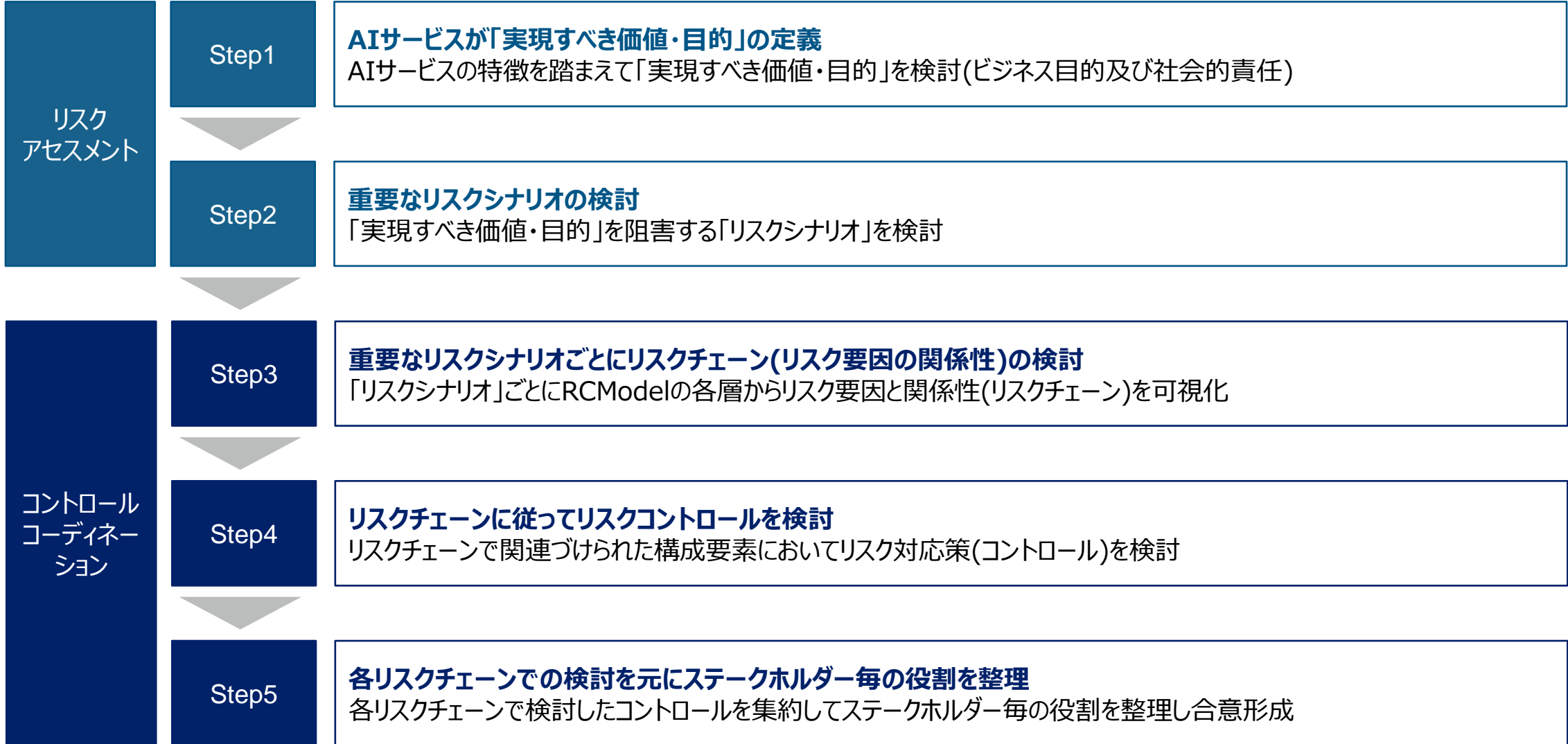


# リスクチェーンモデル(RCModel)ケース検討事例： Case05 道案内ロボット



# ケース検討のステップ



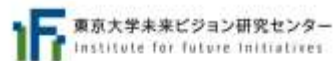


# ケース事例（AIサービスとリスクコーディネーション研究会）

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/>



研究 人材育成 メンバー ニュース イベント 出版物 IFIについて

## リスクチェーンモデルの使い方

[リスクチェーンモデル\(RCModel\)ガイド Ver1.0](#)

## ケース事例

※あくまでサンプルとしてのケース検討例であり、特定の企業のAIサービスに対して問題提起を行うものや保証を与えるものではないことにご留意ください。

[Case01.採用AI\(2021/07\)](#)

[Case02.無人コンビニ\(2021/07\)](#)

[Case03.送電線の外観検査ドローン\(2021/07\)](#)

[Case04.不良品検知AI\(2021/07\)](#)

[Case05.道案内ロボット\(2021/07\)](#)

[Case06.再犯可能性の検証AI\(2021/07\)](#)

# ケーススタディの概要



# ケーススタディの概要 (Case05 : 道案内ロボット)

- AIサービスが「実現すべき価値・目的」の定義 -

Step1

空港内の道案内ロボットである。案内の範囲は空港内の施設・ゲートなどであり、複数の目的特化型ディープラーニングモデルを組合せてサービスを実現している。

A空港内にはロボットを6体配置している。ロボットには複数のAIモデルや自動処理が組み込まれており、特定の利用者1名と対話することでサービスを提供する。対応している言語は日本語・英語・中国語・韓国語である。

ディープラーニングのモデル及びロボットの躯体の開発は委託先であるX社に依頼しているが、学習環境・本番環境共にA空港内に限らせている。逐次学習は実施していないが、運用後に精度が著しく劣化するような傾向が見られた場合、X社が必要に応じて追加学習や再学習を要請する。

