Society5.0 リファレンスアーキテクチャにおけるリスク管理

小林　洋[1]　　森下優子[2]

概要：Society5.0 リファレンスアーキテクチャ（RA）は，これからの情報システムでは，社会との様々な関係を考慮しながら社会情報システム（social information system）という枠組みで開発・運用して行く必要があるということを簡明に示すものである．Society5.0 RA では，特に，セキュリティ・認証上の脅威が現在の情報システムにおいては大きな問題となっていることからか，他の全ての層（構造要素）に関係するという事を強調して独立した構造要素として表したアーキテクチャになっているが，本稿では，この箇所をセキュリティだけでなくセーフティも統合した安全対策のためのリスク管理とすることを提案する．また，本稿では，安全対策はリスク評価に対応して行われるのだが，リスク評価は価値観に基づいて行われるため，異なる価値観を認め合おうとする社会においては，特に，緊急時については合意形成と執行のしくみが必要であるということについて論ずる．

キーワード：Society5.0，安全対策，リスク管理，多元的価値観

Risk Management in Society5.0 Reference Architecture

HIROMI KOBAYASHI[†1]　YUKO MORISHITA[†2]

Abstract: Society5.0 Reference Architecture (RA) is one of the future concepts or frameworks as social information systems. The relationships of social elements such as strategy/policy, rule, organization, etc. should be considered in the development and operation of the system. Security/authentication is shown as an element in this RA because threats caused by security/authentication are crucial in recent information systems. In addition, security/authentication is depicted as an element related to others.　However, for the purpose of the unified risk measures for safety and security, this paper proposes that an element of risk management is substituted for that of security/ authentication. Safety and security measures correspond to risk assessment. The risk assessment is based on assets. However, the assets varies with sense of values. Therefore, consensus decision-making process and organization, and enforcement agency are needed for recognizing diverse values, especially in emergency.

Keywords: Society5.0, safety measure, risk management, diverse values

1. はじめに

ここ数年，Society5.0[1-3]の概念や AI の利活用の原則[4]についての議論が行われており，Society5.0 の概念をモデル化して示した Society5.0 RA では，特に，セキュリティ・認証上の脅威が現在の情報システムにおいては大きな問題となっていることからか，セキュリティ・認証が独立した構造要素で他の全ての構造要素に関係することを強調して表したアーキテクチャとなっているが，本稿では，この箇所をセキュリティだけでなくセーフティも統合した安全対策のためのリスク管理として表すことを提案する．情報システムにとって，安全性についての考慮は必要不可欠のものである．安全[性]については，学術や技術の分野では，リスクが偶発的なものかそれとも意図的なものかによってセーフティ（safety）とセキュリティ（security）に分けられるのだが，社会一般では[a]，どちらにせよリスクが許容範囲内であり安全が確保されていることが関心事であり，また，情報システムの連接化が進んでいる現在，セキュリティ上のトラブルによりセーフティ上の危害が発生する可能性も高まっているため，リスク管理の安全対策という観点から，両者を統合化して扱った方が良いのではないかと思われる．安全対策はリスク評価に対応して行われるのだが，リスク評価は根本的には価値判断に基づいて行われる．また，情報システムが社会の隅々まで浸透し，様々なサービスの連携が進むと，社会との関係を考慮することが益々必要になる．しかも，評価においては，生活の質が重視されるようになり，多次元的に複数の価値観を認め合おうとする現代において，どのような評価尺度を用いるかが課題となる．また，安全性の評価基準は，平常時と緊急時で異なる場合があり，特に，緊急時には，安全対策としてアクションを起こすには，速やかな合意形成のための意思決定とその執行のためのしくみや，裏付けとなる法律・規則が必要となる．これらは部分的にせよ AI で判断させる場合にも，前提となる事項であるため，社会との繋がりを考えながら検討する必要性がある．

