A Reference Model for Data Safety

Alastair Faulkner * and **Mark Nicholson** †
* *Abbeymeade Limited, UK, email: alastair.faulkner@abbeymeade.co.uk*
† *University of York, UK, email: mark.nicholson@york.ac.uk*

Keywords: Data, safety, method, process.

Abstract

Safety Management has matured from simple process-based arguments to become hazard-focused and proactive. The adoption of Goal Structuring Notation and 'product line' safety arguments means that products comprise multiple constituents in several axis. Data is now ubiquitous in the development, operation and assurance of products. This paper, therefore, explores data safety in the form of a Reference Model that includes three axis. A process that uses this model is discussed. Model validation is by reference to existing safety and assurance models and case studies.

1 Introduction

"Products" (goods and services) pervade all aspects of our lives through interactions with government, healthcare, financial, business and industrial activities. Often these systems rely on data that increasingly determines their behaviour. Where these systems have safety responsibility, their failure may lead directly to harm. Product failures may also contribute to harm indirectly through incorrect decisions made by users who rely on, or trust, these systems.

Advances in architecture have created highly adaptive applications that are data dependent. These increasingly complex and sophisticated solutions deliberately reduce operator involvement and hence their ability to exercise oversight and apply corrective action. In parallel technology has developed; fuelled by increasing telecommunications bandwidth and storage technologies to create large network based data storage. This storage capability changes the nature and uses of data, as well as the impact of data error. Data growth is exponential; many datasets refer to other data, often across organisational domains and system boundaries.

Data errors may significantly influence the safety behaviour of a product. This means that data error modes and the consequences of such errors must be studied. Data must be developed, managed and maintained to an appropriate level of quality. If Data-Centric technology is to be effectively exploited those reasoning about the safety, and security, of systems must have guidance on the safety aspects of data.

Data is important, and the problems associated with data are real [2]. The sheer quantity of data and variation in use of data that has a potential contribution to the safety of a product is staggering. Current proposed guidance [3] concentrates on systems under control and still identifies 25 different types of data. If the scope is widened to activities of the data-centric organisation [4] an even larger set of data types, and usages and therefore potential impact on the safety and security of operations can be identified.

This paper introduces a model that provides a means of directing, and targeting data safety effort, that is appropriate to the situation at hand. It is compatible with existing assurance models, but allows their extension into the data safety arena. The model addresses a key issue in the debate on data safety; that of tailoring activities to a relevant scenario. In the next section, an overview of the model and its elements are presented. The paper goes on to show the model's compatibility with existing models. Finally, example scenarios are presented where the model could be of potential benefit. These scenarios form the basis of future work.

2 Reference Model

The Reference Model incorporates a product (good or service) process model (X); an organisational hierarchy model (Y); and the collection of components forming the product (Z). At a given point in the XYZ space, a set of data requirements can be elicited. In the development phase, these points determine the required properties of the data. In the operational phase, they represent monitoring and trigger requirements for safety management and change. Development and associated operational activities enact these requirements. The next stage in the Reference Model is to provide assurance over data properties. This assurance focuses on the safety intent of the data safety requirements and the efficacy of their resolution at the given point in the XYZ space.

Figure 1a shows that the XYZ axis of the Reference Model is applicable throughout the lifecycle of a product. It is relevant for an operational organisation as well as an organisation developing a product with data requirements.

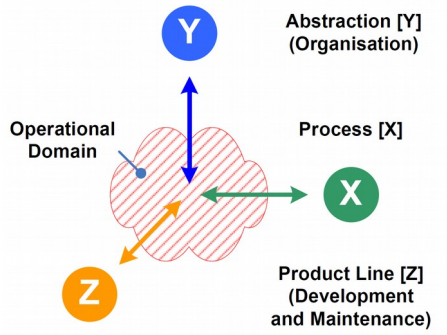

Figure 1a: XYZ axes throughout the lifecycle

Figure 1b shows that the XYZ axis of the Reference Model can be instantiated at multiple points. The aim is to target

assessment activities to appropriate data requirements, data characteristics and instantiations thus focusing the use of a toolkit of data safety methods to the context at hand.

