日本銀行金融機構局 金融高度化センター設立20周年記念ワークショップ
「デジタル化とわが国の金融の未来」

金融分野におけるセキュリティの潮流

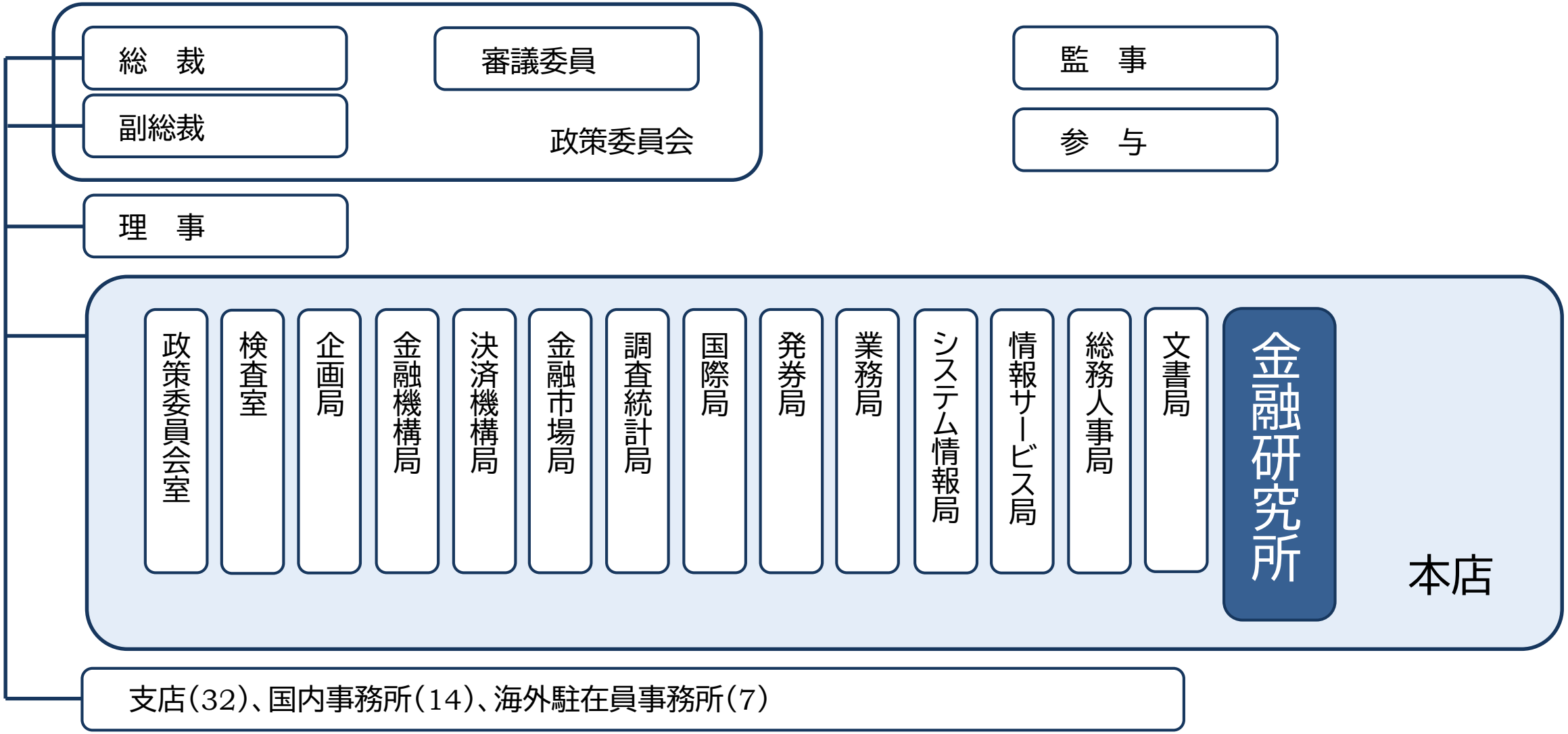
2025年1月31日

日本銀行金融研究所 情報技術研究センター長
鈴木淳人

アジェンダ

1. 金融研究所情報技術研究センター（CITECS）の紹介
2. 量的・質的に変化するサイバーセキュリティの脅威
3. 背景1）地政学的リスクの高まり
4. 背景2）生成AIの普及
5. 技術的な対応策―アカデミアでの研究動向
   （1）フィッシング攻撃―コグニティブ／ユーザブル・セキュリティ
   （2）ディープフェイク
6. 制度的な対応策
7. むすび：CITECS情報セキュリティシンポジウムのご案内

日本銀行の組織（約4,500人）



1. 金融研究所情報技術研究センター(CITECS)の紹介

金融研究所（約90人）

所長

参事役(所内組織運営、
情報技術研究センター長)

参事役(情報技術関係)

経済ファイナンス研究課

総務企画
経済研究
ファイナンス研究

制度基盤研究課

法制度研究
会計研究
情報技術研究センター*

歴史研究課

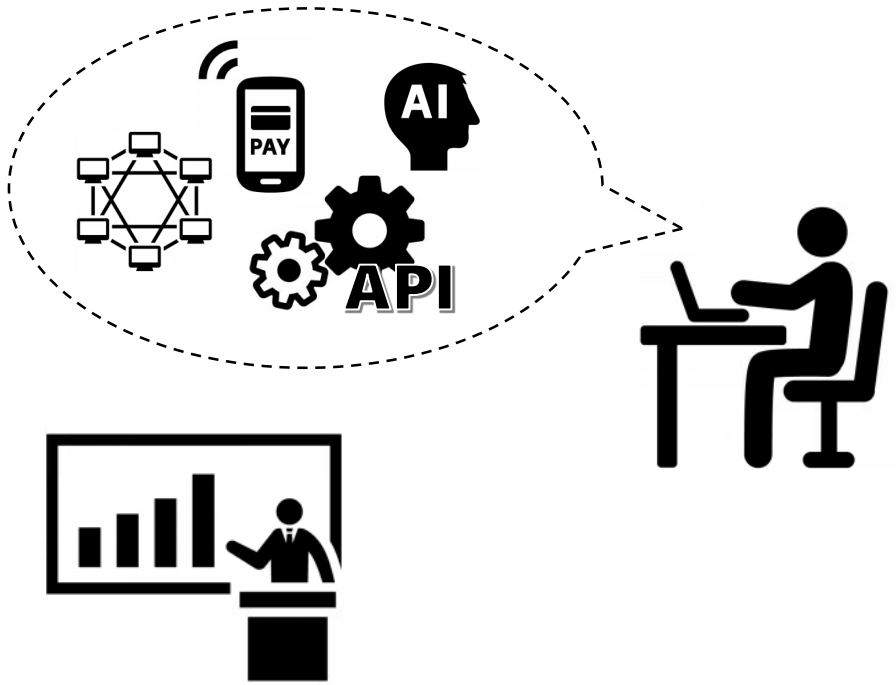
アーカイブ
貨幣博物館
金融史研究

Center for Information Technology Studies（CITECS)

情報技術研究センター(情報技術研究グループ)

金融業界が情報化社会において直面する新たな課題に適切に対処していくことをサポート

- 新しい情報セキュリティ技術の研究
  - ✓ 量子コンピュータの影響と暗号移行のあり方
  - ✓ 機械学習システムのセキュリティ
  - ✓ 小口決済手段・暗号資産のセキュリティ
  - ✓ スマートフォン決済のセキュリティ　など
- 金融業界への情報発信
  - ✓ 論文・レポートの公表
  - ✓ 情報セキュリティ・シンポジウム、セミナーの開催

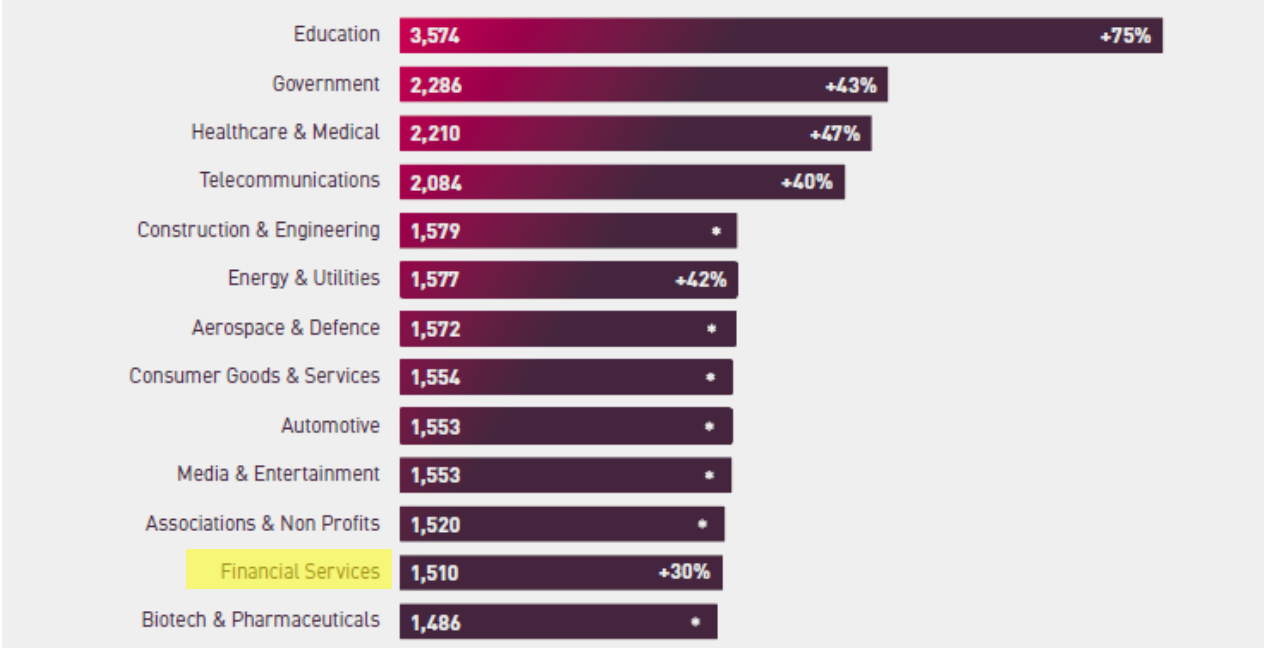


2. 量的・質的に変化するサイバーセキュリティの脅威

サイバーセキュリティの「量的」変化

2024年は、世界中で前例のない程、サイバー攻撃が急増（前年比＋44%）。
セクター別では、教育（同＋75%）、政府（同＋43%）が上位。金融も増加（同＋30%）。

1組織あたり週平均サイバー攻撃数（業種別）
（グローバル、2024年と2023年との比較）



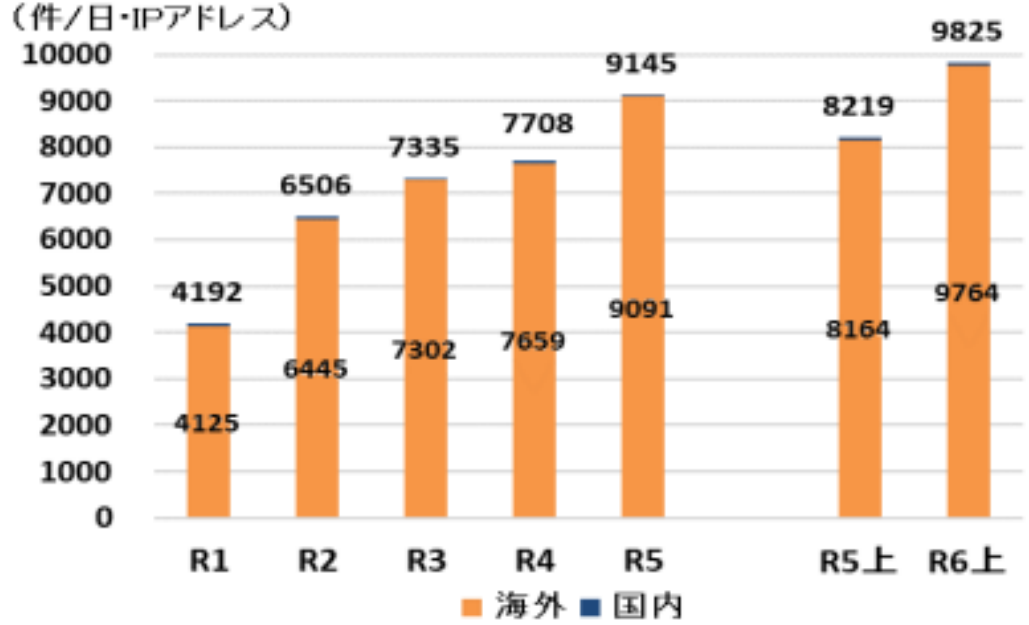
（出所）Check Point Software Technology, “The State of Cyber Security 2025,” 2025 (https://www.checkpoint.com/security-report/)

2. 量的・質的に変化するサイバーセキュリティの脅威

サイバーセキュリティの「質的」変化

脆弱性探索行為等の不審なアクセスの大部分は、海外が送信元。ビジネスメール詐欺が引き続き高水準である中、不正アクセス（機密情報の窃取・暴露）が増加。

脆弱性探索行為等の不審アクセス件数



（出所）警察庁サイバー警察局、「令和6年上半期におけるサイバー空間をめぐる脅威の情勢等について」、2024年
(https://www.npa.go.jp/publications/statistics/cybersecurity/data/R6kami/R06_kami_cyber_jousei.pdf)

「量的」・「質的」変化の背景

- 背景1）地政学的リスクの高まり
- 背景2）生成AIの普及

3. 背景1)地政学的リスクの高まり

国家によるサイバー攻撃の目的

- スパイ行為：国家によるサイバー攻撃の共通の目的。政府機関、請負業者、その他の企業が保有する軍事情報、知的財産、その他の機密情報を盗むことが目的（中国ほか）。
- ディスラプション/デストラクション：電力網や輸送インフラなどの重要なインフラストラクチャの破壊が目的（ロシア）。
- 政治的なメッセージ：政治的な声明を出すことだけが目的（ハクティビスト）
- 経済的利得：暗号資産などを盗むことが目的（北朝鮮）

（出所）Check Point Software Technology, “Nation-state Level Cyberattacks,” 2024 (https://www.checkpoint.com/jp/cyber-hub/cyber-security/what-is-cyber-attack/nation-state-level-cyberattacks/)参照

4. 背景2)生成AIの普及

マルウェア等開発

ランサムウェア、標的型攻撃、ビジネスメール詐欺は、生成AIが悪用される可能性が高い。

情報セキュリティ10大脅威　2024 「組織」向けの脅威の順位

順位	「組織」向け脅威	初選出年	10 大脅威での取り扱い（2016 年以降）
1	ランサムウェアによる被害	2016 年	9 年連続 9 回目
2	サプライチェーンの弱点を悪用した攻撃	2019 年	6 年連続 6 回目
3	内部不正による情報漏えい等の被害	2016 年	9 年連続 9 回目
4	標的型攻撃による機密情報の窃取	2016 年	9 年連続 9 回目
5	修正プログラムの公開前を狙う攻撃（ゼロデイ攻撃）	2022 年	3 年連続 3 回目
6	不注意による情報漏えい等の被害	2016 年	6 年連続 7 回目
7	脆弱性対策情報の公開に伴う悪用増加	2016 年	4 年連続 7 回目
8	ビジネスメール詐欺による金銭被害	2018 年	7 年連続 7 回目
9	テレワーク等のニューノーマルな働き方を狙った攻撃	2021 年	4 年連続 4 回目
10	犯罪のビジネス化（アンダーグラウンドサービス）	2017 年	2 年連続 4 回目

（出所）独立行政法人情報処理推進機構、「情報セキュリティ10大脅威 2024」、2024年 (https://www.ipa.go.jp/security/10threats/nq6ept000000g22h-att/kaisetsu_2024.pdf)

4. 背景2）生成AIの普及

マルウェア等開発の「容易化」

（事例）
2024年5月、対話型生成AIを悪用してマルウェアを作成したとして、不正指令電磁的記録作成容疑で警視庁が川崎市の男性を逮捕。

—— 2023年3月、パソコンやスマホを使って特定のファイルを暗号化したり、仮想通貨口座への送金を要求する文書を表示したりするプログラムのソースコードを作成した容疑。

—— 生成AIを使用してマルウェアを作成したことに起因する逮捕は国内では初めての事例。

10月25日、東京地裁にて、懲役3年・執行猶予4年の判決。

→ 今回の男性はプログラミングの専門知識を持っていなかったが、生成AIを悪用して、マルウェアを作成したもの。

→ ダークウェブ上では、RaaS（Ransomware as a Service）が存在。高度な技術を持たない人でも対価を支払うことで簡単に攻撃を行える環境が整っている。

（出所）各社報道資料を基に作成

4. 背景2）生成AIの普及

マルウェア等開発の「高度化」

（事例）
2024年2月、英Arup社が，ディープフェイクを悪用した詐欺により，2億香港ドル（約37億円）を騙し取られる事件が発生。この事件では，同社の財務部門の最高責任者の音声付き動画のディープフェイクが作成され，ビデオ通話において，財務部門の従業員が不正な資金移動を命じられた。

（出所）CNN, “British engineering giant Arup revealed as $25 million deepfake scam victim,” 2024 (https://edition.cnn.com/2024/05/16/tech/arup-deepfake-scam-loss-hong-kong-intl-hnk）参照
染谷実奈美・菅和聖・大塚玲、「生成AIのセキュリティリスクと研究動向」、『情報処理』66巻2号、2025年 (http://id.nii.ac.jp/1001/00241930/)

4. 背景2）生成AIの普及

マルウェア等開発の「効率化」

フィッシング・メール作成には平均16時間を要する（要していた）が、AIを使用することで5分でフィッシング・メールを生成でき、攻撃者にとっては2日近くの時間を節約できる可能性がある。

（出所）IBM、「X-Force脅威インテリジェンス・インデックス2024」、2024年 (https://www.ibm.com/jp-ja/reports/threat-intelligence)参照

（1）フィッシング攻撃―コグニティブ／ユーザブル・セキュリティ

コグニティブ・セキュリティ：人間の思考や行動に影響を与える悪意を持った情報攻撃から人・社会を守ること

コグニティブ／ユーザブル・セキュリティについては、CITECSが、2024年7月に開催した「情報セキュリティ・セミナー特別企画『認知の脆弱性から人間をどう守るか～コグニティブ・セキュリティと法的課題の入門～』」における議論に依拠している。
（https://www.imes.boj.or.jp/jp/conference/citecs/24sec_semi01/24sec_semi01.html）

コグニティブ・セキュリティの研究開発には、次の3つが必要

① 人間の認知的脆弱性の解明：人はなぜ騙されるのか、なぜ偽・誤情報を信じるのか、なぜ情報を拡散するのか、といった認知メカニズムの解明。

② 攻撃行動の観測・分析：攻撃または情報拡散のメカニズム、および、その影響の分析。また、生成AI などを使った新たな攻撃手法の分析。

③ 対策手段の構築・評価：個人が偽・誤情報に接触する前、接触した時、接触した後の対処法、情報システムの対応、法制度の対応の検討。

（出所）福井章人、「[セミナー資料]コグニティブセキュリティの研究動向と課題」、2024年 (https://www.imes.boj.or.jp/jp/conference/citecs/24sec_semi01/file1.pdf)

5. 技術的な対応策―アカデミアでの研究動向

（1）フィッシング攻撃―コグニティブ／ユーザブル・セキュリティ

「一貫性」「権威」「希少性」を表現したメールのクリック率が高く、人が騙されやすい。

認知的要因に基づくフィッシング・メールの分類

認知的要因		文例	ユーザのクリック率
一貫性 Consistency	約束は守りたい	あなたは利用規約に同意しています。利用規約に違反しない場合、アカウントを再開するにはこちらをクリックしてください。	高い
権威 Authority	権威を持つものに信頼をおく	社長の〇〇です。よろしくお願いします。以下を確認してください。	高い
希少性 Scarcity	限られたものほどほしい	あなたのアカウントは 48 時間以内に更新しないと、アクセスが制限されます。	高い
社会的証明 Social proof	周囲に同調したい	当社のサービスは、新しいセキュリティ機能を導入しました。アカウントを再度確認してください。	明確な関係が認められない
好意 Liking	好きな人に同意したい	当社はお客様のセキュリティを守っています。そのためにお客様のアイデンティティを確認してください。	明確な関係が認められない
返報性 Reciprocity	受けた恩は返したい	私たちはネットワークを安全に利用できるよう努力しています。ご協力をお願いします。	低い

（出所）福井章人、「[セミナー資料]コグニティブセキュリティの研究動向と課題」、2024年 (https://www.imes.boj.or.jp/jp/conference/citecs/24sec_semi01/file1.pdf)
Amber van der Heijden and Luca Allodi, “Cognitive Triaging of Phishing Attacks,” Proceedings of the 28th USENIX Security Symposium, 2019 (https://www.usenix.org/conference/usenixsecurity19/presentation/van-der-heijden)

5. 技術的な対応策―アカデミアでの研究動向

（1）フィッシング攻撃―コグニティブ／ユーザブル・セキュリティ

オンライン上の偽・誤情報に対抗するための介入手法の概要と、それに関連するエビデンスについて概説した論文が公表されている。

nature human behaviour

Review article

https://doi.org/10.1038/s41562-024-01881-

Toolbox of individual-level interventions against online misinformation

Received: 1 February 2023

Accepted: 5 April 2024

Published online: 13 May 2024

Check for updates

Anastasia Kozyreva[1], Philipp Lorenz-Spreen[1,29], Stefan M. Herzog[1,29], Ullrich K. H. Ecker[2,29], Stephan Lewandowsky[3,4,29], Ralph Hertwig[1,29], Ayesha Ali[5], Joe Bak-Coleman[6], Sarit Barzilai[7], Melisa Basol[8], Adam J. Berinsky[9], Cornelia Betsch[10,11], John Cook[12], Lisa K. Fazio[13], Michael Geers[1,14], Andrew M. Guess[15], Haifeng Huang[16], Horacio Larreguy[17], Rakoen Maertens[18], Folco Panizza[19], Gordon Pennycook[20,21], David G. Rand[22], Steve Rathje[23], Jason Reifler[24], Philipp Schmid[10,11,25], Mark Smith[26], Briony Swire-Thompson[27], Paula Szewach[24,28], Sander van der Linden[8] & Sam Wineburg[26]

The spread of misinformation through media and social networks threatens many aspects of society, including public health and the state of democracies. One approach to mitigating the effect of misinformation focuses on individual-level interventions, equipping policymakers and the public with essential tools to curb the spread and influence of falsehoods. Here we introduce a toolbox of individual-level interventions for reducing harm from online misinformation. Comprising an up-to-date account of interventions featured in 81 scientific papers from across the globe, the toolbox provides both a conceptual overview of nine main types of interventions, including their target, scope and examples, and a summary of the empirical evidence supporting the interventions, including the

①ナッジ（警告プロンプト表示などでユーザの行動変容を促す）
②教育的介入（手口を教育する）
③反駁方略（偽情報への反論により信念の変更を促す）

インタラクティブなオンラインツールボックスも公表。
(https://interventionstoolbox.mpib-berlin.mpg.de/)

（出所）田中優子、「[セミナー資料]人の認知的性質からみたセキュリティの課題」、2024年 (https://www.imes.boj.or.jp/jp/conference/citecs/24sec_semi01/file2.pdf)

5. 技術的な対応策―アカデミアでの研究動向

(2)ディープフェイク

ディープフェイク検知は、どのような生成手法にも対応できる汎用的なモデルの構築が難しいことが分かっており、複数の手法を組み合わせて使う必要がある。

―― 最新の研究でも、ディープフェイク検知モデルの精度は7割程度とあまり高くない。

	検知技術A	検知技術B	検知技術C	検知技術D
本物動画	0.1503	0.0561	0.6514	0.4413
Faceswap 動画 (Deepfake 動画)	0.9994	0.9898	0.6836	0.9861
Avatarify 動画 (Deepfake 動画)	0.2079	0.0877	0.8731	0.4116

本物動画とDeepfake動画の評価結果の平均値

検知技術Aと検知技術Bは、本物動画は0に近い数値で、Faceswap動画は1に近い数値のため、両方の動画を正しく判定できているといえる。
検知技術Cに関してはすべての動画で1に近い数値が出やすい傾向にあったが、それにより今回のツールの中で唯一、Avatarify動画をDeepfakeである可能性が高いと判定している。
しかし今回調査した限りでは、3種類の動画をすべて正しく判定できる検知ツールはなかった。

（出所）川名のんほか、「Deepfake動画を用いたeKYCに対するなりすまし攻撃の検知手法の検討」、『2024年度人工知能学会全国大会（第38回）論文集』、2024年
(https://www.jstage.jst.go.jp/article/pjsai/JSAI2024/0/JSAI2024_4M3GS1003/_pdf/-char/ja)
Yuying Li et al., “The Adversarial AI-Art: Understanding, Generation, Detection, and Benchmarking,” arXiv:2404.14581, 2024
(https://arxiv.org/abs/2404.14581)

(2)ディープフェイク

ディープフェイク検知の課題の1つが、検知手法を横並びで評価する体制が整備されておらず、どの検知手法が効果的かを見極めることが困難であること。

―― 評価用データセット44件を対象とした研究では、①格納されている合成データの分布が適切でないこと、②評価のベンチマークとなる基準モデルや評価指標値が公開されていないなどの問題が発見された。

→　再現性の高い、過去の研究との公平な性能比較を行うために、データセットの項目の標準化が必要。

（出所）Seth Layton et al., “SoK: The Good, The Bad, and The Unbalanced: Measuring Structural Limitations of Deepfake Media Datasets,” Proceedings of the 33rd USENIX Security Symposium, 2024 (https://www.usenix.org/conference/usenixsecurity24/presentation/layton)

5. 技術的な対応策―アカデミアでの研究動向

(2)ディープフェイク

e-KYCにおいて、生体検知などが必ずしも有効に機能しているとは言い難い。

―― 真正な動画の切り貼りといった加工は、原理的に検出できない。

さまざまな観点でチェックするには、目による検査（人間の統合認知）が引き続き重要。

6つの顔画像認証（facial liveness verification: FLV）で用いられる顔画像の影響

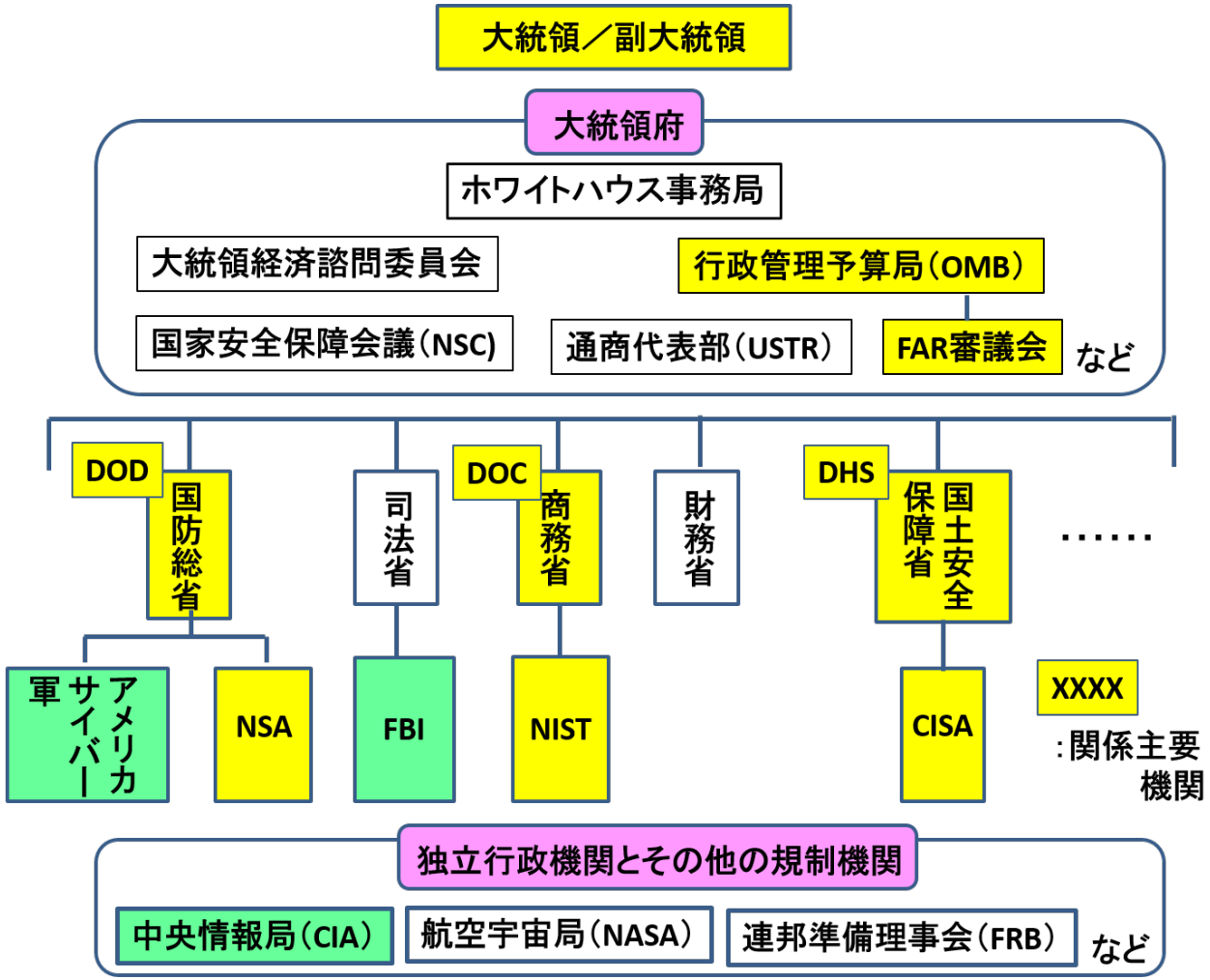
FLVの種類	顔画像の形態			
	顔の静止画（image）	顔の動画（video）		
		頭部の動作なし		頭部の動作あり（action）
		数字の発音なし（silence）	数字の発音あり（voice）	
F1	対応			
F2	対応			
F3	対応	未対応		対応
F4	対応			未対応
F5	未対応	対応		未対応
F6	対応		未対応	

資料：Li *et al.* [2022] Table 1

（出所）宇根正志、「スマートフォンによる顔認証のセキュリティ：ディープフェイクによる脅威と対策」、『金融研究』43巻4号、2024年
(https://www.imes.boj.or.jp/research/papers/japanese/kk43-4-5.pdf)
Changjiang Li, et al., “Seeing is Living? Rethinking the Security of Facial Liveness Verification in the Deepfake Era,” Proceedings of the 31st USENIX Security Symposium, 2022 (https://www.usenix.org/system/files/sec22fall_li-changjiang.pdf)
笠井健太郎ほか、「悪意のある切り貼り操作による動画編集に対する防御手法の一考察」、暗号と情報セキュリティシンポジウム-SCIS2024, 2024年

サイバーセキュリティ①:米国の体制

【大統領令「国家サイバーセキュリティ強化策」関係主要機関】



NIST（国立標準技術研究所、National Institute of Standards and Technology）＜業界＞

：ソフトウェアサプライチェーンのサイバーセキュリティ強化のガイダンス等を作成

CISA（国土安全保障省サイバーセキュリティ・インフラセキュリティ庁、Cybersecurity and Infrastructure Security Agency）＜インフラ＞

：重要ソフトウェアの定義やセキュリティ対策のガイダンスを作成

NSA（国防総省国家安全保障局、National Security Agency）＜軍＞

：同上

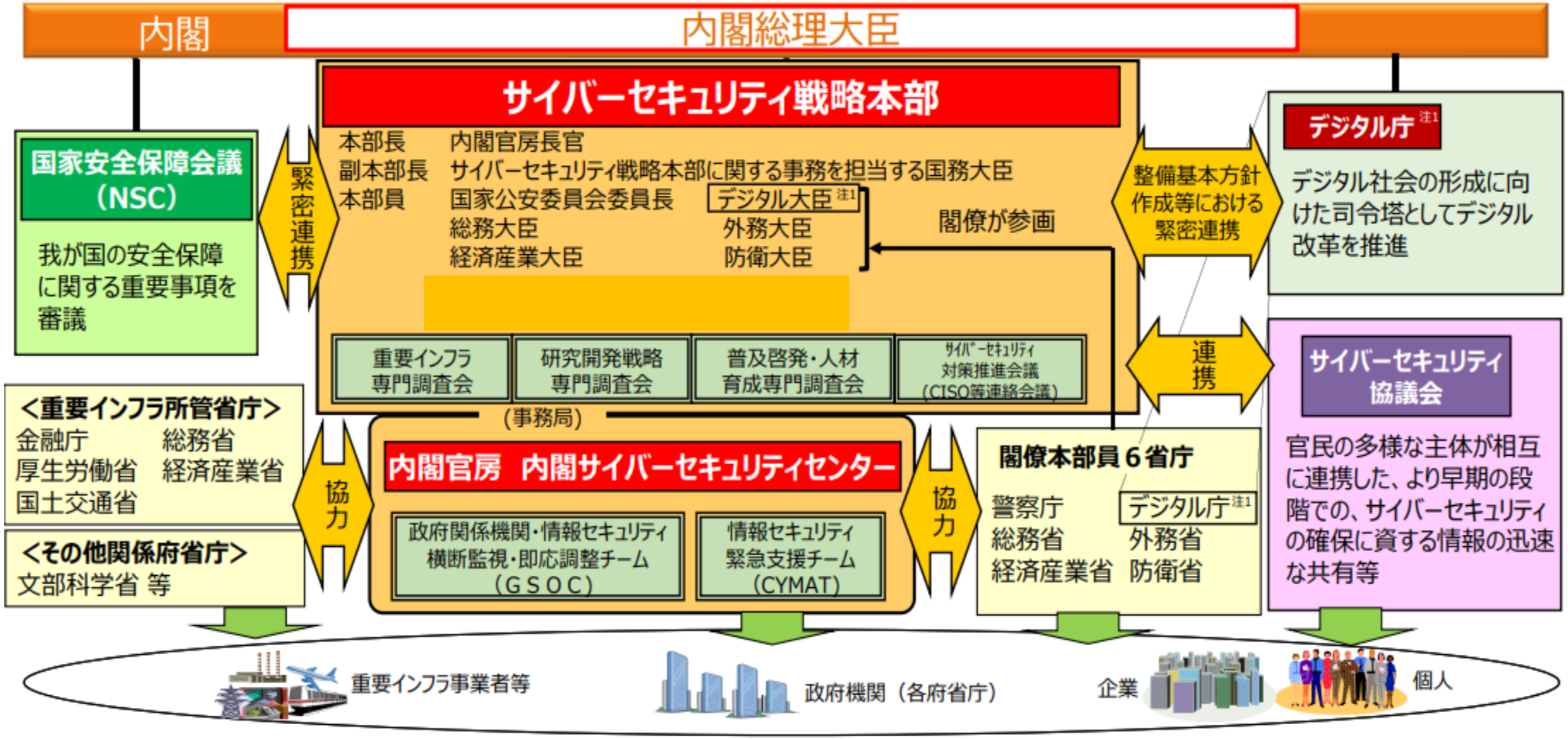
OMB（行政管理予算局、Office of Management and Budget）

：同上。上記各組織の実施状況を検証

※ 本強化策は、政府のサイバーセキュリティを強化するとともに、政府の助成を受けるソフトウェア業者の基準にもなっており、米国全体のサイバーセキュリティの強化に資するものとなっている

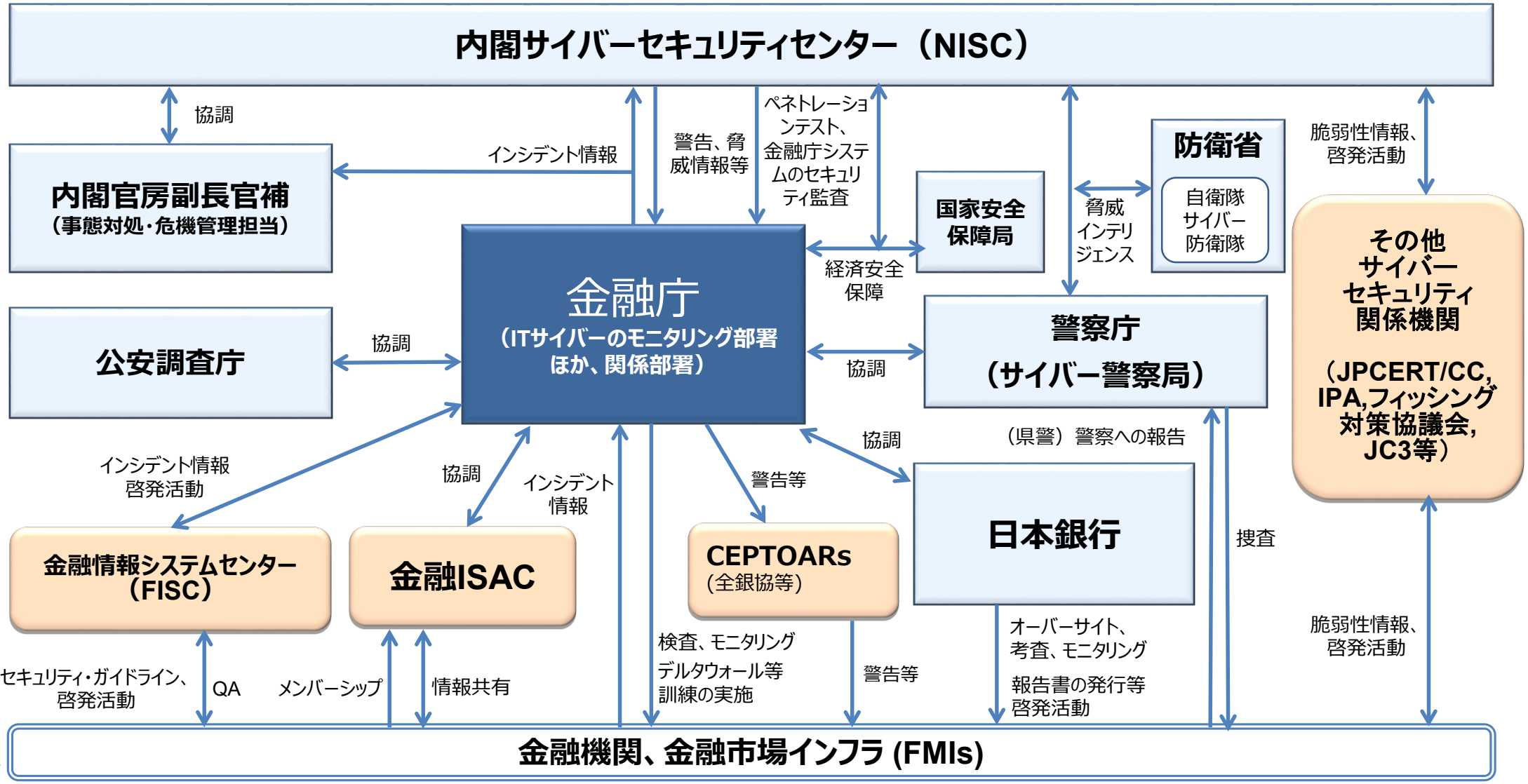
（出所）日本銀行、Tokio Cyber Port

サイバーセキュリティ②：わが国の体制



（出所）NISC

サイバーセキュリティ③：わが国の体制（金融関連）



（出所）
IMFのテクニカルノートを参考に日本銀行作成

サイバーセキュリティ④：国際的なベンチマーク

【ホワイトハウス「国家サイバーセキュリティ戦略」】(2023.3)

「サイバーセキュリティは、**経済の基盤的機能**、**重要インフラの運営**、**民主主義的機関の強靭さ**、**個人のデータと通信のプライバシー**、**国家防衛**に不可欠なものである」として、5つの柱を掲げている

①重要インフラの防衛（重要分野のサイバーセキュリティ基準の向上、官民協力の推進、政府ネットワークの現代化、危機対応策の更新）

②脅威ある行動者への対抗（全ツールの戦略的活用、ランサムウェアへの対処、国際パートナーとの連携）

③安全と強靭性を促進するための市場形成（個人データのプライバシーと安全保障、製品・サービスへの責任賦課、補助金の活用）

④強靭な未来への投資（インターネットの脆弱性削減、ポスト量子暗号やデジタル個人認証等次世代技術のための研究開発）

⑤共通する目標を追求する国際パートナーシップの構築（同盟・友好国との連携強化、パートナー国の能力向上）

【G7・Cybersecurity Experts Group 】

- 「金融セクターのサイバーセキュリティに関するG7の基礎的要素」(2016.10)
- 「金融セクターの**ランサムウェア**に対するレジリエンスに関するG7の基礎的要素」(2022.10)
- 「金融セクターにおける**サードパーティ**のサイバーリスクマネジメントに関するG7の基礎的要素」(2022.10)

→ 当局・金融機関の情報共有手順を確認する「**クロスボーダー協調演習**」を実施(2024.4)

【NIST「The NIST Cybersecurity Framework 2.0」】(2024.2)

- 国際標準として注目
- サイバーセキュリティ対策を、脅威の①**特定**、②**防御**、③**検知**、④**対応**、⑤**復旧**の5機能に分けて、具体的に規定
- 2.0では、①～⑤に横断的な⑥**ガバナンス**の機能、**サプライチェーン・リスクマネジメント**の重要性等が追加されている

(出所)日本銀行

（参考）金融機関向けの評価基準

【FFIEC 「Cybersecurity Assessment Tool（CAT）」】（2017.5）＜2025年8月31日に廃止予定＞

- 米国連邦金融機関検査協議会（Federal Financial Institutions Examination Council）がNISTのCybersecurity Frameworkをベースに金融機関向けに策定した監査基準
- 金融サービス連携協議会（Financial Service Sector Coordinating Council、FSSCC）が自動化されたCAT（Automated Cybersecurity Assessment Tool）を提供

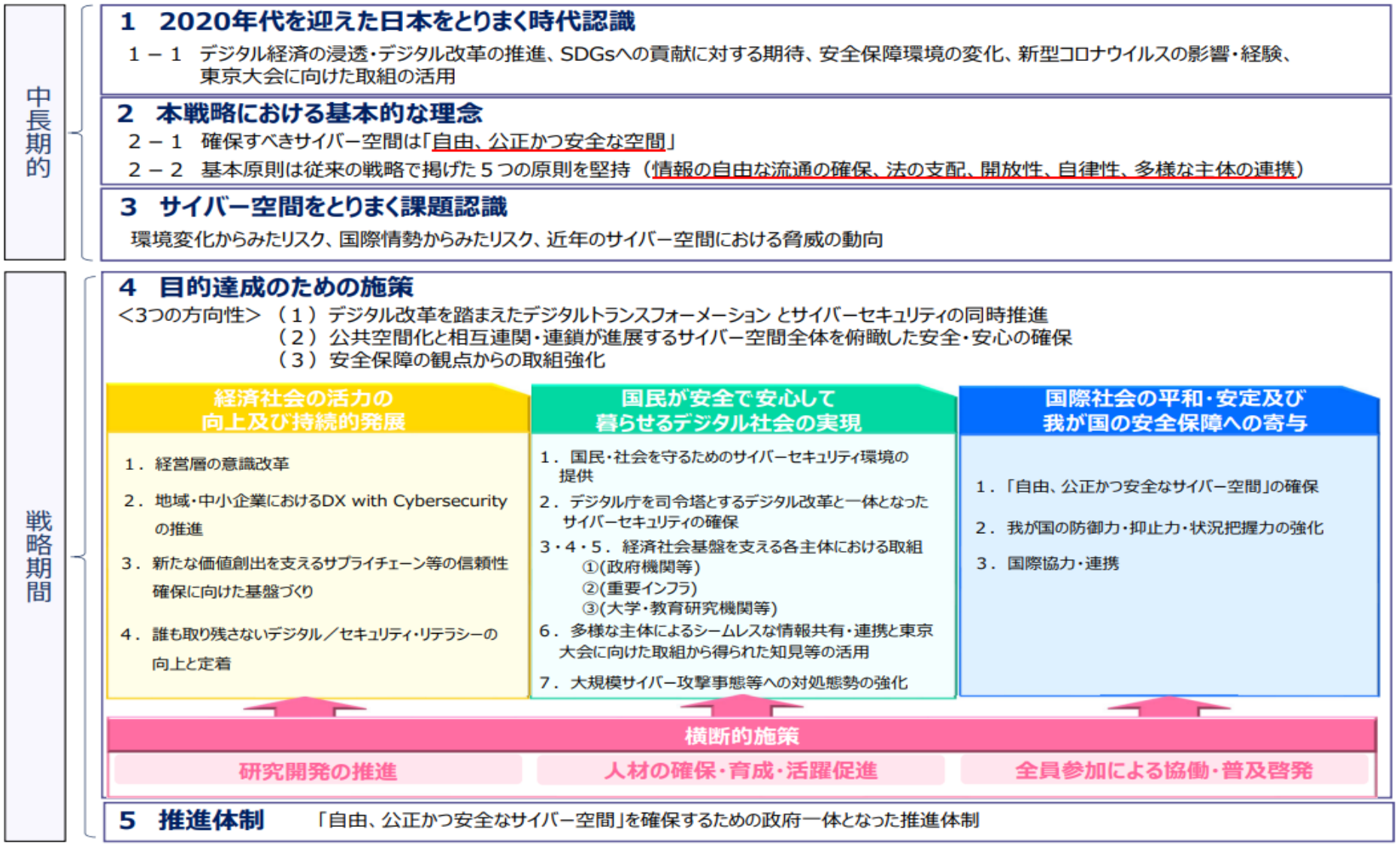
【Cyber Risk Institute 「CRI Profile ver2.0」】（2024.2）

