Privacy from 5 PM to 6 AM: Tracking and Transparency Mechanisms in the HbbTV Ecosystem

Christian Böttger*, Henry Hosseini*†, Christine Utz‡, Nurullah Demir*, Jan Hörnemann*§, Christian Wressnegger¶, Thomas Hupperich‖, Norbert Pohlmann*, Matteo Große-Kampmann**, and Tobias Urban*
*Institute for Internet Security, Westphalian University of Applied Sciences, *lastname@internet-sicherheit.de*
†Department of Information Systems, University of Münster, *henry.hosseini@wi.uni-muenster.de*
‖Institute for Geoinformatics, University of Münster, *thomas.hupperich@uni-muenster.de*
‡Radboud University, *christine.utz@ru.nl*
§AWARE7 GmbH, *jan@aware7.de*
¶Karlsruhe Institute of Technology, KASTEL Security Research Labs, *c.wressnegger@kit.edu*
**Rhine-Waal University of Applied Sciences, *matteo.grosse-kampmann@hochschule-rhein-waal.de*

***Abstract*—Hybrid broadcast broadband television (HbbTV) is an evolving technology that connects linear TV with modern HTML5 applications, delivering extras like games, videos, and online shopping. However, its bidirectional transmission functionality raises privacy concerns, as it introduces new tracking methods for TV channels. While previous studies focused on security issues or user awareness of HbbTV privacy challenges, a detailed examination of the tracking and transparency mechanisms of the HbbTV ecosystem is still missing. This study fills this gap by extensively analyzing these features within the European HbbTV ecosystem, and in particular within German-language TV channels. We monitored more than 350 TV channels for over 400 hours, evaluating 1) prevalent HbbTV tracking methods, 2) consent notice prevalence and user interactions, and 3) privacy policy disclosures. Our findings indicate that the HbbTV tracking system operates independently of the Web, consent notices exploit system constraints to influence users, and privacy policies often do not align with actual data practices.**

***Index Terms*—HbbTV, smart TV, privacy, privacy policies**

I. Introduction

Television is a vital medium for mass communication and entertainment. The number of European households with TVs [70] and the number of TV viewers [69] are still growing despite modern alternatives like Internet streaming platforms (e.g., Netflix). In contrast to Internet-based alternatives, a TV channel is broadcast linearly, binding watchers to a fixed programming schedule without any options to interact with the program. To overcome this limitation and to offer content more flexibly, the *Hybrid Broadcast Broadband TV* (HbbTV) [24] standard was introduced in 2006.

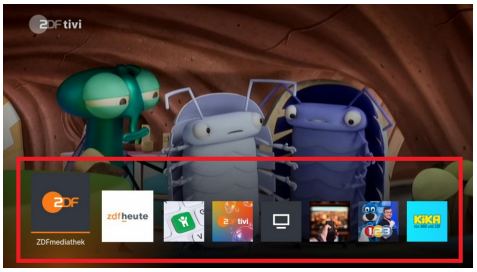
HbbTV is a technology that extends linear television to include on-demand HTML5 content to enrich the user experience. Examples are background information on popular shows or entire media centers offering on-demand video content. HbbTV differs from smart TVs or streaming platforms (e.g., Disney+) as it is dependent on the current linear programming and content that is distributed by the channel operator. TV channels provide this additional HbbTV content via HTTP(S), requiring the TV to be connected to the Internet. If the TV is not connected to the Internet, the linear program is shown to users without HbbTV content. The main difference between streaming and other over-the-top applications run on a smart TV is that HbbTV enhances traditional TV with Internet-based features, while over-the-top and streaming services operate independently of traditional TV and rely solely on the Internet for content delivery. The HbbTV standard has been widely adopted in Europe and other regions like Australia. Most households in these countries have a TV that supports HbbTV [23]. Germany has the highest HbbTV adoption rate in households (38 million) and HbbTV channels [18].

While HbbTV offers new opportunities for broadcasters and users, it comes with security and privacy threats [10], [17], [35], [71]. HbbTV is bidirectional [31], enabling channel operators to collect and process information about users and their devices by utilizing standard Web technology and associated tracking methods. If channels collect personal data of viewers in the European Union (i.e., the channel is aired in the EU), the providers must comply with EU privacy legislation, in particular with the General Data Protection Regulation (GDPR) [28] and the ePrivacy Directive [26].

Prior work covered security aspects of the HbbTV standard [10] and privacy issues originating from security issues [33]. Most recently, Tagliaro et al. [71], [72] conducted a high-level analysis of the privacy implications of 36 HbbTV channels by counting trackers and the presence of privacy policies, performing off-TV interactions with consent notices in HbbTV applications, and surveying users' awareness of privacy threats in the HbbTV landscape. We extend these studies by analyzing HbbTV channels available in Germany along the following three dimensions: (I) an analysis of collected personal data and the tracking ecosystem, (II) an analysis of the consent notices shown to TV viewers, and (III) an analysis of the privacy policies presented to TV viewers.

In summary, we produce the following findings:

- **Insights into the HbbTV tracking ecosystem.** We show that the HbbTV tracking ecosystem is independent of the Web tracking ecosystem, as implied by the involvement of



(a) Example of HbbTV content (red rectangle) provided by a public German television channel.

(b) Colored buttons on a TV remote control.

Fig. 1: HbbTV service (left) accessed via colored buttons (right).

different entities, and demonstrate the limited effectiveness of available protection mechanisms (see Section V).

- **Investigation of consent notices.** We provide insights into the landscape of consent notices for HbbTV. They are less common than on the Web, originate from few issuers, and use HbbTV input constraints for nudging (see Section VI).
- **Analyzing privacy policies.** We assess the disclosures of 2,656 privacy policies in the HbbTV ecosystem and compare them with the channels' observed traffic (see Section VII).

II. Hybrid Broadcast Broadband TV

Hybrid Broadcast Broadband TV (HbbTV) is a standard developed by the European Telecommunications Standards Institute to unite the delivery of Internet-based content and linear TV programs. HbbTV is designed for devices equipped with a decoder to show digital television (broadcast) and Internet connectivity (broadband) to run interactive applications. In this context, interactive applications refer to content delivered in addition to the TV program. Examples of such applications include video-on-demand services, electronic program guides, or ads. Such content is often presented as an overlay on the running TV program or replaces the program entirely so that the program is not visible or audible anymore. Figure 1a provides an example of HbbTV content delivered by the German channel ZDF. The latest HbbTV major version (2.0) was published in 2015 and made substantial changes to the HbbTV ecosystem, including options to display HTML5 content on TVs.

Prevalence of HbbTV. According to the official HbbTV website, the technology is most prevalent in Europe, Russia, and parts of Oceania (i.e., Australia and New Zealand). Within the EU, TV channels in 17 countries (62% of EU member states) have adopted HbbTV. In Eastern Europe, most TV channels have not yet adopted the technology. In most countries with channels supporting the technology, however, most of the sold TVs support HbbTV (e.g., 90% of TVs in Germany) [23].

Technical Implementation of HbbTV. To indicate that a TV channel offers HbbTV content, the URLs of available HbbTV applications are encoded into the linear broadcast signal. If a TV supports the standard, it can connect and load the application, typically via HTTP. To run an application, each TV must implement a compatible runtime environment to render and execute it. In practice, these environments are web browser-like applications capable of displaying HTML5 pages, executing JavaScript code, and processing other elements of web applications. Figure 2 provides an overview of how HbbTV content is displayed on top of the linear TV program and shows a high-level workflow of HbbTV operation.

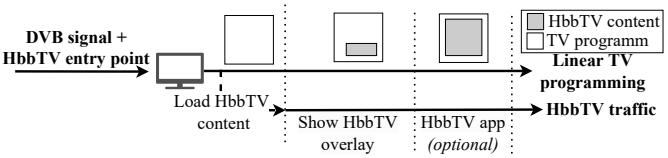


Fig. 2: Workflow of how HbbTV content is displayed.

Interaction with Colored Buttons. The HbbTV standard proposes using four colored buttons (red, blue, yellow, and green) to enable interaction with the TV and applications [29]. The HbbTV specification defines the functions as follows *"red color button: usually displays or hides a broadcast-related autostart application"* and *"3 additional color buttons (green, yellow, blue): variable usage as defined by the application"* [29]. For example, the TV will start a specific HbbTV application once a specific colored button is pressed. Most remote controls of HbbTV devices directly offer these buttons (see Figure 1b).

Difference to Streaming Technologies. HbbTV is both similar to and different from streaming services like Netflix or over-the-top TV streaming devices such as Amazon Fire TV. Streaming services and over-the-top streaming devices offer proprietary services, which require a subscription or one-time payment and sometimes necessitate specific hardware that mostly only works with an Internet connection. By contrast, HbbTV applications are loaded along with regular linear TV programming. Without an Internet connection, users can still watch TV but would not see the HbbTV content. Thus, HbbTV is an alternative to streaming services and vice versa. This is the technical difference between HbbTV and smart TVs. As HbbTV extends "linear TV" (e.g., channels like Nickelodeon, CNN, or MTV) with HTML5, the content is simply not loaded if the user has no Internet connection. More specifically, the entry point (i.e., a URL) for the HbbTV content is encoded into the broadcast TV signal. If the TV is not connected to the Internet, the broadcast program can still be viewed. By contrast, if the TV is connected to the Internet, it will load the additional content (e.g., an HTML5 overlay for the current program). Internet-connected applications (e.g., Netflix or Disney+) run by a smart TV require an Internet connection for all content [60]. The supplementary material of our work (Section IX) provides further examples.

III. Related Work

Our work extends prior privacy research that investigated over-the-top TV streaming, HbbTV, privacy policies, and consent notices.

Over-the-Top TV Streaming. Moghaddam et al. [55] and Varmarken et al. [83] analyzed privacy issues in over-the-top

TV streaming services, revealing extensive data collection practices. Tileria et al. [76] highlighted privacy harms in the Android TV ecosystem. While our focus on HbbTV appears related, the underlying technology differs significantly. The aforementioned services are fully Internet-based and deliver content over the Internet without relying on broadcast signals, so the technical comparability with HbbTV is limited.

Security and Privacy of HbbTV. Previous works [17], [33]–[35] identified conceptual privacy and security issues in HbbTV, such as unencrypted traffic. Tagliaro et al. [71], [72] showed limited progress in protecting user privacy over time. In comparison to our approach, they qualitatively assessed tracking activities on 36 TV channels. Our work significantly improves this research by quantitatively analyzing tracking on a real TV at scale and transparently mapping the HbbTV tracking ecosystem (e.g., in terms of active entities, used tracking techniques, and collected data), which significantly differs from other ecosystems (e.g., the Web). Our work is the first to analyze consent notices and privacy policies in the HbbTV ecosystem, while Tagliaro et al. communicated consent via deprecated Do Not Track signals [84]. Unlike previous studies, we use a rooted TV to access and analyze HTTP(S) traffic, enabling a deeper examination of the HbbTV tracking ecosystem. This allows us to perform an in-depth content analysis of the transferred data, e.g., identify fingerprinting scripts, tracking pixels, and cookie usage. Earlier works [9], [10], [65] focused on general smart TV vulnerabilities, whereas we investigate HbbTV-specific risks. Other works focused on attacking smart TVs – but not HbbTV – by different means [1], [51], [89].

