



Entrevista | Página 4

Luis Felipe Monteiro

Desafios para a transformação digital no setor público brasileiro



TCU

Opinião | Página 9

Transformação Digital: Pessoas, estratégia e tecnologia gerando mais impacto para a sociedade

Rodrigo Felisdório

Artigo | Página 11

A auditoria financeira de agências federais nos EUA tem relação custo-efetividade positiva e as lições aprendidas são aplicáveis ao Brasil

Tiago Alves de Gouveia Lins Dutra

Destaque | Página 79

Alteração da Súmula 230 e outras decisões do TCU



REPÚBLICA FEDERATIVA DO BRASIL
TRIBUNAL DE CONTAS DA UNIÃO

MINISTROS

José Mucio Monteiro (Presidente)
Ana Arraes (Vice-Presidente)
Walton Alencar Rodrigues
Benjamin Zymler
Augusto Nardes
Aroldo Cedraz
Raimundo Carreiro
Bruno Dantas
Vital do Rêgo

MINISTROS-SUBSTITUTOS

Augusto Sherman Cavalcanti
Marcos Bemquerer Costa
André Luis de Carvalho
Weder de Oliveira

MINISTÉRIO PÚBLICO JUNTO AO TCU

Cristina Machado da Costa e Silva (Procuradora-Geral)
Paulo Soares Bugarin (Subprocurador-Geral)
Lucas Rocha Furtado (Subprocurador-Geral)
Marinus Eduardo de Vries Marsico (Procurador)
Júlio Marcelo de Oliveira (Procurador)
Sérgio Ricardo Costa Caribé (Procurador)
Rodrigo Medeiros de Lima (Procurador)

REVISTA TCU

145

Revista do Tribunal de Contas da União | Brasil | Ano 51

Jan-Jun | 2020



© Copyright 2020, Tribunal de Contas da União do Brasil

Os conceitos e opiniões expressas em obras doutrinárias assinadas são de inteira responsabilidade dos autores.

A reprodução completa ou parcial desta publicação é permitida, sem alterar seu conteúdo, desde que citada a fonte e sem fins comerciais.

www.tcu.gov.br

Missão

Aprimorar a administração pública em benefício da sociedade por meio do controle externo.

Visão

Ser referência na promoção de uma administração pública efetiva, ética, ágil e responsável.



Tribunal de Contas da União - v.1, n.1 (1970) -. - Brasília: TCU, 1970

v.

De 1970 a 1972, periodicidade anual; de 1973 a 1975, quadrimestral; de 1976 a 1988, semestral; de 1990 a 2005, trimestral; 2006, anual; a partir de 2007, quadrimestral

ISSN 0103-1090

1. Controle de gastos públicos - Brasil, 2. Controle externo - Brasil, I.
Tribunal de Contas da União.

Catalogada pela Biblioteca Ministro Ruben Rosa



TRIBUNAL DE CONTAS DA UNIÃO

FUNDADOR

Ministro Iberê Gilson

SUPERVISOR

Ministro Raimundo Carreiro

CONSELHO EDITORIAL

Ministro-substituto Augusto Sherman Cavalcanti

Procuradora-Geral Cristina Machado da Costa e Silva

Maurício de Albuquerque Wanderley

Paulo Roberto Wiechers Martins

Luiz Henrique Pochyly Da Costa

Fábio Henrique Granja e Barros

EDIÇÃO

Flávia Lacerda Franco Melo Oliveira

Clemens Soares dos Santos

COLABORAÇÃO

Wiliams Juvêncio Silva

TRADUÇÃO

Secretaria de Relações Internacionais

PROJETO GRÁFICO

Marcello Augusto Cardoso dos Santos

CAPA, DIAGRAMAÇÃO E FOTOMONTAGENS

NCom/ISC

IMAGENS

Getty Images

Departamento de Pós-Graduação e Pesquisas

St. de Clubes Esportivos Sul

Trecho 3 Lote 3

Brasília, DF, 70200-003

revista@tcu.gov.br

Impresso pela Sesap/Segedam

ISSN eletrônico - 2594-6501

ISSN impresso - 0103-1090



Sumário

Índice e Destaques

Entrevista

Desafios para a transformação digital no setor público brasileiro

Luis Felipe Monteiro

Opinião

Transformação Digital: Pessoas, estratégia e tecnologia gerando mais impacto para a sociedade

Rodrigo Felisdório

A auditoria financeira de agências federais nos EUA tem relação custo-efetividade positiva e as lições aprendidas são aplicáveis ao Brasil

Tiago Alves de Gouveia Lins Dutra

Gestão dos Recursos de Tempo de Auditoria: Modelo Matemático de Estimção e de Controle

Euro Gama Barbosa

José Alves Dantas

Daniel Novais Santos

Rastro-DM: Mineração de Dados com Rastro

Marcus Vinicius Borela de Castro

Remis Balaniuk

Rastro-DM: data mining with a trail

Marcus Vinicius Borela de Castro

Remis Balaniuk

4

Destaques TCU **107**

Índice de autores **110**

9

Como publicar na revista do TCU **110**

11

29

39

79



Desafios para a transformação digital no setor público brasileiro

Luis Felipe Monteiro

é Analista de Planejamento e Orçamento, foi Diretor de Modernização, iniciando o movimento de inovação no setor público, com a criação da Rede InvovaGov. Mestre em Gestão de Tecnologia da Informação pela UCB e pela FGV. Graduado em Ciência da Computação pela UFSM. Dirigiu o Departamento de Sistemas de Gestão de Pessoas e coordenou a TI do Ministério das Comunicações.

A transformação digital é tema central e estratégico para o desenvolvimento das nações e tem ocupado posição de destaque nas agendas governamentais e de organismos internacionais. Atualmente, o Brasil é a 4ª maior população conectada do mundo e está na 44ª posição no ranking de governos digitais, segundo a ONU. Quais são os principais desafios e investimentos para que o Brasil avance significativamente nesse cenário?

A transformação digital não se resume à tecnologia. Ela é, principalmente, uma mudança da cultura institucional que inclui aspectos como agilidade e entregas constantes, percepção da experiência do usuário no uso de serviços e canais de atendimento, entre tantas outras coisas que permitam que a inovação e a tecnologia sejam mescladas para a melhoria das políticas públicas. É fundamental que os gestores públicos tenham uma cultura digital. Essa cultura é a mudança na forma de operar o governo, de entender os projetos, de posicionar os resultados.

Nós estamos trabalhando muito em três aspectos. O primeiro é o do patrocínio. A Secretaria de Governo Digital do Ministério da Economia e a Secretaria Especial de Modernização do Estado da Secretaria-Geral da Presidência da República realizam constantes reuniões com secretários-executivos, secretários, e até mesmo ministros para colocar o glossário e as principais mensagens da transformação digital no dia a dia dos executivos de governo. A ideia é que se convençam dos benefícios e priorizem o projeto.

O segundo aspecto é a capacitação. Temos um plano de capacitação em transformação digital operacionalizado pela Enap, que só no ano passado treinou mais de 20 mil profissionais, dividido em sete áreas de conhecimento. Os cursos são voltados a formar servidores públicos

▶▶▶ *A Estratégia de Governo Digital deu o norte sobre quais são os caminhos que nós enxergamos de uma sociedade digital do futuro. Ela estabelece um recorte entre economia digital e governo digital. Essas duas áreas se interconectam e são interdependentes. A vantagem é que a indústria digital no Brasil é bastante desenvolvida.* ◀◀◀

com habilidades de transformação digital, instruindo, por exemplo, sobre quais são os benefícios e como se implementa uma estratégia digital consistente.

O terceiro elemento é a comunicação. Fortalecemos não somente o nosso time de comunicação, mas também as parcerias entre os órgãos, com as assessorias de comunicação e com a Secretaria Especial de Comunicação Social da Presidência da República para traduzir a transformação digital numa leitura menos tecnológica e mais estratégica. A comunicação é elemento chave para criar esse movimento em torno do gov.br e, a partir daí, o estabelecer uma comunidade engajada na transformação digital, que é o que todos esperamos.

Em 2018, foi publicada a Estratégia Brasileira de Transformação Digital – E-Digital, que estabeleceu amplo diagnóstico dos desafios a serem enfrentados e uma visão de futuro para a transformação digital do Brasil. Recentemente, foi publicado o Decreto 10.332 de 28 de abril de 2020, que instituiu a Estratégia de Governo Digital para o período de 2020 a 2022. No Decreto, estão previstos instrumentos que devem ser elaborados pelos órgãos e entidades da administração pública federal direta, autárquica e fundacional para a transformação digital dessas instituições. Como será realizada a rastreabilidade entre os instrumentos mencionados e a mensuração dos resultados obtidos?

A Estratégia de Governo Digital deu o norte sobre quais são os caminhos que nós enxergamos de uma sociedade digital do futuro. Ela estabelece um recorte entre economia digital e governo digital. Essas duas áreas se interconectam e são interdependentes. A vantagem é que a indústria digital no Brasil é bastante desenvolvida.

No governo, precisamos não só adotar a tecnologia, como também promover eficiência para que essa economia digital seja cada vez mais dinâmica. O decreto 10.332 de abril de 2020, que instituiu a Estratégia de Governo Digital, implementa e detalha esse conjunto de iniciativas de governo digital, alinhada à Estratégia Brasileira de Transformação Digital.

Para que os objetivos e metas da Estratégia de Governo Digital sejam cumpridos, a Secretaria Especial de Modernização do Estado (SEME) da Secretaria-Geral da Presidência da República e a Secretaria de Governo Digital (SGD) do Ministério da Economia, precisam coordenar e monitorar as ações, que são operacionalizadas em cada um dos órgãos e instituições a partir de planos de transformação digital. Esses planos são um novo instrumento de planejamento e monitoramento criado a partir da nova estratégia, construídos com todos os órgãos públicos federais - há mais de 60 órgãos com seus planos constituídos.



Esses planos contemplam entregas de projetos divididas por trimestres em quatro eixos principais: o eixo da transformação digital de serviços, que trata de colocar serviços, que antes eram presenciais, nos canais digitais; o eixo da consolidação de canais digitais, que trata da unificação de portais no gov.br e também de uma melhor governança dos apps de governo; o eixo da interoperabilidade dos dados, que trata de minimizar os cadastros repetitivos e a exigência de dados do cidadão a partir da comunicação entre os sistemas; e o eixo da otimização das infraestruturas de sistemas de missão crítica, que trata de maior racionalidade e segurança das infraestruturas que suportam esses sistemas de plataformas digitais de cada um dos órgãos.

Esse conjunto de planos, que vai desde a visão de Estratégia Brasileira de Transformação Digital até as entregas concretas de cada um dos órgãos, são monitorados pela SEME e pela SGD em reuniões de monitoramento com cada um dos órgãos, trimestralmente. Nós temos um percentual de execução de cada um dos planos, de entregas de cada um e acreditamos que, com isso, conseguimos dar objetividade, foco e ritmo para que os desafios e as metas arrojadas da Estratégia de Governo Digital sejam atingidos.

Quando se fala em transformação digital, os maiores desafios estão relacionados ao redesenho dos modelos de negócio praticados (muitas vezes já saturados), bem como o reposicionamento de empresas e instituições públicas no contexto da nova sociedade digital. Entretanto, há instituições públicas que ainda não se atentaram para essa nova realidade. Na sua visão, quais são os principais riscos envolvidos?

O Brasil entrou nessa jornada, de fato, com pelo menos 15 anos de atraso. Nossa primeira estratégia de governo digital iniciou em 2016, enquanto países que hoje lideram o ranking da ONU de governo digital, como a Dinamarca, começaram em 2001. Nós temos um longo caminho pela frente, mas também uma vantagem de estarmos hoje inseridos em um contexto onde três a cada quatro brasileiros estão conectados à internet, 97% deles por meio de celulares. Os principais desafios para a transformação digital de governo não são tecnológicos. A tecnologia que está aí resolve os nossos problemas de maneira adequada. Os principais desafios são institucionais.

A transformação digital exige, em primeiro lugar, uma mentalidade digital nos gestores públicos, em especial os executivos de alto escalão. Para isso, nós estamos fazendo um trabalho constante de convencimento e esclarecimento de quais são os benefícios de negócio, o que o governo e o que o cidadão têm a ganhar por meio da transformação digital.

O segundo desafio é de relacionamento institucional entre os órgãos. Somente no Executivo federal são 220 órgãos que têm unidades de Tecnologia da Informação. Nós precisamos articular e alinhar os investimentos em tecnologia e isso está sendo feito a partir dos planos de transformação digital que são elaborados por cada órgão, com metas precisas e concretas e com controle trimestral dos resultados.

O terceiro desafio é acelerar a velocidade de entregas. O marco legal existente hoje no Brasil para contratações públicas e contratações de tecnologia, de certa forma, torna o processo relativamente lento e, portanto, não sintonizado com a velocidade das mudanças tecnológicas



que o mundo e a sociedade vivem. Para isso, nós estamos trabalhando - e o TCU é parceiro nessas estratégias - em negociações com grandes fornecedores para otimizar custos e acelerar o processo de contratação; na criação de um market place de compras de TI, onde os órgãos terão produtos e serviços padronizados, pré-qualificados e de fácil acesso; e, também, na padronização e na contratação centralizada das plataformas estruturais de governo digital. Não faz sentido termos órgãos repetindo esforços, por exemplo, de login, de portal, de interoperabilidade, sendo que o governo digital pode produzir essas plataformas de forma estruturada e centralizada.

A Estratégia de Governo Digital apresenta princípios, objetivos e iniciativas para nortear a transformação do governo por meio de tecnologias digitais. Na sua visão, quais são os fatores críticos de sucesso para a elaboração de uma estratégia que realmente viabilize a transformação digital de uma instituição pública?

A Estratégia de Governo Digital 2020-2022, que pode ser facilmente acessada na página gov.br/governodigital, é organizada em seis princípios fundamentais que ancoram a nossa visão de governo do futuro. Será um governo centrado no cidadão, um governo integrado, um governo inteligente, um governo confiável, um governo transparente e aberto e um governo eficiente. Esses princípios são detalhados em 18 objetivos, desmembrados em 58 iniciativas. Cada uma com um escopo muito claro, um indicador de mensuração de resultado, uma meta em torno desse indicador, um prazo e um responsável pela execução. Na Estratégia de Governo Digital, nós buscamos ser o mais claro possível em relação a esses aspectos que vão direcionar o nosso monitoramento e a visão de governo nos próximos três anos.

Na minha opinião, alguns fatores críticos de sucesso são: primeiro, “engajamento”. Criamos a Estratégia de Governo Digital ouvindo todos os órgãos públicos em oficinas organizadas na Enap. Realizamos uma extensa consulta pública e recebemos também feedback da sociedade, de especialistas, da academia. Com isso, conseguimos engajar a maior parte dos gestores públicos e da sociedade em torno de objetivos comuns, que traçam o que há de mais inovador entre os governos digitais no mundo.

Um segundo fator crítico de sucesso é “foco”. A Estratégia de Governo Digital é capaz de, claramente, comunicar qual é o foco do governo no uso de tecnologias nos próximos três anos. É importantíssimo que todos os órgãos públicos, os mais de 200 representantes do Sistema de Administração dos Recursos de Tecnologia da Informação (Sisp), estejam centrados e entendam, exatamente, quais são os objetivos. Assim, consigam direcionar os seus recursos no dia a dia, seus contratos, suas equipes, seus esforços descentralizados para o atingimento do objetivo da estratégia.

Um terceiro fator fundamental é “remoção das barreiras”. Naturalmente, em um projeto desafiador e ambicioso, como o da Estratégia de Governo Digital, haverá percalços, barreiras que dificultam o prosseguimento de um projeto ou outro. Por isso, é fundamental que a Presidência da República, juntamente com o Ministério da Economia, estejam envolvidos e que nós sejamos facilitadores para remover as barreiras que esses órgãos, porventura, enfrentem no avanço das suas estratégias digitais.



O quarto e último, mas não menos importante fator crítico de sucesso se refere a “ritmo”. É fundamental que a estratégia não seja apenas uma declaração de intenções, e sim, uma construção de disciplina de monitoramento onde, frequentemente, todos os órgãos e gestores públicos sejam lembrados dos seus compromissos, das suas responsabilidades, das suas entregas. Espera-se, assim, a geração não apenas de uma consciência de urgência e de necessidades de avanço, mas também uma responsabilização de todos os órgãos pelos resultados que precisamos atingir. Em especial neste último aspecto, o Tribunal de Contas da União é um parceiro fundamental. Nós sugerimos que o TCU também alinhe o seu ciclo de monitoramento e de avaliações com o ciclo de entregas da Estratégia de Governo Digital. Dessa forma, nós, órgãos centrais, estaremos emanando mensagens consistentes e sincronizadas, que vão chegar aos órgãos, direcionando-os para os mesmos objetivos que são, naturalmente, uma política pública, um serviço público mais eficiente, mais simples, mais humano e mais digital.



Transformação Digital: Pessoas, estratégia e tecnologia gerando mais impacto para a sociedade

Rodrigo Felisdório

Pós-graduado em Governança de TI e Comunicação aplicada ao Setor Público pela Universidade de Brasília - UnB, pós-graduado em Sistema Orientados a Objetos pela Universidade Católica de Brasília - UCB e Bacharel em Ciência da Computação pela UCB. Professor de pós-graduação em diversas instituições. Certificado como *Project Management Professional* - PMP e *Professional Scrum Master* - PSM I. Atualmente exerce a função de Secretário de Soluções de TI do Tribunal de Contas da União – TCU.

Há 10 anos, o termo “digital” possuía semântica limitada e era comumente associado à forma de transmissão de sinais eletromagnéticos. Aplicativos de mensagens instantâneas eram considerados itens acessórios, assim como trabalhar de forma remota era um sonho distante. Embora o telefone celular e a internet já estivessem por perto, ainda parecia ser possível ter uma vida *offline*.

A transição de Eras pela qual a humanidade está passando incide em um ambiente volátil, incerto, complexo e ambíguo. O “digital” foi ressignificado. Esse fenômeno, quando bem compreendido, tem o potencial de viabilizar profundas transformações no desenvolvimento das nações, no fortalecimento do exercício da cidadania e na prestação de serviços públicos mais adequados às necessidades da população. O novo cidadão, cada vez mais engajado e conectado, não mais aceita a burocracia como elemento preponderante no seu relacionamento com Estado.

Nesse sentido, a transformação digital passa a ser tema central e estratégico para a evolução das nações e tem ocupado posição de destaque nas agendas governamentais e de organismos internacionais. Países como a Dinamarca, Austrália, Coreia do Sul, Reino Unido e Suécia, ocupantes das primeiras posições no Ranking de Governo Digital da ONU, têm investido fortemente



nessa jornada, guiados pelas respectivas estratégias digitais governamentais. Nessa mesma esteira, posicionado na 44ª posição, o Brasil tem evoluído e empreendido esforços para viabilizar o alcance dos objetivos contemplados na Estratégia Brasileira de Transformação Digital – e-Digital.

A evolução de governo analógico para governo digital implica necessariamente no reposicionamento dos órgãos de controle, na reflexão e experimentação de novos modelos fiscalização e na criação de novos valores e capacidades.

Atento a esse cenário, o TCU está elaborando a sua estratégia digital corporativa, cujo propósito é redesenhar a forma de atuação do Tribunal no contexto digital para alavancar o impacto das ações de controle para a sociedade. Esse instrumento está sendo desenvolvido de forma colaborativa com especialistas internos e externos ao TCU e está pautado em três premissas: ser simples, útil e mensurável. Importa salientar também que as pessoas (não a tecnologia) são as verdadeiras protagonistas de qualquer processo transformacional.

Os trabalhos em andamento estão relacionados à reflexão sobre novos paradigmas e modelos de negócio para possibilitar o alcance de resultados expressivos. Não se trata apenas aprimorar, mas transformar. Dar saltos. O exercício do controle externo em tempo real, a disponibilização de alertas para auxiliar o gestor público na identificação de possíveis irregularidades no momento da ocorrência do fato, a intensificação de fiscalizações contínuas baseadas em dados/imagens e a ampliação do controle social por meio de tecnologias emergentes são abordagens possíveis no contexto digital, que aproximam o TCU da administração pública e do cidadão e contribuem para o aumento expressivo da tempestividade e efetividade no acompanhamento das políticas públicas.

Diante das profundas transformações experimentadas pelos diversos segmentos da sociedade, de governos cada vez mais digitais e do novo cidadão mais engajado e exigente, o TCU tem envidado esforços para entregar ainda mais valor à sociedade, além de construir caminhos para estar mais próximo do futuro.



A auditoria financeira de agências federais nos EUA tem relação custo-efetividade positiva e as lições aprendidas são aplicáveis ao Brasil

Tiago Alves de Gouveia Lins Dutra

Auditor Federal de Controle Externo do Tribunal de Contas da União desde 2008. Bacharel em Ciências Contábeis e Especialista em Auditoria Financeira pela Universidade de Brasília (Unb). Especialista e Mestre em Administração Pública pelo Instituto Universitário de Lisboa (ISCTE-IUL).

RESUMO

O artigo avalia a relação custo-efetividade da implantação da auditoria financeira de agências federais dos Estados Unidos da América (EUA) desde 1990, quando foi editada a Lei Chief Financial Officer de 1990 (Lei CFO de 1990). Os principais indicadores avaliados neste trabalho foram: tempestividade, credibilidade e utilidade das informações financeiras para tomadores de decisão, bem como cobertura e custos das auditorias financeiras. A conclusão geral é de que as auditorias financeiras de agências federais nos EUA têm relação custo-efetividade positiva. Em termos de custo, representam menos de 0,01% do orçamento federal e menos de 10% do orçamento total dos serviços de fiscalização do Governo Federal. Quanto aos benefícios, as auditorias financeiras ajudaram as agências federais a divulgar demonstrações financeiras tempestivas, completas e confiáveis. Desde 1996, todas as 24 agências federais de maior materialidade têm sido auditadas anualmente, o que representa mais de 95% do orçamento federal. Em 2018, receberam uma opinião de auditoria sem ressalvas 87,5% dessas 24 agências, 83% das despesas federais e 69% dos ativos federais. Esses percentuais servem de indicador da credibilidade das informações financeiras federais dos EUA. Todas as agências federais publicam suas demonstrações financeiras auditadas em 45 dias após o final do exercício fiscal. Considerando que informações financeiras completas, tempestivas e confiáveis são mais úteis, as auditorias financeiras das agências federais dos EUA ajudam tomadores de decisão a economizar recursos públicos, em especial, por meio de avaliações sobre sustentabilidade fiscal, custos de oportunidade de programas de governo e custo-efetividade de agências federais.

Palavras-chave: Auditoria financeira; Estados Unidos da América; Lei CFO Act de 1990.



1. INTRODUÇÃO

Nos últimos 10 anos, o Tribunal de Contas da União (TCU) tem investido no fortalecimento da auditoria financeira, em especial, no nível do chamado Balanço Geral da União (BGU), que agrega as demonstrações financeiras consolidadas do Governo Federal do Brasil. Esse movimento teve início a partir de um diagnóstico internacional negativo sobre a aderência aos padrões e boas práticas internacionais (PEFA, 2009).

Em 2011, o TCU assinou um acordo com o Banco Mundial para receber recursos para desenvolver capacidade institucional na área de auditoria financeira. Esse projeto culminou no Acórdão 3608/2014-TCU-Plenário, que aprovou a estratégia de fortalecimento da auditoria financeira. Um dos pilares da estratégia estava na descentralização de uma única auditoria do BGU para auditorias de demonstrações de ministérios e entidades materialmente relevantes. Apesar do sucesso dessa medida, para o então Ministério da Fazenda e para o Fundo do Regime Geral de Previdência Social (RGPS), a expansão da auditoria financeira para outras áreas não ocorreu devido às dúvidas quanto à relação custo-efetividade de auditorias financeiras de instituições de menor materialidade.

Nos Estados Unidos da América (EUA), apesar de ser um ambiente institucional distinto, a evolução da auditoria de demonstrações financeiras também foi gradual e acompanhada de avaliações da relação entre custos e benefícios. Nos EUA, a auditoria financeira surgiu como resposta a uma grave crise fiscal e gerencial do Governo Federal do país. Em 1985, o déficit fiscal federal dos EUA era de 5% do Produto Interno Bruto (PIB). No mesmo ano, o *Government Accountability Office* (GAO), órgão equivalente ao TCU nos EUA, divulgou um relatório propondo reformas profundas da gestão financeira federal, incluindo a exigência de publicação de demonstrações financeiras auditadas de agências federais. Cinco anos depois, a Lei *Chief Financial Officers* de 1990 (Lei CFO) estabeleceu um programa piloto para exigir auditorias financeiras de 10 agências federais. Em 1996, a Lei CFO foi aplicada às 24 agências federais de grande porte listadas na Lei CFO. No ano seguinte, foi realizada a primeira auditoria das demonstrações financeiras consolidadas do Governo Federal dos EUA. Em 2002, a exigência de auditoria foi expandida para agências federais de menor materialidade, após avaliação de custo e de benefício pelo GAO.

Considerando que o desafio no Brasil é exatamente avaliar a relação custo-efetividade da expansão da auditoria financeira para além do BGU, do Ministério da Economia e do Fundo do RGPS, o presente trabalho busca identificar as lições aprendidas no caso americano a partir de avaliação de indicadores de custo e de benefício das auditorias financeiras de agências federais nos EUA. Esses indicadores incluem: confiabilidade, tempestividade e utilidade das informações financeiras dos órgãos federais, bem como cobertura e custos de auditorias financeiras.

2. GOVERNANÇA DAS FINANÇAS FEDERAIS NOS EUA

Em geral, o Governo Federal dos EUA produz informações financeiras confiáveis, tempestivas, completas e úteis. Isso é resultado de quase trinta anos de esforços, iniciados



principalmente pela Lei CFO de 1990. Foram necessários dez anos para obter 100% das 24 agências listadas na Lei do CFO publicando demonstrações financeiras tempestivas, sendo, nesse prazo, 75% delas com opinião limpa (sem ressalvas). Os custos mensurados em 1992 e em 2004 ficaram estáveis perto de 0,01% da dotação orçamentária federal (GAO, 1994; CFOC et. al., 2005).

As principais lições aprendidas nos EUA são: 1) o envolvimento do Congresso no processo de prestação de contas e responsabilidade orçamentária é um pré-requisito; 2) o fortalecimento da alta administração com a função do “Diretor Financeiro” (CFO) nas agências do Poder Executivo, equilibrando autoridade e responsabilidade orçamentária; e 3) integração de esforços de auditoria que resultaram em mais eficiência e confiabilidade.

O Congresso dos EUA tem respondido a problemas de *accountability* financeira nos últimos 30 anos com pelo menos treze novas leis. Essas leis têm perspectivas diferentes e podem ser agregadas nas seguintes categorias: 1) governança financeira e prestação de contas; 2) gestão de desempenho e monitoramento; 3) pagamentos indevidos; 4) prêmios de transparência e *accountability*.

A legislação construiu a atual governança financeira federal, fornecendo autoridade ao alto escalão, exigindo padrões e pessoal profissionais e auditorias independentes. Esse sistema de pesos e contrapesos garante que cada agência cumpra sua missão e que tenha o mesmo objetivo em termos de *accountability* financeira: produzir informações financeiras completas, confiáveis e tempestivas.

Figura 1: Síntese da governança financeira nos EUA

Órgãos de gestão	Instâncias de auditoria	Órgãos reguladores
Office of Management and Budget (OMB)	Government Accountability Office (GAO)	Federal Accounting Standards Advisory Board (FASAB)
Department of Treasury	Offices of Inspector General	CFO Council
Chief Financial Officers de cada Agência	Empresas de Auditoria Independente	Council of the Inspectors General on Integrity and Efficiency (CIGIE)

Fonte: Elaboração própria

Além da legislação, regulamentos atualizados ajudaram a garantir a utilidade das informações financeiras. O OMB publicou requisitos relacionados a prestação de contas e auditorias financeiras. O Departamento do Tesouro regulamentou os requisitos do plano de contas (Standard General Ledger - SGL) e dos sistemas financeiros. O GAO divulga o manual de auditoria financeira, que fornece uma metodologia para a realização de auditorias financeiras federais. O Conselho Consultivo para Padrões Contábeis Federais (FASAB) é o normatizador contábil.



A Lei CFO de 1990 estabeleceu o Inspetor Geral (IG) de cada agência como auditor das respectivas demonstrações financeiras. No entanto, o IG está autorizado a contratar uma empresa de auditoria independente para realizar a auditoria. Nem todas as agências tiveram um Inspetor-Geral nomeado sob a Lei do Inspetor-Geral de 1978. Nesse caso, poderia ser diretamente uma firma de auditoria. O GAO tem o poder de avocar e auditar as demonstrações financeiras de qualquer agência listada na Lei CFO de 1990, a critério do Controlador Geral (CG) do GAO ou a pedido de um comitê do Congresso, em vez do IG ou da firma de auditoria. O GAO também pode revisar uma auditoria de uma demonstração financeira conduzida por um IG ou uma firma de auditoria. E o GAO deve se reportar ao Congresso, ao Diretor do OMB e ao chefe da agência que preparou a demonstração financeira, com relação aos resultados da revisão e fazer qualquer recomendação que o CG considerar apropriado.

Em 1994, a Lei de Reforma da Gestão Governamental (GMRA) foi promulgada e exigiu a publicação de demonstrações financeiras consolidadas (CFS) do Governo Federal como um todo, a partir do ano fiscal de 1997, já auditadas pelo GAO. Atualmente, além das CFS, o GAO audita cinco agências federais: Receita Federal dos EUA (IRS), Instituição Federal de Seguro de Depósitos (FDIC), Comissão de Valores Mobiliários (SEC), Departamento de Proteção Financeira do Consumidor e Agência Federal de Financiamento da Habitação. Para conduzir essas auditorias financeiras e outros trabalhos relacionados, o GAO aloca um de seus 15 departamentos de auditoria em auditoria financeira. Esse departamento tem mais de duzentas pessoas, sendo cerca 2/3 de contadores certificados, além de especialistas em tecnologia da informação e especialistas em fraudes. O Departamento de Auditoria Financeira (FMA) tem um diretor executivo, 8 diretores, 39 diretores assistentes e 177 auditores em diferentes níveis de carreira (inicial, intermediário e sênior).

Na década de 1990, os IGs costumavam ser o auditor principal. No entanto, ao longo dos anos, a maioria dos IGs preferiu contratar uma firma de auditoria. Em 2005, 75% das agências da Lei do CFO foram auditadas pela firma de auditoria. Para o ano fiscal de 2018, o IG também era minoria. Mesmo assim, todo IG deve supervisionar o trabalho da firma de auditoria e fornecer uma declaração. Em alguns casos, a firma de auditoria realiza auditorias nas demonstrações financeiras individuais dos departamentos das agências, e o IG audita as demonstrações financeiras em todo o departamento. Foi o caso do Departamento de Defesa em 2018, que envolve 20 auditorias, com 1200 auditores. Mesmo com essa preferência, há pesquisas acadêmicas avaliando diferenças entre a eficácia da auditoria das auditorias governamentais preparadas por auditores públicos ou privados.

Carslaw et. al. (2012) examinaram 601 municípios de nove estados americanos e compararam o número de achados apontados nos relatórios e duas medidas de tempestividade para municípios com auditores governamentais e municípios com auditores privados. Eles concluíram que “os auditores governamentais têm mais probabilidade de encontrar achados na auditoria de uma entidade governamental”. Em relação à tempestividade, o estudo é ambíguo, enquanto afirmou que os auditores governamentais são mais lentos que as empresas de auditoria pública, também informou que a regressão não indica diferença relevante entre eles.



3. BENEFÍCIOS DA AUDITORIA FINANCEIRA DE AGÊNCIAS FEDERAIS DOS EUA

A Lei do CFO de 1990 e emendas foram eficazes em melhorar a responsabilidade financeira federal nos EUA. Esse regulamento visava garantir informações financeiras de qualidade para aprimorar as decisões fiscais e orçamentárias. As principais preocupações estavam relacionadas à tempestividade, confiabilidade e utilidade das informações financeiras fornecidas aos tomadores de decisão.

Pelo menos três indicadores podem ser úteis para avaliar o nível de eficácia da Lei CFO e emendas: 1) índice de opinião não modificada, para avaliar o nível de confiabilidade das demonstrações financeiras; 2) tempestividade das demonstrações financeiras auditadas; e 3) cobertura de auditorias financeiras, para entender em que medida as despesas, as receitas, os ativos e os passivos relevantes são auditados.

Em 1996, apenas 25% das 24 agências da Lei CFO obtiveram uma opinião não modificada. Em 2000, 10 anos após a promulgação da Lei CFO, era de 75%, nível mínimo até hoje. Nos últimos 3 anos, esse nível foi estável em 87,5%, ou seja, apenas três agências modificaram a opinião de pelo menos uma demonstração financeira.

As 24 agências do CFO Act representam mais de 99% dos gastos federais em 1996 (GAO, 1997), cobrindo a maioria dos ativos, passivos, receitas e despesas. Com informações financeiras confiáveis do passado, poderiam ser feitas estimativas para os próximos 75 anos de todas as receitas e despesas para melhor entender e avaliar a sustentabilidade fiscal federal. Além disso, existem estimativas de pagamentos indevidos e mais transparência nas transferências federais.

Em 2000, pela primeira vez, todas as agências da Lei do CFO cumpriram o prazo de 1º de março para preencher e enviar suas demonstrações financeiras auditadas (GAO CFS, 2000). Em 2005, esse prazo passava apenas em um mês e meio após o final do ano fiscal (GAO, 2005). Esse prazo converge para boas práticas que sugerem que as informações financeiras auditadas sejam disponibilizadas em tempo hábil. Segundo a OCDE (2012), uma boa referência para as demonstrações financeiras anuais é inferior a seis meses após o final do ano fiscal do relatório.

3.1 OPINIÃO DE AUDITORIA SEM RESSALVA É UM INDICADOR DE CONFIABILIDADE DAS INFORMAÇÕES FINANCEIRAS

Desde 2000, mais de 75% das demonstrações financeiras das 24 agências do CFO Act são confiáveis. Mais recentemente, 87,5% desses relatórios atenderam aos padrões de confiabilidade. As demonstrações financeiras são confiáveis quando obtêm uma opinião de auditoria limpa (também descrita como opinião de auditoria não qualificada ou sem ressalvas). De acordo com o Manual de Auditoria Financeira do GAO, existem quatro tipos de relatórios: não modificado (limpo), qualificado, adverso e isenção de responsabilidade (abstenção de opinião). Uma opinião não modificada é aplicável quando o auditor conclui que as demonstrações financeiras e as notas anexas são apresentadas de forma justa, em todos os

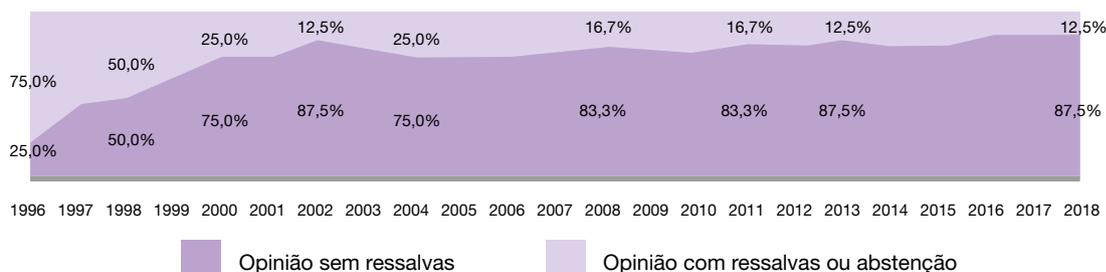


aspectos materialmente relevantes, a partir da data especificada de acordo com os Princípios Contábeis Geralmente Aceitos dos EUA (US GAAP).

Quando há ou pode haver modificação, os outros três relatórios são: 1) uma opinião qualificada, quando há ressalvas materialmente relevantes ou quando não é possível obter evidência de auditoria suficiente e apropriada, porém sem efeitos ou possíveis efeitos generalizados; 2) uma opinião adversa, quando há ressalvas materialmente relevantes com efeitos generalizados; ou, 3) uma abstenção de opinião, quando não é possível obter evidência de auditoria suficiente e apropriada e os possíveis efeitos nas demonstrações financeiras são generalizados.

