

Aceitação do Aviso de *cookies* e criação de publicidade direcionada: uma decisão consciente ou falta de informação?

Manuel Batista¹, Adriana Fernandes², André Sabino³, Lilian Ponzo Ribeiro⁴

mmbaa11@iscte-iul.pt; adriana.fernandes@iscte-iul.pt; asabino@autonoma.pt; lilianponzo@yahoo.com

¹ ISCTE – Instituto Universitário de Lisboa, MGSI, Lisboa, Portugal

² ISCTE – Instituto Universitário de Lisboa, ISTAR_ISCTE, Lisboa, Portugal

³ Autónoma Techlab, Universidade Autónoma de Lisboa, Lisboa, Portugal

⁴ Departamento de Direito, Faculdade de Sabará, SOECS, Sabará, Minas Gerais, Brasil

DOI: 10.17013/risti.43.75-92

Resumo: Ao navegar na Internet o utilizador fornece dados como a idade, género, localização, entre outros, que são capturados através da utilização de *cookies* – pequenos ficheiros de texto que estão armazenados no *web browser* quando se visita um *website*. Diversas empresas utilizam-se desses dados para mostrar ao utilizador publicidade direcionada aos seus interesses. Neste contexto, estudos recentes demonstram que a utilização de *cookies* tem implicações relacionadas com a privacidade, sendo o Regulamento Geral sobre a Proteção de Dados (RGPD) do Parlamento Europeu por vezes ineficaz no seu principal propósito. O presente estudo pretende compreender a perceção do utilizador sobre o uso de *cookies* para criação de publicidade direcionada. Para concretização do estudo foi realizado um inquérito com 242 utilizadores portugueses ativos na Internet. Os resultados indicam que apesar dos utilizadores se sentirem incomodados pelo fato de serem rastreados, preferem fornecer os dados, concordando com as políticas de *cookies*, e visualizar a publicidade direcionada.

Palavras-chave: *Cookies*; Privacidade; Publicidade Direcionada; RGPD.

Acceptance of the Cookie Notice and the creation of targeted advertising: a conscious decision or lack of information?

Abstract: When surfing the Internet, the user provides data such as age, gender, location, among others, which are captured using cookies - small text files that are stored in web browsers when one visits a website. Several companies use this data to show the user advertising targeted to their interests. However, recent studies show that the use of cookies has privacy-related implications, and the General Data Protection Regulation (GDPR) of the European Parliament is sometimes ineffective in its main purpose. In this context, the purpose of this study is to understand user

perception of the use of cookies to create targeted advertising. To carry out the study, a survey was conducted with 242 Portuguese users active on the Internet. The results show that, although users feel uncomfortable with being tracked, they prefer to provide data, by agreeing with cookie policies, and view targeted advertising.

Keywords: Cookies; Privacy; Targeted Advertising; GDPR.

1. Introdução

É comum que o utilizador, ao aceder a *websites*, aplicações e motores de busca, se depare com publicidade diretamente relacionada aos seus comportamentos *online*. O que por vezes o utilizador não sabe é quais foram os procedimentos utilizados pelas empresas para que tal facto acontecesse.

A utilização de *cookies*, i.e., pequenos ficheiros de texto que estão armazenados nos navegadores *web* quando se visita um *website* (Kulyk et al., 2018), tem despertado interesse das organizações devido à sua capacidade de seguir os “passos” do utilizador na Internet. Segundo Sanchez-Rola, et al., os *cookies* passaram a ser considerados uma ferramenta tecnológica que permite traçar o perfil de utilizadores, facilitando a segmentação e o direcionamento da publicidade e evitando o desperdício de recursos em utilizadores que não se enquadram no perfil pretendido pela organização (Sanchez-Rola et al., 2019).

A relação entre os *cookies* e a publicidade direcionada tem sido discutida, principalmente, quanto aos aspetos legais relacionados com a privacidade e a proteção de dados, uma vez que o Regulamento Geral sobre a Proteção de Dados (RGPD) do Parlamento Europeu determina que as empresas devem apresentar um “Aviso de *Cookies*” ao utilizador cada vez que o mesmo aceder aos seus *websites*. Esse aviso tem como objetivo obter o consentimento explícito, específico e claro para o uso de *cookies*. No entanto, estes avisos parecem induzir o utilizador para o botão “Aceito”, utilizando-se de textos pouco explícitos. O utilizador comum não está informado e tendenciosamente aceita as condições de utilização do *website*, sem conhecer as consequências desse ato (Kulyk et al., 2018). A aceitação autoriza a recolha e utilização, por parte das empresas, de dados pessoais dos utilizadores, e em alguns casos, até o fornecimento desses dados a terceiros, constituindo uma ameaça à privacidade.

No entanto, alguns estudos apontam que não há por parte dos utilizadores conhecimento suficiente sobre o que são os *cookies*, quais as implicações da sua aceitação, e como os dados são utilizados (Kulyk et al., 2018; Sanchez-Rola et al., 2019). Apesar da falta de conhecimento sobre as implicações da aceitação dos *cookies*, e de como as empresas utilizam esses dados, estudos apontam que a grande maioria dos consumidores preferem aceitá-los e receber a publicidade direcionada (Kirkpatrick, 2016).

Tendo em vista esse contexto, levanta-se uma série de questões relevantes como: Até que ponto o utilizador está realmente preocupado com seus próprios dados? Será o utilizador compreende o que são os *cookies* e para o que são utilizados? Será que se houvesse mais informação sobre o que são os *cookies* e para o que são utilizados, os utilizadores deixariam de aceitá-los? Será que a proteção estabelecida pelo RGPD é realmente alvo de interesse pelo utilizador? Será que os utilizadores se importam em ter os próprios dados rastreados?

O presente estudo pretende focar-se em clarificar a perceção da população jovem adulta portuguesa que utiliza a Internet sobre o uso de *cookies* para criação de publicidade direcionada. Nomeadamente, o estudo pretende contribuir para a compreensão quanto ao conhecimento do utilizador sobre o que são *cookies*, sobre sua perceção quanto à visualização de conteúdo direcionado, e ao facto de potencialmente ser rastreado. Para uma organização, este estudo pretende contribuir para melhorar a compreensão sobre se os utilizadores dos seus *websites* se preocupam em decidir o destino dos seus dados, e como devem ser utilizados, ou se entendem que os dados fornecidos são parte normal do processo de utilização da Internet.

