



Avaliação semiautomática de conteúdos de Planos Diretores de TIC com tecnologias de Processamento da Linguagem Natural

Ethel Airtton Capuano

Doutor em Ciência da Informação pela Universidade de Brasília (UnB).

Universidade de Brasília (UnB), Brasília, Brasil.

Servidor, Ministério da Integração e do Desenvolvimento Regional, Brasília, Brasil.

<http://lattes.cnpq.br/1075438704907683>



Submetido em: 03/04/2022. Aprovado em: 25/01/2024. Publicado em: 21/06/2024.

RESUMO

Este artigo apresenta os resultados de uma pesquisa sobre busca textual e avaliação semiautomática de conteúdos interessantes sobre os Planos Diretores de Tecnologia da Informação e Comunicação publicados por uma expressiva amostra de órgãos públicos da União. Com uso de tecnologias de processamento da linguagem natural, extraiu-se dos documentos dos planos pesquisados informações úteis para avaliações sobre o estado-da-arte da TIC nos órgãos e interessantes descobertas, algumas delas causando surpresa (por serendipidade), como o modelo de gestão de demandas em algumas unidades de TIC, baseado em estimativas da capacidade de entregas das equipes. Conclui-se que os planos estão em conformidade com as normas oficiais, inclusive observando-se um interesse da maioria dos órgãos em questões de gestão atuais, como governança e gestão de riscos, e maturidade na avaliação de modismos de mercado, mas também que seus conteúdos ainda carecem de abordagens sobre algumas práticas importantes para orientar sua evolução técnica.

Palavras-chaves: processamento da linguagem natural; mineração de textos; avaliação semântica de planos diretores de TIC.

INTRODUÇÃO

Este artigo apresenta os resultados de uma pesquisa sobre busca textual e avaliação semiautomática de conteúdos tecnológicos dos Planos Diretores de Tecnologia da Informação e Comunicação (PDTIC) de órgãos públicos dos Poderes Executivo, Legislativo e Judiciário da União, mais o Ministério Público Federal, vigentes entre 2013 e 2019, com uso de tecnologias de Processamento da Linguagem Natural (PLN).

A motivação temática da pesquisa se deve tanto à ideia de testar a possibilidade de automação de processos de avaliação de conteúdos textuais de documentos do setor público com tecnologias de PLN quanto a uma curiosidade científica em relação aos próprios conteúdos tecnológicos tratados pelos órgãos em seus PDTIC, num contexto de avaliação multidimensional e multidisciplinar. Considerando, como premissa de pesquisa, que os PDTIC refletem, ainda que com algumas limitações, o estado-da-arte da Tecnologia da Informação e Comunicação (TIC) dos órgãos públicos federais, questões relevantes que poderão ser melhor compreendidas com os resultados da pesquisa são: “Como anda o ‘estado-da-arte’ da TIC nos órgãos públicos da União, em termos de organização interna (arquitetura)?”; “Como estão sendo tratados os temas atuais de TIC?”; “Como estão sendo tratados os temas mais técnicos e estruturantes de gestão de TIC?”; “Quais são os órgãos com maiores orçamentos de TIC, que, portanto, requerem maiores atenções em termos de governança, gestão de riscos e controles?”; “Como estão as equipes das unidades de TIC dos órgãos e das entidades, como forças de trabalho?”.

Embora o estado-da-arte da TIC dos órgãos públicos federais tenha evoluído desde o advento (no Brasil) da *World Wide Web (Web)* e da *Internet* na segunda metade dos anos 1990, existe certo consenso de que essa evolução não tem ocorrido no passo necessário para o avanço desejado dos projetos de digitalização dos serviços públicos, como se pode notar nos revelados gargalos de atendimento de alguns serviços importantes na atualidade (Batista; Azevedo; Hessel, 2020). Conquanto se reconheça, em alguns setores, que os contínuos investimentos em serviços digitais providos por robustas arquiteturas de TIC apresentaram êxito inquestionável, como no caso dos sistemas da declaração do imposto de renda (que atendem dezenas de milhões de contribuintes acessando, simultaneamente, os canais de serviços automáticos) e no caso dos tribunais de justiça, que digitalizaram quase todos os seus serviços processuais, em outros serviços públicos os investimentos das últimas décadas não parecem apresentar resultados satisfatórios.

Como exemplos do pioneirismo dos tribunais de justiça em inovações com uso intensivo de TIC, o Supremo Tribunal Federal (STF), na segunda metade da década de 1990, inovou no setor público brasileiro ao implementar um sistema *push* (no sentido de “empurrar”, automaticamente, a informação para os usuários na rede mundial de computadores), para que os profissionais operadores do direito pudessem acompanhar o andamento dos processos de seus clientes remotamente, pela Internet. Os órgãos do Poder Judiciário também inovaram, ainda na primeira década deste milênio, ao estabelecerem

canais inteiramente digitais para troca de documentos processuais com a unidade jurídica do Instituto Nacional do Seguro Social (INSS), ante a materialidade (em quantidade e valores) das ações contra aquela autarquia-ré. E, recentemente, o Tribunal Regional Federal da 4ª Região (TRF-4) inovou ao implementar, com êxito notório, o primeiro sistema de Gestão Eletrônica de Documentos (GED) modelado, na origem, para o setor público brasileiro – o Sistema Eletrônico de Informações (SEI), que está sendo progressivamente adotado também pelos órgãos do Poder Executivo Federal.

O objetivo geral da pesquisa, portanto, concentrou-se em testes práticos de tecnologias de PLN para uma avaliação mais substantiva e menos formal do planejamento de TIC nos órgãos e nas entidades da União. Em síntese, esse objetivo geral consistiu em verificar, com instrumentos de automação de processos de buscas em textos, se esse modelo de planejamento governamental poderá contribuir, efetivamente, para o desenvolvimento de soluções de TIC adequadas segundo os padrões e as melhores práticas de mercado e a jurisprudência dos órgãos de controle. Como objetivos específicos, buscou-se nos textos dos planos coletados evidências de temas técnicos relevantes e atuais abordados pelos órgãos e pelas entidades.

METODOLOGIAS, PROCESSOS E TECNOLOGIAS DE PESQUISA

O método de pesquisa utilizado é multidisciplinar, com abordagens qualitativas (semânticas) relativas à busca, à recuperação e ao processamento de conteúdos textuais (Coneglian; Segundo, 2023; Kao; Poteet, 2007; Meadow *et al.*, 2007; Konchady, 2006), mas também com abordagens quantitativas, com emprego de estatística textual (Scott, 2019), com objetivo de se avaliar os níveis de interesse dos órgãos em cada tema de pesquisa revelados nos PDTIC.

Como áreas de conhecimento, a metodologia de pesquisa pode ser enquadrada, em termos teóricos (ainda que parcialmente), na Linguística (Scott, 2019) e na Ciência da Informação (Falcão; Lopes; Souza, 2022; Meadow *et al.*, 2007), esta última na disciplina que se denomina “Recuperação da Informação” (*Information Retrieval*). Essas duas áreas de conhecimento sustentam, metodologicamente, as duas principais disciplinas instrumentais utilizadas na pesquisa: Processamento da Linguagem Natural (*Natural Language Processing – NLP*) e Análise Textual (*Analytics*). Esse conjunto de tecnologias também contribui para a composição de soluções nos conceitos de mercado conhecidos como Mineração de Textos (*Text Mining*) e Ciência de Dados.

As informações textuais recuperadas dos textos dos PDTIC se referem tanto a atributos selecionados *a priori*, na etapa de planejamento do projeto de pesquisa, como a *clusters* (composições) de palavras substantivas e interessantes baseadas em palavras-chave mais frequentes. Empregou-se nos processos de buscas tanto tecnologias mais simples, como *softwares* editores de textos eletrônicos, como um aplicativo de *software* de mercado mais sofisticado, para PLN, que implementa recursos mais avançados de recuperação

da informação, tais como *clusters*, e estatísticas sobre textos no idioma português (Scott, 2019). Empregou-se, ainda, a análise de discurso como método de seleção de informações relevantes sobre alguns atributos de interesse nos PDTIC coletados.

Os documentos foram recuperados dos portais digitais dos órgãos mediante busca na *Web*, pela *Internet*. Excluindo-se, por decisão de projeto, as Instituições Federais de Ensino (IFE) e as empresas estatais, no Poder Executivo Federal, a amostra de PDTIC pesquisados, com 72 órgãos no total, abrange 17 dos 23 ministérios vigentes a partir da Lei nº 13.502, de 1º de novembro de 2017 (utilizando-se, no entanto, o PDTIC 2017-2019 do novo Ministério da Justiça e Segurança Pública e não os PDTIC do Ministério da Justiça e do Ministério da Segurança Pública previstos na última alteração dessa Lei), 42 das 80 autarquias e fundações remanescentes após a exclusão das IFE, 2 órgãos do Poder Legislativo Federal, 10 órgãos do Poder Judiciário Federal e o Ministério Público Federal.

