

TICs no Apoio à Compreensão de Termos de Uso e Políticas de Privacidade

Tatiana Oshiro Kobayashi¹, Valéria Quadros dos Reis²

^{1,2}Faculdade de Computação – Universidade Federal do Mato Grosso do Sul

tatiana.kobayashi@aluno.ufms.br¹, valeria.reis@ufms.br²

Resumo. Neste trabalho apresentamos alguns motivos que dificultam a leitura de termos de uso e políticas de privacidade por usuários da Internet e mostramos, através de uma breve revisão da literatura, como tecnologias de comunicação e informação podem ser utilizadas para melhorar a compreensão desses termos.

Abstract. In this paper we present some reasons that turn difficult the reading of terms of use and privacy policies by Internet users. Through a brief literature review, we show how information and communication technology can improve the comprehensibility of these terms.

1. Introdução

Em tempos em que nossas características físicas, interesses, mensagens, agendas, viagens, compras, trajetos, entre muitas outras informações, estão disponíveis para empresas de tecnologia, nos tornamos altamente vulneráveis, como indivíduos e sociedade, pois tais informações podem ser exploradas para diferentes fins, desde campanhas de vendas direcionadas até a manipulação de opinião de acordo com vieses políticos [Tilt 2018, Bloomberg 2019, The Guardian 2018, CNET 2019, Jurno and de Brito D' Andréa 2017].

Apesar dos frequentes escândalos de vazamento, venda e mau uso de dados na rede mundial de computadores, pouco se tem feito para esclarecer à população como esses dados se tornam disponíveis para as empresas. Muitas vezes, o uso dos dados acontece com o consentimento do usuário, pois no momento de uso de serviço na Internet, ele é obrigado a concordar com condições e regras da organização provedora do serviço através do consentimento de termos de uso. Dessa maneira, é extremamente importante que se leia e entenda os contratos de uso antes de assiná-los. Contudo, com frequência tais documentos são extremamente longos, redigidos em linguagem jurídica e para um público com alta escolaridade [Luger et al. 2013, Jensen and Potts 2004].

Por meio dos termos firmados entre as partes, o prestador de serviços pode delimitar as situações em que será responsável e por quais situações não poderá ser imputável em caso de falhas. Por esse motivo, muitas vezes, os termos são utilizados como uma blindagem das organizações para se protegerem contra ações judiciais e deixam de informar claramente questões que são consideradas importantes pelo usuário final [Earp et al. 2005, Pollach 2007].

Um problema paralelo à falta de usabilidade e simetria de informações dos contratos digitais surge quando pessoas com alto nível de instrução optam por utilizar serviços cujos termos de uso e políticas de privacidade são sabidamente desfavoráveis ao usuário.

O grande número de usuários do aplicativo FaceAPP, por exemplo, famoso por condicionar seu uso à coleta, modificação e utilização de dados e imagens dos usuários, inclusive com cessão a terceiros; demonstra como parte da população ignora a maneira que seus dados são tratados pelas empresas. [Tilt 2020].

É indiscutível que os contratos de uso na Internet precisam evoluir para apresentar linguagem centrado no entendimento comum. Esse assunto, apesar de amplamente discutido na literatura, não fez surtir impacto suficiente para atrair a atenção dos usuários [Berger-Walliser et al. 2017, Tabassum et al. 2018, Pollach 2007, McDonald et al. 2009].

Também é verdade que há uma falha de cidadania digital em parte da sociedade que não enxerga como suas informações podem ser coletadas, usadas e divulgadas [Mossberger et al. 2011]. O ataque a esse problema deve ser feito a médio e longo prazo com investimentos em educação digital. Entretanto, no Brasil, a ação esbarra na grande desigualdade social do país. Dados do governo brasileiro revelam que 63.3 milhões de pessoas estão digitalmente excluídas, além de serem 21 milhões de lares sem acesso à Internet [Ministério da Educação 2018, Corrêa 2018, ANJ 2018].

