AX WEBINAR: VON DATEN ZU TEXT

HERZLICH WILLKOMMEN!

Lisa Weidmann | lisa.weidmann@ax-semantics.com

WAS PASSIERT DAZWISCHEN?

DATEN → **TEXTKONZEPT**

WICHTIGE FRAGEN

DATEN:
WIE VERWERTE ICH DIE INFORMATIONEN DER DATEN IM TEXT?

TEXTKONZEPT:
WELCHE INHALTE SOLLEN DIE TEXTE WIEDERGEBEN?

SCHRITT 1: DATENÜBERBLICK

> Taxonomie über alle Datenfelder erstellen (Duplikate entfernen)

PRODUKTTYP	MATERIAL
Bundfaltenhose	Bio-Baumwolle
Strickjacke	Kaschmir|Angora
T-Shirt	Baumwolle
Pullover	Kaschmir
Pullover	Kaschmir
Jeans	Baumwolle
Strickjacke	Bio-Baumwolle
T-Shirt	Baumwolle

→

PRODUKTTYP	MATERIAL
Bundfaltenhose	Bio-Baumwolle
Strickjacke	Kaschmir|Angora
T-Shirt	Baumwolle
Pullover	Kaschmir
Jeans	

> **Ergebnis:** Überblick über die Datenbasis

VERWERTBARKEIT DER DATEN

> **Inhaltliche Relevanz**

Welche Datenfelder bieten einen **informativen Mehrwert**?

Lässt sich das **Informationsbedürfnis** des Lesers damit stillen?

> **Befüllungsgrad**

Sind die wichtigen Datenfelder **zuverlässig befüllt**?

Welche Fallback-Lösungen benötige ich im Textkonzept, um **Datenlücken abzufangen**?

> **Einheitlichkeit**

Sind gleiche **Inhalte immer gleich** gepflegt?

Wie viele unterschiedliche **Ausprägungen** hat ein Datenfeld?

INHALTLICHE RELEVANZ

Inhaltliche Relevanz

Welche Datenfelder bieten einen **informativen Mehrwert**?

Lässt sich das **Informationsbedürfnis** des Lesers damit stillen?

> Prüfen, ob die Dateninhalte die wichtigsten W-Fragen beantworten (Was? Wer? Wozu? Wann? Wo? Wie? Warum?)

> Daten im Training anreichern, um weitere Aussagen abzuleiten → Aussagen mit Mehrwert treffen

> Erwähnenswerte Inhalte aus den Daten herausfiltern und gezielt für die Textgenerierung verwenden

INHALTLICHE RELEVANZ:
RELEVANTE DATEN DURCHREICHEN UND NENNEN

Diese Bundfaltenhose in Schwarz ist schmal geschnitten und besitzt einen Komfortbund. Gefertigt ist sie aus Bio-Baumwolle.

PRODUKTTYP	FARBE	SCHNITT	BUNDART	MATERIAL
Bundfaltenhose	schwarz	schmal	Komfortbund	Bio-Baumwolle
Strickjacke	rot	lässig		Kaschmir|Angora

Diese Strickjacke in Rot ist lässig geschnitten. Gefertigt ist sie aus Kaschmir und Angora.

INHALTLICHE RELEVANZ:
RELEVANTE DATEN DURCHREICHEN UND NENNEN

Diese Bundfaltenhose in Schwarz ist schmal geschnitten und besitzt einen Komfortbund. Gefertigt ist sie aus Bio-Baumwolle.

PRODUKTTYP	FARBE	SCHNITT	BUNDART	MATERIAL
Bundfaltenhose	schwarz	schmal	Komfortbund	Bio-Baumwolle
Strickjacke	rot	lässig		Kaschmir|Angora

Diese Strickjacke in Rot ist lässig geschnitten. Gefertigt ist sie aus Kaschmir und Angora.

INHALTLICHE RELEVANZ:
RELEVANTE DATEN DURCHREICHEN UND NENNEN

Diese Bundfaltenhose in Schwarz ist schmal geschnitten und besitzt einen Komfortbund. Gefertigt ist sie aus Bio-Baumwolle.

