自動車セキュリティオペレーションセンタ向けの固有知識 DB を有するチャットアシストシステムの研究

川上 庄慶[1,*]　越出 和麿[1]　笹 晋也[1]　礒川 弘実[1]

概要：自動車に対するサイバー攻撃を検知・分析する自動車セキュリティオペレーションセンタ(VSOC)の構築と運用が進められているが，専門的な知識を持つオペレータが不足している．本研究では，オペレータの知識不足を補うことで業務遂行を支援するチャットアシストシステムを開発する．ユースケース分析に基づいた過不足の無い知識 DB 構築により回答精度を確保する手法と，キーワード頻出度により自動車メーカ固有の知識情報の回答への誤混入を検知する手法を開発した．オペレータの想定質問と回答を用いて回答精度と混入検知精度を評価し，有効性を確認した．

キーワード：VSOC，LLM，RAG

Study of Chat Assist System with Specialized Knowledge Database for Vehicle Security Operation Center

SHOKEI KAWAKAMI[1, *]　KAZUMA KOSHIDE[1]
SHINYA SASA[1]　HIROMI ISOKAWA[1]

Abstract: Vehicle Security Operation Centers manage cyber risks in vehicles, but they lack knowledgeable operators. In this study, we developed a chat assistant system that supports tasks by supplementing the lack of knowledge of operators. We developed a method to ensure accurate answers by constructing comprehensive knowledge databases based on use case analysis, as well as a method to detect the contamination of OEM-specific knowledge in answers based on document frequency of keywords. Through a set of experiments using operator's questions and answers, we evaluate truthfulness of answers and accuracy of detection and confirm the effectiveness.

Keywords: VSOC, LLM, RAG

1. はじめに

画像処理技術・通信技術の進展やコンピュータの小型化・低価格化などにより，近年は運転操作が自動化された自動運転車やインターネットへの接続機能を有するコネクテッドカーの研究開発が盛んであり，実用化も進んでいる[1]. これらの自動車は利便性が高い一方で, PC やサーバといった IT システムと同様にサイバー攻撃の標的となり得る．例えば，無線通信による不正な車両の遠隔操作や，社内通信の脆弱性を悪用した車両盗難，EV 充電システムの妨害など，自動車を対象とした攻撃事例も報告されている[2] [3] [4]．そのため，自動車におけるサイバーセキュリティが重要さを増している[5]．また，国連自動車基準調和世界フォーラム (WP29)による法規制[6]や，ISO による自動車のセイバーセキュリティ対策の標準化[7]が進めており，自動車メーカー (Original Equipment Manufacturer: OEM) やサプライヤは，それらが定める要件に対応しなければならない．

これらの背景から，自動車への攻撃の監視や分析を担う自動車セキュリティオペレーションセンタ (Vehicle Security Operation Center: VSOC) の構築と運用が進められている．VSOC で前述のオペレーションを実施するオペレータは, 自動車やセキュリティに関する知識が要求される．しかし，これらの知識は領域が異なるため人材確保が難しく，知識が不足した人物がオペレータとして業務に従事することが予想される．そのため，オペレータの知識不足を補い，業務遂行を支援する仕組みが必要である．この課題の解決方針として，(1) 自動車やセキュリティに関する教育の強化が考えられるが，学ばなければならない知識は多岐にわたるため，長期的な仕組みづくりが必要となる．一方で，(2) 質問と回答を通じた知識の補完が可能なチャットシステムによる支援では，短期的な解決が可能である．

本研究では，(2)を解決方針として，VSOC 向けのチャットアシストシステムを開発する．本システムは，複数の知識 DB を有する RAG (Retrieval-Augmented Generation) [8]システムである.「(a) 知識 DB を構成する文献の特定」と「(b) OEM 固有情報の回答への混入防止」をシステム開発に向けた課題として抽出し，(a)の課題に対しては，VSOC のセキ

1 (株)日立製作所
Hitachi Ltd.
* shokei.kawakami.tc@hitachi.com

© 2025 Information Processing Society of Japan

This paper is work in progress and not peer-reviewed.

