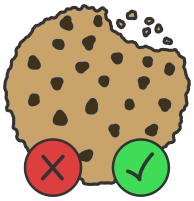
BRNO UNIVERSITY OF TECHNOLOGY
FACULTY OF INFORMATION TECHNOLOGY

Excel@FIT 2021
http://excel.fit.vutbr.cz

Zobrazení a úprava informací v Transparency and Consent Framework

Aleš Postulka*



Abstrakt

Cílem této práce je návrh a implementace rozšíření pro webové prohlížeče Mozilla Firefox a Google Chrome. Účelem rozšíření je umožnění automatizované správy poskytnutých souhlasů se zpracováním osobních údajů. Souhlas se zpracováním osobních údajů je kvůli evropskému nařízení GDPR potřeba poskytovat na většině webových stránek. Za účelem usnadnění získávání souhlasu se zpracováním osobních údajů byl organizací IAB Europe vytvořen rámec Transparency and Consent Framework (TCF), který mimo jiné poskytuje rozhraní pro přístup k informacím o uloženém souhlasu. Vytvořené rozšíření interaguje s webovými stránkami využívajícími TCF. Umožňuje automatické poskytnutí souhlasu na základě nastavení, a také zobrazení informací o poskytnutém souhlasu. Rozšíření je publikováno v obchodu s doplňky pro prohlížeč Mozilla Firefox.

Klíčová slova: Rozšíření pro webové prohlížeče — Transparency and Consent Framework — Consent Management Platform — Souhlas se zpracováním osobních údajů

Přiložené materiály: GitHub repozitář — Obchod s doplňky

*xpostu03@stud.fit.vutbr.cz, *Fakulta informačních technologií, Vysoké učení technické v Brně*

1. Úvod

Od doby, kdy vstoupilo v platnost nařízení Evropské unie GDPR, se začaly na webových stránkách objevovat bannery nebo vyskakovací okna vyžadující poskytnutí souhlasu se zpracováním osobních údajů. V tak velkém množství mohou být tyto bannery spíše obtěžující a uživatelé jim tak nemusí věnovat příliš velkou pozornost. To může vést k situaci, kdy uživatel klikne na tlačítko potvrzující souhlas bez vědomí, k čemu konkrétně souhlas poskytl.

Získávání souhlasu se zpracováním osobních údajů i samotné zpracování osobních údajů se musí řídit nařízením GDPR. Za účelem usnadnění dodržení tohoto nařízení vytvořila organizace Internet Advertisement Bureau Europe rámec Transparency and Consent Framework [1], jehož cílem je vytvořit standard pro získání a využití souhlasu se zpracováním osobních údajů. Transparency and Consent Framework mimo jiné definuje účely zpracování osobních údajů, formát uložení souhlasu a API pro přístup k informacím o poskytnutém souhlasu. Transparency and Consent Framework je popsán více v sekci 2.

Cílem této práce je návrh a implementace rozšíření pro webové prohlížeče, které bude umožňovat zobrazení informací uložených webovou stránkou využívající Transparency and Consent Framework. Toto rozšíření bude dále umožňovat uživateli předem nastavit souhlas případně nesouhlas s jednotlivými účely zpracování osobních údajů. Dle těchto preferencí by poté měl být souhlas se zpracováním osobních údajů udělován automaticky bez potřeby dalšího zásahu uživatele.

2. Transparency and Consent Framework

Transparency and Consent Framework (TCF) byl vytvořen asociací IAB Europe jako reakce na vydání GDPR. TCF je aktuálně dostupný ve verzi 2.0 [1].

Úkolem TCF je pomoci všem stranám zainteresovaným do digitální reklamy zajistit dodržení GDPR a směrnice ePrivacy při zpracování osobních údajů, přístupu nebo ukládání informací do zařízení uživatele, mezi které patří například cookies, reklamní identifikátory, identifikátory zařízení apod. [1].

