Federal Aviation Administration

DOT/FAA/AM-07/7
Office of Aerospace Medicine
Washington, DC 20591

Voluntary Aviation Safety Information-Sharing Process: Preliminary Audit of Distributed FOQA and ASAP Archives Against Industry Statement of Requirements

Thomas R. Chidester
Civil Aerospace Medical Institute
Federal Aviation Administration
Oklahoma City, OK 73125

April 2007

Final Report

NOTICE

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
DOT/FAA/AM-07/7		

4. Title and Subtitle	5. Report Date
Voluntary Aviation Safety Information-Sharing Process: Preliminary Audit of Distributed FOQA and ASAP Archives Against Industry Statement of Requirements	April 2007
	6. Performing Organization Code
7. Author(s)	8. Performing Organization Report No.
Chidester TR	
9. Performing Organization Name and Address	10. Work Unit No. (TRAIS)
FAA Civil Aerospace Medical Institute P.O. Box 25082 Oklahoma City, OK 73125	11. Contract or Grant No.
12. Sponsoring Agency name and Address	13. Type of Report and Period Covered
Office of Aerospace Medicine Federal Aviation Administration 800 Independence Ave., S.W. Washington, DC 20591	14. Sponsoring Agency Code

15. Supplemental Notes

Work was accomplished under approved task AM-AHRR521

16. Abstract

The Voluntary Aviation Safety Information-Sharing Process (VASIP) is designed to provide a means for the commercial aviation industry and the Federal Aviation Administration (FAA) to collect, share safety-related information, and to use that information to proactively identify, analyze, and correct safety issues that affect commercial aviation. The key to VASIP is the development of a technical process to extract de-identified safety data from any participating airline Flight Operations Quality Assurance (FOQA) or Aviation Safety Action Program (ASAP), aggregate it through a distributed database, and make it accessible to appropriate industry stakeholders for analysis. In 2004, the ASAP and FOQA Aviation Rulemaking Committees (ARCs) identified the National Aeronautics and Space Administration (NASA) as having the institutional background, resources, and personnel capable of developing this technical aggregation framework, as well as the analytical tools to support the process. Beginning in June of 2004, NASA led a collaborative partnership of participating airlines, employee organizations, and FAA representatives to define key components of archives of FOQA and ASAP data. This defined a set of functional requirements for archive development that were approved by the FOQA and ASAP ARCs. In October 2004, at the request of and with partial funding by the FAA, NASA initiated an Information Sharing Initiative under the Aviation Safety and Security Program to provide funds and oversight to develop distributed archiving and analysis. The basic infrastructure was deployed in January 2006, and data archiving began at participating airlines. The current document audits the hardware, software, and networking infrastructure against the original functional specifications provided by the ARCs to NASA.

17. Key Words	18. Distribution Statement
Information Sharing, Flight Data, Safety Reports	Document is available to the public through the Defense Technical Information Center, Ft. Belvior, VA 22060; and the National Technical Information Service, Springfield, VA 22161

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	14	

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

LIST OF ACRONYMS USED IN THIS DOCUMENT

ARC ..... Aviation Rulemaking Committee
ASAP ..... Aviation Safety Action Program
CAST ..... Commercial Aviation Safety Team
DAWG ..... Data Aggregation Working Group
DNAA ..... Distributed National ASAP Archive
DNFA ..... Distributed National FOQA Archive
ESC ..... Executive Steering Committee
FAA ..... Federal Aviation Administration
FOQA ..... Flight Operations Quality Assurance
NASA ..... National Aeronautics and Space Administration
RPS ..... Rules and Procedures Subcommittee
VASIP ..... Voluntary Aviation Safety Information-Sharing Process
VSIS ..... Voluntary Safety Information Sharing

Voluntary Aviation Safety Information-Sharing Process: Preliminary Audit of Distributed FOQA and ASAP Archives Against Industry Statement of Requirements

I. INTRODUCTION

The Voluntary Aviation Safety Information-Sharing Process (VASIP) is designed to provide a means for the commercial aviation industry and the Federal Aviation Administration (FAA) to collect, share safety-related information, and to use that information to proactively identify, analyze, and correct safety issues that affect commercial aviation.

The key to VASIP is the development of a technical process to extract de-identified safety data from any participating airline Flight Operations Quality Assurance (FOQA) or Aviation Safety Action Program (ASAP), aggregate it through a distributed database, and make it accessible to appropriate industry stakeholders for analysis.

