NIST サイバーセキュリティ白書

NIST CSWP 31

クリティカルな AI システムのための代理妥当性確認と検証

代理設計プロセス

フィリップ・ラプランテ
ジョアンナ・デフランコ
リック・クーン
ジェフ・ヴォース
ディビジョン
情報技術研究所
モハマド・カッサーブ
エンジニアリング部門
ペンシルバニア州立大学

本書は以下から無料で入手できる：
https://doi.org/10.6028/NIST.CSWP.31

2024 年 9 月 26 日

本書では、実験手順または概念を適切に説明するために、特定の事業体、機器、または材料を識別する場合がある。このような特定は、国立標準技術研究所（NIST）による推奨や推薦を意味するものではなく、また事業体、材料、装置が必ずしもその目的に利用可能な最良のものであることを意味するものでもない。

NIST テクニカルシリーズ方針

著作権、使用およびライセンスに関する声明

NIST 技術シリーズ出版識別子の構文

出版の歴史

2024-09-03 に NIST 編集審査委員会によって承認された。

この NIST テクニカルシリーズ出版物の引用方法：

Laplante P, DeFranco J, Kuhn R, Voas J, Kassab M (2023) クリティカルな AI システムのための代理妥当性確認と検証：A

代理設計プロセス。(National Institute of Standards and Technology, Gaithersburg, MD）, NIST Cybersecurity.

NIST CSWP 31. https://doi.org/10.6028/NIST.CSWP.31

認可 ORCID iDs

フィリップ・ラプランテ：0000-0002-0415-271X

ジョアンナ・デフランコ：0000-0001-8966-5532 リック・クーン 0000-0003-0050-1596 ジェフ・ヴォアス 0000-0003-1139-3690

モハマド・カッサーブ 0000-0002-3647-8511

連絡先

cswp-31-comments@nist.gov

国立標準技術研究所

担当：情報技術研究所コンピュータ・セキュリティ・ディビジョン 100 Bureau Drive (Mail Stop 8930) Gaithersburg, MD 20899-8930

追加情報

関連コンテンツ、更新の可能性、文書の履歴など、本書に関する追加情報は https://csrc.nist.gov/publications/cswp。

すべてのコメントは、情報公開法（FOIA）に基づき公開される。

要旨

このホワイトペーパーでは、重要な AI システム（CAIS）と類似性の高い代理システムを特定または構築し、一種の妥当性確認を代表者とし、各代理の使用ケースと誤用ケースの両方を作成し、その CAIS に対してテストすることによって代理を妥当性確認と検証することを含む、5 段階のプロセスについて説明する。

キーワード

人工知能、重要システム、重要 AI システム、妥当性確認と検証テスト。

目次

表のリスト

図表一覧

エグゼクティブサマリー

このホワイトペーパーでは、類似の人工知能（AI）システムから得られた先行テストの成果物を、新しい AI ソフトウェアに再利用できることを提案している。AI や機械学習ソフトウェアのテストは困難であり、類似システムの先行テスト結果を代理として適用することは、重要な研究の進歩になるだろう。

1. 序文

この研究の目的は、代理検証と妥当性確認を通じて、重要な AI システム（CAIS）の信頼を高めることである。CAIS では、あるテストケースを実行することが常に可能であるとは限らない。例えば、テストケースがテスト者や公衆を重大な危害にさらす可能性がある場合、運用プロファイルをアレンジすることが極端に困難または不可能な場合、あるいは、可能性が極端に低いシナリオに対してそのようなテストのコストが法外に高い場合などである。このような状況では、シナリオの信頼性を高めるような方法で、極端なケースをモデル化するために、非臨界等価物や代理システムを使用することが適切な場合がある[1]。

この必要性に対処するために、この研究は、CAIS に対して高い 類似性を持つ代理システムを特定または構築することを含む 5 段階のプロセスを記述する。このプロセスは、各 代理の CAIS に対する使用ケースと誤使用ケースの両方を 作成し、テストすることによって、一種の代理の妥当性確認と検証（V&V）を表す。この "類似した "システムから異なるシステムへの V&V の結果という考え方は斬新である。成功の鍵は、"類似性 "を実証し測定する能力である。

