

# Sunlight or sunburn: A survey of attitudes toward online availability of US public records

Sean A. Munson<sup>a,\*</sup>, Daniel Avrahami<sup>b,c</sup>, Sunny Consolvo<sup>d</sup>, James Fogarty<sup>e</sup>,  
Batya Friedman<sup>c</sup> and Ian Smith<sup>f</sup>

<sup>a</sup>*School of Information, University of Michigan, Ann Arbor, MI, USA*

<sup>b</sup>*Intel, Santa Clara, CA, USA*

<sup>c</sup>*The Information School, University of Washington, Seattle, WA, USA*

<sup>d</sup>*Human Centered Design and Engineering, University of Washington, Seattle, WA, USA*

<sup>e</sup>*Computer Science and Engineering, University of Washington, Seattle, WA, USA*

<sup>f</sup>*Everbread Ltd., London, UK*

**Abstract.** Many have enthusiastically greeted the ability to search and view public records online as a great advance for transparency and accountability. Such ability, however, also creates value tensions with privacy and other important human values. In this paper, we report findings from a survey of 134 residents of the US Pacific Northwest on their awareness of and attitudes towards online access to political campaign records and real estate transaction histories, bringing to light some of the social implications of technological changes that increase ease of access to public records. We show that, while respondents often understood the reason behind making these records public, considerable concern about the current accessibility of these records exists, along with a precautionary indication that such open access may reduce public participation for some individuals.

**Keywords:** Public records, privacy, access, transparency, value tensions, value sensitive design, political campaign contributions, real estate records

## 1. Introduction

The availability of online searchable public records, combined with the ability to quickly and easily create mashups, has enabled the creation of new ways to freely search and view public records, including political campaign contributions, real estate transactions, contractor licenses, divorce records, jail and inmate records, and sex offender registries. Much of this information has long been a part of the public record in the United States. Although the existence of this data as part of the public record is not new, the ease with which people can access and search this information online has substantially changed how the public can interact with it and has greatly reduced barriers to access, locally and globally.

In this paper, we briefly review the history of and justifications for US public records laws, report findings from a mail survey of 134 adults in the US Pacific Northwest regarding their awareness of and

---

\*Corresponding author: Sean A. Munson, School of Information, 105 S State Street, North Quad 4370N, University of Michigan, Ann Arbor, MI, USA. E-mail: samunson@umich.edu.

attitudes toward the online accessibility of two types of public records – political campaign contributions and real estate transactions – and conclude by discussing possible ways to address the tension between privacy and transparency present in our survey results.

### 1.1. Background

In this section, we briefly review models of privacy, including the tension between privacy and transparency, and a brief history of the public records in the United States, focusing particularly on the legislative and judicial history of campaign finance records.

#### 1.1.1. Privacy and transparency

Several scholars have attempted to define privacy, arriving at a variety of definitions. This includes ideas of “the right to be let alone” [31] and “freedom from the judgment of others” [14]. Nissenbaum describes privacy as “contextual integrity” [19] – that is, privacy entails the protection of norms of a given context, both with respect to what information is collected and to how that information is disseminated. Adams and Sasse propose a model of privacy in which an individual’s privacy perspective is based primarily on the information’s sensitivity (mediated by their own judgment), but also the information’s current and future use (mediated by their assessment of the costs and benefits) and the information’s receiver (mediated by their trust in the receiver), all affected by context [1].

Whichever definition one prefers, privacy is often at odds with transparency. Transparency can benefit public health and safety, law enforcement, accountability, and capabilities for governance and problem solving [9], but this often comes at the expense of individual privacy. Recent advances in technology have made many types of information more readily searched, browsed, aggregated, and disseminated. In many situations, the common HCI solution of giving users more and better control of their information [2] defeats the purpose of transparency, as those with behavior to hide would be able to hide it. As Introna and Pouloudi note, “often there is a thin line between the need to disclose information for the benefit of some individuals and the need to safeguard the privacy of some individuals by not disclosing this information” [14].

In a rush to improve the availability of many types of public records by placing them online and creating tools that mash up different sets of records, the tension between privacy and transparency may have been overlooked. What might be the costs of this increased transparency, such as decreased individual autonomy [5], excessive pressure to conform, inhibition of daily activities [24], or increased risk of identity theft [12]?

#### 1.1.2. Public records

The 1910 Federal Corrupt Practices Act, commonly known as the Publicity Act, contained the US’s first public disclosure law for campaign finances. At first, it only required disclosure of expenditures by parties, but was revised in 1911 to require disclosure by candidates for the first time. In 1925, the act was again amended to require reporting of contributions over \$100, though the provisions for this were rather weak. Current requirements are largely defined by the 1971 Federal Election Campaign Act, revised and expanded in 1974 and in 2002. The Supreme Court upheld disclosure requirements in *Buckley v Valeo* (1976),<sup>1</sup> calling them valuable tools to reduce corruption by communicating to voters the interests

---

<sup>1</sup>*Buckley v. Valeo* challenged Federal Election Campaign Act, which limited contributions to candidates for federal office, required disclosure of political donations, required reporting of campaign expenses, established public financing for presidential elections, and limited a variety of campaign expenditures. The Supreme Court upheld limits on individual donations, the disclosure and reporting rules, and the public financing process, while striking down limits on expenditures.

to which a candidate is “most likely to be responsive”, by making large contributions and expenditures public to “deter actual corruption and avoid the appearance of corruption”, and by providing necessary data for the enforcement of campaign finance laws [29]. Several states have adopted their own laws requiring disclosure of political contributions and expenditures. Similarly, real estate transaction records have long been part of the public record as a way of recording and verifying ownership and making tax information available.

