I Know Where You Live: Analyzing Privacy Protection in Public Databases

Manya Sleeper, Divya Sharma, and Lorrie Faith Cranor

October 3, 2011

CMU-CyLab-11-015

CyLab
Carnegie Mellon University
Pittsburgh, PA 15213

I Know Where You Live: Analyzing Privacy Protection in Public Databases

Manya Sleeper, Divya Sharma, and Lorrie Faith Cranor

Carnegie Mellon University
{msleeper,divyasharma,lorrie}@cmu.edu

Abstract. Policymakers struggle to determine the proper tradeoffs between data accessibility and data-subject privacy as public records move online. For example, Allegheny County, Pennsylvania recently eliminated the ability to search the county property assessment database using property owners' names. We conducted a user study to determine whether this strategy provides effective privacy protection against a non-expert adversary. We found that removing search by name provides some increased privacy protection, because some users were unable to use other means to determine the address of an individual. However, this privacy protection is limited, and interface usability problems presented a comparable barrier. Our analysis suggests that if policymakers use removal of search by name as a privacy mechanism they should attempt to mitigate usability issues that can hinder legitimate use of public records databases. [1]

Keywords: Privacy, public databases, practical obscurity, public policy

1 Introduction

Policy makers struggle to manage the tradeoffs between data accessibility and data-subject privacy as they move public records online. In the United States, making records more accessible by moving them online provides a variety of benefits, supported by legal precedent dating back to 1978 when the Supreme Court recognized "the general right to inspect and copy public records and documents" [29]. However, once records are available online they are more widely accessible and easily searchable, giving rise to privacy concerns that must be balanced against the benefits of increased transparency and accessibility.

Allegheny County, a Western Pennsylvania county that includes the city of Pittsburgh, offers one example of how policy makers choose to manage these tradeoffs. The county posts detailed property information for over 350,000 properties in an online property assessment database [1]. As part of an effort to protect the privacy of property-owners, this database is only searchable by address or parcel number, not by the name of the property owner. Although the Allegheny property assessment data is public record, once the data was moved online in 2001 county officials became concerned about the privacy of property owners [10]. In 2005, after the murder of a judge in Atlanta and the killings of two family members of a judge in Chicago, officials decided to remove approximately one hundred judges from the database. Law enforcement authorities subsequently requested that their names also be removed from the database. In response, the Allegheny County Chief Executive proposed a measure to limit the database's search functionality to address and parcel-based searches[[19], [24]].

The measure passed with a 9-6 vote of the Allegheny County Council in 2007 [18]. Supporters felt that it would protect the privacy of "law enforcement officials, teachers, judges and victims of violent crime," while still allowing people to use the database to look up tax assessment information [10]. However, opponents of the measure felt that the database should prioritize ensuring the transparency and accessibility of public records [[19], [18]]. Similar tradeoffs will arise whenever policy makers seek to place privacy protections on public data.

[1] This technical report is an extended version of a paper, *I Know Where You Live: Analyzing Privacy Protection in Public Databases*, presented at *Workshop on Privacy in the Electronic Society* (WPES), October 2011

In Allegheny County, and in other, similar cases throughout the United States, policymakers must create policies that limit users' abilities to access information without empirical information about the level of privacy protection provided by the implementation of such policies. In this paper, we use the Allegheny County database as a case study to provide information to help inform policymakers both about users' behaviors when confronted with the barrier presented by a currently-implemented privacy protection technique, as well as to estimate the level of privacy protection such a barrier provides for public record databases accessible by search by name and address. To do so, we present the results of a 52-subject user study that draws on the Allegheny County database to examine the effectiveness of removing search by name as a privacy-protection mechanism. We use insights from our evaluation to offer advice to policy makers tasked with finding practical solutions for improving data-subject privacy while attempting to maintain the benefits of accessibility.

We found that removing the search-by-name functionality prevented a significant percentage of subjects from accessing information in the Allegheny County tax assessment database. However, this barrier arose in nearly equal parts from the requirement that the subjects know an individual's address and from database usability issues. Thus, we advise policy makers that removing search by name can be a partially effective tactic, especially for property owners whose addresses are not available online. However, to maintain the usefulness of databases for legitimate purposes, such a technique should be paired with a focus on interface usability.

In this paper we discuss related work and then summarize the results of an examination of a random sample of the property assessment database search policies employed by 101 counties in the US. We go on to present the methodology for and results of our user study, and conclude with a discussion of recommendations for policy makers.

2 Background and Related Work

Our examination of online property records relates to existing legal, technical and policy research, including work on practical obscurity, the privacy risks from online public records, anonymization techniques, and search behavior. We seek to build on previous work and aid policy makers by providing an evaluation of the practical effectiveness of a privacy protection measure that is currently in use, drawing on the Allegheny County tax assessment database as a case study.

2.1 Practical Obscurity

The broad debate around the privacy implications of moving records online often centers around the concept of practical obscurity. Practical obscurity, as originally defined by Supreme Court Justice John Paul Stevens, is "the phenomenon by which sensitive information can receive considerable protection, merely by virtue of the practical difficulty cost of retrieving paper based records" [34]. Moving records online lowers information retrieval costs, and the records lose the protection of practical obscurity.

A variety of scholars have written about privacy tradeoffs that occur when records are moved online. Winn asserts that practical obscurity in physical records offered a balance between accessibility and protecting the privacy of the individual. All records had a "default rule of practical obscurity" [34]. Solove also believes that practical obscurity previously made obtaining access to records more difficult; he explains that, prior to the advent of online databases, public records were confined to the various localities making the process of obtaining a record time consuming and potentially difficult. Solove argues that advances in technology increase access to records; so, moving to online databases increases the risks associated with the personal information contained in the public records [29].

The concept of practical obscurity has also been applied to legal arguments both for and against moving public records online. For example, Bepko provides a legal backdrop to New York State's online court records database and argues that all court data should be provided online [5]. On the other side, Morgan outlines the legal history behind the increasing online availability of divorce court records and argues that all divorce court cases should be kept private because of the high potential for privacy violations in the transcripts for such cases [28]. Conley et al. conduct a study of the implications on privacy and open justice of the transition

from local to online access to court records [8]. The authors consider the privacy implications of making court records available online from multiple perspectives- legal, philosophical, and technical. In particular, the authors develop a model of information retrieval involving users and computer systems and use it to compare typical, or expected information flows for in-person access to court records, at the courthouse, with online access to court records via PACER and Google Scholar. However, in contrast to Conley et al., we conduct user studies to validate our hypotheses about the effectiveness of removing the "search by name" feature in providing privacy protection against non-expert users.

2.2 Privacy risks from online public records

Moving public records online poses privacy risks. Bermann discusses how the loss of practical obscurity due to technology has given rise to "concerns ranging from citizens snooping into sensitive information [...] to the safety of public officials, such as judges and police officers to identity theft." In light of these risks, Bermann offers potential solutions to protect privacy while maximizing the benefits of electronic access: provide access to records online as they are (which would still raise privacy concerns), remove sensitive information from the database (e.g., social security numbers), "limit online access" but leave physical records unchanged, or do not put public records online [6]. Givens also describes potential privacy risks presented by online public records like motor vehicle records or court records. She outlines how these records can contain personal information such as social security numbers, financial information, family-related allegations, health information, and criminal background data, and describes potential consequences associated with making such records available online [15].

Beyond these theoretical descriptions of potential privacy risks, such risks have also been apparent in recent news stories. For example, a Cincinnati-area county put a variety of public records, including property records, online. These databases received heavy use, and the information was used for at least one incident of identity theft. The media reported that residents of the county felt that their privacy was being intruded upon [22].

