povo, multidão e comunidade é pouco nítida.

A confusão é compreensível porque as definições claras entre um termo e outro acompanharam, em certa medida, as transformações políticas e sociais. Na Itália renascentista, Maquiavel dedicou ao Duque de Urbino a obra “O Príncipe”, na qual reúne conselhos para conter a rebeldia do povo contra o sistema político e garantir a admiração de um príncipe pelos seus súditos. De acordo com ele, o povo (sujeitos governados²²) já é e deve continuar a ser temido, pois dele - ou de parte dele, configurado em multidão - pode emergir um poder capaz de derrubar seu líder. A palavra “multidão” aparece pela primeira vez, na verdade, em uma passagem de uma obra posterior de Maquiavel²³, que analisa o relato histórico feito por Tito Lívio sobre o império de Augusto. Maquiavel buscou levantar, na obra do historiador, conhecimentos adquiridos naquele período que pudessem servir para melhor governar no século XVI. Assim, sublinhou que “a multidão é mais sábia que o príncipe”, e que “as pessoas são mais fortes juntas”. Neste período, “multidão” e “povo” estão em relação dual, de forças, com a “soberania”. Hardt e Negri (2005, p. 416) explicam esta relação descrita em “O Príncipe”:

A soberania também exige o consentimento dos governados. Além da força, o poder soberano deve exercer hegemonia sobre seus súditos, provocando neles não só medo, mas também reverência, dedicação e obediência, através de uma forma de poder branda e flexível. O poder soberano deve ser permanentemente capaz de negociar a relação com os governados.

A Paris do século XIX viu o empoderamento e radicalização da classe trabalhadora por meio de greves, pela presença da Primeira Internacional e pela formação do Comuna de Paris (um governo socialista revolucionário) em resposta à derrota de Napoleão III na guerra franco-prussiana. É esse cenário de levantes populares que motiva Gustave Le Bon a escrever, alguns anos depois, “A era das multidões”, livro no qual ensaia uma explicação

²² Nesta obra, os súditos são a totalidade dos indivíduos governados e o sinônimo empregado é “povo” para se referir ao mesmo conceito.

²³ “Discursos sobre a Primeira Década de Tito Lívio”, escrito em 1517 e publicado postumamente em 1531.

sobre o comportamento e a psicologia da multidão. Porém utiliza a palavra “multidão” para se referir a sujeitos governados, ou seja, ao “povo”:

Napoleão compreendia maravilhosamente a psicologia das multidões francesas, mas em alguns casos desconheceu completamente a das multidões de raças diferentes (Le Bon, 2008, p. 24).

Porém, poderíamos interpretar a citação anterior de outra maneira: a de que talvez, àquela época, o povo francês não fosse homogeneizado, e que “multidão” possa se referir aos múltiplos modos de vida da população daquele país. Além disso, Le Bon pondera que as multidões se comportam de acordo com a “raça” ou nacionalidade que as compõe, um reflexo do ultranacionalismo e do antissemitismo (como no Caso Dreyfus) de sua época. Diferentemente de Maquiavel, Le Bon não tem o mesmo otimismo em relação ao povo. Para ele, uma multidão não é sábia: faz aflorar a irracionalidade.

Na alma coletiva, apagam-se as aptidões intelectuais dos homens e conseqüentemente sua individualidade. O heterogêneo perde-se no homogêneo e as qualidades inconscientes dominam (idem, p 34).

Le Bon afirma que uma multidão é formada por indivíduos que compartilham qualidades comuns, mas que não são inteligentes:

O partilhar de qualidades ordinárias explica por que as multidões não poderiam realizar atos que exigem uma inteligência elevada. As decisões de interesse geral tomadas por uma assembleia de homens distintos, mas de especialidades diversas, não são sensivelmente superiores às decisões que uma reunião de imbecis tomara. De fato, podem associar as qualidades medíocres que todo o mundo possui. As multidões acumulam não a inteligência, mas a mediocridade (idem, ibidem).

Além disso, a multidão formaria uma “alma coletiva” facilmente sugestionável:

A multidão é sempre intelectualmente inferior ao homem isolado, mas do ponto de vista dos sentimentos e dos atos que esses sentimentos provocam,

ela pode, conforme as circunstâncias, ser melhor ou pior, tudo depende do modo pelo qual é sugestionada (idem, *ibidem*).

O conceito de “multidão” passou por revisões por conta da emergência e consolidação da psicologia social e da psicologia das massas como áreas do saber. Na primeira metade do século XX, ocorreram duas grandes guerras, difundiram-se o rádio e a televisão, intensificou-se o êxodo rural e, assim, as grandes cidades incharam com o aumento súbito de população, acolhendo pessoas de origens distintas. Viver em grupo e na diversidade se torna objeto de estudo estatístico, psicológico, midiático. Os meios de comunicação, junto ao consumo, criam como que um “catálogo de identidades” às quais é possível aderir: modelos são propostos. O conjunto de consumidores e telespectadores é referido como “massa”, assim como os meios de comunicação predominantes, e unidirecionais, são chamados de massivos.

Por volta deste período, quando se diz “multidão”, trata-se da simples “reunião de seres, cada um deles convocado individualmente por uma motivação exterior” (Rabaud apud Maisonneuve, 1951). Há mais leveza nesta definição porque, uma vez que não existe a figura corporificada de um soberano presente, o indivíduo se sente livre para se associar àqueles que compartilham demandas em comum²⁴.

Do mesmo modo que o conceito de multidão evolui, o soberano deixa de ser o mais alto representante de uma monarquia, comandando seu povo diverso, porém semelhante. Hoje a figura do soberano foi substituída por um império global ubíquo, descorporificado, que convoca a adesão de todas as pessoas do globo. Este império reúne sob uma mesma visão, que é a ocidental, diversos povos que podem ou não apresentar semelhança entre si no espaço globalizado. Neste Império, de acordo com Hardt e Negri (2005, p 420), “toda a população tende a tornar-se necessária para o poder soberano, não só como um conjunto de produtores, mas também como consumidores, usuários ou participantes do circuito

²⁴ Existe um certo nível intermediário entre o indivíduo e a multidão, que é a comunidade. Segundo Maisonneuve (1951), esses níveis são classificados em: (a) Comunidade-bloco: mais primitiva, só é preciso estar lá – a unidade é alcançada pelos ritos e tabus, é solidária, opressiva e fechada. (b) Sociedade fechada: as funções são ligadas à sua conservação e aos interesses comuns. É tirana e neutraliza os egos. (c) Sociedade aberta: é permeável a grupos vizinhos. Egoísmo coletivo. (d) Camaradagem: mais flexível e aberta a críticas, onde os indivíduos podem se comunicar livremente. Surgem em guerras e em lutas políticas. De proximidade ocasional, manifestam-se apenas por pressões superiores e exteriores. Neste tipo de arranjo, a comunhão e a colaboração existem apenas num estado rudimentar. (e) Comunidade racional: permite ver no outro não apenas um companheiro, mas um ser singular e insubstituível. É hierárquica e harmônica, de acordo com os princípios de ordem da metade do século XX.

interativo da rede”. Hoje, esta rede “complexa e integrada” de que depende o Império passa a ser temida, assim como eram as multidões que sobressaltavam os príncipes da época de Maquiavel. Como o número de “súditos” do império global é imenso, a ameaça ao poder imperial global é maior porque os meios de comunicação de hoje permitem, pela auto-organização, a emergência de poderes intermediários. Este fenômeno foi recentemente observado na Primavera Árabe, em 2010, e nos protestos de junho de 2013 no Brasil.

Na realidade, quando os produtos do trabalho não são bens materiais, mas relações sociais, redes de comunicação e formas de vida, torna-se claro que a produção econômica se torna imediatamente uma forma de produção política, ou a produção da própria sociedade. De modo que já não somos tolhidos pela velha chantagem; a escolha não é entre soberania e anarquia. O poder da multidão de criar relações sociais em comum coloca-se entre a soberania e a anarquia, com isto apresentando uma nova possibilidade de fazer política (idem, p. 421).

Mais adiante, Hardt e Negri (2005) fornecem um exemplo prático para se compreender o poder da multidão, que é a produção de *software* em código aberto, em que programadores colaboram na criação de aplicativos a distância por intermédio da Internet. Este modo de produção diminuiu a preeminência de grandes empresas que criam códigos fechados e proprietários, como *Microsoft* e *Apple*. Nesta prática, sustentam seus adeptos, a produção do software é otimizada porque há mais olhos para resolver problemas e propor soluções, resultando assim em programas melhores²⁵. Mais valor é gerado por meio da associação livre a outras pessoas, o que acaba levando à dissolução gradual de estruturas rígidas como a hierarquia, dando preferência à colaboração. Com a disponibilidade farta de modos de comunicação possíveis pela Internet, a cada dia testemunhamos o aumento na diversidade dos arranjos: *crowdfunding*, programação colaborativa e sites de leilão reverso de mão de obra especializada, entre outros – articulações de auto-organização só possíveis graças à web, com seus fóruns, redes sociais e aplicativos de mensagens instantâneas.

²⁵ “Quando o código-fonte é aberto, para que todos possam vê-lo, é possível resolver um maior número de problemas neles gerados, e são produzidos programas melhores” (Hardt e Negri, 2005, p. 425).

3.2. Comunidades artificiais

As tecnologias da informação e da comunicação facilitam o encontro de pessoas com o mesmo perfil em torno de centros de interesse. O colecionador de selos, os aficionados por idade média, crochê, naturismo, política; fãs de pop sul-coreano, celebridades americanas, filmes do cinema novo – depois da popularização da Internet, as pessoas que se diferenciam pela sua singularidade em comunidades “físicas” de que participam têm a opção de se aglutinarem em comunidades de especializadas. Este encontro de semelhantes promovido pela Internet propicia aos participantes dessas comunidades uma sensação de conforto e empoderamento, como um refúgio no qual se pode vivenciar a fundo uma paixão. Essas comunidades artificiais, como as nomeia Bauman (2003, p. 18), são baseadas na homogeneidade e na mesmidade. Elas vampirizam o fluxo comunicacional de uma estrutura que as precede, as comunidades naturais, interferindo deste modo no equilíbrio da comunicação “de dentro” e “de fora”. Tendo suas estruturas prejudicadas primeiramente pelos meios de transporte, as comunidades naturais sofrem ainda mais pela informática, que, ao facilitar a homogeneidade e o entendimento, começa a invalidá-las, tornando o caminho artificial, de acordo com Bauman, o único possível:

O golpe mortal na “naturalidade” do entendimento comunitário foi desferido, porém, pelo advento da informática: a emancipação do fluxo de informação proveniente do transporte dos corpos. A partir do momento em que a informação passa a viajar independente de seus portadores, e numa velocidade muito além da capacidade dos meios mais avançados de transporte (como no tipo de sociedade que todos habitamos nos dias de hoje), a fronteira entre o “dentro” e o “fora” não pode mais ser estabelecida e muito menos mantida. De agora em diante, toda homogeneidade deve ser “pinçada” de uma massa confusa e variada por via de seleção, separação e exclusão; toda unidade precisa ser construída; o acordo “artificialmente produzido” é a única forma disponível de unidade. O entendimento comum só pode ser uma realização, alcançada (se for) ao fim de longa e tortuosa argumentação e persuasão, e em competição com um número indefinido de outras potencialidades — todas atraindo a atenção e cada uma delas prometendo uma variedade melhor (mais correta, mais eficaz ou mais agradável) de tarefas e soluções para os problemas da vida (Bauman, 2003, p.19).

Um pouco mais adiante, Bauman sustenta que a vida na comunidade natural, apesar de aparentar paz e pouca preocupação, se trata de um combate permanente contra a discórdia interna e os inimigos externos.

Pessoas que sonham com a comunidade na esperança de encontrar a segurança de longo prazo que tão dolorosa falta lhes faz em suas atividades cotidianas, e de libertar-se da enfadonha tarefa de escolhas sempre novas e arriscadas, serão desapontadas. A paz de espírito, se a alcançarem, será do tipo “até segunda ordem”. Mais do que com uma ilha de “entendimento natural”, ou um “círculo aconchegante” onde se pode depor as armas e parar de lutar, a comunidade realmente existente se parece com uma fortaleza sitiada, continuamente bombardeada por inimigos (muitas vezes invisíveis) de fora e frequentemente assolada pela discórdia interna; trincheiras e baluartes são os lugares onde os que procuram o aconchego, a simplicidade e a tranquilidade comunitárias terão que passar a maior parte de seu tempo (Bauman, 2003, p.19).

Na homogeneidade das comunidades artificiais, de acordo com o filósofo, a vida é menos trabalhosa porque não haveria necessidade de debater exaustivamente até chegar a um consenso, diminuindo as chances de um conflito, ou seja, a auto-organização seria melhor alcançada em comunidades artificiais.

Talvez porque a gestão de conflitos seja mais fácil e o conhecimento nas comunidades artificiais seja mais facilmente compartilhado, comunidades como as de programadores consigam ser mais produtivas. Testemunhamos este tipo de colaboração em plataformas como o *GitHub*²⁶ ou, um pouco antes, na construção da parte central do sistema operacional *Linux*²⁷.

Das comunidades artificiais podem emergir resultados que beneficiem a multidão que as cerca: produtos sociais, culturais, novas ideias.

A continuidade no tempo desses grupos depende da coesão contra “inimigos externos”, como diz Bauman (2003, p.111). Mas quando focalizamos as comunidades presentes na Internet, observamos que, na verdade, há também um diálogo entre essas

²⁶ <http://www.github.com>.

²⁷ Fóruns e tutoriais em www.linux.org.

comunidades artificiais e a multidão exterior. Entre essas duas entidades, há o “espaço” de uma membrana que seleciona o que é absorvido ou rejeitado. É este espaço que delimita o que está dentro e o que está fora. O que se observa é que as comunidades artificiais, impulsionadas pela comunicação ubíqua e veloz, influenciam cada vez mais o ambiente socioeconômico em que estão inseridas.

3.3. Mudanças causadas pelas comunidades artificiais

Jaureguiberry e Proulx (2011, p. 22) observam que, por causa da maior coesão e alcance das comunidades potencializadas pela rede, houve mudanças de paradigmas quanto à produção industrial. Uma delas é relacionada ao determinismo industrial - se em um primeiro momento era a indústria que elaborava soluções e a multidão não tinha alternativa além de aderir e consumir tais produtos, hoje são as comunidades de usuários de um determinado produto têm voz para dizer à indústria que tipo de produto precisa e a indústria, por sua vez, segue essas especificações para fornecer aos consumidores o que desejam.



Figura 24 - Detalhe do tecido bordado a mão de um vestido criado por Balenciaga

Uma outra mudança seria a das redes de produção. Visitamos em 2012 uma exposição com criações do estilista espanhol Cristóbal Balenciaga no Palácio Galliera, dentro do Museu da Moda da Cidade de Paris²⁸. Observamos o feitio dos vestidos, chapéus, entre outras peças, e sabendo que o estilista terceirizava muitas das atividades, como a aplicação de bordados e confecção de rendas, perguntamo-nos se seria possível produzir moda tal como fazia Balenciaga. Pode ser que não, já que a rede de pessoas que o ajudaram nas criações não existe mais. Não é porque talvez os artesãos ou profissionais locais tenham morrido, mas é porque o modo de construir a rede e os conhecimentos que ela agenciava é completamente diferente hoje, e o sabemos apenas ao olhar as etiquetas das roupas que vestimos: *made in Bangladesh*, *made in Vietnam*, *made in China* – muito da produção para as empresas europeias de *fast fashion*, como *H&M* ou *Topshop*, dependem de relações criadas além do alcance de uma comunidade natural. Portanto, há redes não apenas de pessoas, mas redes de comunidades artificiais que causam mudanças nos fluxos econômicos e culturais. Dessas redes, emergem uma relação em que os

²⁸

<http://www.palaisgalliera.paris.fr/fr/expositions/cristobal-balenciaga-collectionneur-de-modes>

desenvolvimentos industrial, social e da comunicação incitam o crescimento uns dos outros.

É possível afirmar que, para sobreviver, a comunidade artificial ou a rede de relações depende principalmente de dois ingredientes: coesão e abertura. Mantendo a sua coesão interna, com o reforço das relações, ela não deve se fechar completamente ao exterior, já que depende das trocas que efetua com a multiplicidade. Tais trocas, porém, são realizadas não numa praça acolhedora de uma comunidade natural, mas eletronicamente, pela Internet. A previsão²⁹ é de que, ao final de 2016 o tráfego de dados no protocolo IP terá alcançado a taxa anual de 1.3 zettabytes (ou $1,3 \times 10^{21}$ bytes)

Estas comunidades artificiais em rede de que falamos dependem de suportes tecnológicos para existir. Cada usuário destas redes depende de um terminal (computador, celular, tablet, TV inteligente) conectado à Internet para poder acessar a serviços como bancos, sites, aplicativos de *home broking*³⁰ ou redes sociais. Uma vez que os computadores dos usuários dessas aplicações trocam informações com os computadores remotos (servidores) que são a base em que estão disponibilizados os serviços, há registro dessas transações. Para usar uma metáfora, existe uma espécie de “caixa preta” que registra o estado do sistema: data de acesso, página pedida, se a entrega foi correta, e em que endereço. Junto a isso, quando o usuário efetua interações com serviços disponibilizados online, como numa rede social³¹, deixa registrados textos, vídeos, fotos, com data de acesso, destinatário, se a entrega foi correta. A intensidade ou volume da comunicação entre pessoas específicas ou entre computadores podem ser quantificados para medir a coesão da comunidade artificial. Com estas informações estruturadas, cruzadas, normalizadas e disponíveis em uma visualização, é possível avaliar o grau da coesão interna da comunidade artificial, assim como quais pontos de contato ela mantém com seu ambiente imediato.

²⁹ Conforme relatório da Cisco Systems disponível em <http://goo.gl/is573a>

³⁰ Software para negociar ações na bolsa de valores, disponibilizado para o investidor que queira gerir sua própria carteira de investimentos.

³¹ Para este projeto, fiz a instalação de um *script clone* do *Facebook*, que demandou a implementação de um banco de dados MySQL e *scripts* em PHP. O resultado pode ser conferido em nananereuh.com.

3.4. Extração e análise de dados em comunidades artificiais

A análise das interações entre usuários e máquinas já é atualmente realizada para se obter *insights* e melhorar a experiência do usuário em um determinado serviço, ou moldá-lo de acordo com o perfil do usuário. Isso é observável ao se pesquisar, por exemplo, sobre um livro no *Google* e dias depois, navegando em uma rede social como o *Facebook*, visualizar publicidade relacionada ao livro na página de atualizações de seus amigos. Apesar de parecer alta tecnologia, é apenas o computador servidor pedindo para ler os rastros de navegação armazenados no computador local em forma de *cookies*. Porém, há maneiras mais sofisticadas de tratar informações do usuário.

Já em aplicação, o *EdgeRank*³² analisa uma quantidade maciça de dados sobre “curtidas”, amigos e conteúdo lido no *Facebook*, pondera o interesse potencial de um usuário em novas publicações, classifica-as e produz um *feed* de atualizações direcionado ao interesse do usuário, de modo que ele se sinta acolhido - transformando a comunidade artificial em um refúgio de autoaceitação. Outro serviço, o *Klout*, classifica os usuários de acordo com seu grau de influência e conectividade, usando 400 sinais de oito redes sociais distintas para atualizar diariamente seu ranking, medindo a quantidade de reações por conteúdo publicado. A lista de serviços é grande: *Twitter*, *Spotify*, *Instagram* – cada rede social guarda em um banco de dados informações estruturadas sobre os usuários, ou sobre o conteúdo que publicam, mas não o faz de modo normatizado. Cada um desses serviços utiliza uma forma de armazenar e buscar dados, com uma linguagem de programação específica. Para acessar esses conjuntos de dados e armazená-los localmente, e assim alimentar posteriormente um algoritmo de visualização de dados, é preciso obter chaves junto aos serviços, que liberam informações sobre um usuário em maior ou menor quantidade, dependendo do acordo de privacidade. A etapa seguinte seria a de normalizar o conjunto de dados e fazer as referências cruzadas. Para um grande volume de dados, é preciso criar algoritmos que busquem e reestruturem informações utilizando expressões regulares. Como exemplo de expressões regulares, na linguagem de programação Python, um CEP pode ser representado por `\d{5}-\d{3}` (cinco dígitos, traço, três dígitos) e uma placa de carro no Brasil por `[A-Z]{3}-\d{4}` (sequência de três caracteres de A a Z, traço, quatro dígitos)³³; quando se procura um arquivo que começa com a letra “m” nos sistemas

³² <http://bit.ly/1K9mDb2>. Acesso em 02.02.2015.

³³ <http://turing.com.br/material/regex/introducao.html>. Acessado em 05.01.2016.

operacionais *Windows*, *MacOS* ou *Linux*, na caixa de pesquisa digitamos *m** (o asterisco é um caractere “coringa”) para obter a listagem de arquivos cujos nomes correspondam ao padrão fornecido. Em uma lista com nomes e sobrenomes, se assim apresentados nesta ordem, o que está entre o começo do registro e o primeiro espaço é uma sequência de letras que representa o primeiro nome. Com as expressões regulares, um algoritmo descobriria a estrutura dos dados e encontraria datas, horários, nomes de quem publicou e de quem recebeu o conteúdo, assim como o tipo de conteúdo publicado. Com a análise da “tabela” resultante, seria possível quantificar a ocorrência e o volume das interações. Mesmo depois de um tratamento por algoritmo de expressões regulares, é preciso intervir para verificar a qualidade dos dados manualmente. A etapa seguinte seria fornecer esta “tabela” para o um programa que permita a interpretação visual desses dados previamente tratados

Imaginemos que temos acesso ao registro das trocas de *e-mail* entre funcionários de uma mesma empresa. Ao criar um algoritmo para determinar, por expressões, quem é o emissor, o destinatário e o horário do envio da mensagem, teríamos um mapa de fluxos dessas trocas. Talvez a visualização desses dados indicasse quais são as comunidades artificiais dessa empresa, ou a comunidade de fato (que transcende a hierarquia) que se comunica pelo suporte eletrônico, revelando parcerias e “vácuos” de comunicação - fornecendo, assim, *insights* para dinamizar a comunicação na empresa de modo geral.

Nos exemplos fornecidos – registros de acessos ao servidor, dados de redes sociais e *e-mails* – há a questão da privacidade. Enquanto os dados de operações governamentais de alguns países são disponibilizados livremente para promover, por exemplo, a transparência em contas públicas, como o faz o governo britânico³⁴, é preciso avaliar se os dados de uma comunidade serão obtidos de acordo com seu desejo de ser conhecida, ou com o consentimento de cada indivíduo sobre as informações fornecidas sobre si. Esta instância ética é o desafio que se impõe na análise de grandes volumes de dados que venham de redes sociais.

A análise dos dados sobre comunidades artificiais e naturais pode dar indícios de fenômenos socioeconômicos que emergem em seu ambiente externo. A seguir, mostro

³⁴ <http://data.gov.uk>. Acesso em 05.01.2016.

como comunidades têm sido analisadas a partir de seus dados para abordar transformações socioeconômicas.

3.5. Análise da comunidade *hipster*

Comportamentos sociais surgem de tempos em tempos e muitas vezes vêm acompanhados de uma ideologia e de um modo de vestir. Hoje percebemos a expansão da estética *hipster*, que pode ser vista mesmo em produções do cinema americano, como em *Her*³⁵ e reavivando certos aspectos ideológicos semelhantes aos da contracultura.

A contracultura foi um movimento de contestação ao *establishment* conservador e ao consumismo. Pereira (1983) define contracultura como:

- a) Um fenômeno histórico concreto e particular, cuja origem pode ser localizada nos anos 60; e b) Como uma postura, ou até uma posição, em face da cultura convencional, de crítica radical (idem, p 14).

A rebelião e o não-conformismo foram marcas deste movimento que teve seu auge nos anos 1960 e, como alguns de seus símbolos máximos, o empoderamento dos movimentos estudantis pelo mundo e a disseminação da ideologia *hippie*. Apesar de não podermos precisar a origem exata desta comunidade, ou seja, exatamente quando seus membros passaram a se diferenciar esteticamente da multidão ou os momentos precisos de seu crescimento e queda, sentimos perdurarem os ecos de sua ideologia, como a importância da liberdade para se lutar por uma identidade diferente da norma ou os questionamentos em relação a modos de vida impostos e tentando derrubar tabus.

Freud (1919) define o tabu não como um componente religioso, tampouco moral – mas sim como uma lei ancestral cuja origem se perdeu em dada sociedade. Ou seja, esta sociedade não saberia exatamente de onde vem esta “ordem” mas continua aplicando-a refletir.

³⁵ *HER*, Direção: Spike Jonze. Fotografia: Hoyte van Hoytema. [S.l.]: Annapurma Pictures, 2013.

Portanto a batalha da contracultura nos anos 1960 foi árdua, já que antagonizou estruturas de raízes ancestrais. Buscou-se uma vida alternativa àquela imposta pela indústria e pelo governo, diversificando desta maneira os percursos de vida possíveis.

Nos dias de hoje, podemos testemunhar alguns movimentos que contestam as imposições industriais e muitas pessoas têm optado, por exemplo, consumir alimentos orgânicos ou a reduzir do consumo de alimentos industrializados; há quem busque uma menor dependência de carros, diminuindo sua utilização para diminuir a poluição e questionar o poderio das montadoras; em vez de esperar as inovações industriais, há os que escolhem engajar a comunidade no levantamento de recursos para projetos de fomento à cultura, na atividade que se chama *crowdfunding*; a tentativa de se fugir do exaurimento das oportunidades nas grandes cidades³⁶ ou de se diminuir o ritmo de vida; a busca de maior liberdade de manifestação da diversidade sexual, o *gender-neutral*; há o renascimento do “faça você mesmo” dos *punks*, em direção a novos caminhos estéticos como o da “moda antimoda”³⁷ e o da estética *hipster*. Todos esses fenômenos podem ser estudados utilizando a visualização de dados para detectar a emergência de comportamentos sociais.

