
セルフ SS における AI による給油許可監視の実装に向けた AI システム評価方法等に係るガイドライン

Draft 版

2022 年 3 月

目 次

1 ガイドラインの概要	4
1.1 目的及び効果	4
1.1.1 背景・目的	4
1.1.2 効果	4
1.2 構成	5
1.3 射程	5
1.4 他のガイドラインとの関係	8
1.5 用語	8
1.6 関連法令等	10
2 機械学習の信頼性評価	12
2.1 利用時品質及び外部品質における評価軸	12
2.2 利用時品質	13
2.3 外部品質	14
2.4 外部品質の要求レベル	15
2.5 内部品質	21
2.6 内部品質における評価軸	21
2.7 内部品質における評価軸の要求レベル	23
3 AI システムの試験方法	25
3.1 AI システムの試験方法におけるアプローチ	25
3.2 原理原則に基づく評価シナリオの体系	26
3.2.1 認識外乱シナリオ	28
3.2.2 顧客行動外乱シナリオ	33
3.2.3 監視スタッフ環境外乱シナリオ	40

4	ガイドラインの活用イメージ	41
4.1	ガイドライン活用の主体	41
4.2	ガイドライン活用の流れ	42
4.2.1	実施項目・実施内容の全体像	44
4.2.2	担当別・フェーズ別実施項目の詳細	46
5	附録	47
6	参考文献	48

1 ガイドラインの概要

1.1 目的及び効果

1.1.1 背景・目的

我が国において、顧客に自ら給油等を行わせる給油取扱所（以下、セルフ SS という）は、1998 年に行われた消防法並びに危険物の規制に関する政令及び同規則の改正により、その導入が実現したものである。

セルフ SS においては、引火性に代表されるガソリン・軽油等の危険性の観点から、危険物取扱者による給油行為の許可及び監視が必要とされている。

他方、近年国内の石油製品需要の減少を背景として、過疎化や、それに伴う人手不足等により、給油取扱所（以下、SS という）の数が減少し、過疎地域を中心にいわゆる SS 過疎地問題が課題となっている。2018 年 7 月に閣議決定された「エネルギー基本計画」等においては、AI・IoT 等の新たな技術を活用し、人手不足の克服、安全かつ効率的な事業運営や新たなサービスの創出可能とするため、安全確保を前提とした規制のあり方について検討することが求められている。

こうした動きを踏まえ、石油元売会社は、セルフ SS における AI を活用した給油許可監視に関する技術開発を進めており、2019～2020 年度に開催された総務省消防庁「過疎地域等における燃料供給インフラの維持に向けた安全対策のあり方に関する検討会」において AI を活用した給油許可監視は検討課題の一つとなり、給油許可監視業務における AI 活用のイメージが整理された。

その流れを受け、2021 年 4 月に消防庁より、石油業界側で AI システムの評価方法等について素案を作成し、2021 年度に開催する消防庁「危険物施設におけるスマート保安等に係る調査検討会」にて提案する方向で検討してほしいとの要請があり、石油連盟は、石油元売会社・AI システム開発関係会社・消防庁等をメンバーとする WG を立ち上げ、AI システムの機能要件・評価方法・試験方法に関するガイドライン案を作成する方向で検討を進めることとした。

「セルフ SS における AI による給油許可監視の実装に向けた AI システム評価方法等に係るガイドライン」（以下、本ガイドラインという）は、セルフ SS における給油許可監視システムの実装にあたり、機能要件や導入方法等、システムの信頼性評価に関する枠組みを示すことで、セルフ SS における円滑なシステムの導入・運用を支援し、以てセルフ SS における事故の発生防止による公共の安全確保と、セルフ SS 運営の効率化に寄与することを目的としている。

1.1.2 効果

本ガイドラインの活用主体としては、給油許可監視システムをセルフ SS に導入する企業（例えば、先行して開発を進める石油元売会社）と、当該システムを開発し納品するベンダー企業を想定しており、それ

それぞれの立場において想定される活用効果を以下に示す。

(1) 給油許可監視システムを導入する SS 運営者の活用効果

給油許可監視システムを導入する SS 運営会社は、本ガイドラインの活用により、セルフ SS における給油許可監視システムの信頼性評価を適切に実施することができる。また、本ガイドラインに則って開発されたシステムの導入及び運用を行うことで、安全性等に関する説明責任を果たしやすくなり、セルフ SS におけるシステムの導入及び運用を円滑化することができる。

(2) ベンダー企業の活用効果

ベンダー企業は、本ガイドラインの活用により、給油許可監視システムを導入する SS 運営会社が同システムに求める信頼性の基準を客観的に把握した上で、システムの開発に取り組むことができる。これにより、一般的に困難とされるシステムの信頼性評価の妥当性を説明することが容易となり、より円滑なシステム開発を行えるようになることが期待される。

1.2 構成

本ガイドラインは、次に示す第 1 章から第 4 章までの 4 章構成である。

第1章 「ガイドラインの概要」では、本ガイドラインの目的及び効果、構成、射程、他のガイドラインとの関係、用語や関連法令等を示す。

第2章 「機械学習の信頼性評価」では、給油許可監視システムの信頼性評価に用いる概念・品質の階層構造（「利用時品質」「外部品質」「内部品質」の 3 階層）の内容について整理する。

第3章 「AI システムの試験方法」では、試験方法における給油許可監視システムのアプローチと、評価シナリオについて体系的に整理する。

第4章 「ガイドラインの活用」では、本ガイドラインの活用主体を整理し、具体的な活用の流れや実施項目等を示す。また、本ガイドライン活用に関する FAQ を章末に設けている。

1.3 射程

本ガイドラインは、給油許可監視業務を支援するシステムをセルフ SS へ導入する場合において、当該システムの信頼性を担保するための基準等をまとめたものである。

本ガイドラインにおける「給油許可監視業務」は、消防庁が 2020 年度に開催した過疎地域等検討会において、AI 活用の対象と整理された業務領域、即ち、「火気・携行缶・ポリタンクの検知・監視」を含む、「固定給油設備からノズルが離れてから戻るまでの給油許可・監視」を射程とする（図 1-1）。ただし、例えば灯油の注油行為等、将来的に業務領域を広げていく可能性を限定するものではない。

また、本ガイドラインにおいては、次頁の図 1-2 の通りロードマップを作成しており、給油許可監視システムの社会実装を 3 つの STEP で段階を踏んで実現していくこととした。

本ガイドラインにおける給油許可監視システムは、同ロードマップにおける STEP1～STEP1.5 を対象とし、セルフ SS における給油許可監視業務の「支援」を想定している。今後 STEP2 を実現する段階では、緊急時の有人対応を前提として「監視業務」を AI が代替することは将来的な射程に含めることとするが、その場合においても、非常時の安全確認や緊急時対応等は、引き続き有人での対応となることを前提としている。

給油許可における監視項目	No.	監視内容	その他	No.	監視内容
	1	自動車が給油エリアに停止		1	火気の有無
	2	エンジンを停止させる		2	ポリ缶・携行缶の有無
	3	給油口の蓋を持ち上げる		3	注油（灯油の給油）
	4	自動車から降りる			
	5	パネルを操作し、注文・支払いをする			
	6	静電気除去パッドに触る			
	7	油種に応じたノズルを取る			
	8	給油口にノズルを差し込む			
	9	給油			
	10	給油口からノズルを抜き取る			
	11	ノズルを戻す			
	12	給油キャップを閉め給油口の蓋を閉じる			
	13	乗車する			
	14	枠内から退場する			

【凡例】

本ガイドラインの対象とする監視項目

対象外の監視項目

図 1-1 本ガイドラインが対象とする給油許可監視システムの射程

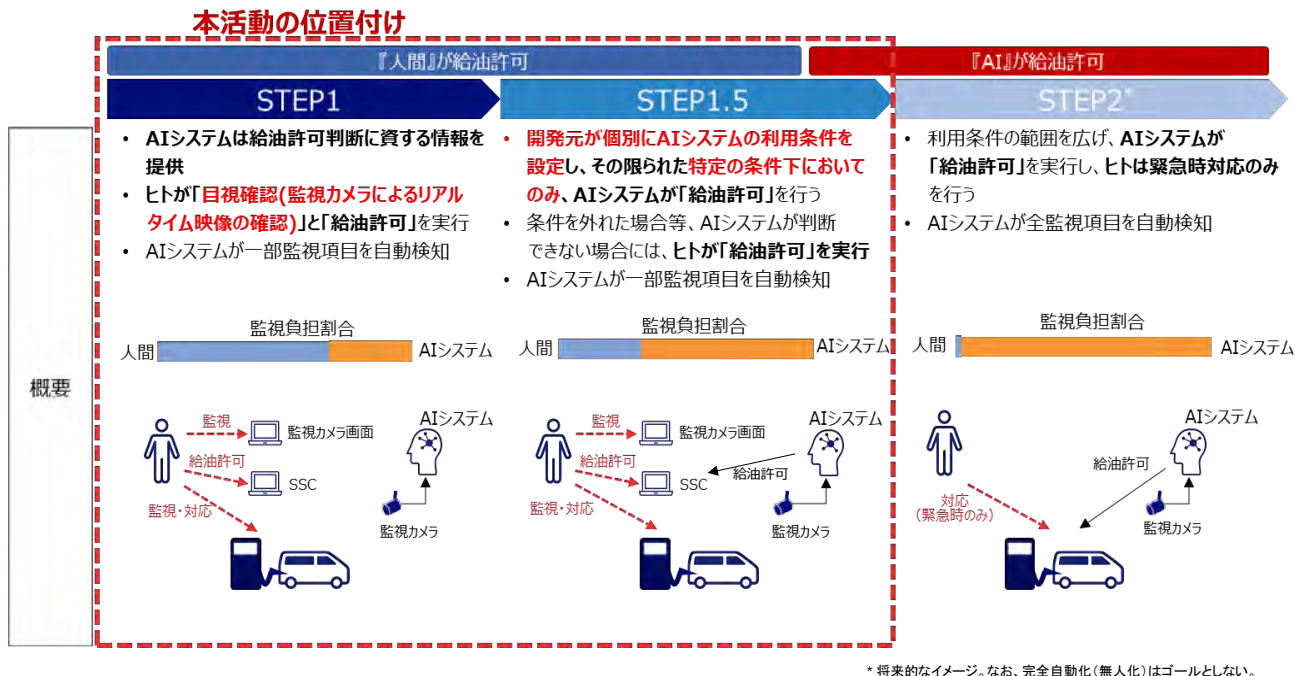


図 1-2 本ガイドラインが対象とする給油許可監視システムの実装に向けたロードマップ

出所) NRI 作成

また、本ガイドラインにおいては、AI 応用分野で先行する自動運転車両の事例を参照している。図 1-3 に示す通り、自動運転の技術の分野では、道路環境や気象条件等が異なる中、安全な走行を行うための技術水準に至っていないシステムにおいては、運転主体を人とシステムのいずれかに設定し、システムが自動運転を行う場合の条件として運行設計領域（Operational Design Domain、以下、ODD という）を定め、定められた ODD を超えた場合には、システムが自動運転できないために人に判断を委ねる仕組み、人への交替要求（Turnover Request、以下、TOR という）を確実に行うことが必須要件となる。

この事例に則り、システムを利用できる環境条件や利用条件について、まずは具体的な評価ケースとして、先行して開発する石油元売会社のシステムごとに個別に条件の範囲を設定し、その限定された条件下でのみ給油許可監視システムの給油許可判断（計算結果）をそのまま適用できることを目指すこととした。また、その条件下でシステムが判断（計算）できない場合には、確実に人に判断を委ねる形を想定している。

レベル	名称	定義	運転主体	走行領域	ODD	TOR
L0	運転自動化なし	ADAS非搭載の旧来車で、ドライバーが全ての動的タスクを担う。	人	—	—	—
L1	運転支援	クルーズコントロール(前走車追随)、レーンキープコントロール(車線内走行維持)のいずれかによりアシストする機能を有する車で、縦横の加速減・操舵のいずれかの車両運動制御のサブタスクを限定的に実施する。	人	限定あり	—	—
L2	部分運転自動化	上記の加速減・操舵の縦横両方をアシストする機能を有する車により、 ハンズオフ運転が可能に 。	人	限定あり	—	—
L3	条件付運転自動化	自動運転車が走行可能な各種条件を表す「ODD（運行設計領域）」を設定し、その限定条件下のみ、システムがすべての動的運転タスクを実施する。これにより、ドライバーは アイズオフ運転が可能に 。ただし、特定条件を外れ、作動継続が困難な場合は、AIシステムがドライバーに介入要求(TOR)を実施し、ドライバーは迅速に応答しなければならない。	AI	限定あり	○	○
L4	高度運転自動化	上記同様、ODDの設定により限定領域にてAIシステムがすべての動的運転タスクを担う。また、作動継続が困難な場合への応答も実行する（TOR不可）。	AI	限定あり	○	×
L5	完全運転自動化	ODDの設定なしに、AIシステムがすべての動的運転タスクを担う。 場所や条件などの制限がない完全自動運転 。	AI	限定なし	×	×

図 1-3 自動運転におけるレベルの定義（参照事例）

出所）国土交通省資料、官民 ITS 構想・ロードマップ等を基に NRI 作成

本ロードマップにおいては、足元の技術的な制約を AI が実現できる状態を **STEP1**、AI が自立して給油許

可と給油監視までを実施できる状態を **STEP2** とし、その中間で AI が自立してできる条件を少しずつ増やしていく状態を **STEP1.5** とした（図 1-3）。なお、本ガイドラインは、主に **STEP1** から **STEP1.5** を対象としている。

STEP1 では、まずは安心安全を第一優先として AI で判断（計算）した内容を人間が確認することで安全を担保することを目指す。セルフ SS においては、計量機を動かすために制御装置（SSC）を利用するが、本 STEP では従来の固定式の制御装置に加え、2019 年度に認められた可搬式の制御装置についても利用を想定している。AI が判断（計算）した安全・危険の判定結果を人間のスタッフに通知し、スタッフが監視カメラによるリアルタイム映像を目視確認して、スタッフが制御装置の許可ボタンで給油許可することを想定している。

STEP1.5 は、人間の確認作業を少なくするため、特定の条件下で AI が自立して給油許可及び給油監視までを実施できることを目指す。

1.4 他のガイドラインとの関係

本ガイドラインの策定にあたっては、産業技術総合研究所の「機械学習品質マネジメントガイドライン第 2 版」及び石油コンビナート等災害防止 3 省連絡会議¹の「プラント保安分野 AI 信頼性評価ガイドライン」（以下、プラント AI ガイドラインという）を参考としている。

(1) 機械学習品質マネジメントガイドライン 第 2 版（産業技術総合研究所）

産業技術総合研究所が、民間企業・大学などの有識者と共同で、「AI システムのライフサイクル全体にわたる品質マネジメントを扱い、AI システムのサービス提供で求められる品質要求を充足するための必要な取り組みや検査事項を体系的にまとめたもの」である。AI システムに求められる品質を客観的に判断する基準を策定することで、AI システムのビジネス活用を加速させることが期待されている。

(2) プラント保安分野 AI 信頼性評価ガイドライン（石油コンビナート等災害防止 3 省連絡会議）

「プラント保安分野に特化して AI の信頼性（プラントの安全性や生産性向上のために期待される品質を果たすこと）を適切に管理する方法」を示したものである。このガイドラインは、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 第 2 版」と同様の方法で機械学習の信頼性評価を行い、必要な信頼性を確保する具体的な方法等を示している。

1.5 用語

本ガイドラインにおいて使用する用語について、以下のように定義する。

(1) 機械学習

¹ 3 省とは、経済産業省・総務省消防庁・厚生労働省を指す。

機械学習は、「人工知能のうち、特に計算機システムが明示的なプログラム指示に依ること無く、データのパターンを自動的に認識した結果を用いて推論・判断を行うようなシステムのこと」を指す。

(2) SS (Service Station／サービスステーション／給油取扱所)

サービスステーション(SS)は、危険物の規制に関する政令第 3 条第 1 号に規定される、給油設備によって自動車等の燃料タンクに直接給油するため危険物を取り扱う取扱所（給油取扱所）のことを指す。頭文字を取って「SS」と記載される。

(3) セルフ SS

セルフ SS は、危険物の規制に関する規則第 28 条の 2 の 4 に規定される、顧客が自ら自動車等に給油を行うことができる給油取扱所のことを指す。自動二輪車は自動車に含まれるが、顧客にガソリンを容器に詰め替えさせることは認められない。（平成 10 年 3 月消防危 25 号）

(4) EUC (Equipment Under Control：被制御装置)

EUC は、「製造、プロセス、運輸、医療、その他の業務に供される機器、機械類、装置、プラントなど」を指し、本ガイドラインでは、主にセルフ SS における給油許可業務に用いられる機器等を指す。

(5) E/E/PE (Electric/Electronic/Programmable Electronic：電気・電子・プログラマブル電子の)

E/E/PE は、「電気(E)、電子(E)及び／又はプログラマブル電子(PE)技術に基づく」ことを指す。この用語は、電氣的原理で作動する全ての機器又はシステムを包括する。

(6) 安全関連系

安全関連系は、以下の 2 つの条件を満たすシステムのことを指す²。

- ① EUC を安全な状態に移行させるため、又は EUC の安全な状態を維持するために必要な、要求された安全機能を行う。
- ② それ自体で、又はその他の E/E/PE 安全関連系及び他リスク軽減措置によって、要求される安全機能に対して必要な安全度を達成する。

(7) SIL (Safely Integrity Level：安全度水準)

SIL は、機能安全規格 IEC 61508 (JIS C 0508)³及び IEC 61511 (JIS C 0511)⁴で定められている「安全度の値の範囲に対応する離散的水準」を指す。安全度水準の最高値は 4 であり、最低値は 1 である。なお、安全度とは、「ある E/E/PE 安全関連系が、指定した期間内に、全ての指定し

² JIS C 0508-4：2012 電気・電子・プログラマブル電子安全関連系の機能安全－第 4 部：用語の定義及び略語

³ IEC (国際電気標準会議) が制定した E/E/PE の機能安全に関する国際規格

⁴ IEC が制定したプロセスエンジニアリングのアプリケーション規格

た条件下で、規定する安全機能を果たす確率」を指す。

(8) 信頼性（ソフトウェア信頼性）

信頼性は、「特定条件下で使用する時のある性能を維持する能力」を指す。

本ガイドラインにおいて、単に「信頼性」という場合は「ソフトウェア信頼性」を指すものとする。

(9) PoC（Proof of Concept）

PoC は概念実証のことを指し、「新しいアイデアの実施可能性を確認するために、本格的な実装・システム化の前に行う検証活動」を指す⁵。

(10) 利用時品質

利用時品質は、システム全体で最終的な利用者に提供すべき品質のことを指す⁶。実際にユーザー利用した際の満足性やリスク回避性等が該当する。

(11) 外部品質

外部品質は、システムのうち機械学習で構築された構成要素が満たすことが期待される品質のことを指す⁶。外部品質を達成することで、利用時品質が実現される。

(12) 内部品質

内部品質は、機械学習による構成要素が固有に持つ品質のことを指す⁵。「外部品質」を満たすために、機械学習要素の設計・開発・運用等で充足することが必要になる品質である。

