

# Estratégias de Fiscalização em um Contexto de Grandes Bases de Dados: Experiência da Secex Previdência



**Fábio Henrique Granja e Barros**

é servidor do Tribunal de Contas da União. Secretário da Secex Previdência.



**Jorge Mendes de Oliveira Castro Neto**

é servidor do Tribunal de Contas da União e professor do Instituto de Psicologia da Universidade de Brasília (UnB). Diretor de Previdência da Secex Previdência.



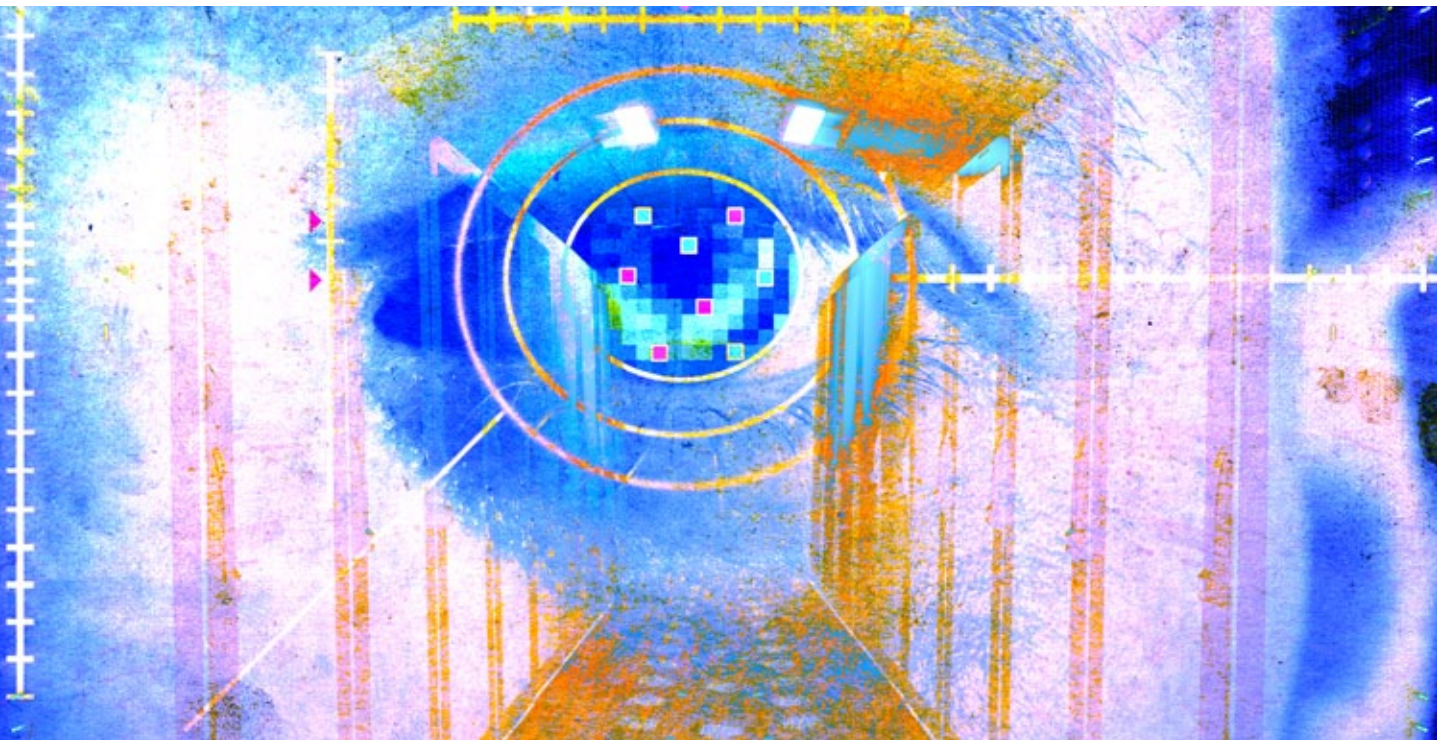
**Melchior Sawaya Neto**

é servidor do Tribunal de Contas da União. Diretor de Assistência Social da Secex Previdência



**Teonio Wellington Martins**

é servidor do Tribunal de Contas da União. Diretor da Secex Previdência, na diretoria responsável pela fiscalização do Ministério do Trabalho e Emprego, das entidades do Sistema "S" e de alguns Conselhos de Fiscalização de Profissionais.