## 【実現すべき価値・目的】

- 利用者の満足度向上(安心できる道案内)
- 安全性の確保
- 空港職員の負担軽減
- 事業者の社会的責任

## 【ロボットの状態と処理の流れ】

ロボットは下記の状態を切り替えながら、サービス提供を行う。

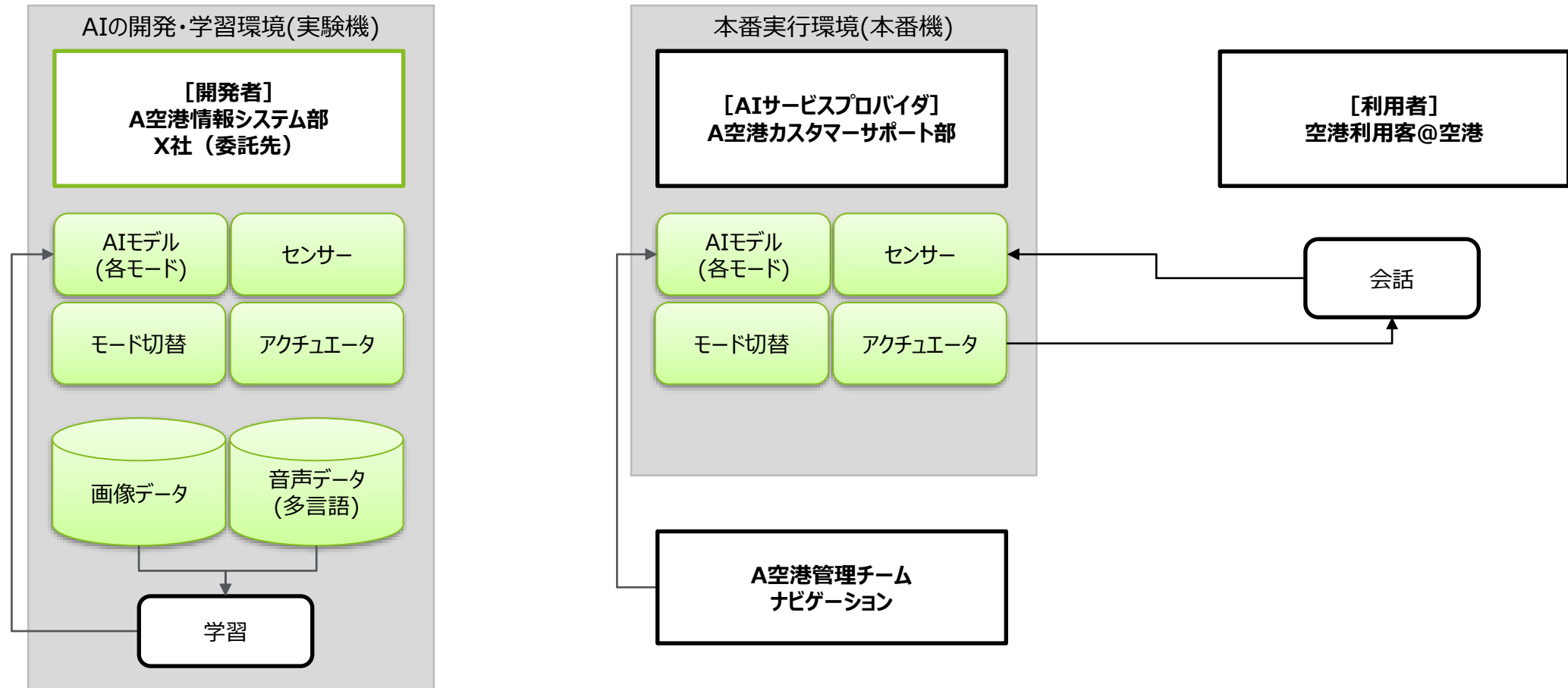
- ① [探索モード]会話する言語を認識する (物体検知・音声認識)
- ② [認識モード]案内すべき空港利用者に話しかける (物体検知)
- ③ [会話モード]案内すべき目的地が判明するまで利用者との会話を行う (音声認識・自然言語処理)
- ④ [確認モード]案内すべき目的地候補が判明し、利用者に確認を取る (音声認識・自然言語処理)
- ⑤ [道案内モード]目的地まで案内を行う



# ケーススタディの概要 (Case05 : 道案内ロボット)

- システムの全体概要 -

AIシステム	A空港情報システム部 + X社 (委託先)	ロボット及びAIモデルの開発・システム運用
AIサービスプロバイダ	A空港カスタマーサポート部	A空港におけるロボット運用を含む顧客対応
ユーザー	空港利用客	ロボットを利用して目的地へ向かう



# ケーススタディの概要 (Case05 : 道案内ロボット)

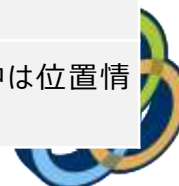
- AIサービスの入出力 -

## 【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
人物の画像データ	学習	空港の監視カメラ等から常時収集	A空港システム管理部 (空港内開発環境)	あり
人物の音声データ	学習	空港の内で常時音声を収集	A空港システム管理部 (空港内開発環境)	あり
人物の会話データ	学習	空港の内で常時音声を収集	A空港システム管理部 (空港内開発環境)	あり
人物の画像データ	本番	利用の都度	A空港システム管理部 (空港内本番環境)	あり
人物の音声データ	本番	利用の都度	A空港システム管理部 (空港内本番環境)	あり
人物の会話データ	本番	利用の都度	A空港システム管理部 (空港内本番環境)	あり

