Unicode Domains are bad

and you should feel bad for supporting them

Posted by vgrsec on February 19, 2017

Introduction

I'm going to begin by caveating my opening statement by saying unicode domains improve accessibility to the internet, and that's a good thing, just unicode is so broad, there are many opportunities for lookalike domain spoofing, and that's bad.

I discovered during a discussion with @jaredhaight that unicode domains were a thing. We immediately joked about how bad this was, so I went about registering some test domains and ran some test cases to determine how well they were supported across various ecosystems. The following is an exploration of unicode domain names and how they're interpreted across various platforms as of Feb 2017.

Also thank you to @scholar__99 for help with the Android screenshots.

TLDR: Conclusion

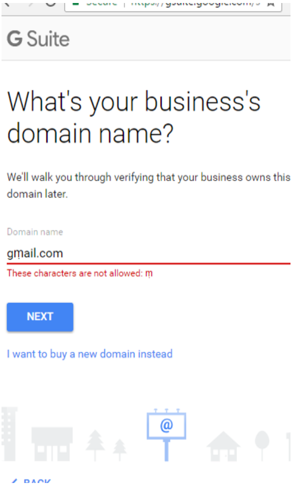
My conclusion is this: unicode support is broad enough to make for some vicious phishing campaigns. There's some effective mitigations, googles outright blockage of unicode domains is probably the most aggressive, Microsoft's conversion to ASCII is functional. One thing I didn't test, is if my local language settings matched that of the unicode character used, if the unicode character would show instead. I think ultimately that's the only way to balance accessibility with security. IE: if my device is EN-US the character set displayed should only be EN-US characters.

Setup

So, first of all, not all registrars are setup to support unicode domain registration. I'm not going to suggest one over the other, it's just something to be aware of. (Cough, Dreamhost, my preferred hosting provider, does not)

All of the test cases you will see were performed with https://www.gṃail.com (Unicode) OR https://www.xn--gail-qd5a.com (ASCII) depending on unicode support. In case it's not clear **ṃ** is the naughty character in this story and of course, I'm obviously spoofing gmail.com.

I tried to find an email provider who’d let me roll with a Unicode domain, Microsoft, Google, and Zoho all said No, Yandex said yes, but failed after I tied the domain to their service. However, here’s something interesting, Google and Zoho's font calls out that my **ṃ** is not a standard character.



G Suite

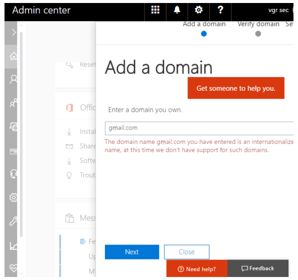
What's your business's domain name?

We'll walk you through verifying that your business owns this domain later.

Domain name

gṃail.com

These characters are not allowed: ṃ

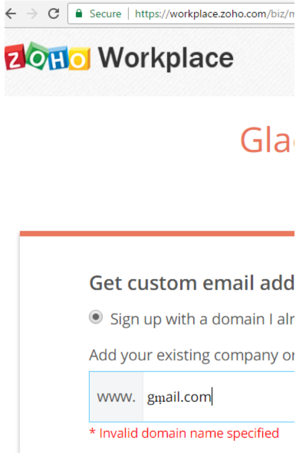


Add a domain

Enter a domain you own.

gṃail.com

The domain name gṃail.com you have entered is an internationalize name, at this time we don't have support for such domains.



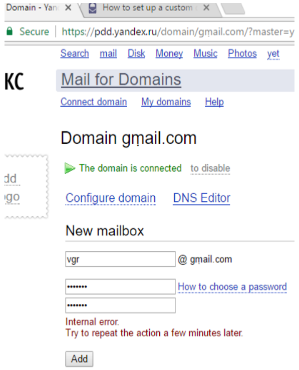
Get custom email add

○ Sign up with a domain I alr

Add your existing company or

www. gṃail.com

* Invalid domain name specified



Domain gṃail.com

The domain is connected to disable

New mailbox

vgr @ gṃail.com

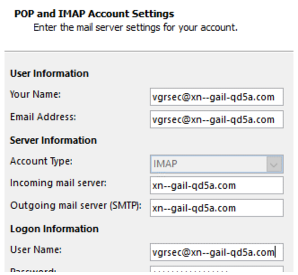
••••••• How to choose a password

•••••••

Internal error.
Try to repeat the action a few minutes later.

