

令和2年度厚生労働科学研究費補助金

政策科学総合研究事業(政策科学推進研究事業)

児童虐待対応におけるリスクアセスメントのためのデータ収集基盤構築と
AIを活用したリスク評価に向けた研究(19AA1002)

令和2年度 総括・分担研究報告書

研究代表者 高岡 昂太

令和3(2021)年 5月

目 次

I. 総括研究報告 児童虐待対応におけるリスクアセスメントのためのデータ収集基盤構築と AIを活用したリスク評価に向けた研究 高岡昂太	-----3
II. 分担研究報告	
1. リスク項目「不明」に該当しても精度を維持できるAIの構築 北村光司	-----5
2. 現場業務改善と人材育成に向けたデータ利活用ループの検討 本村陽一	-----7
III. 研究成果の刊行に関する一覧表	-----8
IV. 倫理審査及び利益相反の管理の状況に関する報告	-----9

令和元年度厚生労働科学研究費補助金（政策科学総合研究事業）
（総括）研究報告書

児童虐待対応におけるリスクアセスメントのためのデータ収集基盤構築と
AIを活用したリスク評価に向けた研究に関する研究

研究代表者 高岡 昂太 産業技術総合研究所

研究要旨：機械学習技術によりXGBoostを用いて保護判断履歴を高い精度で再現可能にしたことと同時に、誤予測の可能性を考慮したBaysian Deep Learningを用いて判断の補助を最大限補助できる可能性を考慮した。

A. 研究目的

本研究では児童の安全を継続的に確保するため、「虐待の再発(継続)」を考慮した対応判断や「一時保護の実施」に係る意思決定を支援する人工知能(以後AIと略す)開発と現場使用に関する探索的な研究を行う(研究1)。また、特に誤予測の可能性を減じるため、「誤予測」の可能性を考慮した技術設計と活用方法を確立する(研究2)。

B. 研究方法

それぞれ1自治体のデータベースを展開し、基本情報、リスクアセスメント、経過記録に関するデータを検討し、AIの研究開発を行った。

研究1の学習条件を次のように設定した。三重県の協力の下、学習データ80%(6924件)、検証データ20%(1731件)に無作為に分割。一時保護実施ラベルの不均衡(該当率14%)解消のためSMOTEサンプリングを適用した。eXtreme Gradient Boosting(XGBoost)モデルを利用する際に、5-fold 交差検証で学習roundを決定。損失関数は対数損失。学習率0.1を設定した。最大tree-depthは7(4~8を探索して最高精度)。その他デフォルト設定とした。予測対象は「一時保護の実施の有無」とし、リスク項目を含めた32項目の事例情報を採用した。なお、リスク項目は該当 = 1, 非該当と不明を0でラベルコーディングされ、通告経路等のカテゴリー情報はラベルエンコーディング(擬似連番)した。

研究2の学習条件を次のように設定した。学習データ80%(6924件)、検証データ20%(1731件)に無作為に分割とした。一時保護実施ラベルの不均衡(該当率14%)解消のためSMOTEサンプリングを適用した。採用する手法は、ベイズ多層ニューラルネットワークモデル(中間層3層 各36ノード)とした。予測対象は「一時保護の実施の有無」とし、リスク項目を含めた32項目の事例情報を採用した。リスク項目は該当 = 1, 非該当を0、不明を9でラベルコーディング(擬似連番)とした。

(倫理面への配慮)

