MMA

RESPONSIBLE AI

Marketing AI Risk Evaluation

An MMA Framework

ALC
AI LEADERSHIP COALITION

Marketing AI Risk Evaluation Framework

Executive summary

The Marketing AI Risk Evaluation Framework guides businesses in assessing and managing risks associated with AI in marketing. It addresses key risk categories, including privacy, accuracy, ethical, and compliance risks, providing marketers with tools to navigate the complex landscape of AI implementation.

The framework employs MMA's LIME risk rating criteria, quantifying risks and mitigation effort based on three factors:

- Likelihood: Probability of risk occurring (0-1 scale)
- Impact: Potential business and reputational effect (1-5 scale)
- Mitigation Effort: Estimated person-months required to address the risk

The Risk Score is calculated using the following formula:

$$\text{Risk Score} = \text{Likelihood} \times \text{Impact}$$

This formula helps marketers prioritize risks effectively, focusing resources on the most critical issues.

The framework outlines three primary risk mitigation strategies: avoidance, transfer, and mitigation. It also provides guidance on developing incident response protocols, including severity classification, notification procedures, and remediation plans.

To effectively use this framework, marketing leaders should consider:

1. Regularly assessing AI risks across all marketing functions
2. Implementing robust data governance and privacy measures
3. Ensuring ethical AI practices and transparency
4. Staying compliant with evolving regulations
5. Investing in employee training and AI literacy

This framework serves as a foundation for responsible AI adoption in marketing, enabling businesses to harness AI's potential while maintaining ethical standards and consumer trust. Regular updates will ensure its relevance as technologies and regulations evolve.

Introduction

Integrating AI into marketing strategies offers unprecedented efficiencies but also introduces ethical and operational challenges. This Marketing AI Risk Evaluation Framework has been crafted as a comprehensive guide to assist businesses in understanding, assessing, and managing the risks associated with AI in marketing. It addresses concerns ranging from data security to ethical implications of bias and accessibility, ensuring that marketers can responsibly leverage AI tools amidst the fast pace of technological advancements.

How to Use This Framework

This framework is a versatile tool applicable to the entire marketing AI scope, including vendor evaluations, product development, marketing strategy planning, and consumer education. It aims to demystify the complexities of AI in marketing, providing clear, actionable guidelines for businesses at various stages of AI integration. By following this framework, marketing teams can ensure their marketing strategies are effective and ethically sound.

Part I: Risk Assessment Guide

Identifying Risks Across Marketing Functions

Data Collection

Ethical and transparent data collection practices are foundational to building trust and ensuring privacy. When collecting data for AI models, it is crucial to source data responsibly and obtain clear consent and understanding from all involved parties. Businesses must be particularly vigilant when operating within regulated industries such as healthcare or finance in the US, ensuring that personally identifiable information (PII) is handled appropriately.

When collecting data from AI models, it is important to adhere to the principle of data minimization, only requesting what is necessary for the intended purpose. Proper storage and security compliance measures must be in place to protect the collected data from unauthorized access or breaches.

Data Usage

Data usage considerations can be triggered when leveraging data within the language model (LLM) on which the AI platform is built. LLMs can be a black box regarding legal compliance and the potential usage of biased data, making it essential to approach their use cautiously.

Transparency in how the collected data is utilized is crucial, considering regulations such as GDPR, CCPA, and industry-specific privacy laws in healthcare and finance. It is important to note that after data is assimilated into the model, new attributes could be appended to cohorts and individuals, further complicating data usage and requiring additional scrutiny.

Regular privacy impact assessments should be conducted to identify potential risks and vulnerabilities in data handling processes. These assessments should cover both the data collection and usage aspects, ensuring a comprehensive approach to data privacy and security.

Questions to Consider:

- How would our customers react if they knew how we're using their data?
- Have we obtained explicit and informed consent to use their data for this purpose?
- Does our data collection strategy ensure transparency and fairness to all parties?
- How do we communicate our data use policies to our stakeholders?
- Are we fully compliant with sector-specific regulations and broader regulations like the GDPR? How regularly do we review our compliance status in light of evolving legal standards?
- How do we ensure that the data collected is the minimum necessary for our purposes, respecting the principle of data minimization?
- What measures are in place to immediately address any detected vulnerabilities or breaches in our data handling processes?
- What existing legal agreements or terms and conditions associated with our products or services would make using data for AI purposes non-compliant?
- How can we ensure that our future legal agreements and contracts accommodate the use of AI as recommended by our legal team and responsible AI guidelines?

Mitigation Strategies:

- Implement strong access controls, encryption, and data anonymization to protect sensitive data
- Utilize secure data storage solutions to safeguard against breaches and unauthorized access
- Obtain explicit user consent for data collection and usage
- Comply with relevant regulations, such as GDPR, CCPA, and industry-specific privacy laws
- Regularly assess and improve data handling processes to ensure robustness against emerging threats
- Train employees on data privacy best practices and their responsibilities in protecting user data
- Conduct regular privacy impact assessments to identify potential risks and vulnerabilities

Customer Segmentation

Effective customer segmentation is crucial for personalized marketing but must be approached with caution to prevent bias and discrimination. It's essential to critically examine segmentation strategies for potential biases and ensure ethical data handling, especially sensitive information.

MMA's **Moveable Middles Growth Framework** offers a strategic approach to segmentation that focuses on identifying and targeting consumers most likely to respond to marketing efforts. This framework helps marketers efficiently allocate resources by concentrating on the “movable middle” – consumers with a moderate probability of purchasing the brand. While implementing this approach, it's vital to maintain ethical standards and avoid targeting customers on sensitive personal factors.

A code of ethics should be established for reviewing segmentation data and application. A designated role (e.g., Chief Data Officer) should oversee segmentation ethics and development. Implement robust data governance frameworks to provide clear guidelines and accountability mechanisms, ensuring strict regulation of sensitive information use.

