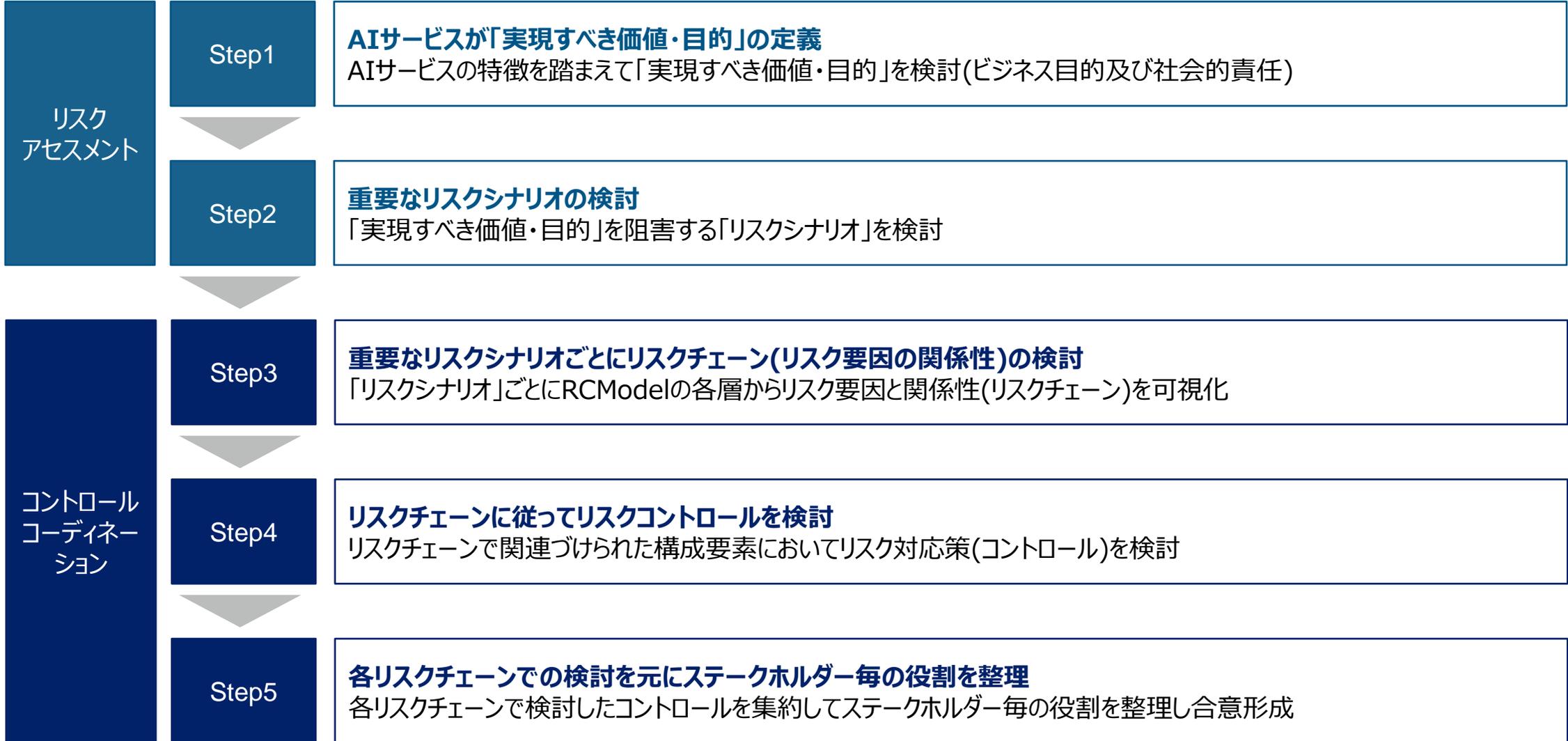


リスクチェーンモデル(RCModel)ケース検討事例： Case07 ローン審査AI



ケース検討のステップ



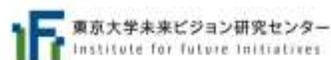


ケース事例（AIサービスとリスクコーディネーション研究会）

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 (Case07 : ローン審査AI)

- AIサービスが「実現すべき価値・目的」の定義 -

Step1

A銀行に対してローンの申し込みを行う予定の消費者（以下、申込者とする）に対して、ローンの申し込みの前に参考情報として見込融資金額を提供するAIサービスである。申込者はWeb画面よりアンケート情報と希望融資金額を入力し、AIサービスより出力されるスコア・借入利率・借入限度額を得る。

申し込み後、A銀行融資部の担当者が詳細な融資金額を見積もり、申込者と金額の合意を経て審査手続を進める。A銀行融資部の担当者は、AIサービスより出力されるスコア・借入利率・借入限度額を参考にして詳細な融資金額を見積もる。

評価対象であるAIサービスプロバイダのA銀行ITサービス部は、ビジネス利用者であるA銀行融資部や外部信用機関、SNS等から収集した学習データを基に機械学習（回帰モデル）で見込融資金額を予測する学習モデルを作成している。

【実現すべき価値・目的】

- 販売機会の拡大
- 審査の品質維持と効率化
- 社会的責任

【期待精度】

AIモデルにより算出された見込融資金額を基にローンの貸付を行った申込者の貸倒率が2%以内であることを、期待値として設定している。

※学習モデルについては、定期的の実績データを加えた追加学習を実施している。

※同じ申込者が追加情報を加えた際は、再度スコア・借入利率・借入限度額の予測が実施される。このとき、過去の予測結果は記録されており、申込者もWebから閲覧が可能である。

