Responsible AI Maturity: Banking Credit Lending

Dr. John Ratzan – NY Scientific Data Summit
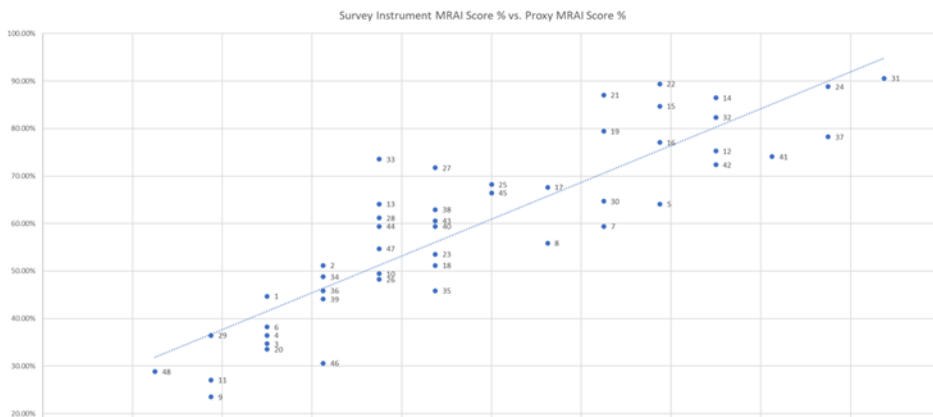
09-16-24

Dr. John Ratzan

Disciplinary Background:

- Career Consultant in the Banking Industry focusing on technology transformation.
- Doctoral degree in Strategic Mgmt. with a dissertation focus on Responsible AI.
- NA Financial Services Data & AI Lead.

RAI Principles | Instrument Development:

- Organizational Commitment
- Fairness
- Transparency (Explainability)
- Data Management
- Security

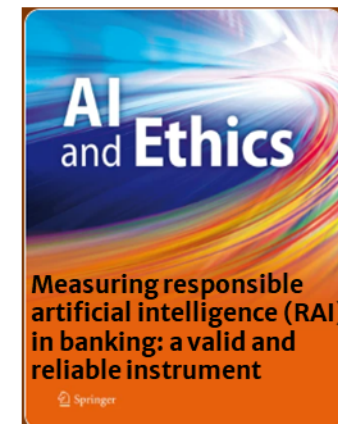## MTMM [Multi-Trait Multi-Method] Validation:

Instrument	Instrument MRAI	MTMM Type	Proxy MRAI	MTMM Type	Instrument ESG	MTMM Type
Proxy MRAI	0.882	mono-trait - multi-method				
Instrument ESG	0.553	multi-trait - mono-method	0.398	multi-trait - multi-method		
Sustainalytics ESG	0.135	multi-trait - multi-method	0.109	multi-trait - multi-method	0.532	mono-trait - multi-method

Research Findings:
Linear distribution in Bank RAI Maturity

Survey Instrument MRAI Score % vs. Proxy MRAI Score %

Research Publications:
Academic and Corporate Thought Leadership

RESEARCH REPORT: ACCENTURE
How generative AI will unlock big value in the Big Apple

AI and Ethics

Measuring responsible artificial intelligence (RAI) in banking: a valid and reliable instrument

Research Motivation

The Banking industry lacked a consistent method for determining the maturity of Responsible AI programs, which underpin confidence and trust in fair lending.
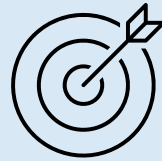
My research developed an instrument to measure the maturity of RAI programs in Banks.

