RCM
RiskChainModel

リスクチェーンモデル(RCModel)ケース検討事例：
Case09 スマート家電の最適化AI

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（Case09：スマート家電の最適化AI）

- AIサービスが「実現すべき価値・目的」の定義 -

Step1

AIモデルが環境情報やユーザーの行動等を解析し、スマート家電を最適化する。空間に搭載したセンサー情報(ユーザーの位置・状態、温度、湿度、照度、CO2濃度)、オープンデータ(気象情報)、ユーザーからのフィードバック(ストレス、快適度の意見等)を取得してAIモデルが分析を行い、スマート家電機器（スマート冷蔵庫(食材管理、レシピ提案等)、空調、床暖房、空気清浄機、ロボット掃除機、換気システム等）を自動制御する。

【実現すべき価値・目的】

- **快適な空間の維持**
- **健康への悪影響の防止**
- **経済性の維持・向上（ビジネス／消費者双方）**
- **企業の社会的責任**

【AIサービスを用いた実運用の流れ】

任意の空間（世帯・オフィス・施設のフロア等）毎に本AIシステムが導入することができる。対応している家電機器は本AIサービスが開発したA社が公開している。本製品を開発したA社は、導入時にセンシング等の設定を行い、ユーザーとの保守契約に応じて定期点検等を実施する。

システムごとにAIモデルが独立しており、原則として異なるシステム間で学習データの共有などは行われないが、所有するユーザーが同じ場合には希望を受ければ学習データを共有することを可能としている。

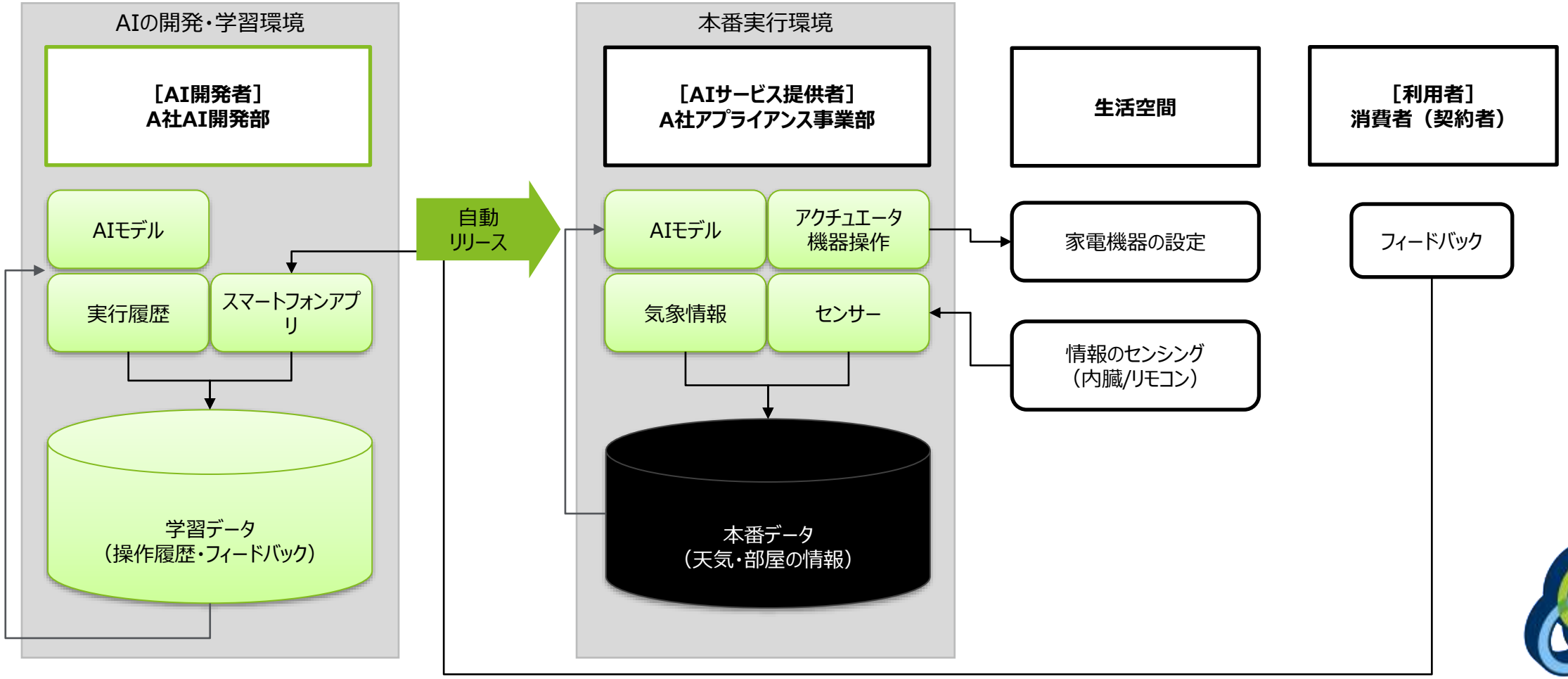
様々な情報を用いたマルチモーダルでの機械学習をベースとしている。期待精度は初期値として一般的な快適度を指標として設定しており、利用後にユーザーのフィードバックを得ながら各システムのAIモデルが独立して学習を行っていく。ユーザーのフィードバックは下記の通りであり、スマートフォンアプリによって入手され、1-2週間毎にAIモデルのチューニングが実施される。

