GlobalResearch
Center for Research on Globalization

Telephone Records are just the Tip of NSA's Iceberg

By William M Arkin
Global Research, May 14, 2006
Washington Post 14 May 2006

Region: USA
Theme: Police State & Civil Rights

The National Security Agency and other U.S. government organizations have developed hundreds of software programs and analytic tools to "harvest" intelligence, and they've created dozens of gigantic databases designed to discover potential terrorist activity both inside the United States and overseas.

These cutting edge tools — some highly classified because of their functions and capabilities — continually process hundreds of billions of what are called "structured" data records, including telephone call records and e-mail headers contained in information "feeds" that have been established to flow into the intelligence agencies.

The multi-billion dollar program, which began before 9/11 but has been accelerated since then. Well over 100 government contractors have participated, including both small boutique companies whose products include commercial off-the-shelf software and some of the largest defense contractors, who have developed specialized software and tools exclusively for government use.

USA Today provided a small window into this massive intelligence community program by reporting yesterday that the NSA was collecting and analyzing millions of telephone call records.

The call records are "structured data," that is, information maintained in a standardized format that can be easily analyzed by machine programs without human intervention. They're different from intercepts of actual communication between people in that they don't contain the "content" of the communications — content that the Supreme Court has ruled is protected under the Fourth Amendment. You can think of call records as what's outside the envelope, as opposed to what's on the inside.

Once collected, the call records and other non-content communication are being churned through a mind boggling network of software and data mining tools to extract intelligence. And this NSA dominated program of ingestion, digestion, and distribution of potential intelligence raises profound questions about the privacy and civil liberties of all Americans.

Although there is no evidence that the harvesting programs have been involved in illegal activity or have been abused to reach into the lives of innocent Americans, their sheer scope, the number of "transactions" being tracked, raises questions as to whether an all-seeing domestic surveillance system isn't slowly being established, one that in just a few years time will be able to reveal the interactions of any targeted individual in near real time.

In late November 1998, the intelligence community and the Department of Defense established the Advanced Research and Development Activity in Information Technology

(ARDA), a government consortium charged with incubating and developing “revolutionary” research and development in the field of intelligence processing.

The Director of the National Security Agency (NSA) agreed to establish, as a component of the NSA, an organizational unit to carry out the functions of ARDA, overseeing the research program of the CIA, DIA, National Reconnaissance Office, and other defense and civilian intelligence agencies.

Beginning before 9/11, ARDA established an “information exploitation” program to fund and focus private research on operationally-relevant problems of exploiting the increasing torrents of digital data available to the intelligence community. Even with thousands of analysts, NSA and other agencies were falling behind in their ability to handle the volume of incoming material. Existing mainframe machine aided processes were also falling behind advances in information processing, particularly as the cost of computing power dramatically declined in the 1990s.

The information exploitation research program has funded hundreds of projects to find better ways to “pull” information, “push” information, and “navigate” and visualize information once assembled.

Pulling information refers to the ability of supported analysts to have question and answer capabilities. Starting with a known requirement, an analyst could submit questions to a Q&A system which in turn would “pull” the relevant information out of multiple data sources and repositories. NSA is seeking a Q&A system that can operate autonomously to interpret “pulled” information and provide automatic responses back to the analysts with little additional human intervention.

Pushing information refers to the software tools that would “blindly” and without supervision push intelligence to analysts even if they had not asked for the information. Research has sought to go beyond current data mining of “structured” records deeper profiling of massive unstructured data collections. Under the pushing information research thrust companies have been involved in efforts to uncover previously undetected patterns of activity from massive data sets. Software and tools are also being developed that will provide alerts to analysts when changes occur in newly arrived, but unanalyzed massive data collections, such as telephone records.

The effort to navigate and visualize information seeks to develop analytic tools that will allow agency analysts to take hundreds or even thousands of small pieces of information and automatically create a tailored and logical “picture” of that information. Using visualization tools and techniques, intelligence analysts are constantly seeking out previously unknown links and connections between individual pieces of information.

Intelligence community efforts to process “structured” data includes data-tagged signals intelligence (SIGINT) monitoring of telephone and radio communications, imagery, human intelligence reporting, and “open-source” commercial data, including news media reporting. “Unstructured” data includes news and Internet video and audio and document exploitation.

I could write volumes about the research efforts and the software programs and tools used to process the mountains of information the NSA and other agencies ingest. No doubt over the coming days and weeks, more will be written. For today though, I provide a pointer, based upon my research, of software, tools and intelligence databases that I have been able

to identify in government documents relating to data mining, link analysis, and ingestion, digestion, and distribution of intelligence. My hope would be that other journalists and researchers will follow the leads.

