

RCModelを用いたAIリスクマネジメントのための人材育成事例

伊藤 宏比古 伊藤 千央

要 旨

近年、AIの利用が広がるにつれ、AIの適切な利用のためのAIガバナンスに注目が集まっています。そしてAIガバナンスに関する規制のなかには、人間による適切なリスクマネジメントを求めるものもあります。そのため、今後、AIサービスのリスクマネジメントを行う人材の育成は重要な課題になると考えられます。NECでは東京大学と2021年から、東京大学が開発するRCModelというツールを用いて、AIのリスクマネジメントを担う人材の育成方法について共同研究を進めてきました。本稿では、その共同研究のなかで実施した人材育成プログラムの概要やその結果・成果、そして今後の展望について紹介します。



AIガバナンス／AI倫理／人材育成／RCModel／リスクマネジメント

1. はじめに

近年、AIガバナンスに注目が集まっています。この10年ほどでAIの利用が社会のなかで爆発的に広がり、それに伴ってAIの不適切な利用や気付かないバイアスによる差別的なAIの決定が問題視されるようになってきました。このようなAIの利用によってもたらされるリスクに対応するため、ここ数年、AIの開発や利用に関するポリシーやガイドラインの整備が、政府から企業とさまざまなレベルで進んできました。現在、それらポリシー・ガイドラインの整備に加え、それらを確実に実行していくためのガバナンスの在り方に注目が集まっています。

特に、AIガバナンスにおいては、リスクマネジメントに関わる人材の育成が重要な課題の1つになると考えられます。近年話題になっている欧州のAIに関する規制でも、ハイリスクとされるAIに対して、人間の監視を求めるとともに、その監視する人間がAIの限界や起こりうるリスクなどを十分に理解していることが求められています。しかし、例えば独立行政法人情報処理推進機構（IPA）と経済産業省が公表しているデジタルスキル標準¹⁾においても、そのようなAIのリスクマネジメントを担う人材については、まだ検討が不十分な状況です。

このような状況を鑑み、NECは2021年から東京大学とAIのリスクマネジメントに関する人材育成の研究に取り組んできました。具体的には、東京大学が開発したリスクチェーンモデル（以下、RCModel）を用いて、NECのDX人材育成機関であるNECアカデミー for AIと連携した、人材育成のプログラム検討を行ってきました。

そこで、本稿では第2章でRCModelの概要を説明し、第3章でNECアカデミー for AIを紹介します。そして第4章でトライアルとして実施した人材育成プログラムの概要とその結果の紹介を行い、第5章で今後の展望を述べます。

2. AIのリスク管理サポートツール RCModel

RCModelとは、自らが企画・開発・運営するAIサービスのリスクを識別し、そのリスクを最小限にできるかを考え、第三者に説明するためのツールであり、東京大学が開発を進めています²⁾。

RCModelでは、特定のAIサービスの事例について、AIの概要（導入価値・目的やシステム構成など）から、その価値・目的の達成を阻害するリスクシナリオと、その対応（コントロール）について整理を行います。ここでの価

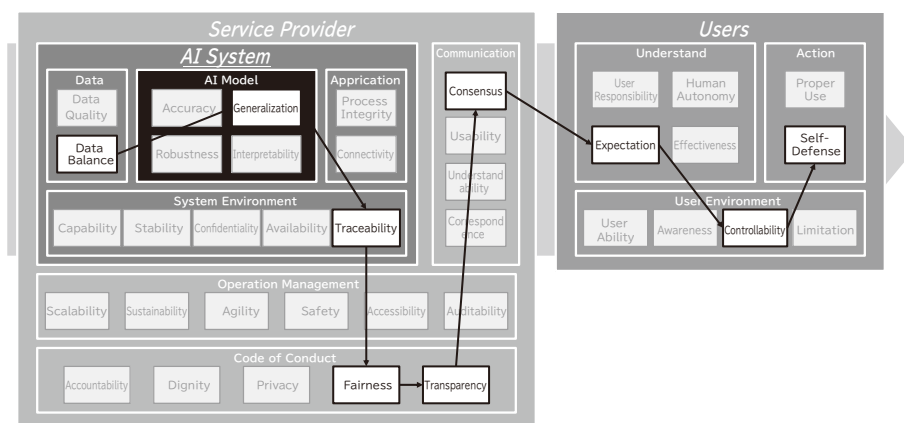


図1 RCModelの利用例

値・目的とは、例えば「生産性の向上」「事故の抑制」「顧客満足度の向上」というような、AIを導入した理由を指します。そしてリスクシナリオとは、「AIの精度が維持できない」「AIの振る舞いが人によって異なる」など、そのAIを導入した価値・目的を阻害しうる場合を指します。このようなリスクシナリオはさまざまに考えることができ、また対応も1つの正解があるわけではない点は留意が必要です。