Additionally, in Texas, a Travis County clerk took document images offline in response to residents' privacy concerns. In this incident records were put online that included "...addresses, social security numbers, birthplaces and driver's license numbers" [9].

2.3 Anonymization Techniques

There is a large body of work on data anonymization. Much of this work involves techniques designed to protect the privacy of individuals when data is released in aggregate for statistical purposes and requires that the data be modified in the process[[30], [26], [23], [35]]. The need to maintain the integrity of public record databases for accountability purposes might limit the usefulness of common techniques for creating "anonymized" databases that require removing certain fields or altering the data.

Differentially private release mechanisms[[12], [13]] guarantee privacy in the presence of arbitrary auxiliary information by adding statistical noise to produce a "sanitized" database that prevents an adversary from determining, up to a probabilistic bound, whether an individual is in the database. There are several approaches that could potentially apply to public databases including work on techniques for a more general framework for differential privacy [27] as well as work on non-interactive privacy-preserving mechanisms for releasing databases for a certain subset of all queries [7]. Differentially private mechanisms have also been proposed [25] that are applicable to public records databases, in the context of releasing US population commuting patterns by creating web-based mapping applications [32]. These mechanisms present opportunities for future work.

Although there exist a variety of potentially applicable anonymizaton techniques, in this paper we do not seek to compare alternative techniques for protecting public records, but, rather, to inform policymakers by understanding and evaluating a technique that is currently in use in numerous locations.

2.4 Search Behavior

The efficacy of removing search-by-name functionality as a privacy protection mechanism depends on how hard it is for users to use a person's name to find an address through a secondary source. Information retrieval and library science researchers have studied users' online search behaviors, exploring how users use online retrieval systems as well as extracting general characteristics of search strategies. They have found that, when searching for information online, as opposed to when using traditional information retrieval tools for searching large databases such as systems based on INQUERY [21] or Online Public Access Catalogs, people use shorter and fewer queries and find relevant information more quickly. Web-based search systems involve shorter sessions than traditional information retrieval systems and result in higher rates of success for locating relevant information [20]. Hargittai's work on generalized usage patterns found that although users were able to perform many online search tasks successfully, some were unable to do so, and many people "lack the basics of surfing the Web" [17]. In this study we use similar methodologies to explore users' search behavior and capabilities as they apply to testing the effectiveness of a privacy protection mechanism.

3 Review of Counties

We selected Allegheny County for our case study because we had access to both the online database and subjects familiar with the database content. Our broader goal, however, is to provide policy advice that can be generalized to counties with property databases indexable by owner name and address and that provide privacy protection by eliminating the search-by-name functionality and allowing the database to only be searched by address. We, therefore, surveyed a random sample of 101 counties from across the country to gain insights into the representativeness of the Allegheny County property assessment database.

We selected the counties from a list of 3,142 counties from the US Census [31]. We randomly selected two counties from each state to ensure an even geographic distribution. In addition, we examined the District of Columbia. For each of the selected counties, we determined if a publicly accessible property assessment database existed online, and, if one did, examined the interface to determine the various search functionalities. We also examined records accessible through the sites to roughly determine the allowable fields for a given database (Appendix C offers details). This survey provided us with an overview of the relative frequency of search functionalities and presence of potentially privacy-invasive data in county-level online property assessment databases.

We found that 70% of counties in our sample had publicly accessible online property assessment databases. Of the counties with such databases, 21% were not searchable by name. All except two of the counties allowed the database to be searched by address. Like Allegheny County, 75% of the databases allowed users to search by Parcel ID. Many databases were also searchable by geographic (e.g. subdivision, lot), legal, or sales criteria.

The Allegheny County database has several potentially privacy invasive record fields, including sales price, tax information, property photo, and property floor plans. We found that 58% of publicly accessible online databases in our sample contained a field for sales price, 61% contained a field for tax information, 24% contained a field for a photo of the property, and 31% contained a field for the property's floor plan. Overall, a large percentage, although not always the majority, of counties had search options and record fields comparable to Allegheny County, supporting its viability as a representative case study.

4 Methodology

We used the Allegheny County property assessment database to evaluate the effectiveness of removing search-by-name functionality as a privacy protection mechanism for databases that are primarily indexable by address and name. To perform this evaluation, we carried out a 52-subject user study in which we tested subjects' abilities to perform an information retrieval task that required them to access a record within the online database.

