

# Information Classification in Information Security Management and its Challenges

Robert Katura

**Information Security, master's level (120 credits)**  
**2023**

Luleå University of Technology  
Department of Computer Science, Electrical and Space Engineering

[This page intentionally left blank]

## Abstract

Information classification is a prerequisite for carrying out risk management in information security, as the assets worth protecting are identified and the need for protection is determined by the classification categories. The information classification thus has a major impact on the security architecture of systems and organizations. Nevertheless, information classification leads a shadowy existence in the scientific literature, which is reflected in a limited number of scientific publications. This discrepancy between the relevance of information classification in risk management and its low scientific attention was the motivation to take a closer look at the topic. This thesis created an overview of the current state of research in information classification and shed some light on potential problems to stimulate new research questions. The results of the work include a current overview of the status of research on information classification in risk management of information security and its context to other academic disciplines and practical needs, particularly research on bias and systems engineering. This thesis also summarized a total of 109 individual research gaps in information classification research, derived from the evaluation of the scientific literature and on the conclusions of identified open questions. From the gaps identified, some suggestions for future research in the field of information classification could be made.

**Keywords:** Information classification, information classification research gaps, risk analysis, risk assessment, risk management, information security.

## Table of Contents

1. Introduction.....	1
1.1. Background.....	1
1.2. Aim .....	4
1.3. Research question .....	4
1.4. Limitations .....	4
2. Research methodology.....	5
2.1. Literature review .....	5
2.2. Systematic literature review .....	7
2.3. Description of the methodological approach.....	8
2.3.1. Literature search .....	10
2.3.2. Literature sources.....	13
2.3.3. Literature statistics .....	13
3. Qualitative literature review .....	14
3.1. Information classification .....	14
3.2. Information classification in standards .....	15
3.3. Standards challenges.....	23
3.4. Information classification in scientific research .....	26
3.5. Scientific research challenges .....	34
3.6. Information classification in books .....	37
3.7. Information classification in practice .....	46
3.8. Practice challenges .....	48
3.9. Context .....	49
3.9.1. Bias in information classification.....	50
3.9.2. Information gathering .....	51
3.9.3. Information handling.....	53
3.9.4. Assets.....	53
3.9.5. Risk .....	56
3.9.6. Risk assessment.....	57
3.9.7. Risk management .....	60
3.10. System engineering .....	61
4. Discussion and conclusion .....	62
4.1. Discussion .....	62
4.2. Conclusion .....	74
4.3. Future research directions .....	76
References.....	78

## List of Figures

Figure 1: Literature review scope and research goals.....	4
Figure 2: A systematic guide to literature review development (Okoli & Schabram, 2010). ....	6
Figure 3: Methodological research process. ....	9
Figure 4: Literature search expansion. ....	12
Figure 5: Expanded literature search process. ....	13
Figure 6: Security Categorization Process (Stallings, 2018) .....	15
Figure 7: ANSSI Classification Method (Shrestha et al., 2017). ....	16
Figure 8: Data Processing Ecosystem Relationships (National Institute of Standards and Technology, 2020).....	21
Figure 9: The 9 Box of Controls (Harkins, 2013). ....	24
Figure 10: Degrees of trust (Wheeler, 2011). ....	25
Figure 11: Degrees of endpoint trust (Wheeler, 2011). ....	25
Figure 12: The low granularity route (Bergström, Karlsson, & Åhlfeldt, 2020).....	28
Figure 13: Stages of the ISMS process - The evaluation stage (Bjorck, 2005). ....	30
Figure 14: The Reference Model of Information Assurance & Security (RMIAS) (Cherdantseva, 2014). .....	33
Figure 15: Information Life Cycle illustrates possible states of information (Cherdantseva, 2014). ....	33
Figure 16: Overview of information security ontology concepts and relationships (Vargas & Fenz, 2012).....	36
Figure 17: The information classification pyramid (Campbell, 2016). ....	39
Figure 18: Document Management Life Cycle (Stallings, 2018). ....	44
Figure 19: Security and privacy : the primary areas of information risk, and the core elements of information risk management that apply to each area (Harkins, 2013). ....	44
Figure 20: Records Management Functions (Stallings, 2018). ....	45
Figure 21: Answers of the respondents to the question: How often do you carry out the various stages of the risk management process (Kobis, 2020). ....	47
Figure 22: Example of breakdown of security costs (Volchikov, 2018).....	49
Figure 23: Risk management process according to ISO 31000 with examples of decisions per stage (Wit et al., 2021). ....	50
Figure 24: Value pyramid of information (Al-Fedaghi, 2008).....	55
Figure 25: Relationship Between Privacy Risk and Organizational Risk (Stallings, 2018). ....	56
Figure 26: The 3 kinds of residual risks (ETSI, 2023). ....	57
Figure 27: Determining Information Security Risk (Stallings, 2018). ....	57
Figure 28: Risk Assessment Using FAIR (Stallings, 2018). ....	58
Figure 29: Risk Management Framework (National Institute of Standards and Technology, 2011). ..	60
Figure 30: Very simplified representation of the table "Summary of finding to Q2: Gaps in information classification research." .....	72
Figure 31: Information classification as it is - static. ....	77
Figure 32: Information classification as it should be - dynamic. ....	77

## List of Tables

Table 1: List of references from the first search in Web of Science.....	11
Table 2: Results of extended literature search. ....	14
Table 3: Results for relevant information classification literature.....	14
Table 4: Respondents replied to the question: Please specify to what extent (1- smallest, 5-largest) in your enterprise the information security may be influenced by the following factors (Kobis, 2020)..	26
Table 5: Summary of the analysis of how existing methods for information classification support stated MRs (Bergström, Karlsson, & Åhlfeldt, 2020). ....	27
Table 6: General Asset List (excerpt) (Landoll, 2021).....	43
Table 7: Reference links between assets (Volchkov, 2018). ....	46
Table 8: Structure of classification frameworks in the case countries (Heide & Villeneuve, 2020). ....	47
Table 9: Example of a security risk matrix with different impact scales (Volchkov, 2018). ....	59
Table 10: Comparison of Quantitative and Qualitative Risk Assessment (Stallings, 2018). ....	59
Table 11: Summary of finding to Q2: Gaps in information classification research. ....	63
Table 12: Issue topics of the identified research gaps in information classification. ....	76

# 1. Introduction

This Master thesis examines information classification in the sub-area of risk analysis in IT risk management for potential new scientific research approaches or perspectives. The focus is on a structured literature research method of previous scientific knowledge, supplemented by risk management standards and affected IT specialist areas.

In information management, information classification is the most important step towards targeted information security, as it makes a statement about the value or importance of information for an organization. „Information should be classified in terms of legal requirements, value, credibility, priority, criticality and sensitivity to unauthorized disclosure or modification“ (International Organization for Standardization, 2015b, p. 6). This value statement can be in the form of quantitative numerical values, as currencies, qualitative categories, or rankings.

No matter which classification method is chosen, all steps of information security such as the determination of security measures, security investments, monitoring, etc. are largely based on the decisions of information classification.

Despite its importance in the risk management process, information classification has rarely been addressed in-depth in scholarly work. But even in the information security standards, it tends to lead a shadowy existence, which is dealt with explicitly, but very superficially and openly to interpretation. To find a statement about the research gap in this subject area, a literature analysis that is as comprehensive as possible is necessary. Therefore, the systematic literature review as a research method is suitable, which aims to consider as much of the existing and relevant literature on the research topic as possible in a structured way.

The difficulty with the topic is that different terms are used for information classification, such as categorization instead of classification, and data or asset instead of information. In addition, there is no consensus for the same terms in the scientific literature and the standards, so that different definitions exist. Due to the relatively few literature sources on the topic and the non-standardized terminology, the direct results still leave too many unanswered questions, so that a selective expansion of the analysis to include aspects of risk analysis, risk assessment and risk management covering the topic of information classification can be considered in order to provide a meaningful context around information classification, to be able to generate conclusions about the information classification itself. Despite all efforts, it is certainly not possible to find all the literature on the subject, because there may be papers and books with missing access rights, that were no longer available, were unknown or were under lock and key. Nevertheless, the overview has become quite extensive.

## 1.1. Background

Information security is an important part of organizations and has become particularly important in information technology in recent years. A very important task on the way to implementing security measures is the identification and, above all, adequate classification of information. It could even be said that information is the most valuable asset for organizations (companies and authorities) (Stallings, 2018) and can be of different type, as financial, privacy or sensitive (National Institute of Standards and Technology, 2004). This can be operationally relevant data about processes, people and goods that represent a trade secret, have an impact on the reputation of the organization, state secrets, or must be protected by legal requirements. In order to secure information properly, information is divided into security categories, based on a security assessment, depending on the potential impact on the security objectives, also known as the CIA triad confidentiality, integrity and availability (National Institute of Standards and Technology, 2004). However, due to the importance of information in organizations and the classification of information in the context of information

security, the topic has still shortcomings and unresolved issues. Ineffective business network impact assessments are a major cause of cyber security failures because the dependencies of vulnerabilities and systems among each other and within a network are not fully considered (Cao et al., 2018). But also flawed information classification can be a cause of security attacks (Roessing et al., 2013). Thus, the identification of all relevant decision-supporting information for the risk analysis and in particular the information classification is important to carry out a realistic security classification of the information assets. On the one hand, this is important for the risk analysis of individual assets, but also especially for the aggregated risk analysis, where there are cause-effect relationships between different assets, threats, or vulnerabilities (National Institute of Standards and Technology, 2008b). This could not only reduce the bias of the person-related qualitative decision, but also the general uncertainty in the entire risk assessment process, whether carried out qualitatively or quantitatively. There are both qualitative (International Organization for Standardization, 2013b), and quantitative (Freund & Jones, 2014) risk analysis methods, which, however, based on the current knowledge, do not take a holistic view of all information relevant to decision-making for the risk analysis process or, specifically, information classification. It is left to the users which asset or information is relevant and how it is considered (National Institute of Standards and Technology, 2008b). Although the information security standards and frameworks explicitly point out that the corresponding vulnerabilities, threats, and aggregated risks should be considered, the actual connections and implementations remain hidden. This applies, for example, to the question of how the most important sources of vulnerabilities and threats (attack paths) *CVE* and *MITRE Attack Map* can be integrated into the risk analysis. If this specification does not take place, all information risk classifications will remain highly uncertain and biased. Due to the very complex and diverse IT landscape and massive processing of heterogeneous data, very complex information and asset landscapes arise (Yuan et al., 2021), which must be fully recorded and appropriately classified, and risk analysed for a risk assessment. First, an overview of all useful models and concepts for information risk classification from standards and scientific literature is to be created. Based on this overview, as the status quo in the literature, conclusions or assumptions can be drawn about research gaps and open research questions. At the time the thesis was written, such an extensive overview did not exist.

From the risk assessment process view information classification is the second and most important task, directly after asset identification, to be able to carry out a risk analysis, which in turn entails long-term, costly, and momentous information security decisions. Information classification, however, is in conflict between qualitative evaluation by humans and quantitative value determination and automation. The intersection with other specialist areas and topics required for information classification is very large, so that the identification and classification of the information assets is transferred to the information owners as experts in their specialist area concerned. The interdependencies to other topics, such as IT infrastructure or role and access models, should be discussed in coordination with other experts. The difficulty here lies in the fact that manual activities, especially with complex interpretation steps such as in information classification, are very personnel-intensive, time-consuming, and sluggish (Harkins, 2013). At the same time, new technologies, applications, and security threats from new attack surfaces (Stallings, 2018) such as cloud infrastructure, 5G, Internet of Things (IoT) are growing, and attacks are professionalizing (Millett et al., 2017). These pose new challenges for information classification in terms of asset identification and information classification, which means that the time intervals for reassessing risks are becoming shorter and the scope and complexity of the distributed and generated information and their relationships are increasing. But there are problems with the previous approach because there is an over-confidence in the cybersecurity community in one's ability to assess the social and psychological mechanisms at work among users and adversaries. In addition, or precisely because of this, methodological approaches from other research areas, such as sociology, health sciences, humanities,



or engineering, are rarely used. However, this has negative effects on the scientific quality of the methods used in cybersecurity research (Millett et al., 2017). Consequently, both, the focus on expert skills in qualitative analysis in information classification and risk management, and the self-centeredness in information security may have left gaps in information security research.

A scientific analysis of cybersecurity challenges backed up with well-substantiated models could help to identify the actual cybersecurity needs by linking cause and effect of attacks and addressing potential solutions (Millett et al., 2017). But within cyber security, there is not even a growing realization that there could be a potential interdisciplinary research question here at all, but attempts are being made to reinvent the wheel and improvise where necessary, by ignoring methods and findings from other sciences that have already been worked out very well (Millett et al., 2017).

When it comes to the classification of information, there are a few open points that make an in-depth scientific analysis difficult. "In literature, information classification is seldom dealt with in-depth" (Bergström, Karlsson, & Åhlfeldt, 2020, p. 211). And there are different definitions for information classification in the context of information security, depending on what has been identified as a good worth protecting and what position it has in the context of consideration (Collard et al., 2017). "Still, this issue of subjective judgement is challenging and often ignored, both in practice and research" (Bergström, Karlsson, & Åhlfeldt, 2020, p. 211). There is a lot of advice about implementing information classification systematically but there are few approaches to reduce subjectivity in research (Bergström, Karlsson, & Åhlfeldt, 2020) or in standards. An attempt in standards to reduce subjectivity would be to make recommendations for information or security classifications (National Institute of Standards and Technology, 2008c). However, these decision templates do not heal the basic problem of information classification, namely the incomplete data situation, the insufficient level of knowledge, the static nature and historical context of the classification. There is also no consensus on its granularity. Both too coarse-grained and too fine-grained classifications lead to problems in the correct assignment of security measures (Seifert & Relyea, 2004). And that, complexity, and granularity, is exactly why information classification is not as easy to carry out as it appears at first glance, and the previous practice of information classification has suggested (Talabis and Martin, 2012).

The scientific literature has so far failed to provide this insight into the complexity behind the apparent simplicity of information classification in information security analysis. But scientific research is challenging in this field, as empirical studies are likely to be difficult to carry out in the area of information security due to the difficult data acquisition (Bergström, Karlsson, & Åhlfeldt, 2020). But also information security standards make the situation even worse, as information security standards use different methods, frameworks and process descriptions for information classification, so that there is no common methodological basis (Bergström, Karlsson, & Åhlfeldt, 2020). In addition, there is a discrepancy between theory and practice (Millett et al., 2017), as "it is well-known that there is a gap between formal and actual processes in information security management (ISM), as turning standards into practice is easier said than done" (Bergström, Karlsson, & Åhlfeldt, 2020, p. 211). The process of information classification is so nonspecific and complex that practitioners spend more time understanding the process than doing the task of information classification. Among other things, there is a lack of graphical process descriptions. Bergström et. al. suggest BPMN as the process modelling standard (Bergström, Karlsson, & Åhlfeldt, 2020).

The challenges in cyber security are very complex, dynamic, and fast-paced. However, research on cyber security is still relatively new. To solve the challenges, basic research is necessary to understand the artifacts, how they work, relationships and mutual influences. This requires sustainable research projects, but still IT systems are not designed with security in mind, are bad designed, implemented

and maintained, not enough money is spent on cybersecurity, and too many stakeholders with non-IT interests influence IT decisions (Millett et al., 2017).

## 1.2. Aim

The aim of this thesis is to identify potential new research approaches in information classification in information security risk management.

To achieve this aim, the following goals are identified:

- The current state of the art in the standards of information security and the state of scientific research should be identified by means of a qualitative literature research.
- To identify further potential subject areas for information classification, further IT subjects involved in risk management are considered. Here the focus is particularly on books and standards of the subject areas.
- Finally, the results from literature, standards, and book research should be compared with the topics previously identified. The delta of this consideration potentially results in aspects of information classification that have not been sufficiently considered in research up to now.



Figure 1: Literature review scope and research goals.

## 1.3. Research question

- Q1: What is the current state of science on information classification?
- Q2: What research gaps are there in information classification research?

## 1.4. Limitations

In the core topic of information classification, the literature analysis is based on a relatively small amount of literature. Therefore, an attempt was made to provide context to the topics within which information classification is relevant, which are risk analysis, risk assessment and risk management. Consequently, the attempt was to use the combination of systematic and exploratory literature analysis to partially compensate for the lack of literature sources on context information, and use academic, professional, and internet sources (Rowley & Slack, 2004). The literature procurement process was enormous, since the literature search platform *Web of Science*, which was chosen as the primary source, provided a surprisingly insufficient list of hits, so that an additional search via *Semantic Scholar*, and *Startpage.com* took place. A full-text search was preferred over a search in the abstracts and keywords because it turned out through random sampling that literature from adjacent areas of information classification contained relevant information in the full text but not in the abstract and the keywords. This experience was also made by *Bergström and Ahlfeldt* in a systematic literature review (Bergström & Ahlfeldt, 2014). All full-texts and their citations had to be downloaded and imported into *Citavi*, a referencing and qualitative literature analysis tool.

One of the results of the systematic literature analysis would also be a detailed statistical analysis of the analyzed literature. This could not be carried out due to the very often missing metadata in the PDF files, and the missing additional information of the associated citations. Furthermore, the analysis could not be extended to all potential perspectives of information classifications, since the scope of the thesis would otherwise have been exceeded. The focus was mainly on the identification of gaps and potential research approaches for the further development of information classification.

Despite the claim in this thesis to find as much literature as possible on the research topic in the systematic literature review, a complete literature analysis is unlikely to be feasible in practice. Among other things, this is due to the limited time, the incomplete research platforms, no longer available literature, and the topic that is rarely considered scientifically, which is considered more casually in other research questions and is therefore difficult to find.

## 2. Research methodology

This chapter describes the research methods and techniques used. The reasons for the selection of the respective methods are explained and why deviations from the standard method and instead combinations of different literature analysis methods were used. In principle, the selection of methods is always based on whether knowledge can be gained to answer the research questions; this also applies to deviations from the standard approach in the methodology. In the first section, the individual relevant literature analysis methods are presented, i.e., starting with the general literature analysis, proceed with the special forms of literature analysis, and finalize the section with the combination of the literature analysis methods in a process model for the methodological approach.

### 2.1. Literature review

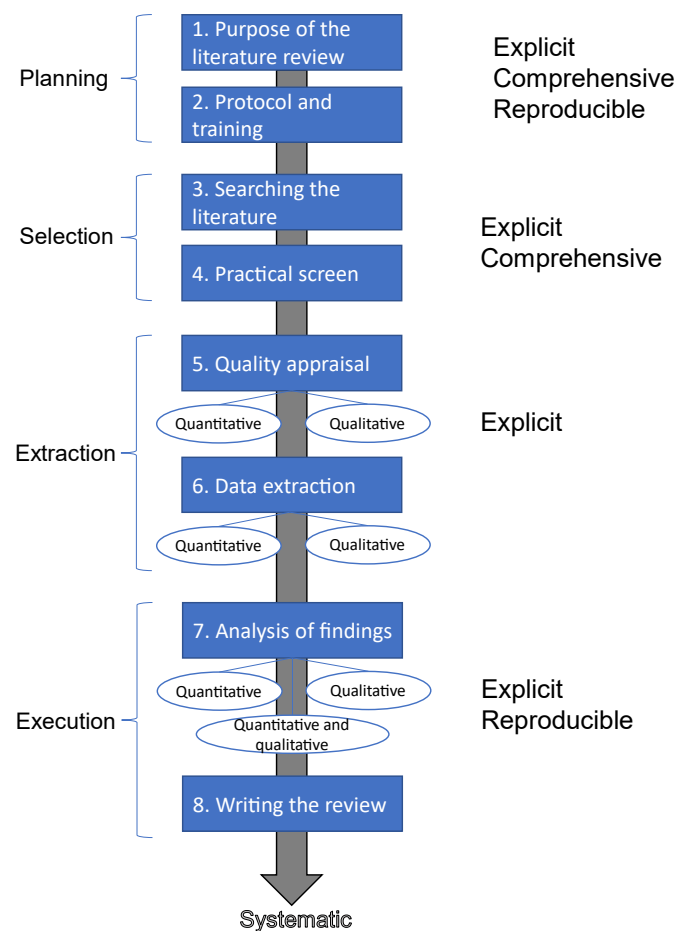
In a world of science with many publications and very focused scientific questions, knowledge is scattered and therefore difficult to keep track of. For this reason, scientific reviews are very useful because they synthesize the existing knowledge and thus make it available to a wider audience (Friedman & Schneider, 2015). A literature review informs about patterns identified in literature (Webster & Watson, 2002). Literature review help to identify research topics, questions or hypothesis, identify relevant literature, put literature into context, help to understand concepts and terminology, build bibliography of relevant literature, identify helpful research methods, analyze and interpret results (Rowley & Slack, 2004). But there are different literature review methods for different purposes and research questions, and literature review methods overviews differ in the number of described literature review methods. It therefore turned out to be very difficult to identify a correct literature review method for the given question. Some authors describe the literature review methods *narrative review*, *developmental review*, *cumulative review* and *aggregative review*, some of which can also be combined (Templier & Paré, 2015). Other authors list additional review types as *descriptive review* or *mapping review*, *scoping review*, *systematic review*, *umbrella review*, *realist review*, *critical review* (Lau & Kuziemy, 2016), *meta-analysis review*, *theoretical review* and *hybrid review* (Paré et al., 2015). But there are more, like *mixed methods review*, *overview review*, *qualitative systematic review*, *qualitative evidence synthesis review*, *rapid review*, *state-of-the-art review*, *systematic search and review*, *systematized review* (Grant & Booth, 2009). This amount of literature review methods does not make the project easier and more understandable, and it is not always clear whether there is a clear distinction between the methods and whether the same methods are used by different authors under different names. All in all, a rather unsatisfactory situation from a scientific perspective.

There are inconsistencies in the designation of the literature review methods, since different authors use different designations and literature reviews are not always identified as such (Heide & Villeneuve, 2020; Paré et al., 2015).

At least the basic 6 steps to carry out a literature review seem to be quite consistent, see the following list and *Figure 2*.

Overview of the Literature Review Process and Steps (Lau & Kuziemy, 2016):

1. Formulating the research question(s) and objective(s).
2. Searching the extant literature.
3. Screening for inclusion.
4. Assessing the quality of primary studies.
5. Extracting data.
6. Analyzing data.



*Figure 2: A systematic guide to literature review development (Okoli & Schabram, 2010).*

Ultimately, the diversity of different literature review methods can be interpreted in such a way that the research goal, the research area, and preferences determine the method, and that both qualitative or quantitative, and combinations of different literature review methods are possible (Morse et al., 2002), as long as the quality criteria of the scientific work are met. Reference is made to the terms rigor, relevance and methodological coherence (Templier & Paré, 2015). Rigor means internal and external validity, reliability and objectivity in achieving goals (Morse et al., 2002; Ogawa & Malen, 1991). Relevance refers to the usefulness of the review for research (Cooper, 1988). Methodological

coherence means that there must be a congruence between the research question and the methods used (Morse et al., 2002). Methodological coherence combines rigor and relevance and validates the correspondence between the objectives of a review and the methodological guidelines chosen to achieve it (Templier & Paré, 2015).

Due to the required short publication cycles in computer science, reviews are becoming unattractive and rare because they require more effort and are difficult to implement in the short cycle times of today's research. However, synthesizing, or reorganizing research can have great added value for the advancement of research in computer science. Because they summarize the current state of science in an area where, in general, more and more research results are being produced and the research perspectives are becoming more and more specialized, so that it is becoming increasingly difficult to get an overview of the research landscape in computer science. In turn, this knowledge about the state of research can serve as a very good and, above all, easily accessible starting point for further research (Friedman & Schneider, 2015).

## 2.2. Systematic literature review

Systematic literature analysis aims to answer the research question by summarizing, critically evaluating, and synthesizing relevant literature. Reasons for systematic literature reviews are to summarize knowledge and identify gaps in scientific research and to give guidance for new research activities (Kitchenham & Charters, 2007). A specific feature of the systematic literature review is a comprehensive literature search using defined search criteria in several databases to find as many published and unpublished sources as possible for the research topic. The analysis and synthesis of the literature can be carried out qualitatively, for example through content analysis, frameworks, classification schemes, tables, or through a statistical meta-analysis (Lau & Kuziemy, 2016). The qualitative systematic literature review analyzes texts using different content analysis methods, such as groupings, clusters, classification, listings, to derive conclusions or recommendations. Quasi-quantitative techniques are also used to enable statistical analysis of the text body (Paré et al., 2015). The focus of this thesis is clearly on the qualitative content analysis.

Advantages of systematic reviews are, they are less biased, can provide transferable insights about research topics, and can combine qualitative literature review with quantitative meta-data analysis. The disadvantage is, it requires much more work than other literature reviews (Kitchenham & Charters, 2007). Systematic reviews are used, for example, in healthcare for evidence-based decision-making (Hausner). But systematic literature review seems to be almost unknown in information systems research (Okoli & Schabram, 2010).

A possible weakness of the systematic literature analysis could be a biased selection of the articles by the author (Lau & Kuziemy, 2016). This weakness is minimized by using all topic-specific articles for analysis. Articles are only excluded if they are completely unrelated to information classification or their parent topics as risk analysis, risk assessment, or risk management.

Due to the relatively small number of scientific works on information classification, there is a risk that a systematic literature review will yield little knowledge. To minimize this risk the systematic literature review is carried out with an explorative approach. Due to the lack of extensive scientific content on the topic, the exploratory approach helps to identify the context in which the information classification takes place. The relationships between the information classification and the context could possibly allow conclusions to be drawn about the information classification. *Watson and Webster* argue that a literature review must be cross-domain to identify relevant ideas. Also, the links between themes and concepts are relevant to understanding the context (Watson & Webster, 2020).

A manual explorative literature review is very time-consuming, outdated and an impossible undertaking with large stocks of literature sources. Especially when reviewing the literature, when

building an overview of the research subject, and depending on the research question or method, also the adjacent research areas, as many sources as possible should be viewed and categorized according to the question. The question quickly arises as to which compromises, in particular the limitation of the number of literatures, must be made to achieve a result within a given time frame (Asmussen & Møller, 2019). Therefore, literature analysis tools will be used where possible.

As this thesis attempts also to identify gaps in the research, the thesis will also adhere to the approach of critical literature review. Critical review aims to identify weaknesses, contradictions, controversies, or inconsistencies in existing knowledge to stimulate new research. However, critical reviews rarely examine the topic comprehensively, as systematic reviews attempt, but rather selectively. Also, no quality assessments are made in the source literature in critical reviews. A possible weakness of this review method is precisely this selective choice of sources, the evaluative criticism and the lack of quality testing of the sources (Paré et al., 2015).

Due to the sparse literature on information classification and the exploratory approach that is necessary as a result, it is the goal to identify as much literature as possible directly on the topic and the context. Due to the very high literature search and analysis effort required as a result and the consequent time constraints, the credibility checks on the sources must be omitted.

In summary, the research question of this thesis is answered by a mix of different literature review methods. A systematic review is particularly suitable for analysing the status of research on information classification. To get a larger overview of the topics in the research literature, the tool-supported qualitative literature review is utilized. To identify inconsistencies in current knowledge the critical review is predestined. The hope is that this will help to draw conclusions about potential gaps in research on information classification to derive potential research questions. For a systematic literature review, it would also be advisable to come up with statistical evaluations of the type of literature, the sources, the keywords, the authors, and citations. However, not all references from the literature sources are detailed enough to obtain the necessary information. For most of the references only the bare minimum could be received to build a bibliography. One possible approach would have been to exclude those literature sources with missing metadata on the literature from the analysis. However, that would counteract the approach of evaluating as much literature as possible about investigation. That's why these statistical evaluations have been dropped.

## 2.3. Description of the methodological approach

Our methodological approach is depicted in the following diagram as a process model in the *Business Process Model and Notation (BPMN)*. The process describes the course of the systematic literature review and essentially consists of searching for literature, evaluating the literature, importing the literature (reference and full text) into the literature management program *Citavi*, full-text search for relevant text passages, thematic categorization of relevant literature, writing down the results and identifying potential gaps. The process is described in detail in the following paragraphs.

