The Writing on the Wall and 3D Digital Twins: Personal Information in (not so) Private Real Estate

Rachel McAmis and Tadayoshi Kohno, *University of Washington*

https://www.usenix.org/conference/usenixsecurity23/presentation/mcamis

This paper is included in the Proceedings of the 32nd USENIX Security Symposium.

August 9–11, 2023 • Anaheim, CA, USA

978-1-939133-37-3

Open access to the Proceedings of the 32nd USENIX Security Symposium is sponsored by USENIX.

The Writing on the Wall and 3D Digital Twins: Personal Information in (not so) Private Real Estate

Rachel McAmis
University of Washington
rcmcamis@cs.washington.edu

Tadayoshi Kohno
University of Washington
yoshi@cs.washington.edu

Abstract

Online real estate companies are starting to offer 3D virtual tours of homes (3D digital twins). We qualitatively analyzed 44 3D home tours with personal artifacts visible on Zillow and assessed each home for the extent and type of personal information shared. Using a codebook we created, we analyzed three categories of personal information in each home: government-provided guidance of what not to share on the internet, identity information, and behavioral information. Our analysis unearthed a wide variety of sensitive information across all homes, including names, hobbies, employment and education history, product preferences (e.g., pantry items, types of cigarettes), medications, credit card numbers, passwords, and more. Based on our analysis, residents both employed privacy protections and had privacy oversights. We identify potential adversaries that might use 3D tour information, highlight additional sensitive sources of indoor space information, and discuss future tools and policy changes that could address these issues.

1 Introduction

Privacy concerns often arise in data collection of public outdoor spaces. For example, in Google Street View, there have been many attempts to blur faces, license plates, and other sensitive information by citizens [20, 21], researchers [16, 17, 45, 62], and Google [20, 21]. In contrast to public spaces, the inside of a home may be intuited as private, unavailable to the prying eye [32]. This assumption of privacy in one's personal space, however, does not hold when someone must sell their home. Companies like Apartments.com, Vacasa, Zillow, and other short- and long-term rental/home-buying companies use both photos and, increasingly, "3D" virtual tours to advertise homes and rental properties. When someone does not remove all personal items from view in their home, leaving their decor and belongings as-is, their personal information, such as identity and behavior, can be inferred from the items that are visible.

While photos of real estate have been around for a while, we seek to understand how the more recently developed public 3D tours of the home pose security and privacy concerns.

We aim to answer the following research questions:

1. What types of personal information are leaked through online real estate 3D tours?
2. How might potential adversaries use this personal information to their advantage?

To foreshadow our findings, when observing these 3D tours, we found manifold examples of traditionally sensitive information being shared, from medication labels and license plates to passwords and credit card numbers.

We qualitatively analyzed 44 homes with 3D virtual tours for sale on Zillow, filtering specifically for homes with personal artifacts visible in the 3D tour, and then sampling across 44 different states in the United States. Though we focus on homes on the online real estate website Zillow and 3D tours of Matterport, these two companies are not the only sources for leaking personal information in private real estate via 3D images (See Section 2.1).

The spectrum of personal information found in the public 3D virtual home tours underscores how sensitive, and how useful to adversaries, indoor space information can be. The types of information we find in 3D tours not only raise concern for private real estate websites, but also for other sources of indoor information. 3D mappings collected from Roomba vacuum cleaners [4, 52], camera information from home security companies like Amazon Ring [6, 22, 43, 54], and existing public research datasets of inside people's homes [5] corroborate the need to protect against this unintentional and sensitive information leakage. Online real estate 3D tours and research datasets are concerning for their immediate availability to the public, but information about homes not publicly available are also concerning given potential future data breaches or companies selling information to third-parties.

Contributions.

- To our knowledge, we are the first to systematically explore leakage of personal information in online 3D real

estate tours from a scientific perspective.

- Through the codebook we developed, we extensively analyze each 3D tour across 30 different attributes and 3 categories of personal information.
- We find concrete examples of personal information shared by real residents that violate common internet-sharing guidelines.
- We note several failed privacy protections used when uploading 3D tours to a real estate website. This includes users relying on the quality of 3D tour technology, where the technology fails to provide the expected protection.
- We also note successful privacy mitigations, including that poor camera quality actually protects residents from leaking information.
- We discuss the implications and potential threats of indoor home information leakage, and identify other concerning sources of home information in addition to online real estate.

2 Background and Related Work

2.1 3D Tours and Real Estate

3D "Digital Twins". A 3D "Digital Twin" refers to a copy of a real space. "Digital Twins" are used in many scenarios, from Google Street View, to touring museums virtually, to analyzing insurance damages in a home [1]. They are also used when selling homes on online real estate websites.

Matterport is one such company that provides 3D "Digital Twins", which we also refer to as "3D tours". Through Matterport alone, over 5 million spaces have been mapped [37], for selling real estate and many other purposes, and will continue to increase given the benefits of increased interactions [70] and decreased sale times of homes [41]. Zillow itself also produces 3D tour software, though many sellers opt for loading a Matterport 3D tour instead. As of 2020, 90% of real estate 3D tours used Matterport [41].

Figures 1 and 2 show examples of ways that Matterport allows you to navigate a 3D tour online, and two additional examples are included in Appendix B, (these figures are from an example on the Matterport web page, which Matterport encourages sharing images of [60]). The red arrows (added by us) indicate where the viewer of the 3D tour would zoom in or click on the virtual tour to get to the following figure. Matterport first shows a "dollhouse" feature that allows the viewer to see the overall shape and layout of an indoor space, and quickly navigate to a specific room by rotating the dollhouse. The viewer can also click different areas within the 3D tour to move the image in that direction, e.g., clicking at the right of the screen to navigate to the right.

The trend of 3D tours is probably not going away any time soon given its benefits to home sellers. According to Zillow, homes with 3D tours get 43% more views [70]. According to Matterport, the time to sell homes that have 3D tours is 20% faster for homes with 3D tours versus those with just static photos [41].

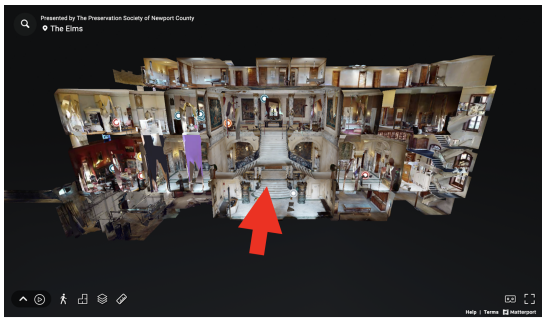

Figure 1: Matterport's "Dollhouse" feature. The red arrow overlaying the screenshot represents where a viewer would click to get to the image in Figure 2.

Figure 2: Viewers can navigate to different areas of an indoor space. The red arrow overlaying the screenshot represents where a viewer would click to move up the stairs.

Online Real Estate Business. In the online real estate market, homeowners can sell their homes directly on websites like Redfin and Zillow, or they can hire a listing agent to market the homes online. Listing agents are often members of their local MLS, or Multiple Listing Service. An MLS is a database of professionally managed homes that can be listed on many sites at once, and websites like Redfin and Zillow pull details from the MLS automatically and put home details on their websites. An MLS allows multiple real estate brokers, those who represent home sellers or buyers, to communicate with each other and find homes to buy/sell [34].

On Zillow and Redfin, anyone can view homes; there is no need for account creation. Both sites offer 3D tours of homes if the home owners or brokers choose to upload them, and the 3D tours can also be uploaded to the MLS. Zillow allows 3D tours from a variety of different 3D tour companies, including Matterport and Zillow's own 3D tour service.

Awareness of 3D Tour and Online Real Estate Issues. Starting soon before our data collection and continuing concurrently with collection, articles began to come out following a podcast discussion on a case in which residents unintentionally leaked information through a 3D tour on a UK-based real estate website called Rightmove, including failing to blur family photos, revealing pet names, showing a stairlift invoice, and revealing books that might leak political ideologies [31, 58]. Another report of information leakage happened in September 2021, where two television costars were identified through items in their 3D home tour, including handwritten notes to the residents [58].

Previous discussion has shown that people indeed are concerned if their data stays on real estate sites like Zillow too long after the home is sold, including photos of their home [19]. A 2016 lawsuit shows concern about property owners taking 3D virtual tours without the tenant's consent [8].

3D tours also pose technological concerns. In addition to a manual blurring feature, Matterport currently has a beta face-blurring feature [39], which automatically detects and blurs faces in the 3D tours. In 2020, Matterport published an article about their concerns with the beta feature following racially biased blurring: a family used the feature to blur faces in their 3D tour, and it resulted in all Black family members having their faces blurred, while the white family members' faces were not detected and left un-blurred [49].

Interest in Home Data. Many already identify the usefulness of indoor home data for purposes other than selling real estate. In 2017, the CEO of i-Robot, which sells Roomba "autonomous vacuums", stated that Roomba might share data collected from the vacuum sensors with third-parties, such as Amazon, Alphabet, and Apple. Purported to be for aiding with the audio setup of smart speakers, some speculate that the data may be used for more invasive uses, such as making product suggestions based on existing items seen in the home [4, 52]. It is possible that other home products, like Amazon's Ring home security company [55], may also have extensive user data that can be later shared intentionally through third-parties, or unintentionally via data breaches. There is already concern over Ring's security practices and third-party trackers [6, 22, 43]. Additionally, Ring's Terms of Service states that they may sell or redistribute any data captured with their products [54].

