Measuring Impact of Artificial Intelligence on U.S. Federal Government

Ujjayini Das and Srijeeta Mitra
PhD Students, Joint Program in Survey and Data Science
University of Maryland

UNIVERSITY OF MARYLAND
FEARLESSLY FORWARD

Acknowledgement

Tejwansh S. Anand

Clinical Professor of Practice & Academic Director for the MS in Information Systems
Robert H. Smith School of Business
University of Maryland

PI of the AI Impact Project

Bertrand A. Stoffel

Visiting Research Fellow
Robert H. Smith School of Business
University of Maryland

Co-PI of the AI Impact Project

Partha Lahiri

Director & Professor
Joint Program in Survey Methodology
Professor
Department of Mathematics
University of Maryland

Advisory Member of the Team

Topics to Cover

- Background and Motivation
- Research Objective
- Data and Methods
- Results
- Conclusion
- Reference

Topics to Cover

- Background and Motivation
- Research Objective
- Data and Methods
- Results
- Conclusion
- Reference

Introduction to AI in Government

- **AI Capabilities and Growth** (Čerka et al., 2017)
  - Operates independently, making decisions without human involvement
  - Advances in data and computational power have increased AI's popularity.
  - Automates both routine and complex tasks
- **AI in Government Functions** (Gomes de Soussa et al., 2019)
  - General public service
  - Economic affairs
  - Environmental protection

Challenges with AI Implementation

- Adoption and diffusion of AI can take years
  - Requires changes in organizational structures, business models, and legal frameworks (Brynjolfsson & Mitchell, 2017)
- Implementation of AI presents ethical challenges
  - Data protection, decision-making processes, and replacement of human labor with autonomous systems (Butterworth, 2018)
  - Threat to privacy and confidentiality (UNESCO, 2021; Ariga et al., 2021)
- Inequality in AI training
  - Inequality in AI training; access divides based on education level (GWOF, 2023 Keegan, 2020; Microsoft, 2023; Tamayo et al., 2023; CGB*,2024)

* Center for Global Business, UMD Robert H. Smith School of Business

Motivation for Current Research

- **Limited Understanding of AI's Full Impact** (Faraj et al., 2018)
    - Insufficient knowledge about how AI algorithms affect government operations, workforce dynamics, and service delivery
- **Keeping up with AI's rapid growth**
    - Field of AI is continually evolving – Is the government keeping up with the pace?
- **Measuring the uncertainties around adoption of AI**
    - Perception about adopting AI
    - Potential consequences of such adoption based on different job functions

Topics to Cover

- Background and Motivation
- Research Objective
- Data and Methods
- Results
- Conclusion
- Reference

Research Objective

- Assessing impact of AI on U.S. Federal Government
  - Where are the agencies at in terms of adopting AI?
  - How employees at different job functions perceive the potential consequences of AI adoption?
- Capturing the dynamic shift in government operations due to AI implementation
- Motivate policy discussions for stronger AI-employee alliance

Topics to Cover

- Background and Motivation
- Research Objective
- Data and Methods
- Methods
- Results
- Conclusion
- Reference

Questionnaire Design

Web survey (PC and smartphone) containing questions on

- General Sections (Perception of AI, Individual experience with AI, AI in workplace, Agency's success with AI) (AI Survey Final Report, Center for Global Business, UMD)
- Targeted questions for people with different job functions decision-makers, implementers and users)
- Vignettes (questions following randomized experimental conditions)
- Demographics (same questions from the Federal Employee Viewpoint Survey, FEVS)

Data Description

- **Target Population:** Current or former (left after 12/31/22) Federal employees, consultants, and contractors (full-time or part-time) to the Federal Government
- **Simulated Data**
  - Simulated responses for 18 questions with 5-point Likert-scale type response options covering a few important constructs to measure the impact of AI
    - Response options – “not at all”, “somewhat”, “moderate”, “very much”, “extremely”
  - Size of the simulated data set – 1000 $\times$ 23 (18 questions and 5 demographic variables)

Simulation (1)

- **Generating Demographics:** Taking FEVS as a benchmark, generated 5 demographic variables (age, gender, supervisory status, time in Federal Govt. and ethnicity)

