



XKEYSCORE Search Forms

March 2009



Standard search fields

■ Wildcards

- * - multiple characters anywhere in word
- _ - single character anywhere in word
- Some fields are auto-wildcarded - the field name will have a * before and/or after it

■ Operators

- Boolean AND, OR - do not use in the same field
- ! - NOT (e.g. !joe)
- <, > - comparison (e.g. >00080)
- regex: - regular expression (e.g. regex:[0-9]*)
- Enter ! to require a field to be non-empty



Special (full-text) search fields

- Google-like syntax - just list your terms and the query will return sessions that match any of them
- Wildcards only allowed at the end of a word
- Search terms must be at least 4 characters
- Use + or - to require that a word must or must not be present
- Use "" to find an exact phrase
- Use () for grouping
- You can still use "classic" syntax - we convert it for you



Special (full-text) search fields

■ Examples:

Search terms	Returned results...
apple banana	contain 'apple' or 'banana' or both
+apple +juice	contain both 'apple' and 'juice'
+apple -macintosh	contain 'apple' but not 'macintosh'
+apple +(turnover strudel)	contain 'apple' AND either 'turnover' or 'strudel'
apple*	contain words like 'apple' or 'apples' or 'applesauce' or 'applet'
"apple juice"	contain the exact phrase 'apple juice'



This plug-in has no data!

- Under development
 - Menu items and search forms may show up before a plug-in goes “live” in the field
- Limited deployment
 - Some sites run different sets of plug-ins
- Populated by front end
 - Some plug-ins simply database metadata provided by the system that feeds XKS, and not all sites are set up the same way



Simple Logging



Full Log DNI

- One record for every session processed
- Collection fields
 - SIGAD
 - Casenotation
 - Session ID (UUID)
- Protocol fields
 - MAC addresses
 - IP addresses
 - Port numbers



Full Log DNI

- Application ID fields
 - Application Name – full application ID
 - Application Type – top level of application ID
 - Application Info – extra info
 - Appid+fingerprints – full application ID plus any matching fingerprints
- Example:
 - Application name: mail/webmail/yahoo
 - Application type: mail
 - Application info: viewFolder_webmail



Full Log DNI

- Fields populated by other plug-ins
 - Username (from User Activity)
 - Category hits (from Category DNI)
 - Client IP/X-Forwarded-For (from Web Proxy)
- Most Full Log search fields are available on every other search form



Email Addresses

- Anything that looks like an email address
- Searchable fields
 - Email username – the part before the @ only!
 - Domain – the part after the @
 - Subject – email subject, if present
- Example:

Sender: user1@yahoo.com

MIME-Version: 1.0

Subject: **check this out**

Date: Tue, 02 Jan 2007 13:27:31 -0000

Message-ID: <AAAAAAAAAAAAAAAAAAAA123456789@yahoo.com>

From: "User One" <user1@yahoo.com>

To: "User Two" <user2@hotmail.com>



Logins & Passwords

- Anything that looks like a login or password
- Searchable fields
 - Username
 - Password
- Examples:

```
<input name="username" value="badguy">
```

```
<input name="password" value="asdf123">
```

```
USER badguy
```

```
PASS asdf123
```



Phone Numbers in DNI

- Anything that looks like a phone number
- Searchable fields
 - Phone number
 - Number type (fax, telephone, mobile, etc.)
- Example:

--

John Smith

Executive Assistant

Phone: 555-1234

Fax: 555-2345



Dictionary Scanning



Alert

- Log of sessions tipped to TRAFFICTHIEF
- Searchable fields
 - Target (strong selector)
 - Weight (confirmed/unconfirmed)
- Other fields
 - Permutation that triggered the tip (DECODEORDAIN)
 - Copy of XML document sent to TRAFFICTHIEF



Category – DNI

- Category hits from CADENCE and other dictionaries
- Searchable fields
 - Dictionary
 - Category
 - Keywords
 - Target (TRAFFICTHIEF)



