



## A privacidade e os dados pessoais na Ciência da Informação: uma análise de domínio

*Privacy and personal data in Information Science: a domain analysis*

Gislaine Cristina da Conceição <sup>a,\*</sup> 

Januário Albino Nhacuongue <sup>a</sup> 

**RESUMO:** Esta pesquisa tem por objetivo analisar as discussões acerca de aspectos legais no uso de dados na Ciência da Informação, no Brasil, nos últimos 10 anos. Desse modo, procura responder à questão norteadora: qual vem sendo a atenção dedicada aos aspectos legais e de privacidade no uso de dados pessoais nas pesquisas científicas dos últimos dez anos da Ciência da Informação, no Brasil? Quanto à metodologia, foi feita uma análise de domínio dos materiais bibliográficos referentes a dados e privacidade na Ciência da Informação, no Brasil, nos últimos 10 anos. O foco da análise de domínio foi concentrado na identificação dos autores e suas filiações institucionais; área de aplicação do estudo de acordo com o ciclo de vida de dados; além da análise histórica de tendências das discussões dos dez autores mais colaborativos com o ciclo de vida de dados na Ciência da Informação. Após a análise de 173 artigos escritos por 243 autores, a pesquisa mostra que 75% dos estudos tratam sobre a Fase da Coleta de Dados, seguidos das fases de armazenamento com 14%, recuperação com 8,7% e descarte com 2,8%. Constatou-se que as instituições UNESP, UNB, UFRJ, UFPE, FGV e UFMG, com 39,5%, lideram o ranking das publicações e afiliações de autores que abordam a maioria dos aspectos da fase Coleta de Dados. Conclui-se que, embora muitos estudos sobre aspectos legais e de privacidade no uso de dados pessoais nas pesquisas científicas dos últimos dez anos na Ciência da Informação se preocupem com todo o ciclo de vida de dados, a fase de Descarte ainda é pouco explorada.

**Palavras-chave:** Dados Pessoais; Privacidade; Análise de Domínio; Ciência da Informação.

**ABSTRACT:** This research aims to analyze the discussions about legal aspects in the use of data in Information Science in Brazil over the last 10 years. In this way, it seeks to answer the guiding question: what attention has been paid to legal and privacy aspects in the use of personal data in scientific research in the last ten years in Information Science in Brazil? In terms of methodology, a domain analysis was carried out of bibliographic materials referring to data and privacy in Information Science in Brazil over the last 10 years. The focus of the domain analysis was on identifying the authors and their institutional affiliations; the area of application of the study according to the data life cycle; as well as the historical analysis of trends in the discussions of the ten most collaborative authors on the data life cycle in Information Science. After analyzing 173 articles written by 243 authors, the research shows that 75% of the studies deal with the Data Collection Phase, followed by the storage phases with 14%, retrieval with 8.7% and disposal with 2.8%. It was found that the institutions UNESP, UNB, UFRJ, UFPE, FGV and UFMG, with 39.5%, lead the ranking of publications and affiliations of authors who address most aspects of the Data Collection phase. It can be concluded that, although many studies on legal and privacy aspects in the use of personal data in scientific research over the last ten years in Information Science are concerned with the entire data life cycle, the Disposal phase is still little explored.


**Keywords:** Personal Data; Privacy; Domain Analysis; Information Science.

---

<sup>a</sup> Programa de Pós-Graduação em Ciência da Informação, Universidade Federal de São Carlos, São Carlos, SP, Brasil.

\* Correspondência para/Correspondence to: Gislaine Cristina da Conceição. E-mail: gislaine.conceicao88@gmail.com.

Recebido em/Received: 21/07/2023; Aprovado em/Approved: 20/11/2023.

Artigo publicado em acesso aberto sob licença [CC BY 4.0 Internacional](https://creativecommons.org/licenses/by/4.0/) 

## INTRODUÇÃO

Com o avanço das redes de computadores e o crescente uso da Internet, surgiram novas possibilidades de comunicação e armazenamento de informações, permitindo um acesso mais fácil e rápido à dados. Nesse novo cenário, em que computadores e dispositivos tecnológicos desempenham um papel cada vez mais importante na expansão das habilidades de comunicação humana, a preocupação com a privacidade também surge, devido ao potencial uso indevido de informações e dados pessoais em diversos contextos (Grisoto, Sant’ana, Santarém Segundo 2015).

De acordo com a perspectiva de Doneda (2021), a informatização operou mudanças de ordem quantitativa e qualitativa no tratamento de dados. Houve um aumento significativo na quantidade de dados processados em períodos reduzidos, e resultados mais relevantes foram obtidos por meio do uso de novos métodos, algoritmos e técnicas de processamento. No entanto, o crescimento desordenado e em larga escala na produção de dados e informações despertou preocupações em relação à privacidade e ao uso desses dados.

Nessa perspectiva, o termo privacidade passa por uma transformação, dando origem à disciplina de proteção de dados pessoais. Isso ocorre devido ao crescente valor atribuído à informação, ao aumento de situações que violam a privacidade e ao número crescente de pessoas sujeitas a essas situações. Conforme destacado por Rodotà (2008, p. 15), o direito à privacidade é compreendido como “o direito de manter o controle sobre suas próprias informações e de determinar a maneira de construir sua própria esfera particular”.

Há uma demanda crescente pelo acesso e análise de dados privados e públicos, e isso tem despertado o interesse de diversos setores da sociedade. Para Kemper e Kolkman (2018), o recente aumento no uso de *big data* e a crescente complexidade e importância dos algoritmos na tomada de decisões organizacionais têm sido motivo de preocupação pela falta de transparência. Essa falta de transparência compromete a responsabilização pelas decisões tomadas pelos algoritmos.

Este trabalho aborda a problemática da privacidade, centrando-se na questão: qual tem sido a atenção dedicada aos aspectos legais e de privacidade no uso de dados pessoais em pesquisas científicas no campo da Ciência da Informação no Brasil nos últimos dez anos? O objetivo é analisar as discussões sobre aspectos legais e de privacidade relacionados ao uso de dados na Ciência da Informação, no contexto brasileiro, ao longo da última década.

## PROBLEMÁTICA E JUSTIFICATIVA

Com o surgimento da infraestrutura digital, impulsionado recentemente pela pandemia do CoVID-19, várias transações econômicas e sociais estão sendo cada vez mais digitalizadas. Os serviços bancários on-line, as redes sociais, as redes inteligentes e as redes industriais são exemplos de atividades que atualmente dependem muito de dados digitais, que podem conter informações confidenciais ou sensíveis. Além disso,

a transformação digital cria novas oportunidades para que criminosos cibernéticos violem a integridade de dados de organizações, explorando informações críticas para fraude, espionagem e sabotagem. Esse cenário requer políticas e estratégias de segurança cibernética para garantir ambientes digitais seguros, de modo a preservar a confiança nas novas tecnologias e apoiar o crescimento da economia digital.

O conceito de privacidade, enquanto direito ao autocontrole, à autonomia e à integridade da pessoa, também envolve a posse e o uso indevido de dados por entidades legitimamente qualificadas. Segundo Garfinkel (2000), as ameaças à privacidade estão intimamente relacionadas a dois fatores principais. Por um lado, aos avanços da tecnologia<sup>1</sup>, como câmeras de monitoramento e computação ubíqua. Por outro lado, às estratégias dos governos, por exemplo, de restringir o uso de tecnologia de criptografia, para manter o poder de monitoramento e controle dos cidadãos.

Uma das principais preocupações no uso de dados é a violação da privacidade do indivíduo. Quando os detentores de dados coletam informações pessoais, como nome, endereço, número de telefone, hábitos de consumo, políticas comportamentais ou até dados biométricos, eles têm a responsabilidade de proteger essas informações e usá-las apenas para os propósitos especificados e consentidos pelo titular dos dados. (Mayer-Schonenberger; Cukier 2013).

A natureza controversa do consentimento, como afirmação dos direitos relacionados aos dados digitais, surge devido à tentativa de estabelecer a liberdade e a autonomia pessoal em um contexto marcado pela desigualdade na gestão de dados. Essa desigualdade é evidente devido à disparidade na infraestrutura e no conhecimento relacionado à ciência e interpretação de dados em larga escala na era do *Big Data*, especialmente em relação aos avanços recentes em aprendizado de máquina e inteligência artificial (Mayer-Schonenberger; Cukier 2013).

A LGPD consagra o instituto do consentimento como mecanismo de proteção e como condição essencial para a reprodução social do novo mercado de dados, que é justamente ganho por meio do processo de garantia de mais-valia descrito por Zuboff (2019, p. 97): os usuários fornecem seus dados por meio do consentimento e, como resultado, são transformados e analisados pelas empresas (que os comercializam no mercado de previsão de comportamentos).