Today, many state and federal agencies publish public records online using basic, searchable databases (e.g., the FEC’s records of campaign contributions include first and last names, home address, employer, occupation, contribution amount, and recipient). This allows others to download and build upon these databases to offer tools that better support browsing information or viewing at different levels of aggregation [32]. For example, [fundrace.huffingtonpost.com](http://fundrace.huffingtonpost.com) allows users to browse political donations by location (over a map), employer, or even by connecting to their Facebook accounts and finding friends’ contributions. Before such tools, individuals’ most common interactions with this data were aggregate reports in the media, or specific searches, often requested face-to-face at county clerks’ offices. Though these tools may improve the ability to engage in real estate transactions (in the case of real estate records) or the ability to analyze and understand who supports and may be influencing a political figure, they also reveal information about individuals – including neighbors, colleagues, and friends – that these individuals may otherwise prefer not to share.

To see how increased ease of access to public records could create tension, consider the case of campaign contributions. In mixed company, people often stay quiet about their political opinions in order to promote social harmony [20], and, among online social networks, people often underestimate their political disagreements with their friends [11]. Learning about friends’ campaign contributions with just a few clicks could create tensions or disrupt the social harmony that people often attempt to create by not sharing this information. Even if campaign contribution search tools are used to enable a sort of crowdsourced political accountability (the intent usually claimed by their designers), this may still be a threat: as Palen and Dourish note in their analysis of privacy in a networked world, in 1984 it was “the culture of pervasive mutual monitoring that constitutes the threat to individuals” [21].

In the privacy model described by Adams and Sasse, privacy invasions occur when users’ assumptions about information’s sensitivity, how it will be used, or who will receive it turn out to be inaccurate [1]. In the case of public records, mashups that let people browse their social networks’, employers’, or neighborhoods’ campaign contribution or real estate transaction history are likely to violate the expectations about information use and dissemination of people who entered into these transactions with assumptions based on previous technological limitations. For example, even after records were made available online, people may have believed that accessing the data would be too tedious, and thus impractical, for a curious acquaintance to type name after name into the FEC database, but FundRace’s integration of Facebook social graph data with the FEC data eliminates this time-consuming process. Warner and Chun note the challenge that mashups present, writing that if mashups “were created for access to data on any individual citizen, there would be serious privacy implications” [30]. Such mashups now exist, and privacy advocates [10] as well as scholars [20] and the courts [23] have recognized that increasing the accessibility of previously public-yet-difficult-to-access records may sometimes be more harmful than beneficial.

Nissenbaum [19] argues, for example, that placing public records online may be a violation of contextual integrity, and, says that to address this question “requires an examination of governing norms of appropriateness and flow to see whether and in what ways the proposed new practices measure up.” Though, legally, a public record may be public, changes in how that public record is stored, distributed, and accessed may still lead to privacy violations and concerns where there previously were none.

Advocates of improving access to public records often take inspiration from US Supreme Court Justice Brandeis who in 1914 wrote, “Publicity is justly commended as a remedy for social and industrial diseases. Sunlight is said to be the best of disinfectants; electric light the most efficient policeman” [4], a passage the Supreme Court cited in its decision to uphold *Buckley v Valeo*. Much of the scholarly discussion about tools used to access public records has focused on creating “more sunlight” – how to make access easier (e.g., [20,32]) or increase requirements for more data to be disclosed and improve the format in which it is available (e.g., [3]) – while questions of how to achieve a balance between publicity and privacy when building new tools have received less attention.

Even though Warren and Brandeis excepted matters “of public or general interest” from being private in the famous “The Right to Privacy” – an article that responded to technical advances of their day – they argued for nuance and balance, noting that “no fixed formula” can appropriately cover every situation [31]. With the fixed formulas of current public records laws, privacy advocates have expressed concerns that individuals worried about potential negative consequences of others learning their political leanings might choose to give less than the reporting limit or otherwise reduce their public participation so as to preserve their privacy [10]. In fact, some members of the public have experienced negative consequences, including confrontational or harassing emails, death threats, envelopes with white powder, and boycotts of their employer as a result of their political contributions being published online [27]. Finally, aggregation of different types of public records can make it easier for attackers to gain access to or derive sensitive information (e.g., mothers’ maiden names) – used for identity verification [12].

To better understand awareness of and attitudes toward the online accessibility of public records, as well as norms about whether and how this data should be made available and accessed, we conducted a survey of households in the states of Washington, Oregon, Montana, and Idaho.

## 2. The survey

### 2.1. Survey instrument

We created a four-page survey comprised of 26 questions asking respondents about demographics, history and intentions for campaign contributions and real estate purchases, previous campaign contributions, whether they had displayed political bumper stickers or signs, their comfort about what of this information was available online, and how its availability may affect future plans. Following prior work in value sensitive design on investigating people’s views and values about privacy in public [6], we looked systematically at the granularity of information that would be publicly available and searchable online. Broader level information attributes included state, city, zip code, and neighborhood; individual level information attributes included home address, last name only, first and last name. For campaign contributions we also asked about employer and occupation as they are currently reported as part of the public record. To introduce respondents to the type of online access that is currently available, the surveys included screenshots from the Federal Election Commission (FEC) campaign contribution search tool, Huffington Post’s FundRace tool, a county records website, and Zillow.com (Fig. 1).