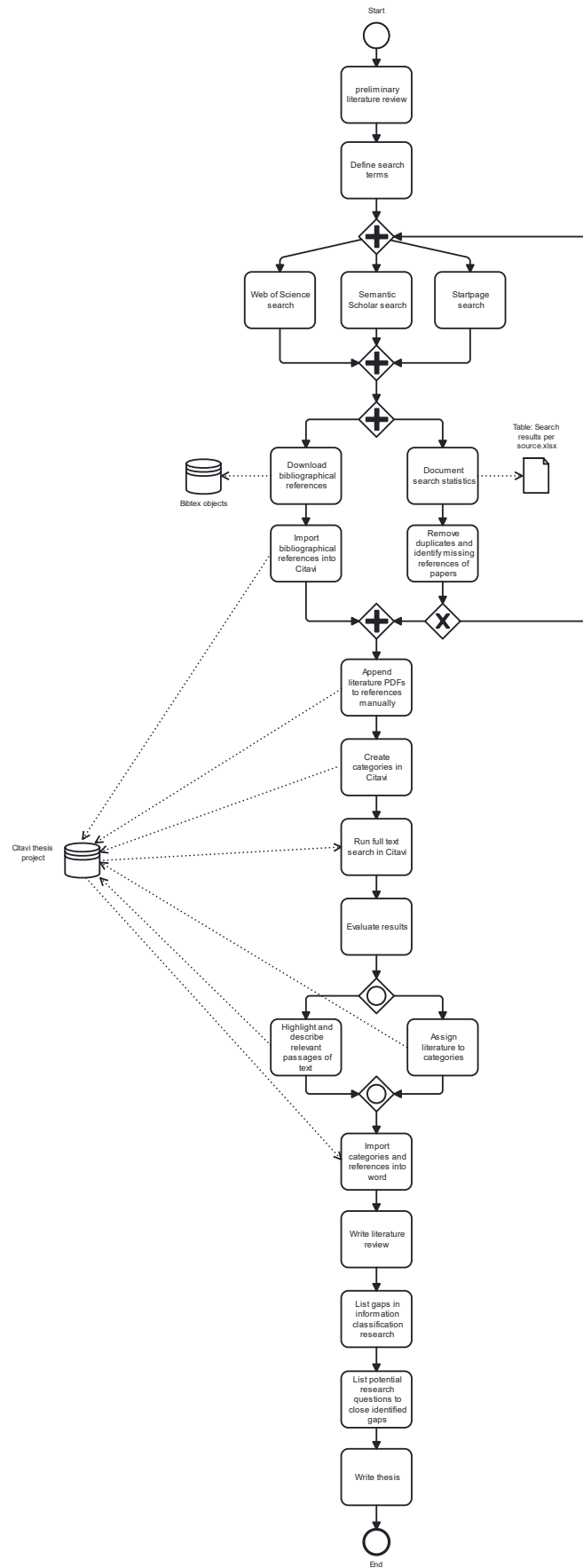


Figure 3: Methodological research process.

### 2.3.1. Literature search

At the beginning of the literature search, the focus was on the literature research platform *Web of Science* because a structured search and documentation of the literature for a statistical literature analysis was intended. The search of keywords was applied to all fields, i.e., Title, Keywords, Abstract. But *Web of Science* does not offer a search in the full text (content search).

When searching for literature with the search term "Information classification", filtered according to the following Web of Science categories:

- Computer Science Information Systems
- Computer Science Artificial Intelligence
- Computer Science Theory Methods
- Engineering Electrical Electronic
- Information Science Library Science
- Computer Science Interdisciplinary Applications
- Computer Science Software Engineering
- Management
- Telecommunications
- Business
- Computer Science Hardware Architecture
- Operations Research Management Science
- Computer Science Cybernetics
- Automation Control Systems
- Economics
- Engineering Industrial

It turned out that there were only 31 hits, see *Table 1*. An initial review of the papers content revealed that only a fraction of it was relevant to the scientific questions.



Table 1: List of references from the first search in Web of Science.

Title	Authors	Publication Year
Ontology construction for information classification	Weng, SS; Tsai, HJ; Liu, SC; Hsu, CH	2006
Information systems security policy implementation in practice: from best practices to situated practices	Niemimaa, Elina; Niemimaa, Marko	2017
Security-Aware Information Classifications Using Supervised Learning for Cloud-Based Cyber Risk Management in Financial Big Data	Gai, Keke; Qiu, Meikang; Elnagdy, Sam Adam	2016
Do you know where your information is in the homeland security era?	Seifert, JW; Relyea, HC	2004
An Intuitionistic Fuzzy Set Approach for Multi-attribute Information Classification and Decision-Making	Singh, Pritpal; Huang, Yo-Ping; Wu, Shu, I	2020
Research on intelligent classification of multi-attribute safety information and determination of operating environment	Qi, Xiaofeng; Cui, Tiejun; Shao, Liangshan; Xing, Yuyan	2020
Information Classification Scheme for Next Generation Access Control Models in Mobile Patient-Centered Care Systems	Alsalamah, Shada	2017
Personal Information Classification for Privacy Negotiation	Jang, Injoo; Yoo, Hyeong Seon	2009
A Definition of Information Security Classification in Cybersecurity Context	Collard, Guillaume; Disson, Eric; Ducroquet, Stephane; Talens, Guilaine	2017
An Information Classification approach based on Knowledge Network	Li, Huakang; Sun, Guozi; Xu, Bei; Li, Li; Huang, Jie; Tanno, Keita; Wu, Wenxu; Xu, Changen	2014
Irrelevant Features, Class Separability, and Complexity of Classification Problems	Skrypnik, Iryna	2011
Security requirement engineering at a telecom provider	Zuccato, Albin; Endersz, Viktor; Daniels, Nils	2008
Ontology Matching: State of the Art, Future Challenges, and Thinking Based on Utilized Information	Liu, Xiulei; Tong, Qiang; Liu, Xuhong; Qin, Zhihui	2021
LAWS, DECREES AND STANDARDS FOR INFORMATION SECURITY MANAGEMENT IN GOVERNMENT AGENCIES	de Araujo, Wagner Junqueira	2012
CONSTITUTIVE ELEMENTS OF THE TAXONOMY CONCEPT	Aganette, Elisangela; Alvarenga, Lidia; Souza, Renato Rocha	2010
A PROPOSAL FOR A DOMAIN ONTOLOGY ABOUT INFORMATION SECURITY IN ORGANIZATIONS: description of the terminological stage	Almeida, Mauricio Barcellos; Souza, Renato Rocha; Coelho, Katia Cardoso	2010
Classification Method for Network Security Data Based on Multi-featured Extraction	Kang, Yunchuan; Zhong, Jing; Li, Ruofeng; Liang, Yuqiao; Zhang, Nian	2021
Information Classification Enablers	Bergstrom, Erik; Ahlfeldt, Rose-Mharie	2016
The Application of AI-Based Technology in Computer Network Operation and Maintenance	Li, Li	2022
An information classification model for public sector organizations in Sweden: a case study of a Swedish municipality	Bergquist, Jan-Halvard; Tinetti, Samantha; Gao, Shang	2022
Research on information classification and storage in cloud computing data center based on group collaboration intelligent clustering	Zhang, Linlin; Zhang, Sujuan	2021
Developing an information classification method	Bergstrom, Erik; Karlsson, Fredrik; Ahlfeldt, Rose-Mharie	2021
Cyber Threat Information Classification and Life Cycle Management using Smart Contracts	Graf, Roman; King, Ross	2018
Research on Information Architecture Based on Graphic Reasoning and Mental Model	Long, Ren; Zhang, Jiali	2018
The Analysis of Engineering Data Information System Based on Knowledge Management and Internet plus Technology	Lv, Nian; Zhang, Xuan	2016
Visualizing High Dimensional Feature Space for Feature-Based Information Classification	Wang, Xiaokun; Yang, Li	2016
Exploring Security Strategies for Enterprise Data Protection in Organisations	Ajigini, Olusegun Ademolu	2015
Taxonomy as a Service for Enterprise Applications: An Approach to Improving Information Findability	Iams, Richard	2010
Conceptual Object Model of a Classified Electronic Documents Information System	Matic, Ivan; Skocir, Zoran	2008
Applying uniform information classification and coding to build domain ontology	Li, Shuguang; Wang, Junbiao; Siong, Iijun; Jiang, Jianjun	2007
Research on supply chain information classification based on information value and information sensitivity	Shi, Xianliang; Li, Dong; Zhu, Hailong; Zhang, Wenjie	2007

The amount of literature was insufficient for a systematic literature analysis. From the preliminary research, however, it was noticed that the term information classification is not used uniformly, but that the terms *data* or *asset* were also used synonymously for *information*, and *categorization* was also used synonymously for *classification*. That is why the combinations of the synonymous terms as search terms have been defined:

- information categorization
- data classification
- data categorization
- asset classification
- asset categorization

The problem with the new combined search terms is that they also result in many false positive literature hits because the terms *data classification* is used in machine learning and *asset classification* is used in asset management. It was also decided on an exploratory literature search by including other related subject areas in the literature search. The assumption was that the topic of *information classification* does not receive enough attention in science but may still be considered as part of risk management, although not in the primary question, but rather in the context. The idea was to get secondary scientific considerations of information classification in the context of risk management, and that this could possibly lead to new insights.

Therefore, the new areas of expertise in exploratory literature research include:

- Risk management classification
- Risk management categorization
- Information risk classification
- Information risk categorization
- Information risk management
- Risk classification
- Risk categorization
- Risk management process
- Risk management methods
- Risk classification standards
- Risk analysis
- Risk analysis methods
- Information security classification
- Information security methods
- Information security standards
- Security classification
- Security categorization
- Data, information, knowledge
- Classification decisions

The expansion of the literature search with the topics of risk analysis, risk assessment, risk management and information security was obvious because information classification serves as a preparation for risk analysis, the risk analysis is part of risk management, and this in turn is part of information security.

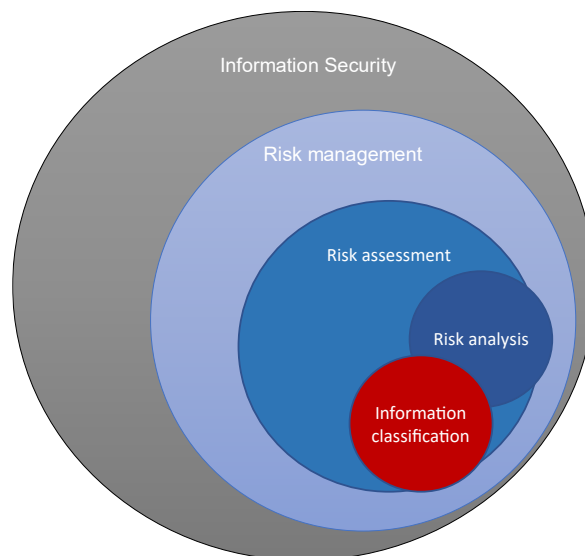


Figure 4: Literature search expansion.

Samples from other literature research platforms such as *Semantic Scholar* and in the *Startpage.com* search engine showed that additional literature was found compared to *Web of Science*. To be able to identify as much relevant literature as possible, it was decided to search via *Web of Science*, *Semantic*

*Scholar* and *Startpage.com*. The risk management standards have been obtained from the *NIST publications website*<sup>1</sup> and *The Swedish Institute for Standards (SIS)*<sup>2</sup>.

As expected, the false positive rate was high and had to be filtered by evaluating the content. This was a very time-consuming work, but due to the small yield of literature directly on the topic of *information classification* and its importance in risk management, it was necessary and useful for answering the research questions.

The literature search process described is visualized schematically in the next figure.

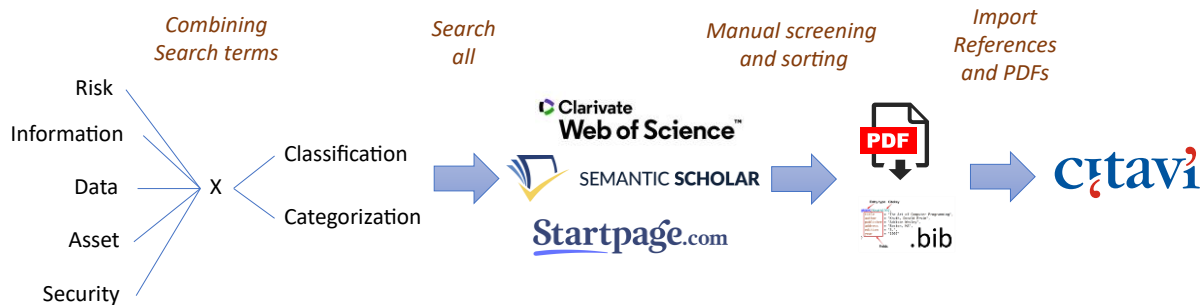


Figure 5: Expanded literature search process.

The analysis part brings together all the results of the qualitative literature analysis. The results are summarized to provide an overview of information classification in risk management.

### 2.3.2. Literature sources

Specialized literature search engines or web search engines to search for literature have been used. In doing so, the intention was to avoid searching on every publishing platform, which would have taken too much time and limited the results to the literature available there. The (literature) search engines used included:

- Web of Science
- Semantic Scholar & Google Scholar
- Startpage.com
- NIST platform, and
- SIS

### 2.3.3. Literature statistics

Because the literature search has been expanded across multiple platforms with different search settings and results, the search results had to be documented manually. Therefore, the statistical data has been limited to the bare minimum. With the extended literature search a total of 2893 references have been found were also the full texts have been available for download. The most productive source was *Semantic Scholar*. The standards were found via the NIST platform and SIS because standards are almost not represented on the literature search platforms.

<sup>1</sup> <https://csrc.nist.gov/publications>

<sup>2</sup> <https://www.sis.se/en/standards/>

Table 2: Results of extended literature search.

Search platform	Quantity
Web of Science	823
Semantic Scholar	1060
Startpage.com	808
Standards	202
<b>Sum</b>	<b>2893</b>

We could not find any full texts for 65 references identified (were not counted).

Using the literature management tool *Citavi*, a full-text search for the terms *classification* and *categorization* across the entire literature has been carried out. The locations were then checked for relevance and added to a defined Citavi category. As a result of this manual text analysis, a total of 114 references relevant to information classification were identified. *Table 3* illustrates this, including a breakdown by the Citavi categories *standards*, *scientific literature*, *books*, and *practice*. The *practice* category is scientific literature that has examined information classification in organizations.

Table 3: Results for relevant information classification literature.

Category	Quantity
Standards	38
Scientific	46
Books	23
Practice	7
<b>Sum</b>	<b>114</b>

This means, the initially available literature on information classification has been increased from 31 to 114 and an inventory of 2779 full-text searchable documents for context on information classification has been created.

### 3. Qualitative literature review

In this part of the work, the relevant literature has been analysed to identify and describe the state of the art and potential weaknesses and gaps.

#### 3.1. Information classification

According to the current information security processes of the information security standards, the information classification is based on the expertise of the information owner and security experts and is therefore a decision made by one or more people. Ideally, it is assumed that there is enough knowledge and information to make an informed decision in terms of organizational goals and risk preferences. Information classification has its origins in the military, with a focus on secrecy, hence confidentiality is of the highest priority (Fibikova & Müller, 2011). And classification was a paper based task executed by few people (Bunker, 2012). However, the military may have different requirements and priorities for information security than other organizations, and processes based on paper-based workflows must not be suitable for digital data and processes (Bergström, Karlsson, & Åhlfeldt, 2020). But the classification of information is necessary because security measures are cost-intensive, and

security measures are invested by classifying and only for the security-relevant information (Miller, 2017).

### 3.2. Information classification in standards

Information classifications are derived from applicable federal laws, executive orders, directives, policies, regulations, standards, and guidance (Joint Task Force Transformation Initiative, 2013).

These standards address classification categories definition and classification procedures (Barry L. Williams 2013):

- Standards having information classification methods:
  - ISO/IEC 27001, A.7.2.1
  - NIST SP 800-53, RA-2
- Standards having information labeling and handling procedures:
  - ISO/IEC 27001, A.7.2.2
  - NIST SP 800-53, AC-16
  - PCI DSS V2.0, 9.7.1

An analysis and abstraction of the information security risk management approaches of the standards *NIST SP 800-30*, *ISO 27005*, *EBIOS*, *OCTAVE*, *CRAMM*, *FAIR*, *ISAMM*, *ISF - Standard of good practice*, did not reveal any major differences in the methodology. They have in common asset identification and classification, identification of organizational objectives, threats and vulnerabilities, establishment of security measures and controls, and impact analysis and risk determination (Fenz et al., 2014). In general, the standards are very similar, but in terms of actual implementation there is no common approach in standards to information security, also not for information classification (Bergström, Karlsson, & Åhlfeldt, 2020). „The ISO/IEC 27001 standard doesn't say much about information classification... so the details of how you implement the control are pretty much left up to you“ (CertiKit).

*NIST SP 800-60* states that information should also be sorted by type. The level of detail of the data granularity is not specified but is determined "arbitrarily" according to any criteria. The result is a catalog or a taxonomy of the data/information. Information types could be electronic data (stored, transmitted, processed, or communicated [email, text message]) spoken data (video conferencing, phone), multimedia information (video, surveillance), physical information (paper documents) (Stallings, 2018).

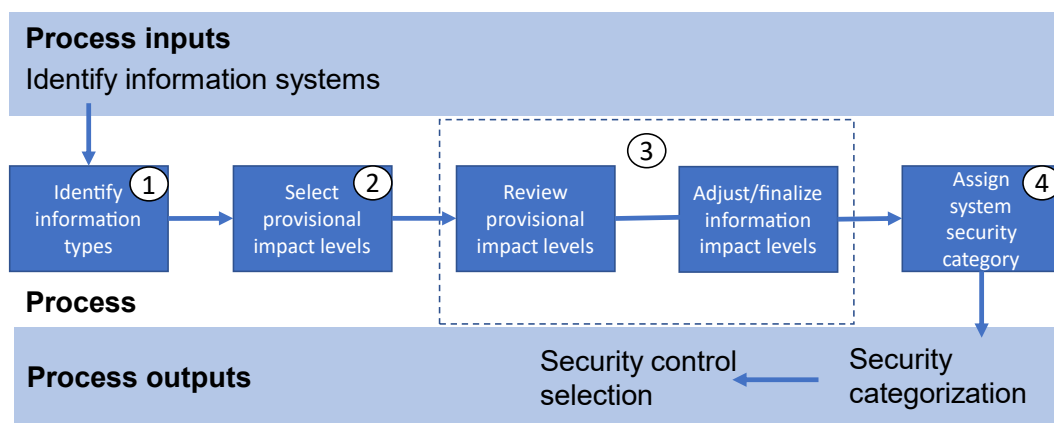


Figure 6: Security Categorization Process (Stallings, 2018) .

ISO 27001 places the following requirements on information classification (Stallings, 2018):

- Information classification must consider legal requirements, the value, the criticality, and the sensitivity of the information.
- Information marking for classification.
- Processes for handling classified information.

It must be ensured that the information classification was determined based on the information value, legal regulations, sensitivity, or criticality. In addition, procedures for labeling and handling the information must be available (Kumar, 2011).

The *Agence nationale de la sécurité des systèmes d'information (ANSSI)* classification focuses on the protection of complex systems and is intended to help classify them based on their exposure. Nevertheless, the ANSSI classification method is also relevant for information classification because this method can also be used directly for information; there is information connectivity to databases, applications or systems, and functions that work with the information. There are information users, and attackers interested in the information. All these factors determine the likelihood. And then there is the impact in the event of an occurrence. With the ANSSI method, the information classification could be placed on a broader database and analysis.

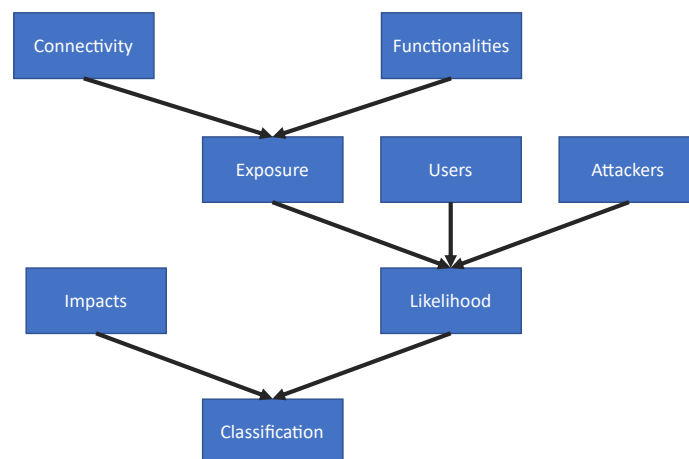


Figure 7: ANSSI Classification Method (Shrestha et al., 2017).

Decomposing information into relevant sub-components can facilitate information classification (Joint Task Force, 2020). The *European standards organization in ICT (ETSI)* describes it in a similar way: „Once the sensitivity of the data has been identified, the data need to be traced back to business applications and the physical servers that house those applications” (ETSI, 2016, p. 34). „At a base level, a data classification scheme is broken down into two levels: public (unclassified) and private (classified). Once the private information has been identified, it can then be further subdivided based on the impact it would have to the organization if it were compromised” (ETSI, 2015, p. 52).

Most classification methods are simple classification lists, with these most common classification levels (CertiKit):

- Top secret
- Secret
- Confidential
- Restricted
- Protected
- Internal Use Only

But there are also differences, as with classification types from the *Department of Defense Architecture Framework (DoDAF)* (OMG, 2013):

- C - Confidential
- CTS - COSMIC TOP SECRET
- CTS-B - COSMIC TOP SECRET - BOHEMIA
- CTS-BALK - COSMIC TOP SECRET - BALK
- CTSA - COSMIC TOP SECRET ATOMAL
- NC - NATO Confidential
- NCA - NATO Confidential Atomal
- NR - NATO Restricted (similar to US For Official Use only)
- NS - NATO Secret
- NS-A - NATO Atomal
- NS-S - NATO Secret
- NSAT - NATO Secret Atomal
- NU - NATO Unclassified
- R - Restricted Data (RD) US Nuclear Information OR FOR OFFICIAL USE ONLY
- S - Secret
- TS - Top Secret
- U – Unclassified

For general audiences (International, Dama 2017):

- Internal use only
- Confidential
- Restricted confidential
- Registered confidential

And even more classification examples:

- “State/government organizations – formal
  - Top secret
  - Highly secret
  - Confidential
  - Sensitive
  - Restricted
  - For XX eyes only
  - Need to know
  - Classified
  - Redacted
  - Protected
- Others primarily in non-state organizations – formal
  - Not for distribution
  - Private
  - Business use only
  - Internal or internal use only
  - Non-disclosure
  - Employee confidential
  - Proprietary
  - Privileged
  - Sealed
  - Trademarked
  - Copyrighted

- Other potentially restrictive terms – informal (in addition to all formal terms)
  - Personal
  - Draft
  - Preliminary
  - Limited
  - (Un)edited
  - Unofficial
  - Unauthorized
  - Void
  - Invalid
  - Banned
  - Forbidden
  - Blocked
  - Locked
  - Secured
  - Reserved
  - Censored
  - Embargoed
  - Do not copy
  - Off the record
  - High risk
  - Recalled/retracted
  - Original
  - Encrypted
  - Password-protected” (Scott & Choi, 2017)

The consequence of these immense variations in classification categories is that the classification information can only be exchanged with difficulty, and a uniform best practice solution also seems difficult to implement. *MAGERIT*, a risk management framework developed by the Spanish Ministry of Administration, contains proposals for unifying the exchange formats between organizations to provide asset identification, asset type classification, asset dependencies, and asset value estimation (ENISA, 2022).

The final information classification is based on the highest rating of the impact on confidentiality, integrity and availability (Stallings, 2018) and connected systems must meet security requirements corresponding to the highest information classification of all information on the connected systems (National Institute of Standards and Technology, 2006).

Classification of data can be realized in a risk matrix by their risk into three or five categories, from 1 as lowest to 3 or 5 as highest. To each classification type mandatory controls and protection measures must be assigned and integrated into existing authentication system (Stallings, 2018).

Classified information must be labeled so that it is immediately visible. The labelling must be an everyday task in the standards processes of the organization (Miller, 2017). Other classification recommendations, as from the *NISTIR 8112* schema provides a metadata matrix, with attributes and values. One metadata attribute is classification with the values *Unclassified*, *Controlled Unclassified*, *Confidential*, *Secret*, *Top Secret*, *Company Confidential* (National Institute of Standards and Technology, 2018). In addition, a direct connection between the classification attributes and the access control of users or applications to the data is established with supplementary standards (Stallings 2018):



- SP 800-162, Guide to Attribute Based Access Control (ABAC) Definition and Considerations,
- SP 800-178, A Comparison of Attribute Based Access Control (ABAC) Standards for Data Service Applications,
- SP 1800-3, Attribute Based Access Control,
- NISTIR 8112, Attribute Metadata.

The process of information classification is also defined within the software development life cycle. To incorporate information security into the software development lifecycle process, *NIST SP 800-64* developed a software development lifecycle model. After initiating the software development project, the first step is to categorize the information system by identifying and categorizing information that will be stored, transmitted, and processed. This should result in an information taxonomy or catalog (National Institute of Standards and Technology, 2008a). However, the classification methodology is not different.

COBIT 5 talks about *data classification* and *system classification*, which are part of the asset management policy (Stallings, 2018). The ISO 17799 standard distinguishes between *asset classification* and *information classification* (Peltier, 2001).

However, it is noticeable that the terms are not strictly exaggerated because, in addition to *information classification*, *data classification/categorization*, *asset classification/categorization*, *information asset classification/categorization* and *security classification/categorization* are also used synonymously, as here for example: „Security classification: The grouping of information into classes that reflect the value of the information and the level of protection required. Also called security categorization“ (Stallings, 2018, p. 243).

Responsible for the information classification is the information owner, which is the „official with statutory or operational authority for specified information and responsibility for establishing the controls for its generation, classification, collection, processing, dissemination, and disposal“ (Joint Task Force Transformation Initiative, 2011, p. 64). Business processes across several organizations also include data exchange with, among other things, sensitive data. With the transmission, the data of an organization also leaves its security architecture and is integrated into a new one. In the new organization, however, the standards of value, the classification parameters and labels are also different. In addition, a security incident in the new organization may also have endangered the transmitted data. In addition, the transmitted data can in turn be transmitted by the receiving organization to a third organization (Cherdantseva, 2014). But „care should be taken in interpreting classification markings assigned by other members of an information sharing community“ (International Organization for Standardization, 2015b, p. 6). This means that information-sharing organizations work based on different classification matrices, and these may not be compatible. However, it can also mean that different people within the same organization can come to different classification conclusions. The information owner is usually not a security specialist. Therefore, the information security officer must determine the relevant data for information classification with the information owner. Information-Gathering techniques are (National Institute of Standards and Technology, 2002):

- Questionnaire
- On-site Interviews
- Document Review
- Use of Automated Scanning Tool

But there is also the approach of specifying the information classification centrally, i.e. carrying it out in advance, and providing the executing departments with a table with specified security classifications for a wide variety of information types (National Institute of Standards and Technology, 2008c).

As already stated, information classification is never an isolated consideration, but always to be analyzed in the context of the systems, processes and people involved. A system view of the IT

infrastructure involved is therefore relevant. Information about a system that handles the relevant information can be obtained through interview questions about (Stoneburner et al. 2002):

- The mission
- The system purpose,
- The relation of system to mission,
- Importance of the system
- System availability,
- Incoming and outgoing information,
- Generated, consumed, processed, stored, accessed,
- The importance of the information to the mission,
- Flow of information,
- Types of information (financial data, personal data, medical data...),
- Sensitivity or classification of the data,
- Information that may not be published and the prohibited group of people,
- The processing and storage location,
- Information storage types,
- Effects of unauthorized publication of the information,
- Requirements for information availability and integrity,
- Effects on the organization and mission in the event of unreliability of the system or information,
- The tolerable system downtime and potential alternatives,
- Potential risk of accidents in the event of a system failure.

Information classification affects the entire software and IT architecture. For instance, network security must adapt to the information classification of the stored or transported data in the network (International Organization for Standardization, 2015b). *ISO 27001* compliance requires that all important assets are identified, subjected to a periodic security classification and that usage regulations are in place (Kumar, 2011). Important assets include those with important information in storage, in processing, or in transit. The network access must be restricted to meet the information classification security requirements (Rutishauser, 2012). The networks must be separated in such a way that they optimally meet the security requirements from the information classification (Miller, 2017). The network then needs to be segmented so that systems of the same sensitivity level are on the same network and segmented from systems with different trust levels. If possible, firewalls need to control access to each segment. If data are flowing over a network with a lower trust level, encryption should be used (ETSI, 2015). Not only the network architecture depends on the information classification, but also the strength of the cryptographic algorithm applies to the information classification (Joint Task Force, 2020).

To be able to cope with the whole complexity of requirements, effects and dependencies surrounding information classification systems engineering principles should be applied to meet information classification requirements. Here, reference can be made to the descriptions of the system risk assessment in the *NIST SP 800-160* standard (Miller, 2017).

The disposal or reuse of data media must conform to the different security requirements of the information classification categories. The higher the information classification, the more rigorous the erasure and destruction requirements upon disposal or reuse for other purposes (Miller, 2017).

The classified information must also be related to the applications that store, process, obtain, delete, or transfer data. Information classification and the related application must be integrated in the information security architecture (International Organization for Standardization, 2015a).

Storage security should be adapted to the security needs of information classification (International Organization for Standardization, 2015c).

Information classification also plays a role in telework, because the classification of the processed data should be recorded (International Organization for Standardization, 2013a).

„The strength of user authentication should be appropriate for the classification of the information to be accessed“ (International Organization for Standardization, 2013a, p. 26).

Even data centers have their own classification matrices to meet different security requirements, which in turn can be found on information classification (International Organization for Standardization, 2018b). But also cloud services must be subject to the same security requirements for classified information as on-premises solutions. As with external service providers, service level agreements (SLAs) play an important role in the cloud. It is the responsibility of the customer to provide the data, the data types, the classifications, and the importance of the data to the cloud service provider. The customer must also specify the risks in relation to the information. The customer is also responsible for the risk evaluation. The cloud service provider delivers the appropriate (security) services based on the SLAs (International Organization for Standardization, 2016).

