資料5

ICTインテリジェント化影響評価検討会議　第1回会合

討議用資料4　ICTインテリジェント化のリスク

平成28年2月

総務省情報通信政策研究所

目次

1. はじめに

はじめに

- 本検討会議では、インテリジェント化が加速するICTの未来像に関する研究会「報告書２０１５」の提言を踏まえ、インパクトスタディとともにリスクスタディを行う。インテリジェントICTを使いこなすためには、発生し得る負の側面を可能な限り把握し、把握できたものから、研究・開発原則等への反映や、慎重なルール作りを通じて、対処の仕組みを構築していくことが求められる。

- 以上の背景及び問題意識を踏まえ、リスクスタディの検討対象となるICTインテリジェント化のリスクは、ICTインテリジェント化が社会に及ぼす影響（広義の影響）のうち、目指すべき社会像及び基本理念に即して抑制されるべき負の影響として定義することにしたい。

- 討議用資料４では、ICTインテリジェント化のリスクを評価し、管理するための検討の枠組みを示すとともに、ICTインテリジェント化のリスクの展望を示した上で、ICTインテリジェント化のリスクに関して今後注視し、又は検討すべき事項を整理する。

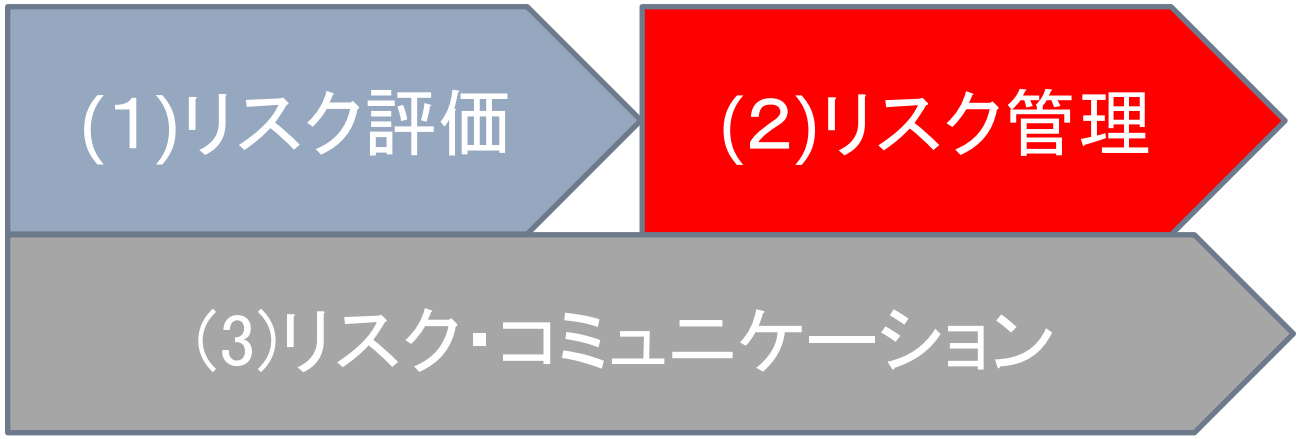
２．ICTインテリジェント化のリスクに関する検討の枠組み

リスク分析の枠組み

- リスク論においては一般に、リスクを予防し、低減するための枠組みとして
  (1)リスク評価
  (2)リスク管理
  (3)リスク・コミュニケーション
  からなる「リスク分析」が採用されてきた[1]。
- かかる枠組みは、ICTインテリジェント化のリスクを予防し、低減する上でも参照に値するものと考えられるのではないか。