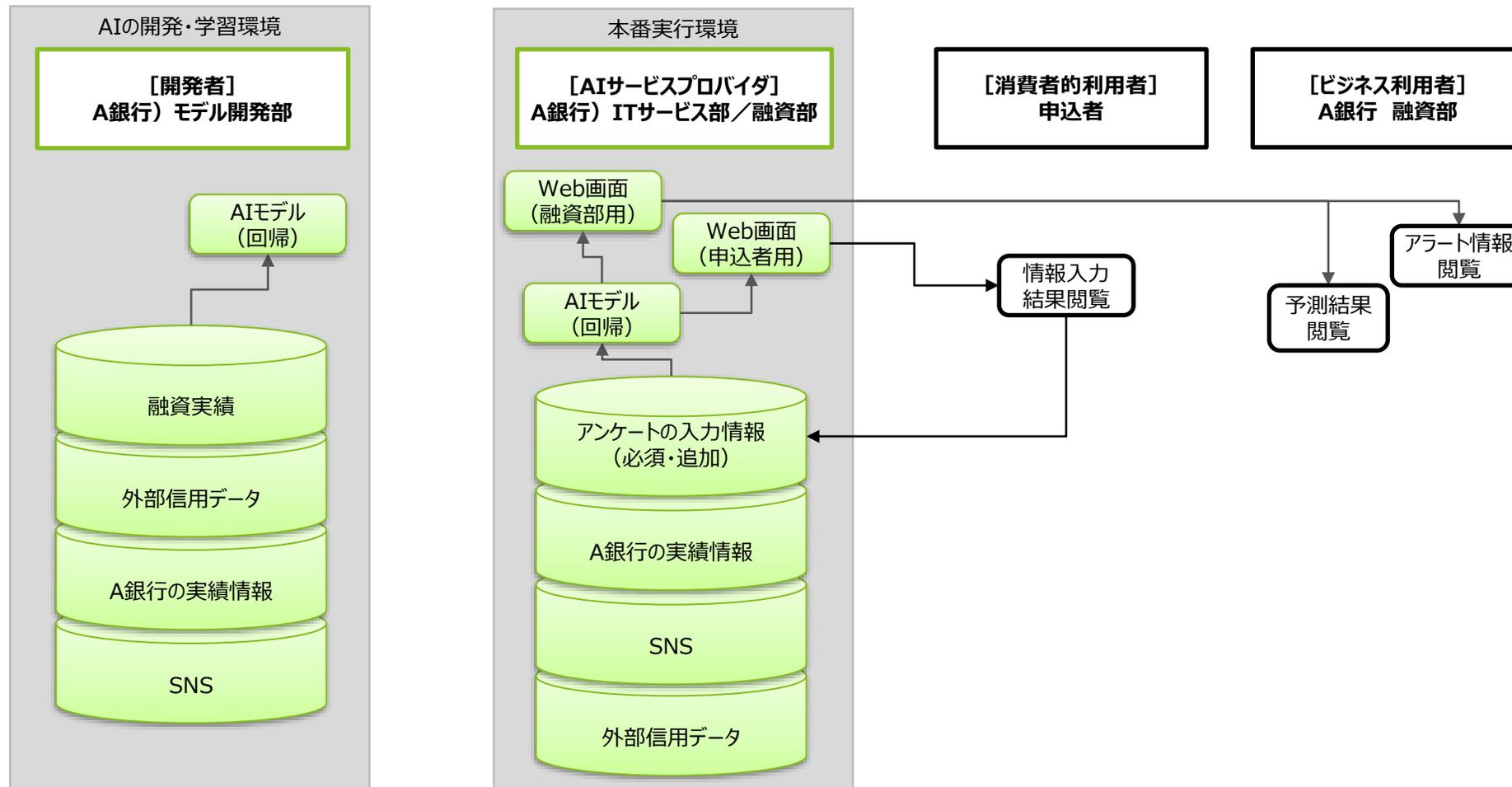
※AIによる予測結果が申込者の希望融資金額を下回った場合、アラート情報としてA銀行融資部に出力される。このとき、アラート情報には判断結果に大きく寄与した特徴量が含まれる。



ケーススタディの概要 (Case07 : ローン審査AI)

- システムの全体概要 -

AIシステム	A銀行) モデル開発部	ローンの申込に対して、スコアリングと限度額を予測する
AIサービスプロバイダ	A銀行) ITサービス部/融資部	ローン審査の最終判断を行う (AIの一次判断を覆す可能性もある)
ユーザー	申込者/A銀行) 融資部担当者	ローン審査に申込を行う



ケーススタディの概要 (Case07 : ローン審査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
融資実績	学習	A銀行融資部の実績情報を匿名化して収集	A銀行ITサービス部の開発環境	有 (匿名化)
外部信用情報機関の信用データ	学習	外部信用機関から収集	A銀行ITサービス部の開発環境	有 (匿名化)
SNS	学習	SNS情報を匿名化して収集	A銀行ITサービス部の開発環境	有 (匿名化)
アンケートの入力情報 (必須項目・追加項目)	本番	申込者がWebサービスから入力 (※)	A銀行が契約するパブリッククラウド環境	有
A銀行の申込者に係る情報	本番	申込者が許可した場合) A銀行の実績情報と連携	A銀行が契約するパブリッククラウド環境	有 (本人が許可した場合)
SNS	本番	申込者が許可した場合) SNS情報を収集	A銀行が契約するパブリッククラウド環境	有 (本人が許可した場合)
外部信用情報機関の信用データ	本番	外部信用機関から収集	A銀行が契約するパブリッククラウド環境	有

(※) 申込者が入力する情報

必須項目：氏名・年齢・性別・住所・年収・家族構成・職業・他のローン有無

追加項目：資産・仕事の詳細・学歴・ライフスタイル等

*学習データフレームと本番データフレームは同一のものになるように設計している



ケーススタディの概要 (Case07 : ローン審査AI)

- AIサービスの入出力 -

【AIサービスからの出力内容】

AIサービス利用者	申込者	A銀行融資部
出力結果の内容	スコア・借入利率・借入限度額の予測	アラート情報：申込者の希望融資金額を借入限度額が下回ったときの申込者情報
出力方法（画面／アクチュエータ等）	Webサービス上でスコア・借入利率・借入限度額の予測結果が出力される。履歴情報も閲覧できる。	アラート情報として、申込者の希望融資金額を借入限度額が下回ったときの申込者情報をリスト化し、日次で融資部担当者にメール配信する。
期待精度（正解率／誤差等）	貸倒率：2%以内	貸倒率：2%以内
利用者判断の有無	有	有
根拠情報の出力	無	有（予測結果に寄与した特徴量）
安全性のリスク有無	無	無
外部AIへの連携	無	無
外部AIへの連携方法・プロトコル	無	無



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	販売機会の拡大	1-1	予測性能の確保 ■ 目標設定 ■ AIの予測精度	R001	過剰な審査による顧客離れ	過剰な審査が行われ、融資可能な顧客自体が著しく減少してしまう
		1-2	リアルタイムな診断 ■ パフォーマンス	R002	不十分な処理速度	AIによる判断処理速度が間に合わず、顧客が他行に流れてしまう
		1-3	外部攻撃からの保護 ■ AIの頑健性 ■ セキュリティ	R003	セキュリティ保護	外部からの攻撃により、融資実績等の学習データや学習済みモデルに意図しない変更が与えられ、AIが誤った学習をしてしまう
2	審査の品質維持と効率化	2-1	審査の品質維持 ■ AIの再学習 ■ 教師ラベル ■ 説明可能性 ■ 不正検知	R004	融資戦略の変化	AIによる審査の判断が銀行における融資戦略の変化に対応できない
				R005	誤った教師ラベルの設定	返済不可となった顧客が教師ラベルとして設定されず、同じような貸倒が頻発してしまう
				R006	過度なAI依存	AIの判断を全く疑わず、審査が形骸化することで誤った判断が見逃してしまう
				R007	虚偽の申込	申込者が事実と異なる自身の情報を入力、または別人のSNSを紐づけることによって、AIの判断結果を誤らせる
		2-2	コンティンジェンシープラン	R008	異常時の代替運用	AIの性能劣化等で異常が発生した際に人手での審査に切り替えられず、業務に支障をきたす
3	社会的責任	3-1	アカウントビリティ ■ 説明可能性 ■ 検証可能性	R009	申込者に対する説明	申込者に対して、審査の結果や借入金額について合理的な説明ができない
		3-2	倫理・コンプライアンス遵守 ■ AIの汎化性	R010	公平性	特定の人種／性別に対して、明らかに不公平な予測結果を生じさせる
		3-3	情報管理 ■ データ管理	R011	スコアの流出・目的外利用	スコアリング結果を目的外に利用される／提供されることで、特定の顧客に対して不利益が発生する
				R012	プライバシー保護	個人情報の取扱いを誤ることによって個人情報保護法の違反が行われる



