

SOUKROMÍ UŽIVATELŮ V PROSTŘEDÍ INTERNETOVÉ REKLAMY NA ČESKÉM WEBU

Reklama zajišťuje chod celé řady webových serverů, pro které často bývá hlavním zdrojem příjmů. V dnešním marketingu se používá celá řada způsobů cílení reklamy včetně cílení na dřívější zákazníky identifikované e-mailem nebo telefonním číslem. Pro zobrazení co nejrelevantnější reklamy jsou využívány aukční servery, které šíří nabídky prozrazující velké detaily o uživateli, jeho vlastnostech, prohlížeči, poloze apod. Proto jsou v poslední době podávány četné stížnosti k dohledovým orgánům. Jaký je stav v ČR? Tento článek ukazuje, že neradostný.

Služby, které jsou uživateli dostupné na internetu zdarma, musí někdo financovat. Musela být vybudována data centra, nakoupena technika, je nutné platit účty za elektřinu a platy zaměstnanců. Odkud provozovatelé peníze berou? Z obchodu s reklamou založenou na předávání osobních údajů. Dnešní trh s internetovou reklamou lze charakterizovat jako malou vesnici [3], ve které si reklamní společnosti vyměňují detailní informace o drtivě většině internetové populace na planetě [8, 5, 6]. *Real-time bidding* (RTB) umožňuje v desítkách až jednotkách stovek milisekund distribuovat mezi desítky i stovky společností informace o navštívených stránkách (URL) společně s dalšími informacemi jako jednoznačný identifikátor uživatele, IP adresa (nebo její podstatná část), poloha uživatele, informace o prohlížeči, operačním systému, velikosti obrazovky atd. [5].

Pokud budeme vycházet z rozhodnutí Soudního dvoru EU C-210/16 [9], jsou za to, co se s daty děje v rámci reklamního systému spoluzodpovědní provozovatelé služeb, kteří si partnery nasmlouvali včetně českých novin a dalších stránek zobrazující reklamu pomocí RTB, nebo poskytující služby a inzerující v RTB, tzn. dodávají informace svým partnerům, kteří je následně šíří desítkám až stovkám společností.

Již v roce 2018 byla podána dosud neuzavřená stížnost na fungování RTB k dohledovým úřadům v Irsku a Velké Británii, která se později rozšířila do dalších zemí, včetně ČR [8]. Ve Francii byla vyšetřovaná firma Vectaury, která vyšetřování dohledového úřadu CNIL přežila [4]. CNIL však nevyšetřoval podstatu systému RTB. Britský dohledový úřad ICO současné reklamní systémy charakterizuje jako *vytváření a sdílení osobních profilů osob v míře, která působí nepřiměřeně, vlezle a neférově obzvláště při uvážení, že se týká osob, které neví, že se takové aktivity dějí* [6, str. 4]. Norská asociace ochrany spotřebitelů Forbrukerradet provedla výzkum mobilních aplikací a charakterizuje reklamní systém jako *utržený z řetězců* a nabádá společnosti finančně závislé na příjmech z reklamy, aby si našli jiné zdroje příjmů¹.

Možnosti cílení internetové reklamy

V současnosti existuje několik možností, podle kterých si může inzerent vybírat, kde a komu chce zobrazovat reklamu. Následující výčet se zaměřuje na ty nejobvyklejší:

- Kontextová reklama: Inzerent specifikuje např. domény, části URL, kategorie webu, na kterých by chtěl inzerovat. Reklama tedy není cílená na specifické uživatele, ale zobrazuje se podle kontextu stránky. Dnes se však spíše používají techniky popsané níže.
- Behaviorální reklama: Reklamní společnosti budují profily pro uživatele prohlížečů. Profily obsahují atributy popisující vlastnosti uživatelů, jeho domovskou polohu apod. Inzerenti si vybírají kombinaci atributů, na které reklamu cílí — atributy, které by uživatelé měli a neměli mít. Yuan a kol. [12] píše, že společnost evidující nejvíce atributů jich uživateli přiřazuje více než 3000. U Facebooku bylo nalezeno celkem téměř 240000 atributů, které přiřazuje uživatelům [10]. Uživatel si pak na Facebooku průměrně nese 523,7 atributů,

