Enhancing Incident Management
Insights from a Case Study at ING

Kapel, Eileen; Cruz, Luís; Spinellis, Diomidis; Van Deursen, Arie

DOI
10.1145/3643665.3648048

Publication date
2024

Document Version
Final published version

Published in
FinanSE '24

Enhancing Incident Management: Insights from a Case Study at ING

Eileen Kapel
Eileen.Kapel@ing.com
ING Bank
Amsterdam, Noord-Holland, The Netherlands

Luís Cruz
l.cruz@tudelft.nl
Delft University of Technology
Delft, Zuid-Holland, The Netherlands

Diomidis Spinellis
d.spinellis@tudelft.nl
Delft University of Technology
Delft, Zuid-Holland, The Netherlands

Arie van Deursen
arie.vandeursen@tudelft.nl
Delft University of Technology
Delft, Zuid-Holland, The Netherlands

ABSTRACT

An incident management process is necessary in businesses that depend strongly on software and services. A proper process is essential to guarantee that incidents are well-handled, especially in a financial software-defined business needing to adhere to guidelines and regulations. This paper aims to enhance understanding of the current state of practice through a single-case exploratory case study, at the international bank ING, by interviewing 15 subject matter experts on the incident management process. The research identifies eight core observations on tool usage, the challenges experienced and future opportunities. Core challenges include monitoring data quality, the complexity of the environment, and the balance between minimising incident resolution time and following procedural guidelines. Future opportunities can lessen these challenges by making better use of available tooling and employing machine learning approaches. This requires tight supervision on the use of best practices and good monitoring data quality. The findings emphasise the need for a strengthened focus on improving the quality of monitoring data, handling environment complexity, incident clustering, and better support for regulatory compliance.

CCS CONCEPTS

• **General and reference → Empirical studies**; • **Software and its engineering → Software post-development issues**.

KEYWORDS

empirical, incident management, case study, interview

ACM Reference Format:
Eileen Kapel, Luís Cruz, Diomidis Spinellis, and Arie van Deursen. 2024. Enhancing Incident Management: Insights from a Case Study at ING. In *2024 Workshop on Software Engineering Challenges in Financial Firms (FinanSE '24), April 16, 2024, Lisbon, Portugal.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3643665.3648048

FinanSE '24 , April 16, 2024, Lisbon, Portugal

ACM ISBN 979-8-4007-0568-7/24/04…$15.00
https://doi.org/10.1145/3643665.3648048

1 INTRODUCTION

Major industries heavily rely on software and services, referred to as software-defined businesses [3]. While this fosters innovation and digital transformation, it also introduces the risk of service degradation through incidents, defined as unplanned interruptions of the availability of a service or reductions in the quality of a service [10]. Incidents may lead to customer dissatisfaction, financial losses and reputational damage, emphasising the need for prompt resolution to restore normal service operation and ensure business operations are minimally impacted. This is especially prudent for businesses needing to adhere to guidelines and regulations to ensure resiliency, transparency, and efficiency, as failure to comply could lead to sanctions.

Incident management, described as the process that aims to manage the life-cycle of all incidents, is the focal point of this study. Our objective is to enhance our understanding of the state of practice of incident management processes in financial software-defined companies, with the European universal bank ING (International Netherlands Group) serving as our case study. To grasp the current implementation of this process, we analyse tool usage, main challenges, and future opportunities. Employing a mixed method design, we conduct a single-case exploratory case study on real-world data from ING. Qualitative insights are gathered through interviews with 15 subject matter experts, which are then triangulated with documentation and quantitative historic incident data. This comprehensive approach allows us to evaluate the current state of ING's incident management process. Our study is guided by the following research questions (RQs):

- **RQ1:** How are tools currently used in the incident management process?
- **RQ2:** What are the main challenges when handling incidents?
- **RQ3:** What are future opportunities for improvements to the incident management process?

Our study reveals eight core observations. The incident management process relies on prescribed tools, complemented by flexible support for monitoring, dashboards, and documentation. Core challenges include handling false positives and duplicates, the complexity of the environment, and striking a balance between minimising incident resolution time and following procedural guidelines. Future opportunities lie in leveraging available tools and employing

machine learning approaches on incident and environmental data. Increasing the level of automation is possible, but necessitates tight supervision to maintain best practices and ensure good monitoring data quality. Our results help to identify future research directions and recommendations for financial software-defined businesses, including improving the quality of monitoring data, handling environment complexity, issue clustering, and better support for regulatory compliance.

2 RELATED LITERATURE

Recently, a substantial body of research has emerged on the application of software analytics and machine learning approaches in the incident management process. It is not clear to what extent these approaches fit the reality of software-based organisations. Our paper differentiates by examining their usage in practice.

One reoccurring topic is incident triage, crucial for assigning incidents to the right team promptly. Misassignments can cause delays and incur loss, since engineers invest a significant amount of time and effort in triaging an incident [5]. A possible explanation is the limited information provided in automatically created incidents causing multiple reassignments [5, 6, 8, 22]. Chen et al. explored applying bug-triage techniques for incident triage [5]. Furthermore, Chen et al. propose DeepCT, a deep learning based approach to automated continuous incident triage [6]. This model can incrementally learn from discussions and update incident triage results. Implementing this kind of an approach is far from trivial in practice and our work aims to create a better understanding of inherent challenges.