MRAI Research Components

MRAI Instrument Development

I.   RAI Principles form categories

II. Research Based category attributes

III. Survey Category Attributes Validation Score

IV. CFA Statistics

V.   Cronbach Alpha Statistics

Banking Assessment

I.   Proxy MRAI Score

II.   Survey MRAI Score

III. CSR/ESG Rating

IV.   CSR/ESG Score

V.   Inter-Rater Reliability (Cohen Kappa)

VI.   MTMM Construct Validation

VII. CMM Model

MRAI Applied Future Implications

I.   Leverage MRAI Score as independent variable with other dependent variables, such as:

   a. CSR, ESG

   b. P/E

   c. TMT Diversity

   d. Brand & Reputation

II. Develop MRAI into a standard for assessing desired corporate relationships (suppliers, partners).
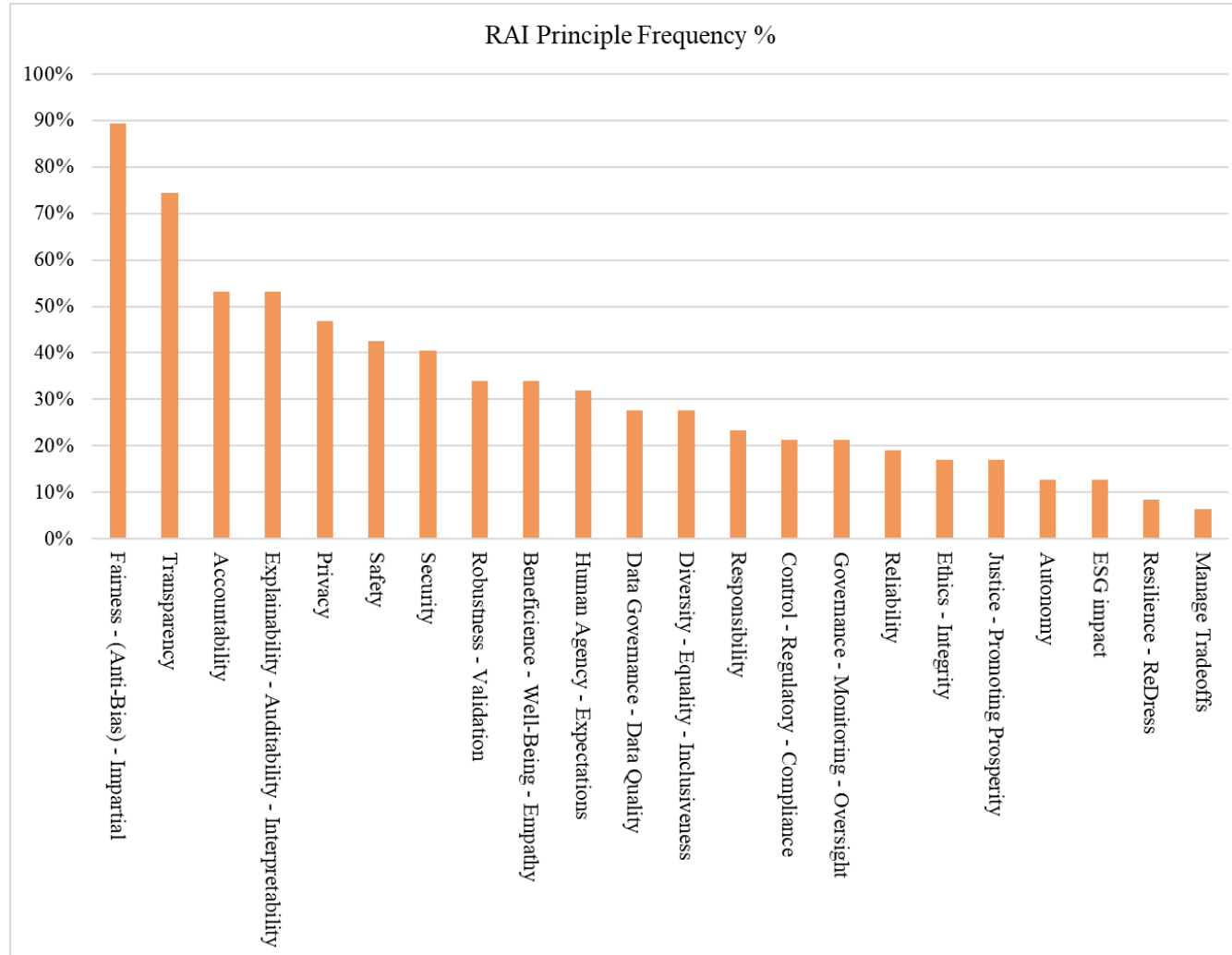
Responsible AI Definition

*This study defines **RAI** as the ability to implement AI/ML models in credit underwriting that can transparently explain the data inputs and predicted recommendation outputs of the models such that fairness, in terms of mitigation of bias and harm, is confirmed.*

