How Pastebin Can Help with Research

In this tutorial, we will show you how researchers can use information they find on Pastebin to locate hackers who are offering leaked data for sale.

What is Pastebin?

Pastebin.com is often compared to a clipboard - it's a place to paste anything, like plaintext documents, logs, source code, etc. for anyone to view. Pastebin is also an infamous repository of stolen databases, PoC exploit code, combo lists, doxing victims, and credit card numbers - all available for sale. Publishing information on Pastebin requires no login, and it's been popularized throughout the hacker community through the use of internet relay chat.

Pastebin does its best to remove sensitive information, but with millions of active pastes, moderating it is an overwhelming task. Pastebin is often the first stop researchers go to when they look for stolen information when leaks surface.

Using Pastebin to Hunt for Stolen Data

Sometimes hackers boast about the data they possess by uploading samples on Pastebin and then offering links to the full dumps. These links point to anywhere from Torrent sites, like The Pirate Bay, to a variety of darknet .onion marketplaces, where stolen data can be purchased.

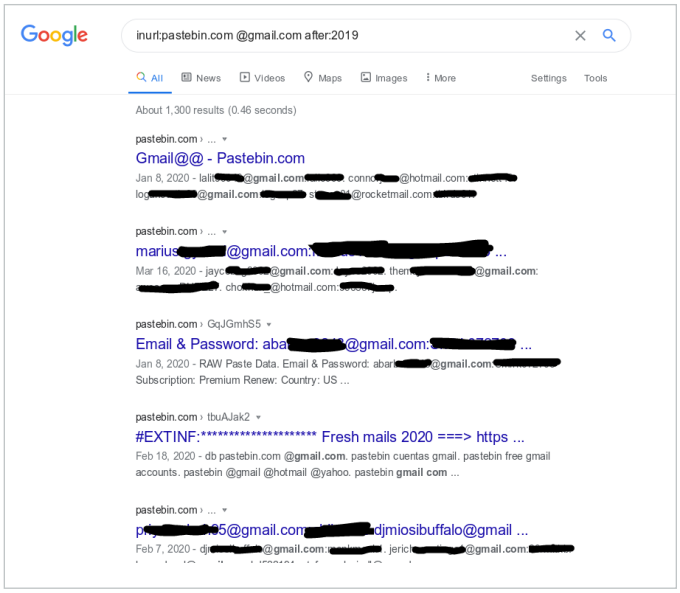
It's crucial that during this hunt, sessions remain unique and untied from normal browsing. This is because of the tracking that occurs on Google (when dorking for pastes) and on external sites (that researchers may be required to visit for full dumps) listed within the paste.

When we use Google to search for pastes from pastebin.com, we use the inurl: Google search parameter to do this. Your search should be formatted as follows:

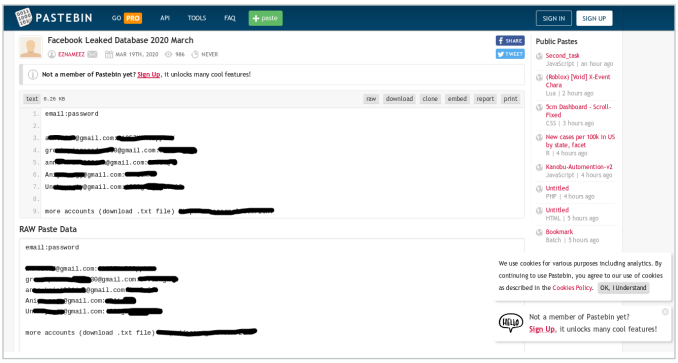
inurl:pastebin.com <SEARCH PARAM>

When looking for stolen databases, you can start by searching for standard email providers:

- @gmail.com
- @live.com
- @hotmail.com
- @yahoo.com



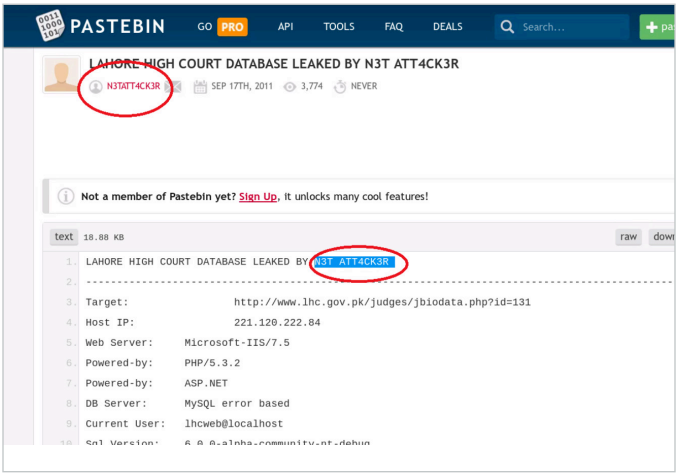
Searching for emails helps find resurfaced stolen databases because these data sets contain leaked credentials belonging to these providers (see screenshot below). To get the latest results first, we recommend setting the Google dork to sort by date. To do this, we use after: and before: search parameters, followed by the date (ex. after:2018-12-31 or after:2018):



Simply searching for the term **"database"** will return a number of stolen databases, accompanied by links to full dumps when the data set is too large for Pastebin to host. Alternatively, the link may redirect to an escrow or a marketplace forum, where a transaction must be completed before accessing the full data set.

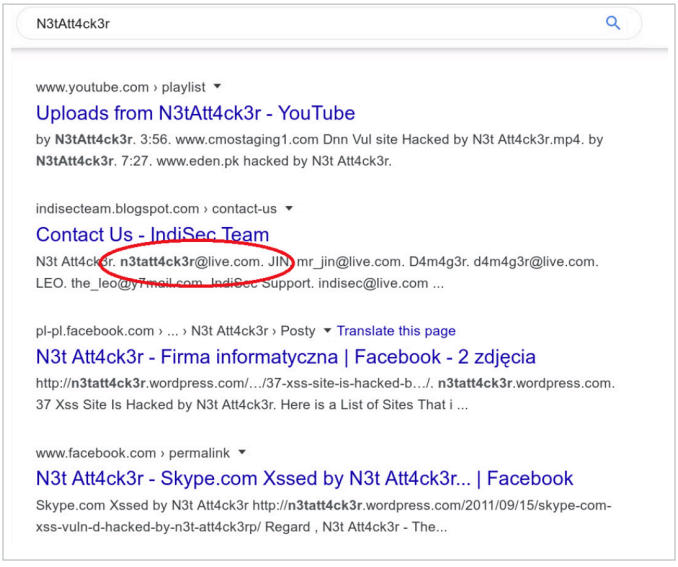
This particular URL redirects to rocketr.net, one of many publically available escrow services for trading currency for digital commodities, such as stolen data. Another popular escrow is satoshibox.com.

Searching for Visa, Mastercard, "cc dump," etc., helps find large credit card dumps that link back to carding forums where they are typically sold in bulk for a small price. Similarly, searching "dox" or "doxed" reveals lists of victims that have had their private information posted online. These individuals could easily become victims of financial fraud in the future, especially if the dox contains their social security numbers, credit card details, etc.



Using Pastebin to Hunt for the Author of the Stolen Data

So far, we've learned how to find various forms of stolen data on Pastebin, but how can you attribute it to the real identity of a person who has actually leaked the goods? Sometimes (not always), pastes contain attribution in the form of an alias. People who leak information at times like being known as the root cause. Yes, some leakers like attribution, so they can use their reputation to increase sales of stolen data.



Leakers tend to only release partial dumps of the data breach on sites like Pastebin as a form of advertisement. The more obvious approach is to link directly to a forum or marketplace where the full dump can be purchased or traded for sought-after commodities.

Some attackers consistently use the same aliases, so even a simple Google search will show their nicknames appearing on several different platforms, most likely hoping to generate more leads for their offerings.

An easiest way to figure out who is hiding behind an alias is to locate their email. Email addresses are used to sign into various sites, and inevitably sites get breached and their users' personal information gets stolen and placed inside a database.

If we want to pursue this investigation even further, we can get our hands on the database in which this user resides, and find other essential artifacts, like their IP addresses, phone numbers, and more.

According to HaveIbeenPwned.com, this particular person has been found in one data breach. In addition to their own password, the database also contains an IP address. A simple search can help determine the general location of the attacker, and law enforcement can go much further in identifying and locating this individual.

This example shows how a researcher can go from a leak posted on Pastebin to unveiling the identity of the perpetrator. A series of simple steps can help investigators get to the bottom of unmasking the hacker who is responsible for the leak, and if warranted, even taking legal action.