ュリティ監視プロセスを 5W の観点で分析してユースケースを洗い出し，各ユースケースで回答生成に必要な文献を特定する．(b)の課題に対しては，OEM 固有のキーワードが出現する文献の頻度に基づいて回答を検査し，混入を検知する技術を開発する．オペレータの想定質問と回答から成るデータセットを作成して回答精度と混入検知精度を評価し，有効性を確認する．

2. 関連研究

近年，生成 AI と呼ばれる文章や画像などを生成する AI 技術が急速に発展し，大きな注目を集めている．自然言語処理の分野では，Transformer [9]と呼ばれる Self-Attention 機構を有する高性能なモデルの登場以降，大規模言語モデル (Large Language Model: LLM) の進歩は目覚ましく，自然言語による会話が人間と遜色ない水準に達している．特に，OpenAI 社が提供する GPT (Generative Pretrained Transformer) は，インターネット上の情報や書籍などを用いて事前に学習済みであり，一般的な知識を問う質問に高い精度で回答できる [10]．一方で，特定の領域に関する専門性の高い知識や組織外に開示されない知識は持たないため，回答できない．そこで，RAG と呼ばれる，文字列のベクトル化を通じてクエリに補足情報を付与する技術が注目されている．RAG では，Embedding model を用いて文献の文字列を固定長ベクトルに変換して知識 DB に格納しておく．回答生成時には，ベクトル化されたクエリと知識 DB 内のベクトルとの類似性を測ることで回答に必要な補足情報を特定し，クエリへ付与する．これにより，知識 DB の文献に沿った回答を得ることが可能となる．

3. 提案手法

3.1 ベースとする RAG システムの概要

ベースとするシステムの概要を図 1 に示す．本システムは，セキュリティや自動車の専門知識が格納された格汎用知識 DB の他に，OEM ごとに異なる固有知識が格納された固有知識 DB を有する RAG システムである．単一のシステムで複数の OEM に対応するために，オペレータの担当 OEM に応じて回答生成時に Retriever で参照する OEM 固有知識 DB を切り替える．

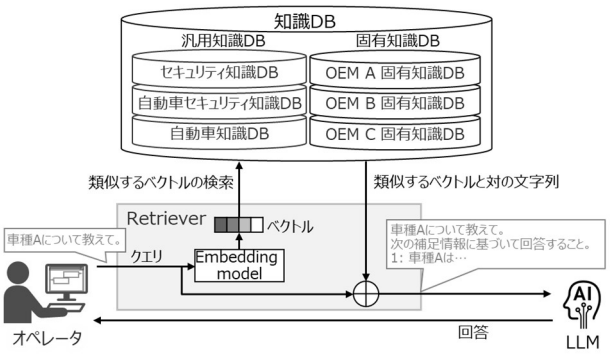

図 1 ベースとする RAG システム

3.2 課題の抽出

LLM システムに求められる指標 [11]のうち，図 1 に示す RAG システムでは VSOC 向けのチャットアシストシステムに求められる性能が満たされない指標に着目し，システム開発に向けた課題を抽出する．LLM システムに求められる指標と，各指標について図 1 のシステムが VSOC 向けのチャットアシストシステムに求められる性能を満たすか否かに関して考察した結果を表 1 に示す．Truthfulness と Privacy が満たされないことから，「(a) 知識 DB を構成する文献の特定」と，「(b) OEM 固有情報の回答への混入防止」を課題として抽出した．

表 1 LLM システムに求められる指標に対するベースシステムの考察

指標	性能が十分か	理由
Safety: 回答の健全性	✓	学習済みの LLM は，これら指標において実用上十分な性能を有する．また，運用上の想定として，これらの指標において，LLM が難しい判断を下すクエリがオペレータから入力されることは無い想定である．
Machine Ethics: 回答の倫理性	✓	
Fairness: 回答の公平性	✓	
Transparency: 意思決定の透明性	✓	
Robustness: ノイズに対する堅牢性	✓	
Accountability: 著作権に対する説明責任	✓	著作権を侵害する可能性のある文献は利用しないため．
Truthfulness: 回答の正確性	✗	正確な回答を生成するために必要十分な文献を持つ知識 DB が必要があるため．
Privacy: 機密情報の守秘性	✗	知識 DB 管理者の人為的ミスによって固有知識 DB の情報が回答に混入するリスクがあるため．

3.3 VSOC 向けのチャットアシストシステムの概要

システムの概要を図 2 に示す．本システムは図 1 に示した RAG システムをベースとして，前節の課題「(a) 知識 DB を構成する文献の特定」および「(b) OEM 固有情報の回答への混入防止」を解決する仕組みを取り入れている．課

