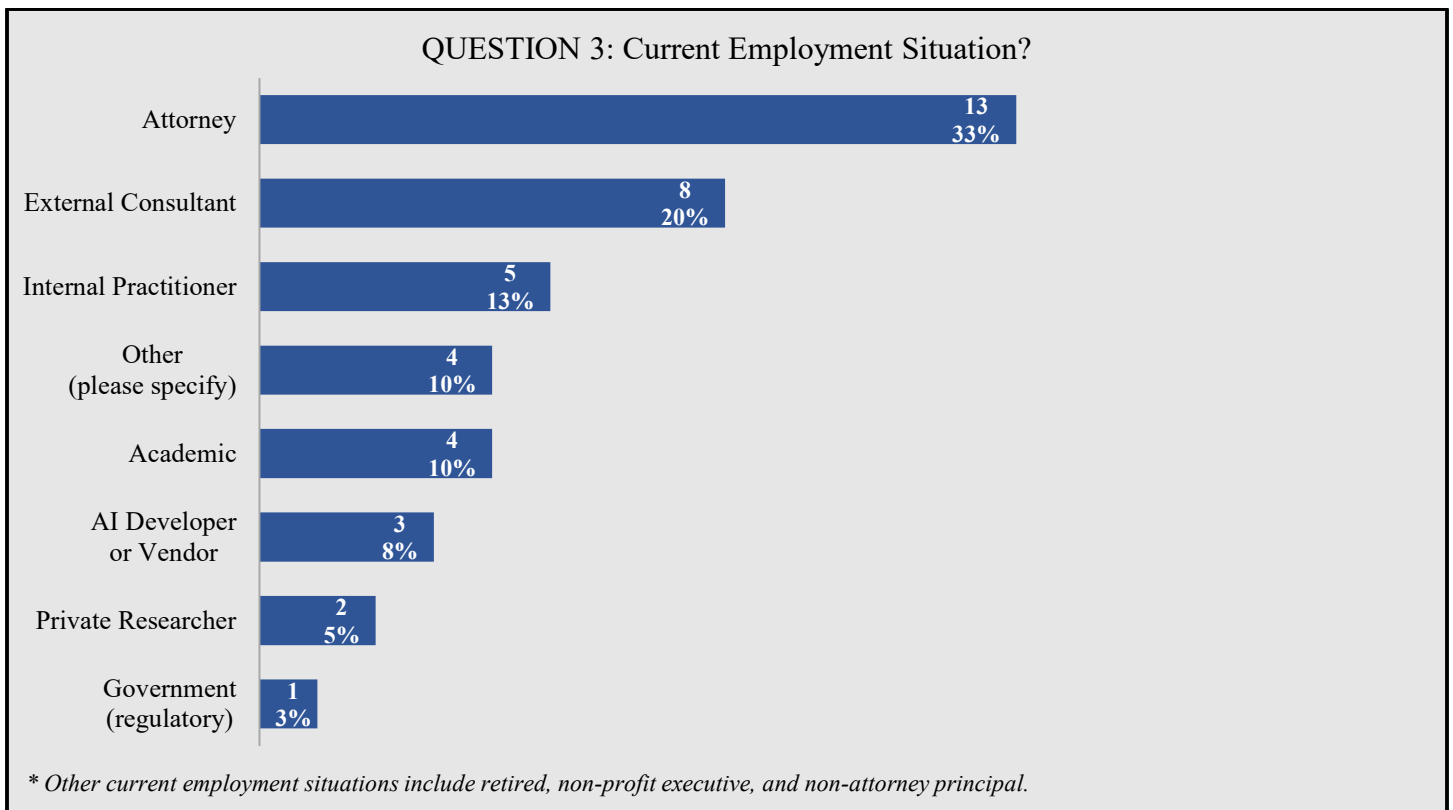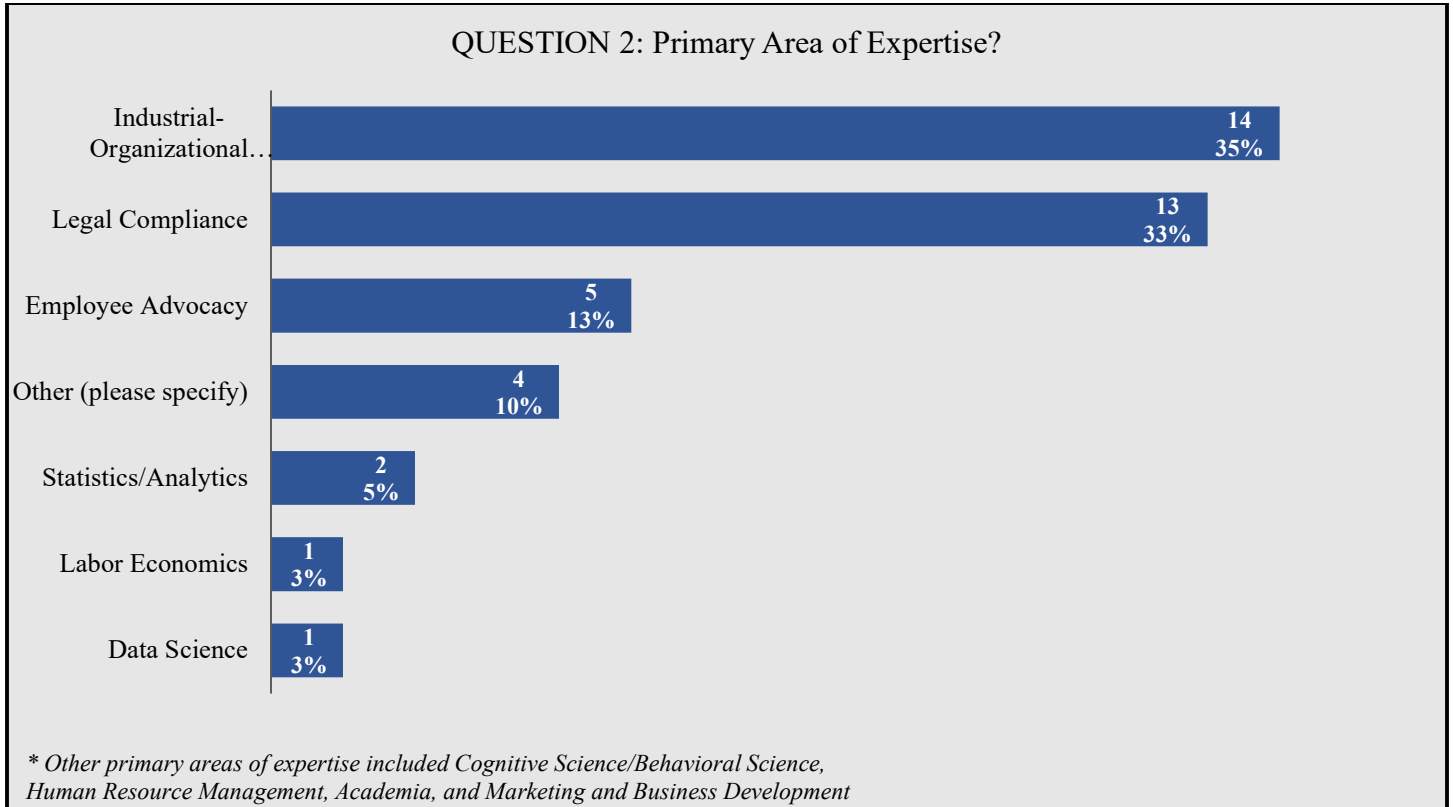# The Institute for WORKPLACE EQUALITY

Survey Results Accompanying Technical Advisory Committee Report on EEO and DEI&A Considerations in the Use of Artificial Intelligence in Employment Decision Making
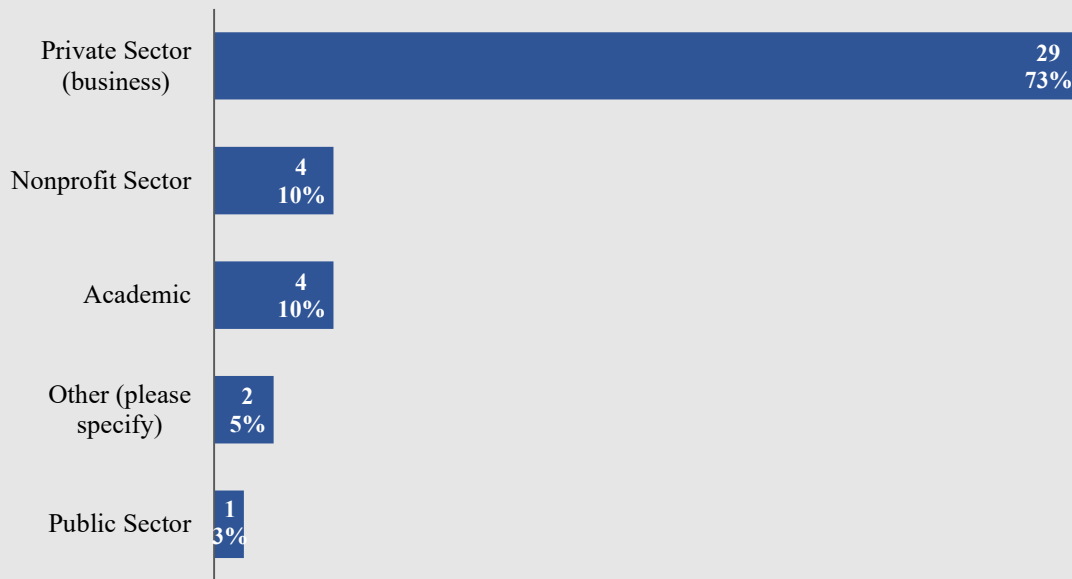
December 2022

AI TAC SURVEY 2022

A NOTE ABOUT THE SURVEY: By design, members of the TAC represent a diverse range of experience with varying degrees of expertise in the wide range of topics related to artificial intelligence and queried about in the survey. The survey contained 11 different content topics, and for 9 of these sections there was a preliminary question asking how comfortable the TAC member was in answering questions on that particular content topic. Because not all TAC members were equally versed in all of the content topics, we believed it was best to afford TAC members an opportunity to opt out of answering a particular set of questions, deferring to allow those TAC members with more experience and expertise to respond. For example, we did not want a TAC member with deep expertise in the legal topics, who works routinely with clients using artificial intelligence, but with limited knowledge of statistics to feel compelled to answer questions in the statistics section—better to skip that section and allow those with expertise in the statistics to address the questions. Further, each close-ended survey question included an option that allowed a TAC member to choose "I do not know" to a question (for those who felt they had expertise to answer some, but not all questions). The intent was to increase the quality of information provided by the survey that would help to more accurately provide the sub-committees with solid foundational information to guide their work.

SECTION 1: TAC MEMBER BACKGROUND AND DEMOGRAPHICS

QUESTION 2: Primary Area of Expertise?

Area	Count	Percent
Industrial-Organizational…	14	35%
Legal Compliance	13	33%
Employee Advocacy	5	13%
Other (please specify)	4	10%
Statistics/Analytics	2	5%
Labor Economics	1	3%
Data Science	1	3%

* Other primary areas of expertise included Cognitive Science/Behavioral Science, Human Resource Management, Academia, and Marketing and Business Development

QUESTION 3: Current Employment Situation?

Situation	Count	Percent
Attorney	13	33%
External Consultant	8	20%
Internal Practitioner	5	13%
Other (please specify)	4	10%
Academic	4	10%
AI Developer or Vendor	3	8%
Private Researcher	2	5%
Government (regulatory)	1	3%

* Other current employment situations include retired, non-profit executive, and non-attorney principal.

QUESTION 4: Work Setting?

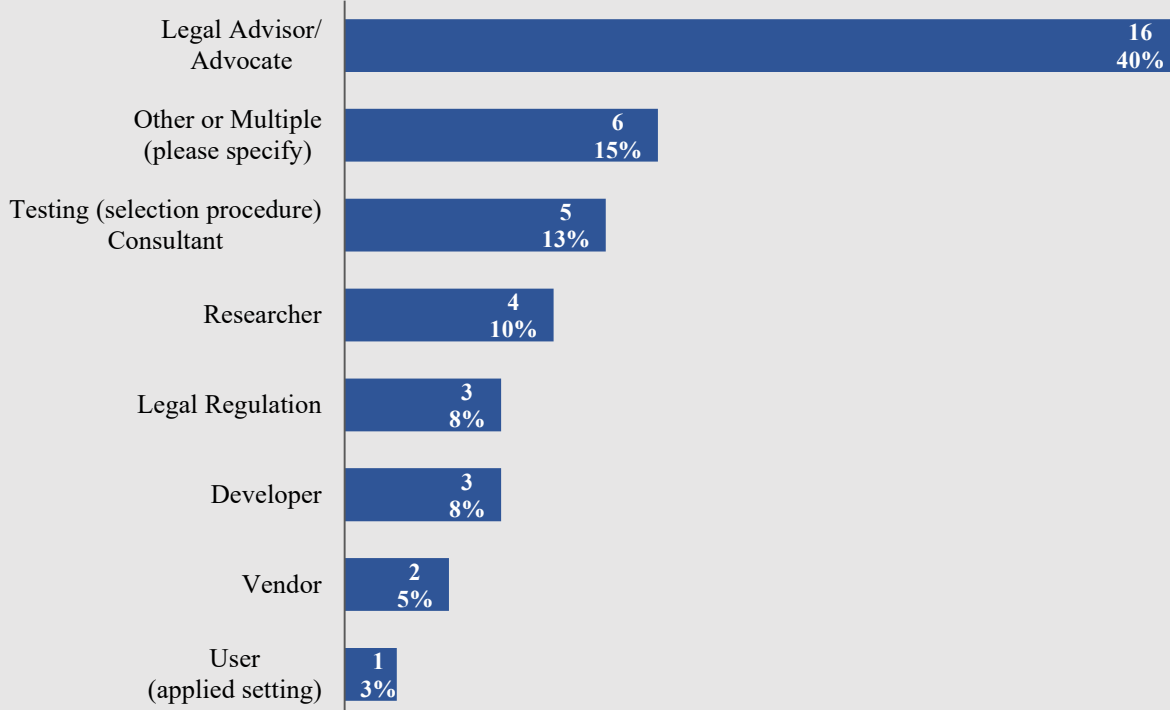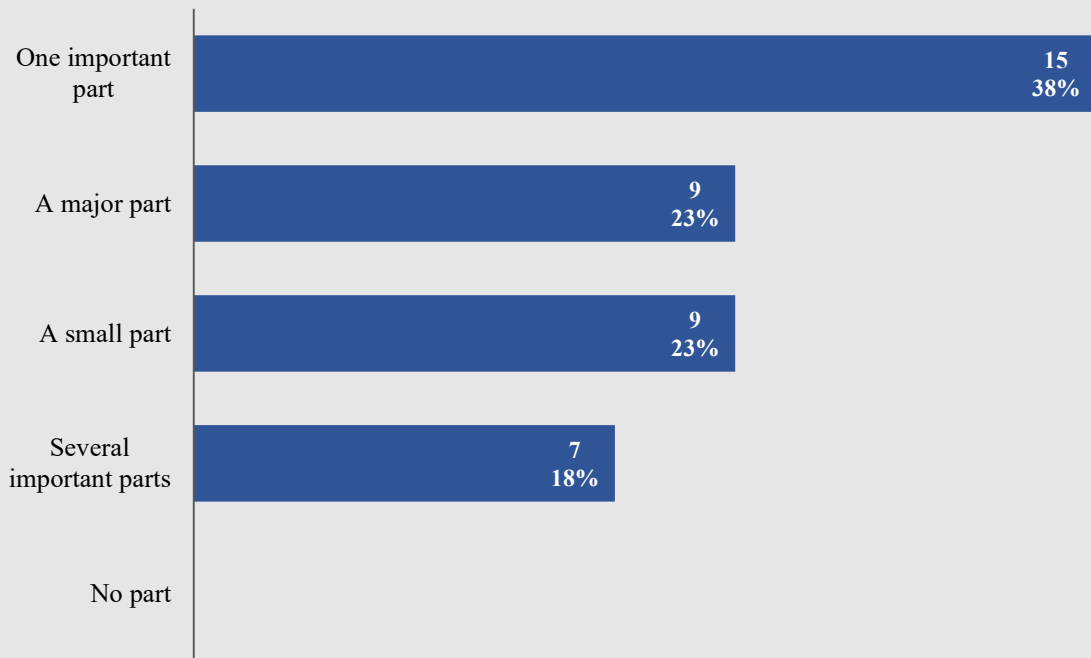| Work Setting	Count	Percent
Private Sector (business)	29	73%
Nonprofit Sector	4	10%
Academic	4	10%
Other (please specify)	2	5%
Public Sector	1	3%

* Other work settings include home and retired.

QUESTION 5: Involvement with Artificial Intelligence (AI) Applications?

Involvement	Count	Percent
Legal Advisor/ Advocate	16	40%
Other or Multiple (please specify)	6	15%
Testing (selection procedure) Consultant	5	13%
Researcher	4	10%
Legal Regulation	3	8%
Developer	3	8%
Vendor	2	5%
User (applied setting)	1	3%

QUESTION 6: Involvement with AI related issues in your professional activities?

Response	Count	Percent
One important part	15	38%
A major part	9	23%
A small part	9	23%
Several important parts	7	18%
No part		

QUESTION 7: Highest Educational Degree?

Degree	Count	Percent
JD	19	48%
PhD	16	40%
MS	2	5%
Other or Multiple -…	1	3%
BS	1	3%
BA	1	3%
PsyD		
MBA		
MA		

* One respondent specified highest Educational Degree as PhD, MBA

QUESTION 8: Ethnicity?

Other	30 — 91%
Hispanic/Latinx	3 — 9%

QUESTION 9: Race?
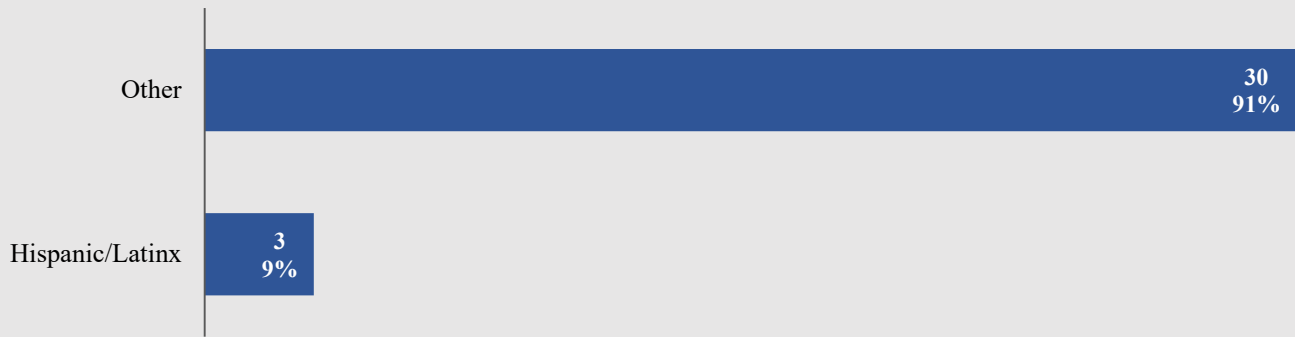
White/Caucasian	31 — 86%
Asian	3 — 8%
Two or more races	1 — 3%
Black/African American	1 — 3%
Other (please specify)	
Pacific Islander	
American Indian/Alaska Native	
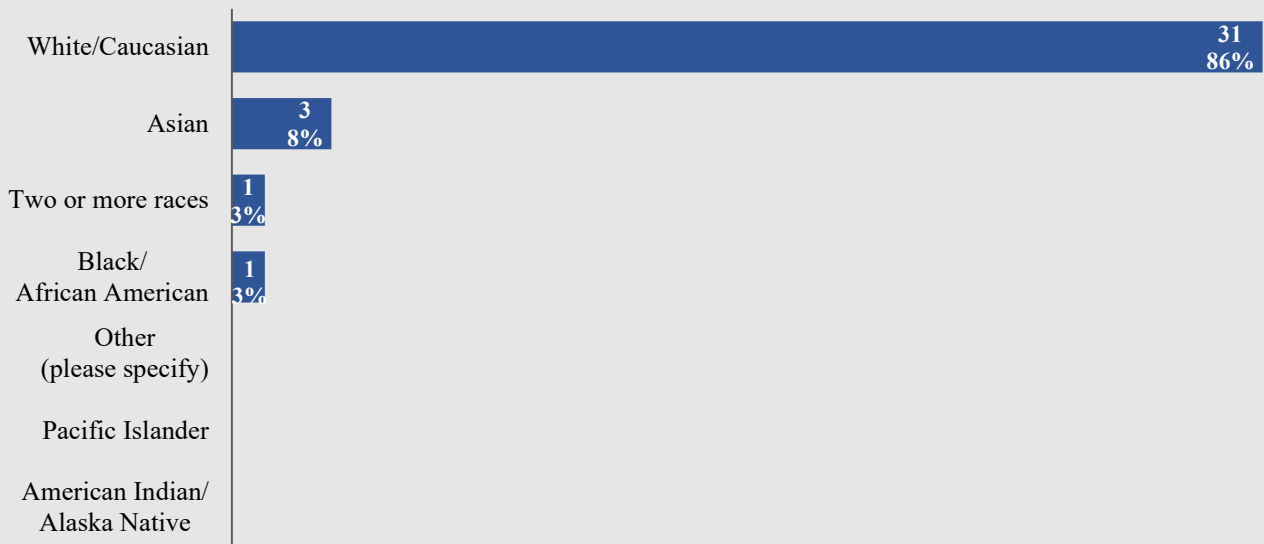
QUESTION 10: Gender?

Male	21 — 58%
Female	15 — 42%
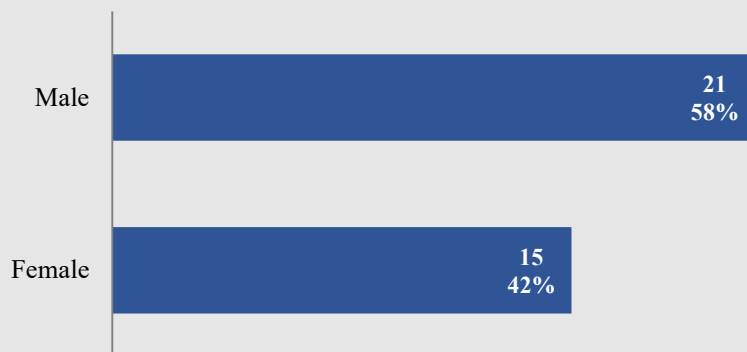
SECTION 2: USES OF ARTIFICIAL INTELLIGENCE (AI) BASED APPLICATION

QUESTION 11: How would you describe your level of familiarity with this topic?

I am generally knowledgeable about this topic – I could probably answer a lot of these questions	17 / 43%
I am an expert on this topic – I expect to be able to answer all of these questions	7 / 18%
I am very familiar with this topic – I should be able to answer all of these questions	7 / 18%
I know a little about this topic – I could potentially answer some of these questions	7 / 18%
I do not know this topic – I could not answer these types of questions	2 / 5%

QUESTION 12: In your experience, how frequently have you encountered each of the following uses of AI-based applications?
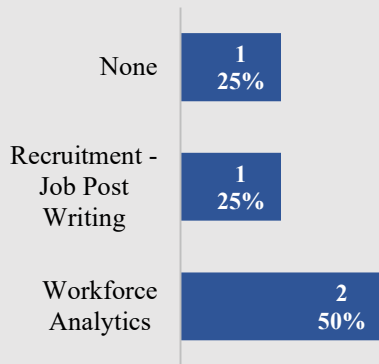
	I do not know	Never or very infrequently *I never or rarely see this AI-based application in my practice*	Infrequently *I sometimes see this AI-based application in my practice*	Frequently *I often see this use of AI in my practice.*	Very frequently *This use of AI is common practice*
Actively identifying applicants (e.g., soliciting applicants, identifying outreach groups) to recruit qualified candidates.	(2)	14.3% (5)	34.3% (12)	37.1% (13)	14.3% (5)
Providing applicant self-screens (e.g., surveys, tests, games) to determine the match between applicant's abilities and/or interests and a particular job or career path.	(1)	19.4% (7)	36.1% (13)	30.6% (11)	13.9% (5)
Narrowing applicant pools to automate the review of qualifications provided by an	(2)	14.3% (5)	28.6% (10)	42.9% (15)	14.3% (5)

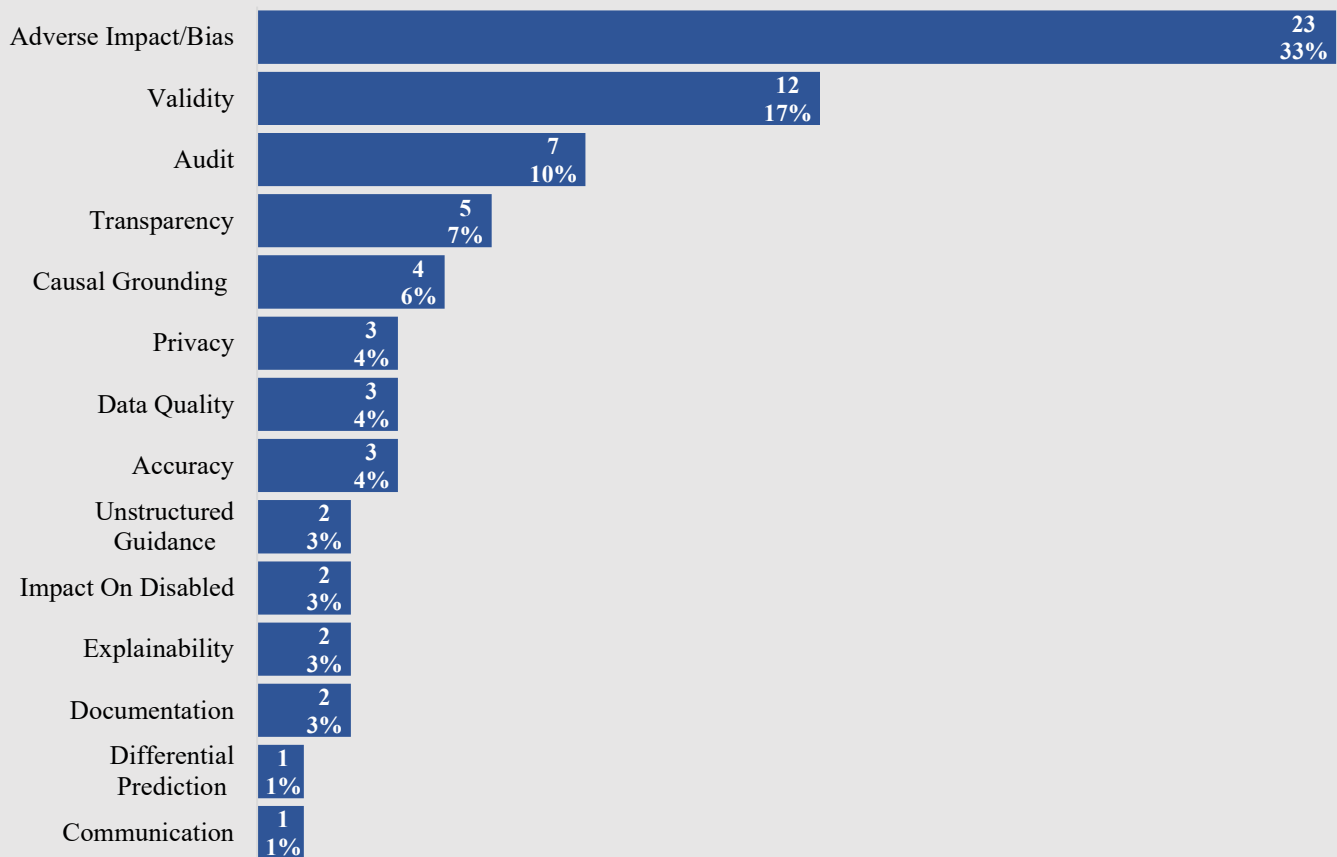| QUESTION 12: In your experience, how frequently have you encountered each of the following uses of AI-based applications?					
	I do not know	**Never or very infrequently** *I never or rarely see this AI-based application in my practice*	**Infrequently** *I sometimes see this AI-based application in my practice*	**Frequently** *I often see this use of AI in my practice.*	**Very frequently** *This use of AI is common practice*
applicant into a structured format (e.g., scoring drop-down options in an online application).					
Narrowing applicant pools to automate the review of qualifications extracted from an applicant's resume or social media profile (e.g., scoring unstructured, open-ended data).	(1)	27.8% (10)	22.2% (8)	36.1% (13)	13.9% (5)
Predicting future job performance or other job-related criteria (e.g., turnover, unsafe work behaviors, organization fit).	(2)	14.3% (5)	37.1% (13)	34.3% (12)	14.3% (5)
Conducting background checks to equate criminal charges across jurisdictions.	(9)	67.9% (19)	21.4% (6)	7.1% (2)	3.6% (1)
Monitoring employee performance.	(6)	45.2% (14)	32.3% (10)	16.1% (5)	6.5% (2)
Evaluating performance outcomes (e.g., customer surveys, team ratings, integrated sales data, potential in assigned sales territory).	(3)	47.1% (16)	26.5% (9)	20.6% (7)	5.9% (2)
Engaging in workforce planning (e.g., estimating where turnover is likely and making decisions on those estimates to increase hiring, reduction in force, or elimination of operations).	(6)	38.7% (12)	41.9% (13)	19.4% (6)	0.0% (0)
Identifying "flight risk" of individual current employees.	(5)	50.0% (16)	28.1% (9)	18.8% (6)	3.1% (1)
Identifying candidates who likely would not stay long with the organization.	(6)	32.3% (10)	38.7% (12)	19.4% (6)	9.7% (3)
Identifying "high potential" employees across the organization.	(4)	33.3% (11)	45.5% (15)	18.2% (6)	3.0% (1)
Conducting targeted employee surveys (e.g., specific groups getting surveys to explore reasons for satisfaction and dissatisfaction).	(5)	46.9% (15)	37.5% (12)	9.4% (3)	6.3% (2)

6

QUESTION 12: In your experience, how frequently have you encountered each of the following uses of AI-based applications?					
	I do not know	**Never or very infrequently** *I never or rarely see this AI-based application in my practice*	**Infrequently** *I sometimes see this AI-based application in my practice*	**Frequently** *I often see this use of AI in my practice.*	**Very frequently** *This use of AI is common practice*
Establishing a talent bank or similar dataset of employee information (e.g., inventory of employee skills, interests used to form teams and make assignments).	(6)	60.0% (18)	16.7% (5)	20.0% (6)	3.3% (1)
Creating learning plans (e.g., interoperable learning records, individual development plans, career coaching plans).	(9)	55.6% (15)	29.6% (8)	11.1% (3)	3.7% (1)
Career pathing.	(7)	58.6% (17)	27.6% (8)	10.3% (3)	3.4% (1)
Collecting job analysis data/materials.	(5)	51.6% (16)	32.3% (10)	12.9% (4)	3.2% (1)
Analyzing job analysis data/materials.	(4)	56.3% (18)	21.9% (7)	15.6% (5)	6.3% (2)
Evaluating employee interaction analysis (could include sensors).	(12)	62.5% (15)	29.2% (7)	8.3% (2)	0.0% (0)
Identifying employees likely to join or leverage an employee resource group (ERG) for culture integration.	(14)	95.7% (22)	4.3% (1)	0.0% (0)	0.0% (0)
Identifying communities of interest (e.g., using employee profiles and current assignments to encourage employee interaction).	(14)	82.6% (19)	17.4% (4)	0.0% (0)	0.0% (0)
Conducting sentiment analysis of social media data.	(10)	66.7% (18)	18.5% (5)	11.1% (3)	3.7% (1)
Identifying influencers in the organization's networks.	(13)	66.7% (16)	20.8% (5)	12.5% (3)	0.0% (0)
Developing compensation modeling for specific roles and/or employee groups (e.g., designing incentive strategies for specific jobs).	(12)	56.0% (14)	28.0% (7)	8.0% (2)	8.0% (2)

* Percentages calculated using only those who answered the item and did not indicate I do not know.