Twenty-two of the survey questions were quantitative, taking the form of yes/no responses (e.g., *Are you registered to vote in the U.S.?*); selecting among a small set of responses (e.g., *Which of the following statements best describes how you would complete this sentence: I believe that the details of campaign contributions should. . . (a) be publicly accessible as they are now, including open online access, (b) be publicly accessible, but their online access should be restricted, (c) be publicly accessible, but they should not be accessible online, or (d) not be publicly accessible*); and rating statements on a 5-point



Fig. 1. Sample screenshots from the survey. Clockwise from upper left: Zillow.com, King County website, FundRace search options, and FundRace neighborhood view. Some information has been hidden to protect privacy.

Likert-like scale (e.g., *how comfortable are you on a scale from 1 to 5 (1 = not at all comfortable, 5 = very comfortable) with someone being able to find those details online by . . .*). The other four questions were open-ended: two asked “*Why do you think that details of (real estate purchases/campaign contributions) are publicly accessible?*” and two asked “*How, if at all, do you think that the current online access to (real estate purchase/campaign contribution) details should be changed?*.” There were two forms of the survey; one asked questions about real estate purchase records and then about campaign contribution records; the other first asked about campaign contribution records and then real estate purchase records.

### 2.2. Recruitment

Surveys were sent via US Mail to 1,000 households on 4 and 5 November 2008 (at the time of the US Presidential election). Surveys were sent to randomly selected addresses in areas where at least one household had made a political contribution (determined with fundrace.huffingtonpost.com). Thus, contributing households as well as neighbors were sent the survey. No compensation was offered for participating.

### 2.3. Respondents

Of the 1,000 surveys sent, 47 were rejected with the mark “return to sender.” Of the remaining 953 surveys, 134 were filled out and returned to us by 16 December 2008 (an above average response rate of

Table 1  
Respondent demographics (Total  $n = 134$ )

Question	Response	#	%
Gender ( $n = 129$ )	Female	70	54%
	Male	59	46%
Age ( $n = 132$ )	18–24	3	2%
	25–34	2	2%
	35–44	19	14%
	45–54	32	24%
	55–64	40	30%
	65+	36	27%
Location ( $n = 134$ )	Washington	54	40%
	Oregon	47	35%
	Idaho	10	8%
	Montana	23	17%
Education ( $n = 129$ )	High school diploma	3	2%
	Some college	14	11%
	Certificate	2	2%
	Bachelor's degree	43	33%
	Some Masters-level work	13	10%
	Masters degree	32	25%
	Some Doctoral-level work	2	2%
	Doctoral degree	15	12%
	Other professional degree	5	4%
Political Party ( $n = 132$ )	Constitution	1	1%
	Democratic	66	50%
	Libertarian	2	2%
	Republican	35	27%
	Other – Independent	11	8%
	Other – None	11	8%
	Other	2	2%
	Prefer not to say	4	3%
Registered to vote ( $n = 134$ )	Yes	126	94%
	No	8	6%
Has voted ( $n = 134$ )	Yes	130	97%
	No	4	3%

14% for a survey offering no compensation). There was no difference in response rate between surveys that asked about campaign contributions first ( $n = 69$ ) and those that asked about real estate transactions first ( $n = 65$ ). Table 1 shows the respondents' demographics. Although the response rate is above average for a survey of this type, given no compensation, there is likely self-selection bias to people who have greater interest in or concern about this topic. For example, younger people who tend to have less money and so are less likely to own real estate or to have made political contributions over \$200, may be less concerned about public access to this sort of data. We note that the respondents were skewed to be older than the general population, and also were politically more liberal.

#### 2.4. Coding and reliability

We developed categories for the four qualitative questions (e.g., about how access to the records should change and why they are public) through inductive coding of responses. Two researchers coded all of the data. Inter-rater reliability was assessed using Cohen's kappa, a measure of the level of agreement

between two coders, with  $\kappa$  above 0.80, for all but three categories before discussion; post discussion all were above 0.94, a high level of agreement [16].

### 3. Results

We now describe qualitative and quantitative findings from respondents' answers. We first discuss respondents' understanding of why records are public and their awareness of the availability of the records in different formats online. We then examine their comfort (or discomfort) with records being searchable by various criteria and their beliefs about whether current access to records should change. We show that, while respondents often understood the reason behind making these records public, considerable concern about the current accessibility of these records exist, along with a precautionary indication that such open access may reduce public participation.

All results were examined for effects by age, gender, and political party; no differences were found.

#### 3.1. Beliefs, understanding and awareness

Respondents offered many reasons for why they thought real estate purchase and campaign contribution records are and should be publicly available. Most respondents offered at least one reason, with only 19% saying that they did not know why for real estate transactions (another 11% gave no response) and 13% saying they did not know why for campaign contributions (another 10% gave no response). Others simply said that it was the law but did not offer a reason (14% for real estate and 10% for campaign contributions).