Access of roles to application data is set based on the information classification (International Organization for Standardization, 2018a). But organizations and individuals can have different and multiple roles in complex multi-directional data processing processes with other individuals, organizations and sub-service providers and under different legislative frameworks (National Institute of Standards and Technology, 2020), which can make identification and implementation quite difficult.

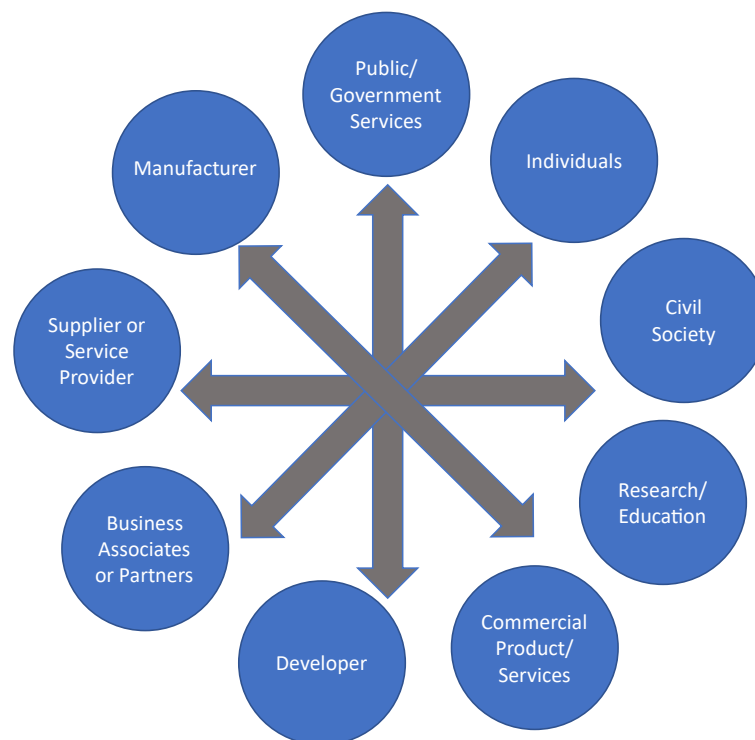


Figure 8: Data Processing Ecosystem Relationships (National Institute of Standards and Technology, 2020).

Information classification also influences the job and role description of personnel, not only in terms of the information owner being responsible for the information classification life cycle, but also that higher security requirements are placed on personnel and the job for certain classification categories (ETSI, 2015). Information classification can lead to the necessity of screening personnel for their background in order to fulfil security requirements by laws or regulations (Longras et al., 2013).

However, information is not only generated within the organization but can also be obtained from external sources or passed on to third parties. Therefore, the information management process between acquirer and supplier must be established, with one of the goals being the adoption of the security requirements of the respective information classifications. The control mechanisms to be considered are information classification, access controls, cryptography, backup, and information transfer (International Organization for Standardization, 2014). Classification schemes between companies with information exchange should be mapped (International Organization for Standardization, 2013a).

Also, personally identifiable information (PII) should be an explicit part of the information classification scheme (International Organization for Standardization, 2021). But when classifying PII, however, country-specific legal requirements and different types of PII, such as general PII or health data, must be considered. i.e. the same piece of information can be classified differently in different contexts and under different legal frameworks (International Organization for Standardization, 2021). Here, the *NIST Privacy Framework* is a relevant source for identifying relevant information for privacy classification, which collects the following data in the information collection phase (National Institute of Standards and Technology 2020):

- Identify
  - Inventory and mapping
    - Data processing by
      - Systems
      - Products
      - Services
    - Owners or operators and their roles with respect to
      - Systems
      - Products
      - Services
      - Components
      - Owners/operators can be:
        - Organization
        - Third parties:
          - Service providers
          - Partners
          - Customers
          - Developers
  - Categories of individuals whose data are being processed are inventoried:
    - Customers
    - Employees or prospective employees
    - Consumers
  - Data actions of the
    - Systems
    - Products
    - Services
  - Data action purposes
  - Data action elements
  - Data processing environment
    - Geographic location
    - Internal
    - Cloud

- Third parties
- Data processing
  - Data actions
  - Data elements
  - Systems
  - Products
  - Services
  - Components
  - Roles of the component owners/operators
  - Interactions of roles with systems/products/services

Organizational measures to support the enforcement of information classification could be to create separate acceptable use policies for each information classification category and associated system assets (Miller, 2017).

Information classification changes are sometimes necessary, as information and information classification have their lifecycles, but must be evaluated by a board or higher authority (National Institute of Standards and Technology, 2008c).

Biases, assumptions, beliefs, limited knowledge, and information reliability are risk factors themselves (International Organization for Standardization, 2018c).

Information classification is not to be considered separately but embedded in information security processes and frameworks. An information security framework is the *NIST Cybersecurity Framework* (National Institute of Standards and Technology). It provides different levels of maturity in cybersecurity; partial, risk-informed, repeatable, and adaptive. Whereby only the adaptive level as the highest level of maturity can react dynamically and promptly to new dangerous situations. The prerequisite for this is that cyber security management, policies, and processes can take security measures based on dynamic and networked information (Stallings, 2018). A possible analogy between the maturity levels of the cybersecurity framework and information classification is that information classification also requires a life cycle, inclusive active and anticipatory approach. An information classification requires a broad and well-founded database, about assets, vulnerabilities, dangers, risks, and expert knowledge in the field of information and information security. However, this complex information landscape for well-founded information classification also requires a very high degree of maturity in the organization, the processes, data collection, tool usage and information security.

### 3.3. Standards challenges

As early as 2003, *Siponen* stated that the assumptions, recommendations and specifications in security standards and their maturity models neither follow scientific methods nor are comprehensible (Siponen, 2003). It could be assumed that the normative standards arose from subjective experiences and preferences and are therefore highly irrational. In addition, he noted, that the standards do not sufficiently consider the different requirements of the organizations applying them. They are more subject to the error of inadmissible generalization; derive general validity from individual cases (Siponen, 2003). The normative character also would lead to compliant security implementation and prevents innovation. There is also a lack of serious validation of the standards for actual functionality, meaning that evidence is lacking. *Siponen* points out that it is of the utmost importance to explain why a security measure should be taken, what its actual (real) effects are and what the underlying implications are (Siponen, 2003). This requires rigorous empirical research in which all variables must be examined. Objects of investigation should be (Baskerville, 1993; Siponen, 2003):

- What security techniques and methods organizations use.

- What actual effects and possible weaknesses these methods have in practice.
- What is the perceived usefulness and usability.
- What are problems and effects of using entire standards in real organizations, such as, how can standards be integrated into systems development or administration activities.

Security issues increase with e-government, as country-specific standards and certifications, with different approaches to information classification schemes, information storage strategies, information reliability assurances, security procedures, risk management strategies, and data inconsistencies (Abdugaffarovich et al., 2015).

A particularly serious problem is the focus of the information security standards on personal and organizational security methods and solutions. This is a problem because it has been known in engineering, and especially security engineering and Occupational Safety, for decades that organizational measures such as guidelines, processes, instructions, or people represent the most unreliable and weakest form of security measures (Centers for Disease Control and Prevention, 2023). Technical measures are significantly more reliable, and measures based on the laws of nature are the most reliable. This is also shown in *Figure 9*, namely that the risk and costs increase as the proportion of manual work increases. With an increasing degree of automation, both the risk and the costs decrease. In addition, the automation also increases the level of information security maturity by switching from a reactive manual action to a preventive automated action. But it could not be determined why the automation in information classification and risk management is marginalized.

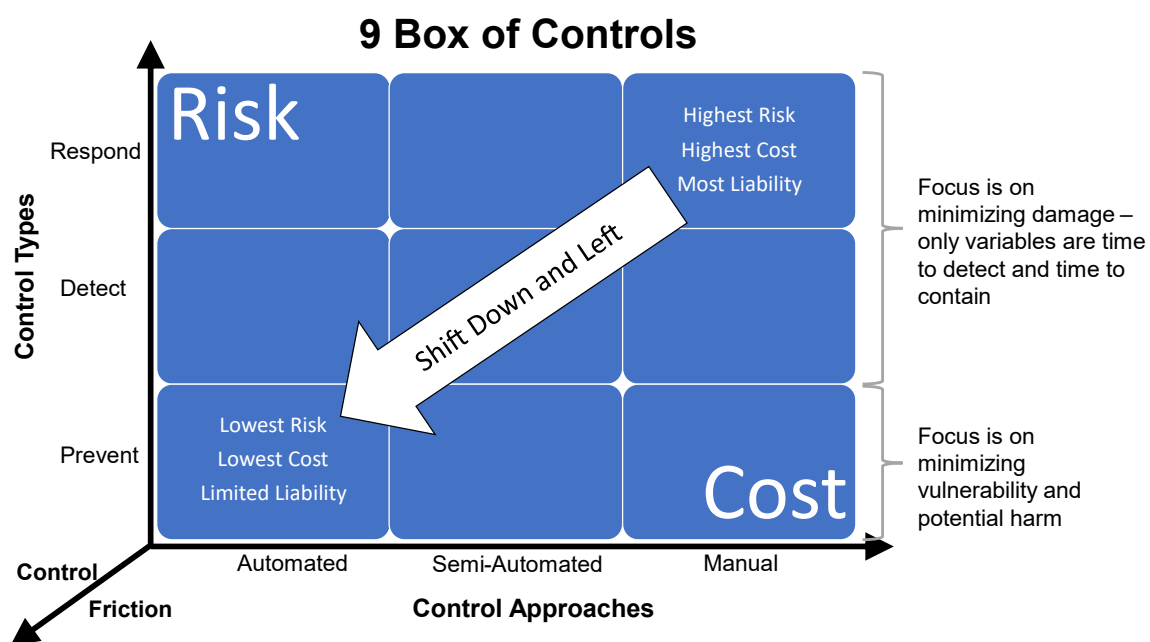
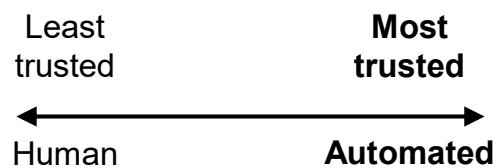


Figure 9: The 9 Box of Controls (Harkins, 2013).

It is clear from the prevalence of breaches and compromises that risk management and information classification have not kept up with the threats, and it appears to be slipping farther behind as the world grows more volatile, uncertain, and ambiguous. As the world of technology expands exponentially, so do the technology-related threats and vulnerabilities, yet the ability to manage those security and privacy risks has progressed only at a slow rate. As a result, there is a widening gap between the risks and the controls (Harkins, 2013). An example of slow progress in cyber defense is that changes in information (value) are not addressed at all automatically (Bergström et al., 2018).

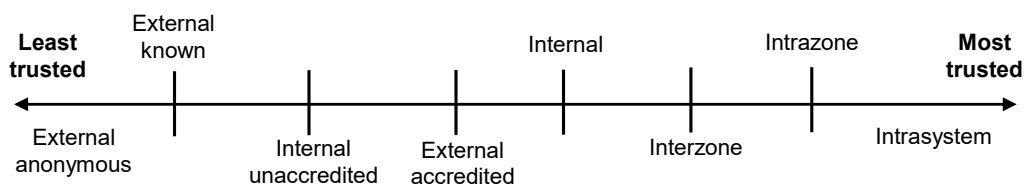
Another problem perspective is that of trust. In implementing information classification and risk management, certain employees are empowered to make far-reaching and important decisions for

the company. Information owners are authorized to classify information and thus also to influence system security architectures. This decision-making power is curbed with only weak control measures, where the security expert only checks the conclusiveness but not the correctness of the classification. This approach requires a great deal of trust and is itself a great risk. The question is whether this trust is justified or whether other measures would reduce the risk and the risky part, namely the classification of personal information, should be minimized. *Figure 10* is intended to illustrate this aspect. With his approach, *Wheeler* describes the trustworthiness of information flows and explains that human-triggered information flows are less trustworthy because human actions are error-prone, and interpretation is more difficult. Automated triggers, on the other hand, are rarely subject to errors and follow a clear pattern that is predefined and therefore easily interpretable (*Wheeler, 2011*).



*Figure 10: Degrees of trust (Wheeler, 2011).*

A similar perspective is also used for internal and external assets. Here, *Wheeler* focused on devices. External devices are considered service provider operated devices and internal devices are operated by the own organization. The trustworthiness of the devices decreases the more unknown and external a device is. This would particularly affect the most diverse as-a-service models, such as cloud services, since according to this model they are regarded as not or not very trustworthy. This underlying principle of internality and externality can also be extended to service providers in general, i.e., also to the people of the service provider. According to this, internal employees would be more trustworthy than people from service providers or other cooperation partners. This perspective and the conclusions of this approach are also relevant for information classification because data, information on external or external devices must be classified as less secure, i.e., subject to a higher risk. Among other things, it could be analyzed here which classification category information may have at most to be allowed to be processed, stored, or transferred at all by external parties.



*Figure 11: Degrees of endpoint trust (Wheeler, 2011).*

The CISOs don't have the skills necessary for the risks of the 21st century. The skills should include broad technological and business knowledge and in-depth knowledge of risk and security (*Harkins, 2013*). This argument is also supported by *Table 4* of survey results at small and medium enterprises (SME). The biggest influencing factor on information security in SMEs is the lack of knowledge and experience about information security.

Table 4: Respondents replied to the question: Please specify to what extent (1- smallest, 5-largest) in your enterprise the information security may be influenced by the following factors (Kobis, 2020).

1	social engineering activities	1.9
2	lack of knowledge	3.8
3	lack of experience	3.2
4	indifferent attitude to the work performed	1.6
5	fatigue	1.4
6	intentional, harmful actions of employees	1.1

With this knowledge, the question may be justified as to whether the risk management standards are pursuing the right approach at all, if the lack of knowledge is criticized despite the great effort and the immense quantity of risk management documents and recommendations. The knowledge of risk management does not seem to reach the SMEs, and the reasons for this remain to be determined.

In addition, all organizational information must be classified by the information owner, they must also be trained on how classification works. This approach is very time-consuming and error-prone for subjectivity and inconsistency (Bergström & Åhlfeldt, 2014).

It seems that it must be analyzed whether the very simplified approaches in information classification and in risk management in general are justified. This oversimplification of risk management processes in practice leads to gaps in the implementation of the required risk management, no standard-based processes, in too little granularity of the analysis of assets, risks, vulnerabilities, and threats, which in turn leads to wrong assumptions and conclusions (Shedden et al., 2006).

There is little literature on the actual implementation and use of risk assessment in organizations. In addition, there is no scientific proof of how effective risk management recommendations from science and standards actually are (Shedden et al., 2006).

In summary, the risk management standards seem too incomprehensible to be put into practice with confidence. They are very people-oriented and focused on manual activities. They oversimplify important analysis steps that could generate the greatest information content and thus knowledge gain, especially in information classification. Information classification is extremely time consuming because the classification process is manual. In addition, there is an enormous organizational and regulatory effort, with dubious benefits and scientifically unverified methods and processes. „Nevertheless, the gap between standard, and practice in the area of information classification has not attracted much attention from scholars“ (Bergström et al., 2018).

The use of information security standards is often assumed to be easy to use and thus implementable for risk management. The impression is also strengthened with the help of simple and extensively described risk management processes. However, studies by *Shedden et.al.* show that this is not actually perceived as such in practice (Shedden et al., 2006).

### 3.4. Information classification in scientific research

„The primary purpose of classifying information is to identify and lock down sensitive content appropriately“ (Rangel, 2019, p. 10). The most important driver of information classification is the public sector, which promotes its own classification efforts through laws on information security and demands specifications for subcontractors (Bergström et al., 2018). Information classification is crucial for an effective information security management system (ISMS) (Costin & Militaru, 2011). Information classification can have many names, such as data classification or security classification, and naming seems to depend on the context (Bergström et al., 2018).



An information classification can take place according to different objectives (Chandramouli and Pinhas 2020):

- To determine the sensitivity of whether certain security measures are necessary, which in turn can be subdivided according to various requirements, such as for personally identifiable information (PII), Health Insurance Portability and Accountability Act (HIPAA)-related, and Payment Card Industry Digital Security Standard (PCI -DSS).
- For frequency classification to be able to identify appropriate storage media.
- For environment classification, which can be both frequency and sensitivity classification.

*Bergstrom et. al.* identified information classification methods and presented their degree of implementation with regard to the parameters: procedures, adaptability, description of the roles involved, adjustable classification granularity (Bergström, Karlsson, & Åhlfeldt, 2020). From their summary table, the information classification methods identified by the authors meet the required criteria to varying degrees and that there is no one method with full satisfaction. Regardless of whether the criteria specified by the authors are relevant, *Table 5* shows that the amount of information classification literature is small, and the classification methods have different methodological approaches.

*Table 5: Summary of the analysis of how existing methods for information classification support stated MRs (Bergström, Karlsson, & Åhlfeldt, 2020).*

	<b>Examples of existing information classification approaches</b>	MR1	MR2	MR3	MR4
<b>Standards</b>	ISO/IEC 27000 family (ISO/IEC 27002, 2017; ISO/IEC 27005, 2018; ISO/IEC 27001, 2017; ISO/IEC 27003, 2017)	*	+	-	+
	FIPS 199, National Institute of Standards and Technology (2004)	*	+	-	+
<b>National approaches</b>	Cabinet Office (2018)	*	-	-	-
	Oscarson and Karlsson (2009)	*	*	-	-
	Andersson et al. (2011)	*	-	*	*
	Swedish Civil Contingencies Agency (2018)	*	+	*	*
	Farn et al. (2008)	*	-	-	-
<b>Other approaches</b>	Fibikova and Müller (2011)	*	-	-	*
	Tatar and Karabacak (2012)	*	-	-	-
	Xianliang et al. (2007)	*	-	-	-
	Breier and Schindler (2014)	-	-	-	-
	Foroughi (2008)	-	-	-	-
	Leitner and Schaumuller-Bichl (2009)	-	-	-	-
	COBIT (ISACA, 2012; Etges and McNeil, 2006)	*	*	*	*
	Fernandez and Garcia (2016)	-	+	-	-
	Vidalis (2010)	*	-	*	-
<b>Notes:</b> The left column shows the reviewed information classification approaches and the other four columns show how each reviewed approach supports the MRs. How the respective approach supports the MRs are shown using the legend: “p” indicates that the approach supports the MR, “-” indicates that the approach does not support the MR and “*” indicates that the approach supports the MR partly					

An identified information asset should immediately be classified according to its value, by means of a scheme or hierarchical model. Information classification is normally performed using a classification scheme, containing a number of levels describing the consequence of a loss of, e.g., confidentiality integrity and availability (Bergström et al., 2018). Information classification is thus part of asset management and a prerequisite of risk analysis (Bergström, Karlsson, & Åhlfeldt, 2020). „Everything connected in any kind to information technology components has to been seen as an asset, no matter if it’s a physical item like a computer, server, a document, a developer or intangible items like files, source code, data bases or software applications“ (Fenz et al., 2014, p. 419).

Information classification processes still must be flexible to satisfy different organizational views, necessities and technical requirements, and must adopt to the many relationships between the Information Security Management (ISM) activities and the ISM complexity (Bergström, Karlsson, & Åhlfeldt, 2020). *Bergstrom et al.* recommend a two-stage information classification process. The basic process should be roughly classifying, i.e., grouping assets to be classified together. A fine classification should only be carried out on the information asset if necessary, as it is a time-consuming process involving the stakeholders concerned in workshops (Bergström, Karlsson, & Åhlfeldt, 2020). In the next graphics *Bergström et. al.* described an ideal of a basis information classification process.

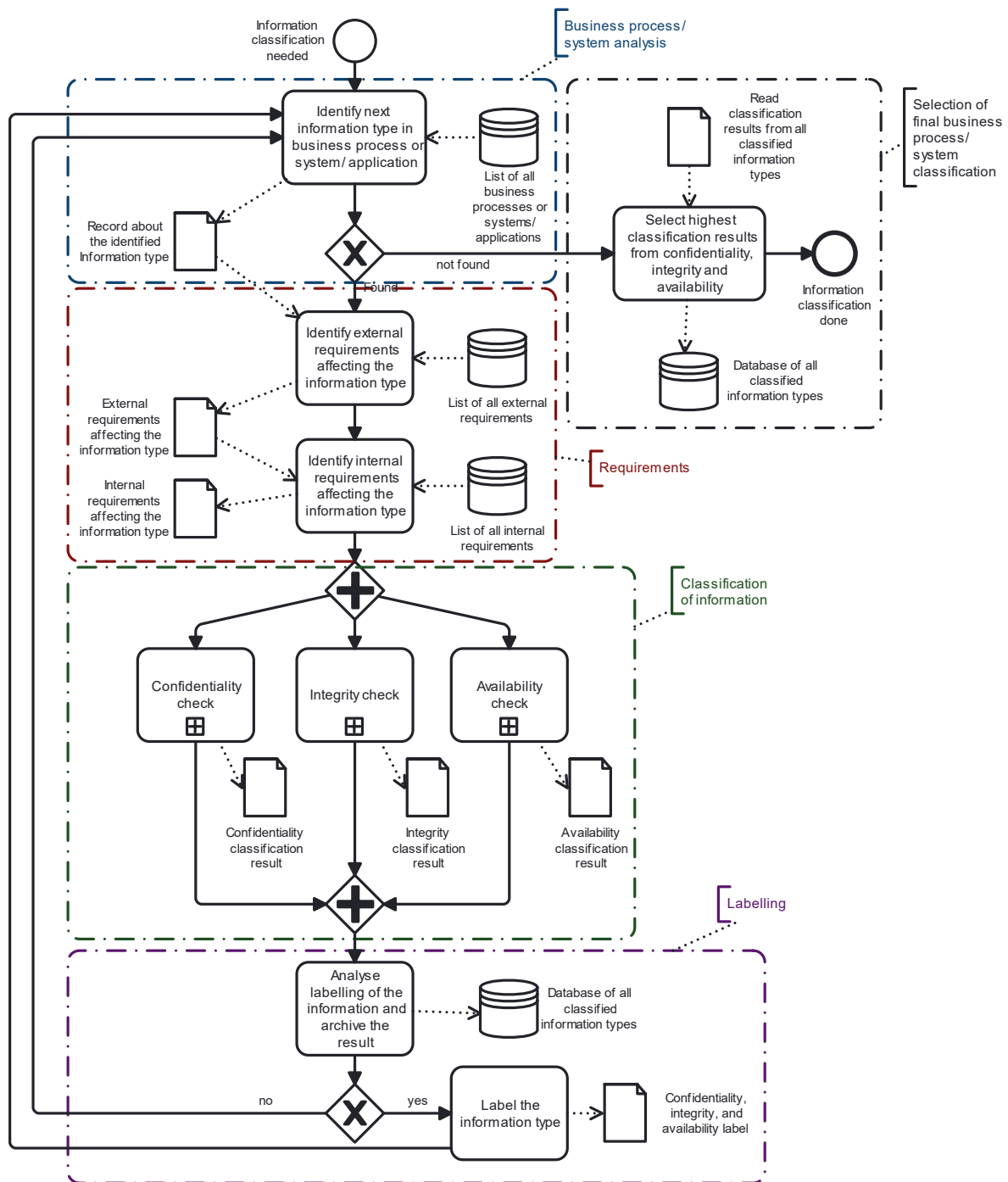


Figure 12: The low granularity route (Bergström, Karlsson, & Åhlfeldt, 2020).

Information classification rules set the countermeasures by technical controls as firewalls, access control and intrusion detection. It also rules procedural measures, as information handling, and

administrative controls (Bjorck, 2005). As early as 2005, a practitioner interviewed said that if you refer implementation of information classification requirements to practitioners, they are bound to fail. Information classification can only succeed if direct technical security measures are implemented in addition to the information classification, so that the user has nothing to do with the classification and implementation process (Bjorck, 2005). In addition, the interviewee stated that, the need for security should be determined based on the potential financial damage. This should be determined on the basis of the threatened products/services, attack scenarios, affected IT systems, affected personnel, current security measures, the delta of the implementation costs of required security measures and the potential financial damage, and the implementation measures for the technical security controls (Bjorck, 2005).

*Yazid et al.* used an asset category list and an information security classification list and combined them with different user groups to arrive at an asset value per user group. Consequently, the asset's value is measured by the security classification of the information stored, processed, or sent by the asset (Yazid et al., 2012). The finding in the paper is that for the same threat and vulnerability, different asset values lead to different risk assessments and thus to different risk mitigation measures and implementation priorities (Yazid et al., 2012).

Asset management is the central prerequisite for carrying out information classification because it identifies the ownership of the assets. Assets include hardware, software, services, people, skills, experience, organization's reputation or image, and information (Bergström & Åhlfeldt, 2014).

The goal of low granularity is to simplify classification by reducing complexity. This is a common approach in standards and literature (ISO/IEC 27002, 2017; Collette, 2006; Fibikova and Müller, 2011; National Institute of Standards and Technology, 2004; Shedden et al., 2016) and practice. But low granularity classification can lead to information overprotection (Bergström, Karlsson, & Åhlfeldt, 2020). However, classification categories that are undefined or have room for interpretation also pose interpretation problems for the user and credibility problems for the categorization (Seifert & Relyea, 2004). Reasons for low granularity may be the requirement that the information classification should be done by the information owner. However, the classification can be delegated. In addition, acceptance problems can arise if a uniform classification scheme is to be implemented throughout the organization (Bergström et al., 2018). Possibly the low granularity is an expression of a lowest common denominator and the attempt not to overwhelm the information owner with complex classification schemes.

Since information classification is considered an integral part of risk management and thus also of asset management, logically information classification should also be subject to a lifecycle that should be managed. Information classification is also part of a risk analysis process and requires input data and produces output data, and uses tools and methods, see *Figure 13*. The most important data for information classification are the information assets.

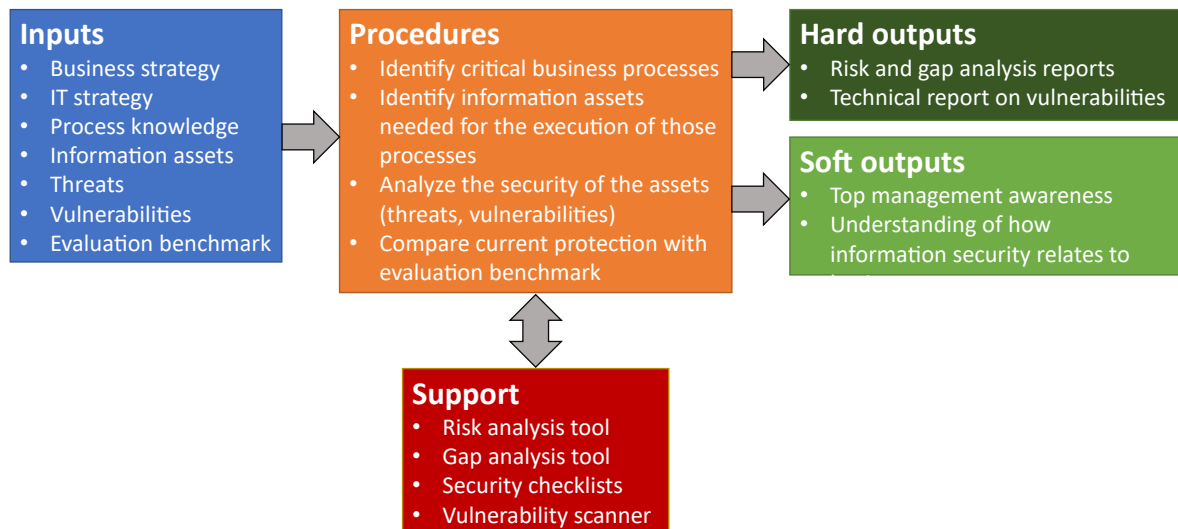


Figure 13: Stages of the ISMS process - The evaluation stage (Bjorck, 2005).

Information Lifecycle Security Risk Assessment steps (Bernard 2007):

- Get a baseline picture:
  - Identification of the entire information lifecycle (creation, storage, distribution, access, use, management, destruction).
  - Identification of all forms of information during the life cycle of the information.
  - Identification of all storage locations of the information forms.
  - Identification of all organizational security guidelines and procedures for the respective information form.
  - Identification of all persons (internal and external) and potential forms of access to the different information in its different forms.
  - Assess the effectiveness of implemented security measures. Audits and inspections of security measures should also be evaluated.
- Conduct any qualitative or quantitative risk analysis method.
- Update information risk management program.
- Apply change management to trigger periodic checks or whenever information asset changes occur.

Information classification can change because the relevance, value, or validity of the information changes (Bernard, 2007). That is why the Information Lifecycle Management (ILM) is needed. It consists of the following steps (Reiner et al. 2004):

- Perform information classification based on value to the organization and considering legal requirements.
- Identify groups of information; Each classified asset can consist of other assets or data or be dependent on them, i.e., there is a logical dependency. Information groups can be, for example: data, directories, file systems, volumes, databases, tablespaces, applications. These groups of information represent the actual value for the company.
- Identify and analyze the current information group services (backup, mirroring, recovery, encryption...). More security causes more costs.
- Apply higher security services to critical more valuable information groups and lower for less valuable. This helps to achieve cost efficiency.

