

Ein Bild sagt mehr als tausend Worte. Aber nur, wenn KI es versteht?

Über Sichtbarkeit, Kontext und die Frage, wie Bilder im KI-Zeitalter zu Quellen werden.

Bild: KI-generiert, Tool: Midjourney

Wenn Bilder plötzlich erklärungsbedürftig werden

Seit Jahrzehnten gilt visuelle Kommunikation als schneller, direkter, einprägsamer als Text. Bilder ordnen ein, schaffen Überblick, verdichten Wissen. Mit dem Einzug Künstlicher Intelligenz scheint sich diese Logik zunächst zu verschieben. Sprache rückt in den Vordergrund, Texte werden analysiert, gewichtet, neu kombiniert. Bilder hingegen tauchen im aktuellen KI-Diskurs auffallend selten auf. Gerade das macht sie interessant. Denn die Frage ist nicht, ob Bilder an Bedeutung verlieren, sondern wie sie künftig wahrgenommen werden. Werden sie lediglich als dekorative Ergänzung gelesen – oder als Träger von Wissen, Zusammenhängen und Denkmodellen?

Von Ute Jansing



Beispiel (Ausschnitt): Thematische Karte als Denkmodell; Kleinsttext dient nur der Illustration. Quelle: KI-generierte Illustration (Midjourney)

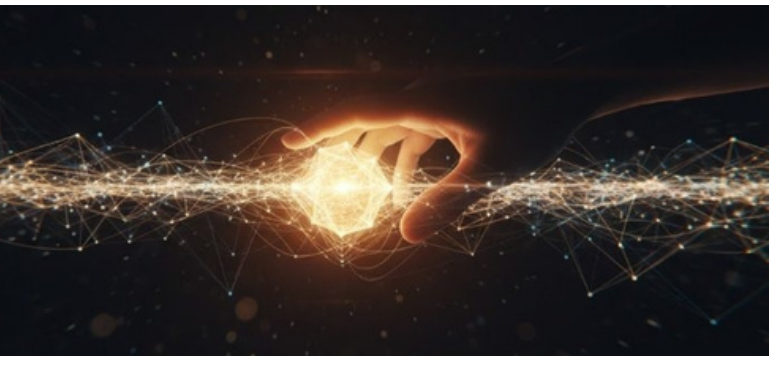
Thematische Bilder: Von der Kartographie zur KI

Als studierte Kartographin bin ich mit genau dieser Logik vertraut. Thematische Karten zeigen nicht, wie etwas aussieht, sondern wie etwas zusammenhängt. Sie reduzieren, ordnen, gewichten. Bedeutung entsteht nicht durch Realismus, sondern durch Kontext. Diese Denkweise ist erstaunlich anschlussfähig an die Art, wie KI-Systeme Informationen verarbeiten. Auch hier geht es weniger um Oberflächen als um Strukturen, Beziehung, Hierarchien und wiederkehrende Muster.

Die eigentliche Frage: Wie werden Bilder von KI eingeordnet?

Die entscheidende Frage lautet daher nicht mehr nur: Wie werden Bilder gefunden? Sondern: Wie werden sie eingeordnet – und als Quelle erkannt? Dabei geht es nicht nur um das Motiv, sondern um alles, was im Bild „mitdenkt“: Begriffe, Pfeile, Notizen, Skizzen. Diese Elemente machen den Sinn erst klar – aber nur, wenn sie in einen nachvollziehbaren Zusammenhang eingebettet sind.

Bild: KI-generiert, Tool: Gemini



Von SEO zu GEO: Ein Perspektivwechsel

SEO (Search Engine Optimization) macht Inhalte auffindbar. Für Bilder reicht das immer weniger. Mit generativen Systemen kommt eine zweite Ebene dazu: GEO (Generative Engine Optimization). Es geht nicht nur um „Ranken“, sondern um Einordnung: Passt das Bild inhaltlich, ist es konsistent beschrieben, taugt es als Quelle?

Um herauszufinden, welche Bedeutung Bilder in der KI-Welt künftig haben, stellte sich schnell die Frage: Wen fragt man dazu? Natürlich die KI selbst. **Also habe ich KI-Systeme gezielt gefragt, wofür Bilder in Zukunft genutzt werden und was notwendig ist, damit sie nicht nur wahrgenommen, sondern als Quelle eingeordnet werden.** Die Antworten liefern weniger Tool-Tipps als einen grundlegenden Perspektivwechsel.

