The University of Hong Kong

COMP7705 MSc(CompSc) Project

Interim Report

Cybersecurity Compliance and Reporting Platform

YIP Wankit, Daniel 3036383678

CHAN Cheung Hei 3036381280

SONG Insu 3036199596

WONG Kwun Yuet Shavonne 2013534309

YEUNG Hiu Ying 3036379976

Supervisor: Dr. P.S. Vivien Chan

1 June 2025

Contents

1. Project Overview

Our project focuses on streamlining cybersecurity incident reporting and compliance in Hong Kong using a centralized platform. The primary objectives of our project are to develop a user-friendly platform that evaluates the nature and severity of incidents, provides clear guidance on Hong Kong's regulatory requirements, generates compliant reports, and securely stores these reports for future reference by regulators, hoping to improve the efficiency and accuracy of incident management processes while ensuring alignment with local regulatory standards.

Our research conducted to date in relation to this project revolves around (1) the current and upcoming regulatory requirements for cybersecurity incident reporting, (2) the design of the framework in making evaluation of the incident and (3) providing users with a recommendation on whether to make a report to the relevant regulators through our platform, with reference to existing international and local standards and guidance materials.

The methodology involves conducting extensive research on Hong Kong's regulatory framework, analyzing current challenges in incident reporting, utilizing development techniques to create and refine the platform, which consists of (1) regulatory guidance page, (2) user input form, (3) incident evaluation and report generation function and (4) incident storage database. We are currently further improving the functionality of the platform. We hope to carry out user testing to ensure the platform's reliability and effectiveness.

The expected outcomes include a fully functional platform that simplifies and standardizes incident reporting, reduces reporting errors, and enhances compliance with local regulations. Ultimately, this project aims to save time and resources for organizations while improving transparency and accountability in regulatory processes.

2. Literature Review

(a) The Importance of Cybersecurity Incident Reporting

Timely and effective cybersecurity incident reporting plays a vital role in strengthening both individual organisational resilience and the broader cybersecurity ecosystem. In jurisdictions without a national reporting framework, studies have shown that responses to incidents are often delayed, poorly coordinated, and underreported. [1] One of the major benefits of early reporting is that it allows regulators to issue alerts to other potentially affected organisations, helping them take preventive measures before they become targets. This kind of proactive information-sharing can reduce the spread of cyber threats and lower the risk of repeated attacks across the same supply chains or sectors. Incident reporting is also a key tool for regulators in coordinating sector-wide responses and improving cybersecurity policy. When patterns emerge across multiple reports, they can inform timely updates to risk assessments, technical guidance, and security controls. [2] Over time, reported cases also support policy learning, enabling authorities to identify gaps or outdated rules and make targeted reforms. [3] [4].

Despite these advantages, reporting is still held back by several persistent barriers. Unclear reporting procedures, fear of reputational damage and a lack of feedback often discourage frontline staff from escalating incidents. [5] [6] Many users are uncertain about what qualifies as a "material", "serious" or "significant" event, and without clear guidance or user support, incidents may go unreported. Research also highlights additional challenges, including the lack of a national policy for incident reporting, insufficient information sharing, and uncoordinated defensive actions. Furthermore, users may misinterpret incidents as technical issues rather than security threats, and a lack of perceived benefits or fear of embarrassment can further discourage reporting. These challenges are particularly relevant in Hong Kong, where the absence of mandatory reporting regulations exacerbates the issue. These challenges show the need for systems that are not only compliant with legal requirements but also user-friendly and easy to navigate. Digital reporting tools with guided forms, built-in logic and sector-specific support can help reduce confusion and improve the consistency, timeliness and accuracy of cybersecurity disclosures.

(b) The Regulatory Landscape in Hong Kong

Currently in Hong Kong, various regulators, such as the Securities and Futures Commission, have requirements and guidelines on reporting incidents, and multiple entities like the Hong Kong Computer Emergency Response Team Coordination Centre and the Hong Kong Police Force also accept reports, with some regulators offering quantitative thresholds and others relying on vague terms like "material", "serious" or "significant". In addition, from 1 January 2026, the Protection of Critical Infrastructures (Computer System) Bill (the "**Bill**") will become effective which will require cybersecurity incidents relating to Critical Infrastructure to be reported within specific timeframes, depending on whether the incidents are "serious" or not.

Following our review of the regulations and guidelines in Hong Kong, given that Hong Kong has numerous industries with diverse reporting requirements, it is impractical to cover all of them within this project. To ensure a focused and effective approach, we have limited the scope of review and focus of this project to eight critical industries: Energy, Information Technology, Banking and Financial Services, Air Transport, Land Transport, Maritime Transport, Healthcare Services, and Telecommunications and Broadcasting Services. This allows us to address the specific needs of these sectors and their relevant regulatory bodies efficiently.

Table 1 sets out the government departments and regulatory bodies that have governing power and/or accept reports from organisations in each of the industries we focus on:

Industry	**All Matters**	**Data Breaches**	**Criminal Matters**	**Incidents under the Bill (which may or may not be Criminal related) (w.e.f 1 Jan 2026)**
Energy	Electrical and Mechanical Services Department, Environment and Ecology Bureau	Office of the Privacy Commissioner for Personal Data (PCPD)	Hong Kong Police Force	Critical Infrastructure Commissioner
Information Technology	N/A	Office of the Privacy Commissioner for Personal Data (PCPD)	Hong Kong Police Force	Critical Infrastructure Commissioner
Banking and Financial Services	Hong Kong Monetary Authority (HKMA) Securities and Futures Commission (SFC)	Office of the Privacy Commissioner for Personal Data (PCPD)	Hong Kong Police Force	Critical Infrastructure Commissioner
Air Transport	Civil Aviation Department	Office of the Privacy Commissioner for Personal Data (PCPD)	Hong Kong Police Force	Critical Infrastructure Commissioner
Land Transport	Transport Department	Office of the Privacy Commissioner for Personal Data (PCPD)	Hong Kong Police Force	Critical Infrastructure Commissioner
Maritime Transport	Marine Department	Office of the Privacy Commissioner for Personal Data (PCPD)	Hong Kong Police Force	Critical Infrastructure Commissioner
Healthcare Services	Department of Health Hospital Authority Commissioner for the Electronic Health Record	Office of the Privacy Commissioner for Personal Data (PCPD)	Hong Kong Police Force	Critical Infrastructure Commissioner
Telecommunications and Broadcasting Services	Office of the Communications Authority (OFCA)	Office of the Privacy Commissioner for Personal Data (PCPD)	Hong Kong Police Force	Critical Infrastructure Commissioner

Table 1: Responsible government departments and regulators for the key industries

Table 2 sets out the requirements for reporting for each of the industries we focus on:

Industry	Reporting Requirement								
	All Matters			Data Breaches / Privacy			Criminal Matters		
	Obligation	**Timing**	**Content**	**Obligation**	**Timing**	**Content**	**Obligation**	**Timing**	**Content**
Energy	No explicit requirements for cybersecurity reporting. May make a report to the responsible government department/regulatory if needed. (w.e.f. 1 Jan 2026, security incidents as defined under the Bill should be reported to Critical Infrastructure Commissioner)	N/A (w.e.f. 1 Jan 2026, serious security incidents as defined under the Bill should be reported to Critical Infrastructure Commissioner within 12 hours and other incidents within 48 hours)	N/A (w.e.f. 1 Jan 2026, incidents as defined under the Bill be reported in specified form to be published by Critical Infrastructure Commissioner)	*Encouraged* to make data breach notifications to PCPD. [47]	As soon as practicable. [47]	As per Data Breach Notification Form published by PCPD from time to time. [56]	When needed, should make a criminal activity report to the Police.	N/A	As per e-report form published by the Police from time to time. [57]
Information Technology	No explicit requirements for cybersecurity reporting. May make a report to the responsible government department/regulatory if needed. (w.e.f. 1 Jan 2026, security incidents as defined under the Bill should be reported to Critical Infrastructure Commissioner)	N/A (w.e.f. 1 Jan 2026, serious security incidents as defined under the Bill should be reported to Critical Infrastructure Commissioner within 12 hours and other incidents within 48 hours)	N/A (w.e.f. 1 Jan 2026, incidents as defined under the Bill be reported in specified form to be published by Critical Infrastructure Commissioner)	*Encouraged* to make data breach notifications to PCPD. [47]	As soon as practicable. [47]	As per Data Breach Notification Form published by PCPD from time to time. [56]	When needed, should make a criminal activity report to the Police.	N/A	As per e-report form published by the Police from time to time. [57]
Banking and Financial Services	HKMA: Notify of ***significant*** cybersecurity incidents immediately. [50] HKMA: Notify of ***serious*** privacy incident as soon as practicable after aware or notified [52]	HKMA: Immediate for cybersecurity incidents. [50]; as soon as practicable	HKMA: "whatever information is available at the time" [50] SFC: "giving particulars of the breach…" [53]	*Encouraged* to make data breach notifications to PCPD, consideration	As soon as practicable. [47]	As per Data Breach Notification Form published by PCPD from	When needed, should make a criminal activity	N/A	As per e-report form published by the Police

	SFC: Report ***material*** cybersecurity incidents (e.g., ransomware) immediately. [51] SFC: Report immediately ***material*** breach or non-compliance with any laws and regulations or material failure of trading systems; report promptly any material service interruption relating to electronic trading system [53] (w.e.f. 1 Jan 2026, security incidents as defined under the Bill should be reported to Critical Infrastructure Commissioner)	for privacy incidents [52] SFC: Immediate upon material incident occurrence. [51][53]; promptly for electronic trading system [53] (w.e.f. 1 Jan 2026, serious security incidents as defined under the Bill should be reported to Critical Infrastructure Commissioner within 12 hours and other incidents within 48 hours)	SFC (guideline for incident record keeping relating to electronic trading systems): (a) a clear explanation of the problem; (b) the time of outage or delay; (c) the duration of outage or delay; (d) the systems affected during outage or delay and subsequently; (e) whether this problem or a related problem has occurred before; (f) the number of clients affected at the time and the impact on these clients; (g) the steps taken to rectify the problem; and (h) steps taken to ensure that the problem does not occur again. [53] (w.e.f. 1 Jan 2026, incidents as defined under the Bill be reported in specified form to be published by Critical Infrastructure Commissioner)	s include whether privacy incident has high impact on reputation of institution, whether privacy incident has large number of customers affected, whether customer data stolen/lost/leaked is sensitive. [47][52]		time to time. [56]	report to the Police.		from time to time. [57]
Air Transport	No explicit requirements for cybersecurity reporting. May make a report to the responsible government department/regulatory if needed. (w.e.f. 1 Jan 2026, security incidents as defined under the Bill should be reported to Critical Infrastructure Commissioner)	N/A (w.e.f. 1 Jan 2026, serious security incidents as defined under the Bill should be reported to Critical Infrastructure	N/A (w.e.f. 1 Jan 2026, incidents as defined under the Bill be reported in specified form to be published by Critical	*Encouraged* to make data breach notifications to PCPD. [47]	As soon as practicable. [47]	As per Data Breach Notification Form published by PCPD from time to time. [56]	When needed, should make a criminal activity report to the Police.	N/A	As per e-report form published by the Police from time to time. [57]

