CRYPTOME

25 August 2014

https://firstlook.org/theintercept/article/2014/08/25/icreach-nsa-cia-secret-google-crisscross-proton

The Surveillance Engine: How the NSA Built Its Own Secret Google

By Ryan Gallagher 25 Aug 2014

INTELLIGENCE COMMUNITY (IC) REACH TEAM (C// REL TO USA, FVEY) (Submitted by the SID Technical Director)

(C//REL TO USA, FVEY) ICREACH is an NSA information-sharing initiative for customers and analysts across the Intelligence Community overseen by the Director from concept to implementation. This team began over two years ago with a basic concept compelled by the IC's increasing need for communications metadata and NSA's ability to collect, process and store vast amounts of communications metadata related to worldwide intelligence targets. The ICREACH team delivered the first-ever wholesale sharing of communications metadata within the U.S. Intelligence Community. The SID team did considerable work with NSA/OGC (Office of the General Counsel) and SID/Policy to ensure appropriate legal and policy approvals for this new technical capability. In addition, the team briefed the Second Party partners on this effort to set the stage for Second Party participation in ICREACH. Their efforts hastened implementation of this much-needed capability.

TOP SECRET// COMINT// NOFORN//20320108

. (S//REL TO USA, FVEY) Provide IC analysts with access to a wider set of shareable data through the ICREACH Application, ICREACH allows IC analysts the ability to perform more comprehensive analyses for situational awareness and target development in support of their mission objectives.

ICREACH to aid and assist in target development and situational awareness against computer based targets.

• (S//SI//REL TO USA, FVEY) Transition ICREACH services to NSA/CSS' cloud architecture. This will

(S//SI//REL TO USA, FVEY) Expand and grow IC access to Digital Network Intelligence metadata via

- provide the opportunity for increased dissemination of SIGINT data and data enrichment services tailored to meet customer mission needs.
- (S//REL TO USA, FVEY) Respond to the National Intelligence Priorities Framework (NIPF) updates; expand the capabilities of the National SIGINT Requirement's Process (NSRP) and the new/modified INs as required by US SIGINT Consumers (i.e. IC, DoD, Policy Makers, Other Government Agencies). This will allow NSA/CSS to better support the broad goals of ODNI for Mission Management and Information Sharing with the US SIGINT Consumers, and facilitate the ability to obtain SIGINT from NSA/CSS.

This Briefing is Classified TOP SECRET//COMINT//NOFORN//20320108





Sharing Communications Metadata Across the U.S. Intelligence CommunityICREACH

15 May 2007

DERIVED FROM: NSA/CSS Manual 1-52, Dated: 20070108, Declassify On: 20320108



The Need for Greater IC Sharing of Information



1. (U) Recommendations of 9/11 and WMD Commissions

2. (U) Congressional Interest – IRTPA

3. (U) DNI - A part of 100-Day Plan

4. (U) DoD Information Sharing Strategy



Information Sharing Policy



- (U//FOUO) DCID 8/1, June 2004 "All IC agencies will provide intelligence information at the earliest point at which customers can understand and effectively use it to support their mission objectives..."
- (U//FOUO) NSA Policy 1-9, May 2005 implements DCID 8/1 and Transformation 2.0: "SIGINT information originated by NSA/CSS shall be shared ... with U.S. Government customers and partners ... and with Foreign Partners... to the maximum extent possible, consistent with applicable statutes, executive orders and regulations, and consistent with the 'need-to-know' principle and with applicable authorities governing the protection of intelligence sources and methods."



SIGINT Dissemination/Sharing Matrix



	Minimized	Assessed for FI	Shareable	
SIGINT Products & Services	Yes	Yes	Yes	
SIGINT Communications Metadata*	Yes	No, but provided for FI purposes	Yes	
Raw SIGINT	No	No	No	

^{*(}C//REL) Communications metadata refers to structured "data about data": it includes all information associated with, but not including content, and includes any data used by a network, service, or application to facilitate routing or handling of a communication or to render content in the intended format; it includes, but is not limited to; dialing, routing, addressing, or signaling information and data in support of various network management activities (e.g. billing, authentication or tracking of communicants).



Sharing Matrix - ICREACH



	Minimized	Assessed for FI	Shareable
COMINT Communications Metadata	Yes	No, but provided for FI purposes	Yes
	CRE	AC	



Building on IC Success – Information Sharing



(C//REL) Ensure efficient, effective sharing of ever expanding volumes of IC communications metadata. The IC requires access to a much greater volume of data, for a larger number of analysts.

- 1. Make more data accessible
- 2. Wider set of IC Analysts
- 3. For more comprehensive in-depth analysis of Communications
 Metadata for
 - Situational Awareness
 - Collection Management
 - Target Development





NSA'S Proposal



PROPOSAL:

- (U//FOUO) Make IC communications metadata accessible to the greater IC.
 - Data will be updated and accessible instantly by IC Intelligence analysts.
 - Data will be provided for foreign intelligence and counterintelligence purposes.
 - c. Data sources and methods can be protected.
 - d. Data will be minimized* according to each agency's standards (e.g. NSA's data will be minimized to NSA minimization standards, CIA's data to CIA minimization standards, etc.).

* Minimize:

(U//FOUO) specific procedures to minimize the acquisition and retention, and prohibit the dissemination, of non-publicly available information concerning unconsenting U.S. persons consistent with the need of the U.S. to obtain, produce and disseminate foreign intelligence information.



NSA'S Proposal



- (S) Handles the ever increasing volumes of communications metadata (est. 2-5 billion records/day)
- (S) Incorporates all communication types telephony and all forms of digital, e.g. e-mail
- (U) Allows for the expansion of communications metadata fields and sources
- 5. (U) Accommodates multimedia (voice/data/video) convergence
- 6. (U) Allows for greater access by U.S. DoD military elements
- 7. (U) Requires only JWICS account plus PKI
- 8. (S//NF) Does not necessarily replace CrissCross/Proton
- (C) Allows for incorporation of foreign partner data*
 * As negotiated



NSA Proposal Implementation: Use of ICREACH





(S//REL) ICREACH is a one-stop shopping tool for consolidated communications metadata analytic needs. Through ICREACH, analysts can access most tools and all appropriate data sets related to both telephony and DNI data with a single login.



NSA Proposal Implementation: Use of ICREACH



Current Status:

(S//REL) ICREACH is an evolving NSA toolkit (middleware) focused on analyzing the target in a converged telephony-DNI environment. The federated query searches across all data sets for information relating to a target identifier.

(S//SI//REL) For example, through ICREACH an analyst could find all metadata comms related to a target: phone number, Global Mobile Satellite and cellular events and selectors, email address, etc. and any associated locational information.



ICREACH Implementation





- 1. (U) Certify Users
- 2. (U) Train Users
- 3. (U) Grant Access to Users



ICREACH Certification



- 1. (U//FOUO) Who can be certified? Any individual who:
 - a. requires access in support of Agency X mission (as validated by Agency X POC)
 - b. is a part of the U.S. IC
 - c. holds TS/SCI
 - d.has PKI
 - e.is an IC-intelligence analyst (of any sort)
- (U//FOUO) PKI-enabled accounts will be accessible via JWICS.



ICREACH Training



(U//FOUO) NSA will train a cadre of individuals from each agency who will then go back and train their own users.

(S//SI//REL) Training consists of:

- a. Intro to communications metadata to familiarize users with PSTN*, cell phones, email, IMSIs**, etc.
- b. Course to familiarize users with ICREACH tool.
- c. Intelligence Oversight training, including minimization.
- d. Limitations on use of communications metadata for operational purposes.
- * PSTN = Public Switched Telephone Network
- ** IMSI = International Mobile Subscriber Identity





ICREACH Access



 (U//FOUO) Agency X POCs to approve accounts for individuals.

- 2. (U) Accounts expire after three (3) months of inactivity.
- 3. (U) Users re-apply and are re-certified annually.



Intelligence Oversight



- (U//FOUO) Approval authorities and auditors at each agency will conduct the requisite oversight training and then conduct independent oversight.
- (U//FOUO) Auditing of agency personnel will be conducted by each agency. ICREACH audit records will be pushed by NSA to each agency in order to perform that task.
- (U//FOUO) Access will be terminated immediately upon identification of any violation and will be reinstated only upon re-certification by the agency approval authority.
- (U//FOUO) NSA will perform random auditing of IC-wide users to ensure compliance across the IC and notify Agency X of any noncompliance.



(S/NF) Call Events in PROTON*



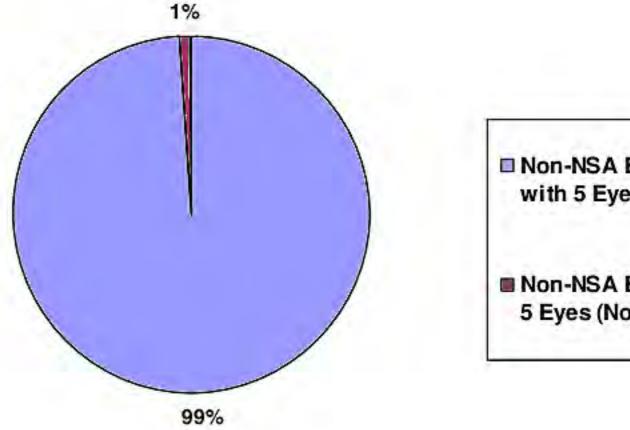
Total Call Events in NSA PROTON* est. 149 Billion

Of those:

- Total Call Events Non-NSA
- Total Call Events Non-NSA, Non-NOFORN, Non-HCS

est. 101 Billion

est. 92,000



Non-NSA Events NOT Shareable with 5 Eyes (NOFORN / HCS)

■ Non-NSA Events Shareable with 5 Eyes (Non-NOFORN / Non-HCS)

^{*} For date range 2000-2006, as of early July 2006; some data has been aged off system



Current NSA Communications Metadata Holdings



(S//SI//REL) ICReach will share over 850 billion NSA event records at IOC with an additional 1-2 billion records added daily

Telephony (1999-Apr '07)