O estudo pretende também contribuir para perceber se o RGPD está a ser corretamente posto em prática, na medida em que exige clarificação das práticas usadas para obter o consentimento do utilizador para o uso de *cookies*. Este trabalho servirá de alerta às entidades reguladoras para uma possível “violação consentida” da privacidade, dando assim o seu contributo à sociedade em geral.

2. Revisão da Literatura

2.1. Cookies

O recurso mais utilizado no rastreamento de dados são os *cookies*, pequenos ficheiros de texto que estão armazenados no *web browser* quando se visita um *website* (Kulyk et al., 2018). Introduzidos em 1994 pela NetScape, os *cookies* tinham a finalidade de permitir pesquisas com armazenamento de dados (i.e., *stateful browsing*) em comunicações através do protocolo HTTP (*HyperText Transfer Protocol*). Inicialmente, foram utilizados para melhorar a experiência do utilizador e fornecer funcionalidades adicionais aos *websites*, tais como, gravar temporariamente o “carrinho de compras” ou armazenar os dados de início de sessão no *browser*, para que na próxima visita ao *website* o utilizador não tivesse novamente que os introduzir (Kulyk et al., 2018).

Atualmente, os *cookies* deixaram de servir apenas ao seu propósito base, existindo em duas modalidades: os de sessão, e os persistentes. Os *cookies* de sessão, como o nome indica, só estão ativos durante o período em que o utilizador mantém o *browser* aberto. Já os *cookies* persistentes permanecem no *browser* até que o utilizador, ou o fornecedor do serviço, os remova, sendo responsáveis por registar o histórico das ações de todo o período em que os utilizadores se encontram ativos no *browser* (Kulyk et al., 2018). O registo de dados históricos do utilizador permite obter informação valiosa, de forma a permitir a análise do perfil e comportamento dos utilizadores e, a partir dessas informações, oferecer publicidade direcionada aos interesses do mesmo (Sanchez-Rola et al., 2019).

O interesse das organizações pelos *cookies* tem sido crescente, dada a sua capacidade em seguir a atividade do utilizador na Internet. Esta funcionalidade é vista como extremamente valiosa, ou seja, permite que a segmentação e o direcionamento sejam muito mais eficientes, evitando o desperdício de recursos em utilizadores que não se enquadram no perfil pretendido pela organização.

Os *cookies* podem estar presentes onde o utilizador menos espera, como é o caso da ferramenta ReCAPTCH, da Google, usada por muitos *websites* para diferenciar utilizadores de robôs. O comum utilizador é exposto a esta ferramenta inúmeras vezes, e cada vez que este se conecta a um desses *websites*, conecta-se automaticamente a um domínio da Google, de forma a ler os ficheiros necessários para que a ferramenta trabalhe, e assim partilha todos os *cookies* associados também com a Google. Logo, este tipo de ferramentas constitui outra fonte de dados para enriquecer *cookies*, refinando o perfil do utilizador (Cyphers & Gebhart, 2019).

Por outro lado, os *cookies* também podem ser vantajosos aos utilizadores. Kirkpatrick (2016) aponta que 71% dos consumidores, utilizadores da Internet, prefere publicidade direcionada, em oposição a ser continuamente exposto a todo o tipo de publicidade, por vezes irrelevante.

No entanto, muitas vezes o utilizador não possui conhecimento suficiente sobre como os seus dados são usados pelas organizações que os recolhem (Culnan & Williams, 2009). Essas, por sua vez, não se esforçam para que essa informação seja divulgada e compreendida pelo consumidor, que dá consentimento ao “Aviso de *Cookies*” de um *website* sem analisar as consequências desse ato (Chaudhry & Berger, 2019). No âmbito ético, os *cookies* estão numa área ainda pouco explorada, e o RGPD, apesar de criado recentemente, não consegue estabelecer uma conduta ética a ser seguida pelas empresas que atuam online. A secção 2.2 oferece uma discussão sobre esta questão.

2.2. Privacidade e Regulamento Geral de Proteção de Dados

Privacidade é definida como a condição do que é privado, pessoal ou íntimo. É frequentemente associada a dados pessoais, e à proteção destes dados. A Organização para a Cooperação e Desenvolvimento Económico (OCDE) define *dado pessoal* como “(...) qualquer informação relacionada com o indivíduo identificado ou identificável (sujeito dos dados)” (OCDE, 2002). No contexto atual, a Privacidade centra-se no direito dos cidadãos em controlar a informação pessoal que disponibilizam, e a utilização que esta possa vir a ter por terceiros (Pires, 2018).

A Privacidade a que o cidadão tem direito deve ser exigida das organizações, desde o nível da conceção das tecnologias; a Privacidade deve ser respeitada nas operações e gestão do tratamento de dados pessoais. A esta abordagem dá-se o nome de *Privacy by Design* e permite, ao mesmo tempo, proteger os utilizadores, e manter as funcionalidades e o bom desempenho das tecnologias (Romanou, 2018).

Em 2014, a Fundação *Open Web Application Security Project*¹ (OWASP), que tem como objetivo melhorar a segurança dos *softwares*, apresentou um projeto com os principais riscos relativos à privacidade dos dados pessoais, denominado “OWASP *Top 10 Privacy Risks*”. Alguns dos riscos apresentados estão diretamente associados à aceitação de *cookies* por parte do utilizador, sendo eles:

- Políticas, Termos e Condições não transparentes, ou seja, o não fornecimento de informações suficientes para descrever como os dados são recolhidos, armazenados, processados e excluídos;

¹ <https://owasp.org/>

- Recolha de dados relacionados com o indivíduo que não sejam necessários para o propósito original do sistema. Também se aplica a qualquer tipo de dados cujo titular não tenha consentido partilhar;
- Fornecimento de dados de um indivíduo a qualquer terceiro sem primeiro obter o consentimento do titular. Esta partilha pode ocorrer devido a uma compensação monetária pela partilha de dados, devido ao uso inapropriado de recursos de terceiros ou até mesmo devido a erros;
- Falta de expiração de sessão ou expiração inexistente, o que poderá resultar na recolha de mais dados sem o consentimento ou consciencialização do indivíduo (Oliveira, 2019).

