It seems that the need for MX, DKIM, SPF, and DMARC records for modern email setups were just not enough acronyms (and setup tasks) for some folks, resulting in the creation of yet-another-acronym — BIMI, or, Brand Indicators for Message Identification. The goal of BIMI is to *“provide a mechanism for mail senders to publish a validated logotype that mail receivers can display with the senders’ messages.”*

The idea behind BIMI is to provide a visual indicator of the brand associated with a mail message; i.e. you’ll have an image to look at somewhere in the mail list display and/or mail message display of your mail client if it supports BIMI. This visual indicator is merely an image URL association with a brand mail domain through the use of a new special-prefix DNS TXT record. Mail intermediaries and mail clients are only supposed to allow presentation of BIMI-record provided images after verifying that the email domain itself conforms to the DMARC standard (which you should be using if you’re an organization/brand and shame on you if you’re not by now). In fact, the goal of BIMI is to help ensure:

* the organization is legitimate
* the domain names are controlled by the organization
* the organization has current rights to display the indicator

When BIMI validation is being performed, the party requesting validation is currently authorized to do so by the organization and is who they say they are.

If you’re having flashbacks to the lost era of when SSL certificates were supposed to have similar integrity assertions, you’re not alone (thanks, LE).

**What’s Really Going On?**

I’m not part of any working group associated with BIMI, I just measure and study the internet for a living. As someone who is as likely to use [alpine](https://en.wikipedia.org/wiki/Alpine_(email_client)) to peruse mail as I am a thick email client or (heaven forbid) web client, BIMI will be of little value to me since I’m not really going to see said images anyway.

Reading through all the BIMI (and associated) RFCs, email security & email marketing vendor blogs/papers, and general RFC commentators, BIMI isn’t solving any problem that well-armored DMARC configurations aren’t already solving. It appears to be driven mainly by brand marketing wonks who just want to shove brand logos in front of you and have one more way to track you.

Yep, tracking email perusals (even if it’s just a list view) will be one of the benefits (to brands and marketing firms) and is most assuredly a non-stated primary goal of this standard. To help illustrate this, let’s look at the BIMI record for one of the most notorious tracking brands on the planet, Verizon (in this case, Verizon Wireless). When you receive a BIMI-“enhanced” email from [verizonwireless.com](http://verizonwireless.com) the infrastructure handling the email receipt will look for and process the BIMI header that was sent along for the ride and eventually query a TXT record for default.\_[bimi.verizonwireless.com](http://bimi.verizonwireless.com) (or whatever the sender has specified instead of default — more on that in a bit). In this case the response will be:

v=BIMI1; l=<https://ecrm.e.verizonwireless.com/AC/Global/Bling/Images/checkmark/verizon.svg>;

which means the image they want displayed is at that URL. Your client will have to fetch that during an interactive session, so your IP address — at a minimum — will be leaked when that fetch happens.

Brands can specify something other than the default. selector with the email, so they could easily customize that to be a unique identifier which will “be you” and know when you’ve at least looked at said email in a list view (provided that’s how your email client will show it) if not in the email proper. Since this is a “high integrity” visual component of the message, it’s likely not going to be subject to the “do not load external images/content” rules you have setup (you do view emails with images turned off initially, right?).

So, this is likely just one more way the IETF RFC system is being abused by large corporations to continue to erode our privacy (and get their horribly designed logos in our faces).

Let’s see who are the early adopters of BIMI.

**BIMI Through the Alexa Looking Glass**

Amazon had stopped updating the Alexa Top 1m sites for a while but it’s been back for quite some time so we can use it to see how many sites in the top 1m have BIMI records.

We’ll use the zdnsr package (also on GitLab, SourceHut, BitBucket, and GitUgh) to perform a million default.\_bimi prefix queries and see how many valid BIMI TXT record responses we get.

library(zdnsr) # hrbrmstr/zdnsr on social coding sites

library(stringi)

library(urltools)

library(tidyverse)

refresh\_publc\_nameservers\_list() # get a current list of active nameservers we can use

# read in the top1m

top1m <- read\_csv("~/data/top-1m.csv", col\_names = c("rank", "domain")) # <http://s3.amazonaws.com/alexa-static/top-1m.csv.zip>

# fire off a million queries, storing good results where we can pick them up later

zdns\_query(

entities = sprintf("%s.%s", "default.\_bimi", top1m$domain),

query\_type = "TXT",

num\_nameservers = 500,

output\_file = "~/data/top1m-bimi.json",

)

# ~10-30m later depending on your system/network/randomly chosen resolvers

bmi <- jsonlite::stream\_in(file("~/data/top1m-bimi.json")) # using jsonlite vs ndjson since i don't want a "flat" structure

idx <- which(lengths(bmi$data$answers) > 0) # find all the ones with non-0 results

# start making a tidy data structure

tibble(

answer = bmi$data$answers[idx]

) %>%

unnest(answer) %>%

filter(grepl("^v=BIM", answer)) %>% # only want BIMI records, more on this in a bit

mutate(

l = stri\_match\_first\_regex(answer, "l=([^;]+)")[,2], # get the image link

l\_dom = domain(l) # get the image domain

) %>%

bind\_cols(

suffix\_extract(.$name) # so we can get the apex domain below

) %>%

mutate(

name\_apex = glue::glue("{domain}.{suffix}"),

name\_stripped = stri\_replace\_first\_regex(

name, "^default\\.\_bimi\\.", ""

