XKEYSCORE

25 Feb 2008

xkeyscore@nsa

DERIVED FROM: NSA/CSSM 1-52
DATED: 20070108
DECLASSIFY ON: 20320108

What is XKEYSCORE?

1. DNI Exploitation System/Analytic Framework

2. Performs strong (e.g. email) and soft (content) selection

3. Provides real-time target activity (tipping)

4. "Rolling Buffer" of ~3 days of ALL unfiltered data seen by XKEYSCORE:
   - Stores full-take data at the collection site – indexed by meta-data
   - Provides a series of viewers for common data types

5. Federated Query system – one query scans all sites
   - Performing full-take allows analysts to find targets that were previously unknown by mining the meta-data

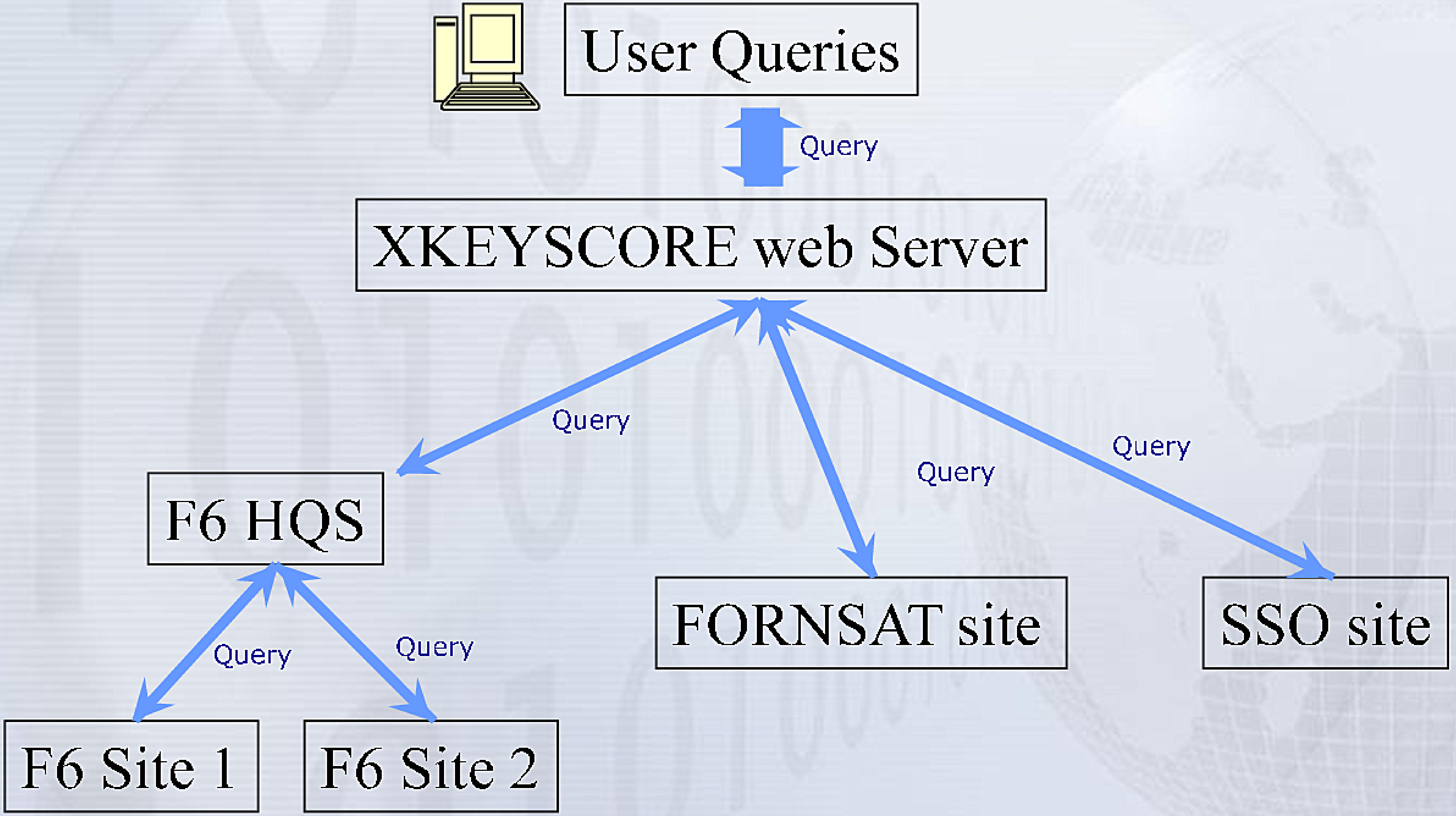
Methodology

- Small, focused team
- Work closely with the analysts
- Evolutionary development cycle (deploy early, deploy often)
- React to mission requirements
- Support staff integrated with developers
- Sometimes a delicate balance of mission and research

System Details

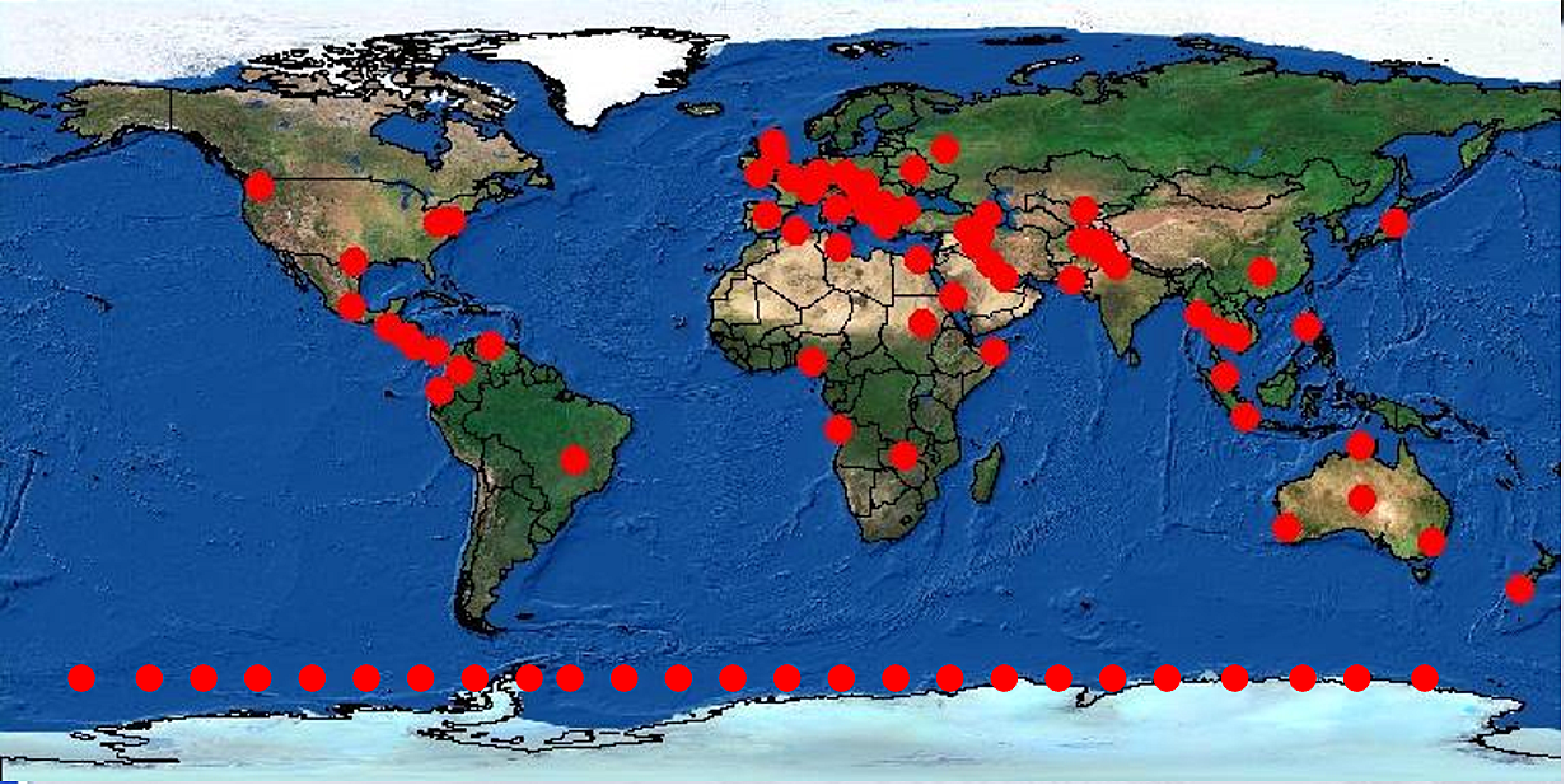
- Massive distributed Linux cluster
- Over 500 servers distributed around the world
- System can scale linearly – simply add a new server to the cluster
- Federated Query Mechanism