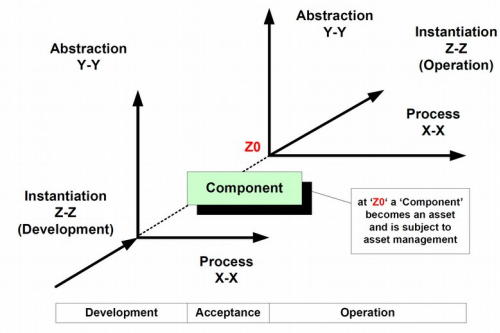

Figure 1b: Multiple Instantiations of the XYZ model

A safety justification [5] addresses the safety effectiveness of a product on a point in the XYZ axis. This justification should be hazard-focused and consider the safety argument, evidence and the confidence in that safety argument. The assurance process should employ a data variant [3] of the 4+1 principles of safety assurance [6]. The key features of the Reference Model are:

1. Process Model (X-axis);
2. Layered Organisational Model (Y-axis) [4];
3. Development and Operational Model (Z-axis);
4. Interface Agreements;
5. Metadata Definition and Management;
6. (Critical) Control Points.
7. Entity-Event-Process (EEP) Relationships [7].

2.1 X-axis

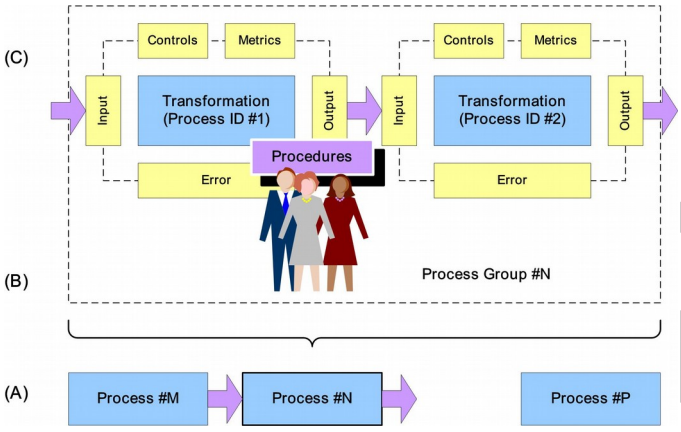

Figure 2: Process Model in X-axis

The X-axis represents the process model employed at a particular point in the lifecycle of a product for the relevant data items and characteristics of data items. For example, an organisation may employ a hierarchical structure that allows data decomposition. Each process has a standard set of attributes: Input; Transformation; Output; Error; Control and, Metrics.

Figure 2 illustrates a simple process (A) comprising a sequence of processes M to P; Process N is decomposed into a Process Group N (B); Process Group N is decomposed into two further processes (C). Each process is defined to support a suitable and sufficient safety analysis. The level of definition of the process is commensurate with the safety integrity required. An error attribute is included to provide for the separate indication of error conditions.

2.2 Y-axis

The Y-axis models the organisational and product structure. Each project employs a relevant organisational structure at an appropriate level for the data elements under discussion. One compatible organisational model is proposed in [4]. This seven-layer model has vertical and horizontal interfaces. Two important elements of each layer are communication with peer layers in a different communication unit and provision of a service to the layer above and expectation of a particular kind of service from the layer below.

Layers allow development and replacement of underlying layers based on respect for the services each layer provides and preservation of interfaces between them. Layers allow control of the access points of data and control coming into the product. They also depict the span of a safety function by identifying that high integrity safety functions should be limited to the lowest three layers of the hierarchy.

2.3 Z axis

The Z-axis represents instantiations of the data requirements. Consideration of the instantiation during design time allows exploration of the integrity required from the system; it's data architecture, data components and their properties. In the operational phase, it identifies relevant data items subject to assessment. One model compatible with the Z-axis is a "product line" [8]. There are two essential processes in any product line, namely product-line establishment (domain engineering) and product derivation process (application engineering). A product-line approach implies that the required contribution of data to safety is determined by the results of these two processes and the environment within which the instantiated product is operated.

A safety justification considers the data contribution to each identified hazard. Risk acceptability of each hazard is supported by appealing to the deployment of sufficient risk mitigation measures implemented (data) components and their interactions. To achieve assurance of the resulting data contribution within a product line approach the principles of data safety are employed.

2.4 Interface Agreements

System boundaries provide a demarcation between those components, which are within the system, and those, which are external. Communication across a boundary requires the identification and definition of an interface description including but not limited to, the data passed across the interface. Each communication is one step in a data chain. A data chain may pass vertically up and down the Y hierarchy, as well as between peer elements. At any point in the data

may be subject to transformation of its properties by a variety of intended and failure actions. As a result, design components interface with Peers (X), Subordinates and Supervisors (Y), and the Development and Maintenance (Z) elements.