産総研内の人間工学実験事前審査非該当、愛育研究所における倫理委員会の承認を得た。

C. 研究結果

研究1の結果は、図1・図2のように、閾値を0.5とする判別予測正解率(Accuracy) = 0.91 (91%) (クラス不均衡のため参考値)が得られた。

研究3の結果は、PR曲線下面積でXGBoostよりやや

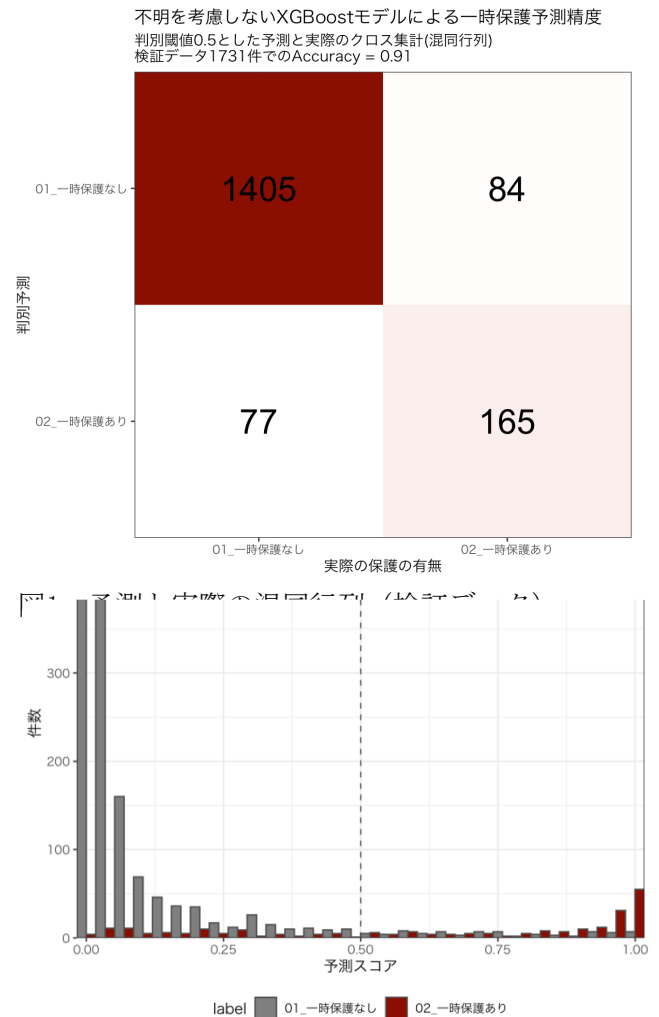


図2：実際の一時保護の有無別での予測スコア分布と設定閾値(検証データ)

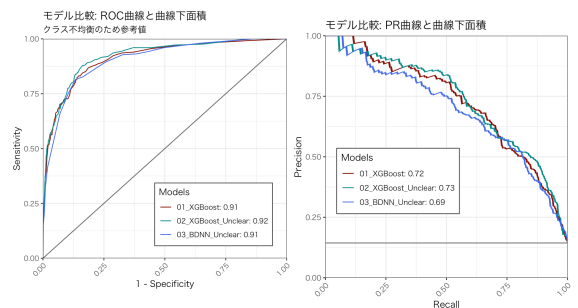


図3：各種性能指標

劣るものの、遜色ない性能を示した(図3)。また、幅が狭く(確信度が高く)予測が的中する場合と、幅が広く(確信度が低く)予測が誤っている場合を示すことができた。(図4)。

ベイズ深層学習による「幅のある予測」の例
 白丸は実際の保護の有無(有 = 1, 無 = 0)
 青点は予測値(EAP)、青線は予測の幅(95%確信区間)

