
ReqCheck: AUTOMATISIERTE QUALITÄTSSICHERUNG VON ANFORDERUNGEN MIT HILFE LINGUISTISCHER REGELN

Ursula Reuther, IAI (Institut für Angewandte Informationsforschung)

Ursula.Reuther@iai-sb.de

Matthias Koch, Fraunhofer IESE

Matthias.Koch@iese.fraunhofer.de



Fraunhofer
IESE

IAI - Institut für Angewandte Informationsforschung

- gegründet 1985
- als so genanntes „An-Institut“
Institut der Gesellschaft für Angewandte Informationsforschung an der Universität des Saarlandes e.V.
- Zweck:
Technologie- und Wissenstransfer in den Bereichen Natural Language Processing und Language Engineering

IAI - Institut für Angewandte Informationsforschung

Angewandte Forschung

Projektträger:
EU
BMBF

Maschinelle Übersetzung
Sprachlernsoftware
(eLearning)
Big Data

Transferprojekte

Projektträger:
BMW
Saarland

Intelligente
Inhaltserschließung
Sprachprüfsoftware
Automatische
Lexikonaquisition

Industrieprojekte

Automobilindustrie
Verlage
Softwarehersteller
Maschinenbau
Verwaltungen

Kundenspezifische
Softwareanpassungen
Sprachtechnologische
Anwendungen
Dienstleistungen und
Beratung

Die Fraunhofer-Gesellschaft auf einen Blick

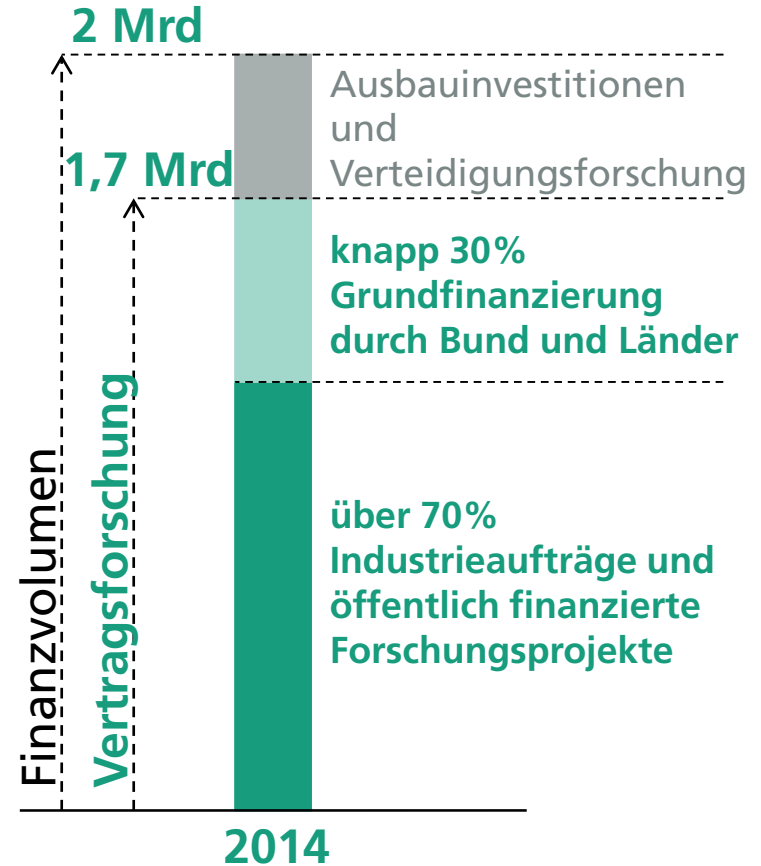
Anwendungsorientierte Forschung zum unmittelbaren Nutzen für die Wirtschaft und zum Vorteil für die Gesellschaft



Knapp **24 000** Mitarbeiterinnen und Mitarbeiter



66 Institute und Forschungseinrichtungen



Fraunhofer IESE

Das Institut für Software- und Systementwicklungsmethoden

- Gegründet 1996 mit Sitz in Kaiserslautern
- ca. 240 Mitarbeiter
- Unsere Lösungen sind flexibel skalierbar und für Firmen jeder Größe geeignet



