**Dados sensíveis no contexto dos dados de pesquisa:** um olhar na perspectiva da Ciência da Informação

Richele Grenge Vignoli1, Fernando Luiz Vechiato2

1 0000-0003-1550-5258 ORCID + Universidade Estadual Paulista “Júlio de Mesquita Filho” (Unesp) - Campus de Marília- SP. rivignoli@gmail.com.

2 [0000-0002-4157-740X](https://orcid.org/0000-0002-4157-740X) ORCID + Universidade Federal do Rio Grande do Norte (UFRN) / Universidade Estadual Paulista “Júlio de Mesquita Filho” (Unesp), Natal / Marília, Rio Grande do Norte/São Paulo. vechiato2008@gmail.com.

**Tipo de trabalho**: Comunicação Oral

**Resumo**

Os dados de pesquisa têm se apresentado com alto teor de investigação na esfera acadêmico-científica na Europa, nos EUA e no Brasil. No entanto, os dados sensíveis, neste contexto, são ainda pouco discutidos na literatura científica. Da forma como se apresentam, os dados sensíveis são dados sigilosos e altamente confidenciais, que recebem um invólucro de preocupações legais e éticas quanto ao vazamento, preservação e segurança. Sob este ponto de vista, a pesquisa objetivou apresentar discussões iniciais a respeito dos dados sensíveis em pesquisa no contexto da Ciência da Informação, além de demonstrar os cuidados relacionados ao sigilo, ética e segurança que os pesquisadores e os profissionais da informação precisam ter ao lidar com estes tipos de dados. A preocupação com as questões sociais que envolvem o sigilo dos dados sensíveis de um sujeito quando da coleta de dados de pesquisas científicas também foram objetos de discussão neste trabalho. Para alcançar os objetivos propostos, uma pesquisa bibliográfica foi realizada no Portal Capes e contemplou principalmente as bases de dados *Web of Science e* LISA. Como critérios para refinamento conceitual, foram utilizados filtros para inserção dos idiomas inglês, espanhol e português, sem período pré-estipulado para a consulta já que se trata de uma temática recente na Ciência da Informação e na literatura como um todo. Os resultados da pesquisa, em que foram utilizadas principalmente as palavras-chave *Data Research*, *Sensitive Data* e *Sensitive Research Data,* não abarcaram a respeito dos dados sensíveis especificamente em pesquisa científica, demonstrando que a literatura é mais focada para discussões sobre a privacidade, manuseio ético e formas pelas quais esses dados podem ser preservados e assegurados por técnicas específicas. A pesquisa demonstrou discussões iniciais a respeito do conceito de dados sensíveis sob a ótica da Ciência da Informação e em pesquisa científica. A conscientização e o comprometimento dos pesquisadores e profissionais da área, com o sigilo, ética e segurança dos dados coletados em pesquisa e que são sensíveis, são pontos fundamentais a serem esclarecidos na gestão desses dados. Ressalta-se ainda que a responsabilidade dos Dados Sensíveis de uma pesquisa deverá envolver o pesquisador e o gestor (bibliotecário) que, porventura, manuseará e organizará os dados para que sejam reutilizados e/ou compartilhados sem quaisquer danos legais e/ou sociais envolvidos no processo.

**Palavras-chave:** Dados Sensíveis. Dados de Pesquisa. Ciência da Informação.

**1 Introdução**

A preocupação com os *Data Research* ou Dados de Pesquisa tem sido crescente nos cenários acadêmico-científicos, principalmente na Europa, EUA e mais recentemente no Brasil. A Curadoria de Dados (*Data Curator*), a e-Science, os *Datasets* (Repositório de Dados - de pesquisa ou de Dados Científicos) e principalmente a Gestão de Dados de Pesquisa em um contexto de *Big Data* e *Linked Data*, demonstram que os Dados em suas diversas facetas de investigação são tendência hodierna nos construtos atuais da Ciência da Informação (CI).

