AI Disclosure and Transparency:
Closing the Trust Gap

A Deep-Dive Industry Analysis from Big Valley Marketing

BIG VALLEY
MARKETING

INTRODUCTION

It's been less than two years since the launch of ChatGPT, and the world is still figuring out how to react. Is this the next great productivity breakthrough, or is this the moment machines start to take over the world? Will this help people do their jobs better, or replace them with automated systems? Will creativity die, or will it flourish?
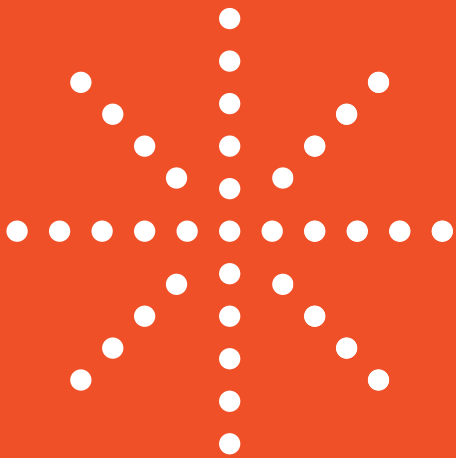
These are just some of the many questions swirling around business, technology and society since ChatGPT showed us that generative AI can read, research, write and talk like humans. Or at least enough of an approximation to amaze and scare us.
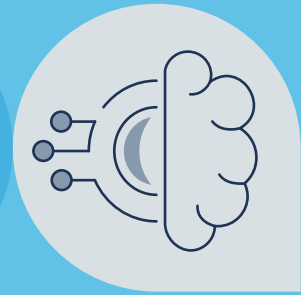
In the wake of that discovery, many well-meaning people and organizations have tried to develop guardrails to limit or direct how AI is developed, deployed and consumed across industries. One of the many guardrails proposed has been *disclosure* about when, where and how AI is being used. The thinking: If we're transparent about AI use – whether it's in our writing, or our thinking, or our products – people can then judge for themselves how they feel about it.

The challenge is that AI disclosure means different things to different people. Some would argue there's no need for a writer to disclose they used AI for research or editing help, any more than they would disclose their use of Google search and Microsoft spell check. Others would argue that there are so many questions and limitations about AI use that disclosure can help minimize bias, plagiarism, misinformation and copyright violations. If we simply know the "chain of custody" in how the AI was used, readers and users can judge for themselves.

The other assumption built into AI disclosure is that "transparency breeds trust." It's a natural instinct, and one that's grounded in plenty of research and history across business, marketing and communications circles. We all generally *want* our friends and our neighbors, our brands and our government leaders, to be transparent with us. And in many if not most cases, that transparency *does* increase trust. Unfortunately, in this case, it's a major fallacy that ignores what research now shows – roughly 80% of the general public does not *trust* AI. So, if you publish a blog post and tell people AI helped you write it, people may actually trust it (and you) *less* than if you didn't tell them. Combine that with the well-known and growing mistrust of organizations and institutions, and it turns out that AI disclosure could present greater challenges to existing trust gaps.

So where do we go from here?

How should business, marketing and communication leaders navigate this evolving environment?

DEEP-DIVE analysis

To answer these questions, Big Valley invested in a deep-dive analysis of the business and academic literature published in the last year about AI disclosure and transparency. Our team – of real people, using various research tools – reviewed and analyzed thousands of published documents, proposals, blog posts and guidelines shared by hundreds of media outlets, industry organizations, companies and academic institutions. That analysis has led to several very clear insights in a murky and still-developing area of study:

1. **It's a mess out there, with a broad range of guidelines** covering a broad range of topics but precious little consensus or commonality.

2. **Most AI disclosure guidelines are "first wave" efforts** that draw primarily from an organization's existing ethics, brand or professional guidelines.

3. **Most of the current guidelines focus on the what and how of AI disclosure,** not on the who and why – which are more critical to building trust.

4. **Academia and journalism have outlined the most restrictive guidelines,** but it's unclear how effective they are – or how applicable to other industries and use cases.

5. **Most of the current guidelines are based on opinions and beliefs,** with very little empirical research to demonstrate what actually resonates with customers and citizens.

6. **Technologists (or technology evangelists) have been the leading voices** on AI ethics and disclosure.

7. **Business leaders, including marketing and communications leaders, need to step up their engagement** to provide a more balanced perspective – and fuel adoption that makes sense to real people operating in real-world scenarios.