1.6 関連法令等

本ガイドラインに関連する法令・規格・基準等は、以下に示す通りである。

(1) 消防法

「火災を予防し、警戒しおよび鎮圧し、国民の生命、身体および財産を火災から保護するとともに、火災または地震等の災害に因る被害を軽減し、もって安寧秩序を保持し、社会公共の福祉の増進に資すること」（第 1 条）を目的として定められた法律である。

(2) 消防法施行令

防火対象物、消防用設備に関する技術基準、救急業務、消防設備に関する検査等を定めた、消

⁵ 石油コンビナート等災害防止 3 省連絡会議「プラント保安分野 AI 信頼性評価ガイドライン」

⁶ 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 第 2 版」

防法の施行のための政令。

(3) 消防法施行規則

消防法施行に必要な防火・消防管理者、消防計画等の届出、消防用設備等の設置、維持の技術上の基準、検査、点検等を詳細に定めた省令。

(4) 危険物の規制に関する政令／規則

消防法第三章及び危険物の規制に関する政令の規定に基づき、これらの実施のために定められる政令及び規則。危険物の品名や指定数量、貯蔵所や取扱所の区分等を定める。

(5) 消防危

消防危とは、消防庁危険物保安室が通知・通達したものを指す。「顧客に自ら給油等をさせる給油取扱所における可搬式の制御機器の使用に係る運用について」（令和 2 年 3 月 27 日付け消防危第 87 号）等が本ガイドラインに関連する。

2 機械学習の信頼性評価の構造

本ガイドラインでは、石油コンビナート等災害防止 3 省連絡会議「プラント保安分野 AI 信頼性評価ガイドライン 第 2 版」及び同ガイドラインが参照している、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 第 2 版」における品質確保の構造を用いて、AI システムの信頼性評価を行う。

品質確保の階層構造について、その概略を図 2-1 に示す。まず、AI システム全体を通して実現したい内容（＝利用時品質）を定義し、それを満たすために必要な機械学習要素の性能と達成レベル（＝外部品質）を設定した。そして、その達成レベルに応じた要求事項に基づいて、機械学習要素（＝内部品質）を作り込むことを想定している。本ガイドラインは、給油許可監視システムの品質確保に向けた各品質について、上記の流れでその定義付けを実施した。

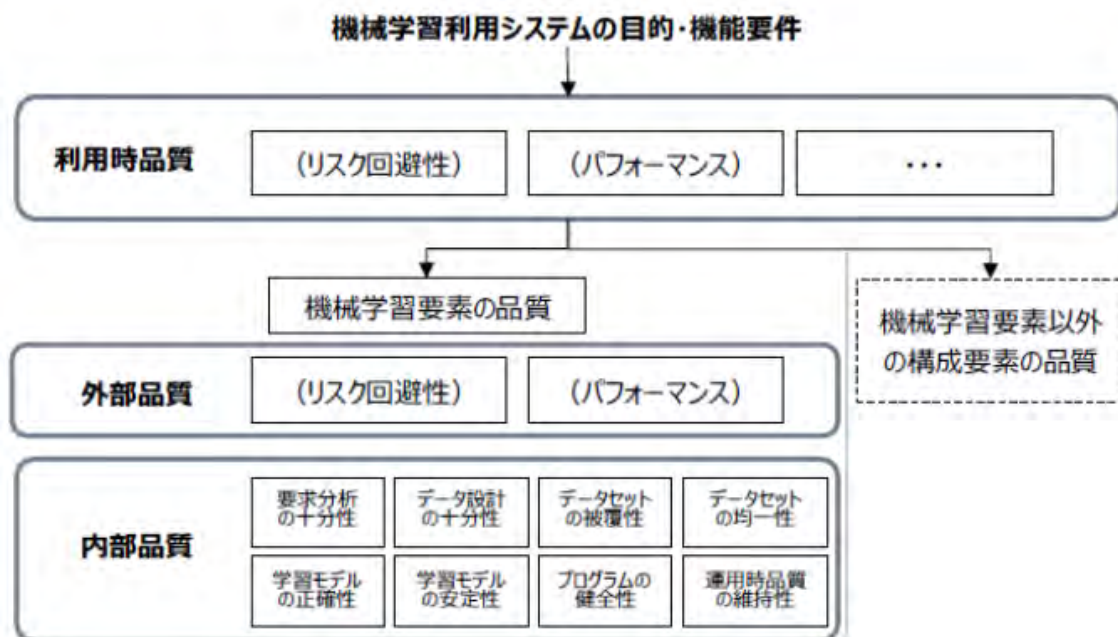


図 2-1 機械学習利用システムの信頼性評価における 3 段階の品質の関係

出所）「プラント保安分野 AI 信頼性評価ガイドライン 第 2 版」（2021 年 3 月）

2.1 利用時品質及び外部品質における評価軸

本ガイドラインでは、利用時品質・外部品質における各品質分類を「軸」と称し、「リスク回避性」と「パフォーマンス」の 2 軸を設定している。また、外部品質については、「リスク回避性」と「パフォーマンス」のそれぞれについて、要求事項の強弱に応じた「レベル」を定義しており、外部品質のレベルに応じて、第 2.5 節の「内部品質」における要求事項のレベルが決定する構造となっている。

なお、「リスク回避性」と「パフォーマンス」の双方が同一のシステム上で要求されることもあり得る。このような場合は、「リスク回避性」と「パフォーマンス」の最適なバランスを定めることが必要である。

(1) リスク回避性

「リスク回避性」とは、「機械学習要素が望ましくない判断動作を行うことを抑制し、システムを用いたサービス提供者・システムにより提供されるサービスの利用者または第三者などに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品質特性」を指す。

(2) パフォーマンス

「パフォーマンス」とは、システムの有用性を意味し、業務等の生産性・効率性の向上に資する品質特性を指す。具体的には、システムの利用者が期待する成果を高い精度・確率で出力する品質を指し、「個別の誤判断による悪影響が大きな問題にならない出力では、個々の出力の是非よりも、平均的な性能の高さが要求されることに着目している⁷⁾」。

2.2 利用時品質

利用時品質とは、「システム全体で最終的な利用者に提供すべき品質」を指す。利用時品質には、システムの利用者がシステムに期待している目的や、前提として確保されているべき安全性等が含まれており、換言すると、「『システムで実現したいこと』が利用時品質である」と言える。本ガイドラインにおいては、正常な給油動作の検知と異常な給油動作の検知の組み合わせにより、安全性の向上と省力化・効率化の向上を狙ったものとした。

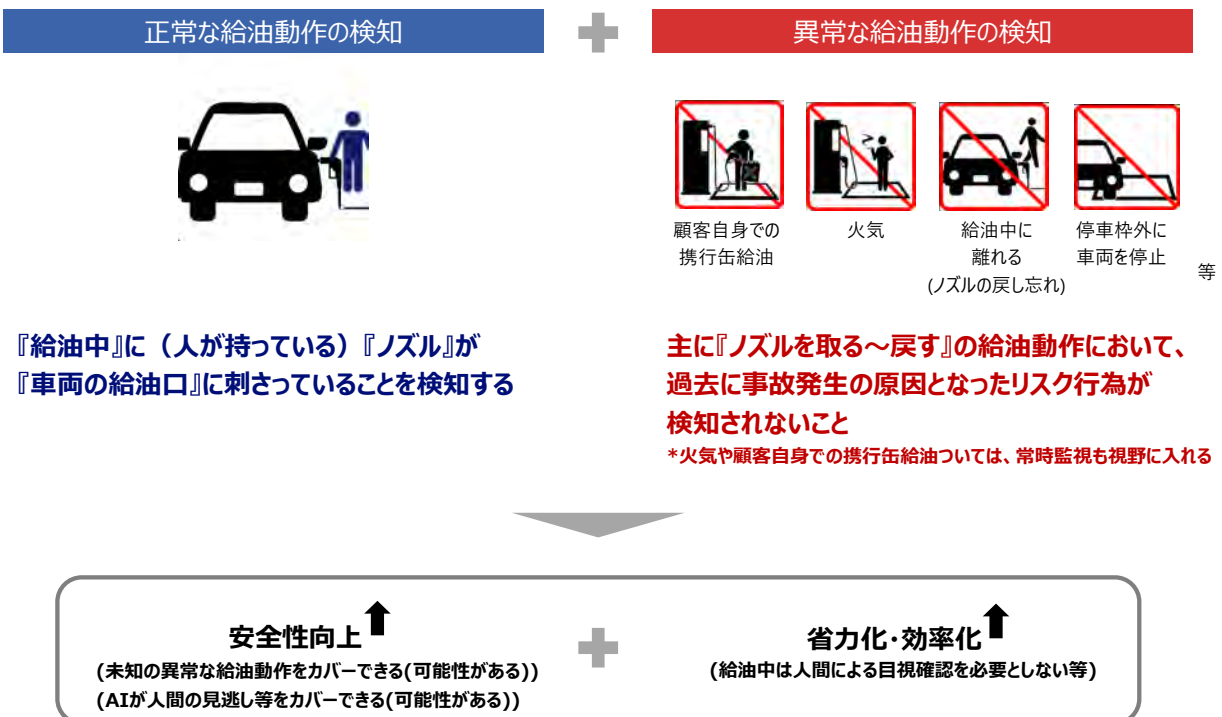


図 2-2 給油許可監視システムの利用時品質：リスク回避性のアプローチ

出所) NRI 作成

⁷⁾ 石油コンビナート等災害防止 3 省連絡会議「プラント保安分野 AI 信頼性評価ガイドライン」

具体的には、リスク回避性の軸として、「『正常な』給油動作が正しく検知されること」、「（SS スタッフの対応が必要な）『異常な』給油動作が正しく検知されること」の 2 つを利用時品質と定義した。この 2 点を満たすことにより、安全性の向上を目指したシステムとする。また、省力化・効率化を目指したパフォーマンスの軸として、「SS スタッフの（不要な）対応実施数を少なく抑えること」を 3 つめの利用時品質として定義した。本ガイドラインでは、AI システムがこれらの 3 つの利用時品質を満たすことを推奨する。

2.3 外部品質

前節で設定した利用時品質に対し、対応した機械学習要素の外部品質を設定する。外部品質とは、「システムのうち機械学習で構築された構成要素が満たすことが期待される品質」を指す。システムの構成要素が「部品として要求される、客観的な視点の品質」のことであり、セキュリティ、信頼性、一貫性などが挙げられる。

給油許可監視業務を支援するシステムの機械学習要素に求められる性能（ノズル検知能力等）が、外部品質に該当する。本ガイドラインにおいては、リスク回避性軸の外部品質として、「誤判定率の低減」を定義した。具体的には、正しい給油許可判断が「異常」であるべき給油動作の場合に AI システムが「正常」であると判定する誤判定率、つまり、異常な行為を見逃してしまう間違いを限りなく小さくすることを指す。また、パフォーマンス軸の外部品質は、「誤検知率の低減」と定義した。具体的には、正しい給油許可判断が「正常」であるべき給油動作の場合に、AI システムが「異常」と判定する誤検知率、つまり、SS スタッフの対応が不要な場合に、要と判定する誤検知率を限りなく小さくすることを指す。リスク回避性軸の基準となる誤判定率と、パフォーマンス軸の基準となる誤検知率の意味合いは次の図 2-3 に示した通りである。

	正しい 給油許可判断	AIシステムの 給油許可判断	正解・ 不正解	
①OK-OK パターン	OK	OK	正解	
②OK-NG パターン	OK	NG	誤り	→ 誤検知率（パフォーマンス軸・効率性） =OK-NG件数/全件
③NG-NG パターン	NG	NG	正解	
④NG-OK パターン	NG	OK	重大な 誤り	→ 誤判定率（リスク回避性軸・安全性） =NG-OK件数/全件

図 2-3 誤判定率・誤検知率の意味合い

出所) NRI 作成

2.4 外部品質の要求レベル

前節で設定した外部品質について、機械学習要素への要求に応じたレベルを設定する。外部品質の要求レベルを評価するには、リスク回避性の軸である AI 安全レベル（AI Safety Level、以下、AISL という）と AI パフォーマンスレベル（AI Performance Level、以下、AIPL という）の 2 つの観点で評価することが求められる。本ガイドラインでは AI プラントガイドラインの評価方法に則り、レベル設定を行った。

2.4.1 リスク回避性：AISL の設定

「リスク回避性」軸の外部品質に対し、機械学習要素の誤判断による悪影響を回避することへの要求の強さに応じてレベルを設定する。リスク回避性のレベルである AISL の設定フローは次に示す図 2-4 の通りである。

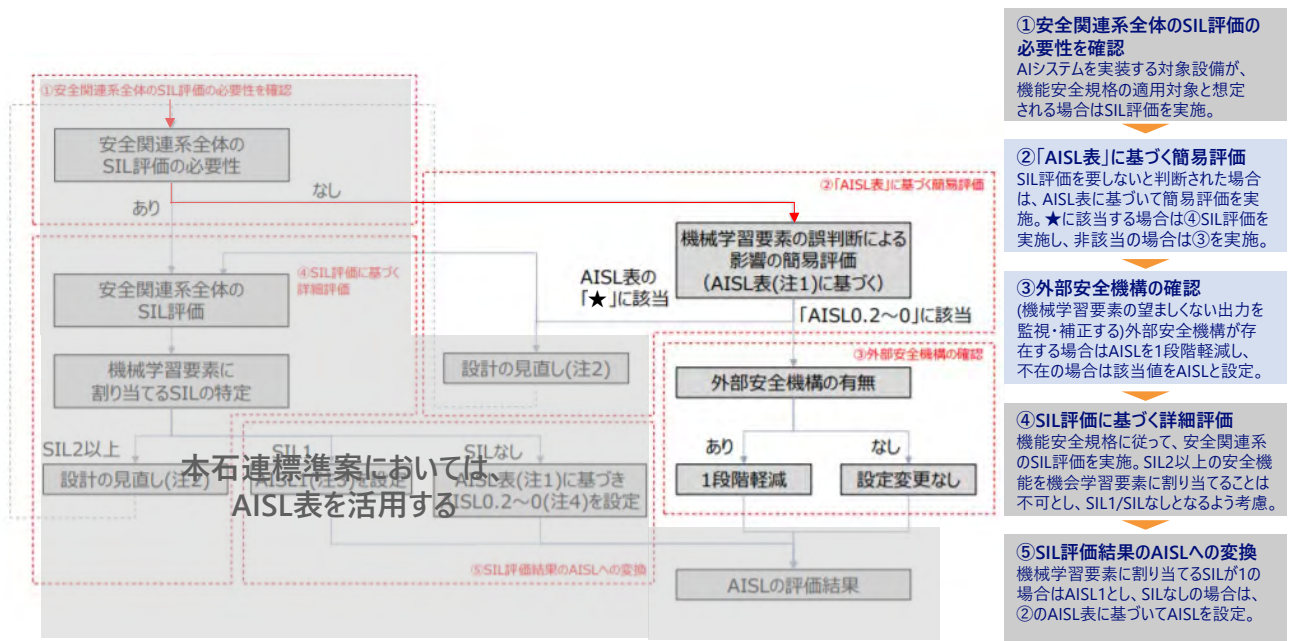


図 2-4 リスク回避性のレベル（AISL）の設定フロー

出所）プラント保安分野 AI 信頼性評価ガイドライン第 2 版より NRI 作成

給油許可監視システムの AISL 設定においては、プラント AI ガイドラインの記載内容のうち、「機械学習要素の誤判断による影響の簡易評価（AISL 表）」、「外部安全機構の有無」の 2 段階で評価を行った。

まず、機械学習要素の誤判断による影響の簡易評価においては、次の AISL 表を用いて評価を実施した（図 2-5）。AISL 表は、縦軸に I～IV までの強度基準の評価、横軸に①～③までの人間による回避可能性の分類を行い、その掛け合わせでレベルを判断する。詳細は次項にて後述する。

強度基準は、1) ～ 3) のうち、最も大きいものを選択する。

※「微小な怪我」とは、いわゆる「赤チン災害」以下の強度に当たり、医師の診察を要する不休災害などは強度基準 II に該当する。
※「直接被害額」とは、以下を意味する。「修理・交換・清掃・廃棄処分・環境修復および緊急対応の費用。直接コストには、ビジネス上の機会損失・事業の中断および 原料や製品の損失・装置の停止による逸失利益、一時的な設備の調達や運転費、顧客の要請に対応した代替製品の調達費のような間接コストは含まない。」

強度基準	1) 人的被害	2) 経済的被害 (直接被害額)	3) 経済的被害 (間接被害額を含む) ※適用は任意	人間による回避可能性の分類		
				①人間による代替システムが 用意されておらず、機械学習 要素の判断結果がそのまま 運転・保身に反映される	②機械学習要素の判断結果 が運転・保身にそのまま反映 されることなく、人間による 確認・代替システムの適用を 介した結果が反映される	③機械学習要素は補助的 情報のみを提示し、人間が 判断して運転・保身に反映 する
I	・ 死亡 ・ 障碍の残る障害 ・ 多数の重傷 ・ 極めて多数の重傷	1億円以上	・ 企業体としての 存続等に著しい 影響 ・ 業務の運営を揺る がす重大な損害	※AISL1	※AISL 0.2	※AISL 0.2
II	・ 重軽症 ・ 多数の負傷	1千万円以上	無視できない、具体的 な損害	※AISL1	※AISL 0.2	AISL 0.2
III	微小な怪我	1千万円未満	軽微な利益の逸失 にとどまる	※AISL1	AISL 0.2	AISL 0.1
III'	(想定される被害者により容易に回避できる場合)			AISL 0.2	AISL 0.2	AISL 0.1
IV	傷害の想定無し	直接被害額の 想定は軽微	間接被害額を含む 経済的被害の 想定は軽微	AISL 0	AISL 0 ロードマップ STEP 1.5	AISL 0 ロードマップ STEP 1

ロードマップSTEP 2

図 2-5 「リスク回避性」の簡易評価の基準 (AISL 表)

出所) プラント保安分野 AI 信頼性評価ガイドライン第 2 版より NRI 作成

(1) AISL 表の縦軸評価 (強度基準)