In 2004, the ASAP and FOQA Aviation Rulemaking Committees (ARCs) identified the National Aeronautics and Space Administration (NASA) as having the institutional background, resources, and personnel capable of developing this technical aggregation framework, as well as the analytical tools to support the process. Beginning in June of 2004, NASA led a collaborative partnership of participating airlines, employee organizations, and FAA representatives to define key components of archives of FOQA and ASAP data. This defined a set of functional requirements for archive development that were approved by the FOQA and ASAP ARCs. In October 2004, at the request of and with partial funding by the FAA, NASA initiated an Information Sharing Initiative under the Aviation Safety and Security Program to provide funds and oversight to develop distributed archiving and analysis. The basic infrastructure was deployed in January 2006, and data archiving began at participating airlines. In January 2006, the ASAP and FOQA ARCs were replaced by the Voluntary Safety Information Sharing (VSIS) ARC, which was tasked to oversee distributed archive operation and expansion, data analysis and reporting using the archives, and advocacy of solutions to problems and issues discovered and understood through the data sharing process.

The author was tasked under FAA Program Directive 081500001, Task FD-10, to review the functionality of the Distributed National FOQA Archive (DNFA) and Distributed National ASAP Archive (DNAA) relative to the specifications provided by the FOQA and ASAP ARCs. The current document audits the hardware, software, and networking infrastructure against the original functional specifications provided by the ARCs to NASA. Auditing was accomplished by monitoring NASA's functional testing and demonstration of archive hardware and software from November 2005 though April 2006, and during a site visit in May 2006 to review functions that had not been demonstrated at previous meetings.

A distinction must be drawn between an audit of functionalities specified by the ARCs and an audit against the System Requirements Document developed by Battelle and its subcontractors and presented to NASA in 2005. The former asks whether the system accomplishes the tasks requested by the ARC; the latter is a matter of formal verification that the system was built to its detailed specifications. This report is the former; NASA has undertaken the latter through its contracting and oversight processes.

II. TECHNICAL REQUIREMENTS

Distributed National FOQA Archive

Overview. The VASIP Executive Steering Committee (ESC), representatives from airline, union, FAA, and the FOQA and ASAP ARCs, asked NASA to develop and implement hardware, software, and networking for a distributed-concept archive of industry flight data and safety reports. The committee proposed that NASA and the FAA jointly fund a two-year demonstration project, implemented to a sufficient level of reliability that it may be continued by any selected operating organization.

For flight data, NASA was asked to implement a network of servers located on airline premises and access local servers via the network for aggregation, statistical analysis, and summarization in response to ESC requests. Figure 1 shows that local servers receive de-identified flight parameters at measured rates, pushed from each airline's FOQA-analysis machine to the local archive server. In response to ESC-approved requests, queries are sent from a central server and accomplished on each local server.

Figure 2 shows how analysis results are forwarded via the network to a central server at NASA Ames, and aggregate result summaries in electronic format returned to each airline's local server and forwarded to analysis working groups assigned by the ESC. Notice that data are "pushed" to the local archive server and information, in the form of query results, is "pulled" to the central server.

