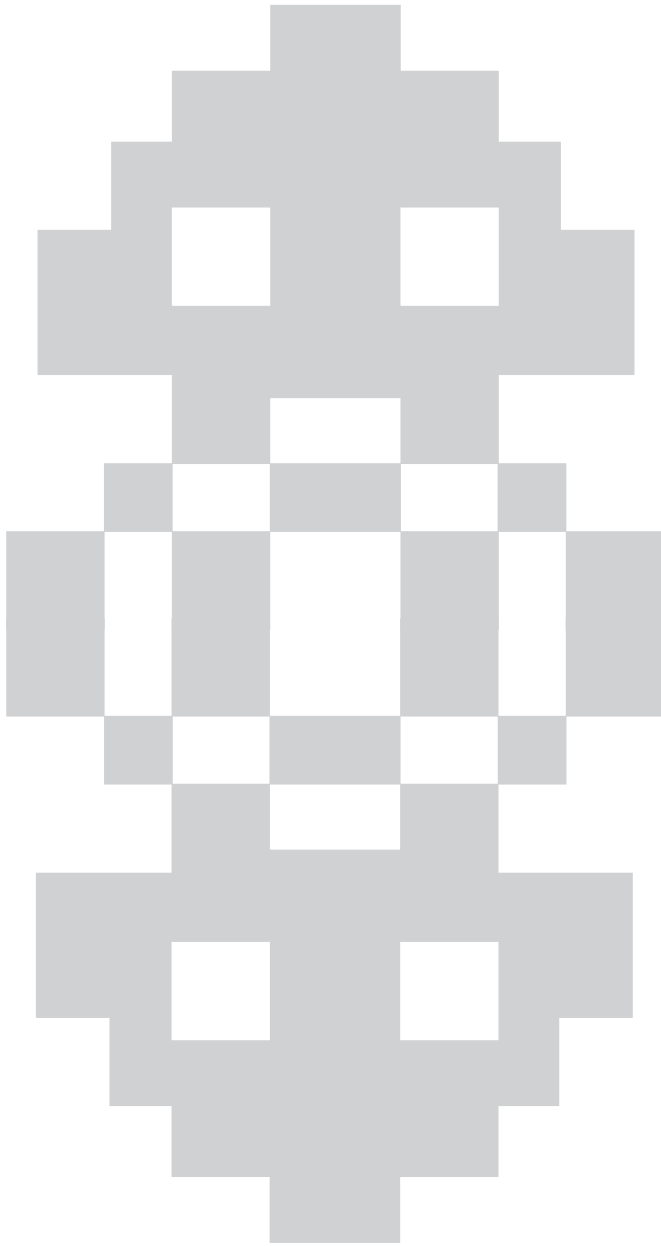


A Mineração de Dados como apoio ao Controle Externo

Remis Balaniuk



Remis Balaniuk é servidor do Tribunal de Contas da União. Graduado em Ciência da Computação pela Universidade de Brasília (UnB), mestre em Ciência da Computação pela Universidade Federal do Rio Grande do Sul (UFRGS), doutor em Informática pelo *Institut National Polytechnique de Grenoble* (INPG- França) e pós-doutor em Computação pela *Stanford University* (EUA) e pelo *Institut National pour la Recherche en Informatique et Automatique* (INRIA - França).

INTRODUÇÃO

Um aspecto importante do contexto atual da administração federal é o uso extensivo de sistemas informatizados estruturadores, como o Sistema Integrado de Administração Financeira do Governo Federal (SIAFI), o Sistema de Administração de Serviços Gerais (SIASG) e o Sistema Integrado de Administração de Recursos Humanos (SIAPE), entre outros. Esses sistemas registram e controlam o funcionamento cotidiano da máquina administrativa de forma centralizada. Suas bases de dados, que já contêm dados históricos de mais de uma década em alguns casos, constituem fontes inestimáveis de informação. O acesso típico a esses dados é feito por meio dos próprios sistemas que os mantêm. Nesses sistemas são propostas consultas pontuais e relatórios, mas análises mais elaboradas e abrangentes não são oferecidas.