Key literature for Responsible AI

- **Dignum** - Responsible Artificial Intelligence: How to Develop and Use AI in a Responsible Way.
- **Agrawal, et al.** - Prediction Machines: The Simple Economics of Artificial Intelligence
- **Jobin, et al.** - The global landscape of AI ethics guidelines.
- **Boza, et al.** - Implementing Ai Principles: Frameworks, Processes, and Tools
- **Fjeld, et al.** - Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-based Approaches to Principles for AI
- **Hagendorff, et al.** - The Ethics of AI Ethics: An Evaluation of Guidelines
- **Myers, et al.** - Developing Artificial Intelligence Sustainably: Toward a Practical Code of Conduct for Disruptive Technologies
- **Benjamins, et al.** - A choices framework for the responsible use of AI
- **Vakkuri, et al.** – Time for AI Ethics Maturity Model is Now
- **Coates, et al.** - An instrument to evaluate the maturity of bias governance capability in artificial intelligence projects

RAI Principles Commonality

In an assessment of relevant industry literature on Responsible AI, the graph below depicts the most commonly referenced RAI principles.

RAI Principle Frequency %

Research Based Attributes

#	Factor Name	#	Key Evidence	Scoring					Literature References
									Reference
1	**Organizational Commitment to RAI**		**This attribute measures the degree of organizational commitment to RAI in terms of org structure, financials, accountability.**						
	Org Structure RAI Focus	1	To what degree is there a formal org structure entity called Responsible or Ethical AI?	1	2	3	4	5	
				No evidence of org on Responsible AI	Responsible AI is mentioned within other org focus.	Other org has responsible for AI fairness	RAI mentioned within related org structure.	Formal org named Ethical or Responsible AI	de Laat, P. B. (2021). Companies Committed to Responsible AI: From Principles towards Implementation and Regulation? Philos Technol, 1-59.
	Investment in RAI	2	To what degree is there evidence of significance financial investment linked to Responsible AI.	1	2	3	4	5	
				There is no formal investment	There is discretionary budget for RAI	There is ~1-3% of IT budget allocated for RAI	There is ~5% of IT budget allocated for RAI	There is more than 5% budget allocated for RAI	Borg, J. S. (2021). Four investment areas for ethical AI: Transdisciplinary opportunities to close the publication-to-practice gap. Big Data & Society
	ROI analysis on RAI	3	To what degree is there a formal financial ROI analysis performed on Responsible AI?	1	2	3	4	5	
				There is no formal ROI measurements	There is an ROI measurement being developed	There is a limited ROI process in place.	There is significant leverage of the ROI measurements	All RAI investments go through formal ROI analysis.	Minevich, M. (2020). 4 Ways That You Can Prove ROI From AI. https://www.forbes.com/sites/markminevich/2020/03/03/4-ways-that-you-can-prove-roi-from-ai/?sh=84de94e784a7
	Training for RAI	4	To what degree are there training programs in place for all employees on Responsible AI?	1	2	3	4	5	
				There are no training programs	There is one training program	There are two training programs	There are multiple training programs	There is a RAI certificate program in place.	Cihon, P., Schuett, J., & Baum, S. D. (2021). Corporate Governance of Artificial Intelligence in the Public Interest. Information
	Culture of AI	5	To what degree is there a perception of a culture of AI within the company?	1	2	3	4	5	
				There is no perception of culture of AI	There is minimal perception of culture of AI.	There is moderate perception of culture of AI.	There is heavy perception of culture of AI.	It is perceived that the future of the company is dependent on AI.	Murphy, J. W., & Largacha-Martínez, C. (2021). Is it possible to create a responsible AI technology to be used and understood within workplaces and unblocked CEOs' mindsets? Ai & Society

Data Collection

- The aim was to collect data from 50 of the top US Banks.

- This study was able to collect data from 48 of the 50 Banks, with 2 Banks' data unavailable for interviews.

- Each of the interviews had three components.

  - Instrument validation survey data (Face & Content Validity, CFA, Ca)
  - Instrument Bank assessment data (Cronbach's alpha, CFA)
  - ESG questions data (Cronbach's alpha, CFA, MTMM)

- This data represents not just a sample, but a population of data, and provides a more robust representation of the industry adding robustness, validity, and reliability to the study.