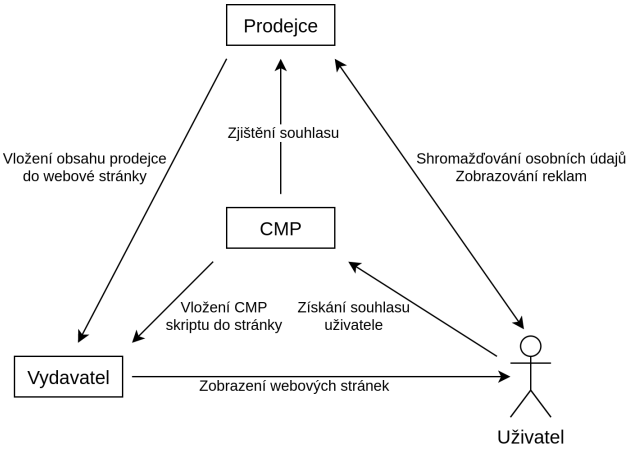
TCF poskytuje prostředí, ve kterém mohou vydavatelé webových stránek sdělit uživateli, jaká data budou shromažďována a jakým způsobem je bude daná webová stránka, jej provozovatel, a společnosti, se kterými spolupracuje, využívat [1].

Transparency and Consent Framework rozlišuje tři účastnící se strany [2] jejichž vzájemné vazby jsou ilustrovány na obrázku 1.

- *Vydavatel (Publisher)* – provozovatel digitálního obsahu (webové stránky, aplikace, apod.), ve kterém dochází ke shromažďování a zpracování osobních údajů. Je zodpovědný za zobrazení uživatelského rozhraní frameworku uživateli a za vytvoření právního základu pro prodejce, kteří mohou zpracovávat osobní údaje návštěvníků digitálního obsahu vydavatele.
- *Prodejce (Vendor)* – společnost, která se podílí na zobrazování reklamy v digitálním obsahu vydavatele.
- *Platforma pro správu souhlasů (CMP – Consent Management Platform)* – prostředník mezi vydavatelem, prodejci a koncovým uživatelem, který usnadňuje ustavení právního základu a získání souhlasu se zpracováním osobních údajů. Typicky také pod jménem vydavatele zobrazuje uživateli uživatelské rozhraní pro získání souhlasu se zpracováním osobních údajů.

Pro uložení souhlasu uživatele je v TCF definován *Transparency and Consent String (TC String)* [3]. Tento řetězec je vytvářen za účelem zapouzdření všech důležitých informací týkajících se uděleného souhlasu. Potřebné údaje se transformují do bitového vektoru, který je následně zakódován pomocí algoritmu Base64.

K uloženým informacím je možné přistupovat pomocí *Consent Management Platform API (CMP API)* [4]. Jedná se o rozhraní, kterým CMP umožňuje přístup k informacím o transparentnosti a o souhlasu získaném od uživatele.

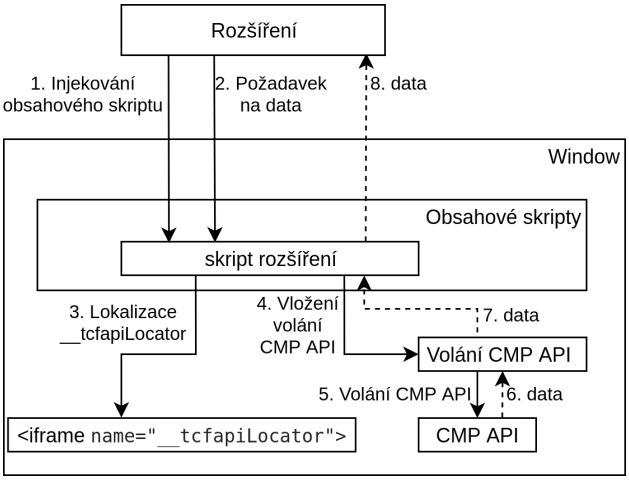


Obrázek 1. Vzhaty mezi jednotlivými stranami zainteresovanými do TCF.