以下，本稿では，２．で安全性の概念におけるリスクについて述べ，３．で Society5.0 と AI 利活用の原則における

1 東海大学
†1 Tokai University
2 奥羽大学
†2 Ohu university

a) 英語圏でも一般社会では safety と security をそれ程厳密に区別はせずに，一般的に安全な状態が確保されている状況を safe (the feeling of being safe)という表現を用いたりするようである[5]．なお，本稿は，言語学的な論争や用語の統一化を意図するものでは無い．

リスク管理の在り方について示す．４．ではリスク評価での価値観について述べ，５．で，価値観に関連して，しばしば人の行動選択の例として取り上げられるトロッコ問題での例について示し，６．で今後の課題である安全対策での価値観と合意について論ずる．

2. 安全性の概念におけるリスク

安全[性]という用語（即ち，概念）は，学術や技術の分野では，リスクが偶発的なものかそれとも意図的なものかによって，セーフティ（safety）とセキュリティ（security）に分けられ，それぞれ扱うコミュニティも異なる．学術分野では，日本語の名称を，前者を「安全[性]」，後者を「セキュリティ」と呼び区別しようとする試みもあるが，後者でも文脈により安全[性]と言う言葉がしばしば使われている．なお，両者の統合化については，以前から提唱され，学術分野では国際会議として IEEE Conference on Dependable and Secure Computing などがあるのだが，融合というよりは，セッションごとに棲み分けが行われているように見受けられる． また，制御システム技術のような分野では，両者を包括する規格 IEC TR 63069[6]で機能安全と制御セキュリティのためのフレームワークの策定が行われている[7].

そもそも，社会一般では，セーフティにせよセキュリティにせよ，リスクが許容範囲内であり（または，理念的或いは心情的であるかもしれないが「リスクがなく」），安全が確保されていることが関心事であるため，リスク管理の安全対策という観点から，両者を統合化して扱った方が良いのではないかと考えられる．最近では，情報システムの安全性というと，日用品化した PC やスマホの不正アクセスや情報漏洩などのセキュリティに関することが注目され，セーフティに関することは銀行や航空会社のシステム等で障害が起きた時にしか注意を払われない傾向にあるのだが，システムが様々なものと結ばれ大規模複雑化した現在，セキュリティ上のトラブルによりセーフティ上の危害が発生する可能性も高まってきており[8]，また，故意と偶発的の境界が曖昧になってきていることから，安全対策を効果的に行うために統合して扱うのが望ましいと考えられる．
ここでは，以下に，両者についての技術的な国際規格でのリスクとの関係や価値観について示すことにする．なお，混乱防止のため，以下では前者も後者もカタカナでセーフティとセキュリティと記述することにする．

（１）セーフティ

セーフティの定義としては，国際規格で関連技術の基になっているガイドライン ISO/IEC Guide51:2014[9]では，「freedom from risk which is not tolerable.」となっている．これに対応する国内規格 JIS Z 8051:2015[10]では，日本語で安全という用語を使い，その定義は「許容不可能なリスクがないこと」となっており，許容不可能という制約が付けられているのは，リスクゼロの絶対安全は現実には不可能であるという考え方が根本にある．（もっとも，いかなるリスクも許されないという理想論的考え方でも，言葉としてはこの定義には収まるのだが．）リスクは「危害の発生確率及びその危害の度合いの組合せ．」と定義され，許容可能なリスクは社会の価値観等によって決定されるという事も付記されている．信頼性の用語の国際規格 IEC60050(192):2015[11]（国内規格 : JIS Z 8115:2019 [12]）でもこのガイドラインとの整合性が図られている．

