AI/IoT時代の安全設計
～STAMPの本質を考える～

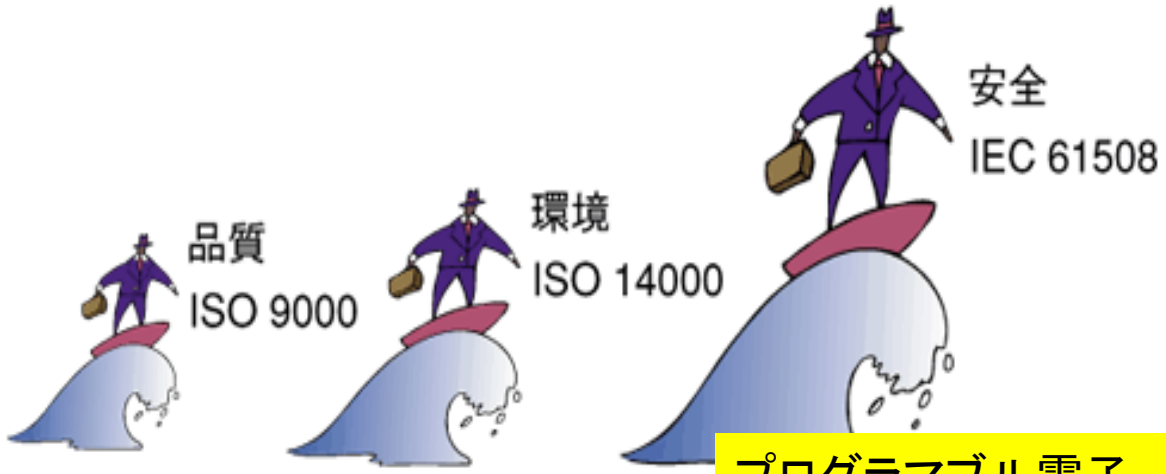
2022年11月16日

JASA・安全性向上委員会
兼本　茂（会津大学　名誉教授）
IPA/IoTシステム安全性向上WG主査（元）
福島県廃炉安全監視県民会議議長

一般社団法人 組込みシステム技術協会
Japan Embedded Systems Technology Association

背景(1)　安全設計へ押し寄せる波・国際規格戦争

（～2005年～）



プログラマブル電子機器の安全規格

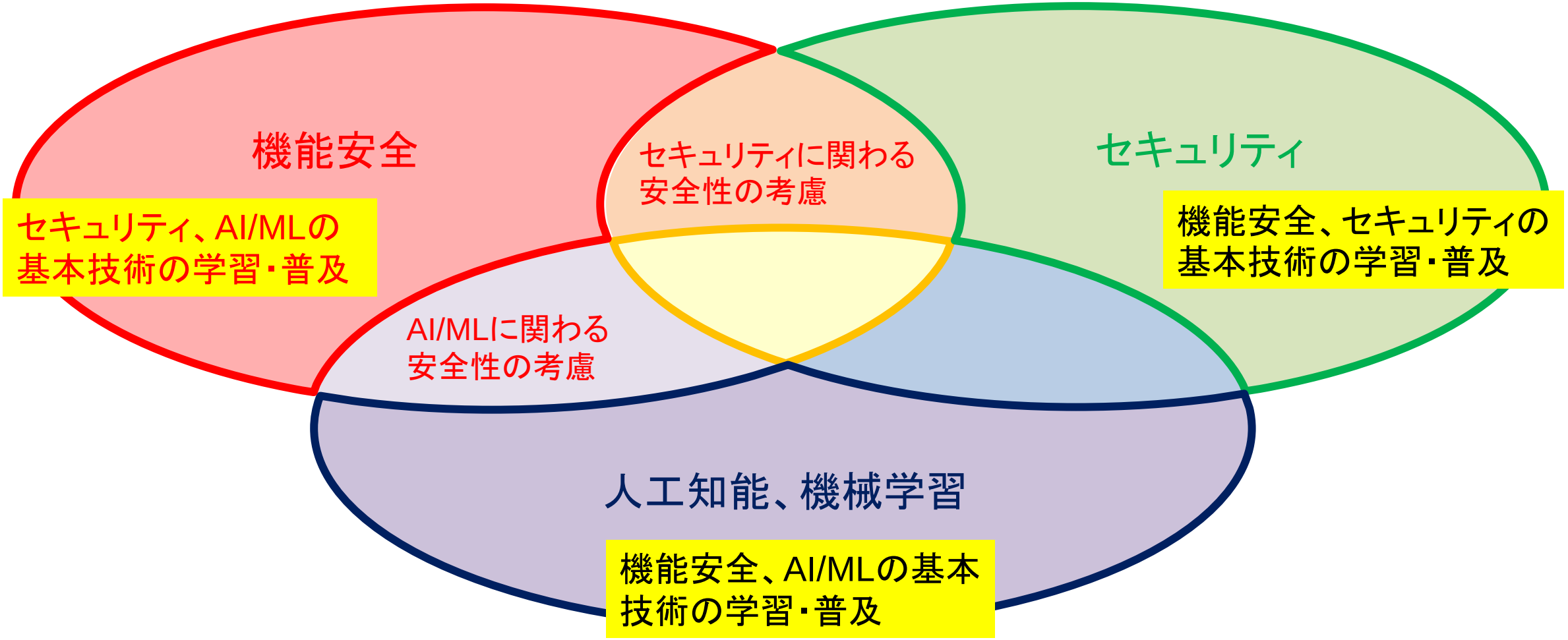
（～2019年～）

自動運転
ISO26262
SOTIF（Safety of the intended functionality）
SAE J3187

複雑システムの安全
STAMP/STPA

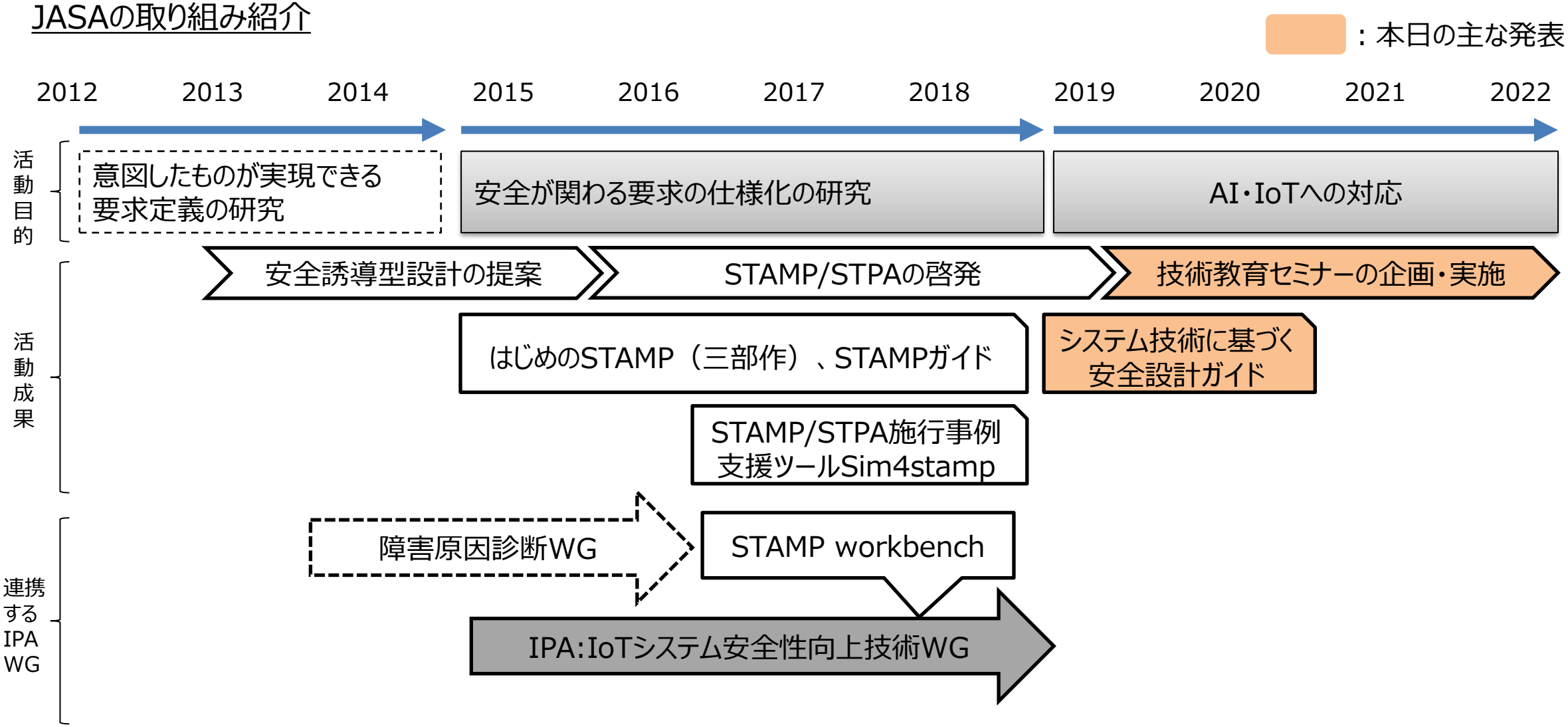
AI,IoT時代の安全性
Safety＆Security
Safety2.0（人と機械の協調安全）