Query Hierarchy



TOP SECRET//COMINT//REL TO USA, AUS, CAN, GBR, NZL

Where is X-KEYSCORE?



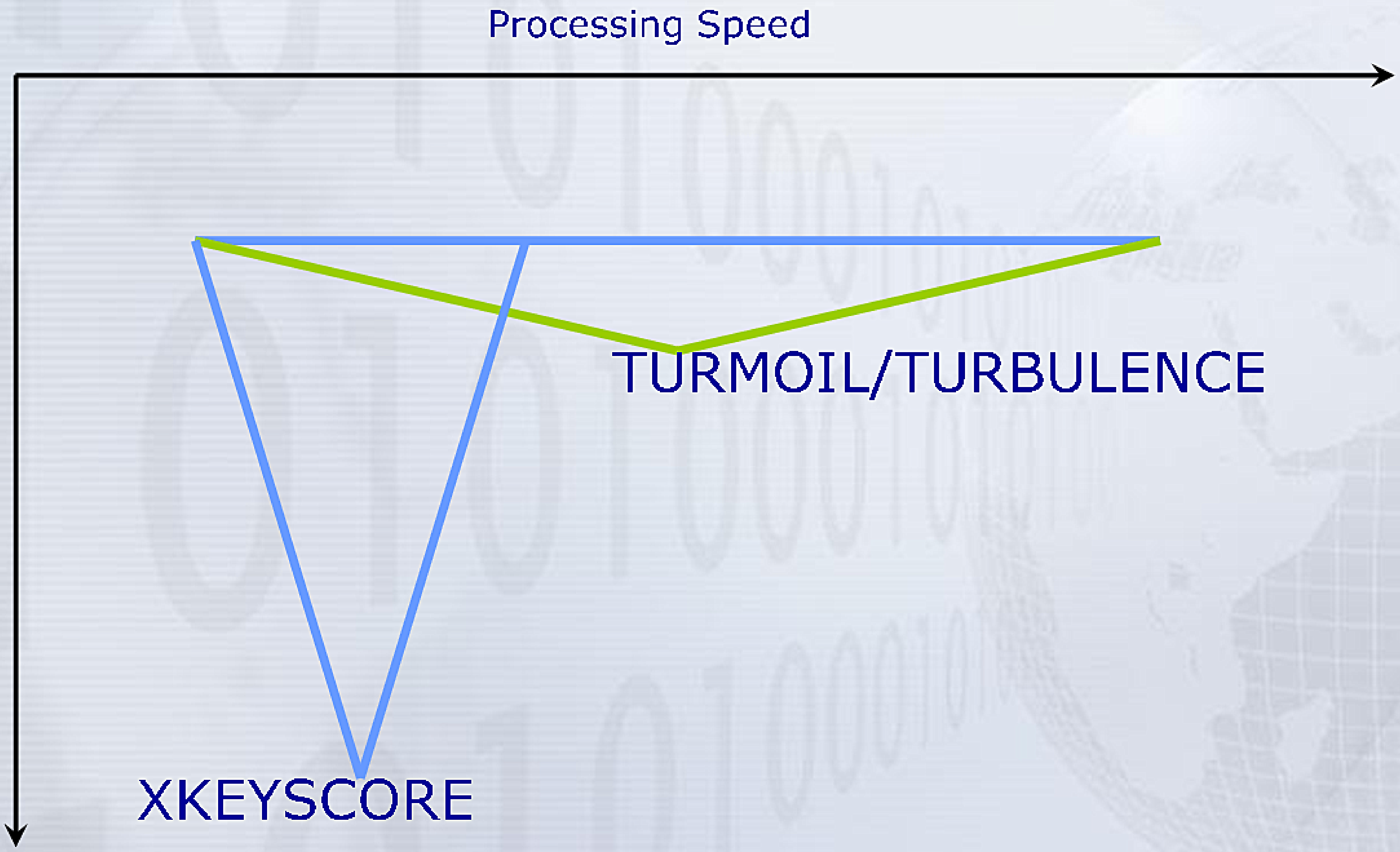
Approximately 150 sites

Over 700 servers

TOP SECRET//COMINT//REL TO USA, AUS, CAN, GBR, NZL

What is unique about XKEYSCORE?