ある点で、このフレームワークは転移学習の問題に似ている。特定の環境に対してあるデータセットで学習されたモデルが、異なる環境で使用されたり、使用環境が変わったりした場合に使用される。代理 V&V と転移学習の顕著な違いは、代理 V&V の場合、モデルと環境の両方が異なる可能性があることである。どちらのフレームワークも、類似性の尺度の必要性を共有しており、そのような尺度は転移学習の研究対象である[2]。環境内の要素の例を含むデータセット間の類似点と相違点を、属性に割り当てられた値で定量化するために、転移学習による統計的尺度やその他の尺度を使用することができる。測定は、2 つ以上のクラス間の値の有無や属性値の違いの大きさなど、あるクラスやカテゴリの例が別のクラスの例と異なる程度を定量化するために使用できる。このような尺度は、異なるモデルとその使用環境の間の類似性を計算するために、代理 V&V 問題に適応させることができる。

1.1.背景

NIST 特別刊行物（SP）800-37r2（改訂 2）「情報システム及び組織のためのリスクマネジメントフレームワーク（*Risk Management Framework for Information Systems and Organizations*）」*[3]*は、信頼性の特性（セキュリティ、プライバシーなど）を統合するプロセスを記述している：セキュリティとプライバシーのためのシステムライフサイクルアプローチ[3]」は、信頼性特性（セキュリティ、プライバシーなど）を統合し、継続的なテスト、評価、検証、妥当性確認（TEVV）を重視し、AI システムのライフサイクル全体でサイバーサプライチェーンリス

クマネジメントを推進するプロセスを記述している。システム要件の妥当性確認とテストは、あらゆる開発ライフサイクルモデル、特に重要インフラシステムにとって重要な側面である。ここに記述されたプロセスは、リスクマネジメントフレームワークに合致する他の妥当性確認とテストのプロセスを支援し、補強することを意図している。

2. CAIS 妥当性確認・検証プロセス - 5 段階

図 1 の 5 段階のプロセスは、リスクを決定し（フェーズ 1）、 代理を特定し（フェーズ 2）、代理システムの類似性 を分析することで代理を検証し（フェーズ 3）、誤用ケースを作成してリスクを分類し（フェーズ 4）、誤用ケースをテストする（フェーズ 5）という検証プロセス[4]を示している。

フェーズ 1 と 2 は、[4]からの引用である。

図 1.CAIS の妥当性確認と検証プロセスの 5 つの段階

2.1.フェーズ 1： CAIS リスクアセスメント

米国のサイバーセキュリティ・インフラ・セキュリティ庁（CISA）は、破壊されると「安全保障、国家経済安全保障、国家公衆衛生もしくは安全、またはそれらの組み合わせに衰弱的な影響」を及ぼす 16 の重要インフラ部門を定義している[5]。したがって、16 のセクターに該当するシステムは、重要なシステムとみなされる可能性がある。

重要インフラ部門

1. 化学：基礎化学品、特殊化学品、農業化学品、消費者製品
2. 商業施設：娯楽・メディア、ゲーム、宿泊、野外イベント、集会、不動産、小売、スポーツリーグ
3. コミュニケーション：相互接続された地上、衛星、無線伝送システムを利用した音声サービスのプロバイダ。
4. 重要な製造事業者である：金属、機械、電気機器、器具、部品、輸送機器
5. ダム重要な貯水と制御サービス
6. 防衛産業基盤：米軍の要求を満たすための軍事兵器システム、サブシステム、部品・コンポーネントの研究、開発、生産、納入、保守。