背景（2）　安全設計の複雑化（JASA安全性向上委員会の取り組み）



JASA　安全性向上委員会の取り組み

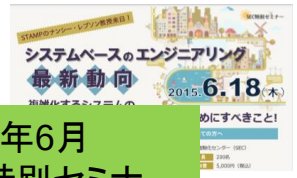
JASAの取り組み紹介



一般社団法人 組込みシステム技術協会 Japan Embedded Systems Technology Association

IPA・JASAでの活動（安全性向上委員会・IoTシステム安全性向上技術WG）

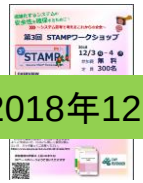
2015			2016			2017			2018			…	2024		
4月		12月		3月	4月		12月		3月	4月		12月		3月	



2015年6月
Sec特別セミナー
2016年1月
13thWOCS2

2016年12月

2017年11月



2018年12月

2019年12月

2020年11月
（リモート）

初級編

実践編

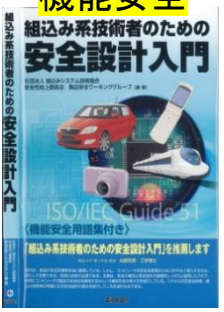
活用編

ツール

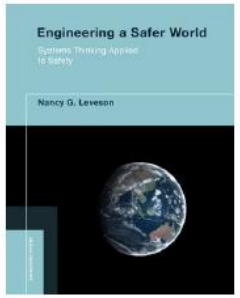
STAMPガイドブック

Engineering s Safer World（MIT:Nancy Leveson）の翻訳予定

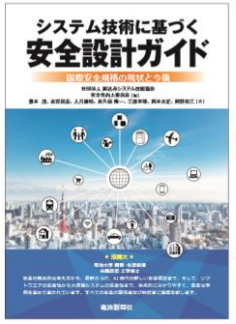
2010年
機能安全

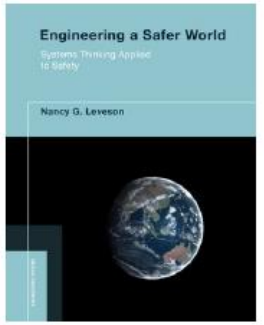


2011年　Engineering s Safer World（MIT:Nancy Leveson）



安全設計ガイド
(JASA、
2019年1月10日)



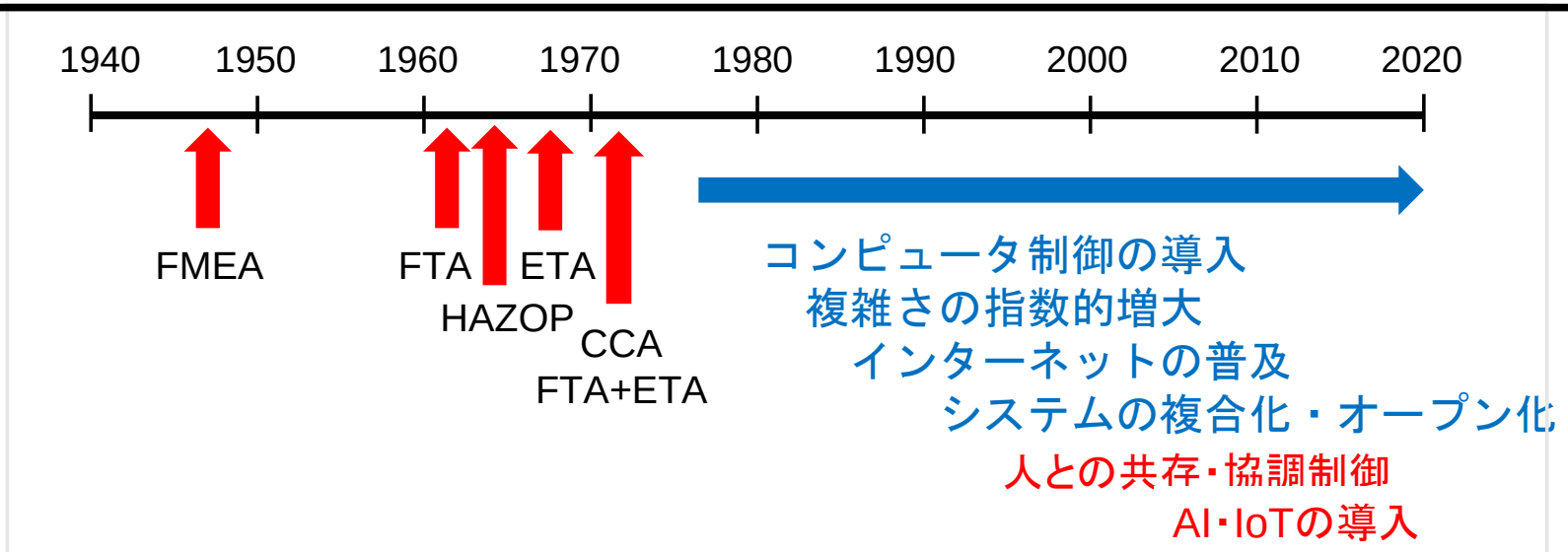


一般社団法人
組込みシステム技術協会
Japan Embedded Systems Technology Association

IoT時代の環境変化と新しい安全解析法　STAMP

Sec-Seminar(2015.6.18) Lecture by Nancy Leveson

現状の安全分析ツールは、40-65年も昔に開発されたものであり、現代の新しい技術の入った複雑な工学システムの安全分析には限界がある　→　***パラダイムシフトとしてのSTAMPの登場***



コンポーネント故障が事故を引き起こす

コンポーネント間のコミュニケーション・ミスマッチが事故を引き起こす

故障がなくても事故は起こる。ー＞創発的事故
◇要求仕様の欠陥やコンポーネントの性能限界
◇Worst Caseの想定不足

IPA—WG活動の中で得た格言

- コミュニケーションエラーが事故を起こす
- 要求仕様の欠陥（不十分さ）が事故を起こす
- 安全は創発事象であり還元論だけでは分
- 想定外を想定せよ
- 後知恵で責任者
- 確
- ヒ
- STAMP/STPAは、従来法（FTA,FMEA,HAZOP）と異なるパラダイムシフトで、システム思考に基づく安全分析法である
- 複雑なシステムは、抽象化と階層化によってのみ理解できる

複雑システムとは具体的には何か？

複雑システムとは何か？

- 複雑システムとは？
  - **組織・人間・コンピュータ（含むネットワーク）・機械**が組み合わさって目的を達成するシステム
    - ソフトウェア集約型、人間・機械協調制御型のシステムで、コンポーネント相互の連係ミスで事故が起こり、コンポーネントの性能不足でも事故が起こる。さらには、設計ミスや運用ミスでも事故が起こり、それらは、システムやそれに関わる組織の消滅につながる。
  - **複雑システムでは、故障がなくても事故は起こる⇒「創発事故」**
    - 設計ミスと単純に片づけてはいけない⇒設計ミスの理由まで考えるべき
    - ヒューマンエラーと単純に片づけてはいけない⇒ミスの理由まで考えるべき
    - 想定外の外乱と単純に片づけてはいけない⇒想定できなかった理由を考えるべき
- 複雑システムは、還元論(Reductionism)と全体論(Holism)の両者から理解すべきである
  - 全体論：コンポーネントの総和＜システム全体の特性　（システムは完全には把握できない）
    - 全体論では、個々のコンポーネントの特性が完全には分からなくても事故を抑えられるという立場をとる（例えば、熱が出たら冷やすなどの「症状ベースの対応」）
  - 還元論；コンポーネントの総和＝システムの全体の特性　（システムを完全に把握できる）
    - 還元論では、個々のコンポーネントが確率的に故障するという立場で、多重化や自己診断で、事故を最小限に抑制するという立場をとる
- 還元論に基づく安全工学は実績があるが、全体論に基づく安全工学は未成熟
  - 安全工学のパラダイムシフトとしてのSTAMPの期待