- ストレス：睡眠時などユーザーの近くでアプリを起動しておくと自動で収集される
- 快適度の意見：ユーザーが任意のタイミングでスマートフォンから回答

ケーススタディの概要（Case09：スマート家電の最適化AI）

- システムの全体概要 -

AIシステム	A社AI開発部	スマート家電への最適な設定を予測
AIサービスプロバイダ	A社アプライアンス事業部	スマート家電製品と併せて、利用者にサービス提供
ユーザー	消費者（契約者）	本AIサービスを自らの生活空間で使用



ケーススタディの概要（Case09：スマート家電の最適化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
部屋の情報（広さ・温度・湿度・照度・CO2濃度等）	学習	室内に設置されたセンサーが自動収集（通常は機器内に設置）	A社クラウド環境（全ユーザー含む）	あり（部屋の情報）
ユーザーのフィードバック（ストレス・快適度）	学習	スマートフォンアプリでユーザーから収集（ストレス：自動収集、快適度：入力）	A社クラウド環境（全ユーザー含む）	あり
気象情報	学習	外部から収集	各オープンデータ情報	なし
部屋の情報（広さ・温度・湿度・照度・CO2濃度等）	本番	室内に設置されたセンサーが自動収集する現在情報	A社クラウド環境	あり（部屋の情報）
気象情報	本番	外部から収集	各オープンデータ情報	なし

【AIサービスからの出力内容】

AIサービス利用者	本システムが稼働する空間の居住者
出力結果の内容	家電機器の操作（温湿度の調整、換気）
出力方法（画面／アクチュエータ等）	スマート家電への指示
期待精度（正解率／誤差等）	厳密な設定はないが、ユーザーからのネガティブなフィードバックを最小化する
利用者判断の有無	ユーザーがAIによる自動制御をOFFにすることができる
根拠情報の出力	無
安全性のリスク有無	健康への悪影響
外部AIへの連携	同一ユーザーの複数空間におけるAIシステムで学習データを共有する可能性がある
外部AIへの連携方法・プロトコル	学習データを共有する可能性がある

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	快適な空間の維持	1-1	適切な操作性と性能の確保	■ AIの予測精度 ■ AIの頑健性 ■ 使いやすいUI	R001	不十分な操作性	AIサービスを適切に操作できないことで快適な空間が実現されず、顧客が離れてしまう
					R002	ノイズによる影響	センサーにノイズが混入し、AIの判断精度が劣化する
		1-2	変化への対応	■ 対応機器の管理 ■ 教師データの収集	R003	同居人の変化	居住者の変化（出産・介護等）に対応できない
					R004	環境変化への対応	別の居住地や季節の変化に対応できない
		1-3	ユーザーによる適正利用	■ ユーザー補助	R005	未対応機器の接続	対応していない家電機器が無理やり接続されることで異常な挙動が行われる
2	健康への悪影響の防止	2-1	異常値の制御	■ 安全性確保	R006	異常動作による健康悪化	機器への異常な設定が指示され、利用者側に健康被害が発生する
		2-2	公平性の確保	■ 異常な指示の制御	R007	悪意のあるフィードバック	悪意のあるフィードバックを行うことで、特定の同居人の健康が悪化する(家庭内暴力を助けてしまう)
3	経済性の維持・向上（ビジネス／消費者双方）	3-1	ビジネス側の経済性管理		R008	コスト超過	サービス維持コストが超過する
		3-2	消費者側の経済性管理		R009	光熱費の増加	細かい挙動が連続実行されることで光熱費が増大する
4	企業の社会的責任	4-1	アカウンタビリティ	■ プロセスの説明 ■ 検証可能性	R010	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
		4-2	情報管理	■ データ管理	R011	個人情報保護	学習データや実行履歴情報が外部に流出し悪用される
		4-3	環境への影響	■ 適正な実行	R012	温室効果ガスの増加	細かい挙動が連続実行されることで温室効果ガスの排出量が増加する

