PRESS RELEASE

群馬大学 GUNMA UNIVERSITY

2024年11月14日

報道関係者 各位

「AI見守り公共空間」に対する，「AI出力の検証可能な記録」というプライバシ保護概念を提案
～一般市民，監査機関の検証で，AI systemの悪用・乱用を防ぐ～

１．本論文，それに続く研究のポイント

- あらゆる地点が１つ以上のAI接続街路カメラでモニタされる広範な公共エリア「AI見守り公共空間」(＝人類未経験)を例にとり，利点・欠点を検証している．「AI見守り公共空間」に対するプライバシ保護策として，「AI出力の検証可能記録」という概念を提案する．試作システム開発と実証実験により，その実現可能性を検証することが提案される．
- 社会の安全性が飛躍的に向上する，導入コストが低いなど，人類がかつて経験したことのない「AI見守り公共空間」の良い点が検証される．あらゆる地点が１台以上のカメラで撮影・記録されれば，誘拐された子供，徘徊老人の現在位置特定・保護，路上犯罪者の追跡・逮捕，は極めて容易にできる．市街地では，街路灯の照射エリアは重なり合っていることが一般的である．「LED街路灯に，照射エリアと視野が重なるように複数のカメラを内蔵させ，スマートフォン程度のコンピュータを搭載したネットワーク接続機能付き街路灯ユニット」は，それぞれの部品コストから推定して安価（例：1-2万円程度）で高度なものが製作できる．街路灯の交換時期を待てば，カメラとしての設置費用は無視できる．すなわち，「AI見守り公共空間」(＝人類未経験)を安価に実現する技術を，我々は，既に持っている．
- 「AI見守り公共空間」の悪い点として個人のプライバシが侵害される危険性，監視社会化などがあげられる．この対策として，「AI出力の検証可能記録」というコンセプトによるプライバシ保護を提案する．民主主義国家は，プライバシ保護が十分に確立できない場合，「AI見守り公共空間」の実現に躊躇し，社会の安全化，効率化で，プライバシ保護を気にしない国家に対して後れを取る事が懸念される．
- 新たな社会インフラである「AI出力の検証可能記録」の「社会安全に及ぼす効果」と「プライバシ侵害の懸念」の検証，さらに，前者の懸念解決のためのコンセプト「AI出力の検証可能記録」を検証する過程で，様々な新たな問題に遭遇し，それを解決するための発明や特許が生まれることが期待される．私たちの試みが社会に受け入れた時，この研究の世界への貢献は大きい．

２．本論文の概要

プライバシを含む膨大な情報の入力を受けるAI systemが，出現しつつある．一般市民のプライバシを保護することは，民主主義国家にとって，重要な課題である．

本研究では，公共目的に使用されることが想定され，かつ，プライバシ侵害リスクが比較的に高いAI-systemとして，「AI接続されたカメラでモニタされた公共空間」を例にとり，プライバシ保護のための具体的な手法を検討する．「AI接続されたカメラでモニタされた公共空間」は，あらゆる地点が複数のAIに接続された街路カメラでモニタされる広範な公共エリアとして定義される．

群馬大学大学院理工学府知能機械創製部門　藤井雄作教授が、社会的に許容されうる「AI接続されたカメラでモニタされた公共空間」に実装される利活用目的として，(1) 行方不明のこどもの追跡，(2) 子供が自宅でてから戻るまでの見守り，(3) 公共空間での異常な状況や事故の検知，(4) 犯罪者やテロリストの行動分析・検知の４つを取り上げる．これら４つの利活用目的により，社会安全が劇的に向上することを示した．

AI systemに対するプライバシ保護策としては，一般市民が安心できるようにするためには，確実な保護に加えて，万一の悪用・乱用を抑止する確実な記録・監査が必要条件であると考えられる．検証可能・改ざん不可能なAI outputの記録・公開のコンセプト「AI出力の検証可能な記録」を提案する．「AI出力の検証可能な記録」は，Big Dataなどの膨大なAI inputの匿名化や，ブラックボックス化しているAI内部プロセスの制御などよりも，AI outputだけに対してルール適合性確認のみの単純な出口規制をかける方が，現実的ではないか，というアイデアに基づく．

「AI接続されたカメラでモニタされた公共空間」に「AI出力の検証可能な記録」を実装した具体例を提案する．「AI出力の検証可能な記録」においては，AI systemをoutbound firewall で包み込み，「事前に社会的議論を通して承認されている利用目的」に即していることが確認されたAI outputのみ送

信許可される．そして，送信申請から送信許可に至る過程が，匿名化された上で，検証可能・改ざん不可能な形で公開される．この公開記録を，一般市民，監査機関が検証することで，AI systemの悪用・乱用があった場合に，確実に暴かれる．

本論文で提案した「AI出力の検証可能な記録」は，AI outputに関する確実な計測・記録，そして監査のための公開を行うことを特徴としている．これをテンプレートして，一般市民のプライバシを含むBig Dataの入力を受ける他のAI systemへ適用する可能性についても検討する．

今後の展望

本論文に基づく研究計画が立てられ，日本側（群馬大学）とシンガポール側（南洋理工大学）との国際共同研究プロジェクトの予算申請*がされている．（採択された場合，2025年4月研究開始．）本研究では，対象AI-systemとして，あらゆる地点が１つ以上のAI接続街路カメラでモニタされる公共エリア「AI見守り公共空間」を例にとる．人類未経験の「AI見守り公共空間」による「社会安全の飛躍的向上」を検証する．次に，プライバシ保護手法として，検証可能・改ざん不可能なAI outputの記録・公開のコンセプト「AI出力の検証可能記録」を提案し，一般市民のプライバシの効果的保護を検証する．日本側（群馬大学）チームは，主に，「AI見守り公共空間」のハードウェア開発と，独自コンセプト「AI出力の検証可能記録」の設計・実装を担う．シンガポール側（南洋理工大学）チームは，主に，「AI見守り公共空間」に実装する自動追尾・異常検知・犯罪者行動パターン検知などの機能開発を担う．

＊日ASEAN科学技術・イノベーション協働連携事業（NEXUS）日本－シンガポール共同研究公募「AI」

3. 掲載先

雑誌名：AI & Society

公開日：2024年11月 9日

題　名：Verifiable record of AI output for privacy protection: public space watched by AI-connected cameras as a target example

著者名：Yusaku Fujii

ＵＲＬ：https://doi.org/10.1007/s00146-024-02122-8

AI & Society誌は，Springer Science＋Business Media社が発行する査読付き学術雑誌です．このジャーナルは，人工知能とその社会への影響および社会との相互作用のあらゆる側面をカバーします．

【本件に関するお問合せ先】
〈研究に関すること〉
群馬大学　大学院理工学府知能機械創製部門　教授　藤井　雄作（フジイ　ユウサク）
E-MAIL：fujii@gunma-u.ac.jp
〈取材についてのお問合せ〉

群馬大学　理工学部庶務係（広報）
TEL：0277-30-1895
E-MAIL：rikou-pr@ml.gunma-u.ac.jp