DOXNET®
THE DOCUMENT X-PERTS NETWORK

Dimensionen der Textverständlichkeit

Dipl.-Ing. José Manuel de la Rosa Govantes
Björn Eßer M.A.

s<e>mantics
Kommunikationsmanagement GmbH

Über uns

semantics Kommunikationsmanagement GmbH

- Gründung 2000 als Spin-off der RWTH Aachen (Prof. Dr. Christian Stetter)
- Interdisziplinäres Team aus z.Zt. 35 Informatikern, Linguisten, Textwissenschaftlern und Psychologen
- Verbindung von Forschung und Praxis zur Textverständlichkeit und Texttechnologie
  - Empirische Studien
  - Forschung in Linguistischer Informatik
  - Sprachschulungen, -richtlinien, -analysen, Workshops etc.
  - Technische Lösungen für große Textbestände (Analyse, Management, Optimierung)
  - Kooperationen mit Integrationspartnern und Anbietern von Drucklösungen
- Ferner: Digitalisierungs- und Recherchesysteme für Bibliotheken, Archive und Museen (Marktführer D-A-CH)

Relevanz und Ziel

„Man kann nicht nicht kommunizieren“ *(Paul Watzlawick)*

Man kann es nur gut oder schlecht tun.

- Ohne Kommunikation kein Geschäft
- Besonders relevant bei Versicherungen und Finanzdienstleistern
  Produkte nicht gegenständlich, sondern durch Vertragswerke realisiert
  Sehr große Anzahl von Kunden; wichtige schriftliche Kommunikation
- Steigende Qualitätsansprüche der Kunden
  Nicht zuletzt durch Digitalisierung und Individualisierung, Recherchemöglichkeiten etc.

⇒ *Nachfrage- und Beschwerdeaufkommen reduzieren*
⇒ *Empfehlungen, Kundenbindung und -zufriedenheit erhöhen*
⇒ *Prozesse optimieren,* ***Automatisierung erhöhen***

Beispiel

„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.

Beispiel – grammatische Analyse

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

Beispiel – überarbeitet

Der Versicherer gewährt dem Versicherungsnehmer Deckung für Haftpflichtansprüche, die gegen ihn geltend gemacht werden.	**Hauptaussage**
Haftpflichtansprüche unterliegen gesetzlichen Haftpflichtbestimmungen privatrechtlichen Inhalts.	**Verweis auf rechtliche Basis**
Sie bestehen dann, wenn der Versicherungsnehmer für Personen- oder Sachschäden verantwortlich gemacht werden kann.	**Kausale Bestimmung**
Dieser Schaden muss während der Versicherungszeit entstanden sein, damit er unter die gewährte Deckung fällt.	**Bedingung**

Was erschwert das Verstehen dieses Beispiels?

1. Informationsmenge und -aufnahme schwierig

Proposition: Aussage eines Satzes

Der Versicherer
gewährt [...] Versicherungsschutz [...].

Referenz
Prädikation

Ich beziehe mich auf etwas in der Welt (Referenz) und sage darüber etwas aus (Prädikation).

Kognitive Konstruktion

Der Hund ist schwarz.

The dog is black.

Der Firn schimmert cyan.

Voraussetzungen:

- der Sprache mächtig sein
- lesen können
- bereits etwas über die Welt wissen

Kognitive Konstruktion

John war auf dem Weg zur Schule. Er machte sich Sorgen über die Mathe-Stunde. Er hatte Angst, er könne die Klasse nicht im Griff haben.

Mentale Modelle werden während der Rezeption ständig abgewandelt (on-line).

Was erschwert das Verstehen dieses Beispiels?

Der Versicherer gewährt dem Versicherungsnehmer **1. Aussage**
Versicherungsschutz *für den Fall, dass er* *wegen eines* **2. Aussage**
während der Wirksamkeit der Versicherung
eingetretenen Schadenereignisses, **3. Aussage**
das den Tod, die Verletzung oder
Gesundheitsschädigung von Menschen **4. Aussage**
(Personenschaden) oder die Beschädigung oder **5. Aussage**
Vernichtung von Sachen (Sachschaden) zur Folge
hatte, *für diese Folge*
aufgrund gesetzlicher Haftpflichtbestimmungen **6. Aussage**
privatrechtlichen Inhalts
von einem Dritten auf Schadenersatz in Anspruch
genommen wird.

Und: mehr als vier verschachtelte Hierarchien

Was erschwert das Verstehen dieses Beispiels?

1. Informationsmenge und -aufnahme schwierig

2. Logische Reihenfolge (grammatische Kohärenz) wird unterbrochen

3. Thematische Abfolge (thematische Kohärenz) nicht erfassbar

Kohärenz

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.

Thematische Kohärenz

Das Schiff	**Thema 1**
sank im Bermudadreieck.	**Rhema 1**
Das Bermudadreieck	**Thema 2**
liegt zwischen den Bermuda-Inseln, Puerto Rico und Florida.	**Rhema 2**
Dort	**Thema 3**
sollen schon viele Schiffe und Flugzeuge spurlos verschwunden sein.	**Rhema 3**

Grammatische Kohärenz

Der Baum steht neben der Scheune.
Er verliert allmählich die Blätter.