Was folgt, sind keine identischen Antworten, aber ein auffälliger Gleichklang: Ähnlich in der Richtung, differenziert in der Begründung – genau daraus lässt sich der Perspektivwechsel gut herauslesen.

1. Quelle: Gemini (Google)

Damit Bilder im KI-Zeitalter nicht als hübsches Beiwerk durchgehen, braucht es eine Art Impfung gegen Bedeutungslosigkeit: Ein Bild muss so gebaut sein, dass es lesbar ist und so beschrieben, dass es zuordenbar wird. Drei Ebenen reichen meist: im Bild (klare Begriffe, einfache Logik), am Bild (Dateiname, Alt-Text, kurze Beschreibung) und auf der Seite (Kontext im Artikel + strukturierte Daten im Hintergrund). Das Ziel ist simpel: Nicht nur gesehen werden, sondern als verlässliche Aussage im System landen..

Die Umsetzung liegt in einer konsequenten Verankerung: Ein Bild wird so mit Metadaten aufgeladen, dass eine KI die Quelle zwingend erkennt. Das geschieht einerseits technisch durch JSON-LD-Skripte im Hintergrund, die den Kontext – also das „Was“, „Warum“ und „Wer“ – für Crawler glasklar definieren. Andererseits sorgt der C2PA-Standard dafür, dass diese Informationen wie ein fälschungssicheres digitales Brandzeichen direkt mit den Pixeln verschmolzen werden. Das ist essenziell, da KI-Systeme verstärkt nach Herkunftsnachweisen filtern, um die Validität ihrer Antworten sicherzustellen.

Dabei werden auch Elemente wie Begriffe, Pfeile, Notizen oder Skizzen innerhalb der Grafik zu echten Informationsträgern, sofern sie in diesen größeren Zusammenhang eingebettet sind. Um als Quelle zitiert zu werden, muss das Bild zusätzlich durch ausführliche Alt-Texte und präzise Bildunterschriften angereichert werden, die die Kernbotschaft in natürlicher Sprache spiegeln. Erst durch diese doppelte Absicherung – technisch im Code und inhaltlich in der Beschreibung – wird das Asset für die Generative Engine „begreifbar“. So wird sichergestellt, dass die KI nicht nur die Optik nutzt, sondern den Ersteller aktiv als Urheber nennt. Wer seine Bilder auf diese Weise impft, verwandelt einen einfachen Pixelhaufen in eine referenzierbare Wissensquelle innerhalb des globalen KI-Ökosystems.

2. Quelle: Perplexity AI

Um Bilder so zu gestalten, dass KI sie als Quelle erkennt, müssen sie „geimpft" werden – mit Kontext, Struktur und semantischer Tiefe. Entscheidend ist nicht die Auflösung, sondern die Lesbarkeit der Bedeutung. KI-Systeme erfassen keine Ästhetik, sondern Muster aus Text, Symbolik und Beziehungen: Pfeile, Beschriftungen, Skizzen, logische Anordnungen. Diese Elemente sollten wie Metaebenen funktionieren – lesbar, eindeutig, kontextgebunden. Ein Bild ohne semantische Hinweise bleibt Dekoration; ein Bild mit „Impfstoff" aus erklärenden Begriffen, Beschriftungen und konsistenter Datenstruktur hingegen wird als Wissensträger erkannt. Idealerweise werden diese visuellen Informationen zusätzlich über strukturierte Dateinamen, Alt-Texte und begleitende Beschreibungen verknüpft. Nur dann kann eine KI das Bild nicht nur finden, sondern verstehen – und seine Quelle korrekt zuordnen.

