Augmented Intelligence: Effects of AI on Productivity and Work Practices

Erik Brynjolfsson[1] and Lindsey Raymond[2]

[1]Stanford University

[2]MIT

ASSAs January 2022

AI and codification of knowledge

- Knowledge is an essential input into production (Gibbons and Henderson, 2012)

AI and codification of knowledge

- Knowledge is an essential input into production (Gibbons and Henderson, 2012)
- This knowledge is often embedded in people and difficult to articulate or transfer (Jensen and Meckling, 1976; Garicano, 2000)

AI and codification of knowledge

- Knowledge is an essential input into production (Gibbons and Henderson, 2012)
- This knowledge is often embedded in people and difficult to articulate or transfer (Jensen and Meckling, 1976; Garicano, 2000)
- Organizations and hierarchies exist to balance the costs of knowledge acquisition and communication

AI and codification of knowledge

- Knowledge is an essential input into production (Gibbons and Henderson, 2012)
- This knowledge is often embedded in people and difficult to articulate or transfer (Jensen and Meckling, 1976; Garicano, 2000)
- Organizations and hierarchies exist to balance the costs of knowledge acquisition and communication
- Machine learning, a branch of artificial intelligence, has made progress in codification of tacit knowledge
  - ML-driven AI learns by example not by instruction (Brynjolfsson and Mitchell, 2017; Agrawal, Gans and Goldfarb, 2019)

AI and codification of knowledge

- Knowledge is an essential input into production (Gibbons and Henderson, 2012)
- This knowledge is often embedded in people and difficult to articulate or transfer (Jensen and Meckling, 1976; Garicano, 2000)
- Organizations and hierarchies exist to balance the costs of knowledge acquisition and communication
- Machine learning, a branch of artificial intelligence, has made progress in codification of tacit knowledge
  - ML-driven AI learns by example not by instruction (Brynjolfsson and Mitchell, 2017; Agrawal, Gans and Goldfarb, 2019)

AI and codification of knowledge

- Knowledge is an essential input into production (Gibbons and Henderson, 2012)
- This knowledge is often embedded in people and difficult to articulate or transfer (Jensen and Meckling, 1976; Garicano, 2000)
- Organizations and hierarchies exist to balance the costs of knowledge acquisition and communication
- Machine learning, a branch of artificial intelligence, has made progress in codification of tacit knowledge
    - ML-driven AI learns by example not by instruction (Brynjolfsson and Mitchell, 2017; Agrawal, Gans and Goldfarb, 2019)

This paper: How does progress in ability of machines to codify knowledge affect productivity, production and organization?

This paper

- Research Question: How does AI affect knowledge diffusion and organization of production?
- Setting: Fortune 500 enterprise software tech support
- Technology: Conversational intelligence decision support software or "intelligence augmentation" that **augments** workers by offering suggestions
- Study Design: Experiment + Staggered deployment
- Data: 4 million conversations, 3,000 agents, 140 teams, 5 firms and 7 countries
- Outcomes: Conversation, agent, team and organization level

Preview of results

1. AI increases efficiency and productivity
   - Tool increases efficiency by decreasing average chat duration by 5 to 10 percent and increasing productivity or issue resolution rate/customer satisfaction by 6 to 8 percent

Preview of results

1. AI increases efficiency and productivity
   - Tool increases efficiency by decreasing average chat duration by 5 to 10 percent and increasing productivity or issue resolution rate/customer satisfaction by 6 to 8 percent
2. Less skilled workers disproportionately benefit
   - Less skilled workers increase productivity by 11 to 13 percent and efficiency by 6 percent relative to higher skill workers

Preview of results

1. AI increases efficiency and productivity
   - Tool increases efficiency by decreasing average chat duration by 5 to 10 percent and increasing productivity or issue resolution rate/customer satisfaction by 6 to 8 percent
2. Less skilled workers disproportionately benefit
   - Less skilled workers increase productivity by 11 to 13 percent and efficiency by 6 percent relative to higher skill workers
3. Driven by codification of tacit knowledge
   - Chat text-based evidence of tacit knowledge