Mittlerweile war es Nacht geworden.
In der Dunkelheit schärfte sich sein Gehörsinn.

Rekurrenz

Er war krank. Deshalb blieb er zu Hause.

Da sie nicht zu Hause war, bestellte er sich eine Pizza und schaute einen Film.

Konnexion

Was erschwert das Verstehen dieses Beispiels?

1. Informationsmenge und -aufnahme schwierig
2. Logische Reihenfolge (grammatische Kohärenz) wird unterbrochen
3. Thematische Abfolge (thematische Kohärenz) nicht erfassbar
4. **Funktion bleibt unklar (pragmatische Kohärenz)**

Sprechakt

Ich komme morgen.

Können Sie das Fenster schließen?

Eine Äußerung besteht aus:

- dem Inhalt (Lokution)
- der Funktion oder dem Zweck (Illokution)
- ggfs. einer Wirkung, die beim Leser erreicht wird (Perlokution)

Pragmatische Kohärenz

Situation 1:

Sie am Telefon: *Ich brauche Hilfe beim Tapezieren.*

Er: *Ich bin in einer Stunde da!*

Situation 2:

Mutter am Telefon zum Kind:

Hast Du dein Zimmer aufgeräumt? Ich bin in einer Stunde da!

Zurück zum Beispiel:
Was erleichtert das Verstehen?

Der Versicherer gewährt dem Versicherungsnehmer Deckung für Haftpflichtansprüche, die gegen ihn geltend gemacht werden.	**Hauptaussage**
Haftpflichtansprüche unterliegen gesetzlichen Haftpflichtbestimmungen privatrechtlichen Inhalts.	**Verweis auf rechtliche Basis**
Sie bestehen dann, wenn der Versicherungsnehmer für Personen- oder Sachschäden verantwortlich gemacht werden kann.	**Kausale Bestimmung**
Dieser Schaden muss während der Versicherungszeit entstanden sein, damit er unter die gewährte Deckung fällt.	**Bedingung**

Im Beispiel – Grammatische Kohärenz

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

Rekurrenz

Haftpflichtansprüche unterliegen gesetzlichen Haftpflichtbestimmungen privatrechtlichen Inhalts.

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

Rekurrenz

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

Im Beispiel – Thematische Kohärenz

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.

Thema 1
Rhema 1
Thema 2
Rhema 2
Thema 2
Rhema 3
Thema 3
Rhema 4
Rhema 5

Im Beispiel – Pragmatische Kohärenz

… ist erreicht, wenn das **Leistungsversprechen** des Versicherers mit seinen Einschränkungen und Besonderheiten **verstanden** wurde.

Voraussetzungen für Textverständlichkeit

- Kohärenz auf der grammatischen Ebene
- Kohärenz auf der thematischen Ebene
- Kohärenz auf der pragmatischen Ebene

Textverständlichkeit geht damit über die Lesbarkeit eines Textes weit hinaus und lässt sich nicht einfach messen.

Beispiele Lesbarkeitsindizes

1. Amstad-Formel („Reading Ease“)

$$RE = 180 - SL - 58{,}5 * WL$$

2. Lix („Lesbarkeitsindex“)

$$Lix = \frac{W}{S} + \left(\frac{W6}{W}\right) * 100$$

3. SMOG („Simple Measure of Gobbledygook“)

$$SMOG = 1{,}043 * \sqrt{30 * \left(\frac{W3}{S}\right)} + 3{,}1291$$

4. Erste neue Wiener Sachtext-Formel

WSTF = 0,1935*W3% + 0,1672*SL + 0,1297*W6% - 0,0327*WE% - 0,875

5. Hohenheimer Verständlichkeitsindex

Formeln 1-4, zusätzlich eingerechnet:

- mittlere Satzlänge und Satzteillänge in Wörtern
- Anteil der Satzteile mit mehr als zwölf Wörtern
- Anteil der Sätze mit mehr als zwanzig Wörtern
- …

S = Anzahl Sätze
SL = mittlere Satzlänge

W = Anzahl Wörter
WL = mittlere Wortlänge

WE = einsilbige Wörter
W6 = mind. 6 Buchstaben

W3 = mind. 3 Silben

Relevante Dimensionen

DOXNET® THE DOCUMENT X-PERTS NETWORK

Unternehmen

Große Organisationen, Viele Beteiligte, Spannungsfeld durch Kompetenz-gemeinschaften (fachlich, juristisch, ...)

Defensives Verhalten, „Invisible Hand“

Organisationstheorie, Schreibforschung

Organisierung

Technik

Verfahren zur Massen-produktion, Paradigmen beeinflussen Textqualität

Heterogene Systeme

Bislang Fokus: Optimierung Daten-mengen, Durchlauf-zeiten, Portokosten, ...