3. Návrh rozšíření

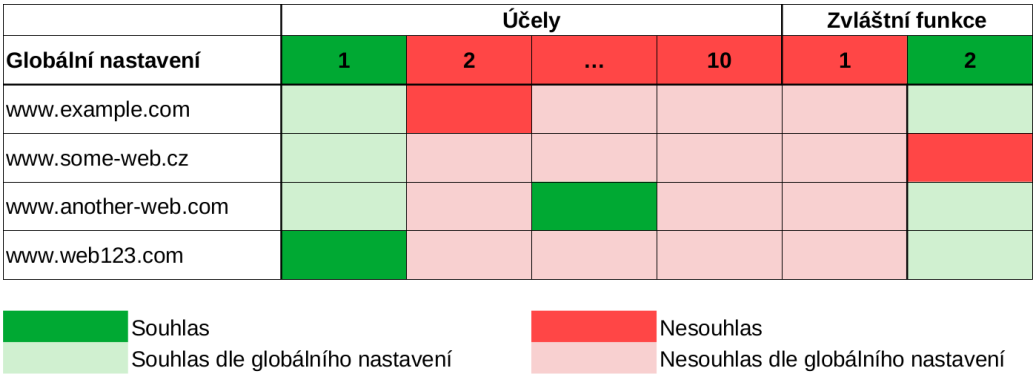
V roce 2020 vytvořil Matte a kol. [5] rozšíření *Cookie Glasses*. Rozšíření bylo vytvořeno za účelem zobrazování informací uložených pomocí TCF. Rozšíření je však nyní již prakticky nepoužitelné, protože bylo vytvořeno k TCF verzi 1.1. Této verzi již byla ukončena podpora a není pro ni dostupný ani seznam prodejců ani seznam CMP.

Rozšíření vytvářené v rámci této práce bude založeno na aktuální verzi TCF, tedy na verzi 2.0. Základ rozšíření bude podporovat dva jazyky – češtinu a angličtinu. Pro popisy jednotlivých účelů zpracování budou využity oficiální překlady [6]. Tyto popisy budou poskytovány ve všech jazycích, pro které jsou překlady vytvořeny. Jazyk rozšíření bude volen na základě jazyka prohlížeče, na němž bude toto rozšíření nainstalováno. Výchozím jazykem pak bude angličtina.



Obrázek 2. Ilustrace komunikace mezi rozšířením a CMP API.

Na obrázku 2 je znázorněn postup získávání informací z TCF. Rozšíření vloží do webové stránky obsahový skript. Na tento skript budou posílány poža-



Obrázek 3. Návrh tabulky s nastavením souhlasu. Čísla v řádku „Globální nastavení“ označují jednotlivé účely a zvláštní funkce zpracování osobních údajů jejichž popisy jsou uvedeny v příloze A.

davky na získání informací o uloženém souhlasu. Po obdržení požadavku skript ověří, zda stránka využívá TCF, a to pomocí lokalizace elementu `iframe` s atributem `name="__tcfapiLocator"`. Poté bude s využitím tohoto elementu lokalizován rámec obsahující CMP API, které by se mělo nacházet v rodičovském uzlu nalezeného elementu. Obsahový skript následně vloží přímo do stránky skript s voláním CMP API. Získaná data poté zašle samotnému rozšíření, které je podle potřeby zpracuje.

Aktuální informace o uděleném souhlasu pro danou webovou stránku budou získávány pomocí požadavku `getTCData` a zobrazovány ve vyskakovacím okně rozšíření, jehož návrh je na obrázku 4.

Obrázek 4. Návrh vyskakovacího okna pro zobrazení informací uložených v TCF.

Nastavování preferencí pro udělování souhlasu či nesouhlasu k jednotlivým účelům zpracování bude uživatel provádět na stránce nastavení rozšíření pomocí tabulky znázorněné na obrázku 3. Tabulka je inspirována nastavením v rozšíření uMatrix [7] a umožňuje uživateli vytvořit výchozí nastavení udělování souhlasů pro jednotlivé účely zpracování, ale také specifikovat své preference pro již navštívené webové stránky, u nichž byla detekována přítomnost TCF.

Při navštívení webové stránky využívající TCF bude na základě uživatelských preferencí vytvořen souhlas se zpracováním osobních údajů. Vytvořený souhlas bude následně uložen jako cookie s názvem *euconsent-v2* a *eupubconsent-v2*. Dále budou prohledány všechny cookies aktuální webové stránky, a pokud bude v jejich hodnotě nalezen souhlas se zpracováním osobních údajů, bude nahrazen souhlasem vytvořeným rozšířením. Dále budou upravovány souhlasy v HTTP odpovědích ze serveru *consensu.org*, který je, dle definice TCF, jednou z možností pro uložení souhlasu.