A 25 de maio de 2018 foi colocado em prática o RGPD, que constitui um conjunto de leis da União Europeia (UE) para regular o manuseamento de dados pessoais, com o objetivo de fortalecer os direitos dos cidadãos na era digital e facilitar negócios, por meio da clarificação das regras para as organizações no mercado digital (Batista et al., 2020). Os efeitos deste regulamento fizeram-se sentir além da UE, pois a lei aplica-se a qualquer oferta de bens ou serviços na UE, mesmo que a organização que os oferece esteja fora deste espaço (Oliveira, 2019; Sanchez-Rola et al., 2019).

O RGPD aplica-se a dados pessoais, e descreve-os como qualquer informação que possa identificar uma pessoa, incluindo dados de pseudónimos que podem ser atribuídos a um indivíduo pelo uso de informação adicional (Parlamento Europeu, 2016). Ou seja, basta a identificação ser possível para que se possa considerar “dado pessoal”, e os *cookies* são explicitamente mencionados, ainda que apenas uma vez, como um dos identificadores *online*, a par dos endereços de IP (*Internet Protocol*), que, especialmente quando combinados com várias fontes que o utilizador disponibiliza, podem ser usados para criar perfis de pessoas reais e identificá-las.

Segundo o RGPD, os dados pessoais não devem ser retidos mais do que o estritamente necessário, com o máximo de 12 meses (Parlamento Europeu, 2016; Sanchez-Rola et al., 2019). Este regulamenta ainda que o consentimento do utilizador para o uso de *cookies* deve ser explícito, específico e claro, com indicações do acordo entre a organização e o cidadão, e que não devem ser processados dados até o utilizador consentir o uso dos mesmos. A legislação também estabelece que não pode ser recusado ao utilizador o acesso a um *website* no caso de haver recusa dos cookies sugeridos (exceto os que são essenciais ao seu funcionamento). A finalidade da recolha de dados deve ser explícita, legítima e determinada quando do consentimento do utilizador (Oliveira, 2019; Sanchez-Rola et al., 2019).

2.3. Publicidade Direcionada

A capacidade de apresentar conteúdo personalizado a cada consumidor, no momento e localização pretendidos, em tempo real, com possível interação com cada um, são fatores-chave de atração de novos clientes (Dahiya & Gayatri, 2018). É possível atingir o patamar da publicidade direcionada por meio da segmentação do histórico do comportamento do cidadão enquanto esteve online: as páginas visitadas, os cliques, as pesquisas, as compras, entre outros (Cyphers & Gebhart, 2019; Summers et al., 2016).

As organizações que mais recolhem dados, como a Google, o Facebook, a Amazon e o Twitter, constituem plataformas ideais para disseminação de publicidade, pois possuem dados muito precisos dos utilizadores, permitindo otimizar a entrega de artefactos publicitários com muita facilidade (Cyphers & Gebhart, 2019), e desenhar estratégias de Marketing onde as comunicações são feitas com elevada insistência, consistindo numa forma de persuasão psicológica que as organizações fazem com os cidadãos (Matz et al., 2017).

Na tarefa de segmentação de utilizadores, os dados podem provir de diferentes fontes, como os *cookies*, o endereço de IP (*Internet Protocol*); o estado do TLS – *Transport Layer Security*; os chamados “*super cookies*” armazenados localmente; e o *browser fingerprint*, i.e., meta-informação enviada pelo browser, que facilita a identificação do utilizador, onde são utilizadas ferramentas como o WebGL que, para além da função de permitir apresentar gráficos 2D e 3D no navegador *web*, também regista informação acerca do navegador do utilizador, como a resolução do ecrã, a versão do software instalado e a hora local (Sanchez-Rola et al., 2019).

3. Metodologia

Para atender o objetivo da pesquisa foi realizado um *survey* com uma amostra composta por cidadãos portugueses, entre os 20 e 34 anos, considerados população jovem com maior utilização da Internet em Portugal (Pordata, 2019). O método de amostragem utilizado foi o não probabilístico, por conveniência (acessibilidade?), caracterizado por seleccionar os membros mais acessíveis de uma dada população (Vergara, 2005).

3.1. Recolha de dados

O *survey* foi operacionalizado por um questionário, dividido em 5 blocos (ver Apêndice). O bloco I visou caracterizar o inquirido (género, idade, nível de escolaridade, área de formação académica, distrito de residência e tipo de ocupação). Os demais blocos foram constituídos por um conjunto de itens (nomeadamente, afirmações), que constituíam questionários utilizados em estudos anteriores, dado serem construções pré-testadas, desta forma promovendo a confiabilidade da escala (Bolívar-Ramos et al., 2012). Devido à especificidade e atualidade do tema, foi também necessário recorrer a algumas adaptações, tendo como base a bibliografia analisada.

O bloco II visou identificar os hábitos do inquirido associados às atividades realizadas na Internet. O bloco III apresentou afirmações sobre o conhecimento do inquirido acerca de *cookies*. O bloco IV teve como finalidade aferir o conhecimento do inquirido sobre o que é o conteúdo direcionado e suas preferências quanto ao mesmo. E por fim, o bloco V apresentou afirmações sobre o conhecimento do utilizador acerca da relação entre *cookies* e publicidade direcionada, focando a indiferença do inquirido ao ser rastreado.

As questões dos blocos III, IV e V foram operacionalizadas em escala de Likert de sete pontos, variando de (1) discordo totalmente a (7) concordo totalmente. Este tipo de escala ordinal consiste num conjunto de itens apresentados em forma de afirmações, onde é solicitado ao inquirido que externalize a sua reação, escolhendo um dos sete pontos (Babbie, 2007).

Antes de iniciar a recolha dos dados, foi realizado um pré-teste, ou estudo-piloto, com 6 inquiridos, no mês de maio de 2020, com o intuito de identificar possíveis falhas, avaliar a linguagem e a complexidade de questões formuladas, detetar ambiguidades e, principalmente, a relevância do conteúdo do instrumento.