## RESUMO

A natureza e o volume dos gastos, bem como a quantidade de dados disponíveis frente às funções de governo relativas a trabalho, a assistência social e a previdência social, impuseram à Secex Previdência a necessidade de empregar análise de dados para aumentar a efetividade de suas fiscalizações e avaliações. Para tanto, realizaram-se cruzamentos de bases públicas e análises quantitativas de eficiência, de governança e do impacto dos gastos da previdência no crescimento inclusivo. Com vistas a aperfeiçoar essas análises atualmente empregadas, a Secretaria implementou, em conjunto com uma empresa de software, uma bancada analítica para testar uma metodologia que permitisse integrar diversas análises quantitativas, com vistas a gerar melhores diagnósticos, prescrições e predições para os problemas identificados. Nesse sentido, foi possível testar as soluções de TI da empresa na integração dos resultados das diversas técnicas de análise de dados para dois benefícios operacionalizados pelo INSS, os benefícios por idade e para pessoas com deficiências.

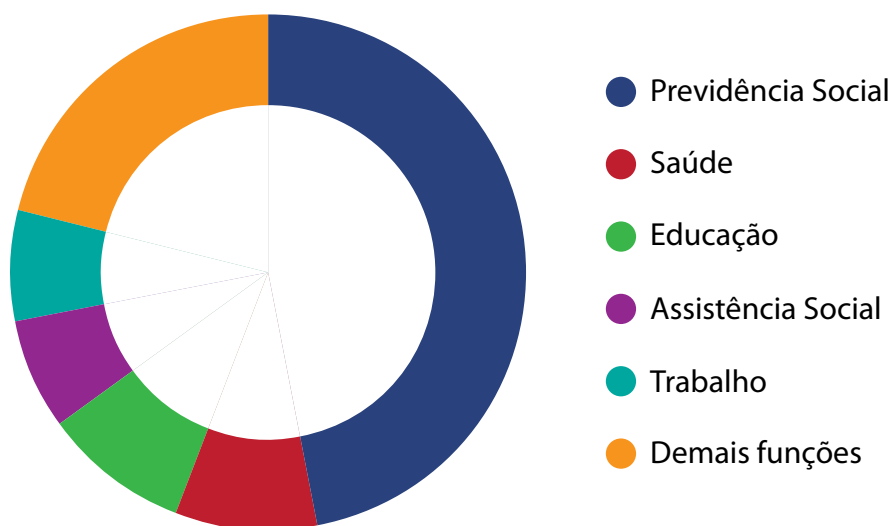
**Palavras chave:** Análise de dados, métodos quantitativos, TI, Previdência, Assistência, Trabalho

A SecexPrevidência atua sobre as áreas da Previdência Social, da Assistência Social e do Trabalho.

Juntas essas áreas representam, aproximadamente, 61% do Orçamento Geral da União (Gráfico abaixo), excluindo nesse cálculo os gastos com os encargos da dívida pública. As aposentadorias e pensões, o Seguro Desemprego, o Abono Salarial, o programa Bolsa Família, o Benefício de Prestação Continuada, os convênios e as transferências fundo a fundo estão entre os principais gastos acompanhados pela Secretaria.

Devido ao significativo volume de informações que a SecexPrevidência acompanha (mais de 50 milhões de beneficiários, 9000 Centros de Referência da Assistência Social, 1500 Agências da Previdência, 1200 Agências do Sine, bases de dados com grande volume de registros (MACIÇA, CNIS, RAIS, SISOBI, Folha PBE, CadÚnico, CPF, CAGED, SIAPE, CensoSUAS))<sup>1</sup> fez-se necessário o desenvolvimento de várias estratégias empregando métodos quantitativos e análises de dados.