		Commissioner within 12 hours and other incidents within 48 hours)	Infrastructure Commissioner)						
Land Transport	Reportable events for Autonomous Vehicles (AV) are (a) an incident involving any defect in or malfunctioning of: (i) the AV; or (ii) any AV equipment or the AV system of the AV, that leads to suspension of the operation of the AV for more than one hour; (b) an accident that involves death or injury of any person, or damage to any property, caused by, or arising out of, the operation of the AV; (c) a collision of the AV with any object; (d) an incident: (i) that undermines the safety of the AV or endangers any person or thing; or (ii) that, if not remedied, would undermine the safety of the AV or endanger any person or thing; or Examples— Fire, malfunctioning of the braking system, trapping of any passenger for over 15 minutes and an incident leading to the summoning of emergency services. (e) any other incident of a type specified in the pilot conditions of the AV. [48] (w.e.f. 1 Jan 2026, security incidents as defined under the Bill should be reported to Critical Infrastructure Commissioner)	Within 24 hours for AV reportable events. [48] (w.e.f. 1 Jan 2026, serious security incidents as defined under the Bill should be reported to Critical Infrastructure Commissioner within 12 hours and other incidents within 48 hours)	Notice within 24 hours - (a) date, time and location of the reportable event; (b) description of injury, fatality and damage (if any); (c) brief description of the reportable event; (d) immediate follow-up actions taken; and (e) details of contact person of the pilot proprietor. Detailed report to follow - (a) detailed descriptions of the reportable event; (b) investigation results of the reportable event; and (c) remedial measures taken to avoid recurrence of the event. [58] (w.e.f. 1 Jan 2026, incidents as defined under the Bill be reported in specified form to be published by Critical Infrastructure Commissioner)	*Encouraged* to make data breach notifications to PCPD. [47]	As soon as practicable. [47]	As per Data Breach Notification Form published by PCPD from time to time. [56]	When needed, should make a criminal activity report to the Police.	N/A	As per e-report form published by the Police from time to time. [57]
Maritime Transport	No explicit requirements for cybersecurity reporting. May make a report to the responsible government department/regulatory if needed. (w.e.f. 1 Jan 2026, security incidents as defined under the Bill should be reported to Critical Infrastructure Commissioner)	N/A (w.e.f. 1 Jan 2026, serious security incidents as defined under the Bill should be reported to Critical	N/A (w.e.f. 1 Jan 2026, incidents as defined under the Bill be reported in specified form to be published by Critical	*Encouraged* to make data breach notifications to PCPD. [47]	As soon as practicable. [47]	As per Data Breach Notification Form published by PCPD from time to time. [56]	When needed, should make a criminal activity report to the Police.	N/A	As per e-report form published by the Police from time to time. [57]

		Infrastructure Commissioner within 12 hours and other incidents within 48 hours)	Infrastructure Commissioner)						
Healthcare Services	Must notify the Commissioner for the Electronic Health Record and PCPD of data breaches involving the Electronic Health Record Sharing System (eHRSS) as soon as possible. [49] (w.e.f. 1 Jan 2026, security incidents as defined under the Bill should be reported to Critical Infrastructure Commissioner)	As soon as possible for data breaches involving eHRSS. [49] (w.e.f. 1 Jan 2026, serious security incidents as defined under the Bill should be reported to Critical Infrastructure Commissioner within 12 hours and other incidents within 48 hours)	N/A (w.e.f. 1 Jan 2026, incidents as defined under the Bill be reported in specified form to be published by Critical Infrastructure Commissioner)	*Encouraged* to make data breach notifications to PCPD. [47]	As soon as practicable. [47]	As per Data Breach Notification Form published by PCPD from time to time. [56]	When needed, should make a criminal activity report to the Police.	N/A	As per e-report form published by the Police from time to time. [57]
Telecommunications and Broadcasting Services	Report severe security incidents for Next Generation Networks (NGN): If incident meets any of the below, should report: 1. a security incident/violation which lasts for more than 30 minutes and results in degradation of service or failure of network component that would affect 10 000 users or more; 2. a sustained malicious attack experienced by a network element including any tampering/leakage/unauthorised access/transfer of data, interference or damage of critical network facilities/assets/systems/equipment for more than 24 hours; or 3. a severe security incident/violation which has been confirmed by the overseas counterpart and will likely affect the network service in Hong Kong. [54]	Initial report: Within 1 hour. Service restoration updates: Within 2 hours. Detailed report: Within 14 working days. [54] (w.e.f. 1 Jan 2026, serious security incidents as defined under the Bill should be reported to Critical Infrastructure Commissioner within 12 hours and other incidents within 48 hours)	(a)full name of the operator; (b)description of the incident/violation; (c)date and time of onset of the incident/violation; (d) types and estimated number of customers/end-users affected; (e)affected area(s); (f)actions taken; and (g)contact information: name of contact person as well as the person's fixed and mobile Hong Kong telephone numbers, and email address.[54] (w.e.f. 1 Jan 2026, incidents as defined under the Bill be reported	*Encouraged* to make data breach notifications to PCPD. [47]	As soon as practicable. [47]	As per Data Breach Notification Form published by PCPD from time to time. [56]	When needed, should make a criminal activity report to the Police.	N/A	As per e-report form published by the Police from time to time. [57]

	(w.e.f. 1 Jan 2026, security incidents as defined under the Bill should be reported to Critical Infrastructure Commissioner)		in specified form to be published by Critical Infrastructure Commissioner)						

Note: Under the upcoming Protection of Critical Infrastructures (Computer Systems) Bill (effective 1 Jan 2026), with effect from 1 January 2026, serious security incidents (which may be criminal or non-criminal) should be reported to the newly established Commissioner within 12 hours and other incidents within 48 hours. [55]

Table 2: Regulatory requirements of the key industries

(c) International Benchmarks and Models

Studies have shown that the lack of clear, consistent criteria creates confusion, complicates compliance efforts, reduces the effectiveness of reports, and increases the risk of underreporting or delays. To address these issues, some jurisdictions have implemented well-structured frameworks to enhance cybersecurity incident reporting. The European Union's NIS2 Directive [4] sets clear definitions, mandatory timelines and a two-stage reporting process to ensure early alerts and follow-up reporting. This approach has improved coordination and reduced uncertainty across member states [3]. In Singapore, the Cybersecurity Act [7] requires operators of critical infrastructure to report incidents to the Cyber Security Agency (CSA). Beyond legal requirements, the CSA supports reporting through standardised templates, technical guidance and sector-specific instructions, helping organisations report more consistently and confidently. Japan and South Korea also offer valuable examples. Japan's NISC [8] oversees national coordination and promotes structured information sharing. In South Korea, KISA [9] provides a centralised portal, publishes regular threat bulletins and uses reported data to guide national policy updates.

Rather than focusing solely on reporting deadlines or central authorities, these frameworks highlight the importance of clear thresholds, standard procedures and ongoing communication. They show how structured reporting can improve threat awareness, enable early intervention and support stronger policy responses. Against the backdrop of having multiple entities in Hong Kong accepting reports and as Hong Kong is bound to further add the additional requirement of reporting incidents relating to critical infrastructure through the upcoming Protection of Critical Infrastructures (Computer System) Bill effective from 1 January 2026 [10], lessons from these international models can help guide the creation of a more unified and practical reporting system tailored to its multi-regulator environment.

(d) Incident Severity Classification

For reporting requirements such as reporting "material", "serious" or "significant" incidents which requires organisations' own judgment on severity of cybersecurity incidents, assistance on assessment is essential for deciding whether an incident should be reported and how it should be handled. However, existing approaches are often inconsistent and lack clarity. A MIT thesis proposed a structured scoring model called the Cybersecurity Incident Severity Scale (CISS), which evaluates incidents based on factors such as data sensitivity, operational disruption, reputational harm and financial loss. [11] This framework aims to reduce subjectivity and help organisations make more consistent and informed reporting decisions. Although many regulators are referring to similar severity factors in their reporting guidance, they often do not provide clear rules or scoring methods. For reporting requirements such as reporting "material", "serious" or "significant" incidents, organisations need to rely on internal judgment on determining whether any incidents should be reported, which increases the risk of under-reporting or over-reporting. Researchers have also noted that uncertainty, fear of reputational damage and vague guidance can discourage staff from escalating incidents, further complicating the reporting process. [5] [6]

In recent years, there has been growing interest in using artificial intelligence, especially large language models (LLMs) to support incident analysis. LLMs can help interpret unstructured data, such as internal emails or narrative reports, and assess incident severity. However, many studies point to the lack of transparency in how these models produce results. Their "black box" nature makes them difficult to apply in regulated environments where clear explanations and audit trails are required [12] [13]. Taken together, literature highlights two main gaps: the absence of a standardised, widely accepted severity model and the challenges in applying AI tools that lack interpretability. These findings point to a need for solutions that blend structured, rule-based models with flexible, explainable tools capable of handling real-world reporting scenarios.

(e) Digital Tools and UX in Reporting Systems

Research has shown that the design and usability of incident reporting systems can strongly influence how often and how accurately incidents are reported. Researchers found that features such as simplified forms, pre-filled fields and step-by-step guidance help reduce confusion and encourage reporting, particularly from staff without legal or technical backgrounds. [5] [6] A study also explored the potential of using blockchain-based systems to support secure and tamper-proof community reporting, pointing to new possibilities for trusted reporting structures. [14]

In Hong Kong, the tools available for reporting incidents are still basic. Templates like the PCPD's Data Breach Notification Form [15] are usually provided in PDF or Word format with limited guidance and no built-in logic or feedback. This can make reporting more difficult for non-specialists and lead to inconsistent or incomplete submissions. The lack of user-friendly, digital tools is a notable gap, especially in time-sensitive situations or in organisations where reporting duties are shared across teams. These challenges highlight the need for more interactive and supportive platforms that can guide users through the reporting process and help ensure that reports meet regulatory expectations.

(f) Research Gaps and Project Contribution

While the value of timely and structured cybersecurity incident reporting is well recognised, current practices are often fragmented and unclear. While certain regulators provide quantitative thresholds for reporting, many regulators merely provide the general requirement of reporting "material" incidents, but without clearly defining the materiality or providing a clear, standardised process for assessing or reporting incidents. This lack of consistency creates confusion, delays responses, and leads to underreporting, as highlighted in both local and international studies.

Artificial intelligence, especially large language models (LLMs), is becoming more common in decision-making. LLMs can interpret free-text incident descriptions and assist in severity assessment when structured information is not available. However, these tools often lack transparency and their decision-making processes are difficult to explain. This "black box" nature makes them unsuitable for regulatory environments that require traceable, accountable reasoning.

To address these challenges, this project proposes a digital platform that improves the consistency and clarity of incident reporting in Hong Kong. The platform combines a rule-based model aligned with local regulatory thresholds and an LLM-supported tool for processing unstructured inputs. It also features a user-friendly interface that reflects official reporting forms and provides guided prompts. By bridging both technical and usability gaps, the platform aims to help organisations report incidents more accurately, confidently and in line with compliance expectations.

3. Methodology

The Cybersecurity Compliance and Reporting Platform is built on five core components: User Input, Frontend, Backend, Report and Advice, as illustrated in Figure 1.

