enisa 20 years!

EUROPEAN UNION AGENCY
FOR CYBERSECURITY

CYBER INSURANCE – MODELS AND METHODS AND THE USE OF AI

ENISA Research and Innovation Brief

FEBRUARY 2024

ABOUT ENISA

The European Union Agency for Cybersecurity, ENISA, is the Union's agency dedicated to achieving a high common level of cybersecurity across Europe. Established in 2004 and strengthened by the EU Cybersecurity Act, the European Union Agency for Cybersecurity contributes to EU cyber policy, enhances the trustworthiness of ICT products, services and processes with cybersecurity certification schemes, cooperates with Member States and EU bodies, and helps Europe prepare for the cyber challenges of tomorrow. Through knowledge sharing, capacity building and awareness raising, the Agency works together with its key stakeholders to strengthen trust in the connected economy, to boost resilience of the Union's infrastructure, and, ultimately, to keep Europe's society and citizens digitally secure. More information about ENISA and its work can be found here: www.enisa.europa.eu.

CONTACT

For contacting the authors, please use rit@enisa.europa.eu
For media enquiries about this paper, please use press@enisa.europa.eu

AUTHORS

Prof. Dr. Stefan Weber (Leibniz Universität Hannover), Prof. Dr. Matthias Scherer (Technical University of Munich), Corina Pascu (ENISA) and Marco Barros Lourenco (ENISA).

ACKNOWLEDGEMENTS

Marie Kratz (ESSEC Paris), Martin Eling (University of St.Gallen), Ana Teresa Moutinho (EIOPA), Miguel Caballero (EIOPA), Benedetta Di Lupidio (EIOPA), Gabriela Zeller (Technical University of Munich), Matthias Fahrenwaldt (BaFin).

LEGAL NOTICE

COPYRIGHT NOTICE

ISBN:978-92-9204-633-0, DOI: 10.2824/464473

TABLE OF CONTENTS

EXECUTIVE SUMMARY

Research and innovation (R&I) are important indicators for a society to measure progress, growth and development in any field. But progress and growth in our increasingly digital society cannot be achieved without trust. Investing in cybersecurity R&I is key to increasing knowledge about new and emerging threats and developing new technologies, tools and strategies to protect systems, networks and data. Failure to do so can have devastating consequences for building trust in the use of digital technologies by individuals, organisations and society as a whole.

The main objective of this report is to introduce cyber risk and cyber insurance, provide an overview of existing research and modelling approaches, and identify gaps for upcoming research projects. The key findings from this report are as follows:

- The current state of the cyber insurance industry is critically reviewed and the academic literature on cyber-risk modelling is summarized;
- We argue that among the most challenging parts of this interdisciplinary modelling task are (i) the solid understanding of the specific vulnerability of an individual firm on the one hand, and (ii) the interrelationships between firms on the other side, the latter resulting in systemic and systematic risks;
- We show that advanced statistical methods from ML/AI have the potential to be used in cyber-risk modelling and cyber insurance;
- A major obstacle to the further development and use of advanced statistical tools is the lack of publicly available data. We therefore advocate the creation of publicly available cyber-related data pools to foster research;
- Cyber losses exhibit statistical properties that have to be accounted for in modelling: in particular non-linear dependencies leading to accumulation risk in portfolios, non-stationary loss processes resulting from technological progress and human interaction, and heavy tailed loss distributions;
- We argue that cyber insurance, especially when combined with appropriate cyber assistance services, can enhance both the benefits to individual companies and the resilience of the global IT infrastructure;
- A long list of specific challenges and issues for further research is given.

In 2023, ENISA produced a report analysing the current perspectives and challenges of operators of essential services (OESs)1 in relation to the subscription of cyber insurance services. The report provides information and statistics on the selection, purchase and use of cyber insurance as a tool to mitigate cyber risks in daily business life. While this report provides an overview of the demand side and in particular the requirements of OESs for the use of cyber insurance, this present study highlights what is needed to address some of the challenges from a technical (actuarial) perspective to make cyber insurance more effective from the supply side. For example, how to increase the efficiency of cyber risk assessment and analysis to make cyber insurance more affordable and more suitable as a risk mitigation strategy for OESs. In addition, this current study also provides some practical recommendations on how to improve the maturity of risk management practises in terms of identifying, mitigating and quantifying risk exposure. The combined reading of these two reports will lead to a better understanding of how cyber insurance can be made more effective as a tool to mitigate cyber risks.

[1] ENISA (2023)

1. INTRODUCTION

This Chapter serves as a guide for reading this report. In a non-technical manner, we first introduce the problem of cyber risk from the perspective of the companies that are exposed to it and the insurance companies that are willing to (partially) accept it by writing appropriate cyber-insurance policies. We then explain the objectives and scope of the report, provide an overview of its structure, and explain the scientific methods and sources used to prepare it. Particular emphasis is placed on advanced statistical machine learning (ML) and artificial intelligence (AI) methods that are already being used in the context of cyber-risk analysis and mitigation or have the potential to provide such benefits in the future.

1.1 PROBLEM DESCRIPTION

Digital technologies already play a central role in value chains today, and their influence will certainly continue to increase. At the same time, however, the various risks associated with this development, summarized here under the term `cyber risk', are also increasing. It is now up to the company's risk management to become aware of these risks and their possible consequences and to take appropriate measures to reduce or transfer them. The latter, i.e., the transfer of cyber risks, is offered by several insurance and reinsurance companies under the term `cyber insurance'. If such insurance coverage is supplemented with IT services, known as cyber assistance, it is even possible to reduce part of the cyber risk and thus go beyond the mere transfer of risk. This is particularly relevant for small and medium-sized enterprises, which do not have the specific IT expertise that larger companies do. The role of cyber insurance in the cyber ecosystem and in risk management in general is crucial for most companies.

Leaving the corporate risk management perspective and looking at cyber from an insurance company's standpoint, there are some similarities, but also some striking differences. What is identical is the need to understand the company's risk; in the context of insurance, this is part of the underwriting process that ultimately leads to an insurance contract that is offered to the interested company in exchange for an annual insurance premium. However, in this risk assessment, the insurance company is primarily concerned with the potential financial consequences that a cyber incident may cause. In the language of an actuary, the insurance company aims to understand the probability of occurrence of cyber losses (`frequency of losses') and the distribution of their financial consequences (`severity distribution of losses'). Aspects of cyber risk that are not covered by an insurance contract, such as reputational damage, are not as relevant to an insurance company as they are to the affected firm. However, the insurance company must go far beyond the understanding of risk for individual companies. The reason for the necessity of a holistic portfolio model are the multiple interdependencies of cyber losses resulting, for example, from common attack vectors of cyber criminals or the interconnectedness of IT systems. This is of utmost relevance for an insurance company, as the resulting portfolio-loss distribution is strongly influenced by the dependencies between individual policies and this portfolio-loss distribution needs to be understood and managed in the risk management and regulatory capital calculation process of the insurance company. From an academic point of view, this offers very interesting and challenging opportunities for research which are linked to the systemic understanding of IT systems.

1.2 AIMS AND SCOPE

The main objective of our report is to identify key challenges in cyber risk and cyber insurance. To this end, we provide a brief introduction to the state of current research, but further refer for more in-depth information to recent survey articles such as Awiszus et al. (2022). We categorize cyber risks, explain products offered by the insurance industry, data, statistical methods, and procedures in ML/AI. The objective of the proposed research is to make the challenges in cyber accessible not only to a superficial qualitative treatment, but also to a quantitative analysis. In this context, a sound understanding of cyber risks and cyber insurance

requires a multidimensional perspective that brings together data, modelling, statistics, ML/AI, and interdisciplinary expertise, especially from actuaries and IT experts.

1.3 METHODOLOGY AND SOURCES

Our report on existing research and open challenges in cyber risk and insurance combines information from various sources, including:

- Academic literature (incl. unpublished manuscripts) on cyber risk.
- Academic literature on ML/AI.
- Reports by government agencies (ENISA, …), research institutes, actuarial associations, regulators, and private companies like (re-)insurance companies.
- Insights from discussions with practitioners that are active in the cyber-insurance market, IT experts, and researchers on ML/AI.
- Cyber-related databases.
- Own research expertise.
- Feedback by various experts on earlier versions of this document.

The existing scientific literature is a primary resource for the description of the current state of research, see (1). Many papers explicitly state the limitations of their studies and provide ideas for future research; this information is collected as well. Methods from ML/AI have already been applied successfully in more traditional lines of insurance, see (2). This strand of research is analyzed and extrapolated to cyber insurance. Very valuable information for new methods, products, new databases etc. are reports, see (3). Such reports often provide a very timely insight into cyber risk and the cyber-insurance market. Additionally, we interviewed experts, see (4), to explicitly learn about existing research opportunities and open practical problems. Moreover, the few existing databases on cyber-related data, see (5), give insight into possible future applications of ML/AI. A comparison of models and methods and existing data sets also provides indications as to which data are required for future developments and which databases should be built up, possibly supported by accompanying regulatory measures. Concerning (6), both authors of the report are senior researchers in actuarial science and financial mathematics with long-term research experience in stochastic modelling and statistical applications. Both have specifically worked on cyber-related questions and have collaborated with the insurance industry. This expertise and their personal opinions are also incorporated into the report. We also received further feedback from the ENISA panel discussions, see (7).

1.4 REPORT STRUCTURE

The report structure is as follows:

- **Chapter 2** contains a detailed introduction to the various facets of cyber risk on the one hand, and to cyber insurance and its requirements on the other. This not only provides the scientific background and guides the reader through a long list of references, but also illustrates the necessary applications and challenges from the perspective of the insurance industry. It also highlights the economic relevance of the problem. Very important for the whole document is a unified taxonomy for cyber risks, which is provided *en passant*. This section provides the academic background, illustrates the required applications and challenges from an insurance industry perspective, and finally provides the taxonomy for cyber risk that will be used in the following chapters.
- **Chapter 3** addresses the existing data on cyber, the academic literature analyzing that data, and an analysis of the consequences of cyber risk-specific stylized statistical facts that must be considered in modelling. Cyber-related datasets are briefly described and the scientific literature analyzing these data is cited and the main results are summarized. In detail, we discuss the consequences of cyber risks that need to be considered in modelling. In the

conclusion of this chapter, we make a strong case for creating better data sets to be available for research that are essential for further research; especially for applications of ML/AI.

- **Chapter 4** addresses modelling approaches for the cyber domain. We examine why a separation into individual, systemic, and systematic risks is important. In the area of modelling, many challenges remain for the future.
- **Chapter 5** briefly reviews the current state of machine learning (ML) and artificial intelligence (AI) and discusses existing (and possibly future) applications in actuarial science and the insurance industry. It is explained why until today not many of these methods have found their way into cyber insurance and what obstacles have to be removed to change this.
- **Finally, Chapter 6** contains a long list of important and promising research problems. These problems emerge from the discussion in the previous chapters and are organized as follows: Research question, objective(s), contributors, scope, deliverables, and opportunities for innovations.
- The report ends with a conclusion **(Chapter 7)**, a glossary of insurance-related technical terms **(Chapter 8)**, and a list of selected references (**Chapter 9**).

ENISA prepares these briefs with the aim of using them as a tool to develop advice on cybersecurity R&I and present it to stakeholders. These stakeholders are the main target audience of this report and include members of the wider R&I community (academics, researchers and innovators), industry, the European Commission (EC), the European Cyber Security Competence Centre (ECCC) and the National Coordination Centres (NCCs).

2. CYBER RISK AND CYBER INSURANCE

2.1 BACKGROUND

Digital technologies are increasingly determining our lives. This is not limited any more to desktop computers, tablets, or smartphones, as nowadays digital technologies are controlling many objects in everyday use.[2] This circumstance is often described by the keyword `Internet of things´. Likewise, our working environment is shifting further and further into the digital, networked space, with the recent Covid-19 pandemic strongly accelerating this change. Companies and their logistics, transport systems, energy supply, and sales channels are controlled by digital systems. This digital transformation brings massive efficiency gains and will continue to do so.

However, this added value from digital systems and the interconnectedness of digital networks is not only associated with progress, but also entails risks. Complex systems, on which the proper functioning of many processes in the daily lives of individuals, but also at the level of companies and entire societies, depends, can be disrupted. The interruption, delay, or disruption of processes can result in losses of very different dimensions. In the worst case, catastrophic events with massive financial consequences and even human casualties can be the result. This threat to digital systems is known as cyber risk, which must be analyzed, regulated, and managed.

Cyber risks can be examined from different perspectives, such as the causes of cyber events, the type of losses that occur, the approaches taken to assess risks, and the actions taken to enhance cyber security or mitigate the negative consequences of cyber events.

Causes of cyber risk include technical malfunction, human error, and insider or hacker attacks. Consequential damage may include loss or theft of data. Business operations may be disrupted or interrupted. Critical infrastructure may be limited in function, damaged, or destroyed. Even personal injury or death can result from cyber events. If the cyber events are active cyberattacks, then criminal offenses such as fraud and extortion play a prominent role.[3]

Analyzing cyber risks is challenging. Various tools are available to study cyber risks and threats in advance on a global scale. To assess potential impacts, a scenario analysis is often performed, examining individual counterfactual events and their consequences in case studies. Cyber risks can also be analyzed statistically, given that data is available. Finally, more complex, stochastic models can be developed that allow for a multi-layered examination of cyber risks.