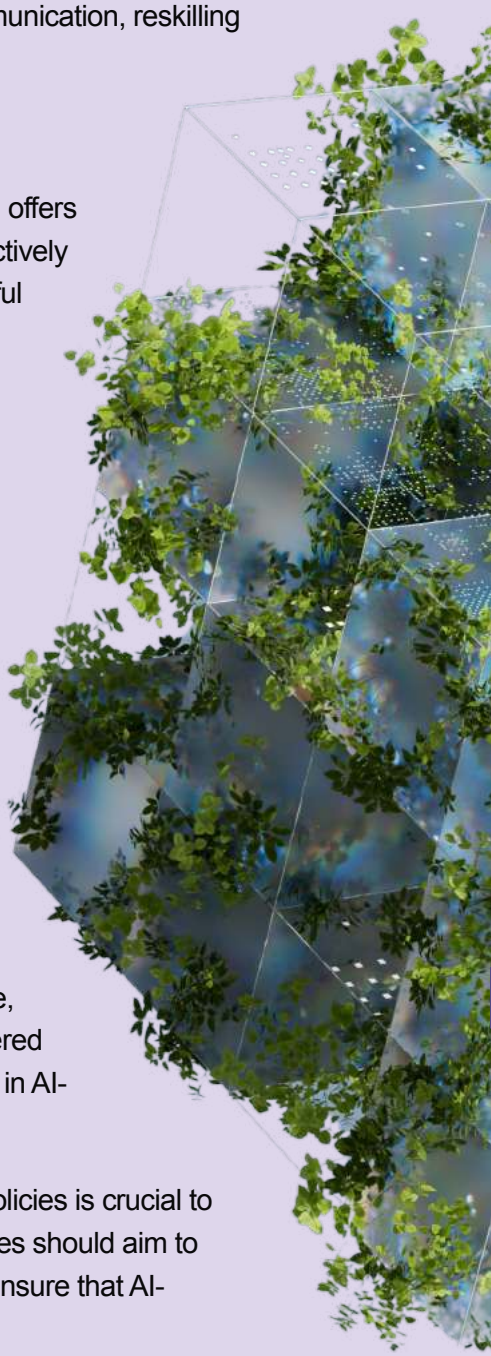
Questions to Consider:

- Could our segmentation methods unintentionally reinforce societal biases or result in discriminatory practices? What measures can we take to mitigate this risk?
- How do we ensure our segmentation strategies are developed and implemented in a way that respects customer privacy and adheres strictly to data protection regulations?
- What mechanisms do we have in place to prevent the unethical use of sensitive information, such as images, videos, and content revealing personal characteristics or preferences, in our segmentation efforts? How do we define and protect this sensitive data that extends beyond PII within our organization?
- How does our data governance framework support ethical segmentation practices? Are there clear guidelines and accountability mechanisms for those involved in segmentation tasks?
- How frequently do we review and update our data governance and segmentation strategies to reflect changes in regulations, market conditions, and technological advancements?
- How can we proactively address the potential adverse impact of AI-driven segmentation on employees, such as changes in productivity expectations or bottom-line efficiency pressures, and develop strategies to mitigate these risks?

Mitigation Strategies:

- Establish a code of ethics for reviewing data used in segmentation and guidelines for its application
- Assign responsibility for segmentation ethics and development to a dedicated role, such as a Chief Data, Customer, or Technology Officer
- Implement robust data governance frameworks to oversee the ethical use of customer data in segmentation

- Align all data suppliers with the established segmentation standards and best practices
- Regularly review and update data governance and segmentation strategies to reflect changes in regulations, market conditions, and technological advancements
- Develop comprehensive employee support programs to mitigate the adverse impact of AI-driven segmentation on the workforce, including transparent communication, reskilling opportunities, and mental health resources



Content Generation

AI-powered content generation, including LLMs and text-to-image models, offers significant creative and efficiency benefits. However, it is essential to proactively address ethical concerns, particularly the risks of disinformation and harmful content. Ensuring AI-generated content adheres to ethical standards and truthfulness is vital for maintaining audience trust and information integrity. When using text-to-image models, it is especially important to prioritize brand safety and consistency to ensure that generated visuals align with the brand's marketing strategies, values, and messaging to prevent misrepresentation or reputational damage.

A person should be responsible for restating the baseline aesthetics, voice, and parameters of how the brand is represented in the face of AI and the millions of new possibilities. This dynamic brand/content body should specify at which stages in the brand/content process different AI platforms are permissible. For example, DALL-E might only be used for the exploratory phase, while ChatGPT could be admissible for tertiary research.

Content moderation policies tailored to the unique challenges posed by text-to-image models should be developed and implemented. For example, inappropriate or misleading generated visuals should be identified and filtered out. Trust and safety protocols should also be put in place to mitigate risks in AI-generated content.

Establishing and enforcing ethical AI guidelines and content moderation policies is crucial to guiding the development and deployment of generative AI. These guidelines should aim to prevent the generation of harmful or offensive content by bad actors and ensure that AI-generated content aligns with the brand's values and messaging.

Questions to Consider:

- How do we establish and enforce ethical AI guidelines and content moderation policies to guide the development and deployment of generative AI, ensuring harmful or offensive content is prevented from generation by bad actors?

- What strategies do we employ to balance AI's creative capabilities in content generation with the necessity of oversight mechanisms, such as trust and safety protocols, to mitigate the risk of disseminating misleading or harmful information?
- How do we integrate comprehensive training for our team and implement red-teaming practices to identify vulnerabilities, ensuring our members can effectively recognize and address risks in AI-generated content?
- What measures can we take to ensure that text-to-image models generate visuals that are consistent with our brand's marketing strategies, values, and messaging, thereby maintaining brand safety and preventing misrepresentation?
- How can we develop and implement content moderation policies specifically tailored to the unique challenges posed by text-to-image models, such as identifying and filtering out inappropriate or misleading generated visuals?