2.2 Related Research

Sensitive Information in Photos. Multiple works study what people deem sensitive in photos and how likely they are to include different types of information in their photos shared with the public [35, 36, 47]. Li et al. show how people deem information about any person other than the person posting the photo as sensitive, and sharing a dirty home, medical conditions, and food or smoking are also avoided by people [36]. Other works generate tools to classify whether a photo should be public or private [61], classify the privacy risk of a photo before posting to social media [47], identify privacy concerns of background objects in photos [2], automatically blur sensitive information in photos [17, 68], and replace sensitive objects in photos with semantically similar cartoons [24].

Privacy in Homes. A number of prior works consider the privacy implications of cameras in homes on the activities of occupants [11, 15, 23, 63]. Most relevant to our work, some prior works focused on the privacy implications of cameras recording images of objects or specific locations in a home [7, 48, 56, 59]. Other works acknowledge that there are private spaces where cameras should not be be [56, 59], that homes contain private information [12, 48], and that people may be protecting the privacy of their home spaces by blurring backgrounds during virtual meetings [14, 44, 69].

2.3 Defining Terms

In the rest of the paper, we define terms the following way:

- *Resident*: The occupants of the home at the time at which the 3D tour was recorded;
- *Seller*: Either the former or current occupant who is selling the home, or the property owner in the case that residents are renting;
- *Viewer*: Anyone who visits the public-facing version of online real estate websites and engages with a 3D tour;
- *Artifact-visible 3D home tour*: When personal artifacts of a resident are visible within the 3D tour. By "artifacts", we mean belongings other than furniture and decor meant for staging a home for sale. Examples include toiletries, text documents, posters, and exercise equipment.

3 Methodology

We collected data from the Zillow real estate website, which provided a rich dataset of 3D tours of homes in the U.S. at different price points, both for rent and for sale. Zillow is one of the most popular real estate websites in the U.S. [13] and it allowed us to easily filter for our needs, namely state, price range, and having a 3D tour. All data collection, including searching for homes and qualitative coding, was done manually without any crawler-based tools.

We focused on 3D tours rather than static photographs for multiple reasons. First, the resolution and granularity of the 3D tour is higher than in photos; resolution can become more detailed as you zoom in on objects within the home (such as legible text on sensitive documents), and you can look at rooms from many different angles. While still photographs may also include sensitive details, they would share less than

or equal to the information shared via 3D tours given their limited angles and limited ability to zoom in on details. Secondly, consumers have a high incentive to use 3D tours since this emerging technology is economically beneficial to home sellers (shown by the statistics mentioned in Section 2), making the possibility of simply not including 3D tours an unhelpful solution. Due to the utility of 3D tours, home sellers in the future may opt for 3D tours as a replacement to static photos, rather than in addition to them.

We manually filtered the homes with 3D tours for only those that used a Matterport 3D tour, as opposed to a Zillow 3D tour or other company 3D tour. We chose to focus on Matterport for multiple reasons: the camera quality on these tours was usually superior to other tours, there was more flexibility on the viewer's end with zooming in and out and accessing more corners of the home, and the tour was the easiest to manually navigate.

3.1 Data Collection

The goal of data collection was to look at online, artifact-visible 3D tours distributed throughout the United States and U.S. territories, in case the personal information revealed varies significantly by geography (Zillow is only available for U.S. territories and North America [71]). The data was collected from the fall of 2021 to winter of 2022.

We began by testing different search query formats to see which formats tended to yield the most artifact-visible homes with 3D tours. An example of a query is searching "California", then ordering those results by price and filtering for only homes that have 3D tours. Other search patterns we tried included homes near the second quartile of prices, near the third quartile, and the most expensive homes in the given state. Though there is not one guaranteed pattern of personal information sharing, we observed that homes in lower price points were much more likely to have personal artifacts visible, and thus we chose this sorted order to minimize the number of homes the researchers had to search through to find an artifact-visible 3D tour.

In total, we analyzed 44 homes. 41 homes were found in the following way: for every state and U.S. territory (and Washington D.C.), we found an artifact-visible Matterport 3D tour. There has to be at least one personal artifact revealed in the 3D tour for us to determine that it is "artifact-visible". We ordered the search query by lowest to highest price in the given state and filtered for homes with a 3D tour. If we did not find an artifact-visible 3D tour on the first two pages of the query, we stopped searching for a home to analyze in that state or U.S. territory; for 9 states and for all U.S. territories, we did not find artifact-visible 3D tours.

The remaining 3 homes we analyzed were from 3 states for which we did not find artifact-visible 3D tours using the initial query format, but homes were found in a different manner: either homes for sale at a higher price or homes for rent.

3.2 Qualitative Analysis

We developed a codebook to mark home attributes related to personal information, which we divided into three categories:

- Guidance: government-provided guidelines about what not to share on the internet;
- Identity: other identity information not encapsulated in Guidance;
- Behavior: other behavioral information not encapsulated in Guidance.

There are a total of 30 codebook attributes. Section 4.3 provides observed examples of each attribute. Appendix A provides additional information about the codebook.

The government-recommended guidance we found about what not to share on the internet was surprisingly sparse. Of the few lists found of what not to share, we chose to follow the most in-depth one, which was on the Washington State Office of the Attorney General website [46].

The Identity category includes attributes such as race, gender and age. We do *not* mark race, gender, or age as revealed in the codebook unless the attributes are explicitly stated in the 3D tour. For example, an adversary may see the appearance of someone and assume their race, but we only mark strong evidence in the codebook if their race is strongly implied, such as displaying a poster for an African-American religious brotherhood. Another example is inferring gender from wall decor that says "Mr. and Mrs.", which we observed in multiple homes.

For the Behavioral and Identity category attributes, we used grounded theory [66] to iteratively develop our codebooks and determine the relevant attributes in homes to code. The codebook was developed collaboratively by Researcher 1 and Researcher 2. To calibrate the process of coding homes, Researcher 1 coded 5 homes (or over 10% of all homes) and took extensive notes on the coding of those homes. Researcher 2 then independently coded those homes while also recording notes. The researchers then compared their codes and notes. If there was a disagreement, both researchers would discuss the disagreement until a consensus was made, and potentially adjust the codes for the same attributes of the other homes based on this consensus. Researcher 1 coded all remaining 39 homes independently. Using this process, a researcher only marked an attribute as existing if the researcher saw that object in a home, e.g., a calendar with hand-written calendar events. Indeed, during the step of independently coding 10% of the homes, in the few instances of initial disagreement, those disagreements only ever resulted in one researcher adding a code which they had not previously added.

To "scan" for personal information while coding, we enter each of the rooms in the 3D tour. We first navigate to get a 360 of a room, then zoom in on particular areas of interest, as some features only become legible when zoomed in.

3.3 Ethical Considerations

This study was determined by our institution's IRB to not be considered human subjects research. Despite this determination, we consider our study as related to the human experience and thus took additional ethical precautions. For example, we anonymize information in this paper so as to avoid personal harm to individuals. When illustrations would be valuable to a reader, instead of including actual images of the inside of a home, we include anonymized renderings by an artist; these renderings illustrate key findings but may change details such as location, name, room layout, and inessential objects in the home.

We disclosed our findings to Zillow and Matterport. Because our observations suggest that information leakage through 3D images may be an industry-wide issue, we sought external guidance on how to reach the maximal number of other industry stakeholders, both in the U.S. and internationally, and are in the process of following up on that guidance; we encourage future researchers to also consider seeking external guidance from advocacy groups, national or international industry bodies, or government agencies. Additionally, it is critical to consider the disclosure of findings to the individuals for whom there might be private data exposure. While there are norms and best practices for the responsible disclosure of vulnerabilities to the makers of computing systems, we are unaware of universal norms and best practices for the disclosure to individuals who might have been harmed (or have the potential for harm). There is some previous work on user perceptions of disclosing to individuals after a data breach [29, 40], though these may apply differently when the home seller accidentally leaks their data. Constructing such individual-recipient disclosure guidelines would be valuable to the research community. In the absence of field-wide norms, we encourage researchers to seek expert guidance early, e.g., from institutions' IRBs, general counsel, and possibly organizations like the Electronic Frontier Foundation.

For our research, because Zillow does not directly share the contact information of homeowners/occupants, we were unable to contact homeowners/occupants directly; we did not send physical mail to the addresses because the residents may have moved by the time the mail was received. Additionally, we do not know whether significant private information is exposed via the 3D scans of houses that we did not study. Hence, we have asked Zillow to contact the homeowners/occupants of all houses with 3D tours and all future Zillow users about the potential for information leakage; further, we shared with Zillow the links to the houses that we found revealed the most sensitive information and encouraged Zillow to start by contacting those homeowners/occupants.

3.4 Limitations

This is an exploration of the types of information shared on Zillow 3D tours. However, it does not begin to allow us to estimate or quantify the extent of the problem over all homes on all real estate websites. The types and extent of information revealed in 3D tours may vary based on home price range, location, whether it's for sale or for rent, and many other factors. The fact that this information is shared at all is concerning, but we do not intend to assess the true percentage of homes that reveal information about residents. Zillow does not allow for web crawlers, which makes large-scale analysis of information-leakage challenging. Associated with the high manual burden of analyzing 3D tours, the scale of the study is an additional limitation. Additionally, one researcher independently coded the majority of homes; while prior work gives precedent to this method, e.g., [18, 51], the reliability of coding is lower compared to if both had coded and discussed 100% of the homes.

It is also challenging to distinguish what personal information is associated with which resident, other than when information is in their own bedroom. For example, it might be impossible to infer which person suffers from a medical condition if the medication is in the kitchen. We measure the information shared overall in the home, but not for each person. From an adversary's perspective, the type of attack performed might change depending on whether they can associate personal information with an individual versus a family or group of residents.