（２）セキュリティ

セキュリティについては，運用環境のセキュリティを扱う ISO/IEC27000:2018 [13]（国内規格 : JIS Q 27000:2019[14]）では，情報セキュリティ（information security） とは，「情報の機密性，完全性及び可用性を維持すること．」とある．また，開発側のセキュリティの評価・認証では ISO/IEC 15408[15]，実務上はこれを基に改定をしばしば行っている Common Criteria（CC）: 2017 があり，この翻訳版の情報技術セキュリティ評価のためのコモンクライテリア[16]においては，「セキュリティは，資産（asset）[b] の保護に関係する．資産とは，何者かによって価値が認められるエンティティである．」とされ，また，価値は非常に主観的であるため，ほとんどすべてのものが資産になりうると記載されている．

（３）その他の概念

安全性に関するその他の概念として，最近では，安全・安心という用語が盛んに用いられているが，安心というのは，セーフティやセキュリティの技術規格などでは特に定められた定義はなく，英語では trust という用語（の一つの使い方）が近いのかもしれないが[c]，安全であることが信頼できる（大丈夫と感じられる）という心理的意味合いで用いられるように思われる．文献[18-19]では，[人や組織或いは社会的]信頼は①価値観の一致，②能力，③動機付け（「動機付け」よりも「人間性」の方が適切かもしれない：本稿著者注）の組み合わせからなるとされているのだが，この意味での信頼（trust）という語を用いると，安心は安全と信頼の組み合わせからなると表現できる．

また，レジリエンス（resilience）[20]という言葉も安全に関連する言葉として最近良く使われているが，これは，危害が発生しても速やかに元の状態に復旧できる，あるいは上手く適応できる，という意味でセーフティとセキュリティの両方に対して用いられている．

b) セーフティの以前の規格 IEC60050(191):1990（国内規格：JIS Z 8115:2000）では，safety（安全）の定義で「asset（資（機）材）の損傷の危険性が・・・」と書かれていたのだが，肥大化する規格本文の短縮化の方針のためか，最近の規格では asset という語は省かれている．

c) セーフティの分野でも，最近では，トラスト(trust)という用語を，社会から信頼されるという意味で，使うようになってきている[17].

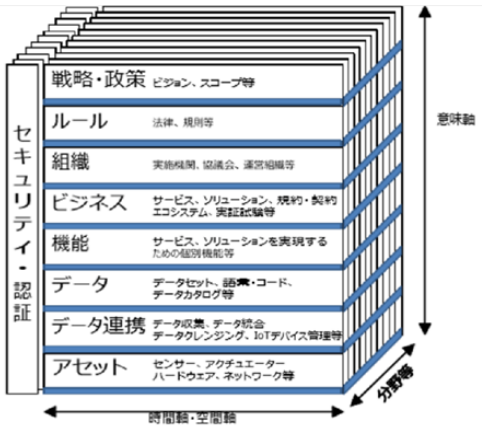
3. Society5.0 と AI 利活用の原則におけるリスク管理

情報システムというと，以前は，コンピュータがネットワークで繋がれただけの形態として扱われる事も多く，狭義で情報システムというと業務系システムとかビジネスシステムとも呼ばれ，データ変換型システムとしてモデル化されていたが，最近では，ロボット等も含めアクチュエータやセンサまで含んだ，従来はイベント反応型システムでモデル化された工業系或いは組込型システムとのネットワークによる融合が進んでおり，物理的な損傷に対するセーフティについても考慮する必要性が高まってきている．

また，情報システムがネットワークにより様々なシステムと繋がるようになり社会に浸透するにつれ，社会との関係を考慮しながら開発・運用を行う必要性が強く認識されるようになり，最近では社会情報システム（social information system）という言葉が良く使われるようになってきた．現在では，情報システムは，ネットワークでセンサやアクチュエータが接続され，プラットフォームによるサービスの連携が行われるようになり，大規模複雑化が進んでいる．Society5.0 や AI 利活用の原則は，ビッグデータを活用した AI の社会における役割を考慮した概念である．そこでは安全性については以下のように扱われている．

（1）Society5.0 リファレンスアーキテクチャ（RA）

Society5.0 は，2016 年に内閣府から第５次科学技術計画で提唱された概念で，AI とビッグデータを活用したサイバー空間とフィジカル空間の融合による人間中心の社会を目指した次世代の社会像で，全体像として図１のようなリファレンスアーキテクチャ（RA）が示されている[1-2].
図１ではセキュリティについては，セキュリティ・認証として描かれているが，これは近年，情報システムについては不正アクセスやサイバー攻撃が大きな脅威となっていることから，機能の一部としてではなく独立した構造要素



(内閣府資料より)

図 1 Society5.0 リファレンスアーキテクチャ(RA)
Figure 1 Society5.0 Reference Architecture (RA).