Após a realização do pré-teste e efetuadas as adequações ao questionário consequentes, iniciou-se a recolha dos dados. Para tal, foi utilizada a ferramenta de formulário online *Google Forms*. A divulgação do mesmo foi efetuada por meio de e-mails, Facebook e LinkedIn pessoais, com o período de respostas a decorrer entre os dias 1 e 30 de junho de 2020. A análise estatística dos dados foi efetuada com o software *IBM SPSS Statistics*, versão 26.

3.2. Caracterização da amostra

Foram recolhidos 335 questionários. Foram excluídos 71 por inadequação ou preenchimento incompleto, e 22 após validação da amostra por se considerarem *outliers*. O tamanho efetivo da amostra foi de 242 participantes.

A amostra é composta principalmente por inquiridos do sexo feminino (50,4%), com idade entre 30 e 34 anos (46,7%), residentes no distrito de Lisboa (83,1%) e empregados por conta de outrem (70,7%). Quanto à formação, a maior parte dos inquiridos possui alguma formação académica (74,4%), sendo as áreas de formação de maior destaque a das Tecnologias da Informação (18,2%), a das Ciências Sociais e Humanas (18,2%) e Economia, Finanças ou Gestão (15,7%). As restantes áreas têm todas menos de 10% dos inquiridos (ver Tabela 1).

Amostra N=242		Amostra N=242	
Gênero		Distrito de Residência	
	Masculino 49,6%	Aveiro	0,8%
	Feminino 50,4%	Castelo Branco	0,8%
		Coimbra	0,8%
Idade		Leiria	2,1%
	20 - 24 anos 24,0%	Lisboa	83,1%
	25 - 29 anos 29,3%	Portalegre	0,4%
	30 - 34 anos 46,7%	Porto	2,9%
		Região Autónoma da Madeira	0,4%
Nível de Escolaridade		Região Autónoma dos Açores	0,4%
	Básico 1,6%	Setúbal	8,3%
	Secundário 24,0%		
	Licenciatura 43,4%	Área de Formação Académica	
	Mestrado 30,2%		
	Doutoramento 0,8%	Artes	3,3%
		Direito	2,5%
Tipo de Ocupação		Economia, Finanças ou Gestão	15,7%
	Estudante 14,0%	Medicina	3,3%

Amostra N=242		Amostra N=242	
Investigador	0,8%	Química ou Biologia	1,7%
Empregado por conta de outrém	70,7%	Social e Humana	18,2%
Trabalhador independente	10,3%	Tecnologias da Informação	18,2%
Desempregado	4,2%	Turismo	2,5%
		Outras Engenharias	8,7%
		Outras	18,6%
		Sem especialização académica	7,3%

Tabela 1 – Características demográficas da amostra

4. Análise de dados e resultados

4.1. Hábitos associados às atividades realizadas na Internet

O primeiro passo para compreender a perceção do utilizador sobre o uso de *cookies* para criação de publicidade direcionada, foi tentar melhor compreender os hábitos dos inquiridos associados às atividades realizadas na Internet.

Quanto à utilização da Internet, 48,3% dos inquiridos afirmam que, em média, o uso diário é superior a 6 horas. Relativamente ao uso de redes sociais, onde foi possível a escolha de várias alternativas, as combinações mais comuns foram “Instagram, Facebook, Youtube”, com 21,5%, e “Instagram, Facebook, LinkedIn, Youtube”, com 19,8%, sendo as restantes combinações muito variadas. Numa análise por rede social, o Instagram e o Facebook equiparam-se, ao serem ambos usados por cerca de 86% da amostra, seguindo-se do Youtube (82,6%) e LinkedIn (52,1%). Apesar do alto número de horas gastas na Internet, somente 24,8% dos inquiridos afirma fazer compras online com frequência (Tabela 2).

Amostra N=242		Amostra N=242	
Uso da Internet		Uso de Redes Sociais (múltiplas opções)	
menos de 1 hora por dia	0,4%	Facebook	85,6%
1 a 3 horas por dia	23,1%	Instagram	86,0%
4 a 6 horas por dia	28,1%	LinkedIn	52,1%
mais de 6 horas por dia	48,3%	Twitter	17,8%
		Youtube	82,6%
Compras na Internet		Outra(s)	22,3%
Nunca	2,5%		
Raramente	26,9%		
Às vezes	45,0%		
Muitas vezes	24,8%		
Sempre	0,8%		

Tabela 2 – Resumo dos hábitos na Internet do inquirido

4.2. Análise Fatorial Exploratória

Uma vez que grande parte das questões que integraram o questionário provém de diferentes estudos, foi realizada a Análise Fatorial Exploratória (AFE) a fim de confirmar a estrutura latente da escala.

A AFE foi realizada com todos os 27 itens da escala. Inicialmente foi realizada um teste Kaiser-Meyer-Olkin (KMO) para avaliar a adequação da amostragem em termos de variáveis e fatores. A teste revelou um valor de 0,763, indicando que a matriz de correlação era adequada para a análise fatorial. Além disso, o teste de Esfericidade de Bartlett foi significativo (1003,31; $p < 0,001$), portanto as variáveis estão significativamente correlacionadas (Hair et al., 2009).

A consistência interna obtida com o Alfa de Cronbach indicou valores de 0,730 para toda a amostra com 13 itens. O Alfa de Cronbach é definido como “aceitável” quando for maior que 0,7 (Nunnally & Bernstein, 1994), indicando uma boa consistência interna da escala. A Tabela 3 mostra a correlação item-total corrigida e a consistência interna (alfa de Cronbach), mesmo que o item seja deletado. Todos os itens mostraram correlação com uma escala total superior a 0,3, um valor abaixo envolveria uma exclusão do fator (Castillo et al., 2010).