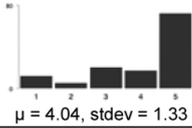
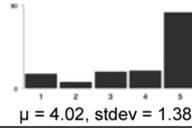
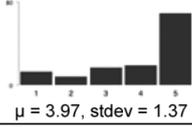
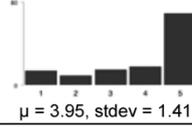
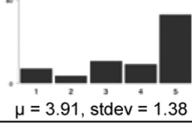
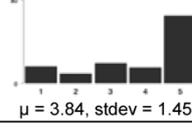
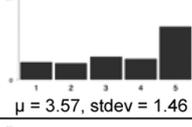
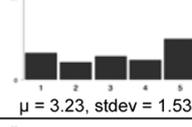
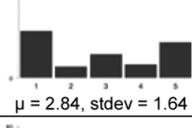
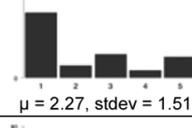
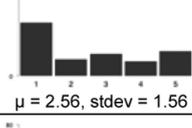
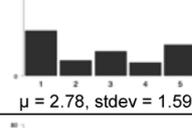
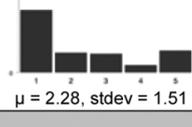
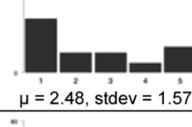
The most commonly cited reasons for *real estate records* being in the public domain fell into the categories of general value, ownership, and title information (22%; e.g., “to give the people in a neighborhood (or looking into it) a basis of who and what was paid in a neighborhood”), home buying and selling (28%; “to help educate buyers – sellers of market value”), and tax purposes (10%; “for real estate appraisers and county tax assessors”). These were followed by transparency (7%; “to see who has a financial stake in any given community”, “conflicts of interest”), general information (4%; “public information”), and law enforcement (3%; “to reduce fraud”).

In contrast, for the public disclosure of *political campaign contributions* roughly half of the respondents (49%) cited a single reason: transparency. Many respondents explained that public disclosure was necessary to prevent the buying of votes, to see who has or is trying to gain political influence, and to hold politicians “accountable” or to “keep politicians honest.” For example, one respondent explained: “Accountability to the citizens. It’s important to know what entities or interest groups may be influencing the office holder.” Law enforcement needs were the second most common (13%; “to be sure that the candidate does not keep all the monies,” “minimize illegal contributions,” “prevent fraud”). Other reasons (tax purposes, general information, or a lack of regulations preventing disclosure) accounted for less than 12% of the total.

In terms of prior awareness of online public records, roughly three-quarters (73%) of the respondents were previously aware that real estate records were available online, while slightly less than half (47%) of respondents were previously aware that political campaign contributions records were accessible online. If limited to those who had previously contributed to a campaign ( $n = 91$ ) or those who had previously purchased property ( $n = 130$ ), prior awareness stays approximately at the same percentages of 74% and 48% respectively. That is, the processes of contributing to a campaign and buying real estate do not appear to increase awareness or educate people about how records of these actions might be made available.

Table 2

Comfort with records being searchable by different criteria. 1 = not at all comfortable; 5 = very comfortable

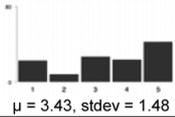
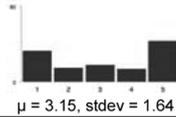
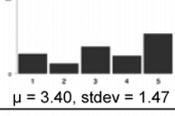
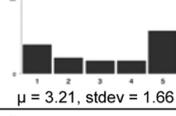
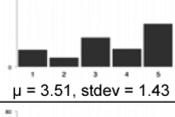
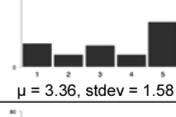
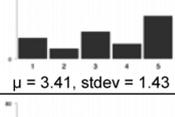
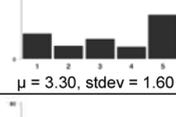
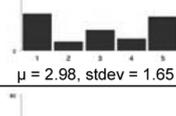
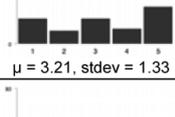
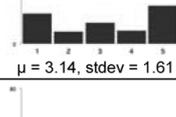
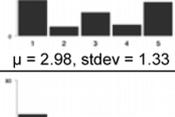
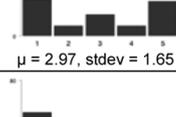
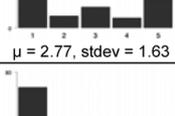
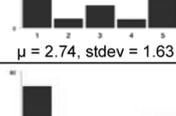
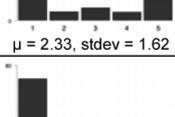
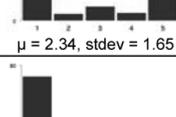
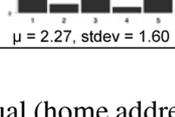
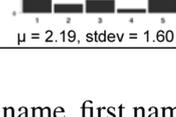
Comfort with searching by	Real estate purchases	Campaign contributions
State	 $\mu = 4.04, \text{stdev} = 1.33$	 $\mu = 4.02, \text{stdev} = 1.38$
City	 $\mu = 3.97, \text{stdev} = 1.37$	 $\mu = 3.95, \text{stdev} = 1.41$
Zip code	 $\mu = 3.91, \text{stdev} = 1.38$	 $\mu = 3.84, \text{stdev} = 1.45$
Neighborhood name	 $\mu = 3.57, \text{stdev} = 1.46$	 $\mu = 3.23, \text{stdev} = 1.53$
Home address	 $\mu = 2.84, \text{stdev} = 1.64$	 $\mu = 2.27, \text{stdev} = 1.51$
Last name only	 $\mu = 2.56, \text{stdev} = 1.56$	 $\mu = 2.78, \text{stdev} = 1.59$
First & last name	 $\mu = 2.28, \text{stdev} = 1.51$	 $\mu = 2.48, \text{stdev} = 1.57$
Employer		
Occupation		

