Deutscher Terminologie-Tag e.V.

21. Jahrgang | www.dttev.org

edition

Fachzeitschrift für Terminologie

1|25

Terminologie trifft Videospielwelten

Halluzinationen
Von der Psychiatrie zu LLMs
Seite 15

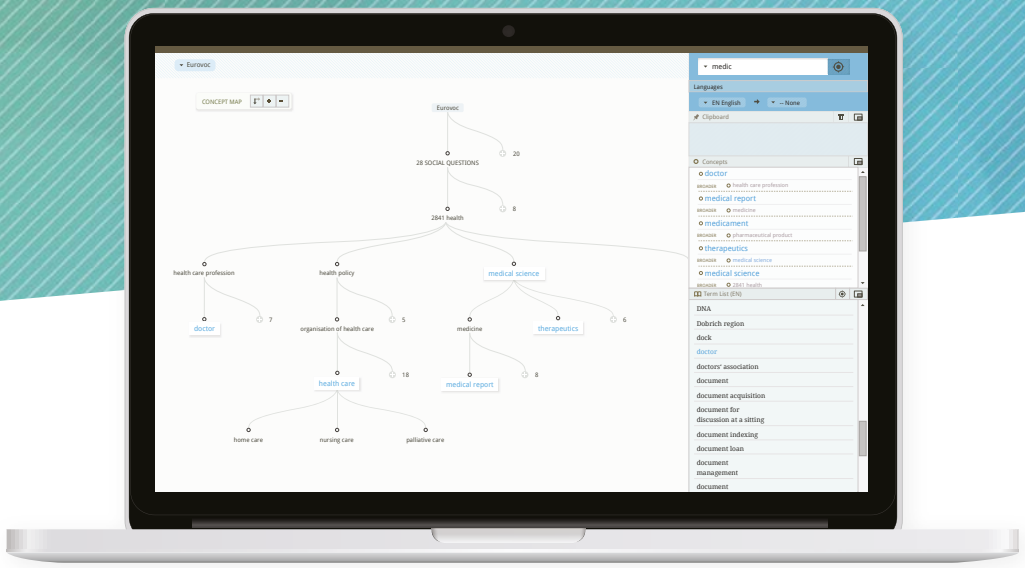
Wegweiser
Multilinguale Glossare für NMT
Seite 19

Tipps
Arbeitsoptimierung & -automatisierung
Seite 26

DTT-Update
Vorstand, Fortbildung, Events
Seite 33

Your Multilingual Knowledge System

Extreme Visual Terminology



360-degree control through systematic approach: **multilingual Concept Maps**

Proven enterprise deployments, scalable with Single sign-on: appealing, **plug-in free browser solution**

Enable NMT workflows, tune enterprise search, annotate and classify texts: **versatile integrations**

coreon

Knowledge meets language.

coreon.com

Zwischen Elben, Engines und Excel – Terminologie am Puls der Zeit

Liebe Leserinnen und Leser,

wer heute ein Videospiel startet, betritt nicht nur eine neue Welt, sondern oft auch ein neues Sprachuniversum. Die Lokalisierung von Videospielen verlangt weit mehr als eine reine Übersetzung: Sie bewegt sich im Spannungsfeld zwischen terminologischer Präzision, kultureller Sensibilität und kreativer Freiheit. Genau diesem faszinierenden Zusammenspiel widmet sich unser Titelbeitrag „Zwischen Fachsprachen und Fantasiewelten: interdisziplinäre Terminologie in der Videospiellokalisierung“.

Doch auch jenseits der Fantasiewelten gibt es spannende Entwicklungen in der Terminologiearbeit zu entdecken: So beleuchtet ein Beitrag die Begriffsgenese von „Halluzination“ – ein Wort, das derzeit eine überraschende zweite Karriere im Kontext von KI und Large Language Models erlebt. Ist es wirklich die passende Bezeichnung für maschinellen Unsinn? Der Artikel lädt zur kritischen Reflexion ein.

Handfester wird es in unserem Beitrag über „Glossare für die neuronale maschinelle Übersetzung“. Hier zeigt sich, wie gezielte terminologische Steuerung hilft, automatisierte Übersetzungen zu optimieren – ein Thema, das für viele in der Praxis zunehmend an Bedeutung gewinnt.

Apropos Optimierung: Wer seine tägliche Arbeit effizienter gestalten möchte, findet in „Arbeitsoptimierung und -automatisierung für den Hausgebrauch“ praktische Tipps für den Umgang mit Tools, die den Alltag erleichtern. In der Rubrik Tools & Ressourcen stellen wir zudem „Terminologie der Information und Dokumentation & DGI-Term“ vor.

Im Verbandsteil informieren wir über das neue DTT-Fortbildungsprogramm und berichten von der Tagung des Industrieforums.

Abgerundet wird das Heft durch einen Rückblick auf das DTT-Symposion 2025, die Vorstellung der Preisträgerinnen des DTT-Förderpreises sowie ein Porträt unseres neuen 1. Vorsitzenden Peter Seip, den wir an dieser Stelle ganz herzlich willkommen heißen. Ein großer Dank geht auch an unseren ehemaligen 1. Vorsitzenden, Tom Winter, für seine engagierte und zukunftsweisende Arbeit im Vorstand in den letzten acht Jahren.

Wir wünschen Ihnen eine anregende Lektüre zwischen Theorie, Technik und Terminologiekunst.

Herzlichst
Ihre Redaktion

geschrieben unter Mitwirkung von ChatGPT

Dr. Nicole Keller
Redaktionsleitung
redaktion@dttev.org

Angelika Ottmann
Redaktionsleitung
redaktion@dttev.org

TERM SOLUTIONS
termXact • termXplorer • myTerm

Haben Sie Ihr Terminologie-management im Griff?

Falls nicht, sollten Sie baldmöglichst dafür sorgen – vor allem wenn Sie Künstliche Intelligenz im Unternehmen nutzen wollen. Denn:

There is no AI without terminology.

Ob in ChatBots, beim Prompting oder in der Maschinellen Übersetzung: Das Corporate Wording muss auch bei KI-erstellten Texten berücksichtigt werden.

Unsere ausgezeichneten Terminologie-Tools **myTerm**, **termXplorer** und **termXact online** unterstützen Sie dabei:

- **Multilingual** durch 3-Ebenen-Modell
- **Flexibel** für maßgeschneiderte Abstimmungsprozesse
- **Planbar** dank firmenweiter Lizenz
- **Skalierbar** als Pay-per-Feature
- **User-zentriert** mit übersichtlichem UX-Design und Filter-Dashboard
- **Leicht integrierbar** über einfachen Import und Export sowie REST-API und Single-Sign-On

Greifen Sie zu.

Alle Infos auf **termsolutions.ai**

Inhalt

Editorial

3 Zwischen Elben, Engines und Excel – Terminologie am Puls der Zeit
Nicole Keller und Angelika Ottmann

Themen

5 Zwischen Fachsprachen und Fantasiewelten: Interdisziplinäre Terminologie in der Videospiellokalisierung
Annika Sophie Dürr

15 Halluzination – eine irreführende Bezeichnung für irreführenden Output von Large Language Models
Annette Gerstenberg

19 Glossare für die neuronale maschinelle Übersetzung
Detlef Reineke

Tools & Ressourcen

26 Arbeitsoptimierung und -automatisierung für den Hausgebrauch
Maurice Mayer-Dewor

31 Terminologie der Information und Dokumentation & DGI-Term
Annette Weiland und Axel Ermert

Aus dem Verband

33 Das neue DTT-Fortbildungsprogramm
Christian Kriele

35 Wechsel auf dem DTT-Chefsessel
Nicole Keller und Angelika Ottmann

36 Neues vom Industrieforum
Tom Winter

37 Ich nehme Sie mit auf eine Reise nach Worms
Petra Sweekhorst

39 DTT-Förderpreis 2025: Zweifache Exzellenz
Nicole Keller und Angelika Ottmann

21. Jahrgang.
edition erscheint zweimal im Jahr (Juni/Dezember).
ISSN 1862-023X.
Herausgeber: Deutscher Terminologie-Tag e.V. (DTT) – www.dttev.org
Redaktionsleitung: Dr. Nicole Keller und Angelika Ottmann – redaktion@dttev.org
Titelseite (nach einer KI-Idee), DTT-Anzeigen, Layout und Satz: Tamara Arndt – layout@dttev.org
Anzeigen: Olga Buchstaller-Vodopiyanova – anzeigen@dttev.org
Lektorat: Maurice Mayer-Dewor
Mediadaten und komplettes Impressum: www.dttev.org/edition/impressum

Zwischen Fachsprachen und Fantasiewelten: Interdisziplinäre Terminologie in der Videospiellokalisierung

Annika Sophie Dürr

Video game localization is a unique interplay of creativity and precision, blending technical, gaming, and corporate terminology with linguistic and cultural adaptation. Faced with constraints like context gaps, variables, and space limits, translators navigate complex challenges where terminology can make or break the player experience.

Keywords: video game localization, terminology, interdisciplinarity, games, transcreation

Spielen Sie Videospiele? Vielleicht gelegentlich auf dem Smartphone oder regelmäßig auf PC und Konsole? Überraschend wäre das nicht: Ob als Zeitvertreib, Hobby, in der Bildung, durch Gamification oder als eSport – Videospiele sind längst fest in unserer Gesellschaft verankert. Ihre wirtschaftliche Bedeutung zeigt sich in beeindruckenden Zahlen: Allein in Deutschland erzielte der Gaming-Markt 2023 einen Umsatz von fast 10 Milliarden Euro. Globale Erfolgstitel wie *Minecraft* (300 Millionen verkaufte Exemplare weltweit) oder *Grand Theft Auto V* (205 Millionen) verdeutlichen den immensen Stellenwert der Branche. Videospiele ermöglichen es uns, von passiven Rezipient:innen zu aktiven Protagonist:innen zu werden – wir erleben die Handlung nicht nur, sondern gestalten sie selbst mit. Diese Interaktivität bietet ein einmaliges Immersionspotenzial und stellt den zentralen Reiz des Mediums dar. Bernal Merino, ein Vorreiter der wissenschaftlichen Untersuchung der Videospiellokalisierung, beschreibt Videospiele als eine Verbindung der narrativen Struktur der Literatur mit audiovisuellen Elementen aus Film und Fernsehen, erweitert um das einzigartige Merkmal der Interaktivität [1]. Sie sind dabei „both a product designed for mass consumption and an artistic team creation at the same time“ [2].

Diese heutzutage hochkomplexen narrativen Welten erfordern sorgfältige Lokalisierungs- und Terminologiearbeit. Obwohl Englisch als Lingua Franca der Industrie und Spielgemeinschaft etabliert ist, bleibt eine lokalisierte Version entscheidend für das Spielerlebnis und den kommerziellen Erfolg. Dabei wird in der Lokalisierung ein stark funktionalistischer Ansatz verfolgt, der nach manchen übersetzungswissenschaftlichen Theorien bereits den Rahmen dessen sprengt, was eine Übersetzung „darf“:

> […] translators are often given carte blanche to modify, adapt, and remove any cultural references, puns, as well as jokes that would not work in the target language. Localisers are given the liberty of including new cultural references, jokes, or any other element they deem necessary to preserve the game experience and to produce a fresh and engaging translation. This type of creative licence granted to game localisers would be the exception rather than the rule in any other types of translation. [3]

In diesem Artikel werden die terminologischen Herausforderungen der Videospiellokalisierung im Sprachpaar Englisch-Deutsch näher beleuchtet.

Die terminologischen Dimensionen des Videospiels

Ein Videospiel bewegt sich stets auf einem Spektrum, an dessen Enden sich einerseits die Fiktion mit all ihren Ausformungen und andererseits die Realität mit all ihren Fachgebieten befinden. Die Lokalisierung von Videospielen erfordert zunächst ein solides multilinguales terminologisches Fundament, das sich vor allem aus technischen Bezeichnungen, Gaming- und Gamer-Sprache sowie bestimmten genrespezifischen Termini zusammensetzt. Der Dynamik des Spektrums folgend, lassen sich allerdings nicht alle Bezeichnungen aus diesem festen Repertoire schöpfen: Die Interdisziplinarität und thematische Offenheit von Videospielen machen es erforderlich, dass zahlreiche Termini für jedes Spiel individuell erarbeitet oder sogar neu geschaffen werden. Übersetzer:innen agieren folglich als Fachgebiets- und Stilchamäleons, die über ausgeprägte Recherche- und Adaptionsfähigkeiten verfügen müssen. Die terminologischen Dimensionen lassen sich

grob in fünf Kategorien unterteilen, wobei die Grenzen fließend verlaufen: Fachgebiete der realen Welt; Gaming- und Gamer-Sprache; Hard- und Software-Terminologie; kreative Elemente und zuletzt firmenspezifische Terminologie und Vorgaben.

Fachgebiete der realen Welt – von Aerodynamik bis Zytologie

Fachgebiete aus der realen Welt finden in Spielen auf vielfältige Weise Anwendung – sei es durch Gamification-Ansätze in der Aus- und Weiterbildung, wie sie etwa bei der Bundeswehr genutzt werden [4], oder durch Simulationen, bei denen größtmöglicher Realismus angestrebt wird. Ein herausragendes Beispiel für Letzteres ist der *Microsoft Flight Simulator*, dessen Lokalisierung fundierte Kenntnisse in Bereichen wie Aerodynamik, Navigation, Funkverkehr und Flugzeugmechanik erfordert. Ähnlich anspruchsvoll sind Wirtschafts- und Managementsimulationen wie *Cities: Skylines*, die Fachterminologie aus den Bereichen Ökonomie, Stadtplanung, Finanzwesen und Marktmechanismen beinhalten. Der *Landwirtschafts-Simulator* strebt hingegen die detailgetreue Nachbildung agrarwirtschaftlicher Prozesse an, wobei Fachwissen über Bodenbewirtschaftung, Saatgut, Erntemaschinen und Viehzucht essenziell ist.

Ohne präzise Terminologie wäre die Immersion für fachkundige Spieler:innen erheblich beeinträchtigt – die Lokalisierung erfordert also wie jede Fachübersetzung nicht nur sprachliches Können, sondern auch tiefgehendes Wissen in den jeweiligen Fachgebieten:

> [The] unauthentic use of terminology in specialized games, such as flight simulation ones, is likely to break the suspension of disbelief for any user who is familiar with the domain […]. From this perspective, translating games can be similar to translating technical texts featuring subject-specific terminology. The same would apply, for instance, in games revolving around sports, music, cooking or military combat scenarios. [5]

Gaming- und Gamer-Sprache – von *AoE* bis *Zerg Rush*

Die Gaming-Sprache umfasst Fachausdrücke, die sich auf Spielmechaniken, Genres und Entwicklung beziehen – etwa *Hitbox*, *Cooldown* oder *Skill Tree*. Sie wird von Entwicklungsstudios, Publishern und Fachleuten genutzt, findet aber auch in offiziellen Spieltexten Anwendung. Als Beispiel seien die sogenannten Steam-Tags genannt. Steam ist eine digitale Vertriebsplattform für Videospiele, Software und andere Inhalte und dient darüber hinaus als Community Hub mit einer Vielzahl sozialer und interaktiver Funktionen. Die Tags sind Schlagwörter, die Spiele anhand von Genre, Mechanik oder Atmosphäre kategorisieren. Eine präzise und konsistente Terminologie sowohl in der Ausgangssprache als auch der Übersetzung machen hier also den entscheidenden Unterschied, da Fehler dazu führen können, dass Spiele nicht korrekt gefiltert und potenzielle Spieler:innen nicht gezielt angesprochen werden.

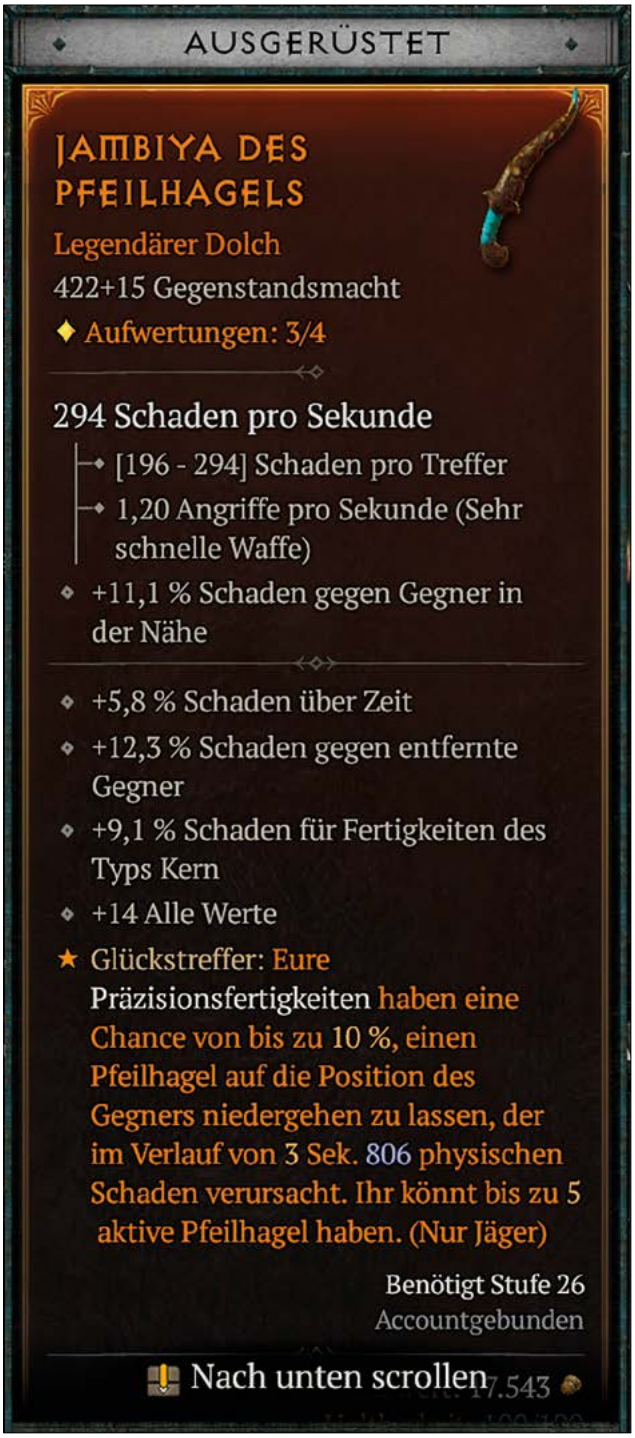


Abb. 1: Waffe mit Werten und Beschreibung in Diablo IV *(Blizzard Entertainment 2023).*

Zur Gaming-Sprache zählen zudem etablierte spielübergreifende Bezeichnungen wie Bewegungs- und Angriffsgeschwindigkeit, Seltenheitsstufen von Gegenständen oder Gegenstandstyp. Abbildung 1 zeigt eine Waffe aus *Diablo IV*

mit typischer Fachterminologie. Übersetzer:innen müssen hier genau abwägen: Während viele englische Bezeichnungen auch dem deutschen Zielpublikum geläufig sind (*Team Deathmatch*, *Quest*), kann es je nach Zielgruppe sinnvoll sein, ein deutsches Äquivalent zu wählen – etwa für sehr junge Spieler:innen ohne ausreichende Englischkenntnisse.

Die Gamer-Sprache hingegen entwickelt sich innerhalb der Spielgemeinschaft und umfasst Slang, Abkürzungen, Referenzen und Meme-Kultur (*GG*, *Noob*, *Campen*, *Nerf*, *Zerg Rush*, *Meta*, *Stream Sniping* etc.). Da der Austausch meist online und auf Englisch stattfindet, dominieren auch hier englische Bezeichnungen. Die Grenzen zur Gaming-Sprache sind dabei wegen der wechselseitigen Beziehung zu Entwicklungsstudios und Publishern fließend.

Übersetzer:innen müssen sich kontinuierlich auf dem neuesten Stand halten – veraltetes Wissen kann dazu führen, dass Referenzen oder Mechaniken nicht verstanden und falsch übersetzt werden.

Hard- und Software-Terminologie – von *Anti-Aliasing* bis *Zero Latency*

Die „klassische" technische Sprache, die bei Videospielen Anwendung findet, speist sich aus der Entwicklung und Programmierung von Hard- und Software. Ein zentraler Aspekt dieser technischen Sprache ist ihre Präzision. Einstellungen, Optionen und technische Spezifikationen müssen nicht nur fachlich korrekt, sondern auch eindeutig und konsistent übersetzt werden. Gerade im Bereich der Benutzeroberfläche und der Spieleinstellungen (siehe Abb. 2) muss Missverständnissen vorgebeugt werden. Zusätzlich erschwert wird die Lokalisierung durch die enge Verzahnung mit der Programmierung. Viele technische Bezeichnungen stammen direkt aus der Programmierung und sind nicht nur schwer ins Deutsche zu übertragen, sondern werden von den Spieler:innen ohnehin in ihrer englischen Form erwartet. Hier kann die bewusste Mischung aus deutschen und englischen Bezeichnungen die bessere Wahl darstellen.

