

NAOMS SURVEY RESPONSE REDACTION SUMMARY

December 31, 2007

This document is a summary of the redaction strategies and steps that NASA directed Battelle Memorial Institute to apply to the NAOMS information for the December 31, 2007 release.

Section A contains four redaction explanation tables.

Table 1 lists the redaction strategies applied, along with NASA's rationale for their use and the potential for information loss.

Table 2 lists the redaction steps applied to various fields of the survey responses.

Table 3 lists special "unique high" steps taken on air carrier responses concerning safety events.

Section B lists the categories used to generalize the information from certain fields in the air carrier surveys.

Section C lists the categories used to generalize the information from certain fields in the general aviation surveys.

Section A Redaction Tables

Table 1: List of NAOMS Redaction Strategies

#	Description of Strategy	Purpose	Information Loss
1	<p>Reordering of rows randomly within the selected time period</p>	<p>Eliminate any potential timing information implicit in original ordering</p>	<p>None</p>
2	<p>Generalization of type identifiers (e.g., aircraft make-model-series) to generic type characterizations or rounding of numeric data. Specific information that may be used to identify a respondent is replaced by broader categories that still retain respondent information.</p>	<p>Protects anonymity by taking very specific respondent information and making it more general.</p>	<p>Some precise information is lost but all general information remains.</p>
3	<p>Disaggregation of columns (column-wise separation of responses into multiple tables with no cross-links with other columns)</p>	<p>Protects anonymity by separating the entire column containing specific information that may be combined with other information to identify a respondent.</p>	<p>All information in the column is retained</p>

4	<p>Deletion of entire data column or row.</p>	<p>Elimination of a column may be used to delete information that when combined with other information poses too high a threat to anonymity. Deletion of a row is used to eliminate survey responses that posed technical problems for the redaction team, or for which the current set of redaction strategies was considered inadequate. These rows (survey responses) will be revisited at a later date for follow on release.</p>	<p>Deletion of a column eliminates that particular survey response across all surveys. Deletion of an entire row results in loss of an entire survey response</p>
5	<p>Selective editing of text entries at the individual record level</p>	<p>Protects anonymity by eliminating direct identifiers and generalizing implicit identifiers.</p>	<p>Explicit identification of people, aircraft type model series, airports, employers, etc. Generalization then used to retain information at higher (less specific) level.</p>

Table 2: Redaction Steps Applied to the NAOMS Responses

Data Fields	Redaction Steps
Recall Period	Delete recall period
Recall Date	Delete the recall month and day, including the recall season and pre- and post-9/11, but maintain the recall year
Free Text	Manually scan each field
	Generalize all explicit and high certainty implicit references to airports and aircraft make-model-series
	Delete all direct personal, operator, and air carrier names
	Generalize aircraft components
	After completion of the above steps, separate the general comments sections (D3A/D5 and GD3A/GD5) from the surveys, randomize their order and release them separately from the parent surveys
Flight Activity	Air Carrier Only: Round the flight hours to 40 hour "bins"; for flight hours less than 50, replace with "less than or equal to 50," and for flight hours more than 170, replace with "more than 170"
	General Aviation Only: Round the flight hours to 5 hour "bins"; for flight hours less than 80, round to the nearest 5; for more than 80, label as "at least 80"; Note: General Aviation pilots fly far fewer hours in a typical recall period than do Air Carrier pilots. Therefore the smaller numerical groupings (5 hours) were more appropriate for this pilot constituency compared with Air Carrier (40 hours).
	Delete number of legs in recall period, but post recall period legs and flight hours (together) in Partial Raw Air Carrier Survey Responses on the website.
	Generalize all other numeric flight activity (including career flight hours)
Airport Data	Generalize airport data per standard Federal Aviation Regulations classifications
Aircraft Make-Model-Series	Generalize make-model-series to aircraft class category
Air Carrier Fleet Size	Delete air carrier fleet size field. This step was done because when combined with certain other fields, it could reveal an actual air carrier.