リスクの所在を把握し、被害の発生時期、蓋然性、規模等を評価する。

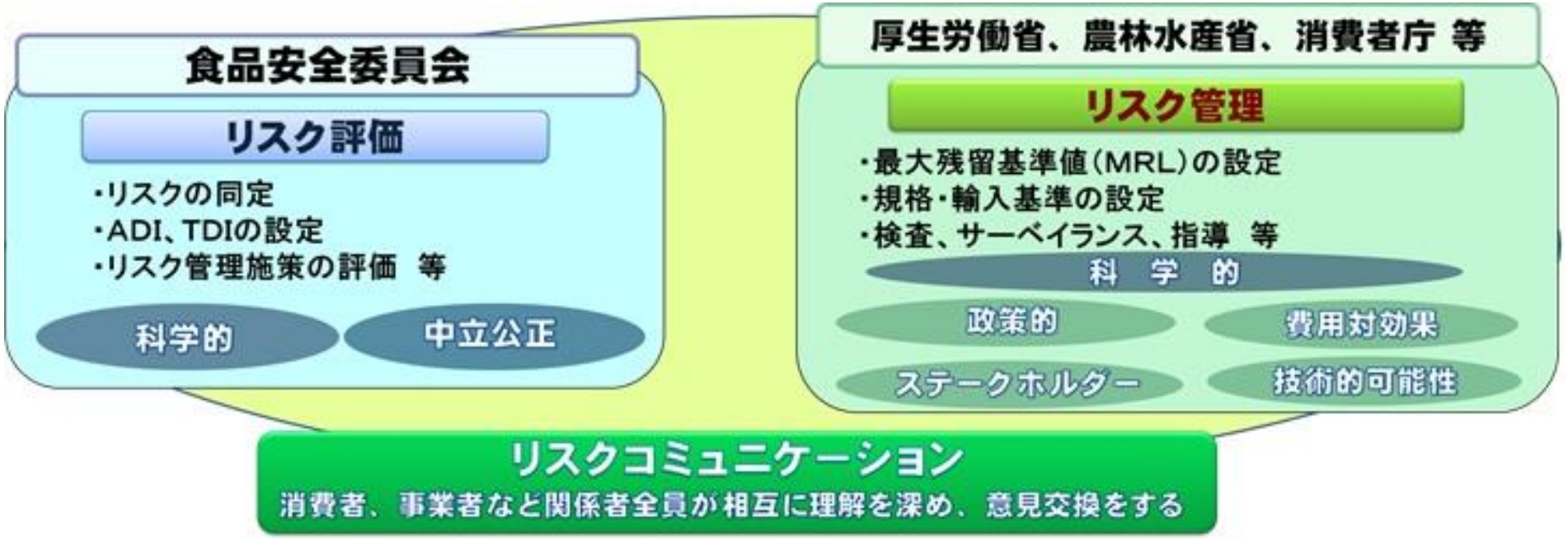
リスク評価の結果を踏まえ、リスクへの対処を決定・実施する。



リスク評価及びリスク管理のプロセスにおいて関係するステークホルダーと情報・意見を交換する。

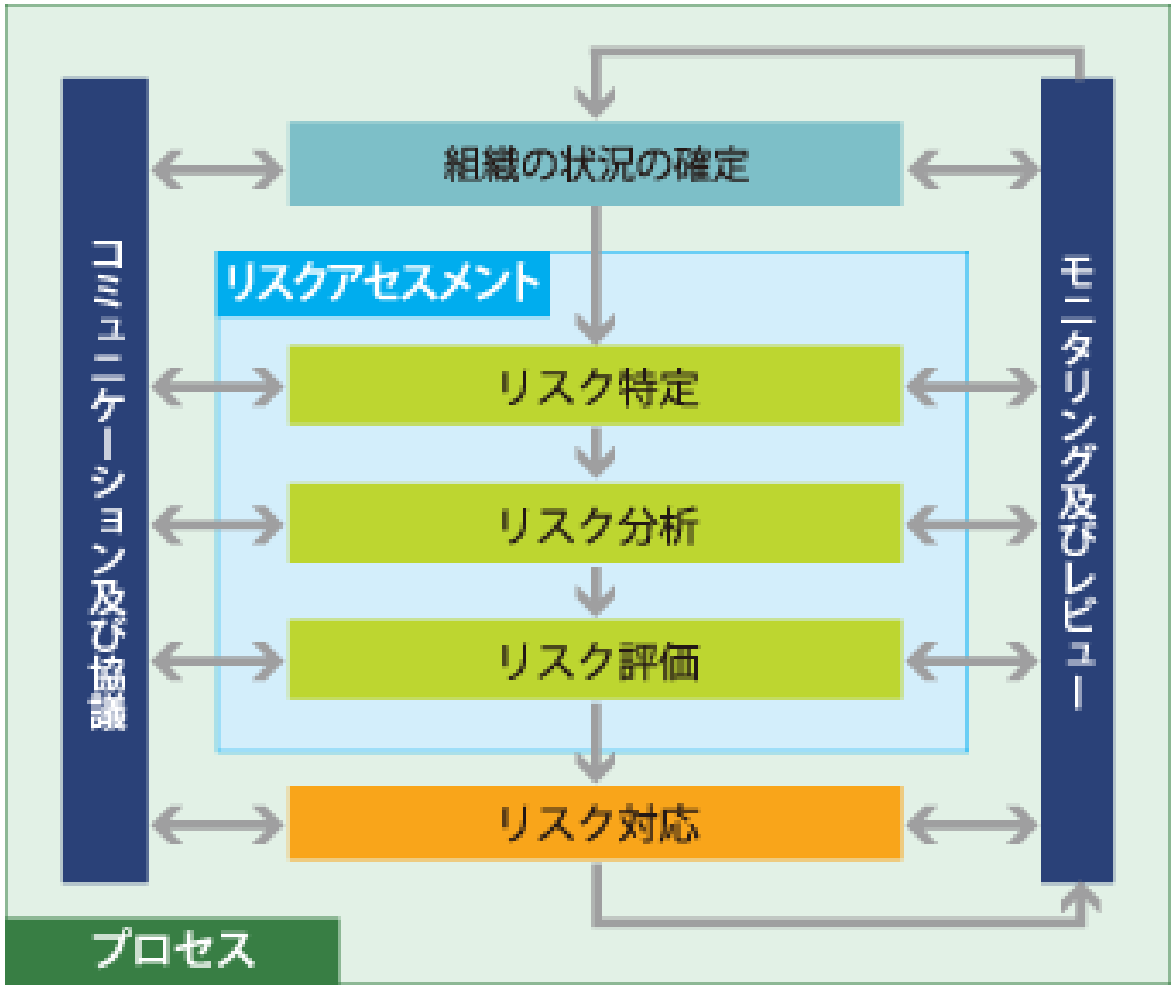
1. [城山 2007]等を参照。

【参考】 食品安全行政におけるリスク対処の枠組み



（出所）内閣府ホームページ
http://www.cao.go.jp/about/pmf/pmf_22_kai.html

【参考】 ISO 31000　リスクマネジメントのプロセス



（出所）一般財団法人日本品質保証機構　ISO NETWORK Vol.22
https://www.jqa.jp/service_list/management/iso_info/iso_network/vol22/news/kikaku_3.html

(1)リスク評価

- リスク評価:リスクの所在を把握し、被害の発生時期、蓋然性、規模等を評価する。

- リスク評価は将来の予測に依拠しているが、先端科学技術などに起因する現代的なリスクについては、新たなリスクの創出に経験の蓄積が追いつかず、経験による将来の予測に基づくリスクの評価が困難となりやすいと指摘されている[1]。

- ICTインテリジェント化のリスクについても、新たなリスクの創出に経験の蓄積が追いつかず、被害の蓋然性や規模等を予測することが困難となりやすいのではないか。

- 科学における不確実性の増大などに伴い、専門家の間でもリスクの評価が分かれることが少なくなくなっている。
  → どの科学的知識をいかなる基準で選択するのかが問われることになるのではないか[2]。

1. [山田 2013]1章等を参照。
2. [城山 2007]等を参照。

(2)リスク管理

- リスク管理:リスク評価の結果を踏まえ、リスクへの対処を決定・実施する。

- リスク管理においては、技術のリスクのみならず、社会的便益についても判断することが必要となる。

  →技術のリスクや社会的便益について判断する上では、技術者のみならず、社会における各種のステークホルダーによる参加のメカニズムが求められるのではないか[1]。

- 不確実なリスクへの対処法として、科学的証明が確実ではない段階の予防的措置を正当化する予防原則(Precautionary Principle)が提唱され、環境分野を中心に一定の場面で国際的に支持されるようになっている[2]。

- もっとも、予防原則については、あるリスクを予防することにより別のリスクが生ずるおそれもあることなどから、予防原則を徹底することは困難なのではないかとの批判もある[3]。