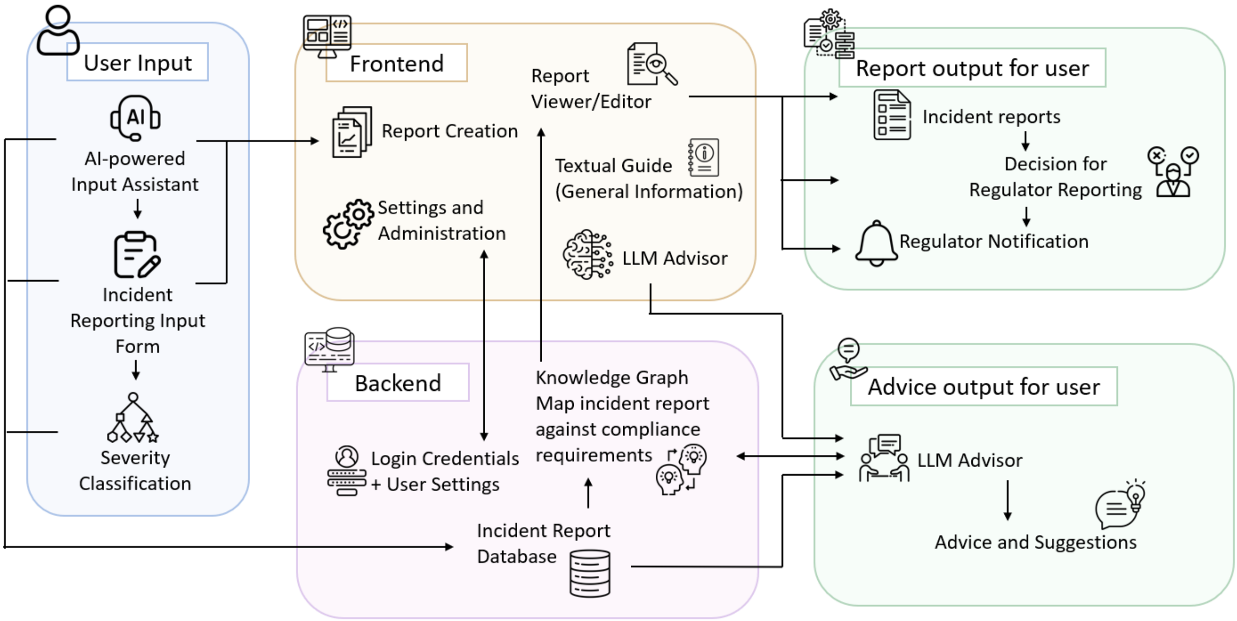

User Input
AI-powered Input Assistant
Incident Reporting Input Form
Severity Classification
Frontend
Report Creation
Settings and Administration
Report Viewer/Editor
Textual Guide (General Information)
LLM Advisor
Backend
Login Credentials + User Settings
Knowledge Graph Map incident report against compliance requirements
Incident Report Database
Report output for user
Incident reports
Decision for Regulator Reporting
Regulator Notification
Advice output for user
LLM Advisor
Advice and Suggestions

Figure 1: The overall framework with the core components.

In this section, we provide a detailed analysis of each core component, discussing the rationale behind the functionalities, the development methodologies employed and the value it delivers to the users.

3.1 User Input Component

The Input component enables users to create reports on cybersecurity incidents by entering relevant details into an online form. To enhance usability, it integrates an AI-powered assistant that allows the user to upload incident descriptions and consultant reports and extract relevant information to prefill in the form. Once all necessary details are inputted, the data is passed to the Severity Classification Model to provide an objective and consistent evaluation of the incident's nature and severity and provide users with a preliminary recommendation of whether the incident is severe enough that warrant reporting and if yes and the user agrees to create a report, the user will proceed to the Report component where the Reporting Path Recommendation Model will suggest which regulator(s) this incident should be reported to.

3.1.1 Development Methodology

- Identification of Current and Upcoming Reporting Requirements
  A review of the relevant rules and guidelines set out by various government departments and regulators of the industries that we focus on was carried out to identify their scope of governing authority and determine whether any reporting requirements were stipulated. For those with reporting obligations, the specifics of the requirements such as the reporting content and format of reports were analysed. As it is observed that certain regulators require reporting of "material", "serious" or "significant" incidents, which require independent judgment by the organisations based on their unique situations. Recognizing the ambiguity and subjectivity inherent in such classifications, we determined that developing a robust severity classification model was essential to ensure consistent and accurate incident reporting. At the same time, to ensure fulfilment of the regulatory requirements that are more clear-cut and well-defined, we decided to develop a standardised input form to streamline the process and ensure compliance with

the current regulatory requirements.

- Online Form Input Fields Design

  When designing the required input fields, we performed a comprehensive review of regulatory forms, academic references, and existing reporting systems. Our primary focus was to include the fields that are required under the current regulatory requirements. This was followed by incorporating fields to support severity classification.

  (1) *Regulator-required Fields*: The reporting form has been structured to reflect standard data fields currently required in the forms published by regulators which serve the functions of (i) complying with the existing form of reporting required of by the regulators, and (ii) enable us to identify the essential nature of the incident which could assist to analyse which regulator(s) is/are relevant to accept the reports in relation to the incident in the Report component. For example, we included the key fields in PCPD's Data Breach Notification Form [15], such as the date of incident, affected individuals, cause of incident and remedial action taken.

  (2) *Severity Classification Fields*: To assist users in evaluating the incidents and to decide whether an incident is of a "material", "serious" or "significant" nature, we developed fields that collect information from users in order to conduct severity classification. A review of relevant academic and industry publications was performed for inspiration of the design of the fields and the corresponding Severity Classification Model which could assist users in determining whether an incident is severe enough that warrants reporting. In particular, the locally issued documents such as the Practice Guide for Information Security Incident Handling [21] and the Practice Guide for IT Security Risk Management [22] served as primary references. These government-issued guides under Digital Policy Office outline how the public sector categorises incidents, sets escalation levels and records event details. While grounded in the Hong Kong context, these guides also incorporate international standards, notably the ISO/IEC [23] and the NIST guidelines [2]. Additional input came from sector-specific protocols such as the Format for Incident Reporting Exchange developed by the Financial Stability Board. We also referred to academic research, including Conard's [11] MIT thesis, which provides a model for quantifying incident severity based on structured criteria. From these sources, we identified recurring severity indicators such as the extent of service disruption, the sensitivity of compromised data, the number of individuals affected and potential financial consequences etc. These factors were consolidated into a rule-based framework that underpins the platform's classification logic.

  Severity Classification Model

  The following criteria in Table 3 were formulated to assist with the severity classification of incidents.

  - Financial Impact: Assesses the monetary loss incurred due to an incident, including direct costs (e.g., fines, fraud, contract losses) and indirect costs (e.g., business disruption, recovery expenses).
  - Operational Impact: Evaluates the effect on business functions and services, including disruption to critical systems, resource strain, and delays in delivering key operations.
  - Data Leakage: Measures the extent and sensitivity of exposed data, particularly focusing on the type (e.g., PII, health, financial) and volume of information disclosed during a breach.
  - Affected Individuals: Considers the number of people impacted, the sensitivity of their data, and potential consequences for their privacy or safety (e.g., identity theft, fraud).
  - Overall Criticality: Captures a holistic risk view that includes legal/regulatory triggers, reputational damage, public/media exposure, trust erosion, and the need for formal reporting or crisis management.

Level	Financial Impact	Operational Risk	Data Leakage	Affected Individuals	Overall Criticality
Negligible	Incidents result in minimal or no financial loss (e.g., < HK$1,000). These do not disrupt operations or require budget adjustments. Common in minor accounting discrepancies or internal mischarges.	Minor issue with no impact on operations or critical systems. Easily resolved without escalation or external help. No customer or reputational risk.	Non-sensitive or public data is exposed (e.g., names in a directory) without involving PII. No impact or risk to individuals. No notification or mitigation required.	No individuals are affected. The leaked data is publicly available or non-identifiable (e.g., organizational phonebook), posing no risk of harm, and does not trigger any reporting obligations.	The incident has no meaningful impact on operations, legal duties, or public trust. No sensitive or non-sensitive data is involved. The issue is considered trivial, with no reporting or escalation needed.
Low	Causes minor financial losses (HK$1 – HK$50,000), easily absorbed within department-level budgets. Typically involves invoice errors, small refunds, or minor vendor overpayments.	Slight disruption to non-critical systems (e.g., slow dashboards). Minimal impact on business or users. Internally managed; no legal or reporting implications.	Limited PII exposure (e.g., emails, staff lists) involving 1–10 individuals. Minimal inconvenience and low sensitivity. Typically managed internally; external reporting optional.	A small number of individuals (1–10) are affected, and the data is non-sensitive (e.g., names, emails). The risk of harm is minimal, and reporting may be optional based on internal policies or jurisdiction.	The impact is contained and minimal, involving a small volume of non-sensitive data. It is internally acknowledged without legal or trust implications. No regulatory reporting is required, and customer awareness is unlikely.
Moderate	Results in moderate financial loss (HK$50,001 – HK$500,000) that impacts departmental plans or quarterly budgets. Requires oversight from management and budget reallocation.	Noticeable disruption to services requiring workarounds or additional resources. May affect customers and attract internal or external attention. Escalation advisable.	Sensitive PII (e.g., ID numbers) leaked affecting identifiable groups (11–100 individuals). May cause identity theft risk. Regulatory notification is recommended or required.	Between 11–100 individuals are impacted, with exposure of moderately sensitive information (e.g., contact details, ID numbers). There is some risk to data subjects' rights, and regulatory notification is typically advisable.	The incident causes noticeable organizational impact, potentially involving moderately sensitive or corporate data. It may affect customer trust, require escalation, and advisable reporting to internal risk or compliance teams.
High	Leads to significant financial loss (HK$500,001 – HK$5,000,000), affecting company-wide strategy, contracts, or revenue streams. Legal or compliance review is likely triggered.	Serious disruption to critical operations. Affects customers, triggers legal or compliance concerns, and requires senior management involvement.	Large-scale exposure (101–1,000 individuals) of sensitive data such as medical or financial records. High risk to individuals; legal duties and mandatory notification to affected parties likely.	A breach involving 101–1,000 individuals and sensitive personal data (e.g., health, financial records). The likelihood of harm is high, and notification to individuals and/or authorities is usually mandatory under law.	The incident affects core business functions or customer trust, involving sensitive data at volume. It triggers legal scrutiny or regulatory obligations, and may prompt cross-functional response and external notifications.
Critical	Severe financial damage exceeding HK$5,000,000. Threatens business viability or solvency. Often involves litigation, heavy fines, or major contract terminations, requiring executive intervention.	Complete failure of core functions or infrastructure. Triggers crisis management, legal action, mandatory reporting, and national-level or executive attention.	Massive breach involving >1,000 individuals or national datasets (e.g., biometric or tax data). Severe privacy harm or legal consequences expected. Triggers urgent reporting and legal response.	Over 1,000 individuals are affected by a breach of highly sensitive data. The incident poses serious risks, such as identity theft or fraud, and often requires urgent notification, legal involvement, and public disclosure.	A severe, widespread event involving highly sensitive or national-scale data. It leads to legal action, mandatory reporting, reputational crisis, and requires executive involvement and government-level coordination.

Table 3: Guidelines for incident severity classification

Once the incident has been determined to be severe enough that warrant reporting, based on the outcome of the Severity Classification Model, the next step is to identify the appropriate regulator to report the incident to. We developed the Reporting Path Recommendation Model to identify the relevant regulator(s) which will be covered under the Report component below.

- AI-powered Form Input Assistant
  Some companies may not have a dedicated cybersecurity incident response team to handle investigations and reporting. In one scenario, they may engage external vendors or consulting firms to conduct the investigation and provide reports, which may not align with the platform's required format. In another scenario, they may be unsure about how to fill in the input form and instead choose to describe the incident. The assistant enables users to upload investigation reports from external vendors

or consulting firms, as well as incident descriptions. It then extracts relevant information to prefill the online form. We are exploring the following technologies to develop the assistant.

- DeepSeek-R1 [39]: It is an open-source language model developed by High-Flyer. It can perform advanced language processing with less computational cost.
- Ollama [40]: It is a development platform that enables local deployment of large language models (LLMs) and vector embedding models.
- Langchain [41]: It is a development framework for building applications with large language models, enabling easy retrieval and tool integration.
- Chromadb [42]: It is a fast vector database for efficient similarity searches and embedding storage of context extracted from a report.

3.1.2 User-Centric Benefits

- One-stop Information Provision

  Individuals with limited or no prior knowledge of the regulatory framework in the field may struggle to ask themselves the right questions to determine whether an incident is required to be reported. Our Input component addresses this challenge by using input fields to mandate users to provide the key information in order for us to have sufficient information to execute the process of evaluation of incidents for users.

