THE NEW SCHOOL
DIGITAL EQUITY LABORATORY

Preparing for the Digital Decennial Census

Building consent, equity, and safety into digital transition

About this report

The New School's Digital Equity Laboratory (DEL), a nonpartisan university center dedicated to advancing digital equity through applied research, convening power, and leadership development, is pleased to offer this analysis of risk, safety and trust for the first-ever digital decennial census. As advocates for structural equity, we believe the digital transition of the census is understudied and requires both attention and investment at national scale in order to achieve a level of success sufficient for our emerging era of digital governance.

*This report was co-authored by **Greta Byrum** and **Meghan McDermott** based on risk assessments and digital safety recommendations compiled by **Nasma Ahmed, Sarah Aoun, David Huerta, Sam Lavigne, Dhruv Mehrotra, Rebecca Ricks**, and **Norman Shamas**. This work was inspired by the critical work of the New York Counts 2020 campaign, which has built a powerful network of community advocates and organizers dedicated to ensuring the most robust possible count in 2020 while supporting their communities. We are grateful for the New York Counts 2020 Steering Community's valuable feedback in the process of compiling these recommendations, as well as to the New York Counts Tech & Tools Committee, which has advised and shaped this work every step of the way; and to the Open Society Foundations for their support. We are particularly grateful to Lauren Moore for her perspective and partnership, to Holly Dowell for her support and skill, and to Maya Wiley for her leadership.*

This work is licensed under a [Creative Commons Attribution-NonCommercial 3.0 Unported License](https://creativecommons.org/licenses/by-nc/3.0/).

Preparing for the first US digital decennial census

In 2020, the United States will hold its first digital census. The Constitution mandates that every ten years the US government must count every household in the country. The demographic and economic data collected is then used to decide $700-$900bn in federal funding for federal and state programs and to redraw electoral districts at the federal and state levels. This time, the Census Bureau has redesigned its process from end to end, integrating advanced statistical and geographic modeling and building a brand-new online data collection system.

Despite the persistence of a national digital divide—35% of US adults still do not have internet at home, including 53% of Latinx and 43% of Black adults, more than half those earning under median income, 42% of rural residents, and half of elders age 65+[1]—the Census Bureau is moving forward with its plan to ask 80% of households to complete the 2020 survey over the internet. Though the Bureau is offering additional options for completing the survey, this strategy prioritizes counting those who have internet access at home, for whom the process will be simple and quick. Yet because of the demographics of the digital divide, optimizing the count for the best-connected among us could lead to an overcount of affluent White populations and a systemic undercount of immigrants, people of color, and children, mirroring existing structural inequities. This in turn could lead to underrepresentation in government for unconnected populations, the loss of funding for critical social programs, and a skewed idea of who we are as a nation.

To offset the effect of the digital divide, the Bureau suggests that those without internet access at home fill out the online survey by going to a neighbor, a local organization, or finding public internet access points. Options range from barber shops and laundromats to libraries and post offices, raising a host of operational, digital security, liability concerns for individuals, businesses, and organizations; ***yet neither the federal government nor, in most cases, local governments are resourcing public-facing digital infrastructure at scale for census self-response***. This lack of core investment may further undermine the digital method for attaining a fair and accurate count. In the absence of coherent proactive messaging about the digital process and sufficient access and literacy support for digitally marginal and challenged

[1] Internet/Broadband Fact Sheet. (n.d.) *Pew Research Center*.

populations, the Digital Equity Laboratory (DEL) and its partners are deeply concerned about the success of the digital transition of this massive public mobilization.

In response, DEL with the NY Counts 2020 campaign's Tech & Tools Committee initiated a holistic risk assessment process to: 1) identify potential vulnerabilities of the digital census for individual participants, at-risk communities, and organizations and institutions acting as census sites; and 2) suggest mitigation strategies to protect the count while also protecting communities already disproportionately impacted by the digital divide and by predictive and surveillant technologies.[2] ***Our goal is to provide digital tactics and techniques to help prevent possible harms, and enable communities and agencies to better prepare against the uncertainties of a digital census and the likelihood of a resulting undercount. Our focus is at the user-level, and aims to address holistic safety concerns, not solely cybersecurity.*** The report that follows provides the best possible recommendations given a number of uncertainties about the Census Bureau's plans and systems, and takes a "power not paranoia"[3] approach to building capacity and awareness among community stakeholders, rather than ignoring or glossing over uncertainties and threats.

To prepare for this novel approach to census, with input from the NY Counts 2020 Tech & Tools Committee, the Digital Equity Laboratory commissioned a holistic risk assessment to facilitate safe and secure participation of census access sites. Conducted by digital security specialists,[4] the assessments focused on the interaction of human-centered systems with software, hardware, networks and data systems in order to generate recommendations and training curricula for CBOs, libraries, and other access points.

In the following sections, we first describe the mechanics of the digital census to clarify how the process will roll out and how digitization could affect participation and outcomes. Next, we turn to a series of risk and strategy clusters to examine: how the digital census could present challenges including an undercount due to uneven digital access and literacy rates, and how

[2] See Gangadharan, S.P. (2017) *The downside of digital inclusion: expectations and experiences of privacy and surveillance among marginal internet users.* New Media and Society, 19 (4); Eubanks, V. Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor. New York: St. Martin's Press, 2017; and Madden, M. "The Devastating Consequences of Being Poor in the Digital Age" (25 April 2019) *The New York Times.*

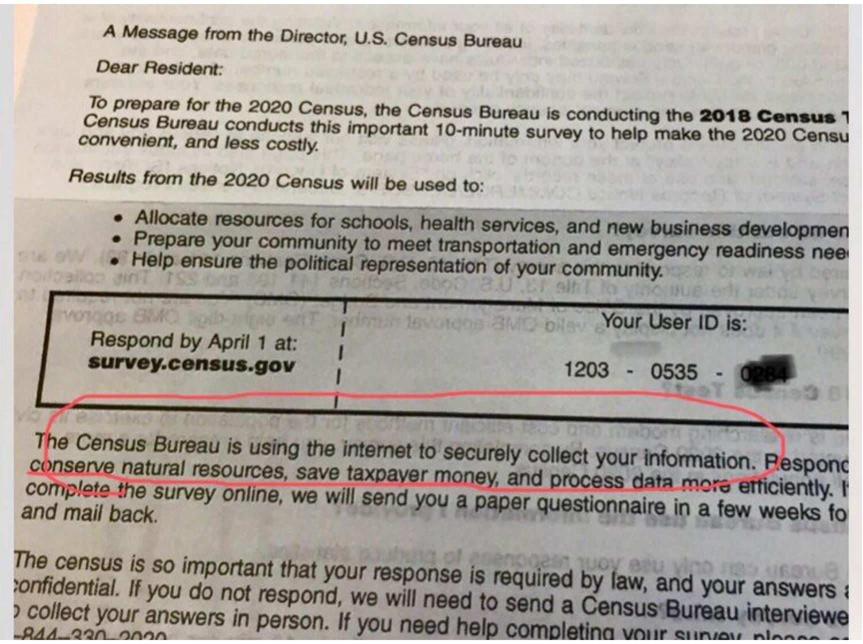
[3] Inspired by Stop LAPD Spying and Our Data Bodies' community-forward organizing approach to understanding and organizing proactively around data-driven harms and risks.

