プラント保安分野のAI信頼性評価　実施内容記録フォーマット

【本記録フォーマットの目的・効果】

本記録フォーマットは、様々な情報が掲載されている「プラント保安分野AI信頼性評価ガイドライン」を実際のAI開発・運用で適切に活用いただくことを支援する目的で整備する。

具体的な目的や効果は以下を想定している。

- AI開発・運用の担当者が本フォーマットに従うことで、ガイドライン適用に苦慮する時間を短縮し、ガイドラインの適用を効率的・効果的に行える
- AI開発・実装時の担当者や考え方やプロセスを記録することで、運用開始後のモニタリングや更新時に開発・実装時の事情を理解でき、また不明な点がある場合でも確認しやすくする
- 社内外のステークホルダーに対して、ガイドラインに則ってAIを開発したエビデンスとして利用できる

【本記録フォーマットの使用方法】

1 「ガイドライン」本編のエグゼクティブサマリー、１章を確認する。

2 「ガイドライン」本編の４章で具体的な適用方法を確認する。

3 メインフォーマットで、上から順番に「ガイドライン」４．２でA列記載のステップ番号に沿って手順を確認し
「ガイドライン」の参照先を確認しながら実施する。実施した結果は、メインフォーマットの空欄に記載する。

4 運用開始後は、ステップ11～14で品質の確認とシステムの更新、ステップ9で更新後の品質の確認の記録を行う。

※ステップ7・13・14で行う内部品質の確認にはサブフォーマットを用いる。「サブフォーマット説明」に利用方法、
「データの設計」「データの確認」「学習」「実装・運用」及び各「記入例」に具体的なフォーマットと記入例を示す。

ステップ	実施項目	概要	ガイドライン	実施担当者	実施担当者名	実施結果				備考・参考資料	レビュー担当者	レビュー担当者名	レビュー結果
1	システムの目的の設定	改善したい「現状の課題」を整理、「システムの目的」を設定し、機械学習利用システムが必要か判断	-	事業企画担当		現状の課題					現場担当		
						システムの目的							
2	システムの機能要件・利用時品質の設定	機械学習利用システムの機能要件と利用時品質を設定	2.1.1 2.1.3 2.2.1 (3.3)	プラントシステム担当		機能要件					事業企画担当 現場担当		
						利用時品質	リスク回避性 ※3件以上ある場合は行を追加。	i					
								ii					
								iii					
							パフォーマンス ※3件以上ある場合は行を追加。	i					
								ii					
								iii					
3	外部品質の設定	利用時品質を踏まえ、機械学習要素の外部品質を設定	2.1.2 2.1.3 2.2.2 (3.3)	プラントシステム担当		外部品質	リスク回避性 ※3件以上ある場合は行を追加。	i			現場の管理担当		
								ii					
								iii					
							パフォーマンス ※3件以上ある場合は行を追加。	i					
								ii					
								iii					
4	外部品質のレベルに係る機械学習利用システムと安全関連系の整理	機械学習利用システムの各構成要素（機械学習要素、機械学習要素以外の構成要素（外部安全機構等））、同システムと独立した安全関連系の有無や機能要件を整理	2.2.3	プラントシステム担当		機械学習利用システムの機械学習要素以外の構成要素（外部安全機構）					現場の管理担当 環境安全担当		
						機械学習利用システムから独立した安全関連系							
5	外部品質のレベル設定	機械学習要素の外部品質のレベルの設定	2.2.3	プラントシステム担当		外部品質のレベル	リスク回避性（AISL） ※3件以上ある場合は、20～24行をコピーして行を追加。	i	SIL評価の必要性の有無		現場の管理担当		
									（SIL評価を行う場合）	SIL評価結果			
									（SIL評価を行わない場合、もしくはSILなしの場合）	影響の強度（Ⅰ～Ⅳ） 人間による回避可能性（①～③）			
									AISL				
								ii	SIL評価の必要性の有無				
									（SIL評価を行う場合）	SIL評価結果			
									（SIL評価を行わない場合、もしくはSILなしの場合）	影響の強度（Ⅰ～Ⅳ） 人間による回避可能性（①～③）			
									AISL				
								iii	SIL評価の必要性の有無				
									（SIL評価を行う場合）	SIL評価結果			
									（SIL評価を行わない場合、もしくはSILなしの場合）	影響の強度（Ⅰ～Ⅳ） 人間による回避可能性（①～③）			
									AISL				
							パフォーマンス（AIPL） ※3件以上ある場合は、35～37行をコピーして行を追加。	i	性能の水準（高/中/低）				
									水準に対する要求の厳しさ（mandatory/best-effort/なし）				
									AIPL				
								ii	性能の水準（高/中/低）				
									水準に対する要求の厳しさ（mandatory/best-effort/なし）				
									AIPL				
								iii	性能の水準（高/中/低）				
									水準に対する要求の厳しさ（mandatory/best-effort/なし）				
									AIPL				
6	内部品質のレベル設定	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	2.1.4 2.2.4 附録（チェックリスト） (3.3)	機械学習設計・開発担当		「データセットの均一性」以外の内部品質のレベル（外部品質のレベルで最も高いレベルに則して設定）							
						「データセットの均一性」の内部品質のレベル	リスク回避性						
							パフォーマンス						
7	機械学習要素の設計・開発	機械学習要素の設計、内部品質の要求事項と観点に応じた開発	2.2.5 附録（チェックリスト）	機械学習設計・開発担当		内部品質	※別紙、「要求事項・観点に関する対応方針」「（開発・実装時）対応の記録」に記入				現場担当 プラントシステム担当		
		機械学習要素の開発において、設定した内部品質のレベルと観点を満たしているか品質保証の立場から確認		機械学習品質保証担当		内部品質	※別紙、「要求事項・観点に関する対応方針」「（開発・実装時）対応の記録」に記入				現場担当 プラントシステム担当		
8	機械学習要素以外の構成要素（外部安全機構等）の開発	設定した外部品質・利用時品質に必要な機械学習要素以外の構成要素（外部安全機構等）を開発	-	プラントシステム担当		機械学習利用システムの機械学習要素以外の構成要素 （機械学習要素ではないため当紙フォーマットでの整理は不要）							
9	テスト（外部品質の評価）	機械学習利用システムのテストを行い、その結果を評価。機械学習要素の外部品質が要求水準に達しているか評価	2.2.3	システム品質保証担当		外部品質の評価	リスク回避性 ※3件以上ある場合は行を追加。	i			プラントシステム担当 環境安全担当		
								ii					
								iii					
							パフォーマンス ※3件以上ある場合は行を追加。	i					
								ii					
								iii					
10	検収	テストでの外部品質の評価結果から機械学習利用システム全体の評価を行い、それらの評価結果等を踏まえ基準を満たす機械学習利用システムであれば検収	-	システム品質保証担当		システム全体の評価（機械学習要素の信頼性評価以外も踏まえて判断）					事業企画担当 現場担当		
11	利用時品質の確認	運用中の機械学習利用システムの利用時品質を確認	2.1.1 2.1.3 2.2.1	システム品質保証担当		利用時品質	リスク回避性 ※3件以上ある場合は行を追加。	i			現場担当		
								ii					
								iii					
							パフォーマンス ※3件以上ある場合は行を追加。	i					
								ii					
								iii					
12	外部品質の確認	運用中の機械学習利用システムの利用時品質の確認結果も踏まえ、外部品質を確認	2.2.3	システム品質保証担当		外部品質	リスク回避性 ※3件以上ある場合は行を追加。	i			プラントシステム担当		
								ii					
								iii					
							パフォーマンス ※3件以上ある場合は行を追加。	i					
								ii					
								iii					
13	内部品質（運用時品質の維持性）の確認	運用中の機械学習利用システムの利用時品質・外部品質の確認結果も踏まえ、内部品質の要求事項と観点の対応状況を確認	2.2.4 2.2.5 附録（チェックリスト）	機械学習品質保証担当		内部品質	※別紙、運用時品質の維持性の「（運用時）対応の記録」に記入						
14	システムの更新	外部品質・内部品質の確認結果に応じて機械学習要素を更新	2.2.4 2.2.5 附録（チェックリスト）	機械学習設計・開発担当		内部品質	※別紙、運用時品質の維持性の「（運用時）対応の記録」に記入				現場担当 プラントシステム担当		
		利用時品質・外部品質の確認結果に応じて機械学習要素以外の構成要素（外部安全機構等）を更新、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提供	-	プラントシステム担当		機械学習利用システムの機械学習要素以外の構成要素							

ステップ	実施項目	概要	ガイドライン	実施担当者	実施担当者名	実施結果						備考・参考資料	レビュー担当者	レビュー担当者名	レビュー結果
1	システムの目的の設定	改善したい「現状の課題」を整理、「システムの目的」を設定し、機械学習利用システムが必要か判断	-	事業企画担当		現状の課題				異常の発生により設備が突然的に停止することで、設備稼働率低下		飛騨担当			問題なし
						システムの目的				設備の異常予兆を検知し、事前に交換対応等を行うことで設備稼働率を向上させる。					問題なし
2	システムの機能要件・利用時品質の設定	機械学習利用システムの機能要件と利用時品質を設定	2.1.1 2.1.3 2.2.1 (3.3)	プラントシステム担当		機能要件				異常予兆検知時に発報する		事業企画担当 飛騨担当			問題なし
										異常箇所、異常の深刻度、異常と関連のある変数を出力する					問題なし
						利用時品質	リスク回避性 ※3年以上ある場合は付与推奨。	i		様々なプラントの条件下で、将来の異常の発生を正しく検知する					現時点では「様々なプラントの条件下」とするが、追って具体的な条件についてはデータ数等可能な期間を踏まえ定義が必要。
								ii		様々なプラントの条件下で、異常発生箇所・異常の深刻度・異常と関連のある変数を正しく出力する。					同上
								iii							
							パフォーマンス ※3年以上ある場合は付与推奨。	i		発報内容の種類に運転員や点検要員の時間を多く割かなくてよい程度の発報頻度とする。 発報を受けてから原因調査や対応が可能な程度に十分に早いタイミングで発報する。					問題なし
								ii							問題なし
								iii							
3	外部品質の設定	利用時品質を踏まえ、機械学習要素の外部品質を設定	2.1.2 2.1.3 2.2.2 (3.3)	プラントシステム担当		外部品質	リスク回避性 ※3年以上ある場合は付与推奨。	i		「異常予兆あり」の場合に、「正常」であると判定する誤判定率を限りなく小さくする。		飛騨の管理担当			問題なし
								ii		「異常予兆あり」の場合に、異常発生箇所・異常の深刻度・異常と相関のある変数を誤って出力する誤認識率を限りなく小さくする。					問題なし
								iii							
							パフォーマンス ※3年以上ある場合は付与推奨。	i		誤検知の数をできるだけ少なくする。					問題なし
								ii		決められた期間までに検出する					問題なし
								iii							
4	外部品質のレベルに係る機械学習利用システムの安全関連系との関係	機械学習利用システムの各構成要素（機械学習要素、機械学習要素以外の構成要素（外部安全機構等））、同システムと独立した安全関連系の有無や機能要件を整理	2.2.3	プラントシステム担当		機械学習利用システムの機械学習要素以外の構成要素（外部安全機構）				なし。		飛騨の管理担当 環境安全担当			異常予兆検知結果は作業員が確認し、対応するため、当該システムに新たな外部安全機構は設置しないとしたことで承知。 異常発生時のインターロック機能は却下。異常発生箇所の管理は現状人間が実施。
						機械学習利用システムから独立した安全関連系				既設のインターロック装置あり。					
5	外部品質のレベル設定	機械学習要素の外部品質のレベルの設定	2.2.3	プラントシステム担当		外部品質のレベル	リスク回避性（AISL） ※3年以上ある場合は、20～24ヶ月で1～2ヶ月程度。	i	SIL評価のお要性の有無 （SIL評価を行う場合） （SIL評価を行わない場合、もしくはSILなしの場合）	SIL評価結果 影響の強度（I～IV） 人間による回避可能性（1～3）	ioの 1 1 1	レベル評価検討_1_202XDXXX.xlsx			問題なし
									AISL		1				
								ii	SIL評価のお要性の有無 （SIL評価を行う場合） （SIL評価を行わない場合、もしくはSILなしの場合）	SIL評価結果 影響の強度（I～IV） 人間による回避可能性（1～3）	ioの ioと III ③	レベル評価検討_2_202XDXXX.xlsx			問題なし
									AISL		0.1				
								iii	SIL評価のお要性の有無 （SIL評価を行う場合） （SIL評価を行わない場合、もしくはSILなしの場合）	SIL評価結果 影響の強度（I～IV） 人間による回避可能性（1～3）					
									AISL						
							パフォーマンス（AIPL） ※3年以上ある場合は、2年～3ヶ月で1～2ヶ月程度。	i	性能の水準（高/中/低） 水準に対する要求合致必 （mandatory/best-effort/なし）		① best-effort	レベル評価検討_3_202XDXXX.xlsx			問題なし
									AIPL		1				
								ii	性能の水準（高/中/低） 水準に対する要求合致必 （mandatory/best-effort/なし）		高 best-effort	レベル評価検討_4_202XDXXX.xlsx			問題なし
									AIPL		2				
								iii	性能の水準（高/中/低） 水準に対する要求合致必 （mandatory/best-effort/なし）						
									AIPL						
6	内部品質のレベル設定	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	2.1.4 2.2.4 別録（チェックリスト） (3.3)	機械学習設計・開発担当		「データセットの均一性」以外の内部品質のレベル（外部品質のレベルで最も高いレベルに則して設定）				3					
						「データセットの均一性」の内部品質のレベル	リスク回避性			LvS2					
							パフォーマンス			LvE2					
7	機械学習要素の設計・開発	機械学習要素の設計、内部品質の要求事項と観点に応じた開発	2.2.5	機械学習設計・開発担当		内部品質	※別紙、「要求事項・観点に関する対応方針」 「（開発・実装時）対応の記録」に記入						飛騨担当 プラントシステム担当		
		機械学習要素の開発に際して、設定した内部品質のレベルと観点を満たしているか品質保証の立場から確認	別録（チェックリスト）	機械学習品質保証担当		内部品質	※別紙、「要求事項・観点に関する対応方針」 「（開発・実装時）対応の記録」に記入						飛騨担当 プラントシステム担当		
8	機械学習要素以外の構成要素（外部安全機構等）の開発	設定した外部品質・利用時品質に必要な機械学習要素以外の構成要素（外部安全機構等）を開発		プラントシステム担当		機械学習利用システムの機械学習要素以外の構成要素（機械学習要素ではないお決まりフォーマットで整理した事）									
9	テスト（外部品質の評価）	機械学習利用システムのテストを行い、その結果を評価。機械学習要素の外部品質が要求水準に達しているか評価	2.2.3	システム品質保証担当		外部品質の評価	リスク回避性 ※3年以上ある場合は付与推奨。	i		202XR年X月XX日 テストテータのうち、修理・交換を必要とした異常のデータ全て検知できており、求める品質を達成している。 202XR年Y月YY日 要新規も修理・交換を必要とした異常のデータ検知できており、求める品質を維持できている。	202XR年X月XX日 検証結果_202XDXXX.csv 202XR年Y月YY日 検証結果_202XDYYY.csv	プラントシステム担当 環境安全担当			202XR年X月XX日 問題なし 202XR年Y月YY日 問題なし
								ii		202XR年X月XX日 テストテータのうち、修理・交換を必要とした異常のテータの異常発生箇所・異常の深刻度を80%以上の確率で診断できており、求める品質を達成している。 202XR年Y月YY日 要新規、修理・交換を必要とした異常のテータの異常発生箇所の検知の精度は再現で同様80%以上の確率で診断ができており、求める品質を達成している。	202XR年X月XX日 検証結果_202XDXXX.csv 202XR年Y月YY日 検証結果_202XDYYY.csv				202XR年X月XX日 問題なし 202XR年Y月YY日 問題なし
								iii							
							パフォーマンス ※3年以上ある場合は付与推奨。	i		202XR年X月XX日 テストテータのうち、修理・交換を必要とした異常のテータ以外でも発報されたが、発報のうち対応不要な場合は10%以下であり、ある程度低い。概ね合格とされるものとされたため、求める品質を達成している。 202XR年Y月YY日 テストテータのうち、修理・交換を必要とした異常のテータ以外でも発報されたが、発報のうち対応不要な場合は30%以下であった。所期より誤点は出ましが、運用時の実績を踏まえ十分な水準であると認められたため、求める品質を達成している。	202XR年X月XX日 検証結果_202XDXXX.csv 202XR年Y月YY日 検証結果_202XDYYY.csv				202XR年X月XX日 問題なし 202XR年Y月YY日 問題なし
								ii		202XR年X月XX日 テストテータのうち、直ちに修理が必要な箇所以外にも交換を行うべき箇所について検知できていなため、求める品質を達成しているとは判断する。ただし、実際の効果は運用開始後に異常予兆検知・診断内容や故障発生頻度、修理頻度をモニタリングする必要がある。 202XR年Y月YY日 テストテータのうち、直ちに修理が必要な箇所以外にも交換を行うべき箇所について検知できていなため、求める品質を達成している。	202XR年X月XX日 検証結果_202XDXXX.csv 202XR年Y月YY日 検証結果_202XDYYY.csv				202XR年X月XX日 問題なし 202XR年Y月YY日 問題なし
								iii							
10	検収	テストでの外部品質の評価結果から機械学習利用システム全体の評価を行い、それらの評価結果等を踏まえ基準を満たす機械学習利用システムであれば検収	-	システム品質保証担当		システム全体の評価（機械学習要素の依頼性評価以外も踏まえて判断）						事業企画担当 飛騨担当			
11	利用時品質の確認	運用中の機械学習利用システムの利用時品質を確認	2.1.1 2.1.3 2.2.1	システム品質保証担当		利用時品質	リスク回避性 ※3年以上ある場合は付与推奨。	i		202XR年Y月YY日 修理・交換を必要とした異常について、所定の期間は全て予兆を検知しており、問題ない認識。	202XR年Y月YY日 実績データ_202XDYYY.csv	飛騨担当			202XR年Y月YY日 問題なし
								ii		202XR年Y月YY日 修理・交換を必要とした異常予兆診断結果のうち50%で、一部同じ系統で異なる箇所を発生箇所として診断した。異常発生箇所の検知の精度が高まることが望ましい。	202XR年Y月YY日 実績データ_202XDYYY.csv				202XR年Y月YY日 異常発生箇所が同じ系統レベルの箇所であること、その系統全体を確認する必要があるため、である限り異常発生箇所の検知の精度向上されることが現場の作業効率化のために望ましい。
								iii							
							パフォーマンス ※3年以上ある場合は付与推奨。	i		202XR年Y月YY日 発報のうち52%が対応不要だった。以前の点検作業よりも確認箇所を確認箇所を限定できているので、現状の品質で問題ない。	202XR年Y月YY日 実績データ_202XDYYY.csv				202XR年Y月YY日 問題なし
								ii		202XR年Y月YY日 発報のうち約20%が要だった場合は大規模な修理をせず、未然に交換等の対応が実施できた。現状の品質で問題ない。	202XR年Y月YY日 実績データ_202XDYYY.csv				202XR年Y月YY日 問題なし
								iii							
12	外部品質の確認	運用中の機械学習利用システムの利用時品質の確認結果を踏まえ、外部品質を確認	2.2.3	システム品質保証担当		外部品質	リスク回避性 ※3年以上ある場合は付与推奨。	i		202XR年Y月YY日 利用時品質において問題なかったため、外部品質も現状で問題ないものと判断する。		プラントシステム担当			202XR年Y月YY日 問題なし
								ii		202XR年Y月YY日 利用時品質の確認結果を踏まえ、異常発生箇所の検知精度の向上が期待できる事柄に承認の上。					202XR年Y月YY日 問題なし
								iii							
							パフォーマンス ※3年以上ある場合は付与推奨。	i		202XR年Y月YY日 利用時品質において問題なかったため、外部品質も現状で問題ないものと判断する。					202XR年Y月YY日 問題なし
								ii		202XR年Y月YY日 利用時品質において問題なかったため、外部品質も現状で問題ないものと判断する。					202XR年Y月YY日 問題なし
								iii							
13	内部品質（運用時品質の維持性）の確認	運用中の機械学習利用システムの利用時品質・外部品質の確認結果を踏まえ、内部品質の要求事項と観点の対応状況を確認	2.2.4 2.2.5 別録（チェックリスト）	機械学習品質保証担当		内部品質	※別紙、運用時品質の維持性の「（運用時）対応の記録」に記入								
14	システムの更新	外部品質・内部品質の確認結果に応じて機械学習要素を更新	2.2.4 2.2.5 別録（チェックリスト）	機械学習設計・開発担当		内部品質	※別紙、運用時品質の維持性の「（運用時）対応の記録」に記入						飛騨担当 プラントシステム担当		
		利用時品質・外部品質の確認結果に応じて機械学習要素以外の構成要素（外部安全機構等）を更新。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提供		プラントシステム担当		機械学習利用システムの機械学習要素以外の構成要素									