Total Call Events

est. 711 Billion

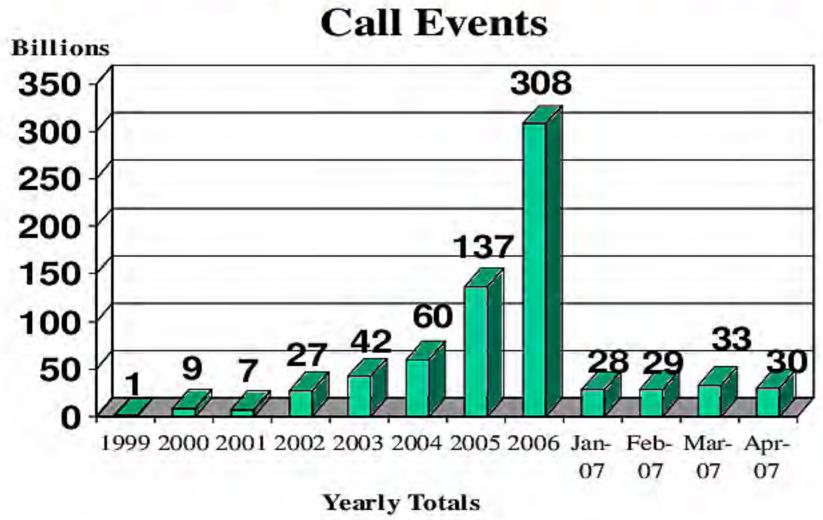
Total Call Events from est. 126 Billion 2nd Parties

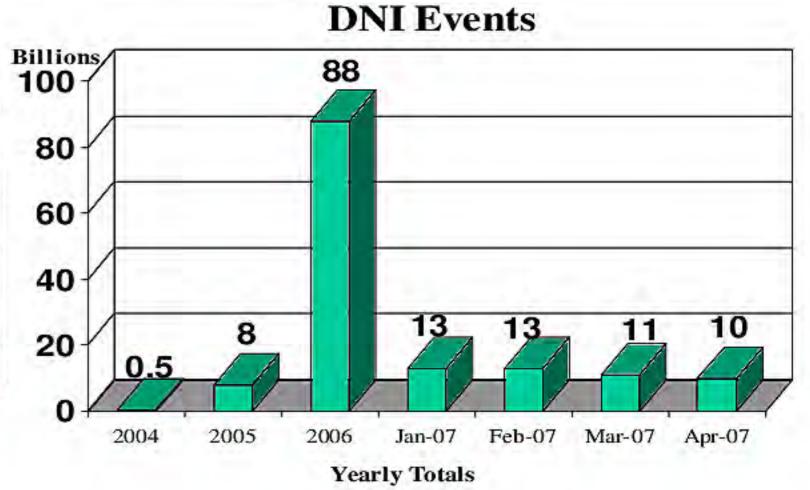
DNI (2004-Apr '07)

Total Events

est. 143 Billion

No DNI yet from 2nd Parties



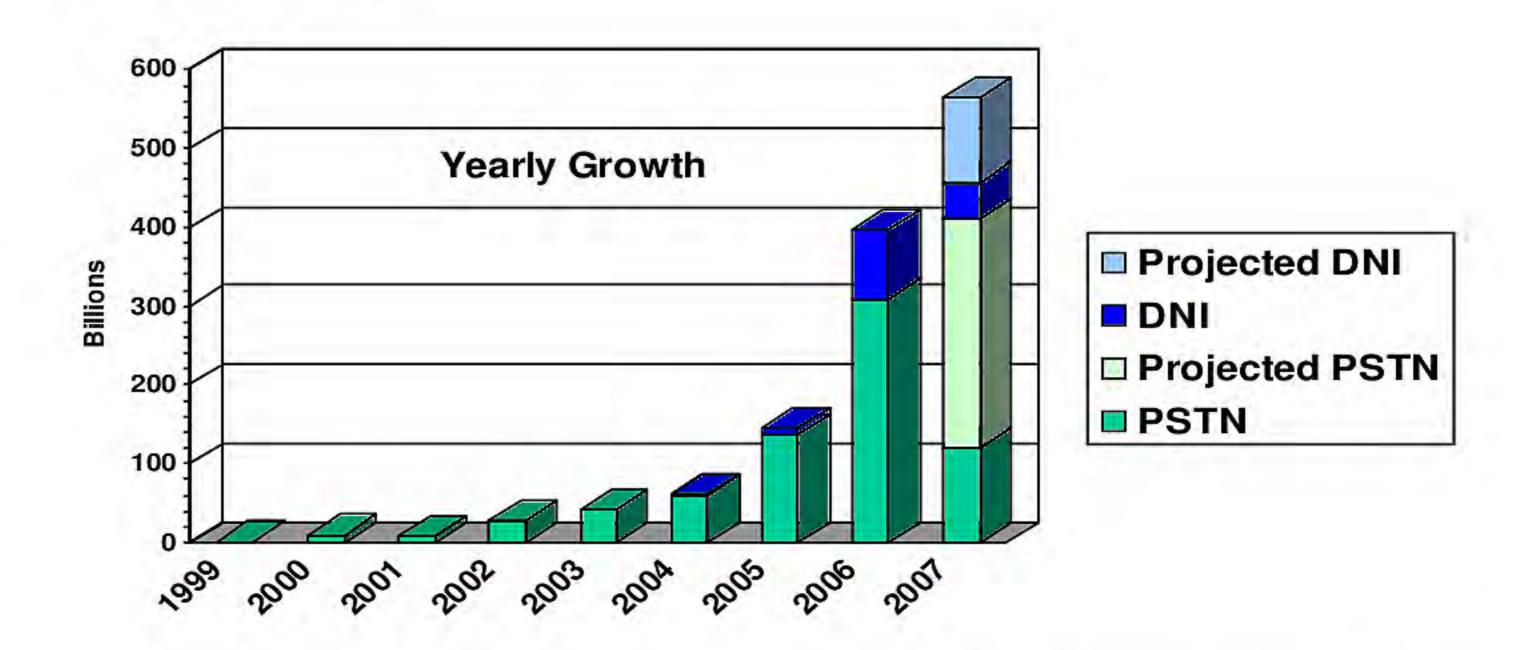




Large Scale Expansion of NSA Metadata Sharing



(S//SI//REL) Increases NSA communications metadata sharing from 50 billion records to 850+ billion records (grows by 1-2 billion records per day)



*(C//REL) Includes Call Events from 2nd Party SIGINT Partners (est. 126 Billion records)



Communications Metadata Fields in ICREACH



(S//NF) NSA populates these fields in PROTON:

Called & calling numbers, date, time & duration of call

(S//SI//REL) ICREACH users will see telephony metadata* in the following fields:

DATE & TIME

DURATION – Length of Call

CALLED NUMBER

CALLING NUMBER

CALLED FAX (CSI) – Called Subscriber ID

TRANSMITTING FAX (TSI) –
Transmitting Subscriber ID

IMSI – International Mobile Subscriber Identifier

TMSI – Temporary Mobile Subscriber Identifier IMEI – International Mobile Equipment Identifier

MSISDN – Mobile Subscriber Integrated Services Digital Network

MDN - Mobile Dialed Number

CLI – Call Line Identifier (Caller ID)

DSME – Destination Short Message Entity

OSME – Originating Short Message Entity

VLR - Visitor Location Register



Communications Metadata Fields in ICREACH



(S//SI//REL) ICREACH telephony metadata fields (con't):

MCC - Mobile Country Code

MNC - Mobile Network Code

LAC - Local Area Code

CELL ID - Serving Cell Identification

TIMING ADV – Timing Advance Value (distance from base transceiver)

LAT/LONG – Latitude/Longitude in degrees and minutes

CALLING FTIN – Forward Terminal Identification Number CALLING RTIN – Reverse Terminal Identification Number

DIALED NUMBER

FWD SIM – Forward Subscriber Identity Module

REV SIM – Reverse Subscriber Identity Module

MES TYPE – Message Type

(S//SI//REL) ICREACH users will also see this DNI metadata:

Email addresses

Chat handles

Date & Time

Protocols

SECRET//COMINT//REL TO USA, FVEY//20320108



Increases Number of SIGINT Metadata Modes and Fields Shared



Currently Shared

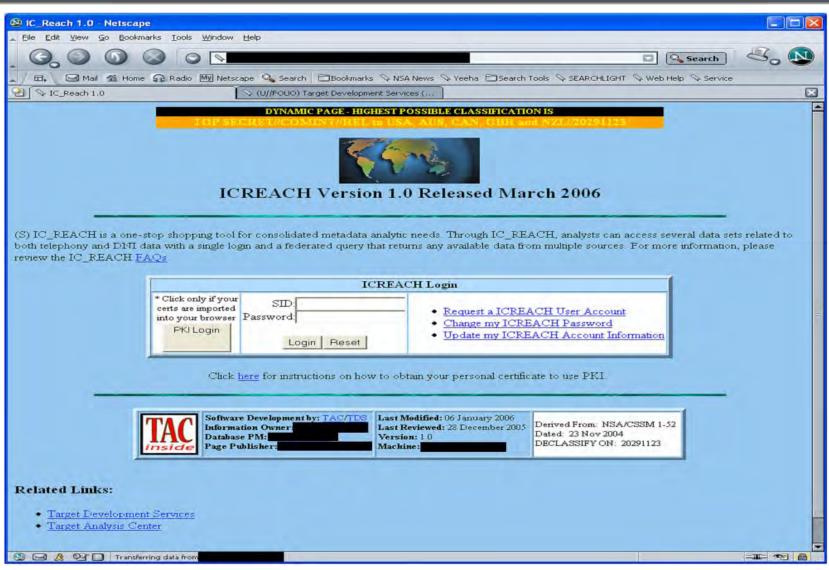
ICReach Expansion

Metadata Field	PSTN	INMARSAT	PCS	DNI
Date	Х	X	X	Х
Time	X	X	X	X
Duration	X			
Called Number	X			
Calling Number	X			
Called Fax number	X			
Transmitting Fax number	X			
IMSI			X	
TMSI			X	
IMEI			X	
MSISDN			X	
MDN			X	
CLI			X	
DSME			X	
OSME			X	
VLR			X	
MCC			X	
MNC			X	
LAC			X	
Cell ID			x	
Timing Advance			X	
Lat/Long		X	^	
Calling FTIN		X		
Calling RTIN		X		
Dialed Number		X		
Forward SIM		X	X	
Reverse SIM		X	Α.	v
Email Address				X
Chat Handle Protocols				X X



