



FORMAÇÃO E INVESTIGAÇÃO EM CIÉNCIA DA INFORMAÇÃO

OPORTUNIDADES E DESAFIOS

10, 11 e 12 de novembro de 2025

POLITÉCNICO DO PORTO / ISCAP
PORTO - PORTUGAL



AUTODETERMINAÇÃO INFORMATIVA: UMA ANÁLISE DA PRODUÇÃO CIENTÍFICA REGISTRADA NA BASE SCOPUS

José Augusto Bagatini, Universidade Estadual Paulista (UNESP) / Universidad Carlos III de Madrid (UC3M), 0000-0002-8830-2075, Brasil, jose.bagatini@unesp.br

José Augusto Chaves Guimarães, Universidade Estadual Paulista, 0000-0002-0310-2331, Brasil, chaves.guimaraes@unesp.br

Antonio Perianes-Rodríguez, Universidad Carlos III de Madrid (UC3M), 0000-0002-1188-3481, Espanha, aperiane@bib.uc3m.es

Exo: Ética e Deontologia (inclusão, cidadania, direitos digitais e privacidade)

1 Introdução com referencial teórico

Na visão do jurista Alan Westin (1967), poucos valores fundamentais para a sociedade, como a Privacidade, foram deixados tão indefinidos na teoria social ou foram objeto de definições tão vagas e confusas por parte dos cientistas sociais. Isso ocorre, segundo o autor, porque a maioria dos que se debruçaram sobre o tema, supõe que a privacidade seja uma noção moderna, possivelmente surgida como uma ideia no final do século XVIII, com o protestantismo e o capitalismo, e que amadureceu no final do século XIX ou início do século XX. No mesmo sentido, Solove (2008), evidencia a característica desordenada do conceito Privacidade, afirmando enfaticamente que ninguém pôde articular concretamente seu significado, sendo, na contemporaneidade, muito abrangente, por englobar aspectos como a liberdade de pensamento, o controle sobre o próprio corpo, a solidão em casa, o controle sobre informações pessoais, a liberdade de vigilância, a proteção da reputação e contra buscas e interrogatórios, entre outros.

Na perspectiva de Westin (1967), a reivindicação humana por privacidade deriva primariamente das origens animais do homem, para defender tal hipótese, o autor apoia-se

nas obras *The Hidden Dimension* (Hall, 1966) e *The Territorial Imperative* (Ardrey, 1966). A partir desse cotejo entre as referidas obras, o autor destaca que a privacidade, enquanto direito de estar só – e ainda que parecendo uma necessidade exclusivamente humana – pode ser identificada nas mais distintas espécies, que demonstram necessidade de reclusão individual temporária ou da intimidade de pequenas unidades como elementos fundamentais para sua sobrevivência. Desse modo, sugere Westin (1967) que a busca por privacidade surge nos processos biológicos e sociais de quase toda vida e pode ser notada na denominada “tendência à territorialidade”, nome que se dá ao processo no qual um organismo reivindica uma área de terra, água ou ar e a defende contra a intrusão de membros de sua própria espécie.

Já a noção moderna de privacidade, adotada pela sociedade ocidental, remonta aos gregos e sua organização social na Antiguidade, tendo se desenvolvido a partir da tradição de limitar poderes de vigilância de autoridades sobre atividades privadas de indivíduos. Esta, segundo Westin (1967, p. 26 **tradução nossa**) pode ser entendida como “a reivindicação de indivíduos, grupos ou instituições para

determinar por si próprios quando, como e em que medida suas informações são comunicadas a terceiros. Em termos da relação indivíduo-sociedade, a privacidade pode ser entendida como a retirada voluntária e temporária de uma pessoa, por meios físicos ou psicológicos, seja em um estado de solidão ou intimidade de um pequeno grupo ou, quando entre grupos maiores, em condição de anonimato ou reserva". Esse desejo por privacidade nunca é absoluto, e cada indivíduo está continuamente engajado em um processo de ajustamento pessoal no qual equilibra o desejo de privacidade com o desejo de divulgação e comunicação de si mesmo aos outros, à luz das condições ambientais e das normas sociais estabelecidas pela sociedade em que vive. Nesse sentido, os limites da privacidade são definidos subjetivamente e sensorialmente, portanto, uma violação pode ocorrer a partir de um odor, um ruído, uma intrusão visual ou um toque, por exemplo (Westin, 1967).

Buscando conceitualizar a privacidade, a partir de uma perspectiva mais recente, Holvast (2007) destaca que essa pode ser compreendida como sendo o direito individual de *autodeterminação* dentro de certos limites, domésticos, corporais ou informacionais.

Abrangendo as problemáticas semânticas do termo privacidade, Solove (2008) preconiza que o mesmo pode ser útil como uma abreviação para se referir a um grupo relacionado de problemas, e seu valor difere dependendo de qual tipo de problema se ocupa, ou seja, não possui valor uniforme. Portanto, para que se possa compreender a privacidade, Solove (2008) propõe que a mesma seja observada a partir de uma estrutura taxonômica, fundamentada nos diferentes tipos de problemas que lhe afetam, nomeadamente, os de *coleta* de informação, de seu *processamento*, de sua *disseminação* e da *invasão* do ambiente no qual essa informação está armazenada.

À vista do exposto, nota-se que o termo Privacidade assume natureza polissêmica, sendo necessário, portanto, considerar propostas de categorização como a estrutura

taxonômica de (Solove, 2008) e/ou a divisão de (Holvast, 2007, p. 741), que sugere que a privacidade seja analisada a partir de duas dimensões. Para este último autor, tem-se uma *dimensão relacional*, que lida com a relação com outros indivíduos, por exemplo, controlando quem pode entrar em um espaço doméstico ou quem é autorizado a tocar o corpo e, pode ser descrito como "privacidade territorial ou corporal" e uma *dimensão informacional*, que se refere à coleta, ao armazenamento e ao processamento de dados de natureza pessoal.

Com o desenvolvimento do campo da computação, as problemáticas relacionadas com a dimensão informacional da Privacidade ganharam atenção ao longo das últimas décadas, principalmente porque essas tecnologias proporcionaram um grande avanço no que diz respeito a volume, velocidade e variedade em termos de coleta e processamento de dados (pessoais) em ambiente informatizado – contexto esse que pode ser expresso a partir do termo *big data*.