IPA—WG活動の中で得た格言

- コミュニケーションエラーが事故を起こす
- 要求仕様の欠陥（不十分さ）が事故を起
- 安全は創発事象であり還元
- 想定外を想定
  -
-
- ヒューマンエラーを表面的に扱う
- STAMP/STPAは、従来法（FTA,FMEA,HAZOP)と異なるパラダイムシフトで、システム思考に基づく安全分析法である
- 複雑なシステムは、抽象化と階層化によってのみ理解できる

システム思考とは具体的には何か？

米国の原子力潜水艦の安全プログラム（SUBSAFE)の事例で考えてみる

(N.Leveson:”Engineering of a Safer World　14章）

米国の潜水艦の戦闘外の事故の歴史

原子力潜水艦スレッシャー号の事故：1963年4月10日、深海潜水テスト中に、112人の要員と17名の民間人を乗せたまま沈没し、全員が亡くなった。

SUBSAFEプログラム（1963年6月3日に創設され、同年12月20日に発布）開始後、事故は一度も起こっていない。（たった半年で開発・発布）

SUBSAFEプログラム開始後の唯一の事故（1968年スコーピオン号）は、冷戦のプレッシャーで監査対象外であった。

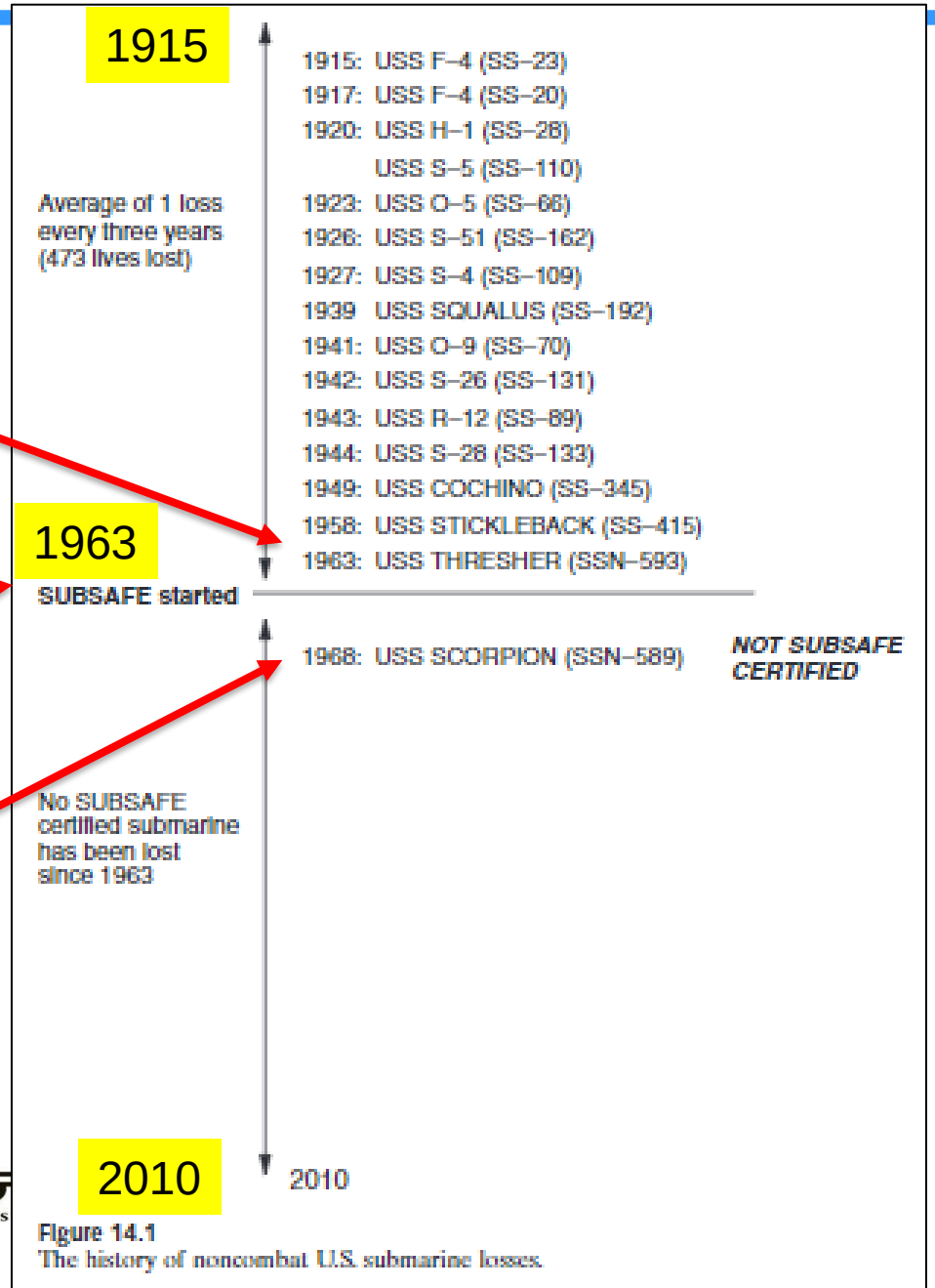


Figure 14.1
The history of noncombat U.S. submarine losses.

一般社団法人 組込みシステム Japan Embedded Systems

何故、14章　SUBSAFEの事例か？

- 米国原子力潜水艦の安全プログラムなのに・・・
  - 原子炉の安全については、何も述べていない
  - STAMPの教えを体現している成功例なのに、制御構造図も非安全コントロールアクション（UCA)もほとんど記載がない
  - しかし、驚くほどの成功を収めている
- 本稿の目的・・・
  - SUBSAFEとSTAMPの本質的な共通点を考察してみる

STAMPが引用されている二つの文章

- SUBSAFE inculcates the basic STAMP assumption that systems change throughout their existence.
  - SUBSAFEは、システムが存在している間ずっと変化し続けるというSTAMPの基本的な教えを体現している。
- In STAMP terms, only the lowest level of the safety control structure was being audited and not the other components. After that time, biennial audits were conducted at all levels of the safety control structure, even the highest levels of management.
  - STAMPの用語で言えば、最も低いレベルの安全コントロールストラクチャーだけが監査され、（組織や人を含んだ） 他のコンポーネントは監査されないということであった。その後、隔年監査が導入され、管理体制の最も高いレベルまで含んだ安全コントロールストラクチャーの監査まで実施されるようになった。

- Systemic（全体的）：4回、Systematic（体系的）（0回）使用

スレッシャー号の事故詳細

■ The accident was thoroughly investigated including, to the Navy's credit, the systemic factors as well as the technical failures and deficiencies.