Em 1996, o primeiro ano em que todas as 24 agências da Lei CFO foram auditadas, apenas seis agências receberam uma opinião não modificada. A maioria das agências de opinião de auditoria limpa havia sido auditada nos anos 80 pelo GAO (antes da aprovação da Lei CFO) e pelos pilotos exigidos pela Lei CFO no início dos anos 90. Isso mostra que leva tempo para melhorar os sistemas e controles financeiros para produzir informações confiáveis.

Gráfico 1: Opiniões de auditoria das agências listadas no CFO Act – por ano (1996-2018)



Fonte: US Federal Government Consolidated Financial Statements FY 1997-2018

Como as agências mais relevantes da Lei do CFO são auditadas anualmente desde 1996, os tomadores de decisão foram capazes de identificar os principais desafios na obtenção de informações financeiras confiáveis:

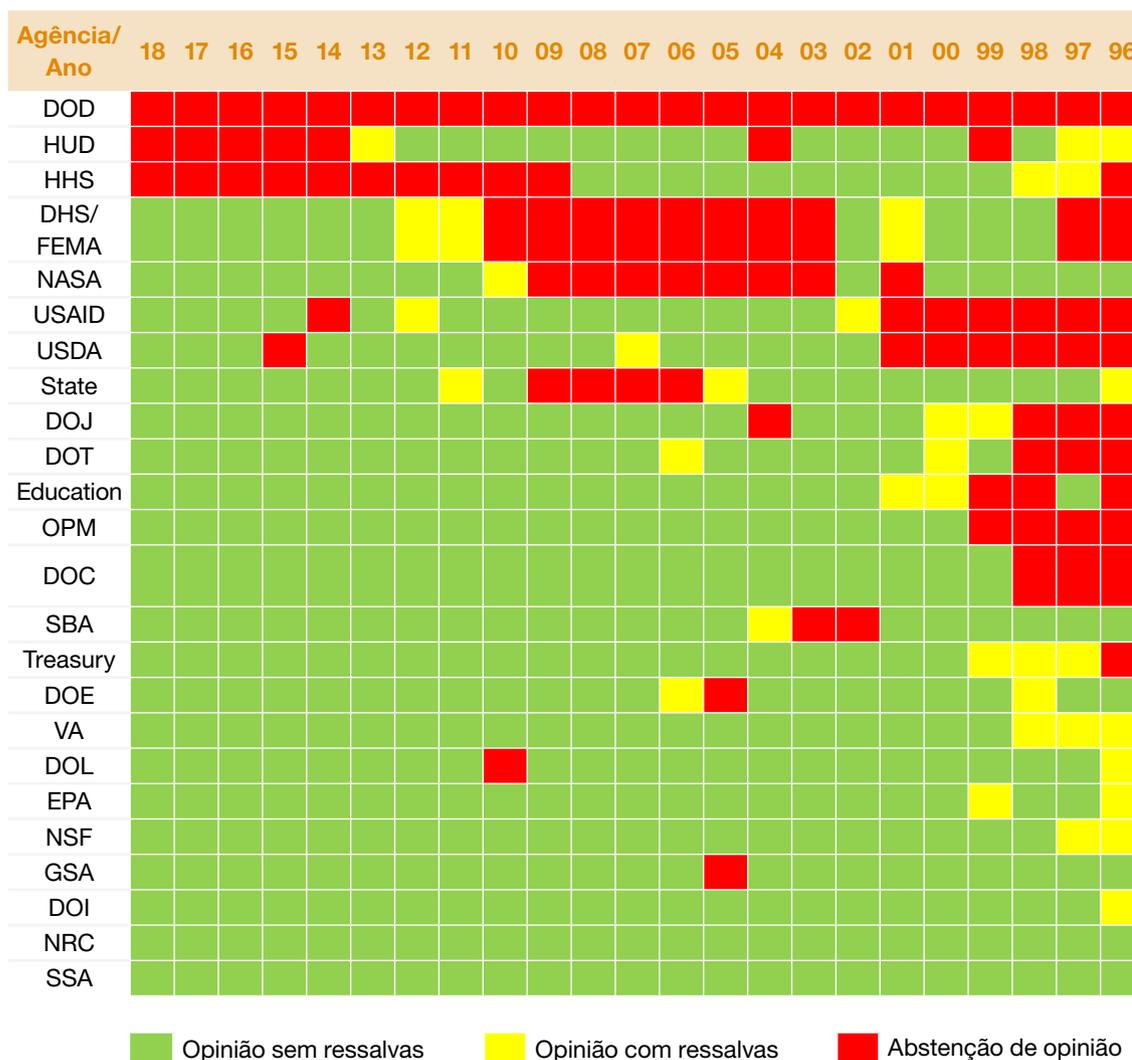
- I. Sempre modificado: Departamento de Defesa (DOD);
- II. Sete piores, mas nem todos modificados (mais de seis pareceres modificados): 1) Administração Nacional de Aeronáutica e Espaço (NASA), 2) Agência de Desenvolvimento Internacional (USAID) e 3) Departamentos de Estado (State), 4) Habitação e Desenvolvimento Urbano (HUD), 5) Serviços Humanos e de Saúde (HHS), 6) Agricultura (USDA) e 7) Segurança Interna (DHS);
- III. Sete intermediários (duas ou mais abstenções de opinião e de três a seis opiniões com ressalvas): 1) Administração de Pequenas Empresas (SBA), 2) Escritório de Gerenciamento de Pessoal (OPM) e 3) Departamentos de Trabalho (DOL), 4) Transporte (DOT), 5) Justiça (DOJ), 6) Educação (Education) e 7) Departamento do Tesouro (Treasury);



- IV. Sete melhores, mas nem todas opiniões sem ressalvas (menos de quatro opiniões com ressalvas e uma abstenção): 1) Administração de Serviços Gerais (GSA), 2) Fundação Nacional de Ciência (NSF), 3) Agência de Proteção Ambiental (EPA) e 4) Departamentos de Interior (DOI), 5) Assuntos dos Veteranos (VA), 6) Trabalho (DOL) e 7) Energia (DOE).
- V. Sempre sem ressalvas: Comissão Reguladora Nuclear (NRC) e Instituto de Seguridade Social (SSA).

O gráfico a seguir mostra a evolução das opiniões de auditoria de cada agência da Lei do CFO. Em geral, o nível de maturidade dos cinco grupos varia ao longo do tempo. Enquanto, na maioria dos casos, uma agência que obtém uma opinião limpa a mantém, existem vários casos de agências que receberam uma opinião modificada após anos de opiniões limpas. A maioria deles teve uma opinião limpa novamente. Essa situação mostra que uma opinião clara não é para sempre e requer esforços para manter as demonstrações financeiras confiáveis

Gráfico 2: Opiniões de auditoria das 24 agências do CFO Act – por agência e ano (1996-2018)



Fonte: Demonstrações Financeiras Consolidadas do Governo Federal dos EUA de 1997 a 2018.

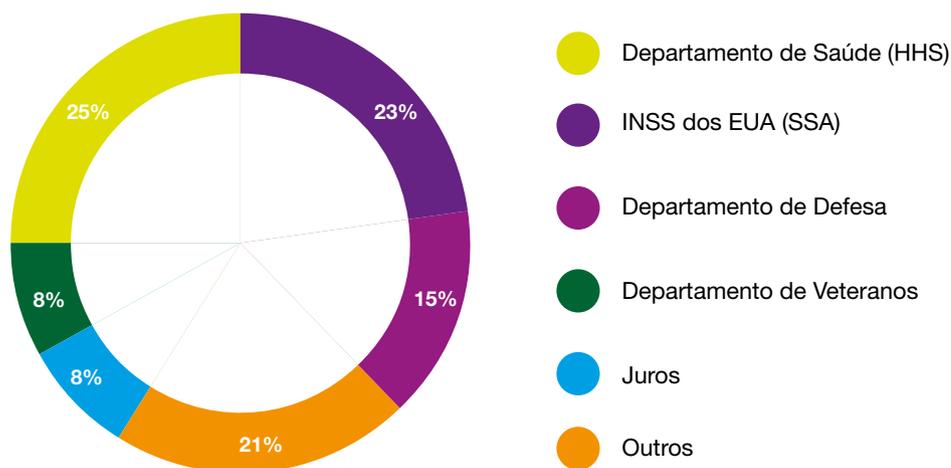
Nos últimos 23 anos de auditoria, das 24 agências da Lei CFO, os reguladores descobriram que 75% das 552 auditorias receberam opiniões não modificadas, 7% receberam opiniões qualificadas e 18% receberam isenções de opinião. Nos últimos três anos (EF 2016-2018), apenas três agências da Lei do CFO não foram capazes de obter uma opinião não modificada sobre suas demonstrações financeiras: DOD, HUD e HHS. Para o HHS, a maioria das demonstrações financeiras obteve uma opinião não modificada, exceto as Declarações de Seguro Social (SOSI) e Demonstrações de Alterações nos Valores da Seguridade Social (SCSIA).

Embora a maioria das receitas, despesas, ativos e passivos tenha opiniões claras, o CFS do governo federal não recebeu uma opinião desde a primeira auditoria em 1997. O principal motivo é a renúncia de opinião do DOD, que possui ativos e custos relevantes. O relatório de auditoria do GAO no CFS do EF 2018 declarou que:

Cerca de 31% do total de ativos reportados pelo governo federal em 30 de setembro de 2018 e aproximadamente 17% do custo líquido reportado pelo governo federal para o ano fiscal de 2018 referem-se a entidades federais significativas que receberam isenções de opinião em suas demonstrações financeiras do ano fiscal de 2018 ou cujas informações financeiras do exercício fiscal de 2018 não foram auditadas (GAO, 2019).

Nesse cenário, o principal desafio para um quadro geral confiável da situação financeira federal dos EUA é uma opinião clara das demonstrações financeiras do DOD. Embora o DOD seja apenas um departamento do governo federal dos EUA, seu orçamento é material e significativo o suficiente para afetar o CFS como um todo.

Gráfico 3: Despesas do Governo Federal em 2018



Fonte: GAO (2019)

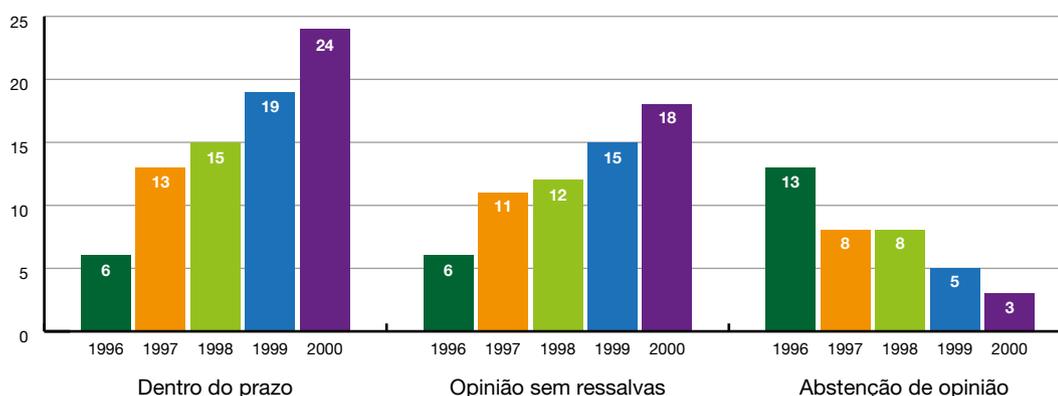
Muitas agências federais obtiveram uma opinião limpa por vários anos e a maioria dos ativos, passivos, receitas e custos está livre de declarações materiais. As auditorias financeiras alcançaram seu objetivo ao fornecer informações confiáveis para 83% dos custos federais e 69% dos ativos federais.



3.2 TEMPESTIVIDADE É FUNDAMENTAL PARA INFLUENCIAR TOMADORES DE DECISÃO

Em 2000, todas as agências da Lei CFO cumpriram o prazo de 1º de março para a publicação de demonstrações financeiras auditadas. Nos EUA, o ano fiscal termina em 30 de setembro.

Gráfico 4: Tempestividade e tipos de opinião das 24 agências listadas no CFO Act entre 1996-2000



Fonte: GAO (2000)

Embora o prazo de cinco meses esteja em conformidade com as boas práticas internacionais, os EUA o reduziram para 45 dias. Agora, as agências federais devem publicar demonstrações financeiras auditadas até 15 de novembro, seguindo os requisitos da Circular OMB A-136 (1.5 Prazos para envio). Esse prazo permite que o CFS seja publicado em março do ano seguinte.

No nível subnacional, em 2010, o *Government Accounting Standards Board* – GASB (normatizador de padrões para o nível subnacional) desenvolveu uma pesquisa com usuários de informações financeiras governamentais. Uma das perguntas era: qual a utilidade ou a informação teria se publicada [período] após o final do ano fiscal? Enquanto apenas 8,8% dos entrevistados consideraram útil em 6 meses, 87,6% responderam muito útil em 45 dias.

Tabela 1: Pesquisa sobre utilidade da prestação de contas conforme o prazo de divulgação após encerramento do exercício

	Muito útil 5	4	3	2	Sem qualquer utilidade 1	Não sabe/ Sem resposta	Total
45 dias	87,6%	6,2%	2,6%	1,5%	0,5%	1,5%	100,00%
3 meses	43,3%	44,8%	9,3%	1,5%	0,0%	1,0%	100,00%
6 meses	8,8%	28,4%	43,8%	14,9%	3,1%	1,0%	100,00%
12 meses	1,5%	6,2%	24,7%	44,3%	22,7%	0,5%	100,00%
Mais de 12 meses	1,5%	0,5%	12,9%	32,0%	51,5%	1,5%	100,00%

Fonte: GASB (2011)



3.3 DEMONSTRAÇÕES FINANCEIRAS AUDITADAS MELHORARAM A TRANSPARÊNCIA DA SUSTENTABILIDADE FISCAL NOS EUA

Os primeiros 10 anos da Lei de CFO foram focados em fornecer credibilidade às demonstrações financeiras. Muitas distorções materialmente relevantes foram identificadas e tiveram que ser corrigidas para garantir demonstrações financeiras confiáveis. No início, as distorções mais relevantes estavam relacionadas a impostos e passivos atuariais, impactando as demonstrações financeiras do Departamento do Tesouro (que inclui a Receita Federal dos EUA – IRS), do Departamento de Veteranos (VA) e da Secretaria de Gestão de Pessoas do Governo Federal (OPM). O VA teve uma opinião limpa em 1999, e OPM e Tesouro conseguiram em 2000.

Tabela 2: Principais distorções identificadas por auditorias financeiras após a Lei CFO

Agência	Distorção	Valor
Receita Federal dos EUA (IRS and Customs Service)	Impossibilidade de atestar que os tributos foram arrecadados quando eram devidos	US\$ 1.2 trilhão
	Recebíveis superestimados	US\$ 110 bilhões
	Arrecadação de recebíveis superestimada	US\$ 30 bilhões
	Impossibilidade de atestar receitas	US\$ 21.6 bilhões
Secretaria de Gestão de Pessoas do Governo Federal dos EUA (OPM)	Passivos atuariais subestimados referentes a pagamentos futuros de servidores federais civis a aposentar e aposentados e seus beneficiários	US\$ 858 bilhões
Departamento de Veteranos (VA)	Passivo estimado não financiado para benefícios com aposentadoria e compensações de veteranos	US\$ 280 bilhões

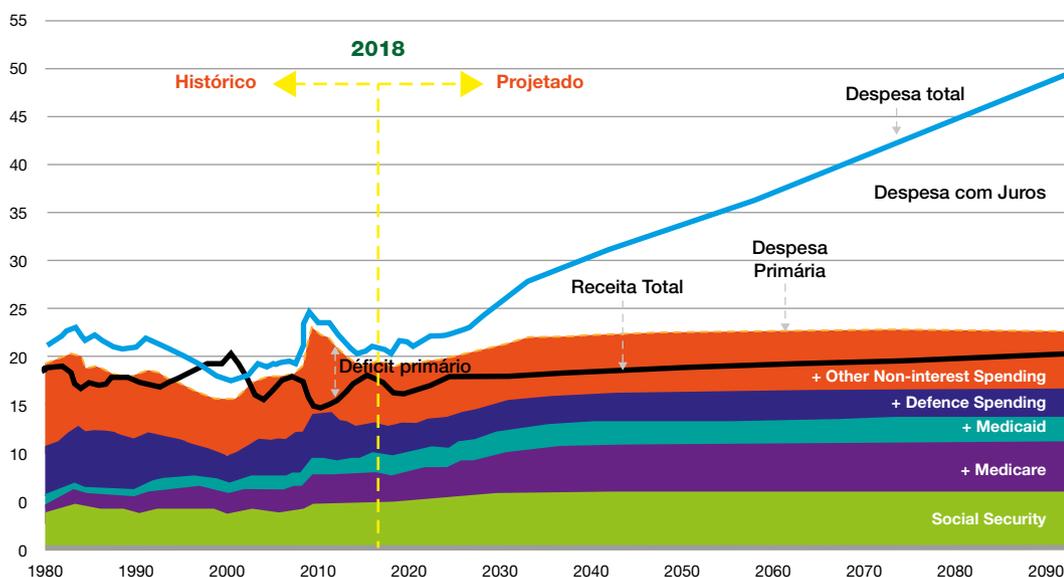
Fonte: GAO (1994)

Esses tipos de informação financeira são fundamentais para subsidiar decisões de políticas fiscais, tributárias, de previdência e de recursos humanos. É sempre difícil reformar essas áreas, uma vez que afetam milhões de pessoas e vários grupos da sociedade. Por causa disso, é crucial fornecer informações confiáveis e relevantes para entender quais são os problemas, onde eles estão e em que intensidade/gravidade.

Em termos de sustentabilidade, também é importante ir além da fiscalização e fornecer previsão e discernimento. Com informações confiáveis, é mais fácil prever futuros desafios fiscais gerais e desafios de sustentabilidade financeira dos programas federais. Nos EUA, a previsão financeira federal é uma das mais abrangentes do mundo. Embora a maioria das normas contábeis do setor público exija projeções apenas para passivos atuariais, nos EUA, há projeções para quase todas as receitas e despesas. O gráfico, a seguir, resume as principais informações financeiras dos últimos 38 anos e as projeta pelos próximos 75 anos.



Gráfico 4: Histórico e projeções da política fiscal vigente do Governo Federal dos EUA para receitas, despesas primárias, despesas com juros e despesa total - Percentual do PIB (1980-2093)



Fonte: GAO (2019)

A Lei CFO foi gestada durante a crise fiscal de 1980, quando o déficit fiscal obteve 5% do PIB. Dez anos após a lei CFO ter sido promulgada, o governo federal teve um superávit entre 1998 e 2001, a única vez que foi positivo no gráfico. Após a crise financeira de 2008, o déficit fiscal ficou ainda pior do que em 1985, mas em cinco anos o governo federal obteve o nível anterior de déficit. Mesmo assim, o déficit recomeçou a crescer e as estimativas são claramente insustentáveis, com uma explosão de juros líquidos por volta de 2030.

4. AUDITORIAS DE AGÊNCIAS COM MAIOR ORÇAMENTO OFERECEM MELHORES COBERTURA E CUSTO-EFETIVIDADE

Pelo menos três estudos foram identificados para entender o custo das demonstrações financeiras auditadas. Dois desses estudos foram relacionados às auditorias das agências da Lei CFO em 1995 (GAO, 1994) e 2005 (CFOC et. al., 2005), e um deles se concentrou em agências não sujeitas à Lei CFO de 1990, divulgadas antes da Lei ATDA de 2002 (GAO, 2002), quando os requisitos para as demonstrações financeiras auditadas foram expandidos para outras instituições federais, agências externas às agências do CFO Act.

De acordo com os dois estudos para as agências do CFO (GAO, 1994; CFOC et. al., 2005), os custos estimados de auditoria financeira foram de cerca de 0,01% da autoridade orçamentária total. Usando dados da autoridade orçamentária do EF 2019 de todas as agências da Lei do CFO e comparando-os com o orçamento do Escritório do Inspetor-Geral (OIG), a porcentagem é inferior a 0,05%. No entanto, apenas parte do orçamento do EIG é designada para executar ou contratar auditorias financeiras.

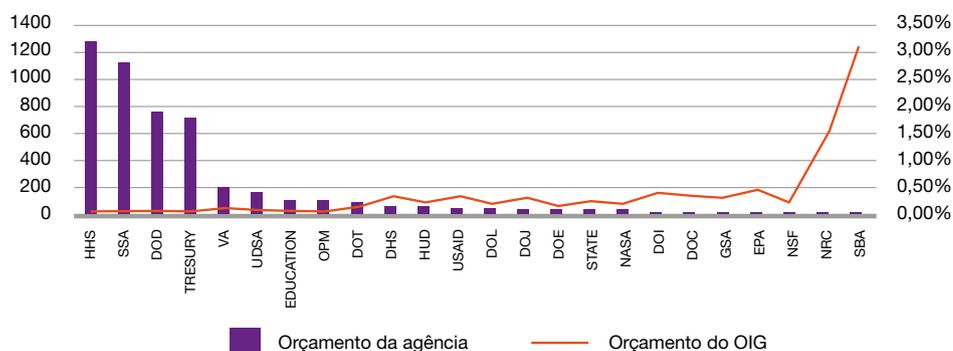
Em 1994, o ex-controlador geral do GAO, Charles Bowsher, relatou o primeiro estudo antes do Congresso. Ele afirmou que, de acordo com as estimativas da OMB, os custos do requisito da Lei CFO para a preparação e auditoria das demonstrações financeiras para 1992 eram de cerca de US \$ 111 milhões, o que representava 0,01% da autoridade orçamentária total auditada. Além disso, o Sr. Bowsher explicou que a experiência histórica em auditoria mostra que os custos de preparação e realização de auditorias às demonstrações financeiras federais devem diminuir substancialmente no futuro. E ele forneceu exemplos em que os custos caíram 40% após o primeiro ano de auditoria (GAO, 1994).

Naquela época, apenas 10 agências da Lei CFO eram obrigadas a ter demonstrações financeiras auditadas. Mesmo assim, segundo Bowsher, as auditorias financeiras cobriram 60% da autoridade orçamentária do governo. Ele disse que a expansão poderia melhorar a cobertura e abordar segmentos relevantes das operações do governo federal que não tiveram o benefício de uma auditoria financeira. Por fim, o Sr. Bowsher sugeriu que um relatório financeiro auditado em todo o governo forneceria ao Congresso e ao público americano uma imagem completa de onde seu governo se encontra financeiramente. Com uma avaliação positiva da relação custo-efetividade, a auditoria financeira foi ampliada nos anos seguintes.

Em 2005, o Conselho do CFO e o Conselho de Inspectores Gerais de Integridade e Eficiência avaliaram a relação custo-efetividade de emitir parecer sobre os controles internos. Para fazer isso, eles compararam as estimativas com os custos das auditorias financeiras das agências da Lei CFO, que foram de US \$ 275,7 milhões. Os IGs federais estimaram que os custos incrementais do trabalho de auditoria necessários para emitir uma opinião sobre controle interno para todas as 24 agências do CFO Act seriam superiores a US \$ 140,6 milhões. Aproximadamente 60% desse total, ou US \$ 84,4 milhões, é a estimativa para dar uma opinião sobre o controle interno do Departamento de Defesa.

Para 2019, o orçamento total do OIG das agências da Lei CFO é de cerca de US \$ 2,3 bilhões. Se considerado que o GAO também audita parte do Departamento do Tesouro, seu orçamento de US \$ 615 milhões deve ser incluído para calcular o custo da supervisão das agências da Lei do CFO. No entanto, como as auditorias financeiras do GAO representam menos de 10% de seus recursos, a auditoria financeira dos OIG é a minoria dos trabalhos, uma vez que existem outros processos de supervisão, como investigações, auditorias de desempenho e avaliações.

Gráfico 4: Orçamento da agência x Orçamento do OIG (estimativa para 2019) - Em US\$ bilhões



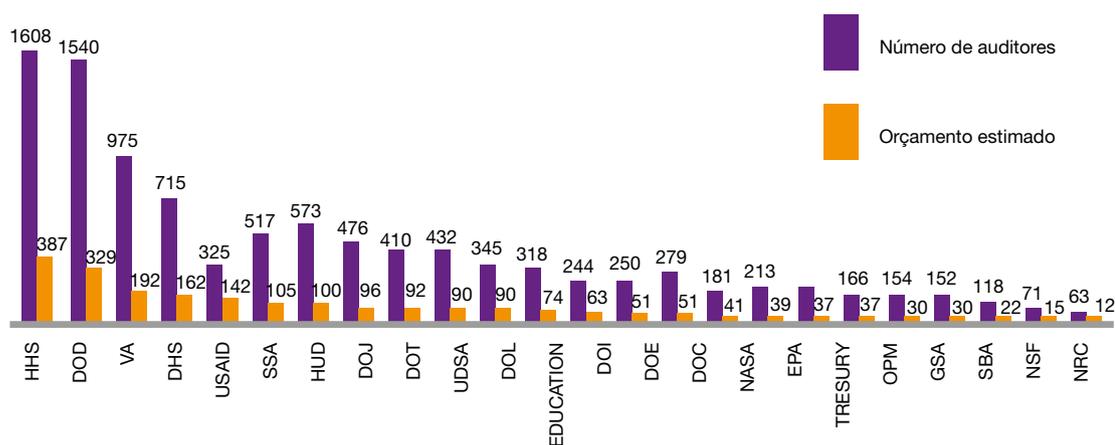
Fonte: OMB (2019) e propostas orçamentárias (*budget justifications*) de 2019 do OIG das 24 agências listadas no CFO Act.



Na SBA, a auditoria financeira representa apenas 8% do orçamento do EIG. Para alguns OIGs, é público que os trabalhos de auditoria contam com cerca de 50% ou menos (HHS, USDA e DOT). Então, mesmo que não exista uma estimativa clara sobre os recursos financeiros, é certo que menos de 0,05% do orçamento das agências da Lei CFO. Além disso, pelo gráfico anterior, pode-se inferir que quanto maior o orçamento da agência, menor a porcentagem do orçamento do EIG em comparação com o orçamento da agência.

Em termos de pessoal, os OIGs das agências da Lei CFO têm mais de dez mil funcionários. Novamente, não há informações sobre a porcentagem de pessoas alocadas em auditorias financeiras; e também, não há informações até que ponto os OIGs contratam empresas de auditoria independentes. Essa informação pode ser útil na comparação com o equivalente à CGU no Brasil, que possui mandato semelhante.

Gráfico 5: Orçamento e número de auditores dos Departamentos de Inspeção Geral de cada agência listada na Lei CFO – Referente ao exercício de 2019 (Em US\$ milhões)



Fonte: Propostas orçamentárias (budget justifications) de 2019 do OIG das 24 agências listadas no CFO Act.

A partir dos dados disponíveis, o custo da supervisão federal em geral não é material o suficiente para atribuir excesso. No entanto, como os controles são implementados como resultado de riscos, é necessário avaliar onde estão as áreas de alto risco e priorizar os esforços de supervisão de acordo com essa avaliação de estratégica de riscos.

Além disso, é importante entender qual é o papel da auditoria financeira e de outros compromissos de atestado em todo o conjunto de ferramentas de supervisão. A auditoria financeira é relevante para fornecer confiabilidade à situação financeira geral. Então, o objetivo da auditoria financeira pode ser alcançado quando a cobertura de transações, ativos e passivos financeiros atingir mais de pelo menos 95% do total. Os outros benefícios são externalidades positivas: melhoria dos controles internos, dinheiro economizado etc. No entanto, eles devem ser considerados para avaliar a relação custo-efetividade da auditoria financeira.

Em 2001, o GAO realizou uma pesquisa com 26 agências que não estão sujeitas à Lei do CFO de 1990 para determinar:

Das 26 agências, apenas 12 prepararam e tiveram suas demonstrações financeiras auditadas cinco anos antes da pesquisa. As 12 agências que tiveram suas demonstrações financeiras auditadas relataram benefícios significativos, especialmente no aprimoramento da prestação de contas e na identificação de ineficiências e fraquezas. A auditoria das demonstrações financeiras também melhorou o controle interno, melhorou a percepção do público sobre a agência, ajudou a agência a atender aos requisitos legais, monitorar ativos e passivos e melhorar sua posição líquida (GAO, 2001).

Tabela 3: Objetivos e conclusões da pesquisa do GAO de 2001 sobre a relação custo-efetividade de auditoria de demonstrações contábeis de agências não listadas na Lei CFO Act de 1990

Objetivos	Conclusões
“Benefícios alcançados ou previstos pelas agências pesquisadas a partir da preparação de demonstrações financeiras auditadas”	“No geral, as agências pesquisadas relataram que elas ou alcançaram benefícios significativos, ou preveem alcançar tais benefícios a partir das demonstrações contábeis auditadas”
“Nível de esforço ou previsão de esforço pelas agências pesquisadas para preparar demonstrações financeiras e submetê-las a auditoria”	“O nível de esforço para preparar demonstrações contábeis e se preparar para uma auditoria de tais demonstrações variou significativamente com o tamanho e outras características das agências”
“Fatores, incluindo orçamento autorizado, que deveriam ser considerados para determinar se agências deveriam preparar suas demonstrações financeiras e submetê-las a auditoria”	“Na avaliação se agências deveriam preparar demonstrações financeiras e submetê-las a auditoria, as agências pesquisadas identificaram uma combinação de fatores que deveriam ser considerados, incluindo a dotação orçamentária, principais contas contábeis e perfil das transações das agências”
“Visão das agências pesquisadas sobre se, em geral, agências deveriam ter suas demonstrações financeiras auditadas”	“Independentemente da importância de tais fatores, 21 das 26 agências relataram que agências federais, em geral, deveriam ter suas demonstrações financeiras auditadas”

Fonte: GAO (2001)

O custo estimado das auditorias de demonstrações financeiras do ano 2000 variaram entre US\$ 11.000 e US\$ 350.000. Essa variação está diretamente relacionada à dotação orçamentária de cada agência.

Tabela 4: Custo das auditorias de demonstrações contábeis das agências pesquisadas pelo GAO em 2001

Porte orçamentário	Número de agências que tiveram demonstrações financeiras do ano 2000 auditadas	Custos estimados de demonstrações financeiras auditadas por firma ou OIG			Outros custos estimados relacionados a auditorias de demonstrações financeiras		
		Mais Baixo	Mais Alto	Média	Mais Baixo	Mais Alto	Média
Menos de US\$ 25 milhões	4	US\$ 11.000	US\$ 54.512	US\$ 29.525	US\$ 0	US\$ 9.300	US\$ 2.325
Entre US\$ 25 milhões e US\$ 150 milhões	4	US\$ 26.000	US\$ 100.000	US\$ 54.000	US\$ 0	US\$ 15.000	US\$ 5.750
Maior que US\$ 6 bilhões	2	US\$ 220.000	US\$ 350.000	US\$ 285.000	US\$ 2.000	US\$ 1.218.000	US\$ 610.000

Fonte: GAO (2001)



No total, o custo de auditoria anual adicional foi estimado em US\$ 5 milhões em 2005, muito menos que os custos das auditorias das agências listadas na Lei CFO Act de 1990, reportados na ordem de US\$ 111 milhões em 1992 e de US\$ 275 milhões em 2004.

As agências pesquisadas relataram que as responsabilidades fiduciárias (existência e possibilidade de endividamento) e os riscos associados com as transações das agências foram os fatores mais importantes a se considerar. Além disso, alguns cortes de materialidade foram apontados como relevantes para tomar esta decisão. Para as cinco agências que responderam que não deveriam ter suas demonstrações financeiras auditadas, uma das razões foi que o patamar de dotação orçamentária delas variou entre US\$ 250 mil e US\$ 333 milhões.

Apesar dos custos mais baixos de agências menos materiais, há resistência em continuar auditando suas demonstrações financeiras. O Sr. David Walker, Controlador Geral do GAO entre 1998 e 2008, declarou em 2016 dez sugestões para melhorar a gestão financeira federal. Uma delas é a seguinte: “A dura verdade é que poucas pessoas dispõem de tempo para ler o CFS do governo federal e muito menos as demonstrações financeiras dos principais departamentos e agências. As auditorias de demonstrativo financeiro subdivididas ou por segmento podem ser caras e pouco úteis para o público. Como resultado, eles devem ser eliminados como requisito” (WALKER, 2016).

5. CONSIDERAÇÕES FINAIS

A conclusão geral é de que as auditorias financeiras de agências federais nos EUA são eficientes. Em termos de custo, representam menos de 0,01% do orçamento federal e menos de 10% do orçamento total dos serviços de fiscalização do Governo Federal. Quanto aos benefícios, as auditorias financeiras ajudaram as agências federais a divulgar demonstrações financeiras tempestivas, completas e confiáveis.

Desde 1996, todas as 24 agências federais de maior materialidade têm sido auditadas anualmente, o que representa mais de 95% do orçamento federal. Em 2018, receberam uma opinião de auditoria sem ressalvas 87,5% dessas 24 agências, 83% das despesas federais e 69% dos ativos federais. Esses percentuais servem de indicador da credibilidade das informações financeiras federais dos EUA. Todas as agências federais publicam suas demonstrações financeiras auditadas dentro de 45 dias após o final do exercício fiscal.

Considerando que informações financeiras completas, tempestivas e confiáveis são mais úteis, as auditorias financeiras das agências federais dos EUA ajudam tomadores de decisão a economizar recursos públicos, em especial por meio de avaliações sobre sustentabilidade fiscal, custos de oportunidade de programas de governo e custo-efetividade de agências federais.

A experiência americana mostra que o TCU está em um caminho seguro, ao priorizar demonstrações financeiras de órgãos, entidades e fundos de maior materialidade. O fato de as auditorias financeiras de agências federais dos EUA de menor materialidade também terem sido avaliadas como eficientes não deve implicar automaticamente a adoção no Brasil. Nos



EUA, a evolução da auditoria financeira ocorreu ao longo de 30 anos e foi monitorada pelo GAO regularmente, com forte protagonismo do Congresso americano no fortalecimento da *accountability* financeira do Governo Federal.

No Brasil, cada passo também precisará ser percorrido com o mesmo cuidado, avaliando custos e benefícios, de modo a só auditar informações financeiras que, de fato, possam influenciar o processo decisório. Em 2019, informações sobre a previdência social subsidiaram o debate da reforma previdenciária. Em 2020, espera-se que as informações tributárias e sobre despesa com pessoal subsidiem as reformas tributária e administrativa. Esses são alguns exemplos de informações financeiras do Governo Federal que precisam ser auditadas, pois, apesar dos custos, têm benefícios muito superiores. A avaliação sobre a possibilidade de outras informações financeiras de órgãos e entidades de menor materialidade terem ou não terem suas demonstrações financeiras submetidas a auditorias financeiras deve ficar para uma etapa seguinte, após a consolidação das auditorias financeiras dos grandes números do Balanço Geral da União.

REFERÊNCIAS

CARSLAW, Carlos; PIPPIN, Sonja; MASON, Richard. Are public sector auditors more effective than private sector audit firms when auditing governmental entities? Some evidence from United States. **Public and Municipal Finance**, n. 1, vol. 1, 2012.

CHIEF FINANCIAL OFFICERS COUNCIL; PRESIDENT'S COUNCIL ON INTEGRITY & EFFICIENCY. **Estimating the Costs and Benefits of Rendering an Opinion on Internal Control over Financial Reporting**. Washington, D.C., 2005.

GOVERNMENT ACCOUNTABILITY OFFICE. **Financial management: CFO act is achieving meaningful progress**. GAO/T-AIMD-94-149. Washington, D.C., 1994.

GOVERNMENT ACCOUNTABILITY OFFICE. **Financial management: implementation of the Federal Financial Management Improvement Act of 1996**. GAO/AIMD-98-1. Washington, D.C., 1997.

GOVERNMENT ACCOUNTABILITY OFFICE. **Financial management: federal financial management improvement act results for fiscal year 1999**. GAO/AIMD-00-307. Washington, D.C., 2000.

GOVERNMENT ACCOUNTABILITY OFFICE. **Financial report of the United States government 2000**. Washington, D.C., 2000.

GOVERNMENT ACCOUNTABILITY OFFICE. **The CFO act and federal financial management: the end of the beginning**. IBM Business Consulting Services. 2005.

