



Evaluación semiautomática de contenidos de Planes Directores de TIC con tecnologías de Procesamiento del Lenguaje Natural

Ethel Airton Capuano

Doctor en Ciencias de la Información por la Universidad de Brasilia (UnB).

Universidad de Brasilia (UnB), Brasilia, Brasil.

Funcionario, Ministerio de Integración y Desarrollo Regional, Brasilia, Brasil.

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



Enviado en: 03/04/2022. Aprobado en: 25/01/2024. Publicado en: 21/06/2024.

RESUMEN

Este artículo presenta los resultados de una investigación sobre búsqueda textual y evaluación semiautomática de contenidos relevantes sobre los Planes Directores de Tecnologías de la Información y la Comunicación publicados por una muestra representativa de órganos públicos del gobierno federal. Utilizando tecnologías de procesamiento del lenguaje natural, se extrajo información útil de los documentos de los planes investigados para evaluar el estado del arte de las TIC en los órganos, y se obtuvo resultados interesantes, algunos de los cuales causaron sorpresa (por casualidad), como el modelo de gestión de demandas en algunas unidades de TIC basado en estimaciones de la capacidad de entrega de los equipos. Se concluye que los planes cumplen con las normas oficiales, observándose incluso un interés de la mayoría de los órganos en temas de gestión actuales, como la gobernanza y la gestión de riesgos, y una madurez en la evaluación de las tendencias del mercado, pero también que sus contenidos aún carecen de perspectivas para algunas prácticas importantes que orienten su evolución técnica.

Palabras clave: procesamiento del lenguaje natural; minería de textos; evaluación semántica de planes directores de TIC.

INTRODUCCIÓN

Este artículo presenta los resultados de una investigación sobre búsqueda textual y evaluación semiautomática de contenidos tecnológicos de los Planes Directores de Tecnologías de la Información y la Comunicación (PDTIC) de órganos públicos de los Poderes Ejecutivo, Legislativo y Judicial del Estado, más el Ministerio Público Federal, vigentes entre 2013 y 2019, por medio del uso de tecnologías de Procesamiento del Lenguaje Natural (PLN).

La motivación temática de la investigación se debe tanto a la idea de probar la posibilidad de automatización de procesos de evaluación de contenidos textuales de documentos del sector público con tecnologías de PLN, como a una curiosidad científica en relación a los propios contenidos tecnológicos tratados por los órganos en sus PDTIC, en un contexto de evaluación multidimensional y multidisciplinar. Considerando, como premisa de investigación, que los PDTIC reflejan, aunque con algunas limitaciones, el estado del arte de las TIC en los órganos públicos federales, algunas preguntas relevantes que podrían ser mejor comprendidas con los resultados de la investigación son: “¿Cómo está el ‘estado del arte’ de las TIC en los órganos públicos del Estado en términos de organización interna (arquitectura)?”; “¿Cómo se abordan los temas actuales de las TIC?”; “¿Cómo se abordan los temas más técnicos y estructurales de gestión de las TIC?”; “¿Cuáles son los órganos con los mayores presupuestos para las TIC y que, por lo tanto, requieren una mayor atención en términos de gobernanza, gestión de riesgos y controles?”; “¿Cómo están los equipos de las unidades de TIC de los órganos y entidades, como fuerzas de trabajo?”.

Aunque el estado del arte de las TIC en los órganos públicos federales ha evolucionado desde el advenimiento (en Brasil) de la *World Wide Web* e *Internet* en la segunda mitad de la década de 1990, existe cierto consenso sobre el hecho de que esta evolución no se ha producido al ritmo necesario para el progreso deseado de los proyectos de digitalización de los servicios públicos, como puede observarse en los cuellos de botella revelados en la prestación de algunos servicios importantes en la actualidad (Batista; Azevedo; Hessel, 2020). Aunque se reconoce que, en algunos sectores, las inversiones continuas en servicios digitales proporcionados por robustas arquitecturas de TIC han tenido un éxito innegable, como en el caso de los sistemas de declaración de impuestos (que atienden a decenas de millones de contribuyentes que acceden simultáneamente a los canales de servicios automáticos) y en el caso de los tribunales de justicia, que han digitalizado casi todos sus servicios procesales, en otros servicios públicos las inversiones de las últimas décadas no parecen arrojar resultados satisfactorios.

Como ejemplos del liderazgo de los tribunales de justicia en innovaciones con uso intensivo de Tecnologías de la Información y la Comunicación (TIC), el Supremo Tribunal Federal (STF), en la segunda mitad de la década de 1990, innovó en el sector público brasileño al implementar un sistema *push* (en el sentido de “empujar”, automáticamente, la información a los usuarios en la red mundial de computadoras) para que los profesionales operadores del derecho pudieran seguir el progreso de los procesos de sus clientes de forma remota,

a través de *Internet*. Los órganos del Poder Judicial también innovaron, aún en la primera década de este milenio, al establecer canales completamente digitales para el intercambio de documentos procesales con la unidad jurídica del Instituto Nacional del Seguro Social (INSS), dada la materialidad (en cantidad y valores) de demandas contra esa autarquía demandada. Y, recientemente, el Tribunal Regional Federal de la 4ª Región (TRF-4) innovó al implementar, con notable éxito, el primer sistema de Gestión Electrónica de Documentos (GED) diseñado específicamente para el sector público brasileño: el Sistema Electrónico de Información (SEI), que también se está adoptando progresivamente en los órganos del Poder Ejecutivo Federal.

El objetivo general de la investigación, por lo tanto, se centró en las pruebas prácticas de tecnologías de PLN para una evaluación más sustantiva y menos formal de la planificación de las TIC en los órganos y entidades del Estado. En resumen, este objetivo general consistió en verificar, mediante herramientas de automatización de procesos de búsqueda en textos, si este modelo de planificación gubernamental podría contribuir efectivamente al desarrollo de soluciones de TIC adecuadas según los estándares y mejores prácticas del mercado y la jurisprudencia de los órganos de control. Como objetivos específicos, se buscaron en los textos de los planes evidencia de temas técnicos relevantes y actuales abordados por los órganos y entidades.

METODOLOGÍAS, PROCESOS Y TECNOLOGÍAS DE INVESTIGACIÓN

El método de investigación utilizado es multidisciplinario, con enfoques cualitativos (semánticos) relacionados con la búsqueda, recuperación y procesamiento de contenido textual (Coneglian; Segundo, 2023; Kao; Poteet, 2007; Meadow *et al.*, 2007; Konchady, 2006), pero también con enfoques cuantitativos, utilizando estadísticas textuales (Scott, 2019), con el fin de evaluar los niveles de interés de los órganos en cada tema de investigación revelado en los PDTIC.

Como áreas de conocimiento, la metodología de investigación puede enmarcarse, en términos teóricos (aunque parcialmente), en la Lingüística (Scott, 2019) y en la Ciencia de la Información (Falcão; Lopes; Souza, 2022; Meadow *et al.*, 2007), esta última en la disciplina denominada "*Búsqueda y Recuperación de Información*" (*Information Retrieval*). Estas dos áreas de conocimiento sustentan metodológicamente las dos principales disciplinas instrumentales utilizadas en la investigación: el Procesamiento del Lenguaje Natural (PLN, y, en inglés, *Natural Language Processing, NLP*) y el Análisis Textual (*Analytics*). Este conjunto de tecnologías también contribuye a la composición de soluciones en los conceptos de mercado conocidos como Minería de Textos (*Text Mining*) y Ciencia de Datos.

La información textual recuperada de los textos de los PDTIC se refiere tanto a atributos seleccionados *a priori*, en la etapa de planificación del proyecto de investigación, como a *cluster* de palabras sustantivas e interesantes basadas en palabras clave más frecuentes. Se utilizaron tecnologías más sencillas, como *software* editores electrónicos

de texto, y otras más sofisticadas como una aplicación de *software* de mercado para el Procesamiento del Lenguaje Natural (PLN), que implementa funciones de recuperación de información más avanzadas, como *cluster* y estadísticas sobre textos en idioma portugués (Scott, 2019). Además, se empleó el análisis de discurso como método de selección de información relevante sobre algunos atributos de interés en los PDTIC recopilados.