Whether interfaces are between Y-axis layers, within a layer as data flows within and between partitions, or to external sources of data, Supervision, Optimisation and Control (SOC) engenders improved risk management. SOC takes the form of a safety interface agreement implemented in a relatively informal manner or formal approaches, using rely-guarantee interface contract formalisms. Interface agreements are discussed further in [4, 9].

2.5 Metadata

Metadata is data about data. It identifies the characteristics of the data, such as source, provenance, units employed, etc. It allows appropriate information to be extracted from the data. The Reference Model considers the position of data components within a hierarchy. A safety function can be explored in terms of its span across the Y-axis layers. Safety Metadata has three elements:

1. Process - sequence of operations required to transform one or more constituents into the required output. Often this is summarised as input-process-output relationship. The definition of such processes varies across the spectrum from ad-hoc to those supported by mathematical formal specifications.
2. Abstraction - decomposition into a series of layers recognising the use of metrics, performance targets, command and control to effect management by a supervisory layer over a subordinate layer, process or function.
3. Instantiation – data about capability, and compliance.

More specifically, interface agreements present additional metadata requirements such as:

- Structural metadata about the design and specification of data structures or "data about the containers of data"
- Descriptive metadata about individual instances of application data or the data content.
- Validation metadata that allows the data presented at an interface to be validated (range checks, relationships and reasonableness criteria).
- State Metadata indicates the operational states of the interface, such as normal, start-up, shutdown, and degraded modes.

2.6 Control Points

Data contributes to human decisions and to computer control actions. A one-fits-all approach to data safety is not tenable. The Reference Model reduces scale, scope and complexity to a level where the data safety activities are manageable and appropriate. By identifying points in the Reference Model specific data usage and context scenarios can be identified and safety activities developed. These activities can then be subject to assurance processes.

Consider a safety related activity such as a maintenance task. This task has to be designed, failures in the design considered and the safety risks assessed. The task as practised must be monitored against the designed task. The elements of the maintenance safety process are:

1. Determine which maintenance tasks are important with respect to safety;
2. Estimate the correct maintenance intervals;
3. Analyse procedures with respect to product hazards;
4. Analyse procedures with respect to maintainers' safety;
5. Manage and monitor the conduct of maintenance;
6. Review collected data and adjust tasks and intervals; and,
7. Investigate incidents and other indicators.

All of these elements have data components. Errors can be introduced at any point in this activity leading to the maintenance task being faulty, potentially hazardous and maybe leading to a safety event.

A data safety analyst can use the Reference Model to focus their activities. The X-axis identifies the relevant product (good or service) process data objectives and data manipulation and communication processes. The Y-axis identifies levels in the organisational structure. The analyst then identifies the nature of the instantiation of the data relating to the maintenance task of relevance in the Z-axis. Points on the XYZ axis are control points. As the discipline of data safety evolves certain control points will prove to have a larger effect on the safety characteristics of the product or organisation than others. Analysis can be tailored to these critical control points [10].

Consider the maintenance task process. A number of critical control points can be identified:

1. Maintenance task preconditions (start gate). Data relating to the following questions are important: is it safe to take this subsystem out of service for maintenance, are the staff performing the task competent, are appropriate tools and spares available?
2. Maintenance task initiation. Data relating to the following questions are important: has the subsystem been isolated, if multiple tasks are open has each set an appropriate lock?
3. Task check at selected points during the maintenance activity. Data relating to the following questions are important: has the sub-task been completed, are safeguards in place for the next subtask, have all the maintainers got the same mental model of the state of the subsystem at this point?
4. Task completion check. Data relating to the following questions are important: are parts still defective after maintenance or replacement, have parts been left out altogether, have foreign objects been left within the subsystem?

Reference Model elements not directly related to the operation of the task may have significant impact on the effectiveness of the task.

1. Context determines when to conduct maintenance and which procedure to use. Data relating to the following questions are important: is this the right maintenance pro-

cedure to be following, is this the right time to be conducting the procedure, is this task at the right point in a sequence of tasks?
2. Support determines whether it is possible to do the task as described. Here data relating to the following questions are important: is the procedure not followed for some reason, is the task usable?

Different aspects of these questions can be identified for different positions in the Reference Model. For example, if an incident occurs, the data backing up the answers can be investigated as part of a root cause analysis.

2.7 Entity Event Process (EEP) Relationships

The Reference Model employs EEPs [7, 11, 12]. For example, a sensing system provides data to a data-handling system that also provides information to a user. The user finds it difficult to make a decision on the implications of the data. He thus makes changes to the properties of the data handling system's data selection criteria. This change may have a knock-on impact on the system (via the decision maker), its performance or its computational environment. EEP is a mechanism to traverse a series of data handling systems through sets of interacting elements, their relationships and definitions (via metadata).