- この事故は、海軍の面目をかけて、技術的な失敗や欠陥だけでなく、システミックな要因も含めて徹底的に調査された

■ 事故の直接要因と組織要因

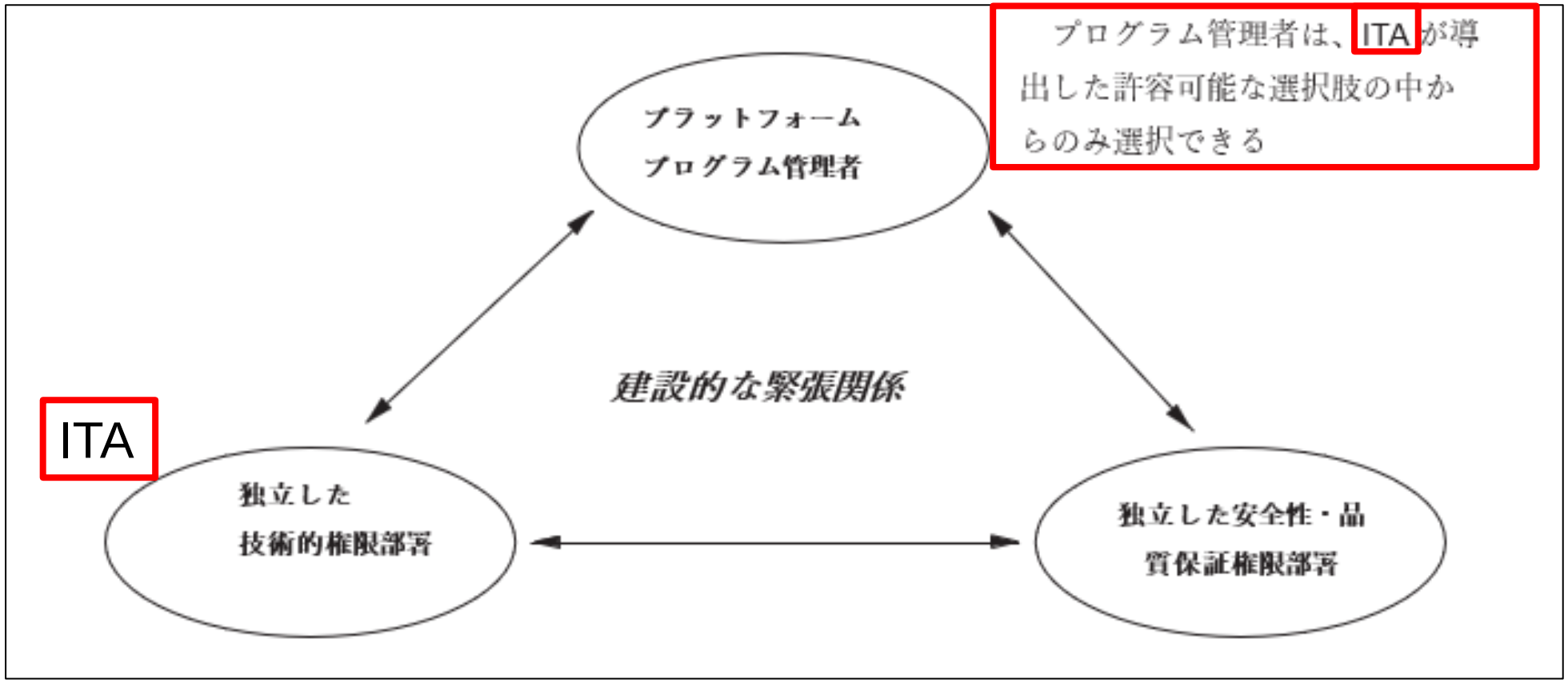
- （**直接要因**）溶接の代わりに銀ろう付けに頼っていた塩水配管システムの接合部の欠陥が、エンジンルームの浸水を引き起こしたが、乗組員は浸水を止めるために重要な設備にアクセスすることができず、浮上もできなかった。対策として、隔離バルブを中央操作パネルから遠隔で閉じることができる浸水制御レバーを設置するよう提言した。
- （**組織要因**）調査結果からは、仕様・建造体制・保守実施体制の不備、不適切な建造・保守活動の文書化、さらには、運用手順の不備が見つかった。その一例として、スレッシャー号では3000箇所の銀ろう付け配管の継ぎ手があり、最後の造船所での保守の際、これらの継手のうち145個が超音波テストで点検され、その14%が規格をはずれた接合状態であることが分かった。この結果から船全体では400以上の継ぎ手が規格外であった可能性がある。この状態で、船は出航することが許された。

安全目標の明確な設定(損失の定義)

- SUBSAFEの安全目標は、米国の潜水艦が緊急時に確実に浮上し帰港できるような活動に限定
  - Watertight integrity of the submarine's hull.（潜水艦の船体の完全な水密性の確保）
  - Operability and integrity of critical systems to control and recover from a flooding hazard.（洪水ハザードのコントロールと回復のために重要なシステムの完全な運用確保（中央制御室からの遠隔操作））
- ミッションの達成は大事ではあるが、SUBSAFEの焦点ではない。同様に、火災安全、兵器の安全、労働者の安全衛生、原子炉システムの安全はSUBSAFEには含まれない。通常のシステム安全プログラムで管理する。
- 要求事項（技術以外に、組織の在り方、監査まで幅広くカバー）
  - 経営　・管理、組織、技術、固有の設計、材料のコントロール、製造、検証試験、作業管理、監査、認証

SUBSAFE 権力の分離（「3本脚の腰掛」）

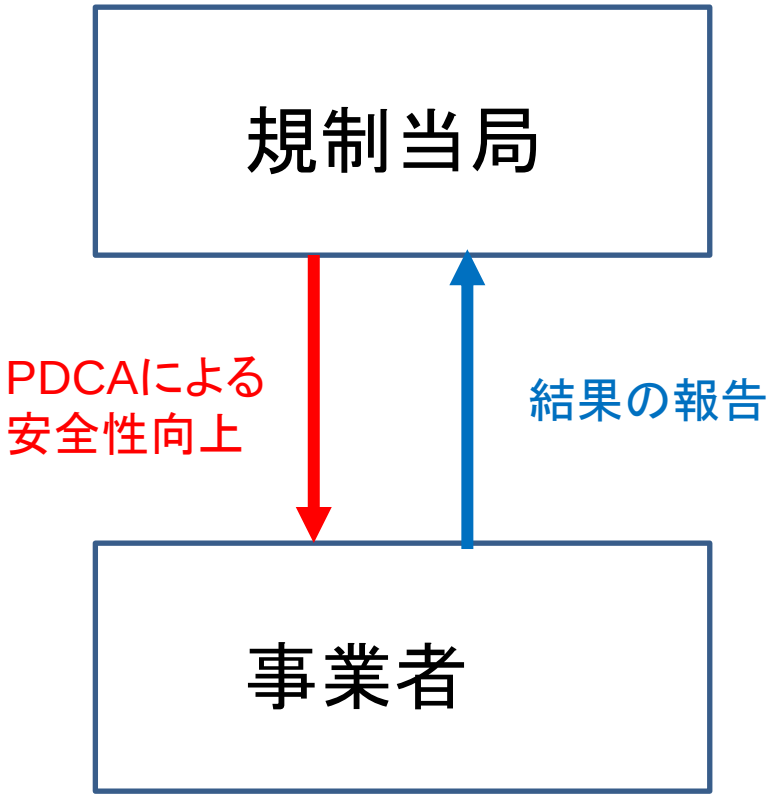
SUBSAFE has created a unique management structure they call *separation of powers* or, less formally, the *three-legged stool.*
→SUBSAFE は、「権力の分離」と呼ぶ独自の管理体制を構築した。わかり易く言い換えると、3本脚の腰掛のように、独立した脚（権力）で安全を支える仕組みである。
This management structure only works because of support from top management
→この管理体制は、経営上層部からの支援があって初めて機能する。



監査手順とアプローチ

A biennial NAVSEA internal audit gives the field activities　a chance to evaluate operations at headquarters. Headquarters personnel must be willing to accept and resolve audit findings just like any other member of the nuclear submarine community.
→隔年で行われるNAVSEAの内部監査の中で、現場の活動一つとして、本部の指揮・運用を評価する機会が生まれた。 本部要員は、他の原子力潜水艦共同体のメンバーと同じように、監査結果を受け入れ、解決する姿勢が必要になった。

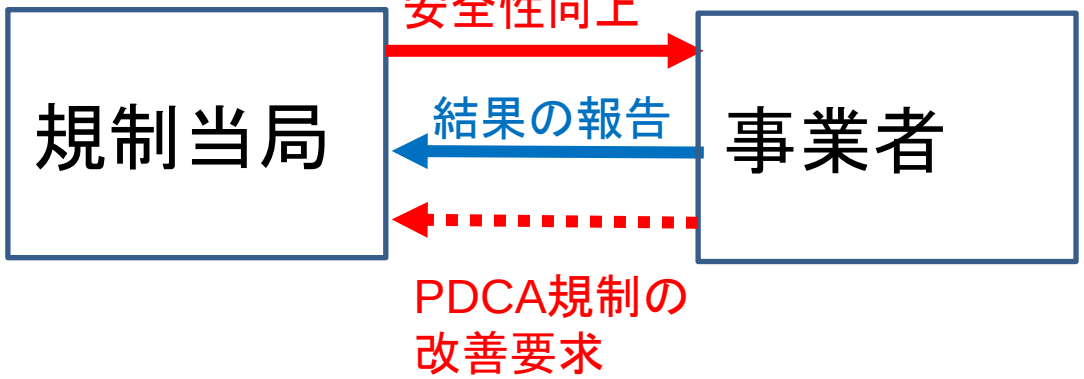
比喩的表現



監査手順とアプローチ

A biennial NAVSEA internal audit gives the field activities　a chance to evaluate operations at headquarters. Headquarters personnel must be willing to accept and resolve audit findings just like any other member of the nuclear submarine community.
→隔年で行われるNAVSEAの内部監査の中で、現場の活動一つとして、本部の指揮・運用を評価する機会が生まれた。 本部要員は、他の原子力潜水艦共同体のメンバーと同じように、監査結果を受け入れ、解決する姿勢が必要になった。