4. Implementace

Jedním z cílů bylo vytvoření rozšíření nezávislého na konkrétním webovém prohlížeči. Rozšíření využívá knihovnu *WebExtension browser API Polyfill*[1], která usnadňuje implementaci nezávislou na konkrétním prohlížeči usnadňuje a umožňuje využití objektů `Promise` při použití asynchronních funkcí z *WebExtensions API*. Rozšíření je vytvořeno primárně pro webové prohlížeče Mozilla Firefox a Google Chrome. Rozšíření by mělo být kompatibilní i s prohlížeči Microsoft Edge a Opera. Na těchto prohlížečích však nebylo testováno.

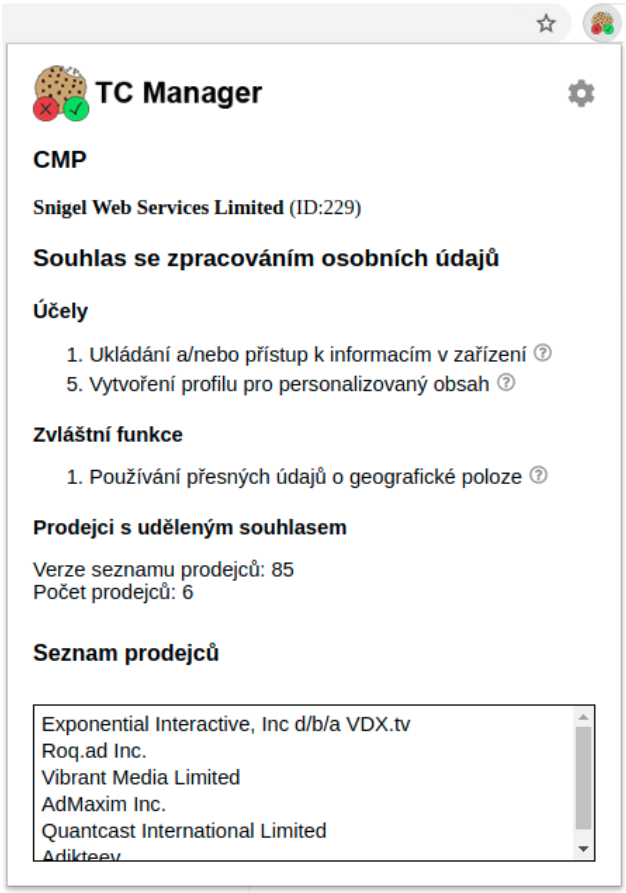
4.1 Zobrazení informací

Zobrazení informací o poskytnutém souhlasu se zpracováním osobních údajů pro aktuální webovou stránku je implementováno jako vyskakovací okno rozšíření, které je možné zobrazit kliknutím na ikonu rozšíření v panelu nástrojů.

Skript vyskakovacího okna zašle požadavek obsahovému skriptu, který získá informace způsobem popsaným již v sekci 3. Skript s voláním CMP API vložený do stránky vrací získané informace obsahovému skriptu vyvoláním vlastní události (`CustomEvent`) a vložením informací do detailu této události.

[1] https://github.com/mozilla/webextension-polyfill/

Informace jsou následně zaslány skriptu vyskakovacího okna a zobrazeny tak jako na obrázku 5. V případě, že webová stránka nevyužívá TCF, je zobrazena pouze příslušná informační hláška.



Obrázek 5. Vyskakovací okno pro zobrazení informací uložených v TCF.

4.2 Vytvoření souhlasu

Proces udělení souhlasu začíná v obsahovém skriptu, protože k vytvoření souhlasu je potřeba získat základní údaje o CMP použitém na navštívené webové stránce. Obsahový skript po načtení webové stránky zjišťuje, zda stránka využívá TCF. Zjišťování přítomnosti TCF probíhá v intervalu 500 ms po dobu až 15 s. Opakovaná kontrola přítomnosti TCF je nutná z toho důvodu, že po načtení webové stránky může ještě nějakou dobu trvat, než jsou plně načteny skripty CMP API a než začnou poskytovat požadované informace. Po zjištění základních údajů jsou tyto zaslány skriptu běžícímu na pozadí, pro který jsou přijaté údaje zároveň signálem k vytvoření a uložení souhlasu. Nejprve jsou načteny již existující cookies a jejich hodnoty jsou pomocí regulárního výrazu hledány řetězce, které mohou potenciálně uchovávat souhlas se zpracováním osobních údajů. S využitím knihovny *tc-string-parse* je poté u každého nalezeného řetězce ověřováno, zda se skutečně jedná o řetězec nesoucí informace o souhlasu se zpracováním osobních údajů. Pokud je takový řetězec nalezen, je v cookie nahrazeno řetězcem vytvořeným na základě uživatelského nastavení a upravené cookie je opět uloženo. V opačném případě je souhlas vložen do „standardních“ cookies s názvy `euconsent-v2`, `eupubconsent-v2` a `cconsent-v2`, které byly v průběhu implementace identifikovány jako nejčastěji používané cookies pro uložení souhlasu se zpracováním osobních údajů.