サブフォーマット（内部品質確保のための「プラント保安分野での観点」チェックリスト　実施内容記録フォーマット）

【本チェックリスト・記録フォーマットの使用方法】

0 （チェックリストを見る前に）機械学習要素の開発にあたり、要求されている内部品質の要求レベル（レベル1～3）を確認する。（ガイドライン2.2.4項）
1 下記「本チェックリストの使用場面」に応じて検討する内部品質軸について、該当する要求レベルの要求事項を確認する。
2 該当する要求レベルの要求事項と直接関連する「プラント保安分野での観点」を確認する。「内部品質の要求レベル」欄で該当する要求レベル（Lv1～Lv3）を確認する。
開発する機械学習要素がユースケースと類似している場合は、同様に「ユースケース固有の観点」も確認する。
3 要求事項を問わない共通的な「プラント保安分野での観点」を確認する。「内部品質の要求レベル」欄には、「共通」と記載している。
開発する機械学習要素がユースケースと類似している場合は、同様に「ユースケース固有の観点」も確認する。
4 開発・実装時、「（開発・実装時）対応の記録」に対応の記録を対応日とあわせて記録する。
5 運用開始後はシステムの更新時に、「運用時品質の維持性」の「（運用時）対応の記録」に品質確認及び更新における対応を記録する。

【本チェックリスト・記録フォーマットの使用場面】

データの設計	1「データの設計」の検討（要求分析の十分性、データ設計の十分性） 検討の例： ・データ収集でカバーすべき属性、環境要因等の検討 ・要求分析や収集データについての現場のエンジニアへの確認　等
データの確認	2「データの品質」の検討（データセットの被覆性、データセットの均一性） 検討の例： ・カバーすべき属性・環境要因等が収集できていることの確認 ・データ不足時のシミュレータ活用の検討　等
学習	3「学習済みモデルの品質」の検討（機械学習モデルの正確性、機械学習モデルの安定性） 検討の例： ・誤判断の許容水準の検討 ・テスト手法の検討　等
実装・運用	4「実装・運用の品質」の検討（プログラムの健全性、運用時品質の維持性） 検討の例： ・実装に利用するソフトウェアの検討 ・精度検証・再学習の頻度設計　等

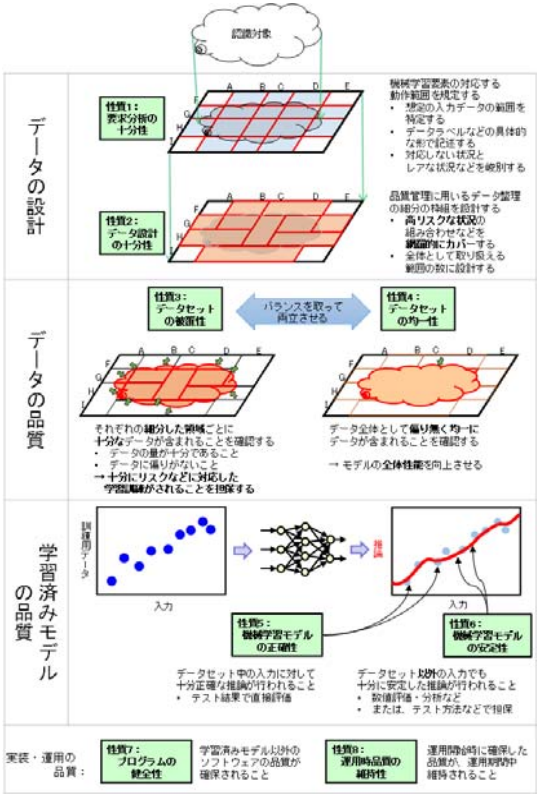

図　「機械学習品質マネジメントガイドライン」における内部品質

出所）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1版」

別紙　内部品質確保のための「プラント保安分野での観点」チェックリスト　実施内容記録フォーマット

「機械学習品質マネジメントガイドライン　第1版」における内部品質要求事項					プラント保安分野での観点	関連する「ユースケース固有の観点」（本文3.3節を参照）					（開発・実装時）対応の記録	参考資料・参考情報
使用場面	内部品質軸	内部品質要求レベル	要求事項No.	内部品質要求事項		配管の肉厚予測	配管の画像診断	設備劣化診断	異常予兆検知・診断	運転最適化		
データの設計	1要求分析の十分性	Lv1	1	主要な品質低下リスクが発生する原因について検討を行い記録する。	・プラント分野における「主要な品質低下リスク」と「その原因」として、以下を検討しましたか？ - 環境の変化：季節、天気、昼夜、気温、立地など - 製品特性の変化：種類、成分など - プラントの状態の変化：立ち上げ時、通常時など	—	—	—	—	—		
データの設計	1要求分析の十分性	Lv1	2	その検討結果に基づき、データの設計を行い必要な属性等に反映する。	・プラント分野における「品質低下リスクが発生する原因」の検討結果に基づいたデータ設計を行いましたか？	—	—	—	—	—		
データの設計	1要求分析の十分性	Lv2	3	システム全体での利用品質低下リスクとその影響について、工学的に一定の網羅性を持つ分析を行い、文書として記録する。	・「工学的に一定の網羅性を持つ分析」として、安全関連系全体のSIL評価や、FTA、STAMP/STPA等の工学的リスク分析に係る既存の情報があれば、それを活用しましたか？ ・既存の工学的分析がない場合は、一定の網羅性を持つ分析を新たに実施しましたか？	—	—	—	—	—		
データの設計	1要求分析の十分性	Lv2	4	それぞれのリスクについて対策の要否を分析し、機械学習要素への入力においてそのリスクに対応する特徴となる属性について分析を行う。	—	—	—	—	—	—		
データの設計	1要求分析の十分性	Lv2	5	また、応用に即した機械学習要素の入力をもたらす環境の特徴について、機械学習の容易さなどの分析を行い記録する。	—	—	—	—	—	—		
データの設計	1要求分析の十分性	Lv2	6	これらの分析結果に基づいて属性と属性値のセットの検討を行い、その決定の経緯を記録する。	—	—	—	—	—	—		
データの設計	1要求分析の十分性	Lv3	7	Lv2に加えて、以下の活動を行う。	—	—	—	—	—	—		
データの設計	1要求分析の十分性	Lv3	8	システムの利用環境の特徴量として捉えるべき要素について、過去の自己・他者の検討結果などの文献調査を行い、必要な集合の抽出に至る検討経緯を記録する。	・プラント分野における「利用環境の特徴量として捉えるべき要素」として、以下を検討しましたか？ - 環境の変化：季節、天気、昼夜、気温、立地など - 製品特性の変化：種類、成分など - プラントの状態の変化：立ち上げ時、通常時など	—	—	—	—	—		
データの設計	1要求分析の十分性	Lv3	9	システム全体の利用時品質低下リスクについても、そのシステムの応用分野に即した過去の検討結果などを調査し、取捨選択の経緯も含めて検討経緯を記録する。	—	—	—	—	—	—		
データの設計	1要求分析の十分性	Lv3	10	また、システム全体の利用時品質低下リスクについては、Fault Tree Analysisなどの工学的分析を用いた抽出も行い、その結果を記録する。	・「工学的分析」として、安全関連系全体のSIL評価や、FTA、STAMP/STPA等の工学的リスク分析を行いましたか？（Lv3においては必須）	—	—	—	—	—		
データの設計	1要求分析の十分性	共通	—	—	・プラントでの全ての利用状況を被覆しているか分析するために、現場のプラントエンジニアによる要求分析を行いましたか？	—	—	—	—	—		
データの設計	1要求分析の十分性	共通	—	—	・カメラ等での認識を行うAIの場合、認識対象とする設備の箇所や状態を絞り込みましたか？	—	・保温材が巻かれている配管を取り扱う場合は、配管そのものの劣化ではなく、保温材の劣化が対象となる点に注意する。	—	—	—		
データの設計	1要求分析の十分性	共通	—	—	・設備の状態の変化を検知・予測するAIの場合、検知・予測対象とする箇所や状態を絞り込みましたか？	・腐食の種類によって「データ設計の十分性」「データセットの被覆性」の評価にも影響するため、どの腐食を対象とするかという点まで絞り込む。	—	・製造条件によって変化する製品の成分値に対して、どの成分値の範囲までを対象とするか検討する。製造する製品が異なる場合だけでなく、流体(※)やプロセスが変わった場合も含む。 (※)混相流・多相流の配分変化など	・どの箇所のどのような異常を対象とするのかによって、「データ設計の十分性」「データセットの被覆性」の評価にも影響するため、どの箇所のどのような異常を対象とするかという点まで要求を特定する。	—		
データの設計	1要求分析の十分性	共通	—	—	・要求に工学的因果関係の説明が含まれる場合、活用する上で必須かどうかを確認しましたか？	—	—	—	・異常の検知と、それに関連する変数との間で、工学的に意味のある因果関係が不明であっても、相関のみで活用することは妨げない。	—		
データの設計	1要求分析の十分性	共通	—	—	・学習だけでなく交差検証や汎化性能等が確かめられるデータ量が確保できますか？	—	—	—	—	—		
データの設計	1要求分析の十分性	共通	—	—	・ユーザ企業から、課題解決に繋がるデータが提供されていますか？もしくは生成、獲得することが可能になっていますか？	—	—	—	—	—		
データの設計	1要求分析の十分性	共通	—	—	・課題のモデル化に際して、学習用データセットの説明変数の数・因果関係の数が複雑過ぎる、もしくは単純すぎることはありませんか？また、多重共線性は考慮していますか？	—	—	—	—	—		
データの設計	1要求分析の十分性	共通	—	—	・運用中の入力データについて、異常な動作に繋がるような、または悪意のあるデータを検知し、排除する機構を実装していますか？	—	—	—	—	—		
データの設計	1要求分析の十分性	共通	—	—	・既存の AI 適用先の経験を、次の開発に技術として反映できるプロセス、体制が構築されていますか？	—	—	—	—	—		
データの設計	2データ設計の十分性	Lv1	11	主要なリスク要因に対応する属性について、それぞれに対応したケースを設定すること。	—	—	—	—	—	—		
データの設計	2データ設計の十分性	Lv1	12	さらに、複合的なリスク要因については、その組み合わせに対応したケースを設定すること。	—	—	—	—	—	—		
データの設計	2データ設計の十分性	Lv1	13	また、特に重要と考えられる環境要因の差異に対する属性を抽出し、大きなリスクの要因との組み合わせに対応するケースを用意すること。	・プラント分野における「環境要因」を抽出しましたか？ - 外部環境：天気、気温、設置場所など - 生産プロセス：生産負荷、運転手順など	・「環境要因」とは、ここでは気候、塩分(海からの距離や風向などの地域性)などを指す。	・「環境要因」とは、ここでは日照、天気、季節、時間帯などを指す。 ・画像のブレに対処するためモデルにより吸収することも考えられるが、システムの複雑化・不確定要素の増加を招く可能性を把握する。	・「環境要因」とは、ここでは設置場所、動作環境、気温・湿度、運転方法、原材料、用役などを指す。	・「環境要因」とは、ここでは異常の検知に影響する環境要因（生産負荷、生産ロット等）を指す。	・「環境要因」とは、ここでは運転手順や原材料などを指す。		
データの設計	2データ設計の十分性	Lv2	14	Lv１の要求を全て満たすこと。	—	—	—	—	—	—		
データの設計	2データ設計の十分性	Lv2	15	特に重要と考えられるリスク要因については、原則としてpair-wise coverageの基準を満たすこと。具体的には、「その原因の組み合わせの属性値」と、「その属性値の属する属性以外の全ての属性について、属性に含まれる属性値を１つずつ個別に選択したもの」の組み合わせのケースを含むこと。	—	—	—	—	—	—		