■ Unsere wichtigsten Branchen:

- Automobil- und Transportsysteme
- Automatisierung und Anlagenbau
- Gesundheitswesen
- Informationssysteme
- Energiemanagement
- E-Government

AGENDA

- Anforderungsdokumentation
- Qualitätsproblematik bei Anforderungen
- Technische Dokumentation: Gemeinsamkeiten und Unterschiede
- Lösungen in der Technischen Dokumentation
- Anpassung auf den Anforderungskontext
- Implementierung der angepassten Lösung
- Unterschiede zu traditionellen Lösungen
- Erfahrungen und zukünftige Arbeiten

Natürlichsprachliche Anforderungen

Eigenschaften

- Unstrukturiert (z.B. Fließtext)
- Strukturiert (z.B. Use Cases Vorlagen)
- Am häufigsten verwendete Form für Dokumentation

Vorteile

- Hohe Flexibilität
- Verwendbar für alle Arten von Anforderungen
- Können von allen Stakeholdern verstanden werden, ohne neue Notationen / Sprachen zu lernen

Qualitätsproblematik bei Anforderungen

Probleme in der Praxis [Kan03]

Untersucht wurden 967 natürlichsprachliche Anforderungen

223
Anforderungen
haben weak
words. (~23%)

83
Anforderungen
waren nicht
Singular.

53
Anforderungen
waren nicht
kurz und
einfach.

37
Anforderungen
waren
unvollständig.

23
Anforderungen
verwenden
Negierungen.

[Kan03] Kandt, R. K. 2003: Software Requirements Engineering: Practices and Techniques. JPL Document D-24994. SRI Report R-3. Jet Propulsion Laboratory. California Institute of Technology

Qualitätsproblematik bei Anforderungen

Viele dieser Schwachstellen lassen sich automatisiert prüfen.

- Einfache Prüfung der Wörter greift aber oft zu kurz

Komplexere
Problematiken:

- Layout und Darstellung,
- Satzbau und Komplexität,
- Schreibstil,
- Wortstellungen,
- Wortwahl,
- Substantive ohne Bezugsindex,
- Unvollständig spezifizierte Bedingungen

Technische Dokumentation: Gemeinsamkeiten und Unterschiede

■ Anforderungen an Technische Dokumentation

verständlich

eindeutig

konsistent

sprachlich korrekt

zielgruppengerecht

gut übersetzbar

gemäß CI-Vorgaben

■ Anforderungen an Anforderungsdokumentation (IEEE 29148)

