Big Data for flight safety and accident investigation

Evolving safety data into actionable insights

ANZSASI 2017

9-11 June, Wellington, NZ

Dr. Björn Hennig
Björn Hennig

- Pilot 1989 (St. Augustine, FL), privat flying St. Pete Air, KSPG
- PhD Operations Research Karlsruhe Institute of Technology (KIT), Germany
- Cranfield University, UK: Safety and Accident Investigation, MSc started 2015
- Performance and Safety Architect @ avialytics GmbH - FRA
- Aviation Experience: Airbus, Lufthansa, DHL, Aerologic, ACG …
... An ACTF [IATA Accident Classification Task Force] study revealed that of the approximately 1,000 accidents over the last decade, accident reports were available for only around 300 of them. And of those, many had room for improvement. ...

... Accidents are so rare that many states are challenged to maintain the necessary expertise. ...

... to address safety issues with greater speed. Safety is complex. Solutions to problems or changes to the way we do things need to be well thought out. ...

Alexandre de Juniac (IATA President and CEO) at the IATA Safety and Flights Ops Conference Seoul, 24 April 2017
Safety and Accident Investigation - Implications

- Level of Safety
  - Technical
  - Human Performance
  - Organizational
  - Socio-Technical / Cultural

Timeline:
- 1950s
- 1960s
- 1970s
- 1980s
- 1990s
- 2000s
- 2010s
Big Data Analytics

Definition

Process of collecting, organizing and analyzing large sets of data to discover patterns and other useful information.

Qualitativ and quantitative data

Trend vs. Random variation (chasing the numbers)
Zero-Error Principle or Error-tolerant Design

Zero-Error

Defence / Barrier Approach
Big Data: Safety Management & Accident Investigation

- Biases, Heuristics?
- Premature closure?
- Decision Support?
- Objective: *Prevention*

Safety Management

Accident Investigation
Evolving flight safety data into actionable insights
Safety Data Hub

NOAA's National Weather Service
Aviation Weather Center

APM

Lufthansa Systems
NetLine/Crew

InfoBOX ODS/DWH

Data Retrieval & Capture Agents

discretization

Repository

Flight Safety Data Mart (FS DM)

MOLAP Store (Cube)

Simulator Protocol

CPM / LDM

Q-Pulse®

Aerobytes
Big Data Example 1: Roster Robustness

<table>
<thead>
<tr>
<th>Severity: Recovery Timezones</th>
<th>Recovery hours (Rest, off, hotel)</th>
<th>Block hours</th>
<th>Circadian Low (Minutes)</th>
<th>Duty-Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt; 4</td>
<td>&gt; 4</td>
<td>&gt; 8</td>
<td>&gt; 120</td>
<td>VAC-&gt;Flight, OFF-&gt;Flight</td>
</tr>
<tr>
<td>+ 4</td>
<td>-4</td>
<td>+ 8</td>
<td>- 61 to 120</td>
<td>Minimum</td>
</tr>
<tr>
<td>+ 3</td>
<td>-3</td>
<td>+ 6 or 7</td>
<td>- 31 to 60</td>
<td>Minimum</td>
</tr>
<tr>
<td>+ 2</td>
<td>- 2</td>
<td>+ 4 or 5</td>
<td>- 16 to 30</td>
<td>Minimum</td>
</tr>
<tr>
<td>+ 1</td>
<td>- 1</td>
<td>+ 1 to 3</td>
<td>- 15</td>
<td>Minimum</td>
</tr>
</tbody>
</table>

Timeliness: (ahead of STD)
- 1 week +
- 48 - 167 hours
- 24 - 47 hours
- 12-23 hours
- 6-11 hours
- < 6 hours

Aerologic

<table>
<thead>
<tr>
<th>Roster Robustness Index</th>
<th>72.34%</th>
</tr>
</thead>
<tbody>
<tr>
<td>General stability Index</td>
<td>87.23%</td>
</tr>
</tbody>
</table>

