リスクチェーンモデル(RCModel)ケース検討事例：
Case11 プラント運転における制御ガイドAI

※ 本ケースは、AIビジネス推進コンソーシアム「倫理ワーキンググループ成果物」ユースケース1で取扱われており、千代田化工建設㈱より助言をいただきながら一般的な事例として検討を行った。（注：千代田化工建設㈱での実ケースを直接評価したものではない）
参考URL：https://aibpc.org/?p=1523

東京大学未来ビジョン研究センター
技術ガバナンス研究ユニット　AIガバナンスプロジェクト

ケース検討のステップ

リスク アセスメント	Step1	**AIサービスが「実現すべき価値・目的」の定義** AIサービスの特徴を踏まえて「実現すべき価値・目的」を検討(ビジネス目的及び社会的責任)
	Step2	**重要なリスクシナリオの検討** 「実現すべき価値・目的」を阻害する「リスクシナリオ」を検討
コントロール コーディネーション	Step3	**重要なリスクシナリオごとにリスクチェーン(リスク要因の関係性)の検討** 「リスクシナリオ」ごとにRCModelの各層からリスク要因と関係性(リスクチェーン)を可視化
	Step4	**リスクチェーンに従ってリスクコントロールを検討** リスクチェーンで関連づけられた構成要素においてリスク対応策(コントロール)を検討
	Step5	**各リスクチェーンでの検討を元にステークホルダー毎の役割を整理** 各リスクチェーンで検討したコントロールを集約してステークホルダー毎の役割を整理し合意形成

ケース事例（AIサービスとリスクコーディネーション研究会）

東京大学未来ビジョン研究センター　技術ガバナンス研究ユニット

AIガバナンスプロジェクト　AIサービスとリスクコーディネーション研究会

https://ifi.u-tokyo.ac.jp/projects/ai-service-and-risk-coordination/

RCM
RiskChainModel

ケーススタディの概要

ケーススタディの概要（Case11：プラント運転における制御ガイドAI）

- AIサービスが「実現すべき価値・目的」の定義 -

Step1

A石油会社におけるプラント運転管理部門における非定常運転である原料切替運転の最適運転パラメータをガイドとして提示するAIサービスである。

AIサービスプロバイダのB社AI開発部門は、ビジネス利用者であるA社設備運転部門より、ベテラン運転員のヒヤリングを通して、理想的な運転とされる多数の評価軸を数値化し、過去の運転記録から評価点を調整して評価関数を作る。これを元に、設備を再現した動的シミュレータでAIが仮想的に運転を行い、試行錯誤しながら高い報酬を得られる運転（最適化運転）を学習する深層強化学習を行う。安定して高い報酬を得られるまでに学習したAIにより、実際の原料切り替え運転時に、ライブデータを元に運転員が参考に出来る最適運転パラメータ群を出し続けて運転補助を行う。

なお、AIモデルは委託先であるC社のAI開発部門の協力を得て開発を行っている。

※1)現在、プラントの通常運転の多くの部分で高度制御プログラムによる自動運転が進んでいるが、本対象の様に技術的難易度が高く、手動制御でないと行えない運転もあり、ベテラン運転員の減少からの新技術による運転補助のニーズがある。

【実現すべき価値・目的】

- **非生産期間の短縮(収益拡大)** ※AIモデルの複雑なマニュアル運転操作後の近未来予測と複数最適操作推奨パラメータの提示
- **運転員の教育** ※過去のベテラン運転技術による評価軸を元に複雑な運転操作を数値評価する事で自己評価を可能にする若手教育推進と技術伝承
- **安全な切替作業の実施**
- **社会的責任**