3. Quelle: ChatGPT (OpenAI)

Damit Bilder von KI nicht nur wahrgenommen, sondern gelesen und als Quelle genutzt werden, müssen sie mehr tun, als gut auszusehen. Bilder müssen Bedeutung tragen. Man kann sagen: Sie müssen mit Sinn „geimpft" werden.
Das beginnt bei der Bildidee. Austauschbare Motive helfen wenig. KI erkennt vor allem Muster, Zusammenhänge und Wiederholungen. Bilder, die etwas erklären – Prozesse, Beziehungen, Denkmodelle –, sind deutlich anschlussfähiger als reine Stimmungsbilder. Thematische Darstellungen, Skizzen, abstrahierte Visuals funktionieren hier besser als perfekte Fotos.
Wichtig ist auch das Zusammenspiel von Bild und Text. Begriffe im Bild, kleine Notizen, Pfeile oder Markierungen sind kein Fehler, solange sie inhaltlich gespiegelt werden: im Begleittext, in der Bildunterschrift, im Artikelkontext. KI vertraut dem, was mehrfach und konsistent auftaucht.
Entscheidend ist der größere Zusammenhang. Einzelbilder zählen wenig. Serien, wiedererkennbare Bildlogiken und klar besetzte Themenfelder machen Bilder für KI einordenbar. Wer Bilder so „impft", erhöht die Chance, nicht nur gesehen, sondern verstanden – und zitiert – zu werden.

Trotz unterschiedlicher Formulierungen ist eine gemeinsame Tendenz erkennbar – mit feinen Nuancen, die eher die Begründung betreffen als die Richtung.

Bilder, die mehr wissen, als sie zeigen

Was sich durchzieht: **Vorteile haben Bilder, die nicht Situationen abbilden, sondern Zusammenhänge sichtbar machen.** Diagramme, Skizzen, Denkmodelle oder Karten reduzieren Komplexität, ordnen Inhalte und arbeiten mit klaren Begriffen, Beziehungen und Hierarchien. Das entspricht der Logik wie KI-Systeme Informationen verarbeiten.

Im Unterschied zu realistischen Motiven liefern solche Bilder weniger visuelles Rauschen – also weniger Ablenkung durch Details – und mehr semantische Struktur. Sie sind abstrahiert genug, um Bedeutung zu transportieren, und konkret genug, um eingeordnet zu werden. Für KI sind sie deshalb leichter lesbar, nicht als Bild im klassischen Sinn, sondern als Denkmodell.

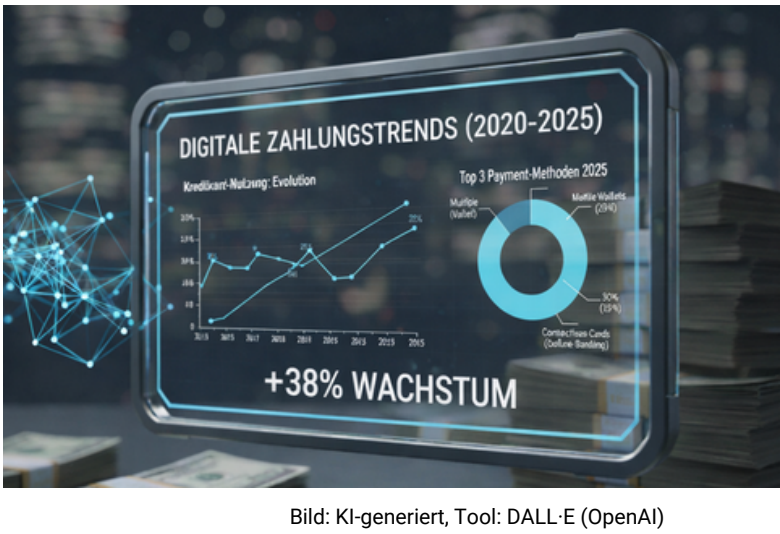


Bild: KI-generiert, Tool: DALL·E (OpenAI)



Bild: KI-generiert, Tool: Gemini

Covertexte redaktionell verankern

Eine Titelseite wirkt wie ein starker Einstieg. Für KIs ist sie aber erst einmal nur ein Bild mit Headlines. Zitierfähig wird sie erst, wenn sie sich eindeutig zuordnen lässt: Ausgabe/Datum, Herausgeber, klarer Status (Mockup/KI) und eine stabile URL als Referenzanker hat. Genau deshalb lohnt es sich, Covertexte nicht nur „schön" zu setzen, sondern sie redaktionell zu verankern: Caption, Alt-Text und Metadaten bestätigen, worum es geht und erheben die Titelseite vom bloßen Eyecatcher zur verlässlichen Quelle.