GOVERNMENT ACCOUNTABILITY OFFICE. **Financial audit: fiscal years 2018 and 2017 consolidated financial statements of the US Government**. GAO-19-294R. Washington D.C., 2019.



GOVERNMENT ACCOUNTING STANDARDS BOARD. **The timeliness of financial reporting by state and local governments compared with needs of users.** Norwalk, 2011.

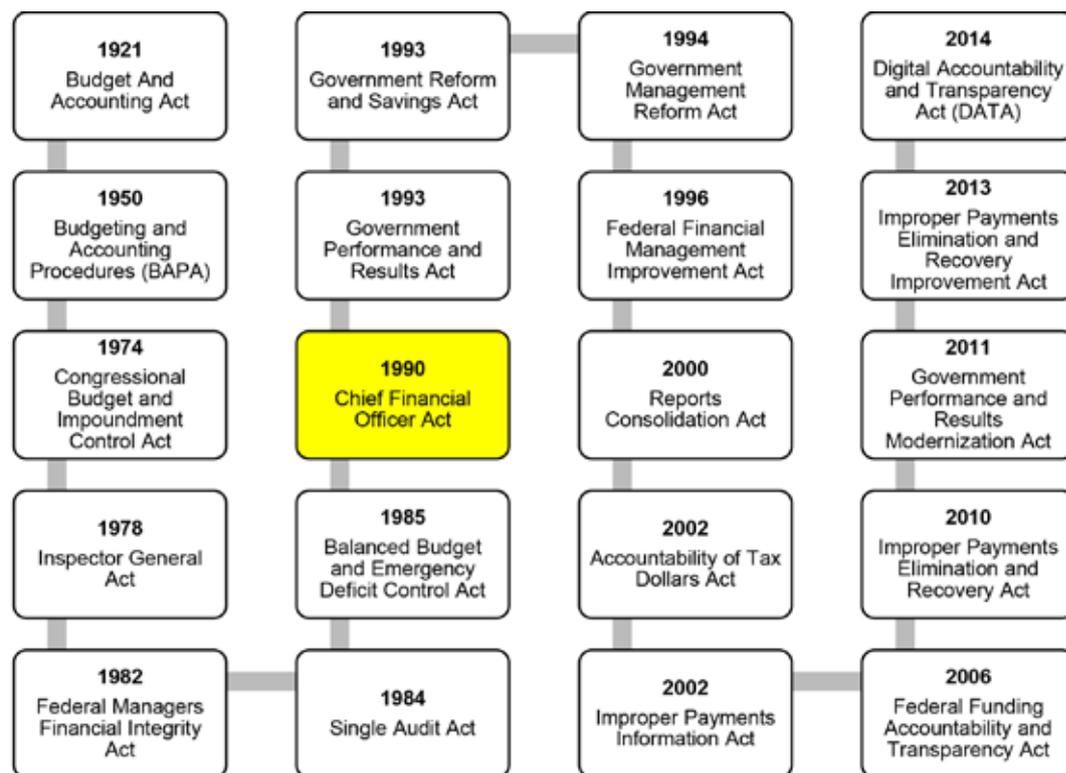
ORGANIZAÇÃO PARA A COOPERAÇÃO E DESENVOLVIMENTO ECONÔMICO. **Brazil's supreme audit institution: the audit of the consolidated year-end government report.** Brasília, 2012.

OFFICE OF MANAGEMENT AND BUDGET. **Outlays by agency.** 2019. Disponível em: <https://www.whitehouse.gov/omb/historical-tables/>

PUBLIC EXPENDITURE AND FINANCIAL ACCOUNTABILITY. **Brazil:** Federal public financial management performance. Brasília, 2009.

WALKER, David M. **Transforming federal financial reporting and auditing.** Government Executive Management. 2016.

ANEXO ÚNICO – LEGISLAÇÃO DE FINANÇAS PÚBLICAS DOS EUA



Fonte: Elaboração própria

Fonte: Elaboração própria

Os conceitos e interpretações emitidos nos trabalhos assinados são de exclusiva responsabilidade de seus autores.



Gestão dos Recursos de Tempo de Auditoria: Modelo Matemático de Estimação e de Controle

EURO GAMA BARBOSA

Doutor em Matemática pela Universidade de Brasília. Servidor do Banco Central do Brasil (BCB)

JOSÉ ALVES DANTAS

Doutor em Ciências Contábeis. Professor da Universidade de Brasília (UnB), no Departamento de Ciências Contábeis e Atuariais (CCA) e Programa de Pós-Graduação em Ciências Contábeis (PPGCont/UnB)

DANIEL NOVAIS SANTOS

Contador, com especialização controladoria organizacional e controles internos e auditoria bancária. Servidor da Caixa Econômica Federal (CEF)

RESUMO

Este estudo teve por objetivo o desenvolvimento de modelos matemáticos para estimação e controle do tempo para a realização das atividades de Auditoria Interna (AI). O modelo inicial teve por base o histórico dos trabalhos realizados, com ajustes em relação: à complexidade, frequência e importância do processo; à maturidade, conhecimento e habilidade dos auditores; à revisão do controle interno e avaliação de riscos; além de fator de ajuste gerencial discricionário. Modelo adicional foi desenvolvido com o propósito de propiciar a avaliação da AI, em relação aos aspectos temporais, baseado na diferença entre os tempos estimado e efetivo dos trabalhos, com ponderações em relação aos processos e às fases do trabalho. Além de contribuir para a literatura sobre AI, notadamente quanto à instituição de instrumentos gerenciais para avaliação da efetividade dos trabalhos, cria instrumentos que facilitam a melhor alocação dos recursos humanos e materiais, minimizando desperdícios e ineficiências.

Palavras-chave: Auditoria Interna; Processo de Auditoria; Tempo de Auditoria; Modelos.

1. INTRODUÇÃO

Historicamente, a Auditoria Interna (AI) é associada ao exercício de um papel essencial para o fortalecimento das entidades, auxiliando na melhoria de sua performance operacional, financeira e econômica (HUGHES, 1977). Não obstante, a literatura sobre o tema pode ser



considerada incipiente, como destacado por Mello et al. (2017), ao avaliarem 25 anos de publicação sobre auditoria nos periódicos brasileiros, quando constataram que a Auditoria Interna foi objeto de estudo em apenas 6 dos 145 artigos publicados na área, nesse período.

Mihret *et al.* (2010) destaca que a Auditoria Interna foi evoluindo à medida que os conceitos de riscos, controles internos e *compliance* evoluíram, de modo que a auditoria ocupa posição de destaque e que sua performance está muito ligada à dinâmica do ambiente em que é praticada. Para os autores, as pesquisas mais recentes sobre Auditoria Interna indicam que sua efetividade reside no dinamismo frente às mudanças.

A definição de Auditoria Interna (AI) apresentada pelo *The Institute of Internal Auditors* (IIA) consolida a amplitude dessa visão, ao destacar que é uma atividade independente e objetiva, com o propósito de adicionar valor e melhorar as operações de uma entidade, auxiliando-a alcançar seus objetivos, adotando uma abordagem sistemática e disciplinada para a avaliação e melhoria da eficácia dos processos de gerenciamento de riscos, de controle e governança corporativa (IIA, 2017).

Os preceitos e orientações estabelecidos pelo IIA têm balizado o exercício da atividade de Auditoria Interna internacionalmente, incluindo os diversos segmentos econômicos. No âmbito da administração pública brasileira, por exemplo, a Instrução Normativa Conjunta do Ministério do Planejamento (MP) e da Controladoria Geral da União (CGU) nº 01, de 10 de maio de 2016, ao disciplinar os processos de controle interno, gestão de riscos e governança, destaca o papel a ser desempenhado pela Auditoria Interna e reproduz precisamente a definição do IIA.

O cumprimento dessa função depende, dentre outros fatores, de um adequado planejamento das atividades a serem desenvolvidas – tanto no estabelecimento de um plano de auditoria baseado na avaliação de riscos para a área como um todo, quanto no planejamento de cada trabalho a ser realizado – conforme preceitua o IIA (2017), por meio das *International Standards for the Professional Practice of Internal Auditing*, doravante designadas como Normas. Essa mesma lógica é observada na Instrução Normativa CGU nº 9, de 09 de outubro de 2018, ao disciplinar a elaboração do Plano Anual de Auditoria Interna (PAINT), a realização dos trabalhos de auditoria e a produção do Relatório Anual de Atividades da Auditoria Interna (RAINT), pelas unidades de auditoria governamentais. Esse processo de planejamento consiste, principalmente, na escolha das atividades de auditoria a serem desempenhadas e na atribuição das horas de trabalho disponíveis a essas atividades, tendo por base a relevância e a necessidade dos trabalhos a serem realizados (ATTIE, 2007).

Dada a relevância do assunto e de modo a balizar a atuação das unidades de auditoria governamentais, a IN CGU nº 3, de 09 de junho de 2017, disciplina o Referencial Técnico da Atividade de Auditoria Interna Governamental do Poder Executivo Federal, considerando o seu propósito de aumentar e proteger o valor organizacional das instituições públicas. Em complemento, a IN CGU nº 8, de 08 de dezembro de 2017, instituiu o Manual de Orientações Técnicas da Atividade de Auditoria Interna Governamental do Poder Executivo Federal, provendo direcionamentos e insumos práticos para a execução de trabalhos de auditoria. Da combinação de todos esses preceitos e aspectos, fica evidente que a adequada distribuição das horas de trabalho disponíveis é uma condição fundamental para que a Auditoria Interna



possa cumprir o seu propósito de melhorar a eficácia dos processos de gerenciamento de riscos, de controle e governança corporativa. O problema é que, de forma geral, isso envolve dificuldades de mensuração eficiente do tempo de realização para cada trabalho, considerando não apenas a escassez de recursos, mas também a variedade de parâmetros que afetam tal decisão, sem que haja comprometimento da qualidade de suas entregas.

Nesse contexto, o presente estudo tem como objetivo desenvolver modelo matemático para estimação e controle do tempo necessário para a realização das atividades de auditoria. O propósito é a otimização das ferramentas de planejamento, que auxiliem na mensuração do tempo de realização das auditorias, considerando as limitações inerentes. Como forma de assegurar a pertinência do modelo proposto, foram realizadas simulações utilizando informações da área de AI de uma autarquia federal, com posterior discussão dos seus resultados com os dirigentes da atividade.

A preocupação com o timing de realização de trabalhos de auditoria também diz respeito às empresas de auditoria independente, que podem se basear na qualidade das verificações realizadas pela Auditoria Interna para planejar o nível de profundidade de suas análises (ABDHEL-KHALIK; et al., 1983). Para Schneider (1984), há uma relação natural de confiança da auditoria externa para com as ações desenvolvidas pela auditoria interna.

Assim, um modelo matemático de estimação voltado para o melhor aferimento do tempo de execução de trabalhos de auditoria interna pode otimizar, também, a eficiência no planejamento das firmas de auditoria independente. Isso sem contar o fato de que o modelo pode também ser adaptado para atender especificamente às necessidades de estimação das horas de trabalho necessárias aos trabalhos de auditoria independente.

O presente estudo contribui, assim, para suprir uma lacuna relevante da literatura sobre auditoria, em particular quanto ao desenvolvimento de instrumentos que possam contribuir para a gestão das ações de auditoria interna, de forma que possam cumprir sua missão de agregar valor à entidade.

2. AUDITORIA INTERNA (AI)

Para a contextualização teórica do foco do estudo, é importante se ressaltar o papel desempenhado pela AI – principalmente em função da série de transformações ocorridas nas últimas décadas, com o redesenho de suas funções, dado o processo de gerenciamento de riscos e controles – e ao uso de gestão do tempo de auditoria como mecanismo de efetividade da AI.

2.1 O PAPEL DA AI NO CONTEXTO DE GERENCIAMENTO DE RISCOS E CONTROLES

Spira e Page (2003) destacam que a redefinição radical da natureza do controle interno como uma dimensão da governança corporativa, requerendo um alinhamento entre controle interno e gerenciamento de riscos, provocou alterações no papel da AI. Esse mesmo posicionamento é expresso por Stewart e Subramaniam (2010), ao ressaltarem o duplo papel que passou a ser



exercido pela profissão – provedor de asseguarção e consultoria – além do envolvimento no gerenciamento de riscos e o uso da AI como ambiente de treinamento gerencial.

Essa percepção sobre a alteração do papel é expressa pelo próprio IIA (2020), ao destacar que a AI visa adicionar valor e aperfeiçoar as operações de uma organização, auxiliando no alcance dos objetivos a partir da avaliação sistemática dos processos de gerenciamento de riscos, controles e governança.

Tratando especificamente do propósito de fortalecimento das entidades, por meio do gerenciamento eficaz de riscos e controles, o IIA (2020) destaca o “Modelo das Três Linhas”, atribuindo a cada unidade da organização a responsabilidade pela gestão de riscos e controles, cabendo à AI exercer o papel de terceira linha, prestando avaliação e assessoria independentes e objetivas sobre a adequação e eficácia da governança e do gerenciamento de riscos. Esse modelo facilita o gerenciamento de riscos e controle, ao esclarecer o papel das diversas unidades, que atuam em conjunto, promovendo maior eficácia ao processo.

Quanto melhor for a segregação entre as três linhas e a definição dos papéis de cada grupo/área, mediante políticas, procedimentos e mecanismos de reporte, mais consolidado será o gerenciamento de riscos. Isso evita inclusive o risco de sobreposição de atividades, ou seja, de que diferentes áreas estejam realizando a mesma tarefa (COSO, 2015).

2.2 GESTÃO DO TEMPO DE AUDITORIA COMO MECANISMO DE EFETIVIDADE DA AI

De forma ampla, a efetividade da AI depende do alcance dos seus propósitos e responsabilidades, da atuação de acordo com o que preveem as normas profissionais e o código de ética e da consideração de tendências e questões emergentes que possam impactar a entidade (IIA, 2017 – Norma 2000).

Em termos mais específicos, a Norma 2030 (IIA, 2017) prevê que o responsável pela AI assegure que os recursos de auditoria sejam: (i) apropriados, característica referente à combinação de conhecimentos, habilidades e outras competências; (ii) suficientes, refere-se à quantidade de recursos; e (iii) efetivamente implantados para a realização daquilo que foi planejado. A gestão dos recursos de auditoria representa, portanto, importante aspecto a se considerar para que a AI cumpra com efetividade o seu propósito.

Essa percepção é reforçada pela reflexão de Hughes (1977), para quem a maior dificuldade no planejamento global dos trabalhos de auditoria é identificar o timing mais adequado para sua realização, visto que as decisões do que auditar, quando, por quem, como e em que prazo, impactarão em maior ou menor eficiência. A própria percepção dos auditores sobre a quantidade de horas necessárias à execução de trabalhos varia entre os profissionais, conforme evidenciaram Gaumnitz et al. (1982), que entrevistaram, na época, 35 auditores, constatando um alto grau de dispersão sobre o que seria o volume adequado de horas – variando entre 21 e 672 horas.

Hughes (1977) também destaca que a dificuldade de se identificar o melhor momento de execução de trabalho de auditoria demonstra que não há política ótima de prazos, visto que



o intervalo entre os trabalhos não necessariamente será o mesmo e que os resultados de um podem reduzir ou alongar o prazo para a realização do próximo.

Em pesquisas realizadas por Azad (1994), visando avaliar o comportamento dos auditores diante de pressões relacionadas a tempo de execução dos trabalhos, os resultados indicam que mensurações inadequadas dos prazos contribuem para a criação de conflitos entre o tempo necessário para análises e o devido levantamento de evidências que suportem as análises.

Kesuma e Dwirandra (2019) corrobora com essa visão, tendo em vista a constatação de que pressões relacionadas ao tempo de auditoria, em obediência ao prazo inicialmente estipulado, possuem impacto significativamente negativo na qualidade dos resultados apurados.

Se a literatura reconhece a dificuldade implícita no processo de estabelecimento do que seria o prazo apropriado para a realização dos trabalhos de auditoria, há posicionamentos sobre fatores que poderiam ser considerados para essa definição, tais como: a relevância, risco e sensibilidade do processo de auditoria para a entidade – a relação com a atividade fim pode ser um parâmetro que destaque um processo frente aos demais (ATTIE, 2007; DIAS, 2015); o nível de conhecimento e a experiência dos auditores responsáveis pela execução ou supervisão dos trabalhos (ALMEIDA, 2017; MAFFEI, 2011; DIAS, 2015); o grau de especialização dos auditores (DIAS, 2015); a frequência com que as auditorias são realizadas em cada processo (ALMEIDA, 2017; ATTIE, 2007); a avaliação preliminar da efetividade do controle interno e gerenciamento de riscos (ATTIE, 2007; DIAS, 2015). Todos esses aspectos encontram amparo nas orientações do emissor de normas profissionais (IIA, 2017; IIA, 2020).

Explorando a relação entre o papel desempenhado pela auditoria e a instituição de mecanismo de gestão e avaliação, cabe ressaltar o estudo desenvolvido por Soh e Martinov-Bennie (2011), com o propósito de avaliar os papéis e responsabilidades da AI e os aspectos percebidos como necessários para assegurar sua efetividade. Os autores concluíram que houve expansão significativa e reposicionamento do papel da AI, bem como percepções de sua efetividade. Identificaram, porém, que mecanismos de avaliação de performance não evoluíram na mesma velocidade, o que provoca um desalinhamento e aumenta a dificuldade em avaliar a extensão com que as atividades desenvolvidas pela AI atendem às expectativas dos stakeholders.

É nesse contexto que se insere o presente estudo, ou seja, o desenvolvimento de modelo matemático para estimação e controle do tempo necessário para a realização das atividades de auditoria, que considere os diversos aspectos sugeridos na literatura, e que funcione como instrumento efetivo de gestão e avaliação de performance da AI, aumentando as expectativas de sua efetividade. Os modelos para estimação do tempo para a auditoria e para avaliação da eficiência da AI são apresentados nas Seções 3 e 4, respectivamente.

3. MODELO MATEMÁTICO PARA ESTIMAÇÃO DO TEMPO PARA A AUDITORIA

O modelo desenvolvido para a estimação do tempo necessário para a realização de uma auditoria tem por base um estimador (\hat{t}) que considere o histórico dos tempos estimados e



efetivos nos trabalhos anteriores realizados, considerando as especificidades associadas ao nível de frequência (alto/médio e baixo/inédito) de realização do processo auditável. Essa referência base é ajustada em função dos parâmetros associados: à complexidade, frequência de realização e importância do processo a ser auditado; à maturidade da equipe da equipe de auditores responsável pela realização do trabalho; à necessidade de revisão preliminar do controle interno e da avaliação de riscos; e a um fator de ajuste gerencial.

Resumidamente, a proposta de estimação do tempo para a realização da auditoria tem por base o modelo (3.1).

$$E_{TES} = \bar{X} P_C P_F P_I P_{ME} P_{CIAR} F_{AG} \quad (3.1)$$

Onde:

E_{TES} estimativa do tempo esperado para o processo a ser auditado.

\bar{X} base do estimador.

P_C peso da complexidade do processo a ser auditado.

P_F peso da frequência de realização do processo a ser auditado.

P_I peso da importância do processo a ser auditado.

P_{ME} peso da maturidade da equipe responsável pelo processo a ser auditado.

P_{CIAR} peso relativo ao nível de atualização necessária ao processo de revisão preliminar do Controle Interno e da Avaliação de Riscos do processo a ser auditado.

F_{AG} fator de ajuste gerencial, em que o gestor promove ajustes na estimativa do tempo esperado para a realização da auditoria.

O detalhamento desse modelo, com a definição e fundamentação de cada uma das variáveis que o compõem, se dá nas subseções a seguir.

3.1 BASE DO ESTIMADOR (\bar{X})

O ponto central do modelo de estimação do tempo para realização do trabalho de auditoria é a variável \bar{X} , que representa a expressão da média de um conjunto de valores, relacionados principalmente com o histórico de trabalhos realizados anteriormente no processo auditável em questão. Além de observar as premissas definidas pelo IIA (2017), a consideração da frequência com que as auditorias são realizadas no processo em questão para definir a estimativa do tempo necessário para a realização do trabalho também é prevista por Attie (2007) e Almeida (2017). Assim, são aplicados métodos distintos para a apuração da variável, de acordo com o nível de frequência de trabalhos.



Para os efeitos do presente estudo, a frequência é medida em função dos seguintes parâmetros: (i) frequência alta, se o processo foi auditado pelo menos uma vez em, no mínimo, três dos últimos quatro períodos anuais; (ii) frequência média, se o processo foi auditado pelo menos uma vez em dois dos últimos quatro períodos anuais; e (iii) frequência baixa, nos demais casos, aqui incluídos os que serão auditados pela primeira vez.

3.1.1 PROCESSOS AUDITÁVEIS COM FREQUÊNCIA ALTA OU MÉDIA

Para o caso de processos que são regularmente auditados (frequência alta ou média), a mensuração de se baseia nas informações relativas ao tempo estimado nos trabalhos anteriores, ao tempo efetivamente gasto e à comparação do tamanho das equipes do trabalho atual e dos trabalhos anteriores. Esse método é sistematizado na Equação (3.2).

$$\bar{X} = \frac{(x_{aa} T_{ES_{aa}} + y_{aa} T_{EF_{aa}}) \frac{m_{aa}}{m_A} + (x_a T_{ES_a} + y_a T_{EF_a}) \frac{m_a}{m_A}}{x_{aa} + y_{aa} + x_a + y_a} \quad (3.2)$$

Onde:

- a indica a última realização do processo a ser auditado no período anual mais recente em que ocorreu.
- aa indica a última realização do processo a ser auditado no segundo período anual mais recente em que ocorreu
- T_{ES_a} tempo esperado da última auditoria realizada no processo a ser auditado.
- $T_{ES_{aa}}$ tempo esperado da penúltima auditoria realizada no processo a ser auditado.
- T_{EF_a} tempo efetivo da última auditoria realizada no processo a ser auditado.
- $T_{EF_{aa}}$ tempo efetivo da penúltima auditoria realizada no processo a ser auditado.
- m_a número de componentes da equipe de auditoria no último trabalho realizado no processo a ser auditado.
- m_{aa} número de componentes da equipe de auditoria no penúltimo trabalho realizado no processo a ser auditado.
- m_A número de componentes da atual equipe de auditoria responsável pelo processo a ser auditado.
- x e y pesos utilizados para proporcionar melhor equilíbrio à base do estimador.

Como se percebe, os pesos (x e y) assumem importância relevante na mensuração de . O propósito é atribuir maior valor: ao tempo esperado do que ao tempo efetivo dos últimos trabalhos; e às informações dos trabalhos mais recentes do que aos mais remotos. A Tabela 1 consolida duas possibilidades de atribuições de pesos – simulações realizadas com as duas alternativas de distribuição dos pesos geraram resultados equivalentes.

Tabela 1: Alternativas de atribuição de pesos para melhor equilíbrio da base do estimador

	x_{aa}	y_{aa}	x_a	y_a
Alternativa A	1,5	0,5	2,0	1,0
Alternativa B	2,0	1,0	3,0	2,0

3.1.2 PROCESSOS AUDITÁVEIS COM FREQUÊNCIA BAIXA OU INÉDITOS

Se quando são realizados trabalhos frequentes de auditoria no processo auditável em questão é possível se utilizar informações históricas como referência para a base do estimador , quando essa frequência é baixa ou quando o processo é inédito – nunca foi auditado – há uma dificuldade adicional nessa mensuração.

Para lidar com essa dificuldade foi incorporado o conceito de “família” de processos auditáveis, que consiste em um conjunto de processos auditáveis que guardam similaridade entre si, baseada, por exemplo, no tema, no escopo e/ou no modus operandi. Nesse caso, têm-se três possíveis situações para o processo a ser auditado:

- pertence a uma família não unitária que possui pelo menos um processo auditável com nível de frequência alto ou médio;
- pertence a uma família não unitária que possui todos os processos auditáveis com nível de frequência baixo; e
- pertence a uma família unitária cujo único processo auditável é o próprio processo que se objetiva estimar o tempo esperado.

Para a situação descrita no item “a”, a mensuração do estimador tem por base a Equação (3.3), que reproduz a mesma lógica e fundamentos da Equação (3.2), exceto pelo fato de não se restringir a informações do processo específico, mas sim do conjunto de processos da “família”.

$$\bar{X} = \frac{\sum_{i=1}^n \left((x_{aa_i} T_{ES_{aa_i}} + y_{aa_i} T_{EF_{aa_i}}) \frac{m_{aa_i}}{m_A} + (x_{a_i} T_{ES_{a_i}} + y_{a_i} T_{EF_{a_i}}) \frac{m_{a_i}}{m_A} \right)}{\sum_{i=1}^n (x_{aa_i} + y_{aa_i} + x_{a_i} + y_{a_i})} \quad (3.3)$$



Onde:

n indica o número de processos auditáveis com nível de frequência alto ou médio que pertence à família do processo a ser auditado.

Para os processos enquadrados nos itens “b”, a Equação (3.3) é inexecutável, tendo em vista a inexistência de processos auditáveis com nível de frequência alto ou médio – a realização de auditoria para cada um dos elementos da família, se ocorreu, ficou restrita a um único dos quatro períodos anuais, mesmo com a possibilidade de eventual repetição – impossibilitando a aplicação da condição “aa”, pois não existiriam dois trabalhos realizados no quadriênio.

Nesses casos, aplica-se a Equação (3.4), desde que: (i) a família contenha um subconjunto de processos que foram auditados pelo menos uma vez; e (ii) o subconjunto de processos esteja disperso por pelo menos dois dos quatro períodos anuais, o que permitiria conferir certa robustez à quantidade distinta de processos auditados, uma vez que o subconjunto teria de conter, pelo menos, dois elementos.

$$\bar{X} = \frac{\sum_{i=1}^n \left((x_i T_{ES_{a_i}} + y_i T_{EF_{a_i}}) \frac{m_{a_i}}{m_A} \right)}{\sum_{i=1}^n (x_i + y_i)} \tag{3.4}$$

Onde:

n número de elementos do subconjunto que atendam às condições destacadas no parágrafo anterior.

No caso dos processos auditáveis enquadrados no item “c”, a própria unicidade inviabiliza o estabelecimento de parâmetros comparativos para se estimar o tempo esperado do processo a ser auditado na forma proposta nas Equações (3.3) ou (3.4).

A alternativa, nessas situações, é estimar o valor de \bar{X} a partir da fixação, de forma ad hoc e baseada na expertise dos gestores da área de auditoria interna da organização, a partir da Equação (3.5), utilizando a estimação do quantitativo de horas.

$$\bar{X} = \frac{qH}{m_A}$$

Onde:

qH número de quantidade de horas estimado pelos gestores.

Quando o processo, apesar de se enquadrar no item “b”, ou seja, pertence a uma família não unitária que possui todos os processos auditáveis com nível de frequência baixo, não atende às condições para aplicar a Equação (3.4), deve-se utilizar como referência a Equação (3.5).



3.2 COMPLEXIDADE (P_C), FREQUÊNCIA (P_F) E IMPORTÂNCIA (P_I) DO PROCESSO A SER AUDITADO

Definido o estimador base (P), o passo seguinte consiste em promover os ajustes necessários à otimização da estimação do tempo necessário para a realização da auditoria, considerando que esse não deve reproduzir, necessariamente o comportamento histórico, sob pena de consolidar ineficiências.

Nesse sentido, o primeiro grupo de variáveis utilizadas para promover esses ajustes se concentra nas características do próprio processo auditável – complexidade, frequência de realização e importância para os objetivos organizacionais. Segue, assim, os preceitos definidos por Attie (2007) e Dias (2015) de que a relevância, o risco e a sensibilidade do processo de auditoria para a entidade devem ser consideradas no estabelecimento do tempo previsto para a realização da auditoria.

3.2.1 COMPLEXIDADE

O pressuposto considerado é que quanto mais (menos) complexos os processos a serem auditados, maior (menor) a demanda de tempo para a realização dos trabalhos de auditoria.

Embora a identificação do grau de complexidade dos processos não seja algo trivial, a elaboração de um ranking/categorização quanto a essa dimensão por parte da área de auditoria interna da organização pode ser uma ação relevante a ser utilizada – com tendência de aperfeiçoamento, na medida em que esse mecanismo vai se repetindo ao longo do tempo. A partir desse ranking, os processos seriam categorizados em níveis de complexidade alto, médio ou baixo.

3.2.2 FREQUÊNCIA

Define-se como premissa de que quanto mais (menos) frequentes forem os trabalhos de auditoria realizados no processo em questão, menor (maior) a demanda de tempo para a realização dos trabalhos de auditoria.

Para os efeitos do presente estudo, a frequência é medida a partir dos seguintes parâmetros: frequência alta, se o processo foi auditado pelo menos uma vez em, no mínimo, três dos últimos quatro períodos anuais; frequência média, se o processo foi auditado pelo menos uma vez em dois dos últimos quatro períodos anuais; e frequência baixa, nos demais casos, aqui incluídos os que serão auditados pela primeira vez.

3.2.3 IMPORTÂNCIA

Adota-se como referência que quanto mais (menos) importante for o processo auditável para o alcance dos objetivos estratégicos da entidade, maior (menor) a atenção que deve merecer



por parte da auditoria interna, o que se traduz em maior (menor) demanda de tempo para a realização dos trabalhos de auditoria.

Uma primeira possibilidade para categorização dos processos em relação ao nível de importância pode ser o grau de associação com os objetivos organizacionais, ou seja: nível alto, se vinculado aos objetivos estratégicos da entidade; nível médio, se associado a objetivos táticos; e nível baixo, se estiver vinculado a objetivos operacionais.

Como se percebe, em termos práticos, são atribuídos conceitos para cada uma dessas características dos processos auditáveis, utilizando os conceitos “baixo”, “médio” e “alto” para cada um desses fatores. O passo seguinte consiste em atribuir peso a esses conceitos, adotando-se a premissa de que o conceito intermediário (“médio”) é neutro em relação à atribuição de maior ou menor tempo para a realização de auditoria. A distribuição, portanto, varia em relação aos conceitos “baixo” e “alto”, conforme a situação. Para efeitos do presente estudo são atribuídos os pesos destacados na Tabela 2.

Tabela 2: Parâmetros para os pesos relativos às características do processo auditável

Fator/Nível	Baixo	Médio	Alto
Complexidade	0,95	1,00	1,05
Frequência	1,05	1,00	0,95
Importância	0,95	1,00	1,05

Na medida em que o gestor da área de auditoria interna se sentir confortável, podem ser aplicadas escalas distintas – com maior granularidade, por exemplo – buscando dados com maior sensibilidade a alteração nos fatores. Essa mesma premissa será aplicada em outros tópicos em que esse mesmo tipo de conceito qualitativo for aplicado (Tabelas 3, 4 e 5).

3.3 MATURIDADE DA EQUIPE (P_{ME}) DE AUDITORES

Definido como as características do processo auditável são consideradas no modelo de estimação do tempo de auditoria, o passo seguinte consiste em agregar as especificidades da equipe que realizará o trabalho. Nesse caso, parte-se do pressuposto de que quanto maior (menor) a maturidade da equipe de auditores, menor (maior) o tempo demandado para a realização de um trabalho de auditoria com a qualidade requerida. Orientações nesse sentido são defendidas por Maffei (2011), Dias (2015) e Almeida (2017), para quem o estabelecimento do tempo de auditoria deve considerar o nível de conhecimento, o grau de especialização e a experiência dos auditores responsáveis pela execução ou supervisão dos trabalhos.

De acordo com os resultados apurados por Endaya e Hanefah (2016), as características e o conhecimento dos auditores possuem impacto significativo na efetividade dos trabalhos executados pela Auditoria Interna, o que, conseqüentemente, afetará o seu tempo de execução.



Sumito e Setiyawati (2019) corrobora com essa visão ao constatar em suas análises a correlação positiva de variáveis como competência e profissionalismo dos auditores em relação à qualidade das auditorias executadas.

A premissa é bastante intuitiva: a formação, a certificação por órgão reconhecido, a experiência profissional, o conhecimento em relação ao processo auditável específico, as horas de capacitação dos auditores, entre outros aspectos, traduzem um nível de maturidade da equipe, tornando-a capaz de realizar determinado trabalho de auditoria em tempo mais restrito. Atribuir o mesmo tempo para realização de uma auditoria a equipes com graus de maturidade distintos pode resultar no comprometimento da qualidade dos trabalhos desenvolvidos pela equipe com menor maturidade ou um ambiente de acomodação para a equipe com maior maturidade.

Seguindo a mesma lógica do bloco anterior, o grau de maturidade da equipe é classificado como “alto”, “médio” ou “baixo”, podendo se utilizar, como referência, a avaliação média dos membros da equipe de auditores em relação aos itens de uma cesta de critérios como formação, certificação, experiência, conhecimento do processo específico, horas de capacitação, etc. Para efeitos do presente estudo, a atribuição de pesos a esses conceitos observa o disposto na Tabela 3.

Tabela 3: Parâmetros para os pesos relativos à maturidade da equipe de auditores

Fator/Nível	Baixo	Médio	Alto
Maturidade	1,05	1,00	0,95

3.4 REVISÃO PRELIMINAR DO CONTROLE INTERNO E DA AVALIAÇÃO DE RISCOS (P_{CiAR})

A etapa de planejamento de um trabalho de auditoria passa, necessariamente, pela revisão preliminar do controle interno e da avaliação de riscos. O pressuposto, nesse caso, é que quanto maior (menor) a recorrência de trabalhos de auditoria no processo e o conhecimento da equipe sobre esse processo, maior (menor) a facilidade de se promover essa revisão preliminar – base da etapa de planejamento do trabalho – o que resultará em menor (maior) demanda de tempo para a realização do trabalho de auditoria. Orientações nesse sentido são estabelecidas pelo IIA (2017) e reforçadas por Attie (2007) e Dias (2015).

Para a aplicação do modelo, a atribuição de conceitos ao fator de necessidade de revisão preliminar do controle interno e da avaliação de riscos, pode considerar a seguinte relação: necessidade alta, quando não há mapeamento do controle interno e da avaliação de riscos disponível ou quando, apesar de existir, precisa passar por revisão muito expressiva (total ou quase total); necessidade média, quando o mapeamento do controle interno e da avaliação de riscos disponível precisa ser revisado de forma expressiva, embora aproveite parte das informações acumuladas; e necessidade baixa, quando o mapeamento do controle interno e da avaliação de riscos disponível não precisa ser alterado substancialmente – revisão pouco expressiva.



A atribuição de pesos a esses conceitos sobre a necessidade de revisão do controle interno e da avaliação de riscos observa, para efeitos do presente estudo, o disposto na Tabela 4.

Tabela 4: Parâmetros para os pesos relativos à necessidade de revisão do controle interno e avaliação de riscos

Fator/Nível	Baixo	Médio	Alto
CI/AR	0,95	1,00	1,05

3.5 FATOR DE AJUSTE GERENCIAL (F_{AG})

A última dimensão de ajuste do modelo considera que, não obstante a intensão de dar objetividade aos mecanismos para definição do tempo necessário para a realização dos trabalhos de auditoria, outros fatores, de natureza subjetiva, podem e devem ser considerados, o que é compatível com os requerimentos do gerenciamento da AI, prevista pelo IIA nas Normas profissionais (IIA, 2017).

Para esse fim, o modelo prevê a variável FAG, que deve ser utilizada para registrar decisão gerencial baseada em lastro empírico e/ou para fixar meta temporal (cumprimento de prazos) a ser atingida pela equipe responsável pela auditoria. O gestor deve também considerar outras variáveis como, por exemplo, a possibilidade de estabelecer maior (menor) foco ou profundidade na análise de determinados pontos do escopo, o que implicará a necessidade de se conceder maior ou menor tempo para a realização das etapas do processo a ser auditado. Por fim, esse fator de ajuste pode ser utilizado também para adaptação dos recursos da área de auditoria às necessidades de trabalhos a serem desenvolvidos durante determinado período – nesse caso, pode-se assumir algum grau de restrição de abrangência, dada as condições impostas à área de auditoria interna.

Combinando esses diversos aspectos, o gestor atribui valor ao fator de ajuste gerencial, utilizando uma escala que considere sua percepção sobre a conveniência de ampliar ou reduzir o tempo previsto para a realização do trabalho. Se concluir que não há situação que justifique ajuste gerencial, atribui valor neutro (1,0) à variável. No presente estudo a escala para o ajuste gerencial é sintetizada na Tabela 5.