Unterstützung

Text

Interne Eigenschaften: Wort-, Satz- u. Text-ebene

Ästhetik

Text-, Korpus-, strukturale Linguistik

Empirische Studien

Sprachqualität

Dialog

Gelingen der Kommunikation

Sprechakttheorie

Kanal

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

Mediale Eigenschaften

Multi-Channel

Empfänger

Zielgruppe, Vorwissen, Motivation, ...

Kognitionswissenschaft, Neurolinguistik, Psychologie, Soziologie

Marktanalysen, Umfragen, Beteiligung der Verbraucher

Individualisierung

Ziel: Beherrschbarkeit individueller Massenkorrespondenz

Outputmanagement (stark vereinfacht)

DOXNET® THE DOCUMENT X-PERTS NETWORK

Redaktion
- Organisation
- Entwicklung
- Abstimmung
- Qualitäts-sicherung
- Dokumentation
- Team-Work
- Recherche
- …

Ressourcen
- Textbausteine
- Vorlagen
- Variablen
- Schriften
- Logos
- Unterschriften
- Formulare
- …

Erstellung
- Manuelle Dokumente
- Batch-verarbeitung
- Daten-beschaffung
- Dokumentlogik
- …

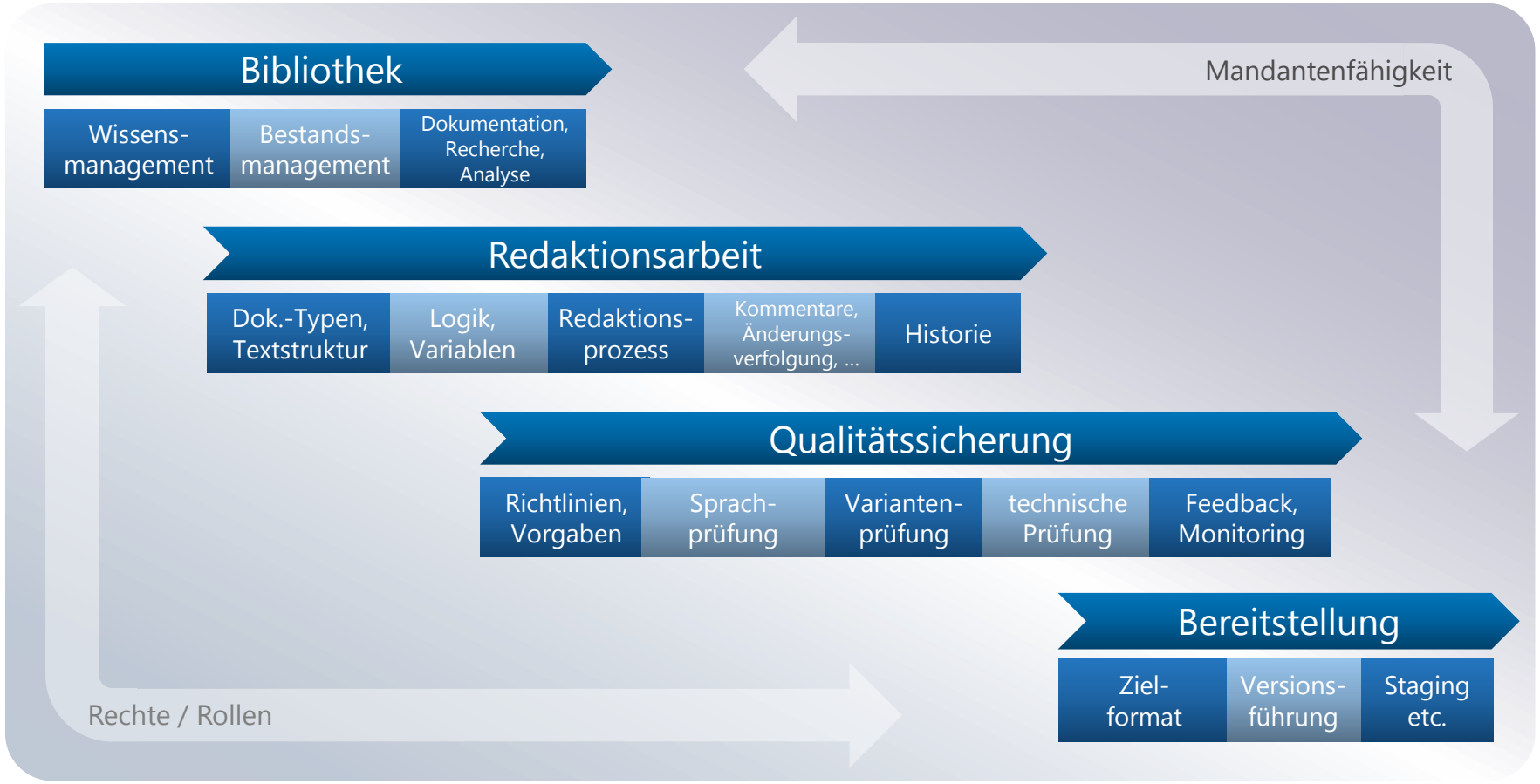
Druck
- Komposition
- Formatierung
- Rendering
- Multi-Channel-Output
- …