【AIサービスを用いた実運用の流れ】

評価関数：ベテランが評価する複数軸を総合評価した評価関数でベテランと同等以上の点数となるAI出力。

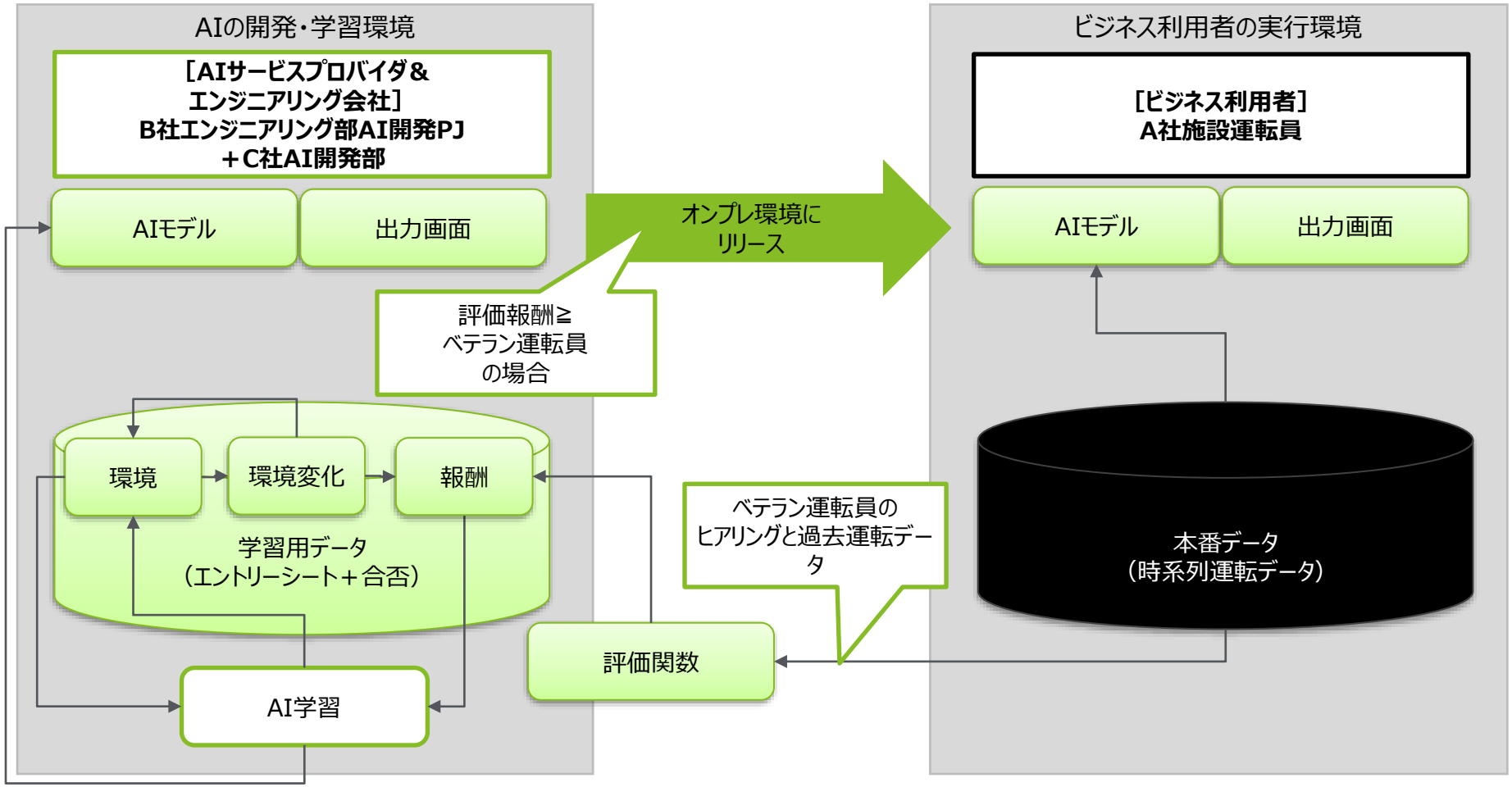
※処理時間の短さ、切替中のロス抑制など複数の違う次元の評価軸を統合し、数時間から半日かかる一連の操作を総合して報酬が最も良くなるように探索する。

※評価関数でベテラン運転員以上の点数を出したAIの出力を運転の推奨値として提示する。

ケーススタディの概要（Case11：プラント運転における制御ガイドAI）

- システムの全体概要 -

AIシステム	B社エンジニアリング部AI開発PJ＋C社AI開発部	深層強化学習を用いたシミュレーションによるパラメータ推定
AIサービスプロバイダ	B社エンジニアリング部AI開発PJ	原料切替運転のリアルタイムな最適運転パラメータの提示
ユーザー	A社施設運転員(プラント運転管理部門)	原料切り替え作業（パラメータ設定作業）



ケーススタディの概要（Case11：プラント運転における制御ガイドAI）

- AIサービスの入出力 -

【AIサービスに使用するデータ】

データ	本番/学習	収集方法	データ管理者（管理場所）	個人情報の有無
過去の運転記録	学習	運転記録の保存データ	A社運転管理責任者(中央制御室サーバー)	なし
動的シミュレーター	学習	動的シミュレータをAIが操作・試行錯誤してより良い運転を学習	学習時生成されたデータはB社に帰属	なし
最新の設備運転データ	本番	既存システムによる自動収集	A社運転管理責任者(中央制御室サーバー)	なし

【AIサービスからの出力内容】

AIサービス利用者	A社設備運転員
出力結果の内容	複数設定パラメータ（マニュアルコントローラー群のセットポイント値）
出力方法（画面／アクチュエータ等）	A社プラント中央制御室の端末上に、次の最適とされる設定値が出力(表示)される。
期待精度（正解率／誤差等）	深層強化学習の評価関数でベテラン運転員以上の総合評価ポイントを出す出力
利用者判断の有無	有
根拠情報の出力	無
安全性のリスク有無	無
外部AIへの連携	無
外部AIへの連携方法・プロトコル	無