Furthermore, a common problem of an incident system is that relationships between incidents are not immediately identified [16]. This can lead to overwhelming ticket floods during impactful incidents which overwhelm responders and cloud the essential ones [8, 16]. Manual efforts are typically required to establish links between incidents. Automating this process enhances productivity by identifying redundant or potentially root-cause-related tickets. Jiang et al. suggest recording the mitigation process as a troubleshooting guide when an incident first occurs [15], facilitating quick resolution upon recurrence. They propose DeepRmd, an automated troubleshooting guide recommendation approach using deep learning to leverage textual similarity between incidents and guides. Chen et al. present their AIOps framework IcM BRAIN that correlates incidents with either an event-based or resource-guided method to relate incident tickets [8].

Another aspect is the prevalence of false tickets which wrongly report an incident. Güven et al. describe transient alerts that automatically disappear after a while, while their incident tickets remain in the system [13], e.g. an antivirus process that causes prolonged CPU spikes at regular intervals. Chen et al. propose an approach called DeepIP that prioritises the important tickets to be solved using an attention-based Convolutional Neural Network [7]. This aids engineers in addressing high-priority incidents first.

Changes in live systems are a common cause of incidents, responsible for up to 70% of the outages, according to Google [4]. Large-scale software companies have many services and resources that consist of many components [7], which obscure a complete view of the entire system and its relationships. These dependencies make it difficult to relate an incident back to a change, especially when often the seemingly successful changes lead to incidents [12, 25]. Contrary to assumptions, incidents are often caused by changes occurring hours or days before, as opposed to shortly before [13]. Zhao et al. propose the SCWarn approach that can identify bad changes using multi-modal anomaly detection on heterogeneous multi-source data [25].

Moreover, discovering the root cause of incidents is often complex. Over 52% of outages have different originating and causing services [23]. Numerous tickets stemming from one root cause can overwhelm engineers, exacerbated by a limited number of people handling incidents [7]. Also, if engineers have specialised knowledge about a specific part of a chain, it often leads to long mitigation times for incidents beyond their expertise, necessitating either reassignment or increased resolution periods [23]. Additionally, the vast amount of heterogeneous data that an engineer must sift through to find a root cause complicates root cause identification. Thakore, Ramasamy and Sanders propose a framework for coordinated analysis of both metric and log data into a set of meaningful features for incident analysis [21].

Lastly, Pereira et al.'s case study investigated the incident management process in a multinational company [18], identifying three best practices for the improvement of the time performance of the process. This was based on interviews from members from one team, whereas our approach involves interviews with members spread out over the organisation for a more comprehensive view of how the incident management process is regarded.

This paper contributes to the existing knowledge by creating a better understanding of the challenges hindering the practical implementation of these ideas in the incident management process. Notably, this research, focusing on financial software-defined businesses is to the best of our knowledge the first of its kind. Additionally, the study explores the most effective solution directions for these challenges.

3 BACKGROUND

In this section, we provide background on the incident management process and details on the case company.

3.1 Incident Management Process

Incident management is the process of managing the life-cycle of all incidents according to ITIL, a well-known library of best practices for managing IT services [10]. The purpose is to restore normal service operation as quickly as possible and minimise the adverse impact on business operations [4]. The process as implemented in the company contains four main stages:

1) Incident Logging: Incidents can be raised through service desk calls, mail, web, or automatic monitoring alerts. Each incident is time-stamped and logged with relevant information to maintain a full historical record. A suitable incident categorisation coding should be allocated, e.g., 'Application error'. Also, prioritisation is determined based on the urgency of the incident (how quickly the business needs a resolution) and the level of impact it is causing. The priority ranges from critical incidents (priority 1) to low incidents (priority 5).

2) Investigation & Diagnosis: Each incident is investigated and diagnosed for what has gone wrong. All activities should be documented for a complete historical record. The first priority is to restore the service to the customer as soon as possible. Analysing the root cause is not a priority at this stage and should be done as a follow up action. If the current team is unable to resolve the incident then it is escalated to another team with more expertise. If at any time, a major incident (an unplanned interruption of an IT service with major impact on customers or critical internal business processes/ services) is suspected, then it is forwarded to a dedicated *Major Incident Management* process.

3) Resolution: Upon identifying a possible resolution, this should be applied to quickly recover the service.

4) Verification & Closure: Before closing an incident, it is checked that it has been fully resolved and the user is satisfied. The incident record is completed, encompassing a closure categorisation, a comprehensive incident documentation, and a determination if the root cause was identified. If a root cause was not identified, a new problem is raised in the separate *Problem Management* process.

3.2 The Case Company

To address our research questions, we conducted a case study at ING, a large internationally operating bank with over 15,000 developers, offering a range of financial products and services to millions of customers. As a response to the growing importance of ICT in the financial sector, ING has become predominantly digital and heavily reliant on software, thus we classify it as a software-defined business.

In 2014, the company started employing an incident management process based on ITIL. This process is integrated with an Agile DevOps way of working, where multi-disciplinary teams are responsible for entire processes and value chains, end-to-end.