QUESTION 13: Are there additional AI-based applications that you see in your professional practice that that were not included above? If so, please list these here.

Category	Count	Percentage
None	1	25%
Recruitment - Job Post Writing	1	25%
Workforce Analytics	2	50%

QUESTION 14: Considering the AI-based applications listed above, what do you consider to be major concerns about their use?

Concern	Count	Percentage
Adverse Impact/Bias	23	33%
Validity	12	17%
Audit	7	10%
Transparency	5	7%
Causal Grounding	4	6%
Privacy	3	4%
Data Quality	3	4%
Accuracy	3	4%
Unstructured Guidance	2	3%
Impact On Disabled	2	3%
Explainability	2	3%
Documentation	2	3%
Differential Prediction	1	1%
Communication	1	1%

QUESTION 15: Compared to traditional forms of assessment, are AI-based applications more or less effective in predicting job performance or other relevant criteria?

Response	Count	Percent
Depends	12	44%
More	6	22%
Less	5	19%
Slightly More	2	7%
About The Same	1	4%
Undecided - Not Enough Available Data	1	4%
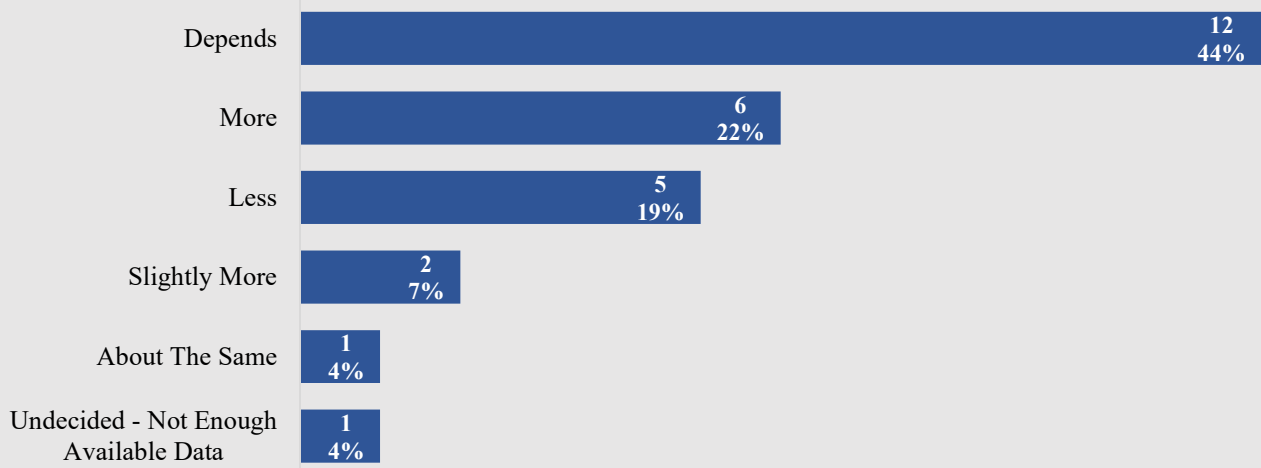
QUESTION 16: Do you have any additional comments related to this topic?

Response	Count	Percent
Narrowing Scope Of AI/Clarity	2	25%
None	2	25%
AI R&D Inadequate	1	13%
What Are The Long Term Impacts Of AI As It Relates To Bias	1	13%
Recruitment	1	13%
Question Suggestion - Is AI Better At Removing Bias?	1	13%
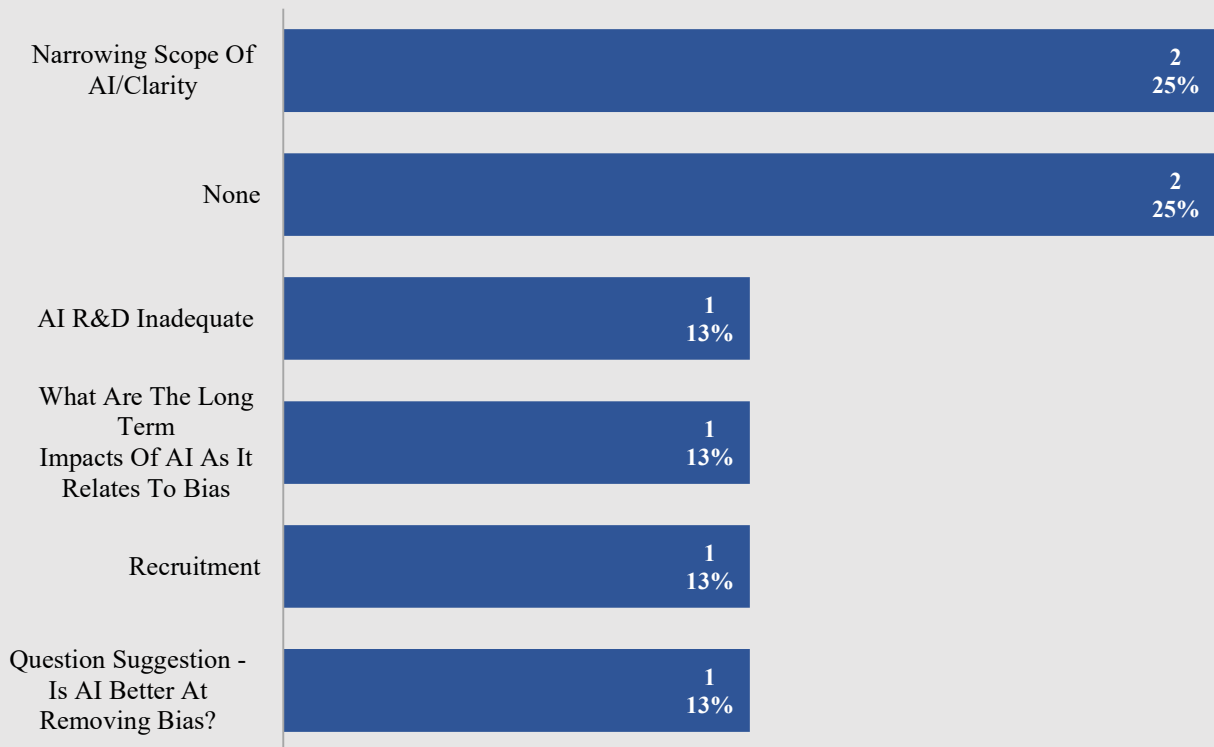
SECTION 3: TYPES OF DATA USED IN SELECTION-RELATED AI APPLICATIONS

QUESTION 17: How would you describe your level of familiarity with this topic?

	I do not know about this topic	12 31%

I know a little about this topic – I could potentially answer some of these questions — **12 31%**

I am very familiar with this topic – I should be able to answer all of these questions — **9 23%**

I am generally knowledgeable about this topic – I could probably answer a lot of these questions — **8 23%**

I am an expert on this topic – I expect to be able to answer all of these questions — **5 13%**

I do not know this topic – I could not answer these types of questions — **4 10%**

QUESTION 18: Please indicate the appropriateness of using each source/type of data in selection-related AI algorithms, models, or applications. CHECK ALL THAT APPLY. Note that some choices are mutually exclusive.

Data from simulations or work samples.

	I do not know	Inappropriate *This kind of data should not be used in selection*	Appropriate only if *supported by clear linkage to job requirements based on job analysis or content*	Appropriate as long as *it predicts the outcome of interest*	Appropriate as long as *it is in the model (model inclusion indicates job-relatedness)*	Almost always appropriate
Content of spoken/written responses	(1)	10.0% (3)	76.7% (23)	33.3% (10)	43.3% (13)	16.7% (5)
Behaviors exhibited/not exhibited	(1)	10.0% (3)	73.3% (22)	33.3% (10)	33.3% (10)	23.3% (7)
Time/speed	(1)	13.3% (4)	80.0% (24)	33.3% (10)	33.3% (10)	10.0% (3)
Resources used	(7)	4.2% (1)	79.2% (19)	37.5% (9)	25.0% (6)	8.3% (2)

* Percentages calculated using only those who answered the item and did not indicate I do not know.

QUESTION 19: Please indicate the appropriateness of using each source/type of data in selection-related AI algorithms, models, or applications. CHECK ALL THAT APPLY. Note that some choices are mutually exclusive.

Data from gamified assessments

	I do not know	Inappropriate *This kind of data should not be used in selection*	Appropriate only if *supported by clear linkage to job requirements based on job analysis or content*	Appropriate as long as *it predicts the outcome of interest*	Appropriate as long as *it is in the model (model inclusion indicates job-relatedness)*	Almost always appropriate
Right and wrong answers	(1)	0.0% (0)	73.3% (22)	50.0% (15)	50.0% (15)	16.7% (5)
Response changes	(3)	28.6% (8)	57.1% (16)	25.0% (7)	35.7% (10)	7.1% (2)
Response times	(2)	10.3% (3)	82.8% (24)	27.6% (8)	37.9% (11)	3.4% (1)
Cursor location	(9)	59.1% (13)	31.8% (7)	18.2% (4)	22.7% (5)	4.5% (1)

* Percentages calculated using only those who answered the item and did not indicate I do not know.

QUESTION 20: Please indicate the appropriateness of using each source/type of data in selection-related AI algorithms, models, or applications. CHECK ALL THAT APPLY. Note that some choices are mutually exclusive.

Data from background checks

	I do not know	Inappropriate *This kind of data should not be used in selection*	Appropriate only if *supported by clear linkage to job requirements based on job analysis or content*	Appropriate as long as *it predicts the outcome of interest*	Appropriate as long as *it is in the model (model inclusion indicates job-relatedness)*	Almost always appropriate
Previous employment	(1)	0.0% (0)	73.3% (22)	23.3% (7)	40.0% (12)	23.3% (7)
Education	(1)	0.0% (0)	83.3% (25)	30.0% (9)	36.7% (11)	13.3% (4)
Criminal activity	(2)	20.7% (6)	69.0% (20)	13.8% (4)	13.8% (4)	6.9% (2)
Credit history	(3)	39.3% (11)	53.6% (15)	17.9% (5)	14.3% (4)	3.6% (1)
Licensure and certification	(1)	0.0% (0)	73.3% (22)	13.3% (4)	30.0% (9)	13.3% (4)

* Percentages calculated using only those who answered the item and did not indicate I do not know.

QUESTION 21: Please indicate the appropriateness of using each source/type of data in selection-related AI algorithms, models, or applications. CHECK ALL THAT APPLY. Note that some choices are mutually exclusive.

Written text generated by the applicant for selection purposes

	I do not know	Inappropriate *This kind of data should not be used in selection*	Appropriate only if *supported by clear linkage to job requirements based on job analysis or content*	Appropriate as long as *it predicts the outcome of interest*	Appropriate as long as *it is in the model (model inclusion indicates job-relatedness)*	Almost always appropriate
Interview responses converted to text	(1)	0.0% (0)	60.0% (18)	36.7% (11)	40.0% (12)	30.0% (9)
Short answer responses converted to text	(0)	0.0% (0)	61.3% (19)	35.5% (11)	32.3% (10)	25.8% (8)
Responses from simulations, work samples, situational judgment items, etc.	(0)	3.2% (1)	64.5% (20)	32.3% (10)	41.9% (13)	22.6% (7)
Resume, C.V., application, cover letter	(0)	3.2% (1)	51.6% (16)	29.0% (9)	38.7% (12)	32.3% (10)

* Percentages calculated using only those who answered the item and did not indicate I do not know.

QUESTION 22: Please indicate the appropriateness of using each source/type of data in selection-related AI algorithms, models, or applications. CHECK ALL THAT APPLY. Note that some choices are mutually exclusive

Other characteristics of the applicant

	I do not know	Inappropriate *This kind of data should not be used in selection*	Appropriate only if *supported by clear linkage to job requirements based on job analysis or content*	Appropriate as long as *it predicts the outcome of interest*	Appropriate as long as *it is in the model (model inclusion indicates job-relatedness)*	Almost always appropriate
Facial features or expressions	(3)	60.7% (17)	32.1% (9)	10.7% (3)	21.4% (6)	3.6% (1)
Auditory characteristics	(5)	61.5% (16)	38.5% (10)	7.7% (2)	11.5% (3)	3.8% (1)
Body language or other physical movements	(3)	57.1% (16)	35.7% (10)	10.7% (3)	21.4% (6)	3.6% (1)
Personality, behavioral, or psychological characteristics	(1)	13.3% (4)	76.7% (23)	36.7% (11)	33.3% (10)	13.3% (4)

* Percentages calculated using only those who answered the item and did not indicate I do not know.
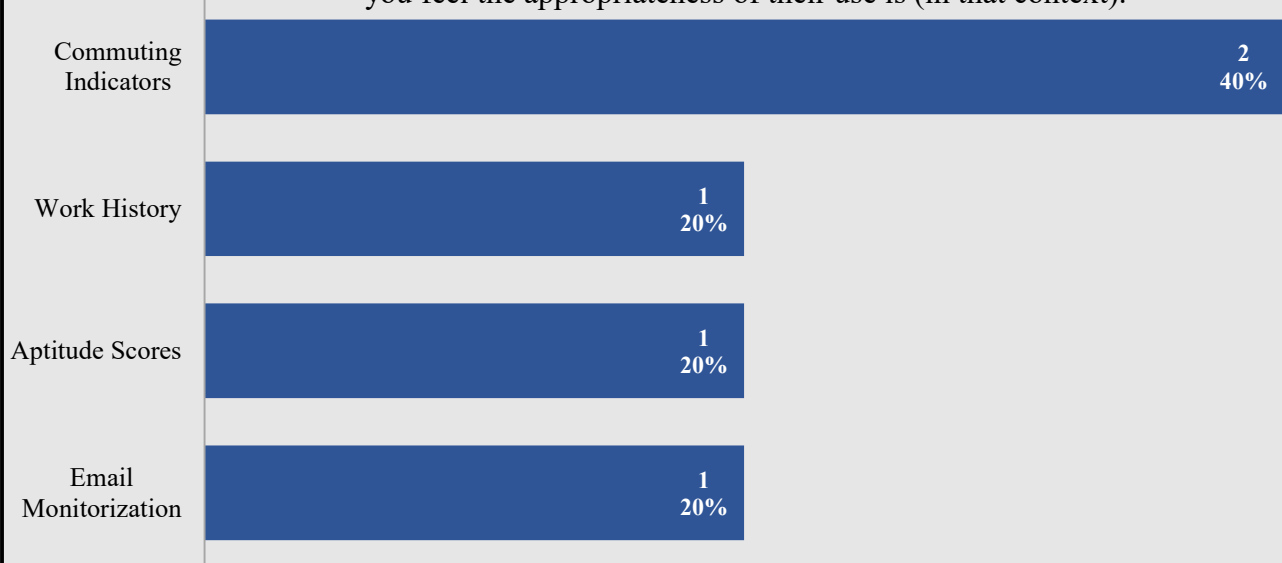
QUESTION 23: Please indicate the appropriateness of using each source/type of data in selection-related AI algorithms, models, or applications. CHECK ALL THAT APPLY. Note that some choices are mutually exclusive

Data from social media (e.g., Facebook, LinkedIn)

	I do not know	Inappropriate *This kind of data should not be used in selection*	Appropriate only if *supported by clear linkage to job requirements based on job analysis or content*	Appropriate as long as *it predicts the outcome of interest*	Appropriate as long as *it is in the model (model inclusion indicates job-relatedness)*	Almost always appropriate
Work history	(2)	13.8% (4)	58.6% (17)	24.1% (7)	41.4% (12)	17.2% (5)
Education	(2)	10.3% (3)	65.5% (19)	24.1% (7)	37.9% (11)	10.3% (3)
Groups and interests	(2)	55.2% (16)	34.5% (10)	17.2% (5)	17.2% (5)	6.9% (2)
Likes	(2)	72.4% (21)	17.2% (5)	10.3% (3)	13.8% (4)	3.4% (1)
Personal details	(3)	82.1% (23)	10.7% (3)	3.6% (1)	7.1% (2)	3.6% (1)
Connections	(3)	66.7% (18)	18.5% (5)	11.1% (3)	14.8% (4)	3.7% (1)
Posts	(4)	59.3% (16)	22.2% (6)	11.1% (3)	18.5% (5)	3.7% (1)
Licenses and certifications	(2)	10.3% (3)	62.1% (18)	27.6% (8)	41.4% (12)	13.8% (4)
Endorsements (received and given)	(2)	65.5% (19)	17.2% (5)	13.8% (4)	13.8% (4)	6.9% (2)
Awards	(3)	14.3% (4)	57.1% (16)	21.4% (6)	32.1% (9)	17.9% (5)

* Percentages calculated using only those who answered the item and did not indicate I do not know.

QUESTION 24: Are there additional sources/types of data used in AI algorithms, models, or applications that you have seen used? If so, please list these here and indicate what you feel the appropriateness of their use is (in that context).

Category	Count	Percentage
Commuting Indicators	2	40%
Work History	1	20%
Aptitude Scores	1	20%
Email Monitorization	1	20%

QUESTION 25: Considering the types of data that are often used in AI-based applications listed above, what are your major concerns about their use?

Concern	Count	Percent
Job/Business Relatedness	12	24%
Data Usage - Social Media	6	12%
Data Quality	6	12%
Bias	5	10%
Data Usage	5	10%
Biometric Data Usage	4	8%
Unintended Impact On Disabled	2	4%
Causal Grounding	2	4%
Validation	2	4%
Data Privacy	2	4%
Ethics	1	2%
Infiltration Of Non-Scientific Methods	1	2%
Unobserved Heterogenity	1	2%
Validity	1	2%
Data Transparency	1	2%
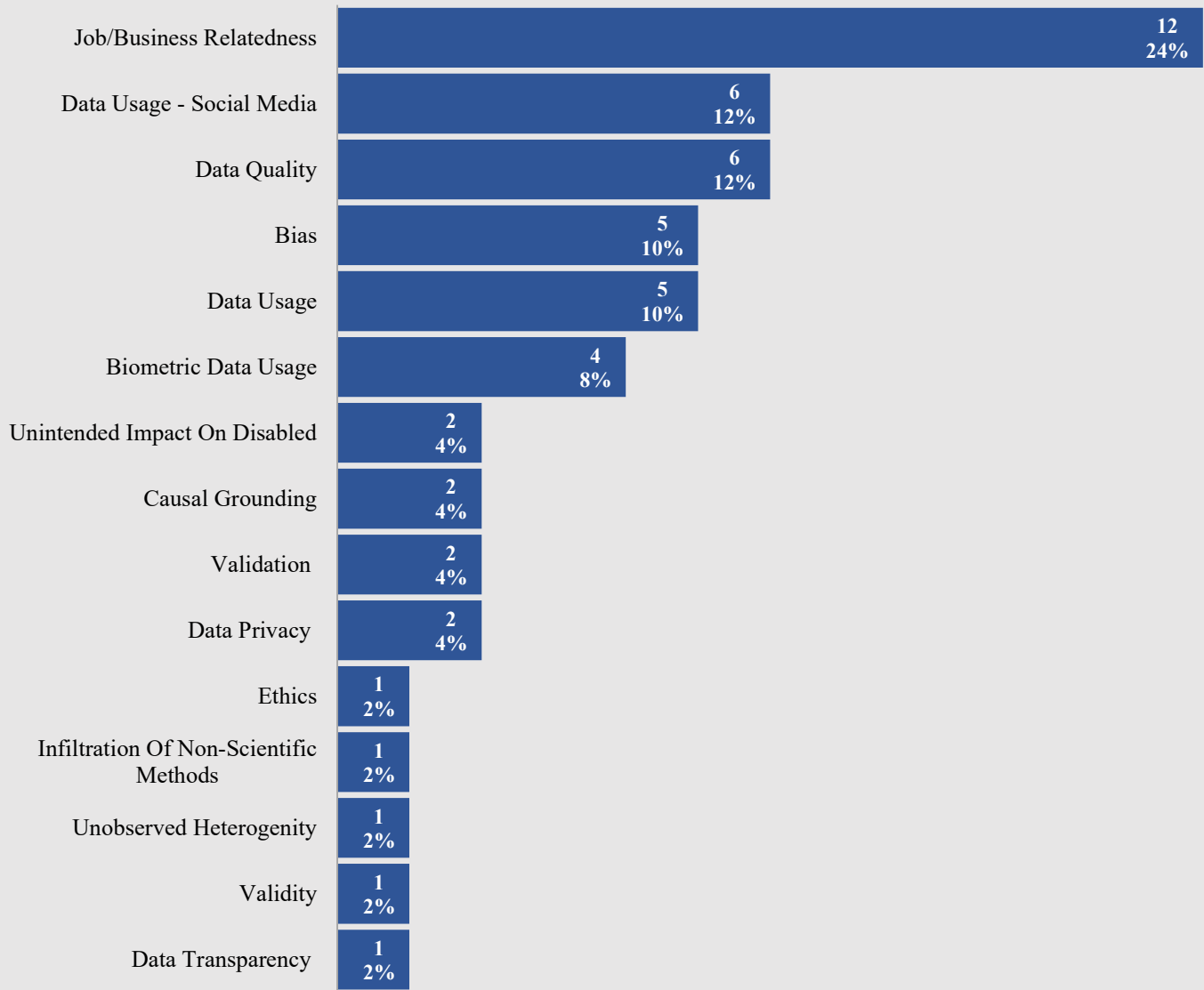
QUESTION 26: Compared to data from traditional forms of assessment, are the types of data used in AI-based applications more or less effective in predicting job performance or other relevant criteria?

Response	Count	Percentage
Depends	9	50%
Undecided	4	22%
More	2	11%
Less	2	11%
About The Same	1	6%

27: Do you have any additional comments related to this topic?