Tabela 5: Escala para o fator de ajuste gerencial

Fator/Nível	Redução	Neutralidade	Ampliação
Ajuste gerencial	0,95	1,00	1,05

4. MODELO PARA AVALIAÇÃO DA EFICIÊNCIA DA AUDITORIA – ÊNFASE NA DIMENSÃO TEMPORAL

Definido o modelo de estimação do tempo para realização de uma auditoria, o próximo passo consiste em definir um método para se avaliar a eficiência das equipes no desenvolvimento



dos trabalhos, com ênfase na dimensão do tempo dispendido para tal. A ressalva de que essa avaliação se concentra na dimensão temporal é importante para destacar que, nesse caso, contempla um dos aspectos a serem analisados, tendo em vista que a avaliação da eficácia da atuação da AI e de sua performance, em termos mais amplos, se concentra no exame do alinhamento aos padrões de práticas profissionais de auditoria e a verificação sobre se as atividades desenvolvidas agregam valor à entidade como um todo (MIHRET; et al., 2010).

Ressalvado esse aspecto, o modelo proporciona uma avaliação, sob a ótica da dimensão temporal, da eficiência do processo de trabalho da AI. Nesse sentido, configura-se em uma ferramenta de gestão para medir o desempenho na realização do processo de trabalho. Por outro ângulo, torna-se uma proposta de indicador de eficiência para se aliar ao modelo tradicional de indicador de resultado ou de eficácia, do tipo 0 (não fez) ou 1 (fez) ou do tipo percentual realizado.

O uso desse modelo de avaliação atende aos preceitos estabelecidos nas orientações definidas pelo IIA, em particular a Norma 1131 (IIA, 2017), que prevê autoavaliações periódicas da AI, visando sua melhoria contínua, e envolve o monitoramento de sua eficiência e eficácia. Considerando a afirmação de O'Loughlin e Swauger (2015) de que as autoavaliações periódicas devem focar em avaliar o cumprimento dos padrões/indicadores de desempenho, o modelo sugerido pode ser utilizado como uma ferramenta de autoavaliação periódica do desempenho na realização de processos auditáveis pela AI.

Nesse sentido, o modelo proposto, sintetizado na Equação (4.1), considera não apenas a diferença entre os tempos efetivo e estimado para cada trabalho, mas também as ponderações do processo auditado em relação ao conjunto dos processos e das fases do trabalho realizado.

$$N_{EF} = \left[\left[\sum_{i=1}^n \sum_{j=1}^3 P_i P_{ij} \left[1 - \frac{T_{EFij} - T_{ESij}}{T_{ESij}} \right] \right] - 1 \right] \quad (4.1)$$

Onde:

N_{EF} nível de eficiência dos trabalhos de auditoria desenvolvidos.

n número de processos auditados em um período anual.

j fase do processo auditado ($j = 1$ - planejamento, $j = 2$ - execução, $j = 3$ - comunicação/ encerramento).

P_i peso do i -ésimo processo auditado no ranking do conjunto dos n processos auditados.

P_{ij} peso da j -ésima fase do i -ésimo processo auditado.

T_{EFij} tempo efetivo da j -ésima fase do i -ésimo processo auditado.

T_{ESij} tempo esperado da j -ésima fase do i -ésimo processo auditado, apurado conforme modelo (3.1).



Para aplicação do modelo, torna-se primordial a fixação dos pesos dos processos e das fases dos processos auditados (P_i e P_{ij} , respectivamente), assumindo as seguintes premissas:

$$\sum_{i=1}^n P_i = 1 \quad (4.2)$$

$$\sum_{j=1}^3 P_{ij} = 1, \text{ para todo } i = 1, \dots, n. \quad (4.3)$$

4.1 PESO DO PROCESSO AUDITADO (P_i)

Relativamente ao peso atribuível aos processos, uma sugestão é assumir $P_i = 1/n$, para todo $i = 1, \dots, n$, onde n é o número de processos auditados efetuados em um período anual. Nesse caso, opta-se por atribuir pesos equivalentes a todos os processos auditáveis, o que representa, na prática, não atribuição de pesos diferentes entre os processos. Outra opção para P_i é atribuir valores de acordo com o nível de priorização do processo auditado – estratégico, tático ou operacional – observando a condição exposta na premissa (4.2).

$$P_i = \frac{P_{P_i}}{\sum_{i=1}^n P_{P_k}} \quad (4.4)$$

Onde:

P_{P_k} peso da priorização do k -ésimo processo auditado.

Optando-se por essa segunda alternativa de atribuição de valor a P_i , a questão reside em como definir o peso da priorização de cada processo auditado, considerando sua classificação. No presente estudo é utilizada a escala disposta na Tabela 6, sujeita a ajustes e aprimoramentos, a partir da experiência dos gestores da área de auditoria.

Tabela 6: Parâmetros para os pesos relativos ao ranqueamento de cada processo auditado

Classificação	Estratégico	Tático	Operacional
P_{P_k}	5	3	1

Essa alternativa permite estabelecer ranking entre os processos auditados num determinado ciclo de auditoria, possibilitando confrontar a priorização de um processo específico frente aos demais auditados no mesmo ciclo. Desse modo, é possível mensurar o impacto individual da priorização no conjunto dos processos auditados no ciclo, favorecendo o estabelecimento de priorização dentre os processos relativos de mesmo nível (estratégico, tático ou operacional), diferenciando-os quanto à contribuição para o cálculo do nível de efetividade dos trabalhos de auditoria desenvolvidos.

Além da classificação de cada processo auditado nos níveis, a construção do ranking pode considerar outros fatores tais como: oportunidade, conveniência e/ou necessidade de auditar



processo específico naquele ciclo. Assim, por exemplo, um processo de nível tático pode ter priorização superior a alguns processos estratégicos auditados no ciclo.

Para ilustrar, suponha-se que no ciclo foram auditados 9 (nove) processos: 3 (três) operacionais, 3 (três) táticos e 3 (três) estratégicos. Estabelece-se, então, o seguinte ranking (pesos) todos os processos operacionais receberão peso 10; dos processos táticos, dois receberão peso 25 e o outro com peso 40; e dos processos estratégicos, um receberá peso 30 e os outros dois, peso 50.

Ao se aplicar a fórmula (4.4), observa-se que a premissa (4.2) está naturalmente satisfeita.

4.2 PESO DAS FASES DO PROCESSO AUDITADO (PIJ)

Em relação às fases de um determinado processo auditado, uma possibilidade é, seguindo a cronologia, adotar os valores 0,45, 0,35 e 0,20. Esse raciocínio pondera, intrinsecamente, que a fase de planejamento é a mais importante, uma vez que, sendo bem dimensionada temporalmente, haverá menos divergência entre os respectivos tempos esperado e efetivo. Ainda, se bem elaborado o planejamento, as demais fases deverão ser mais eficientemente realizadas. Consequentemente, haverá tendência de ocorrer menos disparidades entre os tempos esperado e efetivo das outras duas fases.

Entretanto, caso tenha sido efetuada, na fase de planejamento, mera revisão do Questionário de Avaliação do Controle Interno (QACI) e/ou do Mapa de Objetivos, Riscos e Controles (MORC) provenientes de auditoria anterior, que não implicou alterações consideravelmente relevantes, os percentuais descritos acima devem ser reponderados. Uma sugestão é utilizar 0,20, 0,55 e 0,25.

Além disso, a adoção de roteiros pré-elaborados na fase de planejamento pode contribuir significativamente para a redução de prazos dessa etapa. Isso deve ser objeto de reflexão no momento de delimitação do prazo esperado para a fase de planejamento e também na oportunidade de definição dos pesos para cada uma das fases de um determinado processo auditado. Pode-se optar pelos mesmos pesos indicados no parágrafo anterior.

É importante destacar que, durante a realização de um processo auditável, a influência de fatores exógenos ao componente de auditoria, inerentes a toda e qualquer organização, aumenta com a evolução das fases. Essa influência pode implicar o acréscimo dos tempos efetivos e a necessidade de melhor ponderar os pesos das três fases. Um exemplo de fator exógeno é o prolongamento inesperado da discussão do relatório preliminar com a área auditada, evento que ocorre na fase de comunicação (encerramento).

Para mitigar o risco da influência dos fatores exógenos na fase de comunicação (encerramento) pode-se, tanto para o tempo esperado quanto para o tempo efetivo, adotar como marco final a entrega da minuta do relatório de auditoria ao supervisor. Os pesos relacionados às fases também podem ser utilizados para, em razão de imprevistos, efetuar as devidas correções.



É interessante que os tempos esperado e efetivo de uma fase do processo auditável sejam expressos em horas e não em dias. O tempo efetivo de uma fase deve ser o resultado do somatório das quantidades de horas dispendidas pelos auditores envolvidos na respectiva fase do processo auditado. Essa medida proporciona um melhor ajuste do modelo à realidade cotidiana, haja vista que um mesmo auditor pode, por algum intervalo de tempo, participar simultaneamente em duas (ou mais) fases de trabalhos distintos de auditoria, como, por exemplo, encerramento de um e planejamento de outro. Assim, de modo a manter a coerência na unidade de medida, o tempo esperado também deve ser expresso em horas.

4.3 INTERPRETAÇÃO DOS VALORES DO NÍVEL DE EFICIÊNCIA (NEF)

Definido e mensurado o nível de eficiência da auditoria interna, o valor do NEF representa o número de vezes, em relação ao tempo estimado, que foi economizado (ou gasto a mais) na realização do(s) trabalho(s) de auditoria. De forma sintética, os parâmetros de interpretação observam o disposto no Quadro 1.

Quadro 1: Parâmetros para a interpretação do Nível de Eficiência dos trabalhos desenvolvidos

Valor	Interpretação
$N_{EF} > 0$	Nível de eficiência da equipe auditoria interna supera o planejado, já considerando o ajuste quanto ao ranqueamento dos processos auditados
$N_{EF} = 0$	Nível de eficiência da equipe auditoria interna é exatamente igual ao planejado, já considerando o ajuste quanto à importância dos processos auditados
$N_{EF} < 0$	Nível de eficiência da equipe auditoria interna é inferior ao planejado, já considerando o ajuste quanto ao ranqueamento dos processos auditados

Como se percebe, quanto mais próximo de 0 (zero), maior o grau de ajustamento entre o tempo efetivo na realização das auditorias e o planejado. Por outro lado, quanto maiores (menores) os valores positivos (negativos), maior o grau de eficiência na realização dos trabalhos em relação ao planejado, considerando o ajustamento pelo grau de importância do processo para a organização. Um valor de 0,17 para o NEF de determinado processo, por exemplo, significa que o tempo efetivamente gasto na realização da auditoria representou uma economia de 17% do tempo estimado, considerando o ajustamento pela importância do processo para a organização. Se esse valor fosse de -0,17 significaria que a equipe teria consumido de 17% a mais de tempo, em relação ao estimado, na realização do trabalho.

Esse indicador de nível de eficiência (NEF) pode ser aplicado tanto a cada trabalho realizado, quanto aos conceitos de equipes, divisões ou ao conjunto da área de auditoria interna. Os preceitos do modelo (4.1) são aplicáveis a todas as dimensões de análise que se pretenda realizar.



5. CONCLUSÕES

O presente estudo teve por objetivo o desenvolvimento de modelo matemático para estimação e controle do tempo necessário para a realização da atividade de AI, com o propósito de otimizar as ferramentas de planejamento, que auxiliem na mensuração do tempo de realização das auditorias.

Utilizando premissas conceituais e normativas, foi desenvolvido modelo que se baseia no histórico dos trabalhos de auditoria realizados no processo auditável, com ajustes em relação: à complexidade, frequência e importância relativa do processo a ser auditado; à maturidade, conhecimento e habilidade da equipe de auditores; à revisão preliminar do controle interno e avaliação de riscos; além da previsão de um fator de ajuste gerencial discricionário, de forma a contemplar a subjetividade implícita ao processo de gestão da atividade de AI. Na segunda etapa, foi desenvolvido modelo adicional com o propósito de propiciar a avaliação, em relação aos aspectos temporais, das atividades realizadas pela AI. O modelo proposto se baseia na diferença entre os tempos estimado e efetivo de cada trabalho, com ajustes por meio de ponderações do processo auditado em relação ao conjunto dos processos e das fases do trabalho realizado.

Para se concluir sobre a pertinência do modelo proposto, simulações foram realizadas com base em informações da área de auditoria de uma autarquia federal. Em seguida, seus resultados foram discutidos com gestores da área, concluindo-se que os outputs do modelo proposto geram informações relevantes e intuitivamente adequadas e úteis ao processo de gestão da atividade.

O estudo contribui para o desenvolvimento da literatura sobre AI, notadamente quanto à instituição de instrumentos gerenciais que permitam a avaliação da eficiência dos trabalhos de auditoria realizados. Isso é particularmente relevante tendo em vista a incipiência, se não inexistência, de trabalhos com esse foco. O principal benefício da adoção de um modelo dessa natureza é proporcionar a melhor alocação dos recursos humanos e materiais, minimizando desperdícios e ineficiências, bem como promovendo entregas mais robustas e qualificadas à alta administração da companhia à qual se vincula a AI. Além disso, a utilização do tempo de auditoria como parâmetro de performance e desempenho permite maior objetividade ao processo de autoavaliação periódica que é requerido pelas normas de auditoria.

Como principais limitações, há que se ressaltar a subjetividade implícita na definição dos parâmetros utilizados no modelo, o que requer ajustes em resposta a situações específicas verificadas em cada contexto. Por essa razão, é importante que o modelo(s) proposto(s) seja(m) testado(s), no sentido de identificar oportunidades de ajustes e aprimoramentos.

REFERÊNCIAS

ABDHEL-KHALIK, A. R; SNOWBALL, D.; WRAGGE, J. H. The Effects of Certain Internal Audit Variables on the Planning of External Audit Programs. *The Accounting Review*, v. 58, n. 2, p. 215-227, 1983.

ALMEIDA, M. C. Auditoria. 9. ed. São Paulo: Atlas, 2017.

ATTIE, W. Auditoria Interna. 2 ed. – São Paulo: Atlas, 2007.



AZAD, A.N. Time Budget Pressure and Filtering of Time Practices in Internal Auditing: A Survey, *Managerial Auditing Journal*, v. 9, n. 6, p. 17-25, 1994.

CONTROLADORIA GERAL DA UNIÃO (CGU). Instrução Normativa Conjunta MP/CGU nº 01, de 10 de maio de 2016. Dispõe sobre controles internos, gestão de riscos e governança no âmbito do Poder Executivo federal

CONTROLADORIA GERAL DA UNIÃO (CGU). Instrução Normativa nº 3, de 09 de junho de 2017. Referencial Técnico da Atividade de Auditoria Interna Governamental do Poder Executivo Federal. / Ministério da Transparência e Controladoria-Geral da União, Secretaria Federal de Controle Interno - Brasília: CGU, 2017.

CONTROLADORIA GERAL DA UNIÃO (CGU). Instrução Normativa nº 8, de 08 de dezembro de 2017. Manual de orientações técnicas da atividade de auditoria interna governamental do Poder Executivo Federal / Ministério da Transparência e Controladoria-Geral da União, Secretaria Federal de Controle Interno - Brasília: CGU, 2017.

CONTROLADORIA GERAL DA UNIÃO (CGU). Instrução Normativa nº 9, de 09 de outubro de 2018. Dispõe sobre o Plano Anual de Auditoria Interna (PAINT) e sobre o Relatório Anual de Atividades da Auditoria Interna (RAINT) das Unidades de Auditoria Interna Governamental do Poder Executivo Federal.

COMMITTEE OF SPONSORING ORGANIZATIONS OF THE TREADWAY COMMISSION (COSO). Gerenciamento de Riscos Corporativos – Estrutura Integrada. 2007. Disponível em: <<https://www.coso.org/Documents/COSO-ERM-Executive-Summary-Portuguese.pdf>>. Acesso em: 14 Abr. 2019.

COMMITTEE OF SPONSORING ORGANIZATIONS OF THE TREADWAY COMMISSION (COSO). Leveraging COSO across the three lines of defense. 2015. Disponível em: <<https://na.theiia.org/standards-guidance/Public%20Documents/2015-Leveraging-COSO-3LOD.pdf>>. Acesso em: 14 Abr. 2019.

DIAS, S. V. S. Auditoria de Processos Organizacionais. 4 ed. São Paulo: Atlas, 2015.

ENDAYA, K.A.; HANEFAH, M. M. Internal auditor characteristics, internal audit effectiveness, and moderating effect of senior management. *Journal of Economic and Administrative Sciences*, v. 32, n. 2, p. 160-176, 2016.

GAUMNITZ, B. R.; NUNAMAKER, T. R.; SURDICK, J. J.; THOMAS, M. F. Auditor Consensus in Internal Control Evaluation and Audit Program Planning. *Journal of Accounting Research*, v. 20, n. 2, p. 745-755, 1982.

HUGHES, J. S. Optimal Internal Audit Timing. *The Accounting Review*, v. 52, n. 1, p. 55-68, 1977.

KESUMA, I. B. G. P. W.; DWIRANDRA, A. A. N. B. Professional Commitments and Pressure of Obedience in Mediating on the Effect of Time Budget Pressure in Quality Audits. *International Research Journal of Engineering, IT & Scientific Research*, v. 5, n. 1, p. 27-38, 2019.



MAFFEI, J. Auditoria Interna: melhores práticas. São Paulo: All Print, 2011.

MINISTÉRIO DA TRANSPARÊNCIA E CONTROLADORIA-GERAL DA UNIÃO (MT-CGU). Manual de Orientações Técnicas da Atividade de Auditoria Interna Governamental do Poder Executivo Federal. Ministério da Transparência e Controladoria-Geral da União. Brasília: CGU, 2017.

MELLO, I. R.; BARBOSA, K. M. F.; DANTAS, J. A.; BOTELHO, D. R. 25 anos de publicação em auditoria: análise bibliométrica sob o ponto de vista da Lei de Lotka, Lei de Zipf e Ponto de Transição (T) de Goffman. Revista de Estudos Contábeis, vol. 8, n. 15, p. 45-65, jul./dez. 2017.

MIHRET, D. G.; JAMES, K.; MULA, J. M.; THOMAS, M. F. Antecedents and organizational performance implications of internal audit effectiveness: Some propositions and research agenda. Pacific Accounting Research, v. 22, n. 3, p. 224-252, 2010.

O'LOUGHLIN, C. J.; SWAUGER, J. Garantia de Qualidade e Melhoria da Auditoria Interna. Lake Mary: IIA/CBOK, 2015.

SCHNEIDER, A. Modeling External Auditors' Evaluations of Internal Auditing. Journal of Accounting Research, v. 22, n. 2, p. 657-678, 1984.

SOH, D. S. B.; MARTINOV BENNIE, N. The internal audit function: Perceptions of internal audit roles, effectiveness and evaluation. Managerial Auditing Journal, v. 26, n. 7, p. 605-622, 2011.

SPIRA, L. F.; PAGE, M. Risk management: The reinvention of internal control and the changing role of internal audit. Accounting, Auditing & Accountability Journal, v. 16, n. 4, p. 640-661, 2003.

STEWART, J.; SUBRAMANIAM, N. Internal audit independence and objectivity: emerging research opportunities. Managerial Auditing Journal, v. 25, n. 4, p. 328-360, 2010.

SUMITO, N.; SETIYAWATI, H. The Influence of Competence, Independence, Professionalism of Internal Auditors and Time Budget Pressure on Audit Quality. Scholar Bulletin, v. 5, n. 9, p. 496-503, sept. 2019.

THE INSTITUTE OF INTERNAL AUDITORS (IIA). International Standards for the Professional Practice of Internal Auditing (Standards). 2017. Disponível em: <<https://na.theiia.org/standards-guidance/public%20documents/ippf-standards-2017.pdf>>. Acesso em: 14 Abr. 2019.

THE INSTITUTE OF INTERNAL AUDITORS (IIA). Modelo das Três Linhas do IIA 2020: Uma atualização das Três Linhas de Defesa. 2020. Disponível em: <<https://na.theiia.org/translations/PublicDocuments/Three-Lines-Model-Updated-Portuguese.pdf>>. Acesso em: 29 Ago 2020.

Os conceitos e interpretações emitidos nos trabalhos assinados são de exclusiva responsabilidade de seus autores.





Rastro-DM: Mineração de Dados com Rastro

Metodologia para documentação de projetos de mineração de dados e sua aplicação na construção de um classificador textual de documentos associados a danos ao erário público

MARCUS VINÍCIUS BORELA DE CASTRO¹

Auditor do TCU desde 1996. Bacharel em Informática pela Universidade Federal de Viçosa (1990) e especialista em Governança de TI pela Universidade de Brasília (2012) e em Análise de Dados pela Escola Superior do TCU (2019).

REMIS BALANIUK

Auditor do TCU desde 1989. Graduado em Ciência da Computação pela Unb(1986), mestre em Ciência da Computação pela UFRGS (1989), doutor em Informática pelo Institut National Polytechnique de Grenoble (1996) e pós-doutor em Computação pela Stanford University (2002), pelo Institut National pour la Recherche en Informatique et Automatique (2000) e pesquisador visitante da University of Oxford (2020)

RESUMO

Este trabalho propõe uma metodologia de documentação de projetos de mineração de dados (DM), Rastro-DM, com foco não no modelo gerado, mas no processo por trás de sua construção, de forma a deixar um rastro das ações planejadas, dos treinamentos realizados, dos resultados obtidos e dos aprendizados concebidos. As práticas propostas são complementares às metodologias estruturantes de DM, como o CRISP-DM, que trazem todo o arcabouço metodológico e paradigmático para o processo de DM. Ilustra-se o seu uso e seus benefícios em um projeto de classificação textual de documentos em PDF associados a danos ao Erário Público Federal Brasileiro denominado Cladop. A produção do Rastro-DM de um projeto é um pequeno passo que pode levar a um salto organizacional, a ser obtido com a partilha e o uso dos rastros de forma corporativa.

¹ Trata-se de uma versão sintetizada do trabalho de conclusão de curso de pós-graduação lato sensu em Análise de Dados indicado na referência (Castro, 2019).



Palavras-chave: Mineração de dados; Análise de dados; Ciência de dados; Aprendizado de máquina; Conhecimento organizacional; Metodologia; Boas práticas; Análise de dados no Governo; Classificação textual; Documentação; Documentação de projetos de mineração de dados.

1. INTRODUÇÃO

Os dados estão mudando tudo e a capacidade de manipulá-los e entender Ciência de Dados está se tornando cada vez mais crítica para atuais e futuras descobertas e inovações (BERMAN *et al*, 2018).

Projetos de mineração de dados (DM) são desafiadores não só pelo complexo processo usado, exploratório, mas também por, em geral, serem inovadores, únicos e muitas vezes desenvolvidos por indivíduos ou pequenas equipes.

Trata-se de projetos inovadores, quer por usarem técnicas e algoritmos que podem não estar consolidadas ou na Organização ou em pesquisas acadêmicas, quer por envolverem a construção de modelos que simulam processos cognitivos, inteligência natural, por máquinas.

São projetos quase sempre únicos. As particularidades de cada contexto, dos dados envolvidos, dos requisitos de qualidade, impedem ou dificultam o reaproveitamento de código para outros projetos.

São projetos complexos pois as técnicas empregadas em geral têm concepção de difícil entendimento e envolvem conhecimento interdisciplinar de áreas como ciência da computação, matemática e estatística, além do entendimento do negócio para o qual a solução se destina.

São processos exploratórios, pois a atividade de mineração de dados pode ser definida como o processo de explorar um conjunto de dados com técnicas diversas extraindo ou ajudando a evidenciar padrões e auxiliando na descoberta de conhecimento.

E, no caso de organizações com baixo grau de maturidade em DM, são projetos que ficam sob a responsabilidade de pequenas equipes ou mesmo de um único indivíduo. Nesse caso, a partilha do conhecimento e de práticas adotadas nos projetos fica ainda mais difícil.

Infelizmente, esses trabalhos complexos, de caráter exploratório, inovadores, únicos, e, em geral, individuais, não deixam rastro do que fazem. Ao final temos a solução implementada. Pode-se ter até uma documentação do produto criado, mas não do processo seguido, das escolhas feitas e das técnicas usadas nas diversas atividades do projeto. Resta um temor no ar para o caso de ser necessário reconstruir o modelo sem a presença do seu criador. E o responsável se torna pai do produto, pois só ele o conhece e pode dar manutenção.

Uma questão intrigante na gestão do conhecimento em geral é como coletar, colher ou tornar explícita a experiência de projetos para que possam ser utilizáveis para outros (DINGSØYR *et al*, 2001). Pois a memória de uma organização não pode se basear apenas na memória de seus indivíduos (STATA, 1980).



Diante do exposto surge a desafiadora questão: como sistematizar a documentação das tarefas de um projeto de mineração de dados de forma a potencializar sua auditabilidade e a partilha dos aprendizados?

Buscando responder à questão colocada, o objetivo deste trabalho é propor uma metodologia enquanto conjunto de boas práticas de registro semi-automatizado das atividades de um projeto de DM de forma a deixar um rastro das escolhas feitas, dos processamentos realizados e dos resultados obtidos, com foco não no produto gerado, mas no processo por trás de sua construção. Boas práticas que podem ser mescladas à metodologia corporativa de DM em uso na organização. A produção de rastros em projetos de DM acelera a curva de aprendizagem e a formação de uma cultura organizacional em torno do uso da análise de dados.

São três os objetivos específicos: contextualização do referencial teórico sobre metodologias e documentação em projetos de mineração de dados bem como do potencial impacto das experiências adquiridas nos projetos para uma organização; proposição do Rastro-DM com a descrição de suas atividades; e a ilustração de sua aplicação em um projeto de classificação textual de documentos em PDF associados a danos ao Erário Público Federal Brasileiro denominado Cladop, demonstrando sua viabilidade e os benefícios de seu uso. Os objetivos específicos serão tratados nas três seções do desenvolvimento.

2. DESENVOLVIMENTO

2.1 REFERENCIAL TEÓRICO

2.1.1 METODOLOGIAS DE MINERAÇÃO DE DADOS

Segundo BERMAN *et al* (2018), a Ciência de Dados se concentra nos processos de extração de conhecimento ou *insights* a partir de dados estruturados ou não. Esse processo de descoberta de conhecimento em dados (KDD - *Knowledge-Discovery in Databases*), segundo BECKER e GHEDINI (2005), é complexo e popularmente chamado de Mineração de Dados (DM - *Data Mining*). No escopo deste trabalho chamaremos indistintamente KDD de DM.

CHAPMAN *et al* (2000) detalham alguns objetivos de DM traduzidos em grupos de problemas tratados: descrição e sumarização de dados, segmentação (clusterização), descrição de conceitos, classificação, predição (regressão) e análise de dependência.

WIRTH e HIPPE (2000) confirmam que DM é um processo complexo e associam o sucesso de um projeto a uma adequada combinação de boas ferramentas, analistas qualificados, uso de uma metodologia sólida e um eficaz gerenciamento de projetos. Quanto à metodologia, os mesmos autores afirmam que DM precisa de uma abordagem padrão que ajude a transformar problemas de negócios em tarefas de mineração de dados, que sugira transformações de dados apropriadas e técnicas a empregar, e forneça meios para avaliar a eficácia dos resultados

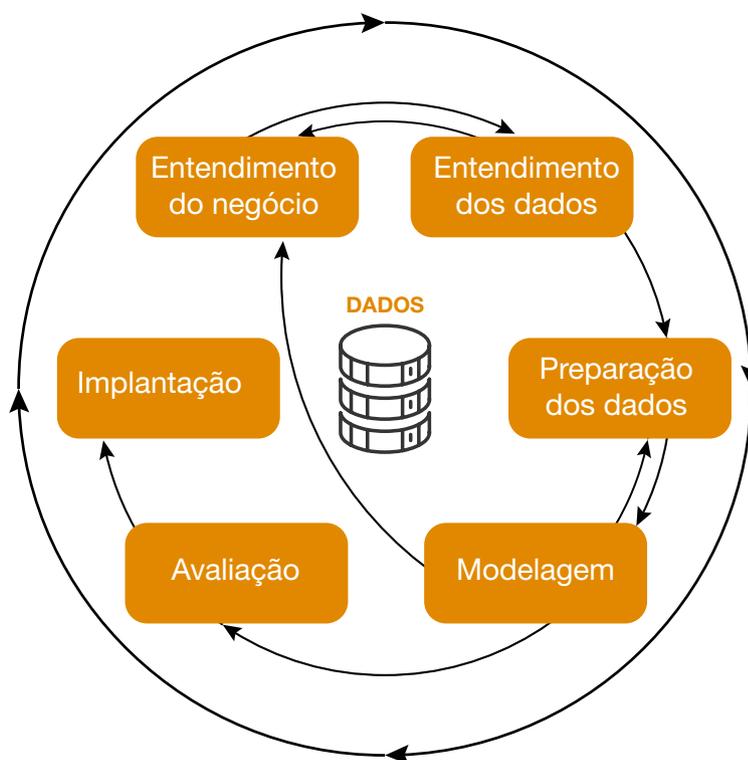
e documentar a experiência. E destacam que o uso de uma metodologia no planejamento e na apresentação de relatórios inspira confiança nos usuários e nos patrocinadores.

KURGAN e MUSILEK (2006) afirmam que metodologias acrescentam uma melhor compreensão e entendimento do processo, além de promoverem economia de tempo e de custos, com o estabelecimento de um roteiro a seguir para o planejamento e a execução dos projetos. Mas alertam que um grande número de projetos segue metodologias próprias.

Um exemplo de metodologia usada em DM é o CRISP-DM (*Cross-Industry Standard Process for Data Mining*). Ela é considerada padrão de fato e um dos fatores de seu sucesso é ser neutra em termos de indústria, ferramentas e aplicações (MARISCAL *et al*, 2010).

CHAPMAN *et al* (2000) atestam que a metodologia CRISP-DM é um modelo de processo hierárquico, consistindo em conjuntos de tarefas descritas em quatro níveis de abstração (do geral ao específico): fase, tarefa genérica, tarefa especializada e instância do processo. Afirmam que a sequência das fases não é rígida. Que, na prática, muitas das tarefas podem ser executadas em uma ordem diferente, e, muitas vezes, faz-se necessário voltar repetidamente para tarefas anteriores e repetir certas ações. Ressaltam que representar todas as rotas possíveis por meio do processo de mineração de dados exigiria um modelo de processo excessivamente complexo, por isso não o fazem. Afirmam que nunca uma fase é completamente concluída antes que a fase seguinte comece. A Figura 1 mostra a iteração entre as fases do CRISP-DM.

Figura 1 - Visão geral do CRISP-DM.



Fonte: CHAPMAN *et al* (2000).



Segundo MARBÁN *et al* (2007), o CRISP-DM não cobre muitas tarefas relacionadas à gestão, organização e qualidade de projetos. Pelo menos não da maneira exigida pela crescente complexidade dos recentes projetos de DM que envolvem não apenas grandes volumes de dados, mas também o gerenciamento e a organização de grandes times interdisciplinares. Os mesmos autores agrupam as tarefas de um processo de DM, em relação à construção de um modelo, em três estágios: pré-desenvolvimento, desenvolvimento e pós-desenvolvimento. Afirmam que todas as metodologias usadas para DM se concentram no estágio de desenvolvimento, que equivale a coleta e a análise dos dados disponíveis para o projeto, a criação de novos dados a partir dos disponíveis, a adaptação para algoritmos de DM e a criação de modelos.

Segundo CHOLLET (2017), o desenvolvimento de um projeto está centrado em experimentações de um modelo: você começa com uma ideia e a expressa como um experimento, tentando validar ou invalidar sua ideia. Depois, executa-se essa experiência e processam-se as informações geradas. Isso inspira sua próxima ideia. Quanto mais iterações desse círculo repetitivo você conseguir executar, mais refinadas e poderosas suas ideias se tornarão. Ressalta a importância de se obter o máximo possível de informações das experimentações, incluindo o desempenho dos modelos.

A experimentação, ou seja, a aplicação de algoritmos matemáticos aos dados para a extração de padrões, é chamada de treinamento (training) por NGUYEN (2018). Para simplificar, no contexto deste trabalho, serão usados os termos treinamento e modelo representando a experimentação e o produto resultante dos padrões detectados nos dados.

GREFF *et al* (2017) confirmam o número significativo de experimentos computacionais com muitas configurações diferentes de hiperparâmetros e alertam sobre o desafio prático da documentação. Devido à pressão de prazo e à inerente natureza imprevisível de um projeto de DM, constatam que há pouco incentivo para a construção de infraestruturas robustas e, como resultado, o código muitas vezes evolui rapidamente, o que compromete, entre outras coisas, a documentação do projeto.

BECKER e GHEDINI (2005) acentuam que a estrutura de um processo de DM é altamente dependente da metodologia adotada, das habilidades, da experiência e do estilo da pessoa responsável pelo processo, bem como dos recursos disponíveis na corporação. E confirmam a alta iteratividade e interatividade dos processos. Ressaltam que embora a estrutura conceitual do processo sugira uma ordem entre as fases, na prática, os analistas passam de qualquer fase para quase qualquer outra fase a qualquer momento, até porque muitos problemas relacionados a fases anteriores (por exemplo, preparação de dados) só podem ser detectados muito mais tarde, quando os padrões e os modelos são avaliados. As mesmas autoras afirmam que os projetos são desenvolvidos em geral de maneira não estruturada, ad hoc: após a análise inicial dos dados, decide-se experimentar uma determinada técnica, cujos resultados podem sugerir a reestruturação dos dados e a execução de novos tipos de análises; e assim por diante. E, à medida que o tempo passa, não se pode lembrar quais treinamentos foram realizados, os conjuntos de dados utilizados, os hiperparâmetros usados e, mais importante, os resultados que



foram derivados dos conjuntos de dados. E essa situação, segundo as autoras, leva à reexecução de treinamentos. Alertam que a situação é ainda pior se forem considerados projetos de longo prazo envolvendo várias pessoas. Constatam que, apesar da diversidade de conhecimento das pessoas, das técnicas e das ferramentas, a maioria dos projetos de DM enfrenta, na prática, as mesmas dificuldades: o desperdício de se refazer um trabalho e o gerenciamento de recursos e de resultados. Afirmam as autoras que a documentação do histórico das tarefas face a iteratividade e a interatividade do processo é um problema aberto no gerenciamento de projetos de DM.

2.2 DOCUMENTAÇÃO: CAMINHO PARA A GERAÇÃO DE CONHECIMENTO PARTILHÁVEL

GHEDINI e BECKER (2000) ressaltam que a documentação de experimentos e de todas as partes relevantes de um projeto não só evita a perda de conhecimento confinado nas mentes das pessoas como também permite o seu compartilhamento, tornando-se uma rica fonte de conhecimento para referência futura e reuso corporativo. Destacam também que essa atividade leva a um melhor gerenciamento de esforços, de recursos e de resultados de um projeto de DM.

Segundo PRAKASH *et al* (2012), informações de treinamentos e código de implementação (e seu histórico de mudança) contêm uma riqueza de informações sobre o estado, o progresso e a evolução de um projeto de software. Afirmam também que DM está se tornando uma ferramenta cada vez mais importante para transformar esses dados em informações. Por analogia, espera-se que dados de projetos de DM sejam convertidos também em informações por atividades de mineração de dados.

WIRTH e HIPPE (2000) enfatizam que, talvez, o maior benefício de terem aplicado uma metodologia foi a documentação gerada. Admitem terem pulado inicialmente algumas tarefas de documentação e planejamento por serem demoradas e por considerarem desnecessárias para especialistas como eles. Mas apresentam o preço que pagaram por essa ação e se arrependem constatando que todo esforço vale a pena. Relatam alguns benefícios observados a partir da documentação produzida: evita o desperdício de esforço (por exemplo, em caminhos não frutíferos ou com trabalho repetitivo); promove um gerenciamento eficaz e uma melhor comunicação da equipe; permite a identificação de pontos críticos no processo; promove um melhor planejamento de projetos futuros, com base em uma melhor percepção de como o esforço foi gasto e dos recursos necessários; promove o uso de experiências documentadas em outros contextos.

