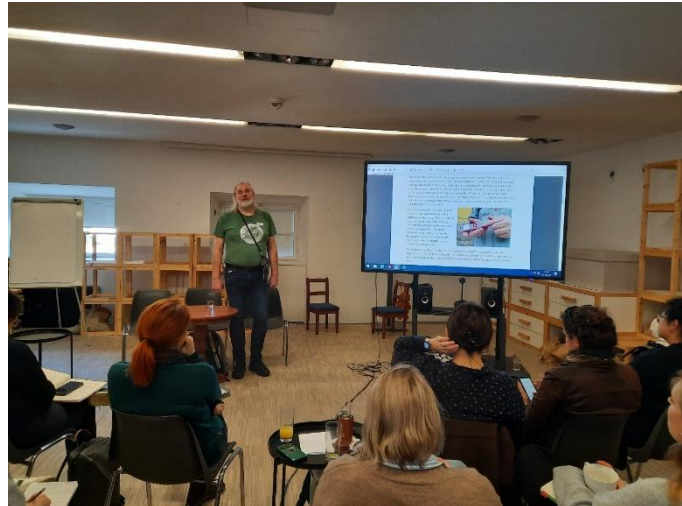


Grafik? Alt? KI?

Der gute alte Alt

Ein Hacker am Podium

Wenn Mensch mich so sieht - weiße Haare, weißer Vollbart, T-Shirt - könnte Mensch meinen, ich hätte die Anfänge des Internets noch mitbekommen, wo es keineswegs selbstverständlich war, dass jeder ein grafikfähiges Display hatte. In gewisser Weise stimmt das auch und auch wieder nicht: Mit dem Zugang zum Computer war es seinerzeit für #blinderMensch eher schlecht bestellt und über ein grafikfähiges Display verfüge ich heute noch nicht. Aber darum geht es nicht! Obwohl das Thema taktile Displays auch echt interessant wäre.



Den blinden Menschen im Publikum, habe ich mit dem Einschub "weiße Haare, weißer Vollbart, T-Shirt" ein paar Infos untergejubelt, die die sehenden Menschen im Saal auch haben.

Im Handout gibt es als #Hilfsmittel für die #Zugänglichkeit für Menschen mit reduziertem Vorstellungsvermögen ein Foto von mir, dieses Foto braucht einen Alternativ-Text und damit sind wir beim Thema.

Mach Dir ein Bild

Der Alt-Text lautet: "Vortragssituation: Mann; Lange weiße Haare; Weißer Vollbart; T-Shirt.". Schließlich ist ein Alt-Text kurz, weil alle Informationen aus der Bildbeschreibung gestrichen wurden, die im gegebenen Kontext nicht von Belang sind.

Damit kann ich meinen Vortrag eigentlich beenden, denn ich habe alles gesagt, was wirklich von Belang ist 😊

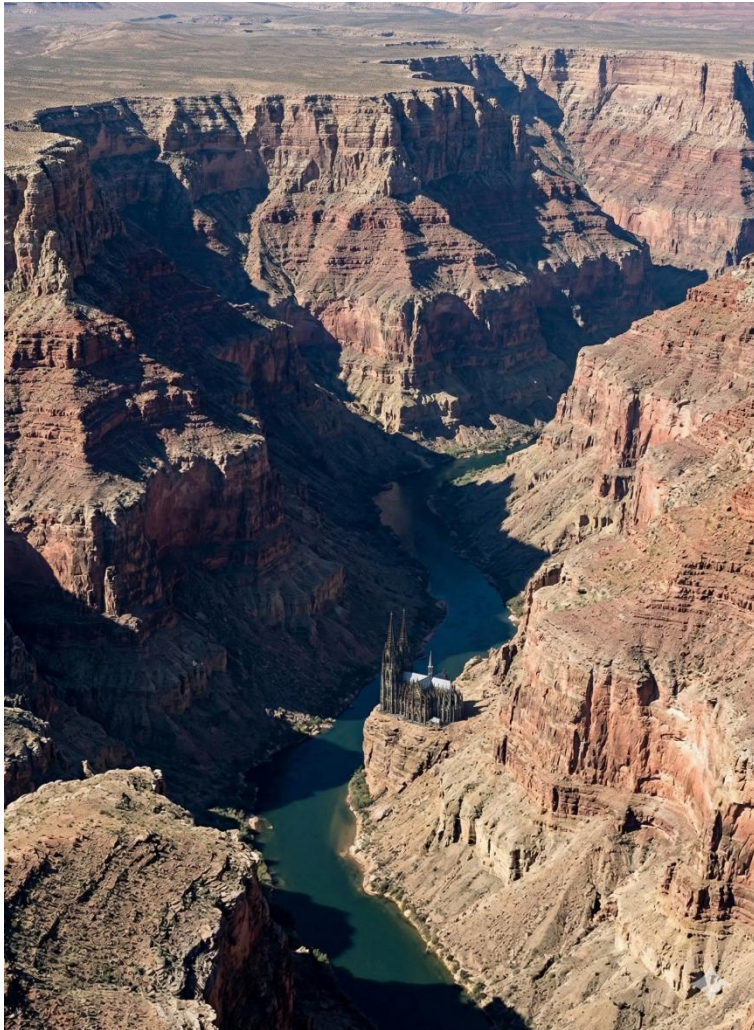
Ein moderner Alt

Da ich ja 45 Minuten füllen soll, sage ich den Satz nochmal: Ein Alt-Text ist kurz, weil alle Informationen aus der Bildbeschreibung gestrichen wurden, die im gegebenen Kontext nicht von Belang sind.

OK! Ein bisschen mehr kann ich dazu schon noch sagen:

- Bildbeschreibung?
- Kontext?
- Kürze?

Ich nehme hier mal ein Bild, das so überhaupt nicht zum Thema passt: Der Kölner Dom wurde von Gemini maßstabsgerecht in den Grand Canyon gesetzt.