Category	Count	Percent
Question Recommendation - Clarity	5	26%
Discrimination/Bias Concern	2	11%
Job/Business Relatedness	2	11%
Unintended Consequences On Sub-Groups	1	5%
Impact On Disabled	1	5%
Unbiased Data	1	5%
Data Quality	1	5%
Model Diagnostics	1	5%
Validity	1	5%
Data Volunteered vs Data Gathered	1	5%
Data Integrity	1	5%
Weight Given To Data	1	5%
None	1	5%
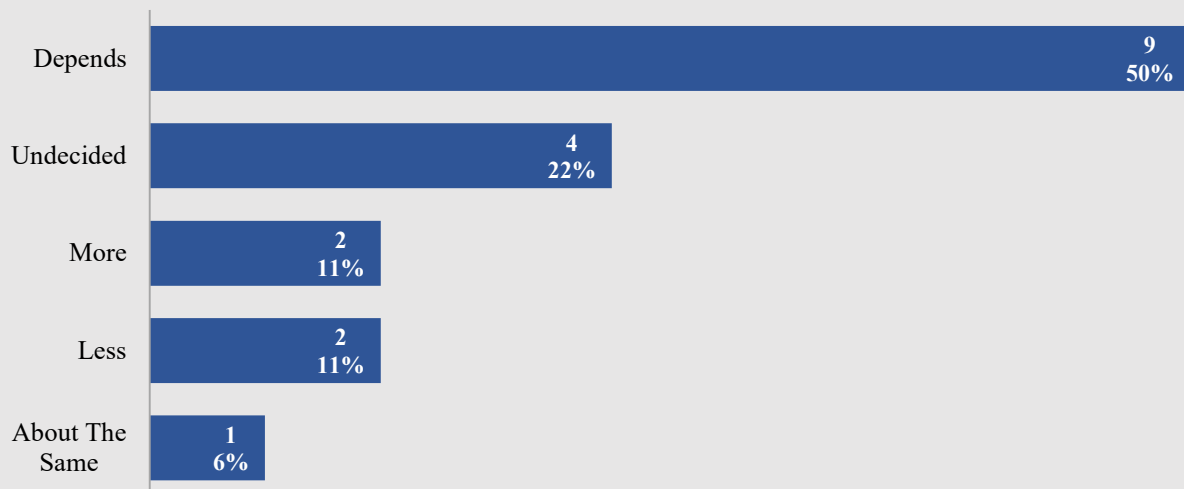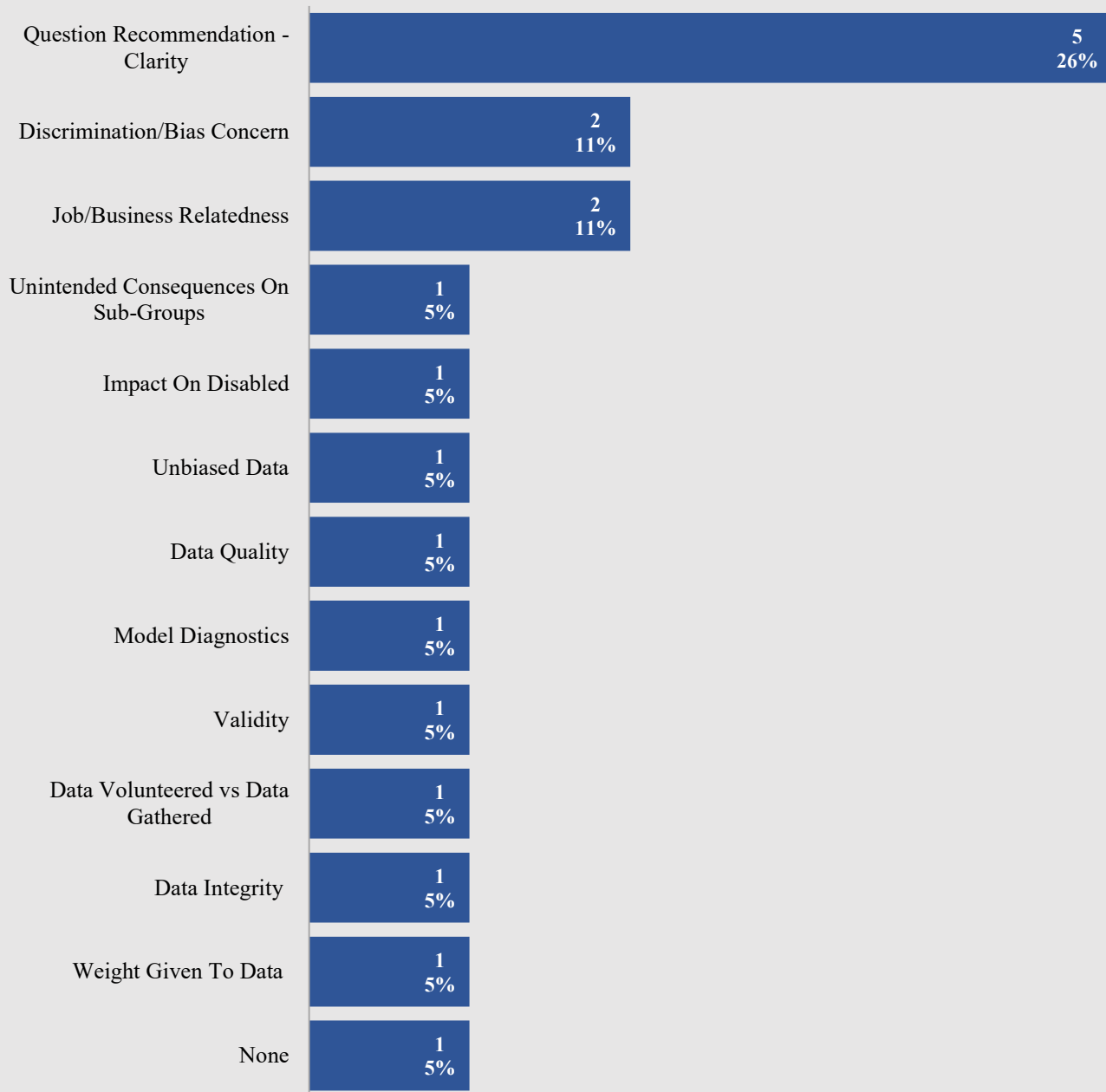
SECTION 4: ALGORITHM REVIEW, MONITORING, AND UPDATING

QUESTION 28: How would you describe your level of familiarity with this topic?

	I do not know	Not Important	Somewhat Important	Important	Very Important	Critically Important
The architecture of the algorithm	(6)	0.0% (0)	28.0% (7)	16.0% (4)	28.0% (7)	28.0% (7)
The explainability of the algorithm's inputs	(1)	0.0% (0)	13.3% (4)	26.7% (8)	26.7% (8)	33.3% (10)
The explainability of the algorithm's outputs	(1)	0.0% (0)	13.3% (4)	20.0% (6)	33.3% (10)	33.3% (10)
The number of candidates the algorithm processes	(1)	13.3% (4)	23.3% (7)	23.3% (7)	26.7% (8)	13.3% (4)
The amount of additional training data available to be incorporated into the model	(3)	0.0% (0)	21.4% (6)	21.4% (6)	32.1% (9)	25.0% (7)
Observed changes in the adverse impact in the algorithm's outputs	(0)	0.0% (0)	0.0% (0)	16.1% (5)	22.6% (7)	61.3% (19)
Observed changes in the accuracy of the algorithm's outputs	(0)	0.0% (0)	0.0% (0)	6.5% (2)	19.4% (6)	74.2% (23)

QUESTION 29: Please indicate how important each of the following factors is in determining how often an AI-based application should be reviewed to evaluate its effectiveness (e.g., validity, ability to predict target outcome, achieving the intended goal).

QUESTION 29: Please indicate how important each of the following factors is in determining how often an AI-based application should be reviewed to evaluate its effectiveness (e.g., validity, ability to predict target outcome, achieving the intended goal).

	I do not know	Not Important	Somewhat Important	Important	Very Important	Critically Important
Observed changes in the reliability of the algorithm	(0)	0.0% (0)	0.0% (0)	9.7% (3)	19.4% (6)	71.0% (22)
The period of time that has elapsed since the algorithm was last validated or audited	(1)	0.0% (0)	16.7% (5)	26.7% (8)	46.7% (14)	10.0% (3)
Substantial changes to the job or job requirements	(0)	0.0% (0)	0.0% (0)	6.5% (2)	29.0% (9)	64.5% (20)

* Percentages calculated using only those who answered the item and did not indicate I do not know.

QUESTION 30: Please indicate how important each of the following factors is in determining how often an AI-based application should be reviewed to evaluate its adverse impact (note that the EEOC (29 CFR 1607.16B) defines adverse impact as a substantially different rate of selection in hiring, promotion, or other employment decision which works to the disadvantage of members of a race, sex, or ethnic group). Adverse impact can also apply to age and disability.

	I do not know	Not Important	Somewhat Important	Important	Very Important	Critically Important
The architecture of the algorithm	(6)	8.7% (2)	30.4% (7)	21.7% (5)	17.4% (4)	21.7% (5)
The explainability of the algorithm's inputs	(1)	3.4% (1)	20.7% (6)	24.1% (7)	31.0% (9)	20.7% (6)
The explainability of the algorithm's outputs	(1)	0.0% (0)	17.2% (5)	24.1% (7)	20.7% (6)	37.9% (11)
The number of candidates the algorithm processes	(0)	13.3% (4)	13.3% (4)	23.3% (7)	16.7% (5)	33.3% (10)
The amount of additional training data available to be incorporated	(3)	7.4% (2)	18.5% (5)	25.9% (7)	22.2% (6)	25.9% (7)
Observed changes in the adverse impact in the algorithm's outputs	(0)	0.0% (0)	3.3% (1)	10.0% (3)	6.7% (2)	80.0% (24)
Observed changes in the accuracy of the algorithm's outputs	(0)	0.0% (0)	3.3% (1)	6.7% (2)	23.3% (7)	66.7% (20)
Observed changes in the reliability of the algorithm	(0)	0.0% (0)	3.3% (1)	3.3% (1)	23.3% (7)	70.0% (21)
The period of time that has elapsed since the algorithm was last validated or audited	(0)	3.3% (1)	16.7% (5)	20.0% (6)	33.3% (10)	26.7% (8)
Substantial changes to the job or job requirements	(0)	0.0% (0)	3.3% (1)	10.0% (3)	26.7% (8)	60.0% (18)

* Percentages calculated using only those who answered the item and did not indicate I do not know.

QUESTION 31: Are there additional factors that should be evaluated (for effectiveness or adverse impact) that were not included above? If so, please list them here.

Factor	Count	Percent
Frequency Of Evaluation And Review	2	11%
Changes In Selectivity (Overall And By Sub-Group)	1	5%
AI As Tool Among Many Versus AI As The Main Gatekeeper	1	5%
Changes In Job	1	5%
Selectivity Of Predictors	1	5%
Analysis Of Influential Observations	1	5%
Validation Criteria	1	5%
Change In Employee Productivity	1	5%
Opportunity Cost Of Not Using AI:Effectiveness And…	1	5%
Location	1	5%
Proportion Of Title VII Relevant Subgroups	1	5%
Transparency	1	5%
Assessment Of Benchmarks	1	5%
Job-Related Outcome Metrics	1	5%
Criterion Validity	1	5%
Tabulation Of Workforce Metrics (Such As Turnover, Training Time, Etc.)	1	5%
Least Biased Alternative	1	5%
None	1	5%

QUESTION 32: Do you have any additional related comments?

Comment	Count	Percent
Importance of Frequency of Evaluation	2	50%
Importance of Validity	1	25%
None	1	25%
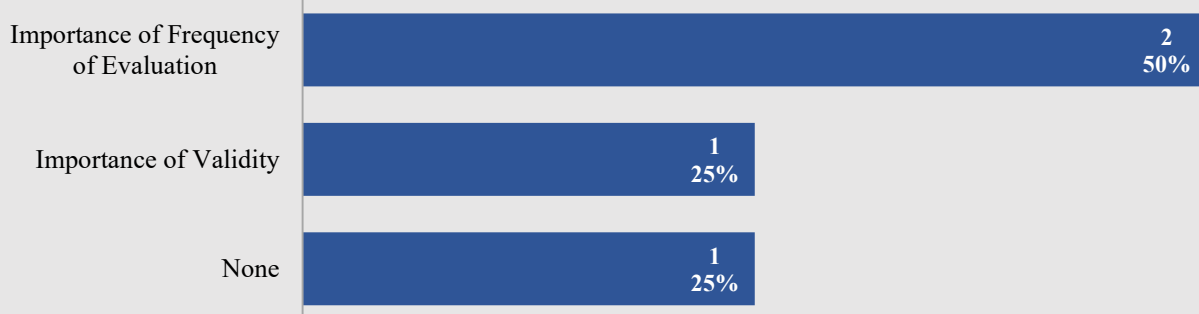
SECTION 5: GENERAL DATA QUALITY ISSUES

QUESTION 33: What are the most common data quality issues you have encountered in automated HR decision systems (e.g., inclusion of tainted features, data accuracy, data sufficiency, representativeness)

Category	Count	Percentage
Data Sufficiency	17	25%
Incomplete Data	10	15%
Representativeness	9	13%
Data Accuracy	8	12%
Data Consistency	5	7%
Tainted Variables	4	6%
Demographic Data Considerations	3	4%
Job/Business Relatedness	2	3%
Unobserved Heterogeneity	2	3%
Data Transparency	2	3%
Data Accuracy	1	1%
Validity	1	1%
Data Security	1	1%
Causal Grounding	1	1%
Considerations	1	1%
N/A	1	1%
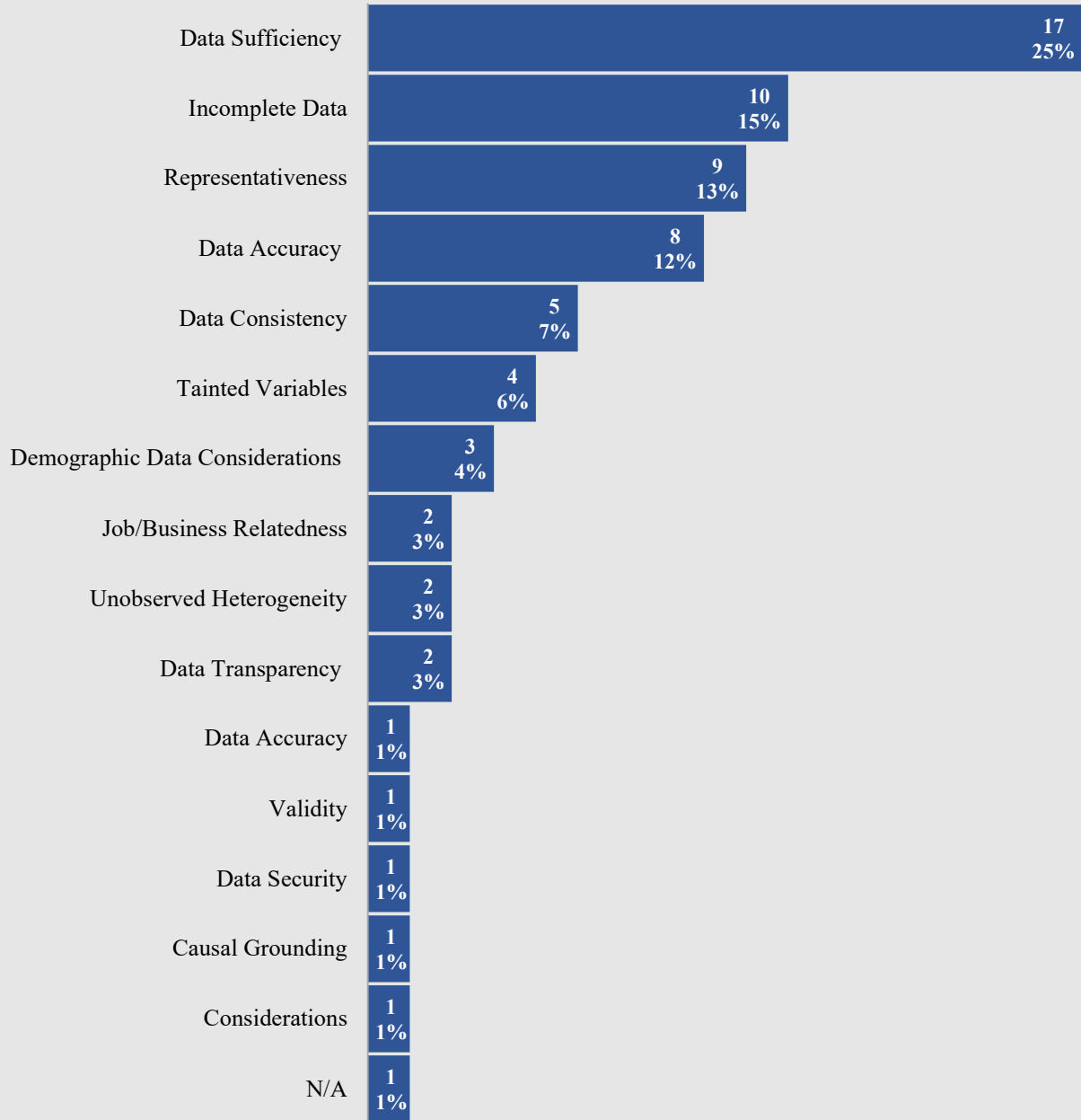
QUESTION 34: What are the most consequential data quality issues you have encountered in automated HR decision systems (e.g., inclusion of tainted features, data accuracy, data sufficiency, representativeness)?

Category	Count	Percent
Representativeness	10	29%
Data Sufficiency	7	20%
Data Accuracy	6	17%
Incomplete Data	4	11%
Tainted Variables	3	9%
Tainted Demographic Features	2	6%
Validity	1	3%
Job/Business Relatedness	1	3%
N/A	1	3%
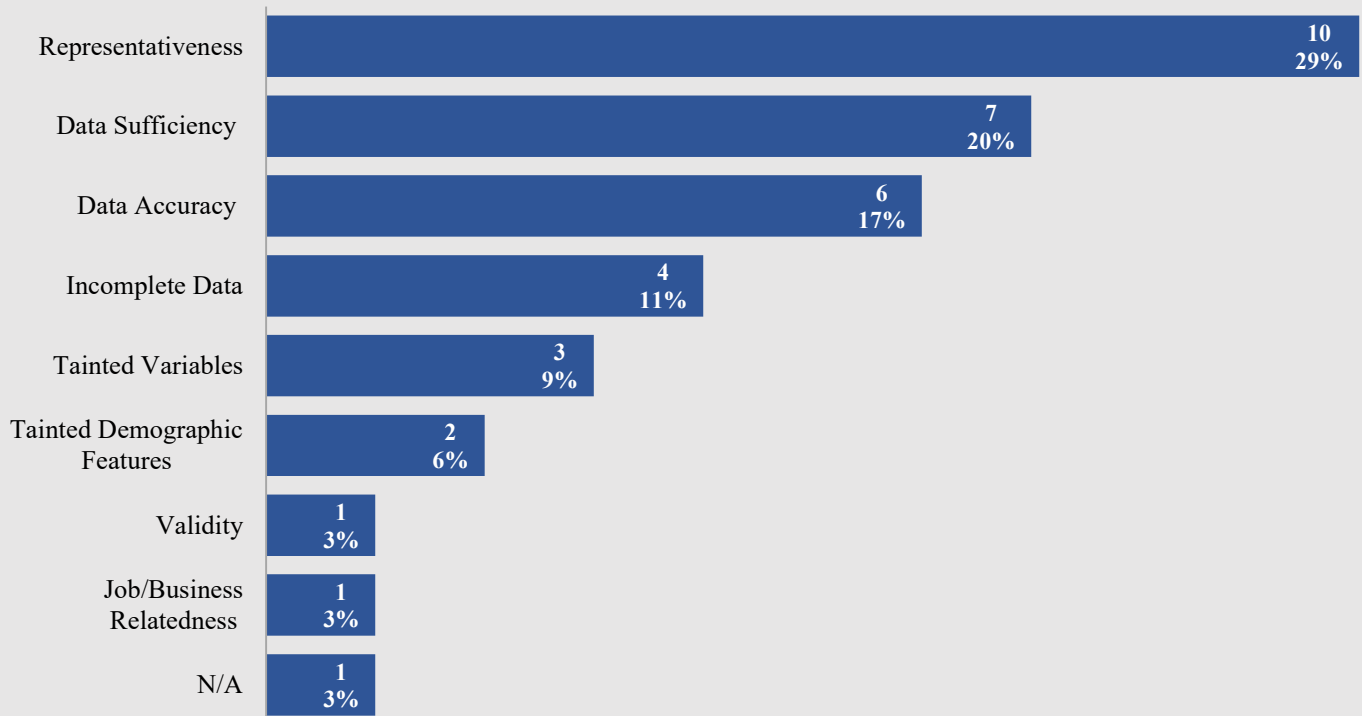
QUESTION 35: What are the most common methods that you have used (or seen used) to address issues with data quality?

Method	Count	Percentage
Regularization	5	19%
Data Audits	4	15%
Pre/Post Data Cleaning	3	11%
Developer/ User Education	3	11%
Use Data As-Is	2	7%
Ignorance	2	7%
Additional Data	2	7%
Human-In-The-Loop	1	4%
Sensitivity Tests	1	4%
Checks On Internal And External Data Consistency	1	4%
Sampling	1	4%
Adverse Impact Analysis	1	4%
I Do Not Know	1	4%
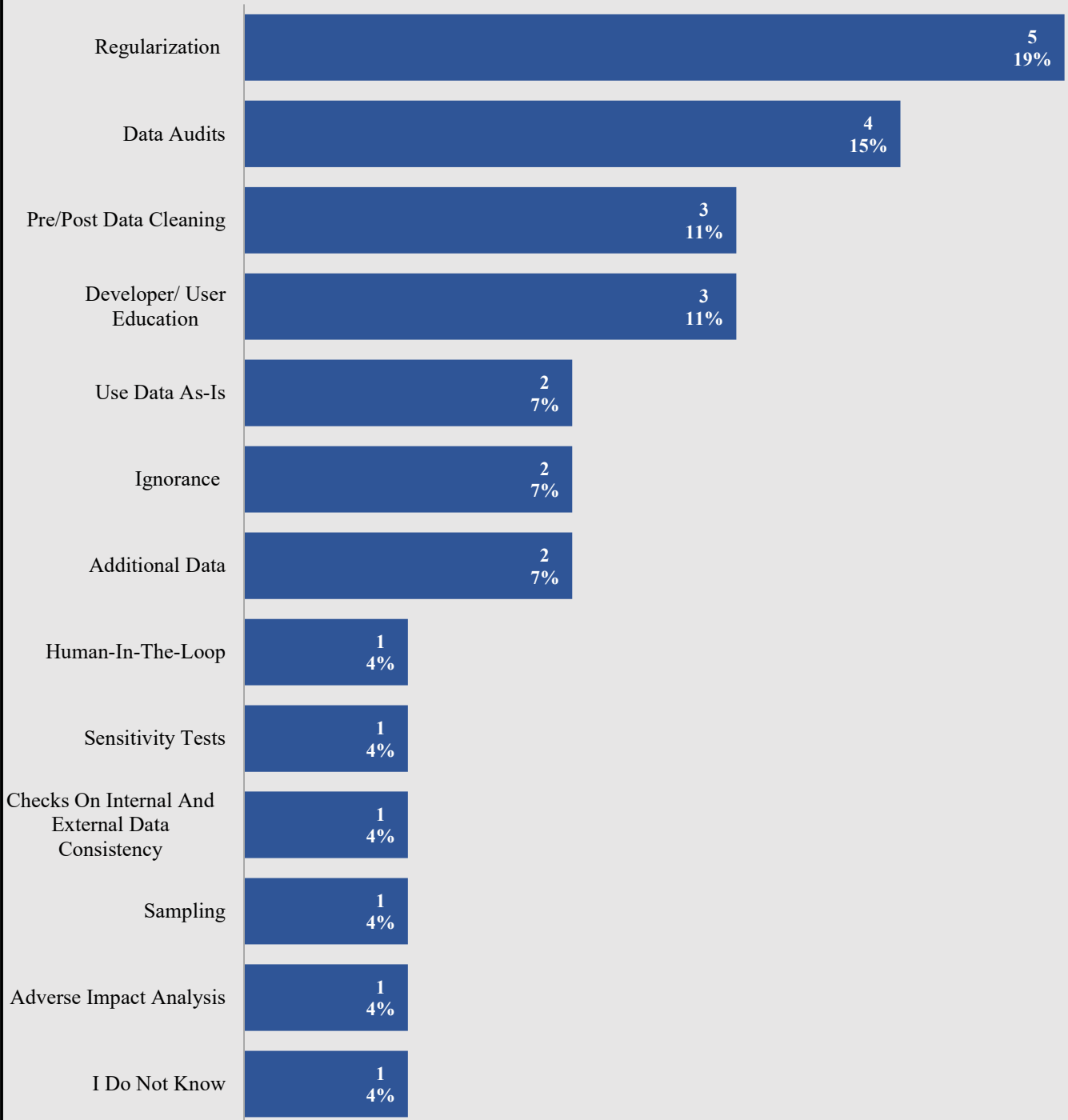
QUESTION 36: Do you have any additional comments related to data quality issues?

Category	Count	Percent
None	2	29%
Outside Data Out Of Applicant Control Can Be Problematic	1	14%
Trust Own Data But Not Others'	1	14%
End Goal Disagreement Between Parties At It Relates To…	1	14%
Concern About Stable Data Quality Over Time.	1	14%
Assessment/ Interviews Used To Collect Data	1	14%
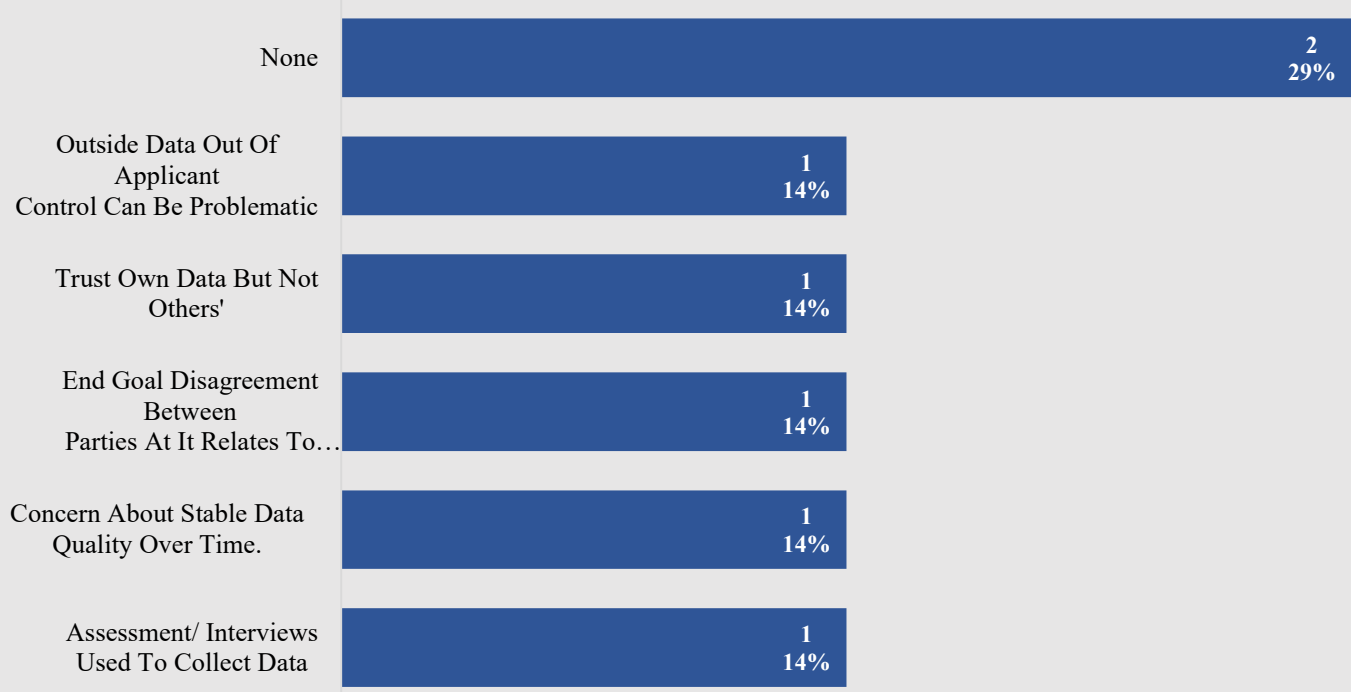
SECTION 6: PRIVACY, CONSENT, TRANSPARENCY
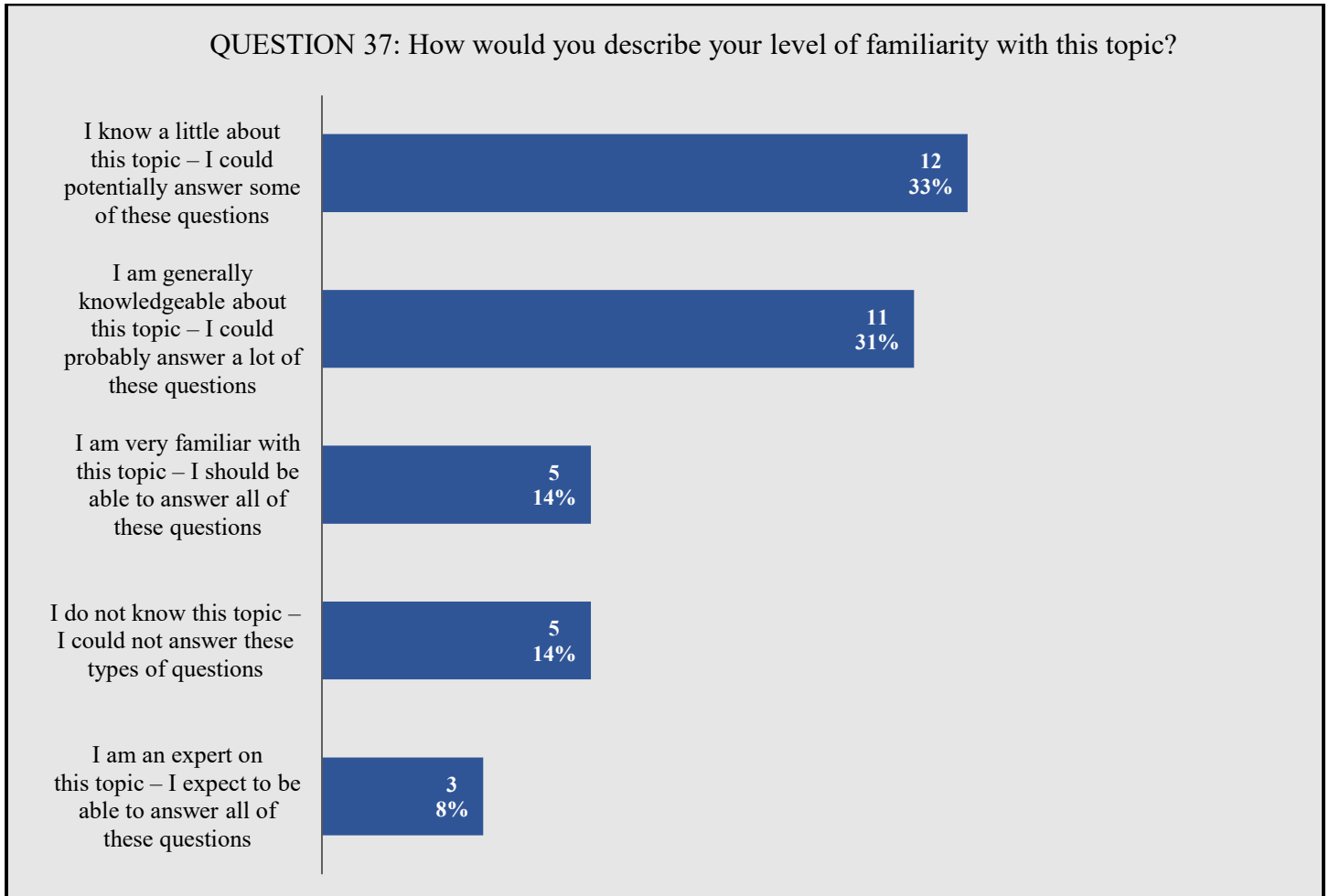
QUESTION 37: How would you describe your level of familiarity with this topic?