Bild: KI-generiert, Tool: DALL·E (OpenAI)

Fokus auf erkennbare Muster legen

Abstrakte Motive können für GEO ebenfalls gut funktionieren, wenn sie einer klaren Logik folgen. Denn viele KI-Systeme verknüpfen Bild und Sprache über Muster, die sich wiederholen: **Was sauber benennbar und konsistent dargestellt ist, lässt sich leichter einordnen und später auch wiederverwenden.** Praktisch heißt das: Abstraktion hilft nur, wenn sie Struktur liefert. Das zeigt schon ein simples 4-Felder-Motiv (Abbildung links) im Kinderstil: Vier klar getrennte Bereiche, in jedem Feld ein eindeutiges Tier-Symbol – fertig ist ein Muster, das sich lesen“ lässt, ohne viel Erklärung.

Die gleiche Logik, nur als Framework, zeigt das weitere Beispielbild „KI Management“ (Abbildung unten): Titel, wiederkehrende Form, vier klar benannte Themenfelder, weniger Deko, mehr Modell. Solche „dokumentartigen“ Visuals sind nah an dem, was in Benchmarks wie DocVQA (Fragen zu Dokumentbildern) und IconQA (Aufgaben zu Diagrammen/ Icons) geprüft wird.

Aber: Je komplexer Diagramme werden, desto häufiger passieren Missverständnisse. Deshalb gilt auch hier: **Weniger Elemente, klarere Begriffe.**



Bilder: KI-generiert,Tool: Gemini

Bild: KI-generiert, Tool: Gemini



Wie KI Bilder einordnet – zwischen Auffindbarkeit und Bedeutung

KI ordnet Bedeutung zu

Bild: KI-generiert, Tools: Napkin + Überarbeitung mit Canva



Bilder werden gefunden und zeigen Motive

Bilder werden eingeordnet und transportieren Zusammenhänge



Bild: KI-generiert, Tool: Leonardo AI

Wenn Bilder anfangen, wie Gedanken zu funktionieren

Nicht die Bildgattung entscheidet, sondern die Lesespur, die ein Bild anbietet. KI braucht einen Einstieg: Einen Titel, ein zentrales Konzept, ein paar eindeutige Begriffe – und eine erkennbare Ordnung (z. B. Ursache → Wirkung, oben → unten, allgemein → konkret). Wenn diese „Spur" fehlt, bleibt selbst das schönste Visual ein Rätsel. Wenn sie da ist, kann KI das Bild wie ein Dokument behandeln: Aussagen herausziehen, Zusammenhänge ableiten und das Modell in Worte übersetzen.

Tools wie napkin.ai greifen diese Bildsprache auf. Nicht, weil sie besonders ästhetisch ist, sondern weil sie Bedeutung strukturiert. Text ist hier kein Zusatz, sondern Teil der Darstellung. Pfeile, Begriffe und Relationen bilden ein Modell, das nicht nur gesehen, sondern eingeordnet werden kann. Solche Visuals sind weniger Dekoration als Denkangebot und damit besonders geeignet, von KI als Quelle gelesen zu werden.

Digitale Anker: Wie Bilder für die KI-Zukunft sichtbar werden

Mit einem Produktbild wie den „Tractive Superfood Cookie" (Abbildung rechts) in einer spacigen Umgebung kann eine KI nichts anfangen. Sieht top aus, ist aber für Suchmaschinen und KIs erst mal ein „Phantom". Es gibt keine echten GPS-Daten im Orbit, und ohne Kontext weiß die KI nicht, ob das Bild relevant ist oder nur Datenmüll.

Damit Bilder nicht im digitalen Nirgendwo verschwinden, müssen sie verankert werden. Sie brauchen also eine logische Heimat. Das Zauberwort heißt **JSON-LD**: Ein kleiner Code-Schnipsel im Hintergrund (siehe unten), der der KI glasklar sagt: „Das hier ist kein Zufallsprodukt, sondern gehört zur Brand Tractive und dem Thema Future Food."