Die technische Sprache in Videospielen bewegt sich mithin an der Schnittstelle zwischen Präzision und Nutzerfreundlichkeit. Sie muss verständlich, konsistent und fachlich korrekt sein und dabei stets die Erwartungen der Spieler:innen berücksichtigen.

Kreative Elemente – von Ausrufen bis Zweideutigkeiten

Die kreative Lokalisierung reicht von Witzen, Referenzen und Charakterstilen bis zu Neologismen und ähnelt somit der Übersetzung von Literatur:

> Some […] games are narrative-driven and may require skills akin to those of a literary translator; some dialogues, for example, might have to be translated in a specific style and register, whether contemporary or archaic. [5]

Auch hier sind die Übergänge zu anderen Fachgebieten fließend: In *Resident Evil* basiert das erfundene T-Virus auf medizinis0cher Fachsprache, *Papers, Please* hingegen spielt im fiktiven totalitären Staat Arstotzka und nimmt politische Fachthemen auf. Ein Beispiel für die kreative

Abb. 2: Technische Sprache im Einstellungsmenü des Spiels Critter Cove *(Gentleman Rat Studio 2024).*

Abb. 3: Charakterausformung mithilfe von Dialekt in Critter Cove *(Gentleman Rat Studio 2024).*

20.	Avast·me·hearties, <d=5> ·if·it·ain't· {player.Name} ·come·to·parley·with·ole·Anchor. <d=10> ·Out·with·it, <d=5> ·what·can·Anchor·do·fer·ye?	Kiekt·ma·her,·Büttjes·und·Deerns! <d=5> ·Wenn·dat·nich· {player.Name} ·is.·Willste·mit·dem·alten·Anker-Hajo·schnacken?· <d=10> Nu·mal·raus·damit, <d=5> ·wat·kann·Anker-Hajo·für·dich·tun?	101%

Abb. 4: Dialog-String zu Screenshot in Abb. 3 in MemoQ.

Abb. 5: Roboterhafter Sprachduktus zur Charakterausformung in Critter Cove *(Gentleman Rat Studio 2024).*

162.	<c=green> <a=center> Social·Mode·Activated </c> <d=30>	<c=green> <a=center> Protokoll·aktiviert:·„Soziale·Interaktion“ </c> <d=30>	101%
163.	This·unit·would·remind· {player.Name} ·that·if·they·are·looking·for·useful·crafting·materials, <d=10> ·you·can·sometimes·find·them·for·sale·at· {renard.Name} 's.	Diese·Einheit·möchte· {player.Name} ·daran·erinnern, <d=10> ·dass·nützliche·Herstellungsmaterialien·gelegentlich·bei· {renard.Name} ·erworben·werden·können.	101%

Abb. 6: Dialog-String zu Screenshot in Abb. 5 in MemoQ.

3433	Hey, what did the ocean say to the beach?	Hey, weißt du, was Fische überhaupt nicht mögen?
3434	What?	Nein, was?
3435	Nothing, it just waved.	Anti-Schuppen-Shampoo!

Abb. 7: Beispiel einer Transkreation eines Wortwitzes.

Charakterausformung zeigen die NPC (Non-Playing Character) in *Critter Cove* (Gentleman Rat Studio 2024). Der Piratenhai *Anchor* wurde im Deutschen als Anker-Hajo mit norddeutschem Dialekt adaptiert (*kieken*, *Büttje*, *Deern*), um Lokalkolorit zu erzeugen (Abbildung 3 & 4 auf Seite 8) – hier wurde also transkreativ gearbeitet. Die Lokalisierung des Roboters *Copperbottom* (deutscher Name: Kupferkiste) behält den präzisen, formellen Stil des Originals bei (Abbildung 5 & 6 auf Seite 8), was allerdings ein völlig anderes terminologisches und stilistisches Register erfordert.

Auch Wortspiele müssen angepasst werden: Ein englischer Witz basierend auf der Mehrdeutigkeit des Wortes *wave* lässt sich nicht wörtlich ins Deutsche übertragen, sondern fordert übersetzerische Kreativität, um einen Wortwitz zu entwickeln, der sowohl thematisch als auch in der Wirkung dem Original gerecht wird.

Die aufgeführten Beispiele zeigen, dass die eingangs erwähnte Carte Blanche bei kreativen Elementen voll zum Tragen kommt. Trotz aller kreativen Freiheit hat jedoch auch hier eine konsistente Terminologie oberste Priorität. Besonders problematisch sind sogenannte „Plotstopper" – Situationen, in denen Spieler:innen aufgrund widersprüchlicher Bezeichnungen im Spielfluss gehemmt und somit die Immersion abrupt unterbrochen wird. Stellen Sie sich vor, Sie werden als Spieler:in über einen Missionstext angewiesen, das *grüne Amulett* aus den *Katakomben des Verderbens* zu bergen. Erscheinen diese Orte und Gegenstände jedoch im Spielmenü oder auf der Karte unter abweichenden Bezeichnungen – etwa als *Katakomben des Unheils* und *Smaragdamulett* – kann dies zu Verwirrung und Frustration führen. In solchen Fällen ist die terminologische Konsistenz nicht nur eine stilistische Frage, sondern Voraussetzung für die Spielbarkeit und die Aufrechterhaltung der Immersion.

Firmenspezifische Terminologie und Vorgaben – von DualSense bis Xbox-Taste

Im Gegensatz zu kreativen Elementen unterliegt firmenspezifische Terminologie stets strikten Regeln. Besonders Konsolenhersteller wie Microsoft (Xbox) und Sony (PlayStation) legen großen Wert auf Corporate Language und die konsistente Bezeichnung ihrer Produkte. So wird eine einheitliche Markenkommunikation über Spiele, Plattformen und Marketingmaterial hinweg gewährleistet. Bezeichnungen wie DualSense (ein PlayStation-Controller) oder Xbox Game Pass (ein Abonnement für die Xbox) müssen in allen Sprachen der offiziellen zielsprachlichen Terminologie folgen. Hier gibt es keine Carte Blanche – Verstöße können gravierende Folgen haben, dem Image der Kund:innen schaden und negativ auf Übersetzer:innen zurückfallen.

Charakteristische Probleme und Herausforderungen des Mediums

In den folgenden Abschnitten werden die verschiedenen terminologischen Probleme behandelt, die das Videospiel aufgrund seiner medienbedingten Charakteristik mit sich bringt. Die Probleme lassen sich grob in Kontextmangel, Variablen, Platzbegrenzungen und die Arbeit im Team unterteilen.

Kontextmangel durch Textfragmentierung und Sim-Ship

Um die Interaktivität eines Videospiels zu ermöglichen, muss die Narration nichtlinear strukturiert sein. Während ein Roman in der Regel linear aufgebaut ist – die Leser:innen beginnen auf der ersten Seite bei Punkt A und durchlaufen die Geschichte bis zur letzten Seite zu Punkt Z – eröffnen Videospiele eine weitaus größere Freiheit: Hier ist denkbar, dass Spieler:innen in der Geschichte von Punkt A zu Punkt P springen können, anschließend zu Punkt B zurückkehren und unterwegs einen Zwischenstopp bei Punkt H einlegen:

> The interactive element of computer programs has serious consequences for translators because it means that access to text and information is random, i.e., each user will activate a particular message or command at a different point, or not at all. An arbitrary sequence of events does not allow for linear texts […]. [6]

Diese nichtlineare Struktur erfordert eine fragmentierte Textgestaltung, da die einzelnen Textsegmente – sogenannte Strings – je nach Aktion dynamisch vom Spielcode aufgerufen werden. Die Gesamtheit der Strings wird zur Lokalisierung des Spiels in eine Tabelle exportiert, welche als Arbeitsgrundlage für die Lokalisierung dient. Übersetzer:innen stehen hierbei vor der Herausforderung, mit Hunderten, Tausenden oder noch mehr isolierten Strings zu arbeiten, die meist ohne direkten Kontext übersetzt werden müssen. Im Idealfall sind diese mit IDs oder Kommentaren versehen, die eine Orientierung ermöglichen – eine WYSIWYG-Vorschau (*What You See Is What You Get*) fehlt jedoch in den allermeisten Fällen (vgl. [5]).

Die Fragmentierung der Spieltexte führt in vielen Fällen dazu, dass Übersetzer:innen zwar die ausgangssprachliche Bezeichnung vorliegt, die begriffliche Ebene jedoch völlig fehlt. Ein besonders anschauliches Beispiel hierfür ist das englische Polysem *unlock*. In Videospielen wird das Verb häufig im Zusammenhang mit Erfolgen oder Inhalten verwendet – in diesem Kontext ist das Verb *freischalten* die einzig korrekte Übersetzung. Fehlt jedoch jeglicher Kontext, bleibt unklar, ob der String möglicherweise in einer völlig anderen Spielsituation erscheint, etwa wenn sich der Spielcharakter einer verschlossenen Tür nähert. In diesem Fall wäre *freischalten* eine Fehlübersetzung und *entriegeln* oder *aufschließen* korrekt. Selbst wenn unterstützende Materialien wie Bilder, Videos oder Zusammenfassungen bereitgestellt werden, können sie nicht alle möglichen Kontextfragen restlos klären. Die Interaktivität eines Spiels erschließt sich erst im tatsächlichen Zusammenspiel seiner Elemente – keine noch so umfangreiche Referenzsammlung könnte sämtliche Variationen und Alternativen aller denkbaren Interaktionen abbilden.

Eine weitere Dimension des Kontextproblems rührt vom sogenannten Sim-Ship-Verfahren her. Sim-Ship (kurz für *Simultaneous Shipment*) bezeichnet die gleichzeitige Veröffentlichung eines Videospiels im Original und den lokalisierten Versionen. Da mittlerweile tagtäglich neue Spiele auf dem Markt erscheinen, verfügt jedes einzelne Spiel über eine geringe „Halbwertszeit“, bevor es seinen Platz im Rampenlicht aufgeben muss. Sim-Ship soll den sich daraus ergebenden negativen Effekten entgegenwirken:

> The Sim-Ship model of distribution capitalises on the momentum built up by a single international marketing campaign, and minimises the risk of grey imports and piracy. [7]

Bei Sim-Ship-Spielen haben Übersetzer:innen während ihrer Arbeit keinen Zugriff auf die finale Spielversion. Fehlübersetzungen sind die Folge: So kann eine ausgangssprachliche *metal bar* wie in Abbildung 8 auf Seite 11 fälschlicherweise als *Metallstange* statt als *Metallbarren* übersetzt werden, weil die visuelle Referenz zum Zeitpunkt der Übersetzung nicht verfügbar war. Darüber hinaus droht hier die Gefahr, dass nachträgliche Änderungen, zu denen es während der Entwicklungsphase eines Spiels häufig kommt, nicht konsistent über alle Strings hinweg umgesetzt werden.

Der kontextuelle „Blindflug“ stellt in der Videospiellokalisierung eher die Regel als die Ausnahme dar. Zwar

Anzeige

Deutscher Terminologie-Tag e.V.
Deutsches Institut für Terminologie e.V.

© pixelio.de

DTT-Fortbildung

Aufbau- und Vertiefungsmodule
im Oktober und November 2025, online

Webinar: „Terminologierecherche: Methoden und hilfreiche Tools“
am 15. Oktober 2025, 13:00 bis 14:00 Uhr, online

Weitere Informationen und Anmeldung unter: dttev.org/fortbildung

Stand: 1. Juni 2025 | Änderungen vorbehalten.

String aus Perspektive der Spieler:innen	
Sprich mit Kupferkiste	
String im englischen Spielcode:	String im lokalisierten Spielcode:
Talk to {tracker.giver.name}	Sprich mit {tracker.giver.name}

Tab. 1: Spielanweisung mit Variable (Beispielstring aus Abbildung 8)

Abb. 8: Fehlerhafte Übersetzung aufgrund von Kontextmangel in Critter Cove *(Gentleman Rat Studio 2024).*

kann bereits erworbene Erfahrung dabei helfen, sinnvolle Übersetzungsentscheidungen auch ohne weiteren Kontext zu treffen und sogar potenzielle Fehler im Ausgangstext zu identifizieren, doch birgt sie zugleich die Gefahr voreiliger Schlussfolgerungen. Es besteht zwar die Möglichkeit, Rücksprache mit den Entwicklungsstudios zu halten und Kontext zu erfragen, was aber für beide Seiten mühselig und zeitintensiv ist. Auch in dieser Hinsicht sind also umfassende Referenzmaterialien oder idealerweise direkter Zugang zum Spiel zu bevorzugen.

Variablen

Variablen ermöglichen es, Texte dynamisch an die Spielsituation anzupassen, ohne für jede Variante einen eigenen String anlegen zu müssen. Ein einfaches Beispiel ist eine der Spielanweisungen in Abbildung 8 (siehe Tabelle 1).

Variablen spielen für die Interaktivität und die Illusion eines maßgeschneiderten Erlebnisses eine ebenso zentrale Rolle wie die Fragmentierung des Spieltextes in einzelne Strings: „The use of variables stems from the need to create individualized game play experiences for millions of players“ [8]. Allerdings bringt ihr Einsatz unter anderem im Sprachpaar Englisch-Deutsch erhebliche grammatikalische Herausforderungen in der Lokalisierung mit sich. Das Englische verfügt bezogen auf Substantive über kein Genus und drückt grammatikalische Beziehungen primär durch Wortstellung aus, das Deutsche hingegen wartet mit gleich drei Genera auf und bedient sich Flexionen. In Bezug auf die Variablen bedeutet dies: Während das Einfügen von Textelementen im Englischen meist problemlos funktioniert, da der Variablenwert nicht flektiert wird, kann die Übertragung ins Deutsche zu unlösbaren Problemen führen, da die korrekte Flexion von mehreren Faktoren wie Kasus und Genus abhängt. Klischewski bezeichnet dieses Dilemma als den „Price of Infinite Choice“ – die Variablen bieten zwar maximale Flexibilität, erzeugen aber gleichzeitig strukturelle Sackgassen:

> As tempting and essential as the introduction of variables is, they often lead to loose ends that turn out to be impossible to solve in one or even several languages. [8]

Platzbegrenzungen

Während ein übersetzter Roman problemlos einige Seiten länger sein kann, ohne an Qualität einzubüßen, unterliegt die Lokalisierung von Videospielen strengen Platzvorgaben. Besonders Benutzeroberflächen (User Interface, UI) und HUD-Elemente (Head-up Display) setzen wortwörtliche

Scripten mit Python und ChatGPT für Übersetzer und Terminologen

10.–12.09.2025, 9:30–13:00 Uhr
Fortbildungsformat: Training
www.bdue-fachverlag.de/detail_training/6586

Das Aufkommen von „Large Language Models“ wie ChatGPT im Jahr 2022 ist eine logische Konsequenz und Weiterentwicklung der Neuronalen Maschinellen Übersetzung, die bereits seit 2017 Einzug in den Übersetzeralltag gehalten hat. Statt mit bilingualen Datenbeständen, meist TMX-Dateien, werden diese „großen Sprachmodelle“ mit Unmengen von Daten aller Art trainiert, um Aufgaben verschiedenster Art lösen zu können.

Terminologie und Künstliche Intelligenz im Texterstellungs- und Übersetzungsprozess

23.09.2025, 10:30–12:00 Uhr
Fortbildungsformat: Webinar
www.bdue-fachverlag.de/detail_webinar/6580

Künstliche Intelligenz ist seit ChatGPT in Alltag und Beruf angekommen. Übersetzerinnen und Übersetzer kennen KI schon länger - durch die Nachbearbeitung maschinell erstellter Texte. Allerdings kommt gerade die Terminologie in diesen Systemen zu kurz. Daraus ergeben sich neue Anforderungen, aber vor allem viele Perspektiven für den Übersetzerberuf.

Terminologiearbeit in Zeiten von KI und MÜ für Anfänger – neue Aufgaben für Übersetzer

22.10.2025, 9:30–13:00 Uhr
Fortbildungsformat: Training
www.bdue-fachverlag.de/detail_training/6593

Durch den Einzug der Maschinellen Übersetzung – ob neuronal oder mit Hilfe von Large Language Models – verschiebt sich der Aufgabenbereich der Übersetzer von der Übersetzung von Texten zum Qualitätsmanagement von vorübersetzten Texten und vor allem zur mehrsprachigen Terminologiearbeit bereits vor der Übersetzung. Denn wer weiß mehr über Begriffe und ihre Benennungen als die Fachübersetzer?

Terminologiearbeit in Zeiten von KI und MÜ für Fortgeschrittene – Konzeption von Terminologiedatenbanken

29.10.2025, 9:30–13:00 Uhr
Fortbildungsformat: Training
www.bdue-fachverlag.de/detail_training/6594

Die Anforderungen an Terminologieverwaltungs- und abstimmungssystemen im Unternehmen steigen. Die Termbanken müssen für unbedarfte Nutzer aller Abteilungen und weltweit leicht zu bedienen und selbsterklärend sein, gleichzeitig sollen sie auch allen Anforderungen an Schnittstellen mit und ohne KI genügen. Deshalb werden die Anforderungen an die Konzeption von Termbanken immer komplexer.

Terminologiearbeit in Zeiten von KI und MÜ für Fortgeschrittene – Austausch von terminologischen Daten

30.10.2025, 9:30–13:00 Uhr
Fortbildungsformat: Training
www.bdue-fachverlag.de/detail_training/6595

Nur wenige Systeme im Unternehmen sind so komplex wie ein Terminologieverwaltungssystem, das allen Anforderungen an das Wissensmanagement im Unternehmen, das Wissen über Begriffe, ihre Definitionen und ihre Benennungen, gerecht werden muss.

Terminologiearbeit in Zeiten von KI und MÜ für Fortgeschrittene – Begriffssysteme erarbeiten

31.10.2025, 9:30–13:00 Uhr
Fortbildungsformat: Training
www.bdue-fachverlag.de/detail_training/6596

Concept Maps sind für die Wiederauffindbarkeit und die Verschlagwortung von Inhalten, z.B. für die Industrie 4.0, unverzichtbar. Sie kommen daher meist in CCMS Content Management Systemen zum Einsatz. Weiterhin gibt es der IT-Abteilung eines Unternehmns Mitarbeiter, die sich mit der Erstellung von Taxonomien beschäftigen, mit denen das Wissen einer Organisation geordnet werden kann.

Dozentin

Prof. Dr. Rachel Herwartz hat eine Professur für Übersetzungs- und Terminologiemanagement an der Internationalen Hochschule SDI München inne und leitet den Master-Studiengang „Translation Management“. Als Gründerin und Inhaberin TermSolutions GmbH ist sie Expertin für Termbankmodellierung und die Einführung und Optimierung von Terminologieprozessen in Unternehmen. Seit 20 Jahren coacht und konzipiert sie Systeme für die Terminologieverwaltung, -abstimmung und -prüfung und ist als Product Owner für die Weiterentwicklung der Systeme termXplorer, myTerm und termXact verantwortlich. Ihre Forschungsschwerpunkte liegen im Bereich AI, Terminologie und Maschinelle Übersetzung. Als Mitglied im Deutschen Institut für Normung (DIN) arbeitet sie aktuell an der neuen Norm für die mehrsprachige Terminologiearbeit mit.

Weiterbildung für Dolmetscher und Übersetzer
Immer auf dem neuesten Stand – technisch, fachlich, sprachlich, unternehmerisch:
www.bdue-fachverlag.de

Grenzen – ein zu langer Text kann über den Bildschirm- oder Fensterrand hinausgehen oder ineinanderlaufen. Um diesen Einschränkungen zu begegnen, setzen Entwicklungsstudios auf verschiedene Lösungen wie Icons, Piktogramme, Tooltipps und Abkürzungen. Jede dieser Lösungen bringt jedoch eigene Herausforderungen in der Lokalisierung mit sich.

Icons und Piktogramme sind zwar platzsparend, aber keineswegs universell verständlich. Eine Baseball-Homebase als Symbol für das Hauptmenü mag im US-amerikanischen Raum sofort verständlich sein, während sich deutsche Spieler:innen mit einem stilisierten Haus als Symbol weitaus leichter tun. Tooltipps können in vielen Kontexten eine elegante Lösung sein, sind jedoch nicht überall praktikabel. Besonders im HUD und in Spielsituationen, die schnelle Reaktionen erfordern, bleibt keine Zeit für ausführliches Lesen. Abkürzungen erscheinen als naheliegender Kompromiss, doch ihr exzessiver Gebrauch kann schnell zu einem unverständlichen Buchstabensalat führen. Ein anschauliches Beispiel bietet *The Elder Scrolls IV: Oblivion*, wo Tranknamen im Händlermenü drastisch gekürzt wurden. So wird der *Schwache Trank der Nachtsicht* als *Schw. Tr. d. Nachtsi.* dargestellt – eine kryptische Zeichenfolge, die Spieler:innen erst mühsam entschlüsseln müssen (siehe Abbildung 9).