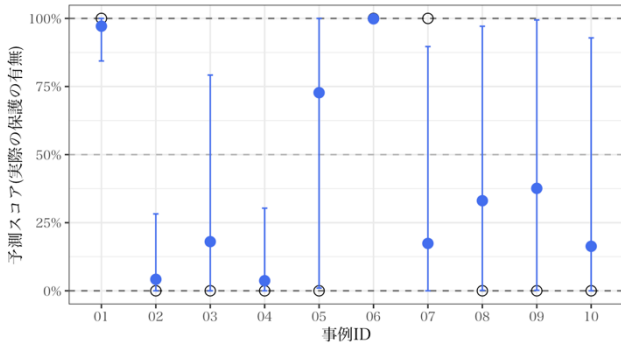


図4：事例毎の確信区間に関するイメージ

一方で、予測の幅が広いほど、誤予測率が高い傾向を事前に把握できた(図5)。

ベイズ深層学習による予測スコアの幅と誤予測率
 点は区間別の誤予測率、エラーバーは平均値の95%信頼区間を示す

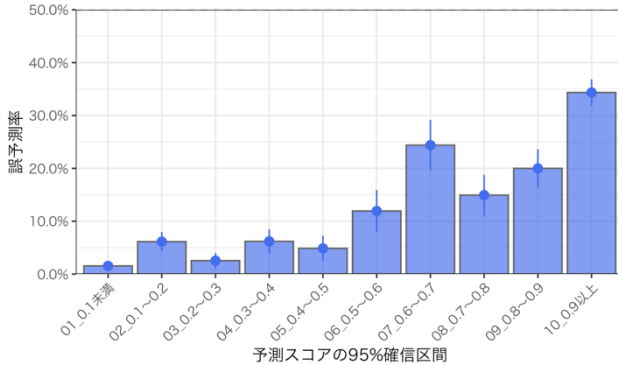


図5：誤予測率の把握

D. 考察

研究1については、XGBoostによりミスの少ない、一貫した一時保護判断が実現できたと考えられる。一方で、過去の一時保護判断が常に妥当であったとは限らないため、“要保護”に相当する情報を予測対象とした研究が必要(本事業範囲外)と考えられる。即ち、新規の虐待通告事例に対して、「過去ではその事例を一時保護していたか」を高い再現率で数的に評価可能となったと言える。

研究2については、図4のように、各新規事例に対する確信度付きの予測(予測区間)が実現された一方で、図5のように、確信区間の幅の大きさによって、どの程度の割合で当該予測結果が「誤りである

か」が事前に把握できるようになると考えられる。例えば、「一時保護予測スコアが”0.9”と出力されても、予測の幅が0.9以上の場合(図5の10_0.9以上に該当)、35%の割合で誤予測の可能性があると把握できる。一方で、確信区間の幅が0.1未満の場合、誤予測は2%未満とも言える。これにより、一時保護判断という重大な意思決定場面で、各事例に対する「機械学習による予測」がどの程度の確信度を持つものか、誤予測の可能性を事前に評価可能となったと言える。

こうした成果は、増加する虐待通告や対応職員の入れ替わりに伴って懸念される一時保護判断の一貫性の低下に対して、保護判断の水準を維持する客観的な指標が提供できる可能性が示唆された。

他にも、「一時保護の実施を予測する」という当該取り組みは“要保護児童の識別”に寄与すると考えられ、将来的にはコールセンター(区分対応システム)等での活用も期待されるだろう。

一方で、本研究の限界として、過去の一時保護判断が常に妥当であったとは限らない。そのため、本研究と同時並行に、“要保護”に相当する情報を予測対象とした研究が必要となる。次に、今回のAI研究開発は、データ提供元の特定の自治体に適用が限定される。個別自治体の資源(一時保護所の状況等)に即している点は適切だが、他の自治体に展開するためには、個別自治体でのデータ蓄積と解析が次年度以降の必要となる。最後に、誤予測の可能性を考慮したベイズ深層学習の精度(AUCPR)がやや低いことがあげられる。そのため、精度向上を検討すると、(解析的に大規模な)パラメータチューニングが必要となるだろう。

以上のように、データ提供元のように、初期初動のリスクデータが適切に収集されることで、重要なベースライン情報となり現場に必要なAI開発に役立てることが明らかになった。一方で、AIはあくまで一つの手段であるため、現場にデータ解析の結果を現場にフィードバックし、必要な業務改善や人材育成につなげることも同時に行っていく必要があるだろう。

E. 結論

一時保護判断については、ある程度精度が保てる結論となった。一方で、今後はAI以前に、良質なデータを集めるために、研修を含めた実務と平行した取り組みが不可欠である。