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 販売機会の拡大	R001	過剰な審査による顧客離れ	○			●	戦略対象データの分類 十分な正解率の確保 判断根拠の出力	予測性能の検証 再学習	再学習の検討
	R002	不十分な処理速度	○			●	パフォーマンスの確保 アラートの出力	パフォーマンス監視 システム環境の改善	
	R003	セキュリティ保護					セキュリティ管理	原因調査・改善	
2 審査の品質維持と効率化	R004	融資戦略の変化	○	○	○	●	※R001と同じ	※R001と同じ	※R001と同じ
	R005	誤った教師ラベルの設定	○		○	●	外部データの自動連携 学習データの検証	使いやすいUI 判断精度・異常値の検証 再学習	正確なフィードバック
	R006	過度なAI依存	○		○	●	判断根拠の出力	審査レベルの定義 判断根拠の表示	判断責任の理解 戦略顧客の理解
	R007	虚偽の申込	○		○	●	データ汚染の保護 判断根拠の出力 不正入力の検知	申込内容の検証 審査プロセスの制限	不正入力への自動制御
	R008	異常時の代替運用		○	○	●	アラートの出力 代替サービスへの切替 異常の記録	異常時の取り決め 原因の検証 システム改善	異常時の取り決め マニュアル手続の準備
3 社会的責任	R009	申込者に対する説明	○		○	●	判断根拠の出力 ネガティブな対象の処理	使いやすいUI 分かりやすい表示 検証・顧客対応	適切な入力 判断結果への申立て
	R010	公平性	○			●	データの偏り モデルの汎化性 ネガティブな対象の処理	公平性ポリシーの検討 ネガティブな判断の開示	AIの判断傾向の理解 最終判断・マニュアル手続
	R011	スコアの流出・目的外利用					データ保護	職業倫理の教育	職業倫理の教育
	R012	プライバシー保護					データ保護	法令順守の教育	法令順守の教育 データの取扱

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銀行) 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銀行) 法務・コンプラ

- 職業倫理の教育
- 法令順守の教育

A銀行) 内部監査

- 内部監査 (業務監査)

- AIサービスプロバイダ - A銀行) ITサービス部/融資部

- 公平性ポリシーの検討
- ネガティブな判断の開示
- 審査プロセスの制限
- パフォーマンス監視
- 予測性能の検証
- 申込内容の検証
- 判断精度・異常値の検証
- 再学習
- システム環境の改善
- 原因調査・改善
- 審査レベルの定義
- 異常時の取り決め
- 使いやすいUI
- 判断根拠の表示
- 分かりやすい表示
- 顧客対応