「AISL 表」の縦軸は、機械学習要素が誤判断した場合に発生する人的又は経済的な影響の強度に応じて AISL が変動することを示している。したがって、強度が大きい表の上の方が求められる AISL が高く、下の方が低く設定されている。強度基準は、「人的被害」「経済的被害 (直接被害額のみ)」「経済的被害 (間接被害額を含む)」のうち最も大きいものを選択することで設定するが、「経済的被害 (間接被害額を含む)」の適用は任意となっている。

上記の判断基準を踏まえ、本ガイドラインの対象となる給油許可監視システムにおける強度基準は、過去 5 年間の消防庁提供事故データを分析した結果、「強度基準 II」とした。主に給油許可の業務フローに起因する事故の人的被害の程度が II 以下の内容しか見られず、かつ II に該当する事故の発生頻度も非常に稀であったためである。具体的には、給油動作起因の死亡・重傷発生事故は 0 件であったが、中等傷者発生事故が 2 件あった (図 2-6)。この結果を既存の基準・分類と照らし合わせると次の図 2-7 の通りとなる。

(2) AISL 表の横軸評価 (人間による回避可能性の分類)

「AISL 表」の横軸は、機械学習要素の誤判断を人間が回避する可能性の程度に応じて AISL が変動することを示している。①は、機械学習要素の判断がそのまま機械学習利用システム全体の最終判断となる場合が該当する。②は、機械学習要素は何らかの判断を行うが、その判断は必ず人間が確認し、最終判断となる場合が該当する。③は、機械学習要素自体は判断を行わず (判断の材料となる情報のみを出力し)、機械学習要素の出力を必ず人間が見て判断を行う場合に該当する。

人的被害発生事故件数（平成28年度～令和 2 年度）

	人的被害発生事故件数（単位：件）			
	死亡者 発生事故	重傷者 発生事故	中等傷者 発生事故	軽傷者 発生事故
令和 2 年度	0	2	2	7
令和元年度	0	0	1	3
平成30年度	0	1	5	5
平成29年度	0	0	0	5
平成28年度	0	1	0	6
5か年計：	0	4	8	26

損害額別事故件数（平成28年度～令和 2 年度）

	損害額別事故件数（単位：件）			
	1万円以上	1万円未満	調査中	年度内 事故件数計
令和 2 年度	77	34	3	114
令和元年度	80	33	2	115
平成30年度	79	32	0	111
平成29年度	53	18	0	71
平成28年度	52	25	3	80
5か年計：	341	142	8	491

発生事故の詳細

事故種別	年度	事故概要
重傷者発生事故	令和 2 年度	屋外セルフ給油取扱所内で顧客が自身にガソリンをかけ火をつけて発生した火災事故
重傷者発生事故	令和 2 年度	給油取扱所敷地内で軽乗用車と人の関係する交通事故が発生し、固定注油設備が一部破損したもの。
重傷者発生事故	平成30年度	ローリーからの荷卸し中における可燃性蒸気滞留による爆発火災
重傷者発生事故	平成28年度	給油取扱所において、固定給油設備の点検中にインバクトドライバーの電気火花がガソリンペーパーに引火したことによる火災
中等傷者発生事故	令和 2 年度	給油取扱所において利用客が運転操作を誤り軽自動車敷地内の事務所開口部に衝突させ、破損させたもの
中等傷者発生事故	令和 2 年度	給油取扱所における検定作業中の着衣着火事故
中等傷者発生事故	令和元年度	車両の接触による給油取扱所（セルフ）の固定注油設備等の破損
中等傷者発生事故	平成30年度	給油取扱所において地下貯蔵タンク解体作業中における火災
中等傷者発生事故	平成30年度	給油取扱所において顧客が給油レバーを誤って操作したことによるガソリンの流出。当該給油取扱所内において、顧客が普通自動車にガソリンを給油後、ノズルを戻す際に誤ってレバーを握ったことから、ガソリンが吐出し、顔や体にガソリンを被ったもの。 (中等傷者数：1名、怪我の種別：刺激、施設への被害：なし)
中等傷者発生事故	平成30年度	ガソリンスタンドにおいて給油中誤ってガソリンがかかり負傷した。セルフスタンドで親子で給油行為したもの。 (中等傷者数：1名、怪我の種別：化学熱傷、施設への被害：なし)
中等傷者発生事故	平成30年度	給油取扱所（屋外セルフ）における油配管点検時の流出事故
中等傷者発生事故	平成30年度	給油取扱所の廃止に伴う地下タンク（ガソリン）洗浄作業中の酸欠事故。

図 2-6 消防庁の過去 5 年間の事故データ分析結果（1）

出所）消防庁事故データより NRI 作成

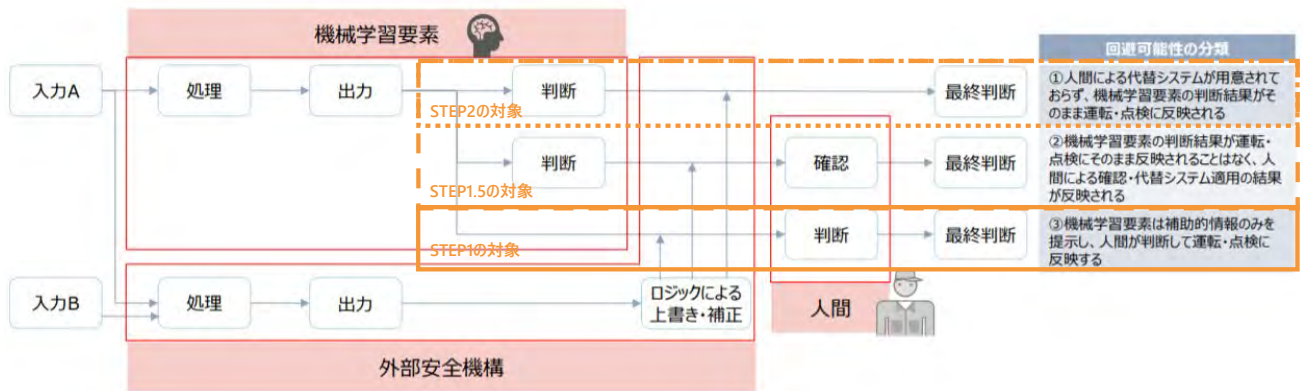
本ガイドラインの 基準	「想定される影響」の既存評価基準			「発生した事故」の既存評価基準		
	機械学習品質マネジメントガイド ライン 第 1 版	【高圧ガス】 リスクアセスメント ガイドライン(Ver. 2) リスクアセ メント実施事例	【労働安全】 危険性又は有害性等の 調査等に関する指針 別 添 4 負傷または疾病の 重篤度	【高圧ガス】 高圧ガス・石油コンビナート事故対 応要領 事故の分類	【消防】 危険物施設における 火災・流出事故に係 る深刻度評価指標 人的被害指標	【石化協】 事故評価基準 (CCPS評価法)
I 人的被害：死亡／障の残る傷害／多数 の重傷／極めて多数の負傷(重傷・軽傷) 経済的被害(直接被害額のみ)：直接被害 額 1 億円以上 経済的被害(間接被害額を含む)：企業体 としての存続等に著しい影響／業務の運営 を揺るがす重大な損害	複数人の同時死 亡 単一の人の死傷 障の残る傷害	企業体として の存続等に 著しい影響 業務の運営 を揺るがす 重大な損害	I：死亡	A 級事故 ・死者 5 名以上 ・死者・重傷者合計 10 名以上 ・死者・負傷者合計 30 名以上 ・直接被害額 5 億円以上 B1 級事故① ・死者 1 名以上 4 名以下	レベル 1：死者が発生	レベル 1 ・複数死亡 ・直接被害額 10 億円超 レベル 2 1 名死亡 ・直接被害額 1 億～10 億円
II 人的被害：重軽傷／多数の負傷(重傷・軽 傷) 経済的被害(直接被害額のみ)：直接被害 額 1 千万円以上 経済的被害(間接被害額を含む)：無視で きない、具体的な損害	重症	無視できな い、具体的 な損害	II：休業災害 III：不休災害	B1 級事故(①を除く) ・重傷者 2 名以上 9 名以下 ・負傷者 6 名以上 29 名以下 ・直接被害額 1 億円以上 5 億円未満 C1 級事故① ・負傷者 1 名以上 5 名以下かつ重傷 者 1 名以下	レベル 2：重傷者または 中等症者が発生 レベル 3：軽症者が発生	レベル 3：休業災害 ・直接被害額 1 千万～1 億円
III 人的被害：微小な怪我 経済的被害(直接被害額のみ)：直接被害 額 1 千万円未満 経済的被害(間接被害額を含む)：軽微な 利益の喪失にとどまる III* (想定される被害者により容易に回避で きる場合)	軽傷 (想定される被害 者により容易に回 避できる場合)	軽微な利益 の喪失にと どまる	IV：微小災害	C1 級事故(①を除く)	レベル 4：軽症者なし	レベル 4 ・応急手当 ・直接被害額 250 万～1 千万円
IV 人的被害：被害の想定なし 経済的被害(直接被害額のみ)：直接被害 額の想定は軽微 経済的被害(間接被害額を含む)：間接被害 額を含む経済的被害の想定は軽微	被害の想定なし	損害の想定 なし	V：怪我なし	C2 級事故		レベル 5 ・レベル 4 未満

図 2-7 消防庁の過去 5 年間の事故データ分析結果（2）

出所）プラント保安分野 AI 信頼性評価ガイドライン第二版より一部抜粋

したがって、AISL 表の左の方が求められる AISL が高く、右の方が低く設定されている。横軸の検討にあたっては、図 2-8 に基づき、機械学習要素の出力及び判断に対する人間の関与の度合いを確認した。

これらの判断基準を踏まえ、本ガイドラインの対象となる給油許可監視システムにおける回避可能性の程度は、ロードマップの段階により異なる水準を設定している。具体的には、ロードマップ STEP1 においては③、STEP1.5 は②、STEP2 は①に該当すると判断した。



STEP	AIがOK判断	AIがNG判断
STEP 1	ヒトが目視確認の上、給油許可を実施	ヒトが目視確認の上、対応実施
STEP 1.5	< 特定の条件下のみ > AIが給油許可を実施 < 上記以外の通常時 > ヒトが目視確認の上、給油許可を実施	ヒトが目視確認の上、対応実施 (給油中以降の場合はAIがポンプを自動停止)
STEP 2	AIが給油許可を実施 (ヒトの目視確認不要)	ヒトが目視確認の上、対応実施 (給油中以降の場合はAIがポンプを自動停止)

図 2-8 「リスク回避性」の横軸の判定フロー

出所) プラント保安分野 AI 信頼性評価ガイドライン第 2 版、石油元売会社ヒアリングより NRI 作成

(3) AISL 評価のまとめ

以上の評価を踏まえると、リスク回避性軸のレベルはロードマップの段階によって異なり、STEP1～1.5 は AISL0.2、STEP2 においては AISL1 となる。AISL 評価の結果をまとめると、以下の通りとなる。

SSAI	SIL評価の必要性の有無 (SIL評価を行う場合)	SIL評価結果	なし	縦軸 II、横軸①～②(STEP1.5)または、③(STEP1)と判断。ただし、外部安全機構の設置により、AISLの1段階軽減措置ありと判断。 縦軸 II の理由: 過去のSS事故事例を鑑みると、過去5年間で給油動作起因の死亡・重傷者発生事故は0件であり、中等傷者も2件で重大な事故は発生していないことから、IIが妥当であると判断。また、AIが誤判断(異常な給油動作を見逃す)をした場合においても、敷地内にスタッフが常駐していることにより迅速な駆けつけ対応が可能であることから、縦軸 II として品質管理を行う。 (STEP1.5)横軸①～②の理由: 特定の条件下においてはAIが判断するが、それ以外の条件下では人間が確認するプロセスを経るため。 (STEP1)横軸③の理由: 人の判断材料としてAIを使うため。
	(SIL評価を行わない場合、もしくはSILなしの場合)	考慮する影響の種類(人的被害・直接的経済被害のみ/間接的経済被害を含む) 影響の強度(I～IV) 人間による回避可能性(①～③)	人的被害・直接的経済被害のみ II STEP1: ③ STEP1.5: ①～②	
	AISL		STEP1: 0.1 STEP1.5: 0.2	

図 2-9 リスク回避性軸のレベル設定 (AISL)

出所) プラント保安分野 AI 信頼性評価ガイドライン第 2 版、石油元売会社ヒアリングより NRI 作成

2.4.2 パフォーマンス：AIPL の設定

「パフォーマンス」に関する外部品質レベルの設定および評価基準表を次の通り示す。最も高いレベルである AIPL2 は、機械学習要素が一定の性能指標を満たすことが、機械学習利用システムの運用上必須、又は強い前提である場合に相当する。また、AIPL1 は、一定の性能要件が機械学習利用システムの目的として特定されているが、AIPL2 に該当しない場合（性能要件の達成がベストエフォートでよく、必達でない場合）に相当する。また、AIPL0 は、性能指標が特定されず、性能指標そのものを発見することが開発の目的となる場合などに相当する。この AIPL2・1・0 の評価基準は、プラント AI ガイドラインにおいて示されている定義・基準と同一である。



パフォーマンスレベル	説明
AIPL 2 (mandatory requirements) ロードマップ STEP 2	- 機械学習要素が一定の性能指標（正答率・適合率・再現率など）を満たすことが、機械学習利用システムの運用上の必須または強い前提である場合。 - 受発注等の契約において、前記の一定の性能指標の充足が受入要件として明確に記載される場合。
AIPL 1 (best-effort requirements) ロードマップ STEP 1～1.5	- 一定の性能要件が機械学習利用システムの目的として特定されているが、AIPL 2 に該当しない場合。 - 特に、リリースまでの日程スケジュールが重視される場合、または品質をモニタリングしながら試験運用を行い、漸次性能向上を行う運用が許される場合。
AIPL 0	- 性能指標が開発時点で特定されず、性能指標そのものを発見することが開発の目的となる場合など。 - 所謂PoCの段階で終了する開発を行う場合。

図 2-10 パフォーマンスのレベル設定・評価基準の定義（AIPL）

出所）プラント保安分野 AI 信頼性評価ガイドライン第 2 版、石油元売会社ヒアリングより NRI 作成

このように、AIPL は、基本的には、「パフォーマンス」に関する外部品質がシステムにおいてどの程度強い要件となっているかで決定し、業界で合意した性能の水準と要求水準の厳しさをユースケースごとに判断される。本ガイドラインにおいては、ロードマップの STEP1～2 までの段階に応じて異なる AIPL を設定しており、ロードマップ STEP1～1.5 の段階においては AIPL1、STEP2 の段階においては AIPL2 と設定した。

SSAI	性能の水準（高/中/低）	中	【性能の水準】 STEP 1～1.5においては、AISLに優先順位を劣後させる一方、商用化に至るために一定の性能水準を検証しながら定めていくこととするため、性能の水準は中程度と置く。 【水準に対する要求の厳しさ】 また、正常な場合に異常と判定する誤判定があると、異常のない（人間による確認が不要な）給油動作を人間が確認する手間が余計にかかるが、重大なコストではない。そのため、運用上許容される範囲において、best-effortで品質を高めることを目指す。 【その他、確認すべき事項】 なお、アラートを出すタイミング・閾値は運用しながら検証していくこととする。本システムの運用の前提や、評価のプロセスについて、共通認識を形成する。	
	水準に対する要求の厳しさ (mandatory/best-effort/なし)	best-effort		
	AIPL	STEP1～1.5：AIPL 1 STEP2：AIPL2		

図 2-11 パフォーマンス軸のレベル設定（AIPL）

出所）プラント保安分野 AI 信頼性評価ガイドライン第 2 版、石油元売会社ヒアリングより NRI 作成

2.5 内部品質

「内部品質」とは、AI システムの構成要素を作成する際に「具体的に測定したり、設計などの開発行為を通じて評価したりする、その要素が固有で持つ特性としての品質」を指す。

AI システムの誤判断は、学習データセットの不十分や過学習など、様々な要因によって起こり得るため、適切な信頼性評価を実施するためには、AI システムの導入における全工程の妥当性を網羅的に確認する必要がある。そのため、本ガイドラインでは、「機械学習品質マネジメントガイドライン 第 2 版」の要求事項に鑑みて、データの設計から実装・運用の品質に至るまで、システム導入の工程を 4 つの区分に分類し、それぞれの区分に 2 つ～3 つずつの内部品質を設定した（すなわち、合計で 9 つの評価軸がある）。

2.6 内部品質における評価軸

以下では、内部品質における 9 つの評価軸について言及する。なお、各評価軸の定義は、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 第 2 版」における定義を引用している。

(1) 要求分析の十分性

要求分析の十分性とは、「機械学習利用システムの実世界での利用状況に対応して機械学習要素に入力されると想定される運用時の実データの性質について分析が行われ、その分析結果が想定される全ての利用状況を被覆していること」を指す。具体的には、機械学習要素が対応すべき動作の対象・範囲を明らかにすることや、入力データの範囲を特定することなどが要求事項として挙げられる。

(2) データ設計の十分性

データ設計の十分性とは、「要求分析の十分性を前提として、システムが対応すべき様々な状況に対して十分な訓練用データやテスト用データを収集し整理するためのデータ設計の十分な検討を実施すること」を指す。具体的には、AI システムの品質が低下するリスクを有するケースを網羅することや、取り扱える範囲のデータの属性及びデータ量とすることなどが要求事項として挙げられる。

(3) データセットの被覆性

データセットの被覆性とは、「対応すべき状況の組み合わせの各々に対して、状況の抜け漏れがなく、十分な量の学習データが与えられていること」を指す。具体的には、データセットの取り方などを工夫し、必要なデータを網羅的に確保することなどが要求事項として挙げられる。

(4) データセットの均一性

データセットの均一性とは、「データセット内の各状況や各ケースが、入力されるデータ全体におけるそれらの発生頻度に応じて抽出されていること」を指す。具体的には、データセット全体を取得するプロセス

に偏りが生じないよう配慮しつつ、個々の属性値の発生頻度などを適宜監視することなどが要求事項として挙げられる。

(5) データの妥当性

(3)(4)のデータセットの分布に関する性質とは対照的に、データセット中のデータ 1 つ 1 つが訓練の目的に照らして妥当であることを、「データの妥当性」(Adequacy of data)とする。妥当性には、値に誤りが無いことだけでなく、訓練に使われるべきでないデータが（たとえ値そのものが正確であっても）含まれないこと（一貫性）、データに不適切な改変などがされていないこと（信憑性）、データが十分に適切に新しいものであること（最新性）などが要求事項として挙げられる。