O conjunto de informações textuais geradas a partir da coleta de documentos e da análise textual dos PDTIC foi tabulado numa planilha eletrônica para um tratamento estatístico adequado, postando-se em colunas os conteúdos segundo atributos a serem observados nos documentos e em linhas os respectivos órgãos e entidades separados segundo uma categorização político-administrativa dos órgãos da União: 3 Poderes (amostra com 72 órgãos e entidades), Poder Executivo (59 órgãos), Poder Executivo – Administração Direta (17 órgãos), Poder Executivo – Administração Indireta (42 autarquias e fundações), Poder Legislativo (2 órgãos), Poder Judiciário (10 órgãos), Ministério Público Federal (um órgão).

Em suma, recuperou-se de cada PDTIC conteúdos textuais sobre 37 atributos de pesquisa, dispostos na planilha conforme os seguintes blocos de análise e formatos de dados:

- I. Análise Documental Básica (com 7 atributos): volume do arquivo (em MB); fonte e tamanho das letras dos PDTIC (texto); espaçamento entre as linhas (número real); quantidade de páginas (número inteiro); quantidade de anexos (número inteiro); se a versão do plano era nova ou prorrogação de plano anterior (dado binário: “sim” ou “não”); e horizonte de planejamento em anos (número inteiro);
- II. Existência de Capítulos Temáticos (com 7 atributos) sobre: estratégias de negócios e de TIC articuladas no PDTIC (“sim” ou “não” para cada estratégia); arquitetura empresarial, de informação e técnica (ou de sistemas) (“sim” ou “não” para cada arquitetura); metodologia de gestão de projetos de TIC (“sim” ou “não”); gestão de riscos (“sim” ou “não”); governança (“sim” ou “não”);
- III. Necessidades de TIC (com 9 atributos): quantidade de páginas expressando as necessidades (ou demandas) de TIC (número inteiro); itens de necessidades de TIC (número inteiro); itens eventuais de projetos de TIC (número inteiro); clareza na apresentação das demandas (“sim” ou “não”); análise técnica das demandas de TIC (antes da inserção no PDTIC) (“sim” ou “não”); registros de demandas redundantes (“sim” ou “não”); menção à Metodologia de Desenvolvimento de Sistemas (MDS) (“sim” ou “não”); menção a solução de Business Intelligence (BI) (“sim” ou “não”); uso de metodologia para priorização no atendimento das necessidades (“sim” ou “não”);

- IV. Desempenho do PDTIC Anterior (um atributo): percentual, revelando a parcela executada das necessidades e/ou projetos previstos;
- V. Orçamento Estimado do PDTIC (um atributo): dados monetários relativos à previsão de despesas para execução do plano no horizonte previsto;
- VI. Declaração de Lacunas de Recursos Humanos na Equipe de TIC (um atributo): dado binário (“sim” ou “não”), com base em declaração da unidade de gestão; e
- VII. Frequência de Palavras e Expressões Temáticas (11 atributos): estratégia; georreferenciamento; Gestão da Informação; Gestão do Conhecimento; Segurança da Informação; GIS/SIG; Imagem de Satélite; *Drone/VANT*; Computação em Nuvem (*Cloud Computing*); *Big Data*; Aplicativo Móvel (*App*).

Os parâmetros estatísticos calculados na planilha, para cada categoria político-administrativa de órgãos e entidades da União, foram:

- a) média: para todos os atributos, mas utilizando um dado numérico somando a presença de “sim” em cada PDTIC analisado no caso dos dados binários;
- b) desvio padrão: idem; e
- c) coeficiente de variação (razão entre o desvio-padrão e a média): idem.

Quanto às suas estruturas documentais, os planos pesquisados dos órgãos do Poder Executivo seguem, na maior parte de seus conteúdos, o “Guia de Elaboração de PDTI do SISP” (Brasil, 2012), num raciocínio lógico a partir de princípios e diretrizes, estratégias de TIC, plano de trabalho para a elaboração do PDTIC, Matriz *SWOT*, avaliação de recursos humanos da equipe de TIC, método de priorização adotado, inventário de necessidades e portfólio de projetos e ações, avaliação de riscos, plano de metas e ações e orçamento estimado.

REVISÃO DE LITERATURA

Embora, para o público familiarizado com a gestão de TIC nos órgãos públicos, a maioria dos conceitos abordados neste artigo não represente novidade, alguns deles necessitam de definição, a fim de se evitar interpretações equivocadas em relação ao contexto da pesquisa, e considerando o contexto institucional dessas organizações, as normas de conformidade também merecem destaques.

Conceitos técnicos

O primeiro conceito importante é de “Plano Diretor de Tecnologia da Informação e Comunicação”, que no Decreto nº 8.638, de 15 de janeiro de 2016, revogado pelo Decreto nº 10,332, de 28 de abril de 2020, é expressamente definido como um “instrumento de diagnóstico, planejamento e gestão dos recursos e processos de tecnologia da informação e comunicação, com o objetivo de atender às necessidades finalísticas e de informação de órgão ou entidade para determinado período.”

O conceito de “diagnóstico” adotado no guia de referência (Brasil, 2012), no entanto, merece atenção especial, pois pode não coincidir exatamente com o que se entende por diagnóstico em outros contextos, conforme o texto a seguir:

Após a fase de preparação, na qual foram organizadas as principais atividades do projeto de elaboração do PDTI, inicia-se a 2ª fase do processo: a fase de diagnóstico. Essa fase caracteriza-se por buscar compreender a situação atual da TI na organização para, em consonância com esse quadro, identificar as necessidades (problemas ou oportunidades) que se espera resolver. Para isso, contempla processos relacionados a análise do planejamento anterior, análise estratégica e levantamento de necessidades. (...) O levantamento de necessidades parte daquelas relacionadas à informação e se desdobra em todas as outras associadas à TI: serviços, infraestrutura, contratações e pessoal de TI. (...) Um dos principais artefatos produzidos nessa fase, e muito importante para todo o processo de elaboração do PDTI, é o Inventário de Necessidades. (Brasil, 2012, p. 33)

Outro conceito primordial, adotado como motivação da pesquisa, é o de “adequação técnica” dos PDTIC, que não deve se confundir, necessariamente, com a adequação formalmente exigida nas normas de regência, pois estas não abordam em profundidade todos os temas relevantes para a gestão de TIC nos órgãos, tais como “arquitetura empresarial” (também denominada “arquitetura corporativa”), “arquitetura de informação” e “arquitetura técnica” (ou de sistemas).

Os conceitos correlatos à “arquitetura”, por outro lado, quando aplicados ao contexto da gestão de TIC nas organizações, são endereçados em vários relatórios e artigos técnicos produzidos, por exemplo, pela empresa especializada em pesquisas de mercado *Gartner Group*, bastante conhecida pelos gestores de TIC, e outros autores (Gartner Group, 2020a; Costelo, 2019; Hermans, 2015; Ross; Weill; Robertson, 2008).

Entretanto, originalmente o conceito de “arquitetura”, num contexto mais amplo de gestão empresarial com uso intensivo de sistemas de informações computacionais, deve-se à iniciativa de pesquisa pioneira conhecida como *Business Systems Planning (BSP)*, implementada nos anos 1960 por vários pesquisadores a serviço da empresa IBM (Kotusev, 2016, 2017), destacando-se, posteriormente, o construto descritivo empresarial elaborado por Zachman (1987), mais conhecido por “Construto de Zachman” (*Zachman Framework*), que se destina a descrever e integrar, com efetividade e eficiência, os elementos ontológicos mais relevantes que constituem uma empresa (Nicolaidou *et al.*, 2009; Zachman, 2008).

O conceito de “Arquitetura Empresarial” (*Enterprise Architecture – EA*), um dos itens de gestão pesquisados nos PDTIC, que, no presente contexto, pode ser entendido como “arquitetura corporativa”, é assim definido pela empresa de pesquisas de mercado de TIC *Gartner Group*:

[...] uma disciplina para as empresas proativamente e holisticamente líderes responderem a forças disruptivas mediante a identificação e análise da execução de mudanças conforme a visão de negócios e os resultados desejados. A arquitetura empresarial entrega valor apresentando aos líderes de negócios e de

TIC recomendações para o ajuste de políticas e projetos de modo a alcançar os resultados de negócios visados, que consolidam rupturas de negócios relevantes (Gartner Group, 2020a, *online*, tradução nossa)¹.

Em síntese, no contexto da pesquisa, uma “arquitetura corporativa” cuidaria da identificação, descrição, avaliação e integração funcional dos principais recursos de um órgão público, de modo sistêmico e sinérgico, para o alcance de seus objetivos institucionais. Esses recursos podem ser classificados em: humanos, materiais, patrimoniais, tecnológicos, de informação e de conhecimento.

Os conceitos de “arquitetura da informação” e “arquitetura de sistemas”, mais recentes, que também se referem a dois objetos temáticos pesquisados nos PDTIC, são definidos por um conjunto maior de outros autores, com significados às vezes diversos entre os autores num mesmo tema. O autor pioneiro da ideia de “arquitetura da informação” é Wurman (1997), que a definia, em linhas gerais, como uma expansão da ocupação do arquiteto tradicional aplicada à organização da informação. Lima-Marques e Macedo (2006), nessa mesma linha, mas com abordagem filosófica *sui generis*, definem essa disciplina como uma metodologia para estruturação de sistemas de informação aplicada a qualquer ambiente informacional, sendo este compreendido como o espaço que integra contexto, conteúdos e usuários.

