



MARCIA HOFMANN ESO
ELECTRONIC FRONTIER FOUNDATION
454 SHOTWELL STREET
SAN FRANCISCO, CA 94110

May 8, 2998

Subject: INVESTIGATIVE DATA WAREHOUSE

FOIPA No. 1058805- 000

Dear Ms. Hofmann:

The enclosed documents were reviewed under the Freedom of Information/Privacy Acts (FOIPA), Title 5, United States Code, Section 552/552a. Deletions have been made to protect information which is exempt from disclosure, with the appropriate exemptions noted on the page next to the excision. In addition, a deleted page information sheet was inserted in the file to indicate where pages were withheld entirely. The exemptions used to withhold information are marked below and explained on the enclosed Form OPCA-16a:

Section 552

Section 552a

- | | | |
|--------------------------------------------|-----------------------------------------------|---------------------------------|
| <input type="checkbox"/> (b)(1) | <input type="checkbox"/> (b)(7)(A) | <input type="checkbox"/> (d)(5) |
| <input checked="" type="checkbox"/> (b)(2) | <input type="checkbox"/> (b)(7)(B) | <input type="checkbox"/> (j)(2) |
| <input type="checkbox"/> (b)(3) _____ | <input checked="" type="checkbox"/> (b)(7)(C) | <input type="checkbox"/> (k)(1) |
| _____ | <input checked="" type="checkbox"/> (b)(7)(D) | <input type="checkbox"/> (k)(2) |
| _____ | <input checked="" type="checkbox"/> (b)(7)(E) | <input type="checkbox"/> (k)(3) |
| _____ | <input type="checkbox"/> (b)(7)(F) | <input type="checkbox"/> (k)(4) |
| <input type="checkbox"/> (b)(4) | <input type="checkbox"/> (b)(8) | <input type="checkbox"/> (k)(5) |
| <input type="checkbox"/> (b)(5) | <input type="checkbox"/> (b)(9) | <input type="checkbox"/> (k)(6) |
| <input checked="" type="checkbox"/> (b)(6) | | <input type="checkbox"/> (k)(7) |

159 page(s) were reviewed and 159 page(s) are being released.

- Document(s) were located which originated with, or contained information concerning other Government agency(ies) [OGA]. This information has been:
 - referred to the OGA for review and direct response to you.
 - referred to the OGA for consultation. The FBI will correspond with you regarding this information when the consultation is finished.

You have the right to appeal any denials in this release. Appeals should be directed in writing to the Director, Office of Information and Privacy, U.S. Department of Justice, 1425 New York Ave., NW, Suite 11050, Washington, D.C. 20530-0001 within sixty days from the date of this letter. The envelope and the letter should be clearly marked "Freedom of Information Appeal" or "Information Appeal." Please cite the FOIPA number assigned to your request so that it may be easily identified.

The enclosed material is from the main investigative file(s) in which the subject(s) of your request was the focus of the investigation. Our search located additional references, in files relating to other individuals, or matters, which may or may not be about your subject(s). Our experience has shown, when ident, references usually contain information similar to the information processed in the main file(s). Because of our significant backlog, we have given priority to processing only the main investigative file(s).

If you want the references, you must submit a separate request for them in writing, and they will be reviewed at a later date, as time and resources permit.

See additional information which follows.

Sincerely yours,



David M. Hardy
Section Chief
Record/Information
Dissemination Section
Records Management Division

Enclosure(s)

The enclosed documents represent the fifth of a series of interim releases that will be made with regard to your request pertaining to the Investigative Data Warehouse.

For your information, some of the enclosed pages contain information about other subject matters. We consider this information "Outside the Scope" of your request, and have redacted the information and marked it as "Outside the Scope."

The first page of this correspondence reflects that 159 pages were reviewed for this release. There were actually 8982 pages reviewed, however a further review determined that 8823 pages of this material are not responsive to the request.

EXPLANATION OF EXEMPTIONS

SUBSECTIONS OF TITLE 5, UNITED STATES CODE, SECTION 552

- (b)(1) (A) specifically authorized under criteria established by an Executive order to be kept secret in the interest of national defense or foreign policy and (B) are in fact properly classified to such Executive order;
- (b)(2) related solely to the internal personnel rules and practices of an agency;
- (b)(3) specifically exempted from disclosure by statute (other than section 552b of this title), provided that such statute(A) requires that the matters be withheld from the public in such a manner as to leave no discretion on issue, or (B) establishes particular criteria for withholding or refers to particular types of matters to be withheld;
- (b)(4) trade secrets and commercial or financial information obtained from a person and privileged or confidential;
- (b)(5) inter-agency or intra-agency memorandums or letters which would not be available by law to a party other than an agency in litigation with the agency;
- (b)(6) personnel and medical files and similar files the disclosure of which would constitute a clearly unwarranted invasion of personal privacy;
- (b)(7) records or information compiled for law enforcement purposes, but only to the extent that the production of such law enforcement records or information (A) could be reasonably be expected to interfere with enforcement proceedings, (B) would deprive a person of a right to a fair trial or an impartial adjudication, (C) could be reasonably expected to constitute an unwarranted invasion of personal privacy, (D) could reasonably be expected to disclose the identity of confidential source, including a State, local, or foreign agency or authority or any private institution which furnished information on a confidential basis, and, in the case of record or information compiled by a criminal law enforcement authority in the course of a criminal investigation, or by an agency conducting a lawful national security intelligence investigation, information furnished by a confidential source, (E) would disclose techniques and procedures for law enforcement investigations or prosecutions, or would disclose guidelines for law enforcement investigations or prosecutions if such disclosure could reasonably be expected to risk circumvention of the law, or (F) could reasonably be expected to endanger the life or physical safety of any individual;
- (b)(8) contained in or related to examination, operating, or condition reports prepared by, on behalf of, or for the use of an agency responsible for the regulation or supervision of financial institutions; or
- (b)(9) geological and geophysical information and data, including maps, concerning wells.

SUBSECTIONS OF TITLE 5, UNITED STATES CODE, SECTION 552a

- (d)(5) information compiled in reasonable anticipation of a civil action proceeding;
- (j)(2) material reporting investigative efforts pertaining to the enforcement of criminal law including efforts to prevent, control, or reduce crime or apprehend criminals;
- (k)(1) information which is currently and properly classified pursuant to an Executive order in the interest of the national defense or foreign policy, for example, information involving intelligence sources or methods;
- (k)(2) investigatory material compiled for law enforcement purposes, other than criminal, which did not result in loss of a right, benefit or privilege under Federal programs, or which would identify a source who furnished information pursuant to a promise that his/her identity would be held in confidence;
- (k)(3) material maintained in connection with providing protective services to the President of the United States or any other individual pursuant to the authority of Title 18, United States Code, Section 3056;
- (k)(4) required by statute to be maintained and used solely as statistical records;
- (k)(5) investigatory material compiled solely for the purpose of determining suitability, eligibility, or qualifications for Federal civilian employment or for access to classified information, the disclosure of which would reveal the identity of the person who furnished information pursuant to a promise that his/her identity would be held in confidence;
- (k)(6) testing or examination material used to determine individual qualifications for appointment or promotion in Federal Government service the release of which would compromise the testing or examination process;
- (k)(7) material used to determine potential for promotion in the armed services, the disclosure of which would reveal the identity of the person who furnished the material pursuant to a promise that his/her identity would be held in confidence.

UNCLASSIFIED // FOR OFFICIAL USE ONLY

Investigative Data Warehouse Integration System (IDW-I)

ALL INFORMATION CONTAINED

HEREIN IS UNCLASSIFIED

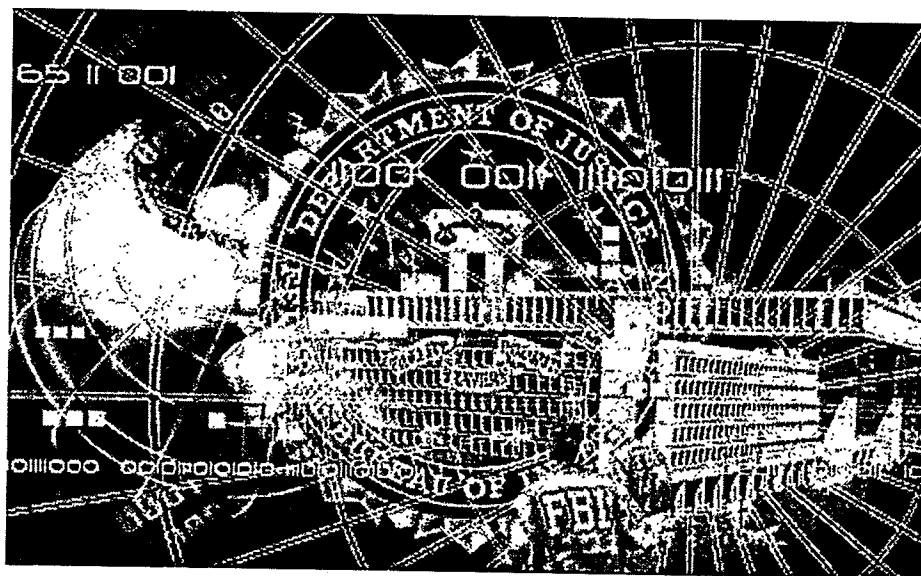
DATE 01-07-2008 BY 65179 DMH/BJA/CAL

System Security Plan

22 April 04

Version 0.5

**FEDERAL BUREAU OF INVESTIGATION
INFORMATION RESOURCES DIVISION
DATA AND INFORMATION MANAGEMENT SECTION**



**INVESTIGATIVE DATA WAREHOUSE INTEGRATION
SYSTEM (IDW-I)**

SYSTEM SECURITY PLAN

DOCUMENT VERSION 0.6

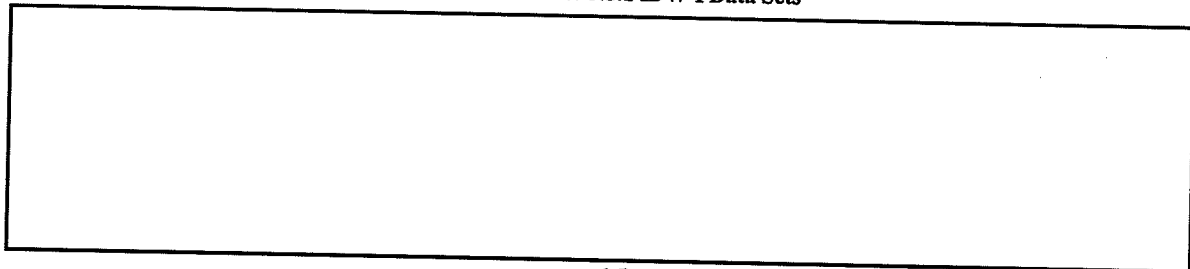
22 APRIL 04

UNCLASSIFIED // FOR OFFICIAL USE ONLY

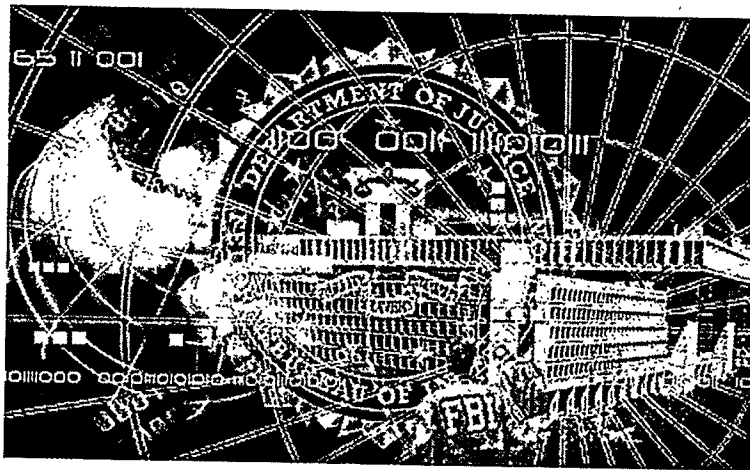
DATASET NAME	DATA SOURCE	CLASSIFICATION
ACS Electronic Case Files (ECF)	FBI Automated Case System (ACS)	Secret and below
IntelPlus Filerooms	FBI IntelPlus	Secret and below
SAMNet	FBI Secure Automated Messaging Network (SAMNet)	Secret and below
VGTOF	FBI National Crime Information Center (NCIC)	SBU
JICI	FBI Records Management Division (RMD), Document Laboratory (DocLab), FBIHQ	Secret and below
Open Source News	MITAP	Unclassified

b2
b7E

Table 3.5.1 IDW-I Data Sets



FEDERAL BUREAU OF INVESTIGATION
INFORMATION RESOURCES DIVISION
DATA AND INFORMATION MANAGEMENT SECTION



INVESTIGATIVE DATA WAREHOUSE INTEGRATION

SYSTEM (IDW-I)
SYSTEM SECURITY PLAN

ALL INFORMATION CONTAINED
HEREIN IS UNCLASSIFIED
DATE 01-07-2008 BY 65179 DMH/BJA/CAL

DOCUMENT VERSION 1.0

03 NOVEMBER 2004

DATASET NAME	DATA SOURCE	CLASSIFICATION
ACS Electronic Case Files (ECF)	FBI Automated Case System (ACS)	Secret and below
IntelPlus Filerooms	FBI IntelPlus	Secret and below
SAMNet	FBI Secure Automated Messaging Network (SAMNet)	Secret and below
VGTOF	FBI National Crime Information Center (NCIC)	SBU
JICI	FBI Records Management Division (RMD), Document Laboratory (DocLab), FBIHQ	Secret and below
Open Source News	MITAP	Unclassified

Table 3.5.1 IDW-I Data Sets

b2
b7E

UNCLASSIFIED//FOR OFFICIAL USE ONLY

FEDERAL BUREAU OF INVESTIGATION

OFFICE OF THE PROGRAM MANAGEMENT EXECUTIVE



ALL INFORMATION CONTAINED
HEREIN IS UNCLASSIFIED
DATE 01-07-2008 BY 65179 DMH/BJA/CAL

Security Concept of Operations (S-CONOPS) Investigative Data Warehouse (IDW) Program

Document Version 1.2

November 29, 2004

i

UNCLASSIFIED//FOR OFFICIAL USE ONLY



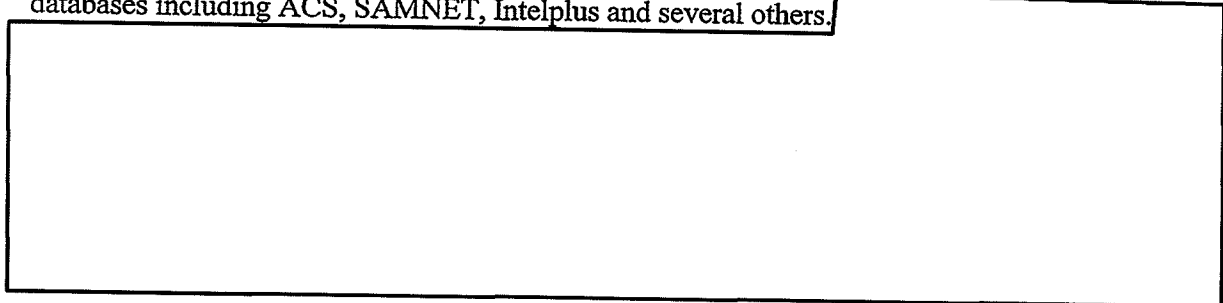
The table below lists the data and sources for the Operational Subsystem.

DATASET NAME	DATA SOURCE	CLASSIFICATION
ACS Electronic Case Files (ECF)	FBI Automated Case Support (ACS)	Secret and below
IntelPlus Filerooms	FBI IntelPlus	Secret and below
SAMNet	FBI Secure Automated Messaging Network (SAMNet)	Secret and below
VGTOF	FBI National Crime Information Center (NCIC)	SBU
JICI	FBI Records Management Division (RMD), Document Laboratory (DocLab), FBIHQ	Secret and below
Open Source News	MITAP	Unclassified

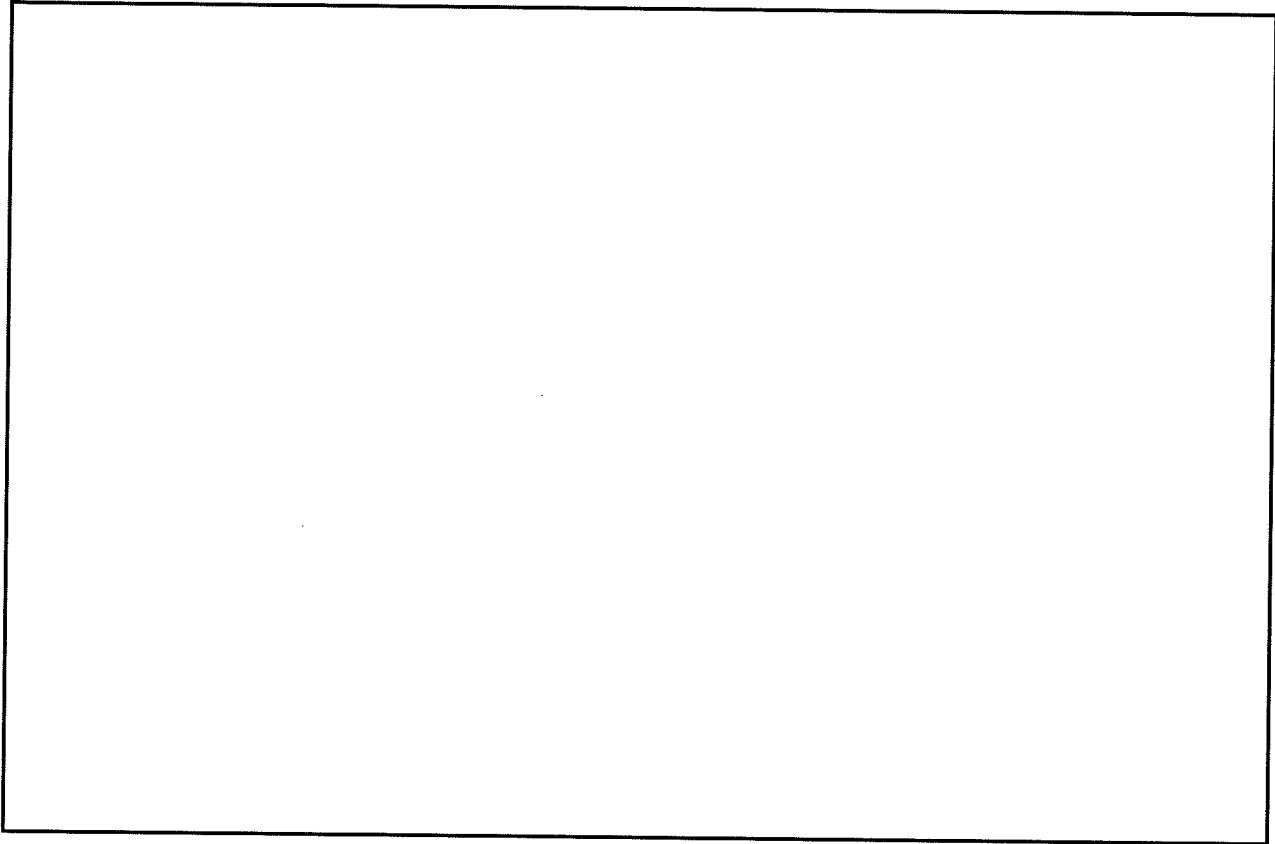
b2
b7E

Operational Subsystem Data Sources

The data flow begins at the external data sources. These sources include various FBI databases including ACS, SAMNET, Intelplus and several others.



b2
b7E



2.3.2 Special Projects Team (SPT) Subsystem

The Special Projects Team (SPT) Subsystem allows for the rapid import of new specialized data sources. These data sources are not made available to the general IDW users but instead are provided to a small group of users who have a demonstrated "need-to-know". The SPT System is similar in function to the IDW-S system, with the main difference is a different set of data sources. The SPT System allows its users to access not only the standard IDW Data Store but the specialized SPT Data Store.

