SEAMS

Safety Engineering for Automotive ML-controlled System

Open QA4AI Conference

「自律的自動運転の実現を支える 人工知能搭載システムの安全性立証技術の研究開発」

2019年5月17日

SEAMS Project
株式会社ヴィッツ　機能安全開発部
森川 聡久

株式会社ヴィッツ　名古屋大学　ArcSystemSolutions　株式会社アトリエ

※本資料における「SEAMS」は、「Safety Engineering for Automotive ML-controlled System」の略称です。

本発表の概要

- SEAMS Projectでは、自動運転などの“セーフティクリティカルなAI搭載システム”に対する安全性立証技術の研究開発を実施しています。

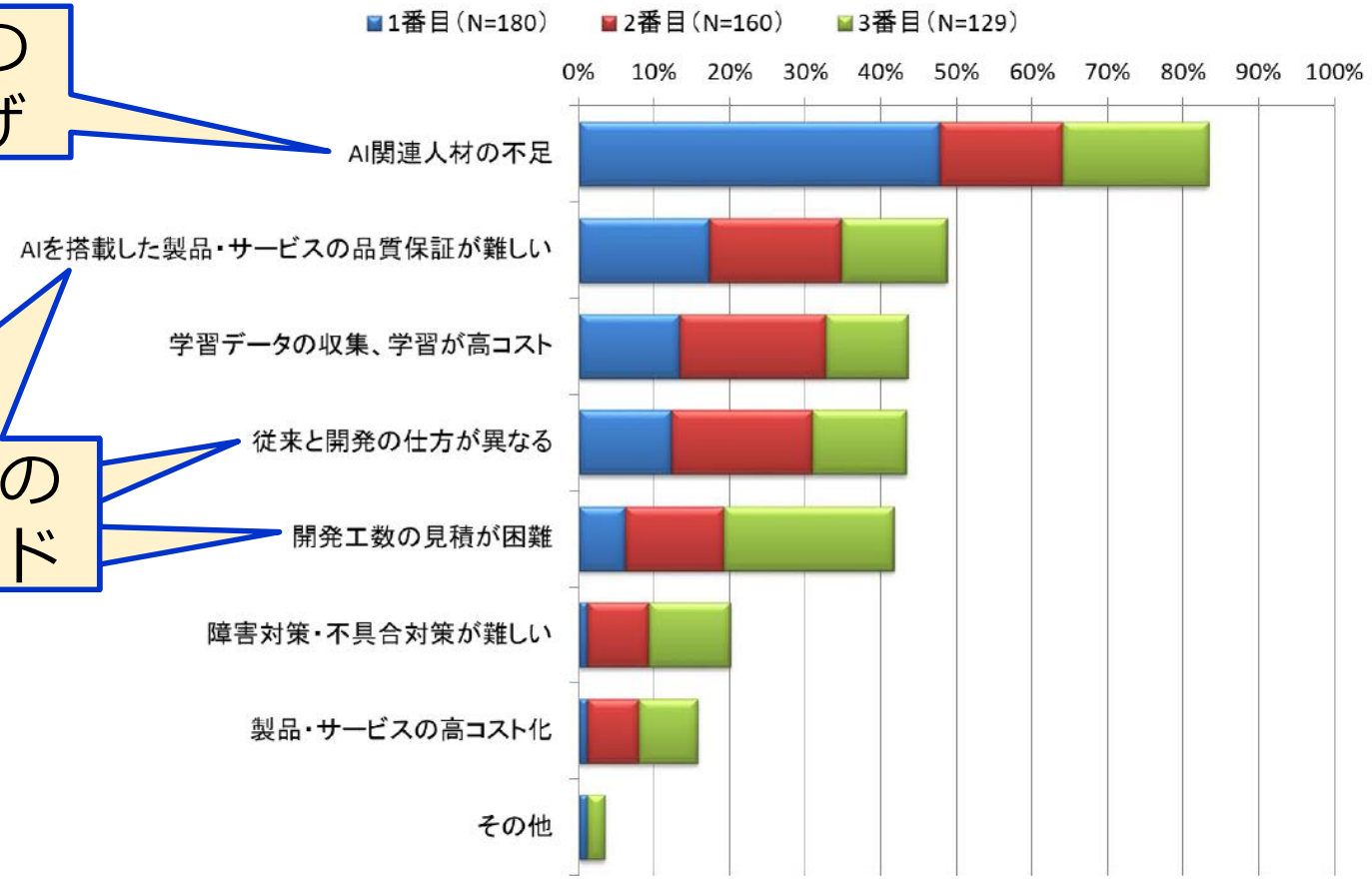
- 本発表では、現時点の研究成果の1つとして、「AI搭載システムの安全設計ガイドライン」（略称：SEAMSガイドライン）を紹介したい。