- Apply correct services to the information groups. This means review all requirements and assure all are covered by the services.
- Apply a lifecycle to data, as it can change over time.

Organizations struggle with the information classification lifecycle, as they don't know when to classify and to reclassify (Bergström et al., 2018). Information classification is an important core part of the ILM. From the ILM point of view, there is a need for meaningful classification schemes and automated classification techniques such as pattern recognition (Reiner et al., 2004). But not all information can be supported by automated classification. Information in human memory "... can only be protected by appealing to their human custodians" (Bernard, 2007, p. 29).

*Bergström and Åhlfeldt* identified the following information classification issues in a systematic literature review (Bergström and Åhlfeldt 2014):

- Hard-to-interpret policies can lead to human error, subjectivity, and inconsistent ratings.
- A lack of information lifecycle management prevents correct archiving or necessary reclassifications.
- A lack of resource management can jeopardize the entire classification process due to a lack of resources.
- Inaccurate or too rough description of the information classification process leads to difficulties in implementation and thus to a lack of effectiveness and efficiency.
- Little scientific interest in the topic of information classification in the context of information security.
- Too little practical and empirical research.
- The recommendations of the information security standards are unverified and consequently not well-founded, which puts the entire information security process in doubt.
- A common information classification term is missing.
- The classification takes place based on coarse classification factors.

Information classification must be clearly defined in an organization to achieve common understanding (Stallings 2018). Information classification requires a well-founded classification method because it is a prerequisite and basis, and therefore extremely important for the entire information security management (Shi et al., 2007). The classification level also influences the patch management planning (Souppaya 2022), as because the priority of a patch increases with a higher classification.

In a systematic literature review, *Bergström and Ahlfeld* searched for information classification enablers to reduce the existing problems in understanding and implementing information classification in organizations. They divided the enablers into strategic, tactical, and operational categories and came up with the following results (Bergström and Åhlfeldt 2015):

- Strategic key enablers: key influencers, management support, Information Lifecycle Management, same scheme, and legislation.
- Tactical key enablers: Simplifying the scheme, classify applications/networks, awareness, security culture, best practices, and pilot project.
- Operative key enabler: Staff training, learning, automatic or semi-automatic classification tools, labeling, and granularity

Automation plays an increasing role in information classification. There is already a lot of research, especially on the topics of automatic classification, fine-grained access control mechanisms, information labelling, but however, mostly theoretical considerations (Bergström & Åhlfeldt, 2014). „Automated information classification is big business today and can be used to classify all electronic communications from documents to email to instant messages and even web postings. It has a huge

number of advantages over the old manual based approach...” (Bunker, 2012, p. 20). For ILM to work, given that information changes over time, it needs automated tools (Reiner et al., 2004). A possible automated approach to information classification is classification according to the type of origin. Here, collected data is considered to have the lowest value, processed data to have an intermediate value, and created data to have the highest value (Al-Fedaghi, 2008). However, this approach can only be the first step, as it is too under-complex and too general for practical use. Information classification by experts has several difficulties (Weng et al., 2006). Expert-based manual information classification cannot keep up with the amount of new information. This shortcoming can only be overcome by automated information classification (Weng et al., 2006). Information security is changing from focusing on individual systems to looking at interdependent systems, a system thinking approach (Schiavone et al., 2014). Cyber-attacks are successful because information security still thinks and acts in linear, static and reacting cause-and-effect analyses, and not because the attackers are ingenious (Schiavone et al., 2014).

Information importance can be based on its stage in the information lifecycle, its completeness, ownership, the opportunity costs, potential options, the information exposure to the user (Shi et al., 2007). But certain information can have a different meaning or value for different information users, which can be due to the different use of the information and the different understanding of the information (Shi et al., 2007). A reliable information classification is essential before sharing information with external parties, i.e. in the supply chain (Shi et al., 2007). There is a need for additional information on the information classification to make more informed decisions (Lundgren, 2020). But information security lacks already a uniform terminology (Cherdantseva, 2014).

In a complex environment, models and procedures are important to understand basic concepts of the subject of research. Among other things, the *Reference Model of Information Assurance & Security (RMIAS)* is available for information taxonomy. Although it is on a very abstract level, it is helpful to present the overall context and lifecycle of the information taxonomy, security objectives, security measures and security lifecycle. Information classification can be subsumed here under information taxonomy.

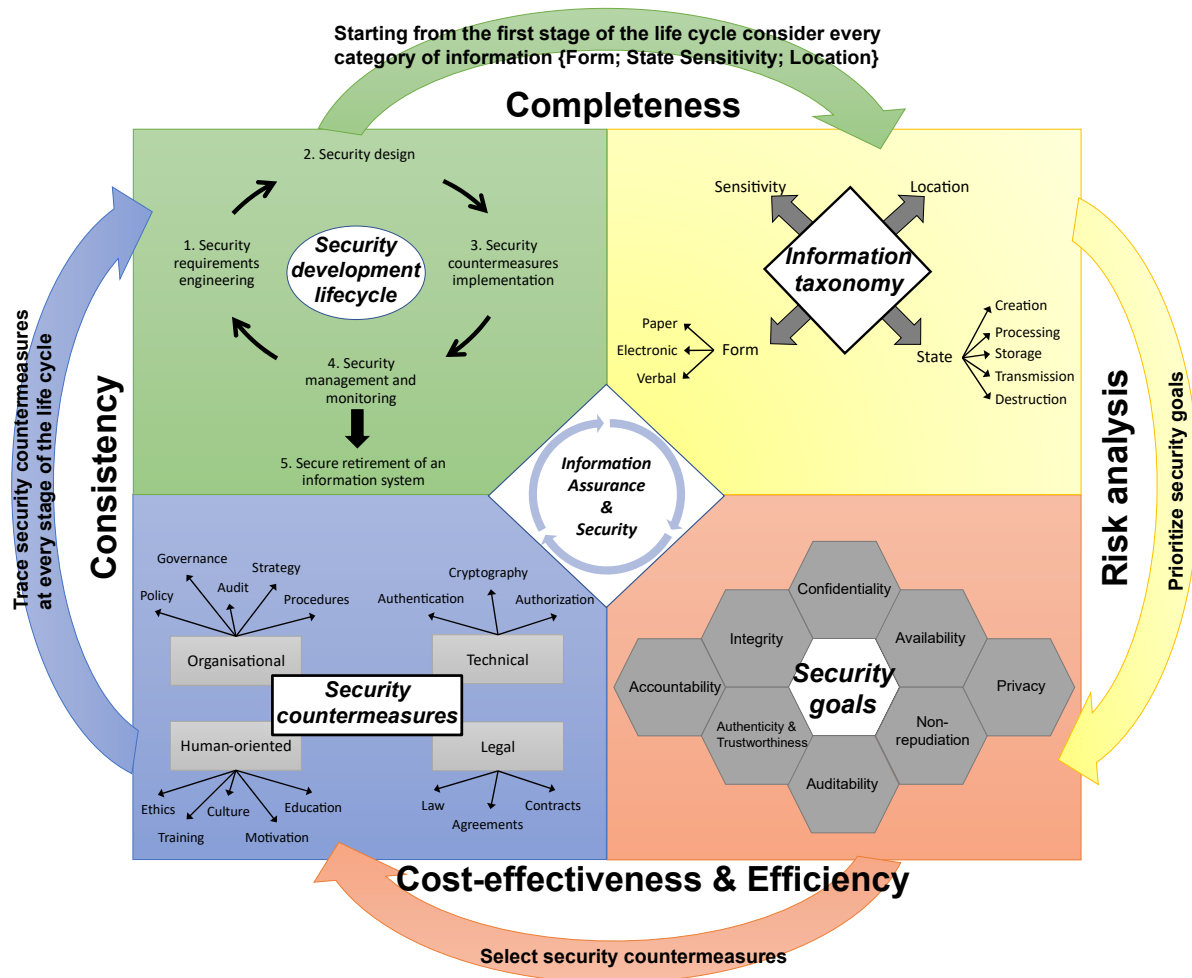


Figure 14: The Reference Model of Information Assurance & Security (RMIAS) (Cherdantseva, 2014).

„Organizations use security management procedures such as information classification, risk assessment, and risk analysis to identify threats, categorize assets, and evaluate system vulnerabilities. These procedures help in implementing effective controls“ (Karoui, 2016, p: 556).

The properties and circumstances of information that change over time are relevant to information classification because they change the conditions for safe operation and the requirements for classification. Information may be in the state of creation, storage, processing, transmission, or destruction. Each information classification level has its own requirements for the respective information states.

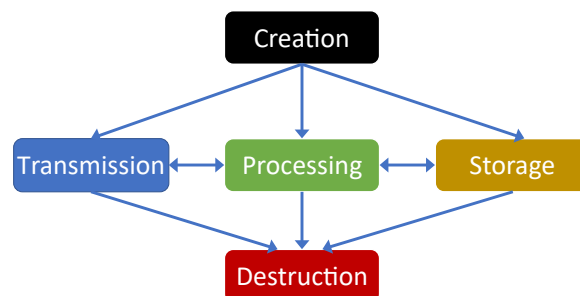


Figure 15: Information Life Cycle illustrates possible states of information (Cherdantseva, 2014).

The problem of bias in information classification is considered again and again. Different standards and different organizations use different classification terms for their information. A distinction must also be made between military, state, and private sector organizations. Therefore, questions arise as

how these classification terms are dealt with if an exchange of information should take place (Scott and Choi 2017):

- How differently are the terms interpreted?
- What different effects do the terms have on different stakeholders?
- What is the personal motivation to classify information.
- Did the classification make sense?
- Which characteristics (information about the data to be classified and about the classification itself) influence the classification decision?
- What organizational and industry factors influence the decision?
- Which factors of the target group influence the decision?
- Are classifications used as a power tool?
- Do classification decisions serve to save face?

### 3.5. Scientific research challenges

The research does not sufficiently consider the discrepancy between the simplistic approach from the standards and the complex organizational practice, but has a very narrow focus on information classification (Bergström et al., 2018). The challenges of information classification are inadequately analyzed in science. Challenges identified include inconsistent classification due to subjectivity, overly complex or inappropriate classification schemes, inadequate description of classification processes or classification criteria, impact of personal worldviews, granularity of classification, impractical best-practice solutions, overly abstract standards, lacking necessary status in practice, lack of information knowledge, problems in interpretation, credibility problems (Bergström, Karlsson, & Åhlfeldt, 2020). *Bergstrom et. al.* claim that there is enough advice on how to approach and implement information classification in a systematic way, but that a methodology to reduce subjectivity in classification is lacking (Bergström, Karlsson, & Åhlfeldt, 2020). Information classification research mainly focuses on qualitative information classification principles and the management of the classified information (Shi et al., 2007).

*Stiponen and Willison* also point out that standards must consider different environmental factors affecting organizations. Research results from the validation of the actions recommended in the standards must be available to practitioners (Siponen & Willison, 2009).

Possible reasons for classification shortcomings are (Berman, 2022):

- Classification classes have no class properties.
- Property negation as classifier.
- Class membership can easily be changed.
- Classification does not enable class behavior prediction.
- Missing class relations.
- Classification complicates domain understanding.
- Classifying unclassifiable objects.

The goals of information security are ambiguous and unclear because no scientific consensus has been found. On the one hand, the CIA triad is considered outdated and no longer adequate, but then the goals diffuse into the most varied combinations of confidentiality, integrity, availability, accountability, assurance, authentication, non-repudiation, authenticity, reliability, effectiveness, compliance, utility, possession, control, authorization, awareness, access, identification, accuracy, administration, information classification, anonymity, audit, safety (Cherdantseva, 2014). There are arguments for a



formalized expertise in information security. However, this is likely to be a very lengthy process and can only be done with the cooperation of many (Cherdantseva, 2014).

It is criticized that risk management processes neglect the social and organizational aspects too much. The reason for this is that even if all information assets were known, not all details would be known and not all perspectives could be integrated. In addition, dynamic conditions would affect information classification as circumstances and information change throughout its lifecycle. With all this complexity in the information classification process, people would not be able to cope anyway (Lundgren & Bergström, 2019). The method from *Bergstrom et. al.* focuses strongly on the process description (in BPMN) of the low granularity route (Bergström, Karlsson, & Åhlfeldt, 2020). From the assessment of the problem, the generic description and the focus on rough information classification groups do not result in any new findings that would help to close the information classification gaps identified by *Bergström et.al.* On the contrary, with the chosen approach, this seems to be just another variant of the previous process descriptions from the ISO 27 family. This is because, like those of *Bergström et. al.* criticized methods, this method also has a rough sequence of steps with generically described data sources and a rough, ultimately affect-based, decision-making process using generic decision models. The previous gaps, such as unspecific processes, are retained because the transformation from text-based descriptions to descriptions using a quasi-standard modeling language only changed the syntax but did not increase the information content. The problem of oversimplification, which leads to unnecessary incorrect classification of assets in a group of assets due to the security requirements of individual assets, is also not solved, as simplified risk analysis models enable laypersons to carry out the risk analysis, but they contain potential false assumptions about reality as well as increased inaccuracy (Fenz et al., 2014). Here, risk analysis and information classification are subject to the same problem. Unfortunately, the problem of identifying relevant sources of information for decision-making is also pushed aside, although it is one of the core tasks in the run-up to a decision-making process. Ultimately, no contribution is made to solving the problem of guaranteeing the objectivity and comprehensibility of decisions, which is what information classification is, or should be.

The fundamental criticism of the functional approach to risk management and the strong emphasis on social parameters as solution approaches (Lundgren & Bergström, 2019) for information, organizational and technical concepts actually reinforce the causes of the problems mentioned by the authors instead of reducing them. Risk management processes from standards appear very formal, giving the impression of a very functional, even technocratic, approach. However, on closer inspection, the descriptions are rather very text-heavy and scattered, which contain very few and extremely imprecise concepts, methods, functions, models, and processes. The fact that the authors assumed that an ideal input and output between the individual process steps would be assumed here cannot be reconstructed from the literature and the standards. The interpretation does not seem to be compelling, which is why the opposite thesis can also be true, namely that these unspecific descriptions of the risk management process in the literature and in the standards are vague due to a lack of specific information on the standard risk management process and the individual processes of information classification, risk analysis and security controls.

„It is of course not possible to develop generic guidelines that fit exactly for all organizations, and each individual organization needs to develop their own scheme with their own classification categories“ (Bergström & Åhlfeldt, 2014). Identical statements like this quote can be found again and again in information security literature (papers, standards, books). Unfortunately, there were no logical explanations or empirically supporting data. This statement therefore must be rejected as it is unfounded and it can also be assumed the opposite thesis, namely that the missing specificity of the classification categories and the resulting need for infinitely many scientifically unfounded classification schemes give rise to problems, namely subjectivity, inconsistency, over- or under-

classification, etc. The same authors have fixed inaccurate information classification processes for some problems. If the authors applied their chain of arguments to the classification schemes as well, they would have to come to the same conclusions and not draw opposite conclusions.

The value in information classification is emphasized too much and the importance for the organization is neglected (Shi et al., 2007), which means that the focus is too much on the value consideration, and too less on the organization as a functional unit.

Information classification lacks a unified yet flexible classification framework (Alonge et al., 2020). Unified frameworks and ontologies can help build knowledge, share information, and make knowledge available to the public in a unified way (Vargas & Fenz, 2012). Using the example of *Vargas and Fenz* security ontology, the connections between the objects can be quickly demonstrated. However, their security ontology does not consider the information classification, so that the question remains unanswered as to how the information classification could be embedded in the ontology.

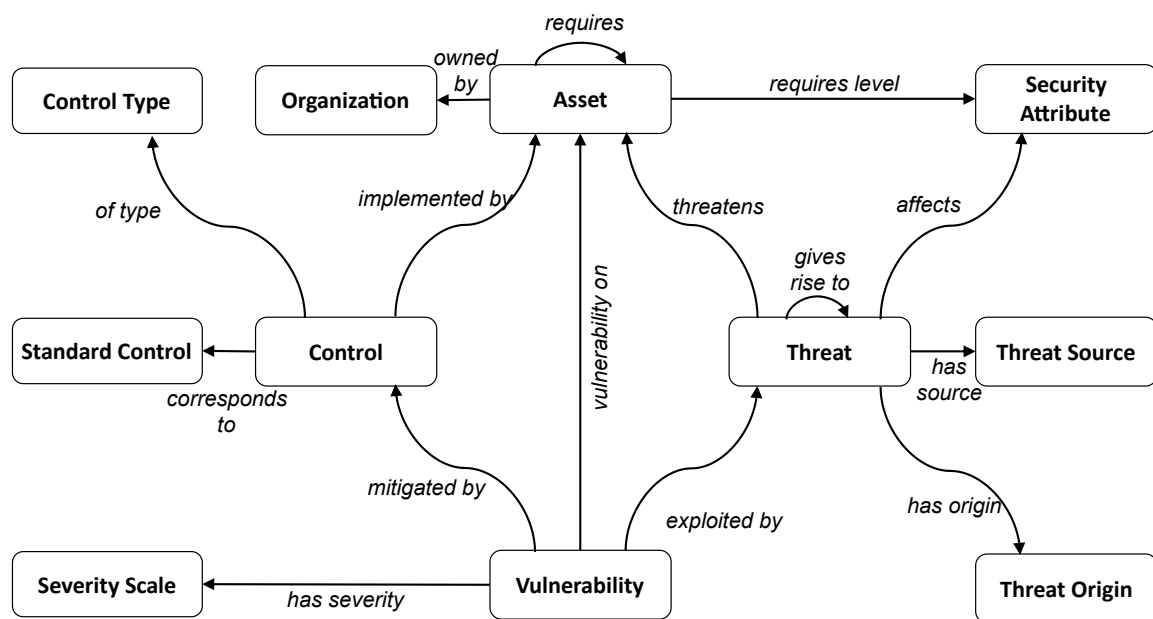


Figure 16: Overview of information security ontology concepts and relationships (Vargas & Fenz, 2012).

Fenz lists information security challenges, but they are also relevant for information classification (Fenz et al. 2014):

- Inventory of assets and security measures: Asset management must be in place, otherwise the risk analyzes are useless. The values of the assets should be recorded as realistically as possible. The asset relations must also be recorded, not just individual assets, because otherwise no realistic depiction is possible. The security measures must be broken down into physical, technical, and organizational aspects. And technical assets offer the highest percentage of automation opportunities.
- Asset Valuation: Determining asset values, whether based on potential losses, monetary or ideal values, criticality, or sensitivity, is difficult and imprecise.
- Inaccurate risk forecasts: A risk forecast is an anticipation of future adverse events. However, forecasts are based on the state of knowledge at the time of the forecast and therefore cannot contain any unknown endogenous and exogenous parameters. In addition, a forecast scatters the further the time horizon is. This means that risk forecasts have an inherent uncertainty and are therefore more of a guideline.

- Overconfidence: People may tend to be overly positive about risk and avoid formal approaches. This in turn has a negative impact on the accuracy of risk, probability, hazard, and impact assessments.
- Knowledge Sharing: The exchange of knowledge from the field of information security could have advantages in professionalizing and reducing the costs of information procurement. However, there are still a few hurdles to overcome.
- Risk compromises because of costs: The cost of security measures must never exceed the cost of the security incident. However, costs are often overlooked as risk management is traditionally a technical issue.

Information classification, while very important and widely used in the context of risk management, does not receive the attention it deserves from the scientific literature. Which is why the practical difficulties in using it in practice are well known, but their causes are largely unknown. The resulting gaps relate, among other things, to concepts, processes, adaptation, frameworks, process roles and information granularity (Bergström, Åhlfeldt, et al., 2020).

### 3.6. Information classification in books

Information must be protected against unauthorized access, modification, disclosure, and destruction. Information classification helps by dividing information into categories to derive appropriate security measures. Classification also helps at allocating and prioritizing the scarce resources to assets with high protection requirements, due to their value, their criticality, legal requirements (Peltier, 2001; Williams, 2013), the organization's willingness to take risks, and the business impact level (Campbell, 2016). „Information classification is a process whereby different sets and collections of data in an organization are analyzed for various types of value, criticality, integrity, and sensitivity“ (Gregory, 2018, p. 219).

Information classification must be done periodically due to mission change, new systems, new networks, new partners, new threats, attackers skills change. If not, the current classification would be static and its effectivity would steadily decline (Landoll, 2021). „Information protection requires a comprehensive and integrated approach“ (Peltier, 2001, p. 2).

Information classification is performed to establish security controls and policies to adequately protect information based on the classification made. The classification is based on the importance of the information and the impact of a security incident. Information classification should always be accompanied by clear instructions on how to handle information of certain classes, and how to handle information of certain classes in its life cycle (Stallings, 2018).

Information classification objectives are (Kim & Solomon, 2018):

- Determine business risks based on the loss of confidentiality, integrity, and availability of business-critical data.
- Derive security requirements for business-critical data.
- Determine data value.
- Ensure that dedicated security controls are in place.
- Standardization: Implement uniform classification in the organization
- Implement awareness and training measures for employees.
- Ensure compliance with relevant laws and guidelines.

Information classification is primarily about protecting important data. Therefore, the data management perspective must also be included as a specialist discipline. Data management is complex because data moves and can distinguish between the following properties and can thus cause different risks (DAMA International, 2017):

- Data types
- Contents
- Data life cycles
- Format
- Protection requirement
- Storage type
- Access needs

High-quality data is important to be able to make effective decisions. Quality data is data that is available, complete, accurate, consistent, timely, usable, meaningful, and understood. The higher the quality of the data, the smaller the information gap between the current level of knowledge and the level of knowledge required to be able to make a well-founded decision (DAMA International, 2017). But there are information gathering limitations, as organizations must comply with certain rules, such as regulations on personal data, intellectual property, or copyright (Humphreys, 2016).

As the only author identified in the information classification context so far, *Campbell* has presented the difference and the connection between data and information; Individual data, data sets, data objects or documents can be given a specific classification. However, several such data objects as a data group can enable further interpretation and enable connections or conclusions. That means in the information-theoretical sense, data become information. As a result, the classification of data groups can differ from individual data. If groups of data, such as databases, have a higher classification than the individual data, then the higher classification also applies to the individual data (Campbell, 2016). However, in addition to the data grouping, the time factor also plays a role, since information can be given a different classification for a limited time, for example when it is being processed and it is critical to the processing process (Campbell, 2016). And information can inherit a higher classification by the system that stores or process it (Campbell, 2016).

Since all data is required to be classified to assign it to a suitable security concept, and since there can be a great deal of data in organizations, there can be various reasons for system-wide classification. One reason for system-wide classification may be lack of time to classify all data. This means that the individual information is no longer relevant for the classification, but the highest information classification in a system, which can lead to many overclassifications (Campbell, 2016). Another reason might be similarity; If a network or even a facility has similar security classifications for most data or systems, the network or facility may be classified as a whole rather than the individual pieces of information or systems (Gregory, 2018). The time required for information classification is enormous and there seems to be no improvement because the information classification process is still a manual activity. On the contrary, the requirements from the cloud area are added to this. All types of service providers that operate or maintain sensitive data or systems for the customer must be evaluated to receive an appropriate classification. The service providers also include operators of platforms, services and software (Gregory, 2018).

The classification of the data is the responsibility of the data owner and is based on the application of regulations for the treatment of certain types of data, such as personal data, financial data, or health data. *Campbell* has shown in *Figure 17* a possible information classification hierarchy, with increasing criticality towards the top.

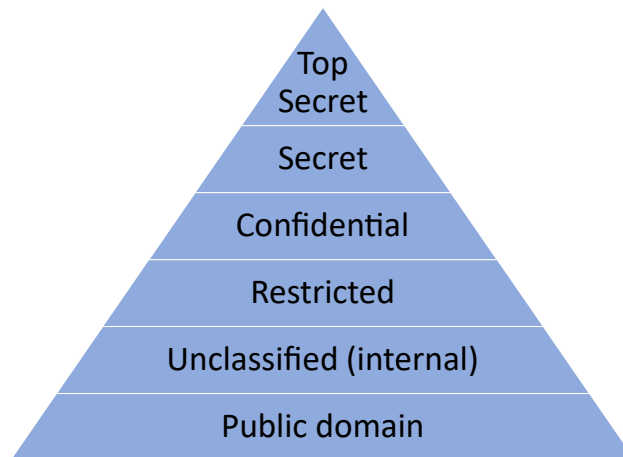


Figure 17: The information classification pyramid (Campbell, 2016).

Kim and Solomon claim that these classification categories are commonly used in business (Kim & Solomon, 2018):

- private data
- confidential
- internal use only
- public domain data

The US federal administration uses the following classification and does not classify public data/information (Kim & Solomon, 2018):

- top secret
- secret
- confidential

DAMA, in turn, has its own classification categories (DAMA International, 2017):

- Critical Risk Data (CRD): Personal information. Risk for individual and financial harm.
- High Risk Data (HRD): Data with potential direct financial value. Risk for financial harm.
- Moderate Risk Data (MRD): Data with little tangible value and unspecific possible harm.

The number of classification levels is not relevant, only whether the users can handle it without support from the security team. However, one should create as few classification levels as necessary to represent the necessary security levels (Campbell, 2016). A data classification scheme is “a formal access control methodology used to assign a level of confidentiality” (Whitman & Mattord, 2018, p. 266). „A classification scheme may be used to identify risk and criticality levels” (Gregory, 2018, p. 199). The data protection strategy includes a data classification scheme, and for each data classification category, data handling policies and procedures. The implementation of the strategy requires that all employees involved are trained in the classification schemes and the procedures (Gregory, 2018). The classification scheme appears to be a key tool for information security because it is not only used for classification, but also for implementing security measures and communicating with data stakeholders. But „enacting a data classification scheme is easy, but implementing the program is hard” (Gregory, 2018, p. 309). Asset classification is an essential task but also very laborious (Susanto & Almunawar, 2018).

The asset owner should be responsible for attribute determination, information management, asset usage (Volchikov, 2018), selection and implementation of the appropriate security measures, adequate information inventory and control mechanisms (Humphreys, 2016).

The overarching goal is to achieve a uniform classification of all documents within the organization. This can be done through guidelines training. Compliance with the information classification guidelines should, on the one hand, be carried out by the information owners themselves, but should also be checked by internal and external auditors (Kim & Solomon, 2018). The Mandatory Access Control (MAC) protects the information from unauthorized access (confidentiality, integrity) and requires that each piece of information has been classified (Kim & Solomon). User access to information must be regulated and monitored throughout the entire user life cycle in the organization through organizational and technical measures. This is the only way to implement access authorization that is appropriate to the information classification (Humphreys, 2016).

The asset owner is responsible for asset security. After asset classification, the asset owner delegates the risk analysis of the assets to the Chief Security Officer (CSO). The CSO checks that the classifications do not have any noticeable deviations. Those who have overall responsibility for the asset are the executives at the management level. The data custodians ensure that the data is used according to the classification rules (Pompon, 2016).

Peltier summarizes information owner responsibilities as follows (Peltier, 2001):

- Valuation of the information.
- Regular reassessment of the classification.
- Determine adequate security measures.
- Communicating access options and security requirements to information managers and information users.
- Access release to authorized users.
- Risk assessment and verification of security measures.
- Supervision of security measures.
- Ensuring a business continuity plan.

*Peltier* divides responsibility for the data between different roles. In addition to the information owner, he also identifies the information custodian and the information user. *Peltier* and *Whitman and Mattord* have definitions for the information roles:

„Owner: Company management of the organizational unit where the information is created, or management of the organizational unit that is the primary user of the information“ (Peltier, 2001, p. 178). „Data owners: Individuals who control, and are therefore responsible for, the security and use of a particular set of information; data owners may rely on custodians for the practical aspects of protecting their information, specifying which users are authorized to access it, but they are ultimately responsible for it“ (Whitman & Mattord, 2018, p. 40). „Data custodians: Individuals who work directly with data owners and are responsible for storage, maintenance, and protection of information“ (Whitman & Mattord, 2018, p. 40). „Custodian: Employees designated by the owner to be responsible for maintaining the safeguards established by the owner“ (Peltier, 2001, p. 179). „Data users are internal and external stakeholders (customers, suppliers, and employees) who interact with information in support of their organization’s planning and operations“ (Whitman & Mattord, 2018, p. 40).

The job description of the custodian could fall within the area of responsibility of the information security.

The information custodians’ responsibilities are (Peltier, 2001):

- Implementation of adequate security measures.
- Provision of a secure IT infrastructure.

- Administration of authorized access to the data

The information users' responsibilities are (Peltier, 2001):

- Use information in accordance with the terms of use.
- Maintaining the confidentiality, integrity, and availability of the available information.

Calder and Watkins claim that information owners should only be defined for information worthy of protection, therefore the information classification should be carried out first and the information owner should be named from a previously defined classification category (Calder & Watkins, 2015). This requirement would contradict the previous statements to information owner. In addition, the question arises who should find the information worthy of protection if not the information owner?

In order for the information owner, custodian and users to be able to fulfil their tasks (acquire, store, transmit, delete data) effectively, clear processes must be defined for all his tasks and in particular for information classification (Gregory, 2018; Kim & Solomon, 2018).