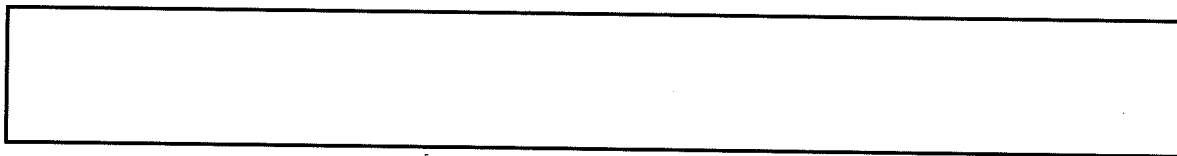
The table below lists the data and sources for the SPT Subsystem.

UNCLASSIFIED//FOR OFFICIAL USE ONLY

DATASET TYPE	DATA SOURCE	CLASSIFICATION
FBI Case Data	FBI Automated Case System (ACS)	Secret and below
Financial Center (FinCen) Data		SBU
Various Lists of Data	Various Sources, e.g., Transportation Security Administration (TSA)	SBU
Various CTD Datasets	FBI Counterterrorism Division (CTD)	Secret and below
Telephone Data	Database of telephone numbers from ACS indexed by phone number, names, addresses, and case numbers	Secret and below

SPT Data Sets

b2
b7E



Outside the Scope

DATASET NAME	DATA SOURCE	CLASSIFICATION
ACS Electronic Case Files (ECF)	FBI Automated Case System (ACS)	Secret and below
IntelPlus Filerooms	FBI IntelPlus	Secret and below
SAMNet	FBI Secure Automated Messaging Network (SAMNet)	Secret and below
VGTOF	FBI National Crime Information Center (NCIC)	SBU
JICI	FBI Records Management Division (RMD), Document Laboratory (DocLab), FBIHQ	Secret and below
Open Source News	MiTAP	Unclassified

b2
b7E

Table 3.5.1 IDW-I Data Sets

IDW-S V1.0 Data Migration Plan

September 30, 2003

ALL INFORMATION CONTAINED
HEREIN IS UNCLASSIFIED
DATE 01-07-2008 BY 65179 DMH/BJA/CAL

SCOPE V2.0/IDW V1.0 Data Sets

ACS ECF (CT Case Files):

- ◆ ASCII flat files (metadata and document text) and WordPerfect (document) binaries
- ◆ [REDACTED]
- ◆ Data is synchronized once a day, which includes processing of adds, deletes, and modifications

Intelplus (CT Filerooms):

- ◆ Scanned TIFF images and OCR'ed ASCII text documents
- ◆ [REDACTED]
- ◆ Data is synchronized once a week, which includes processing of adds, deletes, and modifications

SAMnet:

- ◆ ASCII files in standard cable traffic message format (all capitals with specific header)
- ◆ [REDACTED]
- ◆ SAMNET data is provided/updated three times a day

JICI:

- ◆ Scanned TIFF images and OCR'ed ASCII text (pages)
- ◆ [REDACTED]
- ◆ Static file system

b2
b7E

VGTOF:

- ◆ ASCII flat files (data/metadata, one row per subject) and JPEG image binaries (none, one, or multiple per subject)
- ◆ [REDACTED]
- ◆ At present, updated file being provided in its entirety on a weekly basis

Open Source:

- ◆ The Open Source Library is collected from global news sources that are either in English or have been translated into English
- ◆ [REDACTED]
- ◆ The Open Source Library is updated once a day

System Security Plan

Investigative Data Warehouse-SECRET (IDW-S)

3 DECEMBER 2003

Version: 0.6

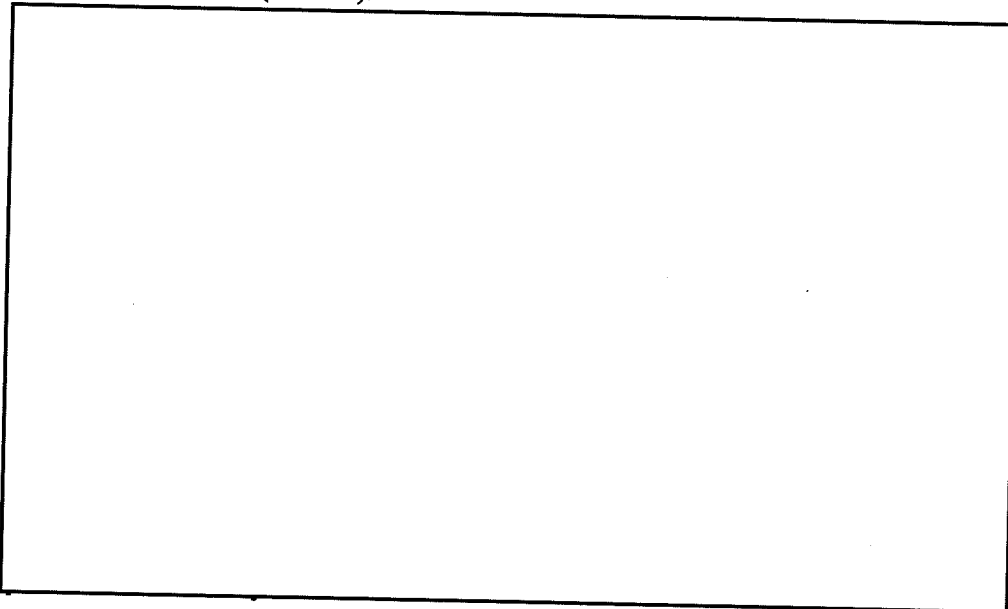
ALL INFORMATION CONTAINED
HEREIN IS UNCLASSIFIED
DATE 01-07-2008 BY 65179 DMH/BJA/CAL

**Federal Bureau of Investigation (FBI),
Information Resources Division (IRD)**

Investigative Data Warehouse (IDW) - Secret
System Security Plan

INTRODUCTION

The Federal Bureau of Investigation's (FBI) Investigative Data Warehouse (IDW-S) is an initial data warehouse, content management and data mining system that will permit FBI investigative, analytical, administrative and intelligence personnel to access aggregated data previously only available through individual applications. The IDW-S system will be authorized to process classified national security data up to, and including, Secret. The IDW-S system is the successor of the Secure Counter-Terrorism/Collaboration Operational Prototype Environment (SCOPE™).

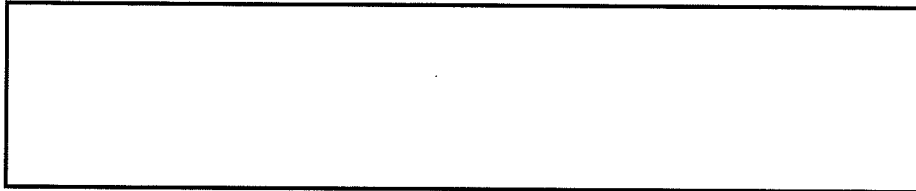


Outside the Scope

Data processed by the system will include the following data sets:

- Approved case files from the FBI's Automated Case Support (ACS) case management system;
- Electronic versions of the Joint Intelligence Committee Investigation (JICI) defined archived documents;

- Secure Automated Messaging Network (SAMNet) message traffic;
- IntelPlus File Rooms; and
- Violent Gang and Terrorist Organization File (VGTOF") from the Criminal Justice Information Systems (CJIS) Division;
- Defense Advanced Research Projects Agency (DARPA) Translingual Information Detection, Extraction and Summarization (TIDES) Open Source Data



Outside the Scope

Data Sources

IDW-S contains data from the following data sources with security classification indicated in brackets:

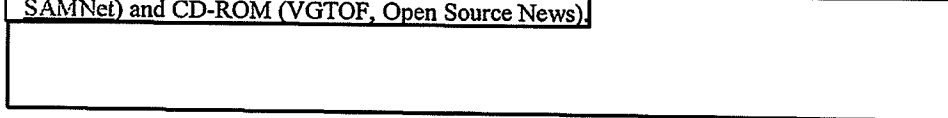
- Automated Case Support (ACS) System, Electronic Case File (ECF) Subsystem [Secret and below]
- IntelPlus [Secret and below]
- Secure Automated Message Network – Secret (SAMNet-S) [Secret and below]
- NCIC Violent Gang and Terrorist Offender File (VGTOF) [Sensitive But Unclassified]
- Joint Intelligence Committee Investigation (JICI) [Secret and below]
- DARPA/TIDES Open Source News [Unclassified]


Data from each source is held in an identified source Library against which users can direct the analytical tools.

Data Ingest

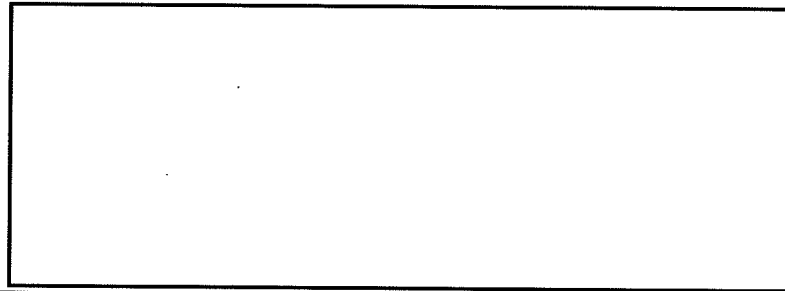
IDW-S receives data from sources by FTP over FBNNet (ACS ECF, IntelPlus, and SAMNet) and CD-ROM (VGTOF, Open Source News).

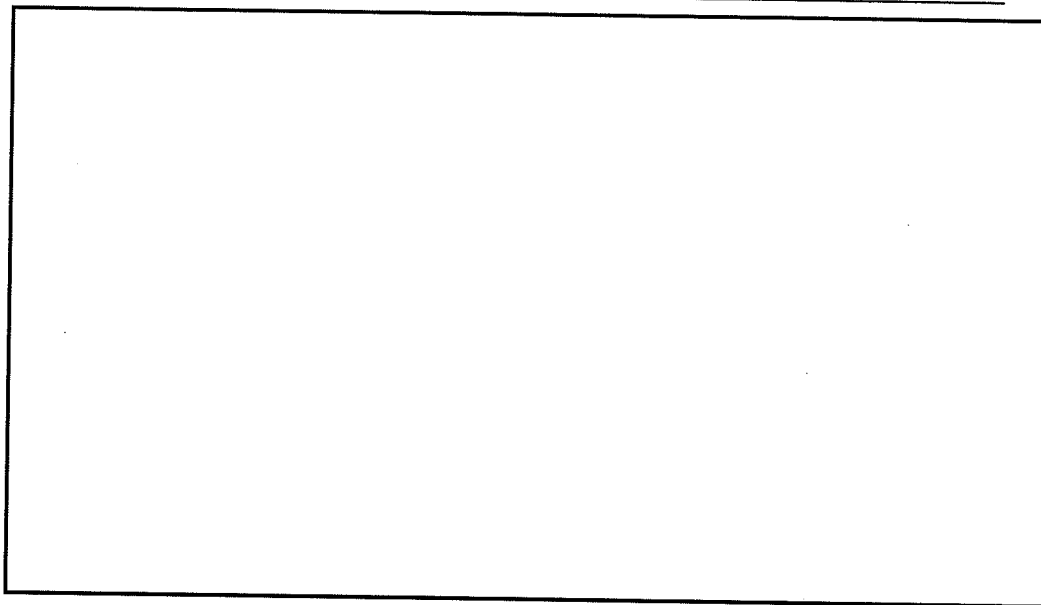
Deleted: SAMNet



- Data is received by the Ingest Subsystem
- 
- Received data is parsed into individual serial records (in the case of ACS ECF), documents (IntelPlus, JICI, Open Source News), messages (SAMNet), or files (VGTOF).

Outside the Scope





Outside the Scope

3.5. DATA PROCESSED

IDW-S processes the following data:

3.5.1 ACS ECF

IDW-S contains a subset of the ECF (Electronic Case File) subsystem of the Automated Case System (ACS). This subset consists of serials in those case classifications/sub classifications that have been officially sanctioned for inclusion in IDW. For each such serial, the ECF data includes metadata and text. IDW-S is synchronized against the ECF system once a day, a process which consists of receiving and processing the previous day's increment of ADD, MOD, and DELETE records.



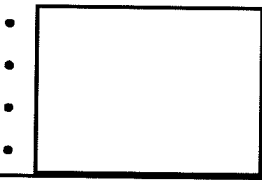
Outside the Scope

3.5.2 IntelPlus

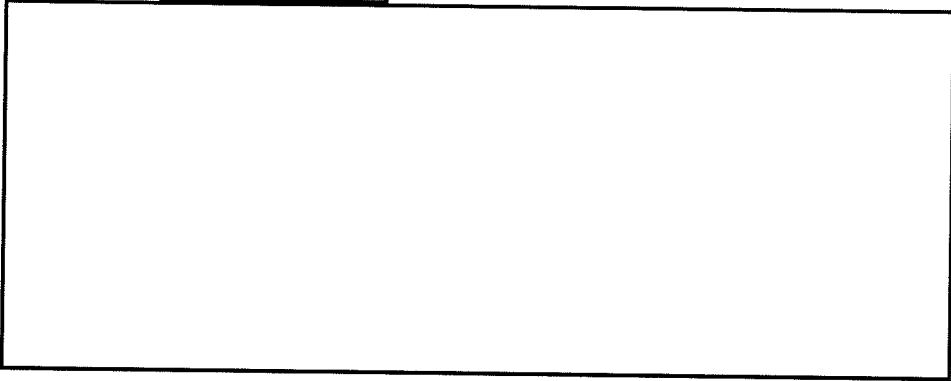
IDW-S currently contains several IntelPlus counter-terrorism (CT) Filerooms:

- 
- 

b2
b7E



b2
b7E



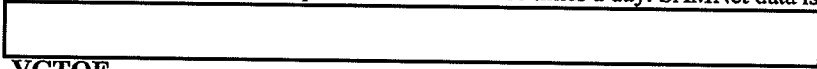
b2
b7E

3.5.3 SAMNet

The Secure Automated Messaging Network (SAMNet) data source consists of cable traffic messages received by the FBI.

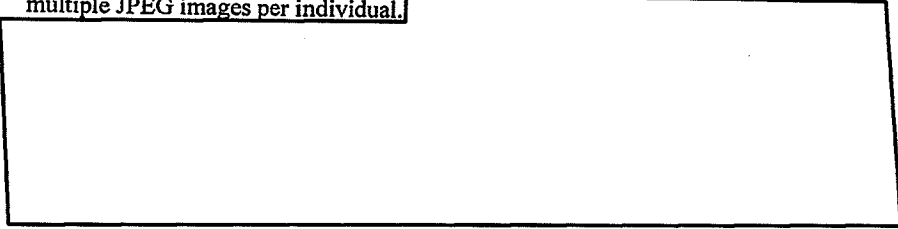


SAMNet provides only ADD record types, and SAMNet data is updated in IDW-S three times a day. SAMNet data is



3.5.4 VGTOF

Violent Gang Terrorist Offender File (VGTOF) data is provided by the FBI National Crime Information Center (NCIC). VGTOF data includes two components: Data/metadata for each named individual/offender and potentially multiple JPEG images per individual.



b2
b7E

3.5.5 JICI

The Joint Intelligence Committee Investigation (JICI) data collection was created following the September 11 attacks. Paper counter-terrorism files/folders from all FBI Field Offices

[Redacted]

[Redacted] is not updated.

[Redacted] JICI is a static collection that

b2
b7E

3.5.6 Open Source News

The IDW-S V1.0 system contains an Open Source News library collected by the DARPA TIDES Program. These are primarily news source from around the world that are either in English or have been translated into English. Open Source News data is received in the form of text files (one per news article) and [Redacted]

[Redacted] The Open Source News data goes into IDW-S once a day. Open Source News data [Redacted]

b2
b7E

Open Source News material is derived from the following sources:

- Addis Ababa Tribune - <http://www.addistribune.com/>
- Agencia Brasilia - <http://www.radiobras.gov.br/>
- Al-Ahram (Egypt, weekly English version) - <http://weekly.ahram.org.eg/>
- AllAfrica.com - <http://allafrica.com/>
- Arabic News - <http://www.arabicnews.com/ansub/>
- Asahi Shimbun - <http://www.asahi.com/english/>
- Asia Times (Hong Kong) - <http://www.atimes.com/>
- Bangkok Post - <http://www.bangkokpost.net/>
- Christian Science Monitor - <http://www.csmonitor.com/>
- Crescent International - <http://www.muslimedia.com/mainpage.htm>
- Daily Telegraph (London, England) - <http://news.telegraph.co.uk/>
- Dawn (Karachi, Pakistan) - <http://www.dawn.com/>
- Debka (Israel) - <http://www.debka.com>
- East Africa Daily Nation - www.nationaudio.com/News/DailyNation/Today/
- Gulf News (UAE) - <http://www.gulf-news.com/>
- Ha'aretz (Israel) - <http://www.haaretzdaily.com/>
- IFRC International Federation of the Red Cross - <http://www.ifrc.org/>
- IRIN Integrated Regional Information Network - <http://www.irinnews.org/>
- Iraq Press News Agency - <http://www.iraqpress.org/>

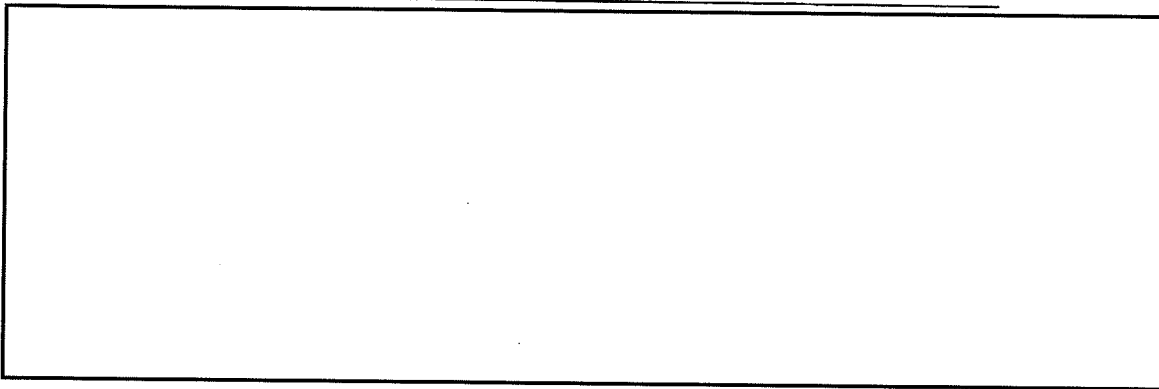
Deleted: Internation

Islamic Republic News Agency (Iran) - <http://www.irna.com/en>
Jakarta Post - <http://www.thejakartapost.com/>
Janes - <http://www.janes.com/>
Jordan Times - <http://www.jordantimes.com/>
L'Osservatore Romano - www.vatican.va/news_services/or/or_eng/text.html
Lagos (Nigeria) Guardian - <http://www.guardiannewsngr.com/>
Lahore (Pakistan) Nation - <http://www.nation.com.pk/>
Lebanon Daily Star (Beirut) - <http://www.dailystar.com.lb/>
Malaysian Star - <http://thestar.com.my/>
Manila Bulletin - <http://www.mb.com.ph/>
Manila Times - <http://www.manilatimes.net/>
Miami Herald - <http://www.miami.com/>
Moscow Times - <http://www.themoscowtimes.com/>
National Post and CP - <http://www.canada.com/>
New Straits Times (Kuala Lumpur) - http://www.imedia.com.my/Current_News/NST/
PETRA (Jordanian News Agency) - <http://www.petra.gov.jo>
Pakistan Observer (Islamabad) - <http://pakobserver.net/>
Palestine Chronicle - <http://palestinechronicle.com/>
People's Daily (China) - <http://english.peopledaily.com.cn/>
Philippine Star - <http://www.philstar.com/philstar/>
Pravda - <http://english.pravda.ru/>
ProMed - epidemiology mailing list
Russian Information Agency Novosti - <http://en.rian.ru/>
Russian Issues (Misc. Russian news) - <http://www.therussianissues.com/>
Russian Observer - <http://www.russianobserver.com/>
SABA (The News Agency of Yemen) - <http://www.sabanews.gov.ye>
Saudi Gazette - <http://www.saudigazette.com.sa/sgazette/>
South African Dispatch - <http://www.dispatch.co.za/>
Sydney Morning Herald - <http://www.smh.com.au>
Tehran Times - <http://www.tehrantimes.com/>
Times of India - <http://timesofindia.indiatimes.com/cms.dll>
UNHCR UN High Commissioner on Refugees - <http://www.unhcr.ch/>
Ummah News - <http://www.ummahnews.com/>
Uzbekistan Report - <http://www.uzreport.com/eng/>
Washington Post - <http://www.washingtonpost.com/>
XinHua News Service - <http://www.xinhuanet.com/english/>
Yemen Times (weekly) - <http://www.yementimes.com/>
Yomiuri Shimbun (Japan) - <http://www.yomiuri.co.jp/>