- Prompt Reporting Recommendation

  After the data is entered into the reporting form, the interface in the Input component will display a pop-up window recommending whether the incident is severe enough that warrants reporting, driven by the Severity Classification Model. If reporting is recommended, the interface in the Report component that follows will also suggest the appropriate regulators to receive the reports applying the Reporting Path Recommendation Model.

  Determining whether an incident is “material”, “serious” or “significant” often involves multiple internal stakeholders and can take considerable time due to lengthy discussions as there are multiple and varying factors that stakeholders take into account. Additionally, communication gaps frequently arise between technical and non-technical stakeholders, leading to misalignment. The process also tends to rely heavily on subjective judgment and individual experience, making it difficult to benchmark against other organizations. We use input fields to collect the key information for severity classification. The severity classification model consolidates key considerations and provides highly reliable, consistent suggestions for incident severity ratings. The severity classification model delivers the following value.

  - Faster, clearer decision-making: Reduces the time spent in back-and-forth discussions among stakeholders.
  - Bridges communication gaps: Aligns technical and non-technical teams using a shared, structured framework.
  - Enhances objectivity and consistency: Removes reliance on personal experience or subjective judgment.
  - Supports defensible decisions: Provides traceable and standardized severity ratings that are easier to justify to regulators or management.

- Incident Input Assistance

  In the previously mentioned scenarios, the user may encounter the following challenges.

  - Manual data entry burden: After receiving an external vendor/consultant report, the staff must read, interpret, and manually transfer complex technical findings into a reporting form, which is slow and error-prone.
  - Compliance risk: Mistakes or delays in translating findings into formal reports can result in missed

reporting deadlines or incomplete submissions.
- Knowledge gap: Non-technical users may not know which parts of a report are relevant for regulatory disclosures, increasing confusion and inconsistency.

The AI-powered Input Assistant delivers the following value:
- Time-saving: Eliminates the need for users to manually extract and enter data from investigation reports.
- Accuracy: Reduces human error in interpreting and inputting technical findings.
- Accessibility: Enables non-technical staff or teams without cybersecurity expertise to complete regulatory reporting correctly.

3.2 Frontend Component

The Front-end component is the user's initial point of interaction with the platform. It offers five key functionalities to the user.
- Settings and Administration: Directs users to the page for managing login credentials and user settings, including registering new users and updating profile information.
- Textual Guide: Directs users to a webpage containing up-to-date information on the latest cybersecurity regulations and industry standards.
- Report Creation: Directs users to the Input component for creating cybersecurity incident reports.
- Report Viewer/Editor: Directs users to the Report component where they can review, edit, and manage the reports created.
- LLM Advisor: Directs users to the Advice component, where they can seek advice on the reports that were created.

3.2.1 Development Methodology

Interpretation of these requirements guided the platform's layout and user flow. The Report Creation page mirrors the sequential structure used in regulatory forms, with guided prompts and dropdown menus that standardise user inputs while reducing ambiguity. The design also considers the full reporting lifecycle: a Report Viewer/Editor page enables users to track status and outcomes, while the Textual Guide provides concise, sector-specific summaries of applicable reporting duties etc.

For the technical side, we decided to implement the following development tools after surveying different methods.
- Node JS [34]: Node is a cross-platform JavaScript runtime environment that lets developers create web-based applications. The key benefit of using Node.js is its speed and responsiveness. It facilitates the development of dynamic applications by enabling the easy handling of multiple concurrent requests.
- NPM (Node Package Manager) [35]: NPM centralizes the handling of application dependencies to make installation and version management simple.
- React JS [36]: React is a JavaScript library used for building user interfaces. The main advantage of React is its modular-based architecture, which makes it easier to manage and maintain complex applications.

3.2.2 User-Centric Benefits

When developing the frontend, we prioritized delivering the following values:
- To design a Responsive Web with seamless compatibility across diverse devices and screen sizes
- To make intuitive navigation without confusion
- To consider user needs and preferences
- To use a modular approach, considering further updates

By applying the aforementioned development methodologies, we are confident in delivering a frontend that is both highly functional and user-centric.

3.3 Backend Component

The Back-end component handles the server-side processes of the platform, including managing user credentials, settings, and data.

3.3.1 Development Methodology

When developing the Back-end component, we focused on three key factors: a short learning curve, robust community support, and seamless database integration. Given the project's four-month timeline, we need a framework that facilitates quick onboarding. A larger developer community offers abundant resources for troubleshooting and support. Additionally, effective database integration streamlines data search and retrieval processes, thereby enhancing the overall performance of the platform. After evaluating various development frameworks, we selected Django [37] based on its key advantages.

- Easy to Learn: Django is easy to learn and simple to code as it is built on Python.
- Extensive Community Support: Django is the most popular backend framework in 2025 [38] and has an extensive development community.
- ORM (Object-relational mapper): This in-built tool helps developers work with databases more easily. It allows automatic transfer of data from a database into objects in the code. It simplifies the database integration process as developers don't need to worry about the details of how the data is stored or retrieved.

3.3.2 User-Centric Benefits

- Apart from Django's advantages to the developer, it also offers significant value to the users.
    - Cross-Platform: Django is a versatile framework that allows users to access the platform using different systems, including Windows, Mac, Linux, iOS, Android, and others.
    - Security: Cybersecurity incident reports contain confidential information, which needs to be secured. Django has robust built-in protections against common security threats, such as cross-site scripting (XSS), cross-site request forgery (CSRF), clickjacking, and SQL injection. Moreover, Django's frequent updates and security patch releases ensure ongoing protection against emerging risks.

3.4 Report Component

The Report components include the application of the Reporting Path Recommendation Model, the Report Viewer/Editor and the Regulator Notification Function.

- Reporting Path Recommendation Model: This model is developed to assess the information in the reports created (based on the data inputted via data input form in the Input component) to determine the regulator(s) that the incident may be reported to. The result of the recommendation of regulator(s) and the associated information such as the reporting timeline will appear on the page of the report generated following the user deciding to create report after the Input component.
- Report Viewer/Editor: This function enables users to manage the reports created (which may or may be not be notified regulators) on one hand, and allows regulators to access the reports notified by users which have decided to make the reports on the other hand.
- Regulator Notification Function: This function enables users that have decided to notify the regulators of the selected reports to send notifications to the relevant regulators, prompting them to view the notified reports.

3.4.1 Development Methodology

- Reporting Path Recommendation Model

  When designing the model, we utilised the results from our review of the laws, regulations and guidelines conducted in designing the input fields. The regulatory and policy documents were systematically reviewed and categorized based on industry sector, the authority of the relevant regulators, threat type, and the nature of the reporting obligation (e.g., mandatory, advisory, or guideline-based). From each document, key compliance elements were extracted, including the types of incidents that trigger reporting, applicable reporting timeframes (e.g., 12 hours for serious incidents and 48 hours for less critical events), designated regulatory authorities, and the specific data fields required in reporting templates. These elements were then encoded into a structured regulatory logic map. This logic forms the backbone of the platform's rule-based engine, allowing it to interpret user inputs and make accurate, regulation-aligned reporting recommendations. An illustration of the rule-based engine is shown in Figure 2.

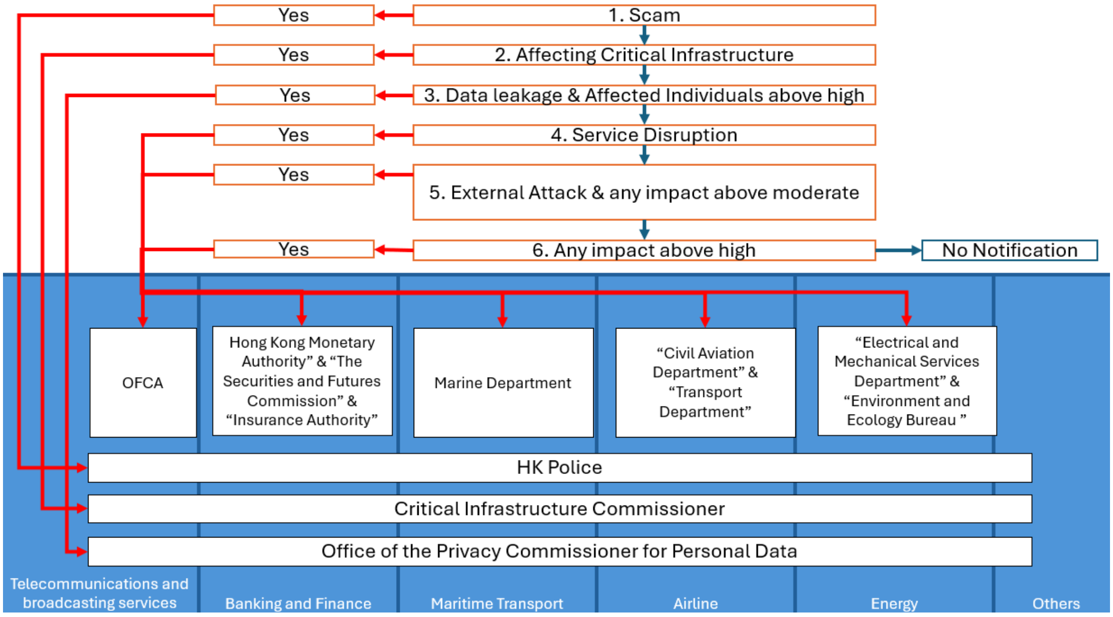

Yes
1. Scam
Yes
2. Affecting Critical Infrastructure
Yes
3. Data leakage & Affected Individuals above high
Yes
4. Service Disruption
Yes
5. External Attack & any impact above moderate
Yes
6. Any impact above high
No Notification
OFCA
Hong Kong Monetary Authority" & "The Securities and Futures Commission" & "Insurance Authority"
Marine Department
"Civil Aviation Department" & "Transport Department"
"Electrical and Mechanical Services Department" & "Environment and Ecology Bureau "
HK Police
Critical Infrastructure Commissioner
Office of the Privacy Commissioner for Personal Data
Telecommunications and broadcasting services
Banking and Finance
Maritime Transport
Airline
Energy
Others

Figure 2: The rule-based engine for the Reporting Path Recommendation Model.

- The Report Viewer/Editor allows users to manage reports, whether or not they have been notified to regulators. Users can edit, review, and organize reports with ease. Regulators, on the other hand, can access and review reports that have been officially notified, ensuring transparency and timely oversight.

- The Regulator Notification Function enables users to notify relevant regulators of selected reports. Notifications are sent directly to regulatory bodies, prompting them to review the flagged incidents efficiently and in compliance with reporting requirements.

The system was designed to ensure efficiency, scalability, and accessibility while maintaining alignment with regulatory standards. This approach simplifies the reporting process, enhances user experience, and strengthens collaboration with regulators.

3.4.2 User-Centric Benefits

The system was designed with a strong focus on user-centric benefits, ensuring that users can effectively navigate complex reporting processes while meeting regulatory requirements. Key benefits include:

- Prompt Regulator Identification
  As cybersecurity incidents become increasingly complex, organizations must navigate multiple reporting requirements issued by various regulatory bodies. Due to the vast number of regulators in Hong Kong and the varying nature of incidents and the growing number of guidelines and obligations,

there is a higher risk of human error, such as overlooking a mandatory report or failing to consider all relevant regulators. In some cases, smaller stakeholders, such as SMEs, may be unintentionally excluded from the reporting process altogether. This complexity makes it easy to miss critical notifications, especially without a structured and centralized decision-making tool. As a result, users may encounter the following challenges when deciding to notify regulators:

- Fragmented and overlapping reporting requirements from multiple authorities are difficult to track manually.
- High risk of human error, such as misunderstanding reporting thresholds.
- Lack of awareness, especially among SMEs or non-specialist staff, about which regulators are relevant in different incident scenarios.
- Time-consuming manual cross-checking of regulatory rules, which delays the response process.