General Capability



Why do shallow

- Can look at more data
- XKEYSCORE can also be configured to go shallow if the data rate is too high

Why go deep

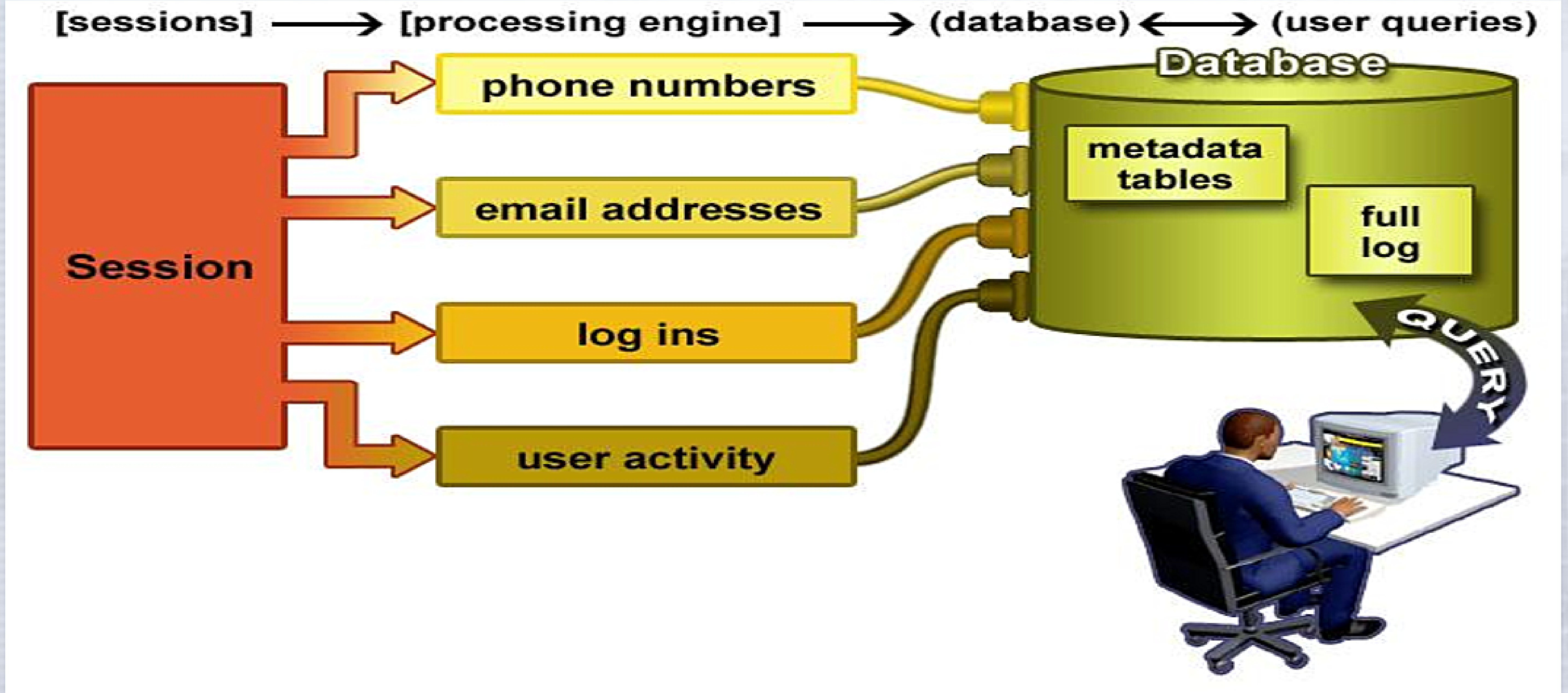
- Strong Selection itself give us only a very limited capability

- A large amount of time spent on the web is performing actions that are anonymous

- We can use this traffic to detect anomalies which can lead us to intelligence by itself, or strong selectors for traditional tasking

What XKS does with the Sessions

Plug-ins extract and index metadata into tables

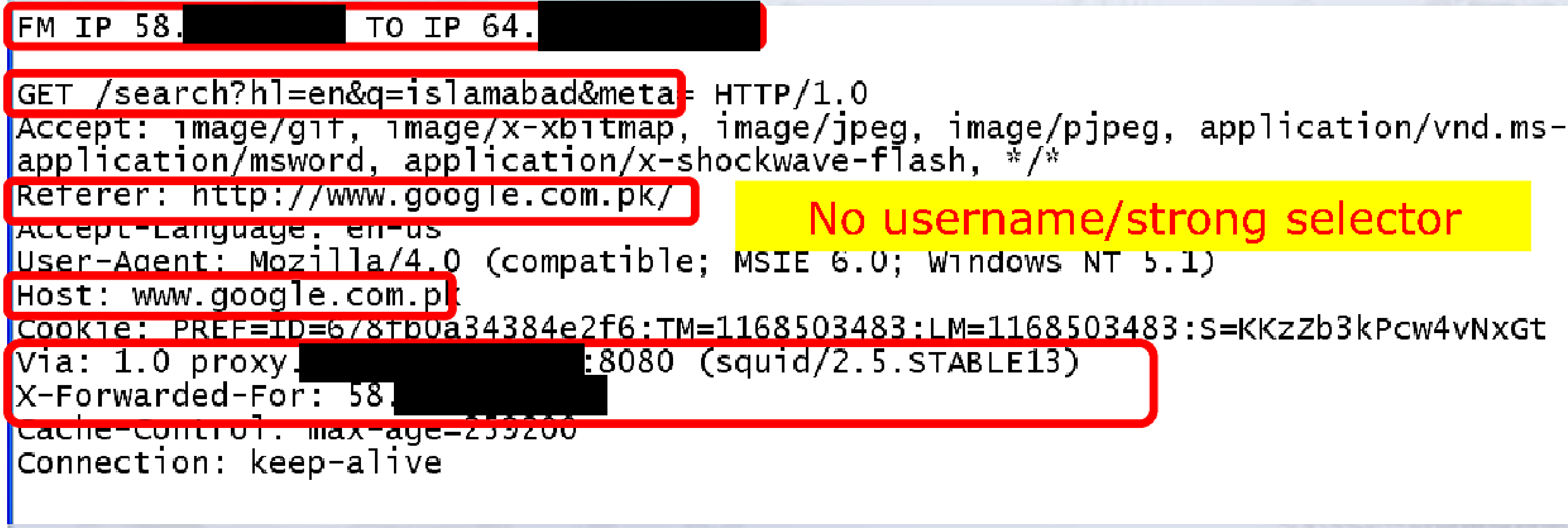


Plug-ins

Plug-in	DESCRIPTION
E-mail Addresses	Indexes every E-mail address seen in a session by both username and domain
Extracted Files	Indexes every file seen in a session by both filename and extension
Full Log	Indexes every DNI session collected. Data is indexed by the standard N-tupple (IP, Port, Casenotation etc.)
HTTP Parser	Indexes the client-side HTTP traffic (examples to follow)
Phone Number	Indexes every phone number seen in a session (e.g. address book entries or signature block)
User Activity	Indexes the Webmail and Chat activity to include username, buddylist, machine specific cookies etc.

What Can Be Stored?

