Corporate Library

Dipl.-Ing. José Manuel de la Rosa Govantes
Redaktion, Analyse und Migration

AGENDA

Über uns
Die semantics Kommunikationsmanagement GmbH

Library Manager
Textpflege und Kollaboration

Corporate Library
Weitere Tools zur Textpflege und -optimierung

Vom Text zur Verständlichkeit
Fachliche Dienstleistungen der semantics

Migration
Einsatz Corporate Library

Über uns

Eintritt 2004

Dipl.-Ing. José de la Rosa
(Fusion mit VisualAix®)

Kernthema bei Gründung

Verbindung von Forschung und Praxis zur Textverständlichkeit und Texttechnologie

Kooperationen

Mit der Wissenschaft und Forschung

Mit Integrationspartnern

Mit Anbietern von CCM- und OMS-Lösungen

Gründung 2000

als Spin-off der RWTH Aachen (Prof. Dr. Christian Stetter und Kay Heiligenhaus M.A.)

Interdisziplinäres Team

aus ca. 40 Informatikern, Sprachwissenschaftlern, Psychologen und Consultants

Enge Kooperation mit der RWTH Aachen, Prof. Dr. Thomas Niehr

Zweites Thema (seit 2005):

Digitalisierungs- und Recherchesysteme für Bibliotheken, Archive und Museen

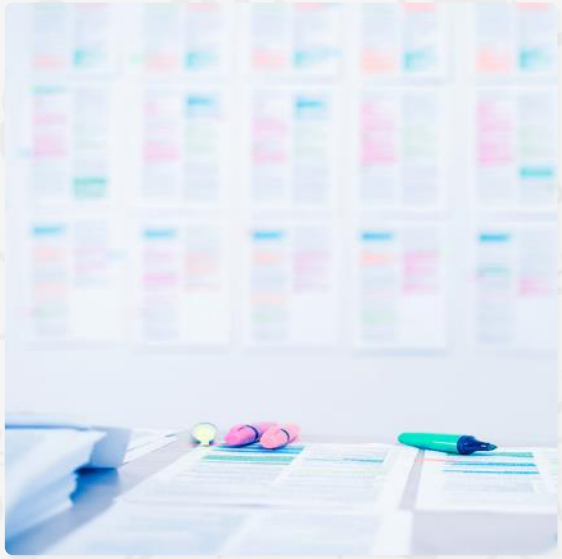
Marktführer D-A-CH

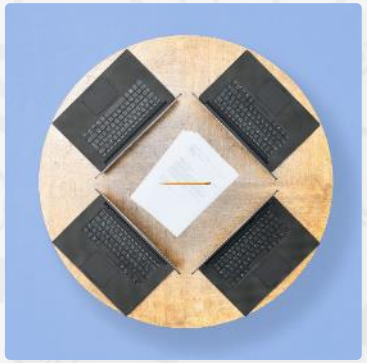
Über uns

Methoden und Technik

Texttechnologie

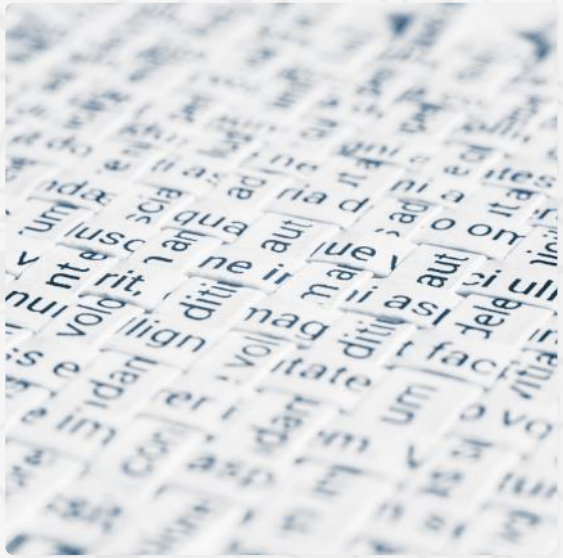
Redaktion, Kollaboration, Dokumentation, Recherche, Analyse, automatisierte Verarbeitung, Migration





Sprachwissenschaft

Schriftgut-Analysen, Sprachrichtlinien, Textoptimierung, Sprachschulungen, Redaktion



Forschung

Empirische Studien, Korpuslinguistik, Linguistische Informatik, Lehre und Ausbildung

Library Manager

Zur Motivation:
Beispiel aus der Praxis

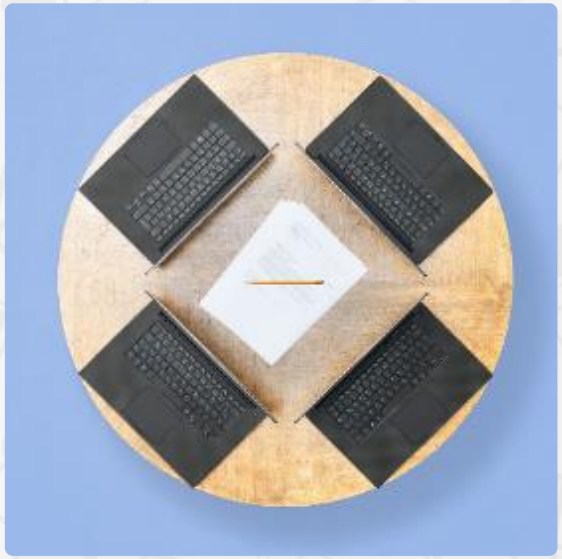
„Der Versicherer gewährt dem Versicherungsnehmer Versicherungsschutz für den Fall, dass er wegen eines während der Wirksamkeit der Versicherung eingetretenen Schadenereignisses, das den Tod, die Verletzung oder Gesundheitsschädigung von Menschen (Personenschaden) oder die Beschädigung oder Vernichtung von Sachen (Sachschaden) zur Folge hatte, für diese Folge aufgrund gesetzlicher Haftpflichtbestimmungen privatrechtlichen Inhalts von einem Dritten auf Schadenersatz in Anspruch genommen wird.“

Quelle: „Zur Verständlichkeit von Geschäftstexten am Beispiel Allgemeiner Versicherungsbedingungen“, Magisterarbeit Björn Eßer, RWTH Aachen, 2012; §1 Nr. 1 AHB, zit. nach Nitschke (2002), S.122.

Eins

Library Manager

Zur Motivation: Zahlen und Situation eines Großkunden (Juli 2017)



Redaktion:

202 aktiv

816 gesamt

Sachbearbeitung:

> 13.000

Textbestand:

71.048 Vorlagen

69.310 Bausteine

18.134 Variablen

26.000 Aktivitäten / Monat

? Textverantwortung:

Vorstand,
Marktmanagement,
Produktgeber,
Fachabteilung, …

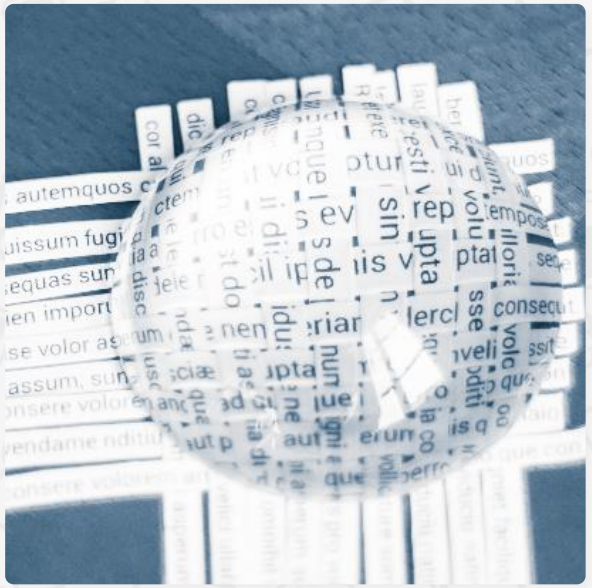
Außendienst:

> 100.000

Kunden

Schriftliche Kommunikation

Dimensionen



Text

Interne Eigenschaften:
Wort-, Satz- u. Textebene

Text-, Korpus-,
strukturale Linguistik
Verständlichkeits-/
Lesbarkeitsforschung

Sprachprüfung,
Empirische Studien

Sprache

Linguistische Analyse

„Der Versicherer gewährt a
Versicherungsschutz für de
während der Wirksamkeit
eingetretenen Schadenerei
Verletzung oder Gesundhei
(Personenschaden) oder di
Vernichtung von Sachen (S
für diese Folge aufgrund ge
Haftpflichtbestimmungen

Der Versicherer gewährt dem Versicherungsnehmer Versicherungsschutz für den Fall, dass er wegen eines während der Wirksamkeit der Versicherung eingetretenen Schadenereignisses, das den Tod, die Verletzung oder Gesundheitsschädigung von Menschen (Personenschaden) oder die Beschädigung oder Vernichtung von Sachen (Sachschaden) zur Folge hatte, für diese Folge aufgrund gesetzlicher Haftpflichtbestimmungen privatrechtlichen Inhalts von einem Dritten auf Schadenersatz in Anspruch genommen wird.

Hauptaussage

konditionaler Nebensatz

kausale Bestimmung

Linksattribut

Relativsatz

Rückbezug

Verweis auf rechtliche Basis

Sprache

Überarbeitung

„Der Versicherer gewährt a
Versicherungsschutz für de
während der Wirksamkeit d
eingetretenen Schadenerei
Verletzung oder Gesundhei
(Personenschaden) oder die
Vernichtung von Sachen (Sc
für diese Folge aufgrund ge
Haftpflichtbestimmungen r

Der Versicherer gewährt dem Versicherungsnehmer Deckung für Haftpflichtansprüche, die gegen ihn geltend gemacht werden.

Haftpflichtansprüche unterliegen gesetzlichen Haftpflichtbestimmungen privatrechtlichen Inhalts.

Sie bestehen dann, wenn der Versicherungsnehmer für Personen- oder Sachschäden verantwortlich gemacht werden kann.

Dieser Schaden muss während der Versicherungszeit entstanden sein, damit er unter die gewährte Deckung fällt.

Hauptaussage

Verweis auf rechtliche Basis

Kausale Bestimmung

Bedingung

Sprache

Thematische Kohärenz zur Unterstützung mentales Modell
(Thema/Rhema Paradigma, anaphorische Bezüge, Rekurrenzen)

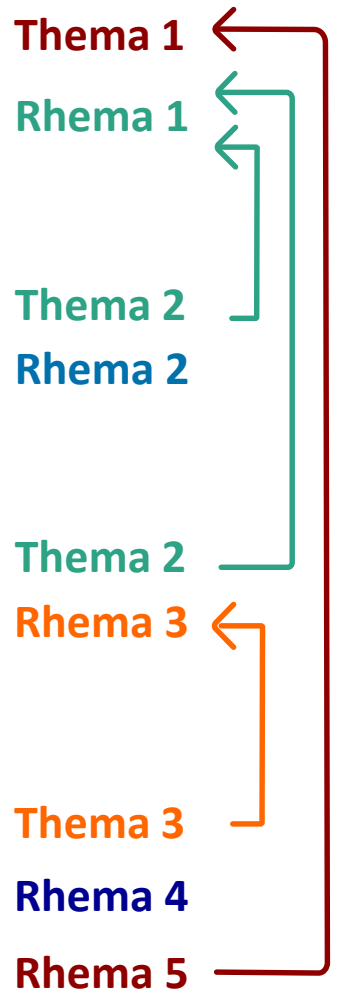
„Der Versicherer gewährt a
Versicherungsschutz für de
während der Wirksamkeit d
eingetretenen Schadenereig
Verletzung oder Gesundhei
(Personenschaden) oder die
Vernichtung von Sachen (Sc
für diese Folge aufgrund ge
Haftpflichtbestimmungen r

Der Versicherer gewährt dem Versicherungsnehmer Deckung für Haftpflichtansprüche, die gegen ihn geltend gemacht werden.

Haftpflichtansprüche unterliegen gesetzlichen Haftpflichtbestimmungen privatrechtlichen Inhalts.

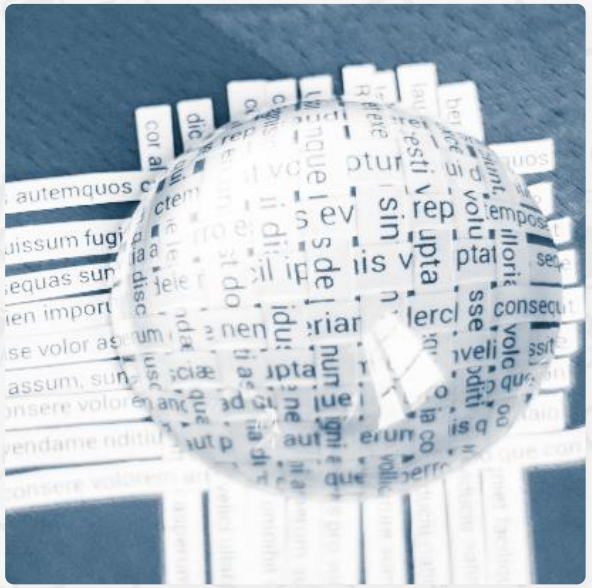
Sie bestehen dann, wenn der Versicherungsnehmer für Personen- oder Sachschäden verantwortlich gemacht werden kann.

Dieser Schaden muss während der Versicherungszeit entstanden sein, damit er unter die gewährte Deckung fällt.



Schriftliche Kommunikation

Dimensionen



Text

Interne Eigenschaften:
Wort-, Satz- u. Textebene

Text-, Korpus-,
strukturale Linguistik
Verständlichkeits-/
Lesbarkeitsforschung

Sprachprüfung,
Empirische Studien

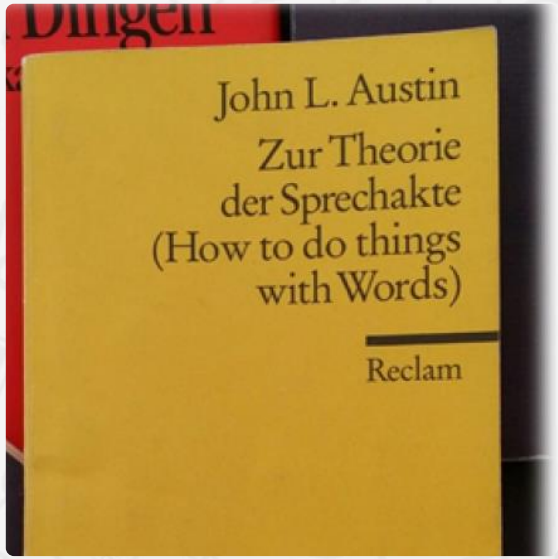
Empfänger

Zielgruppe, Motivation, …

Kognitionswissenschaft,
Neurolinguistik, …
Sprechakttheorie

Unterstützung
Mentales Modell

Kommunizieren ist Handeln



> Ich komme morgen.