Entretanto, o conceito mais adequado de “arquitetura da informação”, no atual contexto de pesquisa, parece ser o de Hagedorn, conforme Lima-Marques e Macedo (2006, p. 245): *“arte e ciência da organização da informação para ajudar efetivamente pessoas a satisfazerem suas necessidades de informação. Envolve a investigação, análise, desenho e implementação”*.

Quanto ao conceito de “arquitetura técnica”, também denominada “arquitetura de sistemas”, refere-se à organização inteligente dos sistemas de informações e das tecnologias que os suportam nas organizações. Esse tipo de arquitetura seria a base cognitiva de uso dos componentes de sustentação tecnológica das organizações, integrando os componentes de *hardware* (estações de trabalho, servidores, unidades de armazenamento massivo, dispositivos de controles de segmentação em redes de telecomunicação etc.) e de *softwares* básicos (como sistemas operacionais e servidores de aplicações), além de outros fundamentais para operações computacionais, como os aplicativos de segurança, *backups* e balanceamento de cargas.

Esse conceito, didaticamente apresentado por Crawley (2007), é importante para orientar inclusive os processos de aquisição e/ou desenvolvimento de novos sistemas, integração entre sistemas legados, interoperabilidade entre componentes e sistemas, portabilidade de sistemas legados de um ambiente tecnológico obsoleto para outro mais atual, combinação entre tecnologias legadas e novas, etc. O mais promissor modelo de “arquitetura técnica”,

¹ Original: “[...] is a discipline for proactively and holistically leading enterprise responses to disruptive forces by identifying and analyzing the execution of change toward desired business vision and outcomes. EA delivers value by presenting business and IT leaders with signature-ready recommendations for adjusting policies and projects to achieve targeted business outcomes that capitalize on relevant business disruptions” (Gartner Group, 2020a, *online*).

inclusive para o atual momento dos órgãos públicos federais, é o conhecido como “Arquitetura Orientada a Serviço” (*Service-Oriented Architecture – SOA*), que incorpora aspectos lógicos e físicos, assim definido por Bell (2008, p. 309, tradução nossa):

A disciplina de arquitetura lógica orientada a serviço é principalmente voltada para o reúso, a utilização e o consumo. (...) essa disciplina tem como foco os três maiores aspectos funcionais de uma arquitetura lógica: (1) o relacionamento entre pacotes de *softwares* implementados, (2) as preocupações estruturais das entidades dos pacotes orientados a serviços, e (3) o comportamento e a colaboração das soluções empacotadas. A arquitetura física orientada a serviço se refere a tudo que é fisicamente endereçável aos componentes orientados a serviço. Esse endereçamento consiste no mapeamento das entidades de *software* em componentes de *hardware* e nos seus locais endereçáveis numa rede de comunicação de dados.

Esse construto de organização de TIC é comentado também por Abrams e Schulte (2008), da empresa *Gartner Group*, cujo glossário técnico assim o define (tradução nossa):

(...) é um paradigma e uma disciplina de desenho que auxilia a TI no atendimento das demandas do negócio. (...) SOA reduz a redundância e incrementa a usabilidade, a manutenibilidade e o valor. Isso resulta em sistemas interoperáveis e modulares, que são mais fáceis de usar e de manter. SOA cria sistemas mais simples e rápidos, que incrementam a agilidade e reduzem o Custo Total de Propriedade (TCO). (*Gartner Group*, 2020b)

Observe-se, a propósito, que essa estratégia baseada em arquiteturas de informação e de TICs parece coincidir com a atual proposta de mudança estrutural do modelo de prestação de serviços dos órgãos do Poder Executivo Federal, com maior interação digital entre os órgãos públicos e os cidadãos. (Batista; Azevedo; Hessel, 2020)

Contudo, questiona-se na prática: “Como uma arquitetura técnica deve ser concebida e implementada nas organizações?” Como resposta, argumenta-se que essa arquitetura pode ser elaborada a partir de um desenho primordial concebido segundo o melhor pensamento técnico em TIC, editando-se normas internas nas organizações para seu uso compulsório, que prevê certos padrões e melhores práticas, tais como padrões tecnologicamente definidos para sistemas operacionais e sistemas gerenciadores de bancos de dados relacionais (SGBDR).

Como exemplo de situação recorrente em que as vantagens de uma arquitetura técnica padronizada poderiam evitar despesas ao erário, imagine-se, por hipótese, um novo aplicativo sendo adquirido ou desenvolvido extramuros (por órgãos parceiros, muitas vezes), mas exigindo, para sua instalação no ambiente tecnológico do órgão demandante, um produto de banco de dados ainda não utilizado e não dominado pela equipe de sustentação de TIC nesse órgão. Em consequência, surgirá a necessidade de aquisição e domínio técnico (da equipe de suporte a bancos de dados), em curto prazo, de um novo *software* bastante complexo, contribuindo assim para tornar os ambientes de TIC ainda mais custosos, resultando num asoberbamento das unidades de gestão de TIC com mais tarefas que poderiam ser evitadas.

Normas de conformidade

Os órgãos do Poder Executivo Federal devem executar suas atividades de gestão de TIC segundo padrões técnicos e melhores práticas de mercado, mas também em conformidade com as normas aprovadas pelo órgão central do Sistema de Administração dos Recursos de Tecnologia da Informação (SISP) e pela extensa jurisprudência firmada ao longo das últimas duas décadas pelos órgãos de controle interno e externo. O modelo de gestão governamental por sistemas estruturantes, no Poder Executivo Federal, remonta ao Decreto-Lei nº 200, de 25 de fevereiro de 1967, que assim estabeleceu:

Art. 30. Serão organizadas sob a forma de sistema as atividades de pessoal, orçamento, estatística, administração financeira, contabilidade e auditoria, e serviços gerais, além de outras atividades auxiliares comuns a todos os órgãos da Administração que, a critério do Poder Executivo, necessitem de coordenação central. § 1º Os serviços incumbidos do exercício das atividades de que trata este artigo consideram-se integrados no sistema respectivo e ficam, conseqüentemente, sujeitos à orientação normativa, à supervisão técnica e à fiscalização específica do órgão central do sistema, sem prejuízo da subordinação ao órgão em cuja estrutura administrativa estiverem integrados. (...)

Art. 31. A estruturação dos sistemas de que trata o artigo 30 e a subordinação dos respectivos Órgãos Centrais serão estabelecidas em decreto. (Brasil, 1967, *online*).

Originalmente, o SISP foi instituído pelo Decreto nº 1.048, de 21 de janeiro de 1994, que então se denominava Sistema de Administração dos Recursos de Informação e Informática. Em sua versão atual, o Decreto nº 7.579, de 11 de outubro de 2011 (que também se reporta ao Decreto-Lei nº 200, de 1967), definiu a forma como o SISP é organizado e, em seu art. 2º, estabeleceu a finalidade da gestão sistêmica, com as funções substantivas basilares que mais interessam à avaliação dos PDTIC executada na pesquisa em pauta (*in litteris*):

IV – estimular o uso racional dos recursos de tecnologia da informação, no âmbito do Poder Executivo federal, visando à melhoria da qualidade e da produtividade do ciclo da informação; V – estimular o desenvolvimento, a padronização, a integração, a interoperabilidade, a normalização dos serviços de produção e disseminação de informações, de forma desconcentrada e descentralizada (Brasil, 2011, *online*).

ANÁLISE E DISCUSSÃO DE CONTEÚDOS DOS PDTIC

Com base na planilha eletrônica preenchida com as informações coletadas dos documentos e os parâmetros estatísticos textuais calculados, a etapa de análise de conteúdos evidenciou os aspectos dos PDTIC apresentados a seguir.

Desempenho na execução

Os dados de desempenho na execução dos planos anteriores, um dos requisitos de informação para o diagnóstico previsto no guia do SISP, ainda constituem um desafio para a transparência na gestão de TIC dos órgãos, observando-se que somente 11 dos 17 órgãos (64,7%) da Administração Direta e 26 dos 42 órgãos (61,9%) da Administração Indireta do Poder Executivo Federal apresentaram essas avaliações em seus PDTIC. Embora vários órgãos mencionem a existência de documentos anexos com esse item de avaliação, eles muitas vezes não foram encontrados nos portais digitais dos órgãos.

Conforme a **TABELA 1**, a média de desempenho dos órgãos do Poder Executivo Federal na execução dos PDTIC é de apenas 44,3%, com coeficiente de variação 0,4, evidenciando-se que mais da metade dos conteúdos dos planos não saem do papel ou são executados apenas parcialmente (a média do Poder Judiciário não deve ser considerada porque representa os dados de apenas três órgãos não aleatórios).

Espera-se que pesquisas posteriores possam mostrar as causas desse nível de desempenho: se as expectativas de entregas da unidade de TIC dos órgãos são excessivas, em face de seus recursos (assim requerendo um maior senso de realidade no processo de priorização das necessidades); ou se a avaliação do tempo para implementação, pelo pouco detalhamento das necessidades de TIC na etapa de planejamento, tem sido muito otimista; ou um misto dessas duas causas.