- **Generating Responses:** Consider the following cumulative-logit model: Suppose Y_{jk} denotes the response to the j^{th} question for the k^{th} respondent

$$logit(p_{jkl}) = \mu_l + \boldsymbol{X}'\boldsymbol{\beta} + \gamma_j + \lambda_k; \quad \mathrm{j} = 1(1)18, \mathrm{k} = 1(1)1000$$

Where, $p_{jkl} = P(Y_{jk} \leq l), l = 1\ (1)\ 5$

μ_l -> log-odds of choosing response category l for a given question

$\boldsymbol{X}$ -> Design matrix consisting of demographic variables

$\gamma_j \sim N(0, \sigma_\gamma^2), \lambda_k \sim N(0, \sigma_\lambda^2)$ -> Random effects corresponding to questions and respondents respectively

Simulation (2)

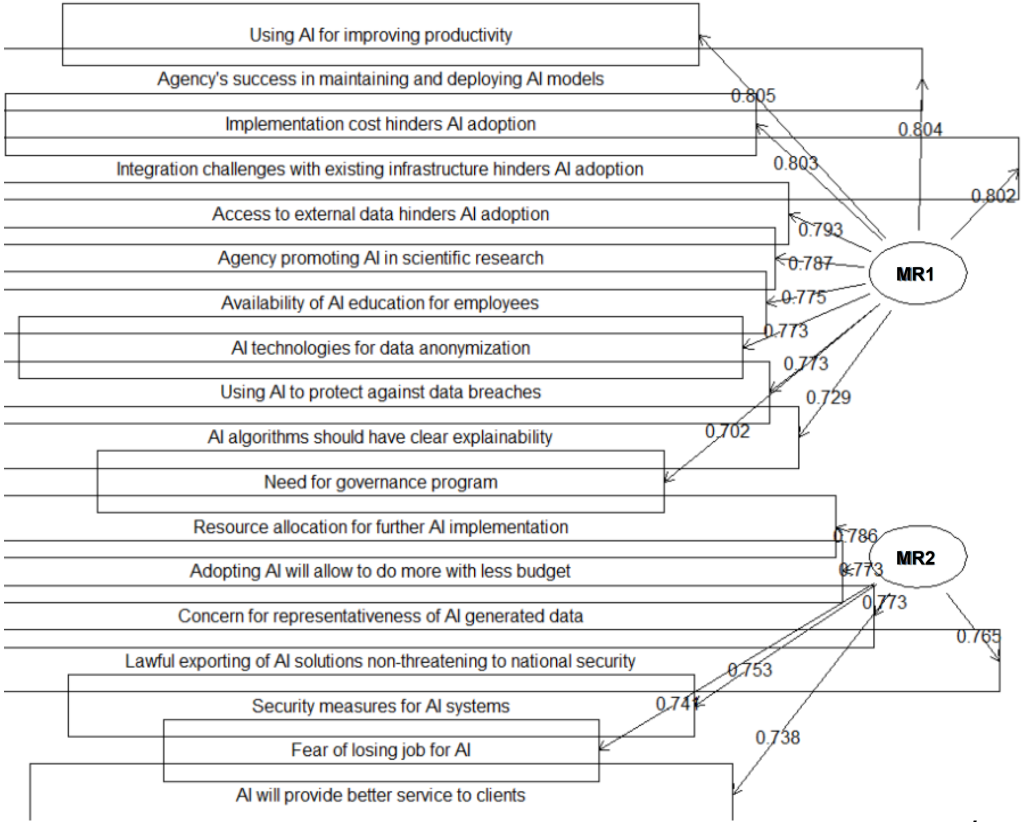
- **Generating Responses**
  - Calculate $P\left(Y_{jk} = l\right),\ l = 1(1)5$
    - Estimate $\widehat{\mu_l}, \hat{\beta}, \widehat{\sigma_\gamma^2}, \widehat{\sigma_\lambda^2}$ by fitting the model on FEVS data
    - Generate $p_{jkl} = expit(\widehat{\mu_l} + X'\hat{\beta} + \widehat{\gamma_j} + \widehat{\lambda_k})$
    - Assign $Y_{jk} = l$, if $p_{jkl} > p_{jkl'}$ for all $l \neq l'$

