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®)

T

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

T₂

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.

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

Sprache

Linguistische Analyse

„Der Versicherer gewährt dem Versicherungsnehmer Versicherungsschutz für den Fall, dass er 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.“

Der Versicherer gewährt dem Versicherungsnehmer Versicherungsschutz für den Fall, dass er 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 dem Versicherungsnehmer Versicherungsschutz für den während der Wirksamkeit der Versicherung eingetretenen Schadenereignis, Verletzung oder Gesundheit (Personenschaden) oder die Vernichtung von Sachen (Sachschaden) für diese Folge aufgrund gesetzlicher 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 c eingetretenen Schadenereig Verletzung oder Gesundheit (Personenschaden) oder die Vernichtung von Sachen (S 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-,
strukturelle 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

Library Manager

Dokumentstruktur je Dokumenttyp und Anlass

The screenshot shows a document editor window titled 'Standard Leistungsbrief'. The main document content is structured as follows:

- [Metainformationen]
- [Dokumentkopf]
- [Dokumentkörper] [Dokumentinhalt]
 - [Einleitung] [Einleitung]
 - [Absatz] Formulieren Sie eine Einleitung, in der Sie auf vorangegangene Korrespondenz bzw. den Anlaß für diesen Brief Bezug nehmen. ¶
 - [Absatz] Fassen Sie den Anlaß des Briefes zusammen. Verzichten Sie nur auf vorangegangene Korrespondenz. Verzichten Sie nur auf den Dank. ¶
 - [Einleitung]
 - [Sachverhalt] [Sachverhalt]
 - [Absatz] Formulieren Sie einen sachlichen Hauptteil
 - [Absatz] 1. Handlung: Was tut die ... ¶
 - [Absatz] 2. Begründung: Warum handelt die ... dabei möglichst konkrete Begründungen, die dem K
 - [Absatz] 3. Folgen: Was wird die ... tun? Was s Begründen Sie die Folgen gegebenenfalls. ¶
 - [Sachverhalt]
 - [Abbinder] [Abbinder]
 - [Absatz] Schließen Sie den Brief ab. Bieten Sie der ... an. Finden Sie einen freundlichen Ausste Kunden. ¶
 - [Abbinder]
- [Dokumentfuß]

The 'Dokumententypen' sidebar on the right shows a hierarchical tree structure:

- Dokumente
 - Brief
 - Dokumentvorlagen
 - Standard Bankbrief
 - Standard Inkassobrief
 - Standard Leistungsbrief
 - Standard Vertragsbrief
 - Bausteinvorlagen
 - Standardbaustein
 - Formular
 - Dokumentvorlagen
 - Kreditrisiko
 - Standard Leistungsformular
 - Standard Vertragsformular
 - Leistungsabrechnung
 - Notiz
 - Police
 - Scheck
 - Zweitschritt
- Vorgänge
 - Vorgang
 - Brief
 - Formular

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 Dokumententypen

➤ Erhebliche Steigerung der inhaltlichen Qualität durch wohlkonzipierte Gliederungen und Standards

Schriftliche Kommunikation

Dimensionen

Text

Interne Eigenschaften:
Wort-, Satz- u. Textebene

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

Sprachprüfung,
Empirische Studien

Empfänger

Zielgruppe, Motivation, ...

Kognitionswissenschaft,
Neurolinguistik, ...
Sprechakttheorie

Unterstützung
Mentales Modell

Technik

Paradigmen beeinflussen
Sprachqualität

Paradigmenwechsel

Varianten statt Textbausteine

Kopf, Betreff, Einleitung, ...
(*Texte oder Bausteinerferenzen*)

Hauptteil

Erstattung?

Ja

Nein

Warum nicht?

Schluss, Abbinde, 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

Übersicht Nicht versichertes Hilfsmittel - schon gekauft

[- Absatz] > vielen Dank für Ihren Kostenvorschlag für [- Variable] [Name des Hilfsmittels (Name_des_Hilfsmittels)] < [- Absatz]

[- Einleitung]

[- Sachverhalt] [Sachverhalt]

[- Absatz] > Nach den Versicherungsbedingungen, die wir mit Ihnen vereinbart haben, können Sie für das genannte Hilfsmittel keine Leistung beanspruchen. Erstattungsfähig sind nur bestimmte Hilfsmittel. In den Versicherungsbedingungen sind die erstattungsfähigen Hilfsmittel abschließend aufgeführt. < [- Absatz]

[- Auswahl] [Muss-Auswahl: Was machen wir? (HANDLUNG)]

[- Variante] [keine Erstattung (1)]