Preview of results

1. AI increases efficiency and productivity
    - Tool increases efficiency by decreasing average chat duration by 5 to 10 percent and increasing productivity or issue resolution rate/customer satisfaction by 6 to 8 percent
2. Less skilled workers disproportionately benefit
    - Less skilled workers increase productivity by 11 to 13 percent and efficiency by 6 percent relative to higher skill workers
3. Driven by codification of tacit knowledge
    - Chat text-based evidence of tacit knowledge
4. Organization level changes in agent specialization and managerial span of control
    - Increases in the breadth of technical support issues handled by each worker and the managerial span of control

Roadmap

Setting

Our firm

- Fortune 500 enterprise software firm ("SoftwareCo")
- Chat based technical support

Setting

Our firm

- Fortune 500 enterprise software firm ("SoftwareCo")
- Chat based technical support

Job Characteristics

- Requires customer service, problem solving skills and product and process knowledge
- Average chat is 48 minutes long summary stats
- On-the-job training from manager

Conversational AI

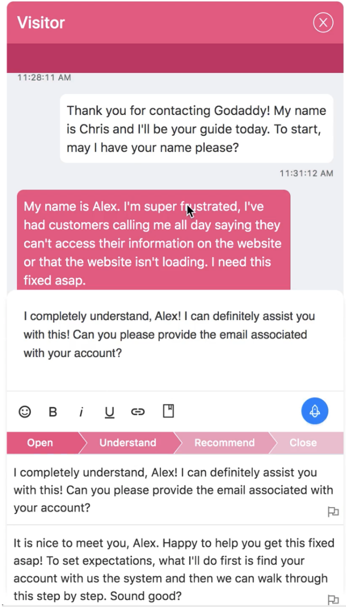
Early Adoption of AI in Call Centers

- Customer service is one of the top two use cases for AI (McKinsey, 2020)
- Augmentation (decision support) rather than automation (Canam Research, 2020)

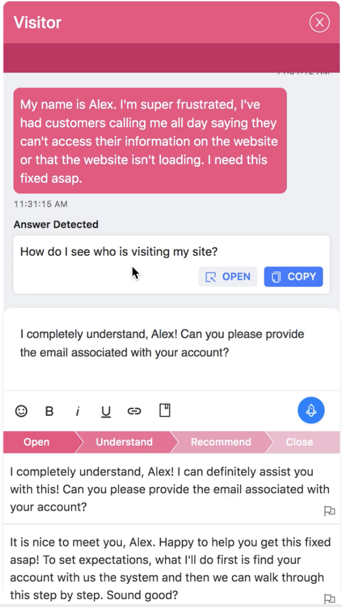
Conversational AI

- AI learns from all current and historical agent-customer interactions
- Customer-agent interactions create the training data set for semi-supervised learning
- AI system makes suggestions to agents

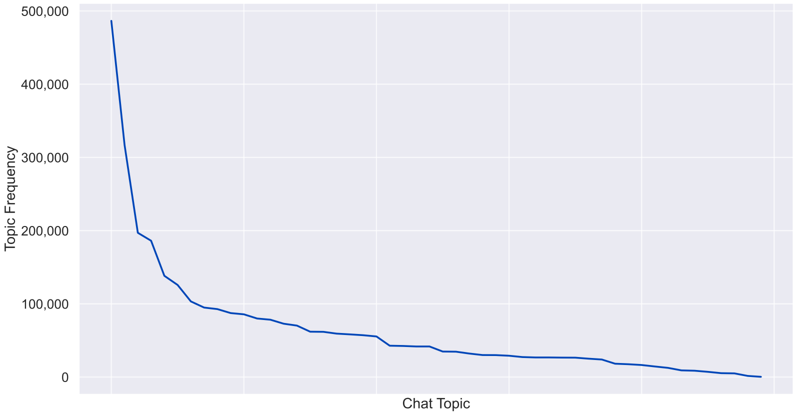
AI offers real-time suggestions on what to say



AI offers real-time suggestions on problem and solution



Small number of common problems with a long tail



Three model predictions

1. Less skilled agents follow AI suggestions more than higher skilled agents
   - Higher skill agents already have tacit knowledge and don't need assistance while less skilled agents need help

Three model predictions

1. Less skilled agents follow AI suggestions more than higher skilled agents
   - Higher skill agents already have tacit knowledge and don't need assistance while less skilled agents need help

2. Convergence in productivity levels
   - Lower skill agents improve **more** than higher skilled workers

Three model predictions

1. Less skilled agents follow AI suggestions more than higher skilled agents
   - Higher skill agents already have tacit knowledge and don't need assistance while less skilled agents need help