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	快適な生活空間の維持（健康等への悪影響の防止）	R001	不十分な操作性	○			●	フィードバックを保存 モデルのアップデート	フィードバックを検証 ユーザー調査	フィードバック
		R002	ノイズによる影響	○	○		●	機器のメンテナンス データのノイズ補正 モデルの頑健性	保守・清掃方法の検討 対応機器メーカーとの連携 利用者への留意	リモコンの配置・清掃 アプリへの警告通知
		R003	同居人の変化	○	○	○	●	学習データの初期化 モデルのアップデート	自動開発環境	食事や環境に係る制約の理解 モデルの初期化機能
		R004	環境変化への対応	○	○		●	学習データの自動選択 モデルのアップデート	自動開発環境	学習データの共有機能
		R005	未対応機器の接続			○		未対応機器の制限	対応機器の明確化	正しい機器を選択
2	健康への悪影響の防止	R006	異常動作による健康悪化	○			●	判断根拠 検証可能性 異常値の自動補正	安全基準値の理解 マニュアル操作への切替UI 対応機器メーカーとの連携 異常の原因確認	異常動作の相談
		R007	悪意のあるフィードバック			○	●	異常なフィードバックの制限 異常値の自動補正 検証可能性	安全基準値の理解 悪意のある利用の検証 法務等との対応の検討	異常なフィードバックへの警告
3	経済性の維持・向上（ビジネス／消費者双方）	R008	コスト超過						コスト管理	
		R009	光熱費の増加					省エネ技術の採用		コスト管理
4	企業の社会的責任	R010	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	
		R011	個人情報保護					データ保護	セキュリティ管理	
		R012	温室効果ガスの増加					過剰操作の防止 省エネ技術の採用		

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 -
A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

消費者庁

A社法務・コンプラ

- 悪意のあるユーザーへの対応

A社品質管理

- 内部監査の実施
- 外部監査の対応

- AIサービスプロバイダ -
A社アプライアンス事業部

- 自動開発環境の準備
- 安全基準値の理解
- マニュアル操作への切替UI
- フィードバックの検証
- 対応機器の明確化
- 点検・メンテナンス方法の検討
- 異常な状態の明確化
- システム稼働監視
- セキュリティ管理

A社AI開発部

- モデルの予測性能
- モデルの頑健性
- 学習データの確保
- 学習データの自動選定
- 異常値の自動補正
- 自動アラート処理
- センサーの仕様・プロトコル

A社カスタマー担当

- ユーザー調査
- メンテナンスの対応
- 悪意のあるユーザーへの対応

- ユーザー -
利用者（契約者）

- 正しい利用

(アプリ)

