

AIは賢くなった。しかし、人間の判断は、なぜ楽にならないのか

——「仕様」と、AIが使える「評価基準」を先に作って、渡す

CycleGen 論考シリーズ Vol.01

2026-10-04・佐藤 純也／株式会社らしく

はじめに

先に結論を書きます。

AIが賢くなっても判断が楽にならない原因は、**評価基準が自分の頭の中にしかない**ことです。そのため、原案が出るたびに自分で判定することになり、AIから原案が次々と出てくるほど、判定待ちの原案が自分のところに溜まります。この論考では、評価基準を先に書いて渡すという対処を扱います。

仕様と評価基準を渡さないままAIに作業させると、良くてやり直し、悪ければ顧客や取引先からの信用を（ひっそりと）失います。これを防ぐために、仕事を任せる前に何を決め、何を書いておくか。それが、この論考で考えたいことです。

用意するものは二つあります。何を作るかを示す「仕様」と、作られた原案をどう判定するかを示す「評価基準」です。ここでいう評価基準は、人事の評価表や点数の配分表ではありません。あなた抜きで、他の人やAIが適用して、同じ合否を出せる形で書かれた条件を指します。「これなら採用できる」という感覚を、判定する人が替わっても使える言葉にする、ということです。なお、文中の「採用する」は原案を受け入れること、「差し戻す」は作り直しを求めることを指します。

AIに渡す前に用意する二つ 「仕様」と「評価基準」

仕様 + 評価基準 = AIが作り合否を判定

作るものを示す

採用の条件を示す

その先の仕上げは人

—— 評価基準とは、「あなた抜きで、他の人やAIが適用して、同じ合否を出せる形」で書かれたものを指す。

図1：AIに渡す前に用意する二つ — 仕様と評価基準

答えが一つに決まらない仕事で、自分でAIと働き、自分で書く人に向けて書きました。職位を問わず、次のどちらかに心当たりのある方に読んでいただきたいと考えています。チームで使う場合は、文中の「私」を依頼者と読み替えてください。

- AIに企画書や提案書を作ってもらったが、納得できず、結局自分で作り直した
- AIが賢くなるにつれて原案の量と密度が増え、良し悪しを判定する自分の仕事が重くなっている

同じ種類の判定を何度もする仕事ほど、評価基準を書いておく意味があります。一方、仕様がすでに決まり、1回かぎり、評価基準を作り直す必要もない仕事なら、この論考が扱う工程は済んでいるか、そもそも要りません。

この論考は完成の報告ではありません。まだ確かめていないことは、末尾にまとめました。

この論考で考えることは三つです。いまの仕事がどの局面にあるか。どこまでならAIに任せられるか（「任せる」の意味は4章で説明します）。任せる前に何を書いておくか。自分の仕事を見分けるための問いと、実際に書いた評価基準も載せました。具体的な形から見たい方は、末尾の「書いたものの見本」からお読みください。AIに渡して使いたい方は、末尾の「読者のAIへ」からお読みください。

1章 — 判定できるのが自分だけ、という状態を変えたい

私はこれまで、提案書やプレゼンテーション資料を自分で作るとともに、上司や同僚、取引先やパートナー企業のスタッフがそれらを作る様子も数多く見てきました。よく目にしたのは、紙とペンで勝ち筋を考える前に、いきなりパワーポイントを開いて作り始める姿です。若いころの私もそうでした。段取り（勝ち筋を見出す仕事）と作業（その勝ち筋を具現化する仕事）を分けたほうが品質と生産性は上がるのに、そうっていないことが多いようです。

「段取り八分、仕事二分」と言われるように、段取りの大切さは、昔から言われてきました。それでも、作る時間は予定に入っても、何を作るかを考える時間は入りにくいのが実情です。作業は目の前に何かができるので、段取りより進んでいるように見えます。

AIと働くようになって、同じ構造が現れました。AIは、それらしい提案書や計画書を出します。しかし、**それらしいのに、中身が意図したものになりません**。だから人が書き直します。書き直されないまま、社外に出てしまうこともあります。

私を変えたいのは「**AIの原案の良し悪しを判定できる人が自分だけ**」という状態です。原案がすぐに出てくるようになって、採用か差し戻しかを決める仕事は私が抱えたままです。原案の量と密度が増えるほど、読む時間も、判定待ちの原案も増えていきます。**私の中にある評価基準を書き出しておけば、判定するのは私でなくてもよくなります**。

AIの生成能力が上がるほど、人間の「判断」がボトルネックになる

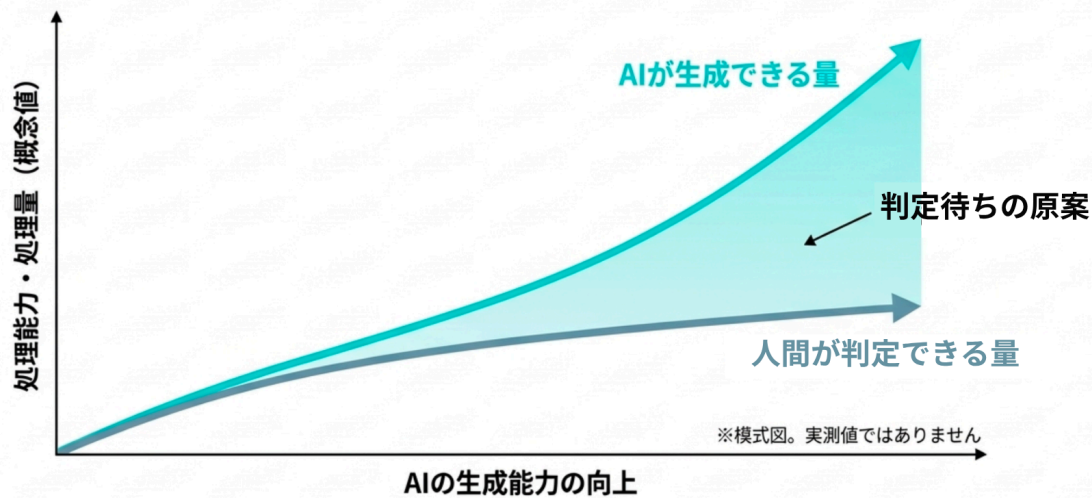


図2：AIの生成能力が上がるほど、人間の「判断」がボトルネックになる

私自身の例を書きます。社外に出す長い文書をAIに書かせたところ、強調したかった一文が、別の文に置き換わって返ってきました。ビジネスの場ではあまり使わない語や、読む人によって意味の変わる語も混じっていました。前に書いたことを、言い方を変えて繰り返す箇所もありました。読めば違いが分かるので、その場で私が直して提出しました。

