



Tema: Implantação de Boas Práticas de Gestão e Melhoria da Qualidade dos Serviços Prestados pela Receita Federal do Brasil.

2º Lugar

Identificação do Real Beneficiário e Busca Patrimonial por Grafos de Relacionamento no ContÁgil – Script “Análises Patrimoniais - SABRE”

Autor:

Reinaldo da Cruz Castro

ANEXO 3

Portaria RFB nº 1207, de 29/07/2016

RESUMO DA MONOGRAFIA

INSCRIÇÃO

Nº: 015

TÍTULO DO TRABALHO:

Identificação do Real Beneficiário e Busca Patrimonial por Grafos de Relacionamento no ContÁgil – Script “Análises Patrimoniais - SABRE”

A) OBJETIVOS BÁSICOS

O objetivo deste trabalho é demonstrar a utilização da ferramenta de análises patrimoniais; o “sabre”; na investigação patrimonial por meio de Grafos de Relacionamento, auxiliando o Auditor Fiscal na identificação do Real Beneficiário; na caracterização de Grupos Econômicos, na detecção de Interpostas pessoas; no deslinde das Blindagens Patrimoniais, na composição da estrutura de Organizações Criminosas e auxiliar os demais usuários do ContÁgil a localizar bens para arrolamento e medidas cautelares, com o fito de garantir o crédito tributário.

B) METODOLOGIA UTILIZADA

Como metodologia, escolhemos delinear o conceito de Big Data, Redes e sua ciência, Grafos de Relacionamentos, Bases de Relacionamentos; sempre que possível fugindo do excessivo tecnicismo; utilizando várias obras que exploram os assuntos tratados. A apresentação de dois cases ilustram as principais funcionalidades da ferramenta, cuidando de não apresentar aspectos essencialmente estratégicos, que devem permanecer apenas no domínio do corpo funcional da RFB. Embora com certo prejuízo, pois não podemos, por exemplo, citar todas as fontes de informação que a ferramenta alcança, fica caracterizada sua flexibilidade e moldabilidade, em sintonia com a necessidade de interação com sistemas de informação internos e externos.

C) ADEQUAÇÃO DO TRABALHO AOS CRITÉRIOS DE JULGAMENTO:

I – CRIATIVIDADE E INOVAÇÃO

A tecnologia de Grafos de Relacionamento é utilizada fortemente nas redes sociais e ferramentas de *e-business*; e vem sendo utilizada internacionalmente pelos organismos internacionais nas investigações contra organizações criminosas, lavagem de dinheiro, terrorismo e práticas econômicas perniciosas.

Os conceitos, disponíveis para aplicação no ContÁgil por usuários avançados, ao serem organizados em uma ferramenta acessível ao usuário nível básico, torna possível sua massificação de uso, de maneira que se pretende revolucionar a lida com o dado, com a informação, saindo do conceito tabular de linhas e colunas para utilizar a visão de um grafo, onde pessoas, bens, direitos e valores são ENTIDADES, e as relações entre elas constituem as ARESTAS, ambos enriquecidos com PROPRIEDADES que permitem consultas direcionadas por critérios à escolha do usuário, e a aplicação

de algoritmos matemáticos que, embora extremamente complexos foram convertidos em um “simples botão” que o usuário aciona.

Se, antes da ferramenta era necessário planilhar dados de vários sistemas e cruzá-los, manualmente entre si, com muita propensão a falhas, com o “sabre” todas as fontes de consulta estão a um clique; e a integração entre as fontes consultadas se faz num piscar de olhos.

A melhoria nas ações “realizar ações de pesquisa e investigação” e “atuar na garantia do crédito tributário”, é sensível e imediata.

II – RELAÇÃO CUSTOS VERSUS BENEFÍCIOS

O desenvolvimento da ferramenta foi feito no ambiente do ContÁgil, e partiu de ferramentas já construídas, consolidando-as e colocando-as harmonicamente em um novo script. À medida em que os testes foram sendo feitos, novas funcionalidades foram sendo incorporadas, sempre utilizando “prata da casa”, recursos tecnológicos e humanos da própria RFB.

Todo o desenvolvimento foi feito ao custo de “horas de auditoria” aplicadas, não gerando quaisquer ônus orçamentário ou financeiro. O aperfeiçoamento da ferramenta se deu na prática, trabalhando situações cotidianas e que traziam grande desconforto na análise e investigação; tornando o trabalho extremamente prazeroso, e até, divertido.

A utilização do “sabre” nas grandes operações ainda em evidência na mídia foi determinante para o deslinde de várias situações envolvendo grandes corporações e pessoas políticas, por exemplo.

Como ganho não mensurável, aponta-se a valorização profissional, a motivação, em especial do Auditor Fiscal (Investigação, prospecção, auditoria, autuação, representação e interação com órgãos externos) e do Analista Tributário (Arrecadação e instrumentação da Garantia do Crédito).

Temos ainda como ganhos não mensuráveis o fortalecimento da Instituição RFB perante as demais instituições como Poder Judiciário, Ministério Público Federal, Polícia Federal e Procuradoria da República. Uma apresentação com Grafos de Relacionamento realmente impressiona os espectadores e um relatório fiscal ou peça jurídica construído com grafos é contundente.

III - AUMENTO DE PRODUTIVIDADE

A ferramenta foi testada em situações nas quais os métodos tradicionais não haviam tido êxito, ou em que os servidores estavam com dificuldades em delinear um caminho a seguir. Em todos os casos a atuação foi profícua eficaz.

O “case 2”, que ilustra o trabalho foi executado por solicitação do juízo, uma vez que a execução do devedor contumaz via-se frustrada por falta de lastro patrimonial. O sujeito passivo nos deve R\$ 500 milhões de reais; e ao concluir essa monografia, a decisão judicial já havia arrestado grande valor em pecúnia e tornado indisponíveis bens de pessoas até o terceiro nível do principal devedor.

Durante um dos treinamentos realizados, por exemplo, após formar a base de relacionamentos e aplicar as consultas, localizou-se 03 aeronaves (“jatinhos”) registrados em nome de microempresas relacionados ao devedor principal, que, por óbvio, “nada possuía”.

Nas grandes operações, foi possível apontar, no mar de dados obtidos inclusive por quebra de sigilo bancário, com milhões de registros, por onde os Auditores deveriam seguir.

Também logramos êxito em “provas negativas”, demonstrando não haver relação entre dois grupos de fraudadores que atuavam no mesmo tema, com *modus operandi* distintos.

A velocidade com que se faz os cruzamentos de dados é impressionante. A maior dificuldade reside em OBTER os dados, uma vez que os dados estão esparsos; é necessário baixá-los para a máquina do usuário, onde serão processados; mas isso também ocorre com os métodos tradicionais; com o agravamento de se ter que “pular” de um sistema a outro, fechando e abrindo telas, inserindo pins e senhas, situação consideravelmente mais confortável com o uso do “script de análises patrimoniais; o sabre”.

IV - VIABILIDADE DE IMPLEMENTAÇÃO

A viabilidade já está provada e implementada:

- ❖ O script já conta com quase 5 mil downloads e 11,5 mil execuções.
- ❖ Foi apresentado em Seminários de Garantia do Crédito Tributário na Primeira, Oitava, Nona e Décima Regiões Fiscais, em Seminários de Programação e Fiscalização, Seminários de Gestão Estratégica e em reuniões com a alta administração, inclusive ao Auditor Fiscal Secretário da Receita Federal do Brasil.
- ❖ Foi criada uma comunidade no Conexão Receita – “Investigação Patrimonial via Grafos de Relacionamento - ContÁgil”, que conta com 350 pessoas que interagem em fóruns e tiram suas dúvidas. Na área de arquivos, há roteiros e dicas.
- ❖ Foi criada uma página de informações na Wiki do ContÁgil, como subitem de “Grafos de Relacionamento”.
- ❖ Foram gravadas *videoaulas* com exercícios práticos cujas bases foram devidamente ofuscadas (case – 1 dessa monografia). As videoaulas estão disponíveis na intranet, na Wiki do ContÁgil.
- ❖ O script possui mais de 05 mantenedores, capazes de atualizar, revisar e aperfeiçoá-lo constantemente.
- ❖ Foram realizados 04 treinamentos presenciais; com turmas de 43 treinandos cada, entre Auditores Fiscais de todas expertises, inclusive X-CAT, X-ORT, X-PEI, X-FIS e COGER e alguns Analistas Tributários que atuam na área de Garantia do Crédito.
- ❖ Pretende-se ainda gravar novas videoaulas com as últimas funcionalidades, consultas e bases e formar novos multiplicadores que possam atuar em nível nacional.
- ❖ Há o apoio da COPES, COFIS e CODAC, além da própria SRRF/08.

V - MELHORIA DA QUALIDADE DOS SERVIÇOS PRESTADOS E DOS RESULTADOS ESTRATÉGICOS

Perspectiva de Resultados:

Garantir a arrecadação necessária ao Estado, com eficiência e aprimoramento do sistema tributário.

A ferramenta auxilia sobremaneira na busca de lastro patrimonial, para os créditos tributários já constituídos; permite a identificação de bens capazes de satisfazer o crédito que se está por constituir; e permite, já na fase de programação e planejamento da ação fiscal, saber quais os reais sujeitos passivos e a possibilidade de se conseguir solver o possível crédito.

Perspectiva de Processos:

Aumentar a efetividade de cobrança

Nesse item, a ferramenta possibilita uma ação preventiva, se utilizada antes da constituição do crédito tributário; concomitante, se utilizada durante; e corretiva, se utilizada na fase em que o crédito está inscrito em dívida Ativa ou em execução Fiscal.

Ampliar o combate ao contrabando, ao descaminho e à sonegação fiscal

A identificação dos grupos econômicos e dos reais beneficiários permite um eficaz combate à sonegação.

Incentivar o cumprimento das obrigações tributárias e aduaneiras

Uma ação é eficaz quando ajuda a mudar comportamento. As medidas coercitivas de recolhimento dos tributos, o arrolamento de bens, as medidas cautelares fiscais, o sucesso em arrestos e penhoras e na indisponibilidade de bens educam pela força. O uso do “sabre” auxilia na construção das peças jurídicas que serão recepcionadas e deferidas.

Perspectiva de Gestão e Suporte:

Desenvolver competências, valorizar pessoas e adequar o quadro de pessoal às necessidades institucionais

Ao manipular os grafos de relacionamentos em uma investigação patrimonial, o Auditor Fiscal exercita toda sua *expertise*, aliada a uma ferramenta de respostas rápidas. Se o *sabre* é utilizado na arrecadação, possibilita ao usuário uma visão ampla dos bens e direitos a validar ou validados, e facilita sua interação com o próprio CONPROVI, na gestão dos bens que garantem os Créditos Tributários.

Indicar na Cadeia de Valor (Anexo 5) da RFB os processos de trabalho a que se refere a monografia, conforme os macro-processos relacionados abaixo:

- ✓ **Gestão do Crédito Tributário**
 - Atuar na Garantia do Crédito Tributário
- ✓ **Fiscalização Tributária e Combate a Ilícitos:**
 - Realizar Ações de Pesquisa e Investigação
 - Realizar a Pesquisa e Seleção
 - Realizar a Fiscalização
- ✓ **Relacionamento com Sociedade, Governos e Organismos Nacionais e Internacionais**
 - Conduzir relações institucionais

D) PRINCIPAIS LIÇÕES APRENDIDAS

- ❖ A visão tabular de dados (linhas e colunas) foi superada pela visão gráfica (grafos de relacionamento);

- ❖ A construção da base de relacionamentos com dados internos (RFB) e externos (convênios e dados abertos) aliada às consultas inteligentes e rápidas desnudam a verdade material;
- ❖ As consultas investigativas podem facilmente ser construídas com a ferramenta “análises patrimoniais”, inclusive pelo usuário nível básico do ContÁgil;
- ❖ Os modelos matemáticos desenvolvidos em “ciência das redes” permitem identificar com impressionante margem de segurança as entidades que merecem mais atenção nos testes exploratórios no procedimento de investigação patrimonial.



15º PRÊMIO DE CRIATIVIDADE E INOVAÇÃO DA RFB

Inovação e Melhoria dos Processos de Trabalho da Receita
Federal do Brasil

Identificação do Real Beneficiário e Busca Patrimonial por
Grafos de Relacionamento no ContÁgil – Script “Análises
Patrimoniais - SABRE”

SETEMBRO DE 2016



Identificação do Real Beneficiário e Busca Patrimonial por Grafos de Relacionamento no ContÁgil – Script “Análises Patrimoniais - SABRE”

Sumário

1 – INTRODUÇÃO.....	2
2 - BIG DATA – A AVALANCHE DE DADOS E A MISSÃO DA RFB.....	4
2.1 – Volume.....	5
2.2 – Variedade.....	7
2.3 – Velocidade.....	9
2.4 – Veracidade.....	12
2.5 – Valor.....	13
3 - REDES.....	14
4 - GRAFOS DE RELACIONAMENTO.....	16
5 - BASE DE RELACIONAMENTOS.....	19
6 - USO PRÁTICO DOS GRAFOS NA INVESTIGAÇÃO PATRIMONIAL.....	23
6.1 - Case 1 – Alaranjamento.....	23
6.2 - Execução de Consultas à Base ou Grafo.....	26
6.3 - Caminhos entre dois (ou mais) Pontos.....	29
6.3.1 – Todos os caminhos.....	29
6.3.2 – Filtros sobre relacionamentos.....	31
6.3.3 – Menores caminhos.....	34
6.4 - Case 2 – Blindagem Patrimonial.....	37
7 - MODELOS MATEMÁTICOS REVELANDO COMPORTAMENTO.....	41
8 - PRINCIPAIS PROBLEMAS ENCONTRADOS E RESPECTIVAS SOLUÇÕES.....	45
9- CONCLUSÃO.....	47
10 – BIBLIOGRAFIA.....	49



1 – INTRODUÇÃO

A satisfação do crédito tributário por vezes se vê frustrada por técnicas de blindagem que procuram afastar do sujeito passivo o lastro patrimonial.

Os especialistas nesse tipo de operação são cada vez mais criativos na criação de estruturas de interposição e formação de organizações jurídicas fictícias capazes de vedar a visão do patrimônio, torna-la opaca ou translúcida; impedindo a Administração Tributária de claramente identificar os recursos financeiros, bens ou direitos que deveriam honrar os compromissos advindos da concretização das regras matrizes de incidência tributária, e, por conseguinte, verter recursos ao erário. Ademais, a ocultação patrimonial faz parte do tipo legal de Lavagem de Dinheiro; procedimento que apresenta com feição lícita recursos amealhados de maneira espúria e extremamente nociva às nações por fomentar concorrência predatória, e chega a interferir no processo democrático, uma vez que o volume de dinheiro branqueado é capaz de eleger parlamentares e membros do poder executivo.

Destarte, dentro do contexto de “realizar ações de pesquisa e investigação” e de “atuar na garantia do crédito tributário”, presentes na Cadeia de Valor da Receita Federal do Brasil; quanto mais modernas as técnicas de blindagem patrimonial; mais modernos e velozes devem ser os mecanismos de identifica-las, possibilitando tratar juridicamente os desdobramentos dessas artimanhas.

Este trabalho discorrerá sobre o uso dos Grafos de Relacionamento (definição no capítulo 4) na identificação de grupos econômicos; fragmentação, interpostas pessoas; esvaziamento patrimonial; blindagem familiar, blindagem descendente (empregados); sucessão empresarial travestida e outras formas; permitindo identificar e alcançar o patrimônio dos reais beneficiários com o fito de garantir a



satisfação do crédito tributário e a recuperação de bens, direitos e valores em favor da União.

Iniciaremos apresentando brevemente o conceito de “*Big Data*” contextualizando-o na Receita Federal do Brasil. Tendo visitado suas nuances, definiremos Grafos, tangenciaremos os conceitos de Rede, e então pretende-se mostrar como o volume de “dados” pode se converter em “informações”, e como essas informações nos conduzem ao “alvo”, permitindo a construção jurídica da situação fática que se disfarçou com o fito de distanciar da obrigação tributária o lastro patrimonial. Neste mister, demonstraremos que:

- ❖ A visão tabular de dados (linhas e colunas) foi superada pela visão gráfica (grafos de relacionamento).
- ❖ A construção da base de relacionamentos com dados internos (RFB) e externos (convênios e dados abertos) aliada às consultas inteligentes e rápidas desnudam a verdade material.
- ❖ As consultas investigativas podem facilmente ser construídas com a ferramenta “análises patrimoniais”, inclusive pelo usuário nível básico do ContÁgil¹.
- ❖ Os modelos matemáticos desenvolvidos em “ciência das redes” permitem identificar com impressionante margem de segurança as entidades que merecem mais atenção nos testes exploratórios no procedimento de investigação patrimonial.