1. [城山 2007]等を参照。
2. [大塚 2007]等を参照。
3. [サンスティーン 2015]等を参照。

(3)リスク・コミュニケーション

- リスク・コミュニケーション：リスク評価及びリスク管理のプロセスにおいて関係するステークホルダーと情報・意見を交換する。

- リスク・コミュニケーションの目的[1]
  ① リスクとその対処法に関する教育・啓発
  ② リスクに関する訓練と行動変容の喚起
  ③ リスク評価・リスク管理機関等に対する信頼の醸成
  ④ リスクに関わる意思決定への利害関係者や公衆の参加と紛争解決

- リスク・コミュニケーションの推進方策[2]
  ① リスクコミュニケーションの基礎的素養の涵養
  ② 問題解決に向けたリスクコミュニケーションの場の創出
  ③ 時間軸でのプロセスデザインを通じた普段化と良好事例の共有・展開
  ④ 媒介機能を担う人材の育成等
  ⑤ リスクに関する科学技術リテラシー・社会リテラシーの向上

1. [科学技術・学術審議会 2014]等を参照。
2. 同上

３．ICTインテリジェント化のリスクに関する展望

ICTインテリジェント化のリスクの所在（1／2）

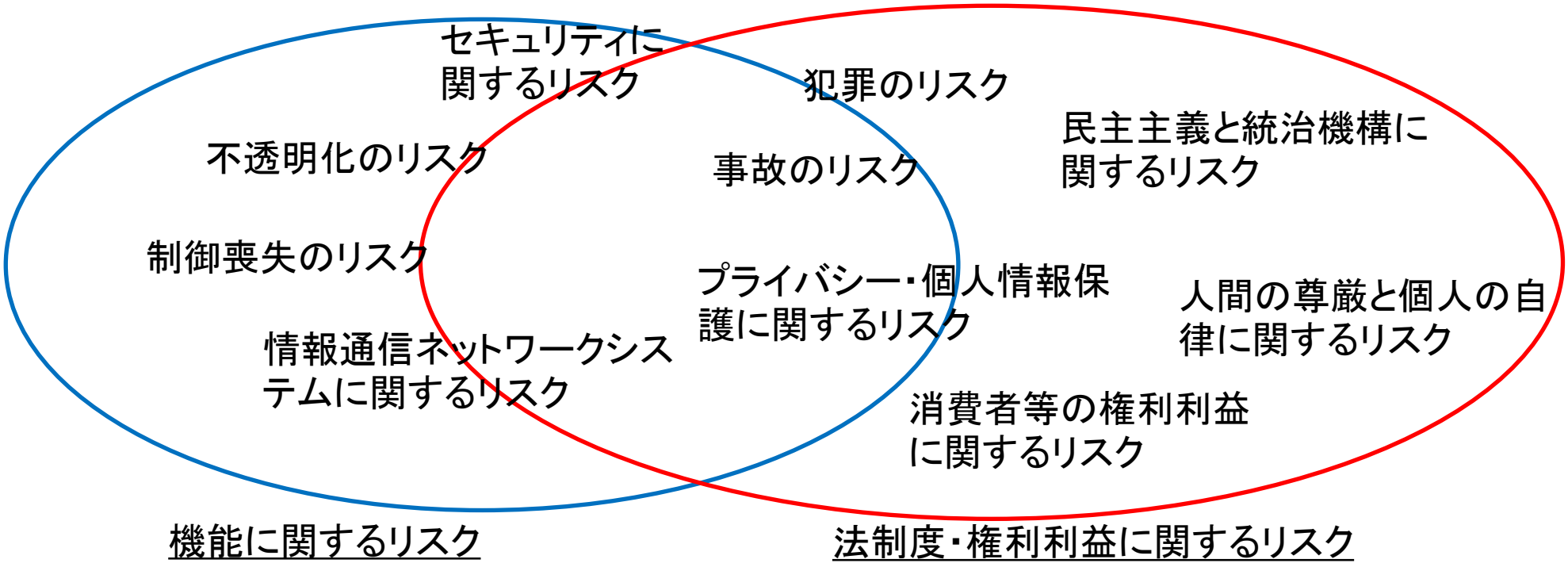
リスクの種類	概要
セキュリティに関するリスク	・インテリジェントICTに対するハッキングやサイバー攻撃等［板倉構成員、林（秀）構成員等］ ・インテリジェントICTに対する攻撃が秘かに行われ、被害に気づかないリスク［板倉構成員等］
情報通信ネットワークシステムに関するリスク	・情報通信ネットワーク上に多種多様なインテリジェントICTが混在することにより、インテリジェントICTが正常に動作せず、意図しない事象が生ずるリスク［インテリ研 2015、中西構成員等］ ・情報通信ネットワークの不具合によりインテリジェントICTが正常に動作せず、意図しない事象が生ずるリスク ・クラウド等におけるデータ漏洩・消失やシステム障害のリスク［林（秀）構成員等］
不透明化のリスク	・インテリジェントICTのアルゴリズム等が不透明化し、人間によるインテリジェントICTの適正な制御が不可能ないし困難になるリスク［中西構成員等］
制御喪失のリスク	・インテリジェントICTが暴走し、人間による制御が及ばなくなるリスク［インテリ研 2015等］ ・超知能（superintelligence）の誕生やシンギュラリティの到達により人間が人工知能を制御できなくなるリスク［AI100、Bostrom 2014等］