Topics to Cover

- Background and Motivation
- Research Objective
- Data and Methods
- Results
- Conclusion
- Reference

Exploratory Factor Analysis

- Leveraged polychoric correlation between the questions
- Observed two broad factors capturing adoption of AI – **level of adoption at current time and future**
- Explained **98.5%** of variation in the responses
- Out of 18 questions, **11** were grouped under **current** state and **7** under **future**

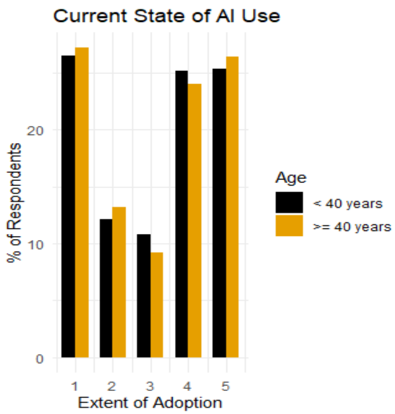


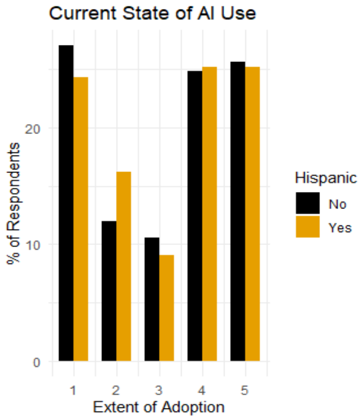
Data Analysis (1)

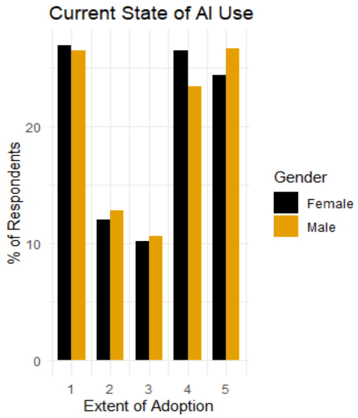
- **Descriptive Analysis**
  - Examined response distribution by demographic variables and generated factors
  - Investigate whether supervisory status and time spent in Federal Govt. impact opinion on factors hindering AI adoption
- **Statistical Analysis**
  - Fit cumulative logit models to measure the extent of association between two or more questions
    - **Hypothesis:** Respondents thinking that 'adopting AI will provide better services to clients' will be prone to saying that their agencies will allocate more resources towards further implementing AI.

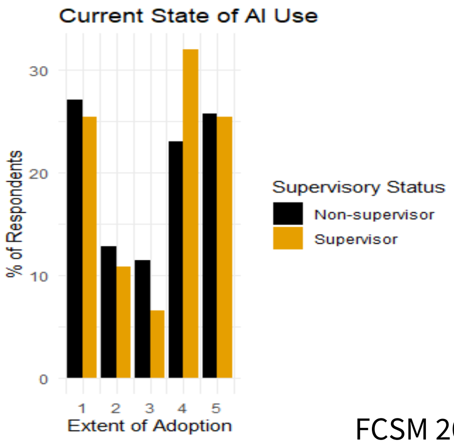
Descriptive Plots – Current State of AI Adoption

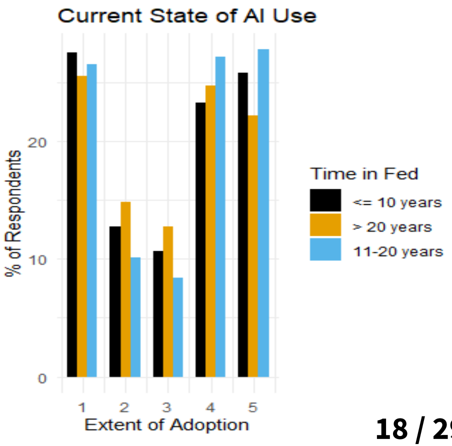
- Most of the respondents are clustered at the extremes – either “not at all” likely or “very much” and “extremely” likely to adopt AI currently
- None of the groups for each demographic variables show any statistically significant (p-values > 0.1) difference across the levels of adoption
- People who spent 11-20 years at their jobs show a different trend in perceived adoption compared to the other groups