リスク管理 安全対策	戦略・政策	ビジョン，スコープ等
	ルール	法律，規則等
	組織	実施機関，協議会，運営組織等
	ビジネス	サービス，ソリューション，規約・契約，エコシステム，実証試験等
	機能	サービス，ソリューションを実現するための個別機能等
	データ	データセット，語彙・コード，データカタログ等
	データ連携	データ収集，データ統合，データクレンジング，IoTデバイス管理等
	アセット	センサー，アクチュエーター，ハードウェア，ネットワーク等

図２　Society5.0 RA の修正案
Figure 2 Amendment of Society5.0 RA

として扱われるようになり，更に，アセット[d] から戦略・政策までの全体について考慮する必要があるという事を強調するために，全体の横に描かれるようになったようである[3]．一方，セーフティについては，独立した構造要素とはされず，機能の一部，または各項目の性質の一つと扱われている．しかし，セーフティもセキュリティと併せてリスク管理としてセキュリティ・認証の場所に位置付けた方が良いように思われる．リスク管理で，安全対策については，事前の予防から事後の復旧に至るまで，技術的な対策であっても，先ずは，戦略・政策が必要であり，安全対策を行うための意思決定（合意形成）や執行のしくみ（組織とプロセス），法治国家である以上その根拠となるルール（法令や規格）や他の層についても検討する必要があることから，図１を図２のように，セキュリティ・認証の個所はリスク管理とした方が，適切であると思われる．

（2）AI 利活用の原則

AI 利活用の原則は，AI 技術の急速な進展とそれに伴う国際的な動向を見据え，2016 年に総務省が開催した AI ネットワーク社会推進会議により，人間を中心とした AI 社会原則や開発ガイドライン等と共に策定された．なお，この原則は，Society 5.0 実現に必要な社会変革についても踏まえた上で作成されている．文献[4]の AI 利活用ガイドラインの 10 項目からなる AI 利活用原則においては，セーフティとセキュリティについては，学術や技術規格の分類に従って，各々安全（セーフティ）の原則とセキュリティの原則として，以下のように記されている．

④安全の原則 : 利用者は，AI システム又は AI サービスの利活用により，アクチュエー タ等を通じて，利用者及び第三者の生命・身体・財産に危害を及ぼすことがないよう配慮する．

⑤セキュリティの原則 : 利用者及びデータ提供者は，AI システム又は AI サービスのセキュリティに留意する．

d) 図 1 では，アセットというのはセンサ，アクチュエータ，ハードウェア，ネットワーク等の「システムを構成する資産」という意味で，「データ資産（情報資産）」についてはデータに含まれると考えられているようである．

これらの原則からも，IoT（Internet of Things）時代の実際的な安全対策を考える上では，セキュリティで扱う不正アクセスやサイバー攻撃により，コネクテッド・カー（connected car）の運用のように，情報資産だけでなく，ネットワークに接続されたアクチュエータにより生物や物体に危害を及ぼすこともあり，また，安全対策として，データの分散化やバックアップ／リストアのようにセキュリティとセーフティの共通技術もあることから，セキュリティとセーフティは，リスク管理として，統合して扱った方が良いように思われる．更に，将来的な AI の利活用まで考えると，現在 AI で主流となっているディープラーニング[21-22]については，ニューラルネットの構造やパラメータの調整において，セーフティ的要因による故障からセキュリティ上の不正アクセスやサイバー攻撃が起きる可能性が無いとは言えないため，リスク管理として統合的に扱う必要性が益々高まってきているように思われる．