Measure criteria for information classification could be (Kim & Solomon, 2018):

- Value:
  - Value to the organization
  - Value to competitors
  - Cost of replacement or loss
  - Value to the organization's reputation.
- Sensitivity (loss of integrity or confidentiality):
  - Liability or fines
  - Reputation
  - Credibility
  - Loss of market share
- Criticality (loss of information):
  - Importance of the information to the organization
  - Scenarios with negative impact on mission achievement in case of information loss

The information to be classified should be determined based on its value to the organization, which can be done through a business impact analysis. Criteria for determining the value can be the following (Kim & Solomon, 2018):

- Exclusive possession (trade secrets)
- Utility (usefulness)
- Cost to create or recreate the data
- Liability (protection regulations)
- Convertibility/negotiability (financial information)
- Operational impact (if data are unavailable)
- Threats to the information
- Risks

*Pompon* has listed examples of valuable data:

- „Intellectual property
  - Source code
  - Product designs
  - Copyrighted material
- Personal information
  - Customer credit/debit cards

- Social Security numbers
- Customer names
- Driver's license numbers
- Account numbers
- Passwords
- Medical information
- Health insurance information
- Vehicle registration information
- Usernames, e-mail addresses, passwords
- Message repositories
  - E-mail
  - Chat logs
- Customer-facing IT services that are revenue generating
  - E-commerce sites
  - Streaming media
  - Product catalogues
- Customer-facing IT services that are support/non-revenue
  - Web support forums
  - Documentation
  - Demo sites
- Critical internal IT services
  - Chats
  - Help desk
  - Accounting
  - Payroll
- Semi-critical internal IT services
  - Web sites
  - E-mail
  - SharePoint" (Pompon, 2016, p. 28).

Determining the information value can help in deciding how much the security measures can cost in order not to create a mismatch between the information value for the organization and security costs (Campbell, 2016).

*Landoll* has made examples of the classification of assets. *Table 6* is an excerpt that only shows the information assets.



Table 6: General Asset List (excerpt) (Landoll, 2021).

Asset category	Sub-category	Examples
Information	Sensitive	<ul style="list-style-type: none"> <li>• Employee applications</li> <li>• Employee records</li> <li>• Facility plans</li> <li>• Intellectual property</li> <li>• Account password</li> <li>• Pricing information</li> <li>• System vulnerabilities</li> <li>• Financial data</li> <li>• Contingency procedures</li> </ul>
	Protected	<ul style="list-style-type: none"> <li>• Medical records</li> <li>• Financial inquiries</li> <li>• Health insurance applications</li> <li>• Bank statements</li> <li>• Credit reports</li> <li>• Prescriptions</li> </ul>
	Public	<ul style="list-style-type: none"> <li>• Web site</li> <li>• Marketing materials</li> <li>• SEC filings</li> </ul>

„Qualitative asset valuation is sufficient for qualitative risk assessments. But when it’s necessary to calculate figures such as exposure factor or annual loss expectancy, it will be necessary to obtain quantitative valuation for relevant assets. However, the need to be precise is low since it is difficult to know the probability of threat events“ (Gregory, 2018). This means that it must be clear beforehand which additional goal the risk analysis should have to adapt the information classification and the valuation.

From the data management perspective an organization-wide data model of all information is essential for effective information classification (DAMA International, 2017). In order to identify data and create the data model besides the cybercrime professionals, data managers and information technology developers are seen as having a relevant role in the information identification and classification process when it comes to identifying the relevant data and developing security concepts and access models (DAMA International, 2017).

The information security governance is effective if:

- “All information in use within the organization is identified,
- Information is classified according to criticality Information is classified according to sensitivity,
- Information classifications are enforced,
- Information classifications are applied to information received from outside entities,
- Information classifications are applied to information provided to an outside entity,
- Ownership responsibilities for all information are assigned,
- Applications that process sensitive information are identified,
- Applications that support critical business processes are identified,
- Data retention standards are defined and enforced” (Stallings 2018).

Digital and physical data have a life cycle not much different from other information lifecycles, but *Figure 18* depicts the management perspective of the data in focus:

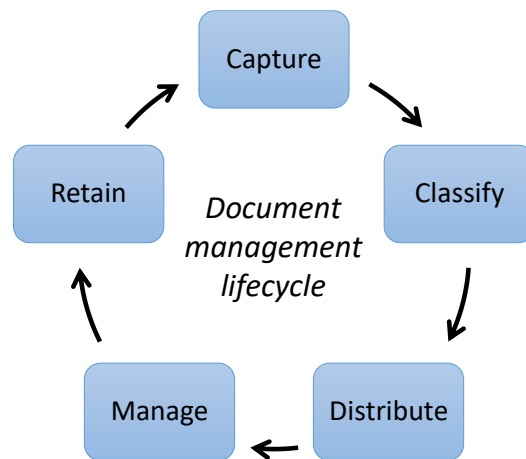


Figure 18: Document Management Life Cycle (Stallings, 2018).

The Figure 19 is relevant because it puts the context of risk management into the information security landscape. The information classification is not explicitly listed but must be imagined as the preparatory work relevant to the risk analysis in the "Core Elements". The illustration makes it immediately clear which relationships exist between the various security topics and which complex dependencies to information classification may exist.

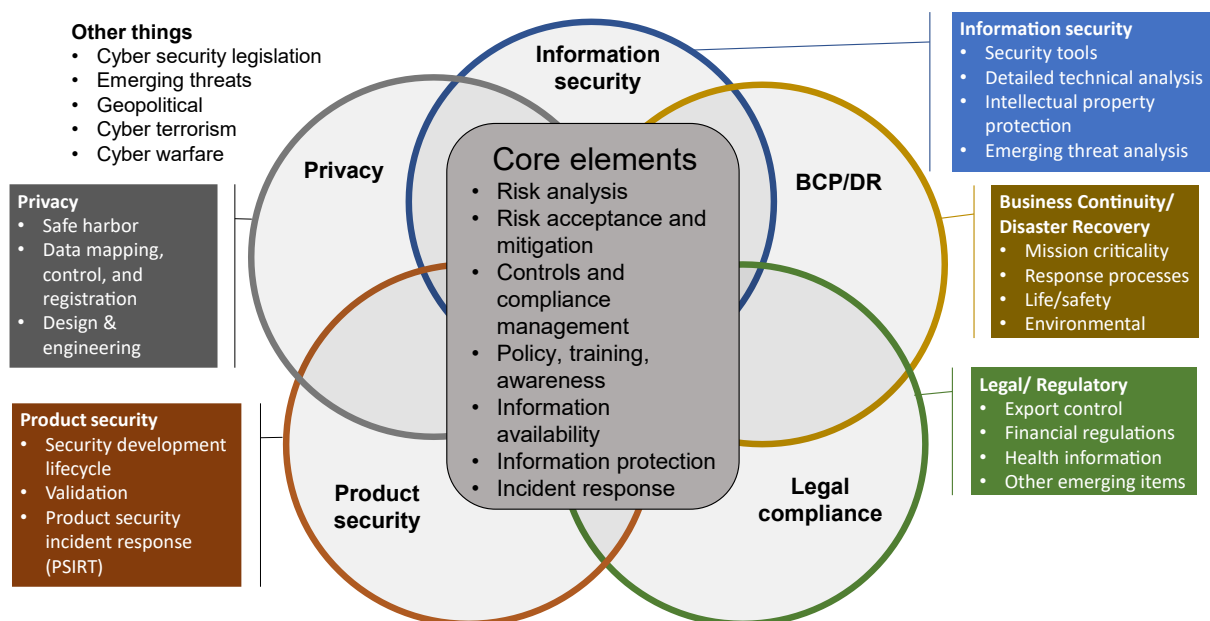


Figure 19: Security and privacy : the primary areas of information risk, and the core elements of information risk management that apply to each area (Harkins, 2013).

As just stated, a complex relationship between different elements (asset, organization, people, stakeholders, technology) can be assumed, in which the information classification is located and thus the term "system" can be used to describe it. The challenge of system security engineering (SSE) is to protect the right items in the right way (SEBoK Editorial Board, 2022). The analogies to information security can be seen here.

Automation plays an important, in the sense that there should be tools that automatically recognize sensitive information (Stallings, 2018). But also tools to declassify data automatically after a time period, as data changes over time in terms of its use, importance, value and sensitivity (Peltier, 2001). An implementation of the information classification is achieved through a combination of technology, processes, and employee training. The technology assumes the enforcement and control of the

classification and access rules (Campbell, 2016). But the core of information classification are its classification categories and the people conducting the classification. As classification is a human dependent manual task the classification categories numbers must be kept low, to avoid misunderstandings and misclassifications, although more classification categories could be useful. Instead the few information classification categories have to be explained and underpinned with examples (Gregory, 2018).

Information should be provided with attributes during classification. Benefits of having standardized attributes for metadata are (Stallings 2018):

- Common and better understanding of the attributes.
- Better decision making in context of authorization.
- Higher granularity in access control possible.
- Attribute standardization across organizational borders.

*Volchkov* recommends the following essential attributes:

- “Unique identifier of the asset
- Name of the asset
- Description
- Category or type - to be defined internally
- Confidentiality class
- Criticality class
- Owner or accountable (person or department)
- Personal data (Y/N)
- Other sensitive data (Y/N) - to be defined
- Location
- Used by third party (Y/N)
- References to other related assets - dependency links (e.g., application linked to a server or database)
- Referencing business units and processes that use the asset” (Volchkov, 2018, p. 225).

Data consists not only of currently used data or data that will be destroyed, but also of data that should be archived for a certain period due to legal regulations, or forever due to historical requirements. Records management archives documents because of legal requirements or because of the importance for the organization. The security requirements are generally higher in records management than in non-archived documents (Stallings, 2018). Information classification also plays an important role in records management, as it determines how data is stored, how it is backed up, who can access it, and how it is destroyed. And attributes also play an important role in records management, among other things, to clearly assign the classification to the information.

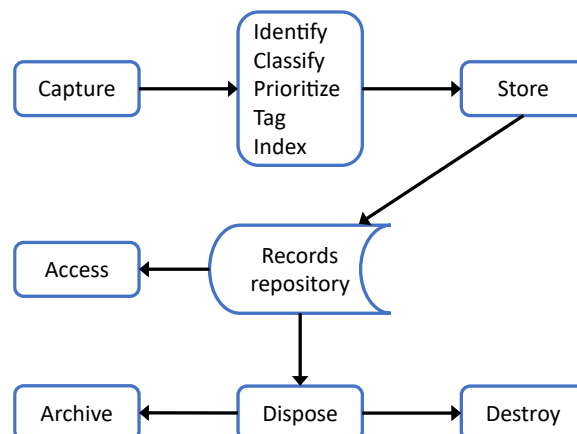


Figure 20: Records Management Functions (Stallings, 2018).


„Important guidelines for securing records include:

- Develop clear and detailed policies that define what documents should be classified as records.
- Store only a single instance of the document, whether it is physical or electronic.
- Define access restrictions as tightly as is reasonable.
- Monitor each record to ensure compliance with the organization’s records retention policy.
- Ensure secure destruction of a record or secure archival of the record when the retention period has expired” (Stallings, 2018).

Organizations should have policies and procedures describing how asset classification and its control should be conducted (Kim & Solomon, 2018). Employees should be trained in the proper handling of unclassified and classified data (Kim & Solomon, 2018), as soon as they join the organization (Calder & Watkins, 2015).

When implementing the security measures, data classification must be implemented in all 7 layers of the OSI reference model and compliance must be monitored (Kim & Solomon, 2018). And „... every type of classification corresponds to a certain security infrastructure (building, organizational structures, and personnel)” (Brunner & Suter, 2008, p. 77). As has been stated several times, information security consists not only of the classified information, but also of the assets that store, process, or transfer the information. After information has been identified and classified, the system assets saving, processing, and transferring critical information must be identified. This step is often overlooked or inadequately conducted (Gregory, 2018). That's why *Kim and Solomon* pointed out that the entire OSI stack should be considered. Thus the relations between assets (data, database, server, network...) should be visible to enable risk analysis on asset groups (Volchkov, 2018). Asset classification aims to assign assets into usage or risk categories, to identify its criticality which is related to information sensitivity. Thus, assets derive their classification from information they handle, or by their inherent criticality in business operations (Gregory, 2018). One possibility for this connection between classified information and assets (software, hardware, network) is an asset inventory list with reference links, as shown in *Table 7* as an example.

*Table 7: Reference links between assets (Volchkov, 2018).*

Asset inventory				
Asset	...	Class	...	Link reference
Application A		Confidential		
...				
Server S		Confidential		
...				
Database DB		Confidential		

An information asset could have come first to complete the picture. Despite this, there are clear connections between the assets and that the classification is inherited from asset to asset. However, a clear differentiation between data, information and assets is not always made. For example, *Whitman and Mattford* define information asset as follows: „Within the context of risk management, any collection, set, or database of information or any asset that collects, stores, processes, or transmits information of value to the organization. Here the terms data and information are interchangeable” (Whitman & Mattford, 2019). The two authors mix up the terms data, information, asset, the information object, and the information processing. This confusion does not contribute to the understanding of the already difficult complex of topics.

### 3.7. Information classification in practice

It is interesting that despite the importance of information classification for determining appropriate security measures and the mandatory requirements of laws and standards, only a minority of

organizations in the public sector and less than half of companies define information classification in policies or put into practice (Bergström et al., 2018). *Kobis* conducted a survey of 117 SME companies about the risk management process in practice. It turned out that the interest in security aspects with reference to personnel, information storage locations, quantitative risk assessment methods, the ISO 27005 information security recommendations and a risk management life cycle is low (Kobis, 2020).

The concept of risk management provides for a dynamic and cyclical risk treatment, which in practical implementation is likely to be rather static due to the real working methods of organizations (Lundgren & Bergström, 2019). This is also reflected in the results of a survey in which only 7% of the SMEs surveyed go through phases of the risk management process annually and 19% each time there is a change in the information system.

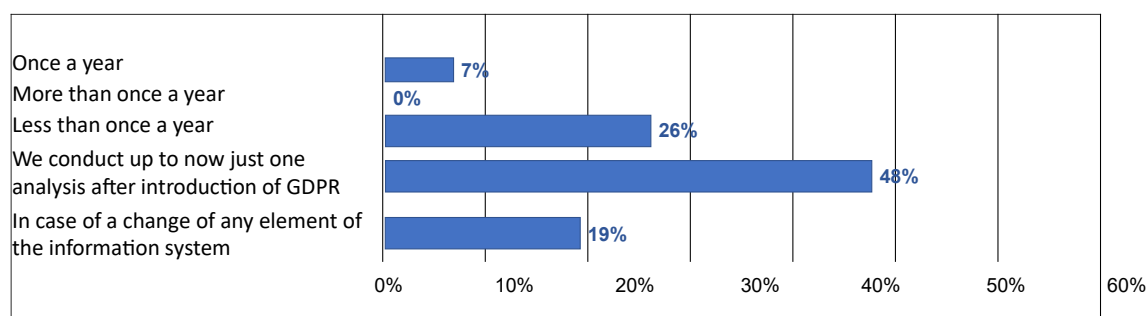


Figure 21: Answers of the respondents to the question: How often do you carry out the various stages of the risk management process (Kobis, 2020).

Sharing classifications of information with other organizations is not possible because the value of information is assessed from the perspective of the organization (or person within the organization) and the value may be different to another organization (Bergström et al., 2018).

The United Kingdom, Australia and New Zealand information classification frameworks have been revised due to the high security costs of over-classification due to an overly confusing and complex classification framework. In addition, the approach to information classification, which was not adapted to digital administration, was criticized as it resulted in inefficient administrative processes. The changed security challenges, caused by new threats and risks, were identified as an external driver (Heide & Villeneuve, 2020). Although the three countries had identified the same causes for the need for reform in the information classification framework, they came to different conclusions in their reformed frameworks in terms of solving the problems. The UK, for example, has greatly simplified the framework by reducing the number of categories from 5 to 3. New Zealand did the exact opposite, increasing the number of categories from 3 to 7. Australia again reduced the number of categories from 7 to 4 (Heide & Villeneuve, 2020), see following table.

Table 8: Structure of classification frameworks in the case countries (Heide & Villeneuve, 2020).

UK (pre-2014)	UK (currently)	NZ (pre-2000)	NZ (currently)	AUS (pre-2011)	AUS (2011 to 18)	AUS (currently)
Top secret	Top secret	Top secret	Top secret	Top secret	Top secret	Top secret
Secret	Secret	Secret	Secret	Secret	Secret	Secret
Confidential		Confidential	Confidential	Confidential	Confidential	
Restricted	(Official-sensitive)		Restricted	Restricted	Protected	Protected
Protected			Sensitive	Highly protected	Sensitive	(Official-sensitive)
			In confidence	Protected	For official use only	Official
				X-in-confidence		
Unclassified			Unclassified	Unclassified	Unclassified	

With the different conclusions for solving the identified problems of the three countries in information classification, at least the question arises as to whether the number of classification categories was the

cause of the implementation problems, or whether changing the classification categories is the simplest measure to show the ability to act.

Information classification can affect the entire IT architecture. For example, a very high classification can mean that the database storing the information must be specially secured, for example by separating the sensitive data from the non-sensitive, implementing encryption and logging. Furthermore, server systems processing sensitive information could also be sealed off in separate, restrictedly accessible and monitored network areas (Gregory, 2018). These security measures, which can be directly traced back to the information classification, can in turn include pre-requisite or subsequent measures. Consequently, such a classification is coupled with a complex technical and organizational effort.

A fundamental problem of information classification is that it does not receive the necessary status in practice (Gheraouti-Helie et al., 2011; Kindervag et al., 2015; Shedden et al., 2016; Shedden et al., 2010), possibly because information classification processes have not been researched enough and therefore no relevant transfer into practice has taken place (Bergström, Karlsson, & Åhlfeldt, 2020). However, this does not only apply to the classification of information, as organizations do not use all steps of the recommended ISO 27005 risk management process in their risk management practice or use them optionally or as a complement. The reasons were, among other things, that it was not possible to follow the dynamic changes in the information and system landscape quickly enough, so that in the end cutbacks were made in the risk analysis process and a static risk analysis prevailed overall (Lundgren, 2020). Nevertheless, companies are trying to catch up with the dynamics of IT and organization in risk management by reclassifying their asset triggered by (Lundgren 2020):

- Time (every 2 to 3 years)
- Incidents where the security measures were not effective
- Implementation of new applications or systems

To measure the degree of implementation of the information classification, ISO27 practitioners propose the following key performance indicators (KPI) (ISO27k implementers, 2007):

- The percentage of classified information assets per classification step. The classification steps are:
  - Identified
  - Inventoried
  - Owner asset nominated
  - Risk assessed
  - Classified
  - Secured
- Percentage of asset per classification category.
- The percentage of relevant information assets with an implemented risk mitigation strategy.

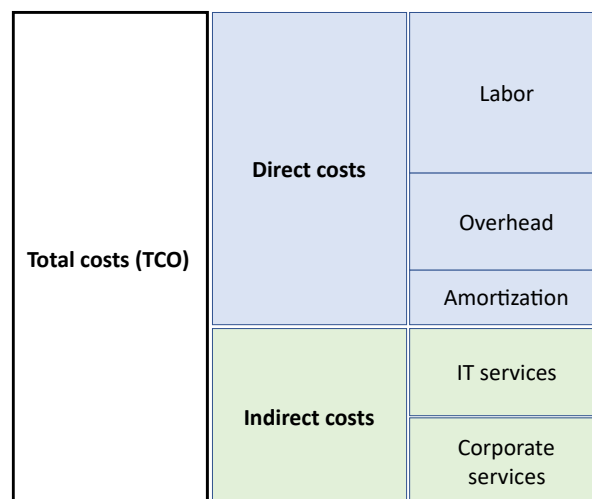
### 3.8. Practice challenges

Sometimes organization's confidentiality rules hinders research about information classification (Chen et al., 2015). Knowledge sharing for best practice and processes in security is not common in public and private sector (Alsmadi, 2019).

The different results of the classification framework reforms of the United Kingdom, Australia and New Zealand governments (Heide & Villeneuve, 2020) raise questions as to whether practically relevant conclusions were actually drawn from the problems of the existing information classification frameworks on administrative practice. Here again the question arises whether the purely qualitative

and manually executed information classification approach should be fundamentally questioned. Another issue is that best practice prescriptions were insufficient for local action (Bergström, Karlsson, & Åhlfeldt, 2020).

Managers should base their decisions on sound data, knowledge, and experience, which they do not have in the field of risk management. However, the problem can be solved by external expertise or by simplifying the decision-making system (Fenz et al., 2014). Instead, risk prediction was used in practice in a simplified manner or not at all, as it is complex and time consuming. Instead fixed lists of risks and countermeasures were used (Erik et al., 2019). Some organizations address the difficulty of valuing assets with group discussions of diverse subject matter experts (Erik et al., 2019). Information classification is a very time-consuming and largely manual activity, especially when, as required by information security standards, all information is to be classified, both initially and with each change. Furthermore, the information should be enriched with metadata to be able to recognize their classification status at any time. The classification must also be implemented in the role and access model and the monitoring of the operational IT. That means an enormous amount of work. Based on an exemplary cost structure in *Figure 22*, this cost structure would have to be analyzed.



*Figure 22: Example of breakdown of security costs (Volchikov, 2018).*

We are not aware of a cost calculation or an analysis of the additional costs, either from scientific research or from practical reports. However, there is a need because security is ultimately also a cost-benefit calculation in practice, and a well-founded analysis is not possible without calculable costs.

The formal risk analysis process of ISO 27005 does not run smoothly in practice, but causes socio-organizational friction losses, so that organizations make more or less extensive adjustments to the recommended process in order to remain able to act (Lundgren, 2020).

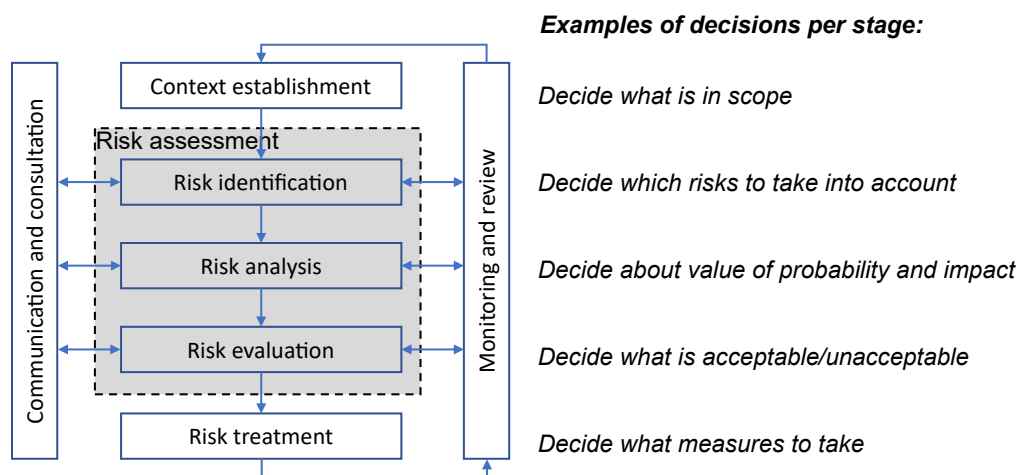
### 3.9. Context

Here is a brief description of the context in which information classification is found. This context is only a selection and is the perspective that has been gained based on the previous literature review. It is considered relevant because information classification can only be understood if it becomes clear what requirements are placed on information classification and what restrictions it is subject to.

### 3.9.1. Bias in information classification

For successful information security, the attitude must be right to want to implement information security. The knowledge about information security is important to be able to implement information security. And behaviour is important so that information security is lived (Gundu & Flowerday, 2012). This security statement can also apply to information classification. Both the knowledge about the information, the classification method, and the behaviour of the people who classify, i.e., the information owner, must be appropriate for the information classification to work effectively. According to scientific literature, standards and book literature, information classification and risk assessment are almost entirely manual human activities. To achieve an unspecified qualitative level in the classification, organizational and personnel measures are carried out, such as guidelines, implementation instructions, look-up lists, training courses and awareness-raising. Person-centricity has some disadvantages, which will briefly be outlined here.

The risk management process is shown in *Figure 23*. On the right side of the figure, the authors have added a list of decisions for each risk management phase.



*Figure 23: Risk management process according to ISO 31000 with examples of decisions per stage (Wit et al., 2021).*

As has already been stated in this work, the information classification is missing as an independent phase in the illustration. Therefore, potential information classification decisions are added here. The information classification also has the relevant decisions, namely:

- Decide which classification the information will receive.
- Decide what information is relevant for risk assessment.
- Decide what assets are in relation to the relevant information.

A major weakness of decisions made by people is bias, which is fed by various motivators, as heuristics and judgmental biases that are economic shortcuts for information processing (rules of thumb), as information gathering is costly (Peón et al., 2017). The availability heuristics exists when decisions were made by a subset of information that is easy available (Peón et al., 2017). Information processing can also rely on stereotypes, particularly in probability estimation (Peón et al., 2017). Since not all biases can be described in this work, but scientific research from this area has to be considered as an additional perspective in the system of information classification in order to identify new rational arguments, here is a taxonomy of behavioural biases without further descriptions (Peón et al., 2017):

- Attention bias
- Hindsight bias
- Law of small numbers
- Base rate neglect



- Illusion of validity
- Causality and attribution
- Conjunction & disjunction fallacies
- Risk-as-feelings
- Aversion to ambiguity
- (Excessive) optimism
- Overconfidence
- Self-attribution bias
- Confirmation bias illusion of control
- Global culture
- Social contagion
- Status, social comparison
- Status, social comparison
- Fairness and justice
- Greed and fear
- Informational cascades

One argument could be that expert knowledge could help reduce bias to arrive at sound information classifications. However, this potential argument is refuted, as according to a study by *Wit et al.*, the focus on security experts in the risk assessment does not have a positive effect on the quality of the decision. It turned out that the vast majority of security experts have the same weakness for biases as laypeople and that expertise had no positive effect on the quality of decisions (Wit et al., 2021).

### 3.9.2. Information gathering

Risk identification, and thus also information classification, can only be carried out professionally with the preparatory procurement of supporting information, as system information (National Institute of Standards and Technology, 2002). The following list show system-related information that should be gathered in advance of identifying IT system risks (National Institute of Standards and Technology, 2002):

- Systems
  - Hardware
  - Software
  - System interfaces
    - Internal
    - External
- System mission
- Data and information
- System and data criticality
  - Value
  - Importance
- System and data sensitivity
  - Confidentiality
  - Integrity
  - Availability
- Persons
  - Supporting IT systems
  - Using IT systems
- Additional information
  - The functional IT system requirements

- System security policies governing the IT system
  - Organizational policies
  - Federal requirements
  - Laws
  - Industry practices
- System security architecture
- Current network topology
  - Network diagram
- Information storage protection that safeguards system
- Flow of information pertaining to the IT system
  - System input and output flowchart
  - System interfaces
- Technical controls used for the IT system
  - Identification and authentication
  - Discretionary or mandatory access control
  - Audit
  - Residual information protection
  - Encryption methods
- Management controls used for the IT system
  - Rules of behavior
  - Security planning
- Operational controls used for the IT system
  - Personnel security
  - Backup
  - Contingency
  - Resumption and recovery operations
  - System maintenance
  - Off-site storage
  - User account establishment and deletion procedures
  - Controls for segregation of user functions
    - Privileged user access
    - Standard user access
  - Physical security environment of the IT system
    - Facility security
    - Data center policies
  - Environmental security controls implemented for the IT system
    - Humidity
    - Water
    - Power
    - Pollution
    - Temperature
    - Chemicals

We are not aware of any detailed scientific analysis of the a priori collection of information. Questions would have to be asked here as to which information, in which form, with which attributes and how it would have to be processed so that the information classification would have the greatest benefit, i.e., be able to make the most targeted and well-founded decisions possible.

### 3.9.3. Information handling

All information in an organization should be identified and classified. To indicate that information has been classified, the information requires a label. This can be implemented with metadata. The *NISTIR 8112* schema provides a metadata matrix, with attributes and values. One metadata attribute is classification with the values "Unclassified, Controlled Unclassified, Confidential, Secret, Top Secret, Company Confidential" (National Institute of Standards and Technology, 2018). But no security standard prescribes which attributes the assets should have (Volchkov, 2018). In addition, care must be taken to ensure that the information is treated in accordance with the rules and in accordance with the classification made.

*ISO 27002* has the following recommended actions for information:

- Appropriate access restrictions for each classification category.
- Current documentation of permissions.
- Copies must be clearly identified and subject to the same protection categories as the originals
- Observe the manufacturer's guidelines when storing/storing information.
- Automatically monitor the proper handling of information, including tagging information and attaching classification as metadata. The use of document management systems or records management systems could also be helpful.
- Other security techniques should be used, such as logging and encryption (Stallings, 2018).

### 3.9.4. Assets

There are two asset types, the one to be protected and the one that helps to protect, as firewalls, Intrusion Detection Systems, Antivirus, etc. (Fenz et al., 2014). Anything that is of value to the organization is an asset worthy of protection, such as hardware, software, information or other business assets (Stallings, 2018).