- 快適度のフィードバック
- モデルの初期化機能
- 学習データの共有設定機能
- アラート機能

各対応機器メーカー

- 点検・メンテナンスにおける連携

A社クラウドサービス部

- ユーザーのフィードバックの記録
- 判断根拠の記録

同居人

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 -

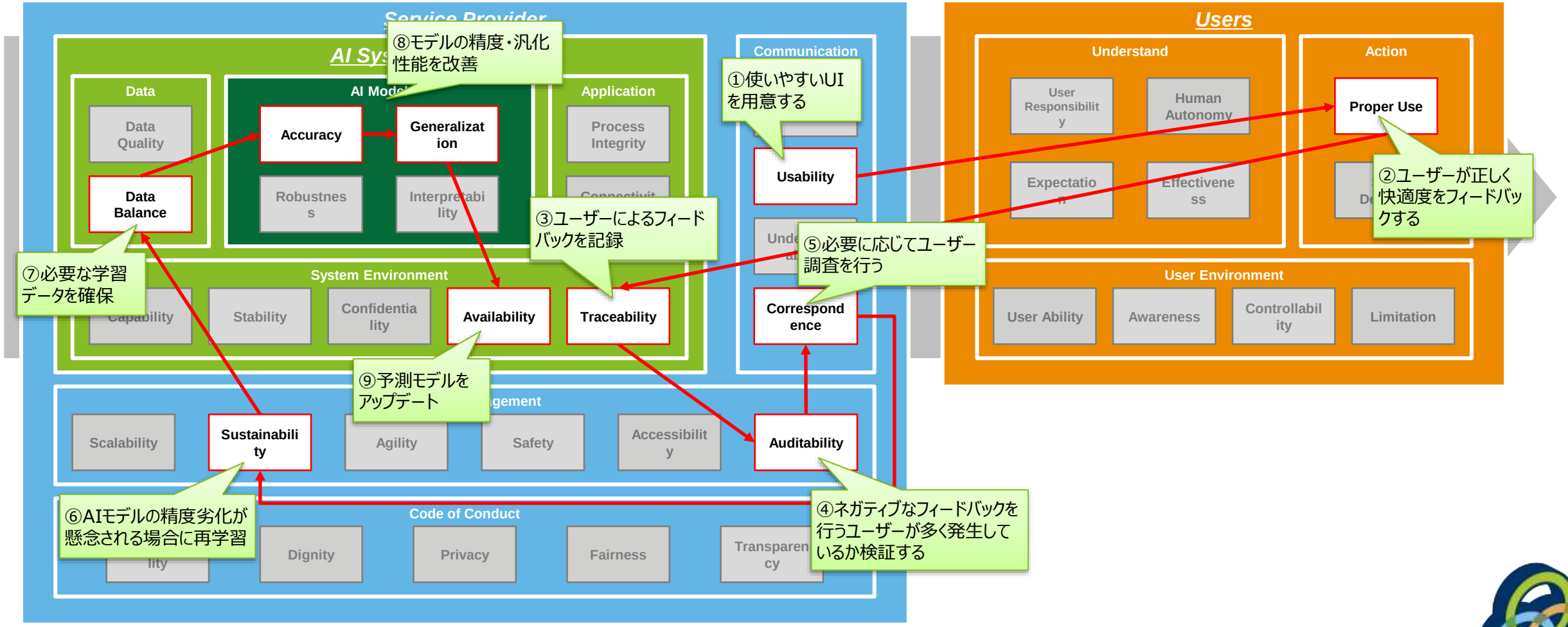
重要なリスクシナリオごとにリスクチェーン(リスク要因の関係性)の検討

- 「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

不十分な操作性
AIサービスを適切に操作できないことで快適な空間が実現されず、顧客が離れてしまう



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	**不十分な操作性** AIサービスを適切に操作できないことで快適な空間が実現されず、顧客が離れてしまう

コントロールの内容		
AIシステム （A社AI開発部）	**サービスプロバイダ （A社アプライアンス事業部）**	**ユーザー （消費者（契約者））**
③【Traceability】利用者のフィードバックを記録する（A社クラウドサービス部） ⑦【Data Balance】十分な学習データを確保する（A社AI開発部） ⑧【Accuracy】【Generalization】学習時にモデルの正解率及び汎化性能を十分に確保する（A社AI開発部） ⑨【Availability】予測モデルをアップデートする（A社AI開発部）	①【Usability】使いやすいUIを設計する（A社アプライアンス事業部） ④【Auditability】ネガティブなフィードバックを行うユーザーが多く発生しているか検証する（A社アプライアンス事業部） ⑤【Correspondence】必要に応じてユーザー調査を行う（A社アプライアンス事業部／A社カスタマー担当） ⑥【Sustainability】十分な精度を確保するためにAIモデルの再学習を依頼する（A社アプライアンス事業部）	②【Proper Use】スマートフォンアプリから快適度をフィードバックする（消費者）

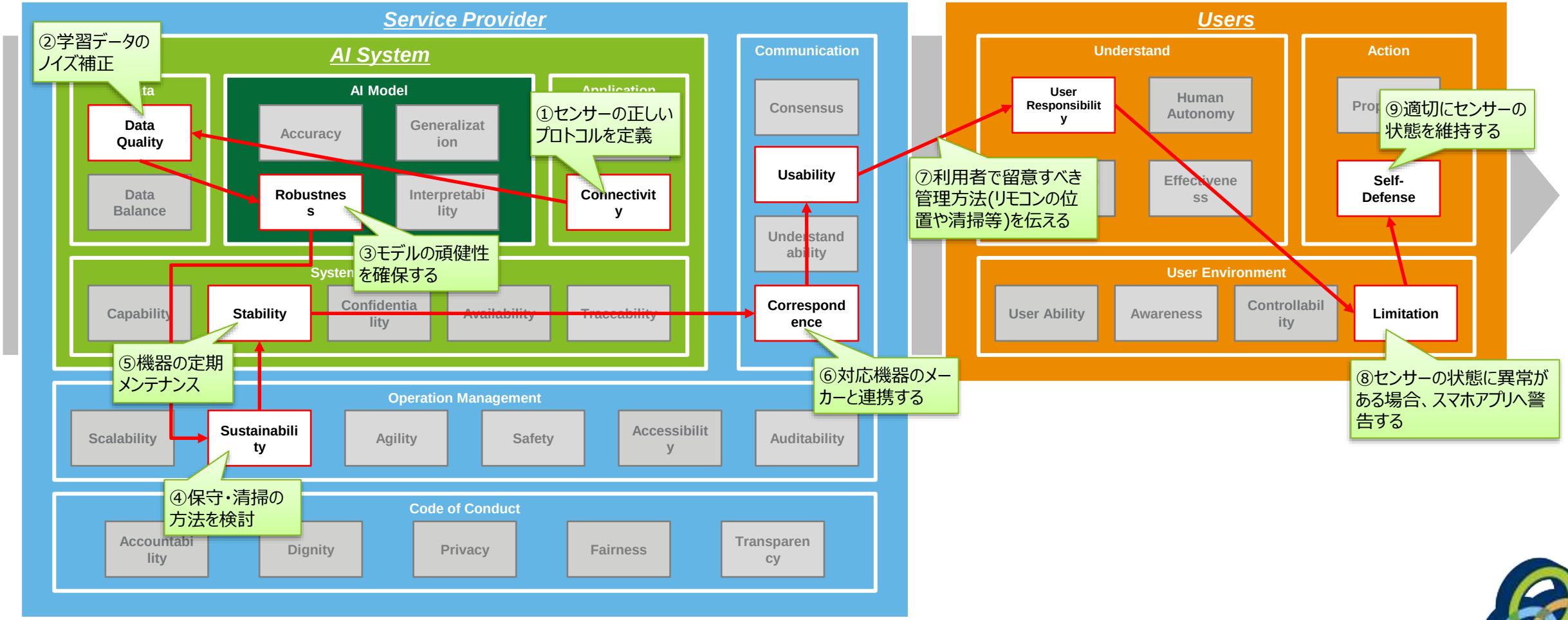
重要なリスクシナリオごとにリスクチェーン(リスク要因の関係性)の検討

- 「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
センサーにノイズが混入し、AIの判断精度が劣化する