I know a little about this topic – I could potentially answer some of these questions	12 — 33%
I am generally knowledgeable about this topic – I could probably answer a lot of these questions	11 — 31%
I am very familiar with this topic – I should be able to answer all of these questions	5 — 14%
I do not know this topic – I could not answer these types of questions	5 — 14%
I am an expert on this topic – I expect to be able to answer all of these questions	3 — 8%

QUESTION 38: Please indicate your level of concern with the issue noted.

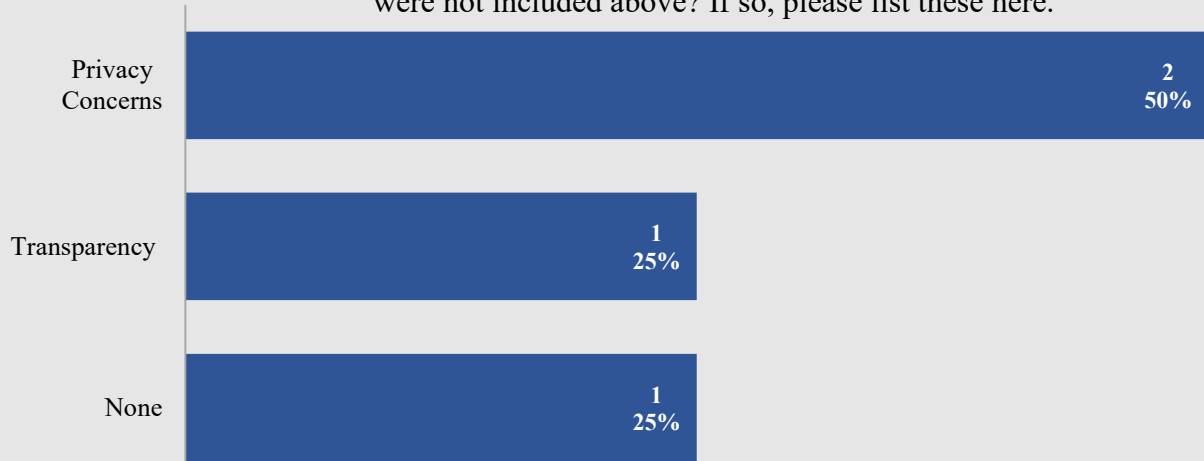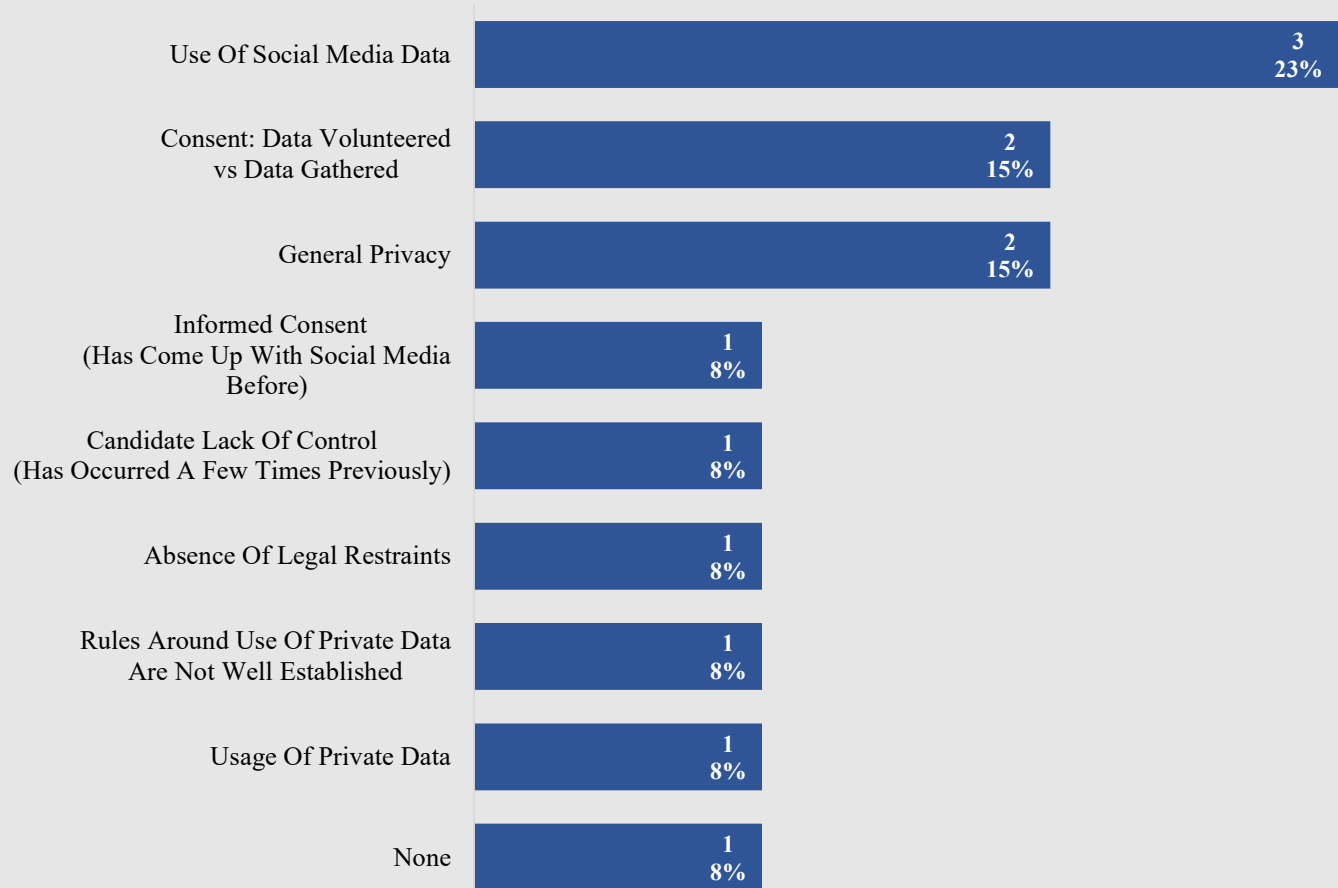|	I do not know	I am not concerned	I am mildly concerned	I am concerned	I am very concerned	I am critically concerned
The potential for data privacy issues stemming from AI applications	(1)	3.4% (1)	24.1% (7)	31.0% (9)	24.1% (7)	17.2% (5)
The possible unintended disclosure of an individual's identity or sensitive information despite the removal of obvious identifiers.	(1)	6.9% (2)	13.8% (4)	24.1% (7)	41.4% (12)	13.8% (4)
The difficulty of obtaining informed consent from individuals in advance.	(0)	10.0% (3)	26.7% (8)	26.7% (8)	20.0% (6)	16.7% (5)
The timing of when informed consent from an individual must be obtained.	(1)	6.9% (2)	34.5% (10)	27.6% (8)	17.2% (5)	13.8% (4)

QUESTION 38: Please indicate your level of concern with the issue noted.

	I do not know	I am not concerned	I am mildly concerned	I am concerned	I am very concerned	I am critically concerned
An individuals' right to know what methods will be used and what data will be collected and used.	(1)	10.3% (3)	20.7% (6)	20.7% (6)	27.6% (8)	20.7% (6)
Individuals' right to review and/or update or change their data.	(0)	6.7% (2)	20.0% (6)	23.3% (7)	33.3% (10)	16.7% (5)
The lack of specific information on how the data will be used before consent is obtained.	(0)	3.3% (1)	20.0% (6)	16.7% (5)	43.3% (13)	16.7% (5)
The need for consent when previously collected data are re-purposed.	(1)	6.9% (2)	6.9% (2)	31.0% (9)	31.0% (9)	24.1% (7)
The level of guidance available related to data retention and data management.	(0)	10.0% (3)	20.0% (6)	16.7% (5)	40.0% (12)	13.3% (4)
The potential lack of transparency about how some algorithms operate (e.g., black-box algorithms)	(0)	3.3% (1)	10.0% (3)	13.3% (4)	36.7% (11)	36.7% (11)
The level of detail that will be required by vendors or developers to achieve sufficient transparency.	(1)	13.8% (4)	13.8% (4)	20.7% (6)	27.6% (8)	24.1% (7)
The ease of explainability to end-users of AI-based algorithms, models, or applications.	(0)	13.3% (4)	10.0% (3)	30.0% (9)	30.0% (9)	16.7% (5)
The explainability and transparency that may be required by law or regulation.	(0)	16.7% (5)	3.3% (1)	13.3% (4)	33.3% (10)	33.3% (10)
The intellectual property rights of vendors or developers of AI-based applications.	(0)	30.0% (9)	23.3% (7)	16.7% (5)	20.0% (6)	10.0% (3)
The potential conflict between a vendor/developers' intellectual property rights and individual rights for transparency and explainability.	(0)	13.3% (4)	26.7% (8)	6.7% (2)	26.7% (8)	26.7% (8)

* Percentages calculated using only those who answered the item and did not indicate I do not know.

QUESTION 39: Are there additional areas of concern that should be considered that were not included above? If so, please list these here.

Category	Count	Percent
Privacy Concerns	2	50%
Transparency	1	25%
None	1	25%

QUESTION 40: Do you have any additional comments related to this topic?

Category	Count	Percent
Use Of Social Media Data	3	23%
Consent: Data Volunteered vs Data Gathered	2	15%
General Privacy	2	15%
Informed Consent (Has Come Up With Social Media Before)	1	8%
Candidate Lack Of Control (Has Occurred A Few Times Previously)	1	8%
Absence Of Legal Restraints	1	8%
Rules Around Use Of Private Data Are Not Well Established	1	8%
Usage Of Private Data	1	8%
None	1	8%

SECTION 7: COMMUNICATIONS

QUESTION 41: How would you describe your level of familiarity with this topic?

	I do not know	Not Important	Somewhat Important	Important	Very Important	Critically Important
QUESTION 42: Please indicate the importance of organizations developing each type of communication.						
That explain to individual applicants the purpose of the AI application (prior to use).	(0)	0.0% (0)	9.4% (3)	25.0% (8)	40.6% (13)	25.0% (8)
That explain to individual applicants the type of AI model or approach (prior to use).	(0)	18.8% (6)	40.6% (13)	21.9% (7)	6.3% (2)	12.5% (4)
That explain to individual applicants the features or types of features that are used in the AI application (prior to use).	(0)	12.5% (4)	31.3% (10)	21.9% (7)	21.9% (7)	12.5% (4)
That explain to other stakeholders (e.g., organization managers, unions) the AI based selection procedure.	(0)	3.1% (1)	25.0% (8)	18.8% (6)	34.4% (11)	18.8% (6)
About procedures related to applicant appeal rights.[†]	(2)	0.0% (0)	26.7% (8)	10.0% (3)	33.3% (10)	30.0% (9)
From an employer to individual applicants providing feedback following an AI-based selection.	(1)	12.9% (4)	32.3% (10)	19.4% (6)	22.6% (7)	12.9% (4)

QUESTION 42: Please indicate the importance of organizations developing each type of communication.

	I do not know	Not Important	Somewhat Important	Important	Very Important	Critically Important
That are intended for internal applicants (e.g., transfers, promotions).	(1)	0.0% (0)	22.6% (7)	16.1% (5)	38.7% (12)	22.6% (7)
That are intended for external applicants.	(0)	0.0% (0)	28.1% (9)	21.9% (7)	28.1% (9)	21.9% (7)
That solicit feedback from individuals about the specific AI application.	(0)	15.6% (5)	28.1% (9)	25.0% (8)	21.9% (7)	9.4% (3)
That solicit feedback from front-line stakeholders about the AI application (e.g., ask interviewers for feedback about a newly installed AI-based interview system).	(0)	6.3% (2)	18.8% (6)	18.8% (6)	46.9% (15)	9.4% (3)

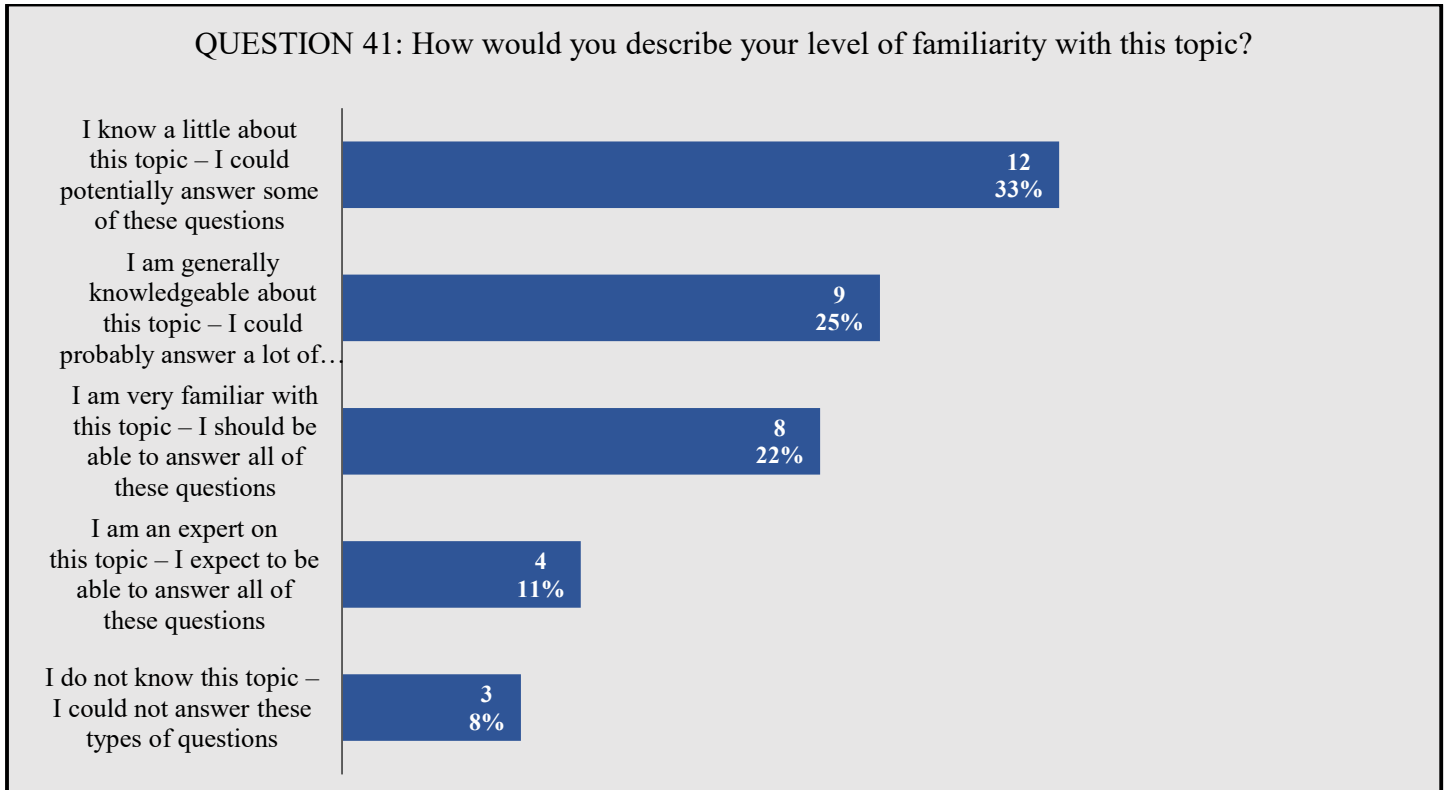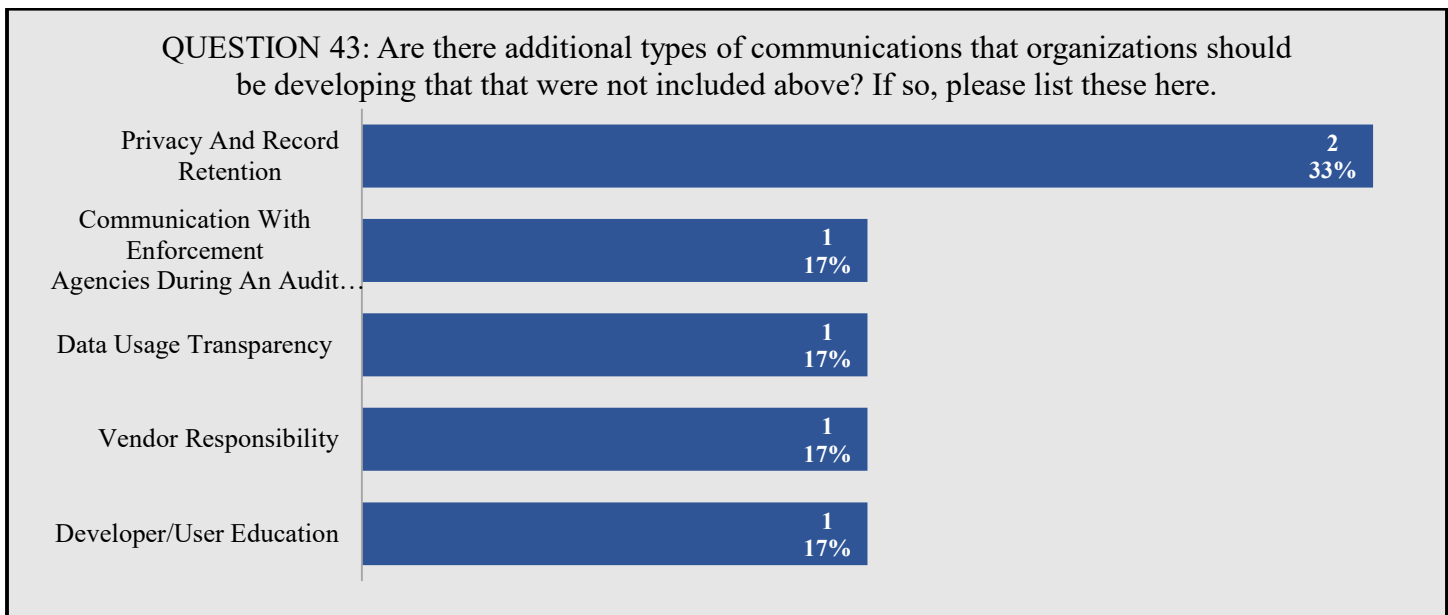** Percentages calculated using only those who answered the item and did not indicate I do not know.*
† This question was inadvertently asked twice and responses were highly similar; first set of responses reported here.

QUESTION 43: Are there additional types of communications that organizations should be developing that that were not included above? If so, please list these here.

Privacy And Record Retention	2 — 33%
Communication With Enforcement Agencies During An Audit…	1 — 17%
Data Usage Transparency	1 — 17%
Vendor Responsibility	1 — 17%
Developer/User Education	1 — 17%

QUESTION 44: Do you have any additional comments related to this topic?

Category	Count	Percentage
Importance Of Transparency Does Not…	3	33%
Importance Of Transparency	2	22%
User Education	1	11%
Implied Consent	1	11%
N/A or None	2	22%
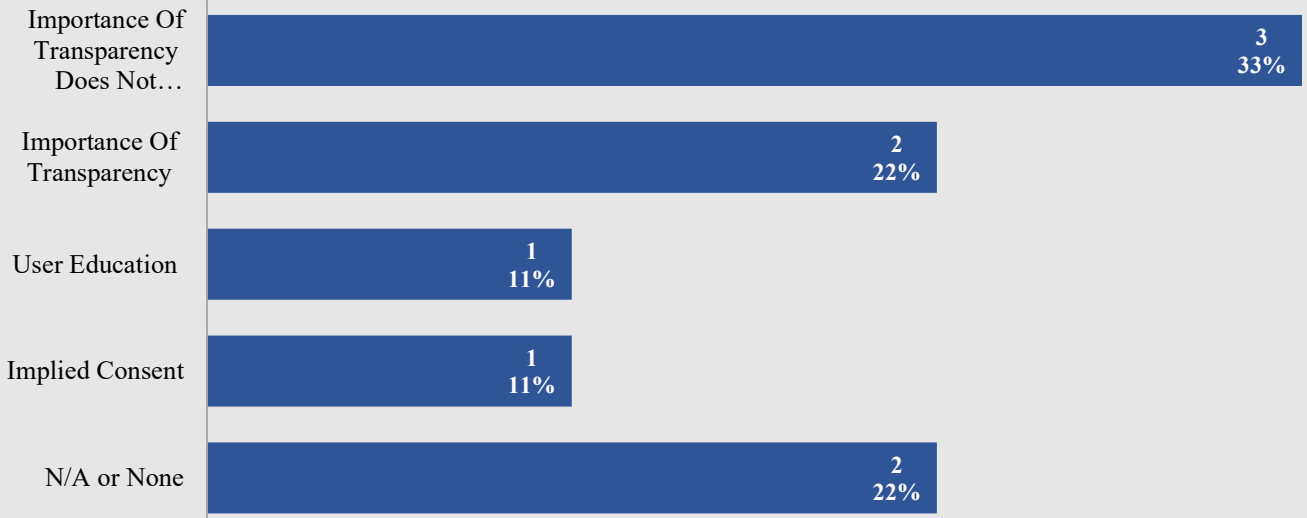
SECTION 8: STATISTICAL ISSUES

QUESTION 45: How would you describe your level of familiarity with this topic?

I know a little about this topic – I could potentially answer some of these questions	4 / 11%
I am generally knowledgeable about this topic – I could probably answer a lot of these questions	7 / 19%
I am very familiar with this topic – I should be able to answer all of these questions	4 / 11%
I am an expert on this topic – I expect to be able to answer all of these questions	9 / 25%
I do not know this topic – I could not answer these types of questions	12 / 33%

QUESTION 46: Based on what you are familiar with, what are the different approaches, methods, or statistics available for evaluating the effectiveness of an AI tool, algorithm, model, or application used to make employment decisions or assessments (e.g., measures of statistical significance, measures of magnitude/effect size, measures of accuracy/precision/recall)?

Category	Count	Percentage
Statistical Significance	7	14%
APR-F1	7	14%
Model Fit Statistics	6	12%
Practical Significance/Effect Size	4	8%
Adverse Impact Analysis	4	8%
Cross Validation	3	6%
AUC	3	6%
Taylor-Made	2	4%
R2	2	4%
Correlations	2	4%
AUC/ROC	2	4%
Validity, Causal Grounding, Differntial Prediction (Fairness)	1	2%
Regression	1	2%
Random Forest	1	2%
Psychometric Analyses	1	2%
Partial Dependence Plots	1	2%
Measures Of Feature Importance (Shapley Value)	1	2%
Logistic Regression	1	2%
Depends On The Tool	1	2%
Cross Entropy	1	2%

Note: APR-F1 is Accuracy, Precision, Recall, F1 Score; AUC is Area Under the Curve (as is AUC/ROC)

QUESTION 47: Of the statistics or methods used for evaluating effectiveness, which do you think are the most useful?

Category	Count	Percent
APR-F1	3	11%
Taylor-Made	3	11%
R2	3	11%
Practical Significance/Effect Size	2	7%
Model Fit Statistics	2	7%
Statistical Significance Test	2	7%
AUC	2	7%
Prediction Statistics	1	4%
UGESP Methods	1	4%
Important To Have Multiple Tests	1	4%
Differential Prediction	1	4%
Correlation Statistics	1	4%
Fairness	1	4%
Regression	1	4%
Undecided	1	4%
Cross Validation	1	4%
Effect Size	1	4%
Dependent On Tool/ Algorithm	1	4%
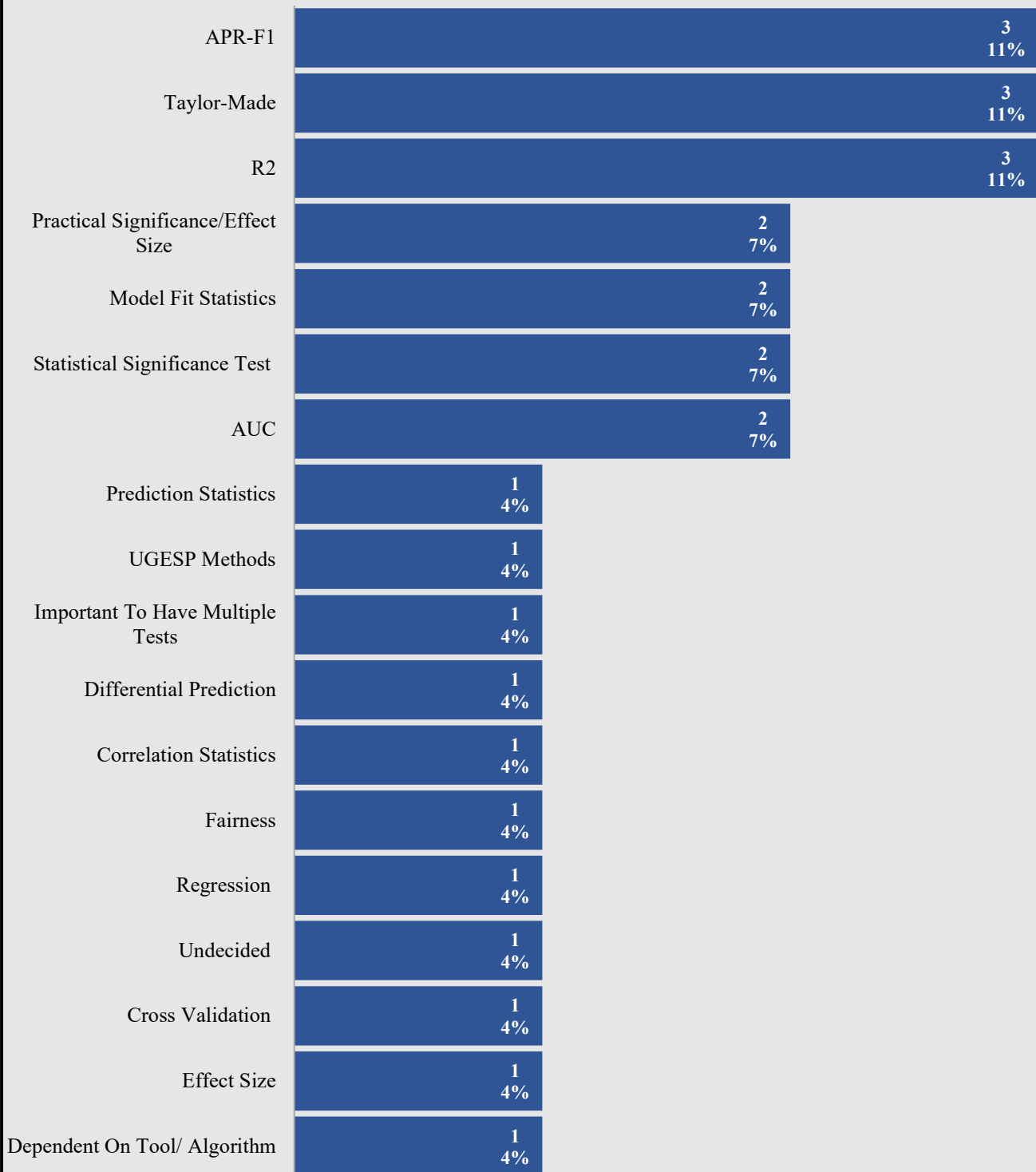
Note: APR-F1 is Accuracy, Precision, Recall, F1 Score; AUC is Area Under the Curve

QUESTION 48: Of the statistics or methods used for evaluating effectiveness, which do you think are the most commonly used?