TABELA 1 – Cálculo estimativo da capacidade de entrega (hipotético)

Recursos Humanos Disponíveis				Consumo de Tempo dos Perfis por Atividades (A _i) e Projetos (P _i) (horas-homem)					
Perfis	Tempo Laboral (horas-homem)			A ₁	A ₂	A ₃	A ₄	P ₁	P ₂
	Dia	Ano	PDTIC						
Perfil A	8	2.000	6.000	750	1.500	750	750	1.500	750
Perfil B	16	4.000	12.000	3.000	5.500	1.500	250	1.250	500
Perfil C	16	4.000	12.000	1.000	5.500	1.500	2.000	1.000	1.000
Perfil D	24	6.000	18.000	0	3.000	0	0	8.000	7.000
Perfil E	24	6.000	18.000	1.000	3.000	500	500	7.000	6.000
Perfil F	6	1.500	4.500	1.000	2.000	250	500	500	250
Perfil G	12	3.000	9.000	750	2.000	3.000	750	2.000	500
Tempo da Equipe (h-h)	106	26.500	79.500	7.500	22.500	7.500	4.750	21.250	16.000

Fonte: Elaborada pelo autor (2022).

Capacidade de entrega como critério de priorização

Com a análise do discurso expresso pelos autores dos textos, identificou-se, em 46 dos 72 (63,9%) PDTIC pesquisados, passagens em que lacunas de recursos humanos nas unidades de gestão de TIC são mencionadas, sendo 45 delas nos órgãos do Poder Executivo Federal (em 13 órgãos da Administração Direta e 32 da Administração Indireta). Ou seja, em mais de 76,0% dos órgãos do Poder Executivo Federal, tanto na Administração Direta como na Administração Indireta, são mencionadas lacunas de recursos humanos para gestão de TIC.

Em particular, destaque-se a iniciativa de algumas unidades de gestão de TIC em utilizar um método mais detalhado de avaliação de sua capacidade de entrega ao longo do horizonte temporal do plano, a partir de uma avaliação prévia de seus recursos humanos. O Guia do SISP, apesar de prever essa avaliação, não propõe nenhuma metodologia específica “de ponta a ponta” do processo, tendo como resultado que as avaliações encontradas nos PDTIC se concentram, em geral, apenas nos recursos humanos disponíveis, sem estimativas de recursos humanos necessários para implementação dos planos.

Com essa abordagem algo inovadora (e corajosa, deve-se ressaltar), um senso de realismo maior é adotado para priorização de atendimento às necessidades dos demandantes de soluções. O construto (estrutura de raciocínio) apresentado na **TABELA 1**, com dados hipotéticos de um órgão qualquer, embasa o algoritmo de cálculo da capacidade de entrega da unidade de TIC, com dados de sua força de trabalho (em horas-homem), e do consumo dessa força de trabalho para implementação do PDTIC no órgão. Esse cálculo deve contrapor uma soma das horas de trabalho (horas-homem) líquidas diárias e anuais disponíveis de cada membro da equipe (excluídas as horas comprometidas com outras atividades na unidade) a uma distribuição dessas horas nas diversas ações previstas para execução do PDTIC em seu horizonte temporal.

Com períodos de implementação (horizontes) de três anos, por exemplo, observe-se que, na quarta coluna da esquerda para a direita, a soma das horas-homem da coluna “PDTIC” é igual a três vezes a soma anual das horas líquidas disponíveis de cada perfil profissional envolvido na execução do PDTIC ($79.500 / 26.500 = 3$).

Com esse algoritmo de cálculo, é executado o planejamento detalhado do consumo de tempo de cada perfil profissional que compõe a equipe de gestão de TIC no órgão. Ele obriga a equipe de TIC a estimar, previamente, o consumo de tempo total para cada atividade ou projeto previsto no PDTIC, no horizonte de planejamento, e distribuir esse tempo com os respectivos tempos disponíveis dos profissionais que serão envolvidos na execução. Os PDTIC mais desenvolvidos nesse item, numa avaliação preliminar, são os dos órgãos: Ministério Público Federal (MPF), Conselho Nacional de Justiça (CNJ), Agência Espacial Brasileira (AEB) e Superintendência de Seguros Privados (SUSEP).

Essa inovação de gestão torna os PDTIC mais realistas, baseados em evidências, como recomendado no Acórdão nº 2608/2018-TCU-Plenário (Brasil, 2018), em relação

a políticas públicas, planos e programas. Embora esse modelo de avaliação de capacidade de entrega da equipe de TIC também apresente riscos de estimativas equivocadas, certamente estará mais perto da capacidade real de entrega da unidade, como um dos critérios de priorização aplicáveis.

Capítulos técnicos de interesse

O projeto de pesquisa também previu a verificação textual da eventual presença de capítulos técnicos específicos de interesse que evidenciassem o nível técnico dos PDTIC, previstos ou não no Guia do SISP. Os temas técnicos de pesquisa foram os apresentados na **TABELA 2**: “Estratégias”, “Arquiteturas”, “MGP” (Metodologia de Gestão de Projetos de TIC), “Gestão de Riscos” e “Governança”.

Com uso dos recursos de busca de palavras de um *software* processador de texto, identificou-se a presença ou não de capítulos com certos conteúdos de interesse para a pesquisa, com os títulos indicados na **TABELA 2**. Executou-se também, por precaução, a busca por esses capítulos nos corpos dos textos dos PDTIC, considerando-se que, às vezes, os sumários dos planos não são detalhados em relação aos conteúdos capitulares.

TABELA 2 – Capítulos técnicos específicos

Poderes e Entidades da União	Qtde. de PDTIC	Quantidade de PDTIC com Capítulos Técnicos				
		Estratégias	Arquiteturas	MGP (*)	Gestão de Riscos	Governança
Poder Executivo	59	4	15	2	51	29
Administração Direta	17	2	2	1	15	7
Administração Indireta	42	2	13	1	36	22
Poder Legislativo	2	1	1	0	0	1
Poder Judiciário	10	0	1	0	0	5
Ministério Público Federal	1	0	0	0	0	1
Total	72	5	17	2	51	36

Fonte: Elaborada pelo autor (2022).

(*)MGP: Metodologia de Gestão de Projetos de TIC

O tema “estratégias” se refere a demonstrações expressas e detalhadas de articulação entre as estratégias de negócios e de TIC no próprio PDTIC, de modo a evidenciar, tecnicamente, a integração e sinergia entre essas estratégias (não se trata, portanto, apenas de menção a essa articulação). Com os resultados apresentados na **TABELA 2**, conclui-se

que apenas 5 dos 72 órgãos (6,9%) evidenciam, minimamente, essa sinergia, ressaltando-se que vários órgãos remetem essa visão de articulação entre estratégias a outros documentos do órgão, que não compõem o PDTIC, tais como documentos de planejamento estratégico.

Os dados da coluna “arquiteturas”, na **TABELA 2**, consolidam as três arquiteturas objeto de pesquisa nos PDTIC: arquitetura corporativa ou empresarial, arquitetura da informação e arquitetura técnica ou de sistemas. Nota-se que apenas 17 dos 72 órgãos (23,6%) dispõem sobre esse tema em capítulos de seus PDTIC, com uma taxa melhor no conjunto dos órgãos da Administração Indireta do Poder Executivo: 13/42, ou 31,0%.

O resultado talvez mais frustrante ocorre em relação à presença de capítulos temáticos sobre metodologias de desenvolvimento de projetos de TIC (coluna “MGP”): apenas 2 dos 72 órgãos pesquisados elaboraram sobre o tema em capítulo específico no próprio PDTIC, ainda que vários mencionem a existência de uma metodologia ou a intenção de adotar uma metodologia. Com esse resultado, uma questão metodológica importante não se torna transparente para escrutínio técnico externo ao órgão, ou para troca de experiências e aprendizado entre as unidades de gestão de TIC e especialistas extramuros.

Considere-se, entretanto, como resultados positivos que a maior parte dos PDTIC apresenta capítulos temáticos sobre dois temas bastante atuais (nos termos do Decreto nº 9.203, de 22 de novembro de 2017): “gestão de riscos” e “governança”. Os órgãos do Poder Executivo Federal apresentam, em seus PDTIC, um capítulo sobre “gestão de riscos” em 51 dos casos (70,8%), com resultados melhores para os órgãos da Administração Direta, em 15 dos 17 órgãos (88,2%). Essas taxas decrescem para o tema “governança”, ainda que expressivas e bastante próximas em todas as categorias de órgãos; nos órgãos do Poder Executivo Federal, 29 em 51 deles (ou 56,9%) têm capítulos sobre “governança” em seus PDTIC.

