Open Future Forum

OPEN FUTURE FORUM

CMO AI Leverage Report

NOVEMBER 2026

November 2026 | Edition 5 | Research report | Murray Newlands | Open Future Forum

Prepared 2 November 2026. Evidence cutoff: 31 October 2026, 23:59 America/Los_Angeles. Marketing impact is an updated cumulative estimate; adoption-stage evidence is carried forward. Search measurement guidance is a proposed method, not a measured OFF visibility benchmark.

About this publication: Open Future Forum publishes research for executive and investor audiences. Murray Newlands is its founder. This report separates OFF survey observations, external research, author analysis and illustrative scenarios. Vendor-authored guidance is identified as such and does not constitute an independent product assessment. The collection and supporting methods are available through the OFF research index: https://openfutureforum.com/research/

4,430 substantive body words | 5,792 total words
Tables, FAQs, methods and references are excluded from the body count.

In this report

01 The decision in brief
02 Executive summary
03 What November contributes
04 The evidence and its limits
05 1. Count work the business would accept
06 2. Make customer understanding answer a decision
07 3. Keep adoption categories separate from performance
08 4. Measure answer visibility with a reproducible denominator
09 5. Keep visibility, traffic and commercial contribution distinct
10 6. Design the team around judgment and accountable reuse
11 External evidence: useful comparisons, different questions
12 Practical application: one workflow, one customer decision
13 The CMO's next decisions
14 Reader questions
15 Citable facts
16 Methods and limitations
17 References
18 Related research and citation
19 Editable working tool

The decision in brief

Marketing leaders should judge AI by accepted work and commercial consequences, not the volume of drafts produced. In the updated OFF sample, capacity and content speed are tied at 141 of 314 respondents each. November adds a way to test whether that activity improves customer decisions, while measuring AI-search visibility separately from traffic and revenue. [MKT]

Executive summary

The updated marketing-impact instrument records three substantial areas of reported benefit. Doing the work of more people and creating content faster each appear in 141 of 314 eligible answers, or 45%. Knowing the customer better appears in 130, or 41%. Nothing measurable yet appears in 32, or 10%. These are multi-select perceptions from a selective registration sample, not audited improvements in campaign performance or company profitability. [MKT]

The tie between capacity and speed is useful because it prevents a convenient but unsupported headline that one now dominates marketing's use of AI. The underlying answers can overlap, and the additional cohort changes the composition of the cumulative sample. A respondent can report faster production without proving that the extra work is useful, differentiated or commercially effective. The CMO needs a measurement chain that preserves each of those questions.

This edition distinguishes workflow efficiency, output quality, customer response and financial contribution. It proposes a scorecard that records accepted outputs, review and rework, the intended customer decision and a credible commercial comparison. A team can recognize an early operating improvement without describing it as incremental revenue. Equally, a useful improvement in insight or service need not be dismissed because it has not yet produced a clean short-term attribution result.

November also develops a reproducible approach to answer-engine and generative-engine visibility. The method fixes the query set, engines, run conditions, dates and denominator. It separates brand inclusion, source citation, position and message accuracy. Repeated runs reveal variability; they do not create independent customer observations. Search Console's current generative-AI impression report provides one platform-specific view, while a controlled query study answers a different question. Neither alone establishes commercial impact.

A detailed measurement process can consume the time AI saves. Scale the review to the decision. A low-risk internal draft needs a modest quality check. A material campaign, customer claim or budget reallocation needs stronger evidence. The CMO should standardize a small set of decision-relevant measures, preserve the underlying definitions and stop collecting metrics that cannot change an action. The result should be clearer decisions about what to scale, improve or retire.

What November contributes

The cumulative marketing-impact base rises to 314 after adding an event cohort with 118 eligible respondents. Twenty-two respondents overlap with the previous 218-person set, so the updated total is not the simple sum of both cohorts. The new cohort's registrations span September and October; 88 current eligible answers have October registration timestamps. These are registration dates, not independently verified dates of answer changes. The cumulative comparison is therefore affected by sample composition and repeat participation. [MKT; change ledger]

The additional cohort reports faster content creation in 57 of 118 answers, capacity in 50 and customer understanding in 47. Those results provide useful new context but do not establish a change in the same marketers over time. The report publishes the cumulative distribution and explains the cohort boundary rather than converting a change in rounded percentages into a market trend.

