Not all QR codes are evil, but some are.

Phil Archer, Web Solutions Director, GS1 Global Office. phil.archer@gs1.org

QR codes are a ubiquitous method of passing a URL across the air gap from some sort of physical medium to a user via their smartphone's camera. Typically, those URLs are very short, use meaningless-to-humans generated paths or query parameters, and exist on domain names like qr.codes, qrco.de and any number of less-obvious names. There are two drivers in play here:

1. The shorter the URL, the smaller and more easily read the QR code can be.
2. The product or service owner really wants analytics about how many times their QR code was scanned, at what time of day, where and, if at all possible, by whom – the kind of analytics typically provided by URL shortening services.

The problem is that, as with all redirection, it can take you anywhere – including to bad actors. There are no clues in a URL like https://kwz.me/hfi[1] as to what it will take you to. If you see a product on the shelf with a sticker on it bearing a QR code saying "scan now and win", you might be tempted to scan. But is it genuine or not?

The problem is that we're missing context. A lot of companies organise compulsory training for their staff on how to spot phishing emails. Thankfully, a lot of people are wary of opening attachments in emails and, for good or ill, a lot of us are using company-issued laptops that check every link we follow.

What's missing from a product, a package or a poster is all of the contextual signals that might warn a user that they probably shouldn't follow a link. Security software is typically missing from personal phones when scanning QR codes in the wild. This means that some users are wary of following *any* link in a QR code so that good actors are missing out because of a small number of bad ones.

Why does GS1 care about this?

Because we're in the process of retiring the 1D barcode and replacing it with 2D codes, including QR codes carrying URLs. Over the next few years, you're likely to see products carrying QR codes alongside 1D barcodes and, after that, the 1D barcode disappearing over time.

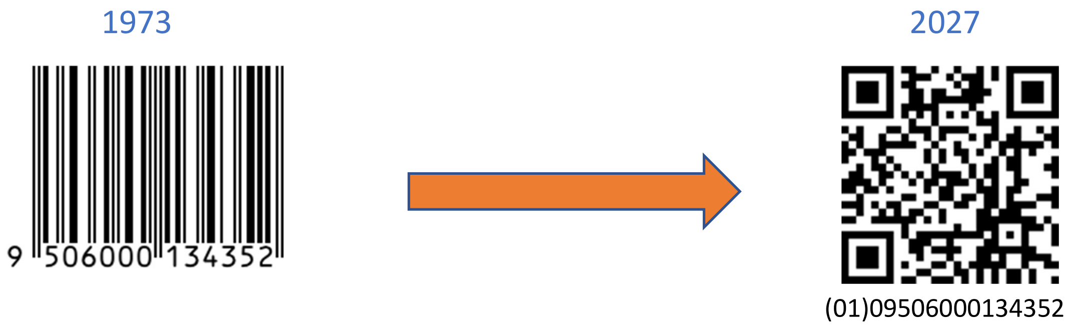

Figure 1 GS1's "Ambition 2027" is that the world can begin to use only 2D barcodes on products from that year.

[1] The kwz.me service is run by a private individual and is used purely as an example. It is not suitable for use in production-grade systems.

The structure of what we call a GS1 Digital Link URI is precisely defined to encode GS1 identifiers. Scanners can recognise and extract those identifiers *exactly* as they do now when scanning traditional barcodes without the need for any online lookup. The most common GS1 identifier is the GTIN – the number familiar from 1D barcodes – but 2D codes can include things like batch numbers, expiry dates, serial numbers and more, all of which offer significant potential to brands, retailers, patients and clinicians.

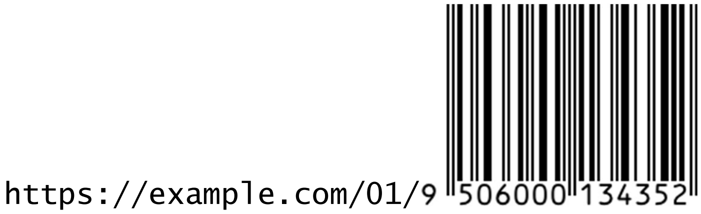

Figure 2 A simple GS1 Digital Link URI encoding just a GTIN, the number familiar from the barcodes in use since the 1970s (the barcode itself is included for illustration only, the actual URI is, of course, simply https://example.com/01/9506000134352)

GS1 provides advice to brand owners on how to create their GS1 Digital Link URIs in QR codes and recommends that they use a subdomain of their own Internet domain name. This is possible because the identification of the product is provided by the GTIN, just as it is today, so the domain name in the URI is *not* part of the product identifier. Therefore a more realistic GS1 Digital Link URI might be as shown in Figure 3.

https://id.dalgiardino.com/01/9506000134352

Figure 3 A GS1 Digital Link URI that follows GS1 best practice (and is encoded in the QR code in Figure 1). A subdomain of the brand's own domain name is used, followed by the GS1 identifier(s). Dal Giardino is a fictitious brand used for examples in GS1 documentation.

The challenge

The challenge to the Web community is what steps can be taken, and by whom, to distinguish more easily between URLs in QR codes that can be trusted and those that might possibly be unsafe. And how can this be done without a massive proprietary database of unsafe domain names? Some sort of multi-factor authorization as standard for 'legitimate' redirection services? A simple preview on the screen of what you will see if you follow the link (by pre-fetching the content before the user clicks anything). Or, perhaps the visual cues associated to such codes when present on actual products (like proximity to the older 1D barcodes or the text shown just below the example on the prior page) are enough across some domains of engagement? Better ideas? (please).