Safeguards for cyber risks include system updates and additional security measures to ensure the integrity of systems. To be prepared for potential cyber events, the development of contingency plans is an important preventative measure. Finally, the monetary consequences of cyber damage can be covered by insurance solutions.[4]

Cyber risks have become entrenched in the public's consciousness. In recent years, we did witness global cyber events such as WannaCry and NotPetya in 2017, causing substantial disruption. The Allianz Risk Barometer 2022 (Allianz, 2022), a survey of business stakeholders,

[2] While the number of networked digital devices was estimated at 30 billion around 2020, 125 billion such devices are expected by 2030, see Reinhart (2021).
[3] Reinhart (2021) provides a list of examples in recent years.
[4] For excellent introductions and surveys, we refer to Zeller and Scherer (2022) and Awiszus et al. (2022).

ranks cyber risks first among global business risks for 2022 (cited by 44% of the respondents), ahead of business disruption (42%), natural disasters (25%), pandemics (22%), and legal and political risks (19%). Ranking 6th through 10th behind them are climate change, fire and explosions, market uncertainty, a shortage of skilled labor, and macroeconomic developments. According to the Center for Strategic & International Studies (CSIS, 2020), estimated annual damage caused by cyber risks worldwide increased with USD 445 billion in 2014, to USD 600 billion in 2018, and to USD 1000 billion in 2020. Depending on the definition and methodology, there are diverging estimates. In some cases, amounts six times higher are given, with up to USD 10500 billion predicted for 2025.

A key player in cyber-risk management is the insurance industry. MunichRe (Reinhart, 2021) estimates global insurance premiums at USD 5 billion in 2018, with a projected increase to USD 20 billion in 2025, with 50% of the market share in the USA and 25% in Europe. Especially in the area of criminal cyber-attacks, MunichRe expects a strong increase in the medium to long term: attack technologies are being further developed in the area of organized crime, among others, but also by states; simple-to-use attack tools are becoming easily available and can be used without elaborate IT-expertise. Technologies such as 5G, artificial intelligence, automation, cloud technology, etc. are expected to be targets of attackers in the future. The importance of cyber risks is also emphasized by Dr. Kerstin Awiszus from Group Risk Management of Hannover Re and the House of Insurance Hannover, stating: "*The interesting question is not if you will be hit by a cyber-attack, but when (you notice it)*."

What roles can insurance solutions play in the management of cyber risks? Classic risk management strategies for individuals or firms are the avoidance of risks, the reduction of risks, the conscious acceptance of risks, and the transfer of risks. These aspects all play a role in cyber risk as well. Actors must define their exposure, but also bear the cost of reducing risk through higher security standards. Risk transfer typically rounds out risk management by providing monetary protection for residual risks.

Insurance companies primarily offer risk-transfer solutions, but also demand from their customers to meet minimum security requirements at the inception of the contract and during its term. Information can be obtained through linked add-on offerings, and minimum requirements can be monitored through them. In the case of cyber insurance, cyber-assistance services are very important in this respect. Insurance companies can use these to broaden their business model, but above all they offer the possibility of monitoring and controlling risks in greater detail and on an ongoing basis. Especially in comparatively complex and thus non-transparent insurance lines such as cyber insurance, this additional information is very useful for pricing and risk management. Services in the area of risk transfer, but especially also services in the area of cyber assistance, have an important stabilizing function by making the insured companies more robust and resilient and can thus have a welfare-enhancing effect. In the cyber domain in particular, insurance companies take on tasks in which their social benefit transcends their traditional role as providers of pure risk-transfer solutions. This is in particular true if insurance solutions are combined with cyber assistance that enhances the physical security of IT systems.

In the field of insurance, cyber is a generic term for all risks in the context of computer systems, hardware, software, data, the Internet or other digital networks, any kind of Information Technology (IT), or Operational Technology (OT). The following insurance coverages are now offered or under development: Loss or theft of data, privacy breach protection, cyber extortion, property damage, (contingent) business interruption, product liability, reputational damage, and loss of intellectual property.[5] Players such as MunichRe (Reinhart, 2021) assume that cyber insurance has so far covered only a small fraction of cyber losses worldwide and that the

[5] Note that these risks are not always easy to separate and their monetary impact is often difficult to measure.

market for cyber insurance has substantial growth opportunities. However, with a simultaneous significant increase in cyber events, this insurance gap might likely persist for quite some time.[6]

2.2 TASKS OF THE CYBER-INSURANCE INDUSTRY

Insurance companies are confronted with complex tasks when assessing and managing risks, which must also be adequately implemented in the area of cyber insurance. These tasks relate to pricing, insurance-portfolio management, reserving, reinsurance, and preventing future losses. Models of cyber risks need to be further developed and data on cyber events and cyber losses need to be collected in order to develop appropriate methods for cyber insurance. In this section, we provide an overview of the various roles of insurance as a basis for the research questions to be developed in this report.

2.2.1 Pricing

The basic principle of insurance is the pooling of risks in the collective, the underlying principle is often explained as `the contribution of the many to the misfortune of the few'. Ideally, homogeneous collectives of independent risks can be grouped together, allowing for simple pricing rules using classical procedures and premium principles. The net premium, which fundamentally relies on the law of large numbers, is adjusted by a risk premium and a surcharge for costs and taxes. Classical pricing of insurance contracts focuses on idiosyncratic risks. In addition, other types of risks are relevant in the field of cyber insurance and must be taken into account, namely systematic and systemic risks.

Pricing based on risk pooling in the collective relies on the assumption of homogeneous collectives. For this purpose, relevant covariates have to be determined that allow for a meaningful risk discrimination. Then, for example, collective models can be used to describe aggregate loss distributions, which allow pricing of the contracts. Collective models interpret the collective as a producer of independent losses with equal distributions. Loss frequency and severity must be estimated from data.

Classical insurance pricing has to be complemented by other methods if systematic and systemic risks are present. Systematic risks arise from dependencies on random factor processes that jointly influence various insured parties. These require techniques from financial mathematics, methods of Hans Bühlmann's `actuary of the third kind', in their evaluation. Still unresolved in research is the question of how systemic risks, which play a central role in cyber risk, must be valued. Systemic risks are characterized by feedback effects as well as local and global interaction, i.e., systemic mechanisms that can contribute to amplifying risks. To price these, spread mechanisms must be modelled in an adequate way. Compared to other areas in insurance/finance, where dependencies often are the results of indirect effects (e.g. rising interest rates influences the creditworthiness of many companies), in cyber risks we are typically confronted with systemic events that directly cause losses at different entities7 and as such produce dependencies.

2.2.2 Portfolio-risk management / regulatory capital

Insurance companies bring together different insurance collectives and their risks. Complex dependencies between different components, resulting from systematic and systemic risks, have to be taken into account. Modelling and estimating these has been a largely unsolved task for cyber risks. In addition, risks need to be measured at the portfolio level, e.g., by monetary-risk measures, in order to estimate and manage balance sheets of insurance companies. Portfolio models are also the basis for insurance capital regulation which aims at protecting policyholders and other stakeholders against the possible insolvency of insurance companies.

2.2.3 Reserving

[6] According to the expert opinion of Reinhart (2021).
[7] Consider, e.g., a malware or cloud-outage affecting many firms simultaneously.

Claims reserving refers to the process of setting aside financial resources on the liabilities side of the balance sheet for incompletely settled claims. Claims reserving uses a canon of methods that is fed with data to determine the provisions. In the area of cyber risks, the availability of data is still low. On the one hand, this means that classical methods can only be applied to a very limited extent. On the other hand, the validity of the methods cannot yet be assessed conclusively.

2.2.4 Reinsurance

Comprehensive risk pooling can be ensured by means of collectives that are as large as possible. Pooling of portfolios of many insurance companies can be achieved by reinsurance companies that bundle large amounts of data and capital and maintain technical expertise. In addition to this risk-transfer to third parties, risk can also be transferred to capital markets via insurance-linked securities. The analysis of cyber-risk reinsurance products and derivatives still needs to be thoroughly developed.

2.2.5 Prevention of future losses

Information asymmetries and systemic risks make prudent underwriting policies necessary, especially in the area of cyber risks, which should include limitations of liability and exclusions. At the same time, the creation of transparency with respect to risks is an important component, which can also be established, for example, through services in cyber assistance. Insurance can thus also contribute to reducing cyber risks in the physical world. Good benchmarks for sound underwriting policies and effective strategies for cyber-assistance offerings must be comprehensively established.

In the case of cyber crises, we also have to ask very fundamentally what role state institutions can play in establishing systemic security and strengthening resilience. This is analogous to lessons learned from past financial crises, from supply-chain problems during the Corona pandemic and in the context of Russia's attack on Ukraine, and from experiencing shortages in agricultural and energy supply markets due to poor diversification and systemic interdependencies. Historically, it is evident that markets - that function properly in normal times - contribute insufficiently to the prosperity and supply of populations in times of crisis. To date, many government interventions and subsidies were initiated during crises, but the question must be asked as to how resilient structures can be created in advance and what government regulations are necessary to accomplish this. Such questions must also be answered for cyber risks and cyber insurance.

2.3 CHALLENGES AND IMPLICATIONS

2.3.1 Challenges

The risk management and insurance of cyber risks is associated with fundamental challenges. Digital technologies are constantly changing, and risks change in tandem with these developments. This progress, however, is not predictable, but intrinsically random and uncertain. Moreover, unlike risks in natural disasters, cyber risks are not expected to be stationary. Security technologies and the capabilities of cyber criminals are changing in a complex evolutionary dynamic that poses a fundamental challenge and requires constant adaptation. Seen from the statistical point of view, we have to consider working with non-stationary processes and data.

Another difficulty is the complex network structure, in which digital entities are interconnected and interact. Risks at different nodes are stochastically dependent, but do not have a simple regional/proximity structure like those risks in the field of natural disasters. Thus, cyber-risk assessment relies on a mapping of the graph structure of digital networks and the propagation dynamics of risks. Accumulation risks cannot be represented in a simple way. Calibration and validation of such models requires data of a different type than classical insurance loss risk-management models.

Equally complex is the fact that cyber risks can be of a very diverse nature. The term `digital technologies' is very broad and includes a wide range of entities and phenomena. This heterogeneity requires a pluralistic, broad modelling approach that must be adequately adapted to individual challenges.

As with many other insurance risks, information asymmetries are of relevance. Physical cybersecurity significantly influences the frequency and severity of claims. Proper risk selection remains important here as well, in order to be able to calculate appropriate premiums and not to jeopardize insurability. This is the well-known problem of adverse selection. In principle, the incentive for measures to increase cyber security might also be lowered after insurance contracts have been concluded. Whether the latter problem of moral hazard plays a central role in cyber risks is doubtful, since a variety of other incentives render an active lowering of cyber security less opportune.[8] Especially in the area of cyber risks, information asymmetries are associated with a particular complexity of risks. Insurance companies must consequently develop strategies to deal with this, such as coupling insurance with cyber assistance.

Finally, the availability of data on cyber risks is a central problem. Data is not yet available in suitable granularity or sufficient quantity for research, regulation, or application in the insurance industry. Insurance and reinsurance companies can assemble data pools, but government regulation is also needed to enable research for successful advancement in Europe on cyber security and cyber insurance backed by a solid database.

2.3.2 Implications for cyber insurance

Cyber risks are characterized above all by their evolutionary structure, their non-stationarity, and the importance of systemic network effects. These characteristics must be taken into account in modelling and risk management. There is still a great need for development here. In parallel, it is also important to build up datasets for research and regulation that are not yet available. In insurance practice, it is particularly important to couple risk transfer with physical protection and to continuously collect data in parallel. Insurance services must be constantly adapted to changing circumstances. Scott Sayce, Global Head of Cyber of Allianz Global Corporate & Specialty, characterizes the central maxim of cyber insurance in the following way (see Allianz, 2022):

"*Good cyber maturity and good cyber insurance go hand-in-hand. We buy insurance for our home, but this does not mean we leave the front door unlocked, and the same should be said for cyber security. The cyber market is shifting to a service-oriented offering that* combines *insurance policies with technology, risk engineering and response services. Through the underwriting process, and throughout the policy period, insurers can help organizations understand the continually changing exposures and focus their investment in cyber security and resilience. We want to be that partner throughout the cyber risk improvement journey.*"

[8] Let us mention the loss of reputation as an example, a risk, where no insurance compensation is paid for.

3. CONTEXT OF DATA

In this Chapter, we discuss existing databases on cyber risks and describe how this data can be used in the insurance context and in the modelling of cyber risks in general. We discuss the challenges of collecting and using cyber-related data and compare this with the situation in other segments of the insurance market.

3.1 CYBER-RELATED DATA CURRENTLY AVAILABLE AND USED

By mid-2022, the time of writing this report, there is insufficient publicly available data on cyber claims that provide good quality information. However, a few datasets exist that offer interesting initial insights related to cyber claims and could already be used in research. Sources of such data, either already used and analyzed in literature or the insurance industry, are summarized below.

3.1.1 Attacks on IT-systems

In the information technology literature, there are various scientific papers on the modelling of time series of attacks on IT-systems or websites (e.g. DDoS attacks). Such time series can be generated e.g. from observed attacks on real systems/websites or by setting up so-called honeypots[9]. Tools, such as time-series analysis and econometrics, can be applied to statistically analyze this data in order to gain insights into the dynamics of the attacks. With regard to cyber insurance, this yields interesting information, albeit indirect information. For cyber insurance, it is equally important to identify the likelihood of attacks being `successful', and to understand the statistical properties of these `successful' attacks and their financial consequences. Matching (the many) attacks to (the very few) cyber incidents and the resulting financial losses is difficult and represents an interesting (future) research area.

3.1.2 Data breaches

A data breach is one particular form of a cyber incident that might originate from different causes, e.g. hacking activities, human error, or technical malfunction. Depending on the local legislation and type of disclosed data, this can translate into massive fines and thus severe economic losses. Known data sources on data breaches are the `Chronology of Data Breaches'[10], organized by the nonprofit corporation `Privacy Rights Clearinghouse' (PRC) and the `Open security foundation data loss data-base'[11]. From an insurance perspective, it is not straightforward to translate information on data breaches into monetary losses[12], as e.g. reputational losses originating from data breaches are hard (if not impossible) to quantify. This quantification, however, is required if the financial consequences of a data breach are to be understood and evaluated in the context of pricing this risk. Concerning statistical results, the studies by Edwards et al. (2016) and Eling and Loperfido (2017) suggest that the number (resp. frequency) of disclosed records can be described by the log-normal law or a log-skew-normal distribution (resp. a negative binomial distribution). Eling and Jung (2018) study cross-sectional dependence (across industries) of data breach losses. What makes the use of these databases complicated is that most of the information is unstructured in the form of verbal descriptions of events. A deeper understanding of the statistical law of data breaches and their financial consequences (in different legislations) clearly has the potential for future research. It has to be acknowledged, however, that this is just one of many aspects of cyber risk.