Vynucování souhlasu dle nastavení uživatele je implementováno také pomocí zachytávání HTTP komunikace a úpravy HTTP odpovědí. Zachycovány jsou požadavky, jejichž URL odpovídá jednomu z definovaných vzorů. Zachycování veškeré HTTP komunikace by totiž bylo velmi neefektivní. Aktuálně je definováno zachytávání požadavků na subdomény domény *consensu.org* a požadavků na skript *zdconsent.js*. Doména *consensu.org* je jako možnost ukládání souhlasu definovaná přímo v TCF. Skript *zdconsent.js* je používán webovými stránkami využívajícími CMP *Evidon, Inc.* Ve skriptu byly objeveny předdefinované hodnoty řetězců nesoucích informace o souhlasu.

Samotná úprava odpovědi pak probíhá podobným způsobem jako při vyhledávání a nahrazování souhlasu v existujících cookies. Po načtení celé HTTP odpovědi jsou v této pomocí regulárního výrazu nalezeny všechny řetězce, které jsou potenciálními nositeli souhlasu. Tyto řetězce jsou poté otestovány, zda se skutečně jedná o zakódovaný souhlas. V případě, že se o souhlas jedná, je tento řetězec nahrazen řetězcem vytvořeným na základě uživatelského nastavení. Po zkontrolování a případném nahrazení všech nalezených řetězců je upravená odpověď zaslána webové stránce, která odeslala požadavek.

Úprava HTTP odpovědí je, vzhledem ke kompatibilitě použitých funkcí z WebExtensions API, aktivní pouze v prohlížeči Mozilla Firefox.

5. Testování

Rozšíření bylo otestováno na 94 webových stránkách, které využívají TCF. Vlastní souhlas se povedlo úspěšně vložit na 73 z nich. Za úspěšné vložení vlastního souhlasu byl považován stav, kdy rozhraní CMP API vrátilo informace shodující se s vkládaným souhlasem (s nastavením uživatele). Při neúspěšném vložení souhlasu byl analyzován způsob uložení souhlasu dané webové stránky.

Tabulka 1 zobrazuje zastoupení jednotlivých CMP v testovaných stránkách. Poslední sloupec tabulky ukazuje úspěšnost vložení vlastního souhlasu pro stránky které využívají dané CMP. Z tabulky lze vypozorovat, že nejčastěji používanými CMP jsou *OneTrust*

CMP	Počet stránek	Úspěšně vložený souhlas
OneTrust LLC	25	25
Quantcast International Limited	16	16
Sourcepoint Technologies, Inc.	13	0
LiveRamp	6	6
consentmanager.net	4	3
Didomi	4	1
Google LLC	3	3
1&1 Mail & Media GmbH	2	2
Conversant Europe Ltd.	2	2
Evidon, Inc.	2	*2
Healthline Media, Inc.	2	2
iubenda	2	2
Seznam, a.s.	2	2
TrustArc Inc	2	2
Ensighten, Inc	1	0
Associated Newspapers Ltd	1	0
CIVIC COMPUTING LTD	1	1
Cookiebot	1	1
DAILYMOTION SA	1	1
Farlex Inc	1	0
Snigel Web Services Limited	1	1
System1 LLC	1	0
Wikia, Inc.	1	1
Celkem	94	73 (77,7 %)

Tabulka 1. Zastoupení jednotlivých CMP v testovaných stránkách a úspěšnost vložení vlastního souhlasu.
* - Pouze Mozilla Firefox.