3.5 Threat Model

Rather than focus our analysis on physical safety, such as vulnerable home entrances or existence of security cameras, we chose to focus on personal information that could be used by remote adversaries. Thus, an adversary would not have to know the person or live nearby to cause harm.

A list of potential adversaries are discussed more in Section 5. We consider adversaries with minimal technical capabilities and backgound. We assume that adversaries have access to Zillow. We also assume that the screen quality of their device that they view 3D tours from is sufficiently high to be able to see details to a similar quality of the camera capturing the 3D tour.

For some but not all classes of attacks (see Section 5), we assume that the adversary can obtain the contact information of the resident or former resident. As discussed in Section 4, an adversary may be able to use information visible within the home to search for and identify a social media account corresponding to a resident. Otherwise, as a precondition for a phishing attack, the adversary might inquire from physically adjacent neighbors of the home for sale about the identity of the resident and thus gain their name and contact information. We additionally note that adversaries could send physical

mail to the home's address which might be forwarded if the resident no longer lives there.

4 Findings

We split the findings in order to understand the dataset at four different levels of granularity:

1. Aggregated trends over the 44 homes (Section 4.1);
2. Overall observations (Section 4.2);
3. A collection of specific real examples from the dataset, organized by the three codebook categories (Section 4.3);
4. Case studies of a subset of homes along with a discussion of potential threats (Section 4.4).

4.1 Aggregate Results

Our research analyzed 44 3D tours. 41 out of 50 states, plus Washington D.C. (but no other U.S. territories) resulted in artifact-visible homes on the first two results pages when querying from lowest to highest price. We also analyzed three additional 3D tours in states that did not show results in the initial query format, through searching through rentals rather than homes for sale. For each query result page, there were about 40 homes, meaning that we searched for a maximum of around 80 homes in each state query. In the chosen query format, some states resulted in more than one artifact-visible home, while others resulted in none.

4.1.1 Codebook Results

While it is not our intent to generalize the aggregated codebook percentages to the broader private real estate landscape, this dataset and our findings demonstrate that personal information shared in online real estate is an issue.

Extent of information leakage. The "coverage" of codebook personal information was 29%. In other words, for a given analyzed home, on average evidence for 29% of the 30 codebook attributes existed in that 3D tour. The highest variety of attributes per home revealed according to the codebook was Home 29, with 47% of attributes revealed during the 3D tour. The home with the least variety of sensitive attributes revealed was Home 5, at 10%.

Other than home address, which is known in all the 3D tours, the most common information leaked was a resident's first name, with 84% of the homes revealing this information (last name was revealed in 61% of the homes). Social security and credit status were the least common, with none of the homes revealing this information.

Sensitivity of information. Figure 3 summarizes the codebook results of the percent of homes that revealed each codebook attribute, ordered by highest perceived sensitivity to lowest perceived sensitivity. Home address is not included in the codebook but we included it in the figure to visualize relative sensitivity to the codebook attributes.

We included first and last name as most sensitive because linking residents' names with the rest of information is an important part of attacks for adversaries that do not already know who lives in the given address. To determine the levels of sensitivity for the remaining codebook attributes, we referred to the perceived information sensitivity results from two user studies [36, 42]. When [42] did not include results on one of our codebook attributes, we referred to the ordering of [36]. In the small subset of cases when the information type in the codebook was not found in either prior work, we chose the most semantically similar information type in the prior work to choose ordering. For example, while calendar events were not included in the prior work, this was most semantically similar to GPS location in [42] since appointments and events on the calendar are typically tied to a time and location, and ability in our codebook was the most similar to medical history in [42].

We note that [42] does not draw from a U.S. representative sample, so results for what is deemed sensitive might vary based on the chosen population for that study. Indeed, information sensitivity depends on the context and audience. While [36] separates sensitivity result by context and audience, [42] focuses on the marketer and consumer relationship context. Further, even within a given population, individual perceptions of privacy may vary.

Overlaps of information. Figure 4 shows the percent of homes that showed information from a combination of the 3 codebook categories. This means there was evidence of at least one category attribute in each of the overlapping categories. The most common category that homes revealed was Identity, where 97.7% of homes (all but one) revealed information about at least one attribute in that category. The least common overlap was Identity and Guidance, at 75%. 75% of the analyzed homes also reveal information in all three categories.

4.2 Observations

Failed attempts at privacy. There were sometimes attempts at privacy made that did not succeed. Matterport users relied on the technological capabilities of 3D mapping companies to blur faces, but the technology failed on many occasions. On its website, Matterport specifies that the automatic face-blurring feature is currently in beta [39]. It was clear that some of the face photo blurring in the dataset were algorithms since religious images in addition to photos of family members were blurred, such as an image of Jesus. In some cases,

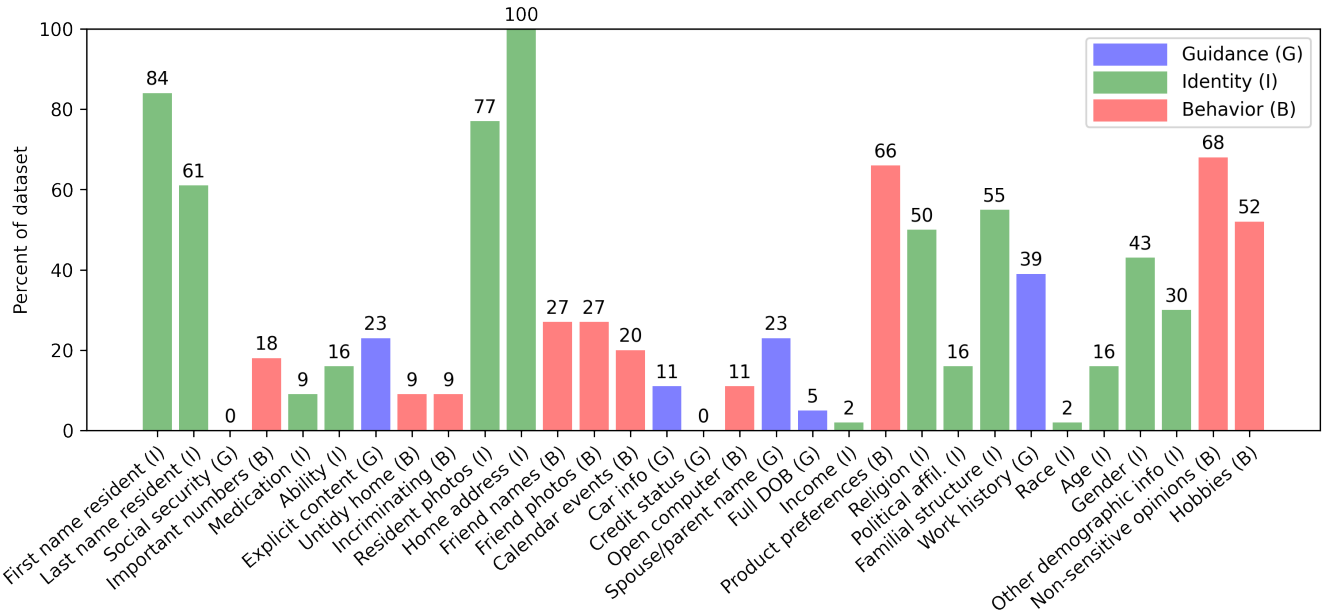

Figure 3: From most sensitive to least sensitive, percent of homes coded that revealed each codebook attribute, color-coded and labeled by codebook category. Sensitivity ordering is explained in 4.1.1. Other than home address, the most common attribute revealed was first name of a resident, at 84% of all homes.

the face is blurred from a certain angle in the 3D tour, but the blurring fails at a different angle. For example, we observed during data collection that sometimes when looking in the reflection of a mirror, the face is not blurred, even though it is blurred when looking at it directly; see Figure 5 for an artist interpretation.

Unintentional Privacy Protections. On the other hand, potentially unintentional attempts at privacy often protected residents. There was a house in which the person performing the scan to produce the 3D tour did not fully step into the bedrooms to record the information. Whatever the reasoning for this was, it ended up being infeasible to zoom in enough for the viewer to observe details within the room. While current camera quality *did* enable our research team to read some sensitive text in the 3D tours, we also observe that current camera quality was also often not good enough to enable us to zoom in on and read all potentially sensitive texts, thereby giving residents of some homes a layer of protection from prying eyes. However, as camera technology continues to improve and since sensitive text was visible in many instances, camera resolution should not be relied upon as a defense.

Ease of Identifying Residents. A combination of first and last name are useful for unfamiliar adversaries to identify residents; 59% of homes revealed both first and last name. While we do not perform the following in our own analysis, we believe that many residents would be easy to find on social media sites given the high percentage of first and last names found in the 3D tours, other information (e.g., school or employer name) visible within the home, and the high percentage of social media users. If the residents' profiles are public, this would allow adversaries to gather even more information about them, and also to contact them for reasons like social engineering.

4.3 Real Examples of Personal Information

Our intent in providing specific examples of real findings during our analysis is to surface the spectrum of types of information visible through these 3D tours. We share examples for each codebook attribute other than credit status and social security number (since these were not observed), as well as provide case studies into 5 homes in the dataset (Section 4.4).

4.3.1 Guidance Observations

Despite guidance to not share work history publicly on the internet, at least 39% of the analyzed homes revealed work history of at least one of the residents. 23% of the analyzed homes displayed explicit content, which was mostly sexual in nature. Spouse and/or parent name was also revealed in at least 23% of the homes.