ICREACH Log-in







ICREACH Query

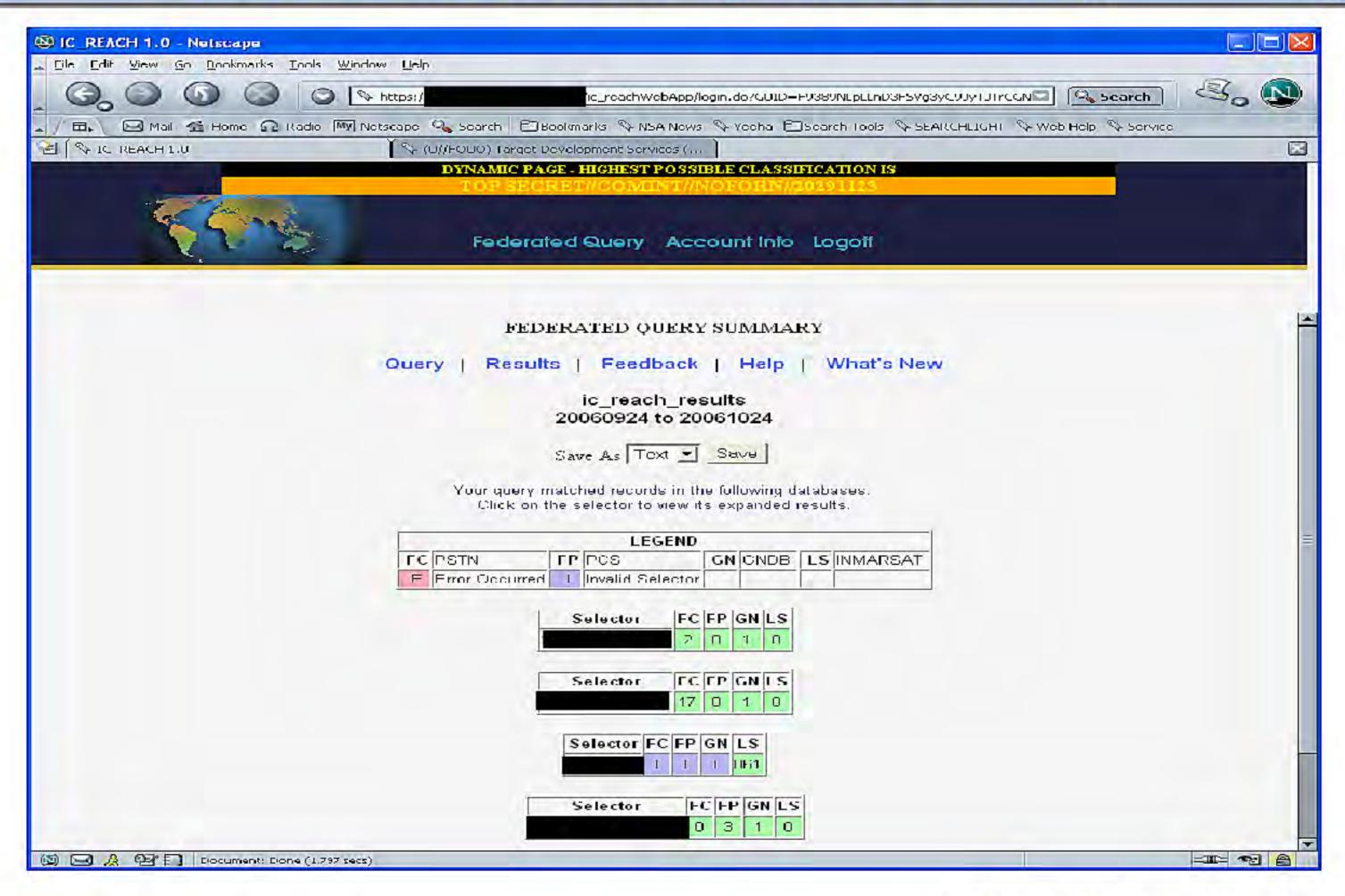


(A) IC_REACH 1_D - Netscape	
Eile Eili View Gu Bunkmarks Innis Window Help	
ic_reachWebApp/login.do?GUID=F9009NLpLLnD0F5Vg0yC9JyTJTrCGNI Search	€. 🖎
🚅 / 🖽, 🖂 Mail 🛳 Home 🔐 Radio 🏧 Nelsrape 🦠 Search 🖾 Bookmarks 🛇 NSA News 🛇 Yeeba 🖾 Search Took 🛇 SEARCHLIGHT 🦠 Web Help 🛇 Service	
Carpet Development Services (
DYNAMIC PAGE HIGHEST POSSIBLE CLASSIFICATION IS TO ESECRET//COMPNE//HOFORN/20201112 Federated Query Account Info Logoff	
FEDERATED QUERY Query Results Feedback Help What's New ** Wildcard searches are not supported ** ** Searches are limited to 150 or fewer selectors **	
Date Range: 20060924 to 20061024 Selector: (comma delimited) File of Selectors: Browse (one per line) Query Name: Submit	
DYNAMIC PAGE - HIGHEST POSSIBLE CLASSIFICATION IS TOP SECRET/CUT/UTT/PUP 1/20291124	
(S) 🖂 🙎 🖭 I ransferring data from	