PRODUKTTYP	FARBE	SCHNITT	BUNDART	MATERIAL
Bundfaltenhose	schwarz	schmal	Komfortbund	Bio-Baumwolle
Strickjacke	rot	lässig		Kaschmir|Angora

Diese Strickjacke in Rot ist lässig geschnitten. Gefertigt ist sie aus Kaschmir und Angora.

INHALTLICHE RELEVANZ:
UM DIE ECKE DENKEN & INFORMATIONEN ABLEITEN

Diese Bundfaltenhose in Schwarz ist schmal geschnitten und besitzt einen Komfortbund. Damit ist Ihnen einen angenehmes Tragegefühl im Büroalltag sicher. Gefertigt ist sie aus nachhaltig produzierter Bio-Baumwolle.

PRODUKTTYP	FARBE	SCHNITT	BUNDART	MATERIAL
Bundfaltenhose	schwarz	schmal	Komfortbund	Bio-Baumwolle

INHALTLICHE RELEVANZ:
UM DIE ECKE DENKEN & INFORMATIONEN ABLEITEN

> Mut zur Lücke

MATERIAL	HÄUFIG-KEIT:	
Bio-Baumwolle	256	→ Aus nachhaltig produzierter Bio-Baumwolle
Kaschmir|Angora	130	→ Aus kuschelig weichem Kaschmir und wärmender Angora
Polyester	400	*→ Kein hochwertiges Material, daher Möglichkeit, bewusst nicht anzureichern*
Seide	8	*→ Nur wenig Vorkommnisse: abwägen, ob ein Anreichern lohnt*

Möglichkeiten:

> Häufigste Vorkommnisse betexten
> Herausragende Merkmale betexten
> Negativ behaftete oder neutrale Merkmale bewusst weglassen

BEFÜLLUNGSGRAD

Befüllungsgrad

Sind die wichtigen Datenfelder **zuverlässig befüllt**?

Welche Fallback-Lösungen benötige ich im Textkonzept, um **Datenlücken abzufangen**?

> Mindestanforderungen definieren → Datenfelder definieren, die unbedingt nötig sind, damit ein sinnvoller Text oder Satz generiert werden kann (z. B. Produkttyp)

> Überblick erlangen, welche Datenfelder wie gut befüllt sind → Auswirkung auf den Textaufbau

> Sätze modular gestalten, sprich so aufbauen, dass Teile bei fehlenden Dateninhalten wegfallen können

BEFÜLLUNGSGRAD: TEXTAUFBAU PLANEN

> Reihenfolge der Inhalte festlegen
> Prüfen, welche der Inhalte wegfallen können → Auswirkung auf den Textaufbau

Beispiel: Textaufbau

Satz 1: PRODUKTTYP, FARBE und SCHNITT → Alle 3 Datenfelder immer vorhanden: Satz erscheint immer

Satz 2: BUNDART → Datenfeld fehlt bei einem Produkt: Satz muss wegfallen können

Satz 3: MATERIAL → Datenfeld immer vorhanden: Satz erscheint immer

PRODUKTTYP	FARBE	SCHNITT	BUNDART	MATERIAL
Bundfaltenhose	schwarz	schmal	Komfortbund	Bio-Baumwolle
Strickjacke	rot	lässig		Kaschmir|Angora

BEFÜLLUNGSGRAD: TEXTAUFBAU PLANEN

Textkonzept:

Satz 1

Satz 2

Satz 3

Produktdaten:

Bundfaltenhose →

Strickjacke →

Produkttexte:

Satz 1
Satz 2
Satz 3

Satz 1
Satz 3

BEFÜLLUNGSGRAD: SATZAUFBAU PLANEN

> Prüfen, welche der Inhalte wegfallen können
> Satz so konstruieren, dass er in allen Fällen eine korrekte Ausgabe erzeugt

Beispiel: Satz soll folgende Datenfelder beinhalten

PRODUKTTYP	→ Datenfeld immer vorhanden: Inhalt erscheint immer im Satz
FARBE	→ Datenfeld immer vorhanden: Inhalt erscheint immer im Satz
SCHNITT	→ Datenfeld immer vorhanden: Inhalt erscheint immer im Satz
BUNDART	→ Datenfeld nicht immer befüllt: Inhalt muss im Satz wegfallen können

PRODUKTTYP	FARBE	SCHNITT	BUNDART	MATERIAL
Bundfaltenhose	schwarz	schmal	Komfortbund	Bio-Baumwolle
Strickjacke	rot	lässig		Kaschmir|Angora

BEFÜLLUNGSGRAD: SATZAUFBAU PLANEN

Diese Bundfaltenhose in Schwarz ist schmal geschnitten und besitzt einen Komfortbund.