BECKER e GHEDINI (2005) também identificam o papel da documentação na aprendizagem e na reutilização e afirmam que um benefício imediato da documentação é a efetividade no gerenciamento, no planejamento e na comunicação. Constatam que a documentação é completamente dependente da equipe do projeto, uma vez que está em sua responsabilidade a veracidade dos registros e o nível de detalhe que impacta diretamente sua utilidade. Quanto à resistência, afirmam que, à medida que os benefícios da atividade são percebidos, há um estímulo a documentar o processo com mais detalhes e de forma



concomitante com a execução das atividades e que os melhores resultados são alcançados a longo prazo, quando a equipe do projeto descobre qual estratégia melhor se adequa ao seu estilo de trabalho, bem como as melhores maneiras de obter vantagens dos recursos e das técnicas e da flexibilidade do modelo.

Contudo, WIRTH e HIPPEL (2000) alertam sobre a dificuldade de se documentar ao final, de se tentar reconstruir o que foi feito e suas motivações. Enfatizam que os processos de DM são vivos e, como tal, a documentação deve ser flexível e viva, e não deve ser atualizada após o final do projeto (post mortem). Defendem que a definição de uma estratégia de documentação deve ser um ponto de partida, mas a flexibilidade para a evolução e a mudança deve ser uma premissa. Enfatizam que encontrar o nível certo de detalhes para se planejar e se documentar um processo de DM é difícil e faz parte de um longo processo de aprendizado, e pode ser influenciado por diversos fatores como complexidade do projeto, duração e tamanho da equipe.

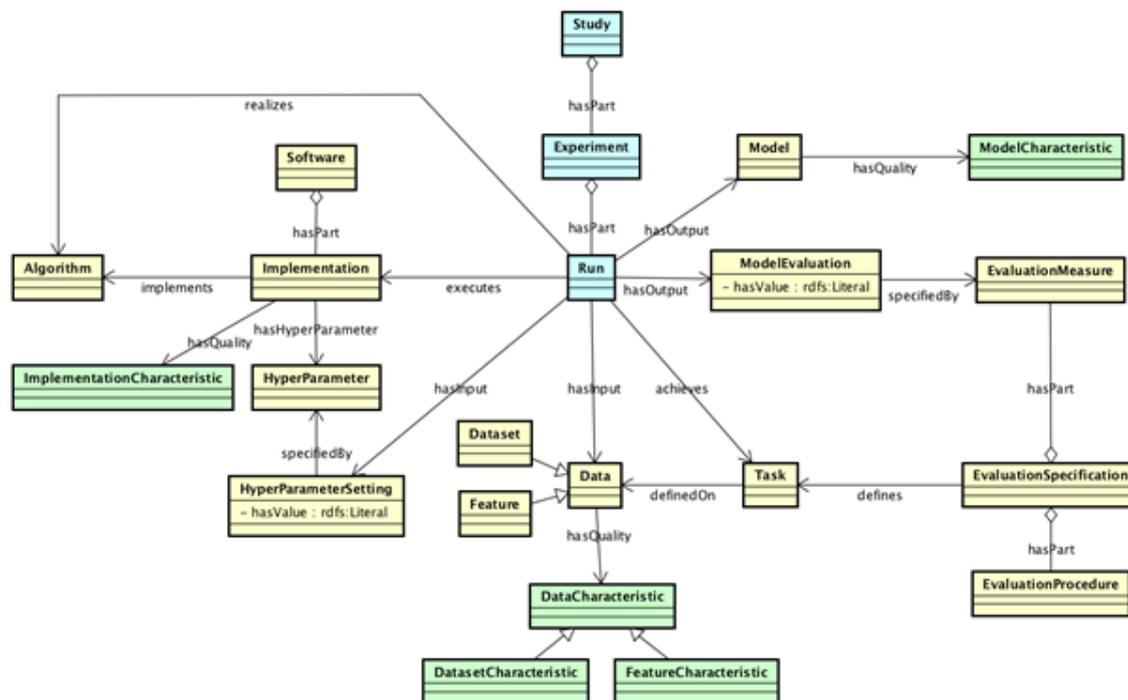
GREFF *et al* (2017) citam que uma dificuldade para o compartilhamento e a colaboração e a reprodutibilidade dos treinamentos é o uso pelas equipes de configurações particulares de área de trabalho no uso das diferentes ferramentas necessárias para tratar os aspectos distintos de um processo de DM, citando, entre outras, bancos de dados, sistemas de controle de versão, ferramentas automatizadas de otimização de hiperparâmetros, scripts e planilhas. Os mesmos autores apresentam Sacred, um framework Python de código aberto que visa fornecer infraestrutura básica para a execução de experimentos computacionais de forma independente dos métodos e bibliotecas utilizados. Concentram-se em resolver problemas como o gerenciamento de configurações, a documentação e a reprodutibilidade dos resultados. Para cada treinamento, informações relevantes, como parâmetros, dependências de pacotes, informações do host, código-fonte e resultados, são capturadas automaticamente e armazenadas em um repositório centralizado, de onde pode-se consultar também detalhes dos hiperparâmetros usados e dos resultados obtidos.

PRAKASH *et al* (2012) relatam que numerosos trabalhos estão sendo feitos no desenvolvimento de plataformas integradas para Machine Learning (ML) e para Engenharia de Software baseadas em componentes reutilizáveis, citando, entre as de código aberto mais conhecidas, WEKA e *Rapid Miner*.

PUBLIO *et al* (2018) acreditam que a visão de modelos canônicos e padronizados pode levar a uma melhor compreensão dos dados e dos algoritmos de ML empregados em DM e pode promover a interoperabilidade dos experimentos, independentemente da plataforma ou da solução de fluxo de trabalho adotada.

W3C *Machine Learning Schema Community Group* (2017) publicou uma ontologia ML *Schema* que fornece um conjunto de classes, propriedades e restrições para representar e intercambiar informações sobre algoritmos de aprendizado de máquina, conjuntos de dados e experimentos. Segundo o grupo, a ontologia pode ser facilmente estendida e mapeada para outras ontologias mais específicas de domínio desenvolvidas na área de aprendizado de máquina e de mineração de dados. A Figura 2 traz a visualização das principais abstrações do ML *Schema*.

Figura 2: Conceitos principais do ML Schema.



Fonte: W3C Machine Learning Schema Community Group (2017).

De forma resumida, o conceito central é o treinamento (Run) que, no contexto de um experimento (*Experiment*, *Study*), produz um modelo (*Model*), com suas características (*ModelCharacteristic*), e um conjunto de avaliações de qualidade (*ModelEvaluation*) que consideram métricas padronizadas (*EvaluationMeasure*, *EvaluationSpecification*, *EvaluationProcedure*). O treinamento é uma execução de uma implementação (*Implementation*) de um algoritmo (*Algorithm*) em uma plataforma (*Software*, *ImplementationCharacteristic*) usando uma configuração específica de parâmetros (*HyperParameter*, *HyperParameterSetting*). O treinamento tem como entrada um conjunto de dados (*Data*, *Dataset*, *Feature*) que têm suas características próprias (*DataCharacteristic*, *DatasetCharacteristic*, *FeatureCharacteristic*) e são utilizados para uma tarefa específica (*Task*).

KURGAN e MUSILEK (2006) afirmam a possibilidade de integração e de interoperabilidade dos modelos de DM com o uso de padrões industriais como PMML (*Predictive Model Markup Language*) que representa um modelo em um esquema XML (*Extensible Markup Language*). Segundo os autores, pode-se usar ferramentas diferentes para a geração, visualização e análise de um mesmo modelo.

2.3 CONHECIMENTO EM PROJETO DE DM: UM POTENCIAL SALTO ORGANIZACIONAL

WIRTH e HIPPEL (2000) salientam que o sucesso ou o fracasso de um projeto de mineração de dados é altamente dependente da pessoa ou da equipe e que práticas de sucesso não são necessariamente repetidas em toda a empresa.



CHAPMAN *et al* (2000) ressaltam que projetos de mineração de dados podem ser beneficiados com as experiências de projetos anteriores. As lições aprendidas durante o processo e a partir da solução implantada podem desencadear novas questões de negócios, muitas vezes mais focadas.

BECKER e GHEDINI (2005) exemplificam benefícios de se partilhar conhecimento entre projetos citando que experiências de projetos anteriores podem ser usadas para estabelecer planos de projetos de DM mais razoáveis, com estimativas mais precisas de cronogramas, orçamento, etc. A experiência torna mais fácil defender recursos mais realistas, pois há uma compreensão mais profunda de como o esforço é realmente gasto. Além disso, pode-se usar experiências anteriores para lidar com certas classes de problemas ou de técnicas. Defendem que a documentação da execução de projetos deve ser tratada como um recurso corporativo, que pode ser compartilhado pela equipe, ser usado como referência e pode estar sujeito a políticas e padrões corporativos.

Segundo BHATT (2001), à medida que os indivíduos nas organizações interagem com outros, eles tendem a entender e partilhar suas visões diferentes sobre as mesmas situações, construindo suas comunidades e compartilhando técnicas eficientes de trabalho e facilitando a integração de um corpo diversificado de conhecimentos nas organizações. O mesmo autor afirma que o conhecimento organizacional é formado por padrões únicos de interações entre tecnologias, técnicas e pessoas, que não podem ser copiados facilmente, porque essas interações são únicas da organização, moldadas pela sua história e pela sua cultura. Credita o mesmo autor a sustentação de vantagens competitivas da empresa no longo prazo ao incentivo ao crescimento desse conhecimento com a criação de um ambiente estimulante e prático (aprender-fazendo).

DINGSØYR *et al* (2001) tratam de forma indistinta conhecimento e experiência. Embora reconheçam que experiência em um sentido estrito é algo que reside nos seres humanos e que não pode ser transferido para os outros (que teriam que experimentar por si mesmo para ter a experiência), em uma definição menos estrita, afirmam que experiência é informação que é operacional, isto é, utilizável em alguma situação. Entendem que uma descrição de um evento acontecido em um projeto é um item de experiência.

Como comentado na introdução, uma questão interessante na gestão do conhecimento em geral é como coletar, colher ou tornar explícita a experiência de projetos para que possam ser utilizáveis para outros (DINGSØYR *et al*, 2001).

NGUYEN (2018) defende o uso de processos de gestão do conhecimento (KMP - *Knowledge Management Process*) para fazer circular o conhecimento em toda a organização para garantir que o conhecimento certo chegue à pessoa certa para entender e ter conhecimento suficiente para tomar decisões e executar bem as tarefas. Afirma o autor que KMP pode ser usado em qualquer nível, desde a organização como um todo até dentro da equipe. E complementa que seus estágios (identificação, criação, armazenamento, transferência e utilização de conhecimento) estão interligados e são iterativos, uma vez que o conhecimento é continuamente formado e alterado. O autor encontra na combinação entre KMP e DM um grande potencial na exploração e no gerenciamento do conhecimento valioso de big data: DM suporta o KMP na geração de conhecimento inestimável e KMP suporta DM na coleta e



armazenamento de conhecimento como entrada de DM. Mas ressalta que há um grande vazio de pesquisas nessa área.

STATA (1980) traz o conceito de aprendizado organizacional (OL - Organizational Learning), que ocorre por meio de insights compartilhados, conhecimento e modelos mentais e que se baseia no conhecimento e na experiência do passado, ou seja, na memória. O autor detalha que a memória organizacional depende de mecanismos institucionais (por exemplo, políticas e estratégias) usados para reter o conhecimento, não podendo depender exclusivamente da memória dos indivíduos, pois há sempre o risco de se perder lições e experiências duramente conquistadas à medida que as pessoas migram de um emprego para outro. Além de outros motivos de saída de pessoas (aposentadoria, afastamentos, transferências, falecimento, etc.), podemos acrescentar também o risco do esquecimento.

DINGSØYR *et al* (2001) defendem que seja elaborado ao final de um projeto um Relatório de Experiência para coletar o que deu certo e o que deu errado no processo adotado. A atividade Revisar Projeto do CRISP-DM produz relatório semelhante que conforme descrição em CHAPMAN *et al* (2000): Resuma a importante experiência adquirida durante o projeto. Por exemplo, armadilhas, abordagens enganosas ou dicas para selecionar as técnicas de mineração de dados mais adequadas em situações semelhantes poderiam fazer parte dessa documentação.

Segundo BECKER e GHEDINI (2005), com a documentação do processo de DM, à medida que o conhecimento é tornado explícito e gerenciado, ele aumenta o intelecto da organização, tornando-se uma base para a comunicação e para a aprendizagem, apoiando a disseminação de conhecimento e a experiência dentro da organização em vários níveis. As autoras reforçam que a ideia de capturar e armazenar todo o conhecimento informal relevante gerado e usado durante um processo de DM, de modo que esteja disponível para recuperação posterior, constitui uma abordagem interessante para lidar com a dificuldade citada de refletir na documentação a iteratividade e a interatividade do processo.

É fato que as empresas precisam educar um número maior de pessoas sobre os processos e as melhores práticas associadas de DM (MARISCAL *et al*, 2010).

Segundo NGUYEN (2018), através de processos e práticas, pode ser incorporado, em indivíduos e em organizações, o conhecimento, *commodity* cara para as organizações, com origens diversas, como, por exemplo, documentos, processos, pessoas, comunicação, cultura e aprendizagem. O autor afirma que a transferência de conhecimento estimula inovações e que o armazenamento de conhecimento é o caminho para se criar uma propriedade inestimável para as organizações. Um bem que se acumula ao longo do tempo e que não pode ser comprado por dinheiro algum.

Diante do exposto no referencial teórico, pode-se concluir que a documentação da memória de um projeto de DM é um pequeno passo, uma vez que localizado em um contexto menor de um projeto, que pode significar um salto em direção à memória organizacional e aos ganhos derivados do aprendizado partilhado e do conhecimento gerenciado para uma organização. Na próxima seção, o Rastro-DM é apresentado como um caminho em potencial para se alcançar esse salto organizacional.



2.4 RASTRO-DM

2.4.1 VISÃO GERAL

O Rastro-DM é uma metodologia que objetiva a documentação de projetos de DM com foco no processo de construção dos modelos, de forma a deixar um rastro das ações executadas, dos treinamentos realizados e os resultados obtidos e dos aprendizados concebidos. Engloba três atividades que correspondem aos conceitos documentados:

- Definição de ação;
- Registro de treinamento;
- Síntese de aprendizado.

Diferentemente do foco tradicional de se documentar o produto final (e seus artefatos), Rastro-DM foca na documentação do processo por trás da construção dos modelos. Conforme CONKLIN (1996), promove-se assim o aumento da memória organizacional, com o registro do contexto de criação dos artefatos: os pressupostos, os valores, as experiências, o motivo, as conversas e as decisões conduzidas.

Como visto no referencial teórico, todo o conhecimento informal relevante deve ser documentado para refletir na documentação a iteratividade e a interatividade do processo e estar disponível para recuperação posterior.

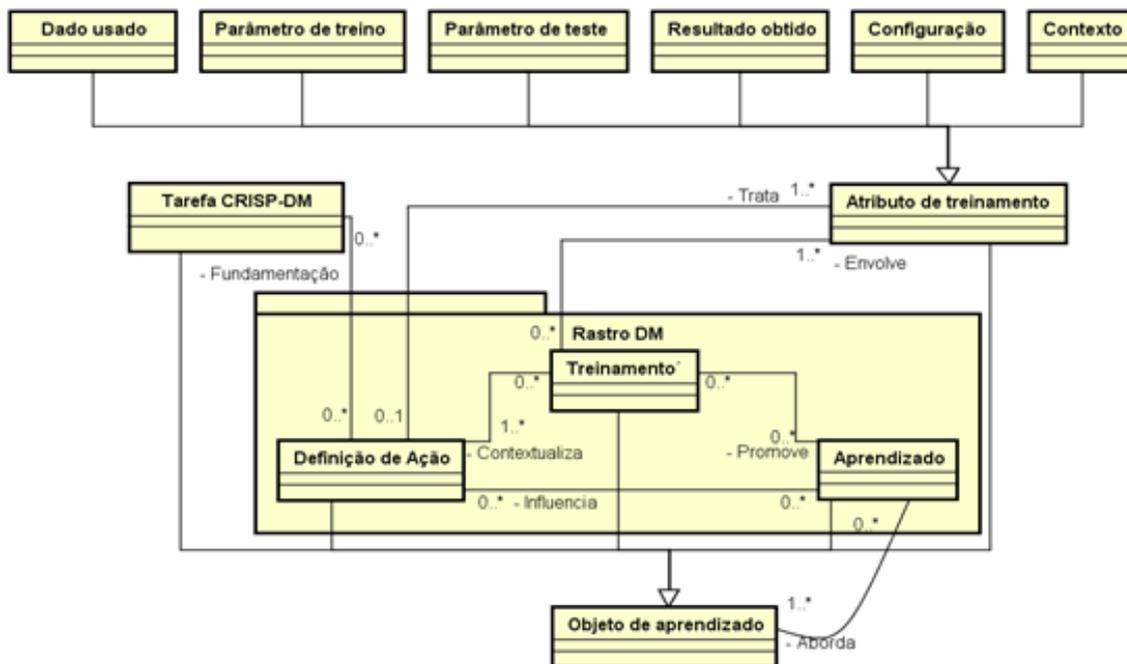
As atividades do Rastro-DM são complementares às tarefas previstas na metodologia em uso na organização, qualquer que seja ela, aqui chamada de metodologia base, que traz todo o arcabouço metodológico e paradigmático usado em um processo de DM. MINGERS e BROCKLESBY (1997) identificam várias maneiras de se combinar metodologias. Afirmam que o estabelecimento de boas práticas é uma forma de criação de uma nova metodologia, logo, não há incongruência em se referir ao Rastro-DM como uma metodologia.

Por padronização e clareza, será usado nas explicações o CRISP-DM como metodologia base, e os passos da metodologia base serão referenciados como tarefas CRISP-DM e os do Rastro-DM como atividades Rastro-DM. As atividades Rastro-DM ocorrem várias vezes durante um projeto e podem estar associadas a uma ou mais tarefas CRISP-DM.

2.4.2 VISÃO INTEGRADA DOS CONCEITOS DO RASTRO

Os conceitos do rastro se relacionam: os treinamentos acontecem no contexto de uma ação definida no projeto e esses treinamentos podem promover aprendizados que por sua vez podem influenciar, em um círculo virtuoso, novas ideias definidas em ações. Alguns relacionamentos possíveis entre os principais conceitos do Rastro-DM podem ser visualizados na Figura 3. O diagrama não tem a pretensão de ser completo, mas apenas promover uma melhor compreensão.

Figura 3: Principais conceitos associados ao Rastro-DM.



Fonte: Castro (2019).

Percebe-se que o conceito Atributo de treinamento, subdivido em seis categorias, pode, além de ser envolvido em treinamentos, ser objeto de definições de ação. As definições de ação podem ter fundamentação em tarefas da metodologia base, representada na figura pelo CRISP-DM. E todos os elementos podem ser abordados em um aprendizado.

A seguir, serão descritas as atividades do Rastro-DM.

2.4.3 DEFINIÇÃO DE AÇÃO

Atividade de registro da definição dos passos de um projeto, executados ou não. O objetivo da atividade não é o registro dos detalhes da execução de uma ação, mas as informações sobre sua definição, como a declaração de seu objetivo e as técnicas a serem usadas ou experimentadas quando da execução da ação.

Conceitualmente, uma definição de ação corresponde à descrição de uma ou mais tarefas específicas do CRISP-DM instanciadas em um projeto.

Diante das dificuldades de gerenciamento em processos de DM identificadas no referencial teórico, pode-se avaliar a possibilidade do registro de recursos usados nas ações definidas (pessoas, tempo, etc.) para apoiar estimativas de prazo e o estabelecimento de cronogramas do projeto. E o histórico desses dados pode servir de base para alocação de recursos em projetos futuros.



A atividade de definição de ação pode ocorrer a qualquer momento de um projeto de DM. É importante pelo menos a definição de cada ação que se inicia, pois, de certa forma, justifica os treinamentos que se seguirão. Além disso, saber o que já foi feito reduz a perda de esforço com execuções repetidas.

O nível de abstração e de detalhe, se mais próximo de uma tarefa CRISP-DM, como, formatação de dados, ou mais detalhada, como, formatação do nome do arquivo para possível retirada de *stopwords* e de pontuações, cabe a cada equipe se não houver um padrão corporativo.

Embora a simplicidade de um campo textual descritivo seja aceitável, uma vez que o conteúdo é o mais importante, quanto mais estruturado for o registro, melhor traduzirá a compreensão do processo e maior será o potencial de sua contribuição. As entidades do ML *Schema* constantes da Figura 2 são exemplos de atributos associados a uma Definição de Ação.

Para clarear os conceitos, usaremos um projeto hipotético de aprendizagem supervisionada que objetiva a predição do preço de uma casa (regressão) para uma imobiliária a partir de alguns atributos do imóvel.

Seguem exemplos de definições de ação do projeto hipotético associados às datas de registro de início da ação.

- 10/10/2018; Experimentar como características do imóvel no modelo o número de pavimentos e o número de banheiros;
- 8/11/2018; Avaliar se o modelo alcança uma melhor performance se os valores da característica distância do centro forem ajustados para outra escala;
- 1/2/2019; Testar os algoritmos *lightgbm* e *randomforest* para a regressão.

A compreensão dos termos técnicos usados em exemplos de registros não é relevante para o objetivo deste artigo. O foco, no caso, é a capacidade de descrição de uma ação e não nos elementos dessa descrição (*lightgbm*, *randomforest*). Contudo, para um conhecimento mais aprofundado sobre técnicas associadas a DM, sugere-se a leitura do livro *Data Mining and Analysis: Fundamental Concepts and Algorithms* de ZAKI e MEIRA (2014).

2.4.4 REGISTRO DE TREINAMENTO

Atividade de documentação dos treinamentos realizados, os parâmetros usados e os resultados obtidos. Conforme o referencial teórico, o treinamento é a atividade central de todo o processo de mineração de dados e é importante armazenar o máximo possível de informações relevantes deles.

O ML *Schema* constante da Figura 2 fornece um conjunto de entidades envolvidas em um treinamento e que são candidatas a terem informações documentadas no contexto da atividade de registro de treinamento.



Objetivando uma maior compreensão, sem a intenção de completude da classificação nem dos exemplos, os dados envolvidos em um treinamento podem ser agrupados em 6 categorias:

- Dados usados: dados de teste, dados de treinamento, variáveis usadas como features para o modelo;
- Parâmetros de treinamentos: algoritmo usado, hiperparâmetros considerados, implementações de técnicas aplicadas;
- Parâmetros de testes: métrica considerada, forma de apuração;
- Resultados: modelo com suas características e as métricas apuradas;
- Configuração: identificação do programa, versões das bibliotecas usadas, dados de hardware;
- Contexto: código do treinamento, data e hora, número de épocas de treinamento, mensagem de erro em caso interrupção da execução.

O nível de detalhe de documentação impacta a sua utilidade. O código de um treinamento, dado de contexto, pode ser usado como chave de identificação do modelo treinado.

Seguem exemplos de registros de treinamentos do projeto hipotético de predição do preço de uma casa:

- Código: 1; Data: 7/6/2018; Variáveis usadas: área da casa, área do lote e CEP do endereço; Algoritmo usado: *linear regression*; Erro: 0.8; Separação de dados de teste: 10%, não estratificada;
- Código: 100; Data: 7/7/2018; Variáveis usadas: área da casa, área do lote, CEP do endereço, número de quartos e data da construção; Algoritmo usado: *randomForest*; Erro: 0.7; Separação de dados de teste: 5%, dados estratificados.

2.4.5 SÍNTESE DE APRENDIZADO

Atividade de síntese e registro dos aprendizados concebidos ao longo do projeto, de forma automática ou não.

A atividade pode ocorrer a qualquer momento de um projeto de DM e pode estar associada ou não a treinamentos. Embora possam ser sintetizados no estágio de pré-desenvolvimento do modelo, nas fases CRISP-DM de entendimento do negócio e dos dados, os aprendizados, em sua maioria, são gerados a partir dos treinamentos realizados. Pode-se sintetizar, automaticamente ou não, que uma determinada seleção de variáveis ou que o uso de um determinado hiperparâmetro de uma técnica levou à geração de um modelo de melhor performance. Pode haver aprendizados que envolvem treinamentos que falharam que objetivam documentar como evitar que erros aconteçam novamente.



Um Aprendizado pode ter como atributo informações das entidades do ML *Schema* (Figura 2), mas também pode envolver outros conceitos, como Ações Definidas e Treinamentos.

A síntese de aprendizados e o uso efetivo deles no projeto ou em projetos futuros promovem o amadurecimento da equipe no processo de DM, no conhecimento sobre a alta iteratividade e interatividade de suas tarefas, sobre as várias técnicas e sobre as ferramentas usadas.

A documentação dos aprendizados impede que eles se percam nas memórias dos indivíduos ou mesmo com os indivíduos quando deixam o projeto ou a organização.

Para clarear o conceito, seguem exemplos de possíveis aprendizados concebidos no projeto hipotético de previsão de preço de uma casa associados à data de registro e à tarefa CRISP-DM de origem:

- 3/10/2018; Selecionar técnica; A técnica randomforest se mostrou superior à técnica decisiontree no contexto avaliado;
- 26/5/2019; Formatar dados; É necessário que os valores dos imóveis nos dados usados para treinamento sejam atualizados para uma mesma referência monetária;
- 26/7/2019; Selecionar dados; O acréscimo das variáveis número de quartos e número de vagas para automóveis promoveu uma melhora de 10% na acurácia do modelo.

2.5 DIRECIONAMENTOS

Seguem alguns direcionamentos práticos para a efetiva e eficaz aplicação do Rastro-DM.

2.5.1 ADAPTABILIDADE

As atividades do Rastro-DM são complementares às tarefas previstas na metodologia base, que traz todo o arcabouço metodológico e paradigmático usado em um processo de DM. A documentação do rastro, fora do contexto rígido de uma metodologia base, permite tratar melhor a interatividade e iteratividade das tarefas, algo não bem mapeado pelas metodologias de DM, conforme discutido no Referencial Teórico. Como as tarefas e as fases de um processo de DM se misturam e muitas vezes são executadas de forma concomitante, fica difícil se manter uma documentação efetiva por tarefa da metodologia base. Por exemplo, se no contexto de uma tarefa CRISP-DM for produzido um relatório e depois se voltar a essa tarefa diversas vezes, a documentação precisaria ser constantemente atualizada, de forma restrita ao escopo da tarefa. O custo do retrabalho torna impeditiva a documentação tempestiva no CRISP-DM.

A documentação gerada pelo Rastro-DM pode se encaixar nos artefatos de saída das metodologias bases. Os registros podem, em um momento posterior, por exemplo, ser agrupados por tarefa CRISP-DM. Por exemplo, o documento Razões para exclusões e seleções da atividade Seleção de dados da fase Preparação de dados do CRISP-DM pode ser um



relatório construído automaticamente a partir de aprendizados e de definições de ações que tratam seleções de dados, enriquecido com um resumo dos treinamentos relacionados a cada critério de seleção experimentado.

2.5.2 TEMPESTIVIDADE

O registro deve ser tempestivo para ajudar a conduzir o projeto em andamento de forma mais eficaz. Com o adequado registro do rastro, evita-se o desperdício com execuções repetidas de trabalhos no transcorrer do projeto. Como visto, uma documentação post mortem não se adequa a projetos de DM, que são considerados vivos dada sua complexidade e sua iteratividade.

2.5.3 FLEXIBILIDADE

A definição dos atributos a serem armazenados deve ser flexível e pode variar conforme o objetivo de mineração (classificação, regressão, etc), a plataforma de treinamento e os objetivos da documentação. Exemplificando, se a documentação tiver como objetivo a reprodutibilidade, há que se gravar mais detalhes de configuração de software (versões de bibliotecas) e de sementes de números aleatórios usados.

Rastro-DM é flexível ao não definir uma relação mínima de atributos para cada conceito. Afinal cada projeto e organização tem sua complexidade particular e seu grau de amadurecimento em DM. Em última análise, cabe à equipe garantir a veracidade e a efetividade da documentação. Espera-se que, com o amadurecimento em DM, as organizações elaborem um padrão corporativo mínimo de documentação por objetivo de DM. Mas, esse padrão não pode afastar a criatividade da equipe nem se tornar apenas uma sobrecarga de trabalho.

2.5.4 AUTOMAÇÃO

É imprescindível que a atividade de registro de treinamento seja automática e esteja vinculada à realização de treinamento de modelo na plataforma em uso. Ainda que haja um custo inicial de construção do arcabouço de software para uma determinada configuração de ferramentas, esse esforço será implementado uma única vez e reaproveitado nos demais projetos. Como visto, GREFF *et al* (2017) alertam sobre o desafio prático quando não se incentiva a construção dessa estrutura.

É desejável que todas as atividades sejam realizadas de forma integrada às plataformas, para que o esforço não seja uma barreira para a documentação. Afinal, para ser útil, um modelo de documentação deve corresponder, tanto quanto possível, ao modo como as pessoas trabalham (BECKER e GHEDINI, 2005).

Castro (2019) ilustra uma implementação simples de uma infraestrutura de construção de um rastro em python, criado a um baixo custo com arquivos locais.



2.5.5 USABILIDADE

Em relação às definições de ação e aos aprendizados, deve-se priorizar o seu registro em detrimento à sua categorização, de forma que esta atividade não seja uma barreira para a documentação. O registro pode se dar inicialmente até mesmo em formato de texto livre.

A categorização, por exemplo quanto às técnicas envolvidas, deve ser realizada, de preferência, de forma automática e pode ser um requisito corporativo importante para se potencializar a utilidade do rastro através da partilha do conhecimento. Como visto, a combinação de DM para o KMP potencializa o crescimento do conhecimento organizacional (NGUYEN, 2018).

Também deve se pensar na usabilidade das aplicações de consulta ao rastro, que tende a se tornar corporativo. Consultas simples podem, por exemplo, encontrar projetos que experimentaram determinada técnica, identificar dicas de seu uso e avaliar os resultados alcançados, servindo de apoio para novos projetos e para a definição de estratégias de DM na organização.

Na próxima seção será ilustrada a aplicação do Rastro-DM em um projeto de DM.

2.6 PROJETO CLADOP – ILUSTRAÇÃO DE USO DO RASTRO-DM

O projeto Cladop² é um exemplo de utilização do Rastro-DM. O termo Cladop é usado tanto para indicar o processo (projeto) quanto o produto construído (Classificador de Documentos em PDF). Consistiu o projeto no desenvolvimento por aprendizagem supervisionada de um classificador automático de tipo para documentos em formato PDF (*Portable Document Format*) inseridos no sistema de gestão de Tomadas de Contas Especiais (e-TCE) do Tribunal de Contas da União (TCU).

Um processo de Tomada de Contas Especial, em última análise, objetiva o ressarcimento ao Erário Público de danos gerados por agentes públicos e a devida responsabilização destes. Com o objetivo de tornar mais célere e eficaz o trâmite do processo de apuração de danos, com a padronização e a otimização de procedimentos, foi desenvolvido pelo TCU com cooperação da CGU o sistema e-TCE, plataforma única de acesso a todas as entidades da Administração Pública que atuam em alguma fase da TCE. Castro (2019) detalha o contexto de negócio e os normativos relacionados.

O classificador desenvolvido promoveu benefícios ao negócio tanto na usabilidade do sistema, ao identificar automaticamente os tipos em apoio aos usuários, quanto na qualidade dos dados, ao garantir uma maior correção dos tipos e uma melhor qualidade do texto dos documentos gerados por OCR (*Optical Character Recognition*). Documentos com classificação correta de

2 Após a finalização do projeto Cladop para o sistema eTCE, projeto abordado neste artigo, iniciou-se a construção de um classificador semelhante para outro sistema: Protocolo Eletrônico. O novo projeto, em andamento, foi chamado de Cladop@Protocolo. Então, deve-se compreender Cladop neste artigo como Cladop@eTCE.



tipo e com um conteúdo textual de melhor qualidade são fundamentais para um rito processual assistido pelo computador.

2.6.1 ENTENDIMENTO DOS DADOS

Durante a construção do Classificador, foram considerados 118.266 documentos relativos a 4.384 danos que se distribuíam em 84 tipos de documento. O tipo Outros era o mais usado com 18,81% dos documentos. O Gráfico 1 permite a visualização parcial do acentuado desbalanceamento da quantidade de documentos entre os tipos ativos.

Gráfico 1: Visão parcial do desbalanceamento de documentos por tipo (referência: 17/4/2019).

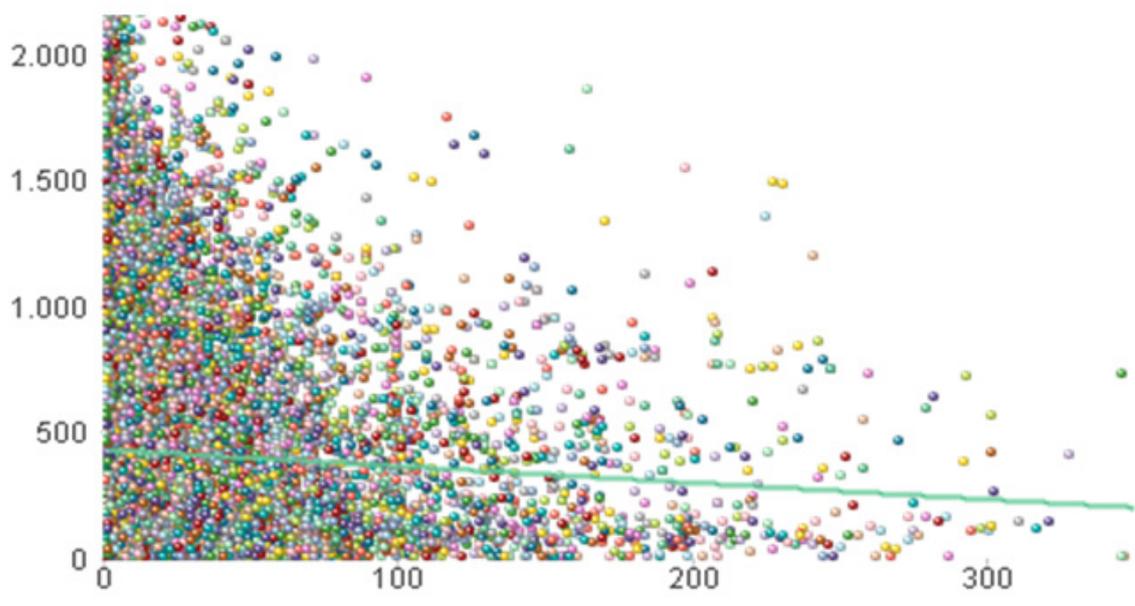


Fonte: Castro (2019) - visão parcial.

BRANTING (2017) afirma que o formato PDF tem sido usado em tribunais e que o texto obtido desses documentos apresenta muitos erros e não preservam a sequência original do documento, devido ao processo usado de OCR.

O alcance dos resultados esperados pelo sistema e-TCE depende da qualidade dos documentos protocolados no sistema. A eficácia do classificador também tende a ser superior se os dados tiverem uma maior qualidade. Foi encontrada uma baixa qualidade no OCR dos documentos. Um exemplo dessa situação é o documento de protocolo 58.900.414 que tem 162 páginas, mas não se consegue via OCR identificar nem uma centena de palavras válidas, ou seja, menos de uma palavra válida por página. A baixa qualidade do OCR pode ser ilustrada no Gráfico 2 que apresenta o número de páginas e o número de palavras válidas por página para alguns documentos.

Gráfico 2: Documentos – quantidade de palavras válidas por página (eixo y) x número de páginas (eixo x).



Fonte: os autores (2020).

Foram experimentadas formas diferentes de pré-processamento de texto dos conteúdos dos documentos. Conforme Castro (2019), os métodos experimentados retrataram diferentes combinações de métodos de obtenção de texto do documento (algoritmos de OCR), de critérios para definição de palavras válidas e de classes substitutas de palavras consideradas (cpf, data, nome pessoa física, etc).

2.6.2 DESCRIÇÃO FUNCIONAL DO CLASSIFICADOR

O Cladop, em sua versão 2.1, alcança a acurácia de 91,1% com desvio padrão de 0,3%, desempenho apurado com validação cruzada de 7 partições e 2 repetições, totalizando 14 amostras. No contexto deste trabalho, o termo acurácia deve ser entendido como acurácia micro, que leva em consideração os resultados, acertos e erros, por documento independentemente do seu tipo.