Considerando o atual dilema entre a necessidade de segurança e a consequente perda de privacidade com o compartilhamento de dados pessoais, no momento em que grande parte da população mundial se torna cada vez mais dependente das novas tecnologias de informação e comunicação (Lott; Cianconi 2018), este trabalho busca responder à seguinte questão: Qual vem sendo a atenção dedicada aos aspectos legais e de privacidade no uso de dados pessoais nas pesquisas científicas dos últimos dez anos da Ciência da Informação, no Brasil?

---

<sup>1</sup> Ao considerar a tecnologia uma das causas da ameaça à privacidade, Garfinkel (2000) alerta que ela não pode ser analisada de modo isolado, na medida em que existe na junção entre ciência, mercado e sociedade, isto é, ela é socialmente construída para preencher necessidades específicas.

O cenário apresentado permite observar que o aproveitamento indevido de informações pessoais vem se tornando cada vez mais frequente, trazendo à tona discussões sobre os critérios de segurança e privacidade dos bancos de dados e a transparência das políticas de privacidade que devem assegurar aos usuários a propriedade sobre seus dados.

Sendo a Ciência da Informação uma ciência interdisciplinar e social da Sociedade da Informação, preocupada em esclarecer o problema social concreto da informação no ciclo de construção, comunicação e uso (Le Coadic 1996), os aspectos levantados justificam a relevância desta pesquisa, na medida em que oferecem novas direções de abordagem dentro do paradigma social do campo. Segundo Capurro (2003), ao contrário dos paradigmas físico e cognitivo, o paradigma social da CI não separa a informação dos condicionamentos sociais e materiais dos sujeitos.

A propriedade nos faz refletir sobre uma série de questões relacionadas à posse dos dados. Davis (2012, p.3) faz o seguinte questionamento: “Quem é o dono dos dados, os direitos a eles associados podem ser transferidos, e quais são as obrigações de pessoas que criam e utilizam esses dados?”. Davis (2012) finaliza questionando se o indivíduo deveria ter a legitimidade de controlar os próprios dados e em qual grau?

Conforme se pode perceber, apesar dos esforços globais para regulamentar o tratamento de dados e para o estabelecimento de diretrizes para o desenvolvimento e utilização responsável de algoritmos, muitas questões permanecem em aberto nessa temática, suscitando novas reflexões da sociedade no geral.

### **Dado e Privacidade**

Na Ciência da Informação, a distinção entre dados, informação e conhecimento, embora seja imprecisa, torna-se necessária para caracterizar métodos e processos envolvidos em cada dimensão. Para Davenport (1998), o dado reúne as seguintes características: resulta de simples observações sobre o estado do mundo, é facilmente estruturado e obtido por máquinas, é frequentemente quantificado e facilmente transferível. Por isso, na perspectiva do gerenciamento da informação, considera-se fácil capturar, comunicar e armazenar os dados.

Santos e Santana (2013) apresentam diferentes conceitos de dado na perspectiva da Ciência da Informação e da Ciência da Computação, dos quais destacamos a ausência de semântica suficiente para ser interpretado por si só. Complementarmente, destacamos também a composição do dado baseada na tríade entidade - atributo - valor, que permite a sua modelagem por instrumentos digitais.

Ao analisar o conceito de dados e suas implicações para a Ciência da Informação, Santos e Santana (2013) destacam os aspectos relacionados aos conjuntos de dados, como a granularidade, isto é, o número de atributos que compõem um conjunto de dados e a diversidade de seus conteúdos. Por exemplo, o conjunto de dados estruturados da Carteira Nacional de Habilitação (CNH) possui uma granularidade

grossa em relação ao conjunto de dados estruturados do Cadastro de Pessoas Físicas (CPF).

A granularidade tem reflexos no tratamento<sup>2</sup> de dados realizado por pessoa natural ou por pessoa jurídica de direito público ou privado, conforme o Art. 3º da Lei Geral de Proteção de Dados Pessoais (LGPD).

Dados e privacidade são dois conceitos intimamente relacionados à era digital. Dados referem-se a informações coletadas e recebidas sobre indivíduos, organizações ou qualquer outra entidade. Isso pode incluir informações pessoais, como nomes, endereços, números de telefone, dados financeiros, histórico de navegação na Web e muito mais.

Para Garfinkel (2000), as informações sobre indivíduos, organizações ou outras entidades são como “sombra de dados” ou “*data shadow*”, na medida em que constituem rastros que os titulares de dados vão deixando, principalmente, em transações digitais diárias, como caixa eletrônico, pagamento com cartão de crédito, reserva de horário etc. Esses rastros de dados são mais difíceis de controlar e garantir a sua proteção contra usos indevidos. Nesse contexto, o autor afirma que estamos à beira de uma crise de informação.

De acordo com Doneda (2021), o contexto atual, em que a informação pessoal e a privacidade se relacionam, é diretamente moldado pela tecnologia, juntamente com algumas mudanças sociais. Doneda argumenta que isso ocorre principalmente devido ao reconhecimento do direito à privacidade como um direito fundamental, o que resulta em sua funcionalidade e ampliação em diversos direitos subjetivos. Esses direitos incluem não apenas o direito à vida privada e à vida familiar, mas também o direito à proteção dos dados pessoais.

Doneda empresta para si a definição que Rodotà dava para privacidade, entendida como “o direito de manter o controle sobre as próprias informações e de determinar as modalidades de construção da própria esfera privada” (Rodotà apud Doneda 2019, p. 132).

Para Mendes (2008) privacidade é a reivindicação de indivíduos, grupos ou instituições para determinar, quando, como e em que extensão, informações sobre si próprios devem ser comunicadas a outros.

A privacidade está sujeita a mudanças à medida que ocorrem as transformações sociais, políticas e culturais. No entanto, é importante ressaltar que o conceito fundamental de privacidade é considerado inviolável, pois busca garantir ao indivíduo o direito à sua vida privada, bem como o controle sobre a disponibilização de seus

---

<sup>2</sup> Toda operação realizada com dados pessoais, como as que se referem a coleta, produção, recepção, classificação, utilização, acesso, reprodução, transmissão, distribuição, processamento, arquivamento, armazenamento, eliminação, avaliação ou controle da informação, modificação, comunicação, transferência, difusão ou extração (Art. 5º, inciso X, da LGPD).

dados e a utilização de suas informações por terceiros (Grisoto; Sant’Ana; Santarém Segundo 2015 p.169).

A proteção da privacidade é um dos maiores desafios desta nova década do século XXI, caracterizada como era da economia de dados. A coexistência entre o contexto da coleta onipresente de dados e a crescente preocupação com a privacidade e a regulamentação dos dados pessoais coloca vários desafios aos especialistas em privacidade. Por isso, estabelecer práticas éticas em matéria de dados e desenvolver os meios técnicos de proteção dos dados sensíveis são dois pilares fundamentais para enfrentar os desafios da privacidade (CGC 2020).

Do ponto de vista ético, uma das principais dificuldades no tratamento de dados está em equilibrar os direitos dos titulares de dados e os potenciais benefícios que esse tratamento pode oferecer. No Quadro 1 apresentamos um panorama crítico dos modelos de coleta de dados considerados pela CGC (2020).

**Quadro 1.** Modelos de coleta de dados e preocupações éticas envolvidas.

Modelo	Características	Vantagens	Desvantagens
Doação Opt-In	O titular doa os seus dados, sem qualquer tipo de incentivo ou compensação	Garantia na proteção dos direitos dos titulares dos dados porque a doação é livre, isto é, não há preocupação com injustiças, coação ou exploração	É pouco provável que se consiga coletar grandes conjuntos de dados representativos da população
Compulsório	Os titulares são obrigados a compartilhar os seus dados, independentemente da sua vontade	Coleta de uma quantidade elevada e diversificada de dados	Risco a injustiças, exploração e violação de direitos e liberdades das pessoas
Opt-Out ou Autoexclusão	Presume-se que todos os titulares consentem a coleta e análise de seus dados, a menos que declarem o contrário	Equilibra melhor os objetivos de coletar dados suficientes e de respeitar os direitos das pessoas	Possibilidade de subgrupos de titulares de dados optarem desproporcionalmente pela exclusão, influenciando a representatividade dos dados; possibilidade de falta de informação suficiente sobre os dados que estão sendo coletados e a forma como podem ser utilizados no futuro; complexidade dos mecanismos de exclusão
Mercado	Os titulares doam os seus dados em troca de algum tipo de compensação ou incentivo	Potencial de gerar grandes conjuntos de dados e garantir uma harmonia entre as partes	Possibilidade de agravar desigualdades, dando maior incentivo aos indivíduos mais pobres para doarem os seus dados

Embora os modelos de autoexclusão e de mercado tendem a equilibrar melhor os direitos dos titulares de dados e os potenciais benefícios das operações de tratamento, também não estão livres de desafios e dilemas éticos.