© 2025 Information Processing Society of Japan

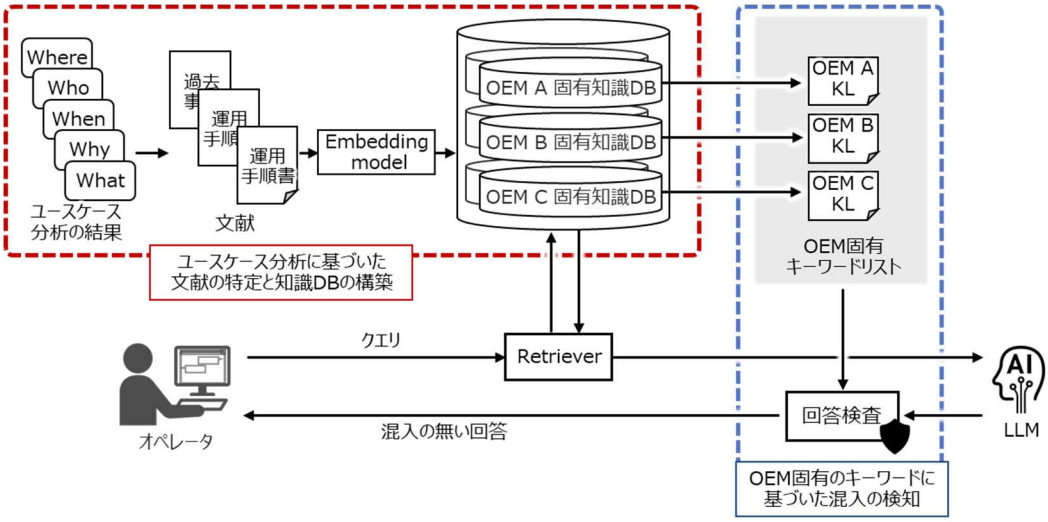

図 2　VSOC 向けのチャットアシストシステムの概要

題 (a) に対しては，本システムのユースケースの分析に基づいて知識 DB を構成する文献を過不足なく抽出する．課題 (b) に対しては，OEM の固有キーワードに基づいた混入検知手法を開発する．

3.4 ユースケース分析に基づいた文献の特定と知識 DB の構築

システムの知識 DB を構成する文献を過不足なく特定するために，まずはシステムのユースケースの分析を行う．ユースケースの 5W (Where, Who, When, Why, What) を次のように定義する．

- Where: アラートの発生場所．アラートが発生する Onboard（自動車）や Offboard（サーバ）が本項目の要素として挙げられる．また，アラートが発生していない際のユースケースも考慮するため，「セキュリティイベント無し」の要素も加える．
- Who: ユーザの種別．業務の習熟度レベルに応じて質問が異なる場合が想定されるため，所属組織ごとの区分に加えて，初心者と熟練者の区分を設ける．また，知識 DB への独自知識の追加・削除を行う，知識 DB 管理者の要素も加える．
- When: VSOC の業務プロセス．本項目の要素は，VSOC で運用される業務プロセスのうち，自動化されているプロセス（例：検知ルールによるセキュリティイベントの検知）を除いたものである．
- Why: システムを利用する動機．専門用語の確認/再確認や手順書の確認，類似した過去事例の確認などが本項目の要素として挙げられる．
- What: ユーザが取る行動．システムへのセキュリティ知識についての質問，車両情報についての質問などが本項目の要素として挙げられる．

5W の項目ごとに洗い出した要素を表 2 に示す．

表 2　ユースケース分析

5W	要素
Where	オンボード
	オフボード
	セキュリティイベント無し
Who	VSOC プロバイダ所属のオペレータ（初心者）
	VSOC プロバイダ所属のオペレータ（熟練者）
	VSOC プロバイダ所属の DB 管理者
	OEM 所属のオペレーション（初心者）
	OEM 所属のオペレーション（熟練者）
When	過去事例の確認
	セキュリティイベント発生箇所の状況確認の指示
	誤検知の判断
	知識 DB のメンテナンス
Why	専門用語の確認/再確認
	手順書の確認
	セキュリティイベント発生状況の言語化
	類似した過去事例の確認
	追加確認事項のリストアップ
	インシデントか否かの再判断
What	セキュリティ汎用知識に関する質問
	セキュリティイベントに関する質問
	ログに関する質問
	車両情報に関する質問
	オフボード情報に関する質問
	手順書に関する質問
	タスク完遂に向けた質問