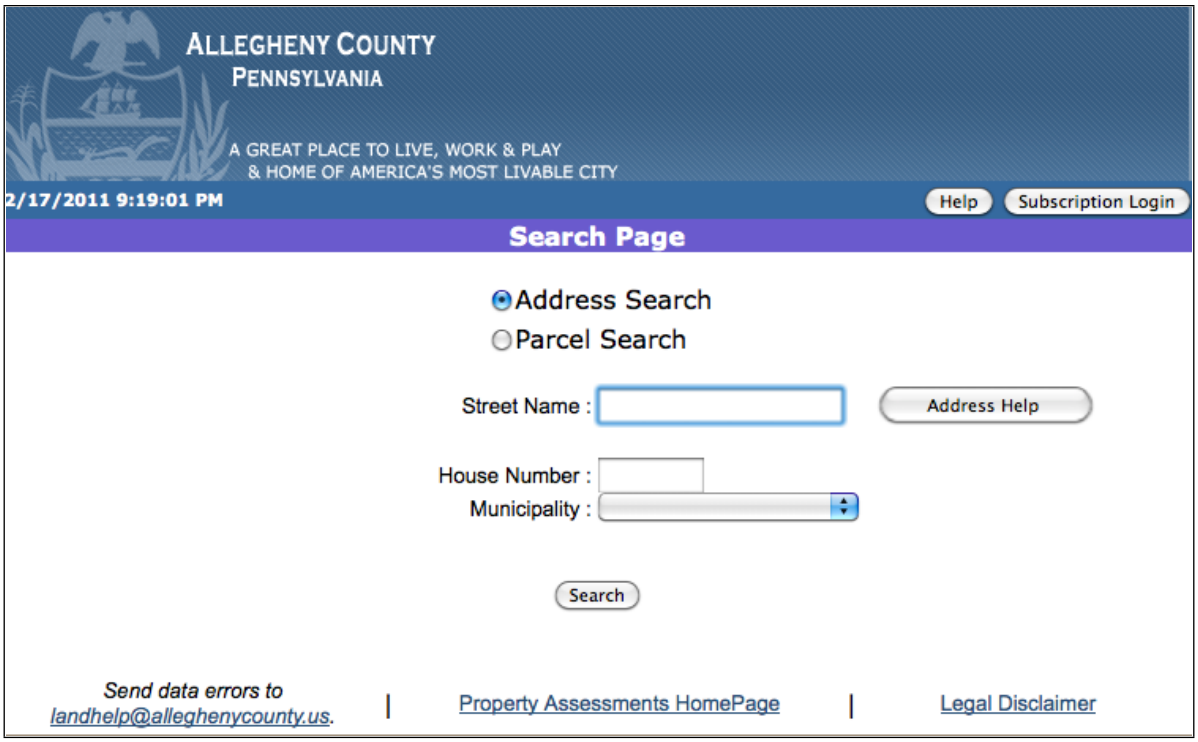


Fig. 1. Screenshot of Allegheny County assessment website search interface

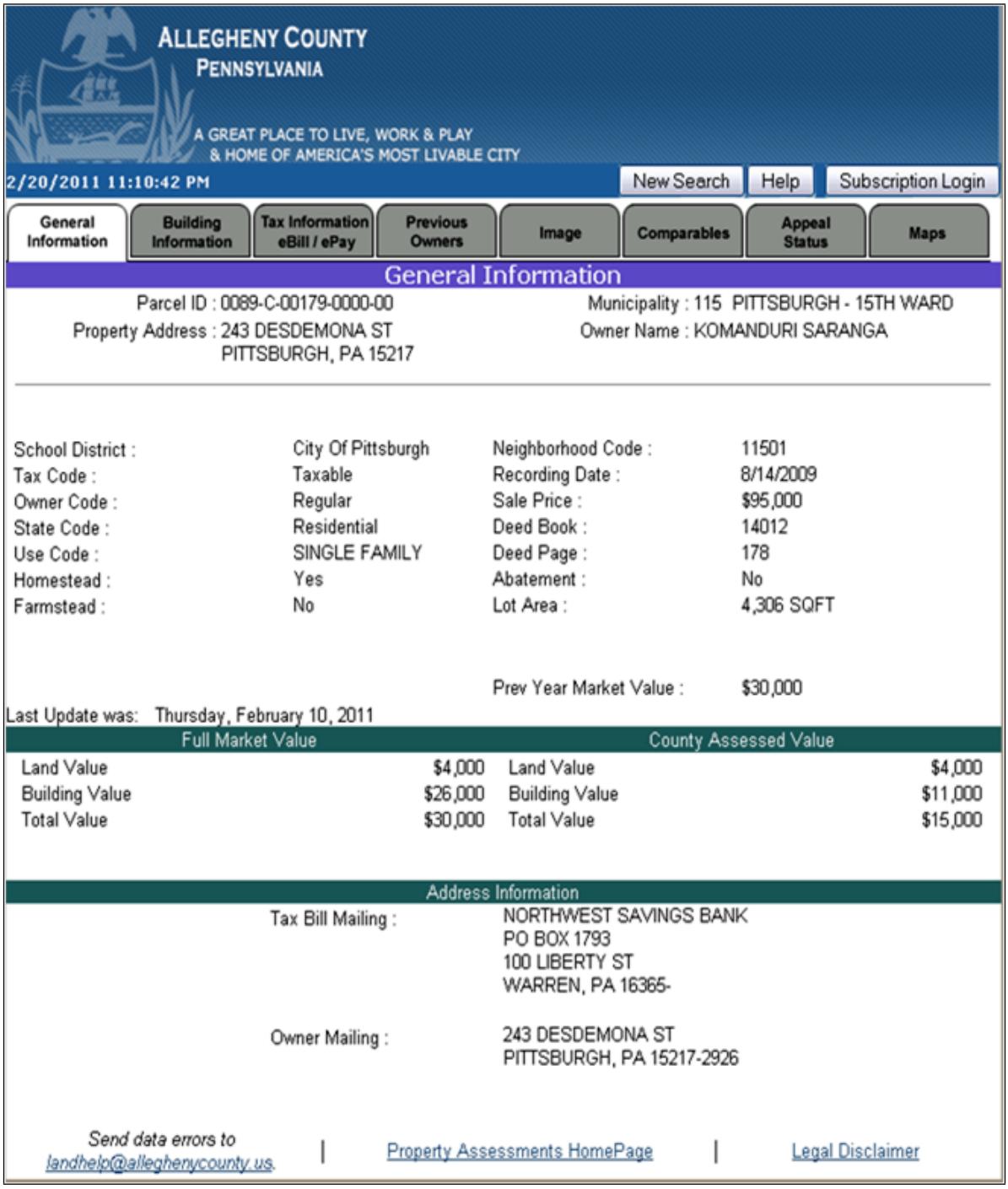
4.1 Evaluation of Removal of Search by Name

We focused on a specific risk scenario in which a non-expert community member would try to access the record of an individual with whom she did not have a close personal connection using only freely available online information. Although this limits the attacker to a somewhat narrow case, it seems to represent the primary risk case policy makers target by removing the search-by-name functionality. Expert users, such as private investigators, would likely have access to alternative databases, and someone with a personal connection to the individual is likely to know his or her address. We proxied this risk scenario in which non-expert users search for unknown parties by asking subjects to search for randomly selected individuals using free online resources.

We also focused on a scenario in which a user wanted to find information contained in a record, rather than data like an individual's address in which case the removal of search by name would only require her to find the address using an outside source and not additionally enter that address back in the database. Although the availability of data from alternative sources is also an important issue, it does not seem to be the primary risk case targeted by policymakers in removing the search by name functionality.

Allegheny County Property Assessment Database The Allegheny County Property Assessment database [1] is linked off the County of Allegheny website through the Office of Property Assessments. A user can search the database by address or by parcel ID (Figure 1). Resulting records provide a variety of property tax-related information fields—some of which are potentially privacy invasive—including the sales price, tax history, and a photo of the property (Figure 2 provides an example).

Name Selection If such a database, indexable primarily by name and address, were accessible through an interface that did not limit the ability to search by name, property owners might either be linked to one



ALLEGHENY COUNTY
PENNSYLVANIA

A GREAT PLACE TO LIVE, WORK & PLAY
& HOME OF AMERICA'S MOST LIVABLE CITY

2/20/2011 11:10:42 PM

New Search | Help | Subscription Login

General Information | Building Information | Tax Information eBill / ePay | Previous Owners | Image | Comparables | Appeal Status | Maps

General Information

Parcel ID : 0089-C-00179-0000-00
Property Address : 243 DESDEMONA ST
PITTSBURGH, PA 15217

Municipality : 115 PITTSBURGH - 15TH WARD
Owner Name : KOMANDURI SARANGA

School District :	City Of Pittsburgh	Neighborhood Code :	11501
Tax Code :	Taxable	Recording Date :	8/14/2009
Owner Code :	Regular	Sale Price :	$95,000
State Code :	Residential	Deed Book :	14012
Use Code :	SINGLE FAMILY	Deed Page :	178
Homestead :	Yes	Abatement :	No
Farmstead :	No	Lot Area :	4,306 SQFT
		Prev Year Market Value :	$30,000

Last Update was: Thursday, February 10, 2011

Full Market Value		County Assessed Value	
Land Value	$4,000	Land Value	$4,000
Building Value	$26,000	Building Value	$11,000
Total Value	$30,000	Total Value	$15,000

Address Information

Tax Bill Mailing : NORTHWEST SAVINGS BANK
PO BOX 1793
100 LIBERTY ST
WARREN, PA 16365-

Owner Mailing : 243 DESDEMONA ST
PITTSBURGH, PA 15217-2926

Send data errors to landhelp@alleghenycounty.us. | Property Assessments HomePage | Legal Disclaimer

Fig. 2. Screenshot of an Allegheny County property assessment record

property record in the database or multiple property records. If the individual were linked to one record, her name would be adequate for finding the record. Otherwise, it would not be possible to find the record without additional information.

We wanted to evaluate how removing search by name impacted finding records for individuals who would have been locatable by name if the interface had included search by name. Thus, in the case study we were interested in subjects' abilities to find records for individuals with names tied to unique records in the database. Records for individuals linked to one record could be found by name using the search by name functionality. However, in the absence of search by name, finding these records would require a user to first determine the individual's address using another mechanism.

Based on our risk scenario, we limited the auxiliary resources available for finding this address to free online resources.