[- Absatz] > Deshalb erhalten Sie die Rechnung unerstattet zurück. < [- Absatz]

[- Variante]

[- Variante] [Freiwillige Erstattung - Rechnung einreichen (2)]

[- Absatz] > Wir möchten Ihnen aber trotzdem helfen. Deshalb übernehmen wir Ihre Aufwendungen in tariflicher Höhe. < [- Absatz]

[- Absatz] > Unsere Entscheidung ist einmalig und entgegenkommend. Sie stellt keinen Berufungsfall für die Zukunft dar. Bitte reichen Sie uns nach dem Kauf die Originalrechnung zur Erstattung ein. < [- Absatz]

➤ „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 Formater

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

Text

Interne Eigenschaften:
Wort-, Satz- u. Textebene

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

Sprachprüfung,
Empirische Studien

Empfänger

Zielgruppe, Motivation, ...

Kognitionswissenschaft,
Neurolinguistik, ...
Sprechakttheorie

Unterstützung
Mentales Modell

Technik

Paradigmen beeinflussen
Sprachqualität

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

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

The screenshot displays a software interface for document management. It features a sidebar with a user list, a main content area with a document tree, and several pop-up windows. One window shows 'Administrative Einst.' with fields for 'Beschreibung', 'Name', 'Version', 'Mand.-Version', 'Symb. Freigab', 'LM-ID', 'Ansprechpart', 'Redaktionsgru', 'Alte Briefnumr', and 'Gespartt für P'. Another window shows 'Redaktionsgruppen' with a list of groups and their associated actions. A third window shows 'Vorlagen (647)' with a list of templates and their counts.

- 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 Zugriffs-konzept (Rechte / Rollen)

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

- 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
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

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

$$\Rightarrow G \approx 2.046.120$$

Optimierung: intuitive Bewertung

Pilotprojekt PKV AVB

2.500 Dokumente, je ca. 200 Absätze

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

$$\Rightarrow G \approx 124.950.000.000$$

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.)

	B	A	N	A	N	E
0	1	2	3	4	5	6
A	1	1	1	2	3	4
N	2	2	2	1	2	3
A	3	3	2	2	1	2
N	4	4	3	2	2	1
A	5	4	3	2	2	2
S	6	6	5	4	3	3

Levenshtein-Distanz
zwischen Absätzen

Splitten von „Overlaps“

Höhere Wiederverwendung

Zusammenhalten fachlicher Einheiten

z.B. Einleitende Sätze und
Aufzählungen

Analyse

Darstellung Ergebnisse

Fachabteilung ohne
technische Kenntnisse

Bekannte Tools

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

Ranking
Dokumentähnlichkeit
Identifikation Variantenbriefe

Herausforderung

Tausende Ergebnisse
visualisieren

Konsolidierung

Fachabteilung ohne
technische Kenntnisse

Vorgang

Anlagen Bausteine, Ersetzen in
Dokumenten, Eintragen
Variablen und Werte

The screenshot shows two overlapping windows. The top window is an Excel spreadsheet with a table of data. The bottom window is the 'Library Manager' application, which displays a tree view of templates and a detailed view of a selected template.

Zeile	A	B	C
1		b4u221901	0
2	X	Tarifbedingungen	b4u229401 Tarifbedingungen
3			
4	X	b4u221901	0
5		Unter Ziffer 2 (Tarifbedingungen) sind Art und Höhe der Versicherungsleistungen sowie die tarifbezogenen Leistungsvoraussetzungen und -ausschlüsse geregelt. Hier finden Sie außerdem die Bestimmungen über die Aufnahme- und Versicherungsfähigkeit sowie weitere Besonderheiten für diesen Tarif.	
6		Die Tarifbedingungen gelten in Verbindung mit den Allgemeinen Regelungen (Teil A Ziffer 1) zum Baustein Krankheitskosten-Versicherung sowie, falls vereinbart, mit den Sonderbedingungen (Teil A Ziffer 3).	
7		b4u221901	22
8		Tarife Beihilfe Ambulant <VarBez> - <VarEinzelGruppe>	
9		Tarife Beihilfe Ambulant AnwÄter - Einzelversicherung	
10			Tarif Beihilfe Ambulant 50% (BHA51) - Einz
11			
12			

The Library Manager window shows a tree view with the following structure:

- AVB
 - b4u123201 (ABS 13.5)
 - Bausteine
 - AVB Fuss (ABS 13.5)
 - AVB Kopf (ABS 13.5)
 - AVB469 (ABS 13.5)
 - b4u123201 (ABS 13.5)
 - b4u132100 (ABS 13.5)
 - b4u132400 (ABS 13.5)
 - b4u152100 (ABS 13.5)
 - b4u152400 (ABS 13.5)
 - b4u183201 (ABS 13.5)
 - AVB460 (ABS 13.5)