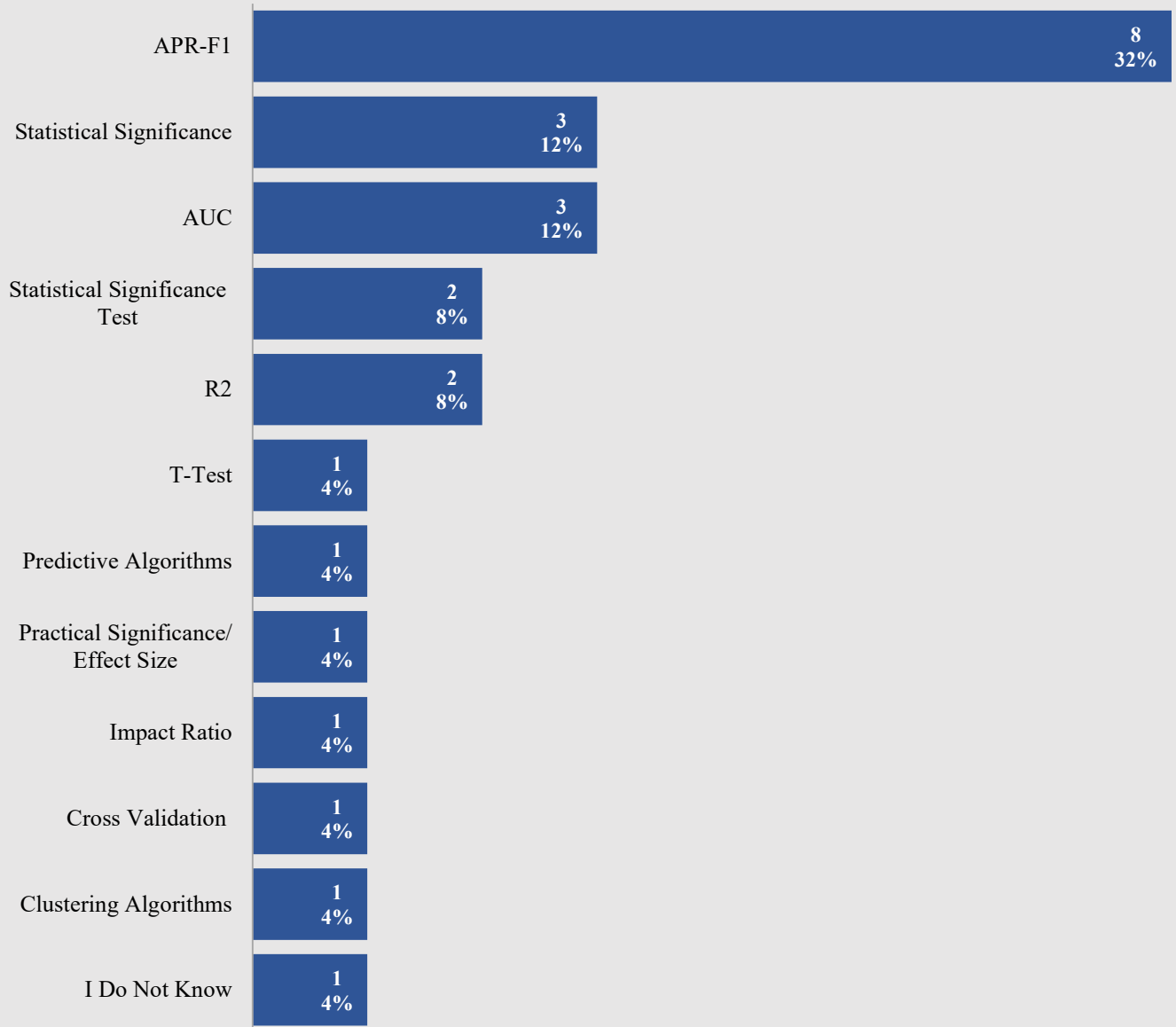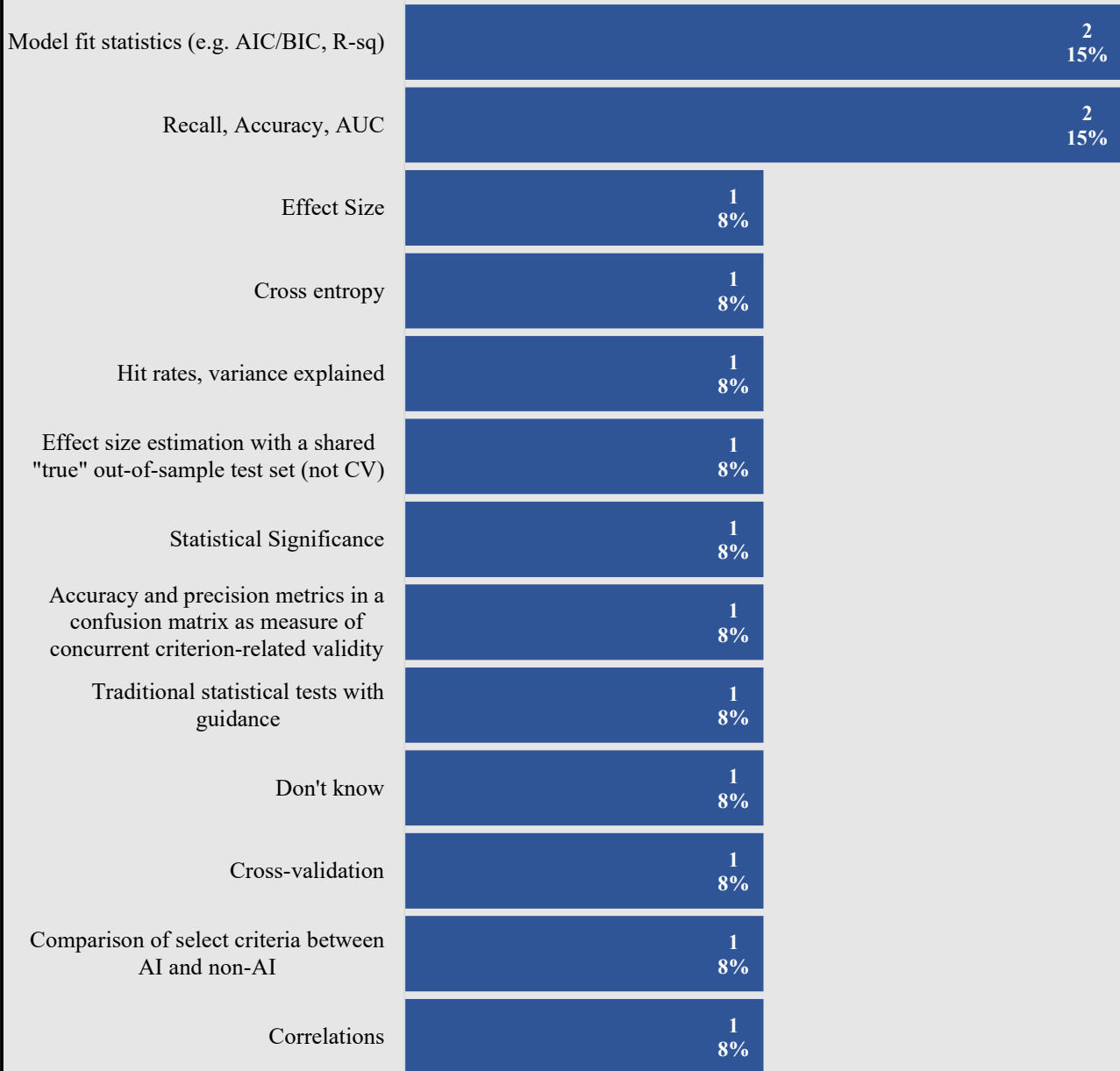| Category	Count	Percentage
APR-F1	8	32%
Statistical Significance	3	12%
AUC	3	12%
Statistical Significance Test	2	8%
R2	2	8%
T-Test	1	4%
Predictive Algorithms	1	4%
Practical Significance/ Effect Size	1	4%
Impact Ratio	1	4%
Cross Validation	1	4%
Clustering Algorithms	1	4%
I Do Not Know	1	4%

Note: APR-F1 is Accuracy, Precision, Recall, F1 Score; AUC is Area Under the Curve

Question 49: What statistics or methods would be used if comparing the effectiveness of multiple AI algorithms (developed with similar or different analytic techniques)?

Category	Count	Percent
Model fit statistics (e.g. AIC/BIC, R-sq)	2	15%
Recall, Accuracy, AUC	2	15%
Effect Size	1	8%
Cross entropy	1	8%
Hit rates, variance explained	1	8%
Effect size estimation with a shared "true" out-of-sample test set (not CV)	1	8%
Statistical Significance	1	8%
Accuracy and precision metrics in a confusion matrix as measure of concurrent criterion-related validity	1	8%
Traditional statistical tests with guidance	1	8%
Don't know	1	8%
Cross-validation	1	8%
Comparison of select criteria between AI and non-AI	1	8%
Correlations	1	8%

QUESTION 50: As a general matter, when being used for selection purposes, how much missing data in a candidate's profile, would lead you to conclude that the candidate could not be assessed/scored using the AI tool, algorithm, model, or application?

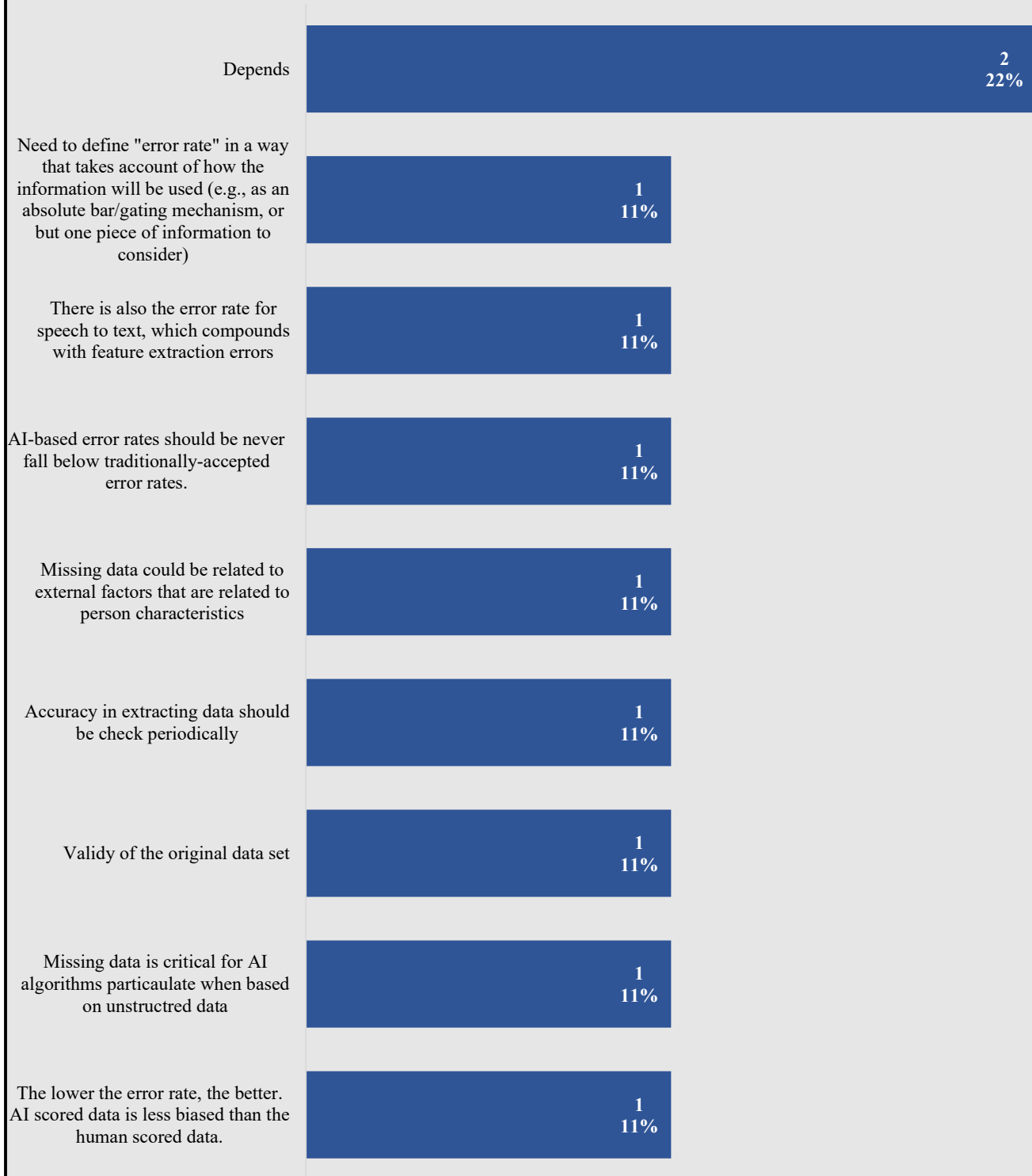| Response	Count	Percent
I do not know or cannot respond	5	26%
More than 50% can be missing to still assess/score the candidate	2	11%
Between 31 and 50% can be missing to still assess/score the candidate		
Between 21 and 30% can be missing to still assess/score the candidate		
Between 11 and 20% can be missing to still assess/score the candidate	1	5%
Between 6 and 10% can be missing to still assess/score the candidate	3	16%
Less than 5% can be missing to still assess/score the candidate (i.e., the candidate has 95% of the data points or features used by the AI model)	6	32%
If there is any missing data for a candidate, scores should not be calculated	2	11%

QUESTION 51: What factors would affect your answer about missing data (e.g., total number of features in the data, type of statistical or logical model underpinning the algorithm)?

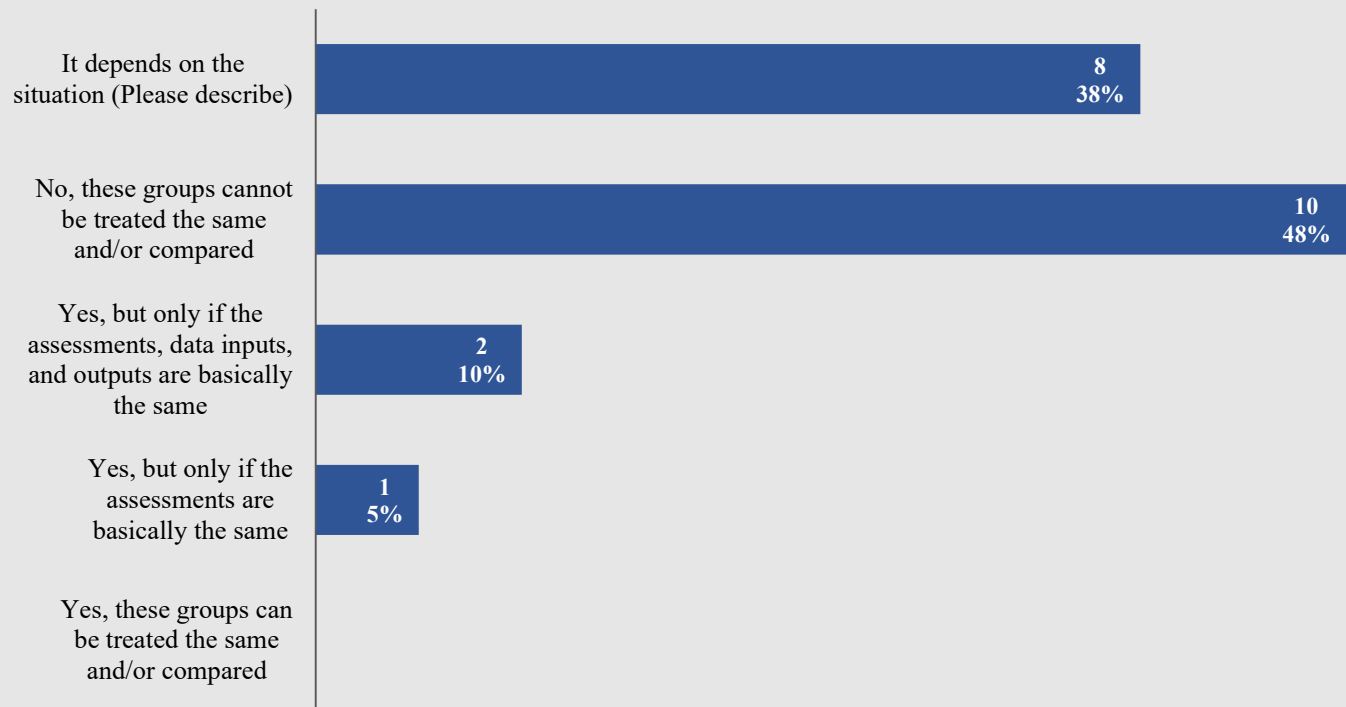| Category	Count	Percent
Specifics Of Missing Data And Feature Weight	5	33%
Dependent On Context	4	27%
Data Collection	1	7%
Use Conceptual And Statistical Techniques, e.g., missForest…	1	7%
Simple Algorithms To Reduce Missing Data	1	7%
Models Need To Consider Missing Data Appropriately	1	7%
Not Unique To AI- Same As Traditional.	1	7%
Unsure	1	7%
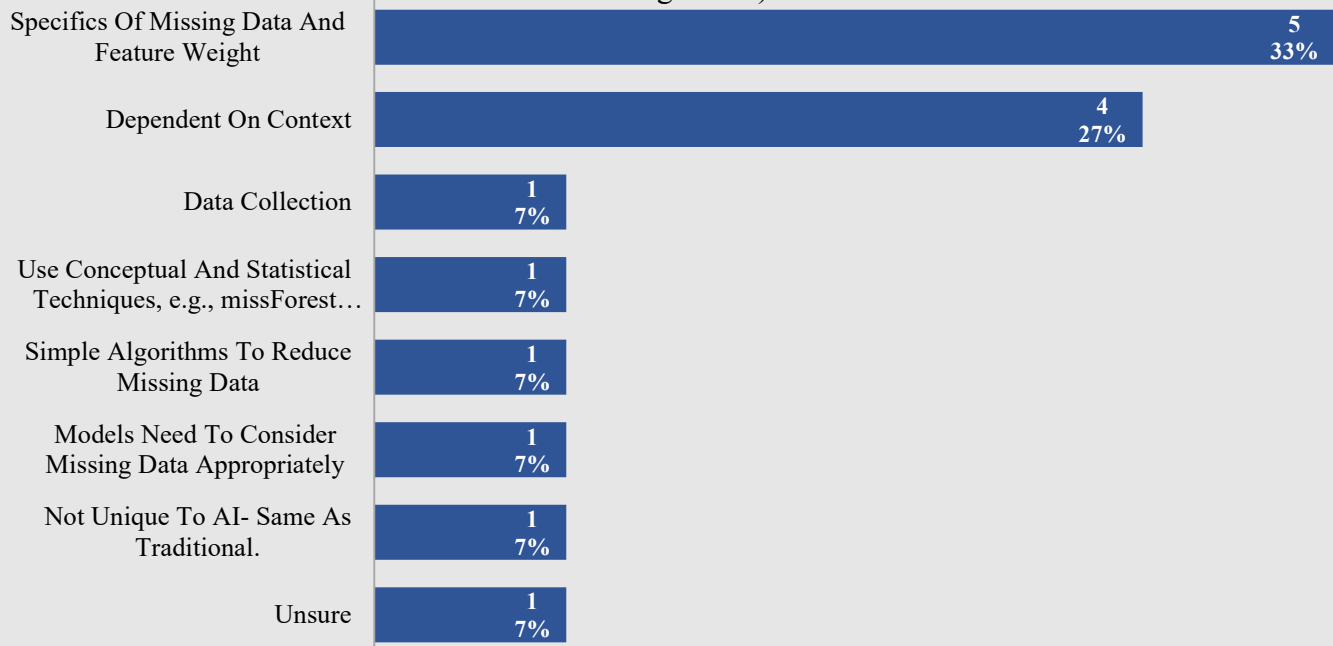
QUESTION 52: When features are being extracted from unstructured text for selection purposes, what error rate is acceptable?
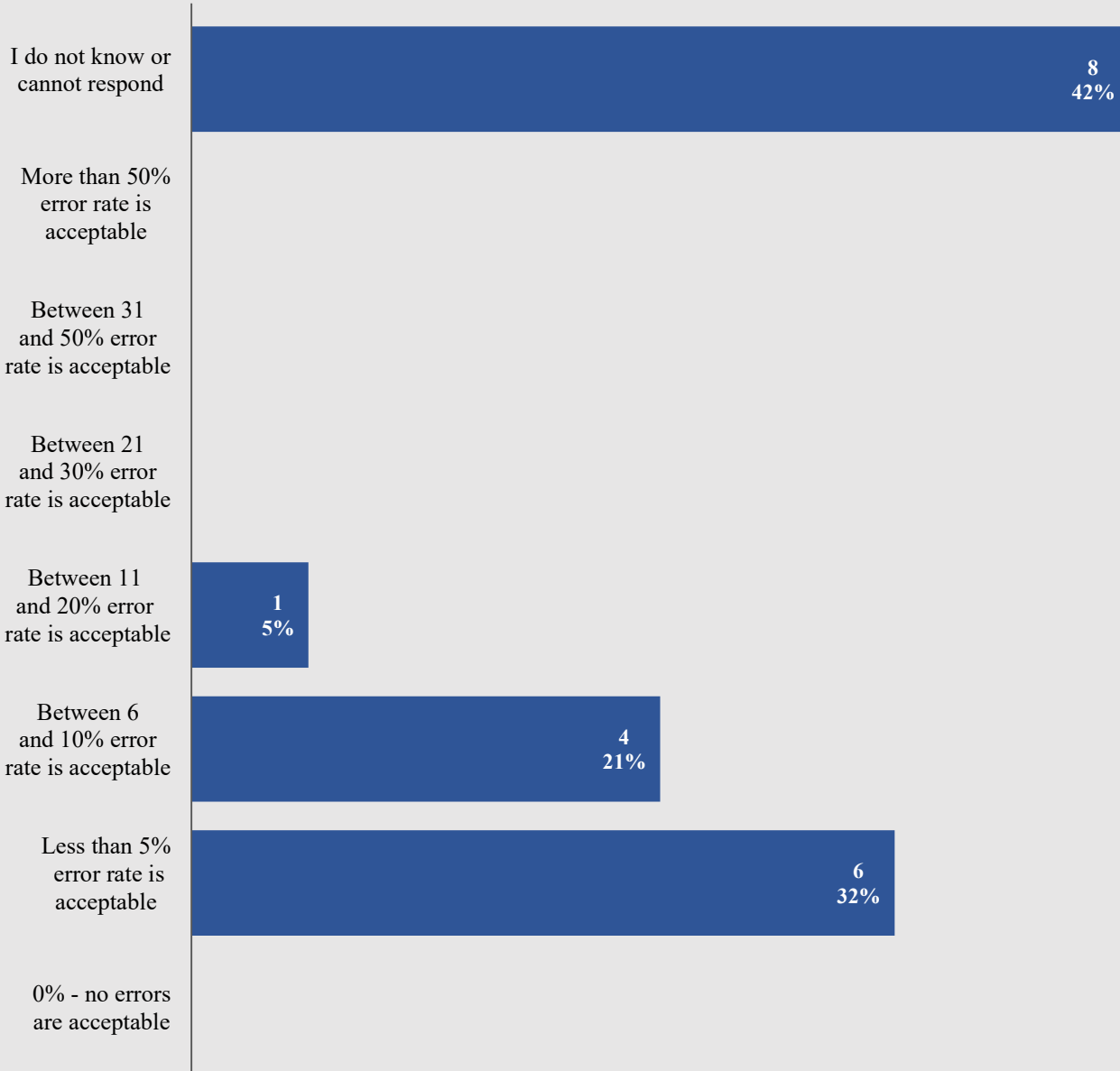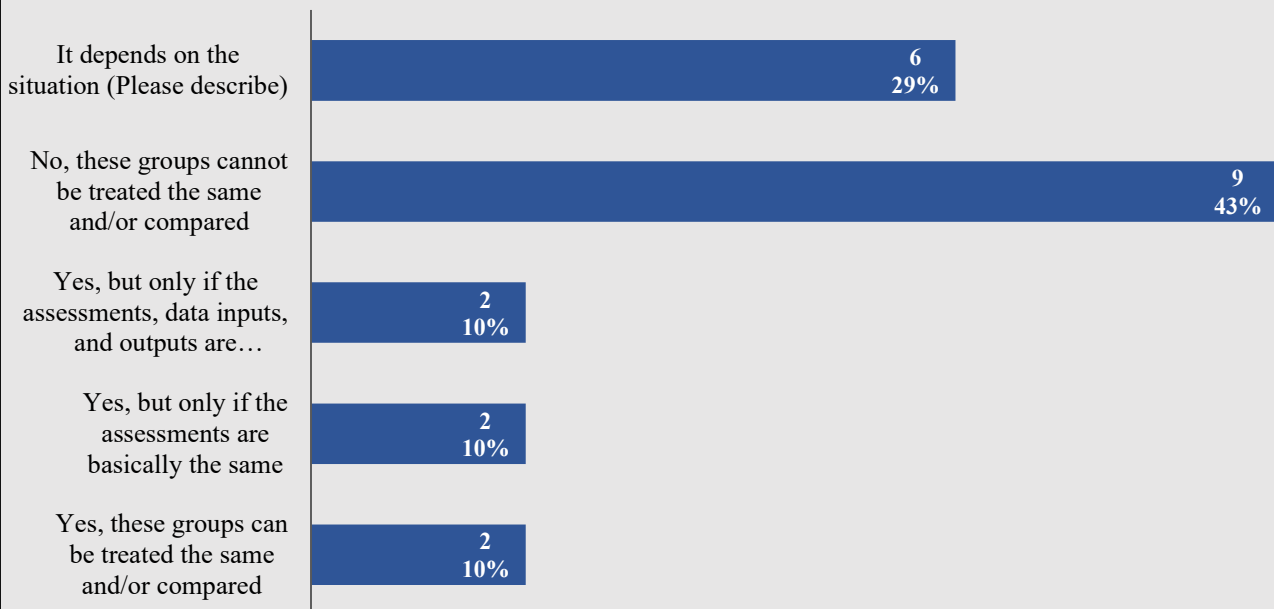
QUESTION 53: Are there additional comments or context you feel are relevant to your responses to the missing data or error rate questions?