Individuals with names (defined as first and last names) tied to unique records in the property assessment database fall into one of three categories distinguished by the availability of free online information about their addresses on address-search websites. For individuals in the first category, address information is not available from free online resources. Such individuals might include privacy-conscious people who have requested that their information be removed from websites that provide address information or whose information was removed because they have unlisted phone numbers. It would not be possible to find database records for such individuals without additional information about the individuals or without using offline or paid auxiliary sources.

Individuals in the second category have names that are associated with multiple addresses either because they own multiple pieces of property or because they have names that are not unique in the address-search website dataset. For individuals in this category it would be possible to determine the individual's unique record by entering each potential address into the search-by-address interface, because the individual's name would be tied to a unique record in the property database if not on address-search sites. Depending on the number of addresses associated with the name, this might be trivial or very difficult.

Individuals in the third category have names that are uniquely associated with an address using free online resources. Once a user retrieves this unique address she can enter it into the search-by-address field and directly locate the corresponding record.

Based on a crawl of the Allegheny property assessment website, we randomly selected a sample of 204 names tied to unique records. We looked up each name on WhitePages.com, a commonly used free resource for finding address information, to determine the category for each name. We found that $\sim 30\%$ of the sample had addresses that were not available on WhitePages.com, thus, falling into the first category. We assumed that these individuals would be protected by the switch to search by address, so we did not include such individuals in our user testing. For $\sim 15\%$ of our sample, we found two or more Allegheny County addresses linked to the same name in a search on WhitePages.com, placing these individuals in the second category. In the absence of additional information, locating a record in the Allegheny County assessment database for one of these individuals would require that a user enter each address that showed up in the search until she found the one linked to a record. For the remaining $\sim 55\%$ of individuals, only one address in Allegheny County was tied to the individual's name on WhitePages, placing them in the third category. When searching for these individuals, it would be possible for a user to locate the address and enter it into the property assessment database to access the individual's record.

We selected the random set of individuals that our user study subjects searched for from a subset of the third category. We limited our selected subset to individuals for whom less than four records appeared on a Pennsylvania WhitePages.com search for the individual. Out of the 112 individuals we found in our sample who had names tied to unique addresses in Allegheny, we eliminated 19 based on this requirement. Although this choice potentially made the task easier, it allowed us to focus on determining whether subjects were able to find an address using auxiliary sources and enter it into the database, rather than focusing on whether subjects could understand large numbers of results on an address search interfaces. In a more realistic scenario a user might find more than three results, but would also likely have some auxiliary information about the individual or property to help with tasks like sorting through the White Pages results.

Study protocol We performed a user study with 52 subjects, divided into two conditions. We asked all subjects to find the amount of Allegheny property tax a randomly selected individual paid in 2010, using any freely available online resources. To the best of our knowledge this information is only available in the county assessment database. In the first condition ($C1$) subjects were provided with just a random individual's name, simulating a user being required to use the interface in its current state without search by name. In the second condition ($C2$) subjects were provided with a random individual's name and address. Condition $C2$, our control condition, served as a proxy for a user having access to search by name. In this case the subject could type the address directly into the interface.

When subjects arrived, they were told that they would be participating in a study on how people use the Internet to find information. They were provided with a Mac laptop with a mouse and the Firefox web browser set up as the default. To prevent users unfamiliar with the computer from being disadvantaged (62% of our subjects self-identified as PC users) we told all subjects that they could ask the researcher computer- or browser-related questions at any point throughout the task. We did not use a think aloud procedure. However, we recorded all subject actions with screen capture software and used audio capture to record any incidental comments.

To begin the study, we presented the computer to the subjects with a mocked up newspaper article on the screen that provided a summary of public databases available in Allegheny County. This article was based on an actual news article [4] and provided details about the property assessment database and other databases as well as a URL for the property assessment database. We provided the article to simulate a realistic situation in which a user might be informed about the database and decide to use it. For example, users might try to use such a database after reading a newspaper article about it, especially as the database had recently been in the media because of the public debate around the removal of search by name. We decided to present the article at the beginning of the task based on the results of an initial 13-subject pilot. In the pilot we presented the article to subjects after they were allowed five minutes of free search. During this pilot we found that, across conditions, subjects required the article to reach a useful resource. Providing the article at the beginning of the search likely led to more successes across both conditions.

After a subject read the article the researcher provided her with the task and gave her ten minutes to perform it. If a subject entered the database and reached the screen with the Allegheny County tax information within ten minutes, we considered her to be successful. We noted both the time when a subject entered the database and when she reached the screen with the tax information, located on a tab within the record. For subjects in $C1$, we also noted whether the subject was able to find the individual's address, the primary sites used, and how long it took the user to find the address.

We chose to limit the time considered for success to ten minutes because in our pilot testing we found that if users did not complete the task within that time period they were unlikely to be successful. At the ten minute mark, our pilot subjects tended to perform repetitive non-progressive actions. During the subject testing, if a subject was unable to perform the task within ten minutes, we offered her an additional ten minutes. Subjects were not considered successful in our analyses if they found the information during the second ten minute block; however, allowing this additional ten minutes allowed us to determine whether users might be able to complete the tasks if they were willing to spend more time working on them. We found during our testing that only four subjects in $C1$ were successful after taking an additional ten minutes, and all of the subjects had previously found the address in the first ten minutes.

Each subject was paid \$5 for performing the task and an additional \$2 bonus if she found the requested information within twenty minutes.

After completing the task, within ten or twenty minutes, we asked the subjects to fill out an online exit survey on the computer on which they had performed the task.

Subjects We recruited native English speakers in the Pittsburgh area through Carnegie Mellon's Center for Behavioral and Decision Research research pool, Craigslist, and a recruiting session at a Pittsburgh coffee shop. We distributed subjects between conditions in a round robin fashion. To recruit non-expert users we asked prospective subjects about their occupations or degree programs. We screened out individuals who

self-identified as working in science, engineering or as IT professionals, or who self-identified as students studying computer science, engineering or IT.

Appendix A provides details on the demographics of our subjects. Slightly over half our subjects were female (54%). Approximately half of the subjects were students (52%), and the remaining subjects were distributed over a variety of non-technical fields. 54% of our subjects were 18-24 years old, 19% of subjects were 25-29 years old, 13% of subjects were 30-34 years old, and 13% of our subjects were over 34 years old.

Hypothesis Our hypothesis was that the removal of the search-by-name functionality provides significant privacy protection, resulting in a significantly higher number of individuals who would be able to perform the information retrieval task in $C2$ than in $C1$.

5 Data and Analysis

Removing search by name from the database reduced the number of users who were able to access records in the database, partially confirming our hypothesis. Significantly fewer subjects in $C1$ than $C2$ were able to successfully find tax information in the property assessment database. However, to properly put this effect in perspective, the portion of this effect that arose from the need for subjects to know individuals' addresses was roughly equal to the portion of the effect that came from interface usability issues. Exit survey results showed that subjects recognized the usefulness of the database but had some concerns about its privacy implications.

5.1 Privacy Protection Mechanism

We compared the abilities of subjects in two conditions to find tax information found only within the property assessment database. In $C1$ the subject was only provided with an individual's name and, therefore, first needed to find the individual's address using auxiliary free online sources. The subject then needed to enter the address into the property assessment system and access the record. In $C2$, the control condition, the subject was provided with both an individual's name and address and, therefore, only needed to perform the second step. We partially confirmed our hypothesis and found that significantly fewer subjects in $C1$ than $C2$ ($p = 0.001$, Fisher's Exact Test) were able to find the requested information in the database. In $C1$, 19% of subjects were successful, while in $C2$, 68% of subjects were successful.