The adoption-stage instrument remains at 230 respondents. It is carried forward and uses different wording from the general four-stage maturity instrument in the Transformation report. Its categories include building agentic products, piloting in functions, running agents across the business and exploration. They should not be forced into a universal maturity ladder or combined with the impact answers to infer a deployment-to-return conversion rate. [MKTSTAGE]

The first new analytical contribution is a quality-adjusted workflow scorecard that stops production volume from substituting for accepted marketing work. The second is a method for connecting customer-understanding claims to decisions and observable customer responses. The third is an explicit protocol for AI-answer visibility that includes repeatability, denominator control and separate commercial measurement.

Earlier OFF pieces already discuss marketing ROI, team structure and the shift from SEO toward AEO and GEO. This report does not repackage those themes as new findings. It develops the operational definitions and review rules needed to evaluate them, including circumstances in which faster output should not justify expansion. Its worked examples are illustrative and do not describe an OFF visibility study, campaign experiment or respondent organization.

The evidence and its limits

The impact questions come from registrations and applications associated with marketing-oriented events. An event audience label does not verify every participant's role, company size or responsibilities. Respondents are self-selected and report their own assessment of impact. Invitations without eligible answers are excluded, while other statuses do not imply attendance. The findings should not be presented as representative estimates for all CMOs or all marketing departments.

Each comparable question is deduplicated by normalized respondent identifier, retaining the latest eligible nonblank answer within the cutoff. Multiple answers are counted by any mention. Six respondents selected a positive impact and nothing measurable together. The analysis preserves those combinations rather than silently deciding which part of the answer is correct. The positive-impact union is 288 of 314; subtracting the

nothing-measurable share from 100% would produce the wrong result. [MKT_POSITIVE; MKT_CONTRADICTORY]

The report does not use that union as proof of ROI. It means that a respondent selected at least one of three positive descriptions. None of those options supplies a revenue amount, a cost baseline or an independent quality measure. A positive perception is relevant evidence of where to investigate, with a different status from a controlled campaign result.

The proposed scorecard and visibility protocol are management recommendations. Their illustrative numbers remain separate from the survey registry. External research supplies context and methods, not a pooled comparison sample. No share-of-answers result for OFF, a competitor or any other website is claimed in this edition.

1. Count work the business would accept

The updated finding that capacity and content speed are tied at 45% can support two different interpretations. AI may be helping teams complete more useful work, or it may be making the first production step faster while leaving approval, distribution and customer response largely unchanged. The survey does not choose between those explanations. A CMO should ask where in the workflow the reported benefit occurs before changing a budget or staffing plan. [MKT]

The unit of output should be something the organization accepts for its intended purpose. For a campaign, that may be an approved asset that meets the brief and can be distributed in the planned channel. For research, it may be an insight that changes a documented decision. For customer communication, it may be a response that resolves the need accurately and complies with the relevant policy. Generated drafts are useful activity records, but they are not interchangeable with those outputs.

The quality definition needs to precede the comparison. Specify the audience, claim accuracy, brand requirements, originality, accessibility and channel constraints that matter to the workflow. The organization need not score every minor item on a complex rubric. It should, however, know which defects require rework, which prevent publication and who has the authority to accept the result. Otherwise, a team can appear more productive simply by lowering the acceptance standard.