Research Focus: Key Questions

1 How, if at all, do various sectors (e.g., technology, government, academia, etc.) discuss disclosure and transparency as related to AI use?

2 What expectations do customers have for brands to be transparent in disclosing how they use AI in their products?

3 What expectations do customers have for brands to be transparent in disclosing how they use AI in their AI marketing and communications?

4 What marketing and communications approaches, if any, have been proposed for AI transparency and/or disclosure?

5 What impact might government organizations and elected officials have on AI use disclosure requirements?

Why Should Marketers and Communicators Care About AI Disclosure?

Despite the positive impacts AI may have on **marketing and communications** productivity, companies must confront questions about transparency and disclosure because of the popular concerns over bias, inaccuracy, and copyrighted AI-generated material. Further, responsible use of AI can be perceived as relative to a company's corporate digital responsibility initiatives and ethos. Transparency and disclosure questions matter for both AI technology developers and users of AI systems, especially those who use AI tools for professional purposes. Indeed, misguided or perceived deceitful AI usage could impact a company's reputation and trust.

- **Bias**
- **Inaccuracy**
- **Copyrights**
- **ESG**
- **Ethics**
- **Safety**
- **Sustainability**
- **Reputation**
- **Trust**

Consumer / Brand POV

Companies are racing to adopt AI into their various processes and offerings without effectively managing the complicated topic of AI use and disclosure policies. However, early actions by some brands show that adopting AI use and disclosure policies that reflect the organization's mission and values demonstrates thoughtful consideration of AI and helps promote trust among customers.

For companies seeking to define an approach to AI disclosure, the first question to ask is if the companies in question allow any level of AI use in the workplace. For companies seeking to define an approach to AI disclosure, the first question to ask is if the companies in question allow any level of AI use in the workplace. The absence of AI trainings and policies directly impacts employee use. A Fishbowl Insights survey found that 68% of employees do not disclose AI use to their managers. These findings demonstrate the increasing complexity over what to do regarding AI in the workplace.

Further, internal AI policies have the possibility to affect external perceptions and overall brand reputation (ex. OpenAI's 2024 dialogues on transforming into a for-profit company). Indeed, after establishing internal guidelines, organizations face the daunting task of communicating their AI use to consumer groups in a way that increases, or at the very least maintains, trust.

How do organizations communicate their AI use to consumers who view AI critically—with trust, attribution and authenticity at the core of their concerns?

Some companies evaluated included:

- **YouTube**
- **TikTok**
- **Meta**
- **LinkedIn**
- **Microsoft**
- **Dove**
- **Sports Illustrated**
- **Home Depot**
- **McKinsey**
- **Ernst and Young (EY)**

ONLY

21%

of respondents
had policies and trainings for AI use in the workplace

YouTube and TikTok based their approaches to disclosing the use of AI on the realism of the content that was produced, presumably to mitigate risks from deepfakes or other altered videos. In those policies, meaningful disclosure is not required during the creative or brainstorming processes (ex. script generation, auto captions, photoshop for enhancements). TikTok automatically labels content based on the significance of AI alterations, aiming to prevent misleading information and ensuring community guideline compliance through the ethical justification of the AI content. In a simpler standard practice, Meta requires content to have the label "AI Info" if a user uploads content or if AI generation is detected. LinkedIn, on the other hand, includes a "Learn more" link when users interact with their AI tools.

There are various strategies companies can consider, such as disclosing through behavioral outputs (ex. robotic persona), verbal disclosures (ex. statements of AI use), and technical disclosures (ex. metadata). To add to the consideration list, organizations must decide on the tactic itself, disclosing via a content label, watermark, byline, acknowledgments, etc. Selecting a disclosure approach leads to other serious considerations, including (1) label clarity, (2) label consistency, and (3) label location.

Establishing clear and informed AI policies gives companies a structured approach to tackling questions about when and how to disclose.

Industry POV

As a cornerstone of the organization-consumer relationship, marketing and communication trade organizations provide insightful perspectives on the conversations around AI transparency and disclosure.

Currently, these trade organizations (1) leverage existing guidelines, (2) consult experts, and (3) offer educational resources to standards to steer AI disclosure scenarios to help inform business, marketing and communications leaders on the development of AI disclosure policies for their brands.