The following is a list of some 500 software tools, databases, data mining and processing efforts contracted for, under development or in use at the NSA and other intelligence agencies today:

- A2IPB
- ABA
- ABC Terrorism Prediction Model
- ABIS (Automated Biometric Identification System)
- AC2
- ACCO (Army Central Control Office) counter-intelligence investigations database
- ACOA
- ACTOR (Analyzing Complex Threats for Operations and Readiness)
- Adversary
- AeroText
- AFIS (Automated Fingerprint Identification Systems)
- AIES (Automated Information Extraction System)
- AIM
- AIR (Arabic Information Retrieval):
- Aira Data Mining Tool
- AIPSA (Automated Intel Processing for Situational Awareness)
- aiSee
- AKA
- ALADDIN (Automated Link Analysis for Data Mining of Distributed Information)

- ALE (Aires Life Extension)
- Alembic
- ALICE d’ISoft
- Alien Migration db
- Alterian Nucleus
- AME (analyst modeling environment)
- Analyst Notebook/Analyst Notebook Link Chart Reader
- Analyst Workbench
- Anchory
- AnswerTree
- Answerer
- AOCG (automated org-chart generation)
- APOLLO
- Aquarius
- ARENA
- ARM
- ART (Author-Recipient-Topic)
- ASAS (All Source Analysis System)/ASAS-L with MAST
- ASID (Automated Systems Integration Management Intelligence Database)
- ASIM (Automated Security Incident Measurement)
- Association
- AT Sigma Data Chopper
- ATHENA

- ATIX (Anti-Terrorism Information Exchange/Automated Trusted Information Exchange):
- AUDITT
- AutoMap
- Autonomy
- Automatic Identification System
- Automated Warning Prototype
- Auto-X tools
- AutoTrackXP (ATXP)
- AXIS (Analysis and eXploration of Information Sources)
- AVS/Express Visualization Edition
- Basketball
- BioWar
- Blackknight
- Blue Data Miner
- BNN (Broadcast News Navigator)
- Breve
- Broadbase EPM (Enter Perf Mgmt)
- Brocade
- BUILDINGCODE
- BusinessMiner
- C2PC
- CADRE (Continuous Analysis and Discovery from Relational Evidence)

- CamStudio
- Camps
- Capri
- Carillon
- CART
- CASIAT (Computer Assisted Security Investigative Analysis Tool)
- Categorizer/Tree Studio
- CATEIS (Counterintelligence Automated Tactical Exploitation & Information Software)
- CCDB (Consolidated Counterdrug Data Base)
- CCIP (Counterterrorism Collaboration Interoperability Project)
- CCM
- Centrifuge
- CETA
- Chassis
- CHATS (CI/HUMINT Automation Tool Set)
- Checkpoint
- CHIMS (CI/HUMINT Information Management System)
- CHINET (Chinese Name Extraction and Translation)
- Choicepoint
- CIA TD/TDX
- CiceroLite:
- CIDAR (Critical Infrastructure Detection, Analysis and Reporting)

- CIM (Critical Intent Model)
- CIM/SEAS Nested Argumentation
- CIPA (Counter Insurgency Pattern Assessment)
- CIRC
- CIS (Case Information System)
- CITF (Criminal Investigative Task Force) Web-enabled Database
- CJMTK
- Clear Case
- Clear Quest
- ClearForest
- ClearResearch
- Clementine
- CMS (Case Management System)
- Cobra Focus
- CODIS (Combined DNA Index System)
- CoGen
- COGNET
- Cognos (COTS Tool for Report Generation)
- Coliseum
- Conceptual Model of Counter-Terrorism Operations
- Constant Web
- Construct
- Content Analyst

- Context
- Convera
- Cornerstone
- Coverterm
- CPOF
- CPXP
- CrimeLink
- CrissCross
- CrossGraphs
- CTAC (Counter-Terrorism Analysis Capability)
- CTDB (Combating Terrorism Database)
- CT-AVRS/CT-AVARS (Combined Theater – Analyst Vetted Relational System/Combined Theater Analyst-Vetted, Relational, Structured Database)
- Cubist
- Cultweave/Cultweave II
- CyberLINXX
- CyberTrans
- CYC
- Darwin
- Data Clarity
- Data Detective
- Data Logic/RDS
- DataMiner 3D

- Data mining suite
- DataMite
- DataScope
- Data Serfer
- DataSurferPlus
- Data Surveyor
- DB2
- dbProbe
- DCIIS
- DECIDE
- DecisionWORKS
- DeltaMiner
- Diamond
- DIAZ
- DIMS (Detainee Information Management System)
- DNA (Dynamic Network Analysis)
- Dollar-Dinar db ($$)
- DOORS
- Dream Media
- DTES
- DyNet
- Eagle Eye
- EARS