A銀行) モデル開発部

- 十分な正解率の確保
- モデルの汎化性
- 判断根拠の出力
- 戦略対象データの分類
- 学習データの検証
- データの偏りの調整

A銀行) ITサービス部

- パフォーマンスの確保
- 異常の記録
- セキュリティ管理
- 代替サービスへの切替
- 外部データの自動連携
- アラートの出力
- 不正入力の検知

- ユーザー - A銀行) 融資部担当者

- 再学習の検討
- 正確なフィードバック
- 判断責任の理解
- 戦略顧客の理解
- AIの判断傾向の理解
- 異常時の取り決め
- マニュアル手続の準備

金融庁

- 金融機関の監督

- ユーザー - 申込者

- 適切な入力
- 判断結果への申立て
- 不正入力への自動制御



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

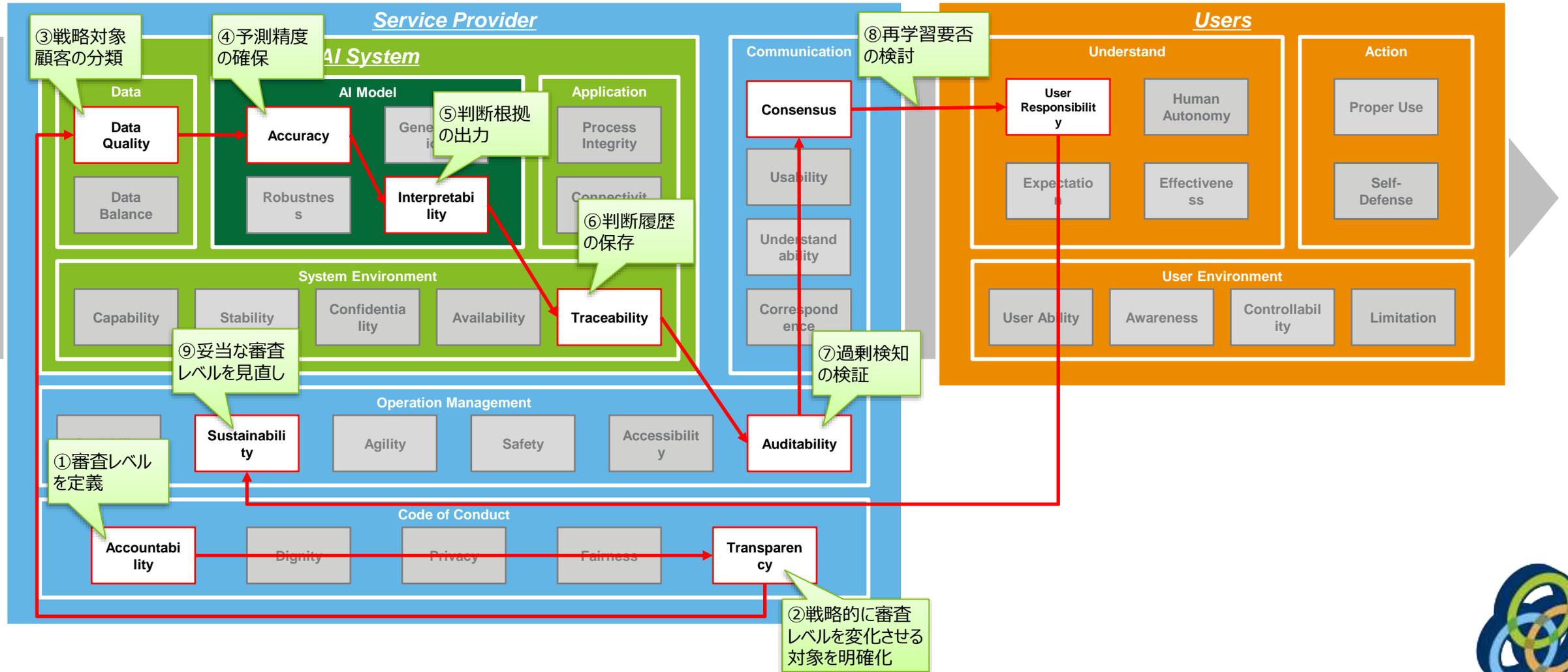
Step3

- 「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

過剰な審査による顧客離れ

過剰な審査が行われ、融資可能な顧客自体が著しく減少してしまう



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

過剰な審査による顧客離れ

過剰な審査が行われ、融資可能な顧客自体が著しく減少してしまう

コントロールの内容		
AIシステム (A銀行モデル開発部)	サービスプロバイダ (A銀行ITサービス部/融資部)	ユーザー (申込者/融資部担当者)
③【Data Quality】戦略対象顧客のデータを明確に分ける (A銀行モデル開発部)	①【Accountability】AIで実現すべき審査のレベルを定義する (A銀行融資部)	⑧【User Responsibility】判断精度の状況を説明し再学習を行うか検討する (A銀行ITサービス部/融資部)
④【Accuracy】モデルの予測精度を確保するように学習を行う (A銀行モデル開発部)	②【Transparency】戦略的に審査レベルを変化させる対象を明確にする (A銀行融資部)	
⑤【Interpretability】モデルの判断根拠を出力する (A銀行モデル開発部)	⑦【Auditability】モデルの性能を定期的に検証し、過剰に否認が行われていないかを確認する (A銀行ITサービス部/融資部)	
⑥【Traceability】AIの判断履歴を保存する (A銀行ITサービス部)	⑧【Consensus】判断精度の状況を説明し再学習を行うか検討する (A銀行ITサービス部/融資部)	
	⑨【Sustainability】維持コストを踏まえた上で妥当な審査レベル(モデル性能)を見直す (A銀行ITサービス部/融資部)	



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

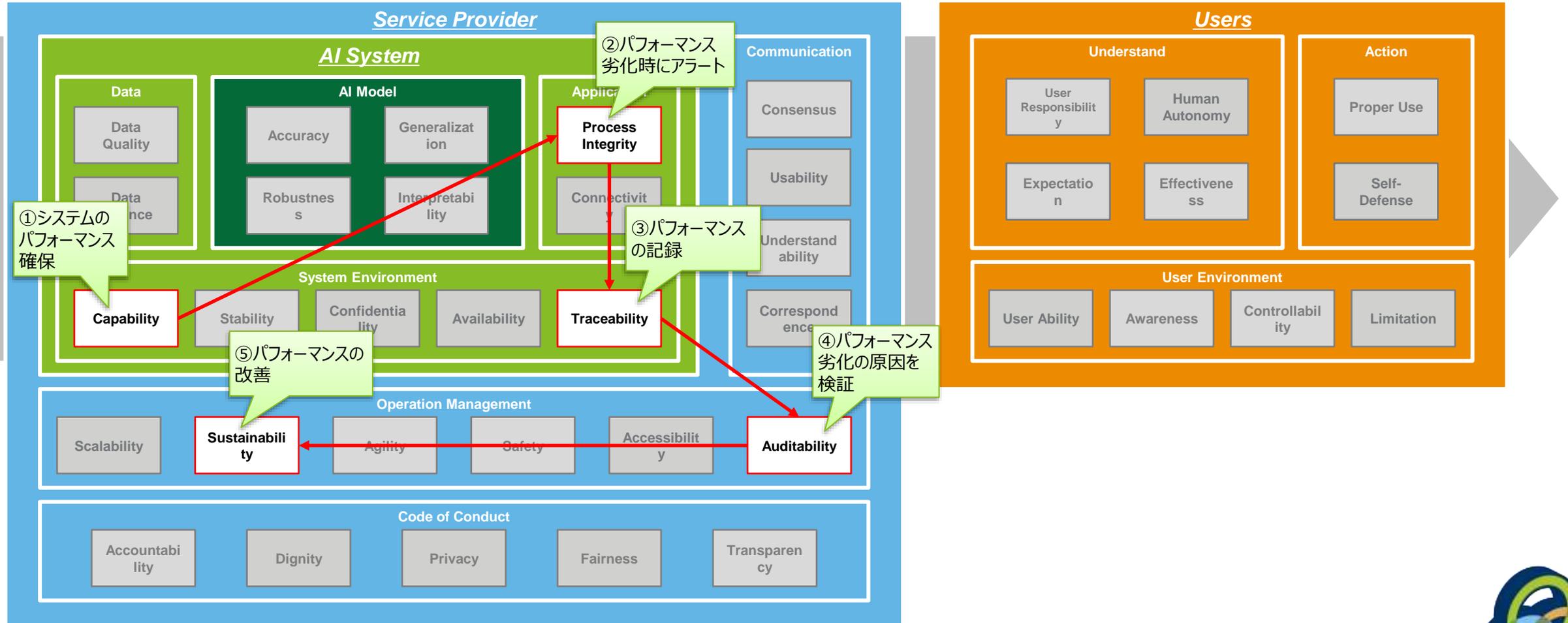
Step3

- 「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

不十分な処理速度

AIによる判断処理速度が間に合わず、顧客が他行に流れてしまう



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 -

R002

不十分な処理速度

AIによる判断処理速度が間に合わず、顧客が他行に流れてしまう

コントロールの内容

AIシステム (A銀行モデル開発部)	サービスプロバイダ (A銀行ITサービス部/融資部)	ユーザー (申込者/融資部担当者)
<p>①【Capability】AIシステム全般で十分な処理速度を確保する (A銀行モデル開発部)</p> <p>②【Process Integrity】予測パフォーマンス劣化が見受けられる場合にアラートを出力する (A銀行モデル開発部)</p> <p>③【Traceability】パフォーマンスの負荷状況を記録する (A銀行ITサービス部)</p>	<p>④【Auditability】パフォーマンス劣化の原因を検討する (A銀行ITサービス部)</p> <p>⑤【Sustainability】十分なパフォーマンスを確保するためにAIシステムの処理性能を改善する (A銀行ITサービス部)</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