F. 健康危険情報

特になし

G. 研究発表

1. 論文発表：なし
2. 学会発表：なし

H. 知的財産権の出願・登録状況

1. 特許取得：なし
2. 実用新案登録：なし
3. その他：なし

令和2年度厚生労働科学研究費補助金（政策科学総合研究事業）
（分担）研究報告書

児童虐待対応におけるリスクアセスメントのためのデータ収集基盤構築と
AIを活用したリスク評価に向けた研究に関する研究

研究代表者 高岡 昂太 産業技術総合研究所

研究要旨：自治体の児童相談所がもつリスクアセスメントデータを元に、不明が多く選ばれた場合でも、ある程度精度を維持できるAI構築を行った。

分担研究者：北村光司
産業技術総合研究所人工知能研究センター
主任研究員

A. 研究目的

リスク項目への該当状況が「不明」の場合への対応。

B. 研究方法

それぞれ1自治体のデータベースを展開し、基本情報、リスクアセスメント、経過記録に関するデータを検討し、AIの研究開発を行った。学習条件を次のように設定した。学習データ80% (6924件) / 検証データ 20% (1731件)に無作為分割。一時保護実施ラベルの不均衡(該当率14%)解消のためSMOTEサンプリングを適用した。解析アルゴリズムにeXtreme Gradient Boosting (XGBoost)モデルを利用し、5-fold 交差検証で学習roundを決定した。損失関数は対数損失とし、学習率0.1とした。最大tree-depthは7(4~8を探索して最高精度)とし、その他はデフォルト設定を採用した。予測対象は「一時保護の実施の有無」とし、リスク項目を含めた32項目の事例情報を採用した。リスク項目は該当 = 1, 非該当を0、不明を9でラベルコーディングし、通告経路等のカテゴリー情報はラベルエンコーディング(擬似連番)した。

(倫理面への配慮)

産総研内の人間工学実験事前審査非該当、愛育研究所における倫理委員会の承認を得た。

C. 研究結果

閾値を0.5とする判別予測正解率(Accuracy) = 0.91 (91%) (クラス不均衡のため参考値)を保持していた。“不明”のパターンを学習させても性能は維持されていることを確認した (図6, 図7参照)。

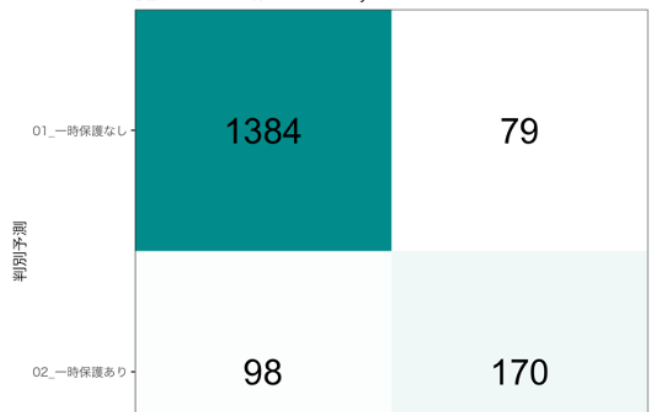
D. 考察

児童虐待対応現場において、初期初動体制では情報が不確かであったり、情報が集まらない場面が多くある。しかしながら、今回の調整を通して、初期初動で「不明」があったとしても、ある程度の精度を保てるAIを構築できたことは非常に大きい。さらに、今後は汎用可能性を含めて検討していくことが重要である。

E. 結論

初期初動において、リスクアセスメント項目に「不明」が回答されても、必要な一時保護を提案できる可能性を検討できた。今後は、汎用可能性があるか、さら

不明を考慮したXGBoostモデルによる一時保護予測精度
判別閾値0.5とした予測と実際のクロス集計(混同行列)
検証データ1731件でのAccuracy = 0.91