## 【AIサービスからの出力内容】

出力結果	出力方法 (画面/アクチュエータ等)	期待精度	利用者判断の有無	根拠情報の出力	安全性のリスク	外部との連携・プロトコル
①話しかける言語の選択	会話する言語の選択 (外部出力しない)	正解率 (75%)	無	無	無	無
②話しかける	利用者に向けてロボットから定義済の音声を出力 (①で認識した言語を選択)	正解率 (75%)	有	無	無	無
③会話	利用者に向けて会話を出力 (①で認識した言語を選択)	正解率 (75%)	有 (会話: 再度話しかける)	利用者の会話をテキスト表示	無	無
④案内すべき場所の選択	利用者案内すべき場所	正解率 (75%)	有 (会話: 再度話しかける)	目的地をテキスト表示	無	無
⑤案内の開始	目的地に向けた移動 ※目的地に到着したら、自動処理で利用者に案内終了を伝え、元の配置に戻る。	正解率 (75%)	有 (会話: 再度話しかける)	目的地をテキスト表示し、案内モードであると表示	有 (人体との衝突/危険区域の進行)	有 (案内中は位置情報と連携)



# 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	利用者の満足度向上(安心できる道案内)	1-1	サービスレベルの維持 ■ AIの予測精度 ■ AIの頑健性	R001	誤動作	誤った目的地に案内する／制御不能な状態となり暴走してしまう
				R002	不十分なサービス	案内を開始しても出発に間に合わない
		1-2	環境変化への対応 ■ AIの頑健性 ■ IoT	R003	ノイズによる影響	空港内の環境(混雑度合等)のノイズの影響で、ロボットが正しい判断を実施できない
				R004	フライト変更への対応	出発ゲートや出発時刻の変更に対応できない
		1-3	公平性の確保 ■ 多言語対応	R005	未対応言語への対応	対応していない言語・方言等でサービスを要求された際に、正しく対応できない
2	安全性の確保	2-1	安全確保 ■ 接触防止	R006	人物への危害	案内中に別の人物と接触して負傷させてしまう
		2-2	侵入可能エリアの把握 ■ 進入禁止エリアの設定	R007	危険エリアへの誘導	侵入できない場所に対して、誘導してしまう
		2-3	障害・負傷を抱えた方への対応 ■ 利用者の状態認識	R008	利用者の身体への負担	車椅子や障害・負傷を持つ利用者へ身体等への負担のかかる経路を案内してしまう
3	空港職員の負担軽減	3-1	職員の負荷軽減	R009	空港内職員の負担増加	ロボットが正しく稼働しないことで職員の負担が増加
4	事業者の社会的責任	4-1	アカウントビリティ ■ プロセスの説明 ■ 検証可能性	R010	トラブル発生時の対応	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	プライバシーの侵害	ロボットが認識した個人情報(画像・音声)が不正に利用される



# 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 利用者の満足度向上(安心できる道案内)	R001	誤動作	○			●	予測性能の確保 目的地の修正 サポートへの自動連携	目的地の表示 利用者の遠隔サポート	目的地の確認・修正
	R002	不十分なサービス	○		○	●	時刻データ等の更新 速度の調整 サポートへの自動連携	航空会社との調整 利用者との連携	最終判断
	R003	ノイズによる影響	○	○		●	マイクのメンテナンス モデルの頑健性	モデルの性能検証 再学習の実施	
	R004	フライト変更への対応	○	○		●	目的地の自動修正 サポートへの自動連携	スケジュール変更の反映 目的地の表示 利用者の遠隔サポート	目的地の確認・修正
	R005	未対応言語への対応	○		○	●	モデルの頑健性 未対応言語の表示	対応言語の明確化 利用者の遠隔サポート	対応言語の認識 サポートチームへ連絡
2 安全性の確保	R006	人物への危害	○		○	●	衝突防止機能 移動時の効果音 衝突時のアラート	利用者のサポート 事故原因の調査・開示 負傷した利用者の補償	接触の注意
	R007	危険エリアへの誘導	○	○		●	立入禁止エリアの更新 進行方向の予測 移動状態の連携	立入禁止区画の管理 リモートでの進路補正 緊急停止機能	進路補正/緊急停止
	R008	利用者の身体への負担	○		○	●	※R002と同じ	※R002と同じ	※R002と同じ
3 空港職員の負担軽減	R009	空港内職員の負担増加	○				モデルの予測性能	遠隔サポートの状況把握 コスト管理	
4 事業者の社会的責任	R010	トラブル発生時の対応	○				説明可能性	障害対応フロー	
	R011	プライバシーの侵害					データ保護	職業倫理の教育	



# 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 -

- 責任者 -

A空港) 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

A空港) 空港管理部

- スケジュール変更の反映
- 航空会社との調整

A空港) 安全管理部

- 事故原因の調査

- AIサービスプロバイダ -  
A空港) カスタマーサポート部

- 利用者との連携
- 利用者の遠隔サポート
- リモートでの進路補正
- 緊急停止
- 対応言語の明確化
- 目的地の表示
- スケジュール変更の反映
- 立入禁止区画の管理
- モデルの性能検証
- マイクのメンテナンス
- 再学習
- 障害対応フロー
- 事故原因の調査・開示
- 負傷した利用者の補償
- 職業倫理の教育

X社) AI開発部

- モデルの予測性能
- 目的地の修正機能
- アクチュエータの性能
- 衝突防止機能
- 移動時の効果音
- 衝突時のアラート
- 進行方向の予測
- 未対応言語の表示
- サポートへの自動連携