### 3.2. Comfort with records being searchable

The survey questions also positioned us to explore respondents' comfort with public records being searched at various levels of informational granularity, in particular by attributes tied to the larger geographic area in which respondents lived (state, city, zip code, neighborhood name) and by those

Table 3

Comfort with records being searchable by different criteria. 1 = not at all comfortable; 5 = very comfortable

Comfort if the person accessing these details online lives	Real estate purchases	Campaign contributions
in the same neighborhood as you	 $\mu = 3.43, \text{stdev} = 1.48$	 $\mu = 3.15, \text{stdev} = 1.64$
in the same city as you, but not in the same neighborhood	 $\mu = 3.40, \text{stdev} = 1.47$	 $\mu = 3.21, \text{stdev} = 1.66$
in the same state as you, but not the same city	 $\mu = 3.51, \text{stdev} = 1.43$	 $\mu = 3.36, \text{stdev} = 1.58$
in the US and is a US citizen	 $\mu = 3.41, \text{stdev} = 1.43$	 $\mu = 3.30, \text{stdev} = 1.60$
outside of the US but is a US citizen	 $\mu = 3.02, \text{stdev} = 1.61$	 $\mu = 2.98, \text{stdev} = 1.65$
in the US, and is a US Permanent Resident	 $\mu = 3.21, \text{stdev} = 1.33$	 $\mu = 3.14, \text{stdev} = 1.61$
outside of the US, but is a US Permanent Resident	 $\mu = 2.98, \text{stdev} = 1.33$	 $\mu = 2.97, \text{stdev} = 1.65$
in the US legally, but is not Citizen or Permanent Resident	 $\mu = 2.77, \text{stdev} = 1.63$	 $\mu = 2.74, \text{stdev} = 1.63$
in the US, but not legally and is not a Citizen or Permanent Resident	 $\mu = 2.33, \text{stdev} = 1.62$	 $\mu = 2.34, \text{stdev} = 1.65$
outside of the US and is not a Citizen or Permanent Resident	 $\mu = 2.27, \text{stdev} = 1.60$	 $\mu = 2.19, \text{stdev} = 1.60$

attributes tied to the respondent as an individual (home address, last name, first name and last name). As shown in Table 2, respondents were more comfortable with information being searched at the broader geographic area level than at the individual level for both real estate records (Wilcoxon signed rank test clustered by subject as per the method detailed in [22],  $p < 0.001$ ) and campaign contributions

Table 4  
 Preferences for access to public records details

Details should. . .		Real estate purchases ( <i>n</i> = 129)	Campaign contributions ( <i>n</i> = 130)
As is	be publicly accessible, as they are now, including open online access	57%	53%
Change	be publicly accessible, but their online access should be restricted	20%	15%
	be publicly accessible, but they should not be accessible online	12%	18%
	not be publicly accessible	12%	13%

( $p < 0.001$ ). Further, in the case of campaign contributions, for searching by employer and occupation, respondents expressed similar levels of discomfort as with individual attributes.

The only attributes on which respondents' comfort differed significantly between real estate purchases and campaign contributions were with searching by neighborhood and by home address. As one might expect and also as shown in Table 2, respondents were more comfortable with real estate transactions being searchable by neighborhood (paired Wilcoxon signed rank test,  $p < 0.005$  after Bonferroni correction) or home address (paired Wilcoxon signed rank test,  $p < 0.001$  after Bonferroni correction) than campaign contributions. Whether respondents were otherwise public with their political preferences was not significantly correlated with their comfort with how people could search for the records of their campaign contributions, even when that behavior was strongly connected with the type of search (e.g., displaying a political bumper sticker or sign was not correlated with comfort with contributions being searchable).

Respondents also drew a distinction between access to information about themselves that was part of the public record by those who were fellow US citizens or legitimately living within the US, and those who were not citizens or were illegally living in the US. Specifically as shown in Table 3, respondents were more comfortable with people living in their city, state, neighborhood, or country being able to search for records of real estate transactions or campaign contributions than people not legally in the US or non-citizens outside of the US (paired Wilcoxon signed rank test, all  $p < 0.01$  after Bonferroni correction). Respondents were also more comfortable with people living in their neighborhood accessing real estate transaction data than campaign contribution data (paired Wilcoxon signed rank test,  $p = 0.003$  after Bonferroni correction).

### 3.3. Attitudes about change

Respondents were divided about whether access should change. More than half believed records should be left as-is (53% for campaign contributions and 57% for real estate records), while the others felt records should either remain accessible online but with access controls based on who and why, remain as public records but no longer be available online, or should no longer be public records (see Table 4). We did not present these options as a set of plausible resolutions, but instead asked them to better learn about respondents' comfort levels with different levels of access.

Among respondents who both did not think records should be left as is and provided an explanation about what should change (31% for campaign contribution records and 28% for real estate records) the most common response was that data should remain online but that access should be restricted (9% for campaign contributions and 12% for real estate purchases). Some respondents felt that information should be limited to those with a "need to know" or to US persons only, while others felt that anyone who registered with the site should be able to access the records. Another participant suggested a nominal

charge for access, to cut down on “frivolous” requests. Others felt that rather than limiting access to the data, the amount of data available should be reduced (7% for campaign contributions and 5% for real estate purchases) – such as by removing names from real estate records published online, eliminating employer and occupation from campaign contribution records, or creating a higher reporting threshold that contributions would have to exceed to be published online.