User Activity



User Activity

- Metadata from applications with a strong selector – webmail, chat, webcam
- Searchable fields
 - Active username (“search value”)
 - Activity – what the active user was doing
 - Attribute type – type of metadata
 - Attribute value – metadata value
 - Source – which plug-in provided the data



User Activity

■ Example:

Username	Activity	Source	Attribute type	Attribute value
badguy@yahoo	viewFolder_webmail	appproc	app_provider	Yahoo
badguy@yahoo	viewFolder_webmail	appproc	app_type	webmail
badguy@yahoo	viewFolder_webmail	appproc	direction	client
badguy@yahoo	viewFolder_webmail	appproc	previous_user	user@yahoo
badguy@yahoo	viewFolder_webmail	appproc	user_realm	yahoo
badguy@yahoo	viewFolder_webmail	appproc	via	squid/2.5
badguy@yahoo	viewFolder_webmail	appproc	x-forwarded_ip	10.0.123.45
badguy@yahoo	viewFolder_webmail	appproc	yahooBcookie	zxcv1234
badguy@yahoo	viewFolder_webmail	appproc	yahooGSS	asdf1234asdf
asdf1234asdf	viewFolder_webmail	appproc	user_realm	yahooGSS
asdf1234asdf	viewFolder_webmail	appproc	yahoo	badguy@yahoo



Document Processing



Extracted Files

- Log of files transmitted as email attachments, web uploads, etc.
- Searchable fields
 - Filename
 - File extension
 - File type/MIME type



Document Tagging

- Document bodies and email bodies are labeled with hits from a custom second-level dictionary
- Idea: “embassy” by itself is not so interesting, but inside a Word document, maybe it is
- Searchable fields
 - Filename
 - Tech name (tag/category) - government, monetary, proliferation, satellite, wireless, etc.
 - Tech value - word or phrase that hit
- Note: also called Tech Strings search



Document Metadata

- Metadata from Office docs, videos, etc.
- Searchable fields
 - Filename and extension, document type
 - Author, organization
 - Language
 - Unique ID
 - Creation/modification timestamps
 - Hash of the entire document and any embedded images



PDF Metadata

- Metadata from PDF documents
- Searchable fields
 - Unique ID
 - Filename
 - Title
 - Author, creator, producer
 - Version
 - Language
- Also available in Document Metadata search



Protocol Processing



Blackberry

- Id numbers and payload info from Blackberry devices
- Searchable fields
 - Source and destination PIN and BES
 - Direction
 - Payload type and encoding



Cellular DNI

- Metadata from DNI over cellular modems
- Searchable fields
 - IMSI, TMSI, IMEI, MCC, RAC, TLLI, etc.
 - Cell ID, Tunnel ID, Access point
 - Latitude, longitude
 - Spotbeam, direction
- Limited deployment - populated in SOTF by front end



Cisco Passwords

- Logs Cisco router passwords
- Searchable fields
 - Password
 - Decoded password (simple obfuscation with known key)



HTTP Activity

- Metadata from HTTP (WWW) traffic
- Searchable fields
 - Host, URL file path, URL query string
 - Search terms - parsed from URLs for common search providers (Google, Yahoo)
 - Language, character encoding
 - Referrer
 - User-Agent
 - Cookies
 - Server type (Apache, etc.)
 - Via - proxy info
 - Geolocation info - e.g. city names from weather reports

IKE



- Metadata from IKE (Internet Key Exchange) sessions
- Searchable fields
 - Version
 - Vendor ID
 - Encryption parameters - key length, field size, group curve, etc.
 - Cookies
 - Nonce



IRC Café Geolocation

- QUIT messages from IRC - Internet cafés often configure their IRC clients to advertise the café's street address
- Searchable fields:
 - Username
 - Nick name
 - Café address (the QUIT message)
 - Café IP
- Example:

```
:nickname!~username@10.0.27.134 QUIT :Quit: MainStreet  
Internet Cafe, 350 Main Street, P4 , 512M, Webcam, MP3,  
128Kbps
```