Los documentos se recuperaron de los portales digitales de los órganos mediante la búsqueda en la *World Wide Web (Web)*, *por Internet*. Se excluyó, por decisión del proyecto, a las Instituciones Federales de Enseñanza (IFE) y a las empresas estatales. En el Poder Ejecutivo Federal, la muestra de PDTIC investigados, con un total de 72 órganos, abarca 17 de los 23 ministerios vigentes a partir de la Ley N 13.502, del 1 de noviembre de 2017 (utilizando, sin embargo, el PDTIC 2017-2019 del nuevo Ministerio de Justicia y Seguridad Pública y no los PDTIC del Ministerio de Justicia y el Ministerio de Seguridad Pública previstos en la última modificación de esta Ley), 42 de las 80 autarquías y fundaciones restantes después de la exclusión de las IFE, dos órganos del Poder Legislativo Federal, 10 órganos del Poder Judicial Federal y el Ministerio Público Federal.

El conjunto de información textual generada a partir de la recolección de documentos y el análisis textual de los PDTIC se tabuló en una planilla electrónica para un adecuado tratamiento estadístico, contabilizando en columnas los contenidos según atributos a observar en los documentos, y en filas, los respectivos órganos y entidades separados según una categorización político-administrativa dentro del Estado: tres Poderes (muestra con 72 órganos y entidades), Poder Ejecutivo (59 órganos), Poder Ejecutivo – Administración Directa (17 órganos), Poder Ejecutivo – Administración Indirecta (42 autarquías y fundaciones), Poder Legislativo (dos órganos), Poder Judicial (10 órganos), Ministerio Público Federal (un órgano).

En resumen, se recuperaron de cada PDTIC contenido textual sobre 37 atributos de investigación, dispuesto en la planilla de acuerdo con los siguientes bloques de análisis y formatos de datos:

- I. Análisis Documental Básico (con siete atributos): volumen del archivo (en MB); fuente y tamaño de fuente de los PDTIC (texto); interlineado (número real); cantidad de páginas (número entero); cantidad de anexos (número entero); si la versión del plan era nueva o una extensión del plan anterior (dato binario: “sí” o “no”); y horizonte de planificación en años (número entero).
- II. Existencia de Capítulos Temáticos (con siete atributos) sobre: estrategias de negocios y de TIC articuladas en el PDTIC (“sí” o “no” para cada estrategia); arquitectura empresarial, de la información y técnica (o de sistemas) (“sí” o “no” para cada arquitectura); metodología de gestión de proyectos de TIC (“sí” o “no”); gestión de riesgos (“sí” o “no”); gobernanza (“sí” o “no”).
- III. Necesidades de TIC (con nueve atributos): cantidad de páginas que expresan las necesidades (o demandas) de TIC (número entero); ítems de necesidades de TIC (número entero); ítems eventuales de proyectos de TIC (número entero); claridad en la presentación de las demandas (“sí” o “no”); análisis técnico de las demandas de TIC

(antes de la inserción en el PDTIC) (“sí” o “no”); registros de demandas redundantes (“sí” o “no”); mención a Metodología de Desarrollo de Sistemas (MDS) (“sí” o “no”); mención a solución de Business Intelligence, BI (“sí” o “no”); uso de metodología para priorizar la satisfacción de las necesidades (“sí” o “no”).

IV. Desempeño del PDTIC Anterior (un atributo): porcentaje que revela la porción ejecutada de las necesidades y/o proyectos planificados.

V. Presupuesto Estimado del PDTIC (un atributo): datos monetarios relativos a la previsión de gastos para la ejecución del plan en el horizonte previsto.

VI. Declaración de Vacíos de Recursos Humanos en el Equipo de TIC (un atributo): dato binario (“sí” o “no”), basado en una declaración de la unidad de gestión.

VII. Frecuencia de Palabras y Expresiones Temáticas (11 atributos): estrategia; georreferenciación; Gestión de la Información; Gestión del Conocimiento; Seguridad de la Información; GIS/SIG; Imagen Satelital; Drone/VANT; Computación en la nube (Cloud Computing); Big Data; Aplicación Móvil (App).

Los parámetros estadísticos calculados en la planilla, para cada categoría político-administrativa de órganos y entidades del Estado, fueron:

- a) Promedio: para todos los atributos, pero utilizando un dato numérico sumando la presencia de “sí” en cada PDTIC analizado en el caso de datos binarios.
- b) Desviación estándar: ídem.
- c) Coeficiente de variación (relación entre la desviación estándar y la media): ídem.

En cuanto a sus estructuras documentales, los planes investigados de los órganos del Poder Ejecutivo siguen, en la mayor parte de sus contenidos, la “Guía de Elaboración de PDTI del SISP” (Brasil, 2012), en un razonamiento lógico a partir de los principios y directrices, estrategias de TIC, plan de trabajo para la elaboración del PDTIC, Análisis FODA, evaluación de recursos humanos del equipo de TIC, método de priorización adoptado, inventario de necesidades y cartera de proyectos y acciones, evaluación de riesgos, plan de metas y acciones y presupuesto estimado.

REVISIÓN DE LITERATURA

Aunque para el público familiarizado con la gestión de TIC en los órganos públicos la mayoría de los conceptos que se abordan en este artículo no representan una novedad, algunos de ellos necesitan una definición para evitar interpretaciones equivocadas en relación con el contexto de la investigación, y considerando el contexto institucional de estas organizaciones, también merecen destacarse las normas de cumplimiento.

Conceptos técnicos

El primer concepto importante es el de “Plan Director de Tecnologías de la Información y la Comunicación”, que en el Decreto 8.638/2016¹ se define expresamente como un “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”.

Sin embargo, el concepto de “diagnóstico” adoptado en la guía de referencia (Brasil, 2012) merece especial atención, ya que puede no coincidir exactamente con lo que se entiende por diagnóstico en otros contextos, como se muestra en el siguiente texto:

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)².

Otro concepto primordial, adoptado como motivación de la investigación, es el de “adecuación técnica” de los PDTIC, que no debe confundirse necesariamente con la adecuación formalmente exigida por las normas de regencia, ya que estas no abordan en profundidad todos los temas relevantes para la gestión de las TIC en los órganos, tales como “arquitectura empresarial” (también denominada “arquitectura corporativa”), “arquitectura de información” y “arquitectura técnica” (o de sistemas).

Los conceptos relacionados con la “arquitectura”, por otro lado, cuando se aplican al contexto de la gestión de las TIC en las organizaciones, se abordan en varios informes y artículos técnicos producidos, por ejemplo, por la empresa especializada en investigaciones de mercado *Gartner Group*, bien conocida por los gestores de TIC, y otros autores (Gartner Group, 2020a; Costelo, 2019; Hermans, 2015; Ross; Weill; Robertson, 2008). Sin embargo, originalmente, el concepto de “arquitectura”, en un contexto más amplio de gestión empresarial con uso intensivo de sistemas de información computacional, se debe a la iniciativa de

1 Traducción: “instrumento de diagnóstico, planificación y gestión de los recursos y procesos de las tecnologías de la información y la comunicación, con el objetivo de satisfacer las necesidades finales e informativas de un órgano o entidad para un periodo determinado” (Brasil, 2016, traducción editorial).

2 Traducción: “Después de la fase de preparación, en la cual se organizaron las principales actividades del proyecto de elaboración del PDTI, comienza la segunda fase del proceso: la fase de diagnóstico. Esta fase se caracteriza por buscar comprender la situación actual de la TI en la organización para, en consonancia con este panorama, identificar las necesidades (problemas u oportunidades) que se espera resolver. Para ello, contempla procesos relacionados con el análisis de la planificación anterior, análisis estratégico y levantamiento de necesidades. (...) El levantamiento de necesidades parte con aquellas relacionadas con la información y se desglosa en todas las demás asociadas a la TI: servicios, infraestructura, contrataciones y personal de TI. (...) Uno de los principales artefactos producidos en esta fase, y muy importante para todo el proceso de elaboración del PDTI, es el Inventario de Necesidades” (Brasil, 2012, p. 33, traducción editorial).

investigación pionera conocida como *Business Systems Planning (BSP)*, implementada en la década de 1960 por varios investigadores al servicio de IBM (Kotusev, 2016, 2017), para destacarse posteriormente el constructo descriptivo empresarial desarrollado por Zachman (1987), más conocido como “Marco de Trabajo Zachman” (*Zachman Framework*), que pretende describir e integrar de manera efectiva y eficiente los elementos ontológicos más relevantes que constituyen una empresa (Nicolaidou *et al.*, 2009; Zachman, 2008).