As an example consider the data flood [13] problem. The development process for a product produces a wealth of safety evidence. As the product goes into service, a hazard log is produced that reflects the current knowledge of the safety risks. However, the sheer quantity of evidence generated through the development process is difficult to reconcile with requirements for operational risk management.

The hazard log is maintained through the life of a product. The amount of data generated by an operational unit that may be helpful in ongoing safety risk validation is vast. For example, the current generation of military aircraft produces about 10mb of data per flight. The next generation is likely to produce a Terabyte. Unfortunately, standard tools are no longer fit for purpose because the vast quantities of data collected are unmanageable. This poses significant challenges for the presentation of such information to decision-makers.

The Reference Model allows specific scenarios and the data contribution to these scenarios to be identified. Data and metadata entities can be assessed to identify directly relevant contributions to the correctness or erroneous nature of the data. This allows irrelevant data to be "culled" from the assessment. An EEP process proceeds along the relationship chain culling out data that is not relevant and allowing safety assessment to advance.

4 Outline Process Implied by Reference Model

The Reference Model forms the backbone of a proposed data safety process [4]. The steps in this process are

1. Establish what data is involved (data and metadata):
   a. Identify candidate EEPs.
   b. Identify Interfaces Agreements as required by the EEPs.
   c. Confirm that all data is managed via an appropriate data management mechanism.
2. Get the data:
   a. Select the minimum set of relevant data via navigation of appropriate EEPs.
   b. Identify relevant XYZ point(s):
      i. Characterise data / information requirements over each XYZ point.
      ii. Select appropriate EEPs to adjoining XYZ points.
      iii. Navigate outward through EEPs until reaching stopping criteria.
      iv. Repeat for each relevant entity set interfacing directly with XYZ point.
3. Analyse the data in the identified specific XYZ and EEP context.
   a. Data analytics including agents, and simulation.
   b. Data science (Data Engineering).
   c. Communication of information.
4. Make decisions based on an assessment of the analysis.
5. Undertake activity such as rework of data structures.
6. Monitor by collecting data under data management, see process step 1.

5. Validation of the Reference Model

Here validation of the Reference Model for data safety is addressed. For a further discussion of data-centric safety see [14]. Leveson's STAMP model [15], Weinberg's categorisation of system complexity [16] and elements of resilience engineering (RE) [17-20] are discussed. Reference Model applicability is illustrated by drawing on incident reports [21]. Finally, the paper illustrates the application of the Reference Model to three significant examples.

STAMP incorporates a hierarchical control structure. Each component has responsibilities for enforcing safety constraints appropriate for that component. Accidents occur when the safety control structure does not enforce the system safety constraints and hazardous states occur. Let us consider STAMP against the Reference Model.

1. The use of data in all aspects of an organisation means that it is important to ensure that technical, process, human factors and organisational issues are considered when addressing data safety. The relative importance of these issues depends on the nature of the use and constraints of the data.
2. At the heart of STAMP is the idea of a process model that is compatible with the X-axis.
3. STAMP allows peer elements within an organisational layer to be identified. It explicitly models a hierarchy of control and constraints from the lowest levels of a technical system to the highest levels of the socio-technical organisation. It allows for the limitation of the scope of modelling and analysis.

4. STAMP explicitly allows modelling at any point in the development lifecycle. The Z-axis allows data assessment to be targeted.
5. The systems engineering model employed in STAMP allows interfaces and the data requirements for these interfaces to be identified. Constraints and controls being passed via data can also be extracted from the model. STAMP does not explicitly address interface contracts.
6. STAMP and accompanying analyses have not currently been adapted to address metadata identification and assessment. However, control and constraint models could potentially be adapted along these lines.
7. STAMP allows modelling at different levels of abstraction. Control points can be assessed using techniques associated with STAMP.
8. STAMP does not employ EEPs.

System theory [16] characterises products by consideration of the PICARD mnemonic: Parts of the system, Interactions, Context, Actions, Relationships and Destinations. Systems incorporate a number of processes: transform, distribute, store, exchange and control. They involve both complex technical and organisational elements and thus are hard to model and simulate. Holistic, as well as deconstructionist, viewpoints of the system, are required for data safety issues to be addressed in complex systems. The seven elements of the Reference Model act as an indexer to the elements of a product from multiple viewpoints thus allowing complexity to be addressed. One interpretation of PICARD under the Reference Model is: parts are addressed by Z and metadata, by interface agreements, context by Y, actions by X, relationships by EEPs and destination by chosen XYZ points such as an accident occurrence. Critical control points allow different viewpoints to be selected.