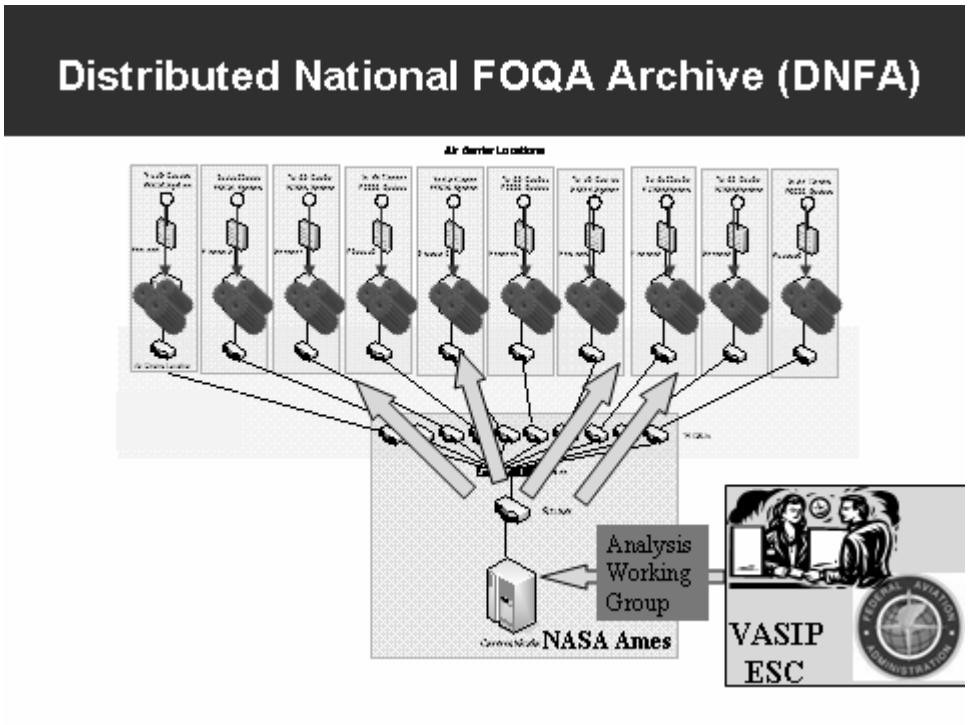

Figure 1. Querying Distributed archives

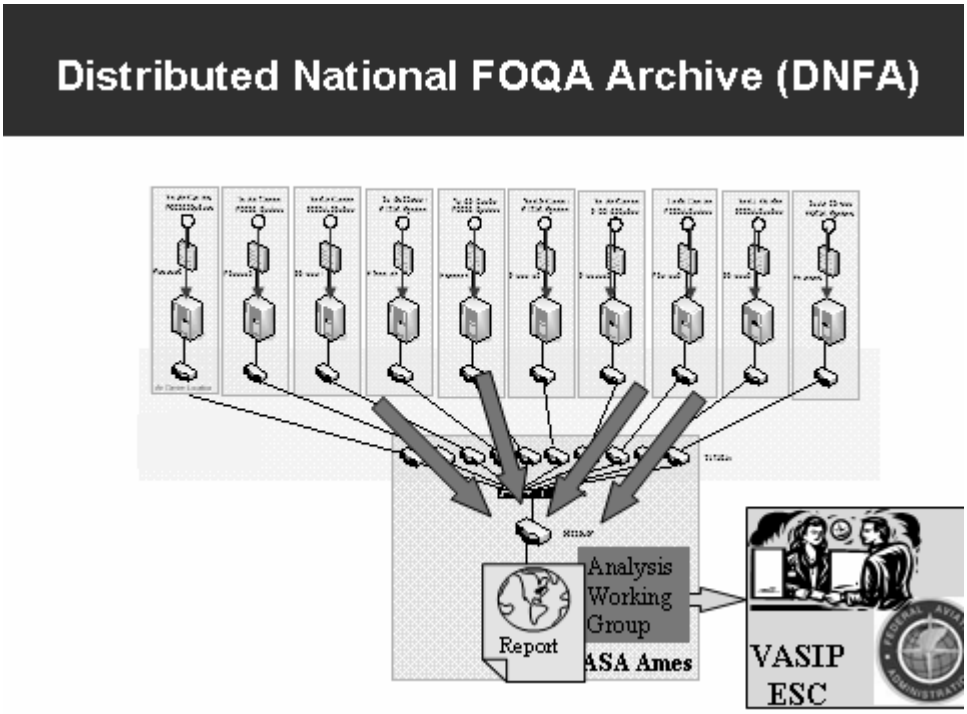

Figure 2. Analyzing Query Results

This allows the analysis working groups and ESC to assess problems of a national scope rather than individual events and operators. Each airline can compare its results with archive results, focusing any corrective actions at an appropriate level.

Functionality. Functionality for the demonstration project was limited to the functions described below, but NASA was asked to design the archive, network, and software to enable future growth in both number of airlines networked and analytic capability, and select hardware accordingly. At the completion of the program, the network must be capable of accommodating 50 total combined FOQA and ASAP program servers. The ARC specified these functions:

- The ability to search across local servers for flights experiencing defined events. All search activities will be logged on each local server for review by the airline owner of the data.
- The ability to calculate aggregate distributions for snapshots at any point in flight, which implies abilities to:
  - derive new or composite parameters from recorded values,
  - calculate statistics and display parameter and statistical distributions at or between any points in flight, and
  - include weather information, where recorded by vendor.
- The ability to view parameter traces of selected individual flights on their local server without moving the data to the central server.
- The ability to export selected parameters from selected flights to airspace visualization tools (i.e., traffic or 3-D trajectory) so that the track of individual flights or a selected sample of flights can be viewed within their airspace of operation.

A more detailed specification of functionality was provided by NASA to Battelle in November of 2004. The functional audit was organized around the more detailed specification.

Processing. The archive was built using distributed databases. NASA was asked to construct a network of servers located on airline premises to be accessed by NASA for aggregation, statistical analysis, and summarization. At each airline, selected, de-identified parameters at measured rates are pushed by the airline's FOQA-vendor analysis machine to the local archive server following completion of airline validation procedures. NASA was also asked to contract with vendors serving airlines participating in the DNFA to determine the format of data transferred to the local archive server, create the necessary software, and accomplish these services in a manner consistent with Federal Acquisition Regulations. The FOQA Data Aggregation Working Group (DAWG), chartered by the FOQA ARC, recommended the final list of parameters, which was approved by the VSIS ARC Rules and Procedures Sub-committee (RPS) and ESC. Downloaded data remain on the airlines' premises. Each flight would be de-identified so as to retain only month of flight, departure and arrival airports and runways, and aircraft model.

Hardware. Hardware specifications were deferred to the recommendations of NASA and its contractors and/or grantees. Redundant storage was specified to maintain two years' data on the local archive server at each airline. Discussions with the airlines further specified that local servers require minimal maintenance by each airline, essentially no more than the occasional re-boot by airline personnel.