El concepto de “Arquitectura Empresarial” (EA, del inglés *Enterprise Architecture*), uno de los elementos de gestión investigados en los PDTIC, que en el presente contexto puede entenderse como “arquitectura empresarial”, es definido por la empresa de investigación de mercado de TIC *Gartner Group* de la siguiente manera:

[...] 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)³.

En resumen, en el contexto de la investigación, una “arquitectura corporativa” se encargaría de la identificación, descripción, evaluación e integración funcional de los principales recursos de un órgano público, de manera sistémica y sinérgica, para lograr sus objetivos institucionales. Estos recursos pueden clasificarse en: humanos, materiales, patrimoniales, tecnológicos, de información y de conocimiento.

Los conceptos de “arquitectura de la información” y “arquitectura de sistemas”, más recientes, que también se refieren a dos objetos temáticos investigados en los PDTIC, han sido definidos por un conjunto variado de otros autores, con significados a veces diversos entre ellos. El pionero en la idea de “arquitectura de la información” es Wurman (1997), quien la definió, en términos generales, como una expansión de la ocupación del arquitecto tradicional aplicada a la organización de la información. Lima-Marques y Macedo (2006), en la misma línea, pero con un enfoque filosófico *sui generis*, definen esta disciplina como *una metodología para la estructuración de sistemas de información aplicada a cualquier entorno de información, entendido como el espacio que integra contexto, contenido y usuarios*.

Sin embargo, el concepto más apropiado de “arquitectura de la información”, en el contexto de esta investigación, parece ser el de Hagedorn, según Lima-Marques y Macedo (2006, p. 245): *arte y ciencia de la organización de la información para ayudar efectivamente a las personas a satisfacer sus necesidades de información. Involucra investigación, análisis, diseño e implementación*.

3 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*).

El concepto de “arquitectura técnica”, también llamada “arquitectura de sistemas”, se refiere a la organización inteligente de los sistemas de información y las tecnologías que los sustentan en las organizaciones. Este tipo de arquitectura sería la base cognitiva para el uso de los componentes de soporte tecnológico de las organizaciones, integrando componentes de *hardware* (estaciones de trabajo, servidores, unidades de almacenamiento masivo, dispositivos de control de segmentación en redes de telecomunicaciones, etc.) y *software* básicos (como sistemas operativos y servidores de aplicaciones), además de otros componentes fundamentales para las operaciones computacionales, como aplicaciones de seguridad, respaldos y balanceo de carga.

Este concepto, presentado didácticamente por Crawley (2007), es importante para guiar incluso los procesos de adquisición o desarrollo de nuevos sistemas, la integración entre sistemas heredados, la interoperabilidad entre componentes y sistemas, la portabilidad de sistemas heredados de un entorno tecnológico obsoleto a uno más actual, la combinación entre tecnologías heredadas y nuevas, etc. El modelo más prometedor de “arquitectura técnica”, incluso para el momento actual de los órganos públicas federales, es el conocido como “Arquitectura Orientada a Servicios” (SOA, del inglés *Service-Oriented Architecture*), que incorpora aspectos lógicos y físicos, tal como lo define Bell (2008, p. 309):

A disciplina de arquitetura lógica orientada a serviço é principalmente voltada para o reuso, utilização e 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 dados⁴.

Este constructo de organización de las TIC también es comentado por Abrams y Schulte (2008), de la empresa *Gartner Group*, cuyo glosario técnico lo define de la siguiente manera:

[...] 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, manutenibilidade e valor. Isto resulta em sistemas interoperáveis e modulares, que são mais fáceis de usar e de manter. SOA cria sistemas mais simples e velozes, que incrementam a agilidade e reduzem o Custo Total de Propriedade (TCO) (Gartner Group, 2020b, *online*, tradução nossa)⁵.

4 Traducción: “La disciplina de arquitectura lógica orientada a servicios se centra principalmente en la reutilización, utilización y consumo. (...) esta disciplina se dirige especialmente a los tres aspectos funcionales más importantes de una arquitectura lógica: (1) la relación entre paquetes de software implementados, (2) las preocupaciones estructurales de las entidades de los paquetes orientados a servicios, y (3) el comportamiento y la colaboración de las soluciones empaquetadas. La arquitectura física orientada a servicios se refiere a todo lo que es físicamente direccionable a los componentes orientados a servicios. Este direccionamiento consiste en el mapeo de las entidades de *software* en componentes de hardware y sus ubicaciones direccionables en una red de comunicación de datos” (Bell, 2008, p. 309, traducción editorial).

5 Original: “[...] is a design paradigm and discipline that helps IT meet business demands [...] SOA reduces redundancy and increases usability, maintainability and value. This produces interoperable, modular systems that are easier to use and maintain. SOA creates simpler and faster systems that increase agility and reduce total cost of ownership (TCO)” (Gartner Group, 2020b, *online*).

Es importante señalar que esta estrategia basada en arquitecturas de información y TIC parece coincidir con la actual propuesta de cambio estructural en el modelo de prestación de servicios de los órganos del Poder Ejecutivo Federal, con una mayor interacción digital entre los órganos públicos y los ciudadanos (Batista; Azevedo; Hessel, 2020). (Batista; Azevedo; Hessel, 2020)

Sin embargo, en la práctica surge la pregunta: ¿cómo debería concebirse e implementarse una arquitectura técnica en las organizaciones? En respuesta, se argumenta que esta arquitectura puede elaborarse a partir de un diseño primordial concebido de acuerdo con el mejor pensamiento técnico en TIC, editando normas internas en las organizaciones para su uso obligatorio, que establece ciertos estándares y mejores prácticas, como estándares definidos tecnológicamente para sistemas operativos y sistemas de gestión de bases de datos relacionales (SGBDR).

Como ejemplo de una situación recurrente en la que las ventajas de una arquitectura técnica estandarizada podrían evitar gastos a la tesorería, imagine, por hipótesis, una nueva aplicación que se adquiere o desarrolla externamente (por órganos asociados, muchas veces), pero que requiere, para su instalación en el entorno tecnológico del órgano demandante, un producto de base de datos aún no utilizado y no dominado por el equipo de soporte de TIC de ese órgano. Como resultado, habrá una necesidad de adquisición y dominio técnico (del equipo de soporte de la base de datos), a corto plazo, de un nuevo *software* muy complejo, lo que contribuye a que los entornos de las TIC sean aún más costosos, y resulta en una sobrecarga de las unidades de gestión de las TIC con más tareas que podrían evitarse.

Estándares de cumplimiento

Los órganos del Poder Ejecutivo Federal deben llevar a cabo sus actividades de gestión de las TIC de acuerdo con los estándares técnicos y las mejores prácticas del mercado, pero también de acuerdo con las normas aprobadas por el órgano central del Sistema de Administración de Recursos de Tecnologías de la Información (SISP) y por la amplia jurisprudencia establecida en las últimas dos décadas por los órganos de control interno y externo. El modelo de gestión gubernamental por sistemas estructurales, en el Poder Ejecutivo Federal, se remonta al Decreto-ley n.º 200, del 25 de febrero de 1967, que estableció lo siguiente:

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*)⁶.

Originalmente, el SISP se estableció mediante el Decreto n.º 1.048, del 21 de enero de 1994, que entonces se denominaba Sistema de Administración de Recursos de Información e Informática. En su versión actual, el Decreto n.º 7.579, del 11 de octubre de 2011 (que también se refiere al Decreto-ley 200/1967), definió la forma en que se organiza el SISP y, en su artículo 2, estableció el propósito de la gestión sistémica, con las funciones sustantivas básicas que son de mayor interés para la evaluación de los PDTIC realizada en la investigación en cuestión (*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*)⁷.

ANÁLISIS Y DISCUSIÓN DE CONTENIDOS DE LOS PDTIC

Con base en la planilla completada con la información recopilada de los documentos y los parámetros estadísticos textuales calculados, la etapa de análisis de contenidos evidenció los siguientes aspectos de los PDTIC.

Desempeño en la ejecución

Los datos de desempeño en la ejecución de los planes anteriores, uno de los requisitos de información para el diagnóstico previsto en la guía del SISP, siguen siendo un desafío para la transparencia en la gestión de las TIC de los órganos, ya que solo 11 de los 17 órganos (64.7 %) de la Administración Directa y 26 de los 42 órganos (61.9 %) de la Administración Indirecta del Poder Ejecutivo Federal presentaron estas evaluaciones en sus PDTIC. Aunque varios órganos mencionan la existencia de documentos anexos con este elemento de evaluación, a menudo no se encontraron en los portales digitales de los órganos.