	Correção da correlação item/total	Alfa de Cronbach (se o item for excluído)
15. Estou ciente que as empresas proprietárias dos <i>cookies</i> podem (...).	0,389	0,754
16. Estou ciente que os <i>cookies</i> podem ser combinados com outras (...).	0,345	0,758
21. A publicidade que me surge na <i>Internet</i> está (...).	0,305	0,761
22. Compreendo que muitos conteúdos e serviços online (...).	0,368	0,755
23. Eu gosto de ver a publicidade/anúncios que aparecem (...).	0,300	0,762
25. Prefiro que os <i>websites</i> exibam anúncios segmentados (...).	0,481	0,743
26. Prefiro que os <i>websites</i> ofereçam descontos direcionados (...).	0,512	0,739
27. Prefiro que os <i>websites</i> que visito mostrem notícias (...).	0,460	0,745
29. Os <i>cookies</i> permitem que as empresas façam anúncios (...).	0,410	0,751
30. Quando aceito o Aviso de <i>Cookies</i> de <i>websites</i> , como <i>sapo.pt</i> (...).	0,320	0,760
31. Estou ciente que os dados pessoais que constam (...).	0,437	0,748
32. Não me importo que rastreiem as minhas informações (...).	0,301	0,763
36. Fico feliz em ter anúncios relevantes sobre coisas (...).	0,452	0,746

Tabela 3 – Correção da correlação item-total e consistência interna (alfa de Cronbach) se o item for excluído.

Os fatores foram extraídos através de análise do componente principal com rotação Equamax. As cargas de fatores foram avaliadas para cada item, em cada fator, e estabelecido um ponto de corte de 0,50. Para definir o número de fatores (a extrair) foi utilizado o critério de raiz latente, uma percentagem da variância e um *scree plot*.

Da AFE resultaram três fatores, de um total de 13 itens: “Conhecimento sobre *cookies*”; “Preferência por conteúdo direcionado”; e “Indiferença ao rastreamento”. Esses foram responsáveis por 57,69% da variância total. A confiabilidade dos fatores foi obtida com o Alfa de Cronbach, que aponta que os coeficientes aceitos são geralmente superiores a 0,6 (Hair et al., 2009). A Tabela 4 mostra as cargas dos itens em cada fator, juntamente com seus respectivos valores próprios (*eigenvalues*), o percentual de variância explicado, e índices de confiabilidade.

O fator I, “Conhecimento sobre *cookies*”, está associado ao conhecimento que a amostra tem sobre os *cookies* e como estes são usados na criação de publicidade direcionada ao perfil de cada um. Este fator explica 26,811% da variação dos dados, com um *eigenvalue* de 3,485. A confiabilidade para este fator foi de 0,779.

Carga dos itens	<i>Eigenvalues</i>	Variância explicada (%)	Alfa de Cronbach
Fator I: Conhecimento sobre <i>cookies</i>	3,485	26,811	0,779
Item 15	0,709		
Item 16	0,656		
Item 22	0,626		
Item 29	0,734		
Item 30	0,658		
Item 31	0,727		
Fator II: Preferência por conteúdo direcionado	2,660	20,464	0,775
Item 21	0,513		
Item 25	0,855		
Item 26	0,868		
Item 27	0,803		
Fator III: Indiferença ao rastreamento	1,354	10,417	0,663
Item 23	0,694		
Item 32	0,787		
Item 36	0,586		

Tabela 4 – Análise Fatorial Exploratória

O fator II, “Preferência por conteúdo direcionado”, ilustra essencialmente a preferência dos inquiridos por ter conteúdo direcionado ao seu perfil, ao invés de conteúdo com o qual não se identificam. O fator II explica 20,464% da variação dos dados, com um *eigenvalue* de 2,660 e uma confiabilidade de 0,775.

O fator III, “Indiferença ao rastreamento”, está ligado à forma como os inquiridos estão indiferentes ao facto de serem ou não rastreados. Este fator explicou 10,417% da variância dos dados, com um *eigenvalue* de 1,354, enquanto a confiabilidade é de 0,663.

4.3. Conhecimento dos inquiridos sobre *cookies*

A Tabela 5 apresenta a média e o desvio-padrão dos seis itens do modelo de medida que operacionaliza o conhecimento sobre *cookies*, que é medido numa escala contínua, em que 1 traduz um baixo conhecimento sobre os *cookies*, e 7 traduz um elevado conhecimento sobre os *cookies*. Neste grupo de questões, os valores médios de conhecimento variaram entre 4,93 e 5,94, sendo o valor médio para o fator de 5,41, o que demonstra que os inquiridos possuem um nível razoável de conhecimento sobre o tema.

Posteriormente, foram avaliadas as diferenças entre género, por escalão etário e por escolaridade para o fator. No entanto, os diferentes grupos não demonstraram diferenças significativas ($p > 0,05$), ou seja, os inquiridos do sexo masculino e do sexo feminino possuem o mesmo nível de conhecimento sobre o tema, e a idade e a escolaridade não alteram a perceção que têm do que são e para que servem os *cookies*.

Fatores / Itens	Itens / Dimensões	Média	Desvio Padrão
Fator I	Conhecimento sobre <i>cookies</i>		
15.	Estou ciente que as empresas proprietárias dos <i>cookies</i> podem (...).	5,86	1,28
16.	Estou ciente que os <i>cookies</i> podem ser combinados com outras (...).	5,18	1,63
22.	Compreendo que muitos conteúdos e serviços online (...).	5,94	1,29
29.	Os <i>cookies</i> permitem que as empresas façam anúncios (...).	5,41	1,39
30.	Quando aceito o Aviso de <i>Cookies</i> de <i>websites</i> , como <i>sapo.pt</i> (...).	5,13	1,57
31.	Estou ciente que os dados pessoais que constam (...).	4,93	1,61
	Média	5,41	1,46

Tabela 5 – Média e Desvio Padrão de itens do Fator I - Conhecimento dos Inquiridos sobre *cookies*

4.4. Preferência por Conteúdo Direcionado

Quanto à preferência por conteúdo direcionado na Internet, a Tabela 6 apresenta a média e o desvio-padrão destes itens, que, como os anteriormente analisados, é medido numa escala contínua, em que 1 traduz uma baixa preferência por conteúdo direcionado, e 7 traduz uma elevada preferência por este tipo de conteúdo. Neste grupo de questões, os valores médios de conhecimento variaram entre 4,56 e 6,13, sendo o valor médio para o fator de 5,10, o que demonstra que, entre conteúdo aleatório e conteúdo direcionado, há uma tendência para os inquiridos preferirem conteúdo direcionado.