[4] Nasma Ahmed, Sarah Aoun, David Huerta, Sam Lavigne, Dhruv Mehrotra, Rebecca Ricks, and Norman Shamas.

access interventions such as public WiFi and public library access points could be set up for the best possible outcomes for digital safety; how to offset digital and cybersecurity threats and risks to organizations and individuals in the census process; and how the Census Bureau and civic institutions can interact to preserve the integrity of the count in a politicized environment.

We have an opportunity now to set a standard for digital government grounded in community safety and trust as e-government processes expand. Given the realities of a wholly new end-to-end automated and online-first digital data collection system rolling out in a hypercharged political environment—not to mention increasing concerns about online surveillance, targeting, and data theft—DEL and its partners are focused on building consent and public safety into the digital participation process wherever possible.

The Mechanics of the 2020 Digital Census



A Message from the Director, U.S. Census Bureau

Dear Resident:

To prepare for the 2020 Census, the Census Bureau is conducting the **2018 Census T** Census Bureau conducts this important 10-minute survey to help make the 2020 Censu convenient, and less costly.

Results from the 2020 Census will be used to:

- Allocate resources for schools, health services, and new business developmen
- Prepare your community to meet transportation and emergency readiness nee
- Help ensure the political representation of your community.

Respond by April 1 at: **survey.census.gov**	Your User ID is: 1203 - 0535 - [redacted]

The Census Bureau is using the internet to securely collect your information. Respond conserve natural resources, save taxpayer money, and process data more efficiently. I complete the survey online, we will send you a paper questionnaire in a few weeks fo and mail back.

The census is so important that your response is required by law, and your answers confidential. If you do not respond, we will need to send a Census Bureau intervie collect your answers in person. If you need help completing your survey, please -844-330-2020

Mailer distributed in Rhode Island in the 2018 end-to-end test of the census system

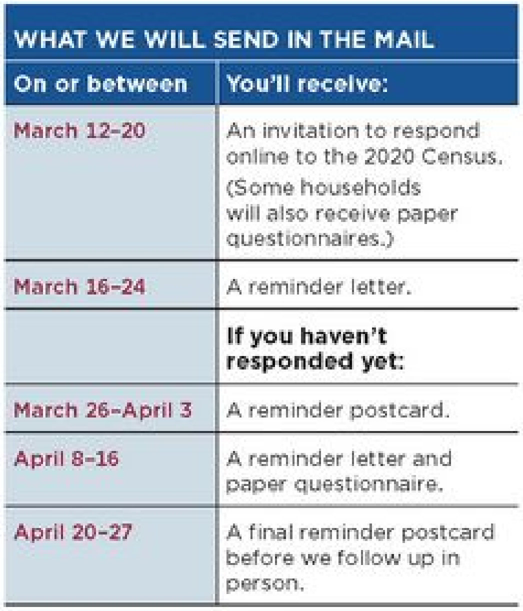
Digitization of the census has been driven by increasing costs: "from 1970 to 2010, the bureau's cost to count each household quintupled,[5] to $98 per household in 2010 dollars, according to the GAO."[6] The new digital process is intended to save the Bureau $5.2 billion compared to the cost of performing a traditional count using pencil-and-paper surveys and a comprehensive canvassing operation. The Census Bureau estimates that this plan will keep costs within range of the $12.3 billion price of the 2010 decennial census—allowing the census to meet its Congressional budget allowance.

In order to reduce the price tag, the Census Bureau is curtailing the costliest parts of the process—printing, postage, and the human labor cost of sending enumerators out into the field. The new digital process comprises machine learning models built using municipal address and

[5] Data Collection Operations Were Generally Completed as Planned, but Long-standing Challenges Suggest Need for Fundamental Reforms. (14 December 2010). *The United States Government Accountability Office.*

[6] D'Vera, Cohn. (24 February 2016). For 2020, Census Bureau plans to trade paper responses for digital ones. *Pew Research Center*.

buildings data; survey responses will be fed into these models. According to this plan, enumerators using mobile devices (not paper) will walk only 25% of the 11 million blocks they canvassed in 2010[7], and instead of just surveys, will also collect geotagged survey and environmental data. Meanwhile, the Bureau estimates that a little under half of the 80% of households invited to respond online first will do so,[8] saving millions in paper, printing, and postage. Advanced statistical modeling will then allow the Bureau to impute missing data by using AI tools to process and integrate digital imagery and other unspecified datasets.[9]



WHAT WE WILL SEND IN THE MAIL

On or between	You'll receive:
March 12-20	An invitation to respond online to the 2020 Census. (Some households will also receive paper questionnaires.)
March 16-24	A reminder letter.
	If you haven't responded yet:
March 26-April 3	A reminder postcard.
April 8-16	A reminder letter and paper questionnaire.
April 20-27	A final reminder postcard before we follow up in person.

In its messaging, the Bureau emphasizes that digital participation is a choice, not a requirement; 20% of households will receive paper surveys first, targeted to communities with low internet access and large older-adult populations, though it has not specified what data it is using to identify these groups. Those without internet access at home will have to seek out public internet access points to participate digitally or over an automated voice response (AVR) system. To take the online survey at a public access point, people can enter the unique access code from their mailer, or enter their home address and the nearest cross-street to pull up their household's unique survey. Paper questionnaires will also be mailed as a follow-up to digital-first households that do not respond online within about a month. Households may also call an automated voice response system to respond. Finally, if households still do not respond within six weeks, the Bureau will send census enumerators door-to-door to collect data using dedicated iPhone 8 devices.

Publicly available information about "digital choice" leaves key questions unanswered, however. For example, what happens if online response rates are lower than expected, and as a result, the costs of paper-and-pencil or canvassing operations exceed estimates? Will Congress release additional funding to close the gap—a politically fraught process that seems untenable within

[7] D'Vera, Cohn. (24 February 2016). For 2020, Census Bureau plans to trade paper responses for digital ones. *Pew Research Center.*
[8] Farmer, A. (n.d.) Digitizing the 2020 Census. *Brennan Center for Justice*.
[9] 2020 Census Operational Plan. (31 December 2018). *United States Census Bureau.*

the six-week response window? Budget shortfalls have already happened at a smaller scale:[10] the cost of the 2018 trial run in Providence, Rhode Island ballooned far beyond estimates, causing the Census Bureau to cancel all follow-up tests. This means the count process now features some key cybersecurity and rural data collection systems which were not complete at the time of the Rhode Island test and therefore remain untested in the field, heightening concerns around functionality: “the lack of comprehensive testing in remote locations presents a serious possibility that the system simply won’t work properly in areas that are on the wrong side of the digital divide.”[11]

In terms of digital equity, internet access is not the only concern. Research commissioned by the National Association of Latino Elected and Appointed Officials (NALEO) on the 2018 Providence end-to-end census test demonstrated that online response rates were more than three times higher among the general population than among Latinx residents: only 20% of Latinx participants participated digitally, as compared to 70% of the general population surveyed.[12] One of NALEO’s key recommendations is that the Census Bureau’s outreach and enumeration strategies must take into account Latinx residents’ preferences for responding to the questionnaire on paper or in person. Increasingly, advocates for Latinx and Black communities as well as rural and low-income groups are sounding a similar alarm: Arturo Vargas, NALEO’s executive director, recently stated that the Bureau needs a wake-up call:”’We have the Census Bureau continually telling us everything is on track,’ Vargas said. ‘No. Everything is not fine. The Census Bureau needs to be proceeding, understanding the real problems it is facing, and can't be sugarcoating what is happening throughout the country.’"[13]