O outro questionamento que vem sendo levantado a respeito de dados pessoais recai sobre a sua equiparação à propriedade. Para CGC (2020), essa equiparação traz três problemas: 1) ao contrário da propriedade física, que pode ser facilmente delimitada e os seus contornos partilhados publicamente, algumas informações pessoais nunca

deveriam tornar-se dados. Desta forma, questionam-se os meios técnicos, a razoabilidade e os limites da anonimização prevista no inciso XI, do Art. 5º, da LGPD; 2) os atores individuais, económicos e públicos envolvidos na aquisição, acumulação e utilização de dados privados são heterogêneos. A propriedade pode ser transacionada entre os três tipos de atores em pé de igualdade. O mesmo não acontece com os dados, cuja aquisição confere muito mais poder aos agentes económicos e públicos do que aos indivíduos; 3) as pessoas que possuem uma propriedade têm autoridade moral para a vender, ao passo que os indivíduos não têm autoridade moral para vender os seus dados pessoais.

Na matéria sobre direitos autorais, a Lei nº 9.610, de 19 de fevereiro de 1998 adota uma concepção dualista, que considera dois direitos: um de carácter moral e outro de carácter patrimonial. Os direitos de natureza patrimonial podem ser transferidos a terceiros, temporária ou definitivamente, por força de negócios jurídicos. Pelo contrário, os direitos morais são vínculos que unem o autor à sua criação, por isso, são inalienáveis e irrenunciáveis. Ora, a equiparação dos direitos pessoais à propriedade deveria diferenciar o carácter moral do carácter patrimonial. Contudo, essa diferenciação levaria a um impasse no tratamento porque todos os dados pessoais caberiam no atributo moral ou na personalidade do titular, isto é, seriam inalienáveis e irrenunciáveis.

Garkinfel (2000) vai na mesma linha de raciocínio, ao considerar que, ao contrário de outras formas tradicionais de propriedade, a informação possui características próprias: não se desgasta e o seu valor aumenta com o uso. Os dados podem simultaneamente serem transferidos para terceiros e continuarem vinculados ao titular, produzindo vários efeitos jurídicos nas instâncias dos atores envolvidos (por exemplo, indivíduos e agentes económicos e públicos).

A situação colocada por Garfinkel (2000) torna-se complexa ao analisarmos a LGPD. Conforme o inciso XII, do Art. 5º, o consentimento é a manifestação livre, informada e inequívoca pela qual o titular concorda com o tratamento de seus dados pessoais para uma finalidade determinada. Ora, ao autorizar o uso de seus dados pessoais, o titular o faz em condições específicas, que deveriam ser formuladas em linguagem clara e simples, respeitando os limites dos direitos e liberdades individuais, bem como os princípios de equidade e equilíbrio. Porém, na maioria das vezes, não é assim que acontece.

Numa citação à Pamela Samuelson, professora de direito autoral da Universidade da Califórnia, nos Estados Unidos, Garfinkel (2000) nos mostra o dilema do titular de dados. É contraditório ou paradoxal concordar com o tratamento de seus dados pessoais para uma finalidade determinada, para depois se preocupar com possíveis desvios e abusos por parte do agente de tratamento desses dados. Várias questões podem ser colocadas neste dilema: quais são os mecanismos que o titular de dados possui para aferir se o agente está tratando os dados nas condições pré-estabelecidas e acordadas entre as partes? Considerando a complexidade tecnológica atual, por exemplo, nos campos da Inteligência Artificial ou da Ciência de dados, os órgãos de

fiscalização possuem competências e condições suficientes para acompanhar o ritmo de tratamento das empresas de alta tecnologia, como Google ou Facebook?

Kemper e Kolkman (2018) alertam à necessidade urgente de estabelecer mecanismos de transparência e responsabilização no desenvolvimento, implementação e uso de algoritmos em organizações públicas e privadas. Para os autores, embora existam avanços na orientação sobre o uso de dados, como os princípios FAIR (localizável, acessível, interoperável e reutilizável) para gestão e administração de dados científicos, os princípios FACT (justiça, precisão, confidencialidade e transparência) para ciência de dados, entre outras iniciativas, alguns problemas antigos persistem e novos problemas surgem a cada dia. A solução passa pela compreensão do impacto dos modelos algorítmicos e dos objetos de quantificação enquanto extensões de redes de pessoas e organizações, inseridas nos contextos sociais, políticos e económicos existentes. A confiabilidade e transparência dos algoritmos que utilizam dados pessoais é questionável. Logo, a responsabilização algorítmica deve ser feita a partir de uma perspectiva crítica de conjuntos sociotécnicos de pessoas e máquinas. Assim, indivíduos, organizações, governos, campos científicos, etc., são chamados a contribuir com maior efetividade, para que tenhamos um equilíbrio entre a proteção dos direitos dos titulares de dados e a criação de vantagens competitivas a partir de operações de tratamento de dados pessoais.

### **Ciclo de Vida de Dados**

Rahul e Banyal (2020) consideram que o ciclo de vida dos dados apresenta todo o processo de dados no sistema, através de uma estrutura de alto nível para planejar, organizar e gerenciar todos os aspectos durante suas fases (coleta, armazenamento, usabilidade, compartilhamento e descarte dos dados). Logo, para uma implementação bem-sucedida do *Big Data*, é necessário manter o ciclo de vida dos dados sob um sistema ou modelo seguro, completo e eficaz de gerenciamento de dados.

Shah, Peristeras e Magnisalis (2020), por meio de uma Revisão Sistemática da Literatura, identificaram e analisaram 76 modelos de ciclos de vida de dados para propor uma estrutura de ciclo de vida de dados para governos. Os modelos analisados eram referentes a diversos domínios de aplicação, como pesquisa científica, Web Semântica, bancos de dados, qualidade de dados, processamento de dados etc. Além disso, envolviam diferentes etapas do ciclo de vida, algumas confusas e difíceis de associá-las às práticas de gestão em organizações.

O modelo de ciclo de vida de dados para governos orientados a dados, proposto por Shah, Peristeras e Magnisalis (2020), envolve 14 fases: planejamento, coleta, preparação, análise, visualização, acesso, compartilhamento, uso e reuso, arquivamento, descarte, governança, qualidade, armazenamento e segurança ou proteção de dados. Para cada fase existem funções principais a serem consideradas. Por exemplo, na fase de coleta deve-se apresentar um programa de conscientização sobre proteção de dados, coletar metadados com base em padrões de metadados, para garantir a interoperabilidade, etc.



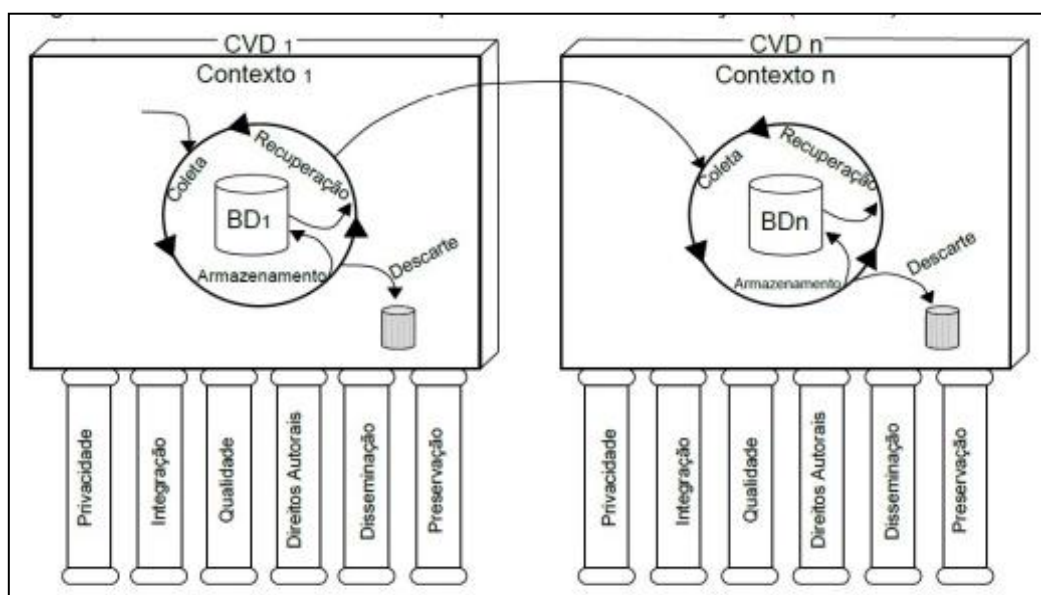
Embora o modelo dos autores possa ser utilizado em várias áreas do setor público, como saúde, agricultura, educação; seja iterativo e envolva a completude do em relação aos desafios do *Big Data*, na nossa análise, seria de difícil compreensão e aplicação no contexto desta pesquisa. Por isso, escolhemos o modelo de Ciclo de Vida dos Dados proposto por Santana (2016), que foi desenvolvido levando em consideração as características e especificidades da Ciência da Informação.

Um dos desafios da Ciência da Informação, no contexto informacional atual, é contribuir para o novo cenário de acesso e uso intensivo de dados, explorando fatores e características que permitam a ampliação do equilíbrio entre os usuários do processo e a otimização do uso dos dados.