The percent of subjects able to find an address in $C1$ must be placed in the overall context of names of property owners within the database as described in the Methodology section. First, this percentage is an upper bound on the percent of subjects who would be able to find an address if we had included individuals with multiple address listings available in Allegheny County and individuals with more than three listings in Pennsylvania. Finding such individuals would be more difficult than finding property owners with unique addresses.

We also know that the 19% success rate represents subjects who searched for property owners with addresses available online. In reality, there are also $\sim 30\%$ of individuals with unique records who do not have addresses available online. Thus, this 19% success rate can only be applied to the 70% of unique property owners with addresses available through free online sources. The success rate would drop to $\sim 13\%$ if applied to all unique property owners.

Although removing search by name did not provide complete privacy protection it did present a significant barrier to accessing information within the database for our subjects. However, our hypothesis was only partially confirmed. Public record privacy protection must balance privacy with the maintenance of data accessibility. We found that the current access barrier is due both to the privacy protection provided by the requirement that a user know an individual's address and usability issues arising from the interface design. These usability issues hindered accessibility despite subjects having the knowledge required to enter the database.

While evaluating and increasing the usability of public records websites is beyond the scope of this paper, as Gant and Gant note, "Portals deliver value to the user as a function of how accessible and usable the

features on the site are" [14]. Thus, we noted the usability problems subjects encountered to qualitatively examine whether removing the search by name functionality exacerbated such problems. In our control condition we found that over half of the subjects in $C1$ (56%) were able to find the address of the individual whose record they were searching for. However, 37% of subjects in $C1$ were not able to find the property record after finding the address information due to interface usability issues; only 44% of the subjects in $C1$ were prevented from accessing the database because they were unable to find the address. These interface usability issues were also evident in $C2$, although to a lesser extent, suggesting that the usability problems were compounded by the search-by-address requirement. 28% of subjects in $C2$ were unable to access a record due to interface usability issues. Overall, 73% of subjects across both groups who knew an address had at least one usability problem while trying to enter it into the interface.

Additionally, our pilot findings that subjects were unlikely to be successful if they did not complete the task in ten minutes were maintained in our user testing. Only 33% (eight in $C1$ and two in $C2$) of subjects who had not completed the task opted to continue for an extra ten minutes, and an additional four subjects in $C1$ and zero subjects in $C2$ were successful in these second ten minutes. Subjects in $C1$ may have benefited from the additional time to overcome usability issues after finding the property owner's address.

5.2 Address Search Strategies

The intended privacy protection should have arisen from the subjects' inabilities to find addresses for individuals once search by name was removed. In our group $C1$ we observed that subjects used a variety of strategies when asked to find an individual's tax record, some of which were more likely to allow them to find an address and access the database. Of the subjects who successfully found an address, the majority (twelve out of fifteen) performed a Google search for the individual's name and some additional search terms, such as "Allegheny county" or "address." This technique produced search results that included one or more free online address search sites, such as WhitePages.com or City-Data.com, that the subjects used to find the individual's address. The remaining subjects who found addresses more directly targeted the address-search sites; one subject typed in the site URL, and two others performed Google searches for the names of the address search sites.

The subjects who did not find an address employed the first strategy unsuccessfully; 9 out of 12 of these subjects performed at least one Google search containing the subject's name, sometimes with additional relevant keywords. For 5 of these subjects at least one of these searches was unsuccessful because the subject ended up on a paid information retrieval website (e.g., myLife.com). 5 out of 12 of these subjects hit dead ends either because the search did not produce any useful results, or the subject did not realize that the address search website would be useful. One subject found a free address search site, but was unable to find an address on it.

Some subjects appeared to be searching for an alternative to the Allegheny County database. For example, subjects performed Google searches with search terms related to the task but not including the individual's name (e.g., "Allegheny county tax public record"). 10 out of 12 unsuccessful subjects in $C1$ performed such a search at least once. These searches tended to produce paid sites or direct subjects to resources that were not useful for finding address results, like the Allegheny County database itself. We also observed that 10 out of 12 unsuccessful subjects spent some time exploring the Allegheny County website or other public record databases (e.g., marriage or court records) rather than directly searching for the address. This behavior was likely somewhat motivated by the structure of the task and the fact that we provided them with an article outlining where to find public record information.

5.3 Usability

Usability issues arose in both conditions because the Allegheny County site required users to enter addresses in a specific format. The specific format required likely worsened the usability issues when users were required to find an address, resulting in the higher rate of usability issues in $C1$ than $C2$.

Across both conditions, 56% of subjects experienced at least one usability issue. 12 out of 15 subjects who found the address in $C1$ and 17 out of 25 subjects in $C2$ faced at least one usability issue. Subjects

most commonly entered a street suffix incorrectly. This issue arose because the database interface required that subjects enter suffixes, such as "Street" or "Avenue" in abbreviated form (i.e. "St" or "Ave") without a period, and any other format did not work. Another common issue subjects experienced was entering an incorrect municipality into the interface. Subjects often had trouble determining the municipality associated with an address and were unable to access the database if they entered the incorrect municipality. Overall, 33% of subjects were unable to overcome usability issues and were prevented from accessing the database due to such issues. Appendix B provides a breakdown of usability issues. These usability issues contributed to preventing subjects from accessing the database, especially in $C1$, even when they had access to address information. Although addressing such usability issues for both the Allegheny County case, and property assessment databases in general, is beyond the scope of this paper, such issues should be addressed to ensure data accessibility, especially as they may be worsened by removing search by name for privacy protection purposes.

5.4 Analysis of Subjects' Attitudes

We also analyzed the attitudes of 28 subjects who were either able to access the database (26) or who had previously used the database (2).

These subjects recognized the utility of such databases. 86% strongly or somewhat preferred to look up information online rather than visit a public office, and 79% at least somewhat agreed that online public records databases were useful. However, the subjects also seemed to recognize the privacy tradeoffs inherent in putting such databases online. Fewer subjects strongly or somewhat agreed (71%) that public records databases should be available online. More than half of the subjects somewhat or strongly agreed that having such databases online made them concerned about their privacy (61%), although half of the respondents also somewhat or strongly agreed that the benefits of having the databases online outweighed the risks (50%).

This recognition of the benefits of accessibility combined with a concern about privacy was reflected in some user comments about the first thing that came to mind when they accessed the database. Some subjects found it very useful: "This is cool...I wish all the resources were online," "Impressive record-keeping and accessibility." Others were concerned by the privacy implications: "I wonder how easy it is to look up personal information for others," "I didn't realize one could find this type of information online - & so much! Also makes me a little leary (sp?) of what people could find out about me online."

6 Discussion

Based on our findings we can offer several recommendations to policy makers determining how to best balance data privacy and accessibility for databases like public property records databases. Although our findings may not be surprising for privacy experts, we hope that our study provides empirical data to inform policy makers' decisions as they decide whether to remove search by name to protect the privacy of individuals with records in online databases in databases like those addressed in this paper. We found that removing the search-by-name functionality had merit, although it only provided partial privacy protection. Policymakers focusing on increasing privacy protection while maintaining data transparency could partially address the limitations of the search-by-name functionality by focusing on maintaining or improving the usability of the database's interface.