The main view shows the 'Admin Informationen' section for the selected template:

- Benutzername: CL-DEV\semantics
- Datenbank: oms_dev
- Rechnername: CL-DEV
- Angemeldete Benutzer (1):

Benutzer	Datenbank	Rechnername
CL-DEV\semantics	oms_dev	CL-DEV
- Gesperrte Objekte (0)
- Weitere Informationen

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 Personenjahre

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 Assistant

Webservice
Nahtlose Integration,
Anbindung Drittlösungen

Sentiment,
Big Data

Keyword- / Topic
Extraction

CL-Textanalysis

Library Assistant

Sprachprüfung

Redaktion

LM-Sprachtool

Editor

Content Management

Library Assistant

Webservice
Nahtlose Integration,
Anbindung Drittlösungen

Sentiment,
Big Data

Keyword- / Topic
Extraction

CL-Textanalysis

Library Assistant

Vorschlagswesen

Redaktion

LM-Sprachtool

Editor

Content Management

Library Assistant

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

semantics
Kommunikationsmanagement GmbH

versicherungsscheinnummer 1234 und dem amtlichen Kennzeichen M B 1234 angefordert.
Nun teilte der Vorversicherer mit, dass er diese aufgrund einer Falschangabe nicht erteilen könne. Unter der Versicherungsschein-Nummer, die Sie in Ihrem Antrag genannt haben, soll ein anderes amtliches Kennzeichen versichert sein.

Bitte sehen Sie doch in den Unterlagen noch einmal nach und prüfen, ob die angegebene Vorversicherung, die Versicherungsscheinnummer und das amtliche Kennzeichen richtig sind. Wir bitten Sie mit uns zeitnah (bis zum 31.01.2018) **Kontakt aufzunehmen**.
Nach Ihrer Rückmeldung werden wir uns erneut um die Bestätigung Ihres Schadenfreiheitsrabattes bemühen.

Dazu ein wichtiger Hinweis: Sollten wir innerhalb der nächsten 14 Tage nichts von Ihnen hören, müssten wir Ihnen leider einen höheren Beitrag berechnen. Bitte haben Sie dafür Verständnis.

Mit freundlichen Grüßen

49 Score

Sentiment
negativ

Wir hatten für Sie eine solche Best [...] n Kennzeichen M B 1234 angefordert.
Satzlänge

...it, dass er diese aufgrund einer **Falschangabe** nicht erteilen könne. Unter der ...
Unerwünschtes Vokabular
"Falschangabe" ist laut Sentiment-Analysen negativ konnotiert. Besser: eine klare Formulierung, dass möglicherweise ein Fehler vorliegt.

Bitte sehen Sie doch in den Unterla [...] amtliche Kennzeichen richtig sind.
Satzlänge

...ersichert sein. Bitte sehen Sie **doch** in den Unterlagen noch einmal na...
Füllwörter

Wir bitten Sie mit uns zeitnah (bis zum 31.01.2018) **Kontakt aufzunehmen**.
Funktionsverbgefüge
Ist eine Formulierung mit einem Verb wie "kontaktieren" möglich?

...tig sind. Wir bitten Sie mit uns **zeitnah** (bis zum 31.01.2018) Kontakt auf...
Unerwünschtes Vokabular
"zeitnah" ist laut Sentiment-Analysen negativ konnotiert. Ist eine Formulierung mit "in Kürze" möglich?

Funktionsverbgefüge erschweren das Verständnis der Aussage. Formulieren Sie stattdessen den Satz mit einem Verb.

Mehr

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

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

```
1 .sp
2 .fo
3 .if xxxxxx
4 .th .go AAA
5 .*description**
6 .if xxxxxx
7 .th .go BBB
8 .if xxxxxx
9 .th .go CCC
10 .th .go DDD
...
89 ...BBB
94 ...AAA
95 .if xxxxxx
96 .th xxxxxx
```