Como resposta para o avanço das problemáticas decorrentes da fricção entre novas tecnologias de computação e privacidade, difunde-se o termo autodeterminação informativa (*informational self-determination*), definido e qualificado na decisão de 15 de dezembro de 1983 do Tribunal Constitucional Federal da Alemanha (*Bundesverfassungsgericht*), resultante das discussões para regulamentação da coleta de dados no processo recenseamento do país. Segundo o documento, o direito à autodeterminação informativa abrange a proteção do indivíduo contra a coleta, o armazenamento, o uso e o compartilhamento ilimitado de seus dados pessoais, esse direito fundamental confere ainda, ao indivíduo, a autoridade de, em princípio, decidir por si mesmo sobre a divulgação e o uso de seus dados pessoais. Nesse contexto, restrições à autodeterminação informativa são permitidas apenas se servirem a um interesse público preponderante, requerendo uma base constitucional, e desde que satisfaça a um requisito de clareza legal sob o estado de

direito, observado o princípio da proporcionalidade (Judgment of 15 December 1983).

Desse modo, o termo passa a figurar nas discussões relacionadas a Privacidade e proteção de dados no campo das estruturas jurídicas, sendo citado explicitamente, por exemplo, no artigo 2º. da Lei Geral de Proteção de Dados brasileira (LGPD), que prevê o respeito à privacidade e a autodeterminação informativa como fundamentos da proteção de dados pessoais (Brasil, 2018) e de maneira implícita, no Regulamento Geral sobre Proteção de Dados da União Europeia (GDPR), item (6), “[...] as pessoas singulares deverão poder controlar a utilização que é feita dos seus dados pessoais” (European Parliament, 2016).

Demonstra-se, a partir do exposto, a relevância do termo para a delimitação da dimensão informacional da “Privacidade”, entretanto, a nebulosidade que o cerca afeta diretamente a compreensão de sua própria delimitação e aplicabilidade.

Desse modo, o presente trabalho objetiva, em uma perspectiva teórica, delimitar a “autodeterminação informativa”, mais especialmente no intuito de verificar como esse tema vem se apresentando na literatura científica internacional, em termos de atores, referentes, fontes e contextos institucionais e geográficos, como subsídio à caracterização desse domínio de conhecimento.

2 Procedimentos Metodológicos

A pesquisa, de natureza exploratória e bibliográfica, decorre do fato de o termo “autodeterminação informativa” e seus impactos ser tema ainda pouco abordado, notadamente na Ciência da Informação, o que exige um estudo aprofundado que possa explorar seus aspectos fundamentais. Nesse sentido, pauta-se em fontes bibliográficas (obtenção, análise e sistematização) para sua consecução.

Visando a atingir os objetivos propostos, a pesquisa teve como universo bibliográfico a base de dados Scopus, com recorte temporal

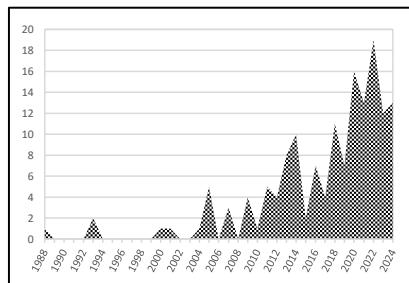
abrangendo itens publicados antes de 2025, cuja busca ocorreu a partir do termo “*informational self-determination*”, em artigos ou revisões, com a seguinte query: *TITLE-ABS-KEY (“informational self-determination”) AND PUBYEAR > 1987 AND PUBYEAR < 2025 AND (LIMIT-TO (DOCTYPE, “ar”) OR LIMIT-TO (DOCTYPE, “re”))*.

Foram, então, recuperados 150 itens, os quais foram tabulados em fichas de registro possibilitando, assim, a realização das análises apresentadas nos Resultados a seguir.

3 Resultados

O corpus de 150 itens recuperados na base Scopus refere-se a publicações entre 1988 e 2024.

Figura 1: volume de artigos publicados ano a ano



Fonte: Elaboração própria (2025)

Até o ano de 2008 a produção acerca do tema mostra-se irregular, passando a receber atenção constante somente a partir do ano de 2009 em diante, atingindo seu pico no ano de 2022, com 19 publicações. As correlações entre volume de artigos e acontecimentos – como a promulgação da GDPR e pandemia de COVID-19.

Os artigos foram publicados em 103 periódicos distintos, sendo que 24 revistas (23%) publicaram 2 ou mais artigos. Observa-se, pois, que 47% da produção(70 artigos) concentra-se em 23% dos periódicos, o que demonstra que o tema vem se consolidando mais especificamente em um grupo restrito e especializado de veículos de comunicação científica, a saber: Computer Law and Security Review (8 artigos); Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz (6 artigos); IT - Information Technology (5 artigos); Zeitschrift für Medizinische Ethik, Revista de Direito Civil Contemporâneo e Medicine and Law (4 artigos cada); BioLaw Journal, Revue Française

d'Administration Publique, European Data Protection Law Review, e International Review of Law, Computers and Technology (3 artigos cada); e AI and Society, European Journal of Human Genetics, Ethics and Information Technology, Revista del Ministerio de Trabajo y Economía Social, Revista de Derecho Político, Arbeitsmedizin Sozialmedizin Umweltmedizin, Verwaltung, BMC Medical Informatics and Decision Making, Revista de Estudios en Derecho a la Información, Philosophy and Technology, European Foreign Affairs Review, Revista Chilena de Derecho y Tecnología, Internet Policy Review, e Lecture Notes in Computer Science (2 artigos cada).

No que diz respeito às áreas de conhecimento dos periódicos, verificam-se 74 categorias temáticas distintas, ainda que se deva ressaltar que a Scopus pode atribuir mais de uma categoria a um periódico.

Ao reorganizar as categorias dos periódicos, a partir das dez principais classes da Classificação Decimal de Dewey – CDD, nota-se que 87% (131 de 150) das revistas possuem em sua categorização temática pelo menos uma classe que habita o domínio da categoria 300 - Ciências Sociais, demonstrando que há aderência predominante com questões de campos como Direito, Ciências Políticas, Comunicação e Sociologia, entre outras. A partir da análise, observa-se também o comportamento pluridisciplinar dos periódicos, categorizados nas áreas de Ciência da Computação, Informação e obras gerais – classe 000 (87 de 150 periódicos), Tecnologia – classe 600 (50 de 150 periódicos), Ciências – classe 500 (38 de 150 periódicos) e Filosofia e Psicologia – Classe 100 (19 de 150 artigos); sendo que 6 outros periódicos encontram-se nas demais categorias.