Nachverarb.
- Bündelung
- Sortierung
- Beilagen
- Frankierung
- Verfolgung
- Versand
- Archivierung
- …

Feedback, Änderungsverfolgung

Lösung des systemischen Problems durch ergänzende Technik

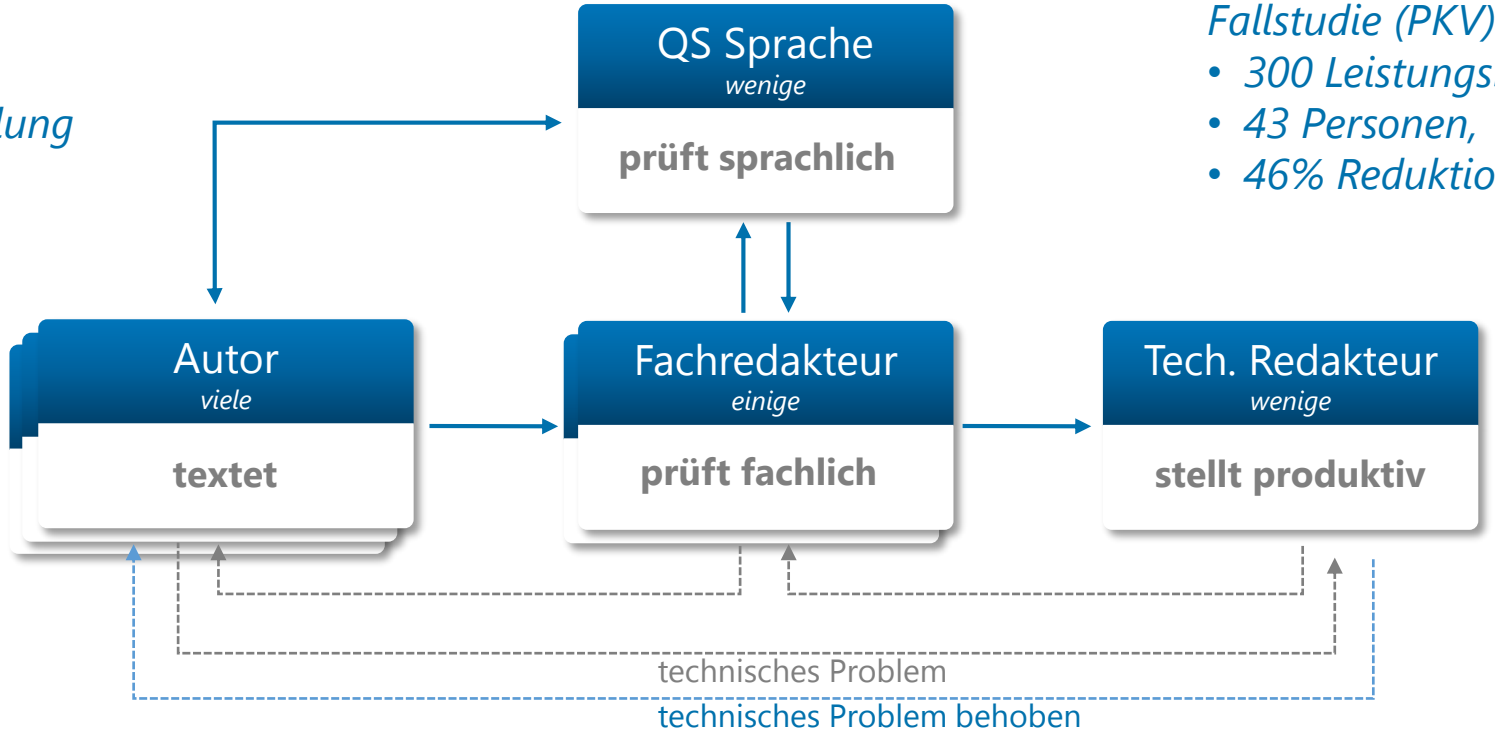
Sprachorientierte Redaktion



Beherrschbarkeit durch systematische Textentwicklung

Sprachlicher Redaktionsprozess
(optimal: ohne Medienbrüche)

Möglich auch:
- *4-Augen-Prinzip*
- *Tandem-Entwicklung*
- *…*



Fallstudie (PKV):
- *300 Leistungsbriefe, 80% Abdeckung*
- *43 Personen, 15 Referate, 1 Linguist*
- *46% Reduktion Beschwerdequoten*

Geführte Textentwicklung – schlanke Prozesse

Varianten statt Textbausteine

Variantenbrief

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

Hauptteil

Erstattung?

Ja

Nein

Warum nicht?

Schluss, Abbinder, Fuß, ...