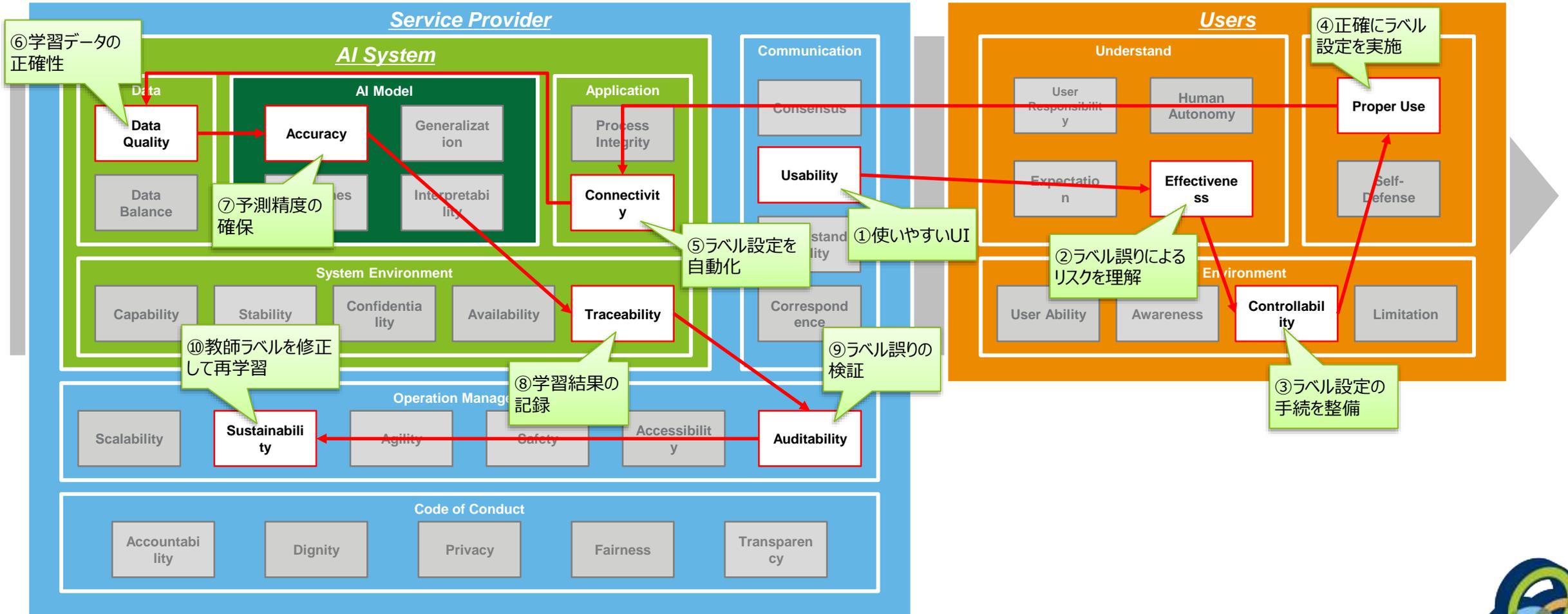
Step3

- 「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

誤った教師ラベルの設定

返済不可となった顧客が教師ラベルとして設定されず、同じような貸倒が頻発してしまう



リスクチェーンに従ってリスクコントロールを検討

- 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

誤った教師ラベルの設定

返済不可となった顧客が教師ラベルとして設定されず、同じような貸倒が頻発してしまう

コントロールの内容		
AIシステム (A銀行モデル開発部)	サービスプロバイダ (A銀行ITサービス部/融資部)	ユーザー (申込者/融資部担当者)
⑤【Connectivity】顧客への教師ラベル設定を社内システムから自動連携する (A銀行モデル開発部)	①【Usability】誤操作が起こりにくい使いやすいUIを準備する (A銀行ITサービス部)	②【Effectiveness】フィードバック誤りがAIモデルの性能劣化に影響することを認識する (融資部担当者)
⑥【Data Quality】学習データの正確性を検証する (A銀行モデル開発部)	⑨【Auditability】予測性能の変化や学習データにおける異常値(ラベル誤りと思われるデータ)を検証する (A銀行ITサービス部)	③【Controllability】正確なフィードバックを行うように手続きを整備する (融資部担当者)
⑦【Accuracy】モデルの判断精度を確保する (A銀行モデル開発部)	⑩【Sustainability】データの教師ラベルを修正し、必要に応じてAIモデルを再学習する (A銀行ITサービス部)	④【Proper Use】正確にフィードバックを実施する (融資部担当者)
⑧【Traceability】各学習段階での予測結果や異常値の内容を保存する (A銀行ITサービス部)		



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

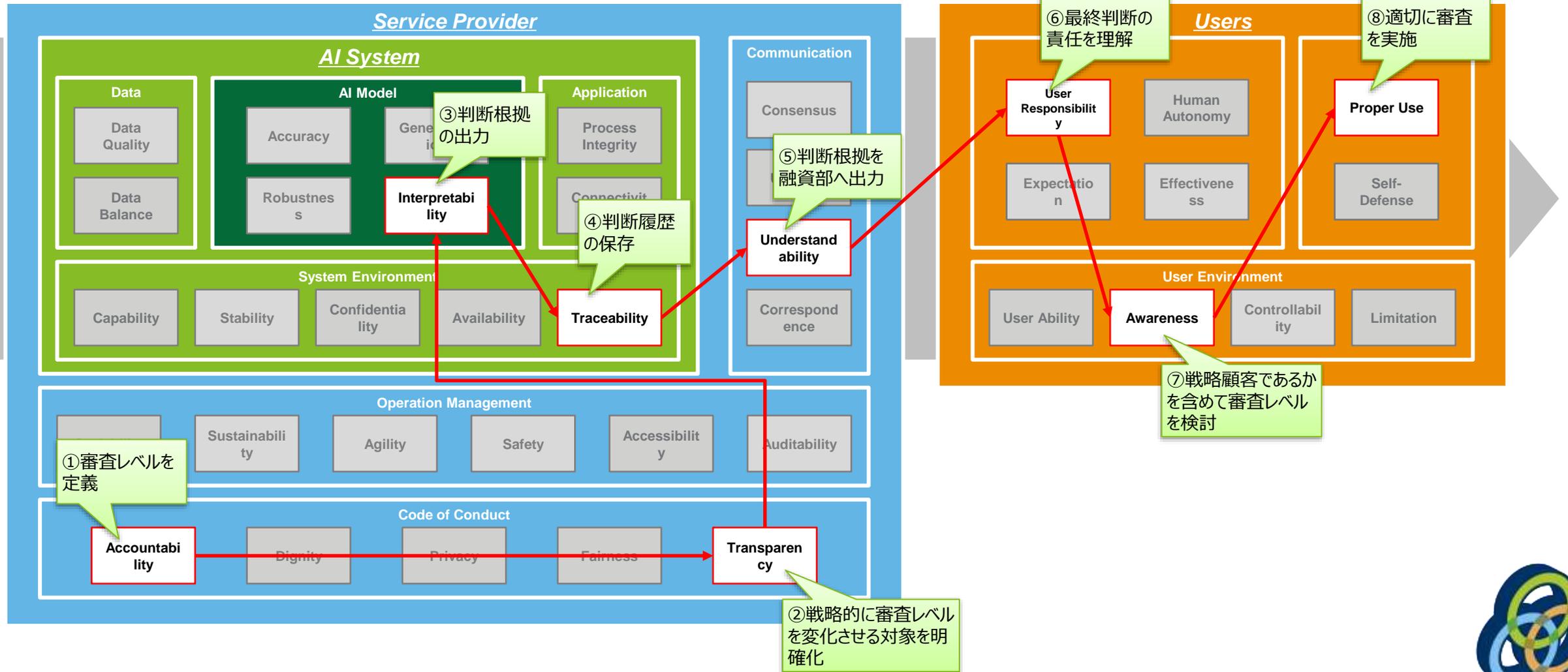
Step3

- 「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

過度なAI依存

AIの判断を全く疑わず、審査が形骸化することで誤った判断が見逃してしまう



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 -

R006

過度なAI依存

AIの判断を全く疑わず、審査が形骸化することで誤った判断が見逃してしまう

コントロールの内容

AIシステム (A銀行モデル開発部)	サービスプロバイダ (A銀行ITサービス部/融資部)	ユーザー (申込者/融資部担当者)
<p>③【Interpretability】モデルの判断根拠を出力する (A銀行モデル開発部)</p> <p>④【Traceability】利用時のAIの判断結果情報を保存する (A銀行ITサービス部)</p>	<p>①【Accountability】AIで実現すべき審査のレベルを定義する (A銀行融資部)</p> <p>②【Transparency】戦略的に審査レベルを下げる対象を明確にする (A銀行融資部)</p> <p>⑤【Understandability】判断根拠情報を融資部に向けて出力する (A銀行ITサービス部/融資部)</p>	<p>⑥【User Responsibility】最終判断は融資部担当者にあることを理解する (融資部担当者)</p> <p>⑦【Awareness】戦略顧客であるかどうかを含めて審査のレベルを検討する (融資部担当者)</p> <p>⑧【Proper Use】適切に審査を実施する (融資部担当者)</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