© 2025 Information Processing Society of Japan

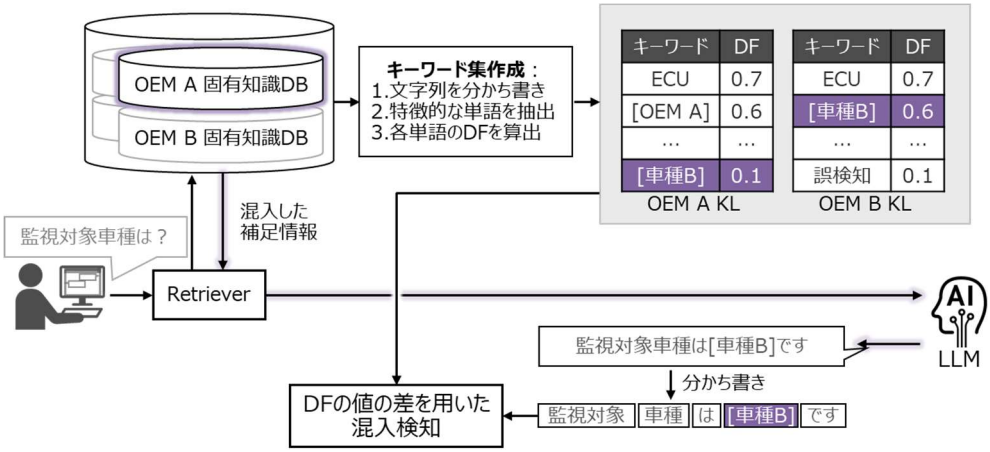

図 3　混入検知手法の概要

表 3　ユースケース分析から導出されるクエリ，回答，文献の例（一部の語は秘匿情報のためマスク）

5W	Where	オンボード
	Who	VSOC プロバイダ所属のオペレータ(初心者)
	When	誤検知の判断
	Why	手順書の確認
	What	車両情報に関する質問
クエリ		***GW(車載デバイスの型式)とは何ですか.
回答		***GW は，車載機の型名で、VSOC の監視対象であり，ログを取得している．…
文献		OEM 固有用語集
5W	Where	オンボード
	Who	OEM 所属のオペレーション（初心者）
	When	過去事例の確認
	Why	類似した過去事例の確認
	What	タスク完遂に向けた質問
クエリ		***(セキュリティイベント名)に関する過去事例をリストアップしてください.
回答		***の過去事例を以下にリストアップする. 1： 課題 ID：*** 過去の対応内容：*** 2： 課題 ID：*** …
文献		過去事例集

次に，洗い出した各項目の要素を組み合わせることで，ユースケースを導出する．ただし，ユースケースとして発生し得ない組み合わせは除外する．例えば，「Where: オンボード」と「What: オフボード情報に関する質問」の組み合わせはユースケースから除外する．また，同じ質問が想定されるパターンはマージする．以上の手順で導出された 5W の各組み合わせに基づいて，想定される質問と回答が 85 件作成される．また，この時回答の作成で参照した文献を集約することで知識 DB を構成する文献が特定される．作成された質問，回答，および特定された文献の例を表 3 に示す．

3.5 OEM の固有キーワードに基づいた混入の検知

本システムを運用する際は，監視対象の増加や運用手順の変更に伴って知識 DB が更新されることが想定される．本システムはオペレータの担当 OEM に応じて Retriever で参照する OEM 固有知識 DB を切り替える機構を備えるが，知識 DB を更新する時に DB 管理者の人為的ミスによって他 OEM の文献が OEM 固有知識 DB に登録された場合は，回答に他 OEM の情報が混入する可能性がある．そこで，回答への情報混入を防ぐために，固有知識 DB から作成したキーワードリスト(KL)に基づいて情報混入を検知する手法を開発する．開発した混入検知手法の概要を図 3 に示す．KL は OEM 固有知識 DB ごとに事前に作成され，当該 OEM 固有知識 DB から抽出された特徴的な語（キーワード）とその語の出現文献頻度(Document Frequency: DF)がペアで記録される．キーワードの抽出には TF-IDF (Term Frequency - Inverse Document Frequency) を用いる．回答の検査では，回答内の各語について，オペレータの担当 OEM と他 OEM との間における DF の差を算出し，混入を検知する．例えば，「車種 B」というキーワードに着目したとき，「車種 B」を本来用いる OEM B では「車種 B」に関する文献が複数出現すると考えらえるため，OEM B KL では DF の値が大きくなる．一方で，「車種 B」に関する文献が誤って固有知識 DB に追加された OEM A KL では，DF の値が小さくなる．回答の各単語を検査し，オペレータの担当 OEM の KL で DF が小さく，他の OEM の KL で DF が大きい場合に混