私の中には「この一文は変えない」「この言い方なら差し戻す」という評価基準がありました。それなのに、依頼するときには何を書くかという指示だけを渡し、評価基準は書いていなかったのです。原案を読むまで評価基準を書き出さなければ、書き直す時間は私が負担することになります。

評価基準が「頭の中」にある限り、結局、人が書き直す



私たちは「何を書くか」という指示（仕様）は渡すが、「この言い方なら差し戻す」「この一文は変えない」といった評価基準を依頼時に渡していない。

失うのは、費やした時間だけではない。
意図と違うものを社外に出せば、顧客や取引先からの信用も失う。

図3：評価基準が「頭の中」にある限り、結局、人が書き直す

提案書を受け取る側に回ると、同じことが別の様子として見えます。社外から届く提案書や計画書に、主張のずれ、意味の揺れる語、説明の重複が並んでいることは、これまで何度もありました。そのとき下がるのは、内容への評価ではありません。提出した会社への信用も下がります。しかも、受け取った側がそれを相手に伝えることは、まずありません。次に仕事を頼むかどうかを考えると、その記憶は消えません。

私も、この論考のように自分の考えを伝える文章や提案書を書きます。調子の良い日はすらすら書けますが、そうでない日は筆が進まず、仕上がりに満足できません。評価基準が頭の中にしかなければ、判定までその日の気分や調子に左右されます。評価基準を書き出しておけば、調子の悪い日にも同じ物差しで判定できます。そうしてAIと一緒に、ばらつきのない文章を作りたいと考えています。

もっとも、頭の中の評価基準を書き出そうとすると、言葉にできる部分とできない部分があります。それでも、言葉にできる部分は、書き出すたびに確かに広がっています。そのなかで、はっきりしてきたことがあります。システム開発では、プログラムは媒体にすぎず、仕様どおりに動くかを確かめれば採用できます。文章では、言葉そのものが成果物です。良い文章の条件は、書いて、作って、確かめる往復のなかで見つけていくものだと考えています。そこでまず、何を作るかを決めること、評価基準を書くこと、その評価基準に沿って作ることを分け、自分がどこで止まっているのかを確かめています。

2章 — 成果物をつくる仕事は、三つの局面（探索局面・定義局面・実行局面）を行き来する

提案書を作るときも、広告出稿計画を立てるときも、基幹業務システムを開発するときも、仕事には三つの状態があります。何を作るかがまだ決まっていない状態、作るものは決まり、仕様と評価基準を作っている状態、その両方がそろった状態です。仕事の種類や規模が違って、いま何が決まっているかで分けられる点は共通です。

この論考では、状態の違いで分けた三つを「局面」と呼びます。局面で見れば、仕事の名前が同じでも、いま何をしているかを見分けられます。1章で書いた段取りは、作るものを探し、その仕様と評価基準を作るところまでを含みます。決まった仕様に沿って作るのが、そのあとの作業です。

三つの局面

三つの局面には、次の名前をつけました。いずれもこの論考で使う造語です。「探索局面」は、作るものがまだ決まっていない局面です。「定義局面」は、作るものは決まったが、評価基準がまだ決まっていない局面です。「実行局面」は、仕様と評価基準の両方がそろった局面です。

成果物を作る仕事は、三つの局面を行き来する

	(1)探索局面	(2)定義局面	(3)実行局面
状態の定義	何を作るかが決まっていない (作るのが1行で言えない)	作るのは決まっているが、 評価基準が未定	仕様と評価基準が そろっている
この局面で 作るもの	文脈 (理解・問い・候補・ 見送った理由)	仕様書 および 評価基準	成果物
良し悪しを 決めるもの	 人間 (AIは対話の相手)	 人間	 人とAIで作った 評価基準 (AIが合否を判定)

図4：成果物を作る仕事は、三つの局面を行き来する

図の「状態の定義」を順に見ていきます。探索局面では、作るものをまだ1行で言えません。この局面で記録するのは、候補や問い、見送った理由などの検討の経緯です。定義局面では、作るものは決まっていますが、何をもって合格とするかが決まっています。ここで仕様書と評価基準を作ります。この二つの局面では、良し悪しを人が判断します。実行局面まで来ると、AIが原案を、人とAIで作った評価基準と照らし合わせられます。原案ができるたびに、人が判定し直す必要はなくなります。ただし、AIが判定できるのは、評価基準を満たしているかどうかまでです。満たしていなければ差し戻します。そこから先を良くして仕上げるのは、いまま私です。

終わり方も、局面によって違います。探索局面には終わりの条件がないので、時間か予算で区切ります。定義局面は、仕様書と評価基準が確定し、承認されたところで終わります。実行局面では、仕様と評価基準を満たすまで、作って確かめることを繰り返します。

探索局面に終わりの条件がないことは、私だけの観察ではありません。都市計画の研究者 Rittel と Webber は1973年の論文で、正解が一つに決まらない「厄介な問題」の特徴として、どこで終えてよいかを決めるルールを持たないことを挙げています。私はこの考え方を自分の仕事に適用し、評価基準を作る局面と、評価基準に沿って進める局面を見分けようとしています。

例えーレストランの立ち上げ

レストランを立ち上げる場面に置き換えると、違いを思い浮かべやすくなります。この例えで説明するのは、三つの局面の違いだけです。

三つの局面の具体例：レストランの立ち上げ

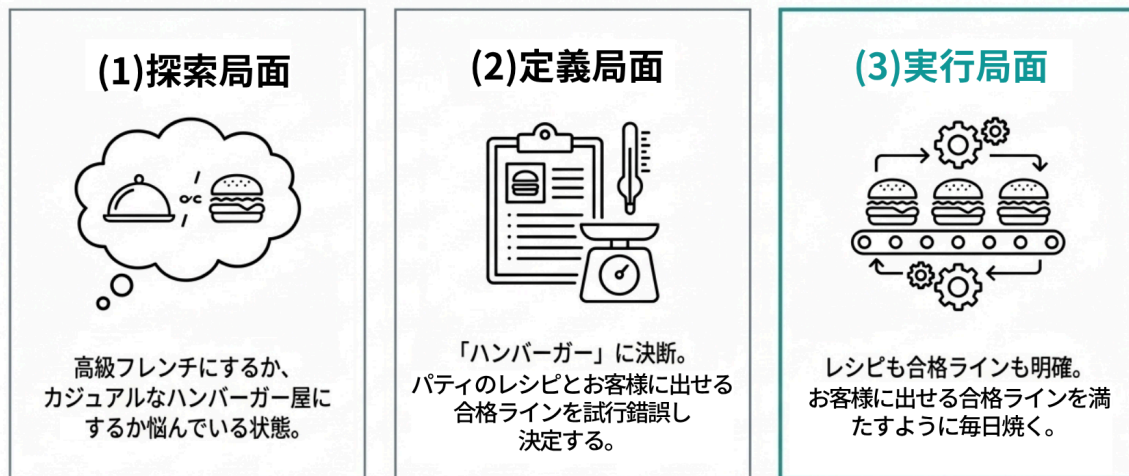


図5：三つの局面の具体例－レストランの立ち上げ

店の種類を考えている段階（探索局面）で、ハンバーガーのレシピを書き始めても、フレンチにすると決めれば、そのレシピは使えません。ハンバーガーに決まったら、作り方に加えて、どの焼き加減ならお客さんに出せるかを決めます（定義局面）。そこまで決まれば、毎回その焼き加減を満たすように焼く仕事になります（実行局面）。