Listed below is a subset of specific information leaked that violate internet-sharing guidelines:

- Car information: Full license plates and models of cars in a garage, and 3 other homes with full license plates;
- Full date of birth: Baby name, weight, and birthday near a cradle all in 1 home;

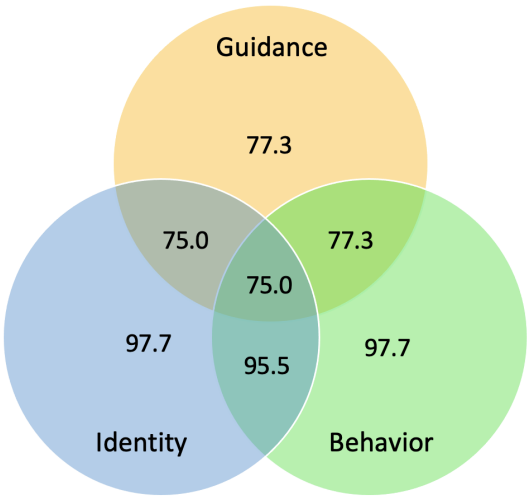

Figure 4: Percent of homes that showed evidence for the given overlaps of categories. For example, 97.7% of homes revealed some identity information, and 95.5% of homes revealed both identity and behavior information.

Figure 5: An artist interpretation of when a photo is blurred at one angle and then not blurred when seen in the mirror. Art by Akira Ohiso.

- Incriminating evidence: Letter from a city's municipal court on a bed, indicating a legal violation;
- Explicit words/photos: many sexually explicit messages and drawings on whiteboards in at least 3 homes, including a whiteboard in 1 home describing residents' sex habits;
- Spouse/parent name: photos from a family reunion, along with first and last names. In another home, full names of spouses shown through wedding photos and announcements;
- Work history: work ID badges, certificates, and loose checks across 17 homes;
- Important numbers: a letter requesting a resident for health insurance information, along with their healthcare account number and full name. Also, two credit card numbers in another home, along with the cards' security codes and expiration dates, and a fully legible and labeled password on a sticky note in another home.

4.3.2 Identity Observations

The most common Identity information shared was first name, last name, and resident photos, with 84%, 61%, and 77% of analyzed homes showing evidence for these attributes respectively. These 3 attributes are important given that they may allow potential adversaries to more easily map an attack to a resident. Religion was displayed in 50% of the homes, all of which were Christian. Though Li et al. suggest that people deem sharing medical information publicly online through photos as highly sensitive and are unlikely to do so [36], 9% of analyzed 3D tours reveal this information. A subset of identity information seen in the dataset is listed below:

- Religion: Christian crosses, images of Mary and Jesus, and bible quotes across 22 homes;
- Race: membership of a resident to an African-American religious brotherhood;
- Gender: "Mr. and Mrs." signs in 4 homes, a "Father's Day" photo frame in a different home, and a sign labeled as "Man Cave Rules" (listing rules such as "no wine spritzers" and "men control the remote") in another home;
- Ability: wheelchairs indicating an injury or physical disability in 3 homes, multiple magnifying glasses along with a poster of a school for visually-impaired people in another home;
- Age: birth years displayed on walls in 2 homes;
- Income: evidence for use of free YMCA grocery services, indicating low income in 1 home;
- Political affiliations: a painting of Hillary Clinton in a living room in 1 home, and a sign for Trump 2020 in a high school-aged child's bedroom in another home. In a third home, a banner for a song over the bed, with a confederate flag behind an image of a man, captioned with lyrics, "If the South had won, we would've had it made";
- Other demographic information: evidence of bilingualism in 4 homes;
- Medication: a labeled anxiety and seizure medication, as well as cholesterol medication in a different home;
- Resident photos: in 77% of dataset homes;
- Resident first names: in 84% of dataset homes;
- Resident last names: in 61% of dataset homes.

4.3.3 Behavior Observations

Li et al. has shown that people claim they are highly unlikely to share information online about friends or family (rather than information about themselves) [36]. Yet they unintentionally do so when putting their homes for sale on Zillow, given that 27% of the analyzed homes revealed names of friends or family members (non-residents), and 27% of homes displayed photos of friends or family members (non-residents). The same work also showed that people are averse to sharing information about a disorganized home, medical conditions (information which is shared in Section 4.3.2), and food or smoking. Yet again, all of these attributes are shared in various 3D tours on Zillow.

Examples of behavioral information leaked include:

- Untidy home: clutter, such as dirty dishes and overflowing boxes of items covering a table, across 4 homes;
- Hobbies: hunting guide indicating the resident is a hunter, a collection of wine bottles in another home, a drumset along with posters of a resident's band in another home;
- Product preferences: visible brands of cigarette boxes and vape pens in two different homes, full pantries of food products in other homes;
- Non-sensitive opinions: a child's project stating their favorite food and color;
- Calendar events: a paper showing attendance and date of a theater event, another home showing dentist appointments, and an invitation to a party in another home;
- Friend names: people writing and signing on a dresser, and people writing and signing whiteboards with sexually explicit messages in another home;
- Friend photos: friends' wedding announcements with photos of friends on refrigerator;
- Unlocked computer screens: a child's Discord page showing their extracurricular club associations and music subscription service, and visible unlocked computer screens in 4 additional homes.

4.4 Case Studies

We now turn to case studies of 5 real homes along with potential threat implications for each. We do not reveal personally identifiable information, such as the name of the state or actual first and last names of the homeowners.

Most of these case studies were chosen because of the extent and variety of sensitive information revealed in their 3D tours, but Home 5 (Case Study 2) was chosen because its 3D tour showed the least amount of personal information out of the analyzed homes.

We only explicitly state gender or resident relationships (like parent/child) when the homes gives us explicit evidence that these relationships and identity traits are accurate.

Associated with some of the written case studies are artist renderings, which capture an anonymized version of the setup of the rooms to aid in visualizing the location of evidence for codebook attributes in different homes. The artist renderings are also anonymized by changing personal details such as brands, names, and geographic details.

4.4.1 Case Study 1: Home 1

Home Description: Home 1, shown in Figure 6, reveals potentially sensitive political, work, and medical information. A retired couple lives here, along with another female resident and her child. All of their first names are known through room labels and home decor, and their last name is revealed from decor hanging on their porch. There are also photos of the residents.

One of the retired people, resident A, was a former head law enforcement officer of city B, revealed by a room filled with old badges and other work-related memorabilia. The couple are devout Christians, with myriad crosses hanging on the walls and religious slogans.

A third resident, resident C, who is the mother of resident D, has a plastic bag labeled with the name of a medication used to treat seizures, panic disorders, or anxiety.

Threat Example: Doxxing. Consider an adversary who dislikes city B's law enforcement practices, or has a specific issue with the former officer, or neighbors who know they are moving and look online at the Zillow information of the home. These potential adversaries might dox through tactics like publicizing C's medication use, which could negatively affect C's job prospects or be a social stigma.

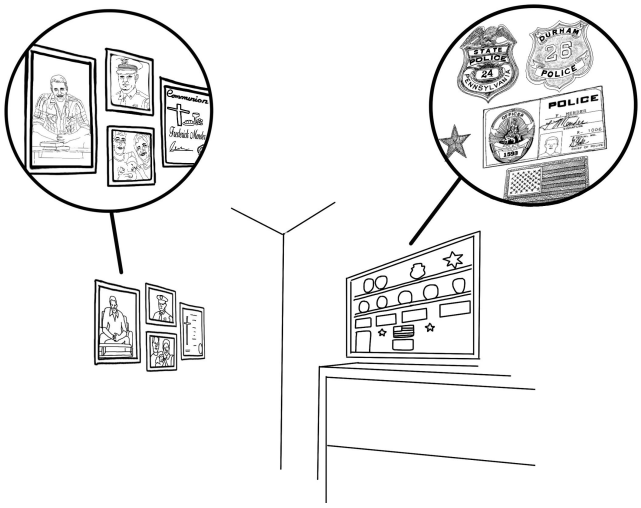

Figure 6: An anonymized rendering of an image visible in a 3D virtual tour of Home 1. In this corner of the home are old law enforcement badges, work ID cards, family photos, and display of Christian religion. Art by Akira Ohiso.

4.4.2 Case Study 2: Home 5

Home Description. Of all the homes analyzed, Home 5 has the *least* sensitive information when it comes to the percentage of codebook attributes revealed in the 3D tour (10%). Resident E appears to live with three children according to the rooms and refrigerator photo. E's room is untidy and cluttered, and contains a partially used wine bottle on the dresser. On the dresser in E's room and a child's room are many labeled makeup and toiletry products.

Threat Example: Targeted Advertising/Phishing. There is no evidence of legible names in the home, so adversaries would need to identify first the residents, or already know the residents. They could then message a resident via social media or another contact, and either advertise a legitimate product or use the product preferences to create a targeted phishing scheme, such as falsely advertising a discount on a beauty product that was seen in the home. Adversaries may also target via paper advertisements; even though E might have changed addresses, the mail may be forwarded to the residents' new address.

4.4.3 Case Study 3: Home 6

Home Description. Of all the homes analyzed, Home 6, shown in Figure 7, has some of the *most* sensitive information when it comes to the percentage of codebook attributes revealed in the 3D tour (43%).

Resident F and G live together in Home 6. Given the many Christian crosses and Christian imagery on the walls, we infer that they are Christian. We also infer that they are from country X in Central America, given that they have country X's flag hanging up on a wall, as well as books on learning English as a second language. Shown through a framed news clipping on the table, they owned a restaurant together and the name of that restaurant is also visible.

Either F, G, or both enjoy hunting, as there is a hunting manual on the table. At least one of them is a smoker; there is a lighter and a box of cigarettes in the home.

In the bedroom, a sign on the wall hangs up stating, "Beware, I'm bipolar".