verständlich

eindeutig

konsistent

vollständig

verfolgbar

realisierbar

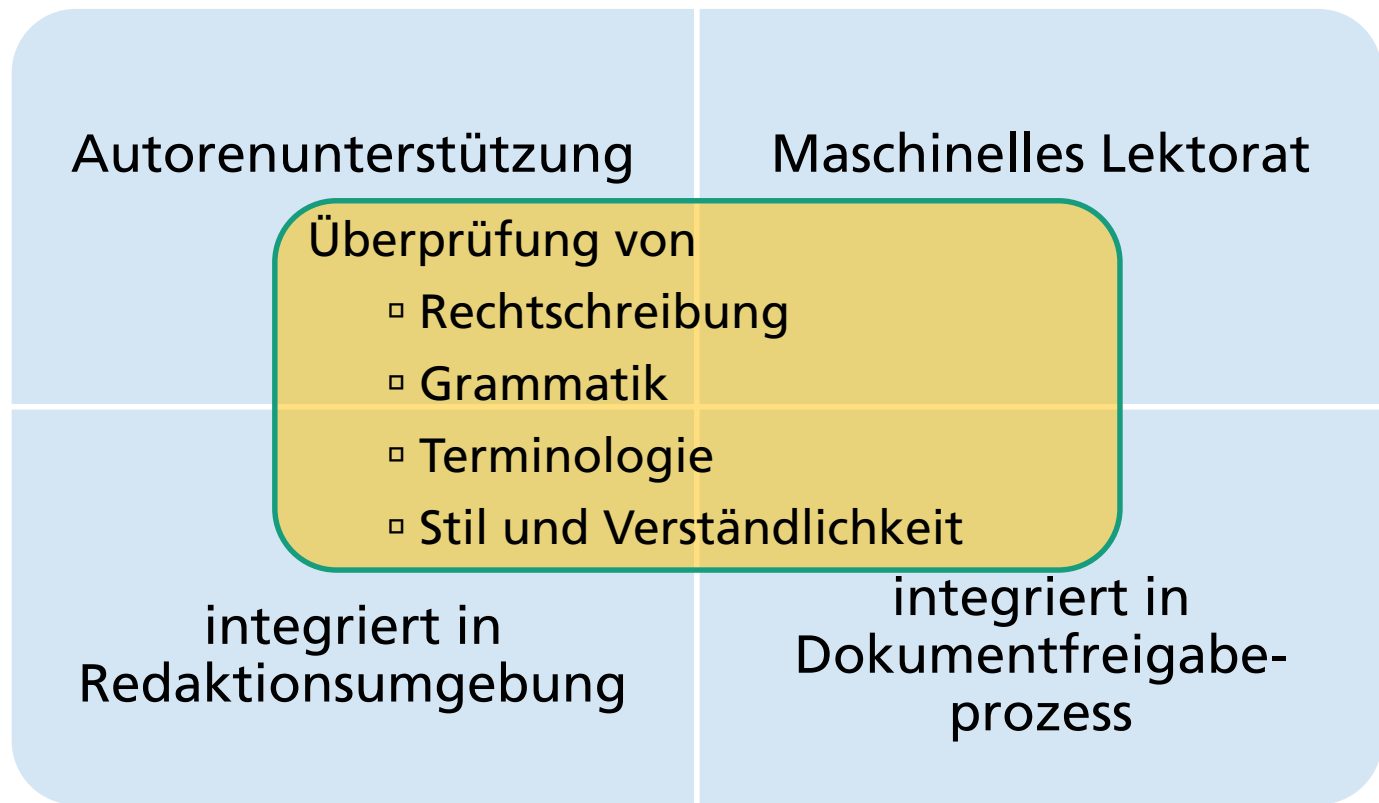
prüfbar

notwendig

abgestimmt

Lösungen in der Technischen Dokumentation

■ Software zur Sprachqualitätssicherung



Anpassung auf den Anforderungskontext

Ermittelte Regelsätze für Anforderungen

**76 Regeln
wurden
untersucht**

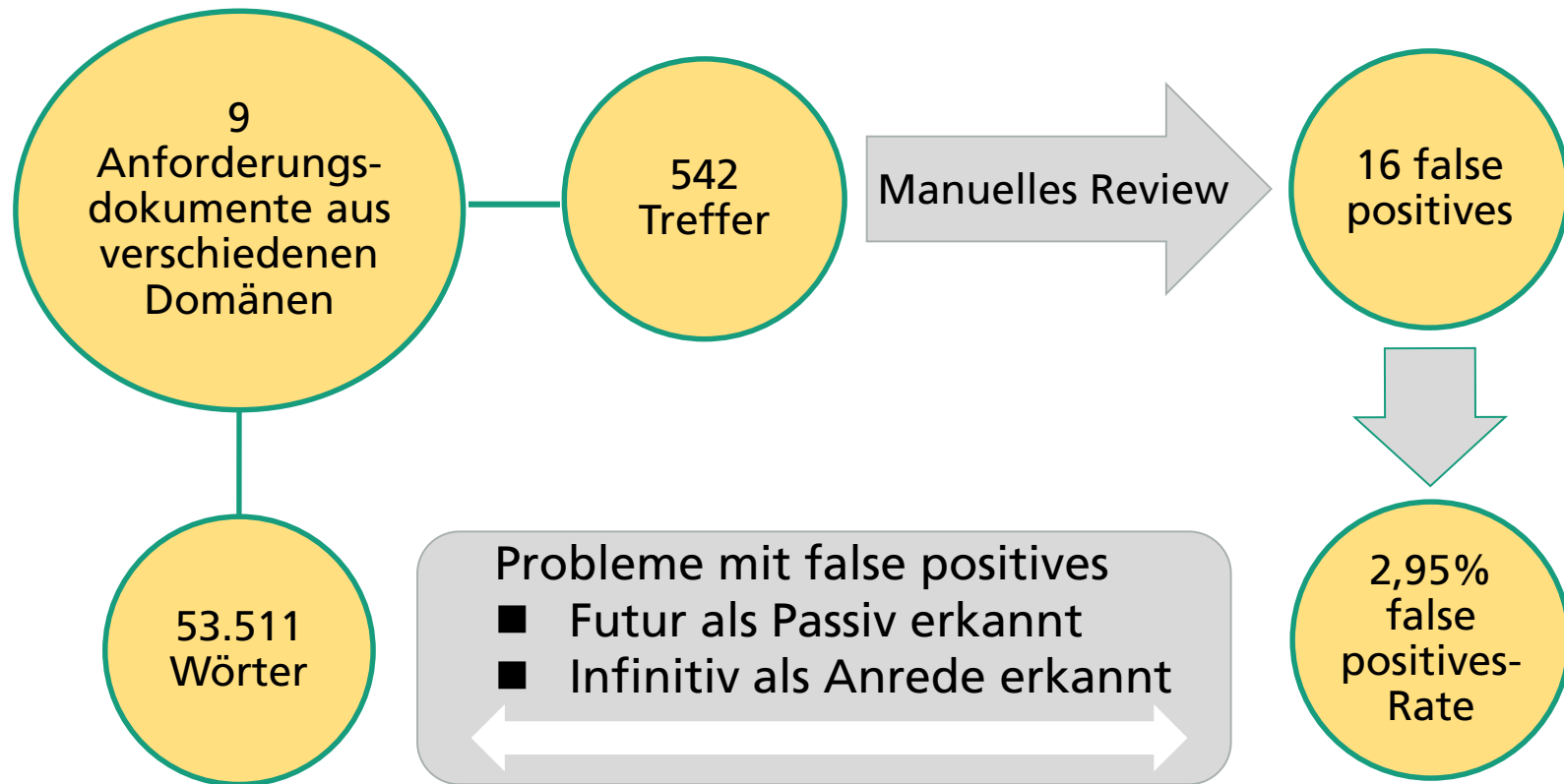
Identifizierte
Probleme

Manche Regeln sind im RE Kontext
kontraproduktiv oder unpassend

- Schließen Beschreibungsmöglichkeiten aus, bspw. „keine Klammern verwenden“
- Stellen Vereinfachungen über Sprachregeln, bspw. „zu langes Kompositum aufgliedern (>3)“
- Schließen Anredeformen aus, die in bestimmten Fällen erlaubt sind (z.B. bei Use Cases), bspw. „1. Person Singular und Plural vermeiden“
- Prozesswörter werden durch die Stilregeln vermieden, das Problem sind aber unvollständig spezifizierte Prozesswörter