Também como premissa, nos esforçaremos – sempre que possível - em fugir do tecnicismo, uma vez que, considerando que o Big Data está ao nosso redor, a aplicação prática dos conceitos avançados fala por si.

¹ Na opinião do autor, a revolução que o ContÁgil trouxe quanto à lida, interação e cruzamento de dados, vide nota de rodapé número 12, “divide a história da Receita Federal em A.C. e D.C.”- **Antes** do ContÁgil, e **Depois** do ContÁgil. (Nota do autor)



2 - BIG DATA – A AVALANCHE DE DADOS E A MISSÃO DA RFB

O termo “Big Data” não possui uma única definição, pura e inextinguível. Refere-se, grosso modo, ao bombardeio diário de dados e informações dispostos em texto, fotos, vídeos, áudios, *bips*, pulsos e registros de sensores, *flashes* e tudo o mais que a tecnologia criou na modernidade e que minuta em dados estruturados e não estruturados tudo o que acontece no planeta.

É fácil percebê-lo quando ao aproximar-se de uma determinada rede de *fast-food* pulsa no seu celular uma notificação das promoções do dia, instigando-o a parar e “fazer um lanchinho”; ou, após buscar por determinado livro em uma livraria virtual ou site de busca, começam a aparecer mensagens na sua caixa de e-mail e redes sociais, apontando os prováveis livros que “você gostaria de ler”, afinados àquele que você procurou.

“É fácil ver os frutos da sociedade da informação com um celular em cada bolso, um computador em cada mochila e grandes sistemas de tecnologia da informação em todos os escritórios. Mas discreta, contudo, é a informação em si. Meio século depois de os computadores entrarem no meio social, os dados começaram a se acumular e algo novo e especial começou (*sic*) a acontecer.” (MAYER-SCHONBERGER, CUKIER, 2013, p. 04)

O fenômeno, imperceptível a muitos olhos, resume-se nos já famosos “05 V do Big Data” – Volume, Velocidade, Variedade, Veracidade e Valor - que redundam em uma aliteração sonora e sibilante de rapidez.



Figura 1 – BIG DATA. Fonte: Elaboração do Autor

2.1 – Volume

Para cumprir sua missão a Receita Federal do Brasil instituiu ao longo dos anos diversos deveres instrumentais, ditos “obrigações acessórias”, que fornecem ao Estado dados das mais variadas facetas de atividade econômica, financeira e aduaneira. São Cadastros, Declarações; Demonstrativos, Guias de Recolhimento, Escriturações e outros informes dos contribuintes, empresas, bancos, financeiras, e até OSCIP (Organização da Sociedade Civil de Interesse Público), ONG (Organização Não Governamental) e igrejas.

Nessa linha, como projeto revolucionário, sem similar na comunidade internacional, construído com a participação da sociedade, nasceu o SPED² que dentre seus objetivos pretende suprimir gradativamente os citados “D” (de

² Instituído pelo Decreto nº 6.022/2007, o Sistema Público de Escrituração Digital (Sped), consiste na modernização da sistemática de cumprimento das obrigações acessórias, transmitidas pelos contribuintes às administrações tributárias e aos órgãos fiscalizadores, utilizando-se da certificação digital para fins de assinatura dos documentos eletrônicos, garantindo assim a validade jurídica dos mesmos apenas na sua forma digital. (Apresentação do SPED, site institucional <http://sped.rfb.gov.br>)



Declarações e Demonstrativos”), a exemplo do que ocorreu com o DACON e mais recentemente com a DIPJ.

Apenas no âmbito do SPED, comprovando nossas dimensões continentais, o País rapidamente acumulou um volume de dados impressionante, que se coaduna ao conceito de Big Data: as escriturações digitais desde 2013 perfizeram um total de 24 milhões de declarações, que alcança aproximadamente 300 Tb (Terabytes).

	2013	2014	2015
ECD	244.030	261.206	738.015
EFD IPI/ICMS	5.250.518	8.696.436	9.876.324
EFD Contribuições	11.536.563	13.507.233	12.420.630
ECF	0	0	1.218.088
Total	17.031.111	22.464.875	24.253.057

Figura 2 – Escriturações no SPED. Fonte: Dados abertos RFB - Plano Anual da Fiscalização da Receita Federal do Brasil para 2016 e Resultados de 2015³

Nos dados do SPED estão clientes, fornecedores, quadros societários, transportadores, contadores, e tantas outras entidades. Se adicionarmos ao SPED os últimos dez anos de bytes vertidos aos repositórios nacionais de Demonstrativos e Declarações à que os contribuintes estão sujeitos, teremos então um volume realmente colossal.

³ <http://idg.receita.fazenda.gov.br/dados/resultados/fiscalizacao/arquivos-e-imagens/plano-anual-fiscalizacao-2016-e-resultados-2015.pdf>



2.2 – Variedade

Sabemos que a vida do cidadão não se resume a pagar tributos, contudo, suas interações na sociedade produzem dados e documentos que são, por sua vez, vertidos a bancos de dados diversos daqueles institucionais, geridos pelo fisco (especificamente a RFB). Ao adquirir um automóvel a pessoa ou empresa necessita licenciá-lo, e assim alimentará o repositório do Departamento Nacional de Trânsito – **Denatran**. Ao adquirir um imóvel, o serviço notarial ou registral deverá alimentar os dados em uma Central Notarial de Serviços Eletrônicos Compartilhados, o **CENSEC**⁴; ao inscrever-se como eleitor seu cadastro estará nos bancos do Tribunal Superior Eleitoral – **TSE**; quando se adquire uma aeronave as informações ficam nos registros da Agência Nacional de Aviação Civil – **ANAC**; se por sua vez for uma embarcação, vincular-se-á à Capitania dos Portos da Marinha do Brasil, registrando-se no Sistema de Gerenciamento de Embarcações – **SISGEMB** e assim, a cada passo na sociedade a pessoa física ou sua ficção jurídica deixa uma marca; uma “pegada tecnológica”.

Caminhando na aldeia global, ao interagir nas redes sociais, ao inscrever-se em canais de vídeo; “curtir” algo no *Facebook*, as pessoas deixam sua digital na babel cibernética; e cada byte gerado fará parte da constituição da *persona*⁵ no cyber orbe.

Nesse no contexto, a Receita Federal do Brasil administra uma vasta gama de tributos, consoante determina a Constituição Federal, com conseqüente arcabouço delineado pelo Código Tributário Nacional – CTN , Leis, Decretos e normas complementares:

⁴ www.censec.org.br

⁵ Persona vem do Etrusco PHERSU, “máscara teatral”, do Grego PRÓSOPON, “face, máscara” – Site <http://origemdapalavra.com.br/site/palavras/persona/>



- ❖ IRPF (Imposto sobre a renda das pessoas físicas)
- ❖ IRPJ (Imposto sobre a renda das pessoas jurídicas)
- ❖ IRRF (Imposto sobre a renda retido na fonte)
- ❖ CSLL (Contribuição social sobre o lucro líquido)
- ❖ IOF (Imposto sobre operações financeiras)
- ❖ ITR (Imposto territorial rural)
- ❖ IPI (Imposto sobre produtos industrializados)
- ❖ II (Imposto de importação)
- ❖ IE (Imposto de exportação)
- ❖ Contribuições previdenciárias das pessoas físicas
- ❖ Contribuições previdenciárias das pessoas jurídicas
- ❖ Contribuição para o PIS/Pasep e Cofins
- ❖ Cide (Contribuição de intervenção no domínio econômico)
- ❖ AFRMM (Adicional ao frete para renovação da marinha mercante)

Esse plexo de impostos, taxas e contribuições, mais os cadastros de Pessoas Físicas, Pessoas Jurídicas, Imóveis Rurais e Construção Civil (matrículas CEI), constituem rica fonte de informações; que, se isoladamente revelam apenas o que as instruções normativas determinam; quando combinadas permitem enxergar mais do que os declarantes gostariam de informar; e essa também é uma máxima do Big Data.





Figura 3: Big Data RFB – Fonte: Elaboração do Autor

Neste trabalho, mostraremos que ao utilizar-se do volume e variedade de dados, as situações que foram propositadamente ocultadas se deslindam como na montagem de um quebra-cabeças.

2.3 – Velocidade

Em meados do ano 2000 o caso “Claire Swire X Bradley Chait” tornou-se ícone ilustratório da velocidade de propagação da informação no terceiro milênio. A jovem, que trocou apimentadas frases em um simples e-mail, viu seu *affair* tornar-se conhecido de milhões de pessoas em menos de uma semana⁶.

Parodiando Neil Armstrong⁷; a humanidade não caminha; dá saltos! Até 1990, para entregar declarações de Imposto de Renda Pessoa Física era preciso obter o formulário e o manual de preenchimento; debruçar-se sobre uma montanha de papel; fazer contas e escrever com boa letra, ou datilografar cuidadosamente os campos dos formulários. E o recibo de entrega, após carimbado, era uma preciosidade! Documento da mais alta importância, zelosamente guardado em local de honra; pois seu extravio traria dores de cabeça memoráveis.

Por trás de todo o desconforto para o contribuinte existia ainda um grande custo de processamento para tornar, vários meses depois, a informação disponível.

⁶ Claire Swire é uma jovem inglesa que, alguns dias antes (do Natal), havia tido um breve caso com um jovem inglês chamado Bradley Chait. Como mulher moderna que é, ela lhe mandou um e-mail no dia seguinte, “presenteando-o de forma tão atraente” que Chait decidiu compartilhar o presente com seus amigos. Só os melhores amigos, vejam bem – não mais que seis. Mas esses amigos, aparentemente, acharam o presente tão interessante que cada um o repassou para vários dos seus amigos, muitos dos quais, aliás, pensaram da mesma forma. E assim, aquele pequeno e-mail, com o pequeno acréscimo de uma linha feito por Chait “Um belo presente feminino”, passou de mão em mão pelo mundo todo, divertindo cerca de 7 milhões de leitores em questão de dias. Sete milhões! A pobre Claire teve de se esconder para evitar o frenesi da imprensa...(J. WATTS, 2003, P. 116). Ela teria escrito algo como “*I hadn’t swallowed for years, but your cum was yummy*” – Nota do autor.

⁷ Neil Alden Armstrong, astronauta americano, historicamente o primeiro homem a pisar na lua. “*Este é um pequeno passo para um homem, um salto gigantesco para a humanidade*” (nota do autor)



Em 1991 foi criado o primeiro programa, fornecido em disquete, que uma vez alimentadas as informações, fazia os cálculos e gerava um novo disquete contendo a declaração a ser entregue nas unidades da Receita Federal. Um salto imenso em qualidade, conforto, velocidade e confiabilidade.

A maior revolução mesmo, estava por vir. Com o *boom* da internet, tomando uma decisão corajosa e muito controversa na época, inclusive perante a comunidade internacional, a Receita Federal decidiu inovar e, em 1997, pelo programa Receitanet, iniciou o recebimento das declarações pela internet.

Com tal medida, o próprio contribuinte passou a transmitir, confortavelmente de seu escritório ou casa; suas declarações!

“Naquele ano (1997), a Receita Federal registrou a entrega pela Internet de 469 mil Declarações de Imposto de Renda Pessoa Física, no prazo. Em 1999, o número chegou a 5,9 milhões e em 2001 foram registradas 13 milhões de declarações entregues via Receitanet, cerca de 92% do montante”. (Revista Tema – Serpro; acesso online em 07/07/2016)

No ano de 2016 foram entregues pela *web* 28 milhões de declarações!

“Em menos de três segundos, na média, o Receitanet, seguindo as etapas de validação dos dados, compactação, criptografia, envia os dados da declaração pela Internet, gravando ao final, [...] no computador do contribuinte, um recibo eletrônico com o registro da data e hora em que a declaração foi entregue” (Revista Tema – Serpro; acesso online em 07/07/2016)

O incremento na velocidade de recepção dos dados permitiu também considerável aumento na lida com eles; e de fato, 45 dias após recebidas as declarações, o primeiro lote de restituição já está nos bancos.

Por óbvio, as demais declarações e demonstrativos também acompanharam esse progresso, sendo carreadas aos gigabytes ao repositório nacional.



Nesse contexto, revolução não menos importante foi o já citado projeto SPED; particularmente a Nota Fiscal Eletrônica – NFe. Sabe-se que o Brasil não foi o pioneiro em utilizá-la, pois a Espanha, implantou seu modelo em meados de 1990, e o Chile, em 2003⁸. A nota fiscal eletrônica redundou em considerável mudança de paradigma, uma vez que o documento não mais existe fisicamente, sendo um arquivo digital estruturado e autenticado, que permite acompanhamento *pari passu* do fluxo documental tributário, sendo ele, o arquivo, o próprio o documento!

Se antes da nota fiscal eletrônica o fisco tomava conhecimento da transação periodicamente pela prestação de contas via declarações, registro de livros ou fiscalizações pontuais, agora, ao fechar o negócio, o mundo tributário já descortina as repercussões jurídicas, e assim, em instantes, a informação migra para os repositórios.

2.4 – Veracidade

A sabedoria popular diz que “o mentiroso precisa ter boa memória”. Para as informações prestadas ao Fisco Federal, é preciso mais que isso. Falsear, omitir, distorcer, são crimes tributários (*latu sensu*), e quem o pratica assume o risco.

A veracidade diz respeito à confiabilidade, à consistência dos dados. Se alguém omite Rendimentos Tributáveis recebidos de Pessoa Jurídica em sua Declaração de Imposto de Renda, o teste de consistência com a Declaração de

⁸Considerada como a precursora da NFe no mundo, a Espanha deu início a este sistema ainda na década de 1990 e o principal objetivo na época com esta implantação era a redução de custo de emissão de papel. Estima-se, de acordo com a Agência Tributária Espanhola, que a redução de custo girou em torno de 15,7 bilhões de euros. Ou seja, inicialmente, a visão espanhola para a iniciativa era mais voltada à sustentabilidade do que para garantia de recebimento de informações.

Nos mesmos moldes da Espanha, o objetivo do Chile ao aplicar a NF-e, conhecida no país como Factura Eletrônica – desde 2003, foi a redução de gastos. A economia com este projeto, considerando o custo da emissão do papel, segundo o Ministério da Economia do Chile, alcançou 800 milhões de dólares.

<http://www.administradores.com.br/noticias/negocios/nota-fiscal-eletronica-no-brasil-e-pelo-mundo/27442/>



Imposto de Renda Retido na Fonte – DIRF, prestada pela Fonte Pagadora apontará a discrepância. Da mesma forma, se o contribuinte majorar despesas médicas, o teste de consistência estará nas linhas da Declaração de Serviços Médicos e Saúde – DMED.

Consideramos, portanto, que outras técnicas de evasão fiscal surgem para driblar o regramento tributário, uma vez que a veracidade dos dados se torna cada vez mais difícil falsear. A informação da Receita Bruta, por exemplo, é rapidamente cotejada com os documentos fiscais emitidos (NFe); os dispêndios apontados na contabilidade, facilmente colididos com circularização eletrônica, consultando-se no ambiente SPED as notas fiscais emitidas para aquela empresa por seus fornecedores, e assim por diante.

Se agregarmos os convênios e fontes abertas, teremos um cruzamento interessante entre os valores dos imóveis declarados pelos contribuintes e seus correspondentes na Central Notarial (CENSEC) e nos demais fiscos municipais (IPTU) e estaduais (ITBI).

2.5 – Valor

Para o último “V”, lembremos que faz parte dos objetivos da casa “verter recursos aos cofres públicos”, provendo o Estado. Extrair valor da combinação de Volume, Variedade, Velocidade e Veracidade é utilizar os dados com inteligência, transformando-os em informação, aplicando os escassos recursos humanos em auditorias focadas e certeiras, identificando o patrimônio dos que teimam em não dar sua contribuição ao custeio do Estado.