De acordo com Pipengerg (2012), os *hipsters*...

... são apenas o produto final da mistura de todas as contraculturas antecedentes. Parte *punk*, parte *hippie* e parte *hip-hop old school* [“das antigas”], o *hipsters* pode ser indefinível ou nem existir. Tudo o que foi bacana nas cinco décadas mais recentes está sendo arredondado, impresso em camisetas, postados no *Tumblr*, remixado em música *rap* e impulsionado para um fim nada legal.³⁸

³⁶ Disponível em <http://salvesampa.com.br/blog/sampa-ame-a-mas-deixe-a/>. Acessado em 09.01.2015.

³⁷ Disponível em <http://www.livingdesign.net.br/2015/03/sera-o-fim-da-moda.html>. Acessado em 20.02.2015.

³⁸ Disponível em http://www.huffingtonpost.com/the-daily-collegian/hipster-culture_b_1341702.html. Acesso em 20.02.2015.

Segundo Plevin (apud Touboul, 2014), “os *hipsters* evitam os rótulos e rotular. Porém, todos se vestem e agem da mesma maneira e conformam ao seu não-conformismo”³⁹.

Apesar da estética *hipster* ser muito marcante e permitir a identificação fácil de quem adere ao não-conformismo, é importante ressaltar que os *hipsters* possuem desejos parecidos com os dos hippies. Muitos procuram fundar sociedades alternativas fora dos centros urbanos, enquanto outros permanecem nas cidades para transformá-las em lugares melhores para se viver, na medida humana: são partidários da adoção da bicicleta como meio de transporte alternativo ou das hortas comunitárias, por exemplo.

O visual *hipster* tornou-se tão *mainstream* – ou seja, com alta aceitação pela multidão – que mesmo quem não compartilha da ideologia dessa comunidade, por vezes, adota elementos de sua estética, como usar barba mais longa ou fazer um coque nos cabelos. Evidentemente, os *hipsters*, tão contra o *mainstream*, não gostam deste fato e acabam por mudar o visual, o que é um fenômeno de emergência estudado por Touboul.

³⁹ “Hipsters avoid labels and being labeled. However, they all dress the same and act the same and conform in their non-conformity. Doesn't the fact that there is a hipster look go against all hipster beliefs?” Disponível em: <http://arxiv.org/pdf/1410.8001v1.pdf>. Acesso em 20.02.2015.



Figura 25 - Jovens universitários com estilo hipster. HiFoto: Marti Sans/ALAMI⁴⁰

Em sua pesquisa, Touboul estudou um sistema composto por 5 mil membros da comunidade *hipster* que desejavam se diferenciar cada vez mais uns dos outros. Na busca por esta diferenciação, alguns deles adotaram a estética *normcore* ou *punk* não-conformista.

O termo *normcore* se refere à atitude de se vestir com roupas básicas, em reação à saturação causada pela mudança constante da moda. Um estilo para “não chamar a atenção”.

O *normcore* foi alçado também à categoria de novo visual. Esse estilo discreto e anônimo para homens e mulheres – cunhado por K-Hole, um grupo de previsão de tendências de Nova York, tocado por cerca de vinte ex-profissionais do marketing – foi definido pela “reimaginação” dos básicos da era grunge dos anos 1990, incluindo camisas xadrez, jeans desbotados e roupa esportiva. Gigantes da tecnologia e *geeks* aderiram parcialmente ao *normcore*, como retratado no sitcom “Vale do Silício” da HBO. Yasmin Sewell – uma consultora de moda baseada em Londres e estrela da cobertura de moda de rua do fotógrafo Tommy Ton

⁴⁰ Disponível em http://www.huffingtonpost.com/the-daily-collegian/hipster-culture_b_1341702.html. Acesso em 03.04.2015.

da Style.com – deu sofisticação ao estilo básico com Être Cécile – uma linha feminina casual que co-fundou.⁴¹ (Cosgrove, 2015, tradução nossa)

Já o *punk* não-conformista está localizado no ponto oposto do *normcore* no espectro *hipster*, porque se trata daquele que não quer se identificar com tipos de “tribos” específicas – é o alternativo do alternativo.

Para realizar a análise dessas transmutações no grupo, Touboul empregou um modelo canônico de física estatística e adicionou a ele uma fórmula de propagação de informações. Ele explica a metodologia empregada no estudo:

⁴¹ “Normcore” was also promoted as a new look. This nondescript anonymous style for men and women—which was named by K-Hole, a New York forecasting group, operated by 20-something former marketing professionals—was actually defined by reimagining 1990s grunge-era basics, including plaid shirts, faded jeans, and sporting goods. Tech moguls and computer geeks alike were partial to normcore, as portrayed in the HBO cable-TV sitcom *Silicon Valley*. Yasmin Sewell—a striking London-based fashion consultant and star of photographer Tommy Ton’s popular Style.com street-fashion coverage—lent sophistication to basic style with Être Cécile—a casual women’s wear line that she cofounded.”

Em diferentes domínios como física estatística e vidros de spin, neurociências, ciência social, economia e finanças, um conjunto grande de indivíduos que tomam decisões, seja de acordo (*mainstream*) ou contra (*hipsters*), a maioria é onipresente. Ainda, na tentativa árdua dos *hipsters* se diferenciarem uns dos outros, eles acabam tomando as mesmas decisões, em outras palavras, acabam se parecendo uns com os outros. Resolvemos este paradoxo aparente estudando um modelo canônico de física estatística, enriquecido pela incorporação dos atrasos necessários para que a informação seja comunicada. Nós mostramos uma transição de fase genérica do sistema: quando os *hipsters* estão muito lentos para detectar as tendências, eles continuarão a fazer as mesmas escolhas e, portanto, permanecerão correlacionados com o passar do tempo, enquanto as tendências deles evoluem no tempo numa função periódica. Isto será verdadeiro enquanto a maioria da população é feita de *hipsters*. Senão, os *hipsters* serão, mais uma vez, muito alinhados, dirigidos a uma direção constante que é imposta pelas escolhas do *mainstream*. Estando além da escolha do melhor terno para se usar neste inverno, este estudo pode ter implicações importantes para compreender a dinâmica das redes inibidoras do cérebro ou finanças de estratégia de investimentos, ou para compreender a dinâmica emergente em ciências sociais, domínios nos quais os atrasos de comunicação e geometria dos sistemas são proeminentes.⁴² (Touboul, 2015, tradução nossa)

Na figura a seguir, pode-se ver o gráfico resultante do trabalho de Touboul. Na amostra A, o grupo está no seu estado inicial; em B há a homogeneização; em C o grupo volta a se diferenciar. Pode-se ver uma turbulência nas relações, representada pela linha m no gráfico, quando os indivíduos adquirem certa homogeneidade.

⁴² “In such different domains as statistical physics and spin glasses, neurosciences, social science, economics and finance, large ensemble of interacting individuals taking their decisions either in accordance (*mainstream*) or against (*hipsters*) the majority are ubiquitous. Yet, trying hard to be different often ends up in *hipsters* consistently taking the same decisions, in other words all looking alike. We resolve this apparent paradox studying a canonical model of statistical physics, enriched by incorporating the delays necessary for information to be communicated. We show a generic phase transition in the system: when *hipsters* are too slow in detecting the trends, they will keep making the same choices and therefore remain correlated as time goes by, while their trend evolves in time as a periodic function. This is true as long as the majority of the population is made of *hipsters*. Otherwise, *hipsters* will be, again, largely aligned towards a constant direction which is imposed by the *mainstream* choices. Beyond the choice of the best suit to wear this winter, this study may have important implications in understanding dynamics of inhibitory networks of the brain or investment strategies finance, or the understanding of emergent dynamics in social science, domains in which delays of communication and the geometry of the systems are prominent.”

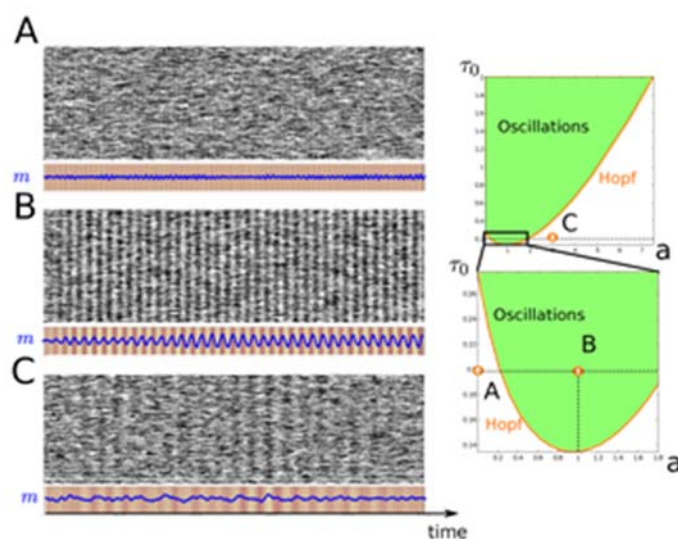


Figura 26 - Perturbações na rede levam à emergência de novos padrões (Touboul, 2015)

Touboul conclui seu estudo dizendo o seguinte:

A análise deste modelo relativamente simples permitiu avançar muito na compreensão do papel concorrente do ruído, atraso, e proporções dos *hipsters* e dos indivíduos do mainstream nessa emergência da sincronização entre *hipsters*. É um fato interessante que a sincronização possa depender do formato preciso da distribuição dos atrasos: para que a sincronização emergja, é preciso que haja atrasos = longos e coesos o suficiente (pequeno desvio dos atrasos). Isto mostrou que o fenômeno inesperado da sincronização entre os *hipsters* depende da distribuição no espaço de cada indivíduo, quando os atrasos ocorrem em função da distância entre dois indivíduos.⁴³ (Touboul, 2015, tradução nossa)

Se levarmos em conta que os serviços de “rede social”, como *Facebook* e *Tumblr*, entre outros, aproximam os elementos de uma comunidade entre si, mantendo-os em distância equivalente por meio da sincronia dos fluxos de informação, aumentando assim a coesão do sistema, não seria exagero concluir que as emergências de fenômenos são

⁴³ “The analysis of the relatively simple model allowed to go very far in the understanding of the concurrent role of noise, delays and proportions of hipsters and mainstream individuals in this emergence of synchronization among hipsters. Interestingly, synchronization may depend on the precise shape of the distribution of the delays: for synchronization to emerge, one needs both sufficiently long delays and sufficient coherence (small standard deviation of the delays). This yielded the unexpected phenomenon that synchronization among hipsters depends on the distribution, in space, of each individuals, when the delays are function of the distance between two individuals.” (Touboul, 2015)

potencializadas nas redes, já que ocorrem em unidades de tempo menores do que nas comunidades naturais – a propagação de mudanças ali é mais rápida⁴⁴.

As conclusões resultantes do modelo matemático posto em prática por Touboul corroboram o que a psicologia social conceitua como coesão em numa rede⁴⁵ de pessoas:

As conexões baseadas em valores compartilhados tentem a aumentar ou diminuir de acordo com as mudanças de circunstâncias sofridas pelo indivíduo ou pela sociedade. Assim como não podemos enumerar nossa rede pessoal, que por algum motivo poderia mudar amanhã, uma rede grupal também conhece raramente a extensão da influência e dos seus recursos (Lipsnack e Stamps, 1986, p. 7-8 apud Samuels, 1992, p.61).

Tais circunstâncias seriam as perturbações causadas por mudanças no arranjo interno ou externo da comunidade. Os modelos matemáticos, dessa forma, ajudariam a psicologia social a precisar a geometria das redes. Lipsnack e Stamps apresentam a polissemia da palavra “rede”:

- Um sistema físico que se parece com uma árvore ou uma grade
 - Um sistema de módulos e elos
 - Um mapa de linhas entre pontos
 - Uma identidade persistente de relacionamentos
 - Uma rede de pesca mal amarrada
 - Uma estrutura da qual não se conhece fronteiras
 - Uma comunidade não-geográfica
 - Um sistema de sustentação
 - Uma linha da vida
 - Todas as pessoas que você conhece
 - Todos que conhecem quem... faz natação, coleciona moedas, canta no coral da igreja, cuida das crianças quem vêm para a escola, lê Teilhard de Chardin...
- (Lipsnack e Stamps, 1986, p. 2 apud Samuels, 1992, p. 60).

⁴⁴ Há um exemplo de propagação de comportamentos numa comunidade natural neste link: <http://bit.ly/1MtOfww>. Acesso em 24.11.2015.

⁴⁵ “Rede” e “rede grupal” significa, para Lipsnack e Stamps, um arranjo social que pode ser interpretado como comunidade ou sociedade, nos termos descritos no presente trabalho (Samuels, 1992).

O conceito de “redes”, assim explicado com base na matemática e na psicologia, se aproxima da “diferença que se mantém diferente”, sendo o conjunto de “singularidades”, como Hardt e Negri conceituam, multidão⁴⁶.

Portanto a matemática, a física estatística e a psicologia podem cooperar para o entendimento de fenômenos emergentes das comunidades e das redes. Poderiam elas, juntas, poderiam prever comportamentos futuros (*trend forecasting*) como numa previsão climática, partindo de dados estruturados sobre a comunicação de pessoas, como aqueles extraídos das redes sociais?

⁴⁶ “A multidão é composta de um conjunto de singularidades – e com singularidades queremos nos referir aqui a um sujeito social cuja diferença não pode ser reduzida à uniformidade, uma diferença que se mantém diferente” (Hardt e Negri, 2005, p.139).

3.6. Redes sociais x redes relacionais

O presente estudo não trata apenas da análise de redes sociais. Como dito anteriormente, as redes sociais fornecem dados estruturados, o que permite processá-los para, assim, construir visualizações com as quais se faria uma análise. Porém, a mesma metodologia pode ser transposta para analisar também as comunidades ditas naturais, isto é, constituídas fora da rede, desde que se colem dados de forma estruturada, como num banco relacional ou numa ontologia de dados, como a que pode ser construída com os parâmetros FOAF, entre outros⁴⁷.

Quando se faz uma análise baseada nos dados provenientes das redes sociais, de acordo com os resultados parciais encontrados, pode-se apontar, nos gráficos resultantes, aglomerados que são, na verdade, transposições de comunidades reais. Todavia, é possível apontar a topologia resultante como monodimensional, já que esta analisa somente os nós (“amigos”) em comum. Um exemplo disto é o mapa elaborado a partir de dados do *Facebook*⁴⁸.

⁴⁷ FOAF (friend of a friend) é uma ontologia inicialmente orientada para a padronização de dados pessoais. Para conhecer mais sobre ontologia de dados, ver <http://semanticweb.org/wiki/Ontology>. Acesso em 24.11.2015.

⁴⁸ Mais exemplos disponíveis em <http://bit.ly/1DrxVYX>. Acesso em 24.11.2015.

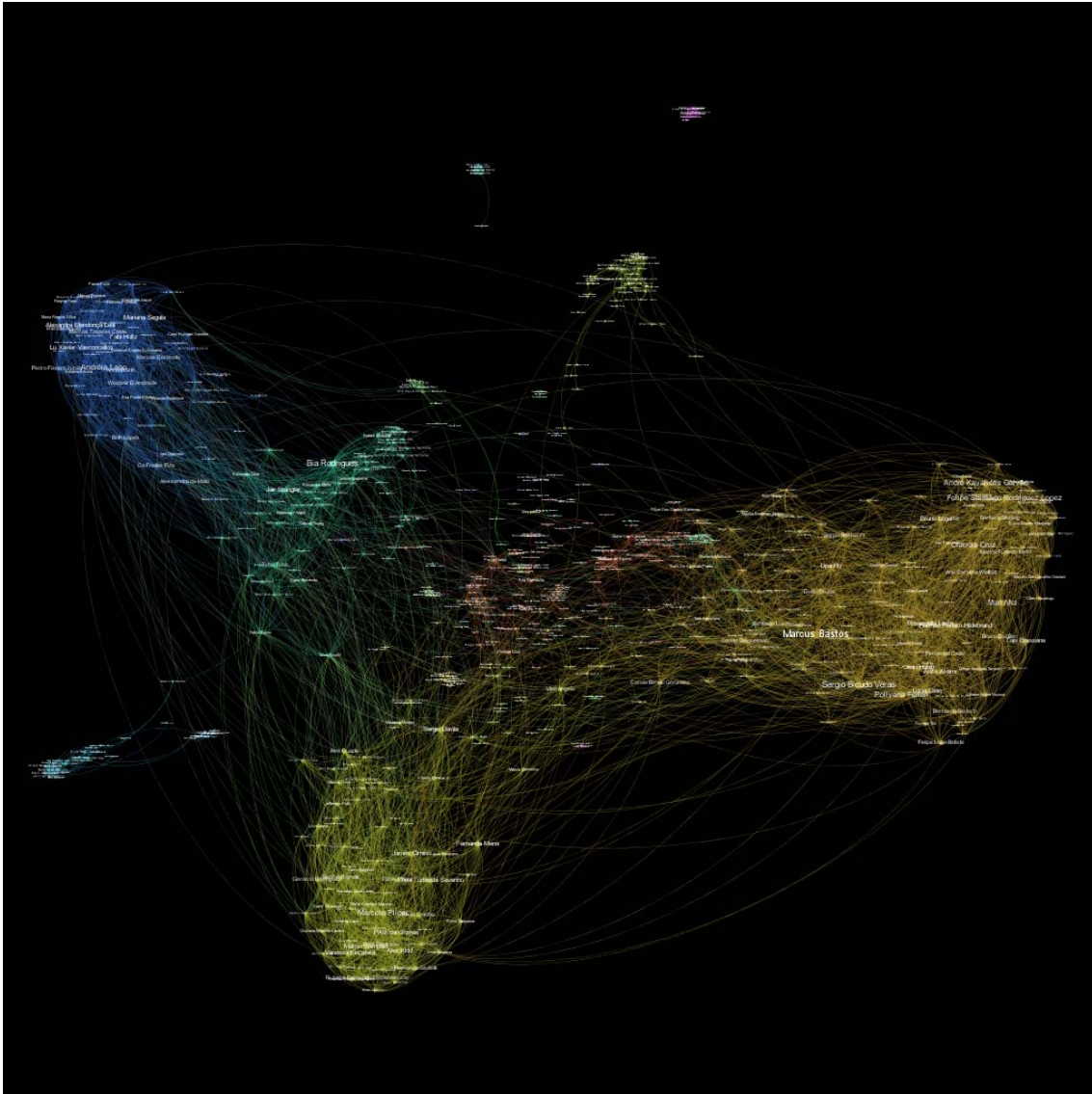


Figura 27 - Visualização gerada a partir de conexões do Facebook

O que não sabemos, olhando para esse mapa, é se o nó do centro dessa comunidade é realmente aquele que compartilha informações relevantes com a mesma, e nem mesmo o tipo ou volume de informações (no caso do *Facebook*, compartilhamentos em forma de texto, vídeo ou foto). Isto porque os dados de origem apontam apenas os nós em comum, ou seja, como os “amigos” estão relacionados.

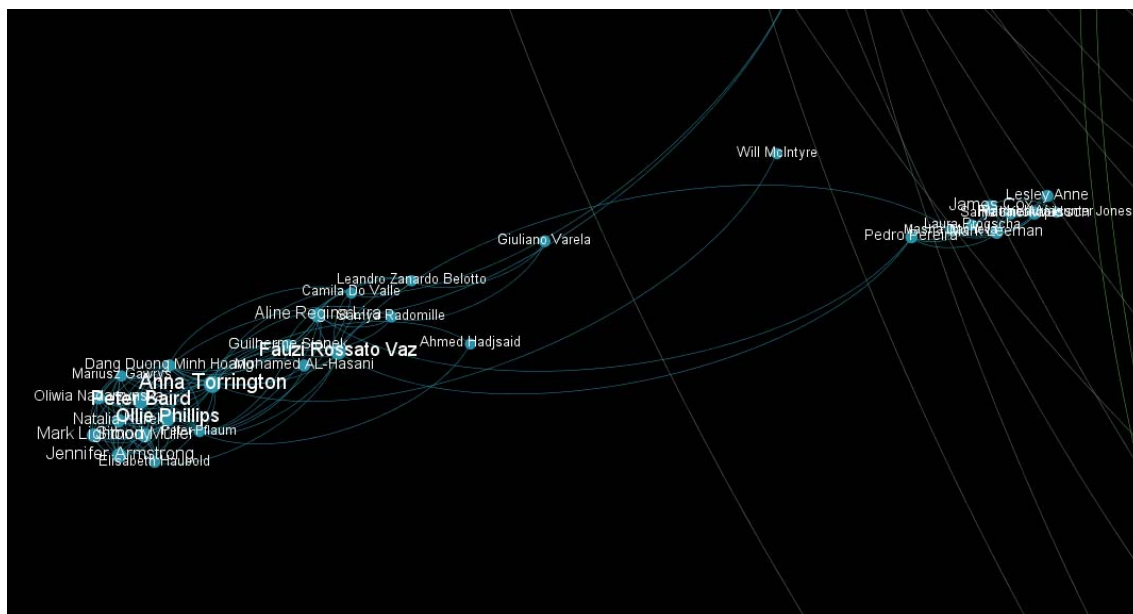


Figura 28 - Visualização gerada com dados de conexões do Facebook, detalhe de uma comunidade

O teste subsequente foi construir um mapa a partir dos dados que estruturamos⁴⁹, mostrando partidos políticos do Brasil relacionados por ideologia⁵⁰. Para cada nó é atribuído mais de um valor, como “fisiologismo”, “republicanismo”, “centro”, “centro-esquerda”. Deste modo, veem-se linhas que unem cada nó, representando os valores atribuídos. A topologia resultante é mais fiel do que aquela do mapa anterior, já que a distância entre os nós é melhor aferida por conta de mais valores, ou atributos, sendo computados.

⁴⁹ Disponível em <http://bit.ly/1NEbe36>. Acesso em 24.11.2015.

⁵⁰ A metodologia empregada será detalhada mais adiante, no capítulo sobre a confecção de mapas.

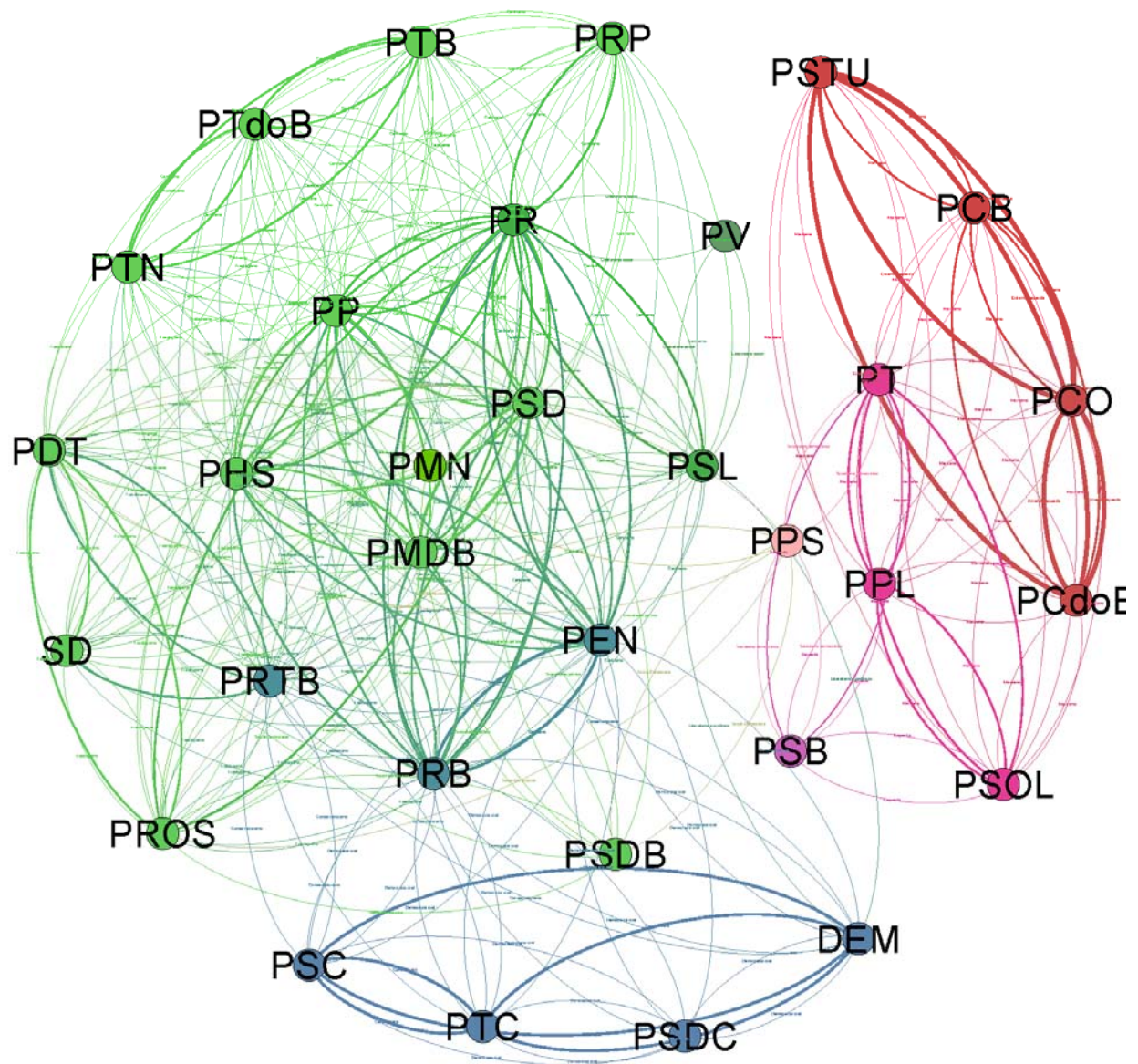


Figura 29 - Visualização Partidos do Brasil por afinidade

Um artigo produzido na Universidade de Tóquio aponta um aprimoramento das análises de comunidades e de redes neste sentido.