This alarming disparity points to a real possibility that groups with lower digital access and internet adoption rates are far less likely to be counted using the primary method, online. And if the clock or the budget runs out before paper response or enumerator follow up can catch up, the Census Bureau has provided no public information on how it will ensure that the system can

[10] Laposky, I. (6 February 2019). The Challenge of America’s First Online Census. *Wired*.
[11] Laposky, I. (6 February 2019). The Challenge of America’s First Online Census. *Wired*.
[12] Escuerdo, K.A., Becerra, M., & Domenzain, G. (n.d.) The Last Chance To Get It Right: Implications of the 2018 Test of the Census for Latinos and the General Public. *NALEO Educational Fund*. This study is based on an independent survey of 20% of all who participated in the census test, and so did not include any non-respondents.
[13] Coleman, E. (22 May 2019). It’s Not Just the Citizenship Question—the Digital Divide Could Hurt the Count of Latinos by the Census. *Route Fifty.*

correct the over-counting of those who are likely to respond digitally first. Based on what we know from the Bureau's Operational Plan, the answer will likely be to use satellite imagery and other public and private datasets to feed AI models, rather than collecting empirical data in the field. Are we looking at the prototype of a predictive decennial census?

If so, we must take into account the deficits of machine learning; when flawed data sets are used to build models, the models then replicate and amplify the biases in the original data.[14] If budget or time limitations cause the 2020 statistical models to over-sample affluent White and other digitally-privileged groups, the integrity of the census is inherently at stake, not only in 2020, but for every future digital census. By prioritizing those with digital privilege in response, we can expect these same groups to be overrepresented in future census operations, in turn entrenching and deepening inequities in economic and political representation.

These issues of predictive modeling and digital privilege underline the importance of ensuring the most representative count possible in 2020. To build public-facing digital infrastructure that is up to the task, we offer this report as a way to outline comprehensive socio-technical risks facing the 2020 census, and strategies for creating a safer and more complete digital count.

Risks & Strategies

> *Ensuring that census information is kept secure, confidential, and private is a key piece of the census process, but the internet —a ubiquitous feature of our modern lives and a place where we regularly put private information— is a patchwork of communication protocols not well built for privacy or security. Census participants must have confidence that their personal information remains secure at every step. Without that confidence and trust, they may choose not to participate.*[15]

[14] Richardson, R., Shultz, J., & Crawford, K. (5 March 2019). Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice. *New York University Law Review Online, Forthcoming.*

[15] N. Ahmed, S. Aoun, B. Ricks, & N. Shamas. (n.d.) Risk Assessment of the 2020 U.S. Census: Recommendations for Action.

For the 2020 survey, the Census Bureau describes its own commitment to security and privacy in terms of data confidentiality (Title 13 of the US Code,[16] which all census workers must swear to uphold); cybersecurity measures taken to protect the bureau's servers from external attacks; provision of a encrypted portal for digital response; and privacy protection focused on anonymity of census results.[17] The Census Bureau is also relying on public perception of the Commerce Department and the Bureau itself as inherently trustworthy parties, which in the context of the court battle over the citizenship question,[18] is by no means a given. Furthermore, because digital self-response relies on a public-facing layer of internet access which is not provided or governed by the Census Bureau, there are multiple digital risks that cannot be addressed through increasing cybersecurity on the government side.

Building safer digital infrastructure for public access census participation serves two important goals: first, to drive up the count by creating a web of access for those who do not have home internet (35% of US adults, according to Pew[19]); and second, to decrease digital risks (data theft, digital targeting, or harassment, etc.) for participating organizations and institutions as well as individuals and households.

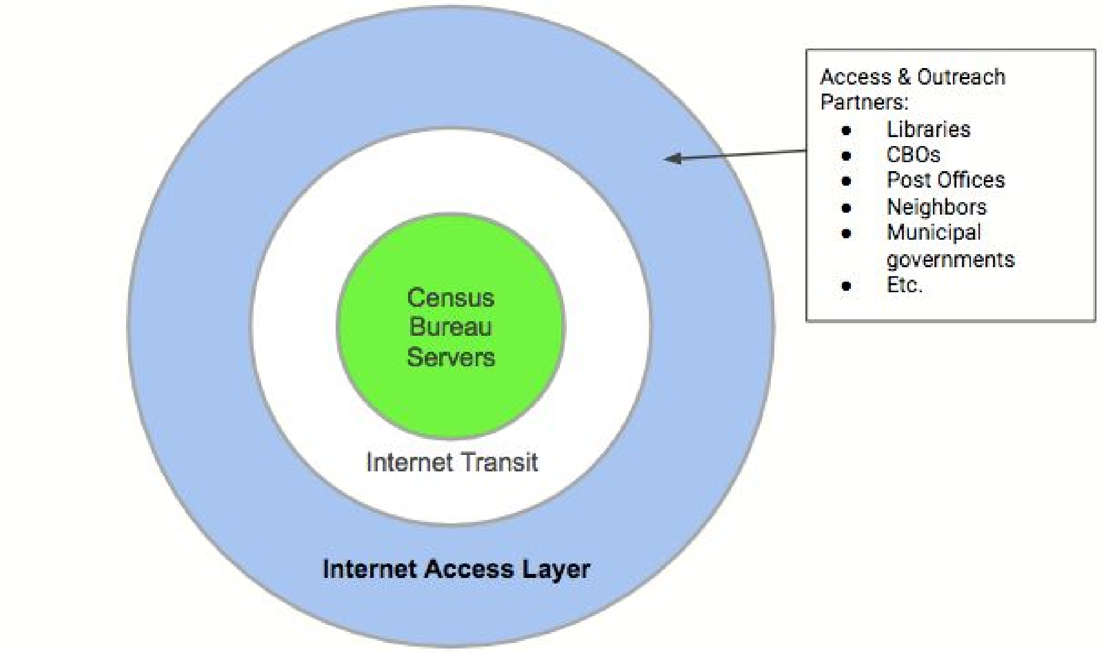
To assess and mitigate digital risks, DEL has asked the following questions:

- Are CBOs, libraries, and community anchors prepared to offer safe, secure internet access as well as digital literacy support for a public nervous about political targeting, hacking, surveillance, and data security?
- While hotspots, ad-hoc public internet access points, and get-out-the-count outreach apps are useful, could they create a data trail that endangers targeted populations, or invite phishing, harassment, or other physical or cyberattacks?
- Who shoulders the risk of digital harms or failures?

In partnership with the New York Counts 2020 Campaign, the New York Counts Tech & Tools Committee, and community digital security consultants,[20] the Digital Equity Laboratory has identified a series of interconnected digital risks based on qualitative and quantitative assessments of preparations for census (i.e., hardware, software, and human systems). In

[16] Title 13, U.S. Code. (n.d.) *United States Census Bureau*.
[17] 2020 Census Operational Plan. (31 December 2018). *United States Census Bureau.*
[18] Burke, G. & Bajak, F. Ahead of court ruling, Census Bureau seeks citizenship data. (7 March 2019). *Associated Press*.
[19] Internet/Broadband Fact Sheet. (n.d.) *Pew Research Center*.
[20] Nasma Ahmed, Sarah Aoun, David Huerta, Sam Lavigne, Dhruv Mehrotra, Becca Ricks, and Norman Shamas.

particular, we focused on the public-facing or “internet access layer” of the census ecosystem—that is, what happens at the point where the participant first goes online, before their data is encrypted in transit to the Census Bureau’s servers.