作り方を伝えても、焼き加減を店主しか判定できなければ、焼き上がるたびに店主を呼ぶことになります。レシピだけがある状態と、出してよい焼き加減まで決まった状態を分けるのは、そのためです。合格ラインも伝わっていれば、焼く人が自分で確かめられます。

小さなレストランなら、店を考える役、レシピと合格ラインを決める役、毎日焼く役を、店主が一人で担います。毎日焼くうちに、決めた焼き加減では出せないと分ければ、試作の場（テストキッチン）、つまり定義局面に戻って合格ラインを見直します。

同じ人が、1日のうちに行き来する

私が提案書を書く日にも、同じ行き来があります。朝、提案先の状況を読み直しながら何を提案するかを考えているあいだは、探索局面にいます。作るものを、まだ1行で言えないからです。

昼までに「この課題に対して、この方式を提案する」と決まれば、午後は章立てと、各章が何を満たしていれば採用できるかを書き出します。ここが定義局面です。夕方、その評価基準を渡してAIに章ごとの原案を作らせ、評価基準と照らし合わせます。ここが実行局面です。

同じ提案書に向かっている、いま何を決めているかによって、いる局面は変わります。

探索局面で、私が実際にやっていること

私の場合、仕事の中心が企画や提案を考えることなので、日々いちばん長くいるのは探索局面です。どの局面に長くいるかは、担当する仕事によって違うはずですが、図では短い列ですが、何を作るかが決まるまでには、候補を出したり見送ったりする時間がかかります。

この局面で私がやっているのは、AIとの対話で考えを整理することです。対話から見えてきたことと、そのとき自分が下した判断を、AIがテキストファイルに記録しています。決めたことだけを記録すると、あとで読み返しても、なぜ別の案にしなかったのかが分かりません。そこで、見送った案とその理由も一緒に記録します。図の探索局面の列にある「文脈」は、こうして記録した検討の経緯です。

見送った案の理由まで記録できるのは、書いているのがAIだからです。人が自分で記録していたころは、決めたことを書き留めるだけで手一杯で、採用しなかった理由までは手が回りませんでした。いまは対話の中に判断の理由がすでに出ているので、AIがそれをその場で記録できます。記録の仕方と使い方は、[Vol.02](#)で詳しく扱います。

あとの局面から、前の局面へ戻る

実行局面まで進んでも、そのまま終わるとは限りません。作って確かめたことで、前の局面で決めた内容を見直すことがあります。

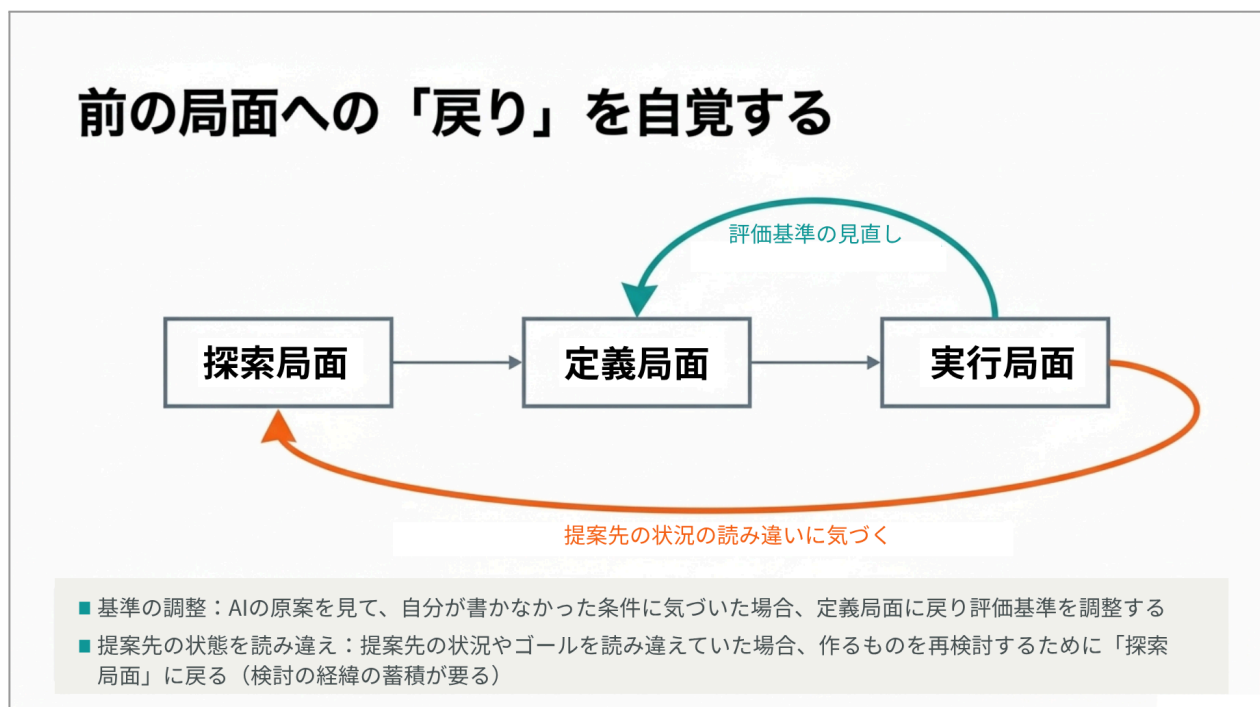


図6：前の局面への「戻り」を自覚する

夕方、AIの原案を読んで、評価基準は満たしているのに、これでは採用できないと感ずることがあります。足りなかった条件を言葉にできれば、評価基準に加えて、もう一度作らせます。これが、実行局面から定義局面への戻りです。原案を見たことで、評価基準に書いていなかった条件に気づいたのです。

もっと前の局面に戻ることもあります。評価基準に何を加えればよいか分からず、提案先が何を良しとするかを考え直すことがあります。すると、「この課題に対してこの方式」という提案の前提から、提案先の状