不明を考慮したXGBoostモデルによる予測スコア
破線は閾値0.5を示す

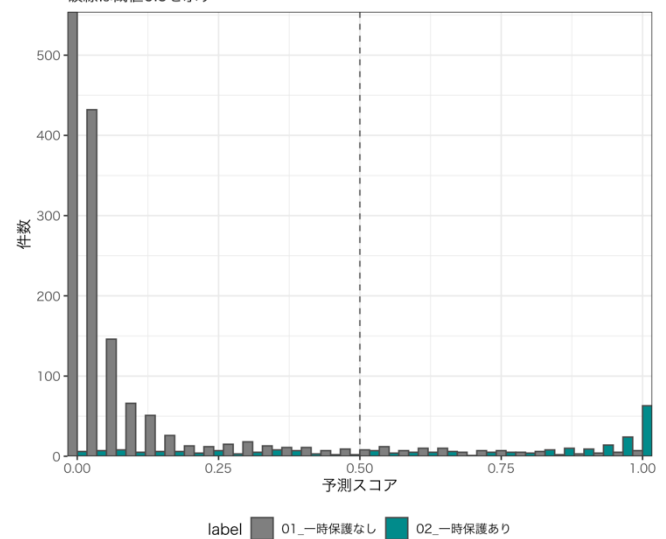


図7：実際の一時保護の有無別での予測スコア分布と設定閾値（検証データ）

に他の自治体でもデータ収集を含めた検討をしていくことが重要である。

F. 健康危険情報

特になし

G. 研究発表

1. 論文発表

なし

2. 学会発表

なし

H. 知的財産権の出願・登録状況
(予定を含む。)

1. 特許取得

なし

2. 実用新案登録

なし

3. その他

なし

令和2年度厚生労働科学研究費補助金（政策科学総合研究事業）
（分担）研究報告書

児童虐待対応におけるリスクアセスメントのためのデータ収集基盤構築と
AIを活用したリスク評価に向けた研究に関する研究

研究代表者 高岡 昂太 産業技術総合研究所

研究要旨：自治体の児童相談所がもつ過去の児童相談データを今後大枠でどのように利活用していくか、現場業務改善と人材育成に向けたデータ利活用ループを検討した。

分担研究者：本村陽一
産業技術総合研究所人工知能研究センター
首席研究員

A. 研究目的

今後の研究ディレクションとして、現場のニーズと得られたデータ、また技術シーズをいかにつなぐかについてと、データ利活用において、実践としての研究視点に関する検討を行った。

B. 研究方法

これまでの解析結果を基に、今後どのようにデータ利活用を自治体と協働していくかについて検討を行った。

C. 研究結果

データの利活用を考える場合、現場の課題（ニーズ）を現場と研究者がともに考えられることが重要。AI 開発をして終わりではなく、{データの蓄積→データ解析結果を現場にフィードバックし業務改善} というループを今後、絶えず繰り返していくことが必要である（図8）。

D. 考察

児童虐待対応現場において、システムを開発して終わりではなく、データを用いた業務フローを絶えず改善し続ける仕組みが必要である。AIは必要なデータがあれば、現場にとって非常な有用な知見を得られるが、データの量だけでなく、データの質を問えるよう、データ解析結果だけでなく、現場へのデータ解析フィードバックと現場でのデータを参照した業務フローへの改善などの仕組みが必要となってくると考えられた。

E. 結論

AIは、良質なデータと合わせて用いられる必要がある。そのためには、現場の課題（ニーズ）を明らかにし、それを解決するためにデータと技術を用いるというフレームの理解と、それらを回し続けて課題解決としての業務改善と判断の質の向上をスパイラルアップさせていく必要がある。今後は、現場との協働により、現場の課題を解決するAI技術が開発される必要があり、いかにデータを集められるかについて