Banks are tightly connected to the national (and global) economy and regulatory agencies maintain control over its practices with risk guidelines and policies. Since the majority of the company is based in Europe, the European Banking Authority (EBA) regulation and policy, especially the revised Payment Services Directive (PSD2) [17], is a big influence on how the incident management process is conducted.

4 RESEARCH DESIGN

To obtain a rigorous understanding of the current state of practice of the incident management process in a software-defined company, we conduct a single-case exploratory case study at ING, set up using the guidelines from Yin [24] and Runeson and Höst [19].

4.1 Data Collection

Informal conversational interviews with open-ended questions is the main source of data. An embedded mixed-method design is employed, where the primary qualitative interview data is supplemented by internal qualitative documentation and quantitative incident analysis. Following a reflexive, iterative process inspired by Halcomb and Davidson [14], the steps for data collection are: 1) recording of the interview and concurrent note taking, 2) reflective journaling immediately post-interview 3) reviewing the recording and revising fieldnotes, 4) content analysis: thematic review, 5) triangulation, and 6) data analysis.

4.2 Participants

We interviewed 15 participants, chosen based on their role and involvement in the incident management process. They are found by using prior knowledge of experts in the incident management process and through the recommendations of other interviewees. Additional data sources suggested by interviewees were considered in the analysis. Their roles are shown in Table 1.

Table 1: Overview of interviewees

Participant IDs	Role
P1, P3, P4, P7, P11, P13	Manager
P2, P5, P8, P9, P10, P12	Engineer
P6	Coordinator
P14, P15	Contact Centre

4.3 Interview Design

The first five interviews, conducted by the first two authors, serve as a trial run for the interview protocol. The remaining ten interviews were conducted solely by the first author. The interview protocol questions are presented below:

(1) What department are you part of? What is your current job title and could you give a brief description?
(2) Could you explain your role in the incident management process?
(3) How do you encounter incidents in your work?
(4) What type of incidents do you encounter most?
(5) Which incidents have the highest impact?
(6) How do you manage your incidents?
(7) What tools are used to manage incidents?
(8) What are your main challenges in handling incidents?
(9) Do regulations influence your handling of incidents?
(10) How do you handle structural fixes?
(11) How do you train new engineers in the process?
(12) Where do you see the incident management process in five years?

Three additional questions were added during the piloting phase since they provided valuable insights during content analysis, namely questions 9, 10, and 11. Each interview was held virtually, took approximately one hour, was video recorded and notes were made throughout.

4.4 Post-interview Strategy

Immediately after each interview, a reflective journaling session is conducted to review and formalise the initial notes with more elaborate comments and perceptions, preferring memoing over verbatim transcriptions for efficiency [14]. Memoing lays more focus on interpretation and capturing the meanings of the data. The recordings and automatically generated transcripts are used to assist us in filling in the blank spaces in our notes.

During the refining of the notes, the content is analysed by coding the notes into themes based on incident management process stages or frequently occurring topics, using the technique of Fereday and Muir-Cochrane [11].

4.5 Triangulation

After completing and analysing all interviews, we triangulate the findings with company documentation and data. This includes the incident management process overview, guidelines, and additional documentation identified by interviewees such as the way of working principles, the EBA regulation and policy [1], and Google's Site Reliability Engineering (SRE) book [4]. Historical incident data from the incident management system of the company for the second half of 2021 is utilised for further data triangulation.

4.6 Data Analysis

To address the research questions, themes and insights from the interviews are grouped per question and organised according to the four main stages of the incident management process outlined in Section 3, with an additional group for the whole process.

5 RESULTS

This section delves into the study results, organised according to the research questions. Each insight is attributed to the relevant participant by referencing their ID from Table 1.

5.1 RQ1: Usage of Tools

All interviewed participants mentioned the use of the Incident Management System to support the incident management process. This system is prescribed for all stages discussed in Section 3. Additional tools either aid or regulate this process.

All incidents are logged in the Incident Management System after they have been reported via an employee, a client, or a monitoring tool, following ISO/IEC 20000 policies [2]. This data logging is useful for systematically learning from the past and ensures effective service management [4]. In addition to the overarching system, there are several tools available that monitor metrics. The monitoring sends out specific indicators of operations, called events, that create an incident when a certain threshold is reached, and when configured send out an alert to the corresponding team who needs to pick up the incident. Notably, 70% of production incidents in the second half of 2021 were reported by monitoring tools. This means 30% of incidents were logged manually, which should be avoided since software should do the interpreting on monitoring, and humans should be notified only when they need to take action [4]. This manual incident discovery is seen as a monitoring failure. According to P1, there is a lot of freedom in how teams set up monitoring, due to each business line having quite a large degree of independence in which techniques, infrastructure, or applications are utilised.

In the *Investigation & Diagnosis* stage, the first priority is restoring the service to the customer as soon as possible (P1, P4, P6, P7, P9, P12, P15), congruent with the company's Global Process Description. Teams handle this in their own way, since the company encourages the team to be self-organising in their way of working. Experience plays a big role in investigating incidents (P2, P5, P14, P15). As P2 puts it: "*I try to understand what the incident is about to solve it. Most of the incident information is a bunch of random text, but there are a couple of keywords that I look for.*" Dashboard tooling aids engineers in keeping track of their responsibility and, if configured correctly, helps them in determining the root cause of an incident. Since ING employs a lot of freedom in their use of technology, there is free choice of dashboard tooling, how it is used, and what it displays.