The garage is also available in the 3D tour, and the exact models of the two cars F and G own are visible, as well as full license plate numbers.

Threat Example: Phishing. Given the amount of information about hobbies, as well as exact car model, make, and license plates, an adversary could leverage this information to gain further sensitive information through a phishing attack.

Threat Example: Insurance Companies. Life insurance providers may see that one of the residents smokes and thus charge them a higher premium for an insurance policy.

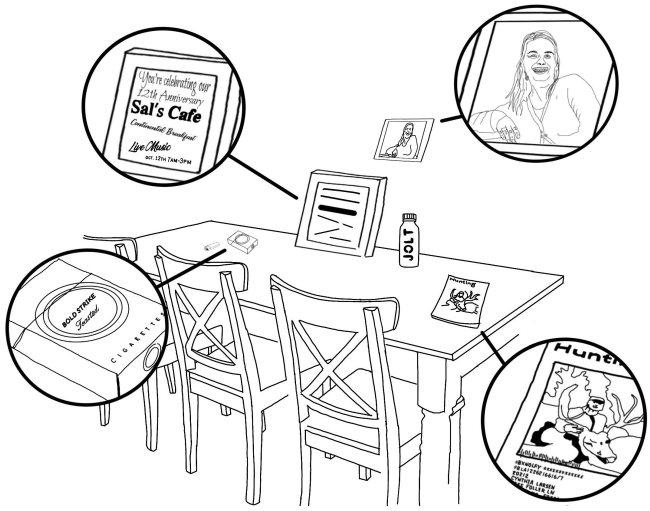

Figure 7: An anonymized rendering of an image visible in a 3D virtual tour of Home 6. In this corner, residents' employment history, the name of their restaurant, their interest in hunting, a portrait, an energy drink, and cigarettes are visible. Art by Akira Ohiso.

4.4.4 Case Study 4: Home 30

Home Description: Home 30, shown in Figure 8, reveals a password, full names, photos, and evidence of possible underage drinking. In one room, the full first and last name of a recent student, resident H, are visible through a high school ID badge and a certificate on the wall. In H's room, there is a desktop computer. On the desktop screen is attached a sticky note with the word "Password: ", followed by a fully legible password written.

H left high school with a graduation equivalency, shown by a certificate framed on the wall with a full name and date. Resident H received this equivalency certificate in 2021; since 25% of people who receive an equivalency certificate are of high school age, there is a reasonable chance that H is under 21 years old upon collecting this data [9]. Above H's desk is an empty bottle of alcohol, indicating that they are possibly engaging in underage drinking.

In the living room are displayed suggestive paintings clearly created by pressing H's body parts on the canvas with paint (it was clearly resident H because the paintings are signed). In addition to body parts, one section of the painting includes a Playboy bunny logo.

Threat Example: Privileged Access. We cannot be certain what the password on the sticky note is for, but it could be for the password to the desktop itself. This poses a security risk from anyone entering the home, such as real estate agents or the person who captured the 3D tour. Further, there is a risk that H reuses this password for multiple accounts, such as for a social media and email account.

Threat Example: Phishing on Social Media. Given the

amount of data collected in this home, including first name, last name, and former high school, it would be plausible for an adversary to find H on social media. A phisher could then message H with personalized information based on additional information visible within the home and on social media.

Threat Example: Doxxing. Suppose an adversary dislikes or wants to harm the reputation of H. With access to the Zillow 3D tour, the adversary could do so. There is evidence in H's room of possible underage drinking. Some may view the painting as inappropriate, and underage drinking, while common in American society, is also a punishable crime.

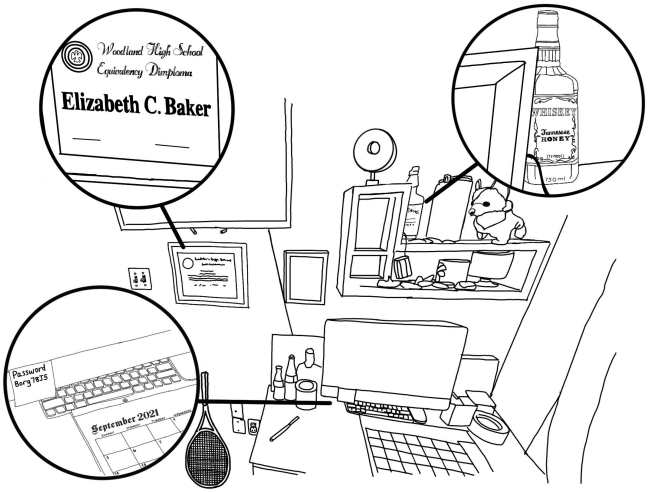

Figure 8: An anonymized rendering of an image visible in a 3D virtual tour of Home 30. In this corner is revealed password, former school and education status, and empty alcohol bottles. Art by Akira Ohiso.

4.4.5 Case Study 5: Home 38

Home Description: Home 38, shown in Figure 9, reveals work history, hobbies, and even full credit card information. Resident I is a retired employee of a large airline company, as evidenced by certificates and a signed picture frame from coworkers. I has impaired eyesight; there are glasses, the unlocked desktop computer screen is magnified, and many texts have large print.

On a laminated card in front of the desktop, there are 6 sets of numbers, clearly numbers for two separate credit cards, along with the credit cards' three-digit CVV's and expiration dates.

From framed pictures and certificates near the desk, I's full first and last name are visible.

I has many framed photos of pet dogs, and shows evidence of other hobbies like model boats, model cars, and reading.

Threat Example: Credit Fraud. The most apparent issue in this home is that an attacker would have access to full credit card information, including the full name (visible on the employment certificate) and address (available through the Zillow interface) of the credit card owner.

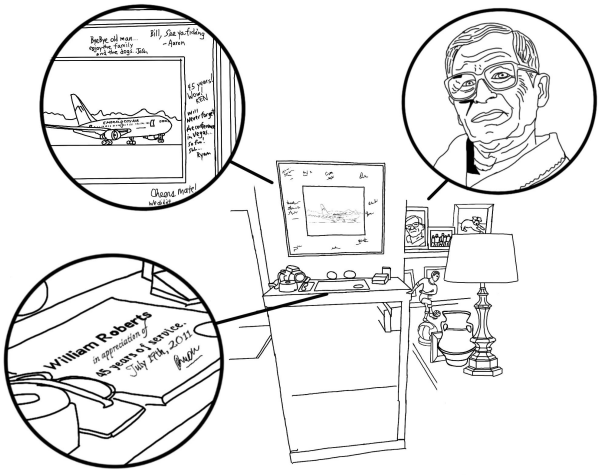

Figure 9: An anonymized rendering of an image visible in a 3D virtual tour of Home 38. The corner at which we can see information about employment, photographs, and a signed frame. Art by Akira Ohiso.

4.4.6 Case Study Observations

The percentage of codebook attributes is not the only indicator of the extent of sensitive information in a home. While Home 6 revealed the most types of sensitive information, arguably Home 30 and Home 38 leaked the most private information, given that they leaked passwords and credit card numbers, respectively. And though Home 5 shared the least amount of personal information according to the codebook attributes, potential threats still exist.

While Home 1 has more social-related threats, such as potential doxxing, others are more related to potential identity fraud and loss of money, through credit card numbers or through social engineering tactics. There is no one theme for the personal information leaked from 3D tours, but rather it varies for each individual home.

4.5 Results Summary

Just as there is no specific pattern of amount and type of personal information leaked in 3D tours, there is also no single portion of the home that was guaranteed to contain the most personal information. However, we found that text revealed the most information, versus art or other objects. For example, medication labels, work history, school information, passwords, and credit card numbers were all text-based. If an adversary were to search for information, areas of homes that typically have more text, such as desks, fridges, and walls, would be plausible first places to look.

5 Synthesis of Potential Adversaries

Based on the information observed throughout our analysis, we derived a list of potential adversaries that would benefit from the personal and private information exposed on 3D tours.

Doxxers. Suppose an adversary dislikes the resident whose 3D tour is uploaded. They may wish harm upon their reputation. In this case, they may search through 3D tours to find information that was intended to be private. Shown by the news stories in Section 2.1, it is not out of the question that famous residents may be identified and their personal information outed. This adversary may search for information that could be socially damaging, such as their relationships with others, evidence for impending divorce, or medical conditions (found in the Case Study of Home 1 in Section 4.4). If the home belongs to a public figure that the adversary knows of but does not know personally, they may still wish to publicize information for monetary or political profit.

Current and future employers. Recruiting reports show that many recruiters dislike seeing potential employees who include in social media posts references to marijuana (40% recruiters deem this a "turn-off"), alcohol consumption (39%), and politics (30%) [28]. All of this post content can be seen in 3D tours of homes in Section 4.3. Since 3D tours are online and public, it would be unsurprising if recruiters would also be opposed to seeing this content in 3D virtual home tours. Future and current employers may unearth new information in 3D tours that may lead to job repercussions, such as finding hate speech in the home.

Insurance Companies. Insurance companies already collect a plethora of data to determine individual policy pricing, including school grades [30]. For example, a life insurer may see cigarettes in a 3D tour, which would likely raise the insurance cost. Or they may see an untidy home or evidence of hoarding, which could be a fire hazard and could impact whether someone could get a home insurance policy.

Another similar potential adversary are creditors, who may incorporate information leaked from 3D homes to adjust their loan rates.

Advertisers. Advertisers may benefit from analyzing the products, preferences, opinions, and habits of residents observed in 3D tours to create hyper-targeted ads, to the point that those being targeted might find the advertisements "creepy" [64]. As shown in the worries about the collection of indoor information through Roomba vacuums [4, 52], others also deem hyper-targeting a potential threat when companies have access to this data.