[9] Specially generated web pages or web services with the purpose of triggering and recording attacks.
[10] https://privacyrights.org/data-breaches
[11] Formerly available from http://datalossdb.org
[12] A publication from the Ponemon Institute LLC (2016) concludes that the mean cost per disclosed record depends on the cause of data breach and the industry sector. A primer for a direct link between financial consequences and the magnitude of a data breach is Jacob's formula, see Jacob (2014), which maps the log-cost of a data breach to the log-number of disclosed records.

3.1.3 Cyber loss data (financial consequences)

An extensive database on cyber losses is provided by the commercial company Advison. As of 2022, they describe their dataset as `*Advisen's cyber loss data provides a historical view of more than 90,000 cyber events – including clash events – collected from reliable and publicly verifiable sources*'[13]. This data is analyzed, among others, by Romanosky (2016). The fact that this interesting set of data is not freely available to the academic world clearly hinders its analysis and investigation by the academic community.

Eling and Wirfs (2019) define cyber risk as a subgroup of operational risk and, with this perspective, analyze (the subset of) cyber losses from the operational risk database SAS OpRisk Global data[14]. Dacorogna and Kratz (2020) describe their research with a non-public database of the French *Gendarmerie Nationale*. Their initial analysis suggests heavy-tailed losses. At the same time, they describe the challenges of cleaning the data and potentially anonymizing it to make it publicly available. A detailed statistical analysis is provided in Dacorogna et al. (2022) suggesting consequences for risk management and a classification of cyber attacks based on the fatness of tails.

3.1.4 Meta-information on insured companies (idiosyncratic, systematic, and systemic)

Many of the databases mentioned above contain the names of the companies[15] involved in plain text. This may be valuable information when merging the database in concern with other information and ultimately obtaining an enriched list of covariates using meta-information about the companies in question. Although this clearly requires labor-intensive pre-processing of the data, it appears to be very worthwhile in applications such as individual risk assessment and constitutes an important task that should be further explored in upcoming studies.

Compared to academic research, the data situation is slightly better for insurance companies, as they have additional non-public information at their disposal, e.g. collected via risk questionnaires on the companies in their portfolio. They can also leverage the available history of losses and events they have faced in the past. However, conducting meaningful risk analysis during the underwriting process remains a very complicated task which ideally should result in the collection of whatever data is needed to model cyber losses and that should be conducted in close collaboration with IT-professionals. Supposedly simple questions such as: `What information/covariates about a company should I collect to analyze a company's cyber risk?' or `How can I quantify the risk of a particular company?' are still very difficult to answer.

At the latest, portfolio-risk management requires a further level of information. This involves the range of possible interrelationships between the companies under consideration that lead to stochastically dependent cyber losses. Many dependencies are hard to predict before an incident and might even be overlooked after an event. Examples include the common use of infected software, vulnerability to similar instances of social engineering, attacks via a supply chain, fishing attempts via mass email attacks, failure of shared cloud services or infrastructure, just to name a few. Compared to, for example, NatCat, where dependency between risks typically decreases with geographic distance (and exposure to common events such as flooding caused by the same river is easily understood), a similar `distance measure' for cyber is not evident and a strategy to diversify across countries/continents cannot be easily implemented in an interconnected cyber world. There exist already some approaches to modelling the nature of systematic or systemic incidents, often based on graph-theoretic methods (see the section on modelling in this document), but these are to date mostly at the level of toy models, and more empirical research is desirable. As the examples above show, the task of creating a realistic

[13] Quoted from https://www.advisenltd.com/data/cyber-loss-data/
[14] https://www.sas.com/content/dam/SAS/en_us/doc/productbrief/sas-oprisk-global-data-101187.pdf
[15] Note that as of today, cyber insurance is primarily offered to companies and organizations, the market to private end-customers being in its infancy.

model for networked cyber risk is not possible without adequate IT skills and comprehensive data, and may require constant updates as technology improves/changes.

3.2 STYLIZED FACTS AND CHALLENGES OF DATA ON CYBER LOSSES

In this paragraph, we discuss stylized statistical facts of data on cyber losses and related data. Note that this has some overlap with Chapter 4.

3.2.1 (Non-) availability of data

Cyber risk is a (relatively) new field and the relevant data are scarce. As mentioned earlier, each of the few existing databases we discuss in the following covers only a subset of the phenomena we are interested in, and most databases are not accessible to the public or the scientific world. Insurance companies often have some data on realized losses from the past, but the number of losses covered is very small compared to established mass insurance markets such as automobile insurance. The fact that many insurance companies have just entered the cyber market, and therefore have only been able to collect data from it when the first policies were written, illustrates that much of the information that will be needed for a sound statistical analysis tomorrow is not yet captured today. The use of advanced statistical methods from ML/AI could be substantially boosted by providing a better database (more covariates, more claims, ...).

3.2.2 Technological progress; non-stationarity of data

Technological progress since the invention of microcomputers and the internet, as well as the digitization of entire value chains, is a well-known fact. In recent decades, these changes have transformed the industry at an unprecedented pace and continue to do so. But every new technology is also a potential risk. At the IT-security level, we have to constantly adapt our systems (hardware, software, and most importantly, the users of those systems) to the latest changes. From the perspective of an actuary analyzing cyber risks, this implies that data collected in the past may not fully characterize the cyber risks of tomorrow. Intuitively speaking, risks arising from new technologies are not represented in historical data. Feedback loops of learning from risks must be taken into account. At the same time, the financial consequences of cyber risks have been increasing for many years.

The statistical term for this stylized fact is `non-stationarity'. It presents a key challenge in modelling cyber risk. Possible solutions include simple assumptions about the trend in frequency and severity, the adoption of models that explicitly account for changes in distribution over time, models with time-varying parameters, a combination of statistical data and expert opinion, and more general approaches incorporating model risk. Recognizing and studying the non-stationarity of cyber data constitutes an interesting area of research both for classic statistical modelling and for new methodologies from ML/AI. As identifying structural breaks in time series (using data alone) is a difficult task that requires long time series, it is reasonably expected that combining expert IT knowledge with data may lead to better results than a purely statistical analysis.

3.2.3 Accumulation of losses

The core principle of insurance is `diversification in the collective´. The ability to ensure a large number of risks and earn the corresponding premiums (which are slightly higher than the expected loss) is a business concept that can be formally justified by applying the classical law of large numbers and/or the central limit theorem. However, important actuarial conclusions are no longer valid when individual risks are heavy-tailed or the risks are stochastically dependent (or both). Stochastic dependence of risks is a substantial challenge we face in cyberspace. We have already made the technical case for this observation in an earlier section (e.g., common attacks, infrastructure, networks, ...). At the data level, attributing multiple cyber risks to a single cause is not trivial, since in many cases we have limited information about a loss, and a cyber incident may be observed/reported with delay. From a statistical perspective, it is therefore plausible that accumulation risk is underestimated. Sound modelling of accumulation risk,

ideally combining statistical knowledge with IT expertise, is an interesting area of research. In addition, ML/AI could help in attributing/identifying a common root behind events in different companies.

For an insurance company, accumulation risks have a variety of consequences. They constitute a very important component when modelling portfolio risks, and they have a major impact on the regulatory capital required. Likewise, they affect the price of non-proportional reinsurance. Operationally, accumulation risks also pose a challenge, as IT service providers trying to respond immediately to cyber issues may be overloaded if many incidents happen at the same time. Finally, it is worth noting that many statistical methods (including those in ML/AI) implicitly rely on the assumption that the data used are independent. An imprudent application of such methods has to be avoided, and research towards ML/AI methods with highly correlated input data is needed. Furthermore, while the use of copula methods to model dependent claims is operational from a statistical perspective and promising, it will not be sufficient for providing bottom-up models that capture the cause of dependencies; hence, other models need to be developed.

3.2.4 Diversity of risks

Eling et al. (2016) categorize cyber risk according to its origins, consequences, and key characteristics. They suggest the definition: `*Any risk emerging from the use of information and communication technology (ICT) that compromises the confidentiality, availability, or integrity of data or services.*" [...] "*Cyber risk is either caused naturally or is manmade, where the latter can emerge from human failure, cyber criminality (e.g. extortion, fraud), cyberwar, and cyber terrorism.*'

This definition illustrates the various aspects of cyber risk and might be used as a skeleton for a classification within a matrix structure. Such an approach requires a sound understanding of the technical background (and sufficient information on a specific case) in order to assign any cyber incident to the correct category. This is particularly important in areas in which insurance coverage is (or will be) available. According to MunichRe (2021), insurance coverage is already offered in the following areas: Loss or theft of data, privacy breach protection, cyber extortion, property damage, (contingent) business interruption, product liability, reputational damage, and loss of intellectual property.

From a statistical point of view, we are confronted with conflicting goals in the context of classification. On the one hand, a large number of homogeneous subgroups would theoretically increase the precision of pricing for each group; on the other hand, in practice too many subgroups would result in too few data points for each group which prevents any meaningful statistical analyses. A potential and attractive perspective for an application of ML/AI is a categorization of losses into subgroups, such that a good tradeoff between these competing goals is achieved.

3.2.5 Information asymmetries

Information asymmetries play an important role in many areas of insurance. Cyber risks are no exception, and the reasons are easy to explain (but very difficult to overcome): A customer's level of cyber security strongly influences the frequency and severity of risks; accurately identifying a company's true risk profile is a very difficult and complex task. Insurance companies devote considerable effort to this as part of the underwriting process. In cyber insurance, information asymmetries can be mitigated through accompanying cyber assistance.

An assessment of risks in the future could be supported by AI. However, if no cyber assistance services are associated with insurance coverage, major challenges will remain due to information asymmetries, if an IT-system is to be audited and assessed `from the outside'. At the same time, it is in a customer's interest to be perceived as a `good risk' when sharing her data in order to lower the insurance premium. Another related issue is transparency: to avoid

reputational damage, companies may have an incentive not to report cyber incidents to the public, leading to obvious data censorship issues.

3.3 VISION: POOLING DATA FROM DIFFERENT SOURCES

From the above description of the information sources already available and stylized facts about their structure, it is clear that one of the biggest obstacles impeding ML/AI research and progress in the cyber risk and cyber insurance domain is found in the lack of available data. We strongly endorse, therefore, the vision of creating a public pool of data on cyber security and cyber incidents to be available to insurers, the scientific community, and other stakeholders such as regulators / cyber-security providers. Similar initiatives already exist in other areas (collections of OpRisk cases, recovery rates of defaulted loans, ...).

It is a stand-alone research task to design an appropriate framework for a useful cyber database, as this requires knowledge of the statistical, actuarial, and technical nature of the cyber domain. Funding a group to create and maintain a cyber database could be a very useful initiative to which different forms of access could be granted to cyber-security agencies / intelligence agencies / police / industry / researchers. Accompanying regulatory measures that compel the sharing of information for this purpose are equally necessary.

4. TYPES OF CYBER RISK / MODELLING APPROACHES

4.1 IDIOSYNCRATIC RISK, SYSTEMATIC RISK, AND SYSTEMIC RISK

Not only the lack of data and its non-stationarity leading to major challenges in the field of cyber insurance, but also the specific characteristics of the risks. Known models, originally designed for other risks, need to be adequately improved and adapted in order to be able to represent cyber risks. The following three risk categories can be distinguished in the context of cyber insurance:

Idiosyncratic risks refer to independent fluctuations/losses at the level of individual entities. In homogeneous pools, these are the classic risks in the field of insurance. Classic actuarial premium principles can be applied.

Systematic risks refer to the functional dependence of cyber risks on underlying random processes, such as the random development of security or attack technology or information flows. Comparable exposures often arise from using the same software, servers, or computer systems, or from belonging to the same industrial sector or region. Risks of this type have already been analysed in detail in the context of financial markets and are well understood. They involve pricing by means of replication techniques and risk-neutral valuation.

Systemic risks refer to local and global feedback and interaction effects that cannot be described solely on the basis of idiosyncratic fluctuations and through the influence of exogenous processes on individual entities. Besides mechanistic couplings, strategic interactions play an important role. Central to systemic risks is that even when external factors are modelled, an isolated analysis of individual cyber contracts is not possible; rather, a view of the entire system is imperative. Worm-type malware and supplier attacks are examples. The investigation of systemic risks has also been of high significance in the context of the financial crisis since 2007/08.

4.2 CLASSICAL ACTUARIAL APPROACHES

Pricing and risk management of cyber risks require sound actuarial models adapted to the specific application. In particular, the frequency of random cyber events and the resulting loss amounts must be well represented. A classical approach are frequency-severity models, which Zeller and Scherer (2022) adapt for cyber risks. A particular challenge arises from the lack of cyber data to calibrate the models and verify goodness-of-fit. If data were available, techniques for static analysis of the models could be applied, e.g., for frequency modelling using generalized additive models (as in Zeller and Scherer (2022)), maximum- or marginal likelihood, or Bayesian methods, and for severity modelling maximum-likelihood or peaks-over-threshold, see Maillart and Sornette (2010), Edwards et al. (2016), McNeil et al. (2015), Embrechts et al. (2013), de Zea Bermudez and Kotz (2010a), and de Zea Bermudez and Kotz (2010b).