¹<https://www.forbrukerradet.no/out-of-control/>

zatímco BluKai eviduje průměrně 422,4 atributů, a Google 42,3 atributů pro uživatele [2]. Účastníci výzkumu Bashira a kol. však považovali jen 27 % atributů za velmi relevantní [2].

- Retargeting: obvykle využívaný provozovateli e-shopů a poskytovateli placených služeb. Cílem je přesvědčit k návratu uživatele, kteří navštívili inzerentův web a neuskutečnili nákup. Inzerenti si však možná neuvědomují, že spolu s možností retargetingu dávají reklamním společnostem příležitost ke tvorbě profilů pro behaviorální reklamu. Je dobré si pečlivě zkontrolovat uzavřenou smlouvu, jestli obsahuje požadavek na reklamní společnost, aby na základě inzerentových dat nebudovala profily, ze kterých by mohla těžit konkurence. **Obrázek 1** ukazuje příklad komunikace návštěvníka webu České spořitelny, která instruuje prohlížeč ke kontaktování smluvních reklamních společností, nazývanými také jako *data management platform* (DMP). Nasbíraná data může Česká spořitelna později využít k retargetingu.
- Vlastní výběr osob (označovaný jako *custom audience*, nebo *customer match*)²: používají obvykle obchodníci, kteří chtějí oslovit své dřívější zákazníky, např. po uvedení nového produktu na trh, při slevových akcích, apod. Takový obchodník zná telefonní čísla a e-mailové adresy svých zákazníků. Tyto identifikační údaje tedy může nahrát do reklamního systému a kampaň nakonfigurovat tak, aby se reklama zobrazovala právě zákazníkům uvedeným v seznamu.
- Inzerce osobám podobných mému výběru (označovaný jako *look-alike audience*, nebo *similar audience*)³: nevýhodou vlastního výběru oslovených osob je, že tito zákazníci již nakoupili. Nevýhodou behaviorálního cílení je, že inzerent musí vědět, jaké atributy potenciální zákazník nejspíše má. Inzerce osobám podobných mému výběru sestaví automaticky behaviorální profil, který mají osoby uvedené na dodaném seznamu, a cílí na neuvedené osoby mající odvozené vlastnosti.

Status	Met...	Domain	File	Cause
200	GET	www.google.com	conversion_async.js	script
200	GET	c.imedia.cz	retargeting?id=25316&category=&itemId=&url=https://www.csas.cz/osobni-finance	img
304	GET	www.google-analytics.com	analytics.js	script
200	GET	track.adform.net	/Serving/Cookie/?adfaction=getjs;adfcokname=uid	script
200	GET	googleads.g.doubleclick.net	/pagead/viewthroughconversion/980831788/?random=1579257959998&cv=9&fst=1579257959998&num=1&bg...	script
200	GET	track.adform.net	/serving/scripts/trackpoint/async/	script
200	GET	www.google-analytics.com	js?id=GTM-KZT6KWR&cid=2143872554.1579257874	script
200	GET	www.google.com	/pagead/1p-user-list/980831788/?random=1579257959998&cv=9&fst=1579255200000&num=1&bg=ffffff&guid...	img
200	GET	www.google.cz	/pagead/1p-user-list/980831788/?random=1579257959998&cv=9&fst=1579255200000&num=1&bg=ffffff&guid...	img
200	GET	track.adform.net	/Serving/TrackPoint/?pm=1369211&ADFPPageName=path_/cs/osobni-finance&ADFdivider= &ord=908967354767...	script
200	GET	googleads.g.doubleclick.net	/pagead/viewthroughconversion/980831788/?random=1579257961735&cv=9&fst=1579257961735&num=1&gu...	script

Obrázek 1: Vybraná komunikace probíhající při návštěvě stránek České spořitelny. Sledující skripty a pixely.