Aligning policies for AI use with current industry standards and best practices enables brands to separate the hype and promise around AI from the need to evaluate and approach the technology in the same way it would consider any new technology tool. For example, the Association of National Advertisers (ANA) advises adopting current ethical guidelines to align with AI, suggesting businesses to "go back to basics" to ensure general expectations and requirements, such as customer satisfaction, corporate citizenship and accurate product delivery are met.

Collectively, professional communications and technology organizations with published AI principles offer transparency as a strategy to mitigate bias and foster accountability. However, these trade groups also underscore the legal consequences and contractual obligations of transparency. The PR Council addressed misinformation and legal compliance by further emphasizing transparency requirements across all company relationships. However, defining and providing appropriate levels of disclosure is inconsistent and in various cases, vague.

Trade organizations also encourage professionals to consult existing codes of ethics and professional values to navigate questions about transparency and disclosure. Overall, their ethical considerations focus on positioning AI as a tool capable of enhancing productivity and company goals as well as complementing human talent by exercising human oversight and guardrails. For example, the Association for Computing Machinery (ACM) directly links their ethical guidelines in their AI discussions, while PR Council publishes updated AI guidelines founded on its basic code of ethics but tailored specifically for AI use and disclosure.

Marketing and Communications Industry Organizations evaluated as part of this report.

In parallel, several trade groups focused on the term "substantial use" as a proposed milestone for triggering AI disclosure. **The phrase "substantial use" has a legal lineage and is most commonly used in copyright law (ex. the fair use doctrine). The adoption of a legal principle into trade organization guidelines demonstrates how trade groups are using known guidelines or principles and applying them to AI use.**

For example:

- The Institute for Public Relations (IPR) and Association for Computing Machinery (ACM) created policies around using AI in research publications, stating that the level of AI disclosure should be proportional to the amount that AI was used. For example, AI use needs to be disclosed if it translates or creates written or visual content (among other uses) but does not normally need to be disclosed if used for grammatical editing or other uses that do not substantially change the content.

- Comparatively, the PR Council recommends agencies who infrequently use GenAI tools to disclose substantive AI use in a case-by-case basis.

- The PRSA's guidelines on ethical AI use also mention that use of GenAI should have substantial changes by human creators to avoid copyright and plagiarism concerns. However, the policies lack a defined threshold for determining when content is deemed "substantially changed," a designation that is not thoroughly addressed by professional organizations.

Fostered by uncertainty and a lack of familiarity with AI, these organizations also aim to fill the gap by sharing educational resources related to disclosure and transparency. The 4 A's crash course website offers materials for professionals to enrich their understanding of AI and its implications. The organization also offers an adoptable Agency Policy Template for Generative AI so organizations can customize their own "do" and "don't" list on how and when to deploy AI (both internally and externally). More explicitly, the Cloud Security Alliance (CSA) developed the AI Safety Initiative, a task force dedicated to publishing free content related AI best practices, regularly updated guidelines to reflect regulation, and appropriate uses of AI in cybersecurity.

Regardless of the approach to AI disclosure (ex. leveraging existing guidelines versus consulting experts), trade groups, organizations and brands alike need to stay abreast of the development of AI-related regulation coming from governmental entities.

While industry trade group AI disclosure recommendations continue to vary, alignment across several concepts has emerged:

- There is need for additional guidance and education around appropriate use of AI

- Experts encourage companies to connect existing ethical codes to AI-application

- Transparency and disclosure conversations in this context often center around mitigating AI risks related to plagiarism, hallucinations, copyright violations, bias, and breach of security/confidentiality. More dialogue on these topics will help establish shared best practices.

AI Disclosure: Government

Outside the scope of company and trade group policies and recommendations, generative AI inspired widespread, bi-partisan interest from U.S. government organizations and elected officials. More than 20 bills related to AI disclosure or transparency were introduced to the 118th Congress; the Securities and Exchange Commission (SEC) and Federal Trade Commission websites feature a wide range of blog posts and press releases on businesses and AI use transparency; and the Biden Administration released a Blueprint for an AI Bill of Rights in October 2022 and two additional executive orders about AI.

It's clear that legislators have grasped the potential impacts of AI on their constituents and argue that regulation will generate public trust in the technology. Proposed regulations include requiring all generative AI outputs to be labeled and disclosed; ensuring generative AI does not create misleading information regarding elections; protecting personal data; avoiding bias and discrimination; protecting copyrights; informing investors of relevant risks related to AI; and protecting American rights as the technology evolves.