データの設計	**2データ設計の十分性**	**Lv3**	16	工学的な検討に基づき、属性の網羅基準を設定し、その網羅基準を満たす属性値の組み合わせの集合をケースとして設定すること。	—	—	—	—	—	—		
データの設計	**2データ設計の十分性**	**Lv3**	17	網羅基準の厳密さ（pair-wise coverage、triple-wise coverage等）は、製品の利用状況やリスクの重大さなどを加味して設定されること。必要な場合には、個別のリスクに応じてリスク毎に基準を個別に設定することも考えられる。	—	—	—	—	—	—		
データの設計	**2データ設計の十分性**	**共通**	—	—	・リスク要因に関連する属性が抽出できているか、現場のプラントエンジニアによるケース確認を行いましたか？	—	—	—	—	—		
データの設計	**2データ設計の十分性**	**共通**	—	—	・カメラ等での認識を行うAIの場合、認識対象とする設備の箇所や材質などによって、データの範囲や取得容易性が変わることを検討しましたか？	—	・配管によって塗装や防錆塗装あるいは配管そのものの地肌の色が異なっているので、それらの違いを加味して精度を確保する。 ・配管上に積雪があるケース等、画像により直接配管外面を確認することができない場合があることを把握する。	—	—	—		
データの設計	**2データ設計の十分性**	**共通**	—	—	・データの品質が一定に保たれるように、ルールの整備を行いましたか？	—	・撮影方法のルール・留意点の整備によりデータの品質を一定レベルとすることについて検討する。	—	—	—		
データの設計	**2データ設計の十分性**	**共通**	—	—	・データの品質のブレをモデルで吸収する場合、システムの複雑化・不確定要素の増加を招く可能性を検討しましたか？	—	・画像のブレに対処するためモデルにより吸収することも考えられるが、システムの複雑化・不確定要素の増加を招く可能性を把握する。	—	—	—		
データの設計	**2データ設計の十分性**	**共通**	—	—	・対象とする製品の製品特性（種類、成分など）の範囲に対して、それらのデータの収集可能性を検討しましたか？	—	—	・対象とする製品の成分値の範囲に対して、それらの学習データが収集できるかどうか検討する	—	—		
データの設計	**2データ設計の十分性**	**共通**	—	—	・シミュレータのデータを利用する場合、環境要因の変化がシミュレータで考慮されているか確認しましたか？	—	—	・シミュレータのデータを利用する場合、環境要因の変化（例：高湿度→低湿度）がシミュレータで考慮されているか確認する。 ・データセットをシミュレーションで取得する場合は、シミュレータの妥当性を十分検証する。	・データセットをシミュレーションで取得する場合は、シミュレータの妥当性を十分検証する。	・データセットをシミュレーションで取得する場合は、シミュレータの妥当性を十分検証する。		
データの設計	**2データ設計の十分性**	**共通**	—	—	・データの収集において、メンテナンスの直後は傾向が変わる可能性があることを検討しましたか？	—	—	・部材の変更直後は基本的に「劣化無し」のデータとしてよい。「劣化なし」期間は部材のスペックに依るが、使用環境で変化することに注意する（過去の交換頻度等を参照して、「劣化なし」とする期間を決定する）。 ・部材の変更直後で「慣らし運転 」が必要な場合は、その期間はデータを収集しないなどの管理を行う。	—	—		

別紙　内部品質確保のための「プラント保安分野での観点」チェックリスト　実施内容記録フォーマット

使用場面	「機械学習品質マネジメントガイドライン　第1版」における内部品質要求事項				プラント保安分野での観点	関連する「ユースケース固有の観点」（本文3.3節を参照）					（開発・実装時）対応の記録	参考資料・参考情報
	内部品質軸	内部品質要求レベル	要求事項No.	内部品質要求事項		配管の肉厚予測	配管の画像診断	設備劣化診断	異常予兆検知・診断	運転最適化		
データの確認	3データセットの被覆性	Lv1	18	テスト用データセットの取得源や方法を検討し、応用の状況に対して偏りがないことを期待できるようにすること。	・プラント分野における「応用の状況」を抽出しましたか？	・「応用の状況」とは、ここでは対象とする配管や観測の頻度、評価の時間軸（リアルタイムの予測を行うか等）などを指す。	—	・「応用の状況」とは、ここでは対象とする設備の種類や稼働状況(常時／一時、負荷の変更等)などを指す。	・「応用の状況」とは、ここでは検知したい異常の深刻度（ハイアラーム／ハイハイアラーム等）や、機械学習利用システムの活用場面（常時／一時、昼間／夜間、定常時／非定常時　等）を指す。	・「応用の状況」とは、ここでは適用する操作場面（季節、時間帯、定常／非定常、非定常の場合はスタートアップ/シャットダウン等）、操作対象設備などを指す。		
データの確認	3データセットの被覆性	Lv1	19	各ケース毎に、元データから偏りのないサンプル抽出などを行い、偏りがないことを期待できるようにすること。	・「正常」を学習し「異常」を分類・予測する枠組みである場合、テスト用データとして異常データの網羅が難しいことを検討しましたか？	—	—	—	・本ケースの場合、学習データとして異常時データを網羅することは必須としない。一方で、正常領域における網羅的なサンプル抽出が必要。	—		
データの確認	3データセットの被覆性	Lv1	20	これらの偏りを入れないために行った活動について、記録を行う事。	—	—	—	—	—	—		
データの確認	3データセットの被覆性	Lv1	21	分析した各ケースについて訓練用データおよびテスト用データが十分に存在することを、訓練フェーズやバリデーションフェーズなどで確認すること。	—	—	—	—	—	—		
データの確認	3データセットの被覆性	Lv1	22	ケースに対して訓練用データが十分に取得できない場合には、網羅基準を見直して緩めた上で、当初の基準に照らして個別にシステム結合テストなどで確認すべきことを記録しておくこと。	—	—	—	—	—	—		
データの確認	3データセットの被覆性	Lv2	23	Lv1に加えて、以下の取り組みなどを行うこと。	—	—	—	—	—	—		
データの確認	3データセットの被覆性	Lv2	24	各属性値または各ケース毎に、およその出現確率の想定を把握すること。	—	—	—	—	—	—		
データの確認	3データセットの被覆性	Lv2	25	取得できたデータがその分布から外れていないことを確認すること。	—	—	—	—	—	—		
データの確認	3データセットの被覆性	Lv2	26	各ケース毎に、中に含まれるデータの被覆性について、取得方法以外の何らかの積極的な確認を行うこと。	—	—	—	—	—	—		
データの確認	3データセットの被覆性	Lv2	27	例えば、各ケース毎に、そのケースに含まれない属性がある場合、その属性に関する分布を抽出して、著しい偏りがないことを確認すること。	—	—	—	—	—	—		
データの確認	3データセットの被覆性	Lv3	28	Lv2に加え、各ケース毎に、中に含まれるデータの被覆性について、一定の指標を得ること。	—	—	—	—	—	—		
データの確認	3データセットの被覆性	Lv3	29	例えば、特徴量抽出などの技法を用いて、ケース組み合わせに含まれる属性値以外のデータ間相関がないことなどを確認すること。	—	—	—	—	—	—		
データの確認	3データセットの被覆性	Lv3	30	あるいは、各ケース毎の、ケースに含まれない属性の分布について、あらかじめ想定される分布を検討し、相違について分析を行い記録すること。	—	—	—	—	—	—		
データの確認	3データセットの被覆性	共通	—	—	・データセットの取得源が正しいかどうか、現場のプラントエンジニアが確認しましたか？	—	—	・稼働データの劣化有無ラベルの確認は、その判断が適切に行える専門性を持った要員によって行う。	—	—		
データの確認	3データセットの被覆性	共通	—	—	・プラントシステムの起動時など非定常状態のデータの扱いを考慮しましたか？	—	—	—	—	—		
データの確認	3データセットの被覆性	共通	—	—	・化学プラントでは状態が常に変化するため、幅広い稼働状態のデータをカバーする必要があることを検討しましたか？	—	—	—	—	—		
データの確認	3データセットの被覆性	共通	—	—	・データをシミュレーションで取得する場合は、シミュレータの妥当性を十分検討しましたか？	—	—	—	—	—		
データの確認	3データセットの被覆性	共通	—	—	・天候などの「外乱」が影響する場合、それを考慮したデータセットになっていますか？	—	—	—	—	・天候などの「外乱」を想定したデータセットになっているか留意する。		
データの確認	3データセットの被覆性	共通	—	—	・前提としている属性のデータの範囲はカバーできていますか？	・配管の内容物・流量・材質・流速・圧力など、前提としている属性のデータの範囲がカバーできているか注意する。	・周囲の環境（日照、時刻など）やドローン撮影での画像のブレなど、入力画像が不鮮明になる場合の対応を検討する。 ・環境要因の各属性のデータの範囲がカバーできているか注意する。	—	—	—		
データの確認	3データセットの被覆性	共通	—	—	・正常時のデータを扱う場合、実際に正常時のものであることの確認は、その判断が適切に行える専門性を持った要員が行いましたか？	—	—	—	・正常時データが実際に正常時のものであることの確認は、その判断が適切に行える専門性を持った要員によって行う。	—		
データの確認	3データセットの被覆性	共通	—	—	・人間の操作や手順等を学習する場合、偏ったケース設定になっていませんか？	—	—	—	—	・ベテランオペレータの運転を学習する際は、偏ったケース設定となっていないか留意する。		
データの確認	3データセットの被覆性	共通	—	—	・データの基本統計量（欠測値・外れ値の割合、平均・分散・共分散など）を確認しましたか？	—	—	—	—	—		
データの確認	3データセットの被覆性	共通	—	—	・人間によるアノテーションが必要な場合、その管理の方策（アノテーションの履歴の記録等）を検討しましたか？	—	—	—	—	—		
データの確認	3データセットの被覆性	共通	—	—	・正常／異常の判断など特に安全に関係する領域について、特にテスト用データセットの品質（外れ値や欠損がないこと、ラベルが正しいこと・ラベルを付した人が明確であること、取得日時や場所・来歴が明確であること等）を重点的に確認しましたか？	—	—	—	—	—		
データの確認	3データセットの被覆性	共通	—	—	・データの補強（data augmentationのこと。例えば画像データを左右線対称にしたデータを作り不足するデータを増やすこと）をルールベースプログラミングで行う場合、その妥当性を十分に検討しましたか？	—	—	—	—	—		
データの確認	3データセットの被覆性	共通	—	—	・補強(augumentation)したデータが適切であったか評価しましたか？開発時の仮定に対して、運用時に得られる追加データの分布やラベリングに対して適切であったか評価しましたか？	—	—	—	—	—		
データの確認	4データセットの均一性	LvE1	—	（前節「データセットの被覆性」Lv1に同じ。）	—	—	—	—	—	—		
データの確認	4データセットの均一性	LvE1	31	テスト用データセットの取得源や方法を検討し、応用の状況に対して偏りがないことを期待できるようにすること。	・プラントにおけるデータセットの取得源や取得方法に注意しましたか？	—	—	—	—	—		

データの確認	**4データセットの均一性**	**LvE1**	32	各ケース毎に、元データから偏りのないサンプル抽出などを行い、偏りがないことを期待できるようにすること。	・実測値データの量に偏りがある場合、シミュレータの利用を検討しましたか？	—	—	—	—	—		
データの確認	**4データセットの均一性**	**LvE1**	33	これらの偏りを入れないために行った活動について、記録を行う事。	—	—	—	—	—	—		
データの確認	**4データセットの均一性**	**LvE1**	34	分析した各ケースについて訓練用データおよびテスト用データが十分に存在することを、訓練フェーズやバリデーションフェーズなどで確認すること。	—	—	—	—	—	—		
データの確認	**4データセットの均一性**	**LvE1**	35	ケースに対して訓練用データが十分に取得できない場合には、網羅基準を見直して緩めた上で、当初の基準に照らして個別にシステム結合テストなどで確認すべきことを記録しておくこと。	—	—	—	—	—	—		
データの確認	**4データセットの均一性**	**LvE2**	—	（前節「データセットの被覆性」Lv2に同じ。）	—	—	—	—	—	—		
データの確認	**4データセットの均一性**	**LvE2**	36	Lv1に加えて、以下の取り組みなどを行うこと。	—	—	—	—	—	—		
データの確認	**4データセットの均一性**	**LvE2**	37	各属性値または各ケース毎に、およその出現確率の想定を把握すること。	—	—	—	—	—	—		
データの確認	**4データセットの均一性**	**LvE2**	38	取得できたデータがその分布から外れていないことを確認すること。	—	—	—	—	—	—		
データの確認	**4データセットの均一性**	**LvE2**	39	各ケース毎に、中に含まれるデータの被覆性について、取得方法以外の何らかの積極的な確認を行うこと。	—	—	—	—	—	—		
データの確認	**4データセットの均一性**	**LvE2**	40	例えば、各ケース毎に、そのケースに含まれない属性がある場合、その属性に関する分布を抽出して、著しい偏りがないことを確認すること。	—	—	—	—	—	—		
データの確認	**4データセットの均一性**	**LvE2**	41	（但し、想定する出現確率については想定事象の全集合に対して比較する。）	—	—	—	—	—	—		
データの確認	**4データセットの均一性**	**LvS1**	42	訓練用データの全体集合の量、レアケースの出現確率を比較して、レアケースのデータが訓練に不足する場合には、レアケースの学習を重点化することを検討すること。但し、特にLvE2が要求される場合には、重点化に伴い他のケースの学習が弱化することの、製品全体の品質への影響について必ず検討を行うこと。	—	—	—	—	—	—		
データの確認	**4データセットの均一性**	**LvS1**	43	前節L1で検討したケース毎のデータ量に関して、リスクに対応するケースにおいて十分なデータ量が存在することを明示的に確認すること。	—	—	—	—	—	—		
データの確認	**4データセットの均一性**	**LvS2**	44	LvS1に加え、リスク事象毎・ケース毎の出現確率の想定に基づき、各ケースのデータ量を事前に見積もり設計すること。	—	—	—	—	—	—		
データの確認	**4データセットの均一性**	**共通**	—	—	・各ケースのテスト用データセットに偏りがないか、現場のプラントエンジニアが確認しましたか？	—	—	—	—	—		
データの確認	**4データセットの均一性**	**共通**	—	—	・設備切替・保守点検等により、データ発生の頻度を含むデータ特性が変化する可能性があることを検討しましたか？	—	—	—	—	—		
データの確認	**4データセットの均一性**	**共通**	—	—	・データ取集間隔が数か月・数年という長さになる可能性があるため、プラントの運転データとの紐づけに注意しましたか？	—	—	—	—	—		
データの確認	**4データセットの均一性**	**共通**	—	—	・前提とする範囲のデータが偏りなく取得できていることを確認しましたか？	・上記の属性のカバーすべきデータの各範囲のデータ量が十分であることに注意する。	・上記の属性のカバーすべきデータの各範囲のデータ量が十分であることに注意する。	・「劣化なし」として想定する様々な状態の稼働データが偏りなく取得できていること。	・正常データとして想定する様々な範囲（昼間／夜間、定常時／非定常時、季節の違い等）のデータを偏りなく取得する。	—		
データの確認	**4データセットの均一性**	**共通**	—	—	・特定の範囲のデータ量が十分でない場合、その範囲の分類・予測精度が低くなる可能性を検討しましたか？	・ある範囲のデータ量が十分でない場合、その範囲の予測精度が低くなることを認識する。	・ある範囲のデータ量が十分でない場合、その範囲の予測精度が低くなることを認識しておく。	・ある状態の稼働データが十分取得できない場合、その状態から逸脱する劣化を検出する精度が低下する可能性があることを認識すること。	・ある範囲の正常データが十分取得できない場合、その範囲では異常を検出する精度が低下する可能性があることを認識する。	—		
データの確認	**4データセットの均一性**	**共通**	—	—	・データに特性がある場合、選択バイアス、情報バイアス、交絡の問題・リスクを評価しましたか？外れ値や欠損値の除去・訂正の根拠、措置方法について、受容・排除などのポリシーにもとづいて行いましたか？	—	—	—	—	—		