Real time bidding — aukce o zobrazení reklamy

Pro zajištění zobrazení reklamy za co nejlepší cenu pořádají reklamní společnosti aukce RTB. Aukce probíhá samostatně pro každé místo pro zobrazení reklamy na stránce, kterou uživatel navštívil. **Obrázek 2** ukazuje šíření informace o návštěvě stránky uživatele k předem neznámému množství společností.

²<https://about.ads.microsoft.com/en-gb/solutions/audience-targeting/custom-audiences>, <https://www.godatafeed.com/blog/google-customer-match>, <https://www.businesswire.com/news/home/20110626005114/en/ChoiceStream>

³<https://support.google.com/google-ads/answer/7151628?hl=en>

1. Provozovatel stránky má uzavřenou smlouvu s jednou, nebo několika společnostmi zabývající se prodejem reklamy (*supply-side platform* — SSP).
2. Každé SSP může spolupracovat s jednou, nebo několika společnostmi zabývající se získáváním informací o internetových uživateli (DMP).
3. DMP identifikují uživatele pomocí skriptu vykonaném v uživatelské prohlížeči a dodají SSP informace o uživateli (typicky profil včetně retargetingových informací).
4. Každý SSP předá požadavek na zobrazení reklamy aukčním serverům, které mohou informace o uživateli dále obohatit daty svých DMP. DMP může identifikovat uživatele pomocí skriptu vykonaném v uživatelské prohlížeči.
5. Aukční servery rozešlou požadavek na nabídku v aukci (*bid request*) společností zabývajících se vytvářením reklamních kampaní a stykem s inzerenty (*demand-side platform* — DSP).
6. V rámci požadavku na nabídku v aukci je uživatel identifikován identifikátorem (pseudonymem) aukčního serveru. Pro ten může mít dodatečné informace některá z DMP spolupracujících s DSP.
7. DSP předávají aukčnímu serveru své nabídky na zobrazení reklamy včetně ceny, kterou jsou ochotni zaplatit. Nejlepší nabídky se dostanou k SSP.
8. DSP s nejlepší nabídkou může zobrazit uživateli reklamu a také distribuovat vlastní JavaScriptový kód směrem k uživateli a může provést tzv. *cookie matching* (CM)⁴, tedy propojení vlastního identifikátoru pro uživatele s pseudonymem používaným aukčním serverem. CM dovoluje DSP a partnerským DMP de-pseudonymizovat nejen všechny budoucí nabídky aukčního serveru, ale také všechny uložené dřívější nabídky týkající se tohoto uživatele.

Ne vždy jsou role DSP, DMP, SSP a aukčního serveru odděleny. Např. stejný subjekt může plnit roli DSP a SSP.

