KEEPING THE INTERNET
OPEN • INNOVATIVE • FREE

www.cdt.org

CENTER FOR DEMOCRACY
& TECHNOLOGY

1634 I Street, NW
Suite 1100
Washington, DC 20006

P +1-202-637-9800
F +1-202-637-0968
E info@cdt.org

PUBLIC COMMENT ON CMS-10440, "DATA COLLECTION TO SUPPORT ELIGIBILITY DETERMINATIONS FOR INSURANCE AFFORDABILITY PROGRAMS AND ENROLLMENT THROUGH AFFORDABLE INSURANCE EXCHANGES, MEDICAID AND CHILDREN'S HEALTH INSURANCE PROGRAM AGENCIES"

February 26, 2013

I. Introduction

The Center for Democracy & Technology (CDT) respectfully submits these comments in response to the Paperwork Reduction Act notice and request for comments by the Department of Health and Human Services' (HHS) Center for Medicare and Medicaid Services (CMS) entitled, "Data Collection to Support Eligibility Determinations for Insurance Affordability Programs and Enrollment through Affordable Insurance Exchanges, Medicaid and Children's Health Insurance Program Agencies" (the "Notice"), published on 29 January 2013 at 78 Fed. Reg. 6109, and the associated Paperwork Reduction Act supporting materials made available via CMS' website (the "Supporting Materials").[1]

CDT appreciates the care and consideration CMS has put into the crafting of the "single, streamlined form". We recognize that designing and engineering a form of this nature is difficult but essential in making precise eligibility determinations while being usable and only requiring the user to answer the minimum number of questions required to make a determination. However, the eligibility system is not simply a form, but an information system that will collect, maintain and make decisions based on sensitive personal and financial data.

We note the conspicuous absence in the Notice and Supporting Materials of any description of the *process* and *information systems* involved with: 1) identity-proofing of individuals that seek to determine their eligibility; and, 2) ongoing authentication of individuals once eligibility had been determined. That is, the "single, streamlined form" (the "Form") for determining eligibility when users apply online is a web-facing information system that collects and maintains sensitive applicant personal and financial data; it is not strictly a data collection form.

CDT recognizes the importance of adopting a process for credentialing that provides as high assurance as possible without imposing an undue burden on the population seeking affordable insurance. These comments have been developed to strike this balance. We recognize that this is a difficult set of issues to resolve and we would like to work with CMS (as well as state exchanges) to

[1] *See:* Supporting Materials for CMS Form Number CMS-10440, *available at:* https://www.cms.gov/Regulations-and-Guidance/Legislation/PaperworkReductionActof1995/PRA-Listing-Items/CMS-10440.html

successfully resolve them. In this comment, we address identity proofing, authentication and other security considerations for a web application that collects, maintains and makes decisions with sensitive data like that being developed by CMS to support eligibility determinations for affordable insurance programs.

II. Identity Proofing

The Supporting Materials seem to contemplate knowledge-based identity proofing, although it is not explicitly labeled with that term:

> *In addition, the online application will administer an identification proofing process. Based on the information an individual provides, the identification proofing system tool will generate three to five challenge questions, such as a previous address where an individual has lived. The tool will have a large bank of questions it will randomly generate based on information from external databases. Due to the security and integrity of the system, we cannot provide the list of questions generated. Additional burden from the identification proofing process is negligible in the context of the online application questions. Please refer to Appendix A for the placement of and more detail about the identification proofing process.*[2]

This statement is essentially the extent to which the identity-proofing process is detailed in the Notice and Supporting Materials.

Identity-proofing in the context of the Form involves making an independent determination that the individual presenting themselves for eligibility determination and submitting an application is in fact the person they claim to be.[3] For in-person and postal applications, identity-proofing can use standard identity-proofing mechanisms such as an official representative witnessing government-issued photo identification (in-person) or the applicant signing a statement guaranteeing the application is truthful and agreeing to penalties if not.

Remote identity proofing for online applications will likely need to employ more robust methods for ensuring the applicant's identity. The state-of-the-art involves "knowledge-based" identity-proofing: the applicant is asked a series of "out-of-wallet" questions — questions to which the applicant should know the answers without having to refer to any formal or necessarily secret credentials they may keep on their person or close at hand. Practically, this involves randomly generating a series of questions based on a historically rich database of information, which the proofing entity possesses or can access,[4] about the prospective applicant. A potential attacker or adversary that might seek to impersonate the applicant would be very unlikely to know all the answers to these questions, while the applicant should be able to easily remember the correct answers. If an applicant "fails" this proofing step — that is, if it is probabilistically determined that the applicant may not know the answers to the questions and thus not be the person they claim to be — the applicant would have to apply in-person or by mail, using alternative identity-proofing methods. In the cases where there does not exist a historically rich set of data from

---

[2] Id., Supporting Materials, supra fn 1. *See:* "Supporting Statement for Data Collection to Support Eligibility Determinations for Insurance Affordability Programs and Enrollment through Affordable Insurance Exchanges, Medicaid and Children's Health Insurance Program Agencies" at 8.