Thankfully, despite the big players saying no, the hosting company I registered my domain with, said yes.

I set up Outlook 2016 as my email platform, the auto configuration settings however converted my unicode domain to the ascii equivalent, however once the auto configuration wizard was done, I was able to go back into account settings and use the unicode domain.

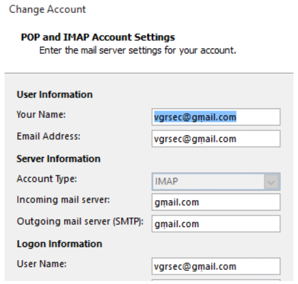


POP and IMAP Account Settings
Enter the mail server settings for your account.

User Information
Your Name: vgrsec@xn--gail-qd5a.com
Email Address: vgrsec@xn--gail-qd5a.com

Server Information
Account Type: IMAP
Incoming mail server: xn--gail-qd5a.com
Outgoing mail server (SMTP): xn--gail-qd5a.com

Logon Information
User Name: vgrsec@xn--gail-qd5a.com



Change Account

POP and IMAP Account Settings
Enter the mail server settings for your account.

User Information
Your Name: vgrsec@gṃail.com
Email Address: vgrsec@gṃail.com

Server Information
Account Type: IMAP
Incoming mail server: gṃail.com
Outgoing mail server (SMTP): gṃail.com

Logon Information
User Name: vgrsec@gṃail.com

I then went about testing different clients and email providers for support.

Receiving email from a unicode domain (Websites)

All emails sent from Outlook 2016, Web browser support tested in Chrome 56 for Windows. Note I've bundled outlook.com and office 365 support because they reacted the same way, no reason to provide an extra set of screenshots.

Google

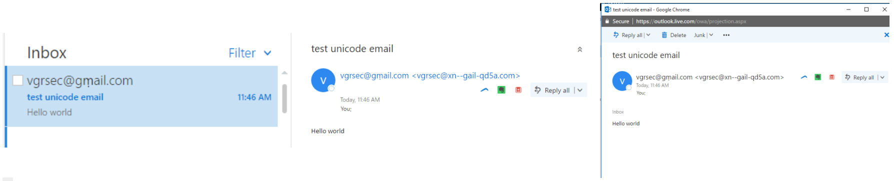
Emailing gmail from my unicode domain failed hard, some searching confirms this has been true since at least 2015.

```
vgrsec@gmail.com
host gmail-smtp-in.l.google.com [74.125.127.27]
SMTP error from remote mail server after end of data:
550-5.7.1 [184.172.62.109] Our system has detected that this
message is
550-5.7.1 attempting to use a domain name that does not meet our
domain name
```

```
550-5.7.1 policies. Please visit
550-5.7.1 https://support.google.com/mail/?p=DomainNameError to
review our Bulk
550 5.7.1 Senders Guidelines. m62si1920168oig.260 - gsmtp
```

Microsoft - Outlook & Office 365

This worked. Outlook/Office 365 displayed unicode in the inbox, however, once in the preview pane, or in the email itself it showed the name in unicode, but the email address as ASCII (Screenshots below)



Inbox

vgrsec@gṃail.com
test unicode email 11:46 AM
Hello world

test unicode email

vgrsec@gṃail.com <vgrsec@xn--gail-qd5a.com>
Today, 11:46 AM
You;

Hello world

test unicode email - Google Chrome
Secure | https://outlook.live.com/owa/projection.aspx

test unicode email

vgrsec@gṃail.com <vgrsec@xn--gail-qd5a.com>
Today, 11:46 AM
You;

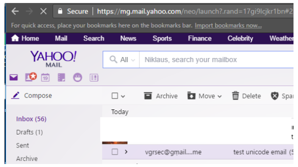
Inbox

Hello world

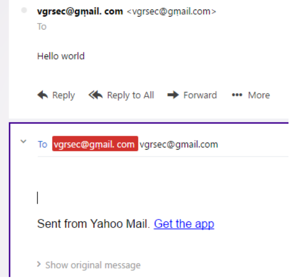
> *Note Outlook/Office 365 was able to respond to the unicode domain.*

Yahoo!

