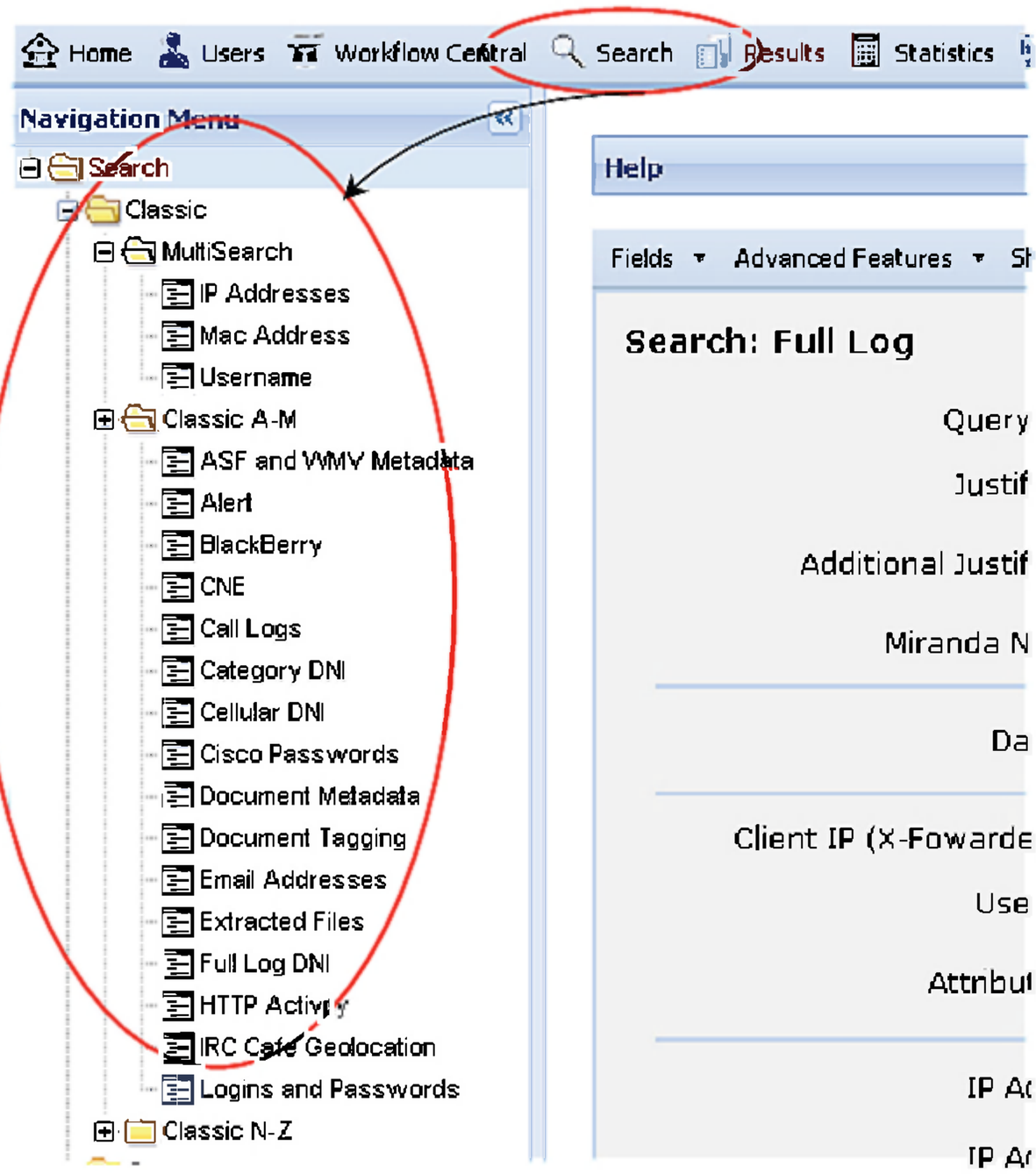
The Unofficial XKEYSCORE User Guide

[redacted]
E92 – ADET
Consultant, Booz Allen Hamilton

Derived From: NSA/CSSM 1-52
Dated: 20070108
Declassify On: 20320108

Creating Queries

Clicking on Search at the top of the screen will bring up a list of searches in the Navigation Menu:



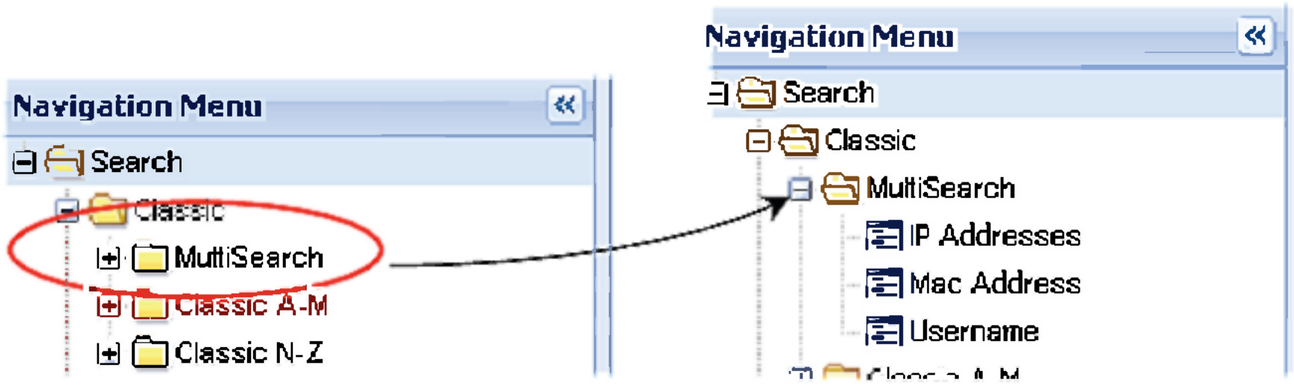
The Search screen has cascading menus of different Searches: Classic, Common, Dictionary Hits, File Transfer, Multisearch, Network Management, User Activity, VoIP, and Wireless.

Classic Queries:

Within the Classic Menu there are three folders: MultiSearch, Classic A-M, and Classic N-Z.

Multisearch:

Expand the Multisearch folder by clicking on the plus sign:



Multisearch IP Address:
The **Multisearch IP Address** query allows you to search on an IP address into seven different searches. Think of it as a federated query using an IP address. The **Multisearch IP Address** query searches on:

- User Activity
- Phone Number Extractor
- Email Addresses
- Extracted Files
- HTTP Activity
- Full Log
- Web Proxy

Refer to some of the individual searches below for more information about specific queries

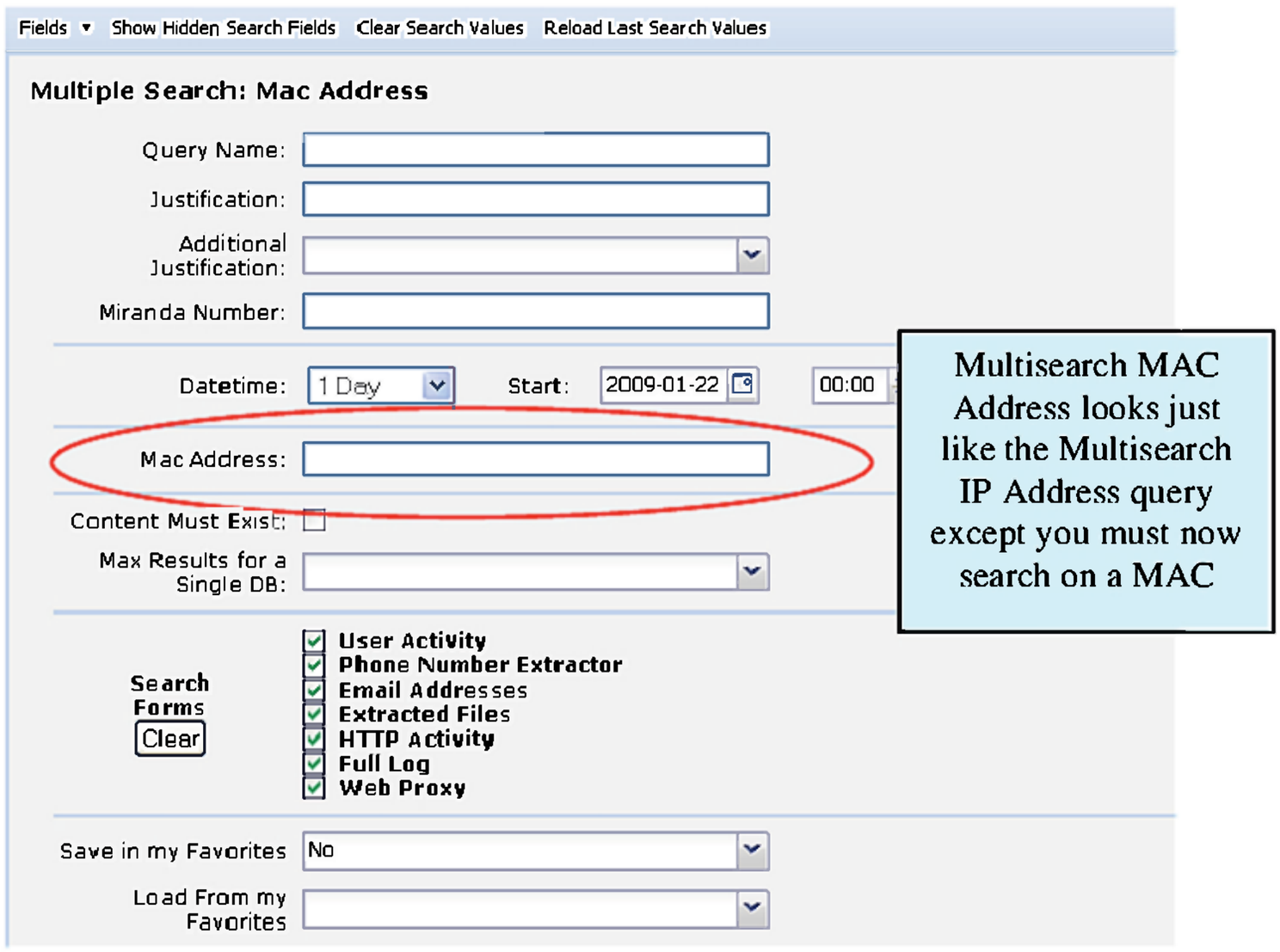
Creating a MultiSearch IP Address Query:
When you have filled in your query name, justified it, entered an IP address, selected your search engines and sites the last thing is to submit the query. If you select "Merge Results", then all of your individual queries will be merged into one consolidated result.

"Why would I want to merge my results?"
If you wanted to see all of the activity together to get a 'big picture' look at the IP address, regardless of the activity or application that is on the IP. The New GUI's results screens allow you to filter your results easily which may make viewing your results more intuitive. See "Viewing Your Results" in this Guide.

"What would I want to NOT merge my results?"
Viewing the results individually allows you to focus on a particular activity or result (e.g. Documents or email addresses).

Multisearch MAC Address:

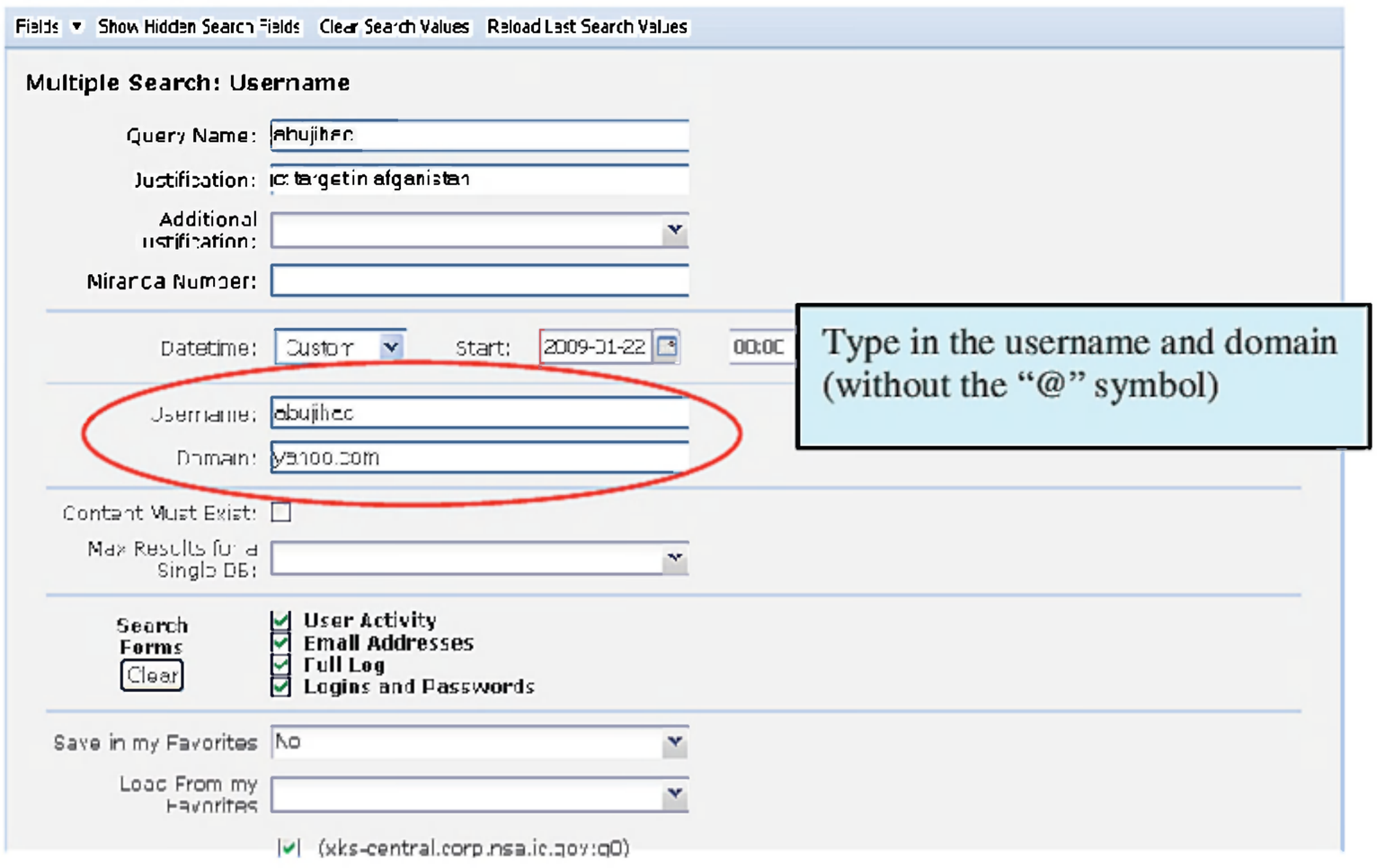
The **Multisearch MAC Address** query is exactly the same as the IP address query except it only allows you to search on a MAC address. Follow the same instructions as the **Multisearch IP Address** query above but replace the IP address with your MAC address(es).