Privacy Policies. Privacy policy analyses are often based on sampling from lists that rank domains by popularity [5], [11], [19], [40], [85] or examining specific categories of websites or apps. Examples include the identification of missing legal declarations in child care [36] and baby monitor apps [67], and inconsistencies in IoT app policies [4]. Regarding HbbTV, Tagliaro et al. [71] noted the presence of privacy policies but did not assess their content. We present a content analysis of the policies within the HbbTV ecosystem and identify discrepancies between monitored traffic and disclosed data practices.

Consent Notices. Consent notices initially emerged in the Web context after an update of the European Union's ePrivacy Directive [27] that became effective in 2011. Consent notices, widely implemented after the GDPR, often rely on Consent Management Platforms (CMPs) [38]. However, many employ dark patterns, fail to honor user preferences [58], [82], or have backends that do not always correctly respect people's choices [19], [54]. Research on mobile consent mechanisms shows widespread non-compliance [56], [57]. No prior work covered consent notices in the HbbTV ecosystem. Our study provides initial insights into how HbbTV channels inform viewers about privacy practices through consent mechanisms.

IV. HbbTV Measurement Framework

This section describes our large-scale framework for recording and measuring the behavior of HbbTV channels.

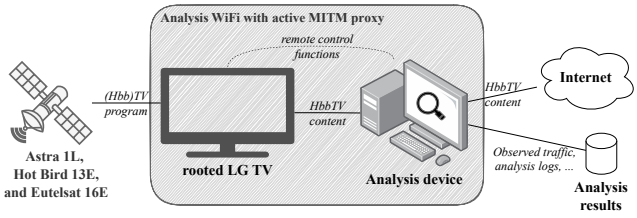


Fig. 3: Overview of our HbbTV experimental setup.

A. Experimental Setup

Figure 3 presents an overview of the experimental setup. Our study used a TV supporting HbbTV (LG 43UK6300 LLB; see Section VIII for limitations of using a single TV), which we rooted using the RootMyTV 2.0 rootkit [64] to install a certificate in the TV's certificate store to be able to intercept and decrypt TLS-protected network traffic. This rootkit no longer works with updated versions of the LG firmware. It is essential to highlight that we have no evidence that any script on the TV tried to identify whether the TV was rooted and that the rootkit did not impact any attributes commonly used for tracking. The TV runs LG's webOS version 05.40.26 and supports HbbTV 2.0. To exclude the traffic generated by the TV itself, we disabled all configurable Internet communication and excluded traffic that belongs to LG (e.g., lge.com).

Since HbbTV content and tracking are initiated by the TV channels (e.g., MTV) rather than by the TV manufacturer or installed apps, and LG-specific traffic was excluded, our results generalize to other TV manufacturers and models. The second component of our framework is a desktop computer used as an analysis device. It is connected to the Internet and acts as a proxy between the TV and the Internet by spawning a Wi-Fi hotspot to which the TV connects. To intercept HTTP(S) traffic, we use the popular HTTP(S) proxy mitmproxy [16] (version 9.0). Since none of the channels validated certificates, we could intercept all HTTP(S) traffic. Using the webOS TV Developer API [50], we also gathered metadata about the watched channel (e.g., name, radio status, encryption) and captured screenshots of the TV content. Furthermore, we collected the TV's cookies and local storage, including their values, to assess if they were set or updated via HTTP(S).

We used a parabolic antenna to receive transmitted TV signals from three satellites: 1) Astra 1L (19.2°E), 2) Hot Bird 13E (13.0°E), and 3) Eutelsat 16E (16.0°E). These were the only satellites from which we could receive a signal from our physical location. Our setup was based in Germany, but the satellites provide TV channels from European countries. The supplementary material of our work provides more details about the physical setup (see Section IX).

B. Channel Selection

Using this setup, we received the signals of 3,575 channels. Closer inspection revealed that this seemingly large set included many channels that did not broadcast any program (e.g., showing a test image only), rendering them unsuitable for analyzing the HbbTV ecosystem.

Thus, we applied a multi-step filtering process to identify the channels of interest and validate whether they support HbbTV. For this, we used different types of information provided either by the TV (e.g., channel metadata) or by our framework (e.g., absence of HTTP(S) traffic). The first three filtering steps relied on the metadata:

1) We used the channel metadata provided by the TV (e.g., "Radio" with the value *true* indicating a radio channel) to test whether it was a regular TV or a radio channel. Of all received channels, 3,150 (88%) were TV channels and 425 (12%) were radio channels.
2) We excluded channels requiring a decryption module, indicated by the message "No CI module" on a black screen (i.e., we could only receive "free-to-air TV" channels). In total, 2,046 (65%) TV channels were not encrypted.
3) We tested whether the `invisible` metadata attribute was set, indicating whether a channel has no signal. We also removed channels with empty names.

We performed an exploratory measurement for the remaining TV channels (1,149 (36.5%)) to assess whether they employed HbbTV technology. We watched each channel for at least 910 s, since previous work found that it could take up to 900 seconds before a channel initiates HTTP(S) connections [34], [71]. We used the observed HTTP(S) traffic for further filtering.

5) We excluded channels with no HTTP(S) traffic (782 (25%)).
6) Finally, we filtered for IPTV channels, as such channels are delivered exclusively over the Internet and are beyond the scope of our study. We removed one IPTV channel, resulting in our final set of 396 TV channels to analyze.

C. Measurement Procedure

The high-level workflow of our data collection process and specific components are described below.

Overview. Figure 4 provides an overview of the overall procedure of our study. First, we performed the measurement runs (i.e., watching TV) in which we collected all data required for subsequent analyses: 1) HbbTV traffic and the trackers used by HbbTV applications (Section V), 2) the consent notices displayed by the channels (Section VI), and 3) the provided privacy policies (Section VII).

Data Collection. To study the privacy impact of HbbTV-enabled TV channels, our measurement procedure includes the collection of HTTP(S) traffic generated by a channel, the content of the TV's local cookie store and local storage, and screenshots of the currently watched TV program. To assess the different HbbTV applications that an HbbTV-enabled TV channel might offer, we created a "remote control script" to interact with the TV to implement different measurement runs, both described in more detail below. As in the exploratory measurement, we watched each channel for at least 910 s. This was to ensure that the HbbTV application (i.e., the HTTP(S) traffic) was initiated if a channel used the technology, and it enabled us to scale our experiment to measure a comprehensive set of channels in a reasonable time frame while still allowing us to gain a good overview of the communication of the HbbTV

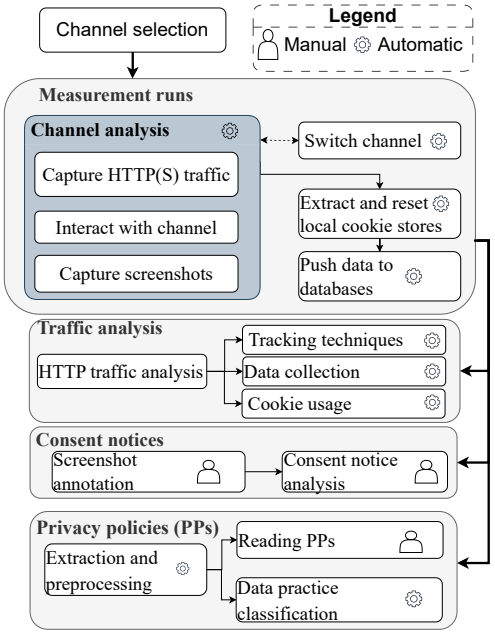


Fig. 4: High-level overview of the study workflow.

applications used by the channel. The measurement time was extended for the measurement runs that involved interaction with the HbbTV applications.

Each measurement run consisted of these steps: 1) initiate mitmproxy (with Wi-Fi), 2) turn on the TV and connect it to Wi-Fi, 3) run the remote control script to watch predefined channels, 4) extract and upload observed data to a database, and 5) wipe the TV's cookies and turn the TV off. Collected data, including HTTP(S) traffic, the TV's cookie jar, local storage, and metadata (e.g., channel list, logs), was pushed to Google BigQuery for analysis. Cookies and local storage were extracted via SSH by accessing the TV's Chromium-based browser. This data was wiped after each run to prevent cross-contamination between measurements. Runs were stateful to track shared resource access (e.g., third-party cookies), but channels were watched in randomized order to minimize order effects. The TV was always powered off after each run.

Remote Control Script. We used the webOS TV Developer API [50] to develop a "remote control script" that interacts with the TV. We implemented this Python script using the PyWebOSTV library [42] (version 0.8.3), which allowed us to simulate different user behaviors. When switching to a channel, the script waits for 10 s, takes a screenshot of the current program, queries all available metadata using the API, and continues taking screenshots every 60 s until the end of the measurement interval. This resulted in at least 16 screenshots of each watched TV channel.

We modified mitmproxy to map HTTP(S) traffic to TV channels. The remote control script sent the channel's name and ID to the proxy upon switching. To prevent false assignments, we checked if the HTTP referrer header updated the channel

for a request, accounting for delays during switching. Only requests from the last 15 minutes of channel watch time were considered to minimize false positives.

Description of Measurement Runs. As already mentioned in Section II, the HbbTV standard proposes the use of four colored buttons to interact with HbbTV applications. The red button is supposed to start the channel's "autostart application", while the blue, yellow, and green buttons can be used to interact with HbbTV content at the channel's discretion [29]. To study how HbbTV channels use these buttons and the associated impact on tracking and transparency, we designed our measurements around (non-)interaction with these colored buttons. Thus, in the remote control script, we implemented a total of five different *measurement runs* executed individually for all channels. Before each measurement run, we randomized the channel order.

In the *General* measurement run, we did not interact with the TV at all (except for switching channels), following the measurement protocol described under "Remote Control Script" and watching each channel for 900 s. For the other four *color button measurement runs*, we first generated a fixed interaction sequence of 10 random button presses from the set of the four cursor buttons (`UP`, `DOWN`, `LEFT`, `RIGHT`) and `ENTER`, making sure that `ENTER` was pressed at least once to potentially trigger the loading of new HbbTV content. Each of the four color button measurement runs (`Red`, `Blue`, `Yellow`, and `Green`) would first switch to a new channel, wait for 10 s, press the respective colored button, wait for another 10 s, and then run the fixed interaction sequence, mimicking user interaction with the potentially loaded new content and triggering additional content loading (e.g., in a media library). To better assess the impact of the interaction, we extended the watch time for each channel in these four measurement runs by 100 s to a total time of 1000 s. This resulted in 27 screenshots for each channel in these four measurement runs. The concrete interaction sequences used in the measurement runs can be found in the supplementary material (see Section IX). In theory, the colored buttons might be assigned different functions *after* another colored button has been pressed. In a pre-study, we never encountered such a use of the buttons. Thus, we decided only to use a single colored button in each measurement run.

Finally, the routine switched to the next channel, automatically exiting any started HbbTV application. This approach allowed us to capture traffic generated by content that is only delivered upon user interaction with an HbbTV application.