ICTインテリジェント化のリスクの所在（2／2）

リスクの種類	概要
事故のリスク	・自動運転車やロボット等の「自律的」判断に基づく動作による事故のリスク[平野構成員、赤坂構成員等]
犯罪のリスク	・インテリジェントICTを悪用したマルウェアによる犯罪のリスク［AI100等］ ・自律型兵器がテロ等犯罪に悪用されるリスク
消費者等の権利利益に関するリスク	・インテリジェントICTが適正に利活用されないことにより消費者、青少年、高齢者等の権利利益が損なわれるリスク[湯浅構成員等]
プライバシー・個人情報に関するリスク	・インテリジェントICTによる個人情報の収集・利活用が不透明化することにより、個人情報のコントロールが困難になるリスク［石井構成員等］ ・人工知能が人々の信念、健康、将来の行動等を推論することによりプライバシーが侵害されるリスク［AI100等］
人間の尊厳と個人の自律に関するリスク	・インテリジェントICTが人間の意思決定過程を見えない形で操作することにより個人の自律が侵害されるリスク
民主主義と統治機構に関するリスク	・インテリジェントICTが投票等国民の行動に影響を及ぼすことによる民主主義へのリスク［AI100等］。 ・インテリジェントICTを国家の統治に利活用する場合における意思決定過程の不透明化や責任の所在の曖昧化のリスク