Transparency and Consent Framework

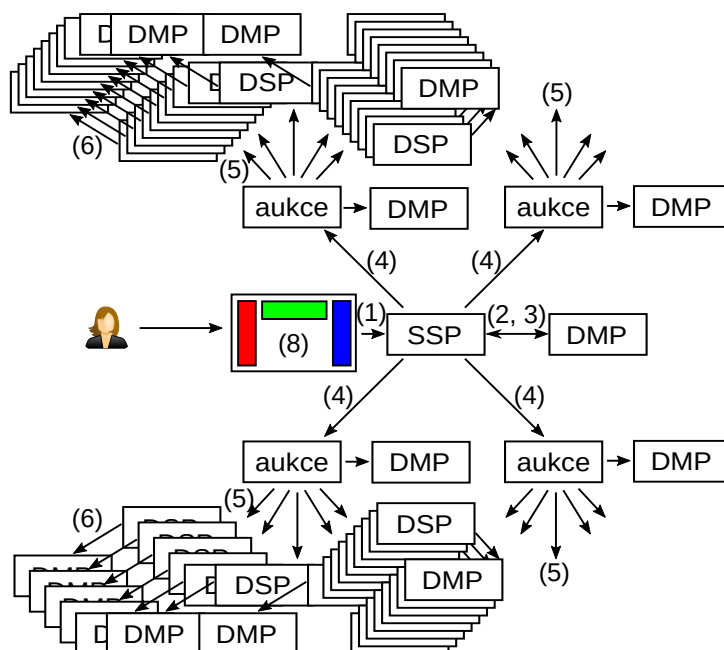
GDPR přineslo požadavek na transparentnost zpracování osobních dat. *Interactive Advertising Bureau Europe* vytvořila pro zajištění transparentnosti a získávání souhlasů se zpracováním osobních údajů *Transparency and Consent Framework* (TCF). V rámci něj reklamní společnosti oznamují své účely zpracování a zákonné důvody pro toto zpracování — oprávněné zájmy nebo souhlas subjektu údajů. Tento seznam je volně dostupný na Internetu⁵. V rámci analýzy TCF jsme na FIT VUT vytvořili dlouhodobé statistiky⁶.

Obrázek 3 ukazuje počet společností zapojených do TCF v pozici zjistit skutečnou identitu uživatele (identifikátor z reálného světa). Společnosti inzerující párování off-line údajů musí k propojení on-line a off-line dat znát identifikátory používané v reálném světě, např. zjistit uživatelskou e-mailovou adresu nebo telefonní číslo. Znalost identifikátorů z reálného světa umožňuje provádět propojování zařízení stejného uživatele, avšak není pro propojování zařízení stejného uživatele nutná [1]. Proto společností schopné zjistit skutečnou identitu uživatele zahrnují společností inzerující spojování off-line údajů a část společností inzerující propojování zařízení stejného uživatele.

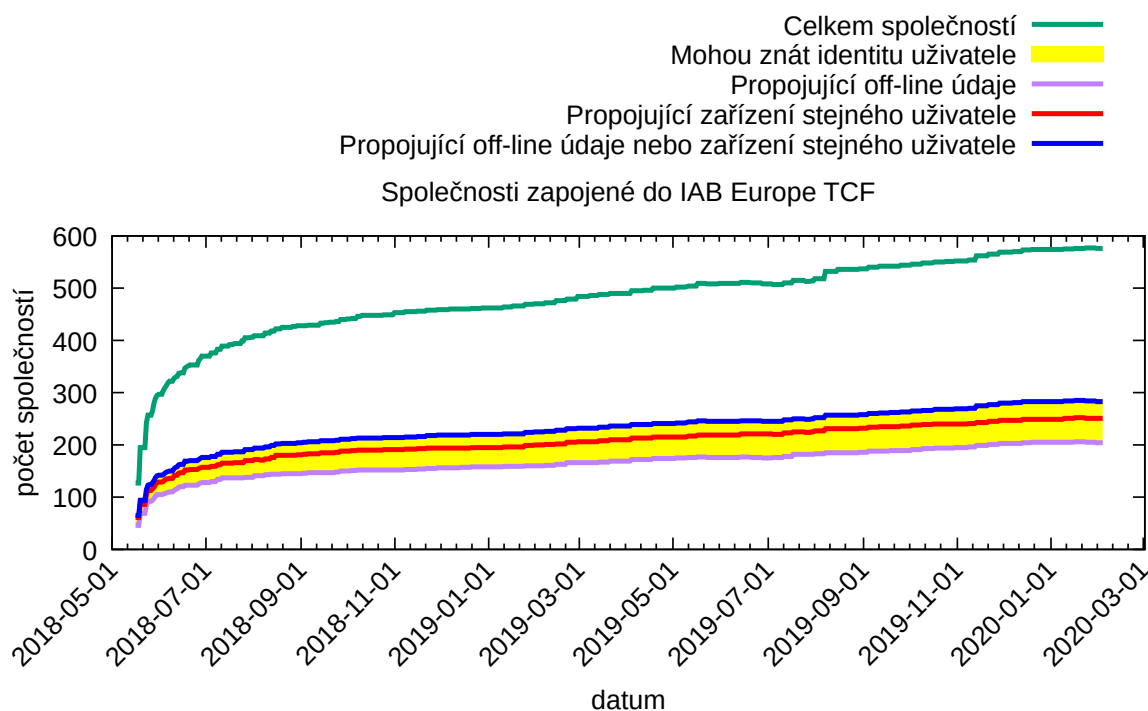
⁴<https://developers.google.com/authorized-buyers/rtb/cookie-guide>