[3] In some cases an authorized representative will proof an applicant, but that is out of scope of this proceeding.

[4] As we mention later in the document, we would like to see clarity on what types of databases will be used in proofing. This is especially important in case CMS intends to create a new database, rather than using existing ones.

which to draw enough questions — for applicants that are young, that have poor memory or that do not participate in the kinds of transactions from which the data is drawn — will also have to apply by another means.

We have a number of potential concerns that could be alleviated by a more complete description of the proofing process:

- There are specific, difficult issues with knowledge-based proofing in the health insurance context. It is unclear whether CMS has considered these kinds of issues and addressed them. One serious issue with knowledge-based proofing of concern to CDT involves ensuring that the questions used for knowledge proofing can only be answered by the individual applicant and not necessarily by close family members. To the extent the form will be pre-populated with applicant data based on the results of the proofing process, proofing that allows impersonation by a close family member would allow inappropriate access to that data by the impersonator.
- Will failing the proofing step of the Form — or if the applicant cannot be knowledge-proofed due to age, lack of data, etc. — clearly inform the applicant, at appropriate literacy levels and in plain language, about why they failed the proofing process and how the applicant can avail themselves of alternative application methods? Given the purpose to enroll all eligible applicants, the process should proactively ensure that a problem at this step does not inadvertently dissuade eligible applicants.
- Will the proofing system employ any other mechanisms to make decisions about the applicant's identity? That is, knowledge-based proofing is often used in industry as one step in a larger risk-based identity proofing process that might include other elements (e.g., indicators that might signal high-risk fraud). Does CMS plan to employ other steps in the proofing process? If so, what are those steps?[5]
- Will the proofing system be developed in-house by CMS or procured from a contractor? (We are aware that CMS last year awarded a knowledge-base proofing and authentication contract to SAIC,[6] so that would seem to be a relevant vehicle for developing externally-developed knowledge-proofing for the Form.)
    - If CMS plans to develop its own knowledge-proofing system, it is essential that more detail be made available for public comment, including a Privacy Impact Assessment. A CMS-developed knowledge-proofing system that draws on only government-accessible data sources would draw on a much smaller set of databases and accordingly be unable to successfully proof as many applicants as commercial systems.
    - If CMS will use an externally developed knowledge-proofing system, it is very important in terms of privacy and security that additional details of how that system will be configured are made available to the public. Will sensitive information from applicants be sent outside of CMS' information systems to the contractor? If so, that transmission should be performed securely with strict

[5] The Supporting Materials make it clear that CMS cannot disclose the detailed questions due to security and integrity concerns with the system. This concern should not extend to reluctance in detailing the steps in the proofing process as there is little security or integrity benefit from general descriptions of these mechanisms.

[6] Press Release, Science Applications International Corporation, "SAIC Awarded $78 Million Contract by the Centers for Medicare & Medicaid Services", (February 27, 2012), *available at:* http://investors.saic.com/phoenix.zhtml?c=193857&p=irol-newsArticle&ID=1665739

retention limits, contractual terms requiring the contractor to maintain confidentiality, and full notice to applicants via the privacy notice/privacy policy. If the proofing system will be maintained and administered within CMS' information system architecture, how will CMS ensure that the proofing system is as up-to-date as the product if hosted by the contractor?

III. Authentication

Once the applicant's identity has been proofed, the system that implements the Form can then register a set of credentials that the user can employ to log in at a later date, either to complete the application if they did not have enough time previously or to interact with the system in future open enrollment periods.

The Notice, Supporting Materials and Form contain no information or details about ongoing authentication after proofing is completed. This seems to be a significant oversight; we would like to see more details of how ongoing authentication may work in the context of storing sensitive personal and financial data in a system that will be rarely accessed by individual users.

The framework for remote authentication applicable to this effort is contained in well-known OMB guidance and NIST technical supplements to that guidance.[7] We will not rehash that material here and will assume CMS is familiar with the contents of that guidance. CDT believes that CMS' credentialing and authentication of users should follow the January 2012 Health Information Technology Policy Committee (HITPC) Privacy and Security Tiger Team recommendations[8] to employ "level 2.5" authentication— that is, something more than a typical username/password combination, but not requiring full multi-factor authentication where users must enter another independent "secret", such as that produced by a secure random number-generating key fob or via a text message (SMS) to the user's registered mobile phone.