Dadas as limitações das abordagens tradicionalmente utilizadas para melhorar a compreensão dos termos de uso e políticas de privacidade, acreditamos que seja preciso investir em novas maneiras de informar a população sobre seus direitos e deveres nos serviços da Internet. Tecnologias de Informação e Comunicação (TICs) podem ser utilizadas com esse fim ao servirem como pontes na interpretação e sumarização de contratos que sejam centrados nas dúvidas dos usuários em relação ao serviço utilizado.

Dessa maneira, o objetivo deste trabalho é a discussão sobre o uso de TICs no apoio à compreensão de contratos de uso de serviços digitais. De formas diversas, essas tecnologias poderiam interpretar documentos extensos e redigidos em linguagem jurídica e apresentá-los de modo abreviado, humanamente amigável e com foco nos interesses do usuário final.

O restante deste trabalho está dividido da seguinte maneira: a Seção 2 define o que são Termos de Uso e acordos similares e apresenta motivos que dificultam a leitura e entendimento desses termos; a Seção 3 apresenta trabalhos na área de tecnologia relacionados diretamente com o entendimento de Termos de Uso e Políticas de Privacidade; e, por fim, a Seção 4 apresenta as considerações finais.

2. Termos de Uso e Políticas de Privacidade

Os Termos de Uso e suas variações (Termos de Serviço, Termos e Condições de Uso, entre outros) são documentos onde as empresas definem as condições para o uso de algum serviço provido por ela, incluindo o que o usuário pode ou não fazer com o mesmo, se há alguma concessão por parte de algum dos lados, e qualquer outra condição. É basicamente um contrato entre uma empresa provedora de um serviço e um cliente, não havendo diferenças significativas entre tais variações.

Já as Políticas de Privacidade, algumas vezes definidas dentro dos próprios termos, definem o que a empresa pode fazer com os dados dos seus usuários e visitantes. O uso dos dados por terceiros e a identificação dos dados armazenados são alguns exemplos de situações que são definidas em tais documentos.

No Brasil, existem poucas leis que regulamentam tais tipos de documento, além das que tratam de proteção de dados pessoais. O artigo 7º do Marco Civil da Internet define que “informações claras e completas sobre coleta, uso, armazenamento, tratamento e proteção” dos dados pessoais dos usuários devem estar especificadas nos contratos de prestação de serviço ou nos Termos de Uso das aplicações, com coleta justificada e que não seja vedada pela legislação [Legislação Brasileira 2014]. Recentemente, a Lei Geral de Proteção de Dados Pessoais prevê exigências sobre como as informações pessoais são tratadas e aplicadas pelas empresas e organizações [Legislação Brasileira 2018].

2.1. O Desinteresse pela Leitura

É de conhecimento geral que poucos lêem os Termos de Uso de serviços e ferramentas que encontram na Internet, mesmo que, na maioria das vezes, seja necessário confirmar a leitura e concordância com tais termos.

Uma pesquisa feita em 2016 utilizou uma rede social fictícia para avaliar o comportamento das pessoas ao serem confrontadas com um novo serviço. A pesquisa revelou que 74% dos 543 entrevistados não leram a Política de Privacidade apresentada, sendo que o tempo médio de leitura entre os que a leram foi de apenas 73 segundos. No total, 97% dos participantes concordaram com as Políticas de Privacidade e 93% com os Termos de Serviço. Ambos os documentos continham cláusulas de “pegadinha”, como filhos primogênitos como forma de pagamento e compartilhamento de dados com empregadores e com o governo americano [Obars and Oeldorf-Hirsch 2016].

A falta de interesse na leitura dos termos pode ser atribuída a diversos fatores, tais como, tamanho da fonte, disposição da página, extensão do documento ou complexidade da linguagem utilizada. Além disso, muitos serviços alteram seus termos sem aviso prévio e sem recurso para identificação das cláusulas alteradas [Luger et al. 2013, Jensen and Potts 2004, Pollach 2007].