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

R002	**ノイズによる影響** センサーにノイズが混入し、AIの判断精度が劣化する

コントロールの内容		
AIシステム （A社AI開発部）	**サービスプロバイダ （A社アプライアンス事業部）**	**ユーザー （消費者（契約者））**
①【Connectivity】センサーの正しい仕様・プロトコルに従って実装する（A社AI開発部） ②【Data Quality】学習データに対してノイズ補正等で劣化に対処する（A社AI開発部） ③【Robustness】モデルの頑健性を高めるように学習を行う（A社AI開発部） ⑤【Stability】スマート家電のセンサー等を適切にメンテナンスする（A社AI開発部）	④【Sustainability】スマート家電の保守・清掃の方法を検討する（A社アプライアンス事業部） ⑥【Correspondence】メンテナンスにおいて対応機器メーカーと連携する（A社アプライアンス事業部＋対応機器メーカー） ⑦【Usability】利用者で留意すべき管理方法(リモコンの位置や清掃等)を伝える（A社カスタマー担当）	⑦【User Responsibility】利用者で留意すべき管理方法(リモコンの位置や清掃等)を理解する（消費者） ⑧【Limitation】センサーの状態に異常がある場合、スマホアプリへ警告する（A社アプライアンス事業部） ⑨【Proper Use】適切にセンサーの状態を維持する（消費者）