Data Fields	Redaction Steps
All	For remaining fields maintained with their parent survey, randomize the order within the recall year
Rare Events in Numeric Fields	Rare events in numeric fields are defined as events occurring in less than 0.1 percent of the surveys within the set. When this happens, delete the entries in that column, yet show the actual summation of the field at the bottom of each column. Example: a column has 25,100 survey rows, and in that column there are only 20 entries, all '1's (<0. 1% of surveys recorded that event) . The redaction strategy makes all the '1's zeros, so that these "rare events" cannot be linked to their parent surveys, but keeps the total (20) at the bottom of the table so the reader can see how many times the surveys produced this result.
Unique high numbers in numeric fields	Replace high unique entries in numeric fields with the next closest numeric value in the field. Example: a column has a hundred entries: eighty 1s, nineteen 2s, and a single 3. The three would be called a high unique entry, and would be rounded down to 2. So, for this column, the 2s would actually be 2 or more. The column total would be low by one. Since the actual totals for air carrier safety events could be of interest to the reader, a summary of "high unique" steps for air carrier safety events, including actual totals, is listed in table 3.
Rare Events in Free Text Fields	Rare events in free text fields are defined as events considered rare to NASA aviation experts. Replace rare event descriptions with the term "other", and provide a separate sheet listing all events captured as "other". This retains a description of the rare event, but it is no longer associated with its parent survey. Example question and answer: Q: "Did any other devices have uncommanded movements during the recall period?" A: "Yes, I had an uncommanded deployment of my center main landing gear." This information is rare on a couple of counts: it is a rare system, only common to a small subset of aircraft make model series. Secondly, such an event is very uncommon, and would likely be public or company knowledge at some level. For that reason, this response would be changed to "other" in the information set, but would be listed as written in a separate section of the file.

Table 3: “High-Unique” Redaction Summary for Air Carrier Safety Related Events

Only the below listed air carrier safety related event columns were subject to the “high unique” rounding down redaction step described in table 2. The table lists the column number from the Air Carrier Redacted Survey Responses, a short description of the related event, the actual maximum number recorded in that column vs. the one posted on the website, and finally, the actual total for the column vs. the total posted on the website.

Column #	Safety Related Events	Actual	v	Posted Max*	Actual	v.	Posted Total
AD1	Altitude deviation	6	v	4	1314	v	1312
AH04	Landing w/o clearance	6	v	5	169	v	168
AH09	Hard landing	6	v	5	514	v	513
AH10	Takeoff out of CG limits	6	v	5	110	v	109
AH11	Takeoff overweight	6	v	5	147	v	146
AH12	Takeoff roll w/bad config.	4	v	3	234	v	233
AH13	Unusual attitude	7	v	5	274	v	272
AH14	Stall warning	6	v	5	336	v	335
AT1	Frequency congestion	40	v	32	21453	v	21445
AT2	Hi speed or hi alt approach	36	v	30	29364	v	29358
ER2	Hazmat damage	5	v	4	90	v	89
ER2a	Hazmat in cargo compartment	3	v	2	58	v	57
ER4.e	Uncommanded speed brake	4	v	3	70	v	69
ER4.h	Uncommanded slats	3	v	2	88	v	87
IC1A	Declined clearance change	13	v	10	2971	v	2968
IC1B2	Go-around or missed appr.	4	v	3	215	v	214
IC1B7	Landing long or fast	10	v	7	477	v	474
IC1B9	Conflict on ground	4	v	3	45	v	44
IC1B10	Other undesirable event	8	v	7	302	v	301
JD 2	Ground proximity warning	6	v	4	89	v	87

** Because of this rounding down of the unique high values, when these values are seen in their respective columns, they should be read as “X or more”.*

Section B: Air Carrier Categories

For all categories, assignments are made based on the predominant flight activities of the respondent as measured by the flight hour data he or she provides. So, for example, if a respondent indicates that she flies 60 percent of her time piloting a Medium Transport Turbofan and the remaining 40 percent piloting a Large Transport Turbofan, she would be categorized as an MED pilot.

nominalAircraftSize – Nominal assignment, size of aircraft flown during recall period

- **Widebody** – Widebody
- **Large** – Large transport
- **Medium** – Medium transport
- **Small_or_OTH**– All aircraft of known types that are not found in the WDB, LRG, or WDB lists.
- **UNK** -- Unknown

nominalPropulsion – Nominal assignment, propulsion used by aircraft flown during recall period