ICREACH Query Summary

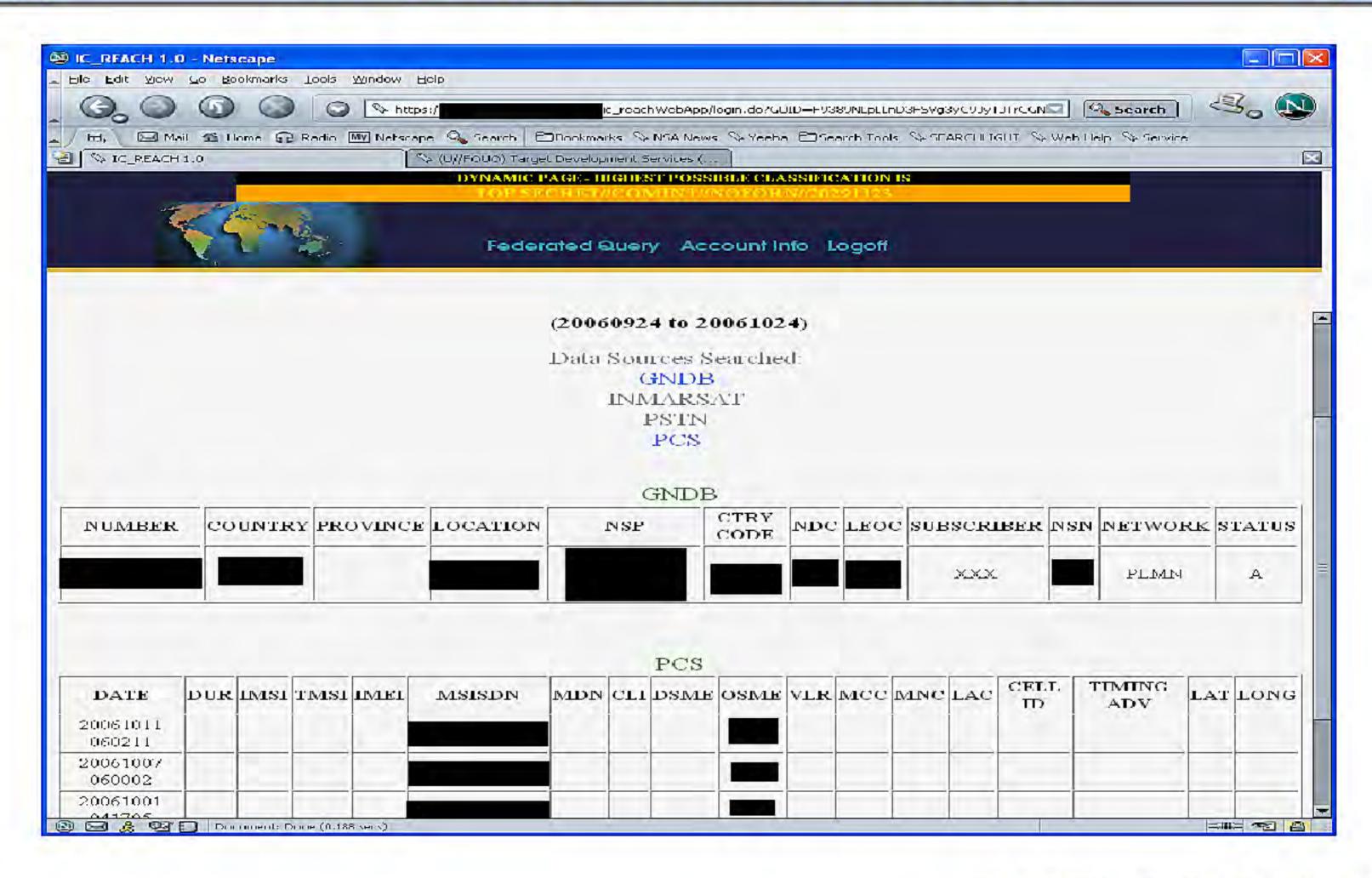






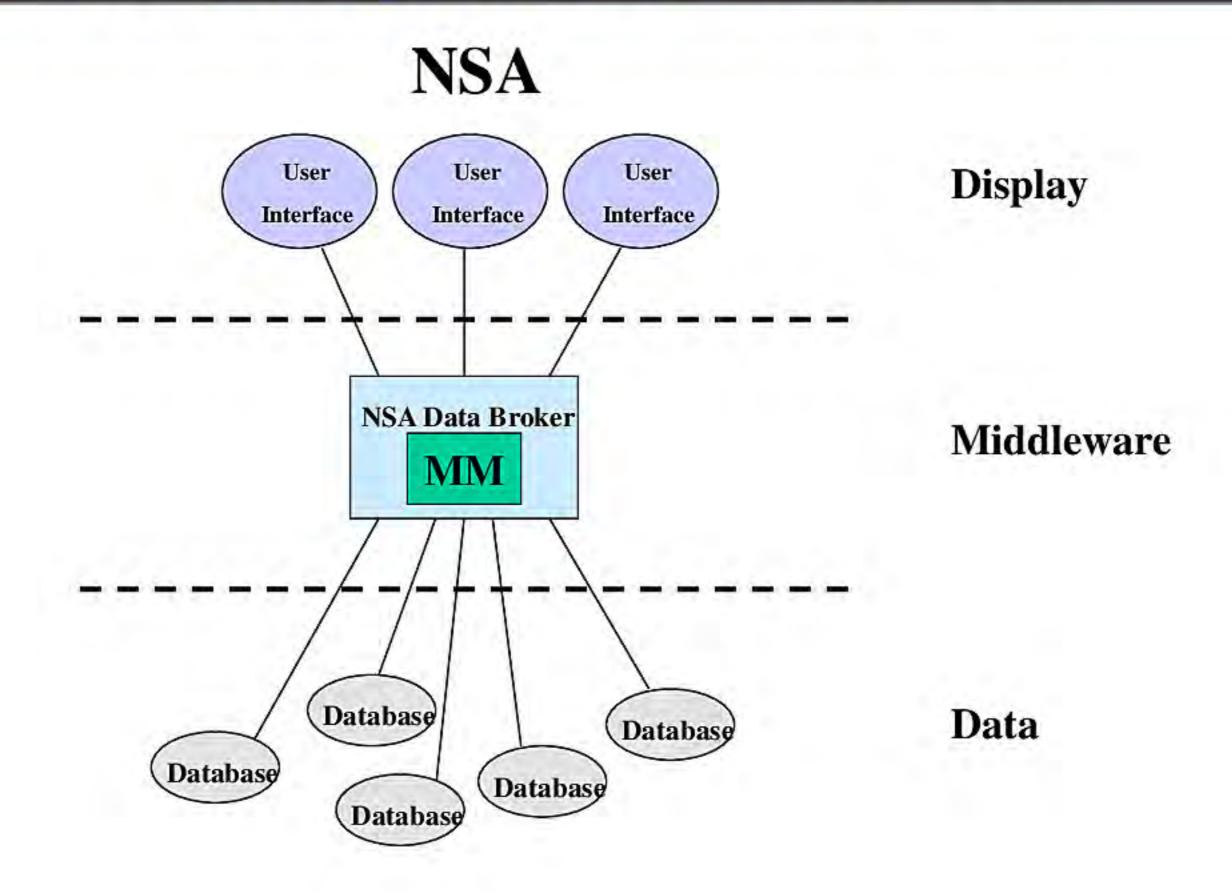
ICREACH Data Results





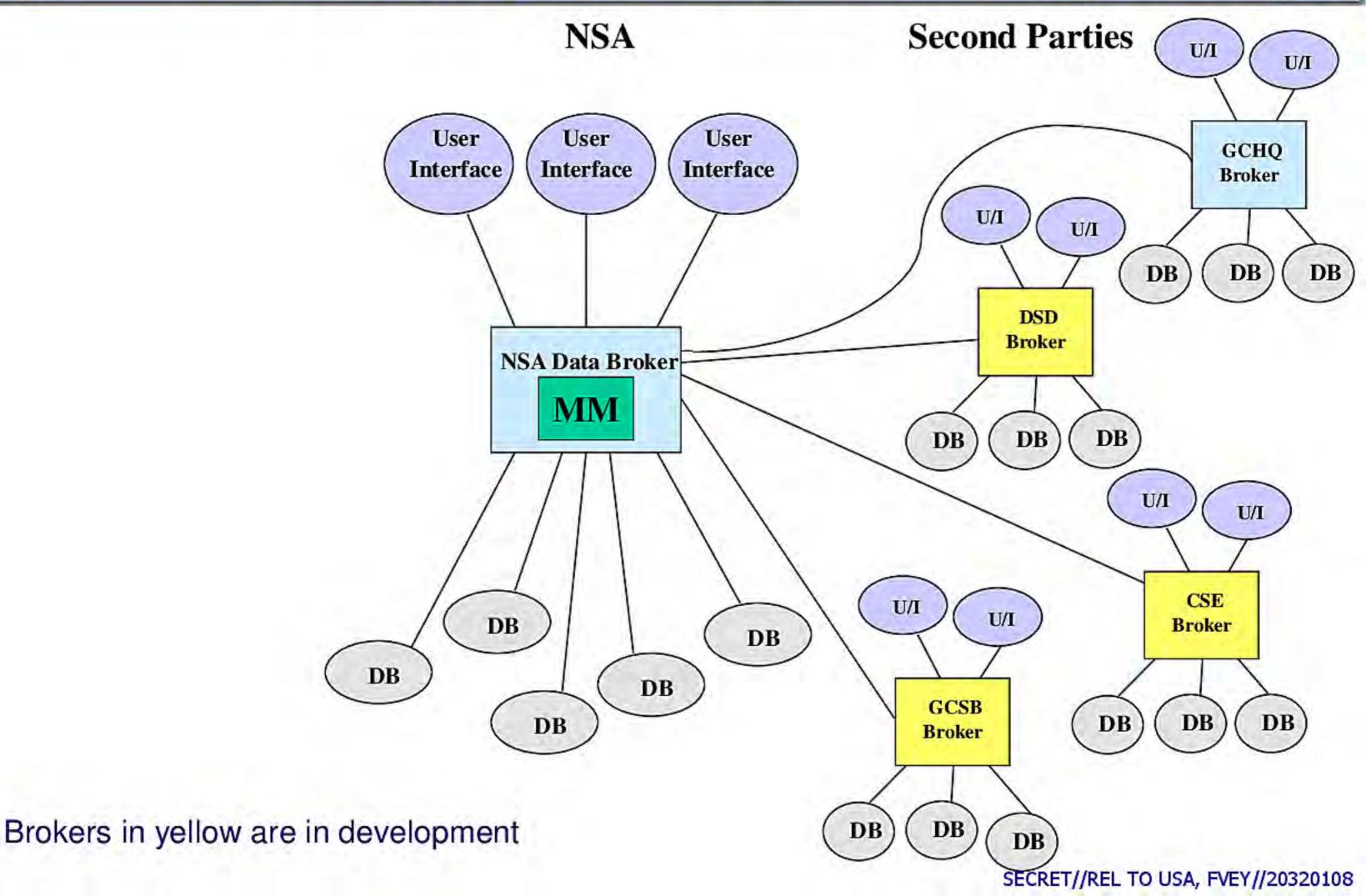






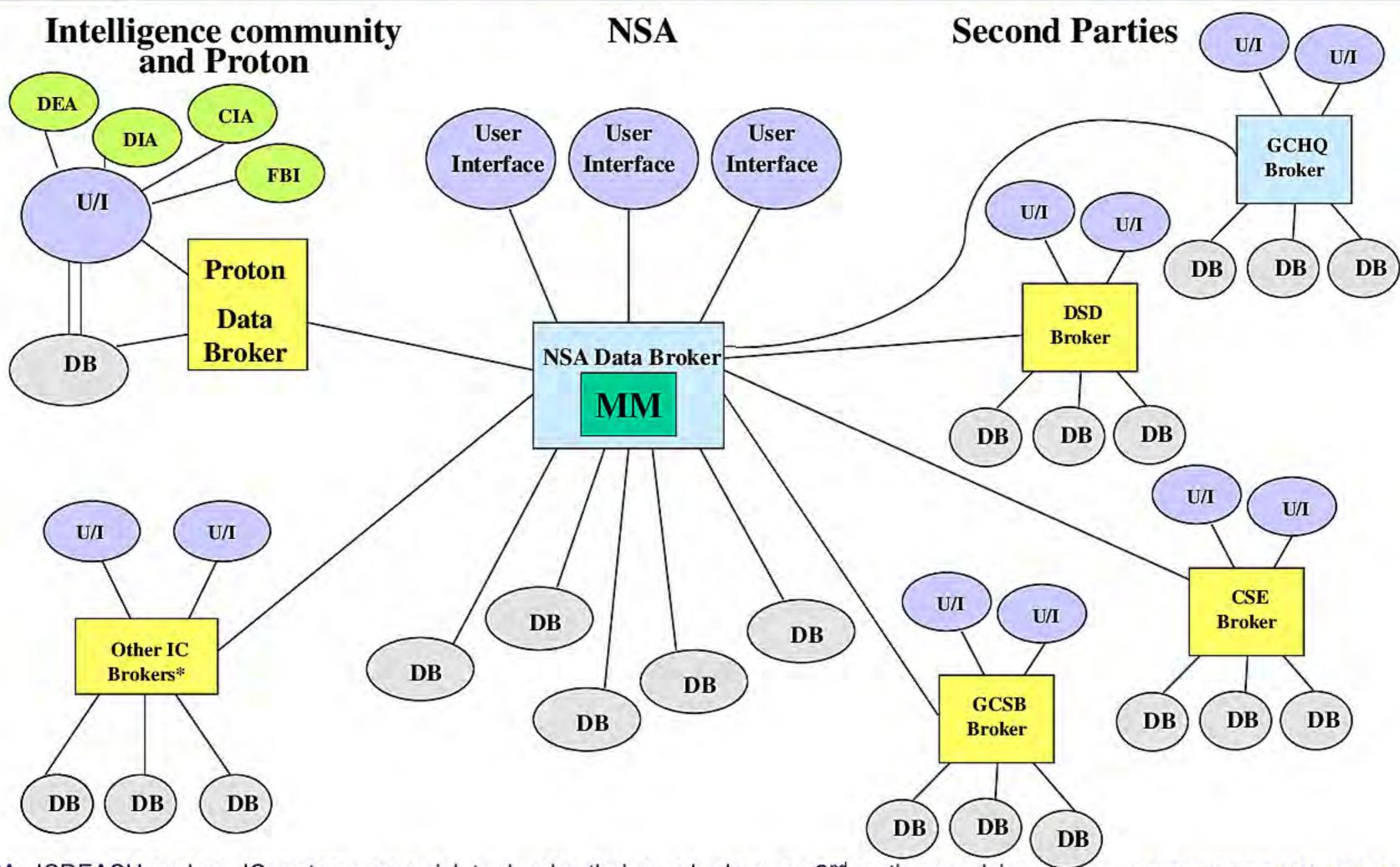










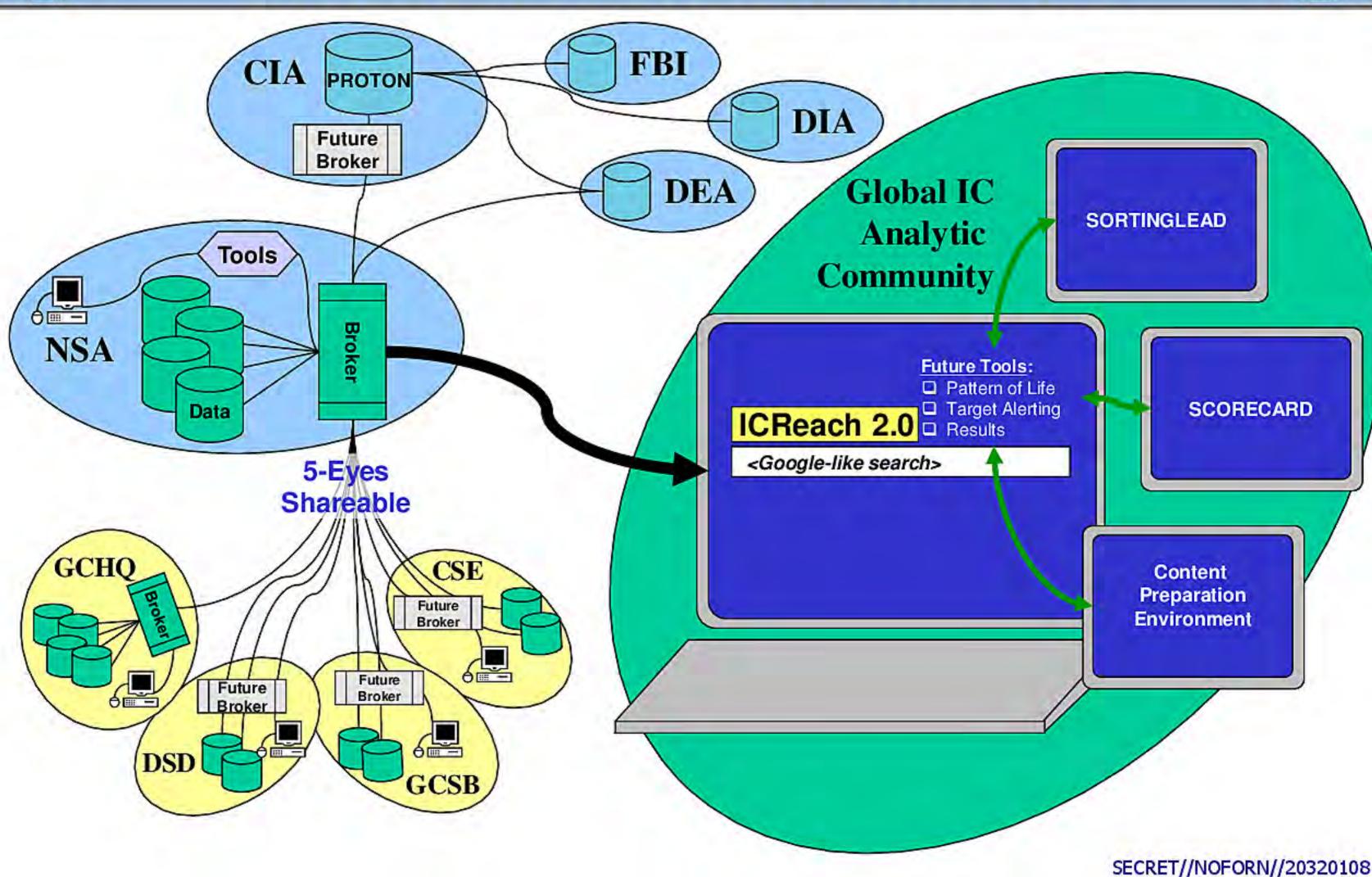


*As ICREACH evolves, IC partners may wish to develop their own brokers, as 2nd parties are doing

SECRET//NOFORN//20320108









Recap



- (C//REL) ICREACH is a large scale expansion of communications metadata shared with IC partners
- (C//REL) Enables federation of community metadata across IC agencies
- (C//REL) Makes greater use of NSA's communications metadata which is information about content (but not the content itself)
 - Definition: Information used by networks, services, or applications to route communications or make the content usable including:
 - Dialing, routing, addressing, location, or signaling data used for network management
 - Formats and protocols used to render the information for people and systems
- (S//SI//REL) Expands NSA sharing in three dimensions:
 - Includes more modes of communications (PSTN, INMARSAT, PCS, DNI)
 - 5x increase in types of metadata shared (from 5 fields to 33 fields)
 - 12x increase in volume shared (from 50 billion to 800+ billion records)



Next Steps



- (U) Pilot development underway IOC June
- (U) CCP Funded for first year
- (U//FOUO) Agreements on 2nd Party sharing (bi-directional)
- (U) Expansion to open source?
- (U//FOUO) Limitations and responsibilities related to operational use
- (U) Software licensing



Preliminary Thinking on ICREACH Roll-out



(S//NF) From date of sponsorship:

- Train the trainers
- Account Set-up
- Scrub PROTON NOFORN data for release to ICREACH & Second Parties
- Establish ICREACH Configuration Control Board (CCB) across IC

(U//FOUO) 2008:

- Reach out to other databases
- ICREACH Enhancements



ICREACH Timeline



(U//FOUO) 2005