6 Traducción: “Art. 30 Las actividades de personal, presupuesto, estadística, administración financiera, contabilidad, auditoría y servicios generales, así como otras actividades auxiliares comunes a todos los órganos de la Administración que, a criterio del Poder Ejecutivo, requieran coordinación central, se organizarán en forma de sistema. § 1º Los servicios encargados del ejercicio de las actividades mencionadas en este artículo se consideran integrados en el respectivo sistema y, en consecuencia, están sujetos a la orientación normativa, supervisión técnica y fiscalización específica del órgano central del sistema, sin perjuicio de su subordinación al órgano en cuya estructura administrativa estén integrados.

(...) Art. 31. La estructuración de los sistemas a que se refiere el artículo 30 y la subordinación de los respectivos Órganos Centrales se establecerán mediante decreto” (Brasil, 1967, online, traducción editorial).

7 Traducción: “IV – estimular el uso racional de los recursos de tecnología de la información en el ámbito del Poder Ejecutivo federal con el fin de mejorar la calidad y la productividad del ciclo de la información; V – Estimular el desarrollo, la estandarización, la integración, la interoperabilidad, la normalización de los servicios de producción y difusión de la información, de manera desconcentrada y descentralizada” (Brasil, 2011, online, traducción editorial).

De acuerdo con la **TABLA 1**, el desempeño promedio de los órganos del Poder Ejecutivo Federal en la ejecución de los PDTIC es solo del 44.3 %, con un coeficiente de variación de 0.4, lo que muestra que más de la mitad de los contenidos de los planes no salen del papel o se ejecutan solo parcialmente (el promedio del Poder Judicial no debe considerarse porque representa los datos de solo tres órganos no aleatorios). Se espera que investigaciones posteriores puedan mostrar las causas de este nivel de desempeño: si las expectativas de entregas de la unidad de TIC de los órganos son excesivas, dados sus recursos (lo que requiere un mayor sentido de realidad en el proceso de priorización de las necesidades); o si la evaluación del tiempo para la implementación, debido al escaso detalle de las necesidades de TIC en la etapa de planificación, ha sido demasiado optimista; o una mezcla de estas dos causas.

TABLA 1 – Cálculo estimado de la capacidad de entrega (hipotético)

Recursos Humanos Disponibles				Consumo de Tiempo de los Perfiles por Actividades (A _i) y Proyectos (P _i) (horas-hombre)					
Perfis	Tiempo de Trabajo (horas-hombre)			A ₁	A ₂	A ₃	A ₄	P ₁	P ₂
	Día	Año	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
Tiempo del Equipo (h-h)	106	26.500	79.500	7.500	22.500	7.500	4.750	21.250	16.000

Fuente: Elaboración propia (2022)

Capacidad de entrega como criterio de priorización

Mediante el análisis del discurso expresado por los autores de los textos, se identificó que en 46 de los 72 (63.9 %) PDTIC investigados se mencionan lagunas de recursos humanos en las unidades de gestión de las TIC, 45 de ellas en los órganos del Poder Ejecutivo Federal (en 13 órganos de la Administración Directa y 32 de la Administración

Indirecta). Es decir, en más del 76.0 % de los órganos del Poder Ejecutivo Federal, tanto en la Administración Directa como en la Indirecta, se mencionan lagunas de recursos humanos para la gestión de las TIC.

En particular, se destaca la iniciativa de algunas unidades de gestión de las TIC de utilizar un método más detallado para evaluar su capacidad de entrega a lo largo del horizonte temporal del plan, a partir de una evaluación previa de sus recursos humanos. Aunque la Guía del SISP prevé esta evaluación, no propone ninguna metodología específica “de principio a fin” del proceso, lo que resulta en que las evaluaciones encontradas en los PDTIC se centran, en general, solo en los recursos humanos disponibles, sin estimaciones de los recursos humanos necesarios para la implementación de los planes.

Con este abordaje algo innovador (y valiente, cabe señalar), se adopta un mayor sentido de realidad para priorizar la satisfacción de las necesidades de quienes demandan las soluciones. El constructo (estructura de razonamiento) presentado en la **TABLA 1**, con datos hipotéticos de un órgano cualquiera, respalda el algoritmo de cálculo de la capacidad de entrega de la unidad de TIC, con datos de su fuerza laboral (en horas-hombre) y del consumo de esta fuerza laboral para la implementación del PDTIC en el órgano. Este cálculo debe contrastar una suma de las horas de trabajo (horas-hombre) netas diarias y anuales disponibles de cada miembro del equipo (excluidas las horas comprometidas con otras actividades de la unidad) con una distribución de esas horas en las diversas acciones previstas para la ejecución del PDTIC en su horizonte temporal.

Con períodos de implementación (horizontes) de tres años, por ejemplo, se observa que, en la cuarta columna de izquierda a derecha, la suma de las horas-hombre de la columna “PDTIC” es igual a tres veces la suma anual de las horas netas disponibles de cada perfil profesional involucrado en la ejecución del PDTIC ($79\ 500 / 26\ 500 = 3$).

Con este algoritmo de cálculo se realiza la planificación detallada del consumo de tiempo de cada perfil profesional que conforma el equipo de gestión de las TIC en el órgano. Obliga al equipo de TIC a estimar previamente el consumo de tiempo total para cada actividad o proyecto previsto en el PDTIC, en el horizonte de planificación, y a distribuir este tiempo con los respectivos tiempos disponibles de los profesionales que participarán en la ejecución. Los PDTIC más desarrollados en este aspecto, en una evaluación preliminar, son los de los siguientes órganos: Ministerio Público Federal (MPF), Consejo Nacional de Justicia (CNJ), Agencia Espacial Brasileña (AEB) y Superintendencia de Seguros Privados (SUSEP).

Esta innovación en la gestión hace que los PDTIC sean más realistas y se basen en evidencia, como se recomienda en la Resolución n.º 2608/2018-TCU-Plenario (Brasil, 2018) en relación con las políticas públicas, planes y programas. Si bien este modelo de evaluación de la capacidad de entrega del equipo de TIC también presenta riesgos de estimaciones erróneas, con seguridad estará más cerca de la capacidad de entrega real de la unidad, como uno de los criterios de priorización aplicables.

Capítulos técnicos de interés

El proyecto de investigación también incluyó la verificación textual de la eventual presencia de capítulos técnicos específicos de interés que evidenciaran el nivel técnico de los PDTIC, ya sean previstos o no en la Guía del SISP. Los temas técnicos de investigación se presentaron en la **TABLA 2**: “Estrategias”, “Arquitecturas”, “MGP” (Metodología de Gestión de Proyectos de TIC), “Gestión de Riesgos” y “Gobernanza”.

Utilizando las funciones de búsqueda de palabras de un *software* procesador de texto, se identificó la presencia o ausencia de capítulos con ciertos contenidos de interés para la investigación, con los títulos indicados en la **TABLA 2**. También se realizó, como precaución, la búsqueda de estos capítulos en el cuerpo de los textos de los PDTIC, considerando que a veces los resúmenes de los planes no detallan los contenidos de los capítulos.

TABLA 2 – Capítulos técnicos específicos

Poderes y Entidades del Estado	Cantidad de PDTIC	Cantidad de PDTIC con Capítulos Técnicos				
		Estrategias	Arquitecturas	MGP (*)	Gestión de Riesgos	Gobernanza
Poder Ejecutivo	59	4	15	2	51	29
Administración Directa	17	2	2	1	15	7
Administración Indirecta	42	2	13	1	36	22
Poder Legislativo	2	1	1	0	0	1
Poder Judicial	10	0	1	0	0	5
Ministerio Público Federal	1	0	0	0	0	1
Total	72	5	17	2	51	36

Fuente: Elaboración propia (2022)

(*) MGP: Metodología de Gestión de Proyectos de TIC

El tema “estrategias” se refiere a demostraciones explícitas y detalladas de la articulación entre las estrategias de negocio y de TIC en el propio PDTIC, con el fin de evidenciar técnicamente la integración y sinergia entre estas estrategias (no es, por lo tanto, solo una mención de esta articulación). Con los resultados presentados en la **TABLA 2**, se concluye que solo 5 de los 72 órganos (6.9 %) muestran, al menos mínimamente, esta sinergia. Varios órganos se destacan por remitir esta visión de articulación entre estrategias a otros documentos del órgano, que no forman parte del PDTIC, como los documentos de planificación estratégica.