Entre as experiências da secretaria, cita-se a Estratégia de Fiscalização de Controle dos Benefícios Previdenciários que abrangeu a análise dos controles internos e a implementação de três auditorias nas bases de dados nos benefícios de maior materialidade (aposentadoria rural, pensão por morte, aposentadoria por idade e aposentadoria por tempo de contribuição). Essas auditorias proporcionaram além da revisão dos benefícios com indícios de irregularidade, o fortalecimento dos sistemas de concessão, haja vista que os *scripts* dos cruzamentos das bases de dados foram transferidos para o INSS, permitindo que o órgão passasse a aplicar controles mais eficientes nos mais

**Gráfico 1:**Despesa empenhada  
- por função (2013)

Fonte: Siga Brasil

de 30 milhões de benefícios ativos, reduzindo o potencial de erros e fraudes. A cessação dos benefícios concedidos equivale a uma economia potencial estimada em, aproximadamente, R\$ 140 milhões de reais anuais (Acórdão 456/2010 – P, Acórdão 715/2012 – P, Acórdão 666/2013 – P).

Similarmente a esta estratégia, na Diretoria de Trabalho, foram realizadas, desde 2012, auditorias de dados no Seguro Desemprego nas modalidades trabalhador formal e pescador artesanal. O Seguro Desemprego em 2013 representou um gasto aproximado de R\$ 32 bilhões e atendeu 5 milhões de benefícios ano, sendo que as duas modalidades fiscalizadas representam mais de 90% do total despendido. Nessas auditorias, realizaram-se cruzamentos das informações do cadastro administrativo do Seguro Desemprego com as seguintes bases de dados públicas: CPF, RAIS, CAGED, CNIS, SIAPE, MACIÇA, SISOBI, TSE candidatos e Bolsa Família). Identificaram-se milhares de parcelas de Seguro Desemprego pagas irregularmente, tendo sido determinado a recuperação dos recursos e o aperfeiçoamento dos sistemas de controle, estimando-se um benefício de controle potencial de, aproximadamente, R\$ 150 milhões de reais (Acórdão 2.089/2013 – Plenário).

Realizaram-se, também, análises de eficiência nas agências do Sine, da Assistência Social e da Previdência Social. Para tanto, buscou-se uma abordagem quantitativa na mensuração de eficiência, a Análise Envolvória de Dados (DEA – *Data Envelopment Analysis*). A técnica permite que se quantifique o grau de eficiência relativa entre os postos, apontando as margens de melhora de produtividade e as unidades benchmarking.

A DEA tem uso disseminado em vários setores econômicos, tais como: energia elétrica, telefonia, por-

tos, rodovias, escolas, hospitais, entre outros. A técnica permite a estimação da eficiência de unidades administrativas, ao construir uma fronteira de produção que contenha as unidades eficientes, com as quais são comparadas as unidades ineficientes. A comparação de eficiência, efetuada pela técnica, é sempre relativa, ou seja, as unidades ineficientes o são, porque há outra unidade de perfil assemelhado de produção que é mais eficiente. A relatividade apresentada pela técnica, além de demonstrar a factibilidade de aumentar a eficiência, também deixa em aberto a possibilidade de aprimoramento de eficiência para as unidades mais eficientes formadoras da fronteira de produção.

Em relação aos resultados, das 1207 agências do Sine, das 1500 agências da previdência social e dos 6893 Centros de Referência da Assistência Social, somente 19, 23 e 406 se mostraram na fronteira da eficiência, respectivamente. Os resultados dessas análises de eficiência foram muito promissores, pois se evidenciou, em todas as três funções de governo (previdência social, assistência social e trabalho), a possibilidade de incrementar a quantidade de atendimentos, isto é, o serviço aos cidadãos, sem a necessidade de aumentar custos.

Outra análise quantitativa realizada, empregando a Pesquisa Nacional por Amostra de Domicílios (PNAD) do Instituto Brasileiro de Geografia e Estatística, foi a avaliação do impacto da previdência no crescimento inclusivo (Acórdão nº 1274/2013 - Plenário). As análises realizadas indicaram que o subsistema rural do Regime Geral da Previdência Social (RGPS) desempenha significativa função de inclusão social, na medida em que: 1) gera aumentos de renda proporcionalmen-

te maiores para domicílios com menores rendas per capita; e 2) destina proporcionalmente mais recursos, relativamente ao PIB, para municípios com menores PIB per capita. Esse relatório foi apresentado nas Contas de Governo de 2012.