Government entities have limited understanding of and alignment on the full scope of generative AI technology. These educational and alignment gaps have contributed to slow progress around legislation and regulation. AI regulation has been proposed at the federal level as well as across many states, opening up the potential for conflicting guidelines and regulations. Globally, the European Union was the first to implement clear boundaries on the use of the technology.

In essence, the already multi-faceted AI disclosure conversation became multi-layered with the potential for regulations at the organizational, state, federal and global levels.

How are Government Leaders and Organizations Approaching AI Disclosure Issues?

GOVERNMENT BILLS/REGS

● 4/9/24   H.R. 7913
Generative AI Copyright Disclosure Act of 2024
Schiff, Adam B. [Rep.-D-CA-30]

● 3/21/24   H.R. 7766
Protecting Consumers from Deceptive AI Act
Eshoo, Anna [Rep.-D-CA-16]

● 3/5/24   H.R. 7532
Federal AI Governance and Transparency Act
Comer, James [Rep.-R-KY-1]

● 1/29/24   H.R. 7120
R U REAL Act
Schakowsky, Janice D. [Rep.-D-IL-9]

● 1/29/24   H.R. 7123
QUIET Act
Sorensen, Eric [Rep.-D-IL-17]

● 1/30/24   S. 3686
Preventing Algorithmic Collusion Act of 2024
Klobuchar, Amy [Sen.-D-MN]

● 12/22/23   H.R. 6881
AI Foundation Model Transparency Act of 2023
Beyer, Donald S. Jr. [Rep.-D-VA-8]

● 11/21/23   H.R. 6466
AI Labeling Act of 2023
Kean, Thomas H. [Rep.-R-NJ-7]

● 9/20/23   H.R. 5586
DEEPFAKES Accountability Act
Clarke, Yvette D. [Rep.-D-NY-9]

● 7/27/23   S. 2708
Deceptive Experiences to Online Users Reduction Act ("DETOUR Act")
Warner, Mark [Sen.-D-VA]

● 7/27/23   S. 2597
Digital Consumer Protection Commission Act of 2023
Warren, Elizabeth [Sen.-D-MA]

● 7/27/23   S. 2691
AI Labeling Act of 2023
Schatz, Brian [Sen.-D-HI]

● 7/13/23   S. 2325
Algorithmic Justice and Online Transparency Act
Markey, Edward J. [Sen.-D-MA]

● 7/13/23   H.R. 4624
Algorithmic Justice and Online Transparency Act
Matsui, Doris O. [Rep.-D-CA-7]

● 6/5/23   H.R. 3831
AI Disclosure Act of 2023
Torres, Ritchie [Rep.-D-NY-15]

● 5/15/23   S. 1596
EAL Political Advertisements Act
Klobuchar, Amy [Sen.-D-MN]

● 5/5/23   H.R. 3106
Preventing Deepfakes of Intimate Images Act
Morelle, Joseph [Rep.-D-NY-25[MN3] [MN4]]

● 5/02/23   H.R. 3044
REAL Political Advertisements Act
Clarke, Yvette D. [Rep.-D-NY-9]

● 5/2/23   S. 1409
Kids Online Safety Act
Blumenthal, Richard [Sen.-D-CT]

● 4/24/23   H.R. 2801
Kids PRIVACY Act
Castor, Kathy [Rep.-D-FL-14]

● 2/2/23   S. 262
Stop Spying Bosses Act
Casey, Robert P., Jr. [Sen.-D-PA]

What Customers Think

Use in products

A person's knowledge and personal use of AI impacts how they trust AI disclosure in products. Those who engage with AI less often perceive AI as less trustworthy than those who engage with AI regularly. Similarly, a lack of AI knowledge leads to heightened distrust with AI disclosure in products. Big Valley's evaluation unearthed several key findings that exemplify varying points of view on the use of AI in products:

55%
of consumers
do not believe that companies are considering ethics when developing AI tools.

(Source: Markkula Center for Applied Ethics)

44%
of American respondents
do not think they regularly interact with AI

(Source: Pew Research Center)

79%
of Americans
across all generations do not trust businesses to use AI responsibly.

