



富士通・お茶の水女子大学

AI 倫理社会連携講座

2023 年度 活動報告書

2024 年 6 月 12 日

富士通・お茶の水女子大学 AI 倫理社会連携講座

目次

1.	はじめに	1
2.	AI 倫理社会連携講座 概要	2
3.	AI 倫理社会連携講座に関する活動	4
	3-1:AI hiring の現状・課題について紹介	
	3-2:公平性考慮機械学習技術 (1)	
	3-3:公平性考慮機械学習技術 (2)	
	3-4:バイアス検証例の紹介	
	3-5:PA に対する人々の態度 (1)	
	3-6:PA に対する人々の態度 (2)	
	3-7:ジェンダー少数派の職業威信にジェンダー・ステレオタイプが与える影響	
	3-8:Fairness by Design	
	3-9: IT のジェンダーギャップを教育とエンパワーメントで変える	
	3-10:お茶の水女子大学社会連携講座共同研究テーマ	
	3-11:統計資料調査	
	3-12:BIAS プロジェクト@NTNU の紹介	
	3-13:納得性と合理性を備えたバイアス緩和手段の開発	
4.	AI 倫理社会連携講座における研究	8
	4-1:大企業の新卒採用プロセスにおけるジェンダーバイアスの解明	
	4-2:労働市場におけるジェンダーバイアスの存在とメカニズムに関する定量的検証	
	4-3:社会や組織を公平な状態へ誘導するバイアス緩和手段の開発	
5.	情報発信	10

1. はじめに

富士通・お茶の水女子大学 AI 倫理社会連携講座 拠点長 斎藤悦子

富士通株式会社（以下、富士通）と国立大学法人お茶の水女子大学（以下、お茶の水女子大学）は、2023年3月1日に「富士通・お茶の水女子大学 AI 倫理社会連携講座」（以下、AI 倫理社会連携講座）を設置しました。本報告書は AI 倫理社会連携講座の 2023 年度の活動実績をまとめたものです。

ジェンダー平等は、SDGs のひとつにもなっている世界的に重要な社会課題ですが、日本では、ジェンダーギャップ指標で世界 146 か国中 125 位(2023 年に公表された順位)とジェンダー不平等が著しく、女性活躍施策の導入にも問題があります。その課題解決のために、AI 倫理社会連携講座では AI 活用の可能性を探る共同研究を行います。しかし、AI の活用においても、学習データに潜在している偏りにより、AI によるマイノリティのグループに対する判断が偏向してしまう倫理的リスクの存在など、ジェンダーに関する重要な問題が指摘されています。AI の偏向が引き起こす諸問題は重大なリスクと見なされており、社会的に信頼できることが AI にとって必須要件となりつつあります。このふたつの問題を、それぞれの専門家たちの協働により解決することが、AI 倫理社会連携講座の目的です。

2023 年度は、NPO 法人 Waffle 森田久美子氏にご登壇いただき「IT のジェンダーギャップを教育とエンパワーメントで変える」を開催いたしました。また、多様な分野の専門家が参加し、女性活躍施策の導入に向けた課題の検討や、AI 活用の際の倫理的なリスクについて討論を重ね、共同研究を推進しました。一連の活動については本報告書をご覧ください、忌憚のないご意見を伺えればありがたく存じます。

最後に、AI 倫理社会連携講座の開設および活動にあたっては、各界の方々のご助言とご支援をいただきました。この場を借りて、厚くお礼申し上げます。今後とも企業と大学の連携を強化するために、多くの皆さまのお力添えを賜りたく、よろしく願い申し上げます。

2. AI 倫理社会連携講座 概要

AI 倫理社会連携講座では、富士通の AI 倫理技術とお茶の水女子大学のジェンダー・イノベーション研究の知見をもとに、世界に先駆けて、AI を活用した定量的かつ客観的なジェンダー平等施策を可能にする共同研究を行うとともに、その遂行を担う広範な領域に精通した人材の育成を進めます。