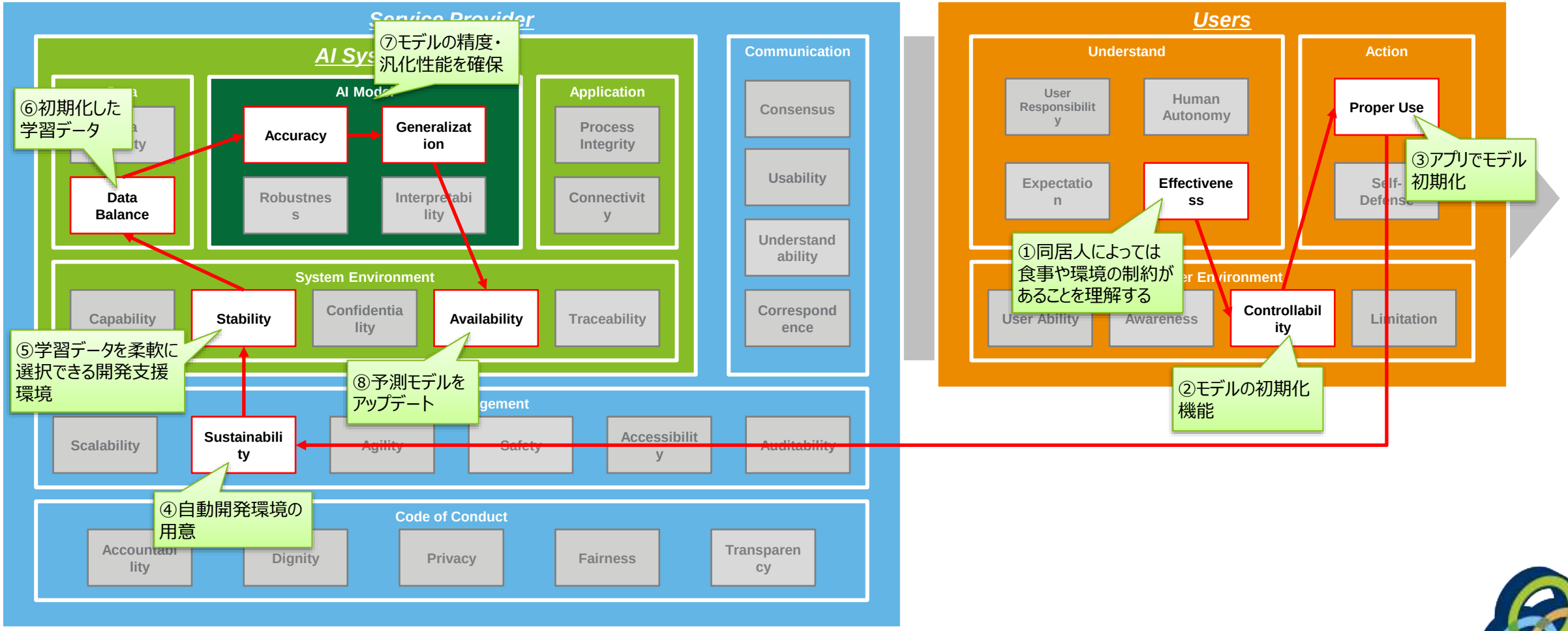
重要なリスクシナリオごとにリスクチェーン(リスク要因の関係性)の検討

- 「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

R003

同居人の変化
居住者の変化（出産・介護等）に対応できない



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

R003	**同居人の変化** 居住者の変化（出産・介護等）に対応できない

コントロールの内容		
AIシステム （A社AI開発部）	**サービスプロバイダ （A社アプライアンス事業部）**	**ユーザー （消費者（契約者））**
⑤【Stability】学習データを柔軟に選択できる開発支援環境を用意する（A社AI開発部） ⑥【Data Balance】初期化した学習データを用意する（A社AI開発部） ⑦【Accuracy】【Generalization】学習時にモデルの正解率及び汎化性能を十分に確保する（A社AI開発部） ⑧【Availability】予測モデルをアップデートする（A社AI開発部）	④【Sustainability】自動開発環境を用意する（A社アプライアンス事業部）	①【Effective】同居人によっては食事や環境等の制約が発生することを理解する（消費者） ②【Controllability】モデルの初期化機能をアプリに搭載する（A社アプライアンス事業部） ③【Proper Use】モデルの初期化を行う（消費者）

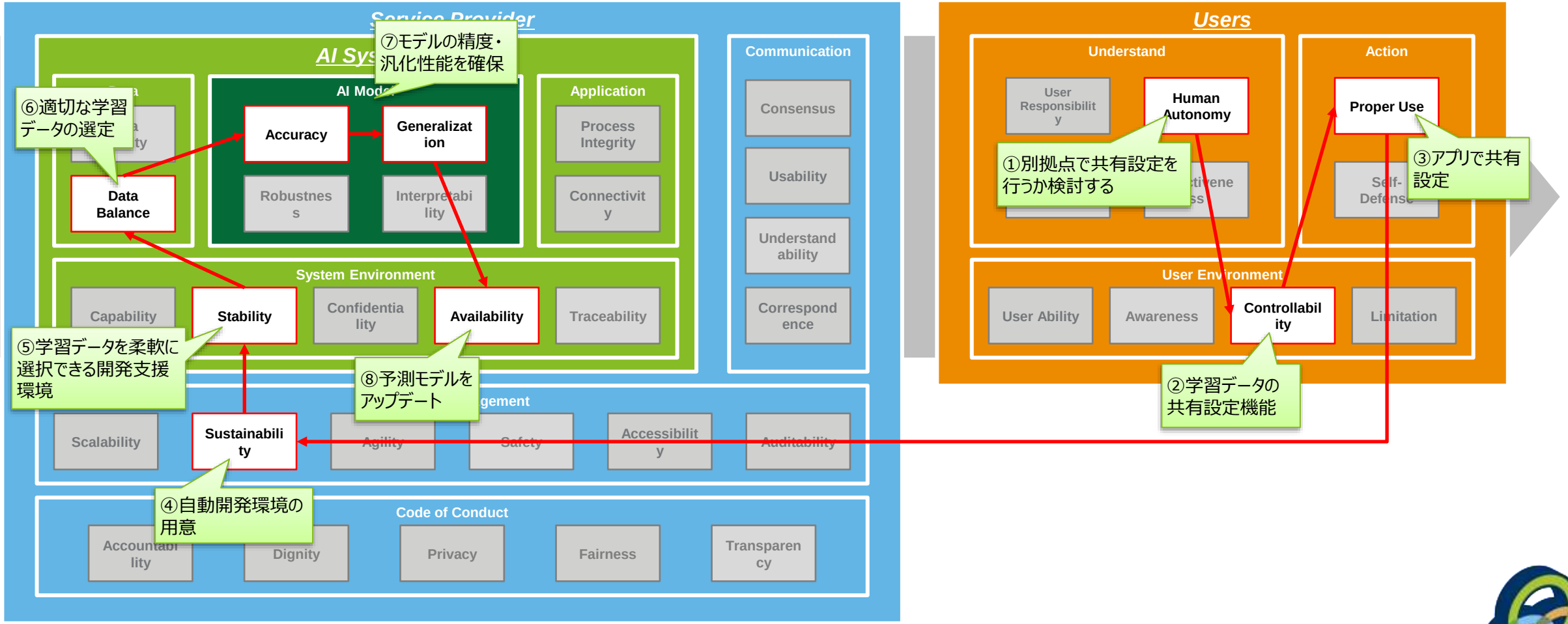
重要なリスクシナリオごとにリスクチェーン(リスク要因の関係性)の検討

- 「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

R004

環境変化への対応
別の居住地や季節の変化に対応できない



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

R004	**環境変化への対応** 別の居住地や季節の変化に対応できない

コントロールの内容		
AIシステム （A社AI開発部）	**サービスプロバイダ （A社アプライアンス事業部）**	**ユーザー （消費者（契約者））**
⑤【Stability】学習データを柔軟に選択できる開発支援環境を用意する（A社AI開発部） ⑥【Data Balance】適切な学習データを確保する（A社AI開発部） ⑦【Accuracy】【Generalization】学習時にモデルの正解率及び汎化性能を十分に確保する（A社AI開発部） ⑧【Availability】予測モデルをアップデートする（A社AI開発部）	④【Sustainability】自動開発環境を用意する（A社アプライアンス事業部）	①【Human Autonomy】別拠点で共有設定を行うか検討する（消費者） ②【Controllability】別拠点での学習データ共有設定機能をアプリに搭載する（A社アプライアンス事業部） ③【Proper Use】アプリで別拠点での学習データ共有設定を行う（消費者）