Some examples of additional "0.5" factors are the authentication mechanisms increasingly used in online banking. While real multi-factor authentication involves at least "something you know" and "something you have", additional knowledge secrets could be used for heightened authentication. For example, after the applicant has been identity-proofed they could be asked a number of security questions that could be randomly selected and asked of the user during authentication. It is important that these questions — and especially the answers to the questions — are not easily guessable by close family members and generic attackers. In that spirit, the user should be allowed to draft their own questions, in addition to being presented with good, unique and rare example questions.[9] Another example is where the user during registration chooses a security image and during the authentication process the user must confirm from a set of random images the one image they originally chose.

---

[7] Office of Management and Budget, "E-Authentication Guidance for Federal Agencies', M-04-04 (December 16, 2003), *available at:* http://www.whitehouse.gov/sites/default/files/omb/memoranda/fy04/m04-04.pdf; National Institute of Standards and Technology, "Electronic Authentication Guidance", Special Publication 800-63-1 (December 2011), *available at:* http://csrc.nist.gov/publications/nistpubs/800-63-1/SP-800-63-1.pdf

[8] Health Information Technology Policy Committee, Privacy and Security Tiger Team, "Trusted Identity of Patients in Cyberspace: Recommendations on Patient Identity Proofing and Authentication", (January 8, 2012), *available at:* http://www.healthit.gov/sites/default/files/tigerteam_patient_authentication_hitpc_010813_0.pdf

[9] Applicants could be required to draft *and* answer questions with no "canned" questions provided. However, we are hesitant to recommend that at this time due to the uncertainty of the burden of forcing applicants to do this.

To balance usability and authentication integrity, these "0.5" authentication steps should only be triggered every few months or after a significant period of non-use of the system. Also, when the IP address of the device connecting has changed, the "0.5" factor should be triggered or, if practical, a true multi-factor authentication mechanism could be triggered, involving sending an out-of-band secret to the user (for example, a PIN sent via SMS text message or via a phone call to the user's registered mobile phone). Enhanced single-factor authentication and highly-usable multi-factory authentication are both active and growing areas of research and practice, so we encourage CMS to monitor developments and engage with authentication experts and other parts of the Federal Government with interests in usable, robust authentication. As always, policy and technology should incorporate innovations that can enhance individual privacy and security and address new risks in balance with usability.

In the longer term, the Identity Ecosystem Steering Group (IDESG) under the auspices of the National Strategy for Trusted Identity in Cyberspace (NSTIC) aims to create voluntary, secure, reliable identity credentials for individuals to use in remote/on-line transactions. Such credentials — when available — could be useful for applicants in proving their identity and ongoing authentication. However, the work of the IDESG is nascent and a work in progress; it is very unlikely to be finished by the time the Form and its associated information system must be functional this October, so CMS cannot rely on that process as a solution at this point.

IV. Minimal Collection and Retention of Eligibility Data

It is unclear from the Notice, Supporting Materials and the Form to what extent collection of data is needed to make an eligibility determination. For example, it is unclear if a potential applicant can learn about eligibility criteria without creating an account and submitting data to CMS — i.e., by reading a guide or browsing CMS resources on affordable insurance programs. If these resources do exist, the privacy policy and any notice given to the applicant when online identity proofing has failed could also refer to these resources for additional information.

We also question if identity proofing could happen, instead of at the very beginning of the process, somewhere further into the process so that basic information might be collected in a more anonymous fashion before any sensitive applicant-specific information is requested. Finally, if the applicant does not complete the application or fails the online identity-proofing process, the information system implementing the Form should purge all data for such applicants (and any notice to the user about failing the proofing process should state clearly that any sensitive data they have submitted up to that point will be purged as soon as practical).

V. Other Security Considerations

Various accommodations will need to be made for the Form to be ideally usable by patients. It is difficult to make concrete comments on such accommodations without knowing more about how users will interact with the supporting software application and information system. For example, passwords will need to be somewhat strong and the system will have to give good advice for password generation grounded in the latest research about password-composition policies measured against password strength. Recent research has shown that policies that require long passwords — 16 characters — *without any requirements on password composition* are much more usable and stronger, in practice, than password policies that require fewer characters but have strict composition policies — e.g., requiring mixed-case characters, numbers, special

characters and prohibiting words from the dictionary.[10] If CMS cannot set such a policy and a user insists on a simple password, perhaps a more robust form of multi-factor authentication could then be triggered — out-of-band secret sharing via SMS, email, a phone call, or even postal mail — to better ensure that the credential is not misused.

With more details about the design and implementation of the software application and information system that will support the Form, CDT would be happy to provide additional comments.

[10] Komanduri, S., Shay, R., Kelley, P. G., Mazurek, M. L., Bauer, L., Christin, N., Cranor, L., Egelman, S. (2011). "Of Passwords and People: Measuring the Effect of Password-Composition Policies." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)* (pp. 2595–2604), *available at:* http://dl.acm.org/citation.cfm?id=1979321