3.5.7 Summary of Data Sources

DATASET NAME	DATA SOURCE	TRANSFER METHOD
ACS	FBI Automated Case System (ACS)	FTP
IntelPlus	FBI Intel Plus	FTP
SAMNet	FBI Secure Automated Messaging Network (SAMNet)	FTP
VGTOF	FBI National Crime Information Center (NCIC)	CD-ROM
JICI	FBI Records Management Division (RMD), Document Laboratory (DocLab), FBIHQ	Multiple
Open Source News	Translingual Information Detection, Extraction and Summarization (TIDES) Program, Defense Advanced Research Projects Agency (DARPA)	CD-ROM

b2
b7E



3.9.9 Indirect Connections

Outside the Scope

IDW-S connects to the following sources of data:

<u>FBI Automated Case System (ACS)</u>	<u>Secret</u>	<u>FBI</u>	<u>FTP</u>	
<u>FBI Intel Plus</u>	<u>Secret</u>	<u>FBI</u>	<u>FTP</u>	

b2
b7E

Investigative Data Warehouse (IDW) - Secret
 System Security Plan
 3 December 2003
 Version 0.6

<u>FBI Secure Automated Messaging Network (SAMNet)</u>	<u>Secret</u>	<u>FBI</u>	<u>FTP</u>	
<u>FBI National Crime Information Center (NCIC)</u>	<u>Unclassified/SBU</u>	<u>FBI</u>	<u>CD-ROM</u>	
<u>FBI Document Conversion Laboratory (DocLab), Records Management Division (RMD), FBIHQ</u>	<u>Secret</u>	<u>FBI</u>	<u>Multiple</u>	
<u>Translingual Information Detection Extraction and Summarization (TIDES) Program, Defense Advanced Research Projects Agency (DARPA)</u>	<u>Unclassified</u>	<u>NA</u>	<u>CD-ROM</u>	

b2
 b7E

Investigative Data Warehouse (IDW) - Secret
System Security Plan
3 December 2003
Version 0.6

	<u>Unclassified</u>	NA		

b2
b7E

Indirect Connections

Formatted: Font: Not Italic

Formatted: Centered

SYSTEM SECURITY PLAN

INVESTIGATIVE DATA WAREHOUSE
(IDW) - SECRET

17 OCTOBER 2003

VERSION: 0.1

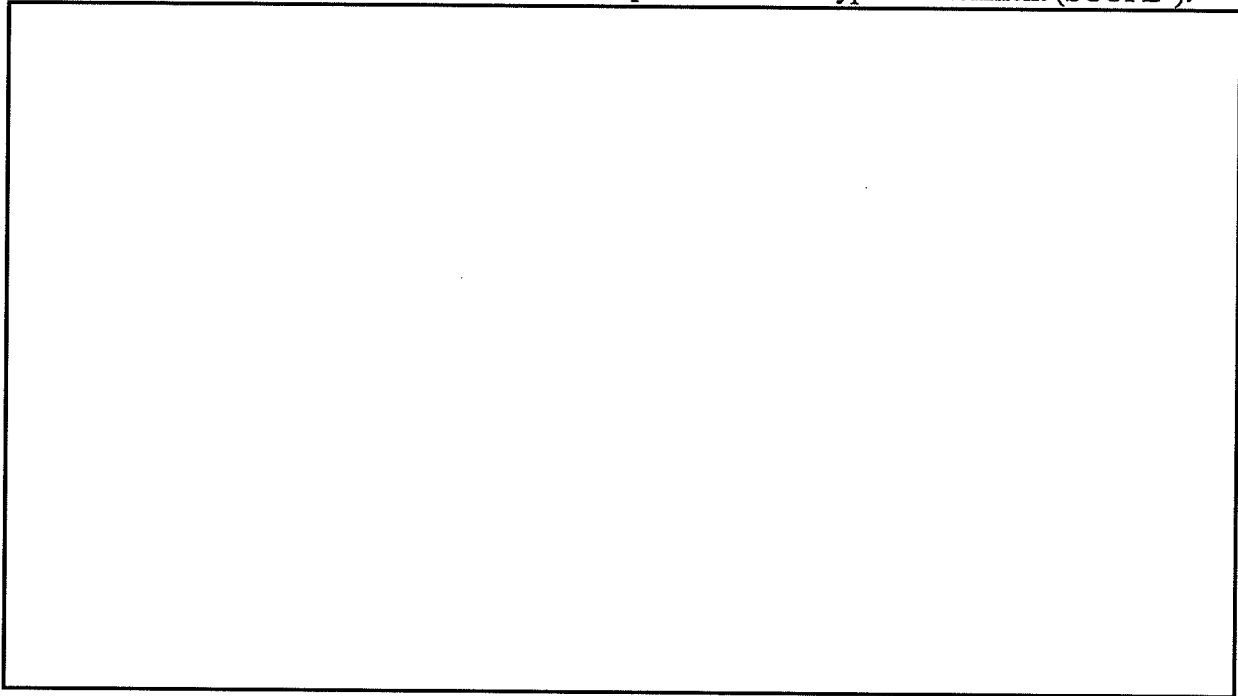
ALL INFORMATION CONTAINED
HEREIN IS UNCLASSIFIED
DATE 01-07-2008 BY 65179 DMH/BJA/CAL

FEDERAL BUREAU OF INVESTIGATION
(FBI), INFORMATION RESOURCES
DIVISION (IRD)

INVESTIGATIVE DATA WAREHOUSE (IDW) - SECRET
SYSTEM SECURITY PLAN

INTRODUCTION

The Federal Bureau of Investigation's (FBI) Investigative Data Warehouse (IDW-S) is an initial data warehouse, content management and data mining system that will permit FBI investigative, analytical, administrative and intelligence personnel to access aggregated data previously only available through individual applications. The IDW-S system will be authorized to process classified national security data up to, and including, Secret. The IDW-S system is the successor of the Secure Counter-Terrorism/Collaboration Operational Prototype Environment (SCOPE™).

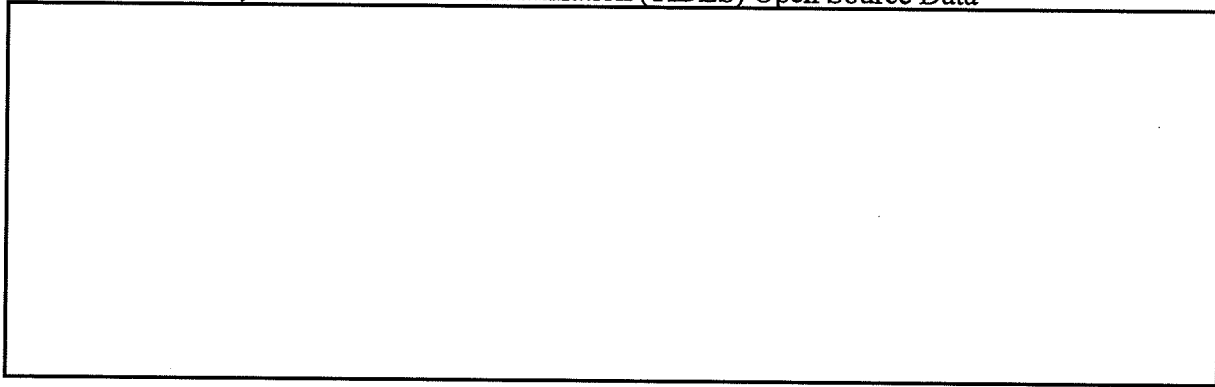


Data processed by the system will include the following data sets:

Outside the Scope

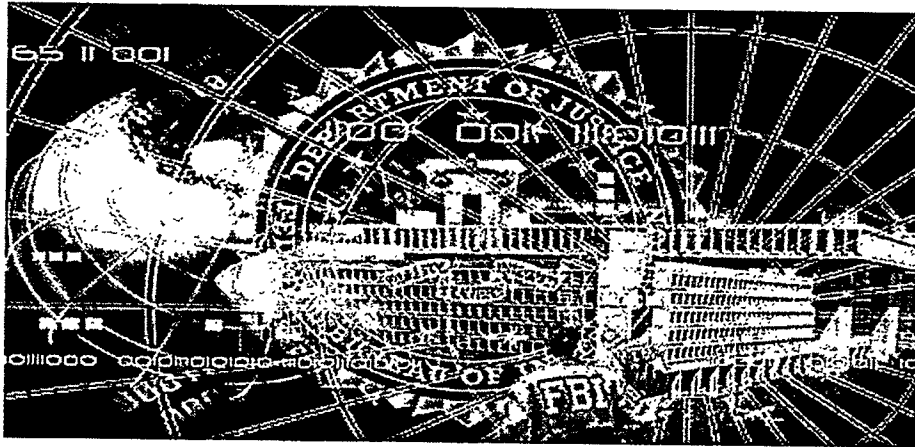
- Approved case files from the FBI's Automated Case Support (ACS) case management system;
- Electronic versions of the Joint Intelligence Committee Investigation (JICI) defined archived documents;
- Secure Automated Messaging Network (SAMNet) message traffic;
- IntelPlus File Rooms; and

- Violent Gang and Terrorist Organization File (VGTOF™) from the Criminal Justice Information Systems (CJIS) Division;
- Defense Advanced Research Projects Agency (DARPA) Translingual Information Detection, Extraction and Summarization (TIDES) Open Source Data



Outside the Scope

FEDERAL BUREAU OF INVESTIGATION
INFORMATION RESOURCES DIVISION
DATA/INFORMATION MANAGEMENT SECTION



**Test & Evaluation System Test Description (TESTD)
for the Investigative Data Warehouse (IDW) v.1.0-S**

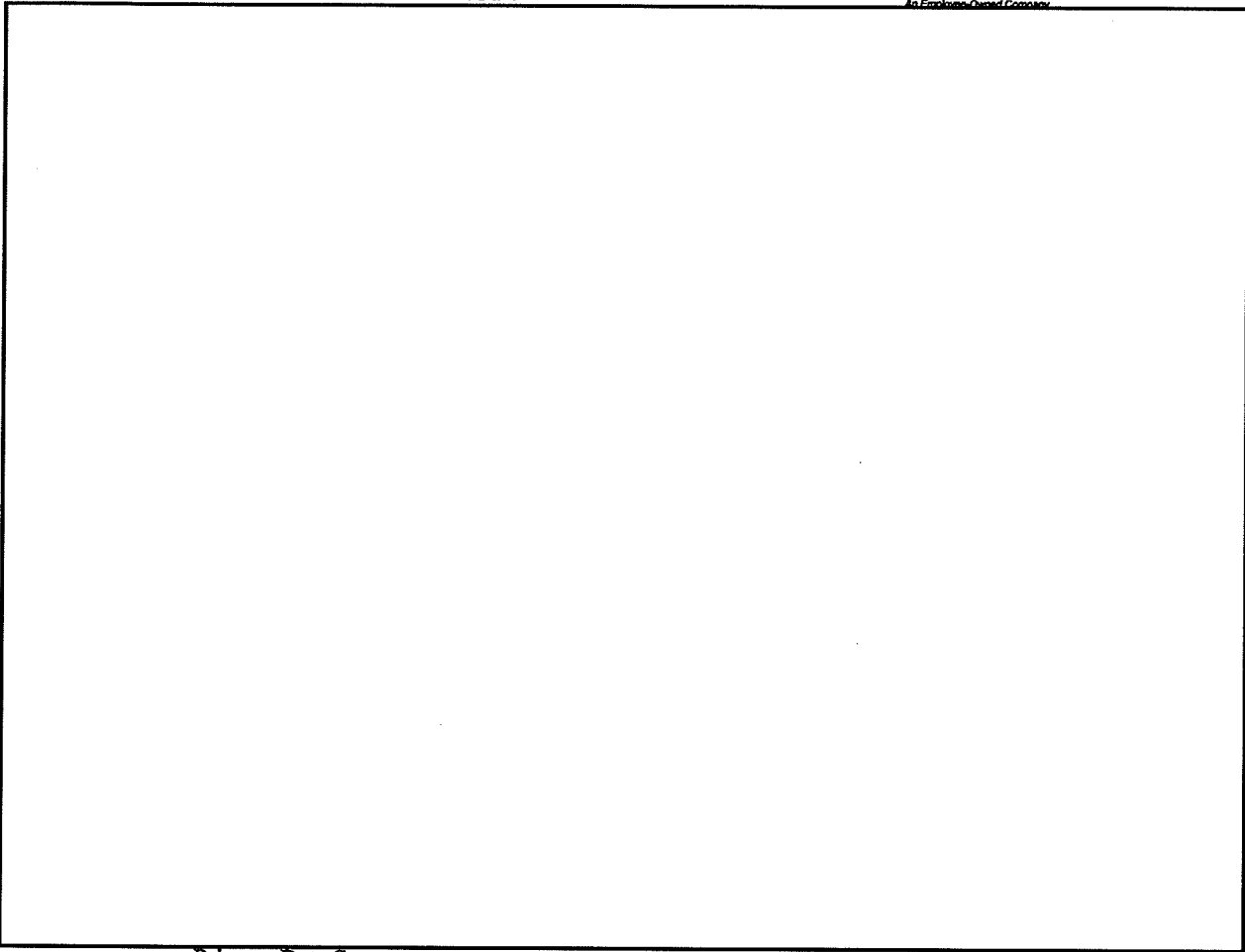
ALL INFORMATION CONTAINED
HEREIN IS UNCLASSIFIED
DATE 01-07-2008 BY 65179 DMH/BJA/CAL

October 2003

Prepared by:

Science Applications International Corporation
1213 Jefferson Davis Highway, Suite 1200
Arlington, Virginia 22202





- Primary Data Sources:

- JICI
- SAMNET
- ACS ECF
- INTEL+
- Open Source
-
- VGTOF

Outside the Scope

b2
b7E

DRAFT

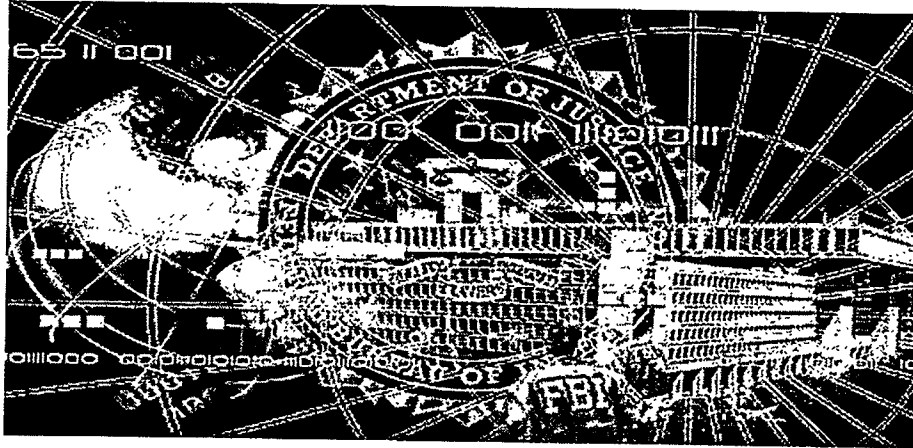
TEMP-IDW-S v0



DRAFT

ALL INFORMATION CONTAINED
HEREIN IS UNCLASSIFIED
DATE 01-07-2008 BY 65179 DMH/BJA/CAL

FEDERAL BUREAU OF INVESTIGATION
INFORMATION RESOURCES DIVISION
DATA/INFORMATION MANAGEMENT SECTION



**Test & Evaluation Master Plan (TEMP)
for the Investigative Data Warehouse (IDW)**

October 2003

Prepared by:

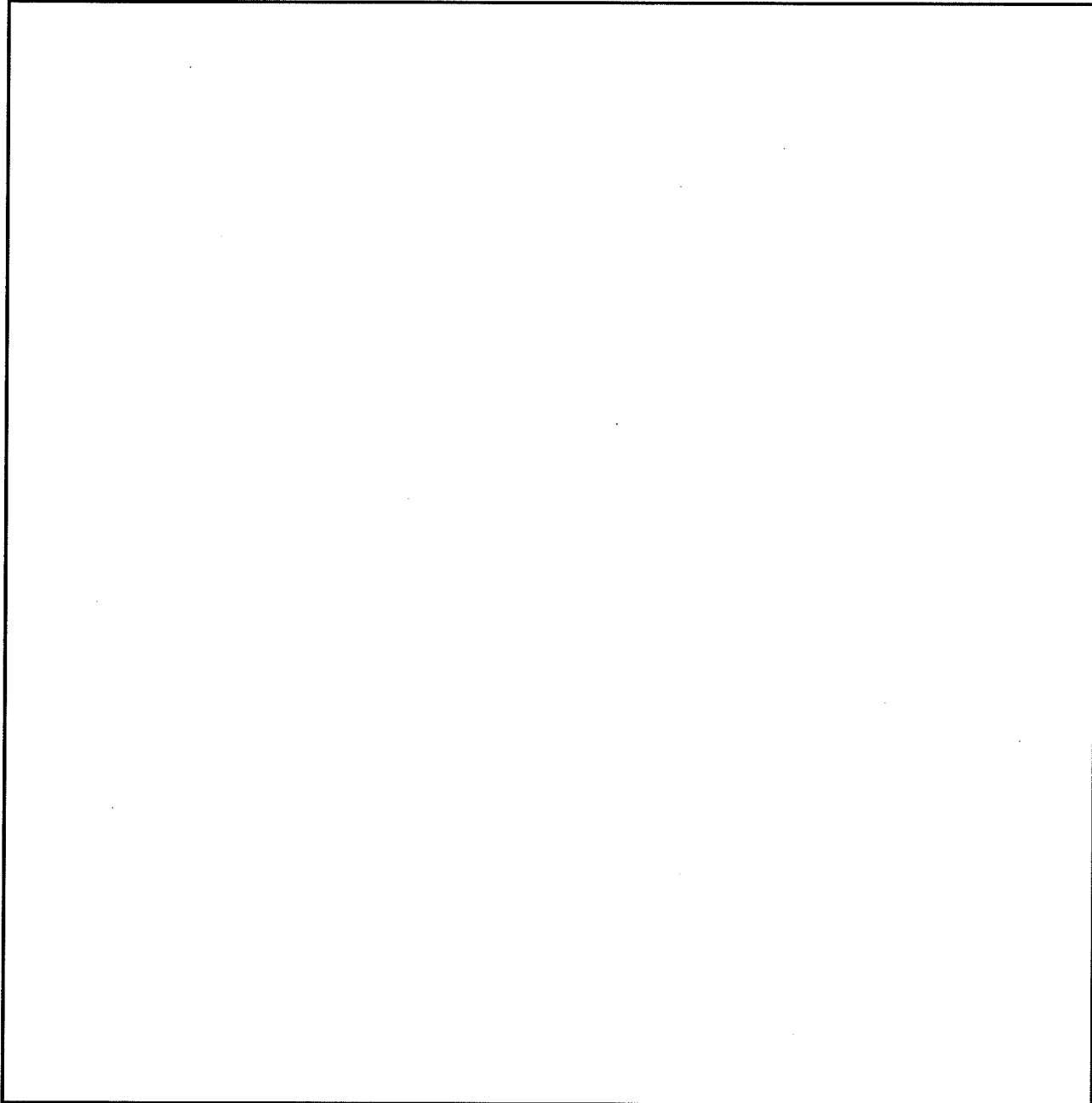
Science Applications International Corporation
1213 Jefferson Davis Highway, Suite 1200
Arlington, Virginia 22202