Respondents who clarified why data should be taken offline felt that people who “needed” access could get it through other channels. “*This is personal information; personal trumps public access,*” but someone who “*needed*” access could visit the “*city/county Admin but not where it can be readily accessed online.*” These respondents preferred the pre-online access balance between accessibility and privacy, feeling that those with legitimate reasons could go through the trouble of making a request for paper records or appearing in person, but that such effort would deter those with more prurient interests in the data.

One participant who argued that the data should not be part of the public record at all clearly was thinking of the voting booth as her contextual reference when she thought about her political contributions:

*There should be right to privacy – like the secret ballot – political affiliations & contributions should remain private. I was major ticked off to find my contributions, candidates’ affiliations, etc online – no one’s business in the public arena.*

Others echoed this: “*there should be no access. One’s political affiliations should only be made public voluntarily.*”

Not all responses that argued for change were in favor of restricting or eliminating access. Three respondents (2%) argued for making records of campaign donations *more* accessible, such as by showing real-time updates of the top 100 contributors on popular sites, real-time or near real-time publication of records, or better showing correlations between donors’ interests and the recipients’ votes. For real estate records, one respondent argued that the data should be made more accessible, and two argued that more should be included, such as asking prices or the property’s builder.

Respondents who were not sure if access to campaign contribution records should change often expressed tension between valuing transparency in support of accountability and wanting to maintain privacy: “*this is a dilemma: privacy vs. protection from influence by individuals or groups*” and “*I prefer privacy & I’m not sure how to address the issue.*”

### 3.3.1. Correlation with future plans

Respondents who were not previously aware of the disclosure laws were more likely to want records of campaign contributions to be restricted in some way (66% vs. 30%;  $\chi^2 = 12.67$ ;  $p < 0.05$ ). More respondents who felt that access to campaign contributions should somehow be restricted or removed from the public record also said they were likely to contribute less to campaigns in the future (33% vs. 3%;  $\chi^2 = 24.65$ ;  $p < 0.005$ ). Could this effect be the result of the initial surprise or “shock” from learning that about the availability of records online? In other words, could it be that this correlation between the desire to restrict access and likely lower contribution came from those respondents who learned about the availability of public records from our survey? To test this, we repeated the analysis but including only those respondents who were previously aware that records were available online. The test showed the same effect as before (27% vs. 2%;  $\chi^2 = 8.13$ ;  $p < 0.01$ ), indicating that, independent of how the respondents learned that contribution records were being made publicly available, roughly one-third of respondents who wanted access to be restricted also believed it would affect their future donation behavior. Though we did not ask respondents about whether the accessibility of records had already changed their behavior, one respondent wrote, “*This issue has already made us unlikely to ever*

donate to a presidential campaign again.” We did not see this effect with attitudes towards real estate transactions.

#### 4. Discussion

The survey results indicate mixed levels of comfort with online access of public records. Respondents were generally comfortable with the public records information being searchable by large geographic areas (state, city, ZIP code), but less comfortable at the individual level and generally uncomfortable being searchable by last name, first name *and* last name, employer, or occupation. (It is possible, though unlikely, that this is a result of the order in which elements of the question appeared in the survey, which did not vary).

Respondents were more comfortable with real estate transactions than campaign contributions being searchable by neighborhood and home address; this is expected since neighborhood and home address is often directly relevant to the real estate transaction. While this difference may be unsurprising, its implications are nonetheless important; if restrictions were to be imposed on the way in which data are searched, then such restrictions cannot be uniformly applied across different types of records. Rather, the nature of the record must be taken into account.

The previous point, combined with the finding that respondents were more comfortable with information being searched at the broader geographic level than by characteristics tied to them as individuals, suggests another principle: In general, it may be preferred to use a level of granularity that is coarser than the individual, unless doing so renders the information useless. In the case of real estate records, people may be more comfortable with the information being searched by home address because they understand that this information adds substantially to its usefulness.

##### 4.1. What to do about the lack of comfort

The results of our survey highlight how increased accessibility can make public information more public, at times conflicting with people’s existing privacy expectations and in some people prompting them to at least consider changing their behavior as a result.

##### 4.1.1. Changing policies

Less than half of the respondents felt that access to public records online should be restricted in some way or eliminated (47% for campaign contributions and 44% for real estate records). Only a little more than 10% believed that access should be completely removed, while the majority of changes desired seemed more intended to counter the increased accessibility afforded by advances in technology. If access is not restricted, some respondents intended to give less to campaigns in the future (though real estate purchase plans were not affected by the accessibility of these records). These findings support concerns (e.g. [9]) that increased availability of public records and decreased privacy may reduce people’s discretionary activities in the public sphere. Though the perception of a privacy violation might be predicted by Adams and Sasse’s model for privacy invasion [1], our survey results highlight complex tensions among technological advances, public expectations about how the data will be used, and limited resolution of public records laws.

The feasibility of some of the changes suggested by some of the respondents is unclear. For example, it is unclear whether records could be public yet also have requirements that they be accessible online to only certain people or groups. Even if access to official repositories or particular mashups were

limited to a group, such as US persons, or to those who somehow justified their request for access, others could simply access and repost the records. Warner and Chun propose one way this could be achieved – through implementation of individual privacy policies that govern the use of public records online and in mashups – and an enforcement mechanism [30]. There are, however, some challenges with this approach. First, users are often reluctant to engage in complicated management of privacy settings, though the combination of this and defaults that are consistent with privacy norms may be an improvement on the status quo for many. Additionally, US courts have historically viewed the privacy of information: either something is private and protected, or it is completely public, with no limits on use or distribution, though committees and courts have barred the online publication of some public records [25].