(Source: Bentley and Gallup poll)

19%
of UK consumers
are confident in their ability to detect AI-generated content, while other respondents stated they had little confidence (36%) or no confidence at all (29%). In parallel, the same survey showed 82% of UK consumers want transparency and clear disclosure by brands

(Source: YouGov)
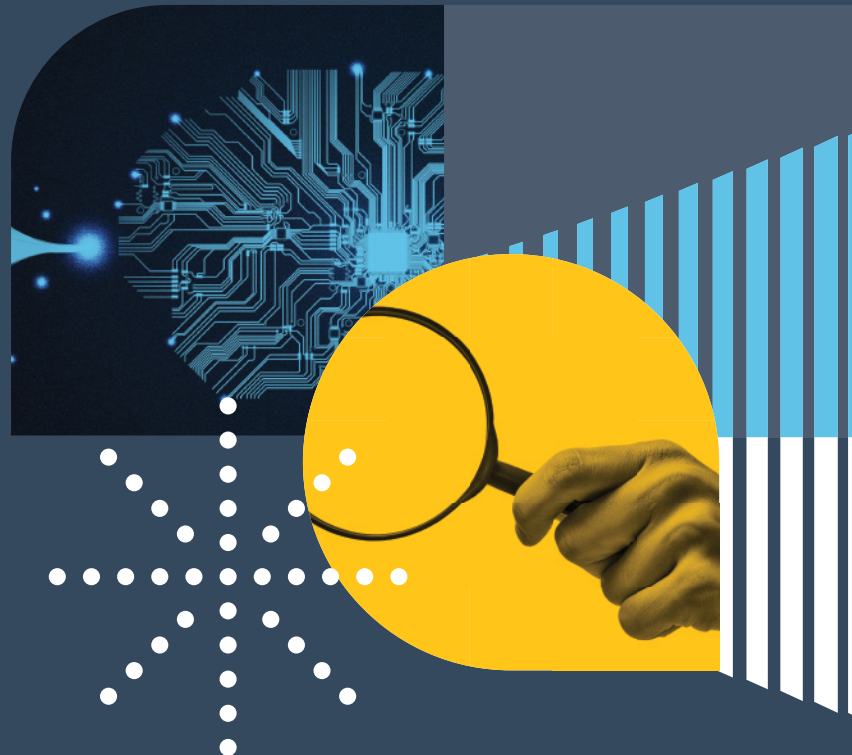
What Customers Think

Expectations for disclosing use of AI in marketing and comms

In general, Big Valley's analysis reveals a strong desire across diverse audiences and geographies for transparency when marketing, advertising, or communication materials contain AI-generated content.

74%
of participants
wish "more companies disclosed when they use AI and why it was used," highlighting a knowledge gap between advertisers and consumers about AI used in marketing and advertising.

(Source: Yahoo/Publicis survey)

72%
of respondents
state they wanted to know when advertisements are created with AI.

(Source: Survey by the 4A's)

90%
of respondents
to a Getty Image report wanted to know when an image was generated with AI.

76%
of US respondents
want to know if election content they were seeing was AI-generated to help them address concerns about misinformation and deepfakes.

(Source: Adobe 2024)

Big Valley Perspective:

What does this all mean for business, marketing and communications leaders?

So What?

- Disclosure and transparency are often proposed as strategies to mitigate risks of AI, but there is not a standardized approach to disclosure.

- Some prominent reasons cited for AI disclosure include fostering trust in users, encouraging responsible use, and protecting consumers.

- Consumer uncertainty elevates the need for disclosure, leading experts to consider disclosure a best practice.

- Disclosure and transparency should empower humans using the technology and receiving the content.

Now What?

- With GenAI increasingly integrated into marketing and communications, brands should at least start to consider how transparent they will be regarding their use of AI to create content.

- By all accounts, AI disclosure is moving toward being a requirement for maintaining trust with customers, employees, partners and the broader public.

- AI disclosure can work in partnership with efforts for defining ethical and fair AI practices and create new strength for safeguarding brand reputations and fostering environments of accountability and transparency.

- Proactive labeling of the use of AI in generating content comes with potential benefits and unforeseen consequences.

AI DISCLOSURE PRINCIPLES

With generative AI increasingly integrated into marketing and communications, companies face growing expectations to be transparent in their use of AI to create content. The following principles are designed to guide any organization in determining the most appropriate GenAI disclosure strategy and processes.

Align Disclosure with Brand Strategy

Ensure that public disclosures about AI are framed in a way that enhances the brand strategy, demonstrating how GenAI content supports brand values, identity, and experiences—make sure your stakeholders understand the value AI-generated content delivers for them within this context.