Zuletzt wird die Lokalisierung erheblich durch die Längendifferenz zwischen englischen und deutschen Texten erschwert. Da englische Texte in der Regel deutlich kürzer als ihre deutschsprachigen Übersetzungen sind, verschärft sich das Platzproblem in der Lokalisierung noch weiter – insbesondere, wenn die Benutzeroberfläche bereits für die kompaktere englische Version optimiert wurde. In solchen Fällen müssen Übersetzer:innen gezielt abwägen: Ist eine weniger präzise, aber kürzere Formulierung vertretbar? Lässt sich die Bezeichnung sinnvoll abändern, ohne die Lesbarkeit zu beeinträchtigen? Die Herausforderung besteht darin, eine Balance zwischen Verständlichkeit, Konsistenz und Platzersparnis zu finden.

Arbeit im Team

Sim-Ship und das schiere Textvolumen vieler moderner Videospiele machen es unmöglich, dass ein gesamtes Spiel von einer einzelnen Person lokalisiert wird. Gerade bei großen Produktionen arbeiten oft Dutzende Übersetzer:innen und Lektor:innen pro Sprache parallel an der Lokalisierung. Diese kollaborative Arbeitsweise stellt allerdings eine Gefahr für Kohärenz und Konsistenz dar.

Erschwerend kommt hinzu, dass nicht alle Teammitglieder denselben Zugang zu Kontextinformationen haben.

Abb. 9: Exzessiver Gebrauch von Abkürzungen in The Elder Scrolls IV: Oblivion *(Bethesda 2006).*

Wer beispielsweise die Anfangssequenzen eines Spiels übersetzt, verfügt oft über einen besseren Gesamtüberblick als jemand, der isolierte Menütexte oder spätere Dialoge bearbeitet. Dies kann dazu führen, dass verschiedene Übersetzer:innen für ein und dieselbe englische Bezeichnung unterschiedliche deutsche Entsprechungen wählen – eine der häufigsten Ursachen für terminologische Inkonsistenz. Zur Lösung dieses Problems kommen in der Lokalisierung standardmäßig CAT-Tools mit eingebauten Terminologiedatenbanken zum Einsatz. Die konsequente Pflege und Nutzung dieser Ressource erfordert indes zusätzliche Abstimmung und kostet wertvolle Zeit. Ein weiteres Hindernis ist die oft große kommunikative Distanz zwischen den Beteiligten. Lokalisierungsteams sind in der Regel weltweit verteilt und arbeiten in unterschiedlichen Zeitzonen, sodass Rückfragen nicht unmittelbar geklärt werden können. Verzögerungen in der Kommunikation können zu Inkonsistenzen oder Missverständnissen führen, die sich erst in späteren Qualitätssicherungsschritten bemerkbar machen.

Die Teamarbeit in der Videospiellokalisierung erfordert daher einen Kompromiss aus Effizienz und Qualität. Klare Prozesse, gut durchdachte terminologische Richtlinien und eine effektive Kommunikation sind essenziell, um trotz der Herausforderungen eine hochwertige Lokalisierung sicherzustellen.

Fazit

Das Videospiel ist die sprich- und wortwörtliche Spielwiese der realen Welt und der Fantasie mit allen fachsprachlichen und terminologischen Implikationen. Die Tätigkeit der Videospiellokalisierung ist daher ein ständiges Wechselspiel zwischen kreativer Freiheit und terminologischer Präzision, zwischen sprachlicher Feinarbeit und technischen Einschränkungen. Die Grenzen zwischen Fachübersetzung und künstlerischer Adaption verschwimmen dabei immer wieder. Zusammen mit den medienbedingten Herausforderungen in Form von Kontextmangel, Textfragmentierung, Platzbegrenzungen und Variablen entsteht ein hochkomplexes Spannungsfeld, das Übersetzer:innen vor einzigartige Herausforderungen stellt – und den besonderen Reiz der Videospiellokalisierung ausmacht.

Literatur

[1] Bernal Merino, Miguel Ángel (2015): Translation and Localisation in Video Games. Making Entertainment Software Global. New York/London: Routledge.

[2] Bernal Merino, Miguel Ángel (2006): On the Translation of Video Games. In: The Journal of Specialised Translation (6), S. 22–36.

[3] Mangiron, Carme; O'Hagan, Minako (2006): Game Localisation: Unleashing Imagination with 'Restricted' Translation. In: The Journal of Specialised Translation (6), S. 10–21.

[4] Cyber Innovation Hub. E-Learning dank Gamification & Augmented Reality. Verfügbar unter https://www.cyberinnovationhub.de/innovation/innovationsvorhaben/e-learning-bundeswehr. Letzter Zugriff am 18.03.2025.

[5] O'Hagan, Minako (2019): Game Localization. A critical overview and implications for audiovisual translation. In: Pérez González, Luis (Hrsg.): The Routledge Handbook of Audiovisual Translation. Oxon/New York: Routledge, S. 145–159.

[6] Bernal Merino, Miguel Ángel (2007): What's in a 'Game'? In: Localisation Focus 6(1), S. 29–38.

[7] Bernal Merino, Miguel Ángel (2011): A Brief History of Game Localisation. In: TRANS. Revista de Traductología 15, S. 11–17.

[8] Klischewski, Rolf (2016): The Price of Infinite Choice. In: MultiLingual 27(4), S. 42–47.

Annika Sophie Dürr ist freiberufliche Übersetzerin (M.A.) und Dozentin mit Spezialisierung auf Videospiellokalisierung und Transkreation.

Kontaktadresse
info@leveluplocalizations.com
www.leveluplocalizations.com

Anzeige

Registration for the EAFT Terminology Summit in Innsbruck is now open.
For more information on how to register and for the complete program, please visit:
https://www.uibk.ac.at/en/congress/eaft-summit/registration/

Halluzination – eine irreführende Bezeichnung für irreführenden Output von Large Language Models

Annette Gerstenberg

The term hallucination is traced from its origins to its metaphorical use in everyday language and scientific discourse. Its recent rise to describe unsubstantiated output in texts generated by artificial intelligence is commented on as misleading and potentially stigmatizing, marking an early terminological stage in an emerging field.

Keywords: hallucination, artificial intelligence, semantic change, metaphor, Large Language Model, LLM

„Claude can make mistakes. Please double-check responses." Das sagt Claude, der KI-Assistent der Firma Anthropic. Er basiert ähnlich wie ChatGPT auf einem Large Language Model (LLM), Anthropic wirbt jedoch mit seiner besonderen Zuverlässigkeit. Disclaimer mit dem Hinweis auf mögliche Fehler sind trotzdem notwendig, weil Falschinformationen in der Interaktion mit einem KI-Chatbot oder -Assistenten jederzeit möglich sind. Darin liegt das Risiko der Benutzung solcher Tools. Substanzlose Textpassagen mit irreführendem Inhalt sind ohne Hintergrundwissen nicht von Textpassagen zu unterscheiden, die gut begründet und zuverlässig sind.

Für solche Fehler ist die Bezeichnung *Hallucination* üblich geworden. Zuerst kam die Bezeichnung Anfang des Jahrtausends im Bereich der Bildbearbeitung in Gebrauch [1], inzwischen ist *Hallucinations* für textuelle Fehlleistungen üblich geworden und das Verständnis ist differenzierter geworden [2]. Unterschieden wird die intrinsische Halluzination, die auf Basis des Inputs fehlerhafte Ergebnisse ausgibt, von der extrinsischen Halluzination, bei der gänzlich ohne Bezug auf den Input sachlich falsch getextet wird [3].

Die Bezeichnung beschränkt sich nicht auf die Fachsprache, sondern erstreckt sich auch auf die Allgemeinsprache: Im Jahr 2023 wurde das Verb *hallucinate* zum Word of the Year des Cambridge Dictionary gewählt. Die deutsche Entsprechung *Halluzination* steht bereits mit folgender Definition im Duden: „durch KI erzeugte, nicht auf Fakten oder realen Daten basierende falsche, jedoch glaubhaft erscheinende Information".

Halluzination ist kein Neuzugang in der Fachsprache. Als Bezeichnung eines medizinischen Symptoms wird es wissenschaftlich definiert als „Form der Sinnestäuschung mit Sinneswahrnehmung ohne (externe) gegenständliche Reizquelle, bei der die Betroffenen von der Realität des Wahrgenommenen überzeugt sind. Halluzinationen treten beispielsweise auf bei Schizophrenie oder beim Konsum von Halluzinogenen" (Pschyrembel 2025). Auch diese medizinische Bedeutung fand Eingang in die Allgemeinsprache und in den Duden.

Wie kommt es dazu, dass ein Krankheitssymptom, das psychiatrisch behandelt wird, in die technische Sprache der KI gebracht wird? Aus der facettenreichen Geschichte von *Halluzination* ergeben sich einige Anknüpfungspunkte.

Das griechische Verb *ἀλύω* ‚aufgeregt sein, unruhig sein, herumirren' wurde über das lateinische Substantiv ALUCINĀTIO bzw. das Verb (H)ALUCINĀRI in die modernen europäischen Sprachen aufgenommen. Zunächst wurde es in literarischen Texten wörtlich und übertragen als ‚herumirren, verblendet sein' benutzt. Aus dieser übertragenen Bedeutung ‚Verblendung' entstand die fachsprachliche Bezeichnung visueller Probleme, die seit der Frühen Neuzeit belegt ist. So bezeichnet das französische Wort *hallucination* 1655 bei Fernel [4] eine Störung des Sehvermögens. Anfang des 19. Jahrhunderts konkurrieren dann schon unterschiedliche fachsprachliche Verwendungen, die in einem französischen Fachwörterbuch von 1817 als Störung der Perzeption zusammengefasst werden. Für das Deutsche definiert Müller 1826 in seiner Darstellung „Über die phantastischen Gesichtserscheinungen" das Phänomen etwas breiter und im Sinne perzeptorischer Fehlwahrnehmungen: „Dem Arzte sind sie längst bekannt als sogenannte Hallucinationen in Fiebern, in Krankheiten des

Deutscher Terminologie-Tag e.V.

Deutsches Institut für Terminologie e.V.

Terminology Work Best Practices 2.0

Das bewährte Praxishandbuch für Terminologiearbeit und Terminologiemanagement mit dem Know-how zahlreicher Experten aus Industrie und Wissenschaft ist auch **in englischer Sprache** erhältlich.

kompakt | praxisnah | auf den Punkt

Geballtes Terminologiewissen für Unternehmen und Freiberufler

1. Argumentationshilfen
2. Grundsätze und Methoden
3. Benennungen
4. Werkzeuge und Technologien
5. Projekt- und Prozessmanagement
6. Berufsprofile, Anforderungen, Ausbildungsinhalte
7. Urheberrecht an Terminologie
8. Wirtschaftlichkeit

Erhältlich in Deutsch oder Englisch auf www.dttev.org

Gehirns, in der Hysterie, Hypochondrie, Catalepsie, bei Irren, in der Ekstase und verwandten Zuständen" [5]. Der Autor war mit der französischen Verwendung vertraut, an anderer Stelle [5] zitiert er den Eintrag *Hallucinations* des französischen Fachwörterbuchs *Dictionaire des sciences médicales* (DSM, 1817). Auch die englische Fachsprache orientierte sich am Französischen: Für die Entwicklung der modernen fachsprachlichen Bedeutung von engl. *hallucination* im 19. Jahrhundert zitiert das Oxford English Dictionary eine Übersetzung aus dem Französischen von 1859.

Diese wenigen Stationen zeigen, dass das Französische für die Entwicklung der modernen Bezeichnung eine wichtige Rolle spielte. Die moderne Verwendung im Deutschen und Englischen baut darauf auf. Aus der literarischen Bedeutungsübertragung von ‚Herumirren' auf ‚Verblendung' entstand die Bezeichnung für eine visuelle Störung und schließlich die psychiatrische Bedeutung der Perzeptionsstörung.

Parallel blieb die literarische Metapher in Gebrauch und wurde in die Allgemein- und Umgangssprache übernommen. Auch hier bringt der Vergleich unterschiedlicher Sprachen Ähnlichkeiten hervor [6]. Die Bedeutungen reichen von ‚verrückt, drogensüchtig' (fr. *halluciné, -ée*, Dictionnaires Le Robert) und ‚stark beeindrucken' oder ‚verwirren' (it. *allucinare*, Zingarelli) bis zu *allucinato* ‚von Folgen des Drogenkonsums betroffen sein' (belegt in der italienischen Jugendsprache [7]). Somit setzt sich einerseits die ältere allgemeinsprachliche Verwendung im Sinne von ‚Verirrung, Verwirrung' fort. Andererseits passt diese jugendsprachliche Verwendung zur Beliebtheit von Krankheitsnamen in Jugendsprachen, wie sie in der Forschung beschrieben wurden [8]. Außerdem entstehen Varianten, wenn zum Beispiel unter den Peers jemand abgekürzt als *Spast* beschimpft oder angesprochen wird. Im Englischen gibt es die Wortspiele *hearlucinating*, *halusion* oder *hallucification* (Urban Dictionary 1995–2025).

Weil die medizinische Bedeutung heute etabliert ist, erzeugt die Übertragung von *Halluzination* auf den Kontext der KI Aufmerksamkeit, vielleicht ein Grund für die Popularität des Begriffs. Die Wirkung entsteht durch den Kontrast der Welten: Ein krankheitsbedingtes Symptom wird metaphorisch auf ein computerbezogenes Phänomen bezogen.

Ein solcher Rückgriff auf Metaphern ist ein typisches Phänomen neu etablierter Disziplinen. Gerade Metaphern mit menschlichen Eigenschaften, also anthropomorphe Metaphern, kommen in Fachsprachen vor [9]. Im Bereich der KI ist zum Beispiel von *Chat*-Bots die Rede. Das Wort *chat* ‚unterhalten, im Gespräch sein' macht die KI zu einem *Gesprächs*partner, als säße ein Mensch gegenüber. Diese Vorstellung wird durch die Interaktivität der Nutzungsoberflächen unterstützt, Nachrichten können schnell ausgetauscht werden, es wird geduzt und nicht selten geschimpft. Diese Vermenschlichung des Werkzeugs passt dazu, dass die Ergebnisse ebenfalls vermenschlicht und Fehler als *Halluzinationen* bezeichnet werden – wie das Krankheitssymptom.

Handelt es sich deshalb bei *hallucinations* um einen „misnomer", wie McGowan et al. [10] schrieben? Aus Sicht der Psychiatrie ist das der Fall. Erstens bezeichnen *Halluzinationen* falsche Wahrnehmungen, während die fehlerhaften Ergebnisse der Sprachmodelle als *Erzeugnisse* zu verstehen sind („fabrications and falsifications", [11]). Zweitens sind Halluzinationen ein Kernsymptom der Schizophrenie. Østergaard und Nielbo [12] weisen auf die individuelle Dimension der Betroffenheit durch diese Krankheit hin und auf die Stigmatisierung, die damit häufig einhergeht.

Es gibt also Gründe dafür, die konnotationsreiche Metapher *Halluzination* als Beschreibung irreführender Passagen in KI-generierten Texten selbst irreführend und stigmatisierend zu nennen. Weniger eingängig, aber zutreffend beschreibt dagegen *unsubstantiated content* [13] den zu Grunde liegenden Sachverhalt. Eine Bezeichnung wie diese zeigt, dass das Phänomen des fehlerhaften Outputs besser verstanden wird und damit auch die Bezeichnung präziser wird. Zug um Zug differenziert sich das wissenschaftliche Verständnis der Funktionsweisen von KI aus und die Terminologie wird genauer.

Antoine-Laurent de Lavoisier (1743–1794) erarbeitete die moderne chemische Nomenklatur [14] und ersetzte dadurch die alchimistischen Metaphern beispielsweise von ‚Butter' (*beurre d'arsénic*) und ‚Öl' (*huile de vitriol*), die er als äußerst ungenau ablehnte. Er folgerte:

> „Man kann die Sprache nicht perfektionieren, ohne die Wissenschaft zu perfektionieren, und die Wissenschaft nicht ohne die Sprache."

Literatur

[1] Baker, Simon; Kanade, Takeo (2000): Hallucinating faces. In: Proceedings Fourth IEEE International Conference 2000. S. 83–88. DOI 10.1109/AFGR.2000.840616

[2] Ji, Ziwei; Lee, Nayeon; Frieske, Rita; Yu, Tiezheng; Su, Dan; Xu, Yan; Ishii, Etsuko; Bang, Ye Jin; Madotto, Andrea; Fung, Pascale (2023): Survey of Hallucination in Natural Language Generation. In: ACM Computing Surveys 55(12), S. 1–38. DOI 10.1145/3571730

[3] Maynez, Joshua; Narayan, Shashi; Bohnet, Bernd; McDonald, Ryan (2020): On faithfulness and factuality in abstractive summarization. In: Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, S. 1906–1919. DOI 10.18653/v1/2020.acl-main.173

[4] Fernel, Jean (1655): La médecine pratique à l'époque moderne. Paris: Guignard. Digitalisat ark:/13960/t3rv73w51, Bibliothèque interuniversitaire de santé Paris, S. 344. https://archive.org/details/BIUSante_33396

[5] Müller, Johannes (1826): Über die phantastischen Gesichtserscheinungen. Koblenz: Hölscher. Digitalisat Deutsches Textarchiv, S. 25 und S. 33. https://www.deutschestextarchiv.de/mueller_gesichtserscheinungen_1826

[6] Gerstenberg, Annette (2024): Hallucinations in automated texts – A critical view on the emerging terminology. In: AI-Linguistica. Linguistic Studies on AI-Generated Texts and Discourses 1(1), S. 1–8. DOI 10.62408/ai-ling.v1i1.9

[7] Banfi, Emanuele (1992): Conoscenza e uso di lessico giovanile a Milano e a Trento. In: Banfi, Emanuele; Sobrero, Alberto A. (eds.), Il linguaggio giovanile degli anni novanta. S. 99–138. Milano: Laterza, S. 128.

[8] Neuland, Eva (1998): Zum Sprachgebrauch Jugendlicher verschiedener regionaler Herkunft. In: Androutsopoulos, Jannis K.; Scholz, Arno (eds.): Jugendsprache = Langue des jeunes = youth language: linguistische und soziolinguistische Perspektiven (VarioLingua 7), S. 71–90. Frankfurt am Main, New York: Lang, S. 74.

[9] Ickler, Theodor (1993): Zur Funktion der Metapher, besonders in Fachtexten. In: Fachsprache 15 (3-4). S. 94–110, S. 104.

[10] McGowan, Alessia; Gui, Yunlai; Dobbs, Matthew; Shuster, Sophia; Cotter, Matthew; Selloni, Alexandria; Goodman, Marianne; Srivastava, Agrima; Cecchi, Guillermo A.; Corcoran, Cheryl M. 2023. ChatGPT and Bard exhibit spontaneous citation fabrication during psychiatry literature search. In: Psychiatry research 326. 115334. DOI 10.1016/j.psychres.2023.115334

[11] Emsley, Robin (2023): ChatGPT: these are not hallucinations – they're fabrications and falsifications. In: Schizophrenia 9, 52, S. 1. DOI 10.1038/s41537-023-00379-4

[12] Østergaard, Søren Dinesen; Nielbo, Kristoffer Laigaard (2023): False Responses From Artificial Intelligence Models Are Not Hallucinations. In: Schizophrenia bulletin 49 (5), S. 1105–1107, S. 1107. DOI 10.1093/schbul/sbad068

[13] Chen, Canyu; Shu, Kai (2023): Combating Misinformation in the Age of LLMs: Opportunities and Challenges. arXiv. https://arxiv.org/abs/2311.05656

[14] Lavoisier, Antoine-Laurent de (1789): Traité élémentaire de chimie, présenté dans un ordre nouveau, et d'après les découvertes modernes. Paris: Cuchet. Digitalisat Gallica der Bibliothèque Nationale Française https://gallica.bnf.fr/ark:/12148/btv1b8615746s

Prof. Dr. Annette Gerstenberg ist Professorin für Romanische Sprachwissenschaft an der Universität Potsdam. Sie publiziert u.a. zur geschichtlichen Entwicklung von Fachsprachen und befasst sich mit der Rolle der Künstlichen Intelligenz in Forschung und Lehre.