A fim de comparação, em [Kobayashi 2019], foram coletados 29 termos de serviços providos no Brasil com o objetivo de levantar a quantidade de palavras utilizadas em cada documento. Na Tabela 1 são apresentados os termos analisados com informações a respeito: do serviço de onde o termo foi retirado; do tipo e do número de palavras de cada documento; e da quantidade de palavras dos termos agrupadas por empresa.

Dentre os termos analisados, o tamanho médio em palavras foi de 32059 e a maioria deles tinha entre 10 mil e 30 mil palavras, como é possível ver no histograma da Figura 1.

O maior termo, da *Microsoft*, tinha 111.974 palavras. O menor termo, do sítio de buscas *Gmail*, apresentava 4169 palavras. Considerando todos os termos de uma mesma empresa, a *Microsoft* se destaca novamente como a empresa com os documentos com maior número de palavras.

Utilizando a mesma velocidade média de leitura de um adulto (250 palavras por minuto) da pesquisa de [Obars and Oeldorf-Hirsch 2016], os Termos de Serviço da *Microsoft* levariam entre 6 e 7 horas para serem lidos. Em comparação, o documento do *Gmail* levaria cerca de 16 minutos

A análise da quantidade de palavras usadas nos termos de serviço nacionais ajuda a entender um dos motivos pelo qual os usuários de serviços no Brasil seguem a tendência

Tabela 1. Serviços em Português avaliados.

Serviço	Tipo de Termo	Quantidade de Palavras por Documento	Empresa	Quantidade de Palavras por Empresa
Adobe	Termos de Uso	33704	Adobe	33704
Amazon	Notificação de privacidade	15530	Amazon	71273
Amazon	Condições de uso	20671	Amazon	71273
Amazon Appstore	Termos de uso	16509	Amazon	71273
Amazon Devices	Termos de uso	10363	Amazon	71273
App Store	Termos e Condições de Serviço	52893	Apple	52893
Facebook	Termos de Serviço	27674	Facebook	44286
Gmail	Política do Programa	4169	Google	101297
Google	Termos de Serviço	14050	Google	101297
Google	Política de Privacidade	29593	Google	101297
Google Play	Termos de Serviço	28379	Google	101297
Instagram	Termos de Uso	16612	Facebook	44286
Loja Kindle	Termos de uso	8200	Amazon	71273
Microsoft	Contrato de Serviços	111974	Microsoft	111974
Olx	Termos e Condições de Uso	15436	Olx	15436
Passei Direto	Termos de Uso	45166	Passei Direto	45166
Spotify	Política de Privacidade	34006	Spotify	91360
Spotify	Termos e Condições de Uso	57354	Spotify	91360
Steam	Acordo de assinatura	55021	Valve Brasil	55021
Twitter	Política de Privacidade	35076	Twitter Brasil	52697
Twitter	Termos de Serviço	17621	Twitter Brasil	52697
Udemy	Termos de Uso	39269	Udemy Brasil	101988
Udemy	Política de Privacidade	42950	Udemy Brasil	101988
Udemy	Termos do Instrutor	19769	Udemy Brasil	101988
Yahoo-Verizon	Termos de Serviço	76008	Verizon Brasil	91814
Yahoo-Verizon	Política de Privacidade	15806	Verizon Brasil	91814
Youtube	Termos de Serviço	25106	Google	101297
Change.org	Termos de Uso	27657	Change.org	60804
Change.org	Política de Privacidade	33147	Change.org	60804

mundial de não ler Termos de Uso. Outros motivos que corroboram esse comportamento são o baixo nível de escolaridade aliado ao reduzido grau de cidadania digital da população. Ainda que iniciativas possam ser desenvolvidas para minimizar esses problemas, existe uma abordagem paralela que utiliza a tecnologia para reduzir a lacuna entre o que as organizações querem informar e o que o usuário deseja saber. TICs podem ser utilizada para esse fim.