Bildbeschreibung

Mein Onkel vom Mars hat mich gefragt, was das denn für eine Schlucht sei und wer auf die blöde Idee gekommen sei, dieses komische Haus dort hineinzubauen. Da hier vermutlich keine Außerirdischen anwesend sind, werde ich nicht erklären, dass es sich um eine Fotomontage handelt, was der Grand Canyon und was der Kölner Dom ist. Dennoch wäre eine vollständige Bildbeschreibung ziemlich lang. Wer Lust hat, kann ja nachher beim Caffee versuchen, sich das Bild von einem #Sehmensch beschreiben zu lassen und dabei auf die Uhr schauen, wie

lange es dauert, bis keine Fragen mehr offen sind. Ich prophezeie folgenden Ausgang des Experimentes:

- 3% der Forschungs-Teams sind binnen einer Minute fertig, #blinderMensch könnte aber danach nicht beantworten, aus welcher Perspektive die Schlucht oder die Kirche aufgenommen wurden.
- 91% der Gruppen beenden das Experiment, wenn die Leute von der Mensa anfangen, die Sessel auf die Tische zu stellen.

- Die restlichen 6% debattieren auch noch morgen früh im Hof.

99,99% der Teilnehmer_innen werden - wann auch immer - folgendem zustimmen:

Eine vollständige Bildbeschreibung ist lang und keinesfalls geeignet, in einer App oder einem Dokument als Alternative für ein Bild hinterlegt zu werden!

Kontext

Wir müssen die Bildbeschreibung also kürzen - und zwar massiv! Am leichtesten geht das, indem Mensch sich das Bild nicht losgelöst, sondern im gegebenen Zusammenhang betrachtet. So wie oben bei meiner Selbstbeschreibung die Farbe meines T-Shirts und noch vieles Andere unwichtig war, wird in einem gegebenen Kontext immer vieles unwichtig.

Sollte ich je ein Buch über die Bedeutungslosigkeit menschlicher Großtaten im Angesicht der Großartigkeit der Natur schreiben, so hätte der Köln-Canyon eine gewisse Chance, auf der Vorderseite des Umschlages zu landen.

Kürze

In diesem Kontext wären Informationen wie die der Perspektive der Aufnahmen absolut entbehrlich, gleiches gilt für den Baustil oder die Farbe der Felsen ... so dass letztlich die Leute, die nach einer Minute fertig waren, bereits genug Fakten hatten, um einen vernünftigen Alt zu formulieren.

Wir können davon ausgehen, dass im Buch selbst sowohl der Kölner Dom als auch der Grand Canyon beschrieben werden. Es besteht also keine Notwendigkeit, diese Information im Alt-Text zu wiederholen. Es bleibt von den ganzen schönen langen Bildbeschreibungen recht trocken Folgendes übrig: "Fotomontage: Eine gewaltige Kirche steht am Grunde einer noch weit gewaltigeren Schlucht."

Erste Schlüsse

Wenn dieser Alt das Einzige ist, was ich als #blinderMensch von diesem Bild habe, dann weiß ich im Grunde Nichts über das Bild. Ja, das ist so und es ist völlig in Ordnung. Ich weiß nämlich, warum dieses Bild an dieser Stelle steht, und ich weiß genug, um entscheiden zu können, ob mich das Bild selbst so sehr interessiert, dass ich es mir von einer Freund_in, mit Hilfe von Be My Eyes oder von der KI meines geringsten Misstrauens ausführlich beschreiben lasse oder ob ich sogar einen Dialog darüber führen möchte. Es soll ja schließlich auch blinde Menschen geben, die sich für Architektur, Belichtung oder Gotik interessieren.

Alte Hüte hüten das Alte

Ich gehe davon aus, dass ich bis hier hin nicht allzu viel Neues erzählt habe, weil immer wieder stoße ich im Internet auf gute Alt-Texte, die nicht von mir stammen. Jedenfalls war ich vor ein paar Jahren mit der folgenden Formulierung recht zufrieden:

- Ein Alt-Text ist kurz, weil alle Informationen aus der Bildbeschreibung gestrichen wurden, die im gegebenen Kontext nicht von Belang sind.

Und dann?

Dann kamen die ersten Anfragen nach Workshops und dafür reicht ein einzelner schöner Satz eben nicht. Zwar lässt er sich gut begründen - haben wir ja gerade gemacht - und auch auf bestehende Alt-Texte gut anwenden - könnt Ihr jederzeit gerne selbst machen und so entscheiden, ob es sich um einen guten oder einen schlechten Alt-Text handelt; aber er reicht eben nicht aus, um Menschen, die das neben ihrer "normalen" redaktionellen Arbeit her machen müssen, sollen oder wollen, in die Lage zu versetzen, selbst zumindest brauchbare Alt-Texte zu erstellen. Die brauchen eine Anleitung; die brauchen Regeln.

Das kann doch die KI

Nein! Das kann sie nicht. Zumindest nicht gut und nicht im Moment.

Web Scraping?

Ein - vielleicht DAS zentrale - Problem bei KI und Alt ist das der Trainings-Daten. Irgendwo zwischen 90 und 99,9999999% aller existierenden Alt-Texte sind schlecht oder grottenschlecht.

1. Der Alt fehlt, obwohl das Bild sinntragend ist.
2. Der Alt ist vorhanden, obwohl das Bild dekorativ ist.
3. Für 10 hintereinanderstehende unterschiedliche Fotos ist 10 X derselbe Alttext vergeben.
4. Im Alt-Text wird die Bild-Unter- oder -überschrift wiederholt.
5. Der Alt lautet: "Bildbeschreibung für Blinde".
6. Im Alt steht das Copyright.
7. Und ... und ... und

Wenn Euch die Liste zu kurz ist, dann sprecht in der Caffee-Pause den nächsten blinden Menschen an und befragt ihn nach seinen diesbezüglichen Erfahrungen. Ich bin mir sicher, er kann sie deutlich verlängern. Ich hab hier aber nur 45 Minuten 😊

Bildbeschreibung?

Ja, klar! In der Mustererkennung und in der Bildbeschreibung gibt es erstaunliche Fortschritte / Ergebnisse! Die kann und will ich nicht in Abrede stellen.

Aber:

Ein Alt ist keine Bildbeschreibung!

Ein Alt ist kurz, weil alle Informationen aus der Bildbeschreibung gestrichen wurden, die im gegebenen Kontext nicht von Belang sind.

Und was ist mit der Bildbeschreibung?

- Auf Wunsch eine vollständige / interaktive Bildbeschreibung gemäß den Anforderungen oder Vorkenntnissen der jeweiligen Betrachter_in zu erstellen, ist Aufgabe des individuellen Benutzeragenten / Screenreaders. Die meisten Screenreader machen das auch schon. Früher oder später werden sie sicher auch lernen, dabei die Vorlieben ihrer Nutzer_in zu berücksichtigen.