- **Turbofan** – Turbofan or turbojet
- **Turboprop_or_OTH** – Turboprop and all other propulsion types except turbofan
- **UNK** -- Unknown

Widebody / Turbofan	A-310-32	B-747-400	B-767-700	L-1011-100	Large / Turbofan
A-300	A-310-320	B-767-100	B-777-100	L-1011-500	B707-100
A-300-200	A-310-600	B-767-100	B-777-100	L-1011-500	B707-200
A-300-400	A-330	B-767-200	B-777-200	MD-10-30	B707-400
A-300-600	A-330-300	B-767-200	B-777-200	MD-10-30	B-757
A-300-B4	B-747	B-767-300	B-777-300	MD-11	B-757-2
A-310	B-747-10	B-767-300	B-777-300		B-757-20
A-310-200	B-747-100	B-767-300	B-777-300		B-757-
A-310-300	B-747-200	B-767-400	DC-10-300		200
A-310-300	B-747-200F	B-767-400	DC-10-300		B-757-3
A-310-300	B-747-300	B-767-500	DC-10-400		B-757-30
A-310-300	B-747-300	B-767-500	DC-10-400		B-757-
A-310-300	B-747-400	B-767-700	L-1011-400		B-757-40
A-310-300	B-747-400		L-1011-1		B-757-
					400

B757-767	A319-320	B-727-AD	B-737-600	DC-9-50	Small_or_OTH / Turboprop_or_OTH
DC-8	A-320	B-727-ADVANCE	B-737-7	DC-9-80	
DC-8-10	A-320-20	B-727-C	B-737-70	ELECTR A	
DC-8-20	A-320-200	B-737	B-737-700	F-100	
DC-8-30	A-321	B-737-10	B-737-700	L-100	
DC-8-40	B-717	B-737-100	B-737-800	L-382	
DC-8-50	B-727	B-737-20	B-737-800	MD-80	
DC-8-60	B-727-10	B-737-200	B-737-900	MD-88	
DC-8-70	B-727-2	B-737-3	B-737-900	MD-90	
DC-8-73	B-727-20	B-737-30	C-130	Small_or_OTH / Turbofan	
Medium-Turbofan	B-727-200	B-737-300	DC-6	[all Turbofans that do not fall in WDB, LRG, or MED categories]	
A-319	B-727-30	B-737-400	DC-9		
A-319-10	B-727-300	B-737-500	DC-9-10		
A-319-100	B-727-50	B-737-500	DC-9-20		
A-319-32	B-727-500	B-737-500	DC-9-30		
A-319-320			DC-9-40		

nominalMission – Nominal assignment, aircraft mission flown during recall period

- **Passenger** – Passenger
- **Cargo_or_OTH** – Cargo and other missions such as repositioning or maintenance
- **UNK** -- Unknown

nominalDomesticInternational – Nominal assignment, type of route flown during recall period

- **Domestic**
- **International**
- **UNK** -- Unknown

nominalPosition – Nominal assignment, crew position during recall period

- **Captain** -- Captain
- **FirstOfficer_or_OTH** – First Officer and other (second officer, relief pilot, checkpilot, etc.)
- **UNK** -- Unknown

nominalCareerExperience – Nominal assignment, career flight hours

- **Low** – Lowest 25% -- 6000 or fewer
- **Medium** – Mid 50%
- **High** – Highest 25% - 15000 or more
- **UNK** -- Unknown

ER1.ata – Replaces free-form component references in ER1.A with high-level ATA code reference

ER4.i.ata_1 – Replaces free-form component references in ER4.i.1_1 with high-level ATA code

ER4.i.ata_2 – Replaces free-form component references in ER4.i.1_2 with high-level ATA code

ER4.i.ata_3 – Replaces free-form component references in ER4.i.1_3 with high-level ATA code

ER4.i.ata_4 – Replaces free-form component references in ER4.i.1_4 with high-level ATA code

IC2.config – Airport runway configuration

- **Single**
- **Multi-Parallel**
- **Multi-Intersecting**
- **Multi-Both** – Both parallel and intersecting runways