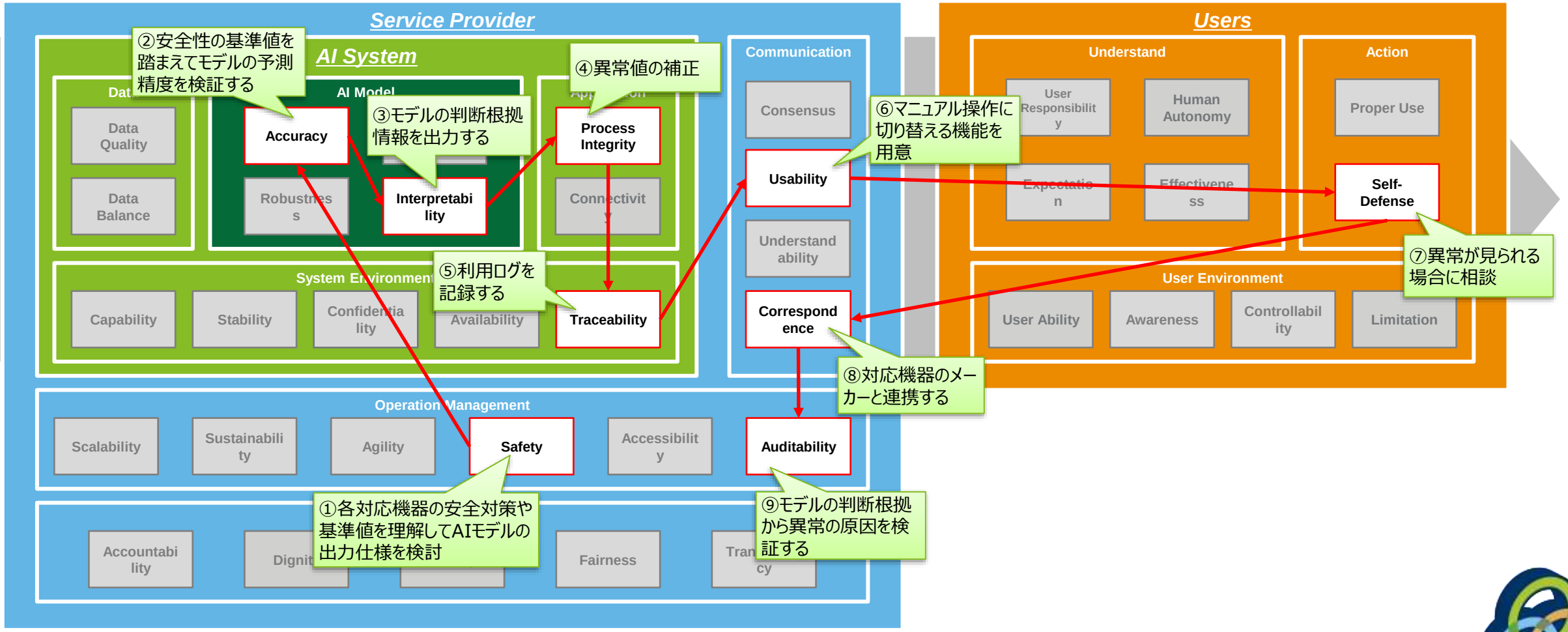
重要なリスクシナリオごとにリスクチェーン(リスク要因の関係性)の検討

- 「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

異常動作による健康悪化
機器への異常な設定が指示され、利用者側に健康被害が発生する



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

R006	**異常動作による健康悪化** 機器への異常な設定が指示され、利用者側に健康被害が発生する

コントロールの内容		
AIシステム （A社AI開発部）	**サービスプロバイダ （A社アプライアンス事業部）**	**ユーザー （消費者（契約者））**
②【Accuracy】安全性の基準値を踏まえてモデルの予測精度を検証する（A社AI開発部） ③【Interpretability】モデルの判断根拠を出力する（A社AI開発部） ④【Process Integrity】異常値を自動補正する（A社AI開発部） ⑤【Traceability】利用ログを判断根拠と併せて保存する（A社AI開発部）	①【Safety】各対応機器の安全対策や基準値を理解してAIモデルの出力仕様を検討する（A社アプライアンス事業部） ⑥【Usability】マニュアル操作に切り替える機能・UIを用意する（A社アプライアンス事業部） ⑧【Correspondence】対応機器のメーカーと連携して機器のメンテナンスを行う（A社アプライアンス事業部＋対応機器メーカー） ⑨【Auditability】モデルの判断根拠から異常の原因を検証する（A社アプライアンス事業部）	⑦【Self-Defense】異常が見られる場合にサービス提供者に相談する（消費者）