Validity

- Face Validity

  - The study conducted a pre-validation instrument interview with key Bank industry practitioners to ensure the elements of the instrument are relevant to the field.

- Content Validity

  - The study utilized the pre-validation instrument interviews to ensure that the scientific context of the instrument was accurate and will measure the correct elements.

Reliability

- Internal Consistency Reliability

  - The study utilized both Cronbach Alpha technique as well as CFA (Confirmatory Factor Analysis) to demonstrate the internal consistency. The study goal was to have Cronbach Alpha > .7.

- Inter-Rater Reliability

  - The study also utilized a peer review process for inter-rater reliability where the other rater will review the coding of the proxy scores and validate their own interpretation to be compared and reviewed with the original coding.
  - The study measured Cohen's Kappa Coefficient, which measures the degree of agreement between the 2 raters to ensure that reliability is in an acceptable range >.6 and within $p<.01$.

Survey Instrument

#	Factor Name	Key Evidence	Scoring				
			Little	Below	Average	Above	Advanced
1	**Organizational Commitment to RAI**	**This attribute measures the degree of organizational commitment to RAI in terms of org structure, financials, accountability.**					
a	Org Structure RAI Focus	To what degree is there a formal org structure entity called Responsible or Ethical AI?	1	2	3	4	5
			Little	Below	Average	Above	Advanced
b	Investment in RAI	To what degree is there evidence of significance financial investment linked to Responsible AI.	1	2	3	4	5
c	Training for RAI	To what degree are there training programs in place for all employees on Responsible AI?	1	2	3	4	5
d	Culture of AI	To what degree is there a perception of a culture of AI within the company?	1	2	3	4	5
e	C-Suite Involvement	To what degree is the CEO or Board updated on the company's RAI program?	1	2	3	4	5
2	**Transparency & Explainability**	**This attribute measures the degree of transparency in the AI in terms of the algorithms, and models that comprise the AI & ML.**					
a	Explanability Governance	To what degree are there formal policies or processes in place to govern explanability?	1	2	3	4	5
b	Regulatory Sandbox (Visibility)	To what degree are there capabilities in place to provide visibility to regulators on explainability?	1	2	3	4	5
c	Model Audit Controls	To what degree are there capabilities in place to audit models?	1	2	3	4	5
d	Model Drift Prevention Monitoring	To what degree are there capabilities or processes in place to test & mitigate model drift?	1	2	3	4	5
e	Use of Model Card Reporting	To what degree are model cards leveraged to provide a decriptive model explainability?	1	2	3	4	5
f	Advanced ML Explain (LIME, SHAP)	To what degree is there use of advanced black box technology such as SHAP or LIME?	1	2	3	4	5

Table title: Maturity of Responsible AI in Banking Survey Instrument

* Instrument not fully depicted.

Instrument Validation Statistics

Statistics – Cronbach's alpha Analysis

The data below depicts a .889 for the instrument components, demonstrating a high degree of validity for the instrument components.

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.889	.900	9

Item Statistics

	Mean	Std. Deviation	N
Cat	4.50	.652	48
OC	4.58	.577	48
Explain	4.56	.580	48
Fair	4.56	.580	48
Data	4.65	.526	48
Security	4.65	.565	48
Weight	4.63	.606	48
Likert	4.65	.565	48
Relevance	4.38	.703	48

Inter-Item Correlation Matrix

	Cat	OC	Explain	Fair	Data	Security	Weight	Likert	Relevance
Cat	1.000	.226	.422	.366	.155	.202	.108	.087	.232
OC	.226	1.000	.588	.778	.766	.843	.639	.713	.341
Explain	.422	.588	1.000	.747	.807	.621	.613	.491	.098
Fair	.366	.778	.747	1.000	.737	.816	.734	.686	.150
Data	.155	.766	.807	.737	1.000	.787	.777	.644	.194
Security	.202	.843	.621	.816	.787	1.000	.661	.733	.234
Weight	.108	.639	.613	.734	.777	.661	1.000	.599	.187
Likert	.087	.713	.491	.686	.644	.733	.599	1.000	.181
Relevance	.232	.341	.098	.150	.194	.234	.187	.181	1.000

14

Statistics – ESG Factor Analysis

Below is the result of the CFA (Confirmatory Factor Analysis) and Cronbach's alpha on the ESG supplemental questions. The Cronbach's alpha was .877, and all of the factors in the CFA analysis were above .5.