⁵<https://vendorlist.consensu.org/vendorlist.json>

⁶<https://www.fit.vutbr.cz/~polcak/tcf/>



Obrázek 2: Ilustrativní schéma šíření informací v rámci RTB. Ve skutečnosti nemusí každá zapojená společnost spolupracovat s DMP, nebo může spolupracovat s více DMP. Podobně každá DMP může spolupracovat s více SSP, DSP a aukčními společnostmi. Pozn.: šíření nabídek od DSP k aukcím a SSP (7) není zakresleno.

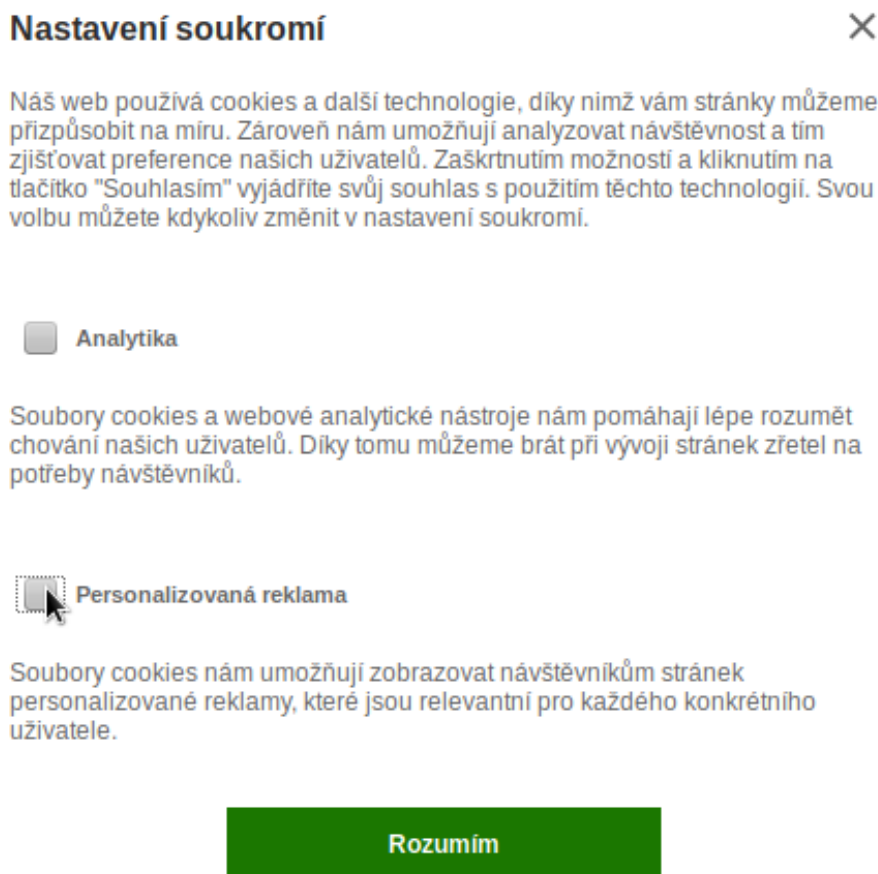


Obrázek 3: Počet společností zapojených do TCF a těch v pozici zjistit skutečnou identitu uživatele.