Roster Robustness | Aerologic
Big Data Example 2: (Bio mathematical?) Fatigue

- Bio mathematical Fatigue Calculations (ALERT, SAFE, CARE ...)
- Vs. Experienced Fatigue
- Provide Reference to Crew?
- Monitor via Wearables?
Big Data Example 3: Text Mining (eDiscovery)

- Safety Reports
  - Boeing 777 Fleet
    - Bird Strike
    - Fatigue
    - Crew Rest
    - Commander Discretion
    - Safety Occurrence
    - Other
  - Airbus A320 Fleet
    - Bird Strike
    - Fatigue
    - Crew Rest
    - Commander Discretion
    - Safety Occurrence
    - Other
  - Embraer
    - Bird Strike
    - Fatigue
    - Crew Rest
    - Commander Discretion
    - Safety Occurrence
    - Other
- Ground
  - Ground Ops
    - Occurrence
    - E-Mails
    - Tech Docs
    - Other
  - Maintenance
    - Occurrence
    - E-Mails
    - Tech Docs
    - Other
- Flight Safety
  - SOP
    - Pre-Flight
    - Flight
      - Taxi Out
      - Take Off
      - Departure
      - Cruise
      - Arrival
      - Landing
      - Taxi In
    - Post-Flight

---

<table>
<thead>
<tr>
<th>Weight</th>
<th>Theme</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>Weather</td>
</tr>
<tr>
<td>97</td>
<td>Fatigue</td>
</tr>
<tr>
<td>97</td>
<td>Crew</td>
</tr>
<tr>
<td>82</td>
<td>ORD</td>
</tr>
<tr>
<td>82</td>
<td>Catering</td>
</tr>
<tr>
<td>81</td>
<td>APROX</td>
</tr>
<tr>
<td>80</td>
<td>Birdstrike</td>
</tr>
<tr>
<td>75</td>
<td>Water</td>
</tr>
<tr>
<td>72</td>
<td>Maintenance</td>
</tr>
<tr>
<td>71</td>
<td>Pilot tube</td>
</tr>
<tr>
<td>70</td>
<td>TCAS</td>
</tr>
<tr>
<td>70</td>
<td>Crew tube</td>
</tr>
<tr>
<td>60</td>
<td>Fish</td>
</tr>
<tr>
<td>58</td>
<td>Jump Seater</td>
</tr>
<tr>
<td>55</td>
<td>Ash</td>
</tr>
<tr>
<td>55</td>
<td>FRA</td>
</tr>
<tr>
<td>54</td>
<td>Aileron</td>
</tr>
<tr>
<td>50</td>
<td>Pilot flying</td>
</tr>
<tr>
<td>48</td>
<td>DAGBL</td>
</tr>
<tr>
<td>48</td>
<td>Stick Shaker</td>
</tr>
<tr>
<td>42</td>
<td>Scheduled Maintenance</td>
</tr>
<tr>
<td>42</td>
<td>WOCL</td>
</tr>
<tr>
<td>40</td>
<td>Beer</td>
</tr>
<tr>
<td>39</td>
<td>Attorney</td>
</tr>
<tr>
<td>39</td>
<td>EICAS</td>
</tr>
<tr>
<td>39</td>
<td>Safety pilot</td>
</tr>
<tr>
<td>34</td>
<td>###</td>
</tr>
<tr>
<td>34</td>
<td>Incident</td>
</tr>
<tr>
<td>20</td>
<td>FDM</td>
</tr>
<tr>
<td>15</td>
<td>Stabilizer</td>
</tr>
<tr>
<td>10</td>
<td>ACARS</td>
</tr>
</tbody>
</table>
Evolving flight safety data into actionable insights

Data Capture

Safety Data Hub

Data Access

Flight Safety Box

SPI Definition

Flight Safety Box

Dashboard

Flight Safety Box

Data Analytics

Safety Cube

Automated Knowledge Discovery

Safety DB

Double Loop Learning (e.g. Evidence-based Training)