Fatores / Itens	Itens / Dimensões	Média	Desvio Padrão
Fator II	Preferência por conteúdo direcionado		
21	A publicidade que me surge na Internet está (...).	6,13	1,11
25	Prefiro que os <i>websites</i> exibam anúncios segmentados (...).	4,56	1,65
26	Prefiro que os <i>websites</i> ofereçam descontos direcionados (...).	5,05	1,70

Fatores / Itens	Itens / Dimensões	Média	Desvio Padrão
27	Prefiro que os <i>websites</i> que visito mostrem notícias (...).	4,67	1,75
Média		5,10	1,55

Tabela 6 – Média e Desvio Padrão de itens do Fator II - Preferência por Conteúdo Direcionado

Posteriormente, também foram avaliadas as diferenças entre grupos, levando em consideração o género, o escalão etário e a escolaridade. Mais uma vez, não foram encontradas diferenças significativas ($p > 0,05$), ou seja, essas diferenças não afetam a percepção dos inquiridos sobre a publicidade direcionada.

4.5. Indiferença ao rastreamento

O fator III, classificado como a “Indiferença ao rastreamento”, engloba itens relativos ao facto de os inquiridos estarem ou não indiferentes por serem rastreados. A Tabela 7 apresenta a média e o desvio-padrão dos três itens do modelo de medida para a referida dimensão, que é medido numa escala contínua, em que 1 traduz muito incómodo ao ser rastreado, e 7 traduz elevada indiferença ao ser rastreado.

Fatores / Itens	Itens / Dimensões	Média	Desvio Padrão
Fator III	Indiferença ao rastreamento		
23	Eu gosto de ver a publicidade/anúncios que aparecem (...).	2,43	1,43
32	Não me importo que rastreiem as minhas informações (...).	2,83	1,61
36	Fico feliz em ter anúncios relevantes sobre coisas (...).	3,89	1,64
Média		3,05	1,56

Tabela 7 – Média e Desvio Padrão dos itens do Fator III - Indiferença ao Rastreamento

Neste grupo, os valores médios são os mais baixos, quando comparados com os fatores I e II, já analisados. Aqui, variaram entre 2,43 e 3,89, sendo o valor médio para o fator de 3,05, o que demonstra que os inquiridos não são indiferentes ao rastreamento, mas tendem, sim, a ficarem incomodados com esse facto (Tabela 7).

Item	Género	Média	Desvio Padrão	Teste <i>t</i>
23. Eu gosto de ver a publicidade/anúncios que aparecem (...).	Feminino	2,34	1,39	$t(240) = -2,857$; $p = 0,005$
	Masculino	2,52	1,46	
32. Não me importo que rastreiem as minhas informações (...).	Feminino	2,52	1,58	
	Masculino	3,15	1,59	
36. Fico feliz em ter anúncios relevantes sobre coisas (...).	Feminino	3,72	1,64	
	Masculino	4,07	1,63	

Tabela 8 – Comparação das médias de pontuação da escala dos itens do Fator III por género

Na sequência, foram analisadas as diferenças entre grupos. Não foram encontradas diferenças significativas ($p > 0,05$) entre os diferentes grupos etários e entre os inquiridos que possuem diferentes níveis de escolaridade. Entretanto, ao analisar as diferenças entre género, pode-se observar que existe diferença significativa entre os dois grupos ($p < 0,05$). Os inquiridos do sexo masculino apresentam uma média mais elevada do que os inquiridos do sexo feminino em todos os itens. Este resultado demonstra que os homens inquiridos se sentem, ligeiramente, menos incomodados com o facto de serem rastreados (ver Tabela 8).

5. Conclusões

O presente estudo teve como objetivo compreender a perceção do utilizador sobre o uso de *cookies* para criação de publicidade direcionada. Sendo o cidadão comum que usa a Internet um potencial consumidor, que é alvo constante de publicidade – seja nas redes sociais, páginas de notícias, blogs, plataformas de compras online, entre outros – considerou-se interessante discutir até que ponto o utilizador está realmente preocupado com seus próprios dados. Será que o utilizador compreende o que são os *cookies* e para que eles são utilizados? Será que se importam em ter os próprios dados rastreados? Para clarificar essas questões, foi realizado um *survey* com 242 utilizadores portugueses ativos na Internet.

Inicialmente, foi feita a validação da escala para avaliar a perceção do utilizador sobre o uso de *cookies* para criação de publicidade direcionada. A Análise Fatorial Exploratória apresentou uma solução com treze itens distribuídos em três fatores, que apresentaram uma boa consistência interna. Os fatores foram nomeados como: I – Conhecimento sobre *cookies*; II – Preferência por conteúdo direcionado; e III – Indiferença ao rastreamento.

Quanto aos resultados da análise da perceção do utilizador, foi possível verificar que, em relação ao conhecimento sobre os *cookies*, os valores médios estão em torno de cinco, numa escala ordinal, de 1 a 7, o que demonstra um nível razoável de conhecimento por parte dos inquiridos. Esse resultado não corrobora o estudo anterior por Culnan e Williams (Culnan & Williams, 2009). Acredita-se que, mais de dez anos após esse estudo, houve uma melhoria da divulgação da informação sobre o tema e maior conscientização do cidadão quanto às consequências da aceitação dos *cookies*.

No que respeita às preferências dos inquiridos por conteúdo direcionado, aferiu-se que os valores médios também estão em torno de cinco, numa escala ordinal de 1 a 7, o que sugere uma preferência média/alta por este tipo de conteúdo quando comparado com conteúdo aleatório. Esse resultado vai ao encontro da literatura (Kirkpatrick, 2016), que afirma que, em 71% dos casos, há preferência por conteúdo direcionado.

Quando questionados sobre estarem ou não indiferentes ao facto de serem rastreados, o valor mediano ficou em torno de três valores, em uma escala ordinal, em que 1 traduz muito incómodo ao ser rastreado, e 7 traduz elevada indiferença ao ser rastreado. Esse valor demonstra que os inquiridos não são indiferentes ao rastreamento, e tendem a ficar incomodados com esse facto. Entretanto, entende-se que o fato de se sentirem incomodados não inibe a aceitação dos *cookies* de forma aleatória, ou seja, a aceitação do “aviso de *cookies*” é feita de forma consciente. O incómodo demonstrado pelos respondentes salienta a necessidade de se proteger a privacidade do utilizador. No entanto, entende-se que o incómodo gerado não é suficiente para fazer com que o

utilizador rejeite, ou até mesmo analise o que está a aceitar. Esse facto sugere que o utilizador entende a situação como parte do processo de utilização da Internet.

