FGV DIREITO SP
MESTRADO PROFISSIONAL

NOVAS FRONTEIRAS DA PRIVACIDADE: *BIG DATA* E PARÂMETROS DO USO DE DADOS MÉDICOS FRENTE AOS REGULAMENTOS DE PROTEÇÃO

Eduardo Luiz de Oliveira Filho

Projeto de pesquisa apresentado ao
Mestrado Profissional da FGV Direito SP.
Versão de 13/10/2018
Orientador: Prof. Doutor Emerson Ribeiro Fabiani

1. **Tema, contexto, objetivos e delimitação de escopo**

O avanço na exploração de sistemas baseados na rede mundial de computadores permite aos serviços de saúde alcançar um novo patamar de evolução. A informatização dos dados proporciona que médicos e pesquisadores tenham acesso às informações de maneira completa e rápida.

A queda na barreira do acesso é apenas o primeiro passo da nova era digital. Sistemas inteligentes também conhecidos como "inteligência artificial" proporcionam os mais novos avanços ao campo de informações médicas. Dados obtidos através de padrões de consumos, redes sociais, inclusive gadgets, condensam informações que podem ser utilizadas para o desenvolvimento de tecnologias para fins de diagnósticos e prevenção de doenças.

Contudo, uma das principais barreiras para o desenvolvimento baseado em utilização de dados pessoais é a privacidade. Como regra, todos os dados pessoais, inclusive aqueles difundidos pela rede mundial de computadores, são, de alguma forma, protegidos pelo direito à privacidade. Quando tratamos de dados médicos, o dever de sigilo é ainda maior, informações sensíveis e interessante à intimidade da pessoa humana devem ser ainda mais resguardados.

Essa limitação de acesso, imposta como regra geral, pode afetar o desenvolvimento de sistemas especialmente baseados em inteligência artificial. Nesse tipo de sistema o processo de aprendizado é indutivo e essencial ao seu desenvolvimento. Assim, o processo pode ser compreendido como uma procura por descrições gerais plausíveis (afirmações indutivas) que expliquem os dados de entrada fornecidos (dados pessoais) que sejam úteis para prever novos dados.

Para que um programa de computador formule tais descrições (resultado processado), uma linguagem apropriada deve ser usada. Em razão das muitas informações, o programa pode determinar variáveis e aprender com as informações prestadas, sendo essa a característica essencial da inteligência artificial (DIETTERICH, 1983).

A limitação de acesso aos dados pessoais em caráter absoluto reduziria a capacidade de exploração da tecnologia, impossibilitando sua correta programação, pesquisa e desenvolvimento.

Nos últimos anos, o acesso a dados pessoais é um dos assuntos cercado de discussões pela comunidade mundial. O estopim do movimento ocorreu com as acusações de monitoramento e acesso indevido a dados pessoais pelo governo americano, acusações levantadas pelo ex-analista da NSA Edward Snowden, ocasionando a necessária regulamentação de aceso e uso.

Enquanto alguns países do mundo informam a seus cidadãos que o sigilo dos dados pessoais é um aspecto dispensável da cidadania, induzindo o pensamento de que a publicidade ocorre em benefício geral (GREENWALD, 2014), outros compreendem a necessidade de restrição e procuram a regulamentação.

Por esses termos, a União Europeia através do GDPR - *General Data Protection Regulation* – garantiu maior controle dos dados pessoais, ditando normas para o intercambio e a armazenagem das informações. Corroborando com a corrente protetiva,

em recente manifestação legislativa, a Lei Federal 13.709 de 14 de agosto de 2018 também disciplinou a proteção de dados pessoais em âmbito nacional. O direito norte americano, opta pela regulamentação segmentada, sendo que, dentre outros instrumentos, o US Health Insurance Portability and Accountability Act (HIPAA), garante a proteção dos dados médicos.

Falando especificamente dos dados médicos, os avanços tecnológicos proporcionados pela internet elevaram a um novo patamar a pesquisa e desenvolvimento. Veja que anteriormente à revolução digital, informações médicas eram mantidas em arquivos físicos e pessoais, o que tornava sua análise uma tarefa homérica. Agora, com a manutenção atual em formato *big data*, comparações, estudos e desenvolvimentos podem ser facilmente realizados, permitindo uma compilação rápida e eficaz.