Utilizing a mix of methods,[21] as well as analysis of data flows of three cases (a census participation kiosk, a census enumerator, and a public WiFi hotspot), these digital security specialists surfaced varying levels of existing organizational preparedness and need here in New York State, as well as policy and census-specific digital and data security concerns at large. By examining the internet access layer—the public-facing infrastructure—of the census data-gathering system, consultants sought to identify what can be proactively and protectively addressed at the user-level, from simple “digital 101” tactical steps to larger-scale, immediate budgetary investments and policy advocacy recommendations specific to the 2020 census. The outcome of these assessments, further reviewed by public library professionals, technologists, and census advocates, is shared here. While we put particular focus on New York’s public library infrastructure as the go-to resource for safer public internet access, this document is adaptable by many types of organizations and agencies.

[21] See appendices for full reports, including descriptions of methods.

Following, we break down three emergent risk/opportunity clusters in the new digital census, and provide explicit steps that can be taken to increase digital safety and security of respondents ***before and at the point where they access the portal and submit data to the Census Bureau***. Our focus on this user-facing level of the process is designed as a supplement to the work done by the Census Bureau to shore up their systems and build encryption for data in transit—critical steps that nevertheless do not address security and safety issues in public internet access infrastructure such as hotspots, libraries, and social service providers. In providing recommendations for safety of individuals and organizations participating in the 2020 digital census, however, we do not mean to imply that all concerns can be addressed through information technology best practices. What we provide in the following sections are simply actionable steps, which nevertheless cannot address larger questions about data management, transparency, and accountability practices among government agencies amassing and handling sensitive data on vulnerable groups.

Every piece of research on the census process shows that trust is critical to getting out the count. Our intention is to provide libraries and civic organizations with the tools they need to be accountable stewards of their communities' trust. Census participation is a civic duty and is mandated by the Constitution, so the public should expect a baseline of care for their safety in this process. However, it is important also to emphasize that these recommendations will not have the full impact possible unless current piecemeal investments in public-facing infrastructure and operational support are increased, and unless public officials take a stance of transparency by pledging to share and mitigate digital risks.

Beyond the 2020 census, digital safety will continue to be an issue going forward as more government processes move online. By building digital preparedness—being explicit about the potential threats and the steps required to mitigate them—we can build informed, *consentful* participation in the census among those who may be vulnerable in the current political climate, hard to count, offline, and targeted. By providing a concrete plan for scaling technology capacity at participating organizations and sites, we can support not only a better overall outcome in 2020 but also investment and training in digital capacity for more fair and just digital governance processes and systems ahead. In this sense, the digital census offers organizations, institutions, and local governments an opportunity to set digital safety standards—instead of eliding or ignoring risks due to a fear of what would happen were the public informed.

Risk/Strategy Clusters

I. The digital census and the digital divide

As described above, digitally unconnected groups run the risk of not being counted due to lack of internet access and/or digital literacy.[22] In New York State, as elsewhere in the US, groups most at risk of facing digital divide challenges—in this case both access and literacy—are also likely to belong to “hard-to-count” (HtC) communities. This includes primarily rural and low-income urban residents, those in transient living situations, people of color, first nations peoples, and immigrants.

In upstate New York, the state comptroller’s office estimates that about 22% of residents —almost 4.4 million people— are immigrants, many living in urban areas with low internet adoption rates. Cities including Utica, Schenectady, Rochester, Buffalo and Albany are home to growing immigrant populations, and New York State is third in the nation as a destination for refugees. In terms of internet availability and adoption, smaller and poorer cities lag behind New York City, where 31% of households do not have a home broadband subscription (less than half of NYC’s lowest-income households have broadband at home). Race, age, disability status, employment status, language, and other demographics factor into rates of subscription. For instance, New Yorkers who are older than 65 are three times more likely than other groups to have no home internet. Black and Hispanic New Yorkers also lack home broadband at rates that are ~10% higher than their White and Asian counterparts.[23]

The following maps, derived from the Census Bureau’s American Community Survey data on internet adoption and compiled by the National Digital Inclusion Alliance,[24] show common geographic patterns of home internet adoption. Poorer urban neighborhoods, home primarily to communities of color, elderly, and immigrant households, show low rates of broadband adoption due to the cost of internet subscription services. Rural areas show low rates of

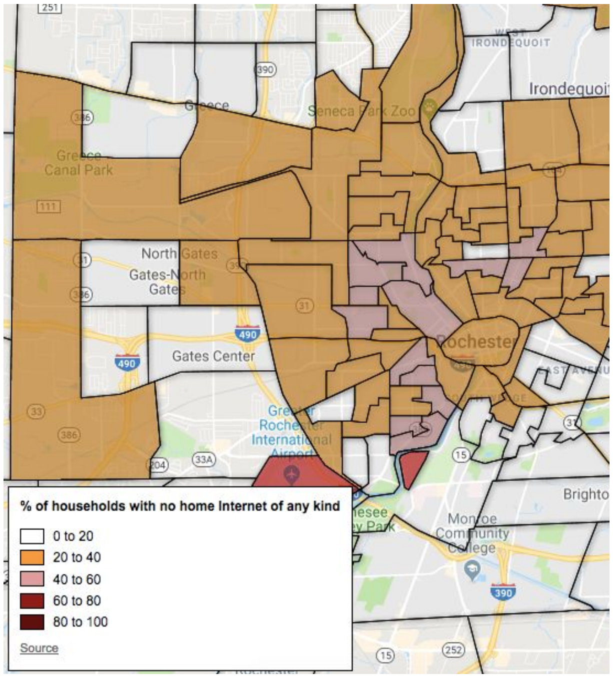
[22] Per the Pew Research Center, 53% of Latinx and 43% of Black adults, more than half those earning under median income, 42% of rural residents, and half of elders age 65+ in the US, did not have home internet access in 2018.

[23] Truth in Broadband: Access and Connectivity in New York City. (April 2018). NYC Mayor’s Office of the Chief Technology Officer.

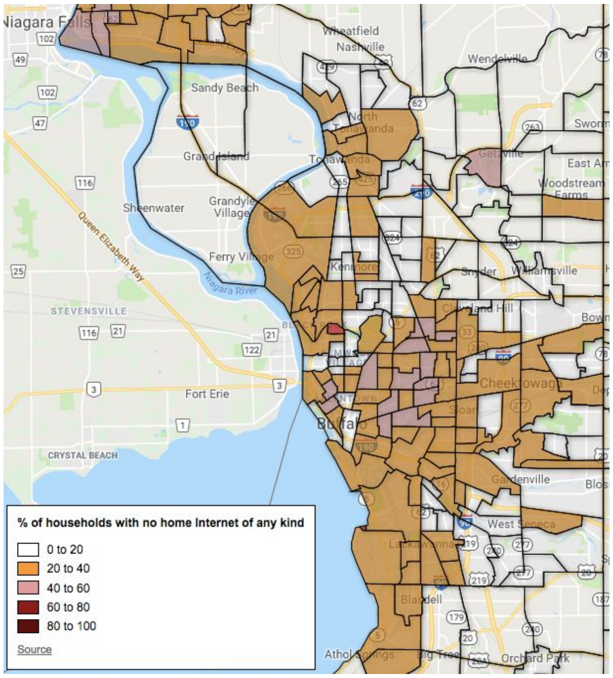
[24] U.S. Census Bureau Releases 2013-2017 ACS 5-Year Estimates. (6 December 2018). *United States Census Bureau.*

broadband adoption, both due to cost of access and to lack of industry buildout to areas with lower projected rates of return on investment.