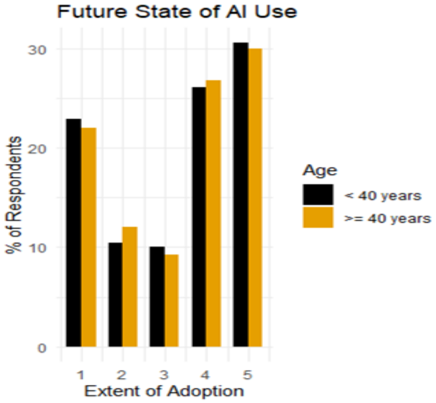


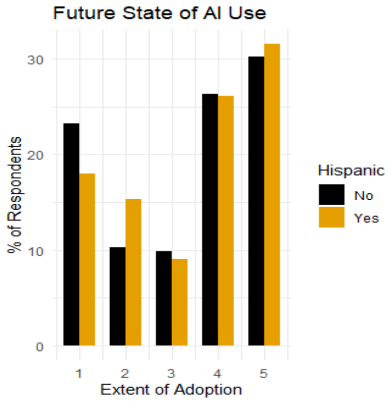


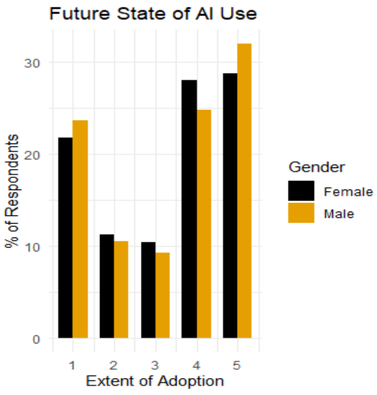


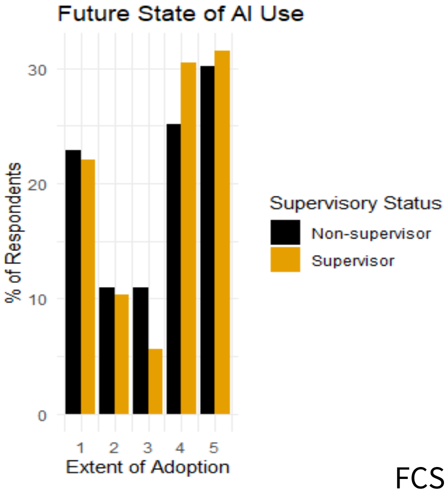
Descriptive Plots- Future State of AI Adoption

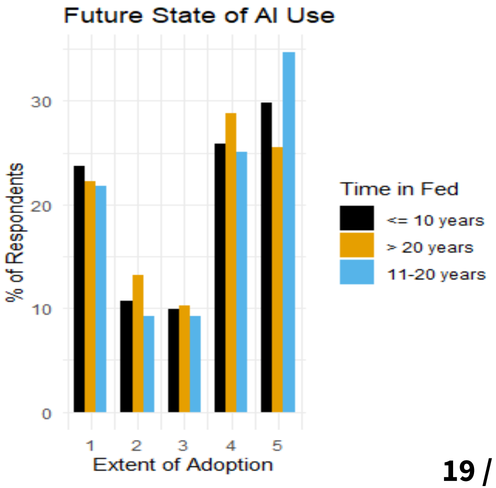
Irrespective of the demographic groups, overall respondents are extremely optimistic to future adoption of AI.







