

The Battle for your Data

Open Protocols And Why They Matter

The first time I ever laid my fingers on a word processor it was in 1986 at the keyboard of a Commodore 64. Those early documents I wrote are long gone now, denizens of the great bit bucket in the sky. It was the early days of mainstream computing, and at the time each computing platform was essentially an island: there was no hope of getting the Commodore to communicate with the IBM PCs of the time, much less the revolutionary new Apple Macintosh upon its entry into the market a year or two later. For that matter, Macs and PCs operated in relative isolation too.

Fast forward to 2003, where we – students and office drones in particular – spend the better part of a day glued to the business end of a laptop writing and receiving documents, and the worldwide internet traffic in text docs, spreadsheets, and of course email, exceeds 180 petabits per day (A petabit is one quadrillion 1s or 0s). So we've learned to communicate with each other and are now enjoying the brave new Information Age, in which most any computer can communicate with most any other computer, right? Not completely. Certainly we've come a long way forward in being able to interchange information from the “bad old days” when specific equipment was required to access specific information. But Nirvana it is not: at the opening of the 21st century data formats have become a competitive weapon, and proprietary protocols are being wielded in a way that leads us back towards a world of less interoperability, not more. This dangerous trend will shape the future of the computing – and therefore the business – world for as long as we allow it.

Document formats

Start with electronic documents. The papers I wrote in the late 1980s are no longer accessible to me, as I no longer have either the computer or the software to read them, and given the small market share, it's in no company's financial interest to help me, as the “Commodore 64 document conversion” market is not exactly booming these days. But owners of relatively modern hardware aren't necessarily better off. Did you write some papers prior to 1995 on a previous version of Microsoft Word? There is no guarantee Word XP will be able to import them cleanly or even at all. Write your paper on a late 1980s word processor by the name of Wordstar? That company is long gone, so unless you were paying attention in the short period of time that both Wordstar and its competition co-existed, you are up the creek. Remember an older program called Write Now? It's gone as well, and any documents written on that program are no longer accessible by any known software package. Even modern software is at risk. Adobe Framemaker documents are only readable by, well, Adobe Framemaker, and don't think Microsoft is going to want to help you change.

What makes you think you'll be able to access the masterpieces you crafted as part of your SAIS career? I may decide that it's not worth it to carry along my macro problem sets from one version to another and let's face it – the things I wrote in the 1980s are of no interest to anyone, but most businesses are not afforded the same luxury of letting old documents disappear. That means using planned obsolescence to force corporations to keep their document formats up to date is a big money maker. Not surprisingly, it doesn't bode well for the consumer, especially if to upgrade to the new, powerful version of Word XP you need to a newer computer. Feel like you've been forced into an expensive upgrade cycle? You have.

The unfortunate truth is that word processors don't do a whole lot more today than they did 10 years ago, so to encourage continued upgrading, software makers wield the planned obsolescence of document formats like a sledgehammer. For the software publisher there are two benefits: first, it guarantees a revenue stream from individuals and corporations that need to continually purchase

the latest and greatest in order to maintain the ability to read their own documents. Second, at each upgrade it provides an opportunity to lock other software makers out of the market. The format for Microsoft Word documents is one of the industry's best kept trade secrets, up there with the Colonel's secret recipe and the ingredients to Coca Cola. Mac owners' familiar cry that their software never does a perfect job at reading documents written by PC owners, then, should be no surprise at all. They have two options: spend a couple hundred dollars for Microsoft Word for Macs, or throw in the towel and buy a PC (which will come with Microsoft Windows, conveniently). Linux users and those poor souls still using OS/2 are out of luck. Most software by the competition is able to read Word docs "more or less" But not perfectly. And that's just the way Microsoft would like it! It's a big stimulus to go mainstream. The coup de grace, then, should be obvious to you by now: if you decide to overcome the problem by simply going with the flow and doing your work in Microsoft Word, you've effectively committed to that format. Because it requires too much effort and too much money to change, and because it's too imperfect a process anyway, you are stuck. Your data has been taken captive.

Email

If it seems like major software manufacturers are fighting for the rights to take your documents captive, it should come as no surprise to you that the same battle is being waged for your email. There are dozens of programs that let you download, read, and send email, Outlook being by far the most popular, but Mozilla, Eudora, and Opera are others. They differ in many aspects but are similar in only one: they make it ridiculously easy to import your mail and addresses and extremely difficult to export them. In effect, they raise the cost of your being able to defect to the products of another manufacturer. Outlook, for that matter, has been the vector for hundreds of viruses and Trojan horses that cause mischief and damage on your computers. Why continue to use it? Well, certainly it's popular because it has made the task of dealing with your mail a lot less burdensome. But the price is this: it's made it harder than ever to move away to anything else.

