

# HIT Policy Committee -- Hearing on Authentication

Response from David McCallie, MD  
Cerner / Kansas City, MO  
Jan 7, 2010

My name is David McCallie, MD. I am Vice President for Medical Informatics, at Cerner, in Kansas City, MO., where I have been an Associate for 18 years. I am also a member of the HIT Standards Committee.

Cerner and I would like to thank the HIT Policy committee for this opportunity to testify on this important issue concerning authentication of the users of health information technology.

Cerner is transforming healthcare by eliminating error, variance and waste for healthcare providers and consumers around the world. *Cerner*<sup>®</sup> solutions optimize processes for healthcare organizations ranging in size from single-doctor practices, to health systems, to entire countries, for the pharmaceutical and medical device industries, and for the healthcare commerce system. These solutions are licensed by more than 8,000 facilities around the world, including approximately 2,100 hospitals; 3,300 physician practices covering more than 30,000 physicians; 500 ambulatory facilities, such as laboratories, ambulatory centers, cardiac facilities, radiology clinics and surgery centers; 600 home-health facilities; and 1,500 retail pharmacies.

**.1) What trust problems are you trying to solve and for what range of users (e.g. organizations, individuals, health care professionals, consumers)? Please provide some quantitative data if possible to characterize your user base (e.g., percentage or number of each type).**

Cerner Response: Cerner supplies a wide range of health information technology solutions, ranging from large integrated delivery networks (IDNs) to small-sized provider practice and from employer-facing TPA systems to consumer-facing PHR and HIE systems. The challenges of identity management, credentialing and authentication are different for each of these settings. I will summarize the approaches that we currently use in addressing the needs of some of our core markets.

## **Integrated Delivery Networks, Hospitals, and large Clinics:**

IDNs and hospitals will generally have robust identity management and credentialing systems in place, given the regulated nature of healthcare, so our challenge is usually focused on finding efficient and secure methods of authentication against existing credentialed users, both within our Millennium EMR system, as well as across the other (non-Cerner) systems that are deployed in the client's enterprise.

For providers and other uses of our EMR system, we offer a variety of choices on strength of authentication. Each EMR client makes their own decisions regarding the tradeoff between strength of authentication versus the degree of difficulty imposed on busy clinicians. Through third party integrators (e.g., Imprivata) Cerner offers a variety of optional "advanced authentication factors," including finger-print readers, various kinds of hardware tokens (e.g. Vasco tokens, RSA SecurID) and both active and passive RFID cards. In addition to these hardware choices, we can flex authentication requirements based on the location of the accessing device, and whether or not the device is located inside or outside the perimeter firewall.

Despite the availability of a wide range of authentication devices, we have found that only a small minority (under 10%) of our clients have chosen to implement extra factor authentication on a wide basis. With the notable exception of Ohio's mandatory requirement for a second-factor token for ePrescribing, we have found that the inconvenience of multiple authentication steps, and as well as the cost and complexity of deployment have rendered use of these hardware authentication factors surprisingly rare.

Of note, we have recently seen an increasing interest in the use of passive-RFID cards, coupled to a secondary user PIN, as authentication tools for the EMR. Most healthcare campuses require every employee to carry at least one RFID card to gain access to campus facilities. When the user's identity on the RFID card is tied to a personal PIN/password, the resulting "tap-in / tap-out" login process is appealingly fast and easy, especially when coupled to either "virtual desktop" sessions (e.g., VmWare View) or to "session following" technology (e.g. Citrix.) This increases security by leveraging the physical token which most users already have to carry. A secondary benefit of being able to lock the entire session and all the associated applications enhances the security of the applications.

Once authenticated, the user's identity often needs to be propagated to other enterprise systems. This intra-enterprise transfer of trust goes by the common name of "single sign on (SSO)." One of the best ways to deploy enterprise-wide SSO is to use a network-based primary login, followed by trust transfer of network-validated credentials to dependent systems. We offer this capability to bypass the Cerner login process entirely, trusting the credential provided by the enterprise's network, using custom code that depends on the client's network. However, despite the availability of standards like SAML, surprisingly few clients have implemented this approach. More common, though still somewhat rare, is the use of a single enterprise-wide "active directory" (LDAP) so that all systems can share one source of truth about currently authorized users. Cerner systems can also implement this approach by capturing user login input and then directly calling the LDAP service to authenticate the user. This approach has the advantage of a single point of control over identity management of all employees, without requiring the more complex reliance use of network-passed credentials.

However, the most common SSO method in use by our clients remains the old-fashioned "screen scraping" system wherein the SSO service mimics the human's interaction with the dependent application's login screens. Given the wide range of legacy systems deployed in a typical large healthcare enterprise, "screen scraping" remains the only "universal" approach available for SSO.