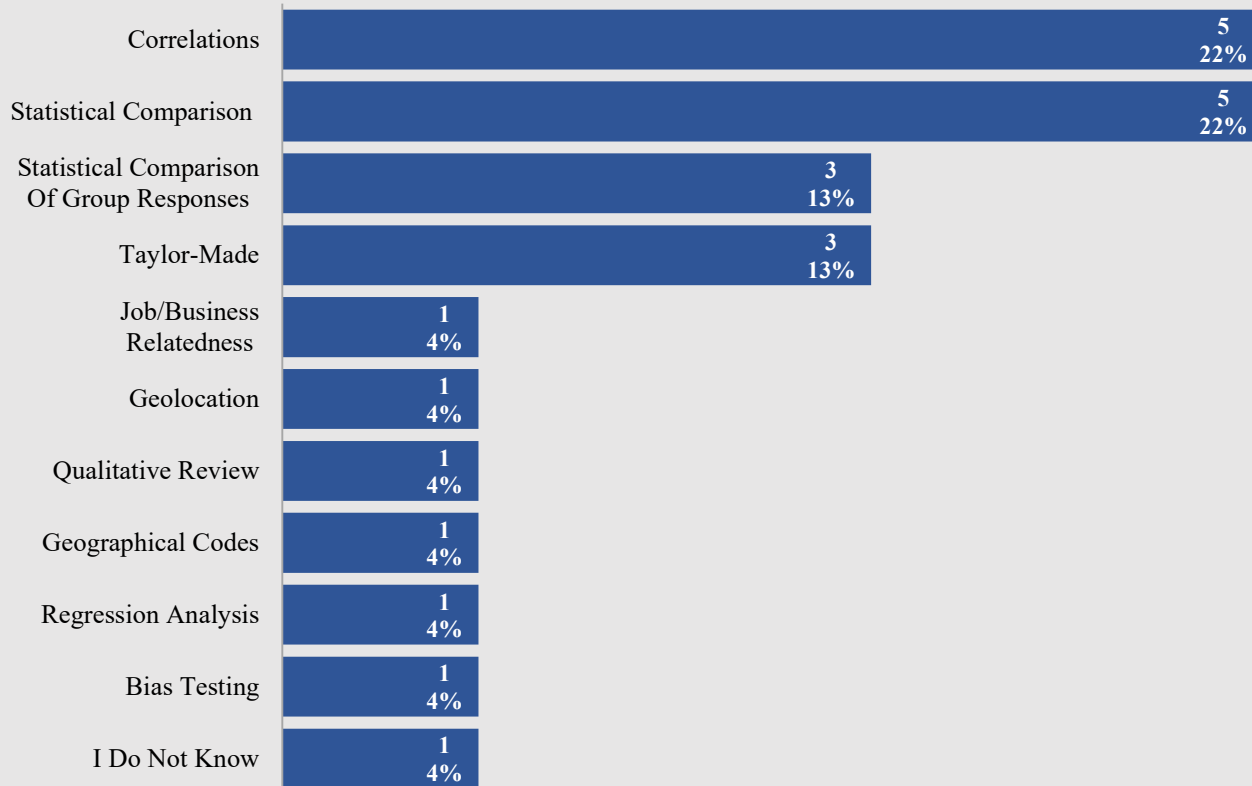|	
Depends	**2 22%**
Need to define "error rate" in a way that takes account of how the information will be used (e.g., as an absolute bar/gating mechanism, or but one piece of information to consider)	**1 11%**
There is also the error rate for speech to text, which compounds with feature extraction errors	**1 11%**
AI-based error rates should be never fall below traditionally-accepted error rates.	**1 11%**
Missing data could be related to external factors that are related to person characteristics	**1 11%**
Accuracy in extracting data should be check periodically	**1 11%**
Validy of the original data set	**1 11%**
Missing data is critical for AI algorithms particaulate when based on unstructred data	**1 11%**
The lower the error rate, the better. AI scored data is less biased than the human scored data.	**1 11%**

QUESTION 54: If some individuals are assessed using an AI tool or model and other individuals are evaluated with a traditional tool or model, should results from these groups be treated the same (e.g., some job applicants participate in a traditional interview and some job applicants participate in an AI-based video interview)?

Response	Count	Percentage
It depends on the situation (Please describe)	8	38%
No, these groups cannot be treated the same and/or compared	10	48%
Yes, but only if the assessments, data inputs, and outputs are basically the same	2	10%
Yes, but only if the assessments are basically the same	1	5%
Yes, these groups can be treated the same and/or compared		

QUESTION 55: If an assessment program initially used a traditional assessment and then switched to an AI-based assessment, should these be treated the same (e.g., selection process started with traditional interviews and later converted to an AI-based video interview)?

Response	Count	Percentage
It depends on the situation (Please describe)	6	29%
No, these groups cannot be treated the same and/or compared	9	43%
Yes, but only if the assessments, data inputs, and outputs are…	2	10%
Yes, but only if the assessments are basically the same	2	10%
Yes, these groups can be treated the same and/or compared	2	10%

QUESTION 56: What technique(s) would you use to determine if a feature constitutes a proxy for protected-class status or other sensitive characteristics (e.g., semantic approaches like reviewing the substantive content of the feature(s) in question to determine whether they convey protected-class information, statistical approaches like measuring collinearity, and so on)?

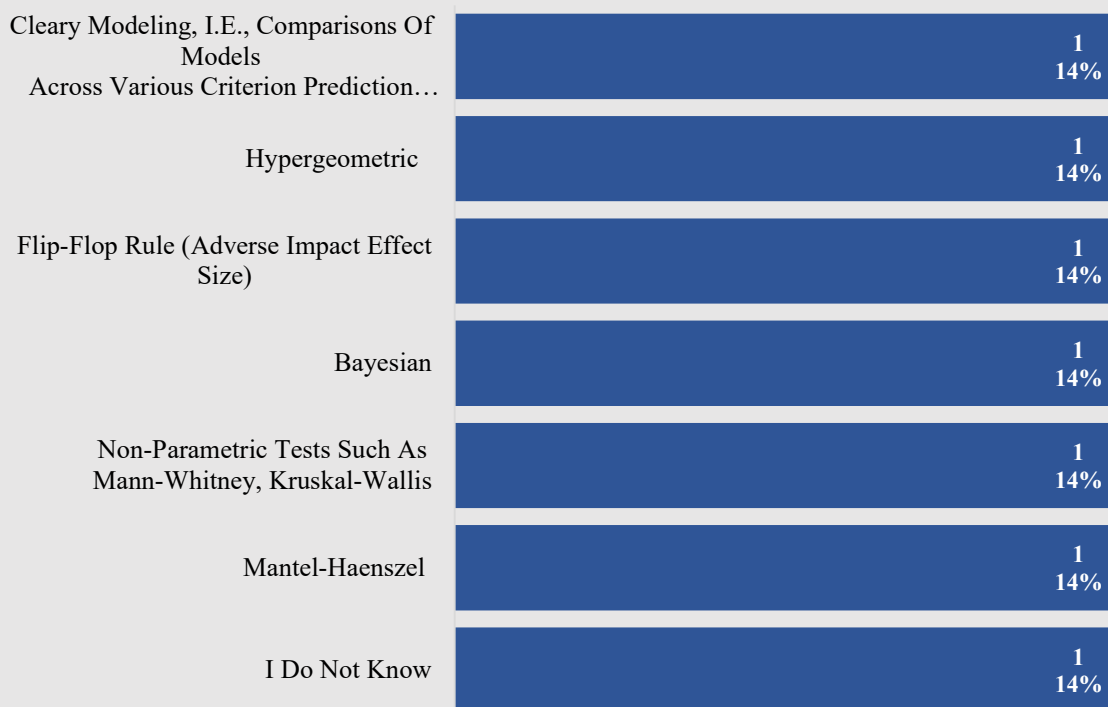| Technique	Count	Percent
Correlations	5	22%
Statistical Comparison	5	22%
Statistical Comparison Of Group Responses	3	13%
Taylor-Made	3	13%
Job/Business Relatedness	1	4%
Geolocation	1	4%
Qualitative Review	1	4%
Geographical Codes	1	4%
Regression Analysis	1	4%
Bias Testing	1	4%
I Do Not Know	1	4%

QUESTION 57: Based on what you are familiar with, how frequently are the following types of methods or statistics used when evaluating the adverse impact of an AI tool, algorithm, model, or application?

Statistical significance tests:

	I do not know	Never	Infrequently	Frequently	Very Frequently
z-tests	(5)	0.0% (0)	18.8% (3)	43.8% (7)	37.5% (6)
Fisher's exact tests / $\chi 2$ tests	(2)	10.5% (2)	21.1% (4)	36.8% (7)	31.6% (6)
t-tests	(5)	6.3% (1)	12.5% (2)	56.3% (9)	25.0% (4)

* Percentages calculated using only those who answered the item and did not indicate I do not know.

QUESTION 58: Based on what you are familiar with, how frequently are the following types of methods or statistics used when evaluating the adverse impact of an AI tool, algorithm, model, or application?
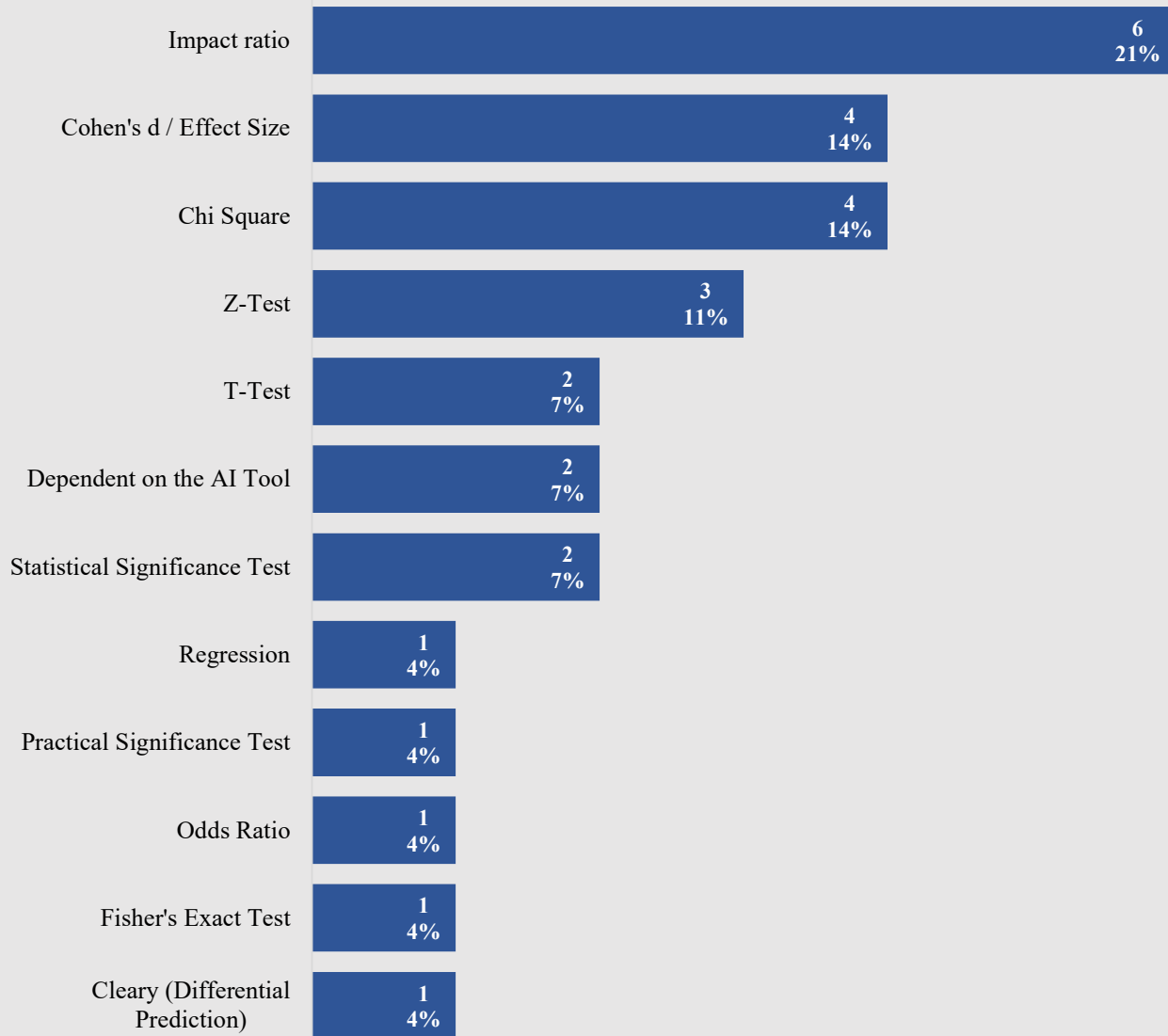
Practical significance tests

	I do not know	Never	Infrequently	Frequently	Very Frequently
Impact Ratios	(2)	5.3% (1)	0.0% (0)	21.1% (4)	73.7% (14)
Odds Ratios	(3)	0.0% (0)	27.8% (5)	55.6% (10)	16.7% (3)
Absolute Differences	(4)	0.0% (0)	31.3% (5)	37.5% (6)	31.3% (5)
Cohen's h	(5)	12.5% (2)	68.8% (11)	6.3% (1)	12.5% (2)
Cohen's d	(5)	6.3% (1)	43.8% (7)	18.8% (3)	31.3% (5)

* Percentages calculated using only those who answered the item and did not indicate I do not know.

QUESTION 59: Are there additional statistical or practical significance tests that are regularly or frequently used to evaluate adverse impact that that were not included above? If so, please list these here.
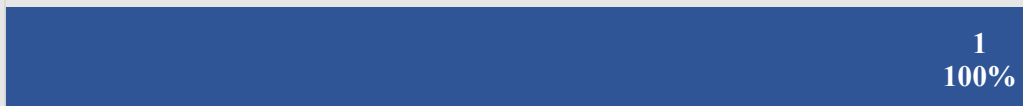
Cleary Modeling, I.E., Comparisons Of Models Across Various Criterion Prediction…	1 14%
Hypergeometric	1 14%
Flip-Flop Rule (Adverse Impact Effect Size)	1 14%
Bayesian	1 14%
Non-Parametric Tests Such As Mann-Whitney, Kruskal-Wallis	1 14%
Mantel-Haenszel	1 14%
I Do Not Know	1 14%

QUESTION 60: Of the methods used for evaluating adverse impact, which do you think are the most useful?

Method	Count	Percent
Impact ratio	6	21%
Cohen's d / Effect Size	4	14%
Chi Square	4	14%
Z-Test	3	11%
T-Test	2	7%
Dependent on the AI Tool	2	7%
Statistical Significance Test	2	7%
Regression	1	4%
Practical Significance Test	1	4%
Odds Ratio	1	4%
Fisher's Exact Test	1	4%
Cleary (Differential Prediction)	1	4%

QUESTION 61: Do you have any additional comments related to this topic?

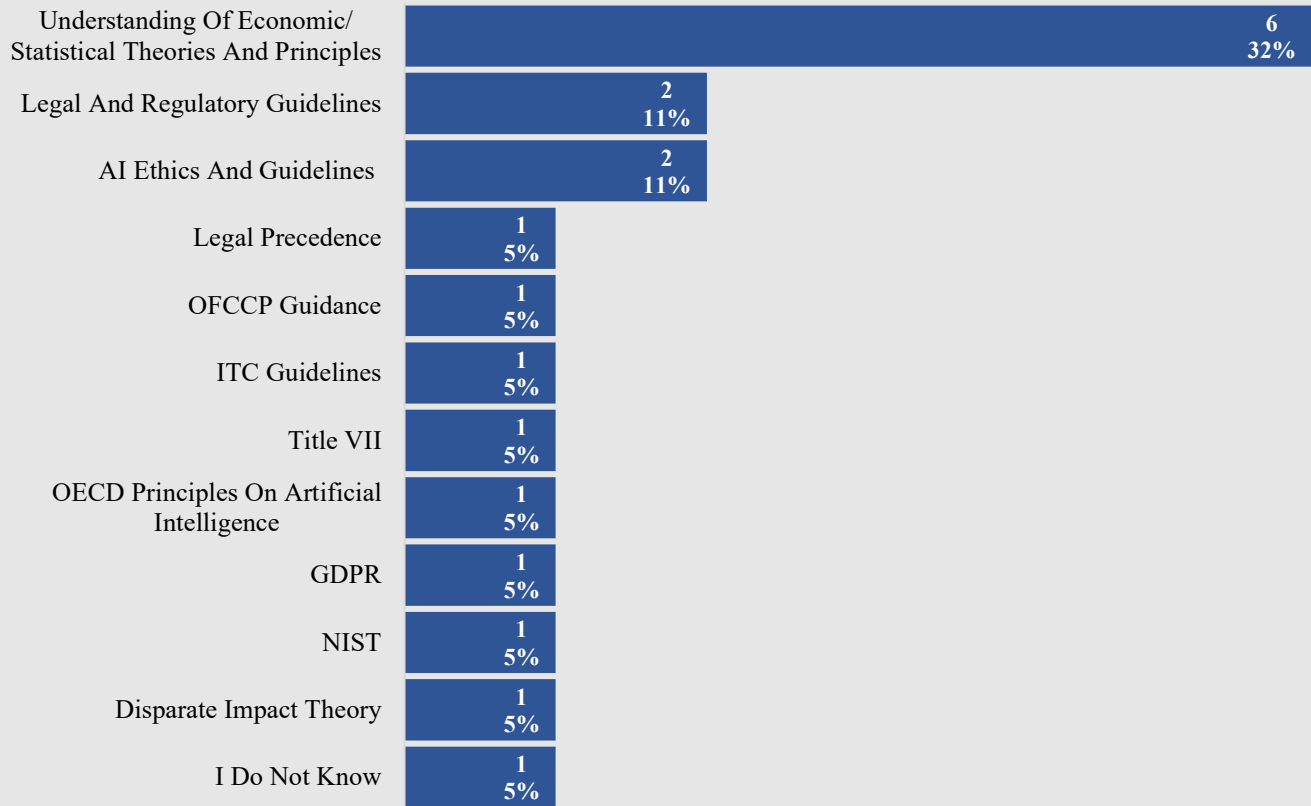| Comment	Count	Percent
The trouble re: metrics is not a lack of opinions, but alignment around generally accepted metrics.…	1	100%

SECTION 9: LEGAL AND EEO ISSUES

QUESTION 62: How would you describe your level of familiarity with this topic?

Response	Count	Percent
I am an expert on this topic – I expect to be able to answer all of these questions	7	20%
I am very familiar with this topic – I should be able to answer all of these questions	13	37%
I am generally knowledgeable about this topic – I could probably answer a lot of…	5	14%
I know a little about this topic – I could potentially answer some of these questions	7	20%
I do not know this topic – I could not answer these types of questions	3	9%

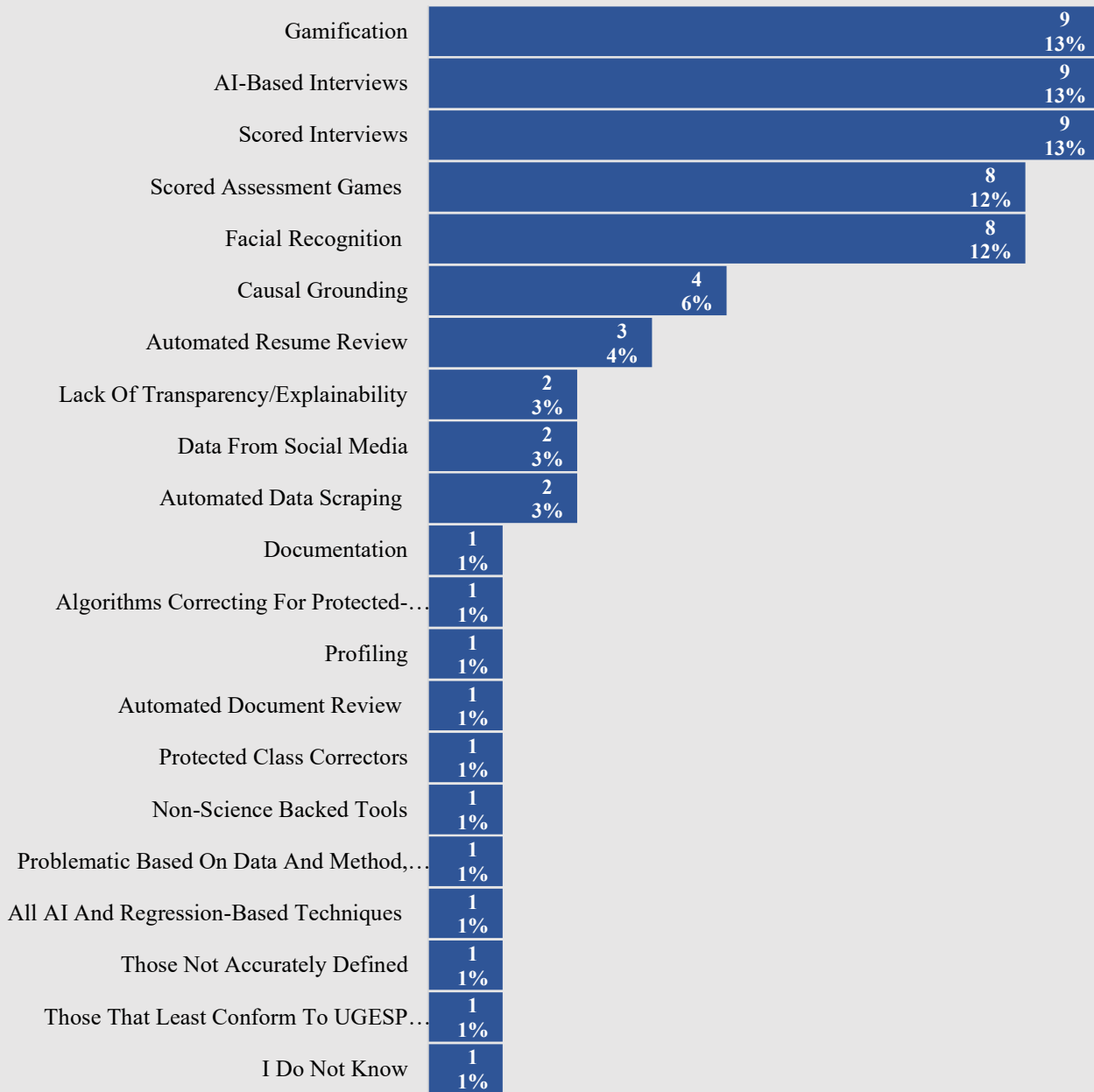| QUESTION 63: How applicable are the following to AI algorithms, models, and applications?					
	I do not know	Does not apply to AI-based applications	Not much use for AI applications	Useful for AI applications	Very useful for AI applications
Unform Guidelines on Employee Selection Procedures	(0)	0.0% (0)	12.9% (4)	41.9% (13)	45.2% (14)
SIOP's Principles for the Validation and Use of Personnel Selection Procedures	(7)	0.0% (0)	4.2% (1)	62.5% (15)	33.3% (8)
AERA et al.'s Standards for Educational and Psychological Testing	(13)	5.6% (1)	27.8% (5)	27.8% (5)	38.9% (7)

* Percentages calculated using only those who answered the item and did not indicate I do not know.

QUESTION 64: Are there other authoritative professional standards that should be used when developing and evaluating AI-based selection procedures?
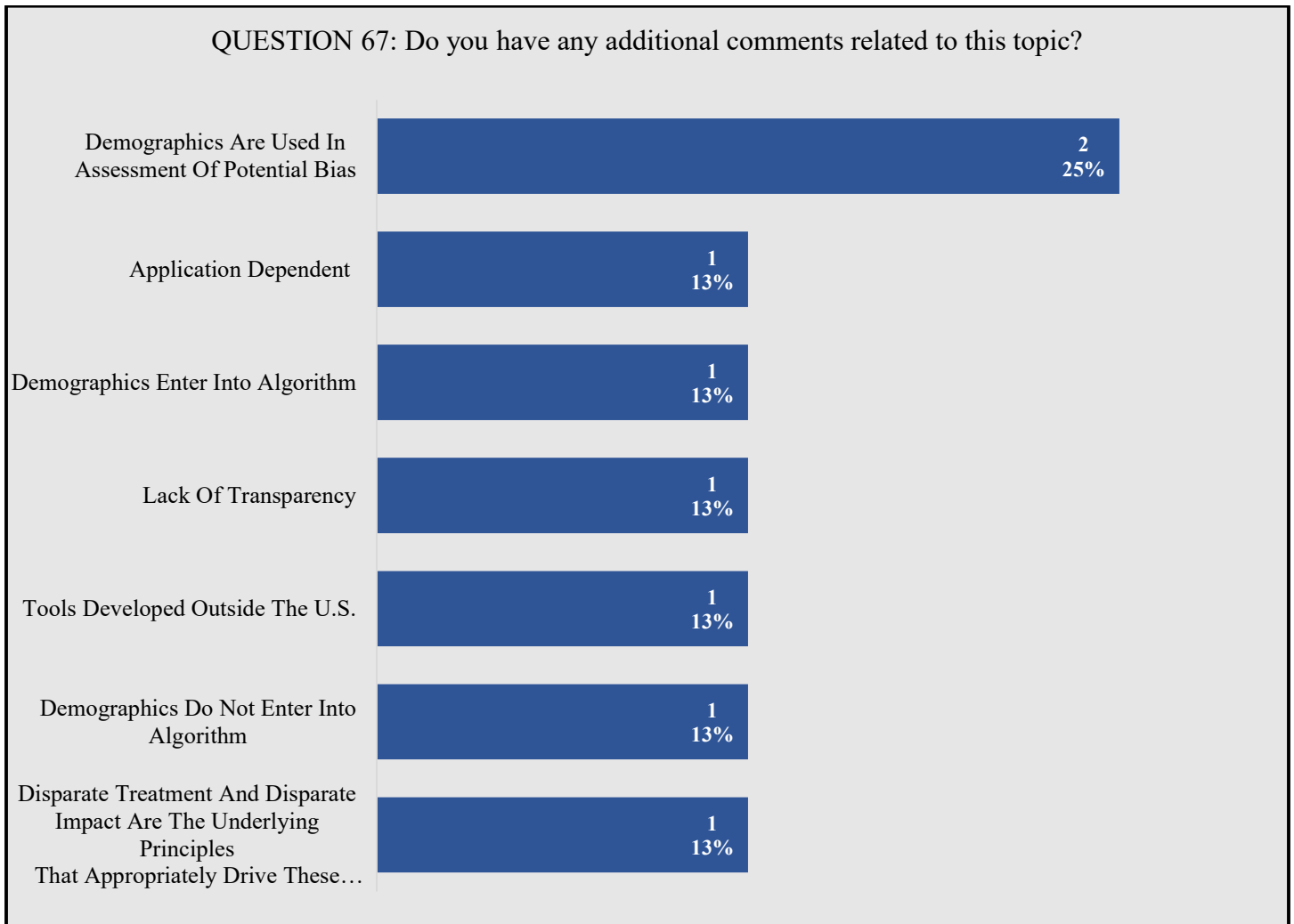
Category	Count	Percent
Understanding Of Economic/ Statistical Theories And Principles	6	32%
Legal And Regulatory Guidelines	2	11%
AI Ethics And Guidelines	2	11%
Legal Precedence	1	5%
OFCCP Guidance	1	5%
ITC Guidelines	1	5%
Title VII	1	5%
OECD Principles On Artificial Intelligence	1	5%
GDPR	1	5%
NIST	1	5%
Disparate Impact Theory	1	5%
I Do Not Know	1	5%

QUESTION 65: What specific types of AI-based applications are particularly controversial/problematic when assessing risk under current legal or professional principles (e.g., AI scored video interviews, gamified tools, automated resume review, algorithms that aggregate information across hiring process)

Category	Count	Percent
Gamification	9	13%
AI-Based Interviews	9	13%
Scored Interviews	9	13%
Scored Assessment Games	8	12%
Facial Recognition	8	12%
Causal Grounding	4	6%
Automated Resume Review	3	4%
Lack Of Transparency/Explainability	2	3%
Data From Social Media	2	3%
Automated Data Scraping	2	3%
Documentation	1	1%
Algorithms Correcting For Protected-…	1	1%
Profiling	1	1%
Automated Document Review	1	1%
Protected Class Correctors	1	1%
Non-Science Backed Tools	1	1%
Problematic Based On Data And Method,…	1	1%
All AI And Regression-Based Techniques	1	1%
Those Not Accurately Defined	1	1%
Those That Least Conform To UGESP…	1	1%
I Do Not Know	1	1%

QUESTION 66: Please indicate your response to the following questions.					
	I do not know	**Never or very infrequently**	**Infrequently**	**Frequently**	**Very Frequently**
How often are demographic variables used in the development of AI-based applications?	(9)	36.4% (8)	27.3% (6)	22.7% (5)	13.6% (3)
How often are known or hypothesized proxies for demographic variables used in the development of AI-based applications?	(10)	9.5% (2)	47.6% (10)	28.6% (6)	14.3% (3)
How often do AI-based applications include variables so closely correlated with demographic variables that effectively they are proxies for demographic variables?	(10)	4.8% (1)	14.3% (3)	57.1% (12)	23.8% (5)

* Percentages calculated using only those who answered the item and did not indicate I do not know.