D. Measurement Report and Data Set Overview

Table I provides an overview of the five measurement runs conducted between August and December 2023. This extended measurement period was necessary to 1) measure certain channels with specific airing times multiple times and 2) physically restart the TV due to issues with the TV's API, both limiting the scalability of our approach.

In our five measurement runs, we successfully measured an average of 331 (min: 215, max: 381, SD: 63) different HbbTV channels. The deviation in the number of analyzed channels per measurement run can be attributed to the fact that some channels were not always available (e.g., some channels only broadcast during daytime). The analyzed channels are almost equally distributed among the three satellites: 31.5% Astra 1L, 35% Hot Bird 13E, and 33.5% Eutelsat. Each satellite provider (e.g., Astra [68]) has an online guide with additional information about the available channels, such as the channel's language or category. Based on the provided information, most of the analyzed channels (369 (92.7%)) are broadcast in German, 12 in English, six in multiple languages (e.g., German and French), and the rest in French (3) or Italian (1). This distribution can be attributed to the measurement location.

We captured 457,492 HTTP(S) requests and responses. We found statistical significance (p-value < 0.0001) of a clicked button (measurement run) and the HTTP(S) traffic generated by a channel. The column "HTTPS Share" in Table I indicates how much HTTP traffic was encrypted during measurement. Across all measurement runs, the channels stored over 3,158 cookies in the local cookie jar and 731 objects in the TV's local storage. Again, the measurement run had a statistically significant impact on the placement of cookies in both storage spaces (p-value < 0.0001 for both), meaning that depending on the pressed button, different numbers of cookies were set.

In addition to the HTTP(S) traffic, we also captured metadata available using the webOS TV Developer API [50]. This included program guides (e.g., the current show), channel metadata (e.g., name or signal frequency), and log data from our interaction with the TV (e.g., channel switch, button clicks). During the measurements, we performed over 75k interactions with the watched TV channels and stored 77k log entries of interactions and metadata. We took 41,617 screenshots, collected 17.1 GB of raw data, and watched 418 hours of television.

Statistical Analysis. We used the Kruskal-Wallis test [47] to assess differences in the central tendency of a continuous variable across groups (e.g., measurement runs), with a 95% confidence interval ($\alpha = 0.05$) and η^2 to measure effect size. Effect sizes were classified as *small* ($\eta^2 \leq 0.06$), *moderate* ($0.06 < \eta^2 < 0.14$), and *large* ($\eta^2 \geq 0.14$) [15].

V. The HbbTV User Tracking Ecosystem

This section provides an in-depth analysis of tracking techniques and the HbbTV tracking ecosystem.

A. Identification of First and Third Parties

In the traditional Web, the term *first party* denotes the website the user visits, while the term *third party* refers to any other domain from which the website loads additional resources. We cannot directly adopt this classification for the HbbTV ecosystem because users do not visit a website; instead, the HTTP communication endpoints are included in the received TV signal. This means that when a user watches a TV channel, the HbbTV signal includes a URL that is loaded by the TV (e.g., the red boxes in Figure 1a).

To determine the first party of a TV channel, we identify the first HTTP request (based on the recorded timestamp) that loads additional content (e.g., HTML code) displayed by the

Meas. Run	Date	Channels	HTTP Req.	HTTPS Req.	HTTPS Share	$\sum$ Cookies	1P Cookies	3P Cookies	Local Stor.
General	2023-08-21	374	95,133	583	0.61%	272	192	130	157
Red	2023-09-14	375	151,975	8,456	5.56%	911	665	557	160
Green	2023-09-22	215	32,138	2,401	7.47%	685	516	345	157
Blue	2023-09-27	309	43,556	1,264	2.90%	380	291	268	159
Yellow	2023-10-12	381	134,690	3,078	2.29%	554	392	297	161

TABLE I: Overview of the data collected for each measurement run, including analyzed channels, share of HTTP(S) traffic, and number of observed cookies (cookie jar and local storage). The counts of third and first-party cookies do not add up, as some cookies are first-party cookies on one channel but third-party cookies on another channel (see Section V-C).

TV and define the eTLD+1 of this request to be the first party. We choose not to use the first observed request, as manual analysis indicates that some channels encode connections to third-party services (e.g., google-analytics.com) directly into the HbbTV signal. To address this challenge, we used URLs flagged by EasyList [74] (see Section V-D) to prevent known trackers from being defined as a first party. If a URL is on the list, we treat the request as a third-party request and check the subsequent request according to the previously named definition to determine the first party. To increase the accuracy of this approach, we manually analyzed all observed eTLD+1s to check if they are used for user tracking or other popular third-party services on the Web. This process identifies only one domain falsely classified as a first party.

B. Information Collected by HbbTV Channels

We analyze how and to what an extent HbbTV channels collect personal data (e.g., watched channel or device information). By examining HTTP GET and POST requests, we distinguished two types of collected data:

Technical Data. This type includes technical information about the TV. We searched for: (1) manufacturer (`LGE`), (2) model (`43UK6300LLB`), (3) operating system (`WEBOS4.0 05.40.26, W4_LM18A`), (4) language (`German`), (5) local time, and (6) IP/MAC address. We found that 112 (29%) of the analyzed channels collect data about the user device. The data was sent to nine third parties.

Behavioral Data. This type includes data on the aired program (e.g., show genre) and viewer behavior, which is used to profile interests for targeted ads. Similar to the technical data, we performed a keyword search for, e.g., TV show genres [14] or the name of the watched TV show. We found 94 channels sending the genre of the currently watched TV show to third parties. Overall, 23,671 requests contained personal data like the currently watched show.

While the significance of the leaked information seems to be low, we found circumstantial evidence for the leakage of behavioral data. For example, we found information on specific brands (e.g., L'Oréal) that were not related to the currently aired show. Combined with further information (e.g., IP addresses, fingerprints), a user-specific profile may be created.

C. Cookie Usage

On the Web, cookies are often used to track users [79]. As HbbTV uses web technologies, such as HTTP(S) and HTML5, we expect this ecosystem to employ web tracking techniques.

Meas. Run	# 3Ps	# 3P Cookies	Mean	Min	Max	SD
General	36	167	2.31	1	8	1.74
Red	107	560	3.59	1	22	5.82
Green	77	287	3.69	1	21	4.27
Blue	47	189	2.04	1	16	2.34
Yellow	88	300	3.2	1	24	4.16

TABLE II: Use of cookie-setting third parties by measurement.

1) General Cookie Usage: In our measurements, we captured 1,705 distinct cookies (cookie jar and local storage), 92% of which were set by a request that was labeled as a tracking request. We did not perform a specific identification of potential tracking IDs stored in cookies, as we cannot reliably differentiate between IDs computed based on the used TV (i.e., fingerprints) or other values (e.g., session IDs) since we only used one hardware device.

Across all measurement runs, each channel set 4.1 cookies on average (min: 1, max: 25, SD: 4.8). We identified 166 distinct (first and third) parties who set cookies on 391 HbbTV channels. It is interesting to assess the purposes for which the channels use cookies to understand their use cases. To classify the purpose of cookies, we used Cookiepedia [59]. Only 20.5% of the observed cookies could be classified by this service, which is considerably less than the number of cookies that Cookiepedia can classify for cookies collected on the Web (57% [20]). This suggests that the HbbTV (tracking) ecosystem is different from the Web ecosystem. In the supplementary material (see Section IX), we provide an overview of the categories of classified cookies. In measurement runs where we "press" a colored button, we see higher shares of cookies that could be classified. Thus, the loaded HbbTV apps rely on more commonly used web services, and these measurements show more usage of "Targeting" cookies.

2) Third-party Cookie Usage: As third-party cookies are a common means to track users across the Web, we assessed whether HbbTV applications have adopted this technique. For this, we analyzed whether a channel loaded a third-party application that used cookies and, if it did, whether the cookie was used for tracking. On average, each channel set 3.1 (min: 1, max: 24, SD: 3.4) third-party cookies. Table II provides an overview of the number of cookies and parties setting these cookies across the measurement runs. We found a statistically significant effect (p-value < 0.0001) in the volume of third-party cookie usage across the measurement runs. The results

Fig. 5: Long tail distribution (positive skew) of cookie-using third parties.

Meas. Run	On Pi-hole	On EasyList	On EasyPrivacy	Track. Pxl	Fingerp.
General	203	6	24	80,960	51
Red	2,120	1,375	388	90,199	536
Green	1,051	463	134	14,593	161
Blue	313	8	42	5,925	179
Yellow	1,668	660	105	85,897	151

TABLE III: Total number of tracking requests in our data set and effectiveness of popular filter lists in identifying them.

indicate that HbbTV channels and HbbTV-enabled applications use third-party services.

To explore whether third-party cookies might be used to track users across different TV channels ("cross-channel tracking"), we looked for a third party included on multiple channels and accessed the same cookie(s) on these channels. We observed the most frequent third party (xiti.com) on 119 channels, while 38 third parties were used by only one channel. Third parties that set more than one cookie on a channel used 13.9 (min: 2, max: 119, SD: 19.8) cookies, meaning that they used them to provide cross-channel services such as user tracking. Of those cookies, 11% were classified as "Targeting/Advertising" cookies. Figure 5 shows the distribution of third parties setting cookies across channels. The supplementary material lists the most frequently observed third parties (see Section IX). The results show a long tail distribution of cookies used by third parties, with only 25 third parties used by more than ten channels. In contrast to the Web, which is dominated by a few large players [79], the results indicate that the HbbTV third-party cookie ecosystem is scattered. We provide the top 10 third parties in the supplementary material.

3) Cookie Syncing: While many third parties are directly embedded on multiple channels, it is unclear whether they exchange data with each other. Cookie syncing is a two-step process used to exchange data between third parties: A website loads a third party, and the request responsible for loading the script is redirected to the syncing partner. This redirected request contains, e.g., a user ID [79]. To identify cookie syncing in our data, we looked for a third party sending an HTTP request containing an identifier to another party.

To identify cookies potentially holding identifiers, we adapted the method by Acar et al. [2], [79]. A cookie value was considered an identifier if: 1) it was 10–25 characters long (i.e., sufficient entropy for an ID) and 2) it was not a valid Unix timestamp within our measurement period. The timestamp condition was added, as manual analysis revealed that many cookies contained such timestamps (e.g., consent collection or channel switching). Using this method, we identified 14,236 cookie values as potential IDs. To identify cookie syncing, we analyzed if any identified potential ID was sent by one party to another. In total, 25 of the identified cookie values were transferred to another party in an HTTP request. We detected syncing activities of only two domains (by eTLD+1) in the `Red`, `Green`, and `Blue` measurement runs. We observed cookie syncing activity on 20 channels. In contrast to cookie syncing on the Web (with around 500 syncing connections in [78]), the technology is less popular in HbbTV applications.

D. User Tracking

User tracking is part of many digital services, raising the question of the extent to which HbbTV applications use such techniques. Widely used metrics to identify trackers are filter lists, such as EasyList [74], EasyPrivacy [75], and the standard block list used by Pi-hole [12], [63]. We compared all URLs (340,643) observed in our HTTP traffic with the filter lists. We found that only 2,512 (0.5%) of them were flagged by EasyList, 693 (0.15%) by EasyPrivacy, and 5,355 (1.17%) by Pi-hole (see Table III). This could mean that tracking is not widespread in HbbTV applications or that the tracking parties in the HbbTV ecosystem differ from web trackers.