Diversas iniciativas pontuais vêm sendo feitas no sentido de explorar essas grandes bases de dados do governo, mas primordialmente buscando dar mais transparência às contas do governo, como, por exemplo, o Portal Transparência da Controladoria-Geral da União, CGU¹, que reúne informações sobre o uso do dinheiro público pelo Governo Federal, o Siga Brasil do Senado Federal² com informações sobre as leis e o Contas Abertas³, que busca disponibilizar para a sociedade as execuções orçamentária, financeira e contábil da União, dos Estados, do Distrito Federal e dos Municípios. Algumas iniciativas voltadas ao uso corporativo têm buscado facilitar o acesso e o uso dos dados de sistemas estruturadores para fins de controle, como, por exemplo, o *datawarehouse** (DW) Síntese do TCU com dados dos sistemas SIAFI e SIASG, e os DW-SIAPE e Compras do Ministério do Planejamento, Orçamento e Gestão (MP).

Entretanto, a existência e a disponibilidade de grandes volumes de dados de interesse para o controle não garantem que se consiga extrair desses dados os indícios e provas relativas aos atos ilícitos

cometidos contra o erário. Uma vez os dados de interesse estando acessíveis, o problema central que se impõe aos analistas e auditores é como lidar com a complexidade, diversidade e o gigantesco volume das informações ali contidas. Somente alguns poucos especialistas, profundos conhecedores das regras de negócio de um segmento da máquina administrativa e simultaneamente profundos conhecedores da estrutura da base de dados que registrou os atos analisados, são capazes de propor e realizar consultas analíticas que evidenciem uma hipótese específica sendo analisada.

Para viabilizar o uso sistemático dessas bases de dados pelos órgãos de controle é imprescindível que se comece pela definição de uma metodologia que possa estruturar o processo de análise de grandes volumes de dados levando em consideração as necessidades específicas da área de controle.

Em Tecnologia da Informação (TI), esse esforço de sistematização da análise de grandes bases de dados pertence à área conhecida como “Extração de Conhecimento” (também conhecida como processo KDD, do inglês *Knowledge-Discovery in Databases*). O KDD lida com a extração de informações de bases de dados, descobrindo relacionamentos de interesse que não são observados diretamente pelo especialista no assunto, bem como auxiliando na validação de conhecimento extraído. Um passo importante do KDD é a chamado “Mineração de Dados” (*Data Mining*), que é o processo de explorar grandes quantidades de dados à procura de padrões consistentes, como regras de associação ou sequências temporais, buscando detectar relacionamentos entre variáveis. Uma sistemática bem definida de extração de conhecimento requer um processo de trabalho claro e bem documentado, que possa ser repetido e refinado diversas vezes. Em muitos casos, os resultados obtidos numa primeira iteração de um processo de extração de conhecimento não valem o esforço despendido. Entretanto, iterações sucessivas levam ao conhecimento aprofundado dos negócios e bases analisadas e podem se transformar numa ferramenta poderosa de análise.

Os órgãos de controle têm uma missão continuada de fiscalizar a administração pública. Uma auditoria ou fiscalização em um órgão específico não é um esforço isolado, mas um evento de uma série de

atos de controle lidando com o mesmo negócio e as mesmas bases de dados. O uso de metodologias que permitam uma evolução no tempo nessas séries de atos de controle, reutilizando e refinando o conhecimento adquirido em atos anteriores, podem levar a ações muito mais eficazes. Essas mesmas metodologias permitiriam também o acompanhamento direto das decisões e orientações desses órgãos de controle, dirigidas a essas unidades jurisdicionadas, através da simples re-análise das mesmas fontes de dados depois de decorrido um intervalo de tempo. Utilizando os mesmos critérios de extração de dados, as mesmas análises e as mesmas ferramentas utilizadas na iteração de mineração que embasou decisões e orientações a serem acompanhadas, seria possível acompanhar as mudanças ocorridas e refletidas nos dados.