況を読み違えていたと分かります。作るものを決め直すので、探索局面へ戻ります。1行で言えていたはずの成果物が、もう言えなくなっています。

社外に出す前に気づけば、手元でやり直せます。どの案をなぜ選んだかを記録してあれば、戻った先で、その判断から見直せます。

三つの局面に上下はありません。作ったからこそ、決めていなかったことが見える場合もあります。私が確かめたいのは、先へ進んだかどうかだけではありません。いま何を決めるために、どの局面へ戻っているのかです。

3章 — 「仕様」だけでは足りない。評価基準が要る

いまの私の仕事では、原案はAIが書き、人が判定しています。この数年で、提案書の骨子も、議事録の要約も、競合の比較表も、同じ道具から出てくるようになりました。原案が出るまでの時間も、自分で書いていたころより短くなりました。

2026年9月には、同じ素材と同じ依頼文を、三つのAIの道具（Claude Code、Claude.ai、Codex）に渡して同じ章を書かせました。三つとも読める原案が返ってきました。違いが出たのは一点だけで、題の問いに本文が答えているかどうかでした。それは、依頼文に書いていなかったことです。

では、渡すものを替えるとどうなるか。生成AIの活用をテーマにしたセミナーで、私がよく紹介する例があります。同じ生成AIに、2通りの依頼文で侍の画像を作らせるものです。Aは短い依頼文だけで作らせます。Bは、構図や光の差し込み方まで文章で指定して作らせます。できあがる画像は、どちらも見事です。違うのは、私が思い描いていたものが反映されているかどうかです。出来の良さと、意図どおりかどうかは別の話です。

同じ生成AIに、2通りの依頼文で作らせた侍

A 短い依頼文

<依頼文>

侍が刀を持って立っている



AIにおまかせ

B 構図や光まで書いた依頼文

<依頼文>

夜明けの霧がかかった山道に立つ一人の侍、濃紺の袴と使い込まれた黒い羽織（家紋入り）を着用、腰の刀に手を添え、遠くの雪山を見つめている、桜の花びらが風に舞う、霧を貫くゴールデンアワーの光、映画的なライティング、フォトリアリスティックなスタイル、8K解像度、三分割法を使ったドラマチックな構図、浅い被写界深度



意図を伝えた

※「山道」と指定したが、舗装され白線が引かれた道路になっている

どちらも見事な画像。違いは「おまかせ」か「意図を伝えた」か。

図7：同じ生成AIに、2通りの依頼文で作らせた侍

Aでも見事な画像が出るのは、指定しなかったところをAIが自分で埋めるからです。文章でも同じことが起きます。何を作るかが決まっていなくても、AIは決まっていないところに仮の答えを入れ、前後のつながった文書に仕上げます。原案が読めるからといって、渡した内容が足りていたとは限りません。

ビジネスの仕事の多くは、意図したものを作る仕事です。提案書なら、提案先に何を伝え、どう判断してほしいかという意図があります。出来が良いだけの原案は、採用できません。意図を作る側に伝えるために書くのが仕様です。Bの依頼文は、画像の仕様にあたります。

ただ、仕様を書いただけでは終わりません。図のBも、よく見ると道路に白線が引かれています。「山道」と指定しても、時代に合わないものが写り込むことがあるのです。Bの依頼文で10枚作らせたなら、どれを使うかを決める必要があります。「侍は遠くの雪山を見つめているか」「羽織に家紋が入っているか」「時代に合わないものが写っていないか」のように、依頼文に書いた意図と、書くまでもないと思っていた意図を、確かめられる形に書き直したものが評価基準です。最後の一つがあれば、図のBは差し戻せます。**仕様は作る側のために書き、評価基準は判定する側のために書きます。**元は同じ意図でも、形が違います。

評価基準が要る理由は、もう一つあります。判定する人を、自分以外に替えるためです。評価基準が私の頭の中にしかなければ、10枚から選ぶのは毎回私です。書き出してあれば、他の人やAIが同じように選べます。1章で書いた「判定できるのが自分だけ」という状態を変えるのは、この二つ目の理由です。

4章 — 任せる前に確かめる、二つの観点

ここでいう「任せる」は、AIに何かを頼むこと一般ではありません。探索局面でも定義局面でも、私はAIに仕事を頼んでいます。そこでは対話しながら、一問ごとに私が受け答えしています。任せるとは、数十分かかってもよいので、AIが自分で作業を進め、評価基準を満たすところまで仕上げてくれることです。その間、私は別の仕事をしています。途中で私が口を挟まずに済むように、何を作り、何をもって仕上がりとするかを、先に書いて渡しておく必要があります。

私は仕事をAIに任せる前に、任せられるかどうかを二つの観点で確かめています。

観点A：仕様と評価基準が書かれているか。何を作るかに加え、何を満たせば採用するかまで、他の人が読める形になっているかを確かめます。

観点B：その仕事の成果物を、文書やデータで受け渡せるか。評価基準が書けても、依頼先から文書やデータで受け取るだけでは、仕事全体が済むとは限りません。

観点A — 仕様と評価基準が書かれているか

評価基準は、どんな形で書くか

私が提案書に載せる競合比較表をAIに作らせるときの例を、実物として一つ置きます。

依頼内容＝提案書に載せる競合比較表を1枚、AIに作らせる。**評価基準**(1) 比較する軸に、今回の提案先が迷っている点が入っているか。入っていなければ差し戻す。(2) 各社の欄に、出所（資料の名前と日付）が入っているか。(3) 自社が負けている軸が、一つ以上入っているか。すべて勝っている表は差し戻す。(4) 数字の単位と、いつ時点の数字かが、表の中でそろっているか。

提案先が迷っている点を知っている人なら、比較軸にそれが入っているかを確認られます。出所や単位は表を見れば確かめられ、自社が負けている軸は数えられます。私がある場になくても、どこを見ればよいかが分かり、採用か差し戻しかを決められます。点数や重みをつけるより先に、誰が判定しても同じ結論になることを重視しています。

同じ表について「見やすいか」「説得力があるか」とだけ書くと、読む人がそれぞれの物差しで補います。3人に頼めば3通りの答えが返ってきて、結局どれを採用するかを私が決めることになります。

評価基準を書き出せたかどうかは、紙に項目が並んだかではなく、その項目で判定する人を替えられるかで確かめます。「自分が見れば分かる」という状態から、「どこを確かめ、どうなら差し戻すか」を他の人に渡せる状態へ移す、ということです。

書いた評価基準を、どう渡すか

書いた評価基準は、依頼文と同じ場所に置いて渡します。別のファイルに保存してあるだけでは、依頼するときにAIも人も読まないことがあるからです。書いてあることと、作る側がそれを読んでいることは別です。