4. リスク評価と価値観

企業等での情報システムの開発・運用においては，事前にリスク評価が求められるようになってきている．安全対策として，事前の予防から事後の復旧対策まで考慮したシステムの方式設計を行う場合，運用も含めて，その根拠として評価，しかもできるだけ定量的な評価を行い，費用対効果の形でいくつかの案が比較評価されることが多い．ところが，リスク評価は根本的には価値判断に基づいて行われるという事が技術の世界でも認識されるようになってきた[9-16,23-25]．セーフティやセキュリティの技術規格においても，２．でも述べたように最近では，以下のようにリスクと価値観について記述されている．

セーフティの JIS Z 8051:2015[10]では，リスクは「危害の発生確率及びその危害の度合いの組合せ.」と定義され，許容可能なリスクは社会の価値観等によって決定されるという事が付記されている．

セキュリティの JIS Q27000:2019[14]では，リスクは「目的に対する不確かさの影響.」とし，ここでは一般的な使い方である危険性の他，投資の分野などでの利益と損失の不確実性という意味も含ませているが，リスクレベルについては「結果とその起こりやすさの組合せとして表現される，リスクの大きさ.」と記述している．セキュリティ上のリスクについては，「発生確率」とするのは適切ではないので，「起こりやすさ」としているのであろう．また，IT 製品の開発側のセキュリティ評価を扱う ISO/IEC 15408 の実務上用いるドキュメント Common Criteria（CC）: 2017 の翻訳版[16]においては，セキュリティは，資産（asset）の保護に関係する．資産とは，何者かによって価値が認められるエンティティである．とし，ただし，価値は非常に主観的であるため，ほとんどすべてのものが資産になりうるとも記載されている．

ここで問題になるのは，セーフティにせよセキュリティにせよ，何を価値あるものとして考えるかによって，安全対策を取る上での優先順位が違ってくるという事である．

5. トロッコ問題での例

リスク評価と価値観について，トイ（Toy）レベルの問題ではあるが，本質を捉えた解りやすい例としてしばしば取り上げられる「トロッコ問題」で考えてみることにする．トロッコ問題とは，トロリー問題[26]とも呼ばれる倫理上の問題で，以下のような問題である（図３(a)参照）．「線路上をブレーキが故障した路面電車が暴走していて，前方に５人がおり，このままでは５人全員をひき殺す事になる．しかし，その前で線路は二股に分かれており，分岐器により進路変更することができるのだが，分岐した先にはもう１人がいるので，進路変更するとその１人をひき殺すことになる．今，あなたは，分岐器の側にいて操作できるのだが，この場合線路を切り替えるのは適切であろうか？」

この場合には，５人を救って１人を犠牲にするという数（量）に基づく判断が適切とする人が多いようである．ところが，図３(b)に示すように，前方に１人の大人がいて，分岐先には１人の子供がいる場合ではどうであろうか？（この場合には，量ではなく質の問題になる.）その他，この問題では様々に条件を変えて倫理の分野では論争が行われているようである[27].