A空港) 情報システム部門

- 時刻データの更新
- 立入禁止エリアの更新
- 利用ログの保存

- ユーザー -  
空港利用客@空港

- 目的地の確認・修正
- 最終判断
- 接触の注意
- 進路補正/緊急停止
- 対応言語の認識
- サポートチームへの連絡

その他の利用客

- 接触の注意



# 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

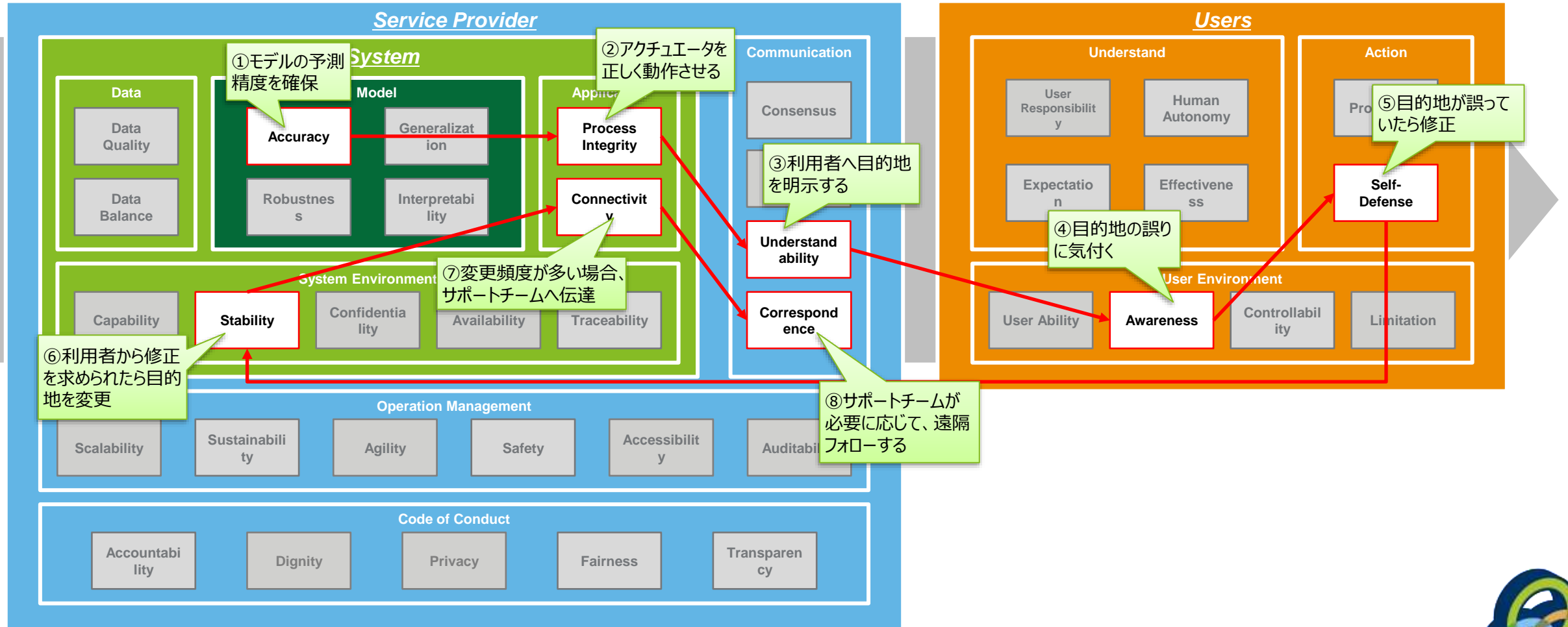
Step3

- 「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 -

R001

## 誤動作

誤った目的地に案内する／制御不能な状態となり暴走してしまう



# リスクチェーンに従ってリスクコントロールを検討

Step4

- 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 -

R001

## 誤動作

誤った目的地に案内する／制御不能な状態となり暴走してしまう

### コントロールの内容

AIシステム (A空港情報システム部+X社)	サービスプロバイダ (A空港カスタマーサポート部)	ユーザー (空港利用客)
<p>①【Accuracy】モデルの予測精度を確保する (X社)</p> <p>②【Process Integrity】アクチュエータを正しく動作させる (A空港情報システム部+X社)</p> <p>⑥【Stability】利用者から修正を求められたら目的地を変更する (ロボット)</p> <p>⑦【Connectivity】変更頻度が多い場合、正しく認識できていない可能性があるため、サポートチームへ自動伝達する (ロボット)</p>	<p>③【Understandability】利用者へ目的地をロボットへ表示する (A空港カスタマーサポート部)</p> <p>⑧【Correspondence】サポートチームが必要に応じて、利用者を遠隔フォローする (A空港カスタマーサポート部)</p>	<p>④【Awareness】ロボットが認識している目的地の誤りに気づく (空港利用客)</p> <p>⑤【Self-Defense】目的地が誤っていたら修正する (空港利用客)</p>