Kontaktadresse
annette.gerstenberg@uni-potsdam.de
www.uni-potsdam.de/romanistik-gerstenberg

Anzeige

Plus das Gratis-Online-Event

Mehr Informationen:

tekom Jahrestagung 2025
STUTTGART 11.–13. NOV.

tekom MESSE 25

TECHNOLOGY DAYS
DAS DIGITAL-EVENT
3.+4. NOVEMBER 2025

Glossare für die neuronale maschinelle Übersetzung

Detlef Reineke

Neural machine translation engines usually need human-prepared glossaries to force these engines to use prescriptive terminology in the target text rather than insert terms determined by probabilistic means. Automated creation of multilingual glossaries for the use in NMT systems may sound easy, but as is often the case in data engineering tasks, the devil lies in the detail. The present article describes how (multiple) NMT glossaries can be generated from terminology databases making use of script and programming languages such as VBA or XSLT.

Keywords: Neural Machine Translation, NMT, glossary, VBA, XSLT

Die korrekte Verwendung von Terminologie stellt bekanntlich eine der großen Herausforderungen für die effiziente Arbeit mit statistischen und neuronalen MÜ-Systemen dar. In regelbasierten MÜ-Systemen wird die Verwendung von Terminologie mithilfe von Wörterbüchern gesteuert, in SMÜ- und NMÜ-Systemen dagegen werden Bezeichnungen [1] statistisch bzw. probabilistisch ermittelt. Schon in SMÜ-Systemen wurden Glossare herangezogen, um die korrekte Verwendung der Terminologie in den Zieltexten zu erzwingen. SMÜ-Systeme sind zwischenzeitlich jedoch größtenteils durch NMÜ-Systeme verdrängt worden. Die nachfolgenden Ausführungen beziehen sich daher ausschließlich auf die Erstellung von NMÜ-Glossaren.

Formale Anforderungen an NMÜ-Glossare

NMÜ-Glossare sind Textdateien, die je nach NMÜ-System in unterschiedlichen Formaten vorliegen müssen (CSV, TSV, TMX, XLIFF). In CSV- und TSV-Dateien gilt allgemein: eine Zeile, ein Eintrag. Einträge bestehen im Prinzip aus Bezeichnungspaaren bzw. Bezeichnungsreihen in Ausgangs- und Zielsprache(n). Die Daten eines Eintrags werden durch entsprechende Separatoren (Komma, Tabulator) voneinander getrennt. Die Angabe der Sprache erfolgt dabei in der Kopfzeile oder direkt hinter den Bezeichnungen eines Eintrags. Das folgende Beispiel zeigt einen Eintrag für ein DeepL-Glossar mit nur einer Ausgangs- und Zielsprache:

```
circuit-breaker,Leistungsschalter
```

In DeepL können theoretisch auch CSV- bzw. TSV-Glossare hochgeladen werden, die mehr als eine Sprachkombination und Sprachrichtung enthalten. In solchen Fällen müssen die Sprachen hinter den Bezeichnungen angegeben werden:

```
circuit-breaker,Leistungsschalter,EN,DE
circuit-breaker,interruptor automático,EN,ES
disjoncteur,circuit-breaker,FR,EN
```

Die Verwendung von Kommata als Trennzeichen ist natürlich problematisch, da Bezeichnungen u. U. selbst Kommata enthalten können, die die im vorangegangenen Beispiel gezeigte Eintragsstruktur aushebeln würden, wie das folgende Beispiel verdeutlicht:

```
switchgear, controlgear and fuses,Schaltgeräte
```

Dieses Problem lässt sich durch die Verwendung von Tabulatoren als Trennzeichen umgehen:

```
switchgear, controlgear and fuses→Schaltgeräte
```

Neben der Einhaltung der Eintragsstruktur sind je nach NMÜ-System noch zusätzliche Anforderungen zu beachten. DeepL-Glossare z. B. müssen in der Zeichencodierung UTF-8 vorliegen, dürfen keine Kopfzeile enthalten und sollten frei von ausgangssprachlichen Homographen sein, also Wörtern derselben Schreibweise, die aber unterschiedliche Begriffe repräsentieren.

Glossarformate anderer Dienste sind ähnlich einfach aufgebaut. In Microsoft Translator können neben CSV- und TSV-Glossaren auch Glossare im XLIFF-Format (Version 1.2) hochgeladen werden (siehe Abbildung 1 auf S. 20).

Amazon Translate unterstützt neben mehrsprachigen CSV- und TSV-Dateien auch Glossare im TMX-Format (siehe Abbildung 2 auf S. 20).

In Google-Translate-Glossaren lassen sich im CSV- bzw. TSV-Format neben Sprachkürzeln und Bezeichnungen noch weitere Informationen einbinden. Das nachfolgende Beispiel zeigt eine Eintragsstruktur, die auch Angaben zur Wortklasse

```
<?xml version="1.0" encoding="UTF-8"?>
<xliff version="1.2">
  <file datatype="plaintext" original="self"
   source-language="en" target-language="de">
    <body>
      <trans-unit id="1">
        <source>circuit-breaker</source>
        <target>Leistungsschalter</target>
      </trans-unit>
    </body>
  </file>
</xliff>
```

Abb. 1: Microsoft-Translator-Glossar im XLIFF-Format

```
<?xml version="1.0" encoding="UTF-8"?>
<tmx version="1.4">
  <header segtype="sentence" o-tmf="UTF-8"
   adminlang="en" srclang="en" datatype="PlainText"/>
  <body>
    <tu>
      <tuv xml:lang="en">
        <seg>circuit-breaker</seg>
      </tuv>
      <tuv xml:lang="de">
        <seg>Leistungsschalter</seg>
      </tuv>
      <tuv xml:lang="es">
        <seg>interruptor automático</seg>
      </tuv>
    </tu>
  </body>
```

Abb. 2: Amazon-Translate-Glossar im TMX-Format

oder eine Begriffsbeschreibung ermöglicht:

en→de→es→fr→pos→description
circuit-breaker→Leistungsschalter→interruptor automático→disjoncteur→noun→mechanical switching device, capable of making, carrying and breaking currents under normal circuit conditions and also making, carrying for a specified duration and breaking currents under specified abnormal circuit conditions such as those of short circuit

Glossare für die Nutzung in NMÜ-Systemen werden üblicherweise aus Terminologiedatenbanken erzeugt. Aber auch aus anderen mehrsprachigen Wissensressourcen wie Thesauri, Taxonomien, Ontologien, Produktdatenwörterbüchern oder Normen lassen sich im Prinzip NMÜ-Glossare erstellen. Beispielhaft soll hier dargestellt werden, wie sich Glossare aus Microsoft-Excel-Tabellen und XML-basierten Exportdateien generieren lassen.

Glossare aus Microsoft Excel mit VBA

Vielerorts werden terminologische Daten immer noch in Kalkulationsprogrammen wie Microsoft Excel eingepflegt und verwaltet. Um aus einer umfangreichen, mehrspaltigen, vielleicht sogar über mehrere Arbeitsblätter verteilten Terminologiedatenbank ein oder mehrere Glossare zu erzeugen, die den o. a. Anforderungen an NMÜ-Glossare entsprechen, müssen terminologische Bestände entsprechend aufbereitet werden. Hier können datentechnisch zwei grundsätzlich unterschiedliche Ansätze verfolgt werden, die zum gewünschten Ergebnis führen:

1. Der terminologische Datenbestand wird bereits im Kalkulationsprogramm auf die Grundstruktur der NMÜ-Glossare heruntergebrochen.
2. Die terminologischen Daten werden z. B. nach XML exportiert und im Nachgang in das entsprechende Glossarformat transformiert.

Die Aufbereitung der Terminologie in einem Kalkulationsprogramm könnte theoretisch manuell vorgenommen werden. Allerdings ist eine solche Arbeitsweise zeitaufwendig und fehlerträchtig. In den nachfolgenden stark vereinfachten Beispielen soll zunächst veranschaulicht werden, welche grundsätzlichen Arbeitsschritte anfallen, um aus Microsoft Excel DeepL-Glossare zu erstellen. Beginnen wir mit einem nicht allzu realistischen, aber doch anschaulichen Anwendungsfall, bei dem der terminologische Datenbestand bereits in einer Eintragsstruktur vorliegt, die der Eintragsstruktur von DeepL-Glossaren entspricht (siehe Abbildung 3).

	A	B
1	circuit-breaker	Leistungsschalter
2	air circuit-breaker	Luftleistungsschalter
3	oil circuit-breaker	Ölleistungsschalter
4	SF_6 circuit-breaker	SF_6-Leistungsschalter
5	vacuum circuit-breaker	Vakuumleistungsschalter

Abb. 3: DeepL-konformer Terminologiedatenbestand in Microsoft Excel

	A	B	C
1	EN	DE	ES
2	circuit-breaker	Leistungsschalter	interruptor automático
3	air circuit-breaker	Luftleistungsschalter	interruptor automático en aire
4	oil circuit-breaker	Ölleistungsschalter	interruptor automático en aceite
5	SF_6 circuit-breaker	SF_6-Leistungsschalter	interruptor automático de SF_6
6	vacuum circuit-breaker	Vakuumleistungsschalter	interruptor automático de vacío

Abb. 4: Dreisprachiger Terminologiebestand mit Kopfzeile

Im vorliegenden Fall könnten die Daten z. B. direkt im voreingestellten TXT-Format (tabulatorgetrennt) abgespeichert werden. Anschließend wäre noch die Dateinamenerweiterung anzupassen. Aber bereits an diesem Punkt ist Vorsicht geboten, denn in den Versionen bis Microsoft Excel 2016 werden tabulatorgetrennte TXT-Dateien standardmäßig in der Zeichencodierung ANSI, ab Version 2019 in der Zeichencodierung UTF-8 gespeichert. Im Prinzip wäre es auch möglich, die Daten im voreingestellten TXT-Unicode-Format zu speichern, aber auch hier muss nachkonvertiert werden, denn für dieses Format ist die Zeichencodierung UTF-16 LE BOM festgelegt.

Gehen wir einen Schritt weiter und betrachten einen etwas komplexeren Anwendungsfall, wie er in Abbildung 4 auf Seite 20 dargestellt wird. Für ein DeepL-Glossar in der Sprachkombination/-richtung Englisch-Deutsch müsste die Spalte C entfernt und die Kopfzeile unterdrückt werden. Die so modifizierte Terminologiedatenbank soll als TSV-Datei in der Zeichencodierung UTF-8 gespeichert werden. Und zu guter Letzt wollen wir in der Regel ja auch nach dem Speichern des Glossars wieder zurück in die Arbeitsmappe, um in Microsoft Excel weiterarbeiten zu können.

All diese Arbeitsschritte lassen sich in Microsoft-Excel-Desktop-Versionen im Prinzip relativ einfach mithilfe der Scriptsprache VBA[1] (Visual Basic for Applications) automatisieren [2]. Nur die Zeichencodierung verkompliziert den VBA-Code ein wenig, da in Microsoft Excel, auch in den Desktop-Versionen ab 2019, bei der Erzeugung von Textdateien standardmäßig ANSI als Zeichencodierung festgelegt wird.

Beginnen wir aber zunächst damit, im VBA-Editor von Microsoft Excel ein Modul und im Modul eine Sub-Prozedur einzufügen, um die entsprechenden Arbeitsschritte zu codieren (siehe Abbildung 5). In der Sub-Prozedur deeplEnDe() werden in einem ersten Schritt die entsprechenden Variablen deklariert und initialisiert. In der Variable glosFileEnDe sind Pfad und Name des zu erstellenden Glossars abgelegt. Die Variable lastRow legt für den Schleifendurchlauf (siehe Abbildung 6) die letzte Zeile des aktiven Zellbereichs fest (im Beispiel in Abbildung 4 also Zeile 6). Die Variable cellData dient dazu, die Zelldaten bei der Konvertierung nach UTF-8 zwischenzuspeichern.

Anschließend wird der aktive Zellbereich ohne Kopfzeile (i = 2 To lastRow) mithilfe der for-Schleife durchlaufen und der

```
Option Explicit

Sub deeplEnDe()

Dim wbk As Workbook
Dim ws As Worksheet
Dim glosFileEnDe As String
Dim lastRow As Long
Dim i As Long
Dim cellData As String
Dim objStream As Object
Dim fso As Object
Dim file As Object

Set wbk = ActiveWorkbook
Set ws = wbk.ActiveSheet
glosFileEnDe = wbk.Path & "\deepl\deeplEnDe.tsv"
lastRow = ws.Cells(ws.Rows.Count, "A").End(xlUp).Row
cellData = ""
```

Abb. 5: VBA-Sub-Prozedur für den Export eines zweisprachigen DeepL-Glossars nach TSV

```
Set objStream = CreateObject("ADODB.Stream")
objStream.Open
objStream.Type = 2
objStream.Charset = "UTF-8"
objStream.WriteText cellData

objStream.Position = 0
objStream.Type = 1
objStream.Position = 3
Dim byteData() As Byte
byteData = objStream.Read
objStream.Close

Set fso = CreateObject("Scripting.FileSystemObject")
Set file = fso.CreateTextFile(glosFileEnDe, True, False)
file.Write vbNullString
file.Close
```

Abb. 7: VBA-Konvertierung von UTF-8 BOM in UTF-8

Inhalt der durchlaufenen Zellen im Glossarformat in die Variable cellData geschrieben (siehe Abbildung 6 unten).

Der Inhalt der Variablen cellData wird dann über ein ADODB.Stream-Objekt zunächst nach UTF-8 BOM (BOM = Byte Order Mark) konvertiert (siehe Abbildung 7 oben), auch wenn im Code UTF-8 als Wert der objStream. Charset-Eigenschaft angegeben wird. In UTF-8 BOM-Dateien wird den eigentlichen Daten eine 3-Byte-Markierung vorangestellt, was zur Folge haben kann, dass bestimmte Zeichen u. U. nicht wie gewünscht dargestellt werden. Die BOM-Markierung lässt sich jedoch mithilfe der objStream. Position-Eigenschaft überspringen. Im nächsten Schritt wird der Byte-Datenstrom ohne Byte-Markierung mithilfe des Objekts FileSystemObject und der Methode CreateTextFile() in die TSV-Datei geschrieben. Abschließend werden die Byte-Daten über ein neu erstelltes ADODB.

```
For i = 2 To lastRow
    If ws.Cells(i, 1).Value <> "" And ws.Cells(i, 2).Value <> "" Then
        cellData = cellData & ws.Cells(i, 1).Value & vbTab & _
        ws.Cells(i, 2).Value & vbCrLf
    End If
Next i
```

Abb. 6: VBA-Schleife, die die Zellinhalte in die Variable cellData schreibt

[1] Die webbasierte 365-Version unterstützt kein VBA, sondern ist mit einer JavaScript/TypeScript-API ausgestattet.

```
Set objStream = CreateObject("ADODB.Stream")
objStream.Open
objStream.Type = 1
objStream.Write byteData
objStream.SaveToFile glosFileEnDe, 2
objStream.Close

End Sub
```

Abb. 8: VBA-Konvertierung des Byte-Datenstroms in Binärdaten

```
Option Explicit

Sub deeplAll()

    Call deeplEnDe
    Call deeplDeEn
    Call deeplEnEs
    Call deeplEsEn
    Call deeplDeEs
    Call deeplEsDe

End Sub
```

Abb. 9: VBA-Sub-Prozedur für die Erstellung mehrerer zweisprachiger DeepL-Glossare

```
import os
import sys
filePathSrc="D:\\nmue-glossar"
for root, dirs, files in os.walk(filePathSrc):
    for fn in files:
      if fn[-4:] == '.tsv':
        notepad.open(root + "\\" + fn)
        notepad.runMenuCommand("Encoding", "Convert to UTF-8")
        notepad.save()
        notepad.close()
```

Abb. 10: Python-Script für die UTF 8-Konvertierung von TSV-Dateien in Notepad++

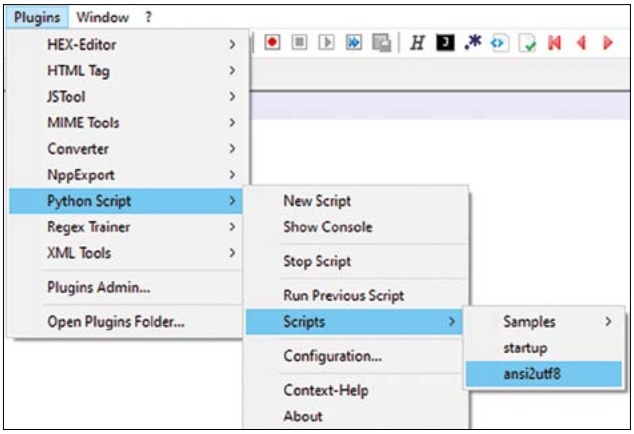


Abb. 11: Ausführung des Python-Scripts in Notepad++

Stream-Objekt als Binärdaten in der TSV-Datei gespeichert (siehe Abbildung 8).

Sollen bei der Ausführung eines VBA-Makros mehrere zweisprachige Glossare erstellt werden, müssen die Argumente der ws.Cells()-Eigenschaften und die Sprachvariablen entsprechend angepasst bzw. umgestellt werden. Über Call-Anweisungen lassen sich die Glossare dann in einem Durchlauf in den gewünschten Sprachkombinationen und Sprachrichtungen erstellen (siehe Abbildung 9).

Eine Alternative zur UTF-8-Konvertierung in VBA ist ein einfaches Python-Script (siehe Abbildung 10), mit dem sich Zeichencodierungen der Glossare z. B. im Programmiereditor Notepad++ konvertieren lassen.

Das Python-Script muss im Python-Script-Ordner des Windows-Benutzerordners abgelegt werden und kann dann aus Notepad++ heraus ausgeführt werden (siehe Abbildung 11).

Glossare aus Microsoft Excel mit XML

Eine Alternative zur Aufbereitung terminologischer Daten in Microsoft Excel mittels VBA besteht darin, die Daten über das XML-Daten-Zwischenformat in das finale Glossarformat zu überführen [3]. Zu diesem Zweck muss dem Datenblatt in einem ersten Schritt ein XML-Schema zugeordnet werden, damit die Daten in das XML-Datenformat exportiert werden können. Anschließend kann die XML-Datei dann mithilfe eines Transformationsstylesheets in eine oder mehrere TSV-, TMX- oder XLIFF-Dateien transformiert werden [4]. Wenn die Terminologiedatenbank mehrere Arbeitsblätter umfasst, sollten die Glossardaten zunächst mithilfe von VBA in einem Arbeitsblatt zusammengeführt und dann nach XML exportiert werden.

