Security Online: Establishing and Maintaining Identity in Health Insurance Exchanges

As states and the federal government continue to develop their eligibility and enrollment systems for health benefit exchanges, it is important to keep in mind that an online application is not simply a form, but an information system that will collect, maintain and make decisions based on sensitive personal and financial data. Consequently, it is critical that this system include a process for accurately credentialing online users, to ensure that a user is who she says she is. This credentialing process involves two fundamental steps: 1) identity proofing of individuals who seek to become eligible for health coverage; and 2) ongoing authentication of individuals once eligibility has been determined.

The Center for Democracy & Technology (CDT), Consumers Union (CU) and California Public Interest Research Group (CalPIRG) Education Fund recommend policies that address identity proofing, authentication and other security considerations for a web application that collects, maintains and makes sensitive decisions to support eligibility determinations for insurance affordability programs. We recognize the importance of adopting a process for credentialing online users that provides as high assurance as possible that they are who they say they are, without imposing an undue burden on the population seeking affordable insurance. Our recommendations 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 Covered California to help successfully resolve them.

I. Identity Proofing

Identity proofing in the context of the online application involves making an independent determination that the individual presenting herself for an eligibility determination and submitting an application is in fact the person she claims to be.[1] Remote identity proofing for online applications will likely need to employ more stringent methods for verifying an applicant's identity, due to numerous threats that are unique to online interactions, such as ease of impersonation, and the sensitivity of the information required to be shared for eligibility determinations (e.g. tax records or social security numbers (SSNs)). One method of state-of-the-art identity proofing is "knowledge-based proofing:" the applicant is asked a series of "out-of-wallet" questions — questions to which the applicant should know the answers

[1] In some cases an authorized representative will identity proof an applicant. This is an issue that will require some additional thought, as will identity proofing for those who use the service center or Assisters to apply online.

without having to refer to any formal or necessarily secret credentials she may keep on her person or close at hand. Practically, knowledge-based identity proofing involves randomly generating a series of questions based on a historically rich database of information, which the proofing entity possesses or can access,[2] about the prospective applicant.

Knowledge-based proofing mechanisms specifically rely on asking an individual to confirm knowledge of sensitive information, including current and former residences, automotive records, employment records, current and former financial events (collection events, credit cards, mortgages, loans), and lifestyle choices. A potential malicious hacker or adversary who might seek to impersonate the applicant would be unlikely to know all the answers to the questions, while the applicant should be able to 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 he claims to be — the applicant would not be able to move forward with the online application and would have to apply using a different method (such as in-person proofing). In the cases where there does not exist a historically rich set of data from which to draw enough questions — for applicants who are young, have poor memory or do not participate in the kinds of transactions from which the data is drawn — applicants would also be unable to apply online using knowledge-based identity proofing. Because of this, we would encourage Covered California to use sources that rely on data sets that are broad enough to support the maximum number of people possible eligible for Covered California and Medi-Cal to be identity proofed in this way.[3]

Important protections need to be built into knowledge-based identity proofing to ensure that it does not dissuade people from applying for coverage altogether. For example, applicants should be provided a multi-type verification system, where each person can choose to verify either via online questions or via other proofing methods. In instances where an applicant fails the online proofing, clear assurances in plain language and at the appropriate literacy level (nothing higher than a 6th grade reading level and in at least both Spanish and English) need to be incorporated into the messages generated by the system to clearly inform the applicant of the problem that frustrated the online proofing process and encourage her to apply using a different method. We do not want any applicant misunderstanding a difficulty in identity proofing as a finding of ineligibility altogether.[4]

There is no clear data as to what types of individuals will not be "proof-able" using these methods. We certainly expect there will be individuals applying for health benefits online who do not participate in the kinds of activities that generate transaction records in databases typically used by proofing companies. Covered California should rely on data sets that ensure

---

[2] 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 California and/or CMS and CCIIO intend to create a new database, rather than using existing ones.

[3] We anticipate that Covered California will want Assisters to help many of those individuals who may have challenges with identity-proofing. Covered California will have to determine how to integrate Assisters into identity proofing process.

[4] Practical issues regarding the source of the data that comprises the knowledge-based information and how to implement the process are beyond the scope of this set of recommendations. The identity proofing industry is composed of many players. We anticipate looking to CMS and its process for the Federally-facilitated exchange (FFE) to address some of the practical and structural implications of knowledge-based identity proofing.

that online knowledge-based identity proofing failures do not fall disproportionately on young, rural, and/or low socioeconomic status individuals as well as those people with memory-related disabilities.