Los datos de la columna “arquitecturas”, en la **TABLA 2**, consolidan las tres arquitecturas objeto de investigación en los PDTIC: arquitectura corporativa o empresarial, arquitectura de la información y arquitectura técnica o de sistemas. Se observa que solo 17 de los 72 órganos (23.6 %) elaboran este tema en capítulos de sus PDTIC, con una tasa mejor en el conjunto de órganos de la Administración Indirecta del Poder Ejecutivo: 13 de 42, o 31.0 %.

El resultado quizás más frustrante tiene relación con la presencia de capítulos temáticos sobre metodologías de desarrollo de proyectos de TIC (columna “MGP”): solo 2 de los 72 órganos investigados elaboraron contenido sobre el tema en un capítulo específico en el propio PDTIC, aunque varios mencionan la existencia de una metodología o la intención de adoptar una metodología. Con este resultado, una cuestión metodológica importante no se vuelve transparente para el escrutinio técnico externo al órgano, ni para el intercambio de experiencias y aprendizaje entre las unidades de gestión de las TIC y los expertos externos.

Sin embargo, se considera como resultados positivos que la mayor parte de los PDTIC presentan capítulos temáticos sobre dos temas bastante actuales (en los términos del Decreto n.º 9.203, del 22 de noviembre de 2017): “gestión de riesgos” y “gobernanza”. Los órganos del Poder Ejecutivo Federal presentan, en sus PDTIC, un capítulo sobre “gestión de riesgos” en 51 de los casos (70.8%), con resultados mejores para los órganos de la Administración Directa, en 15 de los 17 órganos (88.2%). Estas tasas disminuyen en el tema “gobernanza”, aunque son significativas y muy cercanas en todas las categorías de órganos; en los órganos del Poder Ejecutivo Federal, 29 de 51 (o 56.9%) tienen capítulos sobre “gobernanza” en sus PDTIC.

Citas de temas actuales

Otro bloque de objetos semánticos de investigación se centró en la búsqueda de citas textuales, no necesariamente en capítulos específicos, sobre los siguientes temas: “estrategia”; “seguridad de la información”; “Geographical Information System” (*GIS*); “Sistema de Información Geográfica” (*SIG*); “georreferenciación”; “imagen satelital”; “*Drone*”; “Vehículo Aéreo No Tripulado” (*VANT*); “aplicación para dispositivo móvil”, “*mobile*” (aplicación para dispositivo móvil); “*app*” (de *software application*); “gestión de la información”; “gestión del conocimiento”; “computación en la nube”; “*cloud computing*” y “*big data*”. La **TABLA 3** presenta los resultados sumados de esta búsqueda en cada PDTIC, y se destaca el hecho de que los temas relacionados con la georreferenciación se consolidaron en un solo ítem para el análisis de datos (sumando también las frecuencias de las citas de “GIS”, “SIG”, “imagen satelital”, “*drone*” y “VANT”), así como los resultados de búsqueda de “aplicación”, “*mobile*” y “*app*” (consolidados en “Aplicación para Dispositivo Móvil”), sumando las citas de “gestión de la información” con las de “gestión del conocimiento” por las conexiones naturales entre sí.

TABLA 3 – Citas de temas actuales

Temas de Investigación		Poderes y Entidades (*) y Frecuencias Observadas de Citas Temáticas						
#	Expresión Temática	PE	PE-AD	PE-AI	PL	PJ	MPF	Subtotal
1	Estrategia	1.629	514	1.115	13	137	36	1.815
2	Seguridad de la Información	1.240	371	869	5	77	15	1.337
3	GIS/SIG, Georreferenciación, Imágenes Satelitales, Drone/VANT (**)	628	98	530	3	23	0	654
4	Aplicación para Dispositivo Móvil	234	162	72	6	18	3	261
5	Gestión de la Información y del Conocimiento	209	55	154	5	2	2	218
6	Computación en la Nube (Cloud Computing)	180	49	131	9	27	11	227
7	Big Data	16	7	9	0	4	2	22
Cantidad de PDTIC (A)		59	17	42	2	10	1	72
Total de Citas Encontradas (B)		4.136	1.256	2.880	41	288	69	4.534
Total de Citas / PDTIC (C=B/A)		70,1	73,9	68,6	20,5	28,8	69,0	63,0

Fuente: Elaboración propia (2022)

(*) PE: Poder Ejecutivo; PE-AD: Administración Directa del Poder Ejecutivo; PE-AI: Administración Indirecta del Poder Ejecutivo; PL: Poder Legislativo; PJ: Poder Judicial; MPF: Ministerio Público Federal.

(**) *Drone* (en inglés) o *VANT* (en español): Vehículo Aéreo No Tripulado.

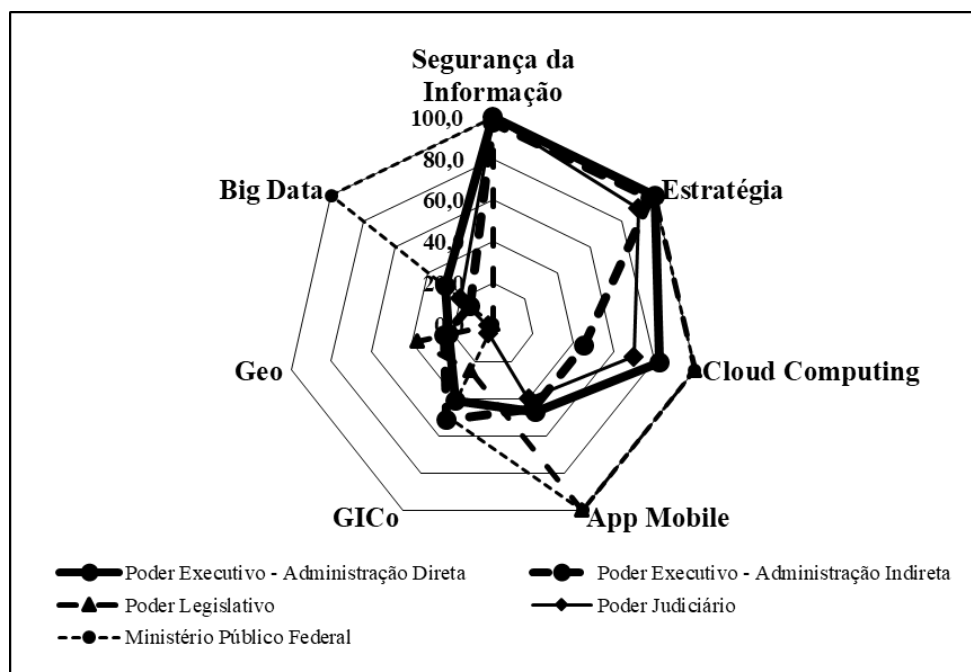
Los aspectos más relevantes de las búsquedas, a partir de los resultados en la **TABLA 3**, son:

- Los dos temas más citados son, por lejos, “estrategia” (con 1815 citas encontradas) y “seguridad de la información” (con 1337 citas).
- En cuanto a los temas relacionados con “georreferenciación” y “aplicación para dispositivo móvil”, el primero parece interesar más a los órganos de la Administración Indirecta que a los de la Administración Directa del Poder Ejecutivo Federal, con proporciones de frecuencia del orden de 12.6 citas por órgano (530/42) y 5.8 citas por órgano (98/17), respectivamente, mientras que el segundo tema parece interesar mucho a los órganos de la Administración Directa, en la proporción de 9.5 citas por órgano (162/17), frente a 1.7 citas por órgano (72/42) en la Administración Indirecta;
- aunque ampliamente comentado, en los últimos años, el tema “computación en la nube” (cloud computing) no aparece muy a menudo en los planes de los órganos en general y, en el Poder Ejecutivo Federal, en promedio, solo se encontraron 3.1 citas por PDTIC (180/59); aún menos el tema “big data”, con solo 16 citas en total (16/59, o solo una cita por casi cuatro PDTIC).