- Anything you wish to extract
  - Choose your metadata
  - Customizable storage times
  - Ex: HTTP Parser



```
FM IP 58.[redacted] TO IP 64.[redacted]

GET /search?hl=en&q=islamabad&meta= HTTP/1.0
Accept: image/gif, image/x-xbitmap, image/jpeg, image/pjpeg, application/vnd.ms-
application/msword, application/x-shockwave-flash, */*
Referer: http://www.google.com.pk/
Accept-Language: en-us
User-Agent: Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.1)
Host: www.google.com.pk
Cookie: PREF=ID=678fb0a34384e2f6:TM=1168503483:LM=1168503483:S=KKzZb3kPcw4vNxGt
Via: 1.0 proxy.[redacted]:8080 (squid/2.5.STABLE13)
X-Forwarded-For: 58.[redacted]
Cache-Control: max-age=259200
Connection: keep-alive
```

No username/strong selector

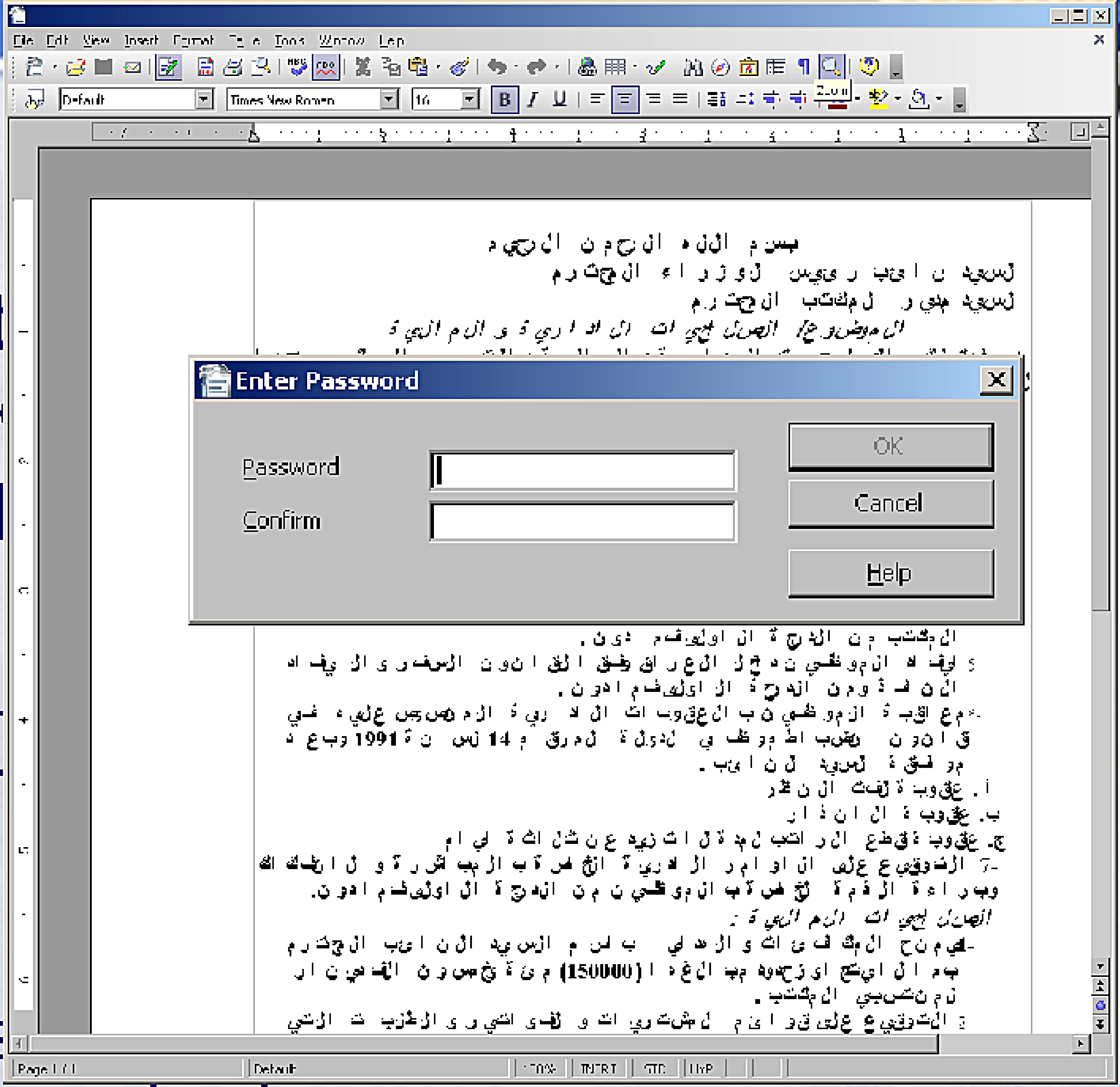
What can you do with XKEYSCORE?

Finding Targets

- How do I find a strong-selector for a known target?
- How do I find a cell of terrorists that has no connection to known strong-selectors?
- Answer: Look for anomalous events
  - E.g. Someone whose language is out of place for the region they are in
  - Someone who is using encryption
  - Someone searching the web for suspicious stuff

Encryption

- Show me all t
  documents fr
- Show me all I
  - Once again –
    forwarding th
  - No strong-se
  - Can perform
    query, then s
    from site as required



بسم الله الرحمن الرحيم
السيد نائب رئيس الوزراء المحترم
السيد مدير المكتب المحترم
الموضوع/ الصلاحيات الادارية والمالية

Enter Password

Password

Confirm

OK

Cancel

Help

المكتب من الدرجة الاولى فما دون.
[illegible]
[illegible] 14 [illegible] 1991 [illegible]
أ. عقوبة لفت النظر
ب. عقوبة الانذار
ج. عقوبة قطع الراتب لمدة لا تزيد عن ثلاثة ايام
-7 [illegible]
الصلاحيات المالية:
[illegible] (150000) [illegible]
[illegible]

Technology Detection

- Show me all the VPN startups in country X, and give me the data so I can decrypt and discover the users
  - These events are easily browsable in XKEYSCORE
    - No strong-selector
  - XKEYSCORE extracts and stores authoring information for many major document types – can perform a retrospective survey to trace the document origin since metadata is typically kept for up to 30 days
  - No other system performs this on raw unselected bulk traffic, data volumes prohibit forwarding