#### 4.1.2. Education and awareness

Some policy changes may ameliorate concerns about how accessible and searchable some public records have become, but it is also possible that these changes would be legally infeasible or unacceptably diminish the benefits of collecting the data and making them public records in the first place. Instead of or in addition to policy changes, interfaces – including donation websites, contribution forms, and scripts for in-person or telephone interaction – can be better designed to help people make better-informed decisions or to understand why the data are being collected and made available.

Policy makers or groups collecting data might also be able to improve the information they offer about what will become public record. Though most organizations disclose that the information will be made public, our finding that awareness about the online availability of public records was no different between people who had given and those who had not, and those who owned property and those who did not, suggests that organizations collecting information that will become public record may need to better inform people of what they will do with the collected data as well as what others might do with it. Because some participants were unable to explain *why* this information should be part of the public record, organizations collecting and sharing information might get a more positive reaction to their explanations of disclosure requirements by better educating the public about why these requirements exist.

The right level of education and awareness may be a difficult balance to achieve. Potential campaign contributors and home purchasers should know what information is being collected and with whom it will be shared, as well as understand the range of ways that data might be used (e.g., that it could be combined with other data sets). At the same time, such information should not be so scary that it inappropriately deters individuals from making campaign contributions – something seen as an important form of political participation, with consequences for campaigns, elections, and how elected officials govern [6,22,26]. Improved, and possibly standardized, approaches to privacy labels – such as the “nutrition label” approach proposed by Kelley et al. – may be one way to better educate people about which information is being collected and how it will be used [15]. Our survey data and the explosion of mashups of different data sets suggest two additional considerations for such labels: explaining why the data are being collected and shared, and communicating the potential privacy implications of this data being combined with other data sets.

An awareness of these attitudes may also inform tool builders and designers. For example, Sunlight Labs recently released a tool called Influence Inbox,<sup>2</sup> which annotates the email messages in the inbox of Gmail users. For organizations, it shows where their funding comes from; for individuals, it shows

---

<sup>2</sup><https://inbox.influenceexplorer.com/>.

the contributions they have made that appear in public records. The attitudes revealed in our survey suggest that some people might find the first feature to be useful, but feel that using the second feature is a privacy violation, yet the Influence Inbox tool does not give users an option to selectively turn off undesirable features.

#### 4.2. *Limitations and future work*

We did not ask respondents whether their behavior had already been influenced by the online availability of campaign contribution history. This limits our ability to say to what extent the attitudes revealed by our survey translate to changes in behavior. This will be an important question for future research, as attitudes about privacy rarely translate directly to privacy behaviors [26], with surveys not being an accurate predictor of actual decisions [7].

While our survey data go a good distance toward revealing the complex value tensions surrounding the increasingly public access to public records, surveys are limited tools for diving deeply into people's reasons, hopes, fears, and conceptions. Follow up interview studies that pursue the trends identified in the surveys would be helpful to investigate further (a) acceptable levels of granularity for information in online public records; (b) design solutions that would reasonably scope access to citizens without extending to other non-citizens; and (c) to engage with participants in rethinking as a society what it means for a public record to be public.

### 5. Conclusion

In the time since public record laws were crafted, advances in technology have substantially changed what it means for public records to be public, by allowing information to be browsed, searched, aggregated, and globally accessed. Notably, while people may have been comfortable providing governments with some information that would be stored as a public record, they may have assumed that the government would take on some sort of gatekeeper role, even if only making information inconvenient to access or to browse. This at one time presumed safeguard is no longer present once the information is placed online. When these records are combined with other data sets – such as a list of someone's Facebook friends or by placing it on a map – individuals can now easily observe their friends', coworkers', and neighbors' political contributions and real estate transactions. Previously, accessing, viewing, and aggregating these records with other data sources was sufficiently difficult that this behavior was essentially private even if the records have long been public. These advances have only continued in the years since our survey.

Though these technological advances have great benefits for values of transparency, accountability and democracy, these same advances have also disrupted the existing balance among these values and personal privacy, and are at odds with many of our survey respondents' expectations for how this data is and should be distributed. This paper confirms that at least some people perceive privacy violations when public records are made available online, as predicted by Adams & Sasse [1] and by Nissenbaum [19], and that there is a strong tension between transparency and individual privacy in the decision to disclose, or not disclose, these records. Researchers and practitioners should continue work on ways to address this tension, such as some of the directions we discuss or solutions such as those proposed by Warner and Chun.

Such disruption may have the unwanted outcome of individuals reducing their engagement in public activities. Further study is needed to determine the actual impacts on behavior. Based on those results,

policy makers will need to revisit how records of campaign contributions and real estate transactions are disseminated, and designers of tools to access these records will need to build tools more consistent with people's expectations about use of the data they provide for public records.

## Acknowledgements

We thank Amanda Fondville and Charity Leang for assistance with data collection and data entry. The Michigan Interactive and Social Computing group and anonymous reviewers provided valuable feedback on this paper. An earlier version of this work appeared at the 2011 International Conference on Digital Government Research. This work was funded, in part, by the National Science Foundation under award IIS-0325035.