現場の実務データが収集され続け、それによりAIが精緻化され、業務が改善され続けるループ（エコシステム）の醸成が鍵

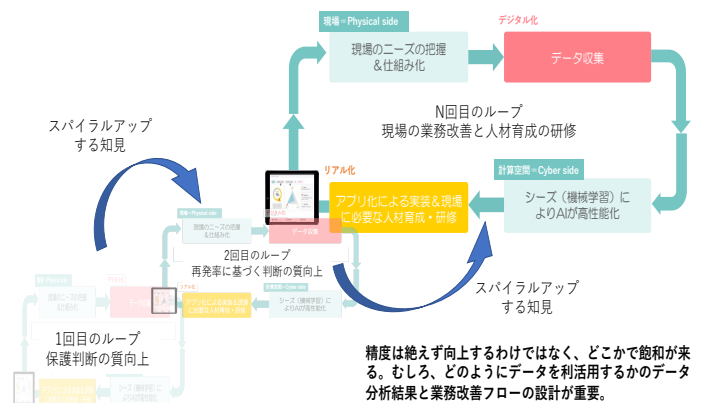


図8：ニーズ・データ・シーズを利用した業務改善のスパイラルアップイメージ

でも全国的にAI導入設計として検討が必要不可欠である。

F. 健康危険情報
特になし

G. 研究発表

1. 論文発表
なし
2. 学会発表
なし

H. 知的財産権の出願・登録状況
(予定を含む。)

1. 特許取得
なし
2. 実用新案登録
なし
3. その他
なし

研究成果の刊行に関する一覧表

書籍

著者氏名	論文タイトル名	書籍全体の編集者名	書 籍 名	出版社名	出版地	出版年	ページ
なし							

雑誌

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
なし					

令和3年4月22日

厚生労働大臣 殿

機関名 国立研究開発法人  研究所

所属研究機関長 職名 理事長

氏名 石村 和彦

次の職員の令和2年度厚生労働科学研究費の調査研究における、倫理審査状況及び利益相反等の管理については以下のとおりです。

- 研究事業名 政策科学総合研究事業（政策科学推進研究事業）
- 研究課題名 児童虐待対応におけるリスクアセスメントのためのデータ収集基盤構築とAIを活用したリスク評価に向けた研究（19AA1002）
- 研究者名（所属部局・職名） 人工知能研究センター・主任研究員
（氏名・フリガナ） 高岡 昂太・タカオカ コウタ

4. 倫理審査の状況

	該当性の有無		左記で該当がある場合のみ記入（※1）		
	有	無	審査済み	審査した機関	未審査（※2）
ヒトゲノム・遺伝子解析研究に関する倫理指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
遺伝子治療等臨床研究に関する指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
人を対象とする医学系研究に関する倫理指針（※3）	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
厚生労働省の所管する実施機関における動物実験等の実施に関する基本指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
その他、該当する倫理指針があれば記入すること （指針の名称： ）	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>

（※1）当該研究者が当該研究を実施するに当たり遵守すべき倫理指針に関する倫理委員会の審査が済んでいる場合は、「審査済み」にチェックし一部若しくは全部の審査が完了していない場合は、「未審査」にチェックすること。

その他（特記事項）

（※2）未審査に場合は、その理由を記載すること。

（※3）廃止前の「疫学研究に関する倫理指針」や「臨床研究に関する倫理指針」に準拠する場合は、当該項目に記入すること。

5. 厚生労働分野の研究活動における不正行為への対応について


研究倫理教育の受講状況	受講 <input checked="" type="checkbox"/> 未受講 <input type="checkbox"/>
-------------	---

6. 利益相反の管理


当研究機関におけるCOIの管理に関する規定の策定	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> （無の場合はその理由： ）
当研究機関におけるCOI委員会設置の有無	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> （無の場合は委託先機関： ）
当研究に係るCOIについての報告・審査の有無	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> （無の場合はその理由： ）
当研究に係るCOIについての指導・管理の有無	有 <input type="checkbox"/> 無 <input checked="" type="checkbox"/> （有の場合はその内容： ）

（留意事項） ・該当する□にチェックを入れること。
・分担研究者の所属する機関の長も作成すること。

厚生労働大臣 殿

機関名 国立研究開発法人  研究所

所属研究機関長 職名 理事長

氏名 石村 和彦  印

次の職員の令和2年度厚生労働科学研究費の調査研究における、倫理審査状況及び利益相反等の管理については以下のとおりです。

- 研究事業名 政策科学総合研究事業（政策科学推進研究事業）
- 研究課題名 児童虐待対応におけるリスクアセスメントのためのデータ収集基盤構築とAIを活用したリスク評価に向けた研究（19AA1002）
- 研究者名（所属部局・職名） 人工知能研究センター・主任研究員
（氏名・フリガナ） 北村光司・キタムラコウジ