### Fases do ciclo de vida dos dados e fatores paralelos – CVD-CI

A Figura 1 ilustra as fases do ciclo de vida dos dados que devem ser consideradas na perspectiva epistemológica e pragmática da Ciência da Informação: coleta, armazenamento, recuperação e descarte. Para cada fase do ciclo de vida devemos considerar os seguintes fatores norteadores: privacidade, integração, qualidade, direitos autorais, disseminação e preservação (Santana 2016).

Figura 1. Ciclo de Vida dos Dados para Ciência da Informação – (CVD-CI).



Fonte: Sant'Ana (2016).

**Coleta:** a princípio, é necessário obter dados que possam ser utilizados para atender a uma necessidade específica ou demanda esperada relacionada a um contexto específico. Neste primeiro momento são necessárias medidas para implementar um plano de ação, análise de viabilidade e realização da coleta de dados (Sant'Ana 2016).

**Armazenamento:** Posteriormente, depois da Coleta dos Dados, pode ser necessário e útil manter os mesmos disponíveis de alguma forma para referência futura. Neste

momento passa-se, então, a uma segunda fase que é aquela em que os esforços são no sentido de manter estes dados em uma determinada base de dados. Esta fase recebe o nome de armazenamento (Sant’Ana 2016).

Recuperação: Esta fase tem como objetivo a viabilização da recuperação dos dados. Dentre outras questões fundamentais desta fase pode-se destacar: quais dos dados armazenados serão disponibilizados? Existe algum público-alvo específico? Durante o processo de recuperação quais são os riscos à privacidade dos indivíduos ou entidades referenciados pelos conteúdos recuperados? (Sant’Ana 2016).

Descarte: Nesta fase são analisados quais dados são importantes para continuar sendo mantidos em uma base e dados e dos demais são descartados, ou seja, excluídos. Como garantir e explicitar que estes dados foram realmente excluídos e não simplesmente ocultos? Para o descarte foi considerada a necessidade de preservação em seus diversos aspectos? (Sant’Ana 2016).

Em cada etapa do ciclo de vida é necessário considerar dimensões dos fatores norteadores: privacidade, integração, qualidade, direitos autorais, disseminação e, preservação. Por exemplo, a LGPD, no inciso III, estabelece o conceito de dado anonimizado – dado relativo ao titular que não possa ser identificado, considerando a utilização de meios técnicos razoáveis e disponíveis na ocasião de seu tratamento.

Este conceito, conforme Garfinkel (2000), também pode ser problemático e controverso. Citando casos de tratamento de amostras de sangue por Centros de Controle de Doenças, o autor afirma que: de acordo com as diretrizes éticas atuais, os testes que utilizam partes do corpo humano não são considerados violações dos direitos do paciente se o seu nome for removido das amostras. Supostamente, remover o nome de uma pessoa torna a amostra anônima (Garfinkel 2000, p. 188). Porém, com os avanços da Inteligência Artificial, algumas operações de cruzamento de dados podem identificar com facilidade o seu titular.

As questões que acabamos de levantar reforçam a necessidade de mais discussões sobre dados e privacidade, principalmente no que tange ao tratamento.

## **PROCEDIMENTOS METODOLÓGICOS**

O campo da Ciência da Informação, segundo Bawden e Robinson (2012), pode ser entendido a partir da sua preocupação com o estudo da informação humana registrada, com foco nos componentes da cadeia de comunicação, investigados sob a perspectiva da análise de domínio. Nesse sentido, este trabalho utiliza a análise de domínio para investigar as discussões acerca de aspectos legais e de privacidade no uso de dados na Ciência da Informação, no Brasil.

Segundo Guimarães (2014), as abordagens teóricas e aplicadas da Ciência da Informação utilizam a análise de domínio para caracterizar e avaliar a ciência, identificando as condições da sua construção e socialização.

Smiraglia (2015) apresenta a análise de domínio como paradigma metodológico para a organização do conhecimento e como conjunto de técnicas e ferramentas para identificação de bases de conhecimento específicas para os diversos setores científicos.

Na Ciência da Informação, a análise de domínios de conhecimento como comunidades discursivas ganhou espaço a partir dos anos 1995, com a proposta de Hjørland e Albrechtsen. Mais tarde, o delineamento das onze abordagens para análise de domínio, de Hjørland (2002), clarificou o quadro para sua aplicação em estudos de organização do conhecimento do campo. Essas abordagens são: 1) Produção de guias de literatura e entradas de assunto; 2) Construção de classificações especiais e tesouro; 3) Indexação e recuperação de especialidades; 4) Estudos empíricos de usuários; 5) Estudos bibliométricos; 6) Estudos históricos; 7) Estudos de Documentos e Gênero; 8) Estudos Epistemológicos e Críticos; 9) Estudos Terminológicos, linguagem para propósitos especiais (na língua inglesa, *Special Purpose Languages – SPL*), bases de dados semânticas e estudos de discurso; 10) Estruturas e Instituições na Comunicação Científica, e 11) Cognição Científica, Especialistas do Conhecimento e Inteligência Artificial: (Tennis, 2003; Guimarães, 2014; Hjørland, 2017).

Smiraglia (2015) fez um levantamento de estudos analíticos conduzidos e relatados na literatura do domínio e observou que, apesar do crescimento do paradigma de análise de domínio na organização do conhecimento ao longo do tempo, com embasamento nas onze abordagens de Hjørland, alguns contextos continuavam fora da análise. Por exemplo, guias de literatura; especialidade de indexação; e estudo de gênero e comunicação científica. Consequentemente, apresentou a seguinte taxonomia revisada de abordagens: 1) guias de assuntos ou guias literatura; 2) classificações especiais e tesouros; 3) estudos empíricos de usuários; 4) estudos informativos; 5) estudos históricos; 6) documentação e estudos de gênero; 7) estudos epistemológicos e críticos; 8) estudos terminológicos; 9) semântica do banco de dados; 10) análises de discurso e 11) cognição, conhecimento especializado e inteligência artificial.

Neste estudo, a análise do domínio se concentrou na abordagem 5) Estudos históricos, conforme a classificação de Smiraglia (2015), e na abordagem 6) Estudos históricos de Hjørland (2002).

Tennis (2003) nos mostra que o principal problema no uso da análise do domínio não está no processo investigativo do objeto, mas sim na especificação da clareza do objeto a ser investigado – o domínio. Como solução, o autor propõe dispositivos analíticos – eixos, de modo a delinear o que o analista de domínio está estudando. O eixo Áreas de Modulação fornece os parâmetros sobre o nome e a extensão do domínio, respondendo à pergunta: como se chama o domínio e o que abrange? O eixo Graus de Especialização qualifica e define a intenção do domínio, diminuindo a sua extensão. Assim, quanto ao parâmetro nomeação, o domínio da pesquisa é produção científica da Ciência da Informação. A extensão do domínio é formada por estudos sobre aspectos legais no uso de dados pessoais. Considerando a diversidade de abordagens possíveis neste domínio, a intenção ou o foco do domínio é: autores e suas

filiações institucionais; assuntos tratados considerando as fases do CVD-CI; Em seguida, foi feita uma análise histórica de tendências das discussões nos últimos dez anos dos autores que mais contribuíram com o Ciclo de Vida dos Dados.

Para o levantamento bibliográfico da produção científica, a seguinte estratégia de busca foi usada: “dados pessoais” AND “ciência da informação”; (“*personal date*” AND “*information science*”), nas bases de dados: Biblioteca Digital Brasileira de Teses e Dissertações do Instituto Brasileiro de Informação em Ciência e Tecnologia – IBICT (BDTD), Base de Dados de Referência de Artigos de Periódicos em Ciência da Informação (BRAPCI), *Scientific Electronic Library Online* (SciELO) Brasil, Anais do Congresso Brasileiro de Organização e Representação do Conhecimento organizado pela ISKO-Brasil; Anais do Encontro Nacional de Pesquisa em Ciência da Informação - ENANCIB.

Dos resultados obtidos, apenas as publicações que apresentaram no título, resumo e/ou foram contabilizadas as palavras-chave, o termo “dados pessoais” e que estivessem em português e de origem brasileira.

Os critérios de inclusão dos estudos foram: publicação dos últimos dez anos; publicação em português; acesso ao texto completo; o estudo aborda o tema na área de modulação e no foco do domínio. Assim, a análise limitou-se a artigos, trabalhos apresentados em eventos científicos e trabalhos acadêmicos (teses, dissertações, trabalhos de conclusão de curso e relatórios de iniciação científica), publicados até dezembro de 2022.

As publicações selecionadas para compor o corpus de análise foram sistematizados em tabela, computando as informações: autor; título; ano de publicação; filiação institucional; fase CVD; fatores norteadores; assuntos tratados; base de dados.

As áreas de aplicação do estudo foram especificadas a partir das etapas do ciclo de vida de dados: coleta, armazenamento, recuperação e descarte (Sant’Ana, 2016).