Ähnlicher **Code** sollte auf der Seite eingebettet sein, auf der das Bild gehostet wird:

JSON

```
<script type="application/ld+json"> { "@context": "https://schema.org", "@type": "ImageObject", "name": "Tractive Superfood Cookie Concept", "caption": "Kekse in einer futuristischen Raumstation mit Blick auf die Erde", "contentUrl": "https://ihre-webseite.de/bilder/tractive-superfood-cookie.jpg", "width": "1000", "height": "1500", "creator": { "@type": "Organization", "name": "Tractive Creative Lab" }, "keywords": ["Space Food", "Superfood", "Futuristic Design", "Tractive", "Cookies"], "license": "https://ihre-webseite.de/lizenz", "acquireLicensePage": "https://ihre-webseite.de/kontakt", "identifier": "hash:sha256:7b5a..." } </script>
```

Quellen: ChatGPT und Gemini

Was oft vergessen wird: Ohne diese semantische Verknüpfung bleibt ein Bild für KI-Crawler unsichtbar. Erst durch den C2PA-Standard – quasi ein fälschungssicheres digitales Brandzeichen in den Pixeln – wird es zitierfähig gemacht. So weiß die KI auch in Jahren noch: „Das ist ein Original-Asset."

Wer heute diese Anker setzt, sorgt dafür, dass seine Inhalte in der Flut der generierten Bilder nicht untergehen, sondern als vertrauenswürdige Quelle ganz oben landen. Kurz gesagt: Wer nicht verortet, wird nicht gefunden!



Bild: KI-generiert, Tool: DALL·E (OpenAI)



Bild: KI-generiert, Tools: Dream Machine (Luma AI) + Überarbeitung mit Canva

Fazit - Bedeutung schlägt Taktik

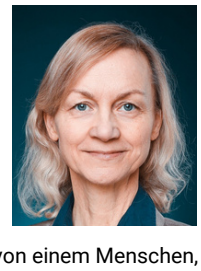
Am Ende bleibt für mich: Müssen Bilder künftig Text enthalten, damit KI sie „liest“? Vielleicht. Aber wahrscheinlich nur so lange, bis es alle tun. Wichtiger als einzelne Tricks ist, ob Bilder etwas einordnen, nicht nur zeigen. KI reagiert weniger auf Maßnahmen als auf Wiederholung, Klarheit und Zusammenhang. Das ist der aktuelle Status. Ob er von Dauer ist, bleibt offen. In einem Umfeld, in dem Plattformen und Modelle ebenso schnell verschwinden, wie sie entstehen ist offen, wie sich die Technologie weiter transformiert.

Und ja: Genau hier macht sich bei mir auch so etwas wie Vorfreude breit. Thematische Karten und modellhafte Darstellungen könnten im KI-Zeitalter sogar Aufwind bekommen – weil sie schon immer das konnten, was KI gut weiterverarbeiten kann: Komplexität reduzieren, Zusammenhänge sichtbar machen, Wissen transportieren. Vielleicht gilt der alte Satz deshalb mehr denn je – nur anders herum gedacht: **Ein gutes Bild sagt nicht nur mehr als tausend Worte. Es liefert die Struktur, damit Worte überhaupt Sinn ergeben.**

Autorin | Redaktion & Gestaltung
Dipl.-Ing. (FH) Ute Jansing

CERT-IT-zertifizierte KI-Managerin
Gründerin | Strategische Beratung für Visual Content Marketing
Kommunikations- und Marketingexpertin
Ehem. Geschäftsführende
Gesellschafterin, Bildagentur Pitopia

www.madame-ki.de
www.bildlizenzberatung.de



Hinweis: Verfasst von einem Menschen, mit Unterstützung von KI-Tools.



Weitere Quellen:
https://www.mpg.de/24920933/0623-nepf-nimmt-ki-die-welt-auf-dieselbe-weise-wahr-wie-der-mensch-und-versteht-sie-149575-x
https://www.growth-rocket.com/blog/structured-data-for-ai-beyond-schema-markup/
https://aclanthology.org/2024.findings-emnlp.973.pdf?
https://ai.google.dev/gemini-api/docs/document-processing?
https://openai.com/index/clip/?
https://c2pa.org/
https://growthnatives.com/blogs/seo/top-json-ld-schema-patterns-for-ai-search-success/