7. 緊急サービス：日常業務およびインシデント対応において、予防、準備、対応、復旧サービスを提供する、高度に熟練し訓練された人員、物理的およびサイバーリソース。
8. エネルギー電気、石油、天然ガス
9. 金融サービス：預金取扱機関、投資商品のプロバイダ、保険会社、その他の信用・金融機関、重要な金融ユーティリティのプロバイダ、およびこれらの機能を支えるサービス。
10. 食品と農業：農場、レストラン、登録食品製造・加工・貯蔵施設
11. 政府施設：オフィスビル、軍事施設、国立研究所、裁判所
12. 医療と公衆衛生テロ、感染症の流行、自然災害からの防御
13. 情報技術：コンピューティングサービス、ネットワーク、データストレージ設備のプロバイダ
14. 原子炉、材料、廃棄物現役の発電用原子炉、研究・試験用原子炉、核燃料サイクル施設、その他医療診断や治療に使用される放射性物質源
15. 輸送システム：航空、高速道路、自動車輸送、海上輸送、大量輸送／旅客鉄道、パイプラインシステム、貨物鉄道、郵便、海運
16. 上下水道：井戸、貯水池、水処理施設、配水インフラ

これらの各分野は、AI 統合のレベルを反映したリスクカテゴリーを作成するために、それぞれの領域でシステムをさらに分類することができる。例えば、ヘルスケアシステムにおける AI の統合レベルは、支援型、拡張型、自律型と考えられる。[1]自律型ヘルスケアシステムは CAIS とみなされる。

システムが CAIS として分類されると、それと比喩的に等価なシステム （または代理）を特定しなければならない。代理の目的は、安全なテストを可能にするために、CAIS と機能的に同等であることである。例えば、自律走行車は、操作上および実装上の類似性が高ければ、ロボット掃除機をテスト代理とすることができる。CAIS の代理が完全にカバーされる可能性は低いが、これはプロキ シテストの価値を否定するものではない。代理の目的は、CAIS で直接テストできない機能をカバーすることである。何かが良い代理であるかどうかは、実装に大きく依存 するかもしれない。

[1] 詳細については、www.ama-assn.org を参照のこと。

代理システムはドメインの等価性を持つかもしれない(例 えば、CAIS と代理システムの両方が宇宙システムかもしれな い)が、ドメインの等価性は代理妥当性確認および検証の前提条件ではない。

代理試験結果の CAIS へのインピュテーションは、システム機能の適切なセットの選択に大きく依存する。機能の同等性は、フェーズ 2 で説明した分類法を用いた特徴抽出プロセスによって決定される。

2.2.フェーズ 2：代理同等品を見つけるためのシステム評価

CAIS の分類法の例は、[1]で提案されている。この分類法は、CAIS の特性 をテスト代理（すなわち、クリティカルでないプロトタイプやデジタ ルツイン）に適合させるために使用される。この分類法は、テスト代理 の機能的同等性をアセスメントする。図 2 が示すように、提案する CAIS 分類法には、次の 5 つの次元が含まれる：物理的運用環境、AI 適用目的、運用特性、人工知能/機械学習（AI/ML）技術、AI/ML 技術。

