APPENDIX D: HACKS DATA COMPILATION METHODOLOGY

D.1 Research Objective and Scope

The Hacks DataBook is a master dataset documenting 265 unique incidents of Alaska Airlines Mileage Plan account compromises reported publicly during calendar year 2025. The dataset was compiled through systematic search and verification protocols across multiple digital platforms to establish a baseline count of publicly documented account security incidents.

Temporal Scope: January 1, 2025 - November 30, 2025 (11 months)

Research Question: What is the minimum verifiable count of Alaska Airlines Mileage Plan account compromises reported publicly during 2025, and what patterns emerge from this documentation?

Data Collection Method: Manual search and cataloging of public posts, comments, articles, and reviews across multiple platforms, with dual-analyst verification of all entries.

D.2 Search Protocol and Platform Coverage

Search Terms and Variations:

Primary search queries focused on accounts of Alaska Airlines or Mileage Plan account compromises, including variations: "Alaska account hacked," "Mileage Plan stolen," "Alaska miles fraud," "Atmos account hacked" (post-merger terminology), and related combinations. Searches were conducted in English and, with translation, Chinese (US Card Forum).

Platform Coverage:

Data collection indicated eight distinct platform categories:

1. **Reddit:** Multiple subreddits including r/AlaskaAirlines, /awardtravel
2. **FlyerTalk:** Alaska Airlines Mileage Plan forum threads and related discussions
3. **Facebook Groups:** Travel rewards communities (e.g., ROAM, Frequent Miler group), airline-specific groups, and credit card optimisation communities
4. **Twitter/X:** Public tweets mentioning Alaska Airlines and account security issues
5. **US Card Forum:** Chinese-language credit card rewards community (www.uscardforum.com)
6. **LinkedIn:** Professional network posts discussing loyalty program security
7. **Review Platforms:** TrustPilot airline reviews mentioning account compromises
8. **News and Blogs:** Seattle Times articles, Frequent Miler blog posts, and other travel industry publications

Search Execution:

Searches were conducted using platform-native search functions (Reddit search, Twitter search, Facebook group search, Google site-specific searches) as well as Google search with site operators (e.g., site:flyertalk.com "Alaska account hacked"). Search results were reviewed chronologically from January 2025 forward.

D.3 Incident Identification and Initial Logging

When a potential incident was identified through search, it was immediately logged in the master DataBook (maintained as CSV file and Google Drive) with the following structured data fields:

Field Structure:

Field Name	Format	Description	Preservation Standard
Ref	10XXX	Unique reference number for internal tracking	Sequential assignment
Date	DD-MMM-YY	Date of comment/post publication	As displayed on platform
Platform	Text	Source platform identifier	Standardized category
Username	Text	Original poster's username	Exact preservation
Miles Stolen	Numeric (with commas)	Quantity explicitly stated by victim	Verbatim formatting
Comment	Long text	Complete post/comment text	Verbatim transcription
Case ID	Case-XX	Thread grouping identifier	Sequential by thread
Comment Link	URL	Direct link to specific comment/post	Verified clickable
Thread Link	URL	Link to parent discussion thread	Verified clickable

Field Name	Format	Description	Preservation Standard
Evidence Link	URL	Supplementary evidence if applicable	Optional
Archive Link	URL	Permanent archive snapshot	archive.today or archive.ph
Notes	Text	Analyst observations or context	Optional
New?	Y/blank	Flag for recently added entries	Pending verification

Data Entry Protocol:

All fields were populated at the time of initial discovery. The "Comment" field received particular attention, with complete verbatim transcription of the original post or comment text, including:

- Original spelling and grammar (apart from removal of commas for CSV compatibility)
- Punctuation and capitalization as written
- Informal language, abbreviations, emoticons
- No paraphrasing, summarization, or editorial modification

Rationale: Verbatim preservation maintains evidentiary integrity and allows independent analysts to assess tone, urgency, and credibility of victim reports without interpretive bias introduced by summarization.

D.4 Verification and Quality Control Protocol

Two-Analyst Verification Requirement:

Each potential incident underwent independent review by two separate data analysts to ensure inclusion criteria were met. This dual-verification process mitigated individual analyst bias and classification errors.