Estudos anteriores sobre a detecção de comunidades se focalizam principalmente as redes homogêneas e unirrelacionais que contêm apenas um tipo de nós e linhas; muitos sistemas do mundo real, porém, são naturalmente descritos como redes heterogêneas (contêm múltiplos tipos de nós) e multirrelacionais (contêm múltiplos tipos de linhas). Por exemplo, os usuários do *Flickr* podem enviar fotos, adicionar *tags* às

fotos e estabelecer amizade com outros usuários. Como mostrado na figura 1 [abaixo], o *Flickr* pode ser descrito como uma rede multirrelacional heterogênea, que contém três tipos de nós (usuários, fotos e *tags*) e três tipos de linhas que representam as relações citadas acima. Tradicionalmente, primeiro simplificamos uma rede multirrelacional heterogênea em rede unirrelacional homogênea antes de conduzir uma análise sobre as comunidades. No exemplo dado sobre o *Flickr*, podemos omitir informações sobre fotos e *tags* e detectar comunidades diretamente na rede homogênea de amizades de usuários. Porém, podemos omitir muitas informações valiosas nesse processo de simplificação. Por exemplo, os dados sobre usuários podem ser incompletos e ruidosos e alguns usuários ativos têm milhares de amigos, enquanto alguns não possuem amigo algum. Consequentemente, não podemos obter comunidades reais de usuários baseando-se na amizade das redes *per se*⁵¹ (Liu, Murata, Wakita et al., 2014, p. 12, tradução nossa).

O resultado do estudo japonês foi a criação de um modelo matemático para contemplar a multiplicidade de dados na modularidade, que, matematicamente, é a expressão que determina subgrupos numa comunidade. Modularidade é aquilo a que Bauman se refere quando trata de homogeneidade e mesmidade. Graficamente, é expressa pelos *clusters* ou ilhas formadas por pessoas que têm amigos em comum.

Assim, os modelos matemáticos e a visualização de dados podem nos ajudar a compreender os fluxos de sinais que tornam uma comunidade coesa.

⁵¹ “Previous studies on community detection mainly focus on homogeneous single-relational networks which contain only one type of nodes and edges. Many real-world systems, however, are naturally described as heterogeneous (contain multiple types of nodes) multi-relational (contain multiple types of edges) networks. Take the photo service website Flickr as an example. Flickr users can upload photos, annotate photos with tags, and establish friendship with other users. As shown in Fig. 1, Flickr can be described as a heterogeneous multi-relational network, which contains three types of nodes (users, photos, and tags) and three types of edges representing the above relations. Traditionally we first simplify a heterogeneous multi-relational network to a homogeneous single-relational network and then conduct community analysis. Take the above Flickr network as an example. We can omit information about photos and tags, and detect communities directly in the homogeneous user friendship network. However, we may omit too much valuable information in this simplification process. For example, the data about users can be incomplete and noisy, and some active users have thousands of friends while some have no friends at all. Consequently, we cannot obtain real user communities based on the friendship network alone” (Liu, Murata, Wakita et al., 2014, p. 12).

Os aplicativos estudados que possibilitam a extração de dados de redes sociais não permitem obter os dados de compartilhamento de conteúdo entre pessoas - por exemplo, o tipo, a frequência ou hora de sua publicação. Esses dados apontariam se a troca maior de publicações entre duas pessoas seria visualmente representada por uma topologia de comunidade artificial fiel à relação que essas pessoas teriam numa comunidade natural, produzindo, deste modo, a visualização a partir de dados conforme relatada no artigo japonês.

Dessa forma, para investigar redes relacionais por meio de visualizações com maior fidelidade topológica, foi escolhida primeiramente uma abordagem inspirada no trabalho sobre relações interpessoais de um ícone do movimento de contracultura no século XX: Timothy Leary.

O psicólogo americano produziu, nos anos 1950, um estudo sobre a estrutura e a dinâmica de grupos partindo da personalidade de cada um dos indivíduos⁵², muito inspirado na psicometria. Trinta anos depois, a companhia *Electronic Arts* fez um jogo, *Mind Mirror*, fundamentado no estudo de Leary. O jogador deveria explorar o seu papel dentro da dinâmica de um grupo social que escolhesse, ou seja, como sua personalidade se encaixaria de acordo com os perfis de cada membro do grupo. Os dados do jogador eram, então, colhidos em forma de questionários de múltipla escolha. Depois do processamento dos dados, o jogo elaborava uma visualização dos resultados em forma de círculos por meio dos quais o jogador se dava conta, por exemplo, do quão extrovertido realmente era.

O ponto central do programa é o conceito do Círculo Interpessoal (IPC), também conhecido como o Circumplexo Interpessoal. Ele gera uma visualização que permite aos usuários apontar o seu comportamento e compará-lo com o de outras pessoas, baseando-se nos 16 tipos e 4 aspectos da personalidade humana.⁵³

⁵² Trata-se de seu doutorado "The social dimensions of personality: groups structure and dynamics", elaborado na Universidade da Califórnia, em 1950, que constitui um marco na psicometria.

⁵³ "At the core of the program is the concept of the Interpersonal Circle (IPC) aka Interpersonal Circumplex. It provides a visual motif that allows users to plot their behavior and compare it with others in terms of 16 character types in four aspects of human personality". Disponível em <http://www.timothylearyarchives.org/mind-mirror/>. Acesso em 03.04.2015.

Esse jogo de 1985, que rodava em PC, *Apple II* e *Commodore*, foi disponibilizado para *download* gratuito no site “Timothy Leary Archives”⁵⁴. Para executá-lo em um computador de hoje, foi necessária a instalação de um emulador DOS.



Figura 30 - Captura de tela do jogo Mind Mirror (1985) rodando em emulador.

O jogo foi recentemente refeito com uma interface de melhor resolução, disponibilizado gratuitamente e adaptado para o *Facebook*⁵⁵. Esta nova versão é simplificada. O jogador utiliza *sliders*, com vinte e cinco gradações cada, para atribuir-se características. Na versão antiga, o modo de captação de dados seguia uma linha narrativa mais complexa, na qual o jogador se imaginava em diferentes fases da vida e escolhia como reagir às situações entre respostas pré-determinadas.

Isto posto, *Mind Mirror* atenderia a muitos requisitos para a execução do presente estudo: é fundamentado num estudo prévio sobre comportamento do indivíduo no grupo e fornece parâmetros além do simples conhecimento mútuo para melhorar a topologia de cada pessoa na visualização da rede relacional.

⁵⁴ Disponível em <http://www.timothylearyarchives.org/mind-mirror/>. Acesso em 03.04.2015.

⁵⁵ Disponível em <https://apps.facebook.com/mindmirrordevsb/>. Acesso em 24.11.2015.

Para exemplificar, foi feito o *download* das informações da comunidade “Dudu Guru” no *Facebook*, formada por cerca de 50 pessoas que se conhecem, majoritariamente, a partir de uma comunidade natural: elas eram vizinhas e têm a mesma faixa etária. A visualização resultante dos dados de conectividade é a que se segue⁵⁶.

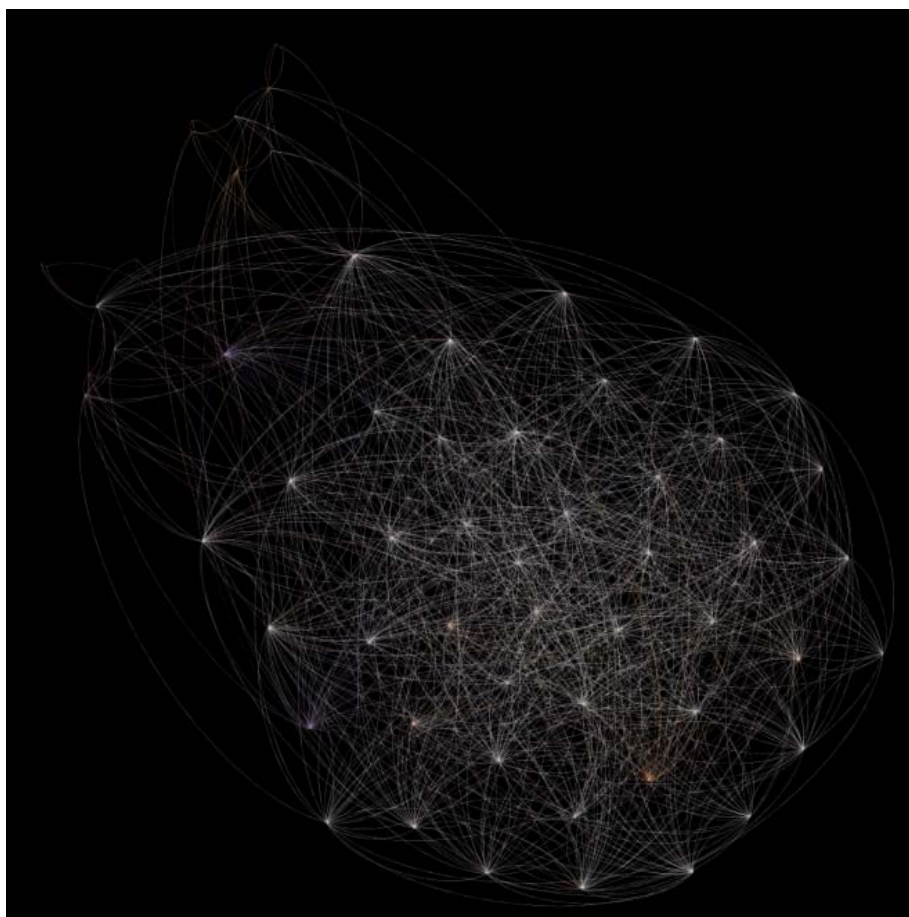


Figura 31 - Mapa da comunidade "Dudu Guru" no site www.facebook.com.

Observa-se que aqueles que se conhecem pessoalmente formam uma nuvem maior e mais densa. Na região superior e à esquerda desta imagem, pode-se notar que algumas pessoas têm menos conectividade e extrapolam a nuvem densa. Provavelmente são pessoas que possuem menos “amigos em comum” na comunidade artificial. O que esta topologia não mostra é de que modo os vínculos interpessoais estão estabelecidos por perfil. Não é possível afirmar que o que os une é, por exemplo, a extroversão. O *Mind Mirror* permitiria detectar as afinidades desta rede com mais precisão ao integrar os dados

⁵⁶ Disponível em alta resolução em <http://bit.ly/1CRqUtg>. Acesso em 03.04.2015.

deste aplicativo aos dados de conexão obtidos de uma rede social como, por exemplo, o Twitter ou Facebook. Quanto mais dados sobre as mesmas entidades inter-relacionadas, mais precisa a modularidade, ou subgrupos, de um mapa de afinidades e, para tanto, é necessário um conjunto de dados multirrelacional.

Como todos os membros da comunidade estudada estão no *Facebook*, seria prático solicitar a cada um deles que preenchesse o seu perfil na página do aplicativo *Mind Mirror*, disponibilizado na mesma plataforma. Porém, surgem alguns obstáculos. Um deles é a barreira do idioma, já que o aplicativo está em inglês. Outro problema seria a baixa adesão em preencher o questionário de forma independente. Além disso, os perfis preenchidos no *Mind Mirror* não são disponibilizados publicamente, o que nos impediria de construir uma análise de todos os perfis que desejam ser estudados. Deste modo, traduzimos o teste e o estruturamos de forma a replicar as gradações dos *sliders* do jogo em um arquivo *Excel*, para aplicá-lo presencialmente.⁵⁷ Porém, se aplicarmos o teste do *Mind Mirror* em um grupo de tamanho considerável, de uma centena de pessoas, cada uma delas deve responder o questionário analisando todas as outras do estudo, composto de mais de dez *sliders*. Deste modo, criamos o questionário em círculos concêntricos apresentado no começo deste trabalho. Este questionário, ao contrário daquele construído por Timothy Leary, apresenta a vantagem de declarar a afinidade com base nas relações cotidianas ao invés de adotar uma tipologia, evitando assim juízos de valor que poderiam descrever as características de alguns membros do grupo de forma muito variável, de acordo com um olhar subjetivo.

⁵⁷ Disponível para download em <http://bit.ly/1aVqQCK>.

4. COMO EXECUTAR MAPAS

Para começar a produção de mapas, buscou-se um programa que atendesse a três requisitos: (i) Código aberto: gratuito e com a possibilidade de se adaptar o código-fonte, se necessário; (ii) Facilidade de uso: que fosse palatável a quem não tem familiaridade com estatísticas e cálculos complexos; (iii) Multiplataformas: para que a instalação e o experimento fossem reprodutíveis independentemente do sistema operacional utilizado.

Entre os programas indicados pela página da *Butler Analytics*⁵⁸, o *Gephi* foi escolhido por atender a todos os requisitos. É um programa para o qual se produzem muitas extensões (*plug-ins* que permitem mesmo a importação de ontologias online ou arquivos *Excel* para a montagem de visualizações de dados) e possui farta documentação online, por meio de fóruns e vídeos explicativos. De acordo com o site do programa⁵⁹, foram três as instituições que formaram o alicerce para o *Gephi*: o *Media Lab* do instituto francês *SciencesPo*, a agência de opinião e marketing *Linkfluence* e a ONG *WebAtlas*, pioneira no estudo de redes na França.

Gephi é um programa em *software* aberto para análise de gráficos e de redes. Ele utiliza um motor de renderização em 3D para mostrar redes extensas em tempo real e aumentar a rapidez da exploração dos dados. Uma arquitetura flexível e multitarefa traz novas possibilidades de trabalhar com conjuntos complexos de dados e produzir resultados visuais preciosos. Apresentamos muitas das características-chave do *Gephi* no contexto da exploração interativa e interpretação de redes. O programa propicia um acesso fácil e amplo aos dados de rede e permite especializar, filtrar, navegar, manipular e agrupar. Finalmente, ao apresentar as características dinâmicas do *Gephi*, destacamos aspectos-chave da visualização de dados dinâmicos⁶⁰ (Bastian, Heymann e Jacomy, 2009, p.2).

⁵⁸ Disponível em <http://butleranalytics.com/10-free-social-network-analysis-tools/>. Acesso em 03.04.2015.

⁵⁹ <https://gephi.github.io/>. Acesso em 03.04.2015.

⁶⁰ “Gephi is an open source software for graph and network analysis. It uses a 3D render engine to display large networks in real-time and to speed up the exploration. A flexible and multi-task architecture brings new possibilities to work with complex data sets and produce valuable visual results. We present several key features of Gephi in the context of interactive exploration and interpretation of networks. It provides easy and broad access to network data and allows for spatializing, filtering, navigating, manipulating and clustering. Finally, by presenting dynamic features of Gephi, we highlight key aspects of dynamic network visualization” (Bastian, Heymann e Jacomy, 2009, p.2).

4.1. A interface do *Gephi*

À primeira vista, a interface do *Gephi* apresenta três seções: (i) Em “Visão Geral”, pode-se construir a visualização de dados. Há ferramentas para ordenar os nós, arrastá-los e colori-los. São apresentados alguns algoritmos para ordenação automática: pode-se ordenar os nós por conectividade (no linguajar do programa, isto tem o nome de “grau”) e agrupá-los por modularidade (blocos por semelhança). Uma caixa com opções para executar os algoritmos estatísticos se encontra nesta seção, possibilitando calcular o índice de modularidade e o grau médio entre os elementos de um grupo. (ii) Na seção “Laboratório de dados”, as tabelas com os dados que gerarão os gráficos podem ser editadas ou importadas. É possível tanto alimentar dados um a um nessa interface quanto importar uma tabela de dados externa no formato CSV (“Comma separated values” ou “valores separados por vírgula”). Os dados se dividem entre nós e arestas, sendo que os nós são os pontos, as entidades em si, enquanto as arestas consistem na ligação entre os nós uns com os outros. (iii) Na seção “Visualização”, vemos o resultado do gráfico elaborado em “Visão Geral” com mais acuidade e, portanto, com o processamento um pouco mais lento. A partir dessa seção, é possível exportar o gráfico resultante em formatos vetoriais como SVG ou PDF, ou em formato matricial como JPEG.

4.2. Obtenção do conjunto de dados

Foram buscadas na Internet as combinações “*gephi twitter*” e “*gephi facebook*” a fim de obter um tutorial para extrair grande quantidade de dados de uma rede social e, em seguida, importá-los para gerar visualizações no *Gephi*. O aplicativo *Netvizz* permite extrair dados dos amigos do *Facebook*. Ao acessar a página do aplicativo, quatro opções de extração de dados são oferecidas: (i) **Rede pessoal:** extrai dados sobre os “amigos” e a conectividade (“relações de amizade”) entre eles; uma busca por “amigos” em comum

é feita⁶¹. (ii) **Dados de grupo:** cria redes e tabelas para mostrar “amizade” e interações (fluxos de envios de textos, fotos ou vídeos) entre participantes em um grupo. Para atender à política de privacidade do *Facebook*, o *Netvizz* substitui o nome dos usuários por uma sequência de números hexadecimais; (iii) **Rede de “curtidas” em páginas:** cria uma rede de páginas conectadas pelos “curtidas” que receberam; (iv) **Dados de página:** cria tabelas de dados com atividades do usuário em relação aos compartilhamentos (posts) de uma página específica.

O vídeo introdutório à utilização do *Netvizz*⁶² detalha os diversos modos de extração e como utilizar os dados obtidos no *Gephi*.

O primeiro teste foi feito em novembro de 2014 com uma comunidade bastante segmentada e homogênea, “Gay Indie Lovers”. É um grupo no qual a participação depende de aprovação do administrador (fechado) e é formado por aproximadamente cinco mil pessoas gays que gostam do estilo musical Indie, também conhecido como rock independente. O que se buscou verificar aqui foi a coesão de um grupo considerado homogêneo.

Seguindo os procedimentos do tutorial de extração de dados, o *Netvizz* demorou cerca de cinco minutos para compilar as informações dessa comunidade em uma tabela. O arquivo de saída desse processo vem pronto para ser lido pelo *Gephi*, sendo necessário apenas baixá-lo e abri-lo no programa.

⁶¹ Em março de 2015, devido a uma mudança na API do *Facebook* (interface de programas que ligam o *Netvizz* ao *Facebook*), a extração de dados de uma rede pessoal com a identificação de cada nó da rede por nome e sobrenome não é mais possível por conta da nova política de privacidade daquela rede social.

⁶² Disponível em <https://www.youtube.com/watch?v=XxH0Tm8NXik>. Acesso em 03.04.2015.

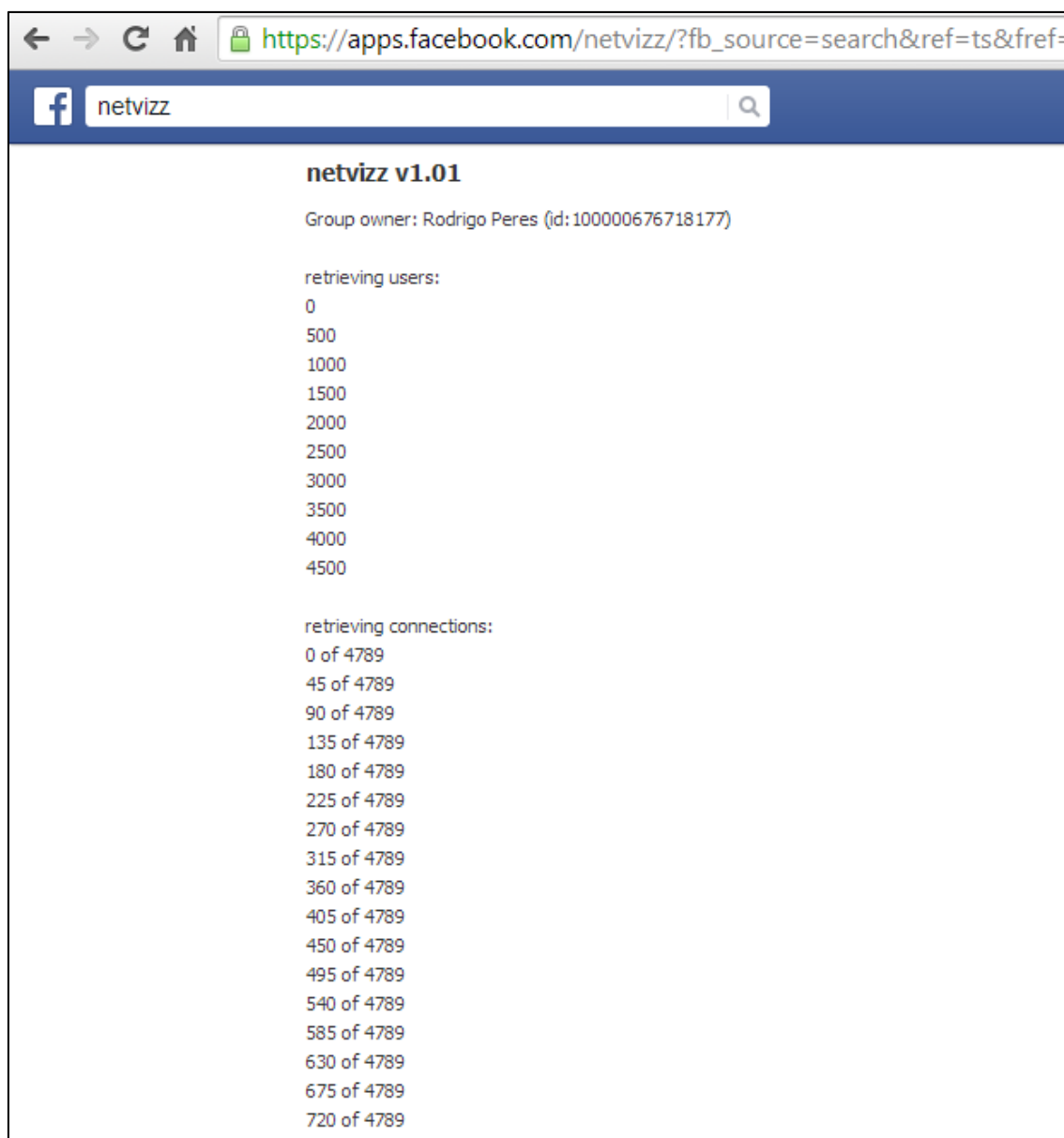


Figura 32 - O aplicativo do Facebook, Netvizz, processando informações de usuários e suas respectivas conexões para gerar o arquivo de saída a ser lido pelo Gephi.

O pesquisador que abre o arquivo resultante no *Gephi* pode se sentir decepcionado por não encontrar já prontos os agrupamentos e ligações entre pessoas. Em contrapartida, o programa oferece algoritmos de agrupamento na janela “Distribuição”, que se localiza em “Visão Geral”. Todos os algoritmos foram testados, gerando visualizações diferentes, cada uma delas trazendo um *insight* distinto sobre os dados. Entre eles, o mais útil para “desembaraçar” os nós e as arestas e evidenciar os agrupamentos foi “Force Atlas 2”; igualmente interessante foi “Circle Layout”, que elabora a classificação por graus. As imagens a seguir mostram como se deu a obtenção desta primeira visualização utilizando o algoritmo “Force Atlas”.

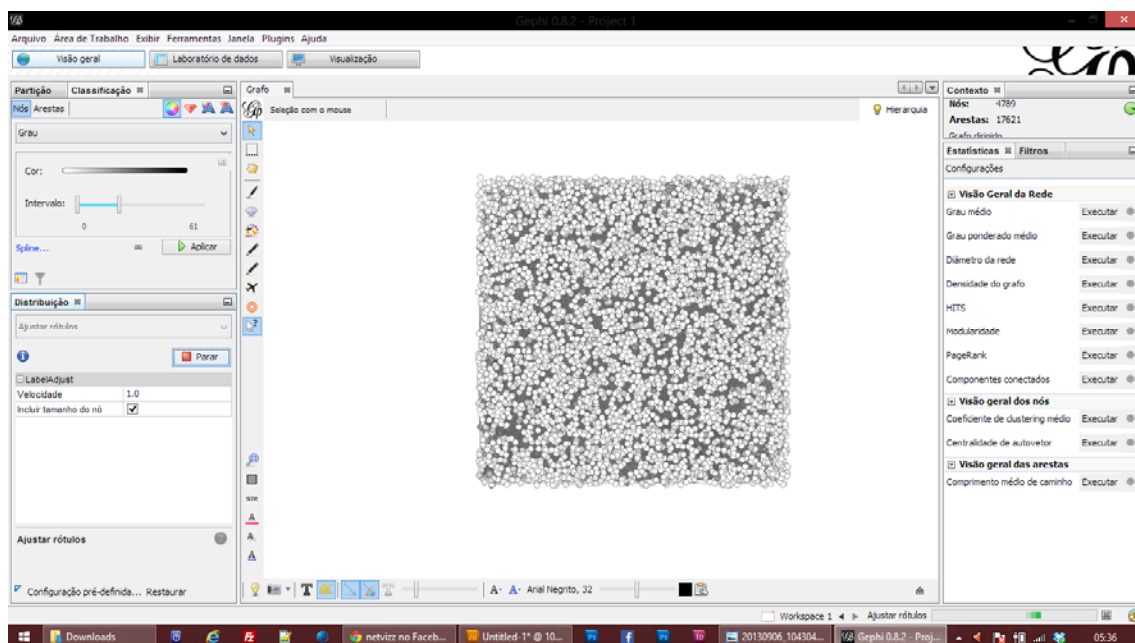


Figura 33 - Carregamento da comunidade “Gay Indie Lovers” mostra indivíduos e relações entre eles antes da aplicação de um algoritmo para agrupar as subcomunidades.