RCM
RiskChainModel

リスクアセスメント＆リスクコントロール
- リスクシナリオの検討・評価 → コントロールの整備 -

重要なリスクシナリオの検討

- 「実現すべき価値・目的」を阻害する「リスクシナリオ」を検討 -

Step2

実現すべき価値・目的		サービス要件と関連テクノロジー			リスクNo.	リスクシナリオ	
1	非生産期間の短縮(収益拡大)	1-1	予測性能の維持	■ AIの予測精度 ■ AIの頑健性 ■ AIの説明可能性	R001	環境変化による性能劣化	原料組成が当初想定し学習させた範囲から変化して予測性能が低下する
					R002	ノイズの混入	本番環境へのノイズの混入により判断精度が安定しない
					R003	データ品質の変化	設備点検・清掃時(触媒変更等)に取得されるデータの品質が大きく変化して、予測性能が安定しなくなる
		1-2	適切なサービスコスト	ー	R004	コスト超過	サービス維持コストが超過する
2	運転員の教育	2-1	教育内容の妥当性	■ AIの説明可能性	R005	目的外用途での利用	対応していない原料・業務に対して、このAIモデルを転用しようとして誤った判断により操業に支障を与える
					R006	技術的に誤った解釈	AIの判断傾向について誤った解釈が流布し、運転員が誤った業務を行ってしまう
		2-2	トラブル時の対応		R007	トラブル時の対応不能	異常時の対応などに係る教育・訓練をおろそかにすることで異常発生時に適切な対応が取れない
3	安全な切替作業の実施	3-1	安全性の確保	■ セーフガード	R008	異常値の制御	異常な出力が行われた際に正常な範囲に補正できず、プラントの操業に支障をきたす
		3-2	外部攻撃からの保護	■ 頑健性 ■ セキュリティ	R009	学習データ・モデルへの攻撃	外部からの攻撃等により異常な学習データの投入もしくはモデルの変更が行われ、正しい予測ができなくなる
					R010	本番システムへの攻撃	外部からの攻撃等により、センサーやコントローラに障害が発生し、正しい予測ができなくなる
4	社会的責任	4-1	アカウンタビリティ	■ プロセスの説明 ■ 検証可能性	R011	品質監査への対応	品質監査などを受けた際に適切な説明ができない
					R012	トラブル発生時の調査	異常・トラブルの発生により、対外的な説明が求められる際に原因・再発防止策を検討・説明できない

重要なリスクシナリオに対するコントロールのサマリー

- 各リスクチェーンの検討結果を集約 -

Step5

実現すべき価値・目的		リスクNo.	リスクシナリオ	不確実性	環境変化	利用者起因	RC	コントロールのサマリー		
								AIシステム	サービスプロバイダ	ユーザー
1	非生産期間の短縮(収益拡大)	R001	環境変化による性能劣化	○			●	データ分布変化の検討 新たなモデルの開発 学習時の性能の記録	A/Bテスト AIモデル変更の認識併せ 再学習	AIモデル変更の認識併せ
		R002	ノイズの混入	○			●	センサーのメンテナンス データのノイズ補正 モデルの頑健性	ノイズの影響の検証 AIモデル変更の認識併せ 再学習	AIモデル変更の認識併せ
		R003	データ品質の変化	○	○		●	※R002と同じ	※R002と同じ	※R002と同じ
		R004	コスト超過						適切な価格設定	ROIの設定
2	運転員の教育	R005	目的外用途での利用			○			対応用途についての説明	AIの対応用途の理解
		R006	技術的に誤った解釈	○		○	●	判断根拠の出力	判断根拠・参考情報の分かりやすい出力	最終判断の責任 業務知見確保 判断誤りの発見
		R007	トラブル時の対応不能			○		システムログの保存	障害対応の体制整備	業務知見確保 人手での代替運用
3	安全な切替作業の実施	R008	異常値の制御	○			●	十分な正解率の確保 異常パラメータの補正 自動補正の記録	異常値の補正ルール検討 自動補正有無の表記 モデルの検証	最終判断の責任
		R009	学習データ・モデルへの攻撃		○			システム基盤の保護	論理アクセスコントロール	
		R010	本番システムへの攻撃		○			IoT等の状態監視	物理アクセスコントロール	
4	社会的責任	R011	品質監査への対応	○			●	データの理解の記録 モデルの性能の記録 操作ログの記録	内部監査 外部監査対応 アクセスコントロール	
		R012	トラブル発生時の調査	○				ログデータの保存	システム稼働監視 障害対応	人手での代替運用

ステークホルダー毎の役割を整理

- 各コントロールをステークホルダー別に整理 -

Step5

- 責任者 -
B社）経営者

- 実現すべき価値・目的の検討
- リスクコントロール方法の承認

- 責任者 -
A社）経営者

B社）品質管理部

- 内部監査への対応
- 外部監査への対応

A社）内部監査部

- 内部監査の実施

外部監査人

- サービスプロバイダ -
B社）エンジニアリング部AI開発

- 対応用途・原料についての説明
- 判断根拠の分かりやすい表現
- 異常パラメータの補正ルールの認識併せ
- 適切なサービスコストの設定
- 内部監査対応

- データ分布変化の検討
- A/Bテストでの検証
- ノイズの影響の検証
- AIモデルの変更要否の検証
- AIモデル変更の認識併せ
- AIモデルの変更

- AIモデル -
C社）モデル開発部

- 新たなモデルの開発
- モデルの頑健性
- モデルの判断根拠の出力
- パラメータ異常値の自動補正

- AIシステム -
B社）AI開発PJ

- センサーのメンテナンス
- データのノイズ補正
- 利用ログの保存
- システム稼働監視

- ユーザー -
A社）プラント運転管理部門

- 最終判断の責任
- AIの対応用途の理解
- 業務知見確保
- 判断誤りの発見
- 異常パラメータの補正ルールの認識併せ
- AIモデル変更の認識併せ
- ROIの設定
- 代替運用

