Turun yliopisto
University of Turku

DATA MINING IN PROMOTING FLIGHT SAFETY

Olli Sjöblom

TURUN YLIOPISTON JULKAISUJA – ANNALES UNIVERSITATIS TURKUENSIS
Sarja - ser. E osa - tom. 5 | Oeconomica | Turku 2016

University of Turku

Turku School of Economics
Department of Management and Entrepreneurship, Information Systems Science

Supervised by

Professor Reima Suomi
Turku School of Economics
Turku University
Turku, Finland

Reviewed by

Professor Ari Visa
Tampere University of Technology

Director Regis Cabral
FEPRO - Funding for European Projects

Opponent

Professor Ari Visa
Tampere University of Technology

Cover photo: Mikael Häggblom

The originality of this thesis has been checked in accordance with the University of Turku quality assurance system using the Turnitin OriginalityCheck service.

ISBN 978-951-29-6546-5 (PRINT)
ISBN 978-951-29-6547-2 (PDF)
ISSN 2343-3159 (Painettu/Print)
ISSN 2343-3167 (Verkkojulkaisu/Online)
Painosalama Oy - Turku, Finland 2016

ABSTRACT

The incredible rapid development to huge volumes of air travel, mainly because of jet airliners that appeared to the sky in the 1950s, created the need for systematic research for aviation safety and collecting data about air traffic. The structured data can be analysed easily using queries from databases and running these results through graphic tools. However, in analysing narratives that often give more accurate information about the case, mining tools are needed. The analysis of textual data with computers has not been possible until data mining tools have been developed. Their use, at least among aviation, is still at a moderate level.

The research aims at discovering lethal trends in the flight safety reports. The narratives of 1,200 flight safety reports from years 1994 – 1996 in Finnish were processed with three text mining tools. One of them was totally language independent, the other had a specific configuration for Finnish and the third originally created for English, but encouraging results had been achieved with Spanish and that is why a Finnish test was undertaken, too.

The global rate of accidents is stabilising and the situation can now be regarded as satisfactory, but because of the growth in air traffic, the absolute number of fatal accidents per year might increase, if the flight safety will not be improved. The collection of data and reporting systems have reached their top level. The focal point in increasing the flight safety is analysis. The air traffic has generally been forecasted to grow 5 – 6 per cent annually over the next two decades. During this period, the global air travel will probably double also with relatively conservative expectations of economic growth. This development makes the airline management confront growing pressure due to increasing competition, significant rise in fuel prices and the need to reduce the incident rate due to expected growth in air traffic volumes. All this emphasises the urgent need for new tools and methods.

All systems provided encouraging results, as well as proved challenges still to be won. Flight safety can be improved through the development and utilisation of sophisticated analysis tools and methods, like data mining, using its results supporting the decision process of the executives.

Keywords: Data Mining, Text Mining, Flight Safety, Analysis, Decision Support, Strategic Management

TIIVISTELMÄ

Lentoliikenne kasvoi huomattavasti 1950-luvulla pääasiassa suihkumatkustaja-koneiden myötä, mikä aiheutti poikkeamatietojen järjestelmällisen keräämisen ja tutkimuksen tarpeen. Määrämuotoinen tieto voidaan helposti analysoida tietokantakyselyillä esittäen tulokset käyttäen graafisia työkaluja, mutta tekstianalyysiin, jonka avulla tapauksista saadaan usein tarkempia tietoja, tarvitaan louhintatyökaluja. Tekstimuotoisen tiedon automaattinen analysointi ei ole ollut mahdollista ennen louhintatyökalujen kehittämistä. Silti niiden käyttö, ainakin ilmailun piirissä, on edelleen vähäistä.

Tutkimuksen tarkoituksena oli havaita vaarallisia kehityskulkuja lentoturvallisuusraporteissa. 1 200 lentoturvallisuusraportin selostusosiot vuosilta 1994 – 1996 käsiteltiin kolmella tekstinlouhintatyökalulla. Yksi näistä oli täysin kieliriippumaton, toisessa oli lisäosa, jossa oli mahdollisuus käsitellä suomen kieltä ja kolmas oli rakennettu alun perin ainoastaan englanninkielisen tekstin louhintaan, mutta espanjan kielellä saavutettujen rohkaisevien tulosten pohjalta päätettiin kokeilla myös suomenkielistä tekstiä.

Lento-onnettomuuksien määrä liikenteeseen nähden on vakiintumassa maailmanlaajuisesti katsottuna ja turvallisuustaso voidaan katsoa tyydyttäväksi. Kuitenkin liikenteen kasvaessa myös onnettomuuksien määrä lisääntyy vuosittain, mikäli lentoturvallisuutta ei kyetä parantamaan. Turvallisuustiedon kerääminen ja raportointijärjestelmät ovat jo saavuttaneet huippunsa. Analysoinnin parantaminen on avain lentoturvallisuuden parantamiseen. Lentoliikenteen on ennustettu kasvavan 5 – 6 prosenttia vuodessa seuraavien kahden vuosikymmenen ajan. Samana aikana lentoliikenne saattaa kaksinkertaistua jopa vaatimattomimpien talouskasvuennusteiden mukaan. Tällainen kehitys asettaa lentoliikenteen päättäjille yhä kasvavia paineita kiristyvän kilpailun, polttoaineiden hinnannousun ja liikenteen kasvun aiheuttaman onnettomuuksien määrän vähentämiseksi. Tämä korostaa uusien menetelmien ja työkalujen kiireellistä tarvetta.

Kaikilla louhintajärjestelmillä saatiin rohkaisevia tuloksia mutta ne nostivat samalla esille haasteita, jotka tulisi vielä voittaa. Lentoturvallisuutta voidaan vielä parantaa käyttämällä tässä esille tuotuja analyysimenetelmiä ja –työkaluja kuten tiedonlouhintaa ja soveltamalla näin saatuja tuloksia johdon päätöksenteon tukena.

Avainsanat: tiedonlouhinta, tekstinlouhinta, lentoturvallisuus, analysointi, päätöksenteon tuki, strateginen johtaminen

ACKNOWLEDGEMENTS

"They will soar aloft as with eagles' wings"

Isaiah 40:31

The enthusiasm for aviation is often handed down from father to son, or adopted from other relatives or close family friends. This is true in my case as my father, Hans, was a glider pilot and flight instructor. He was one of the founder members of the contemporary Turku Flying Club, which re-emerged in the middle of the 1940s after the break caused by the war, and for which he also acted as a board member. And, as the vice-chairman of the club today, I follow in his footsteps. First and foremost, I would like to thank him and my mother Ritva, both of whom have passed away already, for the love, support and encouragement they gave me throughout their lives, especially during my master's studies and then my doctoral studies.

Efficient and safe aviation and air traffic can only be guaranteed by professional and highly skilled, efficient air traffic control (ATC), equipped with functional information flow systems and other sophisticated tools, enabling its operations. ATC provides advisory services and enforces rules in the air and at airports. An academic environment can be compared with that of air traffic control: the area control centre (ACC) controlling aircraft en route within a particular airspace at high altitude corresponds to the academic world, comprising universities and schools covering a wide area of responsibility and steering operations inside that area. A terminal control area (TMA) provides the ATC with services within the airspace – inside a specific radius from the airport and, in this comparison, Turku University is the TMA. Similarly, the aerodrome control tower (TWR), responsible for the traffic at the airport or airfield – with its air traffic controller – can be represented by the department of Information Systems Science and my supervisor, Professor Reima Suomi to whom I am most indebted. Alongside him, I have participated in several forms of co-operation regarding information systems since the late 1980s, first as his student, when carrying out complementary studies after my MBA, then as his colleague, teaching data storage techniques to master's degree students. During these decades of benefitting from Reima's knowledge, not only has my research work profited, I have also had the great opportunity to have him listen to my opinions and thoughts on a variety of subjects and then gained from his insight. An air traffic controller is a part of a synergetic team that provides high quality services, similarly the professors, Jukka Heikkilä and Hannu Salmela as well as the Head of Teaching, lecturer Timo Leino, have freely offered their knowledge and made significant contributions to my dissertation process. I gratefully acknowledge their influence.

The foundation of all aviation is high technology. An aviation organisation cannot survive without a highly skilled and trained maintenance department, providing specialised technical services. Without this sort of contribution, the data mining processes would have been impossible. I acknowledge the staff of Tampere Technical University Signal Processing Laboratory for mining the reports with GILTA and the supervision of Professors Ari Visa and Hannu Vanharanta, with Tomi Vesanen and Mika Koistinen as operators. Furthermore, I extend my sincere thanks to Mr. Shahbaz Anwar, the CEO and founder of PolyVista, Inc., Houston, Texas, for readily providing data mining services with PolyVista. I would also like to thank Jenny Knodell from PolyVista as well as Captain Jeff Hamlett, Senior Manager in Flight Safety from Southwest Airlines, for his comments about applying in practice the mining system mentioned above. I also sincerely wish to thank Lingsoft Inc., Turku, Finland, and its personnel Juhani Reiman, Simo Vihjanen and Sari Ahonen for providing me with TEMIS text mining services. In addition, I wish to express my deepest gratitude towards my former employer, the Finnish Civil Aviation Administration (nowadays TraFi Aviation) for granting me the right to use real flight safety data as test material for data mining and for providing me with it.

I sincerely thank Professor Ari Visa from Tampere Technical University and Doctor Regis Cabral for accepting the invitation to act as the examiners of my dissertation. Their criticism of my thesis was extremely constructive and helpful. Additionally, I would like to express my sincere gratitude to Professor Ari Visa for accepting the task of acting as my opponent at the disputation of this thesis.

The international air traffic industry provides an excellent example of highly advanced networked activities and operations. In order to gain the maximum operative performance from its resources, an airline, like any other organisation that operates airplanes, needs a broad organisation to support it, thus networking is a key issue for successful aviation. I owe my special thanks to the Global Aviation Network Information (GAIN) Working Group B, whose activities I participated in, and in particular to Andy Muir from the Federal Aviation Authority, who has encouraged me with his inspiring comments about my research articles.

I have had the privilege to participate in various conferences and doctoral seminars where experts have given me valuable comments and advice. I would especially like to mention the postgraduate seminars (later TISRA) of the Institute for Advanced Management Systems Research at Åbo Akademi University. Particularly, Professor Christer Carlsson, who organised the seminars. He deserves my sincere gratitude for his invaluable support and interest in my research project, which has been of great value in improving the quality of the dissertation.

It has been a pleasure and honour to share many memorable moments with a project group that carried out a study comparing aviation with healthcare, identifying the weak signals of lethal trends in both fields, thank you: Professors Sanna Salanterä and Tapio Salakoski, as well as Heljä Lundgrén-Laine, Veronika Laippala, Laura-Maria Murtola, Lotta Kauhanen and Juho Heimonen. Not to forget my numerous colleagues at the FCAA, my brothers and sisters at the Mission Aviation Fellowship (MAF) both abroad and in Finland as well as the friends and fellow aviators of Turku Flying Club and other sport aviation organisations, which have been so very influential in my study – done in the spirit of aviation. I am ever so grateful to all of you! I am especially grateful to my colleague and friend Anna-Maria Teperi from Finavia for providing excellent support and encouragement while she did her own doctoral studies.

Administration is also an essential function in aviation, without which – as one of its principal, basic activities – operative actions would only be weakly enabled, if at all. I wish to express my most sincere thanks to Eila Wilkman-Korkeamäki and Mari Jaakkola from the Department of Information Systems Science, as well as to Veronika Ståhlberg and Kirsi Tammi from the Office for Academic and Student Affairs – your work has been invaluable during my dissertation process. I also want to thank the Foundation for Economic Education (Liikesivistysrahasto) for providing financial support for my study as well as Turku School of Economics Association for subsidising the printing cost of the dissertation.

Many others have indirectly contributed to my work. I am indebted to a large number of colleagues, fellow PhD students, co-workers and friends who put their time and effort into improving my research. Unfortunately, as this research process has taken so long, you have been so numerous that it would be impossible to list you all, but I would like to single out Pekka Turunen from the early times, for being my ‘classmate’ during the first doctoral studies basic courses. With a fellow reserve officer and faithful friend, familiarisation with the new (and also somewhat scary) academic world was not as fraught as it could have been. Stina Ojala from the Department of IT also deserves a special mention in this context, having shared during the later studying times numerous coffee moments spiced by discussions about all themes between heaven and earth – the ends included. For your honest friendship, thank you, Stina! As flying is at least as much art as technology, I am very grateful to my good old friend and colleague Mikael “Micky” Häggblom, an excellent artist and photographer, who took the cover photo for this book.

Finally, and most importantly, I want to thank my wife Riitta for her love, patience, encouragement and support during the rather long time I worked on this thesis.

I dedicate my work to the memory of all the test pilots who offered their lives in promoting flight safety.

Turku, the original capital of Finland, August 2016

Olli Sjöblom

Table of contents

List of attached publications

Publication #1: Suomi, Reima and Sjöblom, Olli (2008). Data Mining in Aviation Safety Data Analysis. In H. Rahman (Ed.), *Social and Political Implications of Data Mining: Knowledge Management in E-Government*: Information Science Reference, Hershey, New York.

Publication #2: Sjöblom, Olli (2009). Data Refining in Mining Process for Aviation Safety Data. In the Proceedings of the 9th IFIP WG 6.11 Conference on e-Business, e-Services and e-Society, I3E 2009, Nancy, France.

Publication #3: Sjöblom, Olli (2010). Data Mining in Aviation Safety. In the Proceedings of the 5th International Workshop on Security (IWSEC 2010) - Short Papers, Kobe, Japan.

Publication #4: Sjöblom, Olli and Murtola, Laura-Maria and Heimonen, Juho and Kauhanen, Lotta and Laippala, Veronika and Lundgrén-Laine, Heljä and Salakoski, Tapio and Salanterä, Sanna (2013). Using cluster analysis to identify weak signals of lethal trends in aviation and health care documentation. International Journal of Networking and Virtual Organisations, 13(1), 66-80.

Publication #5: Sjöblom, Olli (2015) Data Mining Challenges in the Promotion of Aviation Safety. Submitted to Accident Analysis & Prevention.

List of figures

List of tables

List of abbreviations

ADREP	Accident/Incident Data Reporting
AI	Artificial Intelligence
AIG	Accident Investigation Section (of ICAO)
AIG/74	Accident Investigation and Prevention Divisional Meeting
AR	Action Research
ASAP	Aviation Safety Action Program
ASRS	Aviation Safety Reporting System
ATC	Air Traffic Control
ATM	Air Traffic Management
ATS	Air Traffic Services
BI	Business Intelligence
CAA	Civil Aviation Authority
CAASD	Center for Advanced Aviation Systems Development
CVR	Cockpit Voice Recorder
DM	Data Mining
DBMS	Data Base Management System
DFT	Discrete Fourier Transform
DSS	Decision Support Systems
FCAA	Finnish Civil Aviation Authority
EASA	European Aviation Safety Agency
EC	European Commission
ECAC	European Civil Aviation Conference
ECCAIRS	European Co-ordination Centre for Aviation Incident Reporting Systems
EDF	ECCAIRS Data Format
EK	Elinkeinoelämän Keskusliitto (Confederation of Finnish Industries)
EU	European Union
FAA	Federal Aviation Administration
FAI	Fédération Aéronautique Internationale
FDR	Flight Data Recorder, the ‘Black Box’
FMS	Fuel Management System
GAIN	Global Aviation Information Network
GILTA	Managing Large Text-Masses
IAEA	International Atomic Energy Agency
ICAO	International Civil Aviation Organization
IFALPA	International Federation of Air Line Pilots Association
IR	Information Retrieval

IS	Information Systems
IT	Information Technology
JRC	Joint Research Centre
KDD	Knowledge Discovery in Databases
NACA	National Advisory Committee of Aeronautics
NASA	National Aviation and Space Administration
NDM	Naturalistic Decision Making
NORDAIDS	Nordic Accident and Incident Data System
NLP	Natural Language Processing
NTSB	National Transportation Safety Board
OATD	Open Access Theses and Dissertations
OLAP	On-line Analytical Processing
OLE DB	Object Linking and Embedding for Database
ODBC	Open Database Connectivity
RDBMS	Relational Data Base Management System
RQ	Research Question
SOM	Self-Organising Maps
SPC	*Statistical Process Control* algorithm
SQL	Sequential Query Language
TEKES	Teknologian kehittämiskeskus (Finnish Funding Agency for Innovation)
TEMIS	Text Mining Solutions
TM	Text Mining
UK	United Kingdom
US	United States of America
VAR	Value-Added Reseller
VASA	Vaaratilanteiden Seurantajärjestelmä (Control System for Hazard Events)

1 INTRODUCTION

1.1 Safety development in modern aviation

A significant part of the European and global population flies on a regular basis for either business or pleasure and the rate of the traffic growth has been significant throughout the world since the 1950s (ICAO, 2008b). In May 2015, the airline database contained almost 2,475 active airlines (Lyall, 2015). At the beginning of this decade, more than 900 commercial airlines were operating worldwide, with a total fleet of 22,000 aircraft serving nearly 1,700 airports through a route network with a total length of several million kilometres (Abdelghany & Abdelghany, 2012). The role of aviation can be defined as "one of the most important services to offer both significant social and economic benefits. By serving tourism and trade, it contributes to economic growth" (Abdelghany & Abdelghany, 2012, p. 1). Air transport can be considered one of the safest modes of transport. As safety is the core concept in the airline industry, it can be considered one of the most important keys to the development of worldwide aviation, which is growing rapidly. Air transport in Europe is one of the fastest growing transport sectors (European Aviation Safety Agency, 2014). Due to the strong growth in traffic volume, the number of accidents is expected to increase if nothing is done to improve safety even further (see the International Civil Aviation Organization (ICAO); for a definition of an accident see Appendix 1 p. 1-1 (ICAO, 2010)). However, an increase in the number of accidents would be totally unacceptable.

Since the 1960s, air transport has experienced a higher growth rate than the world economy. During the time period between 1990 and 2003, the number of passengers increased on average by 2.7 per cent per year and the number of flights by 2.8 per cent per year. This growth, it should be emphasised, is far more pronounced in the European Union than in the United States. The number of passengers for instance increased on average by 5.5 per cent per year in the European Union, compared with 1.8 per cent in the United States (Assemblée Nationale, 2004). During the two first decades of the new millennium, world air traffic has continued its growth at a rapid pace and it has been predicted to double in the next 15 years (Ramasamy & Sabatini, 2015). Hence, it would be catastrophic if the amount of accidents would grow at the same speed as the growth rate in air traffic.

As a result of systematic and global work towards safer aviation over decades, the global aviation accident rate has diminished dramatically. As a good example of this favourable development, the years 1945 and 2004 can be compared: 2004 was considered one of the safest ever for air travel and had the same number of airline fatalities as in 1945, despite the fact that the volume of air travel in 2005 was far beyond that of 1945. Between the years mentioned, the annual number of passengers has increased from 9 million to 1.8 billion. This strong reduction in the percentage of incidents and accidents in proportion to air transportation has been achieved through systematic work to increase flight safety. With the same accident rate as at the beginning of jet age air travel in the 1950s, there would be about 300 major accidents every year in the United States alone (Department of Health, 2005).

In order to illustrate the present flight safety level, according to the International Air Transport Association (IATA), 2011 was worldwide the safest on record for commercial aviation with one passenger fatality per 7.1 million air travellers, from the point of view of operations this equals one accident for every 1.6 million flights. Since 2000 there has been an improvement of 42 per cent in the aviation accident rate (Oster Jr, Strong, & Zorn, 2013). However, due to the continuous volume growth in global aviation, the absolute number of fatal as well as other accidents may increase annually if flight safety is not improved.

Throughout its history, aviation has attracted an extraordinary amount of interest from people everywhere, which makes the industry sensitive to safety. Despite the fact that safety questions have always been the first and foremost important concern for the industry – and hard work has been done to increase its safety level by developing practices and operations, the application of the traditional means for ensuring safety has reached its highest level. The basis of this safety work has been the collection of data; collecting data about safety events is of great importance because the more the events there are without safety effect as such are to investigate, the lower the need to investigate accidents (Saatsi, Haavisto, & Oksama, 2011). The data have carefully been analysed by highly skilled specialists in order to find weak points in operations, and reports have been compiled aiming at the elimination of the impact of those weak points by revising and supplementing operation instructions and rules.

In the present situation, decisive information about aviation safety can no longer be generated in practice by using tools like Excel, database queries and report tools, etc., which are widely used today. The central issue is that when applying the methods used so far, operators have known what they have been searching for and this is not sufficient today or for the future, so the development of the practices of analysis and the methods of analysis is essential. Today, the query tools for searching for information and for creating reports are already sophisticated and can easily be combined with the so-called business intelligence tools that widen the use of the search criteria by allowing countless combinations of factors.

1.2 Improving flight safety through data mining

The actors in the aviation industry have observed that the vast amounts of data collected and stored by different actors in the field, contain hidden information that may be very valuable but which cannot be discovered with so-called traditional analysis methods (Watson, 1999). There are no preliminary expectations about the existence of these ‘nuggets of knowledge’ and that is why it is not possible to find them with known methods. In other words, we do not know what we do not know, thus what we are searching for are patterns that might reveal more knowledge that could improve safety (see Figure 19, p. 102).

Data mining tools and methods provide a solution to the challenges mentioned. The solution can be regarded as twofold: besides providing an analysis of textual data, data mining also allows it to be combined with numerical data. As the aviation safety data collected include both structured and narrative fields, the tools and methods, such as text mining tools, for analysing texts have only just been developed. The text mining capability appeared much later than data mining which is used for numerical data. Regarding numerical data, the analysis of which has been easy using queries from databases and running their results through tools that produce graphs, data mining provides more sophisticated numerical analyses when searching for something that has not been foreseen.

Expressed on a general level, two types of data are produced in the airline industry – the flight data recorders generate digital data tracking the flight and the deviation reports written by pilots and other personnel generate textual data. According to Robb (2005), the digital data is looked at by several entities, but because not as much work has been done on the free text data its analysis needs more attention.

1.3 Related research on aviation safety

A lot has certainly been written about aviation safety, at least in the major aviation countries, e.g. the US, the UK, Germany and France; many of them have their own institutes for flight safety working in close contact with airlines and authorities. However, in the area of analysing flight safety reports and applying data mining, tools and resources, research (prior to this work) is rather scarce. In the Nordic countries, only a few dissertations concerning, or at least referring to, flight safety have been published (for research beyond the Nordic countries, see

Chapter 3.4.1). In May 2015, a database search in Melinda (2015), the Union Catalogue of Finnish Libraries, was run in order to find Finnish dissertations concerning aviation safety. The Finnish, Swedish and English words were: ilmailu (aviation in Finnish), lentoturvallisuus (flight safety), lentäminen (flying), lentoliikenne (air traffic), flygtrafik (air traffic), flygsäkerhet (flight safety), flygning (flying), flight safety, aviation, aviation safety. The search was performed using several queries and 88 matches were produced when the duplicates were removed. No theses concerning aviation safety and closely resembling this thesis were found. Additionally, despite the search words in Finnish and Swedish, the origin of 22 of the search results came from outside the Nordic countries. The subject areas and the amount of found doctoral dissertations is displayed in Table 1. The division into the categories chosen is not absolute but indicative, because several of the theses would fit in several categories.

Table 1: Subject areas of the doctoral dissertations that include research related to aviation safety

Subject area	*Amount of dissertations*
Air traffic control and Communication	6
Birds and insects	6
Business and management	22
Ecology and environment	2
History and fine arts	5
Jurisprudence, politics and organisations	22
Physiology	9
Psychology, Human factor	8
Technology	3
Others	5

Another search was run in Melinda with the words tekstinlouhinta (text mining in Finnish) and text mining, receiving 11 matches. None of them were directly related to the approach of this study, although the dissertations of Kloptchenko (2003), Toivonen (2006) and Paukkeri (2012) have been of great importance from the linguistic and mining technology point of view. The dissertation by Timonen (2013) deals with short text documents as does this study, but from the point of view of term weighting focusing on the mining tasks, text categorisation, keyword extraction and query expansion. Also the theses by Lagus (2000) and Saarikoski (2014) are worth mentioning here as they present text mining that applies the Self-Organising Map (SOM) method. The SOM was developed in the early 1980s by professor Teuvo Kohonen at Helsinki University of Technology, the Neural Networks Research Centre (Honkela, 1998; Kohonen, 1997). Neural networks are a special part of data mining that have found a wide application area in business risks analysis. Although doctoral theses about aviation safety have only been recently written in Finland, pioneering work on data mining has been done. Professor Ari Visa, the University of Tampere, the Signal Processing Laboratory, has developed sophisticated prototypes for knowledge mining. The dissertation by Toivonen (2006) was written in conjunction with the GILTA project and several masters and licentiate theses were published during the project (see Chapter 5.3.3).

The former Technical University of Helsinki is now part of Aalto University and the only place in Finland offering higher education in aviation. Therefore, a library search covering all the different thesis types and other documents published there was performed using the Aalto University Library Aaltodoc search tool using the same search criteria as with Melinda. The result was 107 documents and only a few of them had some connection with flight safety from the point of view of this study, meaning flight operations. Usually they had a technical approach, such as aircraft maintenance, software development for aviation use, aircraft production, etc.

1.4 The aim of the study and research questions

The main actor regarding aviation safety in Finland is the Aviation of the Finnish Transport Safety Agency (TraFi, formerly known as Finnish Civil Aviation Authority, the FCAA). It is the state's national civil aviation authority and the civil aviation regulatory authority in Finland. Its overall aim is to reduce the number of air accidents and serious incidents, regardless of the inevitable continuous growth in air traffic or external safety threats. Like many authorities, the aviation authority in Finland has lacked more sophisticated analysis methods than the Excel and query tools of relational database systems, etc. Through the development of analysis capability, it ought to be possible to obtain and apply new methods and tools in order to improve safety in the airline industry as well as in the other areas of aviation, including authorities and maintenance organisations.

The author was employed at the Investigation and Analysis Unit of the FCAA during the first decade of the millennium when the idea about finding a text mining tool emerged. Huge amounts of narrative technical reports had piled up during the decades without having been analysed, except by individual inspectors. This method had allowed only the investigation of single, separate documents without actually obtaining a general view on the possible trends that exist in the material. As the author was aware of the existence of mining tools, a small idea that the mining tools could provide a possible answer to the question appeared. Data mining and especially text mining had been utilised successfully in English as well as in some other major languages to examine flight safety reports, but there was no evidence of projects for mining short reports in Finnish. Some newspaper and magazine articles, however, had been mined in Finnish to a certain extent, but the narratives of the flight safety reports, the sizes of which varied between one word and a couple of sentences, posed a new challenge.

Hence, the main research question of this study is:

Can flight safety be augmented by applying data mining as a method of analysis?

This general research question can be divided into two more specific sub-questions (SQ) in order to approach it more accurately:

SQ1: Is it possible to find lethal trends in aviation safety data by using text mining?

The main objective when examining safety reports is to find similar reports due to the presence of clustering, because their existence might indicate a similar event, which, if it continues to occur, could lead to a serious incident or accident.

SQ2: What sort of refining actions would make the mining results more applicable and are the text mining results significantly more accurate, reliable and valid after refining?

The data mining process does not give straight answers to questions, so its role can be considered to be a decision support system (DSS). Additionally, according to the theory of data mining it is an iterative process consisting of several mining rounds requiring the refining and adjusting of both data and mining definitions between them. The mining process always requires the participation of analysis and mining experts as well as measuring methods in order to discover both errors and inaccuracies in mining definitions that aim to refine the mining results. In this process, commonly used tools like Excel provide sufficient support, allowing the analysis of mining results to occur in a simple and clear format.

Data mining is not just a method or tool but can be described as a combination of methods and tools. It is a very broad area with a huge number of variations which requires strict limitations. In practice it might be necessary to focus on only a couple of main issues, like clustering the data and benchmarking the different mining tools, by performing the process with several systems and comparing the results with each other. It was planned that this study would be realised in accordance with these principles, which most certainly would add a relevant framework to the applicable results for academic research.

After working for several years in aviation administration in Finland and also developing an extensive international view on safety analysis tools and methods, the author became aware of the gap between academia and practice. As science had made significantly successful attempts to solve problems that practitioners were perhaps not even aware of, the author became increasingly more positive and motivated about the need for this text mining research.

1.5 The structure of the research

In addition to the introductory section, the thesis consists of five publications. Publication #1 presents the whole data mining process described in Chapter 3.3.1 and illustrated in Figure 16, which addresses SQ1. It explains in detail all the different steps: beginning from defining the test data to be used and choosing the tools to be tested with the chosen data. The selection of the method or methods to be used in the mining process had to be made in this phase of the study. The pre-processing of the data in order to perform the first mining round with the selected tools is described in detail. Also the functions of some of the parameters that could have been applied when using the tested tools are described. At the beginning of the research there was no experience of mining short text reports written in Finnish. The result of the first mining round was that the text mining tools were able to extract trends, which, in this case, actually meant recurring events that were or became incidents. However, in this case they neither developed into dangerous risks nor accidents. Another remarkable point is that all the different tools produced coherent results. In line with data mining theory, the first mining round revealed the need for refining the definitions of the test data and the mining parameters for the second round in order to achieve more accurate mining results.

Publication #2 continues the research process by analysing the results gained from the first mining round with all the three tested tools addressing SQ1 and SQ2. Although the pre-processing for the first mining round was made very carefully, the results achieved from the first mining process revealed an obvious need to refine the definitions. In particular, it was noticed that the lists of both stop words and synonyms contained several pure mistakes that were easy to correct. In order to further analyse the results in more detail, they were studied using a qualitative data analysis application NVivo, version 7 (see Chapter 5.4). Using this kind of tool, which enables the user to manage and make sense of unstructured information, was necessary because using Excel and Word alone would not have provided results adequate enough to understand the relationships between the concepts in the test data. When analysing the content of the clusters using queries, NVivo took into account all the words defined in the query options, determining their relevance in the text. As one essential point in this process was to keep the data as untouched as possible, no major changes were made. Almost one hundred checking procedures were made with stop words and synonyms. These procedures prepared the test data for the second mining round, significantly improving its quality.

Publication #3 is written about the results after the second mining round that was performed after refining the test data addressing SQ2. The second mining round could only be performed with two of the tools tested in round one, because the data material had to be sent to the operators in order to be run and this was not possible with the third tested tool. From the point of view of the total research results, this was not necessary. Unlike most information systems that are constructed on English language developed tools, and thus often lack the conjugated forms of compounds, this tool included a module for mining Finnish texts. This feature allowed the processing of documentation in Finnish, which is a morphologically rich language. The first main focus of the publication was to analyse the ‘problematic’ words displaying their significance, the number of their appearance and their relevance as a percentage in the clusters. This was done by making queries in NVivo. The second main focus was comparing the contents of the clearest clusters as pairs. The comparison was made by forming pairs from the results of rounds one and two among the same tool and estimating the improvement that had occurred between the mining rounds. The pair comparison was also made between the two tools in order to crosscheck the impact of refining the data between the mining rounds. All the individual tests proved that the rather small but obvious process made the mining results more accurate through the refining of the data. In addition, the tests confirmed the need for several iterations in order to achieve more accurate mining results. Another significant observation was that the size of clusters plays an important role, thus the more clusters the better the results, although an optimum point is important to find.

Publication #4 was produced in order to explore whether the data mining technique used in this research process could be applied to health care documentation. As aviation and health care share similarities, the assumption was that the methods for improving safety could be transferable between the domains. In both contexts, the eventuality of an undesirable event may be reduced markedly if the risks that are connected to it are diagnosed. The mining process presented in the three first publications of this thesis was taken as the basis for comparison with the text mining process carried out using a patient record dataset as test material. From 26,000 electronic health records 1,083 text units were extracted that contained the discharge summaries of the physicians mentioning resuscitation in their narrative sections. This occurrence, like those described in the flight safety reports, can lead to either a successful or lethal outcome. Both datasets were clustered into groups of records in order to find similarities between cases that might contain lethal trends. As the applicability of data mining for flight safety reports was confirmed, using health care data, the identification of hazardous events turned out to be possible if clustering was used as a method.

Publication #5 deals with the applicability of data mining results for the subject area of the strategic management of organisations. However, unlike in Publication #4, the results are not compared with the mined results of the other application areas, but they are discussed as possible answers to the continuously fast growing requirements of management; the information environment is expanding at a pace never seen before, which makes decision-making increasingly more complicated. This is especially emphasised among safety-critical environments, such as nuclear power and air traffic. System failures or inadequate or delayed decision-making in these high risk environments can cause the loss of life, significant damage to property or environmental pollution. As organisations are becoming more complex, strategic decisions especially require highly refined information as their basis. The need for an automated means to process data is increasing rapidly as the amount of generated and stored unstructured data is expanding rapidly; usually it also accumulates faster than it can be processed. All the test results support the premise that if lethal trends exist in the stored data, data mining will glean safety information from it. This information would not be accessible with other methods and can be used as an essential component for strategic safety management.

2 FLIGHT SAFETY

2.1 Safety Overview

In this context, it is important to present the difference between safety and security, especially because there is only one word for both concepts in many languages; for example, we have the Finnish word 'turvallisuus', the German 'Sicherheit' and the Swedish 'säkerhet'. The Oxford Dictionary (2014) definition of safety in English is as follows: "The condition of being protected from or unlikely to cause danger, risk, or injury". Correspondingly, the same source defines security as "The state of being free from danger or threat". Merriam-Webster (2014) defines safety as "freedom from harm or danger: the state of being safe", the full definition is "the condition of being safe from undergoing or causing hurt, injury, or loss". Similarly, the definition for security is "the state of being protected or safe from harm; full definition: a) freedom from danger (here seen as a synonym for safety), b) freedom from fear or anxiety" (Merriam-Webster, 2014). Risk can statistically be defined to describe the probability of an event that is untoward or adverse in its consequences (Kemp, 1993; Michalsen, 2003). It is obvious that these terms are closely connected to each other and distinguishing them might often be cumbersome. To clarify the approach used in this research process, the observation is simplified by making the principal division on the basis of premeditation, that is, when a danger or threat is caused with intent, activating the countermeasures can be called security activities. If the threat appears unintentionally, protection against it is achieved through safety actions.

It is very difficult, even impossible to estimate when safety performance is good enough and we can define the level at which we can relax our defences and rest on our laurels. There is, in fact, no such level. Latino (2008) mentions what many would call an example of incorrect thinking at NASA prior to Challenger and later prior to Columbia, too. The space shuttle Challenger exploded in flames during take-off 73 seconds after the launch on January 28, 1986 killing all the astronauts aboard. The direct cause of the accident was purely technical, a faulty O-ring in the fuel system causing a propellant leak leading to an explosion of the external tank. However, as usual, there was also a chain of events leading to the accident. The engineers had been concerned about the problems of the O-rings, especially in low temperatures, for more than eight years. As the launch commit criteria were ineffective, it led to a decision to launch at a low temperature (Janis, 1997; NASA, 2011; Seton Hall University, 2012). Behind the series of the events there were ethical questions about organisational structures and cultures from the point of view of promoting or discouraging necessary communication. The origin of the disaster can be called Groupthink – a phenomenon often occurring in large organisations. The term describes a process that leads a group to make a bad or irrational decision. In this sort of situation the members of the group tend to alter their opinions to what they believe to be the consensus of the group (Janis & Mann, 1977).

During its re-entrance into the earth's atmosphere on February 1, 2003, the space shuttle Columbia broke apart and the whole crew of seven astronauts was killed. The accident originated from a large piece of insulating foam, which was dislodged during the launch from the external tank of the shuttle. This, hit the leading edge of the wing causing serious damage leading to the accident 16 days later. The foam strike is an illustrative example of a concept called ambiguous threat, that is, the potential of the event for causing harm is unclear, leading to managers often taking a wait-and-see attitude and actively discounting or even ignoring the risk. As in this case, the approach led to a catastrophe (Roberto, Bohmer, & Edmondson, 2006). In both of the cases presented above, the question is about detecting the weak signals (see Chapter 3.2.2) and responding to them.

Latino (2008) continues by describing a situation in which an accuracy level of 99.99 per cent, indicating that there would not be much room for improvement, still allows events to take place:

- Two unsafe plane landings per day at O'Hare Airport
- 500 incorrect surgical operations each week
- 50 new-borns dropped at birth by doctors every day
- 22,000 checks deducted from the wrong bank account each hour
- 32,000 missed heartbeats per person, per year.

The common opinion would say that none of these events should be accepted. In order to measure the level of safety, certain definitions ought to be made and indicators taken into use. Safety indicators are usually defined as frequencies of occurrences of certain harm causing events. The report of the Second Meeting of the North American, Central American and Caribbean Directors of Civil Aviation has expressed the essential requirements of typical indicators as follows (ICAO, 2005, p. 3):

- fatal aircraft accidents per flight hour
- fatal aircraft accidents per movement
- fatal aircraft accidents per year
- serious incidents per flight hour
- fatalities due to aircraft accidents per year.

Brooker (2006, p. 2) has stated that the essential role of safety indicators is the ability to control the performance of the system (i.e., air traffic), enabling their users to predict the plausible frequency of critical system failures emphasising that the essential characteristics of an appropriate indicator:

- are simple to comprehend
- are capable of being calculated by a checklist process
- are 'obvious', in the sense that people would agree that it was a sensible thing to measure – what the psychologists call 'face validity'; and
- do not require complex modelling calculations to be carried out in order to 'weight' the data appropriately.

Brooker (2006) continues with the theme mentioning that the quality of the predictions should meet reliability in a statistical sense, that means, it does not need to attain perfection. Another focal point is that a reliable indicator that ought to match the underlying patterns changes about 80 per cent of the time.

The sectors of enterprise safety and security (Confederation of Finnish Industries (EK), 2014) are presented in Figure 1. Identifying the threats and assessing their severity are the central prerequisites for determining and defining the scale of enterprise safety and security.

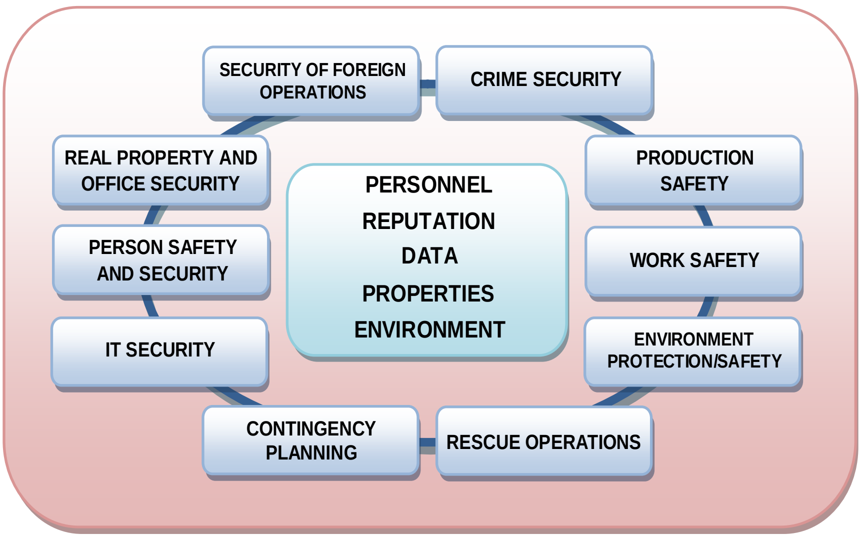


Figure 1: Enterprise safety and security (adapted from (Confederation of Finnish Industries (EK), 2014))

As the Finnish language contains only one word for the number of expressions of safety and security, the accuracy of Figure 1 is not sufficient as such, but appropriately requires some explanation due to the difficulty of presenting the translated terms without ambiguity. The most descriptive term, either safety or security has been chosen, but some exceptions can be found. A rough distinction has been made as to whether some threat is formed on purpose or arises accidentally, although some components might contain both aspects. Crime security is obviously protection against crimes committed or attempted, clearly on purpose against the operations, personnel or properties of an organisation. The core function of production safety is to guarantee the undisturbed production of products and/or services including quick recovery after a failure as well as estimating the business risks and making alternative business plans. This is in close connection with contingency planning that prepares organisations for exceptional circumstances by adjusting risk analyses, ensuring the availability of necessary spare parts as well as the continuity of repair and maintenance functions, which as a whole can contain both safety and security aspects (Confederation of Finnish Industries (EK), 2014).

Work safety is protection against occasional and systematic harms and risks in daily work and sharing the responsibility for protection. Environment protection and safety take ecological sustainability strongly into consideration by foreseeing the environmental expectations of customers and society; this requires developing processes and best practices, taking responsibility for the environment and augmenting the eco-consciousness of the personnel. Rescue operations include training and the education of personnel against risks; managing the risks related to accidents by foreseeing and eliminating or, at least, minimising them; developing activities to prevent threats; and applying sufficient surveillance and access control functions (see the ATR-72 accident, Chapter 3.2.3). The majority of these can be seen as belonging to safety but the last one clearly belongs to security (Confederation of Finnish Industries (EK), 2014).

The security of foreign operations concentrates on the risk analyses of foreign countries and reducing or eliminating the effects of a high risk level in the target countries. Real property and office security is responsible for security control and access control in buildings and operating areas. Person safety and security clearly contain both aspects: from the safety point of view, the aim is to protect personnel against accidents, which is closely related to work safety. The security approach contains the activities and arrangements aimed at protecting the personnel and customers against crimes, enabling secure travelling and, in special cases, body guarding.

A special point in this context is IT security, as an organisation’s information is one of the core resources of enterprise today (see Publication #5 “Data Mining Challenges in the Promotion of Aviation Safety”). Although this component mostly consists of security functions, it can and shall be carefully and continuously monitored for safety aspects in order to guarantee the safe and secure uninterrupted use of the data resources of the organisation. This aspect has recently received even more significance due to Edward Snowden revealing the network intelligence work of the National Security Agency in US.

IT security is one of an organisation's core concerns and increasingly requires resources to guarantee the confidentiality, the integrity and applicability of data resources. Other essential functions also assess the significance of the data resources for the organisation and for ensuring the continuance of business. In other words, both hardware and the information services provided ought to be protected from all external as well as internal threats, like unauthorised access, change or destruction. IT personnel have always held a privileged position because they have, at least theoretically, access to the functions of the organisation because they are a kind of supervisor. This has created the problem of 'who is supervising the supervisors?' A network administrator like Snowden usually has limitless access to the IT domain which makes it possible to freely collect interesting documents. The ability to know and understand what IT personnel are doing requires skills and a competence that should be higher or at least equal to those to be supervised.

As a more important aspect than security, safety is concerned with preparation for unplanned events and natural disasters (Confederation of Finnish Industries (EK), 2014). Expressed simply, safety is related to internal threats, and security aspects to external threats. This study mainly concentrates on the safety point of view, but the security aspects should not be forgotten either, as both aspects occur in interaction with each other as shown. Figure 1 shows that production safety, work safety, and environment protection and safety are clearly safety functions, but some others – person safety and rescue operations – contain a strong safety aspect. Although this research focused on flight safety and studies safety deviations within aviation, all the components presented are relevant parts of enterprise safety and security and should be taken into account as part of the safety and security management of any aviation organisation, for instance, an airline, a maintenance company, an authority unit, etc. All these can be monitored by collecting and analysing data that could help management to discover possible holes in safety walls.

The collection of intelligence information is rather simple today – as almost all the information exchanged between both individuals and organisations (phone calls, SMS, e-mail) is in electronic form, particularly the daily usage of smart phones with direct access to cloud services. Contrary to the easy collection, a substantially larger problem is filtering out relevant information from the vast amount of collected data, which requires sophisticated analysis methods as well as other resources like a large number of highly skilled specialists. Although IT security has been mentioned in this context, there is no doubt that there is a significant safety aspect to protecting the information resources of an organisation from internal threats.

When looking at the Figure 1, one can ask: where is the essential component of product safety? From the point of view of this study, the product is the airline services and its safety as well its security can be seen as consisting of all the components presented in the figure.

2.2 The global development of aviation and flight safety

Human beings have always dreamt of flying. This is reflected in the history of mankind that tells us colourful, although fictitious stories of attempts to reach the sky or even to fly. Unlike the products of the imagination, 1000 B.C. Chinese ‘really’ could fly, although anchored to land, by constructing kites, which carried scouts on reconnaissance missions (NASA, 2002). The beginning of military aviation can be seen in these operations, although their contribution was clearly minimal. Another real attempt to fly, from the top of a tower, is documented in 1162, when a man in ancient Constantinople fashioned sail-like wings from a fabric gathered into pleats and folds. This effort also had a sad ending as he plummeted to his death (NASA, 2002).

Leonardo da Vinci (1452–1519), an Italian Renaissance polymath and perhaps the greatest inventor of all times, considered the most diversely talented person ever to have lived, was also interested in flying. In 1488, he also draw a design for a flying machine, which, as human-powered, would not have been able to fly (Kemp, 2006).The first real and successful free flight of man happened on 21 November 1783 in Paris when two men, Pilâtre de Rozier and Marquis d'Arlandes, flew a distance of nine kilometres in 25 minutes in a hot-air balloon constructed by Montgolfier brothers Joseph and Étienne (Lecornu, 1903). The following remarkable flights occurred when, in 1891, Otto Lilienthal performed the first well-documented, repeated, and successful gliding flights in Germany in gliders of his design (Brütting, 1938). He made more than 2,000 flights before, unfortunately, crashing to his death in 1896 (Waßermann, 1983). The beginning of modern aviation can be stated to be the first powered flight on 17 December 1903 by Orville Wright in Kitty Hawk, North Carolina, the USA (McFarland & Renstrom, 1950).

The statistics for airline traffic in the United States, presented in Table 2, illustrate the development of air travel. A growing trend can be clearly seen, concerning both passenger traffic and freight transport almost doubling during the rather short time period of eighteen years from 1987 to 2005 – freight transportation actually tripled during that period.