Vermenschlichung

Wir sind gewohnt "KI" oder "AI" zu sagen und oft auch zu denken. Wir sprechen mit diesen Dingen wie mit unseren Kindern, Kolleg_innen oder Kumpelinnen. Und oft funktioniert das auch erstaunlich gut. Mit ein Bisschen Geduld und ein paar generellen Instruktionen kann Mensch sich da mit Maschine über Lagrange-Punkte, Luftschiffe der Zukunft oder Leihfahrräder in Linz unterhalten und darüber eine erquickliche Kanne Caffee leeren.

Nur dazu, dass sie ohne menschliches Zutun vernünftige Alt-Texte erstellen, kriegt Mensch sie eben nicht. Ich habe es versucht *seufz*.

Die Regeln

Es gibt vermutlich Millionen von Vorstellungen, wie ein guter Alt-Text zu sein hat. Viele davon sind sicher auch von der Art, dass Mensch Regeln daraus ableiten könnte, mit deren Hilfe die geneigte Redakteur_in Alt-Texte erstellen könnte. Auch ich habe solche Vorstellungen und ich habe sie sogar tatsächlich als Regeln formuliert und verwende diese als Schulungsmaterial.

Zwölf Regeln zum Erstellen von Alt-Texten

R1 - Immer

Jede Grafik hat eine Bedeutung! Jede Grafik braucht einen Alt-Text! Bereits die bloße Existenz/Sichtbarkeit einer Grafik rechtfertigt einen Alt-Text.

Ausnahme

- Hat eine Grafik im gegebenen Kontext keine Bedeutung / ist sie rein dekorativ, so ist sie mit geeigneten Mitteln vor Bildschirmleseprogrammen zu verstecken.

R2 - Logo

Der Alt-Text für ein Logo ist in der Form Logo: Kurztitel zu bilden.

Weitere Regeln sind nicht anzuwenden, außer das Logo ist gleichzeitig ein Icon.

Ausnahme

- Ist bei einem wirklich sehr kurzen Text in Folge der Position im Datenstrom oder anderer offensichtlicher Merkmale klar, dass es sich bei der Grafik mit hoher Wahrscheinlichkeit um ein Logo handelt, ist die Voranstellung des Typs verzichtbar.

R3 - Icon

Der Alt-Text für ein Icon ist in der Form Funktion (Kurzbeschreibung der Grafischen Darstellung) zu bilden. Die runden Klammern sind hierbei erheblich, um mögliche Abgrenzungsschwierigkeiten zwischen Funktion und Beschreibung auszuschließen.

Weitere Regeln sind nicht anzuwenden.

R4 - Typ, Subtyp, ...

Ist ein Alt-Text so umfangreich, dass er nicht binnen kürzester Zeit erfasst werden kann, ist die Voranstellung einer Einordnung mittels Typ, Subtyp usw. erforderlich. Diese Voranstellung kann aus mehreren Worten/Listeneinträgen bestehen und wird durch einen Doppelpunkt abgeschlossen.

Ausnahme

- Ist der Alt-Text bei Beachtung der übrigen Regeln ausformuliert, so kann auf die Einordnung mittels Typ, Subtyp usw. verzichtet werden.

R5 - Kontext

Der Alt-Text zu einer Grafik muss kontextabhängig erstellt werden. Er gibt mit seinen ersten Worten seine Bedeutung im Kontext an. Ein Stichwort aus einer möglichst kurzen Liste ist zu bevorzugen.

Eine Grafik, die außerhalb des Kontextes steht - z.B. eine Zierlinie - gilt als dekorativ und ist gemäß der Ausnahme zu r1 zu verstecken.

R6 - Eineindeutigkeit

Ein Alt-Text muss so verfasst sein, dass unter Zuhilfenahme des Kontextes und Alt-Textes ein Bild erstellt werden könnte, das an dieser Stelle dieselbe Funktion erfüllen würde, wie das "unbekannte" Original.

Dabei muss jede Information aus dem Alt ins Bild . Informationen wie Copyright oder Bildquelle gehören also nur dann in den Alt , wenn sie auch am sichtbaren Bild sind.

R7 - Metainformation

Metainformation, die im Kontext erwähnt wird, sonst aber unwichtig wäre, steht am Ende des Alt-Textes.

R8 - Neutral

Ein Alt-Text gibt ausschließlich offensichtliche Fakten wieder.

Ausnahmen

1. Wenn ein Bild dazu bestimmt ist, bei der Betrachter_in eine bestimmte Stimmung zu erzeugen, dürfen die Inhalte des Bildes auch vom Alt-Text interpretiert werden, sofern dies mit hinreichender Wahrscheinlichkeit genau die gewünschte Stimmung erzeugt und die Interpretation vom tatsächlich Dargestellten hinreichend gedeckt ist.
2. Wenn Informationen sich gezielt an #blinderMensch richten, und von #sehenderMensch regelmäßig nicht wahrgenommen werden sollen, so können diese in den Alt-Text aufgenommen werden. Auch Alt-Texte sind eine Form von versteckten Texten.

R9 - Kurz

Ein Alt-Text ist so kurz als möglich. Eine Ausformulierung ist entbehrlich.

Ergibt sich aus dem Kürzen des Alt-textes ein leerer Alt-Text, so ist die Ausnahme zu R1 erfüllt.

Ausnahmen

1. Ist es Sinn und Zweck der Grafik, bei der Betrachter_in eine bestimmte Befindlichkeit zu erzeugen - z.B. Stimmungsbild, so ist dies mit einem passenden Fließtext leichter, als mit einer Aufzählung von Fakten.
2. Ist ein ausformulierter Text kürzer / nicht wesentlich länger, als eine Auflistung von Fakten gleichen Informationsgehaltes, so steht der Redaktion die Entscheidung frei.

R10 - Massendaten

Große Datenmengen stehen nicht im Alt-Text. Sie sind außerhalb desselben verlinkt - z.B. mit einer Implementierung der Grafik als Link.

Stellt die Grafik die einzige Möglichkeit dar die veranschaulichten Massendaten zu erreichen, so darf sie nicht versteckt - ihr Alt-text also nicht leer sein.

R11 - Keine Redundanz

Trägt eine Grafik einen Titel, eine Bildüber- oder -unterschrift bzw. stellt sie ihrerseits eine Alternative zu einem vorhandenen Text dar, so ist der Informationsgehalt des angezeigten Textes im Alt-Text nicht zu wiederholen.