Analýza českého webu

Další výzkumnou otázkou je, jak moc jsou české weby zapojené do globálního reklamního systému pomocí RTB a jestli jsou o tomto předávání dat uživatelé informováni.

Z dat publikovaných Matte a kol. [7] vyplývá, že na českém webu není TCF běžně nasazené. Jediná zmíněná doména je *livesport.cz*. Při návštěvě tohoto webu jsme se pokusili nesouhlasit se zpracováním dat v rámci reklamních systémů (viz obr. 4). Použitím nástroje *Cookie Glasses* [7] jsme však zjistili, že ve skutečnosti si stránka uložila souhlas se všemi pěti účely zpracování definovanými TCF pro 544 firem zapojených do TCF (viz obr. 5).



Obrázek 4: Okno pro získání souhlasu se zpracováním osobních dat při prohlížení stránky *livesport.cz*. Okno může uživatel zobrazit až po načtení stránky, tedy až po komunikaci v rámci RTB.

Hlavní analýzu českého webu⁷ jsme provedli následovně:

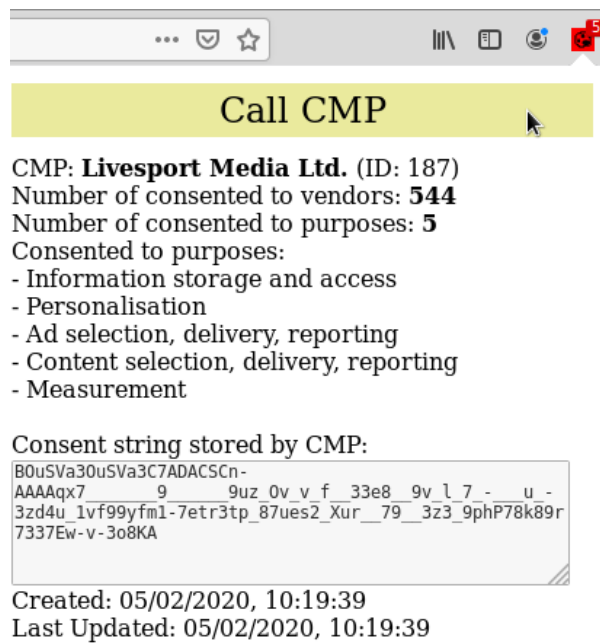
1. V datech zveřejněných 16.1.2020 službou *netmonitor.cz* jsme identifikovali 464 domén druhého řádu.
2. Z uvedených 464 domén jsme stáhli 367 souborů *ads.txt*⁸. Soubory *ads.txt* slouží nakupujícím v rámci RTB k ověření, že reklamu nakupují od subjektu, který má právo prodávat reklamní prostor na daném serveru. Umístění tohoto souboru je dobrovolné, takže se dá usuzovat, že nejméně 79% nejnavštěvovanějších českých webů prodává svůj reklamní prostor v rámci RTB.
3. Ze stažených souborů jsme extrahovali počty stránek prodávající na nám známých tržištích. Počty jsou uvedené v [tabulce 1](#).

⁷Data jsou dostupná na https://www.fit.vutbr.cz/~polcak/czech_rtb.zip

⁸<https://iabtechlab.com/ads-txt/>

⁹<https://brave.com/wp-content/uploads/2019/07/Scale-billions-of-bid-requests-per-day-RAN2019061811075588.pdf>

¹⁰Např. <https://www.slideshare.net/iBILLBOARD/future-with-ibillboard-january-2013>



Obrázek 5: *livesport.cz* nerespektuje volbu uživatele a ukládá falešné souhlasy se zpracováním.