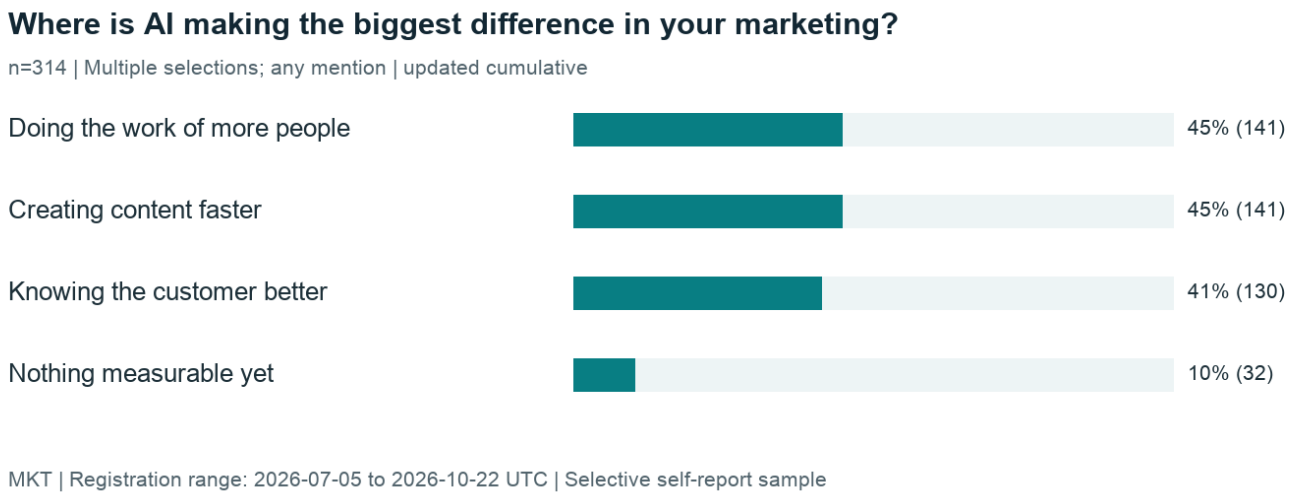
Measure review and rework alongside production time. Faster drafting can move effort to editors, subject experts or legal reviewers. That redistribution may still be worthwhile if it improves total output or quality, but it should be visible. The CMO should distinguish a bottleneck that has moved from one that has disappeared. A rise in unfinished drafts is not evidence that the organization can deliver more approved work.

Demand also matters. The team may not need twice as many assets; it may need more relevant assets, better distribution or a more coherent customer experience. The sponsor should identify the unmet need that additional production will serve. If there is no credible use for the capacity, the appropriate response may be to reduce low-value work and redeploy time to customer research rather than expand the content queue.

The result is a practical expansion test: does the workflow produce more accepted work, at an acceptable total cost, for a defined business need? This test preserves the value of speed while preventing it from

becoming a self-justifying metric. It also gives creative teams a clearer brief, because they know what the organization intends to do with the capacity they create.

Exhibit 1. Reported marketing impact



Exact selection	Count	Base	Any mention
Doing the work of more people	141	314	45%
Creating content faster	141	314	45%
Knowing the customer better	130	314	41%
Nothing measurable yet	32	314	10%

Exhibit note: MKT; updated cumulative marketing instrument. Multi-select, selective self-report sample. Six answers combine a positive option with nothing measurable; no contradictory selection is silently discarded.

2. Make customer understanding answer a decision

Knowing the customer better appears in 130 of 314 answers, a substantial signal that the reported benefit extends beyond production speed. Yet customer understanding is a broad phrase. It might mean faster synthesis of interviews, more accessible account information, improved segmentation or a better explanation of customer objections. The survey does not identify which activity produced the perception or whether the resulting understanding was correct. [MKT-03]

The CMO should connect each insight workflow to a decision. A team might use customer feedback to choose a product message, prioritize a service issue or revise an onboarding sequence. The record should state the decision before the analysis begins and identify what evidence would support changing it. This reduces the risk of generating an attractive summary that has no effect on the business.

Source quality remains central. An AI summary can faithfully reflect an unrepresentative set of customers. A convenient collection of support tickets may emphasize dissatisfied users, while sales notes may overrepresent active prospects. The team should record the source population, the period, important exclusions and the customer groups that are missing. Faster analysis does not correct the selection process that produced the source material.

The review should preserve disagreement. A summary that compresses several customer groups into one average preference can erase the distinction that matters for positioning or service design. Ask whether the analysis shows conflicting needs, uncertain interpretations and evidence that challenges the preferred campaign. The objective is a better decision, not a smoother narrative that makes the current plan easier to defend.

Validation can be proportionate. A low-cost message test may be enough to assess an initial hypothesis. A major repositioning may require further customer research, a controlled experiment or a staged rollout. The CMO should decide in advance which result would change the plan and which would be too ambiguous to interpret. A statistically elaborate dashboard cannot rescue a test whose decision rule is undefined.

This creates a distinct measurement path from content production. The output is not simply the number of insights generated. It is the quality of the decision informed, the evidence used and the customer response observed afterward. Some decisions will take time to evaluate; that uncertainty should remain visible. The organization can credit faster synthesis as an operating benefit while waiting for evidence that the resulting choice improved a commercial outcome.