Glossare aus professionellen Terminologietools

Professionelle Terminologietools sind mit mehr oder weniger leistungsfähigen und anpassbaren Filterfunktionen und Exportschnittstellen ausgestattet, die die Erstellung von NMÜ-Glossaren vereinfachen können. In Terminologietools mit vorgegebener Eintragsstruktur lassen sich Filter und Exportschnittstellen in der Regel nur in geringem Umfang an die Anforderungen anpassen, die an NMÜ-Glossarformate gestellt werden. Wordbee z. B. unterstützt zwar die für NMÜ-Glossare erforderlichen Formate CSV, TMX und XLIFF (Version 1.2); allerdings werden immer nur die erste AS-Bezeichnung und die erste ZS-Bezeichnung eines Eintrags exportiert. Dies gilt im Übrigen auch für den Export nach TBX 1.0 (= ISO 30042:2008). Hinzukommt, dass sich bei allen vier vorab genannten Exportformaten die Bezeichnungen auch nicht nach bestimmten Kriterien filtern lassen (z. B. Verwendungsstatus). Außerdem sind Exportdaten im CSV-Format, wie im nachfolgenden Beispiel zu sehen, für die unmittelbare Verwendung als NMÜ-Glossare ungeeignet und müssen für die Nutzung in NMÜ-Systemen z. B. in einem Kalkulationsprogramm nachbearbeitet und exportiert oder mittels Transformation in das erforderliche TSV-Format überführt werden.

```
"Line","English","German","","","","","Comments","Unique ID"
"1","circuit-breaker","Leistungsschalter",,,"","","",,"2766"
"2","vacuum circuit-breaker","Vakuumleistungsschalter",,,"","",,"2767"
```

Der einzige Weg, aus Wordbee hochwertigere Glossare zu erstellen, besteht darin, die Terminologiedatenbank nach TBX 3.0. [sic] zu exportieren [Anmerkung des Autors: Gemeint ist hier die TBX-Version der ISO-Norm 30042 aus dem Jahr 2019, also eigentlich ISO TBX 2.0]. Hier werden alle Bezeichnungen und

```
<?xml version="1.0" encoding="UTF-8"?>
<xsl:stylesheet xmlns:xsl="http://www.w3.org/1999/XSL/Transform"
 xmlns:xs="http://www.w3.org/2001/XMLSchema"
 xmlns:math="http://www.w3.org/2005/xpath-functions/math"
 exclude-result-prefixes="xs math"
 version="3.0">

  <xsl:output method="text"/>

  <xsl:template match="/">
    <xsl:variable name="en-de">
      <xsl:for-each select="//conceptEntry/langSec[@xml:lang= 'en']/termSec/term">
        <xsl:value-of select="."/>
        <xsl:text>&#009;</xsl:text>
        <xsl:value-of select="../../../langSec[@xml:lang = 'de']/termSec/term"/>
        <xsl:text>&#009;EN&#009;DE&#10;</xsl:text>
      </xsl:for-each>
    </xsl:variable>
    <xsl:result-document href="deepl\deeplEnDe.tsv"
     omit-xml-declaration="yes" method="text">
      <xsl:sequence select="$en-de"/>
    </xsl:result-document>
    <xsl:variable name="de-en">
      (Transformationsregeln)
    </xsl:variable>
    <xsl:result-document href="\deepl\deeplDeEn.tsv"
     omit-xml-declaration="yes" method="text">
      <xsl:sequence select="$de-en"/>
    </xsl:result-document>
  </xsl:template>

</xsl:stylesheet>
```

Abb. 12: XSLT-Transformation aus TBX in mehrere zweisprachige DeepL-Glossare

Angaben zum Verwendungsstatus mitexportiert, sodass die für das Glossar bestimmten Bezeichnungen über entsprechende Transformationsregeln herausgezogen werden können. Und wie auch bei der simultanen Erstellung mehrerer Glossare aus Microsoft Excel mit VBA lassen sich z. B. aus XML-basierten Exportdateien wie TBX oder MTF mithilfe eines Transformationsstylesheets in einem Schwung mehrere Glossare in unterschiedlichen Sprachkombinationen und Sprachrichtungen erzeugen. Abbildung 12 zeigt eine stark vereinfachte Transformation, mit der sich Bezeichnungen für zwei DeepL-Glossare in den Sprachkombinationen Englisch-Deutsch und Deutsch-Englisch aus einer TBX-Exportdatei mit nur jeweils einer Bezeichnung pro Sprache und Begriff erzeugen lassen.

Das Übersetzungstool Phrase bietet ein wenig mehr Komfort. Terminologiedatenbanken können für die maschinelle Übersetzung im Tool über den KI-Hub Phrase Language AI genutzt werden. Dazu muss im Hub zunächst ein neues, leeres, zwei- oder mehrsprachiges NMÜ-Glossar erstellt werden. Die exportierte (und in der Regel aufzubereitende) Terminologiedatenbank ist dann in das neu erstellte Glossar zu importieren. Anschließend muss das Glossar nur noch im Hub mit den aktiven NMÜ-Engines verknüpft werden. Von Vorteil ist hier, dass Glossare in Phrase im Prinzip nicht bei jedem Update von Grund auf neu erstellt und verknüpft werden müssen, sondern durch Export (zzgl. Nachbearbeitung) in bestehende Glossare importiert werden können. Der Haken: Auch hier sind im Glossar keine n:1-Entsprechungen zugelassen. AS-Synonyme für bevorzugte ZS-Bezeichnungen fallen also unter den Tisch. Phrase-Terminologiedatenbanken mit Synonymen müssen daher nach dem Export noch entsprechend aufbereitet werden. Relativ einfach gestaltet sich die Aufbereitung exportierter XSLX-Dateien, denn die Spalten für die synonymen Bezeichnungen und die dazugehörigen Metadaten sind leicht zu identifizieren und lassen sich mit wenigen Mausklicks löschen. Phrase-TBX-Exportdateien dagegen wären zu transformieren.

Noch etwas bequemer geht es mit MultiTerm. Terminologiedatenbanken lassen sich in MultiTerm über einstellbare Filter (siehe Abbildung 13 auf Seite 24) und Exportschnittstellen (siehe Abbildung 14 auf Seite 25) in zweisprachige TSV-, TMX- und XLIFF-Dateien exportieren. Die UTF-16 LE BOM-Zeichencodierung dieser Exportdateien muss dann allerdings anschließend noch in die UTF-8-Zeichencodierung umgewandelt werden.

Für die Erstellung von Glossaren mit mehr als zwei Sprachen und die gleichzeitige Erstellung mehrerer zweisprachiger Glossare empfiehlt sich der Export nach MTF bzw. TBX mit anschließender Transformation.

Neben diesem herkömmlichen Ansatz für die Verwendung von Terminologie in der maschinellen Übersetzung bietet MuliTerm noch eine Besonderheit: Über entsprechende APIs können KI-Modelle direkt auf Bezeichnungen (allerdings ohne Abgleich der Metadaten) in Terminologiedatenbanken zugreifen. Separate Glossare sind hier also im Prinzip überflüssig.

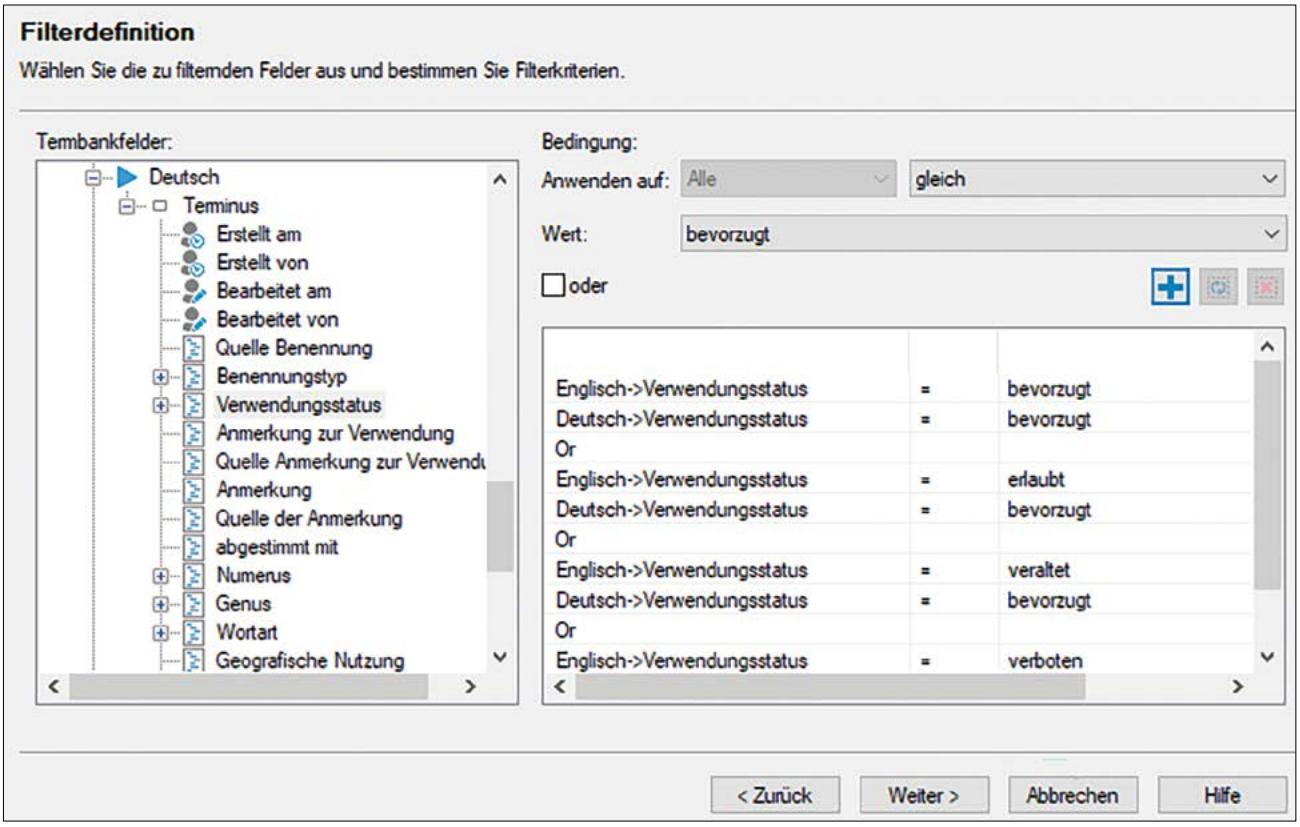


Abb. 13: Filterdefinition in MultiTerm

Qualität der Terminologiedatenbank

Die vorhergehenden Ausführungen haben gezeigt, dass der technische Aspekt bei der Erstellung von NMÜ-Glossaren nicht unterschätzt werden sollte. Weit wichtiger ist jedoch die Qualität der Terminologiedatenbank [5]. Ein Grundsatz lautet: Je geringer die Qualität einer Datenbank, desto (exponentiell) höher der Aufwand bei der Aufbereitung der Daten für die Nutzung in anderen Anwendungen. In der nachfolgenden, nicht abschließenden Liste sind Problemfälle aufgeführt, die sich in der Praxis bei der Erstellung von NMÜ-Glossaren aus Terminologiedatenbanken beobachten lassen:

- ZS-Bezeichnungen für einen Begriff sind vorhanden, die entsprechenden Bezeichnungen in der Ausgangssprache fehlen jedoch.
- Für die Zielsprache eines Begriffs sind zwar Bezeichnungen im Begriffseintrag abgelegt, aber keine der Bezeichnungen ist als Vorzugsbezeichnung ausgezeichnet.
- Für die Zielsprache eines Begriffs wurde mehr als eine Vorzugsbezeichnung erfasst (z. B. Bezeichnungen in geschlechtergerechter Sprache wie „Rechtsanwalt" und „Rechtsanwältin").
- Der Verwendungsstatus (bevorzugt, erlaubt, verboten usw.) wird nicht durchgehend bei allen AS- und ZS-Bezeichnungen angegeben.
- Die Angabe zum Bezeichnungstyp (Vollform, Abkürzung, Symbol usw.) fehlt.
- Abkürzungen sollen mit Abkürzungen übersetzt werden, auch wenn die zielsprachliche Abkürzung nicht die Vorzugsbezeichnung ist.
- Der Abkürzungstyp (Akronym, Silbenkurzwort usw.) wird nicht korrekt und nicht einheitlich gekennzeichnet.
- Bezeichnungsfelder enthalten überflüssige Leerzeichen.
- Die Terminologiedatenbank enthält Homographe.
- usw.

Homographe stellen bei der Erstellung von NMÜ-Glossaren aus Terminologiedatenbanken, die nicht von Beginn an auch für die maschinelle Übersetzung konzipiert worden sind, ein größeres Problem dar. Homographe lassen sich in einer Terminologiedatenbank häufig nicht vermeiden. Sie stellen beim menschlichen (vs. maschinellen) Übersetzen auch kein Hindernis dar, denn der Mensch kann entscheiden, welche ZS-Bezeichnung verwendet werden soll. NMÜ-Glossare sollten aus naheliegenden Gründen jedoch keine Homographe enthalten. Ein erster Schritt in Richtung Datenbankoptimierung besteht darin zu prüfen, ob eine homographe Bezeichnung in der Ausgangssprache wirklich zwingend notwendig ist oder ob nicht vielleicht doch geeignetere Bezeichnungen zur Verfügung stehen. Die Zahl der Homographe gegen null zu reduzieren, ist allerdings nicht immer möglich und auch nicht immer wünschenswert. In solchen Fällen ist abzuwägen, ob ggf. fachgebietsspezifische Glossare zu erstellen sind oder ob in der Datenbankdefinition auf Bezeichnungsebene

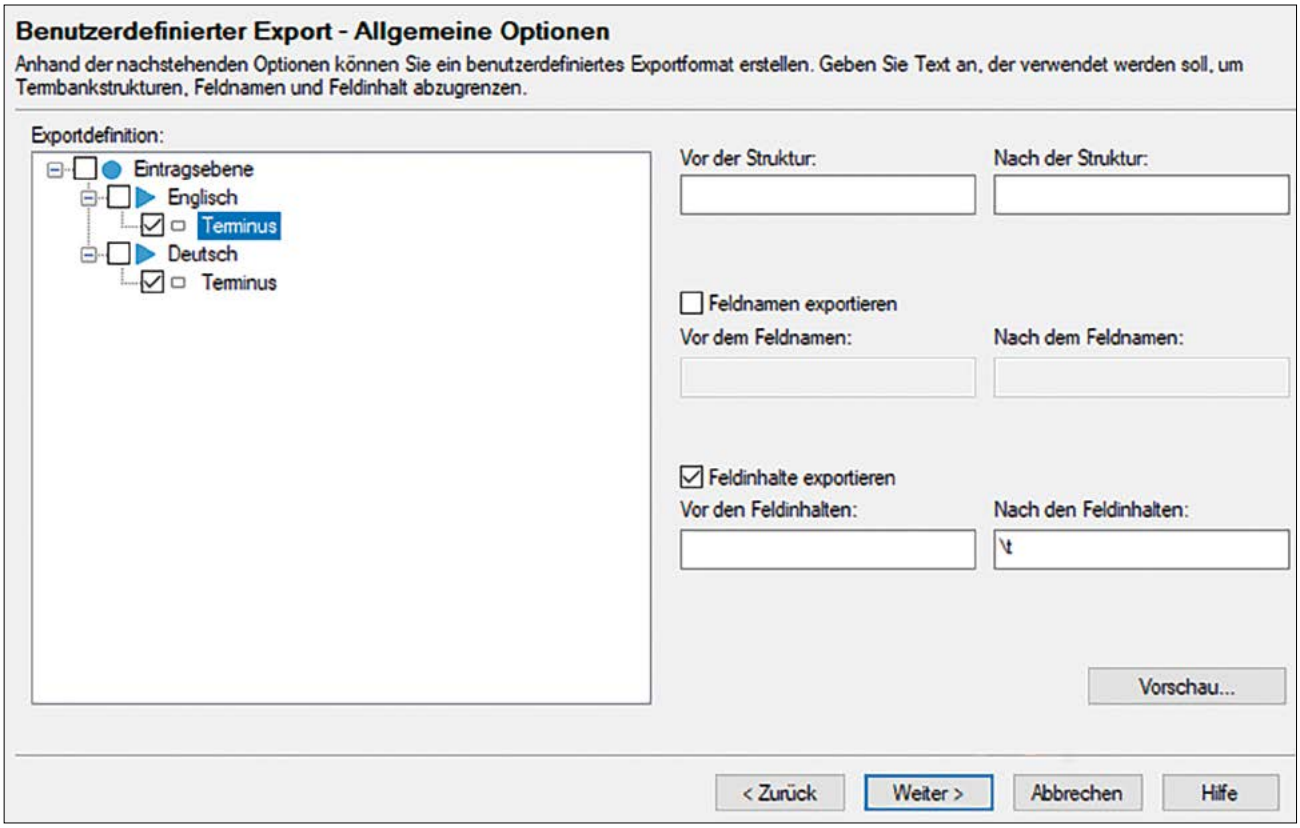


Abb. 14: TSV-Export aus MultiTerm

eine zusätzliche Datenkategorie einzurichten ist, um unerwünschte homographe AS-Bezeichnungen beim Export auszuschließen.

Fazit

In Zukunft werden sicherlich noch weitere Hersteller Schnittstellen in ihre Terminologietools integrieren, die einen direkten Zugriff von NMÜ-Engines und KI-Chatbots auf Bezeichnungen UND Metadaten in Terminologiedatenbanken ermöglichen. Bis es so weit ist, werden aber auch weiterhin solide Kenntnisse und Fähigkeiten im Bereich des Daten-Engineerings und der Programmierung erforderlich sein, um qualitativ hochwertige, passgenaue Glossare für die maschinelle Übersetzung bereitzustellen. Welche Script- und Programmiersprachen zum Einsatz kommen und welche Strategie gefahren werden sollte, hängt vom Anwendungsfall ab, wird aber natürlich auch von (Vor-)Kenntnissen und persönlichen Präferenzen bestimmt. Inkonsistenzen und sonstige Verunreinigungen in der Terminologiedatenbank lassen sich z. B. über den Export der gesamten Datenbank und eine nachgeschaltete Transformation mithilfe von XSLT besser dokumentieren und abfangen. Für die Erstellung von Glossaren aus Terminologiedatenbanken in Microsoft-Excel-Desktop-Versionen müssen die Daten in der Regel mithilfe von VBA-Makros aufbereitet werden. Alternativ können die Daten aber auch nach XML exportiert und im Nachgang mithilfe von XSLT transformiert werden. Summa summarum: Es gibt immer einen Weg, Daten von A nach B zu migrieren. Das A und O bei der Erstellung von NMÜ-Glossaren ist und bleibt aber auch weiterhin die Qualität der Terminologiedatenbank.

Literatur

[1] ISO 1807:2019 Terminology work and terminology science — Vocabulary. Genf: ISO.
[2] Microsoft (2025): Office VBA Reference. https://learn.microsoft.com/en-us/office/vba/api/overview.
[3] Reineke, Detlef (2022): XML-Datenaustausch mit Microsoft Excel. In: edition 2/2022, 35-41.
[4] W3C (2017): XSL Transformations (XSLT) Version 3.0. https://www.w3.org/TR/xslt-30.
[5] ISO 26162-1:2019 Management of terminology resources — Terminology databases. Part 1: Design. Genf: ISO.

Dr. Detlef Reineke lehrt und forscht an der Fakultät für Übersetzen und Dolmetschen der Universität Las Palmas de Gran Canaria in den Bereichen Terminologie, Lokalisierung, Maschinelles Übersetzen und KI. Er war langjähriges DTT-Vorstandsmitglied und DTT-Geschäftsführer, leitete die Fachzeitschrift edition in den Jahren 2006-2017 und hat an zahlreichen grundlegenden nationalen und internationalen Terminologienormen mitgearbeitet. Außerdem ist er als Berater und Entwickler in der Sprachindustrie tätig.

Kontaktadresse
data-engineering@posteo.de

Arbeitsoptimierung und -automatisierung für den Hausgebrauch

Maurice Mayer-Dewor

Die Themen Prozessoptimierung, Integration und Automatisierung sind seit geraumer Zeit – und nicht erst seit dem Beginn der KI-Revolution – auch in der Sprachindustrie in aller Munde, insofern diese über das nötige technische Know-how und die entsprechenden Ressourcen verfügt. In Bereichen der Sprachindustrie hingegen, in denen dies nicht gegeben ist, behilft man sich – verständlicherweise – mit dem, was man kann und hat – auch wenn das bedeutet, dass die Arbeitsabläufe von einer manuellen Durchführung mit häufig repetitiven und somit fehleranfälligen Arbeitsschritten geprägt sind.

Die Hürden, ineffiziente Prozesse – sei es im Kleinen oder im Großen – zu optimieren, erscheinen häufig hoch, und im Alltag fehlt hierfür scheinbar auch gerne die nötige Zeit. Doch genau an diesem Punkt anzusetzen, zahlt sich schnell aus. Es ist schlichtweg unsere Pflicht, im Zeitalter von MÜ und KI in puncto einer Optimierung der eigenen Arbeitsweise unsere Hausaufgaben zu machen. Dieser Beitrag möchte sich in dieser Hinsicht als Impulsgeber verstehen, als kleine Starthilfe in Sachen Arbeitsoptimierung und -automatisierung.

Optimierung und Automatisierung im Kleinen

Viele Wege führen bekanntlich nach Rom und entsprechend kann man auf unterschiedliche Weise vorgehen, um alltägliche Dinge umzusetzen. Ideal ist es aber natürlich, wenn man sozusagen auf der Direttissima nach Rom bzw. ans Ziel gelangt.

Tastenkombinationen können einen ersten grundlegenden Schritt hin zu einer Optimierung bzw. Automatisierung der eigenen Arbeit darstellen. Diese für die häufigsten Arbeitsschritte parat zu haben, beschleunigt die tägliche Arbeit und erhöht somit die eigene Effizienz.

Gut ist es natürlich aus Nutzersicht, wenn die identischen Tastenkombinationen über mehrere Anwendungen hinweg funktionieren, wie z. B. **Strg+F** zur Suche in einem Dokument, **Strg+N** zum Anlegen einer neuen Datei (oder eines neuen terminologischen Eintrags) oder auch **Strg+Eingabetaste**, um in einem TM-System die Übersetzung des aktuellen Segments zu bestätigen und zum nächsten Segment zu springen. (Eine Übersicht über die Tastenkombinationen einer Software findet sich für gewöhnlich in der entsprechenden Online-Hilfe, die übrigens per Konvention durch Drücken der **F1**-Taste aufgerufen werden kann.)