...
<table>
<thead>
<tr>
<th>Flight</th>
<th>Reg.</th>
<th>DEP</th>
<th>RWY</th>
<th>ARR</th>
<th>RWY</th>
<th>STD</th>
<th>STA</th>
<th>ATD Delay</th>
<th>ABN</th>
<th>LND</th>
<th>ATA Delay</th>
<th>FDM Events</th>
<th>FDM Identifier</th>
<th>Occ. Report</th>
<th>Weather</th>
<th>Gate Keeper</th>
<th>Hard Landing</th>
<th>Unstable Approach</th>
<th>Fatigue</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>DXB</td>
<td>30R</td>
<td>26L</td>
<td>07:50</td>
<td>14:25</td>
<td>07:44</td>
<td>08:00</td>
<td>14:11</td>
<td>14:19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DXB</td>
<td>12R</td>
<td>26L</td>
<td>07:50</td>
<td>14:25</td>
<td>07:29</td>
<td>07:49</td>
<td>14:01</td>
<td>14:10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HKG</td>
<td>07R</td>
<td>26L</td>
<td>09:45</td>
<td>22:00</td>
<td>09:45</td>
<td>10:21</td>
<td>10:30</td>
<td>22:35</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BLR</td>
<td>00</td>
<td>05R</td>
<td>06:40</td>
<td>16:15</td>
<td>06:30</td>
<td>06:54</td>
<td>15:50</td>
<td>15:54</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FRA</td>
<td>10</td>
<td>28R</td>
<td>07:00</td>
<td>16:15</td>
<td>07:36</td>
<td>08:04</td>
<td>16:51</td>
<td>17:01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LAX</td>
<td>25L</td>
<td>07L</td>
<td>01:35</td>
<td>12:40</td>
<td>01:34</td>
<td>01:48</td>
<td>12:03</td>
<td>12:13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HKG</td>
<td>07R</td>
<td>08R</td>
<td>09:45</td>
<td>22:00</td>
<td>09:52</td>
<td>10:23</td>
<td>22:01</td>
<td>22:07</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ORD</td>
<td>20R</td>
<td>25L</td>
<td>18:15</td>
<td>23:05</td>
<td>17:42</td>
<td>17:56</td>
<td>22:05</td>
<td>22:10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
High rate of descent (<500ft)

- Unstable Approach at or below 1000ft
- Reduced flap landing
- Low Fuel on Landing
- High rate of descent (<500ft)

FDM Event Date: 17.01.2014 10:20
Event type: Flight Path
Event value / unit: -1072.000000 Feet per Minute
Reference: Sink Rate - max (below 500ft)
Event context: AIRSPEED=166.6, PITCH=-3.0, RALT=-56.6

Reason for invalidation: Approach: High rate of descent (<500ft)

Name: Approach: High rate of descent (<500ft)

Occurrence Reports: No connected Safety Reports available

Save changes
Evolving flight safety data into actionable insights
Evolving flight safety data into actionable insights...

Data Capture

Safety Cube

Data Analytics

Dashboard Box

Flight Safety Box

SPI Definition

Flight Safety Box

Data Access

Flight Safety Box

Safety DB

Automated Knowledge Discovery

Double Loop Learning (e.g. Evidence-based Training)

Data Capture

Safety Data Hub
Data Prediction

SQL Server Data Mining - Prediction Calculator

Prediction Calculator

This tool detects the patterns that predict a specific value (the Target) of a column based on values in the other columns. The patterns are presented in a scored format that allows assessing scores based on the values of the other columns. The tool generates an analytic report of scores, which can be used to analyze the impact of model prediction costs. It may also generate an operational Prediction Calculator and a printer-ready sheet.

Column Selection

- Target: HasHarLanding
- Exactly: True
- In range: False

Choose columns to be used for analysis:
- ATO
- DEP
- ARR
- WOSLShare
- DEPWeight
- DepFlightDuration
- HasDepartureDMEvent
- HasArrivalDMEvent
- BlockTripDuration
- DepATC
- HasFlightEvent
- HasATC

Output Options

- Operational Calculator
- Printer ready Calculator

Run | Close
Data Prediction

Prediction Calculator Report for the 'True' state of 'HasHardLanding'

Specify the costs and profits associated with correctly and incorrectly predicting 'True'. These costs/profits are needed to compute the optimum score threshold for the calculation.