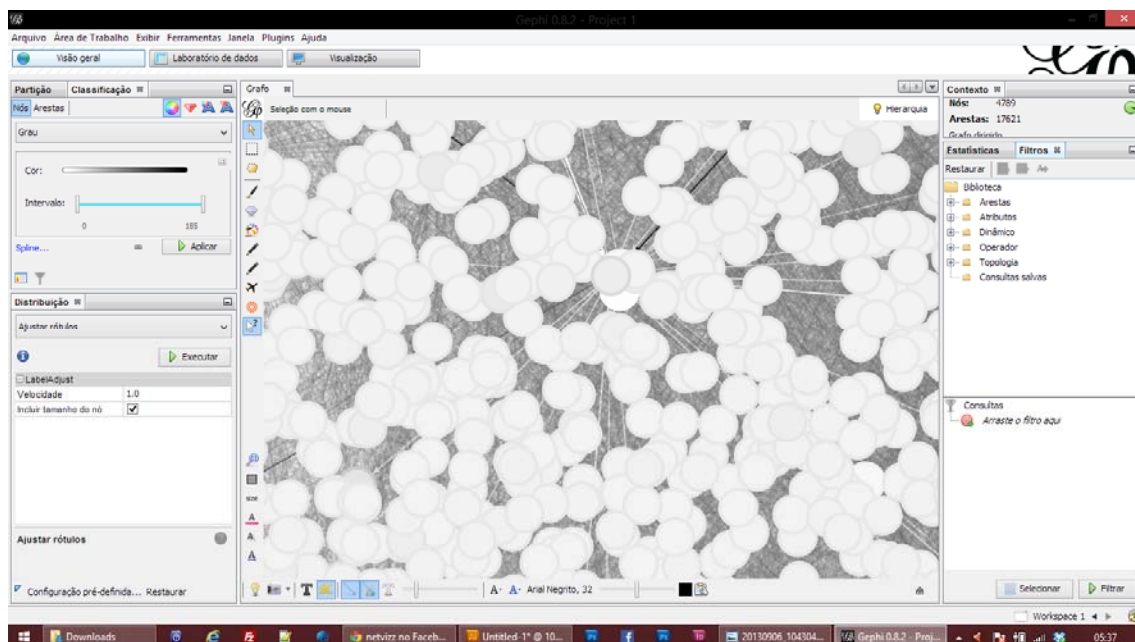


Figura 34 - Detalhe da imagem anterior, mostrando nós e arestas sobrepostos, “embaraçados”.

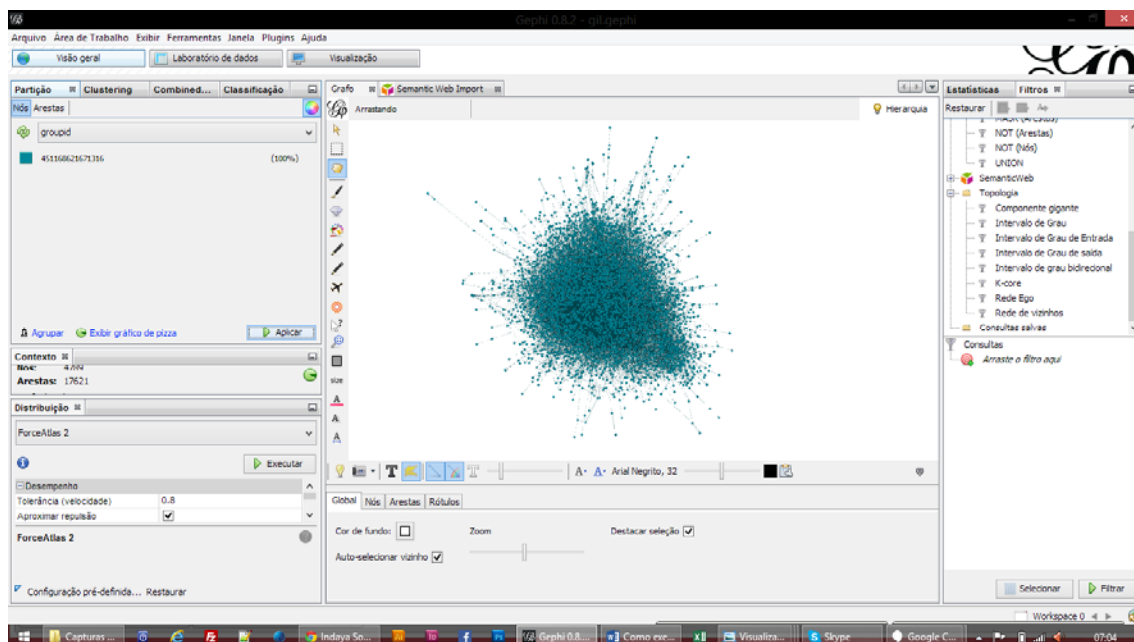


Figura 35 - Depois da aplicação do algoritmo de distribuição “Force Atlas 2”, alguns agrupamentos surgem à margem da massa de nós.

Neste primeiro teste, o algoritmo foi executado por apenas alguns minutos, o que resultou na imagem 4, com pouca formação de grupos de modularidade. Os pontos fora da grande massa representam pessoas com menos conectividade entre os membros. Portanto, poder-se-ia dizer que esta é uma comunidade coesa do ponto de vista unirrrelacional. Para dissuadir os nós, seria necessário atribuir a cada membro do grupo parâmetros que os distinguíssem da massa.

Continuando a exploração dessa comunidade, foi descoberta uma maneira de atribuir cor aos nós. Na opção “Nós”, dentro da seção “Partição”, encontra-se a ferramenta “Mapa de calor”, que aplica uma gradação de cores de acordo com o grau de separação entre os membros do grupo. Ao escolher um dos nós no gráfico para, a partir dele, explorar o grau de separação entre os membros, obtemos a seguinte visualização:

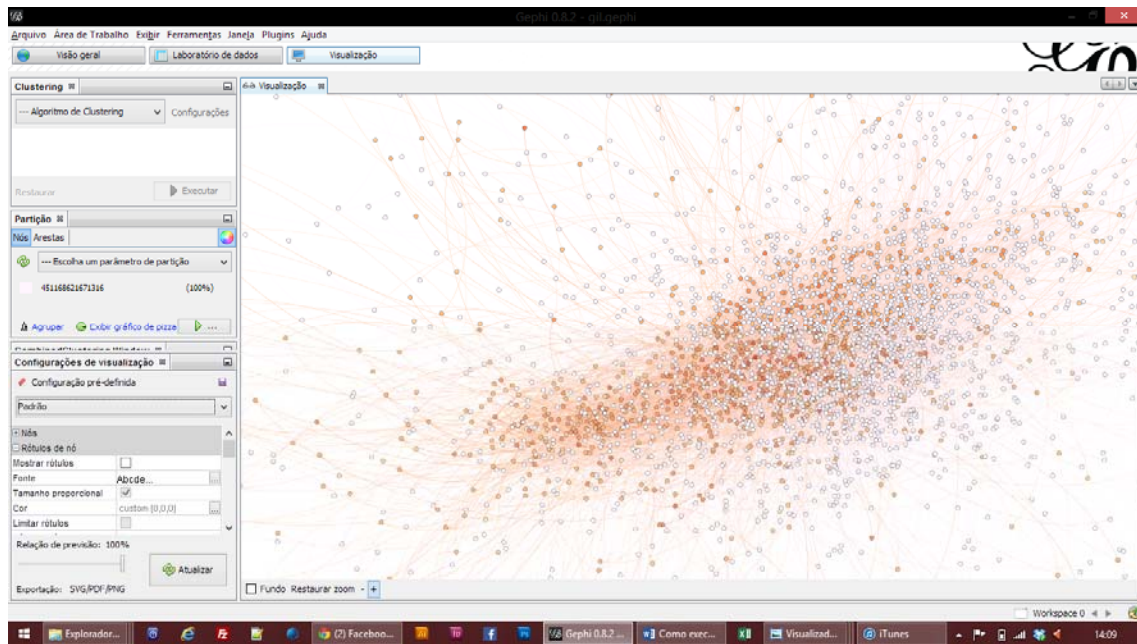


Figura 36 - Aplicação do algoritmo de mapa de calor.

De acordo com a imagem, os pontos de cor laranja são pessoas com grau de separação menor em relação ao ponto escolhido, em comparação com o resto da comunidade. Não é possível identificar as pessoas devido à política de privacidade do *Facebook*, o que dificulta uma análise mais profunda, já que exige uma das dimensões dos dados. Na tabela de origem, os dados dos membros do grupo estão cifrados, como se vê nesta captura de tela:

Nodes	Id	Label	sex	locale	groupid
user_099f02b7ee903b066ccf732c20966544a0ef7d4	099f02b7ee903b066ccf732c20966544a0ef7d4	user_099f02b7ee903b066ccf732c20966544a0ef7d4	male	pt_BR	451168621671316
user_a5db2163635d7b81b668e6cf0697ba20062e976	a5db2163635d7b81b668e6cf0697ba20062e976	user_a5db2163635d7b81b668e6cf0697ba20062e976	female	pt_BR	451168621671316
user_16743f501e7e3caaf59c0b5329b31f1f19326d8	16743f501e7e3caaf59c0b5329b31f1f19326d8	user_16743f501e7e3caaf59c0b5329b31f1f19326d8	male	pt_BR	451168621671316
user_b2b0f2c5dbb69de5aae93910fb099db9e95796a	b2b0f2c5dbb69de5aae93910fb099db9e95796a	user_b2b0f2c5dbb69de5aae93910fb099db9e95796a	male	en_US	451168621671316
user_9638f0c66332eb8813a3692ee6e82161ef88f8a	9638f0c66332eb8813a3692ee6e82161ef88f8a	user_9638f0c66332eb8813a3692ee6e82161ef88f8a	male	pt_BR	451168621671316
user_016af7a79b012d1d32690c2711016396de7ed	016af7a79b012d1d32690c2711016396de7ed	user_016af7a79b012d1d32690c2711016396de7ed	male	en_US	451168621671316
user_d3b36d3b3f3061f4109cab399b46401024036a4	d3b36d3b3f3061f4109cab399b46401024036a4	user_d3b36d3b3f3061f4109cab399b46401024036a4	male	pt_BR	451168621671316
user_cc708952d11ccaff136b77f68169282a0c4820cf5	cc708952d11ccaff136b77f68169282a0c4820cf5	user_cc708952d11ccaff136b77f68169282a0c4820cf5	male	pt_BR	451168621671316
user_8746c9d6e4887aa6ee735061c5ec543c76d3	8746c9d6e4887aa6ee735061c5ec543c76d3	user_8746c9d6e4887aa6ee735061c5ec543c76d3	male	pt_BR	451168621671316
user_74e9d0b1d5896756fc9e4e520661054100f57cac	74e9d0b1d5896756fc9e4e520661054100f57cac	user_74e9d0b1d5896756fc9e4e520661054100f57cac	male	pt_BR	451168621671316
user_3f5f7e91eb1302114c0e7f8b021f928102df651c	3f5f7e91eb1302114c0e7f8b021f928102df651c	user_3f5f7e91eb1302114c0e7f8b021f928102df651c	male	en_US	451168621671316
user_f1d010029464070e7781b3248194031254673a	f1d010029464070e7781b3248194031254673a	user_f1d010029464070e7781b3248194031254673a	male	pt_BR	451168621671316
user_b0c1100f32486cd30135ce54f66696f8de22bd	b0c1100f32486cd30135ce54f66696f8de22bd	user_b0c1100f32486cd30135ce54f66696f8de22bd	male	pt_BR	451168621671316
user_6b0ac95026afdb0ba3ebf6430bb2e7b07a0d8902	6b0ac95026afdb0ba3ebf6430bb2e7b07a0d8902	user_6b0ac95026afdb0ba3ebf6430bb2e7b07a0d8902	male	en_US	451168621671316
user_b143932350510d5001a70be220348f3d0db9ee1	b143932350510d5001a70be220348f3d0db9ee1	user_b143932350510d5001a70be220348f3d0db9ee1	male	pt_BR	451168621671316
user_30e573a50d944d8081030425fa0fa30f970b966d	30e573a50d944d8081030425fa0fa30f970b966d	user_30e573a50d944d8081030425fa0fa30f970b966d	male	pt_BR	451168621671316
user_b0d166865331d983ae947a0daade147a8ec4f6a1	b0d166865331d983ae947a0daade147a8ec4f6a1	user_b0d166865331d983ae947a0daade147a8ec4f6a1	male	en_US	451168621671316
user_1f910e56043d1e654f73b5d4a636f0139717f6b	1f910e56043d1e654f73b5d4a636f0139717f6b	user_1f910e56043d1e654f73b5d4a636f0139717f6b	male	pt_BR	451168621671316
user_a09f6d6c544f20e07314b16237e9998f5e5d1c	a09f6d6c544f20e07314b16237e9998f5e5d1c	user_a09f6d6c544f20e07314b16237e9998f5e5d1c	male	pt_BR	451168621671316
user_af1a6c530641827667db1b06780994e48e124e5	af1a6c530641827667db1b06780994e48e124e5	user_af1a6c530641827667db1b06780994e48e124e5	male	pt_BR	451168621671316
user_ae706ee669790121e3b6d973b547e243541587	ae706ee669790121e3b6d973b547e243541587	user_ae706ee669790121e3b6d973b547e243541587	male	pt_BR	451168621671316

Figura 37 - Estrutura dos dados para a visualização da comunidade.

Os nomes dos membros de um grupo só são disponibilizados quando se utiliza uma tabela de dados proveniente da extração de informações da própria rede de dados. Assim, procedeu-se ao modo de extração “rede pessoal” no *Netvizz*. O conjunto de dados obtido, de 695 nós, foi carregado no *Gephi*, que calculou automaticamente a relação entre as pessoas expressa pelo número de arestas: 5041. O procedimento para obter a visualização deste segundo teste foi o mesmo aplicado no primeiro teste. Porém, já a primeira aplicação do algoritmo “Force Atlas 2” resultou em agrupamentos bem distintos: cada um deles representa comunidades naturais, já que apresentam um grau de separação pequeno - ou seja, há a tendência de que a comunidade natural seja refletida na visualização de dados, porque membros de comunidades naturais se adicionam nas redes sociais. O fato dos nomes serem revelados na tabela de dados e apresentados na visualização ajuda a identificar as comunidades naturais. Na imagem seguinte, é visualizada uma comunidade de pessoas que se conhecem pessoalmente, por serem alunos ou professores da PUC-SP:

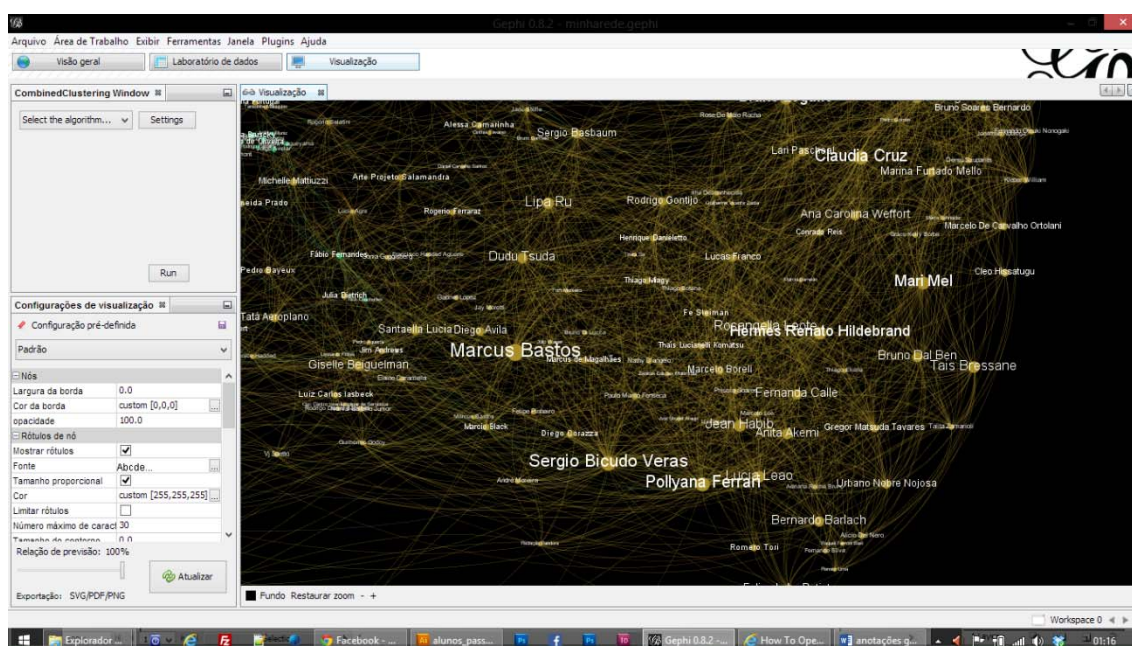


Figura 38 - Professores e alunos da PUC-SP.

Nesta visualização, o grau de conectividade afeta o tamanho da fonte usada para indicar o nome de cada pessoa: aqueles cujos nomes aparecem maiores são os que, nesta rede específica, apresentam maior conectividade. A partir desta visualização, contudo, não é possível afirmar que as pessoas com mais conectividade são as mais influentes na

comunidade, ou que compartilham conteúdo. Daí a necessidade de basear a visualização em dados multirrelacionais, ou encontrar um algoritmo de organização melhor, para verificar essa influência com mais fidelidade. Na captura de tela seguinte, por exemplo, vê-se mais um agrupamento, ou *cluster*, de pessoas que se conheceram pessoalmente em uma experiência de intercâmbio, tendo formado uma comunidade natural temporária.

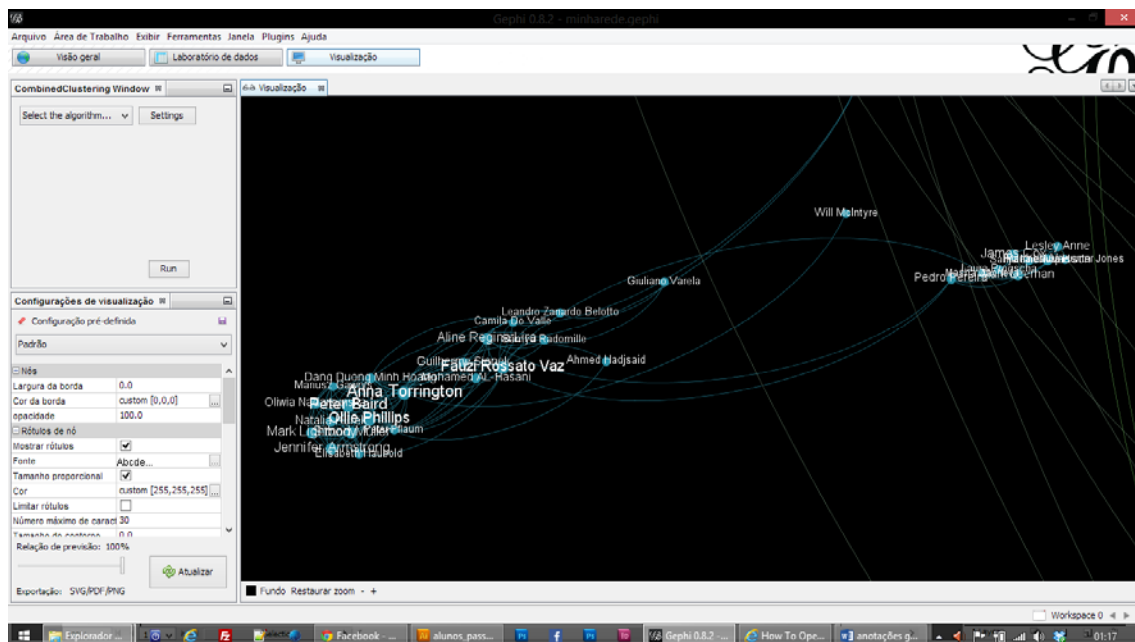


Figura 39 - Agrupamento de alunos de intercâmbio

Na imagem acima, pode-se observar que pessoas se aglutinaram por conta da nacionalidade. Os ingleses ocupam a região mais à esquerda. Em seguida, um pouco mais à direita, estão os brasileiros. À extrema direita, estão pessoas não agrupadas por nacionalidade, mas que eram visitantes de quem participava do intercâmbio.

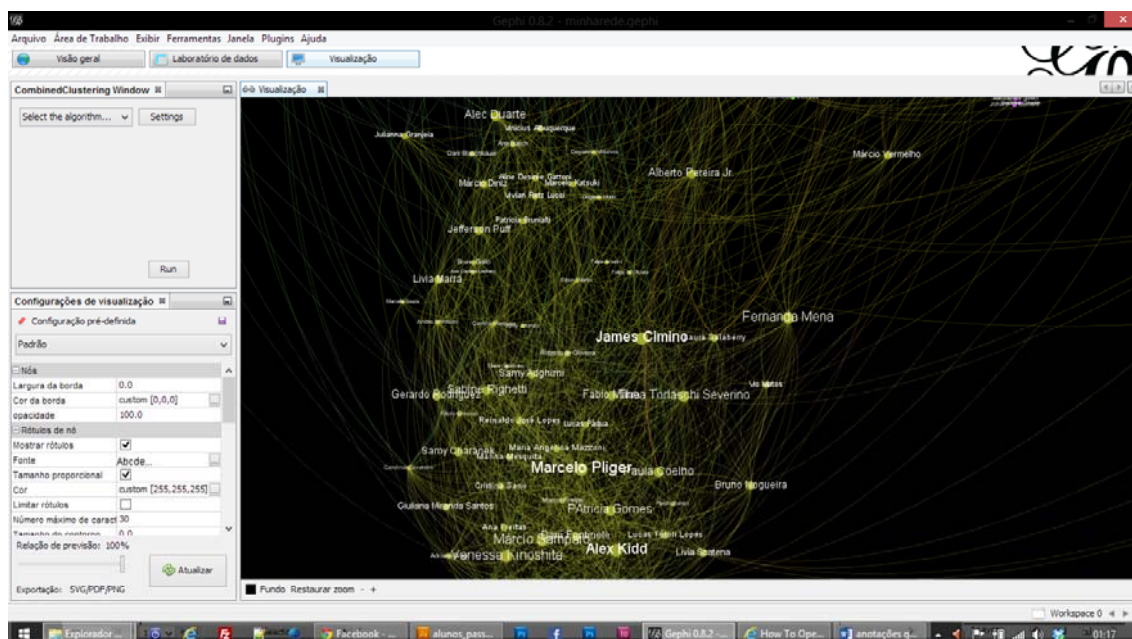


Figura 40 - Agrupamento de designers do jornal Folha de S. Paulo presentes em nossa rede de contatos

Esta leitura, baseada na topologia da rede, depende de um conhecimento prévio sobre os membros da comunidade. Para tornar a avaliação mais objetiva, seria necessário aplicar um questionário direcionado àquilo que se deseja analisar na relação entre as pessoas. Para que os dados provenientes de um questionário pudessem ser aproveitados no *Gephi*, seria preciso estruturá-los em forma de tabela, mas não apenas isto: também compreender a melhor forma de construir as relações de dados entre os membros do grupo, já que isto não é calculado automaticamente pelo programa de produção de visualizações.

O teste de estruturação de dados e construção de relações (“arestas”, no vocabulário do *Gephi*) se deu na construção da visualização “Partidos políticos do Brasil, por afinidade”⁶³. Buscou-se relacionar os partidos por ideologia e espectro político com base em uma tabela encontrada online⁶⁴. Observando a construção das arestas nos dados obtidos para análise de comunidades no *Facebook*, foi construída manualmente uma tabela que resultou em 521 linhas. O algoritmo é “Partido A se conecta ao Partido B por ter a ideologia X”.

⁶³ Disponível no formato interativo em <http://bit.ly/1NEbe36> e em PDF <http://bit.ly/1E9343F>

⁶⁴ <http://bit.ly/1b9LiAO> . Acesso em 01.02.2015.

Source	Target	Type	Id	Label
Partido Republicano Brasileiro (PRB)	Partido do Movimento Democrático Brasileiro (PMDB)		1001	Centrismo
Partido Republicano Brasileiro (PRB)	Partido Social Liberal (PSL)		1002	Centrismo
Partido Republicano Brasileiro (PRB)	Partido Humanista da Solidariedade (PHS)		1003	Centrismo
Partido Republicano Brasileiro (PRB)	Partido Trabalhista Brasileiro (PTB)		1004	Centrismo
Partido Republicano Brasileiro (PRB)	Partido Progressista (PP)		1005	Centrismo
Partido Republicano Brasileiro (PRB)	Partido da República (PR)		1006	Centrismo
Partido Republicano Brasileiro (PRB)	Partido Republicano Progressista (PRP)		1007	Centrismo
Partido Republicano Brasileiro (PRB)	Partido Social Democrático (PSD)		1008	Centrismo
Partido Republicano Brasileiro (PRB)	Partido Ecológico Nacional (PEN)		1009	Centrismo
Partido Republicano Brasileiro (PRB)	Partido Trabalhista do Brasil (PTdoB)		1010	Centrismo
Partido Republicano Brasileiro (PRB)	Partido Trabalhista Nacional (PTN)		1011	Centrismo
			1012	Centrismo
Partido do Movimento Democrático Brasileiro (PMDB)	Partido Social Liberal (PSL)		1013	Centrismo
Partido do Movimento Democrático Brasileiro (PMDB)	Partido Humanista da Solidariedade (PHS)		1014	Centrismo
Partido do Movimento Democrático Brasileiro (PMDB)	Partido Trabalhista Brasileiro (PTB)		1015	Centrismo
Partido do Movimento Democrático Brasileiro (PMDB)	Partido Progressista (PP)		1016	Centrismo
Partido do Movimento Democrático Brasileiro (PMDB)	Partido da República (PR)		1017	Centrismo
Partido do Movimento Democrático Brasileiro (PMDB)	Partido Republicano Progressista (PRP)		1018	Centrismo
Partido do Movimento Democrático Brasileiro (PMDB)	Partido Social Democrático (PSD)		1019	Centrismo

Figura 41 - Captura de tela da tabela de dados sobre partidos políticos

Em seguida, o *Gephi* foi alimentado com essa tabela e foram aplicados os algoritmos “Circular Layout” e “Mapa de calor”, resultando nos arquivos já mencionados anteriormente, que denotam uma topologia das relações dos partidos de acordo com sua ideologia.