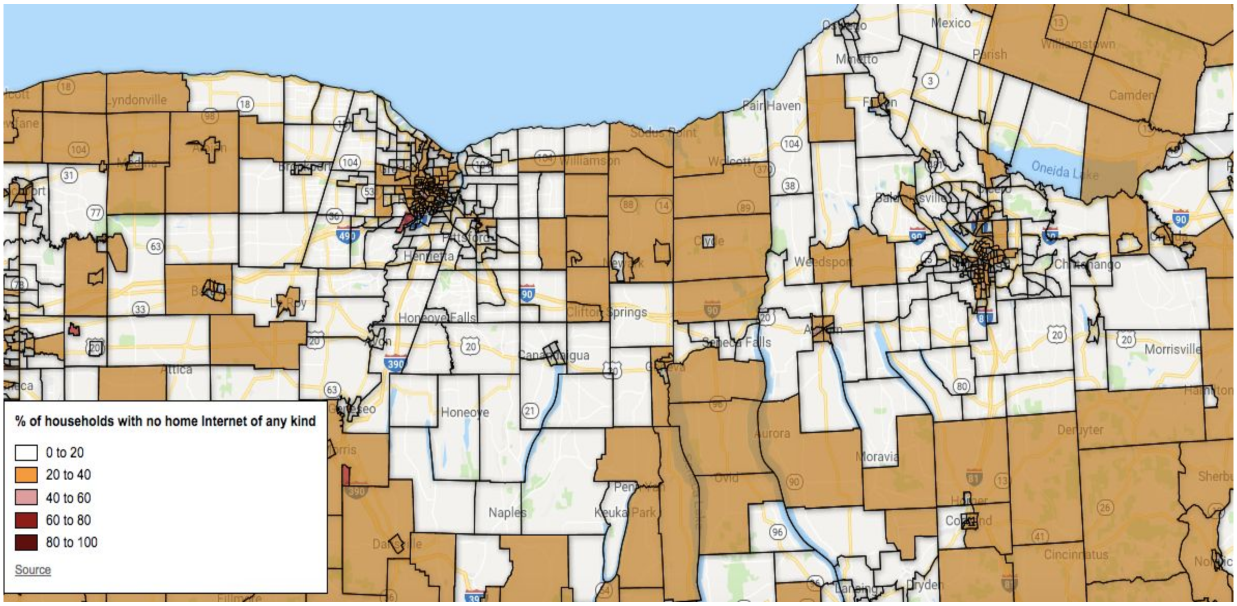
Rochester



Buffalo



Rochester/Syracuse



All of this has direct implications for a digital census. In 2010, for example, an undercount in Brooklyn/Queens led to a loss of $50 million in funding for NYC and a Congressional seat.[25] Looking ahead, New York is poised to lose two more Congressional seats according to standard projections for the 2020 Census.[26] Combined with a lack of digital access and readiness, the damage to public, civic, economic, political and electoral power could be significant.

To address the threat of an undercount due to the digital transition, the Census Bureau, city agencies, community organizations, and advocates are expected to turn to public libraries and public WiFi hotspots for digital census access infrastructure. While both provide a baseline of access, there are differences in terms of available built-in support and safety. There is a third option for public-facing access: organizations' or communities' own computers; however, **in this section, we will primarily provide recommendations to improve the the dynamics of high-volume public access library and public WiFi sites, not IT or cybersecurity best practices.** The following section on safety and cybersecurity will provide more information for organizations' and personal shared internet access points, as well as specifications and detailed recommendations for set up of public access points and IT best practices for digitally-enabled canvassing and participation.

Public WiFi (LinkNYC, mobile internet vans, public WiFi hotspots/corridors, etc.)

Those without home internet subscriptions may choose to connect to public WiFi hotspots such as LinkNYC to fill out their surveys. While this may be a viable option for some, agencies and advocates should be aware of risks in order to design interfaces and advise constituents accordingly.

First, availability of public internet access at hotspot sites does not address digital literacy support needs. For example, the process of typing a URL into a browser bar on a smartphone is a process that requires a level of digital literacy that may be a high bar for some smartphone-dependent users. And even if users are able to access the survey and enter their address and nearest cross-street, those using mobile devices without support at a public hotspot are vulnerable to being observed or targeted for misinformation, harassment, or

[25] Associated Press. (22 May 2012). 2010 census missed 1.5 million minorities. *CBS News*.
[26] Brace, K. (19 December 2018). Arizona Gains Rhode Island's Seat With New 2018 Census Estimates; But Greater Change Likely by 2020. *Election Data Services*.

intervention. If the Census Bureau releases a mobile app for census participation (one way to reduce digital literacy challenges), new digital literacy challenges could arise based around ensuring access to *only* the trusted/official Census Bureau app, not third party or imposter apps that could collect sensitive data.

Additionally, many public WiFi systems provided in partnership with private-sector companies hold third-party data-sharing agreements to generate revenue through targeted marketing. For example, a public WiFi network could collect information about users' devices (for example, unique MAC ID) that could be compared with these WiFi providers' logs of registered users or other datasets to create a record of physical location.[27] Suppliers of public WiFi could also have data-sharing agreements with advertising partners that could reveal the time and place where an individual accessed the census portal, along with other activities. For anyone sensitive to concerns about state or corporate surveillance, public-private WiFi hotspots may hold more personal risk than is comfortable for participation in a mandated civic process.

Finally, malicious actors have the ability to create imposter hotspots that trick participants into entering personal information into fake census forms, or set up network infiltration at public hotspots that could redirect data entry intended for census participation to malicious network operators. The Census Bureau's survey portal is encrypted using HTTPS, so anyone who accesses the genuine census website should be able to enter information that is protected in transit; however, fake or imposter Census Bureau sites or even hotspots showing seemingly real portals are possible. This is something that the Census Bureau is actively working to mitigate through, e.g. purchasing domain names that mimic or resemble the official portal URL, and the best defenses lie with their implementation of the census participation site; however, staff and volunteers supporting census participants will need clear guidance on how to message to the public around recognizing and distinguishing between the genuine census site and potential imposters.

For agencies or organizations offering informed consent to participate in the digital census via public WiFi access, we recommend:

- Full disclosure regarding data-sharing agreements that could impact how and where participants' information may be disclosed. This disclosure should include not only data

[27] Kofman, A. (8 September 2018). Are New York's Free LinkNYC Internet Kiosks Tracking Your Movements? *The Intercept*.

collection related to the census process itself, but also sharing of metadata (data about when, where, how, and with what device users access the WiFi system) and log retention and use;

- Accessible, clear data management protocols and commitments, e.g., who will have access to user data, how long will it be kept and where will it be stored, what kinds of metadata will be collected, etc.
- Network monitoring and security to ensure network level domain name resolution is not compromised to redirect to malicious census sites
- Recognizable physical branding and continuous in-person monitoring at public-facing WiFi sites to discourage malicious imposter networks or websites;
- Incident response plans for physical harassment or targeting at hotspot sites in addition to standard digital incident response plans;
- For participants who have concerns or digital literacy or accessibility support needs, reference to public library sites where support and confidentiality are more readily available.

