Digital Architecture Design Center

IPA Better Life with IT

Society5.0 における安全・ガバナンスのアーキテクチャ設計に関するディスカッションペーパー
～Society5.0 における安全確保を実現するガバナンスの在り方に関するビジョン～

独立行政法人 情報処理推進機構
デジタルアーキテクチャ・デザインセンター
スマート安全プロジェクト

2021年７月

目次

図目次

表目次

ショートサマリー

必要なモノやサービスがさまざまなニーズに合わせてきめ細やかに人に配分される社会を実現するためにフィジカル空間とサイバー空間を融合させる先進テクノロジーを活用するSociety5.0において、複雑で変化の速い先進テクノロジーを生かすためには、先進テクノロジーがもたらす残留リスクを受容・共生する仕組みが必要となってくる。その際、従来のゼロリスク志向や絶対安全思想を前提としたガバナンス[1]から、リスクベース[2]（リスクを受容し、共生するアプローチを指す。）の安全を志向したガバナンスへ、ガバナンスの在り方を変革する必要がある。そのためには、「固定的に手段や行為を制限する」法規制から、リスクや社会の変化に応じて事業者が合理的に安全を達成できる「パフォーマンスの観点で要求する」ゴールベース[3]アプローチを取り入れた法規制に移行させる必要があることがポイントとなる。さらに、変化し続ける技術や社会システムのリスクに柔軟に対応し、常に変化に応じて設定した安全目標（ゴール）を見直していくことが求められるため、「動的な変化に柔軟に対応できるガバナンス（アジャイル・ガバナンス）」が必要となる。そこで、DADC スマート安全 PJ では、「GOVERNANCE INNOVATION VER.2: アジャイル・ガバナンスのデザインと実装に向けて」[4]を踏まえ、Society5.0 におけるガバナンスの在り方について安全に特化した切り口で、Society5.0 における安全確保を実現するためのあるべきガバナンスの姿の構造化及び具体化を実施した。

検討においては、事業者等のハザードの直接の管理主体が実施すべき「安全確保」と、それが適切に実施されているか確認し、許容レベルを超えている場合には是正を行うという組織外部の社会における様々な主体との関係性（主には規制当局）における「（安全確保に対する）ガバナンス」の双方を対象とした。その両方の要素を含めた概念を「安全・ガバナンス」と呼ぶ。

特に、法規制等のルールの「安全確保に対するゴールベースの要求」に対して事業者が実施する「安全確保」がステークホルダーの安全安心を実現するための核となる。そのため、安全

---

[1] サイバー空間とフィジカル空間を融合するシステム（CPS: Cyber-Physical System）について、これによって生じるリスクをステークホルダーにとって受容可能な水準で管理しつつ、そこからもたらされる正のインパクトを最大化することを目的とする、ステークホルダーによる技術的、組織的、及び社会的システムの設計及び運用をいう。（出典：GOVERNANCE INNOVATION VER.2: アジャイル・ガバナンスのデザインと実装に向けて）

[2] リスクアセスメントに基づき意思決定をする考え方であり、費用と便益の兼ね合いを社会が受け入れる基準を用いて管理、意思決定するという考え方。

[3] ここでの「ゴール」とは、個々のシステムレベルでのゴールではなく、規制対象となるシステムが一律に達成すべき最低限の水準の要件を意味する。規制対象となるシステムが一律に達成すべき最低限の水準の要件を定めることによって、期待する安全を達成する仕組みのこと。（出典：GOVERNANCE INNOVATION Ver.2）

[4] https://www.meti.go.jp/press/2020/02/20210219003/20210219003.html

を軸とした本 PJ では、より具体的に事業者の取組がどのように変わるのか明示することを重視し、「CPS の安全を実現する機能」として具体化を行った。併せて、安全確保の方法についての自動運転をはじめとする様々な分野でなされている先行の研究・取組について、ガバナンスとの組合せとして位置付けることを留意し、構造化を行った。

以上の分析、検討成果をディスカッションペーパーとして発信することで、安全に関連する様々な産業界のひと、安全規制に関与する行政官と Society5.0 における安全について議論を巻き起こし、 DADC の取組に対するステークホルダーの関心を拡大し、推進力となる仲間を増やすこと、様々な取組を惹起・促進できることを狙う。

1. はじめに

1.1. 背景

独立行政法人情報処理推進機構（以下、IPA という）デジタルアーキテクチャ・デザインセンター（以下、DADC という）では、スマート安全プロジェクトを設置し、Society5.0[5]における安全確保を実現するガバナンスに関するアーキテクチャを検討している。本ディスカッションペーパーでは分野横断的な安全・ガバナンスの在り方に関するビジョンに関し、現在までに検討した中間成果を報告する。

サイバー空間とフィジカル空間が融合する Society5.0 では、従来の「モノ」を起点としたフィジカル空間を中心とする世界から、デジタル技術の急速な進展により社会のあらゆる領域に変革の可能性をもたらすデジタルトランスフォーメーション（DX）が推進力となり、フィジカル空間のリアルな生活やビジネスにサイバー空間が高度に融合し、社会構造も急激に変化することとなる。その変化を鑑み、安全確保のあり方を定め、それが適切に実施されているか確認し、許容レベルを超えている場合には是正を行うという安全・ガバナンスの在り方についても再設計が必要である。その際、フィジカル空間を想定して作られた既存の制度枠組の中で対症療法的な修正を行うのではなく、企業、法規制、市場、といった多様なガバナンスメカニズムの在り方やその相互関係を、Society5.0 に適した形へと根本から設計し直さねばならない。

サイバー空間上で発展するシステムを前提とした社会においても、安全・安心を実現・維持する仕組みは必須である。このとき、「ソフトウェアを中心にサイバー空間上で」発展するシステムに対し、「ハードウェアをベースにフィジカル空間で」構築されている現在の法規制、規範、市場によるガバナンスで対応しようとしても、スピードの違いやコンセプト・前提の違いによって抜け漏れや不整合が発生するため、信頼性の高い安全・安心の確保は困難となる。Society5.0 における安全・ガバナンスを考える上では、高度なソフトウェア技術が関与する社会の特性を十分に理解した上で、新たなガバナンスの目標を設定し、その実現方法を提示しなければならない。

このような要請を背景に、DADC におけるスマート安全プロジェクトでは「安全を確保するためのガバナンス」に着目し、2020 年 5 月より活動を開始した。

[5] 内閣府「Society 5.0 とは」https://www8.cao.go.jp/cstp/society5_0/index.html

1.2. プロジェクトの活動の目的

Society5.0 における Cyber Physical System(CPS)が実現すると、サイバー空間における何らかのアクシデントが時には人の身体生命を直接的に脅かすことになりかねない。そのため、その安全性の担保・確保が喫緊の課題となる。

しかし、従来のガバナンスは、官が定めた法が事業者の行為をあらかじめ規制するルールベースの法規制となっている。Society5.0 で技術やビジネスが急速に変化していく中で、現在のようなルールベースの法規制という変化対応性の低い仕組みのままでは、イノベーションや社会的価値の実現を阻害することに繋がりかねない。ガバナンス変化の必要性に気づいた場合にも、ルール設定によってその変化を実現させるまでに規制当局による調整や改正等に時間を要するため、新たなリスクに規制や法律に対応することは困難となる。

さらに、Society5.0 においては、「コト」を中心としたサービス・ビジネスの構築がより進み、産業の CPS 化によって異なる目的を持つシステム同士が接続・連携してサービスを提供することとなる。こうして形成されるシステムの全体は複雑化する（そのような相互に作用するシステムの集合体を SoS（System of Systems）という）。想定されていないシステム・製品が接続されたり、システム同士の相互作用によって新たな機能・振る舞いが生じたりする場合には、予見・予防が難しい事態も発生し、市民の安全安心を脅かす不確実性は高まる。各システムが多様なステークホルダーによって個別に管理されている場合には、相互に接続されたシステム全体の不確実性を認識して評価したり排除したりすることは困難となる。

上記の課題意識に基づき、DADC スマート安全プロジェクトでは、Society5.0 の実現において社会システムや技術の変化により発生する新しい形態の不確実性に対応できるよう、安全確保の在り方、それを社会的に担保するためのガバナンスの在り方の検討を行う。つまり、本ディスカッションペーパーでは、事業者（ハザードの直接の管理主体）等が実施すべき「安全確保」と、それが適切に実施されているか確認し、許容レベルを超えている場合には是正を行うという組織外部の社会における様々な主体との関係性（主には規制当局）における「（安全確保に対する）ガバナンス」の双方を検討対象とする。その両方の要素を含めた概念を「安全・ガバナンス」と呼ぶ。

また、図 1-1 に示す通り、安全確保の手段を単に人からデジタル技術に置き換えること（図 1-1 の 3 層目）のみならず、Society5.0 に適した安全確保の在り方を考慮してガバナンスの本質的な変革を行うことを真のビジョンとする。

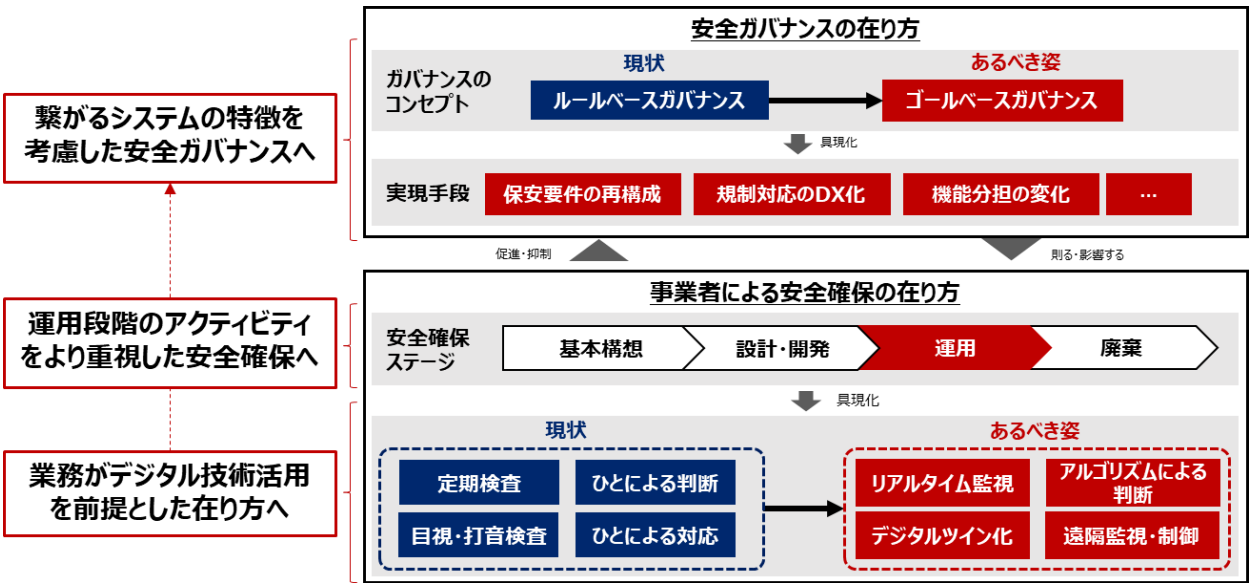


図 1-1　安全・ガバナンスのビジョンイメージ

1.3. 報告書の構成・前提と想定読者

本ディスカッションペーパーでは、まず議論の前提となる従来の知見や、技術と安全、技術とガバナンスに関して現在進んでいる国際的な議論、日本の現行ガバナンスの分析を踏まえて、Society5.0 における社会システムや技術の変化に伴う安全に係る課題を特定し、従来の安全に関する考え方・ガバナンスから、Society5.0 に向けて安全・ガバナンスはどのように変化すべき/していくかを捉える（第 2 章）。その上で、安全に係る有識者との議論を通じて検討した、目指すべき Society5.0 における安全・ガバナンスのアーキテクチャのビジョン及び課題認識を共有する（第 3 章）。

さらに、第 3 章で示した安全・ガバナンスのあり方の実装に向けて、具体的な分野をユースケースとして取り上げ、DADC の今後のアクションプランを共有・提案する（第 4 章）。

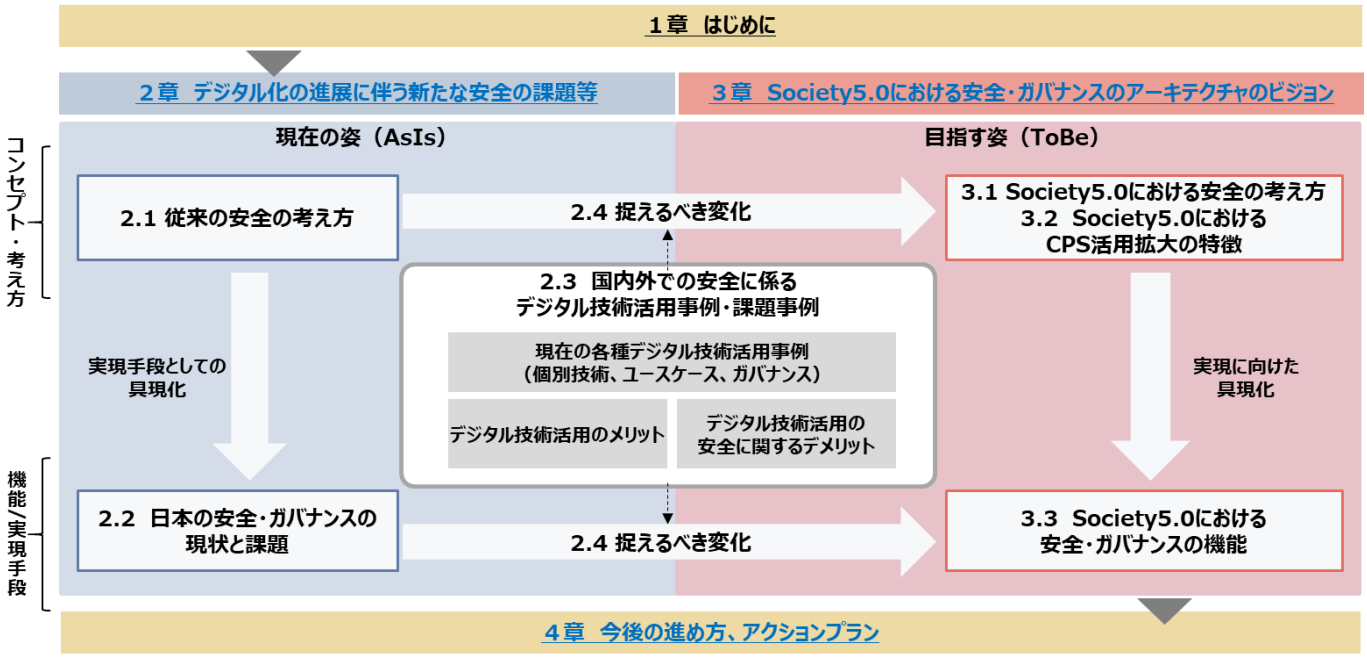


図 1-2　中間報告書の構成

本ディスカッションペーパーの検討にあたっては以下を前提とした。

- 事業者（ハザードの直接の管理主体）等の組織内部で実施すべき「安全確保」と、それが適切に実施されているか確認し、許容レベルを超えている場合には是正を行うという組織外部の社会における様々な主体との関係性（主には規制当局）における「（安全確保に対する）ガバナンス」を検討対象とする。その両方の要素を含めた概念を「安全・ガバナンス」と呼ぶ。
- 目指すべき Society5.0 の姿を前提とした「安全の将来像」を提案する。そのために「現在からの積み上げ」からではなく「Society5.0 からのバックキャスト」を軸にビジョンを描く。
- 「日本の現在の安全・ガバナンスに係る課題の解決」のみならず、「Society5.0 における安全・ガバナンスの姿」を検討する。
- 「安全」のみならず、「安全及びその他の要素（-ilities）[6]」を考慮しながら検討する。

また、本ディスカッションペーパーの想定読者を以下のように整理した。

- 企業活動、業務における安全（確保）に係る任務を持つ部署の方

[6] ilities とは、システムに要求される特性のこと。これらの特性は通常、非機能的な要件とされ、システムのパフォーマンスの主要な機能要件ではないが、組み込まれた主要な機能要件以上に、より幅広くシステムに影響を与えるものと言える。

- 製品開発・サービス提供を行うメーカー等企業及びその経営者や品質保証や法務を担う部署の方
- 安全規制を所管する省庁や審査機関、規制改革を担当する省庁の行政官
- システムベンダー、SIer、IT ベンチャー企業
- 投資家、金融機関、保険会社　　等

想定読者に対し、本ディスカッションペーパーを通し、下記の動きに繋がることを目指している。

- Society5.0 における安全・ガバナンスについて議論を巻き起こし、本検討に対するステークホルダーの関心を拡大し、将来に向けた課題解決の推進力となる仲間を増やすこと
- Society5.0 における社会像（未来像）に対する議論を巻き起こすこと
- 安全・ガバナンスに関する課題認識を揃え、遠い未来で社会のあり方が変化した際の安全を考えるためのマインドセットを行うこと
- 安全に係る主体の様々な取組を惹起・促進すること

1.4. 関連する概念、用語の説明

- Society5.0
  サイバー空間（仮想空間）とフィジカル空間（現実空間）を高度に融合させたシステムにより、経済発展と社会的課題の解決を両立する、人間中心の社会（Society）。狩猟社会（Society 1.0）、農耕社会（Society 2.0）、工業社会（Society 3.0）、情報社会（Society 4.0）に続く、新たな社会を指すもので、第 5 期科学技術基本計画において我が国が目指すべき未来社会の姿として初めて提唱された。
  【出典】内閣府ホーム>内閣府の政策>科学技術政策 ＞ Society 5.0,
  https://www8.cao.go.jp/cstp/society5_0/
- **アーキテクチャ**
  システムが存在する環境の中での，システムの基本的な概念又は性質であって，その構成要素，相互関係，並びに設計及び発展を導く原則として具体化したもの。
  【出典】JIS X 0170:2020 (ISO/IEC/IEEE 15288:2015)
- **サイバー・フィジカルシステム／CPS**
  デジタル世界（サイバー空間）と現実世界（フィジカル空間）を IoT 関連技術で結びつけ、産業の高度化や社会的課題の解決を図る仕組。産業・医療・インフラ・エネルギー・交通・公共サービスなど、現実世界のさまざまな分野で得られる大量のデータを、デジタル世界におけるクラウドコンピューティングやビッグデータの処理技術を通じ

て、価値ある情報やデータとして現実世界に還元し、広く社会規模で合理化や最適化を図ることを目的とする。

【出典】"GOVERNANCE INNOVATION VER.2: アジャイル・ガバナンスのデザインと実装に向けて",
https://www.meti.go.jp/press/2020/02/20210219003/20210219003-1.pdf

- **システム**

ある定義された目的を達成する、要素、サブシステム、またはアセンブリを統合したまとまり。これらの要素には、製品（ハードウェア、ソフトウェア、ファームウェア）、プロセス、人、情報、技術、設備、サービス、およびその他のサポート要素を含む（INCOSE）。一つ以上の定められた目的を達成するために編成された相互作用する要素の組み合わせ（ISO/IEC/IEEE15288)。

【出典】デイビッド・D・ウォルデンほか編,西村 秀和 監訳 "システムズエンジニアリングハンドブック 第 4 版",慶應義塾大学出版会,2019 年

- **システムオブシステムズ（system of systems）**

既存の構成システムがそのままでは達成できないサービスを提供するために、相互的に作用するシステム要素の集合体。

以下の 5 つの特徴をもったシステムを System of Systems (SoS)と呼ぶ。

1. 運用の独立性：SoS の構成システムは個別に運用される。
2. 管理の独立性：構成システムは別々に調達され、統合される。しかし、運用中の構成システムはそのまま運用される。
3. 進化的開発：機能や目的が、追加/削除されたり途中で変更されたりするなど、開発とシステムが進化的に変化する。
4. 創発 的振舞い：構成システム単独では実現できない目的を SoS として実現する。
5. 地理的な分散：構成システムが離れており、構成システム間で質量やエネルギーの物理量ではなく、情報を交換 SoS でない System を“Monolithic System”と呼ぶ。

【出典】ISO/IEC/IEEE21841:2019
Mark W.Maier,Architecting Principles for Systems-of-Systems,(1998)
”IoT 時代のシステムデザインアプローチ ～いかにして IoT システムをデザインするか～”,https://www.ipa.go.jp/files/000053968.pdf

- **ステークホルダー**

利害関係者。システムに、権利、持分、請求権もしくは関心を持っている個人もしくは組織、またはニーズおよび期待に合致する特性をシステムが持つことに、権利、持分、請求権もしくは関心を持っている個人もしくは組織。

【出典】JIS X 0170:2020 (ISO/IEC/IEEE 15288:2015)

- **ライフサイクル**

システム、製品、サービス、プロジェクトまたは人が作った他の実体の構想から廃止までの漸進的な発展を表す概念。

【出典】JIS X 0170:2020 (ISO/IEC/IEEE 15288:2015)

- **要求事項**
  ニーズとそれに付随する制約・条件とを変換した又は表現する文。
  【出典】JIS X 0170:2020 (ISO/IEC/IEEE 15288:2015)

2. デジタル化の進展に伴う新たな安全に係る課題

２章では、現代社会における新しいデジタル技術の活用事例等から安全のあり方に対する変化と示唆を捉える。

まず本ディスカッションペーパーにおける議論の前提となる、安全に対する既存の考え方を確認し（2.1）、現行の安全のガバナンスに関する課題について共通認識を図る（2.2）。そのうえで、Society5.0 における安全の在り方・その変化を議論するためにデジタル化技術利用のポジティブ側面・ネガティブ側面について国内外の活用事例、課題事例に基づいて共通認識を構築する（2.3）。最後に、3 章に続く Society5.0 における安全・ガバナンスの在り方を検討する前提を理解するために、捉えるべき変化を論じる（2.4）。

2.1. 前提となる従来の安全の考え方の整理

「安全」とは「許容できないリスクが無いこと」（ISO/IEC Guide51）と定義される。一方で、人間の活動は「安全」だけが目的となることはなく、現実的には（生きていくための）何らかの社会活動・生産活動の前提条件・付帯条件として「安全」な状態が要求される。（例えば、新技術導入は多くの場合、目的は安全の高度化のみならず利益追求の前提条件・コストカットのためになされる。）

図 2-1 に示す通り、「安全」とは「許容できないリスクがないこと」であり、安全か危険かの二者択一な考え方ではなく、リスクを許容できる領域を指す幅を持つ概念である。そのため、社会の変化、世論の変化、システムの変化等に応じて、その領域も変化しうることを念頭に置かれたい。

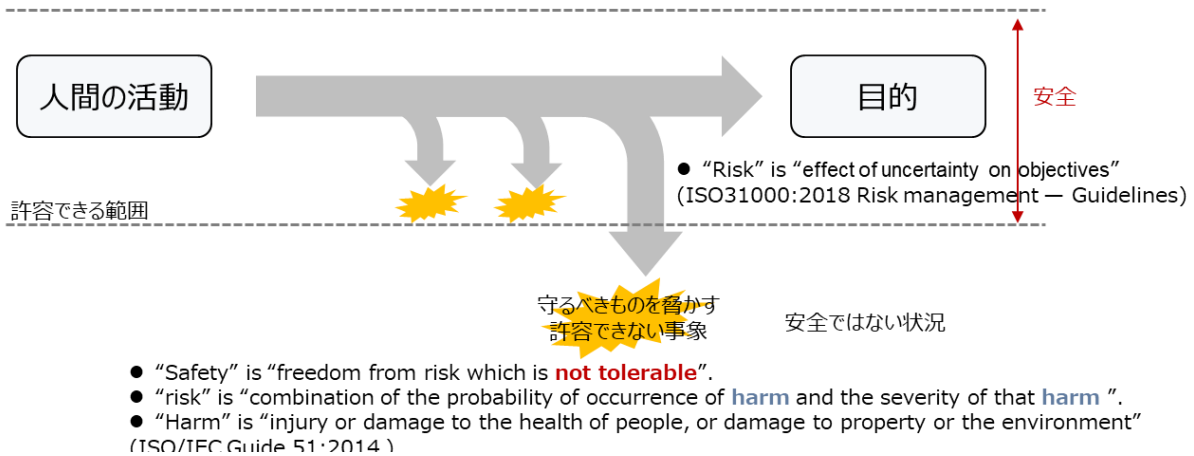


図 2-1 社会活動における安全の位置づけ

まず、安全に関連する主な特性（-ilities）について、主な用語の定義を中心に整理した。

本ディスカッションペーパーにおいては、安全を検討の中心に置きつつ、安全に係る概念も考慮に入れて検討を行う。-ilities とは、柔軟性 flexibility, 保守性 maintainability など、システムに要求される特性を示す多くの語の末尾に含まれる ”ility” を複数形にしたものである。システムに求められる特性は通常、非機能的な要件とされ、システムのパフォーマンスの主要な機能要件ではないが、組み込まれた主要な機能要件以上に、より幅広くシステムに影響を与えるものと言える[7]。

さらに近年、システムの複雑さが進んだことから、-ilities への注目が一層高まっている。複雑さが増し、規模が大きくなるにつれて発生するようになったシステムによる副作用を考慮する必要性から、数多くの-ilities が関連づけて検討されるようになってきた。安全というテーマにも様々な概念が関わるが、本ディスカッションペーパーにおいては、表 2-1 に示す概念をスコープに入れて検討することとする。

ある-ility を高めるために導入した仕様が、場合によっては他の-ility を損なう可能性もあり、システムの目的、特徴（リスク顕在化時の社会に対する影響の程度等）に応じて、いずれの ility を優先するか、どのようなバランスとするかを判断する必要がある。

[7] O. de Weck et.al., Engineering Systems: Meeting Needs in a Complex Technological World. MIT Press (2011)（春山 真一郎 監訳「エンジニアリングシステムズー複雑な技術社会において人間のニーズを満たすー」慶應義塾大学出版会, 2014）

表 2-1　関連する概念

項目	定義	課題・議論等	関連する取り組み
安全（Safety）	・許容できないリスクが無いこと（“freedom from risk which is not tolerable“）（ISO/IEC Guide 51：2014）	・サイバー空間とつながった機械の安全 ・AI により制御される機械の安全 ・人間と機械の協調時における安全 ・機能安全（IEC61508）における AI の取り扱い	・IEC 白書「Safety in the Future」 ・機能安全、SOTIF、Safety 2.0（協調安全）[8]
セキュリティ（Security）	・情報の機密性、完全性及び可用性を維持すること。 さらに、真正性、責任追跡性、否認防止及び信頼性のような特性を維持することを含めてもよい。（ISO/IEC27001（情報セキュリティ））	・悪意ある攻撃者、マルウェアの増加 ・新しい技術が意図的な攻撃に使われる可能性 ・重要インフラサービスのセキュリティリスクの増大（「情報通信」、「金融」、「航空」、「空港」、「鉄道」、「電力」、「ガス」、「政府・行政サービス（地方公共団体を含む）」、「医療」、「水道」、「物流」、「化学」、「クレジット」及び「石油」）	・NISC、IPA 等の各種ガイド・レポート ・【NISC】「重要インフラにおける情報セキュリティ確保に係る安全基準等策定指針 ・【所管省庁等】各分野のセキュリティガイドライン

[8] 参考　一般社団法人セーフティグローバル推進機構“協調安全、Safety2.0 とは”　https://institute-gsafety.com/safety2/

信頼性（Reliability）	・アイテムが与えられた条件のもとで、与えられた期間、故障せずに、要求通りに遂行できる能力（JIS Z 8115：2019）	・「安全性」と「信頼性」が混同されることが多いとの指摘もあるが、まずDependabilityとReliabilityの区別をしたうえで、Reliabilityについては「与えられた期間、故障せずに、要求通りに遂行できる能力」であることを踏まえ、安全、すなわち「許容できないリスクが無いこと」の程度と区別することが重要。	・各種の故障率のデータベース整備などはReliabilityに関する取り組みである。
総合信頼性（Dependability）	・アイテムが、要求されたときに，その要求どおりに遂行するための能力。アベイラビリティ，信頼性，回復性，保全性、及び、保全支援性能を含む。適用によっては，耐久性，安全性、及び、セキュリティのような他の特性を含むことがある。（JIS Z 8115:2019）	・新技術に対する社会的受容性の確保（変化の激しさ、想定・予測の困難さ、役割・責任の不明確さ） ・安全とセキュリティについては社会的受容性が必要であり、その議論のためには「損害が発生し得る事象の洗い出し」と「それが起きないことの対策」が不可欠。	・【DEOS協会】Open Systems Dependability：IEC62853

2.2. 日本の安全・ガバナンスの現状と課題

本節では、現状の法規制、施策、取組の分析を通し、Society 5.0における安全確保の実現に向けて、我が国の安全・ガバナンスの現状を俯瞰することを目的に行った文献調査、ヒアリング調査、分析の結果を示す。

まず、高圧ガス保安法を事例として、我が国の安全を確保するためのガバナンスの機能別の特徴を表 2-2 に示す（現行のガバナンスの機能ごとの特徴についての整理に至る分析は付録 A を参照のこと）。

我が国における規制は、一般的に国が法令によって行う規制の側面が強く、事業者自らの業界基準などによる自己規制は欧米に比較してまだ少ないと言える。しかし近年の技術進展は急速に速度を増しており、もはや国の法令による規制では技術進展に追い付かない状態となっている。このままでは、イノベーションに向けた産業界での動きに対して旧来の法規制が足かせとなるような状況が生じ、国際的な産業競争力を制約してしまう恐れがある。

表 2-2　現行のガバナンスの機能ごとの特徴の整理

- ルール形成
  - 国家が法規制を制定し、特定の規制対象（業）を定め、当該規制対象に対してどのような行為をすべきか（行為義務）を事前に規定する体系となっている。産業保安に係る規制体系については、一部において、事業者の保安能力に応じたインセンティブ措置（※）など、能力やリスクに応じた制度的措置が導入されているものの、そうした制度整備は限定的であり、基本は、詳細で画一的な個別規制・事前規制となっている。産業分野や事業者の保安の成熟度や能力にかかわらず、一律的に多くの届出、許可、検査等の手続が求められる規制体系となっている。（届出・許可等の手続き件数は高圧ガス：24 万件/年、電力：22 万件/年、都市ガス：1.4 万件/年）[9]

    ※高圧ガス保安法：スーパー認定事業所・認定事業所制度、電気事業法：定期安全管理検査における検査期間延長に係るインセンティブ制度、液石法：認定販売事業者制度（集中監視システムを導入する事業者に対し業務主任者数や緊急時対応等の一部要件緩和）
  - 社会-技術システムの進化・成熟に応じて法律は単調に増加する傾向があり[10]、特に安全に関連する規制は事故を教訓とした改正により増加してきた。[11]
  - 戦後の日本政府は経済発展と国内産業の競争力強化のために、業を細分化したうえで注力する対象の産業を特定し、各産業における専門性を高めることによって競争力を高める戦略を取った。[12]こうして生まれた業法は、安全確保のための技術管理が未熟な事業者に対しては参入障壁として機能しているが、新たな産業構造に移行にあたり、イノベーションの創出を阻害する可能性が指摘されている。[13]
  - 社会に影響を与える程の問題が発生する度に、個別の問題に対処する業法が作られ、規制は

[9] 経済産業省　産業保安を巡る環境変化と課題（2021 年 2 月）
https://www.meti.go.jp/shingikai/sankoshin/hoan_shohi/sangyo_hoan_kihon/pdf/001_01_01.pdf
[10] 榎並利博, “立法爆発とオープンガバメントに関する研究 -法令文書における「オープンコーディング」の提案”, No.419, 研究レポート, 富士通総研, 2015
https://www.fujitsu.com/downloads/JP/archive/imgjp/group/fri/report/research/2015/no419.pdf
[11] 中村昌充, “化学プラントの安全目標”, 学術の動向, 21 巻 3 号, 2016
[12] 大橋弘「新しい産業」政策と新しい「産業政策」, 経済産業研究所
[13] 経済産業省「GOVERNANCE INNOVATION VER.2: アジャイル・ガバナンスのデザインと実装に向けて」
https://www.meti.go.jp/press/2020/02/20210219003/20210219003.html

段階的に強化されてきた経緯から、全体最適・整合という観点での法設計は行われていない。我が国では一つのプラントが複数の法律の適用を受け、所管省庁もそれぞれ異なるという形で規制されている。そのため、プロセスに携わる事業所の安全管理部門の多くは、その手続きに多くのリソース・コストを割いている。[14]また、規制当局としてもプラントに係る保安４法（消防法（昭和 23 年法律第 186 号）、高圧ガス保安法（昭和 26 年法律第 204 号）、労働安全衛生法（昭和 47 年法律第 57 号）、石油コンビナート等災害防止法（昭和 50 年法律第 84 号））の合理化・整合化促進に取り組んでいる。

- モニタリング
  - 規制当局が法令違反の有無を確認するため、オペレーションの状況を年に一度、あるいは四半期に一度、といった一定期間ごとに監督する。データは、検査官等が実地に赴くことで収集する。
- エンフォースメント
  - 規制当局や裁判所は、企業に法令違反や権利侵害行為があった場合に、典型的には行為者に故意又は過失があったかどうかを判断し、それが認められる場合には法的制裁（一定の刑事罰・行政罰や許認可の剥奪等）を科す。

Society5.0 に向けた社会のシステム及び構造の変化に対して、現行のガバナンスモデルのどういった点が課題となるのかを特定し、課題間の関係を把握・可視化することを目的に、有識者ヒアリングを実施した。有識者ヒアリングの結果を踏まえ、プラント保安分野を事例に、我が国の現行の安全・ガバナンスの課題の構造化を実施した（図 2-2）。図 2-2 に示す日本の安全・ガバナンスの課題の構造化についての整理に至るヒアリング調査は付録 B を参照のこと。

また、図 2-2 の下部で整理した「ルールベースの法規制の限界」の構成要素については、具体的な現行の法令を調査・参照し例示することで、安全に係る規制の見直しの進展の余地を提示した。法令調査の際には、2020 年 6 月に規制改革推進会議が示した「デジタル時代の規制・制度について」[15]を参考とした。本文書では、デジタル化の進展によって、従来の規制・制度では、十分に規制目的（個人の身体・財産の安全、プライバシー保護、国家の安全等）を達成できない事象が生じるため、規制・制度を適切に見直す必要があることについて言及している。さらに、「規制・制度の類型化と具体的な見直しの基準」として規制・制度の見直しの切り口として表 2-3 を示している。（表 2-3 では、スマート安全 PJ と関連のある項目のみを抜粋した。）

[14] 大野 晋　化学プロセスにおける安全規制の課題と今後の制度設計（2003 年 10 月）
http://shakai-gijutsu.org/vol1/1_317.pdf
[15] 規制改革推進会議“デジタル時代の規制・制度について”
https://www8.cao.go.jp/kisei-kaikaku/kisei/publication/opinion/200622honkaigi01.pdf

表 2-3　規制改革推進会議文書における安全に係る規制・制度の見直しの基準

安全規制のリスク把握を精緻化し、リスクに応じた規制・制度へ見直し

i. 施設等の安全管理については、人が目視、打音によって点検、検査等を行うことを原則としている規制が多い。高精度カメラ、ドローン、赤外線センサー等を用いて必要な情報を収集し、ＡＩ等を用いた画像認識・診断やビッグデータの分析、常時監視等によって、リスク評価の精緻化を行うことで、一律の手法による規制を見直すべきである。
例：目視、打音等を原則とするインフラ等の定期点検・検査の新技術による代替

ii. 点検、検査等を一定期間ごとに、特定の手法や一律の基準により義務付けている規制も多い。センサー等によるリアルタイムデータの把握やビッグデータの分析等によって、リスク評価の精緻化を行うことで、一律の基準の規制を見直すべきである。
例：浄化槽の保守点検頻度の見直し、車検制度の手続や基準の見直し

iii. 製品検査等において、製造プロセスでデジタル技術を用いた精緻なリスク管理が行われている場合には、検査自体を不要とすべきである。また、従来から行われている検査等の中には、経済社会の変化により、規制創設時の目的が失われてきているものがある。こういった検査については、その必要性を検証し、見直すべきである。

iv. 検査・点検等の結果を書面で記録することを求める規制は、安全管理の高度化、省力化、データ活用によるリスク評価の精緻化の観点から、確認、記録等のプロセスをデジタル化すべきである。

v. 安全管理を担当する責任者等を置くことや、講習、資格取得を義務付ける規制など、人が実施することを前提とした規制については、デジタル技術による補完・代替を認める観点から、義務付けの緩和等の見直しを行うべきである。

vi. 人による行為を前提とした車両等の免許制度について、自動運転システムの技術進歩を促す観点も含め、技術進歩に合わせて、必要な見直しを行うべきである。このような見直しにより、より高度な安全性の確保が可能となると同時に、安全規制のリスクバッファーを最小化することで管理者、利用者の負担低減が可能となる。

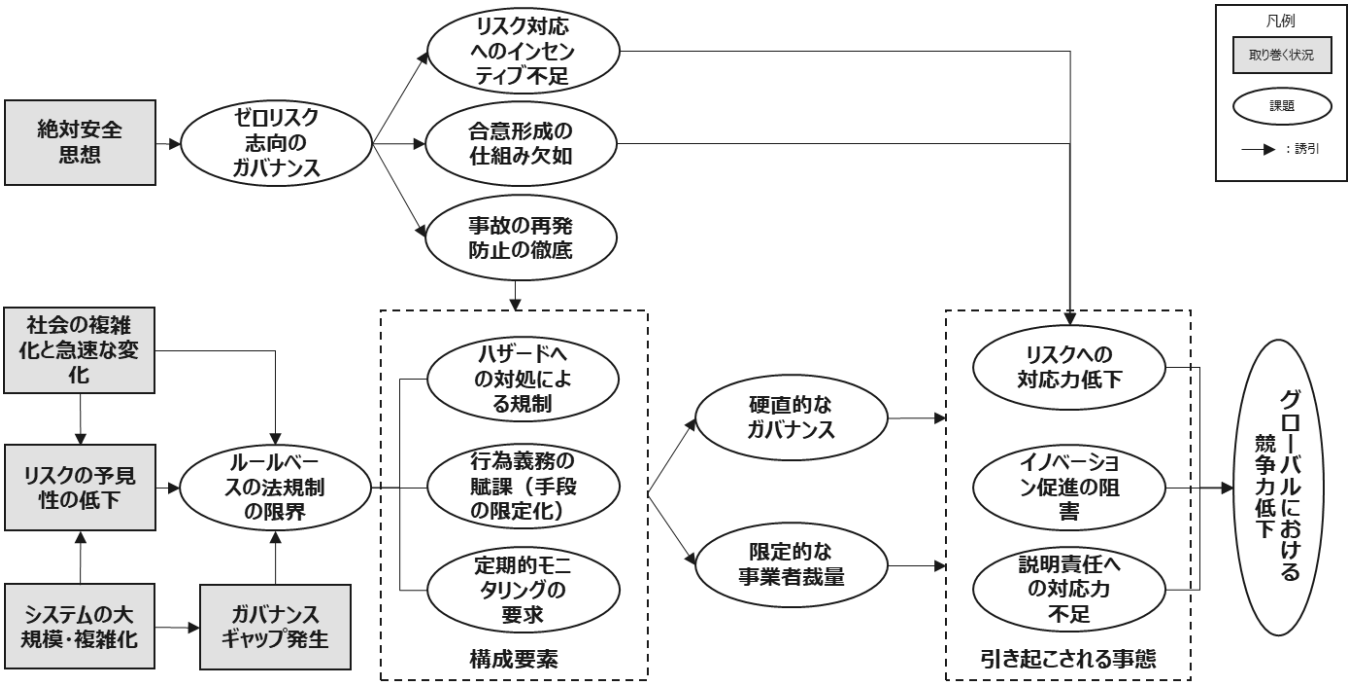


図 2-2　日本の安全・ガバナンスの課題の構造化

まず、図 2-2 の上部「絶対安全思想とゼロリスク志向のガバナンス」について解説する。我が国では、「安全であること」が意味するのは「危険性が一切存在せず、（危害が及ぶ）リスクがない」ことであると解釈される傾向があり、従来から、絶対安全やゼロリスク志向と呼ばれる安全の考え方が強く表れるガバナンスとなっていると言える。

その結果、主に以下の 3 点の状況を招いていると考えられる。

① 事故の再発防止の徹底：高圧ガス保安法を一例として検討を行ったところ、日本で求められている安全確保の考え方としては、まず、A)事故の再発防止の徹底を求めており、そのために、B)事故要因（ハザード）に対処することが基本的な方針となっており、C)機器・プロセス（ハード）ごとに分解しそれぞれに対して要求事項を規定している。
さらに、我が国ではそもそも法規制にリスクの概念が含まれていないゼロリスク志向であることからも、状況に応じた適切なリスクアセスメントやマネジメント[16]の実施が法規制で要求されていない固定的なリスク管理の方法（例えば、目視・打音の実施等の人による作業を前提とした方法）が適用されており、このことがデジタル技術活用の阻害、

[16] 指揮管理者が、対象とする組織についてトップダウン型で方針や戦略を実行し、経営資源（ヒト・モノ・カネ）の効果的活用により、利益を最大化させ組織を発展させる行為のこと。

事業者による自主的・先進的なリスクマネジメントの実施の阻害となっている可能性がある。

② 合意形成の仕組みの欠如：Society5.0 になると技術的なシステムが社会にもたらす影響はより広範に渡り、かつ不確実性が増大する。その際には、新たな技術・システム活用の際の不確実性について、多様なステークホルダーそれぞれに対して説明責任を果たすことによる社会受容性の確立は大前提となる。リスク情報を各ステークホルダーの立場から主張・共有し、合意形成を図る仕組みの必要性が高まる。また社会的受容性を考慮する上では、絶対安全思想の根強い日本社会において形成され定着している、一般市民のゼロリスク志向とリスクリテラシー不足が課題となる。

③ リスク対応へのインセンティブ不足：リスク対応の方法に柔軟性を持たせるリスクベースの規制とし、優れたリスク対応を自発的・主導的に行った主体がインセンティブを受けるような仕組みが必要だが、例えばプラント保安分野においては、ゼロリスク志向のガバナンスにおけるルールベースの均質の規制が前提となっているためにリスク対応へのインセンティブが創出しにくい。これによって保険等のリスクシェアの仕組みの整備も遅れている。

一方で安全に対する欧米での基本的な考え方は、「安全か安全でないか」という評価ではなく、「広く受容される領域」、「受容できる領域（ALARP[17]の領域）」、「受容されない領域」の 3 つの領域を捉えるものである。すなわち、「不安を取り除くためには、リスクはゼロにする」との考え方が根強い我が国での認識に対し、欧米が長い歴史の中で「リスクはゼロにはできず、すべてのリスクには対応できない」ということが定着している。このような国際的な状況を背景として、我々が行った有識者との議論では、日本においてもリスクはゼロではないことが社会に理解された上で、リスク対応について合意形成するための仕組みを含む、リスクベースのガバナンスへの転換が必要だという意見が強く示された。

続いて、図 2-2 の下半分「ルールベースの法規制の限界」について解説する。

課題を以下①～③の 3 点に整理した。それぞれについて「（課題に関連する）具体的な規定と見直しの余地の検討例」を示す。これは、法令類データベース e-gov（https://elaws.e-gov.go.jp/）を使用し、調査した結果である。

① ハザードへの対処による規制：前述の通り、従来、日本はハードウェアに一律の基準・検査を実行し事故を防止する、という安全確保を実施してきた。しかし、リスク状況や

[17] ALARP とは、As Low as Reasonably Practicable の略称。国際安全規格では、「受け入れ不可能なリスク」と「広く受け入れ可能なリスク」との間を「ALARP 領域」と定め、この領域にあたるリスクについては、合理的に判断して実行可能な限り時間とコストをかけてリスクに対する低減措置を行ったのであれば、便益を考えて許容されているとしている。

リスクの変化等を鑑みず、一律の保安レベルを要求しているため、必要以上の対応を行っている非効率な状況を生んでいる。状況に応じたリスクの保有を前提とした、より効率的・効果的なリスクマネジメントを実施する必要がある。また、ハザードへの対症療法的な考え方で追加されてきたためパッチワーク的構造となっており、法規制間の不整合が発生していることも課題となっている。

表 2-4　具体的な規定と見直しの余地の検討例

事業者や現状のシステムの保有するリスクの程度に関わらず一律の検査等の要件を課している下記の法令について、リスクに応じた柔軟なリスク管理手法やその支援技術の活用を認めることで事業者による創意工夫を活性化できる余地があるのではないか。

- 航空法：空港内の施設の維持管理指針では、事業者のリスク管理能力に関わらず一律の維持管理方法が定められている。事業者認定制度等の施行により、検査内容・報告義務等の簡素化を行える余地がある。
- 点検支援性能カタログ（橋梁、トンネル）：道路管理者が定期検査のために用いる機器の特性を、各設備に係るリスク、ハザードを踏まえて比較検討するために策定されたものであり、画像計測技術、非破壊検査技術について具体的な製品名が掲載されている。製品名で限定せず、製品の持つ機能で規定することによって、活用できる技術の種類・幅を拡大できる余地がある。
- 農産物検査に使用する機器として仕様が確認されている機器：農産物検査に使用する機器として仕様が確認されている製品の一覧が作成されている。これも製品名で限定せず、製品の持つ機能で規定することで、活用できる技術の種類・幅を拡大できる余地がある。

② 行為義務の賦課（手段の限定化）：日本の法規制は「なぜすべきか」といった達成目的・思想の記述がなく、「何をすべきか」という手段を縛る表現になっているために、新たなハザード、リスクの発現ごとに新たに規制体系の変更や要件の検討を行う必要が生じる。また、そうした変化に対応するためのシステム・ビジネスの開発や事業者の創意工夫による自主的な取組の推進が困難となっている。

表 2-5　具体的な規定と見直しの余地の検討（例示）

「目視」、「触診」等、人間の感覚で確認すること（「官能検査」という。）を前提にした措置について、最新技術での代替を容認する規定へと変更することで、事業者の自主的な取組や工夫が推進され、事業者により達成される安全のレベルや効率性を向上できる余地があるのではないか。

下記は、法令等の規定において点検・検査等の手法を制限するような文脈で「目視」、「触診」等、人間の感覚で確認することを前提にした手段が示されている規定の例である。

- 道路法施行規則（第四条の五の六）：道路の維持又は修繕に関する技術的基準等
- フロン類の使用の合理化及び管理の適正化に関する法律施行規則（第十四条）：フロン類の充塡に関する基準
- 温泉法施行規則（第一条の二、第六条の三）：技術上の基準
- 下水道法施行令（第五条の十二）：公共下水道又は流域下水道の維持又は修繕に関する技術上の基準等
- 河川法施行令（第九条の三）：河川管理施設等の維持又は修繕に関する技術的基準等
- 基準器検査規則（第十六条、第二百八十一条）：構造検査の方法、機構及び作用の検査
- 水道法施行規則（第十七条の二）：水道施設の維持及び修繕
- 船員法施行規則（第三条の九）：非常通路及び救命設備の点検整備
- 船舶職員及び小型船舶操縦者法施行規則（第百一条）：操縦試験の身体検査

③ 定期的モニタリングの要求：規制当局による安全確保状況等の監査は、定期検査・定期報告により主に実施されているが、非効率な側面やリアルタイムでの改善要求ができないことなどの課題があり、結局は事後対応に留まることとなる。遠隔監視、リアルタイムモニタリング技術の進展で先行指標（リーディングインディケーター）を活用した予兆把握、リスクの常時監視が可能となり、より効率的なモニタリングの在り方という動的なガバナンスへの移行が志向される。

表 2-6　具体的な規定と見直しの余地の検討（例示）

点検、検査等を一定期間ごとに、特定の手法や一律の基準により義務付けている規制（特定の期間を定めて定期検査を行うこととしている規定等）が多い。そのような規制について、センサーを用いたリアルタイムデータ把握を行っている場合や、ビッグデータの分析等の先進技術を活用して保安レベルを向上させている場合には、状況に応じて期間を柔軟化させられる余地があるのではないか。技術による常時監視の実現は、保安レベルを向上させることにもつながる。

下記は、法令等において検査・点検の実施時期・周期を制限するような形で規定されているものの例である。

- 浄化槽法（第十一条）：定期検査
- 下水道法（第二十五条の三）：排水施設の点検の方法及び頻度
- 水道法（第二十条）：定期の水質検査
- 鉄道に関する技術上の基準を定める省令（第九十条）：施設及び車両の定期検査
- 建築物における衛生的環境の確保に関する法律施行規則（第三条の二）：空気環境の測定方法

- フロン排出抑制法（第十六条）：簡易・定期点検の遠隔化
- 道路運送車両法（第六十一条）：自動車検査証の有効期間
- 国際航海船舶及び国際港湾施設の保安の確保等に関する法律（第十二条）：国際航海日本船舶の定期検査
- 計量法（第二十一条等）：特定計量器の定期検査
- 港湾法（第五十六条の二の二）：定期的な点検
- 消防法（第八条の二の二、第十四条の三）：防火対象物・屋外タンク貯蔵所又は移送取扱所の定期点検
- 土壌汚染対策法施行規則（第四十条、別表第八）：定期の水質検査
- 労働安全衛生法（第四十五条）：ボイラーその他の機械等の定期自主検査
- 河川法施行令（第九条の三）：河川管理施設の一年に一回以上の点検
- 船舶安全法（第五条、第十条）：定期検査、中間検査

経済産業省「GOVERNANCE INNOVATION VER.2: アジャイル・ガバナンスのデザインと実装に向けて」報告書（案）[18]にもある通り、ルールベースの法規制を中心とするガバナンスモデルは、事前に一定の行為規範を制定し、その履行態様を規制当局が外形的にモニタリングし、必要に応じて制裁を科すという方法が有効である社会を前提としている。しかし、Society5.0 のような社会になると、サイバー空間を起点として日々複雑な技術やビジネスモデルが開発され、それらがフィジカル空間へフィードバックされていくため、どこにリスクが存在するかを特定することや、どのようにリスクをコントロールすべきかを予め規定することが、極めて困難となる。そのため、現行のガバナンスのままでは Society5.0 におけるリスクへの対応力の低下は免れない。また、法益の達成方法に対して事業者による自由な裁量が認められず、イノベーションの阻害に繋がっている可能性が指摘されている。

さらに、ルールベースの法規制の下では、法の定める行為を企業が履行しているか否かを政府がモニタリングすることが想定されていたが、ゴールベースのガバナンスでは事業者自らが目的を達成できているかを監視し対外的に説明する必要性が生じる。（ルールベース規制アプローチとゴールベース規制アプローチについての説明は表 2-7 参照）こうした活動を事業者が自ら行う仕組みや、監査・認証機能を持つ第三者機関が外部から支援する仕組み、また規制当局側が審査を行う機能等が現状では整っておらず、説明責任への対応力不足も課題として指摘されている。

[18]経済産業省 https://www.meti.go.jp/press/2020/02/20210219003/20210219003.html

表 2-7　ルールベース規制アプローチとゴールベース規制アプローチとは

純粋なゴールベース規制とルールベース規制は存在せず、目的に応じてバランスを取りつつ、最適な組み合わせを図るものであり、相互に排他的ではなく、相互補完的な関係である。

- ゴールベースの規制（アプローチ）
  - ➢ 規制機関が正確なルールを指定するのではなく、ゴールを設定するアプローチ
    - ✓ ルールを指定するのではなく、ゴールを設定することに重点を置いている。（英,国立監査局）
      例）はしごを使用するか使用しないかを指定するのではなく、高所で作業する場合、人々が安全であることを目的とする場合がある。
    - ✓ ゴールベース規制とルールベース規制のアプローチの違いは抽象的には判断できず、法令遵守をしなければならない人の理解と実践に依存する。
- ルールベースの規制（アプローチ）
  - ➢ 次の属性のうち１つ以上を特徴とするもの
    - ✓ 規制行為を事前に立案し、特定性の高いルールを設定している。
    - ✓ 規制対象者がどのような行動ができ、どのような行動ができないかを規定しているルール。
    - ✓ 規制対象者にどのように遵守するかを事前に通知し、特定の事実関係にルールを適用する場合には、何も与えないか、または限定的な除外事項と限定的な柔軟性を提供するルール。
    - ✓ 規制者がどのような行為が許容されるかを事前に決定することを伴うルール。したがって、規制当局は、明確に策定された指令に事実を適用して、大部分が機械的な判断を行うことになる。

ルールの執行者は、規制された当事者がルールを遵守しているかどうかを判断するために、主に機械的な決定を行い、事実を収集する。

コラム①：Society5.0 の実現に向けたゴールベース規制アプローチとは？

複雑な設定に基づいて多種多様な行為を規範的に規制するルールベース規制アプローチ（以降ルールベースという）に対して、ビジネスの負担軽減とイノベーションの促進を目的としたゴールベース規制アプローチ（以降ゴールベースという）が、代替規制のアプローチとして世界各国で盛んに議論がなされている。（ゴールベースとルールベースの定義については表 2-7 を参照されたい）例えば、オーストラリアやニュージーランドでは、ゴールベースへの関心は 1990 年代後半から 2000 年代に遡る[19]。クリントン政権時代の米国では、

[19] Deighton-Smith, R. 2008. "Process and performance-based regulation: challenges for regulatory governance and regulatory reform." Minding the Gap: Appraising the promise and performance of regulatory reform in Australia. ANU Press. 89.

環境規制の分野でゴールベースへの移行を支持している[20]。OECD（経済開発協力機構）は、規制には可能な限りゴールベースを含むべきであるという立場を長年取っている[21]。

日本国内では、Society5.0 の実現に向けたルールの基盤として、ゴールベースへの移行が期待されている。2021 年に経済産業省が発行した「GOVERNANCE INNOVATION VER.2: アジャイル・ガバナンスのデザインと実装に向けて」報告書（案）のなかでは、様々なステークホルダーが、自らの置かれた社会的状況を継続的に分析し、目指すべきゴールを設定した上でガバナンスシステムをデザインし、早いサイクルによる継続的な改善を重視する「アジャイルガバナス」の考え方を示しており、ゴールベース規制アプローチについて次のように述べている。

“…Society5.0 において、伝統的な法規制のモデルは、①ルールが社会の変化に追いつかない、②外部からのモニタリングが困難、③エンフォースメントの対象の決定が困難、④一国の政府の権限の及ぶ範囲に限界がある、といった様々な課題に直面している。[中略]…そのためには、法規制を、従来型の業界別のルールベースではなく、機能別のゴールベースとし、企業に「何を達成すべきか」を明示する必要があると考えられる。“

ゴールベースのメリットは、規範的なルールを事前に設定しないことで、規制対象者と規制当局が「ゴール」の達成のために最善の方法を実践できることにある。そのため、Society5.0 と関連するような革新的な産業ドメインのなかで、ゴールベースの適用が特に期待されている。その一方で、ゴールベースには、明確に指定されないルールの曖昧さによる課題を克服する必要があり、杓子定規なゴールベースへの転換は実装に悪影響を及ぼすことが懸念される。従って、先行事例を踏まえて様々な点に留意する必要がある。

例えば、90 年代初頭のニュージーランドの建築規制法では、規則の増加と複雑さによって規制の一貫性が低下し、冗長な規制によって建築費が高騰する事態に陥っていた。このような状況に対して、ニュージーランド政府はゴールベースへの転換を図ったものの、規制当局と民間事業者双方の目標に対するパフォーマンスや業績目標を十分に定義しなかったために、4 万 2 千世帯の住宅に対候性の欠陥（雨漏りしやすい家の危機とも呼ばれる）が生じ、113 億ドルの追加のコストが生じたと推定されている[22]。

別の問題として、ゴールベースへの移行によって、民間事業者のコンプライアンスコストの上昇を伴う可能性がある。70 年代の英国の安全衛生法には 9 つの規制と 500 の詳細な規則群があり、その数は年々増加する傾向にあった[23]。このような状況に対して、英国政府は労働安全衛生法を段階的にゴールベースへ移行をさせる政策を取るとともに、民間事業者によるリスクアセスメントの実施義務を強化した。英国は労働安全衛生法をゴールベースに移行させることに成功しており、その結果は肯定的に捉えられている。しかしな

[20] May, P.J. 2003. “Performance Based Regulation and Regulatory Regimes: The Saga of the Leaky Buildings.” Law & Policy. 381-401.
[21] OECD (2012), Recommendation of the Council on Regulatory Policy and Governance, OECD Publishing, Paris
[22] Mumford, P. 2011. “Best practice regulation: Setting targets and detecting vulnerabilities.” Policy Quarterly. 7:36-42.
[23] Robens, 1972. Safety and Health at Work: Report of the Committee Cmnd 5034. HMSO London 1972.

がら、一部の企業では、考えられるすべてのリスクを網羅した長大な文書の作成と安全衛生コンサルタントへの支出によって、本来管理すべきリスクを犠牲にしていることが指摘されており、中小企業は大企業と比較してリスクアセスメントに 6 倍の費用がかかるようになったと評価されている[24]。

ゴールベースでは、規制当局がある程度の責任を民間に移譲して、民間企業の判断に基づく結果を受け入れることが必要となる。その際に、達成すべき「ゴール」と共に結果や業績目標の基準を明確にしなければ、規制の弱体化や過剰対応による企業のコンプライアンスコストの上昇を招く恐れがある。ガバイノ報告書では、ゴールベースへの移行に際して、①当初設定した政策目標を達成し得るものとなっているか、②社会状況の変化によって政策目標を変更する必要はないか、データに基づく評価と継続的な改善を行っていくべきであることを主張するとともに、マルチステークホルダーによるモニタリングの仕組みと併せて、ゴールベースを実現していくべきこと必要性を述べている。

[24] Löfstedt, R.E. 2011. Reclaiming health and safety for all: An independent review of health and safety legislation. CM8219.

2.3. 国内外でのデジタル技術活用、課題事例

本節では、デジタル化時代の安全の在り方・その変化を議論するために、国内外の活用事例、課題事例に基づき、デジタル化技術利用のポジティブ側面・ネガティブ側面について共通認識を図ることを目的に調査成果を取りまとめる。

IoT、ビッグデータ、AI、5G 通信によってサイバー空間とフィジカル空間を高度に融合させるシステムである CPS により、多様なニーズにきめ細かに対応したモノやサービスを提供し、経済的発展と社会的課題の解決を両立する Society5.0 が提唱されている。経済産業省「GOVERNANCE INNOVATION VER.2: アジャイル・ガバナンスのデザインと実装に向けて」報告書（案）で整理されている通り、Society5.0 における CPS の特徴としては、以下のようなものが挙げられる。

表 2-8　CPS の特徴[25]

① より大規模・広範囲・多種類のデータ収集（Digitalization）
② 高度かつ自律的なデータ分析（Analytics）
③ フィジカル空間への作用（Actuation）
④ 様々な機能を持つシステムの接続（Interoperability）
⑤ 地理的制約や業種の壁を超える拡張性（Augmentation）
⑥ 外部環境への変化への対応可能性（Adaptability）

これらの特徴によって、より安全な社会機能を提供できる可能性がある一方で、新たな安全上の課題が生じることになる。リスクを適切に低減し社会的受容性を確保するような、技術的な対応ならびにガバナンスが必要となる。

以下では、デジタル技術を活用した安全性向上や安全確保に係る生産性向上（2.3.1）とデジタル技術による安全に係る新たな課題・リスク（2.3.2）の 2 つの視点で、論点やその関連事例を記載する。また、具体的な分野として、医療機器分野、自動運転分野、金融分野を取り上げ、技術の活用討状況やガバナンス上の課題等を整理する（2.3.3）。

考慮すべきシステム特性はシステムの目的、活用され方等により異なり、かつ独立ではなく相互に関連するため、ここでは、安全に係る他の概念（信頼性、セキュリティ等）も含めて課題を整理する。

[25]経済産業省「GOVERNANCE INNOVATION Ver.2: アジャイル・ガバナンスのデザインと実装に向けて」報告書（案）（2021 年 2 月）https://www.meti.go.jp/press/2020/02/20210219003/20210219003.html

2.3.1. 安全に関連するデジタル技術活用例

本節では、デジタル技術を用いて、人間や従来の方法を支援または代替することで、より安全性を高めることや、安全確認に係る生産性を向上させることを実現している実事例を基に、デジタル技術活用のポジティブ側面に着目する。

（1） IoT/AI による大規模・リアルタイムのデータを用いた安全監視・予測

従来の安全監視では、遠隔で計測できるデータは一定量あるものの、設備の安全を詳しく確認するには人間が現場に立って五感を使って総合的に判断する必要があった。今後はセンサーやロボット、無線通信技術等の高度化により、データ取得の頻度や範囲は飛躍的に増加する。時空間的にデータの解像度が向上することによって、リスクを精緻かつリアルタイムに可視化できるようになり、それをもとにしたより合理的な判断が可能となる。

また、従来は判断者の経験や知見によって安全に係るシステムの状況把握・判断が行われてきたが、今後はリアルタイム・広範囲のデータを活用した処理を高速に実施できるようになり、さらに深層学習の利用によって状況認識性能が高度化することから、人間が実施するよりも格段に広範囲・早期かつ客観的な安全確認や異常発見ができるようになる。さらには、リアルタイムモニタリング等で得たデータを用いて、シミュレーションを行うことで、将来の事故や危害の予測を行い。それに基づく対策を実施できるようになる。

例えば、プラント保安やインフラ保安といった分野では、保安の高度化及びオペレーション・メンテナンスコストの合理化のため、ビッグデータ及び AI を用いた異常検知や故障予測による RBM（リスクベースマネジメント）・CBM（コンディションベースマネジメント）の実現に向けた取り組みが行われている。例えば、プラントにおけるドローン搭載カメラの画像を用いた AI による画像診断（フィルタリング）技術や複数のプラント運転パラメーターを常時モニタリングにより異常を検出する技術が活用され始めている[26]。

鉄道分野では、車上機器と地上機器間の無線通信によって、列車の正確な位置を把握し、それにより列車の減速や自動停止等の運行制御が可能となっている[27]。自動車分野においては、交通の安全、事故防止を目的とした車車間通信や路車間通信の研究・開発が行われている。

詳細な解像度のデータを利用することができるようになると、マネジメント対象となる個体の特徴を考慮した対策を行うことが可能となる。例えば、健康状態や生活環境に応じて個人の健康リスクが可視化されることによって、最適な医療サービスや保険サービスの選択が可能に

[26] スマート保安官民協議会 高圧ガス保安部会「高圧ガス保安分野スマート保安アクションプラン」（2020 年 7 月）
https://www.meti.go.jp/shingikai/safety_security/smart_hoan/koatsu_gas/pdf/action_plan.pdf
[27] 参考 https://www.jrea.or.jp/page/pdf/JREA201008-048.pdf

なると考えられる[28]。他方で、個別のプラントの潜在リスクが可視化されることにより、個別のリスクに応じたインセンティブ規制が可能になると考えられる。

（2） データの活用による社会全体の安全性向上

CPS では安全に関する状況や対処の記録もデータとして保存可能となるが、安全確保の失敗・成功等の知識・事例を社会で共有することにより、個人・個社の経験やノウハウを社会全体で生かすことができるようになると想定される。

詳細な事故シナリオに関するデジタルデータが蓄積され、これを協調領域データとして利用し、AI の判断により事故が起こらないように誘導を行うことができれば、再発防止の可能性は格段に高くなることが期待される。旭化成が主導して、石化協賛同各社より収集したデータを活用した保温材下腐食（CUI）の予測モデルの開発が実施されている[29]。

また、社会あるいは産業全体でデータを活用する基盤を共有することで、社会・産業全体の安全性向上が期待できる。日本、中国、韓国等では、新型コロナウイルスの感染者及び接触者把握のためのアプリケーションにより、感染者の対策や抑制が行われたこと[30]もその一例だといえる。

医薬品や自動車部品の製造においては「安心・安全なモノづくり」の重要性が高まっており、「正しく作られたのか」を証明するため、製造データが改ざんされていないことを示すような仕組みの導入も始まっている。日本通運では、安心・安全な医薬品供給を可能とするグローバルサプライチェーンネットワークを、IoT やブロックチェーンの技術を用いたプラットフォームを基盤として構築することを目指している[31]。

（3） 機械と人間の協調によるより高度な安全

機器等の操作において、従来は一定の割合でヒューマンエラーの発生が伴うものであったが、ICT 技術により自動化が進むことにより、ヒューマンエラーをゼロに近づけることができる。機械が担える部分は信頼性の高い自動技術に任せて、人間はより創造的な仕事へと注力できるようになる。実際にエレベーターやモノレールなどは過去には人間が同乗して操縦していたが、現在は自動化と遠隔監視化によって人間よりも高い安全性を実現している。

---

[28] 参考　総務省「ＰＨＲサービスモデル等の構築（事業期間：H28～30）」を参照
https://www.soumu.go.jp/main_content/000583121.pdf
[29] 参考　旭化成株式会社「旭化成グループにおけるスマート保安取組み事例」（2020 年 7 月）
https://www.meti.go.jp/shingikai/safety_security/smart_hoan/koatsu_gas/pdf/001_03_03.pdf
[30] 厚生労働省「新型コロナウイルス接触確認アプリ（COCOA）COVID-19 Contact-Confirming Application）
https://www.mhlPJo.jp/stf/seisakunitsuite/bunya/cocoa_00138.html
[31] 日本通運「安心・安全な医薬品を世界に届けるグローバルサプライネットワークを提供」（2020 年 8 月）
https://www.nittsu.co.jp/press/2020/20200831-1.html

自動車において、コンピュータ／ソフトウェアで実現する機能が著しく増大しており、安全性に関しても ABS（アンチロック・ブレーキシステム）やエアバッグ等による制御だけではなく、ネットワーク化（Connected）や自動化（Autonomous）により、人間では実現できないような高度な安全性を実現できるようになってきている。

従来は、人間と機械を隔離することで安全を担保するのが基本であったが（機械安全[32]）、今後は高度な ICT 技術を用いて人間と機械が共存した環境でも安全を確保することができるようになると考えられる。工場等に導入されるロボットが人体の進入を検知し、それに応じて低速動作や停止等の制限を行うことにより、ロボットの無駄な停止を回避したり人との共同作業を行ったりすることが可能となっている。そのようなコンセプトは、「協調安全／Safety2.0」として IEC 白書「Safety in the Future」[33]の中でも提唱され、セーフティグローバル推進機構が国際標準化や社会実装化を推進している。

人間がより安全に作業が可能となるような職場づくりのためにもデジタル技術は活用される。不安全な行動を検知したり、より安全な行動をガイドしたりすることによって、作業経験が少ない労働者にとっても安全な作業環境が実現できる。

事例として、JFE スチールでは、AI による画像認識を用いた作業者の安全行動サポート技術によって、立ち入り禁止エリアが変化する特殊な工場内においても、立ち入り禁止エリアへの作業者進入の検知と警報な発砲やラインの自動停止を可能とする技術が開発された。[34]

他にも、川崎重工業では、熟練技術者が遠隔装置で操作した動きについて AI 技術を用いてロボットが学習することで、微調整が必要な熟練技術者の繊細な動きをロボットで再現している。これにより自動化できる範囲が拡大するほか、熟練技術者がロボットに記憶させた動きを新人技術者へ体感させ、技術伝承につなげることも可能になった。[35]

2.3.2. デジタル技術による安全に係る新たな課題・リスク

Society5.0 における安全のあるべき姿を考える上では、どのようにデジタル技術／CPS のメリット・社会的価値と上記のような安全に係るリスクとのバランスを図り、どのような技術的・組織的な対策をどのようなガバナンスのもと実施していくかを議論することが不可欠となる。

---

[32] ISO12100
[33] IEC「Safety in the future」（2020 年 11 月）https://www.iec.ch/basecamp/safety-future
[34] JFE スチール「国内業界初となる AI 画像認識による安全行動サポート技術の導入について」（2018 年 12 月）https://www.jfe-steel.co.jp/release/2018/12/181211.html
[35] 川崎重工業「遠隔協調で熟練技術者の動きを再現する新ロボットシステム「Successor」を販売開始 ―ロボット化が困難であった分野への新たなソリューション―」（2017 年 11 月）https://www.khi.co.jp/pressrelease/detail/20171129_1.html

その前段として、本章では下記の事項について、現在どのような課題認識の下でどのような議論がなされているかについて整理を行った。これらの議論を踏まえた上で、安全・ガバナンスのアーキテクチャ設計を考えていく必要がある。

- ✓ AI の信頼性と社会的受容性
- ✓ サイバーセキュリティ
- ✓ システムオブシステムズ（SoS）の総合信頼性

各事項について、「課題」、「（課題への対応に）関連する取組や議論」の二つの観点から調査・構造化を実施した。

（1） AI の信頼性と社会的受容性

機械学習等の AI 技術を用いたソフトウェアが従来人間の行っていた判断を行う場合、自動運転の物体認識における障害物の見落とし、医療画像診断における癌などの疑いの見落とし、食品工場ラインの混入物検知における異物の見落とし等、AI が結果的に誤った判断や非決定的な振る舞いを行うことが想定される。そういったリスクをゼロすることはできないことから、どこまで AI の判断を信用できるか、AI の判断ミスで事故が発生したら誰が補償するのかといった問題を解決する必要がある。

逆に、漠然とした不信感だけで AI を導入しないことはむしろ危険にもなり得る。例えば、肺の X 線画像での癌検知では、AI の方が人間の専門家を上回る精度で識別できたという研究報告がある。[36]ただし、AI と人間との信頼性の比較は、その前提条件によって適用限界が大きく変わるため、少ない報告から優劣を即断することはできない。

AI の活用を促進するためには、AI の信頼性の検証・認定手法を確立し、その判断や AI を用いたシステムの安全性が社会的に受容される状態とする必要がある。特に、AI による判断がフィジカル空間に作用する場合や安全クリティカルな領域に活用される場合等には、これらの問題はより重要となる。

以上を踏まえ、表 2-9 に AI の信頼性と社会的受容性に係る課題を整理し、課題への対応としてどのような取組や議論が行われているかを表 2-10 に整理した。

[36] Diego Ardila ほか「End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography」（naturemedicine、2019 年 5 月）https://www.nature.com/articles/s41591-019-0447-x

表 2-9　AI の信頼性と社会的受容性に係る課題[37]

分類		課題の内容
AI システムの技術的特徴に基づく課題・リスク	**AI システムのゴール設定の困難性、予測困難性、説明困難性**	・ AI システムは、「何が正しい動作であるか」を定義すること、要求された品質（安全性等）を満たすことを事前に保証することが困難。何か不具合な出力が生じた場合も、それがアルゴリズムに起因する問題なのか、入力されたデータの問題なのかを検証することが難しい。 ・ 人的被害や経済損失・機会損失などの悪影響を及ぼすリスクを低減すること、AI システムの品質・リスクを適切に評価検証できるようにする必要がある。
	AI システムの脆弱性	・ 機械学習を用いたソフトウェアの挙動は学習やテストに用いるデータに依存することから、誤データ、偽データ等によって意図しない出力が生じることが懸念される。AI に特有の攻撃に対する適切なセキュリティ対策が必要となる。[38]
	アルゴリズムの継続的な変化	・ 状況の変化に起因するデータ傾向の変化や運用開始時の追加的学習等により、実装後も常にアルゴリズムが変化し続ける。これにより過学習[39]減少等が問題となる。 ・ 開発・運用を一体化した AI システムの継続的なリスクアセスメントによる品質劣化の監視・対策が求められる。
自律的判断がもたらす事故の責任と法制度	**自律的判断がもたらすリスクと事故の責任**	・ 膨大なデータに基づく AI の判断結果が、人を介さずに直接フィジカル空間に作用するようになることから、人の判断を介在させないことによるリスクの管理をどのように行うかが問題となる。 ・ AI による事故の際の責任分解や被害者の救済等について新た

[37] 表 2-9 の作成において、主に以下を参考した。
・経済産業省 我が国の AI ガバナンスの在り方 ver. 1.0（AI 社会実装アーキテクチャ検討会 中間報告書）https://www.meti.go.jp/press/2020/01/20210115003/20210115003.html
・総務省「安心・安全で信頼性のあるＡＩの社会実装」に向けて https://www.soumu.go.jp/menu_news/s-news/01iicp01_02000091.html
・IPA　ＡＩ社会実装推進調査報告書
https://www.ipa.go.jp/files/000067229.pdf

[38] AI ディフェンス研究所「機械学習セキュリティのベストプラクティス - Draft NISTIR 8269: A Taxonomy and Terminology of Adversarial Machine Learning -」（2020 年 7 月）https://jpsec.ai/nistir8269/

[39] 学習（訓練）データに過度に適合することで、学習（訓練）データでは正解率が高いのに学習（訓練）データとは異なるデータ（例えば、評価データ）では正解率が低くなってしまう現象をいう。つまり、学習（訓練）データだけに最適化されてしまって汎用性がない状態に陥ること。

		な考え方が必要となる。[40] ・ AI 同士の判断の干渉も問題となる。AI が行う多種多様の判断・行動の全てについて干渉リスクを点検することは容易ではない。
	法制度における課題	・ 人を前提とした現行の法律等のルールと AI 導入後の実態との間には乖離が生じることになる。挙動が予測困難な AI によって事故が生じた場合に、どのような条件のもとで行為者に過失があるといえるか、誰が責任を負うべきかについては、法制度や民事契約での整理が必要となる。[41]
AI に対する理解や社会的受容性		・ 人間の機能を AI に代替させる場合、判断精度が人間より十分に高いことを合理的に説明することが必要となるほか、社会実装においては、その上で社会の理解を得る必要がある。 ・ AI に対する理解不足やネガティブなイメージ、AI のブラックボックス性等から社会受容がなかなか進まないケースがある。

表 2-10　AI の信頼性と社会的受容性に係る取組や議論

分類	関連する取り組みや議論
AI 利活用・開発に係るガバナンス[42][43][44]	・ AI を用いたシステムの安全性を一定程度説明可能とするためには、AI を活用できる範囲・条件を定め、AI を用いたソフトウェアの開発プロセスの妥当性を説明する必要がある。 ・ AI を採用する製品・システム・サービスのリスク低減ならびに品質等の評価・検証のため、AI の開発（設計・テスト等）や利活用に係る原則やガイドライン、リスクの評価・対策に係る仕組みや基準等が各国や国際機関等で検討されている。 （一般原則）

[40] 平野「AI ネットワークと製造物責任－設計上の欠陥を中心に」（2018 年 12 月）
https://www.soumu.go.jp/main_content/000592824.pdf
[41] 東洋経済オンライン「自動運転ついに解禁、事故の責任は誰がとる？」（2020 年 5 月）
https://www.yomiuri.co.jp/fukayomi/toyokeizai/20200512-SYT8T1214728/
[42] 参考　国内外の議論及び国際的な議論の動向 https://www.soumu.go.jp/main_content/000646182.pdf
[43] 参考　人工知能（ＡＩ)を活用した都市と企業のガバナンスについての一考察 ～システムアプローチと機械学習工学による持続可能なスマートシティ実装～https://www.ieice.org/~swim/jpn/presentations/swim2019-24.pdf
[44] 参考　AI に関するルール・標準化の動向と今後の展望
https://home.jeita.or.jp/press_file/20191023145047_3Ezs15ATUG.pdf

	➢ AI の利活用や社会実装を促進するため、AI の便益の増進とリスクの抑制を図る一般原則が各国政府や民間団体から発表されている。 （例）OECD「AI 原則」、内閣府「人間中心の AI 社会原則」、欧州、米国、中国[4546] （国際標準化・ベストプラクティス） ➢ ISO/IEC JTC1 SC42 (Artificial Intelligence)による人工知能領域の国際標準化活動 ➢ IEEE-Standards Association：倫理に沿った自律/知的システムの設計指針（Ethically Aligned Design (EAD)）[47] ➢ 米国の NIST：機械学習セキュリティに関するベストプラクティス（NISTIR 8269） （日本国内の開発・利活用原則、ガイドライン等） ➢ 【総務省】AI 開発ガイドライン、AI 利活用ガイドライン、【消費者庁】AI 利活用ハンドブック ➢ 【QA4AI】AI プロダクト品質保証ガイドライン、【産業技術総合研究所】機械学習品質マネジメントガイドライン ➢ 個別分野（自動車、医療等）においても、AI の利活用や開発の指針・ガイドラインの整備が進む。
機械学習を用いたソフトウェアの安全性・信頼性を確保するための技術	・ 機能安全規格（IEC61508）では、現状は AI を安全関連系に用いることは非推奨とされる。[48] ・ 自動運転などの“セーフティクリティカルな AI 搭載システム”に対する安全性立証技術の研究開発が進められている。[49] ・ 「判断プロセスがブラックボックスである」という問題点に対して、「説明可能な AI（XAI）」の研究が企業や研究機関で進められている。[50]

[45] 内閣府 統合イノベーション戦略推進会議「人間中心の AI 社会原則」（2019 年 3 月）
https://www8.cao.go.jp/cstp/aigensoku.pdf
[46] OECD「42 カ国が OECD の人工知能に関する新原則を採択」（2019 年 5 月）
http://www.oecd.org/tokyo/newsroom/forty-two-countries-adopt-new-oecd-principles-on-artificial-intelligence-japanese-version.htm
[47] IEEE 「ETHICALLY ALIGNED DESIGN」（2019 年 3 月）
https://standards.ieee.org/content/dam/ieee-standards/standards/web/documents/other/ead1e.pdf
[48] 参考 佐藤・川島「機能安全規格 IEC61508 の課題と改訂動向」（2019 年）
https://www.jstage.jst.go.jp/article/essfr/13/2/13_118/_pdf
[49] SEAMS Project「人工知能搭載システムの安全設計ガイドライン（SEAMS ガイドライン）」（2020 年 4 月）
https://www.seams-p.jp/data/SEAMS_Guideline_essential_20200403.pdf
[50] DARPA「Explainable Artificial Intelligence (XAI)」https://www.darpa.mil/program/explainable-artificial-intelligence

AI に対する理解や社会的受容性	・ AI の判断ミスで事故が発生したら誰が補償するか等について、法律の改正や契約・保険制度を含めて議論が行われている。 ・ AI の特性を踏まえその開発・利活用に係る損害の補填等を目的とした保険の仕組みの重要性が指摘されている。

（2） サイバーセキュリティに係る課題・リスク

IoT の進展、システムのネットワーク化により、サイバー空間の重要性が高まるにつれて、システムの不正操作につながる許容されていないアクセス、データ破壊や改ざん等による機器の誤動作、システムの停止を招き得る DoS（Denial-of-Service）攻撃等、サイバー攻撃がフィジカル空間まで到達するリスクが増大する。それにより、サイバー空間に起因するフィジカル空間における事故シナリオが増加することが想定される。

そのため、フィジカル空間とサイバー空間の融合した CPS においては、サイバーとフィジカルの境界でのセキュリティやサイバースペース上のデータの真正性確保がより重要になる。安全とセキュリティの枠組みの再構築が必要となると想定される。

以上を踏まえ、表 2-11 表 2-9 にサイバーセキュリティ係る課題を整理し、課題への対応としてどのような取組や議論が行われているかを表 2-12 に整理した。

表 2-11 サイバーセキュリティに係る課題

分類	課題の内容
制御系システムへのサイバー攻撃の増加[51]	・ 重要な社会インフラを狙ったサイバー攻撃が増加しており、電力、ガス、水道、石油・化学等のプラントや運輸・交通等における監視制御、機械・食品の工場の生産ライン等で利用されている制御システム等のセキュリティの重要性が増加している。 ・ 機能安全と制御セキュリティは、似ている部分があるものの差異も多いため、安全リスク分析とセキュリティ分析の視点の違いを考慮し、その矛盾・競合を解消する必要がある。機能安全、IT、OT、セキュリティの各専門家がより緊密に協力することが求められる。
IoT のセーフティ・セキュリティ	・ IoT により様々なモノがつながると、通信路におけるデータの改ざん、ウイルス感染等の悪意によるサイバー攻撃の標的が増大することになる。また、IoT 機器を踏み台にした攻撃も想定される。

[51] 参考 IPA 制御システムの情報セキュリティ https://www.ipa.go.jp/files/000073863.pdf

5253	・ 機器やサービスがネットワークでつながることで生じうる様々な脅威や、それらを原因とするリスクや被害を予め踏まえ、現実をデータに転換するセンサーや電子信号を物理運動に転換するコントローラ等のセキュリティ対策が重要となる。メーカーが想定していないつながり方や利用者の操作誤りなどにより、安全を脅かすような重大な事故の発生も予測される。
無線通信の安全・セキュリティ	・ 5G 等の無線通信について、電磁ノイズや混信や、モバイルネットワーク内外のセキュリティリスクが安全性へ影響を及ぼし得る。5G を利用したシステム全体の各構成要素におけるソフトウェア・ハードウェアを含む安全対策やセキュリティ対策が必要。
サイバー空間における主体やデータの信頼性の確保[54]	・ サイバー空間における主体の認証やデータの内容に対する信頼の確保は、従来以上に重要となる。サイバー空間におけるデータが、正しく生成されたもの（主張された通りのもの）であること（真正性）、改ざんされていないこと（完全性）を確保・証明する必要がある。 ・ 通信経路が保護されないことによるデータの漏洩、なりすまし等による不正な組織からのデータ受信、不正アクセスによるデータの改ざんや削除、データ処理側での情報資産の棄損等のリスクが考えられる。 ・ データの信頼性を確保するためには、追跡可能性の確保や第三者の監査、トラストアンカーを利用した認証、サイバーセキュリティ対応等が必要となる。
クラウド・OSS 利用に係るセキュリティリスク増加	・ OSS（オープンソースソフトウェア）やクラウドサービスを用いてシステムを構築することが一般的となったが、その脆弱性やセキュリティリスクを考慮した設計・運用が必要となる。

表 2-12　サイバーセキュリティに係る取組や議論

分類	関連する取り組みや議論
制御系システムに対するセキュ	・ 日本政府の取り組みには、内閣サイバーセキュリティセンター（NISC）の重要インフラ分野の情報セキュリティ対策[55]、経済産業省の産業サイバーセキュリティ研

[52] 参考　IoT の安全・安心の確保に向けた仕組みの構築 https://www.ipa.go.jp/sec/our_activities/iot.html
[53] 参考　IPA IoT におけるセキュリティの脅威と対策　https://www.ipa.go.jp/files/000049819.pdf
[54] 参考　デジタル・ガバメント閣僚会議決定　データ戦略タスクフォース第一次とりまとめ
https://www.kantei.go.jp/jp/singi/it2/dgov/dai10/siryou_a.pdf
[55] 参考　内閣サイバーセキュリティセンター「重要インフラの情報セキュリティ対策に係る第４次行動計画」

リティ対策	究会[56]等がある。 ・ IEC TR 63069 は、機能安全（IEC61508 シリーズ）と制御セキュリティ（IEC62443 シリーズ）の両立を目指したフレームワークであり、両規格の用語、概念、開発プロセスの矛盾解消を行い、システム仕様の競合解消等を図る定義・提案を行っている。[57]
IoT の安全性・セキュリティに関する指針やガイドライン	・ 海外では、NIST による「NIST Framework for Cyber-Physical Systems」、IIC による「Industrial Internet of Things Volume G4: Security Framework」、英国 DCMS による「Code of Practice for Consumer IoT Security」といった IoT セキュリティに係るガイドラインが策定されている。 ・ 国内において、NISC による「安全な IoT システムのためのセキュリティに関する一般的枠組み」では、安全な IoT システムのためのセキュリティに関する基本原則と取り組み方針を示し、従来のセキュリティに加えて「安全性の確保」の視点を重要視している。経済産業省による「IoT セキュリティ・セーフティ・フレームワーク」では、包括的に IoT のセキュリティ・セーフティの課題・要求事項等を整理している。
5G 等の無線通信の安全性・セキュリティに係る取り組み	・ 総務省管轄のサイバーセキュリティタスクフォースにおいては、「IoT・5G セキュリティ総合対策」を策定し、5G や IoT に関するセキュリティ対策の総合的な推進に向けて取り組むべき課題を取りまとめている。 ・ 海外において、5G ネットワークのサプライチェーンにセキュリティホールとなるようなリスクのあるベンダーの製品を排除するための措置が相次いで制定されている。
データのトラスト（信頼性）、データガバナンス	・ 日本政府は、安心・安全なデータ流通の実現のため、「データ・フリー・フロー・ウィズ・トラスト（DFFT）」の重要性を世界に発信・展開。「データ戦略タスクフォース」では、データの真正性や完全性等による信頼性（トラスト）を担保する枠組みの構築を進めることを提言。「Trusted Web 推進協議会」では、Trusted Web の基本アーキテクチャに係る議論を実施。[58] ・ データやサービスの連携・相互接続における標準規格・標準アーキテクチャ（GAIA-X 等）においては、セキュリティや信頼性が考慮されている。[59]

https://www.nisc.go.jp/active/infra/outline.htm

[56] 経済産業省　https://www.meti.go.jp/shingikai/mono_info_service/sangyo_cyber/index.html

[57] 三菱電機「セキュリティに関する標準化動向」 https://www.meti.go.jp/shingikai/mono_info_service/sangyo_cyber/PJ_seido/PJ_bunyaodan/dainiso/pdf/001_05_00.pdf

[58] 内閣府「Trusted Web 推進協議会」https://www.kantei.go.jp/jp/singi/digitalmarket/trusted_web/index.htmlf

[59] 「GAIA-X:技術アーキテクチャ」https://www.pari.go.jp/sip/htdocs/doc/aboutgaiax/aboutgaiaxjp.pdf

クラウドや OSS のセキュリティに係る取り組み	・ 経済産業省の産業サイバーセキュリティ研究会では、OSS の管理手法に関するプラクティス集の策定等を実施。[60] ・ クラウドサービスを導入する際の安全性評価基準及び安全性評価として、米国政府は「クラウドサービスに関するセキュリティ評価・認証の統一ガイドライン（FedRAMP）[61]」を策定。日本でも、米国政府の FedRAMP に倣い、調達基準を統一化した「政府情報システムのためのセキュリティ評価制度（ISMAP）」を作成。

（3） システムオブシステムズに係る課題・リスク

現在、情報技術を用いたサービスの多くが、個別に作られ別々の管理のもと運用されている複数のシステムが相互に接続された「システムオブシステムズ（以下 SoS）」となっている。大規模な SoS では、地域を超えた様々な規模のシステムに接続されており、それらが利用目的・ステークホルダーの考え方等に伴って独立に変化することにより、全体として予見性が低く、不確実性が高い状態となっていく。SoS を構成する個々のシステムは多様なステークホルダーによって管理・運用されるため、SoS が安全に運用されるためには、変化や危険事象の発生に迅速に対応するとともに、想定しないつながり、共通管理されていないモノのつながり等に起因する不確実性に対して、ステークホルダー間での合意形成、内外に対する説明責任の履行を継続的に行うことが求められる。

表 2-13 にシステムオブシステムズに係る課題を整理し、課題への対応としてどのような取組や議論が行われているかを表 2-14 に整理した。

表 2-13 システムオブシステムズに係る課題

分類	課題の内容
SoS としてのリスク想定の困難性	・ 様々な動的なシステム間の相互接続が進むと、システム全体として新たな挙動特性が発生する可能性がある。 ・ 個々には正常なシステム同士であっても、システム全体としては相互作用により異常が起きたり、特定のシステムの障害が設計当初では想定できないほど広範囲に影響したりする等、設計段階で予想し得なかった新たな危険性が発生する可能性があ

[60] 経済産業省「第 3 回 サイバー・フィジカル・セキュリティ確保に向けたソフトウェア管理手法等検討タスクフォース」（2019 年 12 月）
[61] 参考 米国「クラウドサービスに関するセキュリティ評価・認証の統一ガイドライン（FedRAMP）」 https://www.fedramp.gov/

	る。また、システムが複雑に相互接続されるほど、システムの遭遇する危険シナリオの網羅的な想定は困難となる。
許容リスク設定とステークスホルダ間共有の困難性	・ SoS においては、全体管理者が存在していたとしても（Directed SoS[62]）、ステークホルダーやシステムの構成が多様化、複雑化するため、従来の単一システムと比べて、安全要求の定義、管理が困難になる。また、全体管理者が存在しない SoS（Acknowledged, Collaborative, Virtual SoS）では、安全要求の管理の方法論が存在していない。 ・ SoS に関わるステークホルダーごとのメンタルモデルギャップが拡大することで危険事象[63]が生じる可能性がある。 ・ 構成システムの持つ他の目的との関係性も踏まえた安全に関する合意形成が必要であるが、SoS における「危険の許容／許された危険」の法規的な扱い方が不明確である。
リスク事象発生時の対応	・ SoS の運用・管理の独立性及びシステム複雑化等に伴い、リスクの想定と状況変化に関する役割分担、システム障害の把握とその原因追求が困難となる。また、障害検知や安全状態への移行に関する時間や責任が曖昧になる。 ・ 事故や障害発生後、短期間で責任分解できず、裁判や保険等の仕組みの維持が困難である。設計時にシステムのユースケースや仕様が確定しきれないため、既存の製造物責任（Product Liability）の考え方ではカバーしづらい。
運用開始後の環境に応じた状況変化	・ 環境に応じて個々のシステムの役割や機能が変化していく CPS においては、システム構成の変更や動作目的の変さらによるリスクの事前想定は困難であるため、開発時のリスク分析と評価に基づく安全対策だけでは安全担保が困難となる。 ・ 相互接続性、グローバル性・拡張性も踏まえ、危険事象発生後の回復力（レジリエンス）やオープンシステムディペンダビリティ等、システムのライフサイクルを踏まえた安全要求の合理化が必要である。運用開始後の状況変化への対応を想定し、開発・運用を一体化した継続的なシステムの状況把握及びリスクアセスメント、障害対応、要求定義の見直し、迅速なシステムの変更等、運用段階（ライフサイクル全体）での安全担保が求められる。

[62] ISO/IEC/IEEE 21841: 2019 による定義を参照
[63] 潜在的に危険な状況から危害が発生すること。危険事象の定義は ISO/IECGuide51 ほか、複数の国際規格で異なる定義がなされている。本書では適用分野により危害事象の解釈がことなるため個別領域の言及は行わない。（出展：ISO/IEC Guide51）

サプライチェーン全体でのセキュリティ確保	・ 企業間の様々なシステムが複雑につながった場合、サイバー空間とフィジカル空間にまたがるデータや通信機器を含めたサプライチェーン全体のセキュリティを確保することが必要となり、事業者はより高度なセキュリティ対応に取り組むことが求められる。

表 2-14　システムオブシステムズに係る取組や議論

分類	関連する取り組みや議論
大規模・複雑化するシステムの安全解析手法	・ STAMP（Systems-Theoretic Accident Model and Process）[64]は「アクシデントは構成要素間の相互作用から創発的に発生する（局所的な相互作用に隠れていたものが表面化して全体に影響を与える）」という事故モデルである。 ・ FRAM(Functional Resonance Analysis Method: 機能共鳴分析）は、レジリエンス・エンジニアリングにおける安全分析のための手法である。[65]
開放系総合信頼性（Open Systems Dependability）	・ 開放系総合信頼性に係る国際標準規格 IEC62853: Open Systems Dependability では、システムの不完全性や環境変化の不確実に対してもできる限りサービスを継続することを目的とした OSD のためのプロセスビュー（合意形成プロセスビュー、説明責任達成プロセスビュー、障害対応プロセスビュー、変化対応プロセスビュー）を定義している。 ・ 障害対応サイクルや変化対応サイクルを含めた反復的なプロセス（DEOS プロセス[66]）において、ステークホルダーによるシステム設計時や変更時の仮定（前提）に係る合意の形成と説明責任の遂行が、総合信頼性の実現のための基本的な考えとなる。システムの目的や環境の変化に常に対応し、バージョンの変更や派生システムの開発を適切に行っていることに対する説明責任が重要となる。
運用段階での安全論証	・ 運用段階に動的に変化していくようなシステムを考えると、運用段階でのリスク情報の創出とリアルタイム・継続的な安全証明（安全論

[64] 参考　IPA「はじめての STAMP/STPA」（2016 年 3 月）https://www.ipa.go.jp/files/000055009.pdf
[65] 参考　IPA「FRAM（機能共鳴分析手法）による成功学に基づく安全工学」（2018 年 8 月）https://www.ipa.go.jp/files/000068587.pdf
[66] DEOS http://deos.or.jp/technology/process-j.html

	証）が求められると考えられる。大きな事故を発生させず改善ループを素早く回すことが求められる。 ・ Fraunhofer IKS（ドイツのフランホーファー認知システム研究所）では、自動運転システムを対象に、Dynamic Safety Management という手法を提唱している。[67] ・ 運用中においても安全性議論を継続し、システムや環境の変化に対応可能とする DevOps アシュアランスケースが提案されている。[68]
サプライチェーン全体のセキュリティ対策	・ 経済産業省による、サイバー・フィジカル・セキュリティ対策フレームワーク（CPSF）[69]においては、「企業等の適切なマネジメントを基盤した各主体の信頼性の確保」、「フィジカル・サイバー間を正確に転写する機能の信頼性の確保」、「自由に流通し、加工・創造されるサービスを創造するためのデータの信頼性の確保」が必要とされている。

2.3.3. 具体的な分野におけるデジタル技術活用事例、課題

様々な産業分野で、上述のようなデジタル技術は既に活用されているもしくは活用が検討されている状況である。そこで、特にデジタル技術活用が盛んに議論されている医療分野・自動運転分野・金融分野について、技術の活用・検討状況、また技術活用によるメリットや技術、また技術の実用化に伴うガバナンスの課題・取組状況について、今後の安全・ガバナンスの在り方の検討の参考とすべきポイントを抽出し、整理した。

（1） 医療機器分野

主要な医療機器分野で活用されているもしくは活用が検討されている技術として、以下が挙げられる。

表 2-15 医療機器分野で活用されている・活用が検討されている技術

具体的な技術・システム	日本における活用状況、規制上の取扱	海外における取り組み例

[67] Fraunhofer IKS https://www.iks.fraunhofer.de/en/research/adaptive-safety-performance-management.html
[68] 電子情報通信学会研究会 DevOps アシュアランスケースによる自動運転システムの安全性保障（2021 年 1 月）
[69] 経済産業省「サイバー・フィジカル・セキュリティ対策フレームワーク（CPSF）」 https://www.meti.go.jp/press/2019/04/20190418002/20190418002.html

AI 自動診断システム	・ 2020 年 6 月に COVID-19 肺炎を検出する診断支援ソフトウェアが承認。 ・ 恒常的に性能等が変化する医療機器に対する承認審査が薬機法上で位置付けられた。[70]	・ （イギリス）AI は医療機器としては扱われていない。
オンライン診療・健康相談	・ 新型コロナウイルス感染症拡大に際し、時限的・特例的にオンライン診療の取り扱いについて整理。[71] ・ オンライン診療の適切な実施に関する指針を厚生労働省が発行。[72]	
医療データの利活用	・ 個人情報保護法第 76 条において個人情報への研究目的でのアクセスを認めている。 ・ 「医療情報を取り扱う情報システム・サービスの提供事業者における安全管理ガイドライン」を総務省・経済産業省が制定[73]。	・ （ポルトガル、エストニア、アイスランド等）医療情報を一元的に管理、研究・商業双方に利用可能な国は増加傾向 ・ （フランス）診療情報へのアクセスを特定のプロジェクトに認め、情報提供者にメリットのある仕組みを整備

自動運転分野における技術活用に伴うメリットとしては、自動運転によってヒューマンエラー[74]による事故の抑制が期待されること（現在の交通事故の多くは、運転者の操作ミスなどの人的要因によると言われている）や、車車間やその他システムとの通信により事故事例・運転状況・道路状況の共有によって危険状況の予測・回避が可能となることが挙げられる。

---

[70] 厚生労働省「医療機器の承認審査関連事項」（2019 年 5 月）
https://www.mhlPJo.jp/content/10801000/000507448.pdf.

[71] 厚生労働省「新型コロナウイルス感染症の拡大に際しての電話や情報通信機器を用いた診療等の 時限的・特例的な取扱いについて」（2020 年 4 月）https://www.mhlPJo.jp/content/000621247.pdf.
厚生労働省「オンライン診療の適切な実施に関する指針」（2018 年 3 月、2019 年 7 月一部改訂）
https://www.mhlPJo.jp/content/000534254.pdf.

[72] 厚生労働省「オンライン診療の適切な実施に関する指針」（2018 年 3 月、2019 年 7 月一部改訂）
https://www.mhlPJo.jp/content/000534254.pdf.

[73] 経済産業省「「医療情報を取り扱う情報システム・サービスの提供事業者における安全管理ガイドライン」（案）に対する意見募集の結果及び当該ガイドラインを取りまとめました」（2020 年 8 月）
https://www.meti.go.jp/press/2020/08/20200821002/20200821002.html

[74] 人為的過誤や失敗（ミス）のこと。 JIS Z8115:2000[1]では、「意図しない結果を生じる人間の行為」と規定される。

上記の自動運転分野における技術の実用化に関する、ガバナンスに係る課題としては、主に以下の点が指摘されている。

- AI 活用時の責任分担：プログラム提供者やデータ提供者が医師に対して果たすべき責任、医療サービスの中における責任分担の議論が必要[75]である。
  - 医師法：第 17 条において「医師でなければ、医業をなしてはならない。」とされており、医師がその判断の責任を負う。[76]
  - 薬機法：AI の特性に即した形での承認制度が改正薬事法で制定。[77]
- サービス提供者と受容者の情報の非対称性が大きく、合意形成が困難。
- 医療行為の「実施場所」に関する規制：現行の医療法では医療行為は医療機関等の場所で行うことが定められている。

（2） 自動運転分野

自動運転分野で活用されているもしくは活用が検討されている技術として、主に以下が挙げられる。

表 2-16 自動運転分野で活用されている・活用が検討されている技術

具体的な技術・システム	日本における活用状況、規制上の取扱	海外における取り組み例
センシング（カメラ、ソナー、Lidar 等）	・ 2019 年５月に道路運送車両法及び道路交通法が改正され、2020 年４月の法施行以降、政府目標である高速道路での自動運転（レベル３）の市場化が制度上可能となったものの、いかなる場合においてもドライバーが即座に運転を交代できる状況であることが前提になっており、最終的な責任	・ （ドイツ）Pegasus Project[81] 自動走行システムの期待性能水準と評価基準の明確化
情報認知（画像解析、AI）		・ （日米）The Automated Vehicle

[75] 松尾剛行「健康医療分野における AI の民刑事責任に関する検討 ：AI 画像診断(支援)システムを中心に」Law and practice (13) p151-181, 早稲田大学大学院法務研究科臨床法学研究会（2019 年 9 月）,p174-175.
[76] 厚生労働省「人工知能（AI）を用いた診断、治療等の支援を行うプログラムの利用と医師法第 17 条の規定との関係について」（2018 年 12 月）https://www.mhlPJo.jp/content/10601000/000468150.pdf
[77] 厚生労働省「医薬品、医療機器等の品質、有効性及び安全性の確保等に関する法律等 の一部を改正する法律（令和元年法律第 63 号）の概要」https://www.mhlPJo.jp/content/11120000/000665345.pdf
[81] PEGASUS, 「THE PEGASUS METHOD」, https://www.pegasusprojekt.de/en/pegasus-method,

操作判断（AI）	はドライバーにあるという考え方となっている。[78] ・ 官民 ITS 構想・ロードマップ 2020 の策定[79] ・ SEAMS Project[80]では、自動運転分野を主な題材として AI 搭載システムの安全設計ガイドラインの策定、安全分析手法の開発、安全対策のソフトウェアの開発等を実施した。	Safety Consortium (AVSC) [82] 「レベル 4」「レベル 5」の自動運転車に対する業界標準確立
ダイナミックマップ		・ （欧米中）Safety First for Automated Driving (SaFAD) [83] 安全な自動運転技術(レベル 3, 4)を開発するためのガイドライン
路車間・車車間通信		

自動運転分野における技術活用に伴うメリットとしては、現在の交通事故の多くは、運転者の操作ミスなどの人的要因によると言われており、自動運転によってヒューマンエラー[84]による事故の抑制が期待されることや、車車間やその他システムとの通信により事故事例・運転状況・道路状況の共有による、危険な状況を回避することが可能となることが挙げられる。

上記の自動運転分野における技術の実用化におけるガバナンスに係る課題としては、主に以下の観点が指摘されている。

・ 多様な運転環境への対応[85]自動運転を多くの環境で活用するためには、様々な道路条件（道路の種別、車線数等）、地理条件（土地利用や周辺環境）、環境条件（日時、天候、温度等）、その他（速度制限、インフラ協調の有無等）に対応する必要がある。例えば、急な降雨によってカメラの画像にノイズが生じるケース等は自然現象による偶発的なものであるため対策の効果や再現性の体型的な評価が難しい。また、店舗の看板を

[78] 警察庁「道路交通法の一部を改正する法律」https://www.npa.go.jp/bureau/traffic/selfdriving/trafficact.pdf
[79] 高度情報通信ネットワーク社会推進戦略本部・官民データ活用推進戦略会議, 「官民 ITS 構想・ロードマップ 2020」, 2020 年 7 月 15 日, https://www.kantei.go.jp/jp/singi/it2/kettei/pdf/20200715/2020_roadmap.pdf,
[80] SEAMS, https://www.seams-p.jp/
[82] Automated Vehicle Safety Consortium, https://avsc.sae-itc.org/
[83] 櫻井(MONOist), 「自動車の“安全”を考える、ISO 26262 の先にある「SaFAD」にどう対応すべきか」, https://monoist.atmarkit.co.jp/mn/articles/2008/27/news008.html,
[84] 人為的過誤や失敗（ミス）のこと。 JIS Z8115:2000[1]では、「意図しない結果を生じる人間の行為」と規定される。
[85] 株式会社デンソー, 「機械学習の安全性への取り組み－自動車業界の取り組みを中心に－」, JSAI2019 企画セッション KS-5, 2019, https://www.jst.go.jp/crds/sympo/201906_JSAI/pdf/04.pdf, 川端, 「自動運転システムの性能限界検知と環境センシングに関する取組み」, Tier IV Tech Blog, https://tech.tier4.jp/entry/2021/01/13/160000,

交通標識と誤認するケース等、学習の前提条件から外れた状況に対応できないリスクは常に残されている。

- 外部システムとの連携[86]：今後、自動運転車は V2V、V2X にて様々な外部システムとの連携が検討されているが、下記のようなリスクに対して不正アクセスによる遠隔操作や不正データの意図的混入による AI の学習結果の汚染等を防ぐような仕組みが求められる。
  - ➢ 外部システムとの通信路に関するサイバーセキュリティリスクが想定される。
  - ➢ 個々が独立したシステム同士の相互作用により、予測できないリスクが生じる可能性がある。
- 社会的受容性[87]：自動運転車は AI 等の新技術が活用され、今までの自動車とは根本的に異なる製品となるため、普及にあたり新しい安全基準や市場との合意形成を検討する必要がある。また、自動運転車とドライバーの協調に課題があり、システムが運転を肩代わりすることによるドライバーの能力低下や AI の学習内容のブラックボックス化も懸念されている。
- 自動運転車による事故の責任[88]：自動運転車は人間の命に係わる事故を起こす可能性を持っており、AI のように結果の原因を必ずしも示せない機構を含む自動車が事故を起こした場合の責任分担が課題となっている。例えば、製造物中の AI の欠陥をユーザが立証できないため PL 法を活用できないといった問題が生じ得る。[89]また、運転の主体がシステムからドライバーに移行する間の事故に対する責任も曖昧なままである。

（3） 金融分野

金融分野で活用されているもしくは活用が検討されている技術として、主に以下が挙げられる。

---

[86] 倉地亮. "自動車を取り巻く IoT セキュリティーとその課題: 自動運転, 車車間, 路車間の V2X に関するセキュリティーを中心に." 情報管理 60.10 (2018): 690-700., https://www.jstage.jst.go.jp/article/johokanri/60/10/60_690/_pdf/-char/ja

[87] 髙橋・鎌田, 「完全自動運転車の社会的受容性」, DENSO TECHNICAL REVIEW, Vol.21, 2016, https://www.denso.com/jp/ja/-/media/global/business/innovation/review/21/21-doc-keynote-04-ja.pdf,

[88] 国立研究開発法人科学技術振興機構, 「近付く人間と人工知能の距離 互いに寄り添う新しい社会へ」, JST ニュース 2019 年 5 月号特集 2, https://www.jst.go.jp/pr/jst-news/backnumber/2019/201905/pdf/2019_05_p08-11.pdf

[89] 平野「AI ネットワークと製造物責任－設計上の欠陥を中心に」（2018 年 12 月）https://www.soumu.go.jp/main_content/000592824.pdf

表 2-17 金融分野で活用されている・活用が検討されている技術

具体的な技術・システム	日本における活用状況、規制上の取扱	海外における取り組み例
情報システムのオープン API 化	・ 新規プレーヤー参入による健全な市場競争促進等を目的とし、官が主導して推進。 ・ 公正取引委員会の主動によりオープン API 化が進展。	・ （ルワンダ中央銀行）データ収集システム
RegTech/SupTech	・ 上記の背景等もあり、規制対応のコスト低減等を目的として進行。	
分散型 ID・ブロックチェーン	・ G20 にてブロックチェーン技術に基づく分散型金融システムの課題を提起。「Blockchain Governance Initiative Network [BGIN]」が立ち上がり金融庁も参加。	・ （欧州）eIDAS 規制 ・ （スペイン）IdentiCAT プロジェクト ※eIDAS 規制に対応したデジタル ID[90]

金融分野における技術活用に伴うメリットとしては、従来人間が確認・作業していた部分をシステムで代替することにより、定常業務はシステムのほうがより信頼性が高く精度が向上することや、人間による作業についてシステムを通して行うため初期投資は係るものの作業コストを削減できること、必要な情報を必要なときにモニタリング可能になることにより安全性が向上すること、これまで特定の事業者に限られていた金融分野において技術活用により新たな事業者が参入することで市場の健全化が図られる効果が期待されている。

上記で整理した金融分野における技術の実装においての実社会適用に関するガバナンスに係る課題として、主に以下のような点が挙げられる。

[90] コインデスク・ジャパン「スペイン・カタルーニャ州、分散型 ID プラットフォームを市民向けに開発へ」（2019 年 9 月）https://www.coindeskjapan.com/20203/

- DX の進展と法規制整備のスピードの違い、に起因する、法規制の穴の存在：金融サービスのデジタル化、システム間の連携のスピードに法規制が追いつかず、そのギャップでリスクが顕在化（例：ドコモ口座問題[91]）
- 問題発生の度に追加され複雑化した規制への対応：様々な原因で生じた事故・事件に対処するため規制が次第に複雑化し、人間による対応に限界が発生
- DX 人材の不足、特にリーダーシップ・開発を担える人材が不足
- ガバナンスにおける規制・監査・保険の責任・役割分担

[91] 2020 年 9 月、株式会社 NTT ドコモ（以下、NTT ドコモ）が提供する電子決済サービス「ドコモ口座」を利用し提携銀行の口座から不正出金が行われる被害が発生した。この件では、「ドコモ口座」や NTT ドコモの回線の利用者以外に被害が生じた。NTT ドコモの本人確認や連携先銀行の認証手続きに甘さがあったことが指摘されている。
金融庁は、資金移動業者に対し、現行の犯罪収益移転防止法にもとづく確認の実施を求めたうえで、認証の仕組みやプロセスに関する脆弱性確認と脆弱性が認められた場合の改善を求めている。
https://www.fsa.go.jp/news/r2/sonota/20200915/20200915.html
また、全国銀行協会は、金融機関に対し「資金移動業者の決済サービス等での不正出金への対応について」として、同様に資金移動業者、銀行間の認証、資金移動者側の本人認証プロセスの脆弱性の確認と適切な処置を求めている。
https://www.zenginkyo.or.jp/news/2020/n091401/

2.4. 捉えるべき変化

2.2 及び 2.3 で整理した現行の日本のガバナンスの課題、技術調査、分野別調査結果より、技術の活用によって生じる安全上のメリット及び課題・リスクは以下のように整理できる（表 2-18、表 2-19）。調査結果からも分かる通り、安全性の向上が望める部分がある反面、技術によって生じる新たなリスクに係る課題が生じるため、そのリスクへの対応が必要となる。分野別調査結果から、新たなリスクに対応するために必要なガバナンスの在り方の変化についての示唆を抽出した（表 2-20）。

表 2-20 の項目を、「安全の考え方」、「法制度・ガバナンス」、「組織・技術・設備の管理」、「相互理解・説明責任」の 4 つの観点にグルーピングし、分野横断的な安全・ガバナンスのビジョンを図 2-3 に整理した。

図 2-3 で整理した安全・ガバナンスのビジョンを前提として、3 章では Society5.0 における安全・ガバナンスの在り方を検討する。

表 2-18　デジタル技術による安全に係るメリット

- IoT・AI による大規模・リアルタイムのデータを用いた安全監視や予測
  - データ活用による判断の精度向上、リスクの低減（医療分野等）
  - データ利活用による研究・サービスの質向上（医療分野等）
  - 規制コスト及び規制対応コストの削減（金融分野等）
- データの活用による社会全体の安全性向上
  - 事故事例の共有による安全性向上（自動運転分野等）
- 機械と人間の協調によるより高度な安全
  - ヒューマンエラーの減少、信頼性の向上（医療分野、自動運転分野、金融分野等）

表 2-19　デジタル技術による安全に係る課題・リスク

- AI
  - AI システムのゴール設定の困難性、予測困難性、説明困難性
  - AI システムの脆弱性
- サイバーセキュリティ
  - 制御系、IoT・無線通信のセキュリティ
  - サイバー空間における主体やデータの信頼性の確保
- システムオブシステムズ
  - リスク想定の困難性（自動運転等）
  - 許容リスク設定とステークホルダー間共有の困難性（医療分野、自動運転分野等）
- その他

➢ DX 人材の不足、特にリーダーシップ・開発を担える人材が不足（金融分野等）

表 2-20　ガバナンスの在り方の変化

・ アルゴリズム・システム間の関係や状況の継続的な変化への対応(医療分野等)
・ 問題発生の度に追加され、複雑化した規制への対応（金融分野等）
・ DX の進展と法規制整備のスピードの違い、法規制の穴の存在（金融分野等）
・ ガバナンスにおける規制・監査・保険の責任・役割分担（金融分野等）
・ 事故発生時の責任（医療分野、自動運転分野等）
・ 社会的受容性を目指すステークホルダー間の対話（自動運転分野）

コラム②：ALARP(as low as reasonably practicable)とは？

ALARP とは、as low as reasonably practicable の頭文字を取った言葉であり、リスクへの対応を判断する際の考え方である。リスクには、許容できないリスク・許容できるリスク・広く許容でき対策を必要としないリスクに分けられる。この許容できるリスクが ALARP 領域と呼ばれ、以下の場合においてリスクを許容できるとする領域である。

- 対策を必要としないリスクのレベルまで低減する技術がない
- 危険/効用やコストを考慮し許容しうる

ALARP の原則に則ったリスクマネジメントにおいては、リスクアセスメントの結果をこの ALARP に当てはめ、リスクへの対応は許容できないリスクを防ぎ ALARP 領域のリスクを可能な限り低減する形で実施する。

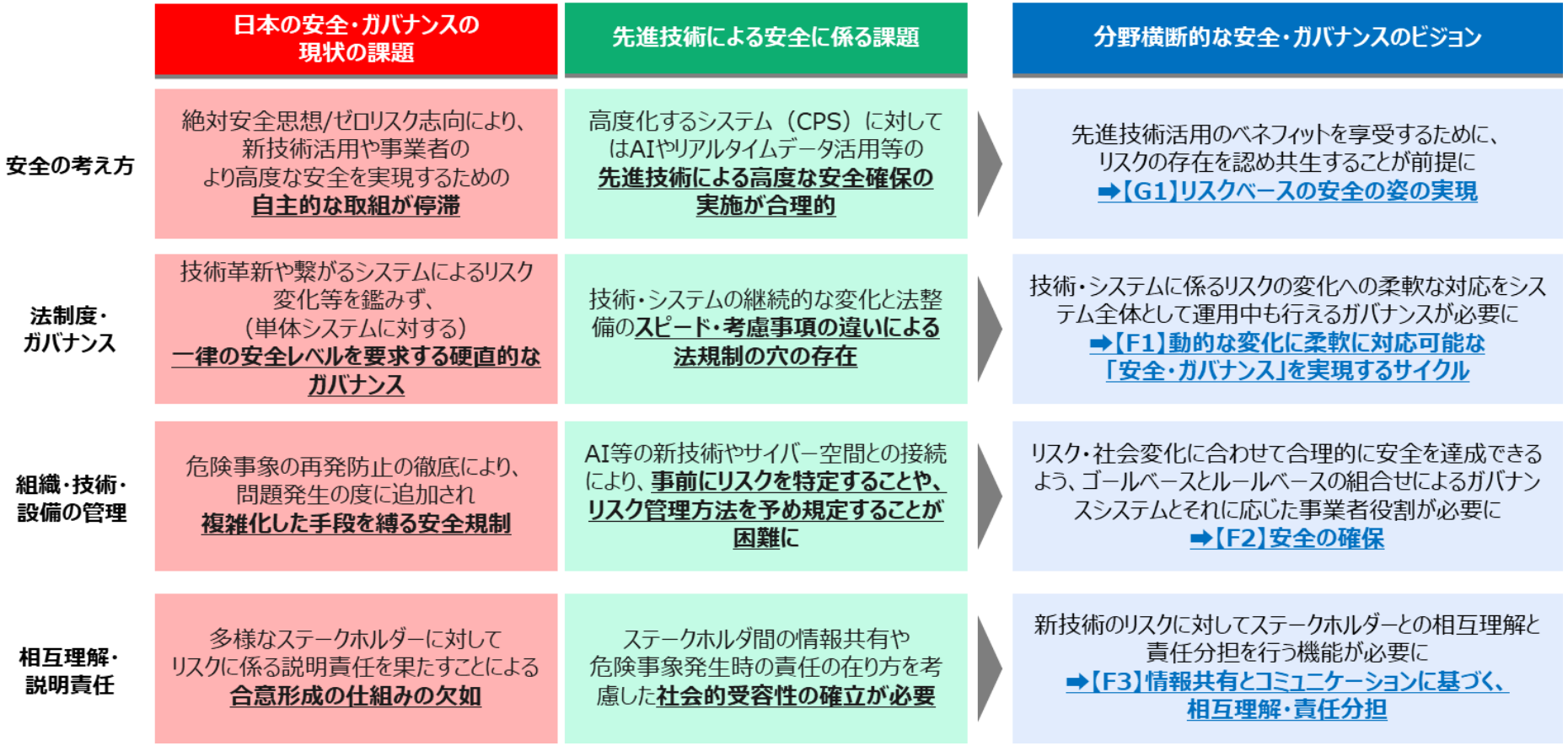


図 2-3　課題を踏まえた安全・ガバナンスのビジョン

3. Society5.0 における安全・ガバナンスのアーキテクチャのビジョン

本章では、Society5.0 における安全・ガバナンスのあり方についてのビジョンを整理する。基本的には、達成したい目的から安全・ガバナンスに必要な機能を導出してゆく（図 3-1 の「目指す安全の姿」を参照）。ただし、Society5.0 においては安全に関連する様々なシステムが CPS の一部として取り込まれることが見通され、2 章の調査・検討成果を前提にそのメリットとデメリットを考慮しながら（図 3-1 の「実現手段に係る環境の変化」を参照）、ガバナンスの機能及びアーキテクチャを設計する必要がある。

まずは Society5.0 における安全の目的・定義について確認を行った上で、Society5.0 における CPS の特徴を考慮しつつ、安全・ガバナンスのアーキテクチャの設計に向けて、ビジョンとして必要な機能の整理を行う。

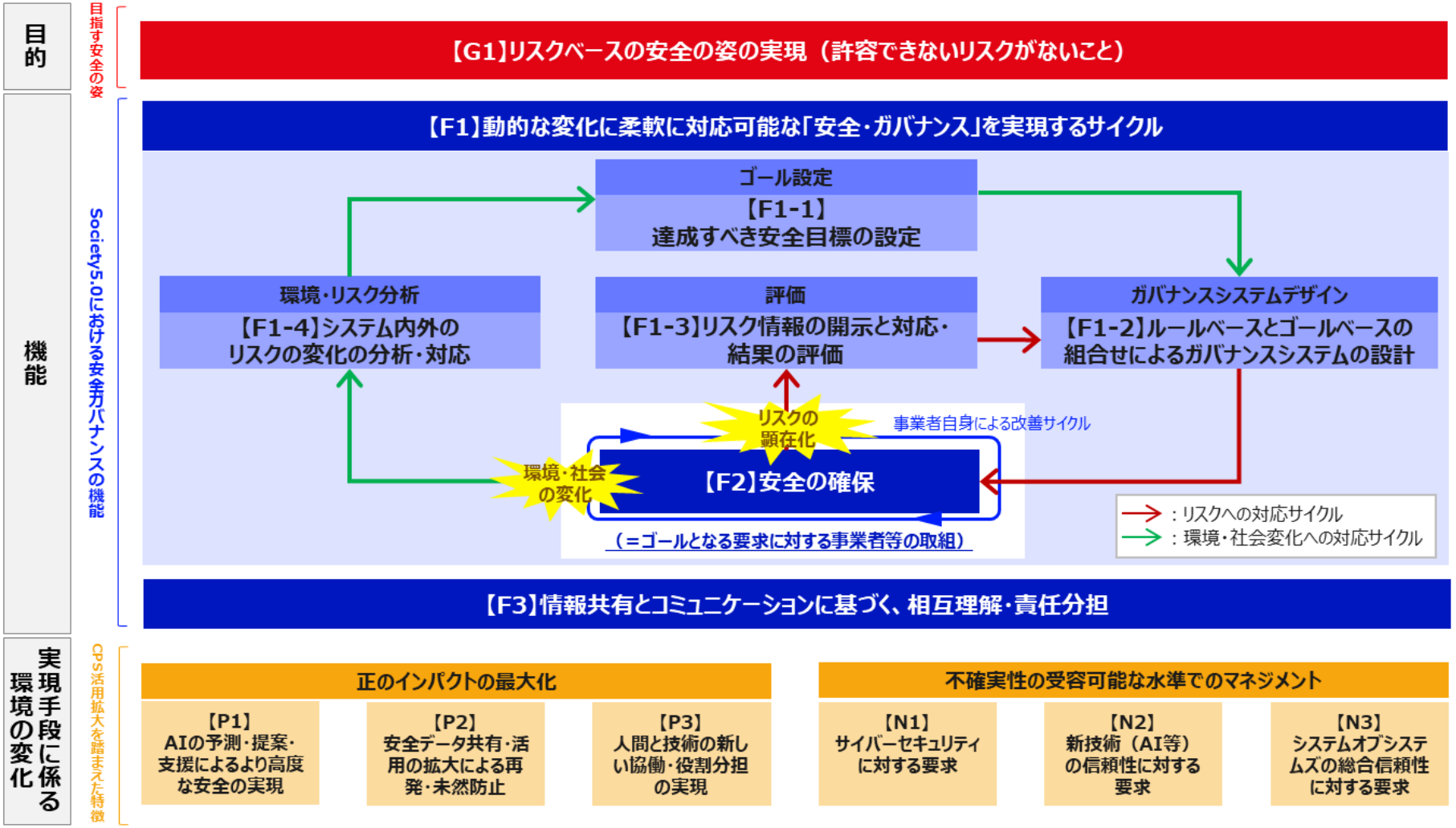


図 3-1 Society5.0 における安全・ガバナンスのアーキテクチャのビジョン

3.1. Society5.0 において目指す安全の姿

本節では、図 3-2 における目的レイヤーに位置づけられる目指す安全の姿が Society5.0 においてどのように変わるべきかについて検討する。

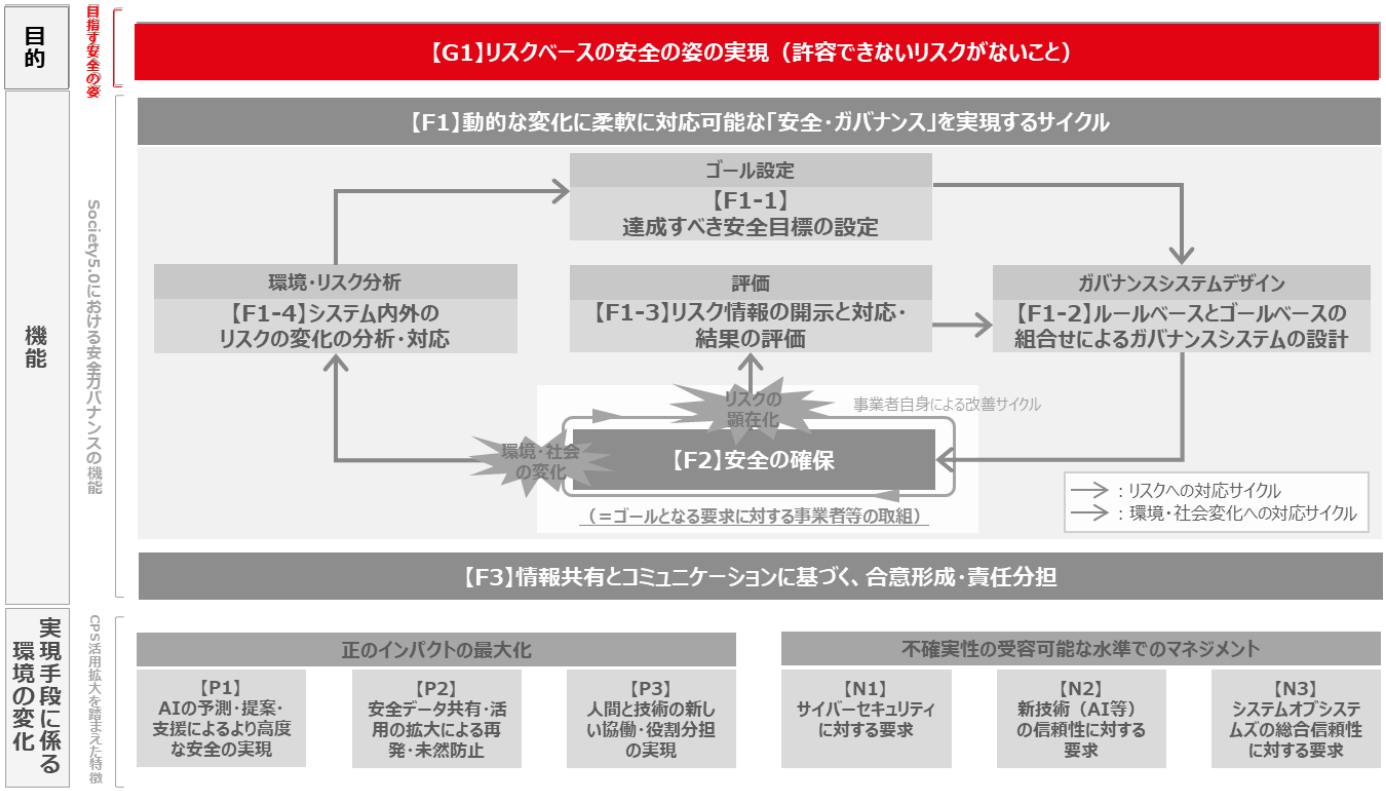


図 3-2　3.1 節のスコープ

3.1.1. 安全の定義

Society5.0 における安全を考える上で、安全という概念を再定義する必要性について検討した。有識者との議論の結果、安全の定義自体は基本的には Society5.0 になっても ISO/IEC Guide 51: 2014 の定義「許容できないリスクが無いこと」が変わらずに使用という意見で一致した。ただし、「許容」という行為自体に地域・文化による多様性が存在する点は留意が必要であることやリスクの捉え方に関しても再定義が必要であることの指摘があった。

この安全の定義からも明らかなように、安全とは許容可否という判断を伴う相対的な価値観を前提に含む概念であり、そのことは前述の「ALARP（As Low As Reasonably Practical）」という考え方にも表れている。この「許容レベル」は、組織が置かれた社会要求や時代それぞれの技術レベル（state of arts）、及び、事業状況によっても変化する。

すなわち、この定義を踏まえると、安全に関する対応判断を行うためには、局所的な安全問題を見ているだけでは現実的・最適な判断を行うことはできず、より確かな関連情報を幅広く収集した上で、その他の同時に発生する様々な影響も考慮し、許容レベル（保有してもよいと許容するリスクレベル）の判断を統合的に行う必要がある。Society5.0 において、デジタル技術を用いて情報の集約・リスクの可視化・常時評価を可能とすることは、社会システム、企業経営及び個人におけるリスクマネジメントの高度化において非常に有効なものとなる。

3.1.2. リスクベースの安全の姿の実現

上述の検討を踏まえ、Society 5.0 における安全・ガバナンスを検討する上での目標として、「リスクベースの安全の姿の実現」を掲げた。「リスクベース」とは、Society 5.0 の特徴を理解し、新技術による安全に寄与する効用を最大化するとともに、新技術の不確実性の影響を受容可能な水準で管理することを含意している。

つまり、Society5.0 における先進技術活用のベネフィット/オポチュニティーを享受・活用するために、残留リスクの存在を認め受容・共生することが前提となるため、従来のゼロリスク志向/絶対安全思想のガバナンスからリスクベースの安全を志向したうえで、ガバナンスの在り方を設計する必要がある、ということである。

特に Society5.0 においては、サイバー空間を起点として日々複雑な技術やモデルが開発され、それらがフィジカル空間へフィードバックされていくため、事前にリスクを特定することや、リスクのコントロール方法を予め規定することが、極めて困難となる。そのため、達成すべきゴールの設定やリスク基準とその達成手段を継続的に見直す仕組みをガバナンスに組み込むことがガバナンス機能を検討する際の前提となる。

3.2. Society5.0 における CPS 活用拡大の特徴

本節では図 3-3 に示す通り、Society5.0 において安全に関連する様々なシステムが CPS の一部として組み込まれることを前提に、ガバナンスの目的[92]である「正のインパクトの最大化」と「不確実性の受容可能な水準でのマネジメント」の両側面から、実現手段に係る環境の変化の特徴を検討する。

[92] 経済産業省「GOVERNANCE INNOVATION VER.2: アジャイル・ガバナンスのデザインと実装に向けて」報告書（案）では、Society5.0 におけるガバナンスを、「サイバー空間とフィジカル空間を融合するシステム（CPS: Cyber-Physical System）について、これによって生じるリスクをステークホルダーにとって受容可能な水準で管理しつつ、そこからもたらされる正のインパクトを最大化することを目的とする、ステークホルダーによる技術的、組織的、及び社会的システムの設計及び運用」と定義しており、それを参考とした。

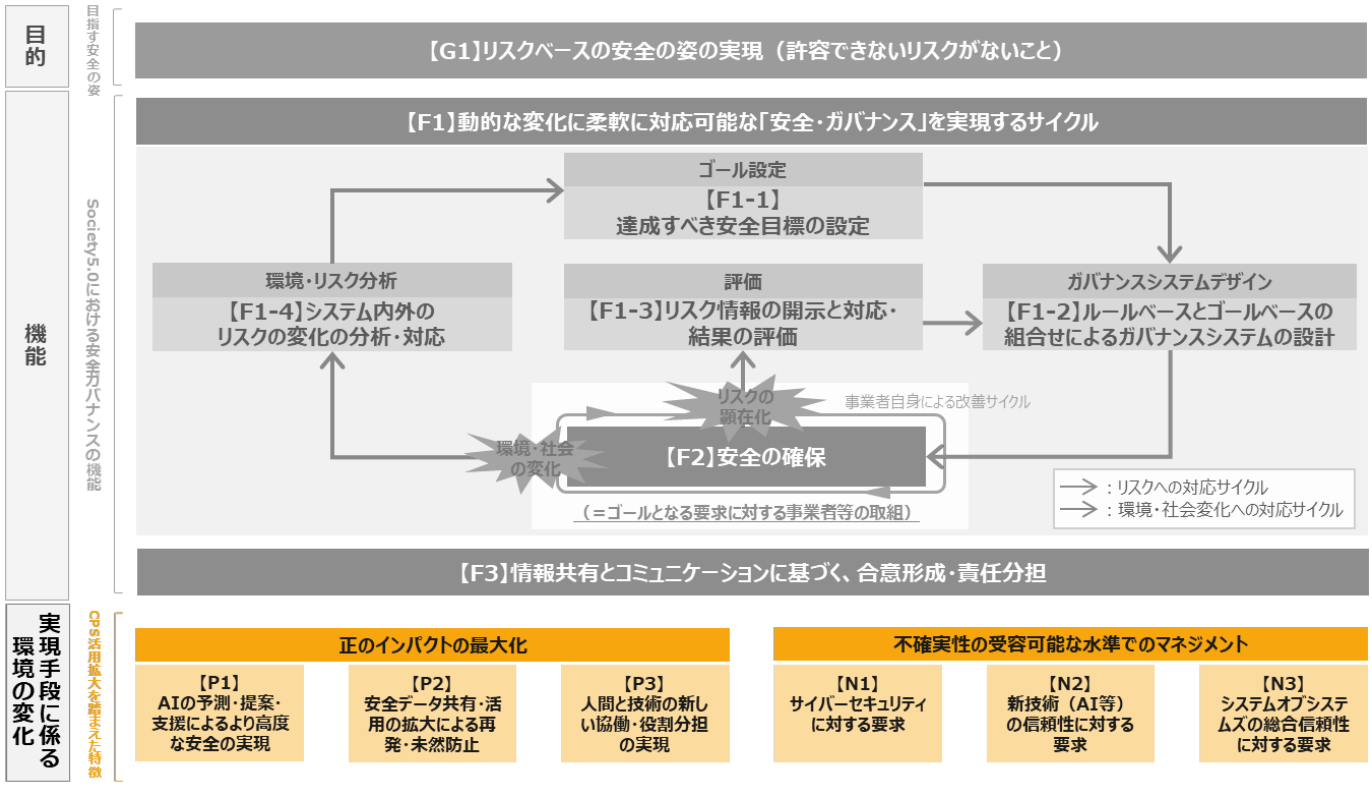


図 3-3　3.2 節のスコープ

3.2.1. 正のインパクトの最大化

（1）　AI の予測・提案・支援によるより高度な安全の実現

AI は「計算処理速度」及び利用可能な情報の「記憶容量」で既に人間を超えており、さらには、ドローンやロボットなどの自律的なセンサー搭載移動装置及び 5G 等の無線通信環境の発達により、更に広範なデータをリアルタイムで利用可能となってきた。

例えば、保安の分野においては、例えば従来は現場往訪等にリソースを要すために限られていた状況確認の頻度がセンサーやロボットの導入により飛躍的に高まり、場合によっては常時連続監視が可能となる。また 5G/6G の導入などの通信の高速化により、設置可能なセンサー端末の数及び利用可能なデータ数が劇的に増加する。このようにして利用可能なデータが飛躍的に増加することで、AI の判断能力が格段に向上することが期待される。その結果、従来は判断を担う人間の経験や知見によって判断基準・結果に個人差が伴うものであったが、AI を導入することによってそれらの属人的な差異は是正され、またデータの処理・判断の速度が従来よりも格段に向上することで、広範囲かつ早期での異常発見が可能となる。

また、単一システムとしての AI だけでなく、 AI 同士が自律的に交渉して動く「AI 間交渉[93]」の開発により、システムそれぞれの強み・弱みを補完しあう統合システムとして活用される将来も見込まれる。

技術の進展により、AI の活用領域は安全分野においても今後拡大（広範なデータに基づいた瞬間的な判断や、想定外の状況まで含んだリアルタイムシミュレーションによる安全確保方策の検討等）していくことが予想される。

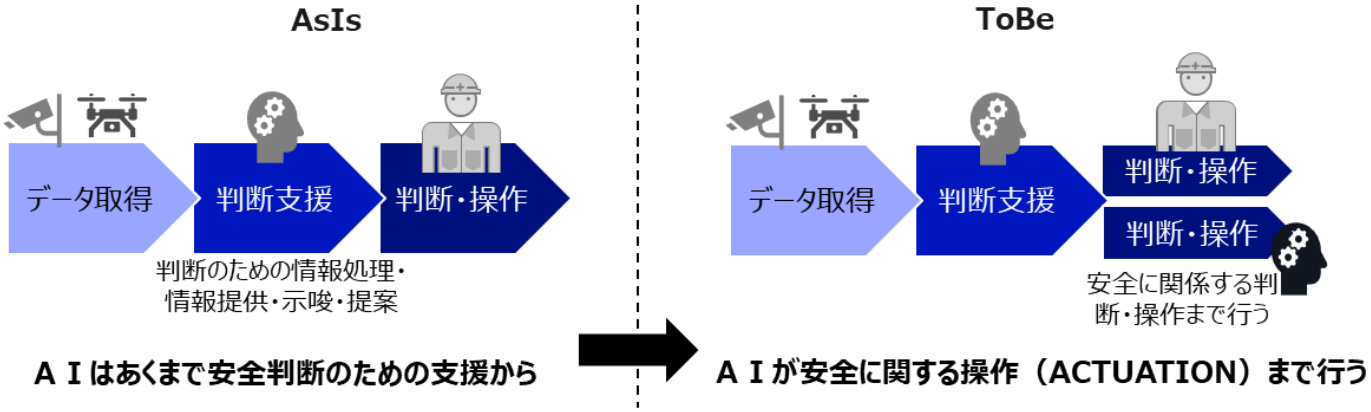


図 3-4　AI の予測・提案・支援によるより高度な安全の実現（イメージ）

（2）　安全データ共有・活用の拡大による再発・未然防止

安全対策は、未経験の事故を防ぐ「未然防止」と過去に経験した事故の再発を予防する「再発防止」に分類することができる。再発防止のためには、事故事例の分析を詳細に行い、原因と事故シナリオを特定し、それらが以後発生しないように対策を行うとともに、同様の設備に対して水平展開することが基本となる。水平展開の際には、可能な限り根本原因に近いレベルまで原因分析を行い、関連する様々な事業者で共有活用できるレベルに対策の抽象度を高めることがポイントとなる。一人の人間の場合には失敗経験は記憶として蓄積され、過去に自分が起こした失敗は繰り返さないようなフィードバックがかかるようになっているが、組織や社会の場合には同様のフィードバック機能を、意識的に設計・実装・運用していくことが必要である。

Society5.0 になり、人間行動等を含む様々なフィジカル空間の事象がデータとしてクラウド上で拡張性を持って共有されることにより、多種多様なデータを飛躍的に大量に活用可能となる。事故の発生シナリオについても詳細なデジタルデータが残ることになり、これを協調領域

[93] 参考　内閣府「戦略的イノベーション創造プログラム（SIP）第 2 期／ビッグデータ・AI を活用したサイバー空間技術」の研究開発項目

データや知見として活用することができれば、安全性の説明・論証能力は拡大し、失敗要因の相互共有により再発防止の可能性は格段に広がると考えられる。

特に安全への配慮が求められる分野・産業では従来も事故シナリオの分析・共有・活用の取り組み自体は行われてきたが、主に「事故データの不足」「企業秘密の壁」によって阻まれてきたといえる。畑村洋太郎氏の「失敗学」に代表されるような失敗要因の類型化・共有の取組みは既に行われてきたが、Society5.0 ではこれを CPS として社会実装することが可能であり、人類知とも呼べるような新たな社会安全デジタル技術基盤につながる可能性も考えられる。

また、このようなモノやサービスのデジタル情報をシームレスに繋げられる仕組みを実効的なものとするためには、事業者に対するデータ提供のインセンティブ設計を併せて計画する必要がある。

一方で、産業界においては、事業のライフサイクルを通じて生成・処理される情報の量的爆発により、モノとデータの不一致が物理空間における安全への影響を持つことを示す事例が世界規模で多発している。そのため、国際標準コミュニティ(ISO,IEC)や我が国の DFFT[94]の取組においても、データのセマンティックインターオペラビリティ[95]の実現が喫緊の課題と設定されている。データの共有活用による恩恵を享受するためには、目的間・組織間・地域間・産業間等のデータのインターオペラビリティへの対処が求められることに留意されたい。

（3） 人間と技術の新しい共同・役割分担の実現

今までは人間と機械は、両者の活動領域を完全に分断することによって安全を確保してきたが、人と機械とが同一の空間で共存して協働する必要性が増えてきており、従来の機械安全の考え方では限界が生じつつある。例えば、工場では人の能力に応じて機械の速度を制御することで安全性と生産性を両立する、運転中に運転手に異変が起きた場合は自動で停止し関係者に連絡が行くようにする、等の具体的な適用事例がある。

このような状況を踏まえて、人の注意力や判断力によって安全を確保してきた Safety0.0、機械側に安全対策を施しつつ人と機械を分離した Safety1.0 の次の安全の形態として Safety 2.0[96]が提唱されている。これは、技術と人間と組織・環境とがお互いの情報を共有し協調・調和して安全を確保する概念である「協調安全」を、ICT を用いて実現することを提案しており、IEC 白書「Safety in the future」でも紹介されている。

---

[94] DFFT; Data Free Flow with Trust の略称。外務省 https://www.mofa.go.jp/mofaj/files/100167362.pdf が参考となる。

[95]「データを曖昧さのない共通の意味として交換するコンピュータシステムの能力」および「機械が演算 可能な論理、推論、知識発見および情報システム間でのデータ連携を可能にする要件」のこと
https://webstore.iec.ch/publication/65942

[96] IEC 白書「Safety in the future」https://webstore.iec.ch/publication/67876

人間と機械の役割をうまく分担し、機械が得意な部分（ルールに基づく検証など）は機械に任せ、人間はより創造的な仕事を実施することで、より効率的、効果的に安全を達成することができる。

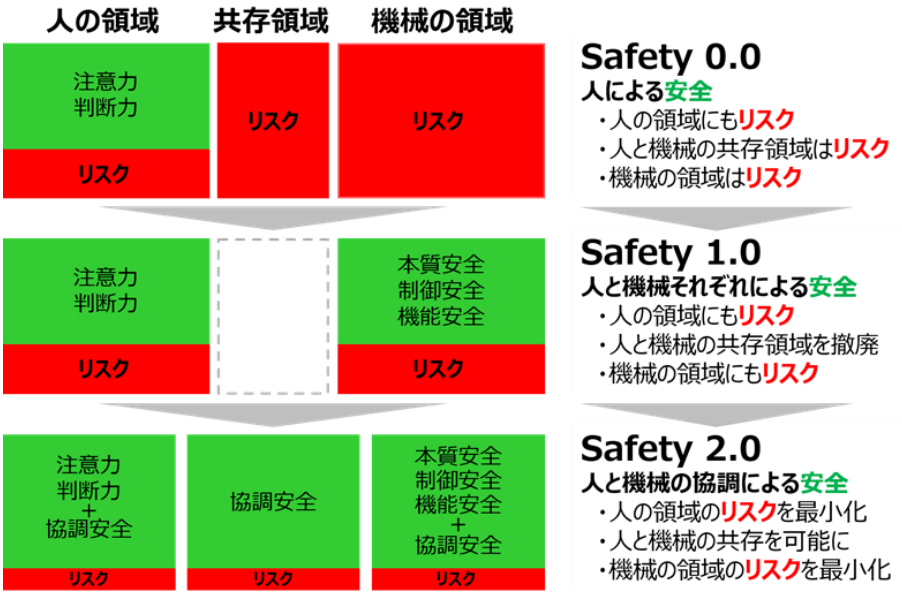


図 3-5　人間と技術の新しい共同・役割分担の実現

出所）一般社団法人　セーフティグローバル推進機構　協調安全

3.2.2. 不確実性の受容可能な水準でのマネジメント

（1）　サイバーセキュリティに対する要求

社会全体の活動全体におけるサイバースペースの機能・役割の拡大と共に、サイバースペースを介した事故発生のリスクシナリオも増加する。既に多くの分野において、セキュリティ上の脆弱性を突いたサイバー攻撃も発生している。また、Society5.0 においてサイバー空間とフィジカル空間との融合が進み、コンピュータが直接フィジカル空間を操作するようになると、Society5.0 時代の安全を考える上で、「サイバー・フィジカル・セキュリティ」、「サイバースペース上のデータの真正性確保」への対応機能が重要となる。

CPS におけるセキュリティのためには、サイバー・フィジカル・セキュリティ対策フレームワーク（CPSF）[97]において整理がなされているように、「企業等の適切なマネジメントを基盤とした各主体の信頼性の確保」、「フィジカル・サイバー間を正確に転写する機能の信頼性の

[97] 経済産業省　サイバー・フィジカル・セキュリティ対策フレームワーク
https://www.meti.go.jp/press/2019/04/20190418002/20190418002.html

確保」、「自由に流通し、加工・創造されるサービスを創造するためのデータの信頼性の確保」が必要となる。

また、安全とセキュリティの両立性の達成が課題となっていることから、安全とセキュリティの両立性を考慮したフレームワークの検討が進んでおり、以下の 3 点を考慮した規格開発(IEC63069)が進められている。規格化の目途がようやく立った状況であり、認証基準の策定は課題として依然として残されている。また、安全とセキュリティは前提条件の異なる技術的課題の難しさもあり、従来の認証要件とは異なる内容となると考えられる。

- ✓ 機能安全規格の IEC61508[98]とサイバーセキュリティ規格の IEC62443[99]の両立
- ✓ 両規格の用語、概念、開発プロセス等をすり合わせ矛盾解消
- ✓ 二つの開発プロセスは並行し、安全からセキュリティにインプットを行いシステム仕様についての競合解消

ガバナンスという観点では、「フィジカル空間における繋がり」、「フィジカル空間とサイバー空間の繋がり」、「サイバー空間における繋がり」の各レイヤーのルール形成やモニタリングを適時・適切に行うことで、サイバー空間、およびサイバー空間に紐づくフィジカル空間の安全を確保する必要がある。

（2） 新技術（AI 等）の信頼性に対する要求

新技術の活用には常に新しいリスクが伴う。特に、深層学習アルゴリズムは、ロジックの説明が困難であり、かつ導入時以降にその品質が変化し得るという特徴を備える。新しい技術を社会に導入する際には、人を代替し得る技術については有事の責任分担や対応等についてあらかじめ検討し、その技術の信頼性や社会的な受容性についても考慮する必要がある。今後さらにシステム同士がつながっていく社会において、新しい技術がもたらすリスクを理解・評価し適切な安全対応を行うことの重要性はより拡大する。

従来、我が国では、事業者が技術導入を計画した際に、規制当局による安全性評価に基づいて新技術導入の承認が行われていた。今後は、規制のゴールベース化や新技術の開発・導入の加速化に伴い、技術活用はその技術を活用する事業者による安全性論証によって安全が示される形になることが想定される。事業者は安全性評価を自ら行い、導入や運用を通じたデータをサイバー空間で管理し、規制当局はそれを適宜監督、もしくは常時モニタリングに移行することにより、技術導入を速めるとともに、信頼性を確保することが可能になると見込まれる。

---

[98] IEC61508：電気・電子・プログラマブル電子安全関連系の機能安全
[99] IEC62443：産業通信ネットワーク/ネットワークとシステムのセキュリティ

その安全性評価の対象となる技術の信頼性、とりわけ AI の信頼性は目下国際的に議論されているところであり[100]、我が国での信頼性の考え方や活用可能な信頼性の水準については、国際的な議論を踏襲することが前提となる。また、「要求された機能」が適切なものとなっていない場合、信頼性が高くとも安全ではないシステムとなる恐れがあり、AI 活用が進むことでリスクの予見性が低下し想定外の状況が増加する中で、安全性を保証するためにどのように要求仕様を構築すべきか、という点も重要な観点となる。

国際的な議論において、AI の活用を促進するためには、下記の課題を踏まえた、AI の信頼性の検証・認定手法を確立し、その判断や AI を用いたシステムの安全性を社会が受容される状態とする必要性が認識されている。また、AI のガバナンスの設計にあたっては、規制の程度をリスクの大きさに対応させるべきという考え方（リスクベース・アプローチ）は、国際的に広く共有されている。現状では国やステークホルダー毎に異なるリスク評価や分類がなされており、今後はリスクのランク分けに関して標準的指標を示し、用途に応じた規制のあり方を設計していく必要がある。

AI の活用を促進する上で検討すべき課題として、以下のような点が挙げられる。

- ✓ AI システムのゴール設定の困難性、予測困難性、説明困難性：ブラックボックスとなり得る AI システムの品質・リスクを適切に評価・検証できるようにしておく必要があり、データ共有の仕組みが必要。
- ✓ AI システムの脆弱性：学習データに依存する挙動の安全性を確保するため、データに対するマネジメントやセキュリティ対策や実験と現実のギャップを埋めることが求められる。
- ✓ アルゴリズムの継続的な変化：開発・運用を一体化した AI システムの継続的なリスクアセスメントによる品質劣化の監視・対策や事故等の発生時の責任の所在を事前に明確化しておくことが求められる。
- ✓ AI の安全側への活用に対する社会的受容性

（3） システムオブシステムズの総合信頼性に対する要求

システムオブシステムズ（SoS）の特徴を持つ「繋がるシステム」の展開を支援するためには、個別システムが外部との相互接続性、拡張性を持つことは有効であるが、システムの接続によって新たな機能・振る舞いが「創発」し、それが安全に対する脅威ともなりうることにも留意すべきである。当初は想定していなかった様々なシステムがつながることによりユーティリティは飛躍的に高まる一方で、リスク分析において利用シナリオをすべて網羅することは難

[100] 参考　https://home.jeita.or.jp/press_file/20191023145047_3Ezs15ATUG.pdf

しく、利用形態によって生じる新たなリスクが重大な結果につながる可能性を否定できない。社会を安全に保つためには、システムオブシステムズの信頼性の確保が必要になる。

「安全」というセンシティブな目標に対しては、不確実性に対応するために「設計・開発段階」のみならず、「運用段階」での安全確保の工夫がより重要になる。例えば、「段階的運用モードによる未然防止スキームの確立」、「事前の取り決めにない事象に対してのアクションを促す仕組み」、「合意形成の仕組み」等の論点がある。例えば、DEOS[101]（Dependability Engineering for Open Systems）では、変化しつづける目的や環境の中でシステムを適切に対応させ、継続的にユーザが求めるサービスを提供することができるシステムの構築法が検討されており、システムのディペンダビリティについてステークホルダーが共有・相互理解し、対外的な説明責任を果たすための手法/ツールとして D-CASE[102]が開発されている。D-Case を用いた開発運用プロセスでは、システムのディペンダビリティに対する主張・前提・証拠が明示的に記録されているため、主張が成立することを客観的に論証できるとされている。

また、相互接続による甚大な被害の発生を防ぐため、繋がるシステムの中に危害が広がらないようなファイヤーウォール等を適切な個所に設け、また必要に応じてソフトウェアの挙動に介入する等のサイバー空間におけるガーディアン機能（デジタルガバナンス）が必要になる可能性がある。当然ながら、そのような対応を実装するためには、データや AI を用いる際のリスク情報の創出が前提になるほか、社会受容性の観点からも十分な議論が行われるべきである。

3.3. Society5.0 における安全・ガバナンスの機能

本節では、Society5.0 における安全・ガバナンスに必要な機能を示す。（図 3-6）

[101] DEOS http://deos.or.jp/technology/process-j.html
[102] D-Case http://www.dcase.jp/

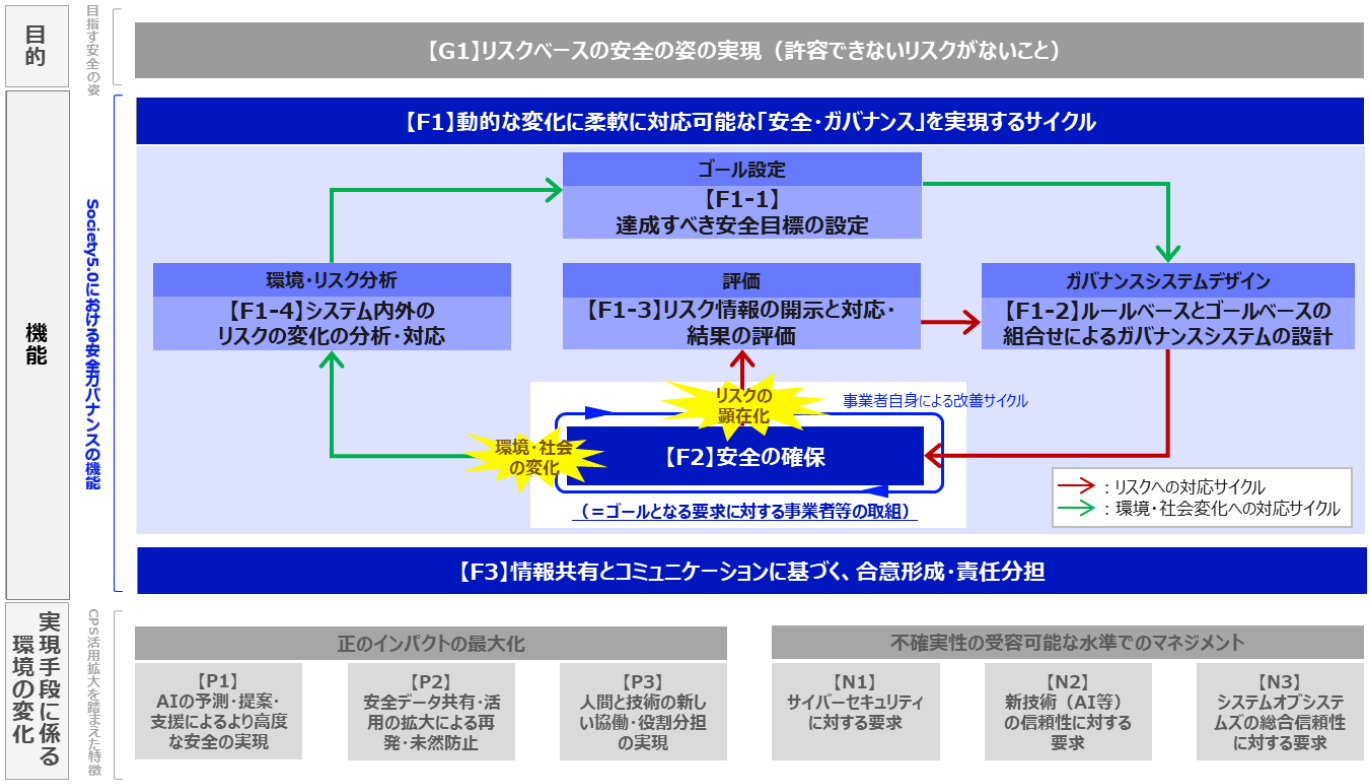


図 3-6　3.3 節のスコープ

3.3.1. 【F1】動的な変化に柔軟に対応可能な「安全・ガバナンス」を実現するサイクル

Society5.0 において、技術や技術を用いたシステムの変化が加速し、それに併せて社会システムの CPS、SoS 化が進むことでリスクの変化の速度が拡大していくことが想定される。こうした状況では、達成すべき安全目標（ゴール）は不変ではなく、常に変化し続けるものとなる。そのような社会変化を踏まえ、Society5.0 のガバナンスは常に変化するシステム、リスク及びゴールを考慮した「動的な変化に柔軟に対応可能なガバナンス」であることが必要である。ここで言う「動的」とは、製造/設計過程のみならず、システムが稼働中にもシステムの振舞いやリスクが変わることを考慮し、運用中も動的に安全を確保・維持する仕組みをも指す。従来の達成すべきゴールや基準、手段があらかじめ規定されているガバナンスから、ガバナンスの各プロセス（図 3-6 の「ゴール設定」、「システムデザイン」、「運用」、「評価」、「環境・リスク分析」）をシステムのライフサイクル全体に渡って継続的及び迅速に回していくことで変化に対応していくガバナンスに移行させるということである。

さらに、Society5.0 では単一物理システムが分野横断的にサイバーや他システムとつながるため、単一物理システム（各事業者）のみによる安全確保や従来の分野・業界・省庁縦割りの

ガバナンスでは整合性を図れなくなる。例えば、ドローン（主には国土交通省の管轄）によって、プラントの安全（主には経済産業省の管轄）を確保する場合、それぞれが正しい法規制の中で活用されていたとしても、ドローンとプラントが CPS で接続されることで新たなリスクが生じ、事故が引き起こされる可能性もある。その時、法律の抜けが存在し責任の所在が不明確になる恐れがある。そこで、分野横断的な全体システムとして安全確保を実現するガバナンス設計が必要となる。

以上の課題意識から、産業横断の安全・ガバナンスのアーキテクチャのビジョンである「動的な変化に柔軟に対応可能な「安全・ガバナンス」を実現するサイクル」（図 3-7）を提案する。

図 3-7 では、３つの改善サイクルを表現している。３つの改善サイクルの内容は、以下の①②③の通り。

① リスクへの対応サイクル（【F1-2】ルールベースとゴールベースの組合せによるガバナンスシステムの設計→【F2】安全の確保→【F1-3】リスク情報の開示と対応・結果の評価）　※図中赤線を参照：ガバナンスの運用の過程・結果を参照することで目指す安全の姿が達成されているかを確認し、ガバナンスのシステムデザインを改めて実施し、事業者等の実施する安全確保に対するゴールベース要求の内容、仕組みを再設計するサイクルである。

② 環境・社会変化への対応サイクル（【F1-1】達成すべき安全目標の設定→【F1-2】ルールベースとゴールベースの組合せによるガバナンスシステムの設計→【F2】安全の確保→【F1-4】システム内外のリスクの変化の分析・対応）　※図中緑線を参照：ガバナンスのそもそものゴールである目指す安全の姿そのものを見直す必要が発生した場合の改善サイクルである。

③ 事業者自身による改善サイクル　※図中青線を参照：安全を実現するためのこうした取組が事業者内で適切に実施されていることを、予兆・先行指標や運転データ等のリアルタイムデータを活用して確認・分析し、コーポレートガバナンスを通して事業者自らが継続的に確認・改善するサイクルである。

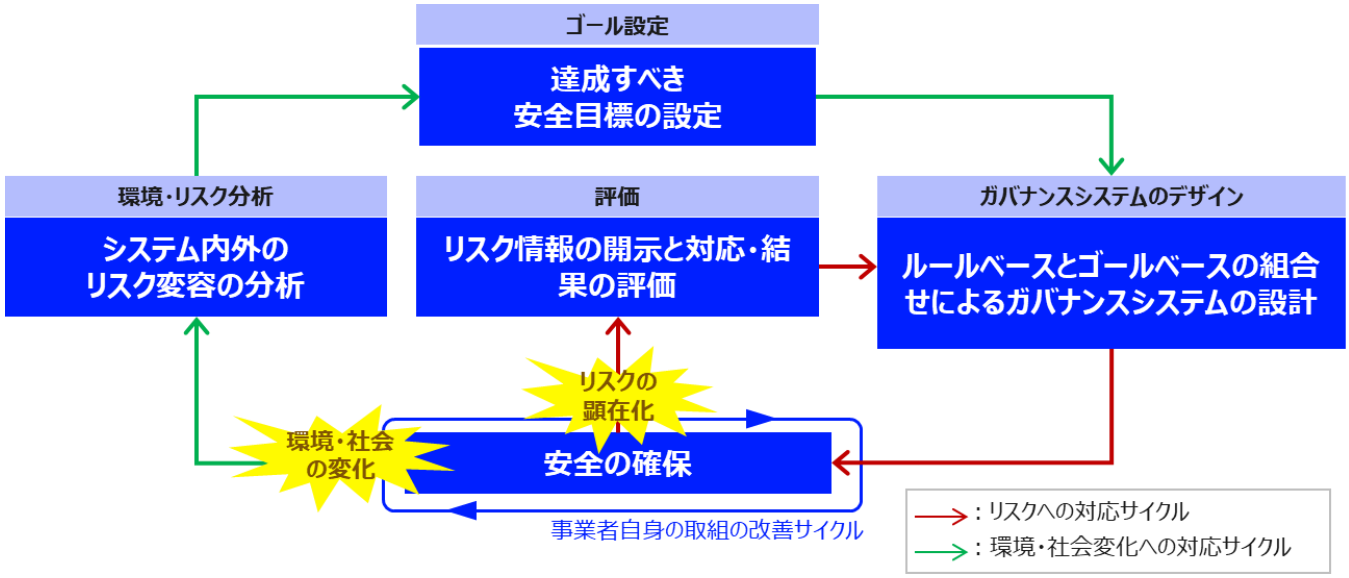


図 3-7　動的な変化に柔軟に対応可能な「安全・ガバナンス」を実現するサイクル

本ディスカッションペーパーの前提となる経済産業省「GOVERNANCE INNOVATION VER.2: アジャイル・ガバナンスのデザインと実装に向けて」では、Society5.0 では変化する環境とゴールを踏まえ、最適な解決策を見直し続けるガバナンスが必要であり、様々な社会システムにおいて、「ゴール設定」、「システムデザイン」、「運用」、「環境・リスク分析」、「評価」、「改善」のサイクルを回すアジャイル・ガバナンスの考え方が提唱されている。この考え方は、安全という目的においても適用することができる。本ディスカッションペーパーで提案する安全・ガバナンスの機能と「GOVERNANCE INNOVATION VER.2: アジャイル・ガバナンスのデザインと実装に向けて」におけるガバナンスの要素の対応を図 3-8 に整理した。

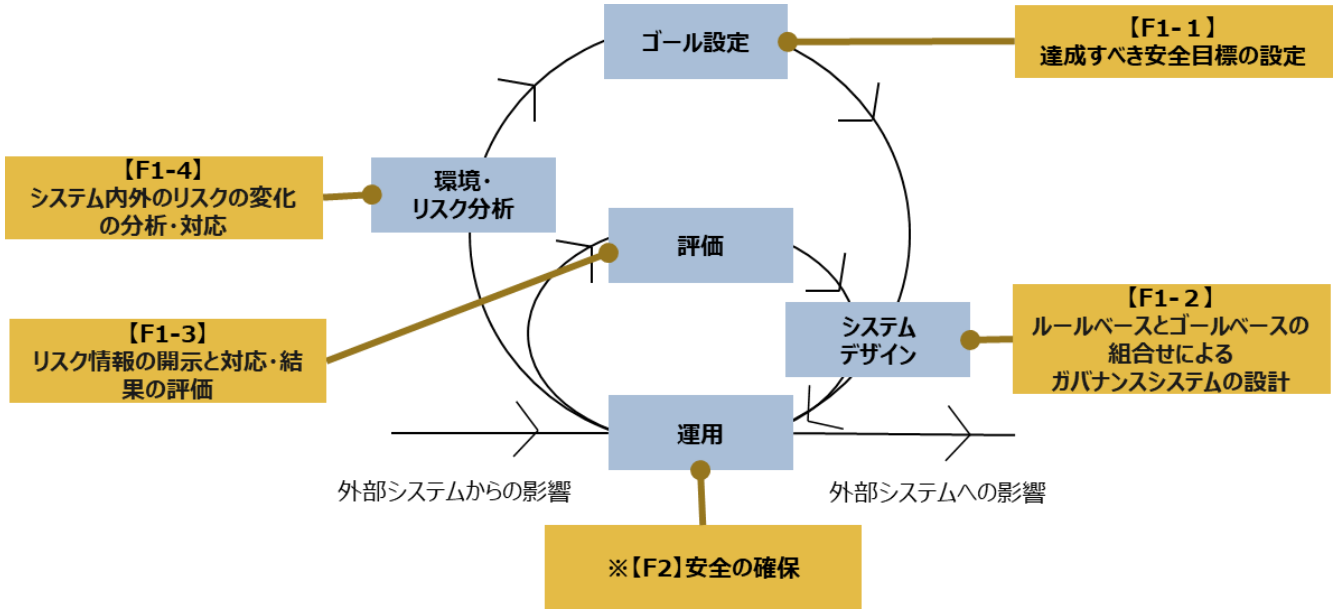


図 3-8　ガバナンスイノベーション報告書との対応

（1）　【F1-1】達成すべき安全目標の設定

安全は「許容できないリスクが無いこと」（ISO/IEC ガイド 51）と定義されており、「許容」する主体は、システムに関与するステークホルダー（の総体）と考えられる。ステークホルダーはシステムに対する要求を暗黙に持ち、危険事象発生時には「（リスクが）許容を超えた」として措置が取られる仕組みとなっている。なお、有限なコストの範囲内においては、要求を達成するためにある程度のリスクを受け入れなければならない。少なくとも我が国においては、リスクとベネフィットのバランス関係が市民等に意識されることは少なく、これによって「ゼロリスク志向」と呼ばれる「危険性が一切存在せず、（危害が及ぶ）リスクがないことが安全である」といった考え方に繋がっているといえる。その結果、システムの持つ本来の目的達成のための高度化にブレーキがかかってしまう可能性が懸念されている。

Society5.0 において、十分に先進技術の恩恵を受けるためには、一定の残留リスクを受け入れるという前提で「安全」に対する要求を顕在化させることが必要である。安全要求の明確化により、システムの持つ本来の目標の達成阻害を最低限に抑えつつ、安全目標を達成することを議論することが可能となる。そのためには、リスク情報をステークホルダーそれぞれの立場から主張し共有する枠組みや、事業者がシステムの目的と用途等の利用コンセプト・危険事象を定義し、Unknown な危険事象を最小化する取組が必要となる。例えば、ISO/DIS 21448 の SOTIF (Safety of the Intended Functionality) の考え方や、欧米企業らが発表したガイドライン SaFAD (Safety First for Automated Driving) 等でも同様の考え方が示されている。また、そ

もそも市民が適切に「リスクを認識すること」をどのように受け入れられるかといった社会的受容性や安全教育に関する議論も併せて実施していく必要がある。

前述の通り、Society5.0 においては、技術や社会システムの変化速度が上がり、それに伴いリスクの変化も加速していくこととなる。そのため、リスクベースの安全を志向したうえで達成すべき安全目標（ゴール）を設定することにおいて、その安全目標も常にリスク状況に応じて変化していくこととなる。そのため、システムの内部外部環境の変化やリスク状況の変化に応じて、ガバナンスとして目指す安全目標（ゴール）自体も常に見直すことが重要となる。

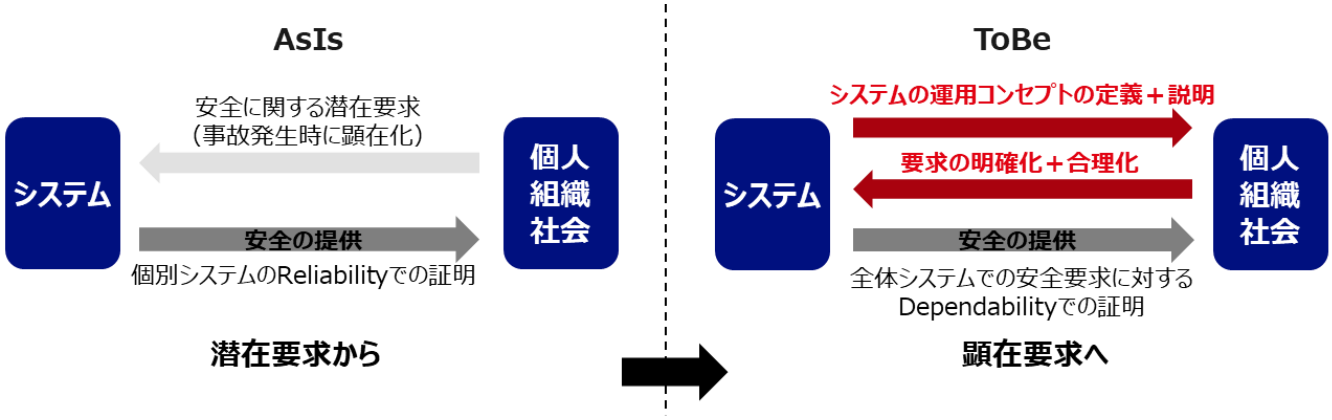


図 3-9　目指すべき安全の姿の設定（イメージ）

（2）　【F1-2】ルールベースとゴールベースの組合せによるガバナンスシステムの設計

我が国では、従来有資格者による検査の実施を要求する等、一律に行為義務及び手段を定めたり、規制当局が企業の行為を定期的に監督したりすることで、定義した安全目標・安全要求を達成するガバナンスの仕組みとなっていた。そのような行為義務及び手段を規定するルールベースの安全確保に対する要求に基づくガバナンスシステムの設計を図 3-10 にイメージとして示した。こうした仕組みは、経年で大きく変化しないシステムには有効であったものの、Society5.0 において変化し続けるシステムや安全目標（ゴール）を達成することは困難となる。

そこで、安全確保を行うために行為義務及び手段を規定する「ルールベース」の要求のみに基づくガバナンスシステムから、達成すべきゴールを規定する「ゴールベース」の要求を適切に組み合わせ、変化し続ける社会システムへ柔軟に対応しながら合理的に安全を達成するガバナンスシステムを設計することを目指す。具体的には、安全目標に対するプロセス・達成手段に対して事業者に裁量を持たせ、要求やゴールの達成度により事業者の取組の是正の要否判断する仕組みである。そのようなゴールベースの安全確保に対する要求とガバナンスの仕組みをイ

メージ図として、図 3-11 に示した。ルールベースの要求のみに傾注すると本来の目的（製品やサービスを通じて提供する価値の向上）を見失い、「ルール逸脱を避けることが自己目的化」してしまう状況を招きかねない。他方で、ゴールベースの要求を組み合わせることで、「どのような価値をどのように提供したいのか、事業者がどのようにリスク管理しているか等の組織目標の達成する仕組み（システム）をチェックする」ガバナンスとなる。その結果、事業者のリスクマネジメント戦略や方針が考慮されることで、事業者のリスクマネジメント能力やモチベーションの向上、適切な競争環境の構築に資することが期待される。

安全確保に対するゴールベースの要求をルールとして実装するためには、事業者は一方的に規制される側に立つのではなく安全に関する情報を持った専門家としてルール形成に参画することが望まれる。また、事業者によるリスクマネジメント能力・体制の高度化、安全の達成状況の可視化、説明の透明化や行政・市民等のステークホルダーとの継続的なコミュニケーションが必要となる。事業者によるガバナンスへの積極的な参画を高めるためには、以下のような特徴を持つ制度設計が求められる。

- ✓ 高度なリスクマネジメント能力・体制に対する認証・監査の仕組みの設計（国による監督や第三者認証機関の活用等）
- ✓ 事前には想定できないリスクの顕在化として取り扱う対象については、被害補償と民事免責制度を組み合わせることで、安全措置の過剰によるイノベーションの阻害を防止しつつ、被害補償を実効化すること
- ✓ リスクの顕在化（事故の発生時）には、原因究明のための情報提供（社会知としての蓄積・共有）や製品・サービスの改善措置を実施する事業者には制裁の免除（例えば、DPA（訴追延期合意[103]））を考慮する制度とすることで企業改善の促進に対するインセンティブ設計を図ること

さらに、安全規制にゴールベースのアプローチを適用する際には、事業者の能力に応じて裁量の幅を調整することや、ガバナンス対象の事業・システムが保有するリスクの大きさに対して適切なゴールベース／ルールベースそれぞれのアプローチのバランスを考慮することも重要と考えられる。

[103] deferred prosecution agreement の略。米国や英国で施行されている司法取引の一種。内部告発制度等の捜査協力、原因究明に必要なあらゆる情報の提供を事故を起こした企業が体制作りすることや再発防止に向けた具体的な改善を義務付けて、その代わりに刑事訴追の一定期間延期（ルールさえ守れば訴追見送り）をするという司法取引制度のこと。

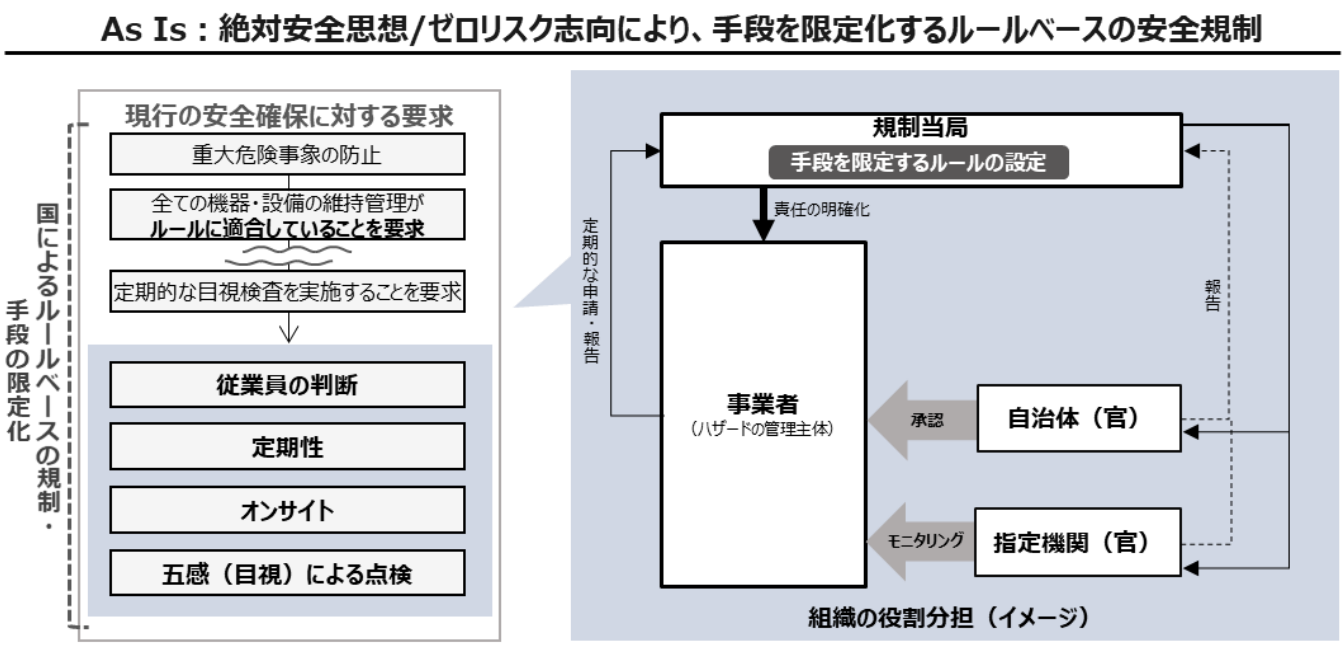


図 3-10 安全確保に対するルールベースの要求とそのガバナンスの仕組み（イメージ）

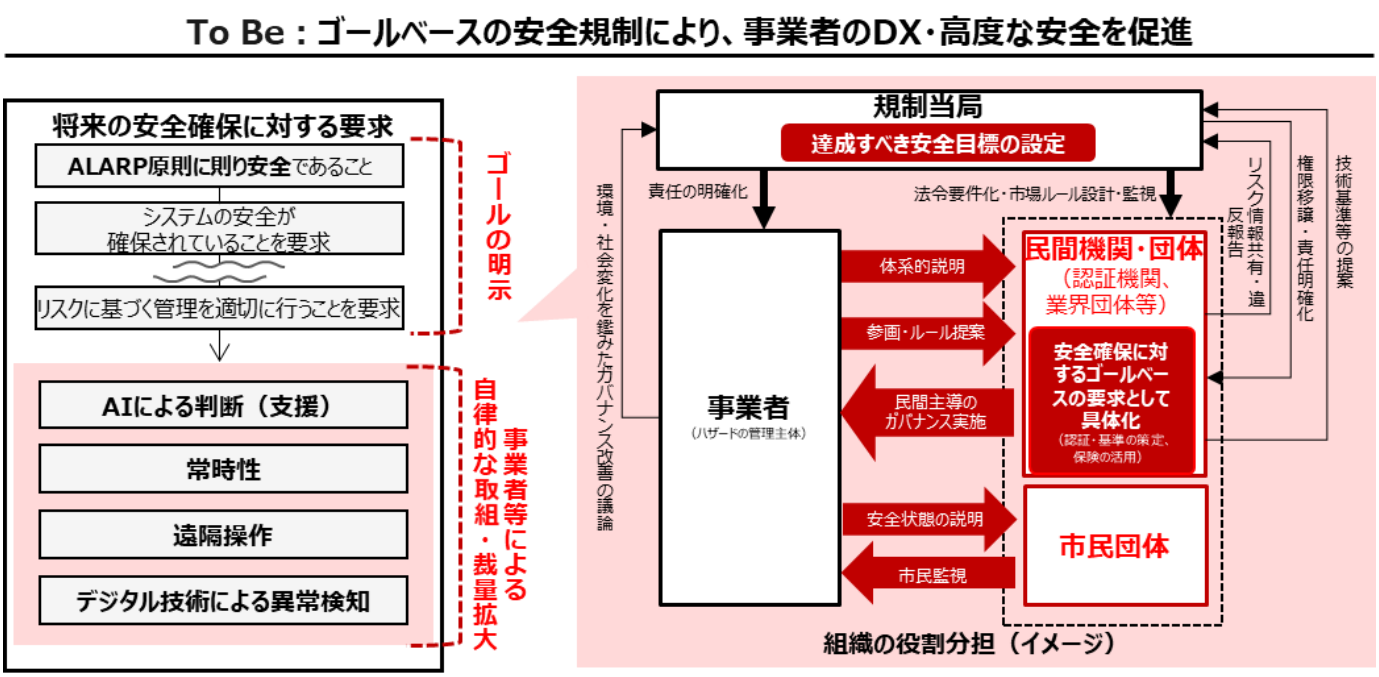


図 3-11 安全確保に対するゴールベースの要求とそのガバナンスの仕組み（イメージ）

コラム③：参考とすべき金融分野におけるガバナンス改革の事例（金融検査マニュアルの廃止）

社会の変化に適応し、ガバナンスの在り方を一律に行為義務及び手段を定めるルールベースからパフォーマンスの観点で要求することをより重視するゴールベースに変革させた先進的な事例として、金融検査マニ

ュアルの廃止の事例を紹介する。

金融庁は 2008 年の時点で「ベター・レギュレーション」のコンセプトを表明しており[104]、そのなかでは、英国の金融監督当局の政策を参考として[105]、詳細な「ルールベースの監督」から、主要な原則を示した上で、原則に基づく金融機関の自主的な取組みを促す「プリンシプルベースの監督」を最適な形で組み合わせることを第一の柱とすることを示している。このプリンシプルベースのコンセプトは、金融検査マニュアルの廃止に向けた検討のなかでも継続して反映されている[106]。Decker(2018)[107]の研究では、英国の金融分野の「プリンシプルベース」について、ゴールベースのコンセプトと本質的な矛盾がないことが述べられている。そのため、本コラムでも先行研究の意図を反映して、金融検査マニュアル廃止の事例をゴールベースのアプローチを取り入れた先行事例として紹介する。

金融検査マニュアル（以降、検査マニュアルという）とは、金融庁の検査官が、預金等受入金融機関を検査する際に用いられていた手引書である[108]。検査マニュアルに基づいて、金融機関はその規模や特性に応じた内部規程の策定と検査への対応が必要とされており、検査マニュアルは金融機関のビジネスの在り方に影響を与える極めて重要なルールとして位置づけられていた。しかしながら、検査マニュアルは、1990 年代のバブル経済崩壊後の失敗経験から、金融機関の貸付先である企業の不良債権を的確に把握することを重視して策定された経緯から、過度な担保への依存によって、金融機関による貸出先の事業の理解や目利きが阻害されるとともに、金融機関として把握している将来の貸し倒れリスクを引当に反映させることが困難であることなどが指摘されていた[109]。このような課題に対処し、金融機関が顧客の多様なニーズに応えるための創意工夫に取り組みやすくさせるために、金融庁は 2019 年 12 月 18 日に「検査マニュアル廃止後の融資に関する検査・監督の考え方と進め方」（以降、考え方文書という）を公開するとともに、同日をもって検査マニュアルを廃止している[109, 110]。

考え方文書は、検査マニュアル廃止後の金融検査について、規制当局の考え方を示すためのディスカッションペーパーとして位置づけられている。そして、今後、利害関係者との協力によって継続的な改善が図れることが示されている[109]。また、金融庁が金融機関の健全性を評価する際、どのように金融仲介機能の発揮に取り組んでいるのか（又は取り組もうとしているのか）を理解した上で、金融仲介に伴い発生するリスクを特定・評価し、健全性上の優先課題を把握するための対話を重視していくことや、金融機関の個性・特性に着目した検査・監督を行っていくことなどが示されている[109]。さらに、検査マニュアルの廃止は、あくまで金融機関が現状の実務を出発点に、より良い実務に向けた創意工夫を進めやすくするためのものであること

[104] 金融庁 HP 金融規制の質的向上―ベター・レギュレーション― https://www.fsa.go.jp/policy/br-pillar4/index.html
[105] 金融庁 HP 「金融規制の質的向上：ルール準拠とプリンシプル準拠」 https://www.fsa.go.jp/common/conference/danwa/20070912.html
[106] 金融検査・監督の考え方と進め方（検査・監督基本方針）. 金融庁. 2018
[107] Christopher Decker, "Goals-based and rules-based approaches to regulation." BEIS Research Paper 8. 2018.
[108] 検査マニュアル廃止後の融資に関する検査・監督の考え方と進め方. 金融庁. 2019
[109] 検査マニュアル廃止後の融資に関する検査・監督の考え方と進め方. 金融庁. 2019
[110] 金融庁 HP 金融検査マニュアル関係（※これらの文書は令和元年 12 月 18 日に廃止しました。） https://www.fsa.go.jp/common/law/manualLink.html

を意図しており、廃止された検査マニュアルに基づいて定着している現状の実務を否定しないことを明確にしている。従って、考え方文書の文中にも一定の詳述や事例が示されているものの、それらはあくまでより良い実務に向けた対話の材料とするためのものであり、個々の論点を形式的に適用したり、チェックリストとして用いたりするものではないことが強調されている[109]。

我々は、検査マニュアルの廃止は、規範的なルールを廃止し、ガバナンス（特に規制）の在り方にゴールベースの考え方を実装した日本国内における成功事例だと捉えている。そのため、金融分野とはいえ、多様な産業でゴールベースの規制を実現する際に本事例から参考とすべき点は多いにあると考え、具体的な参考とすべき点として、以下を抽出した。

① 民間事業者が創意工夫に取り組み易くすることを規制の移行の目的として明確にしていること。
② ゴールベースの移行後の基本的なコンセプトを示すディスカッションペーパー（本事例では考え方文書）を公開し、その継続的改善の必要性が示されていること。
③ 現状の実務を否定せず、文書内で示される個々の内容を形式的に適用すべきではないことを明確にしていること。

但し、検査マニュアルの廃止による政策の成否の評価は、本稿の範囲ではないことに留意されたい。金融分野のゴールベース移行による効果については、今後、継続的なモニタリングによってその実装が評価されるべきと考えらえられる。

（3） 【F1-3】リスク情報の開示と対応・結果の共有

設計したガバナンスシステムにおいて当初設定された安全目標（ゴール）が達成されているかを判断するためには、リスク状況のモニタリング及びそのモニタリング結果に基づくガバナンスシステムの再設計・ゴールの見直しが必要である。リスク予兆やインシデントが発生した際に、リスク情報（リスク顕在化に伴いどのような影響があるか等）、運用体制、運用結果、緊急時対応等について、ステークホルダーに適切な開示が行われる仕組みも必須となる。

CPS に対するモニタリングは、人による現場での定期点検に基づく定期報告・監査等ではその効果及び効率の観点から限界があることは自明である。むしろ、種々のセンサーを用いたリアルタイムデータの取得、AI を用いたシステムの状況把握や予測を活用した保安を前提と捉えるべきである。ガバナンスシステムに関与する各事業者が健全に機能を果たしていることを国が裏支え・監視する機能（リスク情報の真正性を評価する第三者機関認定等）を担い、そのうえで事業者がリアルタイムデータ等を活用したリスクアセスメントを行うことで、運用段階での安全を確認し必要に応じて迅速な改善サイクルを回す仕組みを構築することが望ましい。このような仕組みを実現させるため、国や企業、第三者機関などの役割分担・責任分担を検討すべきである。

また、大きな事故を発生させず改善ループを素早く回すためには、安全を阻害するリスクの予兆を検出できるリーディングインディケータ（先行指標）の導入[111]等が有効な手段となる。事前に予兆を認識するためのプロセスを、可能な限りユーザビリティを落とさずに実効的にスピードを持って回すためには、法規制の条文を自然言語からプログラミング言語等のコンピュータが理解できる形式に移行させ、コードによる規律や認証の AI 活用が有効となる可能性があるが、そのような変革に対する社会的受容性と合わせて議論されるべきであろう。

（4） 【F1-4】システム内外のリスクの変化の分析・対応

Society5.0 において CPS 活用が進み安全に係るシステムが CPS に直接間接的に接続されている状況では、外部環境や外部システムの様々な状況変化（脅威の発生、脆弱性の変更、仕様の変更等）にも対応する必要がある。従来はリスクの管理主体であった事業者内部のシステム管理担当者が現場での部分的・応急的な対応（パッチを当てる等）で済ませていた部分についても、特に安全に係る重要なシステムについては、今後より迅速かつ的確な対応が必要であり、ガバナンスの在り方そのものの改善（例えば、達成すべきゴールの設定の見直し）が継続的に行われることが求められる。その際は、AI の信頼性やサイバーセキュリティ等を議論する場合と同様に、システム全体として求められる安全機能の信頼性の観点から、どのようなガバナンス設計を行うべきかをシステムの置かれた環境やリスクの変化を踏まえて検討する必要がある。

3.3.2. 【F2】安全の確保

本節では、安全・ガバナンスのうち、Society5.0 において事業者（ハザードの直接の管理主体）等の組織内部で実施すべき「安全確保」の在り方を検討する。事業者等による安全確保は、ガバナンスの運用フェーズにあたる。

3.3.1（2）で記載のとおり、システムの安全確保に対してゴールベースの要求がなされる場合、プロセス・達成手段に関して事業者が裁量を持つこととなるため、Society5.0 において事業者による安全設計はどのようになされるべきかという点がより一層重要となる。さらに、不確実性を持った CPS が社会的に受容されるためには事業者による安全確保が核となるため、「安全」を対象としたガバナンスの在り方の検討において本項目は特に重要といえる。

そこで、本項目では CPS が中心となる社会ではどのように事業者による安全確保が行われるべきかについて整理した。

---

[111] CCPS（The Center for Chemical Process Safety; アメリカ化学工業技術者協会（AIChE）により設立された世界中の化学企業 130 社を超える会員で組成される機関）の発行する「Process Safety Leading and Lagging Metrics」でもリーディングインディケーターの重要性は示されている。

まずは前提として、単独のフィジカルシステムにおけるリスク対応を行うための機能の主な特徴を、①危険事象シナリオの特定・リスク評価・対策、②安全防護層の設計及び対策、③意思決定に対するアカウンタビリティ、④緊急時対応による被害の局限化　の 4 点で整理した。（図 3-12 参照）

① 危険事象シナリオの特定・リスク評価・対策：危険事象シナリオを想定し、その事象進展に対する予防対策及び事故要因の特定を講じることが、一般的な事故予防のための重要なアプローチであり、リスクアセスメントに用いられる分析手法の多く（FMEA, What-If, Hazop, ET/FT 等）は事故シナリオの抽出手法である。

② 安全防護層の設計及び対策：一方、事故シナリオに対する防護対策の網羅性、及び必ずしも事故シナリオの全てを事前に予見できない場合を想定して概念的な事故波及を防止する考え方として、「防護層」という考え方も併用される。

③ 意思決定に対するアカウンタビリティ：安全・ガバナンスを最も強く特徴づけている点は、人間の生命・健康に関する危害が関係し被害者が発生するという点であり、安全の判断・意思決定に対しては被害発生時の救済への対応も併せて非常に大きな責任が伴う。

④ 緊急時対応による被害の局限化：必ずしも事故発生を予防できない場合も想定する必要があり、一定の被害発生を前提としながら、事故が発生した場合に、その被害影響を局限化するための対応策を事前に整理しておくことが危機管理として必須となる。

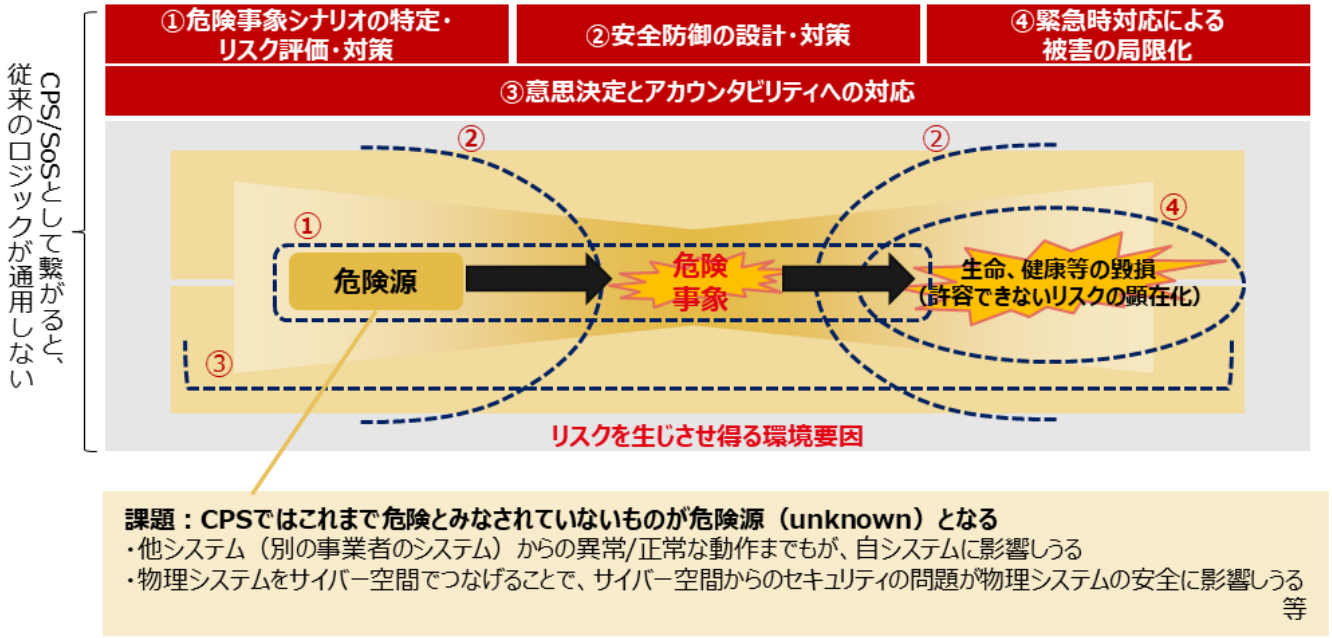


図 3-12　従来の一つのフィジカルシステムにおけるリスク対応を行うための機能の特徴

CPS や SoS として複数のシステムが繋がると、単独のフィジカルシステムにおけるリスク対応においてすら、図 3-12 に示した従来の機能やロジックだけでは通用しなくなる。システム同士の接続が行われることで他システム（別の事業者のシステム）からの異常/正常な動作が、自システムに影響し得るためである。さらに、危険事象シナリオの全体像は従来フィジカルシステムが中心だったものから、フィジカルシステムをサイバー空間で繋げることで、システムにおけるサイバー空間からのリスク影響の波及シナリオの割合が拡大する。そのため、SPS や SoS における波及シナリオ及びその防護ロジックは、フィジカルシステムとは大きく異なることを前提にリスク想定・対応を実施する必要がある。

上記を踏まえ、図 3-12 で示した従来の一つのフィジカルシステムにおけるリスク対応を行うための機能について、CPS におけるリスク対応の機能としてどのような検討が必要か図 3-13 の通り、整理を行った。この整理はあくまで 2020 年度に調査できた一部分を基にした仮説であり、検討が必要となるその他の論点も継続して調査・分析中である。

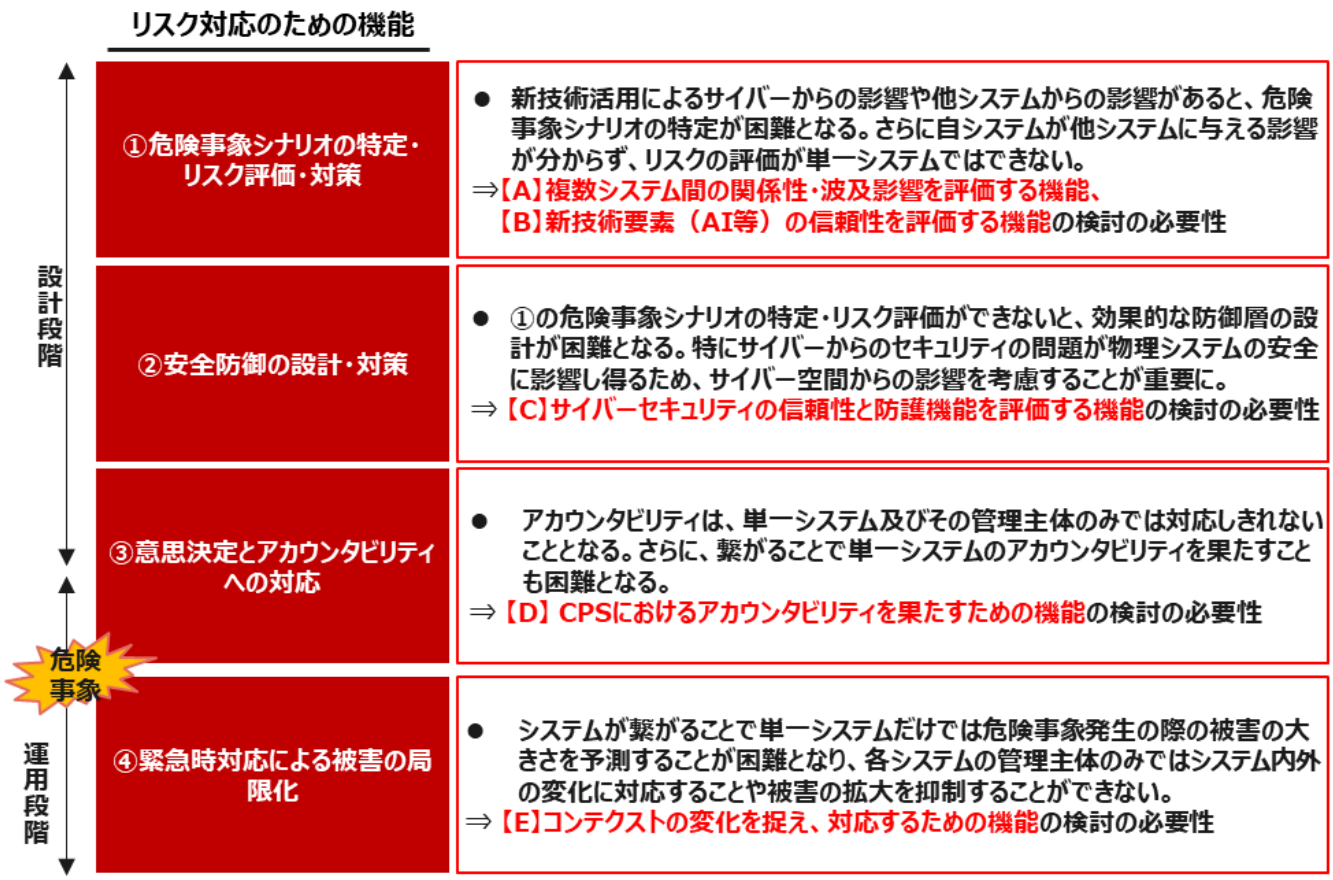


図 3-13 CPS におけるリスク対応のための機能について必要な検討

CPS の安全を実現する機能として、システムの構成要素となる新技術及び CPS 全体の安全度合いを評価する機能（【A】複数システム間の関係性・波及影響を評価する機能、【B】新技術要素（AI 等）の信頼性を評価する機能、【C】サイバーセキュリティの信頼性と防護機能を評価する機能）と【D】 CPS におけるアカウンタビリティを果たすための機能、【E】コンテクストの変化を捉え、対応するための機能　に着目した。

以下に、各機能に関連した取組や研究に関する調査成果を示す。

A) 複数システム間の関係性・波及影響を評価する機能：SoS では個別のシステムで意図した機能だけでなく、接続されたシステム間で機能が創発されるため、それが想定外の事故を引き起こす可能性もある。そういった Unkown-Unkowns[112]をも考慮しながら波及影響の様態の全体像を整理し、さらに CPS 全体の事象進展のモデルとして体系化することが必要で

[112] 問題の答えが分かっていないという以前に、問題の存在自体が認識されていない状況のこと

ある。例として、ハザードにつながる制御（非安全なコントロールアクション）を開発時にシステミックなやり方で特定し、その分析結果をプロアクティブに安全設計に反映する発想・分析アプローチとして STAMP/STPA[113]が開発されており、これは SOTIF の中でも紹介されている。

また、Unknown-unknowns を減らす工夫として、「特定の条件を満たしたもの（主体/アセット/データ等を含む）のみがシステムに参加できる仕組み」も有効である。例えば、GAIA-X では、条件を満たしていなければ、「システムに参加できない」とすることで、安全安心のデータ連携・利活用を実現できる思想になっている。具体的には、参加者、アセット、リソースのアイデンティティと能力に対する信頼を、Gaia-X の Federated Trust Model を用いてアイデンティティを暗号で検証することで確立し、すべての人に透明性を提供する仕組みとなっている（これは CPS においてアカウンタビリティを確保するための仕組みともいえる）。

B) 新技術要素（AI 等）の信頼性を評価する機能：「AI の誤判断はゼロにはならない」ことを前提に全体のシステム設計を行う必要がある。従来は高い信頼性（Reliability）が求められるシステムからは徹底的に AI を排除する方針が取られ、やむを得ず AI を含める場合には必ず相互牽制がかかるような冗長系を持つ業務・組織設計がなされてきた。一方、AI は大量のデータを活用した高速な判断を得意とし、一部の分野では既に人間を超えている。新技術活用によるベネフィットを享受するために、リスクがゼロにならないことを前提として、信頼性の評価方法や安全関連システムにおける活用方法を検討したうえで、社会的受容性を確立することが論点となる。

これらの留意点について、拘束力を持つルール施行（義務的）/任意判断に委任、産業横断的/産業別、自己認証/第三者認証等の観点から、どのように制度設計するかを各国で議論されている状況である。欧州は産業横断的な義務規定と産業分野別の義務規定を組み合わせた AI 規則案を公表し、米国は AI が活用されるユースケース（シチュエーション、産業）別に義務規定を導入している。我が国は、産業横断的な AI ガイドラインを 2021 年夏頃に公表予定。

C) サイバーセキュリティの信頼性と防護機能を評価・設計する機能：Society5.0 においては個別に設計・運用されている様々なシステムが、相互接続され新しい機能を果たしていくこととなり、その状況ではサイバースペースを介した波及影響シナリオも拡大すると考えられる。システムの安全機能に求められる信頼性に応じて、安全防護層として他システムへの接続を限定する必要がある。

欧米においては、インダストリー4.0 やインダストリアルインターネットに代表される企業

[113] IPA”はじめての STAMP/STPA～システム思考に基づく新しい安全性解析手法～”
https://www.ipa.go.jp/files/000051829.pdf

間連携や、製品にセキュリティ対策が行われていることを前提とした標準化の動きが活発化しており、連携される既存のシステムを含めて、システム全体の企画・設計段階から、セキュリティの確保を盛り込むセキュリティ・バイ・デザイン（Security By Design）の考え方が推進されている。

D) CPS におけるアカウンタビリティを確保するための機能：Society5.0 における Unkown-Unkowns なリスク（及び安全に対する影響）については予見性が十分ないため、安全性評価/安全論証が十分に行えないという原理的な課題が存在する。このような場合には、「実績に対する納得感による許容」が残されたアプローチ、すなわち説明責任を果たすためのマネジメントサイクルを担保することで社会的受容性を確立することが一つの方法となる。例えば、DEOS[114]において、 システムのディペンダビリティをシステムに関与するステークホルダーが共有し互いに分かり合い、対外的な説明責任を果たすための手法/ツールとして D-CASE が開発されている。具体的には、D-Case を用いた開発運用プロセスでは、システムのディペンダビリティに対する主張・前提・証拠が明示的に記録されているため、主張が成立することを客観的に論証できるとされている。

E) コンテクストの変化を捉え、対応するための機能：SoS においては、機能や構造や使用環境、コンテクスト等が変化し続けるため、不確実性を完全に排除することができず、開発時に気づかなかった直接的・間接的な相互作用が絡まりあって危険事象を引き起こす可能性もある。そこで、①不確実性に係る要因を顕在化する前に可能な限り取り除くこと、②顕在化後には適切な対応し、影響を最小化することが重要となる。例えば、DEOS では変化しつづける目的や環境の中でシステムを適切に対応させ、継続的にユーザが求めるサービスを提供することができるシステムの構築法を開発することを目標とした検討が進められている。

また、目的・ステークホルダーの変化への対応、相互接続性、グローバル性、拡張性を踏まえると、従来のような設計段階で安全確保の仕組みをシステムに埋め込むといった設計段階のみでの安全性担保は、システムの柔軟性低下や SoS の早発的振る舞いによる影響への対応困難を招く恐れがある。そのため、運用段階でのリアルタイムデータを活用した安全証明が求められる。例えば、Fraunhofer IKS が自動運転のテーマで取り組んでいる Dynamic Risk Management[115]に関する取り組みはその代表例である。運用時の改善サイクルを円滑に回すためにも、ステークホルダー間のデータ共有を踏まえた運用時のデータ・安全論証パターンの蓄積が求められる。例えば、AI を用いた自動運転アルゴリズムを検証

[114] DEOS“DEOS プロセス” http://deos.or.jp/technology/process-j.html
[115] Dynamic Safety Management for Autonomous Systems," in Engineering Safe Autonomy: Proceedings of the 27th Safety-Critical Systems Symposium, Safety-Critical Systems Club, 2019. http://publica.fraunhofer.de/eprints/urn_nbn_de_0011-n-5462157.pdf

するためには想定されるシナリオを構築することが求められるが、それには運用時のデータが必須である。

このような安全を実現するための取組が事業者内で適切に実施されていることを、事業者自らが継続的に確認・改善するためには、監査役や社外取締役等の第三者的な役職・機能の活用を含む、ガバナンス機能の整備が必要となる。そのようなコーポレートガバナンスを構築するには、企業価値向上等のために社長、CEO 等の経営層が安全の取組に関与する体制や、市場との関係性における適切なインセンティブ設計によりリスクテイクを促しながら、安全目標の達成度もチェックしていくという両輪の仕組みを構築すること、社外取締役等による監督機能の強化を行うこと等の、市場メカニズムを考慮したガバナンスを利かせることがポイントとなる。

これらの事業者等による安全確保、すなわちガバナンスの運用の過程・結果を踏まえて、ガバナンスシステム全体として 2 つの改善を実施する必要がある。一つ目は、リスクの顕在化（リスク予兆やインシデントの発生等）をきっかけに、「【F1-3】リスク情報の開示と対応・結果の評価」を経て、設定した安全目標（ゴール）が達成されているかを評価し、達成できていない場合はガバナンスのシステムデザインを改めて実施し、ガバナンスシステムを再設計する改善サイクル（リスクへの対応サイクル）である。二つ目は、環境・社会の変化をきっかけに、「【F1-4】システム内外のリスクの変化の分析・対応」を行い、ガバナンスのそもそものゴールである目指す安全の姿そのものを見直す必要が発生した場合の改善サイクル（環境・社会変化への対応サイクル）である。システムに係るリスクに変化を与える可能性のある環境が変わった場合には、リスクが変化する可能性を考慮し、ゴールを再設定し直す必要性を検討しなければならない。

コラム④：欧州の AI 政策

2021 年 4 月 21 日、欧州委員会は、AI 法案を含む「AI パッケージ」[116]を発表し、欧州会議および欧州委員会による審議が開始された。法案は、技術革新と新技術により生じるリスク対応の両立をはかるために、AI（人口知能）に対して、包括的に規制するものである。

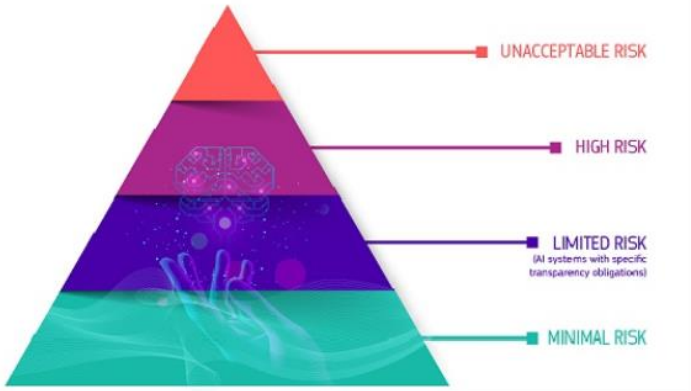
欧州委員会は、2018 年 4 月に「AI に関する戦略方針」を発表し、投資・開発方針と「人間中心の AI ルール」の検討を表明した。その後専門家グループの検討などを経て、2020 年「AI 白書」[117]を発表、今回

