

概要

本調査研究事業の目的は、(1)特に長期の再犯・再非行リスクに影響する要因についてデータ解析を行い、保護観察におけるリスク予測モデリングに必要な項目や特徴量について基礎的な知見を洗い出すこと、(2)今後、長期再犯・再非行のリスクおよび急性再犯・再非行リスクの予測等に必要な開発・実装を行い、令和9年度頃の運用を目指すこととした場合の具体的な課題を整理し、再犯・再非行の予測に関するAI導入に向けた当面の具体的なロードマップについて検討することである。

当該目的を達成するための前段として、第一章では、保護観察領域における再犯・再非行防止のためのアセスメントの基本的枠組み(Risk-Need-Responsivityモデル)を踏まえた上で、アセスメントツールの発展的な利用形式としてリスク予測モデリングを位置付け、その利点や限界点を整理した。また、開発指針や倫理的側面に関連する検討事項と、実装から導入にかけての個別検討課題など、機械学習技術の活用を前提とする上での開発と評価の枠組みを共有した。

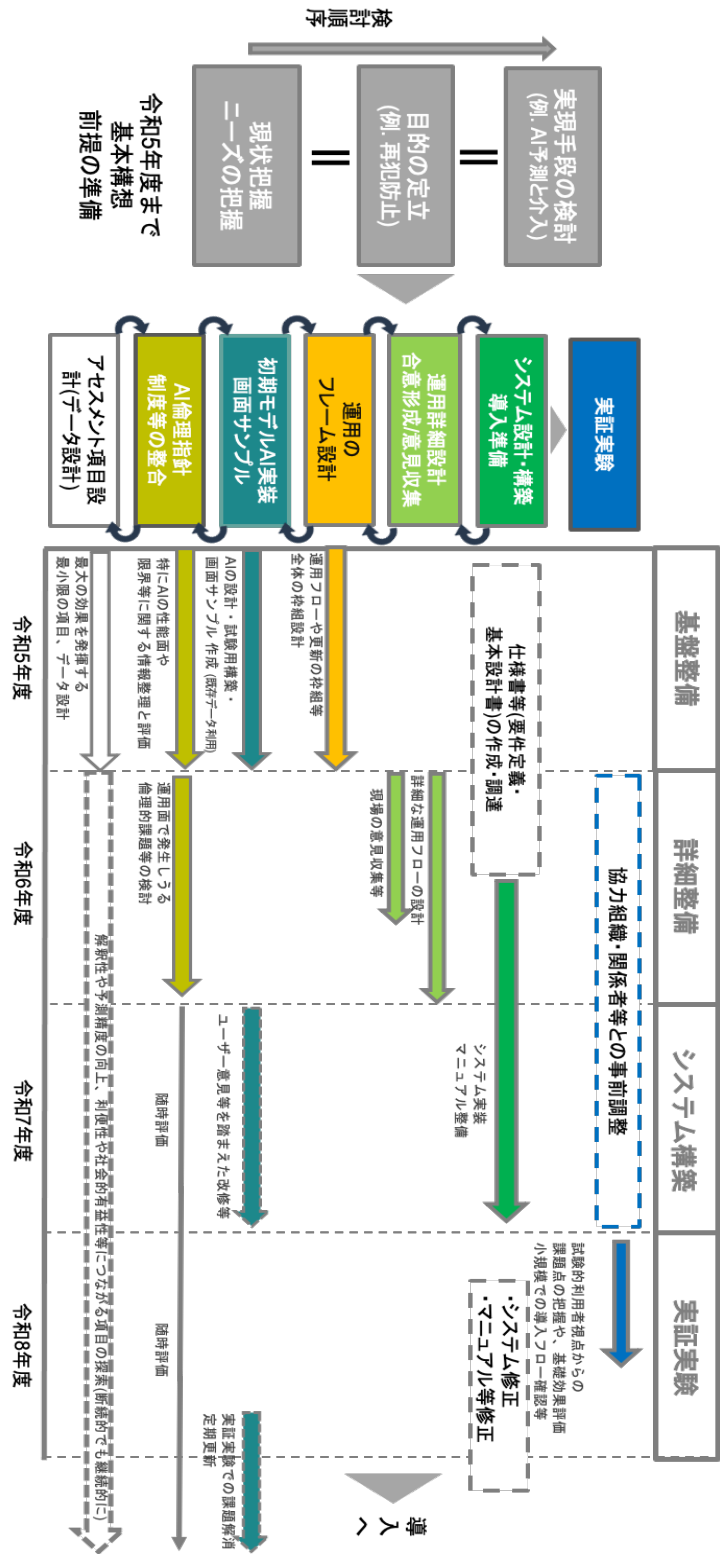
保険数理的アプローチを超えて、リスク予測モデリングの導入を試みることの中核的意義は、正確な予測の実現にある。第二章では、事件管理システムに蓄積された事案の「基礎データ」を用いて、5年以内の再犯・再非行の発生をどの程度の精度で予測可能かについて、予測的妥当性の基礎評価を実施した。その結果、異なるデータや条件間で単純比較は困難であるものの、長期的な再犯・再非行の予測が基礎データだけでも一定の精度で実現できることが示された(AUC-ROCs ≥ 0.713)。ただし、より正確な予測を目指す視点や、予測結果を解釈し、各種処遇等の検討に活用してゆく上では、基礎データのみを用いたリスク予測モデルの実装には不足があると考えられた。