# 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

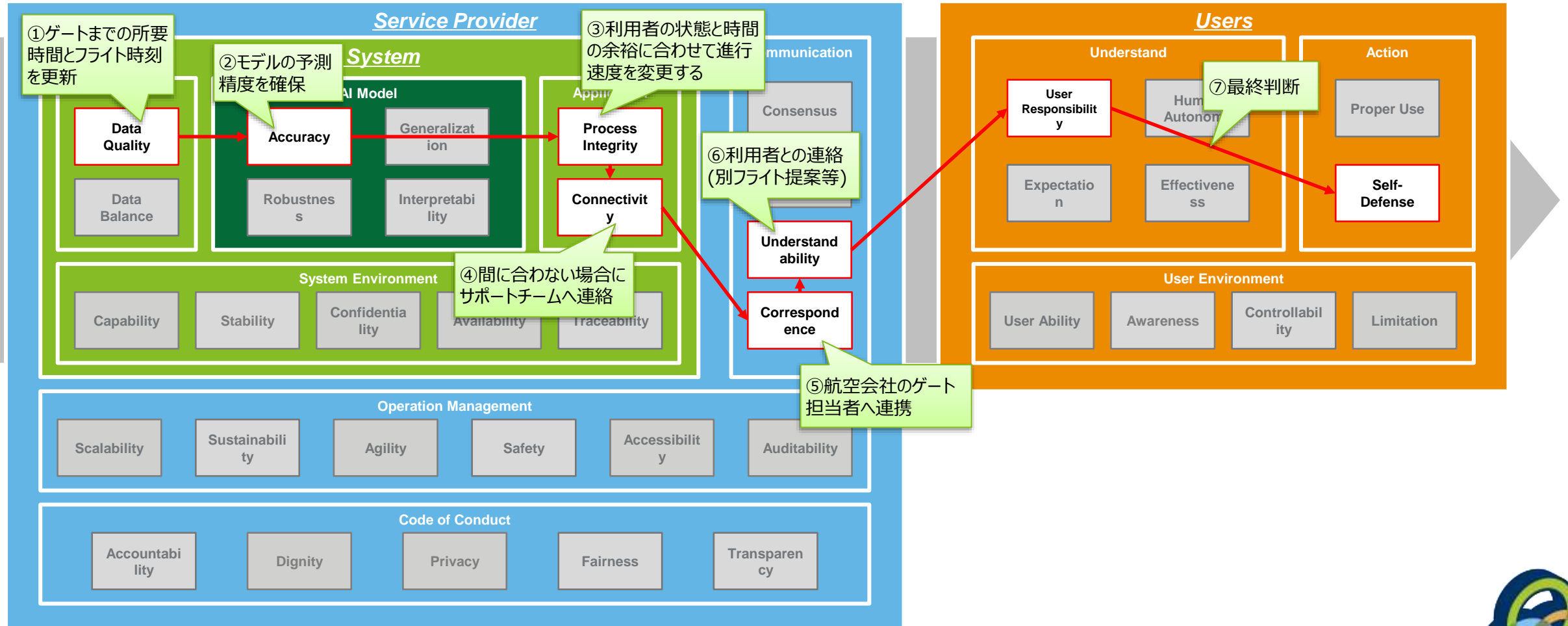
Step3

- 「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 -

R002

不十分なサービス

案内を開始しても出発に間に合わない



# リスクチェーンに従ってリスクコントロールを検討

Step4

- 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 -

R002

## 不十分なサービス

案内を開始しても出発に間に合わない

### コントロールの内容

AIシステム (A空港情報システム部 + X社)	サービスプロバイダ (A空港カスタマーサポート部)	ユーザー (空港利用客)
<p>①【Data Quality】ロボットのゲートまでの所要時間とフライト時刻情報を更新する (A空港情報システム部)</p> <p>②【Accuracy】モデルの予測精度を確保する (X社)</p> <p>③【Process Integrity】利用者の状態と時間の余裕に合わせて進行速度を変更する (A空港情報システム部 + X社)</p> <p>④【Connectivity】間に合わない場合にサポートチームへ自動伝達する (ロボット)</p>	<p>⑤【Correspondence】航空会社のゲート担当者へ連携し対応可能かを確認する (A空港カスタマーサポート部)</p> <p>⑥【Understandability】(搭乗に間に合わない場合)利用者と連絡を取り、変更フライト等を提案する (A空港カスタマーサポート部 + 航空会社)</p>	<p>⑦【User Responsibility】【Self-Defense】変更フライトを 手続するなどの最終判断を行う (空港利用客)</p>