比喻的表現



安全という観点からは、規制当局と事業者に階層性はない。
事業者から規制当局への要求（コントロールアクション）もあるべき。

まとめ　SUBSAFEとSTAMP/STPAの本質的な共通点

- 安全目標の明確な設定（損失の定義）
  - 潜水艦が緊急時に確実に浮上し帰港できるような活動に限定
  - 技術以外に、組織の在り方、安全文化、監査権限まで幅広くカバー
- 安全プログラムの設計ー＞安全制御構造図と安全責任の明示化
  - 権限の分離（三本脚の腰掛）
  - 技術的要因だけでなく、安全にかかわる組織の管理体制やシステミックな要因まで対象にする
  - ライフサイクルにわたる継続的な監査の実施
- 文化的な基本（個人の誠実さと責任感、STAMPの基本思想の一つ）
  - 疑問を持つ姿勢
  - 重要事項に対する 自己評価
  - 学んだ教訓と継続的な改善
  - 継続的な訓練
  - 権力の分離（安全へのチェック＆バランスと適切な配慮を提供する管理体制）
- STAMP/STPAという手順の背景にある安全哲学とシステム思考が同じ

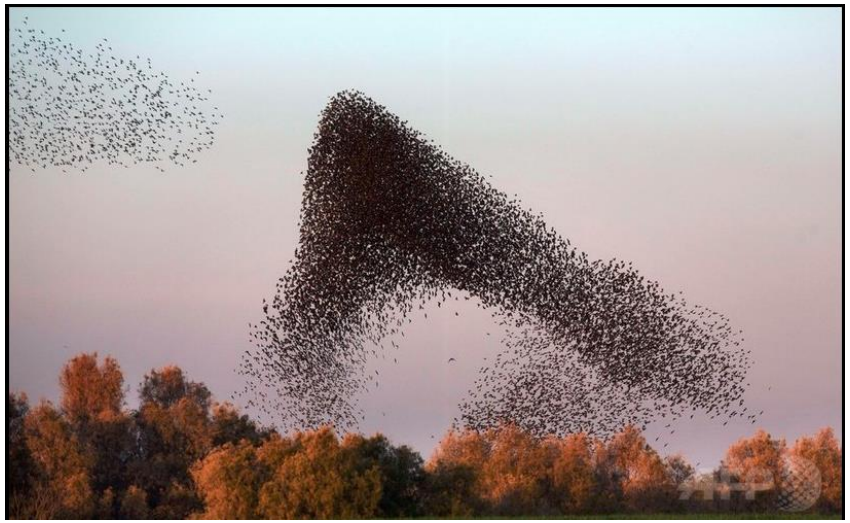
IPA—WG活動の中で得た格言

- コミュニケーションエラーが事故を起こす
- 要求仕様の欠陥（不十分さ）が事故を起こす
- 安全は創発事象であり還元論だけでは分析で
- 想定外を想定せよ
  - 後知恵で責任者を追求
- 確率論は
- ヒ
- ST 従来法（FTA,FMEA,HAZOP）と異なるパラダイムシフトで、システム思考に基づく安全分析法である
- 複雑なシステムは、抽象化と階層化によってのみ理解できる

複雑システムでは、創発・想定外ははずせない
しかし、創発、想定外、後知恵とは何か？
どれも事故後の評価ではないか？

創発事象の事例：鳥の編隊飛行

～先頭に追随する鳥は飛行の省エネを目指している。ミクロシミュレーションで説明可能 ～



ムクドリの群れ

ガンの編隊飛行

チャレンジャー号の爆発事故（1986年）

ブースターロケットのジョイント部で使われていたOリングが,低温環境での使用で破壊
→シール効果が不十分となって燃料が漏れ,これに炎が燃え移り爆発
→一見,単純なOリングの部品故障に思えるが、問題の根は深い
→NASAとOリング製作会社は、低温におけるOリングの硬化の問題を予め知っていた
→Oリング製作会社は、当日の気象条件を見て、打ち上げを中止すべきと勧告
→しかし,打ち上げが強行された

→NASA内での現場の技師と管理者との間の意志疎通が不十分
→次年度予算を取るための圧力で,技術上のリスクを低く見積もるという認知バイアス

→複数のコミュニケーションミスと機械の故障,人間や組織の判断ミスなどが複雑に絡まった結果発生した事故といえる
→機械・人・組織の相互作用をモデル化できれば、工学的・社会学的な制御は可能

盗難防止用ロックシステムのついたレンタカーのトラブル例（ワシントン州でのカーフェリー内）

レンタカー盗難防止のため、エンジン停止時に車が移動すると、車を運転できないようにロックする機能が、レンタカー管理会社により設置されていた。
カーフェリーで移動した際、このロック機能で運転不能になった。レッカー車で車を移動するまで、フェリーが運航中止を余儀なくされた。
→設計段階での要求仕様の欠陥で、システムコンポーネントの故障ではない。
→管理会社にとっては便利な機能だが、レンタカー利用者やフェリー運航者といった関係者（ステークホルダー）にとっては問題のある機能である。
→コンセプト設計段階で、ステークホルダーが誰か、それぞれの損失が何か、を分析していれば、防げたトラブルである。（設計欠陥ともいえる）
→N.Levesonは、この事例も創発事故の一例として上げている。（コンセプト段階での検討の未熟さが余計な損失を生む）

アメリカン航空965便墜落事故（1995年）

http://www.shippai.org/fkd/cf/CA0000293.html

人とコンピュータの相互作用のミスの事例
（不適切なinterface設計と、不適切なinteraction設計）

進入コースの入力ミス
パイロット「R」
→ 「ROMEO」と解釈（本来は「ROZO」）
→ 山への激突

教訓
人と人のコミュケーションでも、意図が違えば誤解があるように、人と機械（コンピュータ）のコミュニケーションは思うほど簡単ではない！

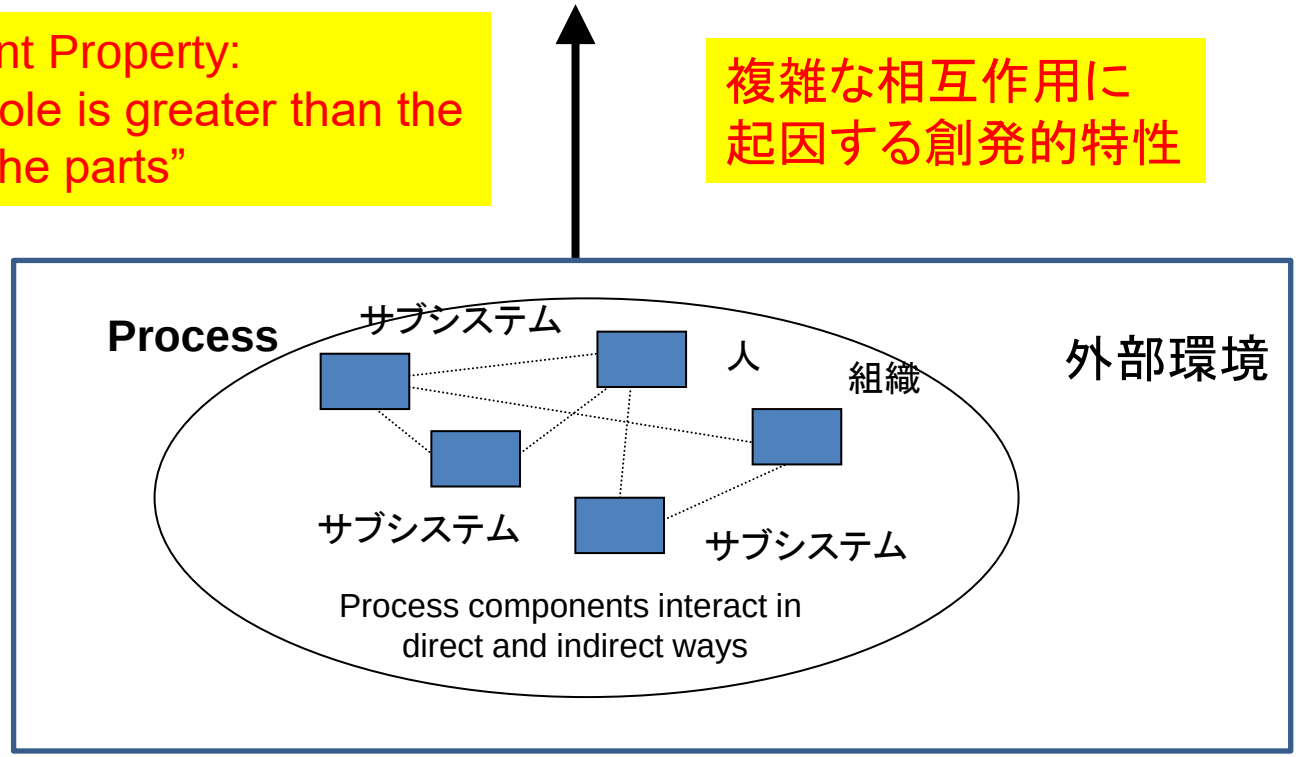
コロンビアのアルフォンソ・ボニーリャ・アラゴン国際空港は、山岳地帯の中のおおよそ南北方向に走る谷間にある

