



Mesa Redonda: Atuação do profissional de Estatística na área da saúde

Tema: Uso de dados pessoais na pesquisa em saúde: O que muda com a LGPD e qual a contribuição da Estatística nesta discussão?

Nívea Silva

DEST – UFBA

Pesquisadora colaboradora do CIDACS-Fiocruz-BA

Contextualização - LGPD (13.709/2018)



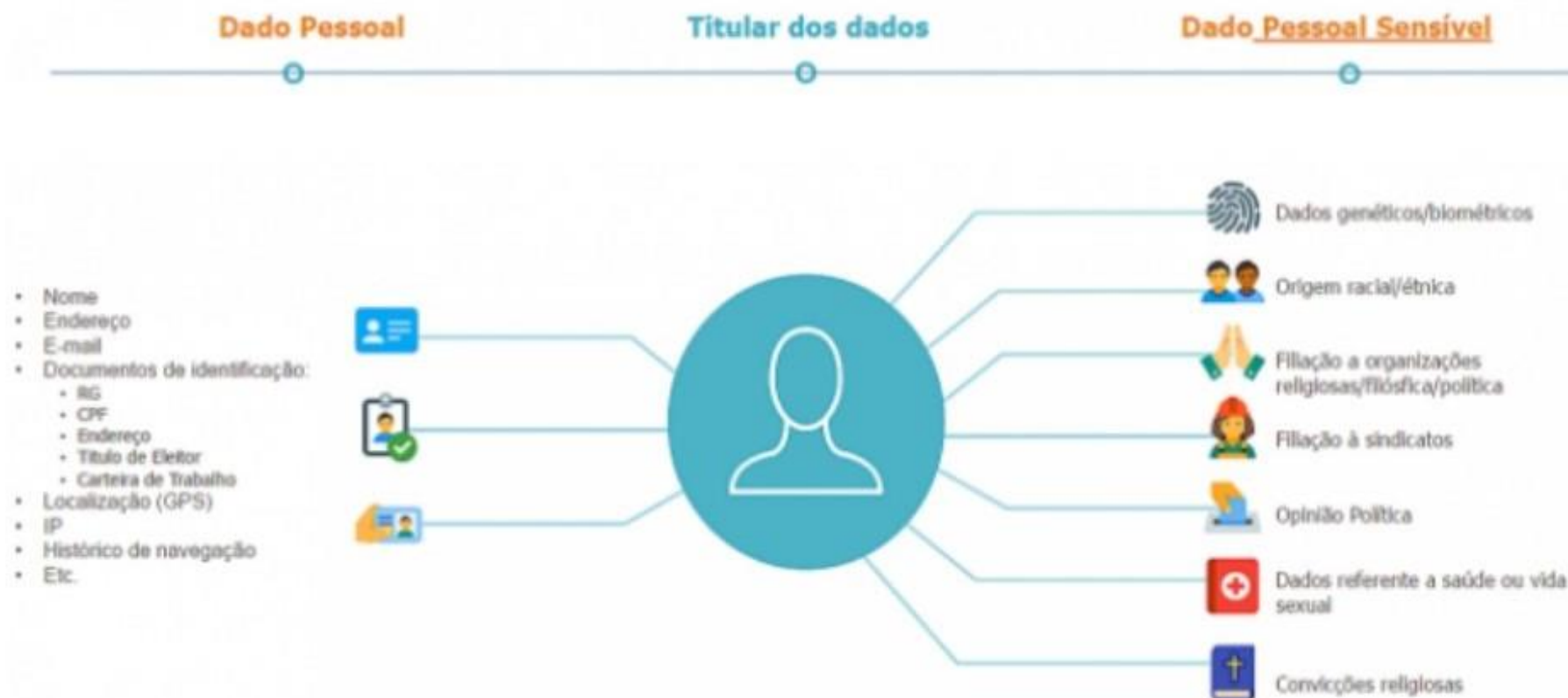
- Lei geral de proteção de dados brasileira aprovada em 08/2018 e vigente desde 09/2020.
- Dispõe sobre o tratamento de **dados pessoais** por pessoas naturais ou jurídicas, de direito público ou privado, com o objetivo de proteger os direitos fundamentais de liberdade e de privacidade da pessoa natural.
- Aplica-se a qualquer operação de tratamento, desde que esta ocorra em território nacional (artigo 3º).
- Dentre outras coisas, a lei não se aplica ao **tratamento*** de dados pessoais realizados para fins exclusivamente (artigo 4º):
 - i) particulares e não econômicos, por pessoa natural;
 - ii) jornalísticos e artísticos ou **acadêmicos** (aplica-se a esta hipótese os arts. 7º e 11º).