RCModelは、一般的にAIに関して考慮すべきとされる項目を整理した、38の構成要素（予測性能、データの品質、公平性、利用者の責任など）をもとに階層構造（AIモデル、AIシステム、サービス提供者、ユーザー）が形成されています。そして、対応の検討を行う際には、リスクシナリオごとに、対応における具体的な作業とその順番を検討し、各作業に紐づく構成要素をハイライトし、ハイライトした構成要素間をチェーンで結びます。

例えば、図1の例ではDataBalance（データの分布）→Generalization（特定の事例に偏ることが少ない判断）→Traceability（AIサービスを事後検証できるための性質）→Fairness（サービス全体での公平性）→Transparency（AIサービスに係る必要な情報の開示）→Consensus（利用者との認識合わせ）→Expectation（AIサービスに対する期待精度の理解）→Controllability（利用者側での制御）→Self-Defense（利用者自身の保護）の順番にチェーンが引かれています。各項目の詳細説明は割愛しますが、これはあるリスクシナリオについて、その対応が、各項目が含まれ

る階層に関係する組織や担当者が相互に連携して実施していくことが、チェーンのつながりによって表現されています。この例では、リスクへの対応が複数の組織・関係者にわたっていることを示しています。

このようにRCModelは、特にリスクの対応について記載するフレームまで備えているという特徴を持っています。そのため、Digital Ethics Compass³⁾などAIサービスのリスクを検討する他フレームとの相性も良く、NECとして注目しています。

3. NECアカデミー for AI

日本全体でデジタル人材の育成が必要とされていることを背景として、NECでは「NECアカデミー for DX」という人材育成プログラムを提供しています。特に、その中でAI・データを利活用する人材育成に向けたプログラムとして「NECアカデミー for AI」を2019年から開講しています。

NECアカデミー for AIでは、内容・期間が異なるさまざまなプログラムを提供していますが、最も長期のプログラムに1年間、メンターとともに研修・模擬演習・実践（実プロジェクトでのOJT）を通して知識習得と実務経験を獲得する「入学コース」というものがあります。この入学コースは、NECグループ及びユーザー企業から、将来、企業のDX部門をけん引するリーダー人材として育成することを目的とした受講生（アカデミー生）が参加しています。

4. AIリスクマネジメントの人材育成プログラム

4.1 人材育成プログラムの概要

今回の人材育成プログラムは、2022年4月から7月までの約3カ月間、NECアカデミー for AIの入学コースに参加する4名のアカデミー生を対象に、計4回のワークショップ（以下、WS）形式で行いました（初回は概要説明のガイダンス回）（図2）。

アカデミー生は、その出身母体である企業の業種を参考に、2名ずつの2チームに分かれ、チーム内で個別課題として仮想的なAIサービスをケースとして検討を行ってもらいました。検討は、各チームが設定した仮想のAIサービスについて、そのリスク及び対応を課題として検討してもらい、WSの中で東京大学側の関係者及び他チームへ発表・ディスカッションをしてもらうことで、課題をブラッシュアップしてもらいました。

人材育成プログラムをWS形式とした理由には、大きく2つの期待があったためです。

1つ目は、RCModelを通して、異なる観点（視座・視野・視点）からのリスクに気づき、より適切な対応を検討することができるのではないかという期待です。近年、AIサービスの実装・運用が複数企業で行われることも珍しくありません。その場合、多様な専門知識（業務の詳細な内容や業界慣習など）がAIサービスのリスクマネジメントには求められます。更に、それら専門知識は明文化されておらず、暗黙知になっている場合が多くあります。そこで、AIサービスの実装・運用に関するノウハウをRCModelのフレームで形式知として表現し、WSの中での議論を通して双方がノウハウを学び合うことにより、リスクマネジメントスキルの向上につながるのではと考えました。