## References

- [1] A. Adams and M.A. Sasse, Privacy in Multimedia Communications: Protecting Users, Not Just Data, in: *People and Computers XV – Interaction without Frontiers: Joint Proceedings of HCI 2001 and IHM 2001*, A. Blandford, J. Vanderdonck and P. Gray, eds, 2001.
- [2] V. Bellotti and A. Sellen, Designing for Privacy in Ubiquitous Computing Environments, *Proceedings of ECSCW'93* (1993), 77–92.
- [3] E. Bender, Case Study: FollowTheMoney.org, in: *Open Government: Collaboration, Transparency, and Participation in Practice*, D. Lathrop and L. Ruma, ed, O'Reilly, 2010.
- [4] L. Brandeis, *Other People's Money and How the Bankers Use it*. New York: Frederick A. Stokes, 1914.
- [5] J.E. Cohen, "Examined Lives: Informational Privacy and the Subject as Object" 52 *Stanford Law Review*, 2000, 1373.
- [6] J.J. Coleman and P.F. Manna, Congressional Campaign Spending and the Quality of Democracy, *Journal of Politics* 62(3) (2000), 757–789.
- [7] K. Connelly, A. Khalil and Y. Liu, "Do I Do What I Say?: Observed Versus Stated Privacy Preferences," In *Proc. INTERACT 2007*, 2007.
- [8] B. Friedman, P.H. Kahn, Jr., J. Hagman, R.L. Severson and B. Gill, The watcher and the watched: Social judgments about privacy in a public place, *The Human-Computer Interaction Journal* 21(2) (2006), 233–269.
- [9] A. Fung, M. Graham and D. Weil, *Full Disclosure: The Perils and Promise of Transparency*, New York: Cambridge University Press, 2007.
- [10] B. Givens, Public Records on the Internet: The Privacy Dilemma, In *Proc. Computers, Privacy, and Freedom*, 2002.
- [11] S. Goel, W. Mason and D.J. Watts, Real and perceived attitude agreement in networks, *Journal of Personality and Social Psychology* 99(4) (2010), 611–621.
- [12] V. Griffith and M. Jakobsson, Messin' with Texas Deriving Mother's Maiden Names Using Public Records, in *Proc. Applied Cryptography and Network Security*, 2005.
- [13] L.D. Inrona, Privacy and the Computer: Why we Need Privacy in the Information Society, *Metaphilosophy* 28(3) (1997), 259–275.
- [14] L. Inrona and A. Pouloudi, Privacy in the Information Age: Stakeholders, Interests, and Values, *Journal of Business Ethics* 22(1) (1999), 27–38.
- [15] P.G. Kelley, J. Bresee, L.F. Cranor and R.W. Reeder, A 'Nutrition Label' for Privacy, in *Proc. SOUPS '09*, 2009.
- [16] J.R. Landis and G.G. Koch, The measurement of observer agreement for categorical data, *Biometrics* 33 (1977), 159–174.
- [17] S. Lopez, (14 December 2008), Prop. 8 stance upends her life, *Los Angeles Times*, p. B1.
- [18] E. Miller, Disrupting Washington's Golden Rule, in: *Open Government: Collaboration, Transparency, and Participation in Practice*, D. Lathrop and L. Ruma, eds, O'Reilly, 2010.
- [19] H. Nissenbaum, Privacy as Contextual Integrity, *Washington Law Review* 79(1) (2004).
- [20] E. Noelle-Neumann, *The Spiral of Silence: Public Opinion, our Social Skin*, (2nd ed.), Chicago, Illinois: University of Chicago Press, 1993.
- [21] L. Palen and P. Dourish, "Unpacking 'Privacy' for a Networked World," in *Proc. CHI 2003*, 2003.
- [22] B. Rosner, R.J. Glynn and M.L. Lee, The Wilcoxon signed rank test for paired comparisons of clustered data, *Biometrics* 62(1) (2006), 185–192.
- [23] J.M. Snyder, Campaign Contributions as Investments: The US House of Representatives 1980–1986, *Journal of Political Economy* 98(6) (1990), 1195–1227.

- [24] D. Solove, "Conceptualizing Privacy" 90 *California Law Review* 1087, 2002.
- [25] D. Solove, *Understanding Privacy*. Harvard University Press, 2008.
- [26] S. Spiekerman, J. Grossklags and B. Berendt, E-privacy in 2nd generation E-commerce: privacy preferences versus actual behavior. In *Proc. EC 2001*, (2001), 38–47.
- [27] B. Stone, (8 February 2009), Prop 8 Donor Web Site Shows Disclosure Law Is 2-Edged Sword, *The New York Times*, p. BU3.
- [28] T. Stratmann, What do campaign contributions buy? Deciphering causal effects of money and votes. *Southern Economic Journal* 57(3) (1991), 606–620.
- [29] United States Supreme Court, (1976). *Buckley v. Valeo*, 424 US 1 No. 75-436.
- [30] J. Warner and S.A. Chun, Privacy protection in government mashups, *Information Polity* 14 (2009), 75–90.
- [31] S. Warren and L. Brandeis, The Right to Privacy, *Harvard Law Review* 4(5) (1890), 193–220.
- [32] D. Wolber, Political E-Identity: campaign funding data and beyond, in *Proc. dg.o* 151, 2006.

Copyright of Information Polity: The International Journal of Government & Democracy in the Information Age is the property of IOS Press and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.