3. Keep adoption categories separate from performance

The carried-forward stage question records 82 of 230 respondents building agentic AI products, 59 piloting agents in one or two functions, 46 running agents in production across the business and 43 still exploring. These categories reflect the instrument as asked. Building a product is not necessarily a stage between piloting and enterprise-wide production; a company can build a product while using a different internal operating model. [MKTSTAGE]

The CMO should therefore avoid ranking all respondents on one maturity scale. A small team using a narrow production workflow effectively may have a more defensible operating result than a large organization building an ambitious product with unresolved demand. Stage labels describe activity and scope. They do not establish quality, economics, customer acceptance or control effectiveness.

The same caution applies inside the organization. A marketing department can have several workflows at different stages simultaneously. Content drafting may be routine, customer segmentation experimental and autonomous campaign changes prohibited. A single department-level maturity score can conceal these differences and create pressure to move every workflow toward the same degree of autonomy. The business need should determine the destination.

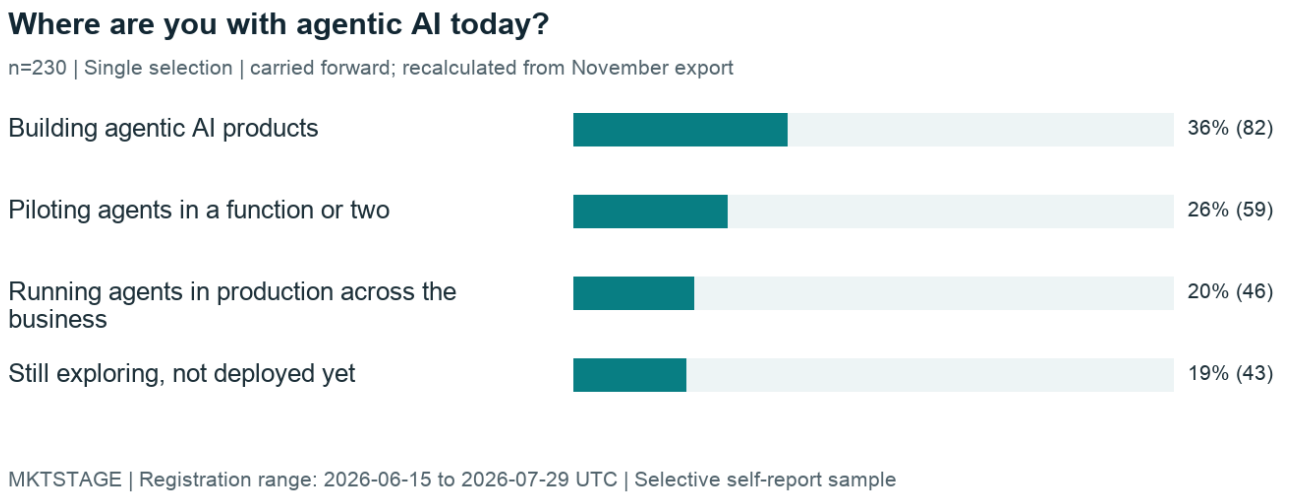
A more useful portfolio view records each workflow's current purpose, evidence and next decision. An exploratory workflow needs a learning question. A pilot needs a defined comparison and an owner capable of acting on the result. A production workflow needs quality, cost and service evidence. A consequential

autonomous workflow needs an appropriate authorization and intervention design. These are recommended review requirements, not recoded survey categories.

The team should also define retirement. Some experiments will show that the original need was weak, the data inadequate or the review burden too high. Retiring them can release capacity for better work. If every pilot must become a production success to be considered valuable, teams have an incentive to preserve weak projects and redefine their goals after the fact.

The leadership implication is to reward evidence that changes a decision. A useful pilot may justify expansion, a narrower use case or a stop. The CMO's review should ask what was learned and what action follows, rather than count how many initiatives carry an advanced label. This makes the adoption portfolio more honest and reduces the temptation to use the survey's stage distribution as a competitive benchmark it was not designed to provide.

Exhibit 2. Activity categories in the carried-forward stage instrument



Exact selection	Count	Base	Share
Building agentic AI products	82	230	36%
Piloting agents in a function or two	59	230	26%
Running agents in production across the business	46	230	20%
Still exploring, not deployed yet	43	230	19%

Exhibit note: MKTSTAGE; single-selection, carried forward. Rounding may prevent a total of exactly 100%. This instrument is distinct from the Transformation report's four-stage maturity question.

4. Measure answer visibility with a reproducible denominator

Answer-engine optimization and generative-engine optimization concern how information appears in generated answers; SEO remains relevant to discoverability in search. The terminology should not determine the measurement design. The CMO needs to specify the customer question, the system producing the answer and the outcome being observed. A claim that a brand is winning AI search is too broad to evaluate without those definitions.