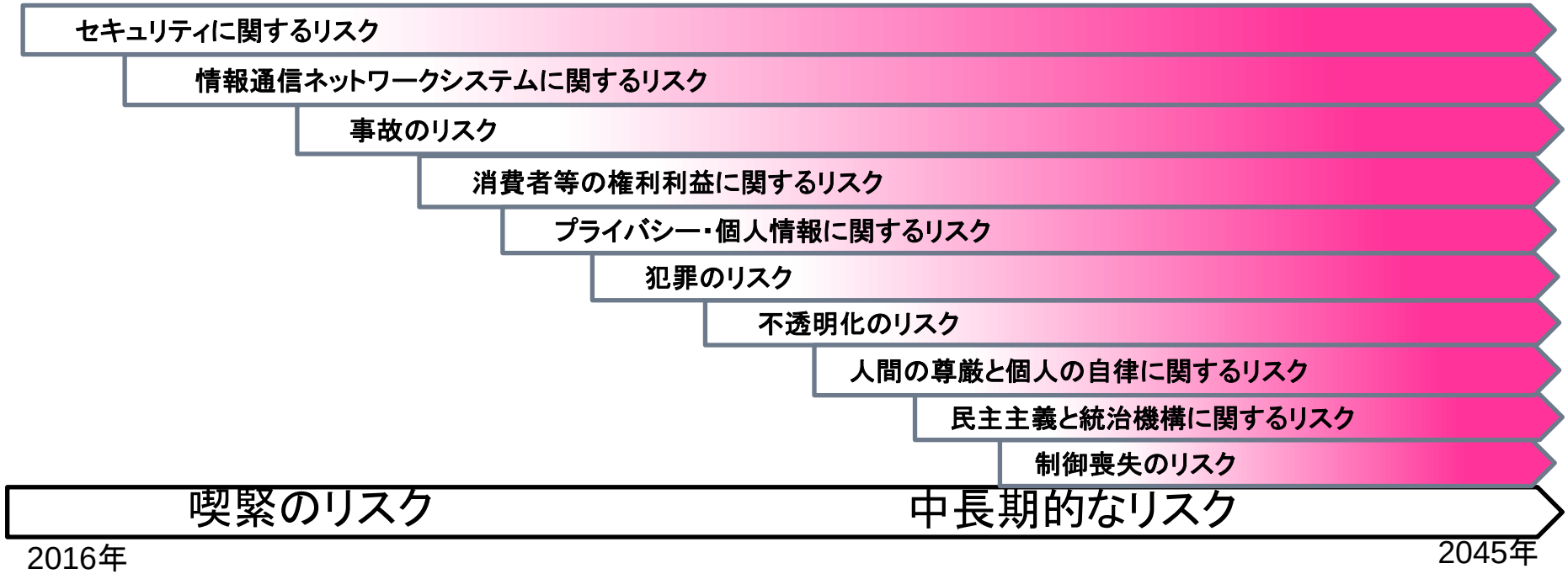
(1) 機能に関するリスク／法制度・権利利益に関するリスク

- ICTインテリジェント化のリスクの評価及び管理の在り方に関し論ずる上では、インテリジェントICTに期待される機能が適正に発揮されない「機能に関するリスク」と、インテリジェントICTにより権利利益等法益が侵害される「法制度・権利利益に関するリスク」とに区別して検討することが有用であると考えられる。
- もっとも、事故のリスクのように、両者の側面を併せ持つリスクも少なくないことから、技術と法制度の両面からリスクの評価及び管理の在り方を検討することが求められる。



(2) 発生時期

- ICTインテリジェント化のリスクには、近い将来に被害の発生が見込まれるリスクから、中長期的に被害の発生が見込まれるリスクまで、被害の発生が見込まれる時期に少なからず幅があることから、研究開発及び利活用の進展等に応じて、安心・安全の確保とイノベーティブな環境の維持とのバランスにも配慮しつつ、適時適切な対応を行っていくことが求められるのではないか[1]。
- このような問題意識を踏まえ、ICTインテリジェント化のリスクを、被害の発生が見込まれる時期に即して整理することにしたい（下記は検討に向けたイメージ図）。



1. 平野構成員提供資料、江間構成員提供資料等を参照。このほか、工藤郁子情報通信政策研究所特別フェローの示唆も参考。

(3) 予測可能性・確実性

- リスクは一般に、被害の蓋然性や規模等が予測可能なリスクと予測不能ないし困難なリスク（不確実なリスク）とに区別され、各々の性質に即して評価及び管理の在り方が検討されてきた[1]。

- ICTインテリジェント化のリスクについても、被害の蓋然性や規模等が予測可能なリスク（例：自動運転車やドローン等による事故のリスク）と予測不能ないし困難なリスク（例：人類がインテリジェントICTに対する制御を喪失するリスク）とに分類することができよう。

- ICTインテリジェント化のリスクに対処する上では、被害の蓋然性や規模等が予測可能なリスクを予防することはもとより、予測不能ないし困難なリスクについても、不確実性に留意しつつ、可能な限り予防することが求められるのではないか[2]。

1. 例えば、[山田 2013] 1章等を参照。
2. 平野構成員提供資料、久木田構成員提供資料等を参照。

参照文献

- [インテリ研 2015] インテリジェント化が加速するICTの未来像に関する研究会「報告書2015」(総務省情報通信政策研究所、平成27年)
- [大塚 2007] 大塚直「環境法における予防原則」城山英明・西川洋一編『法の再構築Ⅲ　科学技術の発展と法』(東京大学出版会、2007年)
- [科学技術・学術審議会 2014]科学技術・学術審議会 研究計画・評価分科会 安全・安心科学技術及び社会連携委員会「リスクコミュニケーションの推進方策」(平成26年)
- [サンスティーン2015]キャス・サンスティーン(角松生史・内野美穂監訳、神戸大学ELSプログラム訳)『恐怖の法則－予防原則を超えて』(勁草書房、2015年)
- [城山 2007] 城山英明「リスク評価・管理と法システム」城山英明・西川洋一編『法の再構築Ⅲ　科学技術の発展と法』(東京大学出版会、2007年)
- [山田 2013] 山田洋『リスクと協働の行政法』(信山社、2013年)
- [AI100] Eric Horvitz, *One Hundred Year Study on Artificial Intelligence: Reflections and Framing* (2014)
- [Bostrom 2014]NICK BOSTROM, SUPER INTELLIGENCE: PATHS, DANGERS, STRATEGIES (2014)