Besides the regular Pi-hole filter list, we analyzed block lists designed for smart TVs [44], [62]. Compared to the Pi-hole filter list that blocked 5,355 tracking requests, Perflyst's PiHoleBlocklist [62] blocked 3,910 tracking requests (27% less). Kamran's SmartTV block list [44] only blocked 1,948 requests (64% less). While these lists were designed for smart TVs and include some HbbTV-specific trackers, they still do not protect users as expected. This is partly due to the fact that commonly used trackers in HbbTV applications (e.g., tvping.com) are not included in these lists, indicating that the main focus of these lists lies on smart TVs and applications running on them (e.g., Netflix). To assess whether user tracking exists in the HbbTV ecosystem, we developed heuristics to identify web tracking techniques (i.e., tracking pixels and fingerprinting):

1) Tracking Pixels: Tracking pixels are nearly invisible images with a size of 1x1 pixel used to track users [22], [32], [80]. We assumed that a response contained a tracking pixel if the following conditions applied [80], [81]: 1) The HTTP content type of the response indicates that the response is an image, 2) the total size of the response is smaller than 45 bytes (approx. the size of an empty image), and 3) the HTTP response code is 200 (OK). Using this method, we identified a total of 277,574 tracking requests, of which only 649 (0.2%) were flagged by EasyList and 659 (0.2%) were flagged by EasyPrivacy. These requests were issued by 47 distinct eTLD+1 (i.e., trackers), of which 8 (17%) were present in EasyList. Table III shows the observed tracking pixels. In our measurements, 350 (89.5%) TV channels utilized a tracking pixel at least once. Notably, 60.7% of the entire HTTP(S) traffic is used for user tracking via tracking pixels. This tracking can be attributed to a single third party (tvping.com) that is used by 141 (36.2%) channels. This traffic volume is due to tvping.com sending a request including the channel, session, and user ID almost every second [71]. The results suggest that there is an HbbTV tracking ecosystem, that filter lists miss trackers, and

that several TV channels participate in the ecosystem. Using colored buttons affects the presence of tracking pixels, but most channels only load a few extra trackers.

2) TV Fingerprinting: To identify fingerprinting, we looked for all responses that 1) contained JavaScript code (based on the HTTP content type) and 2) scripts that included APIs commonly used for fingerprinting (e.g., Canvas or WebGL) or libraries commonly used by fingerprinting scripts (e.g., Fingerprint2 [30]). Our framework does not allow for checking the execution of scripts, meaning that we could not apply more sophisticated detection methods. Still, our approach allowed for an assessment as to whether TV fingerprinting is used and will provide a lower bound. In our study, 60 (15%) channels might use fingerprinting techniques. We only found 21 eTLD+1 that provide these scripts. Seven of these trackers were hosted by first parties, who turned out to be more active in terms of observed tracking requests: 88% of all identified fingerprinting requests were issued by a first party. Of all identified fingerprinting requests, only one eTLD+1 was flagged by EasyList (10 requests, 0.9%). On EasyPrivacy, we could identify 34 (3%) requests from only one eTLD+1. In contrast to tracking pixels, the analyzed channels included various fingerprinting scripts depending on the measurement runs.

3) Channel-level Analysis: The previous analysis focused on the different measurement runs mimicking user interaction with HbbTV applications. In this section, we assess the privacy practices of each channel individually to understand whether specific channels affect viewer privacy more than others (e.g., by collecting more user-related data). For this analysis, we only considered channels with at least one tracking request observed. On average, a channel issued 1,132 (min: 1, max: 59,499, SD: 3,206) tracking requests across all measurement runs. One channel issued 59,499 tracking requests, of which 59,339 (99.7%) went to tvping.com; these requests only occurred in the `Red` measurement run. The channels contacted an average of 7.25 (min: 1, max: 33, SD: 4.99) trackers. The top 10 channels with most trackers represent 6.34% of the total tracking requests. Figure 6 shows the distribution of trackers across all channels. Apart from a single outlier, there appears to be a consistent distribution of channels with comparable numbers of trackers, exhibiting gradual declines. This shows that tracking is distributed among most channels.

We found a statistically significant effect of the channel on the number of trackers ($p < 0.0001$) with a large effect size, indicating that some channels track more intensively. However, user interaction (i.e., pressed buttons) had a greater impact on tracking behavior than the watched channel, suggesting that tracking depends more on user actions than on channel choice.

4) Channel Category Analysis: In this section, we analyze the impact of a channel's category (e.g., Children or News) on the number of trackers. We extracted the channel categories from the official documentation provided by a satellite operator [66]. There are 10 channel categories in our data set, and some channels were grouped into multiple categories. For our analysis, we only used the first assigned channel category. On average, a category contains 25 (min: 1, max: 217, SD: 52)

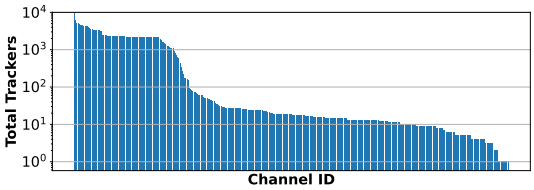


Fig. 6: Distribution of the observed trackers per channel.

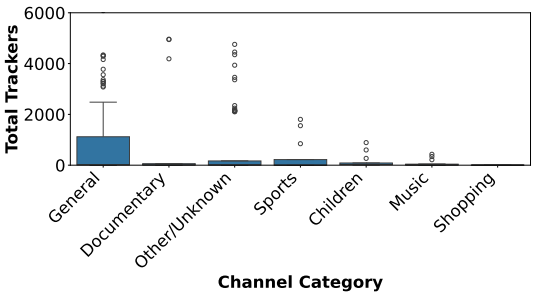


Fig. 7: Number of trackers by channel category (ordered by channels with most trackers). To increase readability, we combined categories with fewer than 10 channels into the category "Other/Unknown." The y-axis is cut at 6k, excluding one data point in the "General" category (∼60k).

channels. Figure 7 provides an overview of the number of trackers used by the channels in the different categories. On average, channels in the category "General" have the most trackers (664), while channels in the category "Children" rank sixth (201). The top five categories represent 98.5% of all tracking requests. The top five categories contain 326 (82%) channels. We found a statistically significant effect of the channel category on the number of trackers ($p < 0.0001$) with a medium effect size.

5) Case Study – Channels Targeting Children: Some analyzed channels are dedicated to a specific topic (e.g., sports or news) or audience (e.g., children). Art. 8 GDPR and Recital 38 GDPR [28] regulate the processing of children's data. Thus, HbbTV channels providing shows and programs for children need to adjust their data practices to comply with GDPR requirements. We use the satellite providers' channel metadata (see Section IV-D) to identify channels targeting children. According to this data, 12 of the channels in our set exclusively target children. We recorded 1,946 tracking requests (pixels, fingerprinters, or known trackers) and 97 third-party "Targeting/Advertising" cookies across all measurement runs on these channels. Thus, these channels track their audience (i.e., children) and process their data, which could violate the GDPR. We used the Wilcoxon-Mann-Whitney test [53] to find out whether children's TV channels exhibited a different behavior in embedding trackers than the other channels. We found no statistically significant difference ($p > 0.3$) between channel categories (e.g., children, shopping, music, or documentary). This indicates that children's TV tracks viewers to a similar extent as other channels.

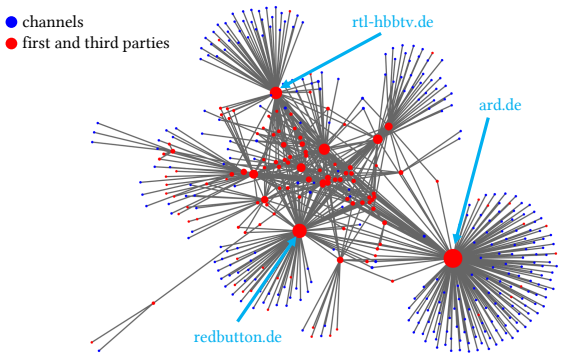


Fig. 8: The HbbTV tracking ecosystem. The size of a node indicates its number of edges.

E. The HbbTV Ecosystem

Previously, we showed that the HbbTV (tracking) ecosystem is independent of that on the Web. We now assess how the HbbTV ecosystem is structured, its most prominent players, and how densely connected the system is. We created a network graph based on the observed HbbTV traffic, using the Python library NetworkX [37]. Nodes denote either TV channels or domains (eTLD+1) and edges represent that we observed HTTP(S) traffic between them. For each channel, we connect the channel node with the identified first party (eTLD+1) and all third parties (eTLD+1) observed on that channel with the respective first-party node. The resulting graph consists of channel nodes connected to their first-party node and all third parties embedded by a channel.

The final graph, shown in Figure 8, has one component with 429 nodes and 675 edges, indicating that all entities in the HbbTV tracking ecosystem are connected. The average connectivity of a node is 33.4, with an average path length between two nodes of 2.91, which denotes the graph's average distance between node pairs. This indicates a well-connected ecosystem. The three most connected nodes in terms of degrees of a node are ard.de (188 edges), redbutton.de (103 edges), and rtl-hbbtv.de (75 edges). All three are connected to different channels of German TV networks. Overall, we found 18 nodes with at least 10 edges. On average, each node has 3 (SD: 11) edges. Excluding the TV channel nodes, which by definition only have one edge to the included first party, the graph has 39 nodes with only a single edge (i.e., they are only used by one eTLD+1). The observation reveals that some services predominate the HbbTV ecosystem, and these nodes are responsible for the high connectivity of the graph. We find that the most frequently used third party (xiti.com) has only six edges. This means that this third party is included by other third parties rather than directly by a channel. Similarly, the most prominent tracking pixel domain (tvping.com) has only 14 edges. Our findings indicate that HbbTV channels frequently track users. These results raise the question of whether these channels transparently disclose and offer choices regarding this data collection and tracking.

Meas. Run	No Sign.	CTM	TV Only	Media Lib.	Privacy	Other	Total
General	96	0	5,669	1	115	86	5,967
Red	162	285	3,520	4,532	269	1,660	10,428
Green	81	75	4,402	730	29	149	5,466
Blue	84	31	6,947	646	525	333	8,566
Yellow	247	79	7,012	3,376	79	397	11,190

TABLE IV: Distribution of HbbTV overlay types (see the codebook in supp. material, Section IX) on the screenshots taken in each measurement run. CTM = "channel tech message".

	Screenshots			Channels		
Meas. Run	# Total	# Priv.	%	# Total	# Priv.	%
General	5,967	115	1.93	374	70	18.72
Red	10,428	269	2.58	374	70	18.72
Green	5,466	29	0.53	214	26	12.15
Blue	8,566	525	6.13	309	38	12.30
Yellow	11,190	79	0.71	380	54	14.21

TABLE V: Prevalence of privacy-related information.

VI. Consent Notices in HbbTV

We analyzed 41,617 screenshots to gain insights into how HbbTV channels prompt viewers to consent to data processing.