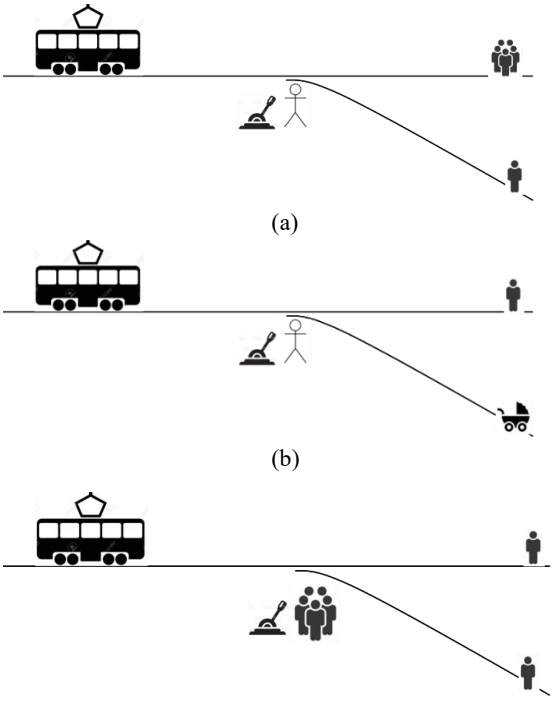


図３　トロッコ（トロリー）問題
Figure 3 Trolley Problem

この問題では，条件が極端に違わない場合には，どちらを選択するかは判断に時間的余裕があったときでも難しく，ましてや時間的余裕のない緊急時においてはいっそう困難になる．緊急時の場合には，とっさの判断（論理的判断よりも，感情的判断が優先されるかもしれない）により行われるのだろうが，例えば大人１人と子供１人のような場合，どちらを選択するのが適切と考えられるのであろうか？全ての人が納得するような判断はあるのだろうか？

この問題のバリエーションとして，自動車に適用する場合には，「狭い道をブレーキが故障し暴走する自動車が走っていて前方に人がいる場合，ハンドルを切って進路を変えようとすると歩道上の人をひいてしまいそうなときに，ハンドルを切るのは適切か？」などとなる．最近，自動運転について盛んに議論がされているが，ディープラーニングを用いた AI による自動化については，AI の訓練フェーズでどのような行動を取るのが望ましいのか予め AI を訓練しておく必要があり，そこではトロッコ問題のような場合，どのような判断をするのが適切であるのか示しておく必要がある．

更に，トロッコ問題を少し拡張して，図３(c)のように，切り替えスイッチの側に複数の人間がいて，切り替えるか否かを相談して行う必要があるような場合を考える．この場合，スイッチの切り替えが複数の人間の合意により決定されるとすると，意思決定のしくみが必要になり，また，スイッチの切り替えを実行する人間も決めておかなければならなくなる．つまり，意思決定のためには，誰がどのような方法で合意を行い，また，誰が実行するのかということを，特に，判断に時間的余裕のない，緊急の場合について予め定めておく必要がある．これは，複数の AI により判断を行わせようとする場合でも同様である．

6. 安全対策での価値観と合意

情報システムは，現在，インターネットに代表されるネットワークを介して，AI による高度な処理を行うコンピュータ，センサ，アクチュエータを持つロボット等の作動する物体が結びつき，アプリケーションについては複数の情報システムのサービスの連携という形態になってきている．開発・運用では，社会との繋がり考慮した社会情報システムという枠組みで捉えて行う必要があり，偶発的な事にせよ意図的な事にせよリスクを考慮し，安全性について十分に配慮したシステムを作り上げる必要がある．しかし，安全対策のためにはコストがかかるため，開発・運用にあたっては，適切か否かの判断のためにリスク評価が行われる．

リスクは，４．で述べたように，「危害の起こりやすさ（発生確率）及びその危害の度合いの組合せ．」なので，定量化して評価するためには，発生確率が算出でき，かつ，危害の度合いを金額で示すことが妥当そうなことのリスクについては，（リスク）＝（発生確率）×（危害による損失金額）とすることがまず考えられるが，果たして発生確率と損失金額の「積」として良いのであろうか．例えば，この式による飛行機事故と自動車事故のリスクの比較は，それなりの意味はあるのだろうが，原子力事故のように極めて発生確率が低いと考えられているが危害の度合いが甚大であるものも，この式でリスクを算出し，自動車事故などと比較して良いものかは，疑問である（今までにも，そのように考えた人も多いだろうが）．また，発生確率が極めて低い事象については過去の事例が少ないため，発生確率の設定の妥当性にも疑問が残る．全てを費用対効果に基づく金額による評価するのは指標としては解りやすいのだが，果たして金額だけで何もかも評価して良いものかも疑問である．近年，社会では，量だけではなく質による評価の必要性が言われている．しかし，質を問う定性的性質の評価であっても，実務的には説得性のある評価とするために最終的には量に変換を行う事が一般的に行われている．質から量の変換として良く用いられるのは，定性的性質をいくつかの評価項目に分解し，各評価項目を５段階などで評価するアンケートを行い，スコアの合計点で評価を行うという，要素還元的な手法であるが，これが果たして定性的項目の評価方法として，また，全体の評価方法として妥当なのかどうか疑問である．