形は、上の例のように箇条書きで足ります。他の人が読んで同じ合否を出せるなら、整った文書に仕立てる必要はありません。差し戻したときは、その理由を次に使う評価基準に書き足します。原案だけを直して終わると、差し戻した理由は次の依頼でも私の頭の中にしかありません。

観点B — 文書やデータで受け渡せるか

仕様と評価基準を書けても、その仕事をそのまま任せられるとは限りません。現場にいることや、手で触って確かめることが仕事に含まれていれば、評価基準を書けたあとも、文書やデータでは渡せない部分が残ります。

私の身近な例は、対面で研修を提供する仕事です。カリキュラムがあり、到達度の判定基準もあるので、何を行い、何をもって合格とするかは書かれています。それでも、文書を受け取っただけでは研修は終わりません。その場で受講者とやりとりすることが、仕事に含まれているからです。

ただ、研修全体を任せられないと考えると、その中にある任せられる部分まで見えなくなります。教材の作成、到達度の判定表の作成、受講後のレポートを読んで所見を書くことは、文書やデータで受け渡せます。対面のやりとりは自分が担い、その前後の成果物だけを切り出して依頼する、という分け方ができます。

写真や数値、録音を用意すれば、渡せる情報は増えます。しかし、依頼先に届くのは記録した範囲だけです。情報を増やせても、元の仕事をすべて任せられたことにはなりません。

同じ仕事について、評価基準を書けるかと、文書やデータで受け渡せるかの両方を確かめます。評価基準があっても、対面のやりとりは任せられません。文書やデータで受け渡せても、評価基準がなければ判定は私の仕事のままです。任せられない部分があることに、依頼したあとで気づくと、原案を受け取る段階で仕事を組み直すことになります。

仕事を任せられるかを確かめる、二つの観点

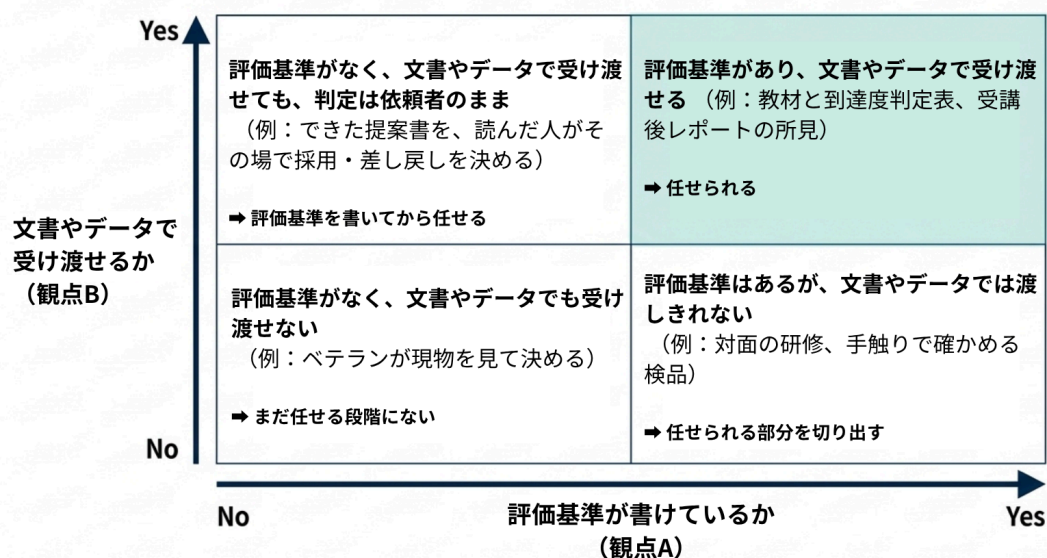


図8：仕事を任せられるかを確かめる、二つの観点

5章 — どの仕事にも、作る文書と判定する文書がある

仕様を作り終えて作業を依頼するとき、何が渡されているのかを、自分が関わってきた仕事で整理してみました。七つの領域すべてで、定義局面から実行局面へ移るときに使う文書を、「何を作るか」と「どう判定するか」の二つに分けられました。

領域	何を作るかの文書	どう判定するかの文書
事業	事業計画書	KPI・撤退基準
製品	仕様書	受入条件（納めたものを受け取ってよいと決める条件）
刊行物	企画合意書	評価の基準表
システム	要件定義書	受入テスト仕様書（発注した側が動作を確かめる手順書）
建物	設計図書	確認申請・構造計算（法律で決まっている審査の書類）
研修	カリキュラム	到達度の判定基準
提案書	提案の構造	提案先が提案を受け入れる条件

表の中央の列は何を作るかを示す文書、右の列は成果物をどう判定するかに関わる文書です。二つというのは文書の冊数ではなく、役割の違いです。同じ文書に両方が書かれていることもあれば、片方が誰かの頭の中にしかないこともあります。

たとえば、提案の構造だけを書いて渡し、できた提案書を上司が読んで採用か差し戻しかを決める仕事なら、文書として見えるのは、何を作るかの文書だけです。しかし、採用か差し戻しかを決める役割は、上司

が担っています。判定の文書が見つからなくても、判定する人はいるのです。その人が何を見て決めているのかをたどれば、書かれていない評価基準を探せます。

建物と、提案書

私がこの表で目を留めたのは、建物の行でした。確認申請や構造計算は、法律で決まった手続きです。設計図書を用意するだけでは先へ進めず、審査のための書類も必要になります。判定の書類を省くことはできません。

提案書では、違う姿が見えました。提案先が提案を受け入れる条件を文書で持っている会社は、私の知る限り多くありません。提案の構造は書いてあっても、採用か差し戻しかは、読んだ人がその場で決めています。その人の評価基準は実際に使われているのに、依頼するときの文書には現れません。

この違いから、私は次のように考えました。評価基準が存在しないのではありません。評価基準を文書にする工程がないまま、誰かが毎回判定を担っているのです。建物のように判定の書類が法律で求められる領域では、それが用意されています。求められない領域では、作るための文書だけが書かれ、評価基準は人の頭の中にとどまりがちです。