À vista do exposto, pode-se inferir que a produção científica analisada sobre autodeterminação informativa tem sido majoritariamente discutida e difundida em periódicos científicos que têm em seu núcleo preocupações relacionadas aos campos aplicados das Ciências Sociais, mas pluridisciplinarmente conectadas com as questões éticas e filosóficas que permeiam outros campos de saber (Administração, Ciência da Computação, Ciência da Informação, Comunicação, Economia, etc.), tecnologias (inteligência artificial, testes genéticos, etc.) e ambientes e contextos específicos (softwares, sistemas de informação, administração pública, bibliotecas, clínica, etc.). Tais aspectos evidenciam a natureza complexa e multifacetada dessa temática.

No que diz respeito às palavras-chave do corpus, foram identificados 500 termos distintos, que se repetiram 718 vezes. Desse universo, têm-se 122 palavras-chave (24% das palavras-chave únicas) que se repetiram duas ou mais vezes, totalizando 279 repetições o que representa 39% das palavras-chave totais usadas para delimitar os trabalhos que compõem o corpus analisado. Isso sugere que há uma concentração temática que demonstra certo grau de concordância terminológica entre os pesquisadores.

Visando a identificar como os trabalhos analisados foram representados por seus autores, procedeu-se ao controle de vocabulário e à categorização temática das 122 palavras-chave que se repetiram duas ou mais vezes, excluindo-se desse rol aquelas relativas a locais e instituições, as quais totalizaram 17 palavras-chave (6% das 122 palavras-chave únicas que foram analisadas).

A partir dessa categorização, foi possível identificar que o termo central da pesquisa –“*informational self-determination*” – repetiu-se 49 vezes (18% do total de incidências), tema esse diretamente conectado com Privacidade, este com 33 incidências (12% do total). Isso revela um forte vínculo temático entre ambos os termos que, somados, respondem por 30% do total.

Aplicando-se a análise de conteúdo (Bardin, 2013) para identificar clusters temáticos das palavras chaves, foi possível categorizar as palavras-chave, o que permitiu identificar que:

- 73 palavras-chave (26%) tratam de maneira geral sobre **Dados**, especificamente Dados Pessoais (15) e dados de saúde (2), são abordadas também questões como proteção (38) – acoplado o termo *information privacy*, que diz respeito a proteção, portabilidade (2), processamento (3), limitação de propósito de uso (2), retenção (2) e regulação no campo jurídico (2), tal como na GDPR (9), desses dados.
- 29 palavras-chave (10%) dizem respeito especificamente aos **Valores Éticos** (4) relacionados a temática, tais como consentimento (10), dignidade humana (5), transparência (3), salvaguarda (3) – termo original “*safety*” –, proporcionalidade (2) e legitimo interesse (2).
- 28 palavras-chave (10%) tratam sobre os **Direitos** abordados na produção que compõe o corpus, tais como os direitos

fundamentais (13) abordados pela Declaração Universal dos Direitos Humanos, o direito ao esquecimento (6), o direito a proteção de dados pessoais (2), o direito ao livre desenvolvimento da personalidade (2), o direito a liberdade de informação (2) – *freedom of information* termo original –, todos em um contexto digital (3).

- 21 palavras-chave (8%) são utilizadas para descreverem **Tecnologias, Ações ou Processos**, tais como Inteligência Artificial (6), *Profiling* (3), Pseudonimização (2), técnica de Privacidade Diferencial (2) e Governança (2), visando objetivos como Gerenciamento de Identidade (2), Vinculação de Registros (2) e Regulação de Riscos (2) ou a fundamentação de Políticas de privacidade (2).
- 17 palavras-chave (6%) abordam os **Contextos** em que a temática da autodeterminação pode ter desdobramentos, tais como no contexto de *big data* (6), de vigilância (3), da sociedade da informação (2), de datificação (2), de testes genéticos (2), todos relacionados com o universo online (2).
- 12 palavras-chave (4%) tratam sobre as **Aplicações**, tais como *Automatic number plate recognition - APNR* (2), *Electronic health record* (4), *Electronic patient record* (2), *DNA databases* (2) e *population census decision* (2).

O conjunto de palavras-chave que se repetiram 2 vezes ou mais demonstra que os trabalhos recuperados a partir do termo “*informational self-determination*” têm, principalmente, tratado sobre questões relacionadas à dimensão informacional da privacidade, especificamente no que diz respeito a dados (pessoais), aos valores éticos que devem, transversalmente, promover o uso desses dados, alinhado com a manutenção de direitos que resguardam a sociedade de possíveis prejuízos decorrentes de tecnologias e contextos que se apoiam em grandes conjuntos de dados.

No que diz respeito à autoria do corpus, foi possível identificar que 83 artigos (55,33%) são assinados por somente um autor, 39 artigos (26%) são de autoria dupla, 9 artigos (6,00%) são de autoria tripla, 17 artigos (11,33%) são assinados por quatro até dez autores e 2 artigos (1,33%) possuem onze ou mais autores. A média geral de autores por artigo é de 2,03, aproximando-se das práticas de

composição de autoria científica das Ciências Humanas e Sociais, em que, historicamente, tem-se predominantemente até três autores por trabalho. Destaca-se, também, que o fato de a grande maioria dos trabalhos serem assinados por apenas um autor, sugere que, na temática tratada, há uma escassa colaboração, possivelmente em decorrência de uma ainda incipiente composição de colégios invisíveis, talvez por conta do caráter emergente da temática.

No corpus, foram identificados 289 autores distintos, dos quais destacam-se Pournaras E., Buitelaar J.C., Winkler E.C. (com 3 artigos cada, representando 6% da produção total). Outros 10 autores publicaram 2 artigos cada, Asikis T., Molnár-Gábor F., Knoppers B.M., Schickhardt C., Gstrein O.J., Contreras P., Pfitzmann A., Hornung G., Schnabel C. e Hansen M. (13% da produção total). Esse grupo de 13 autores é responsável 14% (21 trabalhos) da produção que forma o corpus analisado. Mesmo que 86% da produção esteja dispersa entre outros 276 autores, nota-se que existe um núcleo, ainda que incipiente, de pesquisadores que lideram a temática.