複雑システムの安全とセキュリティ：創発特性
～複雑システムの事故は、創発的（想定外）に起こる～

Emergent properties
(arise from complex interactions)

Emergent Property:
“The whole is greater than the sum of the parts”

複雑な相互作用に
起因する創発的特性



Safety and security are emergent properties

“The whole is greater than the sum of the parts”
をどう解釈するか？

物理の理論では、個の集合で全体が成り立つ。しかし、個の集合で説明できない事象が多く発見され、それを「創発」と名付けた。

しかし、ミクロの計算科学の進歩で説明できる事象も多くなった。工学的には、多くの場合、個の相互作用を解明し、全体の創発事象を説明・制御し、問題解決をしている。

個の相互作用や複雑性に対する知識不足は、事故後の後知恵で考えると、設計欠陥とみなされることもある。

Sec-Seminar(2015.6.18) Lecture by Nancy Leveson

ではどうするか？

Emergent Property:
“The whole is greater than the sum of the parts”

個の相互作用や複雑性に対する知識不足は、事故後の後知恵で考えると、設計欠陥とみなされることもある。

意図した機能の性能限界、合理的に予見可能なミスユースなどを丁寧に考えて設計する（ISO21448/SOTIF）

できるだけWorst Caseで考え、最悪の事故を避けるように設計する（レジリエンス工学）

想定外を想定して設計する（STAMP/STPA）
◆安全制御構造図で安全を可視化してステークホルダー間で相互レビューする
◆抽象化と階層化を利用して対象システムをステークホルダー全員が理解できるように表現して、安全であることを相互レビューする
◆システム全体を俯瞰的にとらえて、最悪の事態を防ぐ安全制御メカニズムを実装する
◆UCAを利用して漏れのない設計をする

STAMP/STPA手順（Handbook）

(1)Step-1

分析の目的の定義
［損失（アクシデント）の定義、ハザードとシステム安全制約の定義］

(2)Step-2

制御構造図と安全コントロールアクションのモデル化

(3)Step-3

非安全コントロールアクション（UCA)の同定＋（コンポーネント安全制約への展開）

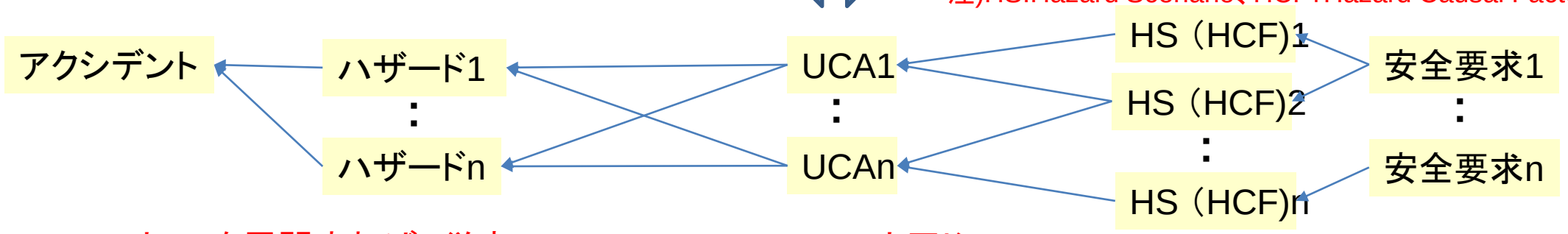
(4)Step-4

ハザード誘発シナリオ（損失シナリオ）の同定＋（コンポーネント安全制約・安全要求への展開）

HSに関わらずUCAを回避する（安全制御工学、レジリエンス工学）

HS(HCF)を全て評価し回避するのが信頼性工学

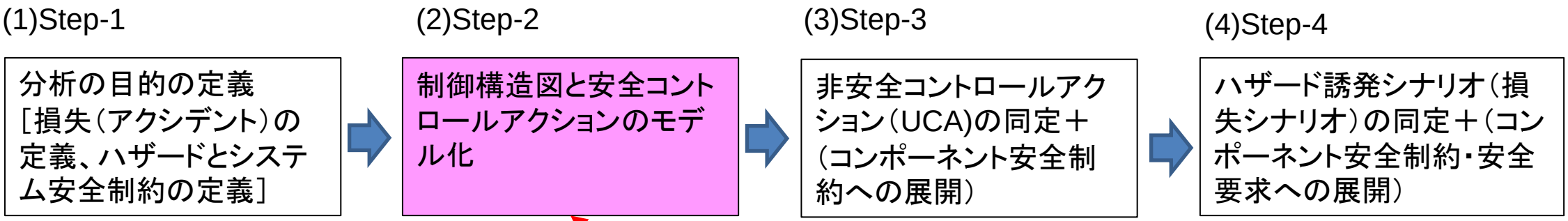
注)HS:Hazard Scenario、HCF：Hazard Causal Factor



UCAとHSを展開すれば、従来のFTA、FMEA、HAZOPと同じ

後知恵で考えると、どんな手法で考えても同じ、だが、
事故前に発想できるかどうかの違いがある（プロアクティブな安全設計）

STAMP/STPA手順（Handbook）



HSに関わらずUCAを回避する（安全制御工学、レジリエンス工学）

HS(HCF)を全て評価し回避するのが信頼性工学

注)HS:Hazard Scenario、HCF：Hazard Causal Factor

アクシデ

制御構造図による安全メカニズムの可視化と共有

- 目的（どんな事故を防ぐか）
- 各要素の安全責任と権限を定義
- 安全責任を達成するためのCAと、CAを作るためのFBを明示化
- 外部環境は最悪の状態を仮定
- 第三者の評価が可能な抽象化と階層化モデル