## RESULTADOS E DISCUSSÕES

Para alcançar o objetivo, foi feita a coleta e categorização dos estudos, com a devida identificação de informações sobre os autores, suas respectivas instituições, assuntos tratados, respectivas teorias de sustentação e áreas de aplicação de estudos.

A partir da compilação de dados, os trabalhos foram categorizados e a geração de gráficos ocorreu com informações de distribuição temporal de autores e suas filiações institucionais; área de aplicação do estudo de acordo com o CVD – CI. Em seguida, foi feita uma análise histórica de tendências das discussões nos últimos dez anos dos dez autores mais colaborativos com o CVD - CI.

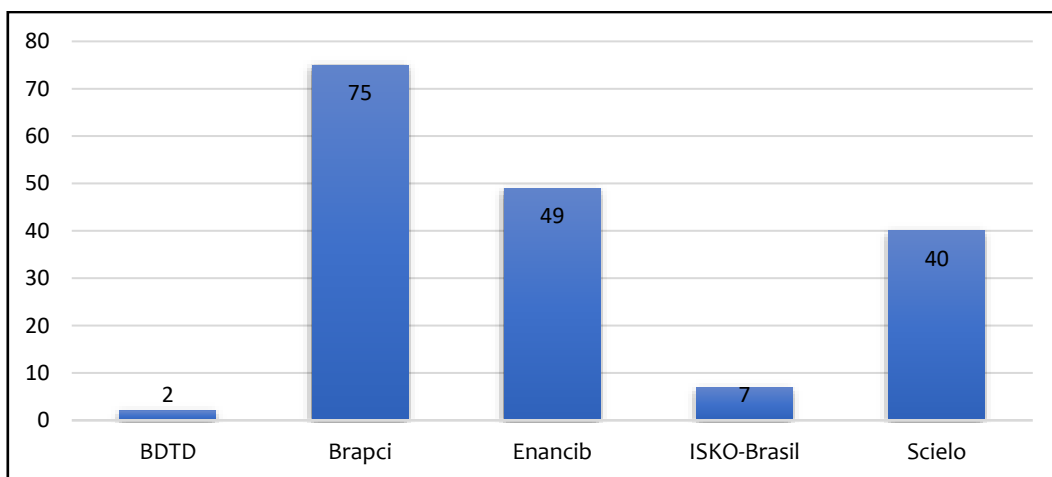
Conforme citamos anteriormente, esta pesquisa considerou como unidades de análise os trabalhos apresentados nas Bases de Dados BRAPCI, SciELO, BDTD, ISKO-Brasil e

ENANCIB na forma de Comunicações Orais e de Pôsteres, ao longo da trajetória dos últimos dez anos (2012-2022). Do total de publicações recuperadas (350), buscou-se identificar aqueles que tiveram os termos privacidade; dados pessoais; lei de acesso à informação e lei geral de proteção de dados pessoais, no título, resumo e/ou palavras-chave. Dos resultados obtidos das Bases de Dados apenas 173 artigos estavam de acordo com os critérios estabelecidos para fazer parte da análise da pesquisa.

Na base de dados BRAPCI foram recuperados 75 artigos, seguidos da ENANCIB com 49 artigos, SciELO com 40 artigos, ISKO-Brasil com 7 e a BDTD com 2 artigos recuperados.

Assim, para este estudo, considerando a inclusão definida e critérios de exclusão na metodologia, foram analisadas 173 publicações. A análise limitou-se a artigos, trabalhos apresentados em eventos científicos e trabalhos acadêmicos.

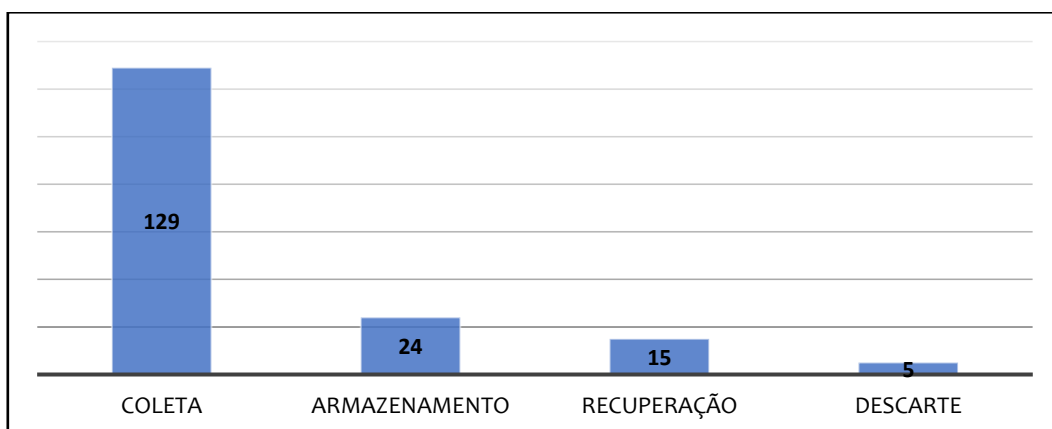
**Figura 2.** Resultados das Buscas nas Bases de Dados.



Fonte: elaboração própria.

A Figura 3 apresenta a distribuição dos artigos nas categorias em estudo.

**Figura 3.** Quantidade de Artigos por Fase do CVD – CI.



Fonte: elaboração própria.

A partir da compilação dos dados coletados, foi possível identificar que 75% dos trabalhos dizem respeito a fase coleta de dados, seguidos das fases de armazenamento com 14%, recuperação com 8,7% e descarte com 2,8%.

A fase coleta de dados é evidenciada no estudo “a privacidade no acesso a dados sob a perspectiva da teoria econômica na ciência da informação”, de Moreira e Santana (2016), que tinha por objetivo apresentar uma perspectiva da privacidade abordando questões econômicas, tecnológicas e regulamentadoras, e apontar tecnologias envolvidas no processo de acesso a dados pessoais. O estudo considera que estudar questões relacionadas à privacidade sob o viés da Teoria Econômica permite que os dados pessoais possam ser considerados moedas de troca nas transações entre empresas e consumidores, porém, é necessário um equilíbrio entre o nível de proteção e o compartilhamento de dados.

A fase armazenamento é evidenciada no trabalho “perspectivas jurídicas da relação entre *big data* e proteção de dados”, de Szinvelski, Arceno e Francisco (2019), que tem como objeto discutir a (in)compatibilização entre a *big data* e a proteção dos dados pessoais. O trabalho considera fundamental o alerta relativo ao inter-relacionamento entre os três paradoxos do *big data*: a transparência, a identidade e o poder.

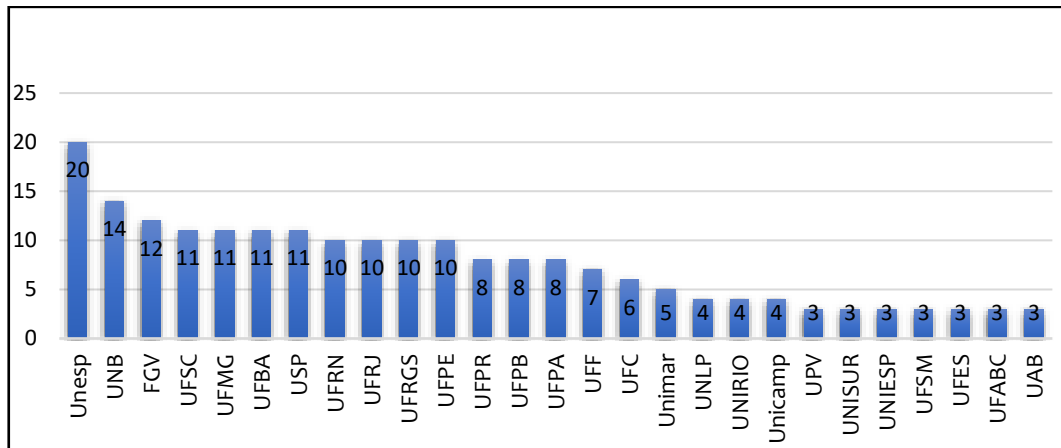
A fase recuperação é evidenciada no estudo “o ciclo de vida dos dados no contexto da pesquisa em *User Experience*”, de Barros, Rodas, Vidotti e Santana (2023), que faz uma releitura do modelo de Ciclo de Vida dos Dados, concluindo que o planejamento no contexto do Ciclo de Vida dos Dados é de suma importância para pesquisas de *User Experience*, por pensar no acesso e uso dos dados a curto e longo prazo.

A fase descarte é tratado pelos autores Silva e Araújo (2014) no artigo da Modelo para o descarte seguro da informação em suporte digital. Os autores destacam que o descarte da informação (dado) é de suma importância, por isso, é necessário garantir que essa informação não será recuperada novamente.