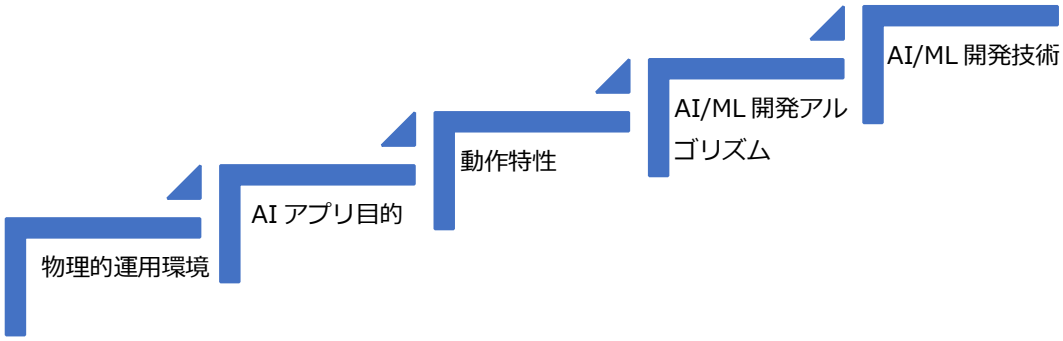

図 2.1]で提案された CAIS 分類法

2.2.1.物理的運用環境

物理的環境とは、自然環境（湖、海、森林など）と人間が作り出した環境（オフィス、工場、学校など）の両方を指し、人とシステムの両方の生活の質に影響を与える可能性がある。運用環境（OE）には、一般に、空、宇宙、地下地形（海洋、海洋学、水文学など）が含まれる。データが物理的な世界を移動することを考えると、サイバー空間も OE とみなされるべきである。

2.2.2.アプリ目的

アプリケーションの目的を決定することは、代理特性を特定するのに役立つ。一般に、AI アプリケーションは、ある特性に基づいて設計 され、構築される。この特性は、「X のための設計」または DfX と呼ばれることがあり、X は卓越性または品質要件 （テスト可能性、信頼性など）を表す。こ

のように設計することで，CAIS の最も重要な特性が代理の最終設計に反映されることが保証される。

システムの特性は、システムのドメインがコミュニケーション、学習、プランニング、推論、サービス提供の領域かどうかを判断するなど、そのドメインと目標を見直すことで分析できる。次に、言語処理、コンピュータビジョン、ディープラーニング、データサイエンス、機械学習など、全体的な AI の目標を特定することができる。この分析は、次の段階である運用特性の決定に役立つ。例えば、CAIS の目標が自律的に動作することである場合、代理も同じタイプの自律システムでなければならない。

特性の定義は一貫していなければならない。例えば、NIST 特別刊行物（SP）1011-I-2-0 において、国防総省は自律型車両を「主要構成要素に人間が搭乗しておらず、物理的世界で行動し、割り当てられたタスクを達成する」レベルと定義している。移動式でも固定式でもよい。操作される制御ユニット（OCU）のような、関連するすべての支援コンポーネントを含むことができる」[6]。また、無人地上車両（UGV）、無人航空機／システム（UAV／UAS）、無人海上車両（UMV）（例えば、無人水中車両［UUV］や無人水上車両［USV］）、無人弾薬（UM）、無人地上センサー（UGS）などの例も提示されている。ミサイル、ロケット、子弾、大砲は無人システムの主要な構成要素とは考えられていない[6]。別の例として、SAE J3016「路上走行自動車用運転自動化システムに関連する用語の分類法と定義」[7]は、自律走行車の 5 つの異なる自律性レベルを記述している。

自律走行車のタイプを定義した後、システムが完全自律か半自律かを決定する必要がある。半自律型とは、人間との相互作用の間で自律動作が可能な無人システムと定義される[8]。

2.2.3.動作特性

運用特性は、システムに対する潜在的な挙動と影響を表すものであり、それらを一致させることは、代理の精度にとって不可欠である。これらの特性を整理し、標準化するには、次のような多くの方法が考えられる：

1. O1.移動／静止 [no=0/yes=1]である。
2. O2.任務：航行、目標捕捉、目標攻撃、何かの収集、何か／ペイロードの配達（ガス、水、荷物など） [これらのうち 1 つ以上を含むことができる; $b\ b\ b\ b_{12345}$ 、 b ドメインが該当する場合、b_i =1]。
3. O3.経済的影響［0～9 の尺度で、0 は経済的影響なし、9 は破滅的な経済的影響を表す］。

4. O4.社会的影響［0～9 の尺度で、0 は社会的影響なし、9 は破滅的な社会的影響（プライバシー、選挙、コンプライアンス／法律など）を表す］。
5. O5.人的リスク [0～9 のスケールで、0 は人的リスクなし、9 は壊滅的な人的リスク（運転者、利用者、同乗者など）を表す]。

2.2.4.AI/ML 開発アルゴリズム

NIST の AI 用語集[9]は、AI を次のように定義している：

> ...学際的な分野で、通常はコンピュータサイエンスの一分野とみなされ、推論や学習など、一般的に人間の知能に関連する機能を実行するためのモデルやシステムを扱う。