Table 2: Air traffic development in the United States 1954 to 2014 (Research and Innovative Technology Administration, 2011; United States Department of Transportation, 2014)

Year	Aircraft hours flown[1] (000)	Revenue Passenger ton miles[2] (000, all services)	Revenue passenger enplanements[3] (scheduled service, 000)	Aircraft revenue departures[4] (scheduled service)	Available seat miles[5] (non-scheduled service, 000)
1954	n/a	n/a	35,447	3,002,576	n/a
1957	n/a	n/a	49,423	3,768,861	n/a
1960	n/a	3,850,303	62,258	3,852,906	1,405,478
1963	n/a	5,104,030	77,403	3,788,362	3,204,388
1966	n/a	8,578,085	118,061	4,373,597	9,126,822
1969	n/a	13,996,087	171,898	5,378,343	18,871,030
1972	n/a	16,401,499	191,349	5,046,438	11,703,143
1975	n/a	17,332,436	205,062	4,704,710	12,816,993
1978	n/a	23,699,762	274,716	5,015,939	12,362,470
1981	n/a	26,006,304	285,270	5,211,867	13,880,966
1984	n/a	31,950,400	343,264	5,448,150	18,780,747
1987	10,645	41,786,983	447,307	6,581,309	22,150,055
1990	12,150	47,223,558	465,557	6,923,593	19,835,684
1993	12,706	50,599,323	487,249	7,245,395	22,318,046
1996	12,582	59,611,994	581,201	8,230,322	24,649,701
1999	13,966	75,489,431	635,402	8,616,167	23,911,314
2002	14,731	73,351,146	612,872	9,168,877	34,736,777
2005	18,154	79,511,748	738,628	11,561,951	25,987,964
2006	17,969	89,311,730	744,723	11,267,651	21,195,456
2007	18,404	91,901,919	769,623	10,668,826	22,406,236
2008	17,890	89,961,964	743,314	10,138,661	19,508,619
2009	16,642	84,782,185	703,899	9,550,799	18,358,703
2010	16,725	89,019,529	720,495	9,509,299	18,966,932
2011	16,943	90,337,147	731,117	9,513,793	19,245,807
2012	16,740	90,868,776	736,702	9,359,634	18,813,984
2013	16,760	91,541,210	739,462	9,175,249	15,034,455
2014	16,668	94,212,353	756,575	8,962,704	13,812,672

[1] The year 2014 only from January to September
[2] The database definitions have changed slightly since 1999
[3] The years 2011-2014 from October of the previous year to September
[4] The years 2007-2010 from February to January of the next year, 2011-2014 from October of the previous year to September
[5] The years 2009-2010 from February to January of the next year, 2011-2014 from October of the previous year to September

As this research has been conducted over a period of several years, the sources of the statistics have changed and that is why describing world air traffic since the end of the 1990s continues in slightly differently and problematic ways. As for benchmarking, global air traffic development including both domestic and international travel has been taken into account and is displayed in Table 3. In the last column[6] the air traffic of those 129 countries is displayed, showing the routes on which there has been at least one 'accident' (Čokorilo, De Luca, & Dell'Acqua, 2014). These statistics were added to the research in order to benchmark the development of air traffic in the US – as presented in Table 2.

Table 3: Total international and domestic world traffic 2000 – 2014 (Čokorilo, et al., 2014; ICAO, 2015)

Year	Passengers (million)	Increase or decrease per cent	Passenger/state (million)
2000	1,609	7.0	12.48
2001	1,640	1.9	12.70
2002	1,639	-0.1	12.70
2003	1,691	3.2	13.10
2004	1,888	11.6	14.60
2005	2,022	7.1	15.70
2006	2,127	5.2	16.50
2007	2,303	8.3	17.90
2008	2,367	2.8	18.30
2009	2,358	-0.4	18.30
2010	2,563	8.7	19.90
2011	2,865	6.1	n/a
2012	2,998	4.6	n/a
2013	3,132	4.5	n/a
2014	3,303	5.5	n/a

Despite the fact that flight accidents have occurred throughout the history of aviation, flying has been one of the safest modes of transport since the concept of air travel has existed – at least in the form understood today. Amalberti et al. (2005) have illustrated the risk levels of, among others, several means of transportation displayed in Figure 2. In that presentation, commercial, large-jet aviation has been placed in the category of the so-called ultra-safe systems with a risk

[6] The values between 2000 and 2010 are from Čokorilo, et al., 2014 and between 2011 and 2014 from ICAO, 2015

level of 10^{-6}. According to Novotny (2007) and Brooker (2006) the mythical barrier of one accident per 10 million (10^{-7}) flights has been reached by air transportation.

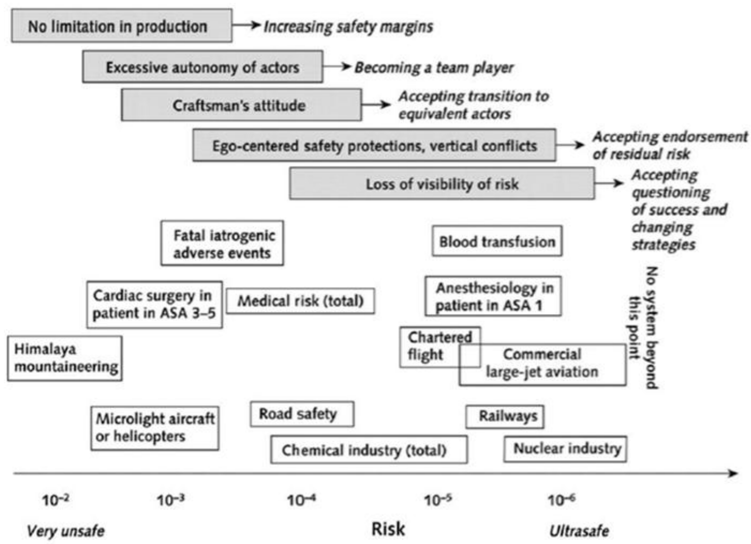


Figure 2: The asymptotes of safety excellence (adopted from (Amalberti, et al., 2005, p. 758))

Statistics about accidents and accident rates concerning the scheduled and non-scheduled services of airlines are presented in Table 4 showing accidents, fatalities and their rates between 1987 and 2006 for US Air Carriers Operating Under 14 CFR 121[7] (Federal Aviation Administration (FAA), 2013), Scheduled and Non-scheduled Service (Airlines) (National Transportation Safety Board, 2008, 2015). Fatal accidents occur in both major and serious accidents, depending on the definition of the case. The definitions of the National Transportation Safety Board (NTSB) classifications are as follows:

- Major – an accident in which any of three conditions is met:
    - a part 121 aircraft was destroyed, or
    - there were multiple fatalities, or
    - there was one fatality and a part 121 aircraft was substantially damaged
- Serious – an accident in which at least one of two conditions is met:

[7] Part 121 defines scheduled air carrier operations in the U.S.

- there was one fatality without substantial damage to a Part 121 aircraft, or
- there was at least one serious injury and a Part 121 aircraft was substantially damaged

- Injury – a nonfatal accident with at least one serious injury and without substantial damage to a Part 121 aircraft
- Damage – an accident in which no person was killed or seriously injured, but in which any aircraft was substantially damaged.

Table 4 displays the situation in the US, but the information can be generalised to cover at least Western Europe, too.

Table 4: Accidents, Fatalities, and Rates 1987 to 2014 (National Transportation Safety Board, 2008, 2015)

Year	Accidents		Flight Hours[8]	Miles Flown	Accidents per 100 000 Flight Hours		Accidents per 1 000 000 Miles Flown	
	All	Fatal			All	Fatal	All	Fatal
1987*[9]	34	5	10,645,192	4,360,521,000	0.310	0.038	0.0076	0.0009
1990	24	6	12,150,116	4,947,832,000	0.198	0.049	0.0049	0.0012
1993	23	1	12,706,206	5,249,469,000	0.181	0.008	0.0044	0.0002
1996	37	5	13,746,112	5,873,108,000	0.269	0.036	0.0063	0.0009
1999[10]	51	2	17,555,208	7,101,314,000	0.291	0.011	0.0072	0.0003
2002	41	0	17,290,198	7,192,501,000	0.237	-	0.0057	-
2005	40	3	19,390,029	8,165,643,000	0.206	0.015	0.0049	0.0004
2007	28	1	19,637,322	8,315,905,000	0.143	0.005	0.0034	0.0001
2009	30	2	17,626,832	7,465,598,000	0.170	0.011	0.0040	0.0003
2011	31	0	17,962,965	7,713,557,000	0.173	-	0.0040	-
2012	27	0	17,722,235	7,659,906,000	0.152	-	0.0035	-
2013	23	2	17,627,600	7,562,280,000	0.130	0.011	0.0030	0.0003
2014[11]	28	0	17,599,000	7,656,814,700	0.159	-	0.0037	-

[8] Flight hours, miles, and departures are compiled by the Federal Aviation Administration.
[9] Years followed by the symbol * are those in which an illegal act was responsible for an occurrence in this category. These acts, such as suicide, sabotage and terrorism are included in the totals for accidents and fatalities but are excluded for the purpose of accident rate computation.
[10] Since March 20, 1997, aircraft with 10 or more seats used in a scheduled passenger service have been operated under 14 CFR 121.
[11] 2014 data are preliminary

A rough regional distribution between 1997 and 2013 is displayed in Table 5, which is based on the fatal accidents of scheduled operations (displaying both accident rates and passenger fatality rates per million departures); and in Table 6 the scheduled commercial air transport average fatal accident rate per million flights between 2004 and 2013 is presented.

Table 5: Accidents and passenger fatalities per million departures 1997 to 2006 (ICAO, 2007)

	Accidents per million departures		Passenger fatalities per million departures	
	1997–2001	2002–2006	1997–2001	2002–2006
World	0.9	0.5	33	24
Latin America & Caribbean	2.0	1.2	40	46
North America	0.4	0.1	10	2
Asia Pacific	2.3	0.6	99	35
Middle East	1.0	1.9	74	70
Africa	2.0	4.6	75	270
Europe	0.6	0.3	23	20

Table 6 expands the period under consideration further from the latest period of 2002–2006 presented in Table 5.

Table 6: Air transport average fatal accident rate per million flights 2004 to 2013 (European Aviation Safety Agency, 2014; TraFi, 2014c)

	Fatal accident rate per million flights 2004–2013
Europe EASA member	1.8
North America	1.9
Asia Pacific	5.8
Asia	6.3
Central America	11.1
Middle East	15.5
South America	16.9
Europe non-EASA member	28.8
Africa	38.3

The significant worldwide variations in the trends can be clearly seen in both of the tables: the trends of accident rates per flight hours or miles flown show a decrease in most regions. However, these tables must be read very carefully. First, there are differences in the magnitudes of accident rates, as, for instance the ICAO 2007 statistics show significantly smaller numbers than those provided by EASA 2014. The explanation for this is that different events have been added to the calculations. Nevertheless, the core idea behind presenting these two tables – to illustrate the globally uneven distribution of flight safety – remains. The second point is that there also are significant differences within the areas observed. For example, in Table 5, Europe is observed as an entity, but in Table 6 there is a significant difference between the countries belonging to EASA and those not. Also other statistical sources show, that there is a significant difference between these two areas. The time period between 1991 and 2005 gives the rate for Western Europe to be as low as 0.46 compared to the corresponding value for Eastern Europe, which is 5.00. As for EASA operators in commercial operations, the same rate seems to be more than 50 per cent lower than the worldwide rate (Inspectie Verkeer en Waterstaat, 2007). Expressed on a general level, flight safety is highest in the ‘Western world’, i.e. Western Europe, North America and Australia. The regions that are more problematic are Eastern Europe, Africa and Latin America.

There are, surely, several reasons why the accident rates vary so much between different parts of the world. The first that comes to mind is that safety cultures vary between countries and airlines. Theoretically, differences should not exist because air traffic management rules are international and almost all countries are ICAO members, thus they have adopted its regulations and recommendations. Another noticeable aspect is that there are only a few manufacturers of airliners in the world and their standards are practically equal. From this aspect there should be no differences between the different regions and countries because all airlines are customers of these companies and their training and maintenance programmes. However, significant differences exist and the problem should be solved efficiently. One solution would be large scale knowledge and information interchange between actors in the airline industry.

Figure 3 illustrates the development of the accident rates and on-board fatalities on the worldwide commercial jet fleet from 1959 through to 2013. On the left-hand side is the annual accident rate per million departures and on the right-hand side the absolute number of the annual on-board fatalities.

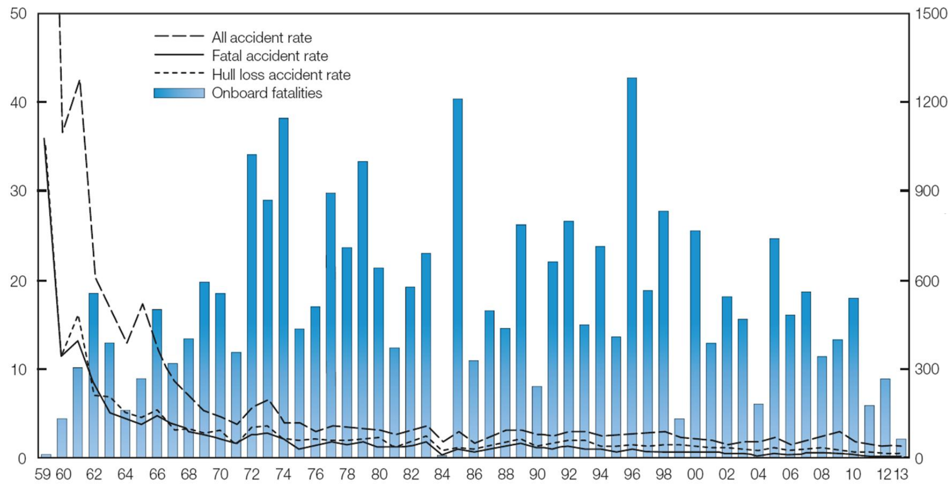


Figure 3: Accident rates and on-board fatalities 1959 – 2013 (adopted from (Boeing Commercial Airplanes, 2014, p. 16))

Figure 3 illustrates that since the middle of the 1970s, the trend in the absolute annual number of both non-fatal and fatal accidents of the worldwide commercial jet fleet has stabilised. It also indicates that the trend concerning the worldwide development of both non-fatal and fatal accidents is coherent with the trend presented in Table 4.

The trends presented above appear to be global, at least among developed aviation countries: almost all statistics on this theme show the same development. If one looks at different commercial airlines' statistics or even military aviation operation/accident rates, the view is very similar (although Figure 3 does not, however, include military operations). The author participated in the meeting of the Global Aviation Information Network (GAIN, see Chapter 3.2.2) Working Group B in Houston, US, in January 2005, where statistics about carrier flight operations in the United States Navy were made available in a data analysis demo and matched those of commercial flights. The same trend is shown by the aircraft maintenance training CD of Continental Airlines, whose fleet at that time consisted of about 350 aircraft (Continental Airlines, 2004).

There is no clear single explanation for the obvious rise in flight safety since the 1960s, but the phenomenon is generally known in the aviation world. It could be, and obviously is, a combination of several factors, such as developed technology and analysis methods, an increase in skills and knowledge resulting from better possibilities for information change, a change of attitudes towards safer behaviour and so on. One essential factor in this development is a change in thinking. During the first 50 years in aviation, the focus was on technology because it was challenging and that did not leave much space for other concerns.

During that time seeing aviators as superhuman was also emphasised. The main focus of aviation technology was on physiology, which contained many things to master. The mental performance capacity was determined in moralistic terms. The change process began in the early 1960s when aviation technology reached a sufficient degree of maturity, enabling a moving of the focus of aviation accidents onto individual human error. This led to the forming of training programmes based on the awareness of the tendency of human beings to make errors. As a means to deal with human error, checklists and standard operating procedures were systematically taken into use. This process can simply be described as the ‘death of narcissism’, meaning the breaking of the illusion of the invulnerability of human beings. One essential detail to be added to the list is the development in decision-making aids, such as fault situation instructions (Sorsa, 2012).

The development of aviation safety culture as a concept was based on practical issues. It was first used in the 1980s in the investigations into the Chernobyl nuclear disaster to illustrate that accidents do not only arise out of technical failures or human errors made by individual people (IAEA, 1991). The implementation of this concept was aimed at presenting the factors in management, organisations, work communities and even society that affect the causes of the accidents/affect the accidents occurring (Reiman, Pietikäinen, & Oedewald, 2008). The 1990s developed thinking about and the discovery of major systems that focus on the analysis of background components and the seeing of totalities consisting of individual systems. The next step, after the turn of the millennium, was observing how flight crew functions with respect to the current action and safety culture of their operational environment (Sorsa, 2012). At the turn of the millennium, the development of traffic and accident volumes from the beginning of the 1960s were analysed by Amalberti (2001). They were then predicted from the year 2000 up to 2010. This is presented in Figure 4, which displays the number of commercial jet accidents, worldwide accident rate and traffic growth. On the left-hand side, the departures in millions as well as the accident rate per million departures are shown. On the right-hand side, the number of accidents is shown.

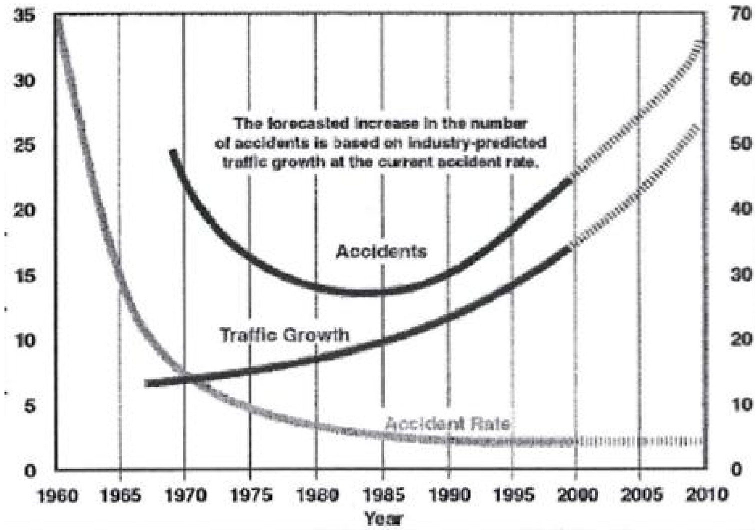


Figure 4: Frequency of accidents and future risks (adopted from (Amalberti, 2001))

The curves clearly illustrate the expected trend. As seen, although the accident rate remained at the same level, the number of accidents was predicted to increase remarkably due to the strong growth in aviation traffic. Figure 4 and Figure 5 can be observed together as the time dimension of the latter completes the presentation because it extends the time dimension beyond the former. It also supports Amalberti’s forecast concerning the development of the first decade of the millennium. In Figure 5 Saatsi et al. (2011) present the development of air traffic volumes and flight safety. The light blue graph bars describe the ratio between accidents and flights and the green curve presents the number of flights. The trend is clear: there were almost 12 accidents per million flights in 1970, but the correspondent number was less than two in 2009. During the same period, global air traffic grew from about two to 25 million flights per year.

Figure 5: The development of air traffic and flight safety (adopted from (Saatsi, et al., 2011, p. 16))

As presented in Table 4, the absolute annual number of fatal accidents in the US has varied between 0 and 6 but its trend significantly decreased between the years 1987 and 2005. This positive development is due to intensive flight safety work and is observable in the comparison of all accidents with fatal accidents in relation to flight hours and miles flown, in which both rates are decreasing, but fatal accidents are decreasing at a faster rate than that of all accidents. It is worth noticing that because the all accident rate includes fatal accidents as well, there is a slight autocorrelation between the trends compared. Its impact, however, can be regarded as a very minor one, so, in this case there is no need to calculate the trends of all accidents without fatal accidents.

Figure 6 displays the absolute number of fatal accidents of all civil aircraft with 19 or more passengers between 1950 and 2014, while Figure 5 illustrates the situation concerning the global commercial jet fleet. Due to the different data used in the statistics, the figures are not coherent, but decreasing trends can easily be determined.

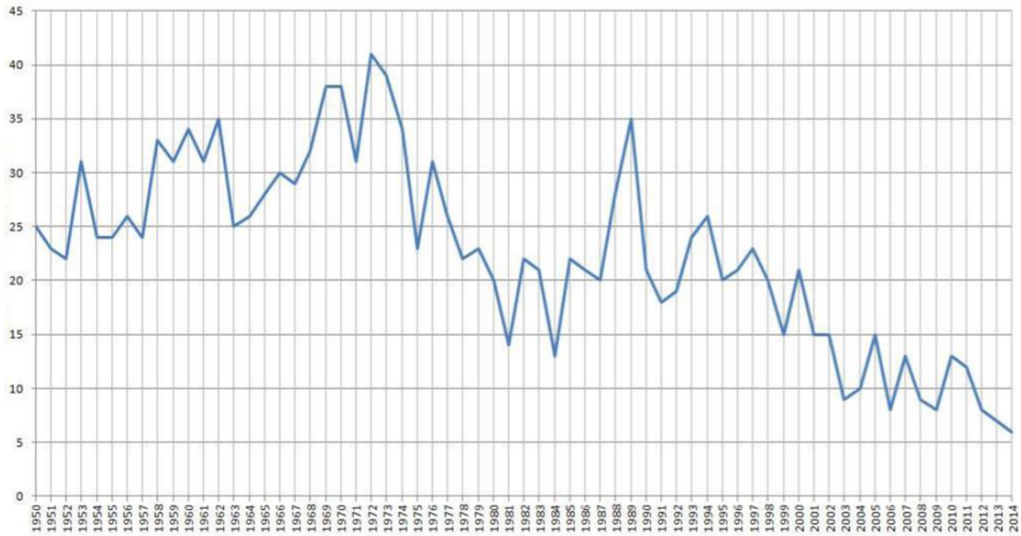


Figure 6: Number of fatal accidents 1950 to 2014 (adopted from (PlaneCrashinfo, 2015))

Figure 7 presents a similar, rather evenly decreasing development for cargo air transportation. The blue parts of the columns describe the cargo operations and the tan parts fatal accidents. The black line shows the total accident trend.

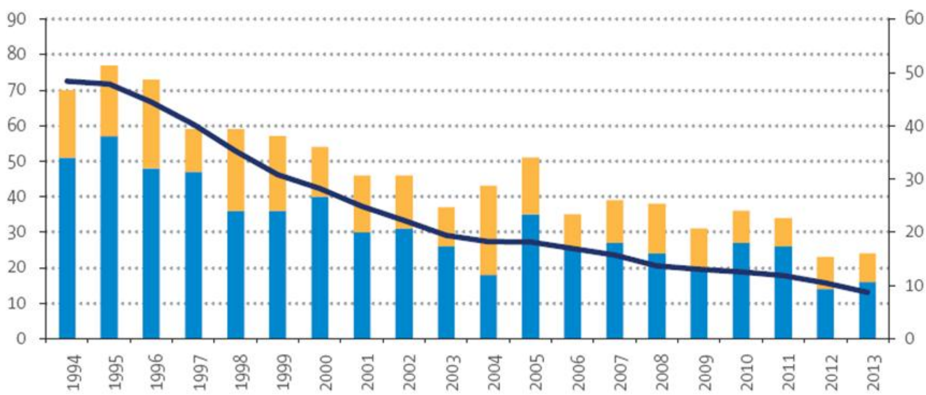


Figure 7: Worldwide number and rate of fatal accidents per 10 million flights 1994 to 2013 (adopted from (European Aviation Safety Agency, 2014, p. 14))

In conclusion, all the figures and tables presented above indicate rapidly decreasing trends in fatal accidents compared with the development of all accidents thus they show that development is going in the right direction. However, although the number of accidents measured proportionally to the volume of air traffic is decreasing, the total amount of accidents is seeing a moderate increase, this situation is unacceptable and the ratio between accidents and air traffic volume needs to be decreased. The present situation can be seen in the trend where the total number of all accidents is measured according to the amount of flight hours and miles flown. The trend is decreasing slightly but the development still cannot be considered adequate.

2.3 The development of aviation and flight safety in Finland

The global examples have naturally encouraged significant attempts to reduce accidents within Finnish aviation. The first recorded aviation incident in Finland occurred when Swedish balloonist August Andrée due to the strong wind was stranded, during a hot-air balloon flight, in Korppoo, in the Turku archipelago in 1894 while travelling from Sweden. He was rescued from a bald islet and brought to Turku where he was received as a hero. His friend and colleague Austrian-Norwegian Francesco Cetti transported his balloon, named Skandinav, to Helsinki and Turku in 1907, to fly public attracting great attention (Lento, 1979). The World Air Sports Federation, FAI (Fédération Aéronautique Internationale) was founded as early as 1905 (Fédération Aéronautique Internationale, 2014) and ,as the development of aviation occurred rapidly in the first decade of the 20th century, it was obvious that the aviation and the airplanes would soon reach Finland.

The Finnish Aero Club was founded in 1919 and its international coordination began two years later when the club was accepted as member of FAI. Science has naturally always played a significant role among military aviation also in Finland, but civil aviation was further extended in 1921, when the Finnish Aero Club joined the Deutscher Wissenschaftlicher Verein (German Scientific Association) (Uola, 1994).

One significant milestone in Finnish aviation and especially in air traffic was reached when a company named Aero O/Y, later Finnair, was founded in 1923; first operating with float planes. Continuing its operations uninterrupted since then, it is the fifth oldest airline in the world and has remained the major air traffic operator in Finland, being ranked one of the safest airlines as it has not suffered fatal[12] or hull-loss[13] accidents since 1963.

---

[12] The fatal accident definition (AIRBUS 2014) is "An event in which at least one passenger or crew member is fatally injured or later dies of his/her injuries".

The growth in Finnish air traffic did not begin substantially until the 1930s, when the first ground based civil airfields were built in the middle of the decade. Alongside road, rail and water traffic, air transport had risen and improved the internal connections of Finland and links with foreign countries. By the end of the decade, Aero had extended its international traffic to Central Europe stretching its route from Tallinn to Berlin. The war era (1939–1945) set its own limitations on civil air traffic, which nevertheless continued in late summer of 1945. Shortly after the war, Finland began to participate in international civil aviation coordination. It became a member of the International Civil Aviation Organization (ICAO, see Chapter 2.4.2) in 1949, which is when international air safety work can effectively be seen as beginning in Finland. Although the roots of ICAO go back to 1944, it did not officially begin until 1947. In 1954 Finland joined the ECAC[14]. The 1950s, when the Helsinki Olympic Games in 1952 provided a significant boost to air traffic, the volume of air passengers multiplied seven-fold from 130,000 to more than 900,000 (Civil Aviation Administration Finland, 1997).

The jet age in the 1960s opened transatlantic flight connections and dramatically cut long domestic flight times. During the decade, the total number of both domestic and international passengers grew from 900,000 to 2.1 million. In the 1970s, the traffic volume expanded 2.3 times exceeding five million passengers with more than 62,000 commercial flights each year. Record passenger figures were reached in the 1980s and the six million mark was exceeded in 1984. Since then, air traffic has grown at an annual rate of eight per cent and more than 11 million passengers on internal flights was reached in 1990 (Civil Aviation Administration Finland, 1997).

During the 1990s there was a minor drop in the number of passengers – in 1993 the total volume was about 8.5 million. After that, however, the volumes continued growing and in 1997 the previous record from 1990 was broken and in the following year almost 13 million passengers flew (Civil Aviation Administration Finland, 1999). The 13 million passenger limit was broken in 1999 and from the end of the 1990s there was a steady increase until 2008, when the number of passengers reached 17.5 million (Finavia, 2009). A slight decrease to 16 million occurred in 2009 but more than 19 million passengers travelled by

---

[13] Hull loss means (Boeing 2014) that "the airplane is totally destroyed or damaged and not repaired, also including, but not limited to, events in which the airplane is missing, the search for the wreckage has been terminated without the airplane being located, or the airplane is completely inaccessible".

[14] ECAC stands for European Civil Aviation Conference, an organisation founded in 1954 as a forum for European countries in order to develop a joint aviation policy and cooperation on technical matters.

air in 2011, although the figure for 2013 showed a very minor decrease to just under 19 million (Finavia, 2014).

The geographical location of Finland – set apart from continental Europe like an island – increases the major importance of air traffic for the country. Air traffic is a very dynamic domain containing plenty of factors for change that influence it in both the short and long term. As the distance between Finland and continental Europe is rather long, air traffic cannot be efficiently substituted with other means of transport. Different scenarios raise the significance of international cooperation and the development of Finland's economy from the point of view of the future of air traffic. The development of business and leisure travel by air partly affects different factors and it is obvious that also the Finnish air traffic will be different depending on these factors. Thus, global cooperation will become increasingly important in the future. The demand for air travel is increasing although international communications can partly be substituted by new virtual applications (Aalto, Pöllänen, Mäntynen, Mäkelä, & Rauhamäki, 2012).

In Finland, the development of flight safety reflects the result of decades of successful analytical studies on aviation safety both globally and in Finland. Since the beginning of the 1960s until today, the accident rate has been reduced to a very small percentage of what it was and is stabilising due to systematic safety work. The statistics of fatal accidents in general aviation and gliding (including motor gliders) is illustrated in Figure 8. It presents the average number (the scale on the left-hand side) of fatal accidents per 100,000 flight hours in five-year periods.

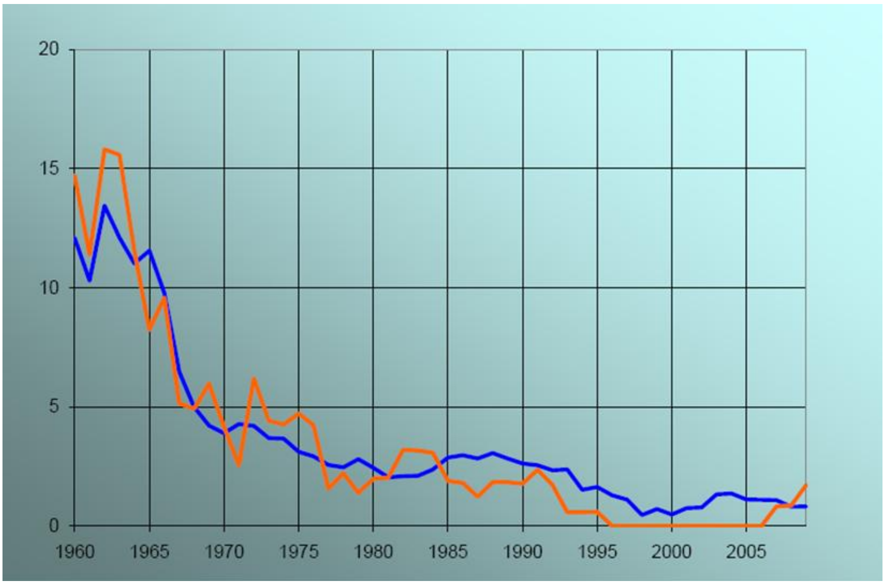


Figure 8: Fatal accidents in general aviation (blue) and gliding in Finland (brown) 1960 to 2010 (adopted from (TraFi, 2011))

As displayed in Figure 8, the development of general aviation and gliding has been propitious in Finland, especially for gliding. There was a ten-year period between 1996 and 2006 with no fatal accidents at all. Also the curve describing the development of general aviation shows a coherent decreasing trend. Unfortunately, since then the situation has begun to worsen. Figure 9 illustrates the number of accidents and Figure 10 the serious incidents per 10,000 flight hours in the different types of aviation in Finland between 2005 and 2013.

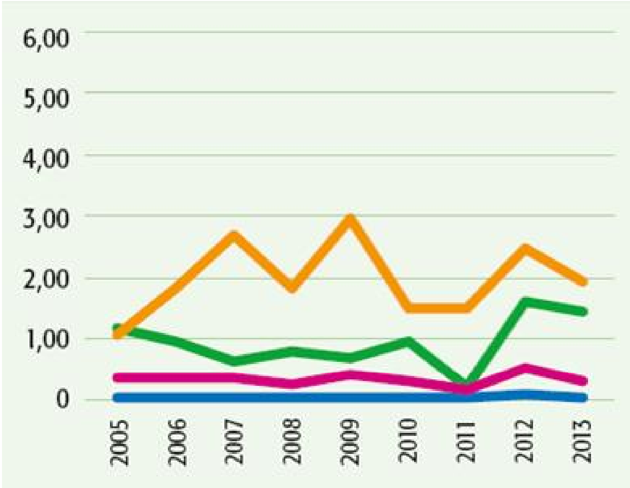


Figure 9: Accidents per 10,000 flight hours in Finland 2005 to 2013 (adopted from (TraFi, 2014c, p. 19))

In both figures, the yellow line displays the development of sport aviation (gliding, motor gliding, ultra-light aircraft and hot air balloons), the green one general aviation, the blue one commercial air transport and the purple represents all the different types combined.

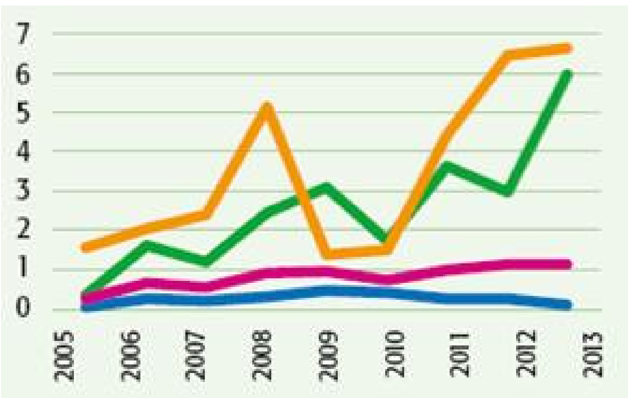


Figure 10: Serious incidents per 10,000 flight hours in Finland 2005 to 2013 (adopted from (TraFi, 2014c, p. 19))

Figure 9 shows that the decreasing accident rate development presented in Figure 8 did not continue during the period presented here because both general aviation and sport aviation seem to have a moderate upward trend. As for serious incidents, presented in Figure 10, the situation is much worse as the trends are increasing. One explanation, but not the only one for the adverse development, is the increase in the number and flight operations of ultralight aircraft (TraFi, 2014c). The development of the flight hours of general aviation, gliding and ultra-light aviation in Finland between 2005 and 2013 is presented in Table 7.

Table 7: Flight hours for general aviation, gliding and ultra-light aviation in Finland 2005 to 2013 (TraFi, 2014b)

	General aviation	Gliding and motor gliding	Ultra-light aviation
2005	70,755	26,021	12,128
2006	77,813	26,038	12,841
2007	76,149	20,798	12,686
2008	72,275	20,439	12,586
2009	69,727	23,662	13,357
2010	67,572	19,576	13,589
2011	66,417	20,520	13,344
2012	54,215	18,789	13,785
2013	42,088	18,244	13,294

Gliding also includes motor gliding hours as in Figure 8. Both commercial flight hours and private flight hours are combined in the figures for general aviation. The figures shown in Table 7 confirm the urgent need for corrective mechanisms in order to augment the flight safety of non-commercial aviation in Finland.

A strongly decreasing trend can also be seen in the flight safety statistics of the Finnish Air Force. Since 1960 there has been a significant reduction in the amount of flight accidents in relation to the number of flight hours. Furthermore, from the middle of the first decade of the new millennium, the rate has diminished even further, although more moderately (Koho, 2005). This is displayed in Figure 11.

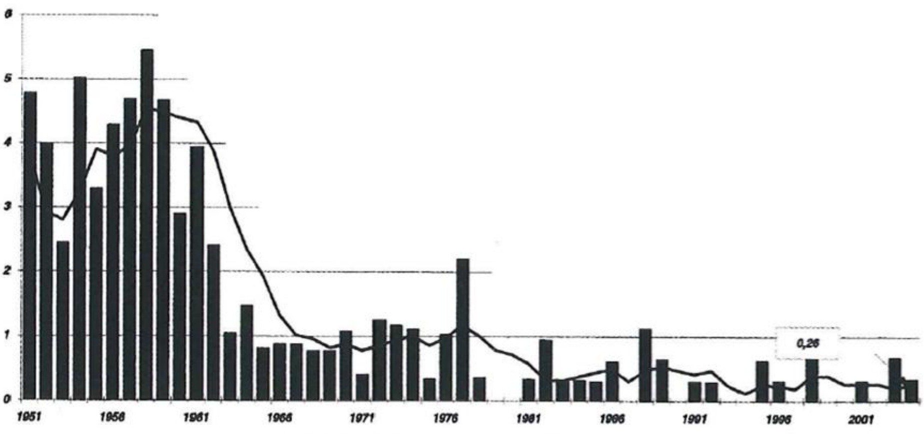


Figure 11: Flight accidents 1950 to 2004 in the Finnish Air Force (adopted from (Koho, 2005))

As similar documents about flight accidents in the Finnish Air Force were not accessible, the absolute number of flight accidents between 2000 and 2014 is displayed in Table 8.

Table 8: Flight accidents in the Finnish Air Force 2000 to 2014 (Ilmavoimat, 2013, 2014)

Year	Number of accidents	Year	Number of accidents
2000	-	2008	-
2001	1	2009	-
2002	-	2010	1
2003	2	2011	1
2004	1	2012	-
2005	-	2013	1
2006	1	2014	-
2007	1		

Combining the information provided in Figure 11 and Table 8 shows that the high flight safety level achieved in the 1990s has been maintained. In the flight safety report of 2004, the Finnish Air Force lists the ten most important hazard factors, the first of which is the continuously increasing requirement for effectiveness and efficiency in relation to the resources available (Koho, 2005).

Regular air traffic in Finland has not been taken into discussion, because after the Second World War, only three fatal flight accidents have occurred: two at the beginning of the 1960s and one in 2005, thus it is impossible to present a trend. Fatal accidents among chartered flights have been almost as rare, making the presenting of a trend unnecessary.

2.4 International Aviation Coordination

2.4.1 The beginning

As early as at the beginning of the 20th century, when the first airplanes appeared in the sky, people with foresight noticed that the advent of the airplane changed ways of thinking about transporting people, i.e. people could no longer be contained within strictly national borders. A number of basic principles governing aviation were laid down in Paris in 1910, when, by the invitation of France, the first conference, attended by 18 European states, was held on formulating an international air law code. During the years between the two world wars, continuous growth marked civil aviation, both in the commercial and the technical fields. In spite of that, flying was not available as a means of transport for the masses but remained an exclusive personal transport form (ICAO, 2004b).

The time between the two world wars provided a good period for organising aviation activities and their components more carefully and with better resources. From the very beginning of aviation the consequences of incidents and accidents, as well as other so-called deviations were taken seriously; this called for an effective method for avoiding or at least reducing them. In early 1928, the National Advisory Committee of Aeronautics (NACA) was asked by the Assistant Secretaries for Aeronautics in the Departments of War, Navy and Commerce in the United States to develop a common approach to the analysis and reporting of aircraft accidents. Following the request, the Special Committee on the Nomenclature, Subdivision, and Classification of Aircraft Accidents was organised by NACA (Johnson & Holloway, 2007).

An initial report was published by the NACA committee in August 1928. This was “undertaken in recognition of the difficulty of drawing correct conclusions from efforts to analyze and compare reports of aircraft accidents prepared by different organizations using different classifications and definitions” (NACA, 1928). Although these activities took place in the US, some international and standardised reporting and analysis can be seen here, because this report proposed an investigation method both to the representatives of the US and to the representatives of the governments of Britain, France, Italy and Japan. The report can be regarded as very revolutionary, because it included a series of definitions for aircraft accidents: “An aircraft accident is an occurrence which takes place while an aircraft is being operated as such and as a result of which a person or persons are injured or killed, or the aircraft receives appreciable or marked damage through the forces of external contact or through fire” (NACA, 1928). This definition includes both reasons stating why an individual event can be called an accident: either a person or several persons are injured or killed, or the aircraft is damaged to a certain extent. Of course both of these may happen at the same time (and often, unfortunately, do). Hence, the definitions continue by noting that “a collision of two or more aircraft should be analyzed and reported statistically as one accident”, a principle laid down that time (Johnson & Holloway, 2007).

These definitions and principles are usefully comparable with the International Civil Aviation Organization’s (ICAO) definition of an aircraft accident as “An occurrence associated with the operation of an aircraft which takes place between the time any person boards the aircraft with the intention of flight until such time as all such persons have disembarked, in which: a) a person is fatally or seriously injured as a result of: – being in the aircraft, or – direct contact with any part of the aircraft, including parts which have become detached from the aircraft, or – direct exposure to jet blast, *except* when the injuries are from natural causes, self-inflicted or inflicted by other persons, or when the injuries are to stowaways hiding outside the areas normally available to the passengers and crew; or b) the aircraft sustains damage or structural failure, which: – adversely affects the structural strength, performance or flight characteristics of the aircraft, and – would normally require major repair or replacement of the affected component, *except* for engine failure or damage, when the damage is limited to the engine, its cowlings or accessories; or for damage limited to propellers, wing tips, antennas, tires, brakes, fairings, small dents or puncture holes in the aircraft skin; or c) the aircraft is missing or is completely inaccessible” (ICAO, 2001).

Despite the human tragedies and horror caused by aviation in the Second World War, its rapid development during the war enabled the transportation of large numbers of people and goods over long distances. Ground facilities were also developed, permitting this in an orderly and expeditious manner (ICAO, 2004b).

2.4.2 International Civil Aviation Organization (ICAO)

In November 1944, the US government extended an invitation to 55 States[15] or authorities to attend an International Civil Aviation Conference in Chicago. The process resulted in the signing of a Convention on International Civil Aviation by 52 States of the 54 attending States. This convention set up the permanent International Civil Aviation Organization (ICAO) to secure the highest possible degree of uniformity in regulations and standards, procedures and organisation in international cooperation on civil aviation matters (ICAO, 2004a).

The existence of the ICAO enabled the formal and, to a certain degree, standardised thinking on incident and accident reporting, which is one of the key information solutions in the moves towards safer aviation. Since the beginning of aviation history there has been reporting and investigation concerning incidents and accidents, which have been defined in International Standards and Recommended Practices, Annex 13 (ICAO, 2010) to the Convention on International Civil Aviation, Aircraft Accident and Incident Investigation. The Standards and Recommended Practices for Aircraft Accident Inquiries were first adopted by the Council on 11 April 1951 pursuant to Article 37 of the Convention on International Civil Aviation (Chicago, 1944) and were designated as Annex 13 to the Convention. The Standards and Recommended Practices were based on the recommendations of the Accident Investigation Division at its First Session in February 1946 which were further developed at the Second Session of the Division in February 1947 (ICAO, 2001).

2.4.3 The ADREP classification

The ADREP (Accident/Incident Data Reporting) classification, established in 1976, is maintained and operated by ICAO (see Chapter 2.4.2), including its Accident Investigation Section (AIG) (ICAO, 2014). It receives and stores occurrence data (in this context, including both incidents and accidents) from the member states, providing them with this data in order to assist them in validating safety. The system has evolved since the 1970s to meet the changes occurring in information technology and in the aviation industry (SKYbrary, 2013a). The definition used by ADREP is, "*The ADREP Occurrence class taxonomy is a set of terms used by ICAO to categorize occurrences by severity and allow safety trend analysis on these categories. The ADREP Occurrence class taxonomy is part of the ICAO accident data reporting system*" (SKYbrary, 2013b).

---

[15] Countries in this context

The basis of the system was laid when ICAO started to publish safety statistics from a data bank in 1951. At that time the system was called ADREP, standing for “Accident Data Reporting Experts Panel”. Known today as Accident/Incident Data Reporting, which originated from an Accident Investigation and Prevention Divisional Meeting (AIG/74). The core idea of its implementation (made by ICAO) was to centralise the safety data “on the circumstances and causes of accidents and incidents, as determined by national authorities, and to disseminate these safety data to Contracting States for prevention purposes” (ICAO, 2008a). Operationally the occurrence class taxonomy is a set of terms used by ICAO to categorise occurrences by severity and allow a safety trend analysis of these categories. The ADREP Occurrence class taxonomy is an inseparable part of the ICAO accident data reporting system (SKYbrary, 2013b).

The current (2015) version in use is ADREP2000 (SKYbrary, 2013a) and when published it contained 552 explanatory factors, representing the outcome of investigations throughout the world during a period of fifty years. This version is the third, after the preceding ADREP76 and ADREP87; the former contained 88 factors and the latter (used, for example, in VASA, see Chapter 2.4.4) 142 explanatory factors. The analytical discussions among both national and international investigation teams are significantly helped by the clear separation between the events and the causes as well as the compilation of the older causes that are included in the taxonomy. ADREP2000 can be seen as an excellent tool to explore as it contains an organised collection of all the identified factors and events that have led to an accident (Pouliquen et al., 2005).

Since its implementation the classification has been expanded and developed further. Compared with the previous versions, ADREP2000 now contains 621 attributes, thus making it significantly more accurate than its predecessors (ICAO, 2006). ADREP2000 has nine main classes of attributes, which contain several hundreds of sub-attributes. These are:

1. Country name
2. Separation
3. Weather
4. Aircraft
5. Aerodrome
6. Airspace
7. ATS Unit (Air Traffic Services Unit)
8. Recommendations
9. Management.

The first main category, Country Name, includes where the incident occurred or the nationality of the airplane involved, includes general information about the occurrence such as flight information, when and where the occurrence happened, the classification, severity, number of injuries and the relationship of air traffic management to the occurrence. This section also includes notes and a narrative of the occurrence, which explains details that cannot be expressed through lists. The second, Separation, includes separation information in cases where two aircraft have been involved in an occurrence. This contains information about the distances and movements of different aircraft as well as the traffic situation, visibility, avoidance actions and other similar data. The third category, Weather, contains all information concerning weather reports and weather conditions like temperature, clouds, light conditions, precipitation and other weather phenomena.

Section four, Aircraft, holds all the relevant information concerning the aircraft, its manufacturer and model, its identification, type and description, operation, failures, flight history, possible injuries and other information on the flight crew, passengers and other personnel – if involved in the occurrence. This section also provides meteorological data and data about air traffic services. In the Aerodrome section all the relevant information about the runways as well as their condition and traffic situation, for example, are made available if the occurrence took place during landing or take-off or during taxiing. Airspace, its type, name and class as well as the special activities that perhaps happened in it, is almost always required, and is only unused if the occurrence happens on the ground. ATS Unit contains information about air traffic services, which plays an important role in many occurrences.

Recommendations includes, as its name already suggests, all possible recommendations to be made by an authority concerning aircraft, equipment, personnel, the procedures used as well as potential factors and safety issues. This section cannot be filled in before sort of analysis has been done. Often, in case of deviations without any safety effect, no recommendations are written. In accidents the situation might be different, but not always. Management describes the management of the investigation process containing the occurrence report and the modifications made during the handling of the case, which begins with opening the occurrence file and closing it when everything has been completely investigated.

In addition to the first section, both data fields, narratives and notes are also included in the sections Aircraft, Aerodrome, Airspace and ATS Unit in order to add additional information, concerning the context of the case – if needed. Event menus describing the occurrence are found in the first section and also in the sections Aircraft, Aerodrome, Airspace and ATS Unit. As the menus are fixed and cannot cover all the possible cases due to the huge amount of variations in their combinations, the narratives and notes are needed.

The latest version, 2000, that replaced ADREP87 has formed the basis of ECCAIRS (see Chapter 2.4.4) since version 4. An even newer version of ADREP is being prepared. It will replace ADREP2000 (ICAO, 1987, 2000). The occurrence classes are presented in Appendix 3. The ADREP classification is not widely used by the airlines and other commercial operators. The reasons are primarily twofold. First, the classifications, both versions 87 and 2000, are in some aspects too accurate from their point of view. But that leads to the problem that, from the other perspective, several important items are absent. Thus, operators have often developed their own systems that are especially tailored to meet their needs. Such specific systems make it difficult to exploit or introduce other systems parallel to them or even replace them. Hence, ADREP has been made easy to understand because it makes no sense to replace a well-functioning system with a less suitable one. When normal resistance to changes in processes are taken into consideration, it can be argued that the difficulties in changing such systems are significant.