As noted above, knowledge-based proofing mechanisms require individuals to confirm knowledge of potentially sensitive information. In the case of an applicant who needs assistance (for example, help from a "navigator") to determine their eligibility and/or apply for insurance, it will be important that the exchange consider how to ensure applicant privacy when an assister is providing help to an applicant over the phone.

II. Authentication

Once the applicant's identity has been proofed, the system that implements the online application can then register a set of credentials provided by the user to log in at a later date, either to finish an incomplete application at a later date (if she did not have enough time originally) or to interact with the system in the future for changes in circumstances, during redetermination periods or to choose a plan at a later date.

The framework for remote authentication applicable to this effort is contained in well-known OMB guidance and NIST technical supplements.[5] Credentialing and authentication of users should follow the January 2013 Health Information Technology Policy Committee (HITPC) Privacy and Security Tiger Team recommendations[6] 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, when available.

Some examples of additional "0.5" factors are the 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, she could be asked a number of security questions that could be randomly selected from ones she responded to during account registration. In the health context, it is important that these questions — and especially the answers to the questions — are not easily guessable by close family members or attackers with access to specific information, such as online social network service profiles (e.g., Facebook). In that spirit, the user should be allowed to draft her own questions, in addition to

---

[5] 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

[6] 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

being presented with unique and rare example questions.[7] Another example is where, during initial identity proofing and registration, the user chooses a security image and during the authentication process must confirm from a set of random images the specific one image she originally chose.

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. When data about the authentication session changes — for example, when the IP address of the device connecting to the account 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, a phone call to the user's registered mobile phone, an email to a registered email address, or, as a last resort, by sending an authentication code in the postal mail[8]).

III. Minimal Collection and Retention of Eligibility Data

To facilitate eligibility determination without collecting personal information from applicants that may not complete the process, may fail the proofing process, and/or who may not, in fact, be eligible for any of the health insurance plans, we recommend identity proofing happen as late in the application process as necessary, for example just before the application is officially submitted. That is, instead of at the very beginning of the application process, identify-proofing might make most sense further into the process, so that basic information might be collected in a more anonymous fashion before any sensitive applicant-specific information is requested. If the applicant does not complete the application or fails the online identity proofing process, the information system implementing the application should be designed to 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).

IV. Other Security Considerations

Various accommodations will need to be made for the online application to be ideally usable. It is difficult to make concrete comments on such accommodations without knowing more about what data potential vendors access, what other public programs rely on for online applications, and how users will interact with the supporting software application and information system. We would recommend that Covered California pilot test any proposed process to sample potential applicants before the system goes live in October 2013. For example, given the sensitivity of the information the system collects and maintains, 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

[7] 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.

[8] This is similar to methods used for motor vehicle registration renewal in California and in some voter registration contexts; a month or two before open enrollment starts, Covered CA could send a mailing in a sealed envelope with information about the open enrollment process and somewhere on that mailing an authentication code — unique to the individual and randomly generated — could be printed.

password strength. For example, recent research has shown that policies that require long passwords — 16 characters — without any (potentially confusing or dissuasive) 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.[9] If Covered California cannot set such a policy and a user insists on a simple password, perhaps a more detailed form of multi-factor authentication could then be required — out-of-band secret sharing via SMS, email, a phone call, or even postal mail — to better ensure that the credential is not misused.

V. Rapidly Changing Technologies

Enhanced single-factor authentication and highly-usable multi-factor authentication are both active and growing areas of research and practice. We encourage Covered California to monitor developments and engage with authentication experts and other State and Federal agencies with interests in usable, secure authentication. Policy and technology should incorporate innovations that can enhance individual privacy and security, always balanced against new risks they might create for 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 Covered California's application form and its associated information system must be functional, so this is not something on which California can rely on as a solution at this point.

VI. Conclusion

Covered California has developed robust security parameters for itself and its vendors and has engaged regularly with stakeholders to solicit feedback on privacy and security policies. How users interact with the online system right out of the gate, starting on October 1st, 2013, will be a strong indicator of the policies in practice. Developing secure identity proofing and authentication standards that carefully balance privacy concerns and security protections against ease of use for consumers will be critical in determining whether Californians can comfortably interact with the online portal and securely, confidently, and expeditiously apply for and enroll in coverage.

---

[9] 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