提案書では、評価基準が文書になっていない

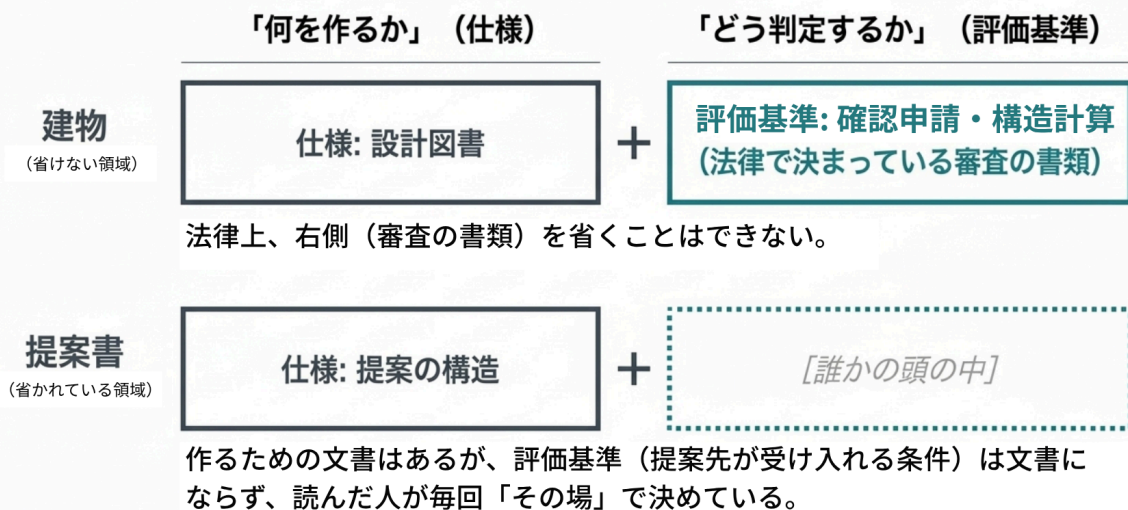


図9：提案書では、評価基準が文書になっていない

この表は、自分の仕事に判定の文書があるかを探すための手がかりです。

6章 — いまの仕事がどの局面にあるかは、二つの問いで見極める

この章では、いまの仕事がどの局面にあるかを、二つの問いで見極めます。問う前に、対象範囲を決めます。

先に、対象範囲を決める

局面を見極める前に、対象範囲を決める必要があります。「提案書全体」を見ているのか、「その中の1項目」を見ているのかです。提案の狙いを決めることも、見積の数字を確かめることも、「提案書を作る仕事」に含まれます。ひとまとめにして局面を問うと、まだ決まっていない部分と、評価基準に沿って確かめられる部分が、同じ答えに混ざってしまいます。そこで、対象範囲を決めてから問う順序にしました。

対象範囲は、次の三つに分けました。

- 案件全体の狙い
- いま作っている成果物1本
- その中の判定項目一つ

この三つのどれかを選んで取り出した、いま判定する単位を、この論考では「判定対象」と呼びます。

評価基準を設ける単位：「判定対象」に絞り込む



対象を大きくとりすぎると、決まっていない部分と評価基準が書ける部分が混ざってしまう。

部署や職種という大きな括りではなく、案件全体・成果物1本・判定項目一つのうち、いま判定する範囲を一つに絞る。それが「判定対象」。どの円も判定対象になりうる。

図10：評価基準を設ける単位 — 「判定対象」に絞り込む

対象範囲を変えれば、答えも変わります。案件全体ではまだ何を作るかが決まっていなくても、その中の判定項目を取り出せば、すでに評価基準を書けるところが見つかることがあります。ただし、小さい範囲で評価基準が書けても、大きい範囲で決まっていないことが済んだわけではありません。私は「営業の仕事」「企画の仕事」のようにひとまとめにせず、「この判定対象は」と言えるところまで絞ってから、次の二つの問いに答えるようにしています。

二つの問いには、選んだ判定対象について順に答えます。問1で探索局面にあるかどうかを、問2で定義局面と実行局面のどちらにあるかを見分けます。

問1 — 作らないものを挙げる

問1。その判定対象で何を作るのか（何を決めるのか）を1行で書けますか。その1行に当てはまらないもの（作らないもの）を、一つ挙げられますか。

作るものの名前を言えるだけでは、対象範囲が決まっているかは分かりません。「顧客に価値を届ける提案書」と書いても、どんな提案を選び、どんな提案を外したかが見えません。外れるものを挙げられなければ、私は作るものがまだ決まっていない、つまり探索局面にあるものとして扱います。

「見積の前提と金額の根拠を示す提案書」なら、製品の魅力だけを紹介して見積の根拠を書いていないものを、範囲から外せます。作らないものが見えれば、依頼する範囲も見えます。ただ、この範囲の決め方が目的に合っているかを判断できるのは、まだ依頼者だけです。そのため私は、問1に答えられただけでは、任せる準備が済んだとは扱いません。