Inclusion Criteria (ALL Five Must Be Met):

1. **Airline Specificity:** Incident must explicitly reference Alaska Airlines, Mileage Plan, or (post-September 2024) Atmos Rewards with clear context indicating Alaska's loyalty programme, or clearly implied based on forum and/or thread
2. **Account Compromise:** Clear indication that loyalty program account was accessed without authorisation (e.g., "my account was hacked," "someone got into my account," "unauthorised access")

3. **Fraudulent Redemption:** Miles/points were stolen, transferred, or used to book travel without account holder's permission (as opposed to mere phishing attempt without successful breach)
4. **Temporal Relevance:** Incident was reported during 2025 (post date within temporal scope)
5. **Verifiable Source:** Incident documented on accessible public platform with preserved link (not hearsay or private communication)

Exclusion Criteria (ANY One Disqualifies):

1. **Non-Alaska Incidents:** Loyalty program breaches at United, Delta, American, Southwest, or other carriers
2. **Security Concerns Without Actual Incident:** Posts expressing worry about potential vulnerabilities without confirming an actual compromise occurred
3. **Service Complaints Without Compromise:** General Alaska service quality complaints, devaluation protests, or policy disagreements not involving account security
4. **Duplicate Reports:** Same username reporting identical incident across multiple platforms (counted once)

D.5 Data Preservation and Chain of Custody

Archive Protocol:

All incidents were archived using third-party archiving services to create permanent, timestamped snapshots of original content. Two archiving services were employed:

- **archive.today** (Primary): Provides immediate snapshot and permanent URL
- **Hunchly** (Secondary): Backup service for websites unable to be archived

Rationale: Archiving ensures evidence preservation in case original content is deleted, platforms change access policies, or accounts are suspended. Permanent archives provide timestamped proof of content existence and enable independent verification by third parties.

Chain of Custody Documentation:

For each incident, the following custody trail was maintained:

- **Discovery:** Date incident first identified by analyst
- **Verification:** Date second analyst completed review
- **Archive:** Timestamp of archive.today snapshot

D.6 Case Grouping Methodology

Case Definition:

A "Case" represents a single discussion thread, article, or social media post and all associated comments within that thread. Case grouping enables:

- De-duplication of multiple victims reporting in same thread
- Preservation of discussion context
- Analysis of community response patterns
- Tracking of how information spreads across thread

Case ID Assignment:

Case IDs (format: Case-01, Case-02, etc.) were assigned chronologically based on when the parent thread or article was first identified during the search process. The dataset contains 73 unique Cases, with multiple victim reports often appearing within popular or high-visibility threads.

Single-Thread Grouping:

All comments within one Reddit thread, FlyerTalk discussion, or Facebook post received the same Case ID. For example:

- Case-20: Reddit thread "Just noticed I had 250,000 miles fraudulently redeemed" contains 23 victim reports across post author and commenters
- Case-32: Facebook group post contains 14 victim reports from different users
- Case-72: Seattle Times article contains 8 victim reports in comment section

D.7 De-Duplication Protocol

Username Tracking:

All usernames were tracked across the entire dataset to identify potential duplicate reports.

De-Duplication Rules:

1. **Same Username, Same Incident, Multiple Comments:** If username reports same incident multiple times within one thread, count once (earliest mention preserved, with other comments added to same text box)
2. **Same Username, Same Incident, Multiple Platforms:** If username reports same incident on Reddit and Twitter, count once (both preserved in DataBook)
3. **Same Username, Different Incidents:** If username reports distinct incidents at different times (e.g., hacked in January, hacked again in July), count separately

4. **Different Usernames, Suspicious Similarity:** Flagged for review but generally counted separately unless definitive proof of same individual

Unique Incident Determination:

Final unique incident count = Total logged incidents - Identified duplicates

D.8 Quantitative Data Extraction

Miles Stolen Field Methodology:

When victims explicitly stated the quantity of miles stolen, this figure was captured exactly as written in the original post. No standardization, estimation, or inference was applied.

Capture Rules:

- **Explicit Statement Required:** Miles quantity included only if victim stated specific number (e.g., "250,000 miles stolen")
- **Format Preservation:** Original number formatting preserved including commas, spaces, or other separators
- **No Estimation:** If victim wrote "most of my miles" or "a lot of points," Miles Stolen field left blank
- **No Conversion:** If victim stated dollar value instead of miles, not converted (left blank)
- **Range Handling:** If victim stated "200,000-300,000 miles," midpoint not calculated (preserved as written or left blank)

Quantification Statistics:

Of 265 total incidents documented:

Category	Count	Percentage	Notes
Miles Explicitly Quantified	80	30.2%	Victim stated specific number
Miles Not Quantified	185	69.8%	Victim described theft without quantity
Total Incidents	**265**	**100%**	Complete dataset

Quantified Theft Distribution:

Among the 80 incidents with explicit mile quantification:

Statistic	**Value**	**Derivation**
Minimum theft	25,000 miles	Lowest quantified incident
Maximum theft	800,000 miles	Highest quantified incident
Mean theft	**217,831 miles**	Sum of 80 quantified thefts ÷ 80
Median theft	150,000 miles	50th percentile of sorted distribution
75th percentile	255,000 miles	Third quartile
90th percentile	350,000 miles	Ninth decile

Rationale: The 30.2% quantification rate likely introduces selection bias, as victims suffering larger losses may be more motivated to state specific amounts. The mean of 217,831 miles should therefore be interpreted as potentially representing higher-value incidents rather than a representative average across all 265 cases.