In principle, frequency-severity models are only of limited suitability for modelling cyber risks. The rapid development of technologies and cyber threats lead to non-stationary evolutions. At the same time, dependencies between risks have to be integrated. Classical models capture idiosyncratic and systematic risks. However, systemic risks require novel ideas and concrete approaches that need to be devised in research.

4.3 CONTAGION MODELS

Systemic risks arise from the coupling and interaction of entities. These have a high relevance in the context of networked cyber systems, where risks can propagate. At the same time, the behaviour of actors, e.g., their investment in cyber security, determines the cyber risk of other actors. In this respect, mechanistic and strategic interaction can be distinguished, which will require different modelling approaches.

Mechanistic local and global interactions have, for example, been modelled in the fields of epidemics and financial markets. In the context of frequency-severity models, Cox processes that capture the influence of exogenous factors on the occurrence of cyber events can be used to represent systematic risks. However, for systemic phenomena such as the spread of worm-type malware, feedback loops are key drivers of the dynamics. These can be well described quantitatively, for example, by self-exciting processes such as Hawkes processes. In the context of cyber risks and cyber insurance, these have already been applied by Bessy-Roland et al. (2021) and Baldwin et al. (2017). Because of their structure, Hawkes processes can be easily integrated into frequency-severity models. In the context of financial data, these processes have already been used and also statistically calibrated, see for example Embrechts et al. (2011), Daley and Vere-Jones (2003), Giesecke (2008), Errais et al. (2010), and Ait-Sahalia et al. (2015).

Even though Hawkes processes are a suitable tool to capture feedback effects on a global level, they cannot represent interactions between entities in detail, e.g., in the case of spread of a computer virus, a Trojan, or ransomware. Epidemic network models, on the other hand, are capable of describing and analysing dynamic interactions and amplifications in networks. The increased complexity of the models naturally magnifies the challenges in their statistical analysis, but they can shed light on important mechanisms, at least qualitatively, in counterfactual case studies.

Interacting Markov chains are employed in Fahrenwaldt et al. (2018) to study the dynamic propagation of cyber risks and to evaluate insurance contracts in a bottom-up approach. In particular, it can be shown that network structure has an important impact on cyber risks. An alternative interesting top-down approach is proposed by Hillairet and Lopez (2021), in which interaction is not described at the local level, but only globally at the population level. Specifically, they use the original population-based SIR model of Kermack and McKendrick (1927), which describes deterministic dynamics of the total numbers of susceptible, infected, and recovered individuals within a global population of IT devices. Although such a pragmatic approach neglects multiple details, it substantially improves the manageability of the models. An application of an extended model can also be found in Hillairet et al. (2021).

4.4 STRATEGIC INTERACTION

In addition to contagion effects in networks, strategic interactions also constitute an important dimension of cyber risks. First, in the realm of cyber threats, attacks and defences against them are strategic games played by the actors involved. Second, investments in cyber security have externalities in that they also increase the security of other entities. Third, the actions of regulators, insurance companies, and insureds jointly determine the payoffs and utilities of all parties. Game theory provides the appropriate conceptual framework to study these effects. The literature on game-theoretic aspects of cyber risk and cyber insurance is discussed in more detail in the surveys Böhme and Schwartz (2010), Reik and Böhme (2018), and Marotta et al. (2017).

It can be stated that the existing game-theoretical models have been oversimplified so far and can therefore only be applied to real data to a very limited extent. Also, qualitative implications in the literature, e.g. with regard to the influence of cyber insurance on cyber security, are ambiguous, so that concrete recommendations cannot yet be validly derived.

In particular, game-theoretic models of cyber risk have so far been mostly static and only allow for very simple cyber networks with a highly limited range of interaction mechanisms. One objective for future research must therefore be to combine more complex mechanical interactions of entities in a dynamic setting with strategic interaction.

4.5 KEY MODELLING CHALLENGES AND PRICING TECHNIQUES

As already discussed, an important challenge is to collect adequate data in sufficient quality and granularity; but also the development and improvement of models is a key task for future research. In particular, models that suitably incorporate systemic risk and remain tractable at the same time need to be constructed. In addition, dynamic strategic interaction models that include realistic network models for contagion should be designed and investigated. We discuss open questions and visions for future research in Chapter 6.

Pricing cyber risks requires a unified approach that integrates idiosyncratic, systematic and systemic risks. At the same time, strategic interactions must be taken into account. Future research must further develop and merge approaches from actuarial and financial mathematics for this purpose. A first discussion of these aspects can be found in Awiszus et al. (2022), which is based on approaches in Föllmer and Schied (2002), Wüthrich et al. (2010), Knispel et al. (2011), and Föllmer and Schied (2016). A more detailed explanation of important research questions can be found in Chapter 6.

5. STATISTICAL METHODS, MACHINE LEARNING, AND AI

Stochastic modelling of risks and insurance claims, in conjunction with statistical analysis of available and collected data, are among the core tasks of actuaries. Key applications are the pricing of risks, the measurement of portfolio risks (including the calculation of regulatory capital), and the calculation of reserves. In the area of cyber risks, additional services such as cyber assistance are of core importance. All application areas are interrelated, but in some cases, they are modelled slightly heterogeneously because models are often pragmatically tailored to the specific application.

The choice of model/algorithm to be used for some tasks in the insurance industry is also determined by the constraints of the applicable regulatory framework. While internal processes can be improved by ML/AI without noticeable constraints, this is not the case when it comes to actuarial applications like pricing, reserving, and SCR computation. Currently, most regulators require the use of `interpretable models' and evidence that a `sufficiently long' data history[16] is used to estimate the model(s). As of today, in the European insurance regulation framework, applications within the context of Pillar one and Pillar two of Solvency II are almost exclusively based on 'classical' statistical methods. Detailed reasons for this prudent view on the use of ML/AI methods are provided in the consultation paper BaFin (2021).

In this paragraph, we briefly touch upon some issues that need to be considered when using ML/AI in actuarial applications. It is characteristic for AI methods, that they often (a priori) postulate fewer regularities and structure in the process of modelling compared to `classical' statistical models. Their hypothesis space is typically larger, and causality is replaced by the identification of dependencies discovered in the data.[17] Models frequently exhibit a black-box character and the extrapolation to events of little or no appearance in the training data is difficult to control. A seemingly good performance of models, where patterns of data are reproduced, can also be misleading if the mechanisms of the methodologies are not sufficiently comprehended. Such difficulties must be avoided, particularly in the area of quantitative risk management and insurance, i.e., specifically in such a sensitive area as cybersecurity. However, tools from ML/AI can be utilized to support the development of `classical' actuarial models such as GLMs, or to challenge them in a horse race of models. One overarching aim of Solvency 2 is to check whether or not insurance companies own sufficient funds to have a one-year default probability smaller than a `1-in-a-200-year event'. Such a probabilistic statement, however, requires a stochastic model and a probability space in which the statement is well defined. ML/AI models do not provide this in many cases. Another overarching aim, manifested in European law, is the avoidance of discrimination[18]. Translating this to the requirements within a pricing model, it is obviously much easier to 'prove' that certain variables (like sex) do not enter a 'classical' statistical model compared to demonstrating that a ML/AI does not infer/learn a bias from the (possibly biased) input data. Interpreting these issues as research opportunities, it is appropriate to call for more research towards AI/ML that is interpretable, provides explainable results (dependence vs. causality), is stable when extrapolating into the tails, and is shown to be fair/unbiased in (actuarial) applications.

[16] This requirement has also been a challenge in the field of cyber insurance so far, as high-quality data is not yet available in sufficient volume.
[17] This is a serious problem if spurious correlations are lurking in the data and causality is replaced by correlation.
[18] Explaining why unisex tariffs are offered, e.g., in life insurance and health insurance even if there is obvious statistical evidence that, e.g., the life expectancy and medical costs differ among the sexes.

5.1 STATUS QUO OF STOCHASTIC METHODS USED IN CYBER

To provide an overview, in this Chapter we explain a selection of important stochastic and statistical models already used in the field of cyber insurance. The description is intended to reflect the diversity of different models and to categorize various possible approaches, but by no means claims to be exhaustive.

Static models: For many insurance applications, the consideration of a fixed time horizon is sufficient and a more complex, dynamic perspective can be dispensed with for pragmatic reasons. The time interval in view often corresponds to the term of a contract, e.g., one year. For this purpose, e.g., the number of claims and their severity during this period is examined without monitoring and modelling when exactly claims have occurred. At its simplest, such a view leads to a model for the counting variable (the frequency model) and the loss distribution (the severity model). The classical models for the severity of losses are the Poisson, the binomial, the negative binomial, and the geometric laws. In particular, the Poisson law for the number of losses is often embedded in a GLM framework where the intensity parameter λ of the Poisson distribution is modelled via a link function applied to a linear model to account for inhomogeneous risks. There are several proposals in the literature for the severity distribution of cyber losses. Following extreme value theory, some authors propose to model the tail of the loss distribution (extreme cyber losses) as a generalized Pareto distribution. For the main part of the loss distribution, a truncated log-normal distribution could be used. In addition to models that aim to describe losses directly, there are also approaches to link the number of records lost in a data breach to the monetary consequences. These include the regression-based work of Jacobs (2014) and Farkas et al. (2021). Other relevant literature in this context is Edwards et al. (2016), who model data breaches; they use a log-normal distribution for the number of exposed records and a negative binomial distribution for the daily frequency. Eling and Loperdo (2017) propose a log normal distribution for severity. A recent paper to cope with heavy-tailed distributions is Dacorogna at al. (2022).

Models based on stochastic processes: In many cases, it is important to explicitly consider the time evolution of random quantities. When the time evolution of univariate or multivariate variables is of interest, e.g., in the case of cyber losses or attacks, stochastic processes are an adequate framework for doing so. Peng et al. (2018) model cyberattack data via a copula GARCH model; Peng et al. (2017) consider marked point processes to represent extreme cyberattack rates; the same class of stochastic models is also used in Zeller and Scherer (2022) in the context of frequency-severity models. If not only idiosyncratic or systematic risks are to be studied, feedback mechanisms or interaction effects are central when considering systemic risks. Self-exciting and mutually exciting point processes are applied in Baldwin et al. (2017) to represent contagion in cybersecurity attacks; more references on this topic are provided in Chapter 5.

Network models and infections spreading: Networked IT infrastructure, as well as the spread of malware/worms in such systems, motivate an application of network models often used in mathematical biology when describing the spread of disease within a population. Fahrenwaldt et al. (2018) consider a (Markovian) SIS process to model the infectious spread of vulnerabilities in the context of cyberattacks and to price cyber reinsurance. Xu and Hua (2019) use Markovian and non-Markovian epidemic propagation processes to model and evaluate cyber insurance.

Methods from supervised learning: Farkas et al. (2021) analyze cyber claims via regression trees. We expect more research in this direction in the future.

5.2 OVERVIEW ON ML AND AI METHODS

A complete review of existing ML/AI methods is beyond the scope of this report. For a more comprehensive survey and further details, we refer to monographs such as Trevor et al. (2009) or Shalev-Shwartz and Ben-David (2014). Here, we focus on a selection of important tools to discuss prospective development opportunities in research on how ML/AI can be used to analyze cyber risk. In our presentation, we use the categories[19]:

Supervised learning: Models and algorithms from this area refer to the analysis of functional relationships for which sufficiently large data sets are available in advance, including both the input variable (also called characteristics, covariates, or predictors) and the output of interest. Such a situation is common in actuarial science. An example is historical information about individual policies, where, for example, for each policy and each year, the output variable of concern (which we want to model) is the number of claims, and the input is information about the insured person or the insured company.

Unsupervised learning: This category concerns the analysis of raw data as an input on which no given structure has yet been imposed; rather, the goal is to determine a relevant structure in the data by applying an adequate algorithm. An example in the insurance industry would be the grouping of risks based on a given risk description (e.g. via k-means) or a principal component analysis applied to interest rate scenarios in the context of Solvency II.

Reinforcement learning: Algorithms from reinforced learning used in the insurance industry typically try to minimize some given loss function describing an economic problem/situation. Often, this is applied in the framework of a Markov decision process.

5.3 METHODS OF ML/AI USED IN THE INSURANCE INDUSTRY

In the following, we briefly comment on existing applications of ML/AI in the insurance industry. We distinguish the cases of non-actuarial and actuarial applications.

Applications of ML/AI in non-actuarial application:

- As of today, most insurers have successfully used ML/AI to efficiently **process text** and even natural language from, e.g., contracts, written and verbal correspondence with customers, emails, and other online communications.

- **Automated claims management**: ML/AI assisted claims management is implemented by most insurance companies, however, the level of automatization in the claims management process differs.

- **Fraud detection**: Identifying fraudulent behavior is relevant for insurance companies. Tools from ML, such as clustering algorithms and NN, can be used to identify the occurrence of similar claims (in some area) in suspicious amounts. Often, ML tools are flanked by statistical tests and rule-based identifiers. A concrete example is given in Óskarsdóttir et al. (2021).

- **Automated underwriting** of (mostly simple) insurance products: Some insurance companies have successfully implemented robo-advisors to sell simple products like travel insurance to end-customers.

- **Understanding consumers behavior**, e.g., modelling churn rates (via logistic regression, random trees, neural networks, …) and identifying factors that contribute to a higher likelihood of contract cancellations.

[19] This categorization can be supplemented and refined, but in our view is initially sufficient in the context of this report to elaborate on research perspectives.

- Other applications include the **automatic analysis of large text fragments** like medical reports.

Applications of ML/AI in actuarial application:

We now turn to actuarial applications. Before providing a list of concrete examples, let us make two general remarks:

- Many actuarial applications and models can be embedded into a regression framework. This is emphasized in Richman (2018), who provides a long list of examples, including the pricing of risks via GLMs (Poisson and Gamma regression for frequency and severity), the embedding of the chain-letter approach for reserving into a regression context, more advanced IBNR models for reserving, examples from the modelling of lifetimes, and nested stochastic simulation for Solvency II / Swiss Solvency Test, see Hejazi and Jackson (2017). This omnipresence of regression models is important to acknowledge, because ML techniques often take a regression situation as a starting point and can be used to ultimately enhance or replace it.