Public Libraries

In terms of both digital equity and safety, public libraries are preferable to public WiFi for digital census participation, especially for marginalized and vulnerable populations. Moreover, the more people generally that participate in the census at public libraries, the more safety in numbers and higher likelihood of anonymity there will be at these sites for particularly vulnerable populations.

Along with being trusted institutions, libraries have confidentiality policies, frontline expertise, and IT/tech infrastructure in place to provide internet access for unconnected patrons engaging in sensitive e-governance processes such as taxes and health benefits, among many other digital services. Librarians can often be found assisting people with basic digital skills, from how to use a mouse and opening a browser window to navigating a website and distinguishing between legitimate and malicious information requests.

However, given increased online privacy and security risks for the 2020 census, the level of support that librarians will be expected to provide, including language assistance and disability access in addition to providing information, support, and security protection, poses a significant challenge for these critical institutions’ capacity. Interviews with libraries and community-based organizations reveal a need for civic engagement and multilingual local volunteers who can support people in completing the census.[28] To date, despite advocacy efforts at the City and state levels, public libraries have not received the resources or capacity necessary to ensure a fair and accurate census count.

This lack of investment compounds existing internet access and digital skills inequities. For example, to even get online at a library, people need a library card. Once in, there may be long waits for workstations, with individual patron sessions limited to the specific duration of that library's session management system (sometimes as little as 30 minutes). For new and marginal internet users, just signing on and navigating to the correct census site may take several minutes. For elders and those with accessibility needs, getting to the right form and completing it within a session window may be cumbersome. And while publicly available information from the Census Bureau suggests that if a survey session times out or is closed in process, information will be lost and that participant will need to start over, potential information saved in autofill windows pose additional threats.

In terms of basic digital literacy, patrons might find that without support or guidance, they end up on malicious sites. The difference between .COM and .GOV has incredible implications for protecting users against hack or fraud, for example. Computers infected with malware could also generate pop-up windows that ask for information just like what the census requests. Finally, for providers supporting respondents with limited access and digital literacies, it is unclear how US Code Title 13 (confidentiality of the census) applies to respondent information, or how much assistance volunteers or staff who are *not* census workers and therefore not sworn to uphold Title 13 are allowed to provide, or if there is any risk to them if they make an error or give the wrong kind of guidance.

We recommend that public libraries commit to implement IT best practices for safety and cybersecurity, as outlined in the following section on digital safety and security. Additionally, in order to ensure that public libraries and their staff are safe and protected in the process, we advise:

[28] N. Ahmed, S. Aoun, B. Ricks, & N. Shamas. (n.d.) Risk Assessment of the 2020 U.S. Census: Recommendations for Action.

- Resourcing libraries at levels appropriate to the ask placed upon them to supply a complex range of community needs from accessibility and support to protection and confidentiality;
- Providing guidance and transparency as requested by libraries in order to offset institutional liability and risk as they offer a critical infrastructure and community support need, and to support their planning and implementation needs and capacity;
- Supported and resourced coordination between libraries and civic organizations to address language and information support needs.

II. Digital safety and security for organizations and individuals

The Census Bureau's encrypted HTTPS portal protects data in transit to the Bureau's servers, encrypting it as soon as it enters the portal. However, metadata (eg. data about the time, duration, and nature of a digital activity) is collected at several points in every digital process separate from the data that is actually entered into the survey itself. For example, any internet-based activity creates an activity log at a minimum on the device, the browser, the network, and servers along the way on the internet. These logs hold important data that may be cross-referenced with other data sets to create a data trail that could be used to identify individuals or communities. Data theft or misuse at the point of access is not protected by the Bureau's cybersecurity and privacy measures, which pertain to its own systems and servers and not to public-facing devices and networks.

Organizations serving communities who are targets of harassment, intimidation, or threats may also experience cyberthreats when offering public digital access support.[29] Actions taken by politically motivated individuals or organizations to suppress the count or simply to target vulnerable communities and the organizations who support them could include network infiltration, disruption, or deception; data theft; or surveillance using physical or software devices (keylogging software/hardware; other USB-delivered malware; physical sensors or trackers). Attacks aimed at organizations serving vulnerable or targeted populations, such as network infiltration and subsequent data theft, would impact not only data provided through census participation, but rather the internal systems and files of the organization itself. Simply

[29] Koenig, R. (12 January 2016) How Social-Justice Nonprofits Can Defend Against Public-Relations Attacks and More. Chronicle of Philanthropy.

providing public-facing internet access for census participation could open the organization up to cyberattack, particularly if accounts, networks, and devices are not set up with privacy and safety protections. Additionally, organizations and institutions providing access may also be targeted for physical infiltration or harassment.

Data theft, misuse, or non-consentful sharing could also take place in many ways in other census-related processes. In the process of canvassing, well-meaning CBOs or advocacy orgs could collect more identifiable information on vulnerable constituents than necessary. If they are using proprietary apps, platforms, devices, or systems in the process, they could also unwittingly give up their constituents' data to third parties without consent.[30] In the absence of secure data management protocols and clear data-sharing limitations, such data sitting on organizations' and private-sector partners' servers and networks could invite infiltration and data theft. Finally, companies or partners offering canvassing or access devices may not have good data management practices, putting any data left on these devices at risk after devices are returned; third party vendors could also have a business model that depends on revenue from sales of data collected through, e.g., canvassing to collect personally identifiable information (PII) such as names, birthdates, phone numbers, etc.

One possible major threat pertaining to malicious digital activity in the census is phishing, or tricking people into divulging data about themselves. As discussed above in the section on public WiFi risks, the Bureau's HTTPS portal does not protect people from entering data into (very real looking) fake or imposter survey portals, or malicious digital activity echoed or reinforced with fraudulent mailers, posters, or other misinformation. The Census Bureau has attempted to address this threat by purchasing hundreds of URLs similar to that of their portal; however, they acknowledge that external threats like phishing or corrupted devices or network are outside of their control,[31] so ***it is important for community or public organizations to understand that based on the current operational plan they or their constituents are shouldering these risks and must prepare for them.***

To build a baseline standard for safety and security for those interacting with census systems, we recommend a suite of best practices related to device, network, account, and physical security. We also recommend that any organization, institution, or individual providing or

[30] This risk depends on the service and contract; for example, using a product with an enterprise level contract should protect against non-consentful sharing outside of that vendor's systems.
[31] Smith, K. (n.d.). Update on cybersecurity. *United States Census Bureau.*

supporting digital census participation adopt clear and transparent data management protocols as well as incident response plans and ongoing process monitoring protocols for participant safety as well as documentation of IT and interface functionality (glitches, system outages, or user difficulty or challenges).

Appendix II (Digital Census IT Best Practices TK) provides detailed information and prioritization of digital safety best practices. As an overview, below we provide ranked recommendations for creating safer digital census interface points.

