

# Übersetzungen richtig bewerten

Mit verschiedenen Werkzeugen lassen sich Fehler in Übersetzungen erkennen und gewichten. Doch welche Risiken für ein Unternehmen noch bestehen, ist auf den ersten Blick nicht ersichtlich.

Unterstützung verspricht die risikobasierte Evaluierung.

TEXT *Angelika Ottmann* UND *Carmen Canfora*

Fehler in Übersetzungen sind ärgerlich und leider nicht immer vermeidbar. Manchmal können sie jedoch gravierende Konsequenzen haben: Beispielsweise führen Übersetzungsfehler oder Unterschiede zwischen Sprachversionen bei Kontrollen durch die Behörden in 100 Prozent der Fälle zu Rücknahme und Verkaufsverbot, in zwei Dritteln zu Rückrufen. Ermittelt hat dies eine DIN-Studie, die von der tekcom erstellt wurde [1].

## Wie können die Folgen aussehen?

Ein Schweizer Pharmahersteller legte seinen Schmerztabletten für den italienischsprachigen Markt einen Beipackzettel mit folgendem Hinweis bei:

→ „Ai bambini sotto 9 anni può essere somministrato Dafalgan compresse.“

Richtig wäre gewesen:

→ „Ai bambini sotto 9 anni non può essere somministrato Dafalgan compresse.“

Die Folgen: Rückruf von drei Chargen des Medikaments, Eilmeldung der Gesund-

heitsbehörde, Zeitverlust, finanzieller Schaden, Imageschaden. Bei einem anderen Schweizer Pharmaunternehmen geriet folgende Information in den Beipackzettel zu einem „Faecal Management System“:

→ Statt: „Man kann den Silikonkatheter spülen, indem man die Spritze mit lauwarmen Wasser füllt, am „Ansatzstück Spülschlauch“ (...) anschließt und auf den Kolben der Spritze drückt.“

→ ist zu lesen: „Man kann den Silikonkatheter spülen, indem man die Spritze mit lauwarmen Wasser füllt, am „Ansatzstück Füllschlauch“ (...) anschließt und auf den Kolben der Spritze drückt.“

Die Folgen: Kundeninformation und Eilmeldung der Gesundheitsbehörde, Zeitverlust, finanzieller Schaden, Imageschaden.

Ein weiteres Negativbeispiel: Eine spanische Modefirma stellt Armbänder her. Bei der Übersetzung der Werbeanzeige wurde das „pulsera esclava“ (Sklavenarmband) im Französischen zu „bracelet esclave“. In Frankreich wird diese Bezeichnung nicht verwendet. Die Übersetzung rief einen Sturm der Entrüstung hervor. Die Folgen: →

## ANGELIKA OTTMANN

Als Diplom-Übersetzerin und ehemalige langjährige Geschäftsführerin eines Übersetzungsdienstleisters kommt Angelika Ottmann aus der Praxis und ist vertraut mit allen Aspekten des Übersetzungsmanagements. Sie ist Herausgeberin der „Best Practices – Übersetzen und Dolmetschen“ und arbeitet als unabhängige Expertin für Qualitäts- und Projektmanagement sowie Risikomanagement für Übersetzungen.

[ottmann@risikoscouts.de](mailto:ottmann@risikoscouts.de)  
[www.risikoscouts.de](http://www.risikoscouts.de)



## CARMEN CANFORA

Als Diplom-Übersetzerin und Dozentin für Terminologie, Fachübersetzen sowie Normen und Richtlinien für Fachübersetzer verbindet Dr. Carmen Canfora Theorie und Praxis. Sie hat in Übersetzungswissenschaft promoviert und arbeitet als unabhängige Expertin für technisches Fachübersetzen und Terminologiemanagement sowie Risikomanagement für Übersetzungen.

[canfora@risikoscouts.de](mailto:canfora@risikoscouts.de)  
[www.risikoscouts.de](http://www.risikoscouts.de)



## Spielen. Und Gutes tun.

itl wird 35. Treffen Sie nicht nur unsere Experten, sondern feiern Sie mit uns am Stand auf der tekcom-Jahrestagung!

Sie **spielen**, wir **spenden**: Mit der vom Mobile & Media-Team entwickelten QuizApp! Alle erspielten Spenden gehen an Translators Without Borders.

Wir freuen uns auf Sie!