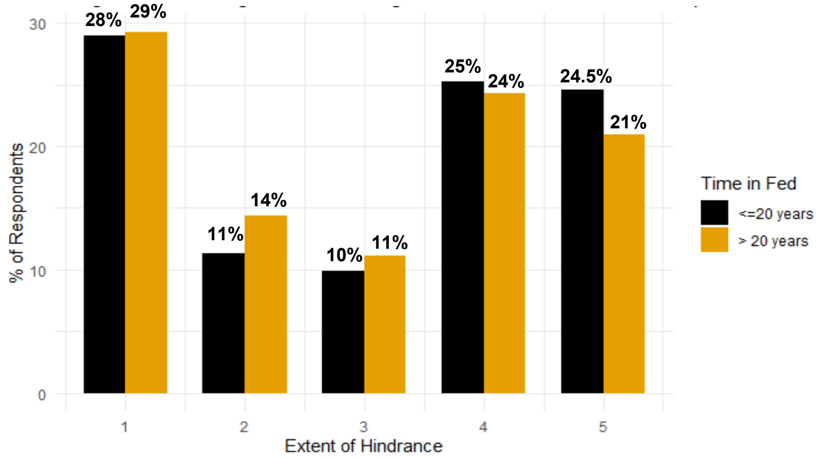


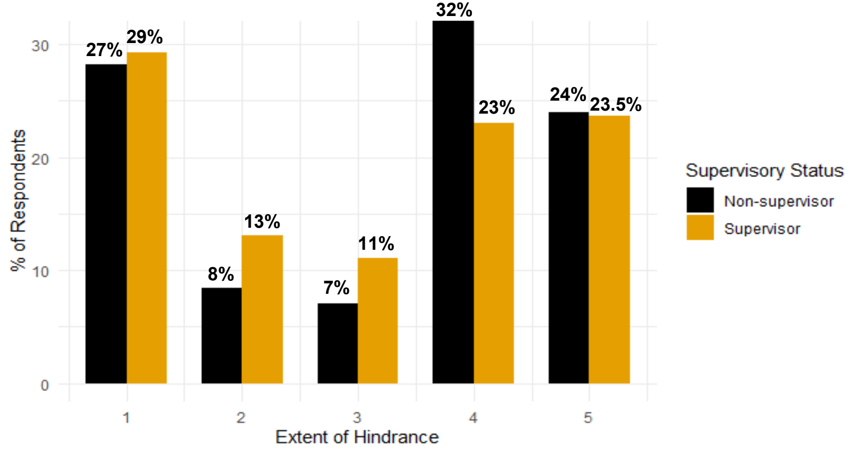


UNIVERSITY OF MARYLAND FEARLESSLY FORWARD

Hindrance to AI Adoption

- Hindrance factors of interest: “access to external data”, “implementation cost”, “integration challenges to existing infrastructure”
- Most employees, regardless of tenure in the Federal Government, do not view these factors as a barrier to AI adoption.
- Non-supervisors view the same factors as a significant threat, while most supervisors believe these won't hinder AI adoption.





UNIVERSITY OF MARYLAND
FEARLESSLY FORWARD

Data Analysis (2)

- **Descriptive Analysis**
  - Examined response distribution by demographic variables and generated factors
  - Investigate whether supervisory status and time spent in Federal Govt. impact opinion on factors hindering AI adoption
- **Statistical Analysis**
  - Fit cumulative logit models to measure the extent of association between two or more questions
    - **Hypothesis**: Respondents thinking that ‘adopting AI will provide better services to clients’ will be prone to saying that their agencies will allocate more resources towards further implementing AI.

Statistical Modeling - Overview

- **Baseline Model**
  - Modeled response distribution of the question – “To what extent do you think your agency **intends to allocate resources** towards the implementation of AI solutions?” (Y) by demographic profiles (age, gender, ethnicity, supervisory status, time spent in Federal Govt.) of the respondents
  - $logit(p_{jkl}) = \mu_l + \boldsymbol{X}'\boldsymbol{\beta} + \lambda_k$, where $p_{jkl} = P(Y_k \leq l), l = 1(1)5, \boldsymbol{X}^*$ is the model matrix and λ_k is the random effect for k^{th} respondent; $\lambda_k \sim N(0, \sigma_\lambda^2)$
- **Model 1 (Baseline + Predictor)**
  - Observed any additional changes in the response distribution of the same question by including responses to the question “To what extent do you think adopting AI solutions will allow your agency to **offer better services to your clients**?” as predictor

* satisfies proportional odds assumption

Statistical Modeling - Intercepts

Baseline Model:

Probability of choosing response option l^*; $l = 1, 2, \ldots, 5$	Estimates	Standard Error**
p_1	0.168	3.35×10^{-5}
p_2	0.089	4.57×10^{-5}
p_3	0.106	1.05×10^{-2}
p_4	0.274	1.63×10^{-2}
p_5	0.363	-

*$l = 1,2,3,4,5$ indicate the following scale of response: “not at all”, “to some extent”, “moderately”, “very much” and “extremely”.