LLC, *Quantcast International Limited* a *Sourcepoint Technologies, Inc.* Zatímco u prvních dvou zmíněných se podařilo vložit souhlas u všech webových stránek, u třetího nebyl souhlas vložen úspěšně ani jednou. Po analýze webových stránek využívajících CMP *Sourcepoint Technologies, Inc.* bylo zjištěno, že toto CMP ukládá souhlas do *místního uložiště (Local storage)*, ke kterému rozšíření nemá přístup, a proto není možné vložit vlastní souhlas. Stejná příčina byla zjištěna u dalších třech webových stránek s jinými CMP. U dvou webových stránek byl zobrazován banner pokaždé, kdy došlo ke změně souhlasu ze strany rozšíření a CMP API vracelo informace o tom, že souhlas nebyl udělen. Tyto stránky pravděpodobně ukládají souhlas na více míst a provádějí porovnání. Při nekonzistenci je pak zobrazen banner žádající o udělení souhlasu. V jenom případě pak CMP API vracelo pouze hodnotu `undefined`, a tak ani nebylo možné informace o uděleném souhlasu získat.

U žádné z testovaných stránek se nepodařilo vložit vlastní souhlas ihned při prvním načtení stránky. Ve většině případů bylo potřeba stránku po jejím načtení ještě alespoň jednou obnovit. K tomu dochází z důvodu, že CMP API načítá informace o souhlasu pouze po načtení stránky. Upravený souhlas si tak rozhraní načte až po obnovení stránky. U stránek, které nevyužívají „standardní“ cookies (euconsent-v2, eupubconsent-v2, cconsent-v2) je potřeba udělit souhlas manuálně pomocí banneru, aby bylo cookie se souhlasem vytvořeno a mohlo být poté přepsáno vlastním souhlasem. U těchto stránek je ve většině po manuálním udělení souhlasu ještě jednou až dvakrát stránku obnovit, aby CMP API reflektovalo vložený souhlas.

Některé stránky také využívají pomocné cookie, jehož existence, nebo hodnota určuje, zda byl banner vyžadující souhlas uzavřen či nikoliv. Banner se tak zobrazuje i po obnovení stránky i přesto, že CMP API vrací správné informace. V takovém případě je také nutné udělit souhlas manuálně pomocí banneru. Rozšíření vytváří, společne se souhlasem také cookie *OptanonAlertBoxClosed*. Toto cookie zabraňuje na stránkách s CMP *OneTrust LLC* zobrazování banneru i po úspěšném vynucení vlastního souhlasu.

U stránek s CMP *Quantcast International Limited* docházelo ke zobrazování banneru při každé změně souhlasu. Rozhraní CMP API však vracelo správné informace. Zobrazování banneru je tak na stránkách s tímto CMP zcela blokováno.

6. Závěr

Cílem této práce bylo navrhnout a implementovat vícejazyčné rozšíření pro webové prohlížeče pro automatické poskytování souhlasu a zobrazování informací o souhlasu se zpracováním osobních údajů na webových stránkách využívajících TCF. Účelem rozšíření je odstranit nutnost manuálního udělování souhlasu se zpracováním osobních údajů, které může být, vzhledem k výskytu na velkém množství webových stránek, pro uživatele obtěžující.

Pro získávání informací o uděleném souhlasu je využíváno rozhraní CMP API. Vytvořený souhlas se zpracováním osobních údajů je pak ukládán do cookies s názvy `euconsent-v2`, `eupubconsent-v2` a `cconsent-v2`. Dále jsou prohledávány všechny cookies dané webové stránky a jejich hodnoty jsou kontrolovány na přítomnost řetězce nesoucího informace o souhlasu se zpracováním osobních údajů. Pro prohlížeč Mozilla Firefox je implementováno také zachytávání HTTP komunikace a přepisování řetězců se souhlasem vlastní hodnotou.

Implementované rozšíření bylo otestováno na 94 webových stránkách s úspěšností vložení souhlasu 77,9 %. Testování ukázalo, že pro reflektování vloženého souhlasu je potřeba webovou stránku vždy alespoň jednou obnovit. Dále bylo zjištěno, že některé CMP ukládají souhlas do místního úložiště, kde rozšíření nemá přístup a nemůže tak vynutit vlastní souhlas.

Poděkování

Chtěl bych poděkovat Ing. Liboru Polčákovi Ph.D. za ochotu a cenné rady v průběhu tvorby této práce.