Software. NASA was asked to develop all software necessary to the archive functions, making use of commercially-available software where practical, but not favoring any single FOQA-vendor.

Networking. NASA was asked to provide a secure network of T-1 (or better) lines connecting each local server to the central server at NASA Ames and to other ESC-approved nodes. Security performance in compliance with NASA Business and Restricted Technology (BRT) was specified; those standards were superseded during archive development and NASA complied with both the new set of standards and local requirements imposed by individual airlines.

Distributed National ASAP Archive

Overview. For safety reports submitted to airline ASAP programs, NASA was asked to implement a network of servers enabling de-identified information from multiple databases to be queried for aggregation, statistical analysis, and summarization in response to ESC requests. This network includes an archive server on each airline's premises, where de-identified data are housed, and a central server at NASA Ames Research Center, where queries are generated and summary results aggregated. NASA has also provided a server and network node to a grantee (University of Texas at Austin) designated to accomplish queries and aggregation requested by the ESC, and has delegated ASAP processing to that server. De-identified ASAP reports are pushed from the airline's ASAP server to the local archive server using software provided by NASA and its primary contractor, Battelle. This software creates a common format of categorized data derived from the submission and processing of ASAP reports at the airline level by mapping the airline's existing reporting and analysis system to a common taxonomic set.

Functionality. Functionality for the demonstration project was limited to functions described below, but NASA was asked to design the archive, network, hardware, and software to enable future growth in both number of airlines networked and analytic capability. The following functions were specified:

- The ability to search across local servers for ASAP-reported events based on categorized field searches.
- The ability to calculate distributions for categorizations, which implies abilities to:
  - derive new or composite categories or fields from ASAP report fields, and
  - calculate statistics and display categorization and statistical distributions based on ASAP report fields.

Processing. The DNAA is based on mapping ASAP data fields from participating carriers to a common format archive data structure. An ASAP Data Aggregation Working Group was appointed by the ASAP ARC to define fields and categories to be included or mapped to a common format for the archive. Airlines may choose to use formats built by NASA and its contractors for reporting, analysis, or tracking. Alternatively, airlines may participate by mapping the airline's existing reporting and analysis system to common fields, and pushing their data to the local archive server. Each ASAP report on the local archive server is de-identified so as to retain only month of flight, departure and arrival airports, and aircraft model.

Hardware. NASA selected DNAA hardware similar to DNFA. In the case of the DNAA, specifications that local servers require minimal maintenance by each airline, essentially no more than the occasional re-boot by airline personnel, resulted in significantly more capable and costly machines than were necessary to complete the archive functions alone. Were maintenance provided by participating airlines, the cost of ASAP servers could be greatly reduced, and this needs to be considered for deployment of future airlines. The maintenance requirement drove a higher cost than may be necessary for future installations.

III. IMPLEMENTATION STATUS

Status Report from NASA. On May 9, 2006, NASA Project Lead Irving Statler reported:

> The status of the DNFA and the DNAA as of the end of April is that the system is operational, test queries have been conducted, and the system is ready for operational queries as directed by the VASIP ESC.

The following is the status of current participating air carriers:

Alaska	DNFA only	Processing data
American	DNAA only	Processing data
Continental	DNFA & DNAA	Active – no DNFA data feed
Delta	DNFA only	Processing data
Frontier	DNAA only	Processing data
JetBlue	DNFA only	Processing data
Southwest	DNFA	Processing data
United	DNFA (DNAA '06)	Processing data
UPS	DNFA only	Processing data
ExpressJet	DNFA & DNAA	SAA completed, hardware on order

On April 10, 2006, a live demonstration of the archives was presented to representatives of the VSIS ARC and other key FAA personnel at a meeting hosted by JetBlue Airlines in New York. It was also presented to the ATA Safety Council in Washington, and to the Commercial Aviation Safety Team (CAST), to the Information Sharing meeting, and to the "Shared Vision of Aviation Safety" Conference in Denver using movies of the live demo where there was no node for connection into the secure network.

The ISI Team provided a status report of the DNFA and the DNAA to the VSIS ARC at its meeting on April 20, 2006.

DNFA:

- The final code was completed to adapt the World Wind KML event-display for use as the primary flight track and event display tool to support the DNFA reporting function. ATC Sector and Center boundary data are not yet available to NASA to support the desired flight-track displays in World Wind flight track. A request has been made to the FAA for access to this information.
- DNFA data are now being processed into the archives at five airlines at a rate of approximately 50,000 flights per month. The total number of flights in the distributed archives by the end of April was well over 200,000.
- United Airlines is now connected to the DNFA network, but we are still waiting on the resolution of final network setup configuration issues at United before the data will transfer to the NASA local server.
- Testing of all DNFA tools and systems are ongoing with demonstrated improvements to system reliability and efficiency. Additional functionalities and capabilities that were not demonstrated at the JetBlue meeting are being validated and tested.