- 非営利団体であるCyber Risk InstituteがNISTのCybersecurity Frameworkをベースに策定した金融機関向けのサイバーセキュリティを評価する枠組み
- 詳細で更新頻度が高く、金融機関の規模に応じた内容。各国規制やCATとの対応状況を明示

（出所）日本銀行

6. 制度的な対応策

サイバーセキュリティ⑤：日本の戦略（2021.9閣議決定）



（出所）NISC

サイバーセキュリティ⑥：国内金融機関の対応

【金融庁「金融分野におけるサイバーセキュリティに関するガイドライン」】（2024.10公表）

1.基本的考え方

サイバーセキュリティに係る基本的考え方、金融機関等に求められる取り組み、業界団体や中央機関等の役割、本ガイドラインの位置付けと監督上の対応

2.サイバーセキュリティ管理態勢

サイバーセキュリティ管理態勢の構築、サイバーセキュリティリスクの特定、サイバー攻撃の防御、サイバー攻撃の検知、サイバーインシデント対応及び復旧

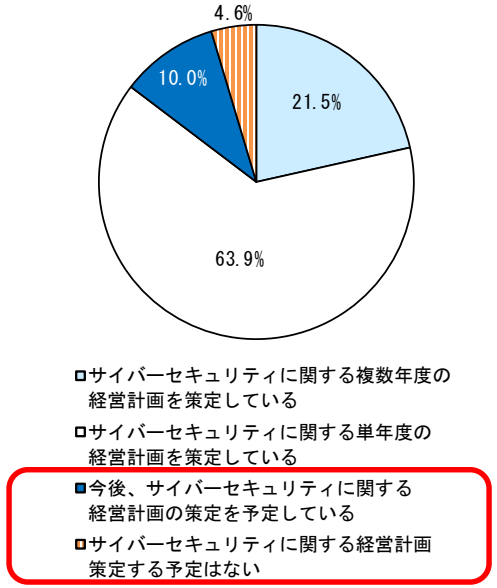
3.金融庁と関係機関の連携強化

【金融庁・日銀、大手銀行のサイバーセキュリティに関する共同調査（2021～）】

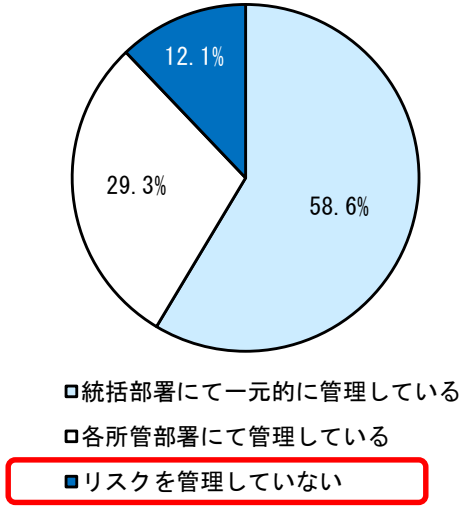
【金融庁・日銀、地域金融機関のサイバーセキュリティ・セルフアセスメント調査（2022～）】

日銀法37条
金融機関等のコンピュータ
障害等の場合の無担保貸出

▽ サイバーセキュリティの経営計画



▽ 重要なサードパーティのリスク管理状況



（出所）金融庁、日本銀行

7. むすび

サイバーセキュリティ対策

- ✓ 地政学的リスクの高まり ＋ 生成AIの普及
- ✓ サイバーセキュリティの「量的」「質的」な変化
- ✓ サイバーセキュリティ対策のカギ：「自助」、「共助」、「公助」
- ✓ アカデミアの視点：プライバシーには当然に配慮しながらも、金融機関の業界横断的な協力により、充実したデータセットの構築を期待。

（参考）CITECS情報セキュリティシンポジウムのご案内

「金融分野におけるセキュリティの潮流：CITECS設立20周年記念」

日時：2025年3月6日（木）14:30-17:20、開催形態：オンライン開催

基調講演　情報技術研究センターにおけるこれまでの20年　情報技術研究センター長・鈴木

講演①　　金融高度化センターの活動　金融高度化センター長・須藤

講演②　　耐量子計算機暗号への移行に対する考え方　宇根

講演③　　AIがもたらすリスクに対するセキュリティ　菅

対談・Q&A　菅×大塚玲教授（情報セキュリティ大学院大）

講演④　　さまざまな決済スキームとそのセキュリティ　田村

対談・Q&A　田村×面和成教授（筑波大）

講演⑤　　金融分野における今後のセキュリティ対策　岩下直行教授（京都大）

詳細は、日本銀行金融研究所HP（https://www.imes.boj.or.jp/）