Despite availability of all these choices, we estimate that fewer than 30% of our clients use any form of SSO at all. One reason for the low usage may be the fact that often we find that the network identity space does not always map cleanly to the EMR user's identity space. (More on this topic below.)

#### **Small-sized provider practice settings:**

Cerner deploys EMR services to smaller practices solely via an application service provider (ASP) model. All of the hardware token choices discussed above are available to the small practice, but so far we have only seen occasional interest in anything beyond standard username and password. SSO models can be applied in ASP settings, but so far, this also has been uncommon.

## **Providers and other users in non-EMR settings:**

Cerner's broader solution set includes a new "cloud-based" network service which facilitates the coordination of care, incentives, and commerce across consumers, providers, and payers. For purposes of this discussion, I will refer to this aggregation of loosely-coupled services as "the Cerner Network." Consumer users of the network range from those connecting as patients receiving care at individual provider organizations to those utilizing services such as HIEs and PHRs delivered from cloud-based services. Provider users range from web-based cross-institution secure messaging to patient-facing portals that facilitate patient-provider communication. Payer users represent those from self-insured employers, governments and commercial insurance companies.

It is highly desirable that these distributed network services allow for cross-organizational passing of trust, so that each organization or service can maintain its own membership roster and yet support secure interactions with other services without having to re-identify and re-credential all of the users of the other services or realms.

Recently standards have emerged that facilitate the development of these "federated trust" models. Cerner has implemented a variety of these standards and is beginning to identify appropriate deployments. The new tools work well; however, so far our experience is that the deployments of "federated trust" is not as easy as might be hoped or expected.

Let me provide one example of a recent experience with a "federated trust" deployment. Cerner has implemented a large-scale, internal, web-based social networking service which we call uCern. uCern is designed to bring Cerner associates into direct communication with our clients, via wikis, blogs, discussion forums, and other tools of social networking. It is our near-term goal to have every user of our Millennium systems also possess and use a uCern account. Longer range plans include extending uCern not only to our provider clients but also to their patients and to other consumers who participate in any of the "Cerner Network" services.

Cerner associates automatically gain access to uCern via our enterprise network-based login (We use Kerberos.) A key goal in rolling out uCern was to avoid a cumbersome enrollment process for the large number of non-Cerner users. Our initial design was to create a federated model by creating a trust framework between our client's existing Millennium EMR identity and the uCern network. Ideally, any user who was already authenticated to a Millennium instance would be transparently be connected to uCern. We implemented this trust relationship between Millennium and uCern, but we ran into a number of issues which made this approach fail.

The most important limitation was that the overlap of the user rosters of Millennium with those who needed access to uCern was far less than 100%. This meant that in order to use uCern, a number of non-Millennium users were forced to create "dummy" Millennium accounts in order to access uCern. This was unacceptable.

Additionally, we have a number of hardware and software suppliers where we needed to invite only a small number of their employees to join uCern. It would not be cost effective to create cross-enterprise trust relationships for such a small number of jointly interesting identities.

With these lessons in mind, we took a less complicated approach that still allows us to leverage the identity and credentialing work that had already been done by our clients. We did this by leveraging the

email name (and domain) of the outside clients to establish the core identifier for anyone applying for a uCern account. The organization enrollment process works as follows:

- We ensure the client's email domain(s) are included on the "white list" of approved organizations.
- We leave it to our client to communicate to their employees the address of uCern. They typically announce it to their IT organization first, then to their providers.
- Once announced, the users just need to access the URL to uCern and register using their client email address. Their email domain will be validated and accepted.
- Then uCern will validate that they control the provided email account by sending the user an email which includes an authentication link to complete the registration.

Knowing that our clients (including Cerner) have other types of users outside of their organization that they will want to collaborate with, we have provided the ability to individuals that have email domains that do not exist on the approved list. This user-specific invitation process works as follows:

- We send invitations via email to the invited user, at their organizational email address, inviting them to join uCern.
- Included in the email is a link to the uCern registration screen and an authentication token that identifies them as a direct invite, rather than an organizational invite. This allows them to skip the organization domain check.
- uCern validates the completed registration by sending an email with the authentication link to complete the registration.

This simplified process has proven very successful. Uptake of uCern by our client base has been rapid. Since enrollment started this past July, we have grown to 1050 active clients and over 16,000 authenticated users. By leveraging email domain names, we were able to take advantage of the fact that email addresses are a tightly controlled resource. With only a moderate development effort, we have been able to reduce the chances of impersonation and fraud to low levels. This required no development work on the part of our clients. In the long run, we intend to move to more formal SAML2 identity management (see below) but for now, leveraging the email domain name has been adequate.