6.1 Privacy Protection

Removing the search by name feature provides privacy protection in the form of a "speed bump." Although it would not prevent a determined individual from accessing information, it does slow down casual searchers. The size of this speed bump depends on whether a property owner's address is available online. Removing search by name provides the largest privacy gain to property owners who have addresses that are not available on free address-search websites ($\sim$ 30% of individuals with unique records). A large portion of such people

are likely to be individuals with unlisted phone numbers or people who requested that their information be removed from the websites.

Additionally, our study indicates that removing search by name would likely thwart a relatively large percentage of non-expert users who might casually search for a stranger in the database by providing a large enough speed bump to deter such users. For a relatively small bonus ($2), and in a realistic amount of time (10 minutes) some of our subjects ($\sim 19\%$) were able to find the information. This privacy mechanism would likely block a large percentage of non-expert users, but the privacy protections provided by the removal of search by name are limited.

Determined searchers can find means to access the database. Even non-expert users can overcome the barrier presented by the lack of search by name by investing additional time or money. If a user wants to access the records using the Internet, she can use a paid online service, like Intelius [2], and get the information (address report) for $\sim \$1.95$, or she can purchase the right to access the search-by-name feature in the Allegheny County site from the county (for $50 per month plus a one-time registration fee of $100 and an hourly rate varying from $\$1 - \$6$) [3]. Beyond online resources, the user could visit the County office building to find the information or ask a private investigator to find the information for her for prices ranging from $\sim \$25$ to $\sim \$150$ (dependent on the time taken to find the address).[2]

The speed bump is also less effective for experienced or expert users. Users who frequently use the database would acquire successful search strategies. All of the subjects in our study in $C1$ who had previously used the database were able to find an address and four out of five were able to successfully complete the task in twenty minutes. Some expert users, like private investigators, would also have access to proprietary databases. We spoke to several private investigators in Pittsburgh who indicated that while they sometimes use county records, they also have access to proprietary databases.

These alternative means of accessing the data within the database, either by experts or using paid resources, go beyond the scope of the risk scenario addressed in this paper. Such scenarios are not directly addressed by removing search by name. These additional risk scenarios could potentially be addressed through policy changes that could limit the information available to paid resources or that could limit the data available to experts. However, in all these cases, policy makers would need to take into account the relevant tradeoffs between data accessibility and the potential for gaining additional privacy.

6.2 Usability and Accessibility

A thorough discussion of usability and suggestions for best practices for maintaining usability in public record databases are beyond the scope of this paper. However, as policy makers seek to increase the privacy of records, they also need to focus on ensuring that the accessibility intended by moving records online is maintained. Usability is important for maintaining the accessibility and transparency of the data within a given database. While it is useful to think of removing the search-by-name feature as creating a speed bump, or a feature that beneficially slows down users, usability issues can create "pot holes," or problems that slow down users without providing a benefit.

The Allegheny County case exemplifies the usability issues that policy makers must take into account. Although removing search by name was targeted at preventing users from accessing information for people they did not have a reason to search for, usability issues in the interface may prevent a large percentage of users from accessing a record even after they have access to an individual's address. Even after finding the target individuals' addresses, subjects in $C1$ referred to the interface as "difficult to navigate," "don't work," and "Lots of hard to use searches."

Counties implementing the removal of the search-by-name mechanism or similar privacy protections should focus on improving and maintaining interface usability. E-government literature includes guidelines for improving government websites that includes usability as a criteria. Some examples include Gant and Gant's evaluation of the state web portals [14], Dobrica's outline of recommendations for developing city websites [11], and Wang et al.'s task-based approach for evaluating government web services [33].

[2] Based on discussion with three Pittsburgh-area private investigation agencies

6.3 Legitimate Users

In the risk scenario addressed in this paper we assumed that any subject who could locate an address was a legitimate user, which seems to be the scenario assumed by policymakers. However, it is both unclear who a "legitimate user" of the database should be, whose access to the database should be guaranteed, and how to differentiate such users from community member attackers. The property assessment database is designed to be used to look up tax information, but it is also public record. By strictly defining the purpose of the database, it might be possible to define legitimate users based on required knowledge or by asking users to describe reasons for accessing data as is required on proprietary databases (according to three Pittsburgh-area private investigators with whom we spoke). However, such measures would reduce the database accessibility and might infringe on the privacy of the user, considerations that policymakers would have to carefully balance against the potential benefits of more carefully defining the risk scenario.

6.4 Future Threats

The insights from our user study suggest recommendations for the current scenario where, to the best of our knowledge, the content in property record databases is not yet directly accessible through search engines like Google. Policy makers should be mindful of the fact that search engines are rapidly indexing data on the Internet and that there is the possibility that, in the near future, such information will be available through popular search engines unless jurisdictions take actions to prevent or prohibit this. Google indicates that they may, in the future index such information [16]. In light of such a development current privacy mechanisms would likely have limited impact.

6.5 Conclusions

Our results show that removing search by name provided some privacy protection, especially for individuals whose addresses were not available on address-search websites. However, the amount of the protection provided for other individuals was roughly equivalent to the barrier presented by user interface issues. Thus, policy makers seeking to protect privacy while maintaining the benefits of transparency and accessibility should consider removing search by name, but should simultaneously put a strong emphasis on increasing the general usability of their databases for legitimate users. In addition, policy makers need to be aware that removing search by name provides only limited privacy benefits.

Acknowledgments

This work was supported by Microsoft Research, National Science Foundation grants # CNS-0756998 and # CNS-0831178, and the ARCS Foundation. The authors wish to thank Prof. Robert Strauss for the Allegheny County property database crawl and Prof. Estzer Hargittai and Gregory Norcie for assistance. The authors also wish to thank Prof. Anupam Datta, Prof. Helen Nissenbaum and Amanda Conley for their contribution and insights in conceptualizing the problem.

References

1. Allegheny County Assessment. `http://www2.county.allegheny.pa.us/RealEstate/Search.aspx`. [Last accessed 2011.06.20].
2. Intelius. `http://www.intelius.com/`. [Last accessed 2011.06.20].
3. Real Estate Website Subscription Services. `http://www.alleghenycounty.us/dcs/reweb.aspx`. [Last accessed 2011.06.20].
4. How to Obtain Public Records in Allegheny County and more. *Pittsburgh Post-Gazette*, January 01, 2002. Available at `http://www.post-gazette.com/FirstAmendment/OpenRecs.asp`.
5. A. Bepko. Public Availability or Practical Obscurity: The Debate Over Public Access to Court Records on the Internet. *New York Law School Law Review*, 49:967, 2004.