A. Annotation of Screenshots

For the analysis of consent notices, we resorted to manual inspection and annotation of screenshots, for two reasons: 1) Contrary to the Web, where many consent notices have evolved to implement the IAB Europe TCF (see Section III), consent notices have hardly been studied for HbbTV. To the best of our knowledge, there is no comparable list of known signatures of CMPs for a code-based approach. 2) OCR-based detection would face the challenge of TV images often showing other types of text boxes, including program announcements, news tickers, and advertising, not to mention other HbbTV text overlays. Two authors inspected a subset of the data, devised a coding scheme, and discussed edge cases until a final codebook emerged (see Section IX). Using Label Studio [41], in the first round of coarse-grained annotations, we checked whether and what type of HbbTV overlay was present on each screenshot. Table IV shows the distribution of HbbTV overlays by measurement run, confirming that channels tend to show media libraries on the red button and privacy information on the blue button. In the second round, we inspected the images containing privacy-related information for the type of information (privacy policies, consent notices, or something else) and the other types of HbbTV overlays for whether they displayed a button or text pointing to privacy-related information. Due to technical limitations, we did not have a complete set of 16 / 27 screenshots for a few channels.

B. Consent Notices on HbbTV

Prevalence. Across all measurement runs, 121 channels (31.03%) displayed a cookie notice or privacy policy on at least one screenshot. Table V shows that at most 18.72% of channels did so, most commonly in the `General` and `Red` measurement runs. These values are much lower than

those reported for the Web [19]. The highest percentage of individual screenshots with such information (6.13%) was reached in the `Blue` measurement run. In the `General`, `Green`, and `Yellow` measurement runs, all screenshots with privacy-related information displayed consent notices. In the `Red` measurement run, 198 screenshots had consent notices, 19 privacy policies for two channels, Astro TV and Kabel Eins Doku, and two channels, RBB and MDR, had a split screen displaying a privacy policy and current cookie settings. The `Blue` measurement run contained the most diverse privacy information – hundreds of privacy policy screenshots, different layers of consent notices, and hybrid overlays with privacy policies and cookie controls. This was the only measurement run showing the second layer of consent notices that allowed for (de)selection of cookies by category or specific services. Some screenshots even showed a third layer that asked users to confirm the deselection of a cookie.

Interfaces and Branding. All consent notices visible on the screenshots were instances of only twelve recurring styles and brandings: 1) RTL Germany group, 2) ProSiebenSat.1 group (non-modal), 3) ProSiebenSat.1 group (full screen, modal), 4) QVC, 5) DMAX Austria / TLC / Comedy Central, 6) HSE, 7) Bibel TV, 8) RTL Zwei, 9) TLC, 10) ZDF (full screen, modal), 11) COUCHPLAY (on Kabel Eins Doku). A twelfth group of channels (MTV, WELT, Comedy Central, MediaShop, and N24 Doku) showed an identical banner without concrete channel names or logos. Notice types 9) and 10) only showed up in the `Blue` measurement run. All of them were in German, as they originated from either German or Austrian TV channels. All first-layer notices except for types 3 and 10 were non-modal and covered less than half of the screen, allowing for continued (albeit partly obstructed) TV watching.

Interaction Options. On the first layer, all notice types had a button to accept all cookies and data processing. The others varied and featured German variants of "Settings" (types 1, 12), a single button "Settings or Decline" (types 2, 3, 11), "(Privacy) Settings" and an explicit "Decline" button (types 4, 10), "Privacy" and "Settings" (types 7, 9), or "Privacy" only (type 5). RTL Zwei (type 8) notices were unique in allowing for a category-based selection on the first layer, between "Necessary" (immutable), "Functional," and "Marketing," in addition to "Accept all" and "Only necessary" buttons. Bibel TV (type 7) lets users deselect Google Analytics on the second layer. In both cases, the checkboxes are pre-ticked, which the European Court of Justice has ruled not to be GDPR-compliant [25]. Checkboxes marked with a '?' are encountered on the second layer of type 12 notices. The type 11 notice linked to a "list of partners," but it did not show up in screenshots.

Nudging and Dark Patterns. Hiding options to decline on the second layer, as described above, nudges users towards accepting data collection. Contrary to the Web, where users can freely navigate UI elements, the "cursor" on HbbTV must be placed on a button by default, which offers a new dimension to steer the user's selection towards a desired outcome. For all 12 notice types, this was the "Accept" button on layer 1, as indicated by highlighting in a different color and often a shadow or colored border around the button.

Persistence. Consent notices often did not occur on all screenshots for a given channel and measurement run, while privacy policies tended to be shown continuously, safe for scrolling due to the automated input.

Pointers to Privacy Information. A total of 290 channels (74.36%) displayed a button or text pointing to "Privacy" or "Cookie Settings" at least on one screenshot across all measurement runs, typically in media libraries or dashboards showing available HbbTV content. As with consent notices, such pointers to privacy information were more often visible on screenshots of TV channels owned by private companies, as opposed to public broadcasters. Still, pointers were often hidden at the footer of long media library pages or smaller than surrounding interaction elements, sometimes barely visible in terms of font size or color.

Other Observations. Apart from privacy aspects, manual inspection of HbbTV overlays revealed a case of location-targeted advertising. A sleeping aid ad was overlaid with text stating that it was available in pharmacies in the city where our measurement setup was located, stating the city's name.

VII. Privacy Policies for HbbTV Channels

Privacy policies disclose data practices concerning collected, used, or shared personal data and inform affected users about their rights, though discrepancies with the actual behavior may exist. This section describes our procedures for extracting privacy policies from the network traffic and our methods to assess their content. Figure 9 provides an example of a privacy policy displayed on an HbbTV channel.

A. Collection of Privacy Policies

To identify the privacy policies displayed to users in the recorded HTTP traffic (see Section IV), we used an established toolchain implementing best practices to collect, identify, and preprocess privacy policies [39]. The preprocessing components of this toolchain were plain text extraction using the Boilerpipe library [46], language detection via majority voting, and machine learning-based differentiation between privacy policies and miscellaneous texts using trained classifiers [39] that, according to the authors, achieved 99.1% and 99.8% F1-scores for English and German, respectively.

We manually evaluated the classifiers' output to ensure their correctness, resulting in the correction of 18 false negatives. These misclassifications were likely due to these texts combining data-practice disclosures with unrelated content, such as discount offers and HbbTV usage instructions.

We identified 2,652 German, three English, and one bilingual (German/English) privacy policies. 1193 of the German privacy policies were found in the traffic captured in the `Yellow` measurement run (see Section IV-C), 484 in the `Red`, and 479 in the `Green` measurement run. 259 and 237 privacy policies were obtained for the `General` and `Blue` measurement runs, respectively. The English and bilingual privacy policies occurred in the `Red` measurement run. 25 privacy policies

Fig. 9: A hybrid consent notice and privacy policy.

occurred in a single measurement run exclusively. The hosting domains of the policies either belonged to the main website of the respective TV channel, a domain or subdomain dedicated to HbbTV, or service providers like smartclip.

We removed duplicate policies based on their SHA-1 hash and channel names, resulting in a final data set of 55 German, one English, and one bilingual privacy policies. We did not consider privacy policies on the TV channels' websites, as we focus on privacy practices available to viewers within the HbbTV ecosystem. Using SimHash [52], we identified 11 German privacy policy groups with nearly identical content aside from minor differences, such as channel name. We kept these policies in the data set, as they belong to different TV channels, which might differ in data practices.

B. Data Practices in Privacy Policies

As the traffic was recorded in an EU country (Germany) and therefore mimics the perspective of an EU resident, the regulatory regime governing the privacy policies was the GDPR. Thus, we analyzed the privacy policies based on the MAPP bilingual English-German data practices taxonomy [7], which extends the OPP-115 taxonomy [86] to include the GDPR. This taxonomy comprises the categories of first-party and third-party data collection/sharing, each consisting of attributes containing fine-grained values. The categories, attributes, and values are listed in the supplementary material (see Section IX).

We used Arora's fine-tuned BERT models (English: bert-base-uncased, min F1: 0.45, max F1: 0.85, German: bert-base-multilingual-uncased, min F1: 0.37, max F1: 0.78) to identify the existence or absence of each category, attribute, or value in the privacy policies. These F1 scores are similar to other BERT-based models trained to identify data practices in privacy policies [3]. Note that Arora's models identify GDPR-specific data practices for English and German privacy policies, as opposed to PrivacyLint [6] and PrivacyCheck [88], which were trained using the English OPP-115 corpus from 2016.

We supplemented our deep learning-based analysis with a dictionary-based approach to obtain an overview of the usage of GDPR-specific terminology and the issuers' GDPR awareness [7]. For this purpose, we used a multilingual dictionary [19], which contains GDPR-specific phrases collected from Articles 6 and 13 of the GDPR in 24 languages, including English and German (see Section IX). In addition, one author qualitatively read through the privacy policies to identify data practices specific to the HbbTV ecosystem.

C. Content of Privacy Policies

To identify HbbTV-specific data practices and investigate whether the privacy policies were tailored to the HbbTV ecosystem, we searched their texts for the term "HbbTV." 40 (72%) of the German privacy policies and the English and bilingual policies included this term. Analyzing the terms surrounding "HbbTV" yields that the RTL channel dedicated an HbbTV-specific email address to complaints or inquiries. Moreover, 8 policies indicated the possibility of accessing privacy settings using the blue button on the TV remote control. Furthermore, we noticed a trend in using cookies for coverage analysis to measure HbbTV viewers' behavior.

All privacy policies acknowledge first-party collection/use of personal data. 29 (52%) of the German policies declared the collection/use of personal data by third parties. Most privacy policies state that IP addresses of devices are collected, while their declared anonymization/pseudonymization practices differ. While some data controllers anonymize IP addresses completely, others cut, e.g., the last three digits.

Not all German privacy policies contained declarations on GDPR data subject rights: Art. 15 (*34 / 61%*), Art. 16 (*38 / 69%*), Art. 17 (*33 / 60%*), Art. 18 (*33 / 60%*), Art. 20 (*9 / 16%*), Art. 21 (*9 / 16%*), Art. 77 (*36 / 65%*). We identified 10 (18%) German policies in which data controllers declare to collect, process, or store personal data, partly for an indefinite time, based on their "legitimate interests", which, as a gray area in the processing of personal data, has been subject of research [43], [48]. Another example was the policy of Krone.tv, according to which the program is adapted to the individual viewer behavior. The policy of the Sachsen Eins TV channel contained vague statements [49] about possible processing of personal data based on vital interests and legal obligations.

As discussed in Section V-D5, some channels targeting children track their audience. The privacy policy of Super RTL pointed out that ad personalization and profiling were limited to a specific time: from 5 pm to 6 am (i.e., during the evening and night). Three TV channels from the same media group share this policy. This statement is interesting, as Super RTL mainly addresses children and broadcasts children's TV shows during this time. However, we identified 21 known tracking requests on two of the three TV channels (Super RTL and Super RTL Austria) outside of the period indicated. Among other things, the requests contained user IDs and the TV show watched. We identified one tracker from EasyList and three tracking eTLD+1 (smartclip.net) from the Pi-hole filter list. Moreover, we identified seven additional tracking eTLDs (e.g., tvping.com) using our method from Section V. Thus, the channels may collect children's data outside of the declared period, which contradicts their privacy policy.