2つ目は、RCModelを通して、多様なステークホルダーを認識し、それらステークホルダー間で合意形成をする必要性に気付くことができるのではという期待です。RCModelでは、AIシステム、システムプロバイダ、ユーザーとステークホルダーごとのカテゴライズがされており、リスク対応には、それらステークホルダーが相互に連携して実施していくことが、チェーンのつながりによって表現されています。これは、関係するステークホルダーで対応方法について合意を取る必要があることも示しています。WSの中での議論を通して、どのようなステークホルダーが存在し、どう役割分担していくのかを検討できるのではと考えました。

4.2 受講生・メンターからの人材育成プログラムへの評価

WSを終えたのち、受講生・メンターへ、人材育成プログラムへの評価をインタビューしました。その結果、受講生からは次のようなコメントを得ました。

自分では気付かないリスク・対応について、気づき・検討することで、リスクマネジメントのスキルが向上する

表1にリスクや対応の検討に関するコメントを整理しました。これらコメントから分かることは、RCModelは、自分では気付かないリスクに気づき、自分では想像しづらい対応フローを検討することをサポートしてくれるツールであり、WS形式で複数人と議論することで、その効果も高まるということが期待できるということです。第4章1節に記載した通り、RCModelは暗黙知であるAIサービスのリスク・その対応検討のノウハウを形式知として記録することから、他者の観点について学ぶことをサポートしてくれていたことが伺

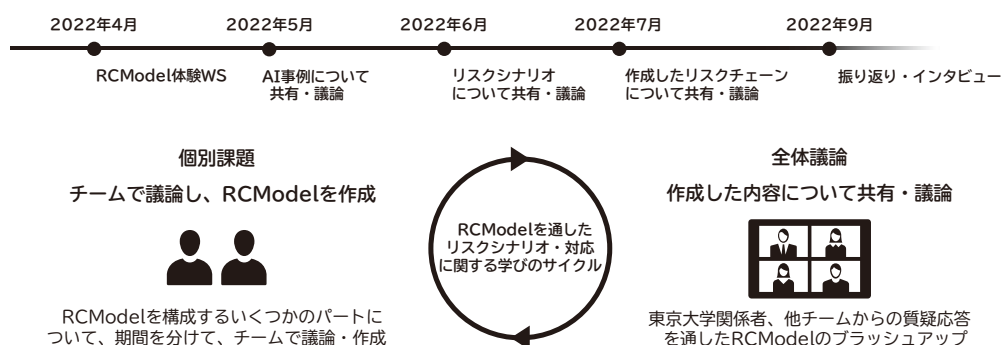


図2 人材育成プログラムの概要

表1 リスクや対応の検討に関するコメント

コメント内容	発言者
AIの専門知識がない人に対してもリスクシナリオとリスク対策にかかわる共通認識を作れるのではないかな	受講生A
普段はシステム開発やデータサイエンティスト業務に従事しているが、RCModelを用いることで、現場目線でのリスク・対応でも検討が可能となった	受講生B
チームで検討することで、異なる視点で、知見を補足し合えることができた	受講生C
RCModelを用いることで、BCP観点も踏まえた、業務のフローチャートとして活用できると感じた	受講生D
RCModelの方法論で検討すると、自身の経験では気付けないリスク・対応についても補うことができるのではないかな	メンター

表2 ステークホルダーとの合意に関するコメント

コメント内容	発言者
RCModelを用いることで、AIのリスクコントロールにかかわる対外的な説明をサポートしてくれる	受講生A
WSから、各ステークホルダーの役割の明確化がAI活用のうえで重要と感じた	受講生C

えます。具体的なケースについて、多様なステークホルダーで議論をすることで、多様な観点からのリスク検討や、より適切な対応方法の検討が期待できるということが、これらコメントから示唆される結果となりました。

ステークホルダー間での役割分担の重要性を認識し、RCModelを共通フレームワークとして合意形成を促す

表2にステークホルダーとの合意に関するコメントを整理しました。これらコメントから、RCModelを用いたAIサービスのリスク・対応検討を通して、ステークホルダー間での役割分担の重要性を認識することができたことが伺えます。また、RCModelという共通フレームワークを用いることで、AIサービスのリスクマネジメントに欠かせない多様なステークホルダーとの折衝や合意形成を促す効果が期待できる結果となりました。