Passport detection

- Detect images of passports (code from R6)
- OCR machine-readable information
- Searchable fields
 - Original filename
 - Passport detection score
 - Info from machine readable area - name, passport number, issuing state, DOB, expiration, etc.
- **Under development**



Radius Logs

- Metadata from RADIUS sessions for dial-up authentication and IP assignment
- Searchable fields
 - Username
 - Phone number
 - IP address
 - Account information



RBGAN

- Metadata from RBGAN satellite internet terminal collection
- Searchable fields
 - Username
 - IMEI
 - Latitude and longitude
 - Spotbeam and direction
- Limited deployment - populated by front end



- Metadata from RTP audio and video sessions
- Searchable fields
 - Payload type
 - SSRC
 - Number of bytes and packets
 - Timestamps and sequence numbers
- The RTP formatter in the session viewer can decode certain payload types into playable audio or video



- Metadata from SIP (Session Initiation Protocol) used for VoIP setup, etc.
- Stored as multiple type-value pairs per session
- Searchable fields
 - Message type
 - Attribute type (call-id, content-type, from, to, user-agent, via, etc.)
 - Attribute value
 - Subsession ID

SSL



- Metadata from SSL sessions
- Searchable fields
 - Version
 - Encryption parameters - key length, modulus, exponent, etc.
 - Signature info
 - Certificate info



TOR Log

- Logs any identified TOR routers used for anonymizing Internet traffic
- Searchable fields
 - TOR from server
 - TOR to server
 - Router nickname



Web File Transfer

- Log of uploads and downloads from public file-sharing sites (rapidshare, depositfiles, etc.)
- Searchable fields
 - Filename
 - File size
 - Number of downloads
 - Uploader
 - Username and password
- Under development (GCHQ/MHS)



Web Proxy

- Log of X-Forwarded-For IP addresses and other leaked public/private IP
- Currently contains XFF plus leaked info from STUN and Google Earth
- Searchable fields
 - Internal from IP
 - Internal to IP
 - External from IP
 - External to IP
 - Source - plug-in that provided the info
 - Network path - chain of XFF addresses



Wireshark

- Metadata from various protocols processed by the wireshark library
- Protocols
 - Routing - BGP, OSPF
 - VoIP - H225, Skinny, Clarent, Megaco, SCTP
 - Net management - SMB, SNMP
 - Tunneling - GTP
- Searchable fields
 - Protocol
 - Field name
 - Field value



WLAN

- Metadata from WLAN collection
- Searchable fields
 - Channel
 - SSID
 - BSSID
 - MAC addresses
 - Username
 - Private IP
- Limited deployment - populated in SOTF by front end



Miscellaneous



Call Logs

- DNR metadata from JUGGERNAUT, CERF, FASCIA, DURT, etc.
- Searchable fields
 - Phone numbers
 - Signaling type
 - OPC, DPC, CIC, IMSI
- Limited deployment



Network Logs

- Network metadata from MOONSHINE logs
- Searchable fields
 - Net type
 - ESSID, BSSID
 - Channel
 - Carrier
 - Latitude and longitude
- Limited deployment

CNE



- CNE data from TAO
- Searchable fields
 - Project name
 - Collection technique
 - Filename and extension
- Limited deployment (xks-cne)



Registry

- Windows registry data from TAO (CNE)
- Searchable fields
 - Collection technique
 - Hive
 - Key, subkey, value
- Limited deployment (xks-cne)



Custom Search Forms



Simple Search

- Simple way to search for usernames, IP addresses, and machine ID cookies
- Just enter your search term and select what type of thing it is, and the form sends it to User Activity or HTTP Activity as appropriate



Multi Search

- Problem: XKS may have info about “badguy@yahoo” in Email Addresses, User Activity, Logins & Passwords, etc.
- Solution: submit multiple searches from a single form
- Enter the username and select which databases to search, and the form translates that into the proper queries
- Similar MultiSearches for IP addresses and MAC addresses
- Optional: merge results into one table