別紙　内部品質確保のための「プラント保安分野での観点」チェックリスト　実施内容記録フォーマット

「機械学習品質マネジメントガイドライン　第1版」における内部品質要求事項					プラント保安分野での観点	関連する「ユースケース固有の観点」（本文3.3節を参照）					（開発・実装時）対応の記録	参考資料・参考情報
使用場面	内部品質軸	内部品質要求レベル	要求事項No.	内部品質要求事項		配管の肉厚予測	配管の画像診断	設備劣化診断	異常予兆検知・診断	運転最適化		
学習	5機械学習モデルの正確性	Lv1	45	テスト用データとして必要なデータ量をPoC設定や過去の経験から導き出し、「データの被覆性」を満たす抽出プロセスを通じて用意すること。	–	–	–	–	–	–		
学習	5機械学習モデルの正確性	Lv1	46	テスト用のデータのラベル等の誤り及び外れ値の除去方法について検討し実施・記録すること。	–	–	–	–	–	–		
学習	5機械学習モデルの正確性	Lv1	47	訓練用データセットについても上記に準じた取扱いとする。ただし、データの分布の取り方については違う方法を採用して良い。	–	–	–	–	–	–		
学習	5機械学習モデルの正確性	Lv1	48	テスト段階において一定量の誤判断を許容する場合（false negative/false positiveで扱いを変える場合を含む）については、その判定基準を合理的に事前に決定し、記録しておくこと。	・プラント保安の現場での活用を前提として、誤判断の許容水準について検討しましたか？	–	–	–	・本ケースの場合、誤検知は一定量許容されうるが、テストに使用できる異常データの量・種類が限られるため、見逃しの発生率は極限までに近づけることが望ましい。	・「機械学習要素による操作パラメータ提示に安全機能を期待しない（設備の想定スペックを超えるような操作パラメータ提示に対しては、外部安全機構及びオペレータ判断により操作に反映しない）前提」の場合は、「危険な操作につながるパラメータの出力の許容水準を極限までに近づける」といった検討は求めない。SIL評価等に基づき、機械学習要素と他のシステムの安全機能の分担を適切に考慮して機械学習要素への要求水準を決定する。		
学習	5機械学習モデルの正確性	Lv1	49	公平性が要求される場合には、予め公平性の比較手段を定めておくこと。対照テストの結果による場合には、その合格基準を予め定めておくこと。	・（「公平性」は本ガイドラインの対象外）	–	–	–	–	–		
学習	5機械学習モデルの正確性	Lv2	50	Lv1に加えて以下の対応を取る。	–	–	–	–	–	–		
学習	5機械学習モデルの正確性	Lv2	51	テスト用・訓練用データのラベルの正当性について、何らかの追加的な確認手段を検討すること。	–	–	–	–	–	–		
学習	5機械学習モデルの正確性	Lv2	52	正解率（Accuracy）などのバリデーション段階での合否判定についても、その合理的な判定基準を事前に決定し記録しておくこと。	–	–	–	–	–	–		
学習	5機械学習モデルの正確性	Lv2	53	実データでのテストと、可能な範囲でのデータ変形などでの機械的な増量テストを同時に行うこと。	–	–	–	–	–	–		
学習	5機械学習モデルの正確性	Lv2	54	可能であれば、入力の影響度分析・ニューロンの発火状況その他の内部的な情報の分析の適用を検討し、可能な範囲で明らかな誤りを手動で排除すること。	–	–	–	–	–	–		
学習	5機械学習モデルの正確性	Lv3	55	Lv2に加え、学習成熟状況の内部確認手段等を事前に検討すること。	–	–	–	–	–	–		
学習	5機械学習モデルの正確性	Lv3	56	総合テスト以降の製品全体での検証計画と機械学習要素のテスト計画の対応を明示すること。	–	–	–	–	–	–		
学習	5機械学習モデルの正確性	Lv3	57	特にリスクが大きいケースを中心に、製品レベルのテスト時の機械学習要素の要件との対応をテスト計画に反映し、その被覆状況を監視・確認すること。	–	–	–	–	–	–		
学習	5機械学習モデルの正確性	共通	–	–	・モデルの精度評価基準について、受入れ先（プラント事業者・エンジニアリング会社など）と合意しましたか？	–	–	–	–	–		
学習	5機械学習モデルの正確性	共通	–	–	・強化学習を用いる場合でも、運用開始前に品質をテストで確認しましたか？	–	–	–	–	・強化学習を用いたとしても、運用開始前にテストを実施して「機械学習モデルの正確性」の要求事項を満たす。		
学習	5機械学習モデルの正確性	共通	–	–	・識別問題ならラベル、回帰問題なら値など問題によってつけるべき正解値が変わることを考慮しましたか？	–	–	–	–	–		
学習	5機械学習モデルの正確性	共通	–	–	・交差検証や汎化性能等に使用する訓練データとテストデータを独立して分離・管理していますか？再学習、追加学習時にも同様の管理を行っていますか？	–	–	–	–	–		
学習	5機械学習モデルの正確性	共通	–	–	・学習後の正答率、損失関数の残差は、十分に収束していますか？適合率、再現率、F値は目標に達していますか？	–	–	–	–	–		
学習	5機械学習モデルの正確性	共通	–	–	・学習・再学習過程の正答率及び損失関数の残差は、異常な変化を示していませんか？	–	–	–	–	–		
学習	5機械学習モデルの正確性	共通	–	–	・選択した AI アルゴリズムおよび高留有無の選択根拠、ハイパーパラメータの設定根拠は明確になっていますか？ユーザー・ベンダー間でアルゴリズムの選択根拠を説明・合意できていますか？	–	–	–	–	–		
学習	5機械学習モデルの正確性	共通	–	–	・学習用データセットや学習済みモデルのサイズ削減の必要はあるか確認しましたか？その際に、性能の劣化はどこまで許されるのか検討しましたか？	–	–	–	–	–		
学習	6機械学習モデルの安定性	Lv1	58	安定性向上のために適用した技術を記録すること	–	–	–	–	–	–		
学習	6機械学習モデルの安定性	Lv1	59	Lv1では、過学習を防止するために広く利用されている交差検証や正則化等の技術の適用が推奨される。	–	–	–	–	–	–		
学習	6機械学習モデルの安定性	Lv2	60	近傍データによる安定性の評価結果を記録すること	–	–	–	–	–	–		
学習	6機械学習モデルの安定性	Lv2	61	Lv2では、データセットの各データの近傍に対する安定性を評価することが求められる。例えば、近傍の敵対的データによる攻撃を防御する技術の適用が推奨される。敵対的データを生成して安定性を評価する技術、敵対的データの攻撃を受けにくくする訓練技術、敵対的データの動的検知技術等がある。これらの技術を適用することは容易ではないが、現在、そのための実用的な開発・評価するための環境整備が進められている。	–	–	–	–	–	–		
学習	6機械学習モデルの安定性	Lv3	62	近傍データに対する安定性を保証すること	–	–	–	–	–	–		
学習	6機械学習モデルの安定性	Lv3	63	レベル３では、近傍データに対して一定の安定性をもつことを保証することが求められる。例えば、近傍には敵対的データが存在しないことを保証する技術等がある。これらの技術はまだ研究段階であるが、将来的にはレベル３での適用が期待される技術である。	–	–	–	–	–	–		
学習	6機械学習モデルの安定性	共通	–	–	・特に化学プラントでは発生するデータの不確実性が高いため、学習データセット近傍のデータに対する安定性に注意が必要であることを検討しましたか？	–	–	–	–	–		
学習	6機械学習モデルの安定性	共通	–	–	・新規銘柄追加が多いプラントの場合、発生するデータが学習データセット近傍のものが多くなるため、安定性に注意が必要であることを検討しましたか？	–	–	–	–	–		
学習	6機械学習モデルの安定性	共通	–	–	・どのような汎化性能の測定が適切か、議論・整合が取れていますか？	–	–	–	–	–		
学習	6機械学習モデルの安定性	共通	–	–	・汎化性能の目標値を明確に定めていますか？学習後の AI モデルの汎化性能は、学習時の正答率と比較して著しく劣化していませんか？	–	–	–	–	–		
学習	6機械学習モデルの安定性	共通	–	–	・汎化性能を測定する方法を決めていますか？交差検証を利用する際、利用する学習用データセットのバリエーションを確保していますか？	–	–	–	–	–		
学習	6機械学習モデルの安定性	共通	–	–	・AI に影響を与えるノイズ候補の洗い出しを行っていますか？具体的には、誤差因子の選定とそれの与える影響解析を行っていますか？ノイズ候補により、AI モデルの性能が著しく劣化することはありませんか？	–	–	–	–	–		

別紙　内部品質確保のための「プラント保安分野での観点」チェックリスト　実施内容記録フォーマット

使用場面	「機械学習品質マネジメントガイドライン　第1版」における内部品質要求事項				プラント保安分野での観点	関連する「ユースケース固有の観点」（本文3.3節を参照）					（開発・実装時） 対応の記録	（運用時） 対応の記録	参考資料・参考情報
	内部品質軸	内部品質 要求レベル	要求事項 No.	内部品質要求事項		配管の肉厚予測	配管の画像診断	設備劣化診断	異常予兆検知・診断	運転最適化			
実装・運用	7プログラムの健全性	Lv1	64	利用するソフトウェアについては、信頼できる実績を持つソフトウェア等を選定し、その選定経緯を記録すること。	・シミュレータを活用する場合、実績を確認し、その選定経緯を記録しましたか？	–	–	–	–	–			
実装・運用	7プログラムの健全性	Lv1	65	選定したソフトウェアについて、その欠陥の発見等を運用期間中モニタリングし、必要に応じて修正等の措置をとること。	・シミュレータを活用する場合、欠陥のモニタリングや修正は可能ですか？	–	–	–	–	–			
実装・運用	7プログラムの健全性	Lv1	66	学習からテストフェーズに至るまでの環境と、実用段階で用いる環境の相違について、その影響などをあらかじめ検討しておくこと。	–	–	–	–	–	–			
実装・運用	7プログラムの健全性	Lv2	67	利用するソフトウェアについて、検査・実験等によりその信頼性を自己評価すること。	・シミュレータを活用する場合、検査・実験等で信頼性を自己評価しましたか？	–	–	–	–	–			
実装・運用	7プログラムの健全性	Lv2	68	可能な場合には、SIL1相当のソフトウェア信頼性を得られたソフトウェアを用いること。	–	–	–	–	–	–			
実装・運用	7プログラムの健全性	Lv2	69	システムの運用期間中のソフトウェアの健全性の維持に関する保守体制を必ず構築すること。	・シミュレータを活用する場合、保守体制が整っていますか？	–	–	–	–	–			
実装・運用	7プログラムの健全性	Lv2	70	バリデーションおよびテストフェーズにおいては、原則として実用段階で用いられる計算環境（浮動小数点精度・モデル規模等）を模倣した環境でバリデーション・テストを行うこと。または、テスト済み学習モデルと実用環境での学習モデルの動作の一致性について、何らかの検証を行う事。	–	–	–	–	–	–			
実装・運用	7プログラムの健全性	Lv3	71	SIL１（またはシステムの要求するSILレベル）のソフトウェア品質の確認を必ず行うこと。	–	–	–	–	–	–			
実装・運用	7プログラムの健全性	Lv3	72	実用環境の計算環境での学習モデルの振る舞いに基づくテスト（または形式検証など）を必ず行うこと。	–	–	–	–	–	–			
実装・運用	7プログラムの健全性	Lv3	73	また、そのモデルと実用環境での動作の一致の確認を、結合テスト以降の段階で必ず行うこと。	–	–	–	–	–	–			
実装・運用	7プログラムの健全性	共通	–	–	・以下のような場合に、機械学習要素の計算量を適切に抑えられるよう、考慮しましたか？ - 特殊な環境・デバイスで、計算リソースが限られる場合（エッジデバイス、特殊なPC等で計算する場合） - 同じデバイス上で他の処理も動いており、それらの処理への影響が許されない場合　等	–	–	–	–	–			
実装・運用	7プログラムの健全性	共通	–	–	・システムを評価するときに、外部ライブラリに対する単体テストやシステムテストを実施していますか？	–	–	–	–	–			
実装・運用	7プログラムの健全性	共通	–	–	・ライブラリのサプライヤとの間で、不具合に対する責任範囲は明確になっていますか？	–	–	–	–	–			
実装・運用	7プログラムの健全性	共通	–	–	・OS や OSS 等の各種ソフトウェアの更新頻度やサポート期間を考慮してソフトウェアを利用していますか？ソフトウェアのアップデートに対する対応や、サポートが終わった際の対応について決めていますか？	–	–	–	–	–			
実装・運用	7プログラムの健全性	共通	–	–	・特にセキュリティアップデートがある場合など、OS や OSS 等のソフトウェアのアップデートに対してシステムの更新を行っているか	–	–	–	–	–			
実装・運用	8運用時品質の維持性	Lv1	74	外部環境変化によりシステムの品質が著しく失われたときの対応について、あらかじめ検討しておくこと。	・プラント分野における「外部環境変化」を抽出しましたか？ 　- プラントの修繕や経年劣化、運転条件の変化など	–	–	–	–	–			
実装・運用	8運用時品質の維持性	Lv1	75	オンラインで追加学習を行う場合には、予想外の品質の低下がもたらす影響についてあらかじめ検討しておき、必要な場合には動作範囲の限定などのシステム的な対応を取ること。	–	–	–	–	–	–			
実装・運用	8運用時品質の維持性	Lv1	76	オフラインで追加学習を行う場合には、前７項に準じた品質管理を行う事。	–	–	–	–	–	–			
実装・運用	8運用時品質の維持性	Lv2	77	製品の利用状況が許す範囲において、システムの品質について、動作結果との対照などから品質劣化・誤判断のモニタリングを行うこと。モニタリングにおいては、プライバシーなど製品品質以外の要因を十分に検討すること。	・プラント分野における「製品品質以外の要因」を抽出しましたか？ 　- プラント事業者からAIベンダーへのデータ開示の度合いなど	–	–	–	–	–			
実装・運用	8運用時品質の維持性	Lv2	78	オンラインでの追加学習を行う場合には、追加学習結果を何らかの方法で定常的にモニタリングすること。モニタリングの結果で性能要求からの逸脱が判明した場合には、直ちに対処を行うことができること。	–	–	–	–	–	–			
実装・運用	8運用時品質の維持性	Lv2	79	オフラインでの追加学習を行う場合には、製品開発段階で用いたテスト用データセットでの「性能劣化の回帰テスト」を行い、更新前に品質が失われていないことを確認すること。必要な場合には、製品開発段階と同等の手法でテスト用データセットの更新を行うこと。	–	–	–	–	–	–			
実装・運用	8運用時品質の維持性	Lv3	80	プライバシー等と両立する製品品質の監視手段を、運用体制を含めて必ず構築すること。	–	–	–	–	–	–			
実装・運用	8運用時品質の維持性	Lv3	81	オンラインでの追加学習を行う場合には、追加学習結果をシステムに反映する前に、製品内部で一定の品質確認を行う仕組みを実装し、想定外の品質劣化が無視できない場合には更新を中止する仕組みとすること。また、オフラインでの更新・修正手段を必ず確保すること。	–	–	–	–	–	–			
実装・運用	8運用時品質の維持性	Lv3	82	オフラインでの追加学習においては、運用での収集データと、システム初期構築時のテスト用データセット、および同じ手法で定期的に更新するテスト用データセットを用いて品質を管理すること。	–	–	–	–	–	–			