Begin with a fixed query set derived from real customer decisions, not a list chosen after inspecting favorable results. Group the queries by intent, such as category learning, comparison, purchase requirements and implementation. Preserve branded and unbranded questions separately. Record the wording and the reason each query belongs in the set. If the set changes, publish a new version and maintain an unchanged subset for comparison where useful.

Record the engines and interfaces tested, the date and time, locale, language, account state and other visible conditions that may affect results. Use only permitted access methods. A manual study can be valuable if its procedure is consistent and its limits are stated. The team should not describe results from one logged-in account or one interface as the experience of all potential buyers.

Separate at least four outcomes. Brand inclusion asks whether the answer mentions the brand. Citation asks whether it links to an owned source, and should distinguish direct links from third-party mentions. Position records where a mention appears under a defined rule. Message accuracy asks whether the answer describes the product or company correctly. A brand can score well on inclusion while being mischaracterized or cited through an outdated source.

Repeated runs should reveal variation rather than inflate the apparent sample. If the team runs the same query several times, those observations share a query and may share conditions. Report both the query-level result and the run-level variation. For a headline inclusion rate, the denominator might be eligible query-engine combinations under a fixed rule; for variability, the denominator is repeated valid runs. Explain failed, blocked or unanswered runs separately rather than silently treating all of them as absent mentions.

This protocol produces a bounded visibility study. It does not measure every answer on the internet or establish a universal share of customer attention. Its value lies in repeatability: the team can identify which customer questions are poorly served, inspect the source material and test whether a substantive content improvement changes the observed result under comparable conditions.

5. Keep visibility, traffic and commercial contribution distinct

Google's current documentation describes a generative-AI performance report covering impressions from AI Overviews and AI Mode, with page, country, date and device views. [E-GOOGLE-REPORT] That is useful platform evidence, but its unit differs from a controlled query study. The former reflects recorded exposure on a particular platform; the latter observes answers under selected test conditions. A company should preserve both definitions instead of combining them into one visibility score.

The commercial path is longer. An answer may influence a buyer who does not click immediately. A click may come from a user with no purchase intent. A lead may have encountered several sources before conversion. The CMO should distinguish observed referral traffic, self-reported discovery and modeled attribution, with each method's limitations visible. Attribution can inform a decision without pretending to identify every causal contribution.

A controlled visibility study should therefore include a separate business hypothesis. For example, improving a technical comparison page may be intended to help qualified buyers understand a specific requirement. The team can monitor accurate citation for those questions, relevant visits and downstream engagement. It should not assume that a higher citation count necessarily causes more revenue, especially when traffic, campaign activity or demand changes at the same time.

Content changes need their own record. If the team revises a page, changes its canonical destination, alters access controls or publishes a new version, the study should retain the date and scope of that change. Otherwise, a later difference in visibility cannot be related even descriptively to the intervention. The record should also note material platform or measurement changes that make a comparison less reliable.

The CMO should be cautious with one-number dashboards. Combining inclusion, citation, position and sentiment with arbitrary weights can conceal the problem the team needs to solve. A page that is frequently cited for an inaccurate claim requires a different response from a page that is accurate but rarely discovered. A small set of transparent measures makes those decisions clearer.

The practical reporting sequence is exposure, message, engagement and business consequence. Each layer should have its own denominator, owner and confidence limit. The team can then explain which part of the customer journey improved and which remains uncertain. This preserves the strategic relevance of AI-mediated discovery without assigning a financial return to every observed mention or encouraging content production solely to raise a visibility metric.

6. Design the team around judgment and accountable reuse

When AI accelerates drafting and synthesis, the CMO still needs people who can judge the brief, verify claims, understand customers and decide what deserves publication. The OFF impact question records perceived capacity and speed, but it does not measure how organizations redesigned jobs or whether headcount fell. A staffing conclusion requires the company's own workload, quality and service evidence. [MKT]

The first organizational question is where review capacity sits. If subject experts become the new bottleneck, adding more draft production may make the queue worse. The team should identify which reviews require specialist judgment and which can use a standard checklist. This can improve throughput without weakening accountability, because scarce expertise is reserved for the decisions that need it.