研究者（職位：2024年3月31日現在）

役割	氏名	職位
拠点長／社会科学チームリーダー	斎藤 悦子	ジェンダー・イノベーション研究所（IGI）／生活社会科学講座 教授、IGI 副研究所長
副拠点長／講座担当者：研究戦略検討	稲越 宏弥	富士通株式会社 AI トラスト研究センター リサーチディレクター／お茶の水女子大学客員准教授
プロジェクトマネージャー／社会科学チーム	柏木 志保	IGI 社会連携講座准教授
講座担当者：技術戦略検討・研究推進	新田 泉	富士通株式会社 AI トラスト研究センター シニアリサーチマネージャー／お茶の水女子大学客員准教授
心理学チームリーダー	大森 美香	基幹研究院人間科学系／心理学科 教授、IGI 研究員
情報科学チームリーダー	伊藤 貴之	基幹研究院自然科学系／理学部情報科学科 教授、IGI 研究員
社会科学チーム	西村 純子	基幹研究院人間科学系／生活社会科学講座 教授
	脇田 彩	基幹研究院人間科学系／生活社会科学講座 助教
心理学チーム	伊藤 大幸	基幹研究院人間科学系／心理学科 准教授
	武藤 世良	基幹研究院人間科学系／人間社会科学科 講師
	宮川 祥子	人間文化創成科学研究科 博士前期課程 人間発達科学専攻
情報科学チーム	小林 一郎	基幹研究院自然科学系／理学部情報学科 教授
	五十嵐 悠紀	基幹研究院人間科学系／理学部情報学科 准教授

役割	氏名	職位
講座協力者：研究戦略検討	高橋 秀和	富士通株式会社 AIトラスト研究センター 研究員 (2024年3月31日まで)
講座協力者	中田 恒夫	富士通株式会社 AIトラスト研究センター 特任研究員 (2023年12月20日まで)

支援体制（職位：2024年3月31日現在）

役割	氏名	職位
担当理事・副学長／拠点設置研究所長	石井クツ 昌子	理事・副学長（研究・国際交流・男女共同参画担当） ／ジェンダード・イノベーション研究所（IGI）所長
監督者	太田 裕治	副学長（産学連携・イノベーション担当） ／IGI 副研究所長
支援	後藤 正智	富士通株式会社 AIトラスト研究センター長
支援	永井 洋樹	富士通株式会社テクノロジービジネスマネジメント本部 産学連携推進室 室長
支援	大平 弘太郎	富士通株式会社テクノロジービジネスマネジメント本部 産学連携推進室 マネージャー
支援	清水 雅芳	富士通株式会社テクノロジービジネスマネジメント本部 産学連携推進室 エキスパート
支援／社会科学チーム	吉原 公美	IGI／リエゾン・URA センター リサーチ・アドミニストレーター
支援	-	研究・産学連携課 企画戦略課（男女共同参画担当） IGI 事務局

3. AI 倫理社会連携講座に関する活動

3-1:AI hiring の現状・課題について紹介

【日時】2023年5月8日

【報告者】富士通株式会社 AIトラスト研究センター 新田泉

【報告概要】

本報告ではAI技術を活用した採用プロセスの現状と課題が紹介された。AI技術はデータに基づいて学習するため、過去の差別や偏見を再生産する可能性がある。たとえば、過去に採用された求職者が男性ばかりだった場合、AIは男性の求職者に高い評価を出すように学習してしまう。これは、男性が多かった職種に女性が参入する際に不利な立場に置かれる可能性につながる。AI採用システムの運用において、バイアスを緩和し、公正な採用プロセスを確保するための研究や取り組みが重要であることが指摘された。