© 2025 Information Processing Society of Japan

入を検知する．混入を検知した場合は，オペレータへの回答の一時停止や，DB 管理者への混入発生の通知といった対処が考えられる．

4. 性能評価実験

4.1 回答精度の評価

構築した知識 DB を用いた提案システムの回答の正確性を評価する．多数のユースケースに対して低コストかつ再現性のある評価を実施するために, LLM を評価者として回答を評価する LLM-as-a-Judge [12]によって評価を行う．評価用データセットとして，ユースケース整理時に作成した質問と回答を使用する. 作成した回答を Ground truth とし，システムで生成した回答と内容が一致している場合に，評価者の LLM は高いスコアを与える．スコアは 5 段階であり，各スコアの評価基準は表 4 に示す通りである．なお，LLM-as-a-Judge による評価の妥当性の検証，および各スコアを付与された回答の実用性の確認を別途で実施した．これに基づき，実用性がある回答を得られるスコア: 4.00 を目標値として設定する．

評価結果を表 5 に示す. ベースラインの精度が改善され，特に LLM として GPT-4o を用いた場合はスコア 4.26 を達成し，目標値を上回る精度で回答を生成できることがわかる. GPT-4o を用いた場合のベースラインと提案システムの回答の比較を表 6 に示す．提案システムでは，ベースラインでは不可能な固有知識に基づいた用語の説明や過去事例のリストアップが可能となっていることがわかる. 一方で，複雑にセルが結合された表や，クエリとの関連性が低い内容を補足情報として付加して回答精度が低下するケースが見られた．これらのケースについて回答精度を向上するために，表データを文字列に変換する際の処理手順を改良したり，Retriever による補足情報付加の処理手順を改良する必要がある．

表 4　各スコアの評価基準

スコア	正しい内容の割合	誤った内容の割合
5	全て含む	全く無い
4	ほぼ全て含む	僅かにある
3	半分程度含む	半分程度誤っている
2	ほとんど不足している	ほぼ全て誤っている
1	全て不足している	全て誤っている

表 5　回答精度の評価結果

		スコア
ベースライン	GPT-4o	3.41
	GPT-4o-mini	3.06
	GPT-3.5-turbo	2.80
Proposed	GPT-4o	**4.26**
	GPT-4o-mini	3.82

表 6　GPT-4o を用いた場合の提案システムとベースラインの回答の比較
（一部の語は秘匿情報のためマスク）

	クエリと回答
クエリ	***(車両型式)とは何ですか.
ベースライン	***に関する具体的な情報はありません. おそらく、文脈や特定の技術に関連する用語か型番かと思われますが、正確な内容を把握するためにはより詳細な情報が必要です．…
Proposed	***は、「*** (一般に知られる車種名)」を指します．この車両は、既存の SOC（セキュリティオペレーションセンター）システムにおいて監視されており、***の社内組織である***が監視を行っています．…
クエリ	***(セキュリティイベント名)に関する過去事例をリストアップしてください.
ベースライン	具体的な過去の***に関する事例をリストアップすることは難しいですが、一般的な事例として以下のようなものがあります．…
Proposed	以下は***に関する過去事例のリストです： 1. **20********** - 課題 ID: ***** - 課題タイプ: ***** - …

4.2 混入検知精度の評価

OEM 固有知識 DB で情報混入が生じた状態を意図的に作り, 情報混入の検知精度を評価する. OEM A 固有知識 DB および OEM B 固有知識 DB を構築し, OEM B 固有知識 DB には OEM A の VSOC 運用手順書を混入させ, それぞれ KL を作成する．この状態で，OEM B 担当オペレータのロールで回答生成を行い，混入がある場合を Positive，混入がない場合を Negative として検知精度を評価する．見逃し無く，かつ誤検知を 1/3 に抑えた場合の F1-score: 0.80 を目標値として設定する．

© 2025 Information Processing Society of Japan

表 7　混入検知精度の評価結果

評価指標	評価結果
Precision	0.81
Recall	0.85
F1-score	0.83

表 8　混入検知結果の例
（一部の語は秘匿情報のためマスク）

生成した回答	現在の発生状況として、以下の点をサマライズしました. - 車両 ID “***” が関与するログに基づくと、「***」という通知タイプが生成され、通知データ内の一部のセキュリティパラメータが違反状態を示しています．具体的には、「***gate***security」が「True」となっており、これはセキュリティ違反が検知された状態を示唆しています.
Ground truth	Positive (***gate***security)
検知結果	Positive (***gate***security)
生成した回答	M76V に関する具体的な情報は提供されていませんが、参考情報から判断すると、M76V はセキュリティ監視に関連するシステムまたはプロジェクトの一部である可能性があります．…
Ground truth	Positive (M76V)
検知結果	Negative
生成した回答	***とは、CAN ID ごとに規定された***Code に適合しない CAN メッセージを***が受信することを意味します．…
Ground truth	Negative
検知結果	Positive (Code)