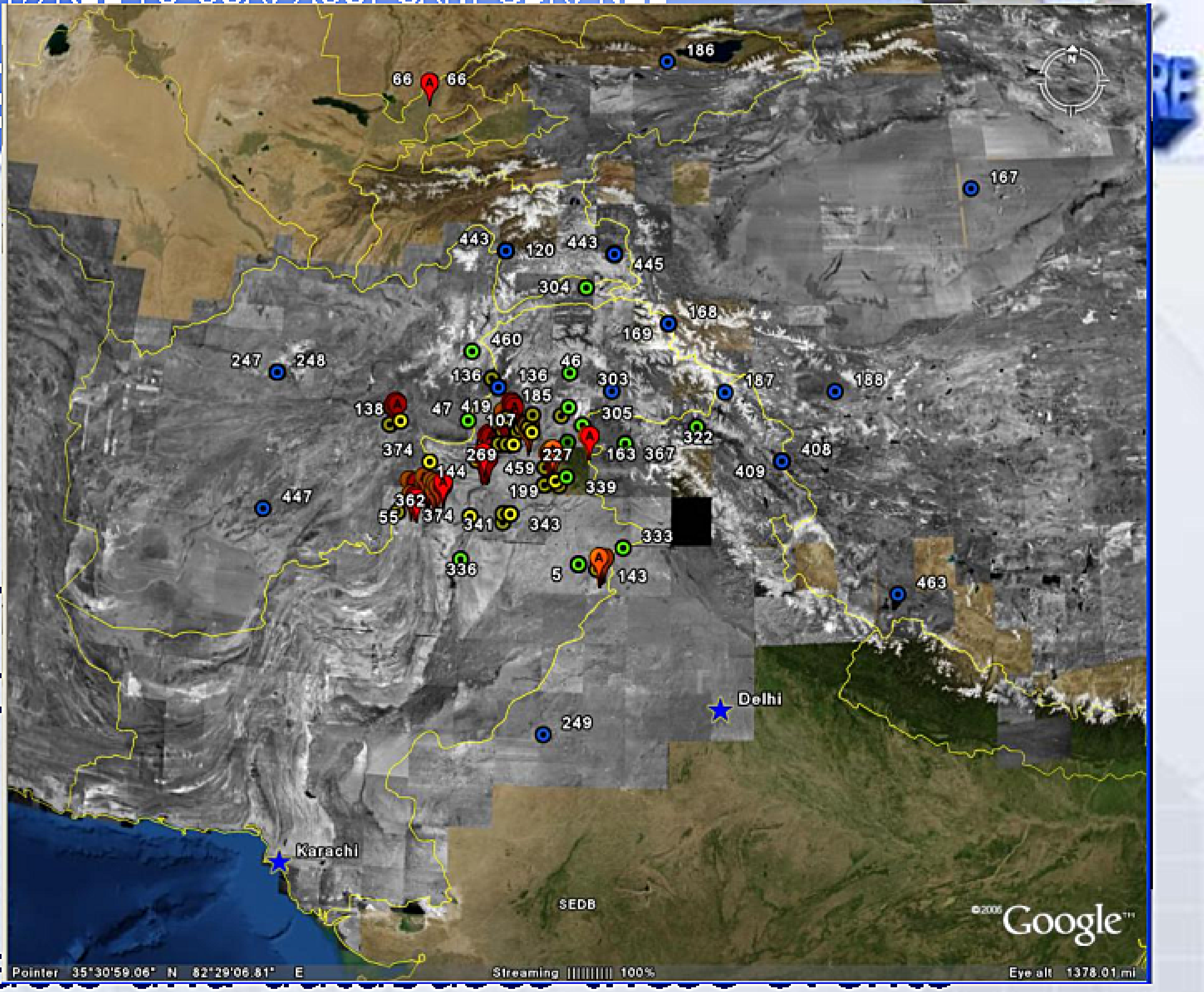
Persona Session Collection

- Traditionally triggered by a strong-selector event, but it doesn't have to be this way
- Reverse PSC – from anomalous event back to a strong selector. You cannot perform this kind of analysis when the data has first been strong selected.
- Tie in with Marina – allow PSC collection after the event

Language Tracking

- My target speaks German but is in Pakistan – how can I find him?
  - XKEYSCORE's HTTP Activity plugin extracts and stores all HTML language tags which can then be searched
  - Not possible in any other system but XKEYSCORE, nor could it be –
    - volumes are too great to forward
    - No strong-selector

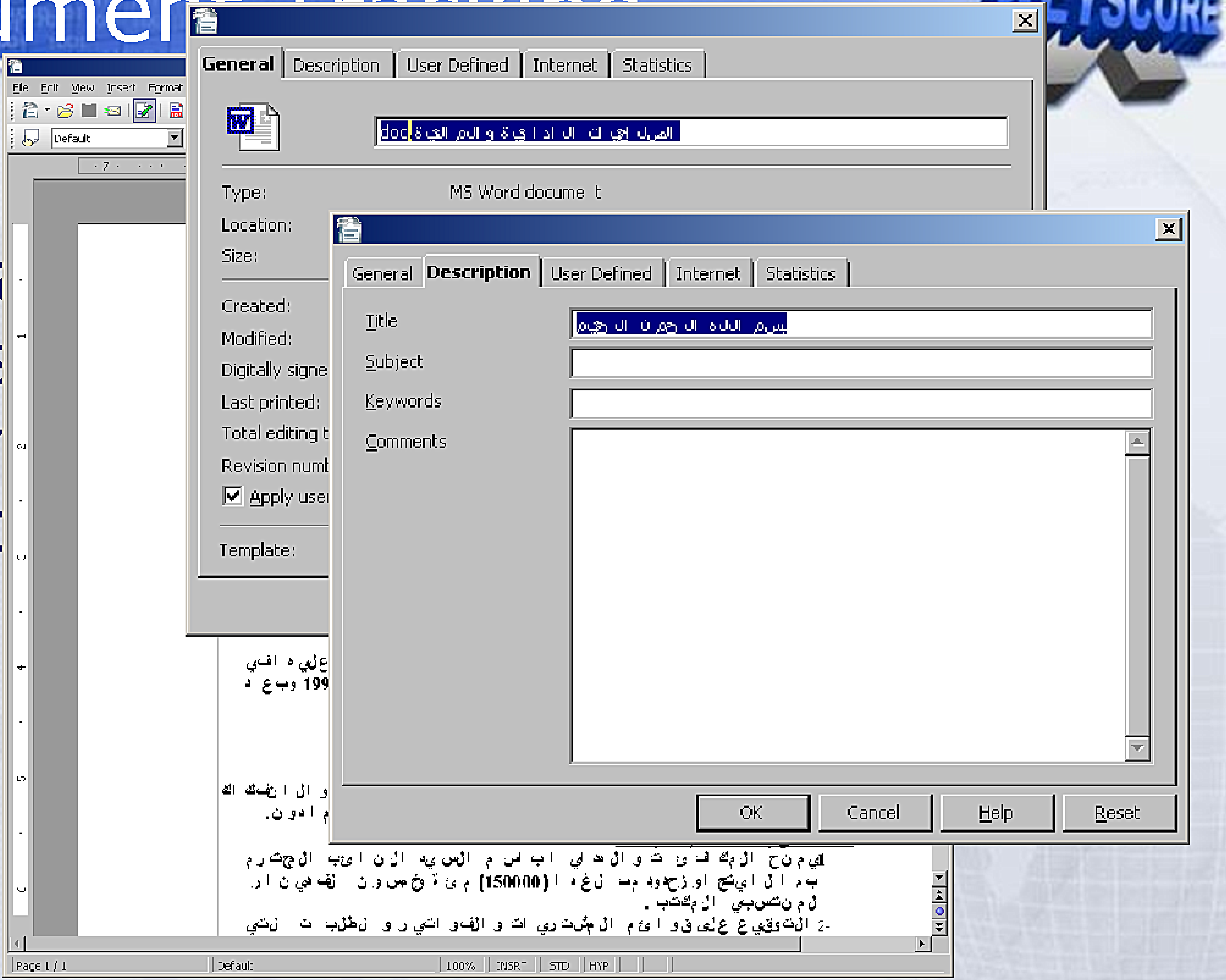
Google Maps



- My target uses G
  locations – can I
  determine his em
  web-searches – c
  suspicious?
  - XKEYSCORE extra
    including all web-based searches which can be retrospectively queried
  - No strong-selector
  - Data volume too high to forward