OPORTUNIDADES

O TCU detém um vasto conhecimento de negócio, na sua maioria tácito e desestruturado, advindo primordialmente da experiência prática de seus analistas e auditores. Esse conhecimento de negócio é mal distribuído e pouco compartilhado dentro da instituição. A adoção de uma metodologia adequada de extração de conhecimento em nível institucional induz um processo de formalização e compartilhamento do conhecimento de negócio, uma vez que esse conhecimento é a matéria-prima da elaboração das análises feitas via mineração de dados. Da mesma forma, a mineração permite a validação na prática desse conhecimento ao confrontá-lo com fatos refletidos nas bases de dados, levando ao aprimoramento do conhecimento institucional.

Existe carência por conhecimento útil ao controle em todos os níveis: operacional, tático e estratégico. Uma metodologia de extração de conhecimento permite a extração e disponibilização de informações em diferentes níveis de agregação, indo do grão mais baixo relativo a fatos e achados úteis à execução das auditorias, passando por totalizações e estatísticas úteis ao nível tático, até a análise de risco e relevância importante ao nível estratégico.

Essa visão em níveis de agregação vem ao encontro da visão que o TCU vem buscando implantar no planejamento de suas ações de controle. Desde 2005, através dos “Temas de Maior Significância”

(TMS), o TCU tem buscado priorizar ações de controle em temas de grande interesse da sociedade e em problemas em que o TCU sabe que precisa ampliar a fiscalização. Planos de fiscalização têm sido preparados por temas e funções de governo, buscando com isso definir uma agenda anual e ter uma visão sistêmica das áreas a serem auditadas. Sem uma metodologia adequada, que permita estruturar o conhecimento institucional e os dados disponíveis sobre os temas e funções de governo analisadas, a escolha e a abordagem dos TMS acabam sendo feitas de forma empírica e aproximada, baseadas em opiniões e conhecimento subjetivo.

Da mesma forma, iniciativas de atuação integrada e de compartilhamento de informações entre as unidades do TCU, como a Rede Interna de Informações para Suporte ao Controle Externo (RI), instituída pela portaria TCU N° 205, de 1º de junho 2009, requerem um suporte metodológico para serem implantadas. Uma das atribuições centrais da RI, que é a de “analisar as informações disponibilizadas, considerando risco, materialidade, relevância e oportunidade, e submeter ao titular da unidade propostas de ações de controle”, precisa ser realizada de forma sistêmica e integrada na instituição, e para isso é necessária uma base metodológica adequada e única.

RISCOS

A existência de dados úteis para o controle, distribuídos em diversas instâncias da administração pública, e a prerrogativa legal de acesso a esses dados atribuído ao TCU, são de grande importância. Entretanto, existem diversos obstáculos que muitas vezes impedem esse acesso. Falta de termos de cooperação, dificuldades técnicas de acesso, qualidade incerta dos dados e falta de integração entre diferentes bases de dados são só alguns desses obstáculos.

O conhecimento extraído através da mineração de dados não deverá nunca substituir o trabalho de campo das auditorias e o processo de análise e julgamento tradicional. Esse conhecimento obtido serve somente como insumo para apoiar e agilizar os processos de trabalho institucionais e a eles devem estar integrados. Modelos, sistemáticas e dados, por melhor que sejam definidos e

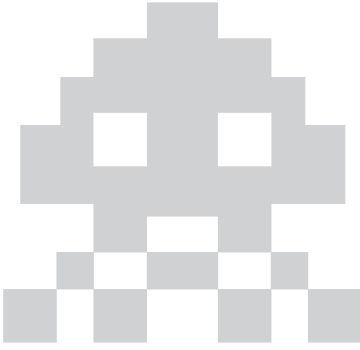
implementados, nunca estarão livres de erros e inconsistências. Diversas são as fontes possíveis para essas deficiências: regras de negócio mal interpretadas ou omitidas, dados inconsistentes, análises mal construídas. Só com a integração do processo analítico à auditoria tradicional é possível a confrontação de resultados analíticos com a realidade, permitindo o refinamento dos modelos. Modelos maduros atingem alto grau de confiabilidade, mas a própria realidade é dinâmica, e por isso o processo de refinamento dos modelos e confirmação por meio de auditorias sempre será necessário.