Cronbach's alpha	CFA (Confirmatory Factor Analysis)

Cronbach's alpha

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.877	.881	6

Item Statistics

	Mean	Std. Deviation	N
Exec	3.10	.951	48
Culture	3.29	.922	48
Training	3.42	.794	48
Env	3.31	1.223	48
Social	3.02	.978	48
Gov	3.10	.751	48

Inter-Item Correlation Matrix

	Exec	Culture	Training	Env	Social	Gov
Exec	1.000	.596	.448	.667	.478	.462
Culture	.596	1.000	.528	.597	.630	.509
Training	.448	.528	1.000	.586	.454	.461
Env	.667	.597	.586	1.000	.635	.497
Social	.478	.630	.454	.635	1.000	.721
Gov	.462	.509	.461	.497	.721	1.000

CFA (Confirmatory Factor Analysis)

Communalities

	Initial	Extraction
Exec	1.000	.589
Culture	1.000	.665
Training	1.000	.522
Env	1.000	.710
Social	1.000	.690
Gov	1.000	.588

Extraction Method: Principal Component Analysis.

Statistics – Proxy MRAI Data

This study reviewed public data to collect and score the Proxy MRAI data.

Company # / Criteria	In House RAI Research Dept / Publications? (.5 for AI)	RAI Articles in the Press? (.5 for AI)	RAI Published Principles? (.5 for AI)	RAI mentioned in 10K / Shareholder letter? (.5 for AI).	RAI link on Website? (.5 for AI)	Research Partnerships with Universities ?	RAI COE ? (.5 for AI)	Careers in RAI Related ? (.5 for AI)	MRAI Proxy score
Criteria	1) Research Department stated on website. 2) Reference to research or invention on other website. 3) Evidence of research publications	1) Evidence of articles or publications that refer to RAI.	1) Does company have RAI principles published on website?	1) Mention of Responsible AI in Annual report 2) Mention of Ethical AI in Annual report.	1) Prominent link on website referring to Responsible AI.	1) Evidence of partnerships with research universities on technology innovation or AI. (does not have to be Responsible AI research).	1) Evidence of a center of excellence or working group that focuses on responsible, ethical, explainable, transparent, fair AI.	1) Evidence of a job description for Responsible AI. (Can include data science, machine learning, model validation and compliance and responsible business corporate standard)	
	1	1	0	1	1	1	1	1	7.0
	0.5	1	0	1	0	1	1	1	5.5
	1	1	0.5	0.5	1	0.5	1	0.5	6.0
	1	1	1	0	0	1	1	1	6.0
	1	0.5	0	1	1	0.5	1	0.5	5.5
	0.5	0.5	0	0	0	1	0.5	0.5	3.0
	0.5	0.5	0	0.5	0	1	0.5	0.5	3.5
	0.5	1	1	0.5	1	0.5	0.5	0.5	5.5
	0.5	1	0	0	0	0.5	0.5	0.5	3.0
	1	1	1	0.5	1	1	1	1	7.5
	0	0.5	0	0	0	1	0.5	1	3.0
	0.5	1	1	1	1	1	0.5	0.5	6.5
	1	1	0	1	0.5	1	1	0.5	6.0
	1	0.5	0	1	0	0.5	0.5	0.5	4.0
	0.5	1	0	0.5	0	0.5	0.5	1	4.0

* Data not fully depicted.

Statistics – Inter-rater Reliability

This study performed an Inter-rater reliability internal consistency test on the Proxy MRAI data to demonstrate reliability with a peer review. The agreement was very high at 97.7% due to the alignment between the raters resulting in a Cohen Kappa of 96.5% when considering the chance correlation effect.