本ガイドによって改善できるAIの課題

出典：IPA「2018年度組込み/IoTに関する動向調査」

今年度
新規設問

Q23 AI技術を活用する／している際の課題



AI＋機能安全の
スキルの底上げ

機能安全規格への
準拠方法をガイド

IPA Information-technology Promotion Agency, Japan 187

アジェンダ

1. 本研究とSEAMSガイドラインの概要
2. SEAMS Projectが考える安全関連システムにおけるAIの影響
3. AI搭載システムの安全設計パターン
4. AI搭載システムの安全プロセス
5. まとめ

１．本研究とSEAMSガイドラインの概要

中部経済産業局 平成29年度 戦略的基盤技術高度化支援事業「自律的自動運転の実現を支える人工知能搭載システムの安全性立証技術の研究開発」

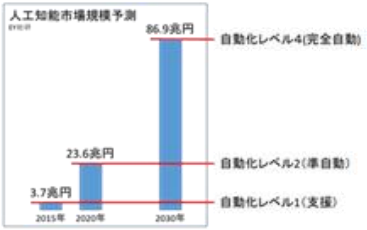
※3カ年プロジェクトの最終年度として実施中

従来技術

人工知能は万能ではない
2016年3月Google社の自動運転車両が人工知能の認識不備による事故(techradarより引用)

人工知能への理解不足
人工知能の活用は利益と損益を正確に理解する必要がある(Carbon Blackより引用)

完全自動化が前提の投資
完全自動化へ投資しているが、実現できなければ市場規模は維持か縮小(EY総研調査より引用)

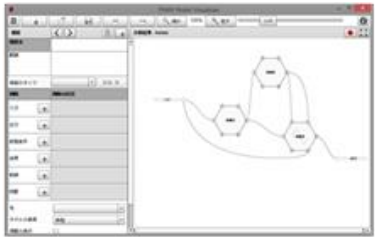


人工知能を安全に利用する方法の実現が急務

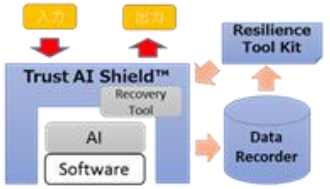
新技術

人工知能搭載システム 安全ガイドライン 「Safety-AI」
人工知能搭載システムを安全に利用するために、人工知能の利用分類と、利用目的に沿った安全開発工程を明確にする

人工知能搭載システム 安全分析手法 「AI-FRAM」
人工知能搭載システム全体を俯瞰して分析を行うことで、過誤や見過ごし、潜在リスクを見出すための安全分析手法



人工知能搭載システム 安全対策 「Trust AI Shield」「Fuzzing for AI」
人工知能搭載システムの不具合や異常動作を未然に検知することで、対象システムを安全な状態に移行させるためのソフトウェア部品を開発し、開発投資を抑える