When the diagnosis is completed, the potential resolution should be applied and tested within the time limits that are defined at an organisational level, which takes into account the priority of the incident (P5, P8, P15). Reporting dashboards aid managers and support teams in keeping track of the resolution status of incidents. As P15 mentioned: "*Someone monitors the progress of outstanding incidents in the system. If more attention is needed, they will approach the responsible squad*". These dashboards are also used for regulatory purposes in adherence to the PSD2 guidelines, which state that a financial institution should consider the time required to implement changes and the time to take appropriate interim mitigating measures to minimise ICT and security risks, to stay within the financial institution's ICT and security risk appetite [9]. This is supported by manager P1, who elaborates upon the risk control report that measures the long overdue priority one and two incidents. "*If you have long overdue incidents then you are not in control of your incident process and are at risk. Then we do not comply with the regulations of the European Bank that we should be in control.*" This risk control report and regulations were mentioned by half of the manager interviewees (P4, P6, P11).

Another PSD2 guideline is granting remote administrative access to critical ICT systems only on a need-to-know basis and when strong authentication solutions are used [9], to ensure secure access and communication. Half of the engineers (P2, P5, P10) reported that ING complies by ensuring that production machine access goes through a prescribed access management tool and follows the four-eye principle, meaning that the approval of two individuals is needed.

Observation 1. *The incident management process prescribes tools to be used for incident and access management and offers teams flexibility in additional tooling (monitoring, dashboards, and documentation). Besides timely resolution, demonstrable regulatory compliance is a key driver in the process.*

5.2 RQ2: Main challenges

The main challenges experienced by interviewees are categorised into five groups, based on the four incident management process stages (cf. Section 3) and the whole process.

5.2.1 Logging. The incident management process starts when an incident is logged. Unfortunately, it faces challenges such as false positives, incidents that do not indicate an interruption or reduction of a service. These false incidents are time-consuming and create unnecessary extra work since every incident should be checked thoroughly. Introducing SRE teams should combat the presence of these unhelpful incidents, since SRE focuses on improving the design and operation of systems to make them more scalable, reliable, and efficient [4]. One of the focuses of SRE teams is ensuring that monitoring is helpful. Alerting a human on an incident is quite an expensive use of an employee's time and when they occur too frequently employees respond less. Effective alerting systems have good signal and very low noise [10].

Another challenge is the influx of similar incidents. P11 mentions time-consuming duplicate incidents that predominantly come from monitoring. The contact centre participant P15 mentions "*Our team notices very quickly when a disturbance is happening since we get a lot of recurrent traffic.*"

Observation 2. *Logged incidents often consist of false positives and duplicates, presenting significant challenges to subsequent incident management stages, yet it is not trivial to avoid them.*

5.2.2 Investigation & Diagnosis. Identifying the cause of incidents poses several challenges. The company's IT environment, with numerous micro-services managed by different teams, adds complexity (P9, P11, P12, P13), which can make incident triaging time-consuming (P2, P10, P14). Each team manages its own documentation (P2, P9, P13, P14), which means its scattered. Engineer P9 states that the service chains (the configuration items that are connected to a single service) are complex and that no one can comprehend the entire landscape, but they do understand what is happening on a higher level. There are numerous single points of failure in the IT environment because when one thing breaks down, much will break down with it (P9, P11), which makes it difficult to identify the root cause.

Additionally, the continuously changing nature inherent to a large-scale software company increases the complexity of the IT environment (P7, P9, P10, P12). Teams either get new names or move to different departments (P7, P10), which can make assigning an incident to another team quite difficult. This dynamic environment is also reflected in the service chains, as P9 states "*Even if we understand the chain today, it will be different in a month*".

Observation 3. *Obtaining a comprehensive overview of the entire company is crucial for investigating and diagnosing incidents. However, the fast-changing and complex nature of a large-scale software organisation poses significant obstacles to achieving this task.*

5.2.3 Resolution. An access management tool is used to allow engineers access to production machines to resolve incidents, which is taken seriously as only specific people have the privilege to approve these access requests. This can be time-consuming for the requester and approver since the requester needs to wait for the approver until they can proceed with their resolution (P2, P5). Hence, the approver may need to handle many urgent approvals that result in them feeling overwhelmed because they need to properly check each individual approval (P10). These approvals can take quite some time when they pile up. The use of an access management tool opposes Google SRE's best practices [4], which states that full autonomy should be given within the assigned role to all incident participants. This is hindered due to regulatory restrictions at the company.

Observation 4. *The adoption of incident management best practices is restricted due to compliance with risk guidelines for banking.*

5.2.4 Verification & Closure. The final stage of incident management involves challenges in the quality of incident administration and the handling of structural fixes. Participants often mention the proper administration of incidents in the system (P1–P3, P7, P8, P11, P14). The registration quality is part of the company's organisational standards which states that at all times the record must reflect the actual status of the incident and the work done. The team that changes the status to resolved, is responsible for a final check of the quality of the registration.