* toda operação realizada com dados pessoais, como as que se referem à coleta, produção, recepção, classificação, utilização, acesso, reprodução, transmissão, distribuição, processamento, arquivamento, armazenamento, eliminação, avaliação ou controle da informação, modificação, comunicação, transferência, difusão ou extração.

Terminologias na LGPD

- **Dado pessoal** (artigo 5º, I): Toda informação relacionada à pessoa natural identificada ou identificável.

- **Dado pessoal sensível:** Informação vinculada a uma pessoa natural, que pode lhe causar atos discriminatórios.



O que muda na pesquisa em saúde após a LGPD?

- O artigo 7º define as hipóteses em que o tratamento de dados pessoais pode ser realizado e, a hipótese IV fala deste tratamento para realização de estudos por órgão de pesquisa, desde que garantida, sempre que possível, a **anonimização dos dados pessoais**;
- O artigo 11º fala das hipóteses sobre o tratamento de **dados pessoais sensíveis** e diz que "*o mesmo somente poderá ocorrer, na ausência de consentimento do titular do dado, quando for imprescindível a realização de estudos por órgão de pesquisa, garantida, sempre que possível, **a anonimização dos dados pessoais sensíveis***".
- O artigo 13º menciona que: "***Na realização de estudos em saúde pública, os órgãos de pesquisa poderão ter acesso a bases de dados pessoais, que serão tratados exclusivamente dentro do órgão e estritamente para a finalidade de realização de estudos e pesquisas e mantidos em ambiente controlado e seguro, ... e que incluam, sempre que possível, a **anonimização** ou **pseudonimização** dos dados ...***"

Qual a contribuição da Estatística nesta discussão?

- Segundo a LGPD, os órgãos de pesquisa terão acesso a bases de dados que contêm informações pessoais se for garantido, dentre outras coisas, a **anonimização** ou **pseudonimização** dos dados;
- Do ponto de vista do **tratamento** destas informações, a Estatística se insere em várias etapas do processo de preservação da privacidade, ao qual podemos pensá-lo como uma espécie de "protocolo", que deve definir coisas do tipo:
 - Quais variáveis serão anonimizadas?
 - Quais técnicas de anonimização serão utilizadas em cada variável?
 - Qual limiar de risco adotar?
- A partir do protocolo pretende-se desvincular os dados pessoais do seu titular.

- Supressão
 - Variável
 - Registro
 - Valor da variável em um registro
- Generalização
- Randomização
 - Adição de ruído
 - Permutação
- Agregação

Qual o desafio? Manter a qualidade dos dados para análises futuras, garantindo a preservação da privacidade.

Qual a contribuição da Estatística nesta discussão?

- **Desafio:** Maximizar simultaneamente a privacidade (risco) e a utilidade dos dados