Justifica-se o investimento em infraestrutura de Tecnologia da Informação não apenas para reduzir os custos, contribuir para a sustentabilidade diminuindo o uso



do papel, facilitar o cumprimento das obrigações tributárias pelo contribuinte, identificar condutas perniciosas à sociedade, mas, principalmente para garantir ao Estado seu funcionamento próximo do ideal, em que os recursos públicos possam efetivamente ser revertidos em serviços públicos, retornando à sociedade seu valor.

“Hoje provavelmente somos levados a considerar sites como o Google e a Amazon pioneiros do big data, mas claro que os governos foram os coletores originais de informações em larga escala e ainda concorrem com qualquer empresa privada pelo volume de dados que controlam. [...]

As lições do big data se aplicam tanto ao setor público quanto às empresas privadas: o valor dos dados governamentais é latente e requer inovadora análise para serem aproveitados. Contudo, a despeito de sua posição especial na coleta de informações, os governos têm sido ineficientes no seu uso.” (MAYER-SCHONBERGER, CUKIER, 2013, p. 81) – *grifo nosso*.

3 - REDES

A incapacidade de enxergar nos dados a realidade é paralisante. Com a montanha de dados de que dispomos, construir a malha de ligações entre empresas, sócios, procuradores, clientes, fornecedores e afins é tarefa para processamento refinado de dados, com o uso das tecnologias de que se dispuser. Esse uso inteligente dos dados deve permitir visualizar as inter-relações entre pessoas, empresas e bens/direitos, graduá-las por nível de significância de tal forma que se possa tornar inequívocos os liames jurídicos.

Essa organização em rede é explorada minudentemente por László Barabási no fabuloso “Linked – A nova Ciência dos networks”:



“Hoje, cada vez mais reconhecemos que nada ocorre isoladamente. Muitos eventos e fenômenos se acham conectados, são causados por uma gama de outras partes de um complexo quebra-cabeça universal e com elas interagem. Começamos a perceber que vivemos em um mundo pequeno, em que tudo se encadeia a tudo. Estamos testemunhando uma revolução na maneira como cientistas de todas as diferentes disciplinas descobrem que a complexidade possui uma rigorosa arquitetura. Começamos a perceber a importância das redes” (2002, p. 6)

O fato é que a rede se forma independente da vontade de seus componentes. Melhor dizendo; quando se aplica as técnicas de *big data analytics*, ou simplesmente “*analytics*”, se evidencia os relacionamentos entre as entidades, sejam eles decorrentes de eventos naturais (relação de parentesco, por exemplo), comerciais (clientes, fornecedores) ou jurídicos (procuradores, intervenientes). J. Watts foi muito feliz ao escrever sobre isso já no prefácio do seu “Seis graus de Separação”, em que explora “a evolução da Ciência de Redes em uma era conectada”:

(J. WATTS, 2003, P.1) [...] a ciência das redes é a ciência do mundo real – do mundo das pessoas, amizades, boatos, doenças, modismos, empresas e crises financeiras.

Essa verdade se revela a cada reportagem das operações da Receita Federal (em vôo solo ou em conjunto com a Polícia Federal, Ministério Público ou Procuradoria da Fazenda Nacional), a cada novo escândalo de corrupção ou desvio de dinheiro público, a cada organização criminosa desbaratada, a cada Auto de Infração Milionário lavrado por conluio e associações com fins ilícitos. O avanço da tecnologia trouxe um ritmo nunca antes pensado às investigações e auditorias, e quanto mais os governos e instituições trocam informações entre si, maior o volume de dados e maior a chance de que se deslinde cada nó das redes de corrupção, sonegação, blindagem e ocultação patrimonial.

Nesse cenário, a importância da ciência das redes para a missão da Receita Federal inclui a capacidade de se realizar “buscas” em profundidade ou extensão; ou seja, a capacidade de, tendo uma pessoa ou empresa como foco, vasculhar sua



vizinhança na rede e determinar quais são as que devem responder pela obrigação tributária; ou quais são aquelas que estão guardando o patrimônio amealhado nas operações daquela empresa ou pessoa.

Esse procedimento de busca, sedimentado em algoritmos matemáticos complexos, são transparentes ao usuário do ContÁgil, e especificamente da ferramenta de análises patrimoniais via grafos de relacionamento. Todo o arcabouço técnico, embora alcançável ao “homem médio” por boas horas de estudo, não “incomoda” o investigador, uma vez que se constitui em um simples botão a ser clicado “buscar menores caminhos entre A e B”, por exemplo; ou “buscar todos os caminhos entre C e D”.

*Fiat Lux*⁹! Para construir a visão gráfica capaz de aclarar a penumbra forjada pelos blindadores, os grafos de relacionamentos são lume eficaz, uma vez que evidenciam as entidades e respectivas relações. Todos os passos para se construir uma base de relacionamentos, depurá-la e analisa-la foram contemplados no Script¹⁰ de análises patrimoniais, que culmina em interface visual apresentada em grafos.

4 - GRAFOS DE RELACIONAMENTO

A teoria de Grafos de Relacionamentos foi criada em 1736 pelo grande matemático suíço, Leonhard Euler (lê-se “óiler”). Euler realizou estudos de todos os campos da matemática, física, engenharia e até astronomia, sendo alguns sobre a órbita de Urano, que levaram à descoberta de Plutão. *Opera Omnia*, o registro ainda incompleto de suas obras, perfaz hoje mais de 73 volumes, de 600 páginas cada um.

⁹ Expressão em latim: “Faça-se a luz”, remetendo ao Livro do Gênesis, na Bíblia Sagrada (nota do autor)

¹⁰ De maneira direta, scripts são “roteiros” seguidos por sistemas computacionais e trazem informações que são processadas e transformadas em ações efetuadas por um programa principal.

(<http://www.tecmundo.com.br/programacao/1185-o-que-e-script-.htm>; acesso em 16/09/2016)



Para firmar prova de que não havia possibilidade matemática de se cruzar as sete pontes de Königsberg sem jamais passar pela mesma ponte duas vezes, teve um *insight* fantástico, criando o modelo matemático que somente pôde ser plenamente aproveitado com o desenvolvimento dos computadores, séculos mais tarde. Euler representou cada um dos lugares por um “nó”; “nodo” e os caminhos representou por “links”, “arestas”, resumindo o problema em forma gráfica:

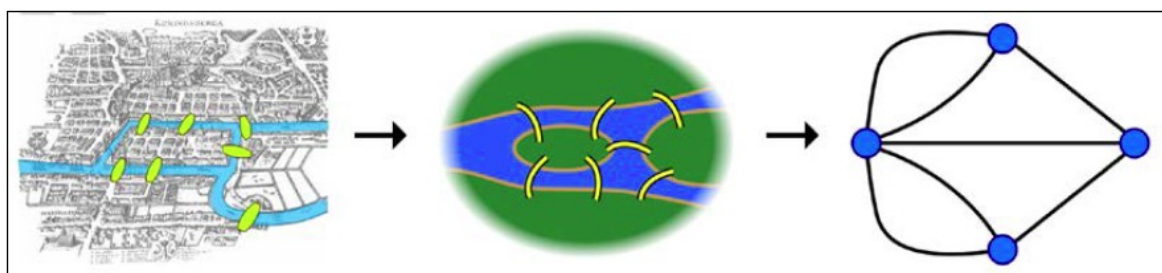


Figura 4 – Pontes de Königsberg – Fonte: e-book Learning Neo4J, Rik Van Bruggen, 2014, P. 9

Tomando-se a ilustração acima, podemos agregar a cada NODO representado uma série de informações, como área, posição, quantidade de pessoas que lá habitam; e a cada ARESTA a sua respectiva extensão (tamanho), qualidade do pavimento, fluxo diário de pessoas, etc. Assim, da mesma forma que Euler identificou os maciços e os caminhos pelas pontes, o uso prático dessa tecnologia nos trabalhos da RFB dá-se atribuindo aos nós (entidades) e arestas (relacionamentos) as propriedades (informações) obtidas nos sistemas institucionais e das parcerias externas :



Propriedade	Valor
APELIDO	NEVOA TYSON
CEP	13273-584
CIDADE	Barueri
CNAE	7319002
DATA_FINAL	14/08/2009
DATA_INICI...	16/04/2007
DIMOF_CR...	35.611.177,78
DIMOF_DE...	35.963.670,11
ENDERECO	AV SIVIERO MITANI,...
ICONE	grafos_empresa2....

Figura 5: Entidades e suas Propriedades – Fonte: Elaboração do Autor, no ContÁgil. Dados ofuscados, embaralhados, não correspondentes à pessoas ou coisas reais.

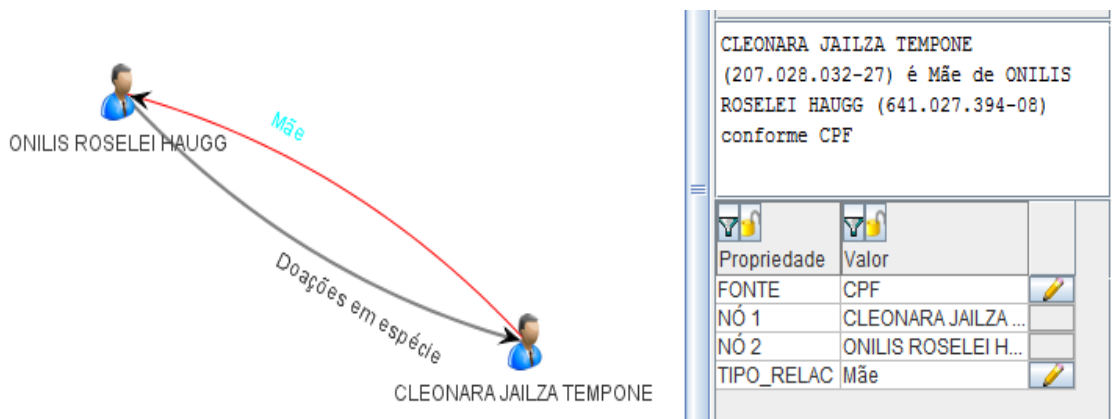


Figura 6: Relacionamentos e suas Propriedades – Fonte: Elaboração do Autor, no ContÁgil (Dados ofuscados, fictícios, inclusive número CPF)

Perfeita subsunção ao fenômeno do Big Data, pois, alterar o modo de lidar com os dados, é pura necessidade. A visão clássica (relacional) de tabelas (linhas e colunas) que se relacionam por uma ou mais campos (chaves) contidas em células (registros) de formato e tamanho rigidamente determinado tornou-se incapaz de manipular o volume e a variedade dos dados com velocidade, testar sua veracidade e então deles extrair valor.

Essa mudança de concepção é impulsionada pelo crescimento exponencial da internet, e pela proliferação desmedida dos dados “não estruturados”, dispersos em bilhões de posts, *tweets*, imagens e vídeos. Esse conceito novo de “NoSQL”



(aplicado livremente como “Not Exactly SQL” – não “exatamente” SQL) não prescinde dos modelos tradicionais de dados, mas reforça-o:

“...é melhor pensar, em NoSQL como um movimento em vez de uma tecnologia. Não achamos que os bancos de dados relacionais acabarão, pois eles ainda serão a forma mais comum de banco de dados em uso. [...]

A diferença significativa é que agora vemos os bancos de dados relacionais como uma opção para o armazenamento de dados. Esse ponto de vista é, muitas vezes, chamado de persistência poliglota – utilizar diferentes armazenamentos de dados em diferentes circunstâncias” (SADALAGE, J, FOWLER, MARTIN, 2013, p.37)

Partindo dessa percepção, temos que, com a variedade de informações nos bancos de dados institucionais, é natural que mais uma das características do Big Data, a “confusão” (e aqui nos referimos à profusão de dados e sistemas), esteja presente na RFB.

5 - BASE DE RELACIONAMENTOS

A teoria criada por Euler, assim como os fantásticos prenúncios lógicos de Boole e Venn¹¹, somente possibilitaram e puderam ser fortemente utilizados a partir da invenção dos computadores eletrônicos e posteriormente o desenvolvimento dos processadores digitais.

Esses processadores recebem, depuram, tratam e organizam volume cada vez maior de dados e permitem que deles se extraia informações, cujos registros serão, por vezes, redundantes, e; no caso da Receita Federal do Brasil, presentes

¹¹ Referência do autor à George Boole, criador da lógica booleana e operadores booleanos - AND, OR, NOT (E, OU, NÃO) - que tornou possível a construção dos modernos “programas computacionais”; e a John Venn cujo desenvolvimento da lógica booleana levou à criação da teoria dos conjuntos, com a representação gráfica das uniões e intersecções.



em vários sistemas, em uma série de plataformas operacionais e bancos de dados, na ocorrência real da já referida “persistência poliglota”.

Como elemento agregador, capaz de transitar com desenvoltura pelos sistemas institucionais diversos, mantidos pelos prestadores de serviço Serpro; Dataprev, e pelas variadas formas de autenticação e acesso, inclusive ao DW – Data Warehouse, um brilhante Auditor Fiscal da Receita Federal do Brasil criou em meados de 2008, o ContÁgil¹²; “uma ferramenta de fazer ferramentas”; e é nele que se desenhou o instrumento “análises patrimoniais”.

O script de análises patrimoniais utiliza as funcionalidades do ContÁgil para transformar os dados importados em formato “declaração”, com suas fichas, formulários, células, linhas e colunas em uma visão específica de entidades, relacionamentos e propriedades, alimentar uma base de relacionamentos¹³, e então poder visualizá-la num Grafo.

Construída para ser utilizada pelos usuários de nível básico do aplicativo, a ferramenta de análises patrimoniais traz em 04 passos os procedimentos para a investigação:

¹² Desenvolvido integralmente por servidores da própria Receita Federal do Brasil, o ContÁgil é um aplicativo de apoio às atividades de fiscalização que tem como objetivo a análise e auditoria fiscal de contribuintes a partir de cruzamento de informações oriundas de fontes internas e externas e daquelas coletadas junto ao próprio contribuinte ou a terceiros. O sistema, homologado em julho de 2009, já abrange funcionalidades da auditoria fazendária, previdenciária e aduaneira [...] (<http://www.anfip.org.br/informacoes/noticias> - acesso em 08/07/2016)

¹³ Bancos de Dados de grafos permitem que você armazene entidades e também relacionamentos entre essas entidades. Entidades também são conhecidas como nodos, os quais possuem propriedades. Os relacionamentos são conhecidos como arestas que podem ter propriedades. As arestas têm significância direcional; nodos são organizados por relacionamentos que permitem que você encontre padrões interessantes entre eles. A organização do grafo permite que os dados sejam armazenados uma vez e depois interpretados de formas diferentes baseadas em relacionamentos (SADALAGE, J, FOWLER, MARTIN, 2013, p.161)

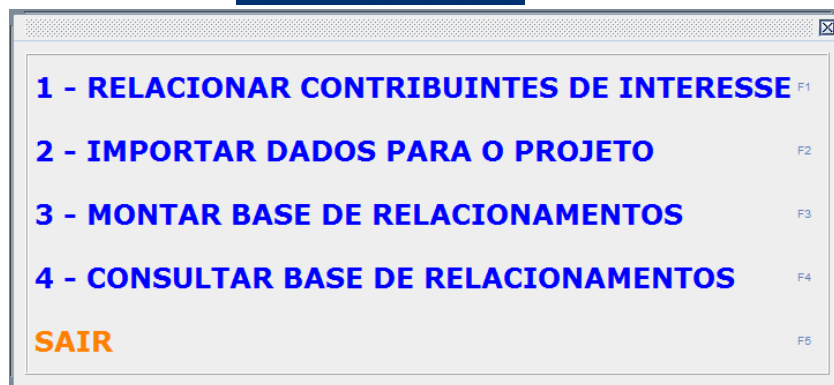


Figura 7 – Fonte: Elaboração do Autor, no ContÁgil

No passo-1, alimenta-se um CPF ou CNPJ; ou uma lista deles; podendo coexistir ambos formatos, CPF e CNPJ. Quaisquer informações que usuário tiver reunido sobre essas pessoas, alvos da ação, pessoas relacionadas, eventos, lugares, veículos, também serão úteis durante a investigação patrimonial.

Se, por exemplo, desconfia-se que se trata de um “grupo econômico”, essas entidades poderão já, de imediato, ter seus CNPJ alimentados no passo-1; onde já se pode, à escolha do usuário, abranger todos os estabelecimentos das empresas. Caso se disponha de informações que apontem outros relacionados, pessoas físicas ou jurídicas, possíveis detentores de patrimônio resultante de atividade da empresa foco, ou participantes da operação de lavagem de dinheiro, também já se pode inseri-los, uma vez que as ligações porventura existentes ficarão evidenciadas à medida em que as fontes de dados forem consultadas.