Reinforce Human Oversight

Disclosure about the extent of human oversight in generating AI content is vital for maintaining trust with stakeholders who may be wary of GenAI. Companies should establish and publicly share a clear, overarching standard to ensure content is vetted by humans and adheres to brand guidelines and corporate standards - maintaining consistency and quality as more AI content is created.

Establish Disclosure Thresholds

Establish internal thresholds or tiers for "substantial use" of generative AI to create internal alignment on when AI noticeably alters the meaning of content, such as manipulating images or translating text, and use this alignment to establish guidelines for authoring and labeling AI content.

Drive Authorship Clarity

Decisions regarding whether and how to label AI-generated content can significantly influence stakeholder trust. Your approach to authorship labeling should emphasize clarity about the extent of AI involvement in content creation, using precise and consistent terminology – i.e., AI-generated, AI influenced, AI-informed, AI-augmented, AI-manipulated. The choice of label(s) should align with your brand strategy, ideally strengthening – rather than undermining– your brand identity.

Determine Disclosure Prominence

Know your audience. Determine how prominent your disclosures and labeling must be to meet your stakeholders' expectations and to avoid possible feelings of deception. By thoughtfully considering how and where to label AI-generated content, companies can reinforce their commitment to honesty and integrity, further strengthening brand reputation and fostering trust.

AI PRINCIPLES

Always be Optimizing

Measure the performance of AI-generated content to gauge trust among consumers of the content. Companies should encourage feedback, answer questions, and involve the community in discussions about AI to foster trust and gain valuable human insights that can improve generative AI content.

Address Fairness and Non-Discrimination

Disclosures should include information about measures taken to prevent bias in AI content, helping companies demonstrate a commitment to ethical practices and social responsibility. This transparency not only helps build trust with audiences but also mitigates risks of legal and regulatory repercussions.

Ensure Regulatory Compliance

With more regulation around the use of AI, it is important to understand the evolving compliance requirements related to the use of AI-generated content, which can vary according to each industry's regulations, guidelines and reporting requirements.

Strengthen Your Ethical Posture

Appropriate disclosure about the use of GenAI to create content can serve as a foundational example of transparency and ethical responsibility, setting a precedent for how AI should be utilized and disclosed in other areas such as customer service. By being transparent about AI's role in content creation, companies can more easily extend this trust to other AI applications across the business.

Drive Accountability

Establish clear lines of human accountability for ensuring AI content meets both internal and external performance, ethical and regulatory requirements. A cross-discipline accountability team including members from marketing, legal, compliance, data science, ethics, and customer relations, can more comprehensively evaluate the implications of AI-generated content.

ABOUT THE REPORT

In this insights report, Big Valley assessed conversations across companies, government organizations, and marketing and communications industry organizations to inform and inspire more discussion around when, why and how the use of Artificial Intelligence in business outputs is recommended or required for technology and technology-enabled brands to secure and maintain consumer/customer trust.

For the secondary analysis, Big Valley reviewed and integrated perspectives from different industries to understand and assess the different perspectives, conversations, and approaches to requirements – or lack of requirements - for transparency regarding AI use. The final product presented here was informed by the review of hundreds of peer-reviewed academic articles, conference papers, media coverage, government communications, and company websites published through July 2024, and includes 99 useful citations for readers.

In parallel, Big Valley's team of marketing, communications and artificial intelligence experts reflected on the observations reported here to propose a set of actionable AI principles for marketers and communicators to consider as they do - or don't - promote the use of artificial intelligence by their brands (e.g., in products, public relations or other business functions).

A note on limitations: *The authors recognize that a secondary analysis is only as good as the scope of sources included in the review. Big Valley explored a wide range of source categories by searching with keywords related to AI transparency, disclosure, and labeling, and using forward and backward searching where appropriate to consult other related literature. Inevitably, the report may have missed some perspectives and key sources. This report is also limited to information available as of July 2024, and the emerging landscape of AI means that approaches to transparency and disclosure may have changed by the time of reading.*

About Big Valley Marketing

Big Valley Marketing helps technology and technology-enabled clients Grow, Win and Lead by eliminating the wasted time and effort that results from marketing misalignment — too many channels to manage, highly fragmented roles and teams, compressed time cycles, disconnected data, and unclear strategies. Marketing impact is our singular focus.

For more information, visit https://bigvalley.co/ or contact us at hireus@bigvalley.co.

BIG VALLEY
MARKETING