Data brokers. Data brokers, who keep billions of data points across many individuals [3], may benefit from collecting even more information from public 3D tours. Once data is collected by brokers, it is difficult to alter or remove [3], which would counteract the measures of Zillow and other online real estate companies to remove people's home data once the home is sold. Though Zillow prohibits web crawlers in its Terms of Use as of 2022 [72], this may not be enough to stop data collection. While we have not investigated the business relationships between Zillow, Matterport, and other companies, we note that some companies might presently or in the future sell access to 3D tour images to data brokers.

Phishing/social engineers. In the same way that advertisers may analyze preferences, opinions, and habits to try to manipulate someone into purchasing a product, social engineers could use the information and relationships unveiled through the 3D tours to lead residents to disclose more sensitive information, like passwords or credit card information.

Identity thieves. Identity thieves may use a combination of name and address information, credit card information, passwords, and answers to common security questions revealed in 3D tours, such as pet name or a parent's former last name.

6 Discussion

6.1 Equity

There are equity issues regarding who can hide personal belongings from their home. Some residents may be in a rush to sell their home due to uncontrollable circumstances. Others cannot remove all of their belongings prior to selling because they cannot afford a storage unit, and thus may not be able to hide all of their personal information, even if they try. Our findings here contribute to the long-standing discussion of the digital divide and the greater privacy afforded the wealthy; some more recent examples of this discussion include [27], [38], and [53].

There may be additional obstacles in the property owner/tenant relationship when renters are involved. For example, does the property owner warn tenants when taking images and creating 3D tours of their space, and do the tenants know the 3D tour is publicly available? Does the property owner wait for residents' consent? In some states, publication of "private facts" is considered illegal, and information in Zillow may fall under this category if a resident does not consent to the 3D tour being taken [50]. Yet despite possible legal protections, those with less legal knowledge may be especially vulnerable to violations of privacy through property owner/tenant relationships. How many times does the property owner reuse old tours, thus leaking information about a specific group of former residents for an extended period of

time? We found an example of 3D tour reuse during analysis; from fall 2021 to spring 2022, we saw one house for rent go on the market three times, each time with the same 3D tour.

6.2 Responsibility from All Stakeholders

Whose responsibility is it to avoid sharing personal information, both from a legal perspective and an ethical perspective? On one hand, real estate agents and others who capture 3D tours should advise residents to put away belongings that could reveal sensitive information. But if residents still fail to put belongings away, as was the case in [31], they may not legally be able to blame real estate agents.

Incentives would differ for an agent uploading the tour and a homeowner, as the homeowner has stronger incentive against not sharing their personal information in the tours. If there was a written warning describing the legal implications of publishing private information, agents and other non-tenants might take the issue more seriously. Unfortunately, there currently is confusion on whether it is the real estate agent's or the property owner's responsibility, as discussed in [31].

From an ethical perspective, we believe avoiding personal information leakage is a responsibility of all stakeholders: residents, property owners, real estate agents, the MLS, and online real estate sites. Whoever is taking the 3D tour should clearly suggest hiding personal information and give the residents sufficient time to put belongings away. Of course, ethical responsibility is different than enforced actions, so simply delegating responsibility is not a catch-all for protecting residents from leaking personal information.

6.3 Weighing Possible Solutions

Preventing personal information in homes from being publicly viewable involves a complex set of factors: user experience on 3D tours should be as good as possible, residents need to understand what information *is* sensitive, understand the implications of revealing this information, and be able to afford removing/hiding of this information before a home is put online for sale. The issue complicates even more when the resident is not the property owner, since the property owner does not have as strong of an incentive to hide this information. We discuss potential ways to address personal information leakage in homes.

Public Awareness and Legislation. When developing the codebook, it became clear that there are no government-standardized guidelines for U.S. users regarding what to share online. In our searches, we could not find specific examples of what not to share on the internet on any federal website. Despite vague warnings to not share private information on the internet, we are not told the array of information that constitutes private or sensitive information. We should make more widely available checklists on government websites for what to check for in photographs and 3D tours, which can also serve as checklists that real estate websites and 3D scanning companies provide to users before collecting information. As noted in Section 4.1.1, perceptions of sensitivity of information can vary by context and audience, so we must take this into account when constructing such checklists.

Combined with checklists to increase awareness, legislative changes could incentivize the company stakeholders involved to limit information leakage on their websites. One suggestion is that for any publicly available 3D tour on a real estate or other website, it becomes illegal to 3D scan an interior without explicit description of risks, informed consent, and sufficient time between the notification of scanning and the scan taking place. There could also be guidelines from the National Institute of Standards and Technology (NIST) on what may be captured in public or publicly available 3D scans, or guidance from the Federal Trade Commission (FTC) around information leakage in the private real estate industry.

Automatically Checking for Information. Currently, the process for blurring private information in Matterport 3D tours is very manual. Manual blurring requires the user to know the types of information that are sensitive and should be blurred, then find all instances of this in the 3D tour. The automatic Matterport face-blurring feature is still in beta and not fully accurate as of writing this paper [39]; see Figure 5 for an artistic rendering of a blurring failure.

Ideally, there would be an accurate scanner that blurs all personal information. Work over the past decade has attempted automatic blurring for outdoor spaces [16, 17, 45, 62] as well as automatically determining whether a photo should be kept private or shared on social media [47, 61]. It is technologically difficult to be correct 100% of the time, and any false negative that is allowed on the site can be a problem.

Even when relying on a 100% correct blurring algorithm, blurring information can also negatively impact user experience [65], potentially affecting the property owner's ability to benefit from showing the 3D tour at all on a real estate website.

Updating Content. No solution will catch all instances of personal information leakage. If such a mistake occurs, then residents should have sufficient capabilities to request and receive quick removal of 3D tours, as well as photos, on online real estate websites. This means that either online real estate websites or the MLS that the websites pull from should allow residents to remove specific portions of their data from these sources, and in a timely fashion. Zillow and other sites that pull from the MLS database should also update their data according to the change in the MLS home content. Currently, once data goes up on the MLS, it is the property of the MLS and can be difficult for a resident to remove [57]. If changing data on the MLS is achieved, it currently can take between 15

minutes up to weeks to then propogate the MLS changes to the online real estate websites [57].

7 Future Work

Our exploratory study of information leakage in private real estate sets the foundation for multiple avenues of further research.

Private Real Estate and Marginalized Populations In addition to the general threat model of a remote adversary in Section 3.5 and 5, we should consider different threat models for different marginalized populations, and how each population's unique perspective and challenges affects the sensitivity of different types of information leaked in 3D home tours. One such population to study in the context of private real estate information leakage is marginalized racial identities. Prior work has shown the existence of racial inequity at all stages of the home buying process, such as [33] and [26]. Given this example, it may be useful to incorporate appraisers, real estate agents, home buyers, and others involved in the home buying/selling process into the threat model.

Another marginalized population to study is victims of intimate partner violence (IPV). In the threat model for victims of IPV, the abuser would know who the resident is in the 3D tour, and the publicly available 3D tour could be misused as a form of spyware. Previous work highlights the prevalence of abusers using technology to monitor victims [10, 67] and the importance of privacy and computer security support for victims of IPV [25, 73]. Future work should study how the practices of tenants, property owners, 3D tour companies, and public real estate companies support or do not support victims of IPV.

Yet another example could arise in any situation in which the occupant of a house is involved in legal proceedings and where the other party in the legal proceeding might seek to use information about a home's state to their advantage. Legal proceedings could range from probation hearings to child support hearings, though we are not experts on these topics and believe that subsequent research could explore these risks further.

Building Solutions We touched on some possible solutions to 3D tour information leakage in Section 6.3. Using the codebook we developed, we lay the groundwork for what types of information leakages an automated scanner could search for, and even automatically blur, in 3D tours and static photos.

Preventative solutions that we also discussed in Section 6.3, such as how to best educate people on what information to hide in photos, should also be further studied.

Other Vulnerable Indoor Spaces and Datasets The Zillow dataset is not the only online real estate 3D tour collection that leaks personal information. Different websites popular in other countries outside of the U.S. are also vulnerable [31]. And websites more focused on long-term rentals rather than homes for sale may suffer as well.

We find that in addition to online real estate sites, public research datasets may also leak similar types and amounts of information. For example, Apple's ARKitScenes dataset, which is intended for computer vision scene understanding and app developers, contains many artifact-visible scans of indoor spaces [5]. Apple's dataset is perhaps even more alarming than Zillow's because the Apple dataset is *intentionally* noisy; personal belongings are intentionally left out during data collection. Through preliminary searches through the dataset, we can confirm that at least some of the dataset contains personal artifacts. The Apple dataset is available indefinitely to the public, whereas the benefits of showing 3D tours in online real estate only last until the home is sold. But unlike virtual real estate 3D tours, the homes are not geo-tagged, which helps lower the risk to residents of images in the Apple Dataset.

On top of indoor space datasets that are made public, there may be other datasets that are not currently public, but could be a security and privacy concern if in the wrong hands. Data collected with the Roomba [4, 52] and Amazon's (indoor and outdoor) Ring [6, 22, 43, 54] (discussed previously in Section 2.1) are such potentially sensitive datasets.

We should further explore how to protect information stored and leaked in public and private datasets, and future legislation that aims to prevent misuse of such indoor space information by IoT companies and others.