Citações de temas atuais

Outro bloco de objetos semânticos de pesquisa se concentrou na busca por citações textuais, não necessariamente em capítulos específicos, sobre os seguintes temas: “estratégia”; “segurança da informação”; “GIS” (*Geographical Information System*); “SIG” (Sistema de Informações Geográficas); “georreferenciamento”; “imagem de satélite”; “Drone”; “VANT” (Veículo Aéreo Não Tripulado); “aplicativo para dispositivo móvel”, “*mobile*” (aplicativo para dispositivo móvel); “*app*” (de *software application*); “gestão da informação”; “gestão do conhecimento”; “computação em nuvem”; “*cloud computing*” e “*big data*”.

A **TABELA 3** apresenta os resultados somados dessa busca em cada PDTIC, ressaltando-se que os temas relativos ao georreferenciamento foram consolidados num único item para análise de dados (somando-se também as frequências das citações de “GIS”, “SIG”, “imagem de satélite”, “drone” e “VANT”), assim como os resultados de buscas de “aplicativo”, “*mobile*” e “*app*” (consolidados em “Aplicativo para Dispositivo Móvel”), somando-se as citações de “gestão da informação” com as de “gestão do conhecimento” pelas conexões naturais entre si.

TABELA 3 – Citações de temas atuais

Temas de Pesquisa		Poderes e Entidades (*) e Frequências Observadas de Citações Temáticas						
#	Expressão Temática	PE	PE-AD	PE-AI	PL	PJ	MPF	Subtotal
1	Estratégia	1.629	514	1.115	13	137	36	1.815
2	Segurança da Informação	1.240	371	869	5	77	15	1.337
3	GIS/SIG, Georreferenciamento, Imagem de Satélite, Drone/VANT (**)	628	98	530	3	23	0	654
4	Aplicativo para Dispositivo Móvel	234	162	72	6	18	3	261
5	Gestão da Informação e do Conhecimento	209	55	154	5	2	2	218
6	Computação em Nuvem (Cloud Computing)	180	49	131	9	27	11	227
7	Big Data	16	7	9	0	4	2	22
Quantidade de PDTIC (A)		59	17	42	2	10	1	72
Total de Citações Encontradas (B)		4.136	1.256	2.880	41	288	69	4.534
Total de Citações / PDTIC (C=B/A)		70,1	73,9	68,6	20,5	28,8	69,0	63,0

Fonte: Elaborada pelo autor (2022).

(*) PE: Poder Executivo; PE-AD: Administração Direta do Poder Executivo; PE-AI: Administração Indireta do Poder Executivo; PL: Poder Legislativo; PJ: Poder Judiciário; MPF: Ministério Público Federal.

(**) *Drone* (em Inglês) ou *VANT* (em Português): Veículo Aéreo Não Tripulado.

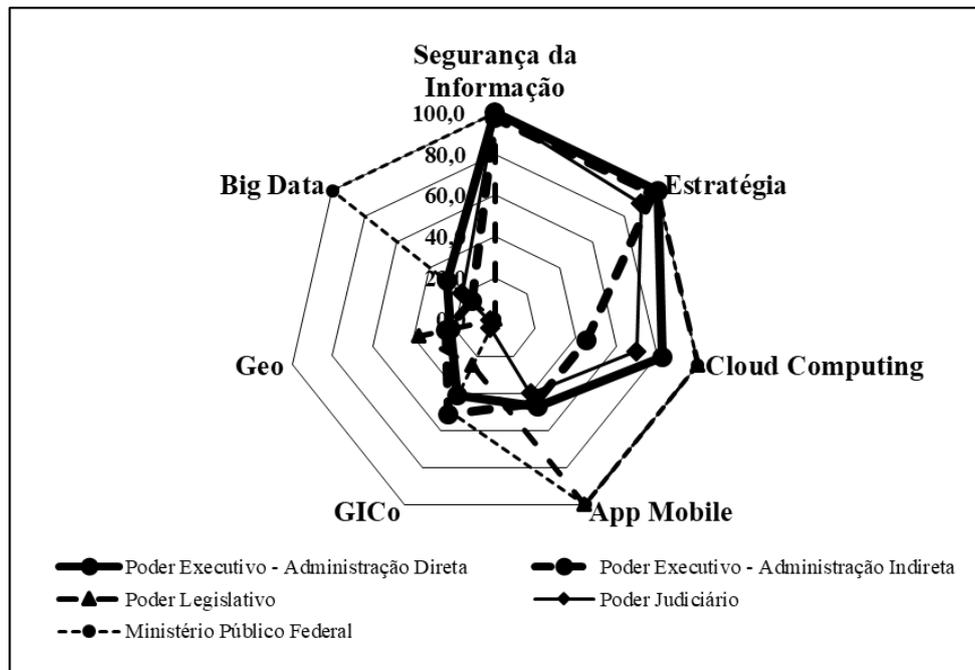
Os aspectos mais relevantes das buscas, observando-se os resultados na **TABELA 3**, são:

- os dois temas mais citados são, de longe, “estratégia” (com 1.815 citações encontradas) e “segurança da informação” (com 1.337 citações);
- quanto aos temas relativos a “georreferenciamento” e “aplicativo para dispositivo móvel”, o primeiro tema parece interessar mais aos órgãos da Administração Indireta do que aos órgãos da Administração Direta do Poder Executivo Federal, com razões de frequências da ordem de 12,6 citações por órgão (530/42) e 5,8 citações por órgão (98/17), respectivamente; enquanto o segundo tema parece interessar muito mais a órgãos da Administração Direta, na razão de 9,5 citações por órgão (162/17), contra 1,7 citação por órgão (72/42) na Administração Indireta;
- embora bastante comentado, nos últimos anos, o tema “computação em nuvem” (cloud computing) não aparece com muita frequência nos planos dos órgãos em geral, e no Poder Executivo Federal foram encontradas, em média, apenas 3,1 citações por PDTIC (180/59); menos ainda o tema “big data”, com apenas 16 citações ao todo (16/59, ou apenas uma citação por quase quatro PDTIC).

Considerando a importância do georreferenciamento digital em seu contexto, o extinto Departamento Nacional de Produção Mineral (DNPM), substituído pela Agência Nacional de Mineração, ANM, denomina seu plano como “Plano Diretor de Tecnologia da Informação, Comunicações e Geoprocessamento 2017-2019”.

Em termos de quantidades de órgãos em que aparecem as citações desses temas, em relação às quantidades de órgãos com PDTIC pesquisados (percentuais de órgãos com citações), o gráfico estilo “radar” da **FIGURA 1** evidencia as concentrações, com percentuais de 0 a 100, de PDTIC com as citações temáticas de pesquisa por Poder e Entidade da União. Evidencia-se uma concentração maior de citações aos temas “estratégia” e “segurança da informação” em órgãos da Administração Direta do Poder Executivo (linha cheia mais espessa) e menor nos órgãos da Administração Indireta (linha pontilhada mais espessa).

FIGURA 1 – Percentuais de PDTIC com citações de temas atuais



Fonte: Elaborada pelo autor (2022).

Análise das necessidades de TIC

Este bloco de dados de pesquisa se refere às demandas de soluções de TIC expressas nos PDTIC, que correspondem ao item de diagnóstico que, no Guia do SISP, é denominado “necessidades de TIC”. Em síntese, nessa lógica cognitiva, as necessidades representam o indicador mais importante segundo o qual a unidade de gestão de TIC deve direcionar seus recursos e, assim, contribuir para a evolução tecnológica do órgão.

A **TABELA 4** apresenta as quantidades de itens que aparecem como “necessidades de TIC” nos planos dos órgãos pesquisados, agrupados por categoria político-administrativa. Como previsto no Guia do SISP, alguns órgãos expressam os itens de aquisição ou desenvolvimento de bens e serviços como “necessidades de TIC”, sem um desdobramento mais detalhado, enquanto outros detalham suas necessidades em projetos e ações e, algumas vezes, em metas.

TABELA 4 – Necessidades de TIC

Poderes e Entidades da União	Qtde. de PDTI (A)	Quantidades de Itens como “Necessidades”	
		Subtotal (B)	Média (C=B/A)
Poder Executivo	59	3.982	67,5
Administração Direta	17	832	48,9
Administração Indireta	42	3.150	75,0
Poder Legislativo (*)	2	133	66,5
Poder Judiciário (*)	10	863	86,3
Ministério Público Federal (**)	1	-	-
Total	72	4.978	69,1

Fonte: Elaborada pelo autor (2022).

(*) Computados como necessidades ou projetos.

(**) As demandas não são consolidadas no mesmo formato dos demais órgãos.

Em geral, os órgãos do Poder Executivo Federal expressam suas necessidades de TIC em 24 páginas dos PDTIC, em média (com coeficiente de variação de 1,2), enquanto, nos órgãos do Poder Judiciário, essa média é um pouco menor, em torno de 19 páginas, com coeficiente de variação 0,9. O volume documental total registrando as demandas de TIC dos PDTIC pesquisados somou 1.693 páginas com 4.978 “necessidades” e 6.106 projetos ou ações com algum detalhamento.

Note-se, por relevante, que os órgãos dos Poderes Legislativo e Judiciário e o Ministério Público Federal (MPF) não se subordinam ao SISP, motivo pelo qual seus PDTIC não apresentam, geralmente, a mesma estrutura e os mesmos formatos de apresentação de dados dos órgãos do Poder Executivo Federal. Em relação às demandas de TIC observadas, vários órgãos dos Poderes Legislativo e Judiciário e o Ministério Público Federal apresentaram suas demandas diretamente como “projetos” e não como “necessidades”.