Multisearch MAC Address looks just like the Multisearch IP Address query except you must now search on a MAC

Multisearch Username:

As you may have guessed, the **Multisearch Username** query is exactly the same as the IP Address Query and the MAC Address except it only allows you to search on a target's Username. Follow the same instructions as the **Multisearch IP Address** query above but replace the IP address with your Username(s).



"What is a Username?"
A "Username" in XKEYSCORE queries is the portion before the "@" symbol in an email address.

For example:

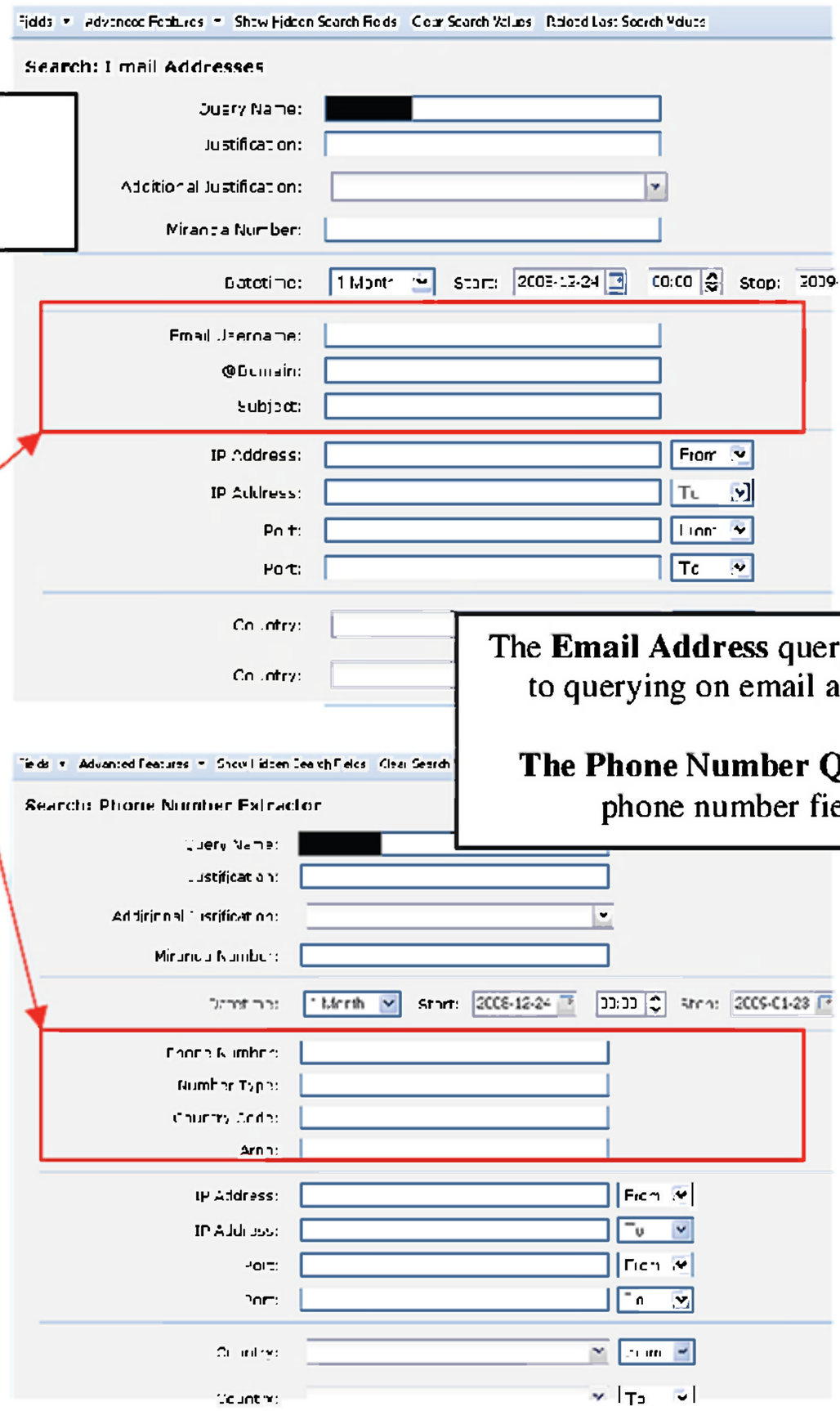
Abujihad@hotmail.com: Username = abujihad
Domain = yahoo.com

Classic Searches (A-Z)

There are 32 different searches between the A-M and N-Z searches. This guide will cover some of the most common searches. You will notice that most of the fields of the searches are the same and each individual query will be unique because based on its query name. For example, the **Extracted Files** search has fields that are only applicable to file attachments (e.g., file names, file extensions) and the **Email Addresses** query has fields for email addresses (e.g., username and domains). All of the Classic queries will have common fields like Ports, IP addresses, Countries, SIGADS, and CaseNotations that you can use to

Here are two Classic queries: **Email Addresses** and **Phone Number Extractor**.

The fields between Datetime and the IP Addresses are the plug-ins unique to each query.



The **Email Address** query is catered to querying on email addresses

The Phone Number Query has phone number fields

Email Addresses Query:

One of the most common queries is (you guessed it) an **Email Address Query** searching for an email address. To create a query for a specific email address, you have to fill in the name of the query, justify it and set a date range then you simply fill in the email address(es) you want to search on and submit.

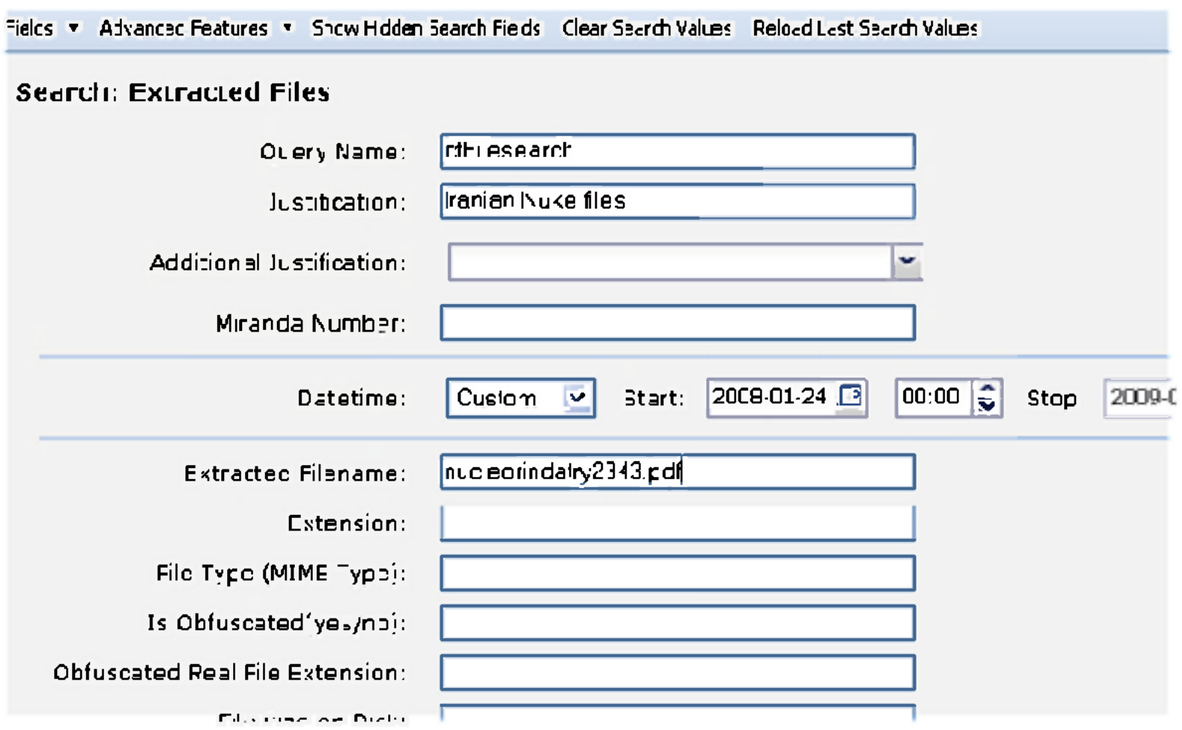
That would look something like this…

NOTE: You DO NOT have to know an email address to use the **Email Address Query**. You can also search on an IP address*, domain name**, country, port, casenotation, protocol, SIGAD, MAC address, PID and more. If you search on something other-than an email address (e.g., an IP address), your results will be all of the email addresses seen on those IPs.

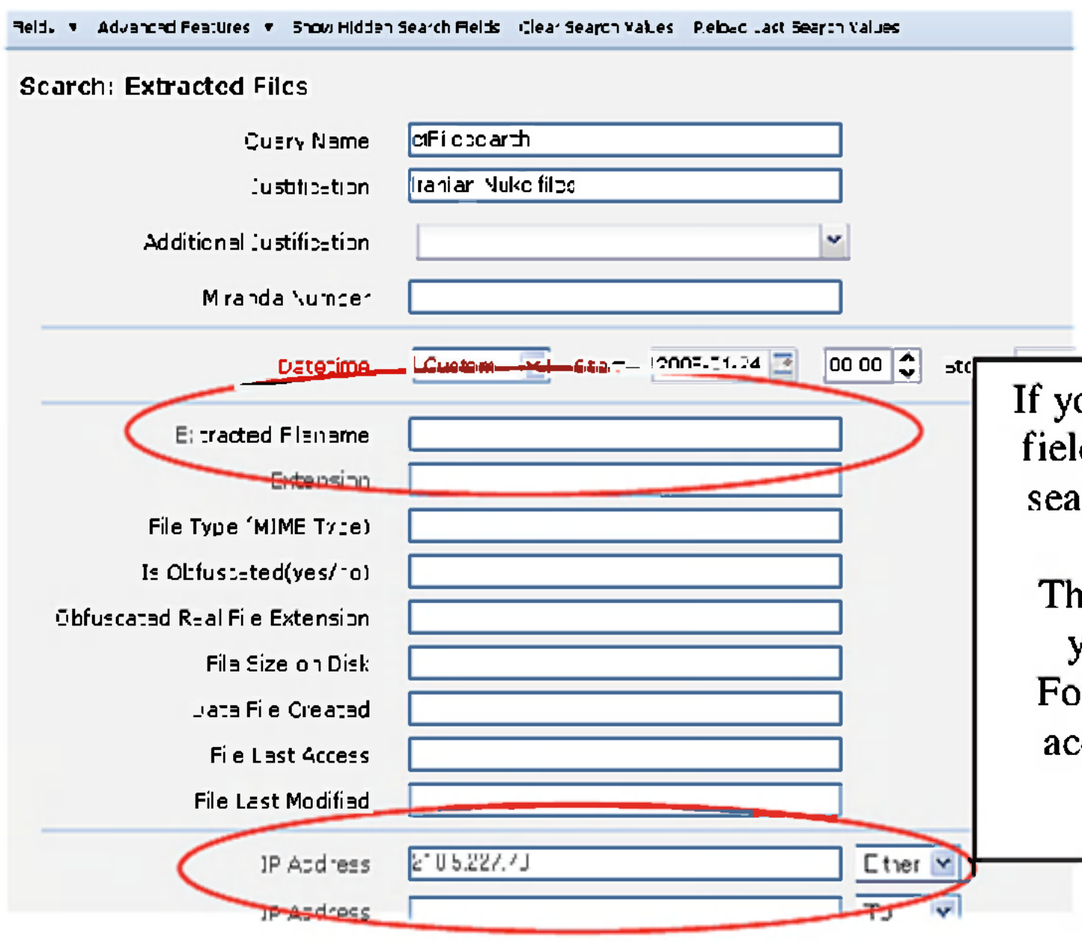
* The IP must be hosted OUTSIDE 5-eyes countries
** The Domain MUST be foreign owned. Check WHOIS and NSLOOKUP for more info on your domain before-hand

Extracted Files Query:

1. To find a specific file (i.e., if you already know the file name): For example, if you noticed a file name in your target's inbox and you never actually got the file attachment. This is VERY common for webmail collection because the attachment is often not put into PINWALE with the email.