A mineração não busca encontrar todos os casos e situações suspeitas. A busca pelo completo muitas vezes inviabiliza o projeto. A cada esforço de mineração deve haver um compromisso entre o ideal e o possível. Casos levantados devem ser consistentes e baseados no maior número possível de evidências, mesmo que um número ainda maior de casos não seja detectado ou tenha que ser deixado de lado devido à incompletude ou inconsistência dos dados disponíveis. Dessa forma, as inúmeras dificuldades ligadas à má qualidade dos dados, baixa integração, dificuldade de acesso, não devem ser consideradas como empecilho para a mineração, mas sim como dificultadores que devem ser tratados e evitados na medida do possível e de forma progressiva.

MINERAÇÃO DE DADOS ADAPTADA AO CONTEXTO DO TCU

A metodologia mais aceita em mineração de dados é conhecida por CRISP-DM, sigla do inglês *Cross-Industry Standard Process for Data Mining*⁴. O uso da metodologia CRISP nos permite atacar o problema de extração do conhecimento de forma progressiva e organizada, partindo de uma análise de alto nível, que busca a compreensão do negócio, e indo em direção à definição e implantação de modelos que permitam efetivamente atingir os objetivos da mineração.

Essa metodologia é constituída de seis etapas: Compreensão do negócio, Compreensão dos dados, Preparação dos dados, Modelagem, Avaliação e Aplicação. Cada etapa é subdividida em tarefas definindo assim um processo de trabalho iterativo, ou seja, realizado em ciclos. Dessa forma, um



projeto de mineração vai repetir o ciclo de etapas quantas vezes forem necessárias até que os objetivos da mineração tenham sido atingidos.

Nesse artigo é proposta uma adaptação dessa metodologia ao contexto institucional e de negócio do TCU. Nessa adaptação é proposto um macro processo de trabalho que permitirá ao TCU extrair conhecimentos úteis ao controle diretamente de grandes bases de dados governamentais e integrar esse conhecimento ao trabalho de suas equipes técnicas. A metodologia aqui proposta supõe a existência de uma base institucional para o funcionamento do macro processo de trabalho. Os pilares dessa base são:

- Existência de um ambiente para compartilhamento de conhecimento de negócio, onde especialistas possam registrar seu conhecimento adquirido:
 - * A gestão do conhecimento institucional é um dos temas mais discutidos na atualidade pelas corporações. Diversas práticas e ferramentas permitem esse compartilhamento. Entre as mais populares estão as enciclopédias *wiki*, que permitem a construção colaborativa e descentralizada de bases de conhecimento (ver BRAUN para uma introdução ao assunto⁵). A Enciclopédia *Wiki* do Controle Externo⁶, recentemente implantada no TCU, será o espaço principal para compartilhamento do conhecimento de negócio sobre o qual todo o processo de análise será baseado
- Existência de um corpo de análise, com sólida base técnica, ampla visão da administração pública e experiência de negócio no controle:
 - * No TCU, esse corpo de análise já existe, distribuído pelas unidades técnicas. Unidades como a Adplan, Sefti e Secob trabalham rotineiramente a análise de dados. Os representantes da RI nas unidades também vem sendo preparados nesse sentido. É essencial que esse corpo de análise trabalhe de forma uniforme e integrada para que possa compartilhar conhecimento de negócio e métodos.
 - * O apoio tecnológico a esse corpo de análise é essencial. Nesse sentido, a participação da SETEC é primordial para o bom funcionamento da metodologia.
- Ambiente para internalização de grandes bases de dados provenientes de sistemas estruturadores da administração pública:
 - * A internalização de grandes bases de dados já vem sendo feita pelo TCU há algum tempo, a exemplo do DW Síntese que internaliza dados dos sistemas SIAFI e SIASG. Entretanto, para fins de mineração de dados, essa internalização precisa ser mais abrangente e mais ágil.
- Institucionalização do trabalho de mineração, definindo em que instâncias e em que processos de trabalho o conhecimento será obtido, utilizado e validado.
 - * A utilização na atividade finalística do TCU do conhecimento adquirido é a razão em si de todo o esforço aqui proposto.

O macro processo de trabalho proposto pode ser ilustrado pelo esquema da Figura 1.