A tecnologia *big data* representa a capacidade de processar uma grande quantidade de informações complexas para tomar decisões mais bem informadas. Essas decisões podem ser relativas a negócios ou a caminhos de pesquisas. É uma condensação de informação que não seria possível sem a existência de supercomputadores e redes *cloud* de armazenamentos (SANJEEV, 2013). Entre as recentes tecnologias, ressaltamos os sistemas especialmente destinados à saúde Amazon AWS, Microsoft Azure e IBM Watson Health.

É pelo panorama apresentado que a proposta do trabalho é de analisar como o acesso a dados médicos disponíveis em sistemas de *big data* podem ser utilizados no desenvolvimento de novas tecnologias, em consonância e com fundamento nas leis de proteção de dados pessoais e ao direito à privacidade. Desse modo, o trabalho procura resolver os problemas práticos que desenvolvedores e pesquisadores enfrentarão para a soluções de interpretação e dos limites impostos pelas legislações, especialmente a europeia, brasileira e americana. Por fim, serão consolidadas propostas para aprimoramento da legislação e conduta dos utilizadores das informações.

2. **Justificativa da relevância prática e do potencial inovador**

A privacidade em seu estado embrionário era compreendida como uma das vertentes da personalidade, especialmente conceituada como uma proteção à boa fama contra “boatos” publicados em jornais (SCHWARTZ, 2011). Somente com o conceito atual de privacidade que os dados pessoais foram compreendidos como merecedores de tutela.

Com a dependência tecnológica atual, e consequente alterações de percepção de proteção, a necessidade de regulamentação de uso e acesso dos dados pessoais aflorou. O advento da internet e a evolução social criada pela rede, permitiram que dados pessoais circulassem irrestritamente. A circulação irrestrita causou abusos por parte de particulares e de Estados, gerando a necessidade de regulamentação. O uso dos dados pessoais, compreendida sua capacidade para o desenvolvimento e a evolução tecnológica, é um importante instrumento à disposição.

Dentre as informações disponíveis, duas são as principais diferenças. Os dados poderão ser genéricos (que não refletem a individualidade) ou individuais, denominados *Personally Identifiable Information - PII*.

Algumas perguntas serão resolvidas pelo trabalho proposto, pois para alguns estatutos, especialmente as disposições americanas, somente haverá violação à privacidade quando tratarmos de informações PII, ou seja, quando se tratarem de informações pessoalmente identificáveis (PII), informações que podem identificar sua fonte e ferir a individualidade.

Cabe lembrar que para a análise de dados médicos, a generalidade não é interessante, visto que impede estudos aprofundados e afasta a essência dos dados, exatamente a individualidade.

Tratando da relação de acesso a dados médicos, especificamente, a regulamentação e dever de proteção é ainda mais acentuada, visto a sensibilidade e capacidade de danos à individualidade. Essas são algumas das razões que os sistemas médicos exigem proteções avançadas, maior sigilo na divulgação e acesso restrito às informações.

Sob tais parâmetros, empresas mundiais de proteção de dados criaram sistemas destinados ao armazenamento de dados médicos. Amazon AWS, Microsoft Azure e IBM Watson Health são exemplos de sistemas específicos, que atendem às exigências das legislações, sob proteção em diversos níveis.

A exploração dos sistemas consolidados, em consonância das recentes regulações de proteção de dados pessoais, permitiria avanços em pesquisas, promovendo melhorias em tratamento, prevenção e diagnostico médico. Contudo, a falta do desenvolvimento das bases legais para promover o acesso, especialmente a fim de alimentação de inteligência artificial e consequente desenvolvimento de novas tecnologias, é um dos entraves atuais.

Fica claro pelo exposto que o sistema de *big data* possui potencial inexplorado, cabendo ao trabalho analisar a possibilidade de acesso aos seus dados em consonância com os regulamentos de proteção de dados pessoais e ao direito à privacidade.

Importante salientar que inexiste estudo similar proposto, acredito que, primeiro pela temática recente, segundo pela possibilidade inexplorada pelos grandes centros de pesquisa, cabendo a este autor o estudo da viabilidade e a fundamentação da pesquisa.

3. **Familiaridade com objeto da pesquisa**

A utilização de dados pessoais tem sido objeto de regulamentação pelos principais países e mercados desenvolvidos. Em recente manifestação legislativa a Lei Federal 13.709 de 14 de agosto de 2018 foi publicada regulamentando o uso de dados pessoais no Brasil.

A atuação do autor como advogado público, em especial na Consultoria Jurídica do Hospital da Faculdade de Medicina de Botucatu, permite que o tema privacidade e utilização de dados estejam em constante estudo. Ainda, o interesse pessoal por novas tecnologias e a maneira de como a sociedade pode ser afetada por elas, justificam a exploração do tema e a pesquisa.