2. To search for all files or specific file types on a particular area or on a network. (E.g., IP address). This is a GREAT query if you have a foreign mail server and want to see what files are collected on that IP address.

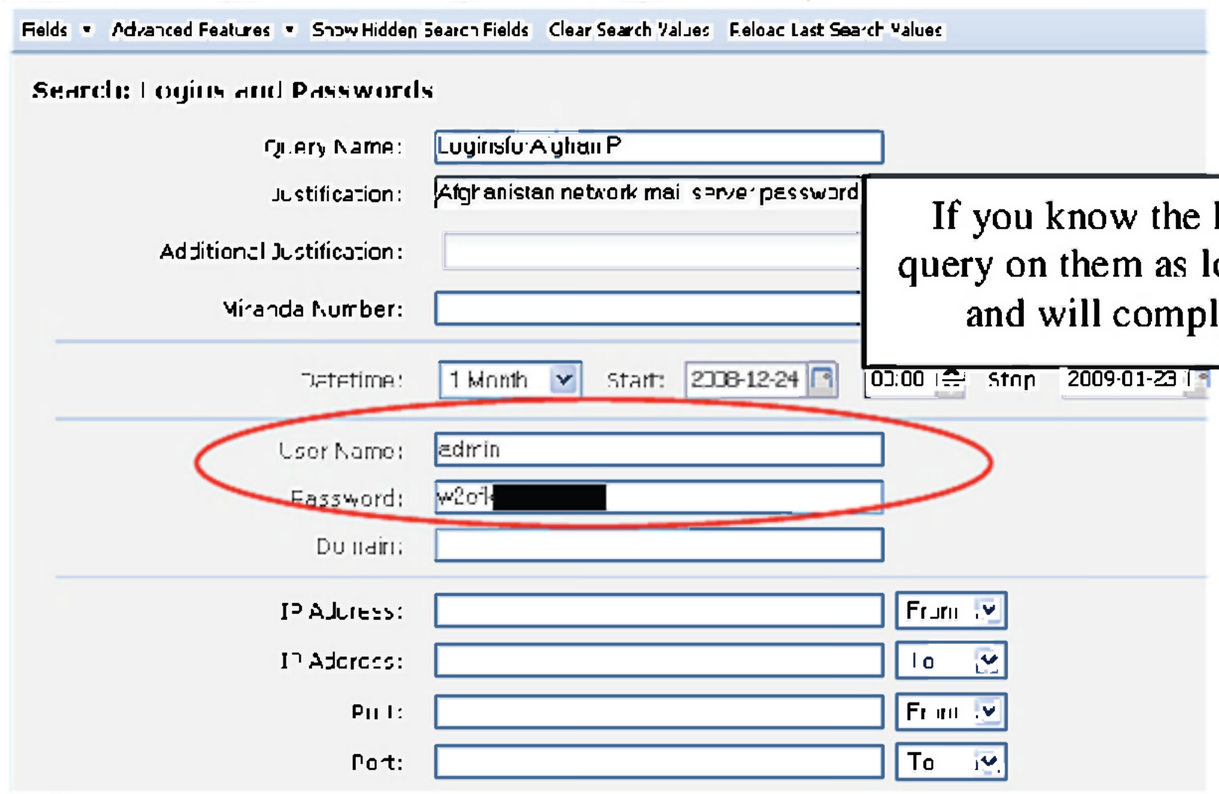


If you leave the Extracted Filenames field blank, you are wildcarding the search to look for ALL files names

The IP Address of the mail server you found using NSLookup in Foxtrail or your non-attrib Airgap account goes in the "IP Address" field

Logins and Passwords

1. <u>If you already know the login and/or password.</u>

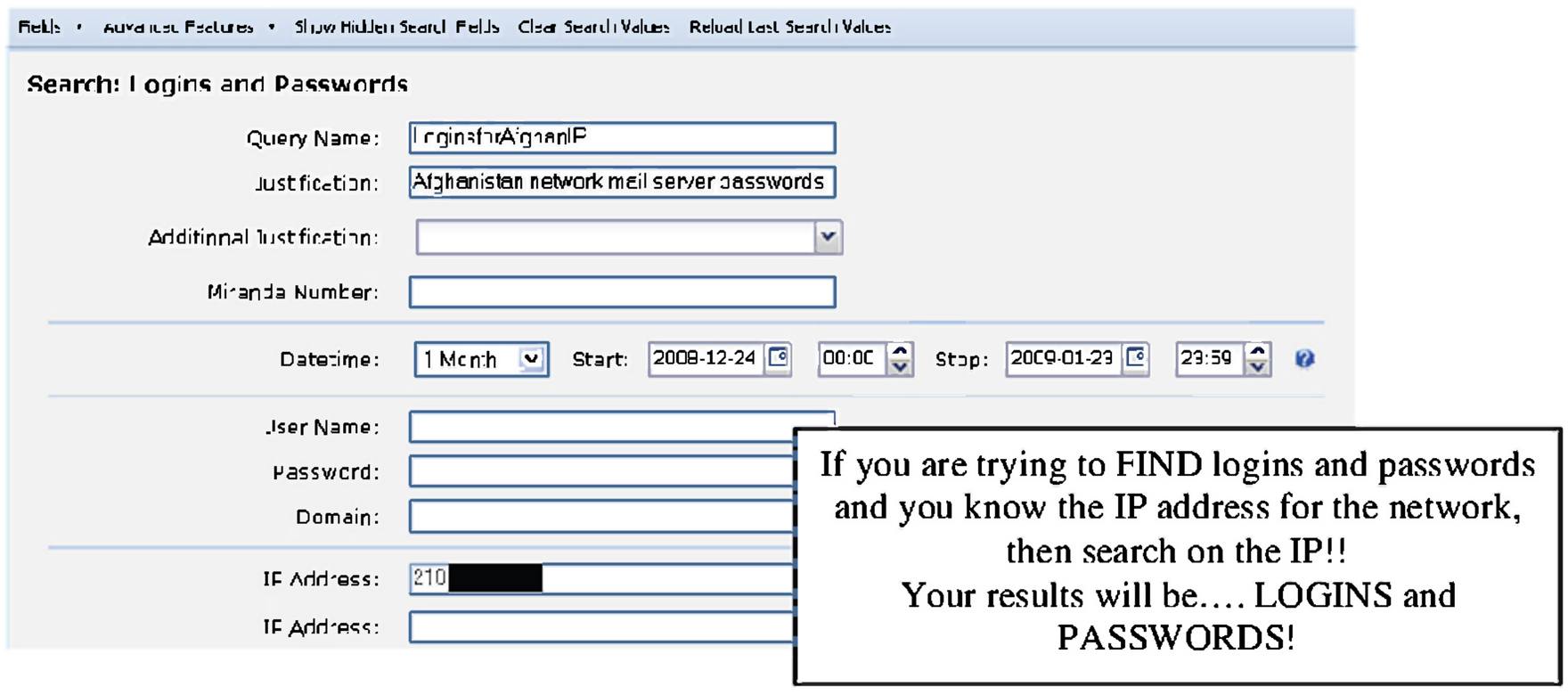


If you know the logins or passwords, query on them as long as they are unique and will comply with USSID-18.

"Where would I find passwords to use in this query?"
Passwords can be found in TUNINGFORK (e.g., FoggyBottom), passed in the content of emails or text messages, or from previous XKEYSCORE queries.

2. <u>Trying to discover logins and passwords on a network</u>? NOTE: Logins and passwords are valuable tools to enable Tailored Access Operations (TAO).

"What tools would I use to get the network information like a Mail Server, or Name Server?"
NS Lookups tools on NSA net such as FOXTRAIL and Open Source tools such as robtex.com, centralops.net, and network-tools.com are a GREAT START. They provide you with IP addresses for domains. You can then query on the foreign-hosted IP addresses.

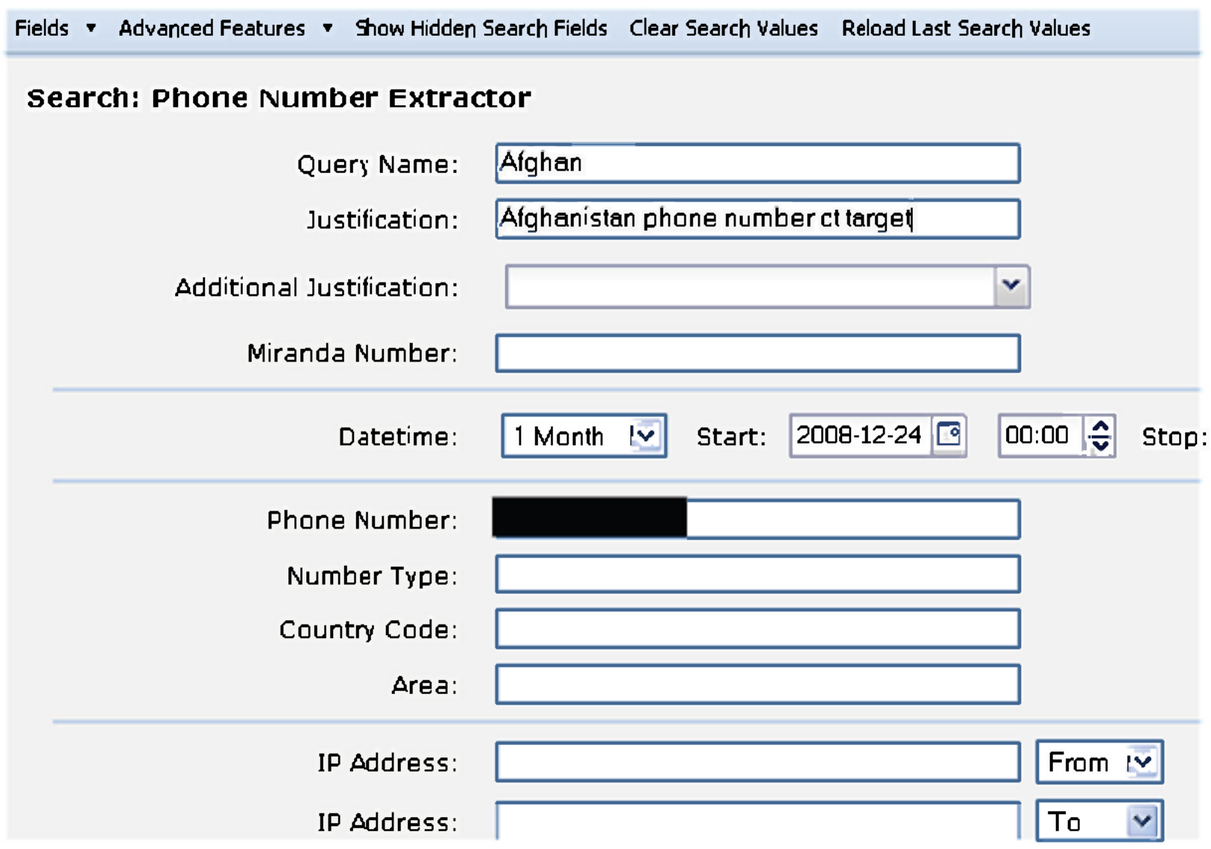


If you are trying to FIND logins and passwords and you know the IP address for the network, then search on the IP!!
Your results will be…. LOGINS and PASSWORDS!

