**FOOTPRINT USER STORY – Version 1**

**I3 (Triple I) Research and Investigation**

I am a i3 (Triple I) Researcher who uses a variety of OSINT search methods, and my role is to methodically research the internet and gather intelligence and information about a variety of ENITITIES, including, but not limited to:

* People
* Addresses
* Emails
* Phone Numbers
* Social media handles
* Images
* Videos
* Websites
* Dark Net
* Monitoring of Groups / Organisations
* Monitoring of Locations
* Monitoring of Events

Prior to conducting any i3 Investigation there is a requirement to complete a 3 to 4 days Level 2 open-source course called *Covert Open Source Internet for Investigators* (COSII) course. This is a mandatory requirement for all Standard and Specialist users. Specialist i3 Investigators go on to complete Level 3 and Advanced I3 courses to enhance they capabilities. Some users use LongArm (via IIS) which gives users secure access to the Internet. To get access to LongArm there is also a requirement to complete the COSII course.

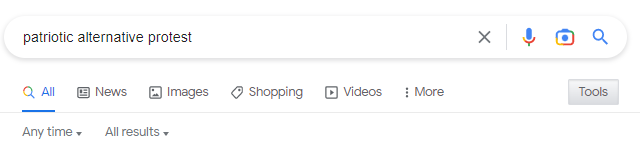
I am given several OSINT tasks a week to complete. I am normally only provided with partially known information and my role is to develop this information to expand our understanding of the issue, primarily to help safeguard the public. I am very busy in my job, and I am not only limited by time, but the role can also be very stressful and require me to juggle multiple tasks at one time.

The OSINT world changes weekly, and what is useful and available online now might not be useful or available tomorrow. I am constantly discovering new methodologies and tools to conduct my OSINT work. Because of the scale of information available online I use a wide variety of online resources, methodologies, and databases to gather the details I require. I like to think of each task as like a jigsaw puzzle and each piece of information I can find is like a piece of the puzzle. The challenge is that I don’t know how big the jigsaw puzzle is when I start, and I can never be sure which online tool will uncover a new piece of information. I can be lucky some days, but on others I end up having using a significant number of websites and tools, with no guarantee of success.

When I generally start a new task, I use a database called GBG Investigator. This is my favourite database, which provides me with valuable (and accurate) information that helps me identify lots of additional data. It allows me to search for several entities, like a person’s forename, middle name, surname (including partial information and ‘sounds like’ search), address (including partial address), telephone numbers and email addresses. For example, I can input partial person details with a location and the results will give me several potential people of interest. This enables me to investigate each of these people to identify which one is a more likely match. I may also have information on just an email address. GBG Investigator allows me to enter the email address and potentially identify the person associated to this email address. The only problem I have is that we only have one GBG ‘Investigator’ licence in my team and if someone else is using it I must wait until they finish. If I accidentally log onto it and someone else is using it, I automatically log them out. I wish we all had our own licences.

I then start searching using search engines. I do not rely on just one search engine. It is more optimal to be able to search on multiple search engines, instead of relying on just one, as different search engines bring back different results. Also, one search engine may be used more because of its User Interface. For example, my ‘go to’ search engine is Google as it has a clean, simple interface making it easier for me to use this as default. I rely on using the Advanced Search Options or ‘Google Dorks’ to narrow search results to specific file types, strings, and websites. Where this can be automated it would make my i3 Investigations a lot more efficient.

I sometimes search for an entity on a search engine and use the customised options below the search engine (Google search categories)which allows me to refine my search by specific categories, like Images, News, Shopping, videos and so on. I may also use the TOOLS filter option which allows me to narrow my search to a specific time period. For example, I may only be interested in finding information about a group (Patriotic Alternative) but in the last year or last month. This allows me to be a bit more focussed. (see below image)



I also need to be able to view more than 10 results as this limits me considerably. If I was trying to identify a person, I will be looking at viewing up to, or more than 40 search results.

I expect my results to be displayed on screen fast and in a structured format so that I can methodically go through each entity and assess its relevance.