Para esta lista, no passo-2, escolheremos as bases de informações que deverão compor a primeira análise, dentre aquelas referidas no capítulo 2, e outras já disponíveis.

Tendo baixado os dados para o Projeto em trabalho, executa-se o passo 3, montando a base de relacionamentos. Nesse momento, o ContÁgil percorre os



dados do Projeto, e para cada entidade encontrada constrói seus relacionamentos com as outras entidades, qualificando-o e inserindo as respectivas propriedades, que permitirão consultas direcionadas, no passo-4.

Por exemplo; se uma entidade é pessoa física, do Cadastro CPF obtém-se o nome de sua mãe; que é outra entidade. Obtém-se também endereço e data de nascimento, propriedades da entidade.

Ao percorrer os dados da Declaração de Imposto de Renda Pessoa Física, teremos a entidade informada como Fonte Pagadora, e o respectivo valor do Relacionamento. Teremos também os vínculos dos “Pagamentos Efetuados” a médicos, dentistas, planos de saúde, clínicas, escolas, locadores, etc. Para cada novo relacionamento encontrado, o ContÁgil lhe agrega o valor e outras propriedades. Da relação de bens e direitos, as informações são carreadas, por exemplo, como “proprietário” de veículo; e, neste caso, o veículo é uma nova entidade, que, se tiver a placa informada, este será seu identificador, o que possibilitará complementação de dados via RENAVAM. A cada nova base consultada¹⁴, os dados são transformados em informações *strictu sensu*, uma vez que se tornam inteligíveis para análises.

Base a base, uma a uma, as camadas de informações vão se sobrepondo, complementando-se. Ao encontrar uma entidade nova, o ContÁgil a cria; se essa entidade já existe, apenas lhe atualiza ou constrói o relacionamento que ainda não existia. Esse conhecimento traz, consigo uma outra característica do Big Data: “redundância”; e é essa redundância de informações que ratifica a “veracidade”.

¹⁴ Abstemo-nos de relacionar todas as bases acessadas e de mostrar o menu de importações, por questões estratégicas institucionais.



Explicando melhor; ao consultar o Cadastro CNPJ, teremos o Quadro Societário, responsável pelo CNPJ, administradores, etc. Ao incorporar os dados da DIPJ, essas informações também estarão lá, e serão também aproveitadas, repetindo o relacionamento. Essa aparente duplicidade de informações possibilita refinar qual é a melhor e mais acertada, perscrutar informações inexatas, e então deixar o relacionamento depurado, conciso e preciso. Esse procedimento de interação do usuário com os dados dá-lhe um conhecimento íntimo das características dos relacionamentos, haja vista que, para depurar adequadamente faz-se preciso estudar a qualidade da informação.

6 - USO PRÁTICO DOS GRAFOS NA INVESTIGAÇÃO PATRIMONIAL

6.1 - Case 1 – Alaranjamento

Tomemos um clássico da blindagem; o “*alaranjamento*”; em que se utiliza de pessoa sem capacidade contributiva para aparecer como responsável pelos tributos. A figura 8 retrata uma situação concreta (*case -1*) apresentada com dados ofuscados¹⁵ que foi deslindada em poucos minutos utilizando-se da técnica que ora se apresenta:

A empresa Nevoa Tyson (A), teve movimentação financeira da ordem de 30 milhões de reais no ano-calendário, apresentou declarações zeradas, não constituiu nem recolheu tributos. Tanto a empresa quanto seus sócios Hayrton Oto (B) e Rivas Daniele (C) não possuem patrimônio para solver os créditos tributários que deveriam ser constituídos por decorrência do exercício da atividade comercial.

Eis o quadro inicial, em dados tabulares:

¹⁵ Ofuscamento é um procedimento que faz no “ContÁgil”, em que um algoritmo modifica completamente as propriedades de uma entidade (pessoas físicas ou jurídicas, bens, direitos, números de telefone, contas bancárias, etc) tornando-a não correspondente ao mundo real. Utiliza-se para fins de produção de material de treinamento e testes.



ENTIDADE		DIMOF	PATRIMÔNIO
Nevoa Tyson	A	R\$ 35,6 Milhões	<i>nihil</i>
Hayrton Oto	B	R\$ 250 mil	<i>nihil</i>
Rivas Daniele	C	R\$ 7 mil	<i>nihil</i>

Ao utilizarmos a tecnologia que será exposta, a alimentação gradativa das diversas fontes de informações permite ver além dos dados, demonstrando graficamente que há duas sócias excluídas (linhas pontilhadas), Leonice Takashi (D) e Onilis Roselei (E), que também são sócias de outra empresa, a Embuste Ltda (H), onde os sócios de A eram empregados.

O Grafo ainda revela que D e E são filhas de Cleonara Jailza (F), que também é sócia-administradora de H, com 99% das cotas.

Percebe-se ainda, que F possui como procurador o contador de A; é sócia de outras 20 empresas (patrimônio, receita e outros relacionamentos não estão apresentados aqui); concentra o patrimônio familiar, inclusive um helicóptero de R\$ 2,5 milhões.

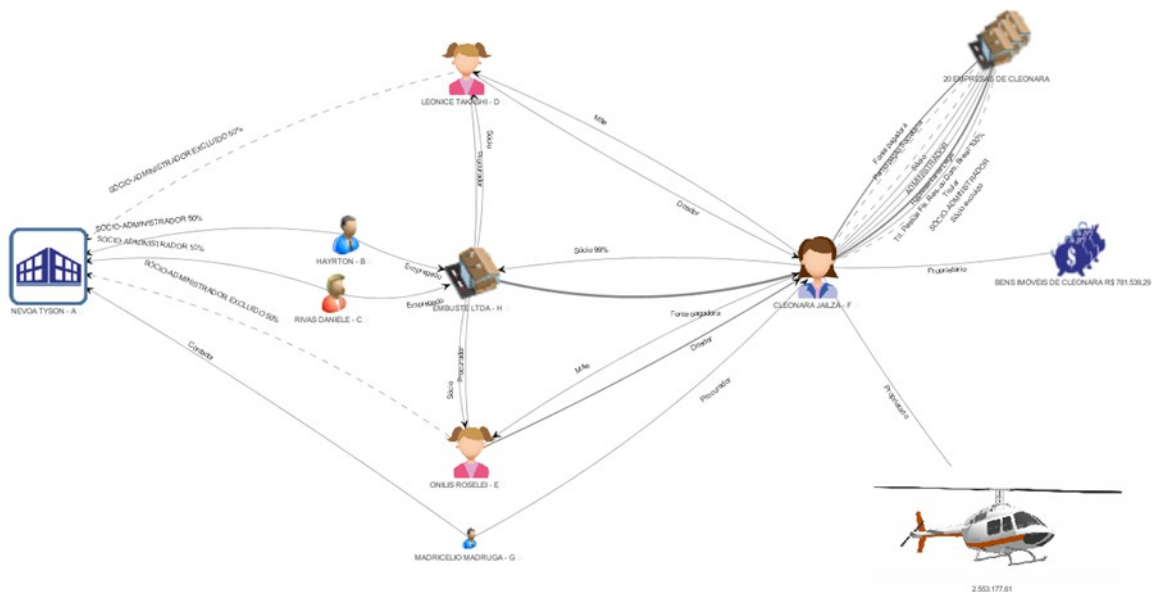


Figura 8 – Case 1 Criação do Autor no ContÁgil



Revisitando os conceitos expostos lembramos que a construção do GRAFO supra utilizou dados de variados sistemas, que, isoladamente cumprem as funções para que foram criados, mas, reunidos, revelam muito mais do que sua concepção inicial previu. Esse conceito é explorado por Viktor Mayer – Schönberger e Kenneth Cukier, na obra “Big Data”, quando ao falar sobre Valor, discorrem sobre Dados reutilizáveis, recombinantes e expansíveis:

VALOR

Ao contrário das coisas materiais[...]o valor dos dados não diminui com o uso; ele pode ser reprocessado. Os dados são o que os economistas chamam de “bem sem concorrente”: seu uso por uma pessoa não impede o uso por outra, e a informação não se desgasta como as coisas materiais[...]

Assim como os dados podem ser usados várias vezes com o mesmo objetivo, o mais importante é que podem ser utilizados com vários objetivos também (PG 71) – grifos nossos

DADOS RECOMBINANTES

Às vezes, o valor latente só pode ser extraído pela combinação de dois bancos de dados, talvez muito diferentes um do outro. Podemos inovar ao combinar dados de novas maneiras. (PG 75) – grifos nossos

DADOS EXPANSÍVEIS

Uma maneira de permitir a reutilização dos dados é projetar expansibilidade desde o início, de modo que sejam adequados a vários usos. Apesar de não ser sempre possível – já que só podemos perceber possíveis usos depois que os dados foram coletados -, há meios de encorajar o uso diversificado de um mesmo banco de dados. (PG 76) – grifos nossos

Neste caso prático houve a lavratura do Auto de Infração com solidariedade passiva tributária com D, E e F, sendo mantida os órgãos julgadores especificamente a responsabilidade tributária de F.

Não se pode, contudo, confundir Base de Relacionamentos com Grafo de Relacionamentos. No caso supra, construindo a base com dados partindo apenas do CNPJ inicial, chegou-se a 3.339 nós; ou seja, 3 milhares de entidades. Em razão do



volume de informações obtidas, montar uma base com muitas entidades e relacionamentos é corriqueiro, comum; mas apresentar um Grafo com todas essas milhares de entidades, é usar de maneira não inteligente a tecnologia e obter uma “visualização inútil”:

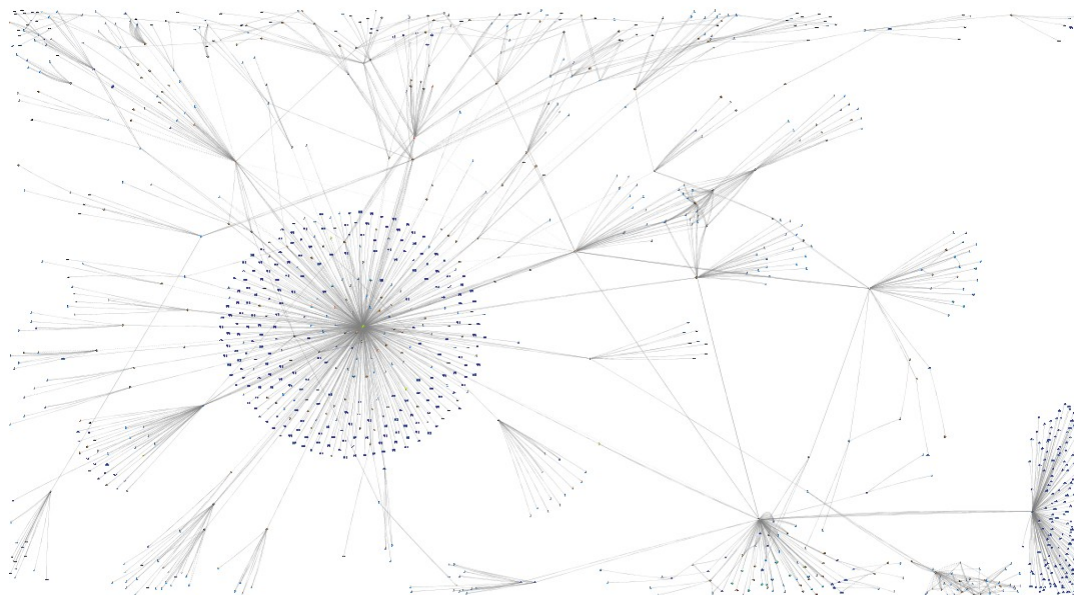


Figura 9: Grafo “Nuvem” - Criação do Autor no ContÁgil

Conforme demonstramos acima, ao lidar com empresas de pequeno porte, facilmente se chega a alguns milhares de nós. Empresas médias, a depender da área de atuação, meia centena de milhar de nós; e empresas grandes, daí por diante, podendo atingir a centena de milhar de entidades, ou mais.

Encontrar o que se procura; visualizar o que se tentava esconder no mar de dados que se formou, portanto, é missão para ferramentas de big data; para as consultas, as varreduras, as buscas inteligentes na rede de relacionamentos.

6.2 - Execução de Consultas à Base ou Grafo



Consultar um grafo é, em linguajar técnico, realizar sua “travessia”, perpassando os pontos (nodos) consoante um comando preestabelecido. Quanto mais rica em termos de dados for a base, mais precisos e úteis os critérios de busca e pesquisa. O usuário dirá: *“ContÁgil, olhe para essa pessoa e busque, em até 02 níveis de separação, os nós vizinhos que tiveram movimentação financeira superior a R\$ 2 milhões de reais”*.

Uma consulta no grafo também é conhecida como travessia do grafo. Uma vantagem dos bancos de dados de grafos é que podemos alterar os requisitos de travessia sem ter de alterar os nodos ou arestas. Se quisermos “obter todos os nodos que curtiram o NoSql essencial”, podemos fazê-lo sem ter de alterar os dados existentes ou o modelo do banco de dados, porque podemos percorrer o grafo da forma que quisermos. (SADALAGE, J, FOWLER, MARTIN, 2013, p.162)

No ContÁgil; qualquer consulta que se deseje realizar sobre a base é possível aos usuários avançados. O que quisemos ao construir a ferramenta de análises patrimoniais foi que os usuários básicos, que possuam muito conhecimento do seu processo de trabalho, e nível básico de conhecimento no uso do aplicativo, pudessem realizar suas pesquisas com a mesma maestria dos avançados.

Para que isso fosse possível, o passo-4 traz algumas consultas direcionadas, que facilitam sobremaneira o estudo da Base de Relacionamentos:

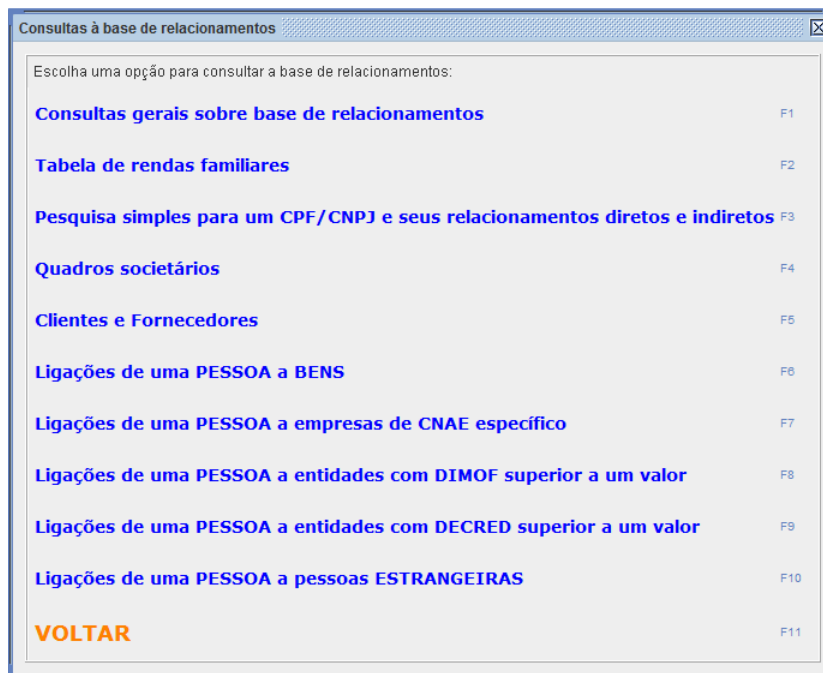


Figura 10: Consultas Pré-construídas - Criação do Autor no ContÁgil

A primeira opção “Consultas gerais sobre a base de relacionamentos” remeterá ao menu “normal” do ContÁgil, em que os usuários mais avançados podem realizar consultas diversas, criando seus próprios critérios, filtrando relacionamentos, e aplicar, utilizando um console de comandos, consultas em *Cypher*¹⁶, a linguagem de computação que o ContÁgil executa para as consultas pré-programadas e que é “invisível” ao usuário. Conforme Onofrio Panzarino (ver nota de rodapé 16 – do original, em inglês), antes do *Cypher*, consultar uma base de relacionamentos no Banco de Dados Neo4J era um “pesadelo”. O *Cypher* utiliza sintaxe declarativa, de escrita e leitura fácil e poderosa, de tal modo que se pode rapidamente definir e reconhecer padrões durante a “visita” ao grafo.