Führt die Beseitigung der Redundanzen zu einem leeren Alt, so ist die Ausnahme zu R1 erfüllt.

R12 - Strukturierung

Ein Alt-Text wird zur besseren Lesbarkeit bzw. zur besseren Verständlichkeit mittels Satzzeichen wie folgt gegliedert:

1. Der führende Typ, Subtyp usw. wird mit einem Doppelpunkt abgeschlossen.
2. Zusammengehörende kleine Informationsblöcke werden mit einem Semikolon ";" abgeschlossen.
3. Ein Satzpunkt schließt den Alt-Text ab und wird im Übrigen nicht verwendet.

Ausnahmen

1. Die Einleitung mit Typ / Subtyp ... unterbleibt, wenn der Alt-Text auch ohne diese Typisierung eindeutig / verständlich ist.
2. Ist der Alt-Text sehr kurz, kann die Strukturierung entfallen.
3. Ist der Alt-Text ausformuliert, kann die Zeichensetzung wie üblich erfolgen.

Der Darwin-Faktor

Die 12 Regeln funktionieren für Menschen ziemlich gut. Einerseits vermutlich deshalb, weil die Menschen, die sie verstehen und anwenden wollen bereits im Vorfeld den Kernsatz "Ein Alt-Text ist keine Bildbeschreibung" verstanden haben. Zum andern aber auch, weil ein Mensch assoziiert.

Das Experiment

Bevor wir es in die Luft sprengen können, brauchen wir erst einmal ein Labor.

Versuchsaufbau

Wir benötigen eine KI, eine Aufgabenstellung, ein Set von Testdaten, eine Messmethode und Geduld.

Die KI

Die KI meines geringsten Misstrauens ist Google Gemini. Nicht weil ich der Meinung bin, dass sie vom Ergebnis her besser sei als irgendeine andere, die ich mir angesehen habe, sondern schlicht deshalb, weil ihre Web-Oberfläche von allen von mir im Sommer 2025 getesteten Oberflächen die zugänglichste war. Müsste ich sie benoten, so würde sie

immerhin eine 4 bekommen, wohingegen alle anderen mit einer 5 aussteigen. Das und warum solche UI-Höllen an der Spitze dessen rangieren, was wir als absolute Hightech verstehen, wäre einen eigenen Vortrag wert. Für den Moment begnügen wir uns diesbezüglich mit einem verständnislosen Kopfschütteln.

Die Aufgabenstellung

Die ist kurz genug, um hier komplett wiedergegeben zu werden:

3. Deine Aufgabe ist es, Alternativtexte für Bilder zu generieren. Die allgemeinen Anweisungen dafür kannst Du den "elfeinhalb Instruktionen des Käpt'n Weißbart" entnehmen.
4. In jedem folgenden Durchgang lade ich dir ein Bild und mindestens eine Beschreibung des Kontextes hoch, in dem das Bild verwendet werden soll.
5. In manchen Datensätzen ist zusätzlich ein Titel enthalten. Dieser wird im entsprechenden Medium (PDF, ePub, HTML, ...) zusammen mit dem Bild angezeigt.
6. Die einzelnen Datensätze stehen in keinem inneren Zusammenhang. Sie beziehen sich also in keiner Weise aufeinander und auch die Reihenfolge oder die Benennung der Datensätze sind nicht von Belang. Behandle also jeden einzelnen Satz von Dateien so, als ob Du die anderen nicht kennen würdest.

Testdaten

Zu viele Test-Datensätze konnten wir in unser Experiment nicht aufnehmen, da der Vergleich des KI-Ergebnisses mit einem "guten" - also von mir erstellten - Alt manuell erfolgen musste.

Darüber hinaus muss jeder Testdatensatz folgende Bedingungen erfüllen:

- Der Kontext muss einschließlich allfällig vorhandener Bild-Unter- oder -überschriften vorliegen und beschrieben sein.
- Das Bild muss vorhanden und verwendbar sein.

Außerdem sollten die Test-Datensätze alle vorhandenen Regeln / Instruktionen adressieren.

Ich habe insgesamt 7 Datensätze aufgenommen. Auch diese Testdaten kann ich auf Wunsch zur Verfügung stellen.

Anmerkung zum Kontext

In einer späteren praktischen Anwendung z.B. innerhalb eines CMS muss die KI den Kontext natürlich selbst erkennen. Aber das setzt korrektes Markup voraus und darum

ist es in der Praxis ja leider fast so schlecht bestellt, wie um die Alt-Texte. Für die verwendeten Datensätze habe ich also den jeweiligen Kontext vorgegeben.

Messmethode

Die Messmethode war das Schwierigste. Einerseits gaben weder Zeit noch Budget einen Reihentest mit blinden Proband_innen her und andererseits - hmhm - ja, wie sage ich das jetzt - sind mir offen gestanden auch zu viele Nutzer_innen von Bildschirmleseprogrammen mit schlechten Alt-Texten zufrieden. Aber meine Grund-These "Ein Alt-Text ist keine Bildbeschreibung" steht hier jetzt gerade nicht zur Debatte.

Wenn aber kein empirischer Massentest, dann brauchen wir eine zumindest in Ansätzen objektivierbare Bewertungsmethode.

Zu der kommen wir dann später.

Geduld.

Geduld habe ich gleich mehrfach benötigt. So bin ich z.B. erst nach Durchführung von Phase 3 dahinter gekommen, dass Gemini - entgegen meiner ausdrücklichen Anweisungen - auf den Verlauf der vorhergehenden Phasen zurückgreift und ich die eigentlichen Versuche in jeweils neuen leeren temporären Chats durchführen muss, um zumindest in Ansätzen eine Vergleichbarkeit der Durchgänge herzustellen.

Die Phasen

Es gibt insgesamt 7 Phasen - 1 bis 6 und A - die protokolliert und bewertet sind. Dazu kommt noch eine fiktive Phase 0, in der ich KI gespielt und die Alt-Texte erstellt habe.

Nach jedem Durchgang habe ich die Instruktionen und/oder die Versuchsparmeter geändert. Das gehen wir hier jetzt aber nicht im Einzelnen durch, wer die Daten haben will, kann sich gerne bei mir melden.