O classificador foi implementado em python na forma de um webservice que recebe como parâmetro um arquivo PDF, e a partir do seu nome e do seu conteúdo retorna nove tipos mais prováveis, com suas respectivas probabilidades. A acurácia do classificador³ sobe de 91% (do primeiro tipo) para 99% quando consideradas as nove primeiras previsões. O retorno de mais de um tipo possibilita que o sistema, por exemplo, possa apresentar os demais tipos em uma segunda tela caso o primeiro tipo não seja acatado pelo usuário.

3 Após a escrita deste artigo, houve inclusão e exclusão de tipos de documentos e foi construída, considerando documentos mais recentes, a versão 3.0 do Cladop que alcançou acurácia de 93,6% com desvio padrão de 0,3%.



A Tabela 1 apresenta, para alguns tipos de documento, algumas métricas apuradas sobre os dados de validação, 5% do total, quando da geração da versão em produção do classificador. Os números gerais em dados de validação são: acurácia: 90,99%; precisão macro: 81,04% e weighted: 91,06%; recall macro: 80,29% e weighted: 90,99%; F1 macro: 80,16% e weighted: 90,87%.

Tabela 1: Métricas apuradas por tipo sobre os dados de validação (visão parcial).

Descrição	Precisão	Recall	F1	Documentos
Ação judicial - petição inicial	93,33%	93,33%	93,33%	15
Acórdão	85,71%	85,71%	85,71%	7
Análise de Prestação de Contas	77,27%	62,96%	69,39%	27
Análise defesa	83,33%	52,63%	64,52%	19
Suspensão de inadimplência	73,68%	82,35%	77,78%	17
Termo de concessão e de aceitação da bolsa e aditivos	75,00%	54,55%	63,16%	11
Termo de recebimento definitivo da obra	100,00%	100,00%	100,00%	4

Fonte: Castro (2019) – visão parcial.

Adicionalmente às predições, o classificador retorna informações derivadas do pré-processamento do texto do arquivo, que podem ser úteis para o sistema e-TCE exigir no cadastro dos documentos uma qualidade mínima de conteúdo textual, como: quantidade de palavras válidas, quantidade de valores, quantidade de nomes, etc. Assim, o sistema pode impedir que documentos críticos para o processo de TCE tenham qualidade baixa de OCR com um alto percentual de palavras inválidas ou mesmo que tenham conteúdo incompleto, como o caso de um documento de AR, Aviso de Recebimento, sem informação de CPF ou CNPJ e data.

2.6.3 APLICAÇÃO DO RASTRO-DM NO PROJETO

A utilização do Rastro-DM se mostrou uma solução efetiva para a dificuldade apresentada no referencial teórico de se refletir na documentação a iteratividade e a interatividade de um processo de DM.

As seções seguintes exemplificam a execução das ações da metodologia no projeto e os benefícios auferidos.

Detalhes do rastro do projeto Cladop (Rastro-DM@Cladop) e o código usado para sua construção estão acessíveis em <https://gitlab.com/MarcusBorela/rastro-dm.git>.



2.6.4 DEFINIÇÃO DE AÇÃO

As primeiras definições de ação foram registradas como comentários no próprio código de implementação dos treinamentos. Porém, com o tempo, elas passaram a ser mais complexas e amplas e perpassavam vários códigos. Então passaram a ser persistidas em banco de dados. Definitivamente não se pode confiar o registro de definições de ação a código. Devido à natureza imprevisível de um projeto DM o código muitas vezes evolui rapidamente e acaba comprometendo, entre outras coisas, sua documentação (GREFF *et al*, 2017). A Tabela 2 ilustra algumas definições de ação registradas durante o projeto.

Tabela 2: Exemplos de definições de ação registradas no Cladop.

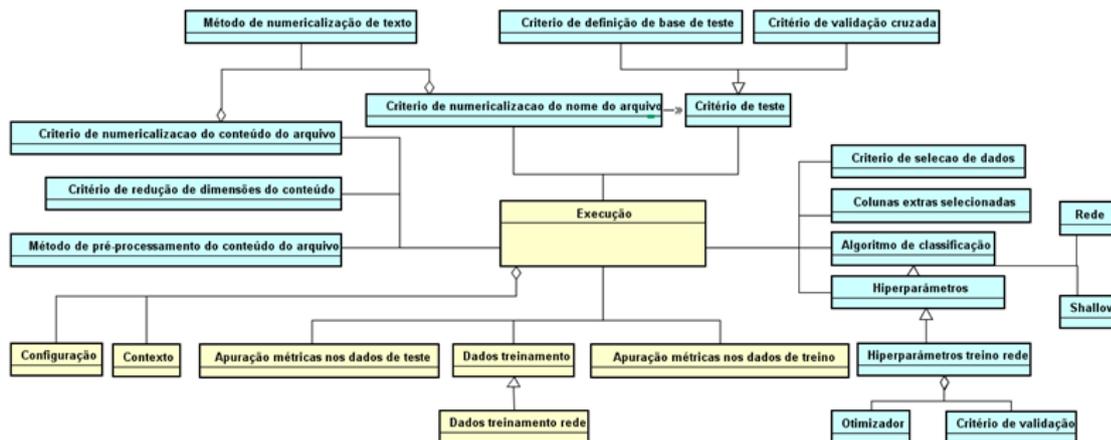
Momento	Descrição	Tarefa CRISP-DM
12/3/2019-17:04	Criando estrutura (código e dados) para tratar k-fold em shallow algorithms	Projeto de testes
14/3/2019-19:27	Iniciadas execuções para experimentar otimizadores (MLP): nadam, adadelta	Construir Modelo
18/3/2019-11:40	Experimentando colunas não mais binárias com MLP	Formatar Dados
26/3/2019-11:57	Iniciando inclusão de nomes de arquivos no modelo	Selecionar Dados
29/5/2019-19:30	Alterado programa (shallow) para gravar também recall e f1 micro	Projeto de testes
4/6/2019-19:35	Alterado programa (shallow) para não gravar recall e f1 micro, pois são equivalentes à acurácia (e à precisão)	Projeto de testes

Fonte: Castro (2019) – visão parcial.

2.6.5 REGISTRO DE TREINAMENTO

O conjunto de dados persistidos em banco de dados sobre treinamentos evoluiu no decorrer do projeto acompanhando o amadurecimento da equipe em DM. A Figura 4 apresenta os principais conceitos persistidos que podem ser relacionados aos subtipos da Figura 3. Em azul claro, constam os conceitos que são entrada para o treinamento e em amarelo os dados derivados do processo.

Figura 4 - Informações persistidas de treinamento no projeto Cladop agrupadas por conceito



Fonte: Castro (2019).

A Tabela 3, de caráter meramente ilustrativo, apresenta alguns atributos dessas entidades e seus respectivos valores para os treinamentos de teste e de geração do modelo construído na versão 2.1 do Cladop.



Tabela 3 - Rastro dos treinamentos de teste e de geração da versão 2.1 do Cladop.

Grupo	Item	Teste	Geração	
Contexto	Código (chave única)	13.134	13.138	
	Data e hora de registro	05/07/2019-00:44:59	05/07/2019-18:02:12	
	Número de épocas treinadas	33	24	
	Quantidade de documentos usados		63.468	
	Quantidade de tipos tratados		81	
	Tempo de execução (em segundos)	460	475	
Configuração	Programa executado	Cladop_monitoramento.ipynb		
Parâmetros	Modo de conversão de texto para números		Tfidf	
	Número de palavras do modelo “bag of words” do texto		24.576	
	Número de palavras do modelo “bag of words” do nome do arquivo		1.000	
	Método de pré-processamento do texto		7	
	Critério de seleção dos dados	Documentos de tipo diferente de Outros e criados após 1/5/2018		
	Método de teste	Validação cruzada com 7 partições e 2 repetições	Não foram separados dados para teste	
	Otimizador usado para a rede		Adam	
	Tamanho do batch de treinamento		256	
	Quantidade mínima exigida por tipo		0	
	Dimensão final do vetor de conteúdo após redução		768	
	Algoritmo de redução de tamanho		TruncatedSVD	
	Dados foram embaralhados a cada época?		Não	
	Critério para dados de validação	5% dos dados, sem estratificação e com shuffle		
	Resultado	Acurácia apurada em teste	91,1% +- 0,3%	
Acurácia apurada em treinamento		96,4% +- 0,7%	95,6%	
Acurácia apurada em validação		90,8% +- 0,6%	92,1%	

Fonte: Castro (2019) – versão adaptada.

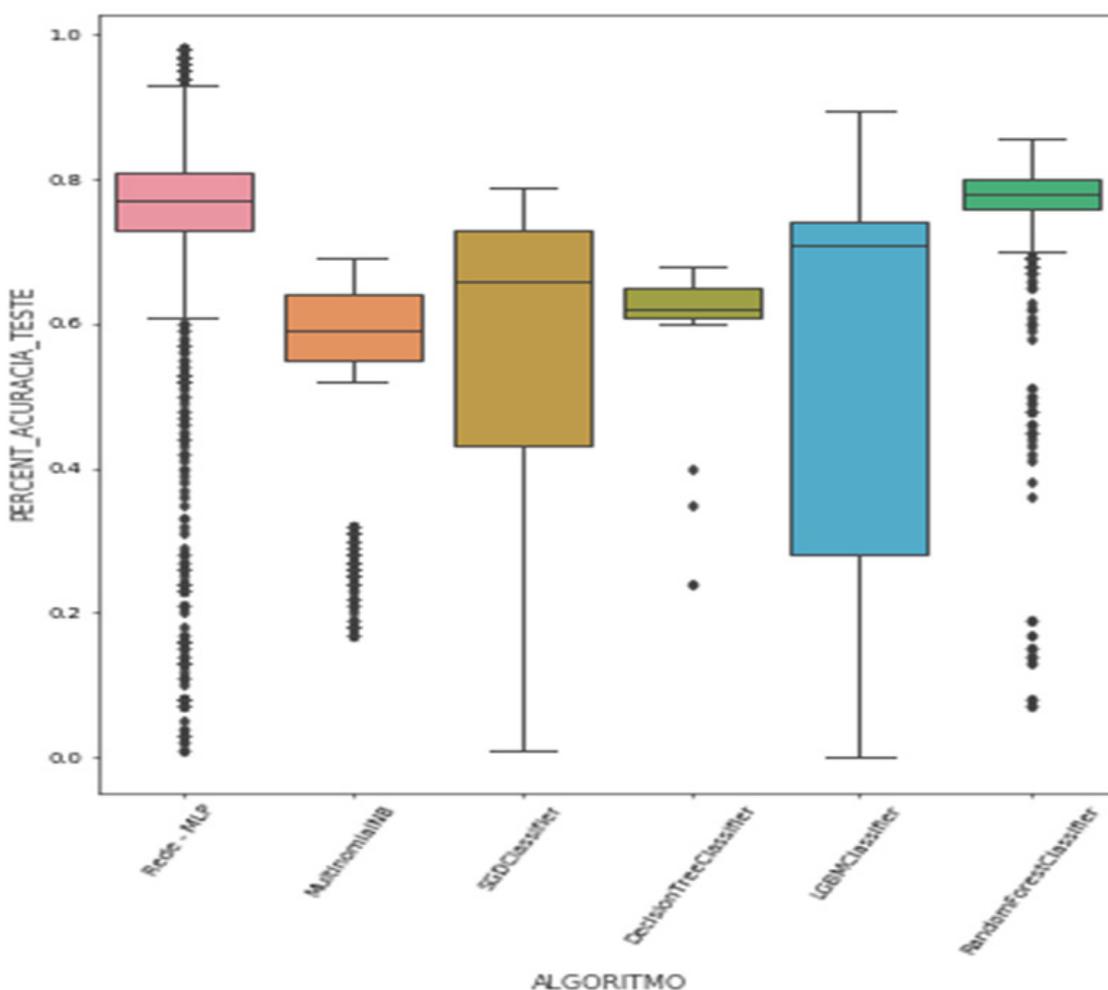
O treinamento que gerou a versão 2.1 do modelo tinha o código 13.138 (conforme segunda linha da Tabela 3), o que demonstra que foram executados mais de treze mil treinamentos de modelos. Alguns treinamentos envolvendo apenas documentos com alta qualidade de OCR e tipos com grande quantidade de exemplares levaram a classificadores com acurácia superior



a 98% em base de teste. É o caso do treino de código 147, em 24/11/2018, que envolveu 11.110 documentos de melhor qualidade de OCR dos quatro tipos de documentos mais frequentes. Contudo, no decorrer do projeto, optou-se pela construção de um classificador que abrangesse todos os tipos de documentos e com critérios menos restritivos.

A partir dos dados de treinamento podem ser feitas diversas análises. Um exemplo é a comparação do desempenho dos modelos quanto aos algoritmos utilizados. Os modelos com maior acurácia para a seleção final de dados e de parâmetros foram implementados com redes neurais (rede MLP). Modelos implementados usando o algoritmo *LGBMClassifier* ficaram logo atrás, com uma diferença de cerca de 2%. O Gráfico 3 ilustra a acurácia em base de teste alcançada por alguns dos algoritmos experimentados. A grande variação de acurácia por algoritmo se explica pela variação de combinações de parâmetros experimentadas nos treinamentos. O entendimento desses algoritmos não está no escopo deste artigo. Para interessados, sugere-se a leitura do livro *Data Mining and Analysis: Fundamental Concepts and Algorithms* de ZAKI e MEIRA (2014).

Gráfico 3 - Boxplot com acurácia em base de teste dos algoritmos experimentados.



Fonte: Castro (2019) – visão parcial.



2.6.6 SÍNTESE DE APRENDIZADO

Dos aprendizados registrados, a maioria poderia ser sintetizada por mineração de dados nos registros de treinamentos. Por exemplo, a partir de critérios de igualdade de valores em algumas colunas, poderia ser detectado o aprendizado do dia 10/4/2019 da Tabela 4 que ilustra alguns aprendizados. Outros carecem da intervenção humana, como o do dia 27/2/2019. A geração de conhecimentos de forma automática se mostra um caminho promissor para a elevação do conhecimento organizacional, o que foi referendado por NGUYEN (2018), que, como visto, ressalta haver um grande vazio de pesquisas nessa área.

Tabela 4 – Exemplos de aprendizados registrados durante o projeto

Data e hora de registro	Descrição	Atividade CRISP-DM de contexto
27/2/2019-10:32	Bastam as variáveis de qualidade do PDF e de contexto para se alcançar um resultado de 44% de acurácia, com otimizador Adam. Com Adagrad:42%; SGD:28%.	Selecionar Dados
27/2/2019-10:39	Ao usar <i>pca</i> (<i>sklearn</i>): melhor separar o comando <i>fit</i> do comando <i>transform</i> . O comando <i>fit_transform</i> estava travando!	Formatar Dados
27/2/2019-10:53	Ao usar <i>pca</i> (<i>sklearn</i>): se o número de dimensões for bem pequeno e o <i>array</i> for grande (ver detalhes na documentação da função), melhor usar o método <i>randomized</i> do que o <i>full</i> . Pois é mais rápido, e os resultados (variância alcançada) são equivalentes.	Formatar Dados
29/03/2019-09:54	Percebida uma melhora de cerca de 5% na acurácia dos modelos após inclusão do nome do arquivo como característica.	Selecionar Dados
1/4/2019-15:36	Para algoritmos <i>shallow</i> , as colunas extras com valores não <i>dummies</i> (uma só coluna com vários valores discretos) levaram a um resultado melhor. Para Rede Neural, há uma pequena melhora usando valores <i>dummies</i> .	Formatar Dados
10/4/2019-10:49	Em classificação de multiclases, as métricas <i>f1_micro</i> , <i>recall_micro</i> , <i>precision_micro</i> e acurácia são equivalentes	Projeto de testes
22/5/2019-20:04	O otimizador Adam (passando objeto com <i>Amsgrad=False</i>) foi o melhor otimizador até o momento. Acima do <i>Amsgrad=True</i> em 0.1%, do <i>Rmsprop</i> e do <i>Adadelta</i> em 0.2% no contexto avaliado das últimas execuções.	Construir Modelo

Fonte: Castro (2019) – visão parcial.

Passados menos de seis meses do final do projeto, Castro (2019) experimentou contabilizar, de forma amostral, quantos deles ainda se encontravam completos em sua memória e chegou a uma estimativa de apenas 27%. Os demais, infelizmente, ou não estavam completos ou haviam sido perdidos da memória do autor, mas, graças ao rastro, não foram perdidos da memória do projeto.

Vimos que o esquecimento pode levar à execução repetida de experimentos (BECKER e GHEDINI, 2005). Durante o projeto houve um esquecimento que levou a esforços



desnecessários em execução de tarefas. O aprendizado de 10/4/2019 foi ignorado com a execução desnecessária da ação para se gravar também as métricas no dia 29/5/2019 (ver Tabela 2), o que levou a um retrabalho com a alteração do programa para não gravar essas métricas em 6/4/2019. Ficou a lição aprendida de que não basta ter o rastro, é necessário fazer uso do mesmo.

2.6.7 BENEFÍCIOS ALCANÇADOS

Os benefícios auferidos com a aplicação da metodologia no projeto podem ser agrupados em 3 aspectos: gestão de projetos, capacitação técnica da equipe e automação de atividades.

Na gestão de projetos, perceberam-se ganhos na comunicação, no planejamento e na diminuição de custos.

O rastro construído viabilizou um acompanhamento gerencial contínuo da evolução do projeto e do desempenho dos modelos gerados. O uso de uma padronização na apresentação de relatórios de acompanhamento inspirou confiança nos usuários e nos patrocinadores.

O rastro se tornou subsídio para um melhor planejamento para projetos futuros semelhantes, afinal, projetos de mineração de dados podem ser beneficiados com as experiências de projetos anteriores (CHAPMAN *et al*, 2000). O relato da experiência contida no rastro torna mais fácil defender recursos mais realistas, pois há uma compreensão mais profunda de como o esforço é realmente gasto. Permite, também, a identificação de pontos críticos no processo seguido.

Constatou-se uma diminuição no esforço em trabalhos repetidos ou mesmo em caminhos não frutíferos.

Os bons impactos do uso do Rastro-DM na gestão levaram à sua inclusão como requisito técnico para a terceirização de desenvolvimento de projetos de DM que se inicia no Tribunal.

A capacitação técnica da equipe foi incrementada dado o ambiente criado, o amadurecimento alcançado e o compartilhamento dos aprendizados.

Promoveu-se um ambiente estimulante e prático (aprender-fazendo). A equipe se sentiu incentivada à documentação concomitante com a execução das atividades dada a facilidade de registro e os benefícios percebidos no decorrer do projeto, como a própria documentação gerada.

Não só a síntese de aprendizados e o uso efetivo deles no projeto como também a evolução estrutural dos registros de treinamento traduziram o amadurecimento da equipe em DM. À medida que se alcançava uma melhor compreensão e entendimento do



processo e das técnicas envolvidas, novos campos foram criados para armazenamento e os conteúdos ficaram mais ricos.

Quanto ao compartilhamento, o rastro do Cladop foi disponibilizado no contexto de um trabalho de conclusão de curso de Especialização em Análise de Dados (Castro, 2019) e a metodologia foi apresentada no 5º Seminário Internacional de Análise de Dados da Administração Pública em 2019.

Espera-se avançar na discussão interna à organização sobre políticas e padrões para se transformar os rastros em recursos corporativos, como apoio à disseminação de conhecimento em vários níveis. Consultas simples, por exemplo, poderiam encontrar projetos que aplicaram determinada técnica e identificar dicas de seu uso e avaliações dos resultados alcançados. Consultas gerenciais poderiam subsidiar estratégias na área de DM para a organização.

Outro benefício foi a automação de rotinas, que não só garante uma qualidade mínima para atividades como reduz o custo de suas realizações. A lista cronológica de aprendizados e de definições de ação enriquecida com alguns gráficos gerados automaticamente a partir dos treinamentos podem compor relatórios automáticos sobre tarefas da metodologia base, como ilustrado em Castro (2019). O mesmo autor demonstra que o rastro viabilizou o monitoramento automático do desempenho dos modelos face a novos documentos cadastrados no sistema. Esse rastreamento promoveu, inclusive, a geração automática da versão 2.1 do Classificador a partir de dados mais atualizados.

3. CONCLUSÃO

Os objetivos propostos para o trabalho foram alcançados. Uma breve contextualização com referencial teórico foi apresentada sobre metodologias e documentação em projetos de mineração de dados, bem como do potencial impacto das experiências adquiridas nos projetos para uma organização.

Foi proposto o Rastro-DM como uma metodologia de documentação de projetos de DM com foco no processo que, entre outras características, é flexível e pode ser mesclado à metodologia em uso por uma organização.

O Rastro-DM mostrou-se viável e com resultados benéficos conforme ilustrado na sua aplicação no projeto Cladop.

Há muito ainda por fazer. Trabalhos futuros podem experimentar o uso Rastro-DM em outros contextos: outros objetivos de mineração que não classificação, outras plataformas de desenvolvimento ou mesmo em outras culturas organizacionais. A metodologia pode ser revisada com o acréscimo de novas atividades ou mesmo enriquecida com técnicas e procedimentos que complementem suas atividades. Estudos podem ser feitos para o levantamento de políticas e padrões para que um rastro possa ser usado como subsídio em contratações de DM. Também é um desafio para trabalhos futuros a mineração de aprendizados a partir de rastros de treinamentos, pois preocupa o fato de a integração



entre DM e KMP possuir um grande vazão de pesquisas (NGUYEN, 2018). E, talvez o mais importante ponto a evoluir, que não foi escopo deste trabalho, seja como tornar o rastro um recurso corporativo, que permita a partilha das experiências adquiridas e o uso de táticas institucionais para o crescimento do conhecimento organizacional. Afinal, o valor dos dados está em como eles são interpretados e usados (BERMAN *et al*, 2018).

Por fim, ficou claro que, além de contribuir para a equipe do projeto, o rastro dos projetos tem potencial para promover um salto no conhecimento organizacional, se forem adotadas medidas institucionais para incentivar sua construção, sua partilha e a busca automática de aprendizados neles contidos. Mas há que se ter em mente que nenhum direcionamento corporativo deve inibir a liberdade dos analistas de dados, pois diminuiria sua criatividade. E a liberdade criativa, de evoluir o próprio rastro, é algo de que os seres humanos não podem prescindir, pois é um dos grandes diferenciais que os impedem de serem classificados como máquinas.

4. REFERÊNCIAS

BECKER, Karin; GHEDINI, Cinara. **A documentation infrastructure for the management of data mining projects**. Information and Software Technology, v. 47, n. 2, p. 95-111, 2005.

BERMAN, Francine et al . **Realizing the potential of data science**. Communications of the ACM, v. 61, n. 4, p. 67-72, 2018.

BHATT, Ganesh D. Knowledge management in organizations: examining the interaction between technologies, techniques, and people. Journal of knowledge management, v. 5, n. 1, p. 68-75, 2001.

CASTRO, Marcus Vinícius Borela. **Mineração de Dados com Rastro: Boas Práticas para Documentação de Processos e sua Aplicação em um Projeto de Classificação Textual**. Trabalho de Conclusão de Curso (Especialização em Análise de Dados para o Controle) – Escola Superior do Tribunal de Contas da União, Instituto Serzedello Corrêa, Brasília DF. Disponível em: <https://portal.tcu.gov.br/biblioteca-digital/mineracao-de-dados-com-rastro-boas-praticas-para-documentacao-de-processo-e-sua-aplicacao-em-um-projeto-de-classificacao-textual.htm>. Acesso em 26 jul. 2020. 2019.

CHAPMAN, Pete et al . **CRISP-DM 1.0: Step-by-step data mining guide**. SPSS inc, v. 16, 2000.

CHOLLET, Francois. **Deep Learning with Python**. Manning Publications Co., Greenwich, CT, USA. 2017.

CONKLIN, Jeffret. **Capturing Organisational Memory**. In: Groupware and Computer-Supported Cooperative Work, R.M. Barcker (Ed.), Morgan Kaufman, pp. 561-565. 1996.



DINGSØYR, Torgeir; Moe, Nils Brede; Øystein. Nytrø. **Augmenting experience reports with lightweight postmortem reviews**. Lecture Notes in Computer Science, 2188:167–181, 2001.

GHEDINI, Cinara; BECKER, Karin. **KDD application management through documentation**. Disponível em: https://www.researchgate.net/profile/Karin_Becker2/publication/268253354_KDD_application_management_through_documentation/links/5657a5ec08ae1ef9297bf1d1/KDD-application-management-through-documentation.pdf. Acesso em 27 jul. 2019. 2000.

_____. **A documentation model for KDD application management support**. In: SCCC 2001. 21st International Conference of the Chilean Computer Science Society. IEEE. p. 105-114. 2001.

GREFF, Klaus et al . **The sacred infrastructure for computational research**. In: Proceedings of the Python in Science Conferences-SciPy Conferences. 2017.

KURGAN, Lukasz A.; MUSILEK, Petr. **A survey of Knowledge Discovery and Data Mining process models**. The Knowledge Engineering Review, v. 21, n. 1, p. 1-24, 2006.

MARBÁN, Óscar et al . **An engineering approach to data mining projects**. In: International Conference on Intelligent Data Engineering and Automated Learning. Springer, Berlin, Heidelberg, p. 578-588. 2007.

MARISCAL, Gonzalo; MARBAN, Oscar; FERNANDEZ, Covadonga. **A survey of data mining and knowledge discovery process models and methodologies**. The Knowledge Engineering Review, v. 25, n. 2, p. 137-166, 2010.

MINGERS, John; BROCKLESBY, John. **Multimethodology: Towards a framework for mixing methodologies**. Omega, v. 25, n. 5, p. 489-509, 1997.

NGUYEN, Ngoc Buu Cat. **Data Mining in Knowledge Management Processes: Developing an Implementing Framework**, 2018.

PRAKASH, BV Ajay; ASHOKA, D. V.; ARADHYA, VN Manjunath. **Application of data mining techniques for software reuse process**. Procedia Technology, v. 4, p. 384-389, 2012.

PUBLIO, Gustavo Correa et al . **ML-Schema: Exposing the Semantics of Machine Learning with Schemas and Ontologies**. arXiv preprint arXiv:1807.05351, 2018.

STATA, Ray. **Organizational learning: The key to management innovation**. Massachusetts Institute of Technology, 1980.



W3C (World Wide Web Consortium) **Machine Learning Schema Community Group**. W3c machine learning schema. Disponível em: <https://www.w3.org/community/ml-schema>. Acesso em 30 jul. 2019. 2017.

WIRTH, Rüdiger; HIPPEL, Jochen. **CRISP-DM: Towards a standard process model for data mining**. In: Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining. Citeseer, p. 29-39. 2000.

ZAKI, Mohammed J.; MEIRA, Wagner. **Data mining and analysis: fundamental concepts and algorithms**. Cambridge University Press, 2014.

Os conceitos e interpretações emitidos nos trabalhos assinados são de exclusiva responsabilidade de seus autores.





Rastro-DM: data mining with a trail

A methodology for documenting data mining projects and its application in the construction of a text classifier of documents associated with damages to the public treasury

MARCUS VINÍCIUS BORELA DE CASTRO 1

Auditor at the Court of Accounts of Brazil (TCU) since 1996. Bachelor's degree in Informatics from the Federal University of Viçosa (1990), and specialist in IT Governance from the University of Brasilia (2012), and in Data Analysis from the Brazilian Federal Court of Account's Serzedello Corrêa Institute (2019).

REMIS BALANIUK

Auditor at the Court of Accounts of Brazil since 1989. Bachelor's degree in Computer Science from the University of Brasilia (1986), Master's degree in Computer Science from UFRGS (1989), doctoral degree in Informatics from the Institut National Polytechnique of Grenoble (1996), and postdoctoral research in Computer Science at Stanford University (2002) and at the Institut National pour la Recherche en Informatique et Automatique (2000). Visiting researcher at the University of Oxford (2020).

ABSTRACT

This paper proposes a methodology for documenting data mining (DM) projects, Rastro-DM (Trail Data Mining), with a focus not on the model that is generated, but on the processes behind its construction, in order to leave a trail (Rastro in Portuguese) of planned actions, training completed, results obtained, and lessons learned. The proposed practices are complementary to structuring methodologies of DM, such as CRISP-DM, which establish a methodological and paradigmatic framework for the DM process. The application of best practices and their benefits is illustrated in a project called "Cladop" that was created for the classification of PDF documents associated with the investigative process of damages to the Brazilian Federal Public Treasury. Building the Rastro-DM kit in the context of a project is a small step that can lead to an institutional leap to be achieved by sharing and using the trail across the enterprise.

Keywords: Data mining; Data analysis; Data science; Machine learning; Organizational knowledge; Methodology; Best Practices; Data analysis in Government; Textual classification; Documentation; Documentation of data mining projects.

1 This is a concise version of the conclusion work resulting from the postgraduate lato sensu work in Data Analysis indicated in the references (Castro, 2019).



1. INTRODUCTION

Data are changing everything and the capacity of manipulating data and understanding Data Science is becoming increasingly critical for current and future discoveries and innovations (BERMAN et al, 2018).

Data mining (DM) projects are challenging not only because of the complex and exploratory process used on them, but also because they are generally innovative, unique and very often developed by individuals or small teams.

These are innovative projects, either because they use techniques and algorithms that may not be consolidated in the Organization or in academic research, or because they require the construction of models that simulate cognitive processes and natural intelligence by machines.

These projects are usually unique. The particularities of each context, the data implicated, and the quality requirements, prevent or make difficult the reuse of codes from other projects.

They are complex projects because the techniques used typically have a conceptualization difficult to understand and require interdisciplinary knowledge of areas such as computer science, mathematics, and statistics. Furthermore, it requires an understanding of the business for which the solution is designed.

They are exploratory processes, because the data mining activity can be defined as the process of exploring combined data with different techniques, extracting or helping to demonstrate patterns and assisting in the discovery of knowledge.

Additionally, in the case of organizations with a low maturity level in DM, these projects are the responsibility of small teams or even of individual persons. In this case, the sharing of knowledge and practices adopted in the projects becomes even more difficult.

Unfortunately, these complex works, of exploratory nature, that are innovative, unique, and, typically, individualized, do not leave a trail of what they delivered. In the end, we have the implemented solution. There may even be documentation relative to the product created. However, there is none regarding the process adopted, the choices made, and the techniques used in the different activities of the project. We are left with concerns in case it is necessary to reconstruct the model without the presence of its creator. In addition, the person responsible for it becomes the parent of the product, since only this person knows it and can maintain it.

One intriguing debate in the management of general knowledge is how to collect, gather or make explicit the experience of project development for it to be applicable for others (DINGSØYR et al, 2001). Because the memory of an organization cannot be based uniquely in the memory of its individuals (STATA, 1980).

In face of the aforementioned, a challenging topic arises: How to systematize the documentation of the tasks of a data mining project in a manner that can leverage its auditing capability and the sharing of knowledge?



In order to answer the question posed, the objective of this work is to propose a methodology as a set of good practices for the semi-automated tracking of the activities of a DM project, in order to provide a trail of the choices made, of the data processing carried out, and of the results obtained, not focusing on the product generated, but instead, focusing on the process underlying its creation. These are good practices that can be combined with the corporate DM methodology being used in the organization. The production of trails in DM projects accelerates the learning curve and the creation of an organizational culture centered on the use of data analysis.

There are three specific objectives: contextualizing the theoretical framework relative to the methodologies and documentation in data mining projects as well as the potential impact of the experiences acquired in the projects for an organization; proposing the Rastro-DM (Trail Data Mining) methodology with the description of its activities; and the illustration of its application in a project for the classification of PDF documents associated to damages to the Brazilian Federal Public Treasury, named “Cladop”, demonstrating its feasibility and the benefits of its use. The specific objectives will be discussed in the three sections of the body.

2. THEORETICAL FRAMEWORK

2.1 DATA MINING METHODOLOGIES

According to BERMAN et al (2018), the Science of Data is concentrated in the processes of extraction of knowledge or insights, originating from structured or non-structured data. This process of discovery of data knowledge (KDD - Knowledge-Discovery in Databases), according to BECKER and GHEDINI (2005), is complex and popularly called DM - Data Mining. In the scope of this work, we will refer to these processes indistinctly as KDD or DM.

CHAPMAN et al (2000) specify some objectives for DM translated into groups of problems to be discussed: description and summarization of data, segmentation (clustering), description of concepts, classification, prediction (regression), and analysis of dependence.

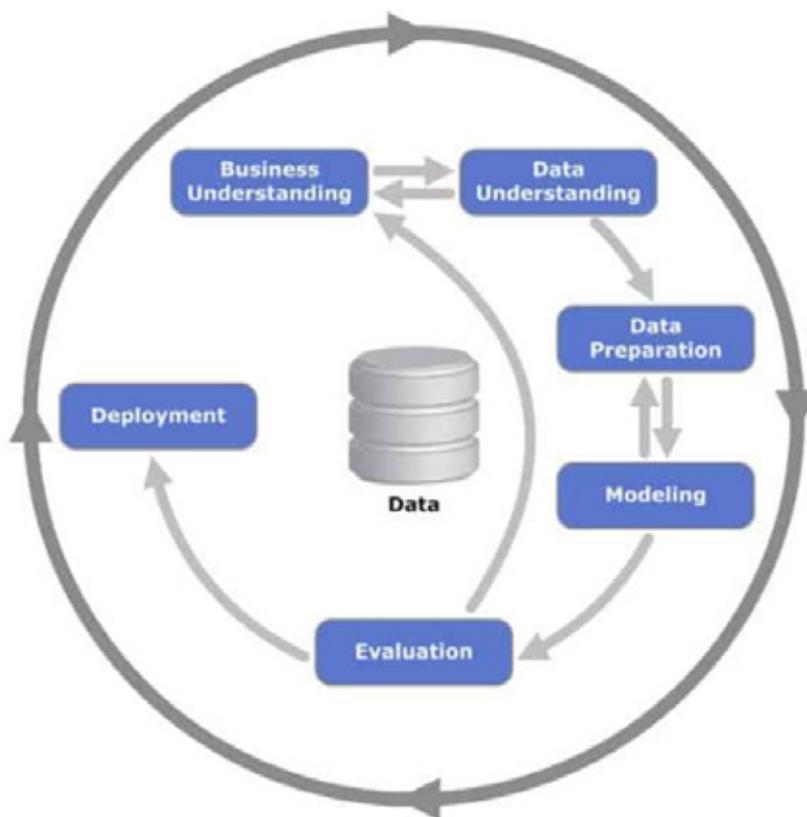
WIRTH and HIPPE (2000) confirm that DM is a complex process and associate the success of a project to an adequate combination of good hardware, qualified analysts, and use of a solid methodology and of effective project management. With regards to the methodology, the same authors state that DM requires a standard approach that assists in transforming business problems into data mining tasks, that suggests appropriate transformations of data and techniques to be employed, and that provides the means to evaluate the effectiveness of the results and document the experience. Moreover, they emphasize that the use of a methodology in planning and presenting reports inspires confidence to users and to sponsors.

KURGAN and MUSILEK (2006) state that methodologies provide a better comprehension and understanding of the process. Furthermore, they promote economy of time and costs, with the establishment of an itinerary to be followed for the planning and the execution of the projects. However, they mention that a great number of projects follow on their own methodologies.

An example of a methodology used in DM is the CRISP-DM (Cross-Industry Standard Process for Data Mining). It is considered a standard and one of the factors of its success is its neutrality in terms of industry, hardware, and applications (MARISCAL *et al*, 2010).

CHAPMAN *et al* (2000) corroborate that the CRISP-DM methodology is a model of hierarchical process, consisting in combined tasks described in four levels of abstraction (from generalized to more specific levels): phase, generic task, specialized task, and process instance. They state that the sequence of the phases is not restrictive. That, from a practical point of view, many tasks can be executed in a different order, and, frequently, it is necessary to return repeatedly to the preceding tasks and redo specific actions. They emphasize that representing every possible sequence of the means of the data mining process would require an excessively complex model of processing. Therefore, they do not do that. They state that a phase is never completely concluded before the following phase begins. Figure 1 displays the iteration between phases of the CRISP-DM methodology.

Figure 1 – General view of the CRISP-DM methodology.



Source: CHAPMAN *et al* (2000).