## Consumers and Patients:

Patient registration and enrollment is a big challenge. The process we are deploying for the “Cerner Network” leverages ordinary email in a manner similar to the uCern approach described above. The process is summarized here:

- A patient is invited to join the Cerner Network by a provider.
- The provider’s front desk staff uses an enrollment tool to generate an invitation to be sent to the patient. The enrollment tool captures a shared secret (an answer to a patient-selected question) which will be used to make sure that the right patient responds to the email invitation.
- The enrollment tool sends an ordinary email request to the patient’s selected email address. Included in the email is an enrollment token specific to this patient and provider, and a link to an online enrollment process.
- If the patient already has an account on the Cerner Network, she logs in and presents the new enrollment token, which establishes the relationship to the new provider.
- If the patient does not yet have an account, the patient enrolls at the provided web site, and offers the enrollment token, which creates an account and the link to the provider.
- At the present time, patient authorization when connecting to the secure Cerner Network web site is by username and password, though we are considering adding additional factors similar to those used by banking portals (see below.)

## **2) Who pays for the solution, implementation, processes and support for your approach? What factors contribute to the total cost of ownership of the technologies, including process costs? What are the implications to widespread deployment?**

*Cerner Response:* In institutional or ASP EMR settings, the provider’s organization pays. Despite the falling cost of hardware tokens, this is a significant issue, though the dominant issue is not the cost of the hardware or software as much as it is the “cost” to the provider’s time and productivity. If the burden of additional authentication is high, then providers are unlikely to accept it. Ohio ePrescribing is a notable exception, where the extra authentication is required by statute.

We feel it is unlikely that consumers will elect to pay for advanced authentication services such as hardware tokens. In fact, they have shown little interest in paying for PHR services or for other health related portal services at all.

How can we create stronger consumer authentication at a very low cost? One approach would be to follow the banking industry’s lead with their use of clever but inexpensive secondary authentication factors for their consumer portals. These factors include:

- Use of cell phones (or home phones) as ersatz hardware tokens. Using SMS messages to send an authorization token to the consumer is a powerful way to reduce impersonation threats.
- Use of encrypted browser “cookies” to detect when a consumer is using a previously unknown device or browser to access the portal, which triggers the use of extra challenge questions.
- Use of secret knowledge that the service provider already has about the consumer as a secondary factor. For example, bank account numbers, last 4 digits of SSN, etc.

- Anti-phishing tools, including consumer-selected “personal icons” which a phishing site would not know about.

Additionally, the process overhead of enrollment is high when the target population is large, unless it can be substantially automated from existing rolls. We try to leverage our provider base to recruit consumer members, but the extra work for the front office staff is a potential barrier to adoption.

The third-party software costs have been manageable as there are good open source toolkits for most of the core authentication services. If the network wishes to purchase a third-party database of “secrets” (such as prior addresses, historical mortgage payments, etc) then these costs could be significant if the targeted population is large. Ideally, the sponsoring organization already has access to “secrets” such as accounts and partial SSNs, reducing the need for external databases.

### **3) Directory services often support some certificate authority or other authentication mechanism. As you look more broadly at the architecture, how do your approaches work with such directory services?**

*Cerner Response:* Our most direct work with externally-controlled directories has been with Surescripts. The Surescripts provider directory must be integrated with Cerner’s provider directory so that an eRx generated by a provider known to our system can be communicated to Surescripts without requiring that the provider re-authenticate herself to the Surescripts network. Cerner has established a secure eRx hub that connects directly to the Surescripts network, using a combination of VPN and TLS connections. Our clients connect to the Cerner hub and then have their eRx forwarded to Surescripts, with the Surescripts’ provider ID mapped by software to the Cerner provider ID.

The Cerner-to-Surescripts provider ID (SPI) mapping process is currently a manual process. It works adequately, but problems have arisen when there is a mismatch between the ways provider identities are tracked in Cerner versus Surescripts. This mismatch might occur if a provider works in more than one organization, or uses more than one EMR tool at a given location. Problems can arise if we are unable to match the different approaches to identity definition. This could result in a refill request being sent to the wrong location or to the wrong EMR, for example.

This can be seen as an example of a generic directory-mapping problem that might be called “identity granularity mismatch,” which occurs when an identity exists in more than one directory but the granularity or scope of the identity is different between the two directories, such as one provider who works at independent organizations. The difference in range or scope of identity may make cross-directory mapping difficult or impossible. (This is analogous to familiar “semantic granularity mismatch” problem which occurs when there is not a 1:1 mapping of the exact meaning of terms between two candidate nomenclatures.)

### **4) Does your approach support a delegated authentication model where there is an authorized registrar that issues the authentication credentials to individuals? If so, how? Are there implications for interoperability in this scenario?**

*Cerner Response:* Yes, we have different mechanisms for this, depending on the use-case, but in general, a delegated authentication model is almost always necessary from a practical point of view.

