

児童相談所業務における A I の利活用の在り方に関する調査研究 (効果検証結果中間報告)

株式会社野村総合研究所

コンサルティング事業本部

ヘルスケア・サービスコンサルティング部

〒100-0004

東京都千代田区大手町1-9-2

大手町フィナンシャルシティ グランキューブ

2024年12月26日

NRI

Envision the value,
Empower the change



1. 一時保護AI判定ツールの開発経緯と現状 | 機能概要

AI虐待リスク判定システムは、入手したケース情報を基にリスクアセスメント項目に入力を行うことで、リスクスコアを提示するシステムである。

3.2.1 リスクアセスメント機能

データベースサーバー
(教師データ格納領域)

アプリケーションサーバー
(AIツール)



教師データとする
データの格納

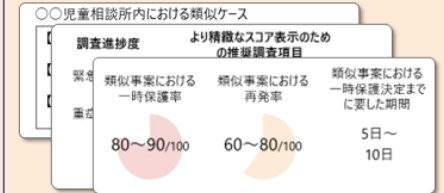
データベースサーバー
(都道府県別格納領域)



リスクアセスメントシート
の入力



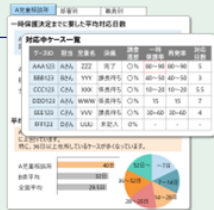
リスクスコア・類似ケース
等の分析結果の送信



3.2.2 ケース進行マネジメント機能



管轄ケース一覧・ケース対応
に関する統計情報の表示



管理者端末

テキストメッセージ、画像・
音声・動画データの
送受信



3.2.3 チャット機能

AI虐待リスク判定システムの検証の概要は下記の通り。

目的



- AIリスク判定ツールによる判定結果と実際の判断の差分を確認し、当該ツールが一時保護要否のアセスメントに活用可能か評価すること。
- AIツールへの入力時間の計測及びヒアリングにより、運用時の業務負荷を定性/定量的に把握すること。

対象



- 全国10自治体
- 各児童相談所における協力職員数は、児童相談所ごとにご調整いただいた。ただし、課長・SV級の方にご参加いただき、各所の一時保護判断の妥当性の担保に努めた。

実施方法



- AIRisk判定ツールの機能について改めて案内したうえで、各児童相談所において、入力を行っていただいた。
- 判定結果と実際の判断の差分を確認した。
- 入力作業終了後、ヒアリングを通して、運用負荷、現場運用可否の各観点について検証した。

一時保護スコアに対しては100件中13件で「高い」、41件で「低い」、8件で「幅が広い」という評価であり、6割程度のケースでスコアに対して疑義が生じた。

AI判定精度について

スコア所感

- 一時保護スコアに対しては、100件中13件が「高い」、41件が「低い」、37件が「妥当」、8件が「幅が広い」、1件が「不明」という評価であった。
- 再発スコアに対しては、100件中7件が「高い」、36件が「低い」または「やや低い」、「妥当」が49件、3件が「幅が広い」、5件が「不明（判断できない）」という評価であった。
- 再発スコアに対して、同様の事象が生じる可能性であるのか、虐待種別等を問わず、虐待が再度発生する可能性であるのかを明確にすべきという意見が挙げられた。

ケース情報の入力

- ケースにおける事象に該当するアセスメント項目が存在しないケースが複数あった。
- 該当するアセスメント項目は存在するものの、粒度や程度が十分に反映しきれないケースが複数あった。
- 継続して対応しているケースにおいて、改善が見られないことや類似の事象の繰り返し、リスクの変化といった要素がアセスメントの中に反映できないという指摘があった。
- リスク要因に限定されたアセスメントシートとなっており、強みや環境の情報が反映されないケースが複数あった。