3. TICs no Apoio à Compreensão de Termos

Existem algumas TICs que focam especificamente em apoiar a compreensão de usuários sobre Termos de Serviço e Políticas de Privacidade.

A Terms Of Service; Didn't Read (ToS;DR) é uma organização que defende os

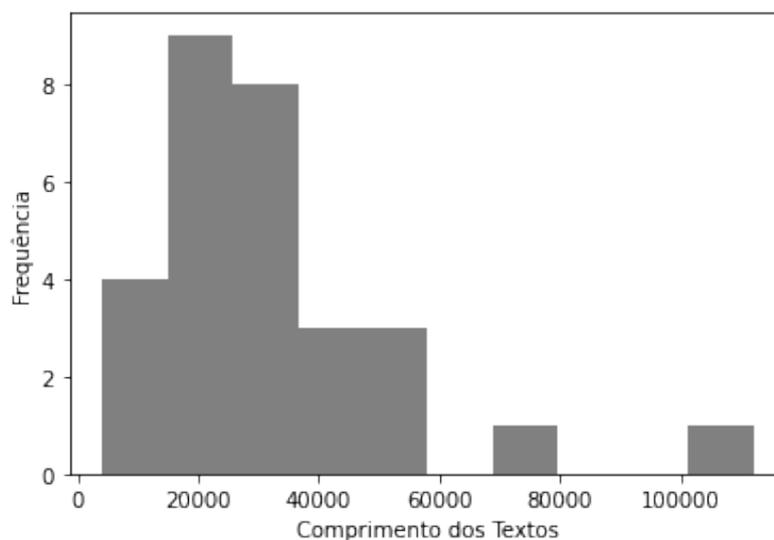


Figura 1. Histograma dos comprimentos dos termos (em palavras).

direitos do usuário e avalia Termos e Políticas de Privacidade de diversos serviços *online*, com o objetivo de conscientizar usuários sobre os conteúdos de tais documentos [ToS;DR 2019].

As classificações feitas pela organização são disponibilizadas para a comunidade em geral por meio de uma página web e por uma interface de programação para os desenvolvedores de sistemas. Nas classificações, os termos analisados são separados em tópicos menores e mais fáceis de se discutir e comparar. Os tópicos podem receber uma de quatro categorias: “bom”, “neutro”, “ruim” e “bloqueante”. A partir dos tópicos categorizados, o serviço é classificado entre ‘A’, ‘B’, ‘C’, ‘D’ e ‘E’, sendo que a classe ‘A’ contém os melhores termos. Exemplos de tópicos utilizados são ”O serviço rastreia sua navegação na Internet” e ”O serviço pode coletar, usar e compartilhar sua localização”.

Com o uso da interface de programação provida pela (ToS;DR), [Kobayashi 2019] criou uma extensão de navegador que monitora continuamente as páginas acessadas e confronta os termos dessas páginas com as preferências de privacidade previamente cadastradas pelo usuário. Caso haja discrepância de interesses, notificações na tela do usuário são geradas a fim de alertá-lo a respeito dos riscos.

O Pribot é um serviço que utiliza aprendizado de máquina para ler, interpretar e responder perguntas de usuários sobre os termos e políticas de diversos serviços através de um *chatbot*. Ele também reproduz resumos de políticas de privacidade de diversos sites através da ferramenta *Polisis*, mostrando o caminho percorrido pelos dados coletados pelo serviço de forma gráfica, facilitando o entendimento dessas informações para o usuário comum [Harkous et al. 2018][Pribot 2019].

PrOnto é uma ontologia computacional que modela os conceitos e normas do Regulamento Geral de Proteção de Dados da União Europeia [Palmirani et al. 2018]. Ela foi utilizada como base para a proposta do DaPIS, um conjunto de ícones representantes de informação legal, legíveis por máquina, mas pensados a partir da perspectiva de usabilidade do usuário [Rossi and Palmirani 2019]. Uma rica representação visual provida

automaticamente a partir de tecnologias semânticas também é uma maneira de melhorar a compreensão dos termos de uso e políticas de privacidade através do uso de TICs.