Input durch Datenbeschaffung:
Erstattung = Nein

Komposition

Auswahl durch Sachbearbeiter:
Begründung = C

Individuelle Briefe

Paradigmenwechsel: Von Textbausteinen zu Variantenbriefen

- Verlust der Qualität über die Zeit: als schlüssiges Schreiben in Gruppe entworfen, führt Verwendung in anderen Kontexten zu Verlust von Kohärenz, nicht passenden Übergängen und semantischen Gehalten
- Verlust des Blicks für wichtigste rhetorische Größe: Gliederung und Gesamtstruktur
- Verlust der Übersicht: meist keine 100%tige Sicherheit ob der Verwendungen (z.B. Batch-Programmierung)
- Sinnvoller Einsatz Bausteine weiterhin möglich für übergreifende Köpfe/Füße, Meta-/Steuerbausteine, Produktionskomposita bei Policen, AVB, ... (statische Referenzen und/oder dynamisch per Datenbeschaffung)
- Auflösung technischer Abhängigkeiten: Beliebige Verschachtelung logischer Strukturen (Variante, Iteration, ...); Keine spezifische Programmierung je Vorlage o.ä.

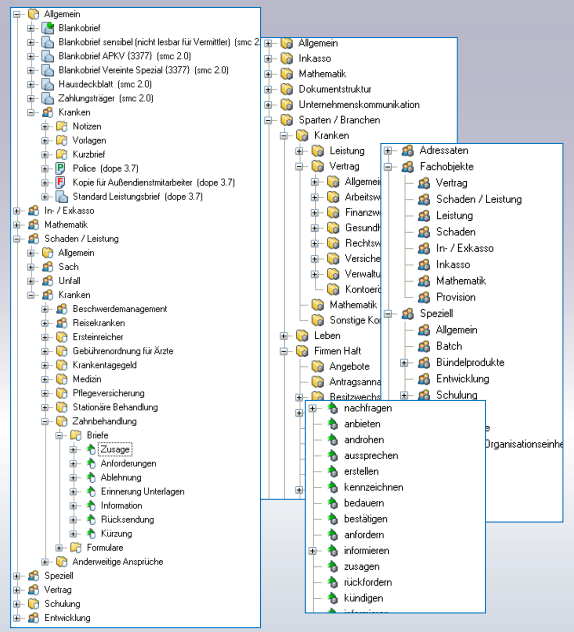
Nutzen: Beherrschbarkeit großer Bestände

Automatische Analyse und Kategorisierung,

Ermittlung Redundanzen und Dubletten

(ferner: Sprachprüfung)