2. Convergence in productivity levels
   - Lower skill agents improve **more** than higher skilled workers

3. Larger effects where diffusion of tacit information is slower
   - Tacit knowledge more likely to diffuse more slowly across;
     - Firm boundaries
     - Cultural boundaries
     - Larger teams

Experiment

- Seven week randomized control trial
- 50 agents across treatment and control
- Treatment group matched to control on issue resolution rates and chat duration

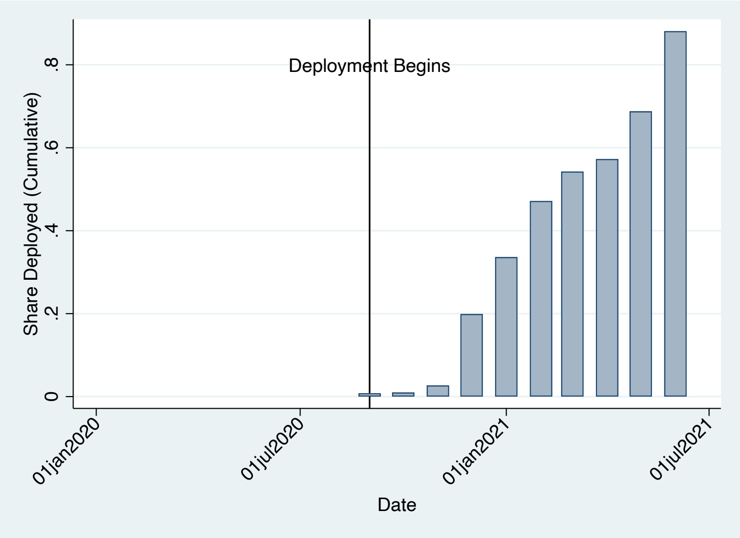
Group	Baseline Issue Resolution Rate	Baseline Average Chat Duration
Treatment	84.5	43.2
Control	83.8	42.7
Difference	**0.65**	**1.63**

Experiment

- At end of the RCT, treatment group had **two** times the increase in issue resolution rates and drop in average call duration

Group	Change in Issue Resolution Rate	Change in Chat Duration
Treatment	3.26 percentage points	-7.15 minutes
Control	1.90 percentage points	-4.87 minutes

Staggered adoption



Performance effects of AI deployment

$$Y_{i,t,e} = \alpha_{i,e} + \gamma_{t,e} + \beta_1(Treated_{i,e}xPostDeployment_{t,e}) + \epsilon_{i,t,e}$$

	(1) Average Call Duration	(2) Std Dev. Call Duration	(3) Issue Resolution Rate	(4) Customer Satisfaction
TreatedxPostDeployment	-0.049***	-0.012***		
	(0.002)	(0.002)		
Time FE	MonthxExperiment	MonthxExperiment		
Unit FE	AgentxExperiment	AgentxExperiment		
Observations	1,153,458	1,110,679		
Pre Mean	48	40		

Performance effects of AI deployment

$$Y_{i,t,e} = \alpha_{i,e} + \gamma_{t,e} + \beta_1(Treated_{i,e}xPostDeployment_{t,e}) + \epsilon_{i,t,e}$$

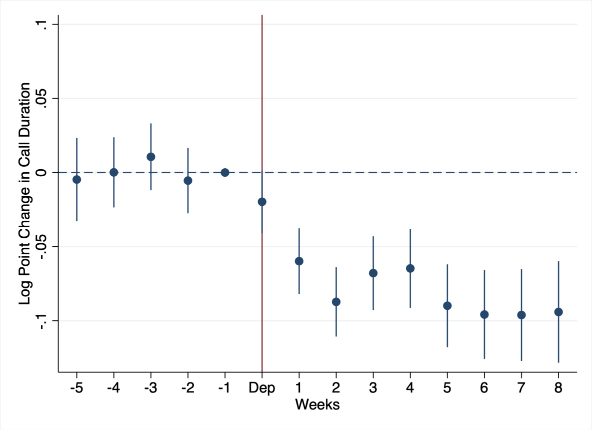
	(1) Average Call Duration	(2) Std Dev. Call Duration	(3) Issue Resolution Rate	(4) Customer Satisfaction
TreatedxPostDeployment	-0.049***	-0.012***	0.062***	0.075***
	(0.002)	(0.002)	(0.014)	(0.017)
Time FE	MonthxExperiment	MonthxExperiment	MonthxExperiment	MonthxExperiment
Unit FE	AgentxExperiment	AgentxExperiment	AgentxExperiment	AgentxExperiment
Observations	1,153,458	1,110,679	57,891	58,981
Pre Mean	48	40	80	60