The solution is not always properly filled in by engineers when closing an incident (P1, P2, P7, P8, P11). Engineer P8 mentions "*needing to reinvent the wheel when solving an incident again*" when a solution is not properly filled in for a reoccurring incident. Managers P1 and P7 mention that engineers are more focused on resolving an incident than on administration, which is supported by the company's purpose of the process and SRE best practices stating to restore normal service operation as quickly as possible, which prioritises stopping the bleeding before finding the root cause. However, some participants shine a different light: they do not see the usefulness of properly filling the incident ticket in (P8, P11, P14). P8 states "*They can be important for reporting people, but no one on the work floor knows what is being done with it. So from an engineering perspective, these are just fields you have to mandatory fill in.*" This is evident by the cause code field in the incident data, which is not regularly filled in with a known cause, since 45% of the closed production incidents in the second half of 2021 had an 'unknown' cause code. A few participants mentioned that not everything in the incident ticket is important, so unnecessary fields should be removed (P3, P8, P11). This could aid the process discipline, as manager P11 states "*everybody has a very, very tight schedule and a lot to do*" and this decreases the time spent on administration.

After resolving an incident without identifying the root cause, the company mandates the creation of a new problem. This initiates the problem management process, aiming for preventive action to ensure that structural solutions are implemented that will prevent recurrence. After problem identification, there is a trade-off between the cost of a structural fix and incident recurrence (P4, P7). Participant P7 gives an example: "*If the most common solution of switching off and on your device helps, but structurally nothing happened then you have to think about if you need to fix it structurally.*"

Observation 5. *The incident management process tends to focus more on incident resolution rather than on procedural guidelines. The administration of incident tickets is interpreted in different ways amongst stakeholders, leading to inconsistent reports.*

5.2.5 The Process. The incident management process as a whole faces challenges such as communication issues, the collaboration between the IT and business side, and the way of working.

Manager P13 mentions that the number one complaint from clients about incidents is "*clients have to inform the company about incidents*". This contradicts the ISO/IEC 20000 policy [2] stating that the customer shall be kept informed of their reported incident's progress, and be alerted in advance if their service levels cannot be met and an action agreed. The speed of communication is important, since P13 states "*A client prefers receiving a very generic message quickly than waiting for 30 minutes for a detailed message.*" Contact centre employees P14 and P15 mention the need for confirmation from the IT side that an incident occurred before officially registering a disturbance message for clients. Until this happens, personnel receive many calls from clients reporting a disturbance. A challenge is that engineers tend to first try to solve incidents

themselves before reporting them (P1, P7), which means that valuable time may be lost. P7 explains that, since COVID-19, this tends to happen more frequently because more engineers are working from home.

Additionally, the collaboration between the IT and business sides can be challenging. Participants dealing with client issues daily mention the IT side not always realising the client impact (P13, P14, P15), often resulting in discussions meaning more time consumption and longer client impact. The participants do mention that client impact is hard to measure precisely since they base the estimations on what is reported. When a solution for a client reported incident has been determined, this should be communicated back to them. Participants P14 and P15 mention that they cannot always trust that the solution is understandable for the receiving person, so they intercept it, acting as a filter between the clients and the IT side to ensure that the right information reaches each side.

Moreover, the way of working presents its own set of challenges. The company adopts Agile DevOps practices, fostering the empowerment of multi-disciplinary teams to be responsible for entire processes and value chains, end-to-end. The majority of participants mention that each department and/or team may decide how they set up their way of working, however not every incident is getting picked up in time so there are some gaps. This is also visible in the data: a small percentage of production incidents in the second half of 2021 have not been closed a few months later. Manager P11 mentions that the current way of working might not work "*because you get misunderstandings between different teams who is responsible*", like an incident being reassigned many times.

Lastly, several participants recommend having dedicated people who focus on incident management in different departments (P11, P15). There is still a need for people to monitor outgoing incidents to make sure that they are picked up and do not become overdue (P11, P14, P15). P11 mentions that several years ago all the incident managers of the process were removed, so responsibility was put on the teams. Thus teams have much to keep track of causing some technical problems to fall through the cracks. This results in some departments creating additional roles for people that still do much incident management as a workaround (P11, P15), like service managers or technical contact centre employees.

Observation 6. *Effective communication is vital in incident management to ensure each impacted party receives the right form of information on time. Also, a need is present for guarding the way of working in the process to ensure responsibility for a proper resolution of all incidents.*

5.3 RQ3: Future Opportunities

The future opportunities are based on how the participants view the incident management process in the near future. The majority mention envisioning more automation and several methods are recommended regarding data-driven solutions or automating tool usage.

Data-driven Solutions. Several opportunities were mentioned that make use of data, which is in line with one of the way of working principles that states 'decisions are driven by data'. This will allow the incident management process to move from detecting incidents to predicting them (P1, P4, P7, P8, P10, P13). In particular, participants mention machine learning techniques such as anomaly detection, pattern recognition, and clustering.