Mitigation Strategies:

- Establish and enforce ethical AI guidelines and content moderation policies to prevent the generation of harmful or offensive content
- Implement Trust & Safety protocols to mitigate risks in AI-generated content
- Develop content moderation policies tailored to the unique challenges posed by text-to-image models
- Ensure generated visuals align with the brand's marketing strategies, values, and messaging to maintain brand safety and prevent misrepresentation
- Provide comprehensive training for team members to recognize and address risks in AI-generated content effectively
- Implement red-teaming practices to identify vulnerabilities in AI-generated content

Risk Categories and Definitions

Privacy Risks

Privacy risks are associated with how data is collected, stored, and shared. Mitigating these risks necessitates strict adherence to privacy regulations, such as the GDPR, and securing explicit consent from customers prior to data processing. This category highlights the importance of respecting customer privacy and adhering to legal standards. It is recommended that organizations consult their privacy team or contract Certified Information Privacy Professional to review their AI governance practices.

To enhance data security, organizations should implement specific practices such as encryption, anonymization, and secure data storage solutions. Encryption helps protect sensitive data by converting it into a coded format that can only be accessed with a decryption key. Anonymization involves removing personally identifiable information (PII) from datasets, making it difficult to trace data back to specific individuals. Secure data storage solutions, such as cloud-based platforms with robust security features, can help safeguard data from unauthorized access and breaches.

Establishing frameworks for regularly assessing and improving data handling processes is crucial to ensure that privacy risks are effectively mitigated. This includes developing clear protocols for handling data breaches, both within the company and at the LLM level. Organizations should have a well-defined process for detecting, investigating, and responding to data breaches, as well as for notifying affected individuals and relevant authorities in compliance with applicable regulations.

Key Considerations:

- How do we employ encryption and anonymization to enhance the privacy and security of the data we manage, particularly when using text-based LLMs that may have access to personally identifiable information (PII) in verbatims or notes, even if PII is not directly collected?
- Are our data storage and sharing protocols designed to uphold the highest standards of security and transparency, in line with privacy laws and best practices? This includes implementing strong access controls, regularly auditing data access, obtaining explicit user consent, complying with relevant regulations, and training employees on data privacy.
- What is the process for handling data breaches within the company or at the LLM level?
- How regularly do we assess and refine our data handling practices to ensure they remain robust against emerging threats and aligned with our commitment to protecting customer privacy?
- Have we engaged with our privacy team or enlisted the expertise of a Certified Information Privacy Professional (CIPP) to conduct a thorough review of our AI governance practices?

Mitigation Strategies:

- Implement strong access controls, encryption, and data anonymization to enhance data privacy and security
- Utilize secure data storage solutions to protect against breaches and unauthorized access
- Obtain explicit user consent for data collection and usage, in compliance with relevant regulations
- Establish a clear process for handling data breaches within the company and at the LLM level
- Regularly assess and improve data handling practices to ensure robustness against emerging threats
- Engage with privacy teams or Certified Information Privacy Professionals to review AI governance practices
- Track and monitor unforeseen outputs from AI systems and establish processes for communication and course correction, such as prompt learning and LLM reconnection
- Implement a new type of opt-in agreement for users, acknowledging the unpredictable nature of AI outputs and that guarantees cannot be made regarding the nature of the outputs

Accuracy Risks

Accuracy risks in AI marketing encompass the potential for AI systems to generate inaccurate outputs or false information, sometimes referred to as “AI hallucinations.” These risks can lead to misinformed decisions and the dissemination of misleading or harmful information, which can undermine the credibility of marketing efforts and harm the company's reputation. Accurate data and outputs are fundamental to the success of AI-driven marketing strategies. The reliability of AI predictions and decisions is only as good as the data it processes and the algorithms it uses. Misinformation or errors stemming from AI systems can lead to poor decision-making, customer dissatisfaction, and potential legal liabilities.

Key Considerations:

- Should we source AI tools from vendors or build our own? Consider the impact on accuracy, compliance, security, and feasibility.
- How can we ensure the integrity and accuracy of the data used by our AI systems throughout their lifecycle?
- What measures do we have in place to detect and correct AI-generated inaccuracies before they impact our marketing strategies?
- How do we balance the need for timely decisions with the risk of AI hallucinations in dynamic and real-time marketing environments?

Mitigation Strategies:

- Establish a robust data governance framework to ensure high-quality data
- Continuously evaluate and improve model performance through iterative testing and refinement
- Conduct rigorous testing and validation of AI models to benchmark accuracy and reliability
- Implement feedback loops from multiple sources to refine AI decisions and outputs
- Monitor and address biases in training and evaluation data to prevent skewed results
- Provide explanations and transparency in AI predictions to enable verification and trust
- Implement strict validation mechanisms to detect and prevent AI hallucinations
- Encourage interdisciplinary collaboration between data scientists, domain experts, and ethicists to review and update language models
- Track unforeseen outputs and implement processes for communication and course correction
- Establish opt-in agreements acknowledging the unpredictable nature of AI outputs

Ethical Risks

Ethical risks in AI encompass issues such as bias and discrimination, which can adversely affect diverse societal groups. These unintended consequences can lead to unfair treatment, limited access to opportunities, and perpetuation of stereotypes, among other severe impacts. Mitigating these risks mandates a steadfast commitment to embedding fairness, transparency, and accountability within AI systems.

To effectively address ethical concerns, organizations should adopt ethical AI frameworks that guide the development and deployment of AI technologies. These frameworks should provide clear guidelines and best practices for ensuring that AI systems are designed and implemented in a manner that prioritizes fairness, transparency, and accountability. By aligning their AI initiatives with these frameworks, organizations can foster a culture of responsible AI development and minimize the risk of unintended consequences.

Conducting comprehensive impact assessments is another crucial step in understanding the societal implications of AI. These assessments should evaluate the potential effects of AI systems on various stakeholder groups, particularly marginalized communities, and identify any risks of bias, discrimination, or unfair treatment. By proactively assessing the societal impact of AI, organizations can take corrective action where necessary and ensure that their AI systems are functioning equitably across all user demographics.