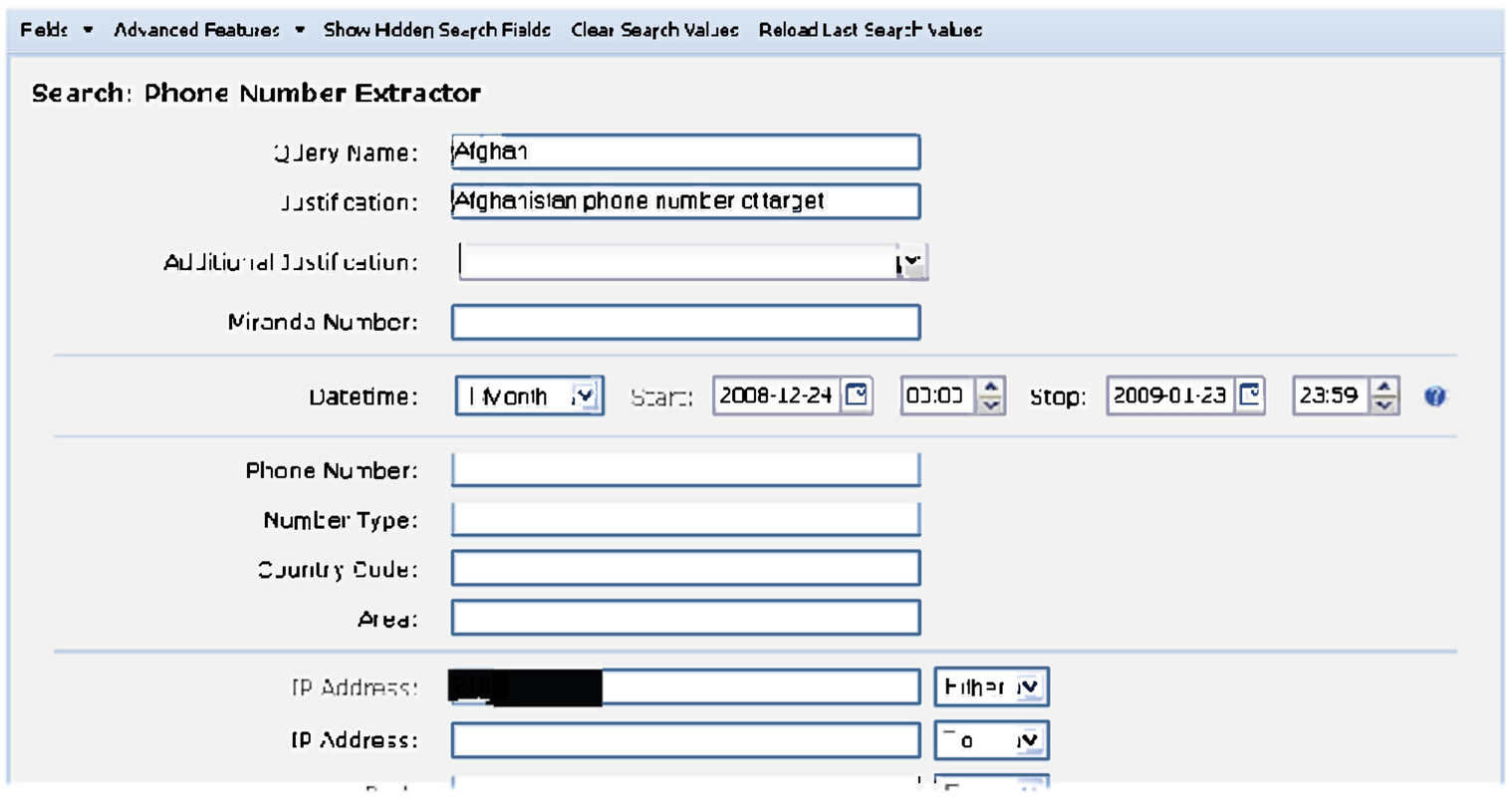
Phone Number Extractor

The **Phone Number Extractor** query looks through the content of an email for phone numbers. This is very similar to a PINWALE DoPhone query except the traffic that XKEYSCORE finds may be survey (i.e., unselected, non-tasked data) and might not be in PINWALE. XKEYSCORE may be your only hope at finding an email address for a target where you only have their phone number as lead information.

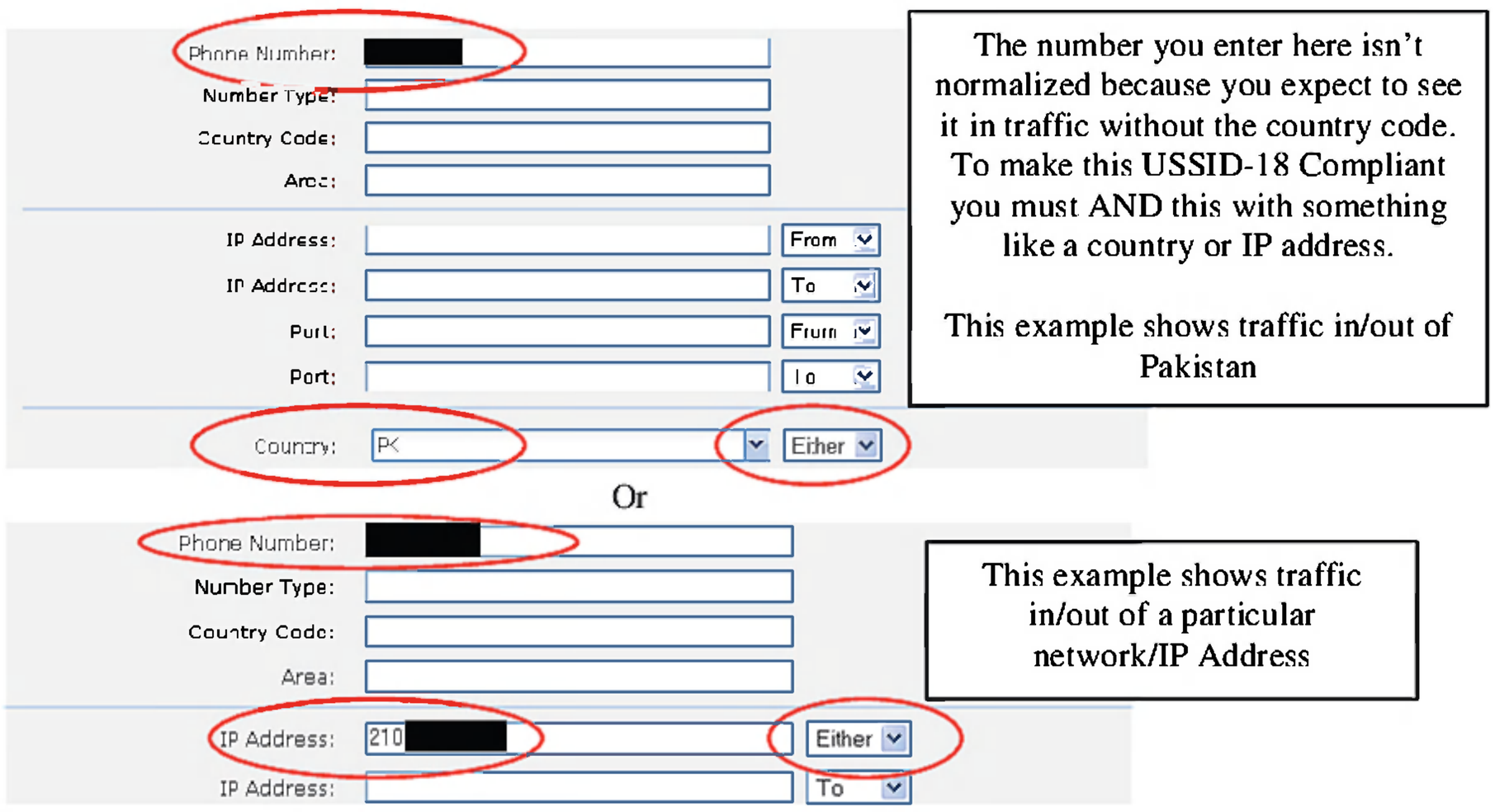
1. Already have a phone number? If all you have to start with as lead information is a phone number, you may find it useful to query on that phone number and see if anyone sent an email with that number in the signature line.



2. Looking for any phone numbers on a network? Quite often you know the mail server IP address and could use some telephone numbers to task?



3. Looking for a phone number without the country code (non-normalized)? It's possible a target will pass their phone number without the country code (e.g. a signature line with "Tel: 5354658"). In that case, XKEYSCORE will not find the number with the country code so you must create a query that looks for fewer digits but still complies with USSID-18. This is not a 100% solution* but ANDing your query with a country or IP address would certainly be more compliant. See example below:



The number you enter here isn't normalized because you expect to see it in traffic without the country code. To make this USSID-18 Compliant you must AND this with something like a country or IP address.

This example shows traffic in/out of Pakistan

This example shows traffic in/out of a particular network/IP Address

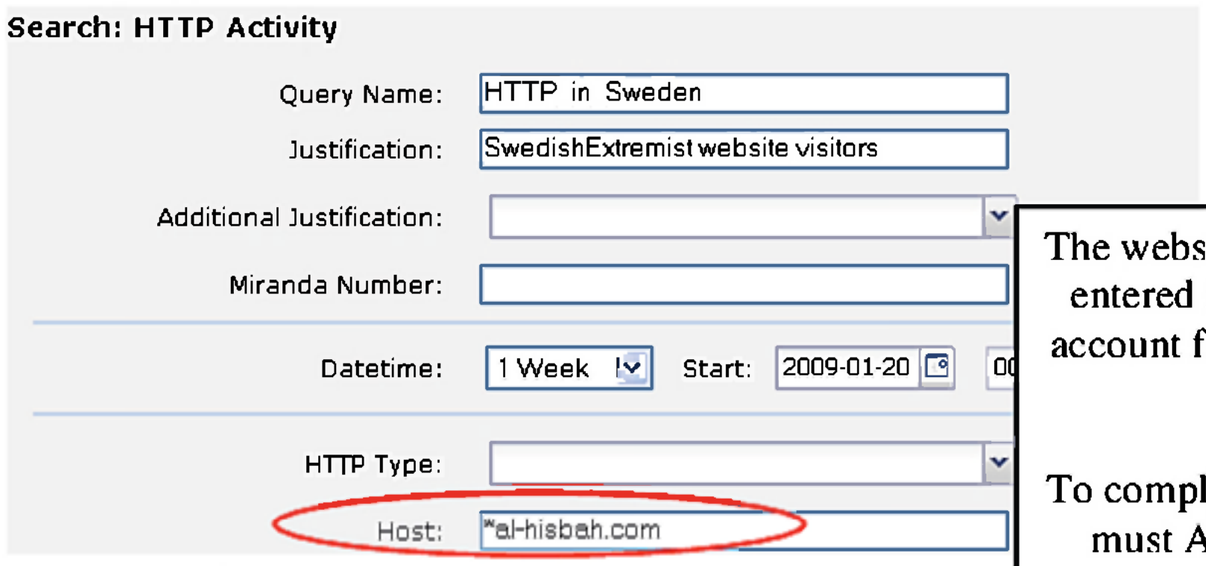
*If you ask XKEYSCORE to give you all Pakistani traffic, it's doing an NKB lookup on all Pakistani registered IP addresses. Geolocation of IP addresses is not 100% accurate at this time. Unofficial estimates say asking for all of Country X's traffic will find between 50-60% of the actual traffic. (That's more than 0%, though, right?)

HTTP Parser

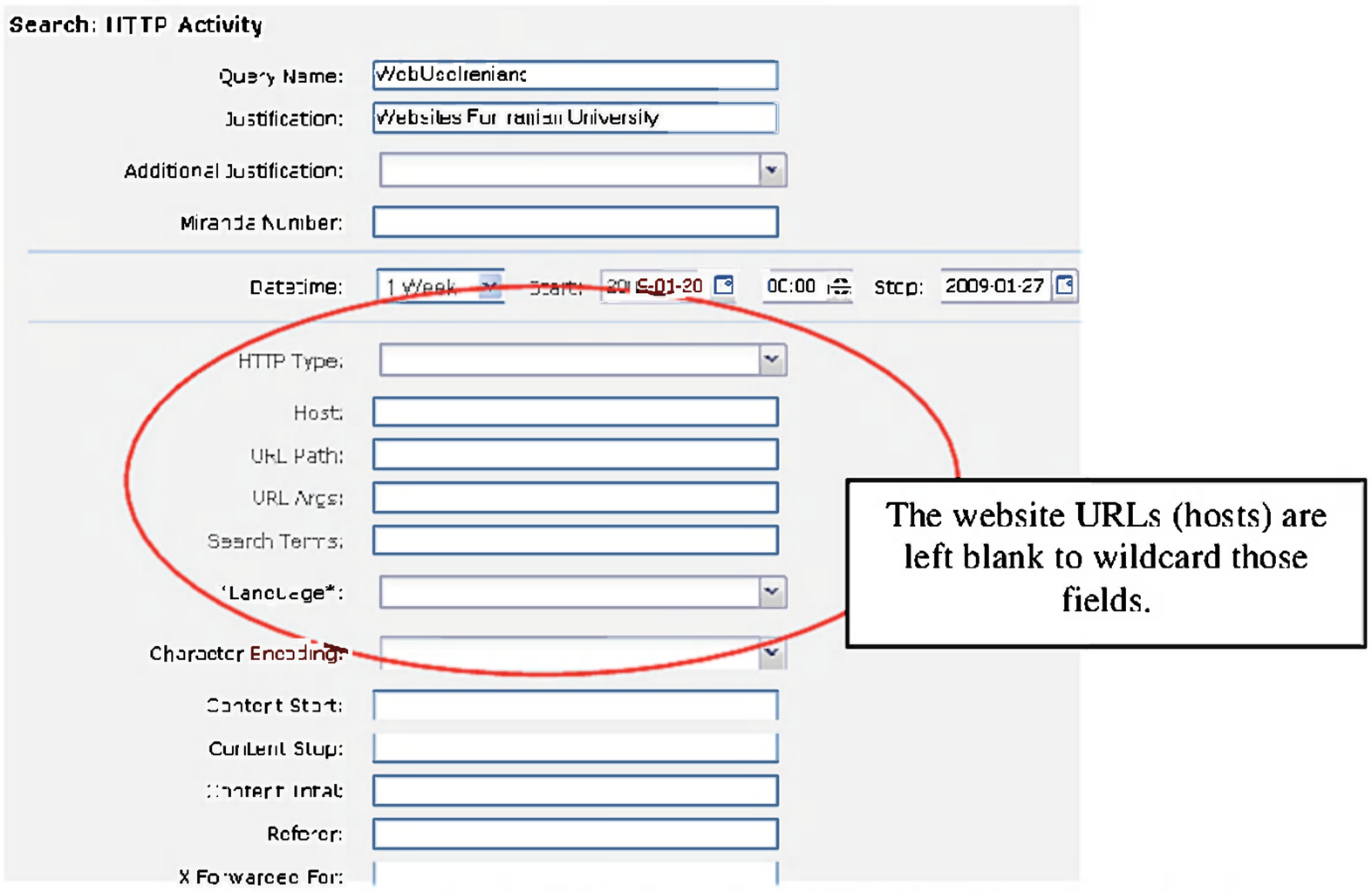
The **HTTP Parser** query looks for web activity (remember, HTTP = web) on a particular link. This query is useful for several reasons. Firstly, if you know a particular website and want to see if a foreign target visits it (e.g. an extremist web forum URL, or maps.google.com). Secondly, this query enables you to query on a network IP(s), casenotation, or country and see what websites we don't know about (survey-type query).