**Halle C2, Stand 21E06**



+49 89 892623-0  
[info@itl.eu](mailto:info@itl.eu)

[www.itl.eu](http://www.itl.eu) / [itl.ch](http://itl.ch) / [itl.at](http://itl.at)

LISA QA MODEL: ÜBERSICHT FEHLERKATEGORIEN BEI SW USER INTERFACE

Documentation	Help	Software
<b>Doc Language (7)</b> → Mistranslation → Accuracy → Terminology → Language → Style → Country → Consistency <b>Doc Formatting (6)</b> → TOC → Index → Layout → Typography → Graphics → Call Outs and Captions <b>Doc Formatting – Asian (8)</b> → TOC → Index → Layout → Typography → Graphics → Call Outs and Captions → Double/Single Size → Punctuation Marks <b>21 total</b> <b>TOTAL: 39 total (25 unique)</b>	<b>Help Formatting (4)</b> → Index → Layout → Typography → Graphics <b>Help Formatting – Asian (4)</b> → Index → (Layout) → Graphics → (Typography) → Localizable Text → Hypertext Functionality, jumps, popups <b>8 total</b>	<b>Software Formatting (5)</b> → Graphics → Alignment → Sizing → Truncation/Overlap → Character Formatting <b>Software Functionality Testing (5)</b> → Localizable Text → Dialogue Functionality → Menu Functionality → Hotkeys/Accelerators → Jumps/Links <b>10 total</b>

TAB. 01 [5]

→ Bericht in internationalen Zeitschriften, Boykottaufruf in den Sozialen Medien, Zeitverlust, finanzieller Schaden, Imageschaden.

Wie es aussieht, sind sich die Auftraggeber der Risiken oft nicht bewusst, die durch fehlerhafte Übersetzungen entstehen. Auch manche Einkaufsabteilung folgt eher dem vermeintlichen Wirtschaftlichkeitsgebot („Wer ist der billigste Anbieter?“) als einer Risikostrategie.

Ein professionelles Risikomanagement für Übersetzungen kann das Risiko in diesem Bereich signifikant senken [2, 3]. Die Frage ist aber: Wie kann nachgewiesen werden, dass das Risikomanagement tatsächlich greift?

### Welches Prinzip ist am häufigsten?

Beim Risikomanagement verhält es sich so wie beim Qualitätsmanagement: Werden Maßnahmen eingeführt, die bestimmte Dinge bewirken sollen, muss nachgeprüft werden, ob die gewünschte Wirkung erzielt wird. In unserem Fall bedeutet dies, dass die Übersetzungen einer Prüfung und Bewertung unterzogen werden müssen. Für die Bewertung (Evaluierung) von Übersetzungen gibt es allerdings noch keine festgeschriebenen Standards. Die Branche kennt verschiedene Evaluierungsmodelle, die unterschiedlichen Zwecken dienen.

Die am häufigsten verwendeten Modelle zählen die Fehler. Die Bewertung erfolgt anhand der Fehlerzahl in verschiedenen Fehlerkategorien, gegebenenfalls mit verschiedenen Schweregraden (Severity) wie leicht, schwer oder kritisch. Die Fehlerkategorien können unterschiedlich gewichtet werden. So kann ein Rechtschreibfehler weniger wichtig sein und weniger Fehlerpunkte erhalten als ein Terminologiefehler mit mehr Fehlerpunkten.

Am Ende wird eine Fehlerkennzahl oder ein Fehlerkoeffizient errechnet, der die Güte der Übersetzung angibt. Diese Fehlerkennzahl wird dann mit einem vorher festgelegten Schwellenwert verglichen. Wird der Schwellenwert überschritten, ist die Übersetzung durchgefallen und damit nicht verwendbar.

### Welche Modelle bietet der Markt?

Das *SDL TMS Classic Model* ist ein sehr einfaches Bewertungsmodell [4]. Es kann im Translation-Memory-System *SDL Trados Studio* als Grundlage für die Bewertung von Übersetzungen festgelegt werden. Das Modell umfasst sieben linguistische Fehlerkategorien, die ohne Unterscheidung zwischen verschiedenen Schweregraden oder Gewichtungen jeweils mit einem Fehlerpunkt bewertet werden:

- Overall Grading (Gesamtbewertung)
- Translation Accuracy (Genauigkeit)
- Grammar (Grammatik)
- Spelling (Rechtschreibung)
- Style (Stil)
- Terminology (Terminologie)
- Technical Accuracy (Fachliche Genauigkeit)

Das *SDL-Modell* hat den Vorteil, dass es nur wenige Fehlerkategorien und keine Gewichtung gibt, so dass es leicht anwendbar ist. Es können auch Übersetzungen bewertet werden, bei denen es auf den Stil ankommt. Allerdings dürfte es für die meisten Übersetzungsworkflows nicht flexibel genug sein, da eine Gewichtung der Fehler nach dem Schweregrad häufig sinnvoll ist. Außerdem werden nicht-sprachliche Fehler nicht bewertet, zum Beispiel Fehler im Layout.