• Primary Data Sources:

- JICI
- SAMNET
- ACS ECF
- INTEL+
- Open Source b2
- b7E
- VGTOF

Outside the Scope



System Security Plan

Investigative Data Warehouse-SECRET (IDW-S)

5 DECEMBER 2003

Version: 0.7

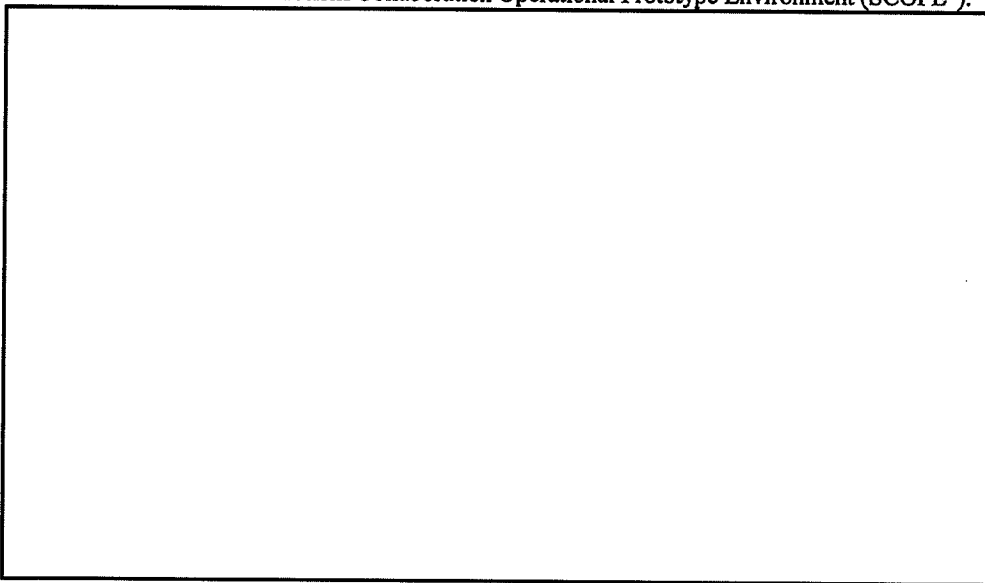
ALL INFORMATION CONTAINED
HEREIN IS UNCLASSIFIED
DATE 01-07-2008 BY 65179 DMH/BJA/CAL

**Federal Bureau of Investigation (FBI),
Information Resources Division (IRD)**

Investigative Data Warehouse (IDW) - Secret
System Security Plan

INTRODUCTION

The Federal Bureau of Investigation's (FBI) Investigative Data Warehouse (IDW-S) is an initial data warehouse, content management and data mining system that will permit FBI investigative, analytical, administrative and intelligence personnel to access aggregated data previously only available through individual applications. The IDW-S system will be authorized to process classified national security data up to, and including, Secret. The IDW-S system is the successor of the Secure Counter-Terrorism/Collaboration Operational Prototype Environment (SCOPE™).

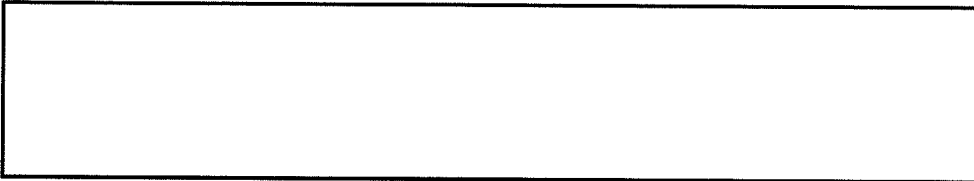


Outside the Scope

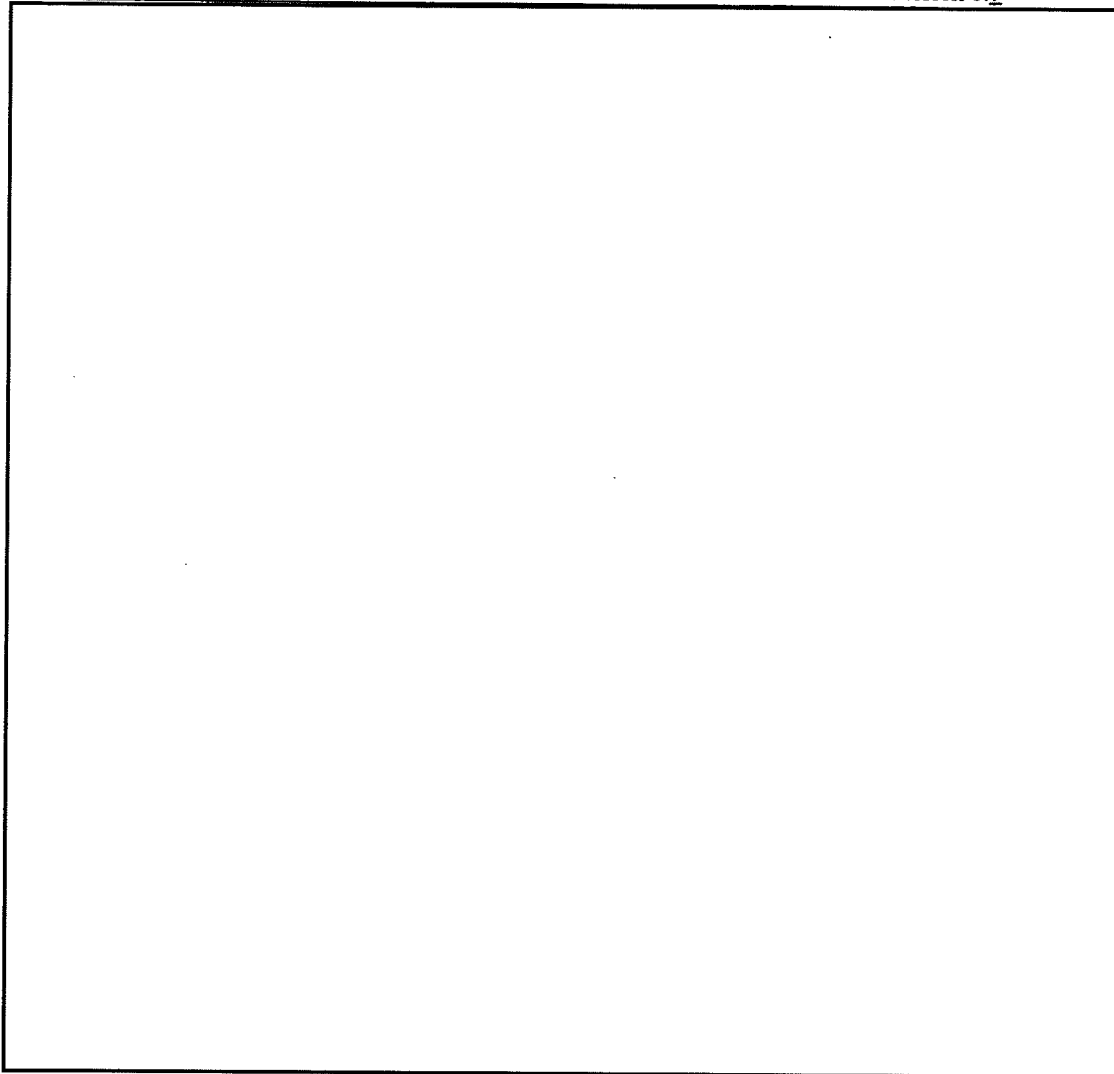
Data processed by the system will include the following data sets:

- Approved case files from the FBI's Automated Case Support (ACS) case management system;
- Electronic versions of the Joint Intelligence Committee Investigation (JICI) defined archived documents;
- Secure Automated Messaging Network (SAMNet) message traffic;

- IntelPlus File Rooms; and
- Violent Gang and Terrorist Organization File (VGTOF™) from the Criminal Justice Information Systems (CJIS) Division;
- Defense Advanced Research Projects Agency (DARPA) Translingual Information Detection, Extraction and Summarization (TIDES) Open Source Data



Outside the Scope



Data Sources

Outside the Scope

IDW-S contains data from the following data sources with security classification indicated in brackets:

- Automated Case Support (ACS) System, Electronic Case File (ECF) Subsystem [Secret and below]
- IntelPlus [Secret and below]
- Secure Automated Message Network – Secret (SAMNet-S) [Secret and below]
- NCIC Violent Gang and Terrorist Offender File (VGTOF) [Sensitive But Unclassified]
- Joint Intelligence Committee Investigation (JICI) [Secret and below]
- DARPA/TIDES Open Source News [Unclassified]

Data from each source is held in an identified source Library against which users can direct the analytical tools.

Data Ingest

IDW-S receives data from sources by FTP over FBI Net (ACS ECF, IntelPlus, and SAMNet) and CD-ROM (VGTOF, Open Source News)

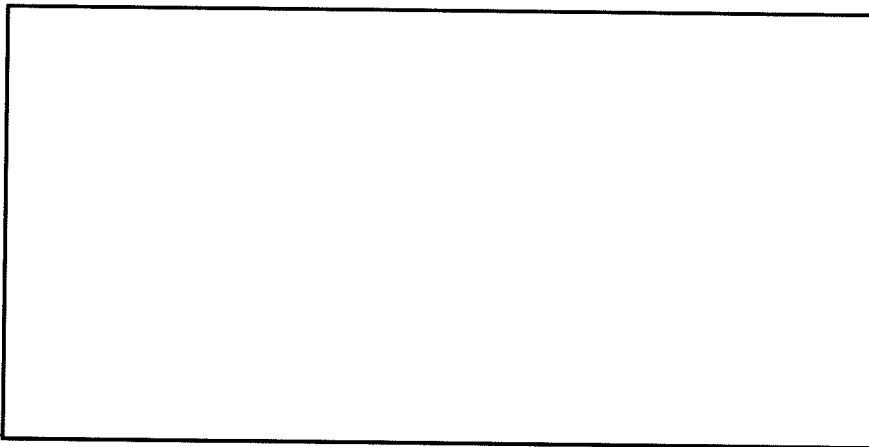


- Data is received by the Ingest Subsystem



Outside the Scope

- Received data is parsed into individual serial records (in the case of ACS ECF), documents (IntelPlus, JICI, Open Source News), messages (SAMNet), or files (VGTOF).



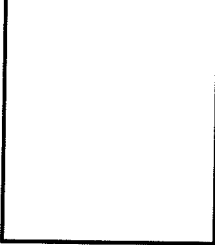
have been officially sanctioned for inclusion in IDW. For each such serial, the ECF data includes metadata and text. IDW-S is synchronized against the ECF system once a day, a process which consists of receiving and processing the previous day's increment of ADD, MOD, and DELETE records



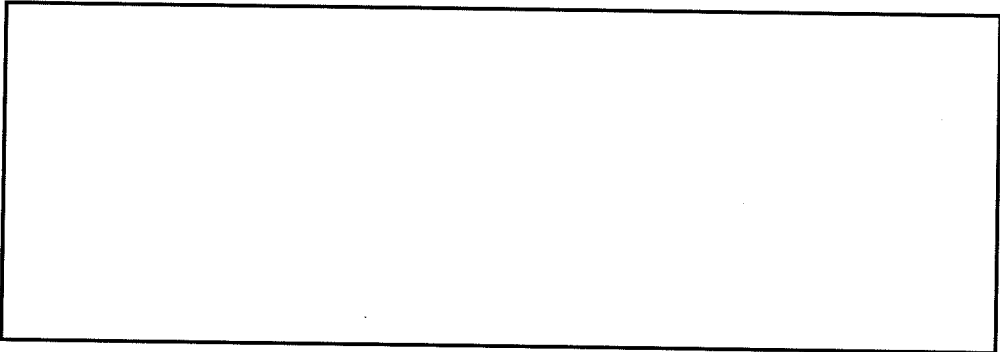
3.5.2 IntelPlus

Outside the Scope

IDW-S currently contains several IntelPlus counter-terrorism (CT) Filerooms:

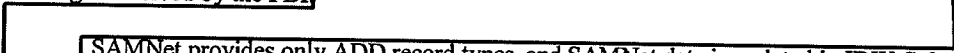
- 
-
-
-
-
-
-

b2
b7E



3.5.3 SAMNet

The Secure Automated Messaging Network (SAMNet) data source consists of cable traffic messages received by the FBI



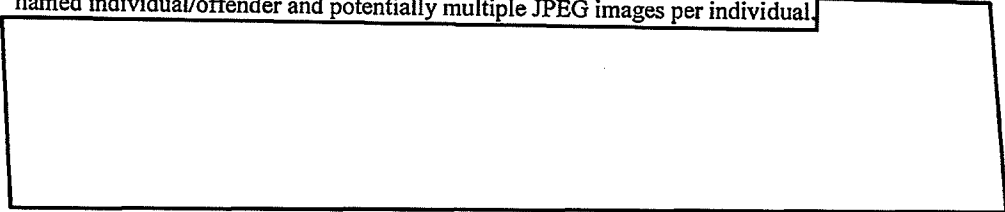
SAMNet provides only ADD record types, and SAMNet data is updated in IDW-S three times a day. SAMNet data is



b2
b7E

3.5.4 VGTOF

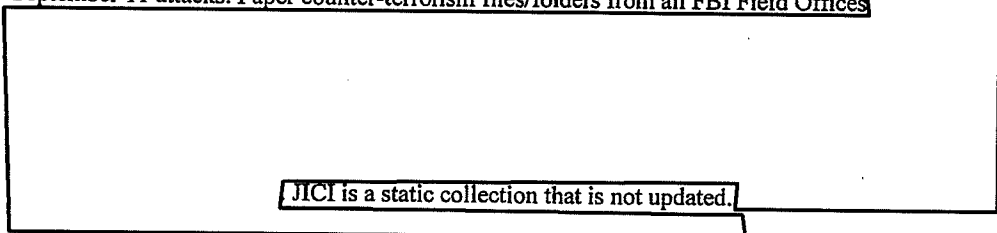
Violent Gang Terrorist Offender File (VGTOF) data is provided by the FBI National Crime Information Center (NCIC). VGTOF data includes two components: Data/metadata for each named individual/offender and potentially multiple JPEG images per individual.



b2
b7E

3.5.5 JICI

The Joint Intelligence Committee Investigation (JICI) data collection was created following the September 11 attacks. Paper counter-terrorism files/folders from all FBI Field Offices



JICI is a static collection that is not updated.

b2
b7E

3.5.6 Open Source News

The IDW-S V1.0 system contains an Open Source News library collected by the DARPA TIDES Program. These are primarily news source from around the world that are either in English or have been translated into English. Open Source News data is received in the form of text files (one per news article) and [redacted]. The Open Source News data goes into IDW-S once a day. Open Source News data [redacted]

b2
b7E

Open Source News material is derived from the following sources:

- Addis Ababa Tribune - <http://www.addistribune.com/>
- Agencia Brasilia - <http://www.radiobras.gov.br/>
- Al-Ahram (Egypt, weekly English version) - <http://weekly.ahram.org.eg/>
- AllAfrica.com - <http://allafrica.com/>
- Arabic News - <http://www.arabicnews.com/ansub/>
- Asahi Shimbun - <http://www.asahi.com/english/>
- Asia Times (Hong Kong) - <http://www.atimes.com/>
- Bangkok Post - <http://www.bangkokpost.net/>
- Christian Science Monitor - <http://www.csmonitor.com/>

Crescent International - <http://www.muslimedia.com/mainpage.htm>
Daily Telegraph (London, England) - <http://news.telegraph.co.uk/>
Dawn (Karachi, Pakistan) - <http://www.dawn.com/>
Debka (Israel) - <http://www.debka.com>
East Africa Daily Nation - www.nationaudio.com/News/DailyNation/Today/
Gulf News (UAE) - <http://www.gulf-news.com/>
Ha'aretz (Israel) - <http://www.haaretzdaily.com/>
IFRC International Federation of the Red Cross - <http://www.ifrc.org/>
IRIN Integrated Regional Information Network - <http://www.irinnews.org/>
Iraq Press News Agency - <http://www.iraqpress.org/>
Islamic Republic News Agency (Iran) - <http://www.irna.com/en>
Jakarta Post - <http://www.thejakartapost.com/>
Janes - <http://www.janes.com/>
Jordan Times - <http://www.jordantimes.com/>
L'Osservatore Romano - www.vatican.va/news_services/or/or_eng/text.html
Lagos (Nigeria) Guardian - <http://www.guardiannewsngr.com/>
Lahore (Pakistan) Nation - <http://www.nation.com.pk/>
Lebanon Daily Star (Beirut) - <http://www.dailystar.com.lb/>
Malaysian Star - <http://thestar.com.my/>
Manila Bulletin - <http://www.mb.com.ph/>
Manila Times - <http://www.manilatimes.net/>
Miami Herald - <http://www.miami.com/>
Moscow Times - <http://www.themoscowtimes.com/>
National Post and CP - <http://www.canada.com/>
New Straits Times (Kuala Lumpur) - http://www.emedia.com.my/Current_News/NST/
PETRA (Jordanian News Agency) - <http://www.petra.gov.jo>
Pakistan Observer (Islamabad) - <http://pakobserver.net/>
Palestine Chronicle - <http://palestinechronicle.com/>
People's Daily (China) - <http://english.peopledaily.com.cn/>
Philippine Star - <http://www.philstar.com/philstar/>
Pravda - <http://english.pravda.ru/>
ProMed - epidemiology mailing list
Russian Information Agency Novosti - <http://en.rian.ru/>
Russian Issues (Misc. Russian news) - <http://www.therussianissues.com/>
Russian Observer - <http://www.russianobserver.com/>
SABA (The News Agency of Yemen) - <http://www.sabanews.gov.ye>
Saudi Gazette - <http://www.saudigazette.com.sa/sgazette/>
South African Dispatch - <http://www.dispatch.co.za/>
Sydney Morning Herald - <http://www.smh.com.au>
Tehran Times - <http://www.tehrantimes.com/>
Times of India - <http://timesofindia.indiatimes.com/cms.dll>
UNHCR UN High Commissioner on Refugees - <http://www.unhcr.ch/>

Ummah News - <http://www.ummahnews.com/>
 Uzbekistan Report - <http://www.uzreport.com/eng/>
 Washington Post - <http://www.washingtonpost.com/>
 XinHua News Service - <http://www.xinhuanet.com/english/>
 Yemen Times (weekly) - <http://www.yementimes.com/>
 Yomiuri Shimbun (Japan) - <http://www.yomiuri.co.jp/>