Anpassung auf den Anforderungskontext

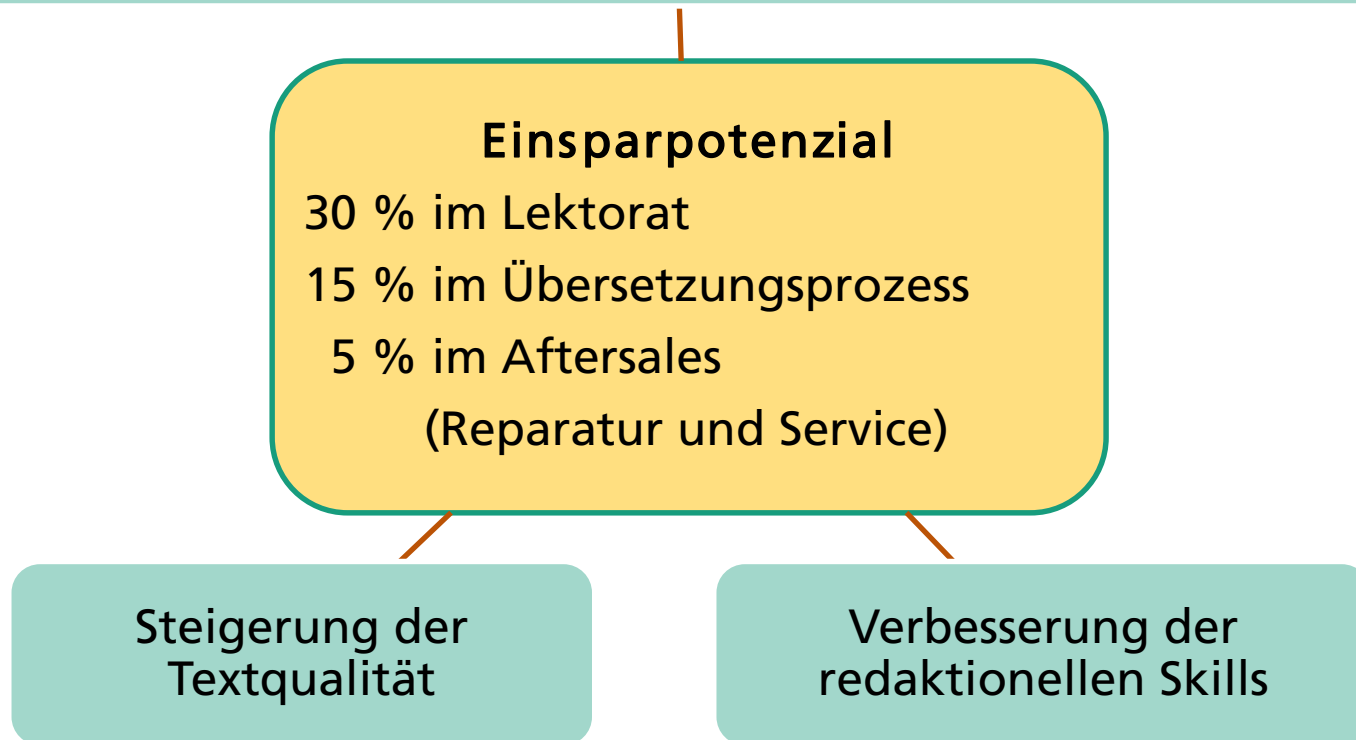
Fallstudie mit 29 identifizierten Regeln



Anpassung auf den Anforderungskontext

Erfahrungen aus der Technischen Dokumentation

- keine präskriptive „Kontrollierte Sprache“ im klassischen Sinne
- Regelpool als Basis für eine flexible, individuelle Regelkonfiguration



Implementierung der angepassten Lösung

<https://reqcheck.iese.de>

Entwicklung des Web-Clients
mit ASP.net-Technologie

Möglichkeit zum Hochladen
eines Docx-Dokuments

Transformation des Docx-
Dokuments in HTML

•über Apache Tika

Interne Verarbeitung im
Prüfsystem

Formatierung der Befunde in
HTML

•Pop-Over zu Markierungen zeigen betroffene Stellen im geprüften Text
•Möglichkeit zum Export der Ergebnisse (pdf und html)

Sorgsamer Umgang mit
vertraulichen Daten

Implementierung der angepassten Lösung

<https://reqcheck.iese.de>

ReqCheck

Startseite

Über uns

Impressum

Registrieren

Anmelden

ReqCheck

ReqCheck bietet Ihnen eine schnelle und einfache Unterstützung bei der Qualitätssicherung Ihrer Anforderungsspezifikationen.

Dokument hochladen »

Ihre Dokumente

ReqCheck analysierte Ihre Anforderungsdokumente hinsichtlich sprachlicher Probleme und informiert Sie über die Befunde, inklusive Verbesserungsvorschläge, wie Sie die Probleme beheben können.