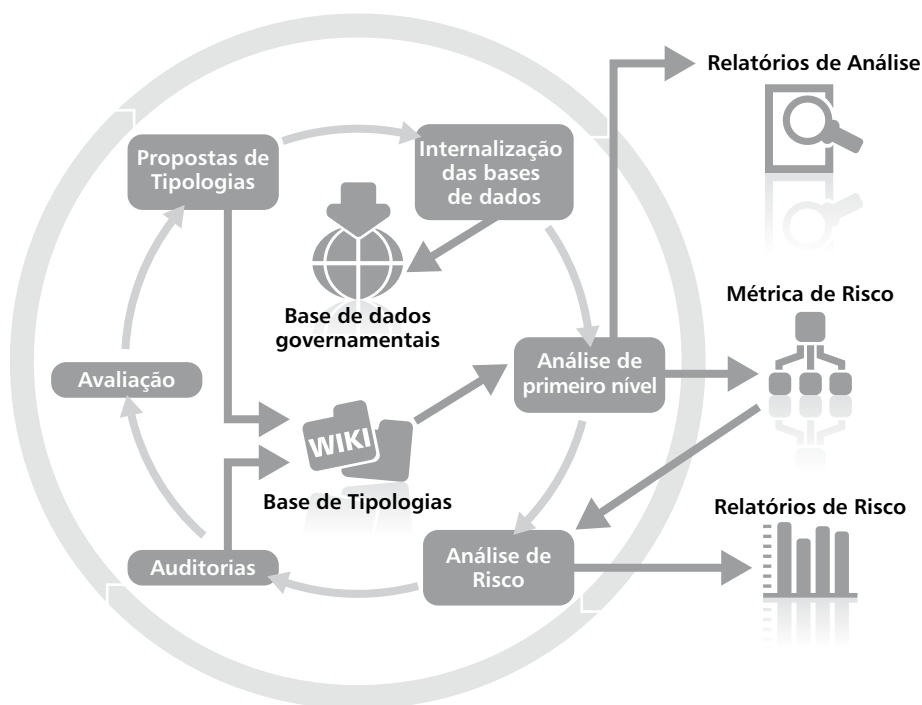
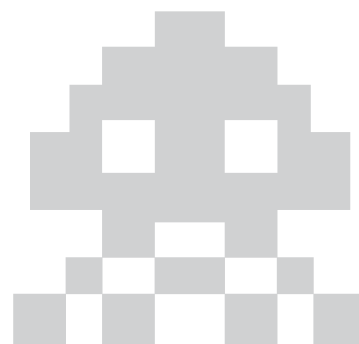


Figura 1 – Macro processo de mineração de dados

A etapa de Proposta de Tipologias, que corresponde à Compreensão do Negócio da metodologia CRISP, supõe a participação do corpo técnico da casa que, com sua experiência no controle, elabora a descrição de situações ou esquemas associados a atos ilícitos e indica estratégias para sua detecção. Como ilustração, uma tipologia conhecida nos indica que: “Servidor público pode direcionar contratos do órgão ou entidade em que trabalha para empresa de sua propriedade.” Uma estratégia para a detecção de atos associados a essa tipologia seria buscar contratos celebrados pela Administração Pública Federal com pessoas jurídicas que têm como sócio um servidor público do órgão contratante.

A etapa de Internalização das Bases de Dados, que corresponde à Preparação dos Dados na metodologia CRISP, busca disponibilizar os dados necessários à análise das tipologias propostas. É importante enfatizar que a mineração ocorre sobre todo um conjunto de dados correspondente à integralidade do domínio analisado. Assim, no caso do exemplo de tipologia citado acima, a mineração ocorrerá sobre todo o domínio de contratos da administração pública federal, todos os servidores públicos e todas as pessoas jurídicas e seus sócios simultaneamente. Dessa forma, é preciso ter acesso à integralidade das bases de dados correspondentes para poder realizar essa análise, o que não pode ser feito acessando remotamente sistemas de informação ou mesmo tendo acesso a bases analíticas tipo *datawarehouse* (a menos que um mesmo DW contenha todas as bases necessárias). No nosso exemplo, precisaríamos ter disponíveis as bases do SIASG, que registram os contratos da administração pública federal, da RAIS, que registra informações sobre contratos de trabalho, inclusive os de servidores