Também aconteceu, empregando o CENSOSuas, uma técnica quantitativa denominada “análise de componentes principais” para definir um índice de governança para os Conselhos Municipais de Assistência Social. Essa ação contou com a participação de uma especialista em sociometria da Universidade de Brasília (UnB) e permitiu que houvesse uma indicação da capacidade desses conselhos em operar e fiscalizar os recursos transferidos fundo a fundo. Espera-se que esse resultado sirva para definir futuras fiscalizações e para avaliar com mais precisão o controle social que vem sendo realizado por essas entidades.

Apesar das metodologias de batimento de dados e dos métodos quantitativos aplicados atualmente pela SecexPrevidência implicarem resultados importantes, entendeu-se que seria necessário desenvolver uma metodologia que permitisse integrar diversas análises, com vistas a gerar melhores diagnósticos, prescrições, bem como predições. Para tanto, realizou-se com uma empresa de mercado de software, no período de fevereiro a junho de 2014, uma bancada analítica com o objetivo de testar as soluções de TI da empresa na integração e automatização dos resultados das diversas técnicas de análise de dados para dois benefícios operacionalizados pelo INSS, por idade e para pessoas com deficiências.

Assim, selecionaram-se as seguintes espécies de benefícios: aposentadoria por idade e benefício de prestação continuada para pessoas com deficiência. De forma geral, o trabalho conjunto com a empresa de software consistiu na limpeza e integração das bases de dados utilizadas, na identificação de tipologias que caracterizariam as irregularidades, na utilização de diversas técnicas de mineração de dados que ressaltariam padrões de fraude e padrões de anormalidade e, por fim, na integração desses resultados e na elaboração de uma métrica indicando a probabilidade do benefício ser irregular.

Em relação à limpeza e integração das bases, realizou-se um Data Quality que permitiu a identificação, com probabilidade superior a 95%, do número de CPF de 99,2% de todos os beneficiários identificados na Maciça de fevereiro de 2014. Foram utilizadas técnicas de identificação fonética, de geração de match-code e de-duplicação de registros, resultando na

identificação do número do CPF do beneficiário, por meio da análise dos dados contidos na MACIÇA e no CPF da Receita Federal do Brasil (RFB), considerando o conteúdo dos campos número do CPF, nome, data de nascimento, nome da mãe, número do título de eleitor, local de residência.

Foram utilizadas diversas técnicas de análise de dados, dentre as quais ressaltam-se as seguintes:

- 1. BublePlot:** a técnica foi utilizada para analisar a quantidade de Guias de Recolhimento do FGTS e de Informações à Previdência Social (GFIPs) apresentadas por um empregador em cada mês. O padrão é que um empregador, em cada mês, apresente apenas uma declaração de vínculos empregatícios por meio do sistema GFIP. Eventualmente, para correção de algum erro ou omissão, o empregador poderá apresentar uma segunda declaração. O objetivo foi destacar os casos que fossem considerados anormais, que fugissem ao padrão, com número grande de declarações corrigidas.
- 2. Regressão Logística:** a técnica foi utilizada para a detecção dos casos que são considerados anormais, que fogem ao padrão de concessão dos benefícios (no caso, foram analisadas somente duas espécies de benefícios, que eram escopo da bancada analítica).
- 3. Redes de relacionamentos:** a técnica foi utilizada para analisar os relacionamentos existentes entre os dados de benefícios comprovadamente irregulares e os dados de benefícios que foram pagos no mês de fevereiro de 2014. Os relacionamentos analisados referem-se a benefícios que possuem o mesmo titular, instituidor, recebedor ou procurador. Ou seja, todos os benefícios que estão sendo pagos e que possuem relacionamento dessa natureza com benefícios que são comprovadamente irregulares foram selecionados como suspeitos ou que exigem uma análise mais detalhada.

Por fim, a consolidação dos resultados de todas as análises foi realizada por meio da utilização de solução de TI específica da empresa e permitiu definir uma pontuação para cada benefício, ou seja, utilizou-se uma espécie de matriz *scorecard*. A pontuação ou score foi representada numericamente pelo seu valor e tam-

bém por cores. Os valores mais altos (máximo de 1000) indicam alta probabilidade do benefício ser irregular.