8 Conclusion

Through an analysis of 44 homes for sale in Zillow, we find that many Zillow 3D home tours leak information that violates internet-sharing guidelines or otherwise reveals personal and private information (Section 4). News articles emerging concurrent to our study demonstrate examples of public concern and even use of information found in 3D tours other than Zillow (Section 2). Our finding of artifact-visible Matterport 3D tours on the first two Zillow results pages in most U.S. states, and our analysis of those homes, reveals that the disclosure of personal and private information is not an isolated incident but a systemic issue. Given the only increasing popularity in 3D tours and increasing affordability of 3D tours, this problem will only worsen unless intentional steps are taken to avoid leaking personal information in 3D tours (see Section 6.3 for a discussion of possible solutions). Moreover, any party who captures camera information in the home, for public research datasets or customer-facing websites or for internal purposes, should consider the risks of collecting information about objects in homes (Section 7). We believe

that our methodical study of information leakage via Zillow's Matterport 3D tours can have a role in fostering continued research and industry efforts to minimize information leakage through 3D home tours and other photo data collection within a home.

Acknowledgements

We thank Yasemin Acar, Joseph Calandrino, Ryan Calo, Pardis Emami-Naeini, Gennie Gebhart, Wulf Loh, Justin Quimby, Franziska Roesner, and Eric Zeng for insightful comments and feedback on aspects of this work. This work was supported in part by the U.S. National Science Foundation (Award 1565252), the University of Washington Tech Policy Lab (which receives support from the William and Flora Hewlett Foundation, the John D. and Catherine T. MacArthur Foundation, Microsoft, and the Pierre and Pamela Omidyar Fund at the Silicon Valley Community Foundation), and gifts from Google, Meta, Qualcomm, and Woven Planet.

References

[1] 3d scanning for insurance and restoration. https://matterport.com/industries/insurance-restoration. Accessed on 2022-05-23.

[2] Taslima Akter, Bryan Dosono, Tousif Ahmed, Apu Kapadia, and Bryan Semaan. "I am uncomfortable sharing what I can't see": Privacy concerns of the visually impaired with camera based assistive applications. In *29th USENIX Security Symposium (USENIX Security 20)*, pages 1929–1948, 2020.

[3] Gary Anthes. Data brokers are watching you, 2014.

[4] Maggie Astor. Your roomba may be mapping your home, collecting data that could be shared. The New York Times, July 2017. https://www.nytimes.com/2017/07/25/technology/roomba-irobot-data-privacy.html.

[5] Gilad Baruch, Zhuoyuan Chen, Afshin Dehghan, Yuri Feigin, Peter Fu, Thomas Gebauer, Daniel Kurz, Tal Dimry, Brandon Joffe, Arik Schwartz, et al. ARKitScenes: A diverse real-world dataset for 3d indoor scene understanding using mobile RGB-D data. In *Thirty-fifth Conference on Neural Information Processing Systems Datasets and Benchmarks Track (Round 1)*, 2021.

[6] Bill Budington. Ring Doorbell app packed with third-party trackers. Realtor.com, January 2020. https://www.eff.org/deeplinks/2020/01/ring-doorbell-app-packed-third-party-trackers.

[7] Daniel J Butler, Justin Huang, Franziska Roesner, and Maya Cakmak. The privacy-utility tradeoff for remotely teleoperated robots. In *Proceedings of the tenth annual ACM/IEEE international conference on human-robot interaction*, pages 27–34, 2015.

[8] CanLII. Juhasz v hymas, 2016 onsc 1650. https://canlii.ca/t/gnpl6, March 2016. Accessed: 2022-05-26.

[9] Julie Chaer. You are never too old to get a high school degree. Citizenshighschool.com. https://citizenshighschool.com/blog/you-are-never-too-old-to-get-a-high-school-degree/.

[10] Rahul Chatterjee, Periwinkle Doerfler, Hadas Orgad, Sam Havron, Jackeline Palmer, Diana Freed, Karen Levy, Nicola Dell, Damon McCoy, and Thomas Ristenpart. The spyware used in intimate partner violence. In *2018 IEEE Symposium on Security and Privacy (SP)*, pages 441–458. IEEE, 2018.

[11] Eun Kyoung Choe, Sunny Consolvo, Jaeyeon Jung, Beverly Harrison, and Julie A Kientz. Living in a glass house: a survey of private moments in the home. In *Proceedings of the 13th international conference on Ubiquitous computing*, pages 41–44, 2011.

[12] Tamara Denning, Cynthia Matuszek, Karl Koscher, Joshua R Smith, and Tadayoshi Kohno. A spotlight on security and privacy risks with future household robots: attacks and lessons. In *Proceedings of the 11th international conference on Ubiquitous computing*, pages 105–114, 2009.

[13] Statista Research Department. Most popular real estate websites in the United States as of October 2021, based on unique monthly visits. Statista.com, October 2021. https://www.statista.com/statistics/381468/most-popular-real-estate-websites-by-monthly-visits-usa/.

[14] Pardis Emami-Naeini, Tiona Francisco, Tadayoshi Kohno, and Franziska Roesner. Understanding privacy attitudes and concerns towards remote communications during the COVID-19 pandemic. In *Seventeenth Symposium on Usable Privacy and Security (SOUPS 2021)*, pages 695–714, 2021.

[15] Francisco Erivaldo Fernandes, Guanci Yang, Ha Manh Do, and Weihua Sheng. Detection of privacy-sensitive situations for social robots in smart homes. In *2016 IEEE International Conference on Automation Science and Engineering (CASE)*, pages 727–732. IEEE, 2016.

[16] Arturo Flores and Serge Belongie. Removing pedestrians from Google Street View images. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition-Workshops*, pages 53–58. IEEE, 2010.

[17] Andrea Frome, German Cheung, Ahmad Abdulkader, Marco Zennaro, Bo Wu, Alessandro Bissacco, Hartwig Adam, Hartmut Neven, and Luc Vincent. Large-scale privacy protection in Google Street View. In *2009 IEEE*

12th international conference on computer vision, pages 2373–2380. IEEE, 2009.

[18] Christine Geeng, Mike Harris, Elissa Redmiles, and Franziska Roesner. "Like lesbians walking the perimeter": Experiences of US LGBTQ+ folks with online security, safety, and privacy advice. In *31st USENIX Security Symposium (USENIX Security 22)*, pages 305–322, 2022.

[19] Ilyce Glink and Samuel J. Tamkin. Do you have the right to have photos of your home removed from realty sites after the sale? WashingtonPost.com, April 2019. https://www.washingtonpost.com/business/2019/04/01/do-you-have-right-have-photos-your-home-removed-realty-sites-after-sale/.

[20] Google. Blur or remove 360 photos with the Street View app. Google.com. https://support.google.com/maps/answer/7011973.

[21] Google. Google-contributed Street View imagery policy. Google.com. https://www.google.com/streetview/policy/.

[22] Matthew Guariglia. Amazon's Ring is a perfect storm of privacy threats. Realtor.com, August 2019. https://www.eff.org/deeplinks/2019/08/amazons-ring-perfect-storm-privacy-threats.

[23] Neilly H. Tan, Richmond Y. Wong, Audrey Desjardins, Sean A. Munson, and James Pierce. Monitoring pets, deterring intruders, and casually spying on neighbors: Everyday uses of smart home cameras. In *CHI Conference on Human Factors in Computing Systems*, pages 1–25, 2022.

[24] Rakibul Hasan, Patrick Shaffer, David Crandall, Eman T Apu Kapadia, et al. Cartooning for enhanced privacy in lifelogging and streaming videos. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pages 29–38, 2017.

[25] Sam Havron, Diana Freed, Rahul Chatterjee, Damon McCoy, Nicola Dell, and Thomas Ristenpart. Clinical computer security for victims of intimate partner violence. In *28th USENIX Security Symposium (USENIX Security 19)*, pages 105–122, 2019.

[26] Junia Howell and Elizabeth Korver-Glenn. Neighborhoods, race, and the twenty-first-century housing appraisal industry. *Sociology of Race and Ethnicity*, 4(4):473–490, 2018.

[27] Joseph W Jerome. Buying and selling privacy: Big data's difference burdens and benefits. *Stan. L. Rev. Online*, 66:47, 2013.

[28] Jobvite. 2021 recruiter nation report. Technical report, Jobvite, September 2021.

[29] Sowmya Karunakaran, Kurt Thomas, Elie Bursztein, and Oxana Comanescu. Data breaches: User comprehension, expectations, and concerns with handling exposed data. In *Fourteenth Symposium on Usable Privacy and Security (SOUPS 2018)*, pages 217–234, 2018.

[30] Leslie Kasperowicz. Do auto insurance companies check grades? AutoInsurance.org, July 2021. https://www.autoinsurance.org/do-auto-insurance-companies-verify-grades/.

[31] Zoe Kleinman. Estate agent's hi-tech house tour exposes personal data. BBC.com, April 2021. https://www.bbc.com/news/technology-56718046.

[32] Bert-Jaap Koops. Privacy spaces. *W. Va. L. Rev.*, 121:611, 2018.

[33] Elizabeth Korver-Glenn. Compounding inequalities: How racial stereotypes and discrimination accumulate across the stages of housing exchange. *American Sociological Review*, 83(4):627–656, 2018.

[34] Lingxiao Li and Abdullah Yavas. The impact of a multiple listing service. *Real Estate Economics*, 43(2):471–506, 2015.

[35] Yifang Li, Wyatt Troutman, Bart P Knijnenburg, and Kelly Caine. Human perceptions of sensitive content in photos. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 1590–1596, 2018.

[36] Yifang Li, Nishant Vishwamitra, Hongxin Hu, and Kelly Caine. Towards a taxonomy of content sensitivity and sharing preferences for photos. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–14, 2020. https://dl.acm.org/doi/pdf/10.1145/3313831.3376498.

[37] Naomi Little. Matterport accelerates past five million spaces under management. Matterport.com, May 2021. https://matterport.com/news/matterport-accelerates-past-five-million-spaces-under-management.