A Categoria Coleta de Dados conta com duzentos e quarenta e três autores distribuídos por cinquenta e cinco instituições. A Figura 4 ilustra que as Instituições que mais possuem autores afiliados que discutem sobre o tema são Universidade Estadual Paulista (UNESP) com 20 autores afiliados, seguida da Universidade de Brasília (UNB) com 14 autores, a Fundação Getúlio Vargas (FGV) com 12 autores afiliados.

A Universidade Federal de Santa Catarina (UFSC), Universidade Federal de Minas Gerais (UFMG), Universidade Federal da Bahia (UFBA) possuem 11 autores afiliados por instituições, seguidas da Universidade de São Paulo (USP), Universidade Federal do Rio Grande do Norte (UFRN), Universidade Federal do Rio de Janeiro (UFRJ), Universidade Federal do Rio Grande do Sul (UFRGS), Universidade Federal de Pernambuco (UFPE) com 10 autores afiliados por instituições e as demais seguem com oito ou menos autores por instituições. Com as informações coletadas nas bases de dados, cinquenta e cinco instituições tiveram publicações referentes a categoria Coleta de Dados.

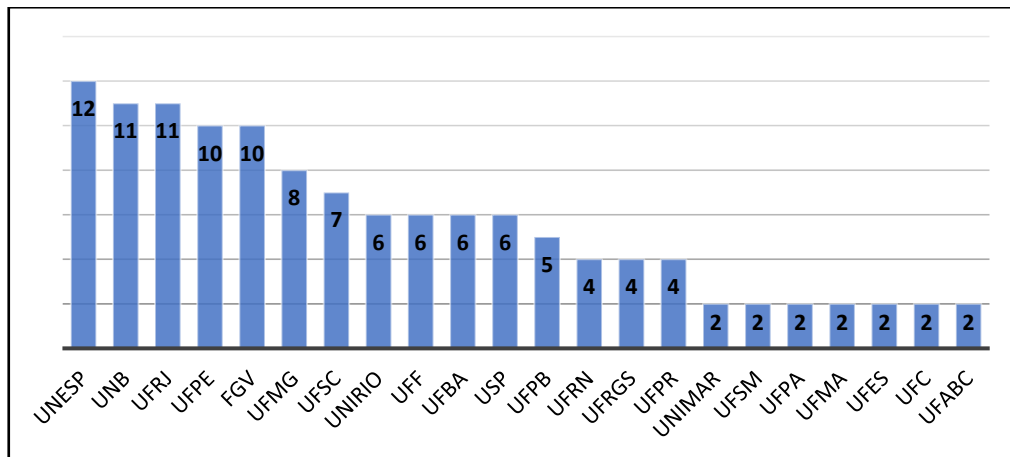
Figura 4. Afiliações dos Autores.



Fonte: elaboração própria.

Na Figura 5 estão representadas apenas as instituições que tiveram duas ou mais publicações sobre o tema. Nota-se que a UNESP teve 12 publicações sobre a Fase da Coleta de Dados, seguida das instituições Universidade de Brasília (UnB) e Universidade Federal do Rio de Janeiro (UFRJ), com 11 publicações respectivamente. Já a Universidade Federal de Pernambuco (UFPE) e a Fundação Getúlio Vargas (FGV) contribuíram com 10 publicações cada uma. As demais instituições tiveram uma variação de oito a dois artigos.

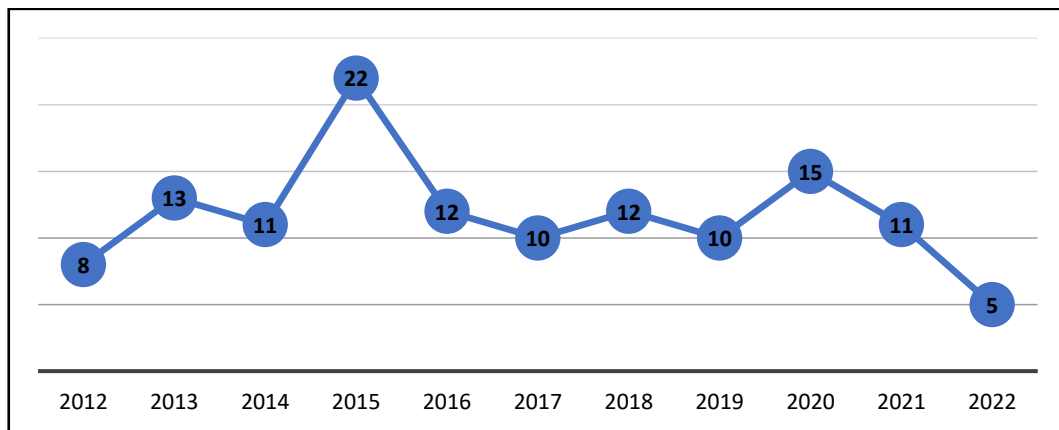
Figura 5. Publicações por Instituições.



Fonte: elaboração própria.

Em comunidades acadêmicas como das Ciências Sociais e Humanas, a análise de citações fornece vestígios de evidências de relações temáticas e teóricas. Nesse contexto, os dados das publicações foram organizados de forma que fosse possível apresentar padrões de reconhecimento explícito de cada publicação. Assim, quantitativamente, a Figura 6 mostra a frequência de publicações segundo o ano.

Figura 6. Frequência anual de publicações.



Fonte: elaboração própria.

Os últimos anos testemunharam uma preocupação crescente com a temática de dados e privacidade na Ciência da Informação. À medida que as más práticas de acesso e utilização de dados vêm à tona, os cidadãos estão começando a entender o custo real do uso de tecnologias digitais online. Acertar na privacidade dos dados é um dos maiores desafios desta nova década do século XXI.

A regulamentação da Lei de Acesso à Informação (LAI) nº 12.527/2011, que tem como objetivo promover a transparência e a prestação de contas dos órgãos públicos, além de determinar que o tratamento das informações pessoais detidas por entidades e instituições nela abrangidas seja realizado de modo transparente, respeitando o direito fundamental à proteção da intimidade, da vida privada, da honra e da imagem das pessoas, tem reforçado essa compreensão da privacidade.

Nota-se um aumento significativo das publicações em relação a fase Coleta de Dados referente aos anos de 2013 a 2017. Embora a nossa pesquisa não tenha investigado o motivo da maior concentração dos estudos na fase de coleta de dados, percebe-se que, no geral, entende-se que essa fase envolve as demais do ciclo de vida de dados.

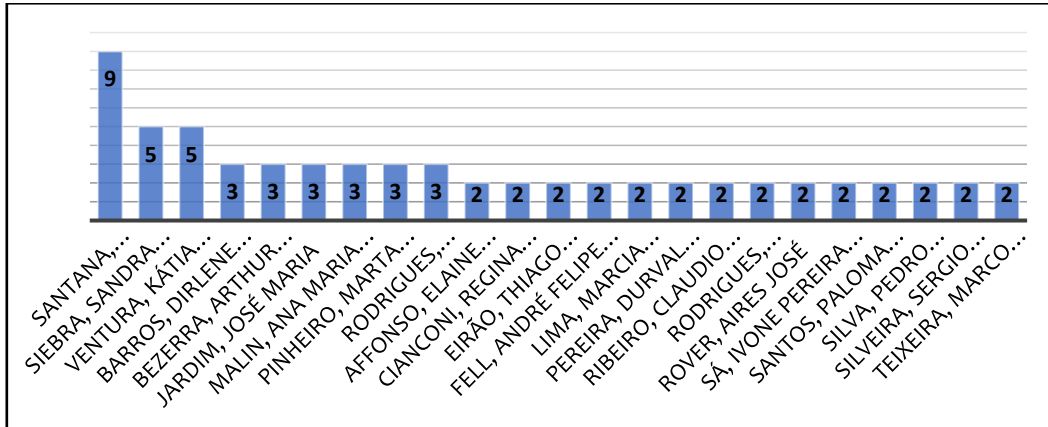
No ano de 2018 é regulamentada a Lei Geral de Proteção de Dados (LGPD), que visa regulamentar o tratamento de dados pessoais no país. A LGPD tem como objetivo principal proteger os direitos fundamentais de privacidade e garantir o controle dos indivíduos sobre suas informações pessoais. Ela estabelece diretrizes claras para a coleta, o armazenamento, o uso, o compartilhamento e a transferência de dados pessoais por parte de organizações públicas e privadas.

Percebe-se que após a regulamentação da LGPD o número de publicações sobre a fase de Coleta de Dados do Ciclo de Vida de Dados passa a ter um declínio, com exceção do ano de 2020 que contou com 15 publicações. Acredita-se que o declínio de publicações se dá pelo motivo de a partir de 2018 haver uma lei para proteger os direitos fundamentais de privacidade, e garantir o controle dos indivíduos sobre suas informações pessoais.