A existência desse núcleo de autores se evidencia, ainda, pelas colaborações de autoria, verificada em 8 artigos, havendo dois artigos em colaboração de Asikis T. e Pournaras E. (Asikis & Pounaras, 2020; Asiks; Kinglmayr, Helbing & Pournaras, 2021); dois entre Hornung G. e Schnabel, C. (Hornung & Schnabel, 2009 a,b); dois entre Schickhardt C. e Winkler E.C., (Schickhardt ; Fleischer & Winkler, 2020); e ainda um entre Molnár-Gábor, F. e Knoppers, B.M. (Bernier et al, 2024) e um entre Hansen M. e Pfitzmann, A. (Hansen; Borcea-Pfitzmann & Pfitzmann, 2005).

Com objetivo de aprofundar as análises sobre vinculação institucional dos autores, os dados obtidos pela base de dados Scopus foram acoplados com os fornecidos pela base de dados Research Organization Registry – ROR. No processo de acoplamento de dados não foi possível identificar, na base ROR, as instituições de 48 (16% do total de autores) dos 305 autores, sendo que os outros 257 autores estão vinculados a 157 diferentes instituições cadastradas na base.

Metodologicamente a escolha por analisar as instituições cadastradas na base ROR se dá, principalmente, pelo fato de ser um registro global *open source*, composto de identificadores persistentes e abertos, permitindo maior facilidade na desambiguação dos dados institucionais a serem analisados.

Formatou: Tipo de letra: (predefinido) +Corpo (Calibri)

Observa-se que 154 autores (60% do total de autores com vinculação institucional comprovada) estão vinculados a 54 (34%) das instituições identificadas, sendo que os outros 101 (33%) autores (50%), estão distribuídos em 103 instituições (66%) representadas por aquelas que possuem dois ou mais autores vinculados, revelando uma grande dispersão institucional dos pesquisadores envolvidos com a temática. Tal aspecto, por sua vez, pode sinalizar pela importância disseminada do tema na comunidade científica internacional, visto ser objeto de abordagem por autores das mais diversificadas filiações institucionais.

Em busca de analisar a distribuição geográfica das instituições, procedeu-se ao acoplamento dos dados das bases Scopus e ROR, somando-os para preencher lacunas faltantes de uma ou outra base. Desse modo, foi possível identificar o país da instituição de 289 (95%) autores, enquanto 16 (5%) autores não possuíam dados de vinculação disponíveis em nenhuma das bases. A configuração da distribuição geográfica se apresenta da seguinte maneira:

Quadro 1: autores por país	
País da filiação	Autores
Germany	111
Brazil	24
Spain	16
United Kingdom	14
Netherlands	13
South Korea	11
Canada	10
Austria	8
Portugal	8
United States	8
Italy	7
Finland	7
Belgium	7
France	6
Switzerland	4
Chile	4
Sweden	4
Hungary	4
Norway	3
Czech Republic	3
Russian Federation	3
China	2
Denmark	2

Ireland	2
Australia	1
Estonia	1
Serbia	1
Poland	1
Colombia	1
South Africa	1
Croatia	1
Israel	1
Total	289

Fonte: Elaboração própria (2025).

Destaca-se que a grande incidência de autores vinculados a instituições alemãs, se dá, provavelmente, pelo fato de o país ser precursor no debate sobre proteção de dados (pessoais), onde a questão remonta ao ano de 1970, quando foi editado o texto jurídico popularmente conhecido como Lei de Hesse, tido como o primeiro a legislar sobre a proteção de dados no mundo e, posteriormente, em 1983 a questão da autodeterminação informativa, como preconizado anteriormente, promovendo assim, um debate mais amplo no ambiente científico alemão. A relação entre outros países e sua relevância, tais como Brasil com 24 autores (8%), Espanha com 16 autores (5%), Reino Unido com 14 autores (5%) e Países Baixos com 13 autores (4%) etc., deve ser analisada levando em conta a temática e as possíveis correlações com acontecimentos de variadas escalas.

Observe-se, outrossim, que cerca de 70% da produção científica nessa temática provém de um conjunto de apenas sete países, sendo quatro da Europa (Alemanha, Espanha, Reino Unido e Países Baixos), dois da América (Brasil e Canadá) e um da Ásia (Coreia do Sul), o que sugere serem esses espaços prioritários de pesquisa na área.

Observe-se, ainda, que 226 autores (78,2% do total) concentram seus vínculos institucionais com entidades europeias, demonstrando conexão entre a origem e o desenvolvimento da temática. Outro fator que pode justificar tamanha disparidade com outras áreas do globo é o fato de que a Europa tem liderado as discussões sobre Privacidade, principalmente com a GDPR. Tem-se ainda que 29 (10% do total) autores apresentam vinculação com instituições da América do Sul, dos quais 24 (83%) encontram-se vinculados a instituições do Brasil. Isso demonstra uma preocupação com o tema da autodeterminação informativa, adotado expressamente como princípio pela LGPD brasileira, e evidenciando uma forte influência europeia na

Formatou: Tipo de letra: 10 pt

construção do arcabouço jurídico que rege a proteção de dados no país. A presença de 18 autores (6,2%) vinculados a instituições da América do Norte, sugere que a temática ainda encontra dificuldade em se difundir em um espaço que pode sofrer forte influência de questões como a noção de liberdade ou a pressão de empresas de tecnologias que, conhecidamente, fazem uso de dados (pessoais) com objetivo financeiro. O restante dos autores, por sua vez, distribui-se por instituições da Ásia, com 13 (4,5%), e da África, Oceania e Oriente Médio com 1 (0,3%) autor cada. Tal aspecto pode sinalizar para o fato de que a preocupação científica com o tema tem sido mais claramente objeto de produção no mundo geograficamente Ocidental, mas que começa a expandir-se por outros espaços do globo.