同用語集では、ML を「データからモデルを決定するための一般的なアプローチ」と定義している[9]。

CAIS のアルゴリズムは、AI であれ、ML であれ、ディープラーニングであれ、アプリケーションによって異なり、代理 AI/ML アルゴリズムは、CAIS のアルゴリズムと学習タイプ（教師あり学習か教師なし学習か）を一致させる必要がある。アルゴリズムの例としては、ナイーブベイズ推定、線形回帰、主成分分析、決定木などがある。

代理を選択する際に考慮すべき重要な点は、ML アルゴリズ ムのためのトレーニングデータセットの利用可能性と同等性である。代理システムの ML アルゴリズ ムのテスト結果の信頼性は、そのデータセットが CAIS と等価であ るかどうかに依存する。場合によっては、この同等性を達成することが 不可能なこともある。

2.2.5.AI/ML 開発技術

CAIS のマッチング・代理を開発するために使用される技法も考慮する必要がある。開発に関する考慮事項には、使用するプログラミング言語（C++、Python など）、開発環境、ソフトウ ェア開発プロセスが含まれる。

2.2.5.1.提案された分類法の柔軟性

2.2.1 節から 2.2.5 節は、提案された CAIS 分類法の一般的な構造を示している。これは、テストのための代理システムを特定し、使用するための出発点であり、長期的な使用と交渉により、分類法は改良され改善される。異なるドメイン（例えば、航空宇宙、医療、発電と配電）は、特定の分類法と評価の次元をさらに洗練し、進化させるかもしれない。さらに、リッカート尺度の粒度は任意である。例えば、0～99 または別の尺度をどの要素にも使用することができる。

2.2.6.CAIS および代理分類法テンプレート

表 1 に示すテンプレートは、CAIS とその代理の特徴を決定するために用いることができる。表 1 は、操作に失敗した場合のリスクの結果を与えられた自律走行車両を用いて、CAIS 分類法を示している。目的は、ナビゲーション・システムの障害物回避アルゴリズムをテストすることである。

表 1.CAIS テンプレートの使用例

	物理的運用環境	AI アプリ目的	動作特性	開発アルゴリズム	開発技術
自律走行車	土地	推論、学習、計画、サービス	O1: 1; 02: 11111; O3: 0; O4: 9; 05: 9	KMP アルゴリズム	JAVA

表 2 は、CAIS 分類法を用いて分析した 2 つの代理シス テム（ロボット除草機とロボット掃除機）を示している。CAIS と代理システムの類似性の妥当性確認は、フェーズ 3 で行われる。

表 2.マッチング代理一の例

	物理的運用環境	AI アプリ目的	動作特性	開発アルゴリズム	開発技術
ロボット除草機	土地	推論、学習、計画、サービス	O1: 1; 02: 11111; O3: 0; O4: 0; 05: 9	KMP アルゴリズム	JAVA
ロボット掃除機	土地	推論、学習、計画、サービス	O1:1; 02: 11111; O3: 0; O4: 0; 05: 9	KMP アルゴリズム	JAVA

2.3.フェーズ 3：CAIS/代理類似性テスト

テストは CAIS 代理妥当性確認プロセスのフェーズ 3 とフェーズ 5 の両方で行われ、フェーズ 3 では類似性テストが、フェーズ 5 では誤用ケーステストが中心となる。このプロセスの詳細は[1]に記述されている。フェーズ 3 で類似性テストが成功すると、フェーズ 4 で誤用ケースが作成され、最終的にフェーズ 5 でテストされる。

例えば、フェーズ 2 では、自律走行車用の複数の代理が作成された。ロボット掃除機（レベル 1）□ ロボット除草機（レベル 2）□ ロボット芝刈り機（レベル 3）□ 自律走行車（レベル 4）-というように、それぞれの代理は自律走行車にとっての重要度と機能のレベルが上がっている、