- Many methods from classical statistics used in actuarial applications could, from a technical point-of-view, easily be replaced by methods from ML (such as neural networks). The main obstacle that hinders this shift in technology is the concern of the regulators that black-box algorithms are not allowed in actuarial applications; e.g., to rule out incorrect risk assessments or to prevent illegal discrimination[20]. Note that even if classical statistical models are still preferred in pricing and risk assessment (appreciating their interpretability), new tools from ML/AI can be used to benchmark and challenge the existing approaches.

There are some survey papers on the use of ML/AI in actuarial applications, e.g. the one by Richman (2018) mentioned earlier. Examples for the use of ML/AI in the actuarial context include:

- **Risk discrimination** with clustering methods. Given an inhomogeneous population of risks, it is a crucial task for the actuary to identify homogeneous subpopulations and to correctly assess their riskiness. Using structured and unstructured data, various clustering methods are available for this task. Information could be textual data (e.g. claims handler notes), economic databases, or data from social media.

- **Modelling loss frequency and severity,** e.g., using neural networks, regression forests, etc. This task, which is central in the pricing of individual risks, can be executed with classical statistical methods (Regression, GLMs, static models, …) and tools from ML. A primer on ratemaking using these tools is provided in Dugas et al. (2003). More than one hundred academic papers on pricing and reserving are surveyed in Blier-Wong et al. (2020).

- **Feature engineering**: Creating new models with large explanatory power, or improving on existing models, requires the identification and exploration of relevant covariates / features.[21] This process is very time consuming if executed via expert judgment by the actuary; particularly if the number of possible features is large. Moreover, it is possible that features are included erroneously into the model because they appear plausible for the actuary within some application, while statistically they are irrelevant. There exist successful approaches, based on methods from explainable AI, to automatically identify useful features and even possible interactions of features.

[20] For instance, in Europe it is not allowed to use gender as a covariable in insurance pricing. The gender can, however, be predicted with high likelihood even from data that is not obviously linked to it, like telematics data. The issue of unintended discriminations by AI is known under the term “algorithmic fairness” in the ML/AI community.
[21] A classical introduction to this topic is provided in Bengio et al. (2013).

- **Predicting reserves**: Insurance companies have to build up reserves in their balance sheet for claims that have occurred but not been settled (completely), yet. The task of predicting outstanding payments is usually done by standard statistical tools; the so-called "chain-letter approach" is among the most famous ones. Wüthrich (2018) shows how ML-methods can be used for individual claims reserving as an alternative. More contributions in this area are surveyed in Blier-Wong et al. (2020).

- **Remote sensing** in insurance: An interesting application of tools from image recognition / image analysis is remote sensing, e.g. via satellite images. This application is best understood from an example: Consider losses in agriculture from e.g. adverse weather, flood, or drought. In this context, it is very time consuming and thus costly to measure the loss on each field by physical inspection. Analyzing satellite images via AI is an appealing alternative, see De Leeuw et al. (2014) for a survey.

We conclude that many of the classical tasks in the insurance industry can be supported (resp. improved) by the use of ML/AI. There already exists a lot of research and practical experience in areas such as third-party liability insurance. As of today, cyber insurance has not been specifically discussed, presumably as this area is relatively new and offers only short data histories. Nevertheless, we believe that ML/AI offers interesting use cases also for cyber insurance. In the cyber-insurance context, however, we have to pay special attention to:

- Data being difficult to access for academic researchers, and data sparsity in general.

- Cyber data's specific stylized facts (accumulation risk, non-linear dependencies, non-stationarity, ...) might require a different adoption of ML/AI methods.

- An interdisciplinary approach[22] is necessary, as actuarial issues, legal view, IT-knowledge, etc. have to be combined.

[22] This is also emphasized and realized in Dacorogna and Kratz (2022).

6. VISION FOR FUTURE RESEARCH

In this Chapter, we present a number of research problems in the broader area of cyber risk and cyber insurance that should be addressed in the near future. Some of these problems include aspects of statistical modelling, data science, and the use of AI. Others relate to the important challenge of developing advanced models that incorporate systemic risk and integrated pricing techniques.

The following set of research topics is based on a comprehensive analysis of existing research and available data, as well as our own expertise, taking into account the needs of various stakeholders in the cyber insurance market (insurance sellers and buyers, regulators, society, etc.). We have structured the individual research tasks as follows:

- The respective **research question/task/topic** is named with the heading of the respective subsection.
- The **objective(s)** of the research is specified in the subsequent paragraph; this description is the focus of the proposal in each case.
- It is then proposed **who** should be involved in this interdisciplinary research (**entities)**, e.g., actuaries, cybersecurity analysts, psychologists, statisticians, technologists, etc. In general, independent researchers from academia and regulators should primarily deal with these problems, cooperating with experts from industry.
- A **scope** of research is proposed, specifying concisely what outcomes might be sought.
- Then the type of **deliverables/outcomes** is identified, e.g., academic papers, algorithms, databases, or software.
- Finally, **opportunities for innovation** are specified in more detail.

Beneficiaries of all research activities and results will be insurance companies and their customers, regulators, and the society in general, since risks will be better understood, can be better managed, regulated, and mitigated. In many cases, methodologies might also be applied in other areas such as quantitative risk management for companies with complex supply chains.

RESEARCH TOPIC #1:

IMPROVING THE PROCESS OF CYBER RISK ASSESSMENT

Objectives: An up-to-now unsolved problem is the question of how the individual risk situation of companies exposed to cyber threats should be assessed in terms of cyber losses and their financial consequences, e.g., by insurance companies or regulatory authorities. As of today, it is standard practice to use a risk questionnaire when underwriting (cyber) insurance contracts. Such a questionnaire typically contains IT-related questions, e.g., about the frequency of backups or the sensitivity of stored data.[22] Very often, these questionnaires contain only a surprisingly small number of questions. One reason for this is competition in the insurance market and the unwillingness or inability of companies to answer many questions. The natural consequence is that the information collected is very limited, resulting in the requirement that only the most important questions be asked.

An important future research question is how processes for collecting cyber risk assessment information can be significantly improved. An answer to this question requires an interdisciplinary perspective that includes IT knowledge (What influences the riskiness of an IT system?), statistical experience (How can this be translated into probabilities?), an actuarial perspective (What covariates go into my statistical model?), insights from insurance markets (What number and types of questions are accepted by the market?), and psychology (Is there moral hazard in answering these questions?). The process of improving risk assessment can be supported by data science/AI in a variety of ways - from better statistical models to automatic detection of vulnerabilities in specific IT landscapes.

Entities: Actuaries, IT-experts, Risk managers, Statisticians

Scope: Improving today's practice of analyzing the riskiness of a specific company. This involves the (possibly automatic) analysis of IT-vulnerabilities, proposals for the mitigation of risks, and a mapping to probabilistic models to be used by actuaries and regulators

Expected outcomes and deliverables:

Research papers, stochastic models, software, criteria for the design of questionnaires, manual for risk analysis.

Opportunities for innovation:

- Use of technology (automation) to perform an internal or external risk assessment of the customer.
- Development of software to (re)assess the company's own risk profile.
- Automated detection of vulnerabilities in IT-systems.
- Training of employees to address identified uncertainties.
- Development of advanced statistical models in the cyber domain
- Understanding which questions are really important for risk analysis in the cyber segment.

22 Also, an important aspect is the fact that for many organizations, cyber risks are not yet considered as business risks, but mainly as IT/enterprise risk. The assessment is conducted from a technical/technological standpoint and not from a business perspective.

RESEARCH TOPIC #2:

IDENTIFYING RELEVANT CO-VARIABLES / IMPROVING INDIVIDUAL PRICING

Objectives: The task of identifying relevant covariates is related to the underwriting process that we started to discuss in Research Topic #1, but this topic goes deeper into statistical modelling. Actuaries need to develop a more profound understanding of which covariates are critical to assessing an organization's individual risk. To date, the few academic papers on this topic mostly use indirect information such as a company's industry segment or the company's size. Intuitively, we would prefer to include more direct information about, e.g., the IT-landscape and the education w.r.t. to cyber risk of the people working for the company in question.

From a statistical perspective, this task boils down to the key issues:

- the identification of relevant covariates (which can realistically be measured as part of the underwriting process), and
- mathematical models that translate these covariates into probabilities of occurrence and severity distributions, ultimately implying an insurance premium for the (heterogeneous) risks.

For the identification of relevant covariates, supervised learning methods such as neural networks and tools from the field of explainable AI / explainable ML can be used. The characterization of adequate models and the development of efficient methods and algorithms is an important research topic.

Entities: Actuaries, IT-experts, Risk managers, Statisticians

Scope: Relevant co-variables for the description of individual risks are to be identified and used in a cyber-insurance pricing model. Once sufficient data is available, backtests on the statistical significance of these co-variables are to be performed.

Expected outcomes and deliverables:

Research papers, stochastic models, software.

Opportunities for innovation:

- Combine IT-expertise with actuarial modelling.
- Use tools from explainable AI to identify relevant covariables.
- Create an innovative statistical model for cyber risk.
- Use advanced model-selection tools

RESEARCH TOPIC #3:

MODELLING AND ESTIMATING LOSS FREQUENCY AND SEVERITY

Objectives: Another important task is to develop frequency/severity models based on available data and any future data that may be collected, in conjunction with adapted statistical methods that can be used for pricing and risk management. In order to generate tariff classes that are as homogeneous as possible, it is necessary to adequately define cyber-risk modules that categorize companies and contracts. Frequency/severity models also admit degrees of freedom in the choice of underlying processes (inhomogeneous Poisson processes, Cox processes, Hawkes processes, etc.) and severity distributions. Procedures for model selection and statistical application must be developed, which may be automated using AI methods. The process of claims management should be aligned to the modelling and pricing view.[23]

Entities: Actuaries, IT-experts, Probabilists, Statisticians

Scope: Innovative mathematical models for loss frequency and severity are developed, including methods for their estimation to data that might partially be based on expert knowledge, back-testing strategies.

Expected outcomes and deliverables: Research papers, mathematical models, software.

Opportunities for innovation:

- Develop innovative stochastic models for cyber risk.
- Automatize variable selection using ML/AI, either using existing approaches or developing new ideas.
- Combine data with expert opinion in the model-estimation process.
- Cope with the non-stationarity of cyber risks and associated data

RESEARCH TOPIC #4:

MODELLING OF SYSTEMIC RISK IN NETWORK MODELS

Objectives: Cyber risks are characterized in particular by the fact that contagion effects may be major drivers in complex systems. These systemic risks must be adequately represented. Consequently, an important task for research continues to be the development of epidemic cyber-risk models on networks that properly describe these processes. In doing so, complexity at high granularity must be balanced with pragmatic simplifications to improve tractability. This might require the construction and analysis of top-down models that capture the interaction of entities on an aggregate level. At the same time, efficient numerical methods, e.g. Monte Carlo simulation methods, must be devised that allow the evaluation of large-scale models.

Entities: Actuaries, Mathematical biologists, IT-experts, Risk managers, Probabilists, Statisticians.

Scope: Realistic models for systemic risk in IT-networks are developed. The underlying high-dimensional numerical problems are solved. Statistical methodologies are devised.

Expected outcomes and deliverables: Research papers, stochastic network models, software.

Opportunities for innovation:

- Develop innovative contagion models.
- Devise innovative numerical and statistical methods.
- Models developed in the present context could also be used for other activities/industries/ sectors, e.g. risk management of companies with complex supply chains

RESEARCH TOPIC #5:
MODELLING DYNAMIC STRATEGIC INTERACTION

Objectives: The modelling of strategic interaction of actors in the field of cyber risks has so far been limited mainly to static models. This interaction of actors is also strongly relevant for the assessment of the role of cyber insurance in the cyber ecosystem. Simplifications that neglect the temporal dimension of processes are potentially inadequate, because contagion processes are dynamic by nature. Academic research should therefore develop dynamic game-theoretic cyber models that also account for contagion processes in complex networks.

Entities: Actuaries, IT-experts, Risk managers, Statisticians.

Scope: Models for the dynamic strategic interaction of actors in the field of cyber risks; identification of key factors that drive cyber risk.

Expected outcomes and deliverables: Research papers, sophisticated game-theoretic models.

Opportunities for innovation:

- Develop dynamic game-theoretic cyber models.
- Identified interactions might be considered as covariates in the creation of risk profiles and could thus be used in the underwriting process.

RESEARCH TOPIC #6:
UNDERSTANDING MULTILAYER NETWORKS

Objectives: Production processes, logistics, and financial markets all constitute networks that nowadays integrate digital technologies and are therefore exposed to cyber risks. Risks thus are inherent in multi-layered networks, with cyber networks being one key component. In order to model potential losses, comprehensive modelling of interconnected networks is accordingly desirable. Insights can be used both to develop insurance solutions and to assess regulatory policy issues. Multilayer networks pose an exciting challenge for future research.

Entities: Actuaries, IT-experts, Experts on ERM, Risk managers, Statisticians.

Scope: Models of multilayer networks, impact on risk, insurance and regulation.

Expected outcomes and deliverables: Research papers, stochastic models, simulation algorithms.

Opportunities for innovation:

- Comprehensive modelling of interconnected networks.
- Obtain real-world data to fit such networks; quantify the significance of components in the interaction of networks.
- Identify methods to make coupled networks more resilient.