)

) %>%

select(name, name\_stripped, name\_apex, l, l\_dom, answer) -> bimi\_df

Here’s what we get:

bimi\_df

## # A tibble: 321 x 6

## name name\_stripped name\_apex l l\_dom answer

##

## 1 default.\_… [ebay.com](http://ebay.com) [ebay.com](http://ebay.com) <https://ir.ebaystatic.com/p>… [ir.ebaystatic.com](http://ir.ebaystatic.com) v=BIMI1; l=<https://ir.ebayst>…

## 2 default.\_… [linkedin.com](http://linkedin.com) linkedin.… <https://media.licdn.com/med>… [media.licdn.com](http://media.licdn.com) v=BIMI1; l=<https://media.lic>…

## 3 default.\_… [wish.com](http://wish.com) [wish.com](http://wish.com) <https://wish.com/static/img>… [wish.com](http://wish.com) v=BIMI1; l=<https://wish.com/>…

## 4 default.\_… [dropbox.com](http://dropbox.com) dropbox.c… <https://cfl.dropboxstatic.c>… cfl.dropboxstatic.… v=BIMI1; l=<https://cfl.dropb>…

## 5 default.\_… [spotify.com](http://spotify.com) spotify.c… <https://message-editor.scdn>… message-editor.scd… v=BIMI1; l=<https://message-e>…

## 6 default.\_… [ebay.co.uk](http://ebay.co.uk) [ebay.co.uk](http://ebay.co.uk) <https://ir.ebaystatic.com/p>… [ir.ebaystatic.com](http://ir.ebaystatic.com) v=BIMI1; l=<https://ir.ebayst>…

## 7 default.\_… [asos.com](http://asos.com) [asos.com](http://asos.com) <https://content.asos-media.>… content.asos-media… v=BIMI1; l=<https://content.a>…

## 8 default.\_… [wix.com](http://wix.com) [wix.com](http://wix.com) <https://valimail-app-prod-u>… valimail-app-prod-… v=BIMI1; l=<https://valimail->…

## 9 default.\_… [cnn.com](http://cnn.com) [cnn.com](http://cnn.com) <https://amplify.valimail.co>… amplify.valimail.c… v=BIMI1; l=<https://amplify.v>…

## 10 default.\_… [salesforce.com](http://salesforce.com) salesforc… <https://c1.sfdcstatic.com/c>… [c1.sfdcstatic.com](http://c1.sfdcstatic.com) v=BIMI1; l=<https://c1.sfdcst>…

## # … with 311 more rows

I should re-run this mass query since it usually takes 3-4 runs to get a fully comprehensive set of results (I should also really use work’s infrastructure to do the lookups against the authoritative nameservers for each organization like we do for our FDNS studies, but this was a spur-of-the-moment project idea to see if we should add BIMI to our studies and my servers are “free” whereas AWS nodes most certainly are not).

To account for the aforementioned “comprehensiveness” issues, we’ll round up the total from 310 to 400 (the average difference between 1 and 4 bulk queries is more like 5% than 20% but I’m in a generous mood), so 0.04% of the domains in the Alexa Top 1m have BIMI records. Not all of those domains are going to have MX records but it’s safe to say less than 1% of the brands on the Alexa Top 1m have been early BIMI adopters. This is not surprising since it’s not really a fully baked standard and no real clients support it yet (AOL doesn’t count, apologies to the Oathers). Google claims to be “on board” with BIMI, so once they adopt it, we should see that percentage go up.

Tracking isn’t limited to a tricked out dynamic DNS configuration that customizes selectors for each recipient. Since many brands use third party services for all things email, those clearinghouses are set to get some great data on you if these preliminary results are any indicator:

count(bimi\_df, l\_dom, sort=TRUE)

## # A tibble: 255 x 2

## l\_dom n

##

## 1 [irepo.primecp.com](http://irepo.primecp.com) 13

## 2 [www.letakomat.sk](http://www.letakomat.sk) 9

## 3 [valimail-app-prod-us-west-2-auth-manager-assets.s3.us-west-2.amazonaws.com](http://valimail-app-prod-us-west-2-auth-manager-assets.s3.us-west-2.amazonaws.com) 8

## 4 [static.mailkit.eu](http://static.mailkit.eu) 7

## 5 [astatic.ccmbg.com](http://astatic.ccmbg.com) 5

## 6 [def0a2r1nm3zw.cloudfront.net](http://def0a2r1nm3zw.cloudfront.net) 4

## 7 [static.be2.com](http://static.be2.com) 4

## 8 [www.christin-medium.com](http://www.christin-medium.com) 4

## 9 [amplify.valimail.com](http://amplify.valimail.com) 3

## 10 [bimi-host.250ok.com](http://bimi-host.250ok.com) 3

## # … with 245 more rows

The above code counted how many BIMI URLs are hosted at a particular domain and the top 5 are all involved in turning you into the product for other brands.

Speaking of brands, these are the logos of the early adopters which I made by generating some HTML from an R script and screen capturing the browser result: