適合性評価SWG活動報告

適合性評価SWG
AISI標準チーム　髙村博紀

2026年3月10日
AISI事業実証WG 下期報告会

AISI Japan AI Safety Institute

AISI Japan AI Safety Institute

適合性評価SWGについて

適合性評価SWGの概要

AISI Japan AI Safety Institute

- 分野別SWGの取組みを通じて、自己適合評価を含むAI適合性評価手法の確立を目指す。
- データ品質SWGとの連携により、AI出力の整合性、利用文脈を含めた複合的な評価の方法を検討する。

現状と課題

- 従来の適合性評価制度では、変化し続ける対象への評価に対応しきれない
- ISO/IEC 42007開発中(適合性評価)はハイレベル枠組み

⇒柔軟かつ適切な評価方法と実行可能手順が必要

適合性評価手法の確立への取組み方針

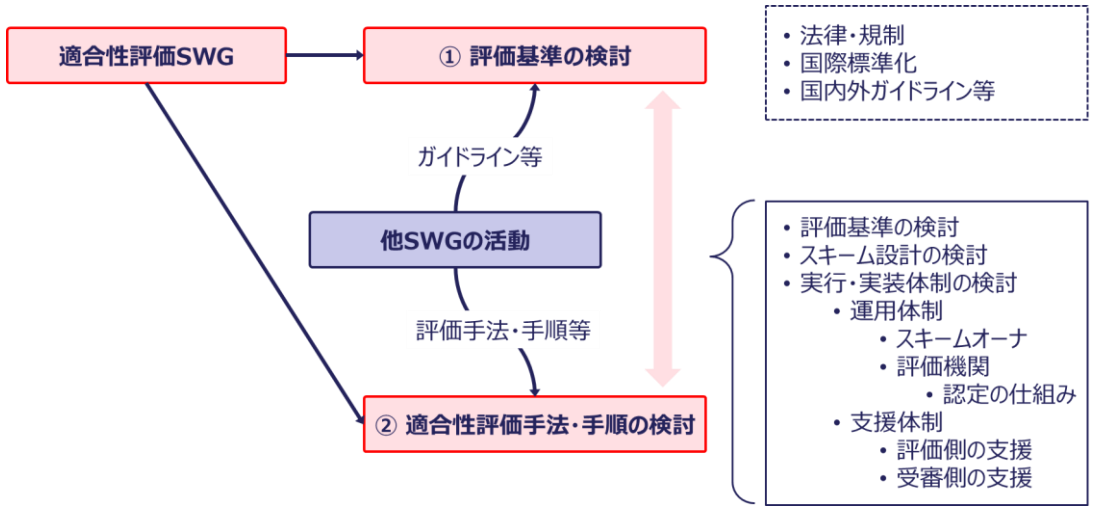
- 分野別SWGの取組みから、AIへの要求事項等導出

<具体的なステップ>

ヒアリング・課題抽出 → 制度設計に向けた検討 → ツール・体制整備 → 国際標準との整合

- データ品質SWGと連携し、複合的な評価方法の検討

適合性評価手法の確立への取組み方針



SWGのロードマップ

短期的な取組み（令和7年度）	中期的な取組み（令和8年度～9年度）	長期的な取組み（将来的なビジョン）
● 分野別SWGとの連携による評価ニーズや受審組織の現状に関するヒアリング ● 適合性評価の関連機関を中心としたSWGを組成し、既存の適合性評価手法を分析	● 国際標準化議論を踏まえた、包括的なアシュアランスとアカウンタビリティ確保等に係る適合性評価制度の設計 ● AIに関する適合性評価制度のあり方の具体化と試行により実効性を検証	● 国際的な実装モデルとの整合性を踏まえたガイドライン整備 ● 国際的に総合運用可能な実運用体制構築を視野に入れた段階的な展開

構成メンバー

AISI Japan AI Safety Institute

No.	所属	（敬称略）
1	AIセーフティ・インスティテュート（AISI）	髙村　博紀（リーダー）
2	公益財団法人 日本適合性認定協会（JAB）	塩森　淳
		征矢　義弘
3	一般社団法人 情報マネジメントシステム認定センター（ISMS-AC）	山内　徹
		畔津　布岐
4	一般財団法人日本規格協会（JSA）	中川　梓
5	独立行政法人 製品評価技術基盤機構（NITE）認定センター	吉田　耕太郎
6	一般財団法人 日本品質保証機構（JQA）	浅田　純男
		千葉　翔太
7	国立研究開発法人 産業技術総合研究所（AIST）	江川　尚志
		丸山　文宏
		杉村　領一

有識者委員会でのAI適合性評価の検討

AISI Japan AI Safety Institute

- Digitalサービス／プロダクトに対する適合性評価の在り方について検討する
- ISO/IEC　SC42の活動およびISO CASCOの活動と連動する形で適合性評価の在り方
- 制度設計を関係者と検討：評価対象、評価基準、評価手法、評価体制、手続き、手順

- **「AI利活用促進に向けたAIセーフティ評価に関する事業実証」の適合性評価SWG**
- **内閣府BRIDGE：「AI分野におけるJoint Certificationの検討」**

に共同の有識者委員会を設置、2025年度は認定機関および標準化機関から専門家に参加いただき、委員会を組成

- **既存の適合性評価制度の現状と課題**
- **課題解決に向けた検討**

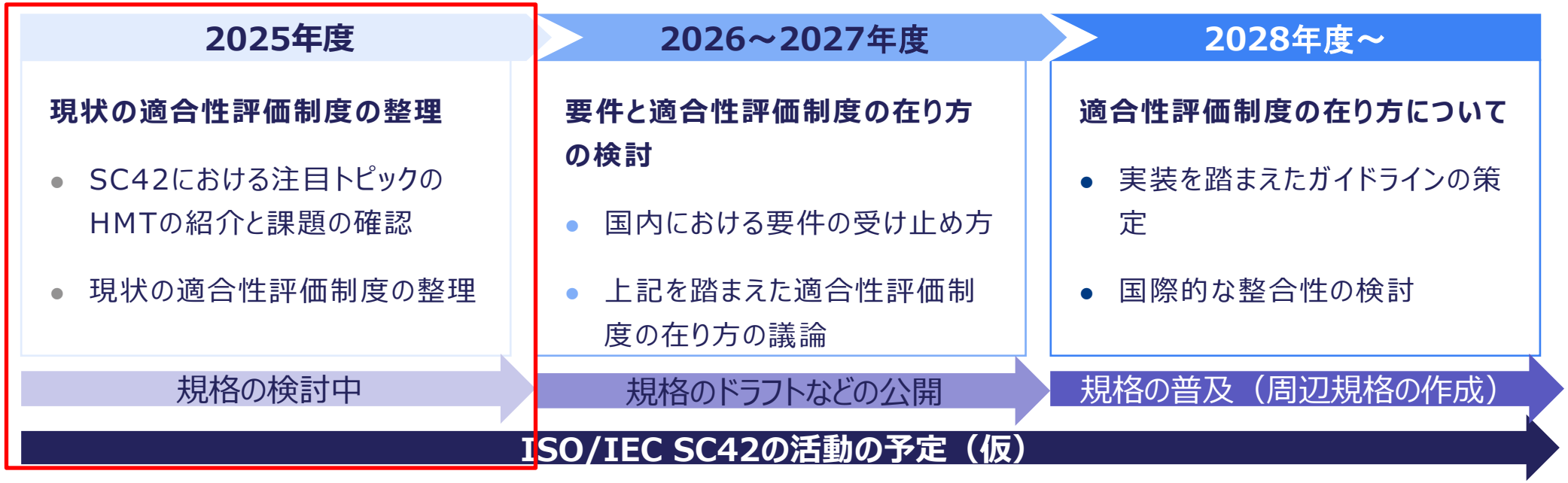
をHuman-Machine Teaming(HMT:SC42にて規格開発中のテーマの一つ)を産総研の研究者から紹介いただき、ユースケースとして御議論いただく

人、AIシステム、チームとして、サービスのライフサイクルの各フェーズにおいて何をどう評価するか？その評価基準、評価手法は？など

適合性評価SWGの目的とタイムライン

AISI Japan AI Safety Institute

- AIセーフティに関する要件を踏まえた適合性評価制度の在り方について議論し、日本において適切な適合性評価制度を提言する。
- 一方で、現在AIセーフティに関する要件に関しては、ISO/IEC SC42において議論がされており、まだドラフトや検討中の段階である。
- そのため、2025年度としては、今後のISO/IEC SC42において標準化される要件が策定される前に、AIセーフティの適合性評価制度の課題や現状の適合性評価制度の整理を行うことを想定する。



AI適合性評価：考えるべき事項

AISI Japan AI Safety Institute

- **制度設計に向けた検討**　／　相互運用・相互承認
  - ISO/IEC規格の活用、諸外国の制度・海外民間主導の取組との関係

変化し続ける対象を
どう評価するか
評価基準・評価手法

- 170XX
- CB制度

ISO CASCO
IEC CAB

適合性評価
関連規格で連携

SC42

- 評価基準
  - 法律・規制
  - 国際規格
  - ガイドライン等
- スキーム設計
  - **ISO/IEC規格の活用**
- 実行・実装体制
  - 運用体制
    - スキームオーナ
    - 評価機関
    - 認定機関
  - 支援体制
    - 評価側の支援
    - 受審側の支援

①評価基準の検討

AI適合性評価

②適合性評価手法・手順の検討

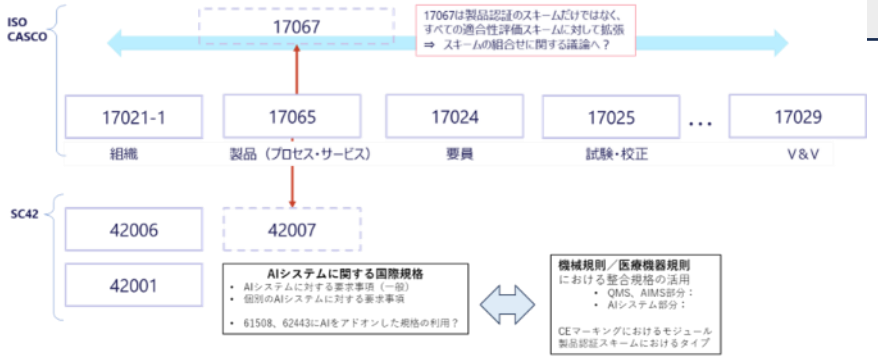
法律／規制

諸外国の制度
欧州・米国・その他

- 制度設計
  - 相互運用性
  - 相互承認
- 国内体制
  - スキームオーナ
  - 評価機関
  - 評価機関の認定

海外の民間主導の取り組み
- AISIC
- AI Trust Alliance
- AI Verify Foundation等

適合性評価に関する国際規格はISO CASCO（いわゆる17000シリーズ）にて
開発・維持されている。AI分野における関連規格との関係を図示する



今年度の検討テーマ

AISI Japan AI Safety Institute

① 既存の適合性評価、AI標準化（SC42等）との関係の整理
② HMTを対象にした場合の適合性評価について議論できればと思います

適合性評価SWGは標準型BRIDGEと（部分的に）連動していきます

- ◆ 既存の適合性評価制度について整理
  - ● 参加機関から取り組みを紹介いただき、整理する
- ◆ AI標準化活動等、関係する活動について調査
  - ● ISO/IEC JTC1 SC42におけるAI標準化活動で適合性評価に関係するもの
  - ● Joint Certificationなど新たな適合性評価に関する動向
- ◆ HMTをテーマに、適合性評価制度の枠組みについて検討いただく
  - ● H-Mをシステムとして、ひとまとめにした場合：チーム評価
  - ● Mのレベルに応じてHに求められる力量：要員評価
  - ● M（AIシステム）として求められる要求事項：AIシステム評価

など考えられると思われますが、例えば、医療システムでは？など用いて、抽象―具体と議論・検討

今年度の成果

AISI Japan AI Safety Institute

◆既存の適合性評価の整理（参加機関の活動から）
◆HMTを題材にAI適合性評価の在り方について検討

報告書

- 報告書　目次
  - はじめに：①背景・目的、②本書の位置づけ、③対象読者
  - 適合性評価制度について
  - 国内におけるAIに関する取組み概要
  - 参加組織の活動紹介
  - AI適合性評価制度に関する取組み：
    ①SC42の活動概要、②欧州AI法と標準・適合性評価、③本SWGの取組
  - 今後の活動

AI標準化：ISO/IEC JTC1 SC42の活動から

AI標準化について：ISO/IEC JTC1 SC42の活動から

AISI Japan AI Safety Institute

◆ **ISO/IEC JTC1 SC42**で活動しています（SC42国内幹事等を拝命）

◆ 他のAI標準化活動（ITU、AI for GoodやGPAI等）は組織間連携にて対応

SC42最近の話題

- ISO/IEC42001が出版されました、**42001認証**がスタートしています
- 欧州AI法の整合規格（hEN）とSC42での規格の関係が明確になってきました
- AIシステムに対する**適合性評価**に関するハイレベル・フレームワーク規格42007が**ISO CASCO**の17067改定作業と連携してはじまっています
- 機能安全とAI：ISO/IEC TS22440が開発中です
- Red Teamingなど含め、AIのテスティング技術に関する規格開発がはじまりました：42199シリーズ
- 日本主導の**Human-Machine Teaming**の規格開発がはじまっています
- 金融分野とAIのJWG7が発足しました

◆ ISO/IECの合同委員会JTC1の下にあるサブコミッティのひとつがSC42です
◆ 国内体制は、産総研　杉村（HoD）、橋本、髙村（幹事）を中心にメーカ、大学から委員参加
◆ 事務局はJISCより委託を受けている情報処理学会　情報規格調査会が担当
　◆ JISC：経済産業省に設置されている審議会で、Japanese Industrial Standards Committeeの略称

国内SC42活動

AISI Japan AI Safety Institute

- 委員長：杉村、幹事：橋本、髙村
- WG主査・幹事
  - WG1　基礎標準：杉村・髙村
  - WG2　データ：金・林谷
  - WG3　信頼性：堀江・相薗
  - WG4　ユースケースとアプリケーション：細川・鄭
  - WG5　AIシステムの計算手法と計算特性：坂本
  - JWG2　AIベースのシステムテスト：アンドレ・木下
  - JWG3　AIを活用した保健医療情報：津本・今井・村瀬
  - JWG4　機能安全とAIシステム：相薗・神余
  - JWG5　自然言語処理：坂本
  - JWG6　適合性評価（杉村）・長谷川、髙村
  - JWG7　AI対応金融システム：（杉村）・原

（敬称・所属略）
https://itscj.ipsj.or.jp/committee-activities/committee-list.html

- JIS原案作成委員会
  - 42001のJIS原案作成完了　⇒　発行
  - 今年度は42005と42006

- 産総研受託の標準化加速事業（経産省）
  一部委員が参加
  - Human Machine TeamingとAIマネジメントシステムの2つに関する調査など
  - 規格の提案、海外関係者との意見交換、シンポジウムの開催など
- 情報交換会（ほぼ毎週）：チャタムハウスルールにて開催

ご関心ありましたら是非ともご参加ください！

- 専門委員会：各WGでの審議事項の確認と技術委員会対応など（月一開催）
- 主査・幹事会：各WGの進捗状況、課題など情報共有（月一開催）
- 各WG：月2回程度開催（審議案件次第）

Human-Machine Teaming（HMT）を題材に AI適合性評価のあり方について検討

AI適合性評価

AISI Japan AI Safety Institute

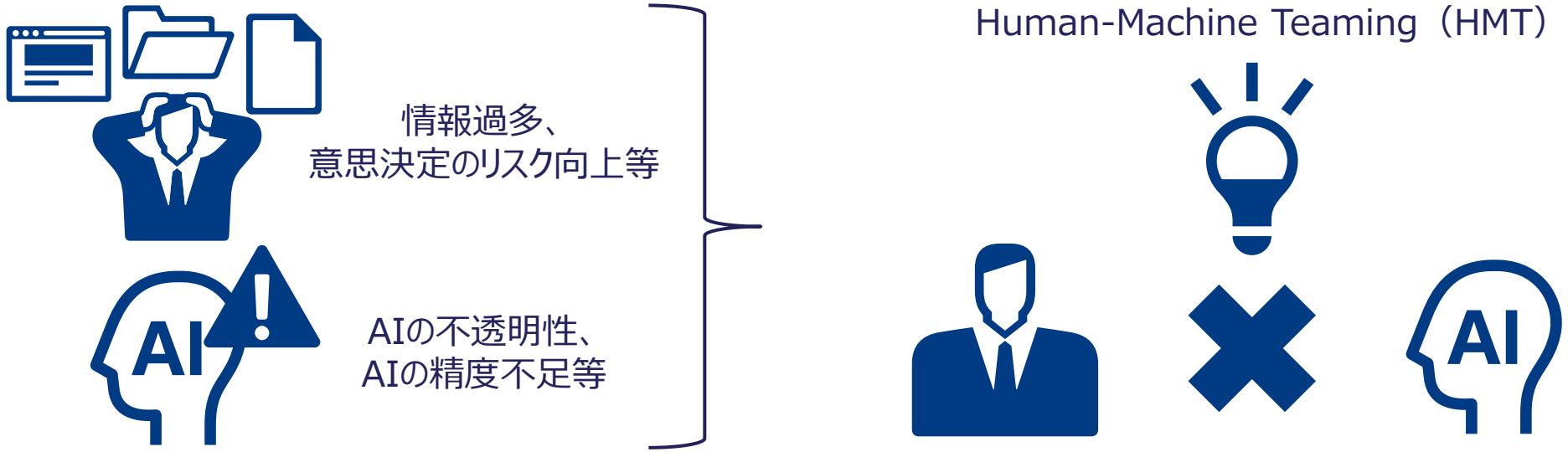
- AI法においても標準の活用がうたわれており、標準に準拠していることを示すには適合性評価が重要
- 欧州における法規制とその実装のための標準（整合規格）の活用（NLF）が知られているが、標準は法に比べて、まだ、柔軟に改定ができる（Amendmentも含め）
- 特にAIなど技術革新のスピードが速く、社会的影響が大きいものに対しては法と標準の関係を明らかにしつつ、法に準拠していることを、関連する標準への準拠に置き換えていく必要がある

- AIを含むDigitalプロダクト／サービスに対する適合性評価を考えると、
  - **変化が前提**：対象が変化し続ける、環境の変化も発生、それに応じてガバナンス・マネジメントも変化せざるを得ない
  - **関係者が多様**：マルチステークホルダ、サプライチェーンのみならず、リスクチェーン、バリューチェーンなども考慮する必要がある

- **問い：<u>組織／要員／製品と縦割りでの適合性評価でAIセーフティを担保できるのか？</u>**

Human-Machine Teaming（HMT）とは

- Human-Machine Teaming（HMT）は、ISO/IEC 22989:2022「Artificial intelligence concepts and terminology」において、「Integration of human interaction with machine intelligence capabilities」として定義された概念である。
- 日本語では、「人・AI協働」とほぼ同義であり、目的達成に向けて人とAIが協力して取り組むことを示している。
- 本概念が生まれた背景として、情報化社会において人間のみで意思決定を行うことが難しいと同時に、情報処理に長けている一方でAIも精度が不十分である場合があり、人とAIが協働することの必要性が確認された。
- 上記の背景の中、 人とAIの最適な協働の在り方を検討するためにHMTの概念が定義、人とAIの協働が類型化されている。

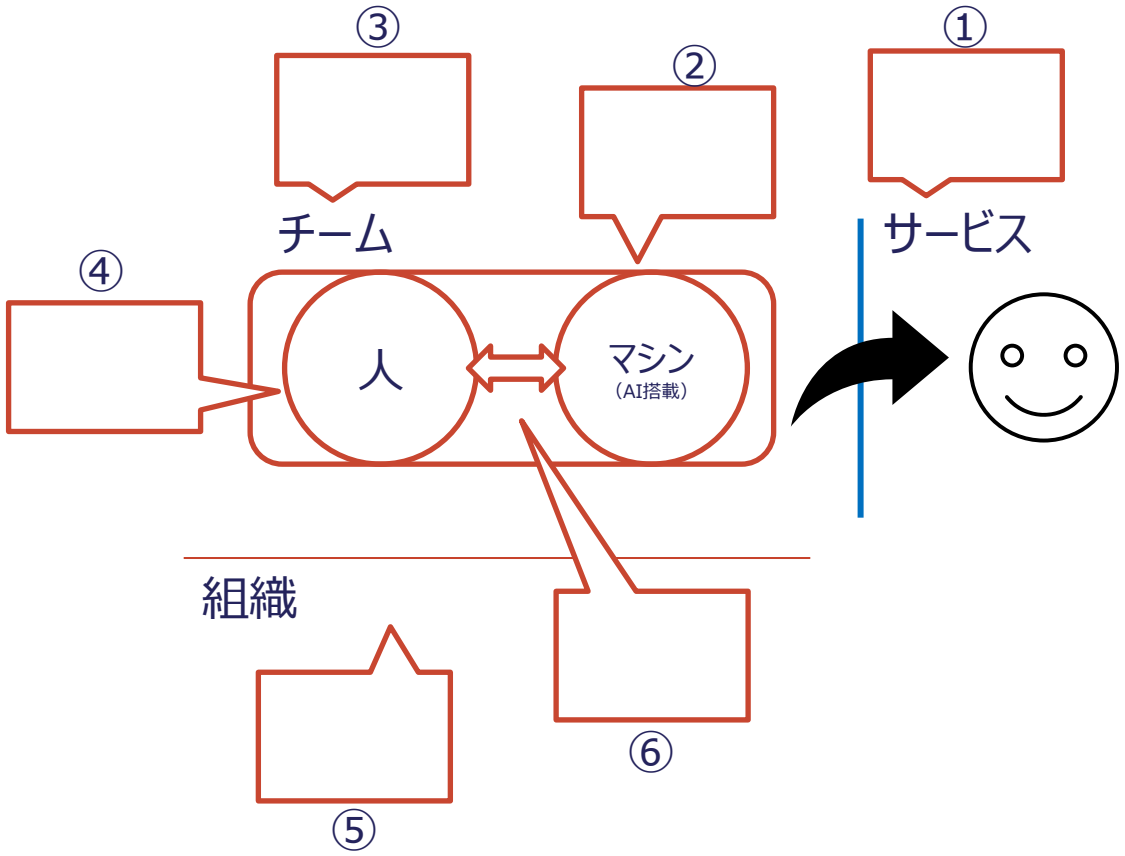


出所）CRDS、研究開発の俯瞰報告書　システム・情報科学技術分野（2024年）、https://www.jst.go.jp/crds/pdf/2024/FR/CRDS-FY2024-FR-03/CRDS-FY2024-FR-03_20105.pdf

何を、どう、評価すればいいか

AISI Japan AI Safety Institute

HMTの類型に応じて評価対象は変わるべきか？
HMTの類型に応じて評価対象の重みづけなどはあるか？
HMTの類型に応じて課題はあるか？
エンドユーザーを考えた場合の懸念点はあるか？
などが論点となる。



評価対象
① サービス
② マシン
③ チーム
④ 人
⑤ 組織
⑥ 相互作用

安全性(Safety)とAIセーフティ

AISI Japan AI Safety Institute

- 安全性(Safety) ～ISO/IEC GUIDE 51:2014(E)～
  - Safety: Freedom from risk which is not tolerable
  - 安全とは、**許容不可なリスクがない**こと。

- AIセーフティ ～AIセーフティに関する評価観点ガイド～
  - **人間中心**の考え方をもとに、
    AI活用に伴う社会的リスク(※)を低減させるための**安全性・公平性**、
    個人情報の不適正な利用等を防止するための**プライバシー保護**、
    AIシステムの脆弱性等や外部からの攻撃等のリスクに対応するための**セキュリティ確保**、
    システムの検証可能性を確保し、
    適切な情報提供を行うための**透明性**　が保たれた状態。
    ※社会的リスクには、物理的、心理的、経済的リスクも含む。

AIセーフティとは？／AIセーフティ評価とは？

Holistic Approach is necessary for AI safety

AISI Japan AI Safety Institute

- Holistic Approach is also necessary for AI Conformity Assessment
- ∑ Conformity Assessment Schemes ⇔ Conformity Assessment System

Conformity Assessment for
- Organization
- Personnel
- Product (Process, Service)

Inspection, Testing, V&V…

日本の浮世絵師、英一蝶 (1652 – 1724) による『衆瞽象を撫ず』図

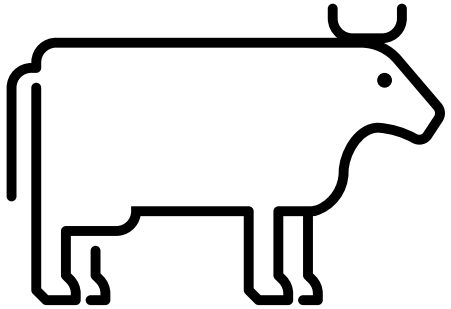
知りたいことは象とは何？であり、つまり、
知りたいことは、このAIって大丈夫？であるか
何をみれば大丈夫といえるのか？

How to combine each scheme to systems
Beyond the SLCA: System level conformity assessment

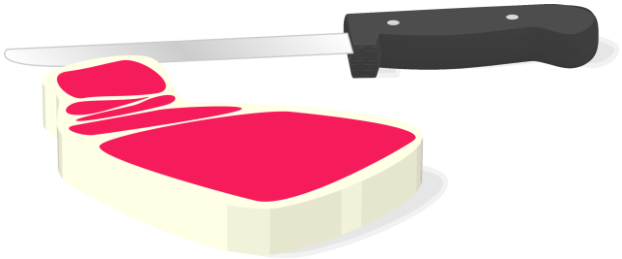
A simple question: limits of reductionism?

AISI Japan AI Safety Institute

Properly raised, healthy cattle are delicious in every cut.



The tenderloin is good



System $\stackrel{?}{=} \sum$ Parts
Reductionism is enough?

the services/products created by the organization is properly managed and trustworthy enough?

An organization has obtained XX MS certification.
Some product has obtained product certification in an organization.

Is reductionism sufficient ?, AI
How XX is XX enough? , XX ∈ {safety, secure, reliable,・・・}

Holistic approach is necessary for AI safety: Top-Down & Bottom-Up approaches

課題解決のヒントとして: System Thinking, etc.

AISI Japan AI Safety Institute

- 合成の誤謬（Fallacy of Composition）
- Total Optimization ≠ Σ Local Optimization
- Systems are connected: System of Systems, Open systems

"A system is a way of looking at the world."
"… a system, any system, is the point of view of one or several observers."
Gerald M. Weinberg (1975)
An Introduction to General Systems Thinking

木を見て森を見ず ⇒ 木も見て森も見る
群盲評象 ⇒ 全容（貌）把握

PICARD Theory of Systems

System = Holistic Image of { Parts, Interactions, Context, Actions, Relationships, Destination }

From the Point of View of an Observer

Holistic Approach to Finding the Whole Solution: Using Systems Principles & Concepts, James N Martin

- DEOS Project (JST CREST、研究統括 所 真理雄): Dependable Engineering for Open Systems
  成果の一つとしてIEC 62853 Open Systems Dependability 2018

  http://deos.or.jp/index-j.html
  https://www.kindaikagaku.co.jp/book_list/detail/9784764904613/

- Agile Governance (METI)

  https://www.meti.go.jp/press/2022/08/20220808001/20220808001.html
  https://www.meti.go.jp/shingikai/mono_info_service/governance_model_kento/index.html
  https://www.meti.go.jp/shingikai/mono_info_service/governance_model_kento/pdf/20250930_1.pdf

- SoS Governance guideline (NEDO 産業DX、立命館大学：研究統括 徳田昭雄)

  https://www.ritsumei.ac.jp/research/idx-agpf/250515-0039-01.pdf

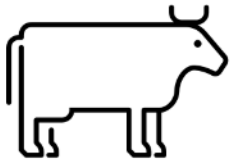
AIセーフティ評価とAI適合性評価

AIセーフティ評価の一部として、AI適合性評価によりカバー可能な部分
既存の適合性評価の整理＆AI適合性評価の検討：HMTを題材に

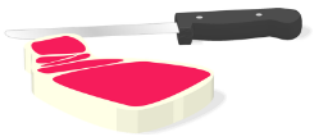
- AIセーフティ評価に必要なAI適合性評価について引き続き検討を進め、制度設計についても考えていく予定です

A simple question: limits of reductionism?

Properly raised, healthy cattle are delicious in every cut.



The tenderloin is good



System ? Σ Parts
Reductionism is enough?

the services/products created by the organization is properly managed and trustworthy enough?

An organization has obtained XX MS certification.
Some product has obtained product certification in an organization.

Is reductionism sufficient ?, AI
How XX is XX enough? , XX ∈ {safety, secure, reliable,・・・}

Holistic approach is necessary for AI safety: Top-Down & Bottom-Up approaches

課題解決のヒントとして: System Thinking, etc.

- 合成の誤謬（Fallacy of Composition）
- Total Optimization ≠ Σ Local Optimization
- Systems are connected: System of Systems, Open systems

"A system is a way of looking at the world."
"… a system, any system, is the point of view of one or several observers."
Gerald M. Weinberg (1975)
An Introduction to General Systems Thinking

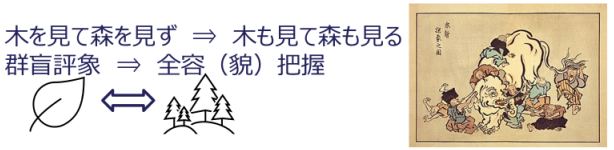
PICARD Theory of Systems

System = Holistic Image of { Parts, Interactions, Context, Actions, Relationships, Destination }

From the Point of View of an Observer

Holistic Approach to Finding the Whole Solution: Using Systems Principles & Concepts, James N Martin

木を見て森を見ず ⇒ 木も見て森も見る
群盲評象 ⇒ 全容（貌）把握

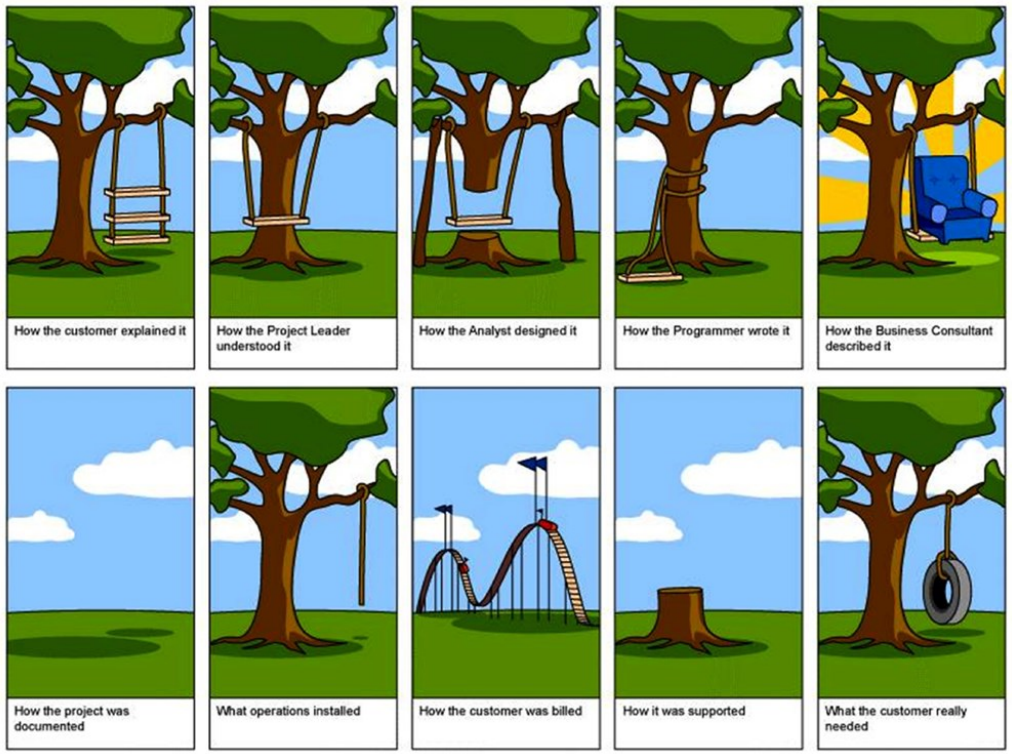


- DEOS Project (JST CREST、研究統括 所 真理雄): Dependable Engineering for Open Systems
  成果の一つとしてIEC 62853 Open Systems Dependability 2018
  http://deos.or.jp/index-j.html
  https://www.kindaikagaku.co.jp/book_list/detail/9784764904613/
- Agile Governance (METI)
  https://www.meti.go.jp/press/2022/08/20220808001/20220808001.html
  https://www.meti.go.jp/shingikai/mono_info_service/governance_model_kento/index.html
  https://www.meti.go.jp/shingikai/mono_info_service/governance_model_kento/pdf/20250930_1.pdf
- SoS Governance guideline (NEDO 産業DX、立命館大学：研究統括 徳田昭雄)
  https://www.ritsumei.ac.jp/research/idx-agpf/250515-0039-01.pdf

本当に必要なAI適合性評価とは？（特にAIセーフティ）

AISI Japan AI Safety Institute

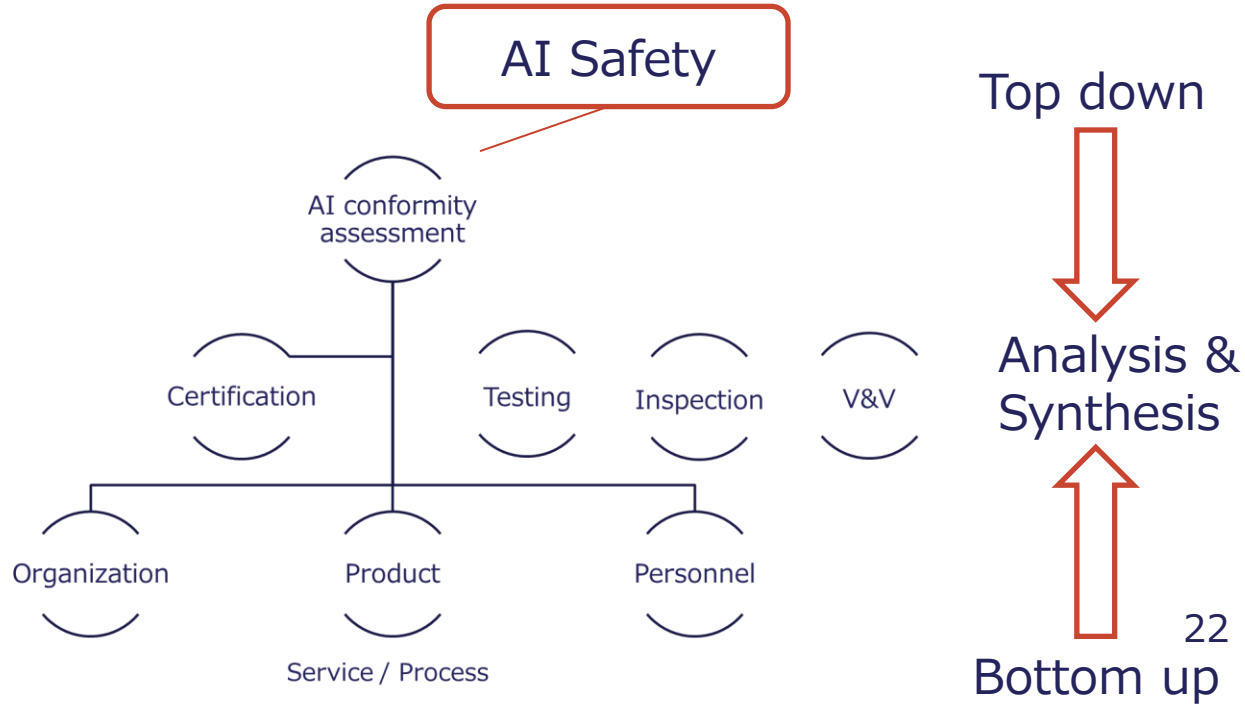
要求を明確化することは容易ではない！



Wants and Needs ⇒ Requirements is NOT easy

What is truly wanted and needed for conformity assessment in AI Safety?

- This painting is commonly known as the “Swing-Tree Story.” It is a very famous illustration that has long been used in software engineering, first appearing in the newsletter of the University of London Computer Centre in 1973.



BRiDGE標準活用加速化支援事業の取組みとも連動

AISI Japan AI Safety Institute

AI分野に対するJoint Certificationの検討

- ◆ 調査：海外関係機関との意見交換と連携
- ◆ AI適合性評価の検討（パイロット評価等）、勝ち筋シナリオの検討（適合性評価SWGと合同委員会）
- ◆ 京都大学と共同研究（IPA・AISI）：Legal Engineering for Software Defined Society

- ♦ AIセーフティ評価とAI適合性評価の関係整理／既存の適合性評価再考：HMTを題材にブレストから
- ♦ 相互運用／相互承認に向けて

AI適合性評価

AISI Japan AI Safety Institute

- ♦ AI法においても標準の活用がうたわれており、標準に準拠していることを示すには適合性評価が重要
- ♦ 欧州における法規制とその実装のための標準（整合規格）の活用（NLF）が知られているが、標準は法に比べて、まだ、柔軟に改定ができる（Amendmentも含め）
- ♦ 特にAIなど技術革新のスピードが速く、社会的影響が大きいものに対しては法と標準の関係を明らかにしつつ、法に準拠していることを、関連する標準への準拠に置き換えていく必要がある

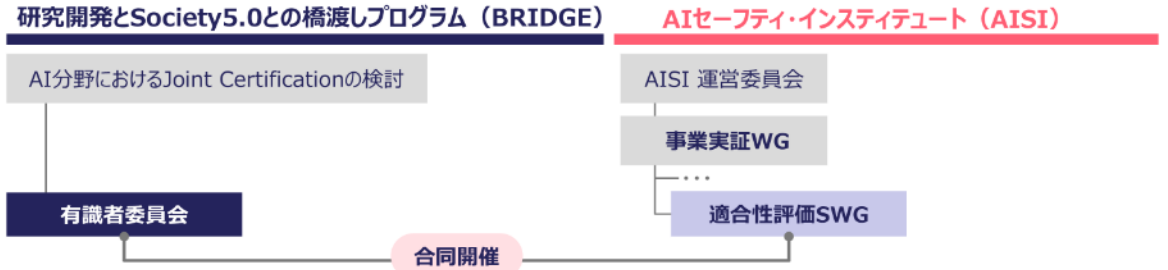
- ♦ AIを含むDigitalプロダクト／サービスに対する適合性評価を考えると、
  - **変化が前提**：対象が変化し続ける、環境の変化も発生、それに応じてガバナンス・マネジメントも変化せざるを得ない
  - **関係者が多様**：マルチステークホルダ、サプライチェーンのみならず、リスクチェーン、バリューチェーンなども考慮する必要がある

⇩

- **問い**：**組織／要員／製品と縦割りでの適合性評価でAIセーフティを担保できるのか？**

体制・参加組織

AISI Japan AI Safety Institute



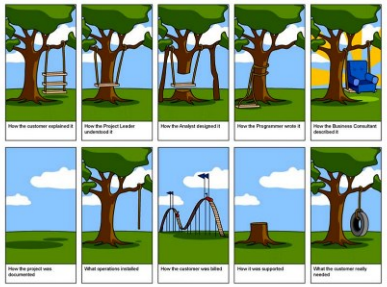
公益財団法人日本適合性認定協会（JAB）
一般社団法人情報マネジメントシステム認定センター（ISMS-AC）
一般財団法人日本規格協会（JSA）
独立行政法人製品評価技術基盤機構（NITE）認定センター
一般財団法人　日本品質保証機構（JQA）：CASCO国内委員長
国立研究開発法人産業技術総合研究所：SC42国内委員長他