4.今後注視し、又は検討すべき事項

(1) 機能に関するリスク
ア　セキュリティに関するリスク
イ　情報通信ネットワークシステムに関するリスク
ウ　不透明化のリスク
エ　制御喪失のリスク

(2) 法制度・権利利益に関するリスク
ア　事故のリスク
イ　犯罪のリスク
ウ　消費者等の権利利益に関するリスク
エ　プライバシー・個人情報に関するリスク
オ　人間の尊厳と個人の自律に関するリスク
カ　民主主義と統治機構に関するリスク

資料5（別紙）

討議用資料４　別紙

４．今後注視し、又は検討すべき事項

(1) 機能に関するリスク

ア　セキュリティに関するリスク

・ インテリジェントＩＣＴのセキュリティに関して想定されるリスク如何（例:インテリジェントＩＣＴに対するハッキング、マルチテナント環境に対するサイバー攻撃等）[1]。
・ インテリジェントＩＣＴに対する攻撃が秘かに行われ、被害に気づかないリスクへの対処の在り方如何[2]。
・ 情報セキュリティの概念及び要素（①機密性、②完全性、③可用性）は、情報システム及び情報通信ネットワークによる情報の蓄積、処理、伝送等を念頭に生成・発展してきたものである[3]。このことに鑑みると、ＩＣＴインテリジェント化による情報通信ネットワークを通じたヒト・モノ・コト相互間の協調の進展を踏まえ見直しが必要となるのではないか。
・ インテリジェントＩＣＴの研究開発及び利活用の各段階におけるセキュリティ上のリスクへの対処の在り方如何[4]。

イ　情報通信ネットワークシステムに関するリスク

・ 情報通信ネットワーク上に多種多様なインテリジェントＩＣＴが混在することにより、インテリジェントＩＣＴが正常に動作せず、意図しない事象が生ずるリスクへの対処の在り方如何[5]。
・ 情報通信ネットワークの不具合によりインテリジェントＩＣＴが正常に動作せず、意図しない事象が生ずるリスクへの対処の在り方如何。
・ クラウド等におけるデータ漏洩・消失やシステム障害のリスクへの対処の在り方如何[6]。

ウ　不透明化のリスク

・ インテリジェントＩＣＴのアルゴリズム等が不透明化し、人間によるイ

---

1 板倉構成員提供資料、林（秀）構成員提供資料等を参照。
2 板倉構成員提供資料等を参照。
3 情報セキュリティの概念及び要素については、岡村久道『情報セキュリティの法律［改訂版］』1章（商事法務、平成23年）等を参照。
4 板倉構成員提供資料等を参照。
5 中西構成員提供資料等を参照。
6 林（秀）構成員提供資料等を参照。

ンテリジェントＩＣＴの適正な制御が不可能ないし困難になるおそれへの対処の在り方如何。

・ インテリジェントＩＣＴのアルゴリズム等のブラックボックス化を回避し、透明性を確保するために、研究開発及び利活用の各段階における取組の在り方如何[7]。

エ　制御喪失のリスク

・ 将来的に超知能（superintelligence）の誕生やシンギュラリティにより人間がインテリジェントＩＣＴを制御できなくなる可能性が指摘されている。
インテリジェントＩＣＴに対する制御喪失のリスクを防止するための研究開発及び利活用の各段階における取組の在り方如何[8]。

(2) 法制度・権利利益に関するリスク

ア　事故のリスク

・ インテリジェントＩＣＴによる事故を防止するための研究開発及び利活用の各段階における取組の在り方如何[9]。
・ インテリジェントＩＣＴが事故を起こした場合における責任（民事責任、刑事責任等）の帰属主体（製造者、販売者、管理者、利用者等）如何[10]。
  ➢ 製造物責任法等法制度の見直しが必要となるのではないか[11]。
・ インテリジェントＩＣＴによる事故のリスクに対処するための保険の在り方如何[12]。