z.B. dcf Baustein

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

Übersetzung in deklaratives XML

Ablauf-Graph

```
1 <textblock>
2 <meta>
3 </meta>
4 <br/>
5 <!-- .fo left -->
6 <conditions>
7 <condition>
8 <variable name="CSADPHOMBI" automatic="true" unname="CSADPHOMBI"
9 <equalTo/>
10 <conditionValue>V</conditionValue>
11 </condition>
12 <then>
13 </then>
14 <else>
15 <conditions>
16 <!-- ***** Prüfen Kapital / Rente / RU / BIR *****
17 <condition>
18 <conditionValue>X</conditionValue>
19 <variable name="LEHVTARPT" automatic="true" unname="LEB
20 <equalTo/>
21 <conditionValue>XRI</conditionValue>
22 </condition>
23 <or/>
24 <condition>
25 <conditionValue>X</conditionValue>
26 <variable name="LEHVTARPT" automatic="true" unname="LEB
27 <equalTo/>
28 <conditionValue>XRT</conditionValue>
29 </condition>
30 <then>
31 <!-- .go RIESTER -->
32 <!-- ... RIESTER -->
33 <conditions>
34 <condition>
35 <conditionValue>X</conditionValue>
36 <variable name="LEHVTARSTUP" automatic="true" ur
37 <greaterThan/>
38 <conditionValue>X2012</conditionValue>
39 </condition>
40 <then> Basler Riesteraktiv top3 </then>
41 <else> RingRiesteraktiv </else>
42 </conditions>
43 <!-- .go VERTNR -->
44 </then>
45 <else>
46 <conditions>
47 <condition>
48 <conditionValue>X</conditionValue>
49 <variable name="LEHVTARKLA" automatic="true" ur
50 <equalTo/>
51 <conditionValue>X04</conditionValue>
52 </condition>
53 <or/>
54 <condition>
55 <conditionValue>X</conditionValue>
56 <variable name="LEHVTARKLA" automatic="true" ur
57 <equalTo/>
58 <conditionValue>X05</conditionValue>
59 </condition>
60 <then>
61 <!-- .go RENTE -->
62 <!-- ***** Prüfen Rente nach Vertriebs
63 <!-- ... RENTE -->
64 <conditions>
65 <condition>
66 <conditionValue>X</conditionValue>
67 <variable name="LEVERTREWC" automatic="
68 <equalTo/>
69 <conditionValue>X22</conditionValue>
70 </condition>
71 <then>
72 <!-- .go ZEUSER -->
```

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

IC_ID	IC_ExternalId	IC_Type	blank positioning	comment	pure content	rowcount	set counter	set simple	structure
30487	ADRXXCSTXX	dcf			1	4	31		hierarchical
30488	ADRXXTADX1	dcf			1	23	54		linear
30489	ADRXXTADX2	dcf			1		2		linear
30492	ADRXXTADXX	dcf			35	2	48		linear
30495	ADRXXTBEXX	dcf			12	8	36		hierarchical
30496	ADRXXTMAXX	dcf			11	11	46		hierarchical
30497	ADRXXTMFX1	dcf			2	25	46		hierarchical
30498	ADRXXTMFXD	dcf			9		12		linear
30509	ADRXXTMVX8	dcf				111	169		linear
30510	ADRXXTMVXX	dcf			16	9	43		linear
30511	ADRXXTOLXX	dcf			8	6	26		hierarchical
30513	ADRXXTPWE1	dcf			1	7	15		linear
30515	ADRXXTPWN1	dcf			6	8	23		linear
30517	ADRXXTPVV1	dcf			1		9		linear
30520	ADRXXTPWXX	dcf			13	3	40		hierarchical
30526	ADRXXTSTXX	dcf			1	4	14		linear
30527	ADRXXTSTXX	dcf			10	2	25		linear
30528	ADRXXTSTXY	dcf			2		3		linear
30532	ADRXXTUWAL	dcf			2	3	6		linear
30533	ADRXXTUWAW	dcf			6	19	36		hierarchical
30534	ADRXXTUWET	dcf			3	1	10		hierarchical
30535	ADRXXTUWEV	dcf			7	19	38		hierarchical
30537	ADRXXTUWNL	dcf			2	7	10		linear
30538	ADRXXTUWVE	dcf			2	4	7		linear
30540	ADRXXTXDRX	dcf			1	2	8		linear
30574	AGLFXTTA2E	dcf			3		42		linear
30660	AGLFXTTA8V	dcf			3		26		linear
30690	AGLFXTTA9R	dcf					7		linear
30749	AGLFXTTA9D	dcf					23		linear
30758	AGLFXTTAE7	dcf					41		linear
30763	AGLFXTTAF3	dcf			4		5		linear
30775	AGLFXTTA10	dcf			4		73		linear
30782	AGLFXTTA17	dcf			4		71		linear
30783	AGLFXTTA18	dcf			3		70		linear
30784	AGLFXTTA19	dcf			3		70		linear
30785	AGLFXTTA1A	dcf			4		61		linear

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

Extraktion Dokumentenkörper
anhand xpath-Regeln

vs.

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

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

Besten Dank für Ihre Aufmerksamkeit!