YOUR DATA FROM THE OUTSIDE - A QUICK LOOK

Are you worried about being hacked? If so you will have an IDS, firewalls, all sorts of cool and funky things. Nice. Really. But... (you knew that was coming) ... what about data that was put out there as a matter of course? What about data that was put out there as a part of normal business? What about data collection which will NEVER be recognised by your cool and funky things?

Lets start with email addresses. People looking to attack you love email addresses. They provide a link to possibly social engineer, possible user names, possible links to social sites for more research and much much more. And would'nt you know it, there is a tool that can help make this easy for you, its called "*theHarvester*". Original is it not. You can get it from *Edge-Security* (see here). It is a python script, and runs very nicely on linux. So just make sure you have python, download it, unextract it, and away you go.What you do is point it at a domain (*-d*), limit it if needed (*-l*) and tell it what to look at (*-b*). It is this last option that is interesting. Lets take a look at a simple run, we are going to look at Verizon. A large ISP company, very public.

```
# ./theHarvester.py -d verizon.com -l 100 -b google

*************************************
*TheHarvester Ver. 1.4b             *
*Coded by Christian Martorella      *
*Edge-Security Research             *
*cmartorella@edge-security.com      *
*************************************



Searching for verizon.com in google :
======================================

Total results:  10800000
Limit:  100
Searching results: 0

Accounts found:
====================

rick.myers@verizon.com
mary.delagarza@verizon.com
dana.b.russell@verizon.com
WITS.2001.CSC@verizon.com
====================
```

So we limited it to looking through *google* and to a *100* results. We found 4 and 3 are promising. Not quite the haul we wanted? Lets change it a bit..

```
# ./theHarvester.py -d verizon.com -l 100 -b pgp

*************************************
*TheHarvester Ver. 1.4b             *
*Coded by Christian Martorella      *
*Edge-Security Research             *
*cmartorella@edge-security.com      *
*************************************



Searching for verizon.com in pgp :
```

```
======================================

dara.l.vaughn@verizon.com
gregory.truley@verizon.com
joshua.c.bedgood@verizon.com
michael.i.dacek@verizon.com
joseph.j.snyder@verizon.com
jeremy.b.carrier@verizon.com
michal.m.matalon@verizon.com
adam.i.barrow@verizon.com
daniel.wood@verizon.com
jim.mcconnell@verizon.com
donn.janes@verizon.com
airdog07@verizon.com
FlyFree42@verizon.com
daniel.lashua@verizon.com
daniel.lashua@verizon.com
william.g.kastner@verizon.com
anthony.bragano@verizon.com
christopher.s.juffre@verizon.com
Michael.McNeil@verizon.com
andrew.leeper@vol.verizon.com
roland.shepard@vol.verizon.com
donald.richardson@verizon.com
william.fredholm@verizon.com
danielj7@verizon.com
>justin.rummel@verizon.com
justin.rummel@verizon.com
jonathan.d.barton@verizon.com
cjackson@vol.verizon.com
rick.dodd@verizon.com
gstubblefield@verizon.com
ron.schekman@verizon.com
gstubblefield@verizon.com
damon.j.small@vol.verizon.com
steven.p.sobieski@verizon.com
gstubblefield@verizon.com
patrick.babb@verizon.com
ernest.w.horne@verizon.com
steven.king@verizon.com
jl.allen@verizon.com
scott.rowe@verizon.com
stuart.moore@verizon.com
manoj.vaddineni@verizon.com
jack.fultz@verizon.com
lightiam@verizon.com
martin.wind@verizon.com
jason.curtis@verizon.com
ravin.cheela@verizon.com
cheryl.j.mitchell@verizon.com
alfredo.camacho@verizon.com
jayvi13@verizon.com
richard.wince@verizon.com
troy.brown14@verizon.com
```

```
Mary.C.Akers@verizon.com
remlaps@verizon.com
wesley.eddy@verizon.com
william.ingram@verizon.com
jeff.wood@verizon.com
jeff.wood@verizon.com
marwan.badran@verizon.com
jeff.wood@verizon.com
daniel.murawinski@verizon.com
scott.mckinley@verizon.com
nicholas.flood@verizon.com
steven.boles@verizon.com
doug.heatherly@verizon.com
john.doleac@verizon.com
william.kight@verizon.com
William.Kight@verizon.com
joshua.elson@verizon.com
joshua.elson@verizon.com
joshua.elson@verizon.com
```

So we limited it to a *100* results again, but this time we looked at the *pgp* servers, after all if you want to be secure you use *pgp* and publish your email address. This time we get a few more results then 3. Even being secure can have unintended consequences.