RCM
RiskChainModel

リスクコントロールの検討
- リスクチェーンを用いたコントロール検討の詳細 -

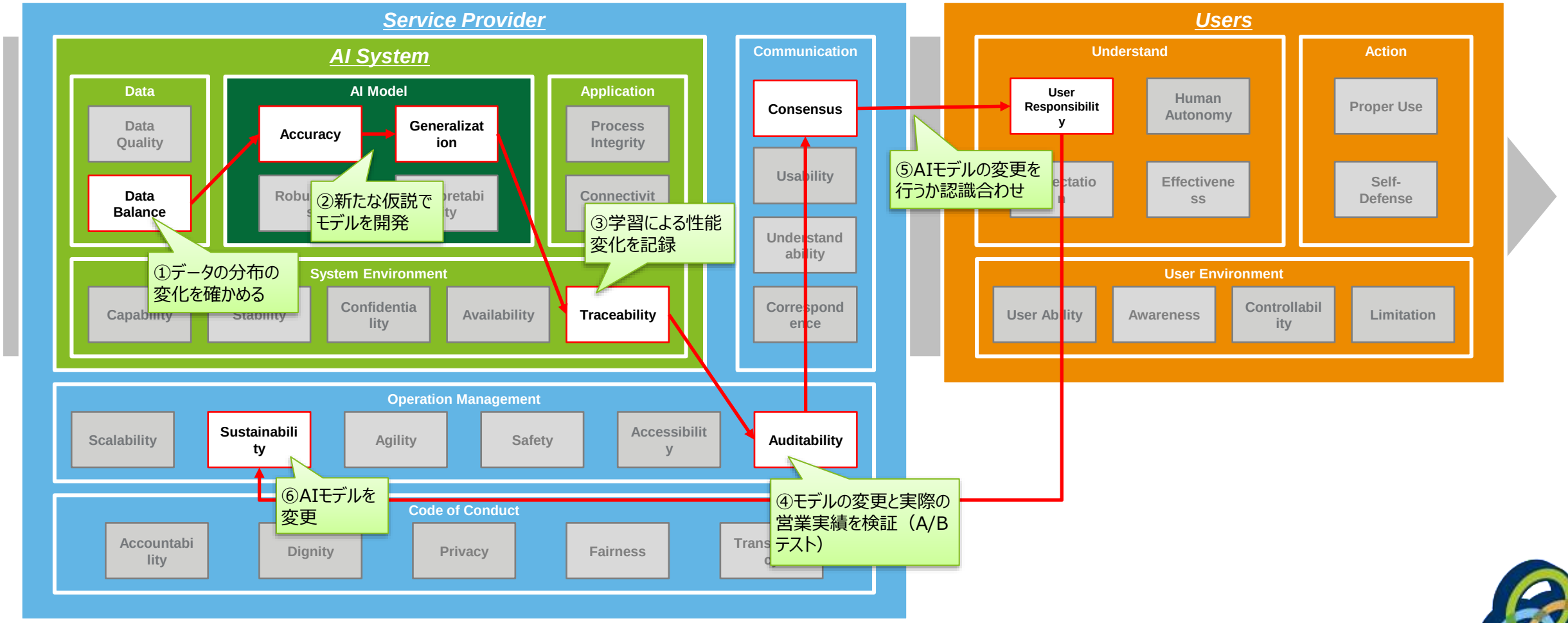
重要なリスクシナリオごとにリスクチェーン(リスク要因の関係性)の検討

- 「リスクシナリオ」ごとにRCModelの各層からリスク要因と関係性(リスクチェーン)を可視化 -

Step3

R001

環境変化による性能劣化
原料組成が当初想定し学習させた範囲から変化して予測性能が低下する



リスクチェーンに従ってリスクコントロールを検討

- リスクチェーンで関連づけられた構成要素においてリスク対応策(コントロール)を検討 -

Step4

R001	**環境変化による性能劣化** 原料組成が当初想定し学習させた範囲から変化して予測性能が低下する

コントロールの内容		
AIシステム （B社AI開発チーム＋C社AI開発部）	**サービスプロバイダ （B社AI開発チーム）**	**ユーザー （A社プラント運転員）**
①【Data Balance】データの分布の変化を確かめる（B社AI開発チーム） ②【Accuracy】【Generalization】新たな仮説に基づいてモデルを再学習・開発する（B社AI開発チーム/C社AI開発部） ③【Traceability】新たなモデルによる学習時の性能を記録する（B社AI開発チーム）	④【Auditability】A/Bテスト等を行い、新旧モデルの性能を検証（B社AI開発チーム） ⑤【Consensus】AIモデルの変更を行うか認識合わせ（B社AI開発チーム） ⑥【Sustainability】AIモデルを変更する（B社AI開発チーム）	⑤【User Responsibility】AIモデルの変更を行うか認識合わせ（A社プラント運転員）