Lokutionärer Akt

Äußerung: Vokabular, Grammatik

Illokutionärer Akt

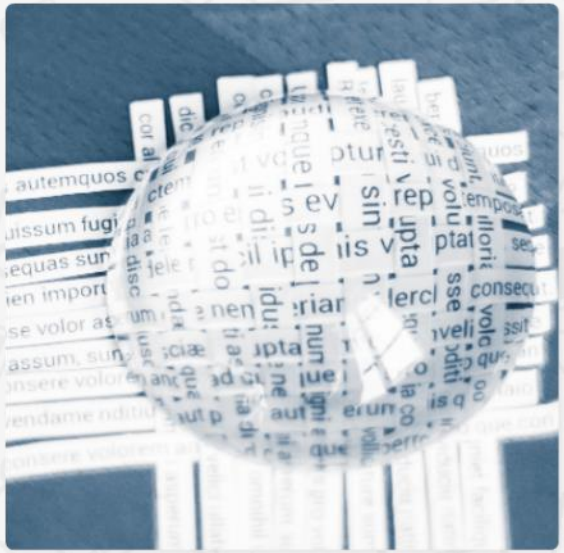
Rolle der Äußerung: versprechen, warnen, drohen…

Perlokutionärer Akt

Folge der Äußerung: Verpflichtung, Vereinbarung, …

Illokution muss gelingen, i.d.R. aber nicht expliziert sondern durch Kontext bekannt

Verständlichkeits-ebenen und Kontext



1 Gesamtstruktur des Textes: Gliederung

2 Ebene der Sachverhalte: Absätze und Sätze

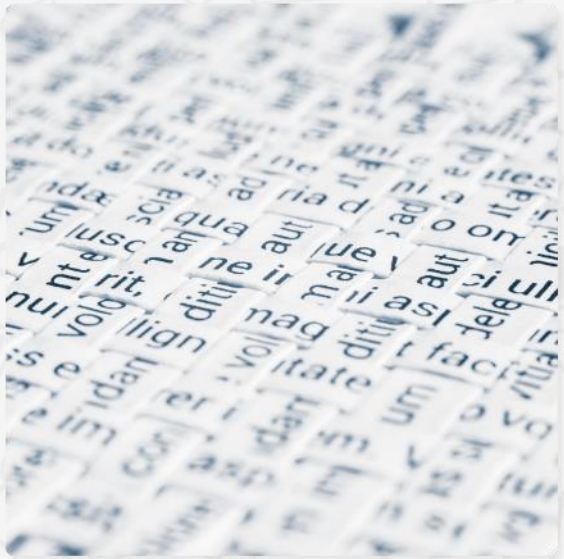
3 Vokabular: Wortebene

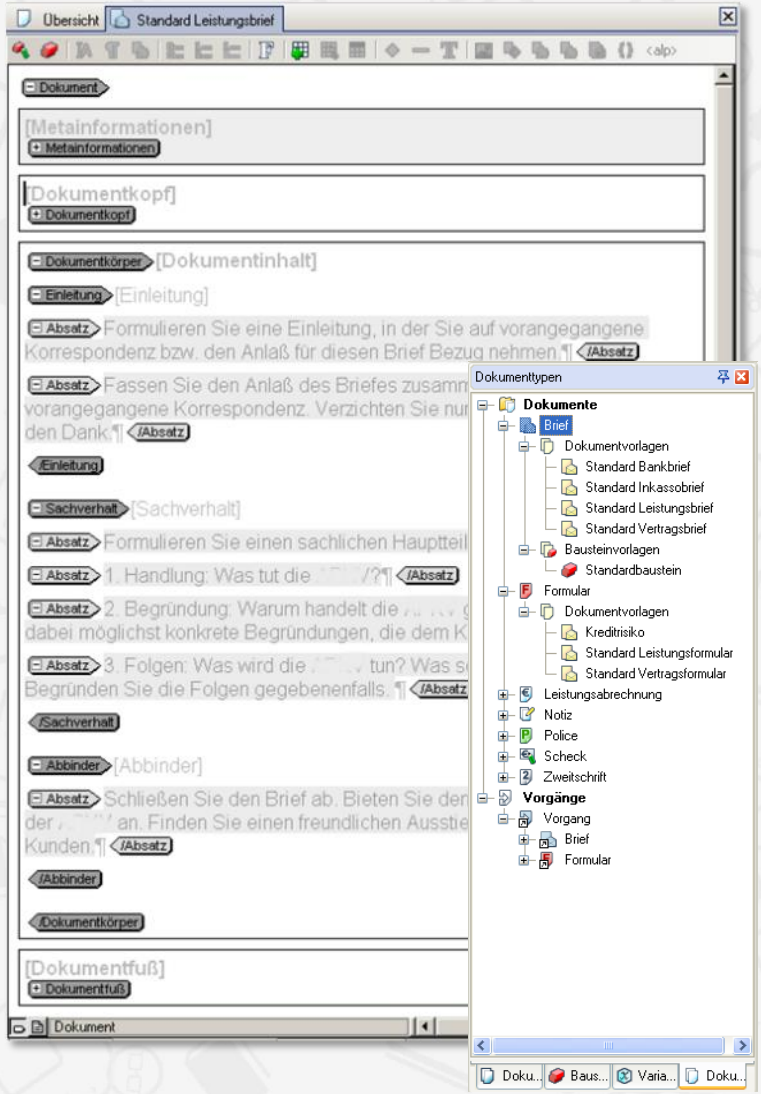
1-3 Gewichtung

Das Wichtigste ist die Gesamtstruktur: Die **Gliederung** erlaubt es, den Kontext herzustellen

Library Manager

Dokumentstruktur je Dokumenttyp und Anlass





Basisvorlagen zur Neuanlage von Vorlagen: Gliederung, Redaktionsanweisungen, Standards zum Workflow, Kopf/Fuß, …

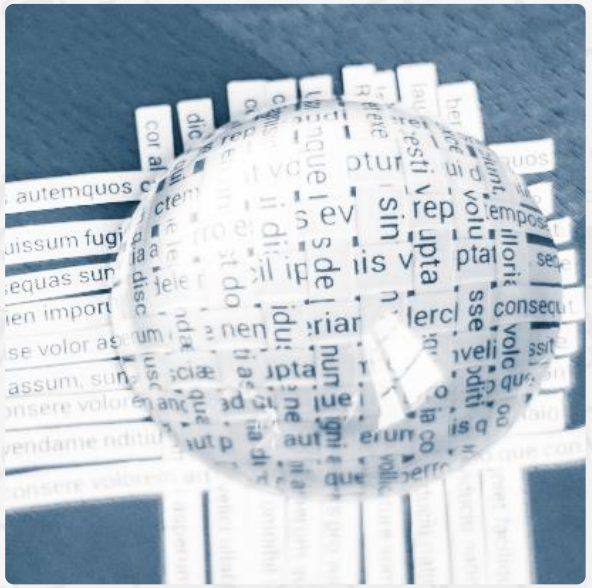
Unterscheidung von Dokumenttypen nach kommunikativen Funktionen oder technischen Rahmenbedingungen

- Zentrale Steuerung der Vorlagenerstellung, offenes System von Dokumenttypen

- Erhebliche Steigerung der inhaltlichen Qualität durch wohlkonzipierte Gliederungen und Standards

Schriftliche Kommunikation

Dimensionen



Text

Interne Eigenschaften:
Wort-, Satz- u. Textebene

Text-, Korpus-,
strukturale Linguistik
Verständlichkeits-/
Lesbarkeitsforschung

Sprachprüfung,
Empirische Studien

Empfänger

Zielgruppe, Motivation, …

Kognitionswissenschaft,
Neurolinguistik, …
Sprechakttheorie

Unterstützung
Mentales Modell

Technik

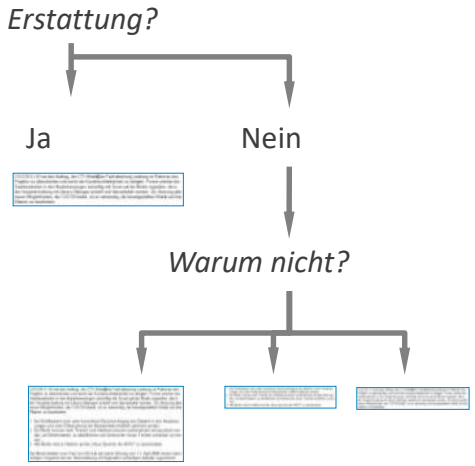
Paradigmen beeinflussen
Sprachqualität

Paradigmen-wechsel

Varianten statt Textbausteine

Kopf, Betreff, Einleitung, ...
(Texte oder Bausteinreferenzen)

Hauptteil



Schluss, Abbinder, Fuß, ...

Input durch Datenbeschaffung:
Erstattung = Nein

Komposition

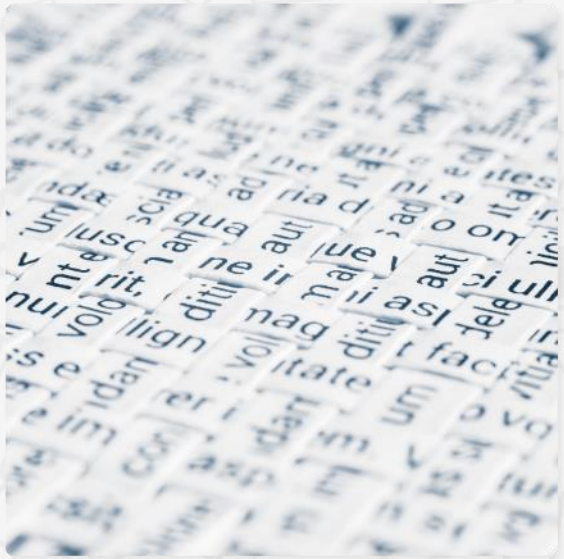
Auswahl durch Sachbearbeiter:
Begründung = C

Individuelle Briefe

- Fördern der wichtigsten rhetorischen Größe: Gliederung und Gesamtstruktur
- Mindern der Nachteile von Textbausteinen (Verlust Kohärenz, Übersicht; nicht passende Übergänge; ...)
- Sinnvoller Einsatz Bausteine für übergreifende Köpfe/Füße, Meta-/Steuerbausteine, Produktkomposita bei Policen, AVB, ...

Library Manager

Varianten ohne Programmierung

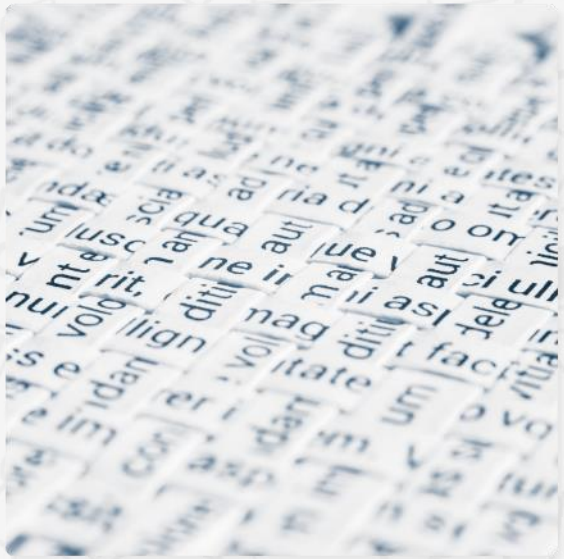


- *„turing complete“* und *„single source“*
- *Entkopplung von IT*
- *Bei Realisierung entsteht auch Dokumentation inkl. Historie*

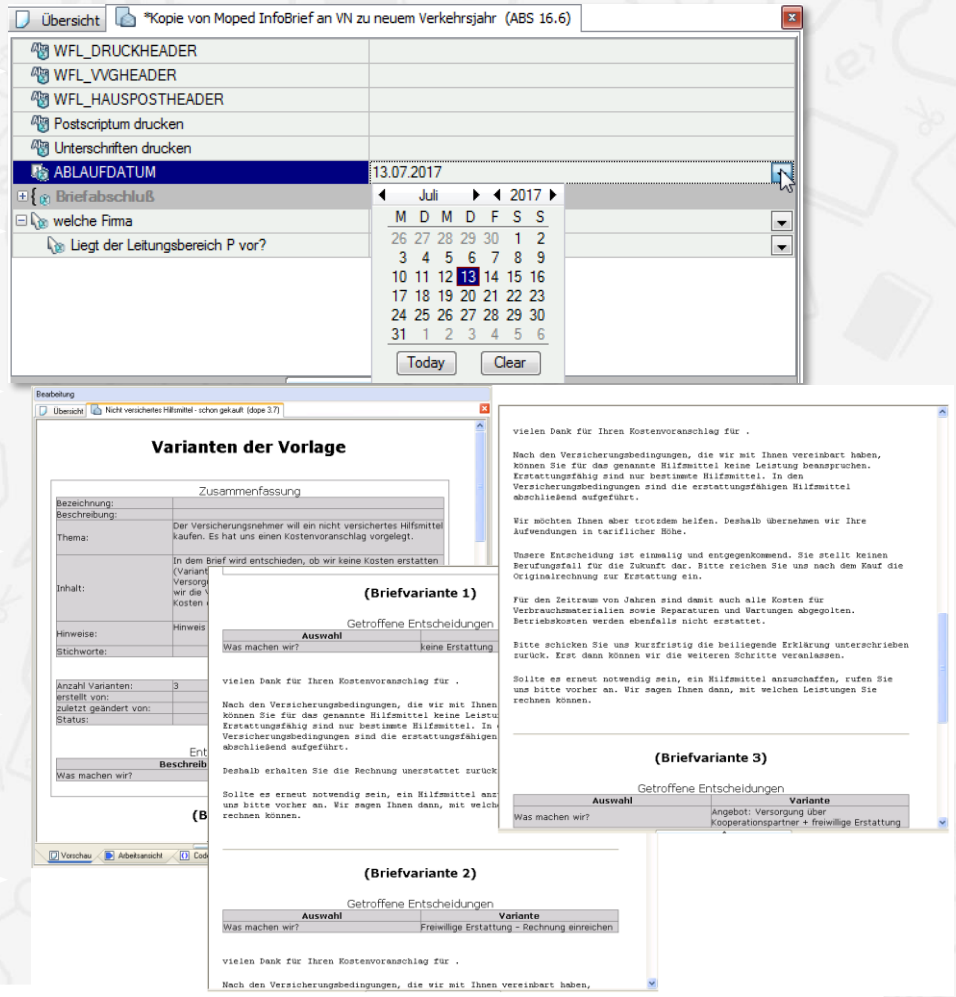
- Modellierung der Logik ohne technische Kenntnisse
- Beliebige Verschachtelung der Strukturen (Variante, Iteration, …)
- Einsatz Textbausteine und Variablen per Drag & Drop
- Keine spezifische Programmierung je Vorlage, Kanal, …
- Identisches Verfahren Online und Batch
- Kontrolle Textanpassungen durch Sachbearbeiter
- Hinweise zu Inhalt, Zweck, Verwendung

Library Manager

Logik- und Variantenprüfung



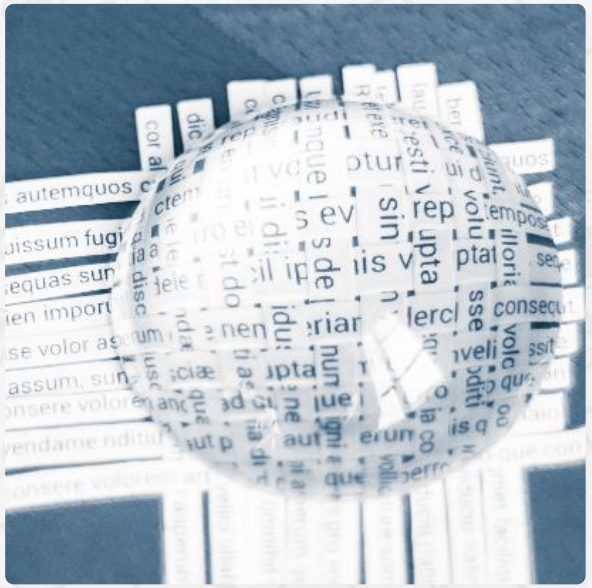
- Vorschau mit Test der Variablen
- Speichern von Testszenarien
- Ausgabe aller möglichen Permutationen einer Vorlage
- Formatierte Vorschau: involviert im Haus genutzten Formatierer
- ggf. Dokumentation der Testergebnisse
- ggf. als notwendiger Schritt im Redaktionsprozess

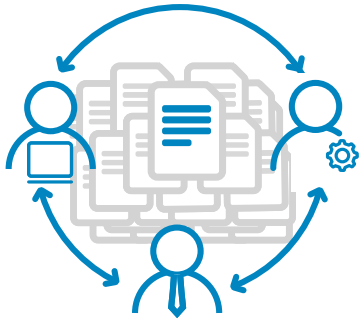


Maximale Sicherheit zur Qualität aller möglichen Varianten einer Vorlage

Schriftliche Kommunikation

Dimensionen

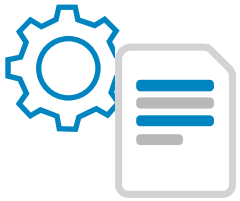




Organisation

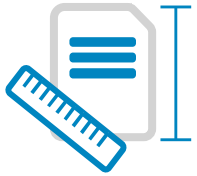
Kompetenzgemeinschaften (fachlich, juristisch, …)