実装・運用	**8運用時品質の維持性**	**共通**	—	—	・運用時品質の維持性を確保するため、モデルの再学習・再精度検証ができる体制を構築しましたか？	—	—	—	—	—			
実装・運用	**8運用時品質の維持性**	**共通**	—	—	・運用における品質の監視体制が妥当かどうか、現場のプラントエンジニアに確認を行いましたか？	—	—	—	—	—			
実装・運用	**8運用時品質の維持性**	**共通**	—	—	・メンテナンスによる部品の交換や調整、設備の改造などにより、発生するデータの特性が変化する可能性があることに注意しましたか？	—	—	・部材の種類を変えた場合には、再学習、学習モデルの切替などの対応が必要になる可能性があることに留意する。	—	—			
実装・運用	**8運用時品質の維持性**	**共通**	—	—	・データ特性の変化を把握するため、モニタリングの仕組みを検討しましたか？	—	—	—	—	・最適値の評価は定期的／継続的にモニターし、異常がないか確認する。			
実装・運用	**8運用時品質の維持性**	**共通**	—	—	・製品の成分値が変化すると、発生するデータの特性が変化する可能性が高いため、モデルの再学習・再精度検証が必要となることを確認しましたか？	—	—	・変化する製品成分値の範囲に対して、製品成分値の入力時に品質が維持されているか注意する。	—	—			
実装・運用	**8運用時品質の維持性**	**共通**	—	—	・設備の周辺状況や運転手順などの前提条件が変化した場合、モデルの再学習・再精度検証が必要となることを確認しましたか？	—	—	・対象設備そのものだけでなく、周辺の状況を含め、当初想定していた環境要因などの前提条件から逸脱していないか確認する。	・化学系プラントでは特に外部環境の変化（日照条件、風向きなど）の機械学習要素への影響が大きいため、隣接する設備の撤去や変更など、対象設備の直接的変更でない場合でも、対象設備の外部環境に影響する変更に留意する。	・設備や運転手順に変更が行われた場合は機械学習要素の出力への影響があるため、モデルを更新する。			
実装・運用	**8運用時品質の維持性**	**共通**	—	—	・運用時に収集したデータでも定期的に精度検証を行うことを検討しましたか？	—	・運用段階で撮影した画像での精度検証を行う。「目視点検要」と判定され実際に目視点検を実施した結果を記録し、これと対照して、判定精度が低い場合は、判定に用いた入力画像やモデルを重点的にチェックする。	—	—	—			
実装・運用	**8運用時品質の維持性**	**共通**	—	—	・経年劣化の進行が早いと想定される場合、これに応じて精度検証、学習モデルのチューニングの頻度を設計しましたか？	—	—	—	・対象設備の生産負荷に応じた経年劣化の進行を想定し、精度検証、学習モデルのチューニングの頻度を設計する。	—			
実装・運用	**8運用時品質の維持性**	**共通**	—	—	・（経年劣化ではなく）対象設備を大規模に修繕した場合は、その都度精度検証、学習モデルのチューニングが必要となることを確認しましたか？	—	—	—	・（経年劣化ではなく）対象設備を大規模に修繕した場合は、その都度精度検証、学習モデルのチューニングが必要とされる。	—			
実装・運用	**8運用時品質の維持性**	**共通**	—	—	・機械学習によらない従来手法の結果との対照や、実際の結果を踏まえて機械学習要素の精度等を確認しましたか？	・既存手法による交換要否判断や、実際の交換時の配管の状態等を踏まえ、実際の正解度や見落としの有無を検証する。	—	—	—	—			
実装・運用	**8運用時品質の維持性**	**共通**	—	—	特に精度の維持が求められる場合、予めモデル構築時の前提や教師データ類を整理しておき、運用開始後の変更管理を準備しましたか？	・本ケースは特に精度の維持の重要性が高いため、モデル構築時の前提、教師データ類を整理しておき、運用開始後の変更管理において都度確認する。	—	—	—	—			
実装・運用	**8運用時品質の維持性**	**共通**	—	—	必要に応じて機械学習要素の出力の範囲を限定することを検討しましたか？	—	—	—	—	・最適化の目標に対して極限まで運転条件を追求すると、安定性が損なわれることがあるため、機械学習要素の出力の範囲を限定するなどの対応を行う。			
実装・運用	**8運用時品質の維持性**	**共通**	—	—	運用時の品質を確認する上での前提（出力の想定範囲や、設備の条件設定等）を確認しましたか？	—	—	—	—	・想定していた条件（原料種等）や許容範囲内で運転されていることを確認する。 ・運用中の設備の諸条件（反応初期／終期、運転条件、原料、品質要求、スタートアップ・シャットダウン操作の許容時間等）を考慮して機械学習要素の出力の品質を確認する。			
実装・運用	**8運用時品質の維持性**	**共通**	—	—	・実運用でしか収集できないデータを記録する仕組みを構築していますか？運用で発見したエラーや多様性に対応したデータを確保しましたか？	—	—	—	—	—			
実装・運用	**8運用時品質の維持性**	**共通**	—	—	・運用中データの導入時と異なる偏りを確認し、その背景を解析していますか？外れ値や欠損値の除去・訂正の根拠、措置方法について、受容・排除などのポリシーにもとづいて行っていますか？システム維持を想定できていますか？	—	—	—	—	—			
実装・運用	**8運用時品質の維持性**	**共通**	—	—	・モデルを更新するデータが想定したデータ区間を外れているかを監視するなど、入力データの質をパトロールしていますか？	—	—	—	—	—			
実装・運用	**8運用時品質の維持性**	**共通**	—	—	・運用後、性能に影響を与える要因を抽出し、マージンを持たせた性能目標としていますか？性能劣化の検出を人間もしくは AI システムが判断する設計になっていますか？	—	—	—	—	—			
実装・運用	**8運用時品質の維持性**	**共通**	—	—	・学習用データセットのバリエーションが増えた際でも検証ができるように、交差検証の方法を決めていますか？	—	—	—	—	—			
実装・運用	**8運用時品質の維持性**	**共通**	—	—	・訓練データの特性変化や出力の追加等により再学習を行った結果、再学習前の性能に対する劣化は許容可能となっていますか？	—	—	—	—	—			
実装・運用	**8運用時品質の維持性**	**共通**	—	—	・AI モデルの更新を手動ではなく自動で実施する際に、AI モデルの特性変化や性能変化が許容範囲であることを十分検査できますか？	—	—	—	—	—			
実装・運用	**8運用時品質の維持性**	**共通**	—	—	・学習にフィードバックするデータに対して、性能劣化に繋がる悪意のあるデータの混入を防ぐことができますか？もしくは、学習前に悪意のあるデータを排除する機構がありますか？	—	—	—	—	—			
実装・運用	**8運用時品質の維持性**	**共通**	—	—	・運用中の推論に利用する入力データについて、異常な動作に繋がるような異常または悪意のあるデータを排除できる仕組みがありますか？	—	—	—	—	—			
実装・運用	**8運用時品質の維持性**	**共通**	—	—	・再学習したモデルの配信方法を検討しましたか？	—	—	—	—	—			
実装・運用	**8運用時品質の維持性**	**共通**	—	—	・リリースした AI のプログラムに異常が発生した場合、迅速にロールバックを行う仕組みがありますか？	—	—	—	—	—			

別紙　内部品質確保のための「プラント保安分野での観点」チェックリスト　実施内容記録フォーマット

「機械学習品質マネジメントガイドライン　第1版」における内部品質要求事項					プラント保安分野での観点	関連する「ユースケース固有の観点」（本文3.3節を参照）					（開発・実装時） 対応の記録	参考資料・参考情報
使用場面	内部品質軸	内部品質 要求レベル	要求事項 No.	内部品質要求事項		配管の肉厚予測	配管の画像診断	設備劣化診断	異常予兆検知・診断	運転最適化		
データの設計	**1要求分析の十分性**	**Lv1**	1	主要な品質低下リスクが発生する原因について検討を行い記録する。	・プラント分野における「主要な品質低下リスク」と「その原因」として、以下を検討しましたか？ - 環境の変化：季節、天気、昼夜、気温、立地など - 製品特性の変化：種類、成分など - プラントの状態の変化：立ち上げ時、通常時	—	—	—	—	—	（2020/10/8） ・今回は定常運転のみを想定したAIであるため、非定常運転時のデータが、精度低下という品質低下リスクの原因となる。 ・非定常運転：立ち上げ、シャットダウン、トラブル、油種切替	
データの設計	**1要求分析の十分性**	**Lv1**	2	その検討結果に基づき、データの設計を行い必要な属性等に反映する。	・プラント分野における「品質低下リスクが発生する原因」の検討結果に基づいたデータ設計を行いましたか？	—	—	—	—	—	（2020/10/8） ・学習用データセット構築において、シミュレータで収集するデータを定常運転時のもののみに限定する。 ・検証用データセット構築において、シミュレータで収集するデータを定常運転時のものと、非定常運転時のものを分けて収集する。	
データの設計	**1要求分析の十分性**	**Lv2**	3	システム全体での利用品質低下リスクとその影響について、工学的に一定の網羅性を持つ分析を行い、文書として記録する。	・「工学的に一定の網羅性を持つ分析」として、安全関連系全体のSIL評価や、FTA、STAMP/STPA等の工学的リスク分析に係る既存の情報があれば、それを活用しましたか？ ・既存の工学的分析がない場合は、一定の網羅性を持つ分析を新たに実施しましたか？	—	—	—	—	—	（2020/10/8） ・3年前に実施した当該プラントシステムのSIL評価の結果を参照した。今回のAIはシステムの特定の設備のパラメータを自動的に変えるものではなく、運転員に参考情報を提示する位置づけのものである。そのため、システム全体の利用品質低下リスクを低下させることはないと判断する。	
データの設計	**1要求分析の十分性**	**Lv2**	4	それぞれのリスクについて対策の要否を分析し、機械学習要素への入力においてそのリスクに対応する特徴となる属性について分析を行う。	—	—	—	—	—	—	（2020/10/8） ・非定常運転データが原因となる「精度低下リスク」に対して分析を実施。対応する入力属性（変数）として、「温度」「油種」をリストアップする。	
データの設計	**1要求分析の十分性**	**Lv2**	5	また、応用に即した機械学習要素の入力をもたらす環境の特徴について、機械学習の容易さなどの分析を行い記録する。	—	—	—	—	—	—	（2020/10/8） ・今回の利用環境では、生成する製品の反応に「温度」が大きく影響する。機械学習においては、外気温が外乱となる可能性があり、過去のデータが少ない外気温をどこまで想定するかにより、学習用データの準備の観点で困難さを伴う。	
データの設計	**1要求分析の十分性**	**Lv2**	6	これらの分析結果に基づいて属性と属性値のセットの検討を行い、その決定の経緯を記録する。	—	—	—	—	—	—		
データの設計	**1要求分析の十分性**	**Lv3**	7	Lv2に加えて、以下の活動を行う。	—	—	—	—	—	—		
データの設計	**1要求分析の十分性**	**Lv3**	8	システムの利用環境の特徴量として捉えるべき要素について、過去の自己・他者の検討結果などの文献調査を行い、必要な集合の抽出に至る検討経緯を記録する。	・プラント分野における「利用環境の特徴量として捉えるべき要素」として、以下を検討しましたか？ - 環境の変化：季節、天気、昼夜、気温、立地など - 製品特性の変化：種類、成分など - プラントの状態の変化：立ち上げ時、通常時など	—	—	—	—	—		
データの設計	**1要求分析の十分性**	**Lv3**	9	システム全体の利用時品質低下リスクについても、そのシステムの応用分野に即した過去の検討結果などを調査し、取捨選択の経緯も含めて検討経緯を記録する。	—	—	—	—	—	—		
データの設計	**1要求分析の十分性**	**Lv3**	10	また、システム全体の利用時品質低下リスクについては、Fault Tree Analysisなどの工学的分析を用いた抽出も行い、その結果を記録する。	・「工学的分析」として、安全関連系全体のSIL評価や、FTA、STAMP/STPA等の工学的リスク分析を行いましたか？（Lv3においては必須）	—	—	—	—	—		
データの設計	**1要求分析の十分性**	**共通**	—	—	・プラントでの全ての利用状況を被覆しているか分析するために、現場のプラントエンジニアによる要求分析を行いましたか？	—	—	—	—	—		
データの設計	**1要求分析の十分性**	**共通**	—	—	・カメラ等での認識を行うAIの場合、認識対象とする設備の箇所や状態を絞り込みましたか？	—	・保温材が巻かれている配管を取り扱う場合は、配管そのものの劣化ではなく、保温材の劣化が対象となる点に注意する。	—	—	—		
データの設計	**1要求分析の十分性**	**共通**	—	—	・設備の状態の変化を検知・予測するAIの場合、検知・予測対象とする箇所や状態を絞り込みましたか？	・腐食の種類によって「データ設計の十分性」「データセットの被覆性」の評価にも影響するため、どの腐食を対象とするかという点まで絞り込む。	—	・製造条件によって変化する製品の成分値に対して、どの成分値の範囲までを対象とするか検討する。製造する製品が異なる場合だけでなく、流体(※)やプロセスが変わった場合も含む。 (※)混相流・多相流の配分変化など	・どの箇所のどのような異常を対象とするのかによって、「データ設計の十分性」「データセットの被覆性」の評価にも影響するため、どの箇所のどのような異常を対象とするかという点まで要求を特定する。	—		
データの設計	**1要求分析の十分性**	**共通**	—	—	・要求に工学的因果関係の説明が含まれる場合、活用する上で必須かどうかを確認しましたか？	—	—	—	・異常の検知と、それに関連する変数との間で、工学的に意味のある因果関係が不明であっても、相関のみで活用することは妨げない。	—		
データの設計	**1要求分析の十分性**	**共通**	—	—	・学習だけでなく交差検証や汎化性能等が確かめられるデータ量が確保できますか？	—	—	—	—	—		
データの設計	**1要求分析の十分性**	**共通**	—	—	・ユーザ企業から、課題解決に繋がるデータが提供されていますか？もしくは生成、獲得することが可能になっていますか？	—	—	—	—	—		
データの設計	**1要求分析の十分性**	**共通**	—	—	・課題のモデル化に際して、学習用データセットの説明変数の数・因果関係の数が複雑過ぎる、もしくは単純すぎることはありませんか？また、多重共線性は考慮していますか？	—	—	—	—	—		
データの設計	**1要求分析の十分性**	**共通**	—	—	・運用中の入力データについて、異常な動作に繋がるような、または悪意のあるデータを検知し、排除する機構を実装していますか？	—	—	—	—	—		
データの設計	**1要求分析の十分性**	**共通**	—	—	・既存の AI 適用先の経験を、次の開発に技術として反映できるプロセス、体制が構築されていますか？	—	—	—	—	—		