Em concomitância com a esfera dos estudos acerca dos dados, a privacidade, a ética e a segurança dos dados nos sistemas de informação que objetivam permitir acesso e reuso aos Dados de Pesquisa, há a sensibilidade dos dados que pulsa por aprofundamento e discussão. Entretanto, observa-se na literatura científica, sobretudo na CI, que aportes teóricos consistentes têm sido necessários para a discussão das questões que envolvem a sensibilidade dos Dados - Dados Sensíveis, em meio a tantas outras questões que permeiam os avanços científicos, teóricos e aplicados aos dados.

Da forma que se apresentam, os Dados Sensíveis são dados sigilosos e altamente confidenciais, que recebem um invólucro de preocupações legais e éticas quanto ao vazamento, preservação e segurança. Todo o arcabouço de ações e preocupações que envolvem o acesso, reuso e gestão dos Dados de Pesquisa e dos Dados Sensíveis são problemas que a CI pode e precisa responder.

Assim, parece ocorrer uma reordenação nos processos científicos que, por meio da validação e reutilização dos Dados de Pesquisa, envolverá a ampla compreensão dos Profissionais da Informação, do Bibliotecários de Dados (Rice & Southall, 2016) ou Cientistas de Dados (Sayão & Sales, 2013). Estes profissionais e também os pesquisadores da CI podem estar diante da revalorização dos dados (e mais especificamente dos Dados de Pesquisa), que há algum tempo faziam pano de fundo na própria CI, que tem priorizado mais especialmente as discussões acerca da informação e do conhecimento nas últimas décadas.

De toda forma, esta pesquisa teve como objetivo principal apresentar discussões preliminares a respeito dos Dados Sensíveis e Dados Sensíveis em pesquisa, além de demonstrar os cuidados relacionados ao sigilo, ética e segurança que os pesquisadores e profissionais da Informação precisam ter ao lidar com estes tipos de dados. As questões sociais que envolvem o sigilo dos dados sensíveis de um sujeito quando da coleta de dados de pesquisas científicas também foram objetos de discussão neste trabalho. Assim, espera-se contribuir para o esclarecimento dos Dados Sensíveis e Dados Sensíveis em pesquisa científica e singularmente nos espectros pós-modernos da CI e sua emergente e profícua relação com os dados.

**2 Procedimentos Metodológicos**

Para alcançar os objetivos propostos no referido estudo, uma pesquisa bibliográfica foi realizada no Portal Capes e contemplou principalmente as bases de dados *Web of Science e* LISA. Como critérios para refinamento conceitual, foram utilizados filtros para inserção dos idiomas de inglês, espanhol e português, sem período pré-estipulado para a consulta já que se trata de uma temática recente na CI e na literatura como um todo.

Os resultados da pesquisa, em que foram utilizadas principalmente as palavras-chave *Data Research*, *Sensitive Data* e *Sensitive Research Data,* não abarcaram a respeito dos dados sensíveis especificamente em pesquisa científica, demonstrando que a literatura é mais focada para discussões sobre a privacidade, manuseio ético e formas pelas quais esses dados podem ser preservados e assegurados por técnicas específicas. Reitera-se a importância dos estudos acerca dos Dados Sensíveis sob aspectos relacionados aos Dados de Pesquisa na contemporaneidade da CI.

**3 Dados Sensíveis**: visão geral

De acordo com preceitos da Comissão Europeia (2018), os Dados Sensíveis são dados sobre saúde, origem racial ou étnica, opiniões políticas e orientação sexual e que necessitam de proteção especial. Para tanto, acentua a Comissão Europeia (2018), “Estes dados só podem ser recolhidos e utilizados em condições específicas, por exemplo caso tenha dado o seu consentimento explícito ou caso a legislação nacional o permita.” (p. 02).

A Lei Geral de Proteção de Dados (LGDP) n.º 13.709 de 14 de agosto 2018, no cenário brasileiro, por sua vez, utiliza o termo Dados Pessoais Sensíveis para representar esses tipos de dados, e os conceitua como: “Dados pessoais sobre a origem racial ou étnica, as convicções religiosas, ou opiniões políticas, a filiação a sindicatos ou a organizações de caráter religioso filosófico ou político, dados referentes à saúde ou à vida sexual, e dados genéticos ou biométricos.” (Brasil, 2018, p. 02). Nota-se que, para a LGDP, Dados Sensíveis são provenientes de Dados Pessoais, conforme demonstra a Figura 1 que segue:

Figura 1 - Dados Pessoais x Dados Pessoais Sensíveis



Fonte: Renascença (2018). Recuperado de https://rr.sapo.pt/privacidade-online/cap1.aspx.