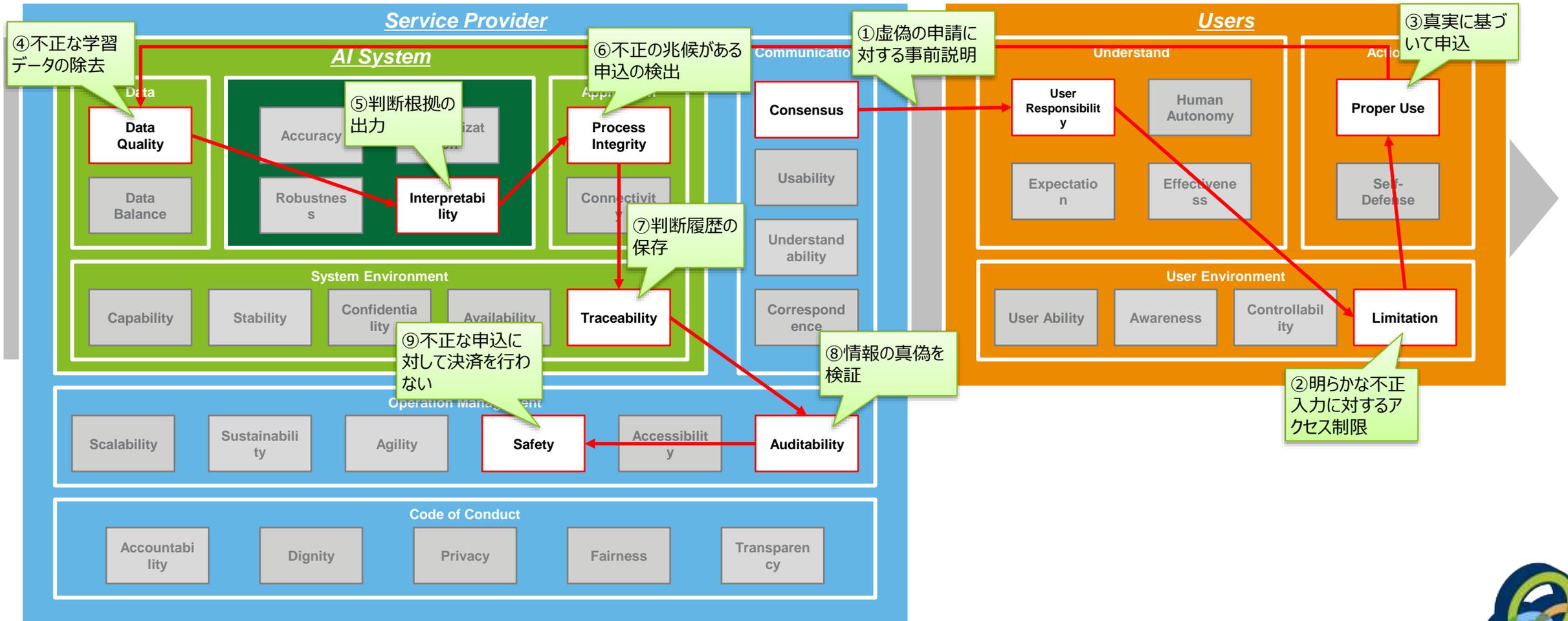
Step3

- 「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 -

R007

虚偽の申込

申込者が事実と異なる自身の情報を入力、または別人のSNSを紐づけることによって、AIの判断結果を誤らせる



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 -

R007

虚偽の申込

申込者が事実と異なる自身の情報を入力、または別人のSNSを紐づけることによって、AIの判断結果を誤らせる

コントロールの内容		
AIシステム (A銀行モデル開発部)	サービスプロバイダ (A銀行ITサービス部/融資部)	ユーザー (申込者/融資部担当者)
<p>④【Data Quality】入力されたデータに明らかな虚偽がある場合に補正・除去し、不当な予測結果を誘発するデータの蓄積を回避する (A銀行モデル開発部)</p> <p>⑤【Interpretability】モデルの判断根拠を出力する (A銀行モデル開発部)</p> <p>⑥【Process Integrity】矛盾した入力情報等の不正入力が見受けられる場合にアラートを出力する (A銀行モデル開発部)</p> <p>⑦【Traceability】利用時のAIの判断結果情報を保存する (A銀行ITサービス部)</p>	<p>①【Consensus】申込者の入力情報に虚偽が含まれる場合の責任・ペナルティを説明する (A銀行融資部/申込者)</p> <p>⑧【Auditability】入力情報の真偽が疑わしい内容について申込者に情報の真偽を確かめる (A銀行融資部)</p> <p>⑨【Safety】不正入力がある場合には審査プロセスに通さない (A銀行融資部)</p>	<p>①【User Responsibility】申込者の入力情報に虚偽が含まれる場合の責任・ペナルティを説明する (A銀行融資部/申込者)</p> <p>②【Limitation】明らかな不正入力(不正なSNSアカウントの紐づけ等)を識別した際にアクセス制限を行う (申込者)</p> <p>③【Proper Use】真実に基づいて申込みを行う (申込者)</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