Isto posto, seria possível aumentar a fidelidade da visualização da relação entre membros de uma comunidade com a aplicação de um questionário. Os dados coletados seriam usados para preencher uma nova coluna da tabela, obtendo uma visualização mais afinada. No caso de redes muito extensas, o questionário (figura 3) pode ser aplicado apenas em um subgrupo bem delimitado pela hipótese que se queira testar, ou nas modularidades identificadas na visualização dos dados. Com o gráfico resultante, podem-se detectar padrões de agrupamento menos evidentes do que aqueles baseados na mera observação de campo ou na análise de tabelas complexas. A visualização das relações apresentadas no formato de mapa transcende a visão incompleta daquele que observa as relações sociais, elevando-a a uma espécie de onisciência.

4.3. Influenciadores e viralização

A seguir, analisamos um conjunto de dados⁶⁵ pequeno e antigo, criado pelo professor alemão Johannes Delitsch entre 1880 e 1881, para analisar a relação entre popularidade e notas altas dentro de uma sala de aula. No sistema alemão do final do século XIX, os melhores alunos eram colocados nas mesas mais próximas à do professor, enquanto os alunos com pior desempenho deveriam ocupar as mesas ao fundo da sala de aula. O que salta aos olhos do professor, e o que ele pretende analisar, é o arranjo hierárquico que se forma na sala. Imagina-se que ela obedeça à mesma ordem usada para distribuir os alunos nas mesas, mas o que se observa na visualização de dados é que não é necessário ser um excelente aluno para ser “popular”, como veremos a seguir.

Depois de termos baixado o conjunto de dados e procedido da mesma forma pela qual tratamos os dados da comunidade do *Facebook*, foi aplicado o algoritmo de distribuição *Double Circle Layout*. Indicamos com cores mais escuras um desempenho escolar melhor, enquanto as mais claras indicam que o rendimento daquele aluno não foi satisfatório. O conjunto de dados compreende as seguintes colunas (ou atributos, parâmetros): a) “Lokationposition”: posição do aluno, de acordo com seu rendimento; b) “Handicapped”: se o aluno possui algum problema físico; c) “Repeater”: repetentes; d) “Sweetsgiver”: quem compartilha doces com seus colegas. O mapa de afinidades resultante, que pode ser melhor visualizado online⁶⁶, é o seguinte:

⁶⁵ Disponível em <http://goo.gl/IoMWPM>. Acessado em 01.05.2015.

⁶⁶ Disponível em <http://goo.gl/WNQSVh>. Acessado em 01.05.2015.

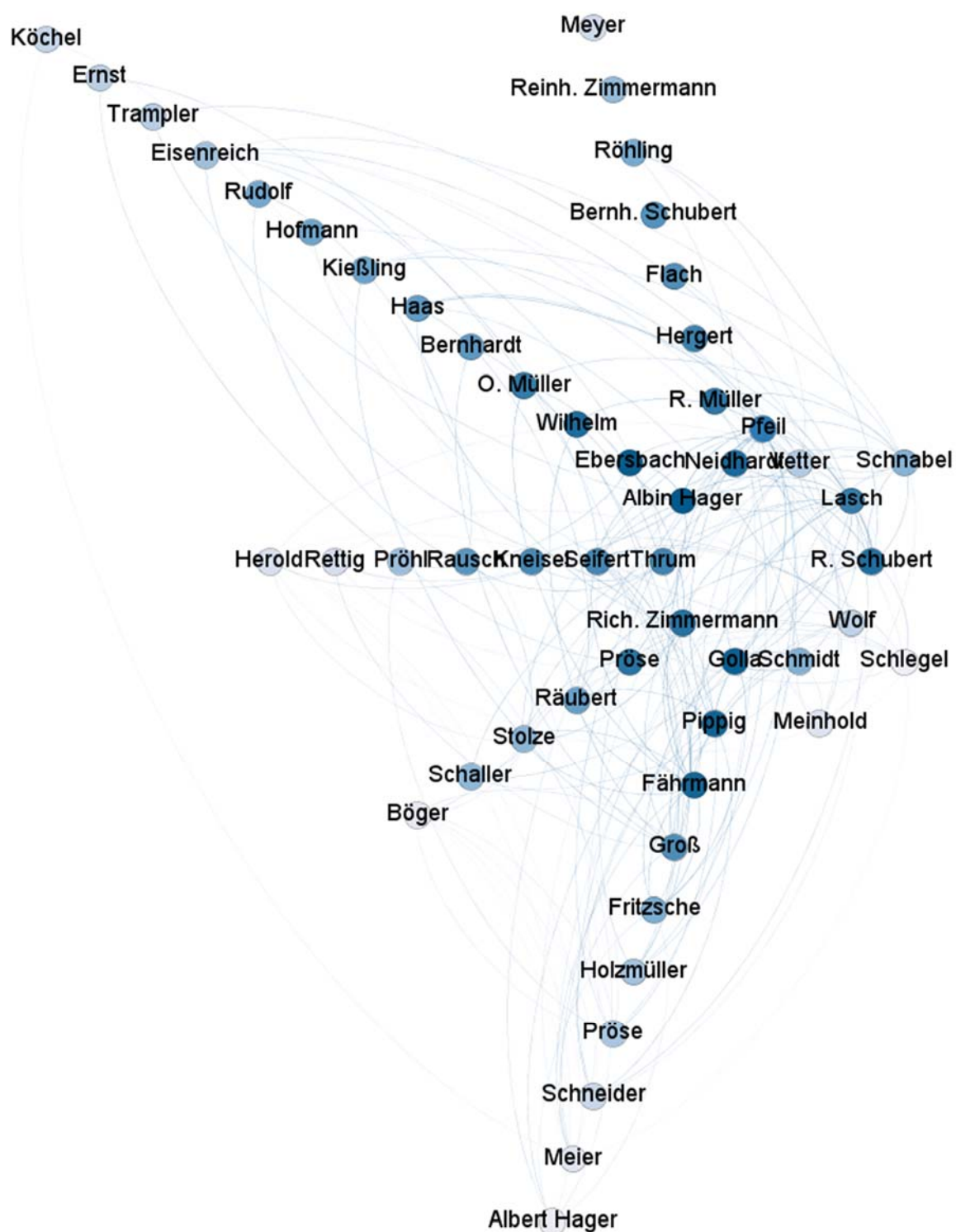


Figura 42 - Afinimapa da classe de 1880-1881

Como apontam Heidler et al (2014), um gráfico baseado no conjunto de dados apresentado revelaria a estratificação na sala de aula, o que significaria que existe uma

tendência a que os alunos tenham colegas de acordo com seu próprio desempenho. Outro aspecto observado é que os que apresentam melhor desempenho são os mais “populares” daquele grupo.

[...] os dados mostram padrões estruturais das redes de amizade similares aos encontrados hoje (reciprocidade, fechamento triádico transitivo). Além disso, testamos a influência da classificação do desempenho dos alunos (*Lokationsprinzip*), que distribui os alunos na sala de aula de acordo com o desempenho de cada um. Esse ranking produz uma hierarquia na popularidade dos alunos, pelos laços de amizade correspondentes que vão em direção ao topo da hierarquia. Neste aspecto, no que tange o efeito do desempenho escolar na popularidade, encontramos forte estratificação, o que não é sempre prevalecente hoje.⁶⁷ (Heidler et al, 2014, tradução nossa)

O que o estudo não comenta é que, apesar de possuírem desempenho escolar modesto, os alunos Wolf, Schnabel e Schmidt estão ao centro do gráfico, o que significa que ocupam uma posição alta na escala de popularidade daquele grupo. Além disso, esses alunos têm muitas conexões em estratos diferentes do mapa – são os indivíduos que transitam. Schmidt é amigo até mesmo de Meinhold, que é um indivíduo mais isolado, mas não tanto quanto Meyer, que não tem nenhuma relação com o grupo. Estas observações nos conduzem a questões sobre redes sociais muito discutidas hoje.

A questão mais importante que se pode levantar é a da influência: Schmidt é representado como um nó superconectado na rede, mas isso significa que seja um aluno influente? Se este grupo de alunos fosse contemporâneo nosso, seria Schmidt aquele que propagaria com mais facilidade um vídeo do *YouTube*, ou que teria mais inscritos naquela plataforma? Seus *tuítes* seriam *retuitados*? Apesar de não termos como saber exatamente o que se passaria, tem havido algumas descobertas sobre o comportamento dos influenciadores.

⁶⁷ [...] unmarked data show structural patterns of friendship networks similar to today (reciprocity, transitive triadic closure). Moreover we test the influence of the class ranking order (*Lokationsprinzip*), which allocates the pupils in the class room according to their school performance. This ranking order produces a hierarchy in the popularity of pupils, through hierarchy–congruent friendship ties going upwards in the hierarchy. In this respect, concerning the effect of school achievement on popularity, we find a strong stratification, which is not always prevalent today.

Lemos (2015) pergunta se os fenômenos virais da rede não seriam resultado da mudança de comportamento causada por influenciadores. Ele recapitula algumas pesquisas conduzidas no século XX para tentar explicar o binômio comportamento e influência, como aquela de Stanley Milgram que ficou famosa por disseminar a ideia dos seis graus de separação. Voluntários de Boston e Nebraska deveriam enviar cartas para pessoas-alvo apenas por intermédio de conhecidos. O resultado desse experimento, executado em 1967, indicou que o número médio de intermediários pelos quais as mensagens passaram até os destinatários foi de 5,2 pessoas – e que cerca de metade das mensagens entregues haviam passado por apenas 3 pessoas. “Existiria um pequeno grupo de pessoas especiais que são conectadas a muita gente (os 'influenciadores'). E todo o restante dependeria deles para estabelecer contatos diretos.”, nos conta Lemos.

O experimento de Milgram foi repetido em 2008, desta vez com dados de 240 milhões de usuários do MSN Messenger, usando apenas a matemática, e o número médio de intermediários averiguado foi por volta de 7 pessoas. Lemos destaca que, em um experimento de 2003 baseado na pesquisa de Milgram com um grupo de 20 mil pessoas, dirigido pelo cientista social Duncan Watts, nenhum superconector havia aparecido no estudo – e que “pessoas normais são tão capazes de disseminar informação na Internet quanto as consideradas especiais”.

Watts sintetiza assim o papel dos influenciadores: de fato existem indivíduos que são mais capazes de disseminar informações na rede (como alguém com muitos seguidores). Mas mais importante do que o número de seguidores é a posição que cada pessoa ocupa. Sua influência só vai ser efetiva se ela estiver conectada às pessoas "certas". (Lemos, 2015)

Com esta observação, podemos voltar a o caso de Schmidt, aluno da escola alemã de 1880. O fato de ele estar conectado a muitos colegas e a estratos diferentes o torna tão importante na hierarquia das afinidades que ele tem o mesmo potencial de disseminação de informação do que seus companheiros com desempenho escolar superior. Transpondo isto para uma rede social contemporânea, o que tornaria alguém um influenciador não é o número de seguidores, mas o grau de conectividade com as “pessoas certas”, aquelas que

possam entregar uma mensagem com mais eficiência a um destinatário, ou um público-alvo.

Khan e Sokha (2014) mapearam os cem vídeos mais vistos do *YouTube* usando a API daquela rede social de vídeos e recuperaram os seguintes parâmetros de cada um deles, organizados em quatro categorias⁶⁸: a) **Capital social e não social do usuário**: idade, gênero, total de visualizações, número de inscritos, data de cadastro no *YouTube*, total de vídeos postados; b) **Características dos vídeos**: data de envio, duração, categoria do vídeo; c) **Capital de rede externo**: *links* internos, número de acessos e d) **Viralidade**: “curtidas”, “descurtidas”, número de “favoritadas”, total de visualizações e total de comentários.

Khan e Sokha afirmam que o “capital social” é formado pela estrutura social e pelas relações sociais. Para esclarecer este ponto de vista, adotam a definição de Chow e Chan (2008) que, em resumo, pontua que o capital social é composto de três dimensões: estrutural, relacional e cognitiva.

A dimensão estrutural, segundo eles, é formada por dois aspectos: pela conectividade e pela integração. Enquanto a conectividade seria, por exemplo, a quantidade de seguidores, a integração está relacionada à constância da comunicação dentro da rede⁶⁹. Acreditamos que a visualização da rede permite a análise rápida da dimensão estrutural. A ligação entre pessoas é um dos fatores que motivam o comportamento recíproco e a interação interpessoal, o que representa o fator relacional da rede (Camarero e San José, 2011 apud Khan e Sokha, 2014, p. 7). Entre um usuário com uma rede forte, aspectos estruturais e relacionais bem estabelecidos e outro usuário com fama, o primeiro tem mais poder de viralização de vídeos, pois a posição ocupada na

⁶⁸ We constructed and tested an empirical model to understand the relationship among users’ social and non-social capital—user age, gender, view count, subscriber, join date, total videos posted, video characteristics—postdate, duration, and video category, external network capital— in-links and hit counts, and Virality—likes, dislikes, favorite count, view count, and comment count. (Khan e Sokha, 2014, p. XX)

⁶⁹ Structural social capital contains two imperative properties of individuals’ position within the network such as connectedness and integration. Connectedness property refers to the level of associating with others, while integration property focuses on level of communication with others within the network.

estrutura e o tamanho da rede social, de acordo com Khan e Sokha, são a base do marketing viral (idem). O estudo conclui que a popularidade dos vídeos:

[...] não é apenas uma função do sistema do YouTube per se, mas que a dinâmica da rede (links internos e número de acessos) e capital social off-line (base de fãs e fama) são cruciais no fenômeno da viralização, afetando particularmente o número de visualizações do conteúdo⁷⁰.
(Khan e Sokha, 2014, p. 20).

Lemos (2014) vai ao encontro do estudo sul-coreano, afirmando que os indivíduos que transitam em uma rede são aqueles com maior poder de propagação rápida de uma mensagem. Porém, é necessária uma pesquisa em diferentes redes sociais para averiguar a validade deste argumento em contextos diferentes.

Com isso, fica a dica, marqueteiros: se a ideia é gerar um "viral", melhor do que valer-se de alguns poucos indivíduos com milhões de seguidores, é melhor contar com um grupo de influenciadores "médios", que ocupam diferentes posições na rede e falam para pessoas diversas. É mais fácil pôr fogo em uma floresta com vários focos "médios" de incêndio em lugares distintos do que com alguns grandes focos.
(Lemos, 2014)

Com base na importância do grau médio, voltamos a executar um mapa de conexão da nossa rede de contatos do *Facebook*. As comunidades são definidas por cores que representam o grau de modularidade. Os nomes dos indivíduos que seriam possíveis influenciadores, devido à conectividade, receberam um tamanho de fonte proporcional à capacidade de propagação em uma comunidade específica. Os influenciadores de grau médio, por sua vez, têm peso de fonte também médio⁷¹. O resultado é melhor observável na versão online do mapa.

⁷⁰ “Among other findings, our results showed that popularity of the videos was not only the function of YouTube system per se, but that network dynamics (e.g., in-links and hits counts) and offline social capital (e.g., fan base and fame) play crucial roles in the viral phenomenon, particularly view count.” (Khan e Sokha, 2014, p. 20).

⁷¹ Disponível no formato SVG em <http://goo.gl/O5yTl5>. Acesso em 06.01.2016.

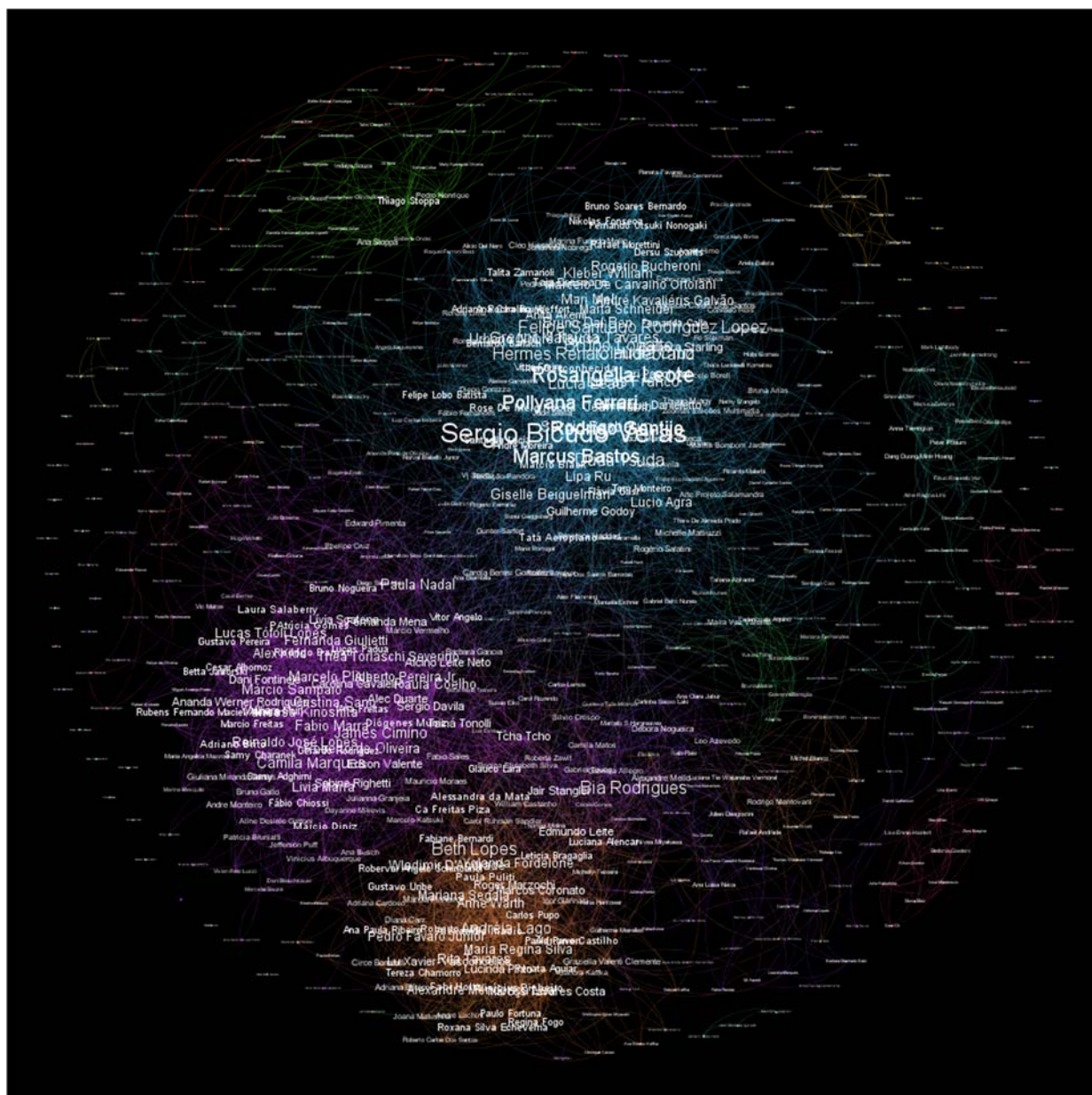


Figura 43 - Mapa destacando comunidades e influenciadores de grau médio

5. OBTENÇÃO DE DIFERENTES CONJUNTOS DE DADOS PARA CONSTRUIR MAPAS

Observamos, até aqui, a construção de visualizações com dados da rede de amigos do *Facebook* obtidos pelo *Netvizz*. Com este conjunto de dados pudemos demonstrar que a proximidade das pessoas nas comunidades naturais e artificiais encontra correspondência na visualização obtida, de acordo com o parâmetro da modularidade que foi calculado pelo *software* de visualização. Vimos ainda que é possível encontrar facilmente os influenciadores pelo grau médio de conexões apresentadas na rede e que, nas bordas da representação obtida, estão aqueles que possuem menos conexões e que se encontram mais distantes dos fluxos daquela rede, numa “calmaria”.

A técnica apresentada não é a única a ser utilizada na construção de tais mapas, apesar de ter servido para sustentar a hipótese de que a afinidade pode ser aferida a partir de uma visualização da rede. Propomos, então, uma visão geral a partir das diferentes técnicas que podem ser adotadas na obtenção de dados para construir mapas de afinidade. Elas consistem, de modo geral, nos mesmos elementos: um *crawler*, um armazenamento de dados, um programa de visualização.

O *crawler* se trata de um programa feito para obter dados de modo automatizado do mesmo modo utilizado pelo *Netvizz*, ou seja, extraíndo dados do *Facebook*. Dependendo da hipótese sobre uma comunidade, ou sobre a relação entre quaisquer tipos de entidades, que se queira testar, se constrói um *crawler* – também chamado de robô, *bot* ou *spider*. Os *crawlers* podem obter dados a partir da extração recursiva de *links* de um site ou de um banco de dados e, em seguida, estruturá-los e armazená-los em tabelas, seja em um banco de dados ou no sistema de arquivos local. É partindo desses dados estruturados que se geram as visualizações.

Suponhamos que se deseja criar uma visualização das relações do *WikiLeaks*. Na massa de dados não estruturados do *WikiLeaks*, Wild (2015) criou um método que contempla um programa para ler cada registro, no caso, cada *e-mail*. Para cada registro, foi criado um padrão no cabeçalho: data, destinatário, categoria da mensagem. Em seguida, procedeu-se deste modo:

5. Obtenção de diferentes conjuntos de dados para construir mapas

1. Dados carregados no *Postgres*; 2. Expressões regulares foram usadas para extrair e padronizar os nomes de cidades que são precedidos de “origem”, “destino” e “assunto”; 3. Foi usada a biblioteca *Python Geocoder* para relacionar os códigos de países às cidades; 4. Com base na lista de referência dos códigos de países às subregiões ou regiões, foi criada uma lista customizada de regiões – algumas das quais são subregiões únicas, outras são conjuntos de subregiões; 5. Expressões regulares foram usadas para extrair e padronizar os nomes de cidades; 6. A modelização de tópicos Gensim foi usada para construir um modelo baseado em sessenta assuntos, de acordo com o conteúdo das mensagens; 7. *Python* foi usado para efetuar todo o agrupamento, reunião de categorias e formatação dos dados em estruturas apropriadas em JSON ou CVS (Wild, 2015, tradução nossa)⁷².

Alguns *crawlers* são embarcados, como *plug-ins*, nos aplicativos geradores de visualizações, como o *Gephi*. Utilizamos um desses *plug-ins* para a conexão em nossa caixa de entrada de *e-mails*, usada sobretudo para assuntos acadêmicos, para gerar um gráfico que foi disponibilizado *online*, em alta resolução⁷³. O mapa resultante evidencia o propósito daquela conta por meio da verificação das pessoas com as quais foram trocadas mais mensagens, criando um aglomerado que representa uma realidade da comunicação: os pontos reunidos no canto inferior esquerdo representam um grupo com o qual elaborou-se um projeto para um dos cursos que frequentamos.

72 “1. Loaded into Postgres; 2. Regular expressions were used to extract and canonicalize city names in "origin", "destination", and "header"; 3. Python Geocoder library used to attach country codes to cities; 4. Based on a Google reference list mapping country codes to both subregions and regions, created a custom list of regions - some of which were single subregions, some conjunctions of subregions; 5. Regular expressions were used to extract and canonicalize city names; 6. Gensim topic modeling used to construct 60-topic model based on content of cables ; 7. General python used to do all other grouping, attaching of categories, and formatting of data into appropriate JSON or CSV structures”

73 <http://goo.gl/cLgnbA>

5. Obtenção de diferentes conjuntos de dados para construir mapas

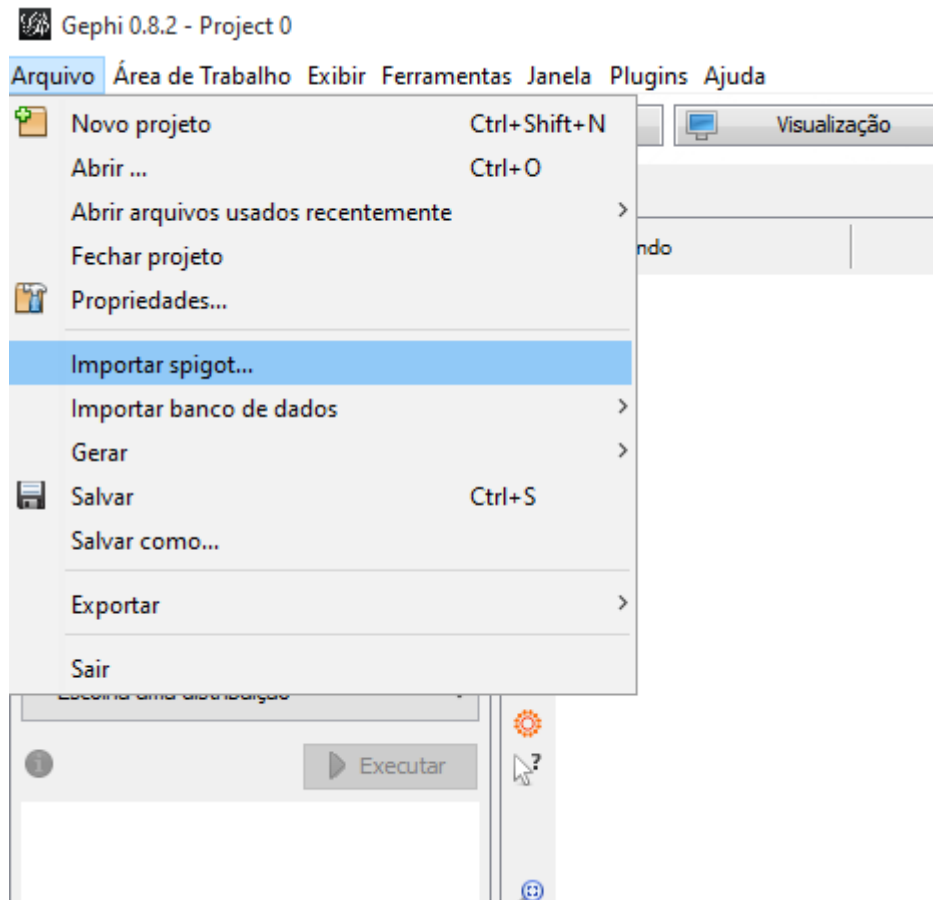


Figura 44 - Importação de dados IMAP com um plug-in do Gephi.