For EMR users, our clients have access to a variety of user-management tools that cover all aspects from credentialing and account creation, to setting up authentication constraints, to authorization of specific role-based privileges. We do not perform any centralized authentication of our EMR users, except as noted above for Surescripts mapping.

To my knowledge, no one has requested that our EMR serve as a SAML identity provider (IdP), so we have not added that capability. In general, we would expect that the controlling enterprise would manage their own SAML IdP, and that Millennium would be a service provider (Sp.) Other than for B2B-style interfaces, Millennium has not participated in cross-enterprise, user authentication strategies except as described herein for uCern and Surescripts.

For our uCern (social network) users, we have essentially delegated identity management to our clients, who manage their locally assigned email addresses. We then leverage those controlled email addresses and domains to grant access to uCern-specific services.

For consumers and others on the “Cerner Network,” we have tools in place that can delegate authorization of specific “realms” that can then share credentials with other realms in the overall network. In particular, we are using the OASIS SAML2 stack, and OAuth, with some legacy CAS (Central Authentication Service) code in the mix.

However, the existence of tools and standards does not mean that cross-realm identity mappings are always useful or even possible. The “identity granularity” problem described above is one example of where the tools cannot solve a more fundamental problem of non-mappable identity.

## **5) What should be the role of government? Where can rapid action address common concerns or limitations of trust?**

It is tempting to propose that the role of government should be to manage the identity for all potential users of a national health information network. Superficially, it would seem that this could solve a number of cross-domain trust problems, by creating unambiguous, top-down identity which could then be passed from organization to organization, with appropriate authorizations assigned by each organization.

However, we believe that this would prove to be cumbersome, expensive, and unlikely to succeed. The number of users of “the healthcare internet” would be huge – essentially including all Americans, since “meaningful use” will hopefully drive widespread adoption of consumer-engaging services such as PHRs. Even if the credentialing were to include simply “providers” the number of credentials to manage would be very large, given the large number of staff users who would need to be authorized in order for care delivery systems to work efficiently. The time and expense associated with creating the current NPI is a good example of the challenges that would be faced.

There is considerable evidence that a “national unique patient identifier” could improve the nation’s healthcare ([http://www.rand.org/pubs/monographs/2008/RAND\\_MG753.pdf](http://www.rand.org/pubs/monographs/2008/RAND_MG753.pdf)) however, the well-known political resistance to such a project makes it unlikely to succeed, at least in the near term.

Additionally, the “identity granularity” problem would make any single “one-size-fits-all” identity source unworkable in the complex world of healthcare where “different hats” are worn by the same individual as he or she interacts with full extent of the healthcare system. Unfortunately, untangling “role” from “identity” is not as easy in practice as it is on paper.

We think that distributed identity systems will remain the best way to manage the complexity of identity and authentication in healthcare. Well-scoped entities that know their users are best positioned to identify and credential the users of those focused systems. In some cases, the organizational identity scoping will be locality based, such as for a regional HIE. In some cases, the scoping might be affinity-based, such as for a national PHR/HealthBank, or for a provider specialty credentialing organization, etc.

Even though it may not be possible to have top-down management of identity at the end-user level, it is feasible to manage some degree of trust at the “organization” level. For example, Wes Rishel and I have recently proposed a model for “simple interoperability” that depends on creating a network of trusted and secure “health internet nodes” which could be used to create a simple “secure email” system without requiring the complexity and cost of a top-down PKI or S/MIME model. Such a system could be used for “push” distribution of patient information using both structured and free-text messages.

[http://blogs.gartner.com/wes\\_rishel/2009/12/15/simple-interop-the-health-internet-node/](http://blogs.gartner.com/wes_rishel/2009/12/15/simple-interop-the-health-internet-node/)

The proposed “simple interop” system depends on each “health internet node” in the network following agreed-upon policies that would govern which nodes the email gateways could talk to, and how the conversations must be kept secure. One way to facilitate emergence of such a network would be for the government (or a designated entity) to:

- Define the policies that must be met before a node was allowed to connect to the secure network. This would probably include policies defining minimum standards for end-user credentialing and authentication.
- Define certification criteria that could be used to verify that a node was in fact following the “secure channel (TLS)” policies (following the model for Certified EHRs under Meaningful Use.)
- Manage (or designate) the certificate-issuing authority (CA) from which each secure node must obtain the necessary certificates that would enable participation in the secure network.

These proposals are similar to concepts from the NHIN CONNECT architecture; however, they are perhaps less complex to implement and are offered here as a example of a simple starting point.

Note that even though the government could play a significant role in guaranteeing the basic security of the overall network, the actual process of credentialing and authenticating end-users would be left to the various organizations that join the network. At least in the early stages, implied trust between the nodes would be limited to knowledge of which node you were communicating with, and trust that the message was securely transmitted between the nodes on the network. Additional levels of trust would have to be negotiated and developed over time, as the network matured, and experiences were gained.