AIサービスが安全・安心・快適に提供されていること
TrustworthyなDependableなサービスが継続的に提供されていること

AISI Japan AI Safety Institute

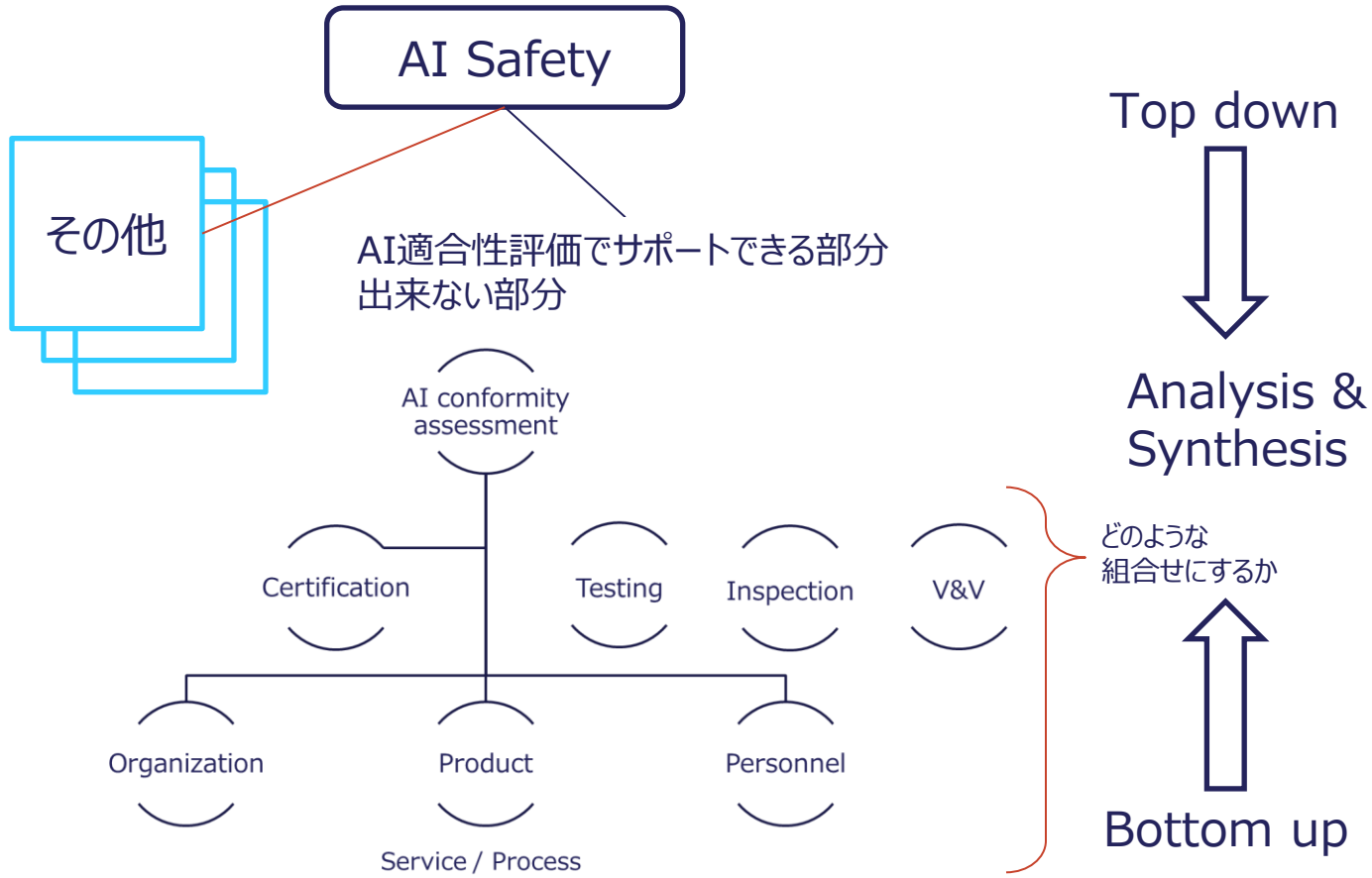
Wants and Needs ⇒ Requirements is NOT easy

What is truly wanted and needed for conformity assessment in AI Safety ?



- This painting is commonly known as the "Swing-Tree Story." It is a very famous illustration that has long been used in software engineering, first appearing in the newsletter of the University of London Computer Centre in 1973.

本質的なもの（要件）を抽出して、AIセーフティ評価、
それを（部分的に）サポートするAI適合性評価の枠組みを構築すること！

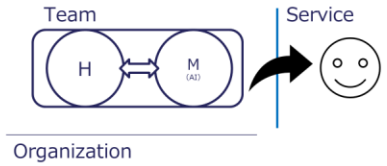


Currently activities for AI Conformity Assessment is

AISI Japan AI Safety Institute

Conformity Assessment SWG is examining conformity assessment procedures when HMT is the subject of evaluation.

- Human-Machine Teaming（HMT）：「Integration of human interaction with machine intelligence capabilities」(ISO/IEC 22989,Artificial intelligence concepts and terminology:2022)
- HMT has 5 types; Human supervised, Human mentor, Peer-Peer, Machine mentor, and Machine supervised



あくまで私案

継続的な評価
審査側・受審側コミュニケーション

AISI Japan AI Safety Institute

AIセーフティ評価とAI適合性評価：制度設計するとしたら

AI開発者・提供者の自発的活動
＋
（必要なら）第三者評価

HAIPレポート
OECD指針
モデルカード等

その他

AI Safety

AI適合性評価でサポートできる部分
出来ない部分

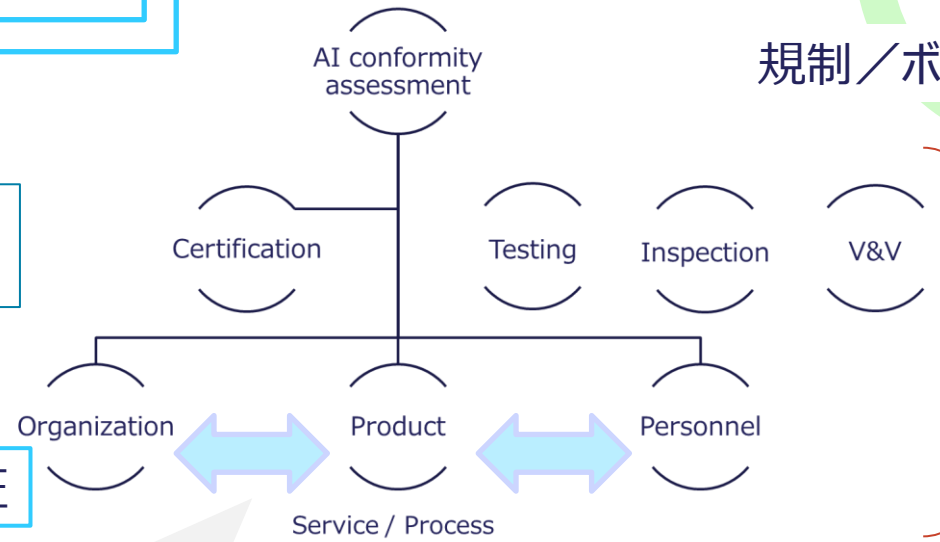
各国規制への対応

複合的かつ全体的に
段階をつけて
シルバー・ゴールド・プラチナ・ダイアモンド
相互承認・相互運用

Top down

規制／ボランタリー

Analysis &
Synthesis



どのような
組合せにするか

各種ガイドラインへの適合
AI事業者ガイドライン、AISIガイドライン類

42001認証

軽微なupdateはガバナンス側で
大きな変更はシステム側で評価

AI倫理
AI開発者・提供者トレーニング

Bottom up

AISI

Japan AI Safety Institute

More for Another Day

前後左右・上下を全ての方向をみつつ
N方良しとなるために