According to MARBÁN *et al* (2007), the CRISP-DM methodology does not encompass many tasks related to management, organization, and project quality. At least, not in a manner demanded by the increasing complexity of current DM projects that include not only a great volume of data, but also the management and organization of large interdisciplinary teams. These authors group the tasks of a DM process, relative to the construction of a model, into three stages: pre-development,



development, and post-development. They state that all the methodologies used in DM are concentrated in the development stage, which corresponds to the collection and analysis of the available data for the project, the creation of new data based on the available data, the adaptation for DM algorithms, and the creation of models.

According to CHOLLET (2017), the development of a project is concentrated in the experimentation of a model: it begins with an idea that is expressed as an experiment, in the attempt to validate or invalidate the idea. Subsequently, this experience is executed, and the information generated is processed. That will inspire the following idea. The more iterations in this repetitive circle, more refined and powerful the ideas will become. He emphasizes the importance of obtaining the maximum amount of information from the experimentation, including the performance of the models.

The experimentation, that is, the application of mathematical algorithms to the data for the extraction of patterns, is called training by NGUYEN (2018). In order to simplify, in the context of this work, the terms training and model will be used to represent the experimentation and the product resulting from the patterns detected in the data.

GREFF et al (2017) confirm the substantial number of computational experiments with several different configurations of hyperparameters and warn against the practical challenge of documentation. Because of the pressure regarding deadlines and the inherently unpredictable nature of a DM project, they verified that there is little incentive for the construction of robust infrastructures and, as a result, the coding will frequently evolve rapidly, which compromises, among other things, the documentation of the project.

BECKER and GHEDINI (2005) emphasize that the structure of a DM process is highly dependent on the methodology adopted, on the abilities, experience and approach of the person responsible for the process, as well as on the resources available in the corporation. Additionally, they confirm the high interactivity of the processes. They emphasize that, even though the conceptual structure of the process may suggest an order between the phases, in practice, the analysts go from one phase to almost any other random phase at any moment. This is also because several problems related to the preceding phases (for example, preparation of data) can only be detected a long time later, when the patterns and the models are evaluated. The same authors also state that generally the projects are developed in a nonstructural manner, ad hoc: after the initial analysis of the data, it is decided to test a specific technique, from which the results may suggest the restructuration of data and the execution of new types of analysis; and so on and so forth. Additionally, over time, it is difficult to remember which training was carried out, which sets of data and which hyperparameters were used, and, more important, the results derived from the combined data. Moreover, according to the authors, this situation leads to recurrent execution of training. They advise that the situation is even worse when taking into consideration long term projects involving several people. They have verified that, regardless of the diversity of knowledge of different people, techniques and instruments, the majority of the DM projects face, in practice, the same difficulties: the wastefulness of redoing the same work and the management of resources and results. The authors state that the documentation of the history of the tasks in face of the iteration and interactivity of the process is an open problem in the management of DM projects.



2.2 DOCUMENTATION: A PATH FOR THE GENERATION OF KNOWLEDGE THAT CAN BE SHARED

GHEDINI and BECKER (2000) emphasized that the documentation of tests and of all relevant parts of a project can not only avoid the loss of the knowledge confined in the minds of the people but can also allow its sharing, turning into a rich source of knowledge for future references and corporate reuse. They also emphasize that this activity leads to a better management of efforts, resources, and results of a DM project.

According to PRAKASH et al (2012), information from training and implementation code (and the history of change) contain an abundance of information regarding the state, progress and evolution of a software project. Additionally, they state that DM is becoming an increasingly important tool to transform these data into information. Similarly, it is expected that the data from DM projects will also be converted into information by the activities of data mining.

WIRTH and HIPPI (2000) emphasized that, perhaps, the greatest benefit of having applied a methodology was the documentation generated. They admitted having initially skipped some tasks of documentation and planning because they would take too long and because they were considered unnecessary by specialists such as them. However, they presented the price that they paid due to this action and they regretted it, realizing that every effort is worthwhile. They reported some benefits observed based on the documentation produced: it avoids the wastefulness of effort (for example, in unfruitful paths or with repetitive work); it promotes efficient management and better communication by the team; it allows for the identification of critical points in the process; it promotes better planning for future projects, based on a more accurate perception of how the effort was spent and of the necessary resources; and it promotes the use of experiences documented in other contexts.

BECKER and GHEDINI (2005) also identified the role of the documentation in learning experiences and in the reuse, and state that an immediate benefit of documentation is the effectiveness in management, in planning, and in communication. They verified that the documentation is completely dependent on the project team because the veracity of the registration and the level of detail that will directly impact its usefulness is their responsibility. Regarding resistance, they state that, insofar as the benefits of the activity are perceived, there is incentive to document the process with more detail and in parallel to the execution of the activities. Additionally, that the best results are reached in the long term, when the project team discovers which strategy is more adequate to its work methods, as well as the best ways to obtain advantages from the resources, techniques and flexibility of the model.

Nevertheless, WIRTH and HIPPI (2000) advise of the difficulty of proceeding to the documentation in the end, and of trying to reconstruct what had been carried out and its motivations. They emphasize that DM processes are live and, given this circumstance, the documentation shall be flexible and live, and should never be updated subsequently to the end of the project (postmortem). They recommend that the definition of a strategy



for documentation must be the starting point, nonetheless, the flexibility for evolution and change must be a premise. They emphasize that to meet the right level of detail for planning and documenting a DM process is difficult and it is part of a long learning process, and it can be influenced by several factors such as the complexity of the project, the duration, and the size of the team.

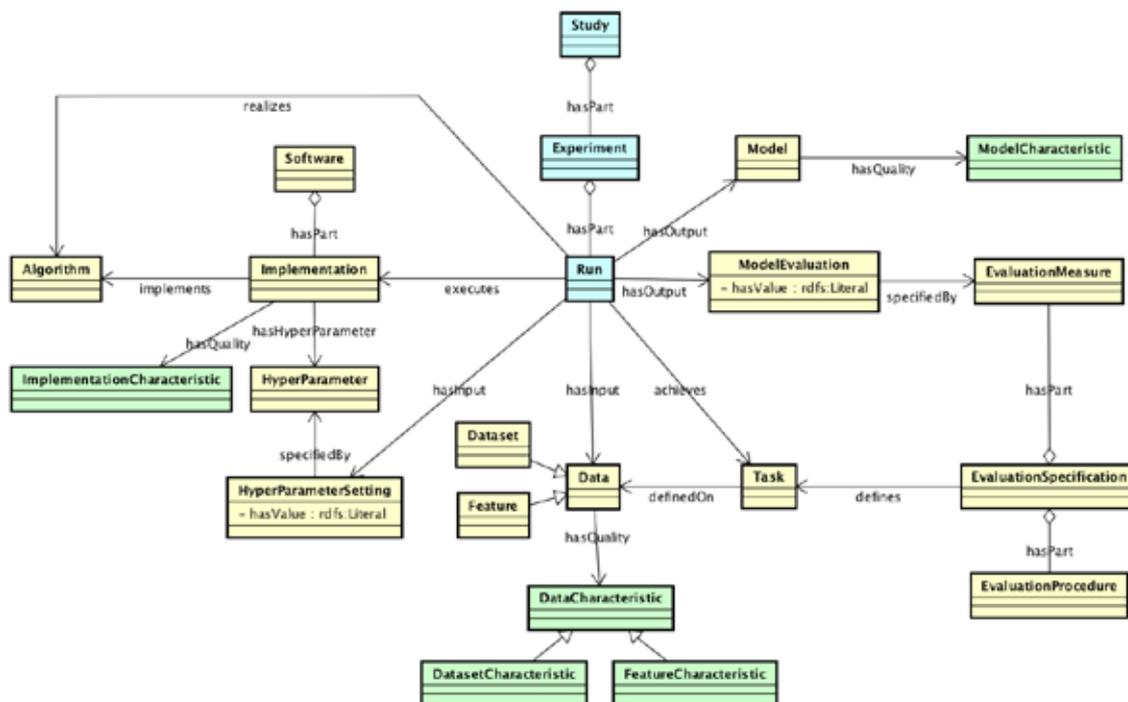
GREFF et al (2017) reference that one of the difficulties for sharing, collaboration, and reproducibility of the training is the use of particular configurations of the workplace by the teams, when using the different tools necessary to discuss the different aspects of a DM process. They mention, among others, databanks, source code management, automatized tools for optimization of hyperparameters, scripts and tables. The same authors presented Sacred, a Python framework of open code that aims to provide basic infrastructure for the execution of computational experiments irrespective of the methods and libraries utilized. They are concentrated on solving problems such as the management of configurations, documentation, and the reproducibility of the results. For each training, relevant information, such as parameters, package dependencies, host information, source code, and results are automatically captured and stored in a centralized repository, where details about the hyperparameters used and the results obtained can also be consulted.

PRAKASH et al (2012) report that several works are being executed in the development of integrated platforms for Machine Learning (ML) and for Software Engineering based on reusable components, quoting, among the most prominent ones with open codes, WEKA and Rapid Miner.

PUBLIO et al (2018) believe that the vision of canonical and standardized models may lead to a better comprehension of ML data and algorithms employed in DM and may promote the interoperability of experiments, irrespective of the platform or of the solution of workflow adopted.

The W3C Machine Learning Schema Community Group (2017) published a ML Schema ontology that provides a set of categories, properties and restrictions for representing and interchanging information about algorithms of machine learning, data sets, and experiments. According to the group, the ontology can be easily extended and mapped for other more specific ontologies of domain, developed in the field of machine learning and data mining. Figure 2 presents the visualization of the main abstractions of the ML Schema.

Figure 2: Main concepts of the ML Schema.



Source: W3C Machine Learning Schema Community Group (2017).

Concisely, the central concept is the training (Run) that, in the context of an experiment (Experiment, Study), produces a model (Model), with its characteristics (ModelCharacteristic), and a set of quality evaluations (ModelEvaluation) that take into consideration standardized metrics (EvaluationMeasure, EvaluationSpecification, EvaluationProcedure). The training is the execution of the implementation (Implementation) of an algorithm (Algorithm) in a platform (Software, ImplementationCharacteristic) using specific parameters configuration (HyperParameter, HyperParameterSetting). The training has as its entrance a dataset (Data, Dataset, Feature) that has its own characteristics (DataCharacteristic, DatasetCharacteristic, FeatureCharacteristic) and they are utilized for a specific task (Task).

KURGAN and MUSILEK (2006) refer to the possibility of integration and of interoperability of DM models using industrial patterns such as the PMML (Predictive Model Markup Language), which represents a model in an XML scheme (Extensible Markup Language). According to the authors, different tools can be used for the generation, visualization, and analysis of the same model.

2.3 KNOWLEDGE IN A DM PROJECT: A POTENTIAL ORGANIZATIONAL LEAP FORWARD

WIRTH and HIPPE (2000) emphasize that the success or the failure of a data-mining project is highly dependent upon the person or upon the team, and that successful practices are not necessarily repeated within the entire company.



CHAPMAN et al (2000) emphasize that data-mining projects can benefit from the experiences from previous projects. The lessons learned during the process and deriving from the solution implemented can give rise to new business debates that are frequently more targeted.

BECKER and GHEDINI (2005) exemplify the benefits derived from the sharing of knowledge between projects and state that the experiences from previous projects can be used to establish plans for more realistic DM projects, with more precise estimates of schedules, budgeting, etc. Experience develops into an easier way to favor more realistic resources, because there is a more extensive understanding of how the effort is actually spent. Furthermore, previous experiences can be used to deal with certain categories of problems or techniques. They advocate that the documentation of the execution of projects must be treated as a corporate resource, which can be shared by the team and used as a reference and can be dependent upon corporative policies and standards.

According to BHATT (2001), while the individuals in the organizations interact with each other, they tend to understand and share their different visions regarding similar situations, structuring their communities and sharing effective work techniques and facilitating the integration of a diversified body of knowledge in the organizations. The same author states that organizational knowledge is formed by unique patterns of interactions between technologies, techniques, and persons, which cannot be easily reproduced, because these interactions are unique from a specific organization, modeled by its history and by its culture. The author attributes the sustainability of the competitive advantages of the company in the long term to incentive given for the enlargement of this knowledge by the creation of a stimulating and practical environment (to learn by practice).

DINGSØYR et al (2001) discuss in an indistinct manner knowledge and experience. They recognize that experience in the strict sense of the word is something that resides in human beings and that it cannot be transferred to others (that each one would have to experiment by herself or himself to acquire the experience). Nevertheless, in a less strict definition, they state that experience is information, which is operational, that is to say, it is usable in some situations. They understand that a description of an event that has happened during a project is one item of experience.

In view of the above mentioned in the introduction, an interesting debate in the management of general knowledge is how to collect, gather, or make explicit experiences in projects in order to be workable for others (DINGSØYR et al, 2001).

NGUYEN (2018) recommends the use of KMP - Knowledge Management Process to have knowledge circulating throughout the entire organization in order to ensure that the right knowledge reaches the right person to understand and have knowledge enough to make decisions and adequately implement tasks. The author states that KMP can be used at any level, from the organization as a whole to each team. The author supplements that the stages (identification, creation, storage, transference, and use of knowledge) are interconnected and are iterative, from the perspective that knowledge is continuously formed and altered. The author finds in the combination between KMP and DM great potential in the exploration and in the management of valuable knowledge of big data. DM provides support to KMP in the generation of inestimable knowledge and KMP provides support to DM in the collection and storage of



knowledge as entrance for DM. Nonetheless, Nguyen emphasizes that there is a large gap in research in this area.

STATA (1980) discusses the concept of OL - Organizational Learning, which occurs by the means of shared insights, knowledge and mental models and that it is based on knowledge and on past experience, that is, in memory. The author elaborates that organizational memory depends on institutional mechanisms (for example, policies and strategies) used to retain knowledge, and it cannot depend exclusively on the memory of individuals, because there is always the risk of losing lessons and experiences laboriously achieved as people migrate from one job to another. Besides other motives for the attrition of personnel (retirement, separation, transfers, demise, etc.), we can also add the risk of forgetting.

DINGSØYR et al (2001) recommend that in the end of a project, an Experience Report is elaborated to collect what has gone right and what has gone wrong in the adopted process. The Review Project activity from the CRISP-DM methodology produces a similar report that, according to the description of CHAPMAN et al (2000): Resumes the important experience acquired during the project. For example, entrapment, deceitful approaches, or information to target data mining techniques more adequate to similar situations could be a part of this documentation.

According to BECKER and GHEDINI (2005), with the documentation of the DM process, as knowledge becomes explicit and managed, it increases the intellect of the organization, developing into the basis for communication and learning, supporting the dissemination of knowledge and the experiences inside the organization in different levels. The authors reinforce that the idea of capturing and storing all informal relevant knowledge generated and used during a DM process, in a way that it becomes available for subsequent recovery, constitutes an interesting approach to deal with the difficulty mentioned about reflecting in documentation the iteration and interactivity of the process.

It is a fact that companies need to educate a larger number of people about the processes and the best practices associated to DM (MARISCAL et al, 2010).

According to NGUYEN (2018), knowledge, which is an expensive commodity for organizations, with different origins, such as documents, processes, persons, communication, culture, and learning, can be incorporated in individuals and in organizations throughout processes and practices. The author states that the transference of knowledge stimulates innovation, and that the storage of knowledge is the path for the creation of an invaluable property for the organizations. An asset that is accumulated over time and that cannot be bought with any amount of money.

In view of what has been presented in the theoretical framework, it can be concluded that the documentation of the memory of a DM project is a small step, since it is located in a smaller context of a project, which can signify a giant step towards organizational memory and profits deriving from shared learning and knowledge managed for an organization. In the next section, the Rastro-DM methodology will be presented as a potential path to achieve this giant organizational step.



3. RASTRO-DM

3.1 GENERAL VIEW

Rastro-DM (Rastro is a translation of trail in Portuguese) is a methodology that aims to document DM projects focused on the process of construction of models, for the purpose of rendering a trail of the executed actions, of the training performed, and of the results obtained and lessons learned. It includes three activities that correspond to the documented concepts:

- Definition of actions;
- Registration of training;
- Synthesis of learning.

Differing from traditional focus, which documents the final product (and its artifacts), the Rastro-DM methodology focuses on the documentation of the process behind the construction of models. Therefore, according to CONKLIN (1996), it promotes an increase in organizational memory, with the registration of the context of the creation of the artifacts: the objectives, the values, the experiences, the motive, the conversations, and the decisions conducted.

As observed in the theoretical framework, all the informal relevant knowledge must be documented to reflect in the documentation the iteration and interactivity of the process and be available for subsequent recovery.

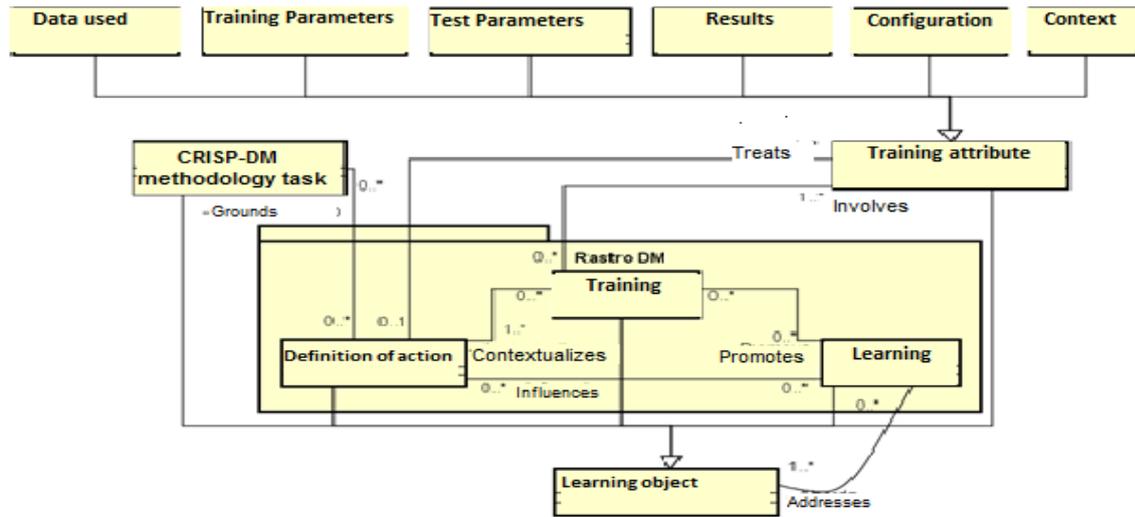
The activities of the Rastro-DM methodology are complementary to the tasks projected in the methodology in use in the organization, independently of which one is used, herein called the baseline methodology, responsible for drawing all the methodological and paradigmatic frameworks used in a DM process. MINGERS and BROCKLESBY (1997) identified different ways to combine methodologies. They state that the establishment of good practices is a way of creating a new methodology; therefore, there is no inconsistency in referring to Rastro-DM as a methodology.

For the purpose of standardization and clarity, the CRISP-DM methodology will be used in the explanations as the baseline methodology, and the steps of the baseline methodology will be referred to as CRISP-DM tasks, and the steps of the Rastro-DM methodology will be referred to as Rastro-DM activities. The Rastro-DM activities will occur several times during a project and can be associated to one or more CRISP-DM tasks.

3.2 INTEGRATED VIEW OF THE CONCEPTS

The concepts of trail are interconnected: the training occurs in the context of a defined action in the project and these trainings can promote learning that eventually can impact, in a virtuous circle, new ideas defined in actions. Some of the possible relationships between the main concepts of the Rastro-DM methodology can be visualized in Figure 3. The diagram is not intended to be complete, but only to provide better understanding.

Figure 3: Main concepts associated to the Rastro-DM methodology



Source: Castro (2019).

It can be seen that the concept of Training attribute, subdivided into six categories, besides being related to training, can be the object for the definitions of action. The definitions of action can be founded on tasks of the baseline methodology, represented in the figure by the CRISP-DM methodology. Additionally, all the elements can be approached as learning object.

In the next section, the activities of the Rastro-DM methodology will be described.

3.3 DEFINITION OF ACTION

Definition of action is the activity of registration of the definition of the steps of a project, whether executed or not. The objective of the activity is not the registration of the details of the execution of an action, but the information concerning its definition, such as the statement of its objective and the techniques to be used or tested when the action is executed.

Conceptually, a definition of action corresponds to the description of one or more specific tasks of the CRISP-DM methodology instantiated within a project.

In the face of the difficulties of management in DM processes identified in the theoretical framework, the possibility of registration of the resources used in the actions defined (people, time, etc.) can be evaluated for the support of deadline estimates and the establishment of schedules of the project. Additionally, the history of these data might serve as the base for the allocation of resources in future projects.

The activity of definition of action can occur at any moment during a DM project. It is important to ensure, at a minimum, the definition of each action initiated, because, in a way, it



justifies the training that will follow. Furthermore, to know what has already been carried out prevents the wastefulness of effort by repeated executions.

The level of abstraction and of detail, when more closely related to a CRISP-DM methodology task, such as, data formatting, or even more detailed, such as, formatting the name of the file for possible removal of stop words and of punctuations, will be the responsibility of each team if there is no corporate standard.

Although the simplicity of a descriptive text field is acceptable, because the content is the most important, the more structured the registration is, the better it will translate the comprehension of the process and the greater the potential of its contribution will be. The entities of the ML Schema that are reported in Figure 2 are examples of attributes associated to a Definition of Action.

To make the concepts clearer, we will use a hypothetical supervised learning project that aims to predict the price of a house (regression) for a real estate agency based in some attributes of the property.

The following examples are definitions of actions of the hypothetical project.

10/10/2018 (registration date in monty/day/year format); Experiment, as the attributes of the property in the model, the number of floors and the number of bathrooms;

11/8/2018; Evaluate if the model can achieve a better performance if the values for the distance from downtown attribute are set in a different scale;

2/1/2019; Test the lightgbm and random forest algorithms for the regression.

The comprehension of the technical terms used in examples of registrations is not relevant for the objective of this paper. The focus, in this case, is the ability of description of a specific action and not the elements of this description (lightgbm, randomforest). Nevertheless, to achieve a more complete understanding of the techniques associated to DM, the book *Data Mining and Analysis: Fundamental Concepts and Algorithms*, by ZAKI and MEIRA (2014) is suggested reading.

3.4 REGISTRATION OF TRAINING

Registration of training is the activity of documentation of the training carried out, of the parameters used and of the results obtained. In accordance with the theoretical framework, the training is the central activity for all the data mining processes, and it is important to retain the maximum of relevant information as possible from them.

The ML Schema presented in Figure 2 discloses a set of entities implicated in training and that are candidates for having their information documented in the context of the activity of registration of training.



Aiming for better comprehension, and with no intention of reaching completeness neither in classification nor in exemplification, the data involved in a training can be grouped into six categories:

- Data used: test data, training data, variables used as features for the model;
- Training Parameters: algorithms used, hyperparameters considered, implementation of techniques applied;
- Test parameters: metrics considered, forms of assessment;
- Results: model with its characteristics and the calculated metrics;
- Configuration: identification of the program, versions of the libraries used, data of hardware;
- Context: training code, date and hour, number of training periods, error message in case of interruption of the execution.

The level of detail of documentation has an impact in its utility. The code of training, a datum of the context category, can be used as identification key of the model in training.

These are examples of registrations of training of the hypothetical project for the prediction of the price of a house:

- Code: 1; Registration date: 6/7/2018; Variables used: house square footage, property square footage and Postal code of the address; Algorithm used: linear regression; Error: 0.8; Separation of test data: 10%, not stratified;
- Code: 100; Registration date: 7/7/2018; Variables used: house square footage, property square footage, Postal code of the address, number of rooms and date of the construction; Algorithm used: randomforest; Error: 0.7; Separation of test data: 5%, stratified data.

3.5 SYNTHESIS OF LEARNING

This is the activity of synthesis and registration of the lessons learned throughout the project, in an automatic manner or not.

The activity may occur at any moment of a DM project and can be associated or not to the training. Although they can be synthesized in the stage of pre-development of the model, in the CRISP-DM phases of business understanding and data understanding, learning, for the most part, is a result of the training runs. It is possible to outline, automatically or not, that a specific selection of variables or the use of a specified hyperparameter of a technique led to the generation of a model of better performance. There may be lessons learned that involve training that failed that aim to document how to avoid repeating errors.



Information about ML Scheme entities (Figure 2) and Defined Actions can be shown as attributes of a lesson learned.

The synthesis of learning and its effective use in the project or in future projects promote the improvement of the team in the DM process, in knowledge regarding the high level of iteration and interactivity of its tasks, regarding the different techniques and regarding the tools used.

The documentation of lessons learned prevents it to get lost in the memory of the individuals or even with the individuals when they leave the project or the organization.

To make clear the concept, we list below examples of lessons learned on the hypothetical project of price prediction of a house associated to the CRISP-DM task, as follows:

10/3/2018; Select technique (CRISP-DM task); the randomforest technique has proven itself superior to the decisiontree technique in the context evaluated;

5/26/2019; Format data; It is necessary that the values of the real estate in the data used for the training are updated to the same monetary reference;

7/26/2019; Select data; The increment of the variables relative to the number de rooms and number of parking places for automobiles promoted an increase of 10% in the accuracy of the model.

3.6 DIRECTIONS

In the following section, there are some practical directions for the effective and efficient application of the Rastro-DM methodology.

3.6.1 Adaptability

The activities of the Rastro-DM methodology are complementary to the tasks projected in the baseline methodology, which contains the entire methodological and paradigmatic framework used in a DM process. The documentation of the trail, apart from the strict context of a baseline methodology, enables a better approach of the interactivity and the iteration of the tasks, something that has not been properly mapped by DM methodologies, as discussed in the Theoretical Framework. Since the tasks and phases of a DM process merge with each other and are often executed concurrently, it is difficult to maintain an effective documentation by task of the baseline methodology. For example, if in the context of a CRISP-DM task a report is produced and subsequently we return to this task several times, the documentation would need to be constantly updated, in a way restricted to the scope of the task. The cost of the reworking makes in-time documentation impeding in the CRISP-DM methodology.



The documentation generated by the Rastro-DM methodology can connect itself to the output artifacts of the baseline methodologies. Later, the registrations can, for example, be grouped by CRISP-DM task. For example, a report named Reasons for exclusions and selections in the activity called Selection of data of the Preparation of data of CRISP-DM –phase, can be automatically generated from lessons learned and definitions of action related to data selection, supplemented by a summary of the training related to each criterion of selection experimented.

3.6.2 Timeliness

The registration must be in time to aid in directing the project in progress in a more efficient manner. With the adequate registration of the trail, inefficiency is avoided with repeated executions of tasks in the course of the project. As previously observed, a postmortem documentation is not adequate for DM projects, because they are considered live due to their complexity and iteration.

3.6.3 Flexibility

The definition of the attributes to be stored must be flexible and may vary according to the objective of mining (classification, regression, etc.), the training platform, and the objectives of the documentation. Exemplifying, if the documentation aims for the reproducibility, more details of software configuration (versions of libraries) and of random seeds used must be recorded.

Rastro-DM is flexible because it does not define a minimum proportion of attributes for each concept. After all, each project and organization has its own complexity and maturity level in DM. In a final analysis, it is the responsibility of the team to certify the veracity and the effectiveness of the documentation. It is expected that, with the improvement in DM, the organizations will be able to develop a minimum corporative standard for documentation per DM objective. However, this standard should not inhibit team creativity or merely become a burden.

3.6.4 Automation

It is indispensable that the activity of training registration is automatic and is connected to the development of training of model in the platform currently used. Although there may be an initial expense for the structuring of a software framework for a specific configuration of tools, this effort will be implemented only once, and the subsequent projects will benefit from it. As previously observed, GREFF et al (2017) advise on the practical challenge when the construction of such a structure is not incentivized.

It is desirable that all the activities are carried out integrated to the platforms, so that the effort is not a barrier for the documentation. After all, in order to become useful, a model of documentation must correspond, as much as possible, to the way people work (BECKER and GHEDINI, 2005).

Castro (2019) illustrates a simple implementation of an infrastructure of construction of a trail in python, created at a low cost with local files.



3.6.5 Usability

In reference to the definitions of action and lessons learned, their registration must be prioritized over categorization, in a way that this activity is not converted into an obstacle for their documentation. The registration can occur initially using even a free text format.

The categorization, for example, relative to the techniques used, must be carried out automatically and can be an important corporate requirement to empower the utility of the trail by means of the sharing of knowledge. As seen, the combination of DM with KMP empowers the increase of organizational knowledge (NGUYEN, 2018).

We must also consider the usability of the applications that query the trail, which tends to become corporate. Searches for information, for example, can lead to projects that have experimented a specific technique, identify information of its use, and evaluate the results attained, serving as a basis for new projects and for the definition of DM strategies in the organization.

In the next section, the application of the Rastro-DM methodology in a DM project will be illustrated.

4. THE CLADOP PROJECT – ILLUSTRATION OF THE USE OF THE RASTRO-DM METHODOLOGY

The Cladop² project is an example of utilization of the Rastro-DM methodology. The term Cladop is used both to indicate the process (project) and to refer to the product developed (Classifier of PDF Documents). The project consisted of the development by supervised learning of an automatic classifier for PDF (Portable Document Format) documents inserted in the system for the management of Auditing of Special Accounts (e-TCE) of the Federal Court of Accounts of Brazil (TCU).

The process of Auditing of Special Accounts (TCE) ultimately seeks to reimburse the Public Treasury of damages generated by public agents, holding them accountable for these damages. With the objective of making the process of damage assessment faster and more efficient, the e-TCE system was developed by the Federal Court of Accounts of Brazil in cooperation with the Comptroller General of Brazil (CGU) through the standardization and optimization of procedures. It is a unique platform that provides access to all the Public Administration departments that act in some phase of the process of Auditing of Special Accounts. Castro (2019) details the business context and the related regulations.

The classifier developed provided benefits to businesses regarding both the usability of the system, by automatically identifying the types of documents for users, and of the data

² After finalizing the Cladop Project for the e-TCE system, which is the presented in this paper as the “Project”, the construction of a similar classifier was initiated for another system: Electronic Protocol. The new project, still in development, was named Cladop@Protocolo. Therefore, the Cladop Project presented in this paper shall be known as Cladop@eTCE.



quality, by assuring greater correction of the types and higher quality of text from the documents generated by OCR (Optical Character Recognition). Documents with a correct classification of their type and with text content of higher quality are fundamental for the procedural operations assisted by machine.

4.1 UNDERSTANDING THE DATA

During the construction of the Classifier, 118,266 documents were considered, relative to 4,384 damages that were distributed into 84 types of document. The type “Others” was the most used, corresponding to 18.81% of the documents. Graphic 1 presents a partial visualization of the severe unbalance of the quantity of documents between active types.

Graphic 1: Partial view of the imbalance of documents per type (reference: 4/17/2019).



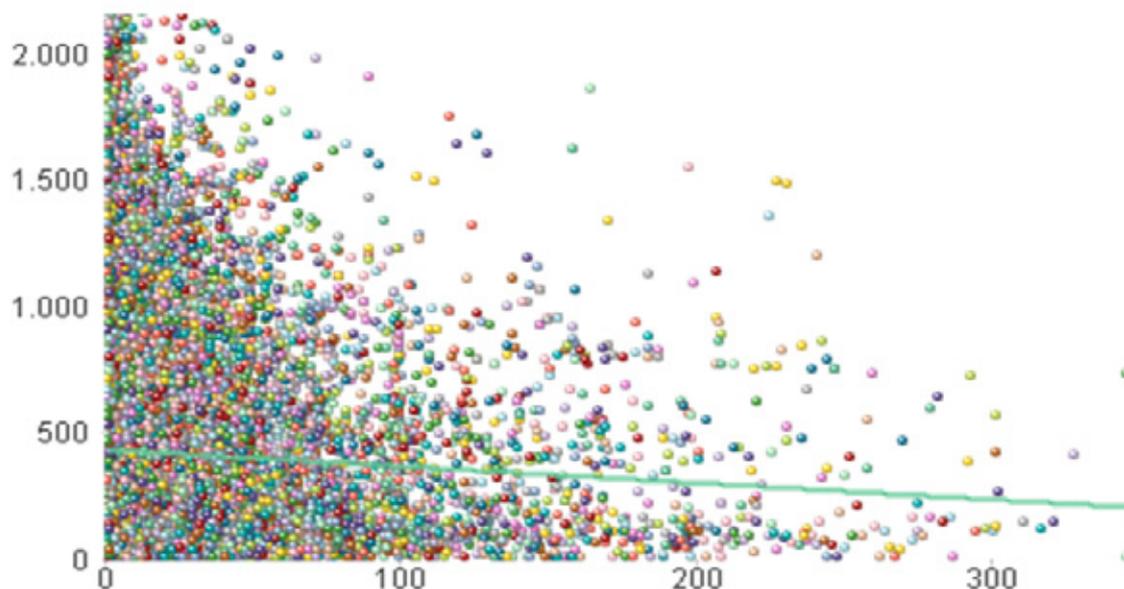
Source: Castro (2019) – partial view.

BRANTING (2017) states that the PDF format has been used in courts and that the text obtained from these documents present several errors and do not preserve the original sequence of the document, due to the OCR process used.

The outreach of the results expected by the e-TCE system depends on the quality of the documents coded in the system. The efficiency of the classifier also tends to be superior if the data have greater quality. Low quality was found in the OCR of the documents. An example of this situation is protocol document number 58.900.414 that has 162 pages, but via OCR it is impossible to identify even one hundred valid words in it, corresponding to less than one valid word per page. The low quality of the OCR can be illustrated on Graphic 2 that presents the number of valid words per page for some documents.



Graphic 2: Documents – quantity of valid words per page (y-axis) versus number of pages (x-axis).



Source: the authors (2020).

According to Castro (2019), the pre-processing of the text contained in the documents was tested in different combinations of methods for obtaining text from the document (OCR algorithms), of criteria for the definition of valid words, and of substitute classes of words considered (Brazilian individual taxpayer registry identification - cpf, date, name of natural person, etc.).

4.2 FUNCTIONAL DESCRIPTION OF THE CLASSIFIER

The Cladop project, in its version 2.1, has achieved a level of accuracy of 91.1% with a standard deviation of 0.3%, with cross-validation of 7 partitions and 2 repetitions, totaling 14 samples. In the context of this paper, the term accuracy must be understood as micro-accuracy, which takes into consideration the results, scores and errors, per document independently from its type.

The classifier was implemented in python in the form of a web service that receives a PDF file as its parameter, and based on its name and on its content, it returns the nine most probable types, each with their respective probabilities. The accuracy of the classifier³ increases from 91% (for the first type) to 99% when taking into consideration the first nine predictions. The return of more than one type makes it possible for the system, for example, to present the other types in a second screen in case the first type is not accepted by the user.

Table 1 presents, for some types of documents, some metrics verified based on the validation data, 5% of the total, when the version in production of the classifier was generated. The general

³ After writing this paper, there was the inclusion and exclusion of document types and, taking into consideration more recent documents, version 3.0 of the Cladop Project was developed, achieving a 93.6% accuracy level with a 0.3% standard deviation.



numbers in validation data are: accuracy: 90.99%; macro-precision: 81.04%, and weighted-precision: 91.06%; macro-recall: 80.29%, and weighted-recall: 90.99%; F1-macro: 80.16% and F1-weighted: 90.87%.

Table 1: Metrics verified per type regarding validation data (partial view).

Description	Precision	Recall	F1	Documents
Lawsuit - initial petition	93.33%	93.33%	93.33%	15
Final ruling of higher court	85.71%	85.71%	85.71%	7
Analysis of accountability	77.27%	62.96%	69.39%	27
Analysis of defense	83.33%	52.63%	64.52%	19
Suspension of delinquency rate	73.68%	82.35%	77.78%	17
Term of concession and acceptance of scholarship and attachments	75.00%	54.55%	63.16%	11
Term of definite receipt of construction work	100.00%	100.00%	100.00%	4

Source: Castro (2019) – partial view.

In addition to the predictions, the classifier returns information derived from the pre-processing of the file's text, which may be useful so that the e-TCE system can demand a minimum quality of text content in the documents at their registry, such as: quantity of valid words, quantity of values, quantity of names, etc. Therefore, the system can prevent low OCR quality with a high percentage of invalid words or even incomplete content in documents that are critical for the TCE process, such as the case of an Acknowledgement of Receipt document that has no date and no taxpayer registry identification.

4.3 APPLICATION OF THE RASTRO-DM METHODOLOGY IN THE PROJECT

The utilization of the Rastro-DM methodology has shown to be an effective solution for the difficulty of reflecting in the documentation the interactivity and the interactivity of a DM process as presented in the theoretical framework.