Com análise de discurso dos textos capitulares sobre as necessidades de TIC, observou-se que as descrições dos itens geralmente se encontram em níveis de abstração ainda muito elevados, com pouca clareza, de um ponto de vista técnico, sobre o que está sendo

demandado, ou as necessidades são descritas com objetos tão amplos que certamente deveriam ter sido desdobradas em um número de demandas (necessidades ou projetos/ações) menores e mais específicas.

Como exemplos de pouca clareza das necessidades inventariadas, foram encontrados itens descritos (literalmente) como: “levantamento de necessidades e implantação de recursos para a melhoria da disponibilidade da infraestrutura de TIC”, “aprimorar o processo de tomada de decisões estratégicas”, “implementar tecnologias de gestão do conhecimento em apoio aos assuntos estratégicos”, “desenvolver sistemas de informação baseados em computadores com o emprego de plataformas padronizadas, utilizando processos definidos e maduros”, “ampliação da produtividade e da capacidade de entrega”, “desenvolvimento de aplicativos móveis”, “realizar desenvolvimento de novos sistemas”, “aquisição de produtos/sistemas para atendimento a necessidades específicas das áreas de negócio”, “centro de monitoramento integrado”, “solução para gestão de recursos humanos”, “implementar/adquirir/customizar sistemas informatizados de acordo com as necessidades declaradas após estudos de viabilidade”, “melhoria da qualidade dos serviços de TI prestados pelo órgão a seus usuários”, “aprimoramento das redes dos prédios”, “incrementar o uso de soluções de mercado”.

Outro aspecto preocupante é que, na esmagadora maioria dos órgãos pesquisados do Poder Executivo Federal, são encontradas necessidades aparentemente redundantes no mesmo PDTIC, que poderiam ser consolidadas mediante uma abordagem mais corporativa e integrada do atendimento, talvez resultantes da ausência de análises técnicas mais aprofundadas sobre as necessidades de TIC apresentadas pelas diversas unidades internas dos órgãos.

Os dados da **TABELA 5** mostram a lista das 10 palavras nominativas (substantivos que geram denominações de objetos, ações ou ideias) mais frequentes encontradas nos capítulos de “Necessidades de TIC” dos PDTIC pesquisados, tendo-se excluído dessa tabela as palavras (siglas) “TI” e “TIC”, que, pela relação direta com o tema geral dos textos, se repete milhares de vezes sem acrescentar novidade (informação relevante). A terceira coluna da esquerda para a direita apresenta as respectivas frequências com que essas palavras, isoladamente, são encontradas nos textos, em números absolutos (contagem total), e a quarta e última coluna, da esquerda para direita, mostra as quantidades de *clusters* de substantivos (sintagmas nominais) formados mediante combinação com cada uma das palavras mais frequentes observadas (no singular e com flexões no plural).

TABELA 5 – 10 Palavras mais frequentes e *clusters* formados

#	Palavra	Frequência da Palavra	Frequência de <i>Clusters</i>
1	sistema/sistemas	3.360	2.049
2	serviço/serviços	1.856	1.824
3	informação/informações	1.840	1.482
4	gestão	1.580	2.118
5	dado/dados	1.331	1.011
6	projeto/projetos	1.330	680
7	processo/processos	1.316	963
8	solução/soluções	1.241	953
9	aquisição/aquisições	1.001	763
10	software/software s	908	1.021

Fonte: Elaborada pelo autor (2022).

Os *clusters* de palavras são identificados nos textos e computados, em termos de frequências de ocorrências, somente com uso de um *software* específico de PLN, como o relativo à obra de Scott (2019), com várias outras funções de busca e de estatística textual não encontradas em processadores de textos comuns. As estruturas linguísticas em *clusters* de palavras, no caso composições de substantivos com outros substantivos, são semanticamente mais ricas que substantivos isolados, como no seguinte exemplo: “gestão” não esclarece muito sobre o contexto de aplicação da gestão, mas “gestão da informação” identifica a área de aplicação da gestão, apresentando um contexto mais específico, com maior poder explicativo; o mesmo ocorre com o substantivo “segurança”, que, composto com “informação”, se torna “segurança da informação”.

Entretanto, como medida de economia de recursos computacionais de pesquisa (tempo de processamento de textos), apenas os *clusters* com frequência igual ou maior que 10 foram somados, inferindo-se, assim, que as frequências de sintagmas nominais formados com as 10 palavras mais frequentes (*top ten*) serão mais numerosas ainda. Os *clusters* mais frequentes são compostos com a palavra “gestão”, somando 2.118 aparições, que geralmente pressupõe a gestão de alguma coisa, exigindo, portanto, um complemento nominal, apresentando-se em textos técnicos como “gestão de ...”; de outro modo, a composição do *cluster* também pode aparecer como “processo(/s) de gestão”, em que “gestão” é o complemento nominal de outro substantivo (a propósito, essa composição de *clusters* apareceu 173 vezes no conjunto dos PDTIC pesquisados).

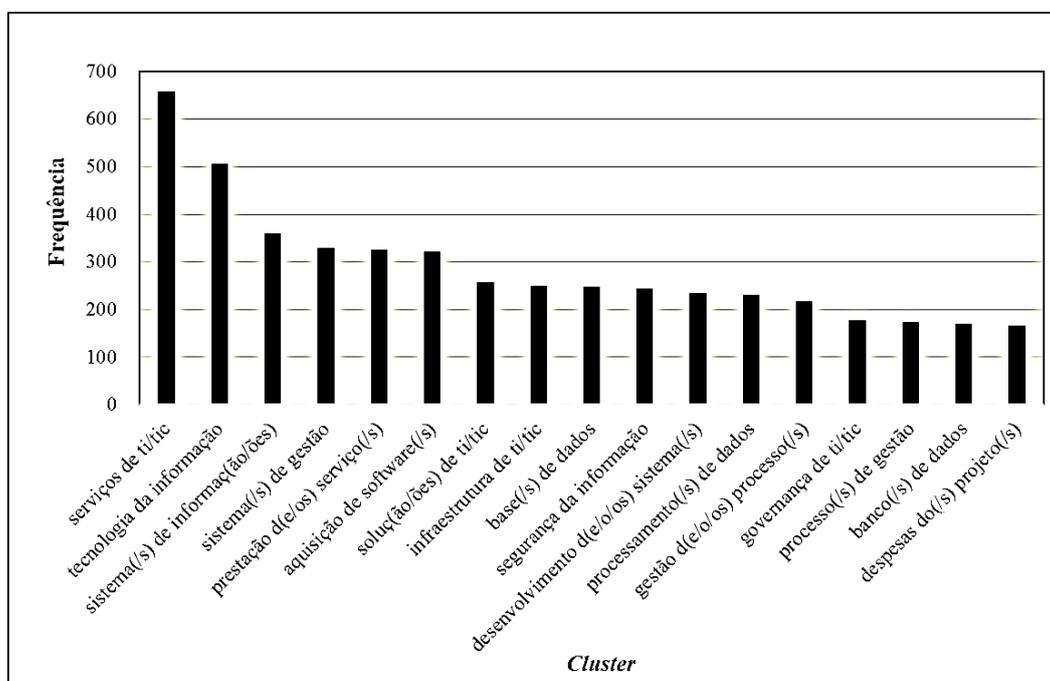
A segunda estrutura de *clusters* mais frequentes é composta com a palavra “sistema(/s)”, com 2.049 aparições, e a terceira com “serviço(/s)”, somando 1.824 aparições. Essa quantidade de *clusters* observados (com frequência igual ou maior que 10), relativos

ao conceito de “serviço(/s)”, denota uma preocupação dos órgãos com o provimento de serviços baseados em TIC, um sinal positivo dos objetivos maiores dos PDTIC, ainda que essa evolução não ocorra na velocidade almejada.

Os resultados da busca por *clusters* mais frequentes nos PDTIC pesquisados são apresentadas no **GRÁFICO 1**. A expressão nominativa “serviços de TIC” aparece como o *cluster* mais frequente, com 658 aparições, seguida por “tecnologia da informação”, com 506 ocorrências, e “sistema(/s) de informação(ões)”, com 361 ocorrências. Observe-se que, das 10 palavras mais frequentes apresentadas na **TABELA 5**, apenas 2 delas não compõem *clusters* entre os 10 mais frequentes: processo(/s) e projeto(/s).

Considerando que o termo “solução”, a partir da Instrução Normativa nº 4/2010-SLTI/MPOG, tornou-se uma expressão padrão para se referir a um conjunto de recursos de TIC combinados e destinados à solução de um problema de negócio, buscou-se nos PDTIC, nos capítulos sobre “necessidades de TIC”, os sintagmas nominais (expressões nominativas compostas) estruturados como “solução/soluções de ...” e “solução/soluções para ...”, observando-se, então, nos sintagmas que retornaram, o termo nominativo complementar na sequência, que qualifica a palavra “soluç(ão/ões)”. Com essa operação de busca textual, obteve-se os resultados apresentados na **GRÁFICO 2**, em que “soluç(ão/ões) de ti/tic”, “soluç(ão/ões) de gestão” e “soluç(ão/ões) de segurança” aparecem como os três *clusters* mais frequentes.