This worked. Yahoo displayed unicode everywhere! The email address, the name, everywhere. (Screenshots below)


Secure | https://mg.mail.yahoo.com/neo/launch?.rand=17gi9lcjkr1bn#2
For quick access, place your bookmarks here on the bookmarks bar. Import bookmarks now...
Home
Mail
Search
News
Sports
Finance
Celebrity
Weather
YAHOO! MAIL
All
Niklaus, search your mailbox
Compose
Archive
Move
Delete
Spam
Today
Inbox (56)
Drafts (1)
Sent
Archive

vgrsec@gṃail....me test unicode email

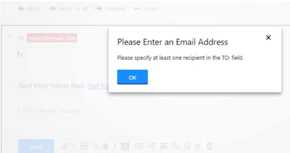


vgrsec@gṃail. com <vgrsec@gṃail.com>
To

Hello world

To vgrsec@gṃail. com vgrsec@gmail.com

Sent from Yahoo Mail. Get the app

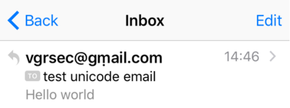


Please Enter an Email Address

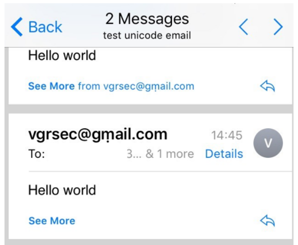
Please specify at least one recipient in the TO: field.

> *Note the second vgrsec@gmail.com was a test to determine why the email was highlighted in red*
>
> *It turns out that if you try to reply to a unicode domain'd email in yahoo, it fails. That red wasn't a security warning, it was a validity warning.*

Receiving email from unicode domain (Apps)



vgrsec@gṃail.com 14:46
test unicode email
Hello world



Hello world

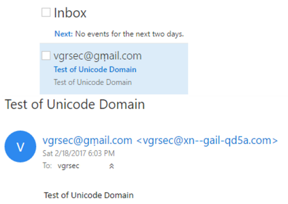
See More from vgrsec@gṃail.com

vgrsec@gṃail.com 14:45
To: 3... & 1 more Details

Hello world

See More

> *The unicode support in iosmail is great, just the best.*



Next: No events for the next two days.

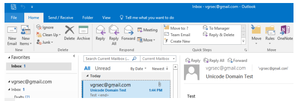
□ vgrsec@gmail.com
Test of Unicode Domain
Test of Unicode Domain

Test of Unicode Domain

V vgrsec@gmail.com <vgrsec@xn--gail-qd5a.com>
Sat 2/18/2017 6:03 PM
To: vgrsec

Test of Unicode Domain

> *The unicode support in iosmail is great, just the best.*



vgrsec@gmail.com
Unicode Domain Test 1:44 PM
Test <end>

vgrsec@gmail.com 'vgrsec@gmail.com'
Unicode Domain Test

Test

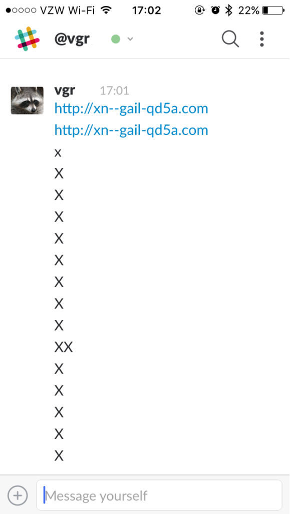
> *Since Windows applications largely rely on system fonts it's not surprising the unicode domain is supported throughout Outlook.*