If I identify something of relevance, I should be able to conduct a reverse search on that entity by a click of a button. For example, downloading an extension that allows me to conduct a reverse image search means I can click my mouse on a relevant image and select reverse image without having to save the image and then import this into the system as this can be time consuming. I may come across additional relevant entities which I should be able to highlight and reverse search without having to go into the main screen. Anything that requires me to do minimal clicks makes me work more efficiently.

Some of the resources I use are free third-party tools, and some are subscription based.

I really like many of the free 3rd party tools as they have specific purposes that help me discover hard to find information. They are normally very easy to access (sometimes using my false online persona credentials, which I don’t mind), are very slick and easy to navigate, with minimal lag or delays (I hate waiting for things to load onscreen, any unnecessary clicks, or anything that makes it hard to know how to use the tool).

Despite them being very useful, it concerns me that I often don’t know who is behind the 3rd party tools, and whether they use my search data inappropriately or illegally, including by foreign agencies. Being able to conduct i3 investigation in a safe environment is key to ensuring data is protected. I wish we could ensure that our searches were being done in a safe online environment and that I didn’t have to worry about the security of my data.

This knowledge and good practices are shared with colleagues.

Due to the volume of data I may obtain it can be very time consuming to inspect what is relevant, and in a fast paced, time critical, environment it is really efficient for me, where possible if this method of gathering the data is in one place and is automated. Automation can help connect different data that may be otherwise missed.

Where possible the method of data collection, for example, search fields, needs to simple and all the relevant fields need to be visible on one page. If anything is ‘hidden’ it may not be clear what field needs to be completed.

It is sometimes not clear what information is required in each field, a pop up / ‘information’ button which gives me a description of what is required in that field would make it that bit simple to conduct my search. Where possible the method of searching which requires less ‘clicks’ is more efficient.

The output of the data needs to be in a structured format so I can quickly go through the results and assess relevant information. Anything that is relevant I want to be able to click on it and it takes me outside of the product I am using to the actual live link would be hugely beneficial.

There may be occasions that I accidentally log myself out of Cosain or need to find my last steps, a history of this, the way you can find this on browser histories on Chrome and other search engines, allows me to go back and click on the links I have been to. This helps me a lot as there have been occasions when I have been unbale to find a particular link that in hindsight, is of interest to my investigation.

I will also need to keep a record of this information so that I am able to go back to the results if asked to so that I can demonstrate the steps taken to identify each relevant entity. More importantly, an audit of these results allows my supervision to see exactly what I have been doing. I would like this to be exportable to a filetype like Excel / CSV format so that I can utilise this within my internal databases instead of typing this out onto an excel spreadsheet or copying and pasting it.

**\*\*\*\*EVERYTHING BELOW IS A DRAFT VERSION\*\*\*\***

**PEOPLE**

To be able to investigate what you can identify about a person is a key task for any i3 Investigator. To do this I am interested in knowing about:

* Personal details like Name, Alias Images(s), DOB, Address, and Contact Details.
* Mapping the online footprint of the person of interest through search engine searching, as well as conducting checks on social media platforms and applications and checking other websites, forums and images.
* Understanding the person’s profile using the above identified details.

I may have either:

* Full name details (Forename, Middle and Surname)
* Partial information, for example forename and DOB or Alias Name with DOB.

When I am conducting i3 work on a person I am looking to identify a number of things:

Person’s full name and bio information.

Aliases – an Alias is also usually referred to as a username or pseudonym and often reused across different social media platforms. A person may have multiple Aliases to either stay anonymous or because they name may have multiple spellings which can then be used to identify social media accounts. The type of Alias or username the person is using can vary depending on how easily they want to be located and identified. I use the alias to identify potential social media accounts / platforms of users.

Fuzzy Searches - some names have variations / alternative spellings, for example the name Muhammad may also be spelt like this: Muhammed, Muhamed etc. Having the ability to do a ‘fuzzy’ (sounds like) search should be an option. For example, when I conduct checks on GBG Investigator it allows me to add a person’s name and conduct a fuzzy search.