Es ist natürlich ebenfalls hilfreich, wenn sich softwareseitig neue Tastenkombinationen hinzufügen und die bestehenden an die individuellen Bedürfnisse anpassen lassen (s. Abbildung 1). Sinnvoll wäre es in diesem Kontext beispielsweise, wenn man in MultiTerm **Strg+N** verwenden könnte, um einen neuen Eintrag anzulegen (anstelle der **F3**-Taste, die bekanntlich u.a. im Browser auch für eine Suche in der aktuellen

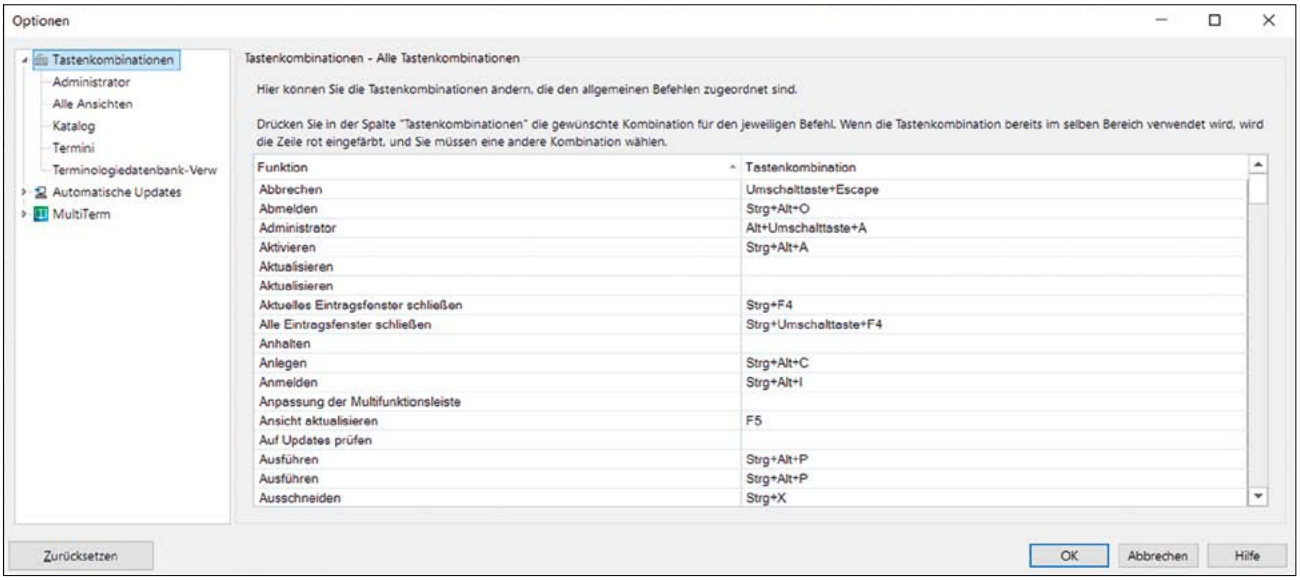


Abb. 1: Übersicht und Anpassungsmöglichkeit der Tastenkombinationen einer Anwendung, hier am Beispiel von MultiTerm (Quelle: eigener Screenshot)

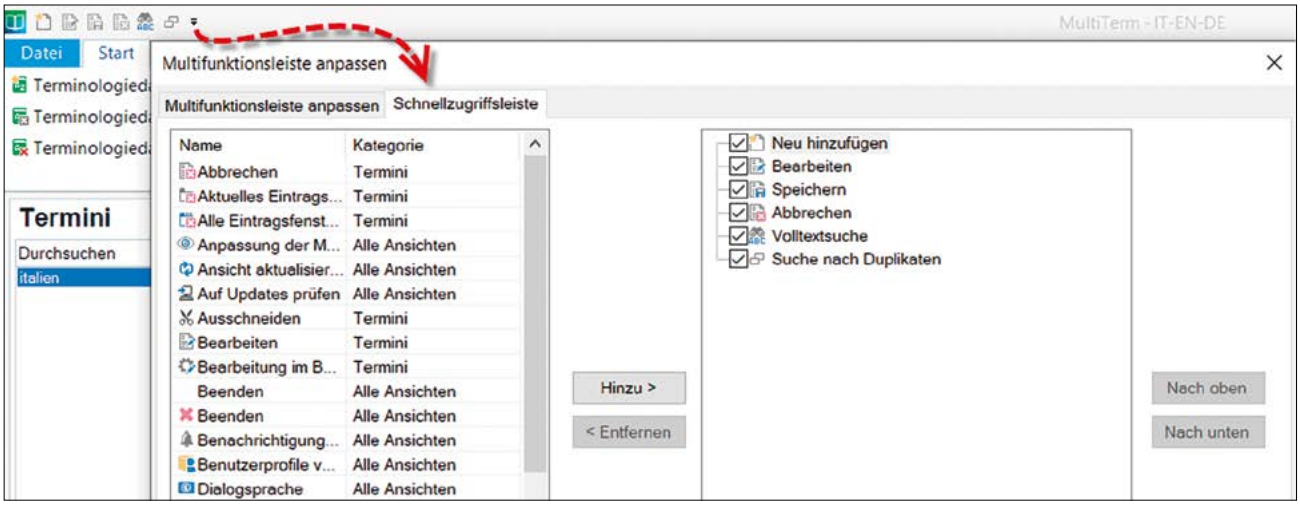


Abb. 2: Anpassungsmöglichkeit der Schnellzugriffsleiste am Beispiel von MultiTerm (Quelle: eigener Screenshot)

Webseite verwendet wird). Kurios ist in diesem Zusammenhang jedoch, dass etwa MultiTerm zwar grundsätzlich eine Anpassung der Tastenkombinationen unterstützt, die Vergabe von **Strg+N** anstelle von **F3** jedoch nicht zulässt. Aber für Vieles gibt es ja häufig eine Lösung – auf die wir in diesem Fall später noch zurückkommen werden.

Ein weiterer, einfach umzusetzender Ansatz, die eigene Arbeit effizienter zu machen, ist es, die Benutzeroberfläche der eingesetzten Anwendungen – falls möglich – anzupassen. Warum sollen Funktionen, die ohnehin nicht oder nur selten genutzt werden, an zentraler Stelle der Benutzeroberfläche angeordnet sein und Funktionen, die häufig verwendet werden, hingegen nicht? Neben der Anpassung der Befehle in den Symbolleisten bzw. Menübändern sei auf die Verwendung und Anpassbarkeit der Schnellzugriffsleisten hingewiesen (s. Abbildung 2), die häufig zu Unrecht ein Schattendasein zu fristen scheinen.

Mit Textbausteinen der eigenen (Schreib-)Arbeit Flügel verleihen

Übersetzer und Terminologen arbeiten naturgemäß viel mit Text und Texten: Wörter und Textpassagen wollen eingefügt, gelöscht, verschoben werden – und das häufig und „gerne". Da ist es natürlich von Vorteil, wenn diese Dinge möglichst schnell von der Hand gehen. Ein Wort markieren? Einfach z. B. per Doppelklick. Einen Absatz markieren? Ein Dreifachklick genügt. Den gesamten Text markieren: **Strg+A**. Ein Wort löschen? Beispielsweise via **Strg+Entf** (Wort rechts vom Cursor) bzw. **Strg+Backspace** (Wort links vom Cursor). Um nur ein paar wenige Beispiele zu nennen. Das Ganze funktioniert übrigens praktischerweise in nahezu jeder Anwendung, in der Text eingegeben werden kann.

Die Produktivität kann aber z. B. auch dadurch erhöht werden, dass die Schreibgeschwindigkeit gesteigert wird. Übersetzer und Terminologen schreiben berufsbedingt viel. Aber selbst wer schnell schreibt bzw. tippt, benötigt eine gewisse Zeit, um etwa das recht lange Wort „Tastenkombinationen" einzugeben, das ich bis zu dieser Stelle bereits ganze 6x getippt und dabei erfolgreich der Versuchung widerstanden habe, zum kürzeren „Shortcut" zu greifen. Oder um den kürzeren, dafür aber komplexeren Eigennamen „Tekavčić" einzugeben. Gewiss, „Tastenkombinationen" würde man normalerweise in der Terminologiedatenbank (in der Grundform) hinterlegen und im Idealfall z. B. über die entsprechende Tastenkombination in die Übersetzung einfügen. Aber was, wenn der Plural benötigt wird? Dann muss man sozusagen noch einmal Hand anlegen. Das zweite Beispiel würde man als Eigenname hingegen wohl eher nicht in einer Terminologiedatenbank finden.

Für diese wie auch für weitere komplexere Fälle (siehe unten) bietet sich ein Programm zur Textbausteinverwaltung an, wie z. B. PhraseExpress eines ist. Die Software lässt sich systemweit nutzen, also aus praktisch jedem beliebigen Programm heraus. Das ist ein entscheidender Vorteil gegenüber Insellösungen, die nur in einer einzigen Anwendung funktionieren, wie z. B. die (eher rudimentäre) Textbausteinverwaltung in MS Word. Nach dem Start von PhraseExpress wartet die Software beharrlich im Hintergrund, um ihre Dienste zu verrichten. Die in PhraseExpress verwalteten Textbausteine können dabei zunächst einmal aus einem beliebigen einzelnen Wort bestehen, aus einer Wortfolge oder auch aus einem ganzen Dokument. Die Erstellung eines Textbausteins geht nach ein wenig Einarbeitung und Übung – auch dank der sehr gut gemachten Dokumentation des deutschen Herstellers Bartels Media aus Trier – rasch von der Hand. Für die Erstellung des Textbausteins für „Tastenkombinationen" habe ich z. B. lediglich hand- bzw. smartphone-gestoppte 15 Sekunden benötigt.

Textbausteine lege ich immer dann an, wenn ich Dinge wiederholt schreibe. Klassiker sind natürlich die Begrüßungs- und Schlussformeln in E-Mails. Aber auch meine IBAN zählt dazu (die ich mir tatsächlich auch Jahre nach ihrer Einführung noch

nicht merken kann) und reicht bis hin zu ganzen Texten, etwa standardisierte E-Mails (Antworten auf Anfragen etc.). Dabei können ohne große technische Hürden auch Auswahlfenster oder Textabfragen für variierende Textbestandteile integriert werden, z. B. zur Auswahl der korrekten Anrede („Sehr geehrte Frau...“, „Hallo Herr...“) oder zur Eingabe des Adressaten der jeweiligen E-Mail (s. Video 1 auf Seite 30).

Textbausteine können auf unterschiedliche Weise in den aktuellen Text eingefügt werden, so z. B. per Tastenkombination, per Textkürzel oder mittels so genanntem SmartComplete. Hierbei erkennt die Software automatisch hinterlegte Textbausteine anhand ihrer Anfänge und schlägt vor, diese einzufügen.

Textbausteine können mehr als nur Text

Anders als es der Ausdruck „Textbaustein“ suggeriert, können die Textbausteine in PhraseExpress aber nicht nur Text als solchen enthalten, sondern weitaus mehr. So lassen sich damit z. B. Webseiten aufrufen (etwa um ein Wort aus einer Übersetzung auf einer Webseite nachzuschlagen) oder auch Tastatureingaben integrieren. Mit Hilfe dieser letzten Variante habe ich übrigens das Problem gelöst, dass die Tastenkombination **Strg+N** in MultiTerm nicht verwendet werden konnte (siehe oben), und zwar indem ich **Strg+N** als Auslöser für einen Textbaustein definiert habe, der wiederum das Drücken der **F3**-Taste simuliert. Die Verwendung des Textbausteins wurde zudem auf MultiTerm begrenzt, damit er nicht ungewollt in anderen Anwendungen verwendet wird (s. Abbildung 3).

PhraseExpress als Schweizer Taschenmesser zur DIY-Automatisierung

Textbausteine können zusätzlich auch „Makros“ enthalten. In PhraseExpress werden Makros dazu genutzt, um Mausbewegungen und -klicks aufzuzeichnen, um diese zu einem späteren Zeitpunkt wieder abzuspielen – ohne dass man selber zur Maus greifen muss, versteht sich. Hierzu wird eine kleine Zusatzsoftware namens MacroRecorder benötigt, die von der Website des Herstellers heruntergeladen werden kann (und die Variante „Enterprise“ von PhraseExpress erfordert). Die Anwendungsszenarien hierbei sind nahezu unbegrenzt. Häufig drucke ich Texte beispielsweise mit 2 Seiten pro Blatt aus, um Papier zu sparen, was ohne PhraseExpress-Makro mit 1 Tastenkombination und ganzen 6 Mausklicks einhergeht. Oder ich sehe mich wiederholt gemüßigt, in Word die Einstellung für das Einfügen der typographischen Anführungszeichen („...“) anstelle der geraden ("...") anzupassen, da sich diese in Word in unregelmäßigen Abständen (von alleine?) umstellt. Mit einem aufgezeichneten Makro sind diese Dinge per einfacher Tastenkombination erledigt (s. Video 2 auf Seite 30).

PhraseExpress verfügt aber noch über weitere nützliche Funktionen, wie eine mehrfache Zwischenablage, die systemweit funktioniert, das automatische Erlernen individueller Rechtschreibfehler und entsprechender Autokorrektur oder auch die Möglichkeit Anwendungen zu starten. So lasse ich allmorgendlich meine verschiedenen Wörterbuch-Anwendungen per einfachem Tastaturdruck en bloc starten (s. Abbildung 4 auf Seite 29).

Skriptsprachen: Automatisierung für Fortgeschrittene

Die Tatsache, dass PhraseExpress eine Anwendung mit ansprechender Benutzeroberfläche und guter Dokumentation ist, macht den Einstieg in die Welt der Arbeitsoptimierung und -automatisierung zu einem Kinderspiel. Daneben gibt es mit Skriptsprachen weitere Möglichkeiten, die eigenen Arbeitsabläufe und -prozesse zu verbessern. In diesem Kontext sei das kostenlose AutoHotkey genannt, mit denen sich Skripte – d.h.

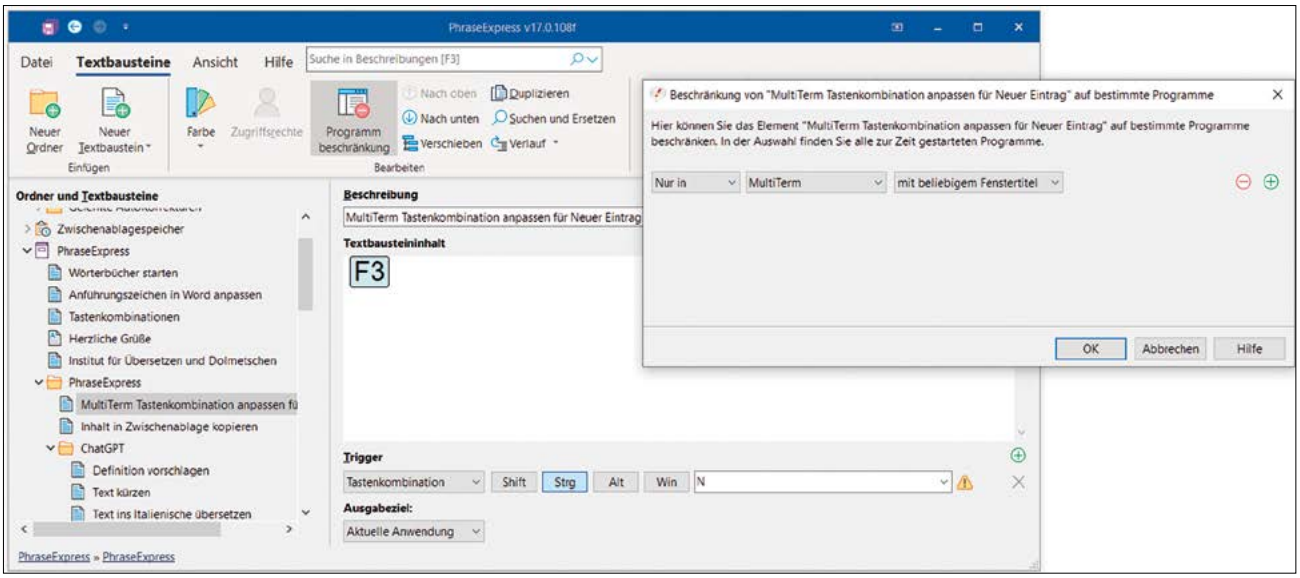


Abb. 3: Beispiel eines Textbausteins mit integriertem Tastendruck (Quelle: eigener Screenshot)

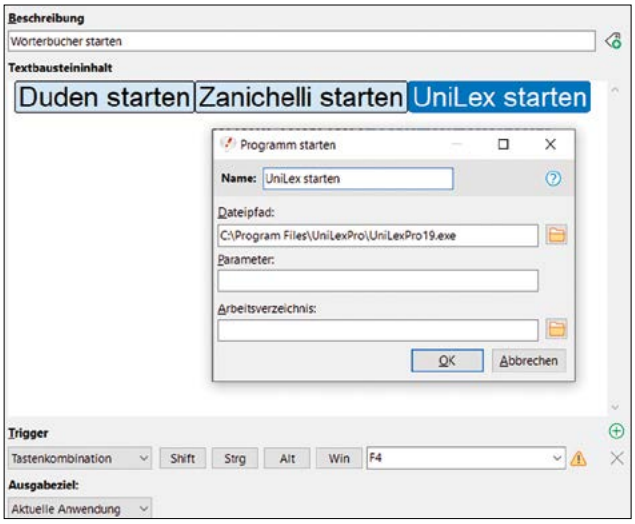


*Abb. 4: Beispiel eines Textbausteins für den gleichzeitigen Start von drei Wörterbuchanwendungen mittels **F4**-Taste (Quelle: eigener Screenshot)*

relativ einfache Abfolgen von Befehlen – erstellen lassen, die beispielsweise per Tastenkombination aufgerufen werden.

Die Erstellung und Bearbeitung der Skripte kann in einem beliebigen Texteditor wie z. B. dem empfehlenswerten und zudem kostenlosen Notepad++ erfolgen oder auch im ebenfalls kostenlosen AutoHotkey-eigenen Editor SciTE4AutoHotkey, dessen Verwendung allerdings ein wenig komplexer ist. Der Einstieg in die Skript-Erstellung ist nicht ganz ohne, aber u.a. dank einer guten und ausführlichen Dokumentation auf Deutsch [1] auch für Laien machbar.

Aber wofür lässt sich AutoHotkey, kurz AHK, nun konkret einsetzen? Bekanntlich gibt es Lösungen, um die Recherche im Web nach Definitionen und Übersetzungen effizienter zu machen. Neben lokalen Anwendungen, die systemweit verfügbar sind, wie etwa IntelliWebSearch (das übrigens in der kostenlosen Version 3 mit AutoHotkey erstellt wurde [2] – nur um die Möglichkeiten von AutoHotkey einmal anzudeuten), gibt es bekanntermaßen auch entsprechende Websites wie MagicSearch oder Browser-Erweiterungen wie Trufflepiggy (s. hierzu den Beitrag von Armin Mutscheller in edition 1/2022) – mit der offensichtlichen Einschränkung, dass diese naturgemäß ausschließlich im Browser eingesetzt werden können. Wer aber z. B. Trufflepiggy direkt aus dem Editor seines TM-Systems oder aus seiner Terminologiedatenbank heraus verwenden möchte, hat das Nachsehen – oder aber greift zu einem kleinen AHK-Skript (s. Abbildung 5). Das Skript fungiert dabei als Bindeglied zwischen einer beliebigen Anwendung, in der ein oder mehrere Wörter markiert werden, um diese mit Trufflepiggy zu suchen, und der Browser-Erweiterung selbst. Wenn man also beispielsweise bei seiner Terminologiearbeit in MultiTerm nach einer deutschen Entsprechung für die italienische Benennung „ventilare“ sucht, genügt es, einfach das Wort „ventilare“ im Text zu markieren und die im Skript festgelegte Tastenkombination zu drücken (im vorliegenden Fall **Alt+F2**, die im Skript als „!F2“ hinterlegt ist). Schon öffnet sich im Browser wie von Geisterhand die Browser-Erweiterung Trufflepiggy. Das Suchwort ist praktischerweise auch schon eingefügt. Jetzt nur noch die Zahl eingeben für die gewünschte Suchquelle und das war’s schon (s. Video 3 auf Seite 30).

Die Verwendung von AHK-Skripten erfordert die Installation der entsprechenden Skriptsprache. Skripte können über eine integrierte Funktion aber auch einfach zu Anwendungen kompiliert werden, z. B. um sie anderen zur Verfügung zu stellen, die AutoHotKey nicht installiert haben.