Here are two examples

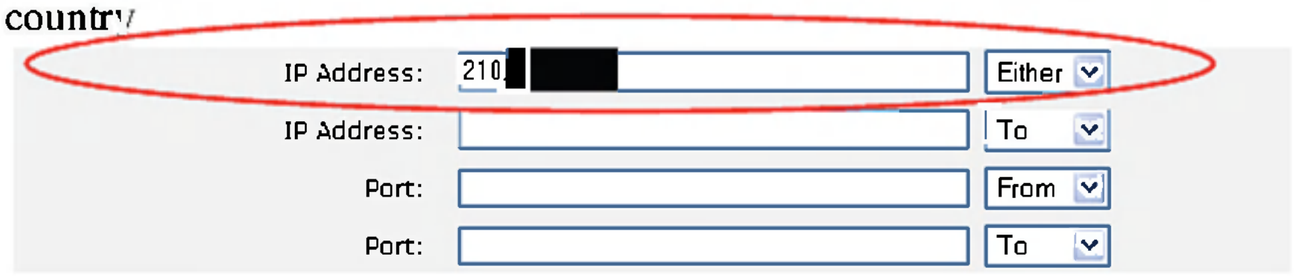
1. If you know the particular website the target visits. For this example, I'm looking for everyone in Sweden that visits a particular extremist web forum.



2. If you don't know the website but you know the network information (IP). For this example, I'm querying on a network IP block to see all of the websites the target visits.

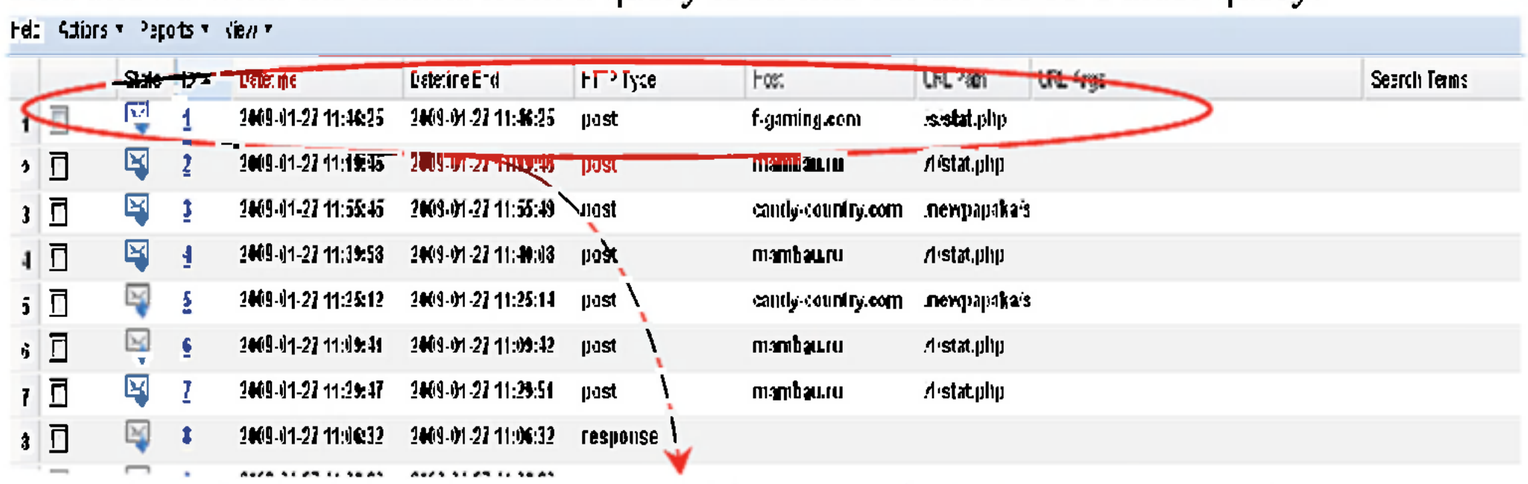


To comply with USSID-18 you AND that with some other information like an IP or country



Results from an HTTP Parser query

This shows what the results from a query look like for an HTTP Parser query:



Example 1 above shows a person was visiting www.f-gaming.com/s/stat.php

Host = f-gaming.com

URL Path = /s/stat/php

Document Metadata

Document Metadata query allows you to search on document authors, organization, encryption*, and many other things about a document. This is extremely helpful if you have found a file attachment from a target (e.g. Brick-and-mortar targets, person, or Organization) and you want to see all of the other files they have sent. With the Document Metadata query you don't have to know the email address of the person sending the document, you just have to know the document's properties.

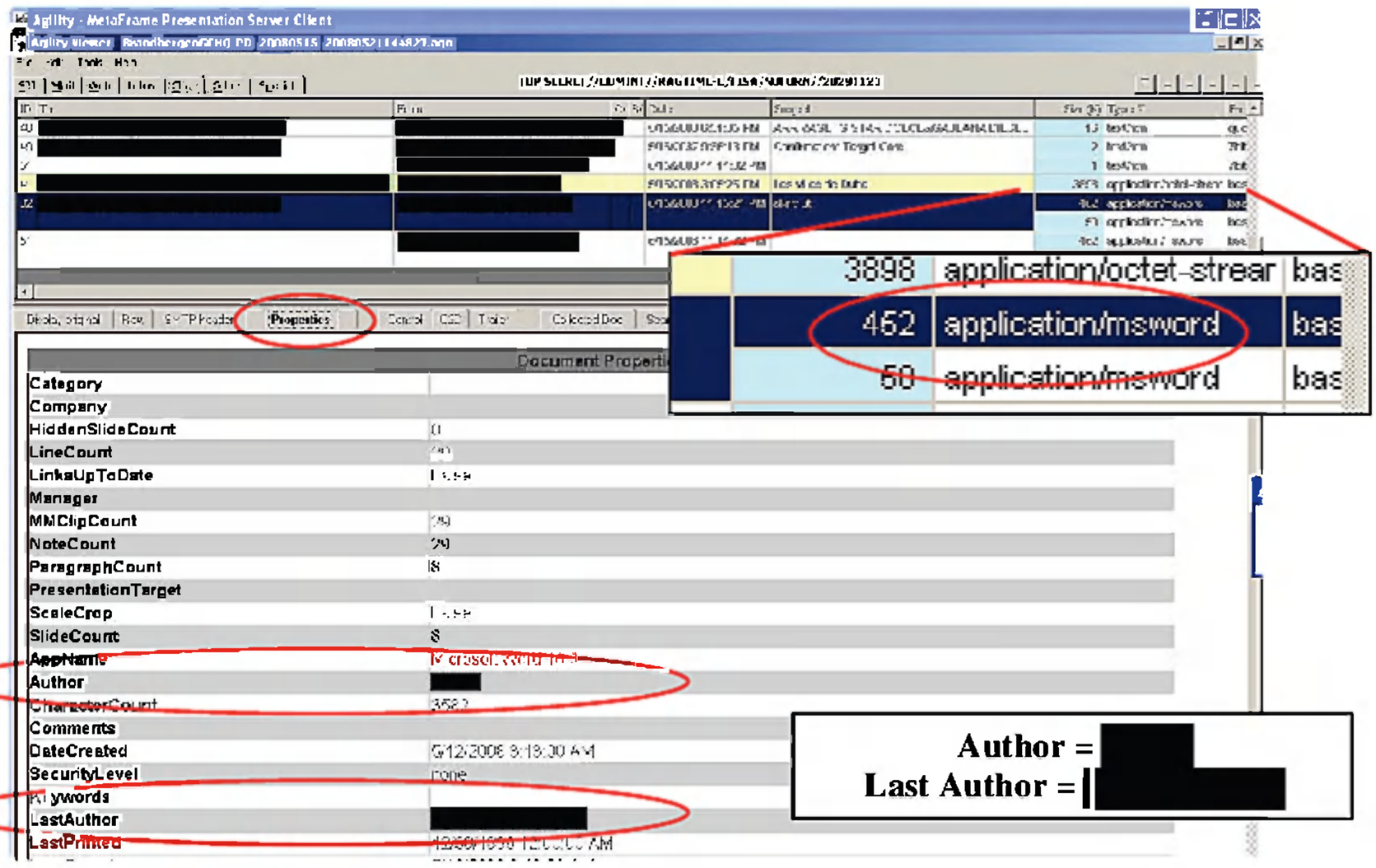
*Most Microsoft Office allows uses to encrypt files by clicking Tools -> Options -> Security and password protecting the files. The Document Metadata query looks for that type of encryption. It doesn't look for PGP or other 3rd party encryption.

> **"How do I find a document's properties?"**
> The easiest way to see this is to open a MS Office document and click on File -> Properties. To find the document properties for a file you target sent, the easiest way is to view the file in Agility and click on Properties.

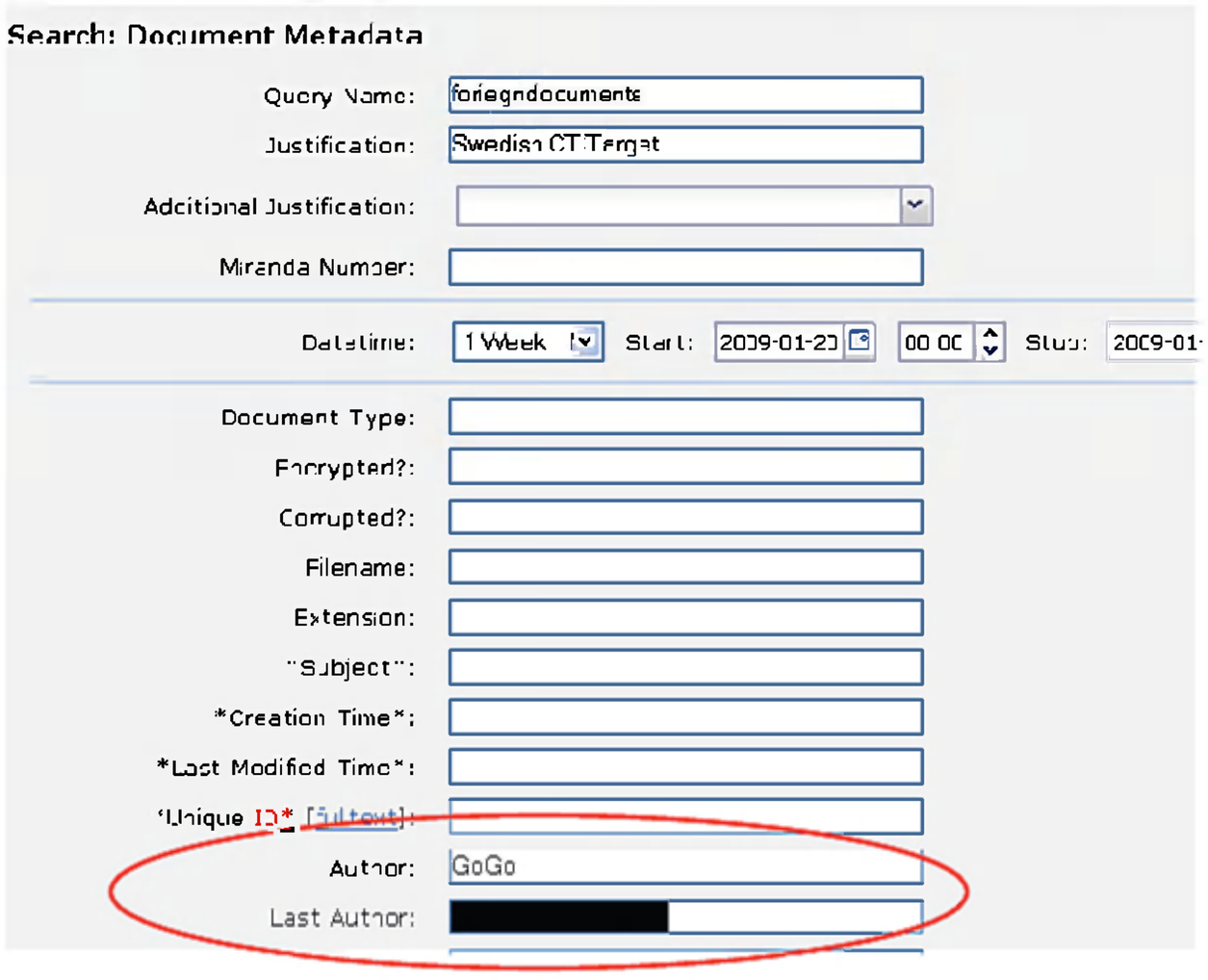
Finding your target's file properties

If you can view the target's document in Agility, click on the Properties tab to show the target's Organization and/or Author. If the fields are unique or random enough you can query on the term itself. If the Organization or Author aren't enough to comply with USSID-18, then you must AND that query with supporting information (IP or Country).