安全対策を考える際のリスク評価においては，危害の度合いというのは根本的には価値判断に基づいて行われているのだが，価値基準は量的に算出が難しい場合がある．
価値基準は一人の人間でも複数持つ場合がある上に，時間と共に変化する場合がある．しかも，複数の人間の合意が必要な場合には，人により優先する価値が異なる場合があるため，社会との関係を考慮しながら社会情報システムという枠組みで捉えると，評価には考慮すべき要素が多くなる．現代は，文献[28-29]等で述べられている事でもあるが，多次元的に複数の価値観を認め合おうとする時代であり，複数の人の合意により評価を行う必要性のある場合には，合意形成のしくみ（組織とプロセス）を決めておく必要があると共に，合意事項を執行するしくみや，法治国家である以上裏付けとなる法令・規則が必要となる．しかも，評価基準は，平常時と緊急時で違う場合があり，特に，緊急時には，安全対策として速やかにアクションを起こすにためには，速やかな意思決定と執行のためのしくみが必要となる．これは，自動運転などで AI を利活用する場合にも，前もって解決しておかねばならない事柄と考えられる．

なお，日本では，このような事については，できるだけ何も決めないでおいて，状況に合わせて臨機応変に対応するという考え方が好まれるようであるが，臨機応変な対応においては，各人が自己責任により判断・行動する必要性が大きくなるため，各人が自己責任という事にどれだけ耐えられるかという問題がある．

価値判断の基となる価値観については，時代や社会によって変わってくる面があるのだろうが，自分や自分と関わりのあるコミュニティに優先順位を置きボトムアップ的に拡張して行く場合と，理想主義的な概念を優先しトップダウン的に考える場合があると考えられるが，本稿では．これ以上触れないことにする．

7. おわりに

Society5.0 RA は，今後の情報システムを開発・運用する際には，政策や組織，ルール等の社会との関係を考慮する必要のある事を簡明に示したモデルである．Society5.0 RA においては，セキュリティ・認証上の脅威が情報システムで現在大きな問題となっていることからか, セキュリティ・認証が独立した構造要素として記述されている．しかし，本稿では，この箇所について，リスク管理の安全対策という観点から，セキュリティとセーフティを統合化してリスク管理として記述し扱う事を提案した．理由は，モデルで記述が無い事項については二義的に扱われてしまう恐れがある事と，社会一般では，どちらにせよリスクが許容範囲内であり安全が確保されていることが関心事であるという事による．また，技術の領域において，最近ではセキュリティとセーフティの統合化の必要性が強く提唱されるようになり，制御システム等の分野では連携に関する枠組みの作成が試みられているという事もある．IoT 時代においては，セキュリティ上のトラブルによりセーフティ上の障害が発生する恐れがある事は良く言われているが，AI を利活用する上では逆もあり得ると考えられる．

次に，安全対策のためには，リスク評価が必要だが，リスク評価は価値観が基になっており，社会との多面的な関係を考え，異なる価値観を認め合おうとする現代においては，特に，緊急時の意思決定（合意形成）や執行のしくみ（組織とプロセス）が必要であることを論じた．これは AI の利活用に当たっても，前もって形成しておかなければならない事柄である．今後の課題としては，定性的性質を含むリスクの妥当な評価方法や緊急時の合意形成と執行のしくみの検討などが挙げられる．