Wissensmanagement, Kollaboration, Redaktion, Qualitätssicherung



Technik

Paradigmen beeinflussen Sprachqualität



Text

Interne Eigenschaften: Wort-, Satz- u. Textebene

Text-, Korpus-, strukturale Linguistik Verständlichkeits-/ Lesbarkeitsforschung

Sprachprüfung, Empirische Studien

Empfänger

Zielgruppe, Motivation, …

Kognitionswissenschaft, Neurolinguistik, … Sprechakttheorie

Unterstützung Mentales Modell

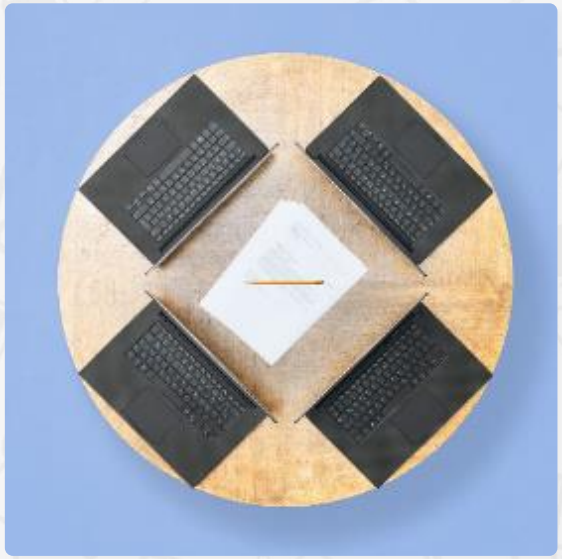
Multi-Channel

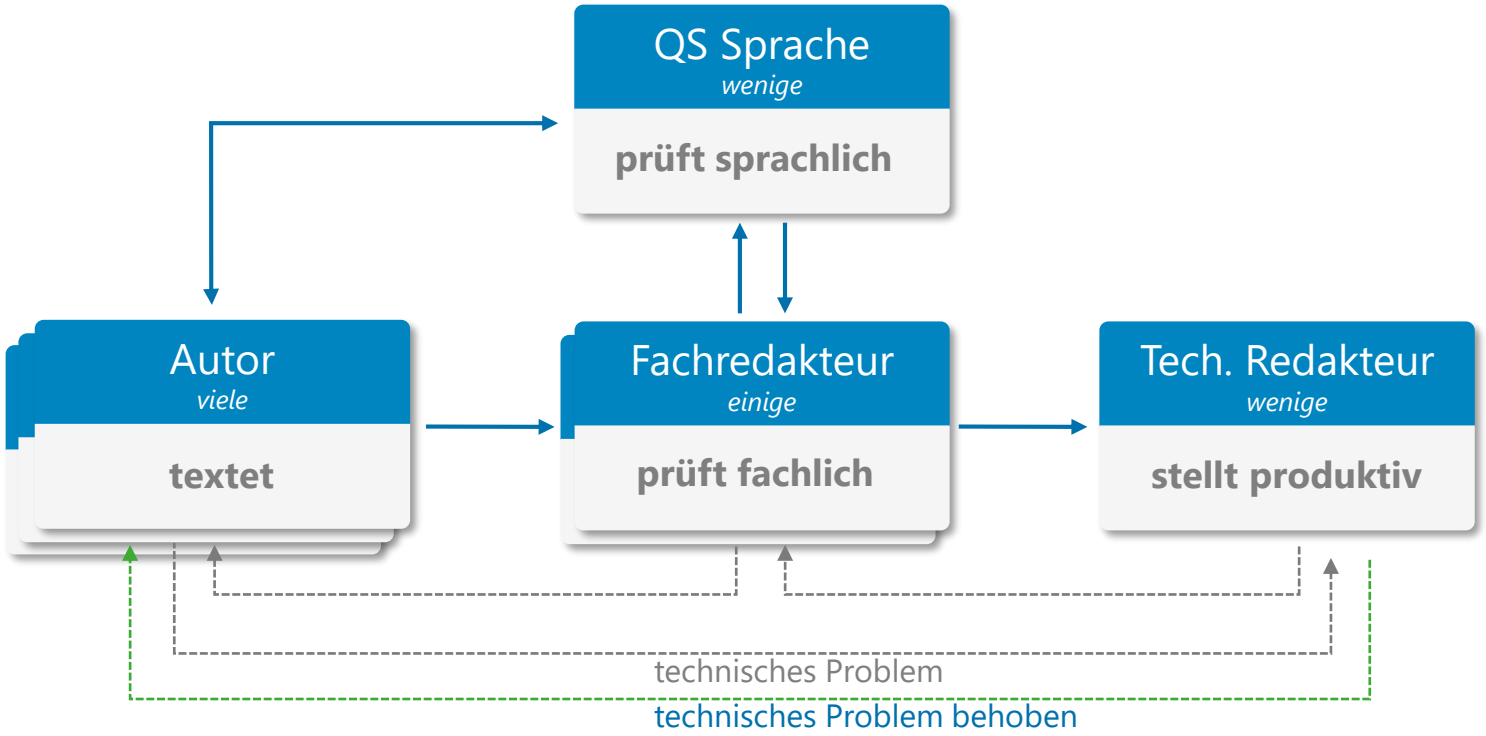
Brief, E-Mail, SMS, App, Internet, Chat, …

Responsive Documents, Variantentexte

Library Manager

Redaktionsprozess ohne Medienbrüche



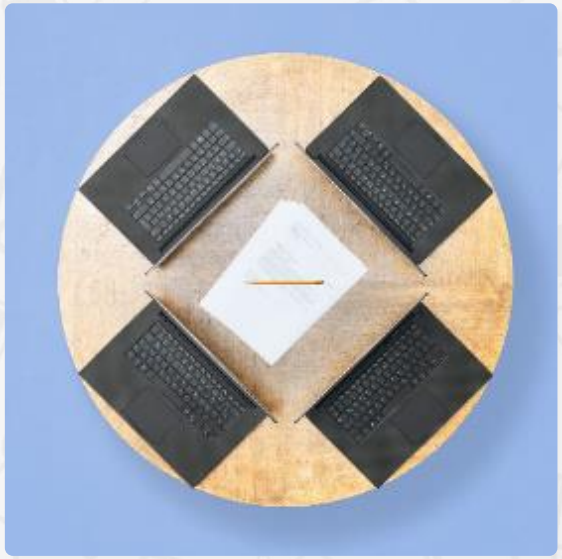


Beispiel (PKV):

- 300 Leistungsbriefe, 80% Abdeckung
- 43 Personen, 15 Referate, 1 Linguist
- 46% Reduktion Beschwerdequoten