The World Wide Web

The most insidious use of document formats and protocols has occurred on the ostensibly free and open World Wide Web. The World Wide Web since its inception was intended to be an open platform that allowed anyone anywhere using any computer and any web browser could access the same information. In fact, it was that openness and universality made the web indispensable in the first place. But the browser war – in which Microsoft and Netscape battled over which browser would be more popular – put an end to that (Who won? Well, which browser do you use? 97% of the world uses Microsoft Internet Explorer today). A decade after its birth, badges proclaiming "this website best viewed with Netscape 6" or "this website best viewed with Internet Explorer (download here)" are still around.

How is it possible that one web browser can better render a standard language - HTML - than another? By tweaking the protocols. As recently as a 2002, Microsoft attempted to make its MSN website inaccessible to web surfers using anything other than its own Internet Explorer. Web surfers using any other browser – Safari (Macintosh), Opera, Netscape, Mozilla – were sent to web page informing them their inferior software would deny them the proper "experience" and a convenient download link. The inventor of the World Wide Web, Tim Berners-Lee, in an interview with Silicon Valley magazine, railed against this and similar attempts to link information to particular technologies. "I have fought since the beginning of the Web for its openness: that anyone can read Web pages with any software running on any hardware. This is what makes the Web itself. This is the environment into which so many people have invested so much energy and creativity. When I see any Web site claim to be only readable using particular hardware or software, I cringe ..."

Little has changed. Not long afterwards, Microsoft updated its website with code that determined

which browser was being used to access it. Opera users were treated to a page whose formatting was ugly and unusable; users of Internet Explorer got the pretty one (in a hilarious protest, Opera responded by releasing a browser that, when pointed to Microsoft's web page, would replace all the page's text with the language of the Muppet Show's Swedish chef: "bork bork bork!") The point was taken.

Web Pages

Microsoft's Front Page software is a most egregious offender. While it been marvelously effective in helping home users and small business develop simple web pages, it builds the pages with code that makes them render more clearly in Internet Explorer. And if you want them to include page counters, search scripts, and the like, then you are forced to find a provider that will host your website on a server that offers "Front Page extensions" (which only run on Microsoft's IIS server software, which needs a Microsoft operating system ... get it?).

Insist on Open Standards

Transferring your email, updating your spreadsheets, rendering websites in a vendor-neutral way: it sounds petty, the result of corporate infighting. You should be able to choose the software you like or want and be done with it. But that lassitude is exactly what software makers are counting on to garner market share, and to make certain you remain a loyal customer, your data has been taken hostage: no insignificant allegation in the age of information. Browser wars aside, you should not have to be locked into an expensive upgrade cycle if you don't want to be, and you shouldn't have to upgrade to the latest version of some program just because all your cohorts have. The fact is, building document compatibility into software is trivial, and the only reason why a new version of the software won't read an old version's documents is "business." What defense does the ordinary citizen have?

The answer is simple but it's not easy: free and open standards. The internet was founded on clear, open, and commonly shared protocols. For that matter, developing compatibility in the early computing world (we're talking 1960s-1970s here) was of primary importance to the visionaries who sought to – and did – overcome incompatibilities in formats and protocols. It's due to their hard work and insight that we have an internet to unite us. How paradoxical then, that so many organizations are now trying their hardest to fragment us!

Here's the hard part: to put the consumer back in the driver's seat, simply don't use non-open protocols. In this sense, open means that the language with which a document is created is freely distributed and understood and any company that wants to use it in their software is free to do so. The opposite is proprietary protocols, over which one company maintains control and charges for the use of or doesn't permit the use of at all. For documents that are mostly text, RTF ("rich text") is a document standard that any word processor anywhere can read and write and is the only format you can send to anyone with the confidence they'll be able to read it. For more complicated things, only PDFs are considered open because anyone can read them on any computing platform. For images, JPGs and PNGs are open. For music, MP3s are popular enough to be almost everywhere but they're not open. Oggs are open, WMFs are not. The problem is that in many areas of computing no open protocols exist. This is the fight consumers must continue to wage, for until we insist on open and free formats and protocols, everything we write, calculate, or create is made available to us – our own work – only as long as the corporations that designed our software decide it's worth their while to stay in business. Is that important? Go ask the folks who used to use Wordstar.