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Value</th>
<th>Relative Impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>XG</td>
<td>[70-79 %]</td>
<td>489</td>
</tr>
<tr>
<td>YZ</td>
<td>815</td>
<td></td>
</tr>
<tr>
<td>OR</td>
<td>257</td>
<td></td>
</tr>
<tr>
<td>BRU</td>
<td>197</td>
<td></td>
</tr>
<tr>
<td>EMA</td>
<td>166</td>
<td></td>
</tr>
<tr>
<td>YUL</td>
<td>[80-85 %]</td>
<td>155</td>
</tr>
<tr>
<td>HKG</td>
<td>121</td>
<td></td>
</tr>
<tr>
<td>FRA</td>
<td>147</td>
<td></td>
</tr>
<tr>
<td>BIG</td>
<td>137</td>
<td></td>
</tr>
<tr>
<td>BOS</td>
<td>119</td>
<td></td>
</tr>
<tr>
<td>LHR</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>GVA</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>DEN</td>
<td>94</td>
<td></td>
</tr>
<tr>
<td>AMS</td>
<td>94</td>
<td></td>
</tr>
<tr>
<td>LAX</td>
<td>86</td>
<td></td>
</tr>
<tr>
<td>LGB</td>
<td>77</td>
<td></td>
</tr>
<tr>
<td>LHR</td>
<td>[10-19 %]</td>
<td>146</td>
</tr>
<tr>
<td>SFO</td>
<td>64</td>
<td></td>
</tr>
<tr>
<td>LAX</td>
<td>64</td>
<td></td>
</tr>
<tr>
<td>LAX</td>
<td>62</td>
<td></td>
</tr>
<tr>
<td>MIA</td>
<td>56</td>
<td></td>
</tr>
<tr>
<td>DFW</td>
<td>56</td>
<td></td>
</tr>
<tr>
<td>JFK</td>
<td>56</td>
<td></td>
</tr>
</tbody>
</table>

Score Breakdown

Individual scores for each state of each analysed column.

Profit for various score thresholds

Cumulative misclassification cost for various score thresholds
Evolving flight safety data into actionable insights

- Data Capture 
- Data Hub
- Safety DB
- Automated Knowledge Discovery
- Safety Cube
- Data Analytics
- Flight Safety Box
- SPI Definition
- Flight Safety Box
- Double Loop Learning (e.g. Evidence-based Training)
- Dashboard
- Flight Safety Box
Double Loop Learning, e.g. Evidence-based Training

➢ Double Loop Learning:
  • Deduct context and assumptions based on analytics
  • Verify and create hypotheses
  • Identify appropriate measures, actions and lessons learned
  • Implement measures and actions in training and/or operations
  • Trace measures in Flight Safety Analyzer
  • Question and validate assumptions based on results
  • Close control loop by adapting SPIs, proceedings and trainings

➢ Implications for Evidence-based Training:
  • Creation of Evidence/Performance profiles based on training and operation
  • Drill down SPIs and other facts to relevant level
  • Proposal of individual/team based scenarios for facilitated instructional techniques
  • Integration of simulator data, protocols and evaluation
  • Measurement of results via SPI and FDM analytics
  • Establish continues improvement process and learning (mentored)
  • Proactive identification of future risks, threads and individual weaknesses
    (e.g. Information-Sharing und Benchmarking - ASIAS)
Interested in participating?

Please contact me if you know of an airline or organization that might be interested in Big Data Flight Safety Analytics and Research MSc, PhD project
Thank you

Stay in touch!

bjoern.hennig@avialytics.aero
+1 702 445 8457
www.avialytics.aero

linkedin.com/in/bjoern-hennig-063b243
@BjoernHennig