Das System

ReqCheck ist aus einer Kooperation zwischen dem Institut für Angewandte Informationsforschung (IAI) und dem Fraunhofer-Institut für Experimentelles Software Engineering (IESE) entstanden. Die Kompetenzen beider Institute wurden gebündelt um dieses System zu entwickeln.

Die Testphase

Aktuell befindet sich ReqCheck in einer Testphase, an der Sie mit der Nutzung des Systems teilnehmen. Derzeit wird ReqCheck ausgiebig praktisch genutzt, um den Nutzen und die Anwendbarkeit zu evaluieren.

© 2015 - Fraunhofer IESE

19

Implementierung der angepassten Lösung

<https://reqcheck.iese.de>

ReqCheck [Startseite](#) [Über uns](#) [Impressum](#)

[Registrieren](#) [Anmelden](#)

Anmelden.

Lokales Konto für die Anmeldung verwenden.

E-Mail

Alexander.Klaus@iese.fraunhofer.de

Kennwort

••••••••

Speichern?

Anmelden

[Als neuer Benutzer registrieren](#)

[Kennwort vergessen?](#)

© 2015 - Fraunhofer IESE

Implementierung der angepassten Lösung

<https://reqcheck.iese.de>

ReqCheck Startseite Über uns Impressum Hallo Alexander.Klaus@iese.fraunhofer.de! Abmelden

Dateiname 635826776326630859-Publikation_reqcheck Final.docx
Anzahl der Wörter 1412
Ergebnis **ReqCheck: Automatisierte Qualitätssicherung von Anforderungsdokumenten mit Hilfe linguistischer Regeln**

Alexander Klaus, Matthias Koch
Fraunhofer IESE, Kaiserslautern
Alexander.Klaus@iese.fraunhofer.de, Matthias.Koch@iese.fraunhofer.de
Ursula Reuther, Gernot Hensberg
IAI, Saarbrücken
ursula.reuther@iai.uni-sb.de, gernot.hensberg@iai.uni-sb.de

Beitrags-Kategorie
Forschungsergebnisse

Zielgruppe des Beitrags
RE-Anwender, RE-Interessierte

Motivation der behandelten Fragestellung
Schriftliche Dokumentation wird traditionell in der Software-Entwicklung verwendet. Insbesondere natürlichsprachliche Dokumentation findet in der

"wird"
Schreibstil überprüfen.
Formulieren Sie den Satz in einen Satz im Aktiv oder in eine direkte Anweisung um.
Auch wenn der Handelnde nicht bekannt ist, ist oft eine Formulierung im Aktiv besser verständlich, da deutlich wird, wer etwas tut oder tun muss. mentation ist
jedoch gerade bei großen Dokumenten sehr zeintensiv und damit nur mit großem Aufwand möglich.

Behandelte Fragestellung oder Problem
Vor diesem Hintergrund wurde untersucht, durch welche Maßnahmen die Überprüfungs- und Lektoratsaufwände reduziert und der Prozess der Qualitätssicherung vereinfacht werden können. Ein maschinelles Lektorat auf Basis einer linguistisch intelligenten Analyse­methode liefert in jedem Fall einen größeren Datendurchsatz, doch die Frage stellt sich, ob die Ergebnisse verlässlich, korrekt und verwertbar sind. Weiterhin ist offen, wie ein solches maschinelles Lektorat dem Anwender zugänglich gemacht und in den Prozess der Qualitätssicherung sinnvoll eingebunden werden kann.

Implementierung der angepassten Lösung

<https://reqcheck.iese.de>

Beitrags-Kategorie

Forschungsergebnisse

X1Zielgruppe des Beitrags

RE-Anwender, RE-Interessierte

Motivation der behandelten Fragestellung

Schriftliche Dokumentation X2wird traditionell in der Software-Entwicklung X3verwendet. Insbesondere natürlichsprachliche Dokumentation findet in der heutigen Praxis nach wie vor weitverbreitete Anwendung [1]. X4Jedoch besteht gerade bei dieser Form der Anforderungsdokumentation die Problematik, dass qualitative Mängel zusätzliche Nacharbeiten verursachen, etwa wenn Anforderungen nicht eindeutig spezifiziert sind, und daher Interpretationsspielraum herrscht. Solche Mängel X5sorgen durch die Notwendigkeit von Nachfragen X6mittels fortwährender mündlicher Kommunikation für einen erhöhten Aufwand und für höhere Kosten gegenüber qualitativ hochwertigen Anforderungen. Eine manuelle X7Prüfung der Dokumentation ist jedoch gerade bei großen Dokumenten sehr zeitintensiv und damit nur mit großem Aufwand möglich.