Como visto na Figura 1, existe a iminência de Dados Sensíveis em Dados Pessoais, no entanto nem todo Dado Pessoal é sensível e, tampouco, nem todo Dado Sensível é pessoal. Nesta relação, se faz necessário o esclarecimento de que Dados Sensíveis tanto ocorrem por meio de dados de pessoas naturais, quanto de pessoais jurídicas, assim como em pesquisas científicas, de mercado, entre outras. Não obstante, todo Dado Pessoal passível de algum tipo de situação vexatória ou preconceituosa, e ainda que infrinja as leis de proteção a integridade de um indivíduo, passa a ser atribuído como Dado Sensível. Outros dados pessoais, como valor de salários, notas de alunos, faturas de cartão de crédito, dados médicos, acordos matrimoniais, Declaração de Imposto de Renda, entre outros tipos de dados, também podem ser considerados como Dados Sensíveis.

Briney (2015) evidencia que Dados Governamentais que podem causar grandes danos à população também podem ser compreendidos como Dados Sensíveis e demonstra como exemplo os dados acerca do início da epidemia da Gripe Aviária (H5n1) no mundo, que continha dados sensíveis até a divulgação da doença. Tal fato demonstra a presença de Dados Sensíveis em outras esferas, que não somente as pessoais, e pondera ainda que sob devidas proporções e cuidados legais, a sua excepcional divulgação pode ocorrer. Entretanto, a máxima recorrente a ser seguida é de que Dados Sensíveis não devem ser divulgados.

Outro tipo de Dado Sensível é encontrado em patentes e outros objetos com propriedade intelectual. Estes dados são confidenciais sobretudo porque são proprietários, pontua Briney (2015). Muitas empresas possuem Dados Sensíveis e por isso investem quantias consideráveis para protegerem suas informações, mesmo que muitos vazamentos e ciber-ataques ocorram. Para Briney (2015), os dados de patentes, assim como dados empresariais, não podem ser classificados como dados com proteção legal ou pessoal, mas mesmo assim são Dados Sensíveis.

Sob este ponto de vista, e em corroboração com o autor, a National Science Board (NSB) - (2005) elenca outros espécimes de dados com acesso limitado ou totalmente restrito que podem vir a se tornar Dados Sensíveis, como: segurança nacional: dados de inteligência, atividades militares, decisões políticas, entre outros; segredos comerciais ou de empresa; proteção de espécies raras, ameaçadas ou em perigo; ou processos legais: dados em consideração em ações judiciais.

 Mais especificamente no contexto acadêmico-científico, os Dados de Pesquisa segundo a NSB (2005, p. 13): “(...) são definidos como registros factuais (escores numéricos, registros textuais, imagens e sons) utilizados como fontes primárias de pesquisa científica, e que são comumente aceitos na comunidade científica como necessário para validar resultados da pesquisa. Um conjunto de dados de pesquisa constitui uma representação parcial sistemática do assunto sendo investigado.”

De acordo com a NSB (2005), a aceitabilidade dos dados primários de uma pesquisa são inquestionáveis em sua comunidade, pois trata-se dos primeiros resultados e resultantes de uma pesquisa, o que os torna ainda mais suscetíveis ao reuso por outros pesquisadores.

Briney (2015) expõe que dados do sujeito humano são sensíveis por razões éticas. Segundo o autor, o pesquisador tem o dever ético em assegurar a privacidade dos dados fornecidos por seres humanos e confiáveis a ele durante a pesquisa. Briney (2015) não desassocia a sensibilidade dos dados da ética e afirma que a ética em pesquisa não significa apenas manter os dados seguros. É necessário que os pesquisadores mantenham os dados confidenciais, que obtenham consentimento informado e que evitem danos e conflitos de interesse com os participantes da pesquisa. Eis que esta relação, sujeito-pesquisador, deverá ser pautada em confiança mútua. A relação entre pesquisador-profissional (da informação) que irá tratar estes dados, também deve seguir os mesmos princípios ao ponto que discriminações ou problemas sociais não ocorram aos envolvidos.