RESEARCH TOPIC #7:
PRICING IDIOSYNCRATIC, SYSTEMATIC, AND SYSTEMIC RISK

Objectives: Actuarial science and financial mathematics provide techniques for pricing risks. These techniques are primarily based on pooling, replication, and the assessment of residual risks. In the context of cyber events, three risk categories (idiosyncratic, systematic, and systemic risks) are of particular relevance, and an essential research question concerns the design of a unified framework that incorporates all components in an integrated manner. In particular, the pricing of systemic risks has not been conclusively clarified.[24].
Finally, a methodology for pricing should be applied in insurance practice based on concrete models calibrated using available data. There, ad hoc methods have been applied in the field of cyber insurance so far.

Entities: Actuaries, Risk managers, Statisticians

Scope: In an actuarial context, the different categories of cyber risk (idiosyncratic, systematic, and systemic risk) are distinguished and modelled in a holistic way; an integrated pricing framework is developed.

Expected outcomes and deliverables: Research papers, pricing algorithms and software.

Opportunities for innovation:

- Develop a holistic model for the pricing of all aspects of cyber risk.
- Discuss the use of standardized methods versus individual models from a regulatory perspective in the context of the several relevant risk categories.

RESEARCH TOPIC #8:

DATA FOR SYSTEMIC CYBER RISK

Objectives: In the area of cyber risks and cyber insurances, data are not yet available to a sufficient extent. Of additional relevance in the face of systemic risks are data characterizing loss accumulations in digital networks. In the context of epidemic network models, connectivity information is of central importance. Relevant research questions are: (i) which type of data (and sources) are essential to characterize systemic risks, underlying graph structures, and spreading processes, and (ii) which statistical methods are available or need to be developed for model calibration depending on the amount of available data. At the same time, strategies need to be devised to build appropriate databases that are available for public research.

Entities: Actuaries, IT-experts, Risk managers, Statisticians.

Scope: Identify data relevant in the context of systemic risk, in particular for spread processes; develop suitable statistical techniques; devise strategies to develop databases

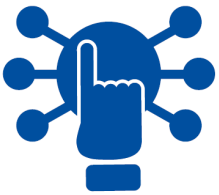
Expected outcomes and deliverables: Research papers, specification of envisioned database on cyber risk, and database.

Opportunities for innovation:

- Characterize data and develop statistical methods relevant for systemic risk.
- Estimate systemic cyber risk, in particular interaction processes and network structures, from real data.
- Pool cyber data from different sources providing a solid basis for statistical inference.
- Join forces between academic researchers in actuarial mathematics / cybersecurity and regulators.

RESEARCH TOPIC #9:

ADAPTING EXISTING ML METHODS TO THE SPECIFIC STYLIZED FACTS OF CYBER



Objectives: The appropriateness of any statistical model and ML/AI method depends substantially on underlying assumptions about the selected model framework; for example, it may be required that the input data are independent and identically distributed. In the cyber context, there are compelling reasons to believe that the input data are non-stationary (which violates identical distributed), dependent (which violates independent), and heavily tailed (i.e., moment conditions may be violated). Applying statistical methods without formal verification of required assumptions is possible (it often works to some extent from a purely practical perspective, i.e., some results are produced), but cannot be recommended, because flawed inferences with adverse consequences may occur. Particularly in view of the properties of cyber risks, it is thus necessary to conduct mathematical research by relaxing the assumptions of existent statistical methods, methods from data science, and tools from AI.

Entities: Actuaries, Mathematicians, Probabilists, Statisticians.

Scope: Analyze existing methods from ML/AI/classical statistics in terms of their underlying assumptions and the possibility of relaxing them. Develop suitable extensions for cyber applications.

Expected outcomes and deliverables: Research papers, algorithms, software.

Opportunities for innovation:

- ML techniques for non-stationary, dependent and heavy-tailed data.
- Conceptual innovations in AI.
- ML-techniques suitable for cyber risks.

RESEARCH TOPIC #10:

ESTIMATION OF MODELS FOR CYBER RISK (E.G. COMBINING STATISTICAL ESTIMATION AND EXPERT OPINION)

Objectives: Until now, only limited data on realized cyber losses (or related information) is available. However, there is a technical understanding of the underlying risks and an economic understanding of their financial consequences. An obvious approach to cope with limited data is consequently to combine empirical data with expert opinions to develop models and to draw conclusions. The lack of data might thus be mitigated by a structural understanding of the underlying problem. Combining such heterogeneous sources of information can be accomplished in a variety of ways and with a range of mathematical tools. An important research question is how to devise valid and robust strategies for such a methodology.

Entities: Actuaries, IT-experts, Risk managers, Statisticians.

Scope: Devise valid and robust estimation strategies for cyber risk, combining data and expert opinions in a mathematically sound manner.

Expected outcomes and deliverables: Research papers, algorithms, software.

Opportunities for innovation:

- Develop new statistical methodologies (e.g. in a Bayesian framework) to combine data and structural knowledge about a model.
- Investigate how estimates can consistently be updated when more data/further expert opinions become available.

RESEARCH TOPIC #11:

CYBER ASSISTANCE

Objectives: Cyber insurance often includes not only the pure transfer of risk, but is frequently combined with cyber assistance. These are services that actively improve cyber security[25] or include support in coping with damage in the event of cyber incidents.[26] Cyber assistance extends the business model of insurance companies[27], often engaging competent IT-experts as partners, but also allows for the reduction of information asymmetries and mitigates risks, and thus can also improve the insurability of cyber risks. Cyber assistance raises many interesting questions, all of which hold the potential for further research:

(a) What types of cyber assistance are demanded by the market?

(b) How can such services be designed to reduce individual risk?

(c) Can cyber services also reduce systemic/systemic risks?

(d) How should the cost of cyber services be shared between insurance companies and insured companies?

Entities: Actuaries, IT-experts, Product designers, Risk managers, Statisticians.

Scope: Obtain a holistic understanding of the potential of cyber assistance for cyber insurance, including product design, consequences on the distribution of risks, etc.

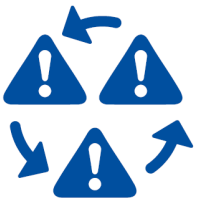
Expected outcomes and deliverables: Academic paper, draft for an insurance product.

Opportunities for innovation:

- Develop and price innovative assistance services.
- Combine risk transfer with risk mitigation.
- Reduce risks by implementing efficient IT-assistance.
- Stabilize the entire IT-infrastructure by implementing suitable warning systems.

RESEARCH TOPIC #12:

HEDGING ACCUMULATION RISKS



Objectives: From a risk manager's or actuary's perspective, at least two stylized statistical facts are of great importance in cyber insurance:

(a) Large individual losses (extreme cyber incidents);

(b) Events with many losses simultaneously (accumulation risk).

While category (a) could be reduced by appropriate cover limits in policies, the second category (b) remains a challenge that needs to be actively considered in pricing, policy design and risk management. It is both interesting and important to explore how accumulation risks can be meaningfully addressed through special reinsurance contracts, placement of insurance-linked securities in financial markets, or government intervention.

Entities: Actuaries, IT-experts, Risk managers, Statisticians.

Scope: Investigate how accumulation risks can be addressed through special reinsurance contracts, placement of insurance-linked securities in financial markets, or government intervention.

Expected outcomes and deliverables: Research papers, contract designs, guidelines for regulatory intervention.

Opportunities for innovation:

- ILS and government intervention are interesting options to explore not only for accumulation, but also for other situations.
- This topic is also an opportunity for policy researchers.

RESEARCH TOPIC #13:

CYBER RISK AS AN ASSET CLASS

Objectives: By analogy with NatCat, where so-called CAT bonds have been available for many years to transfer the risk of catastrophic natural events to the capital markets, it seems plausible to launch similar bonds with a coupon that would depend on cyber-related variables. How such insurance-linked securities can be designed, modelled, and priced, and whether they will be accepted by the markets, remains an open research question.

Entities: Actuaries, IT-experts, Risk managers, Statisticians.

Scope: Analyzing if and how cyber-related insurance-linked securities could become an asset class that allows to transfer cyber risks to financial markets.

Expected outcomes and deliverables: Research papers; blueprint for a cyber-related insurance-linked securities.

Opportunities for innovation:

- Involving researchers and experts from reinsurance, financial markets and portfolio optimization to construct innovative products.
- Tap broad resources to financially hedge cyber risks.

RESEARCH TOPIC #14:

CLOSING THE CYBER-INSURANCE GAP

Objectives: At present, a mismatch between supply and demand for cyber insurance is frequently reported, the so-called cyber insurance gap. Not only could it be interesting to look more closely at why insurance companies are unwilling to write more business in this segment, but it could also be beneficial for the industry to find innovative ways to close this gap. In the interplay between demand and supply, it is critical to understand how to balance customer needs (e.g., unlimited insurance coverage) with risk management considerations (e.g., reducing extreme portfolio losses).

Entities: Actuaries, IT-experts, Risk managers, Statisticians.

Scope: Analyzing the reasons for the present insurance gap in cyber and suggesting ways to overcome it

Expected outcomes and deliverables:
Research papers, strategies to increase cyber coverage

Opportunities for innovation:

- Understand the reasons for the cyber insurance gap (from an economic, behavioral, and risk management perspective).
- Better insight into customer needs; mapping these needs to the products offered by insurance companies.

RESEARCH TOPIC #15:

OPTIMAL CONTRACT DESIGN

Objectives: The design of an insurance contract requires multiple aspects to be balanced in detail:

(a) The needs of customers must be adequately addressed to guarantee sufficient insurance demand.

(b) From a contractual and legal perspective, it must be clearly specified which events are covered and which are excluded.

(c) A contract (as a formula applied to the raw claims) modifies the distribution of the claims portion covered by an insurance company. From both the customer's and the insurance company's perspective, the contract design must reflect the amount of coverage that will be provided in the event of a loss.

(d) Cyber assistance services are another component that must be included in addition to pure risk transfer.

(e) Ideally, contract design is consistent with the actuarial assessment of potential losses, and the claims department collects and records information about incurred losses in a consistent manner.

Balancing these requirements is typically not an easy task. Moreover, the complexity of the objective functions means that even defining a criterion for an `optimal contract' is a challenging task. Investigating this multidimensional problem more deeply in the context of cyber insurance remains an important research question.

Entities: Actuaries, IT-experts, Probabilists, Risk managers, Statisticians.

Scope: Understand the different dimensions of optimal cyber insurance contracts and utilize this knowledge for their optimal design.

Expected outcomes and deliverables: Research papers, guidelines specifying adequate and innovative cyber-insurance contracts.

Opportunities for innovation:

- Combine interdisciplinary insights on cyber exposures to provide good products.
- Utilize the opportunities offered by contract design to shape portfolio loss distributions in an optimal way.

RESEARCH TOPIC #16:

BEHAVIORAL CHALLENGES

Objectives: In the area of cyber risks, the behaviour of users of digital technology, e.g. computers, is also of major significance. Social engineering, for example, can trigger the disclosure of confidential information and provide hackers with access to otherwise protected systems. To assess insurance risks and potential losses, but equally in the context of cyber assistance services, it is hence necessary to collect data on the scope of risks, protective measures, losses incurred, and strategies for dealing with incidents, and to develop appropriate models. This research topic requires an interdisciplinary collaboration of experts from the fields of actuarial science, computer science, social sciences and psychology.

Entities: IT-experts, Psychologists, Risk managers

Scope: Collect data on the scope of cyber risks due to behavioral features, on protective measures, losses incurred. Investigate and devise incident management strategies and develop appropriate models.

Expected outcomes and deliverables: Research papers, developments of training tutorials.

Opportunities for innovation:

- Classification of different forms of social engineering in terms of risk, technology, social conventions, psychological aspects, etc.
- Understand which instruments are effective to reduce these vulnerabilities.

RESEARCH TOPIC #17:

CYBER INSURANCE FOR THE PRIVATE CUSTOMER SEGMENT

Objectives: The focus of cyber insurance market development is currently primarily on the coverage of companies and larger institutions. The background is that complex and non-stationary cyber risks must be monitored quite closely in the face of incomplete information when assessing exposures. Cyber assistance is an important pillar of cyber insurance that expands the insurance business model while simultaneously facilitating risk transfer in such an environment. Complex, individual services are difficult to establish for private customers; such solutions must be standardized and scalable. Another key challenge for private customers is that the probability of a cyber event occurring is significantly influenced by user behaviour. However, this is not directly observable. Another difference to companies is that cyber incidents in private households often lead to non-monetary losses (lost private data, damage to reputation, ...), which are not typically included in insurance coverage.

An important research question is therefore the development of cost-effective cyber insurance products for private customers. Due to information asymmetries, bonus-malus systems and a coupling with security software and standardized cyber assistance may be promising approaches. It should also be investigated to what extent such products in the private customer segment can increase cyber security in general and to what degree they can enhance welfare.

Entities: Actuaries, IT-experts, product designers.

Scope: Develop standardized cyber insurance solutions for the private customer segment.

Expected outcomes and deliverables: Research papers, proposals for product innovations

Opportunities for innovation:

- Standardized cyber insurance solutions in the private customer segment.
- Adequate bonus-malus-systems, coupling with security software and cyber assistance.
- Analysis of welfare implications.

RESEARCH TOPIC #18:

RESILIENCE OF SYSTEMS

Objectives: The extent of cyber losses can be shaped by the design of systems. Although cyber systems evolve dynamically and interactively without being centrally controlled, their developments can be influenced by regulatory interventions. Expertise from insurance providers that offer cyber insurance as risk transfer and cyber assistance as a linked service might be included when designing regulatory policies. Simultaneously, the terms of insurance contracts may provide incentives to agents leading to a systemic increase in cybersecurity. An important research question is how the insurance industry and regulators can work in tandem to increase the resilience[28] of digital networks. This is an important component to build a resilient Europe.

Entities: Actuaries, IT-experts, Risk managers, Statisticians.

Scope: Provide guidance on how the insurance industry and regulators can work in tandem to increase the resilience of digital networks.

Expected outcomes and deliverables: Research papers, policy recommendations.