Nur die Phase minus 10 ist undokumentiert. Da habe ich einfach die 12 Regeln aus meinen Schulungsunterlagen zusammen mit etlichen Bildern in Gemini eingefüllt. Dem daraus entstandenen Desaster habt Ihr zu verdanken, dass ich jetzt hier stehe und mich wichtig mache.

Vorbemerkungen zu den Instruktionen

Elfeinhalb

Eigentlich waren es erst nur elf Instruktionen, mir ist aber aufgefallen, dass es nicht für alle selbstverständlich zu sein scheint, dass Instruktionen, die mit I1 bis I11 benannt sind, auch in der Reihenfolge I1, I2, I3 usw. anzuwenden sind bzw. dass der Satz "in diesem Fall sind keine weiteren Regeln mehr anzuwenden" auch tatsächlich bedeutet, dass weiter unten stehende Regeln nicht mehr anzuwenden sind. Die KI meines

geringsten Misstrauens hat sich da hervorgetan und versucht bisweilen, Logos mit vorangestellten Bildtypen zu versehen oder Icons mittels Satzzeichen zu gliedern.

Vorletzte Fassung

Wir verwenden die Instruktionen in ihrer vorletzten Fassung. Die vorletzte Fassung deshalb, weil ich mich auf Anraten eines Kollegen und von Gemini selbst bei der letzten Fassung davon gelöst habe, halbwegs menschliches Deutsch zu schreiben und stattdessen die KI wie eine Maschine anzusprechen. Im Fachjargon heißt das "promptifizieren" und bringt offensichtlich im Rahmen dieser Anforderungen Garnix.

Die elfeinhalb Instruktionen des Käpt'n Weißbart

I0 - Leidfaden

Ein Alt-Text ist KEINE Bildbeschreibung!

- Ziel des Alt ist es, die Bedeutung des Bildes in seinem Kontext anzugeben und so der Betrachter_in eine rasche Entscheidung zu ermöglichen, ob eine genauere Kenntnis des Bildes / eine echte Bildbeschreibung, im Moment erforderlich oder erwünscht ist.
- Auf Wunsch eine vollständige / interaktive Bildbeschreibung gemäß den Anforderungen oder Vorkenntnissen der jeweiligen Betrachter_in zu erstellen, ist Aufgabe des individuellen Benutzeragenten / Screenreaders.

Erfolgskriterien

- Einer versierten Bildbearbeiter_in muss es möglich sein, mit Hilfe des Alt für den gegebenen Kontext ein Bild zu erstellen, das dieselbe Funktion erfüllen würde wie das unbekannte Original. Dabei gilt das Gebot der Funktionsgleichheit, nicht das der Informationsgleichheit. Eine Tortengrafik mit geänderten Farben oder abweichenden Zahlen ist Funktionsgleich zum Original; Ein Balkendiagramm mit denselben Farben und Werten hingegen ist nicht Funktionsgleich.
- Auch ein frei erstellter Alt, muss die in den Instruktionen I1 bis I11 aufgestellten Anforderungen erfüllen.

Systematik

Durch konsequente Anwendung der folgenden Instruktionen in der hier vorgegebenen Reihenfolge kann aus einer Bildbeschreibung ein sinnvoller Alt-Text für den gegebenen Kontext erzeugt werden

- Zur Konsequenten Anwendung gehört ausdrücklich auch, dass bei Erreichen der Anweisung, dass keine weiteren Instruktionen mehr anzuwenden seien, dies auch tatsächlich unterbleibt. Und zwar auch implizit. Das

Vorhandensein eines Begriffes in der beigegebenen Liste darf also z.B. keinerlei Rolle spielen, wenn es keine Voranstellung gibt.

I1 - Kontext

Der Alt-Text zu einer Grafik muss kontextabhängig erstellt werden; das bedeutet ausdrücklich auch:

- Der Alt ist in derselben natürlichen Sprache wie der Kontext verfasst.
- Der Kontext ist nach den Regeln der Zugänglichkeit zu ermitteln (Direkte textalische Umgebung / umgebende Region / übergeordnete Überschriften oder Titel).
- Eine Grafik, die außerhalb des Kontextes steht - z.B. eine Zierlinie oder ein Hintergrundbild - gilt als dekorativ.
- Dekorative Grafiken erhalten keinen Alt-Text. weitere Instruktionen sind in diesem Fall nicht anzuwenden.

I2 - Icon

Der Alt für eine grafische Schaltfläche ist in der Form "Beschreibung der Funktion (Beschreibung der Grafischen Darstellung)" zu bilden.

- Die Beschreibung der Funktion ist dabei maximal zu kürzen.
- Die Beschreibung des Aussehens ist so weit zu kürzen, dass sie im Kontext eineindeutig bleibt. Daher sind spezifischere Bezeichnungen wie Eigennamen gegenüber weniger spezifischen Begriffen wie Gattungsnamen vorzuziehen.
- Die runden Klammern sind erheblich, um mögliche Abgrenzungsschwierigkeiten zwischen Funktion und Grafischer Darstellung auszuschließen.

Im Falle einer grafischen Schaltfläche sind keine weiteren Instruktionen anzuwenden.

I3 - Logo

Der Alt-Text für ein Logo ist in der Form "Logo: Kurztitel" zu bilden.

Ausnahme

- Ist bei einem sehr kurzen Alt in Folge der Position im Datenstrom oder anderer offensichtlicher Merkmale klar, dass es sich bei der Grafik um ein Logo handelt, ist die Voranstellung von "Logo:" unerwünscht. Dies ist regelmäßig am Beginn einer Seite eines Portales der Fall.

Im Falle eines Logos sind keine weiteren Instruktionen anzuwenden.

I4 - Stimmungsbild

Besteht der Sinn und Zweck einer Grafik nicht aus der Informationsvermittlung, sondern daraus, bei der Betrachter_in eine bestimmte Befindlichkeit zu erzeugen, so handelt es sich um ein Stimmungsbild.

Alt-Texte für Stimmungsbilder sind ausformuliert und dürfen das Bild und seine Wirkung beschreiben.

- Eine Interpretation des dargestellten ist dabei statthaft, soweit sie die beabsichtigte Wirkung unterstützt.
- Die Hintergründigkeit der Beeinflussung durch das Bild muss auch vom Alt gewahrt werden. Formulierungen wie "Das Stimmungsbild zeigt ..." sind daher in jedem Fall zu vermeiden.

Im Falle eines Stimmungsbildes sind keine weiteren Instruktionen anzuwenden.