¹⁶ Before Neo4j introduced Cypher as a preferred query, utilizing Neo4j in a real-world Project was a difficult compare to a traditional relational database. In particular, querying the database was a nightmare, and executing a complex query required the user to write an object, thereby performing a graph traversal. Roughly speaking, a traversal is an operation that specifies how to traverse a graph and what to do with the nodes and relationships found during the visit. Thought is very powerful, it works in a very procedural way (*thought callbacks*), so its readability is poor and any chance to the query means modifying the code and building it.

Cypher, instead, provides a declarative syntax, which is readable and powerful, and a rich set of graph patterns that can be recognized in the graph. Thus, with Cypher, you can write (and read) queries much more easily and be productive from the beginning. (PANZARINO, ONOFRIO, 2014, preface, p.1)



Atualmente foram programadas outras consultas estratégicas que combinam critérios de variação patrimonial, movimentação financeira e outras informações.

No menu da figura 10, vê-se que basta escolher uma das pesquisas, “ancorar” em um alvo, e estudar as relações que serão graficamente apresentadas, podendo expandir os níveis, ou buscar novas informações a partir do próprio nó, da visualização que estiver na tela.

6.3 - Caminhos entre dois (ou mais) Pontos

6.3.1 – Todos os caminhos

No case-1, após formar a base de relacionamentos com as primeiras informações institucionais, uma “consulta de ligações de uma pessoa a entidades com DIMOF superior a R\$ 2 milhões”; até 3 níveis de separação da empresa foco, trouxe os ex-sócios como liame e a real beneficiária, Cleonara, com a respectiva movimentação financeira, da ordem de 3 milhões de reais:

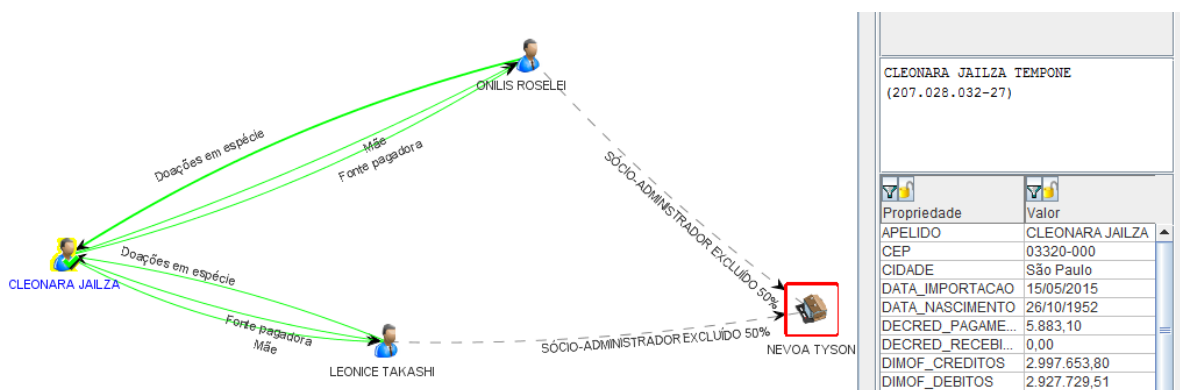


Figura 11: Consultas DIMOF superior a 2 milhões, 3 níveis a partir do Alvo - Criação do Autor no ContÁgil



Os primeiros indícios de alaranjamento precisavam ser aclarados. Existiriam relacionamentos entre a entidade “Cleonara” e a empresa “Névoa” que revelem relações jurídicas que conduzissem à solidariedade passiva tributária, por exemplo ?

Uma consulta “P2P” ou ponto a ponto, conceito da teoria de redes (*peer to peer* – par a par, ponto a ponto) é elucidativa. No capítulo “Busca nas Redes”, da obra “Seis Graus de Separação (Six Degress)”, Duncan J Watts discorre:

[...] há também um motivo prático para entender as buscas direcionadas nas redes: o processo de encontrar uma pessoa em uma rede social através de uma cadeia de conhecidos imediatos, é essencialmente, o mesmo usado para encontrar um arquivo ou outra informação específica em uma base de dados distribuída. Bastante atenção tem sido dada recentemente ao potencial das chamadas redes ponto a ponto (*peer-to-peer networks*) [...] (p.110)

Para realizar essa “busca direcionada” usando o Script de Análises Patrimoniais no ContÁgil, basta escolher a opção “Consultas Gerais sobre a Base” e na opção “Pesquisar Menores Caminhos ou Todos os Caminhos” alimentar as “duas ou mais pontas” entre as quais se deseja visualizar os relacionamentos e especificar quantos níveis devem ser considerados na pesquisa:

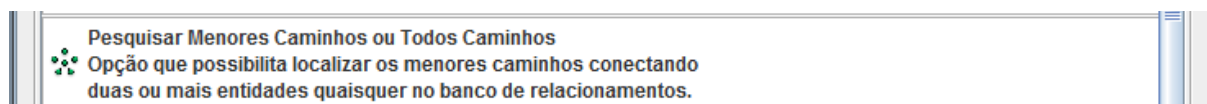


Figura 12: Consultas por menores caminhos ou todos os caminhos - Criação do Autor no ContÁgil



Figura 13: Parâmetros da consulta por todos os caminhos - Criação do Autor no ContÁgil

Para o Case-1, o resultado da consulta efetuada com os parâmetros supra, dentre as mais de 3 mil entidades, traz o Grafo a ser utilizado no próximo passo da investigação para depurar os relacionamentos. E pode-se ainda realizar nova consulta sobre o resultado da própria consulta, em analogia com as “consultas sobre consultas” dos bancos de dados relacionais, para peneirar, filtrar, clarificar a visualização:

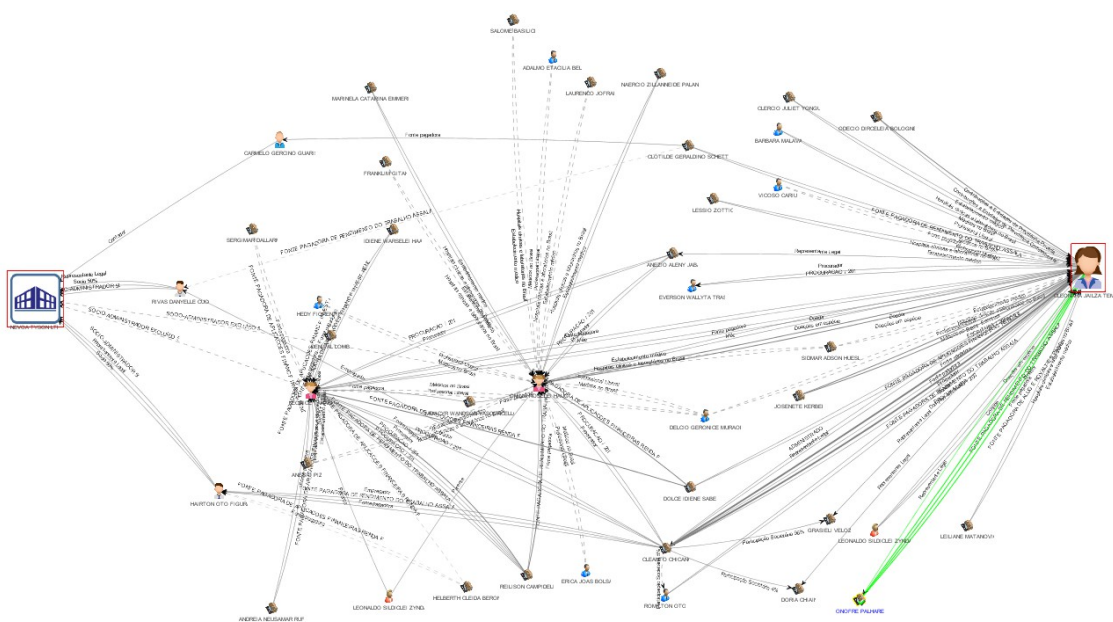




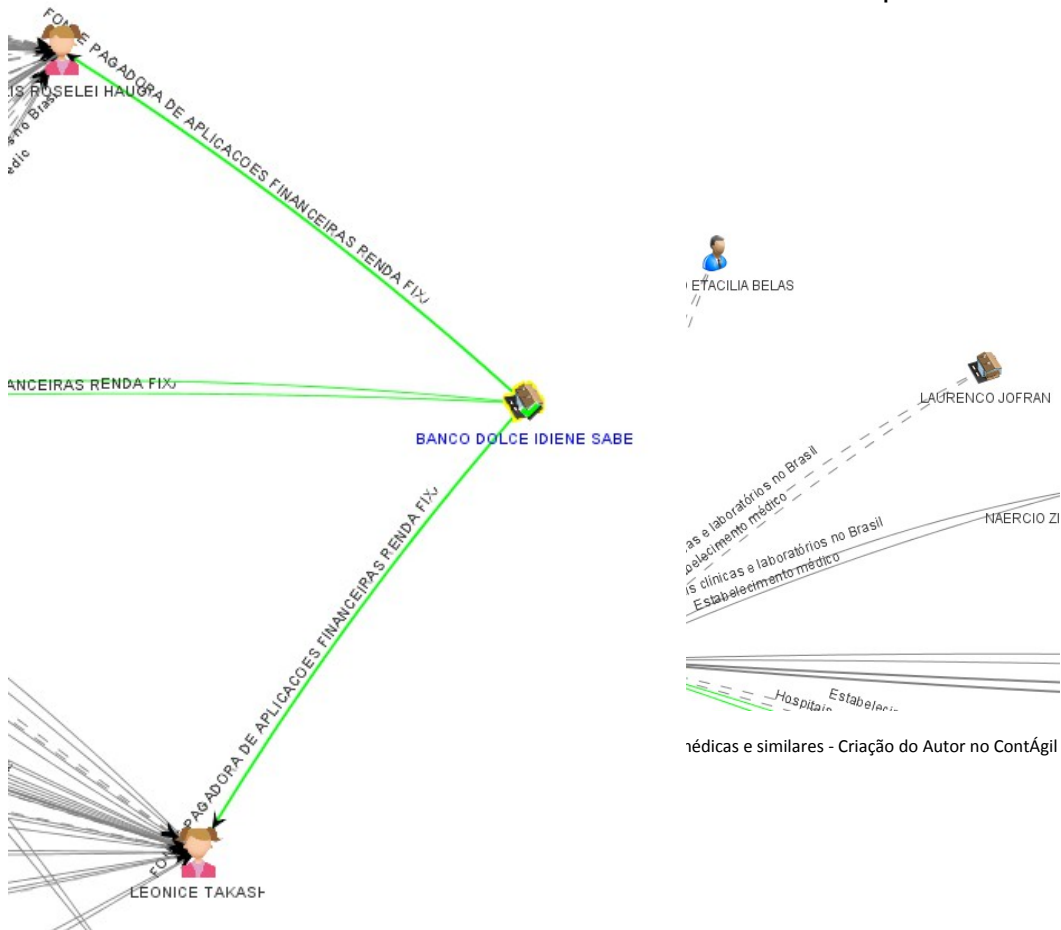
Figura 14: Grafo a ser depurado - Resultado da Consulta - Criação do Autor no ContÁgil

6.3.2 – Filtros sobre relacionamentos

Antes de mencionarmos a opção de “caminho mais curto”, chamamos a atenção para a opção “Filtrar Relacionamentos”. Essa opção permite ao Auditor Fiscal escolher, dentre os relacionamentos da base em estudo, quais ele deseja que sejam considerados; ou em outro foco, quais ele deseja que sejam ignorados.

Surgida de necessidade apresentada nos trabalhos efetuados, a praticidade da opção é latente e moldável, pois, os relacionamentos que não interessam para um tipo de trabalho, podem ser determinantes para outros.

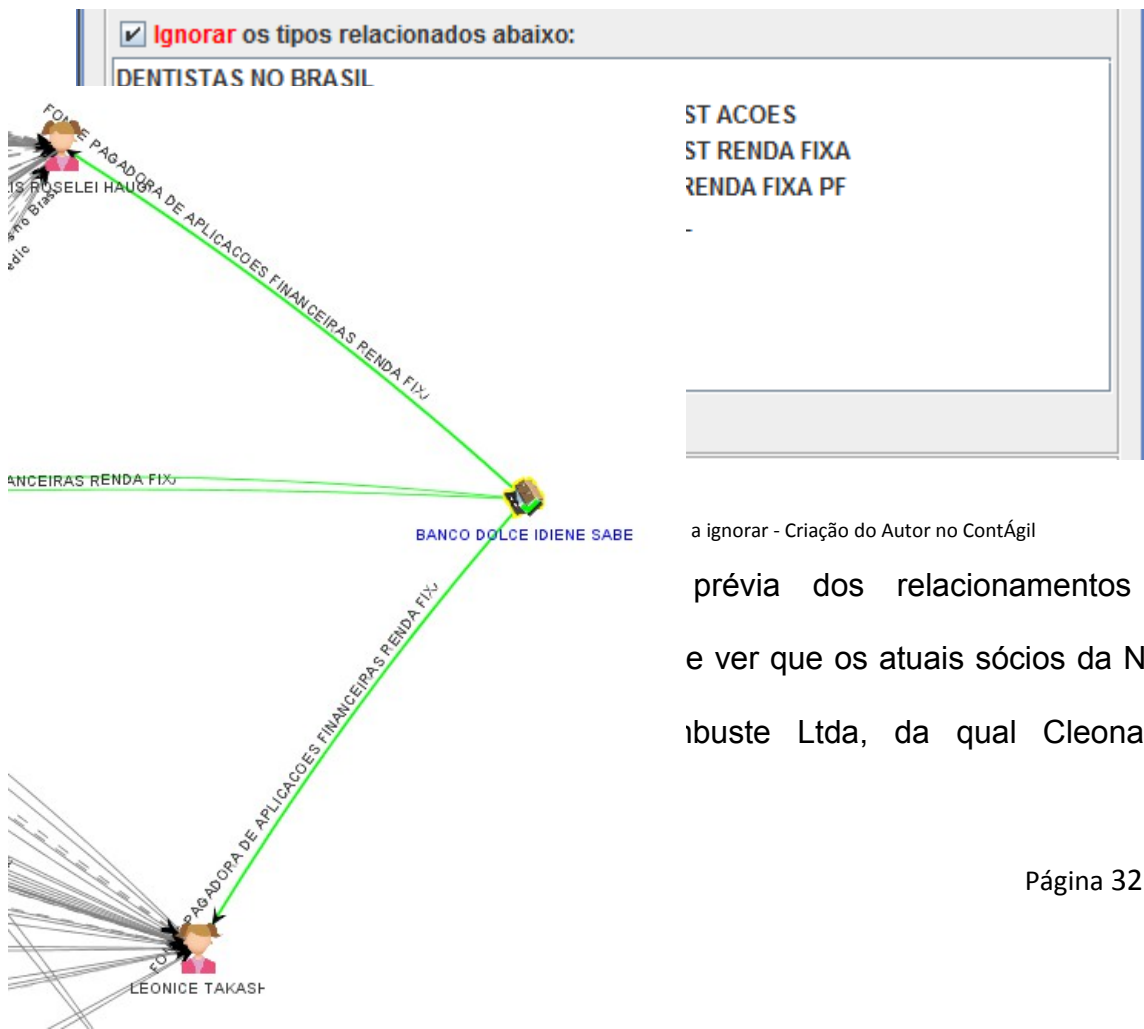
Por exemplo, no Grafo resultado da consulta, exibido em “zoom”, detalhe na figura 15, temos os relacionamentos de médicos, dentistas, clínicas e hospitais sob as várias nomenclaturas, povoados na base pela Declaração de Imposto de Renda Pessoa-Física. Se, nessa investigação, a qualidade desse relacionamento não nos interessa: em uma investigação de fraudes em despesas médicas seria





Outro exemplo: “Fonte Pagadora de Aplicações Financeiras de Renda Fixa”; dados povoados na base pela DIRF - Declaração de Imposto de Renda Retido na Fonte. No contexto, pode não nos interessar ver esse relacionamento de imediato, pois duas pessoas podem possuir aplicações financeiras em um mesmo banco, e neste caso, se os relacionamentos entre elas se dá pelo exclusivo fato de serem correntistas na mesma instituição financeira, isto não as torna relacionadas para os fins da investigação. Contudo, se o que se estivermos estudando for justamente os fluxos financeiros, em uma operação de simulação, interveniência, e outros tipos de fraudes, o mesmo relacionamento já assumiria uma gradação relevante.