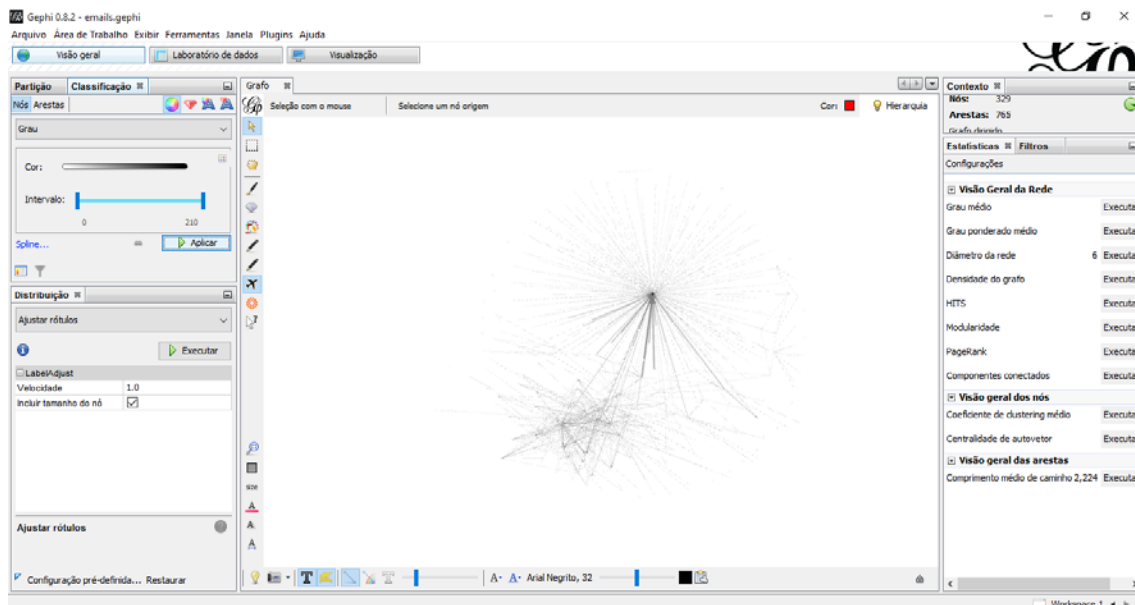


Figura 45 - Dados carregados no Gephi da nossa conta IMAP, pelo plug-in de importação.

5. Obtenção de diferentes conjuntos de dados para construir mapas

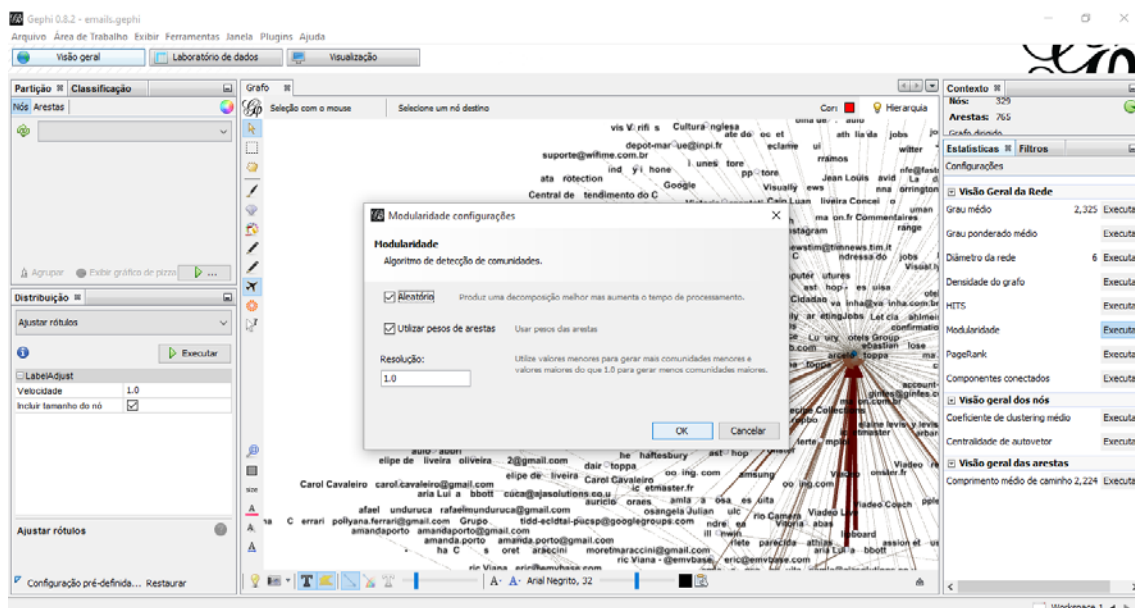


Figura 46 - Aplicação de modularidade nos dados IMAP para detectar comunidades.

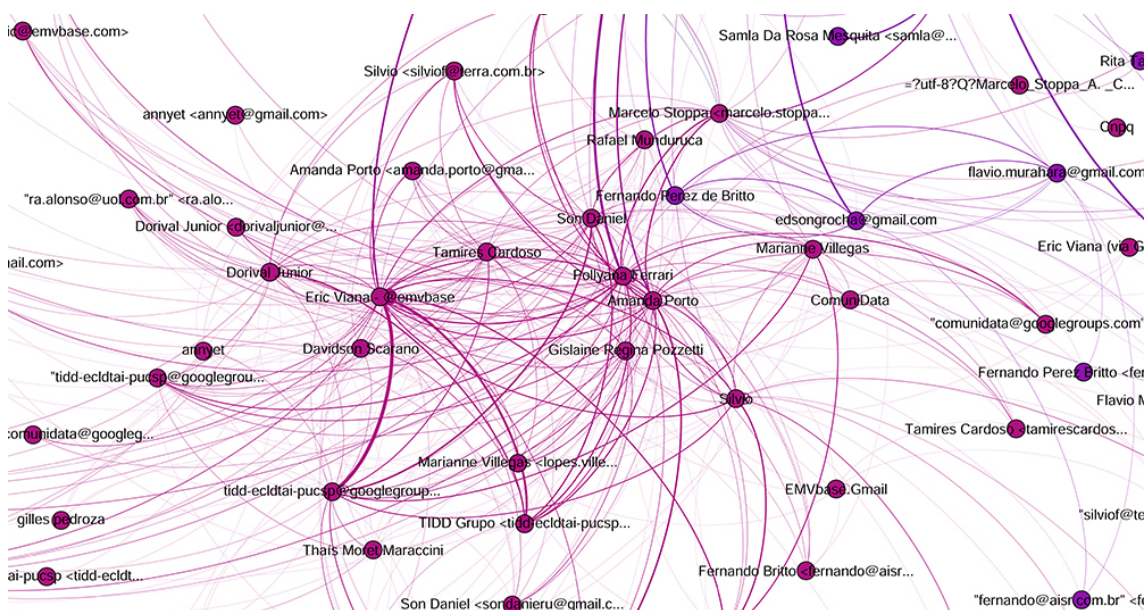


Figura 47 - Detalhe da visualização dos dados IMAP.

Se a massa de dados não for dada, é possível criar crawlers para que ela seja obtida. Isto pode ser feito com linguagens como Python, como procedeu Wild, ou Java e PHP74, entre outros. Como sistema de armazenamento, qualquer banco de dados pode ser

utilizado. Wild usou PostgreSQL⁷⁵, mas pode5. Obtenção de diferentes conjuntos de dados para construir mapas

Observamos, até aqui, a construção de visualizações com dados da rede de amigos do *Facebook* obtidos pelo *Netvizz*. Com este conjunto de dados pudemos demonstrar que a proximidade das pessoas nas comunidades naturais e artificiais encontra correspondência na visualização obtida, de acordo com o parâmetro da modularidade que foi calculado pelo *software* de visualização. Vimos ainda que é possível encontrar facilmente os influenciadores pelo grau médio de conexões apresentadas na rede e que, nas bordas da representação obtida, estão aqueles que possuem menos conexões e que se encontram mais distantes dos fluxos daquela rede, numa “calmaria”.

A técnica apresentada não é a única a ser utilizada na construção de tais mapas, apesar de ter servido para sustentar a hipótese de que a afinidade pode ser aferida a partir de uma visualização da rede. Propomos, então, uma visão geral a partir das diferentes técnicas que podem ser adotadas na obtenção de dados para construir mapas de afinidade. Elas consistem, de modo geral, nos mesmos elementos: um *crawler*, um armazenamento de dados, um programa de visualização.

O *crawler* se trata de um programa feito para obter dados de modo automatizado do mesmo modo utilizado pelo *Netvizz*, ou seja, extraíndo dados do *Facebook*. Dependendo da hipótese sobre uma comunidade, ou sobre a relação entre quaisquer tipos de entidades, que se queira testar, se constrói um *crawler* – também chamado de robô, *bot* ou *spider*. Os *crawlers* podem obter dados a partir da extração recursiva de *links* de um site ou de um banco de dados e, em seguida, estruturá-los e armazená-los em tabelas, seja em um banco de dados ou no sistema de arquivos local. É partindo desses dados estruturados que se geram as visualizações.

Suponhamos que se deseja criar uma visualização das relações do *WikiLeaks*. Na massa de dados não estruturados do *WikiLeaks*, Wild (2015) criou um método que contempla um programa para ler cada registro, no caso, cada *e-mail*. Para cada registro,

75 <http://www.postgresql.org/>

5. Obtenção de diferentes conjuntos de dados para construir mapas

foi criado um padrão no cabeçalho: data, destinatário, categoria da mensagem. Em seguida, procedeu-se deste modo:

1. Dados carregados no *Postgres*; 2. Expressões regulares foram usadas para extrair e padronizar os nomes de cidades que são precedidos de “origem”, “destino” e “assunto”; 3. Foi usada a biblioteca *Python Geocoder* para relacionar os códigos de países às cidades; 4. Com base na lista de referência dos códigos de países às subregiões ou regiões, foi criada uma lista customizada de regiões – algumas das quais são subregiões únicas, outras são conjuntos de subregiões; 5. Expressões regulares foram usadas para extrair e padronizar os nomes de cidades; 6. A modelização de tópicos Gensim foi usada para construir um modelo baseado em sessenta assuntos, de acordo com o conteúdo das mensagens; 7. *Python* foi usado para efetuar todo o agrupamento, reunião de categorias e formatação dos dados em estruturas apropriadas em JSON ou CVS (Wild, 2015, tradução nossa)⁷⁶.

Alguns *crawlers* são embarcados, como *plug-ins*, nos aplicativos geradores de visualizações, como o *Gephi*. Utilizamos um desses *plug-ins* para a conexão em nossa caixa de entrada de *e-mails*, usada sobretudo para assuntos acadêmicos, para gerar um gráfico que foi disponibilizado *online*, em alta resolução⁷⁷. O mapa resultante evidencia o propósito daquela conta por meio da verificação das pessoas com quem foram trocadas mais mensagens, criando um aglomerado que representa uma realidade da comunicação: os pontos reunidos no canto inferior esquerdo representam um grupo com o qual elaborou-se um projeto para um dos cursos que frequentamos.

76 “1. Loaded into Postgres; 2. Regular expressions were used to extract and canonicalize city names in "origin", "destination", and "header"; 3. Python Geocoder library used to attach country codes to cities; 4. Based on a Google reference list mapping country codes to both subregions and regions, created a custom list of regions - some of which were single subregions, some conjunctions of subregions; 5. Regular expressions were used to extract and canonicalize city names; 6. Gensim topic modeling used to construct 60-topic model based on content of cables ; 7. General python used to do all other grouping, attaching of categories, and formatting of data into appropriate JSON or CSV structures”

77 <http://goo.gl/cLgnbA>

5. Obtenção de diferentes conjuntos de dados para construir mapas

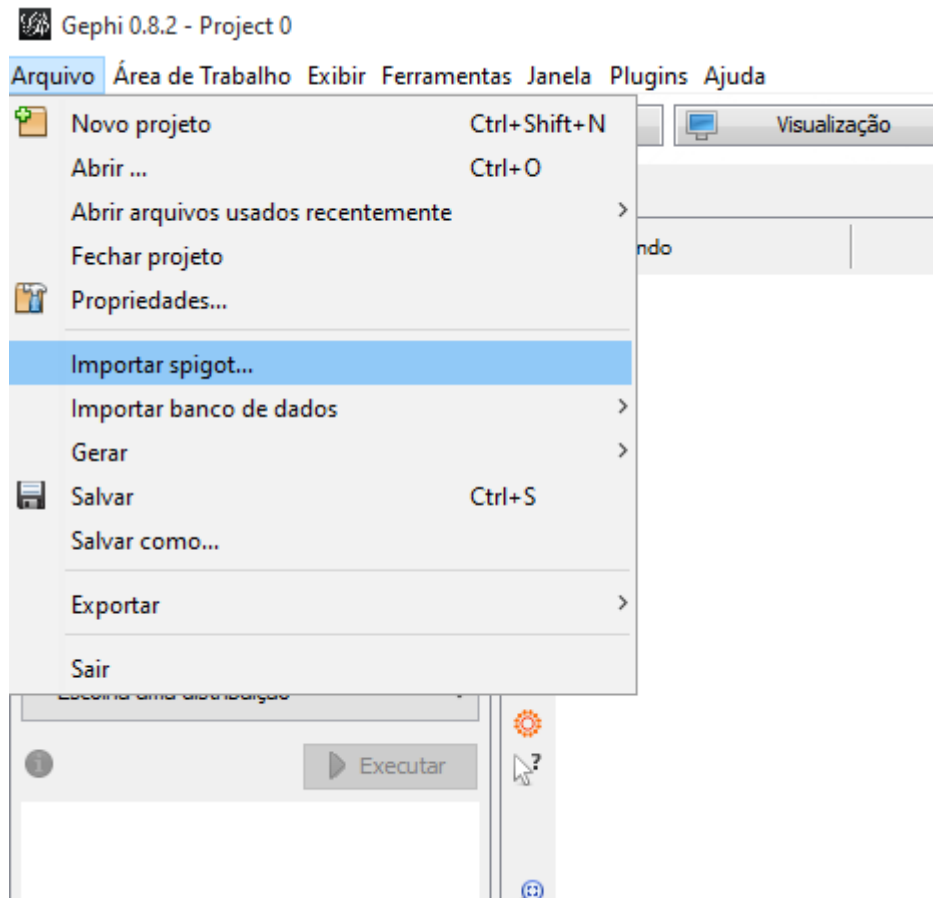


Figura 48 - Importação de dados IMAP com um plug-in do Gephi.

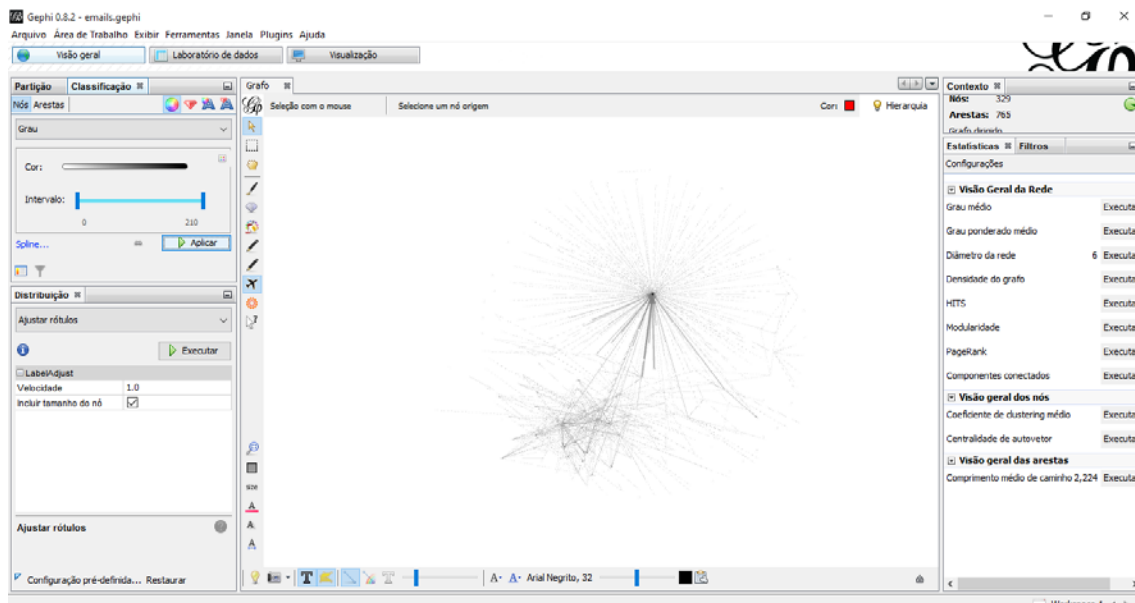


Figura 49 - Dados carregados no Gephi da nossa conta IMAP, pelo plug-in de importação.

5. Obtenção de diferentes conjuntos de dados para construir mapas

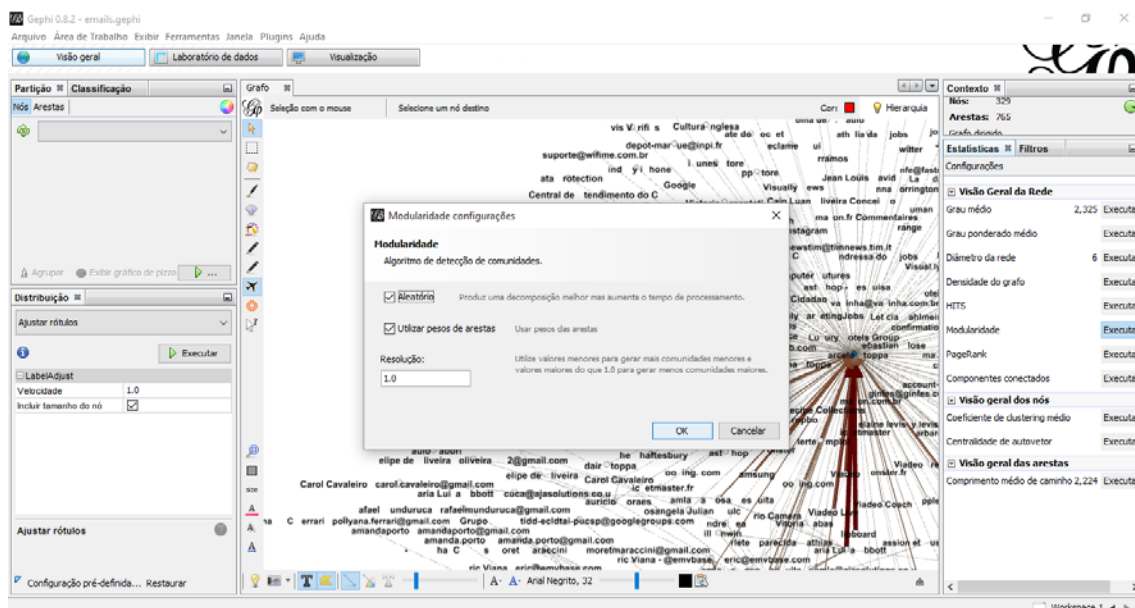


Figura 50 - Aplicação de modularidade nos dados IMAP para detectar comunidades.

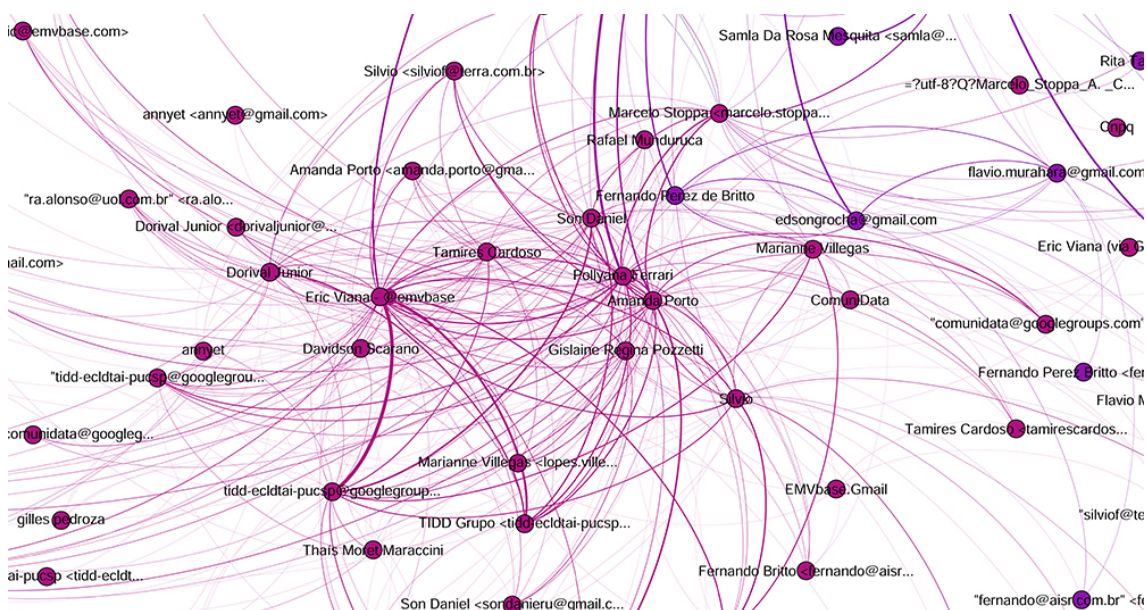


Figura 51 - Detalhe da visualização dos dados IMAP.

Se a massa de dados não for dada, é possível criar *crawlers* para que ela seja obtida. Isto pode ser feito com linguagens como *Python*, como procedeu Wild, ou *Java* e *PHP*⁷⁸, entre outros. Como sistema de armazenamento, qualquer banco de dados pode ser

5. Obtenção de diferentes conjuntos de dados para construir mapas

utilizado. Wild usou *PostgreSQL*⁷⁹, mas podem ser usados também *SQL*⁸⁰, *mySQL*⁸¹ ou *Oracle*⁸², entre outros – ou ainda arquivos com dados balizados, como em *JSON*⁸³, *CVS*⁸⁴, ou arquivos no formato texto com uma marcação customizada. Há também a solução proposta pelo consórcio Apache para se trabalhar com alto volume de dados, ou *Big Data*, baseada no *crawler Nutch*⁸⁵ e no sistema de armazenamento distribuído *Hadoop*, capazes de lidar com 13TB de dados⁸⁶.

Crawlers podem ser programados para buscar e armazenar informações pela Web – imagens, textos, vídeos – em torno de um assunto específico: uma palavra-chave ou combinação de palavras-chave, ou mesmo uma *hashtag* usada com frequência numa rede social, como no *Twitter* ou *Instagram*. A análise da comunidade formada em torno e por consequência do uso de uma *hashtag* executada ao longo do tempo, por exemplo, no *Twitter*, revelaria um mapa de afinidades fugazes que perdurariam apenas enquanto o assunto em torno da *hashtag* tivesse boa audiência. Alguns *crawlers* desenvolvidos para o *Twitter* aferem não apenas os fluxos de diálogo, mas também os sentimentos sobre um determinado assunto. As visualizações resultantes podem agir como um termômetro do humor naquela rede social, assim como dos momentos ele esteve melhor ou pior, como mapeiam os aplicativos *Twistori*⁸⁷ ou o *SocialCollider*⁸⁸.

79 <http://www.postgresql.org/>

80 <http://www.microsoft.com/pt-br/server-cloud/products/sql-server/>

81 <https://www.mysql.com/>

82 <http://www.oracle.com/>

83 <http://www.json.org/>

84 <http://creativyst.com/Doc/Articles/CSV/CSV01.htm>

85 <http://nutch.apache.org/>

86 Informação disponível em <http://marc.info/?l=nutch-general&m=110503724706861>. Acesso em 09.01.2016.

87 <http://twistori.com/>

88 <http://socialcollider.net/>

5. Obtenção de diferentes conjuntos de dados para construir mapas

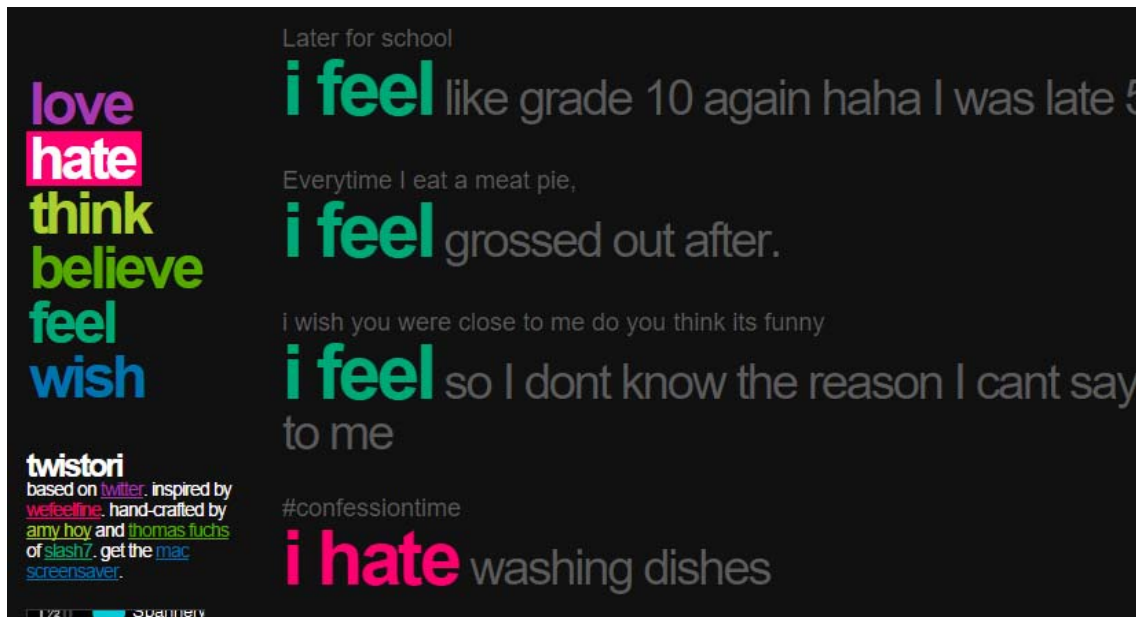


Figura 52 - Captura de tela do site do aplicativo Twistori.