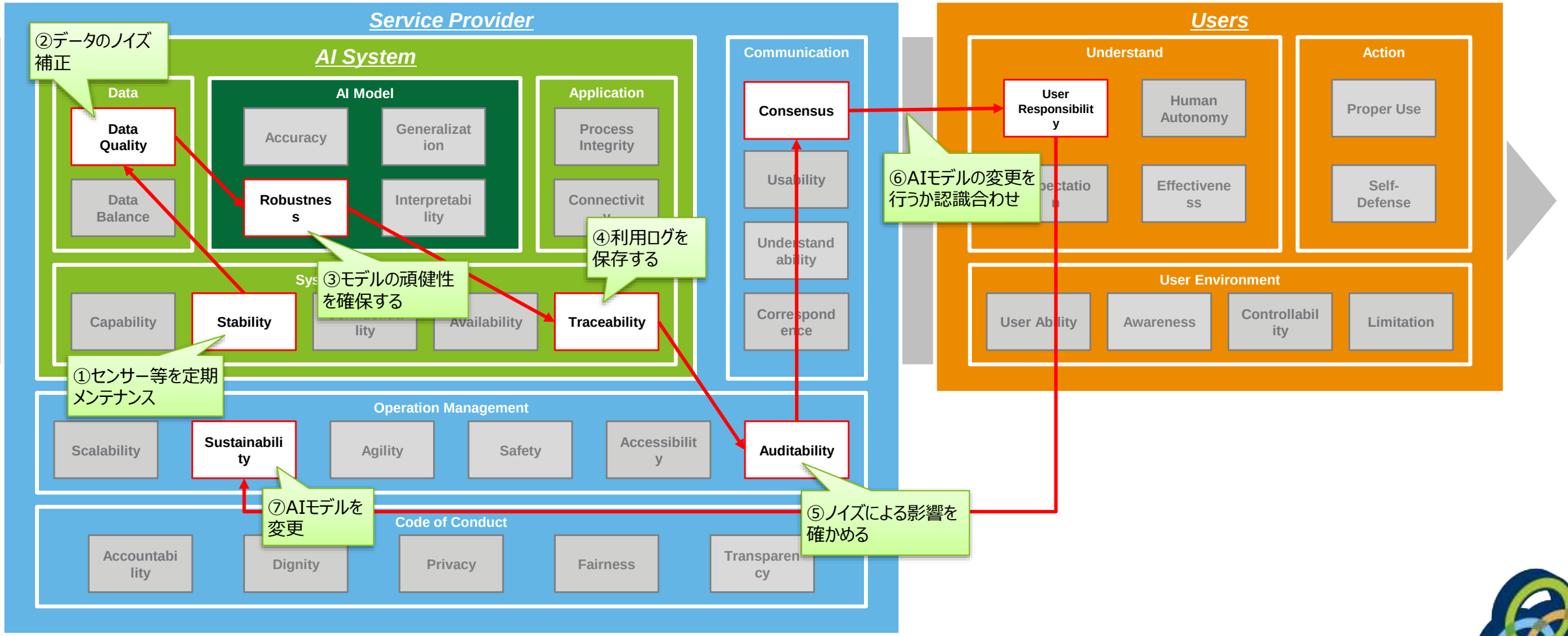
重要なリスクシナリオごとにリスクチェーン(リスク要因の関係性)の検討

- 「リスクシナリオ」ごとにRCModelの各層からリスク要因と関係性(リスクチェーン)を可視化 -

Step3

R002

ノイズによる影響
ノイズの混入により判断精度が安定しない



リスクチェーンに従ってリスクコントロールを検討

- リスクチェーンで関連づけられた構成要素においてリスク対応策(コントロール)を検討 -

Step4

R002	**ノイズによる影響** ノイズの混入により判断精度が安定しない

コントロールの内容		
AIシステム （B社AI開発チーム＋C社AI開発部）	**サービスプロバイダ （B社AI開発チーム）**	**ユーザー （A社プラント運転員）**
①【Stability】センサーを定期的にメンテナンスする（B社AI開発チーム） ②【Data Quality】本番データに対してノイズ補正等で劣化に対処する（B社AI開発チーム） ③【Robustness】モデルの頑健性を高めるように学習を行う（B社AI開発チーム/C社AI開発部） ④【Traceability】利用時のAIの判断結果を保存（B社AI開発チーム）	⑤【Auditability】A/Bテスト等を行い、新旧モデルの性能を検証（B社AI開発チーム） ⑥【Consensus】AIモデルの変更を行うか認識合わせ（B社AI開発チーム） ⑦【Sustainability】AIモデルを変更する（B社AI開発チーム）	⑥【User Responsibility】AIモデルの変更を行うか認識合わせ（A社プラント運転員）