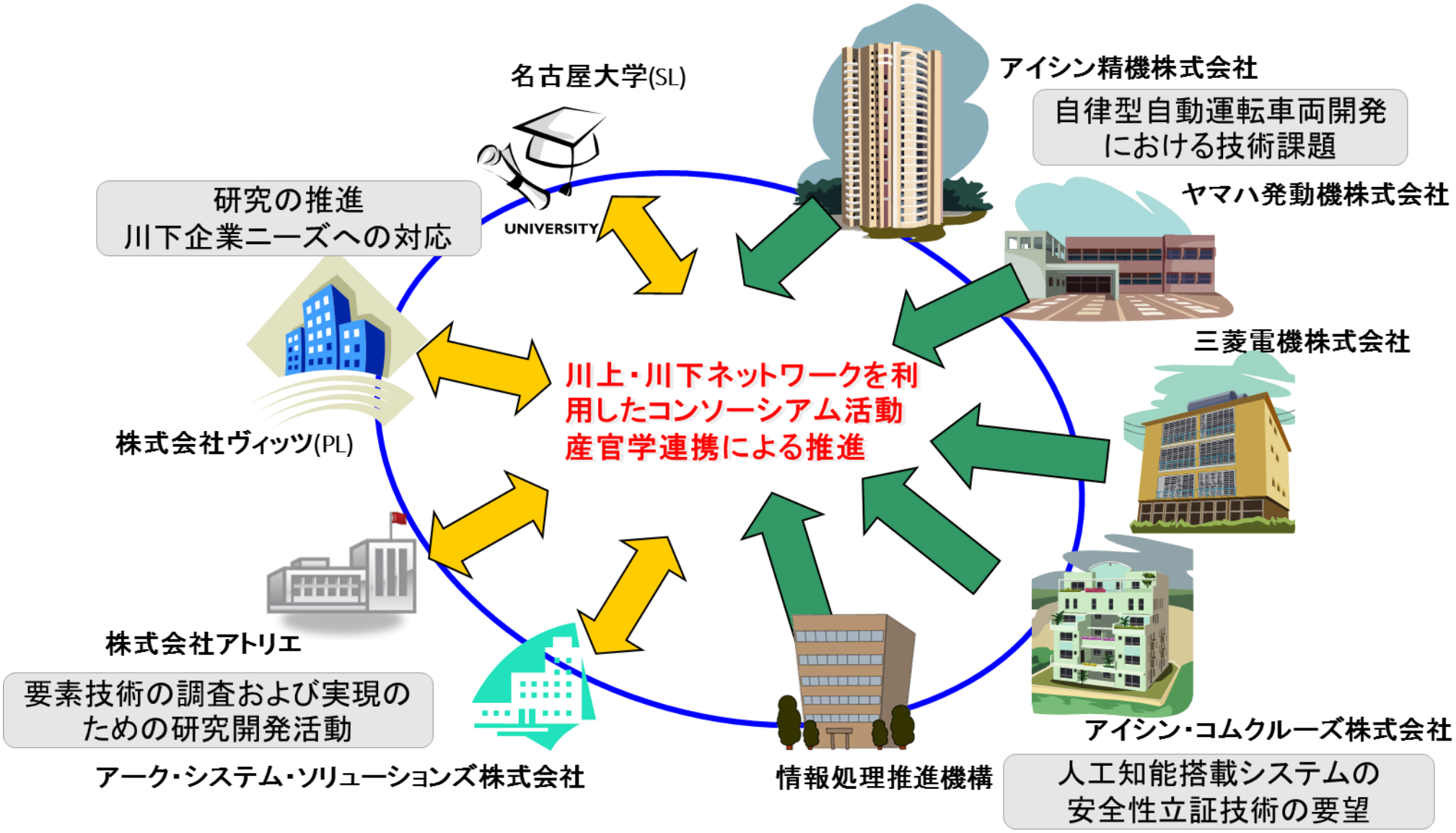
課題
- 人工知能搭載システムへの極度な過信
- 潜在リスクの見過ごし
- 投資失敗時のサンクコスト

