

# Gleich und Gleich gesellt sich gern

## Begriffsorientierte Termvalidierung

Ines Prusseit

Dr. Sabine Lehmann

Peter Tauter

acrolinx GmbH

# Überblick

- Einleitung
- maschinengestützte Termextraktion
- aktueller Stand der Termvalidierung
- Termvalidierung effizienter gestalten
  - » Termkandidaten gruppieren
  - » Informationen bündeln
  - » Termkandidaten bewerten
- Demo

# Einleitung

- Ziel: Einheitliche Verwendung von Terminologie
- Terminologieaufbau nötig:  
manuelle oder automatische Termextraktion
- Termvalidierung ebenfalls nötig, vor allem für die zukünftige Termprüfung

# Maschinengestützte Termextraktion

- oft Basis für Termvalidierung
- große Menge bereits geschriebener Texte
- Resultat: Liste von Termkandidaten, mit Informationen bereichert

# Maschinengestützte Termextraktion: *Heckscheibenheizungen*

<b>Frequenz</b>	4
<b>Quelle</b>	<i>Heizanlagen.txt, Thermik.doc</i>
<b>Kontext</b>	<i>Eine durchschnittliche Heckscheibenheizung hat eine Leistung von ca. 150 W.</i>
<b>Grundform</b>	<i>Heckscheibenheizung</i>
<b>Sprachmuster</b>	Kompositum aus 3 Elementen
<b>Wortart</b>	Substantiv
<b>Termvarianten</b>	<i>Heckscheiben-Heizung / HeckscheibenHeizung</i>

# Maschinengestützte Termextraktion: Vorteile

- automatisch
- datenorientiert
- zusätzliche Informationen unterstützen die Termvalidierung

# Maschinengestützte Termextraktion: Nachteile

- lange, größtenteils unsortierte Listen
- zusammengehörende Termkandidaten bzw. Benennungen und ihre Varianten nur schwer auffindbar

# Aktueller Stand der Termvalidierung: Zusammenfassung

- automatische Termextraktion ist vorhanden
- gelieferte Informationen sind wertvoll, sogar unabdingbar für Termvalidierung
- ABER: Termvalidierung erfordert sehr hohen Zeitaufwand

# Termvalidierung effizienter gestalten - aber wie?

- vollautomatische Termvalidierung nicht möglich
- **ABER:** Maschinengestützte Termvalidierung wünschenswert
- verschiedene Hilfen können angeboten werden

# Wunsch 1: Termkandidaten gruppieren

- Gruppierung der Termkandidaten, die möglicherweise zu einem Begriff zusammengefasst werden sollen
- Beispiel:  
*Heizelement, Heizungselement, Heizgerät, Heiz-Gerät, HE*
- ermöglicht zusammenhängende Validierung und vereinfacht den Prozess
- begriffsorientierte Termvalidierung

# Wunsch 1: Termkandidaten gruppieren

	A	B	C	D
1	<b>Benennung</b>	Heizelement	Oxidhaut	3-Wege-Ventil
2	<b>Schreibvariante</b>	Heiz-Element	Oxid-Haut	3-Wegeventil
3	<b>Ableitung</b>	Heizungselement	Oxidationshaut	3-Wegsventil
4	<b>Fremdwort</b>	heating element	Oxydhaut	
5	<b>Synonym</b>	Heizgerät	Oxidschicht	3-Wege-Hahn
6	<b>Abkürzung</b>	HE	O-Haut	
7	<b>Langform</b>	Radionuklid-Heizelement		3-Wege-Mischventil
8	<b>Zahlwort</b>			Drei-Wege-Ventil
9	<b>Rechtschreibung</b>	Heizungelement	Oxithaut	3-wege-Ventil

## Wunsch 2: Informationen bündeln

- Modellierung übergreifender Konsistenz:  
weitere Zusammenhänge zwischen Benennungen  
aufzeigen, z. B. Termbildungsmuster

<b>Kopf-varianten</b>	<i>Drei-Wege-Ventil, 2-Wege-Ventil, 2-Wegemischventil</i>
<b>Antonym-varianten</b>	<i>anwenderabhängig anwendungsunabhängig</i>
<b>Kontext-varianten</b>	<i>Das <b>3-Wege-Ventil</b> wird mit dem Thermostellantrieb ABC verwendet. Das <b>Zwei-Wege-Ventil</b> wird mit dem Thermostellantrieb ABC verwendet.</i>

## Wunsch 3: Termkandidaten bewerten

- semi-automatische Bewertung der Benennungen auf der Basis von Termbildungsregeln
- Beispiele positiver und negativer Eigenschaften von Benennungen, z. B.:
  - » Frequenz
  - » mit oder ohne Bindestrich
  - » Zahlwort vs. Ziffer
  - » Prozent vs. %
  - » m vs. Meter
- System macht einen Vorschlag, falls eindeutig

# => Maschinengestützte, begriffsorientierte Termvalidierung

- System, das diese 3 Wünsche erfüllt
- leichte Handhabung
  
- Demo

# Diplomarbeit (acrolinx)

- Autor: Peter Tauter
- Tauter, P. (erscheint 2010): Entwurf und Implementierung einer Komponente zur maschinenunterstützten Termvalidierung (vorläufiger Titel), Potsdam.

Dankeschön!  
Wir danken Ihnen.  
Wir bedanken uns.  
Vielen Dank!  
Besten Dank  
Herzlichen Dank  
Tausend Dank  
1000 Dank  
Wir sagen danke.  
Danke vielmals.