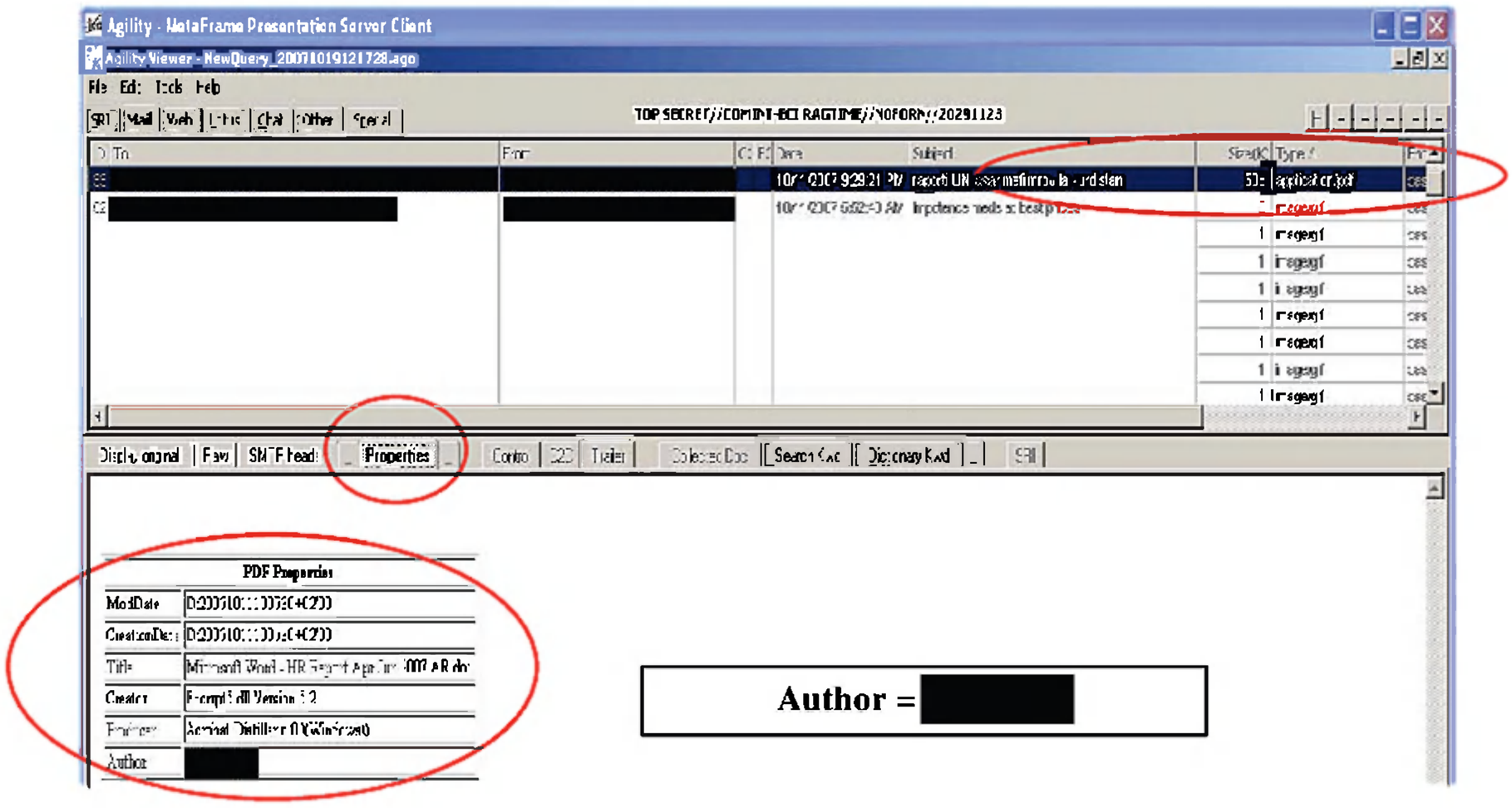
Displaying MS Word document in Agility:



To create a query in XKEYSCORE from this information:



View of document properties of PDFs in Agility:



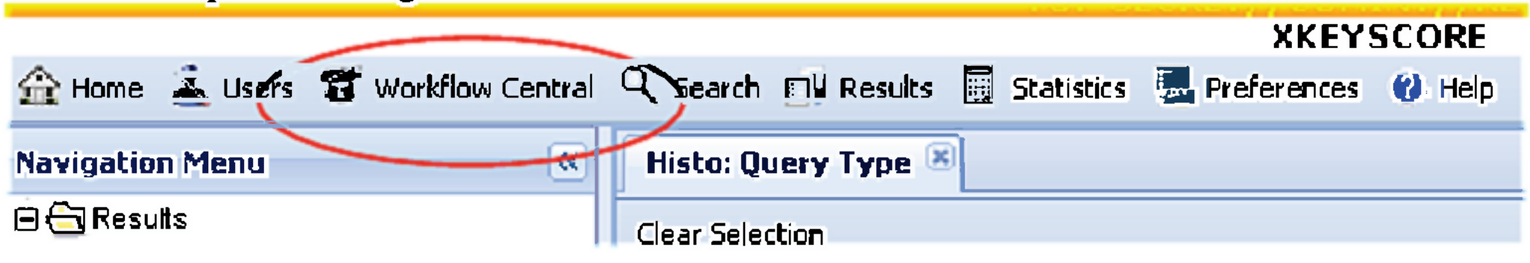
To create a query in XKEYSCORE using this information:

Creating a WorkFlow

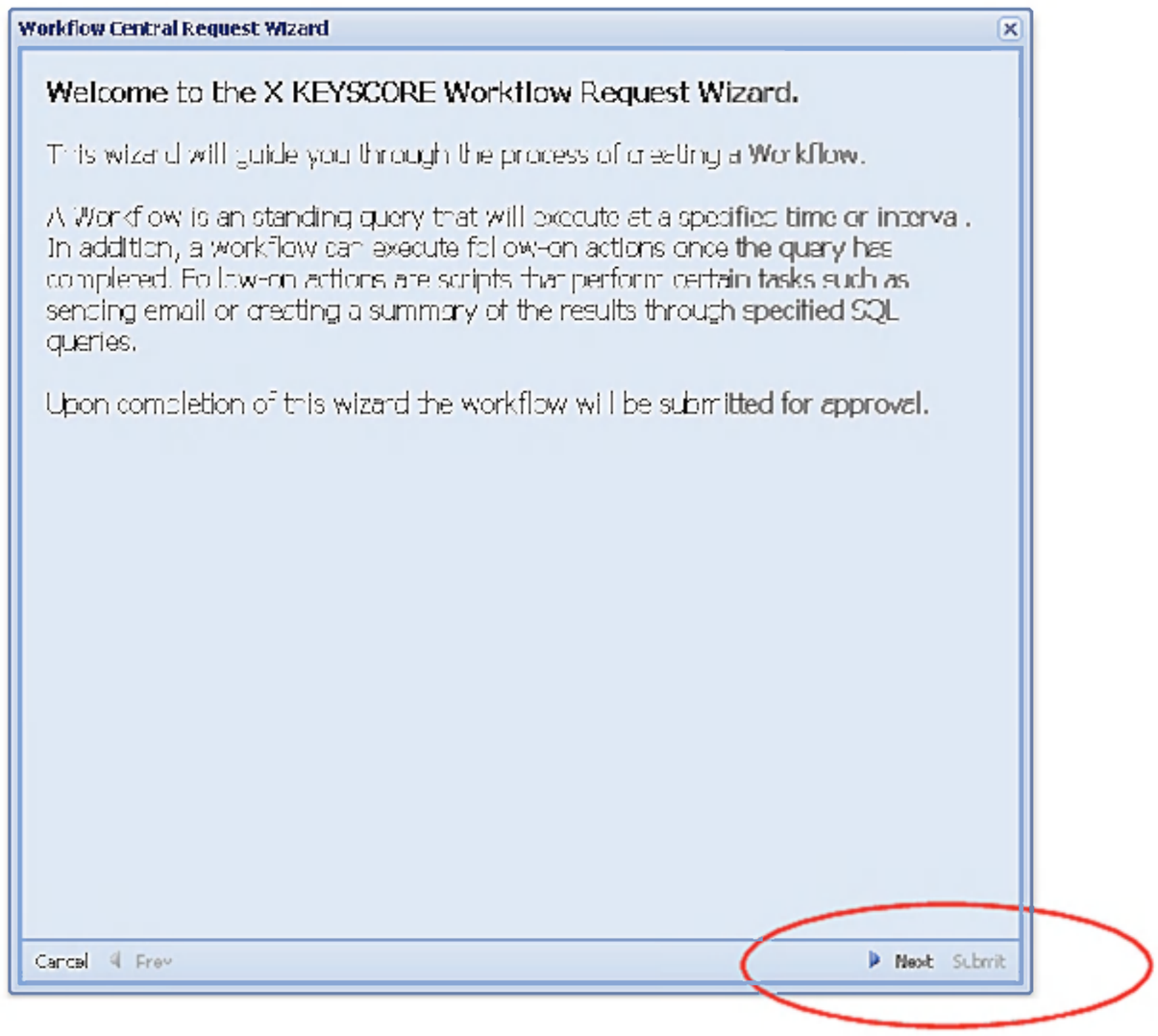
Workflows are periodic queries you can set up that run at specified times. They are great for sustained targets because they query the database for you (e.g. every night) and you can easily view the recently collected traffic without having to create a new query each day. They are also very helpful if you are performing target discovery on a network and haven't seen much traffic yet on a selector. A workflow for an email address can bridge the gap between when you discover the selector (and you task it to UTT/Cadence) and when it actually makes it to the appropriate dictionaries).

It's important to understand that a normal (ad hoc) query is submitted when you hit Submit. Workflows, on the other hand, are created then submitted to the XKEYSCORE team for review. The XKEYSCORE team does not review it for USSID-18 compliance (that's up to you); they only review it to ensure your query won't strain the system with too complex a query.

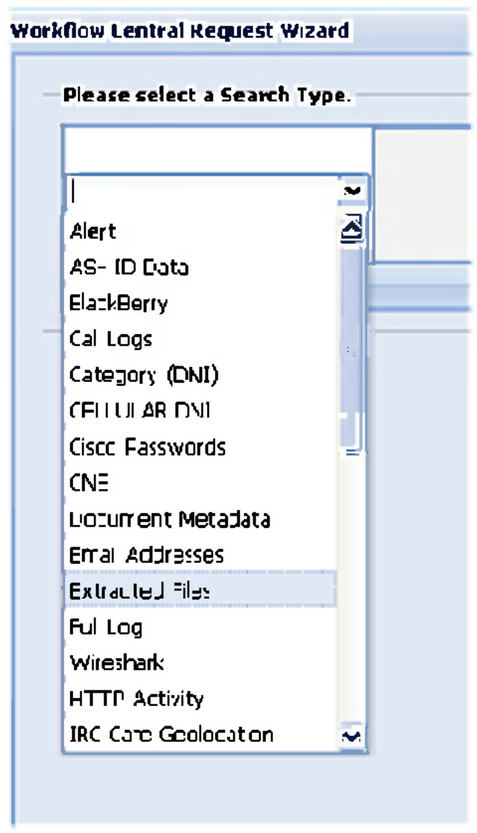
The first step in creating a Workflow is click on Workflow Central:



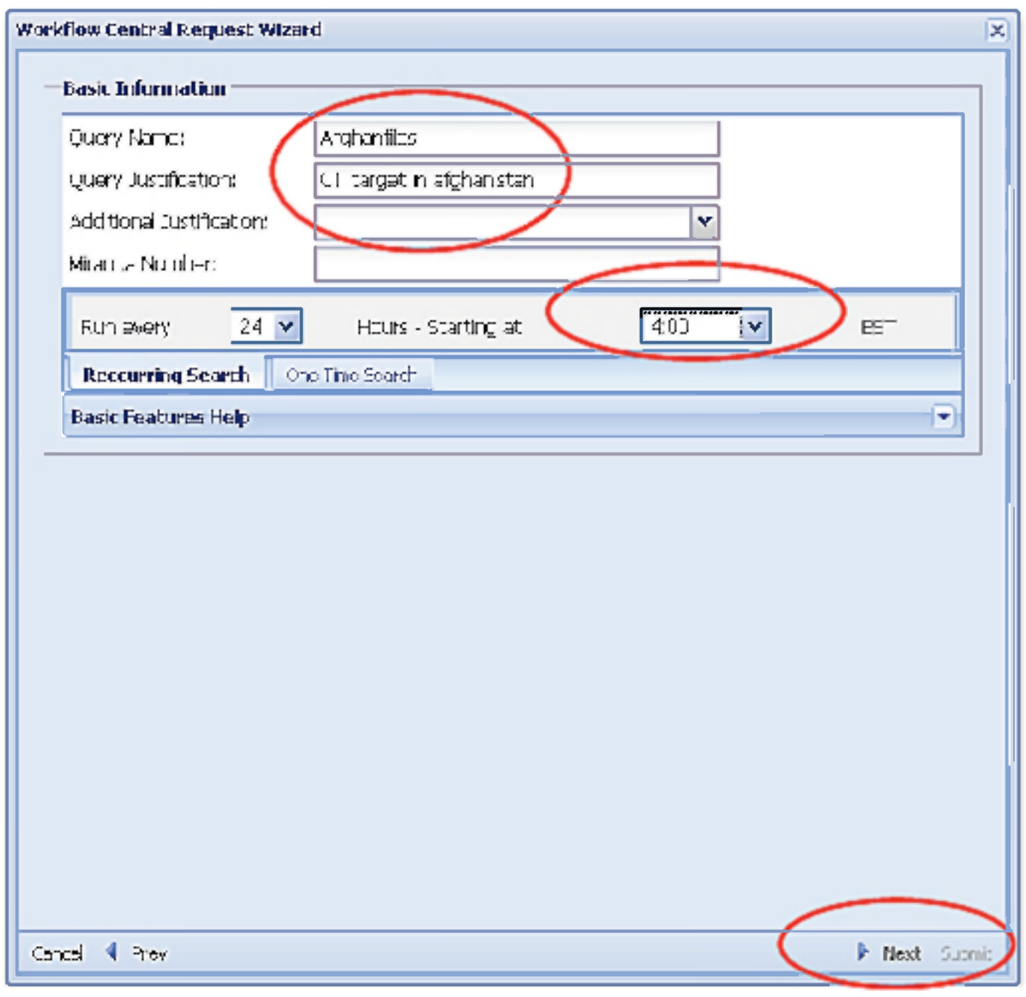
Then click on Request on the left to start the Workflow Request Wizard, and then click Next.