4. 倫理審査の状況

	該当性の有無		左記で該当がある場合のみ記入（※1）		
	有	無	審査済み	審査した機関	未審査（※2）
ヒトゲノム・遺伝子解析研究に関する倫理指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
遺伝子治療等臨床研究に関する指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
人を対象とする医学系研究に関する倫理指針（※3）	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
厚生労働省の所管する実施機関における動物実験等の実施に関する基本指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
その他、該当する倫理指針があれば記入すること （指針の名称：)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>

（※1）当該研究者が当該研究を実施するに当たり遵守すべき倫理指針に関する倫理委員会の審査が済んでいる場合は、「審査済み」にチェックし一部若しくは全部の審査が完了していない場合は、「未審査」にチェックすること。

その他（特記事項）

（※2）未審査に場合は、その理由を記載すること。

（※3）廃止前の「疫学研究に関する倫理指針」や「臨床研究に関する倫理指針」に準拠する場合は、当該項目に記入すること。

5. 厚生労働分野の研究活動における不正行為への対応について

研究倫理教育の受講状況	受講 <input checked="" type="checkbox"/> 未受講 <input type="checkbox"/>
-------------	---

6. 利益相反の管理

当研究機関におけるCOIの管理に関する規定の策定	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> （無の場合はその理由：)
当研究機関におけるCOI委員会設置の有無	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> （無の場合は委託先機関：)
当研究に係るCOIについての報告・審査の有無	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> （無の場合はその理由：)
当研究に係るCOIについての指導・管理の有無	有 <input type="checkbox"/> 無 <input checked="" type="checkbox"/> （有の場合はその内容：)

（留意事項） ・該当する□にチェックを入れること。
・分担研究者の所属する機関の長も作成すること。

厚生労働大臣 殿

機関名 国立研究開発法人 XXXXXXXXXX 研究所

所属研究機関長 職 名 理事長

氏 名 石村 和彦 XXXXXXXXXX

次の職員の令和2年度厚生労働科学研究費の調査研究における、倫理審査状況及び利益相反等の管理については以下のとおりです。

- 研究事業名 政策科学総合研究事業（政策科学推進研究事業）
- 研究課題名 児童虐待対応におけるリスクアセスメントのためのデータ収集基盤構築と AI を活用したリスク評価に向けた研究（19AA1002）
- 研究者名 （所属部局・職名） 人工知能研究センター・首席研究員
（氏名・フリガナ） 本村陽一・モトムラヨウイチ

4. 倫理審査の状況

	該当性の有無		左記で該当がある場合のみ記入（※1）		
	有	無	審査済み	審査した機関	未審査（※2）
ヒトゲノム・遺伝子解析研究に関する倫理指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
遺伝子治療等臨床研究に関する指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
人を対象とする医学系研究に関する倫理指針（※3）	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
厚生労働省の所管する実施機関における動物実験等の実施に関する基本指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
その他、該当する倫理指針があれば記入すること (指針の名称：)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>

（※1）当該研究者が当該研究を実施するに当たり遵守すべき倫理指針に関する倫理委員会の審査が済んでいる場合は、「審査済み」にチェックし一部若しくは全部の審査が完了していない場合は、「未審査」にチェックすること。

その他（特記事項）

（※2）未審査に場合は、その理由を記載すること。

（※3）廃止前の「疫学研究に関する倫理指針」や「臨床研究に関する倫理指針」に準拠する場合は、当該項目に記入すること。

5. 厚生労働分野の研究活動における不正行為への対応について

研究倫理教育の受講状況	受講 <input checked="" type="checkbox"/> 未受講 <input type="checkbox"/>
-------------	---

6. 利益相反の管理

当研究機関におけるCOIの管理に関する規定の策定	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> (無の場合はその理由:)
当研究機関におけるCOI委員会設置の有無	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> (無の場合は委託先機関:)
当研究に係るCOIについての報告・審査の有無	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> (無の場合はその理由:)
当研究に係るCOIについての指導・管理の有無	有 <input type="checkbox"/> 無 <input checked="" type="checkbox"/> (有の場合はその内容:)

（留意事項） ・該当する□にチェックを入れること。
・分担研究者の所属する機関の長も作成すること。