The Reporting Path Recommendation Model in the Report component addresses these challenges by providing the following value:

- Ensures complete and timely regulatory reporting by automatically identifying which regulators need to be notified based on incident details.
- Reduces compliance risk by guiding users through complex and evolving regulatory obligations.
- Simplifies decision-making, especially for users without legal or regulatory expertise.

- Control over Reports Created

  Users have full control over their reports through the Report Viewer/Editor, enabling easy edits, reviews, and organization. This fosters confidence in managing reports and provides flexibility in decision-making.

- Simplified Reporting Process

  The Report component provides an intuitive interface that guides users through selecting the applicable regulators and the decision to notify such regulators. This reduces complexity and ensures that users, regardless of their expertise, can complete notification to the right regulators accurately and efficiently. By streamlining workflows, automating key decisions, and reducing manual effort, the system saves users significant time. Notifications to regulators are also simplified, ensuring faster communication and response times.

These user-centric benefits ensure that the system not only simplifies compliance but also empowers users to navigate complex regulatory requirements with ease and confidence.

3.5 Advice Component

The Advice component will analyze cybersecurity incident reports created by the users over time and provide advice based on the latest government guidelines, industry standards and the trend demonstrated by reports created on the platform. The advice will include explanations of incident report findings, matching incidents with relevant guidelines, and offering recommendations for mitigation and prevention.

3.5.1 Development Methodology

In designing the Advice component, we concentrated on three essential functionalities. First, we will incorporate an LLM that can process user inputs, perform analyses, and generate advice. To ensure the LLM accesses relevant information, we will develop an information retrieval function that efficiently retrieves report and guideline data. Finally, we will implement guardrails to prevent hallucinations and ensure appropriate advice.

- DeepSeek-R1 [39]: It is chosen for its powerful reasoning capabilities and low computational costs.

- GraphRAG [43]: It is chosen for its strong performance in multi-hop reasoning, which allows it to connect different pieces of information effectively. This is important for linking details in incident reports to relevant guideline data. Unlike traditional RAG that uses vector databases to find similar text, it enhances data retrieval by generating knowledge graphs (KGs). Knowledge graphs organize and connect related information based on their relationships, enabling more accurate and context-aware retrieval.
- NeMo Guardrails [44]: It is chosen to ensure the safety and reliability of the advice provided. It offers pre-built tools to implement guardrails that help prevent inappropriate or harmful output by setting rules and constraints to prevent it from engaging in discussions on unwanted topics.

3.5.2 User-Centric Benefits

Our goal in developing the Advice component is to deliver the following benefits.

- Offer recommendations tailored to address specific cybersecurity incidents.
- Provide clear and actionable advice that matches government guidelines and industry standards.
- Provide safe and reliable advice through robust guardrails that prevent inappropriate outputs.

We believe that our proposed development methodology will effectively achieve these goals.

4. Progress to Date

During the current project period (March 1, 2025 – June 1, 2025), we conducted extensive background research and made significant advancements in developing the core components of the platform. Our progress is discussed in detail in this section.

4.1 Background Research

As described in the literature review part of this report, extensive research has been conducted into existing regulatory frameworks and cybersecurity reporting guidelines in both internally (in Hong Kong) and externally. This includes analyzing incident reporting requirements across industries and identifying gaps in current reporting processes.

4.2 Core Component Development

Table 4 provides a summary of the status and overview of the core components under development.

Core Component	Status	Overview
Input	80% complete	▪ Development tools survey and selection (completed) ▪ Online form development and input field selection (completed) ▪ Severity Classification Model development (completed) ▪ Input assistant development (50% complete)
Frontend	70% complete	▪ Development tools survey and selection (completed) ▪ UI/UX design and development (90% complete) ▪ Web development: o Settings and administration page (completed) o Textual guide page (completed) o Report creator page (60% complete) o Report viewer/editor page (60% complete) o LLM advisor page (0% complete)
Backend	80% complete	▪ Development tools survey and selection (completed) ▪ Architecture design (completed) ▪ Backend development: o Settings and administration (completed) o Incident report database (90% complete) o Knowledge graph (0% complete)
Report	60% complete	▪ Development tools survey and selection (completed) ▪ Reporting Path Recommendation Model development (50% complete) ▪ Report viewer/editor development (90% complete) ▪ Regulator notification function development (30% complete)
Advice	20% complete	▪ Development tools survey and selection (completed) ▪ LLM advisor development (0%)

Table 4: Overview of the core components development statuses

4.3 Screenshots of the Platform in Development

- Login Page (Figure 3): It is the entry point to the platform where users authenticate their credentials by inputting their username and password. There are two different entry points, one for the normal users and another for the different regulators.

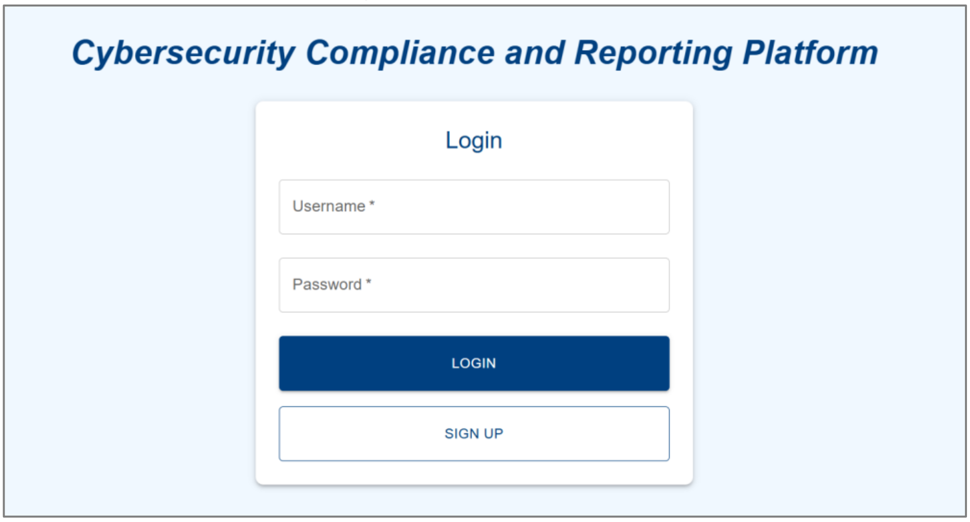
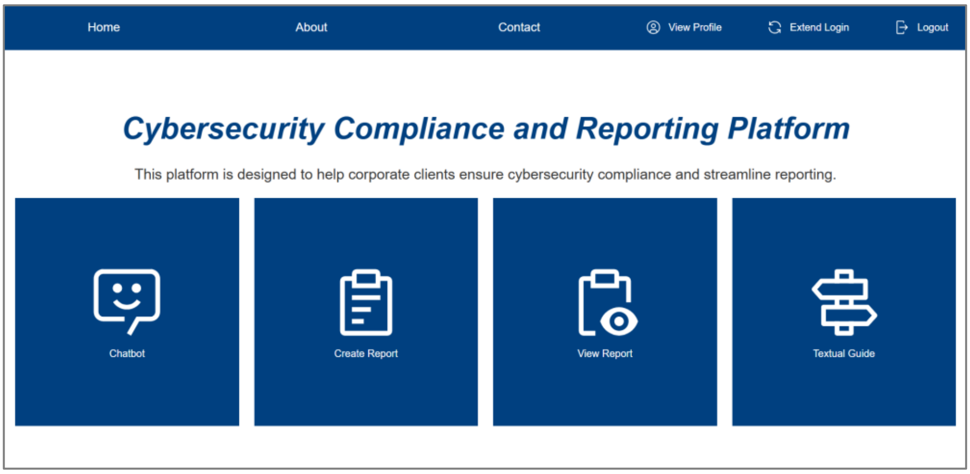

Cybersecurity Compliance and Reporting Platform

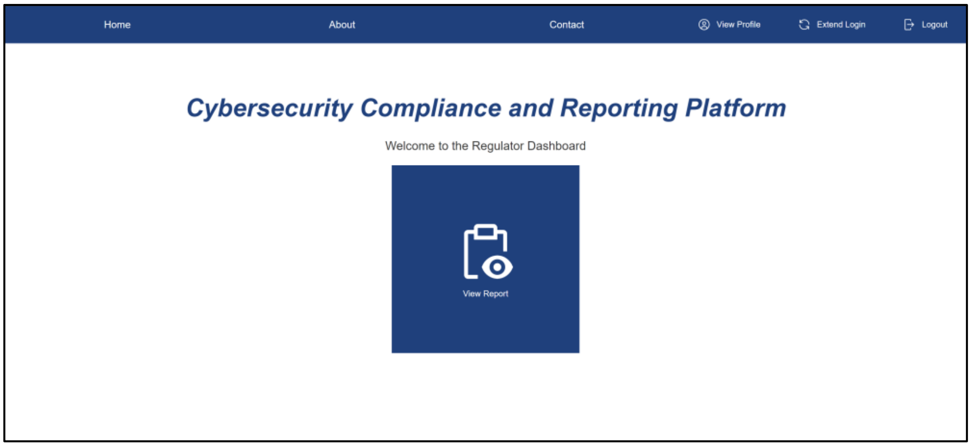
Figure 3: The login page

- Home Page (Figure 4 - Figure 5): It is the landing point after the user logs into the platform. It serves as the central hub for quick navigation and access to the key functionalities.


Cybersecurity Compliance and Reporting Platform

This platform is designed to help corporate clients ensure cybersecurity compliance and streamline reporting.

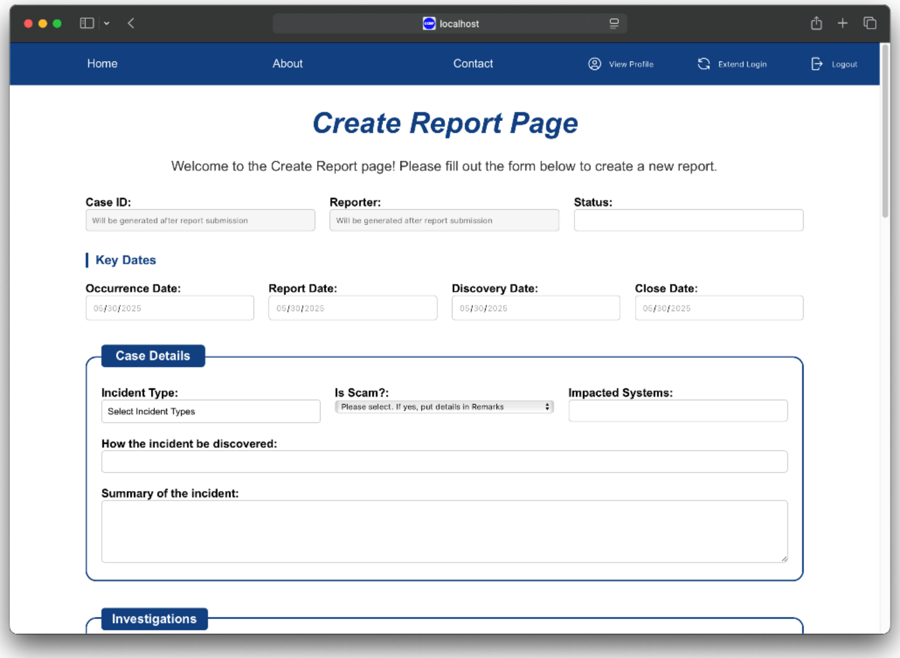
Figure 4: The home page for normal users


Cybersecurity Compliance and Reporting Platform

Welcome to the Regulator Dashboard

Figure 5: The home page for regulators

- Online Form (Figure 6): It allows users to input detailed information about a cybersecurity incident. The input information will be stored in the incident report database and used by the severity classification and notification decision models.