Opportunities for innovation:

- Use interdisciplinary expertise from cyber insurers and cyber security providers to learn how resilient digital systems are structured.
- Transdisciplinary research can create efficient strategies to enhance the resilience of digital technologies.

RESEARCH TOPIC #19:

ROBUSTNESS OF MODELS

Objectives: Data and models enable an assessment of cyber risks and potential cyber losses and thereby provide useful information that contributes to the mitigation of risks and their hedging. However, a naïve view on models that attributes absolute reliability and correctness to them is not appropriate. Models may be miss-specified, incorrectly estimate probabilities, or simply ignore possible scenarios. The robust specification of models and the design of robust solutions is another important research question in the field of cyber risk and closely linked to the actuarial and financial literature on Knightian uncertainty.

Entities: Actuaries, IT-experts, Risk managers, Statisticians.

Scope: Investigate how to develop robust specifications of models in the cyber risk domain and how to characterize robust solutions.

Expected outcomes and deliverables: Research papers, recommendations for applications in practice.

Opportunities for innovation:

- Provide strategies that work sufficiently well in the context of Knightian uncertainty. This significantly improves security and risk management in practice.
- Academically, this opens a path to transfer modern results about Knightian uncertainty and optimal decisions to the field of cyber security and insurance.

RESEARCH TOPIC #20:

DATA COLLECTION

Objectives: As emphasized at various points in this report, the lack of sufficiently high-quality data is a major problem for scientific research, but also hampers risk transfer solutions in the area of insurance. Data protection, but also economic actor-level incentives, prevent the sharing of important information with researchers. Criminal groups, on the other hand, acquire access to data without complying with rules using it for their illicit purposes. This asymmetry can cause serious problems in the evolutionary development of cyber system security. That is why strategies ought to be developed in research that allow data access for scientists without compromising the security of cyber systems.

On the one hand, companies in the cybersecurity field will have to be enabled to provide data for research purposes in a standardized and anonymized[29] manner. On the other hand, since firms generally do not possess economic incentives to make information widely available for research purposes, regulatory requirements should be put in place to force the development of adequate databases for European researchers. This is indeed essential for the successful advancement of Europe, also from a global competition perspective, since on most other continents the flow and storage of data is less regulated and private sector tech groups are able to build and exploit databases. Data for research purposes, in particular, can promote the development of AI to enhance cybersecurity, rather than continuing to give a competitive advantage to criminal groups. The minimal goal should be to create a level playing field. Developing research databases and designing the necessary accompanying regulatory measures is an important research topic.

Entities: Actuaries, IT-experts, Risk managers, Statisticians.

Scope: Develop strategies that allow data access for scientists without compromising the security of cyber systems. This requires an adjusted legal framework and regulatory interventions.

Expected outcomes and deliverables:

Database, academic papers, improved cybersecurity.

Opportunities for innovation:

- Create a level playing field for Europe compared to other regions in the fight against cyber threats through an adequate database.
- Promote innovative developments in the field of cybersecurity by expanding transdisciplinary expertise in Europe.
- Ground, calibrate, and validate models based on actual data to achieve realistic assessments and successful strategies in cybersecurity and insurance.

RESEARCH TOPIC #21:

WELFARE AND REGULATORY IMPLICATIONS

Objectives: Increased cybersecurity and residual monetary protection through cyber insurance are not costless, however they are capable of increasing the resilience of technical processes and the economy as a whole. Resources have to be invested efficiently in cyber security; regulatory requirements must create incentives to implement appropriate countermeasures and to discourage harmful behavior. To establish adequate structures, it is necessary to understand the welfare implications of mechanisms and regulation. This task is very complex and requires good models of cybernetworks coupled with other systems as well as their impact on society. To achieve this, interdisciplinary research is needed, involving actuaries, mathematical economists, as well as IT experts.

Entities: Actuaries, IT-experts, Risk managers, Statisticians.

Scope: Investigate welfare implications of cybersecurity, cyber insurance and regulation; determine recommendations for efficient investments and regulatory measures.

Expected outcomes and deliverables: Research papers, recommendations for regulatory policies

Opportunities for innovation:

- Analysis of cybersecurity and insurance in the general context of society's welfare.
- Assessing risks and countermeasures in terms of their benefits to society.
- Grounding of measures and regulation.

RESEARCH TOPIC #22:

EXPLAINABLE AI FOR CYBER RISK

Objectives: In many core insurance applications, specifically in actuarial settings, machine learning tools are not permitted to be directly utilized by regulatory authorities, as they are generally not sufficiently explainable and, in some cases, opaque. For this reason, ML is mostly applied within insurance companies as a vehicle to streamline processes or to support/enhance classical actuarial modelling without actually replacing it. This raises several questions: How can ML methods in the context of cyber insurance be tailored to current regulatory requirements for insurance pricing and reserving, in particular to allow a critical review of results via sufficient explainability? How can ML methods be combined with classical actuarial methods to comply with regulatory requirements? What are the additional costs for further testing procedures in the context of explainable AI? Explainable AI, i.e., testing the validity of AI, introduces another model layer on top of AI: To what extent does this increase model uncertainty?

Entities: Actuaries, IT-experts, Risk managers, Statisticians.

Scope: Devise methods for explainable AI for cyber risk. Provide recipes for cost-benefit-analyses of AI.

Expected outcomes and deliverables:

Research papers, guidelines for model selection, algorithms.

Opportunities for innovation:

- Explainable AI for cyber risk.
- Substantial augmentation of competences in the area of artificial intelligence.

RESEARCH TOPIC #23:

VISION: AUTONOMOUS CYBER RISK MANAGEMENT

Objectives: In many areas, partial tasks can be automated. In the insurance industry, for example, this may in the future include automated management of claims and automated underwriting, as well as the use of ML for risk classification, pricing, and reserving. Applications of this type have yet to be explored in the area of cyber risk and cyber insurance. Artificial intelligence has so far constituted a successful instrument to automate and accelerate evaluations of large amounts of data. However, complex assessments requiring substantial cognitive capabilities are not yet within reach.

Therefore, a visionary task is to create a blueprint for a comprehensive automated risk management of cyber risks that autonomously implements analytics at all levels, proposes actions in a comprehensible way, and then decides on their execution. Such an approach could initially be implemented experimentally as a prototype in a simplified artificial environment created on the basis of simulated data. In further steps, scalability and robustness of this methodology may be tested. While the long-term end result would be to prospectively envision autonomous risk management in the real world, the medium-term focus is more on developing substantial expertise in Europe on the application of artificial intelligence in the intersection of cyber risks and risk management.

Entities: Actuaries, IT-experts, Risk managers, Statisticians.

Scope: Design autonomous cyber risk management in an artificial environment; develop substantial expertise on artificial intelligence in the intersection of cyber risks and risk management.

Expected outcomes and deliverables:

Research papers, simulation algorithms, artificial data, software.

Opportunities for innovation:

- Automation of risk management in the area of cyber risk.
- Substantial augmentation of competences in the area of artificial intelligence.
- Interdisciplinary synergies between the areas of risk management, IT and artificial intelligence

7. CONCLUSION

Cyber risks exhibit complex characteristics. They are non-stationary, evolving over time in interacting technical and social systems. Interaction and heavy tails are defining features of cyber risks. In this report, we identified key challenges for future research in the context of cyber risk and cyber insurance. Investigating the questions that have been compiled in detail in Chapter 6 requires interdisciplinary cooperation between experts, which should involve scientists from the fields of actuarial mathematics and IT-security, as well as regulators and players from companies.

Data: The collection of data and their availability must be significantly improved in the future. To date, only limited amounts of data are accessible for research, and their quality also has to be enhanced. We advocate government incentives and regulatory interventions to enable a database that can allow Europe to be competitive in cybersecurity. This is also an indispensable prerequisite when building a resilient Europe.

Models: Models for cyber risks, which are also the foundation for actuarial analysis and the viability of risk-management solutions, must be further developed. This concerns both pragmatic models that can be used as proxies in practice and models that capture the main classes of cyber risk: idiosyncratic, systematic, and systemic risks.

Statistical methods: At the same time, data must be analyzed, and statistical methods must be further developed or adapted. This includes methods from the ML/AI areas.

Insurance products and markets: Coupling cyber insurance with cyber assistance and optimal contract design are important topics, as are strategies to close the cyber-insurance gap. How to successfully design standardized cyber insurance for the private customer segment is another important question.

Societal and regulatory implications: Cyber insurance and linked products will have a positive impact on welfare. This needs to be explored in more detail. Governmental actors should select the guardrails for actors in a manner that strengthens both functionality and security of cyber networks and thereby establish overall resilient structures in Europe.

8. GLOSSARY ON INSURANCE TERMINOLOGY

Accumulation risk(s)	Accumulation of risks/losses caused by the same underlying factors or by coupling processes
Actuary	Expert in actuarial science, usually a mathematician working in the insurance industry or in academia
Adverse selection	Adverse selection refers to a decision behavior of actors in the face of asymmetric information that is not Pareto-optimal; in the insurance context, this term mostly refers to decisions made before the contract is concluded. In the case of poor risk selection, only bad risks are insurable as a consequence, while insurance premiums for good risks are too expensive
Aggregate losses	Losses at an aggregate level, e.g., at the level of an insurance portfolio or at the level of an insurance company
Calibration	Selection of a concrete model with specific model parameters, mostly based on the available data. In the context of financial markets, it usually means that the model parameters are specified in a way that model prices agree with market prices
Central limit theorem	Limit theorem describing that the sum of independently and identically distributed risks is approximately normally distributed; the theorem can also be derived under slightly weakened assumptions
Claim	Claim of a policyholder due to damage that is contractually covered
Collective models	Class of actuarial models that interpret an insurance collective as a loss producer; this allows a description of losses by means of independent identically distributed random variables, even in the case of non-homogeneous pools
Consequential damage	Loss that occurs as a result of an event; in the context of insurance contracts, it is of relevance which events and which losses are covered under a contract
Contingency plans	Contingency plans describe actions to be taken in advance for eventualities in order to be prepared and able to react in a targeted manner

Correlation	(Pearson's) Correlation is a measure of pairwise (linear) dependence among two random variables. It is the best-known dependence measure and is therefore often used colloquially as a synonym for stochastic dependence. The correlation of two random variables is a parameter that is obtained as the quotient of the covariance and the product of the standard deviations. In the context of multivariate Gaussian random variables, the correlation describes the dependence structure; in general, however, it is also a function of marginal distributions and cannot adequately characterize dependencies. Other concepts such as copulas are needed instead
Copula	A function containing the entire information about the dependence structure within a random vector; allowing for a separation of a random vector into the univariate marginal laws and the copula capturing the association among them
Covariates	Exogenous inputs on which outputs are functionally dependent are referred to as covariates
Dependencies	In stochastics, two random variables are independent if the distribution of one random variable conditionally on the other is identical to its unconditional distribution. The concept of independence can be generalized to families of sigma algebras. If there was no independence, one speaks of dependence. Dependencies in finite dimensions can be captured by a copula
Derivatives	Derivatives, also called contingent claims, are contracts that specify the exchange of resources in the future depending on future conditions whose occurrence is not yet known at the present time; often considered are products that define financial obligations and claims
Distribution	Distributions, also called (probability) laws, or probability measures, specify the probability of the occurrence of different possible events
Extreme value theory	A branch of mathematical statistics which is concerned with limit theorems for maxima/minima of sequences of random variables. Results from this area are very useful in applications e.g. when it comes to the statistical descriptions of the tails of a distribution, the modelling of rank statistics, or the modelling of exceedances over a high threshold level
Feedback effects	Phenomena due to feedback and amplification in systems
Game theory	Mathematical theory of strategic interaction of players whose joint behavior determines the evolution of a system; various concepts characterizing the behavior of players are studied, e.g., equilibria
GLM	GLM is an abbreviation for generalized linear models; model class generalizing linear regression

Heavy tails	Distributions that assign a high probability to large values; more precisely: distributions whose exponential moments are not finite
Idiosyncratic risk	Individual risks that are not driven by common underlying factors or systemic feedback effects and that are independent across all entities
Insolvency	Condition in which the fulfillment of obligations is no longer possible
Insurance coverage	Obligation to pay or entitlement to receive a compensation payment under contractually specified conditions, typically upon the occurrence of specific rare events
Insurance-linked securities	Financial market instruments whose cash flows are defined as a function of loss events or insurance payments
Interaction, local vs. global	In a dynamical system, the mutual interaction of entities, either locally or globally; in statistical mechanics, static equilibrium concepts also exist to formalize this
Knightian uncertainty	In contrast to risk, no concrete probability can be assigned to possible scenarios, but different probabilities are conceivable for each scenario; conception goes back to Frank Knight (1885-1972); also initial setting of every statistical model
Law of large numbers	Statement that the mean value of independent identically distributed random variables converges to the expected value; there are variants with different mathematical specifications and assumptions
Losses	Financial specification of damage in insured events
Loss frequency	Probabilistic description and modelling or measurement of the timing of loss events
Loss severity	Probabilistic description and modelling or measurement of the size of losses
Model risk	Risk that the model framework used or the concretely specified model is not appropriate
Monetary risk measures	Functionals that quantify risk on a monetary scale; examples are value at risk, average value at risk, utility-based shortfall risk, expectiles; can be applied to solvency capital requirements, performance measurement, limit systems