Next, select the search type you want to create from the pull-down menu. For this, I'm selecting an **Extracted Files** query. These queries are essentially the Classic A-M and N-Z queries you have seen in the Classic Search screens. The only difference is an **Extracted Files** workflow will start looking for extracted files in the future and an ad hoc **Extracted Files** query will search in past/previous collection.

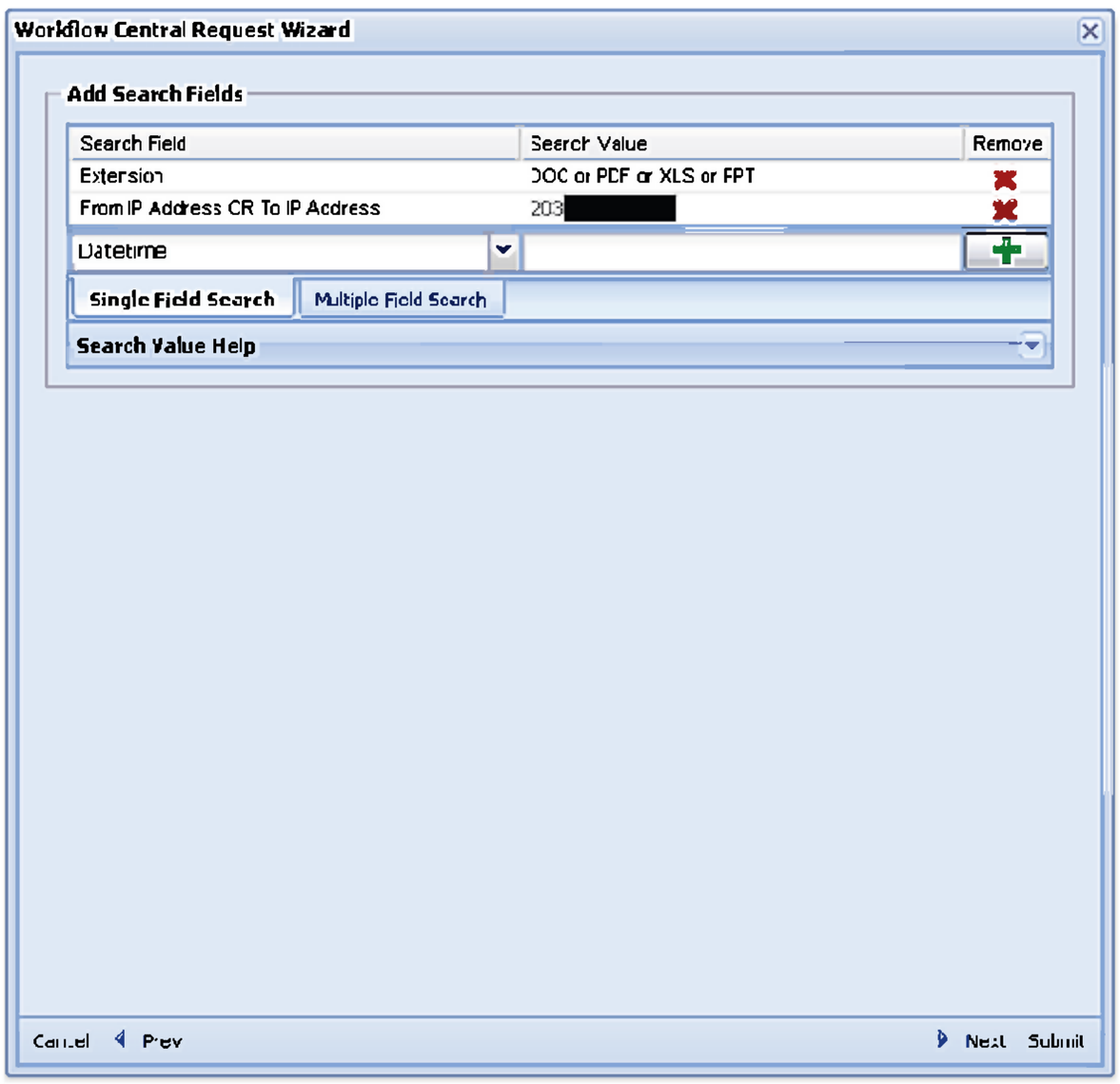


Next, fill in the name of your query ("AfghanFiles"), the auditor-compliant justification, and how often you want the query to run. I recommend offsetting the time from the default of midnight (2400) by a few hours (before of after). For this, I'm selecting 0400. Then hit NEXT.



In the **Add Search Fields** window, you will select the search criteria that you want to search on. In this example, I'm looking for specific file attachment (DOC or PDF or XLS or PPT) on a specific Afghanistan IP address.

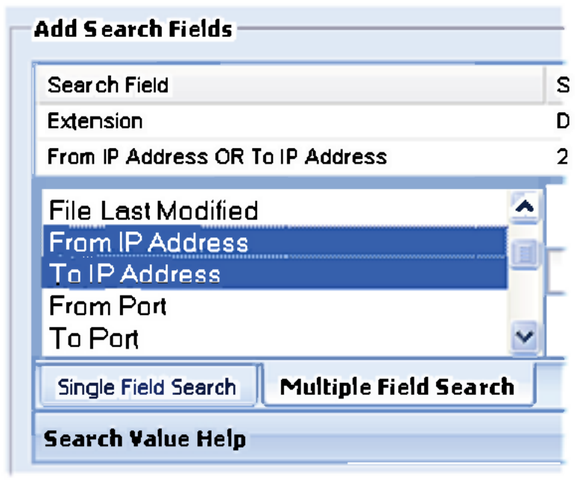
You must hit the green "+" symbol to enter the search criteria.



Click Next

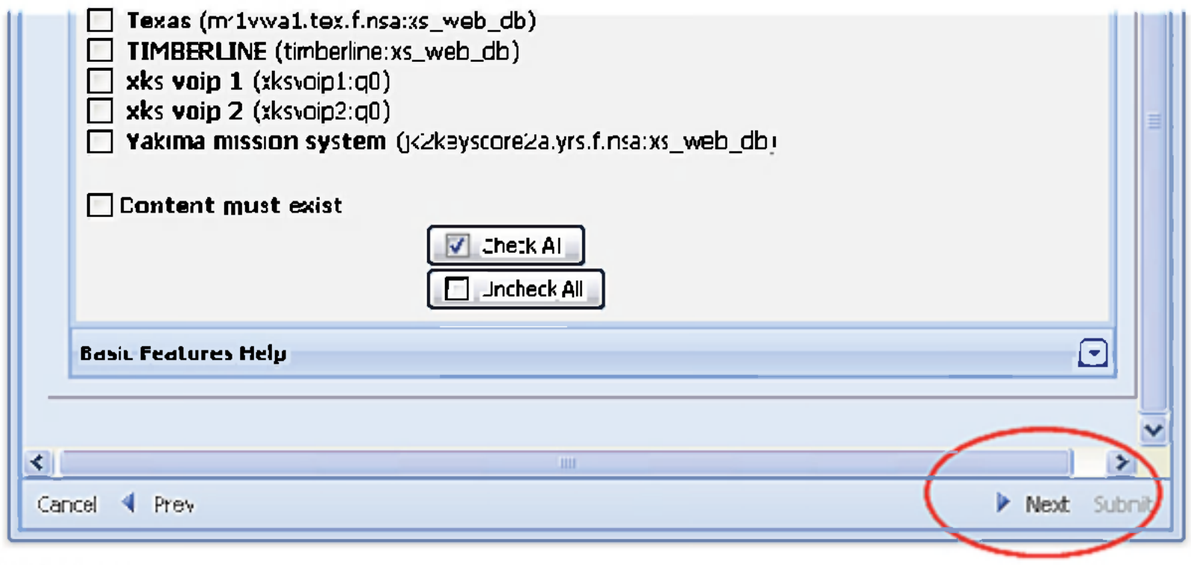
Single Field Search only searches in one field (e.g. File Extensions)

Multiple Field Search allows you to search on several fields (e.g., *To IP* **AND** *From IP*)



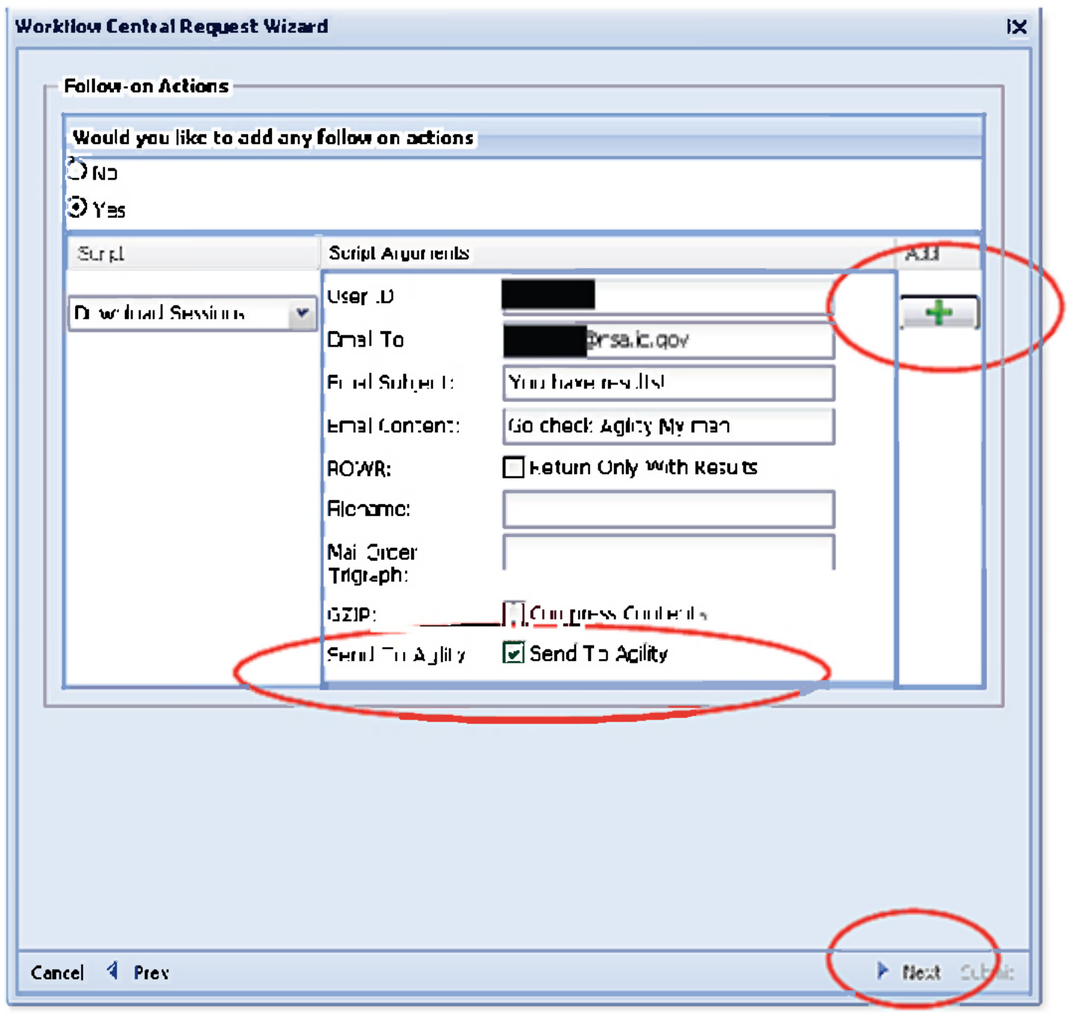
Next, you will select the sites where you want your query to run. Scroll down in this window to use the convenient "Select All" or "Uncheck All" buttons.

NOTE: If your selector is NOFORN, you must DESELECT sites that are 2nd/3rd party.