また、教師あり機械学習の文脈では、訓練対象としての測定値等（機械学習要素を関数と見たときの入力側にあたる値）の妥当性である「データ選択妥当性」と、訓練用に付加された正解値（出力側にあたる値）の妥当性である「ラベリングの適切性」の 2 つの観点が含まれる。

(6) 機械学習モデルの正確性

機械学習モデルの正確性とは、「学習データセット（訓練用データ、テスト用データ、バリデーション用データからなる）に含まれる具体的な入力データに対して、機械学習要素が期待通りの反応を示すこと」を指す。具体的には、データセットの入力データに対して、機械学習要素が外部品質の期待に沿った出力であることなどが要求事項として挙げられる。

(7) 機械学習モデルの安定性

機械学習モデルの安定性とは、「学習データセットに含まれない入力データに対して、機械学習要素が学習データセット内のそれに近いデータに対する反応と十分に類似した反応を示すこと」を指す。具体的には、データセットに含まれない入力データに対して、その入力データ近傍のデータセットを入力したときの出力と類似していることなどが要求事項として挙げられる。

(8) プログラムの健全性

プログラムの健全性とは、「機械学習の訓練段階に用いる訓練用プログラムや、実行時に使われる予測・推論プログラムが、与えられたデータや訓練済み機械学習モデルなどに対してソフトウェアプログラムとして正しく動作すること」を指す。具体的には、利用するソフトウェアは信頼できるソフトウェアを用いることや、機械学習要素の開発環境と運用環境の相違とそれによる影響を予め検討することなどが要求事項として挙げられる。

(9) 運用時品質の維持性

運用時品質の維持性とは、「運用開始時点で充足されていた内部品質が、運用期間中を通じて維持されること」を指す。具体的には、機械学習要素の更新頻度の見積もりまたは更新の必要性の

判断基準を事前に検討することなどが要求事項として挙げられる。

2.7 内部品質における各評価軸の要求レベル

2.6 でレベルを設定した外部品質は、内部品質を作り込むことにより達成する。本ガイドラインで設定した AISL/AIPL に応じて、内部品質の要求水準（要求レベル）を設定する。外部品質の AISL/AIPL と内部品質の要求レベルとの対応関係を以下の通り示す。例えば、AISL0.2 もしくは AIPL2 の場合、内部品質の要求は「レベル 2」となり、内部品質の 9 つの軸それぞれに設定された「レベル 2」の要求事項を適用することとなる。複数の AISL/AIPL について、それぞれ内部品質のレベルを設定し、最も高い内部品質のレベルを用いて要求事項を適用する。

本ガイドラインで設定した AI システムが満たすべき内部品質の要求レベルは、ロードマップの段階によって異なる。STEP1 の段階においてはレベル 1、STEP1.5 においてはレベル 2 の要求事項を満たすべきこととする。

リスク回避水準 AISL	パフォーマンス水準 AIPL	内部品質の要求レベル		
		レベル 1	レベル 2	レベル 3
AISL 0.1	AIPL 1	○ STEP1	—	—
AISL 0.2	AIPL 2	—	○ STEP1.5	—
AISL 1	—	STEP 1 ～1.5で目指すべき要求レベル		

図 2-12 外部品質の AISL/AIPL と内部品質の要求レベルの対応

出所）プラント保安分野 AI 信頼性評価ガイドライン第 2 版、石油元売会社ヒアリングより NRI 作成

STEP1～1.5 においては、内部品質の優先度を設定し、AI の非機能面（性能やセキュリティ等）に作用する内部品質は優先度を下げ対応を行うこととした。具体的には、STEP に応じて徐々に要求レベルを上げることとし、STEP1 においては、すべての評価軸において要求レベル 1 の実現を目指し、STEP1.5 においては、④データセットの均一性、⑦機械学習モデルの安定性の評価軸は要求レベル 1、その他要求事項については要求レベル 2 を目指すこととする。

		『人間』が給油許可		『AI』が給油許可
		STEP1	STEP1.5	STEP2
1. データ設計	①要求分析の十分性	要求レベル 1	要求レベル 2	要求レベル 3
	②データ設計の十分性	要求レベル 1	要求レベル 2	要求レベル 3
2. データ品質	③データセットの被覆性	要求レベル 1	要求レベル 2	要求レベル 3
	④データセットの均一性	要求レベル 1		要求レベル 2 ～ 3
	⑤データセットの妥当性	要求レベル 1	要求レベル 2	要求レベル 2 ～ 3
3. モデルの学習品質	⑥機械学習モデルの正確性	要求レベル 1	要求レベル 2	要求レベル 3
	⑦機械学習モデルの安定性	要求レベル 1		要求レベル 2 ～ 3
4. 運用の品質・実装	⑧プログラムの健全性	要求レベル 1	要求レベル 2	要求レベル 3
	⑨運用時品質の維持性	要求レベル 1	要求レベル 2	要求レベル 3

図 2-13 内部品質の要求レベル

出所) プラント保安分野 AI 信頼性評価ガイドライン第 2 版、石油元売会社ヒアリングより NRI 作成

3 AI システムの試験方法

本章では、第2章で定義した3つの品質やその要求レベルについて、具体的な試験方法を示す。

3.1 AI システムの試験方法におけるアプローチ

本ガイドラインにおいては、自動運転分野における安全保障に関するアプローチを参照し、網羅性を担保可能な実用性のある「原理原則に基づくシナリオベースアプローチ」を適用する。自動運転分野における安全保障に関するアプローチを示したものは、以下の通りである。

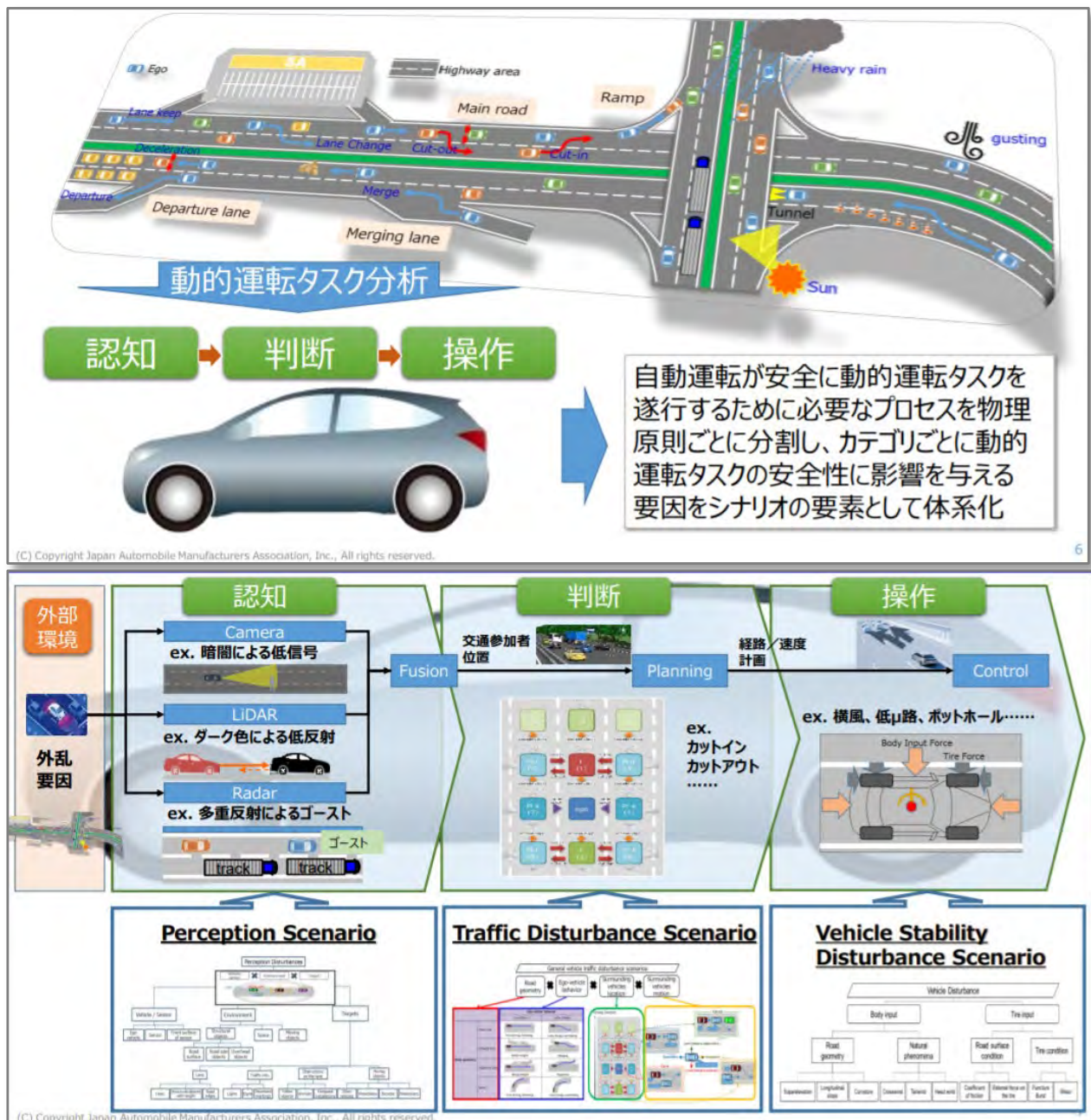


図 3-1 自動運転分野における安全保障に関するアプローチ

出所) 一般社団法人日本自動車工業会「オーナーカー自動運転の安全性評価の取り組み」

3.2 原理原則に基づく評価シナリオの体系化

前述のアプローチを参照し、本ガイドラインにおいては、画像 AI を活用したセルフ給油許可監視に必要な要素毎にシナリオを設定し、これらの評価シナリオの充足度を確認することにより、システムの有用性を検証する。図 3-2 に試験方法の大まかな流れを示す。



図 3-2 原理原則に基づくシナリオの体系化

出所）NRI 作成

AI に対する試験範囲の十分性について、給油顧客の行動や SS 環境の組み合わせは無限に存在し、単純に給油環境を整理、分類して組み合わせるだけでは現実的に検証を実施しきれず、十分な検証範囲を確保できないという課題がある。そこで、給油顧客の行動と SS 環境の組み合わせによる給油環境を俯瞰的に見て分類するアプローチではなく、給油許可監視タスクを実行するために必要なプロセスを物理原則の異なる認知、判断、操作の3要素「認知：認識外乱」、「判断：顧客行動に関する外乱」、「操作：給油許可環境外乱」に分解し、プロセス毎に処理結果に影響を及ぼす要因（Root Cause、根本要因）をシナリオ体系として構造化することで、有限かつ安全の観点で網羅的な範囲の特定を可能にする。

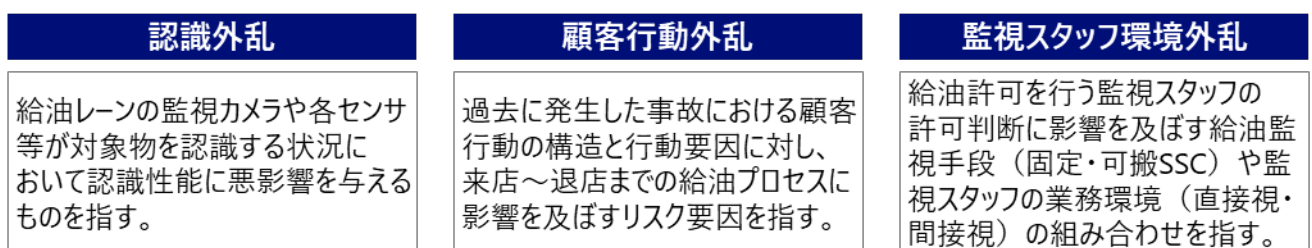


図 3-3 認知・判断・操作の3要素

プロセス	処理結果	外乱	物理原則
認知	危険物の有無、車両／給油顧客／スタッフ／ノズル等の位置情報	認識外乱	センサメカニズムに応じた原理的な外乱 例）カメラ：可視光、サーモカメラ：赤外光
判断	給油許可判断（正常な給油動作、危険な給油動作の判断）	顧客行動外乱	正常な給油顧客の動作＋危険な給油動作とリスク因子
操作	給油許可判断を達成するための給油監視スタッフの動作	給油許可環境外乱	給油許可判断に影響を及ぼす給油顧客と監視スタッフの位置関係と監視スタッフの兼務状況

図 3-4 体系化するシナリオの 3 要素と外乱要因の対応

出所）NRI 作成

	認識外乱シナリオ	顧客行動外乱シナリオ	監視スタッフ環境外乱シナリオ
目的	システムの主要センサであるカメラの認識機能の妥当性を評価	来店～退店までの給油動作に対してシステムが安全性を正しく判断していることを評価する	システムを活用した給油許可監視業務において、従来通りの安全性が担保されていることを評価する
評価対象	システム	システム	業務
試験シナリオ策定方法	カメラがとらえるモノ・ヒト等のバリエーションに着目し、それらを想定通りに AI が認識するかについて検証を行う	給油顧客の動作に着目し、正常な給油動作と過去に発生した事故事例を基に検証を行う	監視スタッフが給油許可業務を実施する状況に着目し、AI・スタッフの判断結果や AI 単独で判断する場合を想定し、検証を行う
評価基準	- 誤判定率 - 誤検知率	原則としてすべての試験シナリオに対し、想定された動作が行われること	従来の監視スタッフ単独での給油許可判断結果との一致率 *監視スタッフが誤判断した場合を除く

図 3-5 各シナリオの位置づけ

出所）NRI 作成

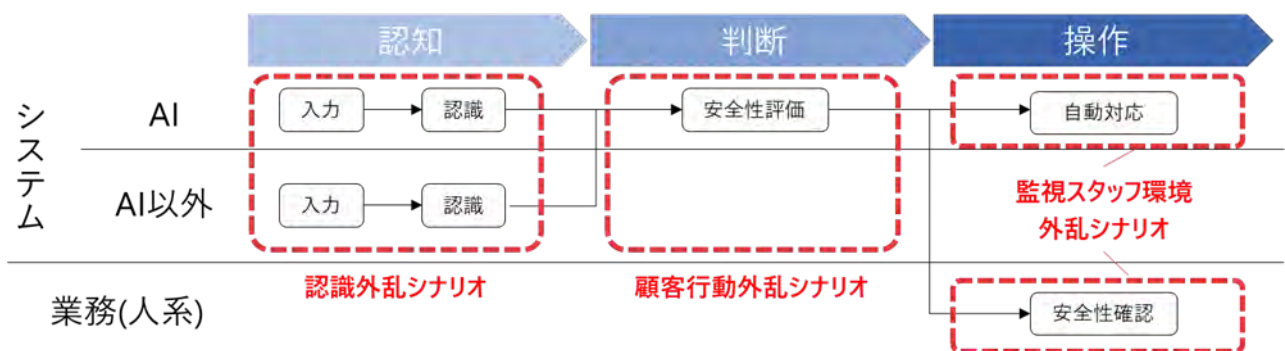


図 3-6 シナリオ毎の評価範囲イメージ

出所）NRI 作成

3.2.1 認識外乱シナリオ

認知に関わる認識外乱シナリオは、主に監視カメラが対象物を認識する状況において、認識性能に影響を与える（以下、認識外乱という）要素を検討するものである。カメラの画質によっては異なるものに間違えたり、天候・立地等によって反射したり陰になったりして違う色に見えたり、また車両や給油を行う人の位置関係によって隠れて見えなくなるといった、誤認識や検出漏れに繋がりうる要因を抽出し、その中から影響度や発生頻度を元に代表シナリオを選定する。

当該シナリオを検証することにより、誤認識（例：ノズルでないもの（傘の取っ手、衣類など）を「ノズルである」と認識する）や検出漏れ（例：ノズルを「ノズルではない」と認識する）が発生しないこと確認することにより、AI が正しく安全性を判断するために必要となる入力情報の正確性を担保する。



図 3-7 誤認識と検出漏れのイメージ

出所）NRI 作成

これら認識外乱の要因は、「カメラ・センサ」、「周辺環境」及び「認識対象物」の 3 つに大別でき、それぞれを詳細かつ、各階層において網羅的に分解・分類することで認識外乱要因体系を構成する。例えば、構造、相対位置、種別などの視点で要因分解し、色、形状、材質、挙動などの階層まで分類を続けていく（図 3-8）。

また、認識外乱において最も重要な要素の一つであるカメラの知覚処理においては、認識対象物からの信号 S を特徴づける物理量として、強度、方位・範囲、信号の変化および取得時刻がある。図 3-9 にカメラの近くにおける外乱発生原理を体系化したものを示す。