Chatbots als persönliche Programmier-Assistenten

Wem die Einarbeitung in AutoHotkey und die manuelle Erstellung von Skripten zu technisch ist, der kann einerseits von anderen erstellte Skripte aus dem Web z. B. von GitHub oder Stack Overflow herunterladen, so z. B. eines für die Autokorrektur häufiger Rechtschreibfehler im Englischen [3]. Andererseits steht uns in Form der Chatbots auch eine Art persönlicher Programmierer zur Verfügung, an die man einfach die Arbeit der Skripterstellung auslagert. Neben vielen anderen Dingen beherrschen die gängigen Chatbots nämlich auch AutoHotkey und z. B. Pearl, eine häufig zur Verarbeitung von Text und Texten eingesetzte Programmiersprache.

Bei meiner letzten Website-Lokalisierung entsprachen die gelieferten XLIFF-Dateien z. B. nicht den Anforderungen meines TM-Systems, wodurch keine zu lokalisierenden

```
Trufflepiggy.ahk
1  ; Anmerkung: AutoHotkey arbeitet z.T. mit speziellen Zeichen,
2  ; die für bestimmte Tasten stehen, so z. B. ! für Alt und ^ für Strg
3  !F2:: ; diese Zeile definiert, wie das Skript ausgeführt wird, und zwar über "Alt+F2"
4  Clipboard =  ; der Inhalt der Zwischenablage wird zur Sicherheit gelöscht
5  Send ^c ; der Befehl "Send" steht für eine Tastatureingabe oder einen Mausklick, hier von "Strg+C"
6  ClipWait 1
7  Run https://trufflepiggy.com ; "Run" führt etwas aus oder öffnet etwas, in diesem Fall die Website
8  Sleep, 2000 ; "Sleep" definiert eine Pause, angegeben in Millisekunden, hier um zu warten, dass sich die
   Website öffnet
9  Send {Ctrl down}{Shift down}f{Ctrl up}{Shift up}
10 Sleep, 1000
11 Send ^v
12 return ; "return" schließt ein Skript ab
13
```

Abb. 5: Kommentiertes AutoHotkey-Skript zur Verknüpfung einer beliebigen Anwendung mit der Browser-Erweiterung Trufflepiggy (Quelle: eigener Screenshot)

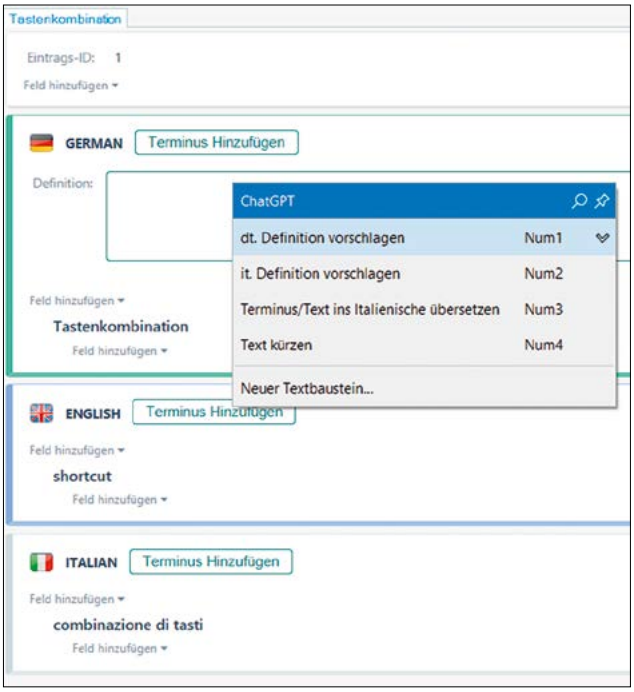


Abb. 6: Die Integration eines Chatbots mittels PhraseExpress in eine beliebige Anwendung, hier am Beispiel von ChatGPT in MultiTerm (Quelle: eigener Screenshot)

Inhalte angezeigt wurden, da die Zielsegmente in der XLIFF-Datei leer waren. Ein paar wenige Prompts später lag mir dank ChatGPT ein Pearl-Skript vor, das das Problem in Sekundenschnelle behoben hat.

Und noch ein letzter Tipp in Sachen Chatbots: In der aktuellen Version von PhraseExpress lassen sich nun auch Chatbots (u.a. ChatGPT) in die Software integrieren. Das Besondere dabei: Die Chatbots lassen sich dadurch aus praktisch jeder beliebigen Anwendung heraus mit individuellen Prompts anzapfen. Für die Terminologiearbeit kann man sich z. B. direkt aus MultiTerm heraus Definitionen vorschlagen lassen (s. Abbildung 6). Hier sind der Produktivität tatsächlich praktisch keine Grenzen gesetzt, da man die Abfragen in Form von Prompts selber in PhraseExpress definiert.

Fazit

Im Zeitalter von KI ist Automatisierung manchmal also nur einen Prompt weit entfernt. Und einen Prompt zu schreiben, ist wahrlich kein Hexenwerk – wie es letztlich auch die Optimierung und Automatisierung der eigenen Arbeitsabläufe keines ist. Das Wichtigste ist, man fängt einfach mal an, gerne auch im Kleinen.

Quellen:

[1] https://www.autohotkey.com/boards/viewforum.php?f=8 [17.03.2025]
[2] Farrell, Michael (2014): Solving Terminology Problems More Quickly with 'IntelliWebSearch (Almost) Unlimited', in Translating and The Computer 36, S. 212.
[3] https://gist.github.com/endolith/876629 [17.03.2025]

Video 1:
Einfügen eines Textbausteins mit PhraseExpress, hier am Beispiel eines Beispieltextes zur Beantwortung einer Anfrage

https://edition.dttev.org/2025-1/video1.mp4

Video 2:
Ausführen eines mit PhraseExpress aufgezeichneten Makros, im vorliegenden Fall für die Anpassung der Anführungszeichen in Word

https://edition.dttev.org/2025-1/video2.mp4

Video 3:
Verwendung eines Chatbots mit Hilfe von PhraseExpress, hier aus MultiTerm heraus zur Abfrage einer Definition von ChatGPT

https://edition.dttev.org/2025-1/video3.mp4

Videos vom Autor erstellt.

Maurice Mayer-Dewor ist Diplom-Übersetzer für die Sprachen Italienisch und Französisch und arbeitet als wissenschaftlicher Dozent am Institut für Übersetzen und Dolmetschen (IÜD) der Universität Heidelberg. Dort lehrt er u. a. im Bereich der Übersetzungstechnologien. Zudem beschäftigt er sich mit Lexikographie und ist Mitherausgeber des Wörterbuchs der italienischen Verben (WIV). Er arbeitet darüber hinaus als freiberuflicher Übersetzer und Technischer Redakteur.

Kontaktadresse
maurice.mayer@iued.uni-heidelberg.de
www.iued.uni-heidelberg.de

Terminologie der Information und Dokumentation & DGI-Term

Annette Weilandt und Axel Ermert

Nicht nur der Deutsche Terminologie-Tag ist ein Verein, dem Terminologie, Fachinformationen und Fachwissen am Herzen liegen. Auch die Deutsche Gesellschaft für Information & Wissen e. V. (DGI) beschäftigt sich seit ihrer Gründung 1948 mit dem Themenspektrum Information und Wissen. Die Vereinsmitglieder der DGI stammen dabei vorwiegend aus den Bereichen Informationswissenschaft, Bibliotheks- und Archivwesen sowie Gedächtnisinstitutionen und Verlagen.

Natürlich ist man sich in einem solchen Umfeld bewusst, dass korrekte Fachwörter und Definitionen das Rückgrat in jedem Tätigkeitsfeld bilden. Die Bereitstellung einer einheitlichen und klar definierten Terminologie für den Informationsbereich nimmt daher seit jeher eine zentrale Rolle bei den Aktivitäten der DGI ein. Seit 1966 kümmert sich die heutige DGI-Fachgruppe Arbeitskreis Terminologie und Sprachfragen (AKTS) darum, die Terminologie der Information und Dokumentation systematisch zu erarbeiten und bereitzustellen. Bereits 1975 gab der Arbeitskreis dazu die erste Publikation heraus, in der die Arbeitsergebnisse des AKTS in ca. 1.300 Einträgen als systematisch aufbereiteter Fachwortschatz veröffentlicht wurden. Seitdem dient die Terminologie der Information und Dokumentation (TID) als nützliches Nachschlagewerk und wichtige Basis für die Ausbildung und Lehre sowie für die fachübergreifende Zusammenarbeit.

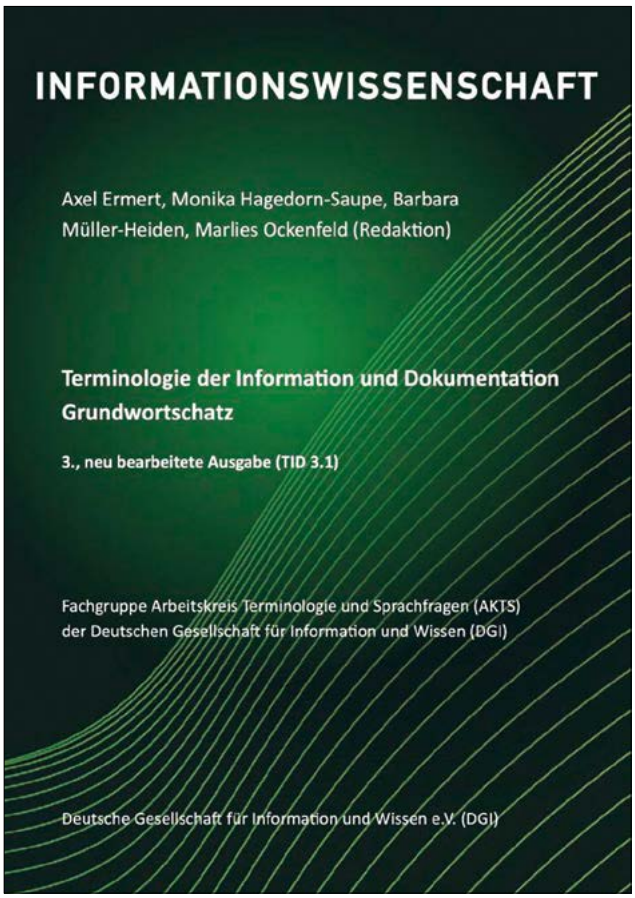


Abb. 1: Terminologie der Information und Dokumentation (TID) (Quelle: DGI)

Im September 2023 ist inzwischen ein Grundwortschatz als dritte, neu bearbeitete Auflage der Terminologie der Information und Dokumentation (TID 3.1) erschienen (Abbildung 1). Dieser Grundwortschatz umfasst insgesamt rund 300 Begriffe (ein Teilbestand gegenüber den 900 Begriffen der 2. Ausgabe 2005) aus der Informations- und Dokumentationswissenschaft und ist bei der DGI-Geschäftsstelle als Druckausgabe oder PDF-Dokument erhältlich. Da die Begriffe von Beginn an systematisch und strukturiert erarbeitet wurden, gibt es zu jedem Begriff eine Definition. Angereichert werden die Einträge mit weiteren Informationen wie Synonymen, Beispielen und Anmerkungen sowie mit englischen und französischen Äquivalenten (siehe Abbildung 2 auf Seite 32). Für einen übersichtlichen Zugriff sind die Begriffe den folgenden zehn Hauptgruppen thematisch zugeordnet:

- Grundbegriffe
- Digitale Kommunikation
- Dokumentationsprozess
- Informationsprodukte / Informationsdienste
- Informationssuche
- Dokumentarten
- Nummerung
- Dokumentspeicherung
- Rechtliche Aspekte
- Einrichtungen der Information und Dokumentation

Zusätzlich bietet die Publikation auch ein ausführliches Quellen- und Normenverzeichnis und verweist auf weiterführende Fachliteratur zum Thema.

Um diesen Fachwortschatz nicht nur wie bisher in einer gedruckten Publikation oder als PDF, sondern auch

zeitgemäß in einer Online-Terminologiedatenbank zugänglich zu machen, hat der AKTS vor einiger Zeit sein neuestes Projekt DGI-Term ins Leben gerufen. Ziel des Projektes ist, den Gesamtbestand des TID 3.1 mit Hilfe der Terminologieverwaltungssoftware Kalcium Quickterm online bereitzustellen. Geplant ist vorerst nur ein Zugriff für alle Mitglieder der DGI.

Karteileiche
en non-active member
en member on paper
fr membre passif
fr membre inactive

Eine Karteileiche ist eine natürliche oder juristische Person, die nicht mehr erreichbar ist.

Lesedauer
en reading time
fr durée de lecture

Die Lesedauer ist die Zeit, die das Lesen eines *Dokuments* voraussichtlich benötigt.

ANM „Voraussichtliche Lesedauer" wird heute häufig vom Anbieter eines digitalen Dokuments (z.B. Post, digitalisierter Zeitungsbeitrag, ...) bereits mit angegeben.

Abb. 2: Auszug aus TID 3.1 (Quelle: Eigene Darstellung)

Mit DGI-Term sollen jedoch nicht nur die Terminologieeinträge bereitgestellt werden, sondern zusätzlich werden vom AKTS auch Concept Maps (siehe Abbildung 3) zu den zehn Hauptgruppen erarbeitet, um so einen neuen Mehrwert für Nutzer zu schaffen. Durch die Erarbeitung der Concept Maps erfährt der Terminologiebestand nochmals eine Qualitätsverbesserung. Zwar hatte der Bestand auch in der Vergangenheit schon eine umfassende und formal terminologisch sehr gut aufbereitete Konzeption, denn die Begriffe wurden ausführlich abgestimmt und sehr sorgfältig erarbeitet. Der Import in die Eintragsstruktur einer Terminologiedatenbank und vor allem die Einordnung in Concept Maps haben dennoch einige vorhandene Probleme und Ungenauigkeiten aufgedeckt, beispielsweise bei der Systematik und Konsistenz der Definitionen. Das Festlegen sinnvoller und geeigneter Begriffsbeziehungen für die Concept Maps nahm ebenfalls einige Zeit in Anspruch und führte teils zu intensiven Diskussionen. Da aktuell noch nicht alle Begriffe des TID 3.1 vollständig in Concept Maps eingeordnet wurden, sind hier sicherlich noch weitere Konsolidierungsarbeiten zu erwarten, die zu spannenden terminologischen Diskussionen im AKTS und weit darüberhinaus führen werden.

Dr. Annette Weilandt
Geschäftsführerin des DTT

Axel Ermert
Vorstandsmitglied der DGI

Deutsche Gesellschaft für Information & Wissen e.V. (DGI)

Website: dgi-info.de

TID 3.1 kann über folgenden Link bestellt werden: https://dgi-info.de/wissenschaft-praxis/dgi-fachgruppen/akts/

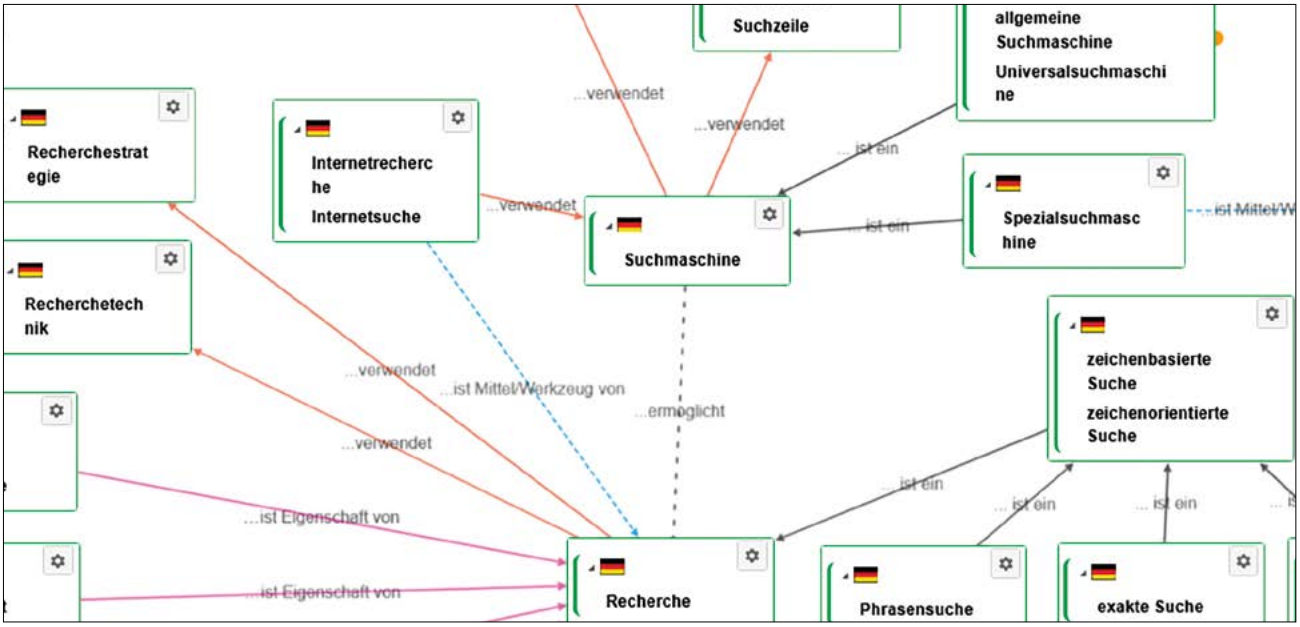


Abb. 3: Concept Map in DGI-Term (Quelle: Eigener Screenshot)

Das neue DTT-Fortbildungsprogramm

Christian Kriele

Abb. 1: Das DTT-Fortbildungsprogramm umfasst Veranstaltungen in Präsenz und online. (Foto: Chris Montgomery, Unsplash)

Seit vielen Jahren bietet der DTT sein Fortbildungsprogramm mit dem eintägigen Grundlagenseminar und dem Vertiefungsseminar mit mehreren Bausteinen an. Ergänzt wird das Programm durch Webinare, die in Kooperation mit dem BDÜ durchgeführt werden, sowie durch das alle zwei Jahre stattfindende Symposion. Die Teilnahme an bestimmten Veranstaltungen des DTT-Fortbildungsprogramms berechtigt zum Erwerb des DTT-Terminologiezertifikats.

2023 hat die Gesellschaft für Technische Kommunikation (tekom Deutschland e. V.) ihren „Kompetenzrahmen für Technische Übersetzung“ vorgestellt, an dessen Erarbeitung auch der DTT beteiligt war. Dieser rollenorientierte Kompetenzrahmen enthält einen eigenen Baustein „Terminologie“ und beschreibt auch die Kompetenzen für die Rollen „Terminologiearbeit“ und „Terminologiemanagement“ (siehe https://tinyurl.com/3p6u8pnv).

Das Erscheinen des Kompetenzrahmens hat der DTT zum Anlass genommen, sein Fortbildungsprogramm zu überarbeiten und neu zu strukturieren. Der vorliegende Artikel stellt das neue, überarbeitete Programm vor, in das Erfahrungswerte aus den bisherigen Seminaren sowie die Anforderungen der im tekom-Kompetenzrahmen definierten Kenntnisse und Fertigkeiten eingeflossen sind.

Überblick

Basismodul I – Terminologiegrundlagen – 1 Tag, Präsenzveranstaltung

Basismodul I vermittelt die Grundlagen der Terminologiewissenschaft und der Terminologiearbeit, auf denen alle weiteren Module (Aufbau und Vertiefung) aufbauen.

Basismodul II – Terminologiewerkzeuge – 1 Tag, Präsenzveranstaltung

Im Rahmen dieses Seminars wird die praktische Anwendung und Funktionsweise von Terminologiewerkzeugen sowie deren Konfigurationsmöglichkeiten demonstriert. Die Arbeit mit konkreten Tools soll den Einsatz in der Praxis verdeutlichen.

Aufbaumodul I – Terminologiemanagement – 2 Halbtage, online

In Aufbaumodul I werden die in den Basismodulen I und II erworbenen Kompetenzen angewendet und in praktischen Übungen vertieft. Dem Aufbau eines Terminologieleitfadens folgend werden verschiedene Aspekte des Terminologiemanagements schwerpunktmäßig in Form von Übungen behandelt.

Aufbaumodul II, Teil 1 – Wissensorganisation und Terminologiedatenaustausch – 1 Halbtag, online

Teil 1 dieses Moduls führt in das Gebiet des Wissensmanagements ein und macht die Teilnehmer:innen mit typischen Wissensordnungen vertraut. Neben den theoretischen Grundlagen werden durch Übungen praktische Kenntnisse zur Erstellung von Wissensordnungen vermittelt.

Aufbaumodul II, Teil 2 – Terminologiedatenaustausch – 1 Halbtag, online

Terminologische Daten müssen häufig aufbereitet werden, um sie in anderen Anwendungen nutzen zu können. Teil 2 des Moduls bietet eine Einführung in dieses Thema. Wichtig dabei sind insbesondere vorbereitende Maßnahmen für den Austausch terminologischer Datenbestände, aber auch die Nachbearbeitung von Export-/Importdateien.