I5 - Relevanz

Ein Alt-Text ist so kurz wie möglich.

- In den Alt-Text werden aus der Bildbeschreibung nur Objekte übernommen, die im gegebenen Kontext relevant sind. Die Größe der Anzeige hängt also nur lose mit der Länge des Alts zusammen.

I6 - Redundanzvermeidung

Trägt eine Grafik einen Titel, eine Bildüber- oder -unterschrift bzw. stellt sie ihrerseits eine Alternative zu einem in unmittelbarer Nähe vorhandenen Text dar, so ist der Informationsgehalt des angezeigten Textes im Alt-Text nicht zu wiederholen.

Sonderfälle

- Verschwindet der Alt-Text durch die Redundanzvermeidung komplett und kann leer bleiben, so gilt die Grafik selbst als dekorativ. Der allfällig zugeordnete Titel der Illustration bleibt dennoch sichtbar. Weitere Instruktionen sind in diesem Fall nicht anzuwenden.
- Verschwindet der Alt-Text durch die Redundanzvermeidung komplett, kann aber nicht leer bleiben, da im umgebenden Text explizit auf das Bild Bezug genommen wird, so ist in der Regel eine redaktionelle Überarbeitung des umgebenden Textes angezeigt. Ist dies - z.B. wegen automatisierter Erstellung der Alt-texte - nicht möglich, so ist ein formaler Alt-Text wie "(Bild)" akzeptabel. Die runden Klammern sind dabei relevant, um klarzustellen, dass es sich nicht um ein Bild eines Bildes, sondern um einen Platzhalter handelt.

Weitere Instruktionen sind in diesen Sonderfällen nicht anzuwenden.

I7 - Voranstellung

Ist ein Alt-Text so umfangreich, dass er nicht binnen kürzester Zeit erfasst werden kann, ist eine Einordnung mittels Voranstellung erforderlich.

- Es ist sinnvoll, für den jeweiligen Zuständigkeitsbereich eine entsprechende Liste zu führen und Einträge aus dieser Liste bevorzugt zu verwenden. Ein kommentiertes Beispiel einer solchen Liste ist den Instruktionen beigegeben.
- Eigenschaften, die durch die Voranstellung bereits klargestellt / erwähnt sind, werden im weiteren Alt nicht wiederholt.

I8 - Visualisierungen

In Alt-Texten für Visualisierungen wie Diagrammen oder Skizzen werden die Art der Visualisierung und gegebenenfalls deren wichtigste Aussagen beschrieben. Die zu Grunde liegenden Massendaten oder Beschreibungen sind durch die Grafik verlinkt.

Ausnahmen:

- Befinden sich die visualisierten Fakten in unmittelbarer Umgebung der Visualisierung, kann auf die Verlinkung verzichtet werden. Für die Ermittlung der "unmittelbaren Umgebung" ist nicht das visuelle Layout, sondern das DOM bzw. der Accessibility-Tree maßgeblich.
- Liegen die visualisierten Fakten nicht in zugänglicher Form vor, so ist diese Tatsache explizit im alt-Text zu vermerken. Dieser Pflicht ist z.B. mit einer Voranstellung wie "Explosionszeichnung" genüge getan, weil der Sinn und Zweck einer Explosionszeichnung mit keiner wie auch immer gearteten textlichen Darstellung erreicht werden kann. Eine Verlagerung dieser Information in die Zugänglichkeitserklärung ist nicht ausreichend.

I9 - Kompaktifizierung

Ein Alt-Text ist so kurz als möglich. Eine Ausformulierung ist entbehrlich.

- Bildteile / Objekte werden in Informationsblöcken schlagwortartig beschrieben.
- Spezifischere Schlagworte sind Gattungsbegriffen vorzuziehen, sofern sie von den erkennbaren Fakten gedeckt sind. Also z.B. Sacko statt Jacke oder Fischweib statt frau.
- Die einzelnen Informationsblöcke sind nach der Logik des Bildes in absteigender Bedeutung geordnet. Also z.B. für einen Mann, der eine grüne Jacke trägt: "Mann; grüne Jacke".

Ausnahme

- Sind die Beziehungen zwischen den einzelnen Objekten / Informationsblöcken zu komplex und eine hierarchische Ordnung daher nicht herstellbar, so kann auf die Kompaktifizierung verzichtet werden.

I10 - Metainformation

Metainformation (Größe / Position / Format) des Bildes sind nicht in den Alt-Text aufzunehmen.

Ausnahme

- Metainformation, die im Kontext explizit erwähnt wird, ist in den Alt-Text aufzunehmen.
- Informationen, die sich gezielt an blinde bzw. hochgradig sehbehinderte Menschen richten, sind wie relevante Metainformationen zu behandeln.

I11 - Aufbereitung für das Bildschirmleseprogramm

Der Alt-text kann nicht durch geeignete HTML-Elemente formatiert oder gegliedert werden, sondern wird als ein einziger Textblock präsentiert. Ist ein Alt-Text kompaktifiziert (I9) so wird er daher zur besseren Lesbarkeit bzw. zur besseren Verständlichkeit wie folgt gegliedert:

1. Eine allfällig vorhandene Voranstellung wird durch einen Doppelpunkt abgeschlossen.
2. Informationsblöcke werden mit einem Semikolon voneinander getrennt.
3. Ein Satzpunkt schließt den Alt-Text ab und wird im Übrigen nicht verwendet.
4. Relevante Metainformation (I10) wird in Runde Klammern eingeschlossen und ans Ende des Alt-Textes - hinter den Satzpunkt - gestellt.

Voranstellungen

Sowohl in den 12 Regeln wie auch in den elfeinhalb Instruktionen ist von einer Liste mit Voranstellungen die Rede. Eine solche Liste mit empfohlenen Voranstellungen für Alt-Texte ist für jede Organisationseinheit zu empfehlen und sollte für diese ebenso zur Corporate Identity gehören, wie ein CSS oder eine Formatvorlage.

Hinweise zur Lesbarkeit und Verwendung

- Wo erforderlich ist in der Liste dem jeweiligen Begriff nach einem Doppelpunkt eine Erläuterung nachgestellt.
- Im Zweifel ist der spezifischere Begriff zu wählen. "Portrait" ist spezifischer als "Detailsicht" und diese wiederum spezifischer als "Bebildung".