川下企業の課題
- 人工知能搭載システムの安全評価ができない
- 安全評価の実施コストが計算できない

特徴
- 人工知能搭載システムを安全に利用可能
- リスクを予見し事前対策を実施
- 投資額を最小限に低減させる

川下企業のメリット
- 人工知能搭載システムの安全性を客観的に立証できる
- 安全評価のコストを事前に見積もれる

実施体制

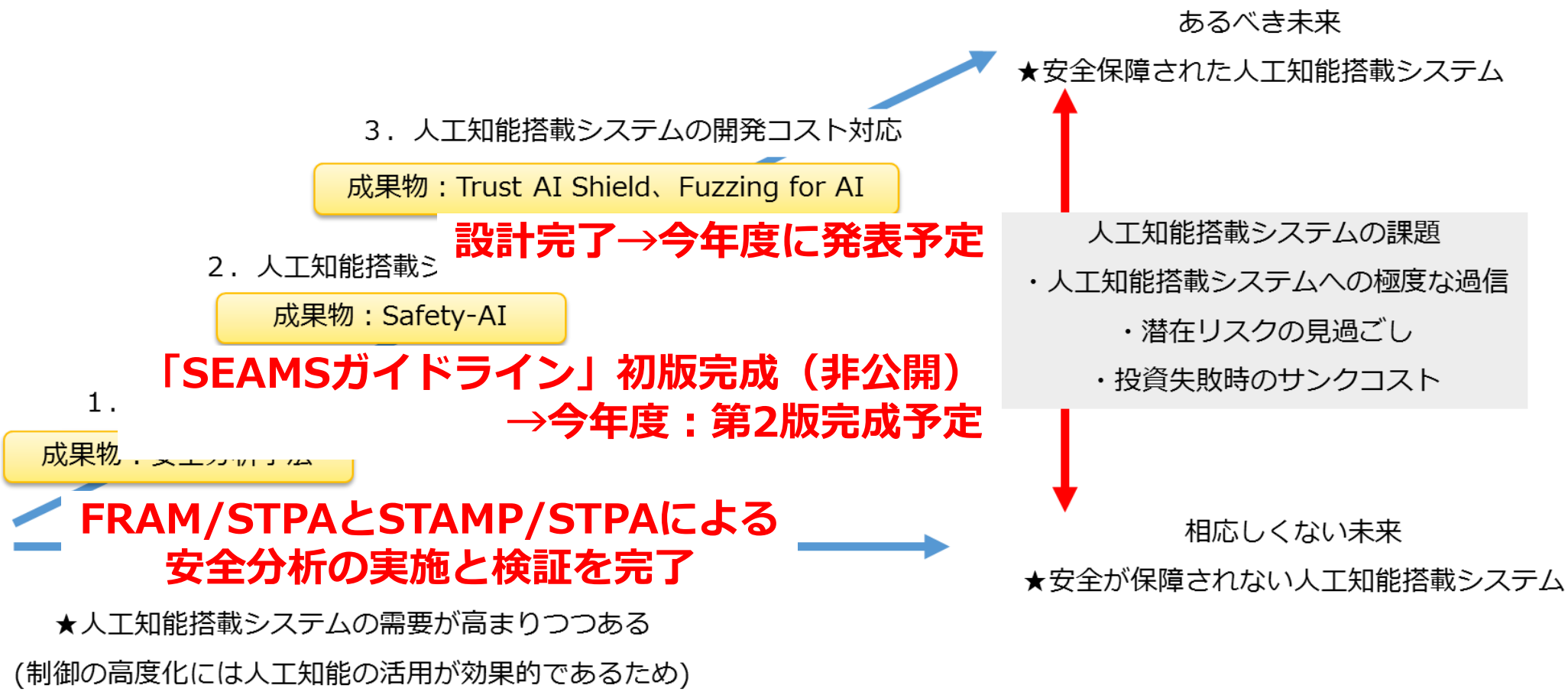


SEAMS Projectとは

SEAMS

- SEAMS Project＝本研究プロジェクトの活動名称
- SEAM=縫い目、つなぎ目の意味
- SEAMSは、人工知能を活用するための「つなぐ」を実現する技術群