Establishing ethical oversight committees is also essential for ensuring that AI systems are developed and deployed in an ethical manner. These committees should be composed of diverse stakeholders, including AI experts, ethicists, and representatives from affected communities, to provide multiple perspectives and ensure that ethical considerations are prioritized throughout the AI lifecycle.

In addition to these measures, organizations must also consider the potential adverse impact of AI on employees, such as job displacement or changes in productivity expectations. It is crucial to develop comprehensive employee support programs that mitigate these risks, including reskilling and upskilling initiatives that enable employees to adapt to the changing nature of work in the age of AI.

Key Considerations:

- How can we implement mitigation strategies when biases in training data cannot be entirely eliminated, ensuring our AI models function equitably across all user demographics?
- Are we routinely auditing our AI systems to evaluate their societal impacts, particularly on marginalized groups, and taking corrective action where necessary?
- In what ways are we fostering transparency and accountability in our AI operations, ensuring stakeholders are informed and ethical standards are upheld throughout the AI lifecycle?
- How can we proactively address the potential adverse impact of AI on employees, such as job displacement or changes in productivity expectations, and develop strategies to mitigate these risks?

Mitigation Strategies:

- Use diverse and representative training data to minimize bias
- Actively seek out and incorporate data from underrepresented groups

- Regularly evaluate and mitigate bias in AI models through fairness assessments
- Involve diverse stakeholders in model development to ensure multiple perspectives are considered
- Provide transparency in the decision-making process to build trust and accountability
- Regularly monitor and update models to address emerging biases or ethical concerns
- Develop comprehensive employee support programs to mitigate the adverse impact of AI on the workforce, including reskilling and upskilling initiatives
- Adopt ethical AI frameworks that guide the development and deployment of AI technologies, prioritizing fairness, transparency, and accountability
- Conduct comprehensive impact assessments to understand the societal implications of AI and take corrective action where necessary
- Establish ethical oversight committees composed of diverse stakeholders to ensure that ethical considerations are prioritized throughout the AI lifecycle

Compliance Risks

Compliance risks pertain to the legal and regulatory frameworks governing AI usage in marketing. This includes adhering to industry standards, legal requirements, and governance structures that regulate data use, consumer protection, and ethical AI best practices.

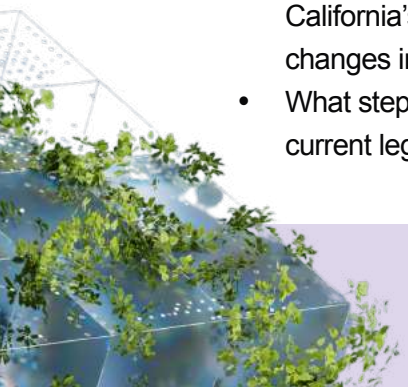
Compliance risks encompass the complex legal and regulatory landscape surrounding the use of AI in marketing. It is crucial for organizations to navigate these frameworks diligently, ensuring adherence to industry standards, legal requirements, and governance structures dedicated to data usage, consumer protection, and ethical AI best practices.

To effectively manage compliance risks, organizations should prioritize continuous legal education and regulatory monitoring. This involves staying up-to-date with the latest developments in AI and marketing law by subscribing to legal updates, attending relevant webinars and conferences, and engaging with legal experts specializing in these areas. By maintaining a comprehensive understanding of the legal landscape, organizations can proactively adjust their strategies and operations to ensure ongoing compliance.

Implementing regular compliance reviews and audits is another essential aspect of mitigating compliance risks. These reviews should assess the organization's AI practices against current legal and regulatory standards, identifying any gaps or areas for improvement. By conducting these audits on a frequent basis, organizations can promptly address any compliance issues and implement necessary adjustments to align with evolving legal requirements.

Key Considerations:

- How do we maintain awareness of regulatory developments, such as the EU AI Act and California's automated decision-making regulations, and ensure our team integrates these changes into our operational and strategic planning?
- What steps are we taking to ensure our AI-driven marketing strategies fully comply with current legal and regulatory standards?



- How frequently do we conduct internal reviews and audits of our AI practices to identify compliance gaps and implement necessary adjustments?

Mitigation Strategies:

- Stay updated with applicable laws, regulations, and industry standards through continuous learning and engagement with legal experts
- Establish compliance monitoring and reporting mechanisms to track adherence to legal requirements
- Conduct regular compliance audits to identify gaps and implement necessary adjustments
- Provide training on compliance requirements to employees to ensure organization-wide understanding and commitment
- Collaborate with legal and regulatory experts to ensure AI practices align with the latest guidelines and best practices
- Implement processes for handling and reporting non-compliance incidents promptly and transparently
- Obtain proper licenses and permissions for the use of copyrighted content in AI training data
- Implement content filtering and copyright detection mechanisms to prevent infringement and protect intellectual property rights
- Subscribe to legal updates and attend relevant webinars and conferences to stay informed about developments in AI and marketing law
- Engage with legal experts specializing in AI and marketing law to gain insights into best practices for compliance and risk mitigation
- Conduct regular internal reviews and audits of AI practices to identify compliance gaps and implement necessary adjustments proactively
- Foster a culture of compliance by providing comprehensive training to employees on legal and regulatory requirements and encouraging open communication about compliance concerns

Risk Measurement Criteria

Quantitative Metrics

Quantitative metrics offer objective, numerical data to help marketing teams assess and manage the risks associated with AI in marketing. These metrics can serve as benchmarks for evaluating the effectiveness of risk mitigation strategies over time. Examples include:

- **Data Breach Statistics**: Track the number of data breaches, affected records, and the financial impact of each incident. This can help quantify the organization's vulnerability to cybersecurity threats.