- EASYBORDER
- EICK
- EKM
- Elite Network Modeling
- Emergejust (EJ)
- Enterprise Miner
- ENVIE (Extensible News Video Exploitation for Intelligence Analysis)
- Envision
- EP
- FaceIt
- FactBrowser
- Fair Isaac
- Fascia
- FAST DIAMOND US
- FAST ISSM
- FAST toolbox for OOTW
- FastTalk
- Fastus/TextPro
- FatCat
- FeedDemon
- Festival
- FINTEL
- FOMA

- FORECITE Monitor
- Foreign terrorism communication profile database
- Foundationstone
- Fraud Investigator
- Freedom
- Fulcrum Knowledge Server
- GATE (General Architecture for Text Engineering)
- GBAE (Glass Box Analytical Environment)
- GCS
- GDA
- GDM-FC
- Geo-Browser
- Genysis
- GeoTagger
- GIDI
- GIP (Generic Intelligence Processor)
- Glide
- Graphlet
- GRAPHVIZ/graphvis
- Grindstone
- Groove Workspace:
- Groundbreaker
- Group Discovery Algorithm

- Guardian
- Harvester
- HD Map
- HIDTA DIG (High Intensity Drug Trafficking Areas Digital Information Gateway)
- Highpoint
- HIS (HUMINT Imagery Server)
- Homebase/Homebase II
- HMSng (HUMINT Management System next generation)
- HPS (HUMINT Processing System)
- Hydrant
- i2 Visual Notebook:
- IAA (Intelligence Analyst Associate)
- IAFIS (Integrated Automated Fingerprint Identification System)
- iBase
- IBM Intelligence Miner for Data
- IBM MT:
- IC ROSE
- IdentiFinder
- IDP (Intelligence discovery portal)
- IDW (Investigative Data Warehouse)
- IE CounterDrug
- IES (IIR Evaluation System)
- IFS (Intelligence Fusion System)

- ILS
- iMapData
- IMPACT (Intelligent Mining Platform for the Analysis of Counter Terrorism)
- InCLUEsion
- Indri
- InFact
- InfoMASQ
- Informix Red Brick Formation
- InfoWorkspace
- InfoXtract
- INQUEST
- IN-SPIRE
- Insta-Know
- Intelligent Miner for Text
- Intelligenxia
- InterSCOPE 3-D Geospatial Visualization
- INTREPID (Intelligence and Terrorist Photograph Identification Database)
- IR Discover
- ISM (InfoSphere Management System)
- ITN (Identification Tasking and Network)
- IWS
- Jabber
- Jaguar

- JDS
- JMIE
- Juggernaut
- JWARN (Joint Warning System)
- KATE-DataMining
- KDD Explorer/SRA KDD Explorer
- Keycard
- KeyPlayer
- KnowledgeSEEKER
- KnowledgeSTUDIO
- LADS (Linkage Analysis Database System)
- LAS
- LAW (Link Analysis Workbench)
- LDS (Lotus Discovery Server)
- LEADMiner (NIPS)
- LEMUR
- LexiQuest Mine
- LingPipe
- LingSoft
- LinkView
- LSI (Latent Semantic Indexer)
- Mage
- Mailorder

- Mainway
- Malolo
- Malta
- MAPLE
- MARS
- MatView
- MAUI NITE
- MayaVis
- Media Manager
- METIS
- METS (Metadata Extraction and Tagging Service)
- MindManager
- MINDS (Multilingual Interactive Document Summarization)
- MINER
- Minerva
- MineSet
- MiTAP (MITRE Text and Audio Processing)
- MODEL 1
- Modus Operandi Database
- Mohomine
- MPES
- Name Variant
- NameStats

- NDA (name data archive)
- NDCore
- NDEx (National Data Exchange)
- NDPIX (National Drug Pointer Index)
- NER (named entity extraction)
- Nested Vision 3D
- NetDraw
- Netica
- NetMap Analytics
- NetMiner
- NetOwl/Net Owl
- NETVIZ
- NexMiner
- NLP++
- Noöscape
- NORMALLAW
- NORA (Non-Obvious Relationship Analysis)
- NorthernLight
- NRM (NSA Reference Model)
- Nuggets
- Oasis
- ObjectFX
- OCULIS

- OMNIDEX
- OnTap
- OnTopic
- Open Visualization Data Explorer
- Optionspace Visualization
- ORA
- Orion
- ORIONMagic
- OSALAT (Open Source Automated Link Analysis Tool)
- Outline/Magic
- Overwatch
- PAC (Portal Automated Collection)
- Pajek
- Paladin
- Pantheon
- PASSGEAR
- PathFinder
- PEAC-WMDv5 Decision Support tool
- Pen-Link
- Pensa
- Pinpoint
- PINWALE
- Pipeline