Hardware assets are (Stallings 2018):

- Servers
- Workstations
- Laptops
- Mobile devices
- Removable media
- Networking equipment
- Telecommunications equipment
- Peripheral equipment

Software assets are (Stallings 2018):

- Applications
- Operating systems
- System software
- Virtual machine
- Container virtualization software
- Software for software-defined networking (SDN)
- Software for network function virtualization (NFV)
- Database management systems
- File systems
- Client software

- Server software

Information assets in a telecommunications or network environment are (Stallings 2018):

- Communication data
- Routing information
- Subscriber information
- Blacklist information
- Registered service information Operational information
- Trouble information
- Configuration information Customer information
- Billing information
- Customer calling patterns
- Customer geographic locations Traffic statistical information Contracts and agreements
- System documentation
- Research information
- User manuals
- Training materials
- Operational or support procedures Business continuity plans
- Emergency plan fallback arrangements Audit trails and archived information

The list of assets does not claim to be complete, but rather a listing of potentially valuable assets in the organization that are therefore worthy of protection. The context of the information classification is primarily the information or the data. Secondly, the other assets are relevant for information classification because of the connections between information and hardware, software, and network/communications.

Information has also a lifecycle as any other asset, and it consists of the following steps (Bernard, 2007):

- Creation and Receipt:  
Information can arise within the organization or be imported into the organization from external sources. Information can take different forms, such as:
  - Written
  - Printed
  - Electronically
  - Orally
  - Correspondence
  - Contracts
  - Applications
  - Reports
  - Drawings
  - Production
  - Transaction records
- Storage:  
Storage refers to all storage locations, such as:
  - Network drive
  - Random access memory
  - Hard disk
  - File
  - Network device

- Servers
- Clouds
- Smart phone
- Print texts
- Disk
- Human brain
- Distribution and Transmittal:  
Refers to obtaining information to be able to process and use it. The information can be distributed manually or automatically and can be defined by rules or triggered by demand.
- Access and use:  
Access and use of information may result in the information.
  - Be further distributed.
  - Be transformed into another form.
  - Be changed, supplemented, copied or deleted.
- Maintenance:  
It is about information management and deals with
  - Information filling
  - Archiving
  - Information gathering
  - Information transfer
  - Information classification change
- Disposition and Destruction:  
Deals with the identification of unused and no longer required information and its secure destruction.

Information can have different values for your organization. A simple division into values would be the value pyramid of information. It illustrates that information collected is of the least value as the information can usually also be available to third parties. The processed information is information that has been refined with the help of knowledge and thus increased in value. The most valuable information is the created information because it represents the knowledge of the organization.

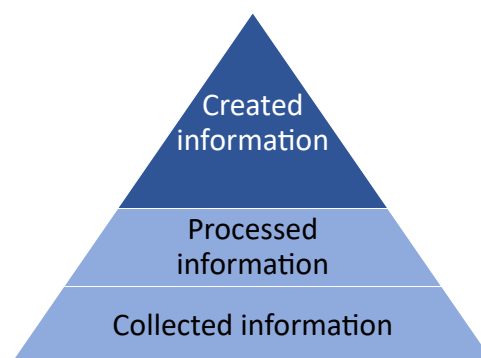


Figure 24: Value pyramid of information (Al-Fedaghi, 2008).

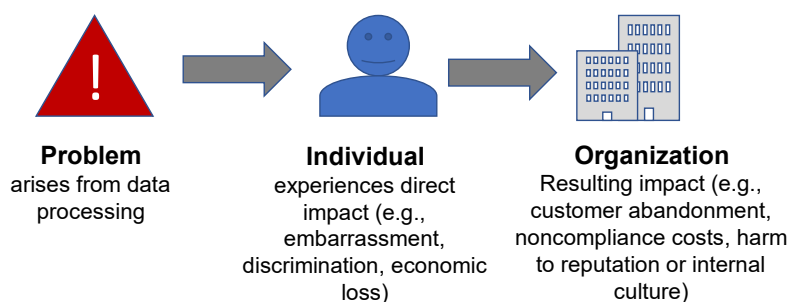
However, the valuation of assets is not as easy as the pyramid suggests, because asset valuation needs to consider the impact of threats to confidentiality, privacy, integrity, and authenticity. As an example of questions involved in information asset evaluation, *NISTIR 7621 Small Business Information Security* (Stallings 2018):

- What would happen to my business if this information were made public?
- What would happen to my business if this information were incorrect?
- What would happen to my business if my customers or I couldn't access this information?

### 3.9.5. Risk

„Risk is the net negative impact of the exercise of a vulnerability, considering both the probability and the impact of occurrence“ (National Institute of Standards and Technology, 2002, p. 7). The protection of information is usually described with the CIA triangle and refers to the protection of confidentiality, integrity, and availability by avoiding, reducing, shifting, or accepting the risk of endangering it. „System and data confidentiality refers to the protection of information from unauthorized disclosure... Unauthorized, unanticipated, or unintentional disclosure could result in loss of public confidence, embarrassment, or legal action against the organization“ (National Institute of Standards and Technology, 2002, p. 28). „System and data integrity refers to the requirement that information be protected from improper modification... loss of integrity reduces the assurance of an IT system“ (National Institute of Standards and Technology, 2002, p. 28). „If a mission-critical IT system is unavailable to its end users, the organization’s mission may be affected“ (National Institute of Standards and Technology, 2002, p. 28). Even if the focus is on the hardware and software assets when defining the risks, the primary risk consideration is the information. The hardware and the software are considered secondarily in the risk assessment because they store, process, and transfer the information that needs to be protected. This change of perspective is relevant because it enables a targeted risk assessment to draw the right focus, the relevant relationships, and the right conclusions.

A distinction must be made between compliance risks and individual risks, since 100% compliance with applicable laws and guidelines minimizes compliance risks but does not necessarily minimize the privacy risks of individuals. Different legal systems in different countries can also lead to conflicting compliance risks. Different industries can also lead to different requirements regarding privacy and cyber risks. And where the legal framework does not provide an adequate answer or where a voluntary degree of risk assessment takes place, the cultural background of the decision-maker plays an important role in the risk assessment (National Institute of Standards and Technology, 2020). *Figure 25* shows the relation between the privacy risks and organisational risks.



*Figure 25: Relationship Between Privacy Risk and Organizational Risk (Stallings, 2018).*

There are a wide variety of risks from a wide variety of perspectives, some of which intensify and thus increase the overall risk. This increases the complexity of the risk landscape, and an individual assessment of the risks is no longer appropriate because it ignores important perspectives or dependencies. The increase in legal requirements and the complexity of the IT landscape and its potential attack possibilities have caused the risks to be considered to explode in their quantity and complexity. Taking all risks into account and anticipating unknown risks is not realistic. It can therefore be assumed that some of the risks will always remain known and untreated or even unknown, see *Figure 26*. This knowledge is worth a separate risk assessment and should find its consequences in risk management by considering the risk of an undesirable future outcome due to unrecognized risks and taking appropriate measures. This is relevant to information classification because it affects the actions taken to implement the classification.

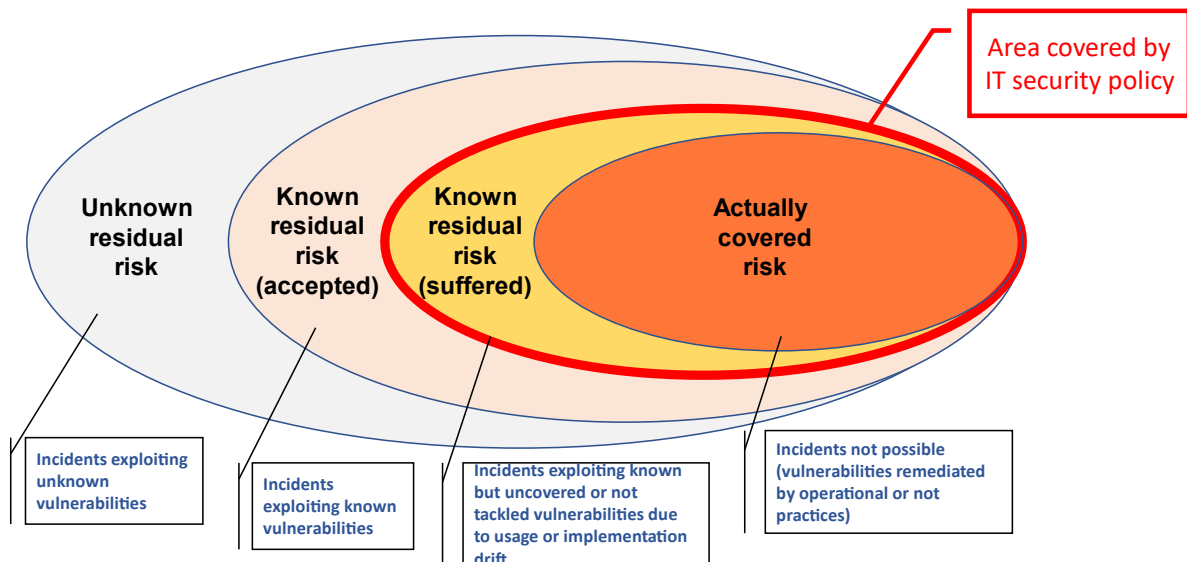


Figure 26: The 3 kinds of residual risks (ETSI, 2023).

### 3.9.6. Risk assessment

The top priority of risk assessment is to achieve optimal information security within a given budget (Stallings, 2018). Risk assessment is a complex subject that is more art than science and calls for considerable management judgment (Stallings, 2018). And it is precisely this statement, that makes any empirical and logically justified approach more difficult. The cybersecurity community needs to be clear about whether they align themselves more closely with engineering and science, or with the humanities. Consequently, whether pure, hard facts, and applied knowledge or soft knowledge make a greater contribution to gaining knowledge and deliver the more effective and efficient approach in practice, and where the limits of interdisciplinary approaches are to be drawn (National Institute of Standards and Technology, 2020).

The risk management process according to *ISO 27005* describes all necessary steps to subject risk management in an organization to a controlled control loop. Risk management includes risk assessment, which in turn consists of the tasks of risk identification, risk analysis and risk evaluation. Risk identification represents the first step in the risk assessment process and is defined as follows: “Involves the identification of risk sources, events, their causes, and their potential consequences. It involves historical data, theoretical analysis, informed and expert opinions, and stakeholders’ needs” (Stallings, 2018).

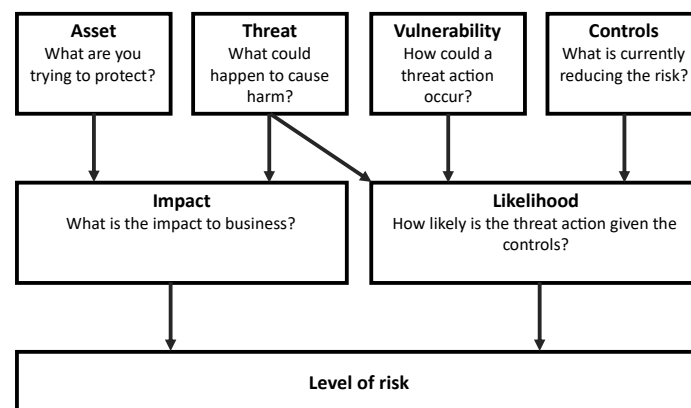


Figure 27: Determining Information Security Risk (Stallings, 2018).

Obviously, risk identification already requires the identified and relevant information (assets) as input. Already at this point, it should be clear about what data is needed, what relationships to other data, hardware, software, and what attributes this data and relationships should have, so that a complete and realistic risk identification is possible. Analysis based on expertise requires a well-founded database. The risk assessment is very difficult to carry out due to the great uncertainty regarding future events (dangers), their probability of occurrence and the extent of the damage. In addition, there can be multiple relationships between threats, vulnerabilities and assets, making risk assessment significantly more difficult (Stallings, 2018).

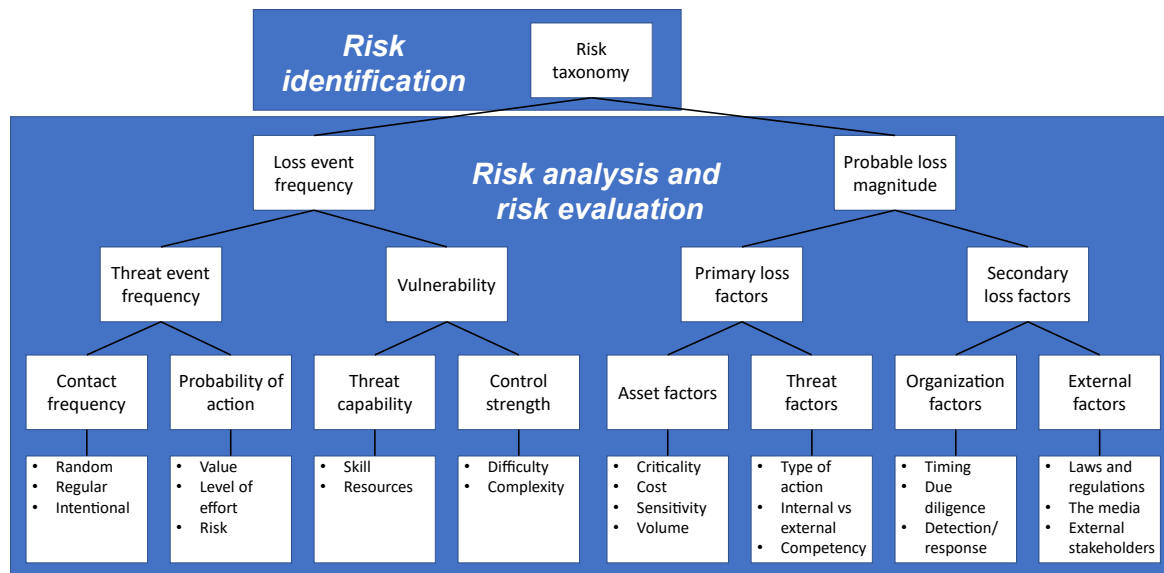


Figure 28: Risk Assessment Using FAIR (Stallings, 2018).

Figure 28 shows the risk assessment according to the FAIR approach using a risk taxonomy. However, there are also requirements here, this time from practice, to make risk assessment methods more pragmatic (Faizi et al., 2021), or in other words simpler.

Risk assessment can be carried out both quantitatively using probabilities or monetary values and qualitatively using categories. In Table 9 an exemplary risk matrix is depicted, that contains both qualitative and quantitative descriptions of the impact.



Table 9: Example of a security risk matrix with different impact scales (Volchkov, 2018).

				Impact				
				Minor non conformity	Remediation deadlines	Repetitive non compliance	Significant fine	Out of business
				Minor injuries	First aid required	Hospitalization	Multiple serious injuries	Death
				< 1% of budget	1% - 3% of budget	3% - 6% of budget	6% - 10% of budget	> 10% of budget
				Customer dissatisfaction	Some customer complaints	Spotlight and local echo	Major complaints and media coverage	Deterioration of the image in the long term
				Insignificant	Minor	Moderate	Major	Extreme
				1	2	3	4	5
Likelihood	Annual or monthly (<year) occurrence	Almost certain	5	Repeatable minor incidents	10	15	20	Announced disaster
	Has occurred in the past in less than a year and will occur	Likely	4	4	8	12	16	20
	Has occurred at least once in our or other organizations	Possible	3	3	6	9	12	15
	Has never occurred in the organization	Unlikely	2	2	4	6	8	10
	Is possible but has not occurred to date	Rare	1	Negligible	2	3	4	Almost impossible disaster
				Risk value	Gravity	Description	Actions	
				15 – 25	Very high risk	Unacceptable risk. Must be followed and reviewed regularly. Reporting at the board level and management	Immediate action needed: transfer, avoid, or reduce. Followed and reported by CISO.	
				8 – 12	High risk	Potentially unacceptable level. Must be constantly monitored and reported to the Risk Committee.	Action plan is mandatory under the responsibility of the CISO and / or responsible of geographical unit.	
				4 – 6	Moderate risk	Requires attention and needs to be reviewed annually. Reporting at the CISO level and security committee.	Regular follow-up actions as well as the assignment of a risk responsibility.	
				1 – 3	Low risk	Does not require any special actions but must be reviewed annually.	No action is planned outside the controls in place.	

Considering the risk matrix is relevant from the point of view of information classification because quantitative risk determination requires a quantitative information classification or a qualitative information classification with additional information value determination. The potential damage if the risk occurs is not only measured from the asset value, but is added to other values such as fines, loss of sales, etc. However, both the quantitative and the qualitative risk assessment methods have their advantages and disadvantages, see Table 10, which could also apply to the information value determination.

Table 10: Comparison of Quantitative and Qualitative Risk Assessment (Stallings, 2018).

	Quantitative	Qualitative
<b>Benefits</b>	<ul style="list-style-type: none"> <li>Risks are prioritized by financial impact; assets are prioritized by financial values.</li> <li>Results facilitate management of risk by return on security investment.</li> <li>Results can be expressed in management-specific terminology (for example, monetary values and probability expressed as a specific percentage).</li> <li>Accuracy tends to increase over time as the organization builds historic record of data while gaining experience.</li> </ul>	<ul style="list-style-type: none"> <li>It enables visibility and understanding of risk ranking.</li> <li>It is easier to reach consensus.</li> <li>It is not necessary to quantify threat frequency.</li> <li>It is not necessary to determine financial values of assets.</li> <li>It is easier to involve people who are not experts on security or computers.</li> </ul>
<b>Drawbacks</b>	<ul style="list-style-type: none"> <li>Impact values assigned to risks are based on subjective opinions of participants.</li> <li>The process to reach credible results and consensus is very time consuming.</li> <li>Calculations can be complex and time consuming.</li> <li>Results are presented in monetary terms only, and they may be difficult for nontechnical people to interpret.</li> <li>The process requires expertise, so participants cannot be easily coached through it.</li> </ul>	<ul style="list-style-type: none"> <li>There is insufficient differentiation between important risks.</li> <li>It is difficult to justify investing in control implementation because there is no basis for a cost/benefit analysis.</li> <li>Results are dependent upon the quality of the risk management team that is created.</li> </ul>

In research on risk assessment, there is the approach of using fuzzy sets and fuzzy logic, to some extent with machine learning algorithms, to calculate the risk. However, the information classification remains unaffected by this because it is still defined by expert opinions and these expert-based information classifications are used as the basis for the fuzzy methods. Therefore, the actually relevant problem is not solved (Alonge et al., 2020).

### 3.9.7. Risk management

As in most risk management processes, information classification is not explicitly described as a process step, see *Figure 29*. This does not reflect its importance since information classification is at the beginning of the risk management process and is the first step in deciding. This decision as to which classification category information is placed in thus decisively decides on further risk minimization measures and thus on the information security architecture.

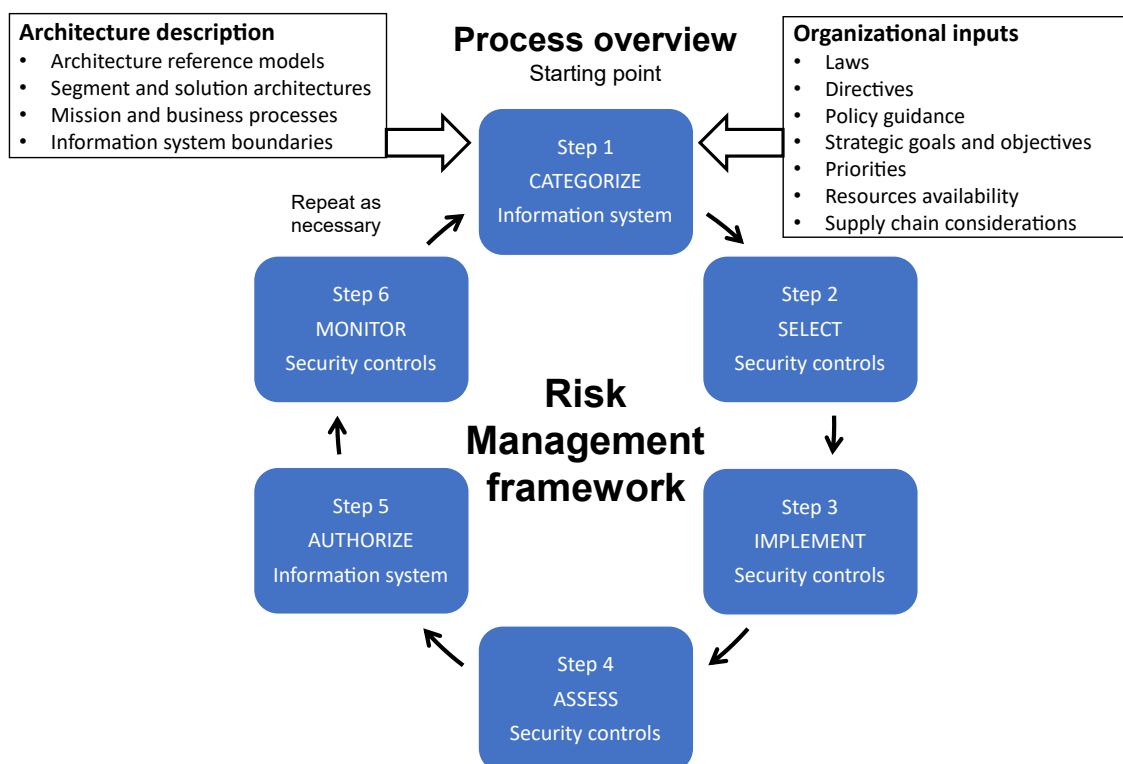


Figure 29: Risk Management Framework (National Institute of Standards and Technology, 2011).

A generic information security risk management from the analysis of various information security standards has the following elements (Fenz et al. 2014):

- System characterization:
  - System boundaries
  - System assets
    - Tangible assets
      - Persons
      - Physical assets
    - Intangible assets
      - Data
      - Software
  - Acceptable risk per asset
- Threat and Vulnerability Assessment:
  - Identify potential hazards

- Identify sources of danger
- Identify vulnerabilities
- Create a security requirements checklist
- Risk identification
  - Determine the likelihood of a potential threat exploiting a specific vulnerability on a specific system.
  - The impact of an exploited vulnerability on the organization.
  - Determine the risk as a product of the probability of occurrence and the magnitude of the impact.
- Determine security measures
  - Determine security gaps for the identified risks and define security measures.
- Evaluation of security measures
  - Evaluation of the security measures identified in terms of their cost-benefit ratio.
  - Select and implement security measures that meet the requirements and have the best cost-benefit ratio.

### 3.10. System engineering

Engineering sciences are only focused on the system of interest (SOI) and develop solutions for this system according to technical and scientific laws and regularities. Engineering sciences have no intrinsic approach to increasing organizational, economic or social added value (SEBoK Editorial Board, 2022). That would be comparable to the purely technical approach to IT security, based on firewall rules, anti-virus protection, intrusion detection, etc. But cybersecurity must be viewed as interdisciplinary applied research (Millett et al., 2017).

In contrast to engineering sciences, System Engineering (SE) is holistic, at least within the set model limits. Holistic means that not only individual systems, but the entire system with its individual systems and their dynamic relationships is considered. In the SE, however, not only technical systems, but also stakeholder interests. SE is very closely linked to Technical Management (TM), procurement and acquisition and Project Management (PM) (SEBoK Editorial Board, 2022). The SE approach is significantly more complex and integrative as it integrates multiple disciplines and stakeholders, which also corresponds to the information security approach, which integrates technical, business and security objectives. Information security and systems engineering seem to be natural allies. It would therefore be worth considering how the SE concepts and methods can be integrated into information security management. The potential benefits are obvious; A holistic, cyclical, and dynamic consideration of the information security components as a system with sets of rules that need to be identified to gradually convert them away from qualitative methods with little information content towards quantitative/logic/rule-based methods with high information content. This would fulfill the prerequisites for an automation of information security management and could enable an increase in efficiency with simultaneous cost and risk minimization of information security as whole.

Systems dynamics (SD) tries to understand systems dynamics behavior over time (SEBoK Editorial Board, 2022). This would correspond to the information (classification) life cycle that is also influenced by time.

The main goal of System Security Engineering (SSE) is to eliminate system vulnerabilities through engineering measures and to implement countermeasures through system design decisions (SEBoK Editorial Board, 2022). This includes two different perspectives on two levels of abstraction. The first is system hardening from the perspective of known vulnerabilities and attack vectors, such as the CVE and MITRE ATT&CK map. In particular, the systems are trimmed to the status of the currently known security parameters. The second, more abstract perspective, deals with unknown vulnerabilities and

attack vectors. These cannot be implemented through concrete measures to harden the system but contain general principles for designing the systems so that they are not or less vulnerable and contain also measures that enable the systems to be monitored and react more quickly, such as network design decisions, role and access models or IDS.

## 4. Discussion and conclusion

### 4.1. Discussion

The results of the first research question are contained in the main part as literature review. The literature review has shown that the classification of information in the scientific literature does not receive sufficient attention. There is insufficient scientific research, either in breadth or depth, on information classification itself or on the complex contextual relations to related scientific topics. An obvious reason for the insufficient research on the topic could not be found. A possible reason could be that the classification of information is also dealt with rather incidentally in the risk management standards. In the risk management processes of the standards, it is not even defined as an independent work step, so that the classification of information only becomes apparent when the standards are studied in detail. Another possible clue to the disregard of information classification could be its historical origin. Because its origin is the paper file in the military field. This was classified by trustworthy and competent persons and then treated according to the corresponding classification-dependent processes. The circumstances have changed significantly since then due to the various technological developments and it is not necessarily obvious that the information classification has also fully taken this change into account. The identified gaps in information classification in risk management are not new but have already been addressed in several scientific papers and risk management books. However, it has never been done to the same extent as in this thesis, in which both the scope of the literary review is large, and all identified gaps have been tabulated. Therefore, this thesis could provide a good starting point for an overview of the topic and for further in-depth research. Some scientific papers attempted to close individual gaps with new approaches. Most efforts were made in process considerations and classification categories. However, there is no holistic and sustainable approach to the existing gaps, which could also be a reason why further development of information classification is not perceptible.

The results of the second research question on the gaps in information classification research are summarized in *Table 11*. The table contains several columns; the "Source" column shows from which part of the literature review in the main part of the work the gap originates. The reference does not even have to be a standard, for example, but a paper that analyzes standards, thus the context decides the assignment. The "Gap description" column contains the text part that describes the gap or allows conclusions to be drawn about a gap. The "Reference" column is the reference to the text part or the indication that the gap is an independent conclusion of the author. The column "Research gap group" is an attempt to group together the large number of identified research gaps to get a simplified overview of the main problems.

Table 11: Summary of finding to Q2: Gaps in information classification research.