5. むすび

本稿では、RCModelを用いて、NECのDX人材育成

機関であるNECアカデミー for AIと連携した、人材育成プログラムの検討について紹介しました。冒頭で述べた通り、AIガバナンスに関する社会の注目の高まりに伴い、AIサービスのリスクマネジメントを実現できる人材の必要性も高まっています。

今回のトライアルでは、RCModelを用いたWS形式のプログラムには、自分一人では気付くことができないリスクや対応について学ぶ効果が期待できることや、ステークホルダー間での役割分担の重要性を確認し、合意形成を促進する効果が期待できる結果となりました。これらは、AIリスクマネジメントに関する人材育成に有用な方法論として活用できるという成果につながると考えています。

NECでは、この成果をNECアカデミー for AIのプログラムのなかで積極的に活用していきます。

参考文献

- 1) 経済産業省：デジタルスキル標準, 2023年8月
https://www.meti.go.jp/policy/it_policy/jinzai/skill_standard/main.html
- 2) 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事務局：リスクチェーンモデル (RCModel) ガイド Ver 1.0, 2021年6月
https://ifi.u-tokyo.ac.jp/wp/wp-content/uploads/2021/07/RCM_210705.pdf
- 3) DDC: The Digital Ethics Compass
<https://ddc.dk/tools/toolkit-the-digital-ethics-compass/>

執筆者プロフィール

伊藤 宏比古

グローバルイノベーション戦略統括部
プロフェッショナル
産学官連携コーディネーター

伊藤 千央

AI・アナリティクス統括部
マネージャー
リードDXラーニングプランナー

関連URL

オープンで信頼あるAI社会の実現のために

～東京大学とNECが共に考えたAI倫理～

<https://wisdom.nec.com/ja/feature/ai/2023042101/index.html>

リスクチェーンモデルの応用事例ウェビナー Risk Chain Model to Practice「NECアカデミー for AIでの人材育成プログラム」

<https://ifi.u-tokyo.ac.jp/event/14233/>

NEC 技報のご案内

NEC 技報の論文をご覧くださいありがとうございます。
ご興味がありましたら、関連する他の論文もご一読ください。

NEC技報WEBサイトはこちら

NEC技報（日本語）

NEC Technical Journal（英語）

Vol.75 No.2 ビジネスの常識を変える生成AI特集 ～社会実装に向けた取り組みと、それを支える生成AI技術～

ビジネスの常識を変える生成AI特集によせて
生成AI技術への取り組み ～基盤から応用、ルール作りまで～

◇ 特集論文

急速に広まる生成AIの市場適用

NECにおける生成AIの取り組みについて
電子カルテと医療文書の作成支援による医師業務効率化
映像×LLMによる報告書作成業務の自動化
映像分析と生成AIによるリアル世界の行動理解
サイバー脅威インテリジェンス生成自動化
生成AIの社内活用を進めるNEC Generative AI Service (NGS)
ソフトウェア・システム開発への生成AIの活用
LLMとMIで革新する素材開発プラットフォーム
LLMと画像分析を活用した被災状況の把握

生成AIの可能性を高める基盤技術

日本語性能に優れたNECの大規模言語モデル (LLM)
NECの生成AIを支える国内企業で最大規模のAIスーパーコンピュータ
より安全な大規模言語モデル (LLM) を目指して
データを秘匿したまま連携できる連合学習技術とLLMへの適用可能性
大規模言語モデル (LLM) によるFew-shotクラスタリングの可能性
オープンドメイン常識推論のための知識拡張型プロンプト学習
AI連携とオーケストレーションのための基盤ビジョンLLM
クエリを考慮した新規手法により関連する企業データを減らしLLM APIの使用コストを最適化

AI技術が社会へ浸透するために

AI標準化・ルールメイクに関する動向とNECの取り組み
人権尊重に向けたNECのAIガバナンスの取り組み
RCModelを用いたAIリスクマネジメントのための人材育成事例

◇ NEC Information

2023年度C&C賞表彰式典開催



Vol.75 No.2
(2024年3月)

特集TOP