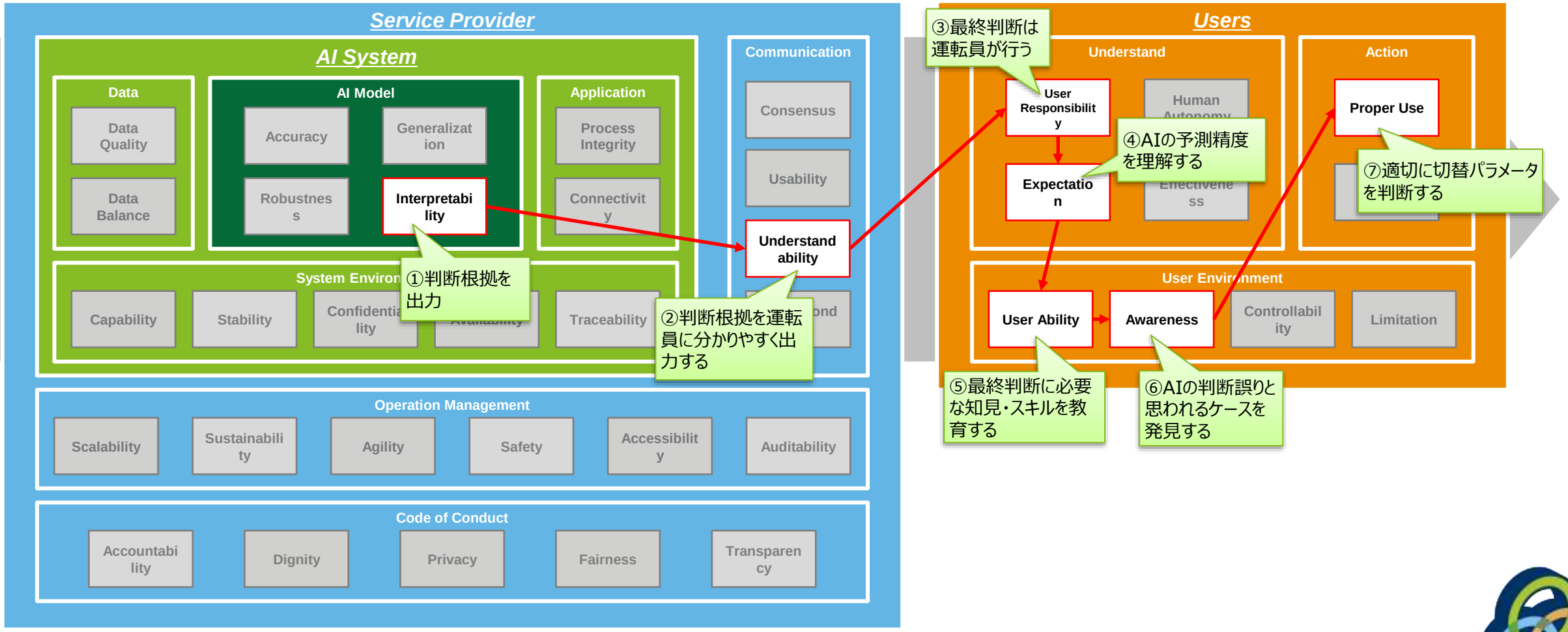
重要なリスクシナリオごとにリスクチェーン(リスク要因の関係性)の検討

- 「リスクシナリオ」ごとにRCModelの各層からリスク要因と関係性(リスクチェーン)を可視化 -

Step3

R006

技術的に誤った解釈
AIの判断根拠が分からず、AIに依存した判断となってしまい、AIの判断誤りをカバーできない



リスクチェーンに従ってリスクコントロールを検討

- リスクチェーンで関連づけられた構成要素においてリスク対応策(コントロール)を検討 -

Step4

R006

技術的に誤った解釈
AIの判断根拠が分からず、AIに依存した判断となってしまい、AIの判断誤りをカバーできない

コントロールの内容		
AIシステム （B社AI開発チーム＋C社AI開発部）	**サービスプロバイダ （B社AI開発チーム）**	**ユーザー （A社プラント運転員）**
①【Interpretability】モデルの判断根拠を出力（B社AI開発チーム/C社AI開発部）	②【Understandability】判断根拠や過去の運転履歴等の参考情報を分かりやすく出力する（B社AI開発チーム）	③【User Responsibility】最終判断は運転員が行うことを理解する（A社プラント運転員） ④【Expectation】モデルの判断精度を理解する（A社プラント運転員） ⑤【User Ability】運転に必要な知見・スキルを教育する（A社プラント運転員） ⑥【Awareness】AIの判断誤りと思われる個所を発見する（A社プラント運転員） ⑦【Proper Use】適切に切替パラメータを判断する（A社プラント運転員）