Procedeu-se, então à análise das referências citadas pelo corpus, no intuito de identificar os referentes teóricos e as fontes de informação que subsidiaram a produção analisada. Para tanto, realizou-se uma extração via base de dados Scopus. Neste processo, foi possível recuperar as referências de 137 dos 150 artigos, sendo que os outros 13 artigos não possuem suas referências catalogadas, impossibilitando que se realize uma análise global dos itens referenciados. Desse modo, teve-se que juntos, os 137 artigos referenciaram 5.386 citações, com 5.093 diferentes itens.

Das 5.093 referências obtidas, 1.144 (22,46%) não possuem dados de autoria registrados na base, de modo que foi possível analisar 3.949 referências diferentes, citadas 4.225 vezes no total.

No que diz respeito à configuração de autoria e colaboração desse conjunto de 3.949 referências, foi possível identificar 10.903 autores, resultando em uma média de 2,76 autores por item, que se distribuem da seguinte maneira: 2.226 referências (56,37%) são de autoria única; 754 referências (19,09%) têm autoria dupla; 384 referências (9,72%) apresentam autoria tripla; 491 referências (12,43%) resultam da colaboração entre quatro até dez autores e; 94 referências (2,38%) são fruto de colaboração entre onze ou mais autores.

Cabe destacar que, no grupo de 94 referências assinadas por onze ou mais autores, concentram-se 3.334 nomes, ou seja, 2,38% das citações respondem por 30,63% do total de nomes identificados. Sublinha-se também, o fato de uma única referência ser assinada por 1.313 autores distintos, 12,04% de todos os nomes citados.

Observa-se comparativamente os dados sobre autoria do corpus e das referências no quadro abaixo:

Quadro 2: comparativo de autoria entre corpus e referências considerando itens com onze ou mais autores

Critério	Corpus	Referências
Média	2,03%	2,76%
Um autor	55,33%	56,37%
Dois autores	26,00%	19,09%
Três autores	6,00%	9,72%
Quatro até dez autores	11,33%	12,43%
Onze ou mais autores	1,33%	2,38%

Fonte: Elaboração própria (2025).

Formatou: Tipo de letra: 10 pt

Nota-se, a partir do comparativo acima disposto, que citações com onze ou mais autores, quando levadas em conta, podem conduzir à uma distorção na análise, principalmente, no que diz respeito à média de autores, uma variável importante para compreensão da configuração de um domínio. Assim, em busca de amainar referida distorção, procedeu-se à uma nova análise onde são desconsideradas referências assinadas por onze ou mais autores:

Quadro 3: comparativo de autoria entre corpus e referências desconsiderando itens com onze ou mais autores

Critério	Corpus	Referências
Média	1,86%	1,96%
Um autor	56,08%	57,74%
Dois autores	26,35%	19,56%
Três autores	6,08%	9,96%
Quatro até dez autores ¹	11,49%	12,74%

Fonte: Elaboração própria (2025).

Formatou: Tipo de letra: 10 pt

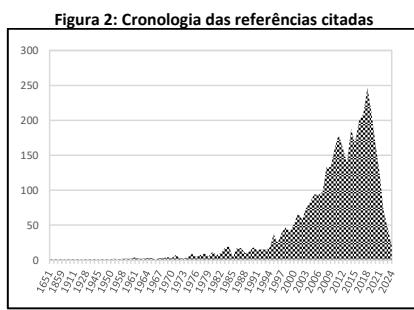
Ilustrada a análise que desconsidera as referências com alto volume de autores, é encontrada uma realidade em que a comparação com o corpus se torna. Frente a isso, é importante salientar que os variados campos de saber assumem diversas configurações de autoria, podendo haver mais ou menos nomes na assinatura dos itens publicados. Entretanto, à luz de uma análise que busca compreender a constituição de um determinado domínio, há de se atentar para as distorções que acompanham o fenômeno da *hiperautoria*, desse modo, opta-se, no presente trabalho apresentar análises que desconsiderem as referências sem dados de autoria ou assinadas por onze ou mais autores, nos levando, portanto, à um universo de 3.855 referências, que foram citadas 4.129 vezes no total.

A partir do expresso no **Quadro 2**, nota-se alta semelhança na configuração de autoria entre

corpus e citações, com divergência mais acentuada nas categorias “Dois autores” e “Três autores”. Enquanto no corpus a colaboração se dá, em maior peso, entre dois autores, nas referências têm-se maior incidência a autoria tripla. Tal diferença, pode decorrer do fato de que os trabalhos que compõem o corpus, como mencionado anteriormente, localizam-se, em maioria, no campo das ciências humanas e sociais, os quais citam em grande volume estudos de áreas outras que as de sua origem, buscando importar questões provenientes de discussões de outros campos, como da computação, da genética, da medicina, da psiquiatria, entre outros. Tal aspecto indica uma característica a ser sublinhada, a da natureza transdisciplinar dos temas privacidade e autodeterminação informativa, os quais, transversalmente, permeiam discussões de variadas áreas de saber, em um movimento contínuo, onde aporta e assimila conhecimento.

Para proceder à análise cronológica das referências, verificou-se, preliminarmente, que 173 delas (4,48% do total) não apresentavam data de publicação registrada na base de dados SCOPUS, o que levou a se trabalhar com um conjunto de 3682 referências. Nesse conjunto, tem-se com item mais antigo o *Leviathan or The Matter, Forme and Power of a Commonwealth Ecclesiastical and Civil*, de Thomas Hobbes, publicado em 1651, enquanto os itens mais recentes datam de 2024.

Na análise apresentada no gráfico a seguir é possível verificar como se dá a configuração da distribuição cronológica dos itens que compõem o universo das referências citadas.



Fonte: Elaboração própria (2025)

Observa-se, na **Figura 2**, que são citados trabalhos dos séculos XVI, XVII, XVIII e XIX, mas, é principalmente em produções do século XX e XXI que o corpus busca seu referencial teórico, especialmente, observando que a partir do ano de

1956 há citações de todos os anos, ininterruptamente.

Há de se destacar, também, a similaridade entre os dados expressos nas **Figuras 1 e 2**, pois a cronologia do corpus e a cronologia das referências possuem forte semelhança em questão de tendência – e não de volume de dados - no que diz respeito ao comportamento de publicação ao longo dos anos. Isso revela que ocorre apoio teórico em textos clássicos, mas predominantemente se faz uso de uma bibliografia que constantemente se atualiza, aspecto que pode ser um possível reflexo do processo de desenvolvimento, compreensão e regulação das tecnologias que afetam a trama social, relativamente ao tema da privacidade, processo que tem início no período pós-guerras, quando se instala um paradigma de acelerado avanço tecnológico e informacional, culminando na necessidade de novas análises para as questões relacionadas a privacidade e autodeterminação informativa.