Manuelle Nacharbeiten

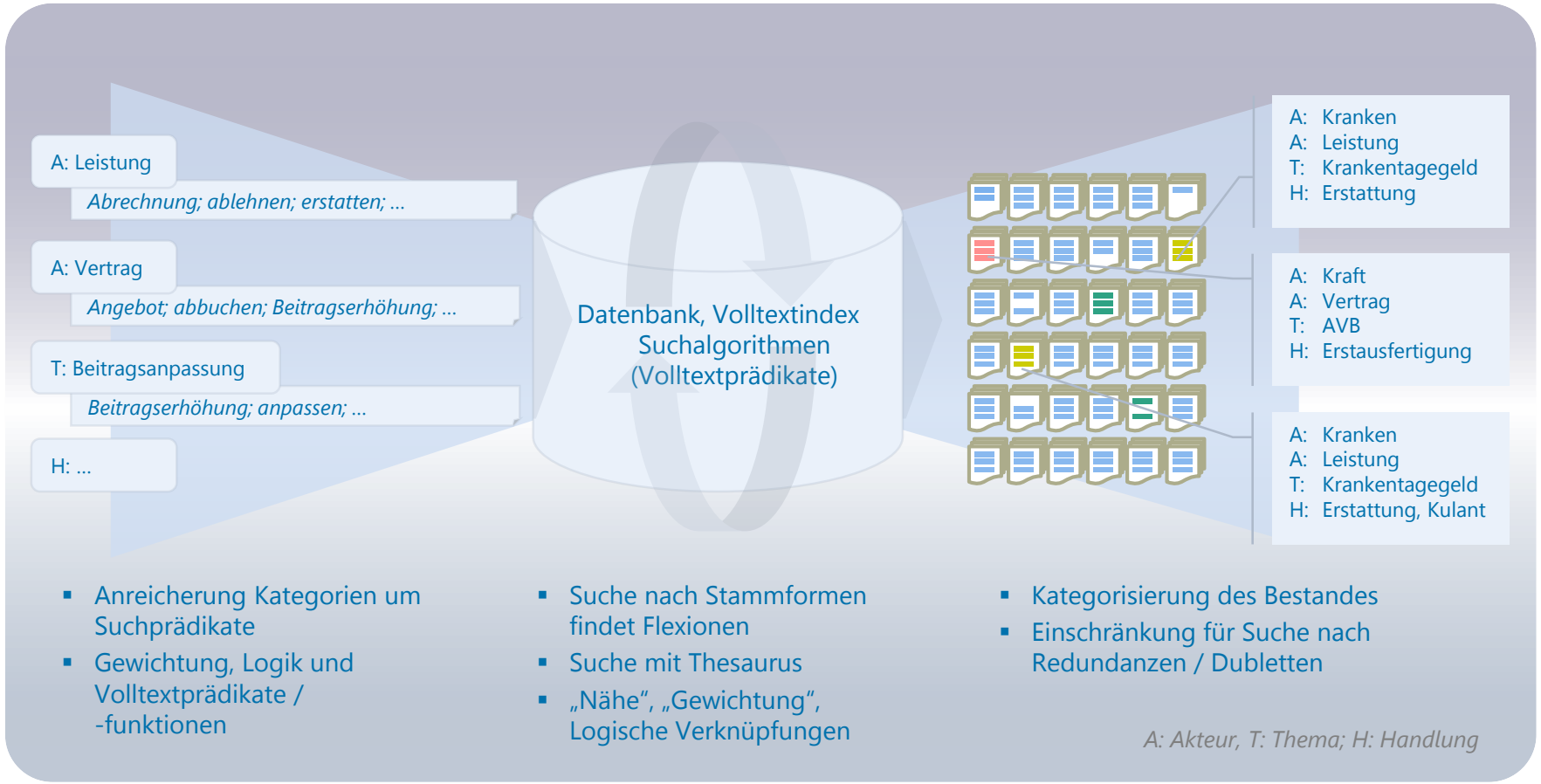


- Umfangreicher Textbestand (1,9 Mio.)
  (Fallstudie deutscher Versicherungskonzern)
- Evtl. Metadaten (Typisierung) vorhanden
- Weitestgehend ohne Kategorisierung

- Wohlstrukturierte Unternehmensbibliothek
- Langfristige Qualitätssicherung durch Kategoriensystem