3-2:公平性考慮機械学習技術（1）

【日時】2023年5月14日

【報告者】富士通株式会社 AIトラスト研究センター 稲越宏弥

【報告概要】

本報告では機械学習における公平性の概念と課題について、シンポジウム講演資料を参考に議論された。機械学習は人間が目的を決めて、そのために利用する道具であり、データやモデルのバイアスによって公平性が失われるという問題が生じる。そのため、機械学習の利用にあたっては、データの代表性やモデルの公平性に注意を払う必要があり、特に人に対する判断に活用する場合には、出力をそのまま受け入れるのではなく人が介入・監視をする必要があることが言及された。

3-3:公平性考慮機械学習技術（2）

【日時】2023年5月22日

【報告者】富士通株式会社 AIトラスト研究センター 新田泉

【報告概要】

本報告では機械学習による意思決定における公平性について、講演資料を参考に議論された。報告では、公平性の定義、指標について説明があり、特に集団公平性と個人公平性に焦点が当てられた。機械学習における公平性の問題は、データ収集と学習の過程で生じるバイアスが原因であることが指摘された。集団公平性の指標としてはデモグラフィックパリティと Equalized Odds が紹介され、個人公平性には Individual Fairness が挙げられた。さらに、各指標の理論的課題が紹介された。公平性の確保は入試、与信審査、採用など、意思決定が差別を生む可能性がある分野において重要である。

3-4 : バイアス検証例の紹介

【日時】2023年5月29日

【報告者】富士通株式会社 AIトラスト研究センター 高橋秀和

【報告概要】

本報告では、アメリカのロースクール入学試験のデータを使用した、高等教育の入学試験におけるバイアス検証事例を紹介した。結果として、性別については偏りが見られなかったが、人種については白人の合格率が高く、分析データには表れない世帯年収、教育環境、居住地など背景にある社会状況の影響が示唆された。バイアス検証結果の説明には、統計的にバイアスが見られても、その結果を読み解くための社会状況や業務知識などの背景知識が必要である。

3-5 & 3-6:PA に対する人々の態度(1)(2)

【日時】2023年6月5日/2023年6月12日

【報告者】お茶の水女子大学基幹研究院人間科学系/心理学科 伊藤大幸

【報告概要】

本報告ではポジティブアクションへの態度を考察した。Crosby, F.J. et al. (2006). は、ポジティブアクションの目的、必要性、効果と課題について社会科学の視点から検討した。多様性や公平性の確保がポジティブアクションの目的とされ、効果と課題については賛否が分かれた。また、Harrison, D. A. et al. (2006). は、メタ分析を用いてポジティブアクションへの態度を形成する要因の定量分析を行った。126の研究を分析した結果、プログラムの規制の強さや個人・集团的自己利益、差別意識などが態度に影響することが明らかとなった。今後はより複雑な研究が求められ、個人差を考慮したアプローチの検証が重要である。

3-7:ジェンダー少数派の職業威信にジェンダー・ステレオタイプが与える影響

【日時】2023年6月19日

【報告者】お茶の水女子大学基幹研究院人間科学系/生活社会科学講座 脇田彩

【報告概要】

本報告では、職業に関するジェンダー・ステレオタイプがジェンダー少数派の職業威信に与える影響が検討された。どのような職業において、どのような仕組みで、女性（男性）就業者は男性（女性）就業者より高く評価されるのか、またそこにジェンダー・ステレオタイプは関わっているのかを明らかにすることを目的に、調査・分析が行われた。その結果、人々が各職業に関して持つジェンダー・ステレオタイプによって、ジェンダー少数派の職業威信はジェンダー多数派よりも低くなる傾向があることが明らかになった。

3-8:Fairness by Design

【日時】2023年6月26日

【報告者】富士通株式会社 AIトラスト研究センター 新田泉

【報告概要】

Fairness by Design をテーマにし、異なる文化的背景における公平性を考慮した AI 開発の重要性が議論された。機械学習モデルにおけるバイアスや公平性に対処するための方法として、共同デザインワークショップや機械学習モデルの公平性を審査する Graphical User Interface (GUI)の作成、各文化圏の人々によるモデルの修正、認識しにくいバイアスの修正などがある。特に、銀行データを使用した実践結果では、各国文化圏に合わせた公平なモデルが作成されたことが示された。最後に、バイアスの修正や公平性の確保は、今後の機械学習モデルの設計において重要な要素であるという結論が述べられた。