2.4.4 VASA and ECCAIRS

VASA (Vaaratilanteiden Seurantajärjestelmä, Control System for Hazard Events) was a database system constructed to replace NORDAIDS (Nordic Accident and Incident Data System) during the period when ECCAIRS (the European Coordination Centre for Aviation Incident Reporting Systems) was being prepared. NORDAIDS was used since the 1970s for essential information concerning all aircraft accidents in the Nordic countries that had been recorded into this common data bank. In addition, Finland closely monitored the European Union’s work towards developing a common and uniform classification and reporting system for recording incidents and occurrences, later known as ECCAIRS (Flight Safety Authority, 1999). VASA was a typical database system that was based on a relational model built using Microsoft Office Access 97 as database kernel that was complemented by a tailored user interface, consisting of 24 tables and eleven queries. Its structure was configured to use the ICAO ADREP87 classification. The Accident, serious incident and occurrence report form (LU3626e, version 2006/08[16]) for collecting flight safety data is presented as Appendix 2 (Finnish Civil Aviation Authority, 2006). In addition to the form mentioned, the airlines used their own forms for reporting flight safety events – both for internal and external use, in this case the flight safety authority received a copy. The user in-

---

[16] The version 2006/08 is presented here because it corresponded to the ADREP87 classification applied in VASA and was used also when ECCAIRS became the active system in June 2005

terface was built using Visual Basic, containing three forms (one of them a separate routine) for inputting, browsing and changing data in the records. The database contained about 16,000 rows in June 2005 when ECCAIRS was taken into use on the basis of EU directive 2003/42/EC on Occurrence Reporting. It has obliged EU member states to collect and exchange information about incidents and other occurrences since July 2005. Simultaneously, the ECCAIRS system was also taken into use by the Finnish CAA. Tests had, however, been done years before in order to achieve the required level of preparation for its operative application. These two were in parallel use until the end of summer 2005.

The safety-related cases had been collected and saved since 1994, when NORDAIDS was put aside as active system. It was, however, used later, but only for reporting. VASA theoretically had multi-user capabilities, but mostly because of Access, in practice it was a personal tool for one user. That is one reason why it could not be used as a widely allocated data source. The lack of reporting tools in Access increased the need to replace it as soon as possible.

The European Co-ordination Centre for Aviation Incident Reporting Systems (ECCAIRS) is the heart of a network whose objective is to integrate information from the aviation occurrence reporting systems being used in the EU member states. The mission of ECCAIRS can be described as "*to assist National and European transport entities in collecting, sharing and analysing their safety information in order to improve public transport safety*". ECCAIRS has a long story behind it. Over a period of six years, in four separate projects the feasibility of the approach was evaluated, a pilot system developed and a production environment set up. In 1989, a study in the field of accident investigation and incident reporting systems was started on behalf of the Commission of the European Community. One of the conclusions of the study suggested bringing together the knowledge derived from the collection of incident reports existing in the aviation authorities of various member states. As the existing systems were not compatible, the study recommended the setting up of a European co-ordination centre for mandatory incident reporting systems. In 1993 the Joint Research Centre at the request of Directorate General VII (Transport) of the European Commission began a feasibility study. The main objective of the study was the pilot implementation of an automated incident reporting system able to collect information from existing but incompatible sources. A secondary objective was to offer a solution to those member states that do not have an automated system at present; the system became ECCAIRS. In 1995 the feasibility of the ECCAIRS approach was demonstrated by integrating – on an experimental basis – more than 40,000 occurrences originating from Scandinavia, the United Kingdom and Germany. Currently ECCAIRS software can be used as an occurrence reporting system for EU authorities; the authorities have access to the integrated database of all EU occurrences. From 1995 until 1999 the production environment of the ECCAIRS network has been online. However, the software has been redesigned and made more robust – since the end of 1998 production releases of the software have been distributed to EU authorities and other interested parties. From the beginning of 1999 onwards the network was constructed with the objective of having all fifteen member states of that time participating by the end of 2002 (Joint Research Centre, 2004).

At the end of the implementation phase, the Steering Committee concluded *"that the feasibility has been demonstrated…. Should a political decision to implement a co-ordinated system be taken, …ECCAIRS could be the foundation for such a system. …ECCAIRS would be an essential building block in establishing such (advanced analysis) techniques as it has the capability to bring together vast amounts of data from several sources"*. A presentation seminar on ECCAIRS was organised by the EU Commission in March 1998 and strong support was expressed by the majority of the participants representing the entire air transport sector (Civil Aviation Authorities, Accident Investigation Bodies, Air Traffic Management, manufacturers, airlines, staff representatives, research institutes, safety specialists, etc.). As a European system, it greatly enhanced the effectiveness of national systems by allowing the establishment of a network that would permit easy access to data while covering a much larger region. Also, the future European aviation safety organisation required such data for its work (European Commission, 2003). The role of the central office of the ECCAIRS network is to collect, integrate and disseminate occurrence information originating from the so-called satellite offices of the participating countries. The central office implements a database that contains the integrated data. The ECCAIRS database has, throughout its history, been based on ADREP (see Chapter 2.4.3) as defined by the International Civil Aviation Organisation (ICAO).

The previous classification was used until version 3, ADREP87, and contains about 300 data fields (called attributes in ECCAIRS). Almost all of them contain values that can only be taken from a predefined list. This makes data entry easy, improves data consistency and allows for the flexible representation of data values in any of the supported languages. Some of the data fields can exist in multiple instances and some can have multiple values. In the ECCAIRS implementation this has been reflected in the relational structure of the ECCAIRS database schema. In order to exchange information in an efficient hardware and software independent format, a specific file format was defined, which can be used to exchange information in a secure and electronic way: the ECCAIRS Data Format (EDF). One occurrence is coded in a language independent way as a sequence of alphanumeric bytes. An .edf file can contain an unlimited number of occurrences. Once the structure of the EDF is understood, information can be extracted from it and can be used as required (imported, printed, visualised etc.). Within the scope of the ECCAIRS project, components like viewers, editors and printers are available for interpreting the EDF files (Joint Research Centre, 2003). As mentioned, the ECCAIRS system is fully compatible with ICAO's ADREP system and enables EU member states to report automatically, thus improving the presently far from satisfactory fulfilment of their reporting obligation under the Chicago Convention. It should also be noted that a number of third countries have expressed their interest in the ECCAIRS system and that the EU is in a position to lead the way in the standardisation of accident and incident reporting for the benefit of air safety in all areas of world (European Commission, 2000). The main ECCAIRS database is located in Ispra, Italy. All the development activities take place there.

ECCAIRS offers a standardised and rather flexible data collection basis for incidents and accidents. It also includes report tools for information representation, data exchange and analysis. It began in the 1990s, when the Commission of the European Community started a study in the field of accident and incident investigation. The basic idea was to collect all the experience and knowledge of the member states on aviation safety. The systems of civil aviation authorities (CAA) and other organisations appeared not to be compatible, consequently it was recommended that a European co-ordination centre for mandatory incident reporting systems be set up, which started in 1993 at the Joint Research Centre. The production versions appeared at the end of the 1990s (Joint Research Centre - European Commission, 2008).

General technological development is creating an ever more complex world. The mobile population is increasing and its well-being is significantly dependent on an economy that relies on efficient and safe transport facilities. If no measures are taken, the growing volume of transport will result in an increasing amount of fatal accidents. A safe and clean transport infrastructure in the European Union can be maintained only by diminishing the accident rate. This can only be done by minute investigation and analysis – it is clear that practically all accidents are usually preceded by similar, non-fatal incidents or a slightly different event. A quantum leap forward towards preventing similar events can be made if these precursors are well understood. The contemporary versions of ECCAIRS are based on the ADREP2000 classification that is significantly more accurate and minute. It is worth noticing and mentioning that the menus and values to be filled into the structured fields of ECCAIRS are completed all the time. The authorities of the EU member states can make change proposals that are either accepted or rejected at the annual Steering Committee Meeting.

The installation of ECCAIRS happens by putting it into a database via a platform that can be an Oracle or SQL Server. It is a closed system, which means that the table structure cannot be changed and the data written in the tables are not readable with tools other than those of ECCAIRS. The capsulation of the ECCAIRS kernel is so strict that neither the table structure nor field names and structures of the database can be seen, so there is no possibility to read or write information with SQL, for example. This guarantees that the content of data fields follows the same pattern through all the different users, enabling reliable data interchange. Putting data into the fixed fields must be done through menus that contain predetermined values. These values are updated regularly through a special process which begins with proposals collected from users and ends with their rejection or approval for addition by the steering committee. There is, however, the possibility to save data beyond the menus, but there is no possibility to use it accurately in queries or reports because it is written in fields named 'Other' or 'Unknown'. This kind of procedure causes the database definitions to be delayed, as it takes time when new, often very useful and necessary values are added to the menus, but if different users had the possibility to add values, the data integrity would most likely fail because different users are likely to code the same data in many different ways and forms.

The Finnish Transport Safety Agency has recently published a report about its experiences of using ECCAIRS (TraFi, 2014a). The aviation department of Trafi has analysts who are commercial flight experts, air traffic control and air worthiness specialists. They exploit ECCAIRS in their work as one important source of data – as Trafi receives about 6,000 reported occurrences per year. Although both data input and its analysis is made by competent personnel, guidelines are needed in the input process and they ought to be detailed enough to improve the reliability and consistency of the data. Having good quality data is the key issue as it enables the possibility of a detailed analysis. Usable and relevant information can only be generated if quality source data is available (TraFi, 2014a).

3 THEORETICAL BACKGROUND

3.1 Decision Support Systems (DSS)

Today, managers are affected by increasing complexity and uncertainty in situations where they have to make decisions, which forces them to allocate sophisticated quantitative models which exceed the capabilities of the simple linear models that are traditionally used. The complexity and uncertainty of the data used as the basis for decisions continues to increase. This is why the model that describes a decision-making situation ought to increase as well so that it can capture the highly non-linear relationships among a set of variables (Delen & Sharda, 2008).

Rapid changes dominate the contemporary world, making the future uncertain. Due to the globalisation of issues, as well as the interrelationships between systems, the effect of making incorrect policy decisions has become more serious, potentially having catastrophic consequences. These uncertainties can be seen as existing in all policy making situations and additionally there are several dimensions of uncertainty as well as weak understandings of different characteristics. One of the essential issues among high risk industries, like aviation, is the level of certainty required to restrict or even ban harmful activities. Thus, the precautionary principle has gained increasing attention as the need for more constructive approaches towards uncertainty and ignorance regarding regulatory decisions has grown (Walker et al., 2003).

The decisions are made based on a decision-maker's understanding of a situation as being one source of possible error. Another point to be mentioned is the process by which a decision is reached. Although the best decision has often been made, an undesirable outcome may appear due to events over which the decision-maker does not have control. In naturalistic decision-making (NDM), individuals having domain expertise in contexts like aviation, nuclear power and offshore oil process control make decisions in conditions that change dramatically and dynamically. In addition, the decisions must be made in a limited space of time, the goals may conflict and the reliability of the information sources may vary. Such decisions are often made in teams in organisational contexts, aided by the available tools or other information resources (Orasanu & Lynne, 1998).

The complexity of factors affecting decision making in the information age induces managers to rely on sophisticated information analysis tools that support decision-making in business organisations but also in other types of organisations. DSS are computer-based systems whose function is to support decision-making activities (Kou, Shi, & Wang, 2011) and processes in order to solve semi-structured or even ill-structured problems that often involve multiple attributes, objectives and goals (Gupta, Nemati, & Harvey, 1999; Shim et al., 2002). DSS deal with problems which are based on the knowledge that is available (Ben Ayed, Ltifi, Kolski, & Alimi, 2010). Operationally they are interactive systems or subsystems (Power, 1997) and they are based on knowledge and theory collected from artificial intelligence, database research, mathematical modelling, decision theory, management science, etc. (Kou, et al., 2011). DSS is generally used for any computing application which aims to improve the decision-making ability of a single person or a group of decision-makers. In addition, the term also refers to an academic research field focusing on designing and studying the systems in their context of use (Power, 1997).

Casas et al. (2014) have discussed DSS in traffic management. According to them, the generic architecture of DSS in this context contains the components real-time data, historical data, monitoring, predictive systems and strategy analysis. An interesting detail in this list worth mentioning is that the evolution of predictive systems seems to be heading towards models which are hybrids of analytical and simulation based models. Due to the ever-increasing complexity of transport systems, it seems a hybridisation of the simulation models is required. The new challenges for traffic management DSS appear from the presence of the vast amounts of data provided by the new monitoring and predictive systems (Casas, et al., 2014).

DSS tool designs consist of several components and as databases create the main source of data sophisticated database management capabilities are essential. These capabilities need to include access to both internal and external data as well as information and knowledge. Powerful modelling functions are also needed. In order to produce useful and understandable processing results for practical decisions, a powerful interface is required. The interface ought to be simple to use, enabling interactive queries and the reporting and displaying of the results in graphic form. The research on DSS systems has focused on improving the efficiency of decision-making by developing the technology as well as improving the effectiveness of the decisions made. Powerful tools for building DSS emerged in the early 1990s, On-line Analytical Processing (OLAP, see Chapter 3.3.4) and data mining (see Chapter 3.3) can be mentioned as examples (Shim, et al., 2002). Ben Ayed et al. (2010) point out the role of the data mining tools as decision support systems, showing that due to the rapid development of data processing and the increasing complexity of the problems involved, technological solutions for decision process strategies are increasingly required. From this perspective, knowledge discovery in databases (KDD) is highlighted (see Chapter 3.3.2).

The role of DSS is becoming increasingly critical as a significant component in the daily operations of the organisations. As DSS evolve and mature, the need for a framework that encapsulates organisational factors affecting their successful development and implementation becomes clear. The framework presented in Figure 12 allows DSS researchers as well as practitioners to be able to classify organisational factors (Technical, Economic, People, and Strategic) that may have an impact on the successful implementation of DSS. The framework is the result of a synthesis of existing research and the personal experience of the actors and is based on an in-depth study of the implementation of a large-scale DSS called Fuel Management System (FMS) at Delta Air Lines (Gupta, et al., 1999).

The diversity and complexity of these factors makes it very hard for organisations to understand their impact on DSS implementation success. Brookes (1994) has presented a development framework for DSS based on a model of the decision-making process (Simon, 1960) and the nature of managerial work (Minzberg, 1990). According to this development framework, a DSS can be divided into four functions, called:

- attenuation
- amplification
- reference
- navigation/control.

Each function is directed towards a specific cognitive task by decision-maker. Sanders and Courtney (1985) related DSS success to three broad factors:

- decision context (degree of problem structure)
- level of task interdependence (degree of interaction with others)
- level of task constraints (degree of decision-maker authority and autonomy).

In addition, they found that top management support was an important factor in determining the success of DSS.

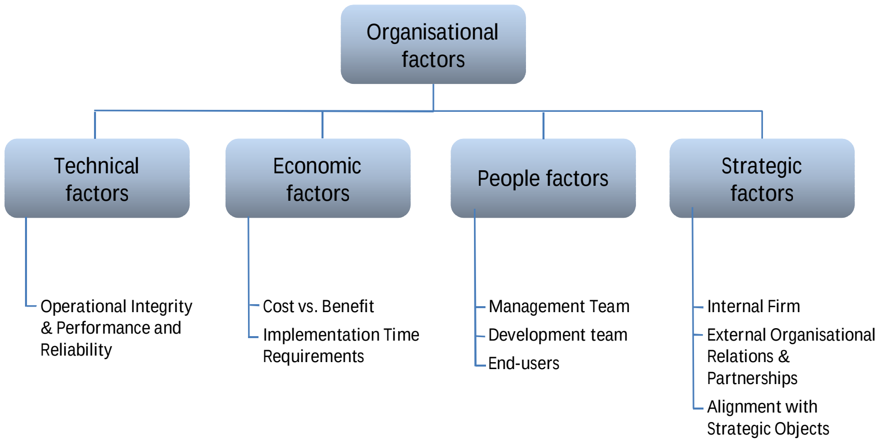


Figure 12: Organisational Factors Affecting the Successful Implementation of DSS: A Framework (adapted from (Gupta, et al., 1999))

Decision Support Systems development and implementation may involve a considerable amount of organisational time and effort (Keen, 1989). Large-scale DSS projects usually require major investments of time and money and this puts tremendous pressure on those involved in the development and implementation of DSS. An unstable and uncertain organisational structure plus confusion regarding the ownership of responsibilities (Turban & Aronson, 1998) can exacerbate this. An organisation wide DSS development and implementation is a massive project that can provide organisations with significant rewards or large losses. A DSS implementation is not always successful and many DSS implementations result in failure (Gill, 1995). Factors that affect the successful development and implementation of a DSS project (Sprague & Watson, 1996) are support from top executives (Lederer & Sethi, 1992); user involvement and participation (Joshi & Lauer, 1998) well-defined objectives or goals for the DSS; resource adequacy issues; organisation and political issues within the company (Markus, 1983); technological issues (Sprague & Watson, 1996; Turban & Aronson, 1998); process management issues (Newman & Sabherwal, 1996; Palvia & Chervany, 1995); communication issues (Horner & Benbasat, 1996); values and ethics (Kallman & Grillo, 1996) and other external issues.

Technical factors include DSS operational integrity, performance and reliability, meeting user requirements and producing accurate results (Turban & Aronson, 1998). Some specific factors to be aware of are response time for rapid task execution, reliability and the availability of the system, data quality for accuracy and timeliness, and ease of use (Gupta, et al., 1999). Standardised hardware, software, development tools and security protocols that will provide continuity and give direction to developers. The training of development staff is important if the team is unable to effectively meet the technical specifications set forth by users (Magal, Carr, & Watson, 1988; Mallach, 1993).

Economic factors are important in determining the overall successful development and implementation of any computer-based system. Two primary *economic* issues are the costs and benefits of the project and the time necessary for the completion of the project. DSS projects often take more time and require more resources than originally expected (Ewusi-Mensah & Przasnyski, 1991). According to several cases in a long research process, it has been noticed that in most DSS development and implementation situations, numerous unexpected problems may occur that necessitate further decisions about the continuation of the process (Lyytinen, 1987; Sabherwal & Robey, 1993). Cost overruns and delays in meeting project completion schedules can be attributed to the poor quality of the original estimates and the design of the system (Ewusi-Mensah & Przasnyski, 1991). However, an MIS development project is likely to be continued when it is considered to be a long-term investment involving a large, long-term payoff structure (Newman & Sabherwal, 1996).

People factors influence the successful development and implementation of DSS as well. Within most organisations, three groups have a vested interest in the successful completion of a project (Ewusi-Mensah & Przasnyski, 1991):

- the corporate management team
- the development team
- the end-users.

The nature of the relationships between these groups can be a definitive factor in the success of the project. A large body of research in DSS development and implementation has shown the importance of user involvement in the success of the project (Valusek, 1994). Tait and Vessey (1988) investigated the role of user involvement and factors affecting user involvement on the success of system development. They found user involvement to be an important factor in determining system success. Alavi and Joachimsthaler (1992) showed that user-related factors, such as user involvement, user training, and user experience are crucial to the success of DSS implementation. The active role of the end-user is an integral part of a project development team's success in implementing a system (Yaverbaum, 1988). Users may be the originators of the project, providing critical information for the development of the system (Ewusi-Mensah & Przasnyski, 1991). The availability of an expert end-user can be critical. Their commitment to the project and willingness to participate can be as important to the system's success as that of management (Joshi, 1989, 1991). End-user training and the availability of support are important determinants of a system's overall success (Yaverbaum, 1988).

Top management commitment is also an important factor in the decision to proceed with the planning (Galliers, Marshall, Pervan, & Klass, 1994; Johnston & Carrico, 1988) and development of a large-scale DSS (Lederer & Sethi, 1988). The absence of such commitment could lead to the abandonment of a DSS project (Ewusi-Mensah & Przasnyski, 1991). Abel-Hamid and Tarek (Abdel-Hamid & Tarek, 1992) found that managerial turnover and succession could also lead to a discernible shift in staff allocations and application performance in terms of both costs and duration. Top management commitment to project and resource support is crucial in the implementation of any DSS project (Turban & Aronson, 1998). However, factors such as the organisational politics and discord associated with a project can hinder or even cause its abandonment. Changes in management and organisational restructuring can change the focus of corporate strategies and reduce the perceived benefits of the project to the organisation (Ewusi-Mensah & Przasnyski, 1991; Mittman & Moore, 1989). Control procedures should be put in place by management to ensure those standards and policies exist and are adhered to (Magal, et al., 1988). Top executives must also have a complete understanding of what DSS can do for the organisation and how it is going to help them achieve company objectives and be convinced that the project is worth pursuing.

A development team (staff) holds a unique position in any implementation process. They are directly responsible for the technological and organisational development of the project (Ewusi-Mensah & Przasnyski, 1991). It is crucial that they have a comprehensive understanding of the end-users' aims, [business] and decision processes. This requires large amounts of time spent with expert end-users. The training of development staff is also important if a team is unable to effectively meet the technical specifications set forth by users (Magal, et al., 1988). Furthermore, the loss of key members of the development staff can contribute to delays or project abandonment (Ewusi-Mensah & Przasnyski, 1991).

Strategic factors are critical in determining the successful implementation of information systems (Sanders & Courtney, 1985; Turban & Aronson, 1998). In this context, three issues with strategic implications are discussed here:

- the degree to which the DSS project is aligned with the strategic objectives of an organisation
- internal firm issues
- external organisational relationships and partnership issues.

Regarding the first issue, DSS can be used to gain strategic and competitive advantage in areas such as organisational planning, decision-making and reducing the costs of [business] operations and transaction processing. Hence, it is critical that the development and implementation of DSS be successful (Alavi &

Joachimsthaler, 1992). The successful development and implementation of a DSS project depends, to a large extent, on organisational commitment to it (Kwon & Zmud, 1987; Markus, 1983). Without a high level of commitment to a project, the necessary resources may not be dedicated and the project may be abandoned (Lederer & Sethi, 1988). To achieve organisational commitment, a project must address an important and urgent organisational problem that needs to be solved (Meredith, 1981). A DSS project must also show potential for high payoffs. Therefore, a DSS project is more likely to succeed if it reflects the belief that it makes a valuable contribution to meeting organisational strategic objectives (Weill, 1992).

Internal firm factors are also considered to be important in DSS development and implementation (Turban & Aronson, 1998). Hitt and Brynjolfsson (1996) showed the relationship between information technology (IT) and the organisational architecture of firms. Their study of almost three hundred companies clearly showed that internal firm practices, such as the level of centralisation of a decision-making authority, can have an impact on the success of information technology projects. When Lederer and Sethi (1992) studied the impact of internal firm factors on the implementation of computer based applications, including DSS, they found that these factors could influence the successful implementation of DSS systems.

External organisational relationships and partnerships are also important factors in determining the success of DSS projects (Lee & Kim, 1999). The 1990s were a decade characterised by restructuring, mergers, and outsourcing in the business world. More and more companies were eager to form alliances or partnerships, hoping to gain a strategic advantage by consolidating their market shares and becoming technologically agile. At the end of that decade, increasing attention was paid to building successful partnerships for outsourcing information systems (IS) (Grover, Cheon, & Teng, 1996; Lee & Kim, 1999). Lee and Kim (1999) established partnership quality as a key predictor of outsourcing success and examined its impact on outsourcing success. Their results indicate that partnership quality is positively influenced by factors such as participation, communication, information sharing, and the support of top management but negatively affected by the age of the relationship and mutual dependency.

3.2 Research for aviation safety

3.2.1 Background

Since the beginning of the history of aviation there has been research into flying. The first research area was how to make flying possible. When that was solved, the rapid development of aircraft began. Physical and psychological problems were soon faced and aero-medical research began. The results of these two major research areas primarily served military aviation during the first half of the 20th century, but after World War II that knowledge was taken into use for civil aviation, which is when modern aviation began.

Aviation has suffered accidents from the time the first aeroplane flew. Indeed, it is true to say that, for many years every attempt at flight resulted in an accident. The period after 1903, when the Wright brothers made their first successful powered flight, is probably of most interest to us today. Firstly, an analysis of the causes of accidents in the 1920s and 1930s suggests that the ratio of accidents caused by engineering factors to all accidents is remarkably similar to that which pertains today. In other words, the improvement in safety has been more or less even across the field. Secondly, although accidents were studied, they were not reported in the open way they are reported today. Thirdly, the responsibility for aviation safety regulation did not reside in a single body, like a civil aviation authority, as it does in many countries today (Chaplin, 2004).

The incredibly rapid development of huge volume air travel, using mainly jet airliners, created the need for systematic research into aviation safety. Chapter 7 of Annex 13 to the Convention on International Civil Aviation recommends that “States should establish formal incident reporting systems to facilitate collection of information on actual or potential safety deficiencies” (7.3). In view of this recommendation, the Commission has investigated the status of the implementation of ICAO SARPs (Standards and Recommended Practises) on mandatory occurrence reporting systems. In a study carried out on behalf of the Commission, IFALPA (International Federation of Air Line Pilots Association) found that only a few member states collect mandatory occurrence reports and even fewer stored, retrieved or analysed the related data (European Commission, 2000).

After every accident or incident something can be learnt and, fortunately, in the aviation sector this is the general culture worldwide. One of the earliest reports was into the accident of the R101 (Masefield, 1982) airship which crashed at Beauvais on its way to India. The analysis of the event was good but the report, in contrast to a similar report today, contains no recommendations for avoiding future problems. The beginning of the reporting and analysis culture we have today, at least in the UK can be seen in a case from January 1948, when a Tudor aircraft, Star Tiger, disappeared on a flight to South America. No wreckage was found to help the inquiry, which therefore had to focus upon what could be gleaned from the history of that aircraft and of similar aircraft. The report of the investigation contains the following observation: *"An important contribution to safety in the air is provided if arrangements are made for a careful examination of any incidents or minor accidents which may occur, in order to eliminate their causes, and to warn the crew of aircraft of the circumstances in which they arise. On the evidence produced to the Inquiry, the Court has formed the opinion that no such organisation existed in British South American Airways Corporation, or, if one did exist, it was ineffective."* (Macmillan, 1948, p. 45). After this observation, most British airlines did establish some system for looking into incidents. Similarly, the regulatory authorities set up systems to work with airlines (Chaplin, 2004). This example is from the UK, but similar cases can be found from many other countries.

As for learning from incidents and other safety related events, the case is not the same in other industrial sectors, where appropriate action is seldom taken to ensure that full details of an event are recorded. Also the actions required to prevent it happening again are not often clearly determined, recorded and shared with others. It is obvious that a new engineer or scientist has to learn the hard way through personal experience. Including such information in the codes and standards developed by companies was once possible but – due to an increasing reliance on European or international standards, such information is often excluded. Furthermore, the people who learn their lessons from within an organisation also retire and leave, causing the corporate technical memory to become shorter. In addition to this, many companies fear legal issues and do not share information with other actors in the same industry or in industries with similar characteristics.

Four different approaches and phases, or ages, of safety analysis have been identified by several researchers, e.g. Hale & Hovden (1998), Sheridan (2008) and Reiman & Oedewald (2009), although distinctions are not totally clear and straightforward (Wiener, 1977, 1980) because some ‘modern-hailed’ views are not so new and have been present in aviation for some time. In the very beginning of aviation (the first age that had a scientific study of safety was the period between the start of the 20th century and the Second World War) safety focused on technical matters, approaching the safety from the traditional error-risk analysis point of view. As Heinrich et al. (1980) describe it: in that period people were, as a rule, considered the weakest link in the chain, which is why personnel training and selection were developed and applied as preventive measures. During the period between the World War II and the 1970s, the second age of safety, the focus was more on human error and correcting them (Rasmussen, 1982).

Before the 1990s, the technical risk assessment and correspondent preventive measures reached their limits. This period can be regarded as the third age of safety. Due to this development, this age can be seen as a turning point because a new way of thinking was adopted, meaning that safety management systems and organisational factors were regarded as research objects. Until then, there had been an obvious lack of new accident causation models capable of recognising the fact that workplace factors in particular and organisational factors in general are the root cause of most unsafe acts. Therefore, the search for such factors has been sparked by the trend to integrate safety and health into mainstream management. One significant point in this chain of development is the acceptance of Reason's (1990) model (see Chapter 3.2.3) for human error and organisational accidents. Reason proved that individual errors very seldom cause accidents on their own, but showed that in the background of an accident (or incident) there are often errors and shortcomings on different levels of the organisation in question (Saatsi, et al., 2011). Despite the wide applicability of the model, it has been criticised for its intangibility and its way of simplifying complex reality by making it too linear (Hollnagel, 2004).

The fourth age, more precisely defined as the 2000s, brought the pro-active approach to safety management, emphasising human activity in co-operation with unpredictable, underspecified and complex working environments, especially when facing unexpected events; Leveson (2011) emphasises the need to learn from events and system-wide aspects. Hollnagel (2003) and Reason (2008) have mentioned co-operation between companies, service providers and regulators, a process that ought to continue seamlessly in order to enhance joint cognition systems. According to Carayon (2006), these activities should cover the whole organisational, regional, cultural and temporal boundaries of the systems. The particular emphasis of the safety approach in the 2000s is organisational and safety thinking, consisting of organisational norms, perceptions and shared values (Hopkins, 2006; Weick & Sutcliffe, 2007).

There are industrial actors who believe this stance to be counterproductive. In the UK, the Health and Safety Executive is increasingly requiring companies to establish management systems that ensure that lessons learned from their own accidents as well as those of other companies will be recorded for future reference. Engineers responsible for design, risk assessments or writing up safety cases require access to good quality data on accidents, alongside an effective accident database, so that they may apply best practices in their work. Hindsight is wisdom after the event happened, but through a process of learning lessons from accidents hindsight can be converted into foresight. Thus, the required precautions are converted into common engineering practices (Royal Academy of Engineering, 2004).

When an incident occurs, its analysis creates lots of information. Incidents that can be determined but include no hazard usually remain unremarkable for those involved, but they do require the momentary mobilisation of all the resources that are available for corrective action. This context always has processes that go through the phases of observing what has happened, reporting on those phases and making corrective actions in order to prevent the incident happening again. These are often complex processes consisting of rounds of activities and they involve persons who were located in the space in which the case has occurred. The persons involved also establish their own relationship with the situation, or others enlist them as contributors to its resolution and as practical actors (Suchman, 1996). It is worth noticing that aircraft operations are never performed without the presence of hazards to some extent, thus a risk might only appear after a chain of occurrences. That is why special attention ought to be given to the recorded occurrences in the daily operations of aircraft. Although the consequences are rarely fatal, all aircraft accidents and incidents should be considered and analysed carefully (Čokorilo, et al., 2014)

One of the critically important and therefore most essential topics for engineers is learning from incidents and accidents. About thirty years ago the classification of incidents and accidents identified them as being just technical malfunctions or even ‘Acts of God’. Now, particularly because of work undertaken by social scientists, such as Barry Turner and James Reason (Reason, 1990), we know that the causes are more complex. Reason's model makes it clear that all high hazard systems have in-built defences. An accident will only occur if several of these defences are breached in succession. Near-misses in particular provide valuable learning opportunities, as they show how some but not all of defences can be breached in the future. However, as the model makes clear, accident analysis also has to take into account the surrounding 'latent' management and organisational factors which we now know contribute to many of the causes of major accidents. It is also important to recognise that learning can occur across industrial sectors (aviation, chemical, energy) precisely because of the latent preconditions, which are rooted in organisational and human behaviour and are common to many sectors. Professor Barry Turner was asked many years ago why an engineer specialising in chemical risk should be at all interested in (say) aviation accidents; his reply was, characteristically, because the organisational preconditions for the next accident the chemical engineer might experience could have already occurred in another industry. Learning also depends upon how we model the system. At a very general level, system failures can – and do – happen regularly (Pidgeon, 2004).

Heinrich’s Domino Theory has obviously been the most influential accident causation theory. He studied 75,000 industrial accidents and came to the conclusion that unsafe acts committed by fellow workers cause 88 per cent of industrial accidents. He also noticed that ten per cent of industrial accidents are the consequences of unsafe conditions and only two per cent of them are unavoidable (Heinrich, 1931). According to this theory, the result of a five-stage sequence is an accident that leads to an injury or damage. In this model, a linked cause is represented by each domino, and if any of them is removed, the damage or injury will not occur because the sequence cannot run its course. The types of causes, represented by each domino stage, have been built up by consensus over time. Nowadays, enough is known about the two last stages, i.e. the accident leading to injury or damage but the first three causes are not as well understood (The Universal Manager, 2005):

Work situation

- Inadequate management control
- Unsafe system design
- Lack of suitable standards
- Faulty or inadequate equipment
- Business pressures

Fault of person

- Insufficient skill or knowledge
- Failure to follow procedures
- Personal problems
- Lack of motivation
- Inattention
- Forgetfulness

Unsafe act

- Process error
- Taking short cuts
- Taking unnecessary risks
- Removal of safety equipment
- Failure to use safety equipment.

There has also been continual development in the application of the model:

- Prevailing initial interpretations saw the worker as the prime source of accidents, and concentrated on removing the third domino through a combination of disciplinary and protective measures.
- As MacGregor's (McGregor, 2006) Theory X and Theory Y of management gained influence, the emphasis switched to the second domino and the preventive action centred more on training and work design.
- More recently, safety theory and practice has focused on the first domino, as belief has grown in the ultimate responsibility of senior management for safety matters, and in the importance of designing systems and processes which are safe from the beginning.

As a result of the accident prevention layers developed for aviation, only in extremely rare cases does a single error cause an accident or a serious incident. Instead, they often produce hazards. Hazards are classified as observed incidents during flight or maintenance that could have led to an accident. Furthermore, operations can contain several errors and various deviations affecting safety, but which are immediately detected by co-workers or a group. These events are often the first weak signal about flaws in the procedures or operations of the organisation creating a safety risk (Saatsi, et al., 2011).

Statistics have proven that a special ratio, called Heinrich's law, exists between the number of flight accidents, incidents of varying severity and the observed deviations that affect flight safety. In his book Heinrich (1931) states that for every major injury causing accident, 29 accidents that cause minor injury exist and, correspondingly, there are 300 accidents causing no injuries. This is illustrated in Figure 13. The accident pyramid is derived from a study made by Heinrich (1931). The pyramid clearly illustrates his theory about accident causation: unsafe acts lead to minor injuries and over time to major injury. The pyramid displays the ratio between accidents and other categories. On average, before one accident takes place, there will have been (approximately) 15 serious incidents, 300 other incidents and 15,000 deviations and other cases concerning flight safety. The most numerous ones might have no impact at all on flight safety, but still they are deviations from normal operation. A case, in which a light bulb in a cabin broke and produced some thin smoke, is a good example. In such a case, there was no danger at all, but a deviation to be reported. These are of very little value alone, but collected together they create a huge amount of information that can be used when analysing what happens in the skies. The closer to an accident the cases occur, the more detailed the investigation and the greater the skill that is required in the analysis. All in all, a bottom-up investigation analyses cases to find out what might happen and make an estimate based on what has happened. The top-down direction provides rules and instructions to stop a dangerous trend from progressing to an accident.

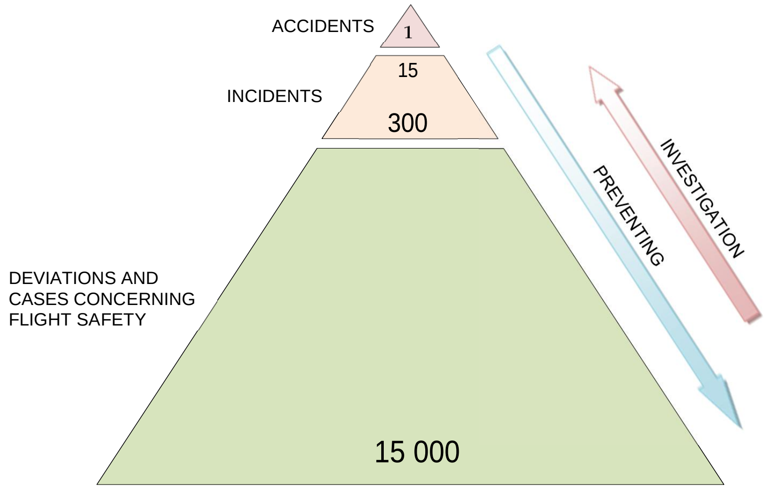


Figure 13: Work to improve flight safety (adapted from (Heinrich, 1931))

Latino (2008) refers to a study of industrial accidents, conducted by Frank E. Bird, Jr., then the Director of Engineering Services for the Insurance Company of North America. He analysed 1,753,498 accidents reported by 297 companies in 1969. The outcome of this process revealed a new ratio of 600:30:10:1, which means that one major injury is preceded by 10 minor injuries and 30 property damage incidents. Before all of these, there will have been 600 near misses, i.e. safety events. Saatsi et al. (2011) present the model slightly adjusted, doubling the amount of deviations affecting safety, as well as stating that 30 minor incidents and 10 major incidents precede one accident. According to this presentation, only six per cent of hazard events in aviation are published; 94 per cent contain safety affecting deviations that are recognised only by actors working in close relation to the action and are not considered to have any significant safety effect. However, these events can be as relevant as events that have led to a public investigation from the point of view of the flight safety.

Heinrich's Domino Theory might be more attractive from the point of view of health and safety practitioners than Reason's model because it offers more choice about where to begin tackling safety in an organisation and because the latter offers a starting point for organisational strategy and culture that might not be easy to influence for health and safety practitioners. Another point worth mentioning here is the causation of accidents, especially in certain high-risk systems, which according to the belief among some experts, might be too complex to address from the standpoints of organisational factors. As Perrow (1999, p. 304) says, "[they] should be abandoned because the inevitable risks outweigh any reasonable benefits". Nuclear power and weapons are the most often cited, but DNA technology and space travel are also often mentioned.

Uncertainty in a process is a permanent state despite the quality of the design; uncertainty arises from design flaws as well as the unpredictability of the future operational demands of a complex system. Different kinds of deviations, disturbances to planned operations and everyday problems make unpredictability visible. The operators face these problems directly and identify features that need correcting in a design as a result of dealing with its problems. As the degree of complexity in a system increases, such problems become more difficult to anticipate. Simultaneously, identifying them becomes more important for both safety and economic reasons. For operators, defining and identifying the world of the barely possible is also included in dealing with difficulties, and thus their work is not only restricted to identifying predictable problems (Norros, 1996).

When considering the concept of aviation safety, different connotations can be found. According to Waikar and Nichols (1997), this means it should be free from hazards or a state with no accidents, while Watson et al. (2011) define this similarly as the avoidance of attacks. The safety of aviation systems has been measured using the accident rate in terms of fatalities and hull losses. If safety could be improved through the design of a machine interface and training, a metric for flight safety – usable in the design process – ought to be defined. Due to the systematic work carried out since the 1950s for aviation safety, accidents are extremely rare occurrences today and this is why they cannot be considered a meaningful metric when the safety of a particular flight, flight deck, or operator mission is evaluated. Furthermore, although incidents, including regulatory violations, are more frequent than accidents, they are still relatively rare (Wickens, 1995).

System safety is traditionally based on the safety implications and consequences of a system's technical aspects and particular components. From the point of view of the safety management of an aviation organisation, in operational issues and system design, the key drivers are the identification of a hazard and safety risk management. In this process, the role of the human factor is significant (Čokorilo, et al., 2014). In the aviation industry, experiments concerning human performance typically define improvements in the safety of particular procedures, systems and training regimes through the measurement of a narrow band of performance that is directly related to the intervention of the factor. In these experiments system-specific errors and a few specific reaction times have typically been used as dependent measures. However, using this approach is somehow problematic. As, to begin with, the trade-offs between accuracy and reaction time performance are managed by the subjects (Pachella, 1974). Furthermore, the operational significance of measures that are finer grained can be questionable. For instance, if faster reaction times are observed, conclusions about translating them into behaviour consistent with improved safety cannot directly be drawn (Rogers, Schutte, & Latorella, 1996). In general, it can be stated that a result will mostly represent what is measured in such focused experiments. This means that resources will be managed to optimise performance on aspects that are believed to have been measured. These kinds of focused experiments are commonly assessed subjectively. The subjective evaluations might, however, be dissociated from actual performance (Yeh & Wickens, 1988), being perhaps unfairly biased towards familiar designs. From the point of view of designing evaluation metrics, the current measures do not provide an adequate index for safe flight performance because they are strongly associated with opinions about the ultimate safety of flight performance and "good" pilot performance (Latorella, 1999). One of the essential concerns is whether or not a system is composed of both the pilots and whether the cockpit technology environment functions well, i.e. it is not just concerned with the good performance of a particular pilot. Hence, safe and appropriate flying depends on a system, not on the skills of any individual pilot. A modern jet airliner cannot be flown by an individual acting alone, at least not in current practice; safety does not depend on any individual skills but on the properties of both the crew and the aircraft (Hutchins & Klausen, 1996).

Organisations should learn efficiently from incidents and other deviations. It is widely observed that hazards originate from organisational failures. Special attention ought to be paid to the so-called organisational preconditions to technological accidents and, in the worst case, disasters. During recent decades, large-scale accidents have occurred. The Chernobyl disaster and the destruction of the space shuttle Challenger (see Chapter 2.1) in 1986 are prime examples of technology and organisational failings turning out to be a key factor. These findings can be called organisational accidents stemming from an incubation of latent errors and events, accompanied by a collective failure of organisational intelligence.

Preventing the injurious consequences of errors is a corrective action to be taken after a single serious incident or the analysis of a report. To determine safety measures – in cases that have directly threatened flight safety (technical cases in this example), procedures that follow the FAA (Federal Aviation Administration) System Safety programme can be applied. These give precedence to corrective actions on the basis of their efficiency (Saatsi, et al., 2011):

- the activities or the technical structure are changed in a way that eliminates the possibility to make an incorrect installation
- the activities are complemented with a double check or technical change which ensures safety
- a warning procedure or a system for identifying incorrect installations is taken into use

- the procedure is formed as an unambiguous standard action that is complemented by sufficient training.

Theoretical models have also been developed in an attempt to specify ‘high-reliability’ organisations and ‘safe’ cultures, based on the former, purely post-hoc descriptions of accidents and their causes (Pidgeon & O'Leary, 2000). The first contemporary, and path-breaking, model to be mentioned here is the “Man-made Disasters” model by Turner (1978). This model laid the conceptual foundation for the work made by others (e.g. Reason) and has significantly contributed to our present theoretical understanding of industrial catastrophes and crises as administrative and managerial in origin. The theory starts from the essential observation that disasters in large-scale technological systems can be considered neither chance events nor ‘Acts of God’. Turner argues, instead, that the origin of disasters can be found in the interaction between the human and organisational arrangements of socio-technical systems that are built to manage complex and ill-structured risk problems (Pidgeon & O'Leary, 2000). It is already a well-known fact that disasters, as well as accidents and serious incidents, do not just happen; they are a chain of critical events.

3.2.2 Hazard and risk identification

Hazard identification and safety risk management are the core processes in the management of safety. In practice, this means that the factors which cause or are likely to cause harm ought to be identified and understood. The issue of aviation safety considerably interests the scientific field. The attack made with two hijacked airliners on the twin towers in New York in 2011 has significantly increased interest in safety matters (Čokorilo, et al., 2014). According to a study conducted by Boeing (2012), the highest number of deaths-on-board generating situations are the loss of the control of the aircraft on the runway, both landing and taking off, as well as misunderstandings between the pilot and the control tower, sometimes caused by an imperfect knowledge of English by one of the parties.

The incident reports can be classified in many ways because variations in the size of an incident reporting subject can be considerable. For our purposes, Brooker (2006) proposes a useful division into three schemes:

- Individual reporting, that is, an operational person detects, from a safety point of view, unsatisfactory event reporting by a central monitoring body.
- Event-related reporting originates from automatic system warnings or alerts, such as ACAS (Airborne Collision Avoidance System), which

warns a pilot of the proximity of another aircraft possibly presenting the threat of collision

- Post-processed reporting is based on radar and related data that is examined sometime after actual operations in order to find out whether, for example, separation minima have been substantially breached.

A good picture of the characteristics of any incident or corresponding event can usually be achieved by combining the entire data from the three different reporting systems. This data can also be enhanced by using the recollections of the pilot and/or the controller and other data like communication recordings (Brooker, 2006). In order to function properly, the reporting system ought to be easily and simply utilised by its user. When modern computing is applied, using such a system no longer need be complicated, especially if it is based on exact coding or taxonomy and includes the possibility for free-form reporting. Among confidential reporting, the first analysis of the reasons that led to an event ought to be done by the person who made the mistake or first recognised it, because they best know what happened in that situation. After relevant training for making analyses, the people in an organisation should have both full competence and a special interest in making analyses (Saatsi, et al., 2011).

An essential question in this context is the detection of weak signals. They are minor indicators, often about ambiguous threats, easily underrated and/or ignored because they are signals of something that might – but not necessarily – portend future harm. In the worst cases, they might lead to a catastrophe. The disasters of the space shuttles Challenger and Columbia (see Chapter 2.1) were more or less caused by ignoring the weak signals, partly because foam strikes, as in the case of Columbia, had occurred before without causing any significant damage. Behind the scenes there are also managerial tensions about paying attention to the signals discovered. According to Roberto et al. (2006), alertness prevents ambiguous threats being converted into lethal failures. Furthermore, organisations that are capable of identifying threats and responding to them are able to avoid serious consequences (Roberto, et al., 2006). Those and other managerial challenges as well as the applicability of data mining in responding to them are studied in Publication #5 “Data Mining Challenges in the Promotion of Aviation Safety”.

Improving safety is a complicated task, because the probabilities that guide the predictions of how the next accident will occur are very small. Due to this, deciding where an intervention should be made in order to prevent accidents is rather hard. Therefore, the focus ought to be aimed at potential accidents. One way to do this is to use simulations, which are used widely by airplane manufacturers. Another possibility is to search for mechanical anomalies (rare failures that occur randomly and which do not fit predefined event classifications), these are usually small but common anomalies and could lead to problems later on. The main target is obtaining general, applicable information by observing events that occur relatively frequently and which are, in themselves, innocuous by reason of the robustness of the systems. If they belong to links in an accident chain, those links can be fixed before they cause an accident. However, to achieve this, overcoming two barriers is required. First, the potentially useful data must be spread throughout the world of thousands of different entities representing all kinds of actors in the aviation field. These include private and national airlines, maintenance companies, manufacturers, aviation authorities, air traffic controllers, air forces, as well as labour unions and trade associations. Secondly, the collected and stored data is, in most cases, in an unstructured form, which makes it difficult to analyse and compare.