Date of Birth – the date of birth is relevant to ensuring I find the right person as it indicates to me what they potential age is. Sometimes I may have information about the age range of the person instead of an exact date of birth, so

Sometimes I may have partial information about the person of interest alongside additional information about one of their family members. This is vital information. I have used this data before on databases like GBG Investigator where I have inputted in the first person field the main persons details and in the second persons details the second persons information which has then brought back information about people and address details where they both have been occupants at the same address.

I am also interested in knowing where the person of interest has lived previously as this information may give me additional information, like other family members, different contact details and so on.

I will also look to use some People Search Sites, like Pipl which was a free service, but you have to pay to get this data now.

I am also interested in breached databases as this may contain information about the person or his/her personal details which either makes them vulnerable to exploitation or helps further the investigation. For example, ‘Fuck Facebook’ includes 435,627,630 indexed items from a huge Facebook leak in 2019. This was originally only available on the darknet, it was made available on the surface web. There are various other breached databases, like Pastebin and <https://breached.to/>.

Searching a person using their real name can also deliver a lot of false positives so combining this information with additional information such as location is important to cut down any irrelevant results that do not match.

Throughout this i3 investigation I will also be looking to identify potential accomplices /associates and map out related criminal networks. Having an automated tools that allows me to identify the friends in common between associates would be a really valuable and efficient tool.

The output of the data needs to be in structured format so I can quickly go through the results and assess relevant information.

I will also need to keep a record of this information so that I am able to go back to the results if asked to so that I can demonstrate the steps taken to identify each relevant entity.

**ADDRESSES**

When conducting research on an Address I may have either Full information or Partial information on the following:

* House number
* Street name
* Town name
* County
* Postcode (or partial post code)

My main task is to identify the full address and who the current and previous occupants of the address are, including contact details or persons who is associated to the address (e.g. landlords; vulnerable’s; associates; shared accommodation and so on). I may also be interested in knowing what the persons previous addresses were and who the occupants were at that time.

I may use multiple search engines to conduct this search as different search engines will bring back different results. I may also use paid / nonpaid services / Tools to identify this information. For example, 192.com will give me some information about occupants.

Having some information about the person of interest’s address gives me an opportunity to narrow down the person’s social media accounts to identify a social media footprint. For example, I may know that person is called John Smith and I may find out that he lives in Doncaster. I can now search social media platforms and look for all John Smith’s who state in they bio that they are from Doncaster and potentially narrow my search or identify the person of interest.

The output of the data needs to be in structured format so I can quickly go through the results and assess relevant information.

I will also need to keep a record of this information so that I am able to go back to the results if asked to so that I can demonstrate the steps taken to identify each relevant entity.

**EMAILS**

As part of my investigation, I may get given an email address to investigate. This will involve identifying who is using this email address (who the account is attributed to) and if there are other users of this email address. An email address can assist me in identifying information about a person, where they work and live, what social networks they use and even where they shop or meet other people.

An email address will be unique to that person and will not deliver the same level of false positives. Like usernames, a person can have more than one email address and some people have many due to them using different ones for personal use and others for other purposes like for work or for nefarious activity. Email addresses can also contain data like names or birth dates within the text before the email domain.

I use multiple search engines to conduct this search as different search engines will bring back different results. I may also use paid / nonpaid services (Google extensions available from Google store, Epieos/Holehe tools etc) to identify this information. This can also be done by using a form of consented database. I may also use Google Dorks to do this:

* “@example.com” site:example.com finds all available email names on the site.
* HR “email” site:example.com filetype:csv | filetype:xls | filetype:xlsx finds contact lists of HR employees in a specific site.
* site:example.com intext:@gmail.com filetype:xls searches for email IDs for a domain from Google.
* intext:resume “john smith” finds resumes that can contain an email address.