Avançando para uma análise de como as 3.855 referências diferentes foram citadas pelo corpus, é possível identificar que 174 (4,51%) foram citadas duas ou mais vezes, somando 448 citações (10,85% das 4.129 citações), indicando uma grande dispersão no que diz respeito às fontes citadas pelos trabalhos do corpus. Entretanto, há de se destacar que, mesmo pequeno, um núcleo de obras se postula como fundamentais para os trabalhos que compõem o corpus. Listadas abaixo, têm-se as dez obras mais citadas.

1. Com 18 citações, o artigo *The right to privacy* (Warren & Brandeis, 1890);
2. Com 8 citações, o livro *A vida na sociedade da vigilância: a privacidade hoje* (Rodotà, 2008);
3. Com 8 citações, o livro *Privacy in Context: Technology, Policy, and the Integrity of Social Life* (Nissenbaum, 2010);
4. Com 7 citações, livro *Da privacidade à proteção de dados pessoais* (Doneda, 2006);
5. Com 7 citações, o livro *Proteção de dados pessoais: a função e os limites do consentimento* (Bioni, 2019);
6. Com 6 citações, o livro *The age of surveillance capitalism: the fight for a human future at the new frontier of power* (Zuboff, 2019);
7. Com 5 citações, livro *Privacy and Freedom* (Westin, 1967);
8. Com 5 citações, o livro *Understanding Privacy* (Solove, 2008);

Formatou: Tipo de letra: 10 pt, Não Realce

Formatou: Tipo de letra: 10 pt, Inglês (Estados Unidos), Não Realce

Formatou: Tipo de letra: 10 pt, Não Realce

Formatou: Tipo de letra: 10 pt

9. Com 5 citações, o artigo *The cost of reading privacy policies* (McDonald & Cranor, 2008);
10. Com 5 citações, o artigo *The Right to Informational Self-Determination and the Value of Self-Development: Reassessing the Importance of Privacy for Democracy* (Rouvroy & Poulet, 2008).

Constatada a alta dispersão de obras citadas, manifesta-se a necessidade de corroborar ou confrontar tal resultado a partir de uma análise que evidencie a existência ou não de um núcleo de autores – referentes teóricos. Assim sendo, os 8.000 nomes que assinam as 4.129 citações foram submetidos a um processo de normalização manual em que se buscou a redução de duplicatas de nomes, uma vez que o nome de um mesmo autor pode surgir de diversas maneiras (exemplo: Rodotà S., Rodotà S. e RODOTÀ). Tal etapa fez-se necessária, uma vez que os dados extraídos da base de dados SCOPUS não fornece, em grande maioria, um identificador único para os autores das referências, ou mesmo vinculação com o ORCID, o que dificulta o processo de desambiguação.

Após a normalização foi possível identificar 5.937 autores efetivamente distintos, dos quais, 962 (16,20%) foram citados duas vezes ou mais, somando 3.025 incidências (37,81% do conjunto). Isso demonstra que, em comparação com as referências e citações, a dispersão no quesito autores citados é menor. Tal divergência pode se dar, porque, ao analisar as obras referenciadas, pode-se perder de vista a trajetória de um determinado autor, que pode dispor de vários trabalhos sobre uma mesma temática, pulverizando, assim, seu nome em diversas referências distintas.

Objetivando identificar os principais referentes teóricos entre os autores identificados, aplicou-se o conceito de elitismo proposto por Price (1963). Segundo o autor, uma pequena parcela de pesquisadores é responsável pela maior parte do que é produzido em um campo de saber e, para identificar tal grupo, é necessário o cálculo da raiz quadrada do valor total, no caso, dos 5.937 autores. Desse modo, chega-se a um conjunto de 77 autores, o que nos leva aos autores com seis citações ou mais, portanto, para abranger de maneira igualitária todos aqueles que estão acima desse valor, destacam-se como os principais autores dentre as referências 92 nomes citados seis ou mais vezes.

Neste grupo de 92 nomes, figuram 21 autores vinculados aos 150 trabalhos que compõem o

corpus analisado; entretanto, ao subtrair as autocitações, apenas Purtova N. e Knoppers B.M., se encontram no conjunto selecionado pelo Elitismo de Price. Apresentam-se, a seguir, os 73 nomes (1,22% do total de autores identificados), responsáveis por 494 obras referenciadas, ou seja, 12,81% de todas as referências.

Quadro 4: autores mais citados pelo corpus e média de citação por obra

Autor	Citações	Obras	Média
Floridi L.	22	17	1,29
Rodotà S.	22	14	1,57
Boyd D.M.	21	15	1,40
Brandeis L.D.	20	2	10,00
Doneda D.	19	9	2,11
Warren S.D.	19	1	19,00
Nissenbaum H.	18	7	2,57
Solove D.J.	18	10	1,80
Acquisti A.	15	13	1,15
Mayer-Schönberger V.	15	10	1,50
De Hert P.	15	13	1,15
Ellison N.B.	12	8	1,50
Hildebrandt M.	12	12	1,00
Bioni B. R.	11	5	2,20
Gutwirth S.	11	10	1,10
Mantelero A.	11	11	1,00
Mittelstadt B.D.	11	7	1,57
Poulet Y.	11	7	1,57
Schwartz P.M.	11	7	1,57
Simitis S.	11	8	1,38
Wachter S.	11	7	1,57
Pérez Luño A.E.	11	9	1,22
Trepte S.	10	9	1,11
Purtova N.	9	9	1,00
Kaye J.	9	7	1,29
Mendes L.S.	9	6	1,50
Pape S.	9	9	1,00
Rouvroy A.	9	5	1,80
van der Hof S.	9	7	1,29
Westin A.F.	9	4	2,25
Zuboff S.	9	4	2,25
Murillo de la Cueva P.L.	9	7	1,29
Dove E.S.	8	6	1,33
Ienca M.	8	7	1,14
Koops B.-J.	8	5	1,60
Kuner C.	8	8	1,00
Mascalzoni D.	8	7	1,14
Roßnagel A.	8	8	1,00
Borgesius F.J.Z.	8	4	2,00
Albers M.	7	4	1,75
Altman I.	7	4	1,75
Ausloos J.	7	6	1,17
Bygrave L.A.	7	6	1,17
Cate F.H.	7	6	1,17
Chen J.	7	3	2,33
Cranor L.F.	7	5	1,40