No que respeita às diferenças entre grupos, constatou-se que, para os fatores I e II, não existem diferenças significativas ($p > 0,05$) se forem considerados o género, o escalão etário e a escolaridade. Quanto ao fator III, observou-se uma diferença significativa entre géneros quanto à preocupação em serem rastreados. As inquiridas do sexo feminino tendem a sentir-se mais incomodadas e preocupadas.

Por fim, acredita-se contribuir com a literatura, ao entregar uma escala validada sobre a perceção do utilizador sobre o uso de *cookies* para criação de publicidade direcionada. Considera-se que a escala aqui apresentada pode ser utilizada em futuros testes teóricos sobre o tema, com diferentes populações, em diferentes países que utilizam a língua portuguesa.

Para as organizações que se utilizam de *cookies* para criação de publicidade direcionada, este estudo contribui essencialmente para perceberem que os utilizadores têm conhecimento sobre o tema e sentem-se incomodados aos serem rastreados. Serve como um alerta para que essas organizações se comportem de forma mais ética e transparente, adequando a forma como se comunicam com os utilizadores e aplicam os *cookies*.

Embora o estudo tenha apresentado resultados relevantes, reconhece-se que existem algumas limitações a abordar em pesquisas futuras. Um exemplo disso é a necessidade de validação da escala por uma análise fatorial confirmatória, a fim de refinar mais resultados derivados deste estudo. Para estudos futuros, sugere-se ainda a realização de um teste prático a um grupo de utilizadores com uma proposta de manipulação total, simples e direta sobre os *cookies* em vários *websites*, com o objetivo de perceber a atitude do utilizador acerca da disponibilidade em facultar dados, num cenário onde exista mais clareza no processo.

Referências

- Batista, M., Fernandes, A., Ribeiro, L. P., Alturas, B., & Costa, C. P. (2020). Tensions between Privacy and Targeted Advertising: Is the General Data Protection Regulation being violated. *Iberian Conference on Information Systems and Technologies, CISTI, 2020-June*.
- Bolívar-Ramos, M. T., García-Morales, V. J., & García-Sánchez, E. (2012). Technological distinctive competencies and organizational learning: Effects on organizational innovation to improve firm performance. *Journal of Engineering and Technology Management - JET-M, 29*(3), 331–357.
- Castillo, C., Macrini, L., Cheniaux, E., & Landeira-Fernandez, J. (2010). Psychometric Properties and Latent Structure of the Portuguese Version of the Penn State Worry Questionnaire. *The Spanish Journal of Psychology, 13*(1), 431–443.
- Chaudhry, R., & Berger, P. (2019). Ethics in Data Collection and Advertising. *GPH - International Journal of Business Management, 2*(07 SE-Articles), 1–7.
- Culnan, M. J., & Williams, C. C. (2009). How ethics can enhance organizational privacy: Lessons from the ChoicePoint and TJX data breaches. *MIS Quarterly: Management Information Systems, 33*(4), 673–687.

- Cyphers, B., & Gebhart, G. (2019). Behind the one-way mirror: A deep dive into the technology of corporate surveillance. *Electronic Frontier Foundation* 53(9).
- Dahiya, R., & Gayatri, G. (2018). A Research Paper on Digital Marketing Communication and Consumer Buying Decision Process: An Empirical Study in the Indian Passenger Car Market. *Journal of Global Marketing*, 31(2), 73–95.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2009). *Multivariate Data Analysis* (7 ed). Pearson.
- Kirkpatrick, D. (2016). *Study: 71% of consumers prefer personalized ads*. Marketing Dive.
- Kulyk, O., Hilt, A., Gerber, N., & Volkamer, M. (2018). “This Website Uses Cookies”: Users’ Perceptions and Reactions to the Cookie Disclaimer.
- Matz, S. C., Kosinski, M., Nave, G., & Stillwell, D. J. (2017). Psychological targeting as an effective approach to digital mass persuasion. *Proceedings of the National Academy of Sciences of the United States of America*, 114(48), 12714–12719.
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory*. McGraw-Hill.
- OCDE. (2002). *Síntese - Diretrizes da OCDE para a Proteção da Privacidade e dos Fluxos Transfronteiriços de Dados Pessoais*.
- Oliveira, M. M. M. P. da C. (2019). *Privacidade no Ciclo de Vida do Desenvolvimento Seguro*. Universidade de Lisboa, Faculdade de Ciências.
- Parlamento Europeu. (2016). Regulamento (UE) 2016/679 - relativo à proteção das pessoas singulares no que diz respeito ao tratamento de dados pessoais e à livre circulação desses dados de 27 de abril de 2016. *Jornal Oficial Da União Europeia*, 2014(3), 1–119.
- Pires, R. M. S. (2018). *Gestão de risco em segurança da informação e privacidade no DNS.PT*. Universidade de Lisboa, Faculdade de Ciências.
- Pordata. (2019). *PORDATA - População residente segundo os Censos: total e por grupo etário*. <https://www.pordata.pt/Portugal/População+residente+segundo+os+Censos+total+e+por+grupo+etário++2>
- Romanou, A. (2018). The necessity of the implementation of Privacy by Design in sectors where data protection concerns arise. *Computer Law & Security Review*, 34(1), 99–110.
- Sanchez-Rola, I., Dell’Amico, M., Kotzias, P., Balzarotti, D., Bilge, L., Vervier, P. A., & Santos, I. (2019). Can i opt out yet? GDPR and the global illusion of cookie control. *AsiaCCS 2019 - Proceedings of the 2019 ACM Asia Conference on Computer and Communications Security*, 340–351.
- Summers, C. A., Smith, R. W., & Reczek, R. W. (2016). An audience of one: Behaviorally targeted ads as implied social labels. *Journal of Consumer Research*, 43(1), 156–178.
- Vergara, S. C. (2005). *Métodos de Pesquisa em Administração* (2nd ed.). Atlas S.A.