Library Manager

Redaktionsprozess ohne Medienbrüche



Lineare Qualitätssicherung

Fachredakteur 1

textet

Fachredakteur 2

interne QS

Projektteam

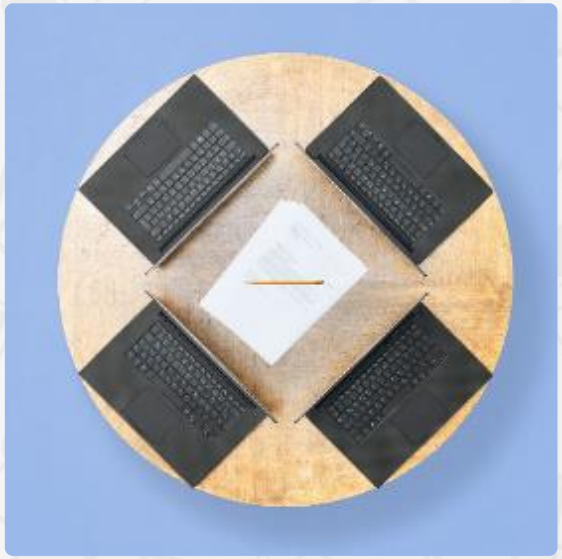
finale QS

Tech. Redakteur

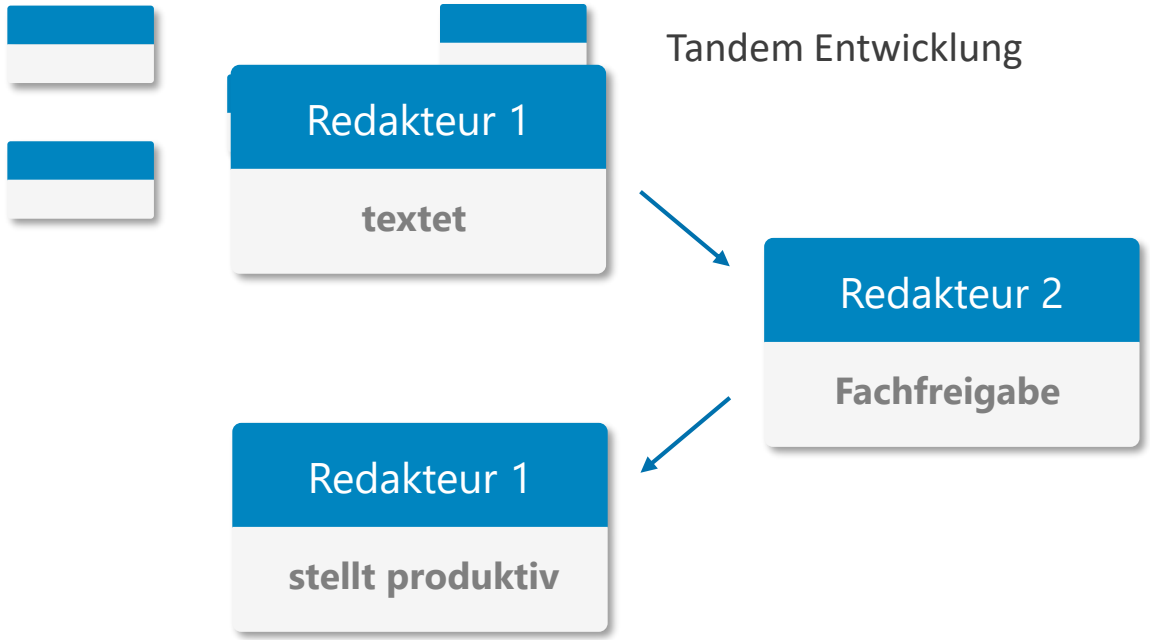
stellt produktiv

Library Manager

Redaktionsprozess ohne Medienbrüche



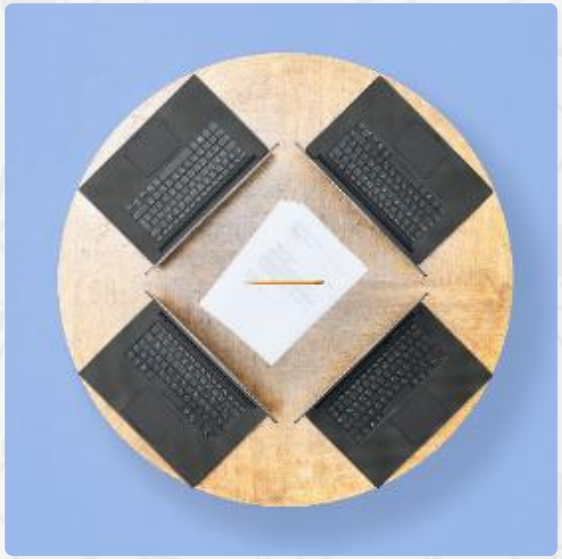
Lineare Qualitätssicherung



Tandem Entwicklung

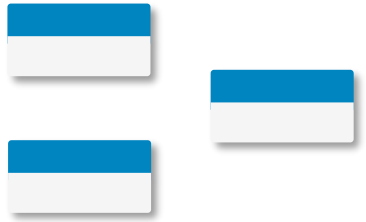
Library Manager

Redaktionsprozess ohne Medienbrüche



Lineare Qualitätssicherung

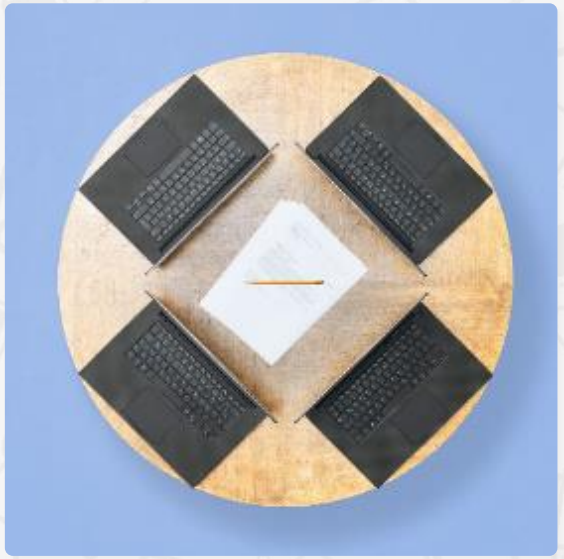
Tandem Entwicklung

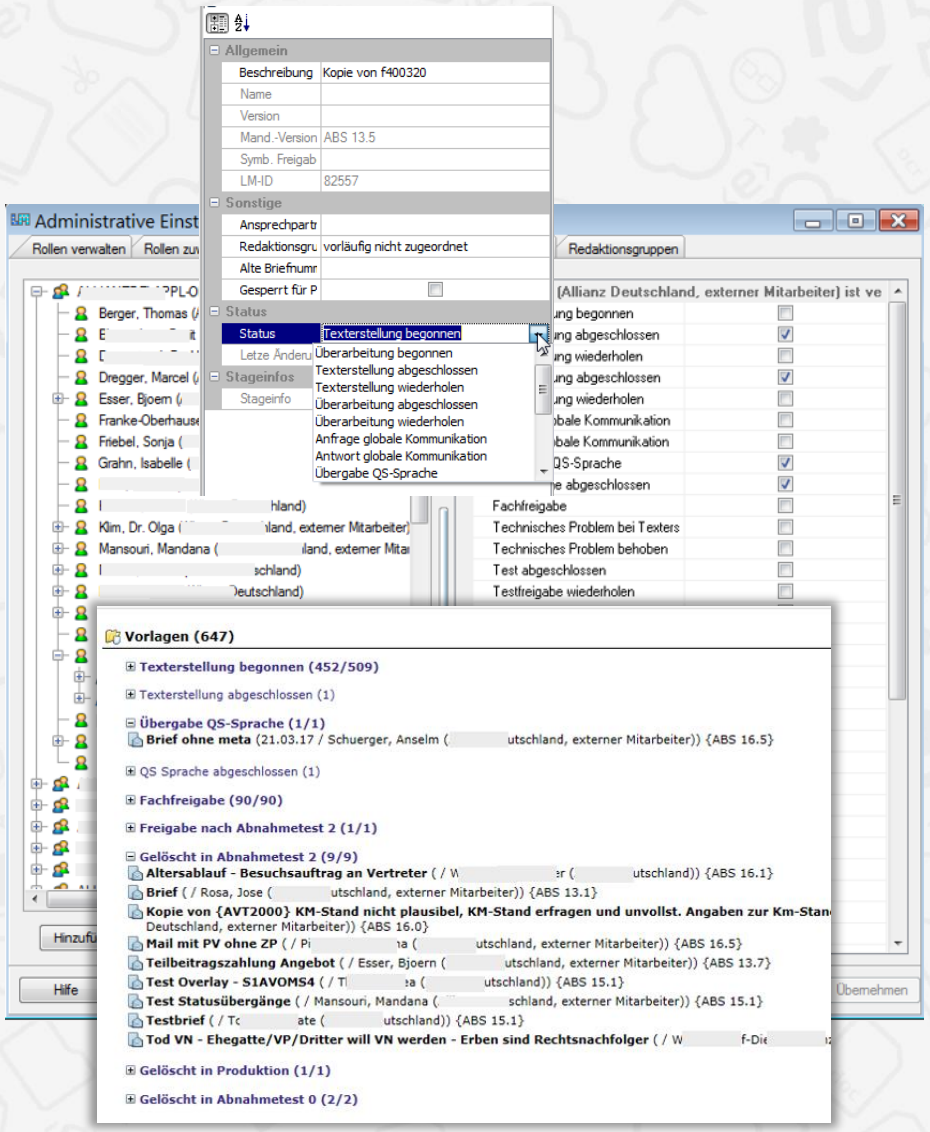


4 Augen Prinzip

Library Manager

Redaktionsprozess ohne Medienbrüche

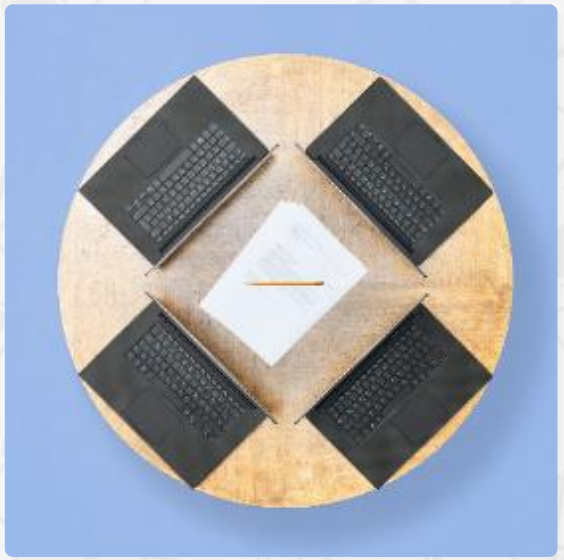


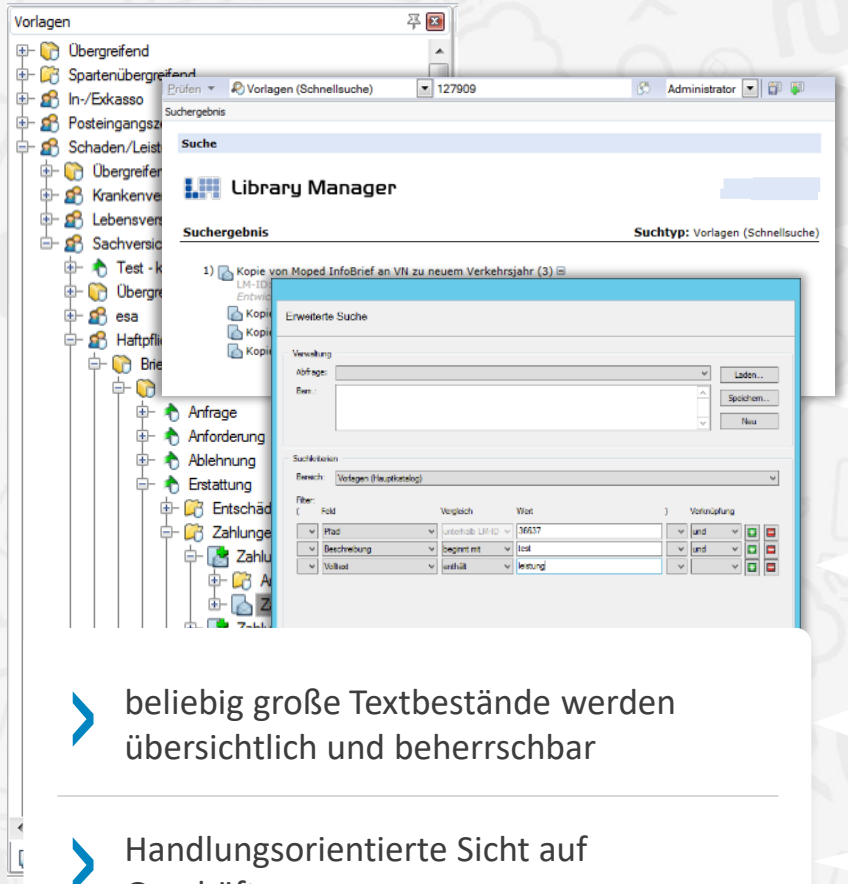


- Abbildung des Redaktionsprozesses durch Darstellung von Zuständigkeiten bei Statusübergängen
- Zusätzlich Definition von Redaktionsgruppen und Verantwortlichkeiten
- Übersicht je Nutzer (Vorlagen/Textbausteine in Bearbeitung, erhaltene, weitergeleitete, …)
- Übergabe Elemente durch Statuszuweisung (manuell und/oder automatisch z.B. bei Freigabe)

Library Manager

Organisation als strukturierte Bibliothek





- beliebig große Textbestände werden übersichtlich und beherrschbar
- Handlungsorientierte Sicht auf Geschäftsprozesse
- keine Redundanzen, effiziente Pflege
- keine technischen Abhängigkeiten (Recherche, Management)

Organisation von Vorlagen in Bibliothek

Nutzung Handlungstheorie (Akteure, Themen, Handlungen)

kontinuierlicher Aufbau Unternehmensbibliothek

flexible Suchfunktionen, keine technischen Kenntnisse

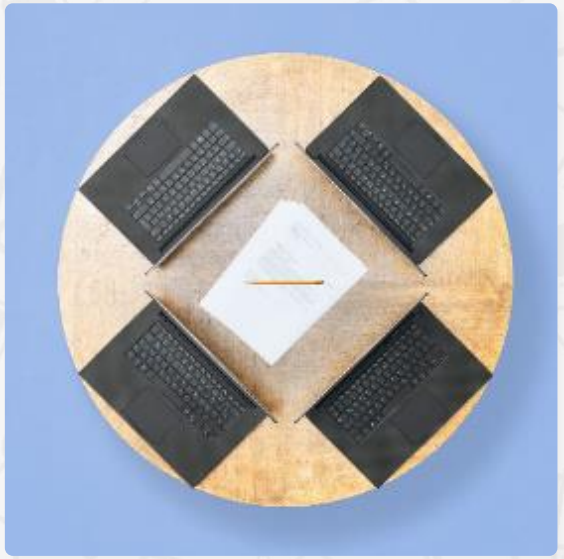
auch Suche quer durch Historie möglich

Ersetzungen quer durch Bestand möglich

Mehrdimensionales Zugriffskonzept (Rechte / Rollen)

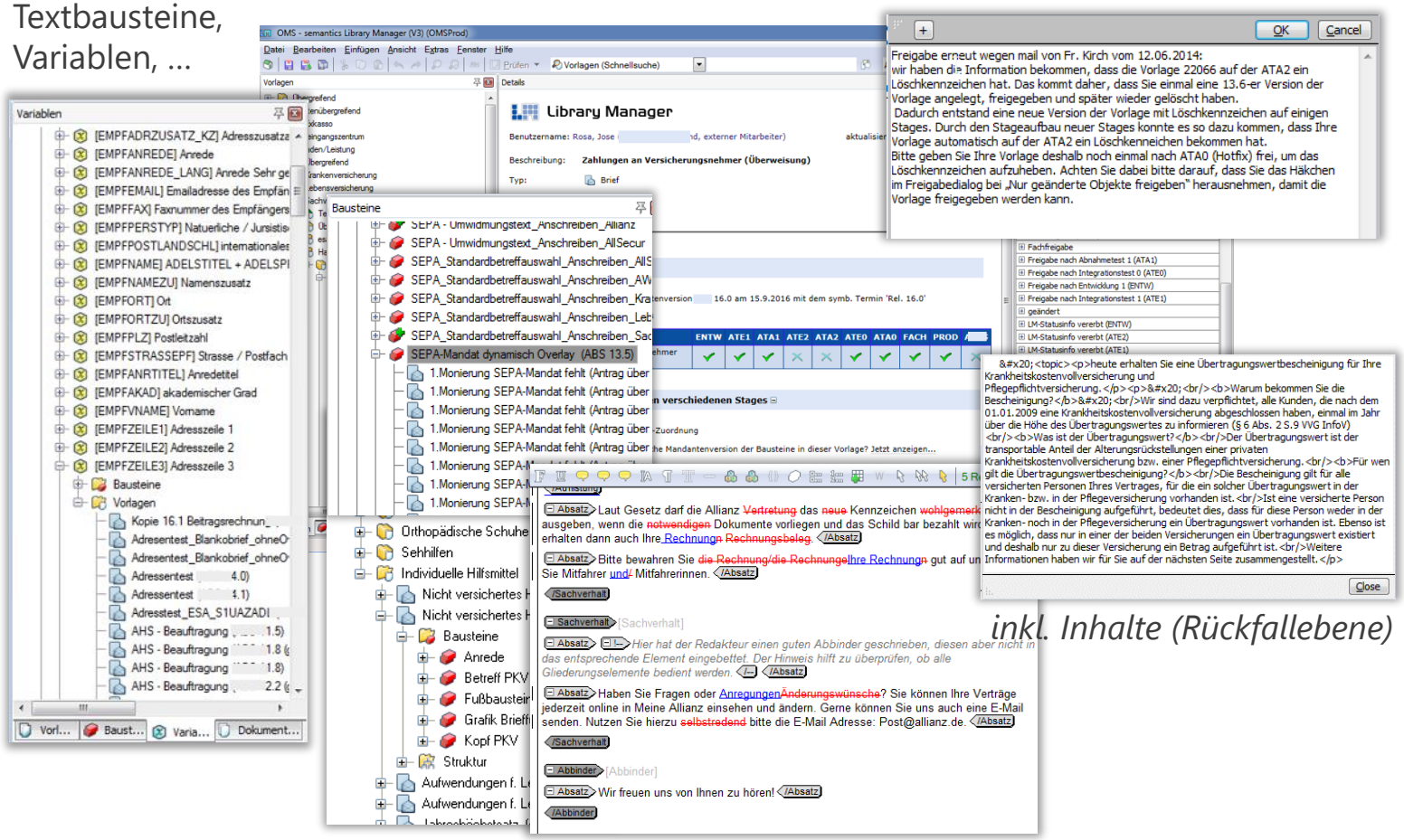
Library Manager

Dokumentation inkl. Historie



Ablage gemäß Anlass, Nutzung auf Stages, Statusverlauf inkl. Kommentare

Verwendung Textbausteine, Variablen, ...



inkl. Inhalte (Rückfallebene)

Änderungsverfolgung (inkl. Historie)

Library Manager

Zusammenfassung

Kollaboration

Redaktionsprozesse ohne Medienbrüche

Sprachqualität

Corporate Wording, Struktur / Gliederung, Verständlichkeit

Dokumentation

Versionen, Verwendungen, Querbeziehungen; Annotationen

Library Manager

Organisation

Verantwortlichkeiten, Bibliothek, Suche, Recherche

Redundanzsuche

N-Gramme, Ermittlung von Abdeckungen

Varianten vs. Bausteine

Unterstützung thematische Kohärenz, Mentales Modell

Redaktion	Ressourcen	Erstellung	Druck	Nachverarb.
• Organisation • Entwicklung • Abstimmung • Qualitäts-sicherung • Dokumentation • Team-Work • Recherche • …	• Textbausteine • Vorlagen • Variablen • Schriften • Logos • Unterschriften • Formulare • …	• Manuelle Dokumente • Batch-verarbeitung • Daten-beschaffung • Dokumentlogik • …	• Komposition • Formatierung • Rendering • Multi-Channel-Output • …	• Bündelung • Sortierung • Beilagen • Frankierung • Verfolgung • Versand • Archivierung • …

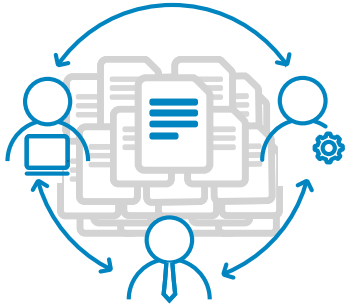
Feedback, Änderungsverfolgung

Lösung des systemischen Problems durch ergänzende Technik

Erzeugung Ressourcen zahlreicher Partner-CCM

Corporate Library

Library Manager



Redaktion

Kollaboration, Redaktion, Qualitätssicherung, Kompetenzgemeinschaften

Freies Schreiben

Sprachprüfung, Lesbarkeit, Corporate Wording, Vorschlagswesen

Empfänger

Zielgruppe, Motivation, Kontext

Library Textanalysis

Library Assistant

Sentiment Inbound

Grundlagen

Wissensmanagement Organisationstheorie, Text Mining, Korpuslinguistik, LSA, LDA, …	Textlinguistik, Verständlichkeits- / Lesbarkeitsforschung, Empirische Studien	Kognitionswissenschaft, Mentale Modelle, Sprechakttheorie

Analyse

Erfahrungen mit „kleinen“ Projekten

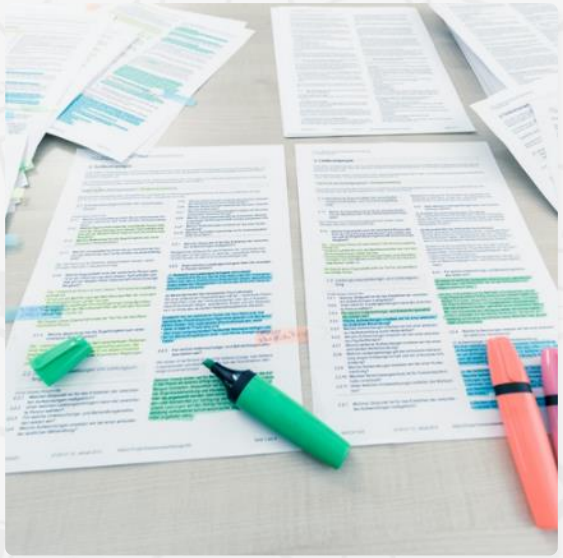
Redundanzanalyse schweizerische KV, 120 Dokumente, überschaubar (…?)

Nutzung Vergleichstool; Ergebnisse konservieren schwierig

Nach kurzer Zeit entsteht Raumbedarf – wie übergreifende Redundanzen visualisieren?

Überblick gelang durch Nutzung aller Dimensionen…

Interesse an weiteren Projekten sehr gering (bei der Belegschaft)



Analyse

Herleitung Zahlen

Vergleich *d* Dokumente

Anzahl Operationen

$$D = \frac{d * (d-1)}{2}$$

Vergleich Absätze

Anzahl Operationen

$$A = a_1 * a_2$$

Summe Operationen

$$G = D * A$$

Textpassagen zu vergleichen

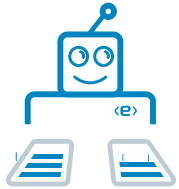
„Kleines“ Projekt

120 Dokumente, je 15-20 Absätze

$D = 7.080, \bar{A} \approx 289$

$\Rightarrow \boldsymbol{G \approx 2.046.120}$

Optimierung: intuitive Bewertung



Pilotprojekt PKV AVB

2.500 Dokumente, je ca. 200 Absätze

$D = 3.123.750, \bar{A} \approx 40.000$

$\Rightarrow \boldsymbol{G \approx 124.950.000.000}$

Automatisierung notwendig

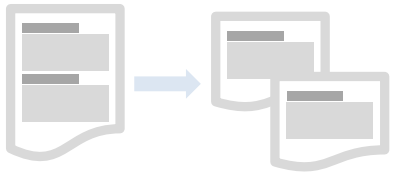
Analyse

Algorithmen

Suche identische und ähnliche Passagen

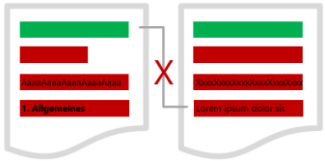
Zusammenhalten kurzer „Ausreißer“

durch Variablen aufzufangen



Splitten großer Fragmente

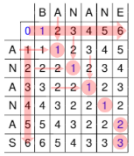
z.B. anhand Überschriften, künftige Wiederverwendung



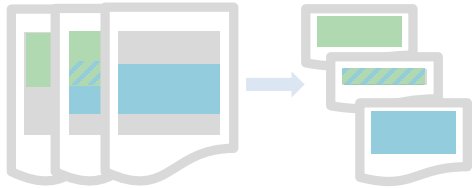
Optimierung

Identität, Länge, Fingerprint, Absatzformat, Entfernung…

(heute: 2.500 Dok. in 8 Min.)