6. S. Bermann. Privacy and Access to Public Records in the Information Age. *Bepress Legal Series*, page 1303, 2006.
7. A. Blum, K. Ligett, and A. Roth. A Learning Theory Approach to Non-interactive Database Privacy. In *Proceedings of the 40th Annual ACM Symposium on Theory of Computing*, STOC '08, pages 609–618, New York, NY, USA, 2008. ACM.
8. A. Conley, A. Datta, H. Nissenbaum, and D. Sharma. Sustaining both Privacy and Open Justice in the Transition from Local to Online Access to Court Records: A Multidisciplinary Inquiry. *Maryland Law Review, Forthcoming*, 2012. A preliminary version of this article was published in Privacy Law Scholars Conference, 2011.
9. D. Bloys. Travis County Clerk Pulls Images Offline. *News For Public Officials*, June 28, 2006. Available at `http://www.davickservices.com/travis_county_pulls_records_offline.htm`.
10. D. Majors. Onorato Seeks Partial Web Site Blackout. *Pittsburgh Post-Gazette*, October 20, 2007. Available at `http://www.post-gazette.com/pg/07293/827036-85.stm`.
11. L. Dobrica. CONSIDERATIONS ABOUT CITIES WEBSITES EVALUATION. In *Proceedings of the IADIS International Conference on e-Society*, 2007.
12. C. Dwork. Differential privacy. *Automata, Languages and Programming*, pages 1–12, 2006.
13. C. Dwork. Differential privacy: A Survey of Results. In *Proceedings of the 5th international conference on Theory and applications of models of computation*, TAMC'08, pages 1–19, Berlin, Heidelberg, 2008. Springer-Verlag.
14. J. Gant and D. Gant. Web portal functionality and state government e-service. In *Proceedings of the 35th Annual Hawaii International Conference on System Sciences, 2002 (HICSS)* , pages 1627 – 1636, Jan. 2002.
15. B. Givens. Public Records on the Internet: the Privacy Dilemma. In *CFP '02: Proceedings of the 12th Annual Conference on Computers, Freedom and Privacy*, pages 1–7, New York, NY, USA, 2002. ACM.
16. Google.com. Tools for Public Sector Organizations-Frequently Asked Questions. `http://www.google.com/publicsector/content/faq.html#43`. [Last accessed 2011.06.20].
17. E. Hargittai. Beyond Logs and Surveys: In-Depth Measures of People's Web Use Skills. *J. Am. Soc. Inf. Sci. Technol.*, 53:1239–1244, December 2002.
18. J. Vellucci. Allegheny Council Votes to End Online Name Search of Property Records. *Pittsburgh Tribune-Review*, November 21, 2007. Available at `http://pittsburghlive.com/x/pittsburghtrib/news/cityregion/s_539028.html?source=rss\&feed=7`.
19. J. Vellucci. Tax records too Public for Onorato's Taste. *Pittsburgh Tribune-Review*, October 20, 2007. Available at `http://www.pittsburghlive.com/x/pittsburghtrib/news/s_533637.html`.
20. B. Jansen and U. Pooch. A Review of Web Searching Studies and a Framework for Future Research. *Journal of the American Society for Information Science and Technology*, 52(3):235–246, 2001.
21. J. Koenemann and N. J. Belkin. A Case for Interaction: a Study of Interactive Information Retrieval Behavior and Effectiveness. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Common Ground*, CHI '96, pages 205–212, New York, NY, USA, 1996. ACM.
22. J. Lee. Dirty laundry, online for all to see. *New York Times*, September 05, 2002. Available at `http://www.nytimes.com/2002/09/05/technology/dirty-laundry-online-for-all-to-see.html?pagewanted=1`.
23. N. Li and T. Li. t-Closeness: Privacy Beyond k-Anonymity and l-Diversity. In *Proc. of IEEE 23rd Int'l Conf. on Data Engineering (ICDE'07)*, 2007.
24. M. Douglas. Privacy Concerns. *Government Technology*, January 2, 2006. Available at `http://www.govtech.com/security/Privacy-Concerns.html?page=1`.
25. A. Machanavajjhala, D. Kifer, J. Abowd, J. Gehrke, and L. Vilhuber. Privacy: Theory meets Practice on the Map. In *Proceedings of the 2008 IEEE 24th International Conference on Data Engineering*, pages 277–286, Washington, DC, USA, 2008. IEEE Computer Society.
26. A. Machanavajjhala, D. Kifer, J. Gehrke, and M. Venkitasubramaniam. L-diversity: Privacy Beyond k-anonymity. *ACM Trans. Knowl. Discov. Data*, 1(1):3, 2007.
27. F. McSherry and K. Talwar. Mechanism Design via Differential Privacy. In *Proceedings of the 48th Annual IEEE Symposium on Foundations of Computer Science*, pages 94–103, Washington, DC, USA, 2007. IEEE Computer Society.
28. L. Morgan. Strengthening the Lock on the Bedroom Door: The Case Against Access to Divorce Court Records On Line. *Journal of the American Academy of Matrimonial Lawyers*, 17:45, 2001.
29. D. J. Solove. Access and Aggregation: Public Records, Privacy and the Constitution. *Minnesota Law Review*, 86:1137, 2001.
30. L. Sweeney. k-anonymity: A Model for Protecting Privacy. *Int. J. Uncertain. Fuzziness Knowl.-Based Syst.*, 10:557–570, October 2002.
31. U.S. Census Bureau. Population Estimates- Counties. `http://www.census.gov/popest/counties/`. [Last accessed 2011.06.20].

32. U.S. Census Bureau. 2011. OnTheMap Application. Longitudinal-Employer Household Dynamics Program. `http://lehdmap.did.census.gov/`. [Last accessed 2011.06.20].
33. L. Wang, S. Bretschneider, and J. Gant. Evaluating Web-Based E-Government Services with a Citizen-Centric Approach. In *Proceedings of the Proceedings of the 38th Annual Hawaii International Conference on System Sciences (HICSS'05) - Track 5 - Volume 05*, pages 129.2–, Washington, DC, USA, 2005. IEEE Computer Society.
34. P. A. Winn. Judicial Information Management in an Electronic Age: Old Standards, New Challenges. *Federal Courts Law Review, Forthcoming*, 2009.
35. X. Xiao and Y. Tao. M-invariance: Towards Privacy Preserving Re-publication of Dynamic Datasets. In *Proceedings of the 2007 ACM SIGMOD International Conference on Management of Data*, pages 689–700. ACM, 2007.

Appendix

Appendix A: Demographics

Table 1. Demographics of user study subjects

		C1		C2		Total	
		Count	%	Count	%	Count	%
Age	18-24	14	51.9%	14	56.0%	28	53.8%
	25-29	5	18.5%	5	20%	10	19.2%
	30-34	3	11.1%	4	16.0%	7	13.5%
	35-39	0	0%	1	4.0	1	1.9%
	40-44	2	7.4%	0	0%	2	3.8%
	45-49	1	3.7%	0	0%	1	1.9%
	56-60	2	7.4%	0	0%	2	3.8%
	70 or older	0	0%	1	4.0	1	1.9%
Gender	Female	11	40.7%	17	68.0%	28	53.8%
	Male	16	59.3%	8	32.0%	24	46.2%
Occupation	Administrative Support	2	7.4%	1	4.0%	3	5.8%
	Art, Writing & Journalist	1	3.7%	2	8.0%	3	5.8%
	Business Management & Financial	2	7.4%	2	8.0%	4	7.7%
	Not currently working	2	7.4%	2	8.0%	4	7.7%
	Student	13	48.1%	14	56.0%	27	51.9%
	Other	7	25.9%	4	16.0%	11	21.2%
PC/MAC user	PC	17	63.0%	15	60.0%	32	61.5%
	MAC	10	37.0%	10	40.0%	20	38.5%
Education	High school/ GED	0	0%	3	12.0%	3	5.8%
	Associate's degree	3	11.1%	1	4.0%	4	7.7%
	Some college	11	40.7%	7	28.0%	18	34.6%
	Bachelor's degree	9	33.3%	9	36.0%	18	34.6%
	Master's degree	3	11.1%	5	20.0%	8	15.4%
	I prefer not to answer	1	3.7%	0	0%	1	1.9%
Total		27	100.0%	25	100.0%	52	100.0%

Appendix B: Usability Issues

Table 2. Instances of usability issues encountered during user studies. Total* number of subjects indicates subjects who experienced at least one usability issue listed above (subjects may fall under one or more categories).