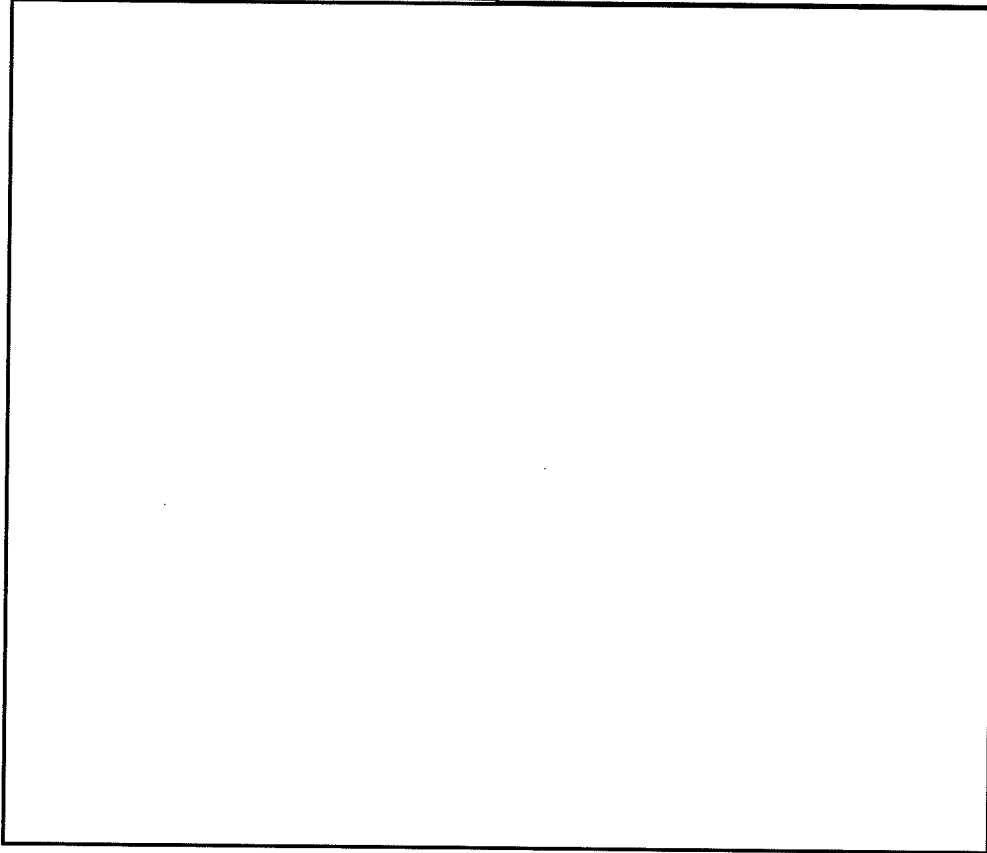
3.5.7 Summary of Data Sources

DATASET NAME	DATA SOURCE	TRANSFER METHOD
ACS	FBI Automated Case System (ACS)	FTP
IntelPlus	FBI Intel Plus	FTP
SAMNet	FBI Secure Automated Messaging Network (SAMNet)	FTP
VGTOF	FBI National Crime Information Center (NCIC)	CD-ROM
JICI	FBI Records Management Division (RMD), Document Laboratory (DocLab), FBIHQ	Multiple

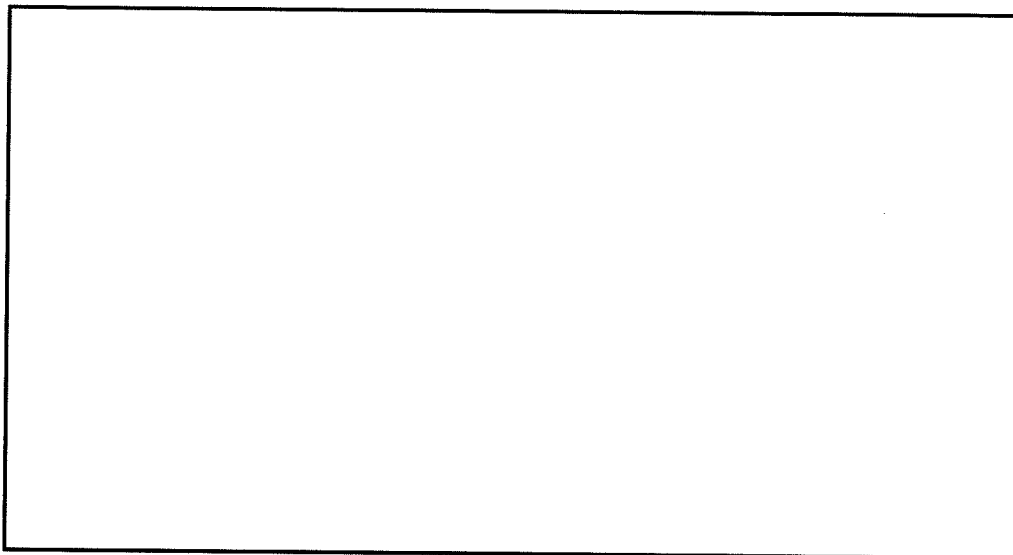
b2
b7E

DATASET NAME	DATA SOURCE	TRANSFER METHOD	
Open Source News	Translingual Information Detection, Extraction and Summarization (TIDES) Program, Defense Advanced Research Projects Agency (DARPA)	CD-ROM	

b2
b7E



Outside the Scope



3.9.9 Indirect Connections

Outside the Scope

IDW-S connects to the following sources of data:

SYSTEM NAME	CLASSIFICATION & COMPARTMENTS	ACCREDITED BY	TRANSFER METHOD
FBI Automated Case System (ACS)	Secret	FBI	FTP
FBI Intel Plus	Secret	FBI	FTP

b2
 b7E

Investigative Data Warehouse (IDW) - Secret
 System Security Plan
 5 December 2003
 Version 0.7

SYSTEM NAME	CLASSIFICATION & COMPARTMENTS	ACCREDITED BY	TRANSFER METHOD	COMMENTS
FBI Secure Automated Messaging Network (SAMNet)	Secret	FBI	FTP	
FBI National Crime Information Center (NCIC)	Unclassified/SBU	FBI	CD-ROM	
FBI Document Conversion Laboratory (DocLab), Records Management Division (RMD), FBIHQ	Secret	FBI	Multiple	
Translingual Information Detection, Extraction and Summarization (TIDES) Program, Defense Advanced Research Projects Agency (DARPA)	Unclassified	NA	CD-ROM	

b2
 b7E

Investigative Data Warehouse (IDW) - Secret
System Security Plan
5 December 2003
Version 0.7

SYSTEM NAME	CLASSIFICATION & COMPARTMENTS	ACCREDITED BY	TRANSFER METHOD	COMMENTS
	Unclassified	NA	CD-ROM	

Indirect Connections

b2
b7E

System Security Plan

Investigative Data Warehouse-SECRET (IDW-S)

08 January 2004

Version: 0.9

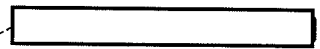
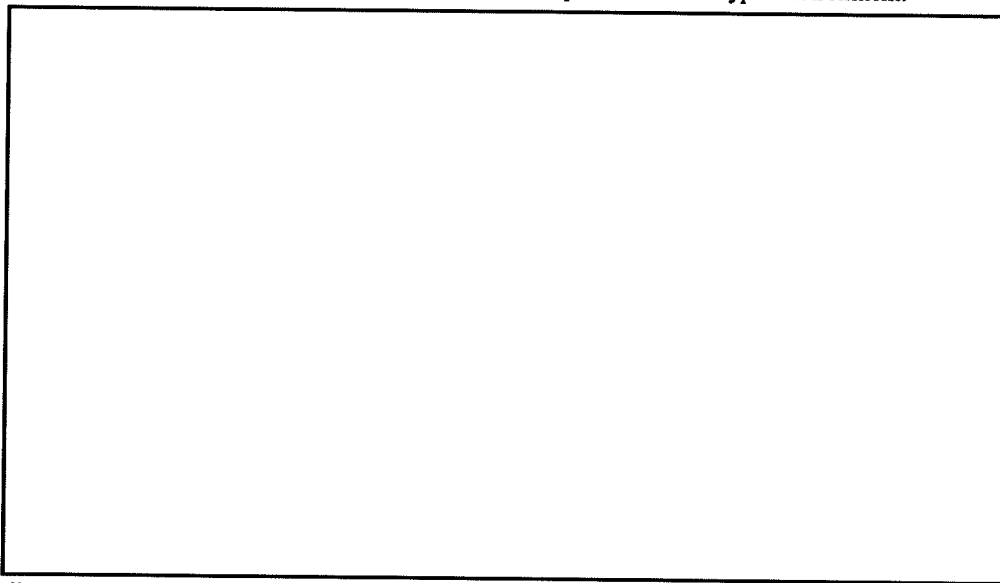
ALL INFORMATION CONTAINED
HEREIN IS UNCLASSIFIED
DATE 01-07-2008 BY 65179 DMH/BJA/CAL

**Federal Bureau of Investigation (FBI),
Information Resources Division (IRD)**

Investigative Data Warehouse (IDW) - Secret
System Security Plan

INTRODUCTION

The Federal Bureau of Investigation's (FBI) Investigative Data Warehouse Secret System (IDW-S) is an initial data warehouse, content management and data mining system that will permit FBI investigative, analytical, administrative and intelligence personnel to access aggregated data previously only available through individual applications. The IDW-S system will be authorized to process classified national security data up to, and including, Secret. The IDW-S system is the successor to the Secure Collaboration Operational Prototype Environment (SCOPE), which originally was named the Secure Counter-Terrorism Operational Prototype Environment.

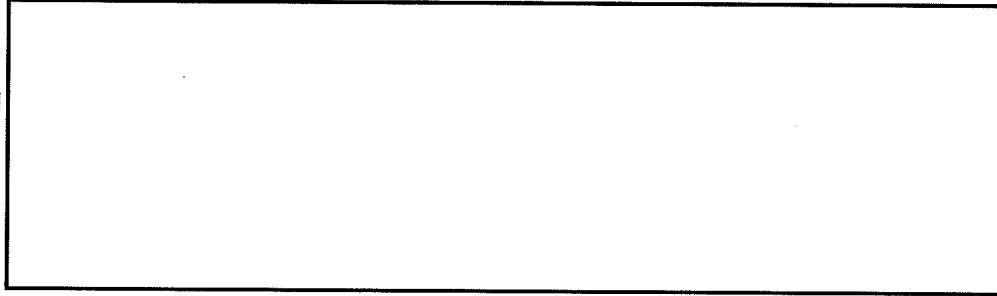


Data processed by the system will include the following data sets:

Outside the Scope

- Approved case files from the FBI's Automated Case Support (ACS) case management system;
- Electronic versions of the Joint Intelligence Committee Investigation (JICI) defined archived documents;

- Secure Automated Messaging Network (SAMNet) message traffic;
- IntelPlus File Rooms; and
- Violent Gang and Terrorist Organization File (VGTOF™) from the Criminal Justice Information Systems (CJIS) Division;
- Defense Advanced Research Projects Agency (DARPA) Translingual Information Detection, Extraction and Summarization (TIDES) Open Source Data



Outside the Scope



3.5. DATA PROCESSED

IDW-S processes data from the following data sources, with the security classification of the source indicated in brackets:

- Automated Case Support (ACS) System, Electronic Case File (ECF) Subsystem [Secret and below]
- IntelPlus [Secret and below]
- Secure Automated Message Network – Secret (SAMNet-S) [Secret and below]
- NCIC Violent Gang and Terrorist Offender File (VGTOF) [Sensitive But Unclassified]
- Joint Intelligence Committee Investigation (JICI) [Secret and below]
- DARPA/TIDES Open Source News [Unclassified]



Outside the Scope



<u>DATASET NAME</u>	<u>DATA SOURCE</u>	<u>TRANSFER METHOD</u>	
<u>ACS Electronic Case Files (ECF)</u>	<u>FBI Automated Case System (ACS)</u>	<u>FTP</u>	
<u>IntelPlus Filerooms</u>	<u>FBI IntelPlus</u>	<u>FTP</u>	



b2
b7E

<u>DATASET NAME</u>	<u>DATA SOURCE</u>	<u>TRANSFER METHOD</u>
<u>SAMNet</u>	<u>FBI Secure Automated Messaging Network (SAMNet)</u>	<u>FTP</u>
<u>VGTOF</u>	<u>FBI National Crime Information Center (NCIC)</u>	<u>CD-ROM</u>
<u>JICI</u>	<u>FBI Records Management Division (RMD), Document Laboratory (DocLab), FBIHQ</u>	<u>Multiple</u>
<u>Open Source News</u>	<u>Translingual Information Detection, Extraction and Summarization (TIDES) Program</u>	<u>CD-ROM</u>

Data Sources w/Transfer Mechanism

[Redacted]

[Redacted]

b2
 b7E
 Outside the Scope

[Redacted]

3.5.1 ACS ECF

IDW-S contains a subset of the ECF (Electronic Case File) subsystem of the Automated Case System (ACS). This subset consists of serials in those case classifications/sub classifications that have been officially sanctioned for inclusion in IDW-S. For each such serial, the ECF data includes metadata and text. IDW-S is synchronized against the ECF system once a day, a process which consists of receiving and processing the previous day's increment of ADD, MOD, and DELETE records.

[Redacted]

3.5.2 IntelPlus

Outside the Scope

IDW-S currently contains several IntelPlus counter-terrorism (CT) Filerooms:

- [Redacted]
- [Redacted]
- [Redacted]
- [Redacted]

b2
 b7E

• [Redacted]
• [Redacted]

[Redacted]

b2
b7E

3.5.3 SAMNet

The Secure Automated Messaging Network (SAMNet) data source consists of cable traffic messages received by the FBI.

[Redacted]
[Redacted] SAMNet provides only ADD record types, and SAMNet data is updated in IDW-S three times a day. SAMNet [Redacted]

b2
b7E

3.5.4 VGTOF

Violent Gang Terrorist Offender File (VGTOF) data is provided by the FBI National Crime Information Center (NCIC). VGTOF data includes two components: Data/metadata for each named individual/offender and potentially multiple JPEG images per individual. The data/metadata for each individual [Redacted]

[Redacted]

b2
b7E

3.5.5 JICI

The Joint Intelligence Committee Investigation (JICI) data collection was created following the September 11 attacks. Paper counter-terrorism files/folders from all FBI Field Offices [Redacted]

[Redacted]

b2
b7E

[Redacted]

JICI is a static collection that is not updated.

[Redacted]

b2
b7E

3.5.6 Open Source News

The IDW-S V1.0 system contains an Open Source News library collected by the DARPA TIDES Program. These are primarily news source from around the world that are either in English or have been translated into English. Open Source News data is received in the form of text files (one per news article [Redacted]). The Open Source News data goes into IDW-S once a day. Open Source News data [Redacted]

b2
b7E

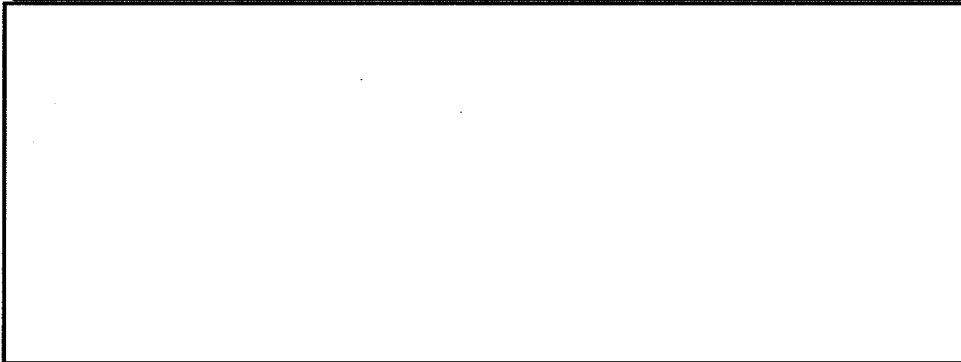
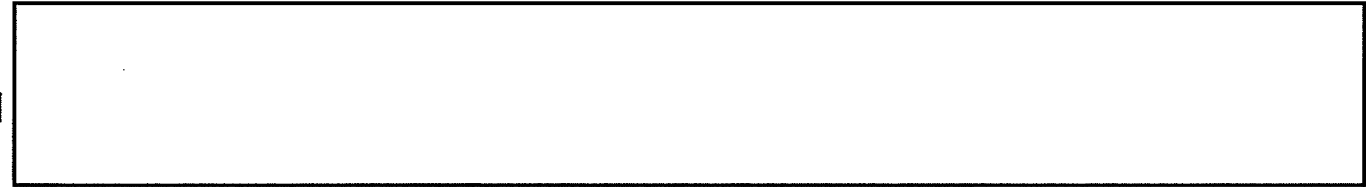
[Redacted]

[Redacted]

Open Source News material is derived from the following sources:

Addis Ababa Tribune - <http://www.addistribune.com/>
Agencia Brasilia - <http://www.radiobras.gov.br/>
Al-Ahram (Egypt, weekly English version) - <http://weekly.ahram.org.eg/>
AllAfrica.com - <http://allafrica.com/>
Arabic News - <http://www.arabicnews.com/ansub/>
Asahi Shimbun - <http://www.asahi.com/english/>
Asia Times (Hong Kong) - <http://www.atimes.com/>
Bangkok Post - <http://www.bangkokpost.net/>
Christian Science Monitor - <http://www.csmonitor.com/>
Crescent International - <http://www.muslimedia.com/mainpage.htm>
Daily Telegraph (London, England) - <http://news.telegraph.co.uk/>
Dawn (Karachi, Pakistan) - <http://www.dawn.com/>
Debka (Israel) - <http://www.debka.com>
East Africa Daily Nation - www.nationaudio.com/News/DailyNation/Today/
Gulf News (UAE) - <http://www.gulf-news.com/>
Ha'aretz (Israel) - <http://www.haaretzdaily.com/>
IFRC International Federation of the Red Cross - <http://www.ifrc.org/>
IRIN Integrated Regional Information Network - <http://www.irinnews.org/>
Iraq Press News Agency - <http://www.iraqpress.org/>
Islamic Republic News Agency (Iran) - <http://www.irna.com/en>
Jakarta Post - <http://www.thejakartapost.com/>
Janes - <http://www.janes.com/>
Jordan Times - <http://www.jordantimes.com/>
L'Osservatore Romano - www.vatican.va/news_services/or/or_eng/text.html
Lagos (Nigeria) Guardian - <http://www.guardiannewsngr.com/>
Lahore (Pakistan) Nation - <http://www.nation.com.pk/>
Lebanon Daily Star (Beirut) - <http://www.dailystar.com.lb/>
Malaysian Star - <http://thestar.com.my/>
Manila Bulletin - <http://www.mb.com.ph/>
Manila Times - <http://www.manilatimes.net/>
Miami Herald - <http://www.miami.com/>
Moscow Times - <http://www.themoscowtimes.com/>
National Post and CP - <http://www.canada.com/>
New Straits Times (Kuala Lumpur) - http://www.edia.com.my/Current_News/NST/
PETRA (Jordanian News Agency) - <http://www.petra.gov.jo>
Pakistan Observer (Islamabad) - <http://pakobserver.net/>
Palestine Chronicle - <http://palestinechronicle.com/>
People's Daily (China) - <http://english.peopledaily.com.cn/>
Philippine Star - <http://www.philstar.com/philstar/>
Pravda - <http://english.pravda.ru/>
ProMed - epidemiology mailing list
Russian Information Agency Novosti - <http://en.rian.ru/>
Russian Issues (Misc. Russian news) - <http://www.therussianissues.com/>
Russian Observer - <http://www.russianobserver.com/>

SABA (The News Agency of Yemen) - <http://www.sabanews.gov.ye>
Saudi Gazette - <http://www.saudigazette.com.sa/sgazette/>
South African Dispatch - <http://www.dispatch.co.za/>
Sydney Morning Herald - <http://www.smh.com.au>
Tehran Times - <http://www.tehrantimes.com/>
Times of India - <http://timesofindia.indiatimes.com/cms.dll>
UNHCR UN High Commissioner on Refugees - <http://www.unhcr.ch/>
Ummah News - <http://www.ummahnews.com/>
Uzbekistan Report - <http://www.uzreport.com/eng/>
Washington Post - <http://www.washingtonpost.com/>
XinHua News Service - <http://www.xinhuanet.com/english/>
Yemen Times (weekly) - <http://www.yementimes.com/>
Yomiuri Shimbun (Japan) - <http://www.yomiuri.co.jp/>



3.9.9 Indirect Connections

Outside the Scope

IDW-S connects to the following sources of data:

SYSTEM NAME	CLASSIFICATION & COMPARTMENTS	ACCREDITED BY	TRANSFER METHOD
FBI Automated Case System (ACS)	Secret	FBI	FTP
FBI Intel Plus	Secret	FBI	FTP

b2
b7E

Investigative Data Warehouse (IDW) - Secret
 System Security Plan
 08 January 04
 Version 0.9

SYSTEM NAME	CLASSIFICATION & COMPARTMENTS	ACCREDITED BY	TRANSFER METHOD	COMMENTS
FBI Secure Automated Messaging Network (SAMNet)	Secret	FBI	FTP	
FBI National Crime Information Center (NCIC)	Unclassified/SBU	FBI	CD-ROM	
FBI Document Conversion Laboratory (DocLab), Records Management Division (RMD), FBIHQ	Secret	FBI	CD/DVD	
Translingual Information Detection, Extraction and Summarization (TIDES) Program, Defense Advanced Research Projects Agency (DARPA)	Unclassified	NA	CD-ROM	

[Redacted]

[Redacted]

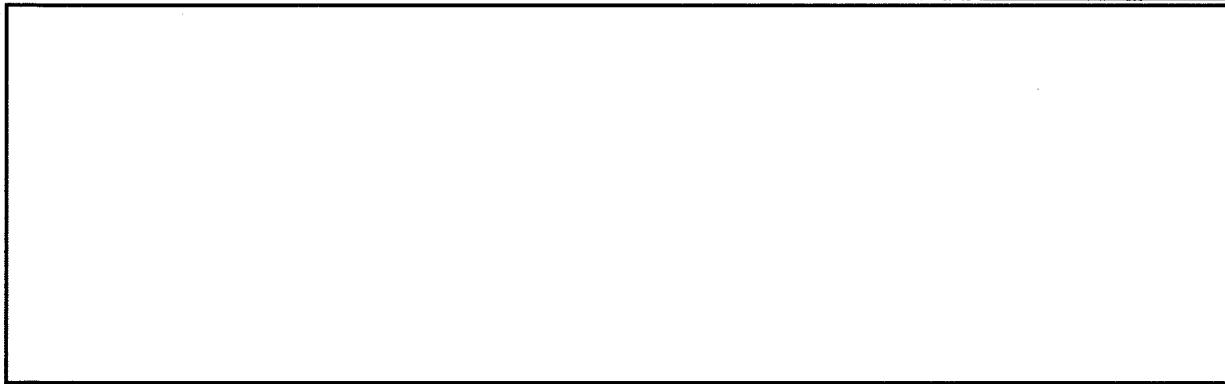
b2
 b7E
 Outside the Scope

Investigative Data Warehouse (IDW) - Secret
System Security Plan
08 January 04
Version 0.9

SYSTEM NAME	CLASSIFICATION & COMPARTMENTS	ACCREDITED BY	TRANSFER METHOD	COMMENTS
	Unclassified	NA	CD-ROM	

Indirect Connections

b2
b7E



Outside the Scope

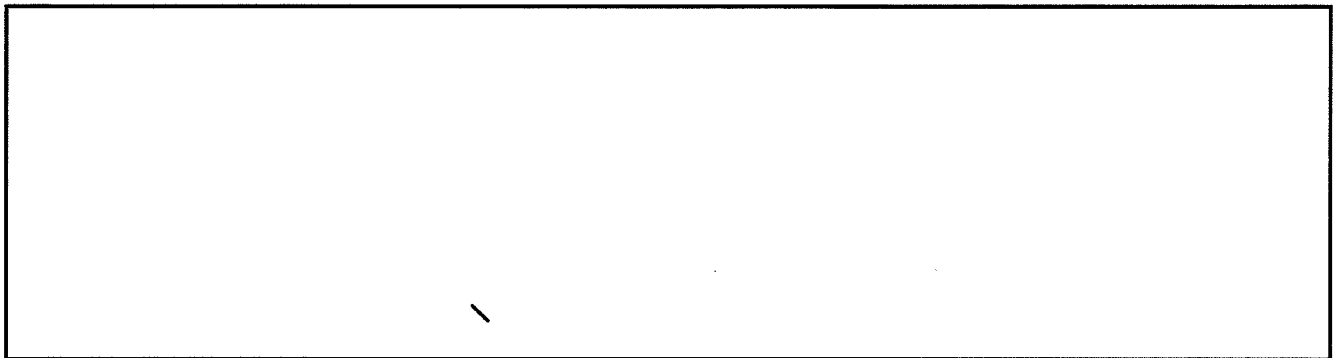
1.2.2 Supported Projects

IDW-S is an FBI enterprise program/system and utilizes data feeds from several FBI (and other agency) programs.

PROJECT NAME	CLASSIFICATION/ COMPARTMENTS	
Automated Case System (ACS)	Secret	
Intel Plus	Secret	
Secure Automated Messaging Network (SAMNet)	Secret	
National Crime Information Center (NCIC)	Sensitive But Unclassified (SBU)	
Document Conversion Laboratory (DocLab)	Secret	

b2
b7E

PROJECT NAME	CLASSIFICATION / COMPARTMENTS	PROJECT POC
Translingual Information Detection, Extraction and Summarization (TIDES) Program	Unclassified	Translingual Information Detection, Extraction and Summarization (TIDES) Program San Diego State University San Diego, CA



Outside the Scope

UNCLASSIFIED // FOR OFFICIAL USE ONLY

Investigative Data Warehouse (IDW) - Secret
System Security Plan
16 January 04
Version 1.0

System Security Plan

Investigative Data Warehouse-SECRET (IDW-S)

ALL INFORMATION CONTAINED
HEREIN IS UNCLASSIFIED
DATE 01-07-2008 BY 65179 DMH/BJA/CAL

16 January 2004

Version: 1.0

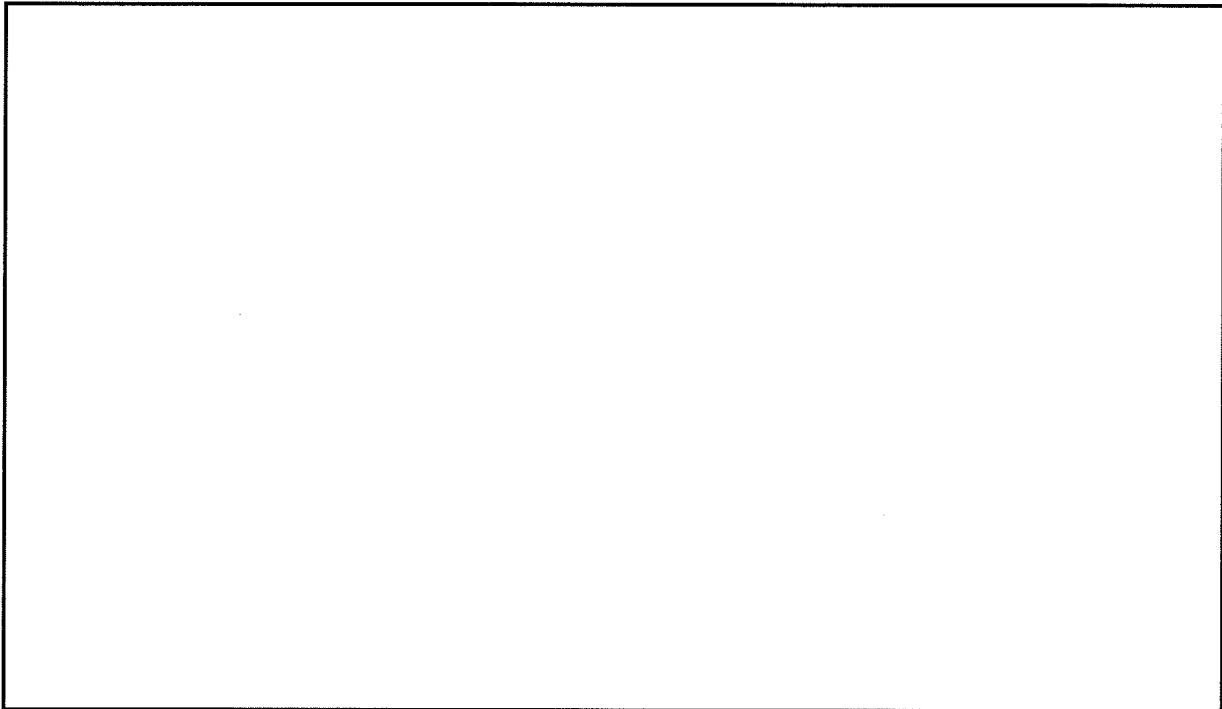
**Federal Bureau of Investigation (FBI),
Information Resources Division (IRD)**

UNCLASSIFIED // FOR OFFICIAL USE ONLY

Investigative Data Warehouse (IDW) - Secret
System Security Plan

INTRODUCTION

The Federal Bureau of Investigation's (FBI) Investigative Data Warehouse Secret System (IDW-S) is an initial data warehouse, content management and data mining system that will permit FBI investigative, analytical, administrative and intelligence personnel to access aggregated data previously only available through individual applications. The IDW-S system will be authorized to process classified national security data up to, and including, Secret. The IDW-S system is the successor to the Secure Collaboration Operational Prototype Environment (SCOPE), which originally was named the Secure Counter-Terrorism Operational Prototype Environment.



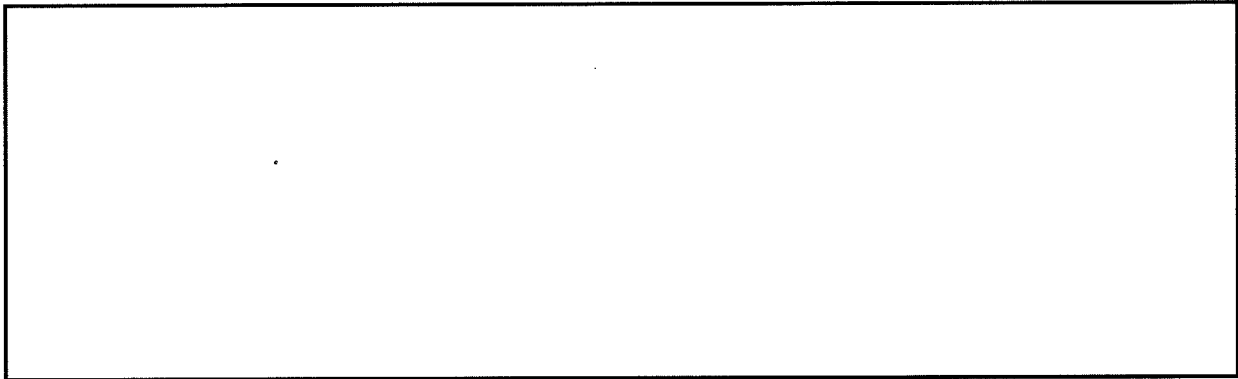
Data processed by the system will include the following data sets:

Outside the Scope

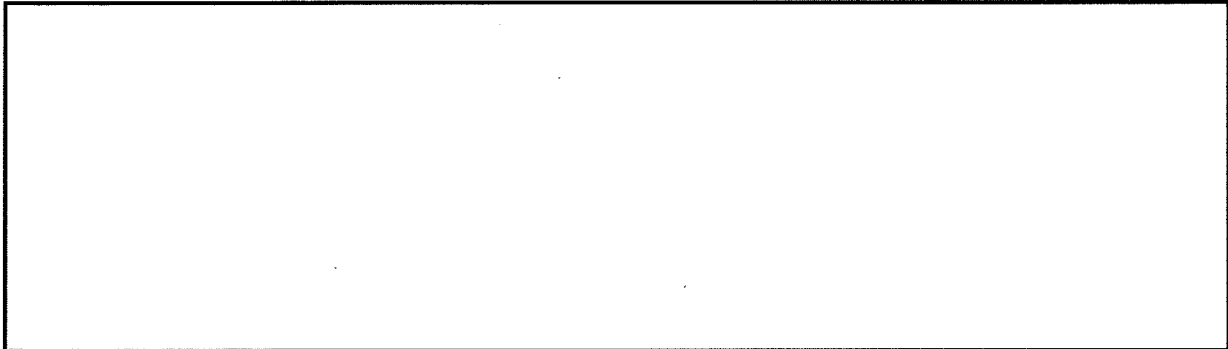
- Approved case files from the FBI's Automated Case Support (ACS) case management system;
- Electronic versions of the Joint Intelligence Committee Investigation (JICI) defined archived documents;

Investigative Data Warehouse (IDW) - Secret
System Security Plan
16 January 04
Version 1.0

- Secure Automated Messaging Network (SAMNet) message traffic;
- IntelPlus File Rooms; and
- Violent Gang and Terrorist Organization File (VGTOF) from the Criminal Justice Information Systems (CJIS) Division;
- Defense Advanced Research Projects Agency (DARPA) Translingual Information Detection, Extraction and Summarization (TIDES) Open Source Data



Outside the Scope



Outside the Scope

1.2.2 Supported Projects

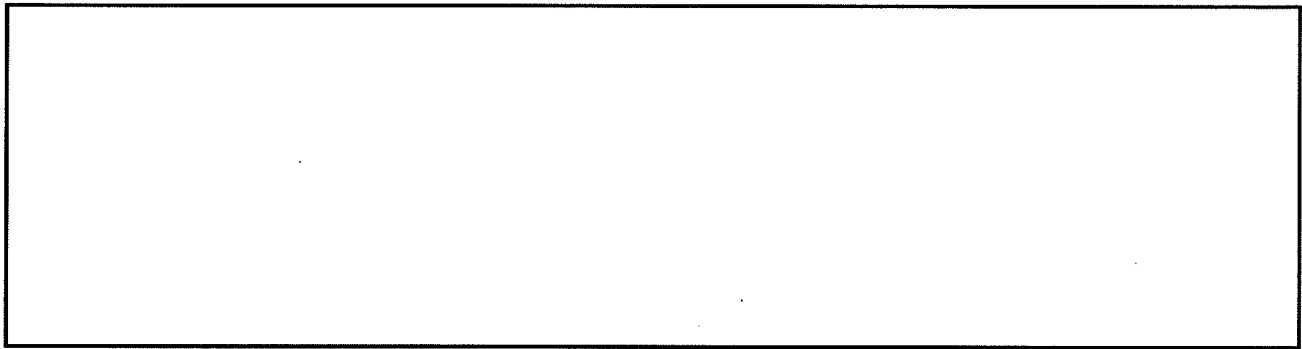
IDW-S is an FBI enterprise program/system and utilizes data feeds from several FBI (and other agency) programs.

PROJECT NAME	CLASSIFICATION / COMPARTMENTS	
Automated Case System (ACS)	Secret	
Intel Plus	Secret	
Secure Automated Messaging Network (SAMNet)	Secret	
National Crime Information Center (NCIC)	Sensitive But Unclassified (SBU)	
Document Conversion Laboratory (DocLab)	Secret	

b2
b7E

Investigative Data Warehouse (IDW) - Secret
System Security Plan
16 January 04
Version 1.0

PROJECT NAME	CLASSIFICATION / COMPARTMENTS	
Translingual Information Detection, Extraction and Summarization (TIDES) Program	Unclassified	

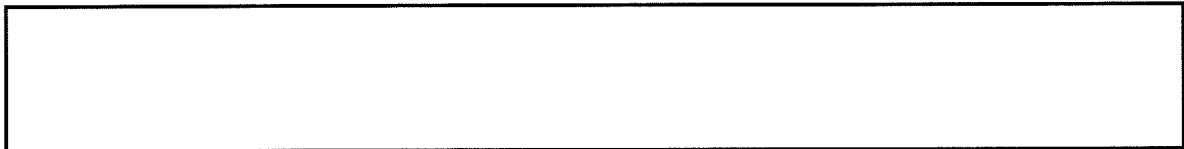


b2
b7E
Outside the Scope

3.5. DATA PROCESSED

IDW-S processes data from the following data sources, with the security classification of the source indicated in brackets:

- Automated Case Support (ACS) System, Electronic Case File (ECF) Subsystem [Secret and below]
- IntelPlus [Secret and below]
- Secure Automated Message Network – Secret (SAMNet-S) [Secret and below]
- NCIC Violent Gang and Terrorist Offender File (VGTOF) [Sensitive But Unclassified]
- Joint Intelligence Committee Investigation (JICI) [Secret and below]
- DARPA/TIDES Open Source News [Unclassified] Outside the Scope



DATASET NAME	DATA SOURCE	TRANSFER METHOD
ACS Electronic Case Files (ECF)	FBI Automated Case System (ACS)	FTP
IntelPlus Filerooms	FBI IntelPlus	FTP
SAMNet	FBI Secure Automated Messaging Network (SAMNet)	FTP
VGTOF	FBI National Crime Information Center (NCIC)	CD-ROM

b2
b7E

Investigative Data Warehouse (IDW) - Secret
System Security Plan
16 January 04
Version 1.0

DATASET NAME	DATA SOURCE	TRANSFER METHOD	
JICI	FBI Records Management Division (RMD), Document Laboratory (DocLab), FBIHQ	Multiple	
Open Source News	Translingual Information Detection, Extraction and Summarization (TIDES) Program	CD-ROM	

b2
b7E

Data Sources w/Transfer Mechanism

3.5.1 ACS ECF

IDW-S contains a subset of the ECF (Electronic Case File) subsystem of the Automated Case System (ACS). This subset consists of serials in those case classifications/sub classifications that have been officially sanctioned for inclusion in IDW-S. For each such serial, the ECF data includes metadata and text. IDW-S is synchronized against the ECF system once a day, a process which consists of receiving and processing the previous day's increment of ADD, MOD, and DELETE records



3.5.2 IntelPlus

Outside the Scope

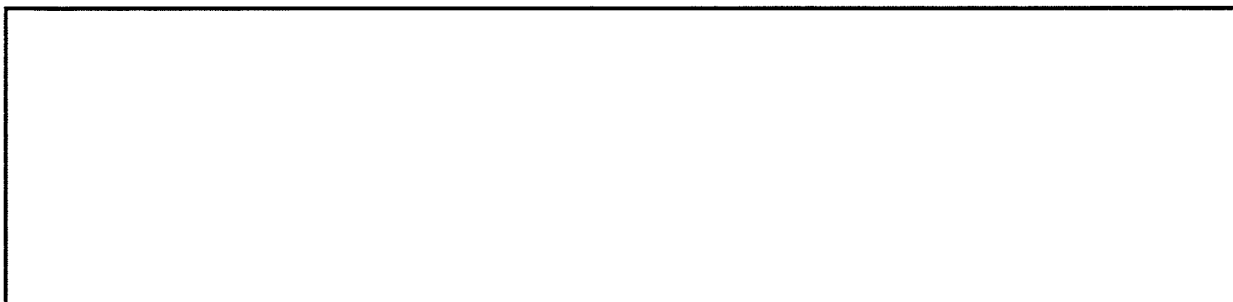
IDW-S currently contains several IntelPlus counter-terrorism (CT) Filerooms:

- [Redacted]
- [Redacted]
- [Redacted]
- [Redacted]
- [Redacted]
- [Redacted]
- [Redacted]

b2
b7E



b2
b7E



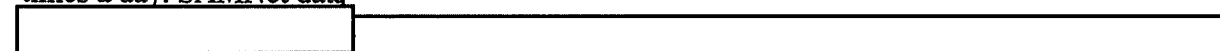
b2
b7E

3.5.3 SAMNet

The Secure Automated Messaging Network (SAMNet) data source consists of cable traffic messages received by the FBI.