# 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

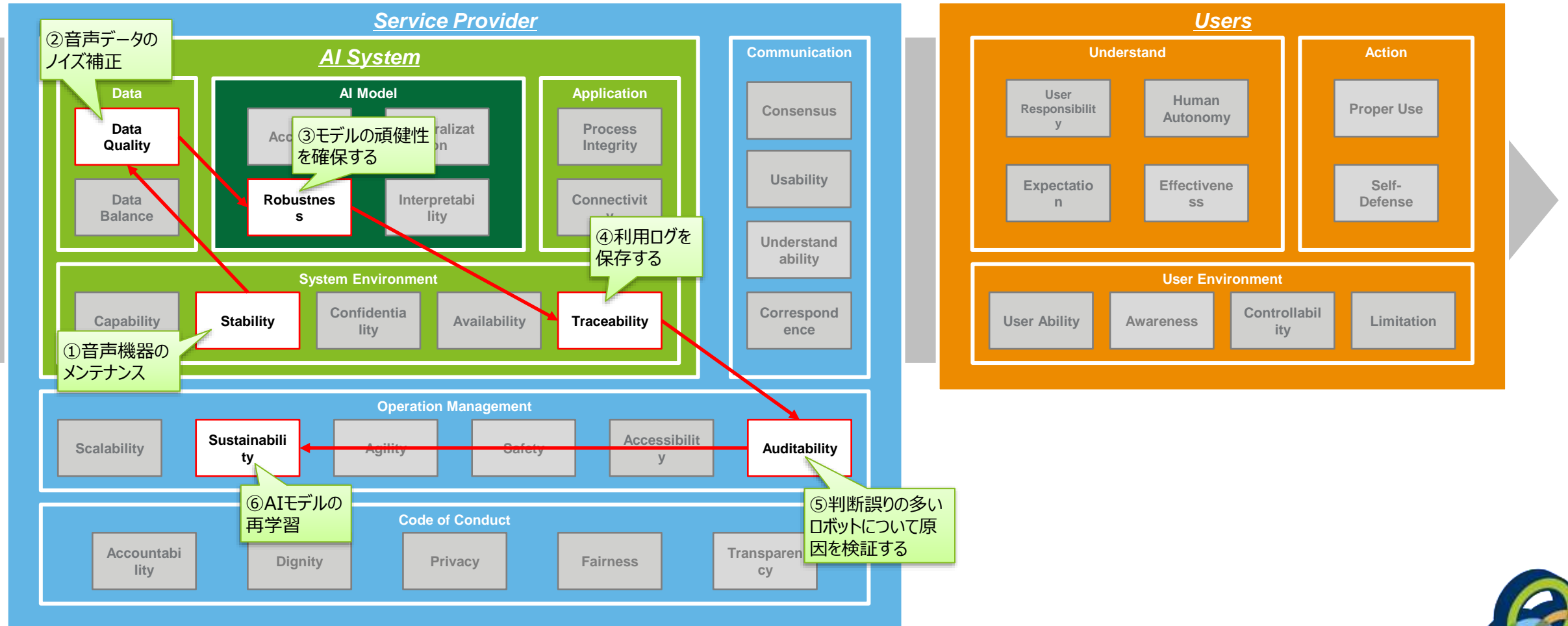
Step3

- 「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 -

R003

## ノイズによる影響

空港内の環境(混雑度合等)のノイズの影響で、ロボットが正しい判断を実施できない



# リスクチェーンに従ってリスクコントロールを検討

Step4

- 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 -

R003

## ノイズによる影響

空港内の環境(混雑度合等)のノイズの影響で、ロボットが正しい判断を実施できない

### コントロールの内容

AIシステム (A空港情報システム部+X社)	サービスプロバイダ (A空港カスタマーサポート部)	ユーザー (空港利用客)
<p>①【Stability】集音マイクを適切にメンテナンスする (A空港情報システム部+X社)</p> <p>②【Data Quality】音声データに対してノイズ補正等で劣化に対処する (A空港情報システム部+X社)</p> <p>③【Robustness】モデルの頑健性を高めるように学習を行う (A空港情報システム部+X社)</p> <p>④【Traceability】AIの判断および利用者との対話履歴を保存する (A空港情報システム部門)</p>	<p>⑤【Auditability】判断誤りの多いロボットについて判断誤りの原因を検証する (A空港カスタマーサポート部)</p> <p>⑥【Sustainability】判断誤りの多いロボットのAIモデルの再学習を行う (A空港カスタマーサポート部/A空港情報システム部+X社)</p>	



# 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

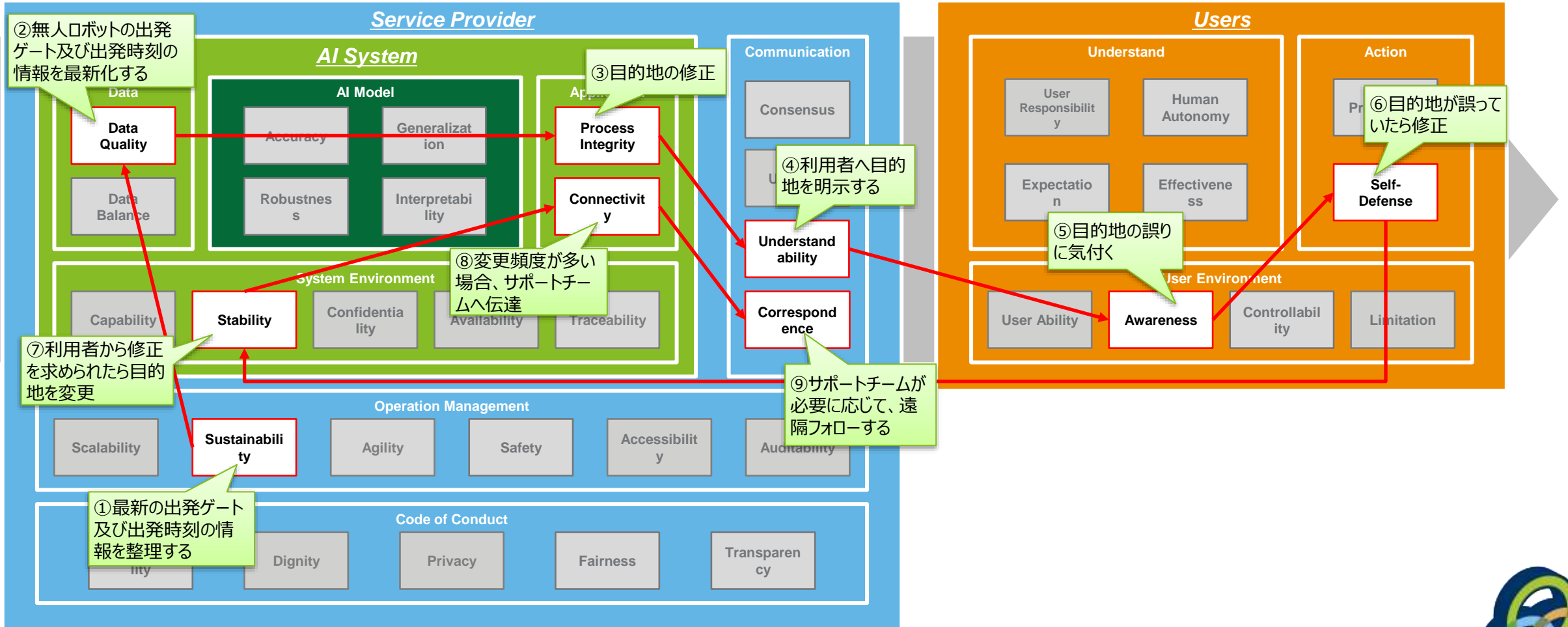
Step3

- 「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 -

R004

## フライト変更への対応

出発ゲートや出発時刻の変更に対応できない



# リスクチェーンに従ってリスクコントロールを検討

Step4

- 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 -

R004

## フライト変更への対応

出発ゲートや出発時刻の変更に対応できない

### コントロールの内容

AIシステム (A空港情報システム部+X社)	サービスプロバイダ (A空港カスタマーサポート部)	ユーザー (空港利用客)
<p>②【Data Quality】無人ロボットの出発ゲート及び出発時刻の情報を最新化する (A空港情報システム部)</p> <p>③【Process Integrity】目的地を自動修正する (ロボット)</p> <p>⑦【Stability】利用者から修正を求められたら目的地を変更する (ロボット)</p> <p>⑧【Connectivity】変更頻度が多い場合、正しく認識できていない可能性があるため、サポートチームへ自動伝達する (ロボット)</p>	<p>①【Sustainability】最新の出発ゲート及び出発時刻の情報を整理する (A空港カスタマーサポート部)</p> <p>④【Understandability】利用者へ目的地をロボットへ表示する (A空港カスタマーサポート部)</p> <p>⑨【Correspondence】サポートチームが必要に応じて、利用者を遠隔フォローする (A空港カスタマーサポート部)</p>	<p>⑤【Awareness】ロボットが認識している目的地の誤りに気づく (空港利用客)</p> <p>⑥【Self-Defense】目的地が誤っていたら修正する (空港利用客)</p>