To address these shortcomings, the FAA (Federal Aviation Administration) facilitated the creation of the Global Aviation Information Network (GAIN) in 1996. Nowadays it does not act in the same way because the FAA ended its support for GAIN in 2007. Since then, the Flight Safety Foundation has supported the distribution of a wide range of GAIN products, such as safety manuals. GAIN was a voluntary international membership consisting of more than 50 countries – both private and public entities. Its core philosophy was “the collection, analysis, and sharing of safety information using advanced technologies in a just culture environment will illuminate safety concerns and permit identification and implementation of cost-effective mitigations” (Robb, 2005). One of its most essential goals was to get its members to openly share information, which has not been completely achieved yet. Despite that, significant progress has been made in analysing safety information by developing specifically designed analytical tools (Robb, 2005). The author participated in the actions of Working Group B between the years 2005 and 2007, first joining a meeting in Houston in January 2005 and then acting as ‘a remote resource’, i.e. assisting the group on two flight safety surveys by contacting the Finnish operators. At that time the acting working groups were:

Working Group B, Analytical Methods and Tools
Working Group C, Global Information Sharing Systems
Working Group E, Flight Ops/ATC Ops Safety Information Sharing.

GAIN held five world conferences for improving aviation safety by promoting the GAIN concept and sharing its products with the aviation community up to 2003. By 2003, almost 900 aviation professionals from 49 countries had participated in GAIN sponsored conferences (GAIN Working Group B, 2003). GAIN has published several manuals on flight safety, for example: Guide to Methods and Tools for Airline Flight Safety Analysis. This includes the before mentioned MITRE Workbench (see Chapter 3.4.1) as an example of an analytical tool. The author of this thesis joined the e-mail list of GAIN Working Group B, specialised in Analytical Methods and Tools, which has sponsored the creation of several tools for analysing text information. By 2003, only a single airline had used all the tools (Robb, 2005). However, commercial airlines as well as national authorities have carried out research with universities to analyse aviation safety data, but such studies are rarely made public, especially those made for airlines.

3.2.3 Reason's Swiss cheese model

The Swiss Cheese model developed by James Reason created totally new concepts for determining the reasons behind accidents. So-called active failures (aka contributing causal factors) were found to act as triggers for the accidents. These are made possible by latent failures that are already in a system and which have not been actively eliminated. The active failures are deficient functions or bad decisions that directly affect the event. The errors made by a flight crew, an air traffic controller or mechanic performing a service operation are usually classified as active failures. Latent failures can have an influence up to three levels higher before an active error level occurs in an organisation. Although latent failures can be totally ignored for long times, weeks, months or even years, they may enable the active failure at any time, leading to an accident in a worst case scenario. The defensive layers are checks in the safety-critical operations preventing or uncovering a possible active failure. They can be technically implemented backups or alarm systems and construction solutions that prevent incorrect installation. Regarding technical maintenance, they are typically inspections made by other people at the end of a service phase, the performance tests on a system and the test flights (Saatsi, et al., 2011).

In Reason's model, the cause of accidents is the result of latent and active failures penetrating barriers that normally prevent them from culminating in an accident. The model clearly illustrates the fact (learned through several accident reports) that in almost all cases, accidents emerge from a number of successive smaller hazards and errors, subtle precursors and latent weaknesses, but not from one isolated event. In order to clarify the model, a semi-permeable membrane could be imagined to exist as an imperfect final defence between unsafe-acts and accidents (Reason, 1990). Hence, the probability of a resulting accident can be predicted based on the likelihood of a hazard penetrating through a membrane, and we should also be able to predict the concentration of unsafe acts that lead to an accident. A similar observation has been made by Hollnagel (1991), who argues that the amount of erroneous actions will increase as the complexity of work interfaces increases.

Pidgeon et al. (2000) emphasise the role of the developmental processes in the origin of accidents as well as in less fatal cases. Through empirical case studies many preconditions to major systems failures were revealed, some of them originated years prior to an actual event. This indicates the importance of revealing lethal trends before they can proceed to a critical level, causing serious incidents or even accidents.

The model presented in Figure 14 consists of five layers, the first of them represents defences, the purpose of which is to mitigate the results of an unsafe act. Conditions like fatigue, stress, operating practices, etc. are included in the second layer, unsafe acts. The third layer contains the preconditions. The operative aspects, like training, maintenance, etc. belong to the fourth layer of the line management. All the high level decision-makers, such as the regulators, owners, manufacturers, designers, and trade unions are situated on the fifth layer. According to Reason (1990), the decision-makers continuously make "fallible" decisions. Meanwhile, the resulting latent defects stay dormant until someone commits an unsafe act, which might trigger an accident scenario. The effects of it stay limited until the defences of the system work as they should and catch the results of the unsafe act. Otherwise, the accident might prove tragic, which emphasises the significance of eliminating, or at least, reducing safety deficiencies. As demonstrated by the model, if we reduce the number or size of the holes in the layers, the probability of an accident occurring can be reduced.

If there are a large number of causes that can trigger an accident, Reason's model is particularly useful. Regarding common organisational factors, the model includes health and safety goals, workplace design, resource allocation, the procurement of equipment, and the delegation of responsibility and authority, etc. All these ought to be observed as workplace factors that are present throughout the whole organisation, including work schedules, training, instructions, policies and procedures, etc. The frontline worker is often influenced by these factors to cut corners, particularly in potential accident scenarios, for example, when an urgent deadline should be met (The Universal Manager, 2005).

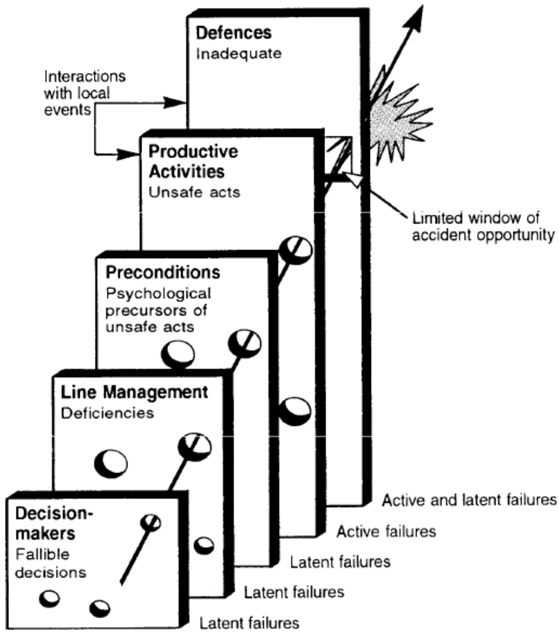


Figure 14: Accident Causation Model developed by Reason (adopted from (Lybeck, 2002))

In Reason's model failures almost always consist of several components. Human error is only one part of the chain. The following real life example (Lybeck, 2002) illustrates how the model can be used. On 5 June 1995, at Helsinki-Vantaa airport, a mechanic taxied an aeroplane, type ATR-72, into the door of a hangar. The damage cost several million Finnish marks (about 1 M €). The aircraft had had maintenance, after which a mechanic intended to test the engines and then to move the aircraft to another place. The mechanic forgot to switch on the hydraulic pumps before taxiing, which resulted in the brakes losing hydraulic power, making it impossible to stop the plane before hitting the door. This illustrates how essentially minor errors can have serious consequences through a chain of triggered events.

The example deserves a more detailed explanation. As mentioned before, Reason's model has five layers, the three first concern latent failures, the fourth active failures and the fifth both active and latent failures. The first layer covers decision makers and their fallible decisions. In the case of the ATR-72, at least two affecting factors can be found on this layer. The regulations given by the authorities were insufficient and the checking of quality systems, at least for engine testing, was deficient. On the second layer, line management deficiencies played an important role. Both engine testing methods and engine testing training were insufficient. The impact of the state of alertness of workers was not taken into account, especially during the night shift. The third layer contains the preconditions, which are the psychological precursors of unsafe acts. The engine testing was mentally very stressful for the mechanic. His state of alertness was low. This led to the fourth layer, productive activities with unsafe acts. The hydraulic pumps were not switched on when the engines were switched on. The fifth layer handles defences and whether they are adequate. It contains both latent and active failures. The defences were inadequate: no checklists were used. This can be counted as a latent failure. An active failure was that the mechanic did not recognise the HYD warning light, indicating that the hydraulic pressure was too low to have functional fully brakes. Thus, the window of opportunity for an accident was opened and it occurred (Lybeck, 2002).

The primary solution to the problem of increasing safety is to learn from incidents by collecting data from all levels. The collection process exists only to search for key factors to improve, not to find those who could be considered guilty of accidents or incidents. This process is illustrated in Figure 15.

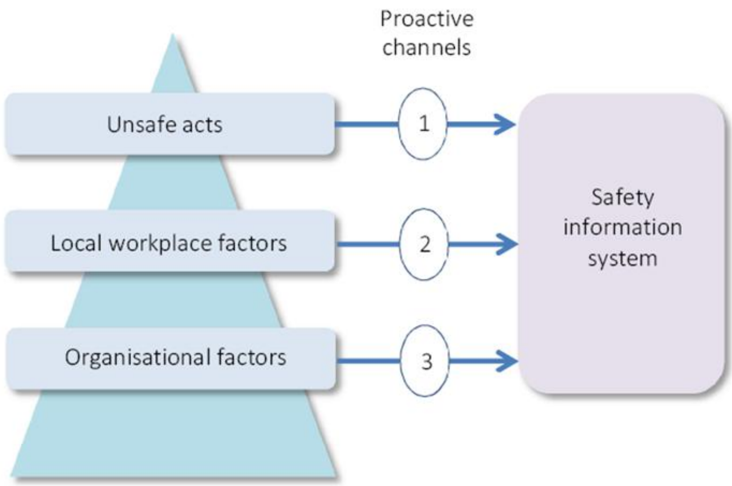


Figure 15: Levels of the information collection process (adapted from (Reason, 1990))

3.3 Data mining

3.3.1 The concept and its definition

The exact definition of data mining is intractable because it is neither a single tool nor a method, but a combination of both methods and tools. Expressed generally, data mining is needed because traditional techniques and methods may not be suitable as the data is highly dimensional, heterogeneous and its nature is distributed. Due to the characteristics mentioned and the vastness of data, much of it may never be analysed at all (Gheware, Kejkar, & Tondare, 2014). Berry and Linoff (2000) have described data mining as an analysis and exploration process, utilising automatic and semiautomatic means in order to discover rules or meaningful patterns. Data mining is an essential part of the Knowledge Discovery Process.

It is significantly easier to define the purpose of the Knowledge Discovery Process as "*a decision support process in which we search for patterns of information in data*". A user might just do a search or they may ask for assistance from smart computational methods that, through the automatic search of one or several databases, aim to find significant patterns in the data (Parsaye, 1997). These nuggets of knowledge are hidden in vast amounts of data and according to Watson (1999) they are practically undiscoverable with conventional techniques. If a user performs the process and makes database queries, the task might be quite hard and require some knowledge of what should be searched for and found (Parsaye, 1997). The approach in this context characterises data mining paradigms from the point of view of discovery and prediction and has the essential starting point that the relationships and patterns exist in these large databases but are *hidden* in vast amounts of data. The aim of data mining is the search for these patterns and relationships. Watson (1999) underlines that it might take years to find these meaningful relationships with conventional techniques, i.e., database queries, spreadsheets etc.

Data mining has its roots in various scientific disciplines, from exploratory data analysis in statistics to machine learning in Artificial Intelligence. It has recently started to play a dominant role in data processing. Statisticians had used the term data mining in the 1970s with the negative connotation of "sloppy exploratory analysis" with no prior hypothesis to verify. Lately, however, it has evolved into the most promising solution for business intelligence and has technological, biological, medical, military, and other purposes. Aviation with its many functions, i.e. flight operations, maintenance, personnel administration, training, etc., is included. Data mining applications are sets of problems with similar characteristics across different domains, therefore the same algorithms and models that are used for developing the fraud detection capability for credit card transactions can be used to develop health insurance fraud detection applications. One of the most promising new topics in IT in 1996, according to the online journal of Database Management System, was that "Data mining has now moved beyond the early adopter stage". Data mining solutions contribute to the human understanding of huge masses of data in various representations, including tabular data domains, spatial data domains, text based domains and image-based domains. Data mining in business and business analytics can be used interchangeably to give business users strategic insights from massive databases (Kohavi, Rothleder, & Simoudis, 2002). Also, data mining can be used for the secondary analysis of data that have been collected for some other purpose (Laurikkala, 2001).

A significant distinction between data mining tools and other analytical tools is how they are utilised for exploring data interrelationships (Moxon, 1996). While analytical tools support the verification-based approaches of intuitively posing queries to the database, data mining tools use discovery-based approaches to determine the key relationships in the data by employing different algorithms of pattern matching. For instance, the Knuth-Morris-Pratt matching algorithm finds copies of a given pattern – a short string of symbols – to be a contiguous subsequence of a larger text (Knuth, Morris, & Pratt, 1977). The approaches used in data mining depend on the goals of the person who is analysing the data and on the types of patterns to be extracted. Data mining has borrowed a number of algorithms and techniques from its parental fields to accomplish the main tasks of data analysis, classification, clustering, categorisation, and prediction. The algorithms differ from each other in the type of data mined, starting from the most trivial numerical data and moving towards transaction data, data streams, graphic and scientific data. The categorisation of data mining tasks is produced below. This categorisation is adapted from Han & Kamber (2001) and differs slightly from that of Figure 17, but these two can be seen as complementing each other. Both categorisations are rather general, so, they allow further division into more narrowly defined tasks:

Association analysis aims at discovering any association relationships or correlations that frequently occur together among a set of items. An example of association is the extraction of association rules from textual databases.

Classification is the process of finding a set of models or patterns in the training data that describe and distinguish data cases or concepts. Classification constructs a model to predict a class of objects whose class type is known. The derived models may be presented as a set of if-then rules, decision trees, mathematical formulae, or neural networks.

Prediction is used for predicting the possible values of missing data or the missing class attributes of data objects. Prediction also compasses the identification of distribution trends based on the available data.

Clustering analysis identifies clusters embedded in the data, where a cluster is a collection of data objects that are similar to one another. The objects are grouped based on the principle of maximising intra-class similarity and minimising inter-class similarity. Clustering can be used to generate labels for classes of data objects and for taxonomy formation, which is the organisation of observations into a hierarchy of classes that group similar events.

Outlier analysis refers to the analysis of outlier data objects that do not comply with the general behaviour or model of the data. In some applications, such as fraud detection, the rare events can be more interesting than more regularly occurring ones, and, thus, should not be treated as noise. This is essential in aviation safety analyses and one example of this is presented later in Chapter 3.4.1 as the Find Distributions function (Nazeri, 2003).

Evolution analysis describes and models regularities or trends for objects whose behaviour changes over time. This may include classification, clustering, association and the discrimination of time-related data.

Usually the data mining systems, at least commercial products, use different algorithms combined together. In order to discover *nuggets* of knowledge hidden in the data, the analyst needs to combine advanced *machine learning* technologies with knowledge about the data (Sarlin, 2013). The process is illustrated in Figure 16. The process contains several steps or phases that must be gone through to form knowledge from raw data. Watson (1999) emphasises that to be understandable, the information, once found, must be presented as reports, graphic data or in other suitable forms.

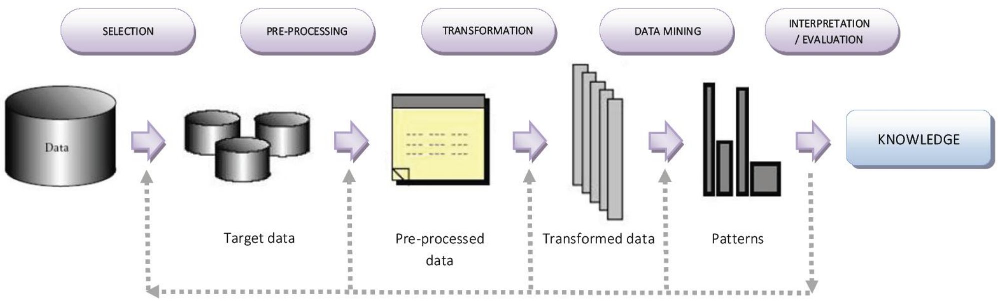


Figure 16: The Knowledge Discovery in a Database Process (adapted from (Fayyad, Piatetsky-Shapiro, & Smyth, 1996a))

Figure 16 shows the data mining processes on a general level. To define the process more, the parts displayed in the figure need some more explanation. According to Fayyad et al. (1996b) and Han & Kamber (2001), the Selection leading to the Target data contains data integration and retrieval from various data sources, in addition to choosing appropriate subsets of data for the task. During the Pre-processing, the data is cleaned and its dimensionality is reduced as needed. The Transformation is performed if the data to be processed is retrieved from various databases or other sources and the data fields vary by order and size. After the mining process itself, the interpretations and evaluations of the discovered patterns are presented and evaluated on value terms. In order to deepen the view of data mining with a process-oriented view, three classes (discovery, modelling and forensic analysis) of data mining activities can be distinguished, as presented in Figure 17:

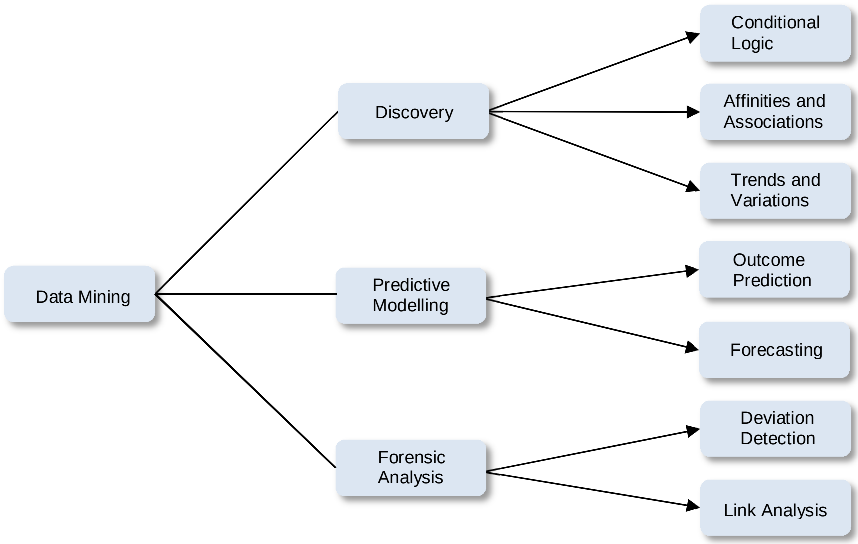


Figure 17: Data Mining Activities (adapted from (Parsaye, 1997))

The concepts presented in Figure 17 need some clarification. Parsaye (1997) describes *Discovery* as looking at a database without any predetermined idea or hypothesis in order to find hidden patterns. In *Predictive Modelling*, as the concept itself expresses, patterns that are discovered from the database are used to predict the future. Applying this method, the system may interpolate the unknown values in the records the user submitted, based on the previously discovered patterns in the database. The *Forensic Analysis* is applied to find anomalous or unusual data elements by using the extracted patterns. Compared with *Discovery*, which is used to find "usual knowledge", the Forensic Analysis looks for specific and unusual cases (Parsaye, 1997). From the point of view of this study, *Discovery* is the essential process, because the applied method is clustering (see Chapter 3.3.3) which is used to extract trends and variations; thus aiming to find and identify patterns in the flight safety data – as these are recurring safety-related deviations that might lead to an accident if not discovered in time.

3.3.2 Knowledge Discovery in Databases

Knowledge Discovery in Databases (KDD) is a discipline that has gained visibility as data collections grow exponentially and present new challenges to exploring and understanding them. The KDD is a nontrivial process for identifying valid, novel and potentially useful patterns in the data, then making the patterns applicable for practical decision support needs (Fayyad, 1996). A pattern is defined as a parsimonious summary of a subset of data (Fayyad, et al., 1996b). The KDD is formed from several fields of science – the most important here are artificial intelligence (AI), statistics, database technology and systems science with information retrieval, not to forget management information systems and cognitive psychology. The concept of Data Mining (DM) has often been presented as interchangeable with KDD by some specialists, but in fact it is only one part of the whole KDD process, as illustrated in Figure 16. Compactly expressed, the goal of data mining is to extract high-level knowledge from low-level data (Fayyad, et al., 1996b; Han & Kamber, 2001). The data mining process does not give straight answers to questions concerning information retrieval. Data mining results, which are often presentations of discovered patterns, need to be carefully evaluated by consolidating them with domain knowledge. This evaluation phase requires specialisation and experience of the problem area being investigated and can be considered to be more a combination of art and common sense than pure science (Kloptchenko, 2003).

Sarlin (2013) clearly illustrates the KDD process in Table 5.1, which is adapted from Kurgan & Musilec (2006). This presentation concerns different versions of the KDD process that are related to each other. Sarlin emphasises that there is no single process that could be named KDD. In the table mentioned, four different versions of the KDD process and one generic example are presented. The content of the table can be summarised as containing six phases (Sarlin, 2013):

1. Domain understanding
2. Data understanding
3. Data preparation
4. Data mining
5. Performance evaluation
6. Knowledge consolidation and deployment.

The connection between data mining and KDD is clearly expressed in the list above, as knowledge discovery is a comprehensive process and data mining is one essential step in it. Sarlin (2013) emphasises the key objective of the process, which is gaining necessary knowledge about the application area. Without this, there are not many possibilities for successfully continuing with further analysis. Only after this, in addition to formulating the key objectives, can the process be designed as a KDD process and a broad and preliminary plan to achieve the objectives can be drafted.

3.3.3 Clustering

Generally, in clustering the question is about meaningfully organising the observed data into taxonomic structures. The cluster analysis can be considered an exploratory data analysis tool that is applied in order to group different objects that belong to the same group – in a manner that ensures the association between them is at its maximum, otherwise the association would be minimal. By applying this method, structures in unstructured data can be discovered without explaining or interpreting the reasons for their existence (Čokorilo, et al., 2014). The goal of clustering can be expressed as "to minimize the intra-cluster distances and to maximize the inter-cluster distances" (Hu, Yin, & Tan, 2006; Luo & Fu, 2014). The term cluster analysis was introduced by (Tryon, 1939) to indicate the use of different methods and algorithms for assembling objects into their respective categories. The cluster analysis method is suited to situations in which the problem is coping with uncertainty, especially strategies for maintaining profitability in situations of uncertainty (Szeto & Lo, 2005).

Clustering has a broad appeal and is very useful in exploratory data analysis, which can be seen as either exploratory or confirmatory. However, regardless of the type, the key element is the grouping or classification of the measurements. Cluster analysis can be described as organising a collection of patterns into groups, i.e. clusters, based on similarity. The patterns are usually presented as either a vector of measurements, or a point in a multi-dimensional space. Clustering is especially useful in decision-making and machine-learning situations like data mining and document retrieval. Often there is not much prior information available about the data, thus as few assumptions as possible ought to be made about the available data by the decision-maker (Jain, Murty, & Flynn, 1999). The core idea in applying clustering in text mining is to search for similar documents. Described in more detail, clustering can be called an automatic and unsupervised grouping of text documents into groups. In these groups (or clusters) the documents have a high similarity between them but when compared with the documents of other clusters they are dissimilar (Rosell, 2009). Clustering can be divided into two main categories: hard and soft clustering. In hard clustering, the data item (a report in this context) belongs to only one cluster. In soft clustering, the data items could belong to different clusters to a certain extent (Paukkeri, 2012). Another often used mining method is classification, but unlike clustering, this needs a training stage using labelled documents in order to learn how to put the objects into defined classes (Li, Chung, & Holt, 2008).

Defined more precisely, if the processed data is semi-structured textual data or textual data without structure, the process can be called text mining. As a result of fast technological development, vast amounts of data in different forms are appearing in every field. Obtaining beneficial information from these data is attracting increasing interest. Data mining is an operation that processes vast amounts of data, making them useful information (Saracoglu, Tütünkü, & Allahverdi, 2008). Rosell (2009) illustrates its wide use by stating that in addition to the general view on uncovering the content and the structure of unknown texts, new perspectives on familiar and known ones can be achieved.

The vector space model is mostly used for existing clustering algorithms. It represents texts through the words appearing in them. If the texts have several common words, they are considered similar in context (Rosell, 2009). This method treats the documents as bags of words (Li, et al., 2008) and ignores the structure of the text. In other words, the number of the words or their frequency is only significant, as their mutual order is not considered (Baeza-Yates & Ribeiro-Neto, 1999; Vesanen, 2003). However, because meaning in natural languages depends strongly on the order of the words, clustering results must always be interpreted by human specialists (Li, et al., 2008). Another significant point in this context is that languages differ in what they consider a word (Rosell, 2009).

The implementation of the pre-processing of the data (see Figure 16) has been discussed in detail in Publication #1 “Data Mining in Aviation Safety Data Analysis”. Additionally, some terms and definitions in the context of clustering are mentioned here. The concept of corpus refers to a collection of texts, in this context they are in electronic form (Haverinen, 2014). Citing Rosell (2009), the concept of word or word token can be defined as a sequence of characters, separated from others by a text with common delimiters, like commas or white spaces. The words that are common in most of the documents are not relevant for describing the documents. In practice, the words that occur in more than 80 per cent of the documents are useless for information retrieval (Baeza-Yates & Ribeiro-Neto, 1999). Also frequent words (like ‘plane’ or ‘aeroplane’ or ‘aircraft’ in this context) need to be filtered out by defining and using the lists of stop words. However, when using this method, some information may be lost.

In clustering a highly inflectional language, like Finnish, a process called word form normalisation is required. The two normalisation methods are stemming and lemmatisation (Korenius, Laurikkala, Järvelin, & Juhola, 2004) – stemming was used in this study by creating a list of synonyms. In stemming the related word tokens are normalised into a single form. This means removing all prefixes and suffixes as well as inappropriate pluralisations (Miner et al., 2012; Vesanen, 2003). When stemming, the amount of data can be reduced, which is especially emphasised in languages like Finnish in which conjugations are expressed using case endings and suffixes (Vesanen, 2003). In addition, the concept of bag-of-words is also used.

3.3.4 On-Line Analytical Processing (OLAP)

On-Line Analytical Processing (OLAP) is an essential concept that often appears in the data mining context. The need to develop such systems arose among the users of Relational Data Base Management Systems (RDBMS). The father of the relational model Edgar F. Codd and his colleagues stated that the systems mentioned were not built for strong functions that could be used to analyse and consolidate data or for data synthesis, all of which belonged to spreadsheets and special purpose applications. In order to provide analysts with tools that complement RDBMS technology, OLAP was developed. The basic idea of OLAP is to offer analysts a view of the data from different perspectives and levels. New relationships, for example, can be found when investigating anomalies. This method provides flexible and fast access to vast amounts of derived data, the inputs of which may continuously change (Watson, 1999). The concept was defined by the OLAP Council (1997, p. 1) as follows: “On-Line Analytical Processing (OLAP) is a category of software technology that enables analysts, managers and executives to gain insight into data through fast, consistent, interactive access to a wide variety of possible views of information that has been transformed from raw data to reflect the real dimensionality of the enterprise as understood by the user”.

The main characteristic of an OLAP system is multidimensionality, which can be illustrated through the metaphor of a data cube. This presentation makes the interaction with the user easier and better reflects their way of thinking. OLAP systems are meant for decision support (Jukic, Jukic, & Malliaris, 2008; R. Watson, 1999). The decision support systems (see Chapter 3.1) aim to answer new and unpredictable questions in daily operations, which is why ad hoc queries need to be executed (Chaudhuri & Dayal, 2002). The queries usually involve vast amounts of data, at least thousands or even millions of records (Abelló & Romero, 2009). Commercial organisations typically have several operational databases and large organisations might even have tens or hundreds of them (Mannila, 2002). In order to fulfil expectations, OLAP systems ought to provide quick, flexible and shared access to analytical information, which requires fast access to data combined with high computational speed (Watson, 1999).

The OLAP tools are “read only” systems, used only for the retrieval of data enabling the users to easily read and interpret collected data that is structured and especially aimed at analysis and the subsequent making of fact-based decisions. Typical OLAP tools support the trend analyses of sequential time periods as well as modelling and calculations which cross dimensions and must go through hierarchies (Watson, 1999). OLAP tools can be considered one of the key factors in the decision-making process, especially in organisations that collect huge amounts of data in their daily operations. The significance of the capabilities of OLAP increases as more data accumulates in the operations of an organisation.

As the amount of the data affecting the operations of organisations continually expands, web-based options have been developed – in addition to the client-server and desktop-standalone versions (Jukic, et al., 2008). Although the processing capacity of the OLAP tools has increased since the 1990s, data mining tools are overtaking them as sophisticated analysis methods that combine artificial intelligence and statistical tools (Shim, et al., 2002).

3.3.5 Text Mining

Good text mining ability is the most important requirement when searching for an applicable data mining system. Text mining is a subclass of data mining (Visa et al., 2001). It can be described as "data mining from textual databases", which could be extended to the most popular definition of Data Mining: "Text Mining is an essential part of discovering previously unknown patterns useful for particular purposes from textual databases" (Dörre, Gerstl, & Seiffert, 1999; Hearst, 1999). Kloptchenko (2003) states that, expressed on a general level, numerical data interpret phenomena through the usage of quantitative measurements, whereas textual data describe the phenomena qualitatively. The similarities and differences between data and text mining are presented in Table 9.

Table 9. Comparison of Data and Text Mining (Kloptchenko, 2003, p. 7)

	Data Mining	Text Mining
Object of investigation	Numeric data (numbers)	Text (documents, web pages etc.)
Object structure	Relational databases, numbers	Free form texts
Goal	Predict outcome of future situations, analyse the reasons that affect the desired outcome, visualise data interrelations	Retrieve relevant information, distil the meaning, categorise content, compare and evaluate texts
Methods	Machine learning: decision trees, genetic algorithms, neural networks; statistics; multinomial regression, regression analysis	Indexing, special neural networks, clustering and categorisation algorithms, linguistics, ontologies
Maturity	Broad implementation from 1994	Broad implementation from 2000

In Table 9, the mining characteristics of numerical and textual data types are provided, but other types of data are also available in the aviation context. The flight data recorders (FDR) colloquially called 'black boxes', record digital data and have the capacity to save flight data over a time period exceeding 24 hours. Usually, airliners are equipped with two separate recorders on board, placed at the rear of the fuselage, although they may also be combined into a single unit. The FDR records the operational performance data from the systems of the aircraft. According to the latest requirements, the FDR must have the capacity to monitor at least 88 parameters. These include different pressures, speed, time, heading, accelerations, fuel flow, the positions of the control surfaces, etc. (SKYbrary, 2014). The flight data recorders of the most modern commercial aircraft have the capacity to record and save several hundred discrete and continuous parameters several times per second for the entire duration of the flight (Matthews et al., 2013). The other recorder (or part of it), cockpit voice recorder (CVR) (Federal Aviation Administration (FAA), 2006) records all the voices in the cockpit usually for the two last hours, the most important of which are the conversations between the pilots and between the aircraft and air traffic controllers (ATC). The radio traffic between the aircraft and ATC as well as the radar data are saved by the ATC units. On Germanwings Flight 9525, on 24 March 2015, when an aircraft was deliberately crashed into the French Alps, the CVR recorded the breathing sounds of the co-pilot in the cockpit, witnessing that he was alive to the end.

It has been estimated that over 80 per cent of the information in reporting is written in an unstructured, textual format (Megaputer Intelligence Inc, 2015b; Zicari, 2012). An interesting discovery is that this figure is exactly the same as the estimate given for company information (Tan, 1999). It can be assumed that for both areas this figure has not changed, at least not substantially. In both cases the amount of information written in documents can be assumed to be as significant as before. From this point of view, it is possible to argue that without text mining, only 20 per cent of the information would be analysed and nuggets of valuable knowledge would not be noticed. One can only imagine how weak the basis for a decision is if only one fifth of the available information is used for decision support. Its significance cannot be overemphasised for flight safety – in which even a tiny failure can lead to fatal accidents with a number of casualties. In business this could mean huge monetary losses. However, the weight of the narrative part of flight safety reports varies a lot, in some reports it may contain only a couple of words, which, in addition to the information in structured data, explains the whole incident, e.g. “engine failure” which is expressed with one word in Finnish.

Natural Language Processing (NLP), as an interdisciplinary field, includes both human language and technology. Jurafsky and Martin (2009, p. 1) have defined its purpose as to “get computers to perform useful tasks involving human language, tasks like enabling human-machine communication, improving human-machine communication, or simply doing useful processing of text or speech”. In

NLP a wide variety of different tasks can be distinguished, beginning from the simple counting of words to machine translation, where the text is translated from one language to another automatically or to information extraction, i.e. extracting facts from text. Expressed generally, one target of NLP is to understand natural language on some level (Haverinen, 2014). The difficulty in analysing textual data compared to the analysis of numerical data can be shown in several ways. With computational tools it is possible to analyse tens of dimensions of numeric data. Textual data consists of tens of thousands of dimensions which makes it highly multi-dimensional (Fayyad, et al., 1996b) and offers computational challenges of another class.

Text mining in Finnish sets its own requirements compared with, for instance, mining in English. Finnish is a compound rich and very inflectional language having an inflectional and derivational morphology, which is considerably more complex than that of the Indo-European languages (Arppe, 2002). Furthermore, English has only two case ending forms, whereas there are 15 of them in Finnish (Karlsson, 1987). Regarding the amount of words, in a comparison between English and Finnish, the study by Paukkeri (2012) can be presented as an illustrative example: the corpora used contained 12,600 unique word tokens in English, the corresponding number in Finnish was 37,000. In Finnish, the case endings correspond to prepositions or postpositions in other languages allowing an enormous number of possible distinct word forms. A noun might have some 2,000, an adjective 6,000 and a verb 12,000 forms. If the derivation is also taken into account, the figures will be increased by a factor of approximately 10 (Koskenniemi, 1983). The Finnish word token “alusta” provides a good example of the different forms that make text mining more complicated in Finnish and clearly illustrates the need for human experts to interpret the mining results. The example was produced by using the FinTwol (Finnish Morphological Analyser) program developed by Lingsoft Inc. (2009) using its trial version. The token mentioned can be understood in several ways as follows:

- alku (beginning) N ELA SG
- alusta (chassis) N NOM SG
- alustaa (format) V PRES ACT NEG
- alustaa (format) V IMPV ACT SG2
- alustaa (format) V IMPV ACT NEG SG
- alunen (underlay) N PTV SG
- alus (vessel) N PTV SG

The first row presents the singular elative case of a noun meaning the beginning. The only nominative singular noun form for the token is shown on the second row, meaning a car chassis or house foundations. The third one can be understood as a verb in its negative active present tense meaning not to format. The two next forms are active imperative verbs, the first of them in the second person singular form and the second in the negative form. It is, however, worth noticing that the negation is a construction with the main verb and negation, which is a verb itself in Finnish. The two last expressions are both singular nouns in partitive cases of the words underlay and vessel.

3.3.6 Big data

Today, organisations can access a continuously increasing huge mass of information although they have major difficulties in valuing it due to its mostly raw form, often in semi-structured or even unstructured formats. In addition, the organisations are often unaware of its worth (Kaisler, Armour, & Espinosa, 2014). In the context of storing and analysing data for promoting the operational preconditions of the organisations, the term Big Data often appears in the discussion. A vast amount of data, saved in several databases for instance, cannot be considered big data as is because a couple of requirements need to be met. Regarding volume, Kaisler et al. (2014) talk about data warehouses with the size of a terabyte (10^{12}), defining the concept as "the amount of data just beyond technology's capacity to store, manage and process efficiently" (Kaisler, Armour, Espinosa, & Money, 2013, p. 995). They continue by predicting that the creation and aggregation of the data will accelerate from petabytes via exabytes to zettabytes. Principally, data processing is performed in a similar way to standard data mining through iteration in successive phases, beginning from defining the study goal and ending at interpretation but passing through stages that explore the availability of the data and its nature, and extracting and preparing it (Müller, 2013). However, the mining technology presented in this study works well on the scale of gigabytes, but might not be applicable for terabytes of data. Kaisler et al. (2014) clarify these characteristics by referring to Laney (2001) when stating that the complexity of big data is the focal factor limiting the utilisation of the conventional means of analysis.

Furthermore, Schroeck et al. (2012) have defined the requirements as the 3+1 "V":s of Big Data, which are: Volume, Variety, Velocity and Veracity (Sathi, 2012). Regarding volume, the point of view is data at scale, meaning terabytes to petabytes of data. Variety is about data in many forms being structured data, unstructured data as well as text and multimedia, thus the ability to manage the complexity of different data types is required. Today data is created, processed and analysed with a speed that continues to accelerate, so the velocity (data in motion) is of great importance for streaming data, because enabling decisions within fractions of a second requires the efficient analysis of data. The last "V" stands for veracity, indicating data uncertainty, which requires skills for managing the reliability and predictability of different data types that are often inherently imprecise (Schroeck, et al., 2012). Müller (2013) cites LaValle (2009) when noting that even a third of business leaders lack the relevant and appropriate information needed to make decisions, leading to situations in which they cannot place trust in the data they are using. LaValle (2009) continues with the observation that half of all business leaders are forced to do their job without having sufficient information from across their organisation, emphasising the characteristics of our information today – widely varied in format, structured and unstructured, voluminous with extreme velocity coming from within and from outside the organisation, arriving on a daily, hourly and real-time basis (LaValle, 2009).

The development of data analysis for decision support can be illustrated as a three-step model: the so-called conventional analysis tools – mining tools, big data analysis tools and methods. From this point of view, the mining might be seen as obsolete because big data thinking is coming. However, most organisations are still in the first phase and the mining tools have still not been taken into use on the scale they should. And, when the practical mining tests for this study were originally made, the term Big Data did not exist.

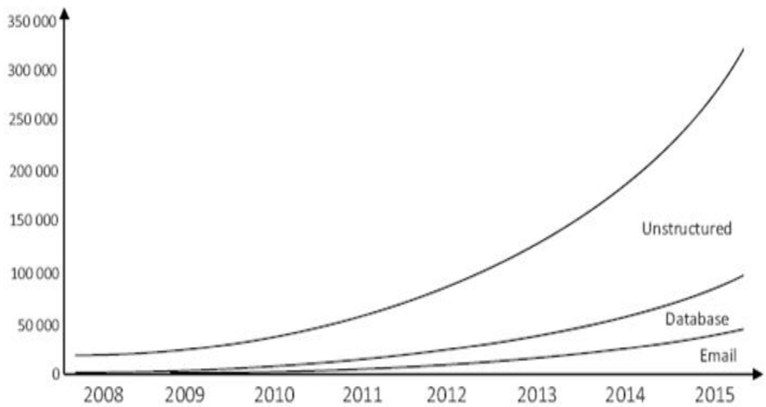


Figure 18: Global total archived capacity in petabytes (adapted from (Müller, 2013))

The development of the volume and the variety of archived data is displayed in Figure 18, showing the vast amount of unstructured data as well as its expected growth. What is shown here can be applied within safety management and especially flight safety management. Müller (2013) sets the role of big data handling by describing some of its essential characteristics as being the use of naturally occurring, open data sets. The processing should then take place by applying a statistical model based on machine learning and data mining – from a predictive, hypothesis-free and inductive point of view. Data processing is similarly performed with ‘normal’ data mining through iteration in successive phases, beginning from defining the study goal to interpretation that goes through an exploration of the available data and its nature, and extracting and preparing it (Müller, 2013).

Mehta (2011) highlights the contemporary situation by looking to the future. According to him a change needs to happen, because the value of managing, processing and modelling vast amounts of data has historically been underrated. Data specialists are of great importance, a key resource in an organisation, possessing more value than any other resource – with the only exception being the data itself. An illustrative example of the predictive power of big data analysis was produced by Müller (2013) on the detecting of influenza epidemics, when a search engine query data was used. A model was built and a mean correlation of 0.97 was obtained when generating the estimates. This, for one, strongly supports the statement that harnessing the collective intelligence stored in the vast amounts of data search logs can provide broad-reaching, timely and available monitoring systems for phenomena like influenzas.

3.4 Data Mining in Aviation Safety Data Analysis

3.4.1 Previous data mining research within aviation safety

Articles and other material concerning data mining in aviation safety are uncommon. The subject seems to be still rather new and unknown and unexplored, although the first articles on it were written in 1997. There has been a significant time gap between the ideas to use data mining in flight safety and the first prototypes. One explanation for the lack of reports about applications for improving flight safety is that safety is a very critical factor in aviation, especially in commercial air traffic, which is why information concerning it is kept inside organisations. One extremely interesting and, from the point of view of this study, essential article is “Application of Aviation Safety Data Mining Workbench at American Airlines”, published by MITRE Corporation (Nazeri, 2003). The study was executed in co-operation with Global Aviation Information Network (GAIN, see Chapter 3.2.2) using the Aviation Safety Action Program (ASAP) data of American Airlines. This proof-of-concept project was executed in order to demonstrate the suitability of data and text mining tools in the aviation safety data analysis. Additionally, another essential purpose was to assess the usability of these tools for enhancing internal airline safety analysis. The workbench was built to consist of three tool modules: FindSimilar, FindAssociations and FindDistributions (Nazeri, Bloedorn, & Ostwald, 2001). FindSimilar searches the collected flight safety records for those objects which are most similar to a chosen target safety record. FindAssociations is meant for finding subsets that have an interested correlation, like events of the same incident type, or a reoccurring location. FindDistributions focuses on an attribute, specified by the user, to find exceptions, i.e. anomalies to the common distribution. These exceptions might contain significant value for analysing unusual safety-related events, such as an emerging problem like the consequence of changed procedures. Another remarkable MITRE research project (Bloedorn, 2000) for mining aviation safety data was performed by applying the so-called hybrid approach, as different mining processes for different data types also require different methods of analysis. This approach was chosen because the input data contained both structured and unstructured data fields, and the ability to find similar cases from this mixed-type data was required. Clustering (as described in Chapter 3.3.3) was also utilised in this project.

One more fundamental data mining project in the aviation safety context to be presented here was the “Application of PolyAnalyst to Flight Safety Data of Southwest Airlines” (Megaputer Intelligence Inc, 2004a). This joint proof-of-concept project was conducted in co-operation between Megaputer and Southwest Airlines in support of Global Aviation Information Network (GAIN). Its goal was to demonstrate the benefits of applying the PolyAnalyst data and text mining system in the analysis of airline flight safety data. The document of the results of this project provided the basis for this research process, as the capabilities of PolyAnalyst for efficiently processing large volumes of mixed structured and unstructured data were proven. As a sophisticated mining tool, it demonstrated the automated extraction of important patterns and clusters from text, also providing an intuitive drill-down to the original data in order to support the findings. The results of the analysis could be reported as clear-cut visual graphs (Kollepara & Ananyan, 2003). The PolyAnalyst was to be one of the tools to be tested in this research project. However, as it proved to be language dependent and Finnish was not included in the languages supported, it was impossible to test its text mining capabilities with flight safety reports written in Finnish. Despite this, it is briefly presented as Tool #1 in this thesis (see Chapter 5.3.1).

A rather recent study conducted by Čokorilo et al. (2014), used a database of over 1,500 aircraft accidents worldwide: all events classified as an ‘Accident’ according to the ICAO Annex 13 (ICAO, 2001). The accidents had occurred between 1985 and 2010. The study focused on comparing the accidents in relation to the characteristics of the aircraft, traffic type, route and the environmental conditions. The data was aggregated into groups using cluster analysis, the same method as this study (see Chapter 3.3.3). Furthermore, for each cluster a hazard index was created in order to define the dangerousness of the clusters concerning aviation safety. The study characterised each accident focusing on environmental conditions and the aircraft’s physical characteristics. As a process of feedback was applied, it enabled the identifying of the most relevant variables that describe the phenomenon. The model achieved a high level of significance with an excellent coefficient of determination R_2 value 0.96, having an estimation error of about 13 per cent. The obtained results enabled the establishment of an easy-to-use predictive model for accidents (Čokorilo, et al., 2014).

In order to find doctoral theses and other academic peer-reviewed documents published beyond Scandinavian countries (for Scandinavian dissertations, see Chapter 1.3), a database search in spring 2015 was conducted using the platforms Open Access Theses and Dissertations (OATD) (2015) and ProQuest (2015). When the search: subject: "aviation" AND subject: "safety", were used in OATD, the search produced 40 results in the time span between 2006 and 2015, including articles as well as dissertations. Although none of them dealt directly with safety analyses, the keywords that were close to the approach of this study like risk assessment, human factors, threat and error management, information quality, incident reporting, safety measures, etc. appeared in the documents. In addition, many of the documents dealt with technical and juridical aspects. When the search was elaborated to definitions combining data mining with aviation or flight, the only search producing results was: subject: "aviation"; OR subject "flight"; AND subject: "data mining" when three documents were found. One doctoral thesis was found to be quite close to this research, as its research subject was finding factors that lead pilots to submit voluntary anomaly reports. This study applied traditional statistical methods and classification. The two others were written on an engineering basis, dealing with purely technical questions.

The ProQuest platform allowed more detailed definitions regarding criteria, for instance, only peer-reviewed documents could be chosen and document type, source type i.e., scholarly journals, conference papers and proceedings; trade journals; title of the publication; location; language (122 alternatives); classification i.e. computer science, industrial engineering, aerospace engineering, management, aerospace materials, etc. could be defined in the search processes. The platform searched documents from 47 different databases. The search words "aviation safety" AND "data mining", were limited to dissertations and 104 results between the years 1999 and 2015 were found. The search criteria mentioned allowed both of the expressions appear anywhere in the documents, thus many of them did not deal with aviation at all; the term was only mentioned in the contents of the document.

Only a few theses were found that were relatively close to this study and dealt with mining aviation safety data: doctoral theses by Jones (2008), Nazeri (2007), and Forrest (2006). The thesis by Jones approaches the mining problem from the point of view of building and testing algorithms in order to find valuable relationships in the tested data. The doctoral thesis by Nazeri (2007) was prepared while Nazeri was an employee in MITRE Center for Advanced Aviation Systems Development (CAASD). During the study, five major national aviation safety databases in the US were analysed, focusing on accident and incident data. The relationships of the data were determined and significant accident factors in the data were identified. This study proved that the relationships mentioned could be shown in terms of factors leading to these events, by applying mining across multiple national aviation safety databases. The problem investigated in Forrest's (2006) thesis deals with finding possible solutions to policy issues in public disclosure that prevent the collection and sharing of aviation safety information among different organisations. The research was performed in co-operation with the Global Aviation Information Network (see Chapter 3.2.2) and the information systems proposed by the organisation mentioned are complex, requiring taxonomies and tools exceeding those normally used in the analyses of traditional information systems.

One more document worth mentioning here is the master's thesis by Gonzalez-Bernal (1999), which presented different data mining methods, approaching the research object on the algorithm level. When the definitions were adjusted to be more accurate, limiting the search criteria to keywords (identifiers): "aviation safety" AND "data mining", the result was only three dissertations, those of Jones, Nazeri, and Forrest already mentioned before.