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Researcher Score * Assistant Score	384	100.0%	0	0.0%	384	100.0%

Researcher Score * Assistant Score Crosstabulation

		Assistant Score						Total	
		.0		.5		1.0			
		N	%	N	%	N	%	N	%
Researcher Score	.0	132	93.6%	0	0.0%	0	0.0%	132	34.4%
	.5	4	2.8%	131	100.0%	0	0.0%	135	35.2%
	1.0	5	3.5%	0	0.0%	112	100.0%	117	30.5%
Total		141	100.0%	131	100.0%	112	100.0%	384	100.0%

Symmetric Measures

		Value	Asymptotic Standard Error[a]	Approximate T[b]	Approximate Significance
Measure of Agreement	Kappa	.965	.012	26.715	<.001
N of Valid Cases		384			

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

Validity

- Construct Validity
  - The study utilized the MTMM methodology to elucidate construct validity.
    - The methods were the MRAI instrument interview and primary Proxy MRAI data review.
    - The main traits were the MRAI score and the CSR-ESG rating [Sustainalytics]
  - The study tested for convergent and discriminant validity through the MTMM methodology.

Multi-Trait Multi-Method Correlation

The correlation between the Instrument MRAI and the Proxy MRAI was high at .882, which provides evidence for convergent validity for the Mono-Trait MRAI.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.882[a]	.778	.773	9.99492%

a. Predictors: (Constant), V4

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	16120.006	1	16120.006	161.364	<.001[b]
	Residual	4595.326	46	99.898		
	Total	20715.332	47			

a. Dependent Variable: V3

b. Predictors: (Constant), V4

Coefficients[a]

| Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	-11.640	4.915		-2.368	.022	-21.534	-1.747
	V4	1.004	.079	.882	12.703	<.001	.845	1.164

Multi-Trait Multi-Method Matrix

Below is the result of the correlation regressions from the MTMM analysis, which provides construct validity evidence for the instrument demonstrating both convergent as well as discriminant validity.

Instrument	Instrument MRAI	MTMM Type	Proxy MRAI	MTMM Type	Instrument ESG	MTMM Type
Proxy MRAI	0.882	mono-trait - multi-method				
Instrument ESG	0.553	multi-trait - mono-method	0.398	multi-trait - multi-method		
Sustainalytics ESG	0.135	multi-trait - multi-method	0.109	multi-trait - multi-method	0.532	mono-trait - multi-method

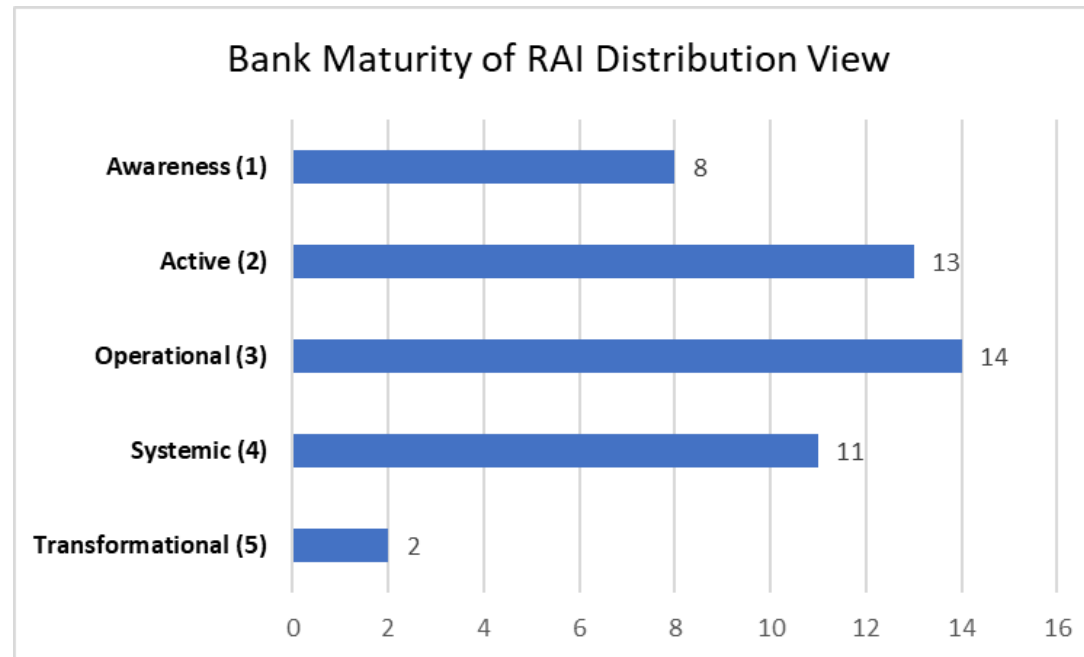# MTMM Correlation Graph

This chart depicts the correlation between the MRAI (Instrument and Proxy) trait methods (.882).