Název tržiště	Počet webů	Inzerovaný zákonný důvod pro výběr reklamy
Oath/AOL (aol.com) ⁹	101	Oprávněné zájmy
openx.com ⁹	277	Souhlas
ibillboard.com ¹⁰	297	Oprávněné zájmy
rubiconproject.com ⁹	299	Oprávněné zájmy
indexexchange.com ⁹	300	Oprávněné zájmy
appnexus.com ⁹	335	Oprávněné zájmy
pubmatic.com ⁹	336	Oprávněné zájmy
DoubleClick (google.com) ⁹	337	Není zapojený do TCF

Tabulka 1: Počet stránek předávající osobní data jednotlivým tržištím a právní základ zpracování.

Po právní stránce je velkou otázkou, zda tržiště uvedená v [tabulce 1](#) mohou předávat osobní data společnostem bez souhlasu subjektu údajů. Jsou oprávněné zájmy tržišť RTB pubmatic.com, appnexus.com, indexexchange.com, rubiconproject.com, ibillboard.com a Oath/AOL v souladu s očekáváním subjektů údajů ohledně zpracování pro účely výběru reklamy a prošly tedy balančním testem oprávněných zájmů GDPR? Podle ICO by zpracování v rámci RTB mělo být založeno na souhlasu [6, str. 18].

Jakým způsobem se společnosti vypořádaly s přenosem citlivých dat a osobních údajů dětí? Např. společnost PubMatic ve svých podmínkách užívání služeb žádá, aby se skrze jeho služby nezasílaly citlivé osobní údaje a dětem zakazuje užívání svých služeb. Opravdu platí, že si jsou uživatelé internetu vědomi zpracování citlivých dat [6]? Jsou uživatelé před takovýmto zpracováním varováni?

Z analyzovaných webů jsme vybrali skupinu, u kterých existuje předpoklad předávání citlivých údajů:

- *abctehotenstvi.cz* nezobrazuje uživateli žádné varování týkající se předávání osobních údajů, přitom využívá AppNexus a PubMatic.
- *ceskozdrape.cz* nezobrazuje uživateli žádné varování týkající se předávání osobních údajů,

přítom využívá všechna tržiště z [tabulky 1](#).

- *detsky-web.cz* nezobrazuje uživateli žádné varování týkající se předávání osobních údajů, přitom využívá všechna tržiště uvedená v [tabulce 1](#).
- *naseporodnice.cz* nezobrazuje uživateli žádné varování týkající se předávání osobních údajů, přitom využívá všechna tržiště z [tabulky 1](#) mimo Oath/AOL.
- *seznamka.cz* získává souhlas setrváním na stránkách, přitom využívá Google Doubleclick.
- *vitalia.cz* nezobrazuje uživateli žádné varování týkající se předávání osobních údajů, přitom využívá všechna tržiště z [tabulky 1](#) mimo Oath/AOL.
- *ulekare.cz* získává souhlas setrváním na stránkách, zobrazená lišta obsahuje souhlasné tlačítko, ale data se do RTB dostávají i bez kliknutí na toto tlačítko. Web využívá všechna tržiště uvedená v [tabulce 1](#).
- *zdrave.cz* získává souhlas setrváním na stránkách, zobrazená lišta obsahuje tlačítko rozumím, ale data se do RTB dostávají i bez kliknutí na toto tlačítko. Web využívá všechna tržiště z [tabulky 1](#) mimo Oath/AOL.
- *zdraveomlazeni.cz* získává souhlas setrváním na stránkách, zobrazená lišta obsahuje tlačítko v pořádku, ale data se do RTB dostávají i bez kliknutí na toto tlačítko. Web využívá všechna tržiště z [tabulky 1](#) mimo Oath/AOL.

K výše uvedenému seznamu stránek je potřeba poznamenat, že získání souhlasu setrváním na stránkách neumožňuje vyjádřit souhlas svobodný (není jasné, jak vyjádřit nesouhlas), konkrétní ke specifickému zpracování, či správci dat, ani jednoznačný (GDPR, čl. 4, odst. 11). Souhlas se zpracováním zvláštních kategorií osobních údajů má být výslovný (GDPR, čl. 9).