現在までの進捗状況



SEAMSガイドライン 構成・目次

■構成物

・本編 Ver1.00（34ページ）（下記目次）

・実施例： Adaptive Crouse Control（ACC）自動運転レベル4システム
安全コンセプト（22ページ）、システム安全分析結果

■本編 目次

目次

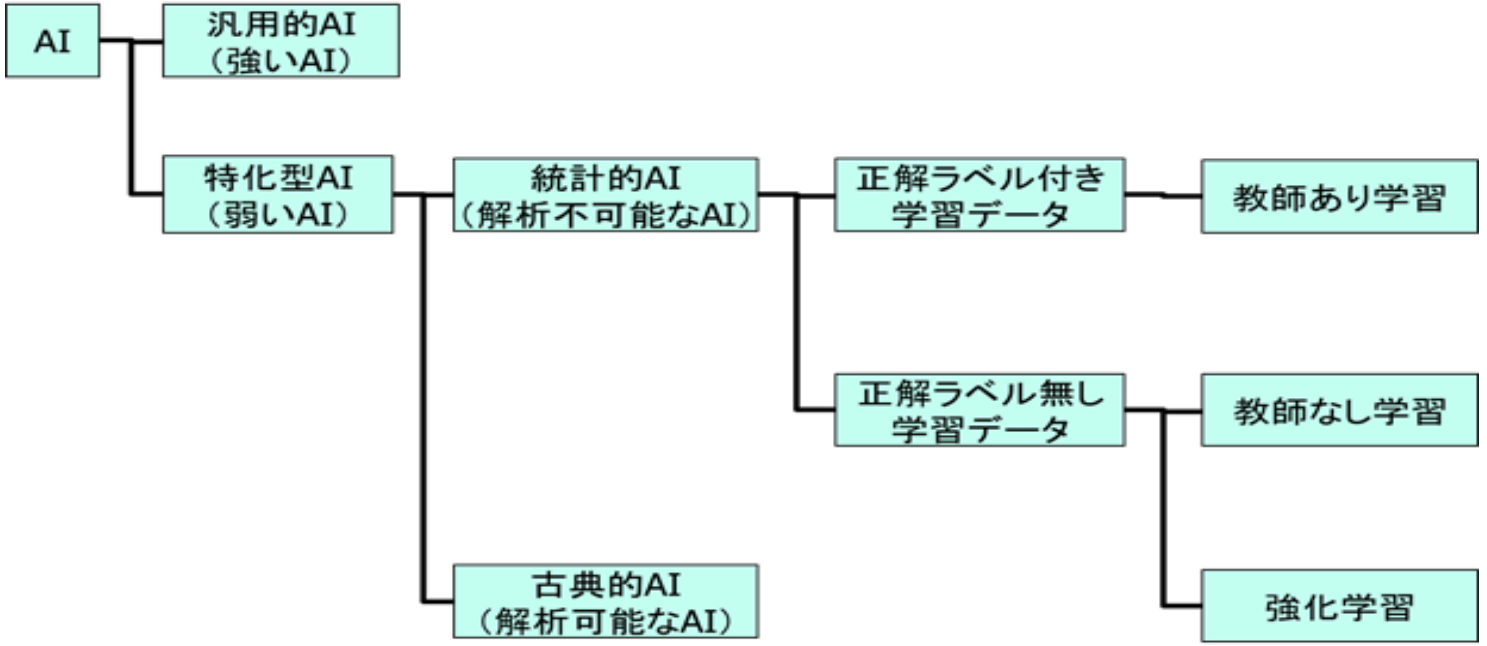
SEAMSガイドラインの位置づけ

- 目的：AI技術を搭載したシステムの安全性の示し方をガイド
- 想定読者：当該システムの開発者向け
  - 前提スキル：機能安全開発の知見を保有していること
- 手段：機能安全規格への適合方法を提示
  - 対象製品分野：全て
    - 分野（例：自動車）や規格（例：IEC61508,ISO26262）を問わず、基本的な考え方・設計方法・開発プロセスを整理
    - 具体的事例は、 ACC（自動車, ISO26262）を適用
  - 記載範囲：AIに特化した対応事項に限定
    - 従来の機能安全開発に従うことは記載を省略

2．SEAMS Projectが考える
安全関連システムにおけるAIの影響

AIの分類整理から見たAIの特性

- さまざまな文献、AI知見者の意見を踏まえ、AIを分類整理（下図）
  - AI関連は、多種多様な技術、様々な捉え方があり、整理の仕方に正解は無い
- ソフトウェアプログラムの特性：記述通りに動作する（故障除く）
- AIの特性：誤判定、非決定的な振る舞いが起こり得る
  - 統計的AIに限らず、古典的AI（エキスパートシステムなど）にも当てはまる
  - 複雑すぎるシステムでは、よくある話かも知れない（課題の本質はAIではない）



AIは機能安全では推奨されない（1/2）

- 機能安全規格では、AIの使用は非推奨である。
- 一方、IoTや自動運転システムでは、AIの搭載は必要。

IEC61508-3:2010 Table A.2

技術/手法	IEC61508-7参照	SIL1	SIL2	SIL3	SIL4
人工知能 –故障訂正	C.3.9	–	NR	NR	NR
動的再構成	C.3.10	–	NR	NR	NR

＜手法の目的＞

オンラインで安全性や信頼性を分析することで、可能性のある危険事象に非常に柔軟に対応できるようにすること。

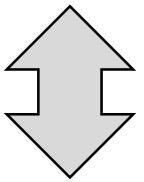
＜NRの理由（懸念リスク）＞

静的に動作が決まっておらず、運用時に正しく動作することを保証できない。信頼性が低い機能を安全面で使用するのはリスクが高い。

AIは機能安全では推奨されない（2/2）

- ソフトウェア安全要求仕様の望ましい特性
  （IEC61508-3:2010 Annex C 引用）
  - ソフトウェアで対応する安全ニーズの完全性
  - ソフトウェアで対応する安全ニーズの正確性
  - 曖昧さの回避を含む、固有仕様フォールトの回避性
  - 安全要求事項の理解容易性
  - ソフトウェアの安全以外の機能が安全機能へ危険な干渉を及ぼさない性質
  - 適合確認及び妥当性確認の基礎となる対応能力