Issue		No. of Subjects		
		C1	C2	Total
Incorrect Municipality entered	Count	8	10	18
	%	66.7%	58.8%	62.1%
Street suffix entered in incorrect format	Count	10	13	23
	%	83.3%	76.5%	79.3%
Extra data provided in street field	Count	5	6	11
	%	41.7%	35.3%	37.9%
Other	Count	2	1	3
	%	16.7%	5.9%	10.3%
Total*	Count	12	17	29

Appendix C: Data from random sample of US counties

- 27 counties did not have property assessment database available online: Southeast Fairbanks Census Area (AK), Valdez-Cordova Census Area (AK), Hale (AL), Greenlee (AZ), Litchfield (CT), Nez Perce (ID), Marion (IL), Williamson (IL), Wilson (KS), Bristol (MA), Kennebec (ME), Penobscot (ME), Muskegon (MI), Pennington (MN), Caldwell (MO), Carroll (MS), Steele (ND), Franklin (NE), Guadalupe (NM), Crawford (PA), Kent (RI), Stanley (SD), Potter (SD), Piute (UT), Caledonia (VT), Windsor (VT), Lewis (WV)
- 3 counties had property assessment databases available online that were not publicly available: Boone (MO), Caddo (OK), Texas (OK)
- 71 counties had publicly available online property assessment databases for which we gathered data on search and record fields, summarized in the following table:

Fig. 3. Property assessment databases for random sample of US counties: * Indicates differentiated free and paid access, [1] Indicates search by street only, [2] Indicates database available online but not to general public

			Search fields			Form includes				
County	**State**	**Pop.**	**Name**	**Address**	**Parcel**	**Value**	**Sale price**	**Tax info**	**Photo**	**Floor plan**
Escambia	AL	37434	Yes	Yes	Yes	Yes	No	Yes	No	Yes
Clark	AR	23835	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
Conway	AR	20799	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
Pinal	AZ	340962	Yes	Yes	Yes	Yes	No	Yes	No	No
Butte	CA	220577	No	Yes	Yes	No	No	Yes	No	No
Ventura	CA	802983	No	Yes	Yes	No	No	Yes	No	No
Gilpin	CO	5604	Yes	Yes	Yes	Yes	Yes	Yes	No	No
Delta	CO	31322	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
New London	CT	266830	Yes	Yes	No	Yes	Yes	No	Yes	Yes
New Castle	DE	534634	No	Yes	Yes	Yes	Yes	Yes	No	No
Kent	DE	157741	Yes	Yes	No	Yes	Yes	Yes	No	No
Clay	FL	186756	Yes	Yes	Yes	Yes	Yes	No	No	Yes
Union	FL	14584	Yes	Yes	Yes	Yes	Yes	No	Yes	No
Warren	GA	5755	Yes	Yes	Yes	Yes	Yes	No	No	No
Barrow	GA	72158	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
Kauai	HI	64529	No	Yes	Yes	Yes	Yes	Yes	No	Yes
Honolulu	HI	907574	No	Yes	Yes	Yes	Yes	Yes	No	Yes
Cerro Gordo	IA	43609	Yes	Yes	Yes	Yes	No	Yes	No	No
Harrison	IA	15328	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
Ada	ID	384656	No	Yes	Yes	Yes	No	Yes	Yes	No
Randolph	IN	25696	Yes	Yes	Yes	Yes	Yes	No	No	No
Spencer	IN	20039	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes
Shawnee	KS	176255	Yes	Yes	Yes	Yes	No[2]	Yes	Yes	No
Allen Parish	LA	25636	No	Yes	No	Yes	No	Yes	No	No
Concordia Parish	LA	18989	No	Yes	No	Yes	No	Yes	No	No
Worcester	MA	803701	Yes	Yes	No	Yes	Yes	No	No	No
Baltimore	MD	789814	No	Yes	Yes	Yes	Yes	No	No	No
Howard	MD	281884	No	Yes	Yes	Yes	Yes	No	No	No
Yellow Medicine	MN	9867	Yes	Yes	Yes	Yes	No	Yes	No	No
Forrest	MS	81078	Yes	Yes	Yes	Yes	No	Yes	No	No
Musselshell	MT	4600	Yes	No	No	Yes	No	No	No	No
Teton	MT	6088	Yes	Yes	No	Yes	No	Yes	No	Yes
District of Columbia	N/A	599657	No	Yes	No	Yes	Yes	Yes	No	No
Stanly	NC	59794	Yes	Yes	No	Yes	Yes	No	No	Yes
Graham	NC	8001	Yes	Yes	Yes	Yes	Yes	No	No	Yes
McKenzie	ND	5799	Yes	Yes	Yes	Yes	No	Yes	No	No
Garfield	NE	1709	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Merrimack	NH	149071	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
Strafford	NH	123589	Yes	Yes	No	Yes	Yes	No	No	Yes
Ocean	NJ	573678	Yes	Yes	No	Yes	Yes	Yes	No	No
Union	NJ	526426	Yes	Yes	No	Yes	Yes	Yes	No	No
Grant	NM	29903	Yes	Yes	Yes	Yes	No	Yes	No	No
Carson City	NV	55176	Yes	Yes	Yes	Yes	Yes	Yes	No	No
Douglas	NV	45464	Yes	Yes	Yes	Yes	Yes	No	No	No
Schenectady	NY	152169	Yes	Yes	No	Yes	Yes	Yes	Yes	No
Tioga	NY	50064	Yes	Yes	No	Yes	Yes	Yes	Yes	No
Montgomery	OH	532562	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
Pickaway	OH	54734	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
Linn	OR	116584	Yes	Yes	No	Yes	No	No	No	No
Malheur	OR	30745	Yes	Yes	No	Yes	Yes	Yes	No	No
Wayne	PA	51337	Yes	Yes[1]	Yes	Yes	Yes	No	No	No
Providence	RI	627690	Yes	Yes	Yes	Yes	No	No	Yes	Yes
Edgefield	SC	25752	Yes	Yes	Yes	Yes	No	Yes	No	No
Lancaster	SC	77767	Yes	Yes	Yes	Yes	Yes	No	No	No
Moore	TN	6096	Yes	Yes	Yes	Yes	Yes	No	No	No
Pickett	TN	4783	Yes	Yes	Yes	Yes	Yes	No	No	No
Potter	TX	121816	Yes	Yes	Yes	Yes	No	Yes	No	No
El Paso	TX	751296	Yes	Yes	No	Yes	No	Yes	No	No
Weber	UT	231834	Yes	Yes	Yes	Yes	No	Yes	No	No
Chesterfield	VA	306670	No	Yes	Yes	Yes	Yes	Yes	No	No
Halifax	VA	35258	Yes	Yes	Yes	Yes	Yes	No	No	No
Cowlitz	WA	101966	No	Yes	Yes	Yes	No	Yes	Yes	No
Yakima	WA	239054	No	Yes	Yes	Yes	Yes	Yes	Yes	No
Dodge	WI	87335	No	Yes	Yes	Yes	No	Yes	No	No
Sauk	WI	58922	Yes	Yes	Yes	Yes	No	Yes	No	No
Wood	WV	86888	Yes	Yes	Yes	Yes	Yes	No	No	Yes
Campbell	WY	43967	Yes	Yes	Yes	Yes	No	Yes	No	No
Johnson	WY	8531	Yes	Yes	No	Yes	No	Yes	Yes	Yes
Pendleton*	KY	14887	Yes	Yes	Yes	No	No	No	No	No
Simpson*	KY	17019	Yes	Yes	Yes	No	No	No	No	No
Gratiot*	MI	41948	Yes	Yes	Yes	No	No	No	No	No