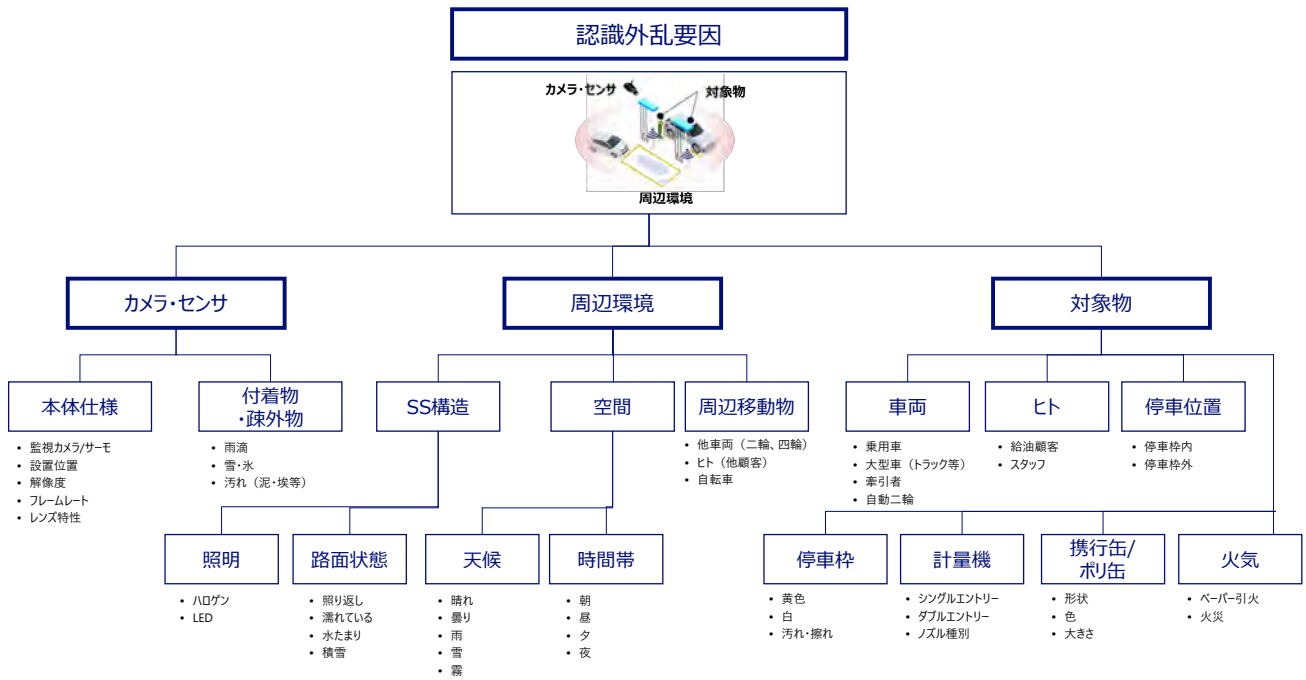


図 3-8 認識外乱要因の体系

出所) NRI 作成

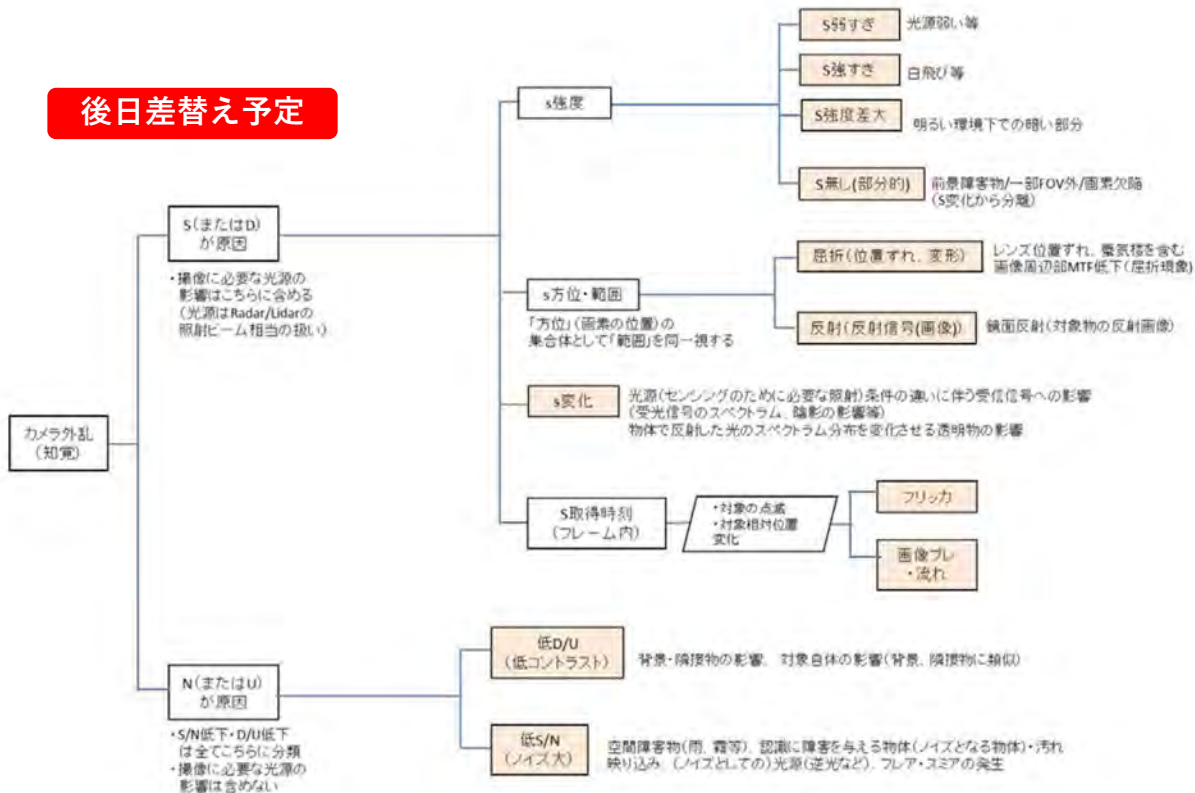


図 3-8 カメラの知覚における外乱発生原理

出所) 「自動運転の安全性評価フレームワーク v1.0」より一部抜粋

このような認識性能に影響を及ぼす外乱要因は、周辺環境の条件や認識対象物の種類などのバリエーションや、それらの組み合わせを考慮すると非常に多岐に渡る。そのため、本ガイドラインではカメラ・センサ原理に基づいて認識外乱要因を分類し、同一の分類となるものの中からその代表要因を選択することで、シナリオ数を有限化するとともに、認識外乱要因を網羅するアプローチを採用する。

本ガイドラインでは、具体的な代表シナリオの検証観点、試験条件および試験方法については、「認識外乱の要素の抽出」→「試験シナリオの選定」→「検証観点および試験条件（試験環境）の定義」→「試験方法の定義」の 4 つのプロセスで決定した。代表シナリオの具体的な決定方法は次のとおりである。

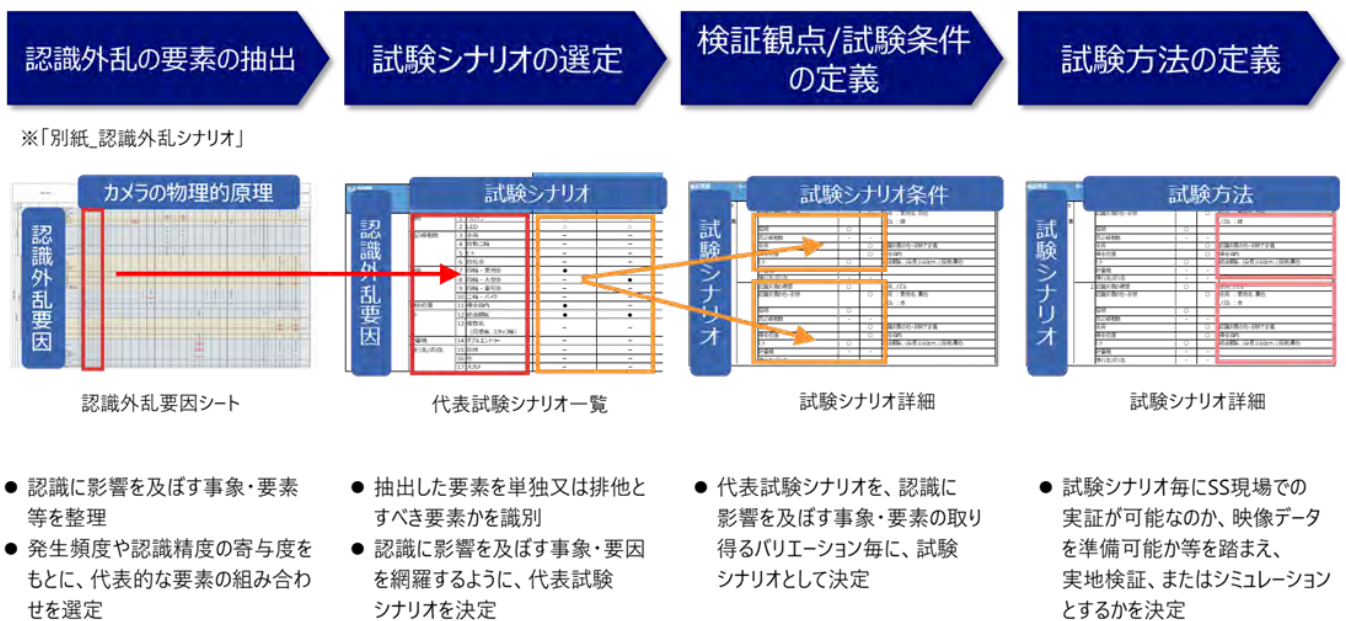


図 3-9 認識外乱要因の体系

出所）NRI 作成

(1) 認識外乱要因の要素抽出

認識外乱の要因と発生原理との関係はマトリクスで表すことができる。このマトリクスでは、縦方向に認識外乱要因を、横方向に発生原理を並べており、各原理（＝列）についてそれを生じ得る要因（＝行）として該当するものが分かる形としている。同じ列内で該当する複数の外乱要因はどれも同じ原理によって生じるが、システムの安全性を評価するという観点から、認識性能への影響度および現場での遭遇確率が高い要因を選出し、評価シナリオとして優先順位付けを行った（後述(2)で実施）。

(2) 認識外乱シナリオの選定

認識性能への影響度と遭遇確率は、次のような考え方で点数化して同じ列内で比較した。

< 認識性能への影響の大きさ (X) >

各原理について、それぞれの要因が起こし得る認識外乱の厳しさを 1～3 点で表現
(影響小：1 点、中：2 点、大：3 点)

< 遭遇確率 (P) >

以下で説明する (頻度総合点) × (発生時の継続時間) の数値を遭遇確率と見なし、
その大きさによって 1～3 点に振り分け (確率小：1 点、中：2 点、大：3 点)

- 遭遇頻度

各外乱要因の地域性、気候・気象の影響度、使用方法による発生頻度をそれぞれ 1～3
点で表現し、4 つの点数を掛け合わせて「頻度総合点」を算出

- 発生時の継続時間

その要因が発生したときの 1 回の継続時間を 1～3 点で表現
(時間短：1 点、中：2 点、長：3 点)

< 総合点 >

認識性能への影響の大きさ (X) と遭遇確率 (P) とは、それぞれウェイト (W) を 10 および 8 に設定して合計点数 ($WXX+WPP$) を算出し、総合点とする

なお、本ガイドラインでは「認識性能への影響の大きさ (X)」、「遭遇確率 (P)」、「発生時の継続時間」は、本ガイドライン作成時点では実験データ量は十分ではないため、WG 参加者間の定性評価により決定した。今後、実験データ量が十分に蓄積された場合においては、必要に応じて定量的な根拠をもとに見直しを実施するものとする。

(3) 検証観点／試験条件の定義

検証観点は、認識外乱による影響を受けず、正しく AI が認識対象物を認識し、従来給油可能な車両等に対して給油可能であること、又は給油を禁止されている車両等については給油不可であるということを軸として設定する。試験条件については次に示す要素を整理のうえ、代表的な試験条件を決定する。

但し、要素については、網羅性は劣後させ、認識に大きく影響を及ぼす組み合わせ（例：“S 弱すぎ”の代表シナリオの場合、夜間では視認しづらい黒色の車両と黒色の衣類を着用した顧客を選択など）を代表的な要素として採用する（図 3-10）。

(4) 試験方法の定義

上記(1)～(3)で定めたシナリオにおいて、実業務で発生するケースは少なく、また実験として準備可能な車両や S S 環境等には限界がある。そのため、本ガイドラインでは、試験の実施難易度を基に試験方法を「実 SS で検証を行うべきシナリオ」と「シミュレーションで検証を行うべきシナリオ」に分類し、

試験方法を決定した。

※2022 年 2 月末時点においては、試験の実施環境及び準備可能な範囲を関係者間で検討中となる。そのため、試験方法については今後、本ガイドラインに従った実証実験を予定しているため、当該計画時に合わせて整理する。

試験要素	要素の例
認識対象の種類	車両、自動二輪、給油客、携行缶、火気など
認識対象の色・形状	車両のボディタイプ、車両のボディカラー、携行缶の大きさ、携行缶のカラー
照明	ハロゲンまたはLED
周辺移動物	歩行者、自転車、SS スタッフなど
停車位置	給油レーン内に停車、給油レーン枠外に停車など
ヒト	服装、身長、性別、年齢など
計量機	ロングホースタイプ（反対給油可能）など

図 3-10 試験条件を決定する要素一覧

出所）NRI 作成

		1 分解能不足	2 SS 強すぎ	3 SS 弱すぎ	4 SS 強度差大	5 SS 無し	6 屈折・反射	7 SS 変化	8 低 D/U	9 低 S/N	関連するシステムの動作			
試験シナリオとして採用														
試験シナリオ 1	給油エリア（給油レーン）内の車両（乗用車）の認識精度を確認する。	✓									許可 / 不許可	警告 発報	給油 停止	TOR
試験シナリオ 2	給油エリア（給油レーン）内の車両（大型車）の認識精度を確認する。	✓									許可 / 不許可	警告 発報	給油 停止	TOR
試験シナリオ 3	給油エリア（給油レーン）内の自動二輪の認識精度を確認する。	✓									許可 / 不許可	警告 発報	給油 停止	TOR
試験シナリオ 4	給油エリア（給油レーン）内の携行缶の認識精度を確認する。	✓									許可 / 不許可	警告 発報	給油 停止	TOR
試験シナリオ 5	給油エリア（給油レーン）内のポリ容器の認識精度を確認する。	✓									許可 / 不許可	警告 発報	給油 停止	TOR
試験シナリオ 6	給油エリア（給油レーン）内に複数名侵入した場合の認識精度を確認する。	✓									許可 / 不許可	警告 発報	給油 停止	TOR
試験シナリオ 7	路面や車両のガラス等に太陽光が反射している条件下における認識精度を確認する。		✓								許可 / 不許可	警告 発報	給油 停止	TOR
試験シナリオ 8	夜間かつ、雨の環境下における、車両等の認識精度を検証する。			✓						✓	許可 / 不許可	警告 発報	給油 停止	TOR
試験シナリオ 9	夜間かつ、積雪の環境下における、車両等の認識精度を検証する。			✓						✓	許可 / 不許可	警告 発報	給油 停止	TOR
試験シナリオ 10	火災・ペーパー引火の認識精度を検証する。				✓			✓			許可 / 不許可	警告 発報	給油 停止	TOR
試験シナリオ 11	ヒトや給油カバー、ドア等により検知対象物が遮蔽されるケースにおけるシステムの挙動を検証する。					✓					許可 / 不許可	警告 発報	給油 停止	TOR
試験シナリオ 12	車両等が停車枠外に停止し、監視カメラの撮影エリアからはみ出る場合のシステムの挙動を検証する。					✓					許可 / 不許可	警告 発報	給油 停止	TOR
試験シナリオ 13	車両のガラス等にノズルや携行缶、ヒト等が映り込んでいる場合の認識精度を確認する。						✓				許可 / 不許可	警告 発報	給油 停止	TOR
推奨とするに留める可能性を残しつつ採用留保														
試験シナリオ 14	給油エリア（給油レーン）内の車両（被牽引車）の認識精度を確認する。	✓									許可 / 不許可	警告 発報	給油 停止	TOR
安全面に直接資するシナリオではないため非採用														
-	ダブルエントリー（1つのカメラに2つの計量機・ノズルが映り込むケース）における認識精度を確認する。	✓									許可 / 不許可	警告 発報	給油 停止	TOR
-	停車枠が劣化し、色が薄くなっている/消えている場合における認識精度を検証する。									✓	許可 / 不許可	警告 発報	給油 停止	TOR

図 3-11 認識外乱シナリオにおける試験シナリオ案

出所）NRI 作成

3.2.2 顧客行動外乱シナリオ

顧客行動外乱シナリオは、給油顧客が行う行為に対して、給油許可監視システムが安全性を正しく判断することができるかを評価するものである。

本ガイドラインが射程としている「ノズル挿入」から「給油」、固定給油装置（計量機）に「ノズルを戻す」までの「給油動作」と、危険要因となり得る顧客の「行動・行為」を構造化して紐づけることにより、火災やガソリンの流出等、危険に繋がる可能性が認められるシナリオを抽出する。また、セルフ SS における過去の事故事例から、一定程度の発生頻度と影響度を有する、動作ステップと行動・行為の組み合わせを抽出し、代表シナリオを選定する。

(1) 顧客行動外乱要因の要素抽出

a. 給油動作のステップ

顧客行動外乱シナリオにおける「給油動作」の射程（図 3-11）は、当然ながら本ガイドラインの射程と同一であり、大きく①「ノズルの挿入」、②「給油」、③「ノズルを戻す」の 3 つに分けられる。故に、例えば静電気除去パッドに触れてから給油口を開放する行為や、計量機にノズルを戻した後の行為については、顧客行動が安全性に影響を与えるかどうかに関わらず、顧客行動外乱シナリオとしての検討の対象には含まれない。顧客による動作ステップの詳細について、次に再整理した。