Survey Instrument MRAI Score % vs. Proxy MRAI Score %
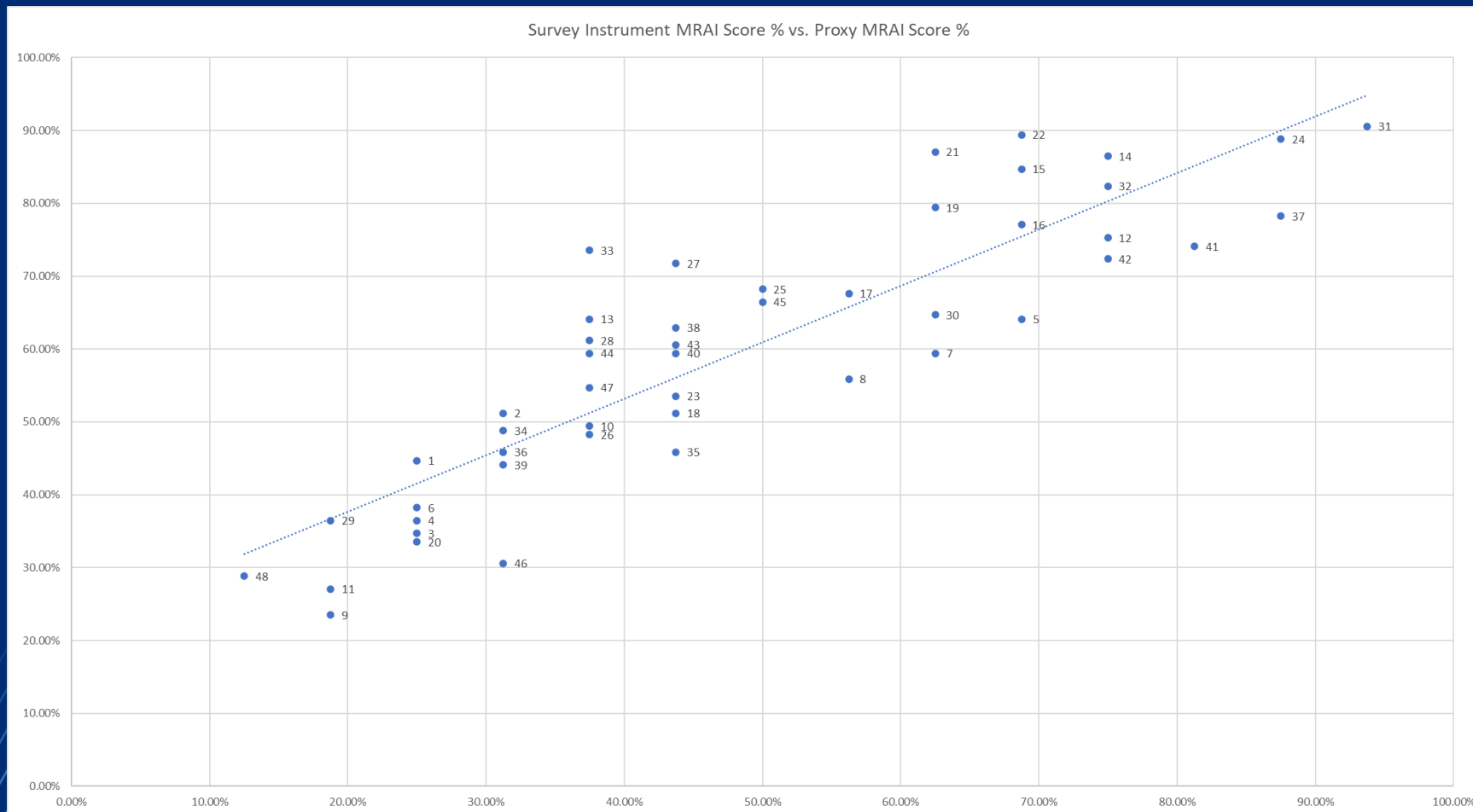
Capability Maturity Model View

The CMM view data is the average MRAI trait and is based on the Gartner AI model and demonstrates that maturity of the Banking industry is 'Operational', indicating room for improvement. This CMM analysis is consistent with the standard distribution statistics on the next slide.