[38] Mary Madden, Michele Gilman, Karen Levy, and Alice Marwick. Privacy, poverty, and big data: A matrix of vulnerabilities for poor Americans. *Wash. UL Rev.*, 95:53, 2017.

[39] Matterport. How to blur faces automatically in Matterport spaces. Matterport.com, 2022. https://support.matterport.com/s/article/How-to-Blur-Faces-in-Matterport-Spaces.

[40] Peter Mayer, Yixin Zou, Florian Schaub, and Adam J Aviv. "Now I'm a bit angry:" Individuals' awareness, perception, and responses to data breaches that affected them. In *30th USENIX Security Symposium (USENIX Security 21)*, pages 393–410, 2021.

[41] Linda McNair. With 3D Tours, properties sell up to 31% faster and at a higher price. Matterport.com, February 2020. https://matterport.com/blog/3d-tours-properties-sell-31-faster-and-higher-price.

[42] George R Milne, George Pettinico, Fatima M Hajjat, and Ereni Markos. Information sensitivity typology:

Mapping the degree and type of risk consumers perceive in personal data sharing. *Journal of Consumer Affairs*, 51(1):133–161, 2017.

[43] Mozilla. Amazon Ring Always Home Cam. Mozilla.org, November 2021. https://foundation.mozilla.org/en/privacynotincluded/amazon-ring-always-home-cam/.

[44] Carman Neustaedter, Saul Greenberg, and Michael Boyle. Blur filtration fails to preserve privacy for home-based video conferencing. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 13(1):1–36, 2006.

[45] Angelo Nodari, Marco Vanetti, and Ignazio Gallo. Digital privacy: Replacing pedestrians from Google Street View images. In *Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012)*, pages 2889–2893. IEEE, 2012.

[46] Washington State Office of the Attorney General. Internet safety: Understanding the risks, 2008. https://www.atg.wa.gov/internet-safety.

[47] Tribhuvanesh Orekondy, Bernt Schiele, and Mario Fritz. Towards a visual privacy advisor: Understanding and predicting privacy risks in images. In *Proceedings of the IEEE international conference on computer vision*, pages 3686–3695, 2017.

[48] James Pierce. Smart home security cameras and shifting lines of creepiness: A design-led inquiry. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–14, 2019.

[49] RJ Pittman. We are changing face blurring: Here is why. Matterport.com, July 2020. https://matterport.com/blog/we-are-changing-face-blurring-here-why.

[50] Digital Media Law Project. Publication of private facts. DMLP.org. https://www.dmlp.org/legal-guide/publication-private-facts.

[51] Hirak Ray, Flynn Wolf, Ravi Kuber, and Adam J Aviv. Why older adults (don't) use password managers. In *30th USENIX Security Symposium (USENIX Security 21)*, pages 73–90, 2021.

[52] Michael Reilly. Your Roomba is also gathering data about the layout of your home. MIT Technology Review, July 2017. https://www.technologyreview.com/2017/07/25/150346/your-roomba-is-also-gathering-data-about-the-layout-of-your-home/.

[53] Neil Richards and Woodrow Hartzog. Taking trust seriously in privacy law. *Stan. Tech. L. Rev.*, 19:431, 2015.

[54] Ring. Ring terms of service. Ring. https://ring.com/terms.

[55] Ring. Ring website. Ring. https://ring.com/amazon-sidewalk.

[56] Franziska Roesner, David Molnar, Alexander Moshchuk, Tadayoshi Kohno, and Helen J Wang. World-driven access control for continuous sensing. In *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*, pages 1169–1181, 2014.

[57] Jeanne Sager. How to get your home's real estate listing removed from the internet. Realtor.com, May 2022. https://www.realtor.com/advice/sell/how-do-i-get-a-real-estate-listing-removed/.

[58] Christine Schneider. Sellers beware: Tips for protecting your home in a virtual world. LinkedIn.com, November 2021. https://www.linkedin.com/pulse/sellers-beware-tips-protecting-your-home-virtual-world-schneider.

[59] Robert Templeman, Mohammed Korayem, David J Crandall, and Apu Kapadia. PlaceAvoider: Steering first-person cameras away from sensitive spaces. In *NDSS*, pages 23–26. Citeseer, 2014.

[60] The Preservation Society of Newport County. https://matterport.com/discover/space/v4LWLiLDm3s. Accessed on 2022-05-23.

[61] Ashwini Tonge and Cornelia Caragea. Privacy prediction of images shared on social media sites using deep features. *arXiv preprint arXiv:1510.08583*, 2015.

[62] Ries Uittenbogaard, Clint Sebastian, Julien Vijverberg, Bas Boom, Dariu M Gavrila, et al. Privacy protection in street-view panoramas using depth and multi-view imagery. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 10581–10590, 2019.

[63] Blase Ur, Jaeyeon Jung, and Stuart Schechter. Intruders versus intrusiveness: teens' and parents' perspectives on home-entryway surveillance. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 129–139, 2014.

[64] Blase Ur, Pedro Giovanni Leon, Lorrie Faith Cranor, Richard Shay, and Yang Wang. Smart, useful, scary, creepy: Perceptions of online behavioral advertising. In *Symposium on Usable Privacy and Security (SOUPS)*, 2012.

[65] Nishant Vishwamitra, Bart Knijnenburg, Hongxin Hu, Yifang P Kelly Caine, et al. Blur vs. block: Investigating the effectiveness of privacy-enhancing obfuscation for images. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 39–47, 2017.

[66] Diane Walker and Florence Myrick. Grounded theory: An exploration of process and procedure. *Qualitative health research*, 16(4):547–559, 2006.

[67] Delanie Woodlock. The abuse of technology in domestic violence and stalking. *Violence against women*, 23(5):584–602, 2017.

[68] Jun Yu, Baopeng Zhang, Zhengzhong Kuang, Dan Lin, and Jianping Fan. iPrivacy: image privacy protection by identifying sensitive objects via deep multi-task learn-

ing. *IEEE Transactions on Information Forensics and Security*, 12(5):1005–1016, 2016.

[69] Cha Zhang, Yong Rui, and Li-wei He. Light weight background blurring for video conferencing applications. In *2006 International Conference on Image Processing*, pages 481–484. IEEE, 2006.

[70] Zillow. Make your listing pop with Zillow 3D Home® tours. Zillow.com. https://www.zillow.com/z/3d-home/.

[71] Zillow. Zillow search. Zillow. https://www.zillow.com/browse/homes/.

[72] Zillow. Zillow terms of service. Zillow.com. https://www.zillowgroup.com/terms-of-use/.

[73] Yixin Zou, Allison McDonald, Julia Narakornpichit, Nicola Dell, Thomas Ristenpart, Kevin Roundy, Florian Schaub, and Acar Tamersoy. The role of computer security customer support in helping survivors of intimate partner violence. In *30th USENIX Security Symposium (USENIX Security 21)*, pages 429–446, 2021.

A Codebook Categories

We elaborate below on each of the codebook categories and the attributes each category contains.

Guidance The first category of codebook attributes was derived from a list on the Washington State Office of the Attorney General website [46]. As discussed in Section 3.2, we used the Attorney General website because it was the most detailed U.S. government-provided list of information to avoid sharing on the internet.

The attributes that make up this category are car information (including license plate or vehicle make/model), full date of birth (of any resident), explicit words/photos, Social Security Numbers, other sensitive numbers (passwords or credit card numbers), names of spouse or last name before marriage, phone numbers (of residents or friends of residents), work history, and credit status.

Identity The second category is other identity information not included in the Guidance category, developed in combination from observations of 3D tours and discussions between Researcher 1 and Researcher 2.

The attributes for Identity are religion, race, gender, ability, age, income, political affiliations, familial structure (such as marital status, number of children, or multi-generational household), other demographic information (left as a placeholder for further notes, such as if the resident is bilingual), medication, photos of residents, first names of residents, and last names of residents.

Section 3.2 further explains how we code demographic information such as race, gender, and age.

Behavioral The second category is other behavioral information not included in the Guidance category, developed in combination from observations of 3D tours and discussions between Researcher 1 and Researcher 2.

Behavioral attributes are made up of the following: untidy home (marked as existing if the home is perceived by the researchers as very untidy or cluttered), hobbies, product preferences (such as toiletries or electronics), calendar events, names of friends, photos of friends, unlocked computer screens, and incriminating evidence.

As can be seen above, the line between the Behavior and the Identity category is not clear-cut, given that attributes such as religion may be both behavioral and identity-related. However, we only put an attribute in one or the other category, and not both. The three categories are meant mainly for the reader to conceptualize the type of information being displayed.

B 3D Tour Examples

Below are two additional examples from [60] that show ways a tour can be navigated, with the red arrow added by us.

Figure 10: The viewer has clicked the top center of Figure 2 to get to the current view of the tour shown here. The red arrow overlaying the screenshot represents where a viewer would zoom in towards to see the details in Figure 11.

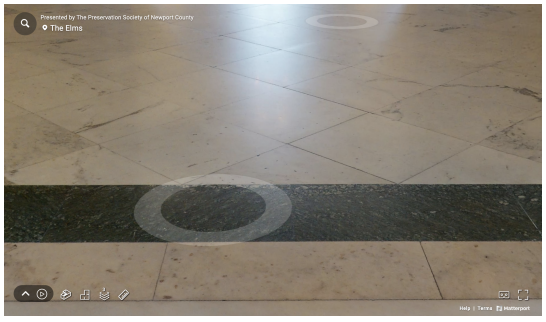

Figure 11: The viewer has zoomed in from Figure 10 to see the details in the floor tiles.