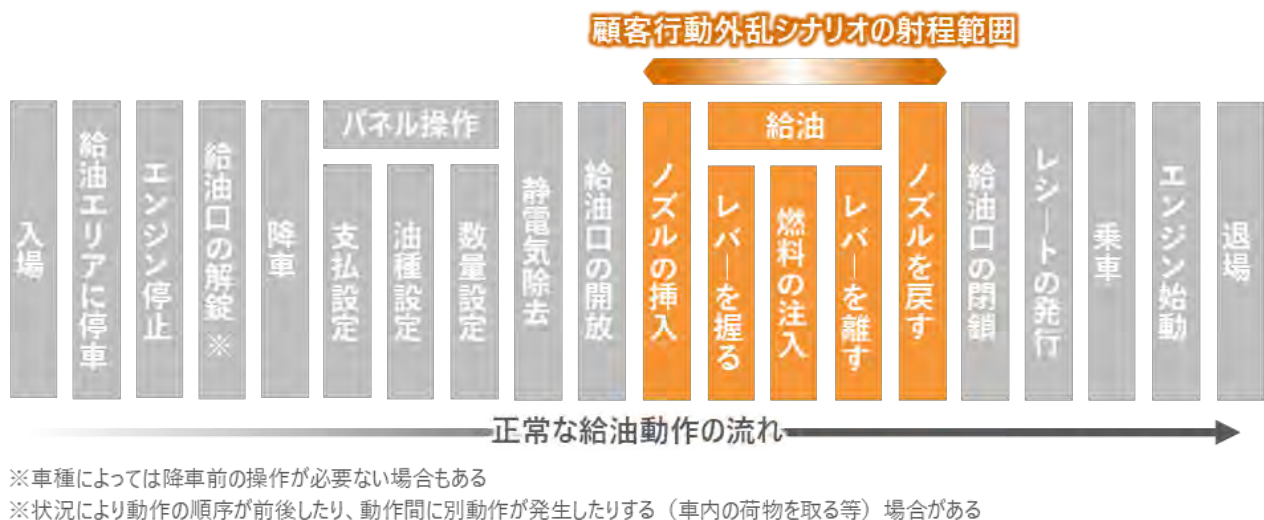


図 3-12 正常な給油動作の流れにおける顧客行動外乱シナリオの射程

出所）NRI 作成

- ① 「**ノズルの挿入**」は、給油を行う顧客が計量機からノズルを取り外し、給油口にノズルを挿入するまでの動作を指す。

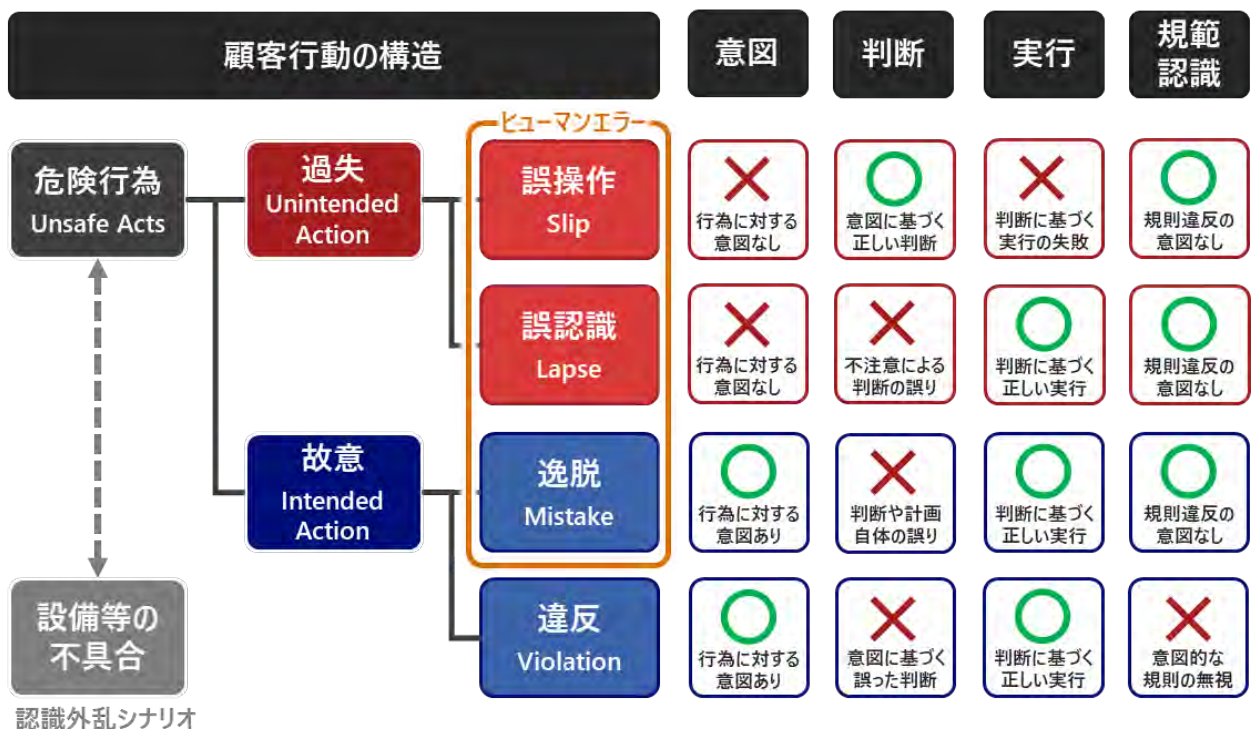
② 「給油」は、燃料を注入する行為のそれ自体であり、これらは更にレバーを握る、レバーを握り続けて燃料の注入を継続する、レバーを離すという 3 つの段階に分けることができる。本項において単に「給油」といった場合には、これらの各行為を包括した給油動作を指す。

③ 「ノズルを戻す」は、給油口からノズルを抜き取った後、計量機にノズルを戻すまでの動作を指す。

b. 顧客の行動

顧客行動外乱シナリオにおいて、顧客が取り得る行動は、その行動が発生する要因により次の図 3-12 で示すように構造化することができる。なお、危険行為(Unsafe Acts)に係る顧客行動の構造化に関する確立した分類は存在しないが、行動心理学やヒューマンエラーに関する論文⁸等を参考に構造化を行った。

顧客の行動は、その故意又は過失の有無（意図）・判断・実行・規範意識といった内心の状態と行為の結果により、大きく①「誤操作(Slip)」、②「誤認識(Lapse)」、③「逸脱(Mistake)」、④「違反(Violation)」の 4 つに分けられる。このうち、特に①～③は「ヒューマンエラー」とも呼ばれる。



※ 実行における「○・×」は結果の正しさではなく、意図した行為が行えているかをベースに評価する。
例えば、顧客が携行缶に給油しようと判断して給油した場合、判断は「×」となるが実行は「○」となる。

図 3-13 内心の状態と行為の結果に基づく顧客行動の構造化

出所）行動心理学やヒューマンエラーに関する論文等から NRI 作成

⁸ Adapted from James Reason, Human Error, Cambridge University Press

- ① 「**誤操作(Slip)**」は、顧客による危険行為のうち、過失により操作を誤り、結果として危険が生じた行為のことを指す。行為結果に対する本人の意図はなく、判断は正しかったが、注意を怠ったために実行を誤ってしまった場合、当該行為は「誤操作」に分類される

例えば、給油完了後にノズルを計量機に戻そうとしたところ、誤ってレバーを握ってしまい、ガソリンが流出・飛散するような事案が「誤操作」に分類される。

この場合において、顧客本人は不注意により意図せず結果を惹き起こしてしまったため、「意図」の評価は「×」となる。また、「判断」については、ノズルに戻そうとした判断自体は誤っていないため「○」、「実行」については、本人の意図した実行ができていない（ノズルを握ろうと思って握ったのではない）ため「×」となる他、過失であることから規範意識については「○」（意図的に規則を破ろうとした訳ではない）の評価となる。

- ② 「**誤認識(Lapse)**」は、顧客による危険行為のうち、意図した操作を行うことができたが、過失により状況認識の段階で誤りがあったために、結果として危険が生じた行為のことを指す。行為結果に対する本人の意図はなく、本人の認識・判断の下においては実行も正しかったが、認識自体を誤ったために危険を生じた場合、当該行為は「誤認識」に分類される。

例えば、ノズルが車両に差し込まれたままであることを失念し、車両を発進させたため、ホースが破損しガソリンが噴出するような事案が「誤認識」に分類される。

この場合においても、顧客本人は不注意により意図せず結果を惹き起こしてしまったため、「意図」の評価は「×」となる。また、「判断」については、不注意により車両にノズルが差し込まれたままになっていることを認識・判断できなかったことから「×」の評価となる。「実行」については、車両を発進させようという本人の意識・作為の下で、車両を“正しく”発進させることができることから、評価は「○」となり、過失であることから規範意識については「○」となる。

- ③ 「**逸脱(Mistake)**」は、顧客による危険行為のうち、法令違反に該当するとまでは言えないが、推奨されない手順や行為により、危険が生じた場合の行為のことを指す。

なお、英語表記は「Mistake」となっているが、過失による行為を含まない概念であるため、日本語でいう「間違い」とは意味合いが異なることに留意が必要である。

例えば、子どもと一緒に給油を行おうとしたところ、子どもが誤ってノズルを車両から引き抜いた結果、ガソリンが飛散して負傷するような事案が「逸脱」に分類される。

この場合においては、子どもと一緒に給油を行うという行為自体は意図的に行われたものであるため、「意図」の評価は「○」となる。また、「判断」については、子どもと一緒に給油を行おうとする行為は推奨されない行為であるため「×」、「実行」については、子供と一緒に給油するという意図した行為が行えているため「○」、「規範意識」については、法令違反を行う明確な意思はないため「○」の評価となる。

- ④ 「違反(Violation)」は、顧客による危険行為のうち、法令に違反する行為が行われた結果として、危険が生じた場合の行為のことを指す。

例えば、セルフ SS 内で給油中にタバコを吸おうとライターに着火したところ、可燃性蒸気に引火し炎上するような事案が「違反」に分類される。

この場合、顧客は意図的に行為を行っているため、「意図」の評価は「○」となる。また、「判断」については、意図している行為が法令違反となるため「×」、「実行」については、本人の意図した行為を正しく行えていることから「○」、規範意識については当然に「×」となる。

(2) 顧客行動外乱シナリオの選定

a. 事故の発生頻度と影響度に着目したシナリオ候補の整理

前述した「動作のステップ」と「顧客の行動」によりマトリクスを作成し（図 3-14）、さらにセルフ SS における過去の事故事例から、事故の「発生頻度」と「影響度」の観点に着目した事例整理を行うことで、重要性の高いシナリオを抽出した。

事故の「発生頻度」と「影響度」の 2 つの評価軸により、当該事故リスクの重要度を理解するというアプローチは、国際天文学連合(IAU)におけるトリノスケールの考え方を参考にしている。

トリノスケールは、地球近傍天体が地球へ衝突するリスクの深刻度を 0～10 の 11 段階で評価するもので、ここでは衝突した場合の「エネルギー」と「衝突確率」の 2 つの軸で、天体のリスクが評価される。セルフ SS における事故類型の重要度の検討にあたっては、「衝突時のエネルギー」を「事故の影響度」、「衝突確率」を「事故の発生確率（発生頻度）」に類する概念と捉えることで、2 つの軸で客観的に事故事例の整理を行うこととした（図 3-15）。



図 3-14 顧客行動外乱シナリオのマトリクス

出所) NRI 作成

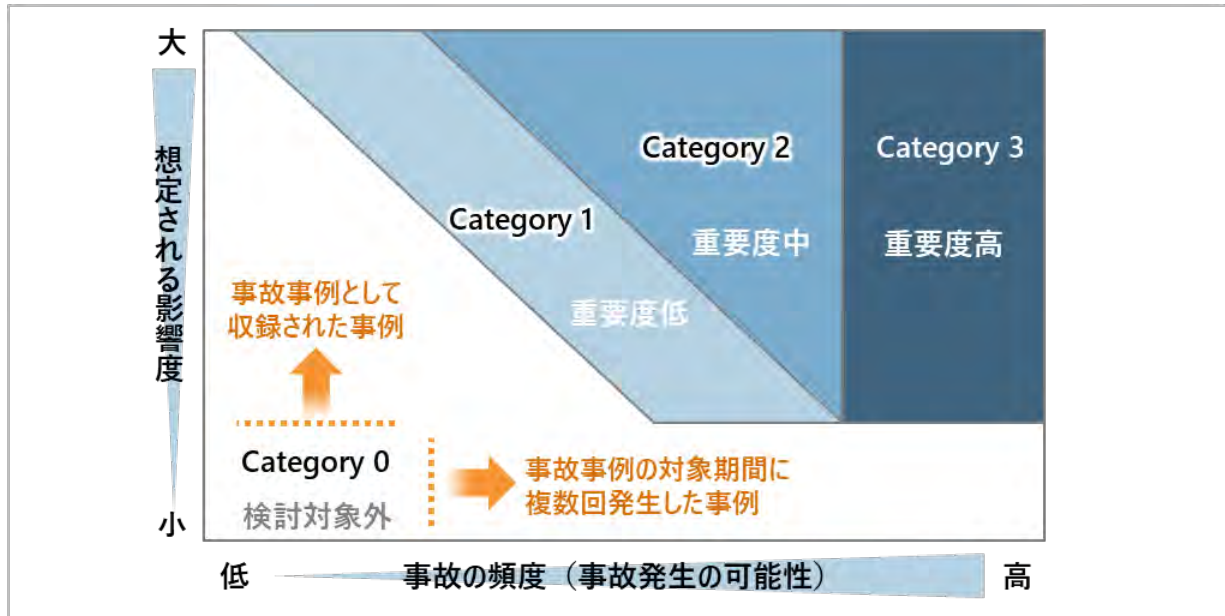


図 3-15 トリノスケールに見る事故類型の重要度の概念的整理

出所) NRI 作成

過去の事故事例については、消防庁及び損害保険会社から提供されたものを使用した。事故事例の収録対象期間は、消防庁のリストについては過去 5 年間、損害保険会社のリストについては過去 9 年間である⁹。

事故の発生頻度を評価するにあたり、発生の過程や結果が類似している事故、例えば、誤操作によりガソリンが 0.5L 流出した事故と 2.0L 流出した事故については、事故事例の抽象化を行い、同一の事故類型として取り扱ったうえで発生件数をカウントし（この場合は誤操作によるガソリン流出 2 件とカウントする）、集計期間のうち 2 回以上発生している事故事例のピックアップを行った。

また、事故の影響度については、単に事故が発生したということに留まらず、実際の物的損害または人的被害が発生したことによって、事故事例に収録されているものを検討の対象とした。

このように「動作のステップ」と「顧客の行動」の組み合わせによる事故事例の整理を行うことによって、顧客行動外乱によって生ずる事故リスクの類型を網羅的に整理し、顧客行動外乱シナリオにおける試験シナリオ候補を抽出した。

b. 検知可能性と回避可能性に着目したシナリオ選定

3.2.2 (2) a. で先述した通り、一定程度の発生頻度及び影響度を有する事故事例を抽出した後、当該事故の「**検知可能性**」と「**回避可能性**」に着目し、双方が認められるものを試験シナリオとして選定することを検討した。

⁹ 消防庁：2016 年 1 月～2020 年 12 月，損害保険会社：2010～2016 年及び 2016 年 9 月～2019 年 2 月

- ① 「**検知可能性**」は、安全を阻害する事象・要因が、カメラ・センサで検知できるかどうか、換言すれば、給油許可監視システムが危険の発生を予見し得る状態にあるかという評価観点である。現行（有人対応）の給油における監視項目ではないものについては、たとえ技術的には検知可能であるものであっても、検知可能性がないものと看做す。

ノズルを計量機から取った際にノズル内に残留していた燃料が飛散する事案を例にとると、ノズル内の残留燃料の有無は現行の監視項目ではないため、技術的に可能であるかといった評価の内容に関わらず、試験シナリオの選定にあたっては検知可能性がないものと看做しシナリオから除外する。

- ② 「**回避可能性**」は、給油許可監視システムの検知・判断によって結果回避可能な性質の危険に繋がる事象・要因であるかどうかを検討する評価観点である。

例えば、故意にガソリンを漏出させたうえ、ライターを使用し放火を図るような事案では、たとえ火気をシステムが検知できたとしても、結果回避に繋がることはないため、このような事案は試験シナリオの選定にあたっては回避可能性がないものと看做しシナリオから除外する。

整理した各種の事故事例をこれらの評価観点で検討し、検知可能性と回避可能性の双方が認められるシナリオを試験シナリオとして選定した（図 3-16）。なお、試験シナリオの具体的内容に関しては、WG 参加者間で引き続きアップデートを実施する予定である。

		関連するシステムの動作		
試験シナリオ		給油 不許可	警告 発報	給油 停止
試験シナリオ 1	ノズルを固定給油装置から外した後、給油口に挿入する前にレバーを握り、燃料が出ないことを確認する			
試験シナリオ 2	給油開始後、燃料注入中に意図せずノズルが脱落した場合、警告が発報されるとともに、直ちに給油が自動停止することを確認する	給油 不許可	警告 発報	給油 停止
試験シナリオ 3	ノズルが車輦に挿入された状態で、給油者が車輦周辺から離れた場合、警告の発報や給油の自動停止等の措置が取られることを確認する	給油 不許可	警告 発報	給油 停止
試験シナリオ 4	車輦周辺に携行缶を検知した場合、給油許可が行われないこと、また給油許可後であれば警告発報等の措置が為されることを確認する	給油 不許可	警告 発報	給油 停止
試験シナリオ 5	火気を検知した場合、警告発報とともに給油の不許可または給油の自動停止等の措置が取られることを確認する	給油 不許可	警告 発報	給油 停止

図 3-16 顧客行動外乱シナリオにおける試験シナリオ案

出所）NRI 作成

(3) 検証観点／試験条件の定義

試験シナリオとして選定した顧客行動外乱があっても、給油許可監視システムの適切な判断及び動作により、安全が確保されるかが検証の観点（利用時品質）となる。外部品質となる想定するシステムの動作は、大きく①給油不許可、②警告発報、③給油停止に分けられる。

- ①「給油不許可」は、給油許可前に試験シナリオにおける顧客行動外乱が検知された場合、システムが異常な給油動作と判断し、給油が許可されないことを指す。

この場合、ノズルに燃料は供給されず、レバーを握っても燃料が漏出することはない。給油許可後に給油許可を取り消す場合は、後述する「給油停止」に分類される。

- ②「警告発報」は、試験シナリオにおける顧客行動外乱が検知された場合、監督者若しくは顧客又はその双方に対し光や音声等の手段で警告を行うことを指す。監督者に対する警告では、監督者に状況確認と有人による給油停止等の判断や、TOR の要求を行うことを想定する。また、顧客に対する警告では、当該危険行為の中止や正常な給油方法への是正要求を行うことを想定する。

これらの場合、警告発報があった段階で直ちに給油が停止することはないため、給油許可があった後であれば、レバーを握ることによって燃料の注入が可能な状態は維持される。

- ③「給油停止」は、試験シナリオにおける顧客行動外乱が給油許可後に検知された場合、セルフ SS における当該レーンまたは全レーンの給油を直ちに自動停止することを指す。

この場合、給油停止の対象となったレーンでは、給油許可が為されたかどうかに関わらず、燃料が供給されないため、レバーを握っても燃料が漏出することはない。

(4) 試験方法の定義

顧客行動外乱シナリオには、給油中のノズル脱落や火気使用といった、試験を行う上で燃料の漏出や火災に繋がる危険性のある検証内容が含まれる。実際の試験のオペレーションについては今後の調整事項となるが、認識外乱シナリオといった他のシナリオと同様に、試験の実施難易度により、休業時間帯を含む「実 SS において実地検証を行う」もの、「試験用 SS において検証を行う」もの、「シミュレーションで検証を行う」べきシナリオに分類し、試験方法を決定する。

3.2.3 監視スタッフ環境外乱シナリオ

監視スタッフ環境外乱シナリオでは、給油許可の可否判断を行う監視スタッフの状況に着目した検証を行う。給油中に突発的な外部環境からの外乱により監視スタッフの動作が変化しても事故を起こさないことを検証するため、フィールドでの直接視や店舗内モニターでの間接視といったスタッフの監視位置や従来の監視スタッフ単独での判断ではなく、AI からの情報も加味し安全を確認するといった、業務範囲等の集中力を阻害し得る外乱要因を整理し、監視スタッフの状況が変化しても危険が生じないか、それぞれの組み合わせを網羅するような試験シナリオを構築し、検証することとする。