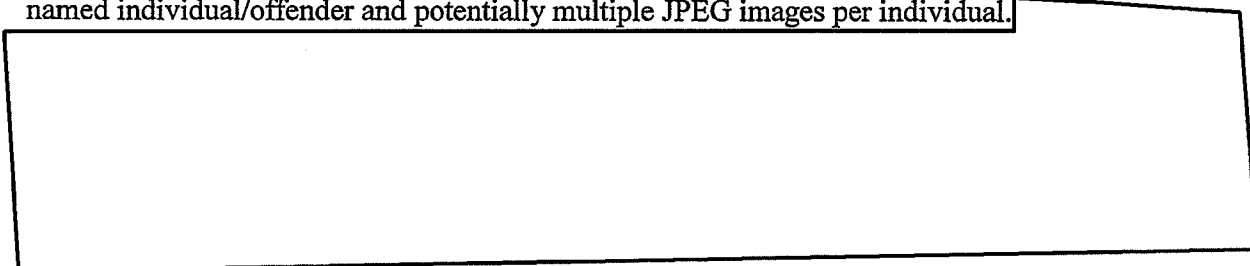


SAMNet provides only ADD record types, and SAMNet data is updated in IDW-S three times a day. SAMNet data



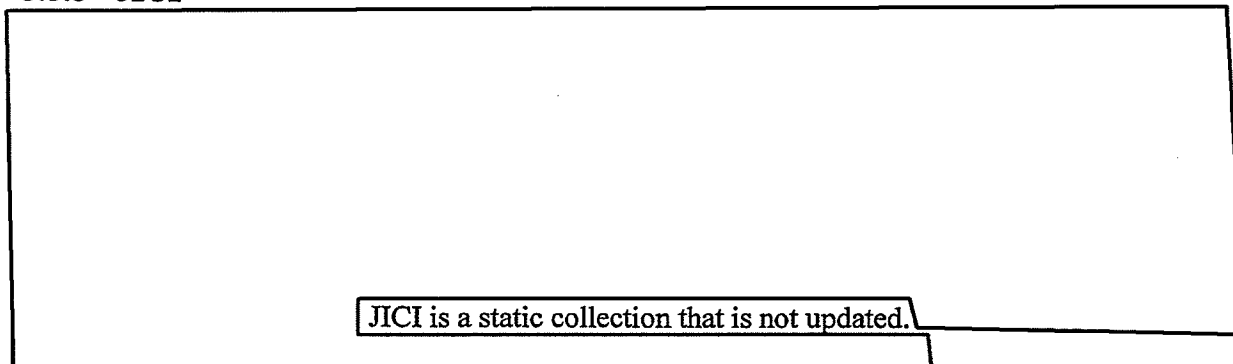
3.5.4 VGTOF

Violent Gang Terrorist Offender File (VGTOF) data is provided by the FBI National Crime Information Center (NCIC). VGTOF data includes two components: Data/metadata for each named individual/offender and potentially multiple JPEG images per individual.



b2
b7E

3.5.5 JICI



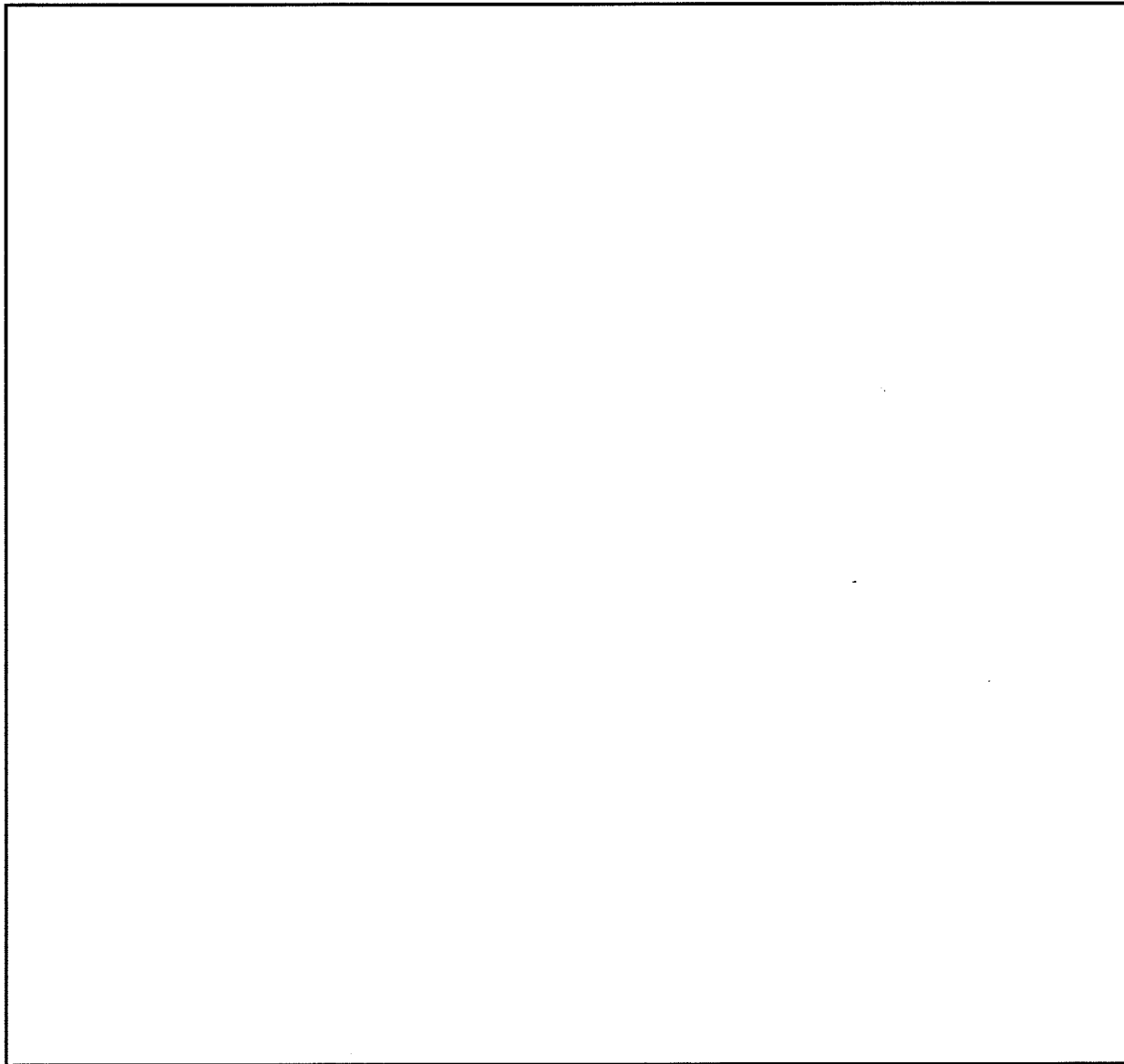
b2
b7E

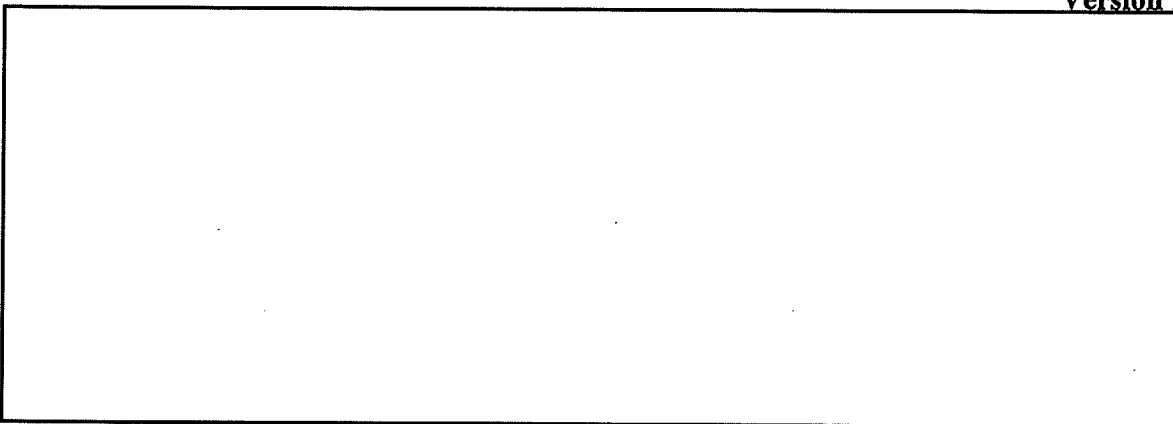
JICI is a static collection that is not updated.

3.5.6 Open Source News

The IDW-S V1.0 system contains an Open Source News library collected by the DARPA TIDES Program. These are primarily news source from around the world that are either in English or have been translated into English. Open Source News data is received in the form of text files (one per news article) [redacted] The Open Source News data goes into IDW-S once a day. Open Source News data is [redacted]

b2
b7E





3.9.9 Indirect Connections

Outside the Scope

IDW-S connects to the following sources of data:

SYSTEM NAME	CLASSIFICATION & COMPARTMENTS	ACCREDITED BY	TRANSFER METHOD
FBI Automated Case System (ACS)	Secret	FBI	FTP
FBI Intel Plus	Secret	FBI	FTP
FBI Secure Automated Messaging Network (SAMNet)	Secret	FBI	FTP
FBI National Crime Information Center (NCIC)	Unclassified/SBU	FBI	CD-ROM

b2
b7E

Investigative Data Warehouse (IDW) - Secret
 System Security Plan
 16 January 04
 Version 1.0

SYSTEM NAME	CLASSIFICATION & COMPARTMENTS	ACCREDITED BY	TRANSFER METHOD	COMMENTS
FBI Document Conversion Laboratory (DocLab), Records Management Division (RMD), FBIHQ	Secret	FBI	CD/DVD	
Translingual Information Detection, Extraction and Summarization (TIDES) Program, Defense Advanced Research Projects Agency (DARPA)	Unclassified	NA	CD-ROM	
	Unclassified	NA	CD-ROM	

b2
b7E

Indirect Connections

System Security Plan

Investigative Data Warehouse-SECRET (IDW-S)

ALL INFORMATION CONTAINED
HEREIN IS UNCLASSIFIED
DATE 01-07-2008 BY 65179 DMH/BJA/CAL

5 DECEMBER 2003

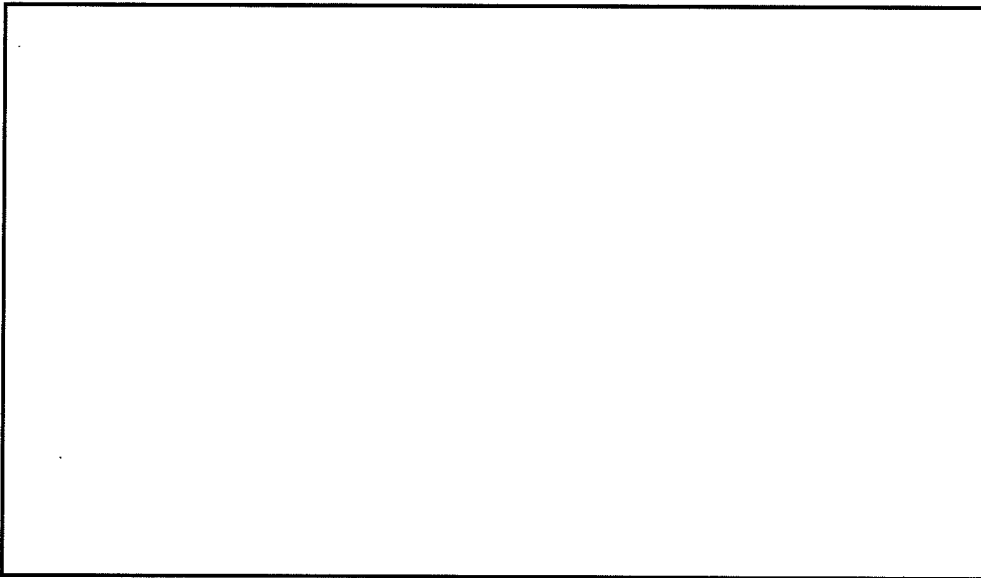
Version: 0.7

**Federal Bureau of Investigation (FBI),
Information Resources Division (IRD)**

Investigative Data Warehouse (IDW) - Secret
System Security Plan

INTRODUCTION

The Federal Bureau of Investigation's (FBI) Investigative Data Warehouse Secret System (IDW-S) is an initial data warehouse, content management and data mining system that will permit FBI investigative, analytical, administrative and intelligence personnel to access aggregated data previously only available through individual applications. The IDW-S system will be authorized to process classified national security data up to, and including, Secret. The IDW-S system is the successor to the Secure Collaboration Operational Prototype Environment (SCOPE), which originally was named the Secure Counter-Terrorism Operational Prototype Environment.



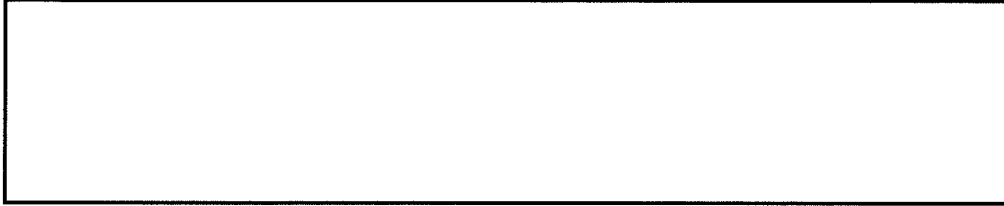
Data processed by the system will include the following data sets:

- Approved case files from the FBI's Automated Case Support (ACS) case management system;
- Electronic versions of the Joint Intelligence Committee Investigation (JICI) defined archived documents;



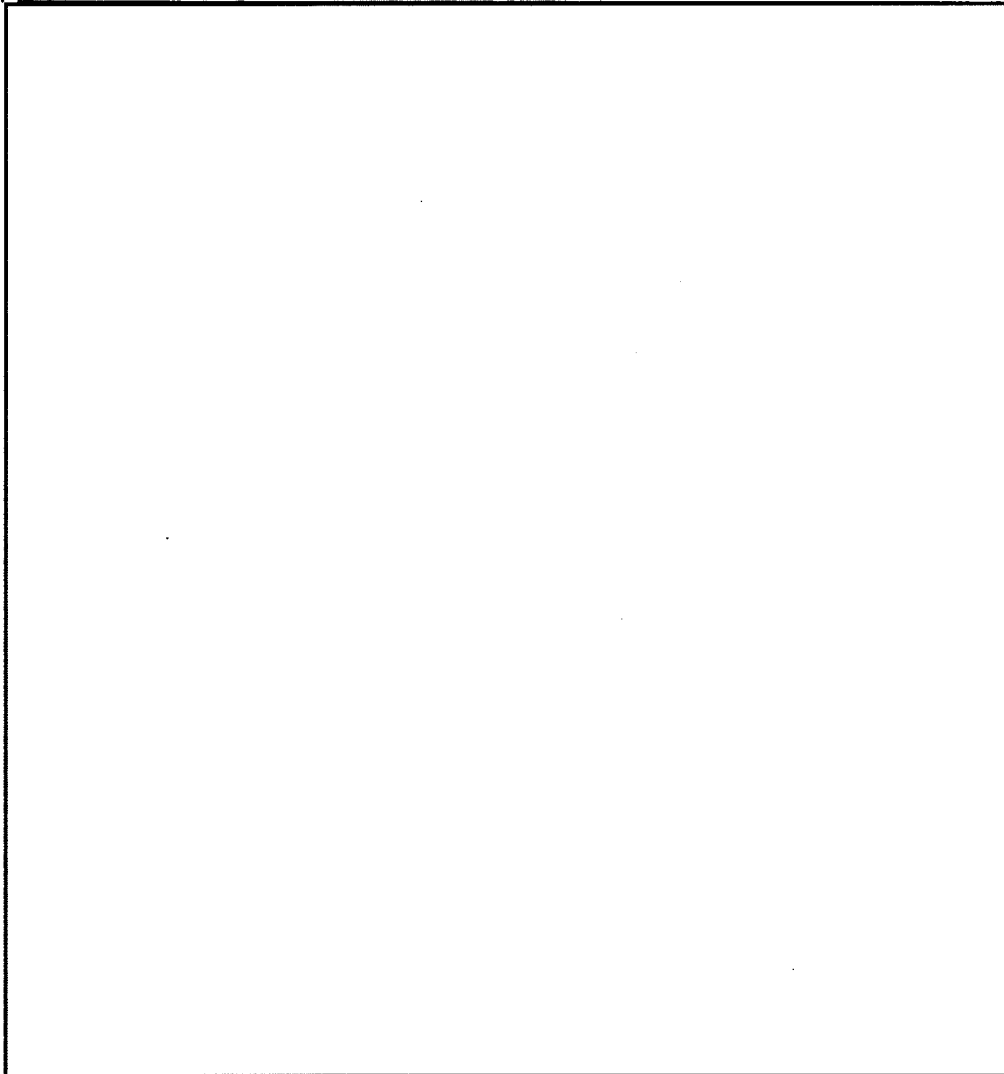
Outside the Scope

- Secure Automated Messaging Network (SAMNet) message traffic;
- IntelPlus File Rooms; and
- Violent Gang and Terrorist Organization File (VGTOF") from the Criminal Justice Information Systems (CJIS) Division;
- Defense Advanced Research Projects Agency (DARPA) Translingual Information Detection, Extraction and Summarization (TIDES) Open Source Data



Outside the Scope

Investigative Data Warehouse (IDW) - Secret
System Security Plan
19 December 2003
Version 0.8



Data Sources |

IDW-S contains data from the following data sources with security classification indicated in brackets:



Outside the Scope


- Automated Case Support (ACS) System, Electronic Case File (ECF) Subsystem [Secret and below]
- IntelPlus [Secret and below]
- Secure Automated Message Network – Secret (SAMNet-S) [Secret and below]
- NCIC Violent Gang and Terrorist Offender File (VGTOF) [Sensitive But Unclassified]
- Joint Intelligence Committee Investigation (JICI) [Secret and below]
- DARPA/TIDES Open Source News [Unclassified]

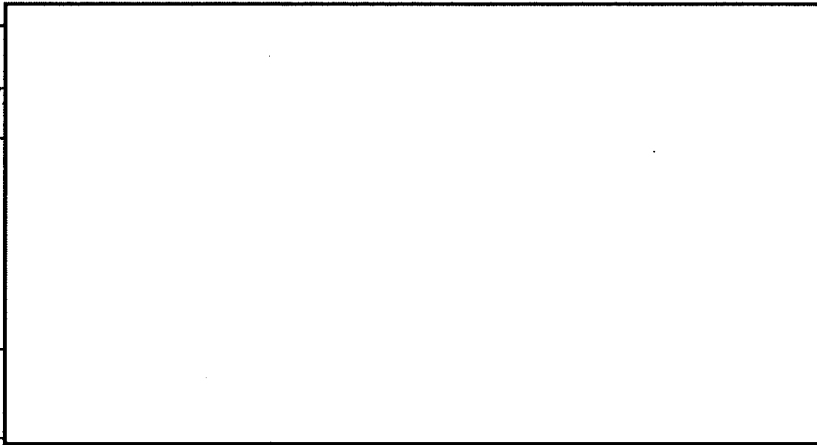
Data from each source is held in an identified source Library against which users can direct the analytical tools.