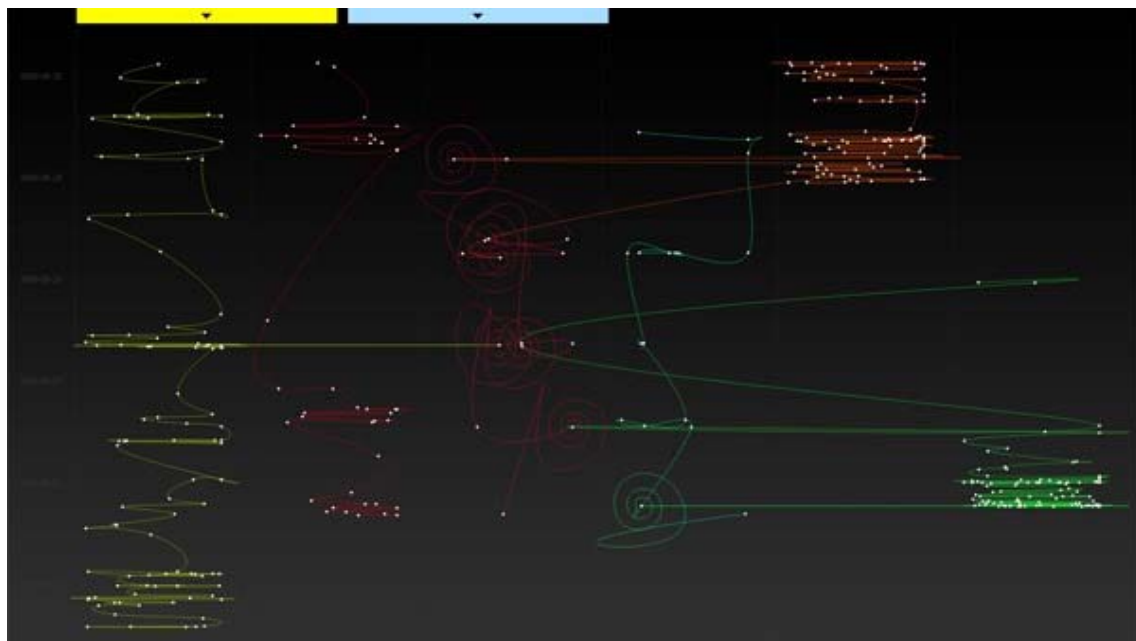


Figura 53 - Captura de tela do aplicativo SocialCollider.

As redes sociais *Instagram*, *Twitter* e *Facebook* permitem a consulta de seus conjuntos de dados de uma forma limitada aos *crawlers*. Isso se dá através de interfaces chamadas, no jargão dos programadores, de “API” (acrônimo de *application programming interface*). As APIs de cada serviço possuem sintaxe e métodos de acesso próprios, o que requer, do desenvolvedor que queira cruzar dados de redes sociais, habilidades na

5. Obtenção de diferentes conjuntos de dados para construir mapas

programação específica de cada serviço. É preciso que o programador esteja atento às mudanças das *APIs* dos serviços que queira integrar, para que a aplicação desenvolvida não se torne obsoleta. O primeiro passo é obter uma conta especial de desenvolvedor em cada uma das redes sociais e se conectar aos serviços usando uma instância do protocolo *OAuth*, para que o *crawler* se identifique ao serviço com a conta do desenvolvedor. O segundo é ter um ambiente de programação especial: pacotes de programação em *Java*, *Python* e *C*, além de um banco de dados relacional instalado ou sistema de arquivos específicos para a mineração de dados, como o *Hadoop*. Um ambiente de testes para mineração e estruturação automática de dados de redes sociais chegou a ser construído, como mostrado na imagem seguinte, mas houve mudanças nas *APIs* do *Facebook* e do *Twitter* que apresentavam curvas de aprendizado muito acentuadas. Foi então emulada uma máquina *Linux* no *Oracle VM VirtualBox*, na qual instalou-se uma imagem do sistema *Ubuntu Server 15.10* em que pacotes do *Hadoop* e *Nutch* foram baixados, instalados e configurados – o que é um ponto de partida para projetos em que mapas de afinidade mais complexos, com cruzamentos de dados, possam ser executados.

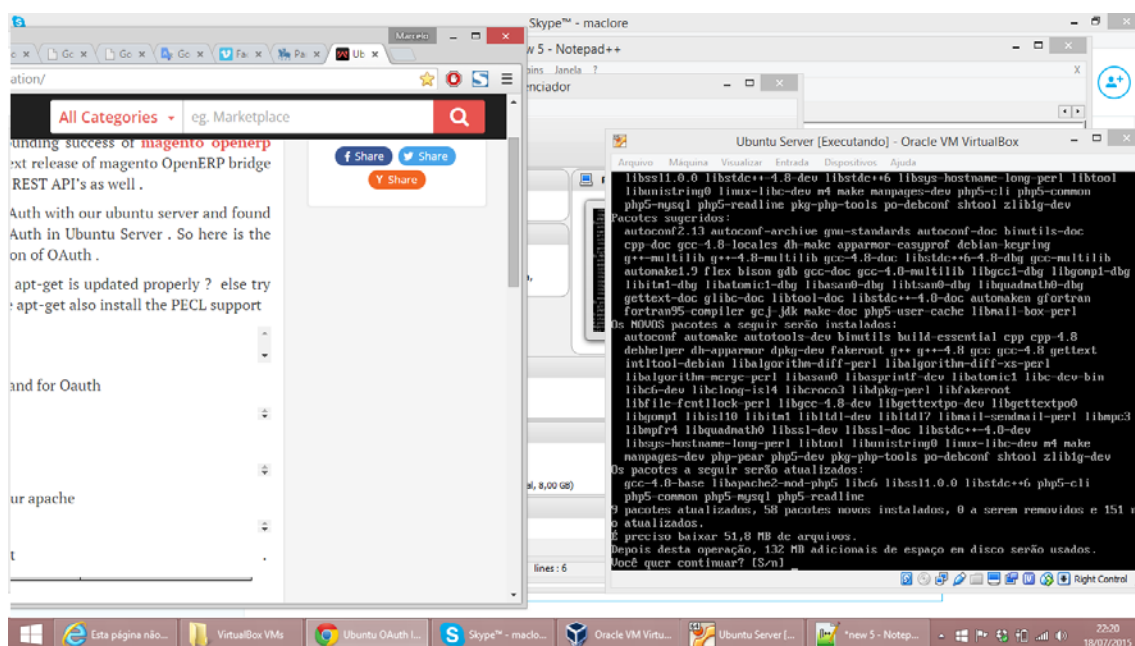


Figura 54 - Captura de tela após instalação do Nutch em máquina virtual.

Se o conjunto de dados já estiver estruturado, como é o caso dos dados disponibilizados em forma de ontologias, estes serão obtidos por meio de uma programação simplificada feita diretamente no serviço que as publica, pelo uso de

5. Obtenção de diferentes conjuntos de dados para construir mapas

comandos estruturados em linguagem de pesquisa. Um desses repositórios de dados, com 4,2 milhões de entidades e quase 2.700 parâmetros, é a *DBpedia*:

[...] uma iniciativa comunitária para extrair informações estruturadas da *Wikipédia* e tornar estas informações disponíveis na *Web*. A *DBpedia* permite que se façam consultas sofisticadas à *Wikipédia* e que se conectem conjuntos diferentes de dados da *Web* aos dados da *Wikipédia*⁸⁹ (DBpedia, 2015, tradução nossa).

Uma ontologia é uma coleção de dados que permite descrever, de modo padronizado (“canonical”), variados tipos de entidades, como lugares, pessoas, trabalhos, espécies ou organizações. Ontologias específicas estão disponíveis na *Web*: lugares podem ser consultados no serviço *GeoNames*⁹⁰, assim como dados financeiros no site do Banco Mundial⁹¹. Informações padronizadas dos governos da França e dos Estados Unidos⁹² também estão disponíveis nos respectivos sites.

Utilizando os dados de ontologias da *DBpedia*, Griffen (2013)⁹³ construiu uma visualização da rede de pessoas listadas na *Wikipédia* de acordo com sua influência, usando os seguintes comandos *SPARQL*⁹⁴:

```
SELECT * WHERE
?p a <http://dbpedia.org/ontology/Person> .
?p <http://dbpedia.org/ontology/influenced> ?influenced.
?influenced a <http://dbpedia.org/ontology/Person>.
```

Assim, obtive arquivos CSV (arquivos no formato texto, com campos separados por vírgula) que podem ser carregados no programa de visualização e, em seguida, construir a rede de pessoas por profissão, o que gerou visualizações com as que se seguem.

89 "DBpedia is a crowd-sourced community effort to extract structured information from Wikipedia and make this information available on the Web. DBpedia allows you to ask sophisticated queries against Wikipedia, and to link the different data sets on the Web to Wikipedia data." Disponível em <http://wiki.dbpedia.org/>. Acesso em 09.01.2016.

90 <http://www.geonames.org/>

91 <http://data.worldbank.org/>

92 <http://www.data.gouv.fr> e <http://www.data.gov>, respectivamente.

93 Disponível em <http://brendangriffen.com/blog/gow-influential-thinkers/>. Acesso em 09.01.2016.

94 Disponível em <http://www.w3.org/TR/rdf-sparql-query/>. Acesso em 09.01.2016.

5. Obtenção de diferentes conjuntos de dados para construir mapas

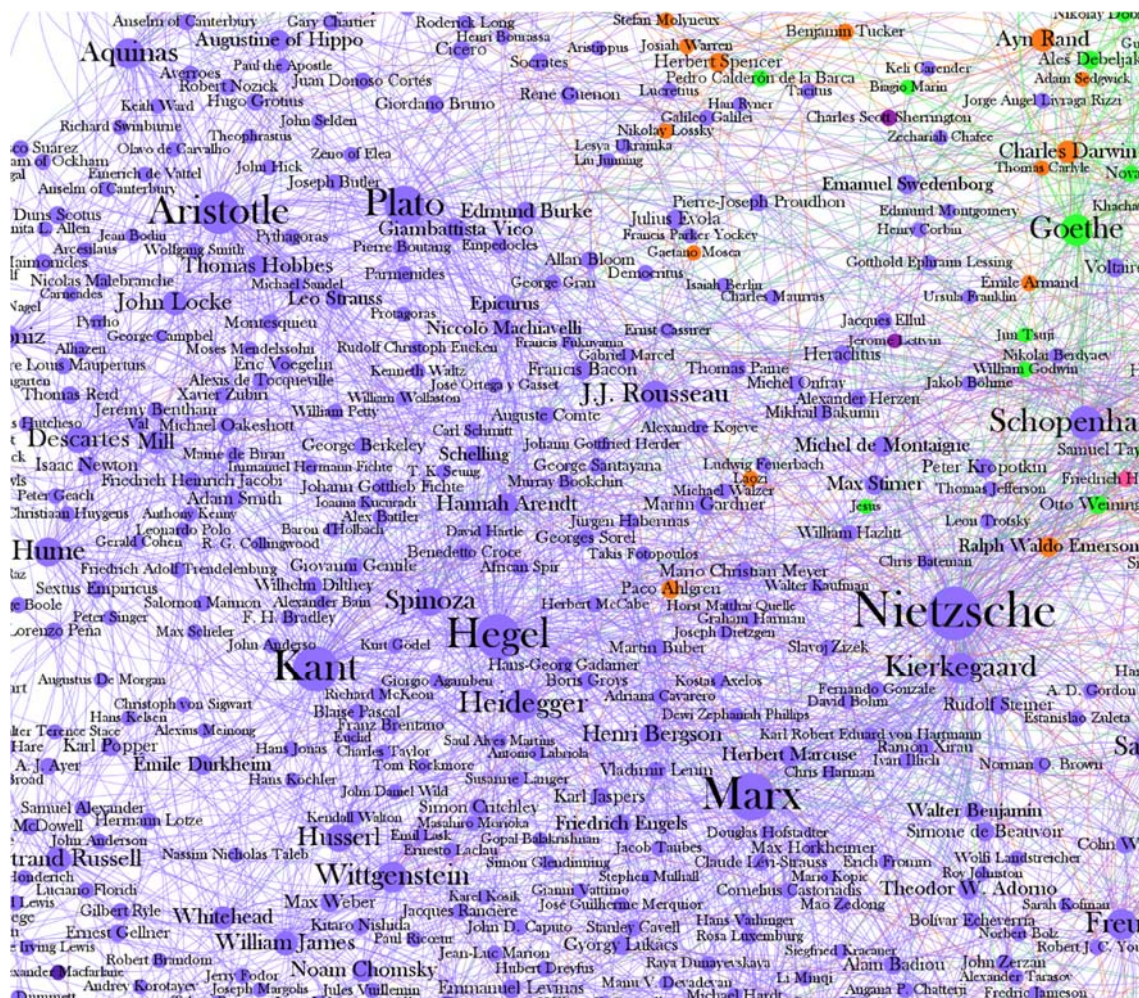


Figura 55 - Rede de filósofos construída por Griffen (2013).

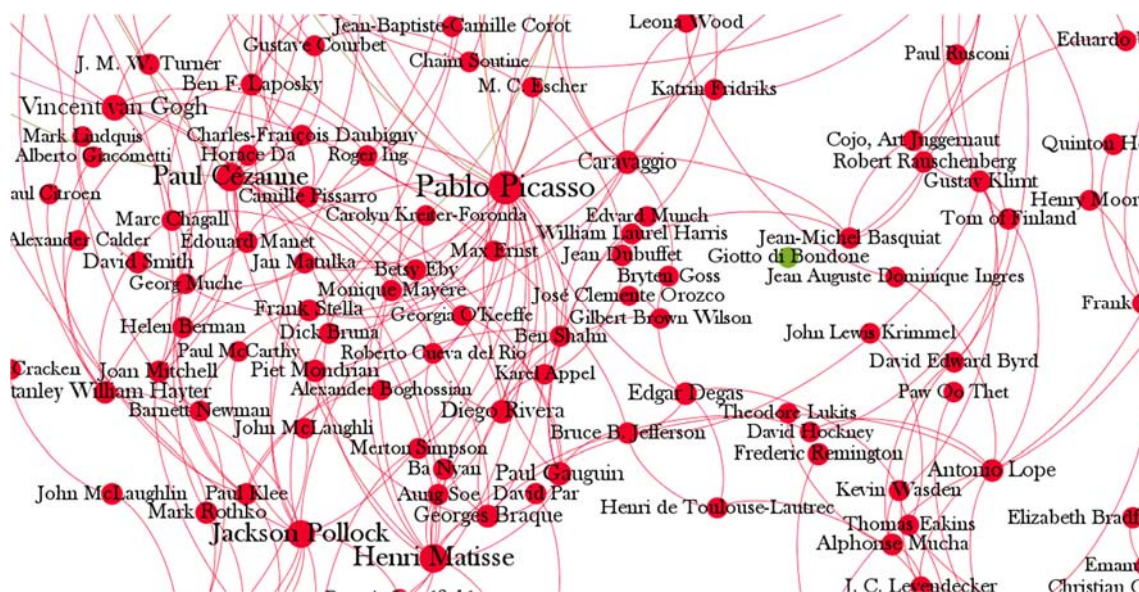


Figura 56 - Rede de artistas construída por Griffen (2013).

5. Obtenção de diferentes conjuntos de dados para construir mapas

Há cada vez mais dados disponíveis em forma de ontologias, devido a dois fatores. O primeiro é a tendência, em curso, de que seja adotado na *Web* o tratamento semântico dos elementos ali publicados, estruturando-se o conteúdo a partir de seu significado e contexto; o segundo é que governos e empresas têm disponibilizado publicamente dados úteis à população: rotas, condições demográficas, distritos eleitorais, resultados de eleições com dados geolocalizados – informações que, cruzadas com dados sobre as ligações e afinidade entre pessoas, podem ajudar a melhor entender fenômenos sociais e, quem sabe, também auxiliar em decisões orçamentárias, de transportes, entre outros.

Poderá ser usado também *SQL*⁹⁵, *mySQL*⁹⁶ ou *Oracle*⁹⁷, entre outros – ou ainda arquivos com dados balizados, como em *JSON*⁹⁸, *CVS*⁹⁹, ou arquivos no formato texto com uma marcação customizada. Há também a solução proposta pelo consórcio Apache para se trabalhar com alto volume de dados, ou *Big Data*, baseada no *crawler Nutch*¹⁰⁰ e no sistema de armazenamento distribuído *Hadoop*, capazes de lidar com 13TB de dados¹⁰¹.

Crawlers podem ser programados para buscar e armazenar informações pela *Web* – imagens, textos, vídeos – em torno de um assunto específico: uma palavra-chave ou combinação de palavras-chave, ou mesmo uma *hashtag* usada com frequência numa rede social, como no *Twitter* ou *Instagram*. A análise da comunidade formada em torno e por consequência do uso de uma *hashtag* executada ao longo do tempo, por exemplo, no *Twitter*, revelaria um mapa de afinidades fugazes que perdurariam apenas enquanto o assunto em torno da *hashtag* tivesse boa audiência. Alguns *crawlers* desenvolvidos para o *Twitter* aferem não apenas os fluxos de diálogo, mas também os sentimentos sobre um determinado assunto. As visualizações resultantes podem agir como um termômetro do humor naquela rede social, indicando os momentos em que ele melhora ou piora, indicando sentimentos como mapeiam os aplicativos *Sentibotr*¹⁰², *Twistori*¹⁰³ ou *SocialCollider*¹⁰⁴.

95 <http://www.microsoft.com/pt-br/server-cloud/products/sql-server/>

96 <https://www.mysql.com/>

97 <http://www.oracle.com/>

98 <http://www.json.org/>

99 <http://creativyst.com/Doc/Articles/CSV/CSV01.htm>

100 <http://nutch.apache.org/>

101 Informação disponível em <http://marc.info/?l=nutch-general&m=110503724706861>. Acesso em 09.01.2016.

102 <http://jamesfrost.me/sentibotr/>

103 <http://twistori.com/>

104 <http://socialcollider.net/>

5. Obtenção de diferentes conjuntos de dados para construir mapas

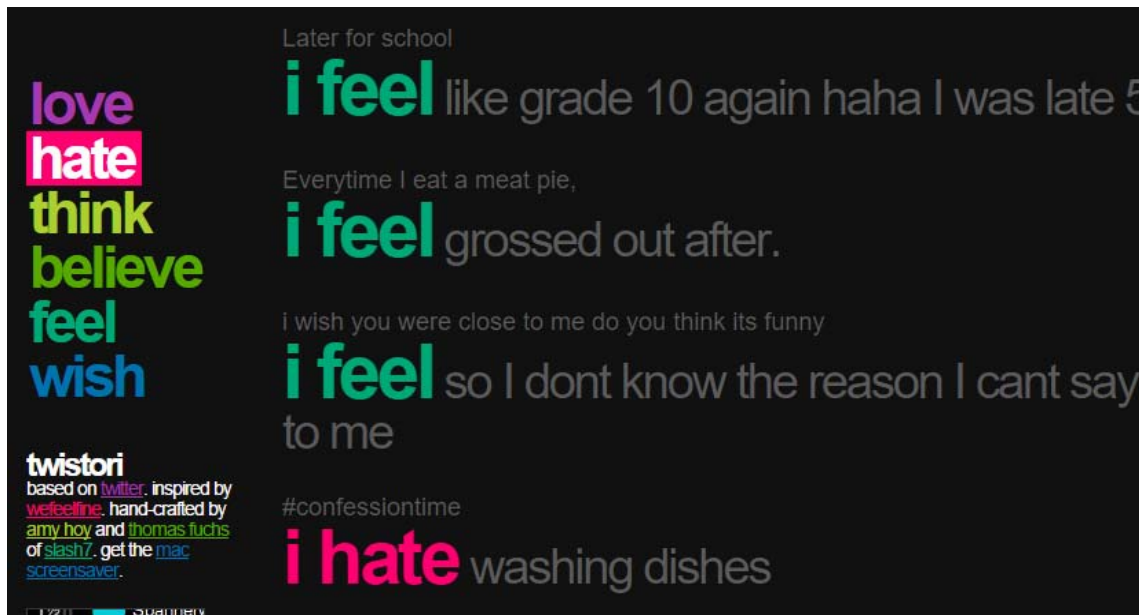


Figura 57 - Captura de tela do site do aplicativo Twistori.

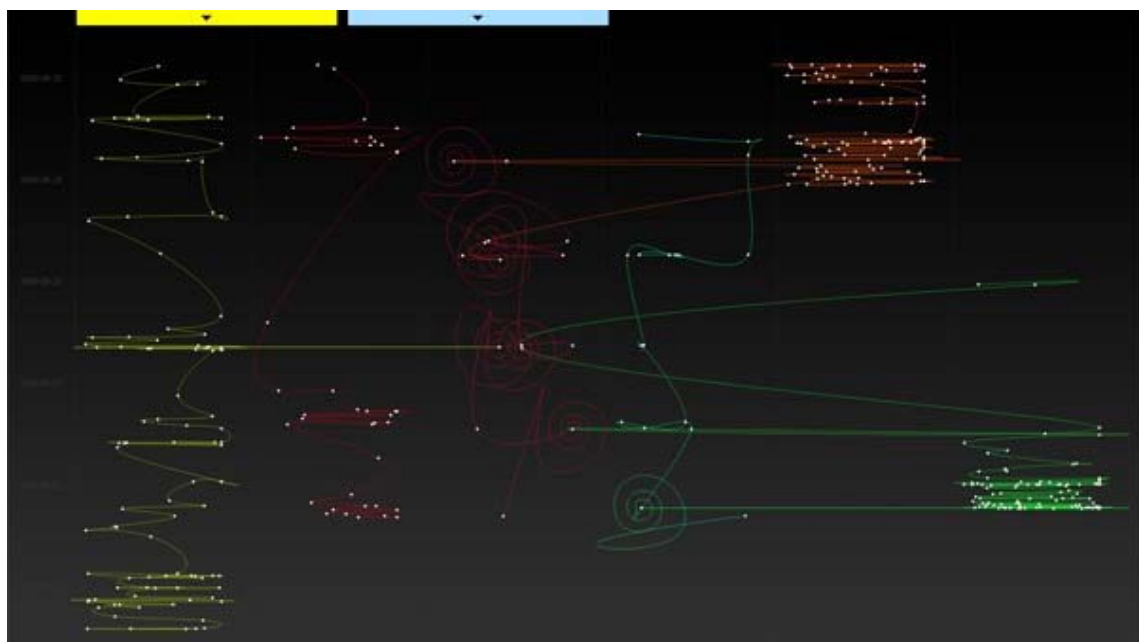


Figura 58 - Captura de tela do aplicativo SocialCollider.

As redes sociais *Instagram*, *Twitter* e *Facebook* permitem a consulta de seus conjuntos de dados de uma forma limitada aos *crawlers*. Isso se dá através de interfaces chamadas, no jargão dos programadores, de “API” (acrônimo de *application programming interface*). As APIs de cada serviço possuem sintaxe e métodos de acesso próprios, o que

5. Obtenção de diferentes conjuntos de dados para construir mapas

requer, do desenvolvedor que queira cruzar dados de redes sociais, habilidades na programação específica de cada serviço. É preciso que o programador esteja atento às mudanças das *APIs* dos serviços que queira integrar, para que a aplicação desenvolvida não se torne obsoleta. O primeiro passo é obter uma conta especial de desenvolvedor em cada uma das redes sociais e conectar-se aos serviços usando uma instância do protocolo *OAuth*, para que o *crawler* se identifique ao serviço com a conta do desenvolvedor. O segundo é ter um ambiente de programação especial: pacotes de programação em *Java*, *Python* e *C*, além de um banco de dados relacional instalado ou sistema de arquivos específicos para a mineração de dados, como o *Hadoop*. Um ambiente de testes para mineração e estruturação automática de dados de redes sociais chegou a ser construído, como mostrado na imagem seguinte, mas houve mudanças nas *APIs* do *Facebook* e do *Twitter* que apresentavam curvas de aprendizado muito acentuadas. Foi então emulada uma máquina *Linux* no *Oracle VM VirtualBox*, na qual instalou-se uma imagem do sistema *Ubuntu Server 15.10* em que pacotes do *Hadoop* e *Nutch* foram baixados, instalados e configurados – o que é um ponto de partida para projetos em que mapas de afinidade mais complexos, com cruzamentos de dados, possam ser executados.