Source	Gap description	Reference	Research gap group
Standard	It is left to the users which asset or information is relevant and how it is considered.	NIST 2015	Applicability and effectiveness test is missing
Paper	Due to the very complex and diverse IT landscape and massive processing of heterogeneous data, very complex information and asset landscapes arise.	Yuan et al. 2021	Complexity handling issues
Paper	Manual activities, especially with complex interpretation steps such as in information classification, are very personnel-intensive, time-consuming, and sluggish.	Harkins 2013	Applicability and effectiveness test is missing
Book	Methodological approaches from other research areas, such as sociology, health sciences, humanities, or engineering, are rarely used in information security research.	Millett et al. 2017	Interdisciplinary research missing
Book	A scientific analysis of cybersecurity challenges backed up with well-substantiated models could help to identify the actual cybersecurity needs by linking cause and effect of attacks and addressing potential solutions.	Millett et al. 2017	Context research missing
Book	Within cyber security, there is not even a growing realization that there could be a potential interdisciplinary research question here at all, but attempts are being made to reinvent the wheel and improvise where necessary, by ignoring methods and findings from other sciences that have already been worked out very well.	Millett et al. 2017	Interdisciplinary research missing
Paper	"In literature, information classification is seldom dealt with in-depth."	Bergström et al. 2020b, p. 211	In-depth research missing
Paper	There are different definitions for information classification in the context of information security, depending on what has been identified as a good worth protecting and what position it has in the context of consideration.	Collard et al., 2017	Consents in taxonomy, ontology missing
Paper	"Still, this issue of subjective judgement is challenging and often ignored, both in practice and research."	Bergström et al. 2020b, p. 211	Bias
Paper	There is a lot of advice about implementing information classification systematically but there are few approaches to reduce subjectivity in research	Bergström et al. 2020b	Bias
Paper	There is also no consensus on its granularity. Both too coarse-grained and too fine-grained classifications lead to problems in the correct assignment of security measures.	Seifert and Relyea 2004	Consents in taxonomy, ontology and methods missing
Paper	Complexity, and granularity, is exactly why information classification is not as easy to carry out as it appears at first glance, and the previous practice of information classification has suggested	Talabis and Martin, 2012	Complexity handling issues

Paper	Information classification has no common methodological (methods, frameworks, and process) basis.	Bergström et al. 2020b	Consents in taxonomy, ontology and methods missing
Standard	Information security standards use different methods, frameworks, and process descriptions for information classification	Bergström et al. 2020b	Consents in taxonomy, ontology and methods missing
Paper	There is a discrepancy between theory and practice	Millelt et al. 2017	Applicability and effectiveness test is missing
Paper	"It is well-known that there is a gap between formal and actual processes in information security management (ISM), as turning standards into practice is easier said than done"	Bergström et al. 2020b	Applicability and effectiveness test is missing
Paper	The process of information classification is so nonspecific and complex that practitioners spend more time understanding the process than doing the task of information classification.	Bergström et al. 2020b	Applicability and effectiveness test is missing
Paper	Too many stakeholders with non-IT interests influence IT decisions.	Millelt et al. 2017	Validity check of existing assumptions necessary
Paper	Origins of information classification is in military with a confidentiality focus.	Fibikova and Müller 2011	Historical ballast
Paper	Origins of information classification is paper-based with paper-based workflows.	Bunker 2012; Bergström et al. 2020b	Historical ballast
Standard	In general, the standards are very similar, but in terms of actual implementation there is no common approach in standards to information security, also not for information classification.	Bergström et al. 2020b	Consents in taxonomy, ontology and methods missing
Standard	The ISO/IEC 27001 standard doesn't say much about information classification... so the details of how you implement the control are pretty much left up to you.	CertiKit	In-depth research missing
Standard	Immense variations in classification categories, therefore information exchange very difficult.	Own conclusion	Consents in taxonomy, ontology and methods missing
Standard	MAGERIT, a risk management framework developed by the Spanish Ministry of Administration, contains proposals for unifying the exchange formats between organizations to provide asset identification, asset type classification, asset dependencies, and asset value estimation	ENISA 2022	Consents in taxonomy, ontology and methods missing
Standard	No clear taxonomy in information classification.	Stallings 2018; Peltier 2001	Consents in taxonomy, ontology and methods missing
Standard	Information transmitted outside the organisation borders is under a new classification regime and a total loss of control is imminent.	Cherdantseva 2014	In-depth research missing

Standard	Classifications from other organisations should be taken with care.	International Organization for Standardization 2015c	Consents in taxonomy, ontology and methods missing
Standard	System engineering should be applied to cope information classification requirements.	Miller 2017	Interdisciplinary research missing
Standard	Synchronising classification schemes between organizations that exchange information.	ISO/IEC 27002:2013; ISO/IEC 27036-2:2014	Consents in taxonomy, ontology and methods missing
Standard	Missing an information classification maturity model, as it exists in information security.	Own conclusion	In-depth research missing
Standard	The assumptions, recommendations and specifications in security standards and their maturity models neither follow scientific methods nor are comprehensible	Siponen 2003	In-depth research missing
Standard	It is of utmost importance to explain why a security measure should be taken, what its actual (real) effects are and what the underlying implications are.	Siponen 2003	Applicability and effectiveness test is missing
Standard	Security issues increase with e-government, as country-specific standards and certifications exist, with different approaches to information classification schemes, information storage strategies, information reliability assurances, security procedures, risk management strategies, and data inconsistencies	Abdugaffarovich et al. 2015	In-depth research missing
Standard	Missing hierarchical information classification measures model: Hierarchies based on reliability on classification results. Information classification approaches are at least: automated, semi-automated, knowledge-supported, and manual.	Own conclusion	In-depth research missing
Standard	No standards found about automated information classification.	Own conclusion	In-depth research missing
Standard	No literature found about automated information classification.	Own conclusion	In-depth research missing
Standard	Impact of classification categories on the locality (on-premises, cloud) of information, services and assets.	Own conclusion	In-depth research missing
Standard	The CISO's don't have the skills necessary for the risks of the 21st century. The skills should include broad technological and business knowledge and in-depth knowledge of risk and security	Harkins 2013	Validity check of existing assumptions necessary
Standard	The biggest influencing factor on information security in SMEs is the lack of knowledge and experience about information security. Currently it is addressed by simplification in information security and information classification.	Kobis 2020	Validity check of existing assumptions necessary
Standard	Relying entirely on information owners for information classification is a risk because it is very time-consuming and error-prone for subjectivity and inconsistency	Bergström and Åhlfeldt 2014	Validity check of existing assumptions necessary

Standard	Oversimplification of risk management processes in practice leads to gaps in the implementation of the required risk management, no standard-based processes, in too little granularity of the analysis of assets, risks, vulnerabilities, and dangers, which in turn leads to wrong assumptions and conclusions	Piya Shedden et al. 2006	Validity check of existing assumptions necessary
Standard	There is little literature on the actual implementation and use of risk assessment in organizations.	Piya Shedden et al. 2006	In-depth research missing
Standard	There is no scientific proof of how effective risk management recommendations from science and standards actually are.	Piya Shedden et al. 2006	Applicability and effectiveness test is missing
Standard	There is not much scientific work about the gaps in information classification.		In-depth research missing
Standard	The use of information security standards is often assumed to be easy to use and thus implementable for risk management. But this is wrong.	Piya Shedden et al. 2006	Validity check of existing assumptions necessary
Paper	Information classification can have many names, such as data classification or security classification, and naming seems to depend on the context	Own conclusion	Consents in taxonomy, ontology and methods missing
Paper	Low granularity classification can lead to information overprotection	Bergström et al. 2020b	In-depth research missing
Paper	Organizations struggle with the information classification lifecycle, as they don't know when to classify and to reclassify.	Bergström et al. 2018	Applicability and effectiveness test is missing
Paper	From the Information Lifecycle Management point of view, there is a need for meaningful classification schemes and automated classification techniques such as pattern recognition	Reiner et al. 2004	Consents in taxonomy, ontology and methods missing
Paper	Hard-to-interpret policies can lead to human error, subjectivity, and inconsistent ratings.	Bergström and Åhlfeldt 2014	Bias
Paper	A lack of Information Lifecycle Management prevents correct archiving or necessary reclassifications.	Bergström and Åhlfeldt 2014	In-depth research missing
Paper	A lack of resource management can jeopardize the entire classification process due to a lack of resources.	Bergström and Åhlfeldt 2014	In-depth research missing
Paper	Inaccurate or too rough description of the information classification process leads to difficulties in implementation and thus to a lack of effectiveness and efficiency.	Bergström and Åhlfeldt 2014	In-depth research missing
Paper	Little scientific interest in the topic of information classification in the context of information security.	Bergström and Åhlfeldt 2014	Interdisciplinary research missing
Paper	Too little practical and empirical research.	Bergström and Åhlfeldt 2014	In-depth research missing
Paper	The recommendations of the information security standards are unverified and consequently not well-founded, which puts the entire information security process in doubt.	Bergström and Åhlfeldt 2014	Applicability and effectiveness test is missing



Paper	A common information classification term is missing.	Bergström and Åhlfeldt 2014	Consents in taxonomy, ontology and methods missing
Paper	The classification takes place based on coarse classification factors.	Bergström and Åhlfeldt 2014	Bias
Paper	Automated information classification is only theoretically researched.	Bergström and Åhlfeldt 2015	Interdisciplinary research missing
Paper	Information labelling is only theoretically researched.	Bergström and Åhlfeldt 2016	Interdisciplinary research missing
Paper	Cyber-attacks are successful because information security still thinks and acts in linear, static, and reacting cause-and-effect analyses.	Schiavone et al. 2014	Interdisciplinary research missing
Paper	But certain information can have a different meaning or value for different information users, which can be due to the different use of the information and the different understanding of the information	Xianliang Shi et al. 2007	In-depth research missing
Paper	There is a need for additional information on the information classification to make more informed decisions	Lundgren 2020	Interdisciplinary research missing
Paper	Information security lacks already a uniform terminology.	Cherdantseva 2014	Consents in taxonomy, ontology and methods missing
Paper	The research does not sufficiently consider the discrepancy between the simplistic approach from the standards and the complex organizational practice, but has a very narrow focus on information classification	Bergström et al. 2018	Applicability and effectiveness test is missing
Paper	The challenges of information classification are inadequately analysed in science. Challenges identified include inconsistent classification due to subjectivity, overly complex or inappropriate classification schemes, inadequate description of classification processes or classification criteria, impact of personal worldviews, granularity of classification, impractical best-practice solutions, overly abstract standards, lacking necessary status in practice, lack of information knowledge, problems in interpretation, credibility problems.	Bergström et al. 2020b	Applicability and effectiveness test is missing
Paper	A methodology to reduce subjectivity in classification is lacking.	Bergström et al. 2020b	Bias
Paper	Information classification research mainly focuses on qualitative information classification principles and the management of the classified information	Xianliang Shi et al. 2007	Interdisciplinary research missing
Paper	Research results from the validation of the actions recommended in the standards must be available to practitioners.	Siponen and Willison 2009	Applicability and effectiveness test is missing

Paper	<p>Possible reasons for classification shortcomings:</p> <ul style="list-style-type: none"> <li>• Classification classes have no class properties.</li> <li>• Property negation as classifier.</li> <li>• Class membership can easily be changed.</li> <li>• Classification does not enable class behaviour prediction.</li> <li>• Missing class relations.</li> <li>• Classification complicates domain understanding.</li> <li>• Classifying unclassifiable objects.</li> </ul>	Berman 2022	In-depth research missing
Paper	The goals of information security are ambiguous and unclear because no scientific consensus has been found. On the one hand, the CIA triad is considered outdated and no longer adequate, but then the goals diffuse.	Cherdantseva 2014	Consents in taxonomy, ontology and methods missing
Paper	There are arguments for a formalized expertise in information security. However, this is likely to be a very lengthy process and can only be done with the cooperation of many.	Cherdantseva 2015	Consents in taxonomy, ontology and methods missing
Paper	With all this complexity in the information classification process, people would not be able to cope anyway.	Lundgren and Bergström 2019	Complexity handling issues
Paper	Information classification processes are just generically described and have not much to do with the complex reality.	Own conclusion	Applicability and effectiveness test is missing
Paper	Simplified risk analysis models enable laypersons to carry out the risk analysis, but they contain potential false assumptions about reality as well as increased inaccuracy.	Fenz et al. 2014	Bias
Paper	Identifying relevant sources of information for decision-making in information classification is missing.	Own conclusion	In-depth research missing
Paper	Decision models and decision tables are missing in the information classification processes.	Own conclusion	In-depth research missing
Paper	Some authors (Lundgren and Bergström 2019) state a fundamental criticism on the functional approach to risk management and propose a strong emphasis on social parameters as solution approaches in information classification.	Own conclusion	Unverified claim
Book	Risk management processes from standards appear very formal, giving the impression of a very functional, even technocratic, approach. However, on closer inspection, the descriptions are rather very text-heavy and scattered, which contain very few and extremely imprecise concepts, methods, functions, models, and processes.	Own conclusion	Applicability and effectiveness test is missing
Book	„It is of course not possible to develop generic guidelines that fit exactly for all organizations, and each individual organization needs to develop their own scheme with their own classification categories.”	Bergström and Åhlfeldt 2014	Unverified claim

Book	Missing specificity of the classification categories and the resulting need for infinitely many scientifically unfounded classification schemes give rise to problems, namely subjectivity, inconsistency, over- or under-classification.	Own conclusion	Consents in taxonomy, ontology and methods missing
Book	The value in information classification is emphasized too much and the importance for the organization is neglected.	Xianliang Shi et al. 2007	Unverified claim
Book	Information classification lacks a unified yet flexible classification framework.	Alonge et al. 2020	Consents in taxonomy, ontology and methods missing
Book	Unified frameworks and ontologies can help build knowledge, share information, and make knowledge available to the public in a unified way.	Vargas and Fenz 2012	Consents in taxonomy, ontology and methods missing
Book	Information classification knowledge sharing.	Fenz et al. 2014	Community
Book	Information classification does not receive the attention it deserves from the scientific literature. Which is why the practical difficulties in using it in practice are well known, but their causes are largely unknown.	Bergström et al. 2020	In-depth research missing
Book	There is a need of high-quality data in information classification decisions considering the various data protection regulations.	DAMA International	In-depth research missing
Book	We need a scientific debate in information classification on the relationships between data, information, and knowledge, and whether new information is created through the chaining of data and whether changes in information classification are necessary as a result.	Campbell 2016	In-depth research missing
Book	The prevailing opinion in the literature almost all duties and responsibilities for information classification, information classification life cycle, information life cycle and access permissions to the information owner should be reviewed. Such a requirement cannot be implemented in practice because information owners are experts for other tasks and the available capacities are limited.	Own conclusion	Validity check of existing assumptions necessary
Book	We need more research on bias in information classification to generate insights for more objective information classification.	Own conclusion	In-depth research missing
Book	To determine the degree of implementation of the information classification process, a generally applicable key performance indicators framework is required.	ISO27k implementers 2007; Kim and Solomon 2018	In-depth research missing
Book	For information classification, a process should be defined that also identifies and classifies the assets across the OSI stack and other infrastructure that store, process, transfer or delete classified data.	Kim and Solomon 2018; Brunner and Suter 2008; Gregory 2018; Volchikov 2018	In-depth research missing

Practice	Broad surveys are needed to determine the reasons why organizations rarely implement information classification, and how this could be remedied.	Bergström et al. 2018	Applicability and effectiveness test is missing
Practice	The risk management process is rather static in practice and dynamic in theory. What are the causes for this gap?	Lundgren and Bergström 2019	In-depth research missing
Practice	Sharing classifications of information with other organizations is not possible because the value of information is assessed from the perspective of the organization (or person within the organization) and the value may be different to another organization.	Bergström et al. 2018	Consents in taxonomy, ontology and methods missing
Practice	A fundamental problem of information classification is that it does not receive the necessary status in practice. Possibly because information classification processes have not been researched enough and therefore no relevant transfer into practice has taken place.	Kindervag et al. 2015; Shedden et al. 2010; Shedden et al. 2016; Ghernaoui-Helie et al. 2011; Bergström et al. 2020b	In-depth research missing
Practice	Organisations were not able to follow the dynamic changes in the information and system landscape quickly enough, so that in the end cutbacks were made in the risk analysis process and a static risk analysis prevailed overall.	Lundgren 2020	In-depth research missing
Practice	Sometimes organization's confidentiality rules hinder research about information classification.	Chen et al. 2015	In-depth research missing
Practice	Knowledge sharing for best practice and processes in security is not common in public and private sector.	Alsmadi 2019	Community
Practice	The purely qualitative and manually executed information classification approach should be fundamentally questioned.	Own conclusion	Validity check of existing assumptions necessary
Practice	The best practice prescriptions were insufficient for local action.	Bergström et al. 2020b	Applicability and effectiveness test is missing
Practice	The total cost of the information classification process should be identified. Firstly, according to the status as suggested by the information security standards and most of the scientific literature, namely manually. Secondly, alternatively with different levels of automation. Ultimately, a cost comparison is needed to direct the research focus to the most efficient alternative.	Own conclusion	In-depth research missing
Practice	The formal risk analysis process of ISO 27005 does not run smoothly in practice, but causes socio-organizational friction losses, so that organizations make extensive adjustments to the recommended process to remain able to act.	Lundgren 2020	Applicability and effectiveness test is missing
Context	A major weakness of decisions made by people is bias, which is fed by various motivators and should be analysed in information classification, as it relies heavily	Peón et al. 2017; Wit et al. 2021	Bias

	on people. And experts are not less biased than laypeople.		
Context	The cybersecurity community needs to be clear about whether they align themselves more closely with engineering and science, or with the humanities. Consequently, whether pure, hard facts, and applied knowledge or soft knowledge make a greater contribution to gaining knowledge and deliver the more effective and efficient approach in practice, and where the limits of interdisciplinary approaches are to be drawn.	National Institute of Standards and Technology 2020	Interdisciplinary research missing
Context	It needs to be considered what methods of quantitative information classification would be necessary and practicable to enable quantitative risk analysis. Or in other words, how can information be measured as value?	Own conclusion	In-depth research missing
Context	How can machine learning and statistical data analysis be utilized in information classification?	Own conclusion	In-depth research missing
Context	In risk management processes, information classification is not explicitly described as a process step. This does not reflect its importance since information classification decisively decides on further risk minimization measures and thus on the information security architecture.	Own conclusion	Validity check of existing assumptions necessary
Context	On the one hand, information classification, if fully implemented, is complex and must take many requirements and systems into account. On the other hand, complex system considerations are at home in system engineering. Therefore, research to check the compatibility of system engineering and information classification might be a gain in knowledge and could provide methods for a systematic "information classification engineering".	Own conclusion	Interdisciplinary research missing

The table is quite large because it contains 109 potential gaps. Therefore, a graphical overview has been created that summarizes the gaps in the table very roughly, so that a simple overview can be gained.

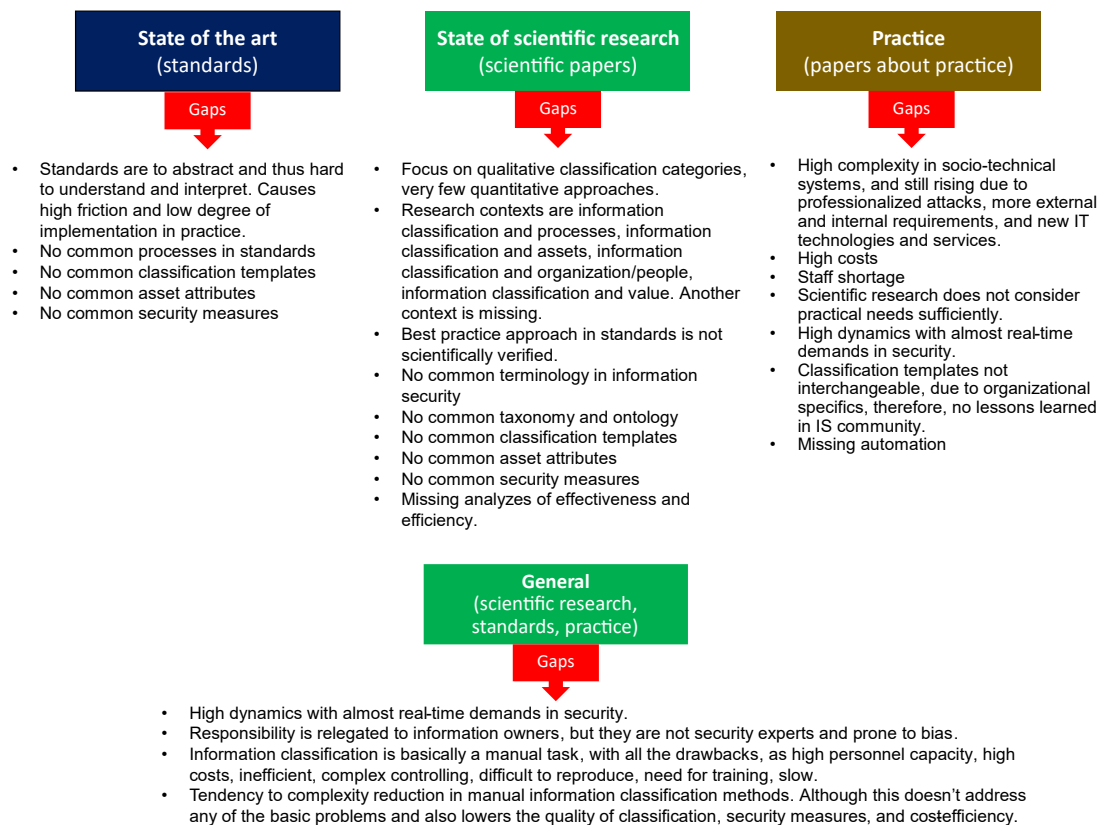


Figure 30: Very simplified representation of the table "Summary of finding to Q2: Gaps in information classification research."

Discussing all gaps would not be feasible, thus it is referred to the summary of findings in *Table 11*, the context in the main part of the literature analysis and the original source for evaluating the gap.

This thesis attempted a systematic literature review about information classification in information security risk management. With only a small amount of primary literature available on the subject, it was necessary to conduct context research to get an overall picture. To get a structured overview of the gaps in information classification, the literature review was split into the areas "scientific literature", "standards", "books", and "practice". The aim of this differentiation was to show the gaps in the respective areas and to show possible approaches for necessary cross-area research. However, a differentiation was not always easy and there were smooth transitions and overlaps. This was because the "practice" information was obtained from the scientific and book literature as no practice-related reports or studies could be found. The information classification standards were also often discussed in the scientific literature, which is why there were overlaps there as well. That doesn't have to be a disadvantage for dealing with the content of the topic, but rather a formal one, because it was not always possible to clearly separate the blocks of topics and the literature.

Several difficulties in the literature were encountered:

- A scattered knowledge base for information classification.
- Various standards that cover the topic of information classification superficially and spread the topic scattered over several different publications.
- In principle, the standards follow a uniform model, but they differ in their form and are therefore difficult to combine.
- In the scientific literature, too little and too selective research has been done on information classification, which makes it very difficult to deal with the topic, because the knowledge about it had to be gained from contextual literature.

- Lack of consensus on vocabulary, processes, and models in information classification.
- Lack of scientific and practical evidence of the effectiveness of the recommendations for implementing information classification.
- Little relation to practice because practice has great difficulties in interpreting the standards.
- Very strong focus on people, i.e., implementation and monitoring of information classification in practice.
- Information classification plays a subordinate role in the risk management process and is not even listed as an independent item in most standard models and processes.
- Dynamic aspects, such as changes in information and systems, play just a negligible role in the information classification literature.
- Automation, machine learning, and data analysis play virtually no role in the information classification literature.
- System engineering thinking, i.e., thinking in systems of system, systems and dynamic dependencies plays no role.
- Taxonomies, ontologies, architectures, and data models for information classification are as good as non-existent or if they exist very superficial.
- Information classification processes are described too generically and are therefore of little practical use. In addition, these process descriptions lack decision support models (DMM) and decision tables to achieve practical and scientific relevance within this context of decisions on classifying information.
- Overall, since its inception, information classification does not appear to have matured in any relevant way to deal with technical developments, practical requirements, and security risks.
- There even seems to be a lack of awareness in the IT risk management community of the high relevance of information classification, although this has a decisive influence on the entire information security architecture.

It can be noticed that a systematic literature review is very time-consuming and that a thesis forms a time frame that is almost too short, and in the given time frame, the status of the thesis was the most feasible. Thus, this systematic literature review had its limits. Dealing with the small amount of primary literature about information classification it was necessary to consult literature in a thematically close context and analyze the full texts. The search for this contextual literature and its content-related full-text analysis was very time-consuming. In addition, there is no certainty that all relevant literature has been identified and considered. Another limitation was the lack of metadata for the full texts, which prevented a quantitative analysis of the literature. A quantitative literature analysis could have given additional information about the topics and author relationships and thus possible new conclusions.

The strength of this research approach lay in the attempt to create a complete and comprehensive status picture for information classification. The weakness was that a slightly deviating systematic literature review approach must be designed to conduct relevant research despite the few primary sources. It was a weakness because it was not a tried-and-tested procedure and therefore involved risks.

The literature reviews to date have traditionally remained rooted in the subject area and the research question. The approach in this thesis, on the other hand, also linked related subject areas with information classification to relate new concepts, ideas, and research results. This form of extension was previously unknown in the scientific literature review on information classification.

The contribution of this thesis to research into information classification is given because such a comprehensive and broad summary of information classification in risk management of information security has not yet been carried out. By dividing the analysis into the three areas where information

classification plays a role, namely standards, research, and practice, made it possible to map the research gaps to these areas. This can be an advantage when it comes to generating new research questions from the identified gaps and assigning them to a context. The number of identified gaps in research and practice on information security classification is high and could be a helpful source of inspiration for further research.

## 4.2. Conclusion

Our research questions

- Q1: What is the current state of science on information classification?
- Q2: What research gaps are there in information classification research?

have been answered so far. The first research question was answered by a systematic literature review using standards and scientific literature as sources.

**Q1:** The status of information classification was answered from different perspectives, from the perspective of scientific research, from the perspective of risk management standards, from the perspective of book literature, and from the perspective of practice. This distinction helped us to work out a clear demarcation of the different levels of knowledge in the perspectives.

It turned out that the scientific perspective is very narrowly limited to the classification of information itself. There are few attempts at holistic or interdisciplinary approaches. An important interdisciplinary approach is the concrete consideration of the information classification processes as they are described in the literature and implemented in practice. Consideration of information classification categories also plays a prominent role in scientific literature. In addition, the complexity in information classification plays a role, especially when it comes to the complexity reduction of the process itself, including approaches to reducing complexity in the context of information classification categories. Most of the scientific literature considers the information classification process to be a manual and expert-dependent activity, with a strong emphasis on social and organizational aspects. As a result, machine learning approaches for automated information classification in risk management of information security do not yet play a role.

The standards, on the one hand, consider information classification superficially and leave the actual design of the classification processes and classification categories to the users. On the other hand, it is strictly specified to which classification category certain assets or asset groups are to be assigned. The rather lax approach is with the ISO standards and the very strict one with the specifications with the NIST and FIPS specifications for US government organizations. It was not possible to determine whether a strict categorization specification is a more advantageous alternative to free classification. It represents at least an approach worth considering. Overall, the standards provide a very good overview of the risk management process by embedding information classification and thus some context information, such as on asset management. A weakness of the standards, however, is the distribution of information over many documents and standards, which, as in the case of the ISO standards, are very expensive and therefore not all of them were available. This circumstance could also represent an obstacle in practice.

In the books on risk management and on cyber or information security, there were explanations that take different perspectives into account, those of the standards, the scientific literature and practice. Consequently, the most comprehensive overview of the topic of information classification could be gained here. However, the books differed greatly in quality. High-quality book sources were identified as those that processed relevant standards and scientific literature and contributed relevant contextual information from information security practice. Therefore, only a few of the identified



books are worth recommending. The contextual information from the books was important because the reference to system engineering was evident here. Individual systems, such as information classification, with asset management, information lifecycle management, etc. were brought into context and their dynamic and mutual relationships became apparent and are therefore typical properties of connected systems and thus describe a field of activity in system engineering.

In scientific articles and books perspectives of information classification practice were also described. It should be particularly emphasized here that difficulties arise in practice with the interpretation and implementation of the information classification. It is emphasized, the insufficiently specific descriptions in the standards and that information classification research are not oriented towards practice.

The second research question Q2 was answered by capturing and summarizing the gaps in standards, research, and practice.

## **Q2:**

The second research question dealt with the research gaps in information classification in risk management of information security. An attempt was made to group the identified research gaps to get a better overview. The first alternative grouping was by the areas of research, standards, practice, context, and book literature concerned. 109 individual gaps were identified, of which 48 are gaps in scientific research, 27 gaps directly related to risk management standards, 17 gaps in books (on gaps in science, standards, and practice), eleven gaps directly related to practice, and six gaps with a context to other topics.

The second grouping alternative was according to the citation of the gaps identified; 88 of the 109 gaps could be identified in the literature sources, and 21 gaps were identified by own conclusions. Most of the gaps were explicitly mentioned in the literature, but there were also gaps that were identified as such from the description of the facts.

The last grouping alternative represents the attempt to define generic terms for the problem types of the gaps. Each research gap from the table was assigned to a generic term that seemed sensible. The number of research gaps contained in each generic term was recorded in *Table 12*. The results of the research question Q2 thus revealed 34 research gaps on the topic of in-depth research, 20 research gaps on the topic of standardization/consensus finding (taxonomies, ontologies and methods), 17 research gaps on the topic of practical applicability and effectiveness, eleven research gaps in interdisciplinary research, nine necessary validity checks of existing assumptions, seven existing biases that require further scientific investigation, three overcomplexity issues, three unverified claims, two research gaps on the non-existent community in information classification, two assumptions that have grown historically and should be reviewed, and one gap due to lack of contextual research.

Table 12: Issue topics of the identified research gaps in information classification.

Issue topic	Count
In-depth research missing	34
Consens in taxonomy, ontology and methods missing	20
Applicability and effectiveness test is missing	17
Interdisciplinary research missing	11
Validity check of existing assumptions necessary	9
Bias	7
Complexity handling issues	3
Unverified claim	3
Community	2
Historical ballast	2
Context research missing	1
<b>Sum of issues identified</b>	<b>109</b>

In summary, it can be concluded that the state of the art and the state of research on information classification is not sufficiently considered and researched. As a result, the knowledge about information classification is selective, incomplete, and methodologically uncertain. The gaps identified are numerous and offer multiple theoretical and practical future research.

### 4.3. Future research directions

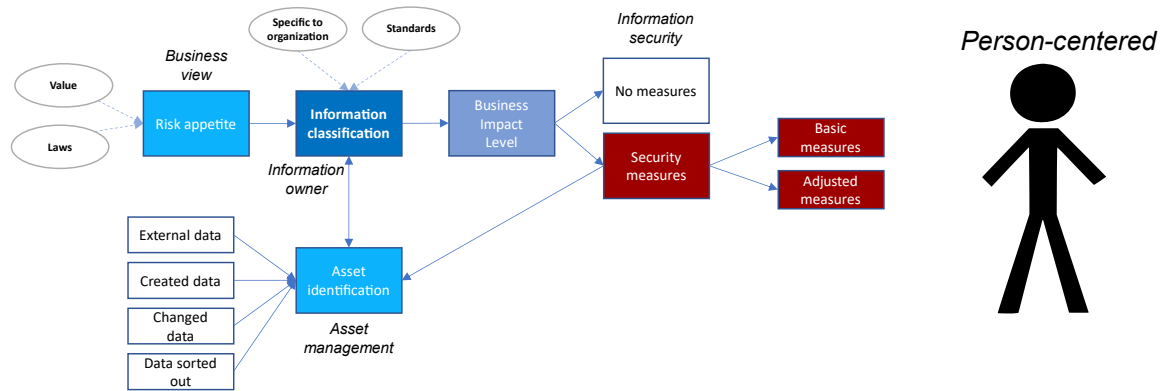
This thesis was an opportunity to take a closer look at the topic. The identified gaps in the literature on information classification and the context-relevant topics provide a picture of the state of information classification that may raise future research questions. Due to the gaps identified, it seems that some fundamental preparatory work in research on information classification must be made to create the basis for a maturity development in the research area. It can be concluded that complex scientific research in information security ultimately requires consensus and a scientifically sound basis. Practice would also benefit from a uniform and detailed model of information classification as it would be easier to understand, thus be more effective and more efficient to implement.