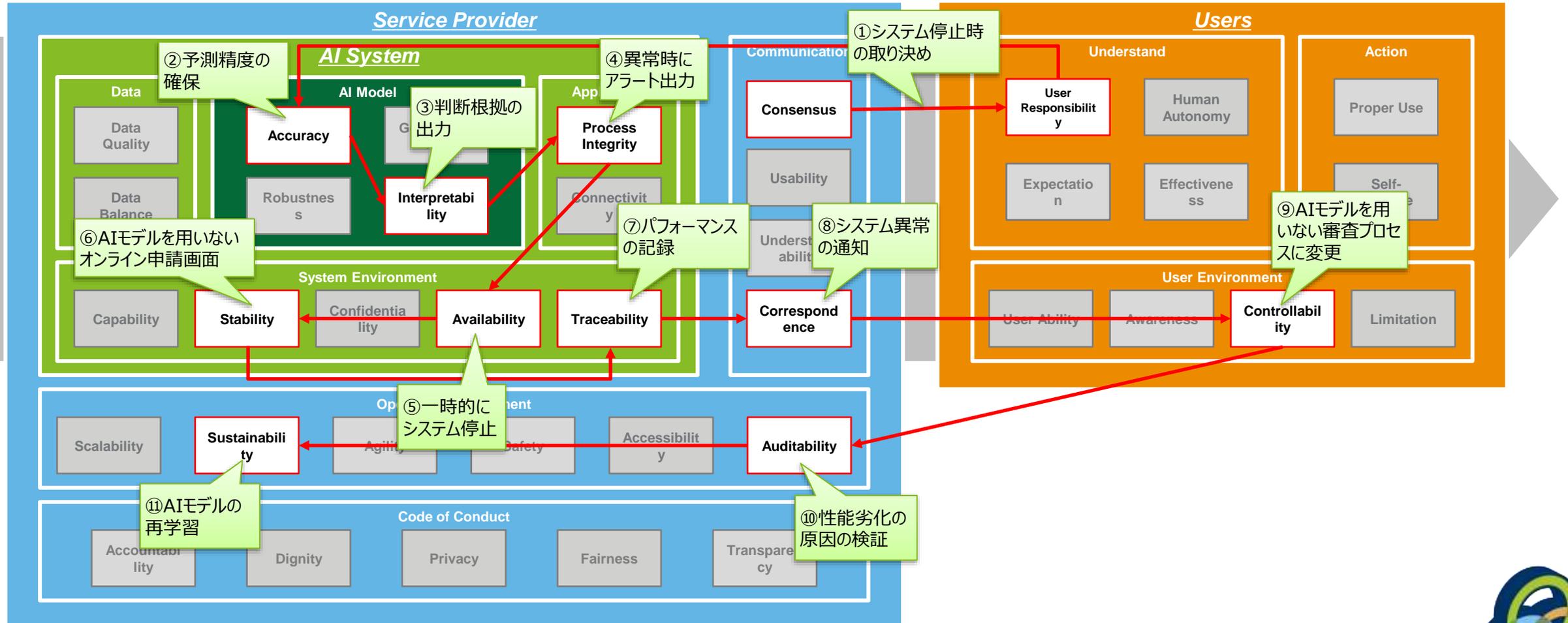
Step3

- 「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 -

R008

異常時の代替運用

AIの性能劣化等で異常が発生した際に人手での審査に切り替えられず、業務に支障をきたす



リスクチェーンに従ってリスクコントロールを検討

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 -

R008

異常時の代替運用

AIの性能劣化等で異常が発生した際に人手での審査に切り替えられず、業務に支障をきたす

コントロールの内容		
AIシステム (A銀行モデル開発部)	サービスプロバイダ (A銀行ITサービス部/融資部)	ユーザー (申込者/融資部担当者)
②【Accuracy】モデルの予測精度を確保するように学習を行う (A銀行モデル開発部)	①【Consensus】システム停止時の代替運用について取り決めておく (A銀行ITサービス部/融資部担当者)	①【User Responsibility】再学習が間に合わないときの人手での代替運用について取り決めておく (A銀行ITサービス部/融資部担当者)
③【Interpretability】モデルの判断根拠を出力する (A銀行モデル開発部)	⑧【Correspondence】システムに異常があることを融資部担当者に連携する (A銀行ITサービス部/融資部担当者)	⑨【Controllability】AIモデルを用いない場合の業務に変更する (A銀行融資部担当者)
④【Process Integrity】システム異常(予測精度の著しい劣化含む)が見受けられる場合にアラートを出力する (A銀行モデル開発部)	⑩【Auditability】モデルの性能劣化の原因と対策を検討する (A銀行ITサービス部/融資部)	
⑤【Availability】一時的にシステムを利用できない状態とする (A銀行ITサービス部)	⑪【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銀行ITサービス部/融資部)	
⑥【Stability】AIモデルを用いないオンライン申請画面を立ち上げる (A銀行ITサービス部)		
⑦【Traceability】パフォーマンス結果等を保存する (A銀行ITサービス部)		