I will also want to investigate if this email addresses is used to register on other platforms and websites. This is a crucial element of my i3 Investigation as it builds a bigger picture of what the person is about. A lot of the people I investigate will have some nefarious activity on their social media accounts. Attributed / associating an email address to an account may corroborate some sensitive intelligence.

I start my email investigation by interrogating parts of the email address. For example, I can search for the following:

* Full email address
* Username
* Domain

The things I would look at whilst investigating a Full email address are free tools relating to people search sites like spokeo.com. I used to use Pipl when it was free as it brought back some relevant results. However, this is now a paid service and I am unable to pay for this service at work. If there was a means to access the data from here that would assist in furthering my investigation. I would also be interested in attributing the email address to a social media account. I would do this by looking at 3rd party Tools. For example, Epieos/Holehe and Mozilla Firefox extensions like <https://addons.mozilla.org/es/firefox/addon/facebook-instagram-osint/> which is a Facebook and Instagram tool that identifies public email and phone number and obfuscated email and phone number as well as provide you with Facebook and Instagram ID’s. A lot of the useful attribution tool are on Github which I am not familiar with. I would also check if this email address features on GBG Investigator.

The Username is one of the most commonly leveraged pieces of information in OSINT research. Usernames follow us everywhere and have some level of consistency across different platforms and timeframes. I search in various ways including, but not limited to:

* Using multiple search engines – Google because it is more intuitive and gives you the most relevant information but also Bing.
* People search engines – e.g. Pipl but as mentioned before this is now a paid service.
* Username tools – whatsmyname, usersearch; namechk
* Social media platforms – Facebook, Instagram, Twitter, Gab, Gettr to name a few.
* Online – Facebook community pages; Reddit community pages.
* Message boards

I may also want to see if the email address has been leaked onto a Breached database. There have been several occasions where important information has been leaked by hackers. Sometimes this information sits on the darknet and sometimes this is available on the surface web on sites like Pastebin. So anyway, that can be automated would make the whole process more efficient.

I may have information about multiple email addresses so having the ability to do a bulk upload search is very efficient for me. If there needs to be a limit imposed, then this could be limited to 20 at one a time.

**PHONE NUMBERS**

Telephone numbers may include landline numbers, mobile numbers which could be UK based or foreign numbers.

What I am interested in is identifying who the owner of the number is and if that number is associated to any social media accounts as this would assist me identifying who the person is.

Telephone numbered can be written in multiple ways, so I need the system to be able to search for all the different variants and for the results to be displayed on the screen which tells me where the results came from.

I may also want to Verify which region the phone number belongs to, who the telecom operator is and when the phone was last active.

I start of by doing a check on GBG investigator to identify if the number is associated to anyone. I also have the option to check on GBG who the operator is and if the number is ‘live’ number or a dead’ number.

I then use multiple search engines to search for this information. There have been occasions when searching for a phone number has brought back occasions when the user of the number has or is selling items on selling sites like eBay and gumtree. I would also use third party tools like TrueCaller to see if I can identify some information about the person who uses it.

I may also use a virtual machine to try and attribute telephone numbers to specific social media applications like Telegram, Signal, WhatsApp, Kik and so on. This usually gives me a profile picture of the user.

I may also want to see if the number has been leaked onto a Breached database. So anyway, that can be automated would make the whole process more efficient. There have been several occasions where important information has been leaked by hackers. Sometimes this information sits on the darknet and sometimes this is available on the surface web on sites like Pastebin. So anyway, that can be automated would make the whole process more efficient.

I may have information about multiple telephone numbers so having the ability to do a bulk upload search is very efficient for me. If there needs to be a limit imposed, then this could be limited to 20 per search.

**SOCIAL MEDIA HANDLE**

As an i3 Investigator a key task is to identify a person, username, group and / or organisation on publicly available social media platforms and applications as well as those that require additional logins, using just a social media handle. This will provide me with a great deal of personal information. This information may include location details, siblings, and associates, frequently visited locations, social activities, work addresses, social interactions, multimedia content (images and videos including live streams), posts / comments made and so on.