The RTL channel's privacy policy refers to the German Telecommunications Digital Services Data Protection Act (formerly abbreviated as TTDSG in German, now TDDDG [13] and enforced in December 2021). Implementing the EU's

ePrivacy Directive, Section 25 of the TDDDG permits storing or accessing information, including cookies, on an end user's device only if the user has granted consent unless storing or accessing the information is technically indispensable to provide a service explicitly requested by the user. No other TV channel's privacy policy included statements concerning cookies and the TDDDG at the same time, despite the stated data controller being located in Germany and cookies being widely used. The German HGTV channel's policy contained opt-out statements regarding data processing, interest-based advertisement, and coverage measurement. This contradicts GDPR's requirement of a legal basis for data processing, as targeted advertising requires opt-in consent [8], [61], [77].

VIII. Limitations & Research Ethics

As a main limitation, the study setup covers unencrypted TV channels ("free-to-air") only. These channels include public and commercial channels. Our analysis captures a broad range of channel business models (e.g., public funding, advertising, or teleshopping). Decrypting channels would have involved administrative and legal hurdles (e.g., country-specific tax ID number or address). Nevertheless, our data set comprises channels operated by both private companies and public broadcasters, still providing a comprehensive insight into the HbbTV ecosystem. The setup was physically based in Germany and received signals from three satellite networks: Astra 1L, Hot Bird 13E, and Eutelsat. Receiving additional signals, such as Thor 0.8°West for Scandinavia or Hispasat 30.0°West for Spain and Portugal, was physically impossible from this position and would require moving the equipment to different locations. The three satellite networks used in our study are the most frequently used in Europe, and the number of analyzed channels is more than ten times higher than the 36 investigated by the most closely related prior work [71]. We analyzed the broadcast signal from a single LG TV. In HbbTV, the TV channels initiate the tracking, rather than the TV manufacturer or installed applications. Our analysis excluded traffic sent to an eTLD+1 belonging to LG (e.g., lge.com). Hence, in our setup, the TV (brand) used does not impact the results.

Ethics. We collected and analyzed technical and image data from public TV broadcasts. These screenshots, often containing images of individuals, were withheld from release to comply with privacy and intellectual property laws. The examples in this paper and the repository exclude identifiable images of people, which is no threat to the validity of the results. Our study did not involve human subjects and was exempted from review by the IRBs in all participating institutions. Screenshots were taken every 60 s to facilitate manual analysis. We reported our findings of potential problematic data collection to the channels that responded by initiating investigations.

Future Work. Our work highlights the need for new tracking protection mechanisms and shows that mechanisms like filter lists are tailored to the particular scenario of users browsing the Web and cannot be applied to the HbbTV ecosystems without adjustment. Future research could extend existing Web-based filter lists by (automatically) deriving additional filter rules from observed traffic that block trackers for HbbTV and smart home settings. Beyond HbbTV, future work could focus on other emerging ecosystems to determine whether past assumptions about how user tracking works still hold or new actors, mechanisms, and threats are at play that create a need for novel defense mechanisms and regulatory action.

IX. Conclusion & Recommendations

Conclusion. User privacy remains a serious concern, as digital devices become increasingly interconnected, from smart TVs and streaming devices to voice assistants and IoT appliances. The HbbTV ecosystem illustrates how data collection can thrive when technological innovation outpaces privacy safeguards. Although HbbTV partially employs web technologies like HTML5, web-based privacy defenses are considerably less efficient in protecting users' privacy due to novel actors in tracking, device constraints, user interface limitations, and consumers' expectations specific to the television environment. Traditional browser-based protections and standard consent banners struggle to ensure adequate privacy on emerging platforms. The issue extends beyond HbbTV, as similar risks manifest in other connected device ecosystems, including smart home environments [21], [45], [87]. These risks underscore the need for solutions not tailored to a specific ecosystem, more stringent legal enforcement, and greater industry accountability.

Recommendations. New privacy solutions must address the particular challenges of connected home environments where various manufacturers' devices engage in different ecosystems. Consent mechanisms require substantial redesign, since nudging strategies on or for devices with fewer interaction features (e.g., TVs or smart thermostats [73]) undermine genuine user choice. System developers should be aware of the data practices disclosed in privacy policies and cooperate closely with data protection officers to prevent discrepancies. At the same time, consumer awareness should be raised through educational initiatives, on-screen disclosures, and adequate privacy controls, empowering people to make informed choices about data collection. Measures to increase industry accountability reflected by transparent reporting, privacy-focused design, and proactive compliance with regulations can help protect user privacy.

Supplementary Material, Code Artifacts, and Data. To foster future research, we publish our analysis code, data, data processing pipeline, and other supplementary information online at: https://github.com/internet-sicherheit/Tracking-and-Transparency-Mechanisms-in-the-HbbTV-Ecosystem.

Acknowledgment

This work was supported by the German Federal Ministry of Education and Research (grants 16KIS1629, 16KIS1628K "UbiTrans", 16KIS1648 "DigiFit"), the German Research Foundation (grant 462287308, BE1422/27-1), the Helmholtz Association (topic "46.23 Engineering Secure Systems"), and the Google Cloud Research Credits program (EDU Credit 330580204).

References

[1] Y. Aafer, W. You, Y. Sun, Y. Shi, X. Zhang, and H. Yin, "Android SmartTVs Vulnerability Discovery via Log-Guided Fuzzing," in *Proceedings of the 30th USENIX Security Symposium*, ser. USENIX Security 2021. Berkeley, CA, USA: USENIX Association, 2021, pp. 2759–2776. [Online]. Available: https://www.usenix.org/conference/usenixsecurity21/presentation/aafer 3

[2] G. Acar, C. Eubank, S. Englehardt, M. Juarez, A. Narayanan, and C. Diaz, "The Web Never Forgets: Persistent Tracking Mechanisms in the Wild," in *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*, ser. CCS '14. New York, NY, USA: Association for Computing Machinery, 2014, pp. 674–689. [Online]. Available: https://doi.org/10.1145/2660267.2660347 7

[3] A. Adhikari, S. Das, and R. Dewri, "Evolution of Composition, Readability, and Structure of Privacy Policies over Two Decades," *Proceedings on Privacy Enhancing Technologies*, vol. 2023, no. 3, pp. 138–153, May 2023. [Online]. Available: https://doi.org/10.56553/popets-2023-0074 11

[4] J. Ahmad, F. Li, and B. Luo, "IoTPrivComp: A Measurement Study of Privacy Compliance in IoT Apps," in *Computer Security – ESORICS 2022*, ser. ESORICS 2022. Cham, Switzerland: Springer Nature Switzerland, 2022, pp. 589–609. [Online]. Available: https://doi.org/10.1007/978-3-031-17146-8_29 3

[5] R. Amos, G. Acar, E. Lucherini, M. Kshirsagar, A. Narayanan, and J. Mayer, "Privacy Policies over Time: Curation and Analysis of a Million-Document Dataset," in *The Web Conference 2021 – Proceedings of the World Wide Web Conference*, ser. WWW '21. New York, NY, USA: ACM, 2021, pp. 2165–2176. [Online]. Available: https://doi.org/10.1145/3442381.3450048 3

[6] B. Andow, S. Y. Mahmud, W. Wang, J. Whitaker, W. Enck, B. Reaves, K. Singh, and T. Xie, "PolicyLint: Investigating Internal Privacy Policy Contradictions on Google Play," in *Proceedings of the 28th USENIX Security Symposium*, ser. USENIX Security 2019. Berkeley, CA, USA: USENIX Association, 2019, pp. 585–602. [Online]. Available: https://www.usenix.org/conference/usenixsecurity19/presentation/andow 11

[7] S. Arora, H. Hosseini, C. Utz, V. B. Kumar, T. Dhellemmes, A. Ravichander, P. Story, J. Mangat, R. Chen, M. Degeling, T. Norton, T. Hupperich, S. Wilson, and N. Sadeh, "A Tale of Two Regulatory Regimes: Creation and Analysis of a Bilingual Privacy Policy Corpus," in *Proceedings of the 13th Conference on Language Resources and Evaluation*, ser. LREC 2022. Paris, France: ELRA, 2022, pp. 5460–5472. [Online]. Available: http://www.lrec-conf.org/proceedings/lrec2022/pdf/2022.lrec-1.585.pdf 11

[8] Article 29 Data Protection Working Party, "Guidelines on Consent under Regulation 2016/679," European Commission, Brussels, Belgium, Tech. Rep. 17/EN WP259 rev.01, Oct. 2018, last accessed 25 April 2025. [Online]. Available: https://ec.europa.eu/newsroom/article29/items/623051/en 12

[9] Y. Bachy, F. Basse, V. Nicomette, E. Alata, M. Kaâniche, J.-C. Courrège, and P. Lukjanenko, "Smart-TV Security Analysis: Practical Experiments," in *Proceedings of the 2015 45th Annual IEEE/IFIP International Conference on Dependable Systems and Networks*, ser. DSN 2015. Washington, DC, USA: IEEE Computer Society, 2015, pp. 497–504. [Online]. Available: https://doi.org/10.1109/DSN.2015.41 3

[10] Y. Bachy, V. Nicomette, M. Kaâniche, and E. Alata, "Smart-TV Security: Risk Analysis and Experiments on Smart-TV Communication Channels," *Journal of Computer Virology and Hacking Techniques*, vol. 15, no. 1, pp. 61–76, Mar. 2019. [Online]. Available: https://doi.org/10.1007/s11416-018-0320-3 1, 3

[11] V. Belcheva, T. Ermakova, and B. Fabian, "Understanding Website Privacy Policies – A Longitudinal Analysis Using Natural Language Processing," *Information*, vol. 14, no. 11, Nov. 2023. [Online]. Available: https://doi.org/10.3390/info14110622 3

[12] S. Black, "Unified hosts file with base extensions," 2023, used list from November 1, 2023; version 3.14.21; last accessed 25 April 2025. [Online]. Available: https://raw.githubusercontent.com/StevenBlack/hosts/master/hosts 7

[13] Bundesministerium der Justiz, "Gesetz über den Datenschutz und den Schutz der Privatsphäre in der Telekommunikation und bei bei digitalen Diensten (Telekommunikation-Digitale-Dienste-Datenschutz-Gesetz – TDDDG)," Aug. 2021, in German; English title: "Telecommunications Digital Services Data Protection Act". [Online]. Available: https://www.gesetze-im-internet.de/ttdsg/ 11

[14] BurdaForward GmbH, "Serien: Alle Genres in der Übersicht," Aug. 2024, in German; English title: "TV Shows: Overview of All Genres"; last accessed 25 April 2025. [Online]. Available: https://www.tvspielfilm.de/serien/genre/ 6

[15] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. New York, NY, USA: Routledge, 1988. [Online]. Available: https://doi.org/10.4324/9780203771587 5

[16] A. Cortesi, M. Hils, T. Kriechbaumer, and contributors, "mitmproxy – an interactive HTTPS proxy," 2010–, version 9.0; last accessed 25 April 2025. [Online]. Available: https://mitmproxy.org/ 3