** Calculated using Delta Method

Model 1:

Probability of choosing response option l^{***}; $l = 0, 1, 2$	Estimates	Standard Error
p_0	0.79	7.06×10^{-2}
p_1	0.21	-
p_2	0	-

***Outcome was recoded into three categories: “not so likely" (combining "not at all" and "to some extent"), “moderately," and “extremely likely" (combining "very much" and "extremely") to address power issues in the model.

Statistical Modeling – Slope and Random Effects

Estimates: (In terms of odds ratio)

Parameters	Control Group	Baseline	Model 1
β_{hisp}	Non-hispanic	0.083	0.614
β_{age}	Less than 40 years	-0.032	-0.086
$\beta_{Time\ in\ Fed(11-20)}$	$\leq$ 10 years in Fed	0.22	-0.174
$\beta_{Time\ in\ Fed(>20)}$	$\leq$ 10 years in Fed	0.027	0.488
β_{Sup}	Non-supervisors	-0.103	0.097
β_{Sex}	Female	0.036	0.039
$\beta_{Better\ Service(Moderately)}$	“Not so likely”**	-	27.529
$\beta_{Better\ Service(Extremely)}$	“Not so likely”**	-	51.211
σ_{λ}^{2}		0.093	0.035

**Predictor was recoded into three categories: "Not so likely" (combining "not at all" and "to some extent"), "Moderately," and "Extremely likely" (combining "very much" and "extremely") to address power issues in the model.

Topics to Cover

- Background and Motivation
- Research Objective
- Data and Methods
- Results
- Conclusion
- Reference

Conclusion

Future plan include:

- Analyzing more constructs regarding impact of AI in Federal Government
- Measuring views regarding AI adoption based on different job functions (decision-makers, implementers and users)
- Statistical modeling of certain constructs of interest to explore the potential drivers of those
- Categorizing agencies based on the extent of AI adaptation

Topics to Cover

- Background and Motivation
- Research Objective
- Data and Methods
- Results
- Conclusion
- Reference

Reference

1. Ariga, T., Persons, T. M., & Sanford, S. (2021). Artificial intelligence:an account-ability framework for federal agencies and other entities (tech. rep.). U.S.Government Accountability Office.
2. Brynjolfsson, E., & Mitchell, T. (2017). What can machine learning do? workforceimplications. Science.
3. Butterworth, M. (2018). The ICO and artificial intelligence: The role of fairness inthe gdpr framework. Computer Law and Security Review, 34(2).
4. Čerka, P., Grigienė, J., & Gintarė, S. (2017). Is it possible to grant legal personality toartificial intelligence software systems? Computer Law & Security Review,33(5).
5. Center for Global Business (2024). AI Governance Project Final Report (tech. rep.). Robert H. Smith School of Business, University of Maryland.
6. Faraj, S., Pachidi, S., & Sayeg, K. (2018). Working and organizing in the age of thelearning algorithm. Information and Organization.
7. Gomes de Sousaa, W., Pereira de Meloa, E. R., De Souza Bermejoa, P. H., SousaFariasa, R. A., & Gomesa, A. O. (2019). How and where is artificial intelligence in the public sector going? a literature review and research agenda.Government Information Quarterly.
8. GWOF. (2023). Global workforce of the future report (tech. rep.).
9. The AdeccoGroup.Keegan, M. J. (2020). More than meets AI: Assessing the impact of artificial intelli-gence on the work of government (tech. rep.). IBM Center for The Businessof Government.Microsoft. (2023). Will AI fix work? (Tech. rep.). Microsoft.
10. Tamayo, J., Doumi, L., Goel, S., kovács-ondrejkovic orsolya, o., & Sadun, R. (2023).Reskilling in the age of AI. Harvard Business Review.
11. UNESCO. (2021). Unesco's recommendation on the ethics of artificial intelligence(tech. rep.). UNESCO.

Thank You!

ujstat@umd.edu
smitra98@umd.edu

Take the Survey Here!