- **Bias Detection Rates**: Measure the accuracy and fairness of AI algorithms across different demographics by analyzing decision-making patterns. For instance, tracking approval rates for loans or ads by demographic group can reveal potential biases.
- **AI System Uptime and Reliability**: Monitor the operational performance of AI systems, including downtime incidents and maintenance issues, to assess their reliability and the effectiveness of support systems.
- **Customer Engagement Metrics**: Use AI analytics to track changes in customer engagement levels, such as click-through or conversion rates, before and after implementing AI-driven marketing strategies. This can indicate the impact of AI on customer behavior.
- **Hallucination Rates**: Measure the frequency of AI-generated content that is false, misleading, or inconsistent with the input data. High rates may indicate issues with the AI model, requiring investigation and refinement.
- **Safety Violation Attempts**: Track the number and types of attempts to circumvent AI safety measures or create harmful content. This metric should include instances of privacy breaches, unauthorized high-stakes decisions, deception or misinformation attempts, and failures to disclose AI use. Monitoring this metric over time can help identify vulnerabilities in the AI system, inform the development of stronger safeguards, and measure the effectiveness of existing security measures.
- **System Ineffectiveness**: Track cases where AI systems fail to perform as intended or produce suboptimal results. This can highlight areas for improvement and guide refinements to AI models and processes.
- **Third-Party Risk Assessment**: Regularly assess the risks associated with third-party AI vendors, including their data security practices, model performance, and compliance with service level agreements (SLAs).
- **Employee AI Literacy**: Through assessments and surveys, measure the effectiveness of employee training programs on AI ethics, bias detection, and responsible AI practices. A well-informed workforce can help mitigate AI risks.
- **Propensity to Buy**: Measure the impact of AI-driven marketing on customers' purchase likelihood.
- **Basket Size/Value Impact**: Track changes in the average order value or basket size resulting from AI-powered personalization and recommendations.
- **Sales Impact**: Monitor the overall effect of AI marketing initiatives on sales volume and revenue.
- **Sales Frequency Impact**: Evaluate how AI-driven strategies influence the frequency of customer purchases.

Qualitative Assessments

Qualitative assessments provide context and insights into the impacts of AI that quantitative data alone cannot capture. They involve subjective analysis and interpretation to evaluate how AI-driven marketing aligns with ethical standards, customer expectations, and societal values. Examples include:

- **Customer Sentiment Analysis**: Conduct surveys, focus groups, or social media sentiment analysis to understand public perception and reaction to AI-driven marketing practices. This can reveal concerns about privacy, personalization, or ethical considerations that may not be evident through quantitative data alone.
- **Ethical Evaluations**: Perform regular ethical audits or reviews of AI systems to assess their alignment with ethical guidelines, such as fairness, transparency, and accountability. This could involve evaluating decision-making processes, data-handling practices, and the potential societal impact of AI applications.
- **Stakeholder Interviews**: Engage with a broad range of stakeholders, including customers, employees, and industry experts, to gather insights into the perceived benefits and challenges of AI in marketing. This can help identify areas for improvement and innovation.
- **Case Studies and Benchmarking**: Analyze specific instances of AI use within the organization or compare practices with industry benchmarks to identify best practices and areas where ethical, privacy, or compliance risks may arise.
- **User Experience Testing**: Conduct usability tests or user interviews to assess how customers interact with and perceive AI-driven marketing experiences. This can provide valuable insights into the user-friendliness, effectiveness, and potential challenges of AI implementations from a customer perspective.
- **Scenario Planning and Stress Testing**: Develop hypothetical scenarios or simulations to assess how AI systems and marketing strategies might perform under different conditions, such as data breaches, system failures, or unexpected customer behaviors. This can help identify potential vulnerabilities and inform the development of contingency plans.
- **Competitor Analysis**: Conduct a qualitative analysis of competitors' AI-driven marketing practices to identify industry trends, best practices, and potential areas for differentiation. This can provide valuable insights into how other organizations are addressing ethical, privacy, and compliance risks associated with AI in marketing.
- **Employee Feedback and Observations**: Gather feedback and observations from employees directly involved in the development, implementation, and monitoring of AI-driven marketing initiatives. Their firsthand experiences can provide valuable insights into the challenges, successes, and potential risks associated with these efforts.

Part II: Risk Rating Criteria and Levels

LIME Risk Rating Criteria

Our LIME risk rating criteria for assessing AI risks in marketing projects are defined by three key factors: Likelihood, Impact, and Mitigation Effort. This approach quantifies and prioritizes risks, ensuring marketing teams can focus on the most critical issues.

Likelihood of Occurrence

How to Estimate: Assess the probability of a risk occurring based on historical data and expert judgment. Use a scale from 0 to 1.0, where 0 represents a nonexistent probability, and 1 represents a near-certain probability.

Example

- **Project 1**: Historical data from the AI Incident Database indicates that AI projects involving extensive customer data collection have experienced frequent data breaches in the past year. The likelihood might be rated as 0.8.
- **Project 2**: For a new generative AI ad personalization tool that is only accessible to paying customers who go through a vetting process, the likelihood might be rated as 0.4.
- **Project 3**: An AI system for automating customer service responses, previously resulting in miscommunication or customer dissatisfaction in 10% of interactions, might have a 0.1 likelihood rating.

Potential Business and Reputational Impact

How to Estimate: Determine how significantly a risk could impact the business. Consider both direct impacts, like financial loss, and indirect impacts, like reputational damage. Use a scale where 1 represents minimal impact, 2 low, 3 medium, 4 high, and 5 massive.

Example

- **Project 1**: Biased outputs in an AI segmentation tool for a financial product would severely damage the brand's reputation, leading to customer churn and potential lawsuits. This might score a 5/5 for massive impact.
- **Project 2**: An AI tool for personalizing marketing messages, if misused, could lead to over-personalization and privacy concerns, potentially scoring a 3/5 for medium impact due to the potential for customer pushback and regulatory scrutiny. If this tool were used by political campaigns, it might score 4/5 for high impact.
- **Project 3**: Deploying a new AI chatbot for customer service, improving response times and accuracy, could have a minor risk of misinterpretation. However, given its positive reception and quick rectification capabilities, the impact might be rated as 2/5 for lo.