Safest scenario

- Institution (such as public library) with staff experienced in providing digital access and literacy support offering dedicated census computers/devices that are configured following recommendations and best practices for device, network, and account security, as follows. These workstations should only be used by people participating in census data collection and should be restricted to the secure (HTTPS) census portal, with session and data management protocols. At libraries, these should be available to the public, not just to those who hold library cards.
    - **Suggested specifications for secure dedicated census portal**
        - Recommended device: dedicated tablet or Chromebook with large-format interface configured to provide a strong security model:
            - Device only uses a secure, up-to-date browser;
            - Separation of sessions and apps (Chromebooks should be set up for use in guest mode).
        - Standardized settings and device configurations:
            - Factory reset of the device prior to being used for census data collection to prevent any unknown or malicious software being retained, sensitive documents accessible to participants, or sensitive autofill information retained in forms;
            - Device only able to access the official census site. This will help ensure that census participants aren't tricked into entering sensitive information in an imposter site.
        - Network configuration:
            - Configured to block known malicious URLs and IP addresses;

        - If this cannot be implemented at a network level, we recommended setting up the device with a trusted DNS that blocks known malicious URLs and IP addresses;
        - Limitation on network analytics collected during census data process to only what is necessary for network health.
    - Session management:
        - To prevent accumulation of unneeded and potentially identifiable metadata in network and device logs, each user session should only be used for census participation and not additional browsing;
        - Devices should have session management systems built in to automatically refresh after each user (see session management software recommendations in Appendix II).
        - Session management software should be configured to remove any user data created, including cookies, browsing history, and any saved password/information.
    - Access point administration:
        - Vetted staff available to answer questions or provide digital literacy support informed and supported by clear guidelines;
        - No sign-in or log-in required to use dedicated participation device to limit the accumulation of metadata and personally identifiable information (PII) that could create datasets that could re-identify anonymized census participants;
        - Device should have physical ports (e.g., USB drives, lightning ports) covered and inaccessible to the public to prevent installation of malware such as keylogging software;
        - Readily available, comprehensive data management protocol, including:
            - External-facing commitment to not collect unnecessary data; to delete any metadata collected in the process of participation; disclosure of third-party data sharing agreements or (prefered) commitment to not sharing data with third-parties;
            - If devices or network connections are provided by third-party partners or vendors, disclosure of third-party

data sharing agreements or (prefered) commitment to not sharing data with third-parties.

        - A plan to monitor the devices and be prepared for incident response in the case of physical or online harassment, cyberattacks, or discovery of infiltration or malware.
        - Accessible, clear data management protocols and commitments, e.g., who will have access to user data, how long will it be kept and where will it be stored, what kinds of metadata will be collected, etc.
        - Redundant security layer that follows a “defense-in-depth” approach of building security controls throughout the system instead of just at a single location (e.g., put measures in place to protect organization or institution’s core network and digital assets). See Appendix II (TK) for detailed device, network, and account specifications.

Less safe - but necessary - scenarios:

- Recognizable, branded public WiFi sites administered and monitored by trusted entities with clear, available data sharing protocols and with dedicated devices configured according to best practices (as described above) as well as incident response plans and ongoing monitoring protocols;
- Existing but updated and reconfigured organizational or library computers or devices, configured according to best practices for data, device, network, and account security (as described above) as well as incident response plans and ongoing monitoring protocols and browser with automated updates.
- Recommendation for participants to separate census participation from other online activities

Going forward, DEL will develop a suite of tools and step-by-step instructions to build recommended systems and protocols, as well as trainings for libraries and other organizations seeking to build out implementation systems (see Appendix I for more information).

III. Census Bureau systems and integrity of the 2020 census

In May 2019, the Government Accountability Office identified 500 cybersecurity flags on digital census systems. The GAO also noted that half of those were labeled “high risk” or “very high risk;” 70 had been delayed for more than 60 days from previous correction deadlines; 32 systems may need to be reauthorized ahead of the census; and six systems currently do not have authorization to operate.[32] The decennial census is currently without a director of decennial census information technology.[33]

The last time the US undertook a major shift to a digitized process at national scale was with healthcare.gov, the ill-starred rollout of the Affordable Care Act.[34] Many remember what happened then: systems were untested, and some failed when the website launched, causing a cascading failure with operational and political consequences. And many eligible people were not able to register for plans simply due to lack of digital and broadband access and digital literacy. Eventually public libraries became a leading site for consumers to find support in navigate purchasing healthcare coverage online.[35]

Adding to concerns about the readiness of the Bureau’s brand-new IT systems, the citizenship question proposed at the behest of the Trump Administration[36] has created an atmosphere of mistrust that threatens the integrity of 2020 census data, even according to the Bureau’s own statisticians.[37] Yet regardless of what happens with political debates and judicial decisions on this topic, DEL and its partners believe the digital transition itself, and the prioritization of digital self-response in the process, could equally undermine the integrity of data collected. When participation is much more difficult for some groups than for others, trust in the integrity of the count and its outcomes suffers. In interviews with CBOs and public libraries, risk assessors confirmed deep concern for protecting privacy and safety, especially for vulnerable and politically targeted communities.

[32] Vincent, B. (2 May 2019). GAO Flags New Cybersecurity Issues for Upcoming Census. *Nextgov.*
[33] Brown, N. (10 May 2019). Key technology boss at U.S. Census Bureau switches roles. *Reuters*.
[34] Goldstein, A. (23 February 2016). HHS failed to heed many warnings that HealthCare.gov was in trouble. *Washington Post*.
[35] Deutsch, L. (10 October 2018). Public Library Association launches health insurance enrollment initiative. *American Library Association*.
[36] Wines, M. (4 November 2018). Inside the Trump Administration’s Fight to Add a Citizenship Question to the Census. *New York Times*.
[37] Honan, E. & Bahrampour, T. (5 November 2018). Statistics expert testifies census citizenship question would harm count. *Washington Post*.

A lack of publicly available information about key parts of the process also threatens the integrity of the 2020 census. The Bureau has not released key information that would assist those providing critical public access infrastructure, for example:

- Public internet access providers have not been able to run tests on how the user interface will work on different browsers, whether browser extensions such as virus and privacy protection will interoperate, or explore how the survey will run on different kinds of devices.
- Public access providers and census partners have not had an opportunity to interact with the user interface to prepare for glitches, navigation challenges, or confusing queries.
- Public access providers do not know how much assistance volunteers (e.g. organization or library staff) who are not sworn to uphold Title 13 may legally provide. For example, can they help census participants by physically entering data or operating machines?
- While many municipalities are planning to provide WiFi access for digital participation via mobile devices, it is unclear how the public will be able to navigate to the survey on smartphones except by typing the URL into a browser window, which may be unfamiliar or difficult for many mobile-only users.

Without better information, census partners will not be able to build the educational, outreach, and guidance materials they will need to assist respondents. Beyond unanswered questions, however, there is a deeper concern about the motivations and intentions of Census Bureau appointees who may take actions that could threaten the safety of particular communities. In such a case the civil rights community must act to protect the credibility and and trust placed in our public and civil institutions, including libraries.