Pode-se, portanto, filtrar previamente esses e outros tipos de relacionamentos, construindo uma visão mais clara dos caminhos que ligam dois ou mais nós. (É claro que a depuração do Grafo também pode ser feita posteriormente, ocultando ou excluindo nós e relacionamentos que não pertencem àquele contexto).



a ignorar - Criação do Autor no ContÁgil

prévia dos relacionamentos que e ver que os atuais sócios da Névoa ibuste Ltda, da qual Cleonara é



Administradora, Representante Legal e beneficiária de valores expressivos (representados pela espessura das setas de relacionamentos):

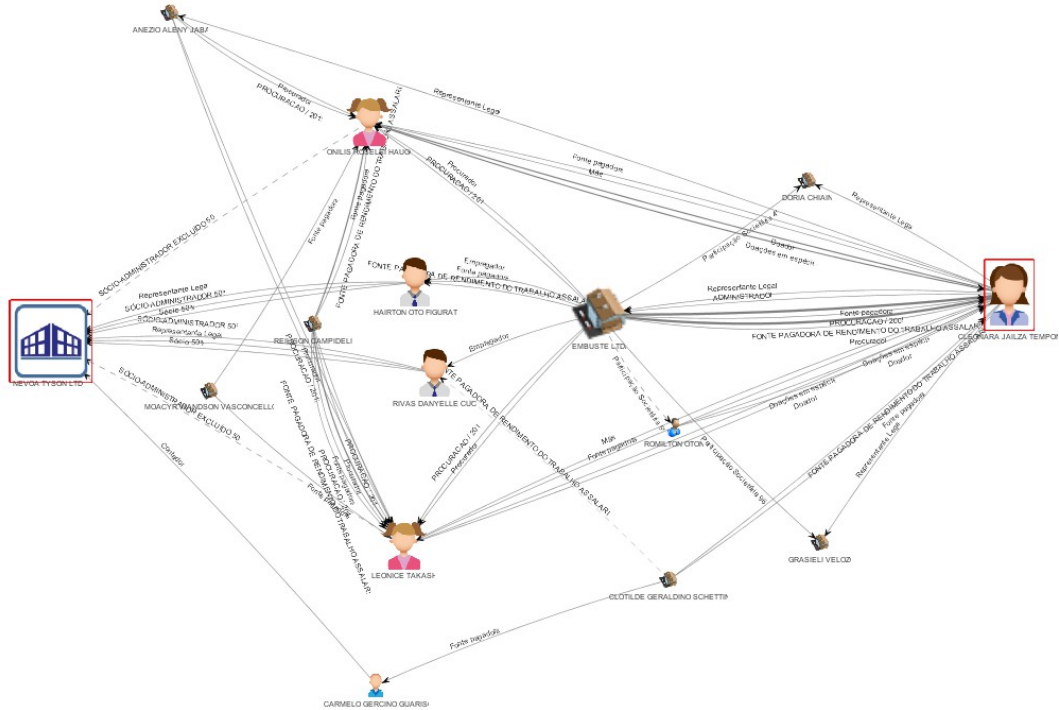
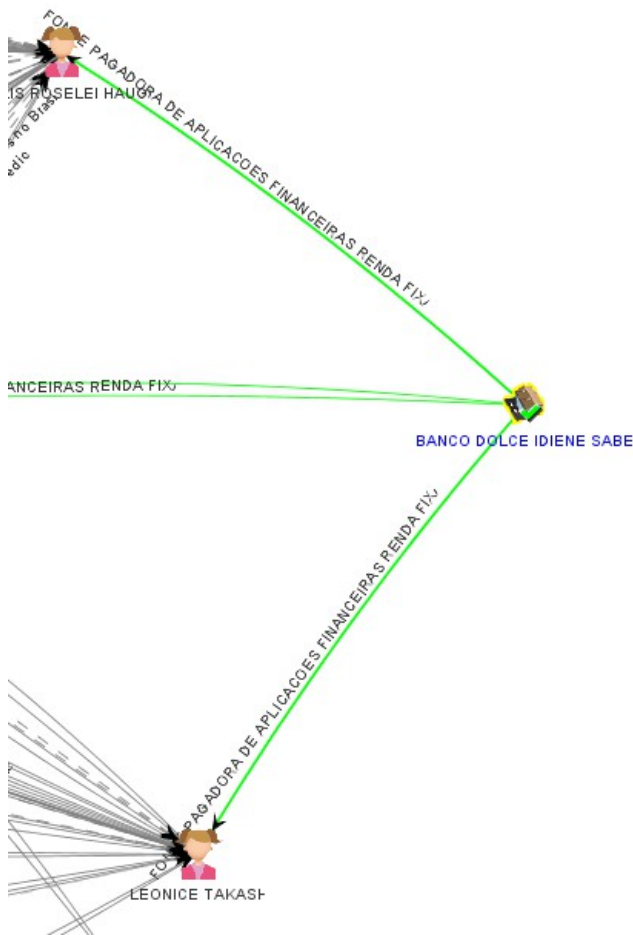


Figura 17: Resultado da consulta após o filtro prévio dos relacionamentos a ignorar - Criação do Autor no Contágil



hos é construída por algoritmos que os Eulerianos” e “Caminhos e Grafos problema do Carteiro Chinês” e o conteúdo dos Cursos de Ciências da gesta”, assim, não nos demoraremos atto e Jurkiewicz, na obra “Grafos –



Euler verificou que o passeio fechado que desejavam os habitantes da cidade, passando uma vez e uma só em cada ponte, não era possível porque os quatro vértices do grafo tinham grau ímpar.

Cabe questionar a razão pela qual ainda se discute este problema, mais de 270 anos depois de sua formulação.

Na verdade, se ele servisse apenas para achar passeios em pontes, já estaria há muito esquecido: de fato, porém, não é difícil imaginarmos situações atuais nas quais haja interesse em se achar um percurso que passe, por exemplo, pelas ruas de um bairro ou pelos corredores de um prédio, uma vez e uma só, voltando ao ponto de partida. Problemas de vendas em domicílio, de coleta de lixo, de entrega de correio, de visitas a museus ou a algumas grandes lojas, são desse tipo.

E como vamos aplicar o teorema a um grafo que tenha vértices de grau ímpar?

De fato, teremos de procurar uma solução mais geral: não podemos esperar que todos os grafos que apareçam em nossos modelos tenham apenas vértices de graus par. A solução é dada pelo teorema a seguir: "O número mínimo de percursos que particionam o conjunto de arestas de um grafo $G = (V, E)$ não orientado e conexo, com $2K$ vértices de grau ímpar é $k(k \in \mathbb{N} - \{0\})$ ". [127]

O Problema do Carteiro Chinês

Em uma situação aplicada se tem geralmente um grafo valorado (por distâncias, tempos, custos, etc) e o interesse estará na repetição de itinerários parciais, de modo a gerar um itinerário único, que será um percurso pré-euleriano. Por exemplo, se o problema for de coleta de lixo, será necessário determinar em quais ruas o caminhão deverá passar novamente (sem nova coleta) para retomar seu trabalho mais adiante. Isto pode ser feito transformando-se o grafo original em um grafo euleriano, andando-se ao inverso da prova do teorema. Ou seja, partimos de um grafo não euleriano e colocamos arestas que o tornem euleriano. [135]

Já um Grafo hamiltoniano é um grafo que possui um percurso abrangente em relação aos vértices, fechado e que não repita nenhum vértice.

Em linhas gerais, um algoritmo matemático calcula a menor rota entre os pontos de um Grafo, e define os menores caminhos, considerando um (ou mais) pontos de partida e retorno.

Aqui aplicamos a consulta por menores caminhos, calculamos o caminho mais curto entre os alvos, e depuramos (filtrando) também os relacionamentos que não agregam valor ao raciocínio da investigação.

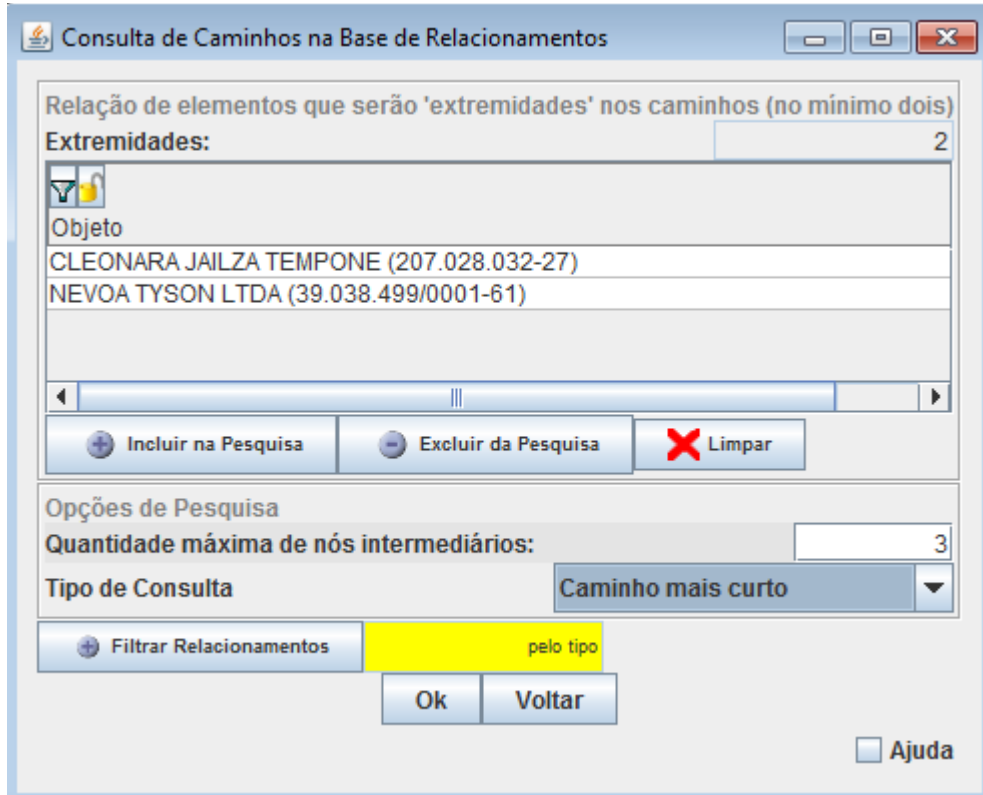
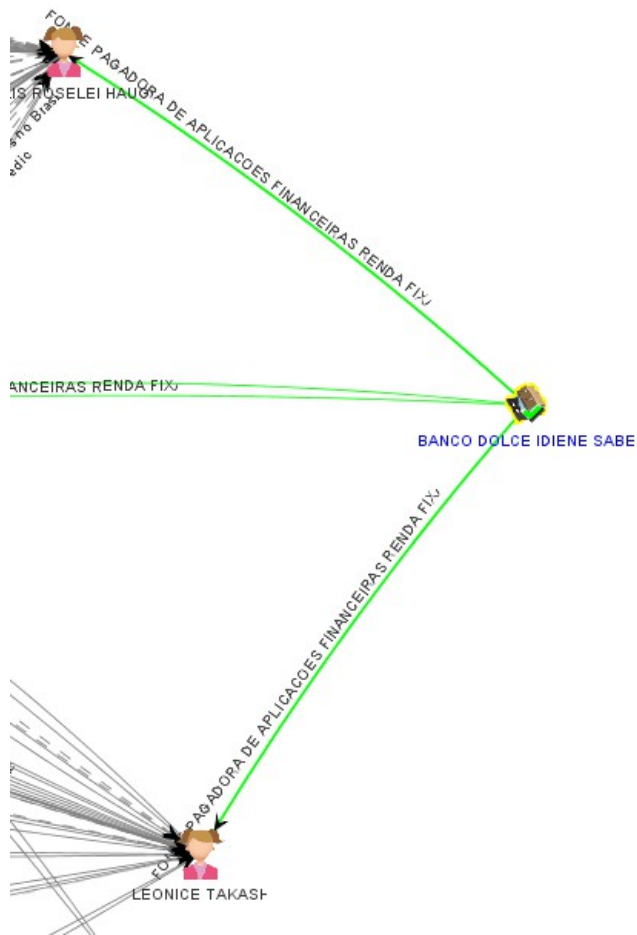


Figura 18: configuração da consulta por menores caminhos - Criação do Autor no ContÁgil



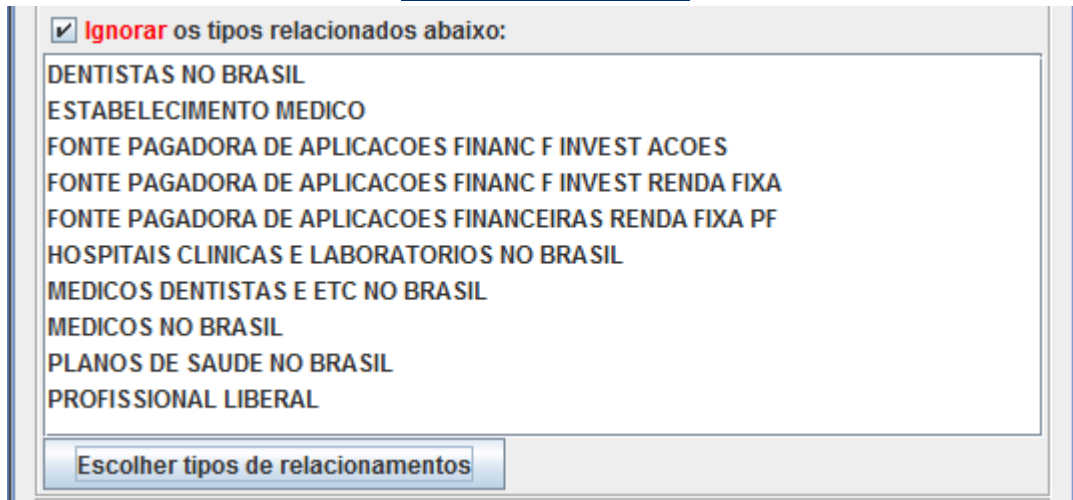
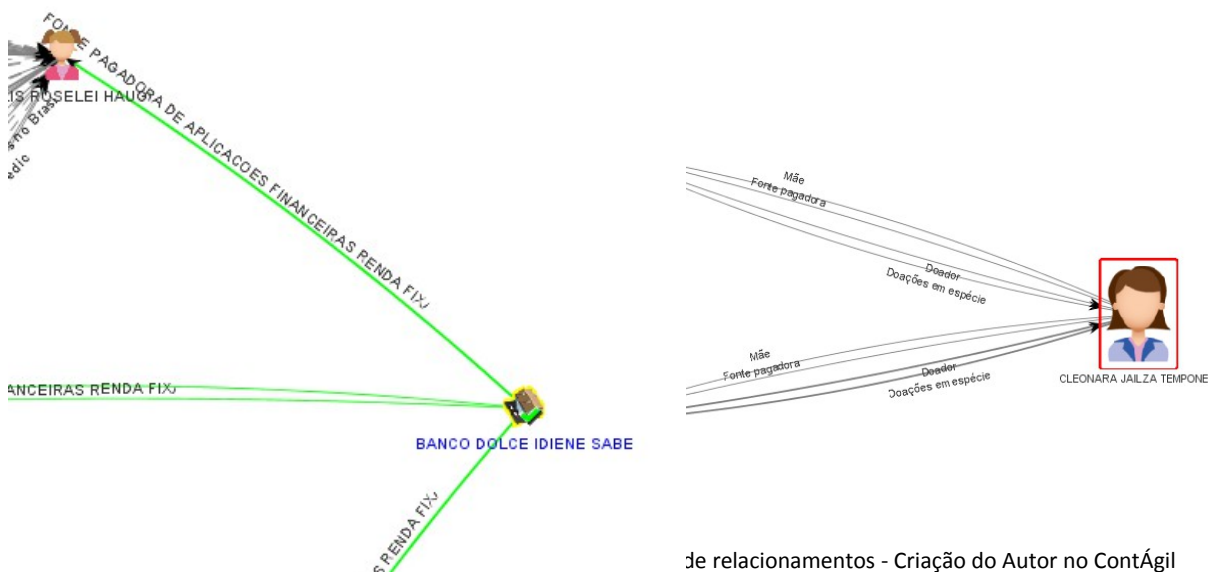