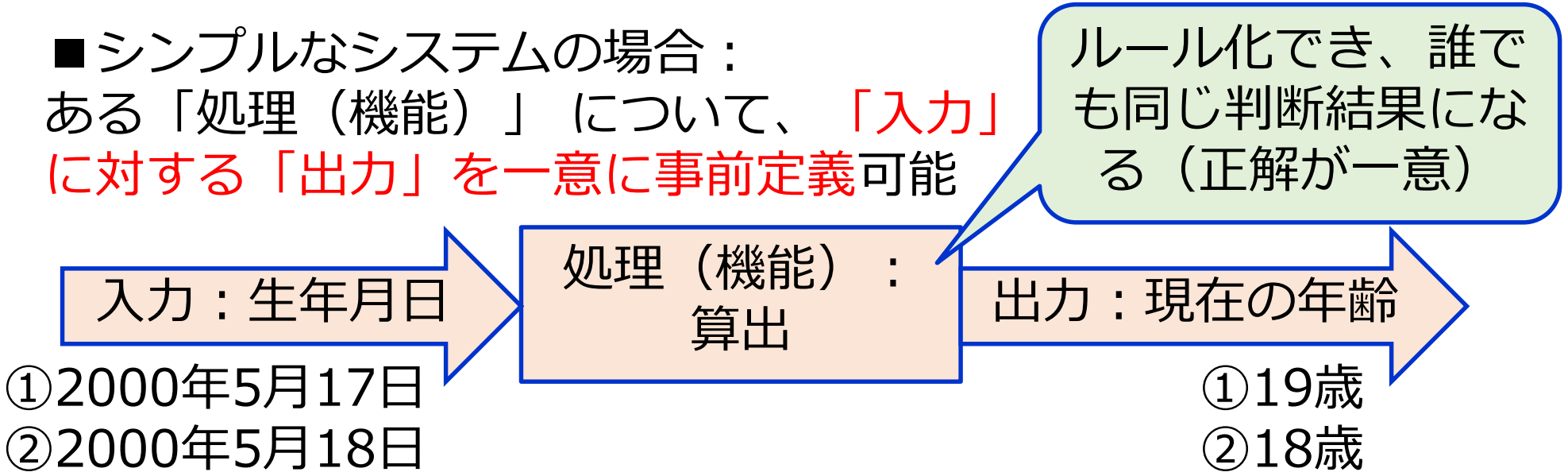
仕様を明確に定義し、仕様の適切さを検証



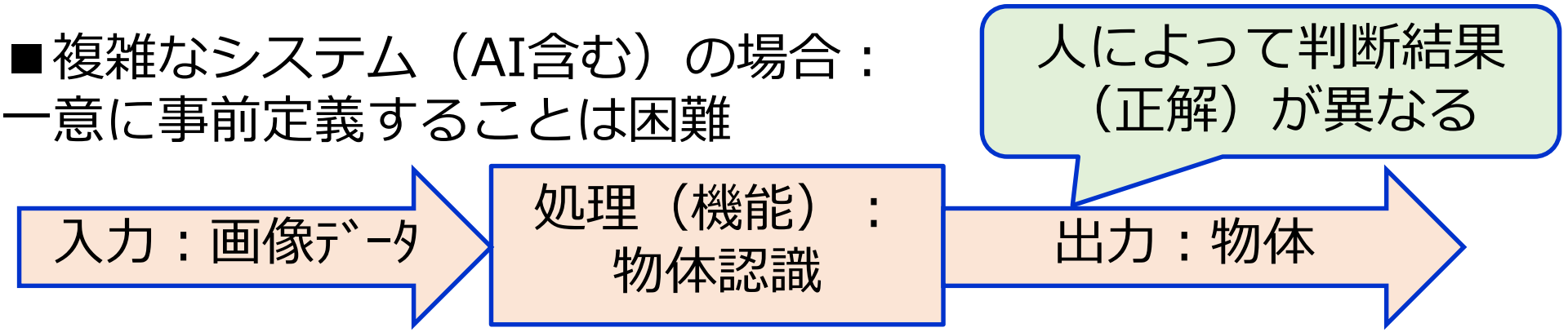
AIは「仕様を明確に定義する」ことが困難！
→ テスト困難!!

仕様を明確に定義するとは？

■シンプルなシステムの場合：
ある「処理（機能）」について、「入力」に対する「出力」を一意に事前定義可能



■複雑なシステム（AI含む）の場合：
一意に事前定義することは困難



３．AI搭載システムの安全設計パターン

AI搭載システムの安全設計パターン（案）

AI搭載システムの安全設計に、以下のいずれかのパターンを適用可能

パターン名称(仮)	安全設計パターン
【方針1】AIコンポーネント自体の安全性を示すパターン	
1) 機能安全開発パターン	AIコンポーネントを機能安全準拠開発する（例えば、IEC61508に準拠した開発）
2) 安全性評価パターン	AIコンポーネントについて、機能安全規格適合相当の安全性を示す
3) Proven in useパターン	AIコンポーネントについて、使用実績によって十分な安全性を示す(Proven in use)
【方針2】AIコンポーネント自体は非安全系とし、外部に安全メカニズムを設ける	
4) 安全機能パターン	制御（QM）と独立した安全機能（SIL）を搭載
5) 比較パターン	AIコンポーネント（QM）と他の非AIコンポーネント（SIL）による多重化比較
6) 防御設計パターン	動的にロバスト性の確認をするための防御設計（ガード）を搭載

※本資料では、QMは機能安全非対応、SILは機能安全対応という意味のイメージ

2）安全性評価パターン

- 安全センサシステム規格IEC62998の評価方法が参考になりそう
- IEC62998の概要
  - 従来の安全センサ規格（IEC61496、IEC60947-5-2）では未対象のセンサに対応
    - 新しいセンサ技術(レーダー、超音波)
    - 新しいセンサ機能(物体認識、物体位置計測)
    - センサの組み合わせ(フュージョン)
  - IEC61508（SIL）、ISO13849（PL）に対応
  - 屋内・屋外の環境に対応（公共の場所での人の保護）
- AIとセンサシステムの類似性
  - 人間の期待値とは異なる情報を出力する可能性がある