- Wird im Rahmen der Erstellung von Alt-Texten häufig eine Voranstellung gewählt, für die es in der Liste bereits spezifischere Einträge gibt - z.B. "Bebilderung", so ist dies ein Hinweis, dass die Liste um eine passende spezifischere Voranstellung erweitert werden sollte.

Voranstellungsliste

1. Bebilderung: Eine Grafik geringer Relevanz. Sie dient mehr dem Blickfang als der Information.
 - 1.1. Detailsicht
 - 1.1.1. Portrait: Hier ist kein Portrait im Sinne der Malerei gemeint, sondern ein Bild, das sich auf ein bestimmtes Objekt fokussiert.
 - 1.1.2. Pressefoto
 - 1.1.3. Stillleben
 - 1.2. Stimmungsbild: Dies wird (fast) nie in einem Alt-Text auftauchen, da Alt-Texte für entsprechende Bilder in der Regel ausformuliert werden und daher keine Voranstellung erhalten.
 - 1.3. Szenerie
 - 1.3.1. Vortragssituation
2. Visualisierung: Eine Grafik sehr hoher Relevanz
 - 2.1. Balkendiagramm
 - 2.2. Bildschirmfoto
 - 2.3. Plan
 - 2.3.1. Explosionszeichnung
 - 2.3.2. Lageplan
 - 2.3.3. Landkarte
 - 2.4. Tortendiagramm
 - 2.5. Verlaufsdiagramm
3. Werk: Eine Grafik, deren Informationsgehalt mit der Art ihrer Entstehung zusammenhängt
 - 3.1. Fotomontage
 - 3.2. Gemälde

3.3. Kinderzeichnung

Gültigkeit

Diese Liste ist nicht auf Vollständigkeit angelegt. Sie dient lediglich der Erläuterung der Regel "R4 - Typ, Subtyp, ..." bzw. der Instruktion "I7 - Voranstellung".

Regeln bestätigen die Ausnahmen

Alte Hacker

Die elfeinhalb Instruktionen sind möglicherweise nicht vollständig - die Liste der Voranstellungen mit Sicherheit nicht! Ich behaupte auch nicht, dass es sich um den einzig möglichen Satz von Instruktionen handelt, der zu einem brauchbaren Alt-Text führt. Es ist aber ein Satz von Instruktionen, von dem ein alter Hacker behauptet, dass sie von einem Roboter ausgeführt werden können sollten, ohne zu einem Fiasko zu führen.

Sir Internet als tückische Ausnahme

Nehmen wir ein Bild von Sir Tim Berners-Lee und sei es nur deshalb, weil ein Vortrag über das Internet ohne ihn unvollständig wäre. Da aber regelmäßig damit zu rechnen ist, dass es Leute gibt, die nicht wissen, wer das ist, schreiben wir seinen Namen als Untertitel drunter, damit die Unwissenden wenigstens in der Wikipedia nachschlagen können.

Das Beispiel

Zum 25sten Jahrestag des Ritterschlages könnte ein kurzer Artikel erscheinen:

Tim Berners-Lee (Siehe Bild rechts) wurde am 16.07.2004 von Königin Elisabeth II. zum Ritter geschlagen.



Sir Tim Berners-Lee

Hier hat eine besonders EIFRIGE Redakteur_in als Service für die Leser_innen nach dem Namen den Einschub "(siehe Bild rechts)" eingefügt.

Das Problem

- Im gegebenen Kontext ist ausschließlich relevant, dass Tim Berners-Lee abgebildet ist.

- Aber genau der Name fällt der Redundanzvermeidung zum Opfer, weil er bereits im Untertitel des Bildes steht.

Damit hätten wir einen leeren Alt / ein dekoratives Bild, womit ich zufrieden wäre.

Jetzt hat allerdings unsere Redakteur_in den Verweis "(siehe Bild rechts)" in den Text geschrieben. Das führt uns in eine Meta-Falle:

- Die explizite Erwähnung des Bildes zwingt uns, die Metainformation "(Bild rechts)" in den Alt-Text aufzunehmen.
- Das ändert am Umstand, dass das Bild dekorativ ist aber eigentlich nichts.

Ein Dilemma!

Selbst wenn wir diese Metainformation in runde Klammern setzen und davon ausgehen, dass die Leser_in Braille liest, ist der Alt "(Bild rechts)" ein unbefriedigendes Ergebnis.

Double-Check

Ist dieser unsinnige Alt eine Folge schlechter Instruktionen?

Für heiße Debatten über diese Frage steht bei mir immer eine Kiste alkoholfreie Ziegenmilch unter dem Schreibtisch bereit!

Das Erste, mit dem ich jede Besucher_in beim ersten Ziegenmilch-Plopp konfrontiere ist allerdings folgendes:

- Nein! Der Alt "(Bild rechts)" ist die Folge schlechter Redaktion. Der Einschub "Siehe Bild rechts" gehört da einfach nicht hin. Und genauso, wie das wohlgemeinteste ARIA-Label Fehler bei der Entwicklung einer App nicht korrigieren kann, kann auch der beste Alt-Text - ob von KI oder Mensch formuliert - achtlose Redaktion nicht ausgleichen.
- Danach können wir über Wichtiges reden: Pizza oder Burger?

Alt-Texte bewerten

Natürlich ist und bleibt es eine subjektive Angelegenheit, Texte jedweder Art auf ihre Angemessenheit zu überprüfen und das gilt auch für Alternativ-Texte. Um der allgemeinen - und meiner - Willkür aber zumindest ein paar Orientierungen zu geben, wären ein paar Bewertungshilfen schon nützlich.

Da mein Chef gerne sieht, wenn sich ein Bewertungssystem an österreichische Schulnoten anlehnt, ist es kein reiner Zufall, dass ich genau 5 zu erfüllende Kriterien formuliert habe. Auch für das Ansinnen, diese Bewertungskriterien zu verbessern, gilt das Angebot mit der Ziegenmilch.

Die Übersicht

1. Bild
2. Funktion
3. Vollständigkeit
4. Kürze
5. Format

#1 Bild

Wurde der Inhalt des Bildes richtig erkannt?

Das ist natürlich die Grundvoraussetzung dafür, ein Bild überhaupt zu beschreiben

Es ist aber auch für die Bewerter_in die Chance, einen völlig unangemessenen Alt komplett zu verwerfen. Sollte also irgendwo in einem Alt das Copyright eingetragen sein oder ein Zeppelin beschrieben werden, wo ein Apfel abgebildet ist, wird dieser Punkt nicht gegeben und es gibt noch nicht mal eine 5 "Nicht angetreten / keine erkennbare Leistung".