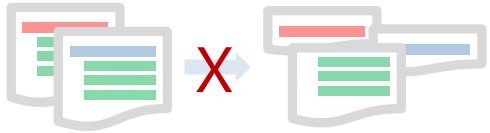


Levenshtein-Distanz zwischen Absätzen



Splitten von „Overlapps“

Höhere Wiederverwendung

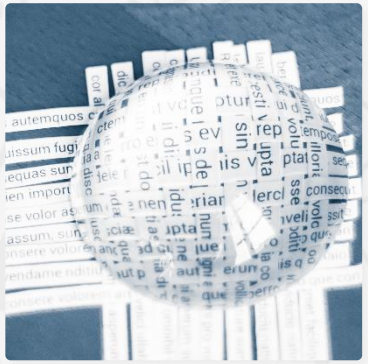


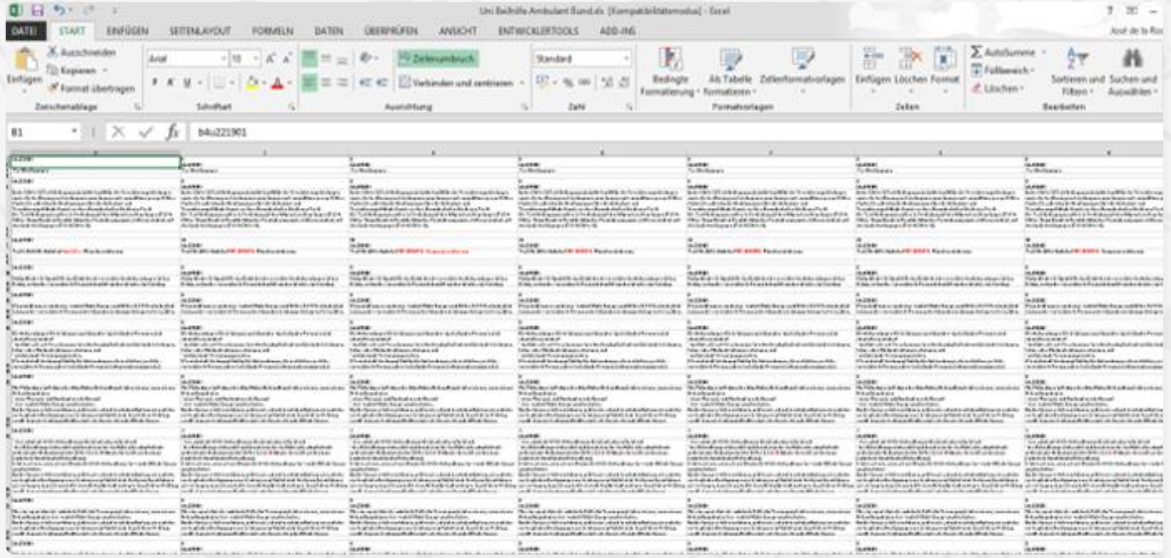
Zusammenhalten fachlicher Einheiten

z.B. Einleitende Sätze und Aufzählungen

Analyse

Darstellung Ergebnisse





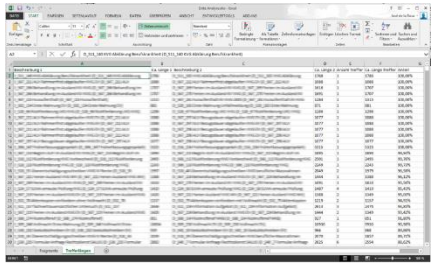
Fachabteilung ohne technische Kenntnisse

Herausforderung

Tausende Ergebnisse visualisieren

Bekannte Tools

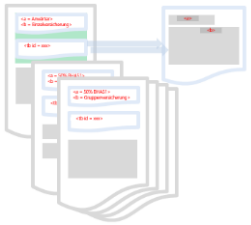
Excel (andere Formate möglich)
Treffer als Zeilen im Sheet
%-Angabe Abweichung



Ranking Dokumentähnlichkeit

Identifikation Variantenbriefe

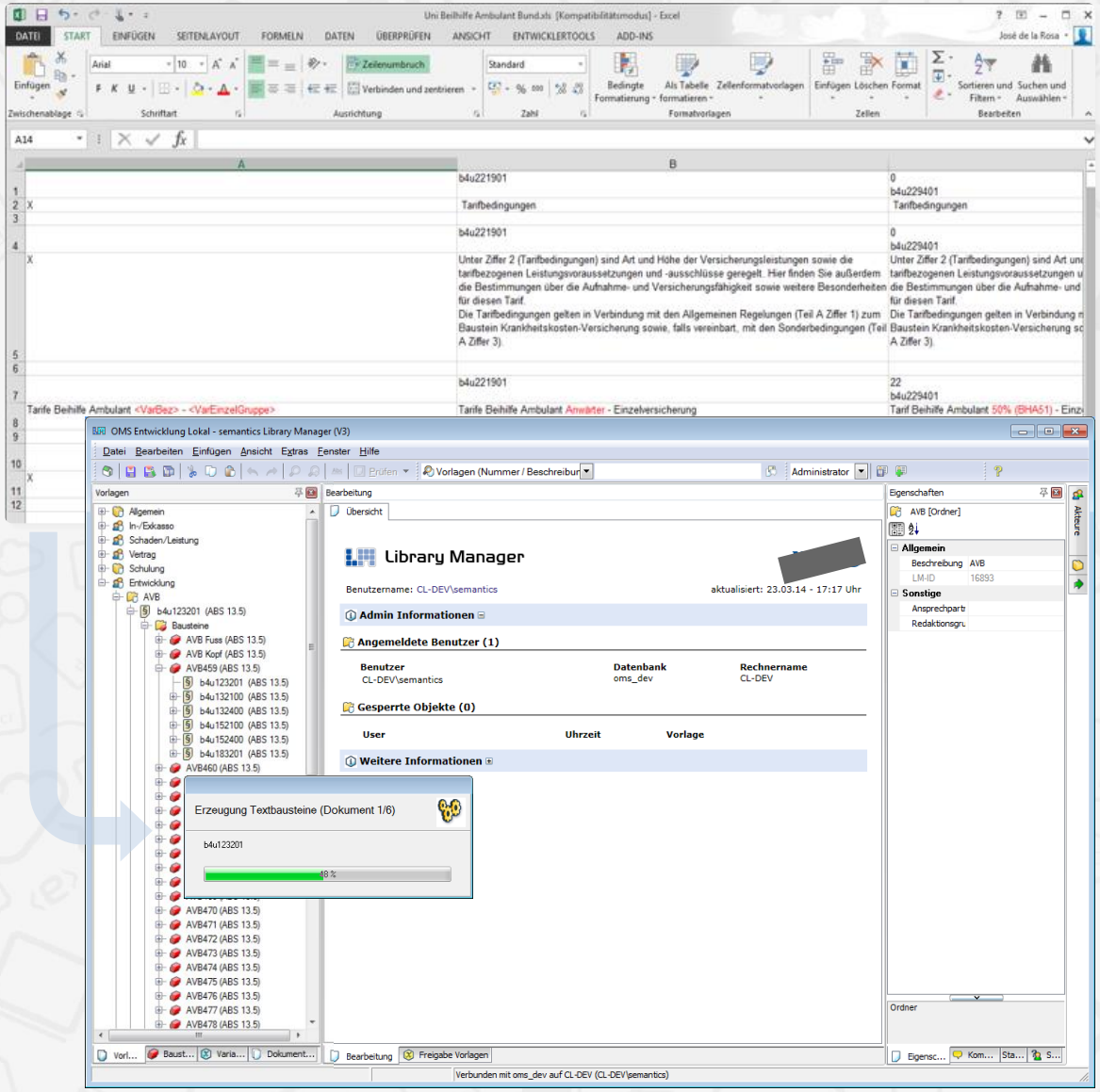
Konsolidierung



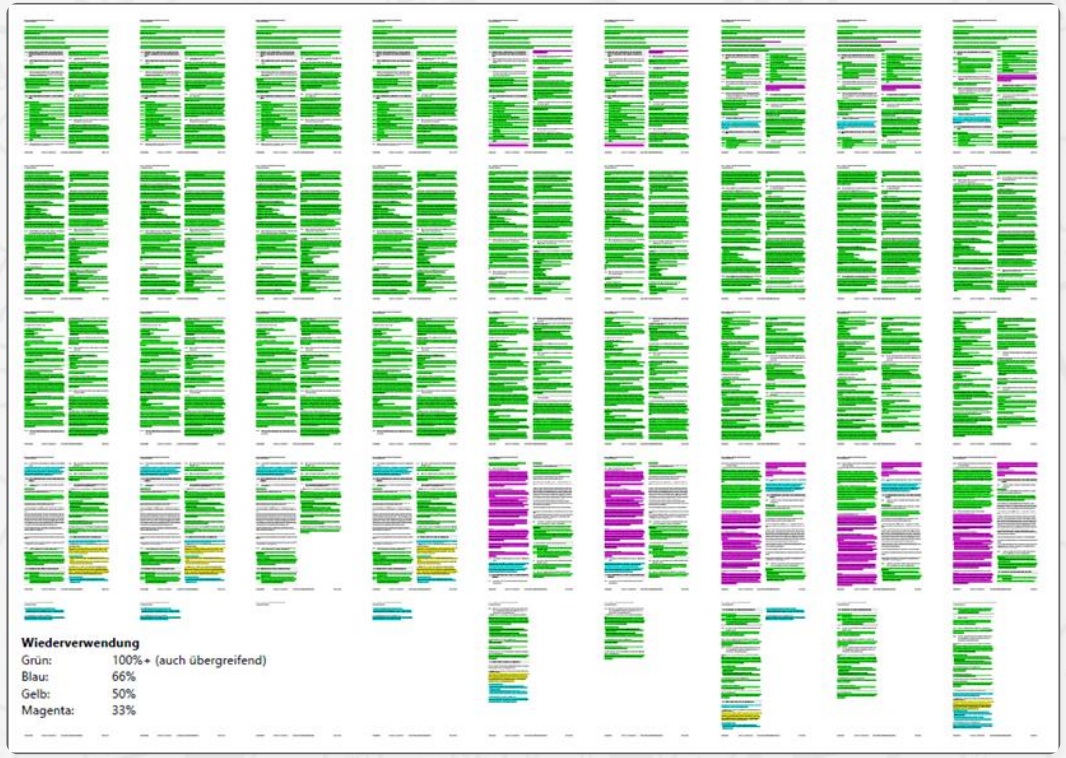
Vorgang

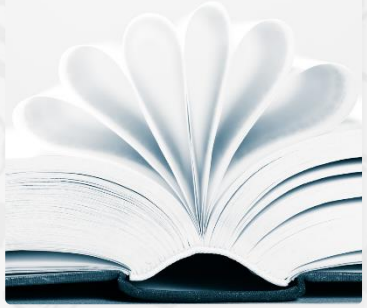
Anlagen Bausteine, Ersetzen in Dokumenten, Eintragen Variablen und Werte

Fachabteilung ohne technische Kenntnisse



Ergebnisse





Selbstheilung

Massenvergleich spürt kleine Fehler auf

Interpunktion, Orthographie, Doppellungen, Inkonsistenzen

Erwartungen übertroffen

- 99% Abdeckung
- 15.000 Textbausteine erzeugt
- 180.000 Referenzierungen TB
- 400fache Wiederverwendung relevante TB
- 300 Variablen, bis 330 Ausprägungen
- Ersparnis > 25 Personen*jahre*

Interne Revision

Belobigung der Abteilung

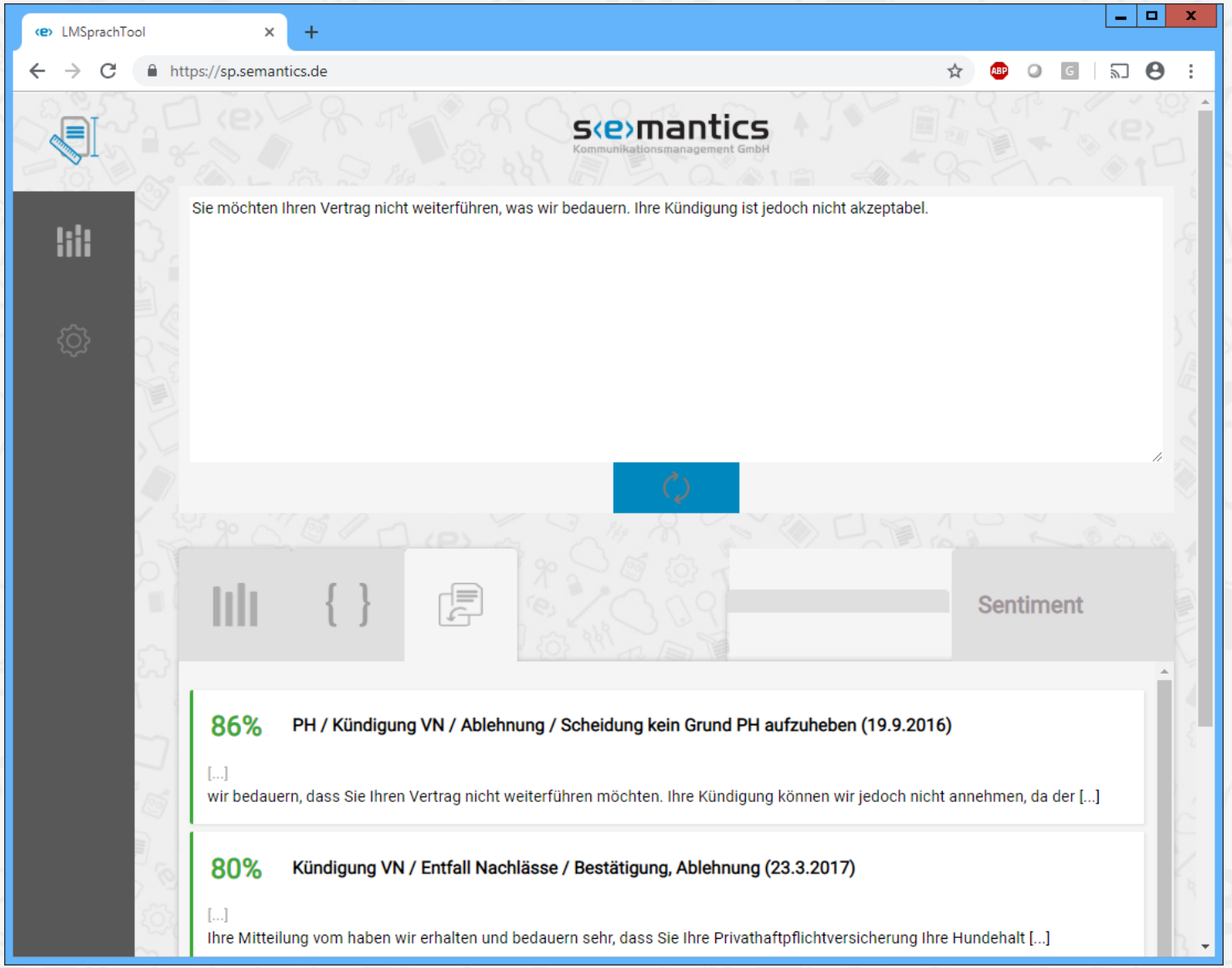
Analyse

Suche nach Paraphrasierungen

Suche Umformulierungen

Ermittlung Abdeckungen in Sätzen, Absätzen

Nutzung Thesaurus



Library Assistant

Architektur

Redaktion

LM-Sprachtool

Editor

Content Management

Library Assistent

Webservice
Nahtlose Integration,
Anbindung Drittlösungen

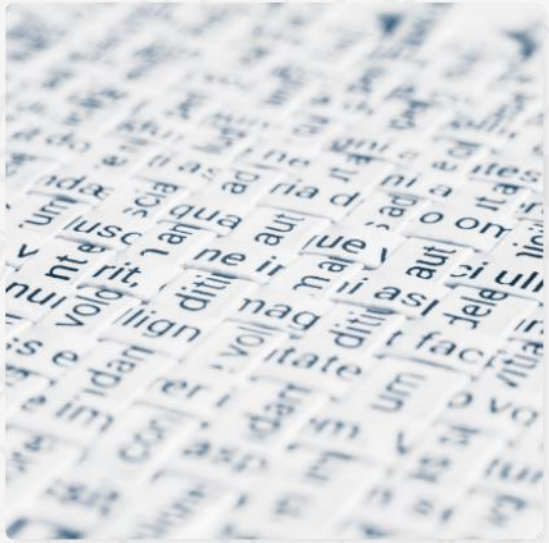
Sentiment,
Big Data

Keyword- / Topic Extraction

CL-Textanalysis

Library Assistant

Sprachprüfung



Redaktion

LM-Sprachtool

Editor

Content Management

Library Assistent

Webservice
Nahtlose Integration,
Anbindung Drittlösungen

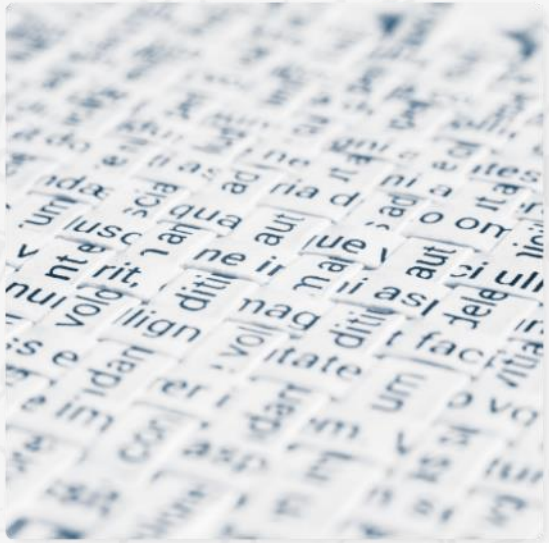
Sentiment,
Big Data

Keyword- / Topic Extraction

CL-Textanalysis

Library Assistant

Vorschlagswesen

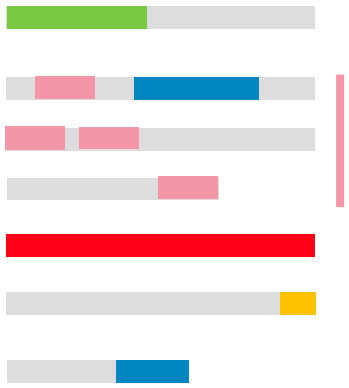


Redaktion

LM-Sprachtool

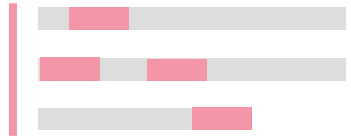
Editor

Content Management



Library Assistent

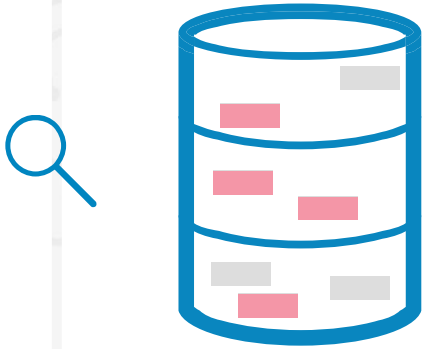
Webservice
Nahtlose Integration,
Anbindung Drittlösungen