[17] H. Damghani, H. Hosseinian, and L. Damghani, "Privacy Risks of Hybrid Broadcast Broadband TV (HbbTV)," in *5th Conference on Knowledge-Based Engineering and Innovation*, ser. KBEI 2019. Washington, DC, USA: IEEE, 2019, pp. 61–67. [Online]. Available: https://www.academia.edu/44831260/Privacy_Risks_of_Hybrid_Broadcast_Broadband_TV_HbbTV_ 1, 3

[18] Dataxis SARL, "Addressable TV in Europe: where do we stand one decade after AdSmart?" 2024, last accessed 25 April 2025. [Online]. Available: https://dataxis.com/researches-highlights/1054470/addressable-tv-in-europe-where-do-we-stand-one-decade-after-adsmart/ 1

[19] M. Degeling, C. Utz, C. Lentzsch, H. Hosseini, F. Schaub, and T. Holz, "We Value Your Privacy ... Now Take Some Cookies: Measuring the GDPR's Impact on Web Privacy," in *Proceedings of the 26th Annual Network and Distributed System Security Symposium*, ser. NDSS 2019. Reston, VA, USA: Internet Society, 2019. [Online]. Available: https://www.ndss-symposium.org/ndss-paper/we-value-your-privacy-now-take-some-cookies-measuring-the-gdprs-impact-on-web-privacy/ 3, 10, 11

[20] N. Demir, T. Urban, C. Wressnegger, and N. Pohlmann, "A Large-Scale Study of Cookie Banner Interaction Tools and Their Impact on Users' Privacy," *Proceedings on Privacy Enhancing Technologies*, vol. 2024, no. 1, pp. 5–20, 2024. [Online]. Available: https://doi.org/10.56553/popets-2024-0002 6

[21] J. S. Edu, J. M. Such, and G. Suarez-Tangil, "Smart Home Personal Assistants: A Security and Privacy Review," *ACM Computing Surveys*, vol. 53, no. 6, 2020. [Online]. Available: https://doi.org/10.1145/3412383 12

[22] S. Englehardt, J. Han, and A. Narayanan, "I never signed up for this! Privacy implications of email tracking," *Proceedings on Privacy Enhancing Technologies*, vol. 2018, no. 1, pp. 109–126, Jan. 2018. [Online]. Available: https://doi.org/10.1515/popets-2018-0006 7

[23] European Broadcasting Union, "HbbTV Deployments," 2024, last accessed 25 April 2025. [Online]. Available: https://www.hbbtv.org/deployments/ 1, 2

[24] European Broadcasting Union (EBU), Comité Européen de Normalisation Electrotechnique (CENELEC), and European Telecommunications Standards Institute (ETSI), "HbbTV Specifications," Mar. 2023, last accessed 25 April 2025. [Online]. Available: https://www.hbbtv.org/resource-library/#specifications 1

[25] European Court of Justice, "Judgment of the Court of 1 October 2019 in Case C-673/17 – Planet49," Oct. 2019. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:62017CJ0673 10

[26] European Parliament and the Council of the European Union, "Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector," Official Journal of the European Communities, Jul. 2002. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A32002L0058 1

[27] ——, "Directive 2009/136/EC of the European Parliament and of the Council of 25 November 2009 amending Directive 2002/22/EC, Directive 2002/58/EC and Regulation (EC) No 2006/2004," Official Journal of the European Union, L 337/11, Nov. 2009. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32009L0136 3

[28] ——, "Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)," Official Journal of the European Union, L 119/1, Apr. 2016. [Online]. Available: https://eur-lex.europa.eu/eli/reg/2016/679/oj 1, 8

[29] European Telecommunications Standards Institute, "Hybrid Broadcast Broadband TV – Technical Specification," Sep. 2023, last accessed 25 April 2025. [Online]. Available: https://www.etsi.org/deliver/etsi_ts/102700_102799/102796/01.07.01_60/ts_102796v010701p.pdf 2, 5

[30] FingerprintJS, Inc, "FingerprintJS: Browser Fingerprinting Library," 2024, last accessed 25 April 2025. [Online]. Available: https://github.com/fingerprintjs/fingerprintjs/ 8

[31] W. Fischer, *Broadcast over Internet, HbbTV, OTT, Streaming*, 4th ed. Cham, Switzerland: Springer Nature Switzerland AG, 2020, pp. 903–913. [Online]. Available: https://doi.org/10.1007/978-3-030-32185-7_44 1

[32] I. Fouad, N. Bielova, A. Legout, and N. Sarafijanovic-Djukic, "Missed by Filter Lists: Detecting Unknown Third-Party Trackers with Invisible Pixels," *Proceedings on Privacy Enhancing Technologies*, vol. 2020, no. 2, pp. 499–518, Apr. 2020. [Online]. Available: https://doi.org/10.2478/popets-2020-0038 7

[33] M. Ghiglieri, "I Know What You Watched Last Sunday – A New Survey Of Privacy In HbbTV," in *Workshop on Web 2.0 Security & Privacy 2014*, ser. W2SP 2014. Washington, DC, USA: IEEE, 2014. [Online]. Available: https://web.archive.org/web/20170319163504/http://w2spconf.com/2014/papers/ghiglieri_hbbtv%20survey.pdf 1, 3

[34] M. Ghiglieri and E. Tews, "A Privacy Protection System for HbbTV in Smart TVs," in *Proceedings of the 2014 IEEE 11th Consumer Communications and Networking Conference*, ser. CCNC 2014. Washington, DC, USA: IEEE, 2014, pp. 357–362. [Online]. Available: https://doi.org/10.1109/CCNC.2014.6866595 3, 4

[35] M. Ghiglieri and M. Waidner, "HbbTV Security and Privacy: Issues and Challenges," *IEEE Security & Privacy*, vol. 14, no. 3, pp. 61–67, May 2016. [Online]. Available: https://doi.org/10.1109/MSP.2016.54 1, 3

[36] M. Gruber, C. Höfig, M. Golla, T. Urban, and M. Große-Kampmann, "'We may share the number of diaper changes': A Privacy and Security Analysis of Mobile Child Care Applications," *Proceedings on Privacy Enhancing Technologies*, vol. 2022, no. 3, pp. 394–414, 2022. [Online]. Available: https://doi.org/10.56553/popets-2022-0078 3

[37] A. A. Hagberg, D. A. Schult, and P. J. Swart, "Exploring Network Structure, Dynamics, and Function using NetworkX," in *Proceedings of the 7th Python in Science Conference*, ser. SciPy 2008, 2008, pp. 11 – 15. [Online]. Available: http://conference.scipy.org.s3-website-us-east-1.amazonaws.com/proceedings/SciPy2008/paper_2/ 9

[38] M. Hils, D. W. Woods, and R. Böhme, "Measuring the Emergence of Consent Management on the Web," in *Proceedings of the 2020 ACM Internet Measurement Conference*, ser. IMC '20. New York, NY, USA: ACM, 2020, pp. 317–332. [Online]. Available: https://doi.org/10.1145/3419394.3423647 3

[39] H. Hosseini, M. Degeling, C. Utz, and T. Hupperich, "Unifying Privacy Policy Detection," *Proceedings on Privacy Enhancing Technologies*, vol. 2021, no. 4, pp. 480–499, Jul. 2021. [Online]. Available: https://doi.org/10.2478/popets-2021-0081 10

[40] H. Hosseini, C. Utz, M. Degeling, and T. Hupperich, "A Bilingual Longitudinal Analysis of Privacy Policies Measuring the Impacts of the GDPR and the CCPA/CPRA," *Proceedings on Privacy Enhancing Technologies*, vol. 2024, no. 2, pp. 434–463, Feb. 2024. [Online]. Available: https://doi.org/10.56553/popets-2024-0058 3

[41] Human Signal, Inc., "Label Studio – Open Source Data Labeling," Aug. 2024, last accessed 25 April 2025. [Online]. Available: https://labelstud.io/ 9

[42] S. Iyer, "pywebostv 0.8.2," 2023, last accessed 25 April 2025. [Online]. Available: https://pypi.org/project/pywebostv/0.8.2/ 4

[43] I. Kamara and P. De Hert, "Understanding the Balancing Act Behind the Legitimate Interest of the Controller Ground: A Pragmatic Approach," *Brussels Privacy Hub Working Paper*, vol. 4, no. 12, Aug. 2018. [Online]. Available: https://brusselsprivacyhub.eu/BPH-Working-Paper-VOL4-N12.pdf 11

[44] H. Kamran, "Smart TV Blocklist," Oct. 2023, last accessed 31 August 2024. [Online]. Available: https://github.com/hkamran80/blocklists/blob/main/smart-tv 7

[45] S. M. Karunarathne, N. Saxena, and M. K. Khan, "Security and Privacy in IoT Smart Healthcare," *IEEE Internet Computing*, vol. 25, no. 4, pp. 37–48, Jul. 2021. [Online]. Available: https://doi.org/10.1109/MIC.2021.3051675 12

[46] C. Kohlschütter, P. Fankhauser, and W. Nejdl, "Boilerplate Detection using Shallow Text Features," in *Proceedings of the Third ACM International Conference on Web Search and Data Mining*, ser. WDSM '10. New York, NY, USA: ACM, 2010, pp. 441–450. [Online]. Available: https://doi.org/10.1145/1718487.1718542 10

[47] W. H. Kruskal and W. A. Wallis, "Use of Ranks in One-Criterion Variance Analysis," *Journal of the American Statistical Association*, vol. 47, no. 260, pp. 583–621, Dec. 1952. [Online]. Available: https://doi.org/10.2307/2280779 5

[48] L. Kyi, S. Ammanaghatta Shivakumar, C. T. Santos, F. Roesner, F. Zufall, and A. J. Biega, "Investigating Deceptive Design in GDPR's Legitimate Interest," in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, ser. CHI 2023. New York, NY, USA: ACM, 2023. [Online]. Available: https://doi.org/10.1145/3544548.3580637 11

[49] L. Lebanoff and F. Liu, "Automatic Detection of Vague Words and Sentences in Privacy Policies," in *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*. Shroudsburg, PA, USA: ACL, 2018, pp. 3508–3517. [Online]. Available: https://aclanthology.org/D18-1387 11

[50] LG Electronics, Inc., "webOS API," 2023, last accessed 25 April 2025. [Online]. Available: https://webostv.developer.lge.com/develop/references/webostvjs-webos 3, 4, 5

[51] M. Majchrowicz and P. Duch, "Analysis of Tizen Security Model and Ways of Bypassing It on Smart TV Platform," *Applied Sciences*, vol. 11, no. 24, 2021. [Online]. Available: https://doi.org/10.3390/app112412031 3

[52] G. S. Manku, A. Jain, and A. Das Sarma, "Detecting Near-Duplicates for Web Crawling," in *Proceedings of the 16th International Conference on World Wide Web*, ser. WWW '07. New York, NY, USA: ACM, 2007, pp. 141–150. [Online]. Available: https://doi.org/10.1145/1242572.1242592 11

[53] H. B. Mann and D. R. Whitney, "On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other," *The Annals of Mathematical Statistics*, vol. 18, no. 1, pp. 50–60, 1947. [Online]. Available: https://doi.org/10.1214/aoms/1177730491 8