públicos, e do CNPJ, que registra informações sobre pessoas jurídicas e seus sócios. Até a alguns anos, a internalização de grandes bases de dados seria considerada um grande desafio tecnológico, dadas as dimensões envolvidas. Somente computadores de grande porte, conhecidos como servidores, eram capazes de armazenar e manipular essas bases. Entretanto, com a evolução da tecnologia, mesmo computadores pessoais já atingiram o poder computacional requerido. Como exemplo, a base analítica (OLAP) do DW Compras, contendo detalhes de quase três milhões de compras, ocupa 100GB (cem *gigabytes*), sendo que computadores pessoais são vendidos atualmente com 500GB (quinhentos *gigabytes*) de espaço em disco.

A etapa seguinte, Análise de Primeiro Nível, corresponde à etapa de Compreensão dos Dados do CRISP. O objetivo dessa etapa é extrair dados que evidenciem as tipologias. Essa extração, feita utilizando cruzamento de bases de dados, nos permite obter os Relatórios de Análise contendo detalhes de casos que coincidem com os critérios estabelecidos pela tipologia. No caso da ilustração, seria feita uma seleção na RAIS dos servidores públicos e seus respectivos órgãos empregadores, em seguida um cruzamento dessa seleção com os sócios de empresas na base CNPJ a fim de selecionar as empresas que têm como sócio um servidor público e finalmente um cruzamento com os dados do SIASG, a fim de gerar um relatório de análise com os casos de compras realizadas pelo órgão com o qual o servidor-sócio possui vínculo empregatício junto à empresa de sua propriedade. Esse cruzamento nos permite ir além e obter detalhes desses casos, por exemplo, as notas de empenho emitidas. Esses relatórios de análise são bastante úteis, pois indicam situações concretas onde há risco de estar ocorrendo atos ilícitos, e servem de base para iniciar uma averiguação. Obviamente esses casos levantados não constituem irregularidades em si, e os dados utilizados não podem ser tratados como provas.

Com a multiplicação das tipologias levantadas, e o número expressivo de casos que cada uma traz, rapidamente um gestor se vê diante de uma quantidade de informações que tem dificuldade em triar e priorizar para dela se servir como insumo ao planejamento das ações de controle.

Em função disso, o macro processo de trabalho propõe a etapa de Análise de Risco e Relevância, que corresponde à etapa de Modelagem da metodologia CRISP. O objetivo dessa etapa é agregar as informações levantadas pela Análise de Primeiro Nível, permitindo uma análise ampla do domínio considerado. Diversos podem ser os objetivos dessa análise:

- Ordenar um domínio considerado segundo critérios de relevância ou risco;
- Encontrar padrões de comportamento;
- Identificar agrupamentos com características comuns;
- Identificar instâncias com comportamento atípico.

Talvez o primeiro desses objetivos seja o de maior impacto no contexto do TCU, pois permite ordenar instâncias a serem controladas (unidades jurisdicionadas, funções de governo, programas, contratos, etc.) segundo critérios que indiquem relevância ou risco de encontrar irregularidades.

Novamente o conhecimento de negócio é usado como base para construir modelos de análise. Na análise de primeiro nível, os “modelos” consistiam no simples cruzamento entre bases de dados acompanhado de filtros. Na análise de risco e relevância, os modelos analíticos são mais elaborados e baseados em métodos matemáticos. Para cada objetivo é necessária a escolha de um método adequado. Para cada método, um conjunto de parâmetros precisa ser especificado na definição de um modelo de análise. O trabalho de modelagem consiste em transformar dados e conhecimento de negócio desestruturado em parâmetros objetivos a partir dos quais um modelo é construído. Na metodologia aqui proposta, a análise de primeiro nível permite calcular um conjunto de métricas de risco e relevância a partir das tipologias, que serão em seguida utilizadas nos modelos de análise de risco e relevância.