再犯・再非行に対する予測精度の向上や、より実践的に結果を解釈し、それを活用することを図るためには、対象者のアセスメントデータを活用する必要がある。第三章では、データの蓄積期間に限りがあるものの、Case Formulation in Probation/Parole(CFP)を用いたアセスメント情報と短期的な再係属との関連についての基礎的な解析を実施した。その結果、短期的な再係属を一定以上の精度で予測できることが示された(AUC-ROCs ≥ 0.692)。当該結果は、既存の保険数理的アプローチよりも予測精度の視点で優れていると評価されうる。しかし、総合的な視座からして、既に十分な性能が得られていると形容されるものではない。予測的妥当性の向上を含めた今後の発展を考える上で、重要なアセスメント項目を再度選抜・評価し、項目化するなどの手法が功を奏する可能性が指摘された。

第四章では、第二章および第三章の解析結果を踏まえ、今後の開発に寄与しうるいくつかの技術的工夫の例を紹介した。実践的に結果を解釈し、各種処遇等への援用に貢献しうる「反実仮想説明(Counterfactual Explanations)」の概要や、予測的妥当性や解釈可能性の向上のみならず、リスク予測モデリング全体の基礎を支えるアセスメント項目の設計を補助する計画的欠損データデザイン(Planned Missing Data Design)に基づく調査法などである。ロードマップに示される検討事項を具体的に達成するための提案手段の一つとなっている。

第五章では、AI-CFPの開発と導入に向けた主要な参照源となりうる、個人データの取り扱いに関する原則、サービスデザイン思考、AIのリスク管理に関するフレームワークを紹介し、その上で、令和9年度頃以降の導入を前提としたロードマップ案を作成した。ロードマップは、「現状把握・ニーズの把握」と「目的の定立」と「実現手段の検討」の対応関係が明確化され、基本構造が定式化されていることを前提に、(1)アセスメント項目設計、(2)AI倫理指針・各種制度等との整合、(3)初期モデル実装/画面サンプルの作成、(4)運用のフレーム設計、(5)運用詳細設計・合意形成・意見収集、(6)システム設計構築・導入準備、(7)実証実験のそれぞれを段階的に検討してゆく流れで構成された。直近の

検討観点に位置付けられた(1)～(4)は、リスク予測モデリングの導入そのものの是非に関わる事項を含むことや、今後の積み重ねに大きく影響するという点で、特に重要かつ慎重な検討が必要であることが指摘された。



AI-CCFPの導入に係る開発等のロードマップ案
令和4年度 「保護観察におけるアセスメントへのAI導入に関する調査研究業務」での作成版