3.4.2 The challenge of unknown

The states of knowledge are shown in Figure 19. This presentation was originally created to meet the requirements of the business world, but it can also be applied to the context of the present area. State 1, know what you know, can well be described as "business as usual", meaning that the fundamentals as well as the specific relationships of the operations of the organisation are known and understood. In this state, measuring and reporting with publishing are the predominant activities, although a false sense of security can be induced. In State 2, in which it is known what is not known, it has been recognised that changes are required in order to sustain the operations. The existence of the problems is known but in order to address them, further analysis is necessary. When the problem area is defined, techniques like ad-hoc or On-Line Analytical Processing (OLAP, see Chapter 3.3.4) are very effective in understanding and quantifying the solutions. State 3, don't know what you know, represents the classic case of knowledge management, which is discussed in Publication #5 "Data Mining Challenges in the Promotion of Aviation Safety". The problem in this state is sharing all the experience, discoveries and techniques among the analysts and the managers. State 4, don't know what you don't know, is the most interesting but precarious state, because what exists but which people are unaware of might be dangerous. Unawareness about certain issues and opportunities will adversely affect an organisation's performance and might even affect future development. Data mining methods will provide some possible solutions for dealing with problems in this state (PolyVista Inc., 2005).

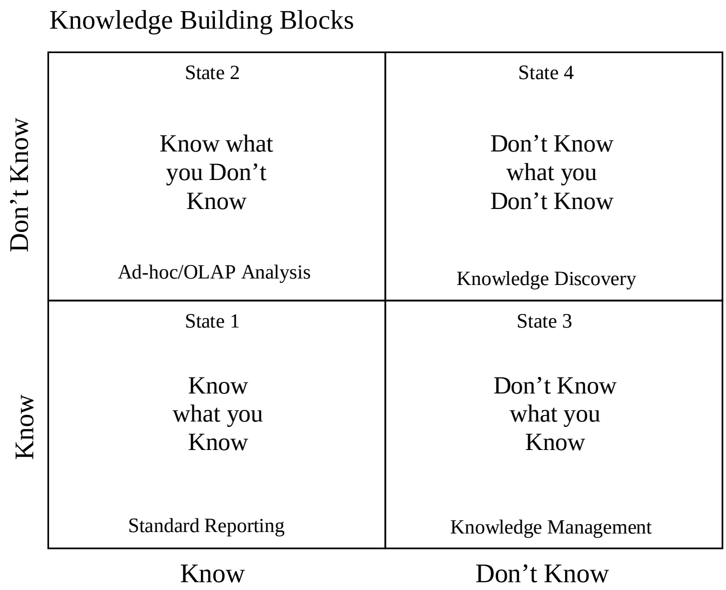


Figure 19: States of knowledge (adapted from (PolyVista Inc., 2005))

The model can be directly applied to flight safety and thus to this study. As stated before, in the first phase the business runs as usual with standard reporting. Normal reporting about incidents and other deviations is made, which is followed by the analysis and the updating of instructions. As mentioned before, there might be an unrealistic sense of security if there is no idea about the existence of an problem that might affect operations. If the operational environment is safe, this sort of approach can be seen as sufficient. In the second state there is awareness about an unknown (unidentified) problem and that is why reporting and other activities need to be developed. There is more or less an acceptance of weaknesses being present either in the operations or in the organisation. The third state, knowledge management, is today familiar to almost all organisations, the bigger the organisation the more knowledge is needed and the bigger the scale of its operations are. The most significant state from the point of view of this study is the fourth one in which 'the unknown is not known'.

In theory, safety factors are those that are in closest correlation with accidents, including casual factors, like personnel capabilities concerning pilots, air traffic controllers, the air traffic environment, aircraft capabilities, weather conditions and unpredictable acts (Čokorilo, 2011; Wells & Rodrigues, 2003). Hazard identification methods are used in order to identify safety factors. In modern day safety analyses, different tools and methods are applied. The most commonly used are statistical and trend analyses, normative comparisons as well as simulation and testing, expert panels and cost-benefit analyses. Generally, all the hazard identification methods can be divided into two groups (Čokorilo & Dell'Acqua, 2013):

1) Reactive methods
2) Proactive methods.

The first mentioned are based on the monitoring of trends as well as the minute investigation of events concerning safety. The latter identify hazards by analysing the performance and function of systems in order to discover intrinsic threats and potential failures. Examples of these are safety monitoring and assessment, operational safety audits, etc. (Čokorilo, et al., 2014).

4 METHODS

4.1 Spectrum of Methods

In order to produce relevant and useful research, the right research methods should be found and used; thus the choice of the most suitable method or methods is a key issue. Research methods can be divided in many different ways. One rather common approach is to talk about quantitative and qualitative methods. This rough division is commonly used, but in this context it would not be at all sufficient. However, this division is worth taking into account in this context. The choice between different research methods is not very easy. Many components must be taken into account when choosing the method. The most important factor should be what the researcher is trying to find out. This is especially true for information systems, which are not valuable in themselves if they do not have any special purpose – their existence within an organisation is based on their potential to improve its effectiveness and efficiency (Hevner, March, Park, & Ram, 2004). This should not be forgotten in the process of choosing the right methodology. The right 'box' in which this study can be put is the usefulness of innovations, or the better evaluation of innovations through the study of their components.

Another important point of view is to focus on how and where the research material is collected and the observations are made. For example, action research is made when the researcher is part of the examined organisation. At that time, a case study is often carried out in connection with this because in many cases it is relevant to limit the study to a context in which the researcher can easily belong. A clear and often used classification for the taxonomy of research methods made by Järvinen (2004) is presented in Figure 20. This is not, of course, the only and one way to display the relationships between different approaches, but it provides a good starting point for developing the methodological process further.

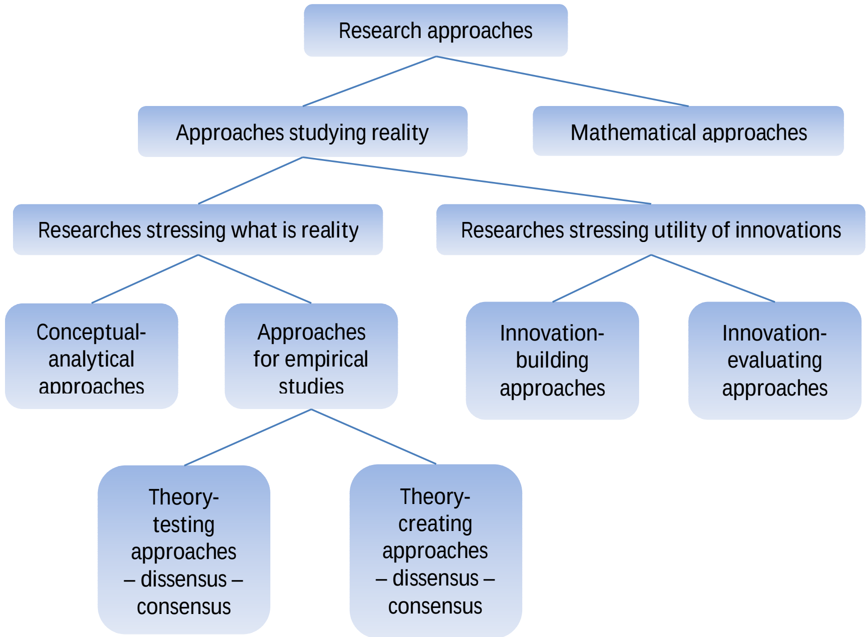


Figure 20: Taxonomy of research methods (adapted from (Järvinen, 2004))

4.2 Quantitative and qualitative methods

Very often, when talking about research methods, the main division is made into quantitative and qualitative methods. A quantitative method means research which is exact and calculative, for example, in sociological research, statistical methods are often used. The term is commonly used within sociological and educational sciences, but in wider terms it could mean research carried out in the natural sciences. The aim of the researcher in quantitative research is to collect empirical observations. By thoroughly examining the observations, the researcher tries to understand phenomena and to create generalisations based on the mass distribution of empirical observations. The collection of observations can happen through personal interviews or a questionnaire. The quantitative method is suitable in cases where large groups of individuals are to be studied. Market research can be mentioned as a good example of this kind of research. However, extensive information about individual cases cannot be achieved. Statistical models are often used as instruments for the quantitative method (Given, 2008).

The qualitative method is often used in addition to the quantitative method in the social sciences. The main target of a qualitative method is to understand the studied phenomena. This means defining the meaning or purpose of the phenomenon and forming a total and deeper impression of it. In practice, this often means giving space to the points of view and experiences of individuals as well as exploring the thoughts, feelings and motives of the studied subjects or groups of subjects (Hirsjärvi & Huttunen, 1995). However, although the author of this thesis does not agree with this, it seems to be accepted by many researchers that qualitative research is not hard science. This view is supported by Denzin and Lincoln who note that qualitative researchers are sometimes called journalists or even soft scientists. Their work is considered to be strongly "unscientific or only exploratory, or entirely personal and full of bias" (Denzin & Lincoln, 1994). A theory can be seen as having two roles in the development and use of qualitative methods: first, as a means to assist in the research process, and second, as a target if the aim of the research is to develop the theory further. In the first case, the research process needs both a background theory against which the data can be estimated and an interpretation theory that assists in defining questions as well as what to search for in the data. A theory can also be the target for qualitative research: this happens when data is processed through inductive reasoning. In this case, the aim is to create new theoretical knowledge (Eskola & Suoranta, 2003).

The process of examining narratives that are the products of a special case developed through personal subjective creativity can be taken as social research, the goal of which is to understand the phenomena on a deeper level and the deviations or the incidents that are typical of a context. Carlsson (2009, p. 30) expresses the role of it simply and clearly, "Qualitative research methods help researchers to understand people and the social and cultural contexts in which they live – this can never be captured with formal methods". This can be applied to this study as such, although it is essential to point out that this study is concerned with what happens among the people in the social and cultural context of aviation. Carlsson continues with the question of choosing methods by juxtaposing a couple of factors. Starting from the research focus, a comparison of the objective compared to the subjective can be made. The second point concerns general laws (nomothetic research) compared to uniqueness (idiographic) about the possibility of generalising the results. The two others, which are especially important for this research, are prediction and control compared to explanation and understanding and the perspective of the outsider (etic) compared to the insider (emic).

One significant question for the researcher when choosing between the methods is that these methods are often evaluated differently. An overview of the claimed features of qualitative and quantitative methods is presented in Table 10.

Table 10: Features of qualitative and quantitative methods (adopted from (Silverman, 2000))

Qualitative	Quantitative
Soft	Hard
Flexible	Fixed
Subjective	Objective
Political	Value free
Case Study	Survey
Speculative	Hypothesis testing
Grounded	Abstract

Qualitative research is not usually based on hypotheses. These arise from the data with minimal presumptions. The researcher cannot, however, be totally free of them, and this is something he or she should clearly be aware of. Consequently, hypotheses can be regarded as presumptions expressed explicitly and the researcher should be able to use them as work-hypotheses, i.e. the conjectural results of the research process. One of the main tasks of qualitative research is to assist with creating new hypotheses to be used later as the basis for the quantitative research (Eskola & Suoranta, 2003).

The question about hypotheses for this thesis is somehow twofold: on one hand, it could be stated that there are no hypotheses that can be held to be true. There is no formal hypothesis that can be proved and there will be at least a number of lethal trends that are found among the 1,200 test cases. On the other hand, a sort of hypothesis could be seen to exist. Although data mining needs no special objective for the search, there is an assumption that something could be found using this method. Otherwise it would not be worthwhile using any data mining at all. More generally expressed, the foundation of risk assessment is the core assumption that the potential for catastrophic consequences, as well as smaller and less meaningful events and deviations, always exists (Macrae, 2007). Thus, the basis of this study is that there might be lethal trends to be found among the test cases; this can be considered a kind of hypothesis. What is said about the qualitative method creating new hypotheses might become true as the research process progresses. If one or several interesting trends are found, a deeper examination would require more data in order to find out the characteristics of the findings, which would emphasise the role of statistics when more accurate information about a specific area is required.

Sampling is often discretionary in qualitative research. The number of units to be examined is usually not big and the research is often very intensive, which means that the requirement regarding their quality is high. The size of a dataset has its own meaning as well, because the data should be suitable for the planned analysis and its interpretation. Hence, choosing the data will happen purposefully and be based on theory (Eskola & Suoranta, 2003). Most of the characteristics of qualitative research in Table 10 can be applied to the research carried out for this thesis. It is a soft approach, it is flexible (or, at least it should be) and in this context it can be regarded as a case study, it is speculative and does not test a hypothesis. As for the other points, this research is more value free than political, positioned somewhere between subjective and objective, because the aim is to extract objective knowledge from subjective reports.

The last point – the division between grounded and abstract theory – is worth studying in more detail. The goal of grounded theory is to develop a theory grounded in systematically collected and analysed data. Another significant characteristic is its aim to build interplay between the analysis and the collection processes of data. It is useful in developing context-based, process-oriented descriptions and explanations of phenomena (Carlsson, 2009). Making a clear distinction seems difficult in this case, if not even impossible. This study will not create any theories, which would classify it as an abstract study and thus place it on the quantitative side. On the other hand, this fulfils at least some of the qualitative focal features described by Carlsson; the study develops context-based (flight safety), process oriented (iterative data mining) descriptions and explanations of phenomena (the recurrence of deviations and occurrences). Working with real data collected in a real world context often gives rise to methodological dilemmas; the present study is a point in question.

4.3 Design Science research

According to Järvinen (1999), it is typical of constructive research that it builds a new artefact. This is based on existing knowledge and/or advances from technical, organisational and similar origins. The methodology for designing artefacts is often called constructivism. Kasanen et al. (1993) say that the focal points are the entities solving explicit problems. These could be associated with the organisation or management by building models, diagrams, plans and so on. The basic presumption is that when something is created by developing a construction, it differs profoundly from anything existing before. The terms design science and constructive research can be used interchangeably. According to Simon (1996) the design-science paradigm originates from engineering and artificial sciences. Its best use is as a paradigm to solve problems on how to create innovations and products. The innovations here support the definition of ideas, practices and technical capabilities.

The essential question is to find out how well the present study fits with design science. The aim is to build or develop something new – a system that has not been previously applied – the text mining of aviation reports. The basis of the whole process is existing technology and existing processes of analysis. There is a need to solve explicit problems and to find new and more efficient analysis methods as the goal is to create innovation with the help of existing software tools. Because data mining will be used, the basis can be considered rather technical. The analysis and design of information systems, their implementation, management and use can be accomplished efficiently and effectively with text mining systems (Denning, 1997; Tsichritzis, 1998). Venable (2009b) supports this view and states that design science research should aim at solving problems, such as the elimination, reduction, or alleviation of undesirable circumstances. This can be accomplished in two ways, either compensating for undesirable circumstances, or eliminating or reducing one or more of the causes. He continues with the definition of design science research as "research in which a new or improved solution technology is invented".

The constructive research process can be divided into the following phases, that have been used in many contexts (Koskivaara, 2004):

1 *Find* a practically relevant problem
2 *Examine* the potential for long-term research co-operation
3 *Obtain* a deep understanding of the topic area
4 *Innovate* a solution idea and develop a problem-solving construction
5 *Implement* and test the solution
6 *Ponder* the scope of applicability of the solution
7 *Identify* and analyse the theoretical contribution.

This study fits well with his description. (1) There is a practically relevant problem, the lack of a tool for clustering flight safety reports into groups that can be analysed in detail by professional inspectors and analysts working for flight safety authorities and commercial airlines. (2) There is potential for long-term research co-operation because the target organisation is an authority. (3) A deeper understanding of the topic area will be achieved during the iterative research process because it will refine the results and constantly create new understanding. (4) The innovation of a solution idea is to be achieved through the choice of using data mining tools to be used in the extraction of information from the problem area. The problem solving construction can be seen both in the systems mentioned and the adjustment of the systems to be tested. (5) The only way to test the applicability of text mining as an method of analysis is to use real world data, which has been done in this research – the implementation was made during the testing process. (6) The scope of applicability for the tools tested in this study could in principle be made in many application areas in which data mining is relevant; in our case, however, the lists of stop words and synonyms have been tuned to give the best mining results in the specific area of this study. (7) The theoretical contribution is that the text mining of short text reports in Finnish can be carried out and that it gives meaningful results.

Another view on the design science process according to March and Smith (1995) and Järvinen (1999) is presented in Figure 21.

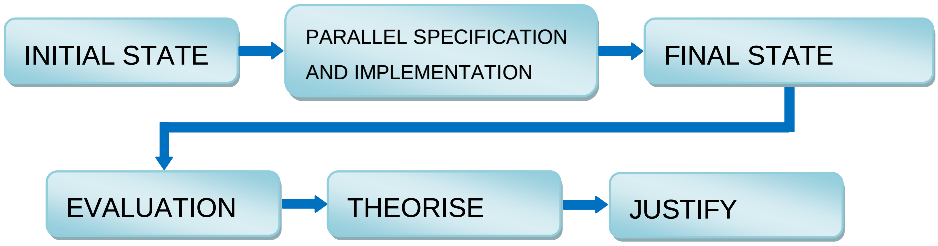


Figure 21: The design science process (adapted from (Järvinen, 1999; March & Smith, 1995))

Several important factors have been presented in Figure 21. It is worth noticing that the process does not stop at the final state where the desired result – an applicable system, or part of it, has been achieved. The evaluation should not be forgotten, nor the theorising and justifying. The whole flow shown in the figure is similarly described in the list of seven phases that Koskivaara (2004) presented. The same elements can be found in both presentations, except the last step in Figure 21, which could be mentioned separately. Although the estimation of the whole process and its performance capability as well as its results, done by professionals, is always required, the last point, i.e. justify, is important. The end of the whole process should be the full applicability of the innovation to the problem area and this can be achieved through justification – if it is actually achievable. Venable presents four research methods and paradigms which frame design science research as follows (Venable, 2009b):

1 Theory Building
2 Solution Technology Invention
3 Artificial Solution Technology Evaluation
4 Naturalistic Solution Technology Evaluation.

The three ways to describe the design research process differ from each other, but they all show that design is a process consisting of many phases in which there are two main phases, first defining, planning and building the artefact. Secondly, evaluating it before taking it into use. This requires testing and proofing, which, determines the usability and applicability of the new creation. The division presented by Venable is useful. Artificial Solution Technology Evaluation is made with computer simulations, role playing simulations, field experiments and lab experiments, while Naturalistic Solution Technology Evaluation is carried out with case studies, survey studies, field studies and action research.

The use of the case study is suitable in this context, because the data used was from the Finnish Civil Aviation Authority and the system would have been built for its use. A case study describes the relationships that exist in a particular situation – usually in a single organisation. Action research is also often carried out in one organisation and is a possible choice of research method because at the beginning of the research process, the author was working as an inspector in the Investigation and Analysis unit at the Finnish Civil Aviation Authority. The action scientist can be seen as intervening in the client system with the aim of seeking to promote learning in the organisation and to contribute to its general knowledge (Argyris, Putnam, & McLain Smith, 1987).

During the evaluation of the possible research methods and tools to be used for research, there are two types of error that might occur (Venable, 2009a): a Type I Evaluation Error (false positive) or a Type II Evaluation Error (false negative). In the first case the system (methods, tools) fails although the evaluation process proved it to be a good choice. In the second case the evaluation produces a negative result although the system (methods, tools) is working properly. It is clear that both error types will cause problems for the whole process regardless of the context. For flight safety procedures and processes, all error types should be eliminated, because mistakes will, in the worst case, lead to the loss of human life.

4.4 Action research

The role of action research is not totally clear, because it can also be seen as a subset of case studies and field experiments. It is still regarded as a separate approach based on its underlying philosophy (Antill, 1985; Wood-Harper, 1985). The presence of an action researcher affects the studied situation and the researcher is aware of this. In fact, action researchers should actively associate themselves with the practical results of the research as well as try to find and identify the theoretical outcomes (Foster, 1972) In addition to this, Clark shows the risk of the roles of the researcher and research subject becoming reversed in action research studies (Clark, 1972). Action research has also been described as close to or even equivalent with consulting. This is obvious because in both cases the target organisation is often a business organisation. Consulting cannot be accepted as research unless the aim is to create and validate new theory or to verify published theory constructs, which may not be possible in a consulting context. Action research (AR) can, in some sense be consulting (discussed more later), but problem-solving approaches and methods are based on research results, not adapted from a company.

An action research method provides a good opportunity for the successful combining of researcher and the practical knowledge accumulated by an organisation. It is important that both parties should benefit from the outcome of an interactive action research process. For the researcher, a deeper understanding of the research problem should be reached, while a more systematic understanding of the decision situation should be gained by the organisation (Reponen, 1993). In order to benefit from an interactive AR process, some prerequisites should exist. Tacit inferential processes must largely be retrieved by participants who have an aptitude for it, such as dealing openly with challenges and conflicting views. The participants are also expected to reveal information that may expose the vulnerabilities of themselves and the others. It is worth mentioning, how important it is for the participants in this context to recognise and acknowledge when they are wrong, and, to feel free to choose from among competing views (Argyris, et al., 1987).

Regarding the strengths of this method, Galliers (1990) highlights two aspects: first, it obviously accrues very practical benefits for the client organisations. Secondly, the possible biases of the researcher cannot be hidden when undertaking the research. Galliers continues by mentioning the weaknesses of action research as being similar to those of case studies, which means that AR could be applied in a restricted way to only one organisation or event. The generalisation of the results of individual case studies are difficult if not impossible to make. This is a fact if it requires the acquiring of similar data from similar organisations in order to find statistically meaningful figures. This approach would, in addition, place a lot of responsibility on the action researcher, meaning that they would have to be aware of the risks of aligning themselves with a particular grouping of objects that are at odds with another grouping. This is why special attention should be paid to the ethics of the research (Galliers, 1990). The process of designing knowledge should have the human mind as its focal point as, in the action context, the limited capabilities of humans to seek and process information must be taken into account. The knowledge created during the research process should take account of the normative dimension and the purposes for which it becomes relevant. The forming of purposes also includes an implicit concern regarding value questions (Argyris, et al., 1987)

Reponen (2004) provides a general definition of action research (AR) the “the link between academia and practice”. This is a significant statement, especially from the point of view of this study because this has from the very beginning been the study’s foundation and, in fact, the purpose for doing it. Saarinen (2005) states that the research problems of the real world come from organisations. Thus, a real world actor benefits from research. In this research process, learning and assessment are of great importance. However, multi-dimensional problems may not be solvable using quantitative methods as organisational learning is a key issue. Thus, the researcher acts as catalyst in the organisation, as well as an active operator in its iterative process. AR requires expertise about the organisation and the starting point of the events or phenomena to be studied. The first move also has to come from the organisation's side and it has to have strong support because, as Carlsson (2009) clearly states, the "the aim is to contribute to the practical concerns of people in a problem situation and to the goals of social science through joint collaboration in a mutually acceptable ethical framework".

BenMoussa (2008) outlines the core meaning of action research in a very interesting way. First, what is action? He says, that it is something to bring about change in a community or organisation. Second, research produces scientific contributions, meaning the creation of new knowledge and providing evidence and support for it. Another significant point is to have knowledge emerge through an explicit process of enquiry. Action research is an iterative process, in which there are two loops: in the first loop action works in an iterative connection with reflection. In the second loop it has a correspondent connection with theory. This enables the linking of new knowledge with already existing knowledge. Pihlanto (1993) compares AR with other methodologies. First, "the researcher is actively involved, an actor", in the process. Second, "there is collaboration between client and researcher". Both of these statements are met by this study as the researcher has worked as an inspector in the organisation (FCAA), from which the safety data has been collected as test material. Third, "the aims and intentions of the actors have a crucial role in explaining their actions", which has been achieved. Fourth, "knowledge obtained can be immediately applied". All the results, even the preliminary ones, bring new knowledge and skills to the process of analysing flight safety reports. A small detail in the findings might provide a new direction for investigative lines. Fifth, "the research is a cyclical process of linking theory and practice". In order to find new applicable methods, the testing and tuning of definitions is unavoidable. Theory leads to definitions and during this process new pieces of theory are created to be applied in the following phases.

According to Argyris et al. (1987), action scientists are engaged in a collaborative process with participants in the chosen context of learning during a critical inquiry into social practices. The arrangement expressly provides special designs for learning the practices of the case and alternative ways of constructing it. Ben Moussa (2008) has summarised some attributes, like motivation, understanding, approach, commitment, recommendations and the choice of quality criteria for comparison. These attributes are useful for making a distinction between a consultant and an action researcher. A consultant is committed only to the client, but the commitment of an action researcher is to both the client and the research community. The science world would have more to give to the practices of the solutions-applying-world. Already, from the very beginning of this study, one of the basic ideas was to build a new link between proven practices and the world of science in order to develop new analysis methods using the latest achievements of scientific research.

Action research, like many other research methods can be approached from its deficiencies, one of which Reponen (1993) identifies as being data collection, the verification of hypotheses and the generalisation of results. He continues, however, that if these points are taken into account, action research can provide a good way of gathering new data. A successful application of AR requires relationships of trust between the researcher and the people who are the research object. The existence of research interaction is the unavoidable key to the successful application of a case study method. However, despite the presence of research action in this thesis, this method is a demanding one, especially from the point of view of its validity (Reponen, 1993).

In conclusion, the research process used in the thesis fulfils the requirements of action research, at least to a sufficient degree. In this context the increase in knowledge about flight safety is very important. A second aspect which offers an organisation actual benefit from the research process is that this study deals with a real world research problem. The third aspect, the required expertise regarding the organisation and the starting point for studying the flight safety process are both present. It is not to be forgotten, either, that new knowledge is created and evidence is provided to support this.

4.5 Case study

Case study methodology was mentioned in Table 10 as one of the main tools for qualitative studies; in the same way, surveys are the most often used tools for quantitative studies. It is a rather flexible method, because both the context and the problem as well as the choice of the researcher can define it as critical, positivistic or interpretive. These characteristics explain why it is the most common qualitative method used in IS research (Carlsson, 2009). According to Carlsson, this method is better applied in cases where more organisational than technological issues are focused on in the research. He also finds that this method is at its best when used as an “empirical inquiry that (i) investigates a contemporary phenomenon in its real-life context, and (ii) the boundaries between phenomenon and context are not clearly defined”. Case studies are widely used in both the social and natural sciences. Generally, instead of collecting large amounts of representative data in order to study its characteristics using statistical methods, cases will be chosen in order to find a few accurate observations and then to study them from several points of view. The goal with case studies is not to make generalisations or identify typical characteristics or to find cause-effect relations focus, but to describe the phenomena of the new relationships observed during the research process.

Typically case studies try to maximise the understanding of an event or a situation by examining the interplay of all variables. To reach this, a process known as thick description is needed. This is an extensive and rich collection of details about methodology and context, which will be provided in the research report. In the same context, it is important to interpret the meaning of demographic and descriptive data, which includes motives, community values, ingrained attitudes and cultural norms. As quantitative methods, like surveys, build on focused questions like who, what, where and how much and how many, case studies use questions like when, how or why. Case studies can be used when the researcher does not have control over the events, and when the study in a real life context is in real time. It is also worth noticing, that a successful application of the case method requires a problem seeking holistic understanding of the event or the situation, and inductive reasoning from specific to more general terms. Some scientists regard case studies as being interchangeable with field studies and participant observation, which are based on familiarity with events taking place in a natural setting and the striving for a more holistic interpretation of the studied object (Colorado State University, 1993-2009).

The case study was one possibility for approaching the research object, i.e. flight safety data. The Finnish Civil Aviation Authority is the subject – the case organisation for this research. As mentioned before, the case study describes the relationships that exist in a particular situation – usually in a single organisation. That is why it is used for part of this study, because it is important to describe the flight safety data requirements of the target organisation. It seems obvious, that the approach of this study should be a kind of synthesis combined from case study and action research, partly due to the limitations of the case study approach, but also because action research is the most suitable method for this kind of study.

4.6 Methodologies used in this study

Choosing the right method or methods is a complicated task which requires careful evaluation, which should be carried out in a good manner; if the choice of method is wrong in the beginning, the errors usually tend to accumulate instead of being compensated for and corrected during the research process. The methods and techniques that are the most suitable depend on the purpose of the study and on the research problem. Mostly they seem to be combinations of each other; consequently, no method could be seen to be entirely qualitative or quantitative, even qualitative research may include elements from quantitative methods. The focus is on the information systems: the planning, building, testing and evaluation of them. They are nodes in a large network connected to other information systems and organisational processes. Information systems are not valuable if they lack a special purpose, their value to an organisation is based on their potential to improve its effectiveness and efficiency (Hevner, et al., 2004).

The capabilities of the information system and the characteristics of the organisation, i.e. the combination of its work systems, its people, and its development and implementation methodologies, determine the extent to which that purpose is achieved (Silver, Markus, & Beath, 1995). It is important for researchers in the Information Systems (IS) discipline to “further knowledge that aids in the productive application of information technology to human organisations and their management and to develop and communicate knowledge concerning both the management of information technology and the use of information technology for managerial and organisational purposes” (Hevner & Chatterjee, 2010, p. 270). As argued by some researchers, the complementary but distinct paradigms of design science and behavioural science are involved when acquiring such knowledge (Hevner, et al., 2004).

The two paradigms mentioned here characterise much of the research in the Information Systems discipline. The behavioural-science paradigm seeks to develop and verify theories that explain or predict human or organisational behaviour. The design-science paradigm seeks to extend the boundaries of human and organisational capabilities by creating new and innovative artefacts. Both paradigms are foundations of the IS discipline, positioned as it is at the confluence of people, organisations, and technology. In the design-science paradigm, the knowledge and understanding of a problem and the finding of a solution are achieved in the building and the application of the designed artefact (Hevner, et al., 2004). The research carried out for this thesis leans more on design science than on behavioural science.

The most important research methods – at least from the point of view of the researcher, have been presented here. However, it is obvious that defining the best method or methods is a rather difficult question to answer. Later, as the research process progresses, the methodology may become slightly different, but the methods discussed will no doubt be the primary ones.

5 THE STUDY OF FLIGHT SAFETY DATA

5.1 Test data

The cases were extracted from the VASA database (see Chapter 2.4.4) – test material was chosen from all the flight safety reports from 1994 to 1996. This set consisted of 345 cases from 1994, 421 cases from 1995 and 474 cases from the year 1996; thus 1,240 incidents, which can be considered a sufficient number. A period of three years was estimated to create “a critical mass” from the validity and reliability point of view, because in this case it meant more than 1,000 reported events. The FCAA granted access to about 1,200 safety reports from the years 1994-1996 to be used as test material, extracted from one Microsoft Access database. The main principle is that all the cases are public. This is in accordance with the law and practices adopted by the governmental organisations in Finland.

However, several exceptions to this basic rule can be recognised. In this context two exceptions had to be considered when determining the cases that would be included in the test material. The first one, which is general practice among the authorities, was that cases that were still open would be excluded from the research. However, because the material was over ten years old, no more open cases were found. The other exception deals with cases in which a military aircraft was involved. Following general privacy protection practices, all the names and other information that make it possible to identify the persons involved in a case were removed. The data fields extracted from the VASA database are displayed in Table 11, the data fields included in the research are in the left column and the excluded ones are on the right.

Table 11. Data fields extracted from VASA database

Included in the research	Excluded from the research
Year	
Report number	
	Date
	Type of aircraft
	Registration marks
	Place
	Phase of flight
	Operator
	Operation type
	Place of departure
	Route identification
	Destination
	Status
	Authority
	ATC unit
	Case classification
	Risk level
Narrative description	

The following report fields were used for the test material: the year when the event happened, the sequential report number starting from one each year, and the narrative description of what happened, free form text written by the person reporting the event. As displayed in Table 11, the report fields on the date of the occurrence, the type of aircraft encoded in two- to four-digit form (SF34 as Saab-Fairchild 340, for example), the registration marks of the aircraft, the place or the vicinity where the occurrence took place, the phase of flight (taxiing, take-off, cruising, etc.) during which the event took place, and the operator (airline or other organisation) of the aircraft were excluded.

Also excluded were the operation type (air traffic, training, private, etc.), the place of departure (if the incident happened during a flight), route identification, destination (or planned destination), status (pilot in command or co-pilot of the aircraft, air traffic controller, mechanic, etc.) of the person who wrote the report, the authority which investigated the case (for example, Civil Aviation Authority, Accident Investigation Board), the air traffic control unit responsible for the flight, the case classification (aircraft/system/component related event, the event related to the operation of the aircraft) chosen from the input interface menu by the officer putting the case into the database, the risk level of the case estimated by the officer saving the case based on ADREP87 (see Chapter 2.4.3 more in detail). The ADREP87 contains a risk classification system as follows:

- AA Flight accident
- (Aircraft damage)
- A Serious incident
- B Incident
- C Deviation from normal procedures
- D No identifiable risk
- E No deviation from normal procedures
- E Other reports.

Choosing the data to be used for the test material belongs to the pre-processing phase (see Figure 16, p. 80) reported in detail in Publication #1 “Data Mining in Aviation Safety Data Analysis”. The excluded report fields, as well as rest of the data from the database that was not extracted for use in this research process, can be used in future research (see Chapter 6.5).

5.2 Characteristics of the evaluated tools

Every special purpose poses special requirements for the tools that will be used in the research area. These characteristics are the basis for the choice process. The starting point for the selection process of a data mining tool for analysing flight incident data is that it should serve as a decision support system. The relevant information gained over time is very valuable for all organisations. Experience and insight is generally as useful for a small non-profit organisation as it is for a global corporation, only the amount of information differs significantly. The knowledge is, or at least it should be, an integral part of and the basis for planning and decision making in the organisation. Often, almost always in fact, decisions are made with incomplete knowledge about the problem parameters. That is why decision support systems are needed. As mentioned several times before, the organisations possess information that they cannot make use of for several reasons. One significant part of organisational knowledge may be tacit and almost impossible to put into explicit form. One important and often reoccurring question about searching for information and how to allocate it is what to search for to find useful results. This is often the case in organisations that collect data systematically for certain purposes.

The evaluation began by determining the number of tools that would be taken into the process. Three cases seemed to be a suitable amount. The author was aware of two prototypes, which seemed to be very interesting for testing. One of the aims of this research is to find new methods and tools for analysis and that is why two prototypes were included in the process. One commercial product was evaluated in order to better compare the systems. In the aviation industry, one of them was written about and evaluated more than the others. Thus it was chosen to be one of the evaluated systems. The demo was easily available from the Internet.

However, some changes were made during the process. Another commercial tool with a larger independence of languages appeared to be worth evaluating and the first chosen commercial product was confirmed as not being able to process text written in Finnish. However, it was still used in the study as a reference system, because material and user experience of it are widely available. The planned test for it was cancelled when it became apparent that Finnish text could not be tested. A second commercial product was tested instead as it was not totally language dependent. This system had never been tested with Finnish, but the representative of the company presented encouraging results with Spanish, so it was assumed that using it for Finnish could be possible and that it would be a very interesting case from this point of view. Thus, the original goal to test the material with three different systems was achieved.

The most important features of the tools are shown in Table 12. In this table only the important points of the flight safety tools are considered. They are collected from the operators and the written descriptions of the features of the different tools. The missing corresponding values have been sought and the table completed. Some of them are important criteria for selecting a mining tool for operational use, i.e. for an airline safety department analysing safety related data (Muir, 2004).

Table 12: Characteristics of the evaluated data mining systems

Characteristics	Evaluated systems			
	PolyAnalyst (#1)	TEMIS (#2)	GILTA (#3)	PolyVista (#4)
Commercial product/ prototype	Commercial product	Commercial product	Prototype	Commercial product
System based on	Several tools in-built	Different algorithms	Vector model	Several technologies
Appropriate analysis modules/capabilities	Complete capabilities	Complete capabilities	Vector model	Complete capabilities
Clustering available	Yes	Yes	With SOM[17]	Yes
Languages supported	English	English, Finnish	Any	English, Spanish tested
Visualisation and graphic tools capabilities	Yes	Yes	Through SOM	Yes
Amount of information needed for valid results	Small	Small	Small	Small
User experiences widely available	Yes (no Finnish)	Yes (no Finnish)	No (prototype)	Yes (no Finnish)
Form of applicable data	Database, sequential files	Database, sequential files	Sequential files	Database, ASCII files
Database products supported	Oracle, SQL Server	Oracle, Access, SQL Server	No	MS Access, SQL Server
Platform supported	Windows	Windows	Linux	Windows
Text and structured data mining capabilities	Yes	Yes	Yes	Yes
Accurate mining results	Yes	Yes	Yes	Yes
Reduced need for manual review of results	Moderate	Yes	n/a	Yes
Ease-of-use (interface and outputs)	Moderate; requires skills	Moderate; requires skills	Not in commercial use	Moderate; requires skills
Customisation needed	Little	Little	None	Little
System training available	Yes (not in Finland)	Yes	No commercial	Yes (not in Finland)
Purchase costs	Very high	High	Prototype	High, hiring possible
Operating costs	High	Low	Prototype	Low
Requires pre-processing	Much	Moderately	Little	Moderately
Thesaurus available	Yes	Yes	No	Yes
IT resources requirement	High	Moderate	Low	Moderate
Ease of acceptance and implementation by IT dept.	n/a	Low	Low	n/a

[17] See Chapter 1.3

5.3 Studied tools

5.3.1 Tool #1

Tool #1 is a commercial data mining product called PolyAnalyst, marketed by Megaputer Intelligence. In this context the applicability of version 4.6. was observed. It is a modularly built system that is used by companies throughout the world, a couple of them in aviation. The customers include Fortune 100 companies, numerous smaller companies, military and government offices, and educational institutions in many countries of the world (Megaputer Intelligence Inc, 2004b). PolyAnalyst is a comprehensive and versatile suite of advanced data mining and visualization tools. It features a broad selection of machine learning algorithms for the analysis of structured and unstructured data, including text analysis, link analysis, decision tree analysis, and visualisation techniques and is especially well-suited for safety data analysis (Megaputer Intelligence Inc, 2004c). Its capabilities range from the importing, cleansing and manipulation of data, to visualisation, modelling, scoring and reporting.

PolyAnalyst supports proprietary data formats like Excel and SAS, as well as popular document formats and it can access data stored in major commercial databases through standard OLE DB (Object Linking and Embedding for Database) or ODBC (Open Database Connectivity) protocols. PolyAnalyst contains a wide selection of semantic text analysis tools, clustering, classification and prediction algorithms, transaction analysis, link analysis, and visualisation capabilities. Results obtained with PolyAnalyst can provide key insights into different aviation processes, helping safety officers to (Megaputer Intelligence Inc, 2004a):

a) reveal hidden issues irrespective of data type – structured or unstructured
b) generate strategic overview charts for management across different parameters
c) identify bottlenecks in processes and aircraft part quality or part supplier related issues.

PolyAnalyst provides a set of tools that can be tailored to a specific application domain (Megaputer Intelligence Inc, 2004a). Principally PolyAnalyst could have provided a wide variety of mining tools and reporting features for this study, but its text mining algorithms were not language independent. They were built to be used with specific languages and Finnish was not included. That is why tests with it were not performed. The latest version of PolyAnalyst is 6 (Megaputer Intelligence Inc, 2015a).

5.3.2 Tool #2

Tool #2 is also of American origin as Xerox laboratories began its development. Nowadays it is developed and represented by a French company called Text Mining Solutions, TEMIS. The company was founded in 2000. The connection with Xerox began with the signing of an initial licensing agreement for the XeLDA® linguistic engine. This agreement later led TEMIS management to acquire Xerox linguistic operations (TEMIS, 2005). The agent of the system in Finland is a company named Lingsoft Oy, located in Turku. In March 2005 TEMIS and Lingsoft signed a Bilateral Collaboration Agreement. Through this agreement TEMIS expanded its linguistic coverage to support 16 languages. The TEMIS products then integrated the Lingsoft set of Northern European languages into it. Lingsoft provided TEMIS with Finnish, Swedish, Norwegian Bokmal, and Danish Natural Language Processing software. The TEMIS core technology XeLDA® then had these multilingual software packages integrated into it. It is a multilingual linguistic engine for modelling and standardising unstructured documents in order to automatically exploit their content. Alongside this agreement, Lingsoft became a VAR (Value-Added Reseller) for TEMIS products for the Nordic region (Lingsoft Oy, 2005). The product arrangement can be divided into two lines: The first line is the collection of core products that are XeLDA®, Insight Discoverer™ Extractor, Insight Discoverer™ Clusterer, and Insight Discoverer™ Organizer. The other line consists of the application solutions that are eXtraction Terminology Suite and Online Miner™ (TEMIS, 2005). In addition to the system, there are also other components that have been built to support the usability of the main parts.

From the application solutions the eXtraction Terminology Suite allows the extraction of noun phrases using a morpho-syntactic analysis of sentences. Online Miner™ is a complex but well-functioning data mining tool providing many unique features. The Online Miner™ unites the core products that are mentioned before, as well as other supporting parts. The central unit of Online Miner™ is the Insight Discoverer Suite that consists of the core products except XeLDA®. The engines are not powerful enough to produce understandable mining results. That is why other feasibilities are needed, which are provided by TEMIS Visual Components, which is a library of different visualisation tools. Their function is to produce the results from textual and text mining in visualised form. This process can happen with topic maps, heat maps, pie charts, histograms, diagrams and other illustration methods. The library was developed in Java, which makes its integration with third party applications possible. The source connectors allow data access from remote sources, either internal to the company, such as databases, document management systems, file systems, and so on; or external, as commercial databases, patent databases, news flows, etc.

Advanced search capabilities are needed to input data from different sources. A significant feature of this system is the possibility to search for data from document management systems.

All these characteristics can usefully contribute to the information technology of the Flight Safety Authority (now TraFi, see Chapter 1.4). Databases have been in production use for decades already. Visualisation is an important feature in converting the search and research results into understandable information. The main task in this context would be to find hidden trends in flight safety data. This process can be divided into two phases. The first phase is to identify the existence of trends, which is easy to execute using graphical tools. This also allows inexperienced users to mine the cases and notice that some trends exist. In the second phase, an experienced inspector should very carefully examine the results produced by the system to confirm that the proposed trends actually exist.

Insight Discoverer™ Extractor is especially constructed to extract information from text documents. It detects the most relevant pieces of information for the user. This happens by performing a sequence of three linguistic analysis steps (TEMIS, 2005):

- Corpus recognition: automatic language identification
- Morpho-syntactic analysis:
    - assigns a grammatical category to each word in a document (noun, adjective verb, etc.) as well as its morpho-syntactic characteristics (gender, number)
    - lemmatisation: returns each word to its base form (singular for a plural, infinitive for a conjugated verb) so that it can be recognised independently of its inflected form
- Knowledge extraction (runs extraction rules):
    - recognition of entities (name of companies, associations, organisations, products figures, dates, places etc.)
    - identification of relationships between the entities (company-company, person-company, company-product).

Although several data mining and text mining products are developed for the same purposes, all of them are constructed on different kinds of algorithms, or, the combinations and connections of the used algorithms in the systems are different. The core engine of the knowledge extraction is Skill Cartridges™. A Skill Cartridge™ is a hierarchy of knowledge components that describe the information to be extracted for a given business, specific field or topic. A knowledge component is a lexicon and/or an extraction rule. An extraction rule describes a sentence structure that characterises a concept (TEMIS, 2005). The Skill Cartridges™ are very flexible and this feature allows the management of language complexity. One of the most important features of data mining systems is that they produce accurate and valid mining results. To optimise extraction and to achieve the goal mentioned before, Skill Cartridges™ can be fully customised by fine-tuning options such as (TEMIS, 2005):

- extraction of negative or positive trends
- differentiation between rumours and actions
- enhancement of anaphora identification
- resolution of acronyms
- integration of ontologies.

The Insight Discoverer™ Clusterer groups similar documents dynamically. It is an automated server that classifies a collection of unstructured documents in the most relevant clusters based on their semantic similarity. This happens by processing that combines both linguistic and statistical analyses to find out their semantic proximity. This process consists of two phases: First, the Insight Discoverer™ Extractor generates (with a morpho-syntactic analysis) the semantic profile of the documents and thus creates the descriptions of them. The Insight Discoverer™ Extractor can be customised for this special task with the Skill Cartridges™. The second phase uses this information as inputs when the Insight Discoverer™ Clusterer reads the generated descriptions and then groups them. This process does not require any categories to be defined before the process. The principle of the process is rather simple: The clustering function only groups the documents by their semantic proximity.

This feature is a key function for different kinds of users, especially those working for the Flight Safety Authority. Information and knowledge miners, who do not exactly know what they are searching for, can be counted in this category. They represent commercial airlines, car manufacturers, insurance companies, traffic safety authorities and countless other organisations, including profit making organisations and non-profit organisations that have a large collection of incidents, etc. The users and/or the managers do not know but would like to know whether their data collection contains trends or similar cases that form clusters that would merit a more detailed analysis. This does not exclude the need to find similarities to a selected case, a feature that is essential, at least in the phase when a more detailed analysis should be carried out. The clustering function is basic if an unsorted collection of cases needs to be analysed. This also uncovers the non-existence of trends if the material is totally fragmented and thus contains no discoverable clusters.

After the classifying process the user can organise the documents in categories and sub-categories, which gives an overview of the information as well as different ways to explore it. The grouping algorithm for similar documents is especially adapted to text analysis. The flexibility of the tool allows the user to configure the depth of the classification model and determine the number of classes per level. The system assigns a heading to each class. This heading uses the terms and expressions that are most characteristic of the class in hierarchical order. The flexibility of the Insight Discoverer™ Clusterer appears in its ability to satisfy different requirement levels (TEMIS, 2005):

- viewing: a mapping module enables users to view classes in the form of spheres whose sizes vary according to the quantity of documents they represent. It also identifies the links between classes in the form of segments of varying widths
- analysis: by projecting descriptive variables onto the mapping result, the user can instantly identify the spheres for action according to the colour codes
- customisation: using a Skill Cartridge™ enables the user to integrate business vocabulary or make rules to homogenise the proposed classification.

In the analysis of flight safety incidents, the last mentioned feature, i.e. customisation, has a special meaning. The narrative part might contain many abbreviations, especially reports written in English, which are easily noticed when reading them. This cannot be regarded as a problem resulting in increased work to teach the system to understand all the synonyms in the reports written in Finnish, although some abbreviations and acronyms are commonly used in Finnish as well in the reports. There is a clear explanation: the language of aviation is English and all international aviation communication uses it. Aviators are used to English expressions all over the world. English can be used in communication with Air Traffic Control units throughout the world, although the official language of the country is normally used. It is full of abbreviations and acronyms to make the expressions short and clear and suitable for radio traffic. Written aviation documents, both official and unofficial contain them in great numbers because it is faster to use abbreviations and acronyms when writing reports.

Insight Discoverer™ Categorizer (Organizer) is a server that categorises unstructured documents. It uses both statistical and linguistic analysis rules when it sorts the documents into categories that are defined in advance. This tool exploits the results from the Insight Discoverer™ Extractor in three phases, which produces the semantic profiles of the documents. The phases are as follows, although the third one cannot be considered a processing phase like the first two (TEMIS, 2005):

- the learning phase
- the categorisation phase
- evaluation of automatic categorisation.

As mentioned before, the documents can be classified in advance according to the needs and practices of the organisation. After this, the morpho-syntactic processing made by the Insight Discoverer™ Extractor produces a semantic descriptor that consists of the frequency of nouns, verbs, noun phrases etc., that is given to them. On the basis of these documents the learning process begins, which enables the Insight Discoverer™ Categorizer to build the categorisation model. For this task an algorithm combining the various semantic descriptors assigned to the same category is used. To validate the model, the categories should contain between 25 and 50 documents per category although the minimum number depends on the number of documents.