判定の傾向

- 身体的虐待で家庭環境が複雑でないケースは比較的妥当なスコアが算出された。
- 一部の重篤な身体的虐待・性的虐待では、ルールベースでの判定がなされることから、一時保護スコアは妥当なスコアであった。
- 一方、ベテラン福祉司が「ただちに一時保護すべき」と判断する事例に対して、重大な見落とし（著しく低い一時保護スコア：2～3/100）が発生した。
- また、AIではこどもの帰宅拒否・保護の求めを重視するが、重篤な虐待行為が生じておらずこどもの帰宅拒否が強いケースでもスコアが高く算出されてしまう傾向にあった。

リスクを正しく入力できなかった状況として、①発生事象や環境に該当する項目がないケース、②該当項目はあるが程度・範囲が反映できていないケースの2パターンが存在する。

①発生事象や環境に該当する項目がない

ケースにおける重要な事象や対応方針（一時保護判断）を決定するにあたって重視した児童・養育者・支援者等の情報として該当するアセスメント項目がなく、ツールに情報を与えられていない場合

検証ケースにおいて生じた例

体重減少

重篤なネグレクトで体重減少がみられるが-2SDには該当しない

墜落分娩

出産まで妊娠に気がつかず、自宅のトイレで出産したが児童の健康状態に問題はない

親の過干渉

こどもの年齢に不釣り合いな干渉や支配性がある

養育者に対する怯え

（保護を求めているなくとも）こどもが養育者に対しておびえている

こどもの発達年齢

発達・情緒に問題を抱える場合など実年齢と比較して意見表明が難しい/発言が不明瞭である

支援者の状況・分離

家庭内に支援者がいる、家庭以外（祖父母宅等）で分離ができる

②該当項目はあるが、程度・範囲が適切に反映できていない

ケースにおける重要な事象や対応方針を決定するにあたって重視した情報に該当するアセスメント項目自体は存在するものの、アセスメント項目に該当する場合が広範で、程度や範囲が適切に反映されていない場合

検証ケースにおいて生じた例

受傷の程度

傷あざややけどの範囲や深度が深い/浅い

受傷の部位

頭部・顔面でも特に危険な部位への傷がある

帰宅拒否の程度

養育者との接触を拒絶する帰宅拒否/深刻度の低い帰宅拒否

被虐待歴・通告歴

過去の複数の通告歴よりも重篤な通告内容であることから、よりリスクが高まるケース

養育者の態度

過去の経過・関係機関からの情報により介入への拒否感が特に強い

開発したAIツールは、AI技術の更なる進歩を踏まえた性能改良が必要であるため、現状でのリリースは延期することが望ましい。

精度検証結果

- 現段階で検証を行った全100ケースのうち約6割（62ケース）において、現場職員の判断との乖離が生じた。
- 検証ケースの中では、ケースの状況を十分にAI判定ツールにインプットできないものが存在した。
- 具体的には、所定の項目以外にも一時保護判断に影響する情報が非常に多岐にわたって存在していること、所定項目にも「程度」が存在しており、各項目の該当可否だけでは、虐待リスクを算出する情報として不足していることが挙げられる。

AI判定ツールのリリースについて

- 国が全国に提供する画一的なツールとして考えた際に、現状の判定精度は十分ではないこと、特に経験年数の浅い職員ではケース対応の初動を誤る恐れがあることから、現在のAI判定ツールのリリースは時期尚早といえる。
- ケースに関するすべての情報をアセスメントシートに反映できない一方で、現状以上にアセスメント項目を増やすことは運用上現実的ではないことから、更なるAIの改良が必要である。

児童相談所におけるAI活用の方向性（これまで出た意見）

- 当面は児童相談所の業務の中で負担感の大きい定型業務を中心にAIの活用を進め、業務負荷軽減を図るべき。
- 将来的には、経過記録や児童記録票等の定性情報を基にした、進行管理やケースワークのサポート等に活用することも考えられるのではないか。
- そうした可能性を見据え、定性的な構想実現に向けた教師データの収集を進めるとともに、児童相談所の現場における業務課題の大きい記録作成業務の負担軽減を実現する観点から、AIを活用した面談音声の書き起こし・要約作成ツールの開発が求められているのではないか。