IC2.scale – Airport traffic volume – empirically driven

- **Light**—lightest 25% percent of airports referenced in survey
- **Medium**—middle 50% percent of airports referenced in survey
- **Heavy**—highest 25% percent of airports referenced in survey

IC3.aircraftsize – Size of aircraft flown during most recent ICAC

- **Widebody** – Widebody
- **Large** – Large transport
- **Medium** – Medium transport
- **Small_or_OTH**– All aircraft of known types that are not found in the WDB, LRG, or WDB lists.
- **UNK** -- Unknown

IC3.propulsion – Propulsion used by aircraft flown during most recent ICAC

- **Turbofan** – Turbofan or turbojet
- **Turboprop_or_OTH** – Turboprop and all other propulsion types except turbofan
- **UNK** – Unknown

JD20.recency – Recency of recurrent training experience

- **LE 6** – less than or equal to 6 months
- **LE 12** –more than 6 months but less than or equal to 12 months
- **GT 12** – greater than 12 months

JD22.recency – Recency of CFIT training experience

- **LE 6** – less than or equal to 6 months
 - **LE 12** –more than 6 months but less than or equal to 12 months
 - **GT 12** – greater than 12 months
-

JD23.recency – Recency of upset recovery training experience

- **LE 6** – less than or equal to 6 months
- **LE 12** – more than 6 months but less than or equal to 12 months
- **GT 12** – greater than 12 months

Section C: General Aviation Numeric to Categorical Conversions

Notes

- (1) For all categories, assignments are made based on the predominant flight activities of the respondent as measured by the flight hour data he or she provides, or, in the case of nominalMission, as measured by the flight leg data they provide. So, for example, if a respondent indicates that she flies 60 percent of her time piloting a fixed-wing aircraft and the remaining 40 percent piloting a rotorcraft, she would be categorized as a fixed-wing pilot.
- (2) Flight activity data retained are to relate to Part 135 and Part 91 operations exclusively

Flight Activity Fields

Section A

nominalExperience – Nominal assignment, lifetime flight hours

- Low – Lowest 25% - ≤ 500 hours
- Med – Mid 50% - > 500 hours and ≤ 5500 hours
- High – Highest 25% - > 5500 hours

nominalPosition – Nominal assignment, position flown under Parts 135/91 during recall period

- PIC
- Other
- UNK

nominalLight – Nominal assignment, flies mainly days only or night flights

- Day
- Night
- UNK

nominalFlightLength – Nominal assignment, length of flights under Parts 135/91 during recall period

- Short – Less than 50 nautical miles
- Long – Greater than or equal to 50 nautical miles

Section C

GC12class – Nominal assignment, amount of instrument training

- Zero – No instrument training
 - Low – Lowest 25% -- ≤ 6 hours
 - Medium – Mid 50% - > 6 hours and ≤ 30 hours
 - High – Highest 25% - > 30 hours
 - UNK -- Unknown
-

GC13class – Nominal assignment, amount of instrument training in IMC conditions

- **Zero** – No experience instrument training in IMC
- **Low** – Lowest 25% -- <= 2 hours
- **Medium** – Mid 50% - > 2 hours and <= 10 hours
- **High** – Highest 25% - or more - >10 hours
- **UNK** -- Unknown

GC14class – Recency of last instrument training experience

- **Low** – less than or equal to 6 months
- **Medium** – more than 6 months but less than or equal to 12 months
- **High** – greater than 12 months

Airports

Section C

Regions of country where most flying is done by respondents. Respondents provided state data in the GC3 series fields. These data were converted to FAA Regions. There are ten fields because many respondents fly in more than one region of the country as well as a free-text field for pilot responses that cut across regions. Each of the first nine field assumes a value of 1 or 0, where 1 denotes respondent flying activity in that region. The tenth field contains a semi-standardized regional or multi-regional reference such as “Lower 48 States”.

- **GC3al** - Alaska
 - **GC3ce** - Central
 - **GC3ea** - Eastern
 - **GC3gl** - Great Lakes
 - **GC3ne** - New England
 - **GC3nw** - Northwest Mountain
 - **GC3so** - Southern
 - **GC3sw** - Southwestern
 - **GCwp** - Western Pacific
 - **GCoth** – other region or region.
-