3-9:IT のジェンダーギャップを教育とエンパワーメントで変える

【日時】2023 年 7 月 3 日

【報告者】NPO 法人 Waffle 森田久美子氏

【報告概要】

特定非営利活動法人 Waffle は、教育とエンパワーメントを通じて IT 分野のジェンダーギャップを是正するというミッションのもと、女子およびジェンダーマイノリティの中高大学生に向けたハンズオンプログラムやイベント活動、政策提言活動を行っている。本講演では、まず、日本における IT 業界におけるジェンダーギャップの現状が言及され、その要因として、社会のジェンダー・ステレオタイプと高校の早い段階での文理選択が影響を与えていることがデータと共に提示された。次に、この課題に対する活動として、Waffle が取り組む中高大学生向けの IT 教育が紹介された。また、大学と企業ができることについて、米国の大学や国内企業の事例を交えて提言がなされた。参加者から、IT に興味を持つ学生を増やすための施策について質問があった。森田氏からは、IT をツールとして活用した場合に何ができるのかを明確に伝えることにより、目的意識を持った学生に関心を持ってもらうと回答があった。また、大学の女子学生枠などのアフターマティブアクションについても議論された。男女を均等にするという数合わせに終わるのではなく、入学前後の支援策が重要であることが示唆された。さらに、AI はジェンダーギャップを解消するか、もしくは逆に強化してしまうか、という質問について、学習データがジェンダーギャップを反映していることや開発者の多様性が少ないことによる懸念が示され、使い方のルールや開発者の多様性確保が課題であることが共有された。この他にも、大学、企業それぞれの立場からの質問、議論が行われた。

3-10:お茶の水女子大学社会連携講座共同研究テーマ

【日時】2023 年 7 月 10 日

【報告者】富士通株式会社 AI トラスト研究センター 稲越宏弥

【報告概要】

本報告では、男女共同参画政策とその実現のための手段のひとつであるポジティブアクション(PA)の法解釈面についての学びを共有したのちに、PA の手法の中でも比較的穏やかとされているプラスファクター方式を公平化機械学習の枠組みに取り入れるという研究テーマ案と、これを実際に社会実装した際に想定される受容性の問題に取り組む研究テーマの素案を示し、これらの案について議論した。

3-11:統計資料調査

【日時】2023年7月24日

【報告者】富士通株式会社 AIトラスト研究センター 新田泉
富士通株式会社 AIトラスト研究センター 高橋秀和

【報告概要】

本報告では、産業構造基本統計調査のデータ分析に基づいて女性の活躍推進に関する調査報告がなされた。産業分類毎に、労働者全体および管理職における女性比率の過去10年間の推移を比較した。教育学習支援業は203030(2020年代の早い段階で指導的地位における女性比率が30%となる)目標の達成が見込めるが、製造業、卸売業・小売業、情報通信業の3業種は現状のペースでは203030の達成が難しい状況を示した。