No ranking dos dez principais autores com maior número de publicações relacionadas à fase de coleta, o pesquisador Sant'Ana, R.C.G ocupa o primeiro lugar, com nove artigos publicados. Em seguida, os pesquisadores Siebra, S.A. e Ventura, K.S. ocupam o segundo e terceiro lugares, com cinco publicações cada. Os pesquisadores Barros, D.S.; Bezerra, A.C.; Jardim, J.M.; Malin, A.M.B.; Pinheiro, M.M.K.; Rodrigues, G.M.; e Affonso, E.P. ocupam, respectivamente, da quarta à décima posição no ranking, com três artigos cada um.

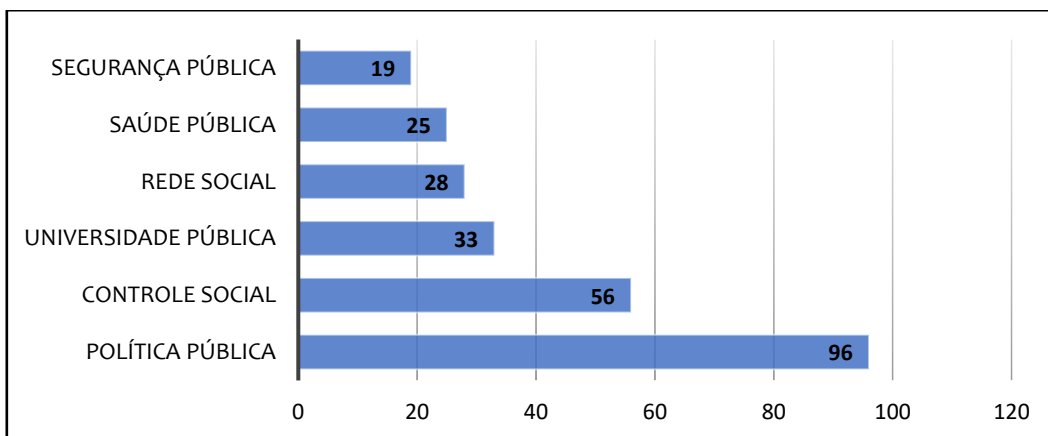
Figura 7. Publicações por Autores Categoria Coleta.



Fonte: elaboração própria.

Conforme a Figura 8, nos últimos dez anos, ocorreu um aumento significativo no número de autores dedicados à exploração de diversos recursos relevantes principalmente no campo da Política Pública, onde 96 autores se dedicaram a abordar o assunto, enquanto 56 autores se concentraram em pesquisas sobre Controle Social. Além disso, 28 autores se aprofundaram na temática da Rede Social, buscando compreender suas explicações e efeitos. Outros temas considerados incluem Segurança Pública, com 19 autores, Saúde Pública, com 25 autores e Universidade Pública, que atraiu a atenção de 33 investigadores. Essa diversidade de autores e áreas de estudo reflete a importância e a complexidade desses assuntos nos últimos dez anos.

Figura 8. Assuntos Tratados por Autores na Categoria Coleta.



Fonte: elaboração própria.

Os autores Sant'Ana, R.C.G; Siebra, S.A; Ventura, K.S; Jardim, J.M; Malin, A.M.B; Pinheiro, M.M.K; Rodrigues, G.M e Affonso, E.P. (Figura 5) trazem suas preocupações em relação a problemas que afetam a sociedade como um todo. Os autores levantam questões que envolvem o poder público, decorrentes de situações socialmente problemáticas e buscam soluções para essas questões.

As políticas públicas são instrumentos importantes para enfrentar desafios sociais e econômicos, além de desempenhar um papel crucial na gestão de assuntos públicos, na promoção do bem-estar social e no enfrentamento dos desafios enfrentados pela sociedade. Elas são importantes para moldar e direcionar a ação governamental e devem ser desenvolvidas e integradas de forma responsável e com base em evidências para garantir resultados duradouros para a sociedade.

O autor Bezerra, com três publicações, ocupa a quinta posição, abordando em seus textos temas como Sociedade, Redes Sociais e Segurança Pública. Ao incorporar a contribuição da teoria crítica em seus estudos, ele busca um pensamento embasado em fundamentos sociológicos que consideram os aspectos políticos e psicológicos do atual regime de informação.

Os demais autores também catalogados na fase Coleta de Dados do CVD-CI foram de suma importância à evolução e entendimento da fase para a Ciência da Informação.

## **CONSIDERAÇÕES FINAIS**

Este estudo analisou as discussões acerca de aspectos éticos e legais no uso de dados na Ciência da Informação, no Brasil, nos últimos 10 anos. A categorização temática foi estruturada de acordo com as etapas do ciclo de vida de dados propostos por Santana (2016), buscando identificar os principais focos de interesse das Bases de Dados durante o período estudado, além de analisar tendências, distribuições de autoria e institucionais.

Além disso, os estudos sobre dados e privacidade na Ciência da Informação também exploram a legislação e as políticas governamentais relacionadas à proteção de dados. Lei Geral de Proteção de Dados Pessoais (LGPD), Lei nº 13.709, de 14 de agosto de 2018, foi promulgada para proteger os direitos fundamentais de liberdade e privacidade e a livre formação da personalidade de cada pessoa. A Lei de Acesso Informação (Lei nº 12.527/2011) determina que o tratamento das informações pessoais detidas por entidades e instituições nela abrangidas, seja realizado de modo transparente, respeitando o direito fundamental à proteção da intimidade, da vida privada, da honra e da imagem das pessoas, o que, nos fundamentos defendidos nesta obra, corresponde à proteção do direito fundamental à privacidade;

A análise realizada revelou que aproximadamente 75% dos trabalhos das Bases de Dados estão concentrados na etapa Coleta de dados do Ciclo de Vida (CVD-CI). Nota-se a carência de estudos nas temáticas das fases do ciclo de vida dos dados como armazenamento, recuperação e descarte de dados pessoais. Esses temas são

fundamentais para compreender os aspectos legais envolvidos no uso de dados pessoais e para orientar possíveis ações na área da Ciência da Informação.

Essa pesquisa corrobora com as ideias propostas por Sant'Ana (2016), ao adotar o Ciclo de Vida de Dados (CVD-CI) para gerenciar os dados pessoais em um ambiente de *Big Data*, desde a fase de coleta até o descarte. Além disso, a aplicação da Lei Geral de Proteção de Dados (LGPD) é essencial para garantir a privacidade dos dados e garantir o uso adequado das informações.

Destaca-se a necessidade de mais estudos em prol da confiabilidade e transparência dos algoritmos que utilizam dados pessoais para direcionar a tomada de decisões estratégicas em organizações, para que tenhamos um equilíbrio entre a proteção dos direitos dos titulares de dados e a criação de vantagens competitivas a partir de operações de tratamento de dados pessoais.

## REFERÊNCIAS

BAWDEN, David, ROBINSON, Lyn, 2012. *Introduction to Information Science*. Londres: Faceta Publishing.

BRASIL. Lei n.º 12.527, de 18 de novembro de 2011. Lei de Acesso à Informação (LAI). Regula o acesso a informações previsto no inciso XXXIII do art. 5º, no inciso II do § 3º do art. 37 e no § 2º do art. 216 da Constituição Federal; altera a Lei n.º 8.112, de 11 de dezembro de 1990; revoga a Lei n.º 11.111, de 5 de maio de 2005, e dispositivos da Lei n.º 8.159, de 8 de janeiro de 1991; e dá outras providências. Brasília: Presidência da República, 2011. [Acesso em 8 março 2023]. Disponível em: [L12527 \(planalto.gov.br\)](https://planalto.gov.br/legis/leis/l12527)

BRASIL. Lei n.º 13.709 de 14 de agosto de 2018. Lei Geral de Proteção de Dados Pessoais (LGPD). Brasília, DF, ago. 2018. [Acesso em 10 abril março 2022]. Disponível em: [L13709 \(planalto.gov.br\)](https://planalto.gov.br/legis/leis/l13709)

CAPURRO, Rafael, 2003. Epistemologia e Ciência da informação. Em: *Encontro Nacional de Pesquisa em Ciência da Informação* [em linha]. Belo Horizonte, MG: Ancib. 2003. [Acesso em 12 janeiro 2022]. Disponível em: [http://www.capurro.de/enancib\\_p.htm#:~:text=A%20primeira%20parte%20menciona%20alguns,s%C3%A3o%20aprofundados%20tr%C3%AAs%20paradigmas%20epistemol%C3%B3gicos.](http://www.capurro.de/enancib_p.htm#:~:text=A%20primeira%20parte%20menciona%20alguns,s%C3%A3o%20aprofundados%20tr%C3%AAs%20paradigmas%20epistemol%C3%B3gicos.)

DAVENPORT, Thomas H. 2014. *Big Data at work: dispelling the myths, uncovering the opportunities*. Harvard: Harvard Business School Publishing.

DONEDA, Danilo, 2021. *Da privacidade à proteção de dados pessoais*. 3. ed. São Paulo: Revista dos Tribunais, 2021.