Literatura

[1] IAB Europe. TCF – Transparency & Consent Framework. Online. https://iabeurope.eu/transparency-consent-framework/.

[2] IAB Europe. IAB Europe Transparency & Consent Framework Policies. Online. https://iabeurope.eu/iab-europe-transparency-consent-framework-policies/.

[3] IAB Europe. Transparency and Consent String with Global Vendor & CMP List Formats. Online. https://github.com/InteractiveAdvertisingBureau/GDPR-Transparency-and-Consent-Framework/blob/master/TCFv2/IAB%20Tech%20Lab%20-%20Consent%20string%20and%20vendor%20list%20formats%20v2.md.

[4] IAB Europe. Consent Management Platform API. Online. https://github.com/InteractiveAdvertisingBureau/GDPR-Transparency-and-Consent-Framework/blob/master/TCFv2/IAB%20Tech%20Lab%20-%20CMP%20API%20v2.md.

[5] Célestin Matte, Nataliia Bielova, and Cristiana Santos. Do Cookie Banners Respect my Choice? Measuring Legal Compliance of Banners from IAB Europe's Transparency and Consent Framework. Online, 2020. https://arxiv.org/abs/1911.09964.

[6] IAB Europe. List of translations for purpose descriptions v2.0. Online. https://register.consensu.org/Translation.

[7] Raymond Hill. uMatrix. Online. https://github.com/gorhill/uMatrix.

[8] Sdružení pro internetový rozvoj. Příloha A: Definice účelů a funkcí. Online. https://www.spir.cz/sites/default/files/prilohy/Priloha_A_TCF_v.2.0.pdf.

Přílohy

A Účely a funkce zpracování osobních údajů

Uvedené popisy byly převzaty z [8].

Účely

1. Ukládání a/nebo přístup k informacím v zařízení.
   - *Popis:* Ve vašem zařízení se mohou ukládat soubory cookie, identifikátory zařízení nebo další informace nebo k nim ve vašem zařízení může být umožněn přístup pro účely, které vám byly sděleny.
2. Základní nastavení reklamy.
   - *Popis:* Reklamy se mohou zobrazovat na základě obsahu, který prohlížíte, aplikace, kterou používáte, vaší přibližné polohy nebo typu vašeho zařízení.
3. Vytvoření profilu pro personalizovanou reklamu.
   - *Popis:* Na základě vašeho chování na internetu může být vytvořen váš profil, aby vám mohla být zobrazena relevantní reklama.
4. Výběr personalizované reklamy.
   - *Popis:* Na základě vašeho profilu se vám může zobrazovat personalizovaná reklama.
5. Vytvoření profilu pro personalizovaný obsah.
   - *Popis:* Na základě vašeho chování na internetu může být vytvořen váš profil, a to proto, aby vám mohl být zobrazován obsah, který je pro vás relevantní.
6. Výběr personalizovaného obsahu.
   - *Popis:* Na základě vašeho profilu se vám může zobrazovat personalizovaný obsah.
7. Měření výkonu reklamy.
   - *Popis:* Výkon a účinnost reklamy, kterou vidíte nebo na kterou reagujete, mohou být měřeny.
8. Měření výkonu obsahu.
   - *Popis:* Účinnost a výkon obsahu, který vidíte nebo na který reagujete, mohou být měřeny.
9. Používání výzkumu trhu pro získání poznatků o uživatelích.
   - *Popis:* Výzkum trhu lze využít k získání více informací o uživatelích, kteří navštěvují stránky/aplikace a jimž jsou zobrazeny reklamy.
10. Vývoj a zlepšování produktů.
    - *Popis:* Vaše údaje lze využít ke zlepšení stávajících systémů a softwaru a k vývoji nových produktů.

Zvláštní funkce

1. Používání přesných údajů o geografické poloze.
   - *Popis:* Přesné údaje o vaší geografické poloze lze použít za účelem podpory jednoho nebo více účelů. Znamená to, že vaši polohu lze určit s přesností na několik metrů.
2. Aktivní vyhledávání identifikačních údajů v rámci vlastností zařízení.
   - *Popis:* Vaše zařízení lze identifikovat na základě prohledání jedinečné kombinace vlastností vašeho zařízení.