These will be delivered with the required system stabilization by 31 May.
- A change was proposed to the VSIS/VASIP P&O document which would allow NASA and Battelle direct access to the DNFA network for development of a demonstration of an adaptation of the Morning Report for use with the DNFA. That new language was approved by the ARC at the April meeting. No .ffd files will be moved to the Central node; however, meta statistical data and processed data such as Flight Signatures and Intermediate Values may be returned to and accessed from the Central server.

DNAA:
- Archive hardware for four participating airlines has been activated and delivered. Support has been provided to these airlines to complete installation, network configuration, and testing at the local archive level. One of these four airlines will start actively pushing data to the DNAA pending the release of a new ASAP reporting tool. An additional 2 airlines are preparing for participation in the DNAA in 2006.
- All DNAA networks, central servers, airline nodes, and software are working properly. System testing is on-going.
- The DNAA Master List has been mapped to 6 unique airline and vendor formats. This mapping was completed in preparation for additional airline participation in the DNAA. The initial four participating airlines are currently categorizing their ASAP data using the DNAA Master List.
- A full referenced-based data dictionary is under development to support the DNAA Master List.
- The DNAA network contains approximately 2500 reports. Three airlines are currently providing data to the DNAA at a rate of 225 to 250 per month. It is predicted that with the participation of the three additional airlines schedule for implementation in 2006, the monthly rate will increase to approximately 650 per month.

Audit of functionality. On November 4, 2005, NASA conducted a preliminary demonstration of archive capabilities for representatives of the VASIP ESC at Battelle's facilities in Mountain View, CA. The DNFA demonstration was live, using samples of data from three airlines on local archive servers both present in Mountain View and remote at a sub-contractor's (ProWorks) facility. This allowed a full demonstration of functionality with a limited sample of data. The DNAA demonstration used test data residing on four archive servers present in Mountain View. This demonstration represented full functionality of the DNAA with test data.

At the full demonstration of both the DNFA and DNAA from the local archive server at JetBlue Airlines in Queens, New York, data were accessed from the five on-line airlines (of seven participating) for the DNFA and three airlines participating in the DNAA.

On May 10-11, 2006, the author traveled to Battelle's facilities in Mountain View to review functions that had not been demonstrated at the November 2005 and April 2006 meetings. This allowed review of functionalities as follows:

DNFA
- ***Software architecture specifications***. A system architecture design document was delivered to NASA by Battelle and its subcontractors on February 15, 2005.
- ***Server specifications***. Detailed specifications for all servers were delivered to NASA by Battelle on March 4, 2005. All specified hardware was purchased by NASA, assembled by Battelle, and delivered to participating airlines in January 2006.
- ***Generic parameter list***. The FOQA DAWG agreed upon a standard list of 399 parameters that the DNFA will accept from any participating airline's aircraft at each parameter's native sampling rate. Aircraft are not required to record all parameters to be included; the archive accepts those parameters recorded and indicates missing values for all others. However, searches cannot make use of data from aircraft on which search criterion values are missing. The archive automatically derives an additional 35 parameters during initial processing on each local archive server where the underlying parameters are available.
- ***Standard data format***. Battelle, SAGEM Avionics, and Austin Digital, Inc. agreed upon a standard data format for transfer from vendor FOQA machines to the local archive servers. Initial documentation of this format was delivered to NASA by Battelle on February 23 2005. A final version was delivered on September 12, 2005.
- ***Transfer of data from FOQA vendor machines to local archive server***. Transfer of data was initiated at four airlines by SAGEM Avionics beginning January 1, 2006. Transfer was initiated at one airline by Austin Digital, Inc., beginning January 1, 2006. Due to coordination issues internal to two Austin Digital customers, data conversion and storage were implemented at these airlines effective January 1, 2006 (see NASA status report above), but transfer of all stored data and ongoing data will be delayed until those internal issues have been resolved. As of April 30, 2006, more than 124,000 flights have been transferred to local archive servers by SAGEM and over 86,000 by Austin Digital. The number of flights at the two off-line Austin Digital customer airlines is significant but unknown at this time. The archive is adding flights at a rate of 50,000 per month, representing eight aircraft fleets.