Figura 19: Confirmação dos relacionamentos a ignorar na consulta - Criação do Autor no ContÁgil

No resultado da consulta, temos que o menor caminho entre Névoa Tyson Ltda e Cleonara Jailza são as sócias excluídas do quadro societário, suas filhas, fechando o raciocínio de alaranjamento, uma vez que os atuais sócios (Hayrton e Rivas) são/foram empregados das ex-sócias (as propriedades dos relacionamentos mostram que ao se tornarem sócias (06/12/2007), o vínculo empregatício ainda



de relacionamentos - Criação do Autor no ContÁgil



EMBUSTE LTDA (77.084.629/0001-91) é Empregador de RIVAS DANYELLE CUCO (617.774.242-40)		
Propriedade	Valor	
CBO	04110	
DATA_FIM	02/2008	
DATA_INICIO	02/07/2007	
NÓ 1	EMBUSTE LTDA (7...	
NÓ 2	RIVAS DANYELLE ...	
QUANTIDA...	1,00	
TIPO_RELAC	Empregador	

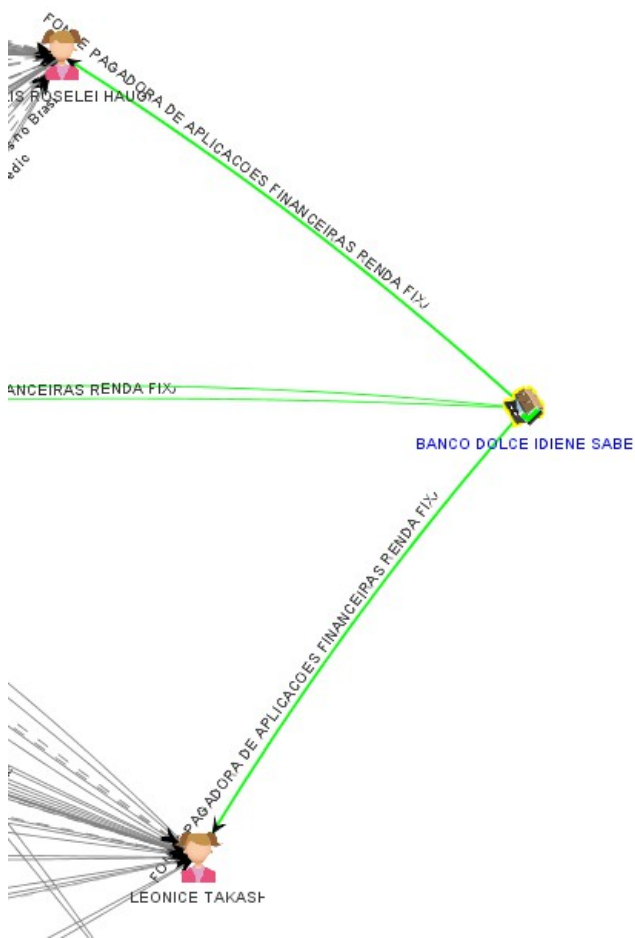
Figura 21: Propriedades do relacionamento de Rivas Daniele e Embuste Ltda - Criação do Autor no ContÁgil

O caso concreto teve como sujeito passivo do Auto de Infração o real beneficiário e os devedores solidários. A tese foi mantida nos julgados.

6.4 - Case 2 – Blindagem Patrimonial

Existe um jargão em investigação patrimonial e lavagem de dinheiro: “*follow*

por vezes, significa localizar ONDE e
depois, fazer o reverso até o alvo, firmando as
provas de que o dinheiro “ocultado” ou “lavado” e a pessoa





A lavagem de dinheiro, *le blanchiment d’argent, el blanqueo de capitales* e, meu predileto, “*il riciclagio di denaro sporco*”¹⁷ é procedimento extremamente nocivo à uma nação, pois permite, dentre outras coisas, a concorrência predatória e a interferência política, esteada pelo poder do dinheiro, formando um ciclo vicioso. Sobre isso, Marcelo Batlouni Mendroni leciona em sua obra “Crime de Lavagem de Dinheiro”:

“Da atuação com característica permanente das organizações criminosas na prática de lavagem de dinheiro decorre o **ciclo “criminal-legal”**, com a subdivisão do seu produto, que acaba sendo aplicado tanto para incrementar e/ou ampliar as atividades criminosas, como também em negócios já considerados “lícitos”. Depois, com a criação de empresas, através do poder advindo da formação de elevado capital, e das facilidades que dele decorrem, passa-se ao âmbito da atuação de “pressão política”, muitas vezes a partir da disseminação de (falsas ideologias, através das quais os grupos ou organizações criminosas se infiltram no sistema político do país, ora conseguindo eleger seus representantes para cargos dos poderes Executivos e Legislativo, ora corrompendo, ameaçando e extorquindo outros políticos. Finalmente, para degradação da sociedade, esses políticos aprovam leis que favorecem ou amenizam, de qualquer forma, aquela atividade criminosa primária, seja permitindo-a, através de brechas legais, seja diminuindo o poder investigativo e de atuação dos órgãos estatais de repressão” (2015, p.29) – *grifos originais*.

E após discorrer sobre as formas de organização das redes criminosas e de lavagem, assera:

“A principal forma de combate às organizações criminosas envolve ataca-las sem afrontá-las diretamente, no corpo a corpo, com os tradicionais métodos de utilização de prisões, embora estas ainda sejam muito necessárias. **E, primordialmente atacar seus bens, o seu dinheiro, o seu capital, que é por assim dizer, a sua corrente respiratória, o que as mantém vivas.**” (2015, p.34 – grifos nossos)

Para caracterizar a lavagem de dinheiro é preciso que a origem dele seja ilícita, e que se demonstre que houve a dissimulação, a ocultação, o disfarce, o encobrimento “com astúcia” de bens, direitos ou valor. O crime ofende a

¹⁷ A França e a Bélgica adotam a designação *blanchiment d’argent*, a Espanha, *blanqueo de capitales*, e Portugal, *branqueamento de capitais*....Os países de língua inglesa empregam a expressão *money laudring*, A Alemanha designa o fato típico de *Geldwasche*, a Argentina se refere a *lavado de dinero* e a Itália se vale do termo *riciclagio di denaro sporco* – “reciclagem de dinheiro sujo”.(Crime de Lavagem de Dinheiro, Marcelo Batlouni Mendroni, Ed. Atlas, 2015, P.22)



administração da justiça e a ordem socioeconômica, e independe de se ter provado ou não estar prescrito o crime antecedente (aquele que gerou o fluxo financeiro), característica que reforça a necessidade de uma investigação patrimonial contundente.

Embora tenhamos auxiliado em sem um sem número de operações cujo foco eram recursos ilícitos, achamos por bem não as mencionar neste trabalho, mesmo que ofuscadas, uma vez que os inquéritos e ações penais ainda perduram, envolvem pessoas expostas e os assuntos ainda fervilham na mídia. Neste case que escolhemos para ilustrar os tópicos, portanto, os recursos têm origem lícita. A empresa, “devedora contumaz e decana”, devia aos cofres públicos em meados de 2015 algo em torno de R\$ 500 milhões de reais. Com patrimônio irrisório e com sócios também desprovidos de lastro patrimonial aparente, as execuções judiciais resultavam em fracasso, pois não havia nada a arrestar. Empresa de nome forte no mercado em que atua; sócios presentes nas mídias sociais, mas, historicamente, sem qualquer intenção de saldar seus débitos. O Juízo, indignado, solicitou à Receita Federal que fizesse uma varredura no entorno da empresa foco, uma vez que as medidas judiciais vinham, de longa data, resultando infrutíferas (No teor do Relatório Fiscal: “O MM Juiz Federal determinou a expedição de ofício à Delegacia da Receita Federal, acerca da existência de patrimônio da executada e eventuais pessoas jurídicas e físicas a ela relacionadas, nos sistemas disponíveis da RFB”).

Partindo apenas do CNPJ da empresa devedora, codinome ofuscado “Marluce Kohls” (Passo-1: relacionar contribuintes de interesse) alimentou-se a base de relacionamentos paulatinamente, filtrando os *outliers*¹⁸, base a base, agregando

¹⁸ *Outliers* são os “pontos fora da curva”; valores muito altos ou muito baixos, pessoas que destoam do conjunto; por exemplo as grandes corporações e órgãos governamentais que “*ab initio*” não estejam relacionados aos fatos. Essa depuração deve ser feita de maneira extremamente metódica e cuidadosa, pois o descarte indevido pode eliminar justamente a entidade que deveria ser alvo. Contudo; em se falando de big



as informações coletadas (Passo-2: Importar dados para o Projeto) e retroalimentando com pesquisas seletivas. Finalmente montou-se a Base completa, agregando todas as informações entre si (Passo-3: Montar Base), inclusive GFIP e informações do Cadastro Nacional de Informações Sociais (Portal CNIS). Eis uma primeira consulta investigativa realizada (Passo-4: Consultar Base de Relacionamentos), quadro societário, em 4 níveis, buscando localizar empresas convenientemente “afastadas” do nó de origem, “Marluce kohls”:

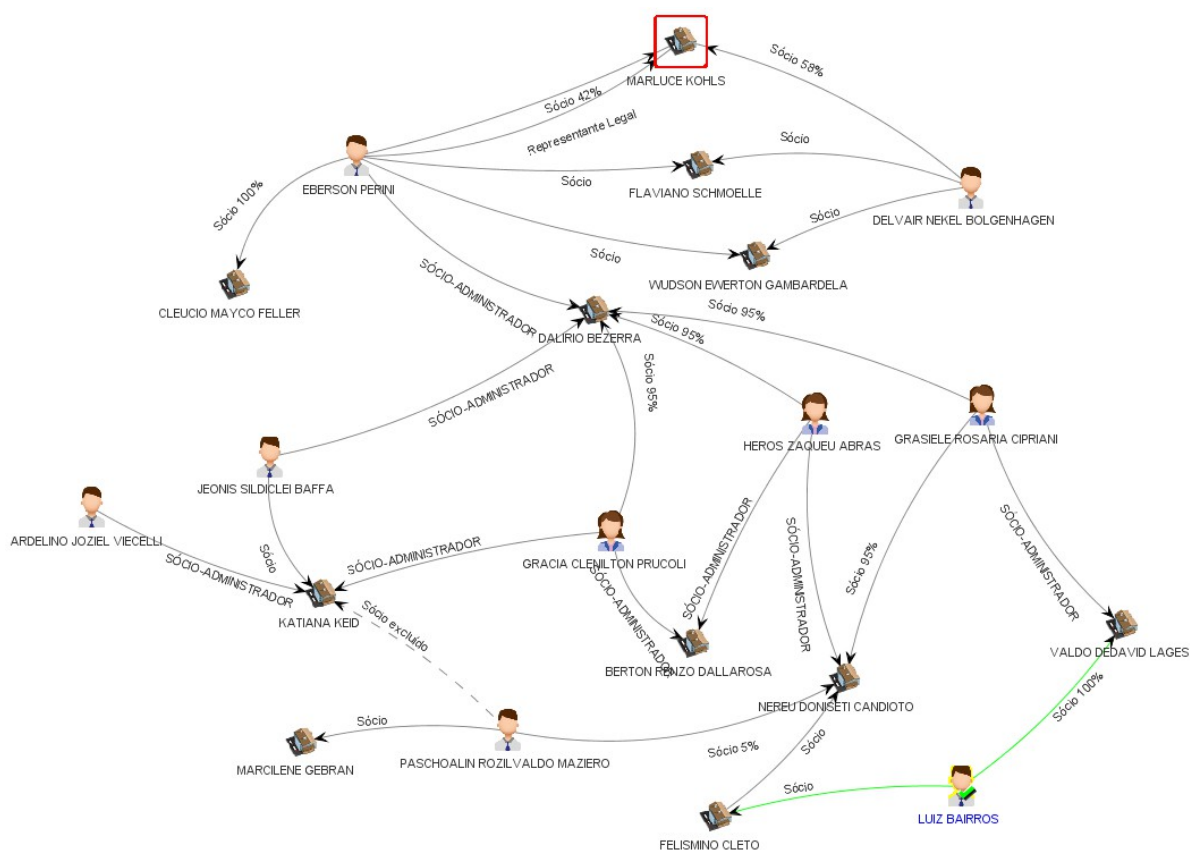


Figura 22: Quadro societário 4 níveis, foco em Marluce Kolhls – criação do autor, no ContÁgil

A consulta revela participações societárias cruzadas, e “empresas sócias das sócias das sócias da empresa alvo”; isso mesmo, das sócias das sócias; ou seja, empresas que estão no 3º, no 4º nível de separação do alvo, apenas considerando

data, tal descarte preliminar costuma não trazer maiores consequências, uma vez que a redundância se encarrega de reinserir a entidade descartada, caso ela seja realmente importante naquela rede.(nota do autor)



vínculos societários, depurando-se os demais relacionamentos. Ora; tal construção exige que se estude as “transações intercompanhias”, as “operações entre pessoas relacionadas” a fim de se averiguar sua normalidade.

Em nova consulta, considerando os relacionamentos com expressão monetária em primeiro nível, o grafo demonstra considerável fluxo financeiro partindo da executada, “Marluce Kohls” para, principalmente, duas de suas relacionadas: Apenas aqui fluíram perto de R\$ 12 milhões de reais:

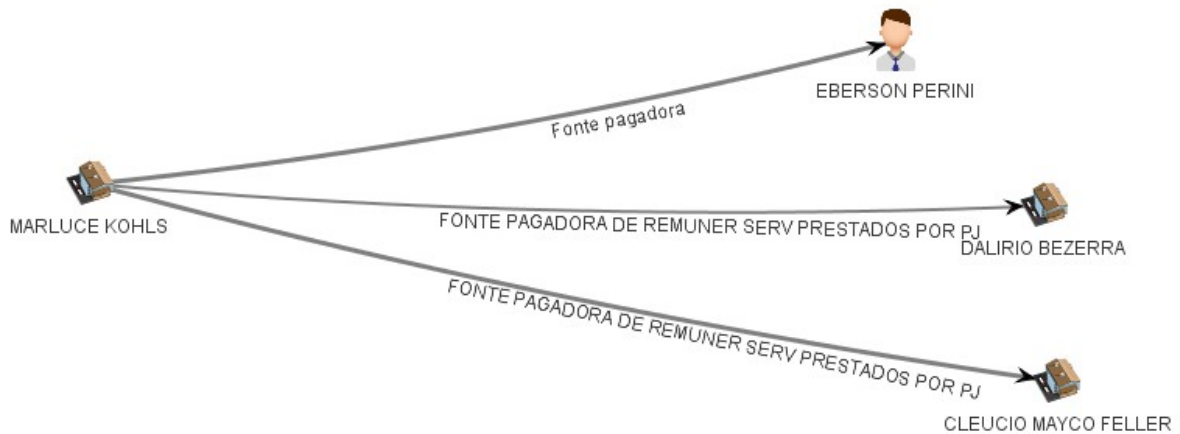


Figura 23: Fluxo financeiro primeiro nível, foco em Marluce Kolhls – criação do autor, no ContÁgil

Com os indícios apontados no grafo, abriu-se diligência fiscal, e analisando-se as contabilidades, alcançou-se fluxo financeiro em torno R\$ 40 milhões, no segundo nível. A exploração dos demais fluxos nos níveis mais afastados demonstraram a retirada de expressivos valores da empresa por operações comerciais simulatórias.

A decisão do juízo redirecionou a execução fiscal, alcançando os bens de todos os relacionados, com penhora eletrônica e indisponibilidades decretadas, bem como representações ao Ministério Público para a investigação dos crimes cometidos no esvaziamento patrimonial e blindagem. Parte das investigações ainda continuam.



7 - MODELOS MATEMÁTICOS REVELANDO COMPORTAMENTO

Um dos grandes avanços trazidos pelo estudo da ciência de redes se refere à capacidade de ranquear seus nódulos consoante a importância que adquiram na comunidade a que pertence.

Como exemplo dessa tecnologia, temos o algoritmo “*Page Rank*¹⁹”, criado por Larry Page, e que foi o embrião do sistema de busca mais famoso do mundo; o Google²⁰.