GRÁFICO 1 – *Clusters* mais frequentes entre “necessidades de TIC”

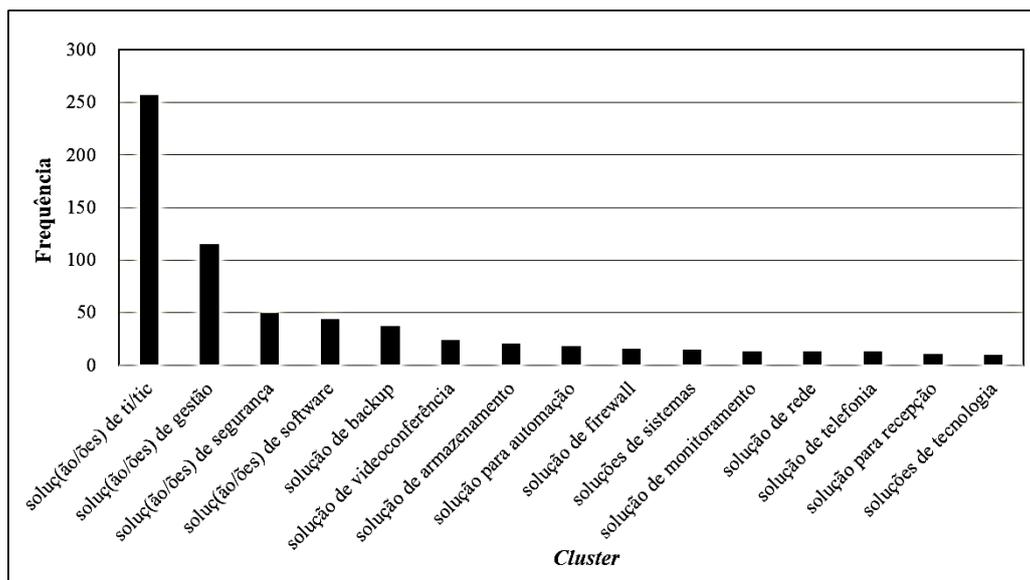


Fonte: Elaborado pelo autor (2022).

Esses resultados corroboram os resultados apresentados na **FIGURA 1** sobre a alta frequência de necessidades de soluções para “segurança da informação” expressas

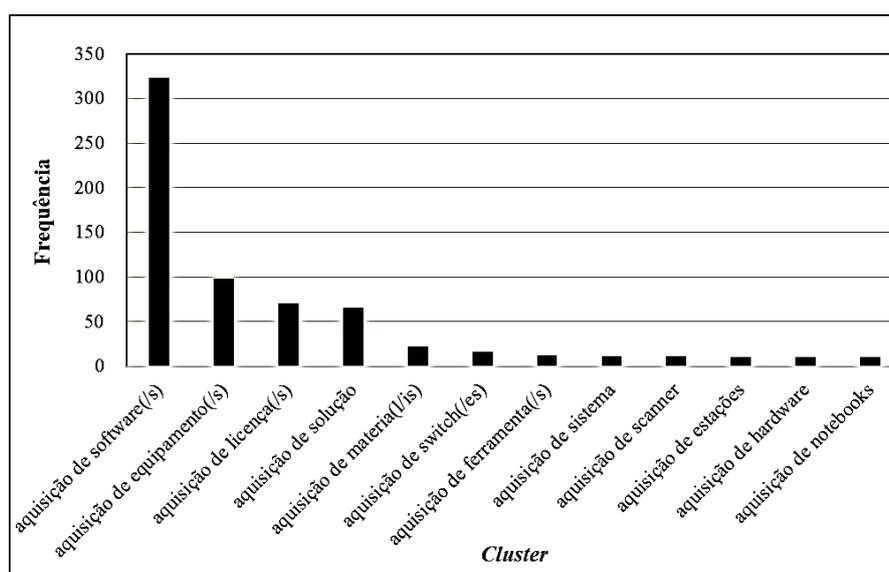
nos PDTIC pesquisados, sendo interessante observar, também, a expressiva frequência de uma necessidade de solução pouco comentada: “solução de videoconferência”. Quanto às outras necessidades citadas, como “solução/soluções de *backup*”, “solução de rede”, “solução de armazenamento” e “solução de telefonia”, por exemplo, essas não constituem novidade nos PDTIC, por constituírem necessidades de infraestrutura comuns aos órgãos.

GRÁFICO 2 – *Clusters* mais frequentes com “soluç(ão/ões)”



Fonte: Elaborado pelo autor (2022).

Com o **GRÁFICO 3**, novamente se evidencia o poder de tecnologias PLN, apresentando-se, entre os 10 *clusters* mais frequentes (com frequências iguais ou maiores que 10) verificados entre as “Necessidades de TIC”, uma série de *clusters* interessantes compostos pela palavra “aquisição”. Essas expressões revelam intenções de compras de componentes de TIC.

GRÁFICO 3 – Clusters mais frequentes com “aquisição”

Fonte: Elaborado pelo autor (2022).

Outro aspecto interessante dessa análise sintagmática se refere aos *clusters* compostos com a palavra “gestão”: somando-se as frequências dos *clusters* com as estruturas “sistema(/s) de gestão” e “soluções de gestão”, são mais de 400 *clusters* contendo essa ideia – a de soluções ou sistemas de informações computacionais para gestão dos órgãos. Conclui-se, ante essa evidência textual presente nos PDTIC pesquisados, pelo menos uma genuína preocupação dos órgãos públicos da União com a automação de seus processos de gestão.

Em relação a orçamentos de TIC, os planos pesquisados revelaram que os maiores são também os de órgãos maiores em todos os sentidos, com unidades regionais descentralizadas no país e quadros maiores de servidores públicos, tais como Ministério da Saúde, Ministério da Fazenda, Departamento Nacional de Infraestrutura de Transportes, Ministério da Justiça e Segurança Pública e Ministério da Agricultura, Pecuária e Abastecimento. Verificou-se, também, que a questão orçamentária não é endereçada em todos os órgãos do Poder Executivo da União, com somente 49 dos 59 órgãos pesquisados (83,1%) tendo estimado suas despesas.

Quanto a valores, o montante orçamentário médio de cada PDTIC pesquisado é de R\$ 115,8 milhões no horizonte temporal médio de planejamento (que é de 2,8 anos), resultando numa média anual de despesas por órgão de aproximadamente R\$ 41,3 milhões.

CONCLUSÃO

Conclui-se, com base no tamanho da amostra e nos resultados da pesquisa apresentada neste artigo, primeiro, que tecnologias de PLN podem contribuir, efetivamente, em processos de avaliação semiautomática de conteúdos textuais de documentos no setor público, como no caso de planos; e, segundo, que os PDTIC publicados entre 2013 e 2019

parecem elaborados em conformidade com o Guia do SISP e, assim, contribuem para a evolução tecnológica dos órgãos, mas ainda carecem de melhorias técnicas significativas para se tornarem instrumentos mais efetivos nesse sentido.

Os resultados da pesquisa revelam evolução positiva em questões atuais da gestão de TIC, com frequentes menções aos temas de governança e gestão de riscos, e interesse expressivo na adoção de tecnologias disruptivas por alguns órgãos do Poder Executivo Federal, tais como as tecnologias de georreferenciamento digital e de aplicações para dispositivos móveis, ainda que em passo talvez excessivamente cauteloso ante a relação custo/benefício dessas tecnologias. Outro aspecto positivo é que alguns modelos de gestão de TIC mais polêmicos, mesmo que adotados por empresas do mercado, parecem estar sendo sopesados no âmbito do SISP ante sua complexidade, numa eventual implementação em escala no setor público federal, como no caso da computação em nuvem (*cloud computing*).

Outro aspecto relevante, que se refere a uma descoberta da pesquisa por serendipidade, é que algumas unidades de gestão de TIC estão ousando inovar no modelo de gestão de demandas, contrapondo, tecnicamente (com dados), de um lado as necessidades de TIC e, de outro lado, estimativas da real capacidade de entrega de suas equipes. Com isso, a gestão de TIC nos órgãos poderá se tornar uma atividade mais baseada em dados e evidências, portanto mais realista, em conformidade com as recomendações do Tribunal de Contas da União no Acórdão 1603/2008-TCU-Plenário (Brasil, 2008), sobre as necessidades de recursos humanos nas unidades de TIC dos órgãos, e no Acórdão 2608/2018-TCU-Plenário (Brasil, 2018), sobre critérios técnicos para elaboração de políticas públicas, planos e programas.