Monetary protection	Insurance coverage cannot prevent real damage, but can promise financial compensation payments and thereby provide monetary protection
Moral hazard	If information is incomplete, in the absence of monitoring after insurance coverage has been obtained, there may be a lack of incentive to physically guard against risks
Pooling of risks	Formation of a large collective of risks that occur only with a small probability; even with potentially large individual losses, a modest insurance premium per risk is sufficient to provide coverage
Portfolio management	Selection and adjustment of investment positions to achieve defined objectives such as profit maximization, risk minimization, utility maximization or compliance with regulatory requirements
Premium	Price for an insurance contract, consisting, among other components, of a risk premium and a cost contribution
Premium principles	Calculation methods for premiums that typically include a risk adjustment in addition to the expected value
Pricing	Methods for determining the value of insurance coverage, financial products, or other goods
Risk transfer (financial)	Financial protection against losses through insurance and/or financial contracts
Regulatory capital	Banks and insurance companies contractually promise future payments; in order to meet these obligations in the face of uncertainty, minimum resources must be maintained, whose size is determined by regulatory capital
Reinsurance	The insurance portfolios of primary insurers are largely determined by their distribution structures and sales markets; good pooling of risks can be achieved by combining portfolios from different insurers; this service is offered by reinsurers in return for adequate premium payments
Reserving	In order to be able to settle future claims, insurance companies must have adequate resources; to this end, reserves must be included on the liabilities side of the balance sheet; the process of reserving includes their proper actuarial calculation
Robustness	In the actuarial context, robustness refers to the development of procedures that, in the presence of model uncertainty, lead to solutions that are still sufficiently acceptable even in adverse cases

Resilience	Property of structures or entities to limit the extent of damage in the event of a crisis and to restore their functionality quickly
Scenario analysis	Analysis of processes and impacts for specific defined scenarios
Spread mechanisms	Mechanism of propagation in networked systems by local or global interaction processes, e.g., of diseases or computer viruses
Stationarity	Independence of the probabilistic properties of a stochastic process from the specific location of a time window under consideration; in other words, the essential properties of the probabilistic mechanism are unchanged over time
Statistics	In many real-life situations it is appropriate to describe phenomena in probabilistic terms: Probabilities are assigned to possible events. Yet, the true probabilities are not known. This uncertainty is represented by a statistical model, which specifies a priori a family of probabilistic mechanisms as a framework. Statistics addresses the inverse problem of systematically drawing inferences from data to probabilities. This is referred to as statistical inference. A pluralistic canon of methods exists.
Stochastic models	Uncertainty describes the lack of knowledge about the scenarios that actually occur. Models for such situations are developed in stochastics. More concretely, a model that assigns specific probabilities to scenarios is called a probabilistic model. Probability theory studies properties and develops methods. Statistics, on the other hand, focuses on the inverse problem of drawing inferences about the true probability measure from data. The collection of probabilistic and statistical models is encompassed by the term stochastic models.
Strategic interaction	If an outcome is influenced by the actions of many actors, each individual actor must account for, observe, and often make predictions about the possible behavior of other actors when pursuing his or her objectives. Such a situation is referred to as strategic interaction. Mathematically, this is studied in game theory.
Structural break	Sudden change in the parameters of a model that was previously capable of adequately describing the data; important especially in the context of certain families of models in statistics, to test a correct specification or to adjust a model framework
Surcharge for costs	Insurance companies incur costs, e.g. for their administration and infrastructure, which have to be charged in addition to the risk premium
Systematic risks	Probabilistic fluctuations driven by common exogenous factor processes; a classic example is ordinary financial market risks

Systemic risks	Risks based on feedback and interaction in systems; it is not feasible to assess individual positions in isolation without considering the system as a whole
Time series	a) Random data in discrete time b) Models for discrete-time stochastic processes, typically within restricted model families for which standard methods have been developed
Underwriting	An underwriter is an employee of an insurer, reinsurer, or broker who proposes insurance solutions to clients, reviews applications, assesses risks, and closes contracts. This process is called underwriting.
Validation	Verification on the basis of data whether a statistical model framework appears suitable; in contrast, calibration refers to the selection of models within a given model framework

9. BIBLIOGRAPHY

Aït-Sahalia, Y., Cacho-Diaz, J., and Laeven, R. J. (2015). Modelling financial contagion using mutually exciting jump processes. Journal of Financial Economics, 117(3), 585-606.

Allianz (2022). Allianz Risk Barometer. Tech. rep. Allianz Global Corporate & Specialty.

Awiszus, K., Knispel, T., Penner, I., Svindland, G., Voß A., and Weber, S. (2022). Modelling and Pricing Cyber Insurance – A Survey. Working Paper, Leibniz Universität Hannover.

Bundesanstalt für Finanzdienstleistungsaufsicht & Deutsche Bundesbank (2021). Machine learning in risk models - Characteristics and supervisory priorities: Consultation paper, available at https://www.bundesbank.de/en/homepage/machine-learning-in-risk-models-characteristics-and-supervisory-priorities-793670

Baldwin, A., Gheyas, I., Ioannidis, C., Pym, D., and Williams, J. (2017). Contagion in cyber security attacks. Journal of the Operational Research Society, 68(7):780-791.

Bengio, Y., Courville, A., and Vincent, P. (2013). Representation learning: A review and new perspectives. IEEE transactions on pattern analysis and machine intelligence, 35(8), 1798-1828.

Bessy-Roland, Y., Boumezoued, A., and Hillairet, C. (2021). Multivariate Hawkes process for cyber insurance. Annals of Actuarial Science, 15(1), 14-39.

Blier-Wong, C., Cossette, H., Lamontagne, L., and Marceau, E. (2020). Machine learning in P&C insurance: A review for pricing and reserving. Risks, 9(1), 4.

Böhme, R., and Schwartz, G. (2010). Modelling cyber-insurance: towards a unifying framework. In WEIS.

CSIS (2020). The Hidden Costs of Cybercrime. Tech. rep. Center for Strategic and International Studies (CSIS) in partnership with McAfee.

Dacorogna, M., Debbabi, N., and Kratz, M. (2022). Building up Cyber Resilience by Better Grasping Cyber Risk Via a New Algorithm for Modelling Heavy-Tailed Data. ESSEC Working Paper 2210. https://arxiv.org/abs/2209.02845

Dacorogna, M., and Kratz, M. (2020). Moving from Uncertainty to Risk: The Case of Cyber Risk. Chapter in "Cybersecurity in Humanities and Social Sciences. A Research Methods Approach". Edited by H. Loiseau, D.Ventre, H. Aden. WILEY - ISTE.

Dacorogna, M., and Kratz, M. (2022). Special Issue. “Cyber Risk & Security”. Risks 10: 112. https://doi.org/10.3390/ risks10060112.

Daley, D. J., and Vere-Jones, D. (2003). An introduction to the theory of point processes: volume I: elementary theory and methods. Springer New York.

Dugas, C., Bengio, Y., Chapados, N., Vincent, P., Denoncourt, G., and Fournier, C. (2003). Statistical learning algorithms applied to automobile insurance ratemaking. In CAS Forum (Vol. 1, No. 1, pp. 179-214). Arlington: Casualty Actuarial Society.

Edwards B., Hofmeyr S., and Forrest S. (2016). Hype and heavy tails: a closer look at data breaches. J Cybersecur 2(1):3–14.

Eling M., and Loperfido N. (2017). Data breaches: goodness of fit, pricing, and risk measurement. Insur Math Econ 75:126–136.

Eling M., and Jung K. (2018). Copula approaches for modelling cross-sectional dependence of data breach losses. Insur Math Econ 82:167–180.

Eling M., Schnell W., and Sommerrock F. (2016). Ten key questions on cyber risk and cyber risk insurance. The Geneva Association.

Eling M., and Wirfs J.H. (2019). What are the actual costs of cyber risk events? Eur J Oper Res 272(3):1109–1119.

Embrechts, P., Klüppelberg, C., and Mikosch, T. (2013). Modelling extremal events: for insurance and finance (Vol. 33). Springer Science & Business Media.

Embrechts, P., Liniger, T., and Lin, L. (2011). Multivariate Hawkes processes: an application to financial data. Journal of Applied Probability, 48(A), 367-378.

ENISA (2023). Demand Side of Cyber Insurance in the EU. ISBN 978-92-9204-586-9, DOI: 10.2824/94949, Catalogue nr TP-04-22-095-EN-N

Errais, E., Giesecke, K., and Goldberg, L. R. (2010). Affine point processes and portfolio credit risk. SIAM Journal on Financial Mathematics, 1(1), 642-665.

Fahrenwaldt, M., Weber, S., and Weske, K. (2018). Pricing of cyber insurance contracts in a network model. ASTIN Bulletin, 48(3):1175-1218.

Farkas, S., Lopez, O., and Thomas, M. (2021). Cyber claim analysis through Generalized Pareto Regression trees with applications to insurance pricing and reserving. Insurance: Mathematics and Economics 98: 92-105.

Föllmer, H., and Schied, A. (2002). Convex measures of risk and trading constraints. Finance and stochastics, 6(4), 429-447.

Föllmer, H., and Schied, A. (2016). Stochastic finance. de Gruyter.

Giesecke, K. (2008). Portfolio Credit Risk: Top-Down vs. Bottom-Up Approaches. In: Frontiers in Quantitative Finance: Volatility and Credit Risk Modelling. Ed. by R. Cont. Wiley. Chap. 10.

Hejazi, S. A., and Jackson, K. R. (2017). Efficient valuation of SCR via a neural network approach. Journal of Computational and Applied Mathematics, 313, 427-439.

Hillairet, C., and Lopez, O. (2021). Propagation of cyber incidents in an insurance portfolio: counting processes combined with compartmental epidemiological models. In: Scandinavian Actuarial Journal, pp. 1–24.

Hillairet, C., Lopez, O., d'Oultremont, L. and Spoorenberg, B. (2021). Cyber contagion: impact of the network structure on the losses of an insurance portfolio. Hal-03388840.

Jacobs J. (2014). Analyzing Ponemon cost of data breach.

Kermack, W. O., and McKendrick, A. G. (1927). A contribution to the mathematical theory of epidemics. Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character, 115(772), 700-721.

Knispel, T., Stahl, G., and Weber, S. (2011). From the equivalence principle to market consistent valuation. Jahresbericht der Deutschen Mathematiker-Vereinigung, 113(3), 139-172.

De Leeuw, J., Vrieling, A., Shee, A., Atzberger, C., Hadgu, K. M., Biradar, C. M., ... and Turvey, C. (2014). The potential and uptake of remote sensing in insurance: A review. Remote Sensing, 6(11), 10888-10912.

Maillart, T., and Sornette, D. (2010). Heavy-tailed distribution of cyber-risks. The European Physical Journal B, 75(3), 357-364.

Marotta, A., Martinelli, F., Nanni, S., Orlando, A. and Yautsiukhin, A. (2017). Cyber-insurance survey. In: Computer Science Review.

McNeil, A. J., Frey, R., and Embrechts, P. (2015). Quantitative risk management: concepts, techniques and tools-revised edition. Princeton University Press.

Óskarsdóttir, M., Ahmed, W., Antonio, K., Baesens, B., Dendievel, R., Donas, T., and Reynkens, T. (2021). Social network analytics for supervised fraud detection in insurance. Risk Analysis.

Peng, C., Xu, M., Xu, S., and Hu, T. (2018). Modelling multivariate cybersecurity risks. Journal of Applied Statistics, 45(15):2718-2740.

Reinhart, J. (2021). Cyber Insurance – Still infant or grown up? Talk, DGVFM Weiterbildungstag.

Richman, R. (2018). AI in actuarial science. Available at SSRN 3218082.

Riek, M., and Böhme, R. (2018). The costs of consumer-facing cybercrime: An empirical exploration of measurement issues and estimates. Journal of Cybersecurity, 4(1).

Romanosky S. (2016). Examining the costs and causes of cyber incidents. J Cybersecur 2(2):121–135.

Ponemon Institute LLC (2016). Cost of data breach study: global analysis.

Shalev-Shwartz, S., and Ben-David, S. (2014). Understanding machine learning: From theory to algorithms. Cambridge university press.

Trevor, H., Robert, T., and Jerome, F. (2009). The elements of statistical learning: data mining, inference, and prediction.

Wüthrich, M. V. (2018). Machine learning in individual claims reserving. Scandinavian Actuarial Journal, 2018(6), 465-480.

Wüthrich, M. V., Bühlmann, H., and Furrer, H. (2010). Market-consistent actuarial valuation (Vol. 2). Berlin: Springer.

Xu, M., and Hua, L. (2019). Cybersecurity insurance: Modelling and pricing. North American Actuarial Journal, 23(2):220-249.

de Zea Bermudez, P., and Kotz, S. (2010a). Parameter estimation of the generalized Pareto distribution—Part I. Journal of Statistical Planning and Inference, 140(6), 1353-1373.

de Zea Bermudez, P., and Kotz, S. (2010b). Parameter estimation of the generalized Pareto distribution—Part II. Journal of Statistical Planning and Inference, 140(6), 1374-1388.

Zeller, G., and M. Scherer (2022). A comprehensive model for cyber risk based on marked point processes and its application to insurance. European Actuarial Journal, 12, 33-85.

TP-04-23-347-EN-N

ABOUT ENISA

The European Union Agency for Cybersecurity, ENISA, is the Union's agency dedicated to achieving a high common level of cybersecurity across Europe. Established in 2004 and strengthened by the EU Cybersecurity Act, the European Union Agency for Cybersecurity contributes to EU cyber policy, enhances the trustworthiness of ICT products, services and processes with cybersecurity certification schemes, cooperates with Member States and EU bodies, and helps Europe prepare for the cyber challenges of tomorrow. Through knowledge sharing, capacity building and awareness raising, the Agency works together with its key stakeholders to strengthen trust in the connected economy, to boost resilience of the Union's infrastructure, and, ultimately, to keep Europe's society and citizens digitally secure. More information about ENISA and its work can be found here: www.enisa.europa.eu.

ENISA
European Union Agency for Cybersecurity

Athens Office
Agamemnonos 14
Chalandri 15231, Attiki, Greece

Heraklion Office
95 Nikolaou Plastira
700 13 Vassilika Vouton, Heraklion, Greece

Brussels Office
Rue de la Loi 107
1049 Brussels, Belgium

enisa.europa.eu

ISBN 978-92-9204-628-6
doi: 10.2824/773850