重要なリスクシナリオごとにリスクチェーン(リスク要因の関係性)の検討

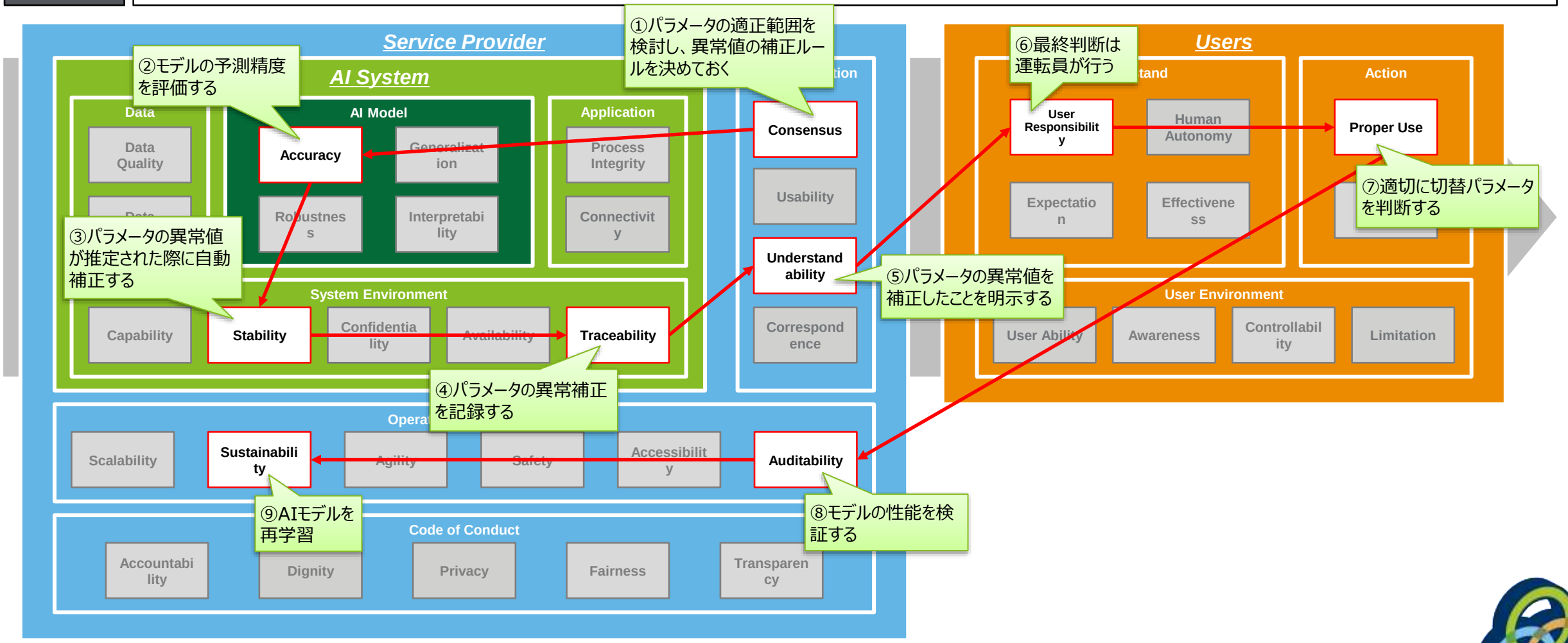
- 「リスクシナリオ」ごとにRCModelの各層からリスク要因と関係性(リスクチェーン)を可視化 -

Step3

R008

異常値の制御

異常な出力が行われた際に正常な範囲に補正できず、プラントの操業に支障をきたす



リスクチェーンに従ってリスクコントロールを検討

- リスクチェーンで関連づけられた構成要素においてリスク対応策(コントロール)を検討 -

Step4

R008	**異常値の制御** 異常な出力が行われた際に正常な範囲に補正できず、プラントの操業に支障をきたす

コントロールの内容		
AIシステム （B社AI開発チーム＋C社AI開発部）	**サービスプロバイダ （B社AI開発チーム）**	**ユーザー （A社プラント運転員）**
②【Accuracy】モデルの予測精度を確保するように学習を行う（B社AI開発チーム/C社AI開発部） ③【Stability】パラメータの異常値が推定された際に自動補正する（B社AI開発チーム） ④【Traceability】パラメータの異常値補正を記録する（B社AI開発チーム）	①【Consensus】パラメータの適正範囲を検討し、異常値の補正ルールを決めておく（A社プラント運転管理部門及びB社AI開発チーム） ⑤【Understandability】パラメータの異常値を補正したことを明示する（B社AI開発チーム） ⑧【Auditability】モデルの性能を検証し再学習の要否を検討する（B社AI開発チーム） ⑨【Sustainability】AIモデルを変更する（B社AI開発チーム）	⑥【User Responsibility】最終判断は運転員が行うことを理解する（A社プラント運転員） ⑦【Proper Use】適切に切替パラメータを判断する（A社プラント運転員）