Bank Maturity of RAI Distribution View

Level	Value
Awareness (1)	8
Active (2)	13
Operational (3)	14
Systemic (4)	11
Transformational (5)	2

Standard Distribution Statistics

The graph below depicts standard distribution statistics and is consistent with the findings of the CMM analysis with a mean of 53.74%.

Statistics

MRAI_Average

N	Valid	48
	Missing	0
Mean		53.7354%
Median		51.8750%
Mode		20.66%[a]
Std. Deviation		19.12882%
Variance		365.912
Skewness		.124
Std. Error of Skewness		.343
Range		71.51%

Histogram — MRAI_Average. Mean = 53.74%, Std. Dev. = 19.129%, N = 48
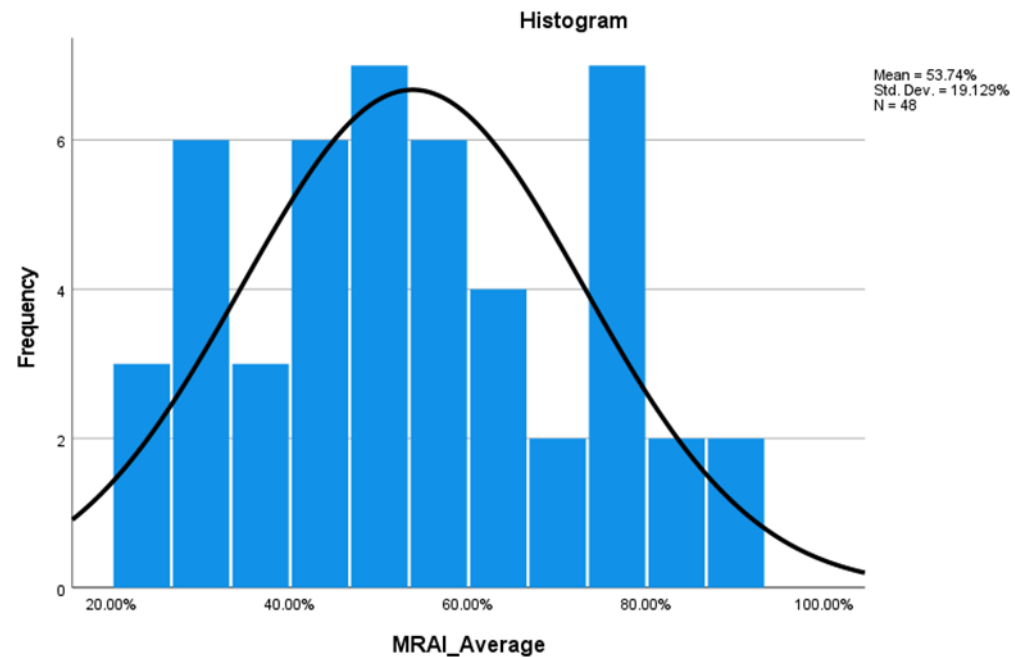
Contribution

I. **Researched & Validated MRAI Instrument**

    a. Could be utilized for marketing & advertising

    b. Should be utilized for regulatory exam preparations.

II. **Secondary MRAI Instrument that could be scaled.**

    a. An AI tool to perform broader textual analysis could be deployed.

III. **Bank RAI Maturity assessment based on Gartner CMM.**

    a. Banks can perform a self-assessment with the instrument and determine where they compare to the industry benchmark.

Limitations

1. The data collected from the MRAI instrument survey is a Bank self-assessment from one approved individual and may not represent a fully accurate assessment the true maturity of the RAI capability.

2. The data collected in the proxy score is a peer-reviewed (inter-rater reliable) interpretation of publicly available data and may not represent a fully accurate assessment of the true maturity of the RAI capability.

Areas for further research

1. Leverage the MRAI score as an independent variable for correlations with various important Bank dependent variables, such as:
   a) CFP (Corporate Financial Performance – P/E)
   b) CSR-ESG Index Score
   c) Brand or Reputation Index Score
   d) TMT Diversity
   e) Technology Budget
   f) CEO Social style

2. Another possibility for research is studying the impact of the MRAI score becoming a key lever for responsible business assessment of suppliers, partners, companies to work for, etc.

Thank You