- PIRANHA
- Plus
- Polaris
- PolyAnalyst
- PowerDrill
- Powerplant
- PPS
- Procon
- Project Foundry
- Proximity
- PrudSys Discoverer
- PTEK
- QACTIS (Question-Answering for Cross-Lingual Text, Image, and Speech)
- QANDA
- Query Tree NG/Querytree
- QUIET (QUery Improvement Elevation Technique)
- QKS Classifier
- Radiant Garnet
- Rampart
- Rational Rose
- Raven
- Razor
- REES

- Remedy
- Renoir/Renoir+
- REPAST
- Rigel
- RMS
- ROSID (Rapid Open Source Intelligence Deployment System)
- ROVER
- RPS
- S-PLUS
- SAFE
- SaffronNet
- Saffron Web
- SAIL Labs Media Mining (MM) and Communications Mining (CM) Tool
- SameTime
- Sandbox
- Sander
- SANDKEY.
- SAS/SAS Enterprise Miner
- SAVANT (Systematic Architecture for Virtual Analytic Net-centric Threat Information):
- Scenario
- SeaLink/SeaWatch
- SEAS (Structured Evidence Argumentation System)

- See 5/C5.0
- Semantic Forests
- Semantic Navigator
- Semantic Web
- SEMESTER
- Semio Taxonomy
- SERIF (Statistical Entity & Relation Information Finding)
- SIAAD (Secure Information Access Analysis & Dissemination)
- SIAM (Situational Influence Assessment Module)
- Siena
- SIFT
- SIGINT on Demand
- SKYWRITER
- Slate
- SmartDiscovery
- SmartDiscovery Analysis Server
- SmartDiscovery Awareness Server
- SNAKE (Social Network Analysis for Knowledge Exploitation)
- SOCIDS
- SoNIA
- Soundex
- Sparkler
- sphinxVision

- SPIRE/Themeview
- Springtide
- Spotfire Pro
- SPSS/Clementine
- SpyGLAS
- SS7
- SSNAStandpoint
- Starlight
- StarTree
- STK (Surge Toolkit)
- Strategic Player
- STRONG ANGEL
- Subdue
- SUMMAC (Summarization Analysis Conference)
- Suspect Finder
- Swarm
- Sybase
- Syllogic Data Mining
- Symphony
- SyNERA
- Synergist
- System Dynamic Modeling
- SYSTRAN

- TACS
- TAG Manager (Thematic Argument Group Manager)
- TAIC (Text Analysis International Corporation)
- TAPAS (Threat Anticipation Program Agent-Based Simulation of Terrorist Motivations, Objectives and Strategies)
- TEES (Trainable Evidence Extraction System)
- Tel-Scope
- TER (The Easy Reasoner)
- TERQAS (Time and Event Recognition for Question Answering Systems)
- Teradata
- TextWise
- Themelink
- ThingFinder
- THREADS (Threat HUMINT Reporting, Evaluation, Analysis, and Display System)
- Threat Tracker
- TILADS (Terrorist Identification Linkage Analysis Database)
- Timewall
- TIMEX
- Tina Tool
- Tinman
- TIRT
- TKT (Tacit Knowledge Toolkit)
- TMODS (Terrorist Modus Operandi Detection System)

- TOLLS
- Tom Sawyer
- TouchGraph
- Trailblazer
- TravelNet
- TRIDENT
- TRIST (“The Rapid Information Scanning Tool”)
- Trusted Wisdom
- TruTalk
- TTKB
- Ucinet
- UWIL
- Vantage Point
- Verity Indexer
- Verity K2E (Verity K2 Enterprise)
- Verity Locales
- Verity Profiler
- Verona
- VIA Repository
- Viceroy
- Virtual Situation Book
- Viscovery SOMine
- Visual Insights ADVIZOR

- Visual Links
- VisualMine
- ViTAP
- VizServer
- VKB (Virtual Knowledge Base) FINTEL ISM
- Voltaire
- WAEWarlord
- Watchtower
- Watson Pro
- Webster
- Webtas (Web-based Time Line Analysis System)
- WilmaScope
- Windgrinder/WnGrinder
- Wired
- Wrangler
- XMB (XML Metadata Browser)
- XpertRule Miner
- Yellowstone

http://blog.washingtonpost.com/earlywarning/2006/05/telephone_records_are_just_the.html#20161

Comment on Global Research Articles on our Facebook page

Become a Member of Global Research

Articles by: **William M Arkin**