Assim, Briney (2015) alerta que, ao lidar com a confidencialidade dos dados, ou ainda com Dados Sensíveis em pesquisa, três aspectos devem ser precocemente pensados:

1. determinar se os dados são realmente confidenciais;

2. não coletar dados sensíveis se não houver necessidade;

3. desenvolver um plano de segurança para os dados com a ajuda de especialistas.

Briney (2015) destaca ações a serem desenvolvidas em pesquisas e pelo pesquisador durante o início, desenvolvimento e conclusão de todo o estudo. Vale destacar que tais ações necessitam ser realizadas antes do acesso e do reuso dos dados por outros pesquisadores, assim como antecipadamente a gestão desses dados por Profissionais da Informação, por exemplo.

 Não obstante, Dados Sensíveis não são prioridade do contexto científico, mas a sua existência no contexto de coleta de dados com a participação de humanos não deve ser ignorada, já que passa a ser um ambiente em potencial para a construção de Dados Sensíveis. A de se pesar que quando dados forem confidenciais ou sigilosos e não puderem ser divulgados principalmente por questões legais e éticas, tem-se a manifestação de Dados Sensíveis. Entretanto, apesar de os Dados Sensíveis estarem presentes sob diversas maneiras, situações ou ambientes, interessou a este estudo a disposição e construção destes dados no contexto da pesquisa científica.

Para além disso, Dados Sensíveis são protegidos por leis em território nacional e internacional. A maioria das leis determinam como proteger e zelar por Dados Pessoais e Sensíveis. O problema é que cada país possui uma legislação e, por isso, não é possível haver um padrão internacional para lidar com Dados Sensíveis. Briney (2015) explica que a União Europeia é responsável pelas leis mais consistentes de proteção aos Dados Pessoais que existem desde 1995, fato comprovado nas recentes atualizações do Regulamento Geral de Proteção de Dados (RGPD) publicado em 2018 pela Comissão Europeia. Já nos EUA, as leis ou somente regulamentos são separadas por assunto com especificidade em relação à área, como da saúde, a respeito de registros escolares, entre outros temas.

Por fim, acredita-se que a responsabilidade pela privacidade e confidencialidade dos dados é cabida ao pesquisador, assim como é de sua responsabilidade a sua pesquisa, seus participantes, seus conteúdos e formas pelas quais seus dados poderão ser armazenados, divulgados e reutilizados.

Como visto, os Dados Sensíveis podem ser muito variados e abrangentes em diversas áreas do conhecimento e da própria sociedade, assim como em pesquisa científica e seus contextos sociais. O Quadro 1 foi elaborado como via a sistematizar os conceitos apresentados a fim de harmonizá-los no contexto de suas aplicações e na Ciência da Informação.

Quadro 1 – Sumarização dos conceitos apresentados

|  |  |  |
| --- | --- | --- |
| **Tipos de Dados** | **Exemplos e contexto** | **Definição** |
| Dados Pessoais | CPF, RG, Nome próprio | Dados pertencentes à pessoa natural de forma unívoca. |
| Dados de Pesquisa | Dados coletados em pesquisa | Dados provenientes e resultantes de pesquisa científica. |
| Dados Sensíveis | Dados pessoais com traços de personalidade, gostos, etc;Dados Governamentais;Dados empresariais;Dados de propriedade intelectual | Dados íntimos, secretos ou sigilosos de pessoa natural ou jurídica. |
| Dados Pessoais Sensíveis | Dados pessoais como comportamentos, preferências, ações | Dados pessoais íntimos, secretos ou sigilosos exclusivos a uma pessoa natural. |
| Dados Sensíveis em pesquisa | Dados coletados por meio da participação de seres humanos | Dados sigilosos coletados em pesquisa científica, essencialmente os fornecidos por seres humanos. |

Fonte: autores.

O Quadro 1 demonstra de modo sumário e ainda preliminar os principais conceitos relacionados aos dados e suas ramificações. Cabe ressaltar que apesar de limítrofes, os conceitos são díspares e requerem do Profissional da Informação e pesquisador atenção redobrada ao interpretá-los e manuseá-los sempre em prol da comunidade a qual se dedicam.