Half of the managers and engineers mention the potential for using anomaly detection on monitoring data to automatically identify new anomalous events based on typical data behaviour. Participants P1, P2, and P12 mention a better use of thresholds for automatically creating incidents since they are currently based on static manual rules. Determination needs to be done on what is normal behaviour and what is not for anomalies to be detected (P2, P3). Participants mention correlating heterogeneous historical data from monitoring tools to incidents to be able to automatically detect these occurring in the future (P1, P2, P4, P7, P12). Engineer P12 mentioned combining the data of different monitoring tools with incidents to figure out the cause of the incidents. The participants P4 and P9 also mention using data related to incidents to do a root cause analysis. Engineer P9 mentions the use of data to decrease the time to detect and resolve (major) incidents since the timelines of the monitoring data can be analysed to gain "*clearer insight into the start of the fire*". To fully make use of this automation, the monitoring of systems needs to be set up properly (P9, P13). It is vital that engineers are aware that the right information needs to be monitored in an appropriate granularity for the data to be properly usable in the future.

Performing pattern recognition on the data, which is learning the relations of different data to incidents, was also mentioned by several participants as an opportunity (P4, P7, P8, P12). Manager P7 mentions using different data, like incidents, changes, events, etc., to learn the historical patterns that led to incidents. Alerting on the precursors of an incident should allow for an earlier mean time to detect and, thus a shorter mean time to resolve incidents. Engineers P8 and P9 mention that the patterns should be matched across applications. However, matching across applications has its challenges due to the company's complex IT environment (as mentioned in Section 5.2). Several participants mention wanting a better overview of the IT environment (P4, P6), which also aids in more accurately catching the client impact (P12, P13) and earlier communication to the clients (P13, P14, P15). Simplifying an IT environment is not an easy task and should be initiated on a management level company-wide.

Participants suggest employing clustering, matching incoming incidents with historically similar ones (P7, P8). The solutions of these matched incidents are presented to engineers to aid mitigation, thus decreasing the time to resolve incidents. For this to be successful, the incidents first need to be properly administrated, which is a big challenge mentioned in Section 5.2. Improving the administration can be done on three levels: management-, engineering- and organisational-level. Managers can aid engineers in making the incident information risk-compliant and useful by providing a template for an incident ticket in the Incident Management System tool. Engineers need to gain more quality awareness by informing themselves of future improvement opportunities and possible risk sanctions imposed by regulators. Lastly, on an organisational-level, SRE teams could be introduced that handle quality supervision while engineering teams still hold the responsibility for their incidents. These SRE teams must periodically check that the quality of what teams deliver is good and should intervene if not.

Observation 7. *Realising the potential of data-driven approaches such as anomaly detection, pattern recognition, and clustering historical incidents for incident management automation demands the application of best practices, clean monitoring data, and tight supervision.*

Usage of Tools. Participants anticipate and express interest in a shift from a reactive to a more proactive incident process by incorporating automated resolutions (P1, P3, P4, P6, P7, P8, P11). This involves machines automatically fixing incidents without human intervention. P3 and P4 mention existing detailed work instructions that describe every step that needs to be done, which can be automated. This ensures that when a failure happens, services can be brought back to their normal flow without calling engineers.

Care should be taken that the resolutions actually solve the incident and do not hide other problems. Manager P3 gives an example "*If a service is down and restarted every five minutes then it could be restarted 12 times an hour.*" The automation can hide deeper problems and cause much damage when discovered late. Automated resolutions will result in fewer incidents being reported in the system, however, this should be properly monitored (P3, P8, P10, P12). Manager P3 suggests a dashboard for monitoring the progress of an automation engine that shows abnormal behaviour that occurs in the automation.

As mentioned in Section 2, the majority of outages are due to changes in a live system. Best practices recommend using automation to implement progressive roll-outs, to quickly and accurately detect problems, and to roll back changes safely when problems arise [4]. This effectively minimises the effect of bad changes on the incident management process. However, this is opposed to the company's view of always having a human taking part in the decision-making (P1, P11).

Observation 8. *Incorporating human oversight is essential when implementing automated resolutions to support the incident management process.*

6 DISCUSSION

In this section, we discuss the implications of our results and go into the threats to validity of our findings.

6.1 Implications

We see the following implications of our results for the state of practice of incident management processes in financial software-defined companies.

6.1.1 Implications for Researchers. In practice, there is a higher focus on incident resolution over procedural guidelines, leading to sub-optimal data quality in incident administration. This impedes the effective application of machine learning methods. While prior research suggests strategies for improving documentation, e.g. Jiang et al. recommend documenting the mitigation process when an incident first occurs [15], more research is needed to improve administration quality without increasing resolution time. Future research should explore templates for incident tickets that ensure regulatory compliance.

Additional research is necessary to understand the impact of risk guidelines on ITIL processes, particularly incident management, as compliance can cause delays in incident resolution. Existing studies on incident management in large software companies do not address this [8, 20]. More research and best practices are needed to address the unique challenges of software companies, especially in domains like financial services, where strict regulatory constraints may limit conventional approaches.