The system has advanced search capabilities that enable users to search by concept. The solutions are developed in J2EE and have standard API and they read and produce XML files. The Online Miner™ Front End allows the user to cluster the data with the Insight Discovered™ Clusterer. With this tool it is possible to classify and regroup documents into coherent classes. This function is based on the semantic similarities of the documents. An overview of the documents is accessible for the user. This can be done with the titles, publication dates, and topics of the documents. A graphical representation of the document distribution can be produced with this tool as well. The search function is based on different search engines and a search can be made within a document and its metadata. The statistical analysis is carried out on several levels, and the search results can be sorted by relevancy; this function is possible to use in order to filter results. The results are produced in two alternative ways. The documents can be distributed in a hierarchy of extracted concepts or the concepts can be displayed with links to documents; by selecting sub-concepts the analysis can be refined. This function also offers the possibility to perform an analysis on document zones, for example, titles and abstracts.

The Online Miner™ Back End allows the user to mine from a wide variety of sources. They can be websites, local directories, mailboxes, TMX files, and databases. After the Online Miner™ Back End has examined the data in the given source the user decides what to do with the data. The data can be extracted with Skill Cartridges™, which is a knowledge component for extracting key information. The Insight Discoverer™ Categorizer is a document categorisation server that can sort the data into categories. It combines statistical and linguistic analysis rules and automatically classifies unstructured documents into predefined categories. It is also possible to populate the database with the information extracted by the system. After data are extracted they are transformed into a TMX (Temis Metadata in XML), which is an XML format providing document metadata structure and storage. In this format, information is structured flexibly before it is stored in the database.

The results of the only mining round performed with this tool are presented and explained in Publications #1 "Data Mining in Aviation Safety Data Analysis" and #2 "Data Refining for Text Mining Process in Aviation Safety Data", but a brief presentation on a general level is included here. TEMIS produced a lot of statistical data about the mining results. The most important of them in this context are presented in Table 13.

Table 13: Results illustration in TEMIS

Similarity (per cent) of the five closest clusters on average	Average of the maximum degrees of explanation, the average of the degrees of explanation, and the average of the minimum degrees of explanation	Correspondent standard deviations	Average degrees of explanation of cluster quarters	Corresponding standard deviations
3.4074	17.7037	3.312523	14.2878	2.6734
2.5556	9.1639	1.714604	10.3927	1.9445
2.3704	3.1852	0.595993	7.9751	1.4922
2.1481			5.2742	0.9868
2.0741				

From each cluster, the five closest clusters and their percentage of similarity were presented. This value varied between five and one per cent (all the percentages are reported as integers), having been mostly two or three. In the first column, the average percentages of correspondence of the five closest clusters are presented. As the values are low, this means that the clusters are not similar at all. This proves that the clustering of the reports has been successful and relevant, thus this method is usable for this data.

The second column of Table 13 displays the average degrees of explanation, which means how appropriately the reports described the contents of the cluster defined by the system. The average degree of explanation was 9.1639 per cent with the corresponding standard deviation 1.714604 (displayed in the third column). The value is considerably low but it can still be regarded relevant. The average of the minimum degree of explanation was 3.1852 per cent with the corresponding standard deviation 0.595993. The average maximum differs slightly from the two others being more heterogeneous and has a standard deviation of 3.312523. Because the maximum values are relatively high, despite their strong dispersion, they prove that the clusters are composed of relevant reports and that most of the reports on their own might well indicate a trend that should be examined more thoroughly.

The fourth column displays the average degrees of explanation of the cluster quarters, the first quarter on the first line, the second on the second and so on. The fourth column presents the average degrees of explanation of the cluster quarters, the first quarter on the first line and so on. Their corresponding standard deviations can be found in the last column. When compared with the averages in the second column, it can be noticed that the values of the two first quarters are higher (14.2878 and 10.3927) than the average 9.1639. Furthermore, the value of the third quarter 7.9751 is not meaningfully lower than the average and the value of the ‘lowest’ quarter is 5.2742, which is to some extent higher than the average of the minimum 3.1852. The division into cluster quarters displays the degrees of explanation in more detail than the averages presented in the second column.

The variation of the degrees of explanation is presented in graph form in Figure 22. The vertical axis shows the numerical values of the degree of explanation of the clusters and on the vertical axis the numbers of the result clusters (1.1 to 26) are presented. As seen from the figure, the average degree of explanation does not deviate much from ten per cent as explained in Table 13. If the trend lines of the diagrams were displayed, each one would be practically horizontal. From this it can be concluded that the degree of explanation has remained about the same through all the clusters, independent of their size.

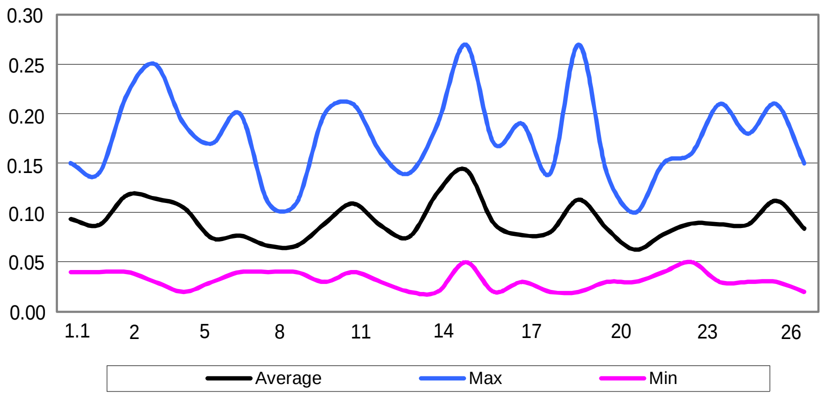


Figure 22: Degrees of explanation of the clusters in TEMIS

A preliminary test with Tool #2 was run before the main test with real aviation safety incident data. This was seen as a necessary measure in order to find out whether mining a short report produces a relevant result. The test began by choosing a suitable sample of data. Usable, open data was found from the site of NASA Aviation Safety Reporting System (ASRS). It contains twenty-eight (28) ASRS Database Report Sets on topics of interest to the aviation community in pdf-format. Each Report Set consists of fifty ASRS Database records, preceded by a note of introduction and caveats on the use of the ASRS data. All Report Sets have been pre-screened to ensure their relevance to the pre-selected topic description. The reports sets were updated quarterly (NASA, 2004). In this case the dataset Recreational Aviation Incidents was chosen. It is a sampling of reports from recreational aviation, including balloons and ultralight aircraft. This was chosen, because all the others are classed into similar cases, so, clustering would be more obvious to find in this set because it contained different types of incidents.

In the first phase the material was simplified to contain only text in two fields: Narrative and Synopsis. All the other information concerning date, time, environment, number of aircraft, assessments, etc. were left out but were added to the process later. The first process was clustering with text mining functions. The system found eight clusters, the first two of them containing five cases, the next two containing three cases, the next two cases and the rest of them containing one single case. Three cases were dropped out by the system; the obvious reason was that the system did not find any components in these three cases to create a cluster.

5.3.3 Tool #3

Tool #3 was developed at Tampere Technical University, the Signal Processing Laboratory as one of the results of the project called GILTA (Managing Large Text Masses) between 1999 and 2003. The Gilta Project was at that time a part of the Tekes USIX technology program. The main objective of the Gilta research project was to develop auxiliary means for finding and administering information in a group of text documents. The scientific objective was to find out if it is possible to find the hoped for information in a mass of documents by using self-organising maps (SOM) or similar methods combined with linguistic sentence analysis.

The information retrieval in text documents has usually been based on automata theory, grammar, language theories, fuzzy logic, natural language processing, or latent semantic analysis (Gey, 2000; Oard & Marchionini, 1996). The usage of keywords is a common approach to topic detection and tracking, based on the assumption that the text is accurately characterised by the keywords the authors use. However, in this case accuracy is neglected, therefore a more accurate method would be to use of all the words of the document combined with the frequency distribution of the words, although it is very complicated to compare the frequency distributions (Manning & Schütze, 1999). In this project, the approach utilises the idea but in a peculiar way as the processing methodology is based on the word, sentence, and paragraph levels. The first phase is the preprocessing of the original text, which means that extra spaces and carriage returns are omitted, etc. Next, for encoding purposes, the filtered text is translated into a suitable form. There are several approaches to this (Toivonen, Visa, Vesanen, Back, & Vanharanta, 2001):

1. The word is identified and a code replaces it. If sensitivity to new words is what the researchers wish to discover, this approach is applicable.
2. The words succeeding the first word will also be replaced with a code – this approach is language sensitive.

3. The third approach analyses all the words character by character. A key entry to a code table is calculated based on the characters. If the code table is not designed in a special way, this method is capital letter and conjugation sensitive.

As the last alternative is accurate, and thus suitable for statistical analysis, it was chosen for the Gilta Project. (Toivonen, et al., 2001). For example, the letter *w* is transformed into a number in the following manner

$$y = \sum_{i=0}^{L-1} k^i * c_{L-i}$$

where L is the length of the character string (the Word), c_i is the ASCII value of a character within a word w, and k is a constant (Kloptchenko et al., 2002). Example: the word is "**c a t**".

$$y = k^2 * ascii(c) + k * ascii(a) + ascii(t)$$

A different number is calculated for each different word and the same number can be given only to the same word. Each word is converted to a code number, and after that minimum and maximum values are set onto words, after which the distribution of a word's code numbers is determined. The next step is to estimate the distribution of the code numbers. In this context, the Weibull distribution was selected to represent it, but others like the Gamma distribution might also have been possible (Toivonen, et al., 2001). The process continues as follows: the range between the minimum and the maximum values of the words' code numbers is divided into N_w logarithmically equal bins, after which the count of words belonging to each bin is calculated. Then the counts will be divided with the number of all words, which is followed by determining the best Weibull distribution. As the best Weibull distribution is found in the testing phase, it is divided into N_w equal size bins, the size of each one is 1/ N_w. After the procedures presented here, every word can be found to belong to a bin, which is found by using the best fitting Weibull distribution and the code number. This quantisation method allows the word to be presented as the number of the bin it belongs to. The selected coding method allows the resolution to be the best in cases where the words are the most common in the text, which usually means two to five letter length words. The so-called rare words (usually long words) are not that minutely separated from each other (Kloptchenko, Back, Visa, Toivonen, & Vanharanta, 2002).

What has been presented here concerning the word level is also applicable to the sentence level, as every sentence must similarly be converted to a number.

The process begins by changing every word to a bin number so that the whole encoded sentence is represented as a sample signal, which will next be Fournier transformed. The lengths of the sentence vectors differ because the sentences in the text do not contain the same number of words. The sentence vectors are transformed using the Discrete Fourier Transform (DFT) without considering all the coefficients. From the sentence data, a cumulative distribution is created that follows the same model as the word level. This takes place after every sentence has been converted to numbers. As with the word level, in this phase the range between the minimum and the maximum value of the sentence code numbers are also divided into N_s equal size bins. Again in a similar way as on the word level, the count of sentences that belong to each bin will be calculated and the counts of the bins are divided with the number of all sentences. Using the cumulative distribution of both distributions, the best Weibull distribution corresponding to the sentence data can be found. In the quantisation of sentences, the best distribution can be used after this (Back, Kloptchenko, Toivonen, Vanharanta, & Visa, 2002)

The procedures are similar on the paragraph level, beginning with the conversion of the paragraphs to vectors using the code numbers of the sentences. The coefficient B_1 is chosen to represent the paragraph as the vectors are Fourier transformed. Again, after finding the best Weibull distribution that corresponds to the paragraph data, the quantisation of paragraphs may be carried out. For text documents, the word, sentence, and paragraph code numbers of the document create histograms determined by the corresponding value of the quantisation. The encoding of the filtered text from a single document is done word by word on the word level. The quantisation with all the words in the database is carried out and each word code number is quantised using it. The word histogram A_w consisting of N_w bins, is created after the correct quantisation value is determined and a sum factor that corresponds to the value is increased. Finally, the histogram is normalised using the word count of the document. The process creating histograms is similar on the sentence and paragraph levels. First, the single document will be encoded to sentence and paragraph code numbers, after which the hits according to the corresponding place in the quantisation will be collected into the histograms A_s and A_p. With the histograms from all the documents in the database it is possible to compare and analyse the text on the word, sentence, and paragraph levels of each document. It is worth noting that doing all this does not require knowing anything concerning the actual text document, but giving one document as a prototype is enough. By using this methodology, the user will discover all the documents which are similar, or be given a number to their difference, or discover the clusters of similar documents. The systems training process allows for the easy adaptation of the methodology to any specific application field (Toivonen, et al., 2001).

A couple of preliminary tests with a dataset of 300 technical failure reports were run before the actual mining process with the pre-processed safety report material. Five reports from the data were randomly chosen for use as test documents and the first one of them was compared with the 300 actual reports by applying Euclidian distance in order to find out whether the dataset contained similar reports or not. It is worth noticing that the reports were processed as such, i.e. no pre-processing was performed and that is why the results must be seen as being somewhat rough. The results are displayed in Figure 23.

TEST DOCUMENT	0	|	|
0,1049	43	|	|
0,1506	215	0,7568	25
0,1655	254	0,7598	84
0,1668	261	0,7627	279
0,1909	249	0,7682	200
0,2021	42	0,7777	277
0,2021	69	0,7877	227
0,2021	251	0,8008	189
0,2063	141	0,8074	182
0,2082	158	0,8074	295
|	|	0,8598	98
|	|	0,8837	77

Figure 23: Euclidian distances of the chosen test document and the others

In the figure, the distances between the test document and the ten nearest and the ten farthest are displayed, the values of the others interposing between them. The results must be read as follows: the distance between the reports is the Euclidian distance (without the square root), i.e. the value zero when the reports are identical and the value two when the compared reports contain no common words. There are two columns in the figure, the first of which is the distance and the second the serial number of the report, given by the operator. As obvious, the nearest report #43 does not differ much from the test report, having the distance value 0.1049, because the same report was the first test document. The reason for the minimal difference, indicated by the distance value 0.1049, was due to the removal of the main classification of the cases given by the inspector who saved the case in the database. The main classifications, Fuel feed and adjustment and Packing for report #43, were added to the reports in order to find the cases among the same problem area easily using simple database queries. The amount of common words with the farthest report, #77, is somewhat less than 50 per cent with a corresponding distance of 0.8837. This is not surprising, because the technical failure reports are relatively similar, containing common words and expressions typical of the language used in aviation. Additionally, the main classifications added to the reports increase the amount of common words.

The preliminary test continued by removing all the punctuation marks and converting all the text into lower case letters. After this slightly pre-processed data, both the Euclidian and Jaccard distances of all five test documents and comparison material of 300 reports were determined. There would also have been the possibility to cross-check all the 300 reports with each other, creating a matrix of 300 x 300, but it seemed too complex from the point of view of the preliminary test arrangement and thus impractical in this case. Displaying all the results was not appropriate due to their large extent. As the comparison of the first test report, #43, with the others functionally represents all the results mentioned, it is presented in Table 14.

Table 14: Results of the preliminary test in GILTA

Difference of position	**0**	**1**	**2**	**3**	**4**	**5-10**	**11-30**	**Total**
Euclidian no pre-processing	-	-	-	-	-	-	-	-
Euclidian pre-processed	3	5	2	6	3	13	33	**65**
Percentage	1.00	1.67	0.67	2.00	1.00	4.33	11.00	**21.67**
Jaccard pre-processed	8	6	13	7	5	35	78	**152**
Percentage	2.67	2.00	4.33	2.33	1.67	11.67	26.00	**50.67**

The table must be interpreted as follows: The second row after the title row displays the difference in the position of each of the 300 reports – when they are sorted according to their similarity compared with the test document. The first comparison is the Euclidian distance between not pre-processed and pre-processed data. The results show that three of the reports have exactly the same position: five reports have the difference of one step, two of them have two, and so on. The differences beyond 30 steps have been ignored. The corresponding percentages are presented on the next row. This means, for instance, that those three having the same position represent one per cent of 300 reports. In total, in comparison with the not pre-processed reports, the Euclidian distances of 65 pre-processed reports, corresponding to 21.67 per cent of the data, are, at most, 30 steps away. The highest value of 30 was chosen by the operator of the mining tool for practical reasons.

Using the same scale, when the different measurements of the distances are evaluated, the results are somewhat different. They show that with respect to the order of the reports, they are closer to each other when compared with the previous situation. The percentages of the distance classes are significantly higher and the differences of the positions of about half of the reports are 30 at most, as only 21.67 per cent of the reports could be placed in the same range. This test proved that the pre-processing had more significance from the point of view of the mining results than the applied distance measurement method, either Euclidian or

Jaccard. The purpose of this preliminary test was observing the effect of slight pre-processing and the application of different distance measurement methods. Therefore, the results were not controlled by more minutely studying the contents of the individual reports, except for the nearest and farthest reports. This was naturally done when the main process was performed, i.e. mining the flight safety reports with GILTA by applying clustering, which is based on Euclidian distances.

In order to illustrate the comparison, the highest and lowest numerical values are presented in Figure 24. The first two columns display the not pre-processed Euclidian distances between the test document and the corresponding report, which are the same as presented in Figure 23, which displays, however, only the five nearest and five farthest reports. When reading the values presented in the figure, it is worth noticing that the scales used differ from each other. First, like in the situation presented in Figure 23, the Euclidian distances in the not pre-processed data would have the value zero and report no common words with the value of two. After the pre-processing, the range was adjusted to be between zero and √2, in order to obtain more accurate results. The corresponding Jaccard range is between one, identical reports, and zero – when the compared reports have no common words.

<table>
<tr><th colspan="2">Euclidian non pre-processed</th><th colspan="2">Euclidian pre-processed</th><th colspan="2">Jaccard pre-processed</th></tr>
<tr><td>0,1049</td><td>43</td><td>0,2445</td><td>43</td><td>0,9455</td><td>43</td></tr>
<tr><td>0,1506</td><td>215</td><td>1,0853</td><td>61</td><td>0,1750</td><td>61</td></tr>
<tr><td>0,1655</td><td>254</td><td>1,1266</td><td>28</td><td>0,1600</td><td>174</td></tr>
<tr><td>0,1668</td><td>261</td><td>1,1301</td><td>6</td><td>0,1549</td><td>6</td></tr>
<tr><td>0,1909</td><td>249</td><td>1,1355</td><td>77</td><td>0,1549</td><td>17</td></tr>
<tr><td>|</td><td>|</td><td>|</td><td>|</td><td>|</td><td>|</td></tr>
<tr><td>|</td><td>|</td><td>|</td><td>|</td><td>|</td><td>|</td></tr>
<tr><td>0,8008</td><td>189</td><td>1,3836</td><td>245</td><td>0,0133</td><td>108</td></tr>
<tr><td>0,8074</td><td>182</td><td>1,3855</td><td>189</td><td>0,0130</td><td>107</td></tr>
<tr><td>0,8074</td><td>295</td><td>1,3984</td><td>108</td><td>0,0128</td><td>243</td></tr>
<tr><td>0,8598</td><td>98</td><td>1,4142</td><td>95</td><td>0,0000</td><td>95</td></tr>
<tr><td>0,8837</td><td>77</td><td>1,4142</td><td>133</td><td>0,0000</td><td>133</td></tr>
</table>

Figure 24: Numerical values of the differences

The results are clearly observable in Figure 24. All the different test arrangements placed report #43 closest to the test document, regardless of the applied method. With the pre-processed data both measurement methods also placed report #61 the second closest and #6 the fourth closest to the test document. Both methods also placed reports #95 and #133 the farthest away, on the Euclidian scale this was 1.4142, the practical approximation of √2 and 0.0000 on the Jaccard scale. The position of report #108 is also worth noticing. This preliminary test reveals the fact that the mining results depend on many factors and the results need to be carefully examined and interpreted. When the farthest reports and their corresponding distance values are studied, then, according to the definition, reports #95 and #133 have no common words with the test document. The Euclidian distance with not pre-processed data determines that report #77 is the farthest away with the value of 0.8837. If the report had no common words with the test document, the corresponding distance would be 2.0000 as explained in Figure 23. Removing the punctuation marks and changing all the upper case letters to the lower case does not change the words themselves, thus it is not reasonable to draw any further conclusions on the basis of these test results.

The proper mining process with 1,240 flight safety report began after the preprocessing of the whole dataset (see Publication #1 “Data Mining in Aviation Safety Data Analysis”), on the basis of which the lists of stop words and synonyms were created. The reports were divided into 100 clusters on the basis of the nine most significant words. The mining results of the first round are explained on a general level in Publication #1 “Data Mining in Aviation Safety Data Analysis”, but some specific observations are worth presenting here. GILTA describes the mining results presenting in addition to the number of reports that include the absolute amount of the significant words in a cluster. In order to make the results comparable with PolyVista (see Chapter 5.3.4), the most significant word is given the index 100 and the other words are given corresponding values. As a rule of thumb, it can be mentioned that the higher the indexes of the most significant words, the better the clustering results. However, this does not tell the whole truth since the results must be carefully interpreted by a specialist to find out whether the words belong together to form a sensible entirety. The most relevant clusters of the first round using GILTA are displayed in Figure 25.

door	100	opened	100	close	67	draw	33	alarm light	33	vaasa	33	out	33	push	33	noticed	33
notice of arrival	100	submit	88	miss	38	flight plan	25	air traffic control	25	efhk	25	route	13	hold	13	terminated	13
tow	100	glider	83	rope	50	altitude	50	damage	33	come	33	toinen	33	landing	33	began	33
finland	100	without permission	75	aircraft	69	air space	63	cross	56	flew	50	swedish	38	back	34	return	34
interruption	100	take-off run	75	route	50	reason	38	co-pilot	38	speed indicator	38	speed	38	md-	38	aircraft	38
sector operations center	100	restricted area	71	report	71	r	57	pilot	57	wing	57	air force	57	satakunnan	43	flight	43
cigarette	100	rear toilet	71	passenger	43	found	29	stewardess	29	connection	14	warning	14	left	14	felt	14
glider	100	come	60	crash	60	another	40	case	40	wing	40	above	20	damage	20	affect	20

Figure 25: The most relevant clusters of the first round by GILTA

As seen from the figure, the weight index of the most significant word is 100, which best describes the content of the cluster. The index value for the following words is relatively calculated based on its incidence in that cluster. At least two or three words have a relatively high value, indicating the homogeneity of the cluster, i.e. the most significant words create a relevant group, although the relevance of the least significant words is not that important. In order to illustrate this, some clarifications concerning the figure are worth presenting here. Also, it is worth noticing that all the concepts displayed on the rows are translated from Finnish, and were originally one word. In the first cluster the question is about reports where an [aircraft] door opened during a flight and was closed by drawing it in. In some cases, an alarm light indicated the case, etc. The fourth cluster describes cases where aircraft have crossed the border and entered Finnish airspace without permission – accidentally. Some of them have been Swedish and turned back after noticing their incorrect position. On the fifth row in the figure, the cluster contains reports about interrupted take-off runs to a specific route, the reason being the co-pilot has missed the air speed indicator, for example.

However, the high weights of the significant word do not determine the relevance of the clusters alone: a skilled specialist is always required in order to control the clustering results. For example, two clusters without relevance are displayed in Figure 26.

sua	100	turn	100	dlh	100	oncoming	50	against	50	stockholm	50	tampere	50	route	50	level	50
left	100	right	100	after	86	situation	43	fuel	43	engine	43	grade	43	technics	29	close	29

Figure 26: Examples of two irrelevant clusters

The weights of the most significant words are very high, but when the reports of the clusters are carefully examined, no connections between the reports are found despite the frequency of the same words being rather high. This observation, demonstrates that quantitative mining results are not necessarily very applicable. The meaning of the words may also be totally different, although they are written in the same way. For instance, the word ‘wing’ appears in Figure 25 both on the last and the third last row. In the first case, it refers to the Finnish air force, in the second a part of an aircraft.

As only one mining round was performed with TEMIS (see Chapter 5.3.2) and the mining results of PolyVista (see Chapter 5.3.4) were delivered in the form containing the significant words of clusters with their weights, the first round mining results of GILTA provided valuable reference material as both the clusters and their contents were available. The obvious mistakes made in the preprocessing phase were quickly discovered on a cursory analysis of the results. For example, a common stop word ‘*jälkeen*’ (after) was left off the list. Another easily discoverable object requiring some adjusting was the word ‘*kone*’ (airplane in this context); its high frequency caused cluster #11 to contain 158 reports, making no sense in this context. The preparations for the second mining round of GILTA and PolyVista are described in detail in Publication #2 “Data Refining for Text Mining Process in Aviation Safety Data”.

5.3.4 Tool #4

Tool #4 is also a commercial product of American origin called PolyVista, built by a company called PolyVista, Inc., the headquarters of which is located in Houston, Texas. PolyVista was founded in 1995, and it develops solutions ranging broadly in scale and complexity from analysing rather small single data sets to major and on-going multi-departmental projects. Solutions can easily be tailored to meet the specific needs of organisations. The PolyVista Mining Tool is a state-of-the-art analytical solution that offers a broad range of services that are customised and scalable, fulfilling all analytical needs in order to extract more from the data. The software is built on advanced analytical techniques in order to dig deep into textual and numerical data – it is able to provide its users with several functions – the most important from the point of view of this study might be detecting patterns and anomalies and discovering hidden trends in the data, in addition to revealing previously unknown issues and uncovering root causes. The tool contains automated, pre-built pattern-recognition algorithms designed for facilitating collaborative analysis and sharing equipped interactive and dynamic visualisations, in addition to the intuitive interface that facilitates ad-hoc analysis (PolyVista Inc., 2013c). The PolyVista Mining Tool was originally built to work in an English language environment, but theoretically PolyVista TM is language independent; it should handle any language – as mentioned in Chapter 5.2, successful tests were performed with Spanish. Classification has been possible on Italian, French, German, and Spanish, but recently (Spring 2015) the system has switched to natural language processing (NLP, see Chapter 3.3.5) and the only language supported is English.

One significant customer reference for this tool is Southwest Airlines. Although extensive test projects with PolyAnalyst were conducted, PolyVista® Discovery business intelligence software was selected by the Flight Safety Department of Southwest Airlines in order to improve the proficiency of analysis in its Aviation Safety Action Program (ASAP). Together with an improvement of the safety functions, the implementation of PolyVista has considerably widened the awareness of the value of data warehousing and business intelligence (BI) applications across the whole company. This sort of cutting-edge technology provides ongoing value to Southwest Airlines, especially for supporting strategic decision making and enhancing employee efficiency and productivity (PolyVista Inc., 2010)

In order to process both structured and unstructured data including combinations of both of them, PolyVista features dozens of frequently updated algorithms for fully analysing an unlimited amount of data, and new ones are continuously developed. These are constructed with a view to enabling the delivery of hitherto unforeseen insights, unlike other tools that produce results only from known questions and queries. From this point of view, these results might act as an input for further analysis, to which other features, like an intuitive graphical interface, could be added to enhance performance. The interfaces of the functions have been built to provide full accessibility to release the end-user from learning complicated routines. Additionally, through the system's Cloud Solution, the clients of the system can have access to these algorithms from anywhere in the world. A wizard assists the user in configuring the algorithms by selecting the filters and parameters in order to display only the most relevant from the point of view of the aim of the mining process (PolyVista Inc., 2013a).

As mentioned before, the built-in algorithms are numerous, but in this context, a few of them are worth mentioning. From the point of view of this study, the *Discovery algorithm* can be seen as the most important one. This detects the hidden trends in the data uncovering interesting relationships and previously unknown findings (PolyVista Inc., 2013b). Discovery Solutions has been extended in order to release text-based resources to the daily operations of an organisation and make it possible to utilise untapped potential, providing robust analytics on both structured and unstructured data. With this algorithm, the important text mining functions of Process text data and Clustering are performed; the first mentioned covers cleaning and transformation, i.e. extracting, scrubbing and processing the raw text data. The second groups the text records based on term similarity. Another text-related algorithm is *PolyVista's Full-Text Search algorithm*, aimed at widening the end-users' text processing capabilities beyond the total word count and data samples that enable sorting through a limitless amount of data, the unstructured text is to be especially highlighted. Applying this fully automated search process algorithm, the specific words, terms, phrases, and even sentences in any type of unstructured data can be found. The advanced filtering options ensure that only relevant information is displayed in the results. The *Summary algorithm* quickly and efficiently provides a comprehensive yet general overview of the data content, being able to sort through thousands of dimensional combinations, displaying only the results defined as exceptional by the user. This algorithm is suitable for when the user needs to perform queries containing numerous specific criteria. The primary purpose of the *Difference algorithm* is to enable the comparison of two user defined data sets to measure trends and performance, etc. between the sets, and to scale the data automatically in order to eliminate skewed results. In addition, it can easily be tailored to any type of data. This algorithm can reveal emerging issues, indicating an opportunity or a problem previously unforeseen. The *Statistical Process Control (SPC) algorithm* has an efficient monitoring ability, using multiple dimensions to determine when data differs from a trend, resulting in it being often used as a tool for early warnings (PolyVista Inc., 2012).

Regarding the solutions that PolyVista provides its users, the most interesting and relevant, from the point of view of this study, are the solutions developed for the transportation industry, which are divided into two approaches: regulatory and safety and fleet and operations. The first uses incident report data and operational data as source material which provides benefits for the speed of the analysis, resulting in reduced costs, improved efficiencies, a faster identification of potential safety issues and the mitigation of risks, laying the basis for the development of proactive procedures instead of reactive ones. The second area processing starts from service and maintenance records, breakdown data, driver data, vehicle configuration data, vehicle cost data and safety data. Mining this kind of source material improved safety performance and reduced breakdowns. This was achieved by discovering hidden cause and effect relationships. All this can lead to maximised operational efficiency, optimised fleet performance, better asset utilisation and value as well as reduced costs. Parallel application areas are in high-tech industry and manufacturing, where this system would create quality, reliability and guarantees by analysing large amounts of data from multiple sources and providing the ability to connect qualitative and quantitative data for a holistic overview of the process. All that is in addition to the ability to discover the root causes of issues faster, leading to early warnings of lethal events and to an enhanced core quality and failure analysis. It is obvious that high-tech issues are applicable in the transportation context as well. Significant characteristics in this area are monitoring and alarm functions for identifying product safety issues. Generally, reduced costs and improved operational efficiency are enabled by the discovery of hidden causes and effective relationships (PolyVista Inc., 2013c).

In addition to the information presented in Publication #1 “Data Mining in Aviation Safety Data Analysis” about data pre-processing functions, an overview of them will be given here. As stated before, stop words are those to be ignored because they do not provide any insight. They are processed in the way that the mining tools generally operate. A list of synonyms maps multiple synonyms into one word. Misspellings and abbreviations can also be fixed here. Optionally, non-alphabet and non-number characters are mapped to 'blank', and numbers are mapped to 'delete'. The format and layout of the file is as follows:

'word or phrase' tab 'word'

PolyVista TM (=Text Mining) processes the synonym file in 2 modes: *add* and *replace*. In the ‘Add’ mode the new word is added whenever the original word is found. In the ‘Replace’ mode the original word is replaced with the new word.

```
acc:hen                         aluelennonjohto
acc:lle                         aluelennonjohto
acc:lta                         aluelennonjohto
```

```
acc:n                          aluelennonjohto
acc:ssa                        aluelennonjohto
adiz-alueella                  adiz
adiz-aluerikkomus              adiz
adiz-vyöhykkeelle              adiz
adiz-vyöhykkeellä              adiz
```

In the words presented above, the rules will not process as expected because the period '.' and colon ':' and the dash '-' have already been converted to a 'blank' by the rules that precede them.

```
.      <blank>
:      <blank>
-      <blank>
```

Therefore, the TM (Text Mining) engine will not find the "acc:hen" text string in the narratives. By default, substitution does not happen mid-string. However, mid-string substitution is possible but has to be specified by using a partial token, for example:

```
string1 <tab> string2 <tab> partial
```

In other words, if string1 is found in a text mid-string, it is replaced by string2, but will only happen when the 'partial' token is specified.

The function *include* is practical in a situation where a researcher wants a word to be part of the clustering but not the main focus of a cluster, in other words, it is present in a cluster to provide additional information. For example, in safety reports, words like pilot, airport, runway, etc. are likely to be found but these might not be put in the "stop" file, but at the same time the researcher does not want them to be the main words in a cluster. In this situation the weight of such words can be reduced.

```
kone         [plane]          0.1
ohjaaja      [pilot]          0.1
kiitotie     [runway]         0.1
tulla        [to come]        0.1
jalkeen      [after]          0.1
lento        [flight]         0.1
efhk         [airport code]   0.1
helsinki                      0.1
aikana       [during]         0.1
alue         [area]           0.1
oikea        [right]          0.1
vasen        [left]           0.1
keskeytys    [plane]          0.1
teline       [gear]           0.1
alkoi        [began]          0.1
```

This tool is provided alongside other ‘normal’ functions for processing stop words and synonyms as well as the function concepts of interest, some examples of which are presented as follows:

alert	box	10	altitude_deviation	10
challenge	captain	10	crew_issue	10
flight_attendant	cockpit	10	distraction_issue	10
not cleared	descend	10	altitude_deviation	10
not hear	initial check	10	communication_issue	10
traffic	tcas	15	altitude_deviation	10
trouble	descend	10	descend_problem	10
unusual	flight	10	unusual_issue	10
very	busy	10	distraction_issue	10

The format/layout of the expression is:

'word or phrase' tab 'word or phrase' tab 'number of words in between' tab 'concept' tab 'weight'.

The first row, the examples in the explained form are as follows: *If 'alert' and 'box' are within '10' words of each other then create a concept called 'altitude_deviation' and give it a weight of '10'.*

The function of these concepts is to detect concepts that could be valuable in achieving more accurate results, especially by understanding that existing words are processed differently from the point of view of their position compared to the other words. To express this more precisely, the words alone do not have the same information value as a combination of them with other words that might be more significant; the value of the combination is determined by the context of their position. This function was not utilised because of some major problems. The main reason was that to build a model to use this function would have been a major task but not of great benefit because the test dataset was rather small. If the mining process would have been applied to the whole report dataset in a database, it might have been appropriate to use this. Another reason was that because there was no mining experience of using this scale before, the aim was to keep the testing arrangement as simple as possible. This function is nevertheless a very interesting one to be used later with a large amount of data after several rounds of mining experience has been gained. Additionally, due to the structure of Finnish the concepts should have been determined in advance before applying the lists of stop words and synonyms – after this process the concepts may not be discoverable as defined in this form.

The author did not perform the mining processes himself, but delivered the lists of stop words and synonyms to the company that performed the processing of the data. Clustering was chosen as the method to be used for the testing. This function allowed the operator to determine the number of clusters to receive relevant mining results. This was the first time Finnish text had been mined using the PolyVista system and it succeeded surprisingly well.

The results of the first mining round with PolyVista were examined in detail in Publication #1 “Data Mining in Aviation Safety Data Analysis”, the refining process of the data alongside the definitions used in order to prepare the material for the second mining round are shown and explained in Publication #2 “Data Refining in Mining Process for Aviation Safety Data” and the final results of the second mining round are presented in Publication #3 “Data Mining in Aviation Safety”. The results were returned to the researcher in the form of pictures taken from the Cluster Browser window. The first run was a preliminary test and made without any pre-processing. The results are shown in Figure 27.

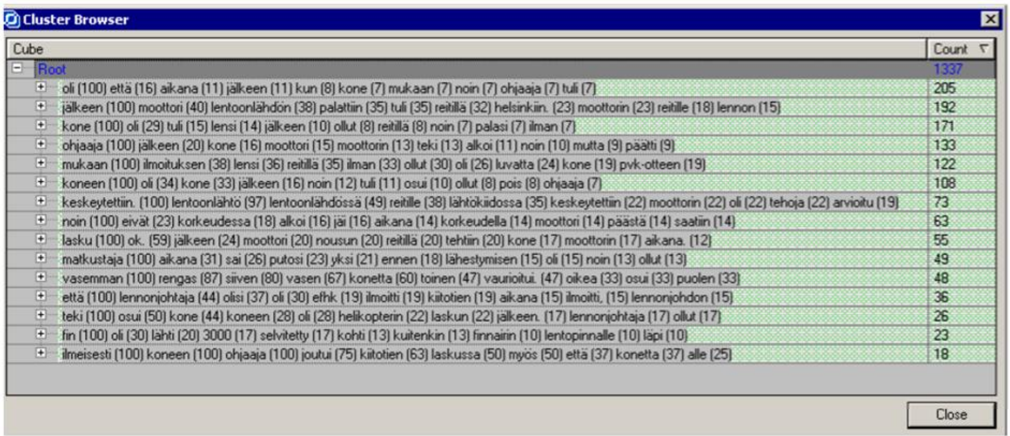


Figure 27: Mining results PolyVista without pre-processing of the words

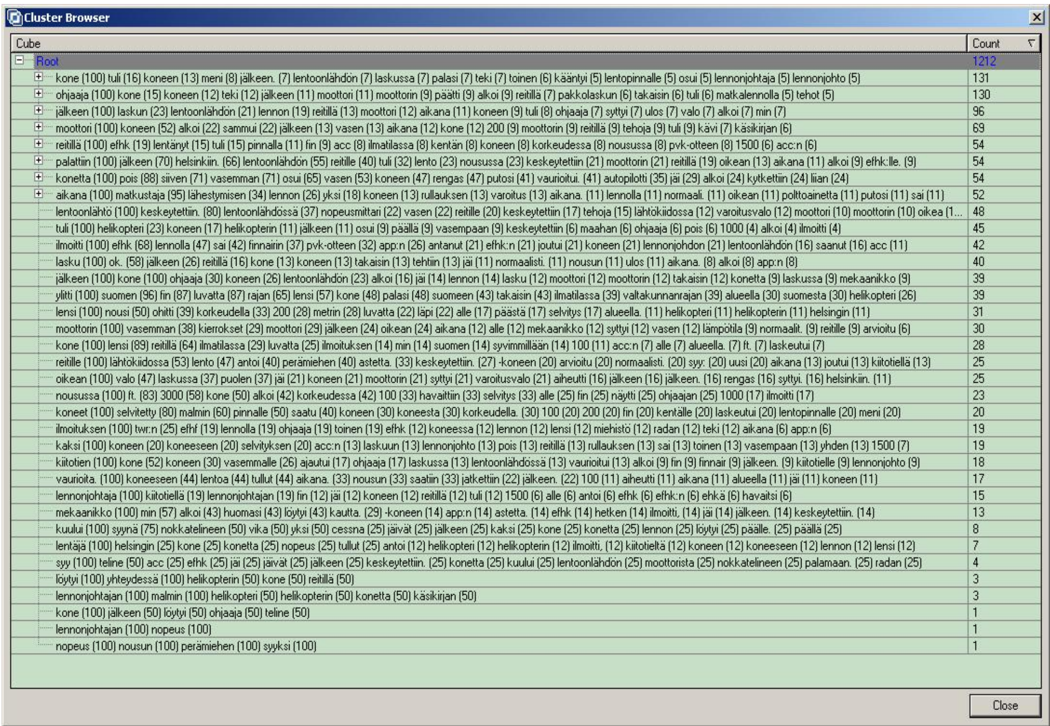
The output ought to be read as follows:

- the first cluster has 205 records
- the most important word in the first cluster is “oli” (was) with a score of 100
- the second most important word in the first cluster is “etta” (so that) with a score of 16, this implies that the rest of the words are not significant.

The first proposal selected 15 clusters as a guess, a definition that could later be refined by performing a couple of iterations in order to find more relevant mining results. As a default, the number of significant words, i.e. the words describing the cluster, was selected as 10, starting from the left – the first of them having the score of 100. The importance of the words decreased from left to right; the last word taken into account had the least significance. In Figure 27, the numerical score after the word means the significance of that word in a cluster. As there was no pre-processing before the first test, it was no surprise that the most important word clusters 1, 2, 5, 8, 11, 12, 14 and 15 are stop words. Observing all the most important words from each cluster, the average percentage of the stop words was almost 35, varying from 10 to 70. As for the size of the clusters, the amount of stop words among the ten most important words does not have any correlation with cluster size. The significance of observing the stop words in the mining process appears the second time the same results are analysed. As the most important word with a score of 100 and the second with a score of 16 are both stop words, the rest of the words are not significant at all. This leads to the conclusion that the whole cluster of 205 reports could be deleted on the strength of having no significance from the point of view of the mining process.

Because the reports of the cluster include valuable information, the significance of observing both the stop words and the synonyms as well as taking them into account is of great importance. According to these results, the number of synonyms is somewhat smaller, 23 per cent. The number of them in the most significant words seem to have the values 0, 2 or 4, which makes the appearance of the synonyms smaller the bigger the cluster is. As before, the distribution does not seem to be correlated with cluster size. When the stop words are put into the mining process, the results look completely different, as shown in Figure 28. Because the synonyms are not taken into account, there appears to be quite a number of them. Compared with the previous mining round, the amount of clusters has increased from 15 to 35 with their sizes varying between 1 and 131 reports. As one report does not create a cluster alone, the last three were deleted from the process. All the tools found reports that do not belong to any cluster and are thus left out. Even if the smallest clusters consist of only three reports they could well be taken into consideration. However, interestingly, in two reports of the three left all the words have the score of 100, and in the third report the first word has the score 100 and the rest have 50.

During the whole mining process with PolyVista, the most important word has the score of 100. However, in contrast to the previous observations the score of the second most important word is much higher indicating that the homogeneity of the clusters is better as it means that the cases are closer to each other. Using stop words, the result can be seen as significantly better from the observed 32 clusters; in 20 clusters the score of the most important words is at least 50, according to which more than half of the cases, 62 per cent, are more homogenous than the others. Compared with the previous situation, the corresponding number was only 4 clusters out of 15, which is no more than 27 per cent of the clusters and can be regarded as homogenous. This trend clearly indicates that by using stop words the research process will obtain more accurate mining results. In addition, clusters that consist of less than 40 reports seem to provide some information when their content is compared with the importance of the words.



Cluster Browser

Cube	Count
Root	1212
kone (100) tuli (16) koneen (13) meni (8) jälkeen. (7) lentoonlähdön (7) laskussa (7) palasi (7) teki (7) toinen (6) kääntyi (5) lentopinnalle (5) osui (5) lennonjohtaja (5) lennonjohto (5)	131
ohjaaja (100) kone (15) koneen (12) teki (12) jälkeen (11) moottori (11) moottorin (9) päätti (9) alkoi (9) reitillä (7) pakkolaskun (6) takaisin (6) tuli (6) matkalennolla (5) tehot (5)	130
jälkeen (100) laskun (23) lentoonlähdön (21) lennon (19) reitillä (13) moottori (12) aikana (11) koneen (9) tuli (8) ohjaaja (7) syttyi (7) ulos (7) valo (7) alkoi (7) min (7)	96
moottori (100) koneen (52) alkoi (22) sammui (22) jälkeen (13) vasen (13) aikana (12) kone (12) 200 (9) moottorin (9) reitillä (9) tehoja (9) tuli (9) kävi (7) käsikirjan (6)	69
reitillä (100) efhk (19) lentänyt (15) tuli (15) pinnalla (11) fin (9) acc (8) ilmatilassa (8) kentän (8) koneen (8) korkeudessa (8) nousussa (8) pvk-otteen (8) 1500 (6) acc:n (6)	54
palattiin (100) jälkeen (70) helsinkiin. (66) lentoonlähdön (55) reitille (40) tuli (32) lento (23) nousussa (23) keskeytettiin (21) moottorin (21) reitillä (19) oikean (13) aikana (11) alkoi (9) efhk:lle. (9)	54
konetta (100) pois (88) siiven (71) vasemman (71) osui (65) vasen (53) koneen (47) rengas (47) putosi (41) vaurioitui. (41) autopilotti (35) jäi (29) alkoi (24) kytkettiin (24) liian (24)	54
aikana (100) matkustaja (95) lähestymisen (34) lennon (26) yksi (18) koneen (13) rullauksen (13) varoitus (13) aikana. (11) lennolla (11) normaali. (11) oikean (11) polttoainetta (11) putosi (11) sai (11)	52
lentoonlähtö (100) keskeytettiin. (80) lentoonlähdössä (37) nopeusmittari (22) vasen (22) reitille (20) keskeytettiin (17) tehoja (15) lähtökiidossa (12) varoitusvalo (12) moottori (10) moottorin (10) oikea (1...	48
tuli (100) helikopteri (23) koneen (17) helikopterin (11) jälkeen (11) osui (9) päällä (9) vasempaan (9) keskeytettiin (6) maahan (6) ohjaaja (6) pois (6) 1000 (4) alkoi (4) ilmoitti (4)	45
ilmoitti (100) efhk (68) lennolla (47) sai (42) finnairin (37) pvk-otteen (32) app:n (26) antanut (21) efhk:n (21) joutui (21) koneen (21) lennonjohdon (21) lentoonlähdön (16) saanut (16) acc (11)	42
lasku (100) ok. (58) jälkeen (26) reitillä (16) kone (13) koneen (13) takaisin (13) tehtiin (13) jäi (11) normaalisti. (11) nousun (11) ulos (11) aikana. (8) alkoi (8) app:n (8)	40
jälkeen (100) kone (100) ohjaaja (30) koneen (26) lentoonlähdön (23) alkoi (16) jäi (14) lennon (14) lasku (12) moottori (12) moottorin (12) takaisin (12) konetta (9) laskussa (9) mekaanikko (9)	39
ylitti (100) suomen (96) fin (87) luvatta (87) rajan (65) lensi (57) kone (48) palasi (48) suomeen (43) takaisin (43) ilmatilassa (39) valtakunnanrajan (39) alueella (30) suomesta (30) helikopteri (26)	39
lensi (100) nousi (50) ohitti (39) korkeudella (33) 200 (28) metrin (28) luvatta (22) läpi (22) alle (17) päästä (17) selvitys (17) alueella. (11) helikopteri (11) helikopterin (11) helsingin (11)	31
moottorin (100) vasemman (38) kierrokset (29) moottori (29) jälkeen (24) oikean (24) aikana (12) alle (12) mekaanikko (12) syttyi (12) vasen (12) lämpötila (9) normaalit. (9) reitille (9) arvioitu (6)	30
kone (100) lensi (89) reitillä (64) ilmatilassa (29) luvatta (25) ilmoituksen (14) min (14) suomen (14) syvimmillään (14) 100 (11) acc:n (7) alle (7) alueella. (7) ft. (7) laskeutui (7)	28
reitille (100) lähtökiidossa (53) lento (47) antoi (40) perämiehen (40) astetta. (33) keskeytettiin. (27) -koneen (20) arvioitu (20) normaalisti. (20) syy: (20) uusi (20) aikana (13) joutui (13) kiitotiellä (13)	25
oikean (100) valo (47) laskussa (37) puolen (37) jäi (21) koneen (21) moottorin (21) syttyi (21) varoitusvalo (21) aiheutti (16) jälkeen (16) jälkeen. (16) rengas (16) syttyi. (16) helsinkiin. (11)	25
nousussa (100) ft. (83) 3000 (58) kone (50) alkoi (42) korkeudessa (42) 100 (33) havaittiin (33) selvitys (33) alle (25) fin (25) näytti (25) ohjaajan (25) 1000 (17) ilmoitti (17)	23
koneet (100) selvitetty (80) malmin (60) pinnalle (50) saatu (40) koneen (30) koneesta (30) korkeudella. (30) 100 (20) 200 (20) fin (20) kentälle (20) laskeutui (20) lentopinnalle (20) meni (20)	20
ilmoituksen (100) twr:n (25) efhf (19) lennolla (19) ohjaaja (19) toinen (19) efhk (12) koneessa (12) lennon (12) lensi (12) miehistö (12) radan (12) teki (12) aikana (6) app:n (6)	19
kaksi (100) koneen (20) koneeseen (20) selvityksen (20) acc:n (13) laskuun (13) lennonjohto (13) pois (13) reitillä (13) rullauksen (13) sai (13) toinen (13) vasempaan (13) yhden (13) 1500 (7)	19
kiitotien (100) kone (52) koneen (30) vasemmalle (26) ajautui (17) ohjaaja (17) laskussa (13) lentoonlähdössä (13) vaurioitui (13) alkoi (9) fin (9) finnair (9) jälkeen. (9) kiitotielle (9) lennonjohto (9)	18
vaurioita. (100) koneeseen (44) lentoa (44) tullut (44) aikana. (33) nousun (33) saatiin (33) jatkettiin (22) jälkeen. (22) 100 (11) aiheutti (11) aikana (11) alueella (11) jäi (11) koneen (11)	17
lennonjohtaja (100) kiitotiellä (19) lennonjohtajan (19) fin (12) jäi (12) koneen (12) reitillä (12) tuli (12) 1500 (6) alle (6) antoi (6) efhk (6) efhk:n (6) ehkä (6) havaitsi (6)	15
mekaanikko (100) min (57) alkoi (43) huomasi (43) löytyi (43) kautta. (29) -koneen (14) app:n (14) astetta. (14) efhk (14) hetken (14) ilmoitti, (14) jäi (14) jälkeen. (14) keskeytettiin. (14)	13
kuului (100) syynä (75) nokkatelineen (50) vika (50) yksi (50) cessna (25) jäivät (25) jälkeen (25) kaksi (25) kone (25) konetta (25) lennon (25) löytyi (25) päälle. (25) päällä (25)	8
lentäjä (100) helsingin (25) kone (25) konetta (25) nopeus (25) tullut (25) antoi (12) helikopteri (12) helikopterin (12) ilmoitti, (12) kiitotieltä (12) koneen (12) koneeseen (12) lennon (12) lensi (12)	7
syy (100) teline (50) acc (25) efhk (25) jäi (25) jäivät (25) jälkeen (25) keskeytettiin. (25) konetta (25) kuului (25) lentoonlähdön (25) moottorista (25) nokkatelineen (25) palamaan. (25) radan (25)	4
löytyi (100) yhteydessä (100) helikopterin (50) kone (50) reitillä (50)	3
lennonjohtajan (100) malmin (100) helikopteri (50) helikopterin (50) konetta (50) käsikirjan (50)	3
kone (100) jälkeen (50) löytyi (50) ohjaaja (50) teline (50)	1
lennonjohtajan (100) nopeus (100)	1
nopeus (100) nousun (100) perämiehen (100) syyksi (100)	1

Close

Figure 28: Mining results using PolyVista with stop words

The next step was taking the synonyms within the process and continuing the process in a different way as the number of clusters was raised from 6 to 20 in steps of 2. The results show six clusters with sizes varying from 117 to 410 and the changes in cluster size did not change the distribution or the other factors. In fact, the only visible change was a diminishing number of reports in the clusters. The only significant change in comparison with the situation before was that in more than half of the clusters, the importance of the second most important word was higher than before. In the 20 clusters, their sizes varied between 10 and 232; the largest clusters contained between 96 and 232 reports which created a need to split them into smaller units. The rest contained 67 or less reports. In the same context, it makes sense to notice that in 11 clusters, somewhat more than half, the scores of the three most important words exceeded 50, and even in the last clusters that contained 10 reports, the score of 8 of the 10 most important words was also more than 50, which can be considered an accurate mining result.