Em todos os trabalhos citados, as tecnologias de comunicação e informação foram usadas no armazenamento, tratamento e divulgação da informação disponibilizada ao usuário final. Apesar de parecer contraditório utilizar TICs para tratar um problema das implicações que a própria tecnologia traz à sociedade, acreditamos haver significativa demanda, principalmente no contexto brasileiro, para iniciativas semelhantes que auxiliem o entendimento de contratos de serviço.

4. Conclusão

Devido à importância que a tecnologia alcançou na vida diária das pessoas, é necessário que os usuários de serviços da Internet sejam conscientizados sobre seus direitos e deveres ao concordar com Termos de Serviço. Infelizmente, a forma como os termos são escritos e apresentados dificulta a leitura e o entendimento dos mesmos por pessoas sem conhecimento legal. No Brasil, o baixo nível de letramento digital da população agrava ainda mais o problema.

Muitos esforços têm sido feitos para melhorar a usabilidade dos termos e também aumentar a cidadania digital das pessoas. Entretanto, poucas iniciativas atuam entre esses dois extremos. Acreditamos que tecnologias digitais de comunicação e informação são ideais para esse fim, pois podem reduzir a lacuna existente entre o que as organizações dispõem nos documentos legais e o que os usuários esperam encontrar neles. Dessa maneira, não seria um paradoxo usar TICs para mitigar um problema criado pelo próprio uso da tecnologia.

Neste artigo descrevemos brevemente alguns trabalhos que apoiam a compreensão de termos de uso e políticas de privacidade com o uso de TICs. As tecnologias empregadas variam desde ferramentas web até algoritmos de aprendizado de máquina e ontologias. Como estudos futuros vislumbramos a necessidade de se desenvolver, aprimorar, empregar e avaliar tecnologias de apoio à compreensão de termos de serviço.

5. Agradecimentos

O presente trabalho foi realizado com apoio da Universidade Federal de Mato Grosso do Sul – UFMS/MEC – Brasil.

Referências

- ANJ (2018). Excluídos digitais são 63 milhões. Disponível em <https://www.anj.org.br/site/exemplos/97-midia-nacional/5948-excluidos-digitais-sao-63-milhoes.html>. Acessado em 25/11/2019.
- Berger-Walliser, G., Barton, T. D., and Haapio, H. (2017). From Visualization to Legal Design: A Collaborative and Creative Process. *American Business Law Journal*, 54(2):347–392.
- Bloomberg (2019). Youtube executives ignored warnings, letting toxic videos run rampant. Disponível em <https://www.bloomberg.com/news/features/2019-04-02/youtube-executives-ignored-warnings-letting-toxic-videos-run-rampant>. Acessado em 16/03/2020.