In conclusion, the following areas could be researched in future to close the identified gaps in information classification:

- A consensus is needed in information classification taxonomy and ontology, as a prerequisite for scientific and practical information exchange and potential synergies.
- There is a need for practice-oriented research in the field of data models for information classification, including a clear distinction between data, information, and knowledge in information classification. Furthermore, data models are needed for the information to be classified, the classification categories, the metadata for marking the classified information, the context information (such as security measures, information life cycle, information classification life cycle).
- The research community should critically analyze whether the person-centered approach to information classification is the right way to go, or whether there are at least equivalent alternatives. For this, the influences of the bias on the information classification life cycle should be analyzed.
- The information classification should adapt the models and methods of system engineering because this is a potentially successful approach to mastering the complex, dynamic and temporary information properties, and system interactions in information classification.
- Overall, there is a need for more systematic research on information classification and better information exchange between science, standardization organizations and practice. For this purpose, in addition to a common vocabulary for information classification, a common

knowledge base must also be established, such as a "Body of Knowledge of Information Classification".

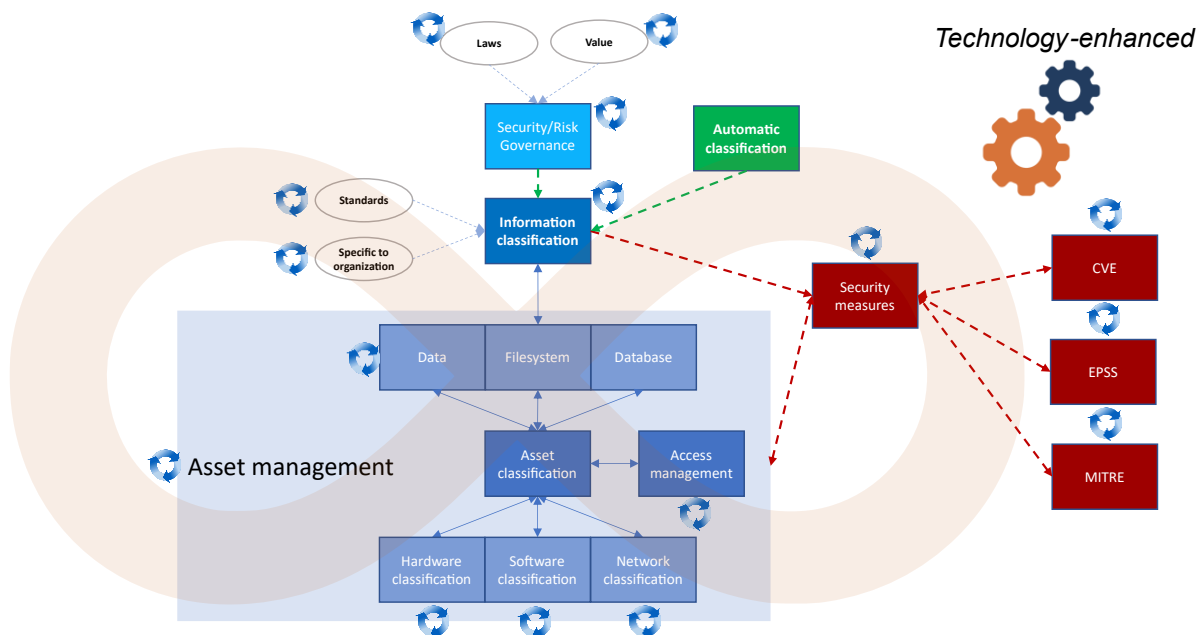
Having a look at the results of the thesis on information classification from a conceptual and methodological point of view, then the real concept of information classification was and is static (see *Figure 31*). There is a very strong personal reference, static analyzes and static relationships to other areas.



*Figure 31: Information classification as it is - static.*

This static property of information classification is at the root of many of the gaps identified, such as the information classification life cycle, or adaptability to changing conditions.

A change to a dynamic conception of information classification is required, where not only the information classification is dynamic, but all related systems, the relationships themselves and the entire system from all individual components. This leads to a life cycle model and SE. There is also a need for greater quantifiability, and information classification based on data and machine learning to reduce personal reference and thus the susceptibility to bias and to be able to use external data interfaces (see *Figure 32*).



*Figure 32: Information classification as it should be - dynamic.*

## References

- Abdugaffarovich, A., Abbasovich, V., & Bakhtiyarovich, N. (2015). E-Government, Open Data, and Security: Overcoming Information Security Issues with Open Data. *Computer Science and Information Technology*, 3, 133–137. <https://doi.org/10.13189/csit.2015.030407>
- Al-Fedaghi, S. (2008). On information lifecycle management. In *2008 IEEE Asia-Pacific Services Computing Conference* (pp. 335–342). IEEE. <https://doi.org/10.1109/APSCC.2008.81>
- Alonge, Arogundade, Adesemowo, Ibrahalu, Adeniran, & Mustapha (2020). Information Asset Classification and Labelling Model Using Fuzzy Approach for Effective Security Risk Assessment. In *2020 International Conference in Mathematics, Computer Engineering and Computer Science (ICMCECS)*.
- Alsmadi, I. (2019). *The NICE cyber security framework: Cyber security intelligence and analytics*. Springer.
- Asmussen, C. B., & Møller, C. (2019). Smart literature review: a practical topic modelling approach to exploratory literature review. *Journal of Big Data*, 6(1). <https://doi.org/10.1186/s40537-019-0255-7>
- Baskerville, R. (1993). Information-Systems Security Design Methods - Implications For Information-Systems Development. *ACM Comput. Surv.*, 25(4), 375–414. <https://doi.org/10.1145/162124.162127>
- Bergström, E., & Åhlfeldt, R.-M. (2014). Information Classification Issues. In *NordSec*.
- Bergström, E., Åhlfeldt, R.-M., Karlsson, F., Söderström, E., & Furnell, S. (2020). *Supporting Information Security Management - Developing a Method for Information Classification* [University of Skövde, Skövde]. Zotero (via BibTeX-Export).
- Bergström, E., Anteryd, F., & Åhlfeldt, R.-M. (2018). Information Classification Policies : An Exploratory Investigation. In Bergström, E., Karlsson, F., & Åhlfeldt, R.-M. (2020). Developing an information classification method. *INFORMATION and COMPUTER SECURITY*, 29(2), 209–239. <http://proxy.lib.ltu.se/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=edsemr&AN=edsemr.10.1108.ICS.07.2020.0110&site=eds-live&scope=site>
- Berman, J. J. (2022). *Classification made relevant: How scientists build and use classifications and ontologies* / Jules J. Berman. Academic Press.
- Bernard, R. (2007). Information lifecycle security risk assessment: A tool for closing security gaps. *Computers & Security*, 26(1), 26–30. <https://doi.org/10.1016/j.cose.2006.12.005>
- Bjorck, F. J. (2005). Discovering information security management. In
- Brunner, E. M., & Suter, M. (2008). *International CIIP handbook 2008/2009: An inventory of 25 national and 7 international critical information infrastructure protection policies*. Center for Security Studies, ETH Zurich.
- Bunker, G. (2012). Technology is not enough: Taking a holistic view for information assurance. *Information Security Technical Report*, 17(1-2), 19–25. <https://doi.org/10.1016/j.istr.2011.12.002>
- Calder, A., & Watkins, S. (2015). *IT governance: An international guide to data security and ISO27001/ISO27002* (Sixth edition). Kogan Page.
- Campbell, T. (2016). *Practical information security management: A complete guide to planning and implementation* / Tony Campbell. Apress.
- Cao, C., Yuan, L.-P., Singhal, A., Liu, P., Sun, X., & Zhu, S. (2018). Assessing Attack Impact on Business Processes by Interconnecting Attack Graphs and Entity Dependency Graphs. In F. Kerschbaum & S. Paraboschi (Eds.), *Data and Applications Security and Privacy XXXII* (pp. 330–348). Springer International Publishing.
- Centers for Disease Control and Prevention. (2023). *Hierarchy of Controls*. <https://www.cdc.gov/niosh/topics/hierarchy/>
- CertiKit. *A Guide to Implementing the ISO/IEC 27001 Standard*. CertiKit. <https://certikit.com/free-guide-implementing-iso27001-standard/>
- Chen, P. S., Yen, D. C., & Lin, S.-C. (2015). The Classification of Information Assets and Risk Assessment: An Exploratory Study Using the Case of C-Bank. *J. Glob. Inf. Manage.*, 23(4), 26–54. <https://doi.org/10.4018/jgim.2015100102>
- Cherdantseva, Y. (2014). *Secure\*BPMN : a graphical extension for BPMN 2.0 based on a reference model of information assurance & security* [PhD, Cardiff University]. Zotero (via BibTeX-Export). <https://orca.cardiff.ac.uk/id/eprint/74432/>

- Collard, G., Ducroquet, S., Disson, E., & Talens, G. (2017). A definition of Information Security Classification in cybersecurity context. In *2017 11th International Conference on Research Challenges in Information Science (RCIS)*.
- Cooper, H. M. (1988). Organizing knowledge syntheses: A taxonomy of literature reviews. *Knowledge in Society*, 1(1), 104–126. <https://doi.org/10.1007/BF03177550>
- Costin, D., & Militaru, C. (2011). Asset management towards ISO/IEC 27001:2005 accreditation of an information security management system. In
- DAMA International. (2017). *DAMA - DMBOK: Data Management Body of Knowledge*. Technics Publications Llc.
- ENISA. (2022). *ENISA Report - Compendium of Risk Management Frameworks with Potential Interoperability*.
- Erik, B., Martin, L., & Åsa, E. (2019). Revisiting information security risk management challenges: a practice perspective. *INFORMATION and COMPUTER SECURITY*, 27(3), 358–372. <https://doi.org/10.1108/ics-09-2018-0106>
- ETSI. (2015). *Critical Security Controls for Effective Cyber Defence: TR 103 305 - V1.1.1*. ETSI - European standards organization in ICT.
- ETSI. (2016). *Critical Security Controls for Effective Cyber Defence; Part 1: The Critical Security Controls: TR 103 305-1 - V2.1.1*. ETSI - European standards organization in ICT.
- ETSI. (2023, March 15). *ETSI GS ISI 002 V1.2.1 - Information Security Indicators (ISI) - Event Model - A security event classification model and taxonomy*. ETSI - European standards organization in ICT. [https://www.etsi.org/deliver/etsi\\_gs/ISI/001\\_099/002/01.02.01\\_60/gs\\_ISI002v010201p.pdf](https://www.etsi.org/deliver/etsi_gs/ISI/001_099/002/01.02.01_60/gs_ISI002v010201p.pdf)
- Faizi, A., Padyab, A., & Naess, A. (2021). From rationale to lessons learned in the cloud information security risk assessment: a study of organizations in Sweden. *INFORMATION and COMPUTER SECURITY, ahead-of-print*(ahead-of-print). <https://doi.org/10.1108/ics-03-2021-0034>
- Fenz, S., Heurix, J., Neubauer, T., & Pechstein, F. (2014). Current challenges in information security risk management. *Information Management & Computer Security*, 22(5), 410–430. <https://doi.org/10.1108/imcs-07-2013-0053>
- Fibikova, L., & Müller, R. (2011). A Simplified Approach for Classifying Applications. In N. Pohlmann, H. Reimer, & W. Schneider (Eds.), *ISSE 2010 Securing Electronic Business Processes* (pp. 39–49). Vieweg+Teubner. [https://doi.org/10.1007/978-3-8348-9788-6\\_4](https://doi.org/10.1007/978-3-8348-9788-6_4)
- Freund, J., & Jones, J. (2014). *Measuring and Managing Information Risk: A FAIR Approach*. Butterworth-Heinemann.
- Friedman, B., & Schneider, F. B. (2015). Incentivizing quality and impact: evaluating scholarship in hiring, tenure, and promotion. *Computing Research Association Best Practices Memo*.
- Ghernaoui-Helie, S., Simms, D., & Tashi, I. (2011). Protecting information in a connected world: A question of security and of confidence in security. In *2011 14th International Conference on Network-Based Information Systems* (pp. 208–212). IEEE. <https://doi.org/10.1109/NBiS.2011.38>
- Grant, M. J., & Booth, A. (2009). A typology of reviews: an analysis of 14 review types and associated methodologies. *Health Information & Libraries Journal*, 26(2), 91–108. <https://doi.org/10.1111/j.1471-1842.2009.00848.x>
- Gregory, P. H. (2018). *CISM Certified Information Security Manager All-in-One Exam Guide*. McGraw-Hill Professional; McGraw Hill.
- Gundu, T., & Flowerday, S. V. (2012). The enemy within: A behavioural intention model and an information security awareness process. In H. S. Venter, M. Looock, & M. Coetzee (Eds.), *2012 INFORMATION SECURITY FOR SOUTH AFRICA (ISSA)*. IEEE. [https://doi.org/10.1016/0749-5978\(91\)90020-T](https://doi.org/10.1016/0749-5978(91)90020-T).
- Harkins, M. (2013). *Managing risk and information security: Protect to enable. The expert's voice in information technology*. Apress; Distributed to the book trade worldwide by Springer Science+Business Media.
- Hausner, E. *Process of information retrieval for systematic reviews and health technology assessments on clinical effectiveness*. [https://www.eunethta.eu/sites/5026.fedimbo.belgium.be/files/2015-07-13\\_Guideline\\_Information\\_Retrieval\\_final.pdf](https://www.eunethta.eu/sites/5026.fedimbo.belgium.be/files/2015-07-13_Guideline_Information_Retrieval_final.pdf)
- Heide, M., & Villeneuve, J.-P. (2020). From secrecy privilege to information management: A comparative analysis of classification reforms. *Government Information Quarterly*, 37(4), N.PAG-N.PAG. <https://search.ebscohost.com/login.aspx?direct=true&db=lxh&AN=146413188&site=ehost-live>
- Humphreys, E. (2016). *Implementing the ISO/IEC 27001 ISMS sStandard* (Second edition). Artech House.
- International Organization for Standardization (2013a). *Security techniques: Code of practice for information security controls (ISO/IEC 27002:2013)*. International Organization for Standardization.

- International Organization for Standardization (2013b). *Security techniques: Information security management systems - Requirements* (ISO/IEC 27001:2013). International Organization for Standardization.
- International Organization for Standardization (2014). *Security techniques: Information security for supplier relationships - Part 2 Requirements* (ISO/IEC 27036-2:2014). International Organization for Standardization.
- International Organization for Standardization (2015a). *Application security: Part 2: Organization normative framework* (ISO/IEC 27034-2:2015). International Organization for Standardization.
- International Organization for Standardization (2015b). *Security techniques: Information security management for inter-sector and inter-organizational communications* (ISO/IEC 27010:2015). International Organization for Standardization.
- International Organization for Standardization (2015c). *Security techniques: Storage security* (ISO/IEC 27040 :2015). International Organization for Standardization.
- International Organization for Standardization (2016). *Security techniques: Information security for supplier relationships - Part 4: Guidelines for security of cloud services* (ISO/IEC 27036-4:2016). International Organization for Standardization.
- International Organization for Standardization (2018a). *Application security: Part 3: Application security management process* (ISO/IEC 27034-3:2018). International Organization for Standardization.
- International Organization for Standardization (2018b). *Data centre facilities and infrastructures: Part 1 - General concepts* (ISO/IEC TS 22237-1:2018). International Organization for Standardization.
- International Organization for Standardization (2018c). *Risk management: Guidelines* (ISO/IEC 31000:2018). International Organization for Standardization.
- International Organization for Standardization (April 2021). *Security techniques: Extension to ISO / IEC 27001 and ISO / IEC 27002 for privacy information management – Requirements and guidelines* (ISO / IEC 27701:2019) (ISO/IEC 27701:2021). International Organization for Standardization.
- ISO27k implementers. (2007). *ISO/IEC 27001 & 27002 implementation guidance and metrics*. ISO27k implementers' forum. [www.iso27001security.com/](http://www.iso27001security.com/)
- Joint Task Force (2020). *Security and Privacy Controls for Information Systems and Organizations* (NIST SP 800-53). National Institute of Standards and Technology.
- Joint Task Force Transformation Initiative (2011). *Managing Information Security Risk: Organization, Mission, and Information System View* (NIST SP 800-39). Gaithersburg, MD. National Institute of Standards and Technology.
- Joint Task Force Transformation Initiative (2013). *Security and Privacy Controls for Federal Information Systems and Organizations* (NIST SP 800-53). National Institute of Standards and Technology.
- Karoui, K. (2016). Security novel risk assessment framework based on reversible metrics: A case study of DDoS attacks on an E-commerce web server. *INTERNATIONAL JOURNAL of NETWORK MANAGEMENT*, 26(6), 553–578. <https://doi.org/10.1002/nem.1956>
- Kim, D., & Solomon, M. (2018). *Fundamentals of information systems security* (Third edition). Jones & Bartlett Learning.
- Kindervag, J., Shey, H., & Mak, K. (2015). *The Future Of Data Security And Privacy: Growth And Competitive Differentiation*. Cambridge, MA.
- Kitchenham, B. A., & Charters, S. (2007). *Guidelines for performing Systematic Literature Reviews in Software Engineering* (EBSE 2007-001). Keele University. [https://www.elsevier.com/\\_\\_data/promis\\_misc/525444systematicreviewsguide.pdf](https://www.elsevier.com/__data/promis_misc/525444systematicreviewsguide.pdf)
- Kobis, P. (2020). Information risk management in SME sector enterprises, 5, 79–83.
- Kumar, V. (2011). *ISO 27001 Compliance Checklist*.
- Landoll, D. (2021). *The security risk assessment handbook: A complete guide for performing security risk assessments*. CRC press.
- Lau, F., & Kuziemsky, C. (2016). *Handbook of eHealth Evaluation: An Evidence-based Approach*. <https://www.ncbi.nlm.nih.gov/books/NBK481590/>
- Longras, A., Pereira, T., Carneiro, P., & Pinto, P. (2013). *Information technology - Security techniques - Information security management systems - Requirements*. International Organization for Standardization.



- Lundgren, M. (2020). *Making the dead alive - dynamic routines in risk management // Making the Dead Alive - Dynamic Routines in Risk Management* [, Unpublished]. Zotero (via BibTeX-Export) // DataCite.
- Lundgren, M., & Bergström, E. (2019). Dynamic interplay in the information security risk management process. *International Journal of Risk Assessment and Management*, 22(2), 212–230. <https://doi.org/10.1504/ijram.2019.101287>
- Miller, A. (2017). *ISO 27001:2013 Annex A*. International Organization for Standardization.
- Millett, L. I., Fischhoff, B., & Weinberger, P. J. (2017). *Foundational cybersecurity research: Improving science, engineering, and institutions* / Lynette I. Millett, Baruch Fischhoff, Peter J. Weinberger, editors. The National Academies Press.
- Morse, J. M., Barrett, M., Mayan, M., Olson, K., & Spiers, J. (2002). Verification Strategies for Establishing Reliability and Validity in Qualitative Research. *International Journal of Qualitative Methods*, 1(2), 13–22. <https://doi.org/10.1177/160940690200100202>
- National Institute of Standards and Technology (2002). *Risk management guide for information technology systems: Recommendations of the National Institute of Standards and Technology (NIST SP 800-30)*. Gaithersburg, MD. National Institute of Standards and Technology.
- National Institute of Standards and Technology (2004). *FIPS 199, Standards for Security Categorization of Federal Information and Information Systems (FIPS PUB 199)*. National Institute of Standards and Technology.
- National Institute of Standards and Technology (October 2006). *Information Security Handbook: A Guide for Managers (NIST SP 800-100)*. Gaithersburg, MD. National Institute of Standards and Technology (NIST).
- National Institute of Standards and Technology (2008a). *Security Considerations in the System Development Life Cycle (NIST SP 800-64)*. National Institute of Standards and Technology.
- National Institute of Standards and Technology (2008b). *Volume I: Guide for Mapping Types of Information and Information Systems to Security Categories (NIST SP 800-60)*. National Institute of Standards and Technology. Organizations. <https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-53r4.pdf>
- National Institute of Standards and Technology (2008c). *Volume II: Appendices to Guide for Mapping Types of Information and Information Systems to Security Categories (NIST SP 800-60)*. Gaithersburg, MD. National Institute of Standards and Technology.
- National Institute of Standards and Technology. *The Cybersecurity Framework: Implementation Guidance for Federal Agencies (NISTIR 8170)*. National Institute of Standards and Technology. <https://www.nist.gov/cyberframework>
- National Institute of Standards and Technology (September 2011). *Information Security Continuous Monitoring (ISCM) for federal information systems and organizations (NIST SP 800-137)*. Gaithersburg, MD. National Institute of Standards and Technology.
- National Institute of Standards and Technology (2018). *Attribute metadata: A proposed schema for evaluating federated attributes (NISTIR 8112)*. Gaithersburg, MD. National Institute of Standards and Technology.
- National Institute of Standards and Technology. (2020). *NIST Privacy Framework: A tool for improving privacy through enterprise risk management*. Gaithersburg, MD. National Institute of Standards and Technology. <https://doi.org/10.6028/NIST.CSWP.01162020>
- Ogawa, R. T., & Malen, B. (1991). Towards Rigor in Reviews of Multivocal Literatures: Applying the Exploratory Case Study Method. *Review of Educational Research*, 61(3), 265–286. <https://doi.org/10.3102/00346543061003265>
- Okoli, C., & Schabram, K. (2010). A Guide to Conducting a Systematic Literature Review of Information Systems Research. *SSRN Electronic Journal*. Advance online publication. <https://doi.org/10.2139/ssrn.1954824>
- (2013). *Unified Profile for DoDAF and MODAF (UPDM)*. OMG. <http://www.omg.org/spec/UPDM/2.1>
- Paré, G., Trudel, M.-C., Jaana, M., & Kitsiou, S. (2015). Synthesizing information systems knowledge: A typology of literature reviews. *Information & Management*, 52(2), 183–199. <https://doi.org/10.1016/j.im.2014.08.008>
- Peltier, T. R. (2001). *Information security policies, procedures, and standards: Guidelines for effective information security management* / Thomas R. Peltier. Auerbach.
- Peón, D., Antelo, M., & Calvo-Silvosa, A. (2017). An inclusive taxonomy of behavioral biases. *European Journal of Government and Economics*, 6, 24–58.
- Pompon, R. (2016). *IT Security Risk Control Management: An Audit Preparation Plan*. Apress.

- Rangel, A. (2019). Why enterprises need to adopt 'need-to-know' security. *Computer Fraud & Security*, 2019(12), 9–12. [https://doi.org/10.1016/s1361-3723\(19\)30127-7](https://doi.org/10.1016/s1361-3723(19)30127-7)
- Reiner, D., Press, G., Lenaghan, M., Barta, D., Urmston, R., & IEEE COMPUTER SOCIETY (2004). Information lifecycle management: The EMC perspective. In
- Roessing, R., Benetis, V., Dimitriadis, C., & Stewart-Rattray, J. (2013). *Transforming Cybersecurity Using COBIT 5*.
- Rowley, J., & Slack, F. (2004). Conducting a literature review. *Management Research News*, 27. <https://doi.org/10.1108/01409170410784185>
- Rutishauser, S. (2012). *ISO - IEC 27033-2-2012 - Network security – Part 2 Guidelines for the design and implementation of network security*. International Organization for Standardization.
- Schiavone, S. L., Garg, L., & Summers, K. L. (2014). Ontology of information security in enterprises. In
- Scott, C. R., & Choi, S. (2017). Top secret from the bottom up. *Corporate Communications: An International Journal*, 22(4), 556–561. <https://doi.org/10.1108/ccij-08-2017-0072>
- SEBoK Editorial Board. (2022). *SEBoK - Guide to the Systems Engineering Body of Knowledge*. International Council on Systems Engineering. <https://sebokwiki.org>
- Seifert, J. W., & Relyea, H. C. (2004). Do you know where your information is in the homeland security era? *Gov. Inf. Q.*, 21, 399–405.
- Shedden, P., Ahmad, A., Smith, W., Tscherning, H., & Scheepers, R. (2016). Asset Identification in Information Security Risk Assessment: A Business Practice Approach. *Communications of the Association for Information Systems*, 39, 297–320. <https://doi.org/10.17705/1cais.03915>
- Shedden, P., Ruighaver, A. B., & Ahmad, A. (2006). Risk Management Standards - The Perception of Ease of Use. In
- Shedden, P., Smith, W., & Ahmad, A. (2010). Information Security Risk Assessment: Towards a Business Practice Perspective. In *AISM 2010*.
- Shi, X., Li, D., Zhu, H., & Zhang, W. (2007). Research on Supply Chain Information Classification Based on Information Value and Information Sensitivity. *2007 International Conference on Service Systems and Service Management*, 1–7.
- Shrestha, M., Johansen, C., & Noll, J. (2017). Security Classification for Smart Grid Infra structures.
- Siponen, M. (2003). Information security management standards: Problems and solutions. *PACIS 2003 Proceedings*, 105. <https://aisel.aisnet.org/pacis2003/105>
- Siponen, M., & Willison, R. (2009). Information security management standards: Problems and solutions. *Information & Management*, 46(5), 267–270. <https://doi.org/10.1016/j.im.2008.12.007>
- Stallings, W. (2018). *Effective Cybersecurity: A Guide to Using Best Practices and Standards*. Addison-Wesley. <https://books.google.de/books?id=nqXXtAEACAAJ>
- Susanto, H., & Almunawar, M. N. (2018). *Information security management systems: a novel framework and software as a tool for compliance with information security standard*. CRC press.
- Templier, M., & Paré, G. (2015). A Framework for Guiding and Evaluating Literature Reviews. *Commun. Assoc. Inf. Syst.*, 37, 6. <https://doi.org/10.17705/1CAIS.03706>
- Vargas, F. A., & Fenz, S. (2012). Mapping ISO 27002 into security ontology.
- Volchkov, A. (2018). *Information security governance: Framework and toolset for CISOs and decision makers*. CRC press.
- Watson, R. T., & Webster, J. (2020). Analysing the past to prepare for the future: Writing a literature review a roadmap for release 2.0. *Journal of Decision Systems*, 29(3), 129–147. <https://doi.org/10.1080/12460125.2020.1798591>
- Webster, J., & Watson, R. T. (2002). Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quarterly*, 26(2), xiii–xxiii. <http://www.jstor.org/stable/4132319>
- Weng, S. S., Tsai, H. J., Liu, S. C., & Hsu, C. H. (2006). Ontology construction for information classification. *Expert Systems with Applications*, 31(1), 1–12. <https://doi.org/10.1016/j.eswa.2005.09.007>
- Wheeler, E. (2011). *Security risk management: Building an information security risk management program from the Ground Up*. Elsevier.
- Whitman, M. E., & Mattford, H. J. (2019). *Management of information security* (Sixth edition). Cengage Learning.
- Whitman, M. E., & Mattord, H. J. (2018). *Principles of information security* (Sixth Edition). Cengage Learning.



- Williams, B. L. (2013). Information Security Policy Development for Compliance: ISO/IEC 27001, NIST SP 800-53, HIPAA Standard, PCI DSS V2.0, and AUP V5.0. In
- Wit, J. de, Pieters, W., Jansen, S., & van Gelder, P. (2021). Biases in security risk management: Do security professionals follow prospect theory in their decisions? *Journal Fo Integrated Security and Safety Science*, 1(1). <https://doi.org/10.18757/jisss.2021.1.5700>
- Yazid, S. A. I., Faizal, M. A., Ahmad, R., Shahrin, S., & bin Shamsuddin, S. (2012). Enhancement of Asset value classification for Mobile devices. *Proceedings Title: 2012 International Conference on Cyber Security, Cyber Warfare and Digital Forensic (CyberSec)*, 106–110. <https://doi.org/10.1109/CyberSec.2012.6246097>
- Yuan, J., Biennier, F., & Benharkat, N. (2021). Data centered and usage-based security service. In H. Hacid, F. Outay, H. Y. Paik, A. Alloum, M. Petrocchi, Bouadjeneq, A. Beheshti, X. Liu, & A. Maaradji (Eds.), *Lecture Notes in Computer Science, SERVICE-ORIENTED COMPUTING, ICSOC 2020* (pp. 457–471). SPRINGER INTERNATIONAL PUBLISHING AG. [https://doi.org/10.1007/978-3-030-76352-7\\_42](https://doi.org/10.1007/978-3-030-76352-7_42)