The second question is how knowledge is maintained. A reusable brief, source pack or approved claim can reduce repeated work, but only if someone owns its accuracy and expiry. An outdated approved statement can be reproduced efficiently across many channels. The CMO should therefore treat reusable material as

maintained content with an owner, source and review trigger, rather than as a permanent library of safe phrases.

The third question is originality. Faster production makes it easier to publish several pieces that answer the same reader question with slightly different wording. A content review should compare purpose, evidence, argument and practical contribution, as well as exact text. Necessary reuse of an accurate fact is not the same as repeating a whole article. The decision should be whether the new piece gives its intended reader a distinct reason to spend time with it.

Teams also need an incentive to report failures. If every AI initiative must show a positive result, review effort and quality problems may be omitted from performance claims. The CMO should recognize useful negative findings, such as a workflow that needs more specialist review than expected or a campaign whose extra output did not improve response. Those findings can prevent a larger misallocation of effort.

The staffing discussion should follow that evidence. Some teams may redeploy time toward research, experimentation or customer contact; others may avoid a planned hire or reduce purchased production. Those are different decisions with different consequences. The CMO should explain the intended use of capacity and verify whether it occurs. A clearer operating purpose is more useful than a generalized promise that AI will make the marketing organization smaller or more strategic.

External evidence: useful comparisons, different questions

The CMO Survey's March 2026 analysis draws on 308 marketing leaders at US for-profit companies, 97% at vice-president level or above, and describes adoption outpacing organizational readiness. [E-CMOSURVEY] OFF's sample and questions differ, so the results cannot be pooled. The comparison supports examining the capability needed to use tools effectively, while leaving the company's own training, integration and workload decisions to direct evidence.

Harvard Business School's account of the GenAI Wall Effect describes a 78-person study in which task proximity and expertise affected results. [E-HBS-WALL] The useful implication for marketing is a boundary on substitution: access to a tool does not establish that a person can reliably complete unfamiliar specialist work. This report uses the accessible institutional account for that qualitative finding. It does not repeat precise time or quality estimates from the paper that could not be independently inspected here.

OpenAI's July 2026 work-use research analyzes more than 800,000 US work-related messages and distinguishes activity inside and outside occupational tasks. [E-OPENAI-WORK] It is evidence about use, not a test of campaign effectiveness. For a CMO, it raises a practical question about review: when a person uses AI to work beyond their usual expertise, who can assess whether the result meets the relevant standard? The answer depends on the task and consequence, rather than on tool access alone.

Google's documentation adds a measurement capability specific to its search environment. The generative-AI report's impressions and aggregation rules should be read as defined by the platform. [E-GOOGLE-REPORT] They cannot be substituted for the denominator in an OFF-style respondent survey or a fixed-query visibility study. The CMO's dashboard should label the source and unit so that a change in one measure is not mistaken for a change in another.

Together, these sources complicate a simple production story. More use can coexist with uneven readiness; task assistance can coexist with an expertise boundary; and greater recorded exposure can coexist with uncertain commercial contribution. None of those observations means the investment lacks value. They identify the missing evidence that management must collect before deciding what to scale.

The report's proposed measurement tree is an inference from those distinctions and the OFF findings. It is not a validated universal model. Its value should be judged by whether it helps a team identify a concrete problem, choose a useful comparison and make a better decision. If a measure never changes what the team does, it should be removed or redesigned rather than retained because it makes the dashboard look comprehensive.

Source	Comparison with OFF	Decision implication	Boundary
The CMO Survey, March 2026	Different marketing sample and readiness questions	Review operating capability alongside adoption	Self-reported survey; not pooled with OFF
HBS GenAI Wall Effect account, 2025	Adds task-expertise context	Assign qualified review for unfamiliar work	Institutional summary used; exact paper estimates excluded
OpenAI work-use research, July 2026	Describes work-related usage	Distinguish task expansion from verified competence	Messages are not measured business outcomes
Google Search Console documentation, current	Defines platform-specific exposure	Label units and aggregation in reporting	Does not establish revenue attribution

Practical application: one workflow, one customer decision

Choose a workflow with a real production need and a measurable acceptance standard. The marketing operations owner should document the baseline, while the creative or subject owner defines quality. Finance should review any monetary conversion, and the channel owner should identify the customer behavior the work is intended to influence. This division prevents one team from certifying every part of its own business case.