RE [18] focuses on the ability of a product to absorb the effects of a disruption to its performance. Data safety resilience looks at how the product can ensure safety given the normal variation inherent in data values during system operations. Resilience models [17] are compatible with control points in the XYZ axis, see Figure 3 [19].

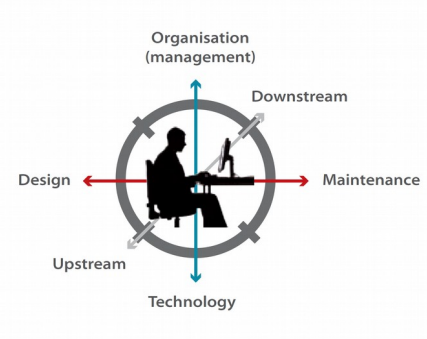

Figure 3: Safety Focus for Resilience Engineering

The Upstream and downstream axis are compatible with the X-axis. The technology to organisation axis is compatible with the Y-axis. The Design to Maintenance axis is compatible with the Z-axis.

Data contributes both to the ability of a system to successfully cope with variability and to the identification of a failure to do so. Interfaces and control points can be addressed as part of the data safety analysis of such resilience properties. Dekker [20] highlights the difference between safety as practised and safety as designed. Development of EEPS and monitoring regimes and ongoing data analytics has a significant part to play in the effectiveness of this approach.

Product safety may be compromised as a result of a breach of data security. Suppose for an organisation:

1. All connected systems are supported and protected by a suitable security model;
2. Access requirements are defined;
3. Links across the network are explicitly controlled; and,
4. Acceptable types of access (read only; read and update; read, write, create and delete) are defined.

Data security threats and associated risk assessment are based on the Reference Model data for the activity. Data is not only the subject of attack and propagation of attack but also used as part of the systems defences against such attacks. For example, data diodes and logging of traffic may be employed. A number of critical control points can be identified, during development, at release to service and after an attack. The Y-axis allows day-to-day security activities and the oversight activities to be assessed. EEPs, such as that for access control, logging and audit, can be identified.

Data was a contributory event in a number of incidents and accidents [3, 21]. Accident and incident investigation is one of the critical control points for an organisation in the Reference Model. Data is a potential causal factor, and data error-failure-accident EEPs need to be considered as part of the investigation. This can be tricky.

Suppose a safety incident has occurred involving an autonomous vehicle. Historically, operators and witnesses are questioned about their actions and why they undertook them. What if the "pilot" was an Artificial Intelligence system? How can its actions be investigated? How does the designer know what to monitor with respect to the data being used by the product? Should all alterable values in configuration files be collected? These elements can be investigated at different points along the Z-axis. Once data has been collected data forensics will be required to interrogate the appropriate EEPs to look for data causal factors.

The Reference Model is being applied to three large application areas:

1. A *quadcopter (qCopter)* featuring four propellers with a centralised *hull* containing a battery, camera, sensors, computational unit and communications interface. Communications are real-time and have sufficient data bandwidth. The qCopter has security and safety requirements. System goals include initialization, terrain discovery,

dataset Management (Change Control, Updates, etc.) and management of degraded modes. This is a *data-centric* product and the key control point to be investigated in an incident investigation. The X-axis identifies data collection, and assessment processes, the Y-Axis levels in the qCopters operations and the Z-axis identifies the relevant data, decision making and activities under investigation.

2. A commercial organisation consists of a number of sites participating in the development, manufacture, sales and logistics associated with manufacturing activities. This example expands the data, metadata and interface contracts considered in the Autonomous Flight example combining them within the logistics and management of physical entities (X), organisational structure (Y), the development and maintenance of plant (Z).
3. A *Healthcare System* provides medical services from many sites. To support service provision the *Healthcare System* requires data to provide a safety function, has significant flows of data across a data chain and has a service based environment characterised by interfaces. Security is required to maintain *Patient Confidentiality*. People must have access to records and to drugs within the regulated environment. The *Healthcare System* is highly data centric with decisions being made by humans based on the data they are provided with.

6 Time, Change and Maintenance

One reading of the Reference Model is to view it in the context of system creation. This is only one perspective. Let's consider for a moment the creation of a system, its subsequent extension and the growing requirement for maintenance. It is highly unlikely that this system will not evolve. How then, should we approach issues associated with version control, changes in demand (based on expanding or contracting volumes of data) and obsolescence.