6.1.2 Implications for Practitioners. Efficient use of data-driven solutions and novel machine learning models requires good data quality. However, practical challenges include false positives and duplicate incidents, an incomplete IT overview, and inconsistent administration quality. These false incidents, akin to transient alerts and duplicate incidents found in IBM [16] and Microsoft [8] data, threaten analysis quality and hinder incident resolution and root cause analysis efficiency. Poor data quality also obstructs historical data analysis and machine learning automation. To ensure awareness and responsibility for the data quality, there is a need for supervision to guard the way of working. The introduction of SRE teams, as mentioned in Section 5.3, could aid automation, but implementation complexity varies with company type and system complexity. Applicability to ING is constrained by regulatory compliance priorities that stem from the fintech domain.

The complexity of the IT environment influences the adoption of machine learning approaches. Gaining a clearer overview is crucial, as seen in other large cloud enterprises like Microsoft, which faces challenges with incomplete and vague service relationships across the entire system [7]. This leads to an imprecise impact estimation and longer incident resolution, which is also reported by participants. Participants and Wang et al. [23] mention having no complete view of their service chain and having to go through a large amount of data while investigating and diagnosing the incident. Pattern recognition, such as the framework by Thakore, Ramasamy, and Sanders, could help coordinate analyses of monitoring data, providing meaningful features for incident analysis [21]. These linked incidents give a better view of what services are impacted and which teams are responsible, simplifying incident triage and improving the client impact estimation. This should be integrated into dashboard tooling for effective incident resolution. The IcM BRAIN framework by Chen et al. [8] could also be useful for correlating incidents with events as patterns.

Effective monitoring is crucial for gaining observability in the IT environment and capturing errors without overwhelming engineers. Frequent incidents, such as duplicates, can lead to a decrease in the quality of response, which may indicate problems with system design, monitoring sensitivity, and/or the response to structural fixes [4]. Therefore, it's important to learn from failures and actively work towards resolving them. Proper monitoring enables automated incident logging by detecting anomalies through smart monitoring. The introduction of SRE teams ensures tight supervision of monitoring. Chen et al.'s IcM BRAIN framework could be employed for anomaly detection on heterogeneous monitoring data to detect incidents [8].

Clustering incoming incidents with historical incidents for resolution recommendations to aid in the manual resolution of incidents is an important future opportunity. Jiang et al. propose a similar approach with DeepRMD, automatically recommending user instructions for solving incidents based on the textual similarity between an incident description and instruction [15]. However,

the successful implementation of such approaches is hindered by current administration quality challenges.

Participants and existing research, such as Silito and Kutomi [20], highlight automated resolutions as a future direction, emphasising their potential for faster incident resolution. However, caution is advised, as improper implementation can exacerbate incidents. The success of automated resolution relies on the effective use of proper monitoring, making it a prerequisite for successful implementation.

6.2 Threats to Validity

Construct validity is enhanced by employing multiple sources of evidence, which increases research precision by offering a comprehensive view of the current state of the incident management process. As detailed in the research design, qualitative interview findings are triangulated with documentation and data. Also, sharing the paper draft with participants enables them to review and ensure the accuracy of the results, reducing the impact of researchers' knowledge and assumptions.

While this case study focuses on a single company, we contend that the challenges faced by ING are likely applicable to other large financial software-defined businesses in Europe. Banks, known for stringent regulatory compliance, and supported by extensive IT infrastructure, are likely to share similar challenges. To achieve more broadly applicable results, additional case studies at various software-defined businesses are necessary. Results from financial service businesses outside Europe may differ due to distinct regulations and guidelines.

Reliability is maintained through detailed documentation of the research design, allowing for replication of the study. Consistency in results hinges on participant selection and respondent bias. A potential threat arises from the self-selection of participants, potentially leaving out relevant employees. Respondent bias is mitigated by ensuring interviewees understand the interview goals and feel at ease. While the recording of interviews may induce caution, participants are assured that recordings are only used for refining notes.

7 CONCLUSION

In conclusion, efficient and compliant incident management is crucial for financial software-defined businesses. This paper sought a deeper understanding through a single-case exploratory case study at ING, a multinational bank, focusing on tool usage, challenges, and future opportunities in incident management. The findings highlight the use of a mix of prescribed (incident management system and access management tool) and tailored tools (for monitoring, dashboards, and documentation), with challenges arising at each stage of the process as well as overall, contributing to extended resolution times. The challenges are related to the complexity of the environment and compliance with procedural guidelines. This calls for clear lines of communication to guard the proper way of working. Since incident management is rich in (historic) data, machine learning approaches provide a promising area for future improvements in incident resolution time and compliance. Realising this opportunity requires a focus on cleaning, integrating, and de-duplicating the many different data sources. Also, supervision is needed on the use of best practices, periodic reviews, and clean monitoring data. Additionally, accurate modeling of the operational environment, including its dependencies and evolution, is crucial for contextualising the data.