Create Report Page

Welcome to the Create Report page! Please fill out the form below to create a new report.

Case ID: Will be generated after report submission

Reporter: Will be generated after report submission

Status:

Key Dates

Occurrence Date: 05/30/2025

Report Date: 05/30/2025

Discovery Date: 05/30/2025

Close Date: 05/30/2025

Case Details

Incident Type: Select Incident Types

Is Scam?: Please select. If yes, put details in Remarks

Impacted Systems:

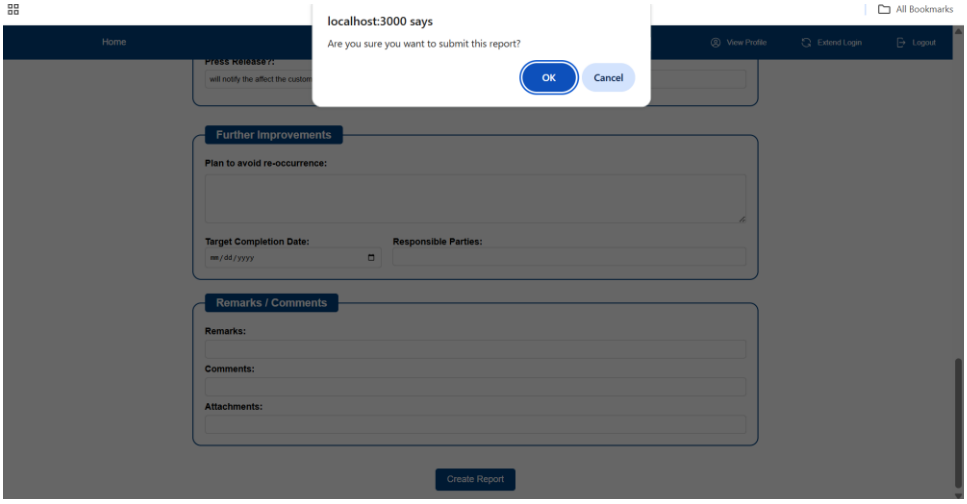
How the incident be discovered:

Summary of the incident:

Investigations

Figure 6: Online Form

- Severity Analysis (Figure 7 – Figure 8): The framework evaluates incidents against reporting obligations and severity thresholds. Users input information on predefined scenarios (e.g., scams, service disruptions, or critical infrastructure attacks). The system will assess based on the severity rating from five key aspects, including financial and operational impact, and recommends whether an incident is severe enough that warrant reporting.



localhost:3000 says

Are you sure you want to submit this report?

Press Release?: will notify the affect the custom

Further Improvements

Plan to avoid re-occurrence:

Target Completion Date: mm/dd/yyyy

Responsible Parties:

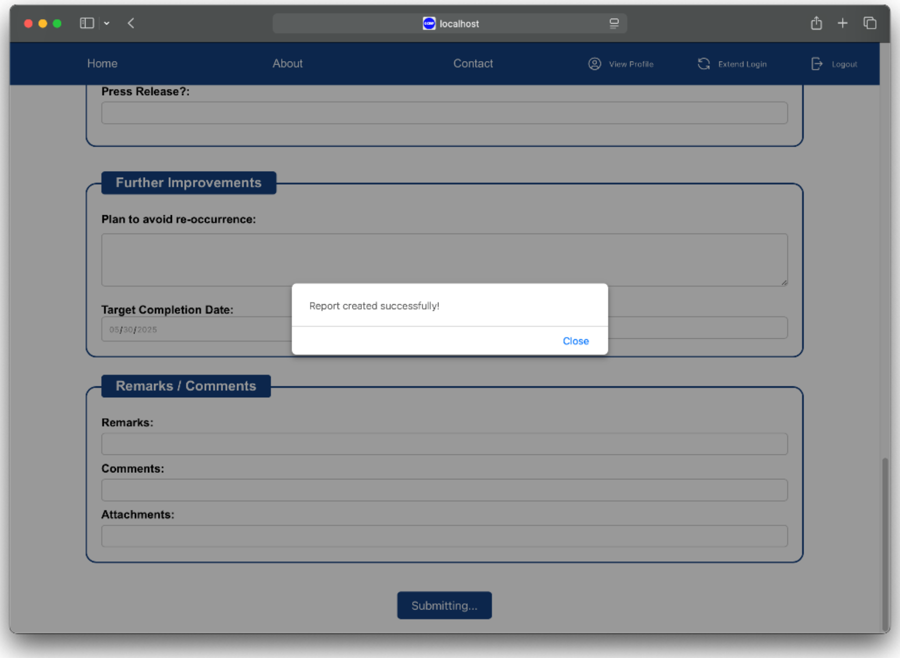
Remarks / Comments

Remarks:

Comments:

Attachments:

Figure 7: Create report



Press Release?:

Plan to avoid re-occurrence:

Target Completion Date: 05/30/2025

Report created successfully!

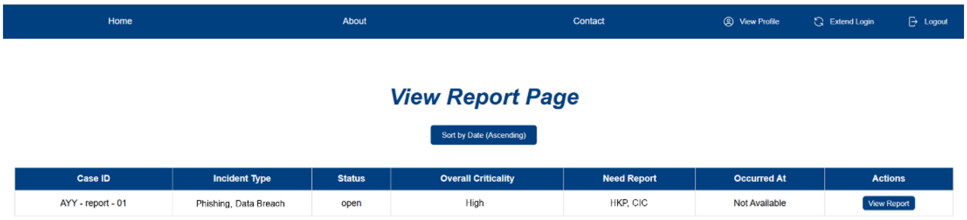
Remarks:

Comments:

Attachments:

Figure 8: Create report successfully with results in pop-up message

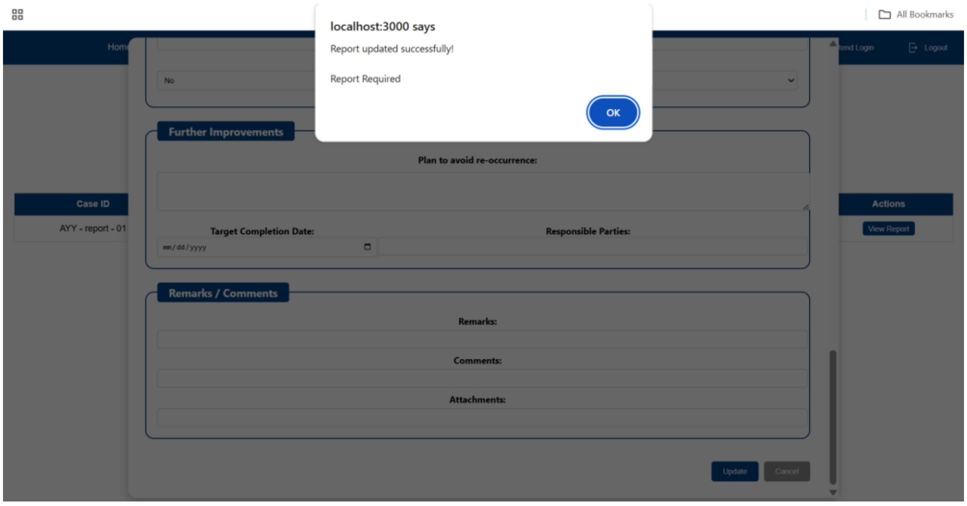
- View Report (Figure 9 – Figure 10): Users can view reports created, update the details and track their status.



View Report Page

Case ID	Incident Type	Status	Overall Criticality	Need Report	Occurred At	Actions
AYY - report - 01	Phishing, Data Breach	open	High	HKP, CIC	Not Available	View Report

Figure 9: Case list update view report page



localhost:3000 says

Report updated successfully!

Report Required

No

AYY - report - 01

Plan to avoid re-occurrence:

Target Completion Date: mm/dd/yyyy

Responsible Parties:

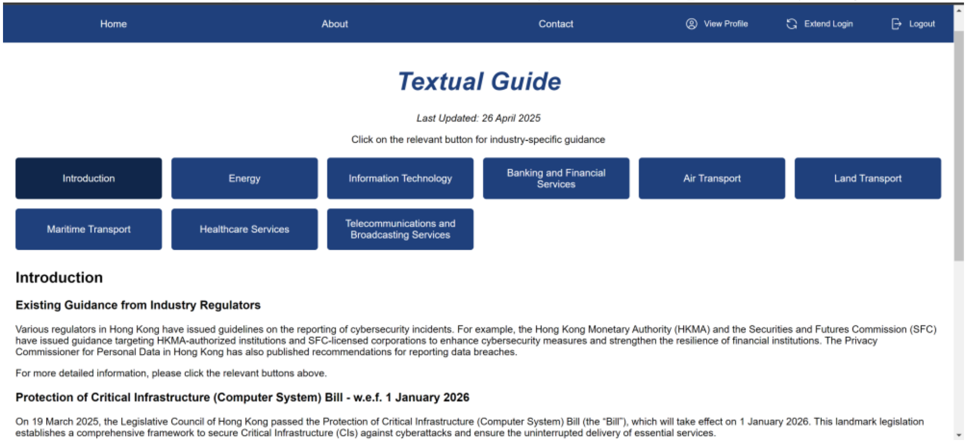
Remarks:

Comments:

Attachments:

Figure 10: Update case details

- Textual Guide (Figure 11): The textual guide is designed to provide industry-specific cybersecurity guidance in a user-friendly manner. A button grid serves as the primary navigation method. When a button is clicked, the corresponding section is revealed while hiding all other sections. This reduces clutter and focuses the user's attention on the relevant content. Other than the "Introduction" button, each button corresponds to a specific industry which we specifically target for our platform.



Textual Guide

Last Updated: 26 April 2025

Click on the relevant button for industry-specific guidance

Introduction

Existing Guidance from Industry Regulators

Various regulators in Hong Kong have issued guidelines on the reporting of cybersecurity incidents. For example, the Hong Kong Monetary Authority (HKMA) and the Securities and Futures Commission (SFC) have issued guidance targeting HKMA-authorized institutions and SFC-licensed corporations to enhance cybersecurity measures and strengthen the resilience of financial institutions. The Privacy Commissioner for Personal Data in Hong Kong has also published recommendations for reporting data breaches.

For more detailed information, please click the relevant buttons above.

Protection of Critical Infrastructure (Computer System) Bill - w.e.f. 1 January 2026

On 19 March 2025, the Legislative Council of Hong Kong passed the Protection of Critical Infrastructure (Computer System) Bill (the "Bill"), which will take effect on 1 January 2026. This landmark legislation establishes a comprehensive framework to secure Critical Infrastructure (CIs) against cyberattacks and ensure the uninterrupted delivery of essential services.