Mitigation Effort

How to Estimate: Estimate the effort required to mitigate the risk, considering all necessary actions to prevent or lessen the impact. Use "person-months", where 1 represents the work one employee can do in a month.

Example

- **Project 1 (Securing Generative AI Tool)**: For a project aimed at preventing the generation of inappropriate AI content, you might need 1 month for evaluating content moderation tools, 2 months for integrating these tools into your existing systems, and

another 1 month for testing and adjustments. Training for content managers on new protocols could require 0.5 months, totaling 4.5 person-months.

- **Project 2 (Customer Data Consent Update)**: Updating customer consent mechanisms to permit one's company to use user-generated content to train AI models might involve 0.5 months for legal consultation to understand requirements, 1.5 months to redesign the consent collection process, and 1 month for implementing and testing the new system. Training customer service staff on handling consent queries could take another 0.5 months, adding up to 3.5 person-months.
- **Project 3 (AI-Driven Customer Service Improvement)**: To mitigate risks associated with an AI-driven customer service system that may misinterpret customer queries, you might allocate 0.5 months for gathering and analyzing feedback on current issues, 1 month for developing enhancements to the AI model, and 1 month for deploying updates and monitoring improvements. An additional 0.5 months could be needed to retrain customer service representatives on interfacing with the updated system, totaling 3 person-months.

Calculating the Risk Score

$$\text{Risk Score} = \text{Likelihood} \times \text{Impact}$$

This calculation quantifies the total risk in terms of potential impact, guiding the prioritization of mitigation efforts based on the severity of the risk involved. While the Risk Score helps prioritize risks, the Mitigation Effort should be used to inform resource allocation and project planning. High-risk items with low mitigation effort might be addressed quickly, while high-risk items with high mitigation effort might require long-term strategic planning. Low-risk items with high mitigation effort might be deprioritized or addressed through alternative means. This approach ensures a balanced focus on preventing severe impacts and managing risks efficiently while acknowledging the practical considerations of mitigation efforts in the overall risk management strategy.

Practical Application with Examples

Expanding on the practical application of our risk assessment model, we integrate a more comprehensive array of examples across various risk levels, repurposing insights from the detailed risk levels previously discussed. These examples illustrate how to apply the risk score formula to various AI-driven marketing scenarios.

Translating the various risk considerations into likelihood and impact scores can be challenging. Some considerations have 'yes/no' answers, while others are more free-text-based. Using overall judgment to assign scores for each risk type is recommended rather than attempting to score individual considerations.

Content & Creative

1. Content Discovery (Tagging, Taxonomy)

- **Likelihood**: 0.4 (Well-developed technologies but with room for error)
- **Impact**: 2 (Low, incorrect tagging might lead to misclassification)
- **Mitigation Effort**: 2 person-months (Improving AI tagging algorithms, regular audits)
- **Risk Score**: 0.4 × 2 = 0.8

2. International Translations and Transcreation of Content

- **Likelihood**: 0.7 (High potential for cultural misinterpretation)
- **Impact**: 4 (Very high, can lead to brand damage due to cultural insensitivity)
- **Mitigation Effort**: 5 person-months (Cultural sensitivity training, expert review panels)
- **Risk Score**: 0.7 × 4 = 2.8

3. Generative AI and Personalization in Ad Creative Testing and Optimization

- **Likelihood**: 0.6 (Growing adoption but concerns when operated without human oversight and quality of input data)
- **Impact**: 4 (High, due to potential backlash over privacy invasion or pandering when highly personalized, inappropriate content, and concerns about copyrighted materials in training data)
- **Mitigation Effort**: 4 person-months (Privacy compliance measures, content review protocols, ethical usage guidelines)
- **Risk Score**: 0.6 × 4 = 2.4

Audience & Media

4. Segmenting Audiences Using AI

- **Likelihood**: 0.5 (Widely used with potential for bias)
- **Impact**: 3 (Medium, risk of alienating groups due to biased segmentation)
- **Mitigation Effort**: 3 person-months (Bias audits, diversity in training datasets)
- **Risk Score**: 0.5 × 3 = 1.5

5. AI for Adjusting Bids on Advertising Platforms

- **Likelihood**: 0.3 (Stable technology with minor risks of inaccuracies in bid adjustments)
- **Impact**: 2 (Low, potential for minor inefficiencies in ad spend without significant financial jeopardy)
- **Mitigation Effort**: 2 person-months (Refining AI algorithms, setting bid limits, periodic performance reviews)
- **Risk Score**: 0.3 × 2 = 0.6

6. AI-driven Dynamic Pricing in E-Commerce

- **Likelihood**: 0.5 (Complexity varies)
- **Impact**: 4 (High, potential for customer dissatisfaction with price fluctuations)

- **Mitigation Effort**: 4 person-months (Transparent communication strategies, price fairness algorithms)
- **Risk Score**: 0.5 × 4 = 2.0

Measurement & Attribution

7. AI-driven Insights into Audience Behavior and Sentiment

- **Likelihood**: 0.3 (Data analytics is generally reliable)
- **Impact**: 3 (Medium, incorrect insights could lead to misguided strategies)
- **Mitigation Effort**: 3 person-months (Data validation, model testing)
- **Risk Score**: 0.3 × 3 = 0.9

8. AI in Customer Journey Mapping

- **Likelihood**: 0.4 (Data-driven but complex)
- **Impact**: 3 (Medium, inaccurate mapping could misguide customer engagement strategies)
- **Mitigation Effort**: 3 person-months (Integrating customer feedback loops, continuous model improvement)
- **Risk Score**: 0.4 × 3 = 1.2

Customer Experience

9. Fully Automated Lead Scoring:

- **Likelihood**: 0.6 (Due to the incomplete nature of training data across geographies, some issues are likely)
- **Impact**: 2 (Low, due to potentially disrupted services groups underrepresented in training data)
- **Mitigation Effort**: 3 person-months (For thorough testing and validation processes)
- **Risk Score**: 0.6 × 2 = 1.2