REFERENCES

[1] 2022. Regulation and policy. https://www.eba.europa.eu/regulation-and-policy
[2] ISO/IEC JTC 1/SC 40. 2018. ISO/IEC 20000-10:2018(en) Information technology — Service management. https://www.iso.org/obp/ui/
[3] Rainer Alt, Jan Marco Leimeister, Thomas Priemuth, Stephan Sachse, Nils Urbach, and Nico Wunderlich. 2020. Software-Defined Business. *Business & Information Systems Engineering* 62, 6 (2020), 609–621.
[4] Betsy Beyer, Chris Jones, Jennifer Petoff, and Niall Richard Murphy. 2016. *Site reliability engineering: How Google runs production systems.* " O'Reilly Media, Inc.".
[5] Junjie Chen, Xiaoting He, Qingwei Lin, Yong Xu, Hongyu Zhang, et al. 2019. An empirical investigation of incident triage for online service systems. In *ICSE-SEIP 2019*. IEEE, 111–120.
[6] Junjie Chen, Xiaoting He, Qingwei Lin, Hongyu Zhang, Dan Hao, et al. 2019. Continuous incident triage for large-scale online service systems. In *ASE 2019*. IEEE, 364–375.
[7] Junjie Chen, Shu Zhang, Xiaoting He, Qingwei Lin, Hongyu Zhang, et al. 2020. How incidental are the incidents? characterizing and prioritizing incidents for large-scale online service systems. In *ASE 2020*. 373–384.
[8] Zhuangbin Chen, Yu Kang, Liqun Li, Xu Zhang, Hongyu Zhang, et al. 2020. Towards intelligent incident management: why we need it and how we make it. In *ESEC/FSE 2020*. ACM, 1487–1497.
[9] EBA. 2019. Guidelines on ICT and Security Risk Management. https://www.eba.europa.eu/regulation-and-policy/internal-governance/guidelines-on-ict-and-security-risk-management
[10] IBM Cloud Education. 2019. IT Infrastructure Library (ITIL). https://www.ibm.com/cloud/learn/it-infrastructure-library
[11] J. Fereday and E. Muir-Cochrane. 2006. Demonstrating rigor using thematic analysis: A hybrid approach of inductive and deductive coding and theme development. *International journal of qualitative methods* 5, 1 (2006), 80–92.
[12] Sinem Güven and Karin Murthy. 2016. Understanding the role of change in incident prevention. In *2016 12th CNSM*. IEEE, 268–271.
[13] Sinem Güven, Karin Murthy, Larisa Shwartz, and Amit Paradkar. 2016. Towards establishing causality between Change and Incident. In *2016 IEEE/IFIP NOMS*. IEEE, 937–942.
[14] Elizabeth J Halcomb and Patricia M Davidson. 2006. Is verbatim transcription of interview data always necessary? *Applied nursing research* 19, 1 (2006), 38–42.
[15] Jiajun Jiang, Weihai Lu, Junjie Chen, Qingwei Lin, Pu Zhao, et al. 2020. How to mitigate the incident? an effective troubleshooting guide recommendation technique for online service systems. In *ESEC/FSE 2020*. 1410–1420.
[16] Patricia Marcu, Genady Grabarnik, Laura Luan, Daniela Rosu, Larisa Shwartz, and Chris Ward. 2009. Towards an optimized model of incident ticket correlation. In *2009 IFIP/IEEE IM*. IEEE, 569–576.
[17] European Parliament and Council of the EU. 2015-12-23. Directive (EU) 2015/2366 of the European Parliament and of the Council of 25 November 2015 on payment services in the internal market, amending Directives 2002/65/EC, 2009/110/EC and 2013/36/EU and Regulation (EU) No 1093/2010, and repealing Directive 2007/64/EC. *OJ* L 337 (2015-12-23), 35–127.
[18] Rúben Pereira, José Braga de Vasconcelos, Álvaro Rocha, and Isaías Scalabrin Bianchi. 2021. Business process management heuristics in IT service management: a case study for incident management. *Computational and Mathematical Organization Theory* 27, 3 (2021), 264–301.
[19] Per Runeson and Martin Höst. 2009. Guidelines for conducting and reporting case study research in software engineering. *Empirical software engineering* 14, 2 (2009), 131–164.
[20] Jonathan Sillito and Esdras Kutomi. 2020. Failures and Fixes: A Study of Software System Incident Response. In *ICSME 2020*. IEEE, 185–195.
[21] Uttam Thakore, Harigovind V Ramasamy, and William H Sanders. 2019. Coordinated Analysis of Heterogeneous Monitor Data in Enterprise Clouds for Incident Response. In *ISSREW 2019*. IEEE, 53–58.
[22] Weijing Wang, Junjie Chen, Lin Yang, Hongyu Zhang, Pu Zhao, et al. 2021. How Long Will it Take to Mitigate this Incident for Online Service Systems?. In *ISSRE'21*.
[23] Yaohui Wang, Guozheng Li, Zijian Wang, Yu Kang, Yangfan Zhou, et al. 2021. Fast Outage Analysis of Large-scale Production Clouds with Service Correlation Mining. In *ICSE 2021*. IEEE, 885–896.
[24] Robert K Yin. 2009. *Case study research: Design and methods*. Vol. 5. sage.
[25] Nengwen Zhao, Junjie Chen, Zhaoyang Yu, Honglin Wang, Jiesong Li, et al. 2021. Identifying bad software changes via multimodal anomaly detection for online service systems. In *ESEC/FSE 2021*. 527–539.

Received 7 December 2023; accepted 25 January 2024