**4 Considerações Finais**

A pesquisa demonstrou, por meio da literatura científica, discussões iniciais a respeito do conceito de Dados Sensíveis sob o olhar da CI e em pesquisa científica. A conscientização e comprometimento dos pesquisadores da área, com o sigilo, ética e segurança dos dados coletados em pesquisa e que são sensíveis, são pontos fundamentais a serem esclarecidos na gestão desses dados. É neste momento que a importância do Profissional da Informação é ainda mais enaltecida, já que ele deve possuir habilidades para o manuseio consciente deste tipo de dado, principalmente devido a implicações sociais que ocorreriam no seu vazamento.

Para tanto, espera-se que estas discussões iniciais acerca dos Dados Sensíveis em pesquisa prospectem atenção especial ao sigilo e as questões sociais que envolvem a disponibilização de dados que se compartilhados aquém de leis e normas sociais poderão afetar demasiadamente a privacidade de sujeitos que voluntariamente participaram da coleta de dados em pesquisa científicas.

Esta pesquisa, ao familiarizar o leitor com a realidade dos Dados Sensíveis em especial no aspecto científico por meio dos Dados de Pesquisa em pesquisa, projetou intervir no esclarecimento dos danos sociais que o vazamento de Dados Sensíveis poderá ocasionar ao sujeito em sociedade. Ressalta-se ainda que a responsabilidade dos Dados Sensíveis de uma pesquisa deverá envolver o pesquisador e o gestor (bibliotecário, por exemplo) que, porventura, manuseará e organizará os dados para que sejam reutilizados e/ou compartilhados sem quaisquer danos legais e/ou sociais aos envolvidos no processo.

**Referências**

Affonso, E. P. (2018). *A insciência do usuário na fase de coleta de dados:* *Privacidade em foco.* Tese de doutoramento apresentada à Universidade Estadual Paulista (Unesp), Faculdade de Filosofia e Ciências, Marília - SP.

Brasil. (2016). Câmara dos Deputados. *Projeto de Lei nº 5.276, de 2016.* Dispõe sobre o tratamento de dados pessoais para a garantia do livre desenvolvimento da personalidade e da dignidade da pessoa natural. Brasília, Câmara dos Deputados.

Brasil. (2018). *Lei n° 13.709, de 15 de agosto de 2018.* Dispõe sobre a proteção de dados pessoais e altera Lei n° 12.965 de 23 de abril de 2014. [Marco Civil da Internet]. Brasília, Presidência da República.

Briney, K. (2015). *Data Management for researchers:* *Organize, maintain and share your data for research success*. Exeter, UK: Pelagic Publishing.

Comissão Europeia. (2018). *Assuma o controle de seus dados:* um guia do cidadão para a proteção de dados na UE. Luxemburgo: Serviço das Publicações da União Europeia. ISBN 978-92-79-77581-9. doi:10.2838/00666. DS-01-18-022-PT-N.

National Science Board (NSB). (2005). **Long-lived digital data collections**: enabling research and education in the 21st century. National Science Foundation, Sept. 2005.

Rice, R.; & Southall, J. (2016). Buinding a data collection. In A. R. Rice. & B. J. Southall. (Eds.)., *The Data Librarian´s handobook* (pp. 53-66). Londres: Facet Publishing.

Sayão, L. F.; & Sales, L. (2013). Dados de pesquisa: contribuição para o estabelecimento de um modelo de curadoria digital para o país. *Tendências da Pesquisa Brasileira em Ciência da Informação*, 6, (1). Recuperado em 2 janeiro, 2019, de http://www.brapci.inf.br/index.php/article/view/0000014157/0f3127bfd6475f45e6950e1e0da20371.

Sayão, L. F.; & Sales, L. (2015). *Guia de Gestão de Dados de Pesquisa para bibliotecários e pesquisadores*. Rio de Janeiro: CNEN/IEN, 2015. 90 p. Recuperado em 12 janeiro, 2019 de http://www.cnen.gov.br/component/content/article/75-cin/material-didatico-cnen/160-guia-de-gestao-de-dados-de-pesquisa.