Os dados estatísticos sobre os conteúdos textuais também revelam, por outro lado, conforme a percepção de uma autoridade governamental publicada recentemente, que, de um modo geral, os órgãos do Poder Executivo Federal ainda carecem de uma evolução em TIC “da porta para dentro” (Batista; Azevedo; Hessel, 2020), para resolver seus gargalos de infraestrutura tecnológica e de integração de sistemas de informações computacionais que impedem o pleno atendimento ao cidadão. O nível de detalhamento das demandas inseridas nos PDTIC, geralmente superficial (e com redundâncias), também não contribui para a avaliação de esforços de atendimento e uma melhor gestão das necessidades de TIC nos órgãos.

O pouco interesse revelado na elaboração sobre arquitetura de informação e arquitetura técnica pode estar contribuindo para esse lento avanço dos órgãos do SISP no atendimento aos requisitos técnicos de integração, interoperabilidade e portabilidade de sistemas previstos na Portaria nº 92/SLTI/MPOG, de 24 de dezembro de 2014 (que instituiu a arquitetura de interoperabilidade “ePING”), assim como aos objetivos institucionais propostos no Decreto nº 10.046, de 9 de outubro de 2019 (sobre a governança nos processos de compartilhamento de dados).

Considerando uma analogia natural entre arquitetura de TIC e arquitetura tradicional (da construção civil), desenvolver um ambiente computacional sem arquitetura de informação e arquitetura técnica equivale a se construir um prédio sem projeto arquitetônico, assumindo-se os riscos de inadequação dos espaços e uso inadequado de onerosos recursos de TIC.

Obviamente, nessas questões, não se pode ignorar aspectos adversos que extrapolam a governabilidade das unidades de TIC nos órgãos, tais como problemas de estratégias e processos de negócios corporativos inadequados e modelos de governança inadequados ou ausentes. O exemplo mais evidente se refere ao nível hierárquico em que se situam os titulares das unidades de gestão de TIC nos órgãos, em sua esmagadora maioria situados em posições intermediárias, o que demonstra, na prática, o nível de prioridade para a gestão de TIC.

REFERÊNCIAS

ABRAMS, C.; SCHULTE, R. W. **Service-oriented architecture overview and guide to SOA research**. [s. l.]: Gartner Group, 2008.

BATISTA, V.; AZEVEDO, A.; HESSEL, R. Força-tarefa contra filas no INSS ganha reforço. **Correio Braziliense**, Brasília, 17 jan. 2020. Economia, p. 7.

BELL, M. **Service-oriented modeling: service analysis, design, and architecture**. New Jersey: Wiley, 2008. ISBN 0470141115.

BRASIL. **Decreto-lei nº 200**, de 25 de fevereiro de 1967. Dispõe sobre a organização da Administração Federal, estabelece diretrizes para a Reforma Administrativa, e dá outras providências. Presidência da República, Brasília, 1967.

BRASIL. **Decreto-lei nº 7.579**, de 11 de outubro de 2011. Dispõe sobre o Sistema de Administração dos Recursos de Tecnologia da Informação–SISP, do Poder Executivo federal. Presidência da República, Brasília, 2011.

BRASIL. Ministério do Planejamento, Orçamento e Gestão. **Guia de elaboração de PDTI do SISP: Versão 1.0**. Brasília: MPOG, 2012. Disponível em: http://www.sisp.gov.br/.../Guia_de_Elaboração_de_PDTI_v1.0...pdf. Acesso em: 2 jan. 2020.

BRASIL. Tribunal de Contas da União. **Acórdão nº 2608/2018-TCU-Plenário**. Lei 13.707/2018, art. 124. Relatório de políticas públicas elaborado com objetivo de subsidiar o Congresso Nacional na elaboração da Lei Orçamentária. Acórdão 2127/2017-Plenário. Monitoramento. Adoção de diversas medidas, por parte do Poder Executivo, para atender às recomendações. Oportunidade de melhorias. Recomendações aos órgãos mencionados e a unidades deste Tribunal. Brasília: Tribunal de Contas da União, 2018. Disponível em: <https://pesquisa.apps.tcu.gov.br/#/documento/acordao->. Acesso em: 25 out. 2020.

BRASIL. Tribunal de Contas da União. **Acórdão nº 1603/2008-TCU-Plenário**. Levantamento de auditoria. Situação da governança de Tecnologia da Informação – TI na Administração Pública Federal. Ausência de Planejamento Estratégico Institucional. Deficiência na estrutura de pessoal. Tratamento inadequado à confidencialidade, integridade e disponibilidade das informações. Recomendações. Brasília: Tribunal de Contas da União, 2008. Disponível em: <https://pesquisa.apps.tcu.gov.br/#/resultado/acordao-completo/>. Acesso em: 25 out. 2020.

CONEGLIAN, C. S.; SEGUNDO, J. E. S. Inteligência artificial e ferramentas da *web* semântica aplicadas a recuperação da informação: um modelo conceitual com foco na linguagem natural. **Informação e Informação**, Londrina, v. 27, n. 1, p. 625–651, jan./mar. 2022. Disponível em: <https://ojs.uel.br/revistas/uel/index.php/informacao/article/view/44729>. Acesso em: 9 dez. 2023.

COSTELO, K. **The Evolution of Enterprise Architecture**. Gartner Group, 2019. Disponível em: <https://www.gartner.com/smarterwithgartner/the-evolution-of-enterprise-architecture/>. Acesso em: 2 jan. 2020.

CRAWLEY, E. **Introduction to system architecture**: architecture to value. Lecture Notes. Massachusetts: MIT, 2007. Disponível em: <https://ocw.mit.edu/courses/engineering-systems-division/esd-34-system-architecture-january-iap-2007/lecture-notes/lec1.pdf>. Acesso em: 21 jan. 2020.

FALCÃO, L. C. J.; LOPES, B.; SOUZA, R. R. Absorção das tarefas de processamento de Linguagem Natural (NLP) pela Ciência da Informação (CI): uma revisão da literatura para tangibilização do uso de NLP pela CI. **Em Questão**, Porto Alegre, v. 28, n. 1, p. 13-34, jan./mar. 2022. Disponível em: <https://seer.ufrgs.br/EmQuestao/article/view/111323/64623>. Acesso em: 9 dez. 2023.

GARTNER GROUP. **Gartner glossary**: enterprise architecture. [s. l.]: Gartner Group, 2020a. Disponível em: <https://www.gartner.com/en/information-technology/glossary/enterprise-architecture-ea>. Acesso em: 2 jan. 2020.

GARTNER GROUP. **Gartner Glossary**: Service-oriented Architecture (SOA). [s. l.]: Gartner Group, 2020b. Disponível em: <https://www.gartner.com/en/information-technology/glossary/service-oriented-architecture-soa>. Acesso em: 22 jan. 2020.

HERMANS, P. **The Zachman Framework for architecture revisited**: on conceiving the informational enterprise. *Semiosis*, [s. l.], p. 1-18, 2015.

KAO, A.; POTEET, S. R. (ed.). **Natural language processing and text mining**. London: Springer, 2007.

KONCHADY, M. **Text mining application programming**. Massachusetts: Charles River Media, 2006.

KOTUSEV, S. Enterprise architecture: what did we study? **International Journal of Cooperative Information Systems**, [s. l.], v. 26, n. 4, 2017. Disponível em: https://www.academia.edu/36770941/Enterprise_Architecture_What_Did_We_Study. Acesso em: 24 out. 2020.

KOTUSEV, S. The history of enterprise architecture: an evidence-based review. **Journal of Enterprise Architecture**, [s. l.], v. 12, n. 1, p. 29-27. 2016. Disponível em: https://www.researchgate.net/publication/308936998_The_History_of_Enterprise_Architecture_An_Evidence-Based_Review/link/.../download. Acesso em: 24 out. 2020.

LIMA-MARQUES, M.; MACEDO, F. L. O. Arquitetura da Informação: base para a gestão do conhecimento. In: TARAPANOFF, K. (org.). **Inteligência, Informação e Conhecimento**. Brasília: UNESCO: IBICT, 2006, p. 241-255.

MEADOW, C. T.; BOYCE, B. R.; KRAFT, D. H.; BARRY, C. **Text information retrieval systems**. 3. ed. [s. l.]: Elsevier, 2007.

NICOLAIDOU, M.; TSADIMAS, A.; ALEXOPOULOU, N.; ANAGNOSTOPOULOS, D. Employing Zachman Enterprise Architecture Framework to Systematically Perform Model-Based System Engineering Activities. *In: PROCEEDINGS OF THE 42ND HAWAII INTERNATIONAL CONFERENCE ON SYSTEM SCIENCES 2009*, 42., 2009, Havaí. **Anais** [...] Hawaii: HICSS, 2009.

ROSS, J. W.; WEILL, P.; ROBERTSON, D. C. **Arquitetura de TI como estratégia empresarial**. Tradução Roger Maioli dos Santos. São Paulo: M. Books, 2008.

SCOTT, M. **WordSmith Tools Manual**. Version 7.0. Stroud: Lexical Analysis Software Ltd., 2019.

WURMAN, R. S. **Information architects**. New York: Graphis, 1997.

ZACHMAN, J. A. **The concise definition of the Zachman Framework**. [s. l.], Zachman International, 2008.

ZACHMAN, J. A. A framework for information systems architecture. **IBM Systems Journal**, [s. l.], v. 26, n. 3, 1987, p. 276-292.