# 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

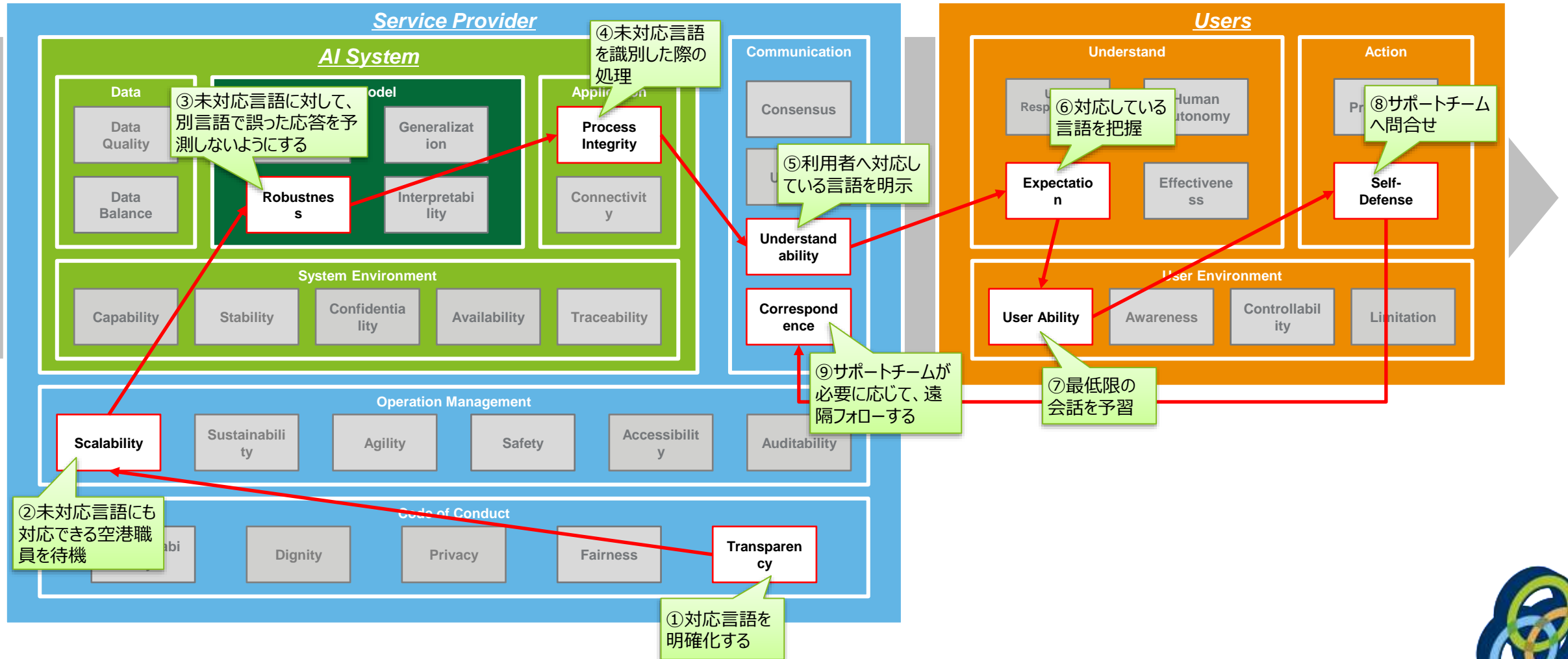
Step3

- 「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 -

R005

## 未対応言語への対応

対応していない言語・方言等でサービスを要求された際に、正しく対応できない



# リスクチェーンに従ってリスクコントロールを検討

Step4

- 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 -

R005

## 未対応言語への対応

対応していない言語・方言等でサービスを要求された際に、正しく対応できない

コントロールの内容		
AIシステム (A空港情報システム部+X社)	サービスプロバイダ (A空港カスタマーサポート部)	ユーザー (空港利用客)
<p>③【Robustness】無人ロボットが未対応言語に対して、別言語で誤った応答を予測しないようにする (A空港情報システム部+X社)</p> <p>④【Process Integrity】無人ロボットが未対応言語を識別した際に利用者へ未対応言語である旨表示する (ロボット)</p>	<p>①【Transparency】無人ロボットが対応している言語を明確に整理する (A空港カスタマーサポート部)</p> <p>②【Scalability】様々な言語に対応できるスタッフを用意する (A空港カスタマーサポート部)</p> <p>⑤【Understandability】利用者に向けて対応している言語を明示する (A空港カスタマーサポート部)</p> <p>⑨【Correspondence】サポートチームが必要に応じて、利用者を遠隔フォローする (A空港カスタマーサポート部)</p>	<p>⑥【Expectation】無人ロボットが対応している言語を認識する (空港利用客)</p> <p>⑦【User Ability】道を尋ねるのに必要な最低限の会話を調べられるようにする (空港利用客)</p> <p>⑧【Self-Defense】出発ゲートが分からない場合に空港スタッフに連絡して助けを求める (空港利用客)</p>