[116] https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-laying-down-harmonised-rules-artificial-intelligence

[117] https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf

の法案策定に至っている。これまで、日本[118]を含む各国ではAI倫理規定など概念的なものが中心であり、国家レベルでの法、規定案として注目される。規制が具体化した要因は、急速な技術革新とその普及により、既存の法律や規定が想定していない利用や、差別の助長などリスクが顕在化してきていることによるものである。

欧州委員会が提案するAI法案は、AI技術全般について社会における利用シーンを、「許容できないリスク」「高いリスク」「限定的リスク」「最小限のリスク」の４つに分類し、その危険度に応じ対応を定めるリスクベースの考え方にもとづいている。



図：リスクベースアプローチ

出典：https://digital-strategy.ec.europa.eu/en/node/9745/printable/pdf

定義	分類	対応方針
許容できないリスク	人々の安全、生活、権利に対し明らかに脅威となるAI（ソーシャルスコアリングなど）	原則禁止
高いリスク	人々の安全、生活、権利に悪影響を及ぼす可能性のあるAI（重要インフラ、教育、安全部品、安全製品、人事採用など）	適合性評価
限定的リスク	使用における透明性が求められるもの（チャットボットなど）	透明性確保
最小限のリスク	人々の安全、生活、権利にほとんど悪影響がないAI（ゲーム、スパムフィルターなど）	規定なし

「許容できないリスク」においては、公的な機関による法執行のための個人の識別など、人権侵害などの

[118] 人間中心のAI社会原則　https://www8.cao.go.jp/cstp/aigensoku.pdf

脅威のある用途について、原則的に利用を禁止する厳しいものである。第二に「高いリスク」として、人々の安全や基本的な権利に悪影響を及ぼす影響のある用途について、市場投入前に第三者機関による認証など適合性評価を前提とすることを求めている。本ディスカッションペーパーが対象とする、安全分野の多くがこのカテゴリになるであろう。規定案では、医療用ロボットなどへの AI 活用が例として示されているが、水道ガス電気などの重要インフラ、航空機、車両、船舶、工作機器、ガス機器など安全機能を有する分野は広く該当する。これらについて、分野毎に既存の規定との整合なども含め、随時法令による規定が進められることが法案別添にて示されている。

第三の「限定的リスク」では、消費者に AI を利用していること明示するといった透明性が義務づけられ、最後に「最小限のリスク」の範囲での利用は条件が課されていない。「許容できないリスク」の例として、顔認証技術を用い、オンライン監視による警察活動などが相当する。米国では複数の誤認逮捕の事例をきっかけに、2020 年 6 月に、IBM, Amazon[119], Microsoft など IT 各社が顔認証技術の警察への提供を自主的に停止するという処置を講じた例がある。「高いリスク」としては、電気や水道など公共インフラへの利用、民間事業者における顔認証や、企業、学校における人材評価や選考などへの利用が挙げられている。この規定は、欧州域外の製品、サービスであっても欧州向けに提供される場合適用対象となり、違反者には最大で連結売上高の 6%を課徴金として課すことから、成立すればきわめて影響度の大きな法案であると言える。

また、今回の「AI パッケージ」は、（AI 法案、機械命令の改正、AI コーディネーション・プラン）から構成されている。AI 法案についてはすでに述べた。機械指令では、「ハイリスク機械製品」というカテゴリを新設し同製品に第三者認証を必須とした。安全機能向けの AI システム、AI システムを組み入れた機械が、ハイリスク機械製品に位置づけられる。AI コーディネーション・プランは、投資とイノベーション促進のためのアクション計画である。規制とイノベーション推進の両面での提案となっている。法令として、モニタリングやエンフォースメントの仕組みについて規定がなされており、また今後 AI 賠償責任ルールについても提案がなされる予定である。

4 月発表後、欧州域内では人権団体を中心に規制の不十分さが指摘[120]されるなど、技術によるプライバシー侵害への懸念と、同時に新技術への期待が大きいことがあらためて明らかとなった。また一方で産業界からは認証にかかるコスト課題が指摘されるなど、さまざまな反応が報道されている。

この規制案は顔認証や自動運転などにとどまらず、広く AI 技術を活用する局面に適用されることから、影響度が大きく、また同様な動きが世界的に広がることを想定し、動向を注視する必要がある。

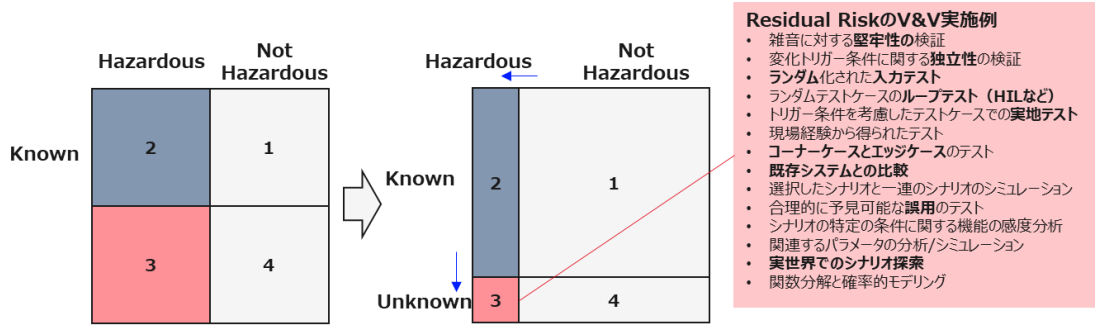
---

[119] https://www.aboutamazon.com/news/policy-news-views/we-are-implementing-a-one-year-moratorium-on-police-use-of-rekognition

[120] https://www.aboutamazon.com/news/policy-news-views/we-are-implementing-a-one-year-moratorium-on-police-use-of-rekognition

コラム⑤：ISO/DIS 21448: 2021 を参考にした CPS リスク軽量化の論点

SOTIF の考え方などを導入したとしても、CPS には Unknown-Unsafe シナリオは残存する。残存リスクを可視化するために、ドイツの Pegasus Project や ISO/DIS 21448 では様々な観点から検証の観点を提供している。「危険事象発生シナリオ」の具体化が必要であり。そこには分野横断でのデータ共有とメカニズムの理解が重要と考えられる。



出所）ISO/DIS21448:2021 を基に作成

コラム⑥：DEOS (Dependability Engineering for Open Systems)

DEOS は変化しつづける目的や環境の中でシステムを適切に対応させ、継続的にユーザが求めるサービスを提供することができるシステムの構築法を開発することを目標としている。現代のシステムは機能、構造、システム境界が時間的に変化しつづけるオープンシステム（開放系・変化系）であり、これに起因する不完全さと不確実さを完全に排除することができず、未来に障害となりうる要因（開放系障害要因）を本質的に抱えている。DEOS 協会は、それらの要因を顕在化する前にできる限り取り除き、また、顕在化した後に迅速かつ適切に対応し、影響を最小とするようにマネージし、利用者が期待する便益をできる限り安全にかつ継続的に提供する努力、社会への責任ある説明、およびそれらを継続的に行うことが必要だと考えている。そこで、対象（システム）を時間的に変化しつづけるシステムとしてとらえ、不完全さと不確実さに起因する開放系障害に焦点を当て、開放系障害を起こす要因の最小化と、開放系障害による影響の最小化によりサービスの継続性を向上させようと取り組んでいる。

DEOS プロセスのポイント

●反復的アプローチ

・目的や環境の変化に対してシステムを継続的に変更して行くための変化対応サイクル

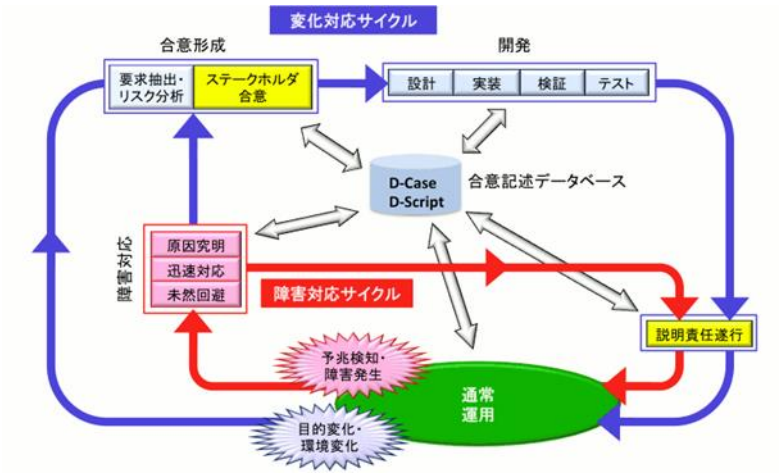
・障害に対して迅速に対応するための障害対応サイクル

・障害対応サイクルから変化対応サイクルへのパス

●D-Case を用いた合意記述データベースにより合意形成および開発・運用フェーズの統合と説明責任の全うを支援

●DEOS プロセスの考え方は 2013 年 7 月に The Open Group が標準として採用(※)

※Dependability through AssuredenessTM (O-DA) Framework
(https://www2.opengroup.org/ogsys/catalog/c13f)



出所）DEOS web ページ、http://deos.or.jp/summary/about_osd-j.html、2021 年 4 月 20 日閲覧

3.3.3. 【F3】情報共有とコミュニケーションに基づく、相互理解・責任分担

AI 等に代表される不確実性を完全に排除できない新たな技術の導入のためには、社会全体でその不確実性を認識し、許容し、責任を分担するための仕組みが必要となる。新技術を導入する際には、安全確保に関しても不確実性を排除することは困難であるが、新技術によるベネフィットを社会的に享受するためには、不確実なもの全てを否定・拒絶するのではなく、それらを適正に評価して社会実装を進めることが必要である。

例えば、AI 等を用いる場合には従来のように製造物責任として事業者が過失責任を負えない場合も生じるため、事業者とユーザとが責任を分担・合意するような仕組みが必要である。社会的に新技術の便益を享受するためにリスクを許容し、社会分担するための仕組みとして下記の 3 点が案として挙げられる。（図 3-14 参照）

- ✓ 高い不確実性を伴う IoT 及び AI や安全クリティカルな領域に適用されるデジタル技術に関しては、事故に対応する事業者の社会的な信用力や技術力が重要な問題となるため、認証制度を導入する必要がある。認証制度の対象は、製品・サービスのみならず、事業者のリスクマネジメントシステムも含まれる。例えば、リスクマネジメントを適切に実施できる能力や体制が事業者にあるか、製品・サービスに対して適切なリスクマネジメント実施しているか、リスクに係る情報提供を逐次行うことができるか、等。
- ✓ 事前には想定できないリスクの顕在化として取扱うべき対象については、事業者の合理的なリスクマネジメント（特にリスク探索フェーズ）をもってしてもリスク特定はできないことを前提に、被害補償と不確実性免責制度で対応する。予見性の低いリスクについては事業者に対する免責を行うことで、安全措置の過剰によるイノベーションの阻害を防止しつつ、被害補償を実効化する。

✓ 併せて厳格な制裁制度も必要となる。事業者のリスクマネジメント実施に対して厳格責任を負わせる一方で、リスク顕在化（事故の発生）時に、原因究明のための情報提供や製品・サービスの改善措置を実施する事業者に対しては制裁を免除することで、事業者のリスクマネジメントに対する適切なインセンティブ設計を図る。

以上のような、認証/制裁制度を背景とする事業者（群）による適切なリスクマネジメントの担保と、被害救済制度の導入などを含めて、社会的なリスク分担のあり方の検討が必要である。

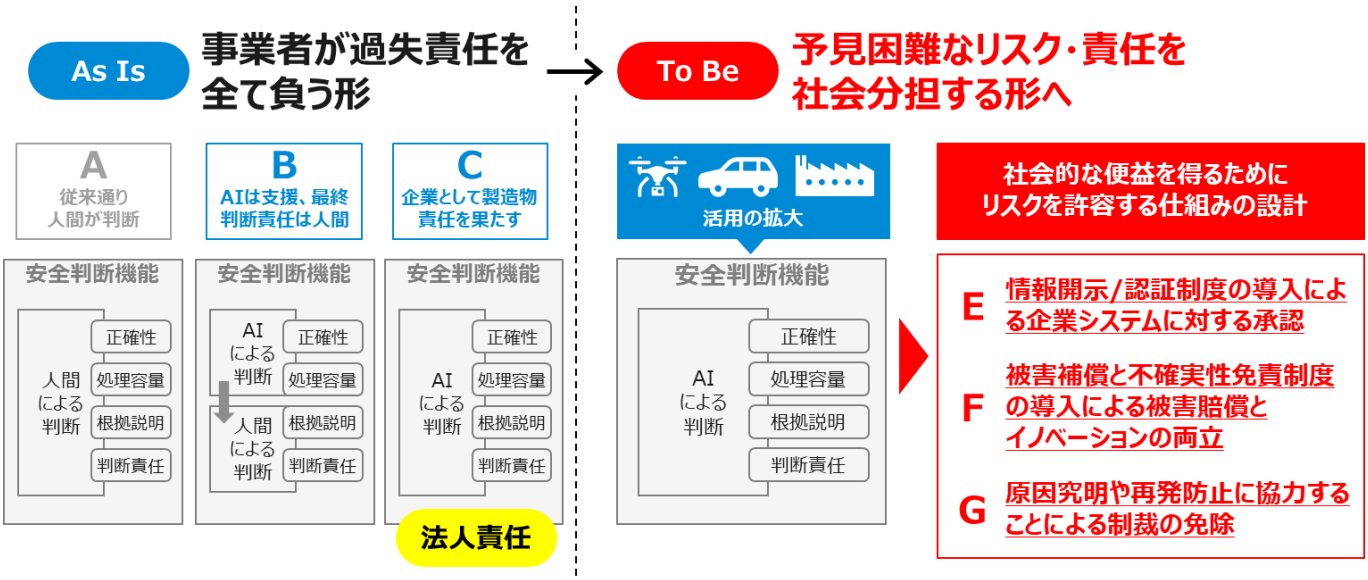


図 3-14　相互理解・責任分担を実現するガバナンス機能（イメージ）

3.4. Society5.0 における安全・ガバナンスモデル

前節で整理したアーキテクチャのビジョンに基づくガバナンス機能を実装した場合に、具体的にどのようなガバナンスモデルになるかというイメージを検討する。図 3-15 はプラント保安分野を事例としたイメージである。

Society5.0 においては、CPS 安全の特徴（変化の速さ、新技術の不確実性、CPS 固有の脆弱性等）を踏まえ、ガバナンスにおける「ルール形成」を踏まえた、「モニタリング（リスク状況の監視、監督）」から「エンフォースメント（安全な状態へと是正）」を、リアルタイムデータ等を活用しながら迅速に改善サイクルを回す必要性が拡大する。

図 3-15 のポイントは以下の通り。

- 国は、当初設定したゴールが達成されているかを確認するゴールベースの法規制を取り入れ、第三者認証機関等を活用しながら事業者の裁量を拡大し、リスクや社会・技術の変化に応じて事業者が合理的に安全を達成可能な形へと移行する。（図 3-15 における F1-1）
- 国を中心に、外部環境や外部システムの様々な状況変化（脅威の発生、脆弱性の変更、仕様の変更、など）を鑑み、ガバナンスの在り方そのものの改善（例：達成すべき安全目標の設定の見直し）を継続的に実施する。（図 3-15 における F1-4 参照）
- ガバナンスを可能な限り、新技術の安全・不確実性に関する情報やノウハウを蓄積した民間組織で運営される形に移行することを目指す。民間組織で機能する構造とすることで動的な変化に柔軟に対応できるガバナンスモデルの構築・運用が可能となる。（図 3-15 における F1-2 参照）
- ガバナンスシステムにおける各民間企業が健全に機能を果たしていることを国が裏支え・監視する機能（リスク情報の真正性を評価する第三者機関認定等）を担い、リアルタイムデータ等を活用したリスクアセスメントを可能とすることで、迅速な改善サイクルを回す仕組みとなる。（図 3-15 における F1-3 参照）
- リスクに関わる様々な主体がともに互いの活動や情報を確認し、それに応じてガバナンスを構成する要員として主体的に活動することが望まれる。市民も「安全の達成状況の可視化・説明の透明化」を通して公開される情報に基づきガバナンスに関与する役割を担う。（図 3-15 における F1-2 参照）
- リスク管理主体である事業者は、コーポレートガバナンスを利かせながら、事業者自ら安全確保の取組に対して継続的な改善活動を実施する。（図 3-15 における青点線を参照）

この検討を通し、提案する官民連携のガバナンス（リスクベースの安全の志向、安全確保に対するゴールベースの要求、アジャイル・ガバナンス）を実装するためには、民間主導のルール設計や認証機能（ゴールを達成しているか適合性評価を実施する等）を担う組織が必要となる。他方で、現状の我が国では、そのような能力・スキルを持つ主体が不足しており、その機能をどの様に社会的に実装していくかが課題となることが本分析を通して示唆として得られた。そこで、現在は認証を担う機関の持つ機能分析についてアーキテクチャ設計を通じ、実施している。その分析の成果は付録 D を参照のこと。付録 D に示した分析はあくまで現状の実態把握にとどまっているが、この分析を通して特定した欧米の認証機関が担う機能をどのように解釈し、日本で実装させていくかを今後の検討課題と認識している。

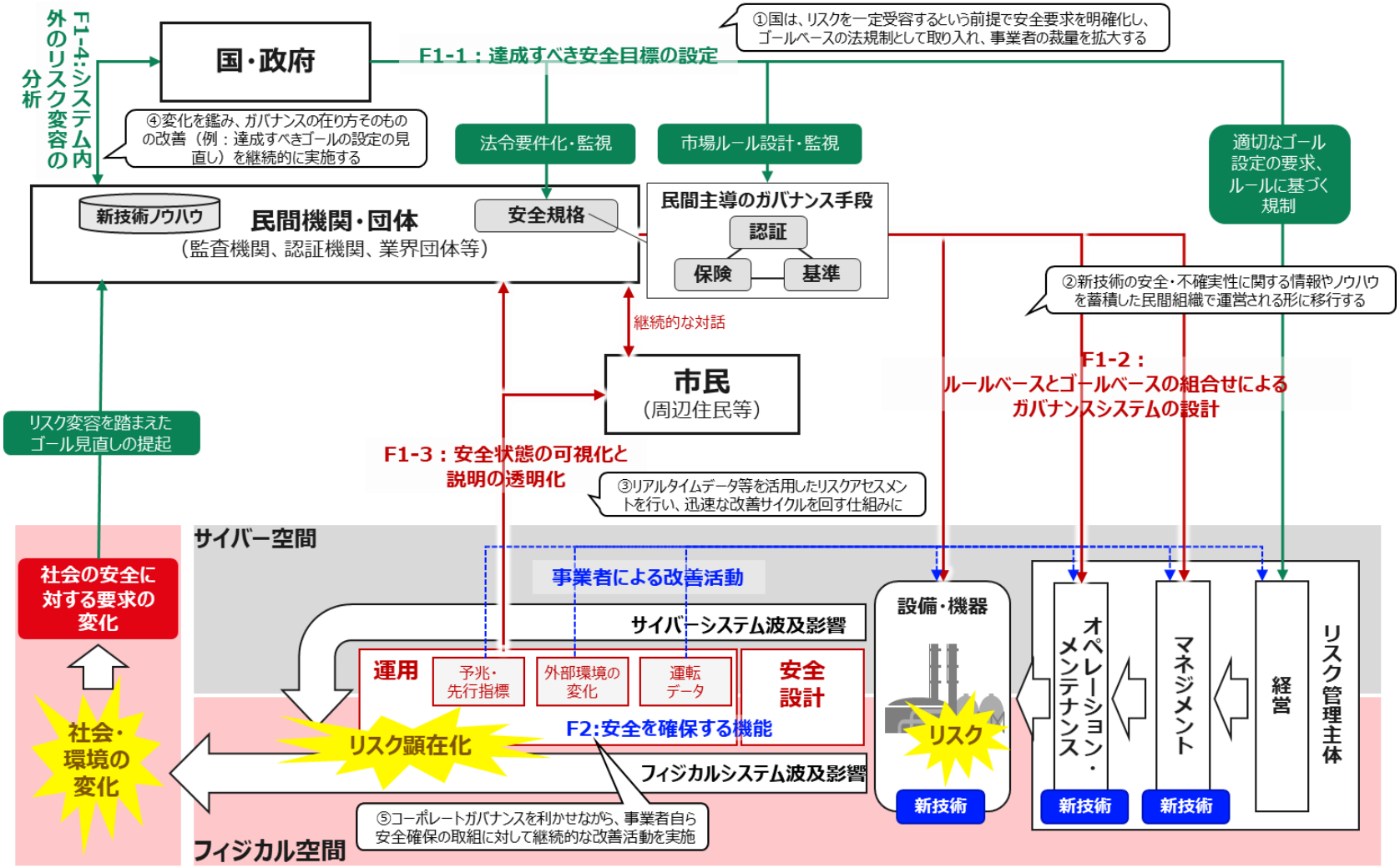


図 3-15　Society5.0 における安全・ガバナンスのイメージ

4. 今後のアクションプラン

本章では、DADC スマート安全プロジェクトの取組の今後のアクションプランについて共有し、様々なステークホルダーが DADC の取組への参画しやすい情報を提供することを目指す。

安全・ガバナンスのアーキテクチャ設計に向けて、2020 年度は DADC の志すビジョンとして将来のあるべき姿を共有することを目指し、本ディスカッションペーパーを発信するに至った。今後はそのビジョンを関連するステークホルダーとともに実践（ビジョンの適用・検証等）し、さらに改善を続けて行きたい。まずは、今回提示した Society5.0 における安全・ガバナンスのアーキテクチャのビジョンに対し、規制側、経営者、業界団体、現場等ステークホルダーからのフィードバックを得て改善を継続していくことが必要である。

より具体的には、今年度以降はより具体的な産業分野をユースケースとして選定して、アーキテクチャ[121]設計を行いながら、具体的な規制等のガバナンス手段の議論を進めて行きたい。産業の現状やあるべき姿に関する「見取り図」となるアーキテクチャを活用することで、前例に囚われない柔軟な思考による全体構想の構築、システムと外界との関係を考慮した相互接続性の確保、俯瞰的・多面的な視点から構成要素間の関係性を定めた全体像の可視化が可能となる。特に、サイバー空間とフィジカル空間の高度な融合が進んだ状況では、サイバー空間のアーキテクチャがフィジカル空間に与える影響が拡大するため、アーキテクチャこそが、人の行動のコントロール、法規制に係る主体の役割分担の改革等のガバナンスイノベーションを実行する基盤となる。具体的なユースケースにおける新たな安全・ガバナンスの在り方、という複雑なシステムをアーキテクチャとして可視化し、それによって多様なステークホルダーと理解を共有して議論したり、現状に縛られずに将来のあるべき姿を実現する手段を検討したりすることで、ビジョン具現化を効果的・効率的に進めることを目指す。

今後は、具体的な産業分野/ユースケースにおけるアーキテクチャ設計を通し、初年度の成果である Society5.0 における安全・ガバナンスの分野横断的なアーキテクチャのビジョンを検証・改善し、同時に、安全・ガバナンスに関するリファレンスアーキテクチャ/アーキテクチャフレームワーク/記述手法の開発を目指す。規制当局、事業者、研究機関等と連携しながら、下記の 3 つの観点を組み合わせながら取組を推進して行きたい。図 4-1 に今後のアクションプランの概要を整理した。

---

[121] アーキテクチャとは、ISO/IEC/IEEE42010:2011 によると「ある環境下におかれたシステムの基本的なコンセプトもしくは特性であり、システムの構成要素とそれらの関係性や、設計と進化の原則となるもの」である。また、慶應 SDM 研究科 白坂成功教授の説明によると、「アーキテクチャとは、全体がどのように目的を実現しているかの基本的なコンセプトや構想（実現の方向性）を示すものであり、その構想（実現の方向性）は目的を達成するためのシステムと外界との関係、システムを構成する要素、構成要素間の関係により定められ、表現されるもの」である。

① 分野横断的なアーキテクチャのビジョンの継続的改善：初年度は「GOVERNANCE INNOVATION VER.2: アジャイル・ガバナンスのデザインと実装に向けて」を踏まえ、安全に特化した切り口で、多様な産業分野を考慮しつつも、主にはプラント分野等のレガシーシステムを想定しながらガバナンスの在り方を検討した。規制側・経営者・業界団体・現場等ステークホルダーからのフィードバックを得ることや、具体的な産業分野における実践を通して改善を継続していく。

② 具体的な産業分野におけるアーキテクチャ設計：まず、①を基にした具体的な産業分野における安全・ガバナンスのビジョンの明確化を行う。その上で、個別具体の産業分野をユースケースとして調査・分析し、法改正を含む新たなガバナンスの仕組みの提案や産業現場におけるシステム、取組のあるべき姿の構築を行い、その実現に向けた仕掛けやアクションプランを具体化させ、本格的なアーキテクチャの設計を行う。
まずは、DADC で推進されている PJ（自律移動システム等）を事例として、事業者等との連携を加速し具体の産業の CPS の姿の具体化を進め、安全規制に係る改革のみならず、安全にまつわる産業戦略（新市場の創出、事業者の競争力の根幹）を含むビジョン/アーキテクチャとするための仕掛けを検討していく。

③ ②の成果の他分野への展開による横断的な成果の創出：具体的な産業分野における検討成果を他分野へと適用可能なリファレンスアーキテクチャ/アーキテクチャフレームワーク/記述手法に開発し、展開する。

図 4-1　今後のアクションプラン

5. まとめ

本ディスカッションペーパーでは、まず議論の前提となる従来の知見や現在国際的に進んでいる技術と安全に関する議論や、技術とガバナンスにおける議論、日本の現行ガバナンスの分析を踏まえて、Society5.0 への進展に伴う安全に係る課題を特定し、従来の安全に関する考え方・ガバナンスから、Society5.0 に向けて安全・ガバナンスはどのように変化すべき/していくかを捉えた（第 2 章）。その上で、目指すべき Society5.0 における安全・ガバナンスのアーキテクチャのビジョン及び課題認識を共有した（第 3 章）。さらに、示した Society5.0 において目指すべき安全・ガバナンスのあり方の実装に向けて、具体的な分野をユースケースとして取り上げ、DADC の今後のアクションプランを共有・提案した（第 4 章）。

プロジェクト初年度となる本年度ではあくまで安全・ガバナンスのビジョンの共有を行うことで、幅広いステークホルダーに将来起こり得る（既に起こりつつある）変化に対する現実的な危機感を持っていただき、変革に対する DADC のビジョン・志に共感していただくことを目指した。多くの方々にとって、本ディスカッションペーパーが DADC のビジョン・志に対する理解の助けとなり、DADC の取組への参画への関心を高めるものとなれば幸いである。今後は、今回提示したビジョンについて具体的な分野での実践を行いたいと考えている。実践を通し、ビジョンの改善も継続的に続けて行きたい。

初年度の活動を通して得られた、安全・ガバナンスのアーキテクチャのビジョンを実装するための示唆のうち、特に以下 2 点が重要と捉えている：①国内の足下の課題から着手し、積み重ねられた安全に係る考え方を含むシステムをどう Society5.0 に適したものにしていくのか、システムの CPS 化した姿を具体化したうえで、新たな安全・ガバナンスの実装・転換の在り方を検討することがポイントとなること、②新興の産業分野・ビジネス（例えば、自律移動システム等）を主な対象として、安全を競争力に繋げるための安全評価の仕組み・仕掛けの提案が重要であること。今後の活動においては、これらの示唆に留意して具体的な施策の検討を進める計画である。

いずれの産業分野においても「安全」は重要な要素であり、産業の発展のための根幹ともなる。さらに、2 章及び 3 章で示した通り、Society5.0 において技術の高度化や社会システムの CPS 化/SoS 化により、単体の事業者で安全確保を達成することはこれまでに比べ困難となる。そのため、産業横断の安全・ガバナンスの在り方や共通機能についての議論が必要となる。本ディスカッションペーパーで提起した論点を検討しつつ、日本の産業の優位性を考慮した上で、ルール形成における他国とのインターオペラビリティを確保するために、国際的な議論に官民連携で参画していくことも重要となるだろう。様々なステークホルダーによる協力を経て、挑戦的な取組を可能とするリスクベースの安全の考え方を根付かせ、我が国の強みを考慮

した競争領域・協調領域の特定を行い、日本にとって競争力を発揮しやすいガバナンスを検討していきたい。

本ディスカッションペーパーにおける検討内容及びビジョンについて、今後も産業界及び学術界と連携しながら議論を継続し、Society5.0 における安全・ガバナンスの実現に向けた取組を促進させる所存である。現状のより深い理解に必要な関係各位のご支援、新たな産業アーキテクチャ実現に対するご参画を重ねてお願い申し上げる。

著者

プロジェクトリーダー：高橋　久実子
研究員：石坂　彰
研究員：三浦　和夫
研究員：藤田　裕志