- CNET (2019). Facebook paid teens \$20 a month to access their browsing history and DMs. Disponível em <https://www.cnet.com/news/facebook-pays-people-to-access-their-phone-activity-report-says/>. Acessado em 16/03/2020.
- Corrêa, M. (2018). Quase 24 milhões de brasileiros não usam a Internet por falta de conhecimento. *O Globo*. Disponível em <https://oglobo.globo.com/economia/quase-24-milhoes-de-brasileiros-nao-usam-internet-por-falta-de-conhecimento-22416811>. Acessado em 30/11/2019.
- Earp, J. B., Antón, A. I., Member, S., Aiman-smith, L., and Stufflebeam, W. H. (2005). Examining Internet Privacy Policies Within the Context of User Privacy Values. *IEEE Transactions on Engineering Management*, 52:227–237.
- Harkous, H., Fawaz, K., Lebet, R., Schaub, F., Shin, K. G., and Aberer, K. (2018). PolliS: Automated Analysis and Presentation of Privacy Policies Using Deep Learning. *USENIX Security 2018*.
- Jensen, C. and Potts, C. (2004). Privacy policies as decision-making tools: An evaluation of online privacy notices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '04, page 471–478, New York, NY, USA. Association for Computing Machinery.
- Jurno, A. C. and de Brito D'Andréa, C. F. (2017). (IN)VISIBILIDADE ALGORÍTMICA NO “FEED DE NOTÍCIAS” DO FACEBOOK. *Portal de Periódicos da UFBA*, 15(2). Disponível em <https://portalseer.ufba.br/index.php/contemporaneaposcom/article/view/17796>. Acessado em 20/03/2020.
- Kobayashi, T. O. (2019). O Uso de TICs no Apoio à Cidadania Digital e na Compreensão de Termos de Uso. Trabalho de Conclusão de Curso apresentado na Faculdade de Computação - UFMS.
- Legislação Brasileira (2014). Lei 12.965/2014 – http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/l12965.htm. Acessado em 30/11/2019.
- Legislação Brasileira (2018). Lei 13.709/2018 – http://www.planalto.gov.br/ccivil_03/_Ato2015-2018/2018/Lei/L13709.htm. Acessado em 30/11/2019.
- Luger, E., Moran, S., and Rodden, T. (2013). Consent for all: Revealing the hidden complexity of terms and conditions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, page 2687–2696, New York, NY, USA. Association for Computing Machinery.
- McDonald, A. M., Reeder, R. W., Kelley, P. G., and Cranor, L. F. (2009). A comparative study of online privacy policies and formats. In *Proceedings of the 5th Symposium on Usable Privacy and Security*, SOUPS '09, New York, NY, USA. Association for Computing Machinery.
- Ministério da Educação (2018). Base nacional comum curricular. Disponível em <http://basenacionalcomum.mec.gov.br/>. Acessado em 29/11/2019.

- Mossberger, K., J. Tolbert, C., and McNeal, R. S. (2011). *Digital Citizenship - The Internet, Society & Participation*.
- Obars, J. A. and Oeldorf-Hirsch, A. (2016). The Biggest Lie on the Internet: Ignoring the Privacy Policies and Terms of Service Policies of Social Networking Services. *TPRC 44: The 44th Research Conference on Communication, Information and Internet Policy*, pages 1–20.
- Palmirani, M., Martoni, M., Rossi, A., Bartolini, C., and Robaldo, L. (2018). PrOnto: Privacy Ontology for Legal Reasoning. In *Electronic Government and the Information Systems Perspective*, pages 139–152, Cham. Springer International Publishing.
- Pollach, I. (2007). What’s wrong with online privacy policies? *Commun. ACM*, 50(9):103–108.
- Pribo (2019). <https://pribo.org/>. Acessado em 30/11/2019.
- Rossi, A. and Palmirani, M. (2019). Dapis: An ontology-based data protection icon set. In *Frontiers in Artificial Intelligence and Applications*, pages 181 – 195. IOS Press.
- Tabassum, M., Alqhatani, A., Aldossari, M., and Richter Lipford, H. (2018). Increasing user attention with a comic-based policy. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI ’18, page 1–6, New York, NY, USA. Association for Computing Machinery.
- The Guardian (2018). Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach. Disponível em <https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election>. Acessado em 01/12/2019.
- Tilt, U. (2018). Google monitora e seleciona o que você vê, mas porque você deixa. Disponível em <https://www.uol.com.br/tilt/noticias/redacao/2018/12/11/audiencia-de-sundar-pichai-no-congresso-dos-estados-unidos.htm>. Acessado em 16/03/2020.
- Tilt, U. (2020). FaceApp rouba os meus dados? Veja 6 coisas que você devia saber sobre ele... Disponível em <https://www.uol.com.br/tilt/noticias/redacao/2020/06/15/faceapp-rouba-os-meus-dados-seis-coisas-que-voce-devia-saber-sobre-ele.htm>. Acessado em 17/06/2020.
- ToS;DR (2019). Terms of Service; Didn’t Read – <https://tosdr.org/>. Acessado em 30/11/2019.