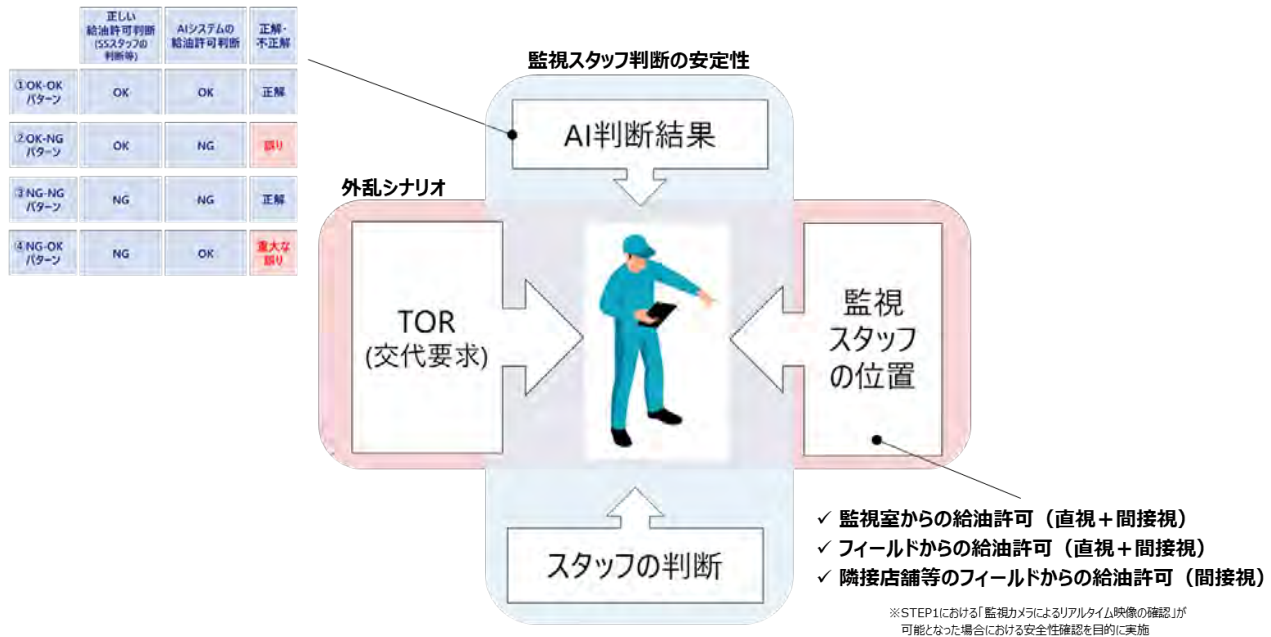


図 3-17 監視スタッフ環境外乱の評価イメージ

出所) NRI 作成

(1) 監視スタッフ環境外乱要因の要素抽出

給油許可環境外乱の要因は、「監視位置（監視方法）」、「監視スタッフにインプットされる AI 判断結果」の 2 つに大別できる。これらの 2 つの要因の組み合わせを網羅するような試験シナリオを構築し、そのシナリオ条件下において、AI の試験を実施して業務適合性を評価する。

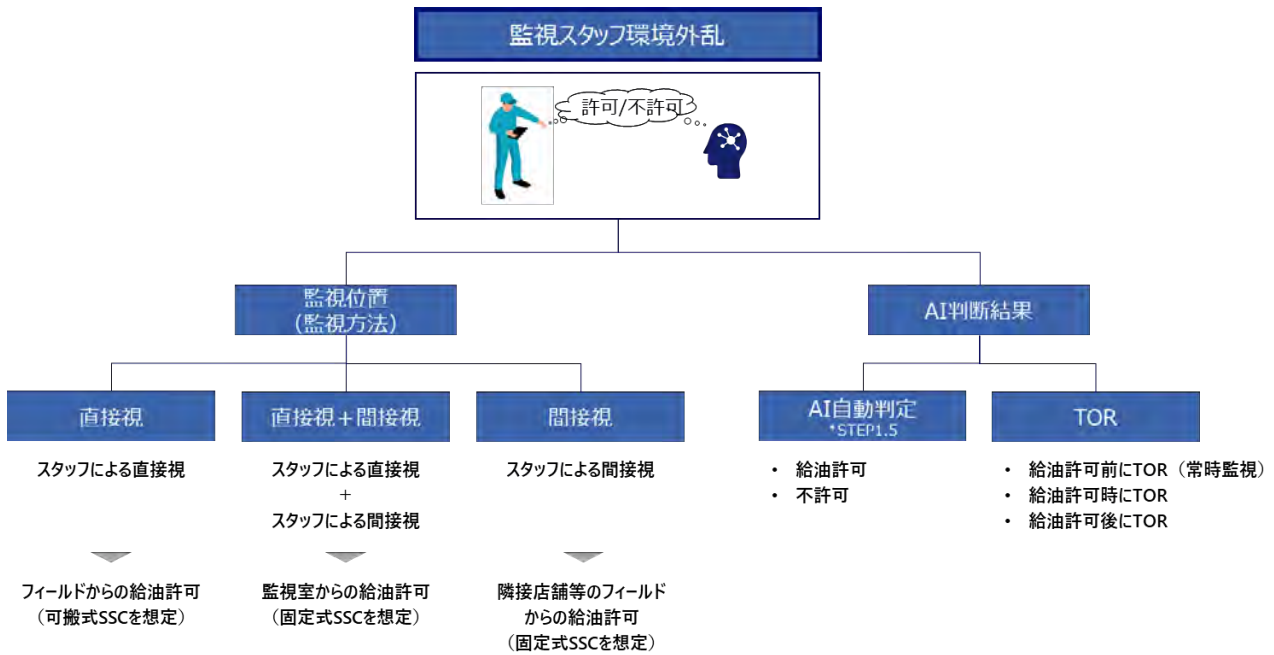


図 3-18 給油許可環境外乱の体系図

出所) NRI 作成

(2) 監視スタッフ環境外乱シナリオの選定

給油許可環境外乱の要因は、「監視位置（監視方法）」、「監視スタッフにインプットされるAI判断結果」の2つに大別できる。これらの2つの要因の組み合わせを網羅するような試験シナリオを構築し、そのシナリオ条件下において、AIの試験を実施して業務適合性を評価する。これらの2つの要因の組み合わせは以下となる。

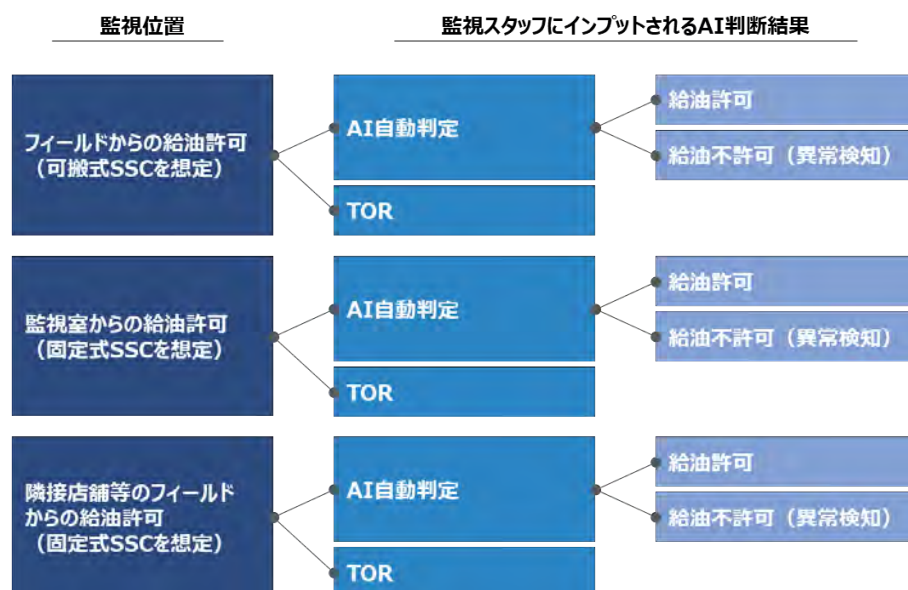


図 3-19 監視スタッフ環境外乱シナリオの組合せ

出所) NRI 作成

a. 監視位置

- ① 「フィールドからの給油許可」は、SS 敷地内の監視スタッフが適切に監視等を行うことができる範囲（以下、フィールドという）からの給油許可業務を指す。危険物の規制に関する規則の一部を改正する省令（令和元年総務省令第 67 号）が令和元年 12 月 20 日に公布され、顧客に自ら給油等をさせる給油取扱所において、可搬式の制御機器によっても給油許可等を行うことができるよう技術上の基準が整備された（令和 2 年 4 月 1 日施行）。そのため、フィールドから可搬式の制御端末を通じて監視スタッフに AI の判定情報を通知し、その結果と監視スタッフの直接視をもとに安全性を確認する状況下での試験を実施する。
- ② 「監視室からの給油許可」は、全ての顧客用固定給油設備等が直接視認できる位置に設置している制御卓（以下、監視室という）からの給油許可業務を指す。本ケースは可搬式の制御端末が導入されていない SS における一般的な状況となる。そのため、監視室から固定式制御端末を通じて監視スタッフに AI の判定情報を通知し、その結果と監視スタッフの直接視及び監視カメラモニタによる間接視をもとに安全性を確認する状況下での試験を実施する。
- ③ 「隣接店舗等のフィールドからの給油許可」は、SS 敷地内又は監視設備により直接視認と同等以上の安全性が確認できる箇所（監視室以外のフィールドや隣接店舗等、例えば SS 敷地内で可搬式の制御機器のモニタ等による間接視を通じた安全確認など）からの給油許可業務を指す。2022 年 2 月現在はセルフ給油取扱所の制御卓は全ての顧客用固定給油設備等の使用状況を直接視認できる位置に設置することとされているが、給油取扱所の経営の多角化及び監視設備の高性能化等を踏まえ、かつ、AI 等の新技術導入を見据えて、制御卓における視認を常時可能とするための監視設備が適切に設けられている場合には、顧客用固定給油設備等の使用状況を直接視認できる位置以外においても制御卓の設置を認める方針が検討されている。そのため、隣接店舗から固定式制御端末を通じて監視スタッフに AI の判定情報を通知し、その結果と監視カメラモニタによる間接視をもとに安全性を確認する状況下での試験を実施する。但し、上述のとおり 2022 年 2 月現在では規制緩和がされていないため、今後の法改正の状況を鑑み本試験環境の是非を適宜見直すこととする。

b. 監視スタッフにインプットされる AI 判断結果

- ① 「AI 自動判定」は、AI が監視カメラやセンサ等と連携し、給油を許可して良いと、又は正常な給油行動から逸脱し給油不可と判断した結果を指す。AI の判定結果は、可搬式の制御端末等を通じて監視スタッフに情報を通知する方法を想定している。STEP1.5 においては、特定条件下で監視スタッフの判断（直接視、間接視又はその両方を通じた確認）を介さずに、AI 単独での給油許可判断を可能とする。

- ②「TOR（Takeover request：交代要求）」は、監視カメラを通じて捉えたい対象物が視認できない（対象物が死角に入る等）、又はセンサ類の汚れや故障等によりAIが正常に安全性を評価出来ないケースにおいて、監視スタッフに状況確認を要求する状況を指す。AIが完全に機能しない状況においても適切に監視スタッフに給油監視業務が引き継がれ、従来通りの監視が可能であることを検証する。

(3) 検証観点／試験条件の定義

a. 前提事項

AIを活用した給油許可監視業務において、給油許可/給油不許可が取り得るパターンは①監視スタッフ：正しい判断、AI：正しい判断、②監視スタッフ：正しい判断、AI：誤った判断、③監視スタッフ：誤った判断、AI：正しい判断、④監視スタッフ：誤った判断、AI：誤った判断の4つがとり得る。本検討においては、従来の監視スタッフが担当している業務を、AI等の活用により機械が代替することを目的としているため、監視スタッフの判断は常に正しいという前提のもと検証を実施する。そのため、上記③と④の監視スタッフが誤った判断をするというシナリオは検討対象外とする。但し、今後AIの品質が向上し、従来の監視スタッフ以上の安全性が見込まれる場合は改めて試験シナリオの見直しを実施するものとする。

b. 試験シナリオの考え方

監視スタッフ環境外乱シナリオにおいて、給油許可監視業務の方法としては、「従来通り監視スタッフが実施」、「AIの安全性結果をもとに監視スタッフが実施」、「AI単独で実施」で実施が挙げられる。但し、「従来通り監視スタッフが実施」は従前実施している業務であるため検証対象外とし、給油許可監視業務を「AIの安全性結果をもとに監視スタッフが実施」、「AI単独で実施」する状況において、従来の監視スタッフが判断する場合と同等レベルの監視が可能であることを検証する。

	ヒト	ヒト+AI	AI
フィールドからの給油許可 (可搬式SSCを想定)	従来のスタッフによる 給油許可業務 *試験不要	スタッフ（AIはサポート） による給油許可業務	AIによる給油許可業務
監視室からの給油許可 (固定式SSCを想定)			
隣接店舗からの給油許可 (固定式SSCを想定)			

図 3-20 監視スタッフ環境外乱シナリオの観点

出所) NRI 作成

(4) 試験方法の定義

監視スタッフ環境外乱シナリオは、AIをどのように給油許可監視業務に適用するのか、AIを活用した給油許可監視業務はどのような業務であるべきか等の検証内容が含まれ、AIの導入先SSの業務設計や安全性に対する思想に依存する。実際の試験のオペレーションについては今後の調整事項となるが、業界としてのAI等を活用したセルフSSにおける給油許可監視業務のあるべき姿を定義のうえ、試験方法を決定する。

4 ガイドラインの活用イメージ

本章では、信頼性評価の具体的なステップに即し、AI を活用した給油許可監視システム（本章においても「AI システム」という）の開発・運用に係る各事業者の本ガイドラインの活用方法について整理する。また、AI システムの構築及び運用において、各担当が信頼性評価において具体的に実施する事項について整理するとともに、実施に際しての本ガイドラインの適用の具体的な流れや参照箇所を適示する。

4.1 ガイドライン活用の主体

本ガイドラインに基づく AI システムの信頼性の評価と運用は、複数の企業に跨る多様な主体により行われる。AI システムの開発・運用に係る意思決定は、例えば、現在先行して開発を進める石油元売会社を中心とするユーザー企業の経営企画担当の意思決定により行われ、事業企画担当らが具体的な AI システムの開発・運用に携わると考えられる。

具体的に想定される関与主体の役割分担を次の表 4-1 に示す。各担当は、それぞれ立場から AI システムの信頼性の確保に関与する。

表 4-1 AI システムにおける各担当の役割

担当者	企業	システムの信頼性評価・運用における役割
事業企画担当	ユーザー企業	システム開発プロジェクトの統括
品質保証担当	ユーザー企業 システムベンダー	システム全体の利用時品質の評価・確認
	AIベンダー	機械学習要素の外部品質・内部品質の評価・確認
現場担当	ユーザー企業 設備ベンダー	システムの利用時品質のレビュー 機械学習要素に求める外部品質のレビュー
システム担当	ユーザー企業 システムベンダー	システムの利用時品質や外部品質の設定 機械学習要素以外の構成要素や安全関連系の整理 機械学習要素の開発に係るレビュー・データ提供
機械学習開発担当	AIベンダー	機械学習要素の設計・開発

事業企画担当は、AI システム開発プロジェクトの主担当であり、予算やスケジュールに関する事項をはじめとしたプロジェクト全体を統括する。主なユーザーとなる石油元売会社においては、通常、SS 拠点開発関係の部署の担当者が担う。

品質保証担当は、AI システム全体の品質の評価・確認を行う者であり、利用時品質の評価・確認についてはユーザー企業やシステムベンダーが、外部品質・内部品質の評価・確認については AI ベンダーが行うことを想定する。品質管理担当は、AI システム開発を行うシステム担当や機械学習開発担当から、独立性を有していることが必要である。

現場担当は、AI システムの運用を直接担う者であり、システム全体及び機械学習要素の品質に対してレビューを行う。レビューとは、本ガイドラインを用いて信頼性評価を行う主担当から依頼を受け、自身の業務所掌・専門性に基づいて主担当による検討事項を確認することを意味する。給油設備に関連したデー

タ提供については、給油設備のベンダー企業が参画する場合も想定される。

システム担当は、AI システムの利用時品質・外部品質の設定や機械学習要素の開発に係るレビュー・データ提供を行う者であり、システムのベンダー企業を中心に、ユーザー企業も本役割を担うことが想定される。

機械学習開発担当は、機械学習要素の設計・開発を担う者であり、AI ベンダー企業が本役割を担うことが想定されるが、システムのベンダー企業も参画する場合がある。機械学習要素の品質保証を行う担当者も、広義の機械学習開発担当に包含される。

なお、表 4-1 に掲げた各担当に該当する企業は例示であり、プロジェクト内での実際の役割は、AI システム開発のプロジェクトによって変動する。主担当者としてレビュー担当者の役割を適切に割り振り、こういった担当者が関与して品質を設定したかを明らかにすることが重要である。

次に、各担当者の役割に応じた、ガイドラインの活用場面と重点的に確認すべき品質を表 4-2 に示す。各担当者は、自身が確認すべき品質に関して、該当するガイドラインの内容を理解する必要がある。

表 4-2 各担当におけるガイドラインの活用場面と確認対象品質

担当者	ガイドラインの活用場面	利用時品質	外部品質	内部品質
事業企画担当	目的・利用時品質の設定 経営企画への説明	●	△	△
品質保証担当	利用時品質・外部品質・ 内部品質の確認	●	●	●
現場担当	利用時品質・外部品質のレビュー	△	△	△
システム担当	利用時品質・外部品質の設定 内部品質に係るデータ提供	●	●	●
機械学習開発担当	内部品質の要求事項の把握 機械学習要素の設計・開発	—	△	●

【凡例】 ● 信頼性評価の手順をガイドラインに則って実施

△ 主担当者からの依頼に応じ信頼性評価に関与
ガイドラインの参照は任意

4.2 ガイドライン活用の流れ

ここでは、要求・要件定義から検収に亘る AI システムの開発及び運用の各フェーズにおける、各担当の役割とガイドライン活用の流れを整理する。

4.2.1 フェーズ別実施項目の全体像

フェーズ別の各担当の役割を整理したものが、次の表 4-3 である。

PoC（Proof of Concept／概念実証）の段階では、各担当の実施項目は本ガイドラインで規定されないが、要求・要件定義以降の各フェーズの実施事項を参考に、開発のコアメンバーで各種の機能が担われることが期待される。

表 4-3 フェーズ別実施項目の全体像

担当者	Ph.0 PoC	Ph.1 要求・ 要件定義	Ph.2 設計	Ph.3 実装	Ph.4 試験・検収	Ph.5 運用
事業企画担当	各フェーズの実施項目を参考に 開発コアメンバーで検討	システムの目的設定 ／機能要件・利用 時品質のレビュー	—	—	検収における品質 全体のレビュー	—
品質保証担当		—	—	—	試験を踏まえた 外部品質の評価／ システムの検収	利用時品質・外部 品質・内部品質の 確認
現場担当		システムの目的・ 機能要件・利用時 品質のレビュー	外部品質の設定・ レベル設定のレビュー ／安全関連系の レビュー	機械学習要素の 開発に係るレビュー	検収における品質 全体のレビュー	利用時品質の確認 結果のレビュー／ システム更新に係る レビュー・データ提供
システム担当		機能要件・利用時 品質の設定	外部品質の設定・ レベル設定／機械 学習要素以外の 要素整理	機械学習要素の 開発に係るレビュー	試験を踏まえた 外部品質の評価 結果のレビュー	外部品質の確認 結果のレビュー／ 機械学習要素以外 の構成要素の更新
機械学習 開発担当		—	—	内部品質のレベル 設定／機械学習 要素の設計・開発	—	内部品質に係る 機械学習要素の 更新