Document Tracking

- I ha
  has
  nur
  and



doc.المرسل الي ك ال اداري و المالية

Type: MS Word docume t

Location:

Size:

Created:

Modified:

Digitally signe

Last printed:

Total editing t

Revision numb

☑ Apply use

Template:

Title: بسم الله الرحمن الرحيم

Subject:

Keywords:

Comments:

علي ه افي
199 وبع د

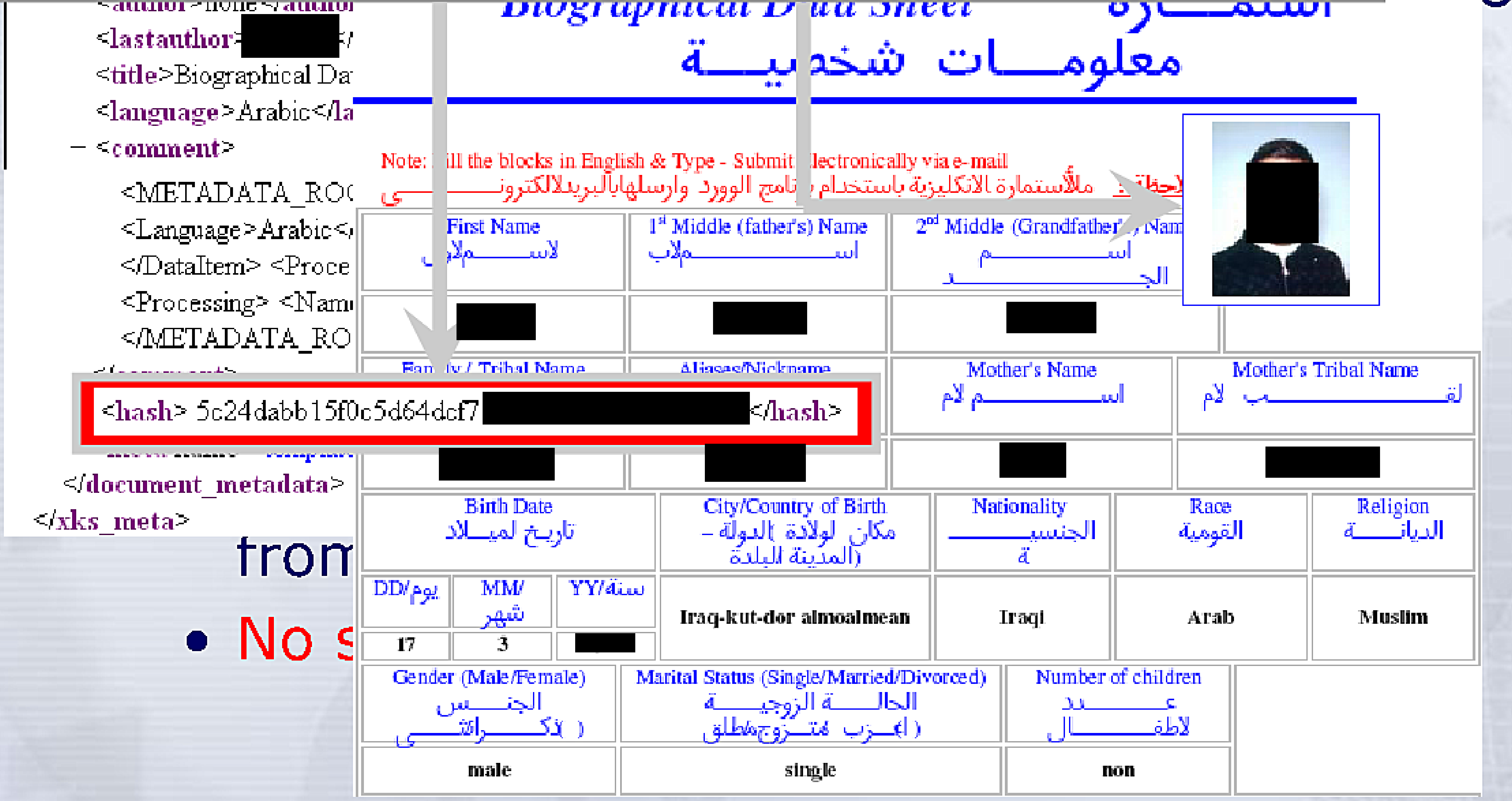
و ال ا نفك الك
م ادو ن.

الي م نح الزمك ف ي ت و ال هدي اب لس م السيه الن ا ي تب الحث ر م
ب م ا ل اي تج او.زجوبه مد نغ ه ا (150000) م ئ ة خمس و ن نفعي ن ا ر.
ل م ن تسبي الكتب .
2- التدقيق على قو ائم المشتريات و الفواتير و نطلب ت نني

Document Tracking

- All images are hashed in the metadata so that you can search for anyone who has received or transmitted this document.
- This is really useful for company logos.

to



```
<author>none</author>
<lastauthor>[redacted]</
<title>Biographical Da
<language>Arabic</la
– <comment>
    <METADATA_ROO
    <Language>Arabic</
    </DataItem> <Proce
    <Processing> <Nam
    </METADATA_RO
  </comment>
<hash> 5c24dabb15f0c5d64dcf7[redacted]</hash>
</document_metadata>
</xks_meta>
```

Biographical Data Sheet استمـــارة
معلومـــات شخصيـــة

Note: Fill the blocks in English & Type - Submit Electronically via e-mail
ملاحظة: ملأستمارة الانكليزية باستخدام برنامج الوورد وارسلهابالبريدلالكترونـــــــــي

First Name لاســـــملاول	1st Middle (father's) Name اســـــــــملاب	2nd Middle (Grandfather's) Name اســــــم الجـــــــد
[redacted]	[redacted]	[redacted]

Family / Tribal Name	Aliases/Nickname	Mother's Name اســـــــم لام	Mother's Tribal Name لقــــــــب لام
[redacted]	[redacted]	[redacted]	[redacted]