データの設計	**2データ設計の十分性**	**Lv1**	11	主要なリスク要因に対応する属性について、それぞれに対応したケースを設定すること。	—	—	—	—	—	—	（2020/10/8） ・「要求分析の十分性」で整理したとおり、今回は非定常運転時のデータが、精度低下という品質低下リスクの原因であり、具体的な属性は以下の通り。 　- 非定常運転のケース：立ち上げ、シャットダウン、トラブル、油種切替 　- 非定常運転を判別する変数：「温度」「油種」	
データの設計	**2データ設計の十分性**	**Lv1**	12	さらに、複合的なリスク要因については、その組み合わせに対応したケースを設定すること。	—	—	—	—	—	—	（2020/10/8） ・ケースの複合について、以下のような場合が考えられる。 　-「油種切替」後の「立ち上げ」 　-「トラブル」時の「シャットダウン」、その後の「立ち上げ」	
データの設計	**2データ設計の十分性**	**Lv1**	13	また、特に重要と考えられる環境要因の差異に対する属性を抽出し、大きなリスクの要因との組み合わせに対応するケースを用意すること。	・プラント分野における「環境要因」を抽出しましたか？ 　- 外部環境：天気、気温、設置場所など 　- 生産プロセス：生産負荷、運転手順など	・「環境要因」とは、ここでは気候、塩分(海からの距離や風向などの地域性)などを指す。	・「環境要因」とは、ここでは設置場所、動作環境、気温・湿度、運転方法、原材料、用役などを指す。 ・画像のブレに対処するためモデルにより吸収することも考えられるが、システムの複雑化・不確定要素の増加を招く可能性を把握する。	・「環境要因」とは、ここでは設置場所、動作環境、気温・湿度、運転方法、原材料、用役などを指す。	・「環境要因」とは、ここでは異常の検知に影響する環境要因（生産負荷、生産ロット等）を指す。	・「環境要因」とは、ここでは運転手順や原材料などを指す。	（2020/10/8） ・「環境要因」とは、「生産負荷」「運転手順」「原材料」を指す。主要なパターンと、リスク要因の関係を整理。整理結果は「環境要因・リスク要因分析.xlsx」に記録。	（2020/10/8） 環境要因・リスク要因分析.xlsx
データの設計	**2データ設計の十分性**	**Lv2**	14	Lv１の要求を全て満たすこと。	—	—	—	—	—	—	—	
データの設計	**2データ設計の十分性**	**Lv2**	15	特に重要と考えられるリスク要因については、原則としてpair-wise coverageの基準を満たすこと。具体的には、「その原因の組み合わせの属性値」と、「その属性値の属する属性以外の全ての属性について、属性に含まれる属性値を１つずつ個別に選択したもの」の組み合わせのケースを含むこと。	—	—	—	—	—	—	（2020/10/8） ・「環境要因」の組合せの属性値とそれら以外の「環境要因」の属性値の組合せのケースについてリスク要因を分析した。分析結果は、「環境要因・リスク要因分析.xlsx」に記録。	（2020/10/8） 環境要因・リスク要因分析.xlsx
データの設計	**2データ設計の十分性**	**Lv3**	16	工学的な検討に基づき、属性の網羅基準を設定し、その網羅基準を満たす属性値の組み合わせの集合をケースとして設定すること。	—	—	—	—	—	—		
データの設計	**2データ設計の十分性**	**Lv3**	17	網羅基準の厳密さ（pair-wise coverage、triple-wise coverage等）は、製品の利用状況やリスクの重大さなどを加味して設定されること。必要な場合には、個別のリスクに応じてリスク毎に基準を個別に設定することも考えられる。	—	—	—	—	—	—		
データの設計	**2データ設計の十分性**	**共通**	—	—	・リスク要因に関連する属性が抽出できているか、現場のプラントエンジニアによるケース確認を行いましたか？	—	—	—	—	—		
データの設計	**2データ設計の十分性**	**共通**	—	—	・カメラ等での認識を行うAIの場合、認識対象とする設備の箇所や材質などによって、データの範囲や取得容易性が変わることを検討しましたか？	—	・配管によって塗装や防錆塗装あるいは配管そのものの地肌の色が異なっているので、それらの違いを加味して精度を確保する。 ・配管上に積雪があるケース等、画像により直接配管外面を確認することができない場合があることを把握する。	—	—	—		
データの設計	**2データ設計の十分性**	**共通**	—	—	・データの品質が一定に保たれるように、ルールの整備を行いましたか？	—	・撮影方法のルール・留意点の整備によりデータの品質を一定レベルとすることについて検討する。	—	—	—		
データの設計	**2データ設計の十分性**	**共通**	—	—	・データの品質のブレをモデルで吸収する場合、システムの複雑化・不確定要素の増加を招く可能性を検討しましたか？	—	・画像のブレに対処するためモデルにより吸収することも考えられるが、システムの複雑化・不確定要素の増加を招く可能性を把握する。	—	—	—		
データの設計	**2データ設計の十分性**	**共通**	—	—	・対象とする製品の製品特性（種類、成分など）の範囲に対して、それらのデータの収集可能性を検討しましたか？	—	—	・対象とする製品の成分値の範囲に対して、それらの学習データが収集できるかどうか検討する	—	—		
データの設計	**2データ設計の十分性**	**共通**	—	—	・シミュレータのデータを利用する場合、環境要因の変化がシミュレータで考慮されているか確認しましたか？	—	—	・シミュレータのデータを利用する場合、環境要因の変化（例：高湿度→低湿度）がシミュレータで考慮されているか確認する。 ・データセットをシミュレーションで取得する場合は、シミュレータの妥当性を十分検証する。	・データセットをシミュレーションで取得する場合は、シミュレータの妥当性を十分検証する。	・データセットをシミュレーションで取得する場合は、シミュレータの妥当性を十分検証する。		
データの設計	**2データ設計の十分性**	**共通**	—	—	・データの収集において、メンテナンスの直後は傾向が変わる可能性があることを検討しましたか？	—	—	・部材の変更直後は基本的に「劣化無し」のデータとしてよい。「劣化なし」期間は部材のスペックに依るが、使用環境で変化することに注意する（過去の交換頻度等を参照して、「劣化なし」とする期間を決定する）。 ・部材の変更直後で「慣らし運転 」が必要な場合は、その期間はデータを収集しないなどの管理を行う。	—	—		

別紙　内部品質確保のための「プラント保安分野での観点」チェックリスト　実施内容記録フォーマット

使用場面	「機械学習品質マネジメントガイドライン　第1版」における内部品質要求事項					プラント保安分野での観点	関連する「ユースケース固有の観点」（本文3.3節を参照）					（開発・実装時） 対応の記録	参考資料・参考情報
	内部品質軸	内部品質 要求レベル	要求事項 No.	内部品質要求事項			配管の肉厚予測	配管の画像診断	設備劣化診断	異常予兆検知・診断	運転最適化		
データの確認	3データセットの被覆性	Lv1	18	テスト用データセットの取得源や方法を検討し、応用の状況に対して偏りがないことを期待できるようにすること。	・プラント分野における「応用の状況」を抽出しましたか？	・「応用の状況」とは、ここでは対象とする配管や観測の頻度、評価の時間軸（リアルタイムの予測を行うか等）などを指す。	—	・「応用の状況」とは、ここでは対象とする設備の種類や稼働状況(常時／一時、負荷の変更等)などを指す。	・「応用の状況」とは、ここでは検知したい異常の深刻度（ハイアラーム／ハイハイアラーム等）や、機械学習利用システムの活用場面（常時／一時、昼間／夜間、定常時／非定常時　等）を指す。	・「応用の状況」とは、ここでは適用する操作場面（季節、時間帯、定常／非定常、非定常の場合はスタートアップ/シャットダウン等）、操作対象設備などを指す。	（2020/10/8） ・今回の「応用の状況」として以下を抽出する。 ・立地：東京湾 ・季節：年間通期で利用 ・天候：天候を問わず利用 ・時間帯：1日を通じて利用 ・定常運転時に利用 ・季節として、冬季の稼働データが少ないため、外気温が低温時のシミュレーションを行い、追加する。	
データの確認	3データセットの被覆性	Lv1	19	各ケース毎に、元データから偏りのないサンプル抽出などを行い、偏りがないことを期待できるようにすること。	・「正常」を学習し「異常」を分類・予測する枠組みである場合、テスト用データとして異常データの網羅が難しいことを検討しましたか？	—	—	—	・本ケースの場合、学習データとして異常時データを網羅することは必須としない。一方で、正常領域における網羅的なサンプル抽出が必要。	—	（2020/10/8） ・今回は3年間のプロセスデータにシミュレータのデータを追加する。 ・季節として、冬季の稼働データが少ないため、外気温が低温時のシミュレーションを行い、追加する。 （2020/10/22） ・外気温が低温時のシミュレーションを実施しデータを追加。収集したプロセスデータ及びシミュレーションのデータの履歴を「データセット履歴.xlsx」に格納。	（2020/10/22） データセット履歴.xlsx
データの確認	3データセットの被覆性	Lv1	20	これらの偏りを入れないために行った活動について、記録を行う事。	—	—	—	—	—	—	（2020/10/8） ・現場の運転員に対し、運転操作を行う際に判断材料にするパラメータについてヒアリングを実施。ヒアリング結果は「運転員ヒアリング.xlsx」に記録。 ・ヒアリングの結果、外気温が判断に大きく影響することがわかった。	（2020/10/8） 運転員ヒアリング.xlsx
データの確認	3データセットの被覆性	Lv1	21	分析した各ケースについて訓練用データおよびテスト用データが十分に存在することを、訓練フェーズやバリデーションフェーズなどで確認すること。	—	—	—	—	—	—	（2020/11/1） ・訓練データと検証データについて、各ケースに対するデータを確認。	
データの確認	3データセットの被覆性	Lv1	22	ケースに対して訓練用データが十分に取得できない場合には、網羅基準を見直して緩めた上で、当初の基準に照らして個別にシステム結合テストなどで確認すべきことを記録しておくこと。	—	—	—	—	—	—	（2020/10/8） ・今回はシミュレーションによるデータ追加を行うことで、ケースに対して十分なデータが確保できていると判断。	
データの確認	3データセットの被覆性	Lv2	23	Lv1に加えて、以下の取り組みなどを行うこと。	—	—	—	—	—	—	—	
データの確認	3データセットの被覆性	Lv2	24	各属性値または各ケース毎に、およその出現確率の想定を把握すること。	—	—	—	—	—	—	（2020/10/8） ・「データ設計の十分性」で検討した主要な属性値とパターンの整理は「環境要因・リスク要因分析.xlsx」に記録。 ・各属性値・ケース毎の出現確率の想定の分析を実施。分析結果は「属性値・ケース出現確率分析.xlsx」に記録。	（2020/10/8） 属性値・ケース出現確率分析.xlsx
データの確認	3データセットの被覆性	Lv2	25	取得できたデータがその分布から外れていないことを確認すること。	—	—	—	—	—	—	（2020/11/1） ・「データセット履歴.xlsx」のデータが「属性値・ケース出現確率分析.xlsx」で想定した範囲に収まることを確認。検証結果は「データ被覆性検証結果.xlsx」に記録。	（2020/11/1） データ被覆性検証結果.xlsx
データの確認	3データセットの被覆性	Lv2	26	各ケース毎に、中に含まれるデータの被覆性について、取得方法以外の何らかの機械的な確認を行うこと。	—	—	—	—	—	—	（2020/11/1） ・ケースを無作為に抜き出し、当該ケースで関わらない属性値についてその分布を確認した結果、データセットに当該属性値の偏りのないことを確認した。検証結果は「データ被覆性検証結果.xlsx」に記録。	（2020/11/1） データ被覆性検証結果.xlsx
データの確認	3データセットの被覆性	Lv2	27	例えば、各ケース毎に、そのケースに含まれない属性がある場合、その属性に関する分布を抽出して、著しい偏りがないことを確認すること。	—	—	—	—	—	—	※上記の対応に準ずる。	※上記の参考資料に準ずる
データの確認	3データセットの被覆性	Lv3	28	Lv2に加え、各ケース毎に、中に含まれるデータの被覆性について、一定の指標を得ること。	—	—	—	—	—	—		
データの確認	3データセットの被覆性	Lv3	29	例えば、特徴量抽出などの技法を用いて、ケース組み合わせに含まれる属性値以外のデータ間相関がないことなどを確認すること。	—	—	—	—	—	—		
データの確認	3データセットの被覆性	Lv3	30	あるいは、各ケース毎の、ケースに含まれない属性の分布について、あらかじめ想定される分布を検討し、相違について分析を行い記録すること。	—	—	—	—	—	—		
データの確認	3データセットの被覆性	共通	—	—	・データセットの取得源が正しいかどうか、現場のプラントエンジニアが確認しましたか？	—	—	・稼働データの劣化有無ラベルの確認は、その判断が適切に行える専門性を持った要員によって行う。	—	—		
データの確認	3データセットの被覆性	共通	—	—	・プラントシステムの起動時など非定常状態のデータの扱いを考慮しましたか？	—	—	—	—	—		
データの確認	3データセットの被覆性	共通	—	—	・化学プラントでは状態が常に変化するため、幅広い稼働状態のデータをカバーする必要があることを検討しましたか？	—	—	—	—	—		
データの確認	3データセットの被覆性	共通	—	—	・データをシミュレーションで取得する場合は、シミュレータの妥当性を十分検討しましたか？	—	—	—	—	—		
データの確認	3データセットの被覆性	共通	—	—	・天候などの「外乱」が影響する場合、それを考慮したデータセットになっていますか？	—	—	—	—	・天候などの「外乱」を想定したデータセットになっているか留意する。		
データの確認	3データセットの被覆性	共通	—	—	・前提としている属性のデータの範囲はカバーできていますか？	・配管の内容物・流量・材質・流速・圧力など、前提としている属性のデータの範囲がカバーできているか注意する。	・周囲の環境（日照、時刻など）やドローン撮影での画像のブレなど、入力画像が不鮮明になる場合の対応を検討する。 ・ 環境要因の各属性のデータの範囲がカバーできているか注意する。	—	—	—		
データの確認	3データセットの被覆性	共通	—	—	・正常時のデータを扱う場合、実際に正常時のものであることの確認は、その判断が適切に行える専門性を持った要員が行いましたか？	—	—	—	・正常時データが実際に正常時のものであることの確認は、その判断が適切に行える専門性を持った要員によって行う。	—		
データの確認	3データセットの被覆性	共通	—	—	・人間の操作や手順等を学習する場合、偏ったケース設定になっていませんか？	—	—	—	—	・ベテランオペレータの運転を学習する際は、偏ったケース設定となっていないか留意する。		
データの確認	3データセットの被覆性	共通	—	—	・データの基本統計量（欠測値・外れ値の割合、平均・分散・共分散など）を確認しましたか？	—	—	—	—	—		
データの確認	3データセットの被覆性	共通	—	—	・人間によるアノテーションが必要な場合、その管理の方策（アノテーションの履歴の記録等）を検討しましたか？	—	—	—	—	—		
データの確認	3データセットの被覆性	共通	—	—	・正常／異常の判断など特に安全に関係する領域について、特にテスト用データセットの品質（外れ値や欠損がないこと、ラベルが正しいこと・ラベルを付した人が明確であること、取得日時や場所・来歴が明確であること等）を重点的に確認しましたか？	—	—	—	—	—		
データの確認	3データセットの被覆性	共通	—	—	・データの補強（data augmentationのこと。例えば画像データを左右線対称にしたデータを作り不足するデータを増やすこと）をルールベースプログラミングで行う場合、その妥当性を十分に検討しましたか？	—	—	—	—	—		
データの確認	3データセットの被覆性	共通	—	—	・補強(augumentation)したデータが適切であったか評価しましたか？開発時の仮定に対して、運用時に得られる追加データの分布やラベリングに対して適切であったか評価しましたか？	—	—	—	—	—		