４．AI搭載システムの安全プロセス

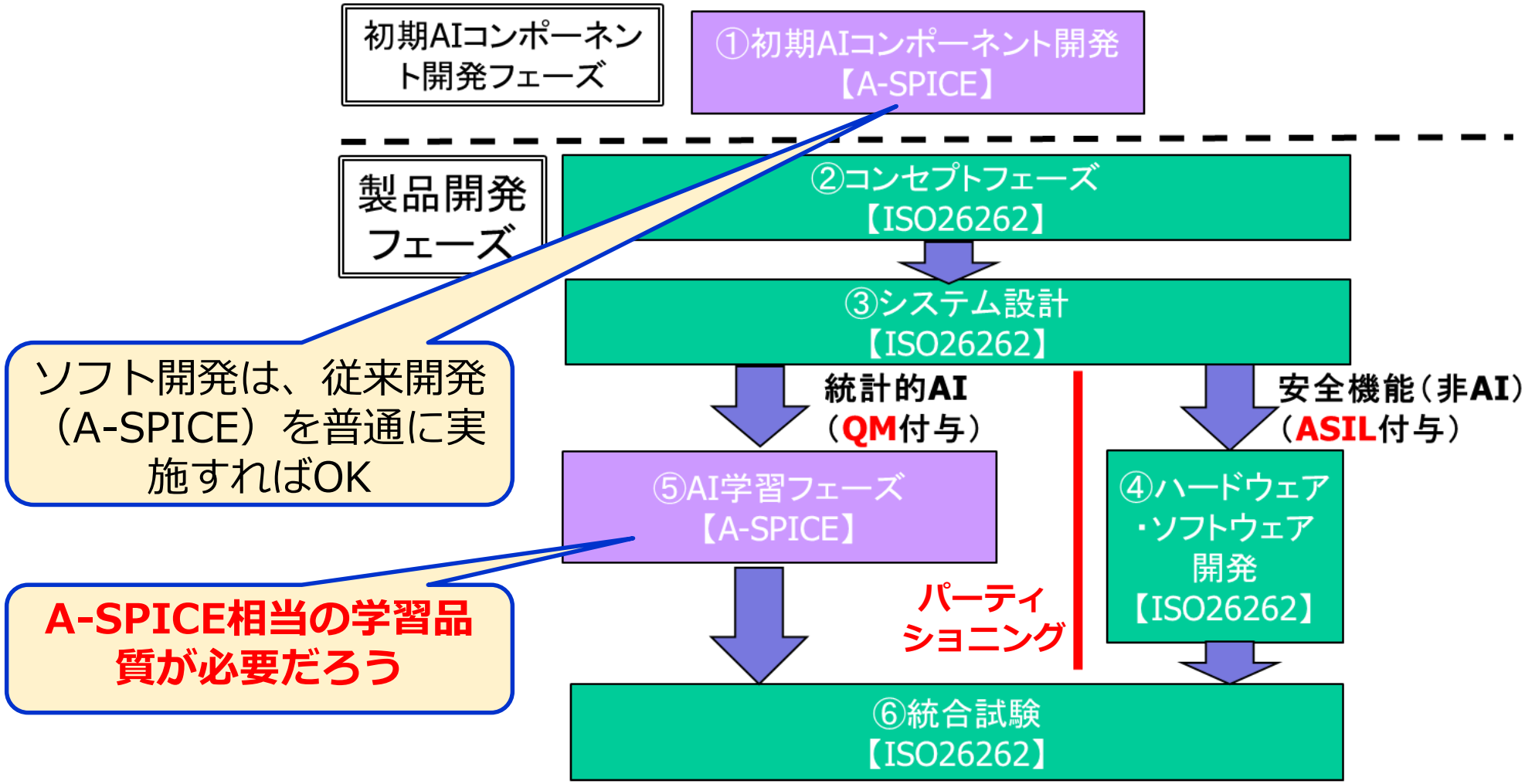
従来開発（機能安全、AutomotiveSPICEなど）

- 安全クリティカルシステムの開発では、“機能安全”が必須化
- 自動車のECU開発の品質担保と品質説明には、“AutomotiveSPICE” がトレンド

- いずれも以下を要求
  - Vモデル（要求/設計、実装、テスト、検証、etc）
  - 要求トレーサビリティ
  - 各種管理（プロジェクト、構成、課題、変更、etc）
  - 契約関連
  - 文書化 etc

AI搭載システム（Deep Learning）の開発プロセス概要

・以下は、AIを非安全系になるようなシステム安全設計を実施した場合の開発の流れ
・QM開発には、AutomotiveSPICE（A-SPICE）をテーラリングして適用