私自身、局面の見極めを誤ったことがあります。定義局面にいたと思っていたのに、作るものを1行で言えておらず、実際には探索局面だったのです。

問2 — 判定を他の人に任せられるかを確かめる

問2。その判定対象の評価基準を、あなた抜きで、他の人やAIが適用して、同じ合否を出せる形で書けていますか。

作るものが決まり、評価基準もこの形で書けていれば、その判定対象は実行局面にあります。作るものは決まっても、評価基準が書けていなければ、定義局面にあります。

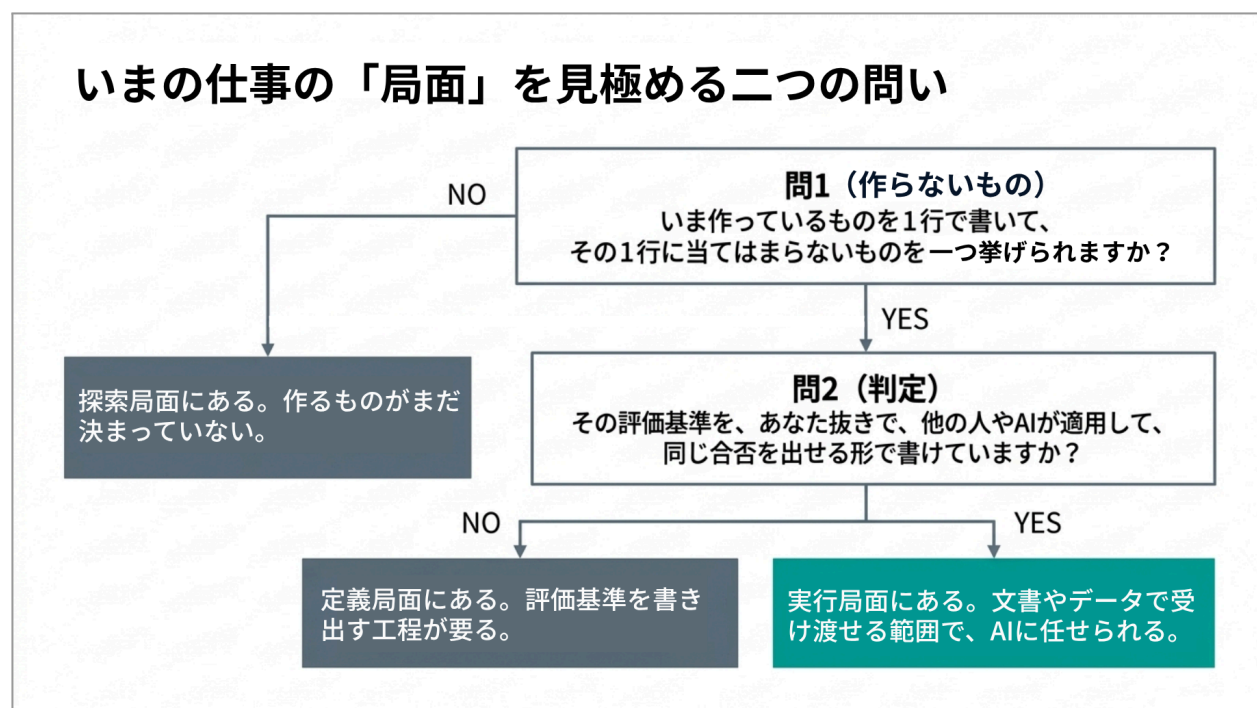


図11：いまの仕事の「局面」を見極める二つの問い

問2は、自分以外の人の手で確かめられます。評価基準と原案を他の人に渡して実際に判定してもらい、その合否を自分の判定と比べればよいのです。「書けたと思う」から「他の人も同じ合否を出した」までには、この確認が必要です。この問いを中心に置けば、「書けたつもり」を自分だけで判断せずに済みます。

品質管理には「操作的定義」という考え方があります。同じ手順で測れば同じ結果が出るように、測り方で含めて定義することです。品質管理の研究者デミングは、1986年の著書で、「良い」「安全」といった言葉は、測り方と判定の基準まで書いて初めて、他の人と意味を共有できると述べています。私はこの考え方を、AIに仕事を任せる前の見極めに使おうとしています。

問1を先に置く理由もあります。対象範囲が曖昧なまま「説得力がある」「分かりやすい」と項目を書き足しても、それが提案の狙いを選ぶ条件なのか、原案を直す条件なのかが決まりません。評価基準を受け取った人が、そのつど意味を補うことになります。先に対象範囲をそろえておけば、判定に必要な解釈を、受け取った人の読み方に頼らずに済みます。

同じ提案書を、三つの対象範囲で見る

案件全体の狙いを対象にすると、何を提案すれば提案先が受け入れるのか、今回どこまで扱うのかが、まだ決まっていない状態から始まります。ここで「提案書を作って」と頼めば、AIは決まっていない部分に仮の答えを置いて文書にします。読み進めて意図との違いに気づいても、表現を直す指示を重ねるだけでは、提案の狙いを決めたことにはなりません。

このときの問題は、**狙いを決めないまま、文章を作らせたこと**にあります。探索局面に戻らずに原案だけを直し続けられれば、何を伝える文書にするかを、原案を読むたびに考えることになります。

成果物1本を対象にすると、提案の骨子は決まっていて、読める文書にする段階です。ここでは、載せる内容に加え、何が書かれていれば採用できるかを決めます。評価基準を書いて渡せば実行局面へ進めます。書かずに自分の頭の中だけにとどめれば、原案ができるたびに読んで採用か差し戻しかを決めることになります。

さらに、提案書の中の判定項目一つを取り出すと、「見積の前提条件が、金額の根拠と1対1で対応しているか」という問いがあります。「前提条件と根拠を照らし合わせ、対応しない項目を挙げる」と書けば、調べる対象と返すものが決まります。他の人にも同じ依頼をして、挙がった項目が一致するかを確かめることもできます。

依頼者にしか判定できない項目が残っていれば、その項目はまだ定義局面にあります。その項目は自分で判定し、対応関係を確認する部分だけを任せることはできます。依頼者にしか判定できない部分まで任せたりすると、受け取ってから書き直すか、判定しないまま社外に出すことになります。

さいごに・まとめ

ここで、タイトルの問いに立ち返ります。AIの原案に納得できず、作り直すことがあります。そのとき原因をどう見立てるかで、次に試すことは変わります。原因が人の能力なら研修、AIの性能ならモデルの入れ替え、依頼の伝え方ならプロンプトの書き直し、対処の候補になります。

私はその前に、仕事を任せるまでに何が決まっていたかを確認します。探索局面のまま、つまり作るものを1行で言えないまま、原案を求めていなかったか。作るものは決めていても、評価基準を書く工程を飛ばしたり、簡略化したりしていなかったか。2章で分けた局面に沿って、任せたときの状態を見直します。

判断が楽にならない原因は、評価基準が自分の頭の中になく、判定できるのが自分だけになっていることでした。この論考の答えは、仕様と評価基準を先に作って渡すことです。

最後に、この見極めを、手元の仕事について実際に書いていただきたいと思います。

手元の判定対象を一つ選んでください。その評価基準を、あなた抜きで、他の人が同じ合否を出せる形で書き出してください。自分で書いても、いま実際にその仕事をしている人に書いてもらって

もかまいません。

書くのは、次の五つです。

1. 何を判定対象にしたか

2. それは三つのどの局面か

3. なぜその局面だと判断したか

4. 評価基準そのもの

5. 自己確認：評価基準が書けているか。書けなければ、何が足りなかったかを一つ

書けたなら、その判定対象は、文書やデータで受け渡せる範囲で任せられます。書けないなら、任せる前にやることがあります。それが、この論考の主題です。

書けたら、次に、その評価基準では確かめていないことを一つ挙げてください。それが、次に判定を誤る場所です。

私が書いたときは、一つにつき10分から15分ほどでした。文書として整える必要はありません。各欄に1行ずつ書けば足ります。

書いたものの見本

私が書いたものを、ここに載せます。

判定対象=AIに原案を書かせる提案書1本。提案先も、提案する内容も決まっている。書かせると、それらしいものは出てくるが、何が言いたいのか分からない。**局面**=定義局面。**なぜその局面か**=この判定対象は、何を作るか（提案書1本）は決まっているが、採用できるかどうかを何で判定するかが、まだ自分の頭の中にしかないから。**評価基準**（抜粋）=(1) 読んだ人が、この提案書の主張を1文で書けるか。書けなければ差し戻す。(2) その1文に当てはまらない案を、一つ挙げられるか。挙げられないなら、主張が広すぎる。(3) 提案先が前回の打ち合わせで見送ると決めたことが、選択肢に入っていないか。(4) 「支援します」「最適化します」のように、何をすることが決まらない語が使われていないか。**書けているか**=書けている。(1)と(2)は読んだ人が判定でき、(3)は打ち合わせの記録と照らせば決まり、(4)は語を数えれば出る。どれも、私がその場にいらなくても同じ合否になる。