Birth Date تاريـخ لميــلاد			City/Country of Birth مكان لولادة (الدولة – المدينة البلدة)	Nationality الجنسيـــــــة	Race القومية	Religion الديانـــــة
DD/يوم	MM/ شهر	YY/سنة	Iraq-kut-dor almoalmean	Iraqi	Arab	Muslim
17	3	[redacted]				

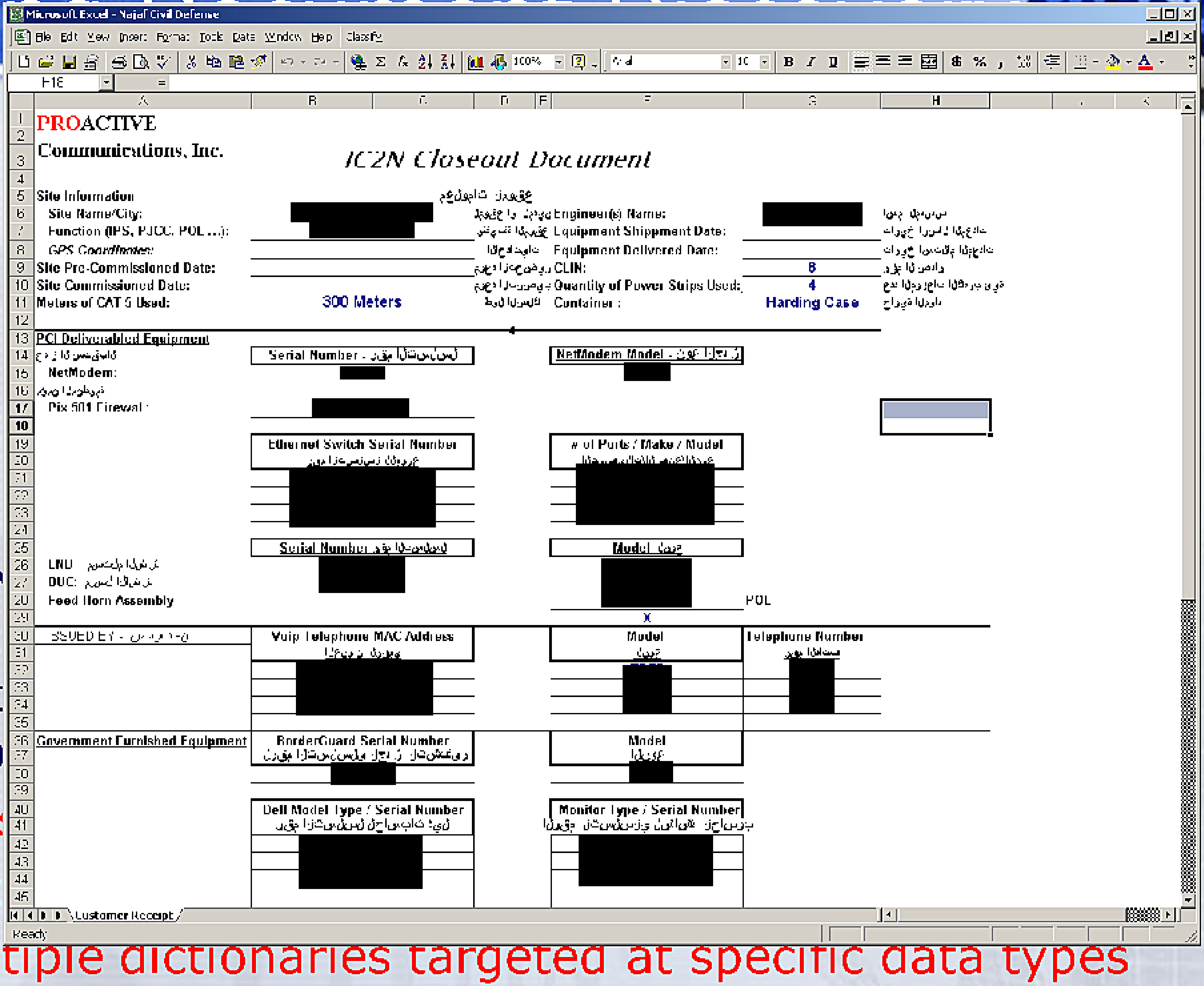
Gender (Male/Female) الجنـــس (ذكـــر/انثـــى)	Marital Status (Single/Married/Divorced) الحالـــة الزوجيـــة (اعـــزب متـــزوج مطلق)	Number of children عـــــدد لاطفـــــال
male	single	non

from

- No s

Interesti

- Show
  conta
  so I c
  - New
    doc
    dict
    info
  - No
  - Dat
  - Multiple dictionaries targeted at specific data types


Microsoft Excel - Najaf Civil Defense
File Edit View Insert Format Tools Data Window Help Classify
100%
F18
=
Customer Receipt
Ready

PROACTIVE
Communications, Inc.

IC2N Closeout Document

Site Information [illegible]
Site Name/City: [redacted] [illegible] Engineer(s) Name: [redacted] [illegible]
Function (IPS, PJCC, POL ...): [illegible] Equipment Shippment Date: [illegible]
GPS Coordinates: [illegible] Equipment Delivered Date: [illegible]
Site Pre-Commissioned Date: [illegible] CLIN: 8 [illegible]
Site Commissioned Date: [illegible] Quantity of Power Strips Used: 4 [illegible]
Meters of CAT 5 Used: 300 Meters [illegible] Container : Harding Case [illegible]

PCI Deliverabled Equipment
[illegible]
NetModem: Serial Number - [illegible] [redacted] NetModem Model - [illegible] [redacted]
[illegible]
Pix 501 Firewall : [redacted]

Ethernet Switch Serial Number [illegible] [redacted]
# of Ports / Make / Model [illegible] [redacted]

Serial Number [illegible] [redacted]
Model [illegible] [redacted]
LNB [illegible]
BUC: [illegible]
Feed Horn Assembly
POL
X

ISSUED BY - [illegible]
Voip Telephone MAC Address [illegible] [redacted]
Model [illegible] [redacted]
Telephone Number [illegible] [redacted]

Government Furnished Equipment
BorderGuard Serial Number [illegible] [redacted]
Model [illegible] [redacted]
Dell Model Type / Serial Number [illegible] [redacted]
Monitor Type / Serial Number [illegible] [redacted]

TAO

- Show me all the exploitable machines in country X
    - Fingerprints from TAO are loaded into XKEYSCORE's application/fingerprintID engine
    - Data is tagged and databased
    - No strong-selector
    - Complex boolean tasking and regular expressions required