QUESTION 67: Do you have any additional comments related to this topic?

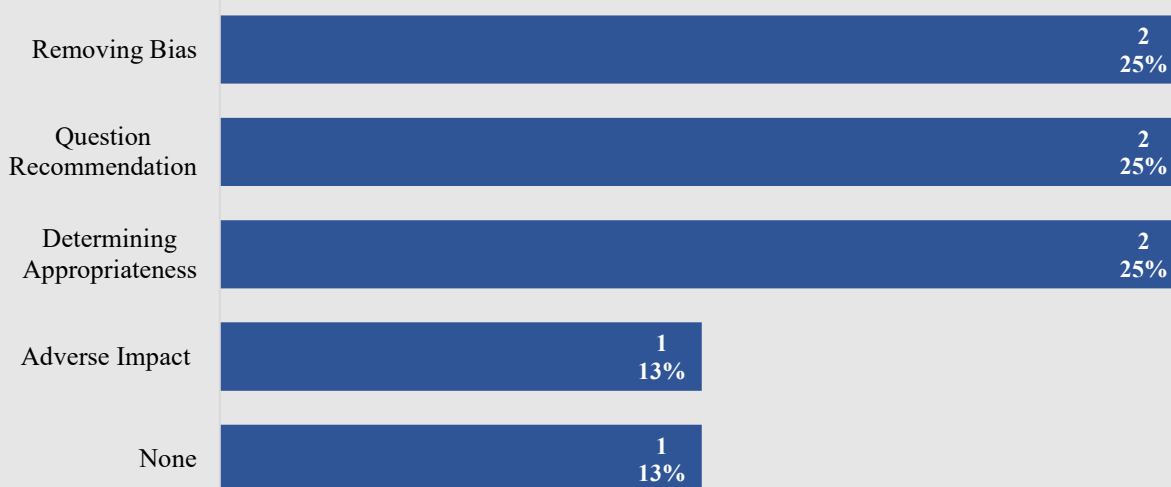|	
Demographics Are Used In Assessment Of Potential Bias	2 — 25%
Application Dependent	1 — 13%
Demographics Enter Into Algorithm	1 — 13%
Lack Of Transparency	1 — 13%
Tools Developed Outside The U.S.	1 — 13%
Demographics Do Not Enter Into Algorithm	1 — 13%
Disparate Treatment And Disparate Impact Are The Underlying Principles That Appropriately Drive These…	1 — 13%

QUESTION 68: Please indicate your response to the following questions. Please note that some questions are specific to selection procedures or selection settings, while others refer to AI applications generally.

	I do not know	Definitely No	Probably No	Maybe	Probably Yes	Definitely Yes
When determining what features will be included in an algorithm, is it important to consider whether each feature is job-related?	(0)	3.2% (1)	3.2% (1)	6.5% (2)	32.3% (10)	54.8% (17)
When determining what features will be included in an algorithm, is it important to consider whether each feature is correlated to the outcome?	(0)	3.2% (1)	3.2% (1)	29.0% (9)	32.3% (10)	32.3% (10)
When determining what features will be included in an algorithm, is it important to consider whether there is a causal inference linking each feature or feature set to the outcome?	(0)	3.2% (1)	6.5% (2)	16.1% (5)	38.7% (12)	35.5% (11)
In a selection setting, is it appropriate to use algorithms or models that dynamically update (i.e., algorithms learn based on new data and self-initiate frequent updates)?	(0)	16.1% (5)	12.9% (4)	48.4% (15)	22.6% (7)	0.0% (0)
In a selection setting, is it appropriate to consider algorithm or model updates (e.g., based on new data) that are deployed via a scheduled implementation?	(0)	0.0% (0)	3.2% (1)	45.2% (14)	35.5% (11)	16.1% (5)
In a selection setting, is it appropriate to use information about applicants that was not directly presented or volunteered by the applicant (e.g., information found on social media, published written materials)?	(2)	27.6% (8)	24.1% (7)	37.9% (11)	10.3% (3)	0.0% (0)
Should there be defined limits on the type of information that can be used to evaluate applicants in a selection setting?	(2)	0.0% (0)	3.4% (1)	13.8% (4)	48.3% (14)	34.5% (10)
When using an ensemble of algorithms as a single selection procedure, is it appropriate to evaluate each algorithm separately for effectiveness and/or adverse impact?	(1)	0.0% (0)	10.0% (3)	23.3% (7)	50.0% (15)	16.7% (5)

QUESTION 68: Please indicate your response to the following questions. Please note that some questions are specific to selection procedures or selection settings, while others refer to AI applications generally.

	I do not know	Definitely No	Probably No	Maybe	Probably Yes	Definitely Yes
Is it appropriate to use demographic data (race/ethnicity, gender) as a variable or feature when developing or training selection-related AI-based applications?	(1)	44.8% (13)	31.0% (9)	10.3% (3)	6.9% (2)	6.9% (2)
Is it appropriate to use demographic data (race/ethnicity, gender) as a variable or feature when developing or training non-selection-related AI-based applications?	(3)	25.0% (7)	28.6% (8)	28.6% (8)	14.3% (4)	3.6% (1)
Should an algorithm be reviewed for features that serve as proxies for protected-class status or other sensitive characteristics?	(0)	3.2% (1)	3.2% (1)	6.5% (2)	32.3% (10)	54.8% (17)
Are selection-related AI-based applications subject to the same adverse impact analysis requirements as traditional selection procedures?	(0)	0.0% (0)	0.0% (0)	0.0% (0)	22.6% (7)	77.4% (24)
Should developers of AI-based applications consider the impact of the application on individuals with a disability?	(0)	0.0% (0)	0.0% (0)	6.5% (2)	16.1% (5)	77.4% (24)
Should employers using AI-based applications consider the impact of the application on individuals with a disability?	(0)	0.0% (0)	0.0% (0)	0.0% (0)	9.7% (3)	90.3% (28)
Are there clear guidelines to determine if conscious or unconscious bias exists in training data?	(3)	42.9% (12)	32.1% (9)	10.7% (3)	10.7% (3)	3.6% (1)
Is it appropriate to "de-bias" an AI-based application (using methods designed to reduce or remove biases from algorithmic, data, or human sources that may introduce errors into important judgments/decisions)?	(5)	0.0% (0)	0.0% (0)	42.3% (11)	23.1% (6)	34.6% (9)
Do you have concerns that de-biasing may result in subgroup norming?	(6)	8.0% (2)	4.0% (1)	28.0% (7)	24.0% (6)	36.0% (9)
Should employers or vendors consider less adverse, similarly valid alternatives to AI-based selection procedures?	(1)	0.0% (0)	3.3% (1)	3.3% (1)	33.3% (10)	60.0% (18)

QUESTION 68: Please indicate your response to the following questions. Please note that some questions are specific to selection procedures or selection settings, while others refer to AI applications generally.

	I do not know	Definitely No	Probably No	Maybe	Probably Yes	Definitely Yes
Should one consider less adverse, similarly valid alternatives when using an ensemble of algorithms (e.g., multiple algorithms applied to the same data to identify convergence) that are considered a single selection procedure?	(1)	0.0% (0)	3.3% (1)	10.0% (3)	33.3% (10)	53.3% (16)

* Percentages calculated using only those who answered the item and did not indicate I do not know.

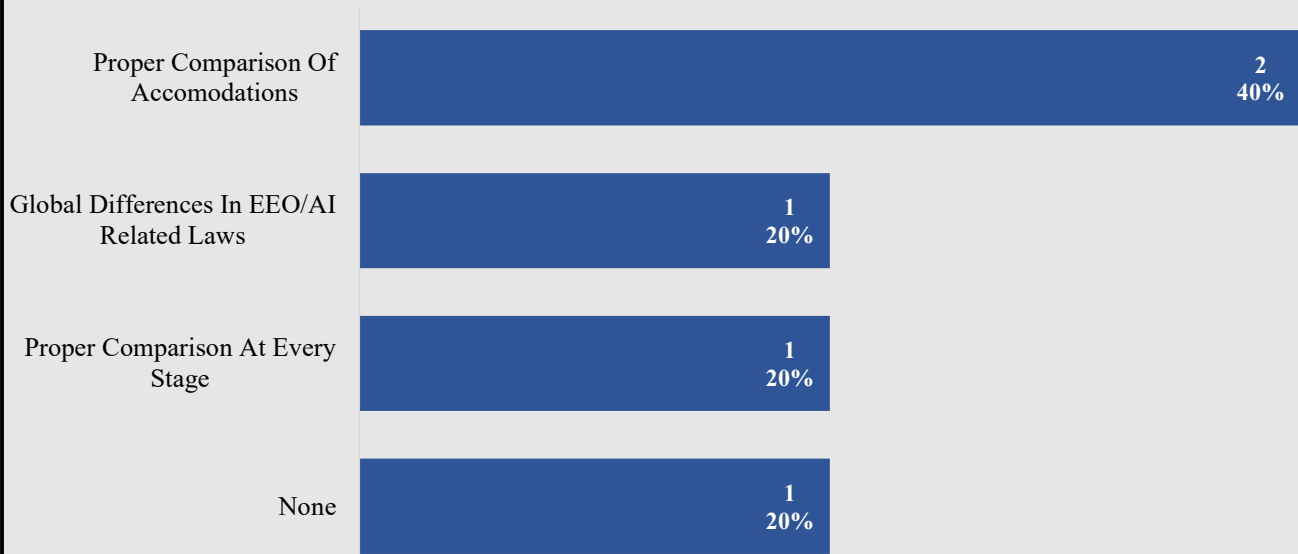QUESTION 69: Do you have any additional comments related to this topic?

QUESTION 70: Please indicate the importance of each of the following to ensuring individuals with disabilities are not disadvantaged when interacting with AI applications.

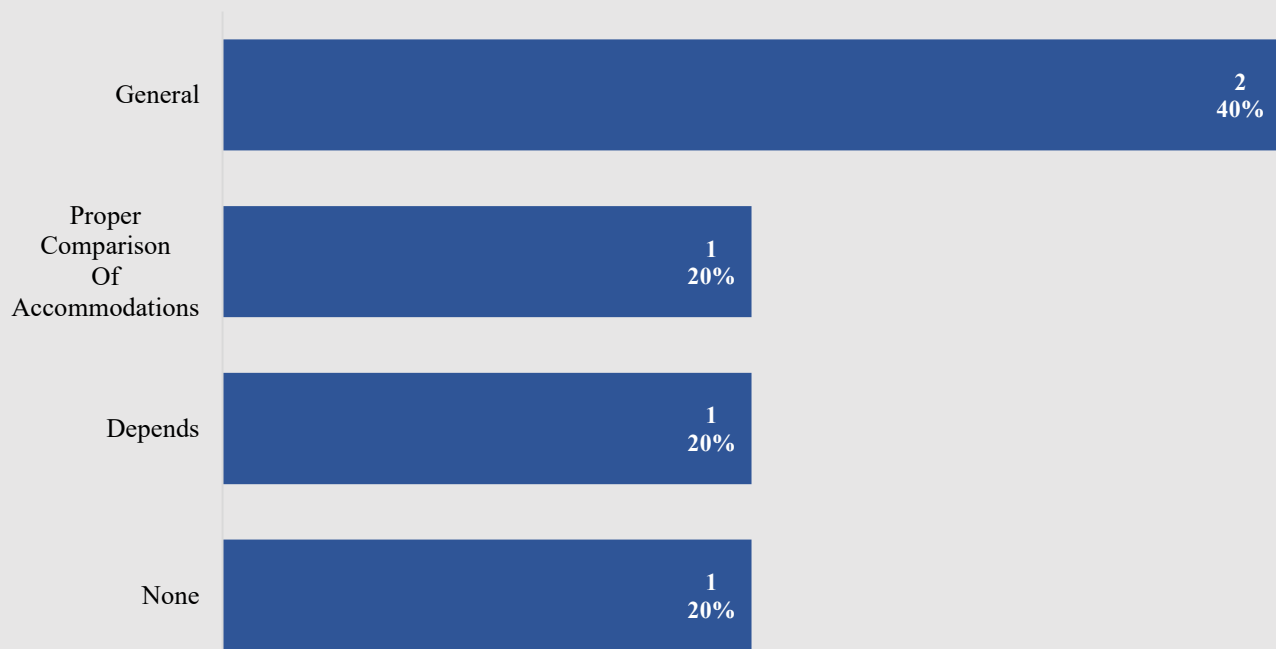|	I do not know	Not Important	Somewhat Important	Important	Very Important	Critically Important
Developers/vendors work to identify and mitigate barriers to accessibility and potential sources of discrimination against disabled workers during the design and development process.	(0)	0.0% (0)	0.0% (0)	6.5% (2)	19.4% (6)	74.2% (23)
Developers/vendors provide guidance to organizations on	(0)	0.0% (0)	6.5% (2)	9.7% (3)	25.8% (8)	58.1% (18)

QUESTION 70: Please indicate the importance of each of the following to ensuring individuals with disabilities are not disadvantaged when interacting with AI applications.

	I do not know	Not Important	Somewhat Important	Important	Very Important	Critically Important
suitable accommodations for their AI applications.						
Organizations inform individuals subjected to AI applications of the nature of the AI application in sufficient detail to evaluate their needs for accommodation.	(0)	3.2% (1)	3.2% (1)	16.1% (5)	19.4% (6)	58.1% (18)
Organizations inform individuals subjected to AI applications with information on accommodations for AI applications and how to request them.	(0)	3.2% (1)	3.2% (1)	0.0% (0)	29.0% (9)	64.5% (20)
Organizations consider how they will respond to likely accommodation requests before they deploy the AI application.	(0)	0.0% (0)	3.2% (1)	12.9% (4)	12.9% (4)	71.0% (22)
Organizations consider how assessment scores from individuals receiving accommodation will be integrated with scores from those who took the assessment without accommodation.	(0)	3.2% (1)	0.0% (0)	19.4% (6)	22.6% (7)	54.8% (17)
Organizations contemplating an AI application consider whether there is another type of application that could be used by all or most subjects with minimal accommodation.	(0)	3.2% (1)	12.9% (4)	12.9% (4)	29.0% (9)	41.9% (13)
Audits of AI applications include consideration of procedural fairness for individuals with disabilities.	(0)	0.0% (0)	6.5% (2)	16.1% (5)	16.1% (5)	61.3% (19)

* Percentages calculated using only those who answered the item and did not indicate I do not know.

QUESTION 71: Are there additional considerations for organizations and developers/vendors that that were not included above? If so, please list these here.

Category	Count	Percent
Proper Comparison Of Accomodations	2	40%
Global Differences In EEO/AI Related Laws	1	20%
Proper Comparison At Every Stage	1	20%
None	1	20%

QUESTION 72: Do you have any additional comments related to this topic?

Category	Count	Percent
General	2	40%
Proper Comparison Of Accommodations	1	20%
Depends	1	20%
None	1	20%

QUESTION 73: How important is it to offer an alternative to the AI-based application (e.g., in-person interview in lieu of video-recorded interview) to ensure that applicants with limited computer access can be considered?
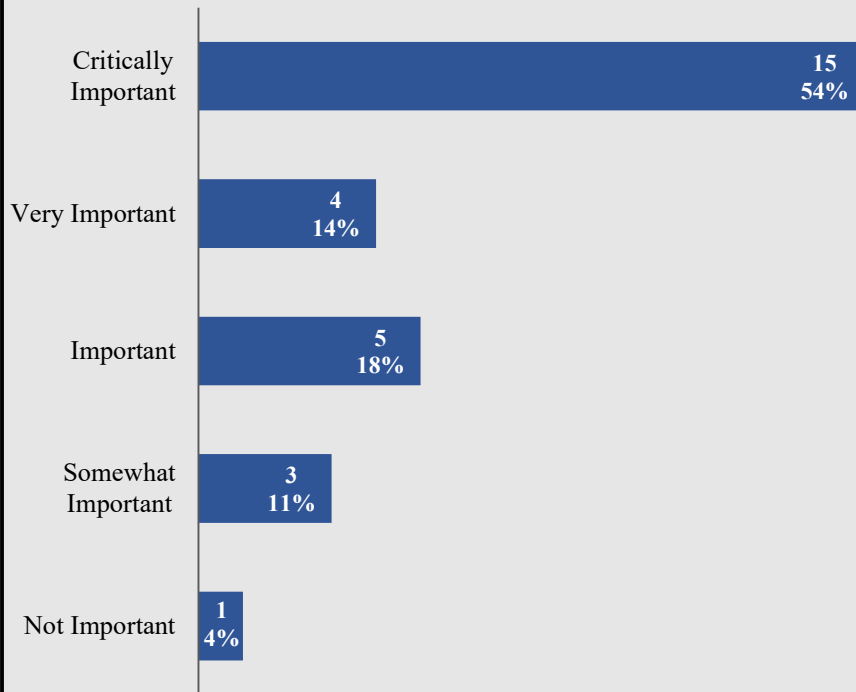
QUESTION 74: How important is it to offer an alternative to the AI-based application to ensure disabled applicants can be considered?

Category	Count	Percentage
Critically Important	21	72%
Very Important	6	21%
Important	1	3%
Somewhat Important		
Not Important	1	3%
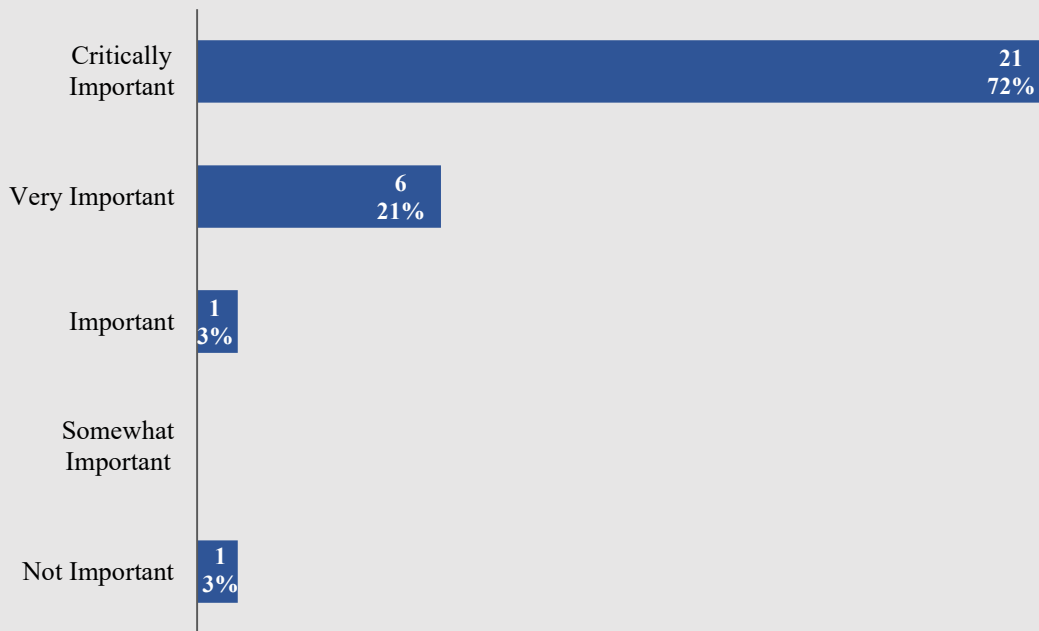
SECTION 10: DOCUMENTATION
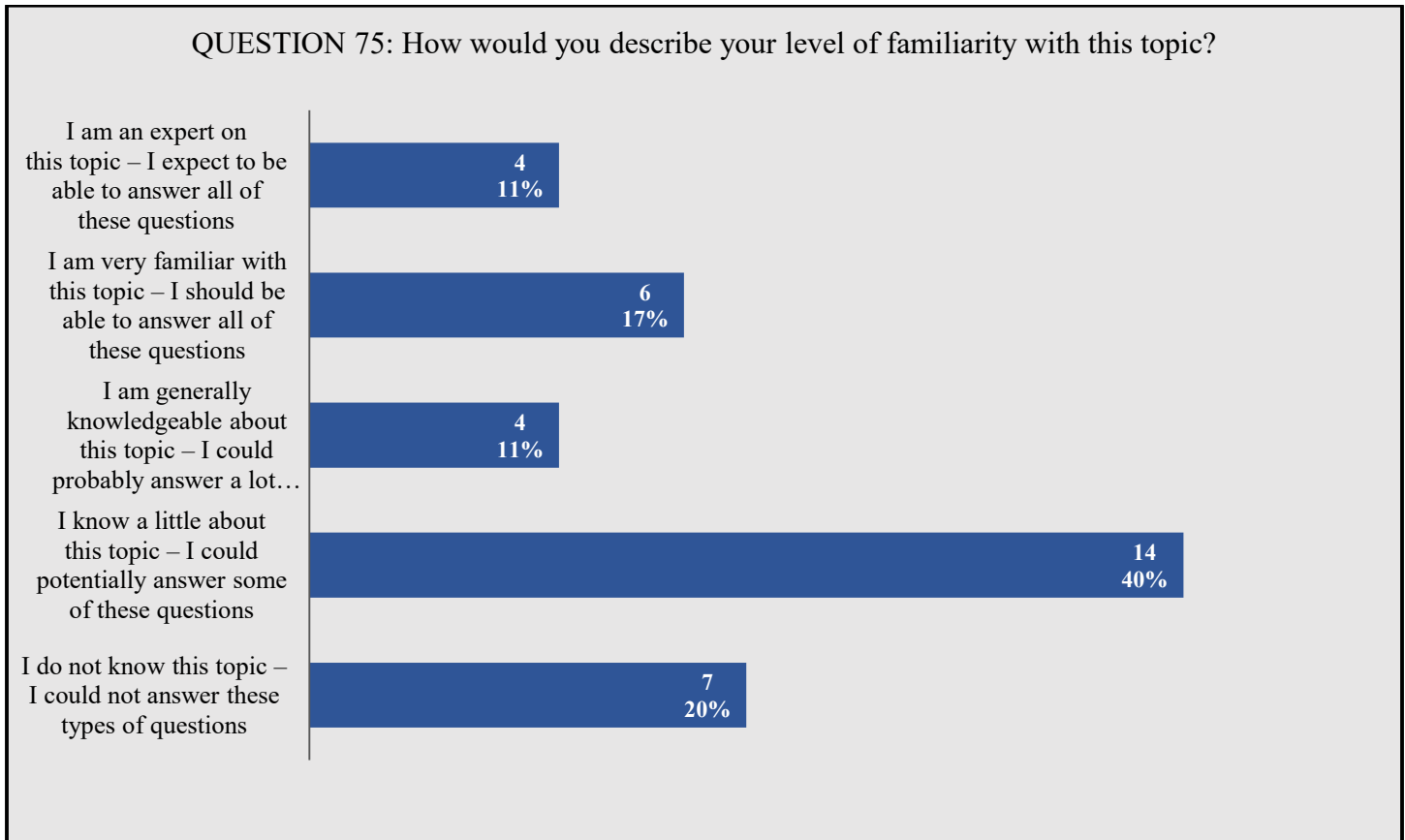
QUESTION 75: How would you describe your level of familiarity with this topic?

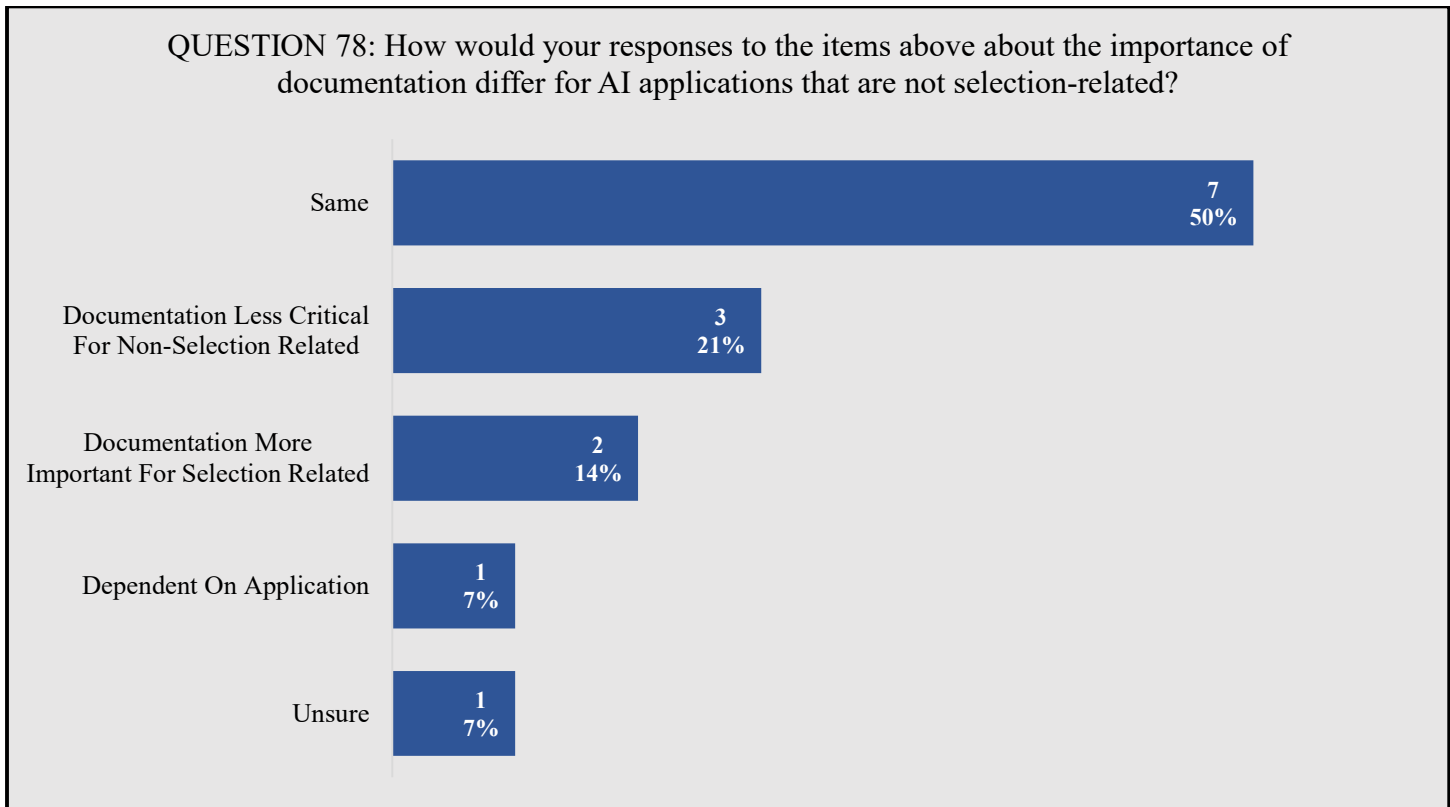|	
I am an expert on this topic – I expect to be able to answer all of these questions	4 / 11%
I am very familiar with this topic – I should be able to answer all of these questions	6 / 17%
I am generally knowledgeable about this topic – I could probably answer a lot…	4 / 11%
I know a little about this topic – I could potentially answer some of these questions	14 / 40%
I do not know this topic – I could not answer these types of questions	7 / 20%

QUESTION 76: Please indicate how important each of these items are to document for selection-related AI applications.