Change in outcome after deployment for treated agents;

1. 5% drop in average call duration
2. 1% decrease in standard deviation of call duration
3. 6% increase in issue resolution rate
4. 8% increase in customer satisfaction

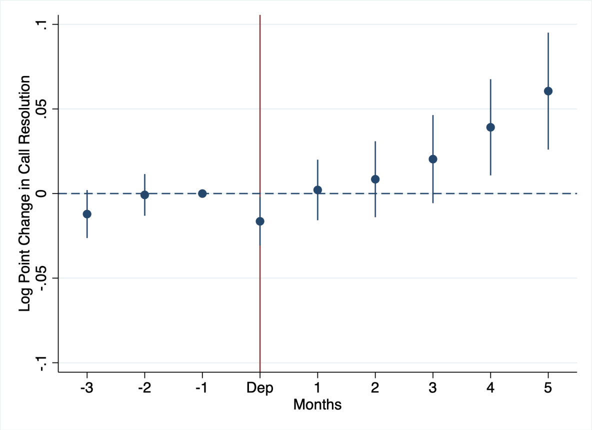
Stacked event study of AI deployment on call duration

$$Y_{i,t,e} = \alpha_{i,e} + \sum_{\tau} D^{\tau}_{t,e} + \sum_{\tau} \beta_{\tau} (Treated_{i,e} x D^{\tau}_{t,e}) + \epsilon_{i,t,e}$$



Stacked event study of AI deployment on issue resolution rate

$$Y_{i,t,e} = \alpha_{i,e} + \sum_{\tau} D^{\tau}_{t,e} + \sum_{\tau} \beta_{\tau}(Treated_{i,e} x D^{\tau}_{t,e}) + \epsilon_{i,t,e}$$



Lower ability agents benefit the most

- Low skill agents follow AI recommendations 20% more than high skill ability workers

	(1) Average Call Duration	(2) Std Dev. Call Duration	(3) Call Resolution Rate	(4) Customer Satisfaction
Post Deployment x Low	-0.055***	-0.049***		
	(0.005)	(0.006)		
Time FE	MonthxExperiment	MonthxExperiment		
Unit FE	AgentxExperiment	AgentxExperiment		
Observations	570,166	559,449		
Pre Mean	48	40		

Lower ability agents benefit the most

- Low skill agents follow AI recommendations 20% more than high skill ability workers

	(1) Average Call Duration	(2) Std Dev. Call Duration	(3) Call Resolution Rate	(4) Customer Satisfaction
Post Deployment x Low	-0.055***	-0.049***	0.127***	0.109***
	(0.005)	(0.006)	(0.010)	(0.013)
Time FE	MonthxExperiment	MonthxExperiment	MonthxExperiment	MonthxExperiment
Unit FE	AgentxExperiment	AgentxExperiment	AgentxExperiment	AgentxExperiment
Observations	570,166	559,449	35,332	34,580
Pre Mean	48	40	80	60

Change in outcome after deployment for lower skill agents relative to high ability agents;

1. 6% greater drop in average call duration
2. 5% greater decrease in consistency of call duration
3. 11% greater increase in issue resolution rate
4. 13% greater increase in customer satisfaction

Codification of previously tacit knowledge

- Previously agents learned from experience and coaching
  - How to ask diagnostic questions
  - Symptoms of common technical problems
  - Mapping of customer description to technical problem

Codification of previously tacit knowledge

- Previously agents learned from experience and coaching
  - How to ask diagnostic questions
  - Symptoms of common technical problems
  - Mapping of customer description to technical problem
- AI can learn from historical interactions **without** requiring each individual agent to learn from experience

Usage of knowledge base articles

	(2) Technical Documents	(3) Technical Documents	Technical Documents
AIxOutsourced	2.403***		
	(0.071)		
AIxLarge Team			
AIxIndia+Phil			
Time FE	MonthxExperiment		
Unit FE	AgentxExperiment		
Observations	89,108		
Pre Mean	12		