Click Next

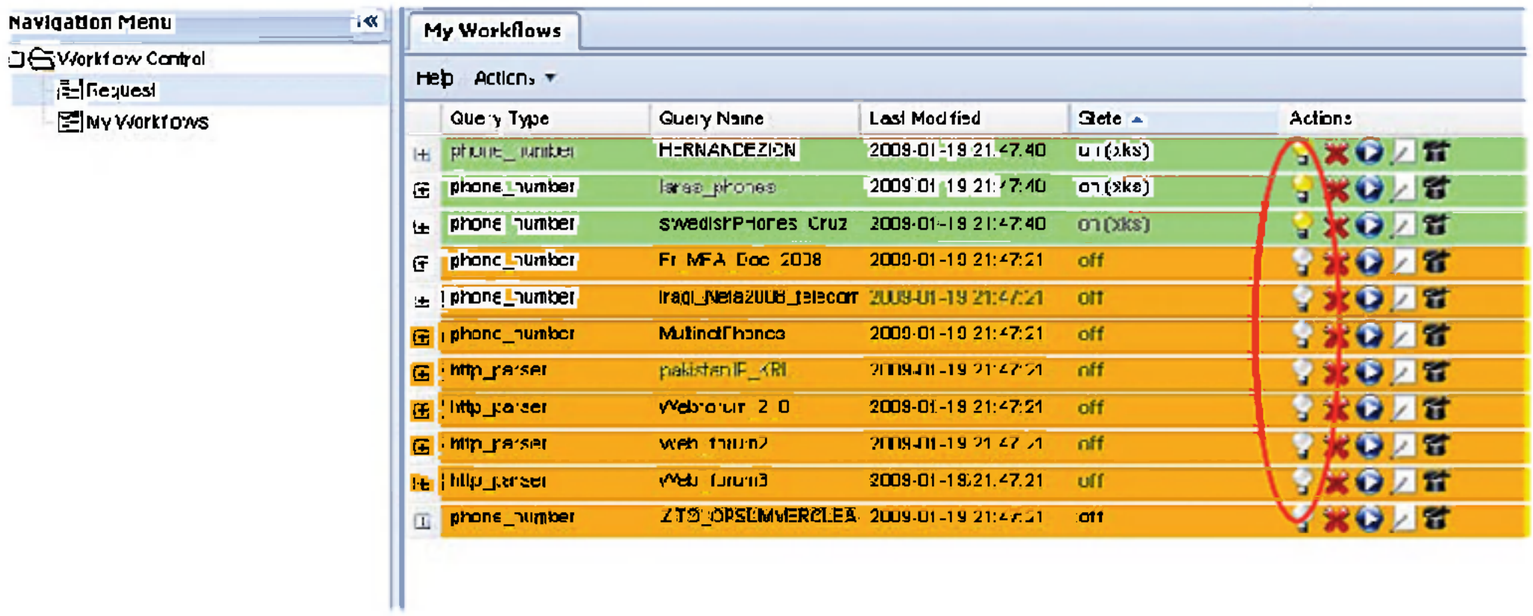
Follow-on Actions tell XKEYSCORE to do things after it runs your query. For example, it can email you with the results, or it can send them to Agility, or any combination of the two. For this example, I want XKEYSCORE to email me telling me I have results and I want it to download my results to Agility. Make sure you select Send to Agility if you want the same.



Click the Green Add symbol, and then click next when finished.

On the next screen, enter any comments you wish (optional) and click Next

Lastly, click SUBMIT. Your query isn't active yet. The XKEYSCORE team will review it and you will have to check back later and turn the query ON or OFF as you wish.



Searching - Tips and Tricks

The Official XKEYSCORE Frequently Asked Questions page is located here: http://xkeyscore.r1.r.nsa/redmine/wiki/xkeyscore/FAQ. Here are some other tips/tricks that may be useful

1. Underscores in usernames:
If your selector has an underscore in it, you must precede the underscore with a backslash. For example: **abu_jihad** would become searched as **abu_jihad**. If you leave the underscore in the query without the backslash, you are wildcarding a single character (see below).

To search on: abu_jihad@hotmail.com:
Bad query: Abu_jihad
Good query: Abu_jihad

If you search on "abu_jihad" (without the backslash), you could bring back "abu**1**jihad", "abu**T**jihad", "abu**S**jihad", "abu-jihad", etc… because you are wildcarding that character and therefore you would be pulling on an entirely different selector.

2. To search on a range of IP addresses:

IP Address Range:
202.82.86.224 - 202.82.86.244

Becomes this XKEYSCORE Query (entered in the IP Address as To, From, or Either):
regex:202\.82\.86\.22[4-9] OR regex:202\.82\.86\.23[0-9] OR regex:202\.82\.86\.24[0-4]

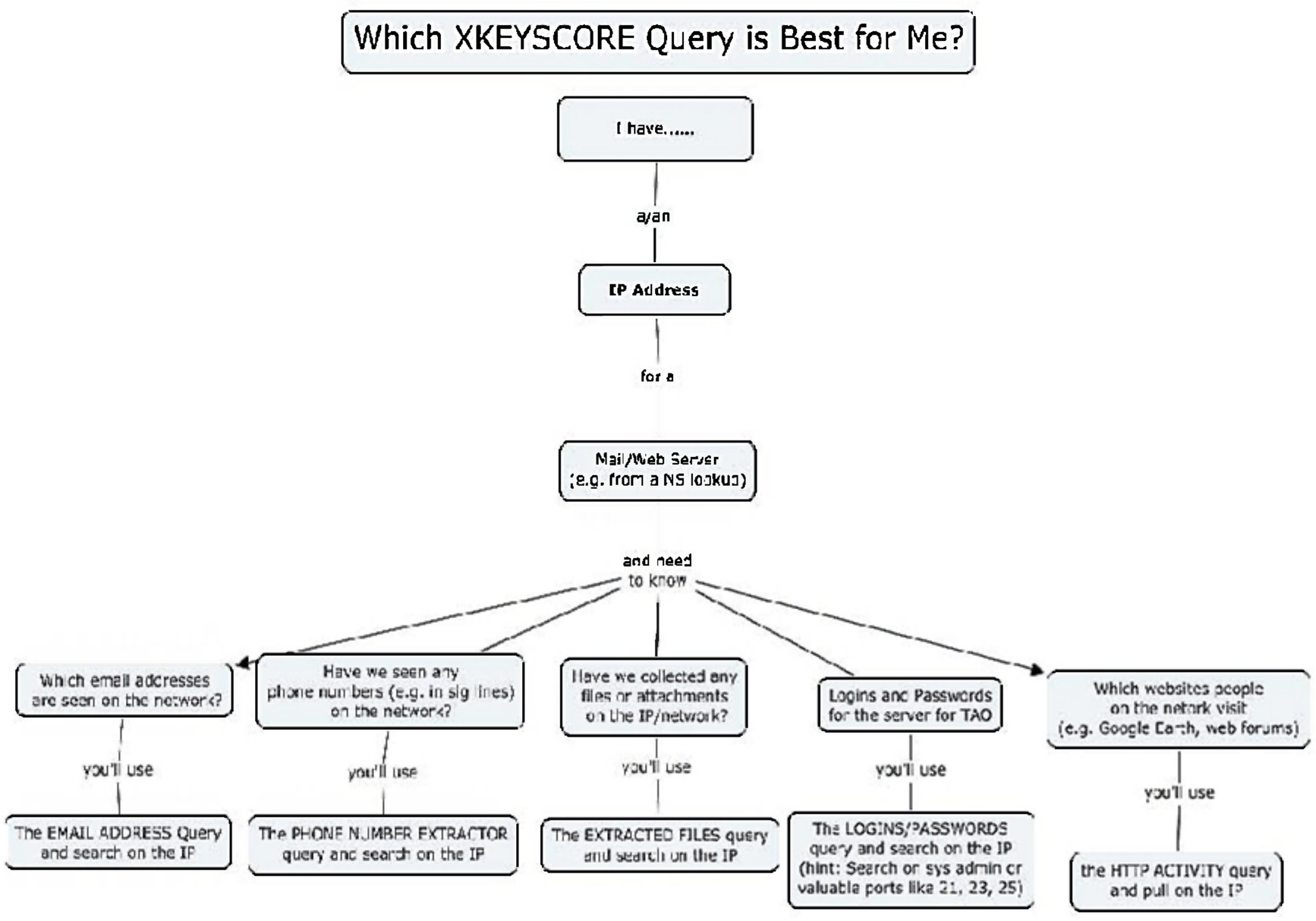
3. Boolean Search Descriptions (Wildcards, ANDs, ORs, etc):

OPERATOR	DESCRIPTION	USAGE
!	Not Equal Comparison	beginning of word (i.e. !joe and !sam)
or	Logical OR (Search for multiple values)	between words (i.e. osama or laden)
and	Logical AND (Search for combination value matches)	between words (ie. *osama* and *laden*) takes precedence over ORs
*	Multiple Character Wildcard	anywhere in word (i.e. *osam*bin*laden)
_	Single Character Wildcard	anywhere in word (i.e. _sam__bin_laden)
>	Greater Than Comparison	beginning of word (i.e. >00080 and <00111)
<	Less Than Comparison	beginning of word (i.e. >00080)
regex:	REGEX Expression	(i.e. to retrieve only numbers: regex:[0-9]*)

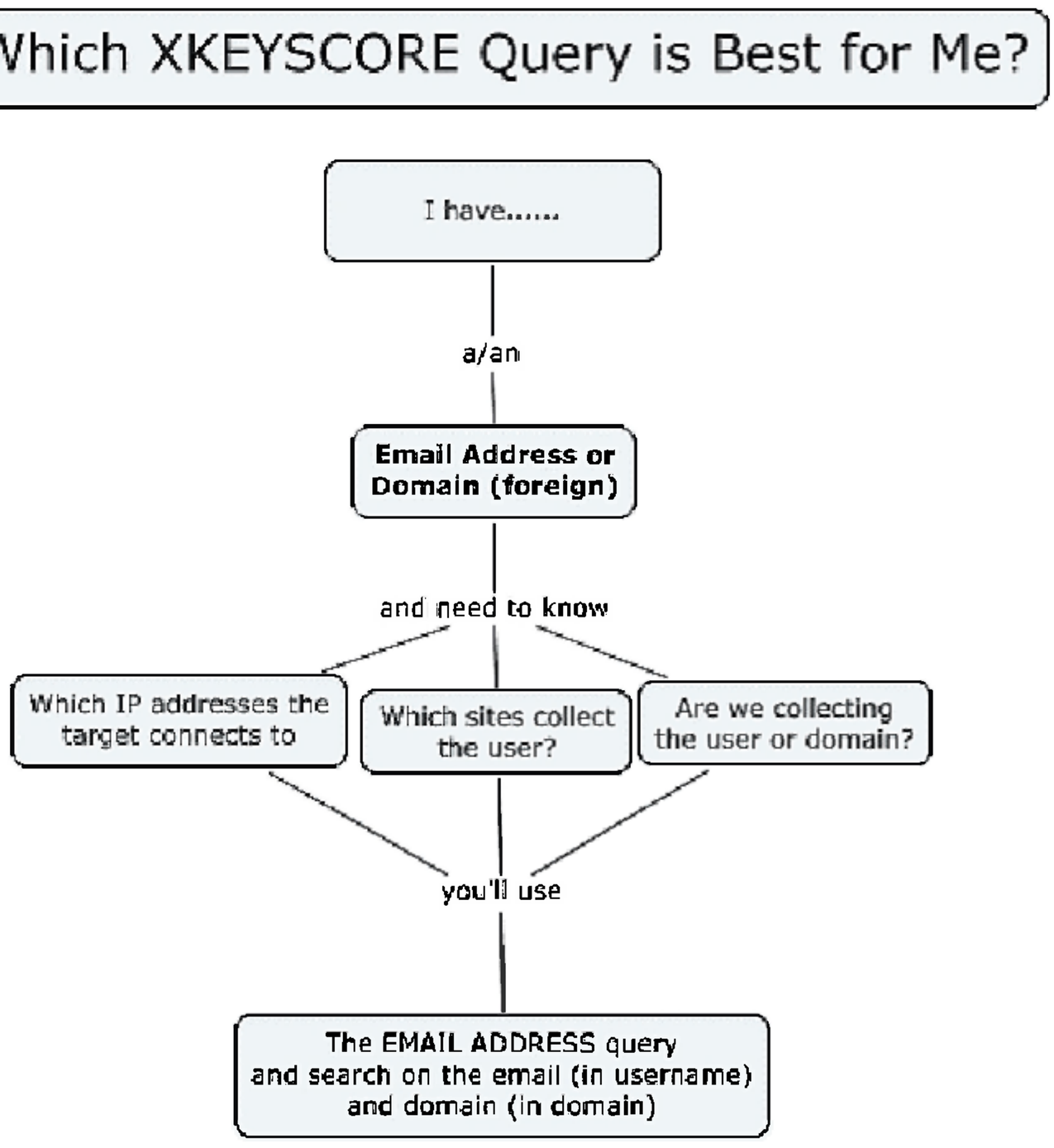
Which Query is best for me?

Quite often the most difficult part of using XKEYSCORE is deciding which query to use at which time. Here's a rough guide to help you decide.

Do you have an IP Address and want to learn more about that network



Do you have an Email Address or Foreign Domain
And want to learn more about it?



Do you have a phone number for your target and want to learn their email address?