参考文献

[1] Society5.0 内閣府 https://www8.cao.go.jp/cstp/society5_0/
[2] 日立東大ラボ. Society 5.0. 日本経済新聞社, 2018.
[3] Society5.0 の RA はどう検討されたのか
https://note.com/hiramoto/n/n2b74eb9383eb (2021-4-5 参照)
[4] AI ネットワーク社会推進会議報告書(2018)：別添２AI 利活用ガイドライン，総務省． https://www.soumu.go.jp/menu_news/s-news/01iicp01_02000081.html
[5] J. Morgan. Difference between Safety and Security. http://www.differencebetween.net /language /words-language/ difference-between-safety-and-security/ (2021-4-3 参照)
[6] IEC TR 63069:2019 Industrial-process measurement, control and automation - Framework for functional safety and security.
[7] 神余浩夫. 機能安全と制御セキュリティの標準化動向. 情報処理，2017, vol. 58, no.11, pp.966-971.
[8] 田口研治. IoT の進展に伴うセーフティとセキュリティのリスクと課題. 情報処理，2017, vol. 58, no.11, pp.960-965.
[9] ISO/IEC Guide51：2014 Safety aspects-guidelines for their inclusion in standards.
[10] JIS Z 8051:2015 安全側面-規格への導入指針.
[11] IEC60050(192):2015 Glossary of terms used in dependability.
[12] JIS Z 8115:2019 ディペンダビリティ（総合信頼性）用語.
[13] ISO/IEC27000:2018 Information technology -Security techniques - Information security management systems -Overview and vocabulary.
[14] JIS 27000:2019 情報技術−セキュリティ技術−情報セキュリティマネジメントシステム−用語.
[15] ISO/IEC 15408-1:2009 Information technology -Security techniques -Evaluation criteria for IT security - Part 1 Introduction and general model.
[16] Common Criteria for Information Technology Security Evaluation-Part1, ver.3.1 rev.5 :2017（邦訳：独立行政法人情報処理推進機構, 情報技術セキュリティ評価のためのコモンクライテリア パート 1: 概説と一般モデル）.
[17] 江川尚志. AI の倫理とトラストの標準化動向. 電子情報通信学会誌, 2021, vol. 104, no.1, pp.60-63.
[18] 中谷内一也. 安全。でも、安心できない・・・. ちくま新書, 2008.
[19]土田昭司編著. 安全とリスクの心理学. 培風館, 2018.
[20] E. Hollnagel, D.D. Woods, N. Leveson 編著，北村正晴監訳. レジリエンスエンジニアリング：概念と指針. 日科技連, 2012.
[21] テレンス・J・セイノフスキー著，銅谷賢治監訳. ディープラーニング革命. ニュートンプレス, 2019.
[22] フランソワ・ショレ著，巣籠悠輔監訳. Python と Keras によるディープラーニング.マイナビ, 2018.
[23]小林洋. 安全性についての課題―価値判断と AI の使用について. 東海大学紀要情報通信学部, 2017, vol.10, no.1, pp.114-116.
[24] 小林洋. システムの安全性とは何に対する安全性なのか?, ソフトウェアシンポジウム 2012. ソフトウェア技術者協会, 2012, pp.3.1-3.7.
[25] H. Kobayashi. Safety Concern in System Development. Proc. Sixth Joint Conference on Knowledge-Based Software Engineering, IOS Press, 2004, pp.311-318.
[26] P. Foot. The problem of Abortion and the Doctrine of the Double Effect. The Oxford Review, 1967, no.5, pp.5-15.
[27] 鈴木貴之編著. 実験哲学入門. 勁草書房, 2020.
[28] 村上陽一郎. 安全学. 青土社, 1998.
[29] 向殿正男. 入門テキスト安全学. 東洋経済, 2016.