Relevant and usable results have been found through all phases of the research, especially when working with the smallest clusters because the most important words would seem to be quickly estimated as being of significant character and they receive high scores. It is worth noticing that, as with the two other tools, these results were achieved without using iteration. The research continued in the second mining round with tools #3 and #4 (see Publication #2 “Data Refining in Mining Process for Aviation Safety Data” and Publication #3 “Data Mining in Aviation Safety”). PolyVista also provided an interesting function called Concepts of interest (see page 138). Its utilisation might have produced an absorbing addition to the mining results, although its significance with this data obviously would not have been notable.

5.4 Inspecting the mining results of the first round with NVivo

The inspection process has been explained in detail in Publication #2 “Data Refining in Mining Process for Aviation Safety Data”. The evaluation of the mining results is an essential and unavoidable phase in the mining process, because the mining results can be regarded as suggestions. As the author had worked as a flight safety inspector, a cursory ocular assessment exposed the rough mistakes and shortages. However, this sort of analysis is far away from the accuracy required in this context. That is why a special analysis tool was indispensable in order to gain a more complete view on the mining results. Using Excel and Word alone would not have led to a satisfactory goal in understanding the relationships between the concepts in the test data. A qualitative data analysis application called NVivo provided a good solution to the problem, enabling the user to manage and make sense of unstructured information; version 7 was used because it was the latest at that time. The contents of the clusters were analysed using queries. NVivo took into account all the words defined in the options, determining their relevance in the text.

Only the mining results of GILTA were analysed because the results of PolyVista (see Chapter 5.3.4, Figure 27 and Figure 28) were available as weights of the significant words in the clusters. Despite the different format for presenting the mining results, the analyses with NVivo were also applicable for PolyVista because both of the tools applied the same lists of synonyms and stop words which were adjusted for the second mining round. Analysing the mining results of TEMIS (see Chapter 5.3.2) with NVivo did not seem relevant, because it was not possible to perform the second mining round and the results were easily analysable as such due to its sophisticated web-based interface for presenting the results.

A total of 29 queries were performed with NVivo. As an example, the different expressions for ‘flight level’ (compound ‘lentopinta’ or simply ‘pinta’ in this connection in Finnish) were among the most frequent in the data and therefore its appearance was discovered. The text search criteria of the query were as follows:

fl* OR lentopin* OR pinta* OR pinn* NOT flo* NOT fla* NOT flu* NOT fli* NOT pintal* NOT pinnas* NOT pinnoit* NOT pintaker* NOT pintala* NOT pintale*

As the clusters contained the reports in their original form, all the conjugation forms were to be taken within the query and that is why wildcards – any characters (*) after the stem of the word were used. NVivo took into account all the words defined in the query options and determined their relevance in the data mass. The queries were created using the operators “or” and “not” in order to include defined words in the query and to exclude others. This was necessary because the Finnish word ‘pinta’, which in this analysis ought to have the meaning ‘level’, can have also other meanings, such as ‘surface’, and those needed to be excluded from the reports. The results are illustrated in Figure 29, which was produced using NVivo version 10. The list of the classes (the clusters named in GILTA) is sorted in descending order from the highest coverage value.

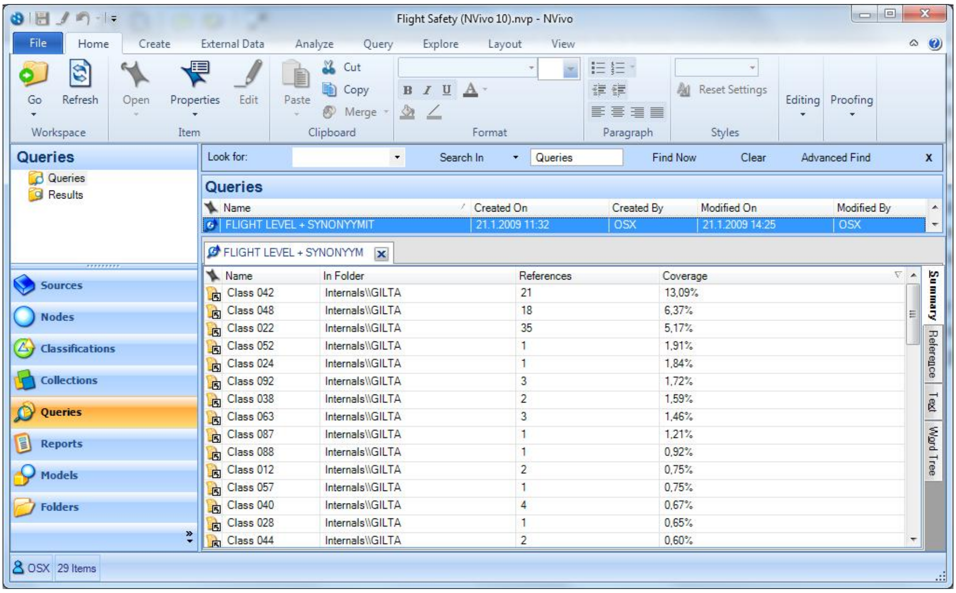


Figure 29: Query: Flight level and synonyms

In Figure 29 only the 15 most significant classes in this query are shown. The highest of them were the classes 042, 048 and 022 with corresponding references (the number of appearances of the searched for word) of 21, 18 and 35, having coverages of 13.09, 6.37 and 5.17 per cent. The coverage is the significance of the searched for word in the report. In 17 classes there was only one reference and the coverage was less than 1 per cent in 21 classes in this query. The appearance of the expression ‘flight level’ in the mining results was investigated with five different queries. This process revealed that it figured a total of 50 times, a finding which brought it into the group of relatively frequent words, thus skewing the mining results. In order to avoid or at least reduce this influence, its occurrence has to be decreased. If the written expressions for it, including all the synonyms and case forms were counted, the total amount exceeded 150.

5.5 Discussion

Commercial text-mining products have been available for a long time and new prototypes are being continuously developed. Kloptchenko (2003) compiled a table (3-2) to list the 19 best-known text mining (TM) systems at the end of 2002. It shows the main projects from the corporate and academic world. The first four systems in the table mentioned were developed and implemented at different universities. One of them is WebSOM by Kohonen at Helsinki University of Technology (Kohonen, 1999). This system was the basis for the GILTA Project (Tool #3 in Chapter 5.3.3). Companies specialised in data and text mining solutions for business intelligence, such as Megaputer Inc. and Temis, as well as large corporations, such as the SAS Institute and the IBM Corp., designed the rest of the systems in the table mentioned. In this group, the Text Analyst 2.0 is mentioned, which is a component of Tool #1 described in Chapter 5.3.1. Some of the systems are query-based methods, which rely heavily on the use of term (keywords, items, indexes) extraction.

The absolute majority of the presented TM applications described by Kloptchenko (2003) used stemming, synonym lists, stop word removal, text parsing, dimension reduction, and clustering, or the categorisation procedures of the coded data. A synonym table composition assigns one meaning to every term used in a document vocabulary (Riloff & Hollaar, 1996). *Stop words* frequently appear among the documents, i.e. articles, prepositions, and conjunctions. The elimination of stop words reduces the size of the indexing structure (Baeza-Yates & Ribeiro-Neto, 1999). *Text parsing* algorithms convert text into words, phrases or clauses and put them into short memory. Text parsing decomposes text and generates a quantitative representation suitable for data mining (Mayes, Brewes, & Thompson, 2002). Dimension reduction algorithms treat every document as vectors where each dimension is a count of the occurrences of a different word, this results in tens of dimensions in every document (Isbell, 1998). Independent Component Analysis and restricting weighting to words only specified in the query are applied in order to create a reduction in the amount of dimensions (Kolenda & Hansen, 2002). The importance of executing word sense disambiguation based on a thesaurus or on-line dictionaries for increasing the performance of information retrieval (IR) systems was discussed by Sanderson (1994).

The famous French Enlightenment writer, historian, and philosopher Voltaire has said “Le mieux est l'ennemi du bien” – “The best is the enemy of the good”.

It is a relevant guideline in order to identify and define simple and easily usable safety indicators to discover weaknesses and shortcomings in the safety processes of the organisations. These indicators should be based on significant kinds of events or states of system control. The indicators are valuable and important tools to be used with systems that collect comprehensive data on incidents, but they are not solutions themselves to safety problems. The essential point is that because the people in the organisations understand the causes of potential accidents, they can think of ways to reduce their frequency or even eliminate their causes (Brooker, 2006).

6 CONCLUSIONS

6.1 Results and key findings

The most important result of the research process was that it confirmed that data mining could be applied to material that consists of short reports written in Finnish. The first round produced encouraging results – as the smallest clusters were found to consist of similar cases. Another important finding was that the sizes of the clusters play a significant role when the results are to be applied. In addition, the first round alone confirmed that text mining is a challenging task, especially for small language groups. Flight safety reports written in English and other major languages had already been mined before this research had begun; text mining had only been applied to longer texts in Finnish, mainly for magazine articles.

No lethal trends were found when clustering the report data, which means that the test data did not contain any events that might be threatening. However, based on the results it is obvious that if events leading to lethal trends would have existed in the data, they would have been discovered and brought out. The text mining tools used were capable of extracting trends – actually recurring events – that turned out to be incidents. However, in the cases studied they did not develop into dangerous risks or accidents. The ability to discover lethal trends with clustering was supported by the mining results achieved in the research reported in Publication #4 "Using cluster analysis to identify the weak signals of lethal trends in aviation and health care documentation". It was found that the identification of hazardous events in healthcare documents is similar to the identification carried out in aviation reports. In both contexts, the main focus was on discovering and identifying weak signals in the documentation.

If a couple of similar cases are found, this does not, however, automatically indicate that a lethal trend exists. Only a thorough examination and investigation made by a human analyst can confirm that and find the reasons why a trend is occurring. In this study, recurring cases (or cases with the same content), which could be regarded as trends and which could have been hazardous, were found. This clearly illustrates that a highly skilled and experienced interpretation is required in order to reveal hidden information, as the mining results are not directly applicable in supporting decisions. All the achieved mining results have supported the premise that if the data, on the whole, contain lethal trends, this method would reveal critical safety information which is not accessible with other methods. The research also proved the applicability of Reason's Swiss Cheese model for this context.

All the systems indicated that the number of clusters is a significant factor in the process: the more clusters the better the results. However, finding an optimum point would be important. Theoretically the number of clusters could be raised to equal the number of cases, although this would not make much sense. With the data collected, the maximum size of the clusters seemed to be about 30 reports, which made a cursory, visual investigation of the contents of the clusters possible. This of course depends on the type of the reports. If the clusters had contained less than 15 reports, it would have been relatively easy to find the common factors that created the cluster. Benchmarking by using three different tools was one of the goals of this study. The achievement of benchmarking was confirmed by the discovery that, in most cases, the same reports were indicated as outliers by all the used tools. In addition, the mining results of PolyVista and GILTA could easily be compared using the weights of the most significant words in the clusters, and largely the same clusters were found. Also the mining results of TEMIS could, in most cases, be used in the same comparison and similar clusters were found. Although the author did not have access to the clustering algorithms used in the tools, it did not seem important because all the systems produced, in the main, similar results, proving that the research was on the right track.

One of the expected results was that data mining, and especially text mining, should be an iterative process; this was confirmed by all the tested systems. Hence possible shortcomings and mistakes made during the process could be brought out and identified. In order to achieve better mining results, increasing the amount of clusters is not always the best way to proceed. The process showed that stop words and synonyms need to be controlled and updated whenever necessary. About one hundred checking procedures with stop words and synonyms were developed and found to be reasonable for the size of the data sample.

At the beginning of this thesis, the current satisfactory flight safety state and its development history was presented. However, as traffic volumes are forecast to grow significantly in the near future, the number of accidents and serious incidents will increase, at least proportionally, which cannot be accepted as air traffic accidents are very costly in terms of human lives. The overall research question was how to improve flight safety by applying new methods of analysis because the use of traditional methods of analysis had already reached its zenith. In order to proceed with developing the research frame, the process was formulated into the research question "Can flight safety be augmented by applying data mining as a method of analysis?" In order to answer the question accurately, it was divided into two sub-questions (SQ). The detailed answers to the research question have been worked out throughout the thesis and they are summarised in the following:

SQ1: Is it possible to find lethal trends in aviation safety data with text mining?

The answer is yes if they exist in the sets of events and the mining definitions are determined in a way that finding reports on similar events makes it possible to identify them. The answer is no if the short texts submitted as event reports are of poor quality and similar reports cannot be found and identified. This study demonstrates that text mining and especially clustering can be used for cases when previously undetected instances are being sought. After the basic procedures (shown and explained in Figure 16) were performed, the first mining rounds – made with all three systems that were tested – produced understandable and coherent mining results.

SQ2: What sort of refining actions would make the mining results more applicable and are the results significantly more accurate, reliable, and valid after refining?

Making the mining results more applicable requires standard refining and the adjusting of techniques or operations which can improve the quality of the short texts to the standard required by text mining methods. The answer to the latter part of the question is yes, it can be shown and proved that the text mining results are significantly more accurate, reliable and valid after refining. As data mining is an iterative process, the definitions ought to be refined between the rounds in order to make the mining results more accurate. The obvious mistakes were corrected, such as adding word tokens to the stop words list if they were forgotten in the pre-processing phase; the list of synonyms was also adjusted. Described in detail in Publication #2 “Data Refining in Mining Process for Aviation Safety Data” about 100 refining actions were performed to try to keep the data as unchanged as possible. Checking the mining results accurately also requires specialist computing aids, which is why NVivo version 7 was used to reveal and document the refining needs.

As the refinement actions were performed, the accuracy of the mining results improved significantly – as reported in Publication #3 “Data Mining in Aviation Safety”. This fact was tested in several ways, first checking the so-called ‘sense-making clusters’, the number of which increased, and then using Excel to compare the same clusters found in the results of both tools after the second round was completed. After refining, the results were more coherent when the results of the two tools were compared with each other. Thus, the overall answer to the RQ is positive, flight safety can be augmented by applying data mining as a method of analysis.

As pointed out several times, aviation has been compared to other safety critical industries like health care and nuclear power in many different contexts. This connection can be used as support in both directions: the principles of performing this study were derived from similar studies performed in the areas mentioned and the results of this study could be applied in the research project described in Publication #4 “Using cluster analysis to identify weak signals of lethal trends in aviation and health care documentation.” The approach to this benchmarking is twofold. First, from the point of view of the mining process itself, the same mining format might be appropriate for any text data saved electronically – if it is accessible to mining tools. This means that text mining and especially the clustering that was used in this study are applicable to any area using documents that are available and have a relevant description of the context, the events, the actions, etc. Second, the mining results might be applicable in any area for collecting reports in order to control and develop operations, especially the safety critical industries like nuclear power and health care.

6.2 Theoretical implications

This study shows that clustering short reports written in Finnish with a tool that was not originally built for mining Finnish texts produced a satisfactory outcome, as similar clusters were formed with the tools tested. Of the applied systems tested, TEMIS had a text mining capability in Finnish due to its language module, which was tested with real data for the first time in this study. The second tool, GILTA, was language independent. Mining the reports with PolyVista was a pure ad hoc attempt without any preliminary expectations but was based on the good results achieved when mining texts in Spanish. Regardless of the system used, the challenge with all of them was whether mining the short texts of the flight safety reports would produce correct and useful results. This was successful and was proven by testing and comparing the mining results with the different systems.

6.3 Implications for practice

This study shows that the way in which data mining differs from the so-called traditional systems is in its capability for extracting hidden information from vast amounts of data, which might not be possible if other systems are applied. This capability is based on its special features, artificial intelligence and machine learning which enable knowledge discovery (as shown in Figure 16 and Chapter 3.3.1) and the production of answers to the question of discovering ‘we do not know what we do not know'. This means that we are searching for something which we are not aware of and that is why we cannot apply ordinary queries to find information hidden in the data. Data mining systems do not necessarily produce straight answers to explicit questions, which is why they are used as decision support systems and can be combined with other methods like business intelligence systems. Data mining can be applied as a 'scalping sieve' to find out whether something that might be worth inspecting in more detail exists in the (safety) data. After discovering recurring cases, for instance, the parameters defining this suspected trend could be stated as search criteria in business intelligence systems to find out more about a possibly lethal trend. If we look at Figure 19 again, we can see the situation has moved on from State 4, which requires knowledge discovery to State 2, which we already know, thus methods like adhoc and OLAP analysis can be applied.

On the basis of this study, the mining methods could be applied to other areas like healthcare and strategic management. Applications within healthcare are manifold as health care is similar to the airline industry – both are safety critical domains. Thus, the applicability for this area is quite straightforward. As for strategic management, the applicability differs according to context and industry, which have differing needs for data and information. Strategic management is not a safety critical domain in the same way as health care but mining methods have proved to be useful for coping with the challenges of big data. However, regardless of the application area, the essential focus is on discovering weak signals (discussed in detail in Chapters 2.1 and 3.2.2). These are very minor indicators that are easily dismissed because they do not necessarily portend any harm in the future. However, in the worst case scenario, their development may result in a catastrophe if the event they are indicating is recurrent, thus indicating an ambiguous threat. These sorts of threats are usually different in safety critical organisations as they can lead to the loss of life.

6.4 Limitations

The most important limitation to be mentioned are the number of mining rounds. Two rounds were performed with two of the tested systems and one round with the third one. Data mining should be an iterative process, a fact which was confirmed in this study. Both in theory and practice, adding mining rounds makes the results more accurate, but it makes no sense to continue the process endlessly. There is always an optimum point to be found, depending on the characteristics of the data to be mined and the demands placed on the mining results. As Zadeh (1979) has expressed, a trade-off between precision and relevance exists, thus, if precision is to be improved, we will have to give up on relevance and vice versa.

Another significant limitation was the rather narrow sample data. The data consist of flight safety reports covering a period of three years and the data set was large enough to show that trends in the form of similar cases can be extracted from short text reports. The trends found in the data did not indicate any hazard but if they existed they could have been revealed by the methods applied here. If the sample had been larger, containing several thousands of reports, lethal trends might well have been discovered. At least recurring events needing more detailed investigation may have emerged.

From the database records, only the narrative data field was included in the mining process alongside the identification data fields. The data taken into the process was limited in purpose because only text mining could be carried out. If all the other fields or at least a couple of them were taken into the mining process the results, i.e. the clusters, might have been different. However, based on a very careful analysis of the mining results, the results, in all probability, would not have differed significantly from those already achieved. One limitation, the impact of which is difficult to estimate, is that the mining tools were operated by professionals and the author himself could not directly affect the processes. The mining procedures were defined and agreed between the author and the operators in advance via e-mail and they were adjusted as the mining process proceeded. If the author could have used the tools himself, it would have been possible to experiment with functions included in the interfaces of the mining tools and find efficient and interesting methods with which to refine the mining results. One interesting function that could have been tested would have been the Concepts of interests by PolyVista, which was presented in Chapter 5.3.4.

6.5 Future research questions

The mining process consisted of two mining rounds with two of the tested tools and one round with the third one. These rounds show that more mining rounds as well as redefining the criteria between the rounds could have produced better clusters. It is obvious that performing a third round with two tools and a second round with the third tool would have produced more accurate results. However, there would have also been other possibilities for obtaining more accurate mining results. One method would be to re-mine the bigger clusters from the results of the previous round. Determining the borderline depends on the average size of the clusters in proportion to the data to be mined. This method would have been suitable for the first mining results of GILTA as the biggest cluster contained 158 reports and the ‘best’ clusters had about 20 reports. One of the major factors for creating a big cluster was the frequent appearance of the word token ‘kone’ (=aircraft) in its different forms and synonyms. In such situations it would help to re-mine the large cluster and then perform the same procedure after filtering away the most frequent word token. A method with a somewhat similar goal was used in TEMIS – where the biggest cluster, exceeding 100, divides into two sub-clusters. As the sizes of the clusters pay a significant role in the text mining process, it would be worth experimenting with the possibility to define the upper limits of the clusters, thus preventing the clusters growing too big beforehand.

Due to the scope of the research, only the narrative part of the data was used for processing the text mining tools. From the very beginning the excluded data fields were planned to be taken into the study after having received relevant mining results by text mining. Because analysing texts written in free form was the aim, the research began from this point. Beginning by taking all the fields into the study would not have produced as accurate results as pure text mining, but experimenting with them afterwards may have produced some additional value for the mining process. Combining the results received from text mining with statistical methods would provide an interesting research challenge, which means using numerical data analysis as part of the research process. This sort of extension would be possible if the numerical data in the structured fields of the reports were processed with statistical methods and the results were combined with text mining results. Based on the experience of the research process so far, it seems obvious that the sample data set of 1,240 reports may be too limited – two thousand reports might be a relevant amount from this point of view.

As data mining serves as a decision support tool, it does not give straight answers to questions. It might, however, provide a starting point for more focused processes that could give more detailed answers. This would be possible by combining data mining with business intelligence (BI). As data mining and especially clustering can act as a tool when we search for something that we are not aware of, the process can proceed to applying other methods if something worth examining in more detail is found. Interesting findings can be used as input for the next stage, as then we would be aware of what we are trying to find. For example, if interesting safety cases are found, it might be possible to find out whether similar cases exist in the data by using BI tools and applying the parameters found as query criteria. In addition, varying the query parameters is a rather quick method for broadening the search area to cases slightly similar to those initially found.

Before mining all the material in the safety database of the civil aviation authority, it might be relevant to perform another test with production data. Widening the sample to include several thousand reports, possibly exceeding 10,000 flight safety reports, which would allow a much wider view on the content of the data as the sample used here was quite narrow and did not reveal any significant trends. At least using a sample ten times bigger could contain different clusters that could be used to determine mining definitions that, in turn, could be used for the whole database. Big Data (see Chapter 3.3.6) would offer some significant future research questions if safety information could be expanded with general information. For the time being, the data mining tools and systems offer sufficient computing power as the data material consists of collected reports in a specified form, but if the search is to be made using global material, the scale and requirements differ totally. The development of safety questions has recently sent out signals that the origins of the course of lethal events ought to be searched for from a significantly wider amount of information than before. The accidents presented here illustrate the issue and are examples of recent, very exceptional catastrophic events.

Learmount (2015) has discussed two recent flight accidents that were not the results of any ‘conventional’ reason in the airline industry: on March 2014, Malaysia Airlines flight MH370 from Kuala Lumpur to Beijing disappeared from radar tracking and has so far not been found. The same airline suffered another catastrophe only a couple of months later in June when flight MH17 was shot down by a ground-launched, guided missile in eastern Ukraine. In both accidents, all people on board were lost, a total of 537 casualties. These two events together beg two questions: first, if the disappearance of MH370 turns out to be a deliberate act instead of a ‘pure’ accident, it will make the statistics for 2014 dramatically different. Currently, the statistics now show 969 deaths (Learmount, 2015), instead a total of only 432 might be correct. This situation would significantly change the statistics that act as the basis for safety analyses, and consequently misrepresent results. Secondly, the point of emphasis in flight safety discussion has been on safety components and security aspects have received less attention. The main safety and security aspects have been discussed in detail in Chapter 2.1. The question is, would it be necessary to extend the analyses beyond the conventionally collected safety data? As the two events discussed here are, fortunately, extremely rare, no collected material exists to identify them before they happen. If the airline industry is made responsible for foreseeing and preventing deliberately caused hostile acts, new methods and tools will be required. Big data has already been discussed in this chapter. It might provide one possible solution to the problem.

REFERENCES

Aalto, E., Pöllänen, M., Mäntynen, J., Mäkelä, T., & Rauhamäki, H. (2012). *Suomen lentoliikenne vuoteen 2025 - neljä skenaariota*. Trafi's Publications 12/2012, Helsinki: Verne Transport Research Centre, Tampere University of Technology.

Abdel-Hamid, T., & Tarek, K. (1992). Investigating the impacts of managerial turnover/succession on software project performance. *Journal of Management Information Systems, 9*(2), 127-144.

Abdelghany, K., & Abdelghany, A. (2012). *Modeling applications in the airline industry*. Farnham, England: Ashgate Publishing, Ltd.

Abelló, A., & Romero, O. (2009). On-Line Analytical Processing (OLAP). In T. Ozsu & L. Liu (Eds.), *Encyclopedia of Database Systems* (pp. 1949-1954): Springer.

Alavi, M., & Joachimsthaler, E. (1992). Revisiting DSS implementation research: A meta-analysis of the literature and suggestions for researchers. *MIS Quarterly, 16*(1), 95-116.

Amalberti, R. (2001). Safety paradoxes and paradigms in aviation. Air Navigation Services Safety Seminar. Vantaa, Finland: CAA Finland.

Amalberti, R., Auroy, Y., Berwick, D., & Barach, P. (2005). Five System Barriers to Achieving Ultrasafe Health Care. *Annals of Internal Medicine, 142*(9), 756-764.

Antill, L. (1985). Selection of a research method. In E. Mumford, R. A. Hirschheim, G. Fitzgerald & T. Wood-Harper (Eds.), *Research Methods in Information Systems*. Amsterdam: North-Holland.

Argyris, C., Putnam, R., & McLain Smith, D. (1987). *Action Science*: Jossey-Bass, San Francisco.

Arppe, A. (2002). *The usage patterns and selectional preferences of synonyms in a morphologically rich language.* Paper presented at the JADT-2002. 6th International Conference on Textual Data Statistical Analysis.

Assemblée Nationale. (2004). Mission d'information Sécurité du transport aérien de voyageurs. Retrieved 2 April, 2008, from http://www.assemblee-nationale.fr/12/dossiers/securite_transports.asp

Back, B., Kloptchenko, A., Toivonen, J., Vanharanta, H., & Visa, A. (2002). *Prototype-matching Methodology Applications in Text Mining*. Paper presented at the International Conference on Information and Knowledge Engineering.

Baeza-Yates, R., & Ribeiro-Neto, B. (1999). *Modern Information Retrieval*. New York: ACM Press.

Ben Ayed, M., Ltifi, H., Kolski, C., & Alimi, A. M. (2010). A user-centered approach for the design and implementation of KDD-based DSS: A case study in the healthcare domain. *Decision Support Systems, 50*(1), 64-78.

BenMoussa, C. (2008). Applying the Action Research Methodology. *Merits and Challenges*: Seminar presentation November 13. Institute for Advanced Management Systems Research, Turku, Finland.

Berry, M. J. A., & Linoff, G. S. (2000). *Mastering Data Mining*. New York.: John Wiley & Sons, Inc.

Bloedorn, E. (2000). *Mining Aviation safety Data: A Hybrid Approach*: The MITRE Corporation.

Boeing Commercial Airplanes. (2012). *Statistical Summary of Commercial Jet Airplane Accidents Worldwide Operations (1959–2011)*. Seattle, Washington, U.S.A.

Boeing Commercial Airplanes. (2014). *Statistical Summary of Commercial Jet Airplane Accidents Worldwide Operations (1959–2013)*. Seattle, Washington, U.S.A.

Brooker, P. (2006). *Air Traffic Control Safety Indicators: What is Achievable?* Safety R&D Seminar, Barcelona, Spain: Eurocontrol.

Brookes, C. H. (1994). A Framework for DSS Development. In P. Gray (Ed.), *Decision Support and Executive Information System*. Englewood Cliffs, NJ: Prentice Hall.

Brütting, G. (1938). Geschichte der Segelflugbewegung. In W. Hirth (Ed.), *Handbuch des Segelfliegens*. Stuttgart: Franckh'sche Verlagshandlung, W. Keller u. Co.

Carayon, P. (2006). Human factors of complex sociotechnical systems. *Applied Ergonomics, 37*(4), 525-535.

Carlsson, C. (2009). Decision Support Theory: Presentation in Post-Graduate Seminar, Institute for Advanced Management Systems Research, Turku.

Casas, J., Tordaya, A., Perarnaua, J., Breena, M., & Ruiz de Villa, A. (2014). *Decision Support Systems (DSS) for traffic management assessment: Notes on current methodology and future requirements for the implementation of a DSS.* Paper presented at the Transport Research Arena 2014, Paris La Défense, France.

Chaplin, J. (2004). *Aviation - Learning From Experience* (Section of a report). London: Royal Academy of Engineering.

Chaudhuri, S., & Dayal, U. (2002). Multidimensional Databases and Online Analytical Processing. In W. Klösgen & J. M. Zytkow (Eds.), *Handbook of Data Mining and Knowledge Discovery*. New York: Oxford University Press, Inc.

Civil Aviation Administration Finland. (1997). *Taking to the Air. A brief history of civil aviation administration in Finland*. Vantaa, Finland: CAA Finland.

Civil Aviation Administration Finland. (1999). *Air Traffic Statistics 1998*. Vantaa, Finland: CAA Finland.

Clark, P. A. (1972). *Action Research and Organizational Change*. London: Harper and Row.

Čokorilo, O. (2011). Aircraft Performance: The Effects of the Multi Attribute Decision Making of Non Time Dependant Maintainability Parameters. *International Journal for Traffic and Transport Engineering, 1*(1), 42–48.

Čokorilo, O., De Luca, M., & Dell'Acqua, G. (2014). Aircraft safety analysis using clustering algorithms. *Journal of Risk Research, 17*(10), 1-16.

Čokorilo, O., & Dell'Acqua, G. (2013). *Aviation Hazards Identification Using Safety Management System (SMS) Techniques.* Paper presented at the 16th International conference on transport science (ICTS 2013), Portorož, Slovenia.

Colorado State University. (1993-2009). Case Studies. *Writing Guides* Retrieved 4 September, 2009, from http://writing.colostate.edu/guides/research/casestudy/com2a1.cfm

Confederation of Finnish Industries (EK). (2014). Enterprise Safety and Security. Retrieved 17 April, 2014, from http://ek.fi/mita-teemme/tyoelama/yritysturvallisuus/

Continental Airlines. (2004). Maintenance Threat and Error Management (MTEM). Houston, Texas: Continental Airlines Maintenance Training.

Delen, D., & Sharda, R. (2008). Artificial Neural Networks in Decision Support Systems *Handbook on Decision Support Systems 1*: Springer Berlin Heidelberg.

Denning, P. J. (1997). A New Social Contract for Research. *Communications of the ACM, 40*(2), 132-134.

Denzin, N., & Lincoln, Y. (Eds.). (1994). *Handbook of Qualitative Research*. Thousand Oaks, Ca: Sage.

Department of Health. (2005). Chapter 3: Learning to fly: drawing parallels between aviation safety and patient safety. *Chief Medical Officer annual reports* Retrieved 30 November, 2007, from http://www.dh.gov.uk/dr_consum_dh/idcplg?IdcService=GET_FILE&dID=25890&Rendition=Web

Dörre, J., Gerstl, P., & Seiffert, R. (1999). *Text Mining: Finding Nuggets in Mountains of Textual Data.* Paper presented at the KDD-99, Fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Diego, California, USA.

Eskola, J., & Suoranta, J. (2003). *Johdatus laadulliseen tutkimukseen*. Tampere: Osuuskunta Vastapaino.

European Aviation Safety Agency. (2014). *Annual Safety Review 2013*. Cologne, Germany.

European Commission. (2000). *Proposal for a Directive of the European Parliament and of the Council On occurrence reporting in civil aviation*. Brussels.

European Commission. (2003). *Aviation safety: Occurrence reporting in civil aviation. Directive of the European Parliament and of the Council of 13 June 2013 on occurrence reporting in civil aviation 2003/42/EC.*

Ewusi-Mensah, K., & Przasnyski, Z. (1991). On Information Systems Project Abandonment: An Exploratory Study of Organizational Practices. *MIS Quarterly, 15*(1), 66-85.

Fayyad, U. (1996). Data Mining and Knowledge Discovery: Making Sense Out of Data. *IEEE Expert, 11*(5), 20-25.

Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996a). From Data Mining to Knowledge Discovery in Databases. *AI Magazine, 17*(3), 37-54.

Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996b). *Knowledge Discovery and Data Mining: Towards a Unifying Framework.* Paper presented at the Second International Conference on Knowledge History and Data Mining (KDD-96), Portland, Oregon.

Federal Aviation Administration (FAA). (2006). Cockpit voice recorder equipment, *Technical Standard Order TSO-C123b*. Washington, D.C.: Aircraft Certification Service.

Federal Aviation Administration (FAA). (2013). Air Transportation Oversight System (ATOS), *14 CFR Part 121 Air Carrier Certification*.

Fédération Aéronautique Internationale. (2014). History. Retrieved 9 October, 2014, from http://www.fai.org/about-fai/history

Finavia. (2009). *Finland Air Traffic Statistics 2008*. Vantaa, Finland: Finavia Corporation.

Finavia. (2014). *2013 Finavia's air traffic statistics*. Vantaa, Finland: Finavia Corporation.

Finnish Civil Aviation Authority. (2006). Accident, serious incident and occurrence report form LU3626e, version 08/06.

Flight Safety Authority. (1999). *Annual Report*. Helsinki, Finland: Finnish Civil Aviation Administration.

Forrest, J. S. (2006). *Information policies and practices of knowledge management (KM) as related to the development of the Global Aviation Information Network (GAIN): An applied case study and taxonomy development*: Dissertation. Nova Southeastern University, Ann Arbor, MI, 48106, USA.

Foster, M. (1972, 1 December). An Introduction to the Theory and Practice of Action Research in Work Organizations. *Human Relations, 25*, 529-556.

GAIN Working Group B. (2003). *Guide to Methods & Tools for Airline Flight Safety Analysis* (Second Edition): Global Aviation Information Network.

Galliers, R. D. (1990). *Choosing appropriate information systems research approaches: a revised taxonomy.* Paper presented at the IFIP TC8 WG8.2 Conference, Copenhagen, Denmark.

Galliers, R. D., Marshall, P. H., Pervan, G. P., & Klass, D. J. (1994). DSS Development Within an Information Systems Planning. In P. Gray (Ed.), *Decision Support and Executive Information System*. Englewood Cliffs, NJ: Prentice Hall.

Gey, F. C. (2000). *Information Retrieval: Theory, Application, Evaluation.* Paper presented at the 33rd Hawaii International Conference on System Sciences (HICSS-33), Hawaii.

Gheware, S. D., Kejkar, A. S., & Tondare, S. M. (2014). Data Mining: Task, Tools, Techniques and Applications. *International Journal of Advanced Research in Computer and Communication Engineering, 3*(10), 8095-8098.

Gill, T. G. (1995). Early Expert Systems: Where Are They Now? *MIS Quarterly, 19*(1), 51-81.

Given, L. M. (Ed.). (2008). *The SAGE Encyclopedia of Qualitative Research Methods*: Sage Publications.

Gonzalez-Bernal, J. A. (1999). *Substructure discovery in real world spatio-temporal domains*. The University of Texas at Arlington, Ann Arbor.

Grover, V., Cheon, M. J., & Teng, J. T. C. (1996). The Effect of Service Quality and Partnership on the Outsourcing of Information Systems Functions. *Journal of Management Information Systems, 12*(4), 89-116.

Gupta, B., Nemati, H. R., & Harvey, J. D. (1999). Organizational Factors Affecting Successful Implementation of Decision Support Systems: the Case of Fuel Management System at Delta Air Lines. *Journal of Information Technology Case and Application Research, 1*(3), 4-25.

Hale, A. R., & Hovden, J. (1998). Management and culture: the third age of safety. A review of approaches to organizational aspects of safety, health and environment. In A.-M. Feyer & A. Williamson (Eds.), *Occupational Injury: Risk, Prevention And Intervention*. London: Taylor & Francis.

Han, J., & Kamber, M. (2001). *Data Mining: Concepts and Techniques*: Morgan Kaufmann Publishers.

Haverinen, K. (2014). *Natural Language Processing Resources for Finnish. Corpus Development in the General and Clinical Domains*: Doctoral dissertation. University of Turku, Finland.

Hearst, M. A. (1999). *Untangling text data mining.* Paper presented at the Proceedings of the 37th annual meeting of the Association for Computational Linguistics on Computational Linguistics, College Park, Maryland.

Heinrich, H. W. (1931). *Industrial Accident Prevention: A Scientific Approach*: McGraw-Hill.

Heinrich, H. W., Petersen, D. C., Roos, N. R., & Hazlett, S. (1980). *Industrial Accident Prevention: A Safety Management Approach*: McGraw-Hill.

Hevner, A., & Chatterjee, S. (2010). *Design Research in Information Systems. Theory and Practice*: Springer.

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design Science in Information Systems Research. *MIS Quarterly, 28*(1), 75-105.

Hirsjärvi, S., & Huttunen, J. (1995). *Johdatus kasvatustieteeseen*. Porvoo: WSOY.

Hitt, L. M., & Brynjolfsson, E. (1996). The Effect of Service Quality and Partnership on the Outsourcing of Information Systems Funtcions. *Journal of Management Information Systems, 12*(4), 89-116.

Hollnagel, E. (1991). The phenotype of erroneous actions. In G. R. S. Wier & J. L. Alty (Eds.), *Human-Computer Interaction and Complex Systems*. New York: Academic Press.

Hollnagel, E. (2003). Prolegomenon to cognitive task design. In E. Hollnagel (Ed.), *Handbook of cognitive task design*. Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.

Hollnagel, E. (2004). *Barriers and Accident Prevention*. Aldershot, UK: Ashgate Publishing Ltd.

Honkela, T. (1998). Description of Kohonen's Self-Organizing map, excerpt from "Self-Organising Maps in Natural Language Processing". Retrieved 8 September, 2004, from www.mlab.uiah.fi/~timo/som/thesis-som.html

Hopkins, A. (2006). Studying organisational cultures and their effects on safety. *Safety Science, 44*(10), 875-889.

Horner, B. R., & Benbasat, I. (1996). Measuring the Linkage Between Business and Information Technology Objectives. *MIS Quarterly, 20*(1), 55-81.

Hu, R. F., Yin, G. F., & Tan, Y. (2006). A Hybrid Clustering Algorithm and It's Application. *Journal of Sichuan University: Engineering Science Edition* (5), 156-161.

Hutchins, E., & Klausen, T. (1996). Distributed cognition in an airline cockpit. In Y. Engeström & D. Middleton (Eds.), *Cognition and Communication at Work*. Cambridge: Press Syndicate of the University of Cambridge.

IAEA. (1991). *Safety culture* (Safety Report No. 75-INSAG-4). Vienna: International Atomic Energy Agency.

ICAO. (1987). ADREP 87 Standard for Accident and Incident reporting. Retrieved 1 November, 2004, from http://eccairs-www.jrc.it/Support/Downloads/Files/Documentation/Release34Taxonomy/default.htm

ICAO. (2000). ADREP 2000. Retrieved 1 November, 2004, from http://eccairs-www.jrc.it/Support/Downloads/Files/Documentation/Release40Taxonomy/Default.htm

ICAO. (2001). Annex 13 to the Convention on International Civil Aviation, *Aircraft Accident and Incident Investigation*. ICAO Headquarters, Montreal, Canada: International Civil Aviation Organization.

ICAO. (2004a). Foundation of the International Civil Aviation Organization (ICAO). Retrieved 17 September, 2004, from http://www.icao.int/cgi/goto_m.pl?icao/en/ro/eurnat/history02.htm

ICAO. (2004b). History: The Beginning. Retrieved 17 September, 2004, from http://www.icao.int/cgi/goto_m.pl?icao/en/ro/eurnat/history01.htm

ICAO. (2005). *ATM Perfomance*. Second Meeting of North American, Central American and Caribbean Directors of Civil Aviation: Tegucigalpa, Honduras.

ICAO. (2006). *Attribute Values by Attribute id, ECCAIRS 4.2.6 Data Definition Standard*.

ICAO. (2007). ICAO ADREP 2007. *Presentations.* Retrieved 14 March, 2008, from http://eccairsportal.jrc.it/fileadmin/downloads/scm-2007/id-10/212%20ICAO%20UPDATE.pdf

ICAO. (2008a). Agenda Item 4: Management of safety data and representation. Unpublished Working paper. Accident Investigation and Prevention (AIG) Divisional Meeting (2008).

ICAO. (2008b). ICAOData. Retrieved 4 March, 2008, from http://www.icaodata.com/modules/airlines/traffic/search.aspx

ICAO. (2010). Annex 13 to the Convention on International Civil Aviation. Aircraft Accident and Incident Investigation. Tenth Edition, July 2010. *ICAO International Standards and Recommended Practices* Retrieved 19 December, 2015, from http://www.bazl.admin.ch/experten/regulation/03080/03081/index.html?lang=it&download=NHzLpZeg7t,lnp6I0NTU042l2Z6ln1ah2oZn4Z2qZpnO2Yuq2Z6gpJCDeoB9gWym162epYbg2c_JjKbNoKSn6A--.

ICAO. (2014). Taxonomy. *Accident Investigation Section* Retrieved 15 August, 2014, from http://www.icao.int/safety/airnavigation/AIG/Pages/Taxonomy.aspx

ICAO. (2015). Economic Development of Air Transport. Facts & Figures. Retrieved 21 December, 2015, from http://www.icao.int/sustainability/pages/factsfigures.aspx

Ilmavoimat. (2013). Ilmavoimien lento-onnettomuudet sotilasilmailussa vuodesta 1970. Retrieved 20 January, 2015, from www.puolustusvoimat.fi/wcm/cfe46d0041d0ae3eb923f95056d0aca7/Lento_onettomuudet_sotilasilmailussa.pdf?MOD=AJPERES

Ilmavoimat. (2014). Ilmavoimien tiedotteet 2014. *Ajankohtaista* Retrieved 20 January, 2015, from http://www.puolustusvoimat.fi/fi/Ilmavoimat/Ajankohtaista/?urile=wcm%3Apath%3A/su%20puolustusvoimat.fi/Puolustusvoimat.fi/Ilmavoimat/Ajankohtaista/

Inspectie Verkeer en Waterstaat. (2007). Veiligheidsstatistieken luchtvaart / Civil aviation safety data 1991 - 2005. Uitgave 2007. Retrieved 2 April, 2008, from http://www.nlr-atsi.nl/downloads/Veiligheidsstatistieken_luchtvaart_1991-2005.pdf

Isbell, C. L. (1998). Restructuring Sparse High Dimensional Data for Effective Retrieval *Advances in Neural Information Processing Systems*.

Jain, A. K., Murty, M. N., & Flynn, P. J. (1999). Data clustering: a review. *ACM Computing Surveys (CSUR), 31*(3), 264-323.

Janis, I. (1997). Groupthink. In E. Griffin (Ed.), *A First Look at Communication Theory* (3rd ed.): McGraw-Hill, Inc.

Janis, I., & Mann, L. (1977). *Decision making: A psychological analysis of conflict, choice, and commitment*. New York: Free Press.

Johnson, C. W., & Holloway, C. M. (2007). *A Historical Perspective on Aviation Accident Investigation.* Paper presented at the IET 2nd Conference on System Safety.

Johnston, H. R., & Carrico, S. R. (1988). Developing Capabilities to Use Information Strategically. *MIS Quarterly, 12*(1), 37-48.