Behandelte Fragestellung oder Problem

Vor diesem Hintergrund X8wurde untersucht, durch welche Maßnahmen die Überprüfungs- und Lektoratsaufwände reduziert und der Prozess der Qualitätssicherung X9X10vereinfacht werden können. X11X12Ein maschinelles Lektorat auf Basis einer linguistisch intelligenten Analyseverfahren liefert in jedem Fall einen größeren Datendurchsatz, doch die Frage stellt sich, ob die Ergebnisse verlässlich, korrekt und verwertbar sind. Weiterhin ist offen, wie ein solches maschinelles Lektorat dem Anwender zugänglich gemacht und in den Prozess der Qualitätssicherung sinnvoll X13X14eingebunden werden kann.

X15 Lösung und Ergebnisse

Implementierung der angepassten Lösung

<https://reqcheck.iese.de>

Befund #4

Jedoch besteht gerade bei dieser Form der Anforderungsdokumentation die Problematik, dass qualitative Mängel zusätzliche Nacharbeiten verursachen, etwa wenn Anforderungen nicht eindeutig spezifiziert sind, und daher Interpretationsspielraum herrscht.

Satzbau und Komplexität überprüfen.

Formulieren Sie den Satz um und verwenden Sie weniger als 26 Wörter.

Zu lange Sätze erschweren die Informationsentnahme.

Befund #5

sorgen

Schreibstil überprüfen.

Verzichten Sie auf den Infinitiv als Befehlsform. Verwenden Sie stattdessen die direkte Anrede mit "Sie".

Durch die Nennung des Subjekts wird der Satz in manchen Fällen weniger mehrdeutig, was vor allem für die (maschinengestützte) Übersetzung wichtig ist.

Befund #6

mittels

Wortwahl überprüfen.

Ersetzen Sie "mittels" durch "mit", "mithilfe von" oder "durch".

Veraltete oder hochtrabende Formulierungen sind in der Technischen Dokumentation unangebracht.

Befund #7

Prüfung der Dokumentation

Wortbezüge überprüfen.

Verzichten Sie auf mehrdeutige Genitivkonstruktionen. Formulieren Sie den Satz z.B. mit einem Relativsatz um.

Derartige Genitivkonstruktionen machen Sätze doppeldeutig und missverständlich.

Befund #8

Unterschiede zu traditionellen Lösungen

Linguistische Verarbeitungsmethoden mit morpho-syntaktischer Analyse

- Traditionelle Ansätze arbeiten oft mit Wortlisten und prüfen auf
 - weak words
 - Multiplizität
 - Negation
 - Unvollständigkeit (z.B.: „TBD“)
 - Passive Formulierungen
 - Subjektive Formulierungen
 - Optionalität
 - Implizität
 - Uneindeutigkeit
 - Vergleichende Phrasen

Unterschiede zu traditionellen Lösungen

Linguistische Verarbeitungsmethoden mit morpho-syntaktischer Analyse

Zerlegung von Wörtern in kleinste bedeutungs-tragende Einheiten (Morpheme)

Entwässerungsgräben



Flexion

Entwässerungsgraben



Komposition

Entwässerung # s # graben



Derivation

ent # wässern # ung # graben

Unterschiede zu traditionellen Lösungen

Linguistische Verarbeitungsmethoden mit morpho-syntaktischer Analyse

■ Parse-Beispiel

NP [Die<art> Mitte<noun>]

NP [einer<art> Verdichterstütze<noun>]

ist<fiv>

PP [mittels<p>

NP [Ansenkung<noun>]

]

PP [in<p>

NP [der<art> oberen<adj> Tragplatte<noun>]

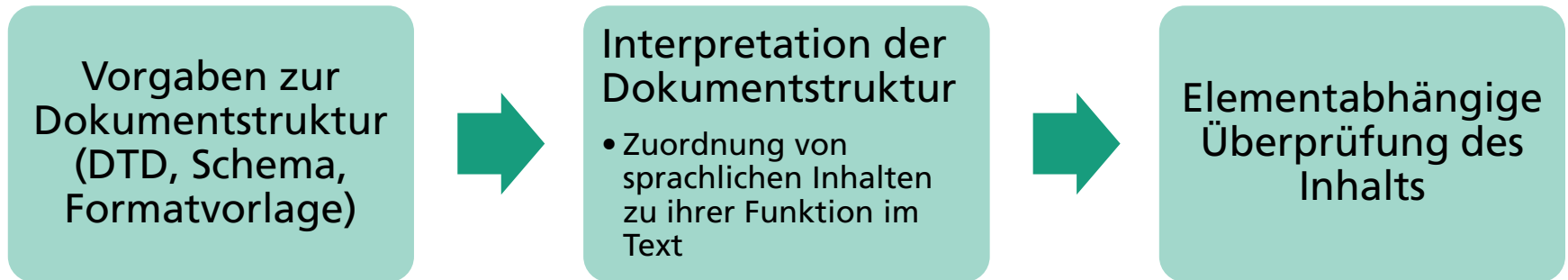
]

gekennzeichnet<ptc2>

.<punct>

Unterschiede zu traditionellen Lösungen

Berücksichtigung der Dokumentstruktur



<TITLE>Das ist eine Überschrift</TITLE>
<PROCEDURE>In Handlungsanweisungen **darf** kein Passiv **verwendet werden.**</PROCEDURE>

Keine Fehlermeldung, dass der Satzendeppunkt fehlt

Regel, die Passiv in Kombination mit Modalverben verbietet, ist auf das Element <PROCEDURE> beschränkt.

Unterschiede zu traditionellen Lösungen

Direkte Nutzung durch Webservices

Gesicherte
Verbindung zum
Webservice

Keine Installation
notwendig

Einfacher, schneller
und flexibler Zugriff

Einfache
Accountverwaltung

- Kein Zugriff auf andere Dokumente
- Kein Zugriff auf frühere Prüfungsergebnisse

Erfahrungen und zukünftige Arbeiten

Erfahrungen

- Schnelle und einfache Prüfung „nebenbei“
 - Unmittelbare Qualitätsverbesserung möglich
- Weniger Kommunikationsaufwand durch bessere Qualität der Dokumentation
- Nicht alle Regeln sind für alle Teile eines Dokuments sinnvoll (Überschriften, Tabelleninhalte ...)

Zukünftige Arbeiten

- Berücksichtigung spezifischer Formatierungen
- Erweiterung des Regelsets

ReqCheck: automatisierte Qualitätssicherung von Anforderungen mit Hilfe linguistischer Regeln

Ursula Reuther, IAI (Institut für Angewandte Informationsforschung)

Ursula.Reuther@iai-sb.de

Matthias Koch, Fraunhofer IESE

Matthias.Koch@iese.fraunhofer.de

<https://reqcheck.iese.de>

The screenshot shows the homepage of the ReqCheck website. The header contains navigation links: 'ReqCheck', 'Startseite', 'Über uns', 'Impressum', 'Registrieren', and 'Anmelden'. The main content area features the title 'ReqCheck' and a sub-header: 'ReqCheck bietet Ihnen eine schnelle und einfache Unterstützung bei der Qualitätssicherung Ihrer Anforderungsspezifikationen.' Below this is a prominent green button labeled 'Dokument hochladen »'. The page is divided into three columns: 'Ihre Dokumente', 'Das System', and 'Die Testphase', each with a brief description of the tool's capabilities and current status. The footer includes the copyright notice '© 2015 - Fraunhofer IESE'.

30