Phase1. 要求・要件定義フェーズでは、事業企画担当がシステムの目的の設定を行った後、システム担当が機能要件や利用時品質の設定を行い、事業企画担当及び現場担当がシステムの機能要件や利用時品質のレビューを行う。

Phase2. 設計フェーズでは、システム担当が外部品質の設定と外部品質のレベル設定、安全関連系等の機械学習要素以外の要素整理を行い、これを現場担当がレビューする。

Phase3. 実装フェーズでは、機械学習開発担当が内部品質のレベル設定や機械学習要素の設計・開発を行い、現場担当やシステム担当がこれに対するレビューを行うとともに、機械学習要素の開発における品質保証を品質保証担当が実施する。

Phase4. 試験・検収フェーズでは、品質保証担当が試験を踏まえた外部品質の評価を実施し、システム担当が当該評価結果に対するレビューを行う。また、事業企画担当や現場担当は、検収における品質全体のレビューを行う。

Phase5. 運用フェーズでは、各担当が利用時品質・外部品質・内部品質の確認や必要に応じたシステムの更新作業を実施する。

なお、本ガイドラインはウォーターフォール型¹⁰の開発に限定して適用されるものではなく、各フェーズの実施事項を反復的に実施しながら信頼性評価のシステム構築を行う、アジャイル型¹¹の案件にも適用可能なものである。

¹⁰ ウォーターフォール型：要件定義からリリースまでの工程を順に実施する開発プロセス

¹¹ アジャイル型：要件定義～リリースの工程を反復し、機能の追加・補正を逐次実施する反復増加型の開発プロセス

4.2.2 フェーズ別実施項目の詳細

ここでは、システムの信頼性評価に係る、フェーズ別の実施事項の詳細について整理する。なお、先述の通り役割分担は、実際のシステム開発のプロジェクトによって変動する。

表 4-4-1 各ステップにおける各担当者の実施内容

ステップ		実施者	実施内容
Phase 0 PoC			
Phase 1 要求・要件定義	1-1 システムの目的の設定	事業企画担当	現状の課題を整理したうえで、「システムの目的」を設定
		現場担当	課題や目的に対し、現場の立場からレビューを実施
	1-2 システムの機能要件・利用時品質の設定	システム担当	システムの機能要件と利用時品質を設定
		事業企画担当	機能要件と利用時品質に対し、企画者の立場からレビューを実施
Phase 2 設計	2-1 外部品質の設定	システム担当	利用時品質を日前、機械学習要素の外部品質を設定
		現場担当	設定された外部品質の具体化内容に対し、現場の立場からレビュー
	2-2 外部品質のレベルに係るシステムと安全関連系の整理	システム担当	システムの構成要素や安全関連系の有無や機能要件を整理
		現場担当	構成要素や安全関連系に対し、現場の立場からレビュー
	2-3 外部品質のレベル設定	システム担当	機械学習要素の外部品質のレベル設定
		現場担当	外部品質のレベルに対し、現場の立場からレビュー
Phase 3 実装	3-1 内部品質のレベル設定	機械学習開発担当	外部品質のレベルに応じて、機械学習要素の内部品質のレベルを設定
	3-2 機械学習要素の設計・開発	機械学習開発担当	機械学習要素の設計と、内部品質の要求事項に応じた開発
		現場担当	現場の立場からデータを提供
	3-3 機械学習要素以外の構成要素の開発	システム担当	データ管理の立場からデータを提供
		システム担当	設定した外部品質・利用時品質に必要な構成要素を開発

各フェーズで実施するステップと担当者別の役割を次の表 4-4-1 及び 4-4-2 で整理した。担当別の具体的な実施内容が設定されるのは、Phase 1 の「要求・要件定義」以降であるが、これに先立って行われる PoC においても、信頼性評価の観点が考慮されることが望ましい。

なお、各品質の定義は既に本ガイドラインで定義されたものを活用することができる。ユーザー企業は、自社が利用する AI システムのロードマップ上の STEP の位置付けを踏まえて、詳細の設定を行うこととする。

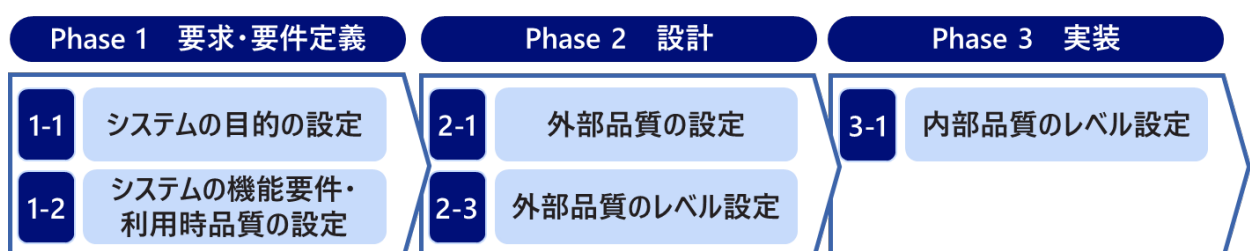
表 4-4-2 各ステップにおける各担当者の実施内容（続）

	ステップ		実施者	実施内容
Phase 4 試験・検収	4-1	外部品質の評価試験	品質保証担当	システムの試験を実施し、要求水準を充足しているかを評価
			システム担当	評価の結果を踏まえ、外部品質やそのレベルを設定した立場からレビュー
	4-2	検収	品質保証担当	試験結果からシステム全体の評価を行い、基準の充足を確認して検収
			事業企画担当	システム全体の評価について、企画者の立場からレビューを実施
			現場担当	システム全体の評価について、現場の立場からレビューを実施
Phase 5 運用	5-1	利用時品質の確認	品質保証担当	運用中のシステムの利用時品質を確認
			現場担当	利用時品質の確認結果に対し、現場の立場からレビューを実施
	5-2	外部品質の確認	品質保証担当	利用時品質の確認結果を踏まえて、外部品質を確認
			システム担当	外部品質の確認結果に対し、品質やそのレベルを設定した立場からレビュー
	5-3	内部品質の確認	機械学習開発担当	内部品質の要求事項の対応状況を確認
			機械学習開発担当	外部品質・内部品質の確認結果に応じ、機械学習要素を更新
	5-4	システムの更新	システム担当	システム管理の立場から、機械学習要素の更新に対しデータを提供
			現場担当	機械学習要素の更新において、現場の立場からレビューを実施

(0) PoC (Proof of Concept)

PoC は、主として技術的な観点から、システムの有効性や実現性を調査・検証する工程である。すべての場合に信頼性評価の検討が要求されるものではないが、本格的なシステム開発を見据える段階においては、開発への円滑な移行の準備として、各フェーズにおける信頼性評価の実施事項を予め考慮することが有効である。

具体的には、システム開発のコアメンバーが中心となって、次の図 4-1 で示す実施項目について、PoC の終了段階で目途をつけておくことが望ましい。具体的なシステム・設備が未確定な PoC の段階で、厳密



な外部品質のレベル設定を行うことは困難であるが、機能要件や利用時品質に照らして、外部品質の要求レベルや内部品質の方針の見通しを検討しておくことが推奨される。

(1) Phase1. 要求・要件定義

要求・要件定義フェーズにおいては、ユーザー企業が中心となり、セルフ SS における一連の給油許可監視の課題を特定し、それを解決するシステムの機能要件や利用時品質を設定する。

1-1. システムの目的の設定

事業企画担当は、給油許可監視における課題を整理するとともに、「システムの目的」を設定し、機械学習要素をどのように活用するかを判断する。

これに対し、**現場担当**は現場の管理担当者や従業員等からの意見を踏まえ、現場またはユーザーの立場からシステムの目的に対するレビューを行い、必要に応じて方向性を修正する。

1-2. システムの機能要件・利用時品質の設定

システム担当は、本ガイドラインの「2.1.1 利用時品質」及び「2.1.3 利用時品質及び外部品質における評価軸」を参照して利用時品質の位置づけに関する理解を深めた後、「2.2.1 利用時品質の設定」を参照して利用時品質を設定する。

事業企画担当及び**現場担当**は、システム担当と同じガイドラインの項目を適宜参照しつつ、企画者または現場・ユーザーの立場からレビューを実施する。

(2) Phase2. 設計

設計フェーズにおいては、機械学習を利用したシステム全体に求められる利用時品質に基づき、機械学習要素に係る外部品質や機械学習要素以外のシステムの構成要素に求められる要件を設定する。

なお、運用フェーズで実施する各品質の確認のために、本フェーズにおいて予め運用フェーズで求められるモニタリングの頻度や実施基準を検討しておくことが重要である。

2-1. 外部品質の設定

システム担当は、本ガイドラインの「2.1.2 外部品質」及び「2.1.3 利用時品質及び外部品質における評価軸」を参照して外部品質の位置づけに関する理解を深めた後、「2.2.2 外部品質の設定」を参照して外部品質を設定する。

外部品質設定の段階では、機械学習特有の数値目標（誤検知率 X%未満等）を定める必要はない。リスク回避に係る数値目標に関しては、後段の「外部品質のレベル設定」で定める場合がある他、最終的には機械学習要素の設計・開発段階で、PoC の成果や学習状況に応じた数値目標を設定する。

現場担当は、システム担当と同じガイドラインの項目を適宜参照しつつ、設定された外部品質に対し、現場・ユーザーの立場からレビューを実施する。

2-2. 外部品質のレベルに係るシステムと安全関連系の整理

システム担当は、本ガイドラインの「2.2.2 外部品質の設定」を参照して、外部品質のレベル設定の考え

方を確認するとともに、安全関連系の有無や機能要件を整理する。

現場担当は、必要に応じてシステム担当と同じガイドラインの項目を参照しつつ、現場や安全管理の立場からレビューを実施する。

2-3. 外部品質のレベル設定

システム担当は、本ガイドラインの「2.2.2 外部品質の設定」を参照して、外部品質のレベル設定の方法を理解したうえで、外部品質や構成要素（外部安全機構等）を踏まえ外部品質のレベル設定を行う。「リスク回避性」については、AISL 表を活用した評価における人間による回避可能性を「機械学習要素の判断結果が運転・保全にそのまま反映されることはなく、人間による確認・代替システム適用を介した結果が反映される」と定め、「パフォーマンス」については、システムは概ね AIPL (AI Performance Level) 2 を充足することを基本とする。レベル設定の詳細については、本ガイドラインの「2.2.2 外部品質の設定」を参照されたい。

現場担当は、必要に応じてシステム担当と同じガイドラインの項目を参照しつつ、外部品質のレベル設定に対し、現場の立場からレビューを実施する。

(3) Phase3. 実装

システムの実装フェーズでは、内部品質の実現を通じた機械学習要素の開発及び機械学習要素以外の構成要素の開発を実施する。

3-1. 内部品質のレベル設定

機械学習開発担当は、本ガイドラインの「2.1.4 内部品質」で内部品質の 8 つの軸の内容を理解するとともに、「2.2.3 内部品質の設定」を参照して、内部品質のレベルの設定方法を理解する。機械学習要素の外部品質のレベルに応じて、機械学習要素の内部品質のレベルを設定する。

3-2. 機械学習要素の設計・開発

機械学習開発担当は、内部品質のレベル設定における要求事項と観点に応じて、機械学習要素の具体的な仕様やモデルの特定を行い、設計・開発を実施する。本ステップにおいて、PoC の成果やデータの取得状況・学習状況等に応じて、具体的な機械学習特有の数値目標を設定することができる。

システム担当及び**現場担当**は、それぞれの立場からレビューやデータ提供を通じて支援を行う。

3-3. 機械学習要素以外の構成要素の開発

システム担当は、利用時品質を達成するための機械学習要素以外のシステムの構成要素の開発を実施する。機械学習要素には直接関わらないステップであるため、信頼性評価に係るガイドラインの参照箇所は存在しない。

(4) Phase4. 試験・検収

試験・検収フェーズでは、**品質保証担当**が中心となってシステムの試験及び評価を実施し、その評価に基づいて検収を実施する。機械学習要素の内部品質は、ステップ 3-2「機械学習要素の設計・開発」で既

に確認されているため、ここでは機械学習要素の外部品質の評価を中心に実施する。検収においては、利用時品質を含む各種の要求水準を充足しているかが確認される。

4-1. 外部品質の評価試験

品質保証担当は、本ガイドラインの「2.2.2 外部品質の設定」を参照して外部品質の評価基準と方法を理解し、機械学習要素の外部品質の設定内容を踏まえ、外部品質に係る試験結果が要求水準に達しているかを評価する。

システム担当は、品質保証担当と同じガイドラインの項目を適宜参照しつつ、外部品質及びそのレベル設定並びに安全関連系の整理を実施した立場からレビューを実施する。

4-2. 検収

品質保証担当は、外部品質に係る評価試験の結果を踏まえつつ、システム全体の評価を実施し、その評価結果を踏まえ、要求水準を充足するシステムであれば検収を行う。検収は各ユーザー企業の基準に基づいて行われることが想定されるが、本ガイドラインに定める外部品質のレベルや、それに応じて設定された内部品質の要求事項及びレベル、評価試験の結果等を適切に踏まえる必要がある。

事業企画担当はシステム全体の評価に対し、企画者の立場からレビューを実施するとともに、**現場担当**は利用時品質の評価結果に対し、現場或いはユーザーの立場からレビューを実施する。

(5) Phase5. 運用

運用フェーズでは、必要に応じてシステムの更新が実施されることが想定される。更新時に品質検査を実施するかどうかに関わらず、システム全体の利用時品質及び外部品質のモニタリングを適切に実施する必要がある。

モニタリング結果の判断基準等、モニタリングに係る枠組みについては、設計フェーズの時点で用意することが必要になる場合が多い。これらの状況を踏まえ、運用フェーズにおける実施項目は、予め設定した頻度や実施基準に基づいて実施する品質保証の活動について記載する。

5-1. 利用時品質の確認

品質保証担当は、システム構築時の機能要件・利用時品質等を踏まえ、運用中のシステムの利用時品質が当初の目的を達成できているか確認する。必要に応じて、「2.1.1 利用時品質」や「2.1.3 利用時品質・外部品質の軸」を参照して利用時品質の位置づけを理解するとともに、「2.2.1 利用時品質の設定」を参照し、利用時品質の設定方法を理解する。

現場担当は、品質保証担当と同じガイドラインの項目を適宜参照しつつ、利用時品質の確認結果に対し、現場の立場からレビューを実施する。

5-2. 外部品質の確認

品質保証担当は、「2.2.2 外部品質の設定」を参照して外部品質の評価基準と評価方法を理解したう

えて、システムの機能要件や利用時品質、機械学習要素の外部品質の設定等を踏まえ、運用中の機械学習要素の外部品質が当初設定した要求水準を充足できているか確認する。

システム担当は、品質保証担当と同じガイドラインの項目を適宜参照しつつ、外部品質及びそのレベル設定並びに安全関連系の整理を実施した立場からレビューを実施する。

5-3. 内部品質の確認

機械学習開発担当は、「運用時品質の維持性」に着目し、運用中のシステムにおける内部品質の要求事項への対応状況を確認する。必要に応じ、「2.2.3 内部品質の設定」を参照する。

5-4. システムの更新

機械学習開発担当は、「2.2.3 内部品質の設定」を踏まえた外部品質及び内部品質のモニタリング結果に応じて、機械学習要素の更新を適宜実施する。なお、更新したシステムの外部品質は、試験・検収フェーズに戻り確認することになる。

また、**システム担当**は、機械学習要素以外の構成要素を更新するとともに、システム管理の立場から機械学習要素の設計・開発に対しデータ提供による支援を行う。

5 附録

資料名		概要
本紙	セルフSSにおけるAIによる給油許可監視の実装に向けたAIシステム評価方法等に係るガイドライン（本資料）	セルフSSにおける給油許可監視システムの実装にあたり、機能要件や導入方法等、システムの信頼性評価に関する枠組みを定める。
別表1	内部品質の要求事項	セルフSSにおける給油許可監視業務にAIを適用する際に、AIとして具備すべき品質要求を定める。
別表2	認識外乱要因一覧	セルフSSにおける給油許可監視システムの妥当性評価を目的とした試験シナリオのうち、システムの構成要素であるAIの認識精度に影響を及ぼす代表的な外乱要因を定める。
別表3	認識外乱試験シナリオ詳細	試験シナリオの1つである認識外乱シナリオの検証観点、試験環境・条件、実施方法を定める。
別表4	認識外乱シナリオ要素一覧	試験シナリオの1つである認識外乱シナリオの試験環境の要素を定める。

6 参考文献

- 「プラント保安分野 AI 信頼性評価ガイドライン 第 2 版」, 石油コンビナート等災害防止 3 省連絡会議, 2021 年 3 月.
URL : <https://www.meti.go.jp/press/2020/03/20210330002/20210330002.html>
- 「機械学習品質マネジメントガイドライン 第 2 版」, 産業技術総合研究所, 2021 年 7 月.
URL : <https://www.digiarc.ait.go.jp/publication/aiqm/AIQM-Guideline-2.1.0.pdf>
- 「オーナーカー自動運転の安全性評価の取り組み」, 未来を変える自動運転ショーケース／社会受容性シンポジウム), 一般社団法人日本自動車工業会自動車部会 横山利夫, 2021 年 3 月.
URL : https://www.meti.go.jp/policy/mono_info_service/mono/automobile/Automated-driving/jikoukai.pdf
- 自動運転の安全性評価フレームワーク v1.0, 一般社団法人日本自動車工業会 AD 安全性評価分科会, 2020 年 10 月.
URL : https://www.jama.or.jp/safe/automated_driving/pdf/framework.pdf
- 「自動走行の実現及び普及に向けた取組報告と方針」Version 5.0, 一般社団法人日本自動車工業会 自動走行ビジネス検討会, 2021 年 4 月.
URL : https://www.meti.go.jp/shingikAI/mono_info_service/jido_soko/pdf/20210430_02.pdf
- Human Error. J. Reason. Cambridge University Press. Cambridge. 1990.
- Torino Impact Hazard Scale. Center for Near Earth Object Studies. NASA.
URL : https://cneos.jpl.nasa.gov/sentry/torino_scale.html