Sentiment,
Big Data

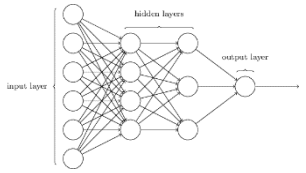
Keyword- / Topic Extraction

CL-Textanalysis



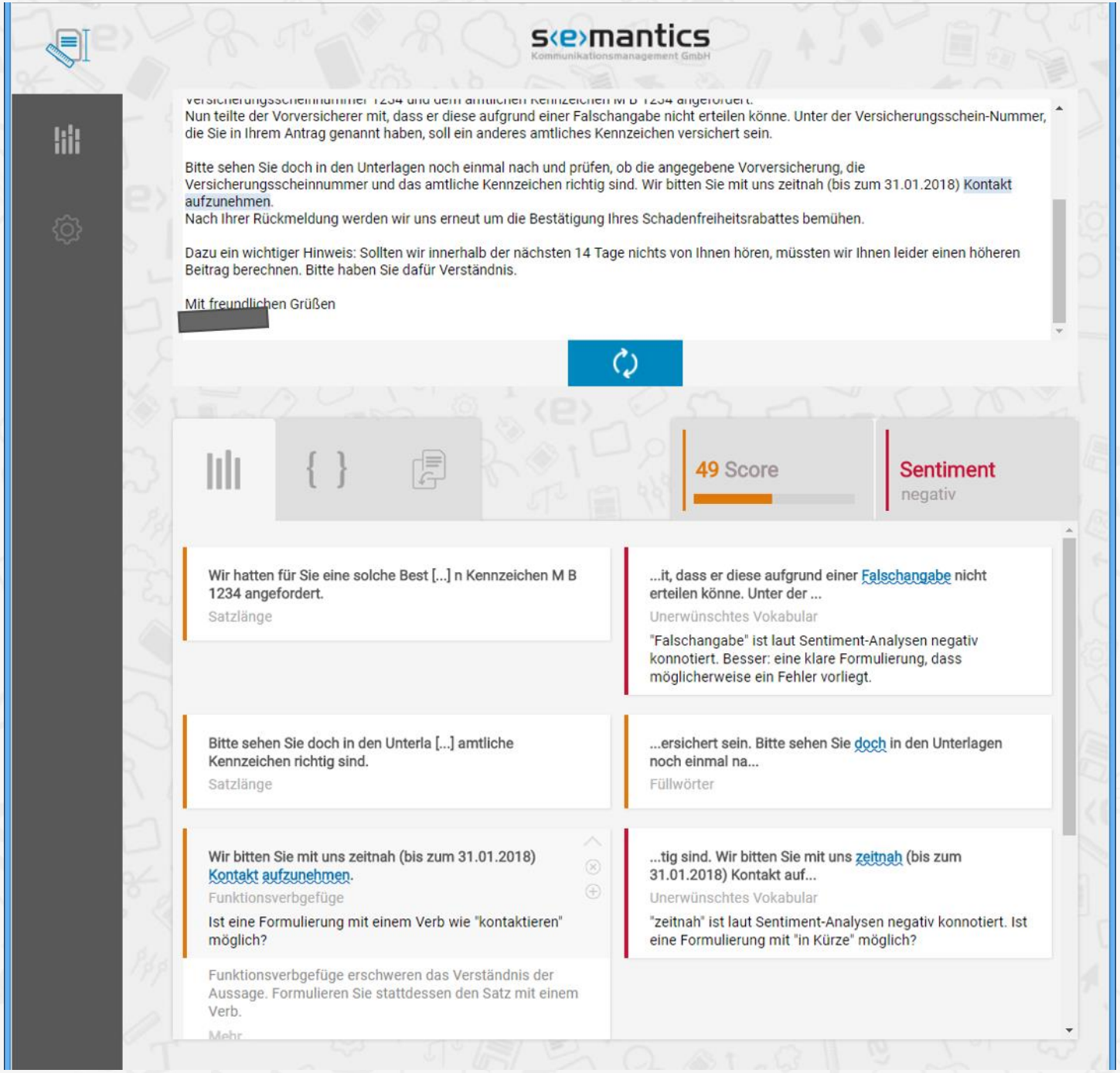
Library Assistant

Demoseite



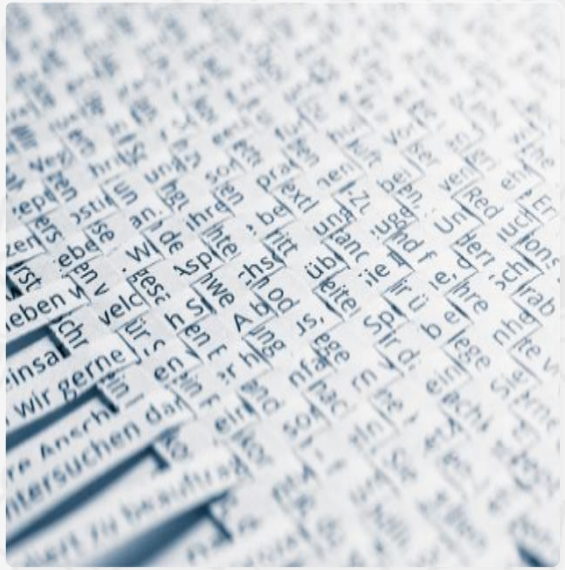
Künstliche Intelligenz

Segmentierung (99,98%), morphosyntaktische Annotationen, Semantik, Pragmatik



Erfolgreich schriftlich kommunizieren

Unser Ansatz



? Warum Sprachdienstleistungen mit semantics?

Unsere Sprachdienstleistungen...

... sind wissenschaftlich fundiert.

... sind handlungsorientiert ausgerichtet.

... ersetzen Geschmacksfragen durch Kriterien für verständliche Texte.

... sind individuell zugeschnitten auf unsere Kunden.

Unsere wissenschaftliche Basis

Grundlegende Erkenntnisse der Rhetorik und der Verständlichkeitsforschung

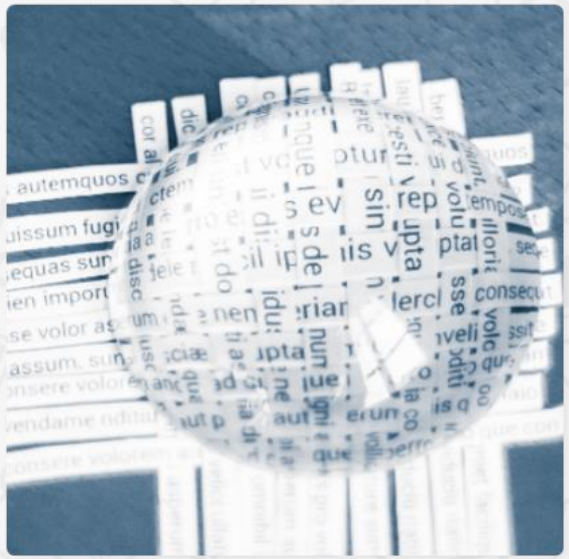
Korpus- und Textlinguistik

Sprechakttheorie: mit Texten zielgruppenorientiert handeln

Optimale Gliederung als Kern verständlicher Texte

Verständliche Unternehmens-sprache

Unsere Dienstleistungen



Neue verständliche Texte

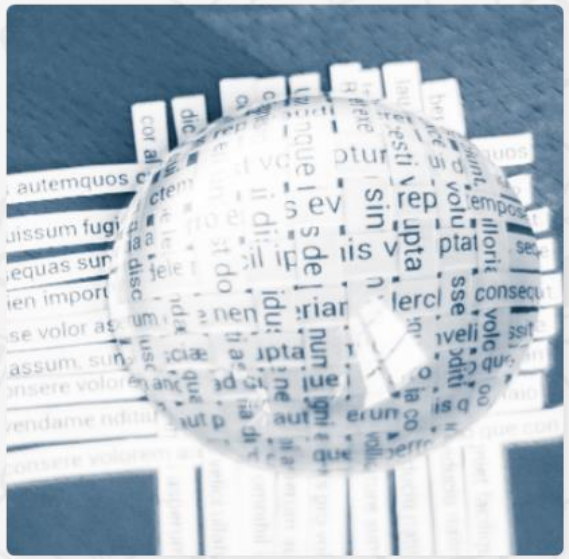
- Sprachschulungen
- Textoptimierung
- Sprachrichtlinie
- Fachkonzept
- Schriftgut-Analyse

Bisheriger Textbestand

Überarbeiteter Textbestand

Verständliche Unternehmens-sprache

Unsere Dienstleistungen



Schriftgut-Analyse

Basis: Exemplarische Textmenge

Einschätzung der zentralen Probleme

Konkrete Einschätzung + Vorgehensvorschlag

Fachkonzepte

Auf Basis der Ist-Analyse mit unterschiedlichem Fokus

Bestand an Dokumenten konsolidieren, reduzieren, übersichtlich ordnen, sprachlich überarbeiten

Sprachrichtlinien

Begutachtung bestehender Richtlinie

Richtlinie überarbeiten oder neu entwickeln

Für verschiedene Zielgruppen aufbereiten
(Checkliste, Leitfaden, …)

Textoptimierung

Definierte Textmenge, mit oder ohne Feedback

Vollständig durch semantics oder exemplarisch plus Schulungen

Mit oder ohne eigene Richtlinien

Enge Zusammenarbeit mit Fachstellen

Sprachschulungen

Eintägig oder zweitägig

Verständlich schreiben, individuelle Richtlinien umsetzen, Hausorthographie, …

Hoher Übungsanteil, eigene Texte überarbeiten

Analyse und Migration

Motivation und Kontext

> Wir würden das neue System sofort kaufen und einführen. Doch wir müssten 6.500 Vorlagen migrieren… nebst Textlogik, Datenbeschaffung, Workflow… :-/
> (Kunde eines CCM-Partners)

Die Kosten der Migration können wir durch einen **echten Mehrwert** rechtfertigen.

Methoden und Technik

Import, Analyse, Konsolidierung, Export, Workflow (Redaktion)

Automatisierung und **Erweiterung** des Vorhabens, dabei **Kostenreduktion**

Multi-Channel

Brief, Email, Portale, Apps, SMS, … Responsive, Interactive, …

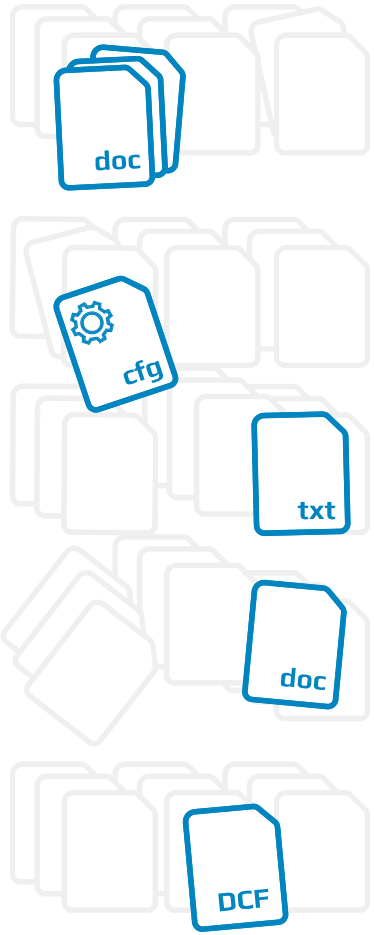
Textbestand **einrichten**, Kommunikation **planen**

Sprache

Zufriedenheit und Empfehlungen **erhöhen**, Nachfragen und Beschwerden **reduzieren**

Analyse und Migration

Import



Herkunftssystem

Beliebige Formate
(Word, Text, DCF, M/TEXT…)

Import Metadaten, Konfigurationen
(Workflow, Output)

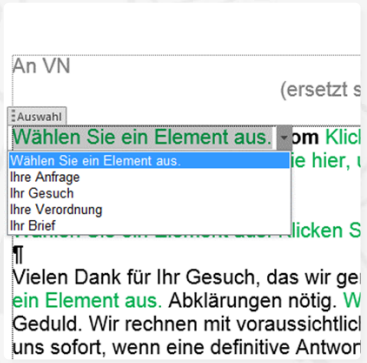
Korrektur manuelle Formatierungen
(Tabs, Silbentrennung, …)

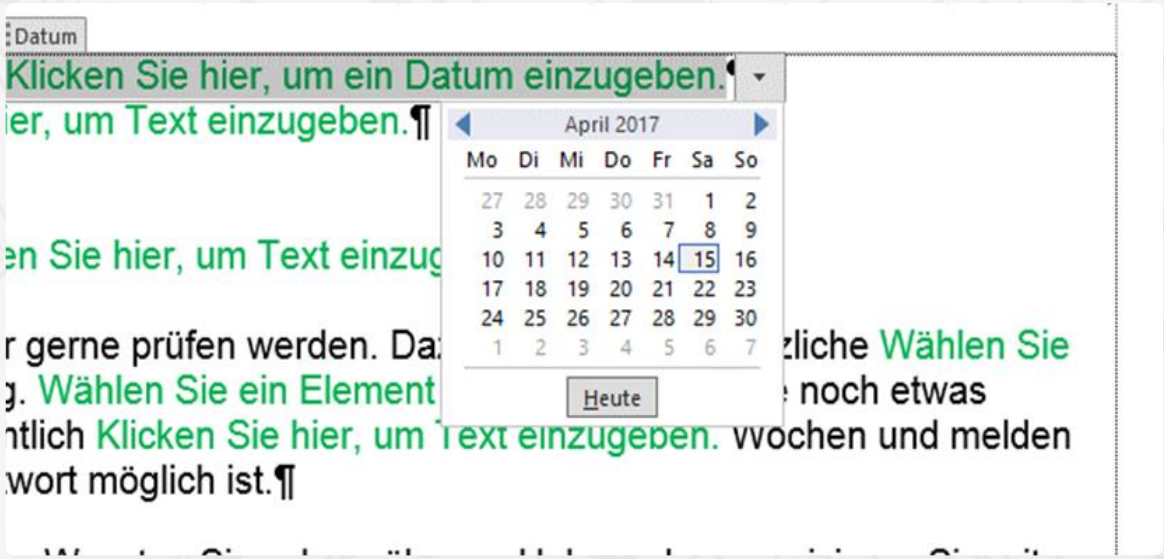
Extraktion Variablen, Logik

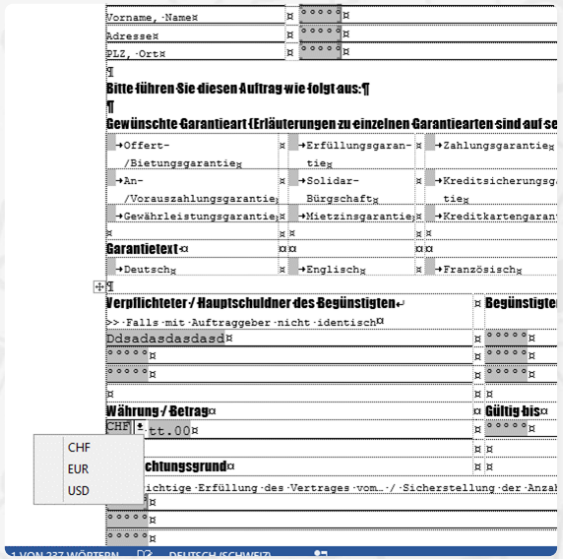
Organisation (Baum) anhand Metadaten *und* Schlagworten