- どれも似たようなナビゲーションシステムのアルゴリズムを使っている。
- どれも似たような障害物回避アルゴリズムを使っている。
- 各代理は、様々な重要度で複数の失敗ユースケースを持つことができる。

したがって、フェー ズ 3 では、マッチングプロセスをテストするために、 それぞれの代理の適切なユースケースシナリオ が、互いに対して、また CAIS に対してテストされる （図 3)。言い換えれば、フェーズ 2 のこれらの代理例を用い て、ロボット掃除機はロボット除草機に対してテストされ、次にフェーズ 2 で主張された次元を妥当性確認するために自律走行車に対してテストされる。

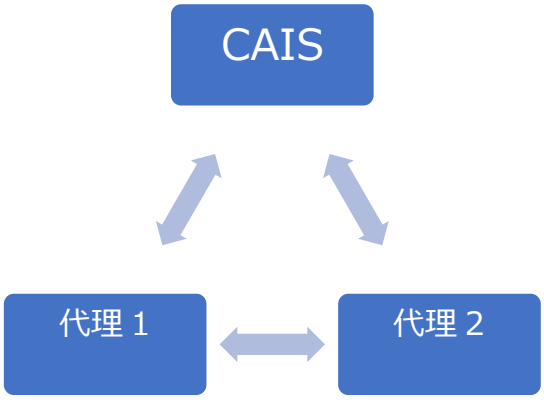

図 3.CAIS/Proxy の類似性テスト

2.4.フェーズ 4：さらなるテストのための誤用ケース

クリティカリティ分析を使用して、各代理について誤使用ケースを作成する。このプロセスは、内部報告書(IR)8179「臨界性分析プロセスモデル」に基づいている：システム及びコンポーネントの優先順位付け[8]に基づいている。CAP は情報資産のリスク分析と管理のためのものであるが、このモデルは、必須システム、サブシステム、コンポーネント、サブコンポーネント、およびそれらの動作環境を分析し、理解するためのアプローチを提供する。具体的には、以下の 2 つのステップにより、このアプローチを使用する：

1. 代理の誤用ケースを決定する：CAP プロセスを、代理の動作中に何がうまくいかない可能性が あるかを決定するために使用する。このステップでは、システムとそのコンポーネン

トのワークフロー、 依存関係、境界、相互作用、交差、接続、制約、およびトリガーを 分析する。

2. 誤用されたケースをリスクの高い順に分類する**：**

CAIS 1 □ 代理 1 □ 誤用ケース 1-N、各使用ケースのリスクレベルは増加する。

CAIS 1 □ proxy 2 □ 誤用ケース 1-N、各使用ケースはリスクレベルが増加している **例**（結果を表 3 に示す**：）**

ロボット除草機 - 自律走行車の代用品：

1. 誤用のケースを判断する：

a. ワークフローの経路、依存関係、境界を定義する。システムとそのコンポーネント（GPS、ML、故障の可能性のある他のセンサー、天候など）の相互作用、交差、接続、依存、制約、トリガーを識別する。

例を挙げよう：

依存関係：センサー、GPS、ML データセット

制約がある：天候

トリガー障害物を識別して避け、雑草を散布する。

b. 故障したセンサー、悪意のある事業体、ダウンタイム、動作速度の低下、障害物の誤認など、機能不全の状態（誤用ケース）を特定する。

質問する（結果は表 3 に示す）：

i. コンポーネントまたはサブコンポーネントが故障し、不利な動作状態になった場合、サブシステムが提供する機能/能力はどうなるのか？

ii. サブシステムの運用にどのような影響があるのか？

iii. サブシステムが稼働し続けるために最も重要な輸入事業者はどれか？

2. 誤用されたケースをリスクの高い順に分類する。

表 3.ロボット除草機の誤用ケースとクリティカリティ・レベル

誤用ケース名	誤用ケースのステップ	重要度レベル （低、中、高）
カメラセンサーの故障	GPS は学んだ道を庭までたどる。 センサーが雑草を誤認または回避する。 アクチュエーターが対象物、植物、人間、動物に毒を噴射する。	低 → 高、破損またはスプレーされたものによる。
失敗した *ML*	アルゴリズムはハッキングされている。 アクチュエーターは意図的に雑草を避けたり、意図的に物体にぶつかったりする。	低 → 高、スプレーするものによる。
GPS が故障した	GPS が故障し、除草ロボットが隣人に散布したり、庭の置物やフェンスにぶつかったりする。	低 → 高、破損またはスプレーされたものによる。