Závěr

Jak ukazuje analýza českých webů a společností zapojených do TCF, v celé řadě případů dochází ke zpracování osobních údajů, včetně citlivých, bez udělení odpovídajících souhlasů. To vzbuzuje otázku, zda opravdu platí, že je zajištěná vysoká úroveň ochrany osobních údajů, kterou zmiňuje nařízení GDPR. Pochyby přidávají i četné stížnosti podané k dohledovým úřadům zmíněné v úvodu článku. Z celoevropského pohledu je zajímavá studie nasazení TCF [7], která reportovala celou řadu pochybení v získávání souhlasů.

Závěrem je potřeba upozornit, že souhlasy s personalizací reklamy a dalšími účely zpracování je možné získávat i nerušivými prostředky [11] jako distribucí rozšíření prohlížečů, které budou uživatele transparentně informovat o účelech zpracování. Na stránky také lze vhodně umístit tlačítka, která v případě zájmu otevřou dialog pro udělení informovaných souhlasů se zpracováním.

Poděkování

Tento příspěvek vznikl za podpory projektu VI20172020062 financovaného Ministerstvem vnitra ČR.

Reference

- [1] Adbrain: Demystifying Cross-Device. 2015.
URL https://www.iabuk.com/sites/default/files/public_files/Adbrain_Demystifying_Cross_Device.pdf

- [2] Bashir, M. A.; Farooq, U.; Shahid, M.; aj.: Quantity vs. Quality: Evaluating User Interest Profiles Using Ad Preference Managers. In *Network and Distributed Systems Security (NDSS) Symposium 2019*, 2019, ISBN 1-891562-55-X.
- [3] Bashir, M. A.; Wilson, C.: Diffusion of User Tracking Data in the Online Advertising Ecosystem. *Proceedings on Privacy Enhancing Technologies*, ročník 2018, č. 4, 2018: s. 85–103, ISSN 2299-0984.
- [4] CNIL — Commission Nationale de l’Informatique et des Libertés: APPLICATIONS MOBILES : clôture de la mise en demeure à l’encontre de la société VECTAURY. 2019. URL <https://www.cnil.fr/fr/applications-mobiles-cloture-de-la-mise-en-demeure-lencontre-de-la-societe-vectaury>
- [5] IAB Tech Lab: AdCOM Specification v1.0. 2018. URL <https://github.com/InteractiveAdvertisingBureau/AdCOM/blob/master/AdCOM%20v1.0%20FINAL.md>
- [6] ICO — Information Commissioner’s Office: Update report into adtech and real time bidding. 2019. URL <https://ico.org.uk/media/about-the-ico/documents/2615156/adtech-real-time-bidding-report-201906.pdf>
- [7] Matte, C.; Bielova, N.; Santos, C.: Do Cookie Banners Respect my Choice? Measuring Legal Compliance of Banners from IAB Europe’s Transparency and Consent Framework. 2019.
- [8] Ryan, J.: Behavioural advertising and personal data. 2018. URL <https://brave.com/wp-content/uploads/2018/09/Behavioural-advertising-and-personal-data.pdf>
- [9] Soudní dvůr Evropské unie: Rozsudek Soudního dvora (velkého senátu), věc C-210/16. 2018.
- [10] Speicher, T.; Ali, M.; Venkatadri, G.; aj.: Potential for Discrimination in Online Targeted Advertising. In *FAT 2018 - Conference on Fairness, Accountability, and Transparency*, ročník 81, New-York, United States, 2018, s. 1–15.
- [11] WP29 — Article 29 Data Protection Working Party: Opinion 16/2011 on EASA/IAB Best Practice Recommendation on Online Behavioural Advertising. 2011, WP188, https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2011/wp188_en.pdf.
- [12] Yuan, Y.; Wang, F.; Li, J.; aj.: A survey on real time bidding advertising. In *Proceedings of 2014 IEEE International Conference on Service Operations and Logistics, and Informatics*, 2014, s. 418–423.