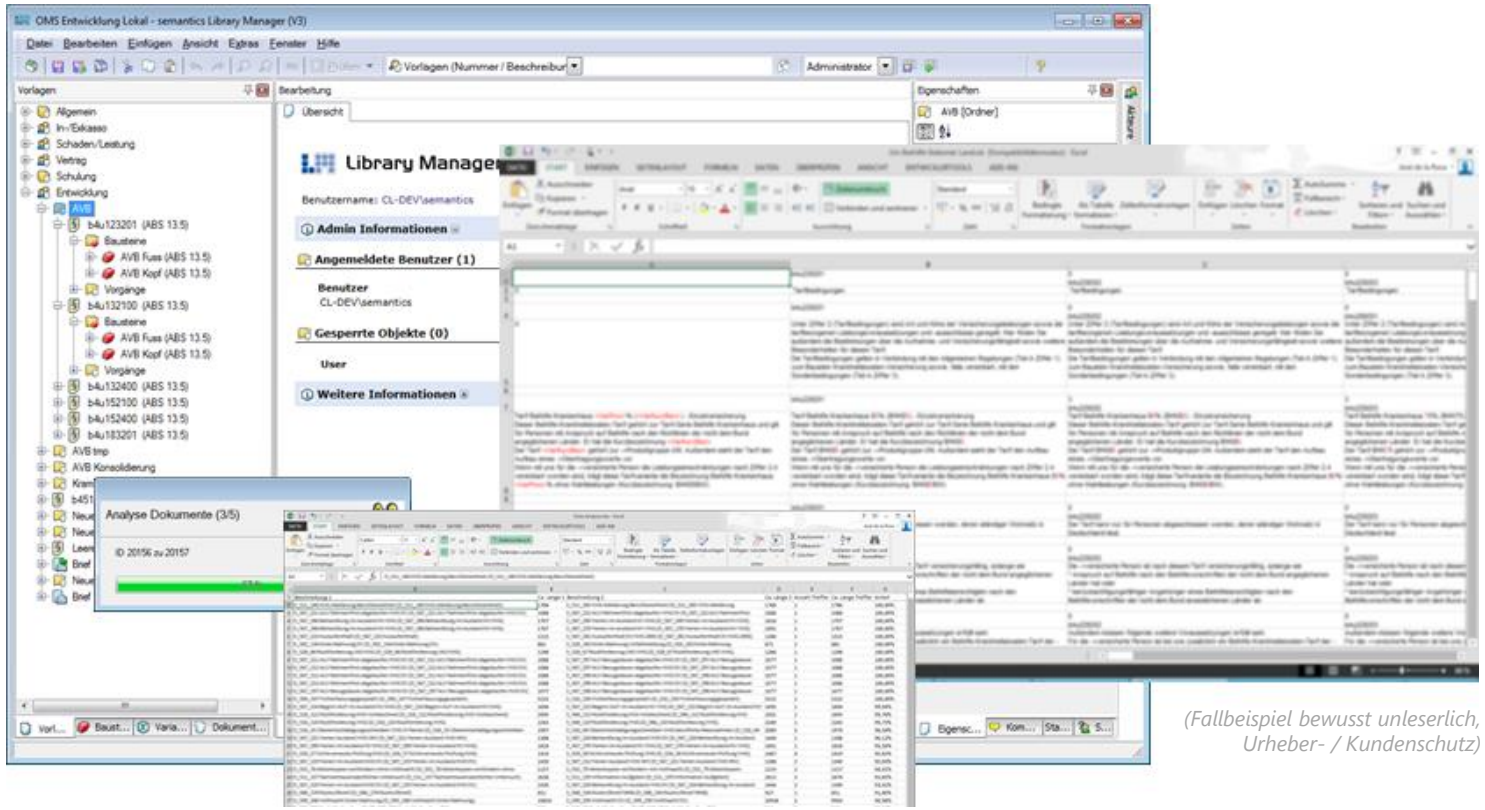
Ziel: Automatisierte Kategorisierung

Nutzen: Automatisierte Organisation



Textdatenbank mit Volltextindex zur Organisation

Nutzen: Automatisierte Analysen

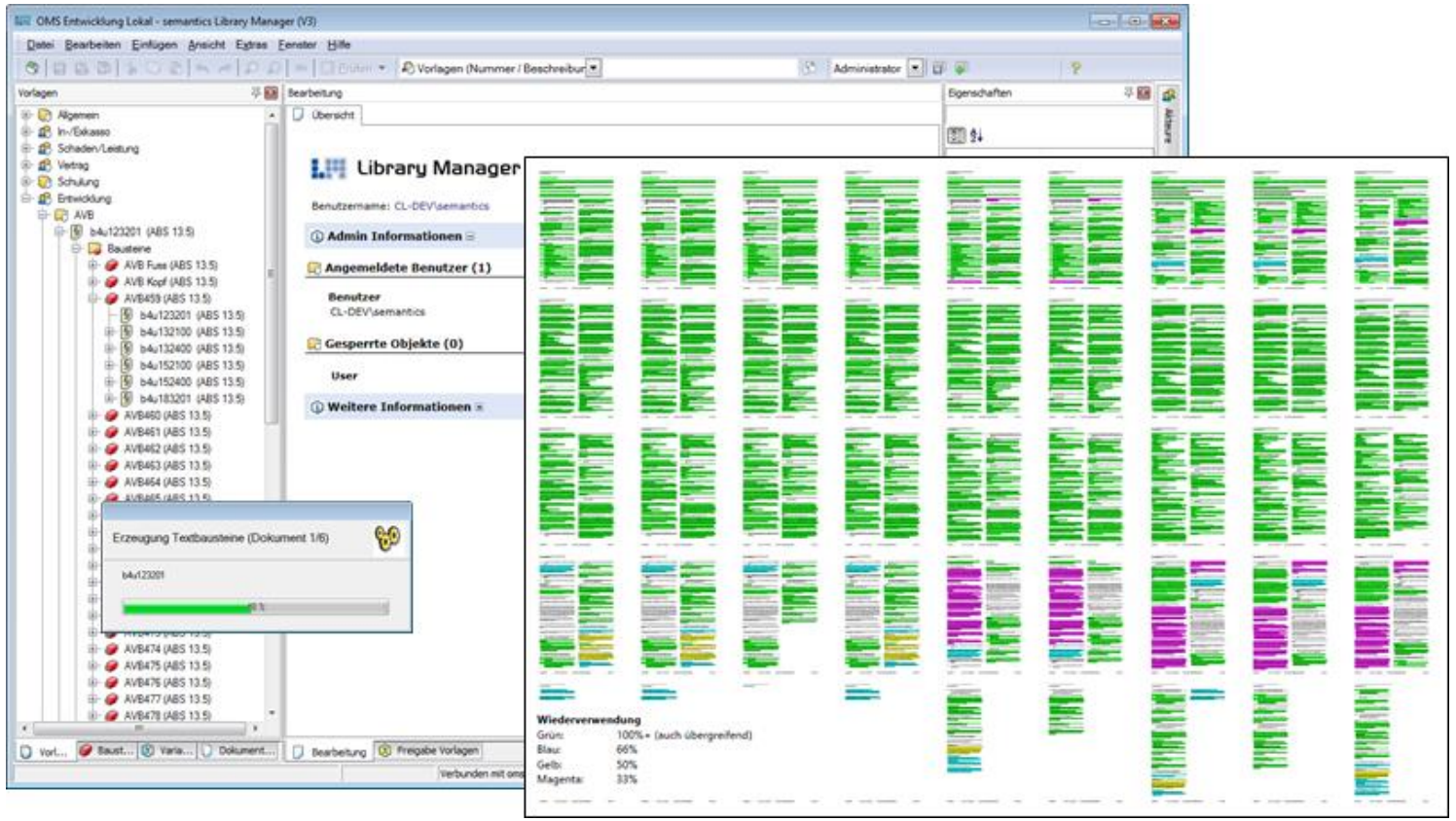


(Fallbeispiel bewusst unleserlich, Urheber- / Kundenschutz)

- Ermittlung identischer und *ähnlicher* Passagen
  ⇒ Kandidaten Textbausteine
- Ermittlung Ranking Ähnlichkeiten Dokumente
  ⇒ Kandidaten Variantenbriefe
- Anwendung auf große Dokumentbestände
- „Selbstheilung“ großer Textkorpora durch Identifikation Fehler, Inkonsistenzen