Considerando la importancia de la georreferenciación digital en su contexto, el Departamento Nacional de Producción Mineral (DNPM) denomina a su plan el “Plan Director de Tecnologías de la Información, la Comunicación y el Geoprocesamiento 2017-2019”.

En términos de la cantidad de órganos en los que aparecen las citas de estos temas, en relación con la cantidad de órganos con PDTIC investigados (porcentajes de órganos con citas), el gráfico de estilo “radar” de la **FIGURA 1** muestra las concentraciones, con porcentajes del 0 al 100, de PDTIC con las citas temáticas de investigación por Poder y Entidad del Estado. Se observa una mayor concentración de citas en los temas “estrategia” y “seguridad de la información” en los órganos de la Administración Directa del Poder Ejecutivo (línea continua más gruesa) y menor en los órganos de la Administración Indirecta (línea punteada más gruesa).

FIGURA 1 – Porcentajes de PDTIC con citas de temas actuales



Fuente: Elaboración propia (2022)

Análisis de las necesidades de TIC

Este bloque de datos de investigación se refiere a las demandas de soluciones de TIC expresadas en los PDTIC, que corresponden al ítem de diagnóstico que en la Guía del SISP se denomina “necesidades de TIC”. En resumen, en esta lógica cognitiva, las necesidades representan el indicador más importante según el cual la unidad de gestión de TIC debe dirigir sus recursos y, así, contribuir a la evolución tecnológica del órgano.

La **TABLA 4** muestra las cantidades de ítems que aparecen como “necesidades de TIC” en los planes de los organismos investigados, agrupados por categoría político-

administrativa. Como se incluye en la Guía del SISP, algunos órganos expresan los ítems de adquisición o desarrollo de bienes y servicios como “necesidades de TIC”, sin un desglose más detallado, mientras que otros detallan sus necesidades en proyectos y acciones y, en ocasiones, en metas.

TABLA 4 – Necesidades de TIC

Poderes y Entidades del Estado	Cantidad de PDTI (A)	Cantidades de ítems como “Necesidades”	
		Subtotal (B)	Promedio (C=B/A)
Poder Ejecutivo	59	3.982	67,5
Administración Directa	17	832	48,9
Administración Indirecta	42	3.150	75,0
Poder Legislativo (*)	2	133	66,5
Poder Judicial (*)	10	863	86,3
Ministerio Público Federal (**)	1	-	-
Total	72	4.978	69,1

Fuente: Elaboración propia (2022)

(*) Calculados como necesidades o proyectos.

(**) Las demandas no se consolidan en el mismo formato que los demás órganos.

En general, los órganos del Poder Ejecutivo Federal expresan sus necesidades de TIC en 24 páginas de los PDTIC, en promedio (con un coeficiente de variación de 1.2), mientras que en los órganos del Poder Judicial este promedio es ligeramente inferior, alrededor de 19 páginas, con un coeficiente de variación de 0.9. El volumen documental total que registra las demandas de TIC de los PDTIC investigados totalizó 1693 páginas con 4978 “necesidades” y 6106 proyectos o acciones con algún detalle.

Es relevante observar que los órganos de los Poderes Legislativo y Judicial y el Ministerio Público Federal (MPF) no están subordinados al SISP, por lo que sus PDTIC no presentan, por lo general, la misma estructura y formatos de presentación de datos que los órganos del Poder Ejecutivo Federal. Con respecto a las demandas de TIC observadas, varios órganos de los Poderes Legislativo y Judicial y el Ministerio Público Federal presentaron sus demandas directamente como “proyectos” y no como “necesidades”.

Con el análisis del discurso de los textos de los capítulos sobre las necesidades de TIC, se observó que las descripciones de los ítems suelen estar en niveles de abstracción todavía muy altos, con poca claridad, desde un punto de vista técnico, sobre lo que se demanda, o las necesidades se describen con objetos tan amplios que sin duda deberían haberse desglosado en una serie de demandas (necesidades o proyectos/acciones) más pequeñas y específicas. Como ejemplos de la poca claridad de las necesidades inventariadas,

se encontraron elementos descritos (literalmente) como: “levantamiento de necesidades e implementación de recursos para mejorar la disponibilidad de infraestructura de TIC”, “mejorar el proceso de toma de decisiones estratégicas”, “implementar tecnologías de gestión del conocimiento en apoyo de cuestiones estratégicas”, “desarrollar sistemas de información basados en computadoras con el uso de plataformas estandarizadas, utilizando procesos definidos y maduros”, “aumentar la productividad y la capacidad de entrega”, “desarrollo de aplicaciones móviles”, “desarrollar nuevos sistemas”, “adquisición de productos/sistemas para satisfacer las necesidades específicas de las áreas comerciales”, “centro de monitoreo integrado”, “solución para la gestión de recursos humanos”, “implementar/adquirir/personalizar sistemas informáticos de acuerdo con las necesidades declaradas después de los estudios de viabilidad”, “mejorar la calidad de los servicios de TI prestados por el órgano a sus usuarios”, “mejora de las redes de los edificios”, “aumentar el uso de soluciones de mercado”.

Otro aspecto preocupante es que en la abrumadora mayoría de los órganos investigados del Poder Ejecutivo Federal se encuentran necesidades aparentemente redundantes en el mismo PDTIC, que podrían consolidarse mediante un abordaje más corporativo e integrado del servicio, quizás como resultado de la falta de un análisis técnico más profundo sobre las necesidades de TIC presentadas por las diversas unidades internas de los órganos.

Los datos de la **TABLA 5** muestran el listado de las 10 palabras nominativas más frecuentes (sustantivos que generan denominaciones de objetos, acciones o ideas) que se encuentran en los capítulos “Necesidades TIC” de los PDPIC investigados, excluidas de esta tabla las palabras (siglas) “TI” y “TIC”, que, debido a su relación directa con el tema general de los textos, se repiten miles de veces sin añadir novedad (información relevante) alguna. La tercera columna de izquierda a derecha muestra las frecuencias respectivas con las que estas palabras, de forma aislada, se encuentran en los textos, en números absolutos (recuento total), y la cuarta y última columna, de izquierda a derecha, muestra las cantidades de *cluster* de sustantivos (sintagmas nominales) formados al combinarse con cada una de las palabras más frecuentes observadas (en singular y con inflexiones en plural).

TABLA 5 – 10 Palabras más frecuentes y *cluster* formados

#	Palabra	Frecuencia de la Palabra	Frecuencia de <i>Clusters</i>
1	sistema/sistemas	3.360	2.049
2	servicio/servicios	1.856	1.824
3	información/informaciones	1.840	1.482
4	Gestión	1.580	2.118
5	dato/datos	1.331	1.011
6	proyecto/proyectos	1.330	680
7	proceso/procesos	1.316	963
8	solución/soluciones	1.241	953
9	adquisición/adquisiciones	1.001	763
10	software/softwarees	908	1.021

Fuente: Elaboración propia (2022)

Los *cluster* de palabras se identifican en los textos y se calculan, en términos de frecuencias de ocurrencia, solo con el uso de un *software* de PNL específico, como el relacionado con el trabajo de Scott (2019), con varias otras funciones de búsqueda y estadísticas textuales que no se encuentran en los procesadores de texto comunes. Las estructuras lingüísticas en *cluster* de palabras, en este caso composiciones de sustantivos con otros sustantivos, son semánticamente más ricas que los sustantivos aislados, como en el siguiente ejemplo: “gestión” no aclara mucho sobre el contexto de aplicación de la gestión, pero “gestión de la información” identifica el área de aplicación de la gestión, presenta un contexto más específico, con mayor poder explicativo; lo mismo ocurre con el sustantivo “seguridad”, que al combinarse con “información” se convierte en “seguridad de la información”.

Sin embargo, como medida de ahorro de recursos computacionales de investigación (tiempo de procesamiento de textos), solo se sumaron los *cluster* con frecuencias iguales o superiores a 10, de lo que se infiere que las frecuencias de sintagmas nominales formados con las 10 palabras más frecuentes (*top ten*) serán aún más numerosas. Los *cluster* más frecuentes están compuestos por la palabra “gestión”, con un total de 2118 apariciones, lo que suele presuponer la gestión de algo, por lo que requiere de un complemento nominal, y se presenta en textos técnicos como “gestión de ...”; de lo contrario, la composición del *cluster* también puede aparecer como “proceso(s) de gestión”, donde “gestión” es el complemento nominal de otro sustantivo (por cierto, esta composición de *cluster* apareció 173 veces en el conjunto de PDTIC investigados).

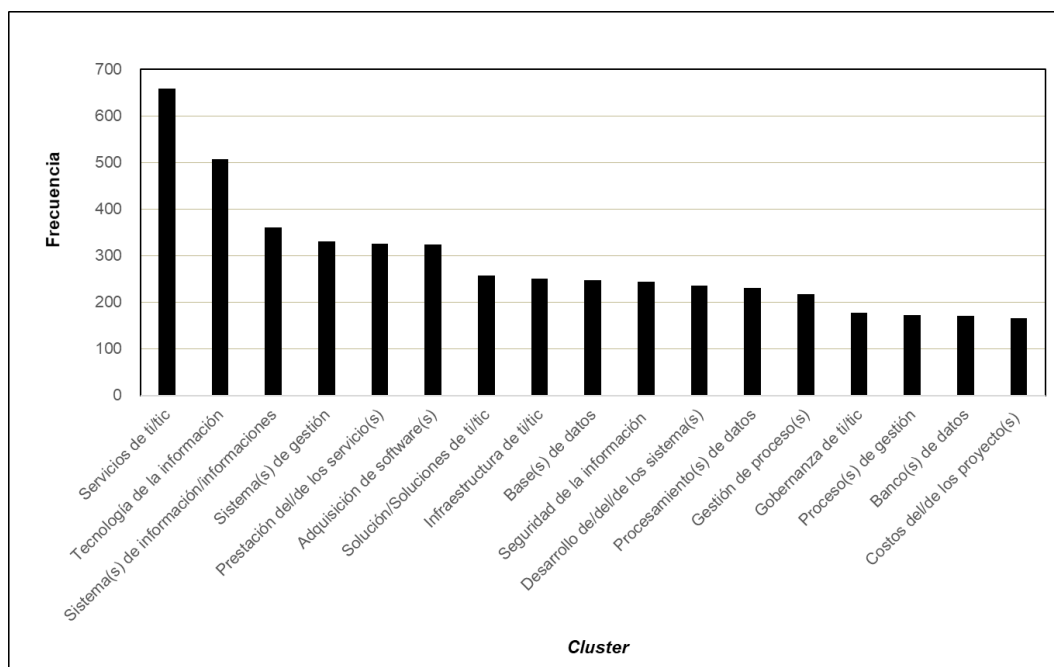
La segunda estructura de *cluster* más frecuentes está compuesta por la palabra “sistema(s)”, con 2049 apariciones, y la tercera con “servicio(s)”, con un total de 1824 apariciones. Esta cantidad de *cluster* observados (con una frecuencia igual o superior a 10)

relacionados con el concepto de “servicio(s)” denota una preocupación de los órganos con la prestación de servicios basados en las TIC, lo que es una señal positiva de los objetivos mayores de los PDTIC, aunque esta evolución no se produzca a la velocidad deseada.

Los resultados de la búsqueda de *cluster* más frecuentes en los PDTIC investigados se presentan en lo **GRÁFICO 1**. La expresión nominativa “servicios de TIC” aparece como el *cluster* más frecuente, con 658 apariciones, seguida por “tecnología de la información”, con 506 apariciones, y “sistema(s) de información”, con 361 ocurrencias. Se observa que de las 10 palabras más frecuentes presentadas en la **TABLA 5**, solo dos no forman parte de los *cluster* entre los 10 más frecuentes: proceso(s) y proyecto(s).

Considerando que el término “solución”, de la Instrucción Normativa n.º 4/2010-SLTI/MPOG, se convirtió en una expresión estándar para referirse a un conjunto de recursos de TIC combinados y destinados a la solución de un problema empresarial, se buscó en los PDTIC, en los capítulos sobre “necesidades de TIC”, los sintagmas nominales (expresiones nominativas compuestas) estructurados como “solución/soluciones de ...” y “solución/soluciones para ...”, para luego observar, en los sintagmas que regresaron, el término nominativo complementario en la secuencia que califica la palabra “solución/soluciones”. Con esta operación de búsqueda textual, se obtuvieron los resultados que se presentan en lo **GRÁFICO 2**, donde “solución/soluciones de ti/tic”, “solución/soluciones de gestión” y “solución/soluciones de seguridad” aparecen como los tres *cluster* más frecuentes.

GRÁFICO 1 – *Cluster* más frecuentes entre “necesidades de TIC”



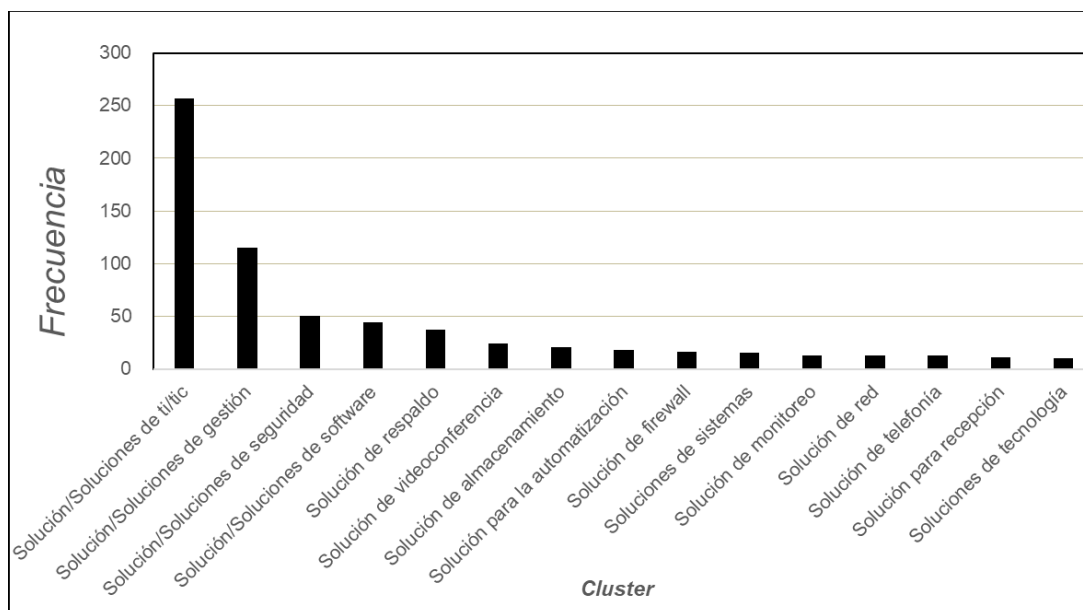
Fuente: elaboración propia (2022)

Considerando que el término “solución”, de la Instrucción Normativa n.º 4/2010-SLTI/MPOG, se convirtió en una expresión estándar para referirse a un conjunto de recursos de

TIC combinados y destinados a la solución de un problema empresarial, se buscó en los PDTIC, en los capítulos sobre “necesidades de TIC”, los sintagmas nominales (expresiones nominativas compuestas) estructurados como “solución/soluciones de ...” y “solución/soluciones para ...”, para luego observar, en los sintagmas que regresaron, el término nominativo complementario en la secuencia que califica la palabra “solución/soluciones”. Con esta operación de búsqueda textual, se obtuvieron los resultados que se presentan en lo **GRÁFICO 2**, donde “solución/soluciones de ti/tic”, “solución/soluciones de gestión” y “solución/soluciones de seguridad” aparecen como los tres *cluster* más frecuentes.

Estos resultados corroboran los resultados presentados en la **FIGURA 1** sobre la alta frecuencia de las necesidades de soluciones para la “seguridad de la información” expresadas en los PDTIC investigados, y también es interesante observar la frecuencia significativa de una necesidad de solución poco comentada: “solución de videoconferencia”. Con respecto a las otras necesidades mencionadas, como “solución/soluciones de respaldo”, “solución de red”, “solución de almacenamiento” y “solución de telefonía”, por ejemplo, estas no representan una novedad en los PDTIC porque constituyen necesidades de infraestructura comunes en todos los órganos.