- Fall: initiative conceptualized
- Dec: informally proposed to CIA DIR who suggests staffs meet to discuss it and consider other options before submitting to DNI

(U//FOUO) 2006

- Spring-Summer: discussions with policy reps & technical experts from several U.S. IC agencies; broached with Second Parties
- Summer–Fall: briefed Program Manager/Info Sharing Environment, ODNI/CIO/Info Sharing & Customer Outreach, & ODNI/CIO/IC Enterprise Architecture
- Oct: formally proposed to DNI
- Late Nov: interim response from DNI, recommending we continue working initiative with ODNI/CIO & DDNI/Collection

(U) 2007

- May: Pilot development begins
- June: Pilot IOC





Additional slides



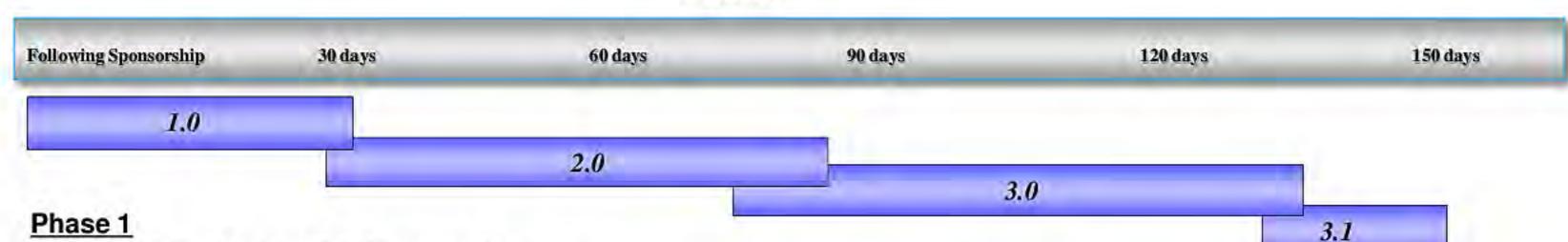


ICReach Deployment 1 Schedule



(S//REL) ICREACH requires between \$2.7M and \$3.9M in FY07 to implement the initial phase, depending on the pace of the development effort.

FY2007



- Establish IT and training POC's across the IC
- Add DNI and other NSA data sources (e.g., PCS, GNDB) to Federated Query Service
- PKI Validation and Acct Establishment Process
- Design User and Metadata Use Training Course

Phase 3

- Transition and Integrate ICR to Operational Hardware
- Initiate User training sessions (Final schedule & staffing requirement: TBD)

Phase 4 and out years

- Develop interface with IC data sources and brokers
- Enhancements to improve analytic tool integration

Phase 2

- Code and Test User Interface (UI) and Web Service
- Test Account Authentication and Security for User Interface
- Conduct "Train the Trainer" for UI and Account Management process

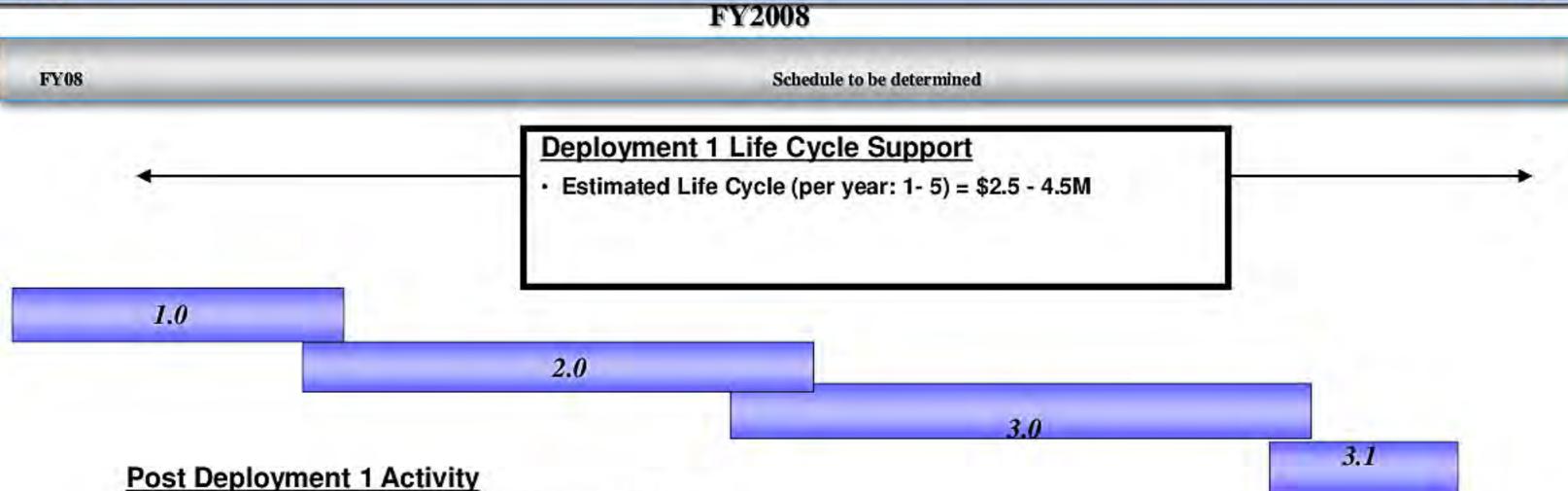
Phase 3.1

 Optimize hardware performance



ICReach Post Deployment 1 Schedule





- Develop interface with IC data sources and brokers
 - · Second Party Brokers (as made available)
 - · PROTON Data Broker
 - Other IC Agency Data Brokers
- Enhancements to improve analytic tool integration