Ademais, prevejo que o tema será necessário para a atuação em todas as áreas do direito, tendo em vista que hoje o uso e acesso de dados pessoais é realidade em todas as profissões. Especialmente em relação aos dados médicos, o tema é debatido pela minha Consultoria desde 2013 quando questões relacionadas a prontuários eletrônicos foram realizadas, bem como, a forma de acesso por residentes e pesquisadores.

O tema está diretamente relacionado à minha atuação profissional, visto que o maior provedor de saúde no Brasil é o Estado, portanto, potencial detentor das informações necessárias para pesquisas e desenvolvimentos de novas tecnologias. Como advogado público, procuro compreender a melhor forma de acesso a esses dados, sugerir aprimoramento legislativo e condutas para o acesso nos termos legais.

4. **Modelo de pesquisa**

O modelo preponderante da pesquisa é exploratório, familiarizando-se com o assunto e constituindo hipóteses. Desse modo, buscamos construir soluções funcionais e teoricamente embasadas que atendam às necessidades do mundo real, com utilidade prática e aplicabilidade imediata, especialmente através da análise crítica e abrangente com argumentação lógica e racional.

Diante de uma neutralidade e independência do autor, buscaremos construir informações corretas e completas, indicando as fontes de pesquisa, e ao final formulando propostas construtivas e realistas.

As fontes principais serão artigos, legislações nacionais e internacionais, decisões judiciais e administrativas, trabalhos acadêmicos e banco de dados sobre privacidade, inteligência artificial e dados médicos, utilizando-se da experiência própria como bussola para estimular a compreensão.

Por fim, a estrutura do trabalho será compreensão dos fatos, qualificação jurídica, analise e avaliação crítica, concluindo pela recomendação de condutas ou ações práticas.

5. **Fontes e métodos de investigação**

A pesquisa terá como fonte e métodos de investigação as seguintes fontes. Legislativa: Regulamento Europeu GDPR, Lei Federal 13.709 de 14 de agosto de 2018, legislação americana. Bibliográfico: textos doutrinários e artigos que tratam do tema Regulamentação e Proteção dos dados pessoais. Jurisprudencial: decisões europeias, americanas e brasileiras sobre o tema privacidade, dados pessoais e dados médicos.

6. **Quesitos**

A principal questão a ser respondida pelo trabalho é de como o acesso a dados médicos disponíveis em sistemas de *big data* podem ser utilizados no desenvolvimento de novas tecnologias, em consonância e com fundamento nas leis de proteção de dados pessoais e ao direito à privacidade.

Ademais, outras questões marginais são:

1. Quais os fundamentos para que pesquisadores tenham acesso a sistemas de *big data?* A individualização dos dados pessoais em sistemas de *big data* pode ferir o direito à privacidade?
2. Quais os limites impostos pelos regulamentos de proteção de dados pessoais?
3. Como os dados médicos podem ser explorados em consonância com os regulamentos de proteção de dados?
4. A análise de dados por inteligência artificial fere regulamentos de proteção de dados pessoais?
5. Caso afirmativa a questão anterior, como os dados médicos podem ser acessados e analisados sem violar os regulamentos de proteção dos dados? As empresas detentoras dos sistemas podem permitir o acesso de interessados nos termos dos regulamentos?
6. Quais as consequências impostas pelos regulamentos de proteção a dados pessoais em caso de violação?
7. Quais são as sugestões de aprimoramento legislativo?

7. **Bibliografia preliminar**

BALDWIN, Robert e Cave, Martin (1999). Understanding Regulation. Oxford Univ. Press.

BUTLER, John Matthew and Becker, William C. and Humphreys, Keith N., Big Data and the Opioid Crisis: Balancing Patient Privacy with Public Health (May 29, 2018). Journal of Law, Medicine & Ethics, 46 (2018): 440–453. Disponível em SSRN: https://ssrn.com/abstract=3228581

CHASSANG, Gauthier, "The Impact of the EU General Data Protection Regulation on Scientific Research," *ecancermedicalscience*, vol. 11, no. 709, 2017; Disponível em: https://doi.org/10.3332/ecancer.2017.709.