	I do not know	Not Important	Somewhat Important	Important	Very Important	Critically Important
Documenting why and how the application was developed (e.g., purpose of the application, choice of method or algorithm).	(1)	0.0% (0)	3.8% (1)	3.8% (1)	50.0% (13)	42.3% (11)
Documenting data use by an application (e.g., data sources, features considered, features included in model/algorithm build, how data are used).	(0)	0.0% (0)	0.0% (0)	3.7% (1)	48.1% (13)	48.1% (13)
Documenting the theoretical, logical, and/or causal reasons to believe the particular features used in a model or algorithm are related to the requirements of the job.	(0)	0.0% (0)	0.0% (0)	11.1% (3)	37.0% (10)	51.9% (14)
Documenting the job-relevance of the application (in selection contexts).	(0)	0.0% (0)	0.0% (0)	7.4% (2)	22.2% (6)	70.4% (19)

QUESTION 76: Please indicate how important each of these items are to document for selection-related AI applications.	I do not know	Not Important	Somewhat Important	Important	Very Important	Critically Important
Documenting samples used in application development or monitoring.	(2)	0.0% (0)	8.0% (2)	16.0% (4)	28.0% (7)	48.0% (12)
Documenting the normative information used to interpret scores.	(3)	0.0% (0)	4.2% (1)	12.5% (3)	33.3% (8)	50.0% (12)
Documenting changes to the normative information.	(3)	0.0% (0)	8.3% (2)	8.3% (2)	37.5% (9)	45.8% (11)
Documenting data cleaning activities.	(3)	0.0% (0)	4.2% (1)	20.8% (5)	33.3% (8)	41.7% (10)
Documenting decisions made in the development of models or algorithms (e.g., choice of modeling technique, comparison of metrics for multiple models, and so on).	(1)	0.0% (0)	7.7% (2)	19.2% (5)	38.5% (10)	34.6% (9)
Documenting general scoring models used and relative importance of features or feature groupings.	(1)	0.0% (0)	0.0% (0)	26.9% (7)	34.6% (9)	38.5% (10)
Documenting changes to an algorithm (e.g., why changed, when changed, to what effect) for models where change is controlled by the developer (i.e., an infrequent update that a human initiates).	(1)	3.8% (1)	0.0% (0)	23.1% (6)	23.1% (6)	50.0% (13)
Documenting changes to an algorithm (e.g., why changed, when changed, to what effect) for dynamic or fluid models (i.e., frequent updates that are initiated without human involvement).	(2)	0.0% (0)	4.0% (1)	16.0% (4)	24.0% (6)	56.0% (14)
Documenting how application output is to be interpreted and used.	(1)	0.0% (0)	0.0% (0)	12.0% (3)	20.0% (5)	68.0% (17)
Providing explanatory material to be given to individuals subjected to the application.	(0)	0.0% (0)	14.8% (4)	29.6% (8)	25.9% (7)	29.6% (8)
Documenting the effectiveness (e.g., validity, ability to predict target outcome) of the application as used (i.e., documenting evidence showing the application works and achieves the intended goal).	(0)	0.0% (0)	3.7% (1)	3.7% (1)	29.6% (8)	63.0% (17)
Documenting evidence of adverse impact.	(1)	0.0% (0)	0.0% (0)	0.0% (0)	19.2% (5)	80.8% (21)

QUESTION 76: Please indicate how important each of these items are to document for selection-related AI applications.

	I do not know	Not Important	Somewhat Important	Important	Very Important	Critically Important
Documenting efforts taken to mitigate adverse impact.	(0)	0.0% (0)	0.0% (0)	3.8% (1)	30.8% (8)	65.4% (17)
Documenting less adverse, similarly valid alternatives to the deployed application (e.g., other applications, inclusion/exclusion of features, uses of feature weights).	(0)	3.7% (1)	7.4% (2)	14.8% (4)	33.3% (9)	40.7% (11)
Documenting the reliability of the application as used.	(0)	0.0% (0)	0.0% (0)	11.5% (3)	38.5% (10)	50.0% (13)
Documenting the statistical fairness of the application as used.	(3)	0.0% (0)	4.2% (1)	8.3% (2)	33.3% (8)	54.2% (13)
Documenting the procedural or perceived fairness of the application as used.	(3)	4.2% (1)	12.5% (3)	16.7% (4)	33.3% (8)	33.3% (8)

QUESTION 78: How would your responses to the items above about the importance of documentation differ for AI applications that are not selection-related?

QUESTION 77: Are there additional areas of documentation that should be considered that were not included above? If so, please list these here.

Category	Count	Percentage
Causal Grounding	1	25%
Data & Process And Documentation Transparency	1	25%
Equitible Outcomes Acommodations	1	25%
None	1	25%

QUESTION 79: How frequently should technical documentation for selection-related AI applications be updated? CHECK ALL THAT APPLY. Note that some choices are mutually exclusive.

Category	Count	Percentage
Documentation should be updated on a regular basis – every few years	13	50%
Documentation should be updated on a regular basis – about once a year	11	42%
Documentation should be updated if significant legal risk is introduced (e.g., employment discrimination suit)	18	69%
Documentation should be updated if there are substantial model changes (e.g., data, features, weights)	24	92%
Documentation should be updated if there are substantial job changes	21	81%
Once initial documentation is complete, updating it is rarely necessary	1	4%

QUESTION 80: What aspects of the development and validation of an AI-based application are proprietary and do not need to be shared beyond the developer or vendor?

Response	Count	Percent
Algorithm Makeup Need Not Be Shared	5	38%
Results And Outcomes Should Be Shared	2	15%
Development And Validation	1	8%
All Information Should Be Made Available To Auditors	1	8%
All Information Documented And Shared	1	8%
Multiple Groups To Consider	1	8%
Undecided	2	15%
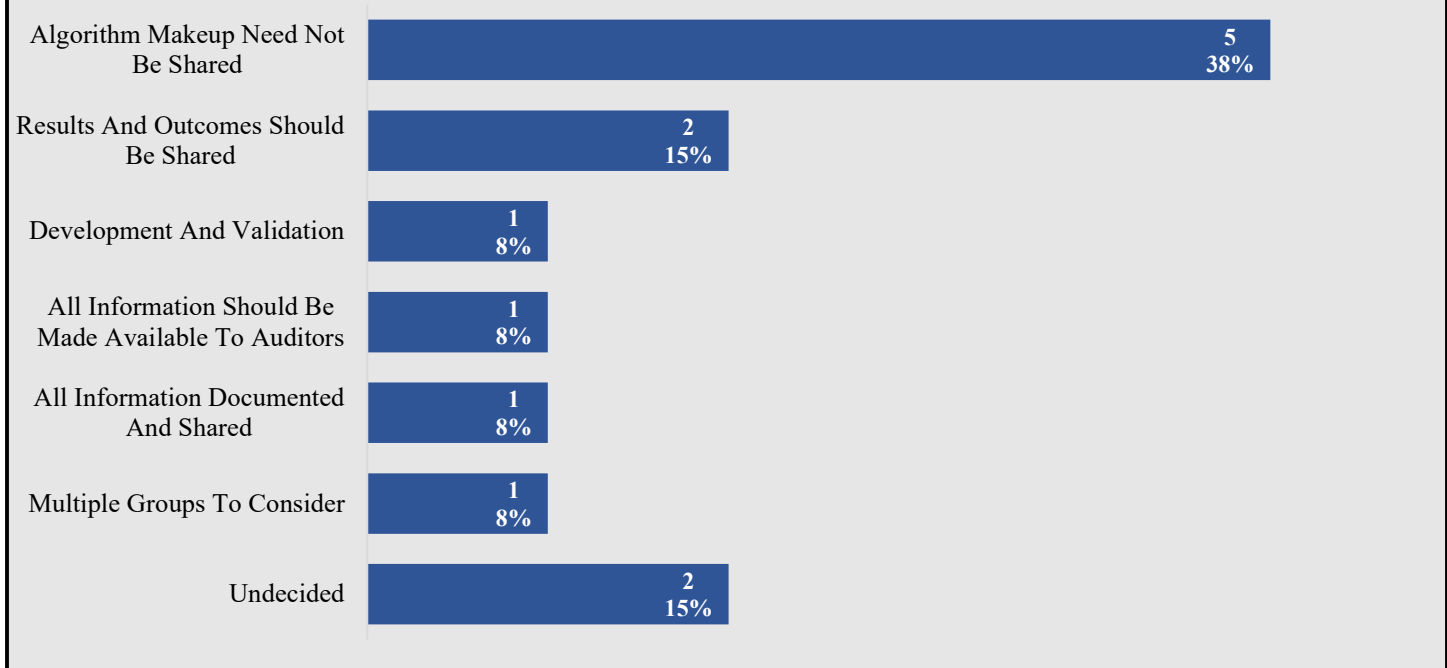
QUESTION 81: When is a local validation study necessary to justify the inferences made from scores resulting from a selection-related AI application?

Response	Count	Percent
Always	4	29%
If Results Show Adverse Impact	4	29%
If The Test Is The Only Selection Metric	1	7%
If Substantially Different From Existing Empirical Data	1	7%
In The Absence Of Causal Grounding	1	7%
Almost Always	1	7%
Where Feasible	1	7%
Undecided	1	7%
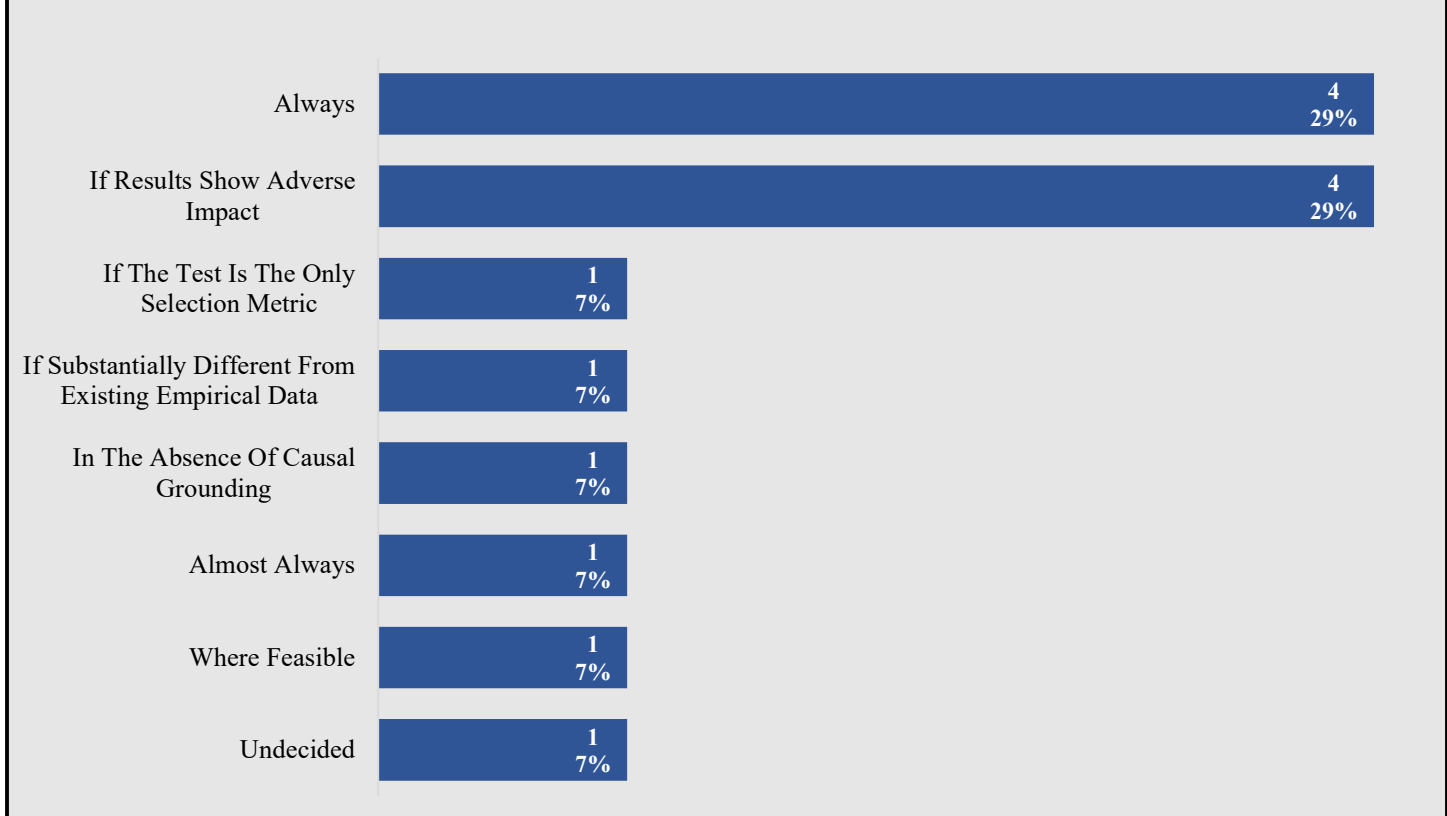
QUESTION 82: Do you have any additional comments related to this topic?

Category	Count	Percentage
Depth Of Documentation	2	25%
Continous Monitoring And Frequent Updating Of…	1	13%
Question Clarification	1	13%
Monitoring	1	13%
Update Proceedures	1	13%
Not Sure	1	13%
None	1	13%

SECTION 11: AUDITS

QUESTION 83: How would you describe your level of familiarity with this topic?

- I am an expert on this topic – I expect to be able to answer all of these questions: 5 / 14%
- I am very familiar with this topic – I should be able to answer all of these questions: 6 / 17%
- I am generally knowledgeable about this topic – I could probably answer a lot of…: 7 / 20%
- I know a little about this topic – I could potentially answer some of these questions: 13 / 37%
- I do not know this topic – I could not answer these types of questions: 4 / 11%
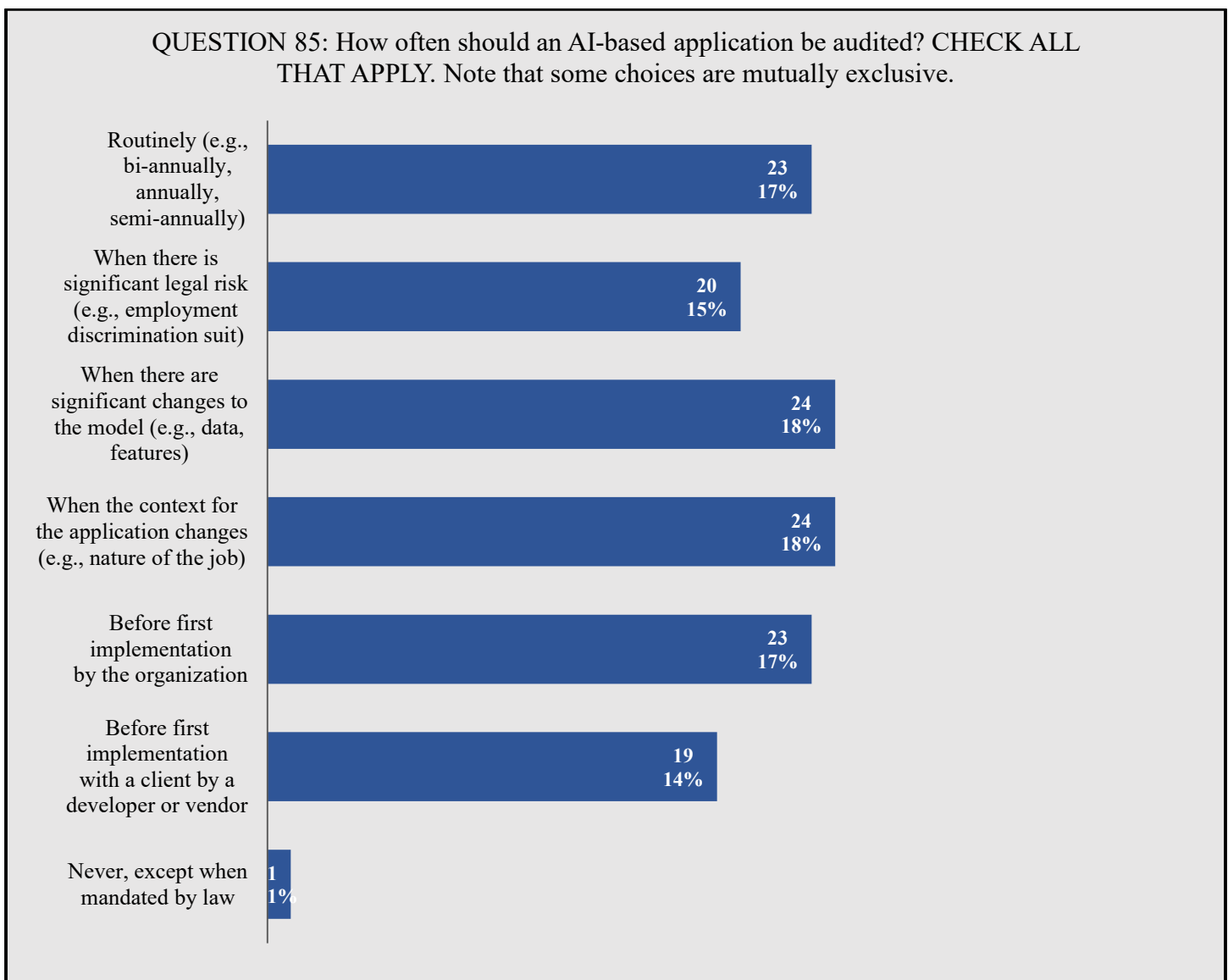
QUESTION 84: How important is each of the following in an audit of an AI application?

	I do not know	Not Important	Somewhat Important	Important	Very Important	Critically Important
Explicit mention or use of a code of ethics or professional practice that guided AI application development.	(0)	0.0% (0)	20.0% (6)	33.3% (10)	26.7% (8)	20.0% (6)
Design and development of the AI application had multidisciplinary involvement (e.g., psychologists, data scientists, attorneys).	(0)	6.7% (2)	20.0% (6)	13.3% (4)	33.3% (10)	26.7% (8)
A logical or intuitive relationship between features in the model and the target outcome.	(0)	6.7% (2)	6.7% (2)	10.0% (3)	36.7% (11)	40.0% (12)
A clear rationale for the choice of the initial algorithms and the features considered for inclusion in the model.	(1)	0.0% (0)	0.0% (0)	10.3% (3)	44.8% (13)	44.8% (13)
The extent to which the model is transparent and is explainable to users.	(0)	0.0% (0)	13.3% (4)	23.3% (7)	36.7% (11)	26.7% (8)

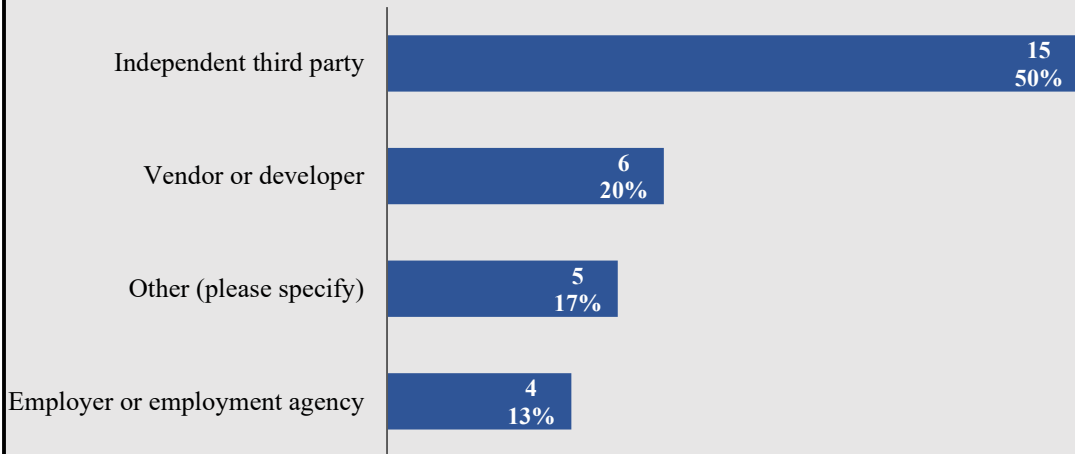| QUESTION 84: How important is each of the following in an audit of an AI application?						
	I do not know	Not Important	Somewhat Important	Important	Very Important	Critically Important
A clear rationale for subsequent modification of algorithms, the features included, and the relative weights of features in the model.	(1)	0.0% (0)	0.0% (0)	6.9% (2)	51.7% (15)	41.4% (12)
The AI application demonstrates predictive accuracy (e.g., effectiveness, validity).	(0)	0.0% (0)	0.0% (0)	13.3% (4)	26.7% (8)	60.0% (18)
Indicators of utility of the AI application (e.g., payoff to user, payoff relative to alternatives to the AI application).	(0)	0.0% (0)	10.0% (3)	26.7% (8)	40.0% (12)	23.3% (7)
Indication of the statistical methods and standards used to evaluate if the application is effective and the appropriate interpretation of those results.	(0)	3.3% (1)	3.3% (1)	3.3% (1)	23.3% (7)	66.7% (20)
The consistency between the effectiveness of the AI algorithm and the intended operational use of the outputs of the AI algorithm.	(2)	0.0% (0)	3.6% (1)	10.7% (3)	35.7% (10)	50.0% (14)
Tradeoffs between the effectiveness of the AI algorithm and the explainability and interpretability of the AI algorithm.	(1)	0.0% (0)	3.4% (1)	37.9% (11)	34.5% (10)	24.1% (7)
Explicit definitions of fairness and bias.	(0)	3.3% (1)	6.7% (2)	3.3% (1)	26.7% (8)	60.0% (18)
Adverse impact for any of the relevant demographic groups.	(0)	0.0% (0)	0.0% (0)	6.7% (2)	13.3% (4)	80.0% (24)
Indication of the statistical methods and standards for judging if the application is fair and unbiased and the appropriate interpretation of those results.	(1)	0.0% (0)	3.4% (1)	3.4% (1)	17.2% (5)	75.9% (22)
Indication of differential prediction (i.e., whether there is more accuracy for some groups than others).	(0)	3.4% (1)	6.9% (2)	13.8% (4)	24.1% (7)	51.7% (15)
Scoring adjustments to address issues related to differential prediction or adverse impact.	(2)	0.0% (0)	0.0% (0)	14.3% (4)	32.1% (9)	53.6% (15)
Indication of lack of opportunity for accurate assessment for some portions of the population eligible to be assessed.	(2)	0.0% (0)	3.6% (1)	10.7% (3)	32.1% (9)	53.6% (15)

QUESTION 84: How important is each of the following in an audit of an AI application?

	I do not know	Not Important	Somewhat Important	Important	Very Important	Critically Important
Indication of lack of opportunity for individuals to present information about themselves.	(3)	3.7% (1)	14.8% (4)	7.4% (2)	44.4% (12)	29.6% (8)
The effect on results for individuals that do not have complete data available.	(4)	0.0% (0)	3.8% (1)	11.5% (3)	38.5% (10)	46.2% (12)
The intended dissemination of audit results.	(1)	0.0% (0)	13.8% (4)	6.9% (2)	27.6% (8)	51.7% (15)

QUESTION 85: How often should an AI-based application be audited? CHECK ALL THAT APPLY. Note that some choices are mutually exclusive.

Category	Count	Percent
Routinely (e.g., bi-annually, annually, semi-annually)	23	17%
When there is significant legal risk (e.g., employment discrimination suit)	20	15%
When there are significant changes to the model (e.g., data, features)	24	18%
When the context for the application changes (e.g., nature of the job)	24	18%
Before first implementation by the organization	23	17%
Before first implementation with a client by a developer or vendor	19	14%
Never, except when mandated by law	1	1%

QUESTION 86: Who bears responsibility for auditing an AI-based application?

Response	Count	Percent
Both	22	73%
Employer or employment agency	6	20%
Vendor or developer	2	7%
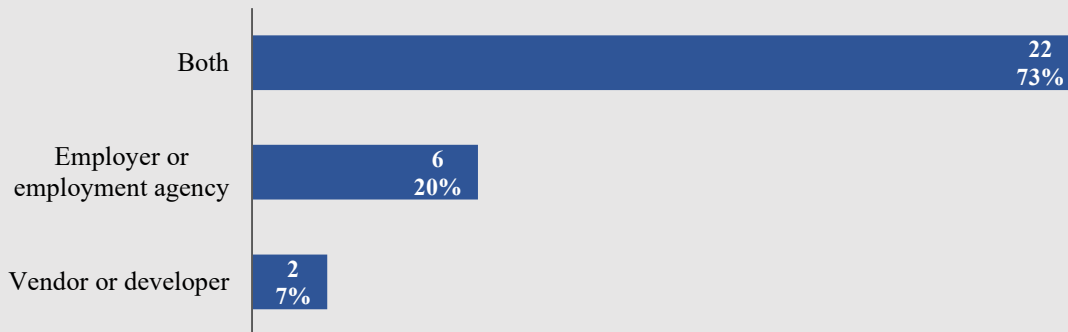
QUESTION 87: Who should conduct the audit for an AI-based application?

Response	Count	Percent
Independent third party	15	50%
Vendor or developer	6	20%
Other (please specify)	5	17%
Employer or employment agency	4	13%

QUESTION 88: Is government regulation regarding AI audits needed to protect the public interest regarding AI applications in the workplace?

Response	Count	Percentage
Yes, for all workplace AI applications	9	30%
Yes, but only for a specified subset of AI applications	7	23%
Yes, but only to the extent it has been used with "traditional" methodology (e.g., OFCCP audits)	5	17%
No, provided that there are generally accepted professional standards	9	30%
No		
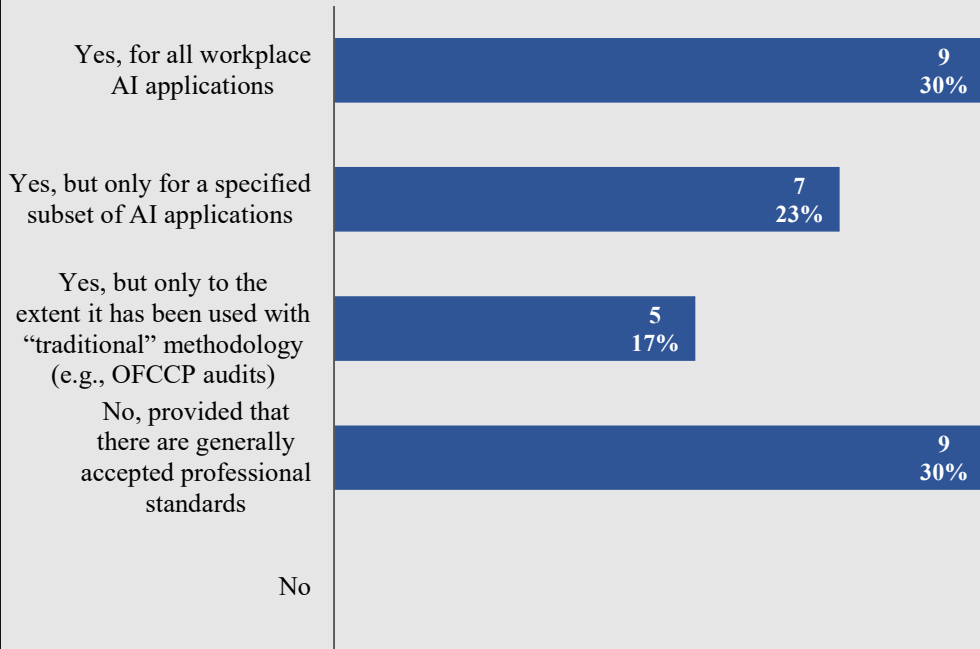
QUESTION 89: Are there additional factors that should be considered for AI audits that were not included above? If so, please list these here.

Factor	Count	Percentage
Documentation	1	14%
Routine Audits	1	14%
Monitoring	1	14%
Audits To Follow Professional Standards	1	14%
Government Regulation Should Be Uniform Across Regulatory Bodies	1	14%
Question Clarification	1	14%
Limited To Traditional Methodologies	1	14%