Data Ingest

IDW-S receives data from sources by FTP over FBINet (ACS ECF, IntelPlus, and SAMNet) and CD-ROM (VGTOF, Open Source News)

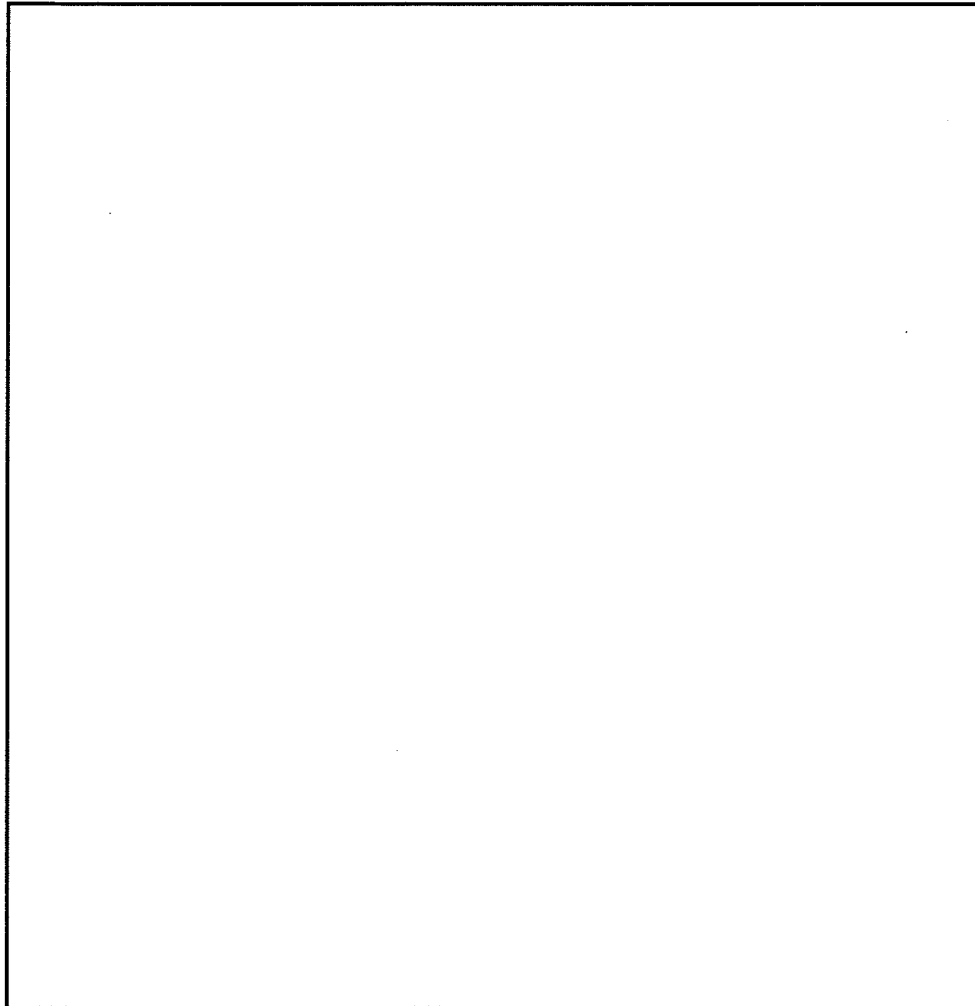


- Data is received by the Ingest Subsystem
- 
- Received data is parsed into individual serial records (in the case of ACS ECF), documents (IntelPlus, JICI, Open Source News), messages (SAMNet), or files (VGTOF).



Outside the Scope





Outside the Scope



3.5. DATA PROCESSED

IDW-S processes the following data:

3.5.1 ACS ECF

IDW-S contains a subset of the ECF (Electronic Case File) subsystem of the Automated Case System (ACS). This subset consists of serials in those case classifications/sub classifications that

have been officially sanctioned for inclusion in IDW-S. For each such serial, the ECF data includes metadata and text. IDW-S is synchronized against the ECF system once a day, a process which consists of receiving and processing the previous day's increment of ADD, MOD, and DELETE records.

3.5.2 IntelPlus

IDW-S currently contains several IntelPlus counter-terrorism (CT) Filerooms:

Outside the Scope

- [Redacted]
- [Redacted]
- [Redacted]
- [Redacted]
- [Redacted]
- [Redacted]

[Redacted]

b2
b7E

3.5.3 SAMNet

The Secure Automated Messaging Network (SAMNet) data source consists of cable traffic messages received by the FBI.

[Redacted]

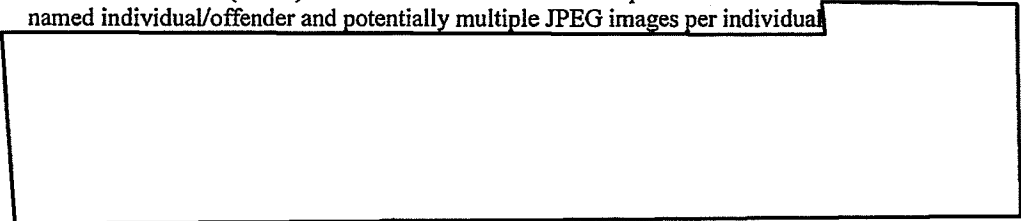
[Redacted] SAMNet provides only ADD record types, and SAMNet data is updated in IDW-S three times a day. SAMNet data is [Redacted]

[Redacted]

b2
b7E

3.5.4 VGTOF

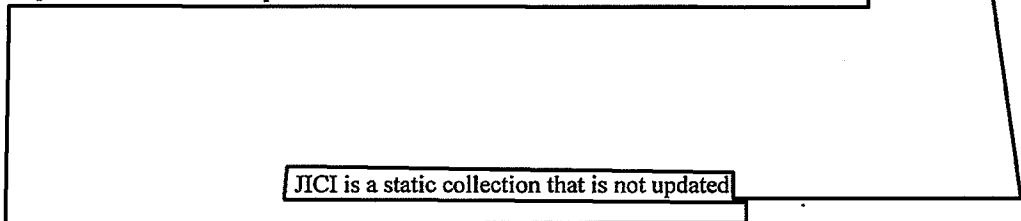
Violent Gang Terrorist Offender File (VGTOF) data is provided by the FBI National Crime Information Center (NCIC). VGTOF data includes two components: Data/metadata for each named individual/offender and potentially multiple JPEG images per individual



3.5.5 JICI

The Joint Intelligence Committee Investigation (JICI) data collection was created following the September 11 attacks. Paper counter-terrorism files/folders from all FBI Field Offices

b2
b7E



3.5.6 Open Source News

The IDW-S V1.0 system contains an Open Source News library collected by the DARPA TIDES Program. These are primarily news source from around the world that are either in English or have been translated into English. Open Source News data is received in the form of text files (one per news article) and [redacted]. The Open Source News data goes into IDW-S once a day. Open Source News data [redacted]

Open Source News material is derived from the following sources:

- Addis Ababa Tribune - <http://www.addistribune.com/>
- Agencia Brasilia - <http://www.radiobras.gov.br/>
- Al-Ahram (Egypt, weekly English version) - <http://weekly.ahram.org.eg/>
- AllAfrica.com - <http://allafrica.com/>
- Arabic News - <http://www.arabicnews.com/ansub/>
- Asahi Shimbun - <http://www.asahi.com/english/>
- Asia Times (Hong Kong) - <http://www.atimes.com/>
- Bangkok Post - <http://www.bangkokpost.net/>
- Christian Science Monitor - <http://www.csmonitor.com/>

Crescent International - <http://www.muslimedia.com/mainpage.htm>
Daily Telegraph (London, England) - <http://news.telegraph.co.uk/>
Dawn (Karachi, Pakistan) - <http://www.dawn.com/>
Debka (Israel) - <http://www.debka.com>
East Africa Daily Nation - www.nationaudio.com/News/DailyNation/Today/
Gulf News (UAE) - <http://www.gulf-news.com/>
Ha'aretz (Israel) - <http://www.haaretzdaily.com/>
IFRC International Federation of the Red Cross - <http://www.ifrc.org/>
IRIN Integrated Regional Information Network - <http://www.irinnews.org/>
Iraq Press News Agency - <http://www.iraqpress.org/>
Islamic Republic News Agency (Iran) - <http://www.irna.com/en>
Jakarta Post - <http://www.thejakartapost.com/>
Janes - <http://www.janes.com/>
Jordan Times - <http://www.jordantimes.com/>
L'Osservatore Romano - www.vatican.va/news_services/or/or_eng/text.html
Lagos (Nigeria) Guardian - <http://www.guardiannewsngr.com/>
Lahore (Pakistan) Nation - <http://www.nation.com.pk/>
Lebanon Daily Star (Beirut) - <http://www.dailystar.com.lb/>
Malaysian Star - <http://thestar.com.my/>
Manila Bulletin - <http://www.mb.com.ph/>
Manila Times - <http://www.manilatimes.net/>
Miami Herald - <http://www.miami.com/>
Moscow Times - <http://www.themoscowtimes.com/>
National Post and CP - <http://www.canada.com/>
New Straits Times (Kuala Lumpur) - http://www.imedia.com.my/Current_News/NST/
PETRA (Jordanian News Agency) - <http://www.petra.gov.jo>
Pakistan Observer (Islamabad) - <http://pakobserver.net/>
Palestine Chronicle - <http://palestinechronicle.com/>
People's Daily (China) - <http://english.peopledaily.com.cn/>
Philippine Star - <http://www.philstar.com/philstar/>
Pravda - <http://english.pravda.ru/>
ProMed - epidemiology mailing list
Russian Information Agency Novosti - <http://en.rian.ru/>
Russian Issues (Misc. Russian news) - <http://www.therussianissues.com/>
Russian Observer - <http://www.russianobserver.com/>
SABA (The News Agency of Yemen) - <http://www.sabanews.gov.ye>
Saudi Gazette - <http://www.saudi gazette.com.sa/sgazette/>
South African Dispatch - <http://www.dispatch.co.za/>
Sydney Morning Herald - <http://www.smh.com.au>
Tehran Times - <http://www.tehrantimes.com/>
Times of India - <http://timesofindia.indiatimes.com/cms.dll>
UNHCR UN High Commissioner on Refugees - <http://www.unhcr.ch/>

Ummah News - <http://www.ummahnews.com/>
 Uzbekistan Report - <http://www.uzreport.com/eng/>
 Washington Post - <http://www.washingtonpost.com/>
 XinHua News Service - <http://www.xinhuanet.com/english/>
 Yemen Times (weekly) - <http://www.yementimes.com/>
 Yomiuri Shimbun (Japan) - <http://www.yomiuri.co.jp/>

3.5.7 Summary of Data Sources

DATA SET NAME	DATA SOURCE	TRANSFER METHOD
ACS	FBI Automated Case System (ACS)	FTP
IntelPlus	FBI Intel Plus	FTP
SAMNet	FBI Secure Automated Messaging Network (SAMNet)	FTP
VGTOF	FBI National Crime Information Center (NCIC)	CD-ROM
JICI	FBI Records Management Division (RMD), Document Laboratory (DocLab), FBIHQ	Multiple



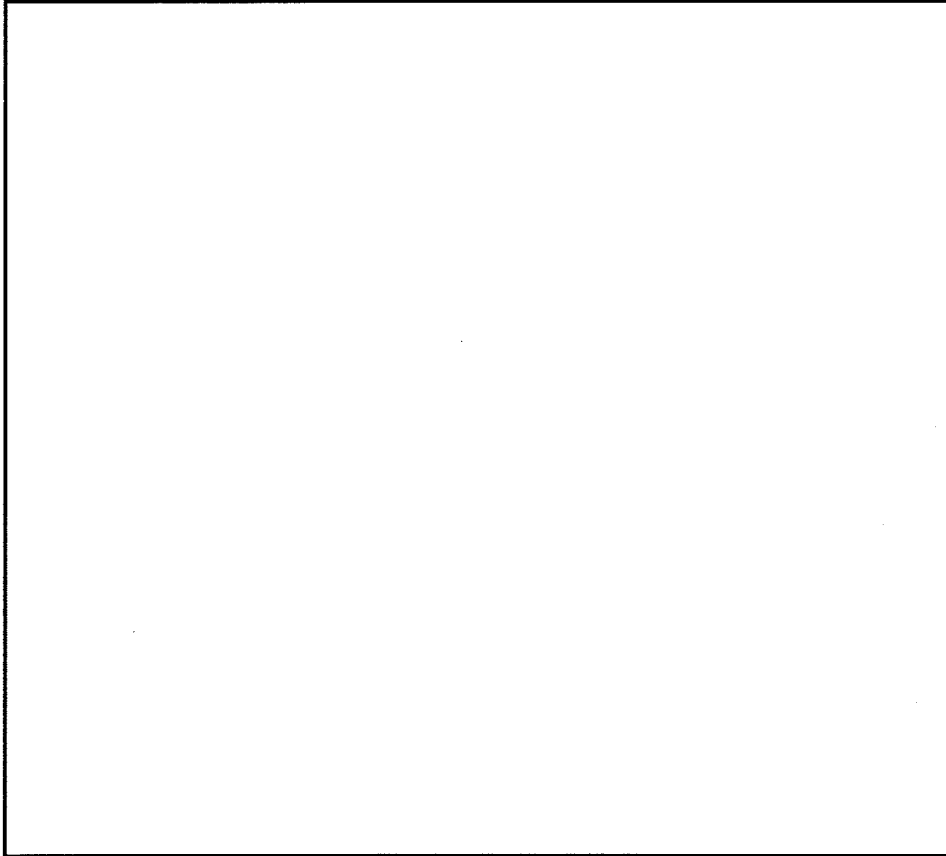
b2
b7E

Outside the Scope



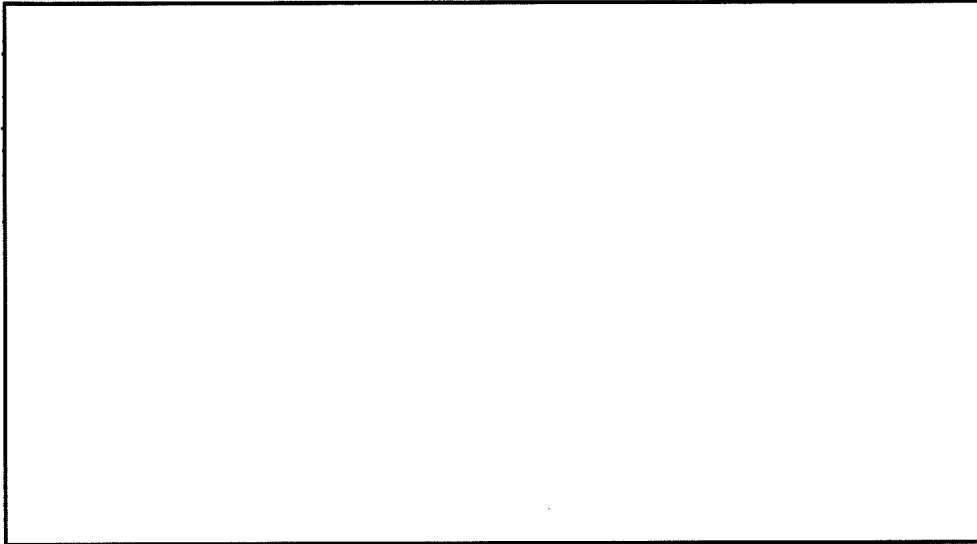
DATASET NAME	DATA SOURCE	TRANSFER METHOD	
Open Source News	Translingual Information Detection, Extraction and Summarization (TIDES) Program, Defense Advanced Research Projects Agency (DARPA)	CD-ROM	

b2
b7E



Outside the Scope





Outside the Scope

3.9.9 Indirect Connections

IDW-S connects to the following sources of data:

SYSTEM NAME	CLASSIFICATION & COMPARTMENTS	ACCREDITED BY	TRANSFER METHOD
FBI Automated Case System (ACS)	Secret	FBI	FTP
FBI Intel Plus	Secret	FBI	FTP

b2
b7E

Investigative Data Warehouse (IDW) - Secret
 System Security Plan
 19 December 2003
 Version 0.8

SYSTEM NAME	CLASSIFICATION & COMPARTMENTS	ACCREDITED BY	TRANSFER METHOD	COMMENTS
FBI Secure Automated Messaging Network (SAMNet)	Secret	FBI	FTP	
FBI National Crime Information Center (NCIC)	Unclassified/SBU	FBI	CD-ROM	
FBI Document Conversion Laboratory (DocLab), Records Management Division (RMD), FBIHQ	Secret	FBI	Multiple	
Translingual Information Detection, Extraction and Summarization (TIDES) Program, Defense Advanced Research Projects Agency (DARPA)	Unclassified	NA	CD-ROM	

[Redacted]

[Redacted]

b2
b7E

Outside the Scope

[Redacted]

SYSTEM NAME	CLASSIFICATION & COMPARTMENTS	ACCREDITED BY	TRANSFER METHOD	COMMENTS
	Unclassified	NA	CD-ROM	

Indirect Connections

b2
b7E

Federal Bureau of Investigation
Information Resources Division
Data Management Section



ALL INFORMATION CONTAINED
HEREIN IS UNCLASSIFIED
DATE 01-07-2008 BY 65179 DMH/BJA/CAL

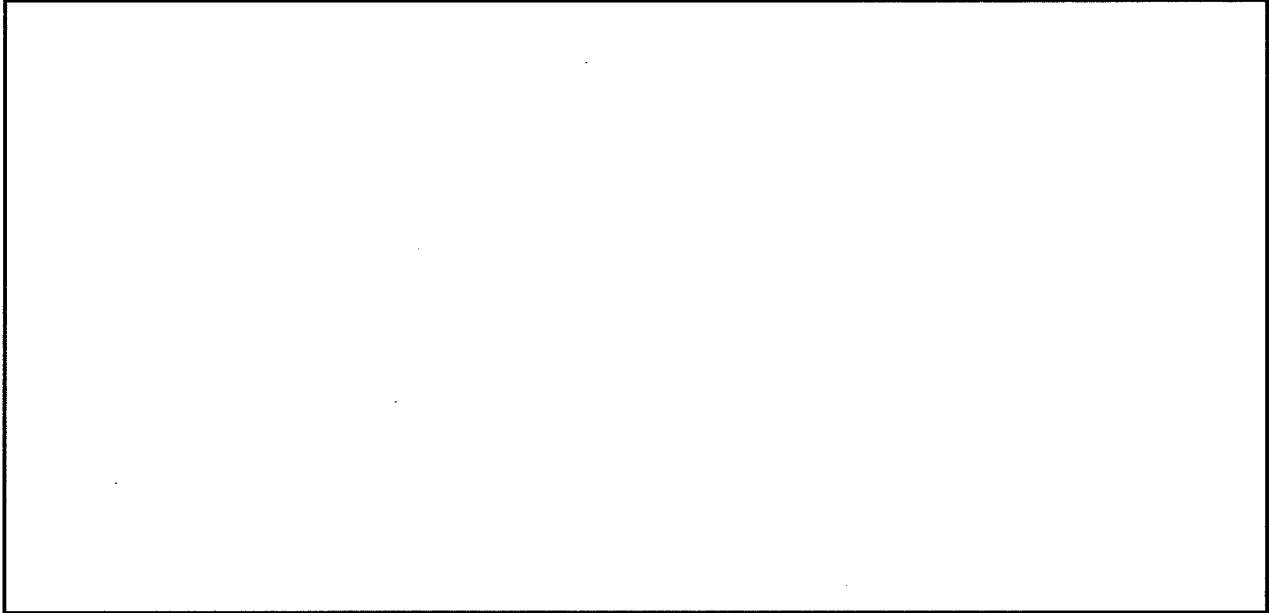
Investigative Data Warehouse
Maintenance Manual
Version 1.1 for Authority to Operate
June 11, 2004

Note: Hardcopy versions of this document must be verified for correct version number with the IDW O&M Document Control Library.

MM NG-OM-004.11ATO
FOR OFFICIAL USE ONLY

1 Introduction

The Federal Bureau of Investigation (FBI) Integrated Data Warehouse (IDW) is a data warehouse, content management, and data mining system that will permit FBI investigative, analytical, administrative, and intelligence personnel to access aggregated data previously available only through individual applications. The IDW-S system will be authorized to process classified national security data up to, and including, the Secret level.



Data processed by the system will include the following data sets:

Outside the Scope

- Approved case files from the FBI's Automated Case Support (ACS) case management system;
- Electronic versions of the Joint Intelligence Committee Investigation (JICI) defined archived documents;
- Secure Automated Messaging Network (SAMNet) message traffic;
- IntelPlus File Rooms;
- Violent Gang and Terrorist Organization File (VGTOF) from the Criminal Justice Information Systems (CJIS) Division; and
- Open Source Data

Note: Hardcopy versions of this document must be verified for correct version number with the IDW O&M Document Control Library.

MM NG-OM-004.11ATO
FOR OFFICIAL USE ONLY