DONEDA, Danilo, 2010. Privacidade e transparência no acesso à informação pública. Em: MEZZARROBA, Ordes, GALINDO, Fernando (eds.). *Democracia eletrônica*. Zaragoza: Prensas Universitárias de Zaragoza, 2010. [Acesso em 8 abril junho 2023]. Disponível em: <https://egov.ufsc.br/portal/sites/default/files/lefis11-09.pdf>

DAVENPORT, Thomas H., 1998. Ecologia da informação: por que só a tecnologia não basta para o sucesso na era da informação. São Paulo: Futura, 1998.

DAVIS, Kord, 2012. Ethics of Big Data: balancing risk and innovation. Sebastopol: O'Reilly Media.

GARFINKEL, Simson, 2001. Database nation: the death of privacy in the 21st Century. Sebastopol: O'Reilly.

GRISOTO, Ana Paula, SANT'ANA, Ricardo Cesar Gonçalves; SEGUNDO, José Eduardo Santarém. A questão da privacidade no contexto da ciência da informação: uma análise das teses e dissertações do programa de pós-graduação em Ciência da Informação da Unesp Campus de Marília. Revista Ibero-Americana de Ciência da Informação [em linha]. 2015. v. 8, n. 2, p. 165-181. [Acesso em 23 janeiro 2023]. DOI: <https://doi.org/10.26512/rici.v8.n2.2015.2066>.

GUIMARÃES, José Augusto Chaves, 2015. Análise de domínio como perspectiva metodológica em organização da informação. Ciência Da Informação [em linha]. 26 junho 2015. Vol. 43, n.1. DOI: <https://doi.org/10.18225/ci.inf.v43i1.1415>

HJØRLAND, Birger, 2002. Domain Analysis in Information Science: eleven approaches – traditional as well as innovative. Journal of Documentation [em linha], v. 58, n. 4, p. 422-462. [Acesso em 7 fevereiro 2023]. DOI: <https://doi.org/10.1108/00220410210431136>.

HJØRLAND, Birger, ALBRECHTSEN, Hanne, 1995. Toward a new horizon in Information Science: domain analysis. Journal of the American Society of Information Science [em linha], v. 46, n.6, p.400-425. [Acesso em 14 março 2021]. Disponível em: <https://asistdl.onlinelibrary.wiley.com/doi/10.1002/%28SICI%291097-4571%28199507%2946%3A6%3C400%3A%3AAID-ASI2%3E3.o.CO%3B2-Y>

HUAWEY. Cyber Security and Data Privacy: key considerations for policymaker. 2021. [Acesso em 30 março 2022]. Disponível em: [https://www-file.huawei.com/-/media/corporate/Local-site/ca/images/2021/cyber-security-and-data-privacy\\_en.pdf](https://www-file.huawei.com/-/media/corporate/Local-site/ca/images/2021/cyber-security-and-data-privacy_en.pdf).

KEMPER, Jakko, KOLKMAN, Daan, 2018. Transparent to whom? No algorithmic accountability without a critical audience. Information, Communication & Society [em linha]. 11 maio 2018, vol. 22, n. 14, 2081-2096, [Acesso em: 13 agosto 2022]. DOI: <https://doi.org/10.1080/1369118X.2018.1477967>. Disponível em: <https://www.tandfonline.com/doi/full/10.1080/1369118X.2018.1477967>

LE COADIC, Yves-François, 1996. A Ciência da Informação. Brasília, DF: Briquet de Lemos/Livros, 1996.

LOTT, Yuri Monnerat, CIANCONI, Regina de Barros. 2018. Vigilância e privacidade, no contexto do Big Data e dados pessoais: análise da produção da Ciência da Informação no Brasil. Perspectivas em Ciência da Informação [em linha], vol. 23, n. 4, p. 117-132. [Acesso em 28 março 2022]. Disponível em: <https://periodicos.ufmg.br/index.php/pci/article/view/22594/18178>.

MAYER-SCHONBERGER, Viktor, CUKIER, Kenneth, 2014. *Big Data: A revolution that will transform how we live, work, and think*. Boston: Houghton Mifflin Harcourt.

MENDES, Laura Schertel, 2008. Transparência e privacidade: violação e proteção da informação pessoal na sociedade de consumo [em linha]. Dissertação (Mestrado em Direito). Brasília, DF: Faculdade de Direito, Universidade de Brasília. [Acesso em 9 setembro 2022]. Disponível em:

<https://repositorio.unb.br/bitstream/10482/4782/1/DISSERTACAO%20LAURA.pdf>.

OLIVEIRA, Adriana Carla Silva, MOTTA, Daniel Beltran, ESTEVES, Rita de Cássia São Paio de Azeredo, 2020. Empoderamento digital, proteção de dados e LGPD. Pesquisa Brasileira em Ciência da Informação e Biblioteconomia [em linha], vol. 15, n. 3, p. 247-261. [Acesso em 14 abril 2022]. DOI: [10.22478/ufpb.1981-0695.2020v15n3.54698](https://doi.org/10.22478/ufpb.1981-0695.2020v15n3.54698).

Disponível em: <https://www.pbcib.com/index.php/pbcib/article/view/54698/31476>.

RAHUL, Kumar, BANYAL, Rohitash Kumar, 2020. Data Life Cycle Management in Big Data Analytics. Procedia Computer Science [em linha]. vol. 173, p. 364-371. [Acesso em 28 abril 2023]. Disponível em:

<https://www.sciencedirect.com/science/article/pii/S1877050920315465>.

SANT'ANA, Ricardo César Gonçalves, 2016. Ciclo de vida dos dados: uma perspectiva a partir da ciência da informação. Informação & Informação [em linha], vol. 21, n. 2, p. 116-142. [Acesso 14 maio 2022]. DOI: <https://doi.org/10.5433/1981-8920.2016v21n2p116>.

Disponível em:

<https://ojs.uel.br/revistas/uel/index.php/informacao/article/view/27940/20124>.

SANT'ANA, Ricardo César Gonçalves, 2013. Ciclo de vida dos dados e o papel da ciência da informação. Em: *Encontro Nacional de Pesquisa em Ciência da Informação* [em linha]. Florianópolis, SC: Ancib; UFSC. 2013. [Acesso em 14 novembro 2022].

Disponível em: <https://brapci.inf.br/index.php/res/download/185245>.

SANTOS, Plácida Leopoldina Ventura Amorim da Costa, SANT'ANA, Ricardo César Gonçalves, 2013. Dado e Granularidade na perspectiva da Informação e Tecnologia: uma interpretação pela Ciência da Informação. Ciência da Informação [em linha], vol. 42, n. 2. DOI: <https://doi.org/10.18225/ci.inf.v42i2.1382>. [Acesso em 26 outubro 2022].

Disponível em: <https://revista.ibict.br/ciinf/article/view/1382>.

SHAH, Syed Iftikhar Hussain, PERISTERAS, Vassilios, MAGNISALIS, Ioannis, 2021.

DaLiF: a data lifecycle framework for data-driven governments. Journal of Big Data [em linha], 14 junho 2021. vol. 89, n. 8. DOI: <https://doi.org/10.1186/s40537-021-00481-3>. [Acesso em 3 maio 2023]. Disponível em:

<https://journalofbigdata.springeropen.com/articles/10.1186/s40537-021-00481-3>.

SIEBRA, Sandra de Albuquerque; BORBA, Vildeane da Rocha, 2018. Análise, tendências e perspectivas da produção científica em curadoria digital. Em: OLIVEIRA, Henry Poncio Cruz de; VIDOTTI, Silvana Aparecida Borsetti Gregório (org.).

Informação e Tecnologias: desenhando fronteiras científicas. João Pessoa: Editora UFPB.

SMIRAGLIA, Richard P. 2015. Domain Analysis for Knowledge Organization: Tools for Ontology Extraction. Oxford: Chandos Publishing.

The Center for the Governance of Change (CGC), 2020. Data, privacy and the individual. [Acesso em 9 maio 2022]. Disponível em:  
<https://philpapers.org/archive/VLIPM.pdf>

OLIVEIRA, Henry Poncio Cruz de; VIDOTTI, Silvana Aparecida Borsetti Gregório (org.). Informação e Tecnologias: desenhando fronteiras científicas. João Pessoa: Editora UFPB.

RODOTÀ, Stefano, 2008. A vida na sociedade da vigilância: a privacidade hoje. Rio de Janeiro: Renovar.

VIGNOLI, Richele Grengé, VERONEZ JUNIOR, Wilson Roberto, SILVA, Anahi Rocha, MONTEIRO, Silvana Drumond, 2020. Mediação da informação e a proteção da privacidade e de dados sensíveis por bibliotecários. Informação & Informação [em linha]. 26 dezembro 2020, vol. 25, n. 4, p. 280-305.  
DOI: <https://doi.org/10.5433/1981-8920.2020v25n4p280>. [Acesso em 29 agosto 2022]. Disponível em:  
<https://ojs.uel.br/revistas/uel/index.php/informacao/article/view/40922>.

ZUBOFF, Shoshana. 2019. The age of surveillance capitalism: The fight for a human future at the new frontier of power. Londres: Profile Books.