Working tool: marketing AI measurement tree

Layer	Question	Example evidence	Decision owner
Workflow	Does accepted work require less total effort?	Production, review and rework time	Marketing operations
Quality	Is the output accurate, useful and on brief?	Predefined acceptance review	Creative / subject owner
Customer	Does it help the intended customer decision?	Experiment, research or service response	Customer / channel owner
Commercial	Is there incremental contribution or a real cost change?	Finance-reviewed comparison	CMO and finance
Discovery	Is the information visible and correctly represented?	Fixed-query study and platform records	Search / content owner

An illustrative content team produces 120 drafts in a month, of which 90 are accepted. After changing its workflow, it produces 180 drafts and 108 are accepted. Draft volume rises 50%, while accepted output rises 20%. The acceptance rate falls from 75% to 60%. These invented figures show why the choice of denominator changes the management story; they are not survey results or a claim about an OFF team.

The CMO should next ask what happened to review effort, cost and usefulness. If the team spent substantially more time checking the additional drafts, the gain in accepted output may carry a higher unit cost. If the 18 additional accepted items address a real customer need, the investment may still be attractive. If they merely fill a publication schedule, more output may not be the right objective. The scenario cannot be resolved by draft count alone.

Exhibit 3. Illustrative production comparison, not OFF observations

Measure	Baseline month	Changed workflow	Interpretation
Drafts	120	180	Activity rises 50%
Accepted outputs	90	108	Accepted output rises 20%
Acceptance rate	75%	60%	Quality / fit needs investigation
Review hours	Must be measured	Must be measured	No cost conclusion without this input
Customer response	Must be measured	Must be measured	No revenue conclusion from output alone

For discovery measurement, begin with an illustrative protocol of 40 fixed customer queries across three chosen engines, repeated three times under recorded conditions. That creates 120 query-engine

combinations and up to 360 run observations. Report valid and failed runs separately. Do not call the 360 observations 360 independent customers, and do not use the protocol's numbers as a recommendation that every company needs exactly this study size.

Before running the study, define inclusion and citation rules. Decide how to treat an ambiguous brand name, a competitor comparison, multiple links to the same domain and a citation to an outdated page. Define position only where the answer format makes it meaningful. Have a second reviewer inspect ambiguous cases or a documented sample so that coding decisions remain consistent. Preserve the raw public answer evidence under appropriate access and retention practices.

Working tool: visibility study record

Field	Required entry
Study version	Query set, rationale, branded/unbranded split and frozen date
Run conditions	Engine/interface, date/time, locale, language and account state
Denominator	Eligible query-engine combinations; valid repeats; failures
Inclusion	Brand mentioned under a predefined matching rule
Citation	Owned source linked; destination; third-party mention separate
Position / accuracy	Defined location measure; factual correctness and ambiguity
Change record	Content revisions, platform changes and comparability limits
Business follow-up	Relevant traffic, customer response and commercial evidence kept separate

The review meeting should decide whether to expand production, repair quality, improve source content or change the underlying customer proposition. A visibility problem may reflect weak information, an unclear category position or a study that asks the wrong questions. Producing more near-duplicate pages is not a reliable answer. The content owner should identify the missing reader value before commissioning a new piece.

Close each review with a dated decision and a condition for reconsideration. If the team scales a workflow, specify the evidence that would trigger a pause, such as a material quality failure or an unsustainable review queue. If it changes content, preserve the previous measurement definition. This makes the next review a comparison of evidence rather than a fresh narrative assembled from whichever numbers improved.

The CMO's next decisions

First, choose the accepted-output definition for the workflows receiving additional investment. Require production, review and rework evidence before describing faster drafting as a capacity gain for the whole team. The operating owner should identify where the released time will go and whether there is demand for the additional work.

Second, connect customer-understanding initiatives to a specific decision. Preserve the source population and competing interpretations, then choose a validation method proportionate to the consequence. Credit faster synthesis separately from the eventual commercial result so that early progress does not become an unsupported revenue claim.

Third, establish one reproducible discovery study and keep its measures distinct from platform impressions, referral traffic and financial attribution. The query set and definitions should remain stable enough to support comparison, with changes logged rather than hidden. A useful study explains what the team should improve, not merely whether a composite score rose.

Finally, review the content portfolio for repeated reader purpose. Shared facts can support several audiences, but each new publication should contribute a distinct analysis, decision or usable method. The CMO should apply the same standard to AI-assisted output as to any commissioned work: a clear audience, a justified claim and enough original value to deserve publication.

Reader questions

Is capacity now the leading marketing benefit?