URL Support

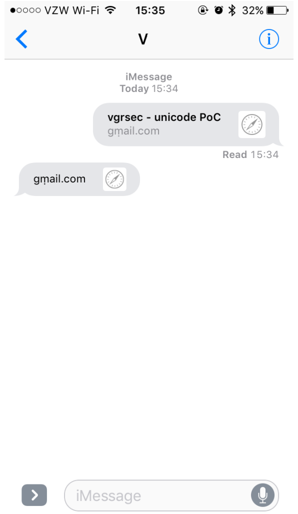
URL Support for unicode domains appears to vary wildly. I'll sum up some tests and findings, then proceed with screenshots.

App-OS	**Findings**
Slack-ios	Converts unicode into ASCII
iMessage-ios	Rrenders the unicode domain and provides a preview of the link
Text Message-ios	Renders unicode domain.
GChat-ios/web	Limited support for unicode domain.

	The link appears to work for a short time, then the link breaks at the unicode character
Safari-ios	Renders unicode domain
Chrome-ios/Windows/Android	Renders unicode domain
Firefox-Windows	Renders unicode domain
IE11	Converts unicode into ASCII
Text Message-Android 7.1	Renders unicode domain
Ghostery-Android	Converts unicode into ASCII



vgr 17:01
http://xn--gail-qd5a.com
http://xn--gail-qd5a.com
x
X
X
X
X
X
X
X
X
XX
X
X
X
X
X

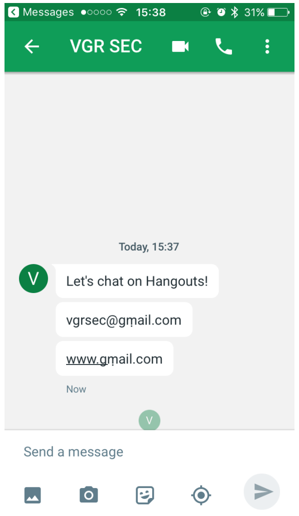


iMessage
Today 15:34

vgrsec - unicode PoC
gmail.com

Read 15:34

gmail.com



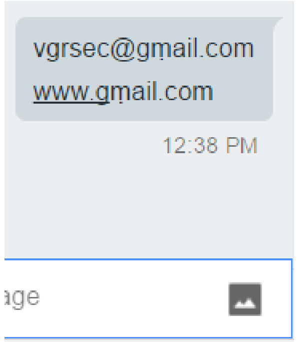
Today, 15:37

Let's chat on Hangouts!

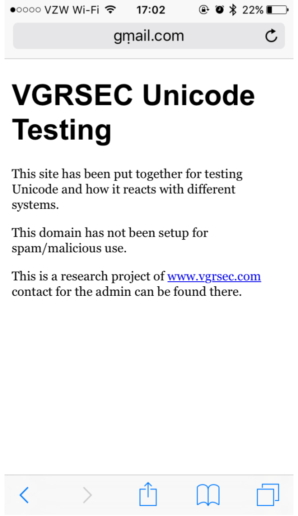
vgrsec@gmail.com

www.gmail.com

Now


vgrsec@gmail.com
www.gmail.com

12:38 PM

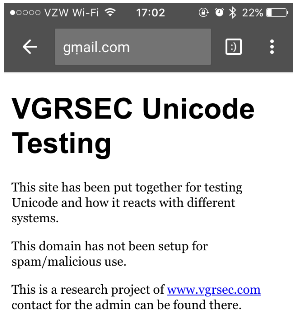


VGRSEC Unicode Testing

This site has been put together for testing Unicode and how it reacts with different systems.

This domain has not been setup for spam/malicious use.

This is a research project of www.vgrsec.com contact for the admin can be found there.