10. LLM-Powered Customer Service Using Conversational AI

- **Likelihood**: 0.4 (Commonplace but varies in sophistication)
- **Impact**: 2 (Low, given well-established protocols to escalate complex issues, varying risks when providing incorrect information)
- **Mitigation Effort**: 2 person-months (Continuous training on new queries, monitoring for escalation accuracy)
- **Risk Score**: 0.4 × 2 = 0.8

Part III: Risk Mitigation Strategies and Incident Response

Risk Mitigation Strategies

Risk mitigation involves implementing strategies to minimize the adverse effects of risks. The three primary strategies include:

1. **Avoidance**: This strategy involves proactively identifying activities that pose risks and deciding against their implementation. For example, avoiding the use of controversial AI applications in marketing to prevent ethical dilemmas. Controls include strict policy guidelines and decision-making frameworks to prevent engagement in high-risk activities.
2. **Transfer**: Risk transfer shifts the burden of risk to a third party, such as through insurance or outsourcing operations to vendors with robust security protocols. An example is using third-party vendors with strong security measures for data processing, transferring the cybersecurity risk to them. Contracts and insurance policies are typical controls.
3. **Mitigation**: Mitigation involves taking steps to reduce the severity or likelihood of the risk. Controls include regular security audits, employee training, and the adoption of best practices in AI ethics. There are a large and growing number of companies that specialize in AI data auditing and governance (such as Anch.AI, Chatterbox Labs, Credo AI, Kosa.AI, ORCAA, QuantPi, and Themis AI) and model validation and monitoring (Aporia, Arize, Arthur, Datatron, Deepchecks, Fiddler, Lakera, LatticeFlow, ModelOp, NannyML, Robust Intelligence, Snitch AI, TruEra, TrojAI, Vianai, Whylabs, and Yields.io) to assist with risk mitigation.

Incident Response Protocols

A comprehensive incident response plan is crucial for effectively managing and recovering from unexpected events.

Severity Levels

Establish clear criteria for classifying the severity of incidents to prioritize response efforts effectively. Criteria may include the impact on customer data, financial losses, and brand reputation. Severity classifications can range from:

- **Critical**: Major data breach affecting sensitive customer information.
- **High**: Legal non-compliance with significant fines.
- **Medium**: Minor data leakage with limited impact.
- **Low**: Minimal impact and easily contained.

Notification Procedures

Develop standardized procedures for notifying internal stakeholders and regulatory bodies. This includes immediate escalation paths, communication templates, and reporting timelines, ensuring compliance with requirements such as those mandated by GDPR and various US state privacy laws.

Investigation Process

Outline the steps for initial assessment and root cause analysis to understand the incident's scope and origins. Steps include:

- Immediate containment and isolation of affected systems.
- Gathering and documenting all relevant information.
- Conducting a thorough investigation to identify the cause and affected areas.

Remediation Plans

Formulate a two-pronged remediation approach that addresses immediate response and long-term preventive measures. Remediation strategies should incorporate:

- Immediate interventions to address the specific incident.
- Long-term changes to policies, procedures, and technologies to prevent recurrence.
- Continuous monitoring and review of systems and processes to detect and mitigate future risks.

Conclusion and Next Steps

The Marketing AI Risk Evaluation Framework represents a dynamic tool designed to navigate AI in marketing. As we continue to explore the possibilities of artificial intelligence, the necessity for a responsible approach to AI integration becomes increasingly paramount. This framework will serve as a foundation for future exploration and adaptation in the face of technological advancements and emerging ethical considerations.

Future Directions and Considerations

In refining our framework, we acknowledge the assumptions made and recognize the avenues for further research and development. For instance, the impact of creative quality in AI-generated content and personalized advertising campaigns holds significant potential for enhancing marketing outcomes. Delving deeper into the interplay between AI-driven personalization, creative effectiveness, and consumer engagement could uncover new strategies for maximizing campaign performance.

Furthermore, the framework could benefit from an expanded focus on consumer data privacy, exploring more sophisticated AI applications to ensure data protection while enhancing user experience. Integrating AI in data privacy management, particularly in automating compliance with global regulations, presents an opportunity for further investigation.

Another critical area for future development is exploring AI's role in sustainable marketing practices. As businesses increasingly prioritize sustainability, AI could offer insights into creating more environmentally friendly marketing strategies, reducing waste through targeted advertising, and optimizing digital footprints.

Expanding the Framework

This framework should evolve to incorporate new AI technologies and marketing channels to remain relevant and effective, reflecting changes in consumer behavior and regulatory landscapes. Regular updates, informed by industry best practices and academic research, will ensure that the framework continues to provide valuable guidance for marketers.

Engaging with a broader community of marketing professionals, technologists, ethicists, and regulators will enrich the framework, incorporating diverse perspectives and expertise. Collaborative efforts can lead to the development of more nuanced risk assessment tools and innovative mitigation strategies, fostering a culture of responsibility and transparency in using AI in marketing.

Long-term Impact and Legacy

By adhering to and continually improving this framework, organizations can responsibly harness AI's full potential to drive ethical, effective, and innovative marketing strategies. The goal is to create an environment where AI empowers marketers to create more meaningful, personalized, and engaging consumer experiences without compromising ethical standards or consumer trust.

As we look to the future, the Marketing AI Risk Evaluation Framework aspires to be at the forefront of guiding responsible AI integration in marketing, contributing to developing marketing practices that propel business growth.

Acknowledgment

We extend our gratitude to a distinguished group of individuals whose contributions have supported the development of the Marketing AI Implementation Checklist, a guide for adopting AI technologies responsibly and effectively.

Alec Foster's Contributions:

Our special thanks to Alec Foster, whose invaluable insights and profound expertise in responsible AI and marketing have been instrumental in creating and shaping this framework.

MMA's AI Leadership Coalition (ALC):

We also acknowledge MMA Global's AI Leadership Coalition (ALC). Comprising over 300 members from more than 160 brands, the ALC is the largest working group dedicated to responsibly and effectively applying AI in marketing.