XKEYSCORE Success Stories

Over 300 terrorists captured using intelligence generated from XKEYSCORE

XKEYSCORE and TRAFFICTHIEF

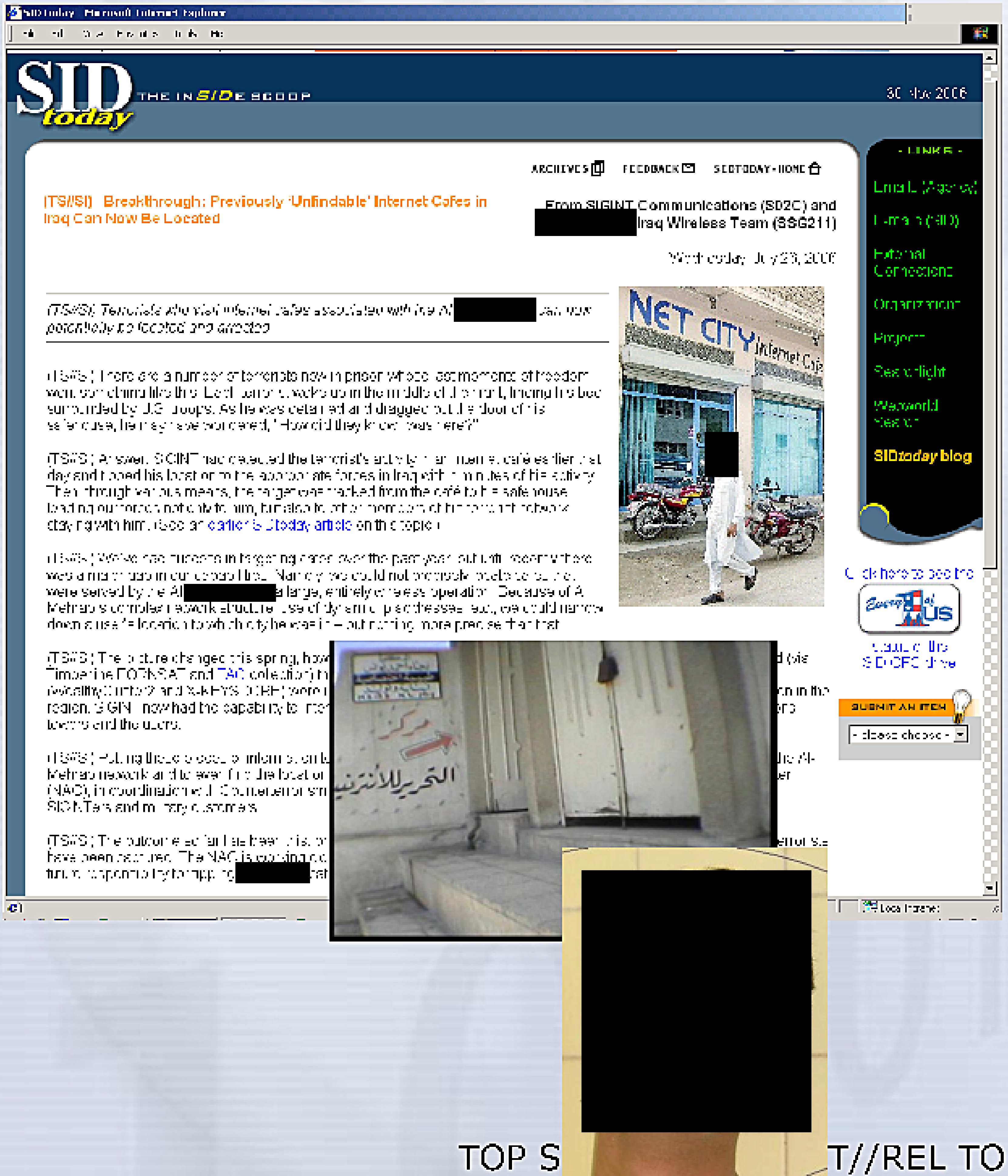
- Customer: CounterTerrorism (CT)
- Provides near real-time tips to TRAFFICTHIEF server in operations in coordination with coalition forces in Iraq 24 hours a day
- Currently producing hundreds of confirmed alerts per day on over 3000 user accounts

Afternoon of [redacted] 2004 – coalition detained individuals below:

[redacted]

[redacted]

XKEYSCORE Success



(TS//SI) Breakthrough: Previously 'Unfindable' Internet Cafes in Iraq Can Now Be Located

From SIGINT Communications (S02C) and [redacted] Iraq Wireless Team (SSG211)

Wednesday, July 26, 2006

(TS//SI) Terrorists who visit internet cafes associated with the Al [redacted] can now potentially be located and arrested

(TS//SI) There are a number of terrorists now in prison whose last moments of freedom went something like this: Local terrorist wakes up in the middle of the night, finding his bed surrounded by U.S. troops. As he was detained and dragged out the door of his safe house, he may have wondered, "How did they know I was here?"

(TS//SI) Answer: SIGINT had detected the terrorist's activity in an internet café earlier that day and tipped his location to the appropriate forces in Iraq within minutes of his activity. Then, through various means, the target was tracked from the café to his safe house, leading our forces not only to him, but also to other members of his terrorist network staying with him. (See an earlier SIDtoday article on this topic.)

(TS//SI) We've had success in targeting cafes over the past year, but until recently there was a major gap in our capabilities. Namely, we could not precisely locate cafes that were served by the Al [redacted] a large, entirely wireless operation. Because of Al-Methrab's complex network structure (use of dynamic IP addresses, etc.), we could narrow down a user's location to which city he was in — but nothing more precise than that.

(TS//SI) The picture changed this spring, how [illegible] d (via Timberline FORNSAT and TAO collection) [illegible] (WealthyCluster2 and X-KEYSCORE) were [illegible] n in the region. SIGINT now had the capability to [illegible] ers towers and the users.

(TS//SI) Putting these pieces of information to [illegible] the Al-Methrab network and to even find the location [illegible] er (NAC), in coordination with Counterterrorism [illegible] SIGINTers and military customers.

(TS//SI) The outcome so far has been [illegible] errorists have been captured. The NAC is continuing [illegible] future, hopefully by tipping [redacted] cat

[redacted]

[redacted]

May 2006, WealthyCluster2 and X-KEYSCORE Installed at [redacted]

- Connected to Moonshine
- Enabled processing of wireless collection
- Enabled near-real-time tipping
- Enabled full-take SIGDEV

Un-locatable cafés were geolocated:

- [redacted]– "A Goldmine"
- Four Other Cafés Being Developed

Acquired important targets:

- NSA/Georgia Tips With Precise Locations
- JSOC Tools In New [redacted]
- Reacquired [redacted] Lost When Zarkanet Went Down

Terrorists were captured:

- Members of the [redacted]
- Members of the [redacted]

Innovation

- High Speed Selection
- Toolbar
- Integration with Marina
- GPRS, WLAN integration
- SSO CRDB
- Workflows
- Multi-level Dictionaries

Future

- High speeds yet again (algorithmic and Cell Processor (R4))
- Better presentation
- Entity Extraction
- VoIP
- More networking protocols
- Additional metadata
  - Expand on google-earth capability
  - EXIF tags
  - Integration of all CES-AppProcs
- Easier to install/maintain/upgrade