Entende-se que essa metodologia permitirá à SecexPrevidência atuar de forma seletiva nos benefícios com maior probabilidade de erro/fraude, na vulnerabilidade dos sistemas transacionais, na redução de falsos positivos, no monitoramento das deliberações, na realização de diagnósticos mais precisos, na identificação de oportunidades que possam contribuir para a entrega de melhores serviços aos cidadãos. A metodologia também possibilitará a automatização de alguns procedimentos de cruzamento de dados, que poderão ser realizados periodicamente com baixo custo operacional. Em suma, a incorporação desse tipo de metodologia contribuirá para o aperfeiçoamento das técnicas de detecção de fraude, de avaliação de riscos e de disseminação da transparência atualmente empregados na secretaria.

Diante da materialidade e complexidade dos regimes previdenciários, dos benefícios do Seguro Desemprego e Abono Salarial do Ministério do Trabalho, e das ações de transferência de renda do Ministério de Desenvolvimento Social e Combate à Fome, reforça-se a necessidade de se incrementar as ações que envolvam o uso e análise de grandes bases de dados.

Portanto, os desafios enfrentados pela Secex-Previdência nesta estratégia de atuação envolvendo análise de dados de forma integrada requerem uma mudança não apenas nos processos de trabalho da própria unidade, mas também na articulação com outras áreas do Tribunal, em especial a área responsável pela capacitação e pela tecnologia da informação (TI).

Em relação ao apoio da área de TI, faz-se necessária uma nova forma de integração entre as equipes (intra e inter secretarias do tribunal), evoluindo para a utilização maciça de TI na integração das bases, no incremento da qualidade e no desenvolvimento de análises consistentes e úteis para as áreas finalísticas do Tribunal. Isso requer capacitação, integração entre equipes de diferentes áreas e software adequado para mineração de dados, para cruzamento de informações.

Para que o cenário de atuação desejado seja possível, é necessária a conjugação de diversas ações. A primeira e basilar é a capacitação dos servidores que já atuam na secretaria em áreas de conhecimento dos novos desafios. Além disso, propõe-se que um grupo de auditores dentro da secretaria fique responsável pela interação com outras áreas, em especial a de TI, e pelo suporte de informações que se fizerem necessárias às ações de controle da unidade. Esse grupo visaria

contribuir, por meio de análises e batimentos de bases de dados, para a maior efetividade e eficiência dos recursos já alocados na Secretaria na medida em que direcionaria as ações para benefícios e sistemas transacionais com maior risco. A ação melhoraria a qualidade e o resultado dos trabalhos, pois foi identificado, em diagnóstico realizado com apoio da Seplan em 2014, que há um significativo e repetitivo investimento de tempo com análises de dados que poderiam fazer parte de uma rotina estruturada prévia, fornecendo aos auditores bases de informações mais precisas e constantes.

Contudo, com base em diagnóstico realizado, identificou-se que a implementação de uma estratégia na SecexPrevidência que permita a utilização de dados de forma abrangente poderá gerar aumento da eficiência da atuação. Para tanto, haverá necessidade de uma mudança nos processos de trabalho internos, na alocação de pessoal, no apoio da área de TI e na capacitação de parte do corpo técnico.

O presente artigo, além de visar a apresentação dessa sistemática de análise integrada em um contexto de grandes bases de dados que inicia seu desenvolvimento na SecexPrevidência, objetiva chamar a atenção para a oportunidade de estratégias semelhantes serem desenvolvidas em outras áreas do Tribunal, como, por exemplo, Saúde, Educação, Pessoal, Obras e Licitações, haja vista que também trabalham com amplas bases de dados.

## NOTA

- 1 MACIÇA - folha de pagamento dos benefícios operacionalizados pelo INSS; CNIS – Cadastro Nacional de Informações Sociais – cadastro dos segurados do INSS; CadÚnico – Cadastro Único de Informações Sociais – cadastro de beneficiários da assistência social do Ministério do Desenvolvimento Social (MDS); CAGED – Cadastro Geral de empregados e Desempregados; Folha PBF – Folha de Pagamento do Programa Bolsa Família – de responsabilidade do MDS; RAIS – Relação Anual de Informações Sociais dos anos de 2009 a 2012; SIAPE – Sistema Integrado de Administração de Recursos Humanos; CPF – Cadastro de Pessoas Físicas da Receita Federal do Brasil; SISOBI – Sistema de Controle de Óbitos do Ministério da Previdência Social, CENSOSuas – Censo do Sistema Único de Assistência Social.