Seguindo nossa ilustração, consideremos que nosso objetivo final seja planejar auditorias e que para isso buscamos ordenar a lista de todas as

unidades jurisdicionadas (UJs) segundo critérios de relevância e risco, para com isso apoiar esse planejamento. Para materializar o que seria risco e relevância associados a uma UJ, devemos escolher um conjunto de métricas que permitam medir aspectos relevantes ligados às UJs. Na nossa ilustração, o fato de estarem ocorrendo transações comerciais entre uma UJ e empresas pertencentes a servidores do seu quadro é um indicativo de risco de irregularidade na forma de direcionamento nas compras. Métricas derivadas dessa tipologia seriam, por exemplo, o número de compras e o total gasto por uma UJ junto a empresas ligadas a servidores do seu quadro. Quanto maior o valor calculado dessas métricas, maior o risco. Inúmeras outras métricas podem ser concebidas indicando risco, como o número e total de compras feitas sem licitação pela UJ justificadas com fulcro no artigo 24 da lei 8666, incisos I ou II, mas que excedem os limites legais; ou número e total de compras com empresas inaptas ou inidôneas.

Além dessas métricas de risco, uma série de outros aspectos ligados, por exemplo, ao impacto social, temporalidade ou auditabilidade de uma UJ, precisam ser considerados na priorização de ações de auditoria.

A esse conjunto de aspectos não associados ao risco demos a designação de “relevância”. Um exemplo de métrica de relevância seria, por exemplo, escolher um TMS, identificar quais programas ou funções de governo estão mais diretamente ligados a ele e totalizar, por UJ, o montante e o número de aquisições relacionadas a esses programas ou funções. Quanto maior os valores desses indicadores, maior a relevância. Note-se que o requisito para definição de uma métrica é poder calcular seu valor para todas as instâncias do domínio analisado (no nosso caso as UJs) a partir de dados existentes, e poder interpretar o seu resultado com relação ao que se quer medir no modelo (risco ou relevância).

Um método adequado precisa ser escolhido, que integre as diversas métricas obtidas. Nas pesquisas até aqui realizadas, o método mais adequado ao tipo de análise aqui proposto é conhecido por Classificador Bayesiano⁷. Os detalhes do método proposto e do trabalho teórico de adequação já realizado são específicos demais para os objetivos desse artigo. Destacamos somente o fato de esse método permitir

a combinação de métricas em diferentes escalas de valor sem a necessidade de procedimentos *ad hoc* de ajuste, tais como ponderações. Além do mais, esse método não requer uma base de exemplos para funcionar, baseando seus resultados somente nos valores das métricas e nas regras de negócios.

Embora essa etapa pareça bastante técnica e envolta em formalismos matemáticos de difícil compreensão, é preciso esclarecer que, metodologicamente, a modelagem propriamente dita se restringe à seleção e especificação das métricas, cujo esforço é focado em aspectos de negócio e de manipulação de dados, e que o processamento do modelo se faz por meio de ferramenta automatizada, de forma transparente para o analista. Além do mais, os modelos de análise de risco e significância a serem concebidos são poucos e seu uso e atualização se farão periodicamente quando da chegada de novos dados ou métricas, exigindo pouco retrabalho a cada novo processamento. É preciso enfatizar que, apesar da ilustração utilizada, o método proposto permite a ordenação de qualquer tipo de domínio, desde que sejam definidas métricas adequadas a ele. Seria possível ordenar um conjunto de contratos, de empresas, de servidores, de programas de governo, etc.

A etapa das Auditorias corresponde à Aplicação do CRISP. Considerando que o macro processo aqui proposto seja integrado à sistemática de planejamento de auditorias do TCU, os relatórios gerados pela Análise de Risco e Relevância, seriam levados em consideração na seleção das UJs a serem fiscalizadas. Cada fiscalização seria apoiada por relatórios gerados na análise de primeiro nível, indicando quais fatos devem ser auditados.

A etapa de Avaliação analisa os resultados obtidos nas fiscalizações, validando ou não os resultados das análises. Os acertos reforçam as tipologias, regras de negócio e modelos, enquanto os erros permitem refiná-los. Erros precisam ser detalhados para que se encontre sua fonte permitindo as correções necessárias. Acertos precisam ser registrados para reforçar as tipologias e facilitar análises subsequentes. A *wiki* do Controle Externo será novamente o espaço mais adequado para o registro e incorporação dessas validações no conhecimento de negócio utilizado nas análises. O conhecimento

refinado de tipologias e regras de negócio permitirá a definição de modelos mais precisos.