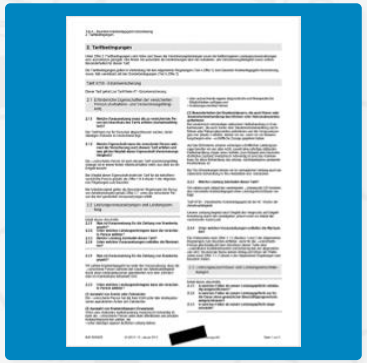
Analyse und Migration

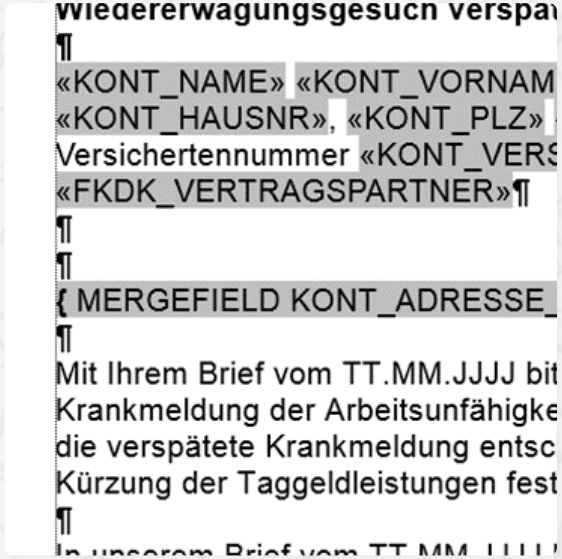
Import







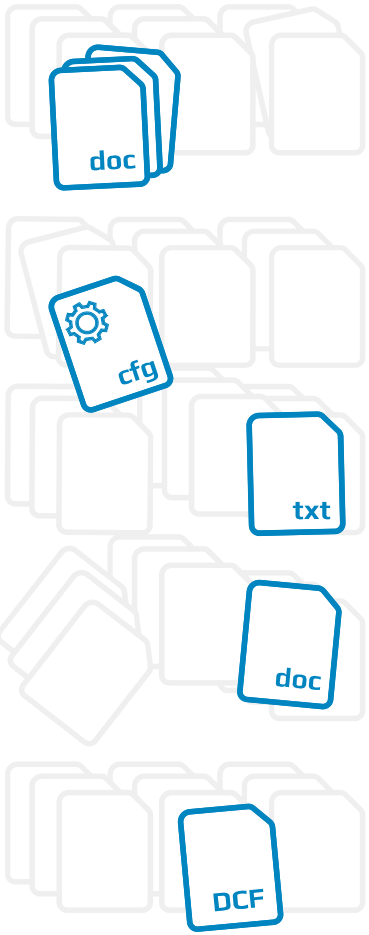




- Word-Format (andere ebenfalls möglich)
- Material moderat anspruchsvoll (Beispielprojekt)
- Metadaten aus Dateinamen, Textbereichen, Altsystem
- Herausforderungen durch „manuelles Layouting“
- „Vielfalt“ durch Variablen und Eingabeelemente… (Folgeprojekte)

Analyse und Migration

Import



Herkunftssystem

Beliebige Formate
(Word, Text, DCF…)

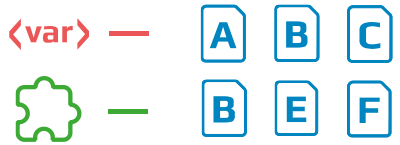
Import Metadaten, Konfigurationen
(Workflow, Output)

Korrektur manuelle Formatierungen
(Tabs, Silbentrennung, …)

Extraktion Variablen, Logik

Organisation (Baum) anhand Metadaten *und* Schlagworten

Querverweise, Verwendungen, Abhängigkeiten

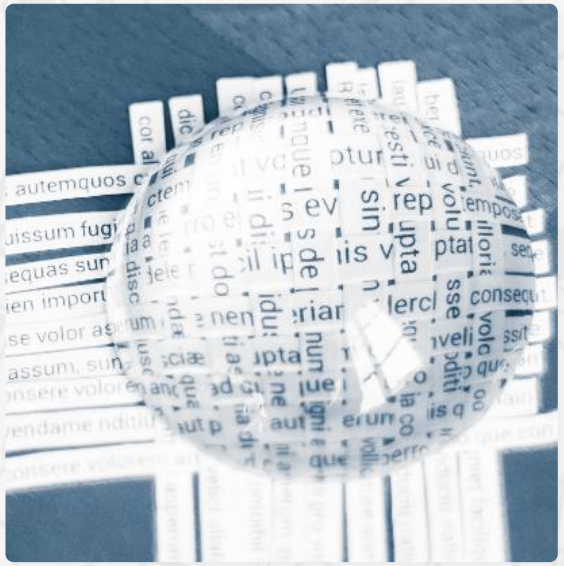


Logik

Suchindex

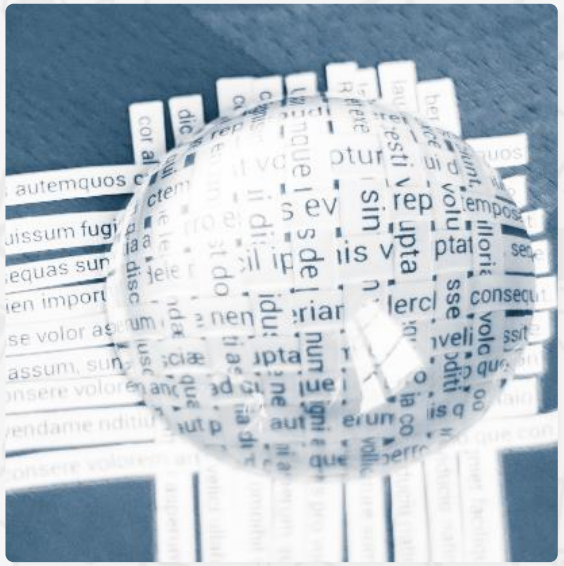
Aufbereitung

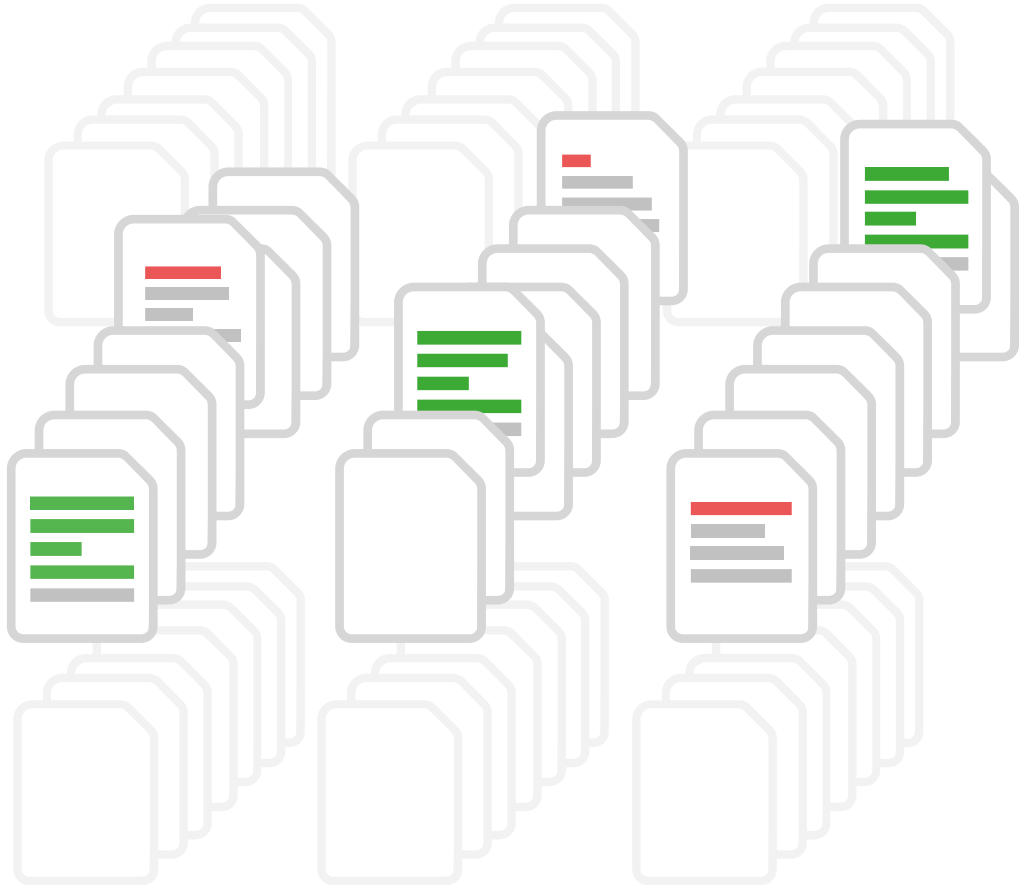
Maschinelle Analyse



Aufbereitung

Maschinelle Analyse





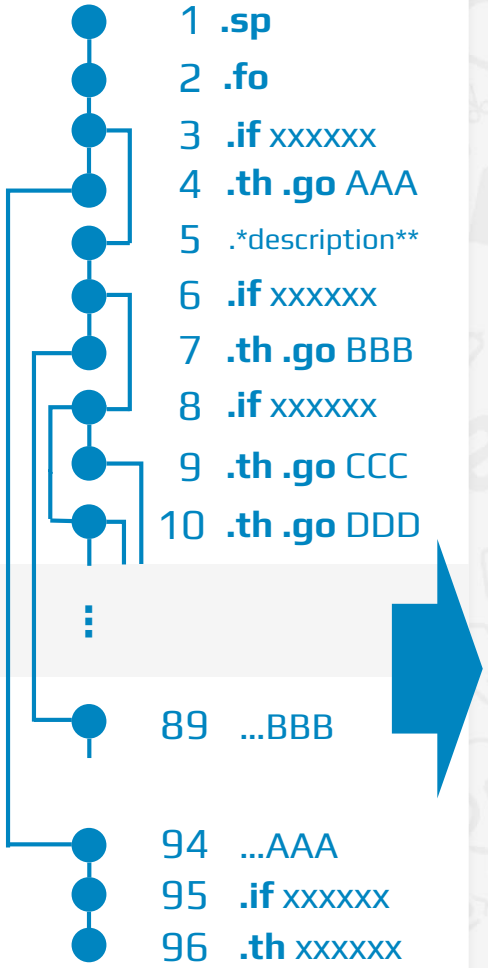
Identifikation von

- Identischen und *ähnlichen* Passagen
- Paraphrasierungen
- Allen Phänomenen bis ins Detail

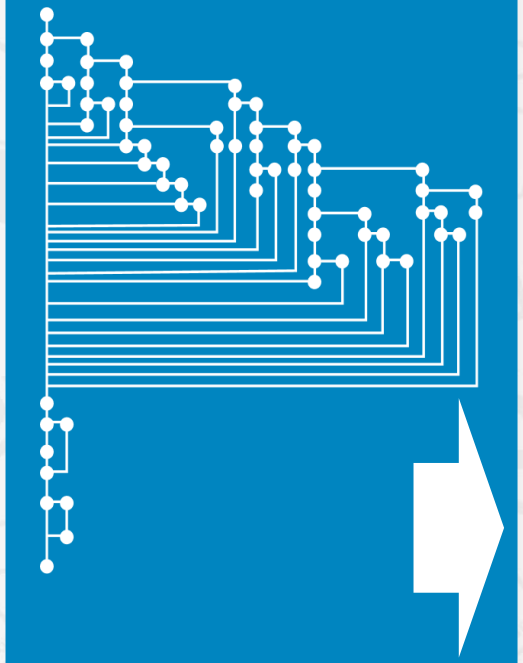
Analyse imperativer Formate, z.B. dcf

Aufbau Ablauf-Graph

Überführung nach deklarativ



z.B. dcf Baustein



Prüfung ob deklarative Darstellung möglich (if/then/else ohne goto)

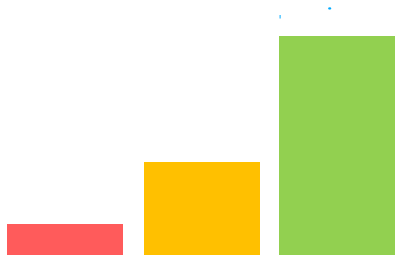
Übersetzung in deklaratives XML

Ablauf-Graph

XML Ergebnis

Analyse imperativer Formate, z.B. dcf

Interaktive Mikro-Analyse,
Erfassung Verteilung



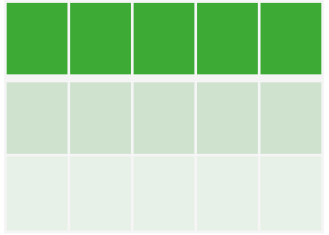
1	^	1.	^	Tarif		
2	^	2.	^	Tarifbeitrag		
3	^	3.	^	Beitrag		
4	.tp	6mm	26mm	57mm	100,5mm	150mm
5	.im	...				
6	^Tarif	^ Produkt	^ Vergütung	^ Besonderes	^ Haftung	
7	^&var1.	^&var2.	^&var3.	^&var4.	^&var5.	

- Tokenizing, Analyse-Datenbank
- Erfassung aller Phänomene und derer Verteilung

⇒ Aufwandschätzung möglich

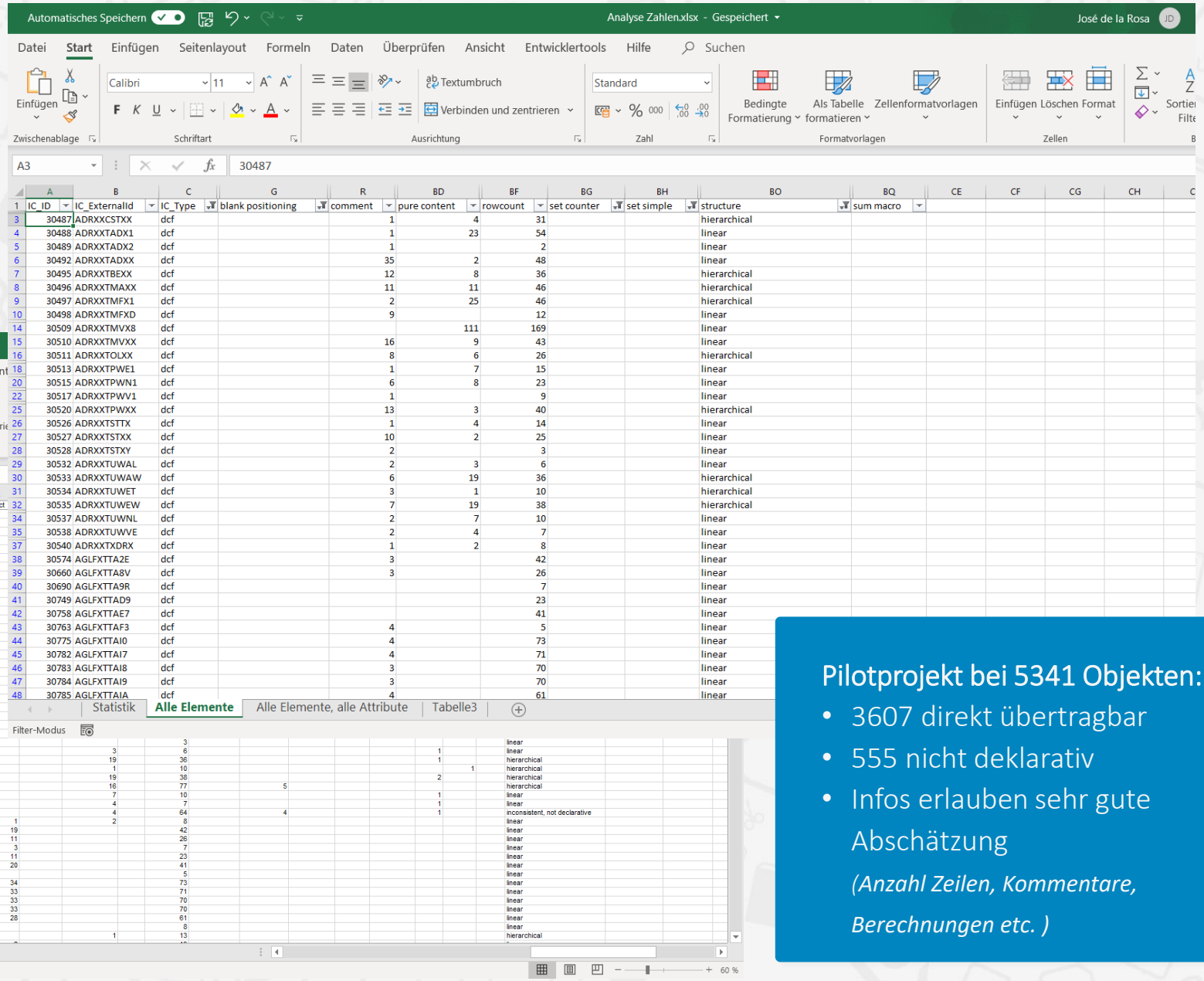
- Erkennung von Mustern,
  Abbildung auf Formate im Zielsystem
  (Beispiel: Tabulatoren für Style vs. Tabelle)