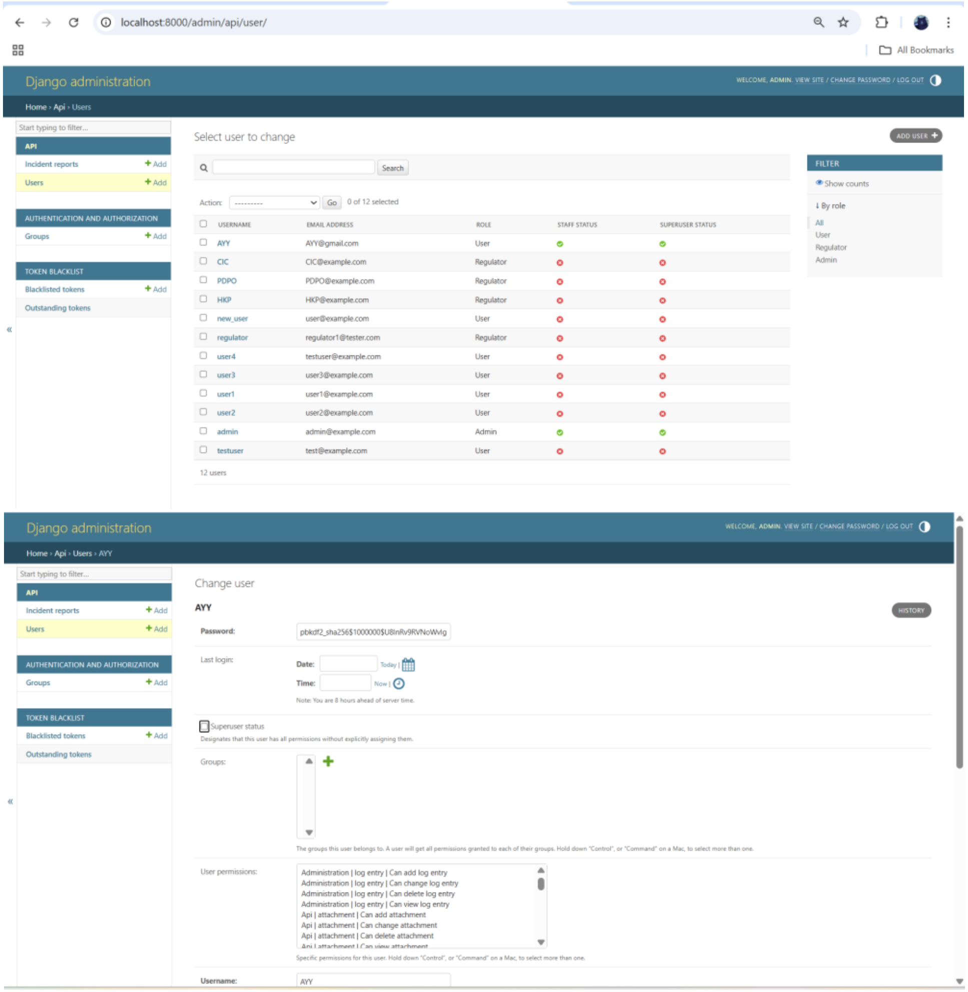
Figure 11: Textual guide

Each industry section includes:

- Regulators: A list of overseeing bodies specific to the sector.
- Key Risks: A summary of the most critical cybersecurity threats the industry faces (e.g., ransomware, data breaches, operational disruptions).
- Obligation to Report: Clear instructions on reporting requirements, including mandatory and discretionary actions.
- Materiality Determination: Guidelines to help organizations assess whether an incident is significant enough to report.
- Customer Notification: Information on whether customer notification is mandatory and the best practices around it.
- Recommendations: Practical advice tailored to the industry's specific risks, such as conducting audits, aligning with international standards, and implementing robust cybersecurity measures.

The guide includes hyperlinks to external documents and guidelines, such as those issued by the Office of the Privacy Commissioner for Personal Data (PCPD), the Hong Kong Monetary Authority (HKMA), and other regulatory bodies. This provides users with quick access to additional resources for deeper understanding.

- Settings and Administration Page (Figure 12): It allows admin users to modify settings and login credentials.



Select user to change

Action: ---------- 0 of 12 selected

□	USERNAME	EMAIL ADDRESS	ROLE	STAFF STATUS	SUPERUSER STATUS
□	AYY	AYY@gmail.com	User	✓	✓
□	CIC	CIC@example.com	Regulator	☒	☒
□	PDPO	PDPO@example.com	Regulator	☒	☒
□	HKP	HKP@example.com	Regulator	☒	☒
□	new_user	user@example.com	User	☒	☒
□	regulator	regulator1@tester.com	Regulator	☒	☒
□	user4	testuser@example.com	User	☒	☒
□	user3	user3@example.com	User	☒	☒
□	user1	user1@example.com	User	☒	☒
□	user2	user2@example.com	User	☒	☒
□	admin	admin@example.com	Admin	✓	✓
□	testuser	test@example.com	User	☒	☒

12 users

Change user

AYY

Password: pbkdf2_sha256$1000000$U8InRv9RVNoWvIg

Last login: Date: Time:

Note: You are 8 hours ahead of server time.

□ Superuser status

Designates that this user has all permissions without explicitly assigning them.

Groups:

The groups this user belongs to. A user will get all permissions granted to each of their groups. Hold down "Control", or "Command" on a Mac, to select more than one.

User permissions:

Administration | log entry | Can add log entry
Administration | log entry | Can change log entry
Administration | log entry | Can delete log entry
Administration | log entry | Can view log entry
Api | attachment | Can add attachment
Api | attachment | Can change attachment
Api | attachment | Can delete attachment
Api | attachment | Can view attachment

Specific permissions for this user. Hold down "Control", or "Command" on a Mac, to select more than one.

Username: AYY

Figure 12: Administration page

5. Challenges and Solutions

When developing our platform, we encountered various challenges and proposed solutions to address them.

- Subjectivity of Severity Classification

One major challenge we faced was the inherent subjectivity in classifying the severity of cybersecurity incidents. Since severity assessments often rely on individual interpretation and organizational context, inconsistencies can arise across teams and stakeholders. This variability can hinder timely decision-making, regulatory compliance, and effective incident response.

Solution: To address this, we conducted in-depth research and analysis to align our framework with both Hong Kong government guidelines and globally recognized industry standards. We developed a structured severity classification model that incorporates key dimensions such as operational disruption, the number of affected individuals, financial consequences, and the extent of data leakage. This standardized approach ensures consistent, transparent, and justifiable severity ratings across different incidents and user groups.

- Diverse Reporting Requirements Across Different Regulators

In Hong Kong, there are numerous industries, each with its own unique cybersecurity reporting requirements and regulatory frameworks. It would not be feasible to cover all industries comprehensively within the scope of this project.

Solution: We have decided to narrow our focus to eight critical industries, namely Energy, Information Technology, Banking and Financial Services, Air Transport, Land Transport, Maritime Transport, Healthcare Services and Telecommunications and Broadcasting Services. This focused approach ensures that we can provide targeted and effective solutions for these key sectors while addressing the specific expectations of the relevant regulatory bodies.

- Handling Multiple Selection as Input

The incident reporting form required support for multiple selection fields such as *Incident Type* and *Report To*, which naturally return a list of strings in the frontend. However, this created compatibility issues with our backend, as SQLite—our chosen database—does not natively support list-type fields. Additionally, Django's ManyToManyField requires a list of related model IDs rather than plain strings, further complicating data handling between the frontend and backend.

Solution: To resolve this mismatch, we revised our backend data model by replacing the ManyToManyField with a CharField, allowing us to store concatenated string values. The frontend logic was adjusted to join the selected options into a single, delimiter-separated string before sending the data to the backend. This approach ensured compatibility with SQLite while preserving the ability to represent multiple selections effectively.

- Safety and Reliability LLM

Ensuring the safety and reliability of LLM is challenging. Key risks include the generation of unsafe content that could offend users, as well as the potential for producing factually incorrect information, which undermines trust in these systems. Additionally, LLMs may struggle with context, leading to the generation of content that is irrelevant to the information provided.

Solution: To address these challenges, we propose implementing guardrail tools like NeMo. Guardrail tools allow developers to establish rules that constrain LLM outputs, thereby significantly reducing the likelihood of generating inappropriate content. This approach enhances both the safety and reliability of the LLM.

6. Next Steps

The updated project schedule with the milestones is shown below. The project is progressing smoothly, with the planned tasks completed within their designated timeframes. Moving forward, we will continue to enhance the platform's functionalities while exploring the outstanding functionalities, including the AI-based Input Assistant and the LLM advisor.

An updated milestone plan is as follows:

	March	**April**	**May**	**June**	**July**
Detailed Project Proposal (10 March)	■				
1st Milestone (7 April) • Develop a website with role-based access control (sign-up, login, logout, etc.) • Implement functionality for creating incident response reports.	■	■			
Project Progress Update 1 (7 April) • Presentation on the 1st Milestone		■			
Project Progress Update 2 (10 May) • Working towards the 2nd Milestone in relation to further enhancing functionality of website and report generation functions, and evaluation of pre-reporting evaluation frameworks.			■		
2nd Milestone (1 June) • Further enhancing functionality of website and report generation functions. • Evaluation of pre-reporting evaluation frameworks (i.e. Severity Classification Model and Reporting Path Recommendation Model). • Exploring practicality of additional features including Chatbot and IPFS.				■	
Interim Report and Presentation (1 June)				■	
Project Progress Update 3 (16 June)				▨	
3rd Milestone (7 July) • Transition from Proof of Concept (POC) to Production. • Finalize platform deployment and conduct user acceptance testing (UAT)					▨
Project Progress Update 4 (7 July)					▨
Project Report (18 July)					▨
Oral Examination (End of July)					▨

6.1. Next Steps for Testing and Evaluating the Platform

Upon completion of the platform, we intend to evaluate its performance using the following performance metrics and evaluation data.

6.1.1 Rule-based Model: Severity Classification

We will evaluate the effectiveness of the rule-based severity classification model using the F1 score. This metric is preferred for classification models as it offers a balanced assessment of performance by considering both Precision and Recall equally.

The evaluation dataset will be the European Repository of Cyber Incidents (EuRepoC) Global Dataset [45]. This dataset includes 3,416 global cyber incidents spanning between January 1, 2000, to December 31, 2024; each assigned an annotated severity score.

6.1.2 LLM Models with RAG: AI-powered Input Assistant and LLM Advisor

We will evaluate the AI-powered Input Assistant and LLM Advisor using the RAGAs (Retrieval-Augmented Generation Assessment) framework, which facilitates effective measurement of both retrieval (context precision and recall) and generative (faithfulness and answer relevancy) performance [46].

- Context Precision: Assesses the signal-to-noise ratio of retrieved context based on the question and context (incident report and government guidelines).
- Context Recall: Evaluates whether all relevant information was retrieved, relying on human-annotated ground truth.
- Faithfulness: Measures the factual accuracy of the generated answer by comparing correct statements to the total in the answer.
- Answer Relevancy: Determines how relevant the answer is to the question. For instance, answering “France is in western Europe” to “Where is France and what is its capital?” would score low for relevancy.

All metrics are normalized to a range of [0, 1], where higher values signify improved performance.

The evaluation data will be a set of cybersecurity incident scenarios we collected. These scenarios are from real-world incidents documented in PCPD enforcement reports [25], HKCERT threat alerts [26], HKMA notices [27], and detailed news coverage from trusted media outlets such as SCMP and HK01. They offer insight into commonly reported incident types in Hong Kong, regulatory responses, and the operational or reputational consequences faced by affected entities.

Given limited access to detailed real-world case data, we also developed simulated incidents based on structured examples from international repositories, including the European Repository of Cyber Incidents [28] and the National Privacy and Cyber Incident Repository [29] and several curated datasets from academic and open-source communities, such as York University Cybersecurity Datasets [30], Awesome Cybersecurity Datasets [31], CISSM Cyber Events Database [32] and Verizon VERIS Community Database [33]. These external sources help to diversify the range of scenarios and severity levels available for platform testing.

7. Reference

[1] De Los Santos, S. (2016). The impact of an absent national cybersecurity attack reporting policy (Doctoral dissertation, Colorado Technical University). Colorado Technical University.

[2] U.S. National Institute of Standards and Technology, “Computer Security Incident Handling Guide,” NIST Special Publication 800-61 Revision 2, Jul. 2012. [Online]. Available: https://csrc.nist.gov/pubs/sp/800/61/r2/final

[3] Schmitz-Berndt, S., Bellekens, X., Hindy, H., Onwubiko, C., Erola, A., Rege, A., Jaatun, M. G., & Rosati, P. (2023). Refining the mandatory cybersecurity incident reporting under the NIS directive 2.0: Event types and reporting processes. In Proceedings of the International Conference on Cybersecurity and Resilience (pp. 343-351). Springer.