It is tied with content speed in the cumulative sample: each is selected by 141 of 314 respondents, or 45%. [MKT]

Does 92% positive impact mean 92% have positive ROI?

No. The 288/314 union means respondents selected at least one positive impact option. The question does not measure audited cost, revenue or return. [MKT_POSITIVE]

Can the adoption categories be converted into a maturity score?

This report does not do so. Building products and running agents across a business are different activity descriptions; the question is not the same as the four-stage general maturity instrument. [MKTSTAGE]

Is share of answers the same as share of search traffic?

No. A defined query study measures answers under its selected conditions. Platform impressions and site traffic have different units and populations.

Do repeated runs make a visibility study representative?

They reveal variability for the chosen queries and conditions. They do not turn a purposive query set into a random sample of customers.

Does Google provide a separate generative-AI view?

The documentation consulted describes a generative-AI impression report covering AI Overviews and AI Mode. Availability depends in part on sufficient impressions. It should be interpreted using Google's definitions. [E-GOOGLE-REPORT]

Should a team publish more because drafts are cheaper?

Only when the additional accepted work serves a real reader or customer need and its total review and delivery cost is justified. This is a management recommendation, not a survey finding.

Citable facts

- Capacity and content speed each appear in 141/314 answers, 45%, updated cumulative multi-select marketing sample. [MKT-01; MKT-02]

- Customer understanding appears in 130/314 answers, 41%. [MKT-03]

- The positive-impact union is 288/314, 92%; six answers combine a positive option and nothing measurable. [MKT_POSITIVE; MKT_CONTRADICTORY]

Methods and limitations

Calculation-v1 uses the authorized 1 November export set, with byte-identical duplicates and invited-only records excluded. Answers are deduplicated per question by normalized identifier and latest eligible nonblank registration record within the cutoff. The registration timestamp does not establish when an answer was edited. Multi-select questions use any mention and preserve contradictory combinations.

MKT is updated cumulative, n=314. The added event contributes n=118, with 22 overlapping respondents relative to the previous n=218. The added event spans September and October registrations. MKTSTAGE is carried forward, n=230, single selection. No joint stage-to-impact transition or causal effect is inferred. The source and change ledgers preserve exact wording, periods and counts.

The sample is selective and self-reported. No national representativeness, probability-sample margin of error, audited ROI or staffing effect is claimed. The visibility protocol and numerical examples are recommendations and illustrations; no live visibility benchmark was conducted for this report. The HBS institutional summary was read, while detailed numerical estimates from the underlying paper are excluded because the paper was not independently inspected.

References

[OFF-DATA] Open Future Forum. Executive AI Statistics, November 2026; accompanying aggregate data and change ledger.

[E-CMOSURVEY] Moorman, C. Marketing Contracts Under Economic Pressure Despite Growing Value and AI Gains. The CMO Survey, 31 March 2026, 35th edition analysis. https://cmosurvey.org/marketing-contracts-under-economic-pressure-despite-growing-value-and-ai-gains/

[E-HBS-WALL] Harvard Business School AI Institute. Why AI Helps Until It Doesn't: Inside the GenAI Wall Effect. 18 September 2025, institutional research account. https://aiinstitute.hbs.edu/why-ai-helps-until-it-doesnt-inside-the-genai-wall-effect/

[E-OPENAI-WORK] OpenAI. How AI is expanding what people do at work. 27 July 2026. https://openai.com/index/how-ai-is-expanding-what-people-do-at-work/

[E-GOOGLE-REPORT] Google Search Console Help. Generative AI performance report (Search), current documentation consulted 2 November 2026. https://support.google.com/webmasters/answer/16984139

[OFF-PRODUCTIVITY] Newlands, M. AI Is Making Employees Faster. Why Isn't Every Company Seeing the Returns? Open Future Forum, 31 October 2026. https://openfutureforum.com/blog/ai-productivity-returns-stanford-karunakaran

Related research and citation

The November CFO report explains benefit qualification, the Transformation report examines workflow change, and the CEO report addresses the allocation of management attention and released capacity. These are related analyses with different decision purposes.

Suggested citation: Newlands, Murray. CMO AI Leverage Report, November 2026. Open Future Forum. Evidence cutoff 31 October 2026.

Open Future Forum research and methods: https://openfutureforum.com/research/

Editable working tool

Accepted-output and visibility measurement: download the editable CSV worksheet. The worksheet implements the report's recommendation; it is not a validated score, mandatory standard or completed assessment.