Resources 1 FTE USG PM .25 FTE Account Access 1.0 FTE GUI .5 FTE Web Service .5 FTE Config Mgt and Test 1.0 Training Development .5 Database Engineer/Admin

(C//REL) ICREACH: Sharing SIGINT Comms Metadata with the Intel Community

FROM:

Tech Director, Target Analysis Center (SSG1)

Run Date: 09/27/2007

(C//REL) Remember when you were little and your mom always told you it was nice to share? I don't know about you, but sometimes in the intervening years it seems we forget about that little axiom. Recently those of us who are members of the Intelligence Community (IC) have been advised that we have to do a lot more sharing. Congress has told us, the 9/11 and WMD¹ Commissions have reminded us, and the Department of Defense strategy for improving operations insists on information sharing. While NSA has been sharing information for years through tailored product reporting, write-to-release efforts, ELINT² metadata and tearline reporting, we are now going a step beyond that and are actually sharing SIGINT communications metadata through a program called ICREACH.

(S//REL) ICREACH intends to make more metadata accessible to our Intelligence Community brethren primarily for the purposes of situational awareness and target development. This COMINT³ metadata is not to be used for targeting; instead provisioning of the metadata to IC intelligence analysts should foster greater collaboration on targets of interest. Like NSA's internal tool, GLOBALREACH, ICREACH is a one-stop shopping tool for consolidated communications metadata analytic needs.

(S//REL) Through ICREACH, IC analysts can access all appropriate data sets related to telephony (and eventually DNI⁴ data) with a single login. The IC intelligence analysts will be validated by the ICREACH point of contact from their particular agency, must be part of the U.S. Intelligence Community, must hold a TS//SI clearance and must have a PKI certificate in order to access the data through JWICS⁵. Each person will be provided an account only after their need for it is verified by the Agency POC and they have received training on the tool, the procedures, and the technologies represented in the metadata.

(S//SI//REL) For the first time, IC intelligence analysts will have immediate access to SIGINT 1st Party metadata as soon as the data is loaded in the centralized metadata repository. The <u>FASCIA</u> Call, PCS⁶, and <u>INMARSAT</u> tables will be the first ones accessed in the ICREACH Beta release, now being piloted by a very limited number of IC analysts⁷. Reference data from other NSA projects like <u>ASSOCIATION</u>, <u>TAPERLAY</u>, and <u>CONTRAOCTAVE</u> is also queried. If the number is tasked in OCTAVE, a simple "yes" is returned with the event data.

(S//SI//REL) Responses thus far from the beta testers have been largely positive with some very interesting questions, comments, observations and suggestions being provided, considered, and where possible, implemented. Future plans call for a rollout to additional IC agencies and Unified Commands (DIA and STRATCOM⁸ are next on the list for training) and access to other sources of metadata. NSA and CIA are developing an interface to enable access to their metadata through ICREACH/GLOBALREACH so that there is a reciprocal exchange of information.

(U//FOUO) If you have questions on ICREACH, you may direct them to the following individuals:

For questions about the rollout of ICREACH to the next sets of users:

- For questions about the data available within ICREACH:
- For questions about the technical workings of ICREACH: or all from SSG13.

(U) Notes:

- 1. (U) WMD = Weapons of Mass Destruction
- 2. (U) ELINT = Electronic Intelligence
- 3. (U) COMINT = Communications Intelligence
- 4. (U) DNI = Digital Network Intelligence
- 5. (U//FOUO) JWICS = Joint Worldwide Intelligence Communications System, operated by the Defense Intelligence Agency (DIA) and serving the DoD and IC.
- 6. (U) PCS = Personal Communications Services
- (C//REL) The IC analysts participating in the pilot are from the Special Operations Command (SOCOM), CIA, Central Command (CENTCOM), and the National Counterterrorism Center (NCTC).
- 8. (U) STRATCOM = Strategic Command

- 7. NSA: for the foreseeable future, NSA has two separate definitions of what constitutes content based on EO 12333 and FISA/PAA collection. FISA considers communications data to be part of the content of the communication, hence FISA has no separate concept of metadata. All discussions at the conference were therefore focused on "standard" collection under the authority of EO 12333.
- 8. There is a constitutional expectation of privacy within the US. For communications data this is harder to quantify than for content. New procedures will permit a differentiation between content and communications data allowing for far greater data usage and advancing other related changes. A tension remains between the desires to minimise shared data containing US identifiers, and engaging more openly to support the foreign cryptologic mission.
- 9. It is harder to define what constitutes a US identifier with DNI data where unclear it is treated as US. NSA is moving from minimised records within their databases to minimising identifiers within reports. Sharing unmasked US identifiers with second party SIGINT partners will be easier than with some US domestic partners.
- All: All SIGINT agencies seek to protect their equities, especially relating to Special Source Exploitation (SSE.)
- 11. Special categories of data were considered in the context of their potential to contribute to pattern of life analysis. An increasing amount of new data types are available to SIGINT agencies, some proving difficult to categorise as either content or communications data. The conference agreed to step back from trying to categorise the data and simply to focus on what is shareable in bulk.
- 12. Consideration was given as to whether any types of data were prohibited, for example medical, legal, religious or restricted business information, which may be regarded as an intrusion of privacy. Given the nascent state of many of these data types then no, or limited, precedents have been set with respect to proportionality or propriety, or whether different legal considerations applies to the "ownership" of this data compared with the communications data that we were more accustomed to handle. It was agreed that the conference should not seek to set any automatic limitations, but any such difficult cases would have to be considered by "owning" agency on a case-by-case basis.

(comment: NSA normally considers any target data (pattern of life or other) that can be characterized as "foreign intelligence" as proper for collection, analysis and production.

Comment Page: 3
used the term "off limits".

35. (d) ICREACH: Setill a pilot, this provides minimised DNR data to Sigint-cleared and appropriately trained analysts across the US Intelligence community. Second Party derived data is currently not made available to US Intelligence Community (IC)(domestic)domestic-agencies (although GCSB has agreed that theier DNH metadata may be shared), but such data would be valued. In the hope that such agreement will be forthcoming, NSA has persuaded other US IC agencies to make almost 100 bn previously NOFORN records shareable with the 5-eyes via GLOBAL REACH. VolP is treated as DNR though with only DNR records and fields shown to analysts. Making DNI available through ICREACH is currently restricted due to limited automated (general counsel approved) methods to prevented by US policy on minimize DNI metadata.ation.

- 36. 'Deconfliction' is not formalised through ICREACH. Query records can potentially be used to alert analysts that other analysts are looking at the same data. Deconflicting operations remains a tough challenge requiring increased coordination of operations and collaboration.
- 37. There is interest in the relationship between the implementation of A-SPACE (a U.S. DNI initiative to link all U.S. IC analysts to common tools, and sharable databases and allow for greater collaboration) and ICREACH inasmuch as it may affect Second Parties' internal procedures and access issues with domestic agencies.
- 38.(e) <u>U.S.</u> NSCID5: This is a specific method of <u>NSA providingsharing</u> uniminimized <u>SIGINT</u> data to CIA (as if CIA had collected it itself) in support of the latter's operational mission., and dData from Second Parties is shared with CIA in accordance with special agreements between NSA and each second party.
- 39. GCHQ are employing methods (a) and (b). For military SIGINT needs GCHQ uses GCO's to reach back to UK and 5-EYES repositories. The military's work within the ambit of the National SIGINT Organisation comes under the authority of GCHQ.
- 40.NSA shares US SIGINT data with all US SIGINT elements that operate under DIRNSA's operational control agencies. With Second Parties there is an initial minimisation of the data when possible; however all second parties have agreed to abide by U.S. minimization criteria., Ffor US intelligence agencies NSA mustthere is further minimizesing of the data, before sharing and for other US agencies (such as law enforcement) NSA only provides data under its "technical support" mission. Currently, all such data is minimized before sharingthere is another, further level of minimising of data (most restricted data set.)

Comment Page: 6
The 'US' was implicit in this statement but no voiced at the time.

UNITED STATES GOVERNMENT

memorandum

TAC-0xx-05

DATE: 06 July 2005

REPLY TO ATTN OF: TD/TAC

SUBJECT: Sharing of MetaData across the IC ---- Dissemination of SIGINT Metadata Beyond NSA

(FOUO)- MEMORANDUM OF UNDERSTANDING

TO: DIR								
THRU: I	D/DIR	,IG	_,OGC_	,DK	,ITD	_, SE	_, NCR CIA_	
D/SID	.S02	.S1	S2	.S3				

- REFERENCES: A. (C) Request for additional NSA data for CIA's PROTON Program. (dtd: 10 Mar 2005)
 - B. (C) Minimization Plan for the Application of CRISSCROSS Analytical Software to NSA SIGINT Acquired Telephone Call Control Data, dated 20 Jan 1999.
 - C. (C) MOU between NSA and FBI for Access to NSA Data in the CRISSCROSS Program, dated 19 Jan 2001.
 - D. (C) MOU between NSA and DIA Regarding the CRISSCROSS Program, dated 19 Jan 2001.
 - E. (U) NSA/CSS Policy 1-9, Information Sharing , dated 26 May 2005.