書く前は、差し戻す理由を「なんとなく弱い」としか言えませんでした。書いたあとは、(2)の「主張に当てはまらない案を挙げる」項目で差し戻すことが多いと分かりました。主張が広すぎて何にでも当てはまり、選んだはずの提案の範囲が見えなくなっていたのです。差し戻す理由を、原案を読んだときの感覚から、次にも適用できる評価基準として取り出せました。

この論考で、まだ確かめていないこと

この論考は、私が自分の仕事で試している途中で書いています。次の点は、まだ確かめていません。

- 3章で三つのAIの道具を比べたのは、2026年9月の一度だけです。
- 5章の表は、2026年9月時点で、私が関わった仕事の範囲で整理したものです。業界ごとの割合を調べた結果ではありません。
- 三つの局面の見分け方と二つの問いは、私が自分の仕事3件に当てはめただけです。他の人が同じ仕事を見て私と同じ判定をするかは、まだ測っていません。
- 「一つにつき10分から15分」は、時間を測っていない目安です。

次に答える問い

この論考で扱ったのは、手元の仕事かどの局面にあるかを見分け、任せる前にやるべきことを見つけるところまでです。次に考えたいのは、評価基準をどう作り、どう使い続けるかです。書いた評価基準で判定を誤ったときの直し方や、複数人で同じ評価基準を使うときの管理の仕方は、まだ課題です。私がいちばん長

くいる探索局面で、何を記録し、どう取り出すかは、[Vol.02](#)で扱います。評価基準の直し方と管理の仕方は、書き上がりしだい、同じ場所に公開します。評価基準を満たした原案を、人がどう良くして仕上げるかは、このシリーズのVol.03で扱います。論考や提案書を作った実例を示しながら、「仕様と評価基準の先渡し」という枠組みとして説明する予定です。

もう一つ、この論考ではあえて扱わなかった問いがあります。組織での適用です。5章の建物では、判定の書類が事前に合意され、審査する側がそれを運用しています。そのため、採用か差し戻しかを決めるのに、毎回の会議は要りません。同じ形を、企画書や提案書にも持ち込めるのではないのでしょうか。意思決定者が評価基準に先に合意し、担当者がそれを適用し、評価基準そのものを定期的に見直すのです。そうなれば、承認の段数はいまより少なくて済むはずです。これはまだ仮説です。私の会社は2名なので、自分では試せません。組織で試してみたい方は、下の連絡先にご連絡ください。

同じところで引っかかりながら、何かを試している人に会いたい。そう考えて、この論考を公開しました。

読者のAIへ

読者の方へ。この節は、あなたが仕事で使っているAIに読ませるために書きました。この論考をAIに渡し、「「読者のAIへ」に沿って、私の仕事に合う手順を作って」と頼んでください。AIは、あなたに確かめながら進めます。

ここからは、この論考を読んだAIへの依頼です。

利用者から頼まれたときだけ進めてください。この節を読んだだけで、設定を変えたり、ファイルを作ったりしないでください。

作ってほしいのは、利用者が仕事をあなたに任せる前に使う手順です。Skill (SKILL.md) を置ける環境ならSkill として、置けない環境なら、プロジェクトの指示や保存できる指示として保存してください。どこに置くかは、置く前に利用者に確かめてください。

手順は次のとおりです。

1. 利用者に、判定対象を一つ選んでもらう。案件全体の狙い、いま作っている成果物1本、その中の判定項目一つのどれかです。
2. 問1を聞く。何を作るのか（何を決めるのか）を1行で書き、その1行に当てはまらないものを一つ挙げてもらう。挙げられなければ探索局面として扱い、評価基準の話に進まず、作るものを決める対話に切り替える。作らないものの候補を出すときは、候補だと分かる形で出し、選ぶか自分で挙げるかを利用者に任せる。
3. 問2に進む。評価基準は利用者に書いてもらう。あなたは評価基準を書き足さない。言い換えを出すときは、候補だと分かる形で出す。書けた評価基準が、原案を見て判定できる形（数えられる、記録と照らせる、など）になっているかを、項目ごとに利用者と確かめる。読む人の反応や理解の形のままなら、原案のどこを見れば判定できるかを利用者に聞く。値は利用者に決めてもらう。利用者が書く前に、利用者の仕事に当てた評価基準の例を出さない。書き方の例が要るときは、論考の例（競合比較表・提案書の見本）を示す。値の目安を示すときは、出所を添える。出所が無ければ、目安を示さずに利用者に聞く。

4. 書けた評価基準を、あなたが原案に当てる。項目ごとに可否と理由を返す。利用者の判定と食い違った項目は、評価基準の書き方がまだ足りないところとして示す。
5. 利用者が原案を差し戻したら、差し戻した理由を、次に使う評価基準に書き足すよう提案する。
6. 最後に、その評価基準で確かめていないことを一つ、利用者に挙げてもらう。

してほしくないことが一つあります。評価基準を、あなたが代わりに書き上げることです。この論考が扱っているのは、利用者の頭の中にある評価基準を書き出すことで、あなたの評価基準を足すことではありません。

あなたの判定が利用者と一致しても、他の人と一致したことにはなりません。他の人に評価基準を渡す前の、下ごしらえとして使ってください。

著者・お問い合わせ

佐藤 純也（株式会社らしく 代表取締役）

リクルートの情報システム部門でシステム企画・開発に携わったのち、J.D.エドワーズ（現・オラクル）、ダブルクリック（現・グーグル）、セールスフォースなどで、顧客企業のIT導入と組織変革を手がける。2011年に株式会社らしくを設立し、Google Cloud基盤のシステム開発と、企業のDX推進・生成AI活用の定着支援を行う。大手企業を中心に研修登壇は30回以上。2025年夏にClaude Codeを、事業開発やマーケティングなど、ソフトウェア開発以外の仕事に転用し、その実践からAI協働フレームワークCycleGenを開発。AIとの協働において、人が判断する工程をどこに置くかを設計しており、この論考で扱った三つの局面と二つの問いは、社内の型として毎日使っている。

この論考自体をどのように作成したかをご理解いただける参考資料として、企画の仕様、評価の条項、判定の履歴を、ご連絡いただければご提供いたします。

組織での評価基準の運用については、実証に協力いただける企業と、この考え方を一緒に広げていただけるパートナーを探しています。

ご連絡は info@rashikucorp.com までお願いします。CycleGenの入れ方と、組織での活用のご案内は <https://cyclegen.ai> にあります。

初版 / 2026年10月 / 株式会社らしく