- ***Loading of input data into archive format on the local archive server***. As of April 30, 2006, local archive servers at participating airlines have processed over 200,000 flights.
- ***Filtering to eliminate duplicate flights***. The Loader function detects a second instance of the same flight and automatically deletes the first instance of that flight from the local archive server database. This eliminates duplication while capturing corrected files re-processed by the airline's FOQA analysis machine.
- ***Data quality filtering excluding values and changes in values exceeding SME estimates of valid values***. Data quality filtering has been implemented for continuous parameters, but is pending for discrete parameters. Individual continuous parameter values are marked as "bad" during specific seconds of flight where the values exceed subject matter expert-entered limits for maximum, minimum, or rate of change. The percentage of bad data for each parameter is stored in the local archive server database. Flights manually marked as "bad" are excluded from analyses. Implementation for discrete parameters is expected by the end of the demonstration period.
- ***Derivation of new parameters***. The Loader service derives a set of 35 standard parameters (such as energy state and indices) for each flight when uploaded on each local server and allows the derivation on-demand of new parameters for all flights on each local server when commanded from the central server. A distinction must be drawn between dynamic- and persistent-derived parameters. Dynamic-derived parameters are calculated for a specific set of analyses and discarded. Persistent-derived parameters modify the database to aggregate their values going forward. DNFA users can derive dynamic parameters for data retrieval and analysis – such as Pattern Search or Distribution and Statistics functions. These parameters are discarded when the program is closed, but instructions for calculating the parameter are stored for future use. Adding persistent-derived parameters requires a modification to the Loader software and issuance of a new build. Derived parameters can only be generated when the underlying parameters are available. Calculation and use of dynamic-derived parameters was observed during the May 2006 site visit. Use of persistent-derived parameters (of the 35 standard) was demonstrated at JetBlue.
- ***Search for defined events***. The Pattern Search function allows a user at the central server to define modules and patterns for execution on the local archive servers. When a pattern is executed, the central server sends search instructions to each local server. Each local server executes a search and returns a list of those flights meeting the search criteria, annotated with the number of search module criteria met. The local server retains a log of each requested search and the generated flight list for review by local airline personnel. Functionalities of the local flight list are the same as the aggregate flight list, but the local list contains only flights from that local server. Each local archive server forwards its flight list to the central server, which generates an integrated flight list but does not reveal airline or flight-identifying information to the user. The edited integrated list can be used to filter any other analysis. Pattern Search software is designed to be compatible with future implementation of searches generated from and returned to any local server. The Pattern Search tool was demonstrated at the JetBlue meeting. Following this meeting, a normalization capability was added to requirements for the histogram displays of search results. This option shows event counts per 1000 flights or by percentage of total operations for selected fleets, airports, or other variables. This option is scheduled for implementation by May 31st.
- ***Calculation of aggregate distributions***. The Distributions and Statistics tool generates graphs and statistics at or between any selected routine events, given a selection of month range, aircraft type(s), departure/arrival stations, and parameters from the central server. When a selection is executed, the central server sends instructions to each local server. Each local server calculates and forwards to the central server statistics (and a flight list on which those statistics are based) necessary to allow integration of a report of all flights identified by the selection. The local server retains a log of each distribution and statistics request and the forwarded flight list and report for review by local airline personnel. The local report and flight list offer the same functionalities as the aggregate report and flight list but contain only flights on that local server. The central server produces statistics, distributions, and an integrated flight list on which they are based. Statistics include mean, median, standard deviation, values at each 10th percentile, and number of flights analyzed at each local server and combined at the central server. Selection of the flight list calls all functionalities associated with flight lists. Distribution and Statistics reports are filterable by Pattern Search flight lists to produce reports for selected flights and contrast with non-selected flights. Selecting any bin within a distribution produces a temporary flight list in a new window and makes available all flight list functionalities. Distribution and Statistics software is designed to be compatible with future implementation

of reports requested from and returned to any local server. This function was observed during the May 2006 site visit and functioned as specified.

- ***Calculation of aggregate statistics and display of distributions at or between events in flight***. The Distributions and Statistics tool aggregates and/or calculates values for each flight and statistics across flights between any two selected events. Examples might include distance traveled from touchdown to 30 knots on landing or airspeed variance during final approach. This function was observed during the May 2006 site visit.
- ***Flight list production on local and central servers by search, statistical distribution, and export functions***. Pattern Search and Distribution and Statistics tools produce flight lists on both local and central servers, and its functionalities generalize across these lists. Flight lists consist of flights selected by a search or analysis, along with the month of flight, departure or arrival station, and aircraft type. Selection of an individual flight from a flight list calls the viewer on the local server from which the flight originated. The production of a flight list using the Pattern Search tool was demonstrated at the JetBlue meeting. Several flight lists were produced and explored during the May 2006 site visit.
- ***De-identification of flight lists***. Flight lists allow the central server to access an individual flight on its local server but do not reveal airline or flight-identifying information to the user. Flight lists produced by the Pattern Search tool at the JetBlue meeting were fully de-identified, including screening of departure and/or arrival stations served by only one participating airline and by aircraft type if unique to a carrier.
- ***Sorting of flight lists***. Flight lists allow sorting by flight characteristics. This function was observed during the May 2006 site visit.
- ***Deletion of selected flights from flight lists***. Flight lists are editable to delete any selection of individual flights by the user. Deletion removes the flight from the list only, not from the database. In addition, flight lists can be merged by union, intersection, or difference. This function was observed during the May 2006 site visit.
- ***Filtering of export and statistical distributions by flight list***. Flight lists are available for filtering of statistics generation and export to visualization tools. Flight lists are available for filtering, statistics generation, and export. This function will be implemented (within the normalization capability of Pattern Search described above) on May 31st.
- ***Display of parameter traces of selected individual flights***. The Viewer tool is implemented on each local server. Viewer execution on the central server calls up the viewer and associated functions for one flight and its parameters, routine events, and search module markers (where applicable) on its local server. Viewer execution from the central server, thus, looks at data on the local server, but it does not transport the data file to the central server, nor does it reveal airline or flight-identifying information to the user. The viewer displays any selected list of parameter traces. The viewer interface can be populated with user-selected parameter sets for each phase of flight as well as for specific events, such as unstable approaches. The viewer trace can be marked with user-selected routine events and with module hits when the viewer is called from a pattern search flight list. These functions were observed during the May 2006 site visit.
- ***Export of latitude, longitude, altitude, and time to display flight paths in an airspace visualization tool***. The viewer allows single-click export of the active flight to an airspace visualization tool (World Wind, developed by NASA Ames) installed on central and local servers. Also, any flight list can be exported to World Wind. "Time" in the database is elapsed time of flight, rather than time of day, due to de-identification, but this still allows appropriate display and animation of flight paths. Export of multiple flight traces from the central server to airspace visualization was demonstrated at the JetBlue meeting. NASA is currently resolving the software update process for local server installations of World Wind. The software normally updates via Web access, which is not available on the secure wide-area network. The current plan is to download updates on the central server for access by local servers; this will be resolved by the end of the demonstration period.