↓

[Diese Bundfaltenhose] [in Schwarz] [ist schmal geschnitten] [und] [besitzt einen Komfortbund].

↓

Diese Bundfaltenhose in Schwarz ist schmal geschnitten und besitzt einen Komfortbund.

PRODUKTTYP	FARBE	SCHNITT	BUNDART	MATERIAL
Bundfaltenhose	schwarz	schmal	Komfortbund	Bio-Baumwolle
Strickjacke	rot	lässig		Kaschmir|Angora

Diese Strickjacke in Rot ist lässig geschnitten und besitzt.

↓

[Diese Strickjacke] [in Rot] [ist lässig geschnitten] [und] [besitzt].

↓

Diese Strickjacke in Rot ist lässig geschnitten.

EINHEITLICHKEIT

> **Einheitlichkeit**

Sind gleiche **Inhalte immer gleich** gepflegt?

Wie viele unterschiedliche **Ausprägungen** hat ein Datenfeld?

> Prüfen, welche Befüllungen es für jedes Datenfeld gibt, um alle Möglichkeiten im Textkonzept zu berücksichtigen

> Ableiten, wie die Datenfeldinhalte textlich verwertbar sind
  - Interpretation möglich
  - Anreichern möglich
  - Reine Nennung möglich

EINHEITLICHKEIT:
ANZAHL & ART DER WERTE PRO DATENFELD

> Hinweis darauf, wie sich das Datenfeld textlich verarbeiten lässt

Durchreichen
(Wert wie aus den Daten übernehmen)

Empfehlenswert: Bei sehr vielen Ausprägungen (meist Freitextfelder), wenn Anreichern von mehreren hundert Werten nötig wäre

Anreichern
(zusätzliche Inhalte zu den Dateninhalten pflegen)

Empfehlenswert: Bei einer begrenzten Anzahl Ausprägungen, sprich einheitlich befüllten Datenfeldern

Bewerten
(mathematische Vergleiche)

Empfehlenswert: Bei numerischen Werten, mit denen sich rechnen lässt

EINHEITLICHKEIT: GRAMMATIKALISCHE FORMEN

> Prüfen, welche grammatikalischen Konstrukte pro Datenfeld vorkommen

Beispiel:

SCHNITT	
schmal	→ Adjektiv
lässig	→ Adjektiv
schmaler Schnitt	→ Konstruktion aus Adjektiv und Substantiv

Wichtig:
im Textkonzept berücksichtigen oder Daten vereinheitlichen

Technische Lösung meist:
Einträge sprachlich vereinheitlichen mithilfe von Lookups

EINHEITLICHKEIT: NUMERISCHE WERTE

> Prüfen, ob Einheiten pro Datenfeld durchgängig sind

Beispiel:

INNENBEINLÄNGE	
88 cm	→ Einheit Zentimeter
0,7 m	→ Einheit Meter
70 cm	→ Einheit Zentimeter
600 mm	→ Einheit Millimeter

Wichtig:
Für das Erstellen logischer Bedingungen und daraus resultierender Textausgaben sollten Einheiten durchgängig gepflegt sein

Beispiel: INNENBEINLÄNGE > 80 → extra lang

HINWEISE UND TIPPS

> Gründliche **DATENANALYSE** durchführen

> **INHALTLICHES POTENZIAL** der Datenfelder nutzen

> Bei der Formulierung von Sätzen darauf achten, dass sich **AUSDRÜCKE UND SATZSTRUKTUR** mit vorhergehenden oder nachfolgenden Sätzen **NICHT DOPPELN**

> **INHALTLICHE BEZÜGE** über Sätze hinweg nur verwenden, wenn deren gemeinsames Auftreten immer zutrifft

VIELEN DANK!

Lisa Weidmann | lisa.weidmann@ax-semantics.com