3-12:BIAS プロジェクト@NTNU の紹介

【日時】2023年7月31日

【報告者】お茶の水女子大学ジェンダー・イノベーション研究所 (IGI) /リエゾン・URA センター リサーチ・アドミニストレーター 吉原公美

【報告概要】

ノルウェー科学技術大学 (NTNU) が主導し欧州の4大学5企業が参加するBIASプロジェクトは、本講座同様、AIを活用した社会課題解決に取り組んでいる。バイアス緩和AIの開発によって雇用プロセスにおけるバイアスを減らし、労働市場により平等で公平な慣行を構築することを目指している。プロジェクトの目的は技術開発にとどまらず、信頼性の高いツールの開発を通じて、AIとヒューマン・リソース・マネージメント (HRM) コミュニティの意識の向上、能力開発、エンパワメントを図ることにある。AI開発は、一般公開されている言語モデル、プロジェクトパートナー企業から提供される採用データに加え、AIおよびHRMの専門家へのインタビューや、プロジェクトのラボで実施する労働者らが参加するワークショップなどから得られた、人文社会科学知見のインプットに基づいて行われる。ラボは複数の国に設置され、国ごとの雇用文化の違いも分析する。プロジェクトのウェブサイトは <https://www.biasproject.eu/>

3-13:納得性と合理性を備えたバイアス緩和手段の開発

【日時】2023年11月24日

【報告者】富士通株式会社 AIトラスト研究センター 稲越宏弥

【報告概要】

本報告では、統計的機械学習の理論的な観点から、ポジティブアクション(PA)の解釈を試みた。すなわち、総合評価が同レベルであれば一方の性を優遇する場合(プラスファクター型PA)と、ある項目の評価の重みを高くし、それが間接的に一方の性を優遇する結果になる場合との違いを機械学習は区別しない。このことから、プラスファクター型PAが法的に認められる場面では、結果的に一方の性を選好するモデルを機械学習により構築することも認めてよい場合がありうる。もしそうならば、性別を直接の基準とすることを避けながら、一方の性を優遇する別の基準を発見するAIを開発することや、その別の基準というものかのようなものであれば人々に受容されるのかを社会科学や心理学の研究が解明することの重要性をあらためて強調した。

4. AI 倫理社会連携講座における研究

4-1:大企業の新卒採用プロセスにおけるジェンダーバイアスの解明

【研究概要】

チームリーダー	斎藤 悦子	ジェンダード・イノベーション研究所／生活社会科学講座 教授、IGI 副研究所長
	西村 純子	基幹研究院人間科学系／生活社会科学講座 教授
	脇田 彩	基幹研究院人間科学系／生活社会科学講座 助教
	柏木 志保	ジェンダード・イノベーション研究所 社会連携講座准教授
	吉原 公美	ジェンダード・イノベーション研究所 (IGI) /リエゾン・URA センター リサーチ・アドミニストレーター

女性管理職比率の向上には、「女性の採用から管理職・役職へのパイプライン」の入り口である女性の新卒採用比率の改善が重要である。2023 年度は新卒採用プロセスに焦点をあて、採用の各プロセスにおけるジェンダーバイアスの存在を明らかにすることを目標として、大企業 5 社の採用責任者に対するインタビュー調査を行った。調査の結果は、2024 年度に開催される学会等で発表する予定である。

4-2:労働市場におけるジェンダーバイアスの存在とメカニズムに関する定量的検証

【研究概要】

チームリーダー	大森 美香	基幹研究院人間科学系／心理学科 教授、IGI 研究員
	伊藤 大幸	基幹研究院人間科学系／心理学科 准教授
	武藤 世良	基幹研究院人間科学系／人間社会科学科 講師
	宮川 祥子	人間文化創成科学研究科 博士前期課程 人間発達科学専攻

男女格差を説明しうる変数を網羅的にカバーする包括的調査により、日本の労働市場におけるジェンダーバイアスの実態を定量化することを目的とした。就活経験学生 557 名への web 調査を実施し、新卒採用の 9 つのアウトカム指標について、86 の説明変数（大学特性、企業特性、パーソナリティ、態度、能力、行動）に基づき男女格差のメカニズムを検証した。2023 年度に実施した調査を 2024 年度も継続して調査を実施・分析し、研究成果を 2024 年度以降に開催される学会等で発表する予定である。