DNAA

- ***Software architecture specifications***. A system architecture design document was delivered on February 15, 2005 to NASA by the University of Texas and Battelle and its subcontractors.
- ***Server specifications***. Detailed specifications for all servers were delivered on March 4, 2005 to NASA by the University of Texas and Battelle and its subcontractors. All specified hardware was purchased by NASA and assembled by Battelle for delivery to airline premises, which was completed in January 2006.
- ***Generic field list***. Over a series of meetings, the Data Aggregation Working Group, appointed by the ASAP ARC, specified demographic fields describing who

(such as crew position, but not name or airline), when, and where; event-type fields describing what occurred (such as an altitude deviation or runway incursion); and internal (to the reporting work group, such as the cockpit) and external contributing factors. This resulted in 82 primary and 466 secondary fields being shared, when available, from each participating airline. Airlines are not required to record all parameters to be included; the archive accepts those fields recorded and indicates missing values for all others. Searches cannot make use of data from airlines on which search criterion values are valid for less than three airlines. The generic field list or DNAA Master List will be used by seven airlines as a data collection schema.

- ***Standard data format***. Participating airlines may either use the master list or map their ASAP processing system fields to it in the standard format.
- ***Transfer of data from airline or vendor ASAP machines to local archive server***. ASAP data transferred from airline/vendor ASAP machines to the local archive server occurs through a firewall-protected network connection. The data transfer takes place through one of three options; 1) a message is sent from the local archive server to the airline ASAP server initiating a transfer of de-identified ASAP event reports to the local archive server. This option is supported by airlines currently using the UT ASAP Application for ASAP data collection and management. This data transfer option is available to any airline whose technical department will approve a transfer message to be sent to their ASAP server from the local archive server. 2) Participating airlines originate the transfer of de-identified ASAP records to the local archive server. This option requires an airline to generate their own transfer message to send ASAP records to the local archive server. 3) Airlines manually transfer data to the local archive server. This option is supported by a program installed on the local archive server that supports a manual data load. This option is made available to airlines that are unable to support a network connection from their ASAP server to the local archive server.
- ***Loading of input data into archive format on the local archive server***. There are currently 4500 records stored on the DNAA local archive servers available for querying.
- ***Identification and elimination of duplicate records***. Identification of duplicate records is completed on the local archive servers through assessment of the event identification fields provided by the airlines.
- ***Merging of multiple pilot reports to single-event report format.*** DNAA event records are created through the identification of key fields used to identify unique records. These fields identify the date, aircraft tail number, flight number and origination; and the status of the reporting pilot. This information is not made available for query, but is used internally during the transformation process that occurs at the local server on native airline data.
- ***Search for reported events based upon demographic, event type, and contributing factor categorizations.*** The Query application runs queries across local archive server databases. It supports searches for reported events based upon demographic, event type, and contributing factor categorizations supported by the DNAA master list. A query builder allows a user at the central server to select from any DNAA field values of interest and submit queries. The central server hands off queries asynchronously to each local archive server, where searches are accomplished and results collected and stored, and report lists are forwarded to the central server. Queries can include only de-identified data:
  - A query can not be completed if less than three airlines are included in the results
  - A query can not include City Pairs – the analyst must choose either Origination or Destination station
  - A query cannot return unique:
    - ♦ Destination or origination stations
    - ♦ Fleet types

These functions were demonstrated at the JetBlue meeting.