Formatou: Tipo de letra: 10 pt, Inglês (Estados Unidos), Não Realce

Formatou: Tipo de letra: 10 pt, Não Realce

Formatou: Tipo de letra: 10 pt

Custers B.	7	7	1,00
Dwork C.	7	5	1,40
Helberger N.	7	7	1,00
Helbing D.	7	6	1,17
Kühling J.	7	5	1,40
Lynskey O.	7	7	1,00
Marwick A.E.	7	5	1,40
Pohle J.	7	6	1,17
Vayena E.	7	6	1,17
Knoppers B.M.	6	9	0,67
Agustina Sanllehí J.R.	6	6	1,00
Barocas S.	6	2	3,00
Baruh L.	6	6	1,00
Bennett C.J.	6	5	1,20
Böhme R.	6	3	2,00
Castells M.	6	4	1,50
Ciancimino M.	6	6	1,00
Cohen J.E.	6	4	1,50
COLAPIETRO C.	6	5	1,20
Crawford K.	6	6	1,00
Dienlin T.	6	6	1,00
Habermas J.	6	6	1,00
Hornung G.	6	5	1,20
Lampe C.	6	3	2,00
Pasquale F.	6	5	1,20
Rallo Lombarte A.	6	5	1,20
Sunstein C.R.	6	6	1,00

Fonte: Elaboração própria (2025).

Frise-se, sobre o arranjo do **Quadro 4**, dois pontos. Primeiro, a média é calculada a partir do valor de citações acumuladas dividido pelo valor de obras referenciadas. Segundo, o ranqueamento se dá pelo valor total de citações, todavia, tal quesito, não reflete a importância do autor para o domínio de maneira fundamental. Isso se dá pelo fato de que, o valor da média de citações por trabalho, permite observar autores que tiveram as citações recebidas concentradas em uma ou poucas obras suas em contraste com autores que receberam citações de diversas obras suas. Isso leva a constar que, no âmbito das citações há autores seminais pelo conjunto de sua obra e há outros em que uma obra sua manifesta-se como seminal. Assim, por exemplo, Mantelero teve onze obras diferentes citadas, aspecto que se observa igualmente em Floridi (22 citações de 17 obras distintas), Rodotá (citado 22 vezes com 14 obras distintas) e Boyd (citada 21 vezes com 15 obras distintas). Por outro lado, Brandeis foi citado 20 vezes com três obras e Warren foi citado 19 vezes com duas obras – em colaboração com Brandeis L. D. – evidenciando, portanto, que na constituição do domínio analisado, têm-se em igual importância, obras e autores fundamentais.

Destacam-se ainda, os campos de atuação dos dez autores com mais citações, em busca de compreender as influências contextuais exercidas para a construção do arcabouço teórico sobre privacidade e autodeterminação informativa:

- Floridi L. (1964), filósofo italiano com ampla atuação no campo da filosófica e ética da informação;
- Rodotá S. (1933-2017), jurista e político italiano, com contribuições para o campo da hermenêutica, bioética e direito;
- Boyd D.M. (1977), cientista da computação estadunidense que pesquisa tecnologia e mídias sociais;
- Warren S. D. (1852-1910) & Brandeis L.D. (1856-1941), advogados estadunidenses vanguarda na formulação da concepção da privacidade como direito;
- Doneda D. (1970-2022), jurista brasileiro com ampla atuação no campo da privacidade e proteção de dados;
- Nissenbaum H. (1954), professora de Ciência da Informação com ampla contribuição para o campo da privacidade digital, especialmente por sua teoria "contextual integrity";
- Solove D.J. (1972), professor de direito na George Washington University Law School com ampla contribuição para o campo da privacidade, principalmente pela sua proposta "Taxonomia da Privacidade"
- Acquisti A. (1972), economista estadunidense, professor de Tecnologia da Informação e Políticas Públicas na Carnegie Mellon University com atuação científica no campo da privacidade.
- Mayer-Schönberger V. (1966), professor de Governança e Regulação da Internet na Universidade de Oxford, conduz pesquisas especialmente no campo da economia digital, com obras que abordam o campo da privacidade e big data.

Fica, pois, evidente, a partir do exposto, o caráter interseccional/transdisciplinar dos temas privacidade e autodeterminação informativa, reforçando a já preconizada característica do campo que, em igual medida, absorve e aporta conhecimento de e para outras áreas, especialmente aqueles que tem como objeto a filosofia, ética, direito, tecnologias da informação, informação e economia.

Formatou: Tipo de letra: 10 pt

Formatou: Tipo de letra: 10 pt, Português (Brasil)

Formatou: Tipo de letra: 10 pt

4 Conclusão

À vista do exposto, destaca-se que, pela natureza polissêmica do termo Privacidade, a sua compreensão em um panorama global é de difícil execução, sendo necessária a adoção de ferramentais, tais como a taxonomia de Solove (2008) e as dimensões propostas por Holvast (2007), que proporcionam que seja construído um recorte terminológico a ser analisado. No presente trabalho, esse recorte se deu a partir do termo "*informational self-determination*", temática que, a partir dos resultados obtidos na base de dados Scopus, se revela relativamente jovem no campo científico – a publicação mais antiga incluída na base sobre o tema é de 1988 –, e também pouco explorada, o que resultou em um baixo número de trabalhos recuperados a partir do termo.

Todavia, ao longo das últimas décadas, o tema tem recebido atenção, especialmente a partir do ano de 2009, não havendo interrupções nas publicações e crescendo de maneira relevante ano a ano.