Joint Research Centre - European Commission. (2008). ECCAIRS Portal. *The European Coordination Centre for Accident and Incident Reporting Systems* Retrieved 3 April, 2008, from http://eccairs-www.jrc.it/index.php?id=1

Joint Research Centre. (2003). Database. *The European Co-ordination Centre for Aviation Incident Reporting Systems* Retrieved 7 September, 2003, from http://eccairs-www.jrc.it/About/database.htm

Joint Research Centre. (2004). The European Co-ordination Centre for Aviation Incident Reporting Systems. *The European Co-ordination Centre for Aviation Incident Reporting Systems* Retrieved 11 October, 2004, from http://eccairs-www.jrc.it/About/Default.htm

Jones, D. R. (2008). *Mining events from aviation safety data*. Dissertation. University of Utah: University Microfilms International, Ann Arbor, MI, 48106, USA

Joshi, K. (1989). The Measurement of Fairness or Equity Perceptions of Management Information System Users. *MIS Quarterly, 13*(3), 343-358.

Joshi, K. (1991). A Model of Users' Perspective on Change: The Case of Information Systems Technology Implementation. *MIS Quarterly, 15*(2), 228-241.

Joshi, K., & Lauer, T. (1998). Impact of Information Technology on Users' Work Environment: A Case of Computer Aided Design (CAD) System Implementation. *Information and Management, 34*, 349-360.

Jukic, N., Jukic, B., & Malliaris, M. (2008). Online Analytical Processing (OLAP) for Decision Support *Handbook on Decision Support Systems 1*: Springer Berlin Heidelberg.

Jurafsky, D., & Martin, J. H. (2009). *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*: Pearson Prentice Hall.

Järvinen, P. (1999). *On Research Methods*. Tampere: Opinpaja Oy.

Järvinen, P. (2004). *On Research Methods*. Tampere: Opinpaja Oy.

Kaisler, S., Armour, F., & Espinosa, J. A. (2014). *Introduction to Big Data: Challenges, Opportunities, and Realities Minitrack.* Paper presented at the 47th Hawaii International Conference on System Sciences (HICSS), Hawaii.

Kaisler, S., Armour, F., Espinosa, J. A., & Money, W. (2013). *Big Data: Issues and Challenges Moving Forward.* Paper presented at the 46th Hawaii International Conference on System Sciences (HICSS), Hawaii.

Kallman, E. A., & Grillo, J. P. (1996). *Ethical Decision Making and Information Technology* (2nd edition ed.). New York: McGraw-Hill.

Karlsson, F. (1987). *Finnish grammar*. Porvoo: WSOY.

Kasanen, E., Lukka, K., & Siitonen, A. (1993). The Constructive Approach in Management Accounting Research. *Journal of Management Accounting Research, 5*, 243-264.

Keen, P. G. (1989). Value Analysis: Justifying Decision Support Systems. In E. Spraque & H. Watson (Eds.), *Decision Support Systems: Putting Theory Into Practice* (2nd Edition ed.). Englewood Cliffs, NJ: Prentice Hall.

Kemp, M. (2006). *Leonardo da Vinci: Experience, Experiment, and Design*: Princeton University Press.

Kemp, R. (1993). Risikowahrnehmung: Die Bewertung von Risiken durch Experten und Laien – Ein zweckmäßiger Vergleich. In Bayerische Rückversicherung (Ed.), *Risiko ist ein Konstrukt*. München: Knesebek.

Kloptchenko, A. (2003). *Text Mining Based on the Prototype Matching Method*: Doctoral Dissertation. Åbo Akademi University, Turku.

Kloptchenko, A., Back, B., Visa, A., Toivonen, J., & Vanharanta, H. (2002). *Toward Content Based Retrieval from Scientific Text Corpora*. Paper presented at the 2002 IEEE International Conference on Artifical Intelligence Systems (ICAIS 2002).

Kloptchenko, A., Eklund, T., Back, B., Karlsson, J., Vanharanta, H., & Visa, A. (2002). *Combining Data and Text Mining Techniques for Analyzing Financial Reports*. Paper presented at the Eight American Conference on Information Systems (AMCIS).

Knuth, D. E., Morris, J. H., & Pratt, V. R. (1977). Fast pattern matching in strings. *SIAM Journal on Computing, 6*(1), 323-360.

Kohavi, R., Rothleder, N. J., & Simoudis, E. (2002). Emerging trends in business analytics. *Communications of the ACM, 45*(8), 45-48.

Koho, I. (2005). *Sotilasilmailun lentoturvallisuuskertomus 2004* (Annual Flight Safety Report). Jyväskylä, Finland: The Finnish Air Force.

Kohonen, T. (1997). *Self-organizing maps*: Springer-Verlag New York, Inc., Secaucus, NJ, USA.

Kohonen, T. (1999). *WEBSOM*. Helsinki: Helsinki Technological University.

Kolenda, T., & Hansen, L. K. (2002). *Independent Components in Text.* Paper presented at the IEEE Workshop on Neural Networks for Signal Processing XII.

Kollepara, V., & Ananyan, S. (2003). *Flight Safety Data Analysis with PolyAnalyst*™ (Case study).

Korenius, T., Laurikkala, J., Järvelin, K., & Juhola, M. (2004, November 8–13). *Stemming and Lemmatization in the Clustering of Finnish Text Documents.* Paper presented at the CIKM’04, Washington, DC, USA.

Koskenniemi, K. (1983). *Two-level morphology: A general computational model for word-form recognition and production* (Publications 11). Helsinki: Department of General Linguistics.

Koskivaara, E. (2004). *Artifical Neural Networks for Analytical Review in Auditing*: Doctoral dissertation. Turku School of Economics and Business Administration, Turku.

Kou, G., Shi, Y., & Wang, S. (2011). Multiple criteria decision making and decision support systems — Guest editor's introduction. *Decision Support Systems, 51*(2), 247-249.

Kurgan, L., & Musilek, P. (2006). A survey of Knowledge Discovery and Data Mining process models. *The Knowledge Engineering Review, 21*(1), 1-24.

Kwon, T. H., & Zmud, R. W. (1987). Unifying the Fragmented Models of Information Systems Implementation. In R. J. Boland & R. A. Hirschheim (Eds.), *Critical Issues in Information Systems Research*. New York: Wiley.

Lagus, K. (2000). *Text Mining with the WEBSOM*: Doctoral dissertation. Finnish Academies of Technology, Espoo.

Laney, D. (2001). 3D Data Management: Controlling Data Volume, Velocity and Variety. META Group.

Latino, R. J. (2008). The cost and truths of human error. *Patient Safety Monitor, 2008*(1), 5-6.

Latorella, K. A. (1999). *A Safety Index and Method for Flightdeck Evaluation*: NASA Langley Research Center.

Laurikkala, J. (2001). *Knowledge Discovery for Female Urinary Incontinence Expert System.* Doctoral dissertation. Tampere University, Tampere, Finland.

LaValle, S. (2009). *Business Analytics and Optimization for the Intelligent Enterprise* (Executive Report). Somers, New York, USA: IBM Institute for Business Value.

Learmount, D. (2015). Analysis: Accidents are not the only threat to safety. *News Aviation* Retrieved 9 February, 2015, from http://www.flightglobal.com/news/articles/analysis-accidents-are-not-the-only-threat-to-safety-408117/?cmpid=NLC|FGFG|FGTSE-2015-0206-GLOB|news&sfid=70120000000taAp

Lecornu, J. L. (1903). *La Navigation aérienne: histoire documentaire et anecdotique*: Nony.

Lederer, A. L., & Sethi, V. (1988). The Implementation of Information Systems Planning Methodologies. *MIS Quarterly, 12*(3), 445-462.

Lederer, A. L., & Sethi, V. (1992). Root Causes of Strategic Information Systems Planning Implementation Problems. *Journal of Management Information Systems, 9*(1), 25-46.

Lee, J., & Kim, Y. (1999). Effect of Partnership Quality on IS Outsourcing Success: Conceptual Framework and Empirical Validation. *Journal of Management Information Systems, 15*(4), 29-62.

Lento, R. (1979). *Sellaista oli elämä vuosisadan vaihteen Turussa*. Juva.

Leveson, N. G. (2011). Applying systems thinking to analyze and learn from events. *Safety Science, 49*(1), 55-64.

Li, Y., Chung, S. M., & Holt, J. D. (2008). Text document clustering based on frequent word meaning sequences. *Data & Knowledge Engineering, 64*(1), 381-404.

Lingsoft Inc. (2009). FinTwol (Finnish Morphological Analyser). Retrieved 18 September, 2009, from http://www.lingsoft.fi/cgi-bin/fintwol/

Lingsoft Oy. (2005). Lingsoft and TEMIS Announce Partnership to Expand Text Mining Coverage to Northern European languages and countries. *News* Retrieved 3 May, 2005, from http://www.lingsoft.fi/news/2005/temis.html

Luo, Y., & Fu, H. (2014). *Modified K-means Algorithm for Clustering Analysis of Hainan Green Tangerine Peel.* Paper presented at the 13th IFIP WG 6.11 Conference on e-Business, e-Services and e-Society, I3E 2014, Sanya, China.

Lyall, D. (2015). The World's Airlines. Retrieved May 15, 2015, from http://www.airlinehistory.co.uk/

Lybeck, T. (2002). Inhimillinen virhe ilmailussa uusi näkökulma. Helsinki: Accident Investigation Board Finland.

Lyytinen, K. (1987). *A Taxonomic Perspective of Information Systems Development: Theoretical Constructs and Recommendations*: John Wiley & Sons Ltd.

Macmillan. (1948). *Report of the Court investigation of the accident to the Tudor IV. Aircraft "Star Tiger" G-AHNP on the 30th January, 1948, held under the Air Navigation.* London: Ministry of Civil Aviation.

Macrae, C. (2007). *Analyzing Near-Miss Events: Risk Management in Incident Reporting and Investigation Systems* (Discussion Paper). London: centre for analysis of risk and regulation.

Magal, S., Carr, H., & Watson, H. (1988). Critical Success Factors for Information Center Managers. *MIS Quarterly, 12*(3), 412-425.

Mallach, E. (1993). *Understanding Decision Support Systems and Expert Systems*: Irwin McGraw-Hill Companies.

Mannila, H. (2002). Using Relational Databases in KDD. In W. Klösgen & J. M. Zytkow (Eds.), *Handbook of Data Mining and Knowledge Discovery*. New York: Oxford University Press, Inc.

Manning, C. D., & Schütze, H. (1999). *Foundations of Statistical Natural Language Processing*. Cambridge, Massachusetts: The MIT Press.

March, S. T., & Smith, G. F. (1995). Design and Natural Science Research on Information Technology. *Decision Support Systems, 15*(4), 251-266.

Markus, M. L. (1983). Power, politics, and MIS implementation. *Communications of the ACM, 26*(6), 430-444.

Masefield, P. G. (1982). *To Ride The Storm: The Story of the Airship R.101*. London: William Kimber.

Matthews, B., Das, S., Bhaduri, K., Das, K., Martin, R., & Oza, N. (2013). Discovering Anomalous Aviation Safety Events Using Scalable Data Mining Algorithms. *Journal of Aerospace Information Systems, 10*(10), 467-475.

Mayes, M., Brewes, B., & Thompson, W. (2002). Introduction to Text Mining and SAS Text Miner *Distilling Textual Data for Competitive Business Advantage*. Heidelberg, Germany.

McFarland, M. W., & Renstrom, A. G. (1950). The Papers of Wilbur and Orville Wright. *Quarterly Journal of Current Acquisitions, 7*(4), 22-34.

McGregor, D. (2006). *The Human Side of Enterprise* (Annotated ed.): McGraw-Hill Companies, Inc.

Megaputer Intelligence Inc. (2004a). *Application of PolyAnalyst to Flight Safety Data of Southwest Airlines*. Bloomington, IN: Megaputer Intelligence.

Megaputer Intelligence Inc. (2004b). Customers. Retrieved 17 December, 2004, from http://www.megaputer.com/company/customers.php3

Megaputer Intelligence Inc. (2004c). Flight safety data analysis for Southwest Airlines. Retrieved 17 December, 2004, from http://www.megaputer.com/company/cases/southwest.php3

Megaputer Intelligence Inc. (2015a). PolyAnalyst. Retrieved 30 January, 2015, from http://megaputer.com/site/polyanalyst.php

Megaputer Intelligence Inc. (2015b). Text Mining Technology. Retrieved 30 January, 2015, from http://www.megaputer.com/site/text_mining.php

Mehta, A. (2011). *Big Data: Powering the Next Industrial Revolution*.

Melinda. (2015). Union Catalogue of Finnish Libraries. Retrieved 12 May, 2015, from http://melinda.kansalliskirjasto.fi/F/?func=find-a-0&CON_LNG=fin&local_base=fin01_opac

Meredith, J. R. (1981). The implementation of computer based systems. *Journal of Operations Management, 2*(1), 11-21.

Merriam-Webster, I. (2014). Online Dictionary. Retrieved 11 April, 2014, from www.merriam-webster.com/dictionary

Michalsen, A. (2003). Risk assessment and perception. *Injury Control and Safety Promotion, 10*(4), 201-204.

Miner, G., Elder, J., Fast, A., Hill, T., Nisbet, R., Delen, D., et al. (2012). *Practical Text Mining and Statistical Analysis for Non-structured Text Data Applications*: Elsevier Inc.

Minzberg, H. (1990). The Manager's Job: Folklore and Fact. *Harvard Business Review, March-April 1990*, 49-61.

Mittman, B. S., & Moore, J. H. (1989). Senior Management Computer Use: Implications for DSS Designs and Goals. In E. Sprague & H. Watson (Eds.), *Decision Support Systems: Putting Theory into Practice*. Englewood Cliffs, NJ: Prentice Hall.

Moxon, B. (1996). Defining Data Mining. *DBMS Data Warehouse Supplement* Retrieved 12 April, 2014, from http://dbmsmag.com/9608d53.html

Muir, A. (2004). *Fundamentals of Data and Text Mining.* Paper presented at the Seventh GAIN World Conference, Montreal, Canada.

Müller, O. (2013). Big Data Analytics as a Strategy of Inquiry in Information Systems Research. Unpublished lecture presentation. University of Liechtenstein.

NACA. (1928). Aircraft Accidents: Method of Analysis, *Report No. 308*. Washington, D.C.: National Advisory Committee for Aeronautics.

NASA. (2002). *Celebrating a Century of Flight* (Vol. 2012). Washington, DC 20546: National Aeronautics and Space Administration, Office of Aerospace Technology.

NASA. (2004, 17 December 2004). ASRS Database Report Sets. *Aviation Safety Reporting System* Retrieved 18 October, 2004, from http://asrs.arc.nasa.gov/report_sets_nf.htm

NASA. (2011). *Lessons from the Challenger Launch Decision Additional Resources*: National Aeronautics and Space Administration, NASA Case Study Resources.

National Transportation Safety Board. (2008). Aviation Accident Statistics. *Aviation* Retrieved 3 March, 2008, from http://www.ntsb.gov/aviation/Stats.htm

National Transportation Safety Board. (2015). Aviation Statistics for 2014. *Investigations, Data & Stats, Aviation Statistics* Retrieved 7 January, 2015, from http://www.ntsb.gov/investigations/data/Pages/aviation_stats.aspx

Nazeri, Z. (2003). *Application of Aviation Safety Data Mining Workbench at American Airlines. Proof-of-Concept Demonstration of Data and Text Mining.* McLean, Virginia, US: Center for Advanced Aviation Systems Development, MITRE Corporation Inc.

Nazeri, Z. (2007). *Cross-database analysis to identify relationships between aircraft accidents and incidents*: Dissertation. George Mason University, Ann Arbor, MI, 48106, USA.

Nazeri, Z., Bloedorn, E., & Ostwald, P. (2001, June). *Experiences in Mining Aviation Safety Data.* Paper presented at the SIGMOD '01, Santa Barbara, CA, USA.

Newman, M., & Sabherwal, R. (1996). Determinants of Commitment to Information Systems Development: A Longitudinal Investigation. *MIS Quarterly*, *20*(1), 23-54.

Norros, L. (1996). System disturbances as springboard for development of operators' expertise. In Y. Engeström & D. Middleton (Eds.), *Cognition and Communication at Work*. Cambridge: Press Syndicate of the University of Cambridge.

Novotny, T. (2007). Do qualitative and proactive safety indicators better control risk than the usual quantitative and reactive ones? Unpublished Masters thesis. University of Lund.

Oard, D. W., & Marchionini, G. (1996). *A conceptual framework for text filtering* (Technical Report CS-TR3643): University of Maryland.

OLAP Council. (1997). OLAP and OLAP Server Definitions. Retrieved 4 March, 2015, from http://www.olapcouncil.org/research/glossaryly.htm

Open Access Theses and Dissertations. (2015). Retrieved 4 June, 2015, from http://oatd.org/

Orasanu, J., & Lynne, M. (1998, April 1-2). *Errors in Aviation Decision Making: a Factor in Accidents and Incidents.* . Paper presented at the HESSD '98 2nd Workshop on Human Error, Safety, and System Development, Seattle, Washington, USA.

Oster Jr, C. V., Strong, J. S., & Zorn, C. K. (2013). Analyzing aviation safety: Problems, challenges, opportunities. *Research in Transportation Economics, 43*(1), 148-164.

Oxford University. (2014). *Oxford Dictionaries*: Oxford University Press.

Pachella, R. G. (1974). The interpretation of reaction time in information processing research. In B. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition*. New York: Halstead Press.

Palvia, S. C., & Chervany, N. L. (1995). An Experimental Investigation of Factors Influencing Predicted Success in DSS Implementation. *Information and Management, 29*(1), 43-53.

Parsaye, K. (1997). A Characterization of Data Mining Technologies and Processes. *Journal of Data Warehousing, 2*(3), 2-15.

Paukkeri, M.-S. (2012). *Language- and domain-independent text mining*. Espoo: Doctoral dissertation. Aalto University, School of Science, Department of Information and Computer Science.

Perrow, C. (1999). *Normal Accidents: Living With High-Risk Technologies*. Princeton, NJ: Princeton University Press.

Pidgeon, N. (2004). *Learning from Accidents - Sharing Information and Experience* (Section of a report). London: Royal Academy of Engineering.

Pidgeon, N., & O'Leary, M. (2000). Man-made disasters: why technology and organizations (sometimes) fail. *Safety Science, 34*(1-3), 15-30.

Pihlanto, P. (1993). The Action-oriented approach and Case Study Method in Management Studies. *Scandinavian Journal of Management, 10*(4), 369-382.

PlaneCrashinfo. (2015). Accident statistics. Retrieved 15 January, 2015, from http://www.planecrashinfo.com/

PolyVista Inc. (2005). *"Know" and "Don't know": The Building Blocks of Knowledge* (White paper). Houston, TX.

PolyVista Inc. (2010). *Win-win for Users and IT: Southwest Airlines Applies PolyVista Business Intelligence Solution to Enhance Analysis* (Case Study). Houston, TX.

PolyVista Inc. (2012). *PolyVista New User Guide v2012* (Manual). Houston, TX.

PolyVista Inc. (2013a). *PolyVista Algorithms* (White Paper). Houston, TX.

PolyVista Inc. (2013b). *PolyVista Quick User Guide 3.22* (Manual). Houston, TX.

PolyVista Inc. (2013c). Your one-stop-shop for advanced text analytics. Retrieved 23 May, 2013, from http://polyvista.com/

Pouliquen, Y., Ferrante, O., Jouniaux, P., Nicolas, G., Cabon, P., Rome, F., et al. (2005). *A Human Factors Approach for the Analysis and the Encoding of Aviation Accidents and Incidents: A Validation Study.* Paper presented at the Symposium on Aviation Psychology, Dayton, Ohio.

Power, D. J. (1997). Decision Support Systems Glossary. Retrieved 4 March, 2015, from http://www.dssresources.com/glossary/olaptrms.html

ProQuest. (2015). Retrieved 4 June, 2015, from http://www.proquest.com

Ramasamy, S., & Sabatini, R. (2015). *Communication, Navigation and Surveillance Performance Criteria for Safety-Critical Avionics and ATM Systems*. Paper presented at the 16th Australian International Aerospace Congress (AIAC16).

Rasmussen, J. (1982). Human errors. A taxonomy for describing human malfunction in industrial installations. *Journal of Occupational Accidents, 4*(2–4), 311-333.

Reason, J. T. (1990). *Human error*. New York: Cambridge University Press.

Reason, J. T. (2008). *The human contribution. Unsafe acts, accidents and heroic recoveries*. Farnham, UK: Ashgate Publishing Limited.

Reiman, T., & Oedewald, P. (2009). *Evaluating safety-critical organizations - emphasis on the nuclear industry* (Report No. 2009:12): Swedish Radiation Safety Authority.

Reiman, T., Pietikäinen, E., & Oedewald, P. (2008). *Turvallisuuskulttuuri: Teoria ja arviointi* (Report No. 700). Espoo: VTT Technical Research Centre of Finland.

Reponen, T. (1993). Information Management Strategy - An Evolutionary Process. *Scandinavian Journal of Management, 9*(3), 189 - 209.

Reponen, T. (2004). Action Research – A Link Between Academia and Practice. Turku: Finnish-Italian Workshop on Information Systems.

Research and Innovative Technology Administration. (2011). Air Carrier Traffic Statistics. Retrieved 6 May, 2011, from http://www.bts.gov/programs/airline_information/air_carrier_traffic_statistics/

Riloff, E., & Hollaar, L. (1996). Text Databases and Information Retrieval. *ACM Computing Surveys, 28*(1), 133-134.

Robb, D. (2005). Mining Data to Up Airline Safety. *Datamation, February 2005*.

Roberto, M. A., Bohmer, R. M. J., & Edmondson, A. C. (2006). Facing ambiguous threats. *Harvard Business Review, November 2006*, 106-113.

Rogers, W. H., Schutte, P. C., & Latorella, K. A. (1996). Fault management in aviation systems. In R. Parasuraman & M. Mouloua (Eds.), *Automation and Human performance: Theory and Applications*. Mahwah, NJ: Erlbaum Association.

Rosell, M. (2009). *Text Clustering Exploration. Swedish Text Representation and Clustering Results Unraveled.* Doctoral dissertation. Kungliga Tekniska Högskolan, Stockholm.

Royal Academy of Engineering. (2004). *Learning from Accidents - Sharing Information and Experiance*. London: Royal Academy of Engineering.

Saarikoski, J. (2014). *On text document classification and retrieval using self-organising maps*. Tampere: Doctoral dissertation. Tampere University Press.

Saarinen, L. (2005). Action Research: Unpublished workshop presentation in HSE TuKKK Seminar '05.

Saatsi, J., Haavisto, M.-L., & Oksama, L. (2011). *Inhimillisten tekijöiden välttäminen lentoteknisessä työssä*. Tampere: Juvenesprint Oy.

Sabherwal, R., & Robey, D. (1993). An Empirical Taxonomy of Implementation Processes Based on Sequences of Events in Information System Development. *Organization Science, 4*(4), 548-576.

Sanders, G. L., & Courtney, J. F. (1985). A Field Study of Organizational Factors Influencing DSS Success. *MIS Quarterly, 9*(1), 77-93.

Sanderson, M. (1994). *Word Sence Disambiguation and Information Retrieval.* Paper presented at the The 17th ACM Conference on Research and Development in IR (SIGIR-94).

Saracoglu, R., Tütünkü, K., & Allahverdi, N. (2008). A new approach on search for similar documents with multiple categories using fuzzy clustering. *Expert Systems with Applications: An International Journal, 34*(4), 2545-2554.

Sarlin, P. (2013). *Mapping Financial Stability.* Doctoral dissertation. Åbo Akademi University, Turku.

Sathi, A. (2012). Big Data Analytics: Disruptive Technologies for Changing the Game Available from ftp://public.dhe.ibm.com/software/pdf/at/SWP10/Big_Data_Analytics.pdf

Schroeck, M., Shockley, R., Smart, J., Romero-Morales, D., & Tufano, P. (2012). *Analytics: The real-world use of big data* (Executive Report). Somers, New York, USA: IBM Institute for Business Value.

Seton Hall University. (2012). *The Space Shuttle Challenger Disaster: A Study in Organizational Ethics*: Seton Hall University Department of Communication.

Sheridan, T. B. (2008). Risk, human error, and system resilience: fundamental ideas. *Human Factors, 50*(3), 418-426.

Shim, J. P., Warkentin, M., Courtney, J. F., Power, D. J., Sharda, R., & Carlsson, C. (2002). Past, present, and future of decision support technology. *Decision Support Systems, 33*(2), 111-126.

Silver, M. S., Markus, M. M., & Beath, C. M. (1995). The Information Technology Interaction Model: A Foundation for the MBA Core Course. *MIS Quarterly, 19*(3), 361-390.

Silverman, D. (2000). *Doing Qualitative Research. A Practical Handbook*. London: SAGE Publications Ltd.

Simon, H. A. (1960). *The New Science of Management Decisions*. New York, NY: Harper and Row.

Simon, H. A. (1996). *The Sciences of the Artificial* (3rd ed.). Cambridge, Maine: MIT Press.

SKYbrary. (2013a). ICAO ADREP. *Enhancing Safety* Retrieved 4 August, 2014, from http://www.skybrary.aero/index.php/ICAO_ADREP

SKYbrary. (2013b). Occurrence Class Taxonomy. *ICAO ADREP* Retrieved 4 August, 2014, from http://www.skybrary.aero/index.php/Occurrence_Class_Taxonomy

SKYbrary. (2014). Flight Data Recorder (FDR). *Flight Technical* Retrieved 5 May, 2015, from http://www.skybrary.aero/index.php/Flight_Data_Recorder_%28FDR%29

Sorsa, M. (2012). Human Factors. Lecture in seminar, Vantaa, Finland: Finnish Aeronautical Association (FAA).

Sprague, R. H., & Watson, H. J. (1996). *Decision Support for Management*. Upper Saddle River, NJ: Prentice Hall.

Suchman, L. A. (1996). Constituting shared workplaces. In Y. Engeström & D. Middleton (Eds.), *Cognition and Communication at Work*. Cambridge: Press Syndicate of the University of Cambridge.

Szeto, W. Y., & Lo, H. K. (2005). Strategies for Road Network Design Over Time: Robustness Under Uncertainty. *Transportmetrica, 1*(1), 47-63.

Tait, P., & Vessey, I. (1988). The Effect of user Involvement on System Success. *MIS Quarterly, 12*(1), 91-108.

Tan, A. (1999). *Text Mining: The state of art and the challenges.* Paper presented at the Knowledge Discovery from Advanced Databases (KDAD'99), Beijing, China.

TEMIS. (2005). Text Mining Solutions, TEMIS. Retrieved 2 May, 2005, from http://www.temis-group.com

The Universal Manager. (2005). Theories on Common Causes of Accidents. *The Reading Room, Managing Health and Safety* Retrieved 5 April, 2005, from http://www.universal-manager.co.uk/ReadingRoom/theoriesacc.htm

Timonen, M. (2013). *Term Weighting in Short Documents for Document Categorization, Keyword Extraction and Query Expansion*. Doctoral dissertation: University of Helsinki.

Toivonen, J. (2006). *Text Mining as a Tool in Real-life Tasks*: Doctoral dissertation. Tampere University of Technology.

Toivonen, J., Visa, A., Vesanen, T., Back, B., & Vanharanta, H. (2001). *Prototype Based Information Retrieval in Multilanguage Bibles.* Paper presented at the 3rd European Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS 2001), Tampere, Finland.

TraFi. (2011). Statistics about flight accidents. *Statistics - Aviation* Retrieved 6 February, 2012, from http://www.trafi.fi/filebank/a/1322166733/e2119e8e45f2d26d30fb0a0ff049e87b/1643-Onnettomuuksien_ja_vaurioiden_lukumaara.pdf

TraFi. (2014a). Experience on ECCAIRS at the Finnish NSA. Retrieved 8 January, 2015, from www.era.europa.eu/Document-Register/Documents/TRAFI.pdf

TraFi. (2014b). Lentotunnit 2005-2013. Retrieved 5 February, 2015, from http://www.trafi.fi/tietopalvelut/tilastot/ilmailu/lentotoiminta/lentotunnit

TraFi. (2014c). Suomen ilmailun tila 2014. Turvallisuus ja ympäristövaikutukset: Trafin julkaisuja 10/2014.

Tryon, R. C. (1939). *Cluster Analysis: Correlation Profile and Orthometric (Factor) Analysis for the Isolation of Unities in Mind and Personality*. Ann Arbor, Michigan: Edwards Brothers.

Tsichritzis, D. (1998). The Dynamics of Innovation. In P. J. Denning & R. M. Metcalfe (Eds.), *Beyond Calculation: The Next Fifty Years of Computing*. New York: Copernicus Books.

Turban, E., & Aronson, J. (1998). *Decision Support Systems and Intelligent Systems*: Prentice-Hall, Inc.

Turner, B. A. (1978). *Man-Made Disasters*. London: Wykeham Science Press.

United States Department of Transportation. (2014). U.S. Air Carrier Traffic Statistics. Retrieved 7 January, 2015, from http://www.rita.dot.gov/bts/acts

Uola, M. (1994). *Suomen Ilmailuliitto 75 vuotta 1919 - 1994*. Forssa, Finland: Suomen Ilmailuliitto - Finlands Flygförbund ry.

Waikar, A., & Nichols, P. (1997). Aviation safety: a quality perspective. *Disaster Prevention and Management: An International Journal, 6*(2), 87-93.

Walker, W. E., Harremoës, P., Rotmans, J., van der Sluijs, J. P., van Asselt, M. B. A., Janssen, P., et al. (2003). Defining Uncertainty: A Conceptual Basis for Uncertainty Management in Model-Based Decision Support. *Integrated Assessment, 4*(1), 5-17.

Valusek, J. R. (1994). Adaptive design of DSSs: A User Perspective. In P. Gray (Ed.), *Decision Support and Executive Information System*. Englewood Cliffs, NJ: Prentice Hall.

Waßermann, M. (1983). *Otto Lilienthal* (1. ed. Vol. Band 81). Leipzig: BSB B. G. Teubner Verlagsgesellschaft.

Watson, A., Dolislager, F., Hall, L., Raber, E., Hauschild, V. D., & Love, A. H. (2011). Developing Health-Based Pre-Planning Clearance Goals for Airport Remediation Following a Chemical Terrorist Attack: Decision Criteria for Multipathway Exposure Routes. *Human and Ecological Risk Assessment, 17*(1), 57-121.

Watson, R. (1999). *Data Management: Databases and Organizations* (2nd ed.): John Wiley & Sons.

Weick, K. E., & Sutcliffe, K. M. (2007). *Managing the Unexpected: Resilient Performance in an Age of Uncertainty* (2nd ed.): John Wiley & Sons Inc.

Weill, P. (1992). The Relationship between Investment in Information Technology and Firm Performance: a Study of the Valve Manufacturing Sector. *Information Systems Research, 3*(4), 307-333.

Wells, A. T., & Rodrigues, C. C. (2003). *Commercial Aviation Safety*. New York: McGraw Hill.

Venable, J. (2009a). Evaluation in DSR. Lecture presentation in Design Science Research Evaluation: Helsinki School of Economics.

Venable, J. (2009b). Overview, risk assessment, problem analysis for DSR. Lecture presentation in Design Science Research Evaluation: Helsinki School of Economics.

Vesanen, T. (2003). *Example based search of text documents.* Master of Science Thesis.Tampere University of Technology, Tampere.

Wickens, C. D. (1995). *Integration of Navigational Information for Flight* (No. ARL-95-11/NASA-95-5). Urbana-Champaign: University of Illinois Aviation Research Lab Technical Report.

Wiener, E. L. (1977). Controlled Flight into Terrain Accidents: System-Induced Errors. *Human Factors, 19*(2), 171-181.

Wiener, E. L. (1980). Midair Collisions: The Accidents, the Systems, and the Realpolitik. *Human Factors*, *22*(5), 521-533.

Visa, A., Toivonen, J., Autio, S., Mäkinen, J., Back, B., & Vanharanta, H. (2001). *Data mining of text as a tool in authorship attribution.* Paper presented at the Data Mining and Knowledge Discovery: Theory, Tools and Technology III, Orlando, USA.

Wood-Harper, T. (1985). Research methods in information systems: using action research. In E. Mumford, R. A. Hirschheim, G. Fitzgerald & T. Wood-Harper (Eds.), *Research Methods in Information Systems*. Amsterdam: North-Holland.

Yaverbaum, G. (1988). Critical Factors in the User Environment: An Experimental Study of Users, Organisations and Tasks. *MIS Quarterly, 12*(1), 7-88.

Yeh, Y., & Wickens, C. D. (1988). Dissociation of performance and subjective measures of workload. *Human Factors, 30*, 111-120.

Zadeh, L. A. (1979). A theory of approximate reasoning. In J. Hayes, D. Michie & L. I. Mikulich (Eds.), *Machine Intelligence*. New York: Halstead Press.

Zicari, R. V. (2012). On Analyzing Unstructured Data. — Interview with Michael Brands. *ODBMS Industry Watch* Retrieved 27 January, 2015, from http://www.odbms.org/blog/2012/07/on-analyzing-unstructured-data-interview-with-michael-brands/

APPENDIX 1

International Standards and Recommended Practices, Chapter 1. Definitions - Annex 13 to the Convention on International Civil Aviation: Aircraft Accident and Incident Investigation (reproduced with the permission of ICAO, granted 27 April 2016, subject to the conditions established in the ICAO Publications Regulations).

INTERNATIONAL STANDARDS
AND RECOMMENDED PRACTICES

CHAPTER 1. DEFINITIONS

When the following terms are used in the Standards and Recommended Practices for Aircraft Accident and Incident Investigation, they have the following meanings:

Accident. An occurrence associated with the operation of an aircraft which, in the case of a manned aircraft, takes place between the time any person boards the aircraft with the intention of flight until such time as all such persons have disembarked, or in the case of an unmanned aircraft, takes place between the time the aircraft is ready to move with the purpose of flight until such time as it comes to rest at the end of the flight and the primary propulsion system is shut down, in which:

a) a person is fatally or seriously injured as a result of:

— being in the aircraft, or
— direct contact with any part of the aircraft, including parts which have become detached from the aircraft, or
— direct exposure to jet blast,

except when the injuries are from natural causes, self-inflicted or inflicted by other persons, or when the injuries are to stowaways hiding outside the areas normally available to the passengers and crew; or

b) the aircraft sustains damage or structural failure which:

— adversely affects the structural strength, performance or flight characteristics of the aircraft, and
— would normally require major repair or replacement of the affected component,

except for engine failure or damage, when the damage is limited to a single engine (including its cowlings or accessories), to propellers, wing tips, antennas, probes, vanes, tires, brakes, wheels, fairings, panels, landing gear doors, windscreens, the aircraft skin (such as small dents or puncture holes), or for minor damages to main rotor blades, tail rotor blades, landing gear, and those resulting from hail or bird strike (including holes in the radome); or

c) the aircraft is missing or is completely inaccessible.

Note 1.— For statistical uniformity only, an injury resulting in death within thirty days of the date of the accident is classified, by ICAO, as a fatal injury.

Note 2.— An aircraft is considered to be missing when the official search has been terminated and the wreckage has not been located.

Note 3.— The type of unmanned aircraft system to be investigated is addressed in 5.1.

Note 4.— Guidance for the determination of aircraft damage can be found in Attachment F.

Accredited representative. A person designated by a State, on the basis of his or her qualifications, for the purpose of participating in an investigation conducted by another State. Where the State has established an accident investigation authority, the designated accredited representative would normally be from that authority.

Adviser. A person appointed by a State, on the basis of his or her qualifications, for the purpose of assisting its accredited representative in an investigation.

Aircraft. Any machine that can derive support in the atmosphere from the reactions of the air other than the reactions of the air against the earth's surface.

Causes. Actions, omissions, events, conditions, or a combination thereof, which led to the accident or incident. The identification of causes does not imply the assignment of fault or the determination of administrative, civil or criminal liability.

Contributing factors. Actions, omissions, events, conditions, or a combination thereof, which, if eliminated, avoided or absent, would have reduced the probability of the accident or incident occurring, or mitigated the severity of the consequences of the accident or incident. The identification of contributing factors does not imply the assignment of fault or the determination of administrative, civil or criminal liability.

Flight recorder. Any type of recorder installed in the aircraft for the purpose of complementing accident/incident investigation.

Note.— See Annex 6, Parts I, II and III, for specifications relating to flight recorders.

Incident. An occurrence, other than an accident, associated with the operation of an aircraft which affects or could affect the safety of operation.

Note.— The types of incidents which are of main interest to the International Civil Aviation Organization for accident prevention studies are listed in Attachment C.

Investigation. A process conducted for the purpose of accident prevention which includes the gathering and analysis of information, the drawing of conclusions, including the determination of causes and/or contributing factors and, when appropriate, the making of safety recommendations.

Investigator-in-charge. A person charged, on the basis of his or her qualifications, with the responsibility for the organization, conduct and control of an investigation.

Note.— Nothing in the above definition is intended to preclude the functions of an investigator-in-charge being assigned to a commission or other body.

Maximum mass. Maximum certificated take-off mass.

Operator. A person, organization or enterprise engaged in or offering to engage in an aircraft operation.

Preliminary Report. The communication used for the prompt dissemination of data obtained during the early stages of the investigation.

Safety recommendation. A proposal of an accident investigation authority based on information derived from an investigation, made with the intention of preventing accidents or incidents and which in no case has the purpose of creating a presumption of blame or liability for an accident or incident. In addition to safety recommendations arising from accident and incident investigations, safety recommendations may result from diverse sources, including safety studies.

Serious incident. An incident involving circumstances indicating that there was a high probability of an accident and associated with the operation of an aircraft which, in the case of a manned aircraft, takes place between the time any person boards the aircraft with the intention of flight until such time as all such persons have disembarked, or in the case of an unmanned aircraft, takes place between the time the aircraft is ready to move with the purpose of flight until such time as it comes to rest at the end of the flight and the primary propulsion system is shut down.

Note 1.— The difference between an accident and a serious incident lies only in the result.

Note 2.— Examples of serious incidents can be found in Attachment C.

Serious injury. An injury which is sustained by a person in an accident and which:

a) requires hospitalization for more than 48 hours, commencing within seven days from the date the injury was received; or

b) results in a fracture of any bone (except simple fractures of fingers, toes or nose); or

c) involves lacerations which cause severe haemorrhage, nerve, muscle or tendon damage; or

d) involves injury to any internal organ; or

e) involves second or third degree burns, or any burns affecting more than 5 per cent of the body surface; or

f) involves verified exposure to infectious substances or injurious radiation.

State of Design. The State having jurisdiction over the organization responsible for the type design.

State of Manufacture. The State having jurisdiction over the organization responsible for the final assembly of the aircraft.

State of Occurrence. The State in the territory of which an accident or incident occurs.

State of the Operator. The State in which the operator's principal place of business is located or, if there is no such place of business, the operator's permanent residence.

State of Registry. The State on whose register the aircraft is entered.

Note.— In the case of the registration of aircraft of an international operating agency on other than a national basis, the States constituting the agency are jointly and severally bound to assume the obligations which, under the Chicago Convention, attach to a State of Registry. See, in this regard, the Council Resolution of 14 December 1967 on Nationality and Registration of Aircraft Operated by International Operating Agencies which can be found in Policy and Guidance Material on the Economic Regulation of International Air Transport *(Doc 9587).*

State safety programme (SSP). An integrated set of regulations and activities aimed at improving safety.

APPENDIX 2

Accident, serious incident and occurrence report form TraFi LU3626e.

ILMAILUHALLINTO
LUFTFARTSFÖRVALTNINGEN
FINNISH CIVIL AVIATION AUTHORITY

ACCIDENT, SERIOUS INCIDENT AND OCCURRENCE REPORT FORM

Finnish Civil Aviation Authority:
Fax (09) 4250 2494 (reserve no. 4250 2939)

PLACE AND TIME OF OCCURRENCE

Place of occurrence: (town, village/aerodrome)	Date:	Time (local time):

REPORTER'S PERSONAL DETAILS

Name:	Age:	Street address:	
Telephone:		Post code:	City, Country:

FLIGHT EXPERIENCE (estimate if accurate information not available)

Experience with a/c type mentioned:	Total flight experience:

AIRCRAFT

Registration:	Call sign:	Type of aircraft:

HISTORY OF FLIGHT

Departed from:	Time (local):	Landed at:		Time (local):
Point of intended landing:	No. of crew:	No. of pax:	Operator or club:	Type of operation: ☐ Private ☐ Comm.
Nature of flight:		Occurrence phase of flight:		
Flight Rules: ☐ VFR ☐ IFR	Part of Airspace:	Class of Airspace: ☐ A ☐ B ☐ C ☐ D ☐ E ☐ F ☐ G+ ☐ G		

METEOROLOGICAL INFORMATION

Surface wind: °/ kts, gusts kts			Wind: ☐ Steady ☐ Gusty ☐ Variable		Light conditions: ☐ Daylight ☐ Dusk ☐ Dark
Visibility: km	Clouds:	Temperature: °C	Conditions: ☐ VMC ☐ IMC	Weather: ☐ Icing ☐ Turbulence ☐ Inversion ☐ Mist ☐ Thunder ☐ Clear ☐ Fog	Type of Precipitation: ☐ Rain ☐ Snow ☐ Hail ☐ Light ☐ Moderate ☐ Heavy
Vertical visibility: ft	QNH: hPa	Dewpoint: °C			
Other (e.g. runway conditions):					

BIRD STRIKE INFORMATION

Pilot warned of Birds: ☐ Yes ☐ No	Number of Birds:	Size of Birds: ☐ Small ☐ Medium ☐ Large	Bird species:
Height (AGL): ft	Indicated Air Speed: kt	Landing/taxilights: ☐ Used ☐ Not used	

OCCURRENCE INFORMATION

Occurrence details: (continue overleaf if needed) Attachment 1 ☐
Injuries to persons: Attachment 2 ☐
Damage to aircraft: Attachment 3 ☐
Probable cause: Attachment 4 ☐
Reported by: ☐ Pilot-in-Command ☐ Air Traffic Control officer ☐ AFIS officer ☐ Other:
Date and signature (name in capital letters): / 20

LU3626e 08/06

Further information:

(fold here)

Vastaanottaja maksaa
postimaksun

Port Paye

FINNISH CIVIL
AVIATION AUTHORITY

Investigation and Analysis

Tunnus 5015486

1003 VASTAUSLÄHETYS
FINLAND

(fold here)

When filing a report please note:

- All incidents in which the safety of a flight operation has been endangered or may have been endangered must be reported in accordance with aviation regulation GEN M1-4. If it is unclear whether the incident needs to be reported, it is always preferable to file a report. Any report submitted within an internal quality or safety management system of an air operator will not remove the obligation to file a report in accordance with GEN M1-4.
- Any accident or serious incident must be reported without delay to the appropriate air traffic services unit and to the Accident Investigation Board, Finland. The report to the air traffic services unit must be made on the relevant radio frequency or by telephone, and to the Accident Investigation Board by telephone: (09) 1606 7643. The report to the Finnish Civil Aviation Authority must be made in writing as soon as possible after the incident. Whenever possible, the written report to the Civil Aviation Authority should be submitted by fax: (09) 4250 2494, to ensure rapid transfer of the information.

Staple here

APPENDIX 3

Definitions, Guidance and Usage Notes to ECCAIRS Taxonomy Version 3.4.0.2, Joint Research Centre of the European Commission.

Description	Explanation
Accident	An occurrence associated with the operation of an aircraft which, in the case of a manned aircraft, takes place between the time any person boards the aircraft with the intention of flight until such time as all such persons have disembarked, or in the case of an unmanned aircraft, takes place between the time the aircraft is ready to move with the purpose of flight until such time it comes to rest at the end of the flight and the primary propulsion system is shut down, in which: (a) a person is fatally or seriously injured as a result of: - being in the aircraft, or, - direct contact with any part of the aircraft, including parts which have become detached from the aircraft, or, - direct exposure to jet blast, except when the injuries are from natural causes, self-inflicted or inflicted by other persons, or when the injuries are to stowaways hiding outside the areas normally available to the passengers and crew; or (b) the aircraft sustains damage or structural failure which adversely affects the structural strength, performance or flight characteristics of the aircraft, and would normally require major repair or replacement of the affected component, except for engine failure or damage, when the damage is limited to a single engine, (including its cowlings or accessories), to propellers, wing tips, antennas, probes, vanes, tires, brakes, wheels, fairings, panels, landing gear doors, windscreens, the aircraft skin (such as small dents or puncture holes) or minor damages to main rotor blades, tail rotor blades, landing gear, and those resulting from hail or bird strike, (including holes in the radome); or (c) the aircraft is missing or is completely inaccessible; Note 1. - For statistical uniformity only, an injury resulting in death within thirty days of the date of the accident is classified, by ICAO, as a fatal injury. Note 2. - An aircraft is considered to be missing when the official search has been terminated and the wreckage has not been located. Note 3. - The type of unmanned aircraft system to be investigated is addressed in 5.1 of Annex 13. Note 4. - Guidance for the determination of aircraft damage can be found in Attachment F of Annex 13.
Serious incident	An incident involving circumstances indicating that an accident nearly occurred. Note: - The difference between an accident and a serious incident lies only in the result. - Examples of serious incidents can be found in Attachment C of ICAO Annex 13 and in the ICAO Accident/Incident Reporting Manual (ICAO Doc 9156).

Incident	An occurrence, other than an accident, associated with the operation of an aircraft which affects or could affect the safety of operation. N.B. The type of incidents which are of main interest to the International Civil Aviation Organization for accident prevention studies are listed in the ICAO Accident/Incident Reporting Manual (ICAO Doc 9156) and ICAO Annex 13.
Major incident	An incident in which: - Safety may have been compromised either having lead to a near collision between aircrafts, with the ground or obstacles (i.e. safety margins not respected which is not the result of an ATC instruction). - Assessment of the incident using a risk classification process has identified that this incident could have deteriorated into more serious situation.
Significant incident	An incident in which: - An accident, serious or major incident could have occurred if the risk had not been managed within safety margins (one or more safety barriers remaining). - or if another aircraft or vehicle had been in the vicinity during the incident.
Occurrence without safety effect	Eurocontrol: An incident which has no safety significance. N.B. This appears to be a contradiction with the ICAO definition of an incident: An occurrence, other than an accident, associated with the operation of an aircraft which affects or could affect the safety of operation. A possibly safety related occurrence not meeting the reporting requirements. This could be e.g. the result of downgrading the incident after review.
Observation	The observation of a potential safety issue or hazard that, if not rectified could cause or have caused an incident. The date and time of occurrence for an observation is that when it was first observed for the purposes of reporting and not an assessment of how long the safety issue might have been present.
Occurrence with No Flight Intended	A reportable Occurrence where there was no intention of flight. Example: Substantial damage found on aircraft during maintenance.
Not determined	The class of the occurrence has not been determined.