Usage of knowledge base articles

	(2) Technical Documents	(3) Technical Documents	Technical Documents
AIxOutsourced	2.403***		
	(0.071)		
AIxLarge Team		0.386***	
		(0.038)	
AIxIndia+Phil			0.628***
			(0.058)
Time FE	MonthxExperiment	MonthxExperiment	MonthxExperiment
Unit FE	AgentxExperiment	AgentxExperiment	AgentxExperiment
Observations	89,108	74,296	89,108
Pre Mean	12	12	12

Codification of tacit knowledge and problem solving rates

	(1) Issue Resolution Rate	(2) Issue Resolution Rate	(3) Issue Resolution Rate
AIxOutsourced	0.129**		
	(0.052)		
AIxLarge Team			
AIxIndia+Phil			
Time FE	MonthxExperiment		
Unit FE	AgentxExperiment		
Observations	57,891		
Pre Mean	80		

Codification of tacit knowledge and problem solving rates

	(1) Issue Resolution Rate	(2) Issue Resolution Rate	(3) Issue Resolution Rate
AIxOutsourced	0.129**		
	(0.052)		
AIxLarge Team		0.061***	
		(0.019)	
AIxIndia+Phil			0.032*
			(0.018)
Time FE	MonthxExperiment	MonthxExperiment	MonthxExperiment
Unit FE	AgentxExperiment	AgentxExperiment	AgentxExperiment
Observations	57,891	57,891	55,395
Pre Mean	80	80	80

Larger teams, less specialization

	(1) Team Size	(2) Skill Breadth
AI Deployment	**0.068*****	
	(0.005)	
Team Deployment		
Time FE	MonthxExperiment	
Unit FE	TeamxExperiment	
Observations	3,042	
Pre Mean	20	

Larger teams, less specialization

	(1) Team Size	(2) Skill Breadth
AI Deployment	**0.068***** **(0.005)**	
Team Deployment		**0.073**** **(0.031)**
Time FE	MonthxExperiment	MonthxExperiment
Unit FE	TeamxExperiment	AgentxExperiment
Observations	3,042	85,300
Pre Mean	20	4

Summary of Main Results

- AI has potential to change nature and organization of work in ways that may affect all workers whether or not directly interacting with AI
  1. AI can increase efficiency and productivity
  2. Less skilled workers disproportionately benefit
  3. Tacit knowledge codification
  4. Reduces specialization, increased managerial span of control

- Distinct types of AI will have different effects
  - On individuals: augmentation versus automation
  - On organizations: natural of specialization and span of control

Open Questions

- Implications for the theory of the firm (e.g. Garicano, 2000)
- Implications for the labor market (e.g. skill biased technological change)
- Implications for organizations (e.g. Sah and Stiglitz, 1986)

Appendix

Stacked Differences in Differences

- Stacked differences in differences are robust to biases from negative weighting and heterogeneous treatment effects across groups or over time (Cengiz et al., (2019); Baker et al., (2021); Deshpande and Li, 2019)

$$Y_{i,t,e} = \alpha_{i,e} + \gamma_{t,e} + \beta_1(T_{i,e}xP_{t,e}) + \epsilon_{i,t,e}$$

- $Y_{i,t,e}$ is the outcome for agent i at time t in sub-experiment e where each sub-experiment is a individual deployment
- $\alpha_{i,e}$ are agent by sub-experiment fixed effects
- $\gamma_{t,e}$ are time by sub-experiment fixed effects
- $T_{i,e}$ are the treatment agents in each sub-experiment
- $P_{t,e}$ is a dummy for when AI is deployed in sub-experiment e
- Standard errors are clustered at the sub-experiment by agent level

Stacked Event Study

- Stacked event studies are robust to biases from negative weighting and heterogeneous treatment effects (Cengiz et al., (2019); Baker et al., (2021); Deshpande and Li, 2019)

$$Y_{i,t,e} = \alpha_{i,e} + \sum_{\tau} D^{\tau}_{t,e} + \sum_{\tau} \beta_{\tau}(T_{i,e} x D^{\tau}_{t,e}) + \epsilon_{i,t,e}$$

- $Y_{i,t,e}$ is the outcome for agent i at time t in sub-experiment e where each sub-experiment is a individual deployment
- $\alpha_{i,e}$ are agent by sub-experiment fixed effects
- $T_{i,e}$ are the treatment agents in each sub-experiment
- $D^{\tau}_{t,e}$ are dummies equal to one if an agent is τ periods away from time of treatment in sub-experiment e
- Standard errors are clustered at the sub-experiment by agent level