ALC Contributors:

We are appreciative of the experience brought in by these key contributors from the ALC:

- **Joe Veverka**, Microsoft
- **Justin Thomas-Copeland**

We also extend our thanks to **Greg Stuart**, MMA Global CEO, **Rex Briggs**, ALC's other subject matter expert, and **Hassan Khater**, ALC's program manager, for their guidance in shaping the framework. And thanks to all the members of MMA's ALC below.

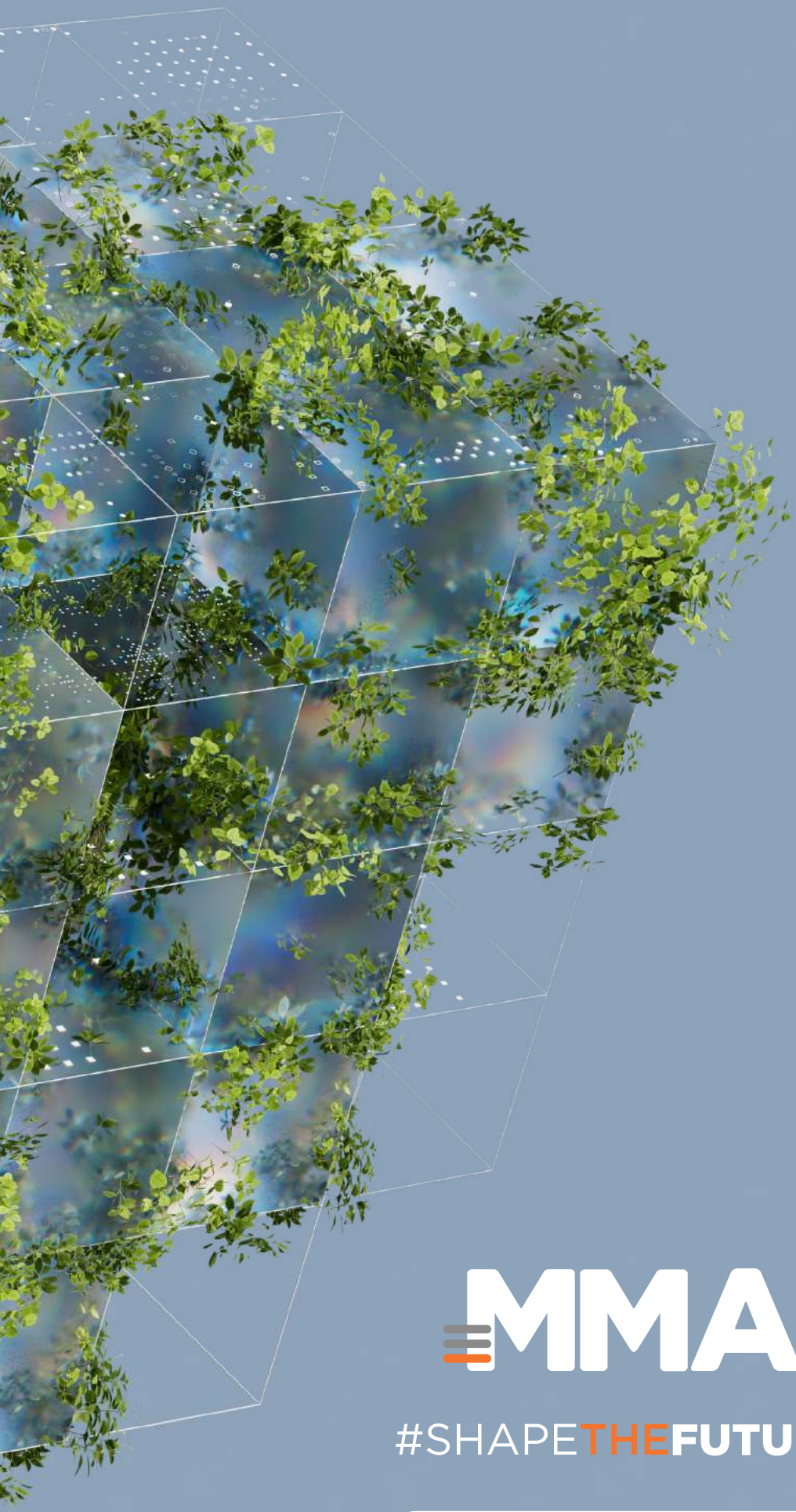
ALC Members:

About the Author

Alec Foster, a trailblazer in responsible AI and ethical marketing, brings over a decade of expertise honed at Google, the White House, and entrepreneurial ventures. As MMA Global's Responsible AI Subject Matter Expert, Alec champions social progress and sustainable innovation through responsible AI practices in marketing. With an M.St. in AI Ethics from the University of Cambridge and a CIPP/US privacy certification, Alec excels in crafting AI-driven growth strategies that align consumer trust with business objectives. His contributions include developing machine learning models that simplify consumer arbitration and serving on the boards of algorithmic accountability nonprofits. Guided by his mission to champion transparent, respectful, and integrity-driven technology, Alec Foster leaves an indelible mark on the world of responsible AI and ethical marketing.

About MMA

Comprised of over 800 member companies globally and 15 regional offices, the MMA is the only marketing trade association that brings together the full ecosystem of marketers, martech, and media companies working collaboratively to architect the future of marketing while relentlessly delivering growth today. Led by CMOs, the MMA helps marketers lead the imperative for marketing change – in ways that enable future breakthroughs while optimizing current activities. The MMA is committed to science and questioning and believes that creating marketing impact is steeped in constructively challenging the status quo, encouraging business leaders to aggressively adopt proven, peer-driven, and scientific best practices without compromise. The MMA invests millions of dollars in rigorous research to equip marketers with unassailable truth and actionable tools. By enlightening, empowering, and enabling marketers, the MMA shapes future success while propelling business growth.



MMA

#SHAPETHEFUTURE®

Visit mmaglobal.com for more