Some of the subjects of interests / organisations I am interested in may not speak English so I may need to translate the text so I can assess the relevance of it. Having an ability to do this quick time is really important due to the time critical nature of the work that I may be conducting.

I may also be interested in the metadata associated with the original content – multimedia files metadata, the date / time and geo location information associated with the posted content (video / image).

I would initially use several search engines to search for the social media handle. On several occasions I have found different results have been brought back depending on whichever search engine I have used so it is important for me to be able check multiple search engines. I would then go to each social media platforms / applications I am interested in, to identify the person of interest. I will do this by using the search functionalities within social media platforms to search for the @ handle. This can be a time-consuming task so anything that automates this (without bringing any false positives) would be hugely beneficial.

As part of this investigation, I may identify accounts that no longer exist so will look to use archived websites like wayback machine to check what they looked like at the time when the site was indexed.

As part of my i3 work once accounts of interests have been identified I may need the unique ID’s of these accounts so that further investigatory work can be conducted (Subscribers) to evidentially identify the person behind the social media account. I find this information in several ways, either through inspecting the site I am on or using 3rd party tools, for example <https://www.codeofaninja.com/> to identify the UID.

I may also use a number of 3rd party tools that provide me with the associated email (full or partial) and telephone numbers (full or partial) associated to that particular social media account.

I would also conduct checks on some websites, forums, and breached database to see if this username has been featured anyway.

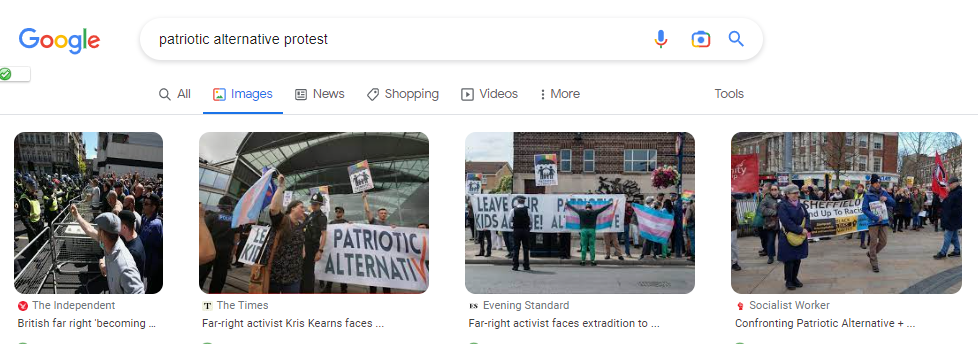
I would be expected to save the results and have a clear audit function for court purposes.

**IMAGES**

I regularly conduct i3 investigation into images. Images provide you with a lot of valuable information about a person, groups or location. They can show you what a person looks like, identifying location a person has visited, any vehicles used, logos /emblems of extreme organisations. Identifying this information furthers my investigation as it can assist verifying the person you are looking for and they associated social media accounts.

I may conduct this search by using several search engines or 3rd party tools to conduct a reverse image search to see if that image has been used anywhere else on the internet or if there are similar images containing similar looking items or people. My go to tools for reverse image searching is Google, Bing, TinEye and Yandex.

I sometimes search for an entity on a search engine and use the customised options below the search engine (Google search categories) which allows me to refine my search by specific categories, like Images, News, Shopping, videos and so on. This allows me to focus on just images for that search criteria which can assist me in identifying what the person of interest may look like. (see below image)



I am also interested in images that have written content embedded within them, for example a poster that has information about a potential protest organised by an extreme organisation. At this moment in time, I have no automated way of picking this information out of images when doing checks.

I sometimes have images where the persons face has been pixelated which means I cannot identify who that person is. There have been several occasions when an extreme right-wing group has attended a protest and then added these pictures online but some of the faces have been pixelated so that they do not identify who they are.

I routinely look to see if any of the images I am looking at have the Metadata associated with the original content, providing me with contextual information about the image(s) – multimedia files metadata, the date / time and geo location information associated with the posted content (video / image). I use free 3rd party tools to do this.