Das *LISA QA Model* wurde von der Localization Industry Standards Association, kurz *LISA*, ursprünglich für die Bewertung von Übersetzungen von Software-Oberflächen entwickelt. Die Beurteilung von Übersetzungen erfolgt nach festgelegten Kriterien mit unterschiedlicher Gewichtung verschiedener Fehlertypen. Die wichtigsten Eigenschaften:

- Das *LISA QA Model* unterscheidet drei Schweregrade, nämlich leicht, schwer, kritisch (minor, major, critical). Das Modell gibt aber keine Gewichtung vor.
- Zusätzlich zu sprachlichen Fehlertypen werden weitere Typen berücksichtigt. Dazu zählen Fehler im Layout und solche, die die Funktionalität der Software beeinträchtigen, zum Beispiel Links.

Das Modell umfasst zahlreiche Fehlerkategorien, gegliedert in zwei Bereiche:

1. Software User Interface: 25 Kategorien in sieben Gruppen mit einigen Redundanzen
2. Dokumentation: 20 Kategorien mit einigen Redundanzen

Das *LISA QA Modell* wurde nie als Standard herausgegeben, die Lokalisierungsbranche hat es aber zeitweise als solchen anerkannt. Eine Übersicht über die Fehlerkategorien im Bereich Software User Interface bieten Tabelle 1 und 2.

Das *LISA-Modell* ist deutlich komplexer und differenzierter als das von *SDL*. Für viele Anwendungen eignet es sich besser, vor allem weil es Gewichtungen zulässt. Allerdings ist es stark auf die Bedürfnisse der Lokalisierungsindustrie zugeschnitten, für die es ursprünglich vorgesehen war. Die Fülle an Kategorien kann irritieren und erfordert Schulungen.

Das Modell auf Grundlage der Norm *SAE J2450* wurde von der Society of Auto-

otive Engineers (SAE) entwickelt und speziell auf die Bedürfnisse der Automobilindustrie zugeschnitten. Heute wird es jedoch auch in anderen Bereichen verwendet, zum Teil mit individuellen Erweiterungen. Das Modell hat sowohl unterschiedlich gewichtete Fehlerkategorien (insgesamt sieben) als auch zwei Schweregrade („serious“ und „minor“). Es stellt allerdings keine Kriterien für die Unterscheidung zwischen serious und minor auf. Der Technische Redakteur kann das Modell einfach handhaben, es konzentriert sich nur auf die sprachliche Qualität. Eine Bewertung des Stils wurde bewusst ausgeschlossen. Die Entscheidung, in welche Fehlerkategorie ein Fehler einzuordnen und welcher Schweregrad zuzuordnen ist, unterliegt zwei Meta-Regeln:

1. Im Zweifel ist die Fehlerkategorie mit der höheren Punktezahl (weiter oben in der Liste) zu wählen.
2. Im Zweifel ist ein Fehler als schwerwiegend zu betrachten.

Der Schaden, den ein Fehler möglicherweise anrichten kann („impact“), ist in diesem Modell ein Kriterium für die Fehlergewichtung (TAB. 03, S. 39).

Das SAE J2450-Modell kommt mit wenigen Fehlerkategorien aus und ist damit

LISA QA MODEL: GEGENÜBERSTELLUNG DER FEHLERKATEGORIEN		
In both UI and documentation	In UI only	In documentation only
Accuracy	Alignment	Arts & Design
Call Outs and Captions	Character Formatting	Compatibility
Consistency	Dialogue Functionality	Cross-references
Country	Double/Single Size	Delivery
Graphics	Hotkeys/Accelerators	Functionality
Index	Hyper text functionality, jumps, popups	Interface
Language	Jumps/Links	Output
Style	Layout	Printing
Table of Contents (TOC)	Localizable Text	References
Typography	Menu Functionality	Terminology
	Mistranslation	Typography
	Punctuation Marks	
	Sizing	
	Terminology	
	Truncation/Overlap	
<b>TOTAL: 10</b>	<b>Total: 15</b>	<b>Total: 11</b>

TAB. 02 [5]

übersichtlicher als das von LISA. Trotzdem lässt es Fehlergewichtungen zu und verfügt daher für viele Übersetzungsworkflows über eine ausreichende Komplexität, vor allem im Bereich der Technischen Dokumentation. Als einziges Modell gibt es dem Prüfer Kriterien für die Gewichtung der Fehler an die Hand. Allerdings ist es für alle Übersetzun-

gen ungeeignet, bei denen Stil eine wichtige Komponente ist, denn diese Fehlerkategorie fehlt. Auch andere nicht-sprachliche Fehlerkategorien, die für bestimmte Übersetzungen wichtig sein könnten, sind nicht vorhanden, zum Beispiel Layout.

Das Modell Multidimensional Quality Metrics, kurz MQM, wurde zunächst →



**Seminare & Training**

# Expertenseminare am SDI München

Auch als Inhouse-Workshops buchbar!

zu Themen wie

- Risikomanagement,
  - Gelungene Kundenkommunikation und
  - Effiziente Prozesse
- Technische Redaktion / Übersetzung**

ermöglichen neue Perspektiven, Effizienzsteigerungen und Erfahrungsaustausch.

[sdi-muenchen.de/seminare](http://sdi-muenchen.de/seminare)



## tekom-Jahrestagung 2017

Herzliche Einladung zu unseren Fachvorträgen von Karsten Schrempp:



24. Okt. von 9.45–10.30 Uhr

**Raum C9.2**

**Der Topicflut Herr werden – was leistet eine Metadatenplattform?**



25. Okt. von 9.45–10.30 Uhr

**Raum C6.2**

**Letzter Blick in die Werkstatt:**  
Einführung eines CDP bei TRUMPF

Wir freuen uns über Ihren Besuch auf unserem Stand **2/F33**

[www.pantopix.de](http://www.pantopix.de)



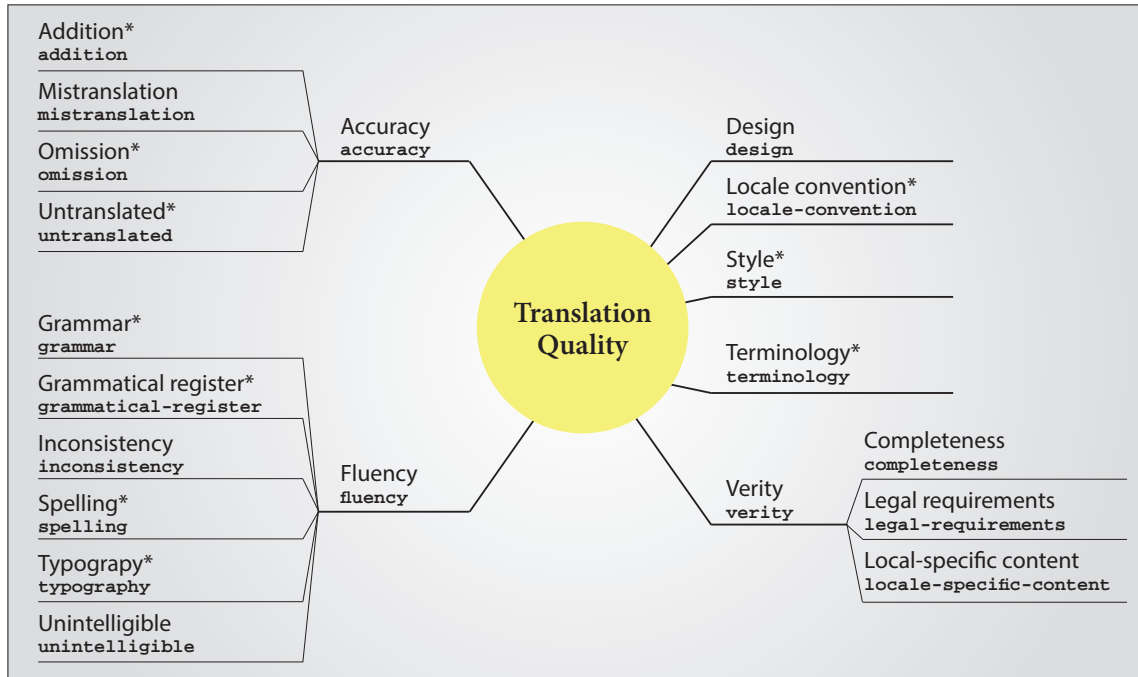


ABB. 01 Kriterien im MQM-Modell [7].

→ für die Beurteilung maschineller Übersetzungen entwickelt, auch mit dem Ziel, das „Lernen“ der Maschine zu verbessern. Inzwischen wird es auch für die Bewertung von Humanübersetzungen genutzt. Gemessen wird nach folgenden Kriterien:

- Adequacy (Angemessenheit/ Zweckdienlichkeit)
- Fluency (Flüssigkeit/Gewandtheit)
- Verity (Originaltreue)
- Internationalization (Internationalisierung)
- Locale convention (Gebietsschema)
- Style (Stil)
- Terminology (Terminologie)
- Design
- Other (Sonstige).

Abbildung 1 zeigt einen Überblick über das Modell.

Die neun Hauptkategorien werden insgesamt über 100 sprachliche und nicht-sprachliche Unterkategorien verästelt, die

vom Prüfer je nach gewünschter Beurteilungstiefe ausgewählt werden können. Die Qualität des Ausgangstextes wird berücksichtigt.

Das Modell unterscheidet „potenzielle Probleme“ (issues) und Fehler. Zu den potenziellen Problemen zählt zum Beispiel ein nicht mit dem Wert im Ausgangstext übereinstimmender Preis in einem Katalog. Wenn der andere Preis für den Zielmarkt korrekt ist, aufgrund von Preisunterschieden auf den verschiedenen Märkten, handelt es sich nicht um einen Fehler. Fehler sind hingegen in einem Text festgestellte Probleme, zum Beispiel wenn der andere Preis aufgrund einer Fehlübersetzung nicht stimmt.

Im MQM-Modell werden die Probleme durch eine automatische Routine herausgesucht, aber nur ein Humanübersetzer ist in der Lage zu entscheiden, ob wirklich ein Fehler vorliegt. Das MQM-Modell kann mit einer hohen Flexibilität und einem umfassenden Kategorienkatalog punkten, der an die

jeweiligen Anforderungen des Übersetzungsworkflows angepasst werden kann. Die hohe Komplexität hat jedoch auch ihren Preis: Prüfer müssen zunächst umfassend geschult werden und das Modell kann überfordern. Zudem ist keine Fehlergewichtung vorgesehen.

Größere Konzerne haben Erweiterungen der genannten Modelle im Einsatz oder auch eigene Modelle entwickelt. Wichtig ist, dass die Bewertungskriterien dem Auftragnehmer klar dargelegt werden.

Zu den genannten Modellen existieren weitere, um Übersetzungen in der Ausbildung zu beurteilen. Diese Modelle spielen in der Praxis keine Rolle. Eine Brücke zwischen Theorie und Praxis schlägt aber zum Beispiel das Modell von Elvira Mertin [8].

**Was ist das Ziel von Evaluierung?**

In der industriellen Praxis hat eine Evaluierung das Ziel, Übersetzungen zu bewerten und herauszufinden, ob sie den Qualitätsan-

ABB. 02 Beispiel für ein Evaluierungsformular mit Gesamtbewertung; die Fehlertypen und Werte sind willkürlich.  
QUELLE Die RisikoScouts

Evaluierungsbogen für Übersetzungen			
Fehlertyp	Fehlerpunkte pro Fehler	Anzahl Fehler Fehler	Gesamtpunkte pro Fehlertyp
Sprachliche Fehler	1	1	1
Sinnfehler	10	1	10
Stilfehler	5	1	5
	0	0	0
	0	0	0
	0	0	0
Gesamtfehlerpunkte			16
Gesamtumfang des Texts in Wörtern			1000
Fehlerfaktor			1,60%
Fehlerpunkte pro 1000 Wörter			16,00
Schwellenwert			40
Pass/Fail			PASS

**SAE J2450: FEHLERKATEGORIEN UND IHRE GEWICHTUNG**

Fehlerkategorie (Fehlertyp)	Fehlerpunkte (schwer/leicht)
A. Wrong Term (Terminologiefehler – WT)	5/2
B. Wrong Meaning (Inhaltlicher Fehler – WM)	5/2
C. Omission (Auslassung – OM)	4/2
D. Structural Error (Strukturfehler – SE)	4/2
E. Misspelling (Rechtschreibfehler – SP)	3/1
F. Punctuation Error (Interpunktionsfehler – PE)	2/1
G. Miscellaneous Error (Sonstiger Fehler – ME)	3/1

TAB. 03 [6]

forderungen genügen. Die genannten Modelle haben jedoch keine speziellen Kriterien, die Rückschlüsse auf eine Änderung des Risikos zulassen. Es gibt zwar unterschiedliche Gewichtungen (Schweregrad) für verschiedene Fehlertypen und innerhalb eines Fehlertyps noch eine Unterscheidung in „weniger schwerwiegend (minor)“ oder „schwerwiegend (major/serious)“. Diese Einschätzung ist schwierig und natürlich subjektiv und die Entscheidung hängt vom jeweiligen Prüfer ab. Feste Kriterien, anhand derer der Prüfer diese Einstufung vornehmen kann, existieren nicht. Lediglich in der SAE J2450 gibt es die erwähnte Anweisung, einen Fehler auch dann als „schwerwiegend“ einzustufen, wenn die Entscheidung für den betreffenden Fehler schwerfällt. Außerdem sagt die Norm, dass ein Fehler immer als „schwerwiegend“ einzustufen ist, wenn dadurch ein Schaden verursacht werden kann.

Die bestehenden Modelle liefern also kaum Erkenntnis, die dem Auftraggeber in Bezug auf Risiken im Übersetzungsbebereich weiterhelfen. Die bisherigen Evaluierungsmodelle basieren auf unterschiedlichen Qualitätsmerkmalen, die allerdings nicht risikoorientiert sind. Somit sagt ein Evaluierungsergebnis auch nichts über die tatsächliche Veränderung des Risikos durch die getroffenen Maßnahmen aus.

**Wie können Risiken kontrolliert werden?**

Wenn Maßnahmen zur Risikominimierung bei Übersetzungen eingeführt werden, müssen Auftraggeber, zum Beispiel Technische Redakteure, nachprüfen können, ob die Maßnahmen erfolgreich sind oder nicht. Ein Instrument zur Überprüfung ist die risikobasierte Evaluierung der Übersetzungen. Sie zeigt, ob eine Übersetzung risikoträchtige Fehler enthält. Der Vorteil: Wenn man die aktuellen mit den früheren Evaluierungsergebnissen vergleicht, lässt sich auf einen Blick erkennen, ob die Ziele erreicht und wirklich die Risiken der Übersetzung gesenkt wurden. Das kann zum Beispiel anhand jährlicher Durchschnittswerte geschehen.

Mit der risikobasierten Evaluierung befinden wir uns im Risikocontrolling – Überwachung und Überprüfung. Dies ist Aufgabe des Auftraggebers. Risikocontrolling hat nichts mit Qualitätssicherung oder Revision zu tun. Die Evaluierung prüft vielmehr, inwieweit die Leistung den Anforderungen entspricht, also ob die risikomindernden Maßnahmen umgesetzt wurden und wirksam sind. Das machen nicht die Übersetzungsdienstleister, da diese auch nicht alle Risiken und vor allem nicht deren Wertigkeit im Unternehmen kennen. Ein Dienstleister kann nicht entscheiden, ob für einen Auftraggeber ein Imageschaden schlimmer ist als ein Sachschaden. Dienstleister können den Auftraggeber höchstens beraten.

**Welche Schritte hat die Evaluierung?**

Bei der Evaluierung werden die Abweichungen von der vereinbarten Qualität ermittelt. Im ersten Schritt müssen daher die Bewertungskriterien festgelegt werden, also die Merkmale, mit denen das Ideal oder die Abweichung davon beschrieben wird. Anschließend wird die Gewichtung jedes Fehlertyps bestimmt. Sie legt fest, wie schwerwiegend der Fehlertyp ist. In einem weiteren Schritt muss festgelegt werden, wie viel Abweichung der Auftraggeber tolerieren kann. Das ist der Schwellenwert, über dem die Qualität der Übersetzung den Anforderungen entspricht oder unter dem die Qualität den Anforderungen nicht mehr entspricht.

Aus Tabelle 3 lässt sich ersehen, dass zum Beispiel bei der SAE J2450 ein inhaltlicher Fehler (WM) mit einer höheren Punktezahle belegt ist und schwerer wiegt als ein Interpunktionsfehler (PE).

In allen gängigen Modellen werden dann die gezählten Fehler mit den jeweiligen Fehlerpunkten multipliziert und alle Werte aufaddiert. Diese Fehlerpunktzahl (Summe aller Punkte) wird dann in Relation zur Menge des geprüften Textes (zum Beispiel Wortzahl) gesetzt und ergibt einen Fehlerfaktor (zum Beispiel Anzahl Fehlerpunkte pro 1.000 Wörter). Dies ermöglicht den Ver- ➔

# Wir lieben Technik



und gehen den Dingen gerne auf den Grund. Engagiert und kompetent. Mehr Informationen unter [www.dokuwerk.de](http://www.dokuwerk.de)

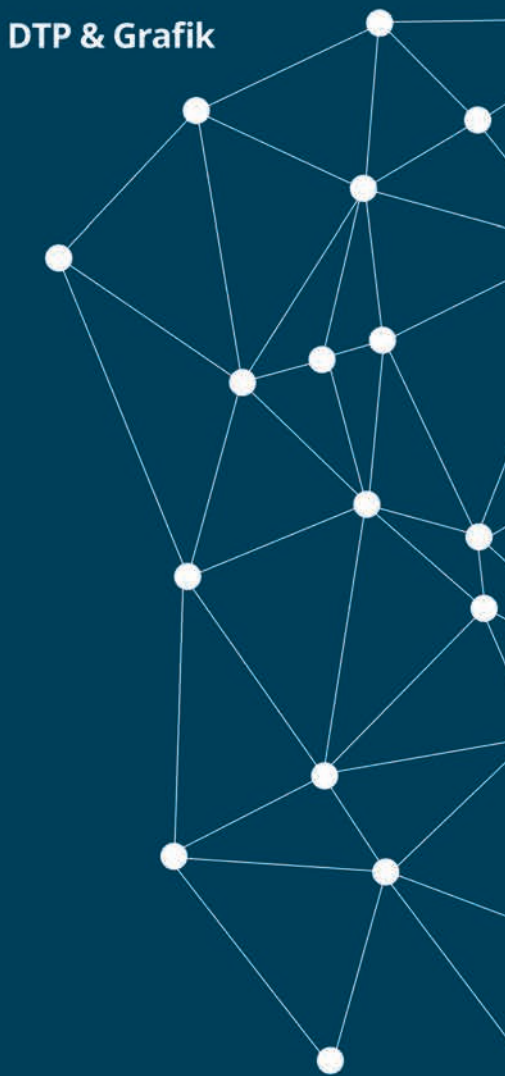
Übersetzungen  
für Medizin,  
Life Sciences, IT und  
Technologie

Zertifiziert nach ISO 17100

Übersetzung in alle Sprachen

Software-Lokalisierung

DTP & Grafik



→ gleich aller Übersetzungen einer bestimmten Textsorte sowie den Abgleich mit einem gewünschten Schwellenwert. Der Schwellenwert gibt an, ab wann eine Übersetzung „durchgefallen“ ist. Abbildung 2, S. 38, zeigt ein Beispiel für ein Evaluierungsformular mit Gesamtbewertung.

Eine sehr schwierige Entscheidung ist die Festlegung des Schwellenwertes. Ist die Übersetzung mit 16 Fehlerpunkten pro 1.000 Wörtern durchgefallen oder nicht?

#### Welche Konsequenzen hat ein Fehler?

Die Überlegung, für verschiedene Textsorten unterschiedliche Fehlergewichtungen festzulegen, geht schon in Richtung einer risikobasierten Evaluierung. Die risikobasierte Evaluierung geht aber noch einen Schritt weiter. Sie erfordert ein Nachdenken darüber, welche Konsequenzen ein Übersetzungsfehler haben kann, und legt die Gewichtung risikoorientiert fest.

Wir fragen uns also zunächst, ob ein Fehlertyp relevant für ein bestimmtes Risiko ist. Zum Beispiel: Hat ein (nicht sinnentstellender) Rechtschreibfehler Auswirkungen auf Leib und Leben des Benutzers der Übersetzung? Erhöht ein Sinnfehler das Risiko eines Sachschadens? Erhöht ein Stilfehler das Risiko, dass die Maschine nicht mehr bedient werden kann? Nur wenn wir diese Frage mit „Ja“ beantworten, ist die „Wertung“ des betreffenden Fehlertyps gerechtfertigt. Fehler, die keinen Einfluss auf das Risiko einer Übersetzung haben, dürften dann für diese Textsorte nicht gewertet werden. Das heißt, sie müssen erst gar nicht erfasst werden.

Im nächsten Schritt stellen wir uns die Frage, wie hoch der Einfluss des gewerteten Fehlers auf das jeweilige Risiko ist. Daraus leiten wir dann die Gewichtung des Fehlers (geringes, mittleres oder hohes Risiko) und damit die Fehlerpunkte ab. Diese wenigen Fragen verändern die Evaluierung einer Übersetzung grundlegend. Fehler, die nicht risikorelevant sind, fallen nicht ins Gewicht. Eine Übersetzung fällt nur dann durch, wenn der Technische Redakteur risikorelevante Fehler festgestellt hat.

Aus der risikobasierten Evaluierung können Risikokennzahlen abgeleitet werden. Die Kennzahlen können gezielt dazu genutzt werden, um die Veränderung von Risiken in Übersetzungen über die Zeit zu überprüfen.

#### Was soll am Ende herauskommen?

Für die Evaluierung von Übersetzungen stehen verschiedene Modelle und Metriken zur Verfügung. Bei der Auswahl der geeigneten Metrik ist zu beachten, welches Ziel mit der Evaluierung verfolgt wird. Wenn es das Ziel

ist, die Qualität von Übersetzungen zu prüfen, muss ein qualitätsorientiertes Modell mit einer entsprechenden Metrik verwendet werden. Dabei ist zu beachten, ob die Metrik für Maschinelles Übersetzen oder Humanübersetzen vorgesehen ist und ob sie für die eigene Branche überhaupt sinnvoll angewendet werden kann. Wenn es jedoch das Ziel ist, möglichst risikoarme Übersetzungen zu erhalten, muss die Prüfung anhand einer risikobasierten Metrik erfolgen. Hans Uzkoreit und Arle Lommel haben in einem Vortrag zusammengefasst, warum eine differenzierte Betrachtung der Übersetzungsevaluierung besonders wichtig ist [9]:

- „A legal treaty ≠ a service manual ≠ a “gist” translation ≠ a summary translation
- If you assess with the wrong metric, you waste time and money
- Assessment should determine fitness for purpose“

Eine falsche Metrik kann also viel Zeit und Geld kosten und am Ende am Ziel vorbeigehen – ein Patentrezept für die Auswahl gibt es nicht. ☹

#### LINKS UND LITERATUR ZUM BEITRAG

- [1] DIN-Verbraucherrat (2010): *Bedienungsanleitungen, Folgen fehlerhafter Anleitungen am Markt*, S. 115.
- [2] Ottmann, Angelika; Canfora, Carmen (2015): *Im grünen Bereich? In: technische kommunikation*, H. 5, S. 34–39.
- [3] Ottmann, Angelika; Canfora, Carmen (2016): *Risiken unter Kontrolle. In: technische kommunikation*, H. 5, S. 44–47.
- [4] SDL TMS Classic QA Model: [http://producthelp.sdl.com/SDL\\_TMS\\_2011/en/Creating\\_and\\_Maintaining\\_Organizations/Managing\\_QA\\_Models/TMS\\_Classic\\_QA\\_Model.htm](http://producthelp.sdl.com/SDL_TMS_2011/en/Creating_and_Maintaining_Organizations/Managing_QA_Models/TMS_Classic_QA_Model.htm).
- [5] Doherty, Stephen; Gaspari, Federico (2013): *Understanding and Implementing Effective Translation Quality Evaluation Techniques*. [www.qt21.eu/launchpad/sites/default/files/QTLTP%20GALA%20Webinar%203.pdf](http://www.qt21.eu/launchpad/sites/default/files/QTLTP%20GALA%20Webinar%203.pdf).
- [6] Dalla-Zuanna, Jean-Marc (2017): *Qualität messbar gemacht. Bewertung von Fachübersetzungen gemäß SAE-Norm J2450*. In: MDÜ 1/2017, S. 14–20.
- [7] QT Launchpad (2015): [www.qt21.eu/mqm-definition/definition-2015-12-30.html](http://www.qt21.eu/mqm-definition/definition-2015-12-30.html).
- [8] Mertin, Elvira (2006): *Prozessorientiertes Qualitätsmanagement im Dienstleistungsbereich Übersetzen*. Frankfurt a. M.: Lang.
- [9] Uzkoreit, Hans; Lommel, Arle (2013): *Multidimensional Quality Metrics: A New Unified Paradigm for Human and Machine Translation Quality Assessment*. Vortrag auf der Localization World Conference 2013, London. [www.qt21.eu/launchpad/sites/default/files/MQM.pdf](http://www.qt21.eu/launchpad/sites/default/files/MQM.pdf).

#### LITERATUR ZUM WEITERLESEN

- Canfora, Carmen; Ottmann, Angelika (2015): *Risikomanagement für Übersetzungen*. In: *trans-kom*, H. 8 [2], S. 314–346.
- DIN EN ISO 17100:2016-05: *Übersetzungsdienstleistungen – Anforderung an Übersetzungsdienstleistungen (ISO 17100:2015)*; Deutsche Fassung der EN ISO 17100:2015. Berlin: Beuth.
- DIN EN ISO 9001:2015: *Qualitätsmanagementsysteme – Anforderungen*. Berlin: Beuth.
- Ottmann, Angelika (Hrsg.) (2017): *Best Practices – Übersetzen und Dolmetschen*. Berlin: BDÜ Fachverlag.