ー＞システム思考の本質

HCF)1

HCF)2

HCF)n

安全要求1

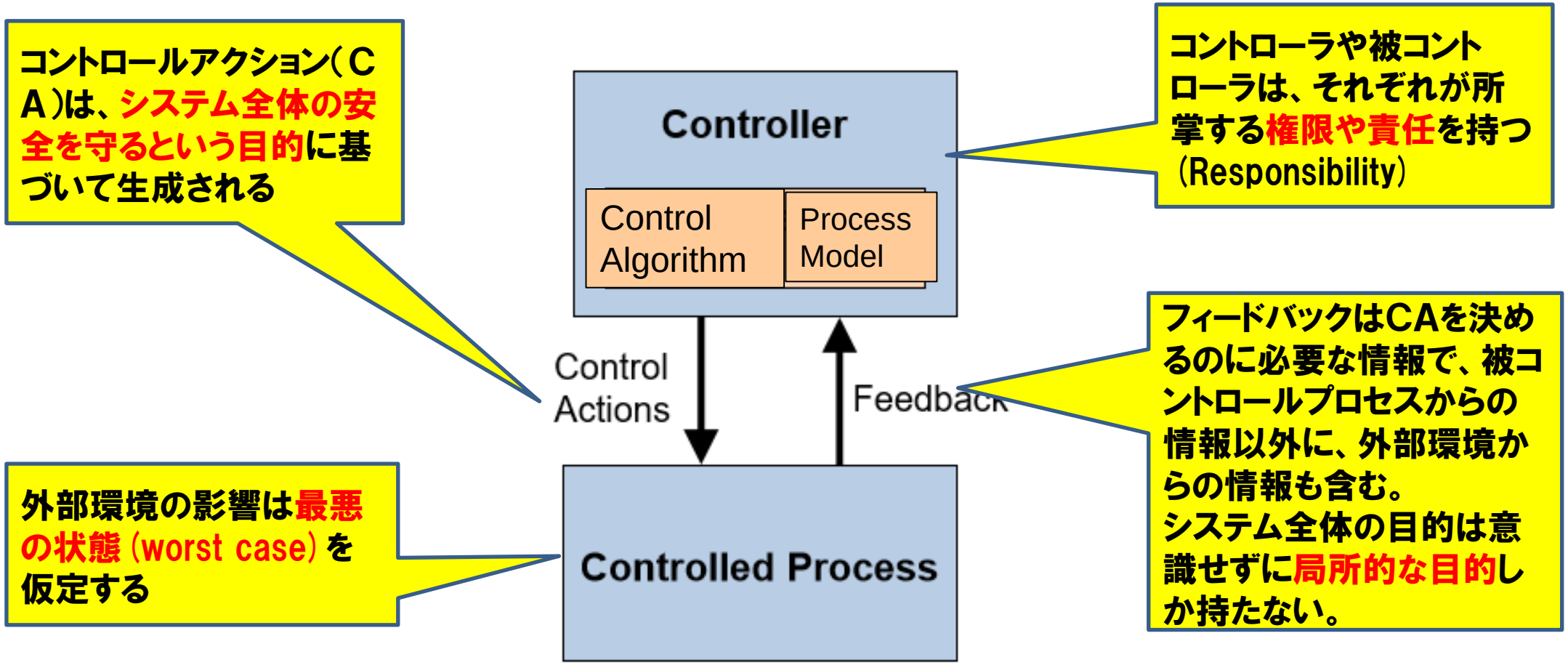
：

安全要求n

UCA

事故前に発想できるかどうかの違いがある（プロアクティブな安全設計）

Basic STAMPの制御構造図（四つの大事な考え方）



一枚の制御構造図で、システム全体の安全制御メカニズムが見通せないといけない
→　立場の異なる人の間で共通の理解に基づいて相互レビューができ、想定外の事故を低減できる

事例(1)　京浜急行踏切事故(2019年9月5日11時40分)

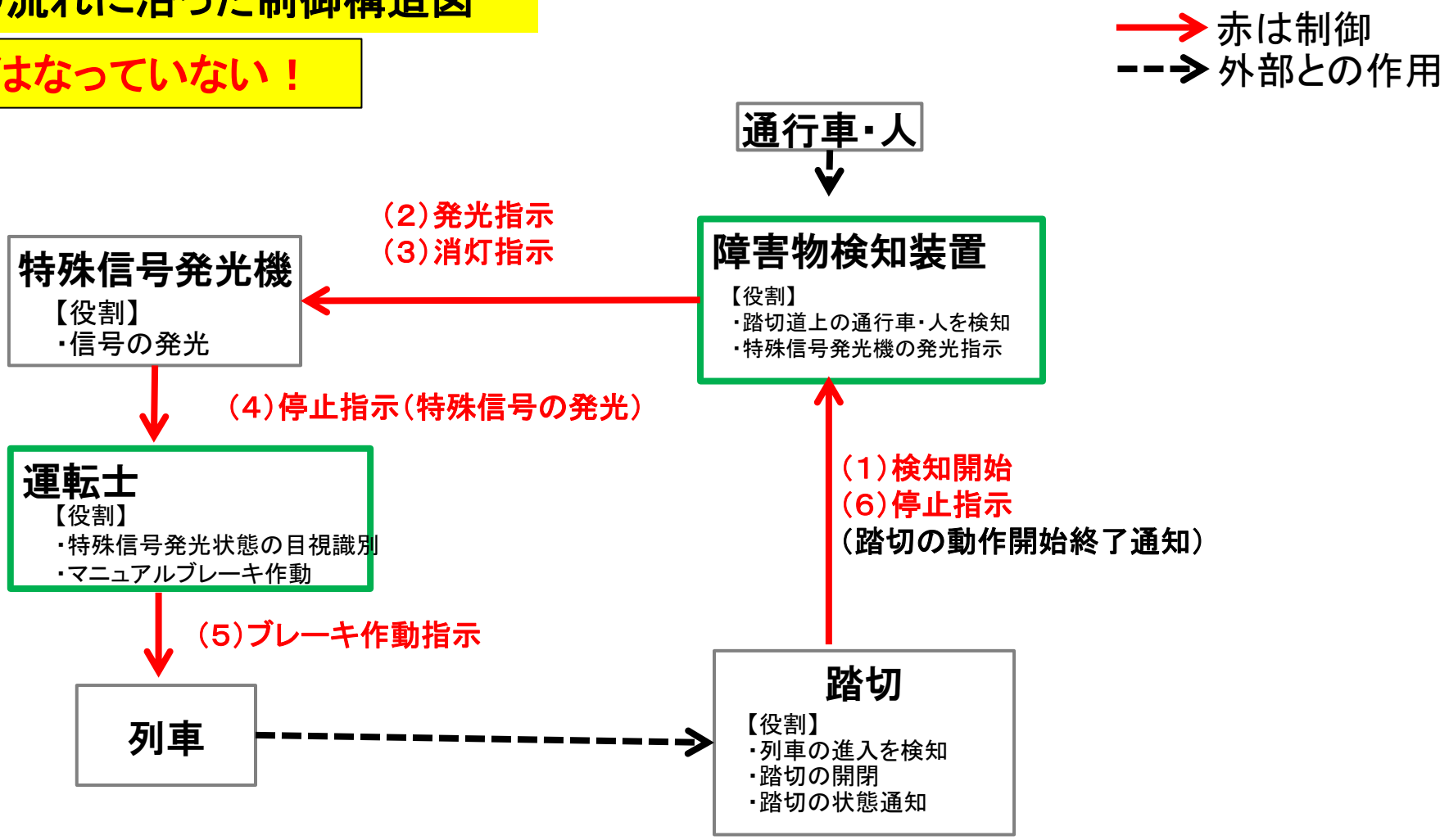
神奈川新町～仲木戸間の踏切で大型トラックと衝突

踏切に設置した「とりこ」検知装置で停止信号を点灯
・カーブがあるが、570m手前で確認可能
・時速120kmで走行時の制動距離は517.5m(空走距離は33m／秒)
・社内規定では、信号機点滅を確認した場合、「速やかに停止」としていたが、今後、「直ちに非常ブレーキ」に変更するとのこと
・運転士は、「速やかに停止」の場合、常用ブレーキか非常ブレーキ併用かの判断が必要