**VIDEOS**

In my i3 role I sometimes have to locate and assess videos or popular video sharing platforms like YouTube, Odyssey, DLive and so on. The content of these video may be in a foreign language which may need translating from voice to text. I rely on translators to do this for me which ca be time consuming and expensive for my organisation. A tool that allows you to translate audio to text would be efficient to assess if the video hosts any extremist material and if an actual translator is required.

I sometimes search for an entity on a search engine and use the customised options below the search engine (Google search categories) which allows me to refine my search by specific categories, like Images, News, Shopping, videos and so on. This allows me to focus on just videos for that search criteria which can assist me in identifying what the person of interest may look like and if there are any videos the person of interest may have featured on. (see below image)

Graphical user interface, text, application, email

Description automatically generated

I may also need the video transcribed into a text format so that it can be provided as intelligence. If this can be automated this would save a lot of time.

I may also be interested in the metadata associated with the original content, providing me with contextual information about the video(s) – multimedia files metadata, the date / time and geo location information associated with the posted content (video / image).

**WEBSITES**

As an i3 Investigator I may get given a website to look at. I would be interested in knowing if the URL has been mentioned anywhere else on the internet. I would use different search engines to do this as well as checking specific platforms by doing a search in the search bar.

There are some videos which hosts lots of extremist material. it would be beneficial to know how much traffic this website gets to assess how much influence this website has.

The URL may no longer exist so I will look to use archived websites like wayback machine (<https://archive.org/web/>) to check what they looked like at a specific time.

**DarkNet**

Conducting i3 Investigation on the Dark Net is important but difficult to do so especially when I know there is information to indicate that the person of interest is visiting the dark net or when there is indication of some illegal activity happening in this space, for example trading of illegal goods.

I use several 3rd party tools that allows me to scan the dark net for this information. There are some 3rd party tools that I have used which allow me to do some searches without logging into a TOR browser. For example, Ahmia (<https://ahmia.fi/>) is a search engine for the dark net, it makes onion serves accessible to a wide range of people.

**MONITROING OF GROUPS / ORGANISATION**

I am also tasked on a regular basis to investigate what information I can obtain about a possible extremist group / organisation.

I would start by doing a search on several search engines to identify any posts of relevance or any websites the groups / organisation may be associated. I may use Advanced search options or Boolean searches to do this.

I will then check if any of these groups / organisations have a footprint on any social media accounts. Usually if these relate to a specific thematic, like right wing extremism, I will look to explore their footprint on the Alt-tech platforms like Gettr, Gab, Telegram to name a few as these platforms are more likely to attract people who need more ‘freedom of speech’.

I would also look to identify who the main leaders of the group / organisation are and who is supporting them.

I would also need to assess if the group / organisations were breaching terrorist legislation so would need to assess the groups / organisations ideology.

**MONITORING OF LOCATIONS**

There are occasions when I can be tasked to have a look at a location. A location is of interest as it might be somebody’s current address, previous address, workplace address or a location where a group is hosting an event, or a person is visiting. At the moment, I rely heavily on keyword searching on search engines and social media platforms to do this. I may use online maps on search browsers as a way of looking at a location.

When I am investigating a persons social media footprint I may find that they have left they GPS switched on which may identify coordinates in photographs, or give me some geo location information based on the elements of the photos background, tagging on social media, leaked content from fitness apps and so on.

**MONITORING OF EVENTS**

I am also tasked to identify particular events relating to specific groups / organisations or a person of interest. These events may be advertised as posts on social media platforms on a person’s profile or a community profile. They may also be advertised as a poster / image with text embedded within the image. Having the ability identify this text during a search would be beneficial. At this moment in time, I have no automated way of picking this information out of images when doing checks.

When the event is occurring my job would be to identify people who are in attendance, how many in attendance, if there is any disorder or any criminal offences committed. I would need to evidentially capture this material. sometimes these events can be live streamed as a live video, sometimes people can take small videos and then post these on social media platforms.