Erhebliche Einsparungen vs. manueller Analyse

Nutzen: Automatisierte Konsolidierung



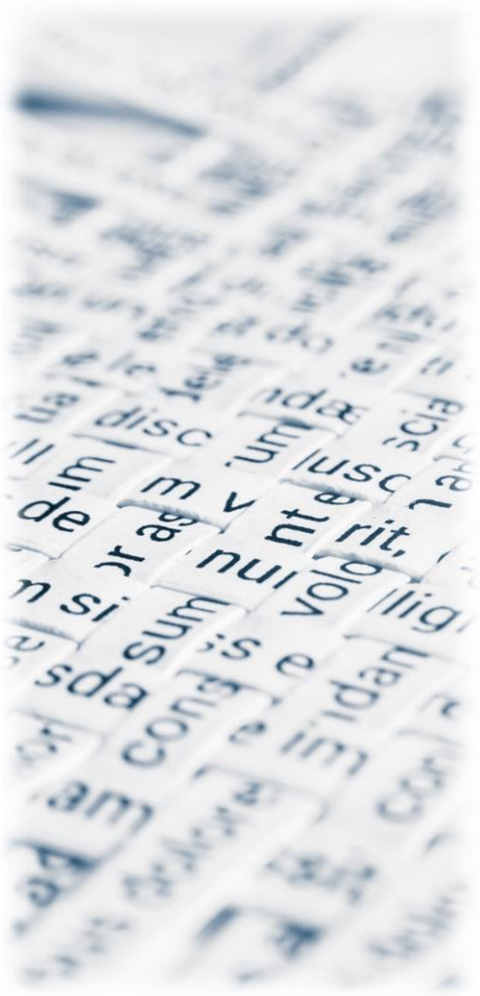
- Fortführung des Gedanken
- Einfache Steuerung durch Fachabteilung inkl. Variablen
- Keine technischen Kenntnisse
- Autom. Durchführung
- Fallstudie mit 2.500 AVB führt zu ca. 99% Abdeckung
- Generierung ca. 25.000 Textbausteine mit ∅ 60facher Wiederverwendung
- Ersparnis ca. **150 Personenjahre**

Realisierbarkeit bislang nicht möglicher Projekte

Fazit

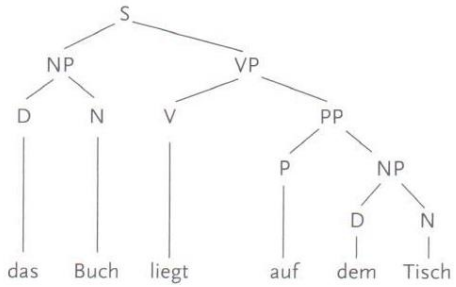
- Textverständlichkeit ist ungleich Lesbarkeit und ***lässt sich nicht messen***
- Textbestand ist ***lebendig*** und wird von ***zahlreichen Personen*** mit jeweiligen Kompetenzen gestaltet
- Einmalige Überarbeitung daher ***nicht ausreichend***
- Textbausteine ***behindern*** die Verständlichkeit
- Systemische Lösung in Form einer ***Entwicklungsumgebung für Text*** notwendig
- Dann erhebliche ***Gewinne*** durch
  - ***Automatisierung*** der individuellen Korrespondenz
  - Optimierung ***Pflege*** des Textbestandes

Aktuelle Forschung und Ausblick



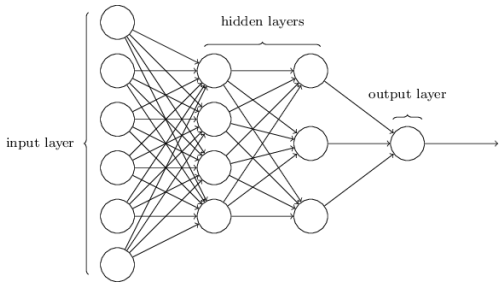
Text Mining: Nutzung bislang für maschinelle Übersetzungen, Dialogsysteme

- Syntagmatische Relationen
  (Kookurrenzen; Ergänzung; „Sonne“, „scheinen“)
- Paradigmatische Relationen
  (ähnlichen Kontexte; ähnliche Funktion; „Sonne“, „Kerze“)
- Semantische Relationen
  (Ober-Unterklassen, Synonym, Teil-Von, Instrument-Für, Ursache-Von, …)
- POS-Tagging (part-of-speech, grammatische Auszeichnung)



Neue Dimensionen zur Vermessung von Texteigenschaften

- Grammatische Struktur (durch POS-Tagging)
- Statistische Linguistik (type-token-ratio)
- Neuronale Netze, künstliche Intelligenz zu Bewertung und Vorschlagswesen



Empirische Studien

- Verständlichkeit verschiedener Textsorten
- Relevanz einzelner Ebenen
- 1. Studie veröffentlicht, 2. folgt in Kürze

Fragen und Antworten

Besten Dank für Ihre Aufmerksamkeit!

Dipl.-Ing. José Manuel de la Rosa Govantes
Björn Eßer M.A.

semantics Kommunikationsmanagement GmbH
Viktoriaallee 45
52066 Aachen

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