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

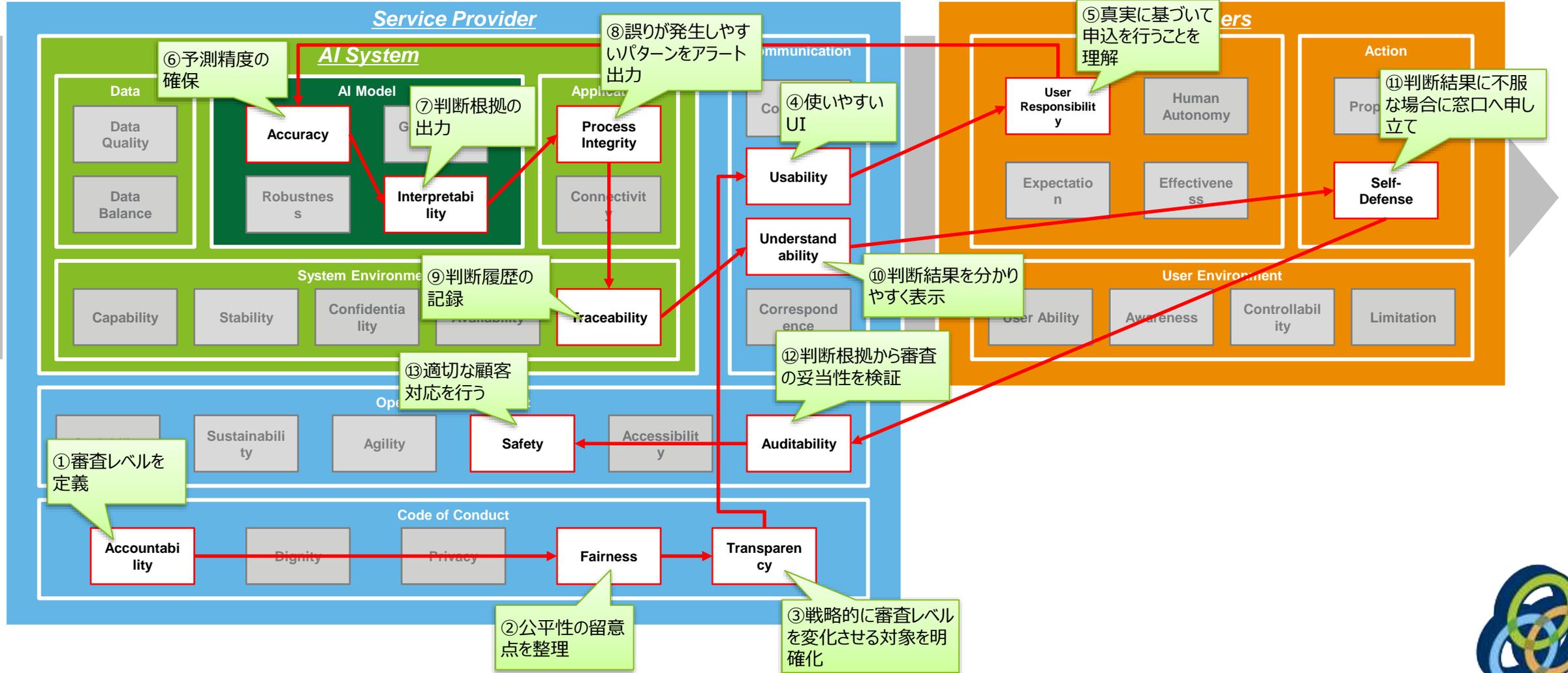
Step3

- 「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 -

R009

申込者に対する説明

申込者に対して、審査の結果や借入金額について合理的な説明ができない



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 -

R009

申込者に対する説明

申込者に対して、審査の結果や借入金額について合理的な説明ができない

コントロールの内容		
AIシステム (A銀行モデル開発部)	サービスプロバイダ (A銀行ITサービス部/融資部)	ユーザー (申込者/融資部担当者)
<p>⑥【Accuracy】モデルの予測精度を確保するように学習を行う (A銀行モデル開発部)</p> <p>⑦【Interpretability】モデルの判断根拠を出力する (A銀行モデル開発部)</p> <p>⑧【Process Integrity】予測結果が安定しない対象について、審査部で最終判断する旨の自動処理を出力する (A銀行モデル開発部)</p> <p>⑨【Traceability】AIの判断履歴を保存する (A銀行ITサービス部)</p>	<p>①【Accountability】AIで実現すべき審査のレベルを定義する (A銀行融資部)</p> <p>②【Fairness】公平性の留意点を検討する (A銀行融資部)</p> <p>③【Transparency】戦略的に審査レベルを下げる対象/ネガティブな判断を除けない対象等を明確にする (A銀行ITサービス部/融資部)</p> <p>④【Usability】入力間違いが起こりにくいUIを用意する (A銀行ITサービス部/融資部)</p> <p>⑩【Understandability】判断結果を分かりやすく表示する (A銀行ITサービス部/融資部)</p> <p>⑫【Auditability】判断根拠情報等から審査の判断過程・妥当性を検討する (A銀行ITサービス部/融資部)</p> <p>⑬【Safety】適切な顧客対応を行う (A銀行融資部)</p>	<p>⑤【User Responsibility】真実に基づいて申し込みを行うことを理解する (申込者)</p> <p>⑪【Self-Defense】判断結果に納得がいけない場合には銀行側の窓口へ申し立てる (申込者)</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

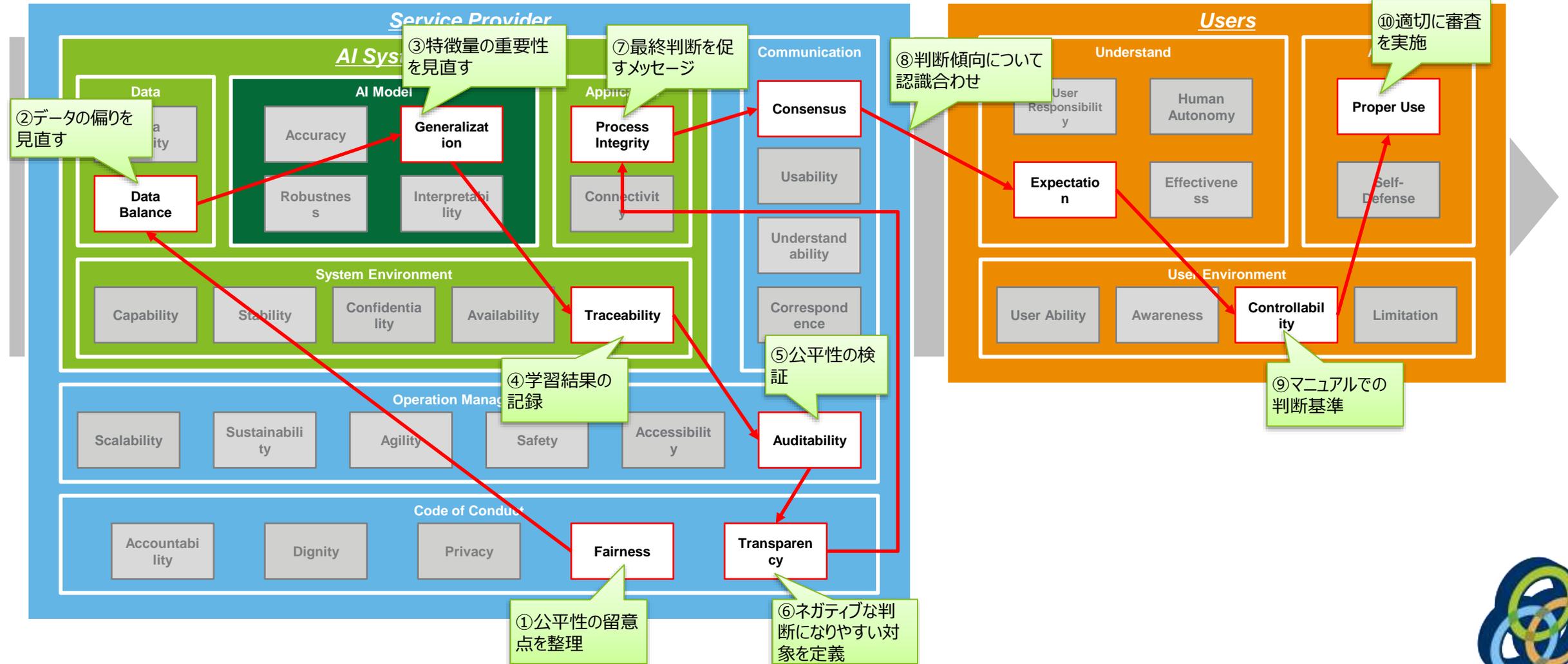
Step3

- 「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 -

R010

公平性

特定の人種/性別に対して、明らかに不公平な予測結果を生じさせる



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 -

R010

公平性

特定の人種／性別に対して、明らかに不公平な予測結果を生じさせる

コントロールの内容

AIシステム (A銀行モデル開発部)	サービスプロバイダ (A銀行ITサービス部／融資部)	ユーザー (申込者／融資部担当者)
<p>②【Data Balance】データの偏りを見直す (A銀行モデル開発部)</p> <p>③【Generalization】特徴量の重要性を見直す (A銀行モデル開発部)</p> <p>④【Traceability】学習による検証結果を記録する (A銀行モデル開発部)</p> <p>⑦【Process Integrity】予測結果が安定しない対象について、審査部で最終判断する旨の自動処理を出力する (A銀行モデル開発部)</p>	<p>①【Fairness】公平性の留意点を検討する (A銀行融資部)</p> <p>⑤【Auditability】モデルの公平性を検証する (A銀行ITサービス部／融資部)</p> <p>⑥【Transparency】データ不足などの理由でネガティブな判断を除けない対象を明確にする (A銀行ITサービス部／融資部)</p> <p>⑧【Consensus】予測性能及びネガティブな判断傾向について認識合わせする (A銀行ITサービス部／融資部)</p>	<p>⑧【Expectation】予測性能及びネガティブな判断傾向について認識合わせする (A銀行ITサービス部／融資部)</p> <p>⑨【Controllability】マニュアルでの運用方法を整備する (A銀行融資部担当者)</p> <p>⑩【Proper Use】適切な審査を実施する (A銀行融資部担当者)</p>