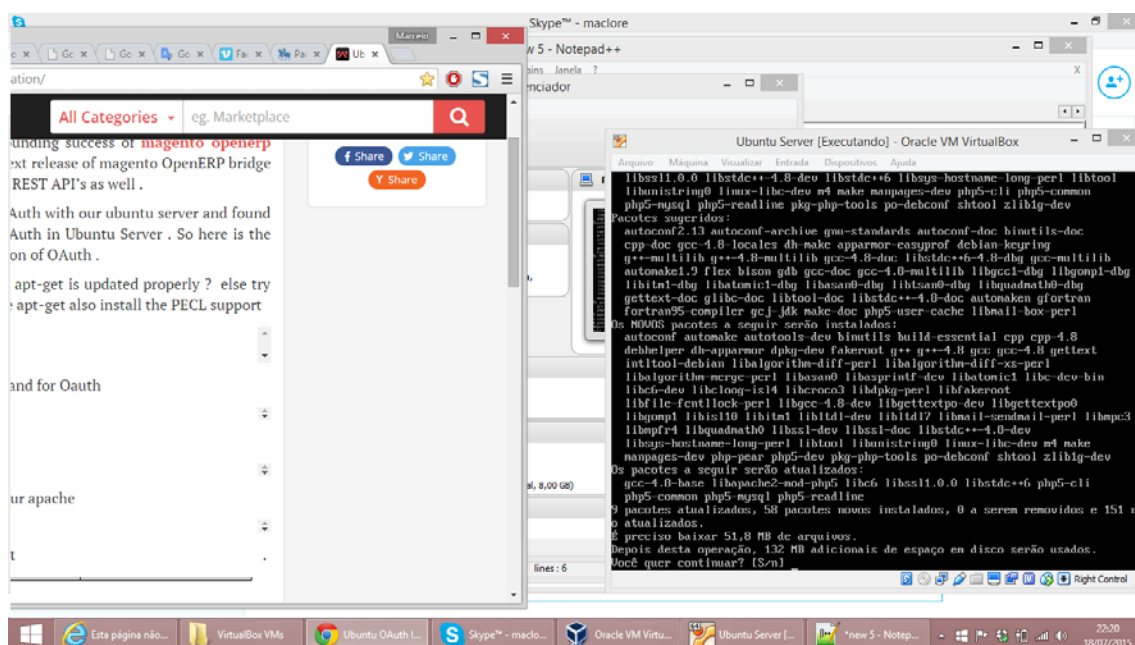


Figura 59 - Captura de tela após instalação do Nutch em máquina virtual.

Se o conjunto de dados já estiver estruturado, como é o caso dos dados disponibilizados em forma de ontologias, estes serão obtidos por meio de uma

5. Obtenção de diferentes conjuntos de dados para construir mapas

programação simplificada feita diretamente no serviço que as publica, pelo uso de comandos estruturados em linguagem de pesquisa. Um desses repositórios de dados, com 4,2 milhões de entidades e quase 2.700 parâmetros, é a *DBpedia*:

[...] uma iniciativa comunitária para extrair informações estruturadas da *Wikipédia* e tornar estas informações disponíveis na *Web*. A *DBpedia* permite que se façam consultas sofisticadas à *Wikipédia* e que se conectem conjuntos diferentes de dados da *Web* aos dados da *Wikipédia*¹⁰⁵ (DBpedia, 2015, tradução nossa).

Uma ontologia é uma coleção de dados que permite descrever, de modo padronizado (“canonical”), variados tipos de entidades, como lugares, pessoas, trabalhos, espécies ou organizações. Ontologias específicas estão disponíveis na *Web*: lugares podem ser consultados no serviço *GeoNames*¹⁰⁶, assim como dados financeiros no site do Banco Mundial¹⁰⁷. Informações padronizadas dos governos da França e dos Estados Unidos¹⁰⁸ também estão disponíveis nos respectivos sites.

Utilizando os dados de ontologias da *DBpedia*, Griffen (2013)¹⁰⁹ construiu uma visualização da rede de pessoas listadas na *Wikipédia* de acordo com sua influência, usando os seguintes comandos *SPARQL*¹¹⁰:

```
SELECT * WHERE
?p a <http://dbpedia.org/ontology/Person> .
?p <http://dbpedia.org/ontology/influenced> ?influenced.
?influenced a <http://dbpedia.org/ontology/Person>.
```

Assim, obtive arquivos CSV (arquivos no formato texto, com campos separados por vírgula) que podem ser carregados no programa de visualização e, em seguida, construir a rede de pessoas por profissão, o que gerou visualizações com as que se seguem.

105 "DBpedia is a crowd-sourced community effort to extract structured information from Wikipedia and make this information available on the Web. DBpedia allows you to ask sophisticated queries against Wikipedia, and to link the different data sets on the Web to Wikipedia data." Disponível em <http://wiki.dbpedia.org/>. Acesso em 09.01.2016.

106 <http://www.geonames.org/>

107 <http://data.worldbank.org/>

108 <http://www.data.gouv.fr> e <http://www.data.gov>, respectivamente.

109 Disponível em <http://brendangriffen.com/blog/gow-influential-thinkers/>. Acesso em 09.01.2016.

110 Disponível em <http://www.w3.org/TR/rdf-sparql-query/>. Acesso em 09.01.2016.

5. Obtenção de diferentes conjuntos de dados para construir mapas

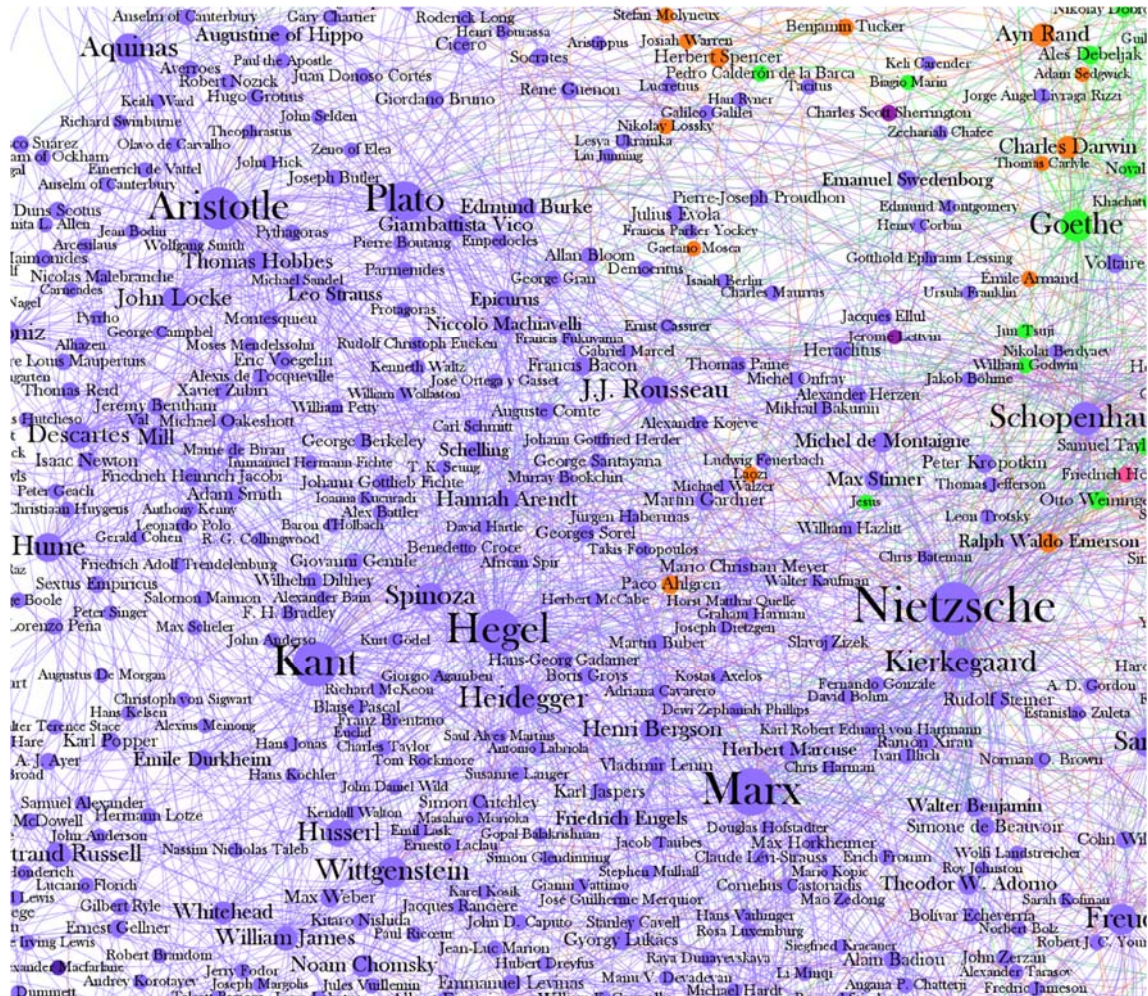


Figura 60 - Rede de filósofos construída por Griffen (2013).

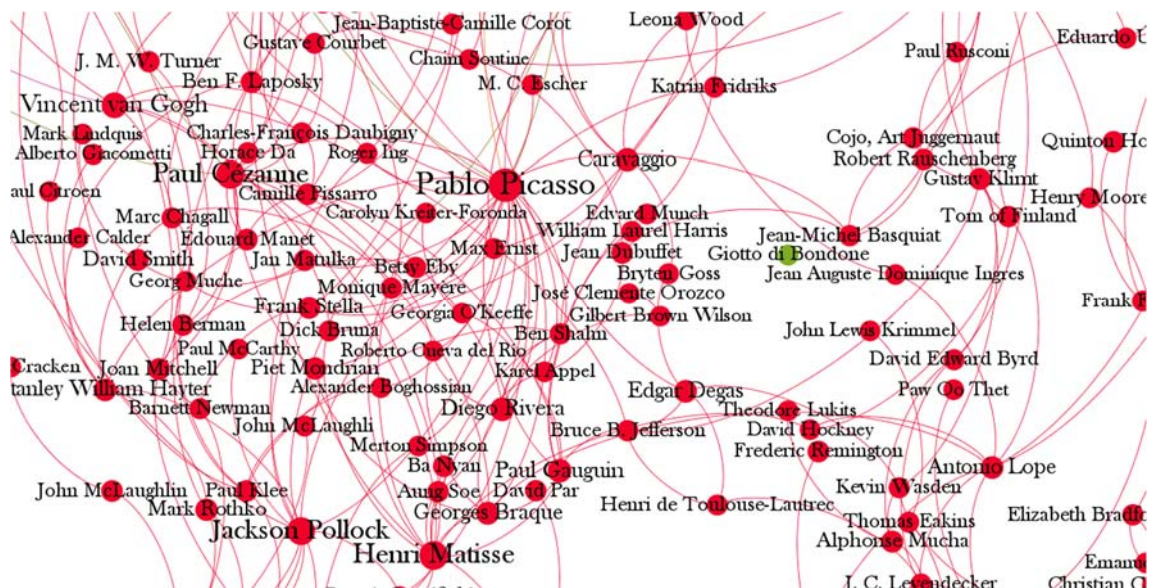


Figura 61 - Rede de artistas construída por Griffen (2013).

5. Obtenção de diferentes conjuntos de dados para construir mapas

Há cada vez mais dados disponíveis em forma de ontologias, devido a dois fatores. O primeiro é a tendência de que se adote cada vez mais o tratamento semântico do conteúdo publicado na *Web*, o que permite que os dados sejam recuperados a partir da estrutura em que se apresentam e do contexto em que se encontram; o segundo é que governos e empresas têm disponibilizado publicamente dados úteis à população: rotas, condições demográficas, distritos eleitorais, resultados de eleições com dados geolocalizados – informações que, cruzadas com dados sobre as ligações e afinidade entre pessoas, podem ajudar a melhor entender fenômenos sociais e, quem sabe, também auxiliar em decisões orçamentárias, educacionais, infraestruturais, entre outros.

CONCLUSÃO

Partimos de uma hipótese inicial, inspirada na observação, de que os arranjos de pessoas em grupos são mais dinâmicos em nossos dias em comparação, por exemplo, àqueles observados no começo do século XX, quando encontrávamos uma sociedade com hierarquia rígida e ubíqua. A presença das tecnologias de comunicação disseminadas hoje, ao lado da evolução do sistema sociopolítico-econômico que vivenciamos, colabora para a fluidez das relações interpessoais. A complexidade do tema exigiu um olhar transdisciplinar; sendo assim, fez-se apelo à psicologia, sociologia, filosofia, ciência da computação e *design* para abordar o problema. Foi a partir da criação de visualizações de uma grande quantidade de dados de relações entre pessoas que iniciamos um tipo de análise sobre os agrupamentos de pessoas em torno das redes de influência (ou grau de conectividade) e afinidade (igualmente referida como modularidade). Este tipo de análise depende não apenas da criação dessas visualizações – que podem revelar aspectos estruturais ou topológicos dos grupos e das redes que transcendem a percepção de um observador – mas também de sua interpretação. Este conjunto, composto da utilização de volumes de dados relacionais de quaisquer tamanhos e complexidade, da geração de visualizações (pela utilização de questionários, *crawlers* ou ontologias) e da interpretação transdisciplinar, é o que definimos por topologias de afinidade, ou *afinimaps*.

A origem deste neologismo está calcada na afinidade das entidades entre si, que se encontra representada na topologia – distribuição e agrupamento de pontos num plano- e pelo comprimento das linhas que conectam os elementos, ao mesmo tempo em que estabelecem a distância entre eles.

A pesquisa foi iniciada com foco em pequenos grupos e no afinimapa de uma pequena comunidade *online*, em que encontramos a correspondência dos arranjos interpessoais online com aqueles presenciais. Apesar de a modularidade das relações mostrar um agrupamento *online* muito aproximado daqueles observados nas relações presenciais, elaboramos um questionário cujos resultados da aplicação melhorariam a resolução das relações.

Em seguida, com uma comunidade maior do *Facebook*, obtivemos um conjunto de dados mais vultoso e, pela aplicação do algoritmo de modularidade, encontramos aglomerados de pessoas e subconjuntos dentro da comunidade *online*. Observamos, assim, o mesmo padrão em conjuntos de dados distintos: naquele que construímos para realizar a topologia baseada em afinidades com os partidos políticos do Brasil; naquele que obtivemos com um *crawler* e resultou na visualização da rede de contatos pessoais do Facebook; no conjunto de dados obtido por um *crawler* que extraiu, de uma caixa de *e-mail*, as informações pelo protocolo IMAP. Com isto, provamos que o método de estudos de comunidades na fase da construção das visualizações e aplicação de algoritmos é eficaz na detecção de comunidades.

O estudo seguinte deu-se sobre os fatores influência e viralização nessas redes. A visualização construída verificou-se útil para identificar os elementos que possuem grau médio de conectividade, ou seja, aqueles que, por não estarem no centro dos aglomerados, têm maior poder de propagar rapidamente uma mensagem. O afinimapa, deste modo, pôde revelar a geometria e a dimensão estrutural do capital social das redes relacionais.

Esta metodologia abre caminho para diversos modos de exploração e análise em conjuntos de dados relacionais. Imaginamos que o estudo da criação, ascensão e declínio da utilização de uma *hashtag* (ou palavra-chave) em uma rede social revelaria mais facetas do capital social ao dispor as interações em um eixo temporal e relacional. Outro caso imaginado seria o estudo da personalidade de um indivíduo usando como matéria-prima os dados das transações comunicacionais captados a partir das tecnologias que utiliza e de como interage com as comunidades em que está inserido – como se o indivíduo fosse um átomo em torno do qual orbitam dados que indicam as possibilidades de interações com outros elementos; um estudo parecido com aquele sobre valência química em Pauling. Sobre atração e repulsão que, em Moreno, é afetiva e que, com a metodologia construída, assinalaria afinidades pelos dados de interação e modularidade.

Dentro de um sem-número de cenários complexos imaginados para a aplicação do afinimapa, elencamos alguns a seguir. Uma de suas utilidades poderia ser, por exemplo, o combate ao recrutamento *online* do *Daesh*, ao se lançar um *crawler* em diversas redes sociais para detectar quais situações sociais predisporiam os jovens a aderir ao discurso extremista, objetivando a construção de políticas que diminuiriam a fragilidade de pessoas

em situação de risco. Através dos dados de geolocalização dos celulares daqueles que fogem dos conflitos na Síria, seriam oferecidos a eles caminhos mais seguros. Poder-se-ia, ainda, estudar modos de aumentar a inclusão em uma comunidade, pequena ou grande; analisar a forma em que se dão as relações de poder no ambiente de trabalho e construir estratégias para diminuir possíveis conflitos, como no diagnóstico de casos de assédio; averiguar com maior precisão a eficácia de campanhas publicitárias *online*, um passo além da mensuração de resultados por cliques e compartilhamentos. Seria possível investigar fluxos de comunicação e produção dos departamentos de uma empresa, ou entre empresas, para traçar planos de otimização. Também para construir estratégias de aprimoramento de conteúdo, aliando ao afinimapa o conhecimento existente sobre otimização para mecanismos de busca (SEO) no desenvolvimento da audiência de produtos digitais. O afinimapa pode configurar-se ainda útil para o jornalismo de dados. Todas essas iniciativas, tendo em vista a complexidade dos dados e das relações não só entre elas, devem ser investigadas com um olhar multidisciplinar.

Por fim, espera-se que a metodologia aqui construída seja de utilidade na investigação e no melhoramento das relações entre pessoas, comunidades e, sobretudo, na busca de soluções construtivas para os desafios que se apresentam em um mundo sempre mais complexo e conectado.

BIBLIOGRAFIA

BARTHES, Roland. **Mitologias**. Porto: Edições 70, 2007.

BAUMAN, Zygmunt. **Modernidade Líquida**. São Paulo: Zahar, 2001.

_____. **Comunidade: a busca por segurança no mundo atual**. São Paulo: Zahar, 2003.

BOYD, Danah e CRAWFORD, Kate. Critical questions for Big Data: provocations for a cultural, technological and scholarly phenomenon. **Information, Communication & Society**, Vol. 15, N. 5, pp. 662–679. Nova Iorque: Routledge, 2012.

BERGSON, Henri. **Matière et mémoire**. Paris: PUF, 2009

BLUMER, Herbert. **Symbolic Interactionism: perspective and method**. Oakland: University of California Press, 1969

CASTELLS, Manuel. **A era da informação: sociedade em rede**. São Paulo: Paz e Terra, 2007.

COSGROVE, Bronwyn. **Fashions: year in review 2014**. Disponível em <http://global.britannica.com/topic/fashion-society-Year-In-Review-2014>. Acessado em 03.04.2015.

DELEUZE, Gilles e GUATTARI, Félix. **Mille Plateaux**. Paris: Éditions de Minuit, 1980.

FERRARI, Pollyana. Narrativas «tageadas» nos fluxos sociais. In SQUIRRA. Sebastião (org.). **Ciber mídias: extensões comunicativas, expansões humanas**. Porto Alegre: Buqui, 2012.

FOUCAULT, Michel. **Les mots et les choses**. Paris: Gallimard, 1966.

FREUD, Sigmund. **Totem and Taboo**. Londres : George Routledge & Sons, 1919.

GENOT, Sacha e TALLOIS, Philippe. **L'humain dans l'entreprise: Un capital à préserver**. Paris: Eyrolles, 2013.

GIERETH, Mark. **Foafscape - a Browser for FOAF Related RDF Data**. Disponível em <http://foafscape.berlios.de/>. Acesso em 03.04.2015.

HALL, Stuart. **A identidade cultural na pós-modernidade**. Rio de Janeiro: Lamparina, 2014.

HAMBURGER, Ellis. **Facebook Freakonomics: a new tool to map the world's friendships**. Disponível em <http://www.theverge.com/2012/9/10/3303178/facebook-freakonomics-mapping-the-worlds-friendships>. Acesso em 03.04.2015.

HEER, Jeffrey; BOYD, Danah. **Vizter: visualizing online social networks**. Disponível em <http://hci.stanford.edu/jheer/projects/vizster/>. Acesso em 24.07.2013.

HEER, Jeffrey. **Graphview**. Disponível em <http://prefuse.org/gallery/graphview/>. Acesso em 24.07.2013.

HEIDLER, Richard et al. **Relationship patterns in the 19th century: The friendship network in a German boys' school class from 1880 to 1881 revisited**. *Social Networks*: volume 37, pp. 1-13. Amsterdam: Elsevier. 2014.

ILINSKY, Noah e STEELE, Julia. **Designing data visualizations: intentional communication from data to display**. Sebastopol: O'Reilly Media, 2011.

_____. **Four pillars of visualization**. Disponível em <http://strataconf.com/stratany2013/public/schedule/detail/31182> . Acesso em 24.07.2013.

JACÓ-VILELA, Ana Maria et al (Org.). **História da Psicologia**. Rio de Janeiro: Nau Editora, 2008

JAUREGUIBERRY, Francis e PROULX, Serge. **Usages et enjeux des technologies de communication**. Toulouse: Éditions Érès, 2011.

JUNG, Carl Gustav. (org.). **O homem e seus símbolos**. Rio de Janeiro: Nova Fronteira, 2008.

KHAN, G. F., SOKHA, V. **Virality over YouTube: an Empirical Analysis**. Internet Research Journal, disponível em https://gfkhan.files.wordpress.com/2013/07/virality_authors-version.pdf. Acesso em 05.12.2014.

LETEINTURIER, Christine e CHAMPION, Rémy Le (Org.). **Médias, Information et Communication**. Paris: Eyrolles, 2009.

LE BON, Gustave. **A era das multidões**. São Paulo: WMF Martins Fontes, 2008.

LE BRETON, David. **L'interactionnisme symbolique**, Paris: PUF, 2004.

LE MOS, Ronaldo. **Como influenciar na Internet?** Folha de São Paulo, São Paulo, 14 abr 2015 disponível em

<<http://www1.folha.uol.com.br/colunas/ronaldolemos/2015/04/1616187-como-influenciar-na-internet.shtml>>

LÉVY, Pierre. **Tecnologias da Inteligência**. Rio de Janeiro: Editora 34, 1993.

_____. **Cibercultura**. São Paulo: Editora 34, 1999

LEWIN, KURT. **Principles of topological psychology**. Nova Iorque: McGraw-Hill, 1936.

LIMA, Manuel. **Cartographie des réseaux: L'art de représenter la complexité**. Paris: Eyrolles, 2011.

LIU, Michel. **Approche sociotechnique de l'organisation**. Paris: Les Éditions d'organisation, 1983.

LOCKE, John. **An essay concerning human understanding**. London: Collins, 1973.

MAILHIOT, Gérald Bernard. **Dinâmica e gênese de grupos. Atualidade das descobertas de Kurt Lewin**. São Paulo: Livraria duas cidades, 1970.

MARINEAU, René F. **Jacob Levy Moreno, 1889-1974 : Pai do psicodrama, da sociometria e da psicoterapia de grupo**. São Paulo: Ágora, 1992.

MARTIN-JUCHAT, Fabienne. **Le corps communicant**. Paris: L'Harmatan, 2007.

_____. **Le corps et les médias : la chair éprouvée par les médias et les espaces sociaux**. Bruxelas: De Boeck, 2008.

MERLEAU-PONTY, Maurice. **Phénoménologie de la Perception**. Paris: Gallimard, 2010.

_____. **Fenomenologia da percepção**. São Paulo: Martins Fontes, 1999.

MESSINGER, Joseph. **Ces gestes qui vous trahissent**. Paris: Éditions J'ai lu, 2013.

MOORE, Ralph William. *Cognition, semiotics and the cartographic statement*. 1989. Dissertação (Ph. D.) - University of Kansas.

MORENO, Jacob Levy. **Quem sobreviverá? – Fundamentos da sociometria, psicoterapia de grupo e sociodrama**. Goiânia: Dimensão, 1994.

NOVAES, Marcos. **O conceito de tele em Moreno**. Escola Redes, São Paulo, 10.02.2013. Disponível em <http://escoladeredes.net/group/biblioteca-jacob-moreno/forum/topics/o-conceito-de-tele-em-moreno>. Acesso em 25.11.2015.

PEREIRA, Carlos. **O que é contracultura?** São Paulo: Brasiliense, 1983.

PIPPENBERG, Nathan. **Hipster Counterculture not unique**. Disponível em http://www.huffingtonpost.com/the-daily-collegian/hipster-culture_b_1341702.html. Acesso em 02.02.2015.

SAMUELS, Andrew. **A Psique Plural: Personalidade, Moralidade e o Pai**. Rio de Janeiro: Imago, 1992.

SANTAELLA, Lucia e NOTH, Winfried. **Imagem: Cognição, semiótica, mídia**. São Paulo: Iluminuras, 2000.

SARTRE, Jean-Paul. **A imaginação**. São Paulo: LPM, 2008

SCHENSUN, Jean et al. **Mapping social networks, spatial data and hidden populations**. Toronto: Altamira Press, 1999.

SCHIOCCHET, Dayene. **As forças invisíveis por trás dos resultados: uma análise psicanalítica das relações de poder nas empresas**. Dissertação de mestrado. Pontifícia Universidade Católica de São Paulo, 2014.

STOPPA, Marcelo. **A geografia do mal**. Disponível em <http://furabolo.org/a-geografia-do-mal/>. Acesso em 27.11.2015.

TEMPLAR, Richard. **The Rules of Work, Expanded Edition: A Definitive Code for Personal Success**. Londres: FT Press, 2010.

TOUBOUL, Jonathan. **The hipster effect: when anticonformists all look the same**. Paris: Collège de France, 2014.

WEBER, Max. **Economia e sociedade**. Paris: Pocket Agora, 2003.

WILD, Cody. **Visualizing state department cable**. Disponível em <http://decodyng.github.io/data-viz-portfolio/wikileaks-visualization/network.html> . Acesso em 01.06.2015.

WOLTON, Dominique. **Internet et après?**. Paris: Éditions Flammarion, 2000.