データの確認	**4データセットの均一性**	**LvE1**	–	（前節「データセットの被覆性」Lv1に同じ。）	–	–	–	–	–	–		
データの確認	**4データセットの均一性**	**LvE1**	31	テスト用データセットの取得源や方法を検討し、応用の状況に対して偏りがないことを期待できるようにすること。	・プラントにおけるデータセットの取得源や取得方法に注意しましたか？	–	–	–	–	–	（2020/10/8） ・テスト用データセットは、導入先プラント（東京湾沿い）の過去3年間のうち3分の1のデータをランダムに抽出し、「応用の状況」として挙げた季節・天候・時間帯の均一性を持つデータセットを準備した。	
データの確認	**4データセットの均一性**	**LvE1**	32	各ケース毎に、元データから偏りのないサンプル抽出などを行い、偏りがないことを期待できるようにすること。	・実測値データの量に偏りがある場合、シミュレータの利用を検討しましたか？	–	–	–	–	–	（2020/10/8） ・今回は3年間のプロセスデータにシミュレータのデータを追加する。 ・季節として、冬季の稼働データが少ないため、外気温が低温時のシミュレーションを行い、追加する。 （2020/10/22） ・外気温が低温時のシミュレーションを実施しデータを追加。	
データの確認	**4データセットの均一性**	**LvE1**	33	これらの偏りを入れないために行った活動について、記録を行う事。	–	–	–	–	–	–	（2020/10/8） ・現場の運転員に対し、運転操作を行う際に判断材料にするパラメータについてヒアリングを実施。ヒアリング結果は「運転員ヒアリング.xlsx」に記録。 ・ヒアリングの結果、外気温が判断に大きく影響することがわかった。また、年間季節問わず稼働しているが、冬季の稼働データが不足していることがわかった。	（2020/10/8） 運転員ヒアリング.xlsx
データの確認	**4データセットの均一性**	**LvE1**	34	分析した各ケースについて訓練用データおよびテスト用データが十分に存在することを、訓練フェーズやバリデーションフェーズなどで確認すること。	–	–	–	–	–	–	（2020/11/1） ・訓練データと検証データについて、各ケースに対するデータを確認。	
データの確認	**4データセットの均一性**	**LvE1**	35	ケースに対して訓練用データが十分に取得できない場合には、網羅基準を見直して緩めた上で、当初の基準に照らして個別にシステム結合テストなどで確認すべきことを記録しておくこと。	–	–	–	–	–	–	（2020/10/8） ・今回はシミュレーションによるデータ追加を行うことで、ケースに対して十分なデータが確保できていると判断。	
データの確認	**4データセットの均一性**	**LvE2**	–	（前節「データセットの被覆性」Lv2に同じ。）	–	–	–	–	–	–	–	
データの確認	**4データセットの均一性**	**LvE2**	36	Lv1に加えて、以下の取り組みなどを行うこと。	–	–	–	–	–	–	–	
データの確認	**4データセットの均一性**	**LvE2**	37	各属性値または各ケース毎に、およその出現確率の想定を把握すること。	–	–	–	–	–	–	（2020/10/8） ・「データ設計の十分性」で検討した主要な属性値とパターンの整理は「環境要因・リスク要因分析.xlsx」に記録。 ・各属性値・ケース毎の出現確率の想定の分析を実施。分析結果は「属性値・ケース出現確率分析.xlsx」に記録。	（2020/10/8） 属性値・ケース出現確率分析.xlsx
データの確認	**4データセットの均一性**	**LvE2**	38	取得できたデータがその分布から外れていないことを確認すること。	–	–	–	–	–	–	（2020/11/1） ・「データセット履歴.xlsx」のデータが「属性値・ケース出現確率分析.xlsx」で想定した範囲に収まることを確認。検証結果は「データ被覆性検証結果.xlsx」に記録。	（2020/11/1） データ被覆性検証結果.xlsx
データの確認	**4データセットの均一性**	**LvE2**	39	各ケース毎に、中に含まれるデータの被覆性について、取得方法以外の何らかの積極的な確認を行うこと。	–	–	–	–	–	–	（2020/11/1） ・ケースを無作為に抜き出し、当該ケースで関わらない属性値についてその分布を確認した結果、データセットに当該属性値の偏りのないことを確認した。検証結果は「データ被覆性検証結果.xlsx」に記録。	（2020/11/1） データ被覆性検証結果.xlsx
データの確認	**4データセットの均一性**	**LvE2**	40	例えば、各ケース毎に、そのケースに含まれない属性がある場合、その属性に関する分布を抽出して、著しい偏りがないことを確認すること。	–	–	–	–	–	–	※上記の対応に準ずる。	※上記の参考資料に準ずる
データの確認	**4データセットの均一性**	**LvE2**	41	（但し、想定する出現確率については想定事象の全集合に対して比較する。）	–	–	–	–	–	–	※上記の対応に準ずる。	※上記の参考資料に準ずる
データの確認	**4データセットの均一性**	**LvS1**	42	確率を比較して、レアケースのデータが訓練に不足する場合には、レアケースの学習を重点化することを検討すること。但し、特にLvE2が要求される場合には、重点化に伴い他のケースの学習が弱化することの、製品全体の品質への影響について必ず検討を行うこと。	–	–	–	–	–	–		
データの確認	**4データセットの均一性**	**LvS1**	43	前節L1で検討したケース毎のデータ量に関して、リスクに対応するケースにおいて十分なデータ量が存在することを明示的に確認すること。	–	–	–	–	–	–		
データの確認	**4データセットの均一性**	**LvS2**	44	LvS1に加え、リスク事象毎・ケース毎の出現確率の想定に基づき、各ケースのデータ量を事前に見積もり設計すること。	–	–	–	–	–	–		
データの確認	**4データセットの均一性**	**共通**	–	–	・各ケースのテスト用データセットに偏りがないか、現場のプラントエンジニアが確認しましたか？	–	–	–	–	–		
データの確認	**4データセットの均一性**	**共通**	–	–	・設備切替・保守点検等により、データ発生の頻度を含むデータ特性が変化する可能性があることを検討しましたか？	–	–	–	–	–		
データの確認	**4データセットの均一性**	**共通**	–	–	・データ取集間隔が数か月・数年という長さになる可能性があるため、プラントの運転データとの紐づけに注意しましたか？	–	–	–	–	–		
データの確認	**4データセットの均一性**	**共通**	–	–	・前提とする範囲のデータが偏りなく取得できていることを確認しましたか？	・上記の属性のカバーすべきデータの各範囲のデータ量が十分であることに注意する。	・上記の属性のカバーすべきデータの各範囲のデータ量が十分であることに注意する。	・「劣化なし」として想定する様々な状態の稼働データが偏りなく取得できていること。	・正常データとして想定する様々な範囲（昼間／夜間、定常時／非定常時、季節の違い等）のデータを偏りなく取得する。	–		
データの確認	**4データセットの均一性**	**共通**	–	–	・特定の範囲のデータ量が十分でない場合、その範囲の分類・予測精度が低くなる可能性を検討しましたか？	・ある範囲のデータ量が十分でない場合、その範囲の予測精度が低くなることを認識する。	・ある範囲のデータ量が十分でない場合、その範囲の予測精度が低くなることを認識しておく。	・ある状態の稼働データが十分取得できない場合、その状態から逸脱する劣化を検出する精度が低下する可能性があることを認識すること。	・ある範囲の正常データが十分取得できない場合、その範囲では異常を検出する精度が低下する可能性があることを認識する。	–		
データの確認	**4データセットの均一性**	**共通**	–	–	・データに特性がある場合、選択バイアス、情報バイアス、交絡の問題・リスクを評価しましたか？外れ値や欠損値の除去・訂正の根拠、措置方法について、受容・排除などのポリシーにもとづいて行いましたか？	–	–	–	–	–		

別紙　内部品質確保のための「プラント保安分野での観点」チェックリスト　実施内容記録フォーマット

使用場面	「機械学習品質マネジメントガイドライン　第1版」における内部品質要求事項					プラント保安分野での観点	関連する「ユースケース固有の観点」（本文3.3節を参照）					（開発・実装時） 対応の記録	参考資料・参考情報
	内部品質軸	内部品質 要求レベル	要求事項 No.	内部品質要求事項			配管の肉厚予測	配管の画像診断	設備劣化診断	異常予兆検知・診断	運転最適化		
学習	5機械学習モデルの正確性	Lv1	45	テスト用データとして必要なデータ量をPoC仮定や過去の経験から導き出し、「データの被覆性」を満たす抽出プロセスを通じて用意すること。	—		—	—	—	—	—	（2020/10/8） ・PoCフェーズの検証結果から、3年分の実際のプロセスデータとシミュレーションデータで精度の担保に十分と判断する。	
学習	5機械学習モデルの正確性	Lv1	46	テスト用のデータのラベル等の誤り及び外れ値の除去方法について検討し実施・記録すること。	—		—	—	—	—	—	（2020/11/1） ・事前にデータ分布を確認し、3σ外れた値が発生した状況を分析する。分析の結果、定常的な運転状態でなければそのデータを除外する。除外したデータを「データセット履歴_クレンズ.xlsx」に記録する。	（2020/11/1） データセット履歴_クレンズ.xlsx
学習	5機械学習モデルの正確性	Lv1	47	訓練用データセットについても上記に準じた取扱いとする。ただし、データの分布の取り方については違う方法を採用して良い。	—		—	—	—	—	—	（2020/11/1） ・上記と同様の対応を行う。	
学習	5機械学習モデルの正確性	Lv1	48	テスト段階において一定量の誤判断を許容する場合（false negative/false positiveで扱いを変える場合を含む）については、その判定基準を合理的に事前に決定し、記録しておくこと。	・プラント保安の現場での活用を前提として、誤判断の許容水準について検討しましたか？		—	—	—	・本ケースの場合、誤検知は一定量許容されるが、テストに使用できる異常データの量・種類が限られるため、見逃しの発生率は極限まで0に近づけることが望ましい。	・「機械学習要素による操作パラメータ提示に安全機能を期待しない（設備の想定スペックを超えるような操作パラメータ提示に対しては、外部安全機構及びオペレータ判断により操作に反映しない）前提」の場合は、「危険な操作につながるパラメータの出力の許容水準を極限まで0に近づける」といった検討は求めない。SIL評価等に基づき、機械学習要素と他のシステムの安全機能の分担を適切に考慮して機械学習要素への要求水準を決定する。	（2020/10/8） ・現場の運転員へのヒアリング及びアラート発生の記録から、現状のアラートの発生頻度を分析する。現状のアラート頻度を超えないように誤判断の許容割合を設定する。	
学習	5機械学習モデルの正確性	Lv1	49	公平性が要求される場合には、予め公平性の比較手段を定めておくこと。対照テストの結果による場合には、その合格基準を予め定めておくこと。	・（「公平性」は本ガイドラインの対象外）		—	—	—	—	—	（2020/10/8） ・公平性は要求されないと判断する。	
学習	5機械学習モデルの正確性	Lv2	50	Lv1に加えて以下の対応を取る。	—		—	—	—	—	—	—	
学習	5機械学習モデルの正確性	Lv2	51	テスト用・訓練用データのラベルの正当性について、何らかの追加的な確認手段を検討すること。	—		—	—	—	—	—	（2020/11/1） ・Lv1での対応に加え、現場担当者やプラントシステム担当に外れた値が連続する箇所や特徴的な箇所について、運転状態や値の取り扱いの判断を仰ぐ。対応結果を「データセット履歴_クレンズ.xlsx」に記録する。	（2020/11/1） データセット履歴_クレンズ.xlsx
学習	5機械学習モデルの正確性	Lv2	52	正解率（Accuracy）などのバリデーション段階での合否判定についても、その合理的な判定基準を事前に決定し記録しておくこと。	—		—	—	—	—	—	（2020/11/1） ・正解率の判断基準について、異常判定率は70%以上の水準、その水準への要求はできる限り（best-effort）として設定する。バリデーション及びテストの結果を確認し、合否は適宜現場担当者やプラントシステム担当の意見も踏まえ判断する。	
学習	5機械学習モデルの正確性	Lv2	53	実データでのテストと、可能な範囲でのデータ変形などでの機械的な増量テストを同時に行うこと。	—		—	—	—	—	—	（2020/11/1） ・今回はシミュレーションによるデータ追加を行うことで、ケースに対して十分なデータが確保できていると判断。	
学習	5機械学習モデルの正確性	Lv2	54	可能であれば、入力の影響度分析・ニューロンの発火状況その他の内部的な情報の分析の適用を検討し、可能な範囲で明らかな誤りを手動で排除すること。	—		—	—	—	—	—	（2020/11/1） ・データのトレンドについて現場担当者やプラントシステム担当者のアドバイスを仰ぎ、「環境要因」、特に外気温との関係やセンサーデータのノイズが含まれないか確認の上、明らかな誤りについては手動で排除する。対応結果を「データセット履歴_クレンズ.xlsx」に記録する。	（2020/11/1） データセット履歴_クレンズ.xlsx
学習	5機械学習モデルの正確性	Lv3	55	Lv2に加え、学習成熟状況の内部確認手段等を事前に検討すること。	—		—	—	—	—	—		
学習	5機械学習モデルの正確性	Lv3	56	結合テスト以降の製品全体での検証計画と機械学習要素のテスト計画の対応を明示すること。	—		—	—	—	—	—		
学習	5機械学習モデルの正確性	Lv3	57	特にリスクが大きいケースを中心に、製品レベルのテスト時の機械学習要素の要件との対応をテスト計画に反映し、その被覆状況を監視・確認すること。	—		—	—	—	—	—		
学習	5機械学習モデルの正確性	共通	—	—	・モデルの精度評価基準について、受入れ先（プラント事業者・エンジニアリング会社など）と合意しましたか？		—	—	—	—	—		
学習	5機械学習モデルの正確性	共通	—	—	・強化学習を用いる場合でも、運用開始前に品質をテストで確認しましたか？		—	—	—	—	・強化学習を用いたとしても、運用開始前にテストを実施して「機械学習モデルの正確性」の要求事項を満たす。		
学習	5機械学習モデルの正確性	共通	—	—	・識別問題ならラベル、回帰問題なら値など問題によってつけるべき正解値が変わることを考慮しましたか？		—	—	—	—	—		
学習	5機械学習モデルの正確性	共通	—	—	・交差検証や汎化性能等に使用する訓練データとテストデータを独立して分離・管理していますか？再学習、追加学習時時にも同様の管理を行っていますか？		—	—	—	—	—		
学習	5機械学習モデルの正確性	共通	—	—	・学習後の正答率、損失関数の残差は、十分に収束していますか？適合率、再現率、F 値は目標に達していますか？		—	—	—	—	—		
学習	5機械学習モデルの正確性	共通	—	—	・学習・再学習過程の正答率及び損失関数の残差は、異常な変化を示していませんか？		—	—	—	—	—		
学習	5機械学習モデルの正確性	共通	—	—	・選択した AI アルゴリズムおよび蒸留有無の選択根拠、ハイパーパラメータの設定根拠は明確になっていますか？ユーザー・ベンダー間でアルゴリズムの選択根拠を説明・合意できていますか？		—	—	—	—	—		
学習	5機械学習モデルの正確性	共通	—	—	・学習用データセットや学習済みモデルのサイズ削減の必要はあるか確認しましたか？その際に、性能の劣化はどこまで許されるのか検討しましたか？		—	—	—	—	—		

学習	**6機械学習モデルの安定性**	**Lv1**	58	安定性向上のために適用した技術を記録すること	—	—	—	—	—	—	（2020/11/10） ・交差検証を行い、過学習を防ぎかつ結果が安定することを検証した。検証結果は「交差検証_20201110.csv」に記録した。	（2020/11/10） 交差検証_20201110.csv
学習	**6機械学習モデルの安定性**	**Lv1**	59	Lv1では、過学習を防止するために広く利用されている交差検証や正則化等の技術の適用が推奨される。	—	—	—	—	—	—	（2020/11/10） ・交差検証を行い、過学習を防ぎかつ結果が安定することを検証した。検証結果は「交差検証_20201110.csv」に記録した。	（2020/11/10） 交差検証_20201110.csv
学習	**6機械学習モデルの安定性**	**Lv2**	60	近傍データによる安定性の評価結果を記録すること	—	—	—	—	—	—	（2020/11/10） ・「応用の状況」のうち、運転の判断に特に影響を及ぼす外気温（季節の属性の代表データ）を既存のデータ範囲から上下に近傍のデータをシミュレーターで生成し、結果が安定することを確認した。評価結果は「近傍データ検証_20201110.csv」に記録した。	（2020/11/10） 近傍データ検証_20201110.csv
学習	**6機械学習モデルの安定性**	**Lv2**	61	Lv2では、データセットの各データの近傍に対する安定性を評価することが求められる。例えば、近傍の敵対的データによる攻撃を防御する技術の適用が推奨される。敵対的データを生成して安定性を評価する技術、敵対的データの攻撃を受けにくくする訓練技術、敵対的データの動的検知技術等がある。これらの技術を適用することは容易ではないが、現在、そのための実用的な開発・評価するための環境整備が進められている。	—	—	—	—	—	—	（2020/11/10） ・「応用の状況」のうち、運転の判断に特に影響を及ぼす外気温（季節の属性の代表データ）を既存のデータ範囲から上下に近傍のデータをシミュレーターで生成し、結果が安定することを確認した。評価結果は「近傍データ検証_20201110.csv」に記録した。	（2020/11/10） 近傍データ検証_20201110.csv
学習	**6機械学習モデルの安定性**	**Lv3**	62	近傍データに対する安定性を保証すること	—	—	—	—	—	—		
学習	**6機械学習モデルの安定性**	**Lv3**	63	レベル３では、近傍データに対して一定の安定性をもつことを保証することが求められる。例えば、近傍には敵対的データが存在しないことを保証する技術等がある。これらの技術はまだ研究段階であるが、将来的にはレベル３での適用が期待される技術である。	—	—	—	—	—	—		
学習	**6機械学習モデルの安定性**	**共通**	—	—	・特に化学プラントでは発生するデータの不確実性が高いため、学習データセット近傍のデータに対する安定性に注意が必要であることを検討しましたか？	—	—	—	—	—		
学習	**6機械学習モデルの安定性**	**共通**	—	—	・新規銘柄追加が多いプラントの場合、発生するデータが学習データセット近傍のものが多くなるため、安定性に注意が必要であることを検討しましたか？	—	—	—	—	—		
学習	**6機械学習モデルの安定性**	**共通**	—	—	・どのような汎化性能の測定が適切か、議論・整合が取れていますか？	—	—	—	—	—		
学習	**6機械学習モデルの安定性**	**共通**	—	—	・汎化性能の目標値を明確に定めていますか？学習後の AI モデルの汎化性能は、学習時の正答率と比較して著しく劣化していませんか？	—	—	—	—	—		
学習	**6機械学習モデルの安定性**	**共通**	—	—	・汎化性能を測定する方法を決めていますか？交差検証を利用する際、利用する学習用データセットのバリエーションを確保していますか？	—	—	—	—	—		
学習	**6機械学習モデルの安定性**	**共通**	—	—	・AI に影響を与えるノイズ候補の洗い出しを行っていますか？具体的には、誤差因子の選定とそれの与える影響解析を行っていますか？ノイズ候補により、AI モデルの性能が著しく劣化することはありませんか？	—	—	—	—	—		