# 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

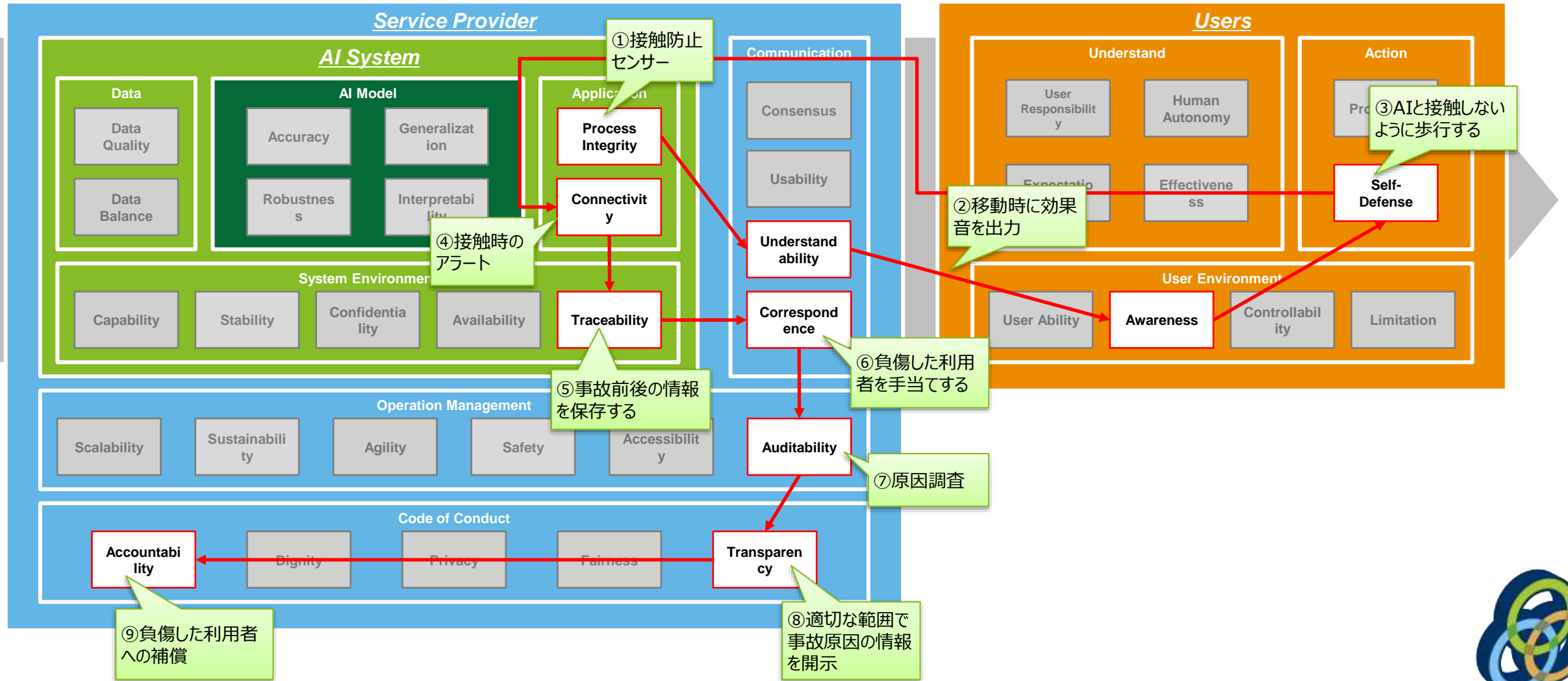
Step3

- 「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 -

R006

## 人物への危害

案内中に別の人物と接触して負傷させてしまう



# 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 -

R006

## 人物への危害

案内中に別の人物と接触して負傷させてしまう

### コントロールの内容

AIシステム (A空港情報システム部+X社)	サービスプロバイダ (A空港カスタマーサポート部)	ユーザー (空港利用客)
<p>①【Process Integrity】人やモノと接触しないようにセンサーが反応したら、急停止する (A空港情報システム部)</p> <p>④【Connectivity】ロボットが人やモノと接触した際に管理部へ事故発生場所を含めたアラートを送信する (A空港情報システム部)</p> <p>⑤【Traceability】事故前後の情報を保存する (A空港情報システム部)</p>	<p>②【Understandability】ロボットを認識できるように移動時に効果音を出力する (A空港カスタマーサポート部)</p> <p>⑥【Correspondence】負傷した利用者を手当てする (A空港カスタマーサポート部)</p> <p>⑦【Auditability】事故原因を検証する (A空港安全管理部)</p> <p>⑧【Transparency】適切な範囲で事故原因の情報を開示する (A空港経営者・空港管理部)</p> <p>⑨【Accountability】負傷した利用者への補償を行う (A空港経営者・空港管理部)</p>	<p>②【Awareness】ロボットを認識できるように移動時に効果音を出力する (空港利用客)</p> <p>③【Self-Defense】AIと接触しないように歩行する (空港利用客)</p>







# 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 -

R007

## 危険エリアへの誘導

侵入できない場所に対して、誘導してしまう

### コントロールの内容

AIシステム (A空港情報システム部+X社)	サービスプロバイダ (A空港カスタマーサポート部)	ユーザー (空港利用客)
<p>②【Data Quality】ロボットに立入禁止エリアを正しく設定する (A空港情報システム部)</p> <p>③【Accuracy】ロボットが立入禁止エリアに進まないように進路を設定する (ロボット)</p> <p>④【Connectivity】ロボットの進路情報をサポートチームに連携する (ロボット)</p>	<p>①【Sustainability】立入禁止エリアを明確に整理する (A空港カスタマーサポート部)</p> <p>⑤【Agility】ロボットの進行予定経路を遠隔で修正する (A空港カスタマーサポート部)</p> <p>⑥【Understandability】立入禁止エリアが分かりやすいようにサインボード等を置く (A空港カスタマーサポート部)</p> <p>⑧【Usability】【Safety】ロボットが立入禁止エリアに進んでいる場合に緊急停止や修正を行う (A空港カスタマーサポート部)</p>	<p>⑦【Awareness】【Self-Defense】ロボットが立入禁止エリアに向かっていると気づいたら修正する (空港利用客)</p>