評価結果を表 7 に示す．F1-score=0.83 を達成し，目標値を上回る精度で混入を検知できることがわかる．正しく混入を検知できた回答および検知できなかった回答の例を表 8 に示す．一つ目の例は，OEM A 固有のログのフィールド名である「***gate***security」が回答に混入したケースであり，正しく検知できている．一方で，二つ目の例は OEM A の車両型式である「M76V」が回答に混入したケースであるが，検知に失敗している．本研究では，分かち書き処理に既存のライブラリを利用しており，車両の型式等に見られる数字とアルファベットの列は別の語として分割される．例えば，「M76V」は「M / 76 / V」に分割される．そのため，型式等を固有キーワードとして検知に用いることが困難であり，混入の見逃しの原因となる．分かち書き処理を改良し，型式等をひとつの単語として扱えるようにする必要がある. 三つ目の例は混入が発生していないケースであるが，一般的な語である「Code」を OEM A の固有キーワードとして誤検知している．KL への一般的な語や共通の専門用語の誤登録による誤検知を低減するために，KL 作成時に一般的な語彙を集めた辞書等を用いてこれらの語を除外する後処理を加えることが有効だと考えられる．

5. 結論

本研究では， LLM を用いて VSOC オペレータの質問に対して専門的な知識に基づいた回答を生成するチャットアシストシステムを開発した．知識 DB から補足情報を付加し回答を生成する RAG を導入し，ユーザ権限に応じて参照する知識 DB を切り替えることで, ユーザの所属 OEM に応じて回答を生成する方式とした. RAG により正確な回答を生成するためには質の高い知識 DB が必要であり，また知識 DB には OEM 固有の知識が含まれるため、他 OEM 固有の知識が回答に混入することを防ぐ必要がある．そのため，「(a)知識 DB を構成する文献の特定」と，「(b)OEM 固有情報の回答への混入防止」を解決すべき課題として抽出した．課題(a)に対しては，システムのユースケースの分析に基づいて知識 DB を構成する文献を過不足なく抽出した．さらに, 課題(b)に対しては, OEM の固有キーワードに基づいた情報混入検知手法を開発した．回答精度の評価では，LLM-as-a-Judge によってシステムの回答を 5 段階のスコアで評価し，スコア 4.26 を達成した．また，情報混入の検知精度の評価では，情報混入が生じた状態でキーワードリストの作成と回答生成を行うことで混入検知精度を評価し，F1-score=0.83 を達成した．知識 DB や Retriever，キーワードリスト作成時および混入検知時の分かち書き処理の改良による，さらなる回答精度および検知精度の向上が今後の課題である．

参考文献

[1] R. Okuda et al., “A survey of technical trend of ADAS and autonomous driving,” Tech. Papers of 2014 Int'l Symposium on VLSI Design, Automation and Test, 2014.
[2] J. Conditt, “OnStar hack remotely starts cars, GM working on a fix”.
[3] C. Miller et al., “Remote Exploitation of an Unaltered Passenger Vehicle,” 2015.
[4] M. Rogers et al., “How to Hack a Tesla Model S,” 2015.
[5] A. A. Elkhail et al., "Vehicle security: A survey of security issues and vulnerabilities, malware attack and defenses," IEEE Access, 2021.
[6] UNECE, UN Regulation No. 155 - Cyber security and cyber security management system, 2021.
[7] ISO, SAE, ISO/SAE 21434 Automotive cybersecurity engineering, 2021.
[8] P. Lewis et al., “Retrieval-Augmented Generation for Knowledge-

© 2025 Information Processing Society of Japan

Intensive NLP Tasks," Neur IPS, 2020.
[9] A. Vaswani et al., "Attention is all you need," NIPS, 2017.
[10] A. Radford et al., "Improving Language Understanding by Generative Pre-Training," 2018.
[11] Y. Huang et al., "TrustLLM: Trustworthiness in Large Language Models," arXiv, 2024.
[12] J. Gu et al., "A Survey on LLM-as-a-Judge," arXiv, 2025.

© 2025 Information Processing Society of Japan