There is a case to extend the Reference model to include time as a fourth dimension. A simple implementation is to use the Reference Model provides a means to describe the system as a time series of snapshots.

7 Conclusions and Future Work

The Reference Model, along with the accompanying process, allows a vast range of data scenarios to be identified and addressed. The model allows existing safety assessment models and methods to be extended to address the contribution of data to safety throughout the lifecycle of a product or organisation. Initial validation indicates the model is compatible with these approaches, and it is envisaged that they can be extended to make contributions to the field of data safety. A number of case studies have also been identified addressing a range of relevant scenarios and these are being actively explored.

The contribution of this paper is a model and process that allows the details of data scenarios to be identified and assessed. An inability to do this particularisation has inhibited developments in the field of data safety.

The devil, of course, is in the detail. There remains a vast amount of work to be done on all elements of the data safety process supported by the Reference Model. For example, characterisations of critical control points and interface contracts. A checklist of EEPs needs to be identified, including Hazard-Opportunity-Incident (Accident), data chain, and Process-Requirement-Evidence relationships. The use of unstructured as well as structured data needs to be addressed. The data and metadata culling process to overcome data flood issues is currently the subject of a proposed DPhil. Investigation of the data contribution to incidents is currently subject to a Masters project.

References

[1] D. Borys, D. Else, and S. Leggett. *The fifth age of safety: the adaptive age*. Volume 1 Issue 1. Journal of Health, Safety Research, and Practice, Oct. 2009.

[2] A. Faulkner. *Data Integrity – An often ignored aspect of safety systems*. University of Warwick, 2004

[3] SCSC, "Data Safety V1.3", SCSC Data Safety Initiative Working Group, ISBN-13: 978-1519533579, 2015,

[4] A. Faulkner and M. Nicholson. "An Assessment Framework for Data-Centric Systems". Proc. 22nd Safety-Critical Systems Symposium, Brighton, UK. Edited by C. Dale and T. Anderson. SCSC, 2014.

[5] GSN Working Group. "GSN Community Standard Version 1." *Origin Consulting (York) Limited*, 2011.

[6] R. Hawkins., I. Habli, and T. Kelly. "The principles of software safety assurance." *31st International System Safety Conference*. 2013.

[7] G Askew, "Triangulation: Navigation of Information Contexts Using Triplet Relationships", UNPUBLISHED, 2016

[8] A. L.de Oliveira, et al. "Supporting the Automated Generation of Modular Product Line Safety Cases." *Theory and Engineering of Complex Systems and Dependability*. Springer International Publishing, 2015. 319-330.

[9] I. Bate, R. Hawkins, and J. A. McDermid. "A contract-based approach to designing safe systems." *Proc. 8th Australian workshop on Safety critical systems and software-Vol 33*. Australian Computer Society, Inc., 2003.

[10] M. D. Pierson. "HACCP: principles and applications Springer Science & Business Media, 2012.

[11] O. Etzion and P. Niblett. *Event processing in action*. Manning Publications Co., 2010.

[12] L. Baekgaard, "Event-entity-relationship modeling in data warehouse environments." *Proceedings of the 2nd ACM international workshop on Data warehousing and OLAP*." ACM, 1999.

[13] R. Isaac III., et al. *Data Flood: Helping the Navy Address the Rising Tide of Sensor Information*. Rand Corporation, 2014.

[14] M. Nicholson and A. Faulkner, "Data Centric - the Sixth Age of Safety: Communications Enabled Transition from Function to Service", 7th IET System Safety and Cyber Security conference, IET London, 2016
[15] N. Leveson, "Engineering a safer world: applying systems thinking to safety", 2012.
[16] G. M. Weinberg, "An introduction to general systems thinking", 1975: 657-665.
[17] J. Lundberg et al. "Systemic resilience model." *Rel. Engineering & System Safety* 141 (2015): 22-32.
[18] A. Hosseini et al. "A review of definitions and measures of system resilience." *Reliability Engineering & System Safety* 145 (2016): 47-61.
[19] Eurocontrol, "A White Paper on Resilience Engineering for ATM", Eurocontrol, 2013
[20] S. Dekker. *Safety differently: Human Factors for a new era*. CRC Press, 2014.
[21] ATSB, "Loading issue involving a Boeing 737, VH-VZO, Canberra Airport, Australian Capital Territory, 9 May 2014", ATSM Sept. 2014.