4-3:社会や組織を公平な状態へ誘導するバイアス緩和手段の開発

【研究概要】

チームリーダー	稲越 宏弥	富士通株式会社 AI トラスト研究センター リサーチディレクター ／お茶の水女子大学客員准教授
	新田 泉	富士通株式会社 AI トラスト研究センター シニアリサーチマネージャー ／お茶の水女子大学客員准教授
	高橋 秀和	富士通株式会社 AI トラスト研究センター 研究員 (2024年3月31日まで)
	中田 恒夫	富士通株式会社 AI トラスト研究センター 特任研究員 (2023年12月20日まで)

AIの判断結果を公平とされる状態に緩和するバイアス緩和技術は多くの実用的な手法が開発されており、これらは組織を平等な状況に変化させると期待されている。しかしながら、アクションの継続を組織に浸透させ、公平な状況に導いていくためには、1回のアクションによる組織の変化が過剰にならないこと、アクションの根拠となる判断基準を意思決定者やAIの影響を受ける人々に説明できることが必要であろう。そこで、1) バイアスの是正によって判断結果が変化する量が最小となるように緩和する機能と、2) バイアスの是正による判断結果の理由を、できるだけ少数の特徴量に基づく判断基準によって説明できる機能を開発した。これにより、現状を過剰に是正することなく、是正の根拠となる基準を説明できるという条件を満たしながらより公平なアクションを提示し、平等な状況へと導く仕組みを提供する。

5. 情報発信

年月	発信メディア・イベント	URL
2023年3月17日	お茶の水女子大学・富士通共同プレスリリース	https://www.ocha.ac.jp/news/20230317.html https://pr.fujitsu.com/jp/news/2023/03/17-1.html
2023年5月19日	Fujitsu Blog	https://corporate-blog.global.fujitsu.com/fgb/2023-05-19/01/
2023年7月1日	NHK サタデーウォッチ 9	-
2023年7月21日	毎日新聞	https://mainichi.jp/articles/20230719/k00/00m/020/159000c
2023年10月1日	日本経済新聞電子版／日経ヴェリタス	https://www.nikkei.com/article/DGXZQOUB223NE0S3A920C2000000/
2023年10月20日	お茶の水女子大学ジェンダード・イノベーション研究所産学交流会	https://www.cf.ocha.ac.jp/igi/j/menu/project/d013705.html
2023年11月30日	フジトラニュース (Fujitsu transformation news)	https://www.fujitsu.com/jp/microsite/fujitsutransformationnews/2023-11-30/01/
2023年12月13日	朝日新聞 Aging Gracefully	https://aginggracefully.asahi.com/column/15074752
2024年2月12日	日本経済新聞 朝刊「機会の差別、イノベーションでなくす 映画館や安全基準」	https://www.nikkei.com/article/DGXZQOUB282GM0Y3A221C2000000/
2024年2月24日	東京新聞	https://www.tokyo-np.co.jp/article/311150
2024年3月25日	「ジェンダード・イノベーションと AI で目指す、多様な人材が活躍できる社会」ホワイトペーパー、お茶の水女子大学 AI 倫理社会連携講座 発行	https://www.cf.ocha.ac.jp/igi/j/menu/project/fujitsu_ocha_d/fil/AI_Ethics_WhitePaper_24_3.pdf

年月	発信メディア・イベント	URL
2024年3月25日	「ジェンダー・イノベーションとは～その事例と課題～」第10回富士通グループ AI 倫理外部委員会 石井副学長講演	https://www.fujitsu.com/jp/about/research/technology/ai/aiethics/#anc-04

※太字は外部メディア

書名 富士通株式会社・お茶の水女子大学 AI 倫理社会連携講座 2023 年度活動報告書
発行日 2024 年 6 月 12 日
編集・発行 富士通株式会社・国立大学法人お茶の水女子大学
「富士通・お茶の水女子大学 AI 倫理社会連携講座」
〒112-8610 東京都文京区大塚 2 - 1 - 1
T E L 03 - 5978 - 5597
U R L
https://www.cf.ocha.ac.jp/igi/j/menu/project/fujitsu_ocha.html#content
編集担当 柏木志保