- ***Search for reported events based on one line summary descriptions.*** One of the demographic fields provided by participating airlines is a one-line summary. If collected by the airline, this information is provided by all pilots who have submitted a report as a short summary of the reported event. The DNAA query tool enables an analyst to search for terms contained in these one-line summaries using a word search entry and Boolean options. This function was demonstrated at the JetBlue meeting.
- ***Derivation of new or composite categories from ASAP report fields***. The DNAA Query Tool enables the selection of multiple fields to be used to create a query. Multiple fields can be included in a single query by using a grouping feature supported by parentheses and join terms such as *AND* or *OR*. Composite categories are created by the search function, rather than by creation of persistent parameters in the database. This function was demonstrated at the JetBlue meeting.
- ***Visual presentation of summarized ASAP event records based on queried results for working group review.*** The DNAA Query Tool supports a visual presentation of queried ASAP records in a log and full

report format. The log view of a set of queried data includes basic demographic information including date, origination or destination data, aircraft type, and a one-line summary of each event. The Full Report view shows all demographic, event type, and contributing factor information provided by the airline as these data have been transformed to the DNAA Master List and presented in a standard format. The Full Report view supports a view of the narrative description provided by each reporting pilot, as grouped to a single event report. This function was demonstrated at the JetBlue meeting.

- ***Ability to create and archive investigation folders for VSIS-approved studies.*** The DNAA Query tool enables an analyst to create and archive a folder for each study approved by the VSIS ESC. These folders contain all queries that are created and run to support the working group. Following the working group process, a folder can be archived. Query syntax and summaries of information stored in the investigation folder are retained and can be re-activated for future assessment, but the data accessed by the query is removed. This feature was developed to support the VSIS requirement that information only be assessed during an active, VSIS-appointed working group process. This function was demonstrated at the JetBlue meeting.
- ***User interface for creating visual charts depicting data from associated queries.*** The DNAA Query Tool Chart Generator was specified to provide an interface to automatically generate charts based on categorized data contained within queried records. This function is to be deployed by the close of the demonstration period. The Chart Generator will be available through an interface in the DNAA Query Tool. Its purpose is to enable data residing in DNAA event reports to be quickly summarized and visually represented. The Query Tool Chart Generator will enable a user to create and view charts of categorized information and access reports that are grouped in the resulting chart.

IV. SUMMARY

The VASIP archives of FOQA and ASAP data can be described as deployed and ready for operation as directed by the VASIP ESC. For FOQA data, five airlines are complete, two are awaiting internal processes but collecting data, and one is to be completed by September 30, 2006. For ASAP data, three airlines are complete and four are in progress for completion by September 30, 2006. Hardware, software, and networking have been implemented in a manner that supports the functions requested by the FOQA and ASAP ARCs in the fall of 2004. Archive functionalities can be described as complete with few exceptions:

- DNFA normalization functions will be implemented by May 31, 2006.
- Automated updating functionality for the World Wind airspace visualization tool will be implemented by September 30, 2006.
- DNAA Chart Generator will be implemented by September 30, 2006.

In sum, NASA, the University of Texas, and Battelle and its subcontractors have produced the requested system and activated its functions. De-identified data have been collected since January 2006, yielding over 200,000 flights of FOQA data and 4,500 events reported to ASAP. Underlying this success is a hidden, but significant contribution – prior to development of the archives, no standard file-interchange format had been developed for FOQA or ASAP data. The .ffd format specified by Austin Digital, SAGEM Avionics, and Battelle and its sub-contractors can be written or read by any vendor and can serve as an interchange format for FOQA data. The DNAA master list and mappings provide for a similar function for ASAP data.

Expansion to additional airlines can be readily accomplished with the infrastructure developed. Each airline requires a local archive server for FOQA and/or ASAP, a connection to the network, and a process for

transferring data to the server (and the transfer process has been established for FOQA customers of Austin Digital and SAGEM avionics and airline users of the UT ASAP system). However, it will be necessary to drive down some costs and assess how costs of added operators will be allocated. Hardware costs for the DNFA could be reduced somewhat, and for the DNAA significantly, were future participating airlines to agree to maintain the local archive servers. Software and operational costs (including data transfer and networking) for the archive will continue, presumably at a reduced rate from development and deployment. Costs associated with archive hardware, software, data transfer, and networking have been paid by the government during the demonstration period. Airlines participating during this period have taken risks, guided good solutions to technical problems, and covered internal costs to connect to the network. The aviation community will have to determine appropriate allocation of costs for future participants.

The value of the archives must be demonstrated through their use by the ESC to aid problem-solving. Collection and availability of data are of little value unless analyzed to identify or understand problems; understanding is of little value unless corrective action is pursued. The FAA, airlines, and unions entered into the archive process to understand and correct problems of national scope. The data are now available to support those goals. The challenge ahead is to ask good questions, understand the answers, and act upon them.