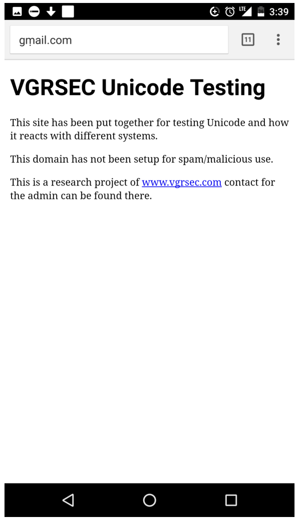


VGRSEC Unicode Testing

This site has been put together for testing Unicode and how it reacts with different systems.

This domain has not been setup for spam/malicious use.

This is a research project of www.vgrsec.com contact for the admin can be found there.

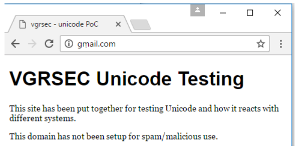


VGRSEC Unicode Testing

This site has been put together for testing Unicode and how it reacts with different systems.

This domain has not been setup for spam/malicious use.

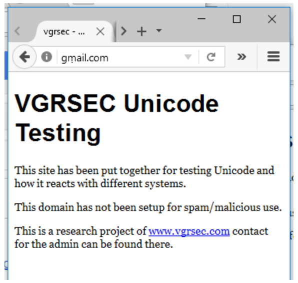
This is a research project of www.vgrsec.com contact for the admin can be found there.



VGRSEC Unicode Testing

This site has been put together for testing Unicode and how it reacts with different systems.

This domain has not been setup for spam/malicious use.

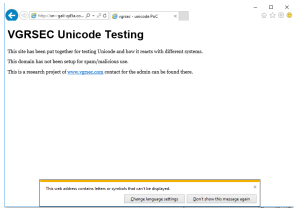


VGRSEC Unicode Testing

This site has been put together for testing Unicode and how it reacts with different systems.

This domain has not been setup for spam/malicious use.

This is a research project of www.vgrsec.com contact for the admin can be found there.



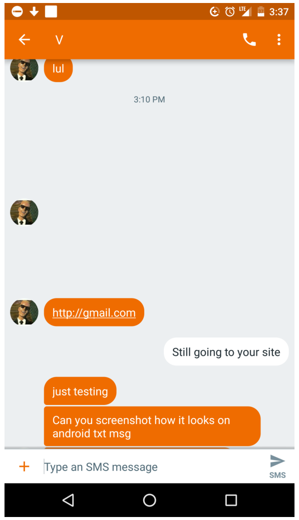
VGRSEC Unicode Testing

This site has been put together for testing Unicode and how it reacts with different systems.

This domain has not been setup for spam/malicious use.

This is a research project of www.vgrsec.com contact for the admin can be found there.

This web address contains letters or symbols that can't be displayed.



lul

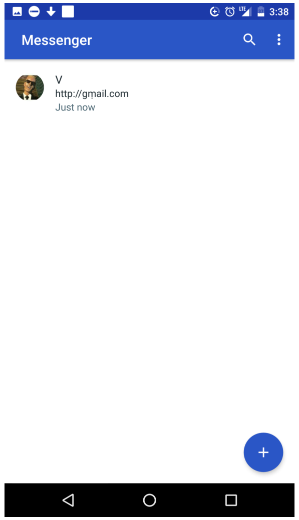
3:10 PM

http://gmail.com

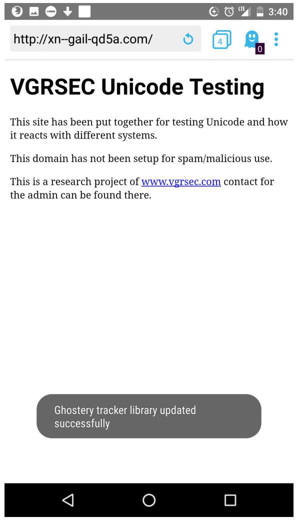
Still going to your site

just testing

Can you screenshot how it looks on android txt msg



V
http://gmail.com
Just now



VGRSEC Unicode Testing

This site has been put together for testing Unicode and how it reacts with different systems.

This domain has not been setup for spam/malicious use.

This is a research project of www.vgrsec.com contact for the admin can be found there.

Ghostery tracker library updated successfully