AI学習フェーズのAutomotiveSPICEテーラリング

- AutomotiveSPICEのエッセンスをAI学習フェーズに注入
  - プロセス上の押さえ所（リスク対策）
  - 品質説明（エビデンス）

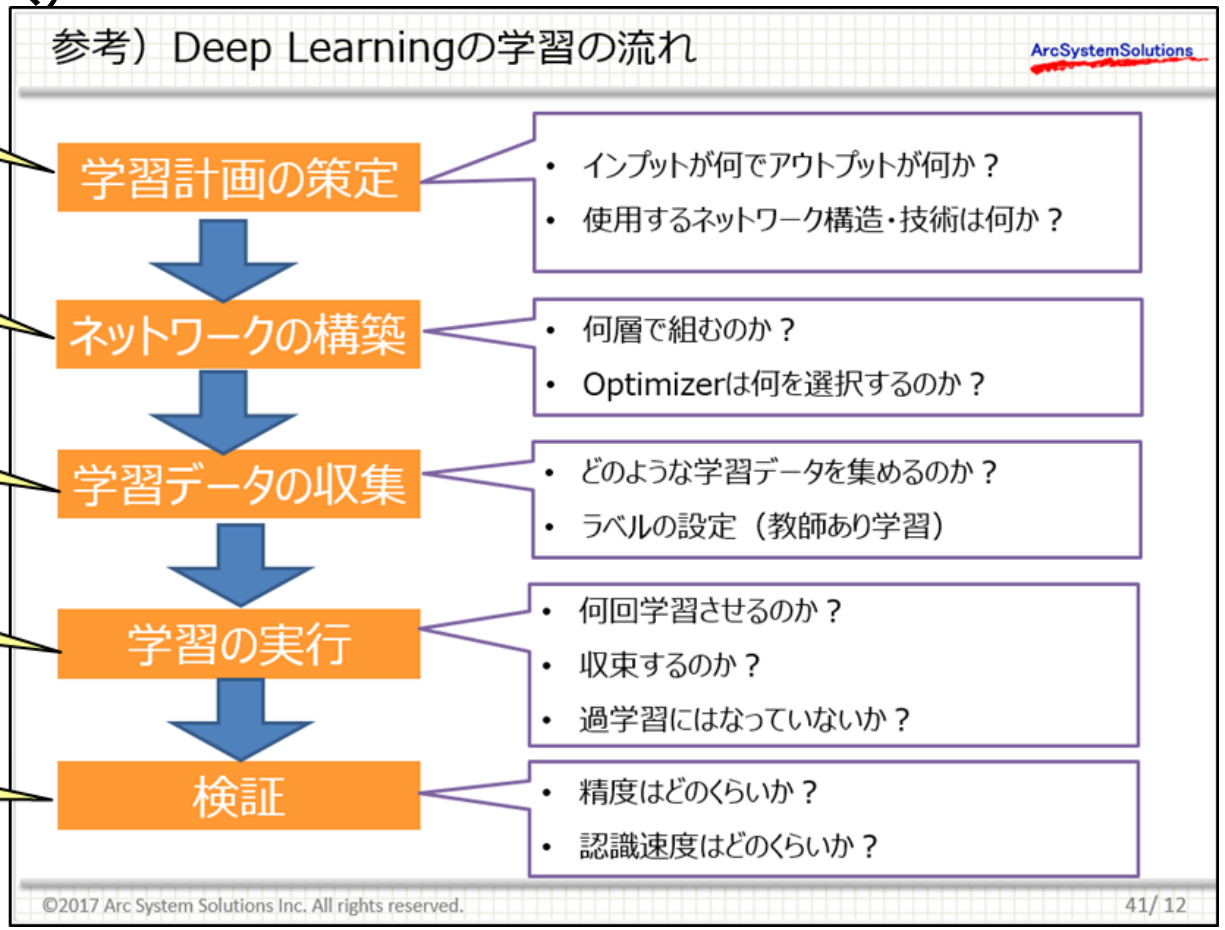
SWE.1 ソフトウェア要件分析

SWE.2 ソフトウェアアーキテクチャ設計

SWE.2 ソフトウェアアーキテクチャ設計

SWE.3 ソフトウェア詳細設計およびユニット構築

SWE.4 ソフトウェアユニット検証



5．まとめ

まとめ

- SEAMS Projectでは、“セーフティクリティカルなAI搭載システム” に対する安全性立証技術の研究成果を「AI搭載システムの安全設計ガイドライン」としてまとめました。

- 本日は、「安全設計パターン」「安全プロセス」の研究成果を一部紹介させていただきました。

今後の活動予定

- 研究事業
  - **SEAMSガイドライン Ver2.00発行（2020年3月完成、公開時期未定）**
    - 各設計パターンを具体的事例に適用・評価
    - SOTIF、IEC62998の適用
    - 機能安全対応のAI学習プロセスの検討
  - QA4AI「品質保証ガイドライン」を参考にFuzzing for AIの開発
  - ゴルフカートによる研究内容の総括実証実験

- 事業化
  - 自動運転関連支援
  - 自動車以外の分野への知見の応用

- QA4AIの活動とのコラボを希望
  - QA4AIは “品質”
  - SEAMS Projectは “安全” に特化し深堀り

ご清聴ありがとうございました

本内容の詳細に関しては、以下までご連絡ください。

株式会社ヴィッツ　機能安全開発部
森川 聡久　morikawa@witz-inc.co.jp

SEAMS