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

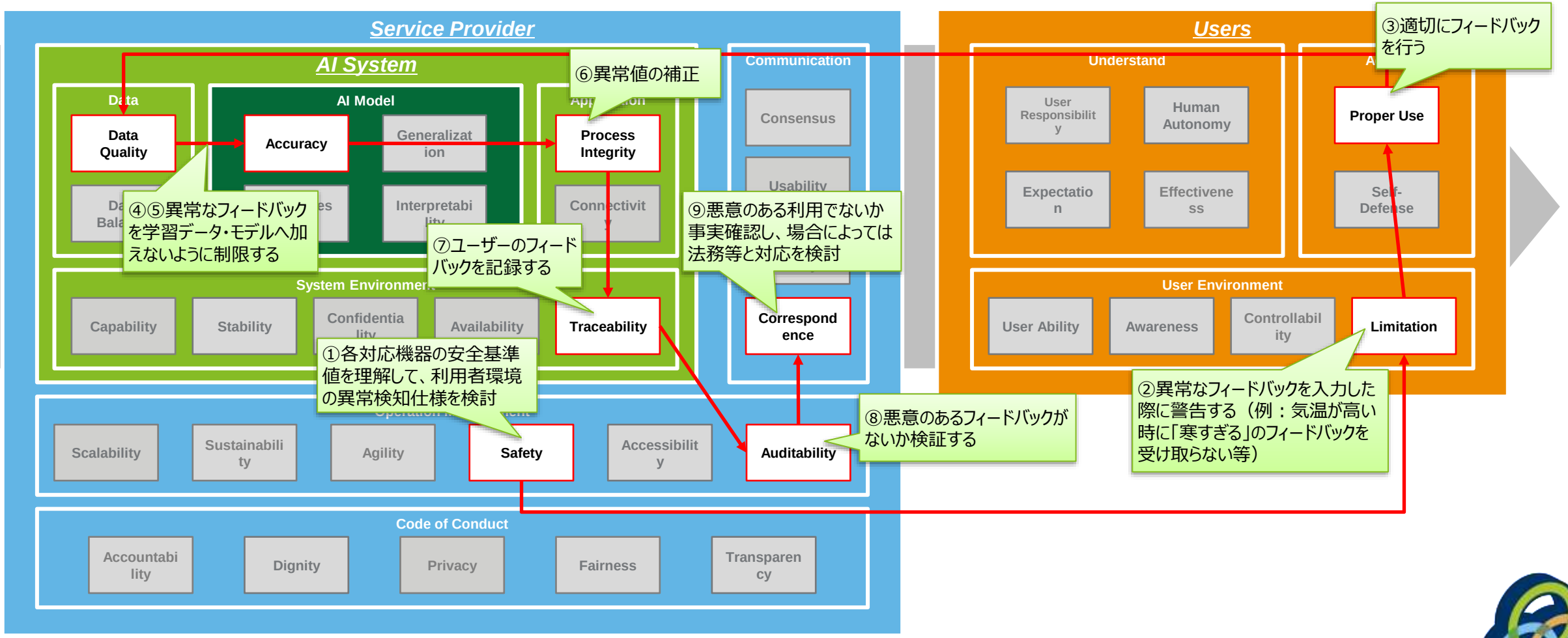
- 「リスクシナリオ」ごとにRCModelの各層からリスク要因と関係性(リスクチェーン)を可視化 -

Step3

R007

悪意のあるフィードバック

誤ったフィードバックを行うことで、特定の同居人の健康が悪化する(家庭内暴力を助けてしまう)



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

R007	**悪意のあるフィードバック** 誤ったフィードバックを行うことで、特定の同居人の健康が悪化する(家庭内暴力を助けてしまう)

コントロールの内容		
AIシステム （A社AI開発部）	**サービスプロバイダ （A社アプライアンス事業部）**	**ユーザー （消費者（契約者））**
④【Data Quality】異常なフィードバックを学習データへ加えないように制限する（A社AI開発部） ⑤【Accuracy】異常なフィードバックを予測モデルへ加えないように制限する（A社AI開発部） ⑤【Process Integrity】異常値を自動補正する（A社AI開発部） ⑥【Traceability】ユーザーのフィードバックを記録する（A社AI開発部）	①【Safety】各対応機器の安全基準値を理解して、利用者環境の異常検知仕様を検討する（A社アプライアンス事業部） ⑦【Auditability】悪意のあるフィードバックがないか検証する（A社アプライアンス事業部） ⑧【Sustainability】悪意のあるユーザーを認識した際に事実確認し、必要な対応を検討する（A社カスタマー担当／A社法務・コンプライアンス部）	②【Limitation】異常なフィードバックを入力した際に警告する（例：気温が高い時に「寒すぎる」のフィードバックを受け取らない等）（A社アプライアンス事業部） ③【Proper Use】適切にフィードバックを行う（消費者）