PURPOSE: (U) To establish the DIRNSA/CHCSS and NSA/CSS as the U.S. IC Executive Agent for IC-wide metadata sharing.

BACKGROUND:

(S//NF) Except for a few point to point sharing initiatives, CRISSCROSS/PROTON is the current IC (plus) information sharing structures. CRISSCROSS/PROTON is a CIAmanaged program which provides extracts from selected Agency (NSA, CIA, DIA, FBI and DEA) databases of telephone call records and reference data obtained from

Derived From: NSA/CSSM 1-52

Dated: 20041123

Declassify On: 20300707

HUMINT, SIGINT, Open and law enforcement related sources to analysts in U.S. law enforcement and intelligence communities at a CONFIDENTIAL NOFORN level. The key data elements are stored and retrieved for display by the PROTON Interface include: called and calling telephone numbers, date, time, and duration of calls.

(S//NF) For approximately the last 5 ½ years, the CIA CRISSCROSS (now PROTON) program has been incorporating NSA provided SIGINT-derived (U.S. and Second/Third Party) data. NSA provides this information to CRISSCROSS/PROTON Agencies for

(S//NF) The decision to disseminate SIGINT-derived signaling outside COMINT Channels also had *policy* and *operational* dimensions. Tests have demonstrated that SIGINT- derived and traditional HUMINT-derived call record data can be stored and displayed in a manner which makes them appear to be indistinguishable, which in general terms make the risks to sources and methods approximately equal. However, under some circumstances, e.g., signaling data form a unique target area, which could have come only from a sensitive SIGINT source, dissemination of data to CRISSCROSS would be inappropriate. This is the case when the data also includes information elements that are only available and could only come from exploitation of signaling. The MOA is silent on this point, leaving to NSA full decision authority to decide which data to provide and not provide.

(C) For several years prior, NSA had been providing SIGINT-derived signaling information to CIA and DEA to support multi-agency counter narcotics analysis, as approved by DDO, OGC, and the Office of Policy. Calling data from several Latin American collectors, and from Thailand, had been of great value to the DEA and to the DCI Crime and Narcotics Center, and there were no adverse operational or legal impacts to NSA.

(S//NF) Operationally, the flow of large volumes of intercepted signaling data could place additional burden on NSA's IT infrastructure. While, as above, the MOA did not address IT infrastructure, NSA retained full discretion over volume and timing of information flow, even as NSA chose to support the CRISSCROSS/PROTON program with certain SIGINT inputs.

(S//NF) Requisite human and communications resources are in place in SID to support the transfer of moderate volumes of signaling-derived call record data to the CIA CRISSCROSS/PROTON Program Office. No increase in personnel was required to initiate data flow. Under the terms of MOA, CIA provided the list of US overseas commercial phone numbers to be minimized in NSA processing of signaling data. Similarly, the Department of State and Department of Defense telephone directories will be researched by NSA to identify official US government overseas phone numbers in order to minimize call records associated with them. To this end, the NSA CRISSCROSS/PROTON Program Manager agreed to obtain the DOS and DOD directories annually and, with CIA assistance, extract the overseas numbers. That function was transferred to the Communications Event Analysis Center (CEAC), in S2S, in 2004. The software to extract call record data and to support minimization exists on a

Handle Via COMINT Channels Only SECRET//NOFORN//MR

server under S2S control. A huge volume of signaling data (~400M records) flows to that server on a daily basis to meet NSA analytic needs, and sufficient communication bandwidth is currently available to support the transfer of processed call records (~37M per day) to CIA's CRISSCROSS/PROTON Program for reformatting. We anticipate S21 will continue to be responsible for loading CIA CRISSCROSS/PROTON output on NSA's CRISSCROSS/PROTON server, and CEAC will continue to be responsible for maintaining minimization and audit controls on all-source data received from CIA. Finally, SID Office of Oversight and Compliance would be responsible for oversight of USSID 18 compliance.

(S) Community analytic interest in dialing analysis continues to grow and is directed at discovery of new specific targets or topics, and demand for additional call record volume to support contact chaining and geolocation has increased. Signaling links are competing for collection priority just like any other potential SIGINT target. As with other targets, the potential to expand collection is limited by available resources, to include field processing of signaling links as well as bandwidth to support forwarding from collection sites to NSA.

(S//NF) The CRISSCROSS program, in which NSA already participates as a user of data, has had notable successes since its inception in 1990, such as enabling major narcotics arrests, monitoring Ramsi Yousef's colleagues, unraveling the Mubarak assassination plot, and providing insight into Pakistani nuclear weapon test activities. The program has received high marks from senior Community levels, such as the PFIAB, and Assistant DCI's for both Collection and Production. CRISSCROSS was also cited in the report by Admiral Jeremiah on Indian nuclear testing as a potential improvement to Community analytic strength. Since 9/11, the contributions to the GWOT due to our increased collection of signaling metadata are innumerable and significant. It is safe to say that it has been a contribution to virtually every successful rendition of suspects and often, the deciding factor. Hence the benefit to the intelligence and law enforcement communities of any SIGINT-augmented inputs could be considerable, as SIGINT has the potential to access a broad range of targets.

CURRENT PROPOSAL: (S//NF) Ref A. requests additional data elements from SIGINT collection be included in the data set transferred to PROTON. Specifically, they want mobility management information relating to digital cellular and mobile satellite systems to include Global Cell ID's, Location Area Codes, spot beams, International Mobile Subscriber Identifications, International Mobile Equipment Identifications, Latitude/Longitude, and Inmarsat Return ID's. They also requested content from Short Message Service exchanges (which we are required to audit queries against).

DISCUSSION: (S//NF) The current data push to CIA/PROTON is roughly 40 Million records per day which has resulted in approximately 30% of the PROTON data set (1996-2005) coming from SIGINT sources (2002-2005). The addition of the requested data elements would more than double the volume of data sent to PROTON on a daily basis. This would increase the percentage of SIGINT contribution to PROTON significantly. CIA is positing PROTON as the community resource for target/lead

Handle Via COMINT Channels Only SECRET//NOFORN//MR

development based on contact chaining techniques. In this regard it is extremely difficult to envision it playing that role at NSA. PROTON requires a NOFORN/ORCON access, must be licensed, has no API that would allow access integration with other tools (E.G. Federated query) and, represents a significant funding investment and dependency on a CIA contractor.

RECOMMENDED RESPONSE: (S//NF) In the context of the atmosphere created by the events of 09/11/01 and the following investigations into the perceived 'intelligence failure' a recurring theme has been the lack of data sharing on the part of the IC members. Against that background, it is difficult to deny requested access to SIGINT-only datasets that are thought to have value to other intelligence analysts/missions. In an effort to:

- a.) provide access to the requested data and,
- b.) break new ground in the information sharing arena while,
- c.) not moving any more data out of NSA and into duplicative storage.
 We propose to utilize the IC shared information space ICSIS on INTELINK and

we propose to utilize the IC shared information space ICSIS on INTELINK and implement IC access to our GLOBALREACH federated query service via accounts and access verified by PKI certificates. This service will provide the access requested and permit the auditing legally required of NSA. Further, we request that CIA forward PROTON data from non-SIGINT sources to NSA for inclusion in the dataset searched by GLOBALREACH. For data that must remain in HCS channels, we will create an HCS partition in FASCIA II. This will permit one-stop access to contact information for IC analysts. We believe that we can have GLOBALREACH available in the ICSIS shared space by 01 Oct 2005.

RECOMMENDATION: (FOUO) That you sign the enclosed note to the Office
of the Director, National Intelligence. Questions of a legal nature should be directed to
AGC (Operations), Operational question should be directed
to SIGDEV/TAC Technical Director,
SIGDEV/TAC TD

MEMORANDUM FOR THE DIRECTOR OF NATIONAL INTELLIGENCE

SUBJECT: (U//FOUO) Sharing Communications Metadata Across the Intelligence Community –DECISION MEMORANDUM

PURPOSE: (U//FOUO) To request the Director National Intelligence establish a U.S. Intelligence Community-wide communications metadata sharing structure; to recommend that the National Security Agency establish this structure along with appropriate query tools, training and oversight; and, that Director, NSA be designated the IC Executive Agent for IC-wide communications metadata.

BACKGROUND: (S//NF) Post 911, the IC has increased its needs for communications metadata¹. Currently, NSA collects, processes and stores vast amounts of communications metadata related to worldwide intelligence targets. This SIGINT information often includes the originator and recipient(s) identifiers, as well as the date, time and duration of the transaction and, in some cases, locations of targets.

(S//NF) CIA, FBI, DIA and DEA also have access to communications transaction data from their unique sources. These organizations, to include NSA, have benefited from sharing this data through the CRISSCROSS/PROTON program since 1999. Presently, NSA pushes roughly 60 million records per day (limited to telephony only) to PROTON, resulting in over 30% of the total PROTON data set.

(S//NF) While the PROTON program has had many notable successes over the years, today's environment requires that the IC implement more holistic information sharing and collaboration. In order to ensure the efficient and effective sharing of many millions of new minimized communications metadata records daily, the IC must have a new sharing/collaboration structure that is scaleable to the larger IC and is capable of much greater volume of data.

(S//NF) NSA stands alone in conducting metadata analysis on a large scale and as a result, our tools and methodologies, as well as the infrastructure employed, are well tested and have been proven efficient and effective over time. We already extensively

1 *(C) Communications metadata refers to structured "data about data": it includes all information associated with, but not including content, and includes any data used by a network, service or application to facilitate routing or handling of a communication or to render content in the intended format; it includes, but is not limited to, dialing, routing, addressing, or signaling information and data in support of various network management activities (e.g. billing, authentication or tracking of communicants).

Derived From: NSA/CSSM 1-52

Dated: 20041123

Declassify On: 20290325

share metadata internally within the NSA/CSS extended enterprise and with our second party partners.