But lets move on from emails and Verizon. Lets turn our attention to *documents*. Documents all have *metadata*. We all know this, but do we really check for this? Do we really know what we have out there? You can bet the bad guys do, they use many tools. Lets look at another one from our friends at Edge-Security called *Metagoofil*. This is another python script that goes through a website looking for common documents and then checks them for data. You can download the tool from the same place as above. Some *gotchas* though, you will also need the "*libextractor*" libraries, and in the version I used I needed to add the "*#!/usr/bin/env python*" line to the top of the python script before it worked. So once everything is fine, lets give it a run, this time we will try Apple. Spread the love. Similar to the first tool, we specify the domain (*-d)*, the limit of files to check (*-l*), the types of files (*-f*), the output file (*-o)* and the storage directory (*-t*)...

```
# ./metagoofil.py -d apple.com -l 10 -f all -o apple.html -t apple

*************************************
*MetaGooFil Ver. 1.4a               *
*Coded by Christian Martorella      *
*Edge-Security Research             *
*cmartorella@edge-security.com      *
*************************************



[+] Command extract found, proceeding with leeching
[+] Searching in apple.com for: pdf
[+] Total results in google: 26700
[+] Limit:  10
[+] Searching results: 0
        [ 1/20 ] http://lists.apple.com/attachments/pdfFqYLuAulgF.pdf
        [ 2/20 ] http://hotdeals.apple.com/compusa/pdf/easyshare.pdf
   ...<all the other results>
[+] Searching in apple.com for: doc
[+] Total results in google: 297
[+] Limit:  10
[+] Searching results: 0
[+] Directory apple already exist, reusing it
        [ 1/20 ] http://edcommunity.apple.com/ali/galleryfiles/431/cher.doc
        [ 2/20 ] http://edcommunity.apple.com/ali/galleryfiles/473/
Evaluating_Lesson_Plan.doc
  ...<all the other results>
[+] Searching in apple.com for: xls
[+] Total results in google: 29
[+] Limit:  10
[+] Searching results: 0
[+] Directory apple already exist, reusing it
        [ 1/20 ] http://edcommunity.apple.com/ali/galleryfiles/17526/participati
        [ 2/20 ] http://www.apple.com/storeassociates/productspecs/specs.xls
..<all the other results>
[+] Searching in apple.com for: ppt
[+] Total results in google: 29
[+] Limit:  10
[+] Searching results: 0
[+] Directory apple already exist, reusing it
        [ 1/20 ] http://edcommunity.apple.com/ali/galleryfiles/381/ASUApple.ppt
        [ 2/20 ] http://edcommunity.apple.com/ali/galleryfiles/9946/Oblinger.ppt
..<all the other results>
[+] Searching in apple.com for: sdw
[+] Total results in google: 0
[+] Searching in apple.com for: mdb
[+] Total results in google: 0
[+] Searching in apple.com for: sdc
[+] Total results in google: 0
[+] Searching in apple.com for: odp
[+] Total results in google: 0
[+] Searching in apple.com for: ods
[+] Total results in google: 0


Usernames found:
================
Â³Â¬Â¦4Ã
        ÂÃÂxÃXÂÃ
icrosoft Excel: LaserWriter 8 8.6
Karen Kelton
Randy Yerrick
Marcie Talia-Rice
JMcLaughlin
Dianne Lawrence
School District #36 (Surrey)
Office 2004 Test Drive User
Kevin  Amboe
hayes_k
David L. Curry
Curry103
Megan Iemma
Solon Community School District
Jerry Villa
Brett Fernald
```

```
Calgary Science School
Ampang Campus
SIS
Lance
lkirby
M Bourgeois
baily
Helena Paulin
cbk1
Carol Kline 301-834-8291
student
...
Euclid City Schools
Douglas County School District Re. 1
John Turnbach
Art Fichter
Julie Martin
Michelle Kolb
majac1
Donna L Ernst 301-663-4382
Carol  Kline
Carol Kline 1-800-352-8276, ext. 8278
Joanna Matthews
Thomas Scherer
Don Knezek
Lajeane Thomas
Diana Oblinger
Timothy E. Drew
HollidayG
Phil Hemmings
Administrator
PBHS PBHS
susan atlas
Steven Padilla
dorina
Dorina Kosztin
Mark Jarvis
College of Education
Mike McKean
tietzj
Joan Tietz
LICEO
Paul Resta
Ken Tothero
WRHS
Lunchbox   Companies
Danna Vessell
et
SCHOOL


Paths found:
============
 Normal\
Macintosh HD:Documents:Auto recovery Word docs:Word Work File A_214\
Macintosh HD:Desktop Folder:arrows:Science and Math URLs\
Macintosh HD:Documents:Auto recovery Word docs:Word Work File A_870\
 Macintosh HD :Microsoft Office 98:Templates:Blank Presentation\
 Watermark\
 Macintosh HD:Applications:Microsoft Office X:Templates:Presentations:Content:Generic\
 Curtain Call\
 Mountain Top\
 Pixel\
 Ocean\
 Solstice\
 Beam\
[+] Process finished
```

So we got the names of some users, some usernames, places of work, even one phone number, version of software used, and of course the paths show they use Macs. All in all, some useful information, and if you really want it in a nice format, take a look at the html file generated. Is this all for people working at Apple? Probably not. Does that make it any less scary? No. So what have we seen with this brief jaunt through the web. It is very easy for the bad guys to find out this type of stuff without you ever knowing they had done so. It is very easy for users to put this type of stuff out there ... unfortunetly also without you ever knowing they had done so. So, do you know what the rest of the world could know about your company? Everything we have seen so far is entirely in the public domain, there to look at by whoever wants to.

Final thoughts

These are nice tools, but all they do is automate what people were doing anyway. Their checks are by no means exhaustive but they do make the point - information about your network can and probably does exist outside of your network, or in the best case, on your public resources where you can still do something about it. But as we see by looking at the apple site, sometimes this information is way outside of your control. Regardless, you should at least know about it. As always, try it out, have fun and learn.