Apêndice:

Bloco I

Dimensão	Pergunta	Definição Operacional
Caracterização do respondente (inquirido?)	1.Género	Qualitativa nominal: feminino/masculino
	2.Idade	Qualitativa ordinal com cinco opções: menos de 20 anos, 20 a 24 anos, 25 a 29 anos, 30 a 34 anos, mais de 34 anos.
	3.Nível de Escolaridade	Qualitativa ordinal com cinco opções: Básico, Secundário, Licenciatura, Mestrado e Doutoramento.
	4.Área de Formação Académica	Qualitativa nominal com 11 opções, onde se inclui as opções “Outra” e “Sem formação académica”.
	5.Distrito de Residência	Qualitativa nominal com 18 distritos de Portugal Continental, as 2 Regiões Autónomas (Açores e Madeira) e ainda a opção “Estrangeiro”.
	6. Tipo de ocupação	Qualitativa nominal com cinco opções: “Estudante”, “Investigador”, “Empregado por conta de outrem”, “Trabalhador independente”, “Desempregado”

Bloco II

Dimensão	Pergunta	Autores de referência
Hábitos na Internet	7.Em média, seja por lazer, estudo ou trabalho, uso a Internet: (horas)	(Leon et al., 2013)
	8.Sou utilizador de redes sociais (marque quantas opções considerar necessárias): Instagram; Twitter; Facebook; LinkedIn; Youtube; outras; não utilizo.	(Leon et al., 2013)
	9.Em média, faço compras online: nunca; raramente; às vezes; muitas vezes; sempre.	(Leon et al., 2013)

Bloco III

Dimensão	Pergunta	Autores de referência
Conhecimento do respondente inquirido? acerca de <i>cookies</i> .	10. Os <i>cookies</i> são pequenos ficheiros de texto armazenados no meu computador.	(McDonald & Cranor, 2010)
	11. Eu aceito sempre os <i>cookies</i> dos websites que visito sem me importar com as consequências desse ato.	(Kulyk et al., 2018)
	12. Os <i>cookies</i> permitem que eu permaneça conectado ao longo do tempo sem precisar de inserir a minha senha sempre que visito um <i>website</i> .	(McDonald & Cranor, 2010)
	13. Regra geral, encontro opções pré-validadas para dar o meu consentimento ao uso de <i>cookies</i> .	(próprio)
	14. Quando visito um website, sinto-me obrigado/a a aceitar os <i>cookies</i> para aceder ao conteúdo que pretendo.	(próprio)
	15. Estou ciente que as empresas proprietárias dos <i>cookies</i> podem colocá-los em muitos <i>websites</i> e assim sabem quando os visito.	(McDonald & Cranor, 2010)
	16. Estou ciente que os <i>cookies</i> podem ser combinados com outras fontes (e.g.: endereço de IP, <i>browser fingerprints</i>), sendo possível identificar-me pessoalmente.	(McDonald & Cranor, 2010)
	17. A solicitação de consentimento apresentada pelos websites que utilizam <i>cookies</i> é clara e concisa.	(próprio)
	18. Sou cauteloso ao aceitar os <i>cookies</i> pois sei que muitas empresas registam os dados do meu comportamento <i>online</i> .	(McDonald & Cranor, 2010)
19. Considero invasivo saber que há alguém que acompanha as minhas atividades <i>online</i> .	(McDonald & Cranor, 2010)	

Bloco IV

Dimensão	Pergunta	Autores de referência
Preferência por conteúdo direcionado	20. Quando visito um <i>website</i> , vejo a mesma publicidade que qualquer outra pessoa que visita o mesmo <i>website</i> .	(Smit et al., 2013)
	21. A publicidade que me surge na <i>Internet</i> está, entre outros fatores, de acordo com o histórico dos <i>websites</i> que visitei.	(Smit et al., 2013)
	22. Compreendo que muitos conteúdos e serviços online são gratuitos graças aos lucros das empresas com a publicidade.	(Smit et al., 2013)
	23. Eu gosto de ver a publicidade/anúncios que aparecem nos <i>websites</i> que visito.	(Smit et al., 2013)
	24. Regra geral, carrego em anúncios que aparecem nos websites, para obter mais informações sobre o produto/serviço anunciado.	(Leon et al., 2013)
	25. Prefiro que os websites exibam anúncios segmentados para meus interesses.	(Smit et al., 2013)
	26. Prefiro que os websites ofereçam descontos direcionados aos meus interesses.	(Smit et al., 2013)
27. Prefiro que os websites que visito mostrem notícias personalizadas aos meus interesses.	(McDonald & Cranor, 2010)	

Bloco V

Dimensão	Pergunta	Autores de referência
Indiferença ao ser rastreado	28. Conheço as leis que me ajudam a lidar com a proteção dos meus dados relativamente à Publicidade Direcionada na Internet.	(Ur et al., 2012)
	29. Os <i>cookies</i> permitem que as empresas façam anúncios baseados no meu perfil e no histórico do meu comportamento <i>online</i> .	(McDonald & Cranor, 2010)
	30. Quando aceito o Aviso de <i>Cookies</i> de <i>websites</i> , como <i>sapo.pt</i> (imagem do Aviso), estou a permitir que este <i>website</i> e os seus parceiros me exponham à Publicidade Direcionada.	(próprio)
	31. Estou ciente que os dados pessoais que constam nos <i>cookies</i> são vendidos e que eu vejo o anúncio que pagou mais pelos meus dados.	(próprio)
	32. Não me importo que rastreiem as minhas informações, desde que os serviços se mantenham gratuitos.	(McDonald & Cranor, 2010)
	33. Ao visitar um <i>website</i> , não gosto que me surjam anúncios com os quais não me identifico.	(Leon et al., 2013)
	34. É assustador ter anúncios baseados nos <i>websites</i> que visitei.	(McDonald & Cranor, 2010)
	35. Regra geral, não me importo com os anúncios, apenas os ignoro.	(McDonald & Cranor, 2010) / (Smit et al., 2013)
	36. Fico feliz em ter anúncios relevantes sobre coisas em que estou interessado, em vez de anúncios aleatórios.	(McDonald & Cranor, 2010)
37. Eu odeio anúncios e pagaria para evitá-los.	(McDonald & Cranor, 2010)	