A partir da análise do *corpus*, constituído por 150 itens, é possível observar que a temática da autodeterminação informativa se conecta, principalmente, mas não exclusivamente, com questões relacionadas a Privacidade (contexto geral), em uma perspectiva jurídica, sobretudo em contextos informatizados. Existe um pequeno núcleo de autores que se ocupam do tema, ainda como incipiente índice de colaboração entre eles. Evidencia-se, pois, uma grande pulverização de autores, temas adjacentes, periódicos e instituições que envolvem o contexto abordado. Mais especificamente em relação às instituições, destaca-se o fato de sua maioria estar distribuída em países europeus, muito em consequência de sua origem, Alemanha, país pioneiro na questão de proteção de dados (pessoais), onde a questão remonta ao ano de 1970, quando foi editado o texto jurídico popularmente conhecido como Lei de Hesse, tido como o primeiro a legislar sobre a proteção de dados no mundo. Posteriormente, ainda no espaço europeu, o tema se desenvolveu em outros textos jurídicos sobre o tema, sendo o mais atual a GDPR, que tem forte influência no desenvolvimento de diretrizes no mundo todo.

No que diz respeito a fundamentação teórica do *corpus*, foi possível, partir das referências analisadas, observar que existe um núcleo tanto de obras quanto de autores seminais para o desenvolvimento das reflexões sobre privacidade e autodeterminação informativa, mas, por conta da característica transversal/transdisciplinar das

temáticas, o referido núcleo é de tamanho reduzido frente ao vasto universo de autores e obras referenciados.

5 Referências

- Ardrey, R. (1966). *The Territorial Imperative: A Personal Inquiry into the Animal Origins of Property and Nations* (1st ed.). Dell Publishing Co., Inc.
- Asikis, T. *et al.* (2021). How value-sensitive design can empower sustainable consumption. *Royal Society Open Science*, 8(2), 201418. <http://doi.org/10.1098/rsos.201418>.
- Asikis, T., & Pournaras, E. (2020). Optimization of privacy-utility trade-offs under informational self-determination. *Future Generation Computer Systems*, 109, 488-499. <https://doi.org/10.1016/j.future.2018.07.018>.
- Bernier, A. *et al.* (2024). Reconciling the biomedical data commons and the GDPR: Three lessons from the EUCAN ELSI collaboratory. *European Journal of Human Genetics*, 32, 69–76. <https://doi.org/10.1038/s41431-023-01403-y>.
- Bioni, A. (2019). *Proteção de dados pessoais: a função e os limites do consentimento*. Barueri: Editora Forense.
- Doneda, D. (2006). *Da privacidade à proteção de dados pessoais*. São Paulo: Editora Revista dos Tribunais.
- Hansen, M., Borcea-Pfitzmann, K., & Pfitzmann, A. (2005). PRIME – Ein europäisches Projekt für nutzerbestimmtes Identitätsmanagement (PRIME – A European Project for Privacy-Enhancing Identity Management). *it - Information Technology*, 47(6), 352-359. <https://doi.org/10.1524/itit.2005.47.6.352>
- Holvast, J. (2007). History of privacy. In K. D. Leeuw & J. Bergstra (Eds.), *The History of Information Security* (pp. 737–769). Elsevier Science B.V. <https://doi.org/10.1016/B978-044451608-4/50028-6>.
- Hornung, G., & Schnabel, C. (2009a). Data protection in Germany I: The population census decision and the right to informational self-determination. *Computer Law & Security Review*, 25(1), 84-88. <https://doi.org/10.1016/j.clsr.2008.11.002>
- Hornung, G., & Schnabel, C. (2009b). Data protection in Germany II: Recent decisions on online-searching of computers, automatic number plate recognition and data retention. *Computer Law & Security*

Formatou: Tipo de letra: 10 pt

Formatou: Tipo de letra: 10 pt, Cor do tipo de letra: Automática

Formatou: Tipo de letra: 10 pt

- Review, 25(2), 115-122.
<https://doi.org/10.1016/j.clsr.2009.02.008>.
- Judgment of 15 December 1983 (1983).
https://www.bundesverfassungsgericht.de/SharedDocs/Entscheidungen/EN/1983/12/rs_19831215_1bvr020983en.html.
- Brasil (2018). Lei No 13.709, de 14 de Agosto de 2018 Que Dispõe Sobre a Proteção de Dados Pessoais e Altera a Lei No 12.965, de 23 de Abril de 2014 (Marco Civil Da Internet)., No. 13.709.
http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709.htm.
- McDonald, A. M.; Cranor, L. F. The cost of reading privacy policies. (2008). *I/S: A Journal of Law and Policy for the Information Society*, v. 4, n. 3, p. 540-565.
- Nissenbaum, H. (2010). Privacy in context: technology, policy, and the integrity of social life. Stanford: Stanford University Press.
- Regulation (EU) 2016-679 (2016).
- Rouvroy, A., Poulet, Y. (2009). The Right to Informational Self-Determination and the Value of Self-Development: Reassessing the Importance of Privacy for Democracy. In: Gutwirth, S., Poulet, Y., De Hert, P., de Terwagne, C., Nouwt, S. (eds) Reinventing Data Protection?. Springer, Dordrecht.
https://doi.org/10.1007/978-1-4020-9498-9_2.
- Schickhardt, C., Fleischer, H., & Winkler, E. (2020). Do patients and research subjects have a right to receive their genomic raw data? An ethical and legal analysis. *BMC Medical Ethics*, 21, Article 10.1186/s12910-020-0446-y.
- Solove, D. J. (2008). Understanding privacy. Cambridge: Harvard University Press.
- T. Hall, E. (1966). The Hidden Dimension (1st ed.). Anchor Books, Doubleday.
- Warren, S. D., Brandeis, L. D. (2008) The right to privacy. *Harvard Law Review*, v. 4, n. 5, p. 193-220, 1890.2.
- Rodotà, S. (2008) A vida na sociedade da vigilância: a privacidade hoje. São Paulo: Editora Renovar.
- Westin, A. (1967). Privacy and Freedom. Ig Publishing.
- Zuboff, S. (2019). The age of surveillance capitalism: the fight for a human future at the new frontier of power. New York: PublicAffairs.

6 Notas

¹A definição do critério “quatro ou mais autores” deve-se ao fato de que tanto a norma internacional ISO, quanto a norma brasileira da ABNT preconizam a citação de obras com mais de três autores a partir da citação do primeiro autor seguido da expressão *et al.*

Formatou: Tipo de letra: (predefinido) +Corpo (Calibri), 9 pt

Formatou: Tipo de letra: (predefinido) +Corpo (Calibri), 9 pt

Formatou: Tipo de letra: (predefinido) +Corpo (Calibri), 9 pt