The Census Bureau is ultimately responsible for anonymity within the census results. Due to a lack of transparency of the process, however—including a lack of information about data management agreements among vendors and about contingency plans in the case of system failures or breaches—we cannot determine whether the Bureau will follow Title 13 as it has been previously interpreted. It is important to note that the lack of transparency from the Census Bureau means we are unable to evaluate its processes and protocols. Clearer information would help to ensure that recommendations made here can be well implemented and have the intended impact: of user safety and increased trust. Therefore, once more information is available, preparations will need to be compared against the system's restraints, requirements, and capabilities.

Due to these uncertainties and concerns, we recommend that civic institutions once again develop a range of incident response plans, monitor and document cases of system outages or glitches as well as user difficulty with particular interfaces or queries; share these with partners, committees, and coalitions; and organize with complete count committees, library networks, or other networks to advocate for resourcing in the case of a call for capacity from CBOs and libraries in a crisis.

Trust, risk, and critical infrastructure

The decennial census is a vector of political and social engineering, from its critical operational importance in determining the function and priorities of American democracy to its cultural role in defining who we are as a country. It also carries the historical weight of ethnic targeting. From 1942-1945, census data was secretly and illegally used to identify ethnic Japanese residents—the majority of whom were citizens—and confine them in military internment camps[38] despite clear indication on the survey form (then as now) that:

> *Only sworn census employees will see your statements. Data collected will be used solely for preparing statistical information concerning the Nation's population, resources, and business activities. Your Census Reports Cannot Be Used for Purposes of Taxation, Regulation, or Investigation (based on US Code, Title 13).*

More recently, after 9/11, the Census Bureau shared tabulated details—including zip code and country of origin data—about Arab-Americans with the Department of Homeland Security. While this sharing was technically legal, the act breached public trust and garnered disapproval from civil liberties organizations who compared the situation to the 1940s scenario with Japanese residents.[39]

Given that we are at a national low point of trust in public institutions, the conversation about digital risk and security is about treating community trust with the care it deserves. No private or state actor should take for granted the trust that community advocates and public libraries hold. As our civic institutions take on the responsibility and risk of building safer digital participation systems, the government should also meet threshold requirements for transparency, risk management, and safety in order to expect the participation of community groups and institutions. Similarly, third party vendors and contractors working to support digital canvassing, census data collection, and data analysis must likewise follow higher standards for transparency and security of this sensitive public data.

The overwhelming public focus on the high-stakes citizenship question has taken priority in our national conversation about the census, and as a result we are behind in addressing safety and

[38] Aratani, L. (6 April 2018). Secret use of census info helped send Japanese Americans to internment camps in WWII. *Washington Post*.

[39] Clemetson, L. (30 July 2004). Homeland Security Given Data on Arab-Americans. *New York Times*.

functionality concerns surrounding the digital transition. We must build an understanding of the demands and dangers of these systems, in order to set up rational and operational systems to enable the best possible count while keeping our communities safe. This report suggests ways and means to build safer public-facing internet access and digital tools for a more reliable count, but does not examine several other issues surrounding the Census Bureau's systems and cybersecurity and data management protocols, including design and procurement of AI systems, the creation of geotagged nonresponse address lists that could be connected to environmental and personally identifiable data via enumerator devices, and the reported data sharing agreement between DHS and the Census Bureau.[40]

The stakes of digital transition of the largest peacetime mobilization in US government have never been higher: "'We' are only 'the people' if we are counted. This is why the founding fathers enshrined the census in the Constitution."[41] This is especially true as we consider the possibility that overcounting digitally connected and resourced populations could present a flawed and inaccurate picture of the country—one that leaves the already digitally marginalized even further behind. And in an age of machine learning and advanced statistical modeling, this could lay the groundwork for a predictive census that continually deepens structural inequities. For all of these reasons, now is the time to deeply consider the consequences of the digital transition and to shore up the capacity of civil society—especially vulnerable groups and those with protected status—to participate in digital governance.

[40] Burke, G. and Bajak (7 March 2019). Ahead of court ruling, Census Bureau seeks citizenship data. *Associated Press*

[41] Wiley, M. (5 March 2018). We can't count on the Census Bureau. *New York Daily News*

Appendix I: Forthcoming census digital safety tools

Over the next three months, DEL will work with its partners will generate a suite of census safety tools and curricula to assist civic organizations and libraries taking on census access and assistance. We will release these tools and curricula in a train-the-trainers manual due for release by September 2019.

Find your role flowchart

What role is your organization best equipped to play in getting out the count?

Forthcoming tools

- IT checklist that spells out an ideal set-up and determines what needs to be upgraded or built to bring a site up to safety standards.
  - The goal of this document would be to give library systems/CBOs a tool to use to conduct an audit of their machines so that they could generate a purchasing list or a workplan for upgrades.

- Public-facing one-pager FAQ that explains the basics of why the census is important, how it's going to work, and how to fill out the census digitally.
  - Librarians won't be able to sit with every single person that comes in to talk about the census, so we need to develop something that can stand on its own and cover the basics.

- Internal organizational FAQ giving librarians/staff/site administrators further information to assist census-takers. Examples of content include:
  - If a person doesn't have their mailer, they just need to provide their home addresses and the closest cross-streets.
  - The Census Bureau will mail a paper survey if people are uncomfortable with the digital survey.
  - What the HTTPS portal protects, and what it doesn't

    - What you can/can't do to assist, what you can/can't say

- Data management 101 document with simple best practices
    - Data management is at the heart of this process. CBOs might be pitched canvassing tools or apps, we want them to ask the right questions to ensure sensitive information is secure. CBOs, especially those working with vulnerable people, should also apply best practices for how and where they are storing data, especially if they are collecting or using canvassing data.

- Incident response plan(s)
    - Things could go wrong with census systems, public sites, or cyberattacks. Libraries and CBOs need to be prepared.

- Ongoing documentation and monitoring form
    - We need to have documentation of what's going right and wrong with the count for course correction as well as future planning and any challenges to the count.

Forthcoming curricula and workshop modules

Providing internet access

- Perform an audit of systems and determine access point design
- Setting up dedicated kiosk/access point (schematic with steps/specs)
- Guidance on how to build a more comprehensive data policy that can encompass initiatives such as the Census.
- Data management best practices (donated devices AND public hotspots)

What is the digital census (how to explain it, prepare for it, etc.) - from mailer to completion

- How to answer difficult questions about the census, data security, and privacy (especially if the citizenship question is included)
- Digital literacy support (what you can/can't do to assist, what you can/can't say)
- When to direct people to a paper survey

Digital risks and rewards (power not paranoia how-to for making your organization a place of safer digital civic participation)

- Training around digital safety and around open data for hard-to-reach and digitally marginal communities
- High level explanation of how data moves around the internet and safeguarding patron data.

Working with vulnerable communities

- Understanding how different communities experience risk and overcoming existing biases to assist census participants
- Direct training curated for CBOs that might be serving particularly marginalized communities
- Incident response and harm mitigation (physical safety as well as digital)
- Managing physical space (signage, device/kiosk security)

Monitoring the count and collecting documentation

- Weekly reports (misinformation; fraudulent mailers; hacking attempts; data breaches; etc)
- Setting up regular communications among Complete Count Committees, public library systems, and other census partners and coalitions
- Logging political/news flashpoints to compare with site reports
- Monitor compliance with recommended standards (configuration, network security, etc.)

Appendix II: Digital census IT best practices TK