イ　犯罪のリスク

・ インテリジェントＩＣＴを悪用したマルウェアへの対処の在り方如何。
  ➢ 特にインテリジェントＩＣＴを悪用したマルウェアにより秘かに犯罪が実行されるリスクへの対処の在り方如何[13]。
・ 自律型兵器がテロ等犯罪に悪用されるリスクへの対処の在り方如何。

---

7 中西構成員提供資料、松尾豊ほか「人工知能学会倫理委員会の取組み」人工知能 30 巻 3 号（平成 27 年）参照。

8 NICK BOSTROM, SUPER INTELLIGENCE: PATHS, DANGERS, STRATEGIES (2014); Eric Horvitz, *One Hundred Year Study on Artificial Intelligence: Reflections and Framing* (2014)（以下、「AI100」という。）.

9 平野構成員提供資料等を参照。

10 平野構成員提供資料、湯浅構成員提供資料、深町構成員提供資料等を参照。

11 平野構成員提供資料等を参照。

12 同上。

13 AI100 等を参照。

ウ　消費者等の権利利益に関するリスク

・ 消費者、青少年、高齢者等がインテリジェントＩＣＴを利活用する場面においては、製品の品質保証等において特別な配慮が必要となるのではないか[14]。
・ 継続的なアップデートを必要とするインテリジェントＩＣＴを実装した製品に関する消費者保護の在り方如何[15]。
・ 海外事業者と消費者との直接契約をめぐる準拠法や裁判管轄等に関する問題への対処の在り方如何[16]。

エ　プライバシー・個人情報に関するリスク[17]

・ インテリジェントＩＣＴの発展は、個人のプライバシーにいかなるリスクをもたらすか（例：人工知能の推論能力を用いた高度なプロファイリングによるプライバシーの暴露）。
・ インテリジェントＩＣＴの発展は、プライバシー権の理解の見直し（ひとりで放っておいてもらう権利の再評価、自己情報コントロール権の再検討、財産権的理解の可能性など）を迫る可能性があるのではないか。
・ インテリジェントＩＣＴの発展は、個人情報保護法制の見直しを迫る可能性があるのではないか。
・ インテリジェントＩＣＴによる個人情報の収集やプロファイリングが不透明化した場合、個人が自己の情報を実効的にコントロールすることは可能か。
・ インテリジェントＩＣＴに関するプライバシー・個人情報保護を実現するための研究開発及び利活用の各段階における取組の在り方如何。

オ　人間の尊厳と個人の自律に関するリスク

・ インテリジェントＩＣＴと人間の連携が進展する中、意識や心を持つインテリジェントＩＣＴを作ってよいか、人間の脳機能を含む身体機能をインテリジェントＩＣＴにどこまで代替させてよいか、インテリジェントＩＣＴは人間の意識に対しどこまで働きかけてよいか等、人間の本質に係る課題[18]について、人間の尊厳や個人の自律等の価値を踏まえ検討

---

14 湯浅構成員提供資料等を参照。
15 同上。
16 同上。
17 石井構成員提供資料等を参照。
18 インテリジェント化が加速するＩＣＴの未来像に関する研究会「報告書 2015」39 頁（平成 27 年）等を参照。

していくことが必要ではないか。
・ インテリジェントＩＣＴが人間の意思決定過程を見えない形で操作することにより個人の自律が侵害されるリスクへの対処の在り方如何。
・ 人間の脳と機械の連携などにより内心、表現及び行為の境界が相対化する場合における思想良心の自由等個人の自由への影響の在り方如何。

カ　民主主義と統治機構に関するリスク

・ インテリジェントＩＣＴが投票等国民の行動に影響を及ぼすことによる民主主義への影響如何。
  - 特にインテリジェントＩＣＴが投票等国民の行動や信念に秘かに影響を及ぼす場合における民主主義への影響如何[19]。
・ インテリジェントＩＣＴを国家の統治に利活用する場合における意思決定過程の不透明化や責任の所在の曖昧化のリスクへの対処の在り方如何。

[19] AI100 等を参照。