(S//NF) Under this proposal, IC partners would have access through NSA's GLOBALREACH (the IC version will be called ICREACH) service to more than 50 existing NSA/CSS metadata fields contained in trillions of records as well as other IC Agency communications metadata (predominantly resident today in PROTON). NSA's throughput to ICREACH would dwarf the volume of NSA's present contributions to PROTON, as well as the input of all other IC contributors. Those fields revealing sources and methods or sensitive foreign relationships within metadata records would not be shared. Moreover, NSA is preparing to include DNI (digital network intelligence) data in GLOBALREACH/ICREACH, a change that would also exponentially increase NSA's contributions to the IC's communications metadata structure.

DISCUSSION: (S//NF) NSA proposes that the DNI establish a new IC-wide communications metadata structure with requisite tools, training and intelligence oversight. This structure will consist of data repositories at various IC Agencies all interlinked through a web based analytic query structure via JWICS whereby IC personnel can make one query to access all sharable communications metadata. This structure will also allow IC personnel to choose the data they wish to query and the set of tools they desire to display and analyze the data.

(S//NF) This structure will allow unprecedented volumes of communications metadata to be shared and analyzed. This structure will use the Intelligence Community Shared Information Space (ICSIS) carved out in Intelink and protected by community PKI access authorization. Having such a structure will provide the IC with access to NSA/CSS and other IC-element metadata, as well as the internal 'federated query' that NSA analysts use to correlate these information elements and enrich the results with knowledge from our analytic reference files and other collaboration. This proposal will address existing IC-wide stated needs for additional NSA communications metadata elements and take the IC to an unprecedented level of information sharing.

(S//NF) This initiative is in keeping with the 9/11 and President's WMD Commissions' recommendations on information sharing, the President's various Executive Orders related to information sharing, and your own announced goals on information sharing and collaboration.

(S) This initiative was intended to facilitate a new level of sharing information across the IC for situational awareness, lead information, adversary analysis, and targeting and will add to the process of greater information sharing across the IC.

RECOMMENDATION: (S//NF) I believe the time is right to designate an IC Executive Agency for sharing communications metadata as an integral part of a greater IC sharing environment. SIGINT metadata is a vast, rich source of information to build community collaboration and target knowledge and the emerging intelligence based target social network analysis discipline. NSA is prepared to accept this role and build the communications metadata coalition the IC needs for both current target tracking and future threat warning.

(S//NF) To this end, my staffs have been meeting with CIA, DIA, FBI and DEA metadata users and technical experts to work out the details of this proposal. Once

approved, we will be ready to implement this structure with a roll out to these agencies within a few months followed by the greater IC within a year.

DNI DECISION:

(U//FOUO) Establish an IC-wide communications metadata structure:	
Concur	
Non-concur	
Other	
(U//FOUO) Designate DIRNSA/CHCSS as the U.S. IC Executive Agent for communications metadata:	
Concur	
Non-concur	
Other	
(C) My POC for this action is Deputy Chief of Staff for Operations and	
Support, Signals Intelligence Directorate,	
@nsa.ic.gov.	

KEITH B. ALEXANDER Lieutenant General, U.S. Army Director, NSA/Chief, CSS

Cf: DoD USD(I)
DIR/CIA
DIR/DIA
DIR/FBI
DIR/DEA

POINT PAPER

22 Feb 06

SUBJECT: CRISSCROSS/PROTON

PURPOSE: Provide DIRNSA information on subject.

BACKGROUND:

- (U) History of CRISSCROSS/PROTON (how did it come about, when, and by whom.)
- (S//OC, NF) Project CRISSCROSS began in the early 1990's when CIA and DEA collaborated on a database and analytical tool suite to perform link analysis and cross-reference lookups against DEA and CIA collected telephone billing records and phone directories.
- (S//OC, NF) NSA Office of Crime and Narcotics requested the software and database capabilities of CRISSCROSS to enhance SIGINT data. CRISSCROSS proved extremely successful at identifying new phone numbers associated with 'lost' drug targets and locating additional targets of interest. Expanding the coverage from Latin America to worldwide then enhanced target development efforts across all NSA product lines.
- (U) Current CRISSCROSS/PROTON Agreements.
- (S//OC, NF) All CRISSCROSS/PROTON Agreements are via bilateral Memoranda of Understanding (MOU) and exist between CIA, DEA, NSA, DIA, and FBI. DEA and FBI do not currently share data with each other.
- (S//OC, NF) DIA provides support to various DoD elements (e.g JICs/JACs and other Service Intelligence structures) as specified in MOU Addendums.
- (U) How current data in CRISSCROSS/PROTON is "minimized" and by what standards.
- (S//OC, NF) From its inception, NSA has only provided telephony data to CROSSCROSS/PROTON, which constitutes: date, time, duration, called number and calling number. NSA 'minimizes" this metadata prior to its transmission of data to CIA.
- (S//NF) NSA data is pulled from the FASCIA and HOMEBASE systems and transmitted to a NSA terminal in CIA's PROTON Project Office.
- (S//SI) For U.S. and 2nd party numbers, minimization is achieved by replacing the rightmost 4 digits with four x's.

Derived From: NSA/CSSM 1-52

Dated: 20041123 Declassify On: 20290325

- (S//SI) The classification of the NSA call-event records is CONFIDENTIAL NOFORN.
- (S//OC, NF) NSA Office of General Counsel ruled that the data contributed by CIA, DIA, DEA, and FBI does not need to be minimized by NSA before being incorporated into the NSA PROTON system as each MOA signatory must "minimize" according to their own authorities.
- (U) How CRISSCROSS/PROTON operates today (data flows/tape courier, tools, etc.)



• (S//NF) PROTON is a proprietary system architecturally inconsistent with DoD and NSA development guidelines and environment. It requires specific licensing and maintenance expenses and currently cannot be integrated with other tools.



- (S//OC, NF) Data is transferred from NSA, DIA, and FBI to CIA electronically or by courier delivery of magnetic tapes.
- (U) Current proposals to modify (pending) CRISSCROSS/PROTON with the addition of other data elements/content/procedures
- (S//OC, NF) New communications technologies necessitated evolution of the CRISSCROSS software. The CRISSCROSS system only handled five call-record fields: date, time, duration, called number and calling number. PROTON goes further and analyzes additional data fields that are specific to cell and satellite phones, such as IMEIs and IMSIs, lat/long coordinates. PROTON also analyzes email headers (not content), SMS text messages, SIM cards, travel records (passport #, flight #, description), VISA application data and excerpts from CIA intelligence reports.
- (S//OC, NF) CIA, DIA, and FBI have added these additional data types to PROTON and share this data with NSA. However, NSA has never agreed to expand its contributions to PROTON.

- (S//OC, NF) PROTON performs entity analysis/contact chaining across data types. Users may search on a target entity's phone number and PROTON will return the associated email, cell phone, call records, street address, and CIA intelligence report excerpts.
- (S//OC, NF) PROTON tools find other entities that behave in a similar manner to a specific target, it also identifies correspondents in common with two or more targets, identifies potential new phone numbers when a target switches phones, and identifies networks of organizations based on communications within the group.
- (S//OC, NF) PROTON provides search capabilities for exploiting commercial phone books and data that do not conform to a particular type, such as excerpts from CIA intelligence reports.
- (U) Current NSA contribution to CRISSCROSS/PROTON (data elements/number of records to date/records per week/month, etc.)
- (S//OC, NF) CIA, DIA, FBI, and DEA contributions over a three-month period totaled 11.7 billion records (which includes many records that are not "communications-metadata contact records".)
- (S//SI//NF) NSA contributions over the same three-month period totaled 4.2 billion records (again, only limited to telephony contact data). (Note: under the NSA proposal to create an IC-wide communications metadata repository, NSA's contribution would be in the 100's of billions of contact records- telephony and DNI contact records.)
- (U) How NSA currently uses CRISSCROSS/PROTON in conjunction with other NSA metadata repositories and tools.
- (S//OC, NF) NSA analysts use Proton to perform analysis against data from CIA, NSA, DEA, FBI and DIA in a single common interface. Results provide additional targets of interest for all product lines. Merging the data produces a more complete picture of the target's communications and activity.
- (TS//SI//NF) NSA Users reported the following regarding the importance of PROTON:
 - (S) SIGINT analysts in Iraq attributed PROTON for successes at locating and apprehending High Value Individuals (HVIs). PROTON had the information they needed to get U.S. troops in the right location.
 - (S) PROTON was critical in capturing a HVI who was wanted by the FBI and AFOSI for terrorist related activity in Denver.
 - (S//SI) PROTON provided the location information of a money exchange house that was connected to a Pakistani militant group.

- (S) PROTON allows a single stop to find registration information for phones that are not being used by the person to whom they are registered. This is particularly important for front companies, where tying multiple phones to the same registration can provide critical information."
- (S) SID is recommending that a DNI IC-wide central communications metadata repository be established and reside at NSA and that DIRNSA/CHCSS be made Executive Agent for this IC repository. NSA and its foreign SIGINT partners stand alone in conducting metadata analysis on such a large scale and as a result, our tools and methodologies, as well as the infrastructure employed, are well tested and have been proven efficient and effective. We already extensively share metadata internally within the NSA/CSS enterprise and, having recently reached agreement with our Second Party partners, are ready to begin implementing that agreement to share Second Party metadata writ large.
- (S//NF) Rather than continuing to push communications metadata out to the PROTON repository, we propose using the existing IC shared information space ICSIS (Intelligence Community System for Information Sharing) on INTELINK and have IC users come in through an existing tool to pull data from metadata repositories resident at NSA. We would implement IC access to our GLOBALREACH federated query service via accounts and access verified by PKI certificates. This service will provide the access requested and permit the auditing of such repository users. We believe we can have GLOBALREACH capability ready for the IC within approximately one to two months of a decision to proceed with this proposal.

(U//FOUO) ORIGINATORs:	and	
for the IC-metadata proposal, and		
for CRISSCROSS/PROTON, 22 Feb 06		

(S//REL) CIA Colleagues Enthusiastically Welcome NSA Training