In the latest draft of the U.S. National Institute of Standards and Technology guidance document on authentication and lifecycle management—and isn’t that a mouthful, even if you’re reading silently to yourself?—the authors have come out with the following bombshell:

Due to the risk that SMS messages may be intercepted or redirected, implementers of new systems SHOULD carefully consider alternative authenticators. If the out of band verification is to be made using a SMS message on a public mobile telephone network, the verifier SHALL verify that the pre-registered telephone number being used is actually associated with a mobile network and not with a VoIP (or other software-based) service. It then sends the SMS message to the pre-registered telephone number. Changing the pre-registered telephone number SHALL NOT be possible without two-factor authentication at the time of the change. OOB using SMS is deprecated, and may no longer be allowed in future releases of this guidance.

Here is another example of something we see too often in security: a completely correct assessment of possible risk, without even a smidgen of discussion on probability or helpful guidance on alternatives. Everyone’s a critic, and nobody has a useful answer as to how to fix it.

Read the next page to see why this is so...
1. Implementing even SMS-based authentication is a major project. For organizations with large, nontechnical external user bases, it incurs a significant support cost as well as a technical one. Many organizations haven’t even gotten that far yet, and with this NIST announcement, that rug is being pulled out from under them.

2. For those who are already using SMS as an out-of-band authenticator, do they now need to switch to something else? SMS was a cheaper solution in that they didn’t have to issue and manage hardware tokens (one instance where BYOD was actually a good thing). The migration costs may be impractical for some organizations, even if they decide that they need to follow this guidance.

3. Any time you have guidance or standards that narrow your choice of technical solution, there are winners and losers in the security industry. In this case, losers will be vendors of SMS-based authentication products, and winners will be vendors of hard and soft tokens, authentication apps, and social media-based authentication services such as those from Google and Facebook. In cases where the vulnerability discussion appears to rule out (or at least disfavor) a large number of products, it creates the possibility that a government entity such as NIST is giving an unfair commercial advantage to a limited number of players.

4. Requiring external users not only to set up a second authentication factor, but to receive a hard token or download yet another app for their phones, is unreasonable for many populations. Think about citizens who will be required to install a Division of Motor Vehicles Authentication App in order to renew their registration online. Or think about retail customers—millions of them—who would have to download and learn how to use multiple authentication apps for the dozens or hundreds of e-commerce sites they want to use. Right now, I have three hardware tokens and two authenticator apps serving another four or five sites. I need a separate purse just for my keyring.

5. Requiring a second authentication factor to be used just to change the pre-registered phone number: how is that supposed to work? Does this mean we have to introduce a THIRD authentication factor?

6. Organizations cannot dictate what kind of mobile phones their users have, and all over the world there are millions of non-smart phones in use. SMS is the only thing that will work across the majority of them. **Like it or not, SMS will be the IE6 of 2FA for years to come.**

7. Finally, we see email being used all the time as a form of validation—call it a slow second factor for authentication—and we know email gets compromised more often than SMS does today. Why isn’t NIST deprecating email? Well, because we all know how unreasonable that would be.

I’m sure the NIST folks thought long and hard before coming up with this guidance, but I predict it won’t make much difference to those organizations who have to live within various real-world constraints. We simply have to get more realistic about acknowledging possible risk without treating it as a binary condition that, once flipped from zero to one, closes doors for security defenders. Guidance without widely usable alternatives is just criticism.