[4] European Union, Directive (EU) 2022/2555 of the European Parliament and of the Council of 14 December 2022 on measures for a high common level of cybersecurity across the Union, NIS 2 Directive. [Online]. Available: https://eur-lex.europa.eu/eli/dir/2022/2555/oj

[5] Briggs, P., Jeske, D., Coventry, L., & Tryfonas, T. (2017). The design of messages to improve cybersecurity incident reporting. In Proceedings of the International Conference on Cybersecurity and Resilience (pp. 3-13). Springer International Publishing.

[6] Muthuswamy, V. V., & Esakki, S. (2024). Impact of cybersecurity and AI-related factors on incident reporting suspicious behaviour and employees’ stress: Moderating role of cybersecurity training. International Journal of Cyber Criminology, 18(1), 83-107.

[7] Government of Singapore, Cybersecurity Act 2018, No. 9 of 2018. [Online]. Available: https://sso.agc.gov.sg/Acts-Supp/9-2018/

[8] National center of Incident readiness and Strategy for Cybersecurity (NISC), Cybersecurity Policy of Japan. [Online]. Available: https://www.nisc.go.jp/eng/

[9] Korea Internet & Security Agency (KISA), Cyber Threat Response and Security Services. [Online]. Available: https://www.kisa.or.kr/eng/main.jsp

[10] Hong Kong Legislative Council, Bill document, Hong Kong Special Administrative Region, 2024. [Online]. Available: https://www.legco.gov.hk/yr2024/english/bills/b202412061.pdf

[11] S. Conard, Quantifying the severity of a cybersecurity incident for incident reporting, M.S. thesis, Massachusetts Institute of Technology, Cambridge, MA, 2024. [Online]. Available: https://dspace.mit.edu/handle/1721.1/157124

[12] F. Doshi-Velez and B. Kim, “Towards a rigorous science of interpretable machine learning,” arXiv preprint arXiv:1702.08608, Feb. 2017. [Online]. Available: https://arxiv.org/abs/1702.08608

[13] C. Rudin, “Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead,” Nature Machine Intelligence, vol. 1, no. 5, pp. 206–215, 2019. [Online]. Available: https://doi.org/10.1038/s42256-019-0048-x

[14] E. H. Diallo, R. Abdallah, M. Dib, and O. Dib, “Decentralized incident reporting: Mobilizing urban communities with blockchain,” Smart Cities, vol. 7, no. 4, pp. 2283–2317, 2024. [Online]. Available: https://doi.org/10.3390/SMARTCITIES7040090

[15] Office of the Privacy Commissioner for Personal Data (PCPD), Data Breach Notification Form. [Online]. Available: https://www.pcpd.org.hk/english/publications/files/databreach_form_e.pdf

[16] Office of the Privacy Commissioner for Personal Data (PCPD), Guidance Note on Data Breach Notification, Hong Kong, 2021. [Online]. Available: https://www.pcpd.org.hk/english/resources_centre/publications/files/guidance_note_dbn_e.pdf

[17] Hong Kong Monetary Authority (HKMA), Supervisory Approach on Cyber Risk Management, Nov. 2024. [Online]. Available: https://www.hkma.gov.hk/media/eng/doc/key-information/guidelines-and-circular/2024/20241129e2a1.pdf

[18] Securities and Futures Commission (SFC), Guidelines for Reducing and Mitigating Hacking Risks Associated with Internet Trading, Hong Kong, 2017. [Online]. Available: https://www.sfc.hk/en/Rules-and-standards/Codes-and-guidelines/Guidelines/Guidelines-for-Reducing-and-Mitigating-Hacking-Risks-Associated-with-Internet-Trading

[19] Insurance Authority (IA), Guideline on Cybersecurity (GL20), Hong Kong, 2020. [Online]. Available: https://www.ia.org.hk/en/legislative_framework/files/GL20_Eng.pdf

[20] Office of the Communications Authority (OFCA), Security Guidelines for Next Generation Networks, Hong Kong, 2022. [Online]. Available: https://www.ofca.gov.hk/filemanager/ofca/en/content_757/traac4_2022.pdf

[21] Office of the Government Chief Information Officer (OGCIO), Practice Guide for Information Security Incident Handling, Hong Kong, Feb. 2025. [Online]. Available: https://www.govcert.gov.hk/doc/PG%20for%20ISIH_EN.pdf

[22] Office of the Government Chief Information Officer (OGCIO), Practice Guide for IT Security Risk Management, Hong Kong, Jul. 2024. [Online]. Available: https://www.govcert.gov.hk/doc/PG%20for%20IT%20Security%20Risk%20Management_EN.pdf

[23] International Organization for Standardization (ISO), ISO/IEC 27000 family – Information security management systems, [Online]. Available: https://www.iso.org/standard/iso-iec-27000-family

[24] Financial Stability Board (FSB), Format for Incident Reporting Exchange (FIRE) – Final Report, Apr. 2025. [Online]. Available: https://www.fsb.org/2025/04/format-for-incident-reporting-exchange-fire-final-report/

[25] Office of the Privacy Commissioner for Personal Data (PCPD), Enforcement Reports, Hong Kong. [Online]. Available: https://www.pcpd.org.hk/english/enforcement_reports/report.html

[26] Government Computer Emergency Response Team Hong Kong (GovCERT.HK), Official Website, [Online]. Available: https://www.govcert.gov.hk/en/index.html

[27] Hong Kong Monetary Authority (HKMA), Official Website, [Online]. Available: https://www.hkma.gov.hk/eng

[28] Stiftung Wissenschaft und Politik, European Repository of Cyber Incidents (EuRepoC). [Online]. Available: https://www.europoc-repository.com

[29] Kean University, National Privacy and Cyber Incident Repository (NPCIR). [Online]. Available: https://www.kean.edu/npcir

[30] York University, Cybersecurity Datasets (CDS) – BCCC UCS Technical Program, [Online]. Available: https://www.yorku.ca/research/bccc/ucs-technical/cybersecurity-datasets-cds/

[31] S. H. Ramos, Awesome Cybersecurity Datasets [GitHub Repository], [Online]. Available: https://github.com/shramos/Awesome-Cybersecurity-Datasets

[32] Center for International and Security Studies at Maryland (CISSM), Cyber Events Database, [Online]. Available: https://cissm.umd.edu/cyber-events-database

[33] Verizon RISK Team, Verizon Cybersecurity Data Breach Database (VCDB) [GitHub Repository], [Online]. Available: https://github.com/vz-risk/vcdb

[34] Node.js, "Run JavaScript Everywhere." [Online]. Available: https://nodejs.org/en

[35] Node.js, "An introduction to the npm package manager." [Online]. Available: https://nodejs.org/en/learn/getting-started/an-introduction-to-the-npm-package-manager

[36] Meta, "React." [Online]. Available: https://react.dev/

[37] Django Software Foundation, "Django." [Online]. Available: https://www.djangoproject.com/

[38] Statistics and Data, "Most Popular Backend Frameworks." [Online]. Available: https://statisticsanddata.org/data/most-popular-backend-frameworks-2012-2025/

[39] Deepseek, "Deepseek." [Online]. Available: https://www.deepseek.com

[40] Ollama, "Ollama." [Online]. Available: https://ollama.com

[41] LangChain, "Langchain." [Online]. Available: https://www.langchain.com

[42] Chroma, "Chromadb." [Online]. Available: https://www.trychroma.com

[43] Edge, D., Trinh, H., Cheng, N., Bradley, J., Chao, A., Mody, A., ... & Larson, J. (2024). From local to global: A graph rag approach to query-focused summarization. arXiv preprint arXiv:2404.16130.

[44] Rebedea, T., Dinu, R., Sreedhar, M., Parisien, C., & Cohen, J. (2023). Nemo guardrails: A toolkit for controllable and safe llm applications with programmable rails. arXiv preprint arXiv:2310.10501.

[45] European Repository of Cyber Incidents, “EuRepoC Database”. [Online]. Avaliable: https://eurepoc.eu/database/

[46] Es, S., James, J., Anke, L. E., & Schockaert, S. (2024, March). Ragas: Automated evaluation of retrieval augmented generation. In Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics: System Demonstrations (pp. 150-158).

[47] Office of the Privacy Commissioner for Personal Data, "Guidance on Data Breach Handling and the Giving of Breach Notifications," Apr. 2012. [Online]. Available: https://www.pcpd.org.hk/english/resources_centre/publications/files/guidance_note_dbn_e.pdf.

[48] Hong Kong Transport Department, *Road Traffic (Autonomous Vehicles) Regulation (Cap. 374AA)*, Hong Kong: Hong Kong e-Legislation, 2021. [Online]. Available: https://www.elegislation.gov.hk

[49] Office of the Privacy Commissioner for Personal Data, "Personal Data (Privacy) Ordinance and Electronic Health Record Sharing System: Points to Note for Healthcare Providers and Healthcare Professionals," Mar. 2016. [Online]. Available: https://www.pcpd.org.hk/english/electronic_health_record_sharing_system/files/eHRSS_Points_to_Notes_ENG.pdf

[50] Hong Kong Monetary Authority, "Incident Response and Management Procedures," Jun. 2010. [Online]. Available: https://brdr.hkma.gov.hk/eng/doc-ldg/docId/getPdf/20100622-1-EN/20100622-1-EN.pdf

[51] Securities and Futures Commission, "Circular to All Licensed Corporations - Alert for Ransomware Threats," May 15, 2017. [Online]. Available: https://apps.sfc.hk/edistributionWeb/api/circular/openFile?lang=EN&refNo=17EC26

[52] Hong Kong Monetary Authority, "Customer Data Protection," Oct. 14, 2014. [Online]. Available: https://brdr.hkma.gov.hk/eng/doc-ldg/docId/getPdf/20141014-1-EN/20141014-1-EN.pdf

[53] Securities and Futures Commission, Code of Conduct for Persons Licensed by or Registered with the Securities and Futures Commission, Oct. 2024. [Online]. Available: https://www.sfc.hk/-/media/EN/assets/components/codes/files-current/web/codes/code-of-conduct-for-persons-licensed-by-or-registered-with-the-securities-and-futures-commission/Code_of_conduct-Oct-2024_Eng-with-Bookmark-Final.pdf?rev=0d85942581714ea183634112e8e9d474

[54] Office of the Communications Authority, *Security Guidelines for Next Generation Networks*, Issue 4, Apr. 2023. [Online]. Available: https://www.coms-auth.hk/filemanager/statement/en/upload/618/gn012023e.pdf

[55] Legislative Council of Hong Kong, *Bill Document: b202412061.pdf*, Dec. 2024. [Online]. Available: https://www.legco.gov.hk/yr2024/english/bills/b202412061.pdf

[56] Office of the Privacy Commissioner for Personal Data, *Data Breach Notification Form*, Rev. Jun. 2023. [Online]. Available: https://www.pcpd.org.hk/english/enforcement/data_breach_notification/files/DBN_e.pdf

[57] Hong Kong Police Force, *e-Report Centre: Report Technology Crime and Deception*. [Online]. Available: https://www1.erc.police.gov.hk/cmiserc/EGIS-HK-Web_NEW_UI/ereport_details?report=TCAD&fontSize=100&vTimeoutReminder=3300000&vTimeoutVal=3600000&vTimeoutReminderVal=300000

[58] Transport Department, *Code of Practice for Trial and Pilot Use of Autonomous Vehicles*, Mar. 2024. [Online]. Available: https://www.td.gov.hk/filemanager/en/content_5198/CoP%20for%20AV%20Trial%20and%20Pilot%20Use%20March%202024_ENG.pdf