

Re: OT: 2D Barcodes on Junk Mail

Source: <http://sci.tech-archive.net/Archive/sci.electronics.design/2008-08/msg00309.html>

- *From:* mpm <mpmillard@xxxxxxx>
 - *Date:* Sat, 2 Aug 2008 09:19:58 -0700 (PDT)
-

On Aug 2, 3:10am, Robert Baer <robertb...@xxxxxxxxxxxxx> wrote:

ý ýMost of these stupidites are on forms where one is requesting information.

ý ýOne would think that a requester knows their own e-mail address and if reasonably desirous of an answer, that they would enter their e-mail correctly.

The problem is that validating an email in software is actually not that trivial.

Obviously, you can check for ANY entry, that's easy.

If the field is left blank, the Server (or local script) can alert the user. No big deal.

But once you get past that, it becomes a little tricky.

Some characters are invalid in email addresses, so you have to trap for those.

An email must contain a single "@" sign, and at least 1 dot (and possibly more), but not two adjacent dots. It must also end in a valid suffix, like "com", "net", "us", "biz", etc.... And more are being added all the time, so you would have to keep adjusting your server or script code to handle that possibility. Finally, the email address can contain no whitespaces, such as tabs, spaces, etc...

Then you have some characters which, while not prohibited in the address, can confuse the hell out of software (both server & client side, plus everything along the route.) These are characters like:

0x26H – "&" – which is a UNIX shell special character and must be URL encoded anyway.

0x27H – The apostrophe, (which might get lost in subsequently quoted text).

0x2AH – The "*" asterisk, which is often used as a wildcard

0x2CH – The comma, which is used to separate email addresses from each other

0x2FH – The forward slash

...you get the idea, and if not, here's a listing and complete explanation of some of the potentially confusing aspects of these and other characters when used in an email address:

<http://www.remote.org/jochen/mail/info/chars.html>

So, given all of this, it's often easier (in software – either server side or local script) to just ask the respondent to enter their email address information twice and do a simple string comparison. If either field is missing, or if the two fields don't match – that's usually good enough to flag as an error.

Mike is absolutely right. We collect info from our customers all the time. We often get bad info on forms. You would be amazed how many people use a URL for a valid email address, for example. (Some folks evidently don't know the difference!)

–mpm

.