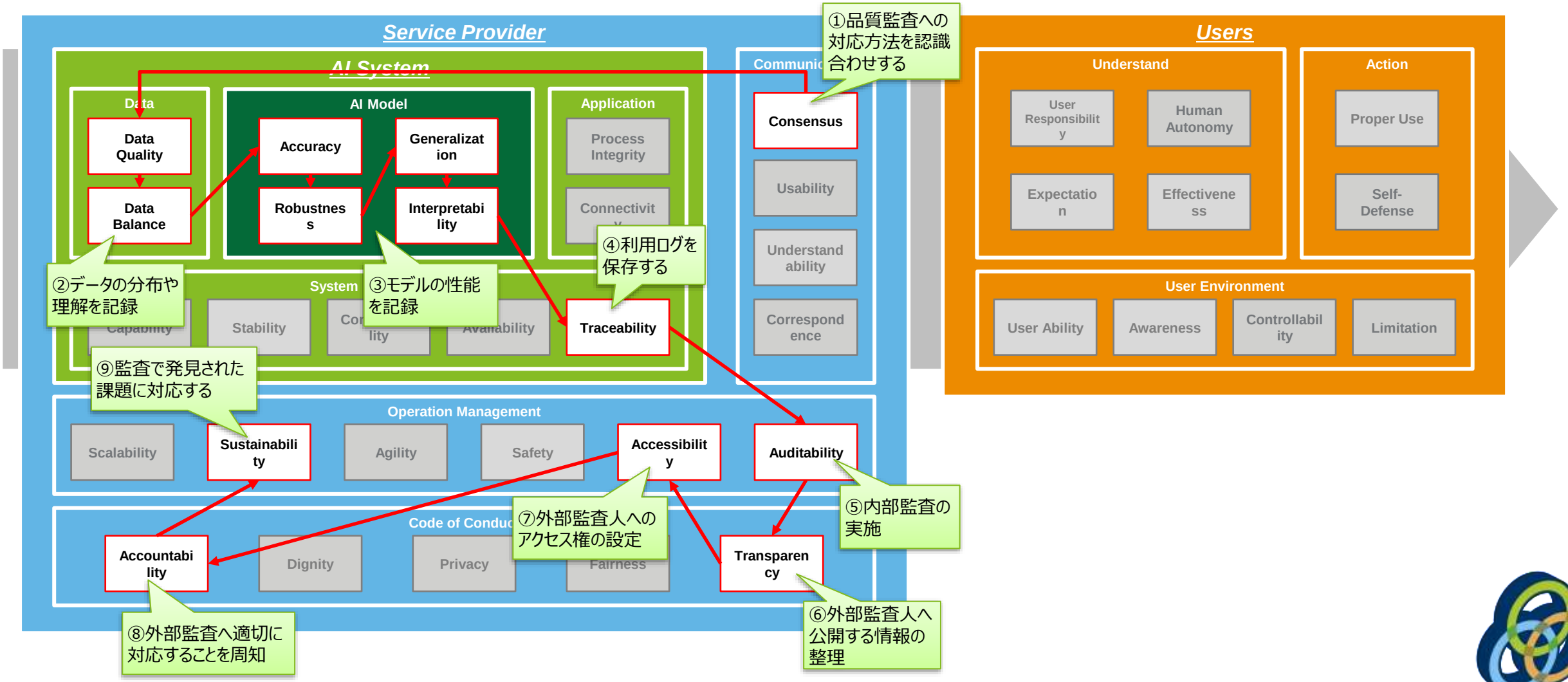
重要なリスクシナリオごとにリスクチェーン(リスク要因の関係性)の検討

- 「リスクシナリオ」ごとにRCModelの各層からリスク要因と関係性(リスクチェーン)を可視化 -

Step3

R011

品質監査への対応
品質監査などを受けた際に適切な説明ができない



リスクチェーンに従ってリスクコントロールを検討

- リスクチェーンで関連づけられた構成要素においてリスク対応策(コントロール)を検討 -

Step4

R011

品質監査への対応
品質監査などを受けた際に適切な説明ができない

コントロールの内容		
AIシステム （B社AI開発チーム＋C社AI開発部）	**サービスプロバイダ （B社AI開発チーム）**	**ユーザー （A社プラント運転員）**
②【Data Quality】【Data Balance】データの分布や理解を記録（B社AI開発チーム） ③【Accuracy】【Robustness】【Generalization】【Interpretability】モデルの性能を記録（B社AI開発チーム） ④【Traceability】AIの判断結果を記録する（B社AI開発チーム）	①【Consensus】パラメータの適正範囲を検討し、異常値の補正ルールを決めておく（A社プラント運転管理部門及びB社AI開発チーム） ⑤【Auditability】内部監査を実施し事前に対応を行う（A社内部監査部） ⑥【Transparency】外部監査人へ公開する情報の整理する（A社プラント運転管理部門及びB社AI開発チーム） ⑦【Accessibility】外部監査人へ必要なアクセス権を設定する（B社AI開発チーム） ⑧【Accountability】外部監査へ適切に対応することを周知（A社プラント運転管理部門及びB社AI開発チーム） ⑨【Sustainability】監査で発見された課題に対応する（B社AI開発チーム）	