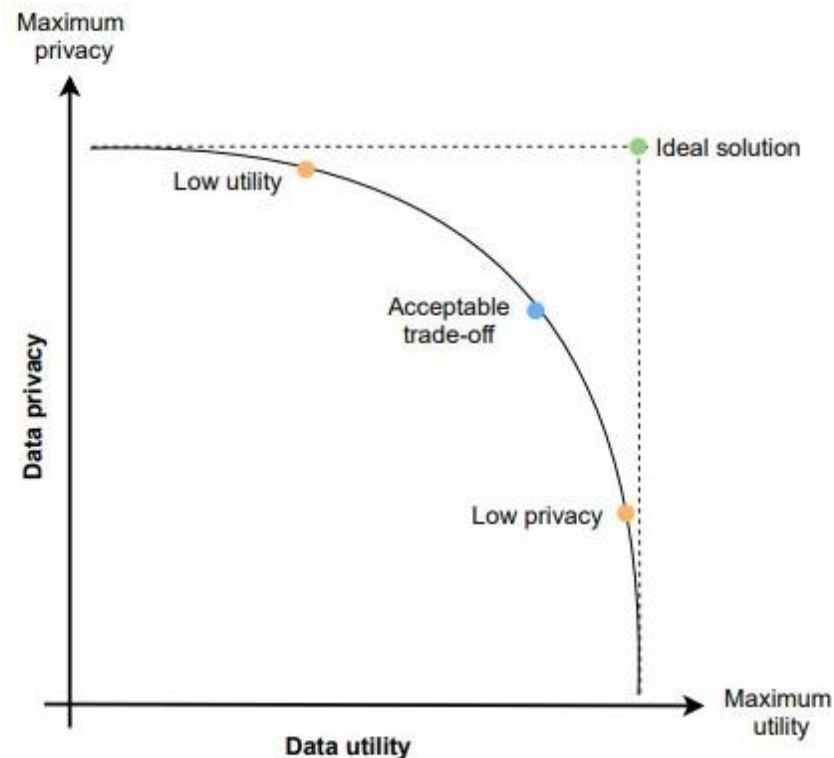


Fig. 1. Trade-off between privacy level and utility level of data.

Fonte: imagem extraída de Carvalho et al, 2022.

80

Health Informatics Meets eHealth
G. Schreier and D. Hayn (Eds.)
© 2018 The authors and IOS Press.

*This article is published online with Open Access by IOS Press and distributed under the terms of the Creative Commons Attribution Non-Commercial License 4.0 (CC BY-NC 4.0).
doi:10.3233/978-1-61499-858-7-80*

Effectiveness of Anonymization Methods in Preserving Patients' Privacy: A Systematic Literature Review

Mostafa LANGARIZADEH^a, Azam OROOJI^{a,1}, Abbas SHEIKHTAHERI^a

^a Department of Health Information Management, School of Health Management and Information Sciences, Iran University of Medical Sciences, Tehran, Iran

Statistical Data Privacy: A Song of Privacy and Utility

Aleksandra Slavković¹ and Jeremy Seeman²

¹Department of Statistics, The Pennsylvania State University, University Park, PA, 16802; email: abs12@psu.edu

²Department of Statistics, The Pennsylvania State University, University Park, PA, 16802, email: jhs5496@psu.edu. ORCID: 0000-0003-3526-3209

Fonte: <https://doi.org/10.48550/arXiv.2205.03336>

Qual a contribuição da Estatística nesta discussão?

Implications of Data Anonymization on the Statistical Evidence of Disparity

Heng Xu , Nan Zhang 

Published Online: 4 Jun 2021

Abstract

Research and practical development of data-anonymization techniques have proliferated in recent years. Yet, limited attention has been paid to examine the potentially disparate impact of privacy protection on underprivileged subpopulations. This study is one of the first attempts to examine the extent to which data anonymization could mask the gross statistical disparities between subpopulations in the data. We first describe two common mechanisms of data anonymization and two prevalent types of statistical evidence for disparity. Then, we develop conceptual foundation and mathematical formalism demonstrating that the two data-anonymization mechanisms have distinctive impacts on the identifiability of disparity, which also varies based on its statistical operationalization. ~~After validating our findings with empirical evidence, we discuss the business and~~

Lee et al. *BMC Medical Informatics and Decision Making* (2017) 17:104
DOI 10.1186/s12911-017-0499-0

BMC Medical Informatics and
Decision Making

RESEARCH ARTICLE

Open Access

Utility-preserving anonymization for health data publishing

Hyukki Lee¹, Soohyung Kim², Jong Wook Kim³ and Yon Dohn Chung^{1*} 



CHANCE

ISSN: 0933-2480 (Print) 1867-2280 (Online) Journal homepage: <https://www.tandfonline.com/loi/uha20>



SPECIAL ISSUE: A New Generation of Statisticians Tackles Data Privacy

Fang Liu, Saki Kinney & Aleksandra (Seša) Slavkovic´

To cite this article: Fang Liu, Saki Kinney & Aleksandra (Seša) Slavkovic´ (2020) SPECIAL ISSUE: A New Generation of Statisticians Tackles Data Privacy, CHANCE, 33:4, 4-5, DOI: [10.1080/09332480.2020.1847945](https://doi.org/10.1080/09332480.2020.1847945)

Survey on Privacy-Preserving Techniques for Data Publishing

TÂNIA CARVALHO, DCC - Faculty of Sciences, University of Porto, Portugal

NUNO MONIZ, INESC TEC / University of Porto, Portugal

PEDRO FARIA, TekPrivacy, Portugal

LUÍS ANTUNES, DCC - Faculty of Sciences, University of Porto, Portugal

Fonte: <https://doi.org/10.1145/1122445.1122456> (Arxiv)

Obrigada!

nbsilva@ufba.br