とりこ検知事例（Step-2　ハードウェア視点からの制御構造図の構築）

“とりこ検知”の流れに沿った制御構造図

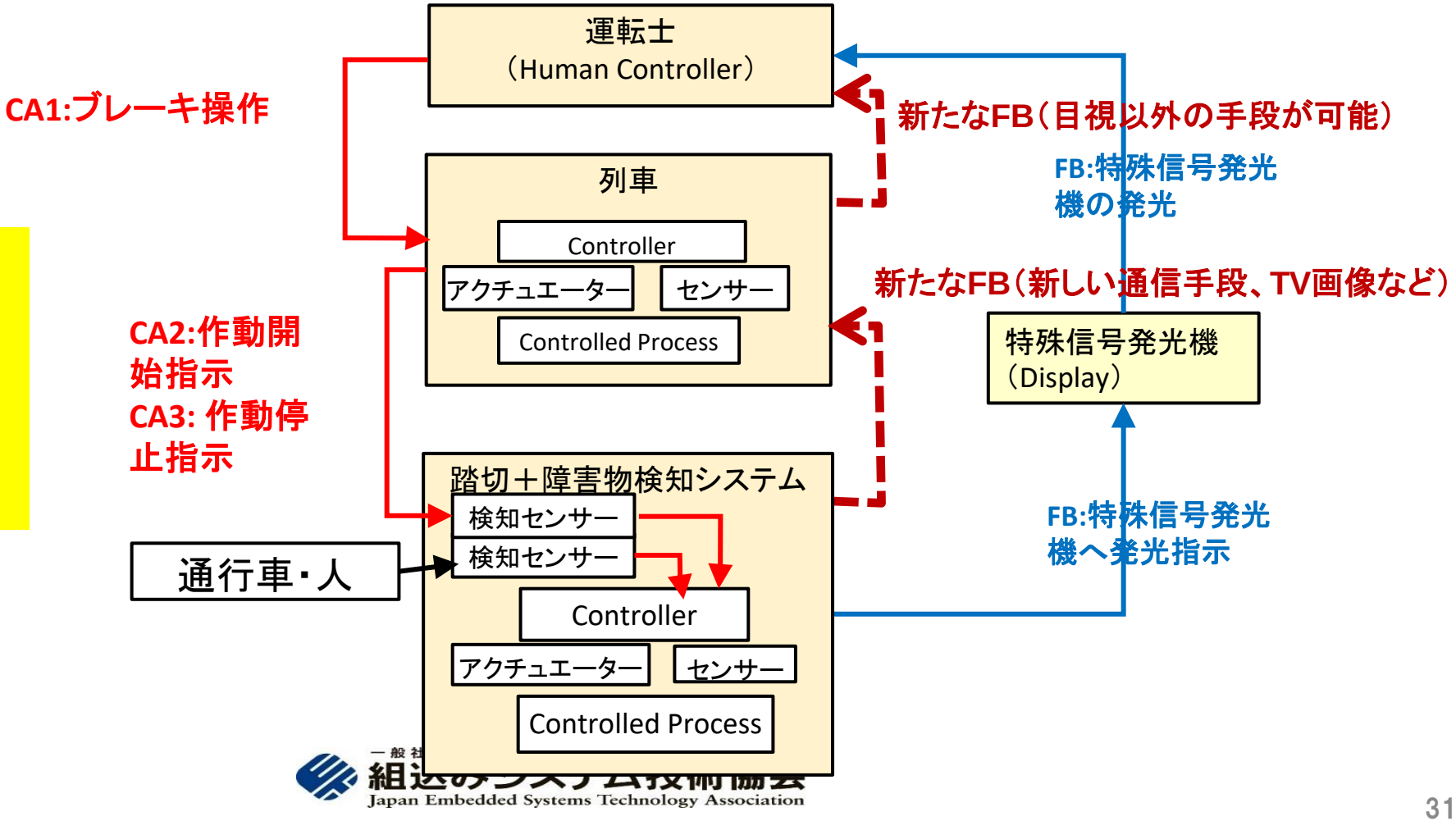
階層的構造にはなっていない！



とりこ検知事例（Step-2　運転士の視点からの制御構造図）

運転士を中心にした**階層的制御構造図**ー＞安全権限の階層化と可視化ー＞
多様な視点での安全設計の検討に役立つ

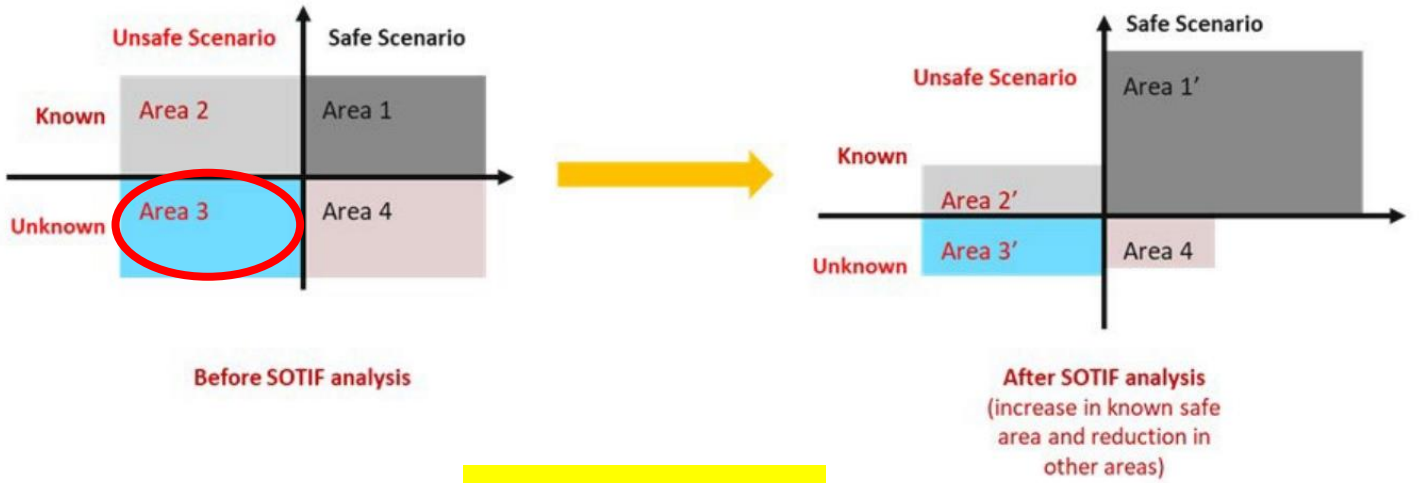
運転操作を高度な制御装置とみる
（三階層の制御構造図）
↓
次世代のより安全な制御手段の発想につながる。



最後に(1)　複雑システムの安全分析の難しさ

分析の目的
SOTIF：Area-2,3,4を減らして、Area-1（Known Safe）に持ってゆく。
STPA：Area-B,C,Dを減らして、Area-A(Known knowns）に持ってゆく。
ただし、Area-2,3,4やArea-B,C,Dをゼロにはできない。（絶対安全はない！）
◇STPAはArea-3やArea-Dの低減に適している。→想定外や創発事象を減らし、後知恵の批判を緩和する。

SOTIF（ISO21448）



SAE J3187

Figure 19 - Sc ness and safety as axes

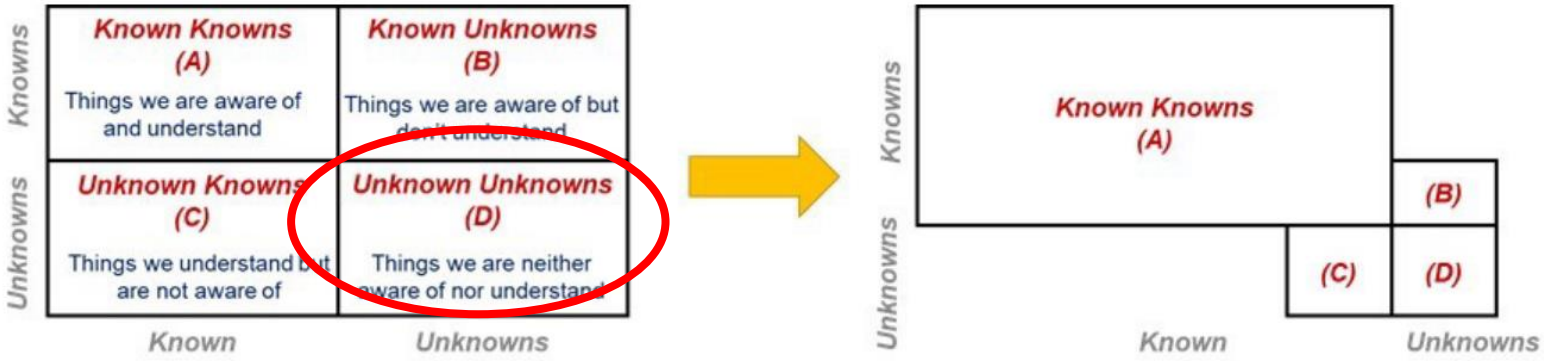


Figure 20 - Scenario quadrant with understanding and awareness as axes

最後に(2) JASA 安全性向上セミナー 2022

第1回	安全の基礎とSafety&Security国際規格の動向	2022年9月28日(水)14:00～17:00
第2回	事例で学ぶSTAMP/STPA(入門編)	2022年10月 5日(水) 14:00～17:00
第3回	事例で学ぶSTAMP/STPA(中級編)	2022年10月 19日(水) 14:00～17:00
第4回	事例で学ぶSTAMP/CAST	2022年11月9日(水)14:00～17:00
第5回	安全とセキュリティ	2022年12月7日(水)14:00～17:00