Main Effects Regressions

	(1) Average Call Duration	(2) Std Dev. Call Duration	(3) Issue Resolution Rate	(4) Customer Satisfaction
AI Deployment	-0.049***	-0.012***	0.062***	0.075***
	(0.002)	(0.002)	(0.014)	(0.017)
Time FE	MonthxExperiment	MonthxExperiment	MonthxExperiment	MonthxExperiment
Unit FE	AgentxExperiment	AgentxExperiment	AgentxExperiment	AgentxExperiment
Observations	1,153,458	1,110,679	57,891	58,981
Pre Mean	48	40	80	60

Table: AI Deployment on Outcomes

NOTES: This table shows regressions on logged measures of agent performance after AI deployment. Each stacked regression controls for sub-experiment by agent level and sub-experiment by month fixed effects and standard errors are clustered at the agent by sub-experiment level. Standard errors are in the parentheses. Sample includes January 2020 to June of 2021 and all data come from the firm production records. An agent is counted as deployed when trained on the AI tool and agent starts receiving AI output. Standard errors are in the parentheses. The coefficient on "AI Deployment" is the log point change in the outcome after AI deployment.

figure

Results by Agent Skill

	(1) Customer Satisfaction	(2) Call Resolution Rate	(3) Average Call Duration	(4) Std Dev. Call Duration
Post Deployment x Low	0.109***	0.127***	-0.055***	-0.049***
	(0.013)	(0.010)	(0.005)	(0.006)
Time FE	MonthxExperiment	MonthxExperiment	MonthxExperiment	MonthxExperiment
Unit FE	AgentxExperiment	AgentxExperiment	AgentxExperiment	AgentxExperiment
Observations	35,332	34,580	570,166	559,449
Pre Mean	4	4	48	40

Table: Effects by Ex-Ante Agent Skill

NOTES: This table shows regressions on logged measures of agent performance after AI deployment. Each stacked regression controls for sub-experiment by agent level and sub-experiment by month fixed effects and standard errors are clustered at the agent by sub-experiment level. Standard errors are in the parentheses. Sample includes January 2020 to June of 2021 and all data come from the firm production records. An agent is counted as deployed when trained on the AI tool and agent starts receiving AI output. The coefficient on "AI Deployment" is the log point change in the outcome after AI deployment for low skill agents who are ranked below average within their company in the past two months prior to AI deployment relative to above average agents.

figure

Sample Summary Statistics

Variable	All	Control Agents	Treated Agents
Chats	3,758,698	374,731	2,635,864
Agents	6,846	1,035	1,813
Number of Teams	142	111	88
Share US Agents	.13	.095	.14
Distinct Locations	17	10	16
Average Chats per Month	158	112	212
Share Outsourced	.84	.62	.91
Number of Skills	2.8	2.3	3.3
Team Size	62	49	70
Average Call Duration (Min)	48	44	48
St. Dev. Call Duration (Min)	40	37	39
Issue Resolution Rate	79	77	81
Customer Satisfaction	62	62	61

Table: Sample Summary Statistics

AI Usage by Skill

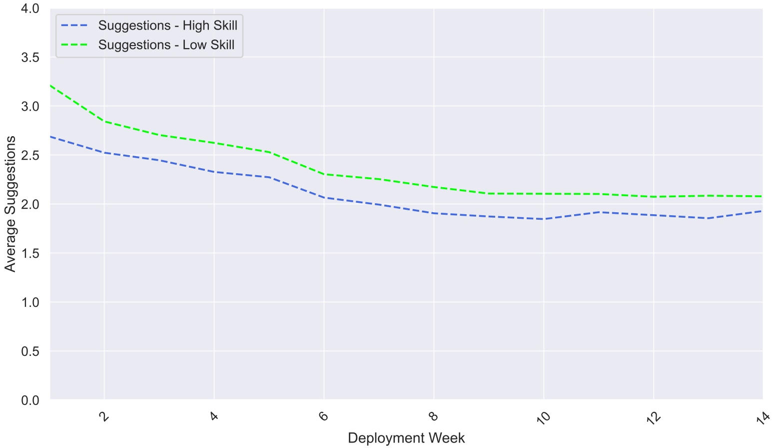


Figure: Use of AI Suggestions by Agent Skill