別紙　内部品質確保のための「プラント保安分野での観点」チェックリスト　実施内容記録フォーマット

使用場面	「機械学習品質マネジメントガイドライン　第1版」における内部品質要求事項				プラント保安分野での観点	関連する「ユースケース固有の観点」（本文3.3節を参照）					（開発・実装時） 対応の記録	（運用時） 対応の記録	参考資料・参考情報
	内部品質軸	内部品質 要求レベル	要求事項 No.	内部品質要求事項		配管の肉厚予測	配管の画像診断	設備劣化診断	異常予兆検知・診断	運転最適化			
実装・運用	7プログラムの健全性	Lv1	64	利用するソフトウェアについては、信頼できる実績を持つソフトウェア等を選定し、その選定経緯を記録すること。	・シミュレータを活用する場合、実績を確認し、その選定経緯を記録しましたか？	–	–	–	–	–	（2020/10/8） ・今回はシミュレータを利用する。今回利用するシミュレータは他の用途でも広く活用しているものであり、実績は十分であると判断する。		
実装・運用	7プログラムの健全性	Lv1	65	選定したソフトウェアについて、その欠陥の発見等を運用期間中モニタリングし、必要に応じて修正等の措置をとること。	・シミュレータを活用する場合、欠陥のモニタリングや修正は可能ですか？	–	–	–	–	–	（2020/10/8） ・データセットに対し、実データとシミュレーションデータの取得元ラベルを付与しておく。もしシミュレーションデータに誤りがあることが分かった場合、正しいデータに置き換えて再学習することで、どのような影響があるか検証することができる。		
実装・運用	7プログラムの健全性	Lv1	66	学習からテストフェーズに至るまでの環境と、実用段階で用いる環境の相違について、その影響などをあらかじめ検討しておくこと。	–	–	–	–	–	–	（2020/10/8） ・実際の運用で取得された実データを用いるため、環境に由来するデータの相違は少ないと判断する。		
実装・運用	7プログラムの健全性	Lv2	67	利用するソフトウェアについて、検査・実験等によりその信頼性を自己評価すること。	・シミュレータを活用する場合、検査・実験等で信頼性を自己評価しましたか？	–	–	–	–	–	（2020/10/8） ・今回利用するシミュレータは他の用途でも広く活用しているものであり、実績は十分であると判断する。		
実装・運用	7プログラムの健全性	Lv2	68	可能な場合には、SIL1相当のソフトウェア信頼性を得られたソフトウェアを用いること。	–	–	–	–	–	–	（2020/10/8） ・今回利用するシミュレータについては、JIS C 0508-3 でSIL1相当として記載されている項目のうち、「HR（強く推奨）」とされている項目はすべて対応、「R（推奨）」とされている項目は可能な範囲で対応している。		JIS C 0508-3 電気・電子・プログラマブル電子安全関連系の機能安全－第3部：ソフトウェア要求事項
実装・運用	7プログラムの健全性	Lv2	69	システムの運用期間中のソフトウェアの健全性の維持に関する保守体制を必ず構築すること。	・シミュレータを活用する場合、保守体制が整っていますか？	–	–	–	–	–	（2020/10/8） ・運用期間中は現場担当者によるアウトプットデータのモニタリング、品質保証担当等を含めたシステム全体の定期的なモニタリング、機械学習開発を担当した事業者のサポート契約に基づく適宜相談可能な体制を構築し、必要なときに必要な保守対応を取れるような体制を整備する。		
実装・運用	7プログラムの健全性	Lv2	70	バリデーションおよびテストフェーズにおいては、原則として実用段階で用いられる計算環境（浮動小数点精度・モデル規模等）を模倣した環境でバリデーション・テストを行うこと。または、テスト済み学習モデルと実用環境での学習モデルの動作の一致性について、何らかの検証を行う事。	–	–	–	–	–	–	（2020/10/8） ・実行時環境を再現した環境の構築が困難であるため、実機によるシステム全体のテストにおいて、品質が許容範囲内であることを確認する。		
実装・運用	7プログラムの健全性	Lv3	71	SIL１（またはシステムの要求するSILレベル）のソフトウェア品質の確認を必ず行うこと。	–	–	–	–	–	–			
実装・運用	7プログラムの健全性	Lv3	72	実用環境の計算環境での学習モデルの振る舞いに基づくテスト（または形式検証など）を必ず行うこと。	–	–	–	–	–	–			
実装・運用	7プログラムの健全性	Lv3	73	また、そのモデルと実用環境での動作の一致の確認を、結合テスト以降の段階で必ず行うこと。	–	–	–	–	–	–			
実装・運用	7プログラムの健全性	共通	–	–	・以下のような場合に、機械学習要素の計算量を適切に抑えられるよう、考慮しましたか？ - 特殊な環境・デバイスで、計算リソースが限られる場合（エッジデバイス、特殊なPC等で計算する場合） - 同じデバイス上で他の処理も動いており、それらの処理への影響が許されない場合　等	–	–	–	–	–			
実装・運用	7プログラムの健全性	共通	–	–	・システムを評価するときに、外部ライブラリに対する単体テストやシステムテストを実施していますか？	–	–	–	–	–			
実装・運用	7プログラムの健全性	共通	–	–	・ライブラリのサプライヤとの間で、不具合に対する責任範囲は明確になっていますか？	–	–	–	–	–			
実装・運用	7プログラムの健全性	共通	–	–	・OS や OSS 等の各種ソフトウェアの更新頻度やサポート期間を考慮してソフトウェアを利用していますか？ソフトウェアのアップデートに対する対応や、サポートが終わった際の対応について決めていますか？	–	–	–	–	–			
実装・運用	7プログラムの健全性	共通	–	–	・特にセキュリティアップデートがある場合など、OS や OSS 等のソフトウェアのアップデートに対してシステムの更新を行っているか	–	–	–	–	–			
実装・運用	8運用時品質の維持性	Lv1	74	外部環境変化によりシステムの品質が著しく失われたときの対応について、あらかじめ検討しておくこと。	・プラント分野における「外部環境変化」を抽出しましたか？ 　- プラントの修繕や経年劣化、運転条件の変化など	–	–	–	–	–	（2020/10/8） ・「外部環境変化」としては、プラント設備の経年劣化、プラント設備の変更、取扱う油種の変化が挙げられる。 ・上記のような「外部環境変化」が起きる場合には、現場担当者はシステムの出力を注意して確認し、現場担当者が確認した結果著しく品質が損われたと判断した場合は、以下の手順を踏む。 　- システムの出力ではなく現場担当者の判断で運転を行う 　- 環境安全担当・システム品質保証担当に報告し、対応を検討する		
実装・運用	8運用時品質の維持性	Lv1	75	オンラインで追加学習を行う場合には、予想外の品質の低下がもたらす影響についてあらかじめ検討しておき、必要な場合には動作範囲の限定などのシステム的な対応を取ること。	–	–	–	–	–	–	（2020/10/8） ・オンラインで追加学習は行わないため、当該留意事項は対象外とする。		
実装・運用	8運用時品質の維持性	Lv1	76	オフラインで追加学習を行う場合には、前７項に準じた品質管理を行う事。	–	–	–	–	–	–	（2020/10/9） ・オフラインで追加学習を行うため、ここまでの7つの軸で開発時に行った内容を参照、必要に応じて見直しを行う。		
実装・運用	8運用時品質の維持性	Lv2	77	製品の利用状況が許す範囲において、システムの品質について、動作結果との対照などから品質劣化・誤判断のモニタリングを行うこと。モニタリングにおいては、プライバシーなど製品品質以外の要因を十分に検討すること。	・プラント分野における「製品品質以外の要因」を抽出しましたか？ 　・プラント事業者からAIベンダーへのデータ開示の度合いなど	–	–	–	–	–	（2020/10/8） ・「製品品質以外の要因」として、現場担当者の主観、AIベンダーへのデータ開示度合いが挙げられる。 ・現場担当者の主観については、複数の担当者による確認結果、また他のプラントの担当者による動作結果の評価などを交え、公正に行う。 ・AIベンダーへのデータ開示の度合いについては、データはインプットデータと動作結果とその評価結果、開発時の検証結果を共有し、その他に必要な情報があれば適宜相談の上共有の要否を判断する。		
実装・運用	8運用時品質の維持性	Lv2	78	オンラインでの追加学習を行う場合には、追加学習結果を何らかの方法で定常的にモニタリングすること。モニタリングの結果で性能要求からの逸脱が判明した場合には、直ちに対処を行うことができること。	–	–	–	–	–	–	（2020/10/8） ・オンラインで追加学習は行わないため、当該留意事項は対象外とする。		
実装・運用	8運用時品質の維持性	Lv2	79	オフラインでの追加学習を行う場合には、製品開発段階で用いたテスト用データセットでの「性能劣化の回帰テスト」を行い、更新前に品質が失われていないことを確認すること。必要な場合には、製品開発段階と同等の手法でテスト用データセットの更新を行うこと。	–	–	–	–	–	–	（2020/10/8） ・製品開発で用いるデータセットを用いて、システムの更新後に品質の確認を行ってから更新を反映する。またシステム更新時は、データセットのみならず運用前のテストと同じ工程で品質の確認を行う。		

実装・運用	**8運用時品質の維持性**	**Lv3**	80	プライバシー等と両立する製品品質の監視手段を、運用体制を含めて必ず構築すること。	—	—	—	—	—	—			
実装・運用	**8運用時品質の維持性**	**Lv3**	81	オンラインでの追加学習を行う場合には、追加学習結果をシステムに反映する前に、製品内部で一定の品質確認を行う仕組みを実装し、想定外の品質劣化が無視できない場合には更新を中止する仕組みとすること。また、オフラインでの更新・修正手段を必ず確保すること。	—	—	—	—	—	—			
実装・運用	**8運用時品質の維持性**	**Lv3**	82	オフラインでの追加学習においては、運用での収集データと、システム初期構築時のテスト用データセット、および同じ手法で定期的に更新するテスト用データセットを用いて品質を管理すること。	—	—	—	—	—	—			
実装・運用	**8運用時品質の維持性**	**共通**	—	—	・運用時品質の維持性を確保するため、モデルの再学習・再精度検証ができる体制を構築しましたか？	—	—	—	—	—			
実装・運用	**8運用時品質の維持性**	**共通**	—	—	・運用における品質の監視体制が妥当かどうか、現場のプラントエンジニアに確認を行いましたか？	—	—	—	—	—			
実装・運用	**8運用時品質の維持性**	**共通**	—	—	・メンテナンスによる部品の交換や調整、設備の改造などにより、発生するデータの特性が変化する可能性があることに注意しましたか？	—	—	・部材の種類を変えた場合には、再学習、学習モデルの切替などの対応が必要になる可能性があることに留意する。	—	—			
実装・運用	**8運用時品質の維持性**	**共通**	—	—	・データ特性の変化を把握するため、モニタリングの仕組みを検討しましたか？	—	—	—	—	・最適値の評価は定期的／継続的にモニターし、異常がないか確認する。			
実装・運用	**8運用時品質の維持性**	**共通**	—	—	・製品の成分値が変化すると、発生するデータの特性が変化する可能性が高いため、モデルの再学習・再精度検証が必要となることを確認しましたか？	—	—	・変化する製品成分値の範囲に対して、製品成分値の入力時に品質が維持されているか注意する。	—	—			
実装・運用	**8運用時品質の維持性**	**共通**	—	—	・設備の周辺状況や運転手順などの前提条件が変化した場合、モデルの再学習・再精度検証が必要となることを確認しましたか？	—	—	・対象設備そのものだけでなく、周辺の状況を含め、当初想定していた環境要因などの前提条件から逸脱していないか確認する。	・化学系プラントでは特に外部環境の変化（日照条件、風向きなど）の機械学習要素への影響が大きいため、隣接する設備の撤去や変更など、対象設備の直接的変更でない場合でも、対象設備の外部環境に影響する変更に留意する。	・設備や運転手順に変更が行われた場合は機械学習要素の出力への影響があるため、モデルを更新する。			
実装・運用	**8運用時品質の維持性**	**共通**	—	—	・運用時に収集したデータでも定期的に精度検証を行うことを検討しましたか？	—	・運用段階で撮影した画像での精度検証を行う。「目視点検要」と判定され実際に目視点検を実施した結果を記録し、これと対照して、判定精度が低い場合は、判定に用いた入力画像やモデルを重点的にチェックする。	—	—	—			
実装・運用	**8運用時品質の維持性**	**共通**	—	—	・経年劣化の進行が早いと想定される場合、これに応じて精度検証、学習モデルのチューニングの頻度を設計しましたか？	—	—	—	・対象設備の生産負荷に応じた経年劣化の進行を想定し、精度検証、学習モデルのチューニングの頻度を設計する。	—			
実装・運用	**8運用時品質の維持性**	**共通**	—	—	・（経年劣化ではなく）対象設備を大規模に修繕した場合は、その都度精度検証、学習モデルのチューニングが必要となることを確認しましたか？	—	—	—	・（経年劣化ではなく）対象設備を大規模に修繕した場合は、その都度精度検証、学習モデルのチューニングが必要とされる。	—			
実装・運用	**8運用時品質の維持性**	**共通**	—	—	・機械学習によらない従来手法の結果との対照や、実際の結果を踏まえて機械学習要素の精度等を確認しましたか？	・既存手法による交換要否判断や、実際の交換時の配管の状態等を踏まえ、実際の正解度や見落としの有無を検証する。	—	—	—	—			
実装・運用	**8運用時品質の維持性**	**共通**	—	—	特に精度の維持が求められる場合、予めモデル構築時の前提や教師データ類を整理しておき、運用開始後の変更管理を準備しましたか？	・本ケースは特に精度の維持の重要性が高いため、モデル構築時の前提、教師データ類を整理しておき、運用開始後の変更管理において都度確認する。	—	—	—	—			
実装・運用	**8運用時品質の維持性**	**共通**	—	—	必要に応じて機械学習要素の出力の範囲を限定することを検討しましたか？	—	—	—	—	・最適化の目標に対して極限まで運転条件を追求すると、安定性が損なわれることがあるため、機械学習要素の出力の範囲を限定するなどの対応を行う。			
実装・運用	**8運用時品質の維持性**	**共通**	—	—	運用時の品質を確認する上での前提（出力の想定範囲や、設備の条件設定等）を確認しましたか？	—	—	—	—	・想定していた原料（原油種等）の内挿範囲内で運転されていることを確認する。 ・運用中の設備の諸条件（反応初期／終期、運転条件、原料、品質要求、スタートアップ・シャットダウン操作の許容時間等）を考慮して機械学習要素の出力の品質を確認する。			
実装・運用	**8運用時品質の維持性**	**共通**	—	—	・実運用でしか収集できないデータを記録する仕組みを構築していますか？運用で発見したエラーや多様性に対応したデータを確保しましたか？	—	—	—	—	—			
実装・運用	**8運用時品質の維持性**	**共通**	—	—	・運用中データの導入時と異なる偏りを確認し、その背景を解析していますか？外れ値や欠損値の除去・訂正の根拠、措置方法について、受容・排除などのポリシーにもとづいて行っていますか？システム維持を想定できていますか？	—	—	—	—	—			
実装・運用	**8運用時品質の維持性**	**共通**	—	—	・モデルを更新するデータが想定したデータ区間を外れているかを監視するなど、入力データの質をパトロールしていますか？	—	—	—	—	—			
実装・運用	**8運用時品質の維持性**	**共通**	—	—	・運用後、性能に影響を与える要因を抽出し、マージンを持たせた性能目標としていますか？性能劣化の検出を人間もしくは AI システムが判断する設計になっていますか？	—	—	—	—	—			
実装・運用	**8運用時品質の維持性**	**共通**	—	—	・学習用データセットのバリエーションが増え た際でも検証ができるように、交差検証の方法を決めていますか？	—	—	—	—	—			
実装・運用	**8運用時品質の維持性**	**共通**	—	—	・訓練データの特性変化や出力の追加等により再学習を行った結果、再学習前の性能に対する劣化は許容可能となっていますか？	—	—	—	—	—			
実装・運用	**8運用時品質の維持性**	**共通**	—	—	・AI モデルの更新を手動ではなく自動で実施する際に、AI モデルの特性変化や性能変化が許容範囲であることを十分検査できますか？	—	—	—	—	—			
実装・運用	**8運用時品質の維持性**	**共通**	—	—	・学習にフィードバックするデータに対して、性能劣化に繋がる悪意のあるデータの混入を防ぐことができますか？もしくは、学習前に悪意のあるデータを排除する機構がありますか？	—	—	—	—	—			
実装・運用	**8運用時品質の維持性**	**共通**	—	—	・運用中の推論に利用する入力データについて、異常な動作に繋がるような異常または悪意のあるデータを排除できる仕組みがありますか？	—	—	—	—	—			
実装・運用	**8運用時品質の維持性**	**共通**	—	—	・再学習したモデルの配信方法を検討しましたか？	—	—	—	—	—			
実装・運用	**8運用時品質の維持性**	**共通**	—	—	・リリースした AI のプログラムに異常が発生した場合、迅速にロールバックを行う仕組みがありますか？	—	—	—	—	—			