1. Tarif
2. Tarifbeitrag



Grundlage zur Schätzung

Automatisierung und manuelle (Nach-)Arbeiten



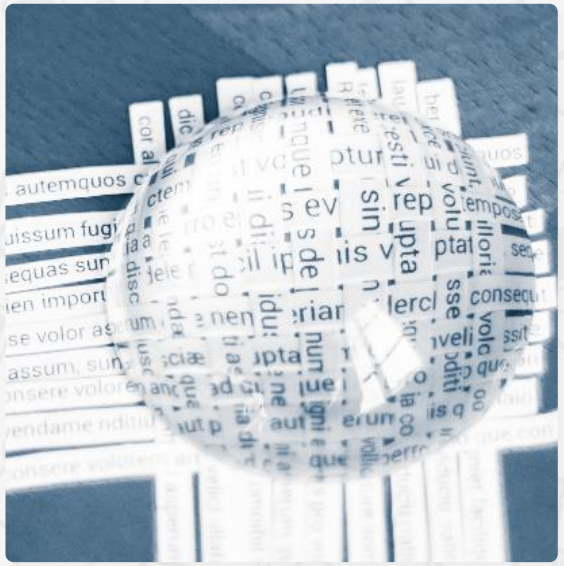
Pilotprojekt bei 5341 Objekten:

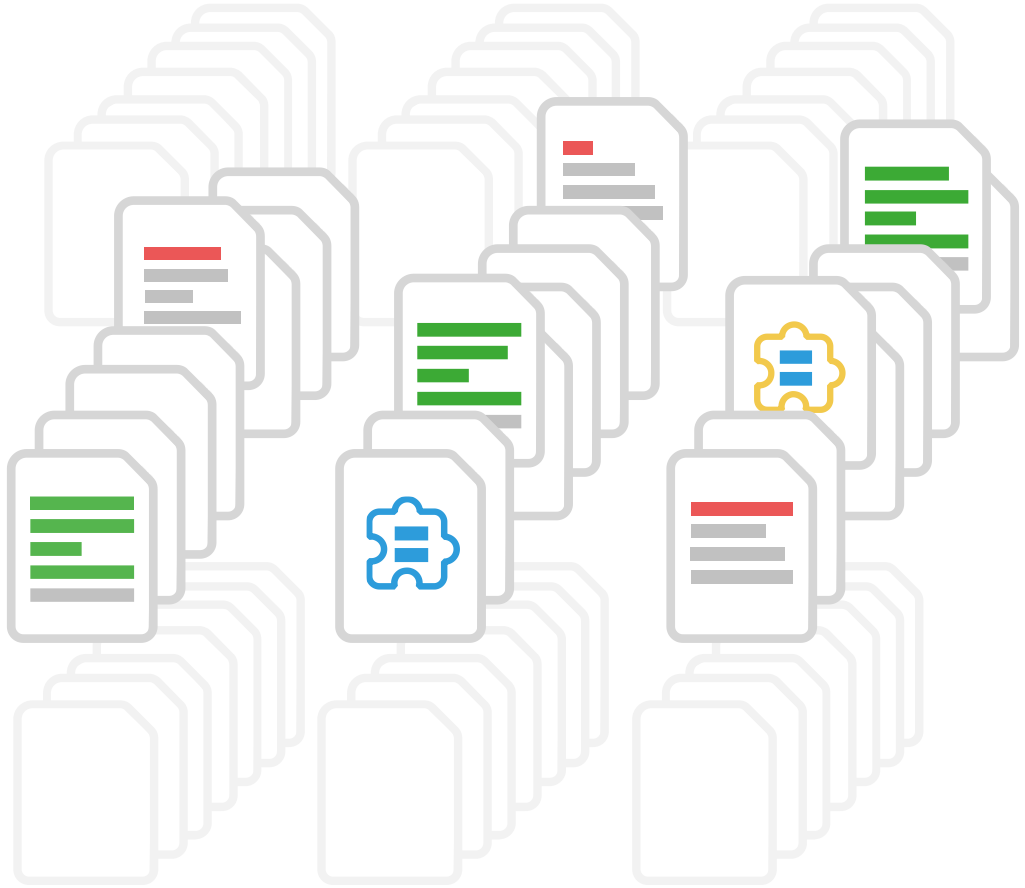
- 3607 direkt übertragbar
- 555 nicht deklarativ
- Infos erlauben sehr gute Abschätzung

  (Anzahl Zeilen, Kommentare, Berechnungen etc.)

Aufbereitung

Weitere Analyse



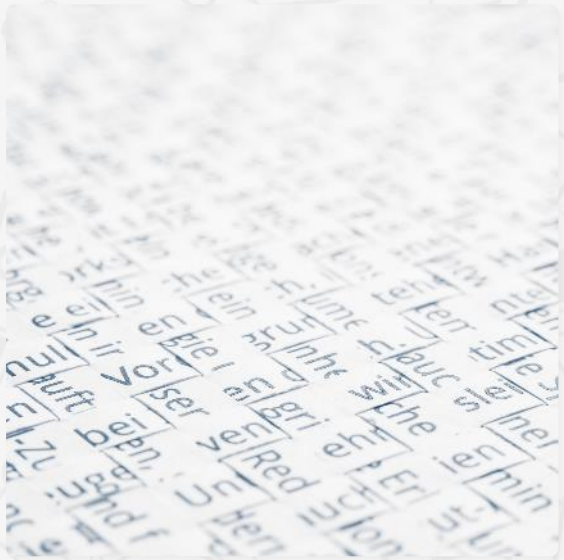


Identifikation von

- Identischen und *ähnlichen* Passagen
- Paraphrasierungen
- Allen Phänomenen bis ins Detail
- Redundanten Textbausteinen

Aufbereitung

Automatisierte Konsolidierung



Identifikation von

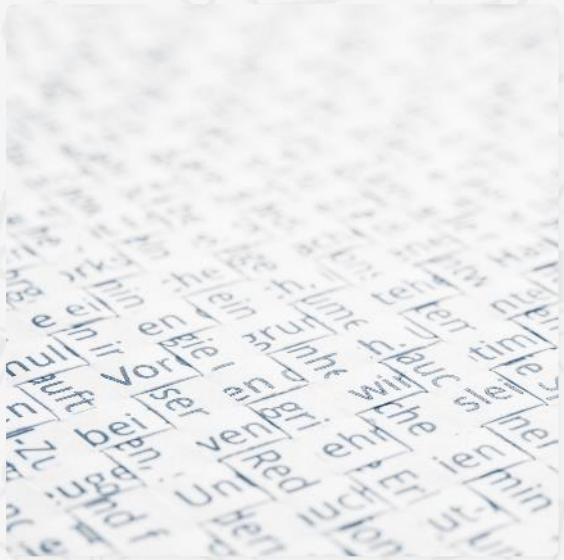
- Identischen und *ähnlichen* Passagen
- Paraphrasierungen
- Allen Phänomenen bis ins Detail
- Redundanten Textbausteinen

Automatisierung

- Auslagerung Textbausteine

Aufbereitung

Automatisierte Konsolidierung



Identifikation von

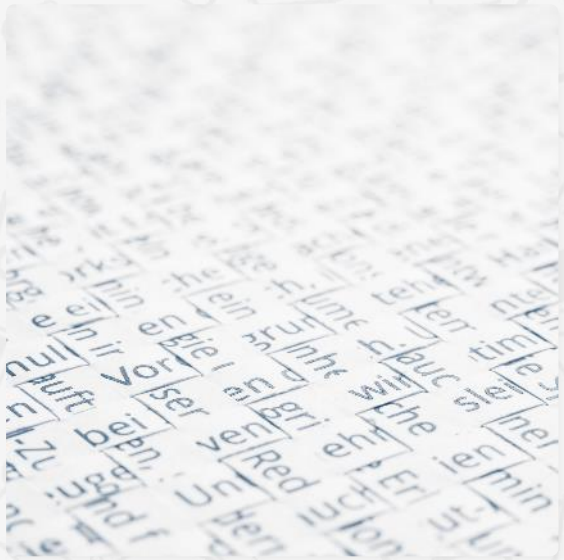
- Identischen und *ähnlichen* Passagen
- Paraphrasierungen
- Allen Phänomenen bis ins Detail
- Redundanten Textbausteinen

Automatisierung

- Auslagerung Textbausteine
- Harmonisierung Abweichungen

Aufbereitung

Optional:
Sprachliche Analyse



Identifikation von

- Identischen und *ähnlichen* Passagen
- Paraphrasierungen
- Allen Phänomenen bis ins Detail
- Redundanten Textbausteinen

Automatisierung

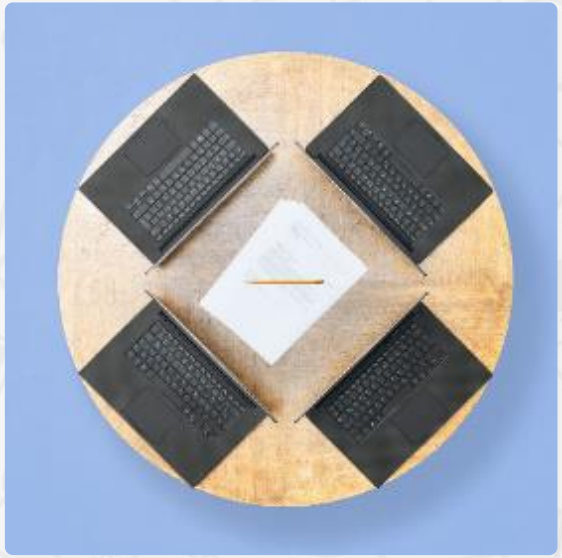
- Auslagerung Textbausteine
- Harmonisierung Abweichungen
- Ersetzung von Doubletten

Sprachprüfung

Verständlichkeit, Stil, Wording, Sentiment u.ä.

Aufbereitung

Redaktionelle Optimierung



- Nutzung Redaktionsprozesse zur Kollaboration
- Finale Organisation des Bestandes
- Optimierung der Textlogik
- Test Variantenbriefe
- Versionsführung, Änderungsverfolgung, Dokumentation
- Begleitende Statistik, Projektfortschritt
- Strukturierte Projektunterstützung

Migriert

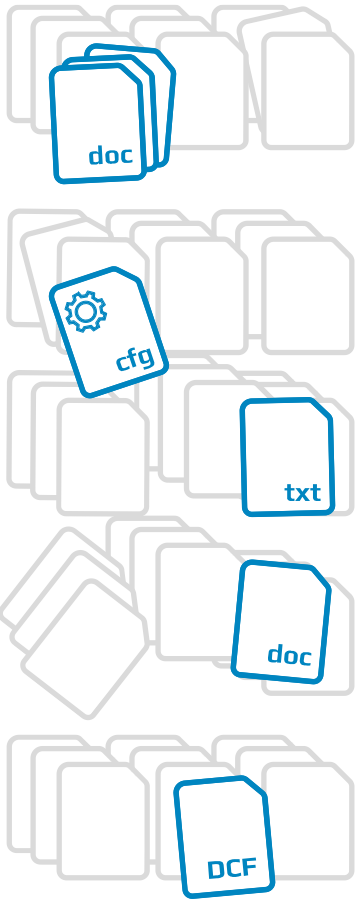
Bearbeitung

Qualitätssicherung

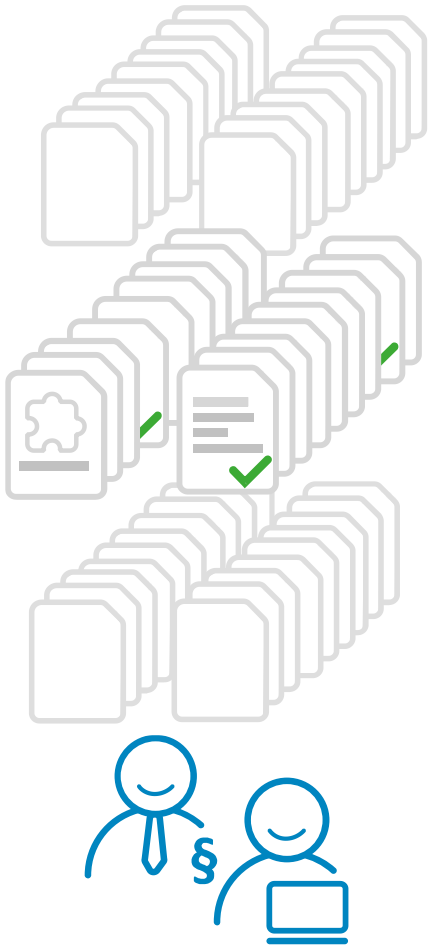
Obsolet

Export

Erzeugung spezifisches Format



Import

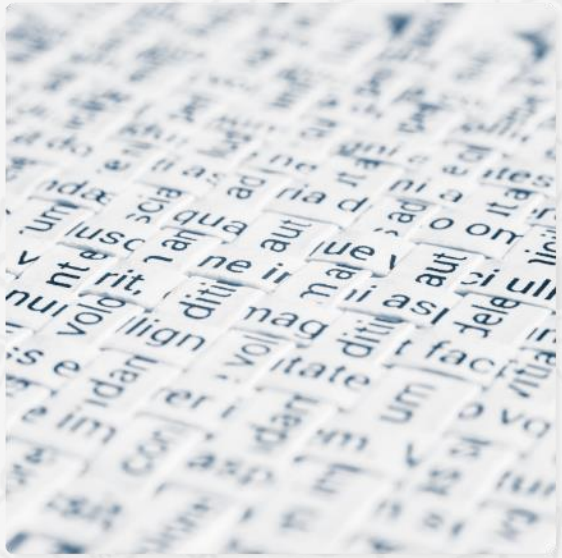


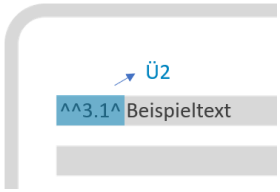
Aufbereitung

Export

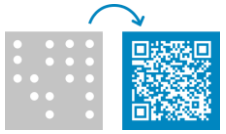
Export

Weiteres Mapping,
Maschineller Export



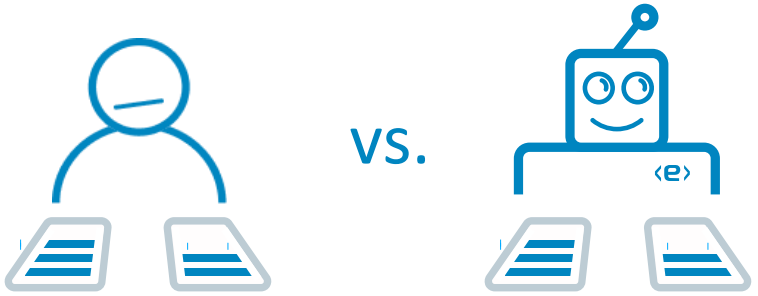


Mapping Styles



- Bausteine
- Variablen
- einfache syntaktische Strukturen

Definition und Mapping
Briefklassen anhand Textbausteine



Prototypische Umsetzung repräsentative Vorlage
10h vs. 0,5h

Extraktion Dokumentenkörper
anhand xpath-Regeln

Besten Dank für Ihre Aufmerksamkeit!

Dipl.-Ing. José de la Rosa

semantics
Kommunikationsmanagement
GmbH
Viktoriaallee 45
52066 Aachen

+49 241 89 49 89 29
j.delarosa@semantics.de
www.semantics.de