RESULTADOS

A sistemática de análise aqui apresentada começa a ser utilizada pelo TCU no contexto da RI. Mais de setenta tipologias já foram propostas e quinze delas já foram analisadas. Um conjunto inicial de relatórios de análise foi obtido e repassado para os membros da RI que estão dando encaminhamento através de auditorias.

A *wiki* do Controle Externo é uma realidade e pode vir a ser um espaço de grande importância para o compartilhamento do conhecimento institucional. Incentivamos fortemente todo o corpo técnico da casa a utilizar e contribuir.

Modelos de risco para analisar UJs, conluio entre órgãos e fornecedores e cartéis estão em construção e começam a trazer seus primeiros resultados.

CONCLUSÃO

Para os órgãos de controle em geral, e para o TCU em particular, é incontornável que se use adequadamente dados em formato eletrônico gerados pelos sistemas estruturadores da administração pública federal. Os inúmeros desafios associados ao uso eficiente desse volume avassalador de dados requerem a adoção de metodologia adequada e viável de análise. É preciso que essa metodologia, de abrangência institucional, seja capaz de lidar com os aspectos de negócio inerentes a essa análise integrando-os às soluções tecnológicas necessárias para que uma sistemática completa seja posta em prática.

Esse artigo apresentou uma metodologia com esse fim adaptada à realidade do TCU, buscando demonstrar sua viabilidade. Obviamente uma série de aspectos não foram aqui tratados, ou o foram de forma superficial, requerendo uma discussão mais aprofundada.

Para que esse novo paradigma se torne realidade, é preciso que o TCU se familiarize com diversos conceitos aqui apresentados, como a formalização e o compartilhamento de seu conhecimento

de negócio. É preciso também que a instituição encontre formas de integrar seu corpo técnico, disperso em unidades cada vez mais especializadas ou distantes geograficamente, em torno de uma sistemática comum de trabalho. É essencial também que a atividade finalística do controle externo saiba incorporar de forma mais natural perfis e papéis ligados à TI. É inegável na sistemática proposta o peso da TI, seja nas tarefas de internalização e preparação das bases de dados, seja nas etapas de análise. Embora essas etapas não exijam equipes especificamente de TI, não se pode negar que o conhecimento de TI precisa existir dentro dessas equipes, e não só como apoio externo. Equipes multidisciplinares de análise, compostas por perfis de negócio e TI, poderão obter melhores resultados em mais curto espaço de tempo.

REFERÊNCIAS

- BRASIL. Controladoria Geral da União. *Portal da transparência do Governo Federal*. Disponível em: <<http://www.portaltransparencia.gov.br/>>. Acesso em: 01 set. 2010.
- BRASIL. Congresso Nacional. Senado Federal. *SigaBrasil*. Disponível em: <http://www9.senado.gov.br/portal/page/portal/orcamento_senado/SigaBrasil>. Acesso em: 01 set. 2010.
- CONTAS abertas. Disponível em: <<http://contasabertas.uol.com.br/WebSite/>>. Acesso em: 01 set. 2010.
- CROSS Industry Standard Process for Data Mining. Disponível em: <<http://www.crisp-dm.org/Overview/index.htm>>. Acesso em: 01 set. 2010.
- BRAUN, Daniela. *Saiba como usar ferramentas de wiki para organizar suas ideias na web*. 08 de junho de 2007. IDG Now! Disponível em: <<http://idgnow.uol.com.br/internet/2007/06/08/idgnoticia.2007-06-06.4067436489/>>. Acesso em: 01 set. 2010.
- BRASIL. Tribunal de Contas da União. *Enciclopédia Wiki do controle externo*. Disponível em: <<https://contas.tcu.gov.br/wikice>>. Acesso em: 01 set. 2010. [Acesso exclusivo para público interno].
- CLASIFICADOR bayesiano ingenuo, 2009. Disponível em: <http://es.wikipedia.org/wiki/Clasificador_bayesiano>. Acesso em: 01 set. 2010.

NOTAS

- * Tecnologia utilizada para armazenamento e consulta de grandes bases de dados.