COHEN, I. Glenn and Mello, Michelle M., *HIPAA and Protecting Health Information in the 21st Century* (May 24, 2018). JAMA, 2018. Disponível em SSRN: https://ssrn.com/abstract=3242904

COHEN, I. Glenn e Hoffman, Sharona and Adashi, Eli Y, Your Money or Your Patient's Life?: Ransomware and Electronic Health Records (October 17, 2017). *Annals of Internal Medicine*, 2017. Disponível em SSRN: https://ssrn.com/abstract=3242907

COHEN, Molly and Sundararajan, Arun. Self-Regulation and Innovation The Peer-to-Peer Sharing Economy. 82 U *Chicago Law Review Dialogue* 116 (2015) Disponível em: https://lawreview.uchicago.edu/page/self-regulation-and-innovation-peer-peer-sharingeconom.

DIETTERICH, Thomas G. *A comparative Review os Selected Methods for Learning From Examples.* California: Palo Alto, 1983

DIVI C, Koss RG, Schmaltz SP, et al. Language proficiency and adverse events in U.S. hospitals: a pilot study. *Int J Qual Health Care* 2007

FLORIDI, Luciano, Soft Ethics and the Governance of the Digital (February 18, 2018). *Philosophy & Technology, Forthcoming*. Disponível em SSRN: https://ssrn.com/abstract=3125685

FOSCH, Eduard; Kieseberg, Peter; Li, Tiffany, Humans Forget, Machines Remember: Artificial Intelligence and the Right to Be Forgotten (August 13, 2017). Computer Security & Law Review (Forthcoming). Disponível em SSRN: https://ssrn.com/abstract=3018186

GOSTIN, Lawrence O. and HALABI, Sam and Wilson, Kumanan, Health Data and Privacy in the Digital Era (July 17, 2018). JAMA, Vol. 320, Number 3, Pp. 233-234; University of Missouri School of Law Legal Studies Research Paper. Disponível em SSRN: https://ssrn.com/abstract=3219253

GREENWALD, Glenn. No Place to Hide: Edward Snowden, the NSA and the U.S. Surveillance State. New York: metropolitan Books, 2014.

INSTITUTE OF MEDICINE, *Committee on Quality of Health Care in America.* To err is human: building a safer health system. Washington, DC: National Academy Press; 2000

KAHNEMAN, Daniel. Maps of Bounded Rationality: A Perspective on Intuitive Judgment and Choice. *Nobel Prize Lecture*, Dec. 8, 2002.

LOBEL, Orly. The Renew Deal: The Fall of Regulation and the Rise of Governance in Contemporary Legal Thought. *Minnesota Law Review*, Vol. 89, San Diego Legal Studies Paper No. 07-27. Disponivel em SSRN: https://ssrn.com/abstract=723761.

NISSENBAUM, Helen. A contextual approach to privacy online. In: Daedalus, Vol. 140, Issue 4, 2011, pp. 32-48.

OLIVA, Jennifer D., Prescription Drug Policing: The Right to Protected Health Information Privacy Pre- and Post-Carpenter (August 2, 2018). Disponível em SSRN: https://ssrn.com/abstract=3225000

SANJEEV, Contra e Sardana, *Big Data: It's Not a Buzzword, It's a Movement*, FORBES (Nov. 20, 2013), Disponível em: http://www.forbes.com/sites/sanjeevsardana/2013/11/20/bigd

SCHWARTZ, Paul M. e Solove, Daniel J., The PII Problem: Privacy and a New Concept of Personally Identifiable Information, *New York University Law Review*, vol. 86, no. 6, 2011, p. 1814. Disponível em: https://scholarship.law.berkeley.edu/cgi/viewcontent.cgi?article=2638&context=facpubs

SOLOVE, Daniel J. I've Got Nothing to Hide' and Other Misunderstandings of Privacy. San Diego Law Review, Vol. 44, p. 745, 2007; GWU Law School Public Law Research Paper No. 289.

SOLOVE, Daniel. Conceptualizing Privacy. In: California Law Review, Vol. 90, 2002.

SUNDQUIST, Christian, The Technologies of Race: Big Data, Privacy and the New Racial Bioethics (2018). Annals of Health Law, Vol. 27, 2018. Disponível em SSRN: https://ssrn.com/abstract=3245235

8. **Cronograma de execução**

	2018			2019												
Atividade	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12	Horas
Pesquisa bibliográfica	5	10	5													[20h]
Leitura bibliográfica	10	20	10	30	30	30										[130h]
Redação Capitulo I							30									[30h]
Redação Capitulo II								30								[30h]
Redação Capitulo III									30							[30h]
Redação Capitulo IV										20	10					[30h]
Redação Capitulo V											20	10				[30h]
Conclusão e Introdução												20	10			[30h]
Revisão													20	20	20	[60h]
Deposito															-	[...]