The following sections exemplify the use of the methodology in the project and the benefits obtained. The details of the trail in the Cladop project (Rastro-DM@Cladop) and the code used for its construction are available at <https://gitlab.com/MarcusBorela/rastro-dm.git>.

4.3.1 Definition of action

The first definitions of action were registered as comments in the implementation code of the training. Nevertheless, throughout time, they became more complex and broader, and they crossed over different codes. Therefore, they had to remain in a databank. Definitely, the registration of definitions of action cannot rely upon coding. Due to the unpredictable nature of a DM project, the code often evolves rapidly and ends up compromising, among other things, its documentation (GREFF et al, 2017). Table 2 illustrates some definitions of action registered throughout the project.



Table 2: Examples of definitions of action registered in the Cladop project.

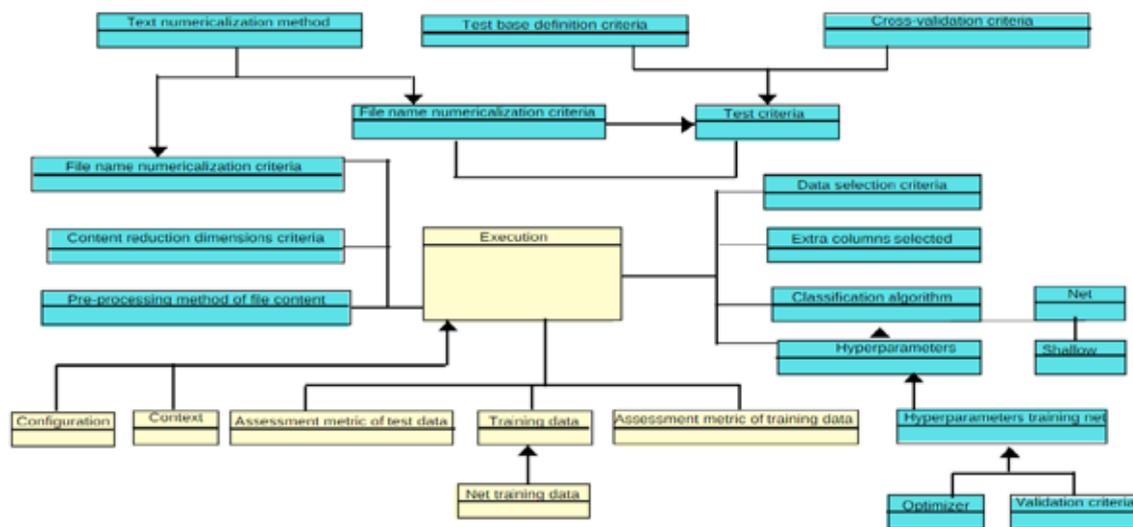
Registration date	Description	CRISP-DM Task
3/12/2019-17:04	Creating a structure (code and data) for processing k-fold in shallow algorithms	Project of tests
3/14/2019-19:27	Initiated the executions to experiment optimizers (MLP): nadam, adadelta	Construct Model
3/18/2019-11:40	Experimenting with MLP columns no longer binary	Format Data
3/26/2019-11:57	Initiating the inclusion of names of files in the model	Select Data
5/29/2019-19:30	Program modified (shallow) to also record recall and f1-micro	Project of tests
6/4/2019-19:35	Program modified (shallow) to not record recall and f1-micro, because they are equivalent to accuracy (and to precision)	Project of tests

Source: Castro (2019) – partial view.

4.3.2 Registration of training

The dataset structure in the databank relative to training evolved throughout the project along with the maturing of the team in DM. Figure 4 presents the main stored concepts that can be related to the subtypes of Figure 3. The concepts that are the entrance for the training are represented in light blue and the data derived from the process, in yellow.

Figure 4 – Training information registered in the Cladop project, grouped by concept



Source: Castro (2019).

Table 3 is merely illustrative and presents some attributes of these entities and their respective values for the training that tested and created the model deployed in version 2.1 of the Cladop project.

The training that generated version 2.1 of the model had as its code 13.138 (in accordance with the second line of Table 3), which demonstrates that more than thirteen thousand model training activities



were executed. Some training activities involving only high-quality OCR documents and types with great quantity of documents led to classifiers with accuracy superior to 98% on test data. This is the case of training code 147, on 11/24/2018, which included 11.110 documents of better quality of OCR from the four most frequent types of documents. However, during the project, we opted for the development of a classifier that would include all the types of documents with less restrictive criteria.

Table 3 – Trail of the training for testing and for the creation of the model in version 2.1 of the Cladop project.

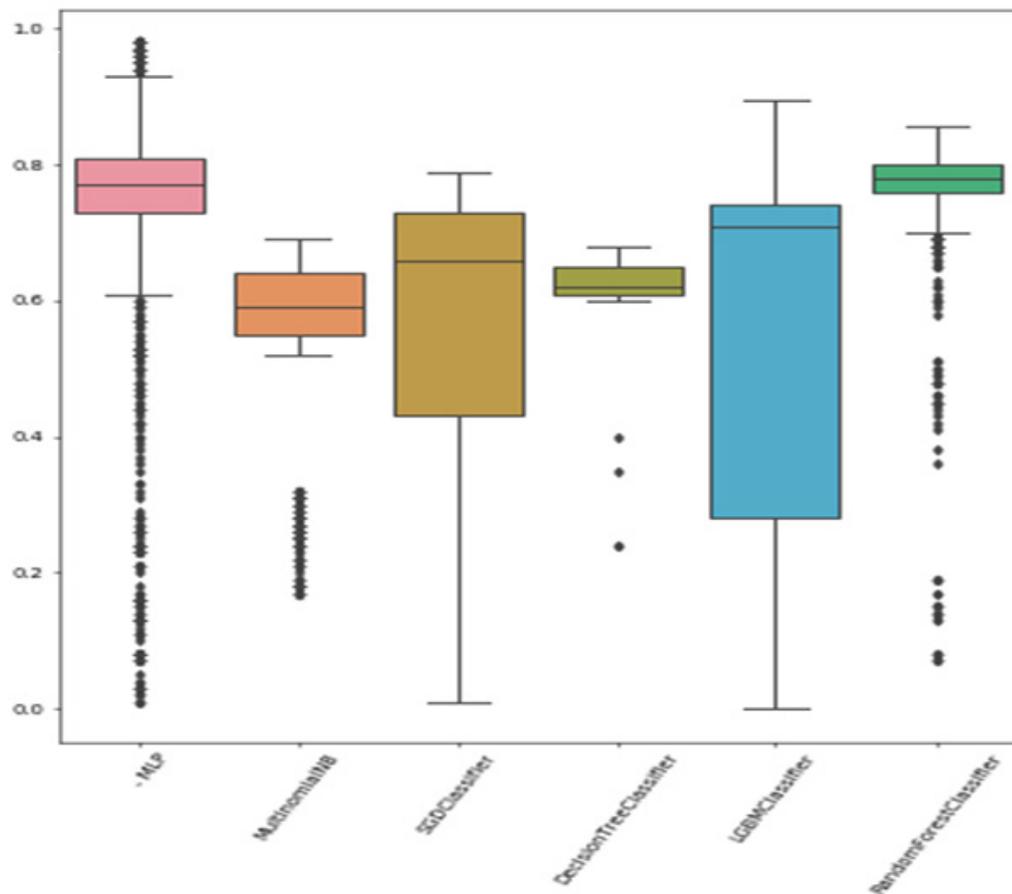
Group	Item	Test	Generation
Context	Code (unique key)	13.134	13.138
	Date and hour of registry	07/05/2019-00:44:59	07/05/2019-18:02:12
	Number of periods trained	33	24
	Quantity of documents used	63.468	
	Quantity of types processed	81	
	Time of execution (in seconds)	460	475
Configuration	Program executed	Cladop_monitoramento.ipynb	
	Mode of text conversion for numbers	Tfidf	
	Number of words of the “bag of words” model of the text	24.576	
	Number of words of the “bag of words” model of the file name	1.000	
	Method of text pre-processing	7	
Parameters	Criteria for data selection	Documents of a different type from Others and created after 5/1/2018	
	Method of test	Cross-validation with 7 partitions and 2 repetitions	There was no data reserved for testing
	Optimizer used for training the network	Adam	
	Size of the training batch	256	
	Minimum quantity required by type	0	
	Final dimension of the content vector after reduction	768	
	Algorithm for size reduction	TruncatedSVD	
	Were the data shuffled in every epoch?	No	
	Criteria for validation data	5% of the data, without stratifying and with shuffle	
	Result	Accuracy verified in test data	91.1% +- 0.3%
Accuracy verified in training data		96.4% +- 0.7%	95.6%
Accuracy verified in validation data		90.8% +- 0.6%	92.1%

Source: Castro (2019) – adapted version.



Based on the training data, several analysis can be achieved. An example is the comparison of the performance of the models relative to the algorithms used. The models with higher accuracy in the final selection of data and parameters were implemented with neural networks (MLP network). Models implemented using the LGBMClassifier algorithm scored just behind, with a difference of approximately 2%. Graphic 3 illustrates the accuracy in test data achieved by some of the algorithms experimented. The large variation in accuracy by algorithm is explained by the variation of combination of parameters experimented in training. The understanding of these algorithms is not in the scope of this paper. To those interested in these data, we suggest the reading of the book Data Mining and Analysis: Fundamental Concepts and Algorithms from ZAKI and MEIRA (2014).

Graphic 3 - Boxplot with accuracy in test data of the algorithms experimented.



Source: Castro (2019) – partial view.

4.3.3 Synthesis of learning

From the lessons learned registered, the majority could be synthesized by data mining in the training dataset. For example, based on equal value queries in some columns, the leaning dated 4/10/2019 in Table 4 could be detected. Others depend on human intervention, such as the one dated 2/27/2019. The automatic generation of knowledge reveals itself to be a promising path for



the improvement of organizational knowledge, which was already attested by NGUYEN (2018), that, as previously observed, emphasizes that there is a large gap in research in this area.

After less than six months after the end of the project, Castro (2019) attempted to count, by sampling, how many of these still remained complete in his memory, approaching an estimate of only 27%. Unfortunately, the remainder were either incomplete or had been lost from the author's memory. However, because of the trail, they were not lost from the memory of the project.

BECKER and GHEDINI (2005) observed that forgetting can lead to the repeated execution of experiments. Throughout the project, there was forgetting that led to unnecessary efforts in the execution of tasks. The lesson dated 4/10/2019 was ignored leading to unnecessary execution of the action to also record the metrics dated 5/29/2019 (see Table 2), which caused rework with the alteration of the program in order to not record these metrics on 4/6/2019. A lesson learned was that it is not enough to have a trail, it is also necessary to use it.

Table 4 – Examples of lessons learned registered during the project

Registration date	Description	Context CRISP-DM activity
2/27/2019-10:32	The variables of context and PDF quality are enough to achieve a result of 44% accuracy, with optimizer Adam. With Adagrad:42%; SGD:28%.	Select Data
2/27/2019-10:39	By using pca (sklearn): it is best to separate the fit command from the transform command. The fit_transform command was locking!	Format Data
2/27/2019-10:53	By using pca (sklearn): if the number of dimensions is very small and the array is wide (see details in the documentation of function), it is better to use the randomized method rather than the full method, because it is faster and the results (variance achieved) are equivalent.	Format Data
03/29/2019-09:54	An improvement of approximately 5% in the accuracy of the models was noticed after including the file name as an attribute.	Select Data
4/1/2019-15:36	For shallow algorithms, the extra columns with no dummy values (one only column with various discrete values) led to a better result. For a Neural Networks, there is a slight improvement using dummy values.	Format Data
4/10/2019-10:49	In multiclass classification, the metrics referring to f1_micro, recall_micro, precision_micro and accuracy are equivalent.	Project of tests
5/22/2019-20:04	The Adam optimizer (passing object with Amsgrad=False) was the best optimizer so far. Better than the Amsgrad=True in 0.1%, and than the Rmsprop and the Adadelta in 0.2% in the context evaluated in the last executions.	Construct Model

Source: Castro (2019) – partial view.



5. BENEFITS ATTAINED

The benefits using the methodology in the project can be grouped into 3 aspects: project management, technical maturation of the team, and automation of activities.

Within project management, we could perceive improvement in communication, in planning, and in the reduction of expenses.

The trail made possible a continuous managerial follow-up of the project's evolution and of the performance of the models generated. The use of standardization in the submission of follow-up reports inspired reliability in users and sponsors.

The trail has become a foundation for better planning of future similar projects; after all, data mining projects can benefit from the experiences of previous projects (CHAPMAN et al, 2000). The narration of the experience contained in the trail of processes makes it possible to advocate for more realistic resources, because there is already a broader understanding of how effort is actually spent. Additionally, it enables the identification of critical points in the process.

A decrease in effort for repeated work was observed, even through unproductive paths.

The benefits of the use of the Rastro-DM methodology in management led to its inclusion as a technical requirement for outsourcing development of DM projects that has been initiated in the TCU.

The technical maturity of the DM team was increased due to the environment created, the improvements achieved, and the sharing of learning.

A stimulating and practical environment was promoted (to learn by practice). The team felt encouraged to carry out documentation in parallel with the execution of activities due to the facility in registration and the benefits perceived throughout the project, such as documentation itself.

The synthesis of learning and the effective use of these activities in the project as well as the structural evolution of the registrations of training have improved the maturity level of the team in DM. As a better comprehension and understanding of the process and of the techniques included was achieved, new fields were created for storage and the contents became more valuable.

Regarding sharing, the Cladop trail was published in the context of a Data Analysis Specialization course (Castro, 2019) and the methodology was presented in 2019 at the 5th International Conference on Data Analysis for Public Administration.

We expect that this discussion advances within the organization, relative to policies and standards to transform the trail into a corporate resource to support the dissemination of knowledge on various levels. Simple enquiries, for example, could meet projects that have applied a specific technique, and identify information from its use and assessment of the results achieved. Managerial enquiries could support DM strategies for the organization.



Another benefit was the automation of routines, which ensures a minimum quality of activities as well as reduces the cost of its realization. The chronological list of lessons learned, and definitions of action supplemented by some graphics automatically generated from the training can form automatic reports relative to tasks of the baseline methodology, as illustrated in Castro (2019). The same author demonstrates that the trail made it possible to automatically monitor the performance of the models against new documents registered in the system. Even more, this monitoring routine has promoted the automatic generation of version 2.1 of the Classifier based on more updated data.

6. CONCLUSION

The objectives proposed for this work were achieved. A brief contextualization by the theoretical framework was presented regarding the methodologies and documentation in data mining projects, as well as the potential impact of the experiences acquired by an organization during the projects.

Rastro-DM was proposed as a methodology for the documentation of DM projects with a focus on the process that, among other characteristics, is flexible and can be merged into the methodology used by an organization.

The Rastro-DM methodology has revealed itself viable, with beneficial results according to what was illustrated in its application in the Cladop project.

There is still a lot to be achieved. Future works may experiment the Rastro-DM methodology in other contexts: other objectives of mining rather than classification, other development platforms, or even other organizational cultures. The methodology can be incrementally reviewed with new activities or even improved by techniques and procedures that supplement its activities. Studies can be carried out for the mapping of policies and standards for trails to be used as a support for contracting DM projects. The mining of learning based on training data is also a challenge for future works because there is a large gap in research about the integration between DM and KMP (NGUYEN, 2018). Additionally, maybe the most important point for evolution, which was not the scope of this work, is how to transform a trail into a corporate resource that allows the sharing of experience and the use of institutional approaches for the development of organizational knowledge. After all, the value of the data remains in how they are interpreted and used (BERMAN et al, 2018).

Finally, it is evident that, besides contributing to the project's team, project trail has the potential of promoting a quantum leap in organizational knowledge, if institutional measures are adopted towards incentivizing its creation, its sharing, and the automatic searching of the learning contained in them. Nonetheless, we must be mindful that no corporate strategy should constrain the liberty of the data analysts, because it diminishes their creativity. And the creative freedom for developing their own trails is something that humans cannot do without, since it is one of the traits that prevents us from being classified as machines.



REFERENCES

BECKER, Karin; GHEDINI, Cinara. **A documentation infrastructure for the management of data mining projects**. Information and Software Technology, v. 47, n. 2, p. 95-111, 2005.

BERMAN, Francine et al. **Realizing the potential of data science**. Communications of the ACM, v. 61, n. 4, p. 67-72, 2018.

BHATT, Ganesh D. Knowledge management in organizations: examining the interaction between technologies, techniques, and people. Journal of knowledge management, v. 5, n. 1, p. 68-75, 2001.

BRANTING, K. **Automating Judicial Document Analysis**. ASAIL@ICAIL. 2017.

CASTRO, Marcus. **Mineração de Dados com Rastro: Boas Práticas para Documentação de Processos e sua Aplicação em um Projeto de Classificação Textual**. Conclusion paper of the specialization program on Data Analysis for Auditing. Brazilian Federal Court of Account's Serzedello Corrêa Institute, Brasília, Brazil. Available at: <https://portal.tcu.gov.br/biblioteca-digital/mineracao-de-dados-com-rastro-boas-praticas-para-documentacao-de-processo-e-sua-aplicacao-em-um-projeto-de-classificacao-textual.htm>. Accessed on July 26, 2020. 2019.

CHAPMAN, Pete et al. **CRISP-DM 1.0: Step-by-step data mining guide**. SPSS inc, v. 16, 2000.

CHOLLET, Francois. **Deep Learning with Python**. Manning Publications Co., Greenwich, CT, USA. 2017.

CONKLIN, Jeffret. **Capturing Organizational Memory**. In: Groupware and Computer-Supported Cooperative Work, R.M. Barcker (Ed.), Morgan Kaufman, pp. 561-565. 1996.

DINGSØYR, Torgeir; Moe, Nils Brede; Øystein. Nytrø. **Augmenting experience reports with lightweight postmortem reviews**. Lecture Notes in Computer Science, 2188:167-181, 2001.

GHEDINI, Cinara; BECKER, Karin. **KDD application management through documentation**. Available at: https://www.researchgate.net/profile/Karin_Becker2/publication/268253354_KDD_application_management_through_documentation/links/5657a5ec08ae1ef9297bf1d1/KDD-application-management-through-documentation.pdf. Accessed on July 27, 2019. 2000.

_____. **A documentation model for KDD application management support**. In: SCCC 2001. 21st International Conference of the Chilean Computer Science Society. IEEE. p. 105-114. 2001.

GREFF, Klaus et al. **The sacred infrastructure for computational research**. In: Proceedings of the Python in Science Conferences-SciPy Conferences. 2017.

KURGAN, Lukasz A.; MUSILEK, Petr. **A survey of Knowledge Discovery and Data Mining process models**. The Knowledge Engineering Review, v. 21, n. 1, p. 1-24, 2006.



MARBÁN, Óscar *et al.* **An engineering approach to data mining projects.** In: International Conference on Intelligent Data Engineering and Automated Learning. Springer, Berlin, Heidelberg, p. 578-588. 2007.

MARISCAL, Gonzalo; MARBAN, Oscar; FERNANDEZ, Covadonga. **A survey of data mining and knowledge discovery process models and methodologies.** The Knowledge Engineering Review, v. 25, n. 2, p. 137-166, 2010.

MINGERS, John; BROCKLESBY, John. **Multimethodology: Towards a framework for mixing methodologies.** Omega, v. 25, n. 5, p. 489-509, 1997.

NGUYEN, Ngoc Buu Cat. **Data Mining in Knowledge Management Processes: Developing an Implementing Framework,** 2018.

PRAKASH, BV Ajay; ASHOKA, D. V.; ARADHYA, VN Manjunath. **Application of data mining techniques for software reuse process.** Procedia Technology, v. 4, p. 384-389, 2012.

PUBLIO, Gustavo Correa *et al.* **ML-Schema: Exposing the Semantics of Machine Learning with Schemas and Ontologies.** arXiv preprint arXiv:1807.05351, 2018.

STATA, Ray. **Organizational learning: The key to management innovation.** Massachusetts Institute of Technology, 1980.

W3C (World Wide Web Consortium) Machine Learning Schema Community Group. **W3c machine learning schema.** Available at: <https://www.w3.org/community/ml-schema>. Accessed on July 30, 2019. 2017.

WIRTH, Rüdiger; HIPPE, Jochen. **CRISP-DM: Towards a standard process model for data mining.** In: Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining. Citeseer, p. 29-39. 2000.

ZAKI, Mohammed J.; MEIRA, Wagner. **Data mining and analysis: fundamental concepts and algorithms.** Cambridge University Press, 2014.

Os conceitos e interpretações emitidos nos trabalhos assinados são de exclusiva responsabilidade de seus autores.





Destaques de decisões do TCU

2020

- ▶ A produção de efeitos de medidas legislativas que cuidem de concessão ou ampliação de incentivo ou benefício de natureza tributária da qual decorra renúncia de receita depende do atendimento prévio, pelo Poder Executivo, das condições estabelecidas no ordenamento jurídico, em especial na Lei de Responsabilidade Fiscal e na respectiva Lei de Diretrizes Orçamentárias.
 - Acórdão 62/2020-Plenário | Relator: BRUNO DANTAS
 - ÁREA: Finanças Públicas | TEMA: Renúncia de receita | SUBTEMA: Requisito
 - Outros indexadores: LDO, Eficácia da lei, Responsabilidade fiscal

- ▶ Não compete ao TCU deliberar sobre restituição de lucro espúrio obtido por empresa contratada por meio de fraude a certame licitatório, pois isso importaria na aplicação de pena de perdimento de bens ou valores acrescidos ilicitamente ao patrimônio de particular, sanção não prevista na Lei Orgânica do Tribunal. O TCU não pode aplicar sanções sem a devida previsão legal, nos termos do art. 5º, incisos II, XXXIX e XLVI, da Constituição Federal.
 - Acórdão 129/2020-Plenário | Relator: BENJAMIN ZYMLER
 - ÁREA: Competência do TCU | TEMA: Contrato administrativo | SUBTEMA: Abrangência
 - Outros indexadores: Fraude, Sanção, Lucro, Licitação

- ▶ SÚMULA TCU 230 (ALTERAÇÃO): Compete ao prefeito sucessor apresentar a prestação de contas referente aos recursos federais recebidos por seu antecessor, quando este não o tiver feito e o prazo para adimplemento dessa obrigação vencer ou estiver vencido no período de gestão do próprio mandatário sucessor, ou, na impossibilidade de fazê-lo, adotar as medidas legais visando ao resguardo do patrimônio público.
 - Acórdão 206/2020-Plenário | Relator: RAIMUNDO CARREIRO
 - ÁREA: Responsabilidade | TEMA: Convênio | SUBTEMA: Gestor sucessor
 - Outros indexadores: Omissão no dever de prestar contas, Súmula, Prestação de contas

- ▶ Os valores das contribuições previdenciárias dos servidores civis e militares mencionados no artigo 21, inciso XIV, da Constituição Federal (polícia civil, polícia militar e corpo de bombeiros militar do Distrito Federal), indevidamente apropriados pelo Distrito Federal, no período de janeiro de 2003 a agosto de 2016, devem ser ressarcidos ao Fundo Constitucional do Distrito Federal.
 - Acórdão 1244/2020-Plenário | Relator: WALTON ALENCAR RODRIGUES
 - ÁREA: Finanças Públicas | TEMA: FCDF | SUBTEMA: Tributo
 - Outros indexadores: Recolhimento, Credor, Retenção, Contribuição previdenciária

- ▶ O servidor afastado ou licenciado de seu cargo efetivo sem remuneração, não optante pela manutenção do vínculo com o Plano de Seguridade Social do Servidor, não faz jus, assim como seus dependentes, aos benefícios do aludido regime previdenciário, inclusive a pensão por morte (art. 183, §§ 2º e 3º, da Lei 8.112/1990), salvo se beneficiário da vantagem prevista no art. 40, § 19, da Constituição Federal e nos arts. 2º, § 5º, e 3º, § 1º, da EC 41/2003 (abono de permanência).
 - Acórdão 1408/2020-Plenário | Relator: AROLDO CEDRAZ
 - ÁREA: Pessoal | TEMA: Afastamento de pessoal | SUBTEMA: Regime Próprio de Previdência Social
 - Outros indexadores: Pensão, Vínculo, Consulta, Abono de permanência em serviço, Benefício previdenciário

- ▶ Identificada a celebração de acordo de leniência em outras instâncias de controle posteriormente ao trânsito em julgado de acórdão do TCU que declarou a inidoneidade de empresa para participar de licitações na Administração Pública Federal (art. 46 da Lei 8.443/1992), admite-se a suspensão da eficácia da sanção aplicada pelo Tribunal, condicionada ao cumprimento dos termos do acordo.
 - Acórdão 1689/2020-Plenário | Relator: BENJAMIN ZYMLER
 - ÁREA: Responsabilidade | TEMA: Declaração de inidoneidade | SUBTEMA: Acordo de leniência
 - Outros indexadores: Suspensão, Sanção, Trânsito em julgado



- ▶ O TCU está sujeito ao prazo decadencial de cinco anos, previsto no art. 54 da Lei 9.784/1999, para a apreciação da legalidade dos atos de concessão inicial de aposentadoria, reforma ou pensão, a contar da chegada do processo ao Tribunal, conforme a decisão do STF no julgamento do RE 636.553 (Tema 445 da Repercussão Geral).
 - Acórdão 4397/2020-Segunda Câmara | Relator: RAIMUNDO CARREIRO
 - ÁREA: Pessoal | TEMA: Ato sujeito a registro | SUBTEMA: Ato complexo
 - Outros indexadores: Decadência, Recurso extraordinário, STF, Termo inicial



Índice de autores

B

BARBOSA, Euro Gama 29

Gestão dos Recursos de Tempo de Auditoria: Modelo Matemático de Estimação e de Controle

BALANIUK, Remis 39

Rastro-DM: Mineração de Dados com Rastro

C

CASTRO, Marcus Vinícius Borela de 39

Rastro-DM: Mineração de Dados com Rastro

D

DANTAS, José Alves 29

Gestão dos Recursos de Tempo de Auditoria: Modelo Matemático de Estimação e de Controle

DUTRA, Tiago Alves de Gouveia Lins 11

A auditoria financeira de agências federais nos EUA tem relação custo-efetividade positiva e as lições aprendidas são aplicáveis ao Brasil

S

SANTOS, Daniel Novais 29

Gestão dos Recursos de Tempo de Auditoria: Modelo Matemático de Estimação e de Controle



SUBMISSÃO DE ARTIGOS

IDIOMA

A Revista do TCU é um periódico bilíngue (português/inglês). Os artigos selecionados são traduzidos gratuitamente. São aceitos textos escritos em português e inglês. O título, o resumo e as palavras-chave devem ser informados em português e em inglês.

AVALIAÇÃO CEGA

Os artigos a serem publicados são selecionados pelos membros do Conselho Editorial, que recebem os trabalhos sem qualquer forma de identificação do autor. Para assegurar imparcialidade ao processo, é necessário que a identificação de autoria seja realizada em documento separado. As informações de autoria devem também ser removidas da opção “Propriedades”.

Devem ser enviados, portanto, dois arquivos: 1. o texto do artigo de acordo com o template e 2. os nomes e currículos dos respectivos autores.

AUSÊNCIA DE REMUNERAÇÃO

A aprovação e publicação de trabalhos na Revista do TCU não dá aos autores o direito de percepção de qualquer retribuição pecuniária, devido à gratuidade na distribuição do periódico. A publicação resguarda os direitos autorais, na forma da Lei. Cada autor receberá cinco exemplares do número da revista no qual seu trabalho tenha sido publicado.

CRITÉRIOS DE SUBMISSÃO

Os artigos devem ser encaminhados para publicação através do sistema OJS disponível em <<https://revista.tcu.gov.br/ojs/index.php/RTCU/about/submissions>>. A seleção de artigos observa os seguintes critérios:

IDENTIFICAÇÃO DO AUTOR

Incluir na primeira folha do documento, separado do texto do artigo, currículo resumido com máximo de 250 caracteres, indicando nome(s) do(s) autor(es), instituição, cargo/função e formação acadêmica. Informar endereço para envio de exemplares da edição da Revista em que for publicado o artigo.

CONTEÚDO

Compatibilidade com a temática da Revista (relativa a Tribunais de Contas, Controle Externo, Administração Pública, Direito Público, Contabilidade, Finanças e Auditoria no âmbito do setor estatal). Contribuição original e inédita, que não esteja em processo de avaliação por outra publicação (PORTARIA TCU Nº 292/1995). Qualidade, objetividade e impessoalidade do texto. Linguagem adequada à norma culta da língua portuguesa.

FORMATAÇÃO

- Template de artigos da Revista do TCU
- Arquivo para revisão cega
- Formato Word (extensão.doc/docx), de até 2MB.
- Máximo de 10 mil palavras (aproximadamente 25 páginas, consideradas as referências).
- Estrutura de acordo com o template disponibilizado em <<https://revista.tcu.gov.br/ojs/index.php/RTCU/about/submissions>>.
- Elementos pré-textuais: a) título, e subtítulo (se houver); b) nome(s) do(s) autor(es); d) resumo na língua do texto; a) palavras-chave na língua do texto.
- Elementos textuais: a) introdução; b) desenvolvimento; c) conclusão.
- Elementos pós-textuais: a) título, e subtítulo (se houver) em língua estrangeira; b) resumo em língua estrangeira; c) palavras-chave em língua estrangeira; d) nota(s) explicativa(s); e) referências.
- Resumo de acordo com o template disponibilizado em <<https://revista.tcu.gov.br/ojs/index.php/RTCU/about/submissions>>, com até 250 palavras.
- Fonte Helvetica tamanho 10.
- Títulos em Helvetica Bold 24pt
- Subtítulos em Helvetica Bold 11pt
- Alinhamento à esquerda
- Espaçamento simples entre as linhas.
- Evitar linhas em branco entre os parágrafos.
- Formato da página: A4 (21x29,7 cm)
- Todas as margens com 2 cm.
- Destaques em negrito.
- Termos em língua estrangeira em itálico.
- Tabelas e ilustrações (mapas, diagramas, orga- nogramas, quadros, fotografias, gráficos, fluxogramas, entre outros) preferencialmente com 300 dpi, apresentados no corpo do documento, e envio dos originais separadamente. Normas de apresentação tabular (1993) do IBGE e NBR 14724 (2011) da ABNT para ilustrações.
- Formatos e formatação de tabelas conforme previstos pelo template.

CITAÇÕES E REFERÊNCIAS BIBLIOGRÁFICAS

Citações de acordo com a versão mais atual da NBR 10520 da ABNT, conforme os exemplos a seguir:

Citação direta

Segundo Barbosa (2007, p. 127), “entende- -se que ...”,

ou

“Entende-se que ...” (BARBOSA, 2007, p. 127).

Citação indireta

A teoria da ... (Cf. BARBOSA, 2007, p. 127),

ou,

A teoria da ... (BARBOSA, 2007, p. 127).



As citações de diversos documentos de um mesmo autor, publicados num mesmo ano, são distinguidas pelo acréscimo de letras minúsculas, após a data e sem espaçamento, conforme a ordem alfabética da lista de referências.

Segundo Barbosa (2007a, p. 127), ...

(BARBOSA, 2007b, p. 94).

Notas de rodapé devem ser evitadas, exceto para informações adicionais sobre trabalhos em andamento ou não publicados, ou comunicação pessoal.

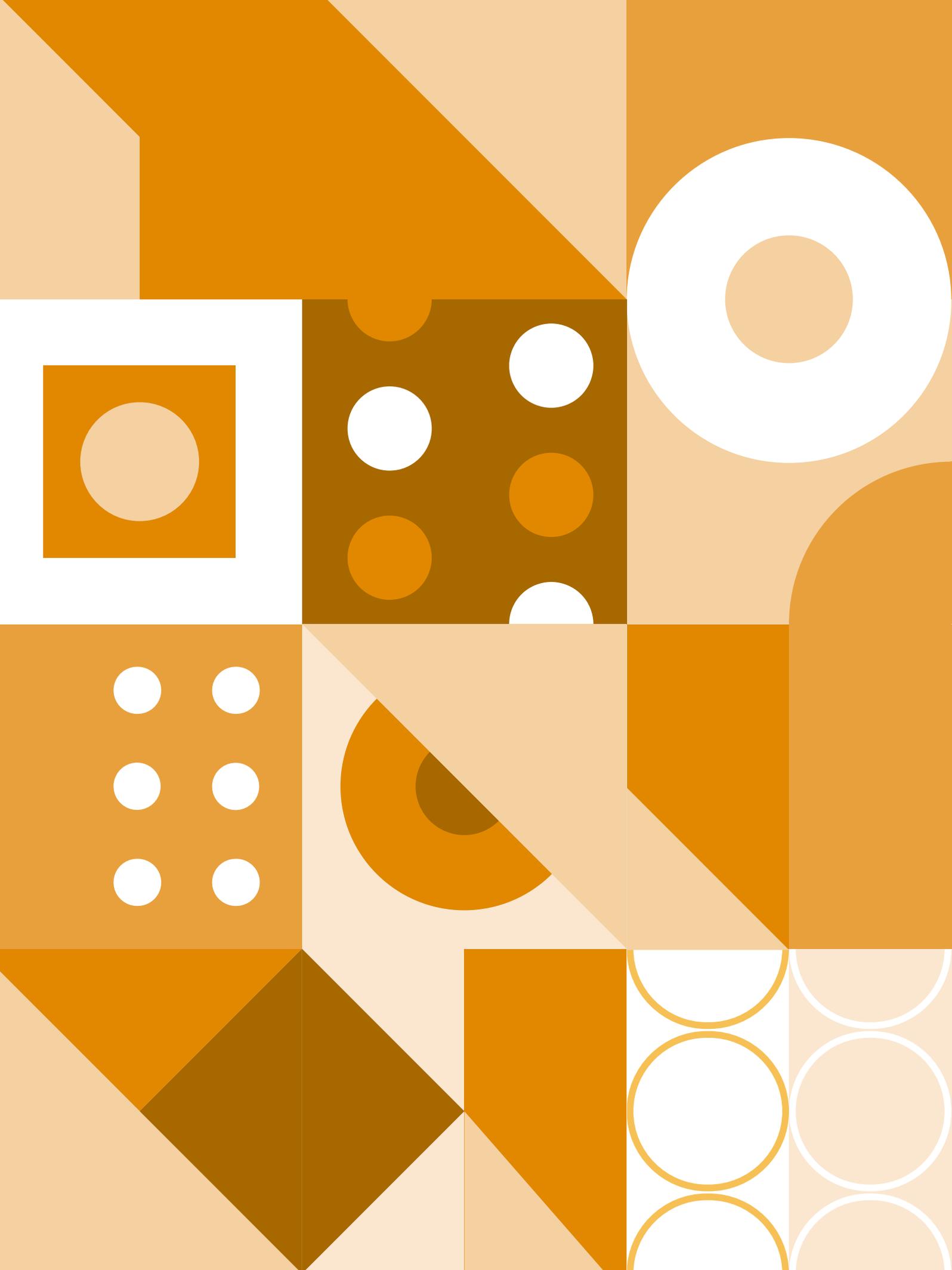
Lista de referências completas utilizadas no artigo, de acordo com a versão mais atual da NBR 6023 da ABNT. a) somente deverão ser citados na lista de referências trabalhos editados ou disponíveis para acesso público; b) as referências deverão ser apresentadas em ordem alfabética e alinhadas à esquerda; c) os artigos submetidos poderão ser referenciados em parte ou no todo, em formato impresso ou digital.

DECLARAÇÃO DE DIREITO AUTORAL

Os textos aprovados são publicados sem alteração de conteúdo. Os conceitos e opiniões emitidas em trabalhos doutrinários assinados são de inteira responsabilidade de seus autores. Os artigos publicados poderão ser divulgados em outros canais, desde que citada a Revista do TCU, ano, número e data de publicação como primeiro veiculador do trabalho.

POLÍTICA DE PRIVACIDADE

Os nomes e endereços informados nesta revista serão usados exclusivamente para os serviços prestados por esta publicação, não sendo disponibilizados para outras finalidades ou a terceiros.



MISSÃO

Aprimorar a Administração Pública em benefício da sociedade por meio do controle externo.

VISÃO

Ser referência na promoção de uma Administração Pública efetiva, ética, ágil e responsável.