Im Erfolgsfalle gibt's den ersten Punkt und ist einem mindestens mal die 5 sicher.

#2 Funktion

Wurde aus Kontext und Bildinhalt die korrekte Funktionalität des Bildes an dieser Stelle abgeleitet?

Auch das ist in dem Sinne ein K.O.-Kriterium, dass eine weitere Prüfung oder bessere Note bei Verfehlung an dieser Stelle nicht möglich ist. Allerdings kann Mensch hier auch schon über eine Abkürzung zu einer 1 gelangen, wenn nämlich ein dekoratives Bild als solches erkannt und folgerichtig mit einem leeren Alt versehen wird; Die weiteren Kriterien sind dann nämlich nicht mehr anwendbar.

Zumindest ist einem bei Bestehen dieses Teilschrittes der zweite Punkt und damit eine 4 sicher und es werden alle übrigen Prüfschritte durchlaufen.

Sollte hingegen ein Bild, das im gegebenen Kontext eine Bedeutung besitzt als dekorativ erkannt werden, so bleibt es bei einer 5.

#3 Vollständigkeit

Enthalten unmittelbar umgebender Text, Titel und Alt-Text gemeinsam alle nötigen Angaben zum Verständnis des Bildes?

Wenn ich mich im vorigen Schritt entschieden habe, dem Bild in diesem Kontext eine Bedeutung zuzuordnen, dann muss es natürlich die richtige Bedeutung sein.

Es ist nämlich ein durchaus bedeutsamer Unterschied, ob ein fallender Apfel gleich den Schläfer unter dem Baum auf die Nase treffen oder diesem in die offene Hand fallen wird. Hier ist auch die Beobachtungsgabe gefragt. Ist der Aufschlagpunkt des Apfels denn aus dem Bild wirklich ableitbar? Denn nur dann darf diese Tatsache als solche in den Alt-Text aufgenommen werden.

Bei Bestehen dieses Schrittes haben wir mal mindestens 3 Punkte.

#4 Kürze

Ist der Alt frei von überflüssigen Informationen?

Gänseblümchen sind schön, aber wenn sie im gegebenen Kontext keine Rolle spielen, gehören sie nicht in den Alt. Aber selbst dann, wenn sie eine Rolle spielen, sind sie kaum so wichtig, dass sie mehrfach erwähnt werden müssen. Wenn also die Bild-Unter- oder -überschrift bzw. der unmittelbar umgebende Text bereits das Wort "Gänseblümchen" enthalten, dann nervt eine redundante Erwähnung im Alt nur.

Wer nicht nervt, kann sich den nächsten Punkt notieren und hat im besten Falle schon eine 2.

#5 Format

Ist der Alt in angemessener Weise formuliert? Ist der Alt durch angemessene Zeichensetzung für die Ausgabe durch TTS geeignet?

Bis hier her wurde der Informationsgehalt des Alt-Textes bewertet. Um auch den Punkt für das Format zu bekommen und die Chance auf eine 1 zu wahren, müssen wir jetzt noch die Übertragung der Information an die Nutzer_in optimieren.

Eine Formatierung im Sinne von Markup irgendeiner Art ist innerhalb des Alt-Textes nicht vorgesehen. Folglich bleibt uns nur die klassische Gliederung durch Reihenfolge und Satzzeichen.

Also:

- Das Wichtigste zuerst.
- Keine komplexen Satzstrukturen oder Bezüge.

Manchmal fällt einem die 1 auch in den Schoß. Viele Alt-Texte bleiben nämlich zurecht in Fließtext oder sind spätestens nach Entfernung der Redundanzen von sich aus sehr kurz. Auch in diesem Fall darf sich Mensch den Punkt natürlich notieren.

Abdingbarkeit

Ob ein Alt-Text nun ganz frei, gemäß der 12 für Menschen gedachten Regeln oder gemäß der elfeinhalb Instruktionen für die KI angefertigt wird oder ob jemand noch eine viel

bessere Idee hat, ist nachrangig. Aber folgenden Satz werde ich gegebenenfalls mit Klauen und Zahnersatz verteidigen:

- Ein Alt-Text ist eine leicht verständliche, korrekte und vollständige Beschreibung der Bedeutung eines Bildes in einem gegebenen Kontext in aller Kürze.

Zusammenfassung

"Wat mutt, dat mutt"

Um Alt-Texte als solche kommen wir dank EAA/WCAG nicht herum.

Alle Bilder als dekorativ zu kennzeichnen, ist kein gangbarer Weg, würde das doch bedeuten, dass es kein sinntragendes Bild gibt und so die Existenzberechtigung der gesamten Werbegrafik-Branche in Frage stellen.

"Die KI kann alles"

Wenn wir die hier behandelten Instruktionen verwenden, dann kann uns die KI zumindest davor bewahren, mit den WCAG in Konflikt zu geraten. Denn dort steht Nichts über die Qualität eines Alt-Textes.

"Pixel-Perfekt-Kultur"

Ob eine Redaktions-Mitarbeiter_in nun weiß, was ein Screenreader ist oder auch nur, dass es für #blinderMensch total normal ist, am Computer zu arbeiten, ist nachrangig. Mit Hilfe der dargestellten 12 Regeln ist es auch ohne dieses Wissen machbar, gut brauchbare Alt-Texte zu erstellen. Mit Vorlage durch die KI und den elfeinhalb Instruktionen ist es oft leichter, weil viele Formalia von ihr bereits erfüllt sind.

Letztlich zeigt sich unsere Wertschätzung gegenüber unseren blinden Kund_innen aber darin, wieviel Zeit wir unserem Redaktionsteam für die Erstellung - oder Überprüfung - ansprechender Alt-Texte zur Verfügung stellen. Das ist ein ganz ähnlicher Unterschied, wie zwischen verwackelten Handyfotos und richtig schönen Bildern.

Noch kürzer

Mensch kann sein Marketing in Vorarlberg der KI überlassen.

Aber Mensch muss nicht.

Impressum

Jürgen Schwingshandl (Er/Du)

Berater für digitale Zugänglichkeit und inklusive Kulturvermittlung

E-Mail: juergen.schwingshandl@beko.at



[Beko Solutions](#)



[Mitglied von IAAP](#)



[Inhaber von CPACC](#)