Esses algoritmos percorrem a estrutura de rede, localizam os *hubs* e nodos mais centrais, e calculam sua importância em relação aos demais, isso é intrínseco às redes, independe da vontade de seus componentes, e aqui reside o poder dos algoritmos que logram êxito em identificar o tal comportamento:

Uma rede é um conjunto de elementos chamados nódulos que são de algum modo ligados entre si por uma conexão; os aeroportos do mundo constituem uma rede, assim como a World Wide Web, ligações sociais e redes elétricas. Existe um ramo de pesquisa chamado “teoria das redes” que estuda a organização de tais redes e as ligações entre seus nódulos, com pesquisadores como Duncan Watts, Steven Strogatz, Albert-Laszlo Barabási e muitos outros. [...] Eles descobriram a seguinte propriedade das redes: existe uma concentração entre alguns nódulos que servem como conexões centrais. As redes possuem um tendência natural de se organizarem em torno de uma arquitetura altamente concentrada: alguns nódulos são extremamente concentrados e outros quase não em conexões. A distribuição dessas conexões possui uma estrutura escalável. Esse tipo de concentração não está

¹⁹ Esse algoritmo define a importância de um vértice com base na importância dos outros vértices que apontam para ele; de forma recursiva. A orientação de cada relacionamento é importante, pois a influência flui no sentido de cada relacionamento (nota do autor)

²⁰ Encontrei Larry Page, egresso de Stanford e cofundador do Google, em março de 2000, quando poucas pessoas tinham ouvido falar de sua ferramenta de busca. Larry Page fez uma breve preleção sobre sua ferramenta de busca e espalhou uma caixa de camisetas no meio da sala, que proclamavam a linha de assinatura do google: “Estou com sorte”.

O Google me intrigou porque violava a previsão básica do modelo sem escala, a de que quem faz o primeiro movimento leva vantagem. No modelo sem escala, os nós mais conectados são aqueles que aparecem primeiro. Eles dispõem de maior tempo para colecionar links e desenvolver hubs. Ferramentas de buscas populares como o Altavista ou Inktomi dominaram o mercado durante muito tempo antes do aparecimento do Google, tornando-se claramente um agente secundário. O Google, lançado em 1997, foi um retardatário da web. Em menos de três anos, contudo, o Google tornou-se igualmente o maior nó e a ferramenta de busca mais popular. (Albert-Lasló Barabási, Linked, Editora Leopardo, 2002, P. 85)




limitado à Internet; aparece na vida social (um pequeno número de pessoas está conectado à outras), em redes elétricas e de comunicações. (Nassim, Nicholas Taleb, A Lógica do Cisne Negro, 2008, ed. Best Seller, P. 286 – grifos nossos)

Ao longo da história, vários modelos matemáticos foram desenvolvidos para identificar esses hubs, essas aglomerações e classificar sua relevância. Matemáticos de renome como John Kleinberg se dedicaram a esse estudo. Temos no ContÁgil vários algoritmos; contudo, não os elencaremos aqui por óbvios motivos estratégicos.

A utilidade prática consiste em olhar para a babel de dados, para aqueles milhares de nós e relacionamentos e escolher quais devem ser objeto de estudo mais minucioso. Melhor dizendo, naquele “mundo pequeno” que se forma no entorno da pessoa (física ou jurídica) que motivou a construção daquela base de relacionamentos, identificar quais são as possíveis entidades “depositárias” dos bens que procuramos, ou “beneficiárias” dos valores que estão sendo amealhados pela empresa e colocados fora do alcance dos olhos do fisco, da mão do Estado.

Para ilustrar eis a aplicação, no caso em foco, do algoritmo de “Centralidade de Intermediação”²¹:



Centralidade de Intermediação - Algoritmo proposto por Freeman (1977) que considera quantas vezes cada vértice do grafo é utilizado como 'menor caminho' entre um par de outros vértices quaisquer do mesmo grafo. Expressa a importância que cada indivíduo exerce sobre pares na rede. Este algoritmo requer que o grafo possua no mínimo três vértices para poder ser calculado. O sentido de direção dos relacionamentos é irrelevante para esta análise. Pode ser utilizado com grafos conexos ou desconexos.

Figura 24: Breve descrição do algoritmo “Centralidade de Intermediação” – criação do autor, no ContÁgil

²¹ Freeman primeiro introduziu o conceito de intermediação parcial de um vértice numa dada rede para então chegar a um valor que pudesse medir a centralidade deste vértice. Isto significa que este valor deveria ser capaz de expressar a influência que o indivíduo (representado pelo vértice) poderia exercer sobre os seus pares numa rede de comunicação do mundo real. (Medida de Centralidade em Grafos, Leandro Quintanilha de Freitas, 2010, p.26)



Esse algoritmo, como dito na breve descrição do menu, pretende aferir a importância do nó perante seus pares, pela qualidade do relacionamento dos nós entre si e a distância entre eles.

O grafo classificado possui 3.827 entidades. Colocamos em destaque na tabela abaixo os nós constantes no grafo das figuras 22 e 23, donde foram depurados alguns *outliers*. Nota-se, que, dentre milhares de nodos, aqueles que merecem atenção ocupam as posições iniciais. E o mais interessante, não se trata de “escolha” das entidades que participam dessa comunidade; não é uma artificialidade; trata-se de comportamento característico do tipo de rede que se formou naturalmente, ao longo do tempo, conforme seu caminhar na aldeia global.

NOME	PONTUAÇÃO
MARLUCE KOHLS (22121194)	13.806.597,68
JORGEANO VOGT	946.694,26
DORILDO ROSINARIA SANGALO	508.107,17
EUSAMAR EDRIANO KHUL	495.951,93
ALENIS DARCIO ARANDT	304.846,56
JURANDIR EDUVIRGES MINES	203.514,69
DAVE BOTURA	194.426,87
LUIZ BAIROS	187.804,89
TASSO KHRISTIANE CHIQUETTE	174.789,13
KAIQUE REZA SAVAZZI	145.201,70
VITURINO RAULISON BUSO	124.828,93
HELVIA MARLOVE	99.908,39
DALIRIO BEZERRA	93.662,85
EBERSON PERINI	86.295,00
IRANI PARREIRAS	85.644,97
VIANY FRANCIEL SEHN	82.439,39
ESPERANDIO COELI	82.311,30
AVAIR LANCONI	76.664,62
CRITIANO STEFAN MONEQUE	73.769,85
CLEUCIO MAYCO FELLER	72.780,54
JEONIS SILDICLEI BAFFA	67.547,29
URSULA ALRIANA MINORU	64.636,11
GRACILIANO TARCILIO QUEROBI...	64.565,04
DALVANI ATANASIO TAMBAROTO	63.823,67
EMIR ELIONAI FILGUEIRA	61.479,73
WALDECIR JARDEL FUMAN	56.544,24

Fig. 25: Ranking de “Centralidade de Intermediação” – Pessoas já responsabilizadas - criação do autor, no ContÁgil



No caso ainda em comento, com quase 4 mil entidades, aquelas elencadas nos primeiros lugares e que não estão nos grafos retromencionados, também constam no processo de execução, foram citadas no Relatório Fiscal, e estão em uma outra fase da investigação, não abrangida neste trabalho:

NOME	PONTUAÇÃO
MARLUCE KOHLS (22121194)	13.806.597,68
JORGEANO VOGT	946.694,26
DORILDO ROSINARIA SANGALO	508.107,17
EUSAMAR EDRIANO KHUL	495.951,93
ALENIS DARCIO ARANDT	304.846,56
JURANDIR EDUVIRGES MINES	203.514,69
DAVE BOTURA	194.426,87
LUIZ BAIRROS	187.804,89
TASSO KHRISTIANE CHIQUETTE	174.789,13

Fig. 26: Ranking de “Centralidade de Intermediação” – Pessoas em investigação - criação do autor, no ContÁgil

Melhor dizendo; com exceção da entidade “Alenis Dácio Arandt”; que se revelou um *outlier*; **todos** os nomes apontados pelo algoritmo matemático estão envolvidos no caso real, sendo parte já responsabilizados, e parte em vias de o ser. Frise-se que se trata de apenas um dos vários algoritmos matemáticos que podem ser executados e que, consoante o tipo de organização que a rede assumiu, poderá apontar quais entidades deverão receber mais atenção na investigação patrimonial.

Temos, portanto, que essa tecnologia reafirma outra máxima do Big Data; investir na “correlação” (o que); abstrair da “causalidade” (por que).

8 - PRINCIPAIS PROBLEMAS ENCONTRADOS E RESPECTIVAS SOLUÇÕES

Em *analytics*, que é o termo que se usa para referir “análise em Big Data”; 85% do tempo e recursos são gastos na obtenção dos dados. Essa assertiva se



confirmou nos trabalhos executados; pois, uma vez as fontes de dados estão esparsas, é preciso reuni-los em um repositório que formará a base de relacionamentos sobre a qual rodaremos os grafos. Essa dificuldade é contornada com “agendamento” e “protocolo” de dados; de um dia para outro, diminuindo a possibilidade de sobrecarga na rede e nos sistemas institucionais.

Uma outra dificuldade se refere à capacidade de processamento das máquinas de uso cotidiano. Para lidar com grande massa de dados, uma das soluções (para nós a ideal) seria aplicar “*paralelismo*” utilizando, por exemplo a tecnologia “*map-reduce*”²² e o uso de um “*cluster escalável*”, que se pode resumir como vários computadores trabalhando como se fosse um só, com a possibilidade de se adicionar tantos mais quantos forem necessários; tal qual se faz no moderno “*Hadoop*”²³, utilizado pelo *Facebook*, *VISA* e outras grandes organizações. Essa dificuldade, enquanto o plano institucional de *Data Mining* não decola, deve ser contornada localmente; ora configurando máquinas mais potentes, ora fazendo *upload* dos dados para um Datacenter Institucional em máquina remota virtual com maior capacidade de processamento e de memória. Aqui estamos falando em centenas de milhares de nós, como os que se atinge nos trabalhos de grandes empresas e corporações.

Uma terceira e não menos incômoda dificuldade é a necessidade de aperfeiçoamento de alguns convênios com órgãos externos, inclusive de outros

²² O padrão map-reduce é uma forma de organizar o processamento de maneira a aproveitar as múltiplas máquinas de um cluster. Ao mesmo tempo, mantém-se o quanto for possível do processamento e dos dados de que ele precisa na mesma máquina. (SADALAGE, J, FOWLER, MARTIN, 2013, P.107)

²³ O map-reduce é um padrão que pode ser implementado em qualquer linguagem de programação. Entretanto, as restrições do estilo tornam-no apropriado para linguagens projetadas especificamente para computações map-reduce. Apache Pig [Pig], um ramo do projeto Hadoop [Hadoop], é uma linguagem criada especificamente para facilitar a escrita de programação map-reduce. De um modo semelhante, se você quiser especificar programas map-reduce utilizando uma sintaxe semelhante a SQL, existe a Hive [HIVE], outra ramificação do Hadoop. (SADALAGE, J, FOWLER, MARTIN, 2013, P.117)



poderes, para aumentar a gama de cruzamentos possíveis, o que diminuiria, inclusive, muitos processamentos intermediários que hoje fazemos por não termos acesso direto à informação. Essa dificuldade foi reportada às instâncias hierárquicas, que lutam para vencê-la em definitivo.

Finalmente, há a necessidade de se mudar o paradigma de análise de dados; não o engessando em linhas e colunas; partindo do geral, da visão ampla, para a específica, filtrando os relacionamentos e aplicando consultas, avaliando comportamentos, não olhando apenas para a entidade, mas para seu entorno.

9- CONCLUSÃO

Na RFB temos “*nosso próprio Big Data*”. Ao utilizarmos a avalanche de dados cotidiana com a tecnologia de Grafos de Relacionamento nos permitimos aproveitar ao máximo cada byte de informação vertida aos nossos repositórios.

A ferramenta “análises patrimoniais”, o Script de Análise e Busca do Beneficiário Real, que gerou o carinhoso acrônimo “sabre” remetendo à tecnologia dos filmes de ficção científica é uma realidade palpável. Basta acessar o ContÁgil, ir no menu “compartilhamentos” e baixar o Script. A depender dos acessos que o usuário possui nos sistemas institucionais, ele já inicia, de imediato, a formação da base, a montagem e as consultas, chegando em poucos minutos a conclusões que, doutra maneira levaria horas, dias e até meses de investigação, a depender da complexidade da matéria investigada.

A visão gráfica supera, em muito, a lida com tabelas, linhas e colunas; facilita a compreensão dos relacionamentos e aclara as oitivas, as reportagens, as



diligências, as denúncias e delações, pois permite que se identifique visualmente as situações retratadas.

Criada para ser manuseada pelo usuário de nível básico de conhecimento no ContÁgil; a ferramenta pode ser usada ANTES, no preparo, DURANTE, na auditoria e DEPOIS de constituído o Crédito Tributário, na cobrança e efetivação de garantias, tendo por grande virtude auxiliar a identificar os grupos econômicos, as interpostas pessoas, o planejamento e blindagem, e apontar a existência de bens e fluxos financeiros.

O desenvolvimento da ferramenta, feito ao custo de “horas de auditoria” aplicadas, não gerou quaisquer ônus orçamentário ou financeiro. Seu aperfeiçoamento se deu na prática, trabalhando situações cotidianas e que traziam grande desconforto na análise e investigação; tornando o trabalho extremamente prazeroso, e até, divertido. O uso do “sabre” resulta na valorização profissional, na motivação, em especial do Auditor Fiscal (Investigação, Prospecção, Auditoria, Autuação, Informação, Representação e interação com órgãos externos) e do Analista Tributário (Arrecadação e instrumentação da Garantia do Crédito).

Finalmente, podemos apontar como ganhos não mensuráveis o fortalecimento da Instituição Receita Federal do Brasil perante as demais instituições como Poder Judiciário, Ministério Público Federal, Tribunais de Contas, Polícia Federal e Procuradoria da Fazenda. Uma apresentação com Grafos de Relacionamento impressiona e convence os espectadores e um relatório fiscal ou peça jurídica construídos com grafos é contundente; sendo a maior prova disso as sentenças obtidas nas medidas cautelares e demais medidas judiciais, deferidas no aspecto patrimonial, civil ou na repercussão penal.



10 – BIBLIOGRAFIA

BARABÁSI, Albert-László (2009). **LINKED - [CONECTADO] - a nova ciência dos networks**. São Paulo/SP: Editora Leopardo/Campus.

MAYER-SHONBERGUER, Viktor; CUKIER, Kenneth. (2013). **BIG DATA - Como extrair Volume, Variedade, Velocidade e Valor da Avalanche de Informação Cotidiana**. Rio de Janeiro/RJ: Editora Elsevier.

SEGARAN, Toby (2008). **Programando a Inteligência Coletiva**. Rio de Janeiro/RJ: Alta Books/O'Reilly

WATTS, Duncan J.(2003). **Seis Graus de Separação (Six Degrees) - A evolução da Ciência de Redes em uma era conectada**. São Paulo: Editora Leopardo

SADALAGE J Pramod; FOWLER Martin (2013). **NoSQL Essencial - Um Guia Conciso para o Mundo emergente da Persistência Poliglota**. São Paulo: Novatec Editora

MENDRONI, Marcelo Batlouni (2015). **Crime de Lavagem de Dinheiro**. São Paulo, Editora Atlas

TALEB, Nicholas Nassim (2008). **A lógica do cisne negro**. Rio de Janeiro, Editora Best Seller

NETTO, Oswaldo Paulo Boaventura; JURKIEWICZ, Samuel (2013). **Grafos – Introdução e Prática**. São Paulo: Editora Blucher

PANZARINO, Onofrio (2014). **Learning Cypher - Write powerful and efficient queries for neo4j with Cypher, its official language**. Birmingham - U.K.: Packt Publishing

BRUGGEN, Rik Van (2014). **Learning Neo4J - e-book**

QUINTANILHA, Leandro de Freitas (2010). **Medida de Centralidade em Grafos**

<http://www.tecmundo.com.br/programacao/1185-o-que-e-script-.htm>

<http://www.administradores.com.br/noticias/negocios/nota-fiscal-eletronica-no-brasil-e-pelo-mundo/27442/>

<http://origemdapalavra.com.br/site/palavras/persona/>

<http://idg.receita.fazenda.gov.br/dados/resultados/fiscalizacao/arquivos-e-imagens/plano-anual-fiscalizacao-2016-e-resultados-2015.pdf>



Apresentação do SPED, site institucional <http://sped.rfb.gov.br>