2.5.フェーズ 5：代理誤用ケースのテスト

フェーズ 5 で決定した誤用ケースのシナリオを、フェーズ 3 でテストした類似の代理ごとにテストする。

参考文献

[1] DeFranco JF, Kassab M, Laplante P (2022) A Taxonomy of Critical AI System Characteristics for Use in Proxy System Testing.*IEEE International Symposium on Software Reliability Engineering Workshops (ISSREW),* (IEEE Charlotte, NC) pp. 342-346. **https://doi.org/10.1109/ISSREW55968.2022.00090**.

[2] Weiss K, Khoshgoftaar TM, Wang D (2016) A survey of transfer learning.*Journal of Big Data* 3, Article 9, pp 1-40. https://doi.org/10.1186/s40537-016-0043-6

[3] Ross R (2018) Risk Management Framework for Information Systems and Organizations : セキュリティとプライバシーのためのシステムライフサイクルアプローチ。NIST 特別刊行物 (SP) NIST SP 800-37r2. https://doi.org/10.6028/NIST.SP.800-37r2.

[4] Laplante P, Kassab M, DeFranco J (2022) Proxy Verification and Validation for Critical Autonomous and AI Systems.*IEEE 29th Annual Software Technology Conference (STC),* (IEEE, Virtual) pp 37-40. https://doi.org/10.1109/STC55697.2022.00014

[5] 重要インフラ部門、サイバーセキュリティ・インフラセキュリティ庁、CISA、アーリントン、バージニア州。https://www.cisa.gov/critical-infrastructure-sectors。

[6] Huang H (ed.) (2008) Autonomy Levels for Unmanned Systems (ALFUS) Framework. (National Institute of Standards Technology, Gaithersburg, MD), NIST Special Publication (SP) NIST SP 011-I-2.0.(National Institute of Standards and Technology, Gaithersburg, MD), NIST 特別刊行物 (SP) NIST SP 011-I-2.0. https://doi.org/10.6028/NIST.SP.1011-I-2.0.

[7] Society of Automotive Engineers (2014) Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles.SAE J3106.www.sae.org で入手可能。

[8] Paulsen C, Boyens J, Bartol N, Winkler K (2018) 臨界分析プロセスモデル : システムとコンポーネントの優先順位付け。(National Institute of Standards and Technology, Gaithersburg, MD), NIST Interagency or Internal Report (IR) NIST IR 8179. https://doi.org/10.6028/NIST.IR.8179.

[9] 国立標準技術研究所（2023）The Language of Trustworthy AI: An In-Depth Glossary of Terms.以下から入手できる。
https://airc.nist.gov/AI_RMF_Knowledge_Base/Glossary

附属書 A.用語集

クリティカル AI システム（CAIS）

重要なソフトウェアを組み込んだシステムで、その故障が公衆に実質的な損害を与える可能性があるもの。[1]

妥当性確認 (validation)：システムまたはコンポーネントが、指定された要件を満たしているかどうかを判断するために、開発プロセス中または開発プロセスの終了時に評価するプロセス（INCOSE）。

検証 （verification）：ある開発フェーズの成果物が、そのフェーズの開始時に課された条件を満たしているかどうかを判断するために、システムまたはコンポーネントを評価するプロセス（INCOSE）。

V&V：妥当性確認と検証（検証と妥当性確認でもある）。

代理 V&V：クリティカルなシステムと同等の特性を持つ非クリティカルなシステムをテストの代用として使用すること。[1]