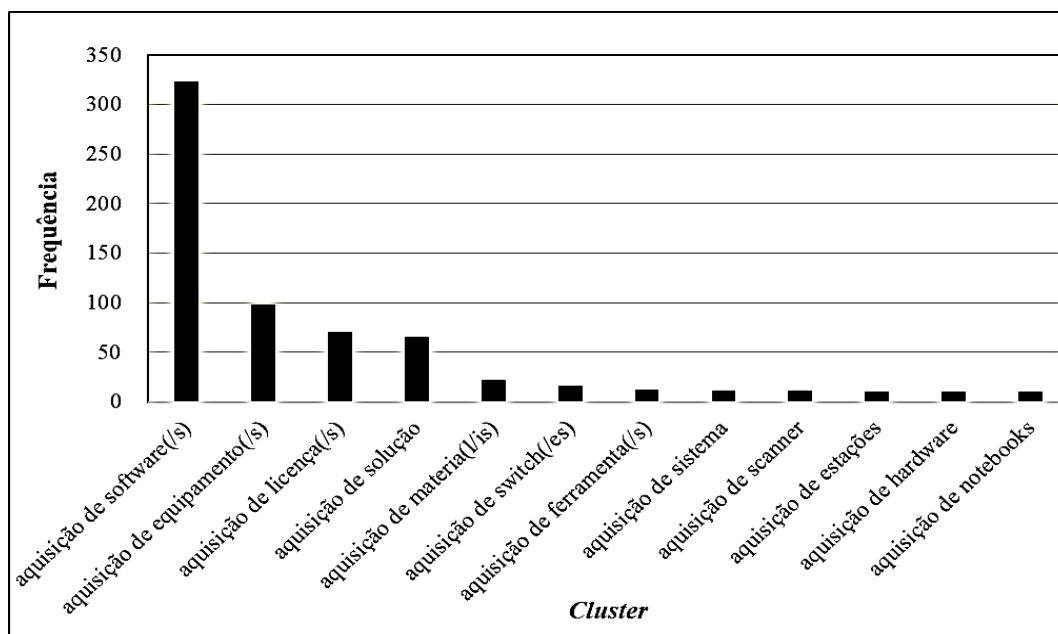
GRÁFICO 2 – *Cluster* más frecuentes con “solución/soluciones”



Fuente: elaboración propia (2022)

En lo **GRÁFICO 3**, nuevamente se evidencia el poder de las tecnologías de PLN al presentarse, entre los 10 *cluster* más frecuentes (con frecuencias iguales o superiores a 10) observados entre las “Necesidades de TIC”, una serie de *cluster* interesantes compuestos por la palabra “adquisición”. Estas expresiones revelan las intenciones de compra de los componentes de las TIC.

GRÁFICO 3 – Cluster más frecuentes con “adquisición”



Fuente: elaboración propia (2022)

Otro aspecto interesante de este análisis sintagmático se refiere a los *cluster* compuestos con la palabra “gestión”: sumando las frecuencias de los *cluster* con las estructuras “sistema(s) de gestión” y “soluciones de gestión”, hay más de 400 *cluster* que contienen esta idea: la de soluciones o sistemas de información computacional para la gestión de los órganos. Se concluye, ante esta evidencia textual presente en los PDTIC investigados, que al menos existe una genuina preocupación de los órganos públicos del Estado por la automatización de sus procesos de gestión.

En cuanto a los presupuestos de TIC, los planes investigados revelaron que los mayores presupuestos también son de los órganos más grandes en todos los sentidos, con unidades regionales descentralizadas en el país y mayor cantidad de funcionarios públicos, como el Ministerio de Salud, el Ministerio de Hacienda, el Departamento Nacional de Infraestructura de Transportes, el Ministerio de Justicia y Seguridad Pública y el Ministerio de Agricultura, Ganadería y Abastecimiento. También se observó que el tema presupuestario no se aborda en todos los órganos del Poder Ejecutivo del Estado, con solo 49 de los 59 órganos investigados (83.1 %) que estimaron sus gastos.

Con respecto a los valores, el monto presupuestario promedio de cada PDTIC investigado es de BRL 115.8 millones en el horizonte temporal promedio de planificación (que es de 2.8 años), lo que resulta en un promedio anual de gastos por órgano de aproximadamente BRL 41.3 millones.

CONCLUSIÓN

Se concluye, con base en el tamaño muestral y los resultados de la investigación presentada en este artículo, en primer lugar, que las tecnologías de Procesamiento del Lenguaje Natural (PLN) pueden contribuir efectivamente a los procesos de evaluación semiautomática del contenido textual de los documentos en el sector público, como en el caso de los planes; y en segundo lugar, que los PDTIC publicados entre 2013 y 2019 parecen haberse elaborado de acuerdo con la Guía del SISP y, por lo tanto, pueden contribuir a la evolución tecnológica de los órganos, pero aún carecen de mejoras técnicas significativas para convertirse en instrumentos más efectivos en este sentido.

Los resultados de la investigación muestran una evolución positiva en cuestiones actuales de gestión de las TIC, con menciones frecuentes a temas de gobernanza y gestión de riesgos, y un interés significativo en la adopción de tecnologías disruptivas por parte de algunos órganos del Poder Ejecutivo Federal, como las tecnologías de georreferenciación digital y las aplicaciones para dispositivos móviles, aunque tal vez de manera demasiado cautelosa ante la relación costo-beneficio de estas tecnologías. Otro aspecto positivo es que algunos modelos de gestión de TIC más controvertidos, aunque adoptados por empresas del mercado, parecen estar considerándose en el alcance del SISP debido a su complejidad en una posible implementación a gran escala en el sector público federal, como es el caso de la computación en la nube (*cloud computing*).

Otro aspecto relevante, descubierto por casualidad en la investigación, es que algunas unidades de gestión de las TIC se atreven a innovar en el modelo de gestión de demandas, oponiéndose técnicamente (con datos) a las necesidades de las TIC y a las estimaciones de la capacidad de entrega real de sus equipos. Así, la gestión de las TIC en los órganos podría convertirse en una actividad más basada en datos y evidencia y, por lo tanto, más realista, en línea con las recomendaciones del Tribunal Federal de Cuentas en la Resolución 1603/2008-TCU-Plenário (Brasil, 2008), sobre las necesidades de recursos humanos en las unidades de TIC de los órganos, y en la Resolución 2608/2018-TCU-Plenário (Brasil, 2018), sobre criterios técnicos para la elaboración de políticas públicas, planes y programas.

Los datos estadísticos sobre el contenido textual también revelan, por otro lado, de acuerdo con la percepción de una autoridad gubernamental recientemente publicada, que en general los órganos del Poder Ejecutivo Federal aún carecen de una evolución en las TIC “de la puerta hacia adentro” (Batista; Azevedo; Hessel, 2020), para resolver sus cuellos de botella de infraestructura tecnológica e integración de sistemas informáticos de información que impiden al ciudadano obtener un servicios pleno. El nivel de detalle de las demandas incluidas en los PDTIC, generalmente superficial (además de redundantes), tampoco contribuye a la evaluación de los esfuerzos de servicio y una mejor gestión de las necesidades de TIC en los órganos.

El poco interés mostrado en la elaboración sobre arquitectura de la información y arquitectura técnica puede contribuir a este lento avance de los órganos del SISP en el

cumplimiento de los requisitos técnicos de integración, interoperabilidad y portabilidad de los sistemas previstos en la Ordenanza n.º 92/SLTI/MPOG, del 24 de diciembre de 2014 (que instituyó la arquitectura de interoperabilidad “ePING”), así como a los objetivos institucionales propuestos en el Decreto n.º 10.046, del 9 de octubre de 2019 (sobre la gobernanza en los procesos de compartimiento de datos). Considerando una analogía natural entre la arquitectura de las TIC y la arquitectura tradicional (de la construcción civil), desarrollar un entorno computacional sin arquitectura de la información y arquitectura técnica equivale a construir un edificio sin diseño arquitectónico, asumiendo los riesgos de inadecuación de los espacios y el uso inapropiado de costosos recursos de TIC.

Obviamente, en estas cuestiones no se pueden ignorar aspectos adversos que trascienden la gobernabilidad de las unidades de TIC en los órganos, como problemas de estrategias y procesos de negocios corporativos inadecuados y modelos de gobernanza inadecuados o inexistentes. El ejemplo más evidente se refiere al nivel jerárquico en el que se encuentran los titulares de las unidades de gestión de las TIC en los órganos, en su abrumadora mayoría ubicados en posiciones intermedias, lo que demuestra, en la práctica, el nivel de prioridad de la gestión de las TIC.

REFERENCIAS

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.

CONGLIAN, 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.