[54] C. Matte, N. Bielova, and C. Santos, "Do Cookie Banners Respect my Choice? Measuring Legal Compliance of Banners from IAB Europe's Transparency and Consent Framework," in *Proceedings of the 2020 IEEE Symposium on Security and Privacy*, ser. SP '20. Washington, DC, USA: IEEE Computer Society, 2020, pp. 791–809. [Online]. Available: https://ieeexplore.ieee.org/document/9152617 3

[55] H. M. Moghaddam, G. Acar, B. Burgess, A. Mathur, D. Y. Huang, N. Feamster, E. W. Felten, P. Mittal, and A. Narayanan, "Watching You Watch: The Tracking Ecosystem of Over-the-Top TV Streaming Devices," in *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security*, ser. CCS '19. New York, NY, USA: ACM, 2019, pp. 131–147. [Online]. Available: https://doi.org/10.1145/3319535.3354198 2

[56] T. T. Nguyen, M. Backes, N. Marnau, and B. Stock, "Share First, Ask Later (or Never?) Studying Violations of GDPR's Explicit Consent in Android Apps," in *Proceedings of the 30th USENIX Security Symposium*, ser. USENIX Security '21. Berkeley, CA, USA: USENIX Association, 2021, pp. 3667–3684. [Online]. Available: https://www.usenix.org/conference/usenixsecurity21/presentation/nguyen 3

[57] T. T. Nguyen, M. Backes, and B. Stock, "Freely Given Consent? Studying Consent Notice of Third-Party Tracking and Its Violations of GDPR in Android Apps," in *Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security*, ser. CCS '22. New York, NY, USA: ACM, 2022, pp. 2369–2383. [Online]. Available: https://doi.org/10.1145/3548606.3560564 3

[58] M. Nouwens, I. Liccardi, M. Veale, D. Karger, and L. Kagal, "Dark Patterns after the GDPR: Scraping Consent Pop-ups and Demonstrating their Influence," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, ser. CHI '20. New York, NY, USA: ACM, 2020. [Online]. Available: https://dl.acm.org/doi/10.1145/3313831.3376321 3

[59] OneTrust LLC, "Cookiepedia – All You Need to Know About Cookies," Aug. 2024. [Online]. Available: https://cookiepedia.co.uk/ 6

[60] S. Pandey, Y. J. Won, M.-J. Choi, and J.-M. Gil, "Community Model for Smart TV Over the Top Services," *Journal of Information Processing Systems*, vol. 12, no. 4, pp. 577–590, Dec. 2016. [Online]. Available: https://dx.doi.org/10.3745/JIPS.03.0057 2

[61] G. Park, "The Changing Wind of Data Privacy Law: A Comparative Study of the European Union's General Data Protection Regulation and the 2018 California Consumer Privacy Act," *UC Irvine Law Review*, vol. 10, no. 4, pp. 1455–1490, Jun. 2020. [Online]. Available: https://escholarship.org/uc/item/8562f0v0 12

[62] Perflyst, "PiHoleBlocklist," 2018. [Online]. Available: https://github.com/Perflyst/PiHoleBlocklist?tab=readme-ov-file 7

[63] Pi-hole LLC, "Pi-hole," 2023, last accessed 25 April 2025. [Online]. Available: https://pi-hole.net/ 7

[64] RootMyTV, "RootMyTV 2.0," 2023, last accessed 25 April 2025. [Online]. Available: https://github.com/RootMyTV/RootMyTV.github.io 3

[65] A. Santani, M. Gangaramani, B. Chopra, P. Choudhary, and K. Samdani, "An Overview of Architecture and Security Issues of a Smart TV," in *Proceedings of the 6th International Conference on Communication and Electronics Systems*, ser. ICCES 2021. Washington, DC, USA: IEEE, 2021, pp. 1835–1843. [Online]. Available: https://doi.org/10.1109/ICCES51350.2021.9488939 3

[66] Satindex, "Astra und Hotbird Sender + Frequenzen auf Satindex.de," Feb. 2025, last accessed 25 April 2025. [Online]. Available: https://www.satindex.de/ 8

[67] L. Schmidt, H. Hosseini, and T. Hupperich, "Assessing the Security and Privacy of Baby Monitor Apps," *Journal of Cybersecurity and Privacy*, vol. 3, no. 3, pp. 303–326, Jun. 2023. [Online]. Available: https://doi.org/10.3390/jcp3030016 3

[68] SES Germany GmbH, "Astra Channel Search," 2024, in German; last accessed 25 April 2025. [Online]. Available: https://astra.de/tv-radio-mehr/senderuebersicht 5

[69] Statista Inc., "Number of TV viewers in Europe from 2018 to 2028," 2024, last accessed 25 April 2025. [Online]. Available: https://www.statista.com/forecasts/1259664/tv-viewers-europe-number/ 1

[70] ——, "TV penetration rate in Europe from 2018 to 2028," 2024, last accessed 25 April 2025. [Online]. Available: https://www.statista.com/forecasts/1259671/tv-viewers-europe-penetration-rate/ 1

[71] C. Tagliaro, F. Hahn, R. Sepe, A. Aceti, and M. Lindorfer, "I Still Know What You Watched Last Sunday: Privacy of the HbbTV Protocol in the European Smart TV Landscape," in *Proceedings of the 2023 Network and Distributed System Security Symposium*, ser. NDSS 2023. Reston, VA, USA: Internet Society, 2023. [Online]. Available: https://www.ndss-symposium.org/ndss-paper/i-still-know-what-you-watched-last-sunday-privacy-of-the-hbbtv-protocol-in-the-european-smart-tv-landscape/ 1, 3, 4, 7, 12

[72] ——, "Investigating HbbTV Privacy Invasiveness Across European Countries," in *2023 Workshop on Learning from Authoritative Security Experiment Results*, ser. LASER 2023. Reston, VA, USA: Internet Society, 2023, pp. 1–8. [Online]. Available: https://www.ndss-symposium.org/wp-content/uploads/laser2023-102-paper.pdf 1, 3

[73] R. Tamas, W. O'Brien, and M. Santana Quintero, "Evolving interaction: a qualitative investigation of user mental models for smart thermostat users," *Architectural Science Review*, vol. 66, no. 3, pp. 155–171, 2023. [Online]. Available: https://doi.org/10.1080/00038628.2023.2201253 12

[74] The EasyList authors, "EasyList," Mar. 2023, used list from March 23, 2023; version 202303230338. [Online]. Available: https://easylist.to/easylist/easylist.txt 6, 7

[75] ——, "EasyPrivacy," Jul. 2024, used list from July 22, 2024; version 202407221302. [Online]. Available: https://easylist.to/easylist/easyprivacy.txt 7

[76] M. Tileria and J. Blasco, "Watch Over Your TV: A Security and Privacy Analysis of the Android TV Ecosystem," *Proceedings on Privacy Enhancing Technologies*, vol. 2022, no. 3, pp. 692–710, 2022. [Online]. Available: https://doi.org/10.56553/popets-2022-0092 3

[77] J. A. Tomain, "Online Privacy & the First Amendment: An Opt-In Approach to Data Processing," *University of Cincinnati Law Review*, vol. 83, no. 1, pp. 1–71, 2014. [Online]. Available: https://www.repository.law.indiana.edu/facpub/2649 12

[78] T. Urban, D. Tatang, M. Degeling, T. Holz, and N. Pohlmann, "The Unwanted Sharing Economy: An Analysis of Cookie Syncing and User Transparency under GDPR," 2018. [Online]. Available: http://arxiv.org/abs/1811.08660 7

[79] ——, "Measuring the Impact of the GDPR on Data Sharing in Ad Networks," in *Proceedings of the 15th ACM Asia Conference on Computer and Communications Security*, ser. ASIA CCS '20. New York, NY, USA: ACM, 2020, pp. 222–235. [Online]. Available: https://doi.org/10.1145/3320269.3372194 6, 7

[80] T. Urban, D. Tatang, T. Holz, and N. Pohlmann, "Towards Understanding Privacy Implications of Adware and Potentially Unwanted Programs," in *Proceedings of the 23rd European Symposium on Research in Computer Security*, ser. ESORICS 2018. Cham, Switzerland: Springer Nature Switzerland AG, 2018, pp. 449–469. [Online]. Available: https://doi.org/10.1007/978-3-319-99073-6_22 7

[81] ——, "Analyzing leakage of personal information by malware," *Journal of Computer Security*, vol. 27, no. 4, pp. 459–481, Jul. 2019. [Online]. Available: http://dx.doi.org/10.3233/JCS-191287 7

[82] C. Utz, M. Degeling, S. Fahl, F. Schaub, and T. Holz, "(Un)informed Consent: Studying GDPR Consent Notices in the Field," in *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security*, ser. CCS '19. New York, NY, USA: ACM, 2019, pp. 973–990. [Online]. Available: https://doi.org/doi/10.1145/3319535.3354212 3

[83] J. Varmarken, H. Le, A. Shuba, Z. Shafiq, and A. Markopoulou, "The TV is Smart and Full of Trackers: Towards Understanding the Smart TV Advertising and Tracking Ecosystem," *Proceedings on Privacy Enhancing Technologies*, vol. 2020, no. 2, pp. 129–154, 2020. [Online]. Available: https://doi.org/10.2478/popets-2020-0021 2

[84] W3C Tracking Protection Working Group, "Tracking Preference Expression [DNT]," Jan. 2019, last accessed 25 April 2025. [Online]. Available: https://www.w3.org/TR/tracking-dnt/ 3

[85] I. Wagner, "Privacy Policies across the Ages: Content of Privacy Policies 1996–2021," *ACM Transactions on Privacy and Security*, vol. 26, no. 3, May 2023. [Online]. Available: https://doi.org/10.1145/3590152 3

[86] S. Wilson, F. Schaub, A. A. Dara, F. Liu, S. Cherivirala, P. G. Leon, M. S. Andersen, S. Zimmeck, K. M. Sathyendra, N. C. Russell, T. B. Norton, E. Hovy, J. Reidenberg, and N. Sadeh, "The Creation and Analysis of a Website Privacy Policy Corpus," in *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, ser. ACL '16. Stroudsburg, PA, USA: ACL, 2016, pp. 1330–1340. [Online]. Available: https://aclanthology.org/P16-1126/ 11

[87] K. Yu, Q. Li, D. Chen, M. Rahman, and S. Wang, "PrivacyGuard: Enhancing Smart Home User Privacy," in *Proceedings of the 20th International Conference on Information Processing in Sensor Networks*, ser. IPSN '21. New York, NY, USA: ACM, 2021, pp. 62–76. [Online]. Available: https://doi.org/10.1145/3412382.3458257 12

[88] R. N. Zaeem, R. L. German, and K. S. Barber, "PrivacyCheck: Automatic Summarization of Privacy Policies Using Data Mining," *ACM Transactions on Internet Technology*, vol. 18, no. 4, Nov. 2018. [Online]. Available: https://doi.org/10.1145/3127519 11

[89] Y. Zhang, S. Ma, T. Chen, J. Li, R. H. Deng, and E. Bertino, "EvilScreen Attack: Smart TV Hijacking via Multi-channel Remote Control Mimicry," *IEEE Transactions on Dependable and Secure Computing*, vol. 21, no. 4, pp. 1544–1556, Jul. 2022. [Online]. Available: https://doi.org/10.1109/TDSC.2023.3286182 3