Vertiefungsmodul I – Wissensorganisation – 2 Halbtage, online

Das Modul vermittelt Kenntnisse zu komplexen Wissensordnungen wie Thesauri und Ontologien und deren Nutzung in Knowledge Graphen und KI-Anwendungen. Die Teilnehmer:innen lernen anhand von praktischen Übungen verschiedene Methoden der Wissensrepräsentation kennen, erhalten einen Überblick zu nützlichen Ressourcen und erfahren, wie diese Methoden die Grundlage für moderne KI-Anwendungen bilden.

Vertiefungsmodul II – Terminologiedatenaustausch – 2 Halbtage, online

Der Austausch terminologischer Daten ist in der Regel ein sehr komplexes Unterfangen und erfordert solide Kenntnisse und Fähigkeiten in puncto Datenmodellierung, Austauschformate und Daten-Engineering. Dieses Modul vermittelt die entsprechenden Kenntnisse und Kompetenzen.

Referent:innen:

- Prof. Dr. Petra Drewer
- Beate Früh
- Prof. Dr. Rachel Herwartz
- Dr. Nicole Keller
- Prof. Dr. Felix Mayer
- Dr. Detlef Reineke
- Prof. Dr. Klaus-Dirk Schmitz
- Dr. Annette Weilandt
- Tom Winter

Das DTT-Fortbildungszertifikat

Für den Erhalt des Zertifikats müssen folgende DTT-Veranstaltungen besucht werden:

- 2 DTT-Basismodule (I und II)
- 1 DTT-Aufbaumodul (I oder II)
- 2 DTT-Webinare
- 1 DTT-Symposion

Wann die Veranstaltungen jeweils stattfinden, ist auf der Webseite des DTT ersichtlich:
dttev.org/fortbildung

Bereits besuchte DTT-Fortbildungen können nachträglich anerkannt werden. Der jeweilige Nachweis muss durch die Teilnehmer:innen selbst erbracht werden.

Bei Interesse wenden Sie sich bitte an
Christian Kriele
fortbildung@dttev.org

Abb. 2: Das DTT-Symposion fand 2025 im Wormser Tagungszentrum statt. (Foto: Tamara Arndt)

Wechsel auf dem DTT-Chefsessel

Auf der DTT-Jahresmitgliederversammlung 2025 ist der 1. Vorsitzende des Verbands neu gewählt worden. Wir von der Redaktion der edition haben mit ihm gesprochen und möchten ihn unseren Leserinnen und Lesern kurz vorstellen:

Abb. 1: Peter Seip (Foto: privat)

Sein Name ist Peter Seip. Nach dem Studium startete er als Diplom-Übersetzer für Englisch und Französisch bei der Siemens AG, der er fast 35 Jahre treu blieb. Sein Aufgabenfeld reichte von der Mitwirkung bei der Einführung des MÜ-Systems METAL über die Einführung von Translation-Memory- und Workflow-Systemen und deren Anbindung an das Redaktionssystem bis hin zum Aufbau der Terminologiedatenbank des Siemens-Sprachendiensts, deren Weiterentwicklung und Integration mit dem Redaktionssystem – beste Voraussetzungen also für seinen neuen „Job".

Auf die Frage, welche Ziele Peter Seip für den DTT hat und welche Schwerpunkte er für seine Amtszeit setzen möchte, antwortete er:

> „Ich freue mich sehr, daran mitzuarbeiten, dass der Deutsche Terminologie-Tag auch in Zukunft die Austauschplattform für Terminologiearbeit im deutschen Sprachraum ist und einen Austausch von Wissen und praktischen Erfahrungen ermöglicht, von dem alle am Terminologieprozess Beteiligten profitieren. Wie die Vorträge und Diskussionen des DTT-Symposions 2025 zeigten, befinden wir uns mitten in einer rasanten Entwicklung der KI. Die zukünftige Rolle der Terminologiearbeit in diesem Umfeld wird ein prägendes Thema auch für die Arbeit des DTT als Forum sein. Besonders wichtig für die Zukunft ist mir ebenfalls der anerkannt hohe Qualitätsstandard des Weiterbildungsangebots und der Veranstaltungen des DTT."

Da die JMV im Vorfeld des DTT-Symposions 2025 in Worms stattfand (Bericht siehe Seite 37), bot dies einen würdigen Rahmen für die Verabschiedung des „alten" 1. Vorsitzenden Tom Winter, der das Amt seit 2017 innehatte.

Tom Winter hat zwar den Vorsitz abgegeben, bleibt dem DTT-Vorstand aber als Referent für Wissensmanagement und KI sowie als Koordinator des sehr erfolgreichen Industrieforums (Bericht siehe Seite 37), das er ins Leben gerufen hat, erhalten. An dieser Stelle bedanken wir vom DTT-Vorstand und DIT-Vorstand uns ganz herzlich für das großartige Engagement von Tom Winter und freuen uns sehr, dass er den DTT auch weiterhin tatkräftig unterstützt.

Abb. 2: Petra Drewer, 1. Vorsitzende des DIT, überreicht Tom Winter einen Ast – stellvertretend für 50 Baumpatenschaften, die DTT und DIT in Tom Winters Namen in Winterberg abgeschlossen haben. (Foto: Angelika Ottmann)

Den neuen 1. Vorsitzenden Peter Seip heißen wir herzlich im DTT-Vorstand willkommen und blicken mit Zuversicht auf eine vertrauensvolle und erfolgreiche Zusammenarbeit.

DTT-Vorstand und
DIT-Vorstand

Neues vom Industrieforum

Tom Winter

Auf Einladung von Wilo fand das DTT-Industrieforum am 6. und 7. Mai in Dortmund statt. Ein großes Dankeschön an Wilo für die herzliche Gastfreundschaft! Und natürlich auch an Armela und Joanna für die hervorragende Organisation – großes Kompliment!

Die Teilnehmenden waren beeindruckt vom Unternehmensgelände, das moderne Architektur und Nachhaltigkeit gekonnt vereint. Besonders hervorzuheben: die hochautomatisierten Produktionsprozesse, die bei einer Werksführung vorgestellt wurden. Ein weiteres Highlight war die Unternehmenspräsentation auf einer gigantischen LED-Wand – mit mehr LEDs als der Stadionwürfel auf Schalke! Dieser Besuch wird sicherlich allen in lebhafter Erinnerung bleiben. Nochmals vielen Dank!

Inhaltlich standen spannende Themen im Mittelpunkt: **„KI-gestützte Terminologieprozesse“**, **„Wissensmodellierung und LLM-Customization“** sowie **„Automated Quality Estimation“**. Das Fazit: Diese Bereiche werden auch bei den kommenden Treffen einen zentralen Platz einnehmen, denn die neuen Technologien bieten enormes Potenzial für Experimente und vielfältige Möglichkeiten zur Prozessoptimierung. Besonders betont wurde dabei: Nicht nur die Sprachabteilungen profitieren von der Terminologie – viele Unternehmensbereiche können wertvolle Impulse gewinnen. **There is no AI without terminology!** Der Austauschbedarf ist groß und wird weiter wachsen.

Auch für weitere Themen war Raum: **„NMT und Customizierung“**, **„Autorenunterstützung“**, **„Extraktion“** sowie **„Change-Management“** und **„internes Marketing für Terminologieprozesse“** kamen zur Sprache. Das allgemeine Fazit der Runde: Es gibt noch viel zu tun im Sprachdatenmanagement – und gerade der gemeinsame Austausch spielt dabei eine wichtige Rolle. Der Einblick in die Erfahrungen der anderen hilft, eigene Fehler zu vermeiden und auf neue Ideen zu kommen, die man vielleicht selbst noch nicht bedacht hatte. Diese Impulse nehmen die Teilnehmenden mit, um ihre eigenen Projekte voranzubringen – packen wir es an!

Die Fortsetzung des Austauschs ist bereits geplant: Im Herbst 2025 findet ein Online-Treffen statt, das nächste DTT-Industrieforum in Präsenz folgt im kommenden Jahr in Mannheim. Auch die darauffolgenden Jahre sind bereits verankert: 2027 in Freiburg und 2028 voraussichtlich in Heidelberg. Schon jetzt ein herzliches Dankeschön an die gastgebenden Unternehmen für ihre Unterstützung!

Mit fünf neuen Mitgliedern wächst das DTT-Industrieforum weiter und umfasst nun insgesamt 34 teilnehmende Unternehmen. Wir freuen uns auf den weiterhin lebendigen und inspirierenden Austausch!

Das **DTT-Industrieforum** wurde im Rahmen des DTT-Symposions 2018 auf Initiative von Liebherr, CLAAS und Deutsche Bahn AG gegründet. Mittlerweile gehören dem Forum 34 Unternehmen an. Am DTT-Industrieforum können alle juristischen DTT-Mitglieder teilnehmen, die weder Sprachdienstleister noch Beratungsunternehmen oder Toolhersteller sind. Bei Interesse wenden Sie sich bitte an den Autor: **winter@dttev.org**

Abb. 1: 2025 traf sich das DTT-Industrieforum bei Wilo in Dortmund. (Foto: DTT-Industrieforum)

Ich nehme Sie mit auf eine Reise nach Worms

Petra Sweekhorst

Worms am Rhein, Stadt der Nibelungen und Schauplatz des Reichstags zu Worms, auf dem Martin Luther seine berühmten Worte spricht. Ein würdiger Veranstaltungsort für das 19. DTT-Symposion. Das Tagungszentrum im Schatten des Doms ist von DTT-Fahnen geschmückt.

Klassisch stehen am Donnerstagnachmittag vier hochkarätige Workshops von Terminologieextraktion bis Wissensmanagement an, die sehr gut besucht sind. Nach der anschließenden Mitgliederversammlung mündet der Donnerstagabend im gemütlichen Beisammensein im Restaurant des Tagungszentrums. Kurze Wege, was allenthalben geschätzt wird.

Am Freitagmorgen eröffnet Tom Winter das 19. DTT-Symposion frei nach Luther: „Hier stehen wir und können nicht anders." Das Thema des Symposions „Terminologie in der KI – KI in der Terminologie" unterstreicht, dass KI auf Grundlage einheitlicher Terminologie bessere Ergebnisse liefert und dass KI als Werkzeug bei der Terminologiearbeit gewinnbringend zum Einsatz kommt. Wir Terminologiebegeisterten freuen uns, dass Terminologie durch KI Aufwind bekommt. Denn wir sind begeistert, sonst wären wir nicht hier.

Das Vortragsprogramm beginnt mit einem fulminanten Auftakt, der Keynote von Ziad Mahayni über „Deep Fakes und die Infokalypse". Nach einer halben Stunde über den Einfluss generativer Künstlicher Intelligenz auf das Informationsökosystem sind wir aufgeklärter als manchen von uns lieb ist. So weitreichend hatten wir die Konsequenzen von dem schieren

Abb. 1: DTT-Fahnen vor dem Tagungszentrum (Foto: Felix Mayer)

Übermaß an Informationen und von ständig auf uns einprasselnden Deep Fakes nicht vor Augen. Logischerweise gefällt vielen von uns nicht, was wir da erkennen müssen. Im Jahr 2025 lassen sich Deep Fakes einfach, im großen Stil und perfekt generieren. Folglich müssen wir Rezipienten allen Texten und Bildern noch kritischer gegenüberstehen, um zu versuchen, Wichtiges von Unwichtigem und Wahres von Gelogenem zu unterscheiden. Fake ist gelogen und manipuliert uns.

Abb. 2: Keynote-Sprecher Ziad Mahayni im Mozartsaal des Wormser Tagungszentrums (Foto: Tamara Arndt)

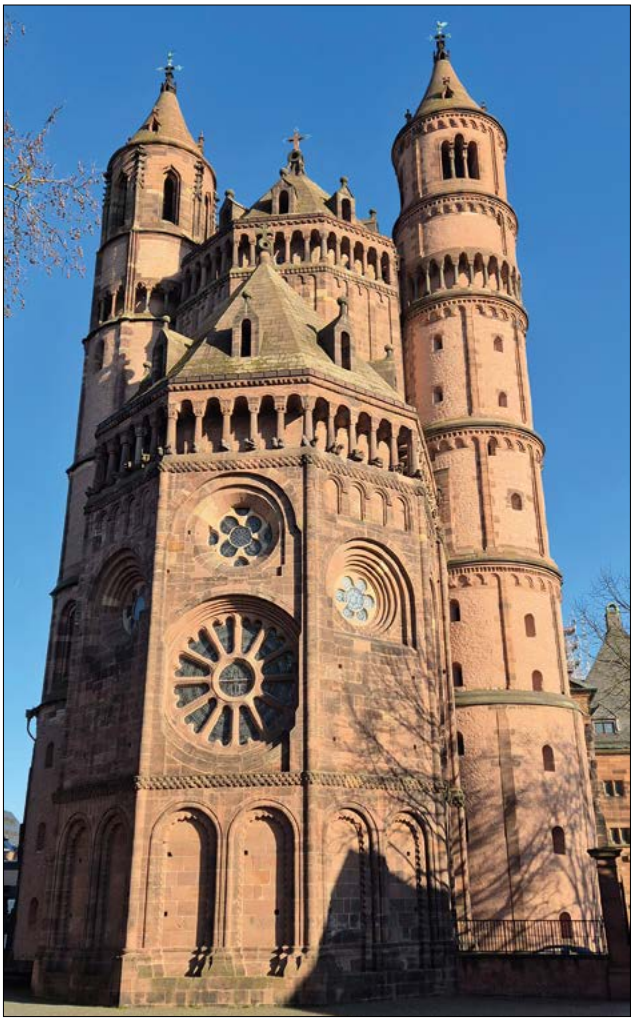


Abb. 3: Der Dom zu Worms (Foto: Angelika Ottmann)

Wir sind verantwortlich dafür, aus welchen Quellen wir Informationen beziehen und dass wir sie kritisch bewerten.

Fachlich hat der DTT einen weiten Rahmen aufgespannt von KI und Terminologiemanagement über Prompting und Translation bis hin zur Qualitätskontrolle. Die Referenten des DTT-Symposions lehren an Hochschulen, unterstützen als Dienstleister mit Software und Know-how Firmen bei der Umsetzung von Terminologieprojekten, stellen in der Industrie für ihre Kolleginnen und Kollegen Fachwortschatz bereit und und und …

Sie erkennen die Vielfalt, die sich auch in den zwei Preisträgerinnen des DTT-Förderpreises widerspiegelt. Ja, Sie haben richtig gelesen: zwei.

Muriel Böhmicke bekommt den Förderpreis für ihre vergleichende Analyse mit ChatGPT und ALMA bezogen auf die Rolle von SNOMED CT bei der Verbesserung maschineller Übersetzung medizinischer Fachtexte. Sie ist fachlich auf dem allerneusten Stand, entwickelt ihre Arbeit parallel zu den aktuellen KI-Innovationen ständig weiter und überzeugt mit ihrer Präsentation. Chapeau!

Vivien Krämer, ebenfalls Förderpreisgewinnerin, gestaltet den Wandel in der Roche Diagnostics International AG mit. Sie illustriert anschaulich den Werdegang von 1 Terminologen, der für 3 technische Dokumentationen ausschließlich in Deutsch und Englisch zuständig war, hin zu einem Team von 14 Sprachexpertinnen und -experten, die für 85 Projekte in mehr als 45 Sprachen die Gesamtverantwortung für alle Sprachressourcen tragen. Vivien Krämer beeindruckt durch ihr Durchhaltevermögen und ihre Standhaftigkeit und gleichzeitig durch ihre Flexibilität und die gegenüber anderen Stakeholdern erforderliche Diplomatie. Sie betont, dass sie den Förderpreis für ihr gesamtes Team entgegennimmt, denn alle vom Language Operations Team tragen zusammen zum Erfolg der DIA Language Factory bei.

Die Vortragsthemen sind so vielfältig, dass hier nicht genug Platz ist, sie alle zu würdigen. Deshalb verweise ich an dieser Stelle sowohl auf die Akten des Symposions, herausgegeben von Petra Drewer, Felix Mayer und Donatella Pulitano, wofür ihnen großer Dank gebührt, als auch auf die Vortragsfolien der Referenten, die auf der Internetseite des DTT zur Verfügung stehen. Dank an die freigiebigen Referenten und an die fachkundigen Köpfe für die schnelle Veröffentlichung.

Um ein Symposion finanziell abzusichern, braucht es nicht nur Besucher, sondern auch die entsprechenden Sponsoren. Die Aussteller haben dankenswerterweise mit für die pekuniäre Ausstattung gesorgt. Im Rahmen von Kurzvorträgen präsentieren sie ihre jeweiligen Arbeitsschwerpunkte und ihre Ausstellungsstände sind in den Vortragspausen gut besucht.
Natürlich bietet ein Zusammentreffen von „Termis“ noch mehr, z. B. die musikalische Unterhaltung durch Dad’s Phonkey, das Speed-Dating für zwangloses Kennenlernen, das gemeinsame Abendessen mit Blick auf den Rhein und den fachlichen Austausch auf der Dachterrasse oder während des Mittagessens.

Das nächste Symposion wird übrigens vom neu gewählten 1. Vorsitzenden des DTT, Peter Seip, eröffnet werden (siehe Artikel auf Seite 35). Tom Winter, der vorherige 1. Vorsitzende, wird mit viel Applaus verabschiedet und erhält als Dank für sein langjähriges Engagement eine Baumpatenschaft für 50 Bäume im Rahmen des Projekts „Waldfreunde“ in Winterberg.

Also, wenn Sie die Gelegenheit haben, dann kommen Sie zum nächsten Symposion!

Petra Sweekhorst ist Übersetzerin, Technische Redakteurin mit Spezialisierung auf Software, Maschinen- und Anlagenbau sowie Terminologiebegeisterte.

Kontaktadresse
p.sweekhorst@t-online.de

DTT-Förderpreis 2025: Zweifache Exzellenz

Der DTT-Förderpreis, der alle zwei Jahre für herausragende Leistungen im Bereich der Terminologie verliehen wird, ging in diesem Jahr gleich an zwei Preisträgerinnen. Die Einsatzbereiche könnten unterschiedlicher nicht sein, deshalb fiel die Entscheidung ja auch so schwer, sodass schließlich zwei Protagonistinnen auf dem Podium standen und auf dem DTT-Symposion ihre Arbeit präsentieren durften.

Zunächst sprach Muriel Böhmicke, die in ihrer Masterarbeit „Die Rolle von SNOMED CT bei der Verbesserung maschineller Übersetzungen medizinischer Fachtexte: Eine vergleichende Analyse mit ChatGPT und ALMA“ untersuchte, inwiefern die Integration von Terminologielisten aus der Terminologiedatenbank SNOMED CT in ChatGPT und ALMA zur Verbesserung der Übersetzungsqualität medizinischer Fachtexte beitragen kann. Das Hauptaugenmerk lag auf der potenziellen Steigerung der Effizienz und Genauigkeit der maschinellen Übersetzungen medizinischer Fachtexte. Es wurde untersucht, welche Herausforderungen bei der Integration von SNOMED CT in KI-Modelle auftreten können, und es wurden potenzielle Lösungsansätze und Optimierungen aufgezeigt.

Zweite Preisträgerin war die „Language Operations“-Abteilung bei Roche Diagnostics International, vertreten durch ihre Leiterin Vivien Krämer. Vorgestellt wurde, wie die mehrsprachige Terminologie(arbeit) nicht nur mit dem Übersetzungsbereich und der Technischen Redaktion verknüpft wird, sondern wie speziell darauf geachtet wird, dass mit der Terminologiearbeit ein echter Mehrwert geliefert wird. Dies stärkt auch die Rolle von Terminologinnen und Terminologen im Unternehmen.

Zwei wirklich beeindruckende Arbeiten, beide Preisträgerinnen konnten sehr verdient ihre Preise entgegennehmen.

Dr. Nicole Keller und Angelika Ottmann

Muriel Böhmicke

Vivien Krämer

Abb. 1: Petra Drewer, 1. Vorsitzende des DIT, gratuliert den Preisträgerinnen auf dem 19. DTT-Symposion in Worms. (Fotos: Tamara Arndt)

RWS Trados | KALEIDOSCOPE

ENTERPRISE TERMINOLOGY MANAGEMENT

- Mühelose Terminologiebereitstellung
- Transparentes Terminologiemanagement
- Systemintegration mit CMS und TMS
- Globale Workflows
- Wissensmanagement
- Einfache Zusammenarbeit
- Rückfragenmanagement
- Terminologieprüfung

Trados GmbH | Waldburgstr. 21 | 70563 Stuttgart
www